ACU Research Bank

A survey on wind power forecasting with machine learning approaches

Item Type	Journal article
Authors	Yang, Yang;Lou, Hao;Wu, Jinran;Zhang, Shaotong;Gao, Shangce
DOI	10.1007/s00521-024-09923-4
Publisher	Springer
Download date	2026-11-19 02:59:57
Link to Item	https://hdl.handle.net/20.500.14802/14026

Neural Computing and Applications (2024) 36:12753–12773
https://doi.org/10.1007/s00521-024-09923-4

REVIEW

A survey on wind power forecasting with machine learning approaches

Yang Yang[1] · Hao Lou[1] · Jinran Wu[2] · Shaotong Zhang[3] · Shangce Gao[4]

Received: 11 August 2023 / Accepted: 28 April 2024 / Published online: 18 May 2024
© The Author(s) 2024

Abstract

Wind power forecasting techniques have been well developed over the last half-century. There has been a large number of research literature as well as review analyses. Over the past 5 decades, considerable advancements have been achieved in wind power forecasting. A large body of research literature has been produced, including review articles that have addressed various aspects of the subject. However, these reviews have predominantly utilized horizontal comparisons and have not conducted a comprehensive analysis of the research that has been undertaken. This survey aims to provide a systematic and analytical review of the technical progress made in wind power forecasting. To accomplish this goal, we conducted a knowledge map analysis of the wind power forecasting literature published in the Web of Science database over the last 2 decades. We examined the collaboration network and development context, analyzed publication volume, citation frequency, journal of publication, author, and institutional influence, and studied co-occurring and bursting keywords to reveal changing research hotspots. These hotspots aim to indicate the progress and challenges of current forecasting technologies, which is of great significance for promoting the development of forecasting technology. Based on our findings, we analyzed commonly used traditional machine learning and advanced deep learning methods in this field, such as classical neural networks, and recent Transformers, and discussed emerging technologies like large language models. We also provide quantitative analysis of the advantages, disadvantages, forecasting accuracy, and computational costs of these methods. Finally, some open research questions and trends related to this topic were discussed, which can help improve the understanding of various power forecasting methods. This survey paper provides valuable insights for wind power engineers.

Keywords Wind power prediction · Time series · Machine learning · Deep learning · Large language models

1 Introduction

Wind power forecasting task can provide predictions of the potential electricity output of wind farms to facilitate scheduling plans. Wind power's instability and fluctuations make it challenging to integrate into the system [1, 2], requiring accurate predictions for stable grid operation [3–5]. Therefore, accurate wind power forecasting is crucial for wind farm operation and grid stability.

Wind power forecasting methods can be categorized on the basis of time scales: short-term, medium-term, and long-term predictions. Short-term wind power forecasting is made within the next few hours and is critical for real-time wind power scheduling [6–8]. Medium-term wind power forecasting is made within the next week and is primarily used for unit mix and standby arrangements

✉ Jinran Wu
ryan.wu@acu.edu.au

Yang Yang
yyang@njupt.edu.cn

Shaotong Zhang
shaotong.zhang@ouc.edu.cn

Shangce Gao
gaosc@eng.u-toyama.ac.jp

[1] Nanjing University of Posts and Telecommunications, Nanjing 210023, Jiangsu, People's Republic of China

[2] Australian Catholic University, Banyo, QLD 4014, Australia

[3] Ocean University of China, Qingdao 266000, Shandong, People's Republic of China

[4] University of Toyama, Toyama-shi, Toyama 930-8555, Japan

[9–11]. Long-term wind power forecasting is generally made in months or quarters and is essential for wind resource assessment and maintenance arrangements [12–15]. The main purpose of all forecasting methods is to improve the accuracy of wind power prediction as much as possible. Traditional wind power forecasting methods are physical and statistical methods [16, 17]; while, intelligent methods are machine learning and deep learning approaches. Physical models attempt to estimate wind speed time series while considering the physical characteristics of the environment [18], utilizing numerical weather prediction (NWP) data to calculate actual wind power generation [19]. However, physical methods suffer from drawbacks such as high computational costs and difficulty in obtaining NWP data from meteorological stations. Statistical models seek relationships among historical data parameters to predict future wind speed and wind power [20]. These methods include autoregressive (AR) [21], vector autoregression (VAR) [22], autoregressive moving average (ARMA) [23], autoregressive integrated moving average (ARIMA) [24], fractional autoregressive integrated moving average (FARIMA) [25], and models based on Kalman filtering [26]. While statistical methods can accurately handle highly volatile noise in wind power time series data with nonlinear input data features, they often lead to significant errors in wind power predictions [27]. Artificial neural network (ANN) forecasting method can provide more information about the uncertainty of wind power prediction [28], with results generated by the output layer [29]. ANN learns the relationship between wind power input and output by training with historical wind power values [30]. Extreme learning machine (ELM) [31] and backpropagation (BP) neural network [32] are the most commonly used ANNs in wind power forecasting. However, due to the inherent randomness of initial network parameters such as weights, thresholds, and smoothing factors, it is challenging to obtain accurate and reliable predictions from ANN. Deep learning approaches generally achieve higher forecasting accuracy but require higher computational resources. Machine learning approaches are simpler to scale up. Intelligent approaches can improve energy effectiveness, reduce energy waste, and make instantaneous decisions in the wind energy sector [33–36]. Intelligent approaches have great potential in wind power forecasting. In addition, large language models (LLMs) are mainly designed for natural language processing and have the ability to capture long-term dependencies in sequences. Due to the sequential nature of textual and temporal data, recent research has extended LLMs to handle time series and spatiotemporal tasks [37].

Several researchers have provided reviews of wind power forecasting techniques over time. In Colak et al. [5] provided a brief overview of data mining and wind power forecasting to identify the top forecasting models for wind power at various time frames. The article found that ANN is suitable and efficient for short-term wind power forecasting; while, multilayer perceptrons are effective for medium to long-term wind power forecasting. In Wang et al. [38] provided a thorough overview of artificial intelligence algorithms in wind farms, summarizing the strengths and limitations of ANN, support vector machine (SVM), and seagull optimization algorithm (SOA) for wind farm applications. The article concluded that numerous applications for wind farms may be created using artificial intelligence algorithms. Hanifi et al. [39] provided a systematic classification of wind power forecastingmethods, which includes physical, statistical, and hybrid methods. The article also reviews factors that affect the accuracy and computation time of predictive modeling efforts. These three review articles reflect the application and development of ANN in wind power forecasting research over the last decade. However, they only address a single forecasting model and do not consider the latest processing and optimization techniques in hybrid models. Additionally, visualizing the historical development of the wind power forecasting problem in a single discussion can be challenging.

Over time, the quantity of published works on wind power forecasting has increased, with more and more researchers joining the field. To comprehensively summarize the literature and identify evidence on challenges, researchers often use a systematic approach known as a systematic review. In wind power forecasting, researchers have attempted to review the research area from various perspectives. For instance, Lu et al. [40] gave a general overview of how meta-heuristic algorithms are used in wind power forecasting and the difficulties they encounter. The article's major focus is on using meta-heuristic techniques to optimize important ANN parameters, such as initial weights and thresholds and smoothing factors, to enhance the accuracy of wind power predictions. Qian et al. [41] conducted a review and discussion of decomposition algorithms used in hybrid models, focusing on wavelet transform (WT) and empirical modal decomposition (EMD). This article also examines techniques employed in recent years to improve decomposition-based models' performance from three perspectives: feature selection, parameter optimization, and predictability analysis and filtering. Bokde et al. [42] focused on the advantages, timely growth, and future possibilities of hybrid EMD or ensemble empirical mode decomposition (EEMD) models for wind power forecasting. Wang et al. [43] presented a thorough analysis of the various deep learning methods applied to the prediction of wind speed and power, including the many steps of data processing, feature extraction, and relationship learning. The article identifies

three challenges to accurate wind speed/power predictions under complex circumstances, namely data uncertainty, incomplete features, and intricate nonlinear relationships. These four articles each focus on a specific aspect of wind power forecasting, such as parameter optimization or data decomposition preprocessing. However, they do not provide a comprehensive overview of the wind power forecasting framework, and as such, cannot fully illustrate current trends in wind power forecasting research.

The aim of this survey is to address the aforementioned gap in the literature. Unlike previous review articles, this article focuses on reviewing the wind power forecasting problem based on existing literature, establishing a systematic knowledge map of research in this area, and identifying the relevant data and findings. Additionally, the survey visualizes the reported results, which have not been done in earlier work. Moreover, a complete forecasting framework involves three essential phases: data pre-processing, model selection, and model optimization. This paper provides a comprehensive overview of hot intelligence approaches in keyword analysis from these perspectives, in the context of a complete process.

This paper focuses on handling the following research questions:

(1) What has been the progress in wind power forecasting research in the last two decades based on the collected literature, with regard to countries, institutions, journals, and authors? Answering this question aims to present a comprehensive understanding of the historical development of the wind power forecasting problem from multiple perspectives, such as time, technology, and research institutions.
(2) How are intelligent approaches and other emerging technologies (non-autoregressive and LLMs) being applied in current research, given their potential in wind power forecasting? This question aims to analyze and evaluate the methods represented by hot co-occurring and bursting keywords in the literature from multiple perspectives, and explore the development trend of wind power forecasting incorporating these approaches.
(3) What are the challenges and opportunities for intelligent approaches in wind power forecasting research, based on the current state of research? By synthesizing the knowledge map of the previous literature in this research area and the current state of research on intelligent approaches, answering this question can inform and open up ideas for future developments. This work provides a comprehensive review of the application of intelligent approaches, specifically machine learning and deep learning, in wind power forecasting. Sections 2, 3 and 4 address the research questions from different perspectives, providing valuable insights into the current state of research. Compared to previous studies, the contributions of this paper can be summarized in three key aspects:

- A novel production of a knowledge map of the literature on wind power forecasting research, including a scientometric analysis and visual presentation of the development status and trends.
- A systematic analysis of the application of intelligent approaches, which are the most important trends in research, from the perspectives of data pre-processing and parameter optimization.
- A detailed analysis of the challenges and future directions for research in wind power forecasting incorporating emerging technologies such as attention mechanisms, non-autoregressive, and LLMs, providing valuable insights for future developments in this field.

The paper is structured as follows: Sect. 2 presents the findings of the CiteSpace-based systematic knowledge mapping and scientometric statistical analysis. It includes a detailed explanation of co-occurring and bursting keywords analysis and highlights highly cited reference authors and journals. In Sect. 3, machine learning approaches, deep learning approaches, attention approaches, and other emerging technologies (non-autoregressive and LLMs) are further analyzed based on the keyword hotspots identified in Sect. 2. Section 4 discusses future trends and recommendations related to wind power forecasting research. Finally, the paper is concluded in Sect. 5. Table 1 gives some abbreviations used throughout the paper.

2 Knowledge map analysis

In this section, we address research question 1 by analyzing and visualizing the knowledge map of wind power forecasting literature in the Web of Science (WOS) database over the last 20 years. We compare the cooperation networks and development channels and examine the number of publications, frequency of citations, publications in journals, and influence of authors and institutions. We also study changes in co-occurring and bursting keywords to reveal shifting research trends.

To guarantee the veracity and credibility of our data, we sourced information from the WOS core database and searched for literature from the last two decades using subject terms such as "wind power prediction/forecasting"

Table 1 All abbreviations used in the article

Abbreviation	Full name
ALO	Ant lion optimizer
AR	Autoregressive
ARMA	Autoregressive moving average
ARIMA	Autoregressive integrated moving average
ANN	Artificial neural network
BERT	Bidirectional encoder representations from transformers
BiLSTM	Bidirectional long and short-term memory
BP	Back propagation
CNN	Convolutional neural network
CSO	Crisscross optimization algorithm
DNN	Deep neural network
DWT	Discrete wavelet transform
EEMD	Ensemble empirical mode decomposition
ELM	Extreme learning machine
EMD	Empirical mode decomposition
ESN	Echo state network
EWT	Empirical wavelet transform
FARIMA	Fractional autoregressive integrated moving average
GPT	Generative pre-trained transformer
GRU	Gated recurrent unit
GSA	Gravity search algorithm
GWO	Grey wolf optimization
HBO	Heap-based optimizer
LLM	Large language model
LSSVM	Least square support vector machine
LSTM	Long short-term memory
MAE	Mean absolute error
MSE	Mean square error
MAPE	Mean absolute percentage error
MTD	Mean trend detector
NREL	National renewable energy laboratory
NWP	Numerical weather prediction
PSO	Particle swarm optimization
RBF	Radial basis function
RMSE	Root mean square error
RNN	Recurrent neural network
SAE	Stacked auto-encoders
SCADA	Supervisory control and data acquisition
SGRU	Stacked gated recurrent unit
SMA	Slime mold algorithm
SOA	Seagull optimization algorithm
SSA	Sparrow search algorithm
SVM	Support vector machine
T5	Text-to-text transfer transformer
TCN	Tme convolution network
VAR	Vector autoregression
VMD	Variational mode decomposition
WD	Wavelet decomposition
WOS	Web of Science
WT	Wavelet transform

and "time series." We searched and screened the data, obtaining 222 articles for visualization and analysis.

2.1 Temporal analysis of literature production

The trend in the number of papers published in a particular research area is a reflection of the level of attention given to that field. In this study, we analyzed 222 pieces of literature and obtained a line graph depicting the number of papers published on wind power forecasting from 2000 to 2022. As shown in Fig. 1, the frequency of citations and the number of published articles have been increasing from 2006 to 2022. The field experienced a phase of slow growth followed by rapid growth, indicating that scholars have continued to prioritize research in this area. According to the statistics of the top three countries in terms of publication volume shown in Fig. 2 from 2006 to 2022, it is not difficult to see that scholars paid less attention to wind power forecasting research before 2014, with a relatively stable and low publication volume. During the period when stable generation sources like thermal, hydropower, and nuclear power were dominant, wind farms were relatively uncommon, which meant that accurate forecasting was not critical due to their low impact on grid power volatility.

From 2014 to 2017, there was a transitional phase in the global wind power development process, with the number of publications increasing steadily, particularly in research related to wind power forecasting. Following 2017, the global growth of installed wind power started to accelerate, leading to a rapid increase in both the number of publications and the frequency of citations. Furthermore, from 2020 to 2022, the number of publications in China surged rapidly. Figure 1 reveals that the change in citation frequency and literature production follows the same trend over time, with 1214, 1444, and 1976 citations in 2020, 2021, and 2022, respectively.

2.2 Analysis of journals and authors' publications and influence

The distribution of articles across journals is an important indicator of the research direction and impact. Table 2 reveals that Energies journal has the highest number of articles and the fifth highest citation frequency. Meanwhile,

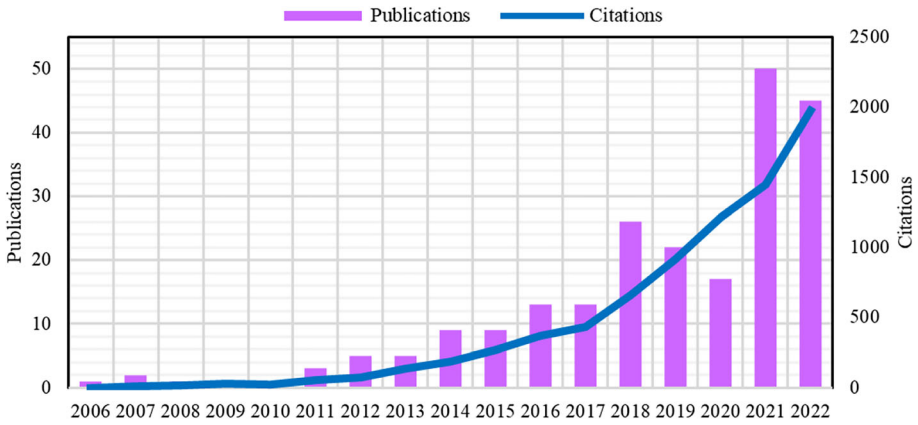


Fig. 1 Number of publications and citations of research on wind power forecasting from 2006 to 2022

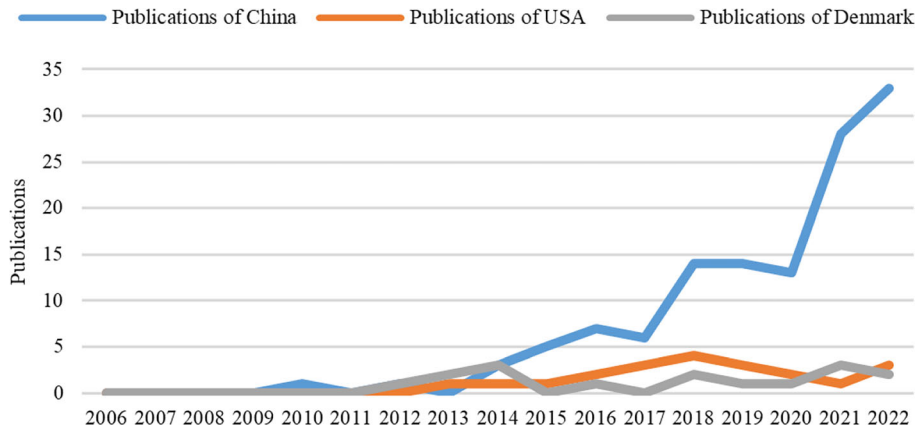


Fig. 2 Number of articles published by the top three countries

Table 2 Top ten journals in the number of publications on wind power forecasting

Journals	Number of publications	Total citations	Average citations per year
Energies	26	424	71.74
Renewable Energy	19	792	149.33
Applied Energy	15	903	153.91
Energy	14	506	94.97
IEEE Access	11	332	72.07
IEEE Transactions on Power Systems	10	1584	153.41
Energy Conversion and Management	10	328	57.89
IEEE Transactions on Sustainable Energy	7	237	38.36
Electric Power Systems Research	6	121	20.72
Journal of Modern Power Systems and Clean Energy	6	75	12.04

IEEE Transactions on Power Systems has a total of 1584 citations and an average of 153.41 citations, suggesting that the content of articles in these two journals aligns with the focus of the journal, and these articles are highly regarded. The high citation rate indicates a growing interest in wind power forecasting research, with topics ranging from energy policy to other relevant areas. The increasing intensity of development and construction also indicates support for relevant policies.

Regarding the top 10 authors with the most articles published on wind power forecasting research from 2000 to 2022 in Table 3, Pierre Pinson had the most published

articles and highest citations, with 10 and 778, respectively. P.K. Dash had the third-highest number of published articles and the second-highest citations, with 378.

2.3 Analysis of country and institution cooperation networks

The country collaboration network in Fig. 3 shows that China, the USA, Denmark, Iran, and Australia have larger nodes, indicating that they have more publications and occupy a leading position in wind power forecasting research. Among them, China (125), the USA (21), and Denmark (16) have significantly more publications than other countries and are ranked in the top three. Table 4 shows that these countries, China, the USA, and Denmark, are also ranked in the top three regarding the frequency of citations of their publications.

Taking into account the 310.629 GW of cumulative onshore wind power installed in China by the end of 2021, which exceeds that of the USA (134.354 GW) and Germany (56.814 GW), as presented in Table 5, the continuous growth of wind power installations serves as a significant driver for research enthusiasm and the level of research conducted in the field. Furthermore, Fig. 3 shows 48 nodes, 82 connections, and a connection strength of 0.0727, which indicates the initial formation of an initial collaborative network for inter-country wind power forecasting research.

The significance of a node in the field can be measured by its high betweenness centrality. China and Denmark have high betweenness centrality, indicating their central role in wind power forecasting research and being at the forefront of research development. Although the USA has more publications (21), its betweenness centrality is 0.08, suggesting that it conducts research independently and does not have close collaborations with other countries.

Figure 4 presents a collaboration map of highly productive institutions in wind power forecasting research. Node circles in the figure represent the number of publications of the institution; while, the number of connected lines indicates the degree of collaboration between the institutions. It is evident from the figure that the circles of all nodes are relatively small, implying that each organization has a limited number of outgoing publications. However, some institutions have larger nodes, indicating that they are high-producing institutions. Furthermore, the higher number of links between high-producing institutions compared to low-producing institutions suggests that high-producing institutions possess significant scientific research and professional authority. This allows them to attract cooperation and academic exchange from other institutions.

Figure 4 displays 207 nodes, 182 links, and a connection strength of 0.0085, indicating the initial formation of a collaborative network among research institutions. On the other hand, Table 6 reveals that Tech Univ Denmark holds the first position with 16 publications and 886 total citations. Notably, four out of the top five institutions in terms of publications are from China, indicating a significant contribution to the field of research from Chinese institutions.

2.4 Analysis of co-occurring keywords

Figure 5 displays a co-occurring keywords map of the wind power forecasting field, which consists of 329 nodes. Each node's size reflects how frequently it appears in the network. Among the top three keywords, "wind power forecasting," "neural network," and "time series" occur more than 40 times. The frequency of "neural network" and "time series" is relatively high, with 55 and 43 occurrences, respectively, indicating that neural network methods are the most common methods applied to time series data for wind power forecasting. After 2011, the keywords "empirical mode decomposition," "wavelet transform," "variational mode decomposition," and optimization algorithms continue to appear, showing that research in wind power forecasting no longer relies solely on

Table 3 Top ten authors in the number of publications on wind power forecasting

Journals	Number of publications	Total citations	Average citations per year
Pierre Pinson	10	778	75.99
Hao Chen	5	83	15.23
Jie Zhang	4	232	28.85
Henrik Madsen	4	256	48.33
P. K Dash	4	378	62.06
Cui Mingjian	3	41	20.5
Liu Hui	3	251	36.66
Dong Zhao Yang	4	173	20.16
Abd elaziz, Mohamed	3	41	20.5
Al-qaness Mohammed A. A	4	279	41.33

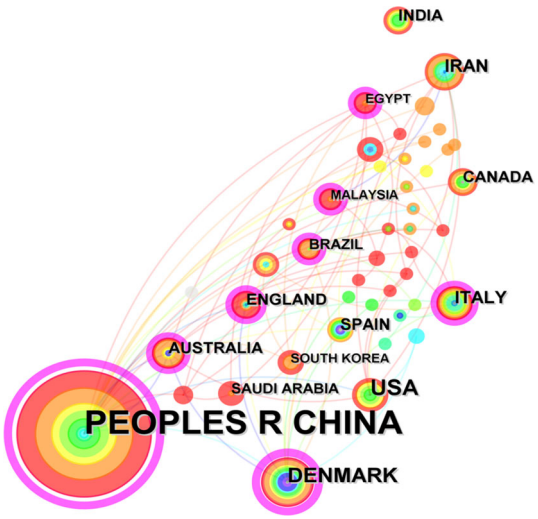


Fig. 3 Visualizing map of national cooperation network

Table 5 Top ten countries regarding the installed wind power in 2021

Country	Wind power installations of 2021/GW	Cumulative wind power installations to 2021/GW
China	30.670	310.629
USA	12.747	134.354
Germany	1.925	56.814
India	1.459	40.084
Brazil	3.830	21.580
France	1.192	19.131
Canada	0.677	14.255
UK	0.328	14.064

intelligent approaches. Instead, it focuses on improving forecasting models through data pre-processing and optimization to enhance prediction accuracy.

In recent years, data-driven artificial intelligence technologies have developed rapidly, with machine learning transitioning from shallow to deep learning. Wind power forecasting models based on artificial intelligence have emerged as a hot topic. Many studies have shown that artificial intelligence-based models exhibit superior forecasting performance compared to traditional statistical models. These models hold the potential to overcome the technical bottlenecks of wind power forecasting and achieve significant improvements in forecasting accuracy. We conducted statistical analysis on various machine learning algorithms used in the literature, and the results are shown in Fig. 6. Artificial intelligence-based wind power forecasting models include traditional machine learning models such as traditional ANN, SVM, and ELM, among others. With the advancement of artificial

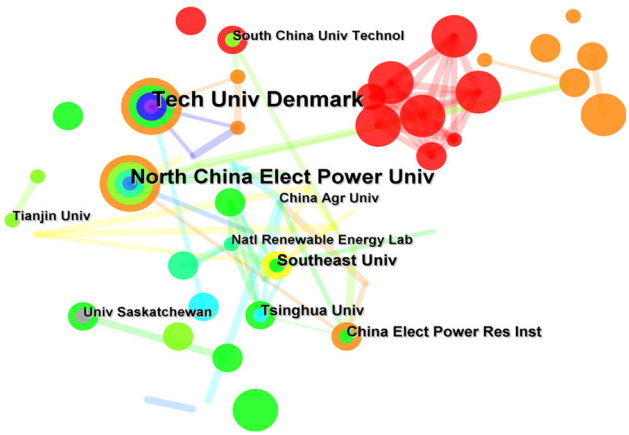


Fig. 4 Visualizing map of publishing institution cooperation network

intelligence, an increasing number of deep learning models, such as long short-term memory (LSTM), convolutional neural networks (CNN), auto-encoders, gated

Table 4 Top ten countries in the number of publications on wind power forecasting

Country	Year of first publication	Number of publications	Total citations	Average citations per year	Betweenness centrality
China	2010	125	3593	711.07	0.84
USA	2013	21	939	148.61	0.08
Denmark	2012	16	906	103.28	0.49
Iran	2014	12	430	64.52	0.1
Italy	2011	11	526	68.15	0.1
Canada	2006	9	381	54.44	0
Australia	2010	9	748	96.34	0.17
England	2014	8	137	33.91	0.15
Spain	2007	8	808	62.63	0

recurrent units (GRU), are being adopted, proving highly effective in addressing complex nonlinear problems.

2.5 Analysis of bursting keywords

Bursting keywords refer to keywords that experience a sudden surge in usage frequency within a specific timeframe, indicating the emergence of hotspots in research during that period. By analyzing the keywords used in wind power forecasting from 2000 to 2022, 20 burst keywords were identified as representing the current forefront of research, as shown in Fig. 7.

In the early years, wind energy and the market were the main factors considered in wind power forecasting research, including the linear variation of wind speed and temperature in different regions. After 2010, the time series approach became the dominant method. In 2016, the terms "artificial neural network" (particularly "extreme learning machine") and "support vector machine" gained attention as methods capable of fitting nonlinear relationships in wind power variation and achieving good forecasting results. In recent years, "deep learning" and "feature selection" have emerged as prominent keywords, with deep learning approaches widely developed and applied to wind power forecasting. Additionally, feature selection has been shown to further improve forecasting accuracy.

3 A review of hot spots in wind power forecasting research

In response to research question 2, we will delve into the research hotspots involved in co-occurring and bursting keywords analysis in Sects. 2.4 and 2.5, namely the intelligent approaches represented by machine learning and deep learning.

3.1 Machine learning approaches based on data pre-processing

Time series data are typically high-dimensional and data-intensive, and using them directly for input modeling can greatly affect the accuracy of predictions [44]. Therefore, data pre-processing is an essential step before machine learning modeling. Researchers have employed various techniques to further analyze time series data for enhancing forecasting accuracy.

According to [45], improved results in wind power forecasting can be obtained by decomposing the original time series data and reconstructing the forecasting model for the sub-series data. They suggested a combined wind power forecasting model using reconstructed data and dynamic adaptive weights, which performs well in short-term forecasting and requires a high degree of adaptability in long-term forecasting. Reconstruction of time series data by chaotic phase space not only achieves dimensionality reduction but also considers the chaotic characteristics in time series data of wind power. Wang et al. [46] also applied the reconstruction of time series data by chaotic phase space and analyzed typical chaotic time series forecasting models accordingly, and constructed a hybrid model using a regime-switching optimal mechanism. This forecasting model outperformed traditional artificial intelligence algorithms, with an average improvement of 11.91% for mean absolute error (MAE) and 35.38% for root mean square error (RMSE).

It is complicated how wind energy and physical weather conditions interact, resulting in inherent stochastic fluctuations, nonlinearity, and non-smoothness in wind power

Table 6 Top ten institutions in the number of publications on wind power forecasting

Institution	Year of first publication	Number of publications	Total citations	Average citations per year	Betweenness centrality
Tech Univ Denmark	2012	16	886	97.53	0.01
North China Elect Power Univ	2015	11	403	84.57	0.01
Southeast Univ	2014	6	156	29.56	0.01
China Elect Power Res Inst	2018	5	198	41.86	0.02
Tsinghua Univ	2016	5	258	38.6	0.02
Natl Renewable Energy Lab	2017	5	279	41.33	0.02
Univ Saskatchewan	2012	4	208	29.91	0
Tianjin Univ	2017	4	284	52.86	0
China Agr Univ	2016	4	184	28.16	0
South China Univ Technol	2019	4	144	32.4	0

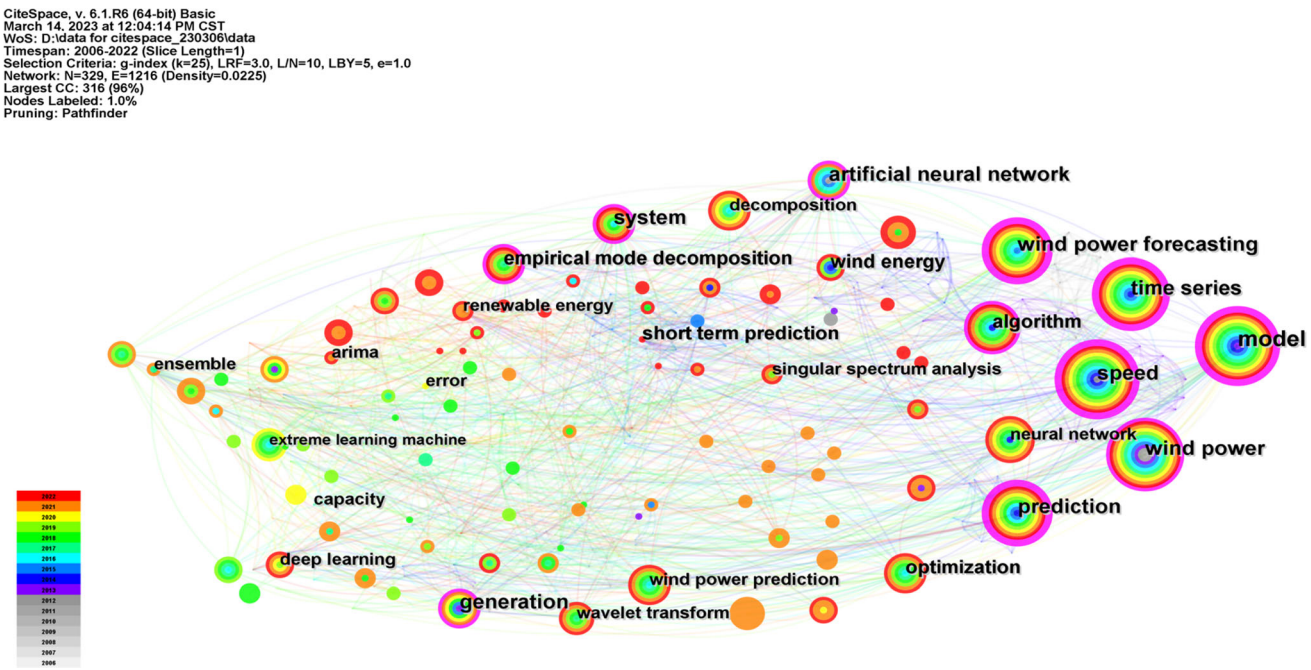


Fig. 5 Visualizing map of co-occurring keywords network

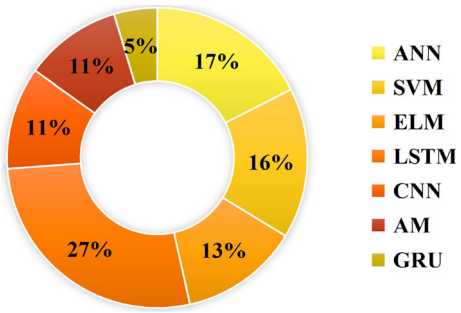


Fig. 6 Proportion of various algorithms in machine learning

data. As a result, predicting wind power with high precision is challenging. To improve forecasting accuracy, many researchers have utilized techniques such as WT [47, 48] or EMD to decompose the time series into multiple levels and then combine them for predictions. Decomposition of wind power data using post-frequency division has been shown to enhance forecasting accuracy. In [49], the authors proposed the use of wavelet decomposition (WD) to decompose the original wind power time series data into a number of sub-signals with improved behavior and contours. EEMD, as used in [50], is suitable for non-smooth and nonlinear signal analysis. EEMD can apply white noise extreme point mutation to improve decomposition accuracy compared to EMD and address the disadvantages of the EMD algorithm in terms of modal confusion phenomenon [51]. EEMD has been utilized to decompose chaotic wind power time series into eigenmodal signals with varying eigenscales, resulting in improved forecasting accuracy. The approach has been examined and evaluated using genuine Chinese wind power statistics [50].

To improve the analysis of wind power data, Zhang et al. [52] reconstructed non-stationary wind power time series through the complete integrated EMD Rumpel-Ziev complexity decomposition, obtaining stable and mutually independent subsequences. Variational modal decomposition (VMD), employed in [53, 54], is used to process the original time series signal, preventing interaction between different modes. Meanwhile, Wu et al. [55] adopted a mean trend detector (MTD) to decompose the time series into a mean trend component and a stochastic component. These two components have a clear physical meaning and can be interpreted as the trend and high-frequency fluctuations in the variability of the time series, respectively. By using this method, a more accurate and stable combined forecasting is obtained.

Based on a synthesis of 222 lature, the use of data pre-processing techniques in wind power forecasting studies has increased significantly from 2012 to 2022, as shown in Fig. 8. The number of published studies has increased steadily, particularly in 2021 and 2022. Figure 9 illustrates the proportion of data pre-processing-based models in all forecasting models in literature studies published from 2018 to 2022. It indicates that the share of wind power forecasting models based on data pre-processing models exceeded 50% in 2022, a 20% increase from 2018.

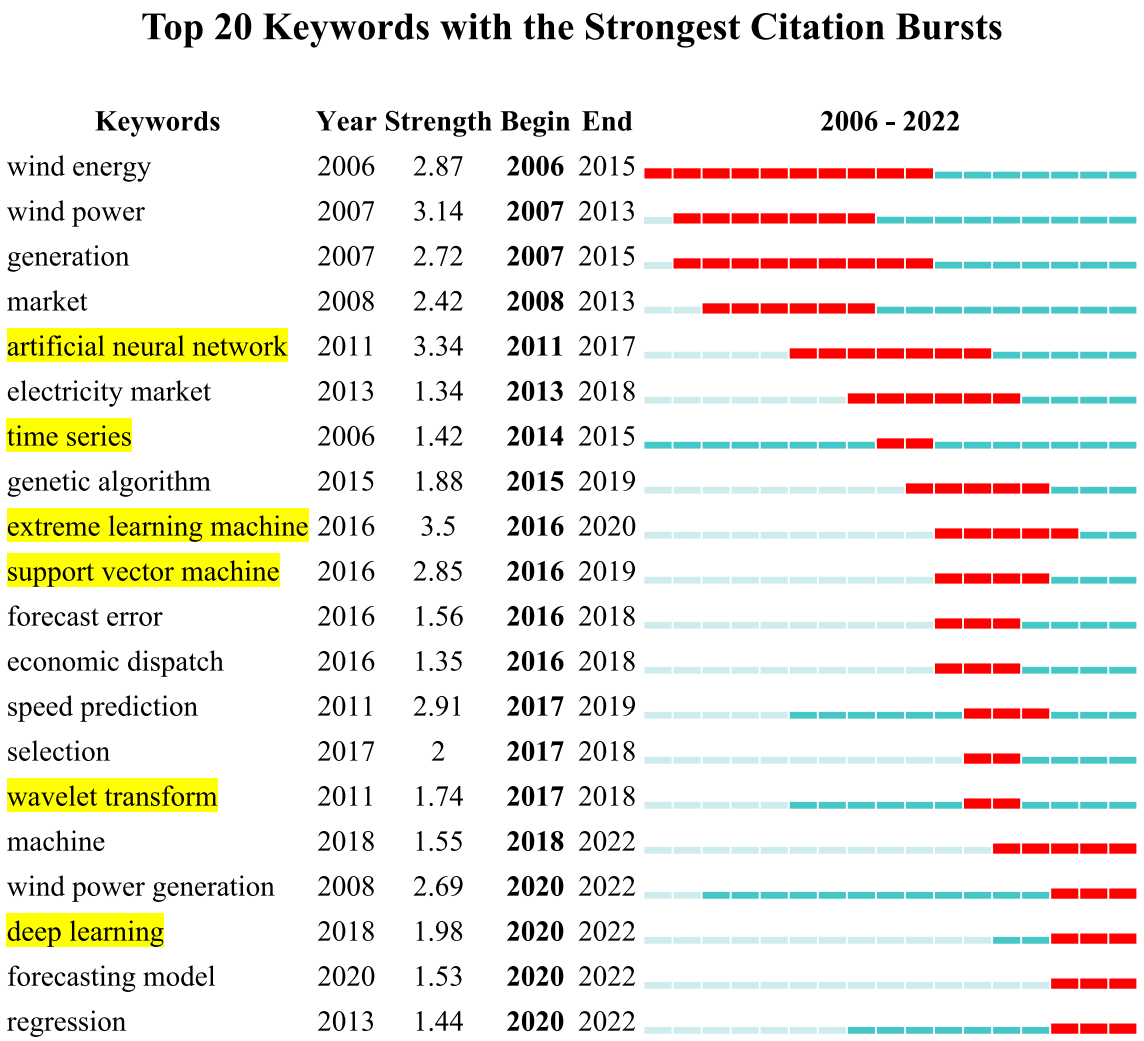


Fig. 7 Top 20 keywords with the strongest citation bursts

3.2 Machine learning approaches based on parameter optimization

Parameter optimization is a crucial step in selecting an optimal set of parameters for machine learning algorithms. The process of parameter selection and optimization strategies can enhance forecasting accuracy and decrease computing time by efficiently training the model [56].

ELM is a popularly applied ANN in wind power forecasting, including an input layer, a hidden layer, and an output layer. According to [40], ELM is the most popular ANN used in wind power forecasting. Although ELM uses initialization parameters to construct the model, it suffers from the disadvantage of redundant nodes in the hidden layer, which hinders the accuracy of the forecasting model to some extent. In their study, Zhang et al. [52] utilized ant lion optimization (ALO) to optimize ELM's parameters. They applied this method to wind power data collected from a wind farm in East China from March 2015 to May 2015. Their findings demonstrate that the suggested model performs well and is especially outstanding for 48-h advance wind power predictions. Meanwhile, Ding et al. [54] presented an improved ELM forecasting model using the gray wolf optimization (GWO) algorithm. The ELM forecasting model is proposed for the sub-time series obtained from the VMD decomposition and optimized with the GWO algorithm. The forecasting results showed that the suggested VMD-improved-GWO-ELM forecasting model performs better on the forecasting curves compared to ELM and BP models. More specifically, both the forward and backward prediction results have significant errors at some outliers, and they do not perform consistently due to the sensitivity of the optimization. In the research by Zhao et al. [57], a backward forecasting model based on ELM with a bidirectional mechanism was proposed to correct prediction errors by comparing the prediction results in both directions. This approach produced beneficial outcomes in medium-term wind power

forecasting ranging from 1 to 6 h. The bidirectional mechanism can also be extended to other algorithmic models such as SVM and gray models to improve their performance.

SVM is based on statistical learning theory, as proposed by [58]. SVM is not only useful for high-dimensional identification but also for forecasting in nonlinear and small-sample scenarios. By mapping wind power time series to a high-dimensional feature space through nonlinear mapping, SVM can be used in the regression problem of wind power forecasting and achieve good results [59]. For high-dimensional data spaces like wind power, the inner product function of SVM is usually replaced by a kernel function. In Wang et al. [60], a hybrid PSO-SVM-ARMA model was suggested to enhance forecasting performance by applying a Gaussian radial basis function (RBF) kernel function and optimizing the parameters with particle swarm optimization (PSO) algorithm. To address the problem of computational complexity in SVM, the least squares support vector machine (LSSVM) was suggested [61], which uses squared errors and equation constraints to transform quadratic programming problems into problems involving linear equations, lowering the amount of computing needed for model learning. An error-corrected LSSVM model was also proposed in [62] for wind power forecasting, which considers the error correction process of the model. The proposed method reduces the average absolute error by 52% compared with a single LSSVM forecasting model. In the studies by Wang et al. [50] and Li et al. [51], the LSSVM parameters are optimized using the improved sparrow search algorithm (SSA) during model construction. This optimization method leads to improved forecasting accuracy and faster convergence speed in the 24-h forecast. Meanwhile, Lu et al. [56] employed a new heuristic optimization algorithm, namely the gravity search algorithm (GSA), to optimize the hyperparameters of LSSVM. The resulting wind power forecasting model exhibits better regression forecasting accuracy when compared to other benchmark models in the validation of the Hebei province wind power dataset.

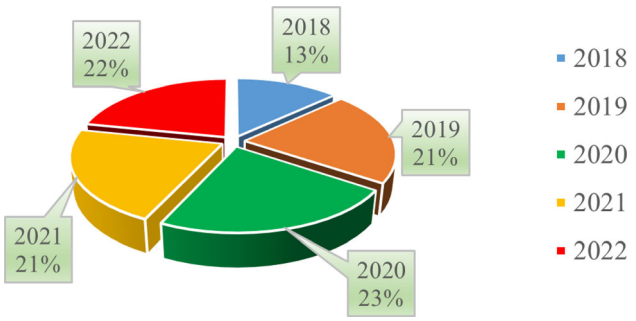


Fig. 9 Proportion of models based on data preprocessing in the total

3.3 Deep learning approaches

In practical applications, traditional machine learning models are usually used for short-term wind power forecasting under ideal conditions. However, in more complex conditions or for long-term forecasting, the mean absolute percentage error (MAPE) can reach between 10% and 17%, which may not satisfy the engineering specifications needed for typical renewable energy projects [63].

To address this issue, deep learning has emerged as a powerful tool with its automatic feature extraction and complex nonlinear relationship fitting capabilities, which have been widely applied in renewable energy exploitation, especially in wind power forecasting [43]. Deep learning models utilize a multilayer neural convolutional architecture with a gradient descent algorithm to minimize estimation errors and are applicable in a wide range of wind power forecasting scenarios [63]. Table 7 lists the top ten articles and models with the most cited frequency of deep learning methods in the reviewed literature. In Yu et al. [64], CNN was employed to simulate and predict spatiotemporal processes of wind farms. The mean square error (MSE) of wind power forecasting results of this

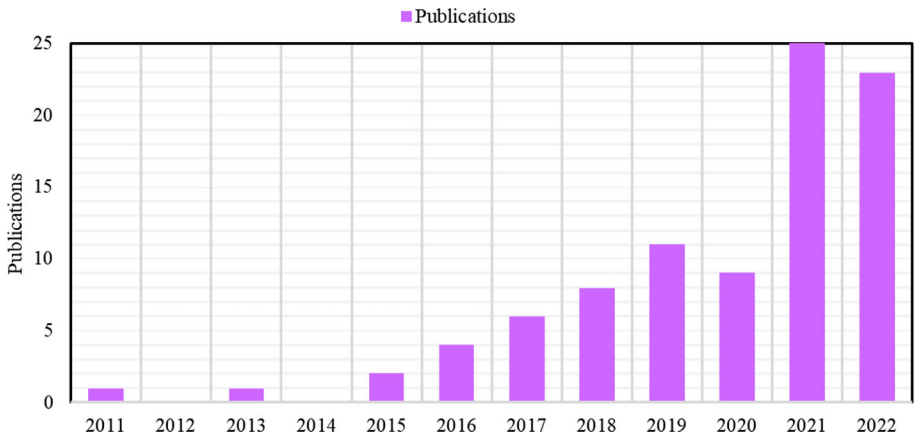


Fig. 8 Number of models based on data preprocessing in 222 publications

Table 7 Top ten articles in cited frequency on wind power forecasting based on deep learning

References	Model	Year of publication	Total citations
[16]	EMD-VMD-CNN-LSTM	2017	188
[70]	STSR-LSTM	2019	127
[33]	CNN-RBFNN	2019	116
[79]	BiLSTM-CNN	2018	72
[72]	GRU-EEMD-CSO	2021	39
[80]	BiLSTM-CNN	2019	39
[81]	CNN-RBFNN	2022	37
[68]	EWT-RKEML-SGRU-SMA	2021	37
[82]	SAE-BP	2022	35
[71]	LSTM-MRMI	2020	35

method is decreased by 49.83% compared with existing methods, and the model's training time is reduced by more than 150 times on average. In Shao et al. [65], the proposed model integrated the benefits of feature selection and recurrent neural network (RNN) by using sorted features as input data to obtain the forecasting results through deep CNN. Validated on the dataset of the national renewable energy laboratory (NREL), the accuracy of wind power forecasting is more than 10% higher than that of traditional methods. For wind power forecasting, Lin et al. [63] developed a feature-based learning model and applied a multi-layer neural convolution architecture time convolution network (TCN) with a gradient descent algorithm to minimize the estimation model error.

LSTM is an enhanced architecture over traditional RNN, which can overcome its limitations in long-term dependency problems and learn dependency information for a long time [66–68]. The addition of a special hidden unit called a memory unit helps to mitigate the gradient disappearance problem and leads to a significant reduction in time series forecasting errors. In Lopez et al. [69], LSTM was combined with an echo state network (ESN), resulting in improved forecasting results compared to traditional models. Furthermore, Yu et al. [70] proposed an optimized LSTM-enhanced forget-gate for wind power forecasting. The model filters the data by clustering to enhance wind power feature extraction through the correlation of turbine data. Devi et al. [71] employed data pre-processing and algorithmic hyperparameter optimization to improve the performance of the LSTM network structure. The resulting LSTM model, which combined EEMD and cuckoo search algorithms, exhibited superior forecasting accuracy compared to BP, SVM, and the original LSTM when tested on given time series data. In another study, Neshat et al. [72] utilized a hybrid optimization approach that combined adaptive differential evolution techniques with sine cosine optimization to optimize the LSTM's hyperparameters. The resulting hybrid model not only demonstrated improvements in forecasting accuracy but also reduced prediction duration. Meanwhile, Liu et al. [73] proposed a simple forecasting method for wind power using discrete wavelet transform (DWT) and LSTM. The study by Ewees et al. [67] proposed a method that utilizes LSTM to divide wind time series data into sub-series and uses the meta-heuristic optimization algorithm heap-based optimizer (HBO) to train the LSTM. The comparison results with existing models indicate that HBO can enhance the predictive performance of the underlying LSTM model.

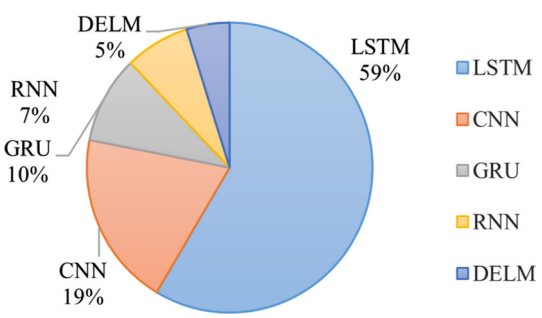


Fig. 10 The proportion of various algorithms in deep learning

As the deep neural network's depth increases, the performance usually improves. However, when the network size exceeds a certain threshold, overfitting may occur, which can negatively impact the overall performance of the DNN [68, 74, 75]. To address this issue, Schuster et al. [76] proposed bidirectional long and short-term memory (BiLSTM), which enables the training in two directions to improve sequence learning's performance [77]. In the paper by Liu et al. [78], a BiLSTM model was utilized and was found to outperform traditional LSTM networks regarding forecasting accuracy. The simulation results demonstrated that the BiLSTM neural network model got a forecasting accuracy improvement of 10.25%, 6.71%, and 12.18% over the LSTM model.

An extensive analysis and synthesis of 222 papers focused on deep learning approaches in wind power forecasting revealed that 44 papers used LSTM and its

Table 8 Recent articles on Transformer and improvements for time series forecasting

References	Year of publication	Model	Improvements
[92]	2023	Crossformer	Crossformer, a Transformer-based model utilizing cross-dimension dependency for MTS forecasting
[93]	2022	Quaterformer	Quaternion transformer, proposing decoupling learning-to-rotate attention with global memory to reduce the computational complexity from quadratic to linear
[94]	2022	FEDformer	FEDformer, proposing to combine transformer with the seasonal-trend decomposition method to capture the global profile of time series
[95]	2022	TACTiS	Transformer-Attentional copulas for time series (TACTiS), estimating joint distributions using an attention-based decoder
[96]	2022	TCCT	Tightly-coupled convolutional transformer (TCCT), proposing the idea of tightly-coupled convolutional Transformer and three TCCT architectures
[97]	2021	Informer	Informer, an efficient transformer-based model for LSTF with a ProbSparse self-attention mechanism, the self-attention distilling, the generative style decoder
[85]	2021	Autoformer	Autoformer, a novel decomposition architecture with an Auto-Correlation mechanism
[98]	2021	Pyraformer	Pyraformer, a Transformer-based model exploring the multi-resolution representation of the time series
[99]	2021	TFT	Temporal fusion transformer (TFT), a novel attention-based architecture that combines high-performance multi-horizon forecasting with interpretable insights into temporal dynamics
[100]	2021	Probabilistic transformer	Probabilistic transformer, proposing deep probabilistic methods that combine state-space models with transformer architectures

improved variants, making it the most commonly used method, followed by CNN. As shown in Fig. 10, LSTM has shown remarkable results in wind power forecasting.

In addition, the main characteristics of time series, i.e., seasonality or periodicity and trend patterns [83], can be efficiently handled by Transformers. Transformers exhibit strong modeling capability for long-distance dependencies and interactions in time series data, thus having high application potential in long-term forecasting [84]. In the research of [85, 86], incorporating sequence periodicity or frequency processing into the time series Transformer can significantly improve performance.

To effectively solve the long-distance information extraction and storage in sequence-to-sequence tasks, researchers have developed the Transformer model [87]. It breaks through the limitation of serialization computation based on the RNN model using parallel matrix operation and achieves high performance in the field of machine translation. Zheng et al. [88] combined Transformer and VMD to validate the effectiveness of the forecasting model in two actual datasets of Chinese wind farms. Wang et al. [89] used three improved encoder–decoder architectures for multi-step wind power forecasting. After comparison, the Transformer performs the best in forecasting accuracy and training efficiency. In wind power data from 12 regions in China, compared to existing models, the Transformer model reduced RMSE by 21.8% and 16.3% in 9-step and 3-step predictions, respectively. Xiao et al. [90] systematically established a hybrid wind power probability density forecasting method with a Transformer network, combining expected regression and kernel density estimation. The proposed method is better than the mainstream RNN model and can more accurately capture the uncertainty in wind power prediction. Huang et al. [91] combined deep learning with Transformer and utilized the unique network structure of Transformer to achieve more effective feature extraction for time series data. Compared with the LSTM forecasting model, the accuracy and efficiency of the forecasting results have been improved. We have collected the latest literature and summarized the latest improvements based on Transformers, as shown in Table 8.

3.4 Other emerging technologies

3.4.1 Non-autoregressive models

When discussing generative models in machine learning, we often think of autoregressive models that use time series for forecasting, such as the ARMA [101] model in load forecasting. These models are known for their ability to generate sequence data incrementally. The advantage of autoregressive generative models is that the generated data are typically well-structured and coherent in terms of syntax. However, a drawback is their slower generation speed because each element must wait for the previous one

Table 9 Non-autoregressive models and its input data variables for wind power forecasting

References	Input data variables	Recording intervals	Models used in the studies
[105]	Wind power, wind speed, wind generator speed, voltage and current in all phases	1-sec	focus time-delay neural network model
[106]	Wind power, wind speed, wind direction, rotor speed, generator torque, blade pitch angle	10-sec	Neural network model, boosting tree, random forest, support vector machine, k-nearest neighbor algorithm
[107]	Wind power, wind speed, wind direction	10-min	Third-order regression model, first-order artificial neural network model
[103]	LV active power, wind speed, wind direction, theoretical power curve	10-min	Homogenized Point Mutual Information and Deep Quantum Reinforced
[108]	Wind power, wind speed, wind direction, air pressure, temperature, humidity	0.5-h	Neuro-fuzzy network model
[109]	Wind power, wind speed, wind direction, air density, potential temperature	1-h	Support vector machine regression, multilayer perceptron, radial basis function, classification and regression tree, random forest
[110]	Wind power, wind speed, wind direction	1-h	Infinite impulse response multilayer perceptron, local activation feedback multilayer network model, diagonal recurrent neural network model

to be generated before proceeding. In contrast, non-autoregressive generative models differ from traditional autoregressive models in that they do not need to generate data sequentially; instead, they can generate the entire sequence simultaneously, resulting in faster generation speeds. This is typically achieved by parallelizing all generation steps.

Non-autoregressive models for wind power forecasting are based on multivariate modeling. Considering that the wind process is influenced by various factors in the atmosphere, more variables may provide additional information for wind power forecasting [45]. Peng et al. [102] utilized a combination of historical wind power data and features extracted and obtained from NWP through encoders and attention layers as composite inputs. The NWP data comprises 24 wind farm features including temperature, momentum flux, and wind direction. Jency et al. [103] employed a homogeneous point mutual feature selection model to design robust wind power forecasting, utilizing input wind turbine data (low-pressure active power, wind speed at turbine hub height, theoretical power curve, wind direction) and related features. Ouyang et al. [104] proposed a novel combined multivariate model, with a research case employing nine variables including wind power generation, and temperature, among others, to enhance wind power forecasting performance. Table 9 summarizes the various input data variables used in non-autoregressive models for wind power forecasting based on recent publications.

3.4.2 Large language models

Recently, in the field of deep learning, LLMs has rapidly developed, bringing changes and opportunities to various disciplines. LLMs are considered to be one of the key technologies for future general artificial intelligence. Initially developed for processing natural language, LLMs leverage large-scale data to learn the structure, semantics, and contextual information of language, enabling them to generate high-quality language outputs based on given instructions. In the context of wind power forecasting, LLMs could potentially be used for the analysis of high-dimensional data related to wind power, thereby enhancing predictions for long sequences and complex scenarios.

LLMs demonstrate powerful capabilities in understanding natural language and solving complex tasks through text generation. Currently, LLMs primarily utilize the Transformer [87] architecture. Bidirectional encoder representations from transformers (BERT) [111] is a deep bidirectional (bidirectional refers to considering context information simultaneously in both directions) representation learning model based on the Transformer architecture. Generative pre-trained transformer (GPT) is a unidirectional Transformer decoder model. Text to Text Transfer Transformer (T5) [112] is a model constructed based on the Transformer's encoder and decoder architecture. Recent studies have shown that LLMs perform excellently in time series analysis tasks, such as prediction, classification, anomaly detection, interpolation, few-shot learning, and zero-shot learning. For instance, Cao et al. [113] utilized GPT as the basis for time series forecasting. Jin et al. [114]

Table 10 A summary of wind power datasets

Dataset	Country	Time range	Interval
SCADA datasets			
Beberibe https://doi.org/10.5281/zenodo.1475197	Brazil	08.2013–07.2014	10 min
Pedra do Sal https://doi.org/10.5281/zenodo.1475197	Brazil	08.2013–07.2014	10 min
Penmanshiel https://doi.org/10.5281/zenodo.5946808	United Kingdom	01.01.2016–01.07.2021	10 min
Delabole https://data.dtu.dk/articles/dataset/Scada_data_from_Delabole_wind_farm/14077004	Great Britain	01.05.1993–30.04.1994	10 min
Kelmarsh https://doi.org/10.5281/zenodo.5841834	United Kingdom	01.01.2016–01.07.2021	10 min
	France	01.01.2013–13.01.2018	10 min
	Denmark	20.01.1988–18.01.1993	10 min
Eolos Wind Research Station https://doi.org/10.13020/1etn-1q17	United States	01.01.2017–31.12.2017	10 min
Aggregated datasets			
Elexon https://www.elexon.co.uk/documents/training-guidance/bsc-guidance-notes/bmrs-api-and-data-pushuser-guide-2/	Great Britain	Since 2011	30 min
New Zealand Datahttps://www.emi.ea.govt.nz/Wholesale/Datasets/Generation/Generation_MD/	New Zealand	Since 1997	30 min
South African Datahttp://redis.energy.gov.za/electricity-production-details/	South Africa	Since 2015	1 h
ENTSO-E https://transparency.entsoe.eu	European Member States	Since 2011	30 min or 1 h
EEM2020 https://pureportal.strath.ac.uk/en/datasets/data-and-code-for-the-eem2020-wind-power-forecasting-competition/	Sweden	01.01.2000–31.12.2001	1 h
Gefcom2014 https://www.dropbox.com/s/pqenrr2mcvl0hk9/GEFCom2014.zip?dl=0/	Australia	01.01.2012–30.09.2012	1 h

Table 11 Evaluation indicators of the forecasting model

Index	Equation	References
Deterministic prediction		
MAE	$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N} \lvert y_{\mathrm{pre(i)}} - y_{\mathrm{act}(i)} \rvert$	[31, 116, 117]
MSE	$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N} \left(y_{\mathrm{pre(i)}} - y_{\mathrm{act}(i)}\right)^2$	[70, 118–120]
RMSE	$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(y_{\mathrm{pre(i)}} - y_{act(i)}\right)^2}$	[31, 79, 80, 121]
MAPE	$\mathrm{MAPE} = \frac{1}{N}\sum_{i=1}^{N} \frac{\lvert y_{pre(i)} - y_{act(i)} \rvert}{y_{act(i)}}$	[79, 122, 123]
NMAE	$\mathrm{NMAE} = \frac{1}{N}\sum_{i=1}^{N} \lvert \frac{y_{\mathrm{pre(i)}} - y_{\mathrm{act(i)}}}{s} \rvert$	[52, 57, 124, 125]
NRMSE	$NRMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \left(\frac{y_{pre(i)} - y_{act}(i)}{s}\right)^2}$	[57, 123, 124]
Probabilistic prediction		
PICP	$\mathrm{PICP} = \frac{1}{N}\sum_{i=1}^{N} e_i, e_i = \begin{cases} 1, y_{\mathrm{pre(i)}} \in [L_i, U_i] \\ 0, y_{\mathrm{pre(i)}} \notin [L_i, U_i] \end{cases}$	[53, 126, 127]
PINAW	$\mathrm{PINAW} = \frac{1}{RN}\sum_{i=1}^{N}(U_i - L_i), R = y_{pre(\max)} - y_{pre(\min)}.$	[53, 126, 127]
CWC	$\mathrm{CWC} = PINAW\left(1 + \gamma e^{-\eta(PICP-\mu)}\right), \gamma(PICP) = \begin{cases} 0, PICP \geq \mu \\ 1, PICP \prec \mu \end{cases}.$	[102, 126, 128]

N is the number of samples, and y_{pre} and y_{act} are the predicted and actual values

s is the operating capacity

U_i and L_i are the upper and lower bound of the prediction interval

η and μ are the hyperparameters for the magnitude of the coverage width criterion

Table 12 A summary of wind power forecasting methods from different perspectives

Strength	Weakness
Machine learning approaches based on data pre-processing	
• Improved data quality for forecasting	• Possible loss of useful information
Machine learning approaches based on parameter optimization	
• Effectively explained the selection of model parameters	• Challenge in comparing various optimization strategies
• Higher forecasting accuracy of models	• Excessively fitting due to parameters optimization
Deep learning approaches	
• More accurate forecasting than shallow machine learning	• Massive complexity in computation
• No need in unsupervised learning or data preprocessing	• Difficulty in applying proper improvement methods

reprogrammed LLMs for time series forecasting while keeping the underlying language model intact. Comprehensive evaluations indicate that the proposed TIME-LLM is a robust time series learner, outperforming state-of-the-art specialized forecasting models. Zhou et al. [115] applied the GPT2 model to key tasks in time series analysis such as few-shot learning, full-data learning, and zero-shot learning. The authors' research demonstrates that the proposed LLM method performs comparably or better than state-of-the-art methods in nearly all time series tasks. LLMs have the ability of universality and logical reasoning, opening up new research directions for time series forecasting, particularly for wind power forecasting.

3.5 Open-source wind power datasets

Open-source datasets play a crucial role in wind power research. However, many studies in the current wind power forecasting literature rely on unique datasets that cannot be replicated by other researchers. To improve standardization in the field and create a uniform evaluation standard, different researchers need to construct forecasting models and compare their forecasting results on the same dataset.

Wind power data are typically generated by individual turbines; while, a large number of monitoring control and data acquisition (SCADA) data are collected by the turbines. An aggregated dataset is a summary of wind turbine data within a wind farm, which generally does not contain specific information such as SCADA, but provides overall wind power data for the entire wind farm, typically at 30-minute to 1-hour intervals. Table 10 provides a summary of some open-source wind power datasets.

3.6 Evaluation indicators of forecasting model

Model evaluation is essential for gauging a model's ability to accurately forecast wind power. It serves as the final step in the modeling process, aiding in identifying the best model to fit the data and estimating its future performance. Table 11 presents a compilation of commonly used evaluation metrics in wind power forecasting.

4 Discussion and future challenges

In answer to research question 3, we have compiled the current research status and knowledge graph of intelligent methods in wind power forecasting to identify the latest development trends and provide future suggestions for forecasting methods. A brief comparison summary of the forecasting techniques is presented in Table 12.

Our analysis of journal publications indicates that because of the irregular and intermittent nature of wind power time series data, many machine learning models rely on data preprocessing techniques to enhance forecasting accuracy. Data preprocessing techniques involve analyzing and processing time series data through techniques such as decomposition, feature de-noising, or feature extraction. By transforming the original time series data into multiple sequences or matrices, the models can identify more distinct features, thus improving their accuracy in predicting wind power.

As previously mentioned, the most commonly used data preprocessing methods are WD and mode decomposition for wind power forecasting. Among the modal decomposition family methods, EMD, VMD, and EEMD are widely used. Our review found that applying appropriate data preprocessing techniques to process time series data before machine learning modeling leads to stronger predictability of data obtained by different preprocessing methods than the original data, which might enhance prediction accuracy.

In addition, parameter optimization is another common practice in machine learning approaches that can improve the performance of forecasting models. For instance, ELM is the most popularly used ANN in wind power forecasting

problems. Research has focused on applying optimization methods to improve crucial parameters such as thresholds and starting weights. The learning capacity of the sample characteristics is only somewhat strong since the neural network is a shallow model. After reviewing various studies, it can be concluded that the application of optimization methods can enhance the forecasting accuracy of almost every machine learning model. However, it is significant to mention that comparing different optimization algorithms can be challenging due to variations in parameters such as population size, the number of iterations, and the number of time scales utilized for wind power forecasting.

Deep learning approaches have shown to be highly effective in capturing complex nonlinear relationships and extracting advanced feature representations from large-scale data. Deep learning models may efficiently extract dynamic characteristics and potential variables from intricate network topologies by thoroughly exploring high-dimensional state spaces. Prediction accuracy can be significantly increased by these elements. The current generation of deep learning methods for wind power forecast comprises CNN, RNN, and their two main variants, LSTM and GRU. Among these models, LSTM networks are capable of predicting and processing time series data with temporal delays and moderate interval lengths thus achieving accurate predictions. As a result, LSTM has the most widespread application in wind power forecasting.

Other emerging technologies have also opened up new research directions in wind power forecasting. The attention mechanism of Transformers can effectively model long-term dependencies in time series data. Capturing the seasonality and periodicity of time series data gives Transformers a significant advantage in long-term wind power forecasting. LLMs were primarily developed for natural language processing but can capture the features of time series data by adjusting the way they learn from textual data. Expanding LLMs for wind power forecasting aligns with two key trends in technological development. Firstly, there is a trend toward greater flexibility, aiming to reduce model size and computational costs. The varying scales of wind farms and fluctuations in the electricity market demand high flexibility in the application of LLMs. Secondly, there is a shift toward multimodality. LLMs need to have a universal interface supporting multimodal input and output to meet the requirements of diverse applications involving different types of wind power and meteorological data.

The following opportunities for future work in wind power forecasting research are outlined below:

(1) Data pre-processing is essential for wind power forecasting. As there is no one-size-fits-all optimal method for data pre-processing, it is worth exploring the application of suitable pre-processing methods in different models to improve forecasting accuracy.
(2) Machine learning algorithms based on parameter optimization are widely used. However, selecting the most suitable optimization algorithm for different machine learning models is critical in wind power forecasting. Hybrid models using data pre-processing and parameter optimization are also common.
(3) Deep learning-based neural networks, such as CNN, RNN, LSTM, and GRU, outperform general machine learning methods in wind power forecasting problems. These algorithms can achieve more accurate predictions because of their capacity to simulate intricate nonlinear functions. However, it is crucial to remember that these methods can be computationally heavy, and proper feature selection, data cleaning and optimizers, and network selection are crucial to enhancing the performance of these models.
(4) For long-term time series data, wind power forecasting is difficult, with forecasting accuracy decreasing as the time scale increases. The success of Transformers and LLMs in natural language processing indicates their strong potential for time series forecasting. However, applying these emerging technologies to complex wind power forecasting requires further refinement and adjustment.

5 Conclusion

Since the turn of the 21st century, wind power has become an increasingly significant source of energy, and wind power forecasting techniques have been fully developed. In this article, we aim to address three research questions related to wind power forecasting by combing through relevant literature and research methods. Through the use of knowledge graph visualization and scientometric analysis, we provide a broad overview of the wind power forecasting problem.

Firstly, we conduct a bibliometric analysis of articles on wind power forecasting in the WOS core collection over the last 20 years using CiteSpace software. By counting the number of publications and collaborations by country, institution, journal, and author, and creating visual knowledge maps of the literature on wind power forecasting research, including a scientometric analysis and visual presentation of the development status and trends, we present the research situation of different institutions and authors in wind power forecasting by country clearly and concisely, providing researchers with valuable data references and insights.

Secondly, co-occurrence mapping of keywords in the literature and analysis of keyword bursts reveal the popular research directions in wind power forecasting, that's machine learning and deep learning. The paper also provides further information and insight into the state of research in the literature on the research methods represented by some of the highlighted words. Machine learning algorithms are simple and easy to use and are often used in conjunction with data pre-processing and parameter optimization methods in applications. On the other hand, deep learning models can explore the deeper information of the data, and appropriate models need to be selected for different wind power time series data. In addition, some emerging technologies, such as Transformers and LLMs, have shown strong modeling abilities for long-range dependencies and interactions in time series data, deserving further studying its application in wind power forecasting.

Finally, it is recommended that future wind power forecasting studies should take into account the observations given in Sect. 4 to get better results.

Acknowledgements The work is supported by the National Natural Science Foundation of China under Grant 61873130 and Grant 61833011, the Natural Science Foundation of Jiangsu Province under Grant BK20191377, the 1311 Talent Project of Nanjing University of Posts and Telecommunications, and "Chunhui Program" Collaborative Scientific Research Project (202202004). This work also is supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP22H03643, the Next Generation (SPRING) under Grant JPMJSP2145, and JST through the Establishment of University Fellowships toward the Creation of Science Technology Innovation under Grant JPMJFS2115.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Wang J, Botterud A, Bessa R, Keko H, Carvalho L, Issicaba D et al (2011) Wind power forecasting uncertainty and unit commitment. Appl Energy 88(11):4014–4023
2. Liu H, Tian HQ, Chen C, Li YF (2010) A hybrid statistical method to predict wind speed and wind power. Renew Energy 35(8):1857–1861
3. Cadenas E, Rivera W (2009) Short term wind speed forecasting in La Venta, Oaxaca, México, using artificial neural networks. Renew Energy 34(1):274–278
4. Tascikaraoglu A, Uzunoglu M (2014) A review of combined approaches for prediction of short-term wind speed and power. Renew Sustain Energy Rev 34:243–254
5. Colak I, Sagiroglu S, Yesilbudak M (2012) Data mining and wind power prediction: a literature review. Renew Energy 46:241–247
6. Foley AM, Leahy PG, Marvuglia A, McKeogh EJ (2012) Current methods and advances in forecasting of wind power generation. Renew Energy 37(1):1–8
7. Catalão JDS, Pousinho HMI, Mendes VMF (2011) Short-term wind power forecasting in Portugal by neural networks and wavelet transform. Renew Energy 36(4):1245–1251
8. Du P, Wang J, Yang W, Niu T (2019) A novel hybrid model for short-term wind power forecasting. Appl Soft Comput 80:93–106
9. Jahangir L, Babbar S (2020) Medium term wind speed using random forest algorithm. Int Res J Comput Sci Technol, pp 47–53
10. Wu GZ, Zhang YB, Su C, Liu YJ (2013) Study on medium-term and short-term wind power forecasting methods. Appl Mech Mater 361:318–322
11. Sankar S, Amudha S, Madhavan P, Lamba DK (2021) Energy efficient medium-term wind speed prediction system using machine learning Models. In: IOP Conference Series: Materials Science and Engineering, vol. 1130. IOP Publishing, p 012085
12. Ouyang T, Huang H, He Y (2019) Ramp events forecasting based on long-term wind power prediction and correction. IET Renew Power Gener 13(15):2793–2801
13. Colak I, Sagiroglu S, Yesilbudak M, Kabalci E, Bulbul HI (2015) Multi-time series and-time scale modeling for wind speed and wind power forecasting part II: Medium-term and long-term applications. In: 2015 International Conference on Renewable Energy Research and Applications (ICRERA), IEEE, p. 215–220
14. Azad HB, Mekhilef S, Ganapathy VG (2014) Long-term wind speed forecasting and general pattern recognition using neural networks. IEEE Trans Sustain Energy 5(2):546–553
15. Borunda M, Rodríguez-Vázquez K, Garduno-Ramirez R, de la Cruz-Soto J, Antunez-Estrada J, Jaramillo OA (2020) Long-term estimation of wind power by probabilistic forecast using genetic programming. Energies 13(8):1885
16. Feng C, Cui M, Hodge BM, Zhang J (2017) A data-driven multi-model methodology with deep feature selection for short-term wind forecasting. Appl Energy 190:1245–1257
17. Landberg L (1999) Short-term prediction of the power production from wind farms. J Wind Eng Ind Aerodyn 80(1–2):207–220
18. Lei M, Shiyan L, Chuanwen J, Hongling L, Yan Z (2009) A review on the forecasting of wind speed and generated power. Renew Sustain Energy Rev 13(4):915–920
19. Andrade JR, Bessa RJ (2017) Improving renewable energy forecasting with a grid of numerical weather predictions. IEEE Trans Sustain Energy 8(4):1571–1580

20. Ramirez-Rosado IJ, Fernandez-Jimenez LA, Monteiro C, Sousa J, Bessa R (2009) Comparison of two new short-term wind-power forecasting systems. Renew Energy 34(7):1848–1854
21. Hill DC, McMillan D, Bell KR, Infield D (2011) Application of auto-regressive models to UK wind speed data for power system impact studies. IEEE Trans Sustain Energy 3(1):134–141
22. Dowell J, Pinson P (2015) Very-short-term probabilistic wind power forecasts by sparse vector autoregression. IEEE Trans Smart Grid 7(2):763–770
23. Flores JJ, Graff M, Rodriguez H (2012) Evolutive design of ARMA and ANN models for time series forecasting. Renew Energy 44:225–230
24. Yuan C, Liu S, Fang Z (2016) Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM (1, 1) model. Energy 100:384–390
25. Sadaei HJ, Enayatifar R, Guimarães FG, Mahmud M, Alzamil ZA (2016) Combining ARFIMA models and fuzzy time series for the forecast of long memory time series. Neurocomputing 175:782–796
26. Hua K, Mishra Y, Ledwich G (2015) Fast unscented transformation-based transient stability margin estimation incorporating uncertainty of wind generation. IEEE Trans Sustain Energy 6(4):1254–1262
27. Bludszuweit H, Domínguez-Navarro JA, Llombart A (2008) Statistical analysis of wind power forecast error. IEEE Trans Power Syst 23(3):983–991
28. Jiang P, Liu Z, Niu X, Zhang L (2021) A combined forecasting system based on statistical method, artificial neural networks, and deep learning methods for short-term wind speed forecasting. Energy 217:119361
29. Chitsazan MA, Fadali MS, Trzynadlowski AM (2019) Wind speed and wind direction forecasting using echo state network with nonlinear functions. Renewable Energy 131:879–889
30. Karabacak K, Cetin N (2014) Artificial neural networks for controlling wind-PV power systems: A review. Renew Sustain Energy Rev 29:804–827
31. Wan C, Xu Z, Pinson P, Dong ZY, Wong KP (2013) Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans Power Syst 29(3):1033–1044
32. Sun W, 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 Manage 157:1–12
33. Hong YY, Rioflorido CLPP (2019) A hybrid deep learning-based neural network for 24-h ahead wind power forecasting. Appl Energy 250:530–539
34. Liu Y, Qin H, Zhang Z, Pei S, Jiang Z, Feng Z et al (2020) Probabilistic spatiotemporal wind speed forecasting based on a variational Bayesian deep learning model. Appl Energy 260:114259
35. Singh P, Dwivedi P (2018) Integration of new evolutionary approach with artificial neural network for solving short term load forecast problem. Appl Energy 217:537–549
36. Wang S (2021) An interview with Shouyang Wang: Research frontier of big data-driven economic and financial forecasting. Data Sci Manag 1(1):10–12
37. Jin M, Wen Q, Liang Y, Zhang C, Xue S, Wang X, et al (2023) Large models for time series and spatio-temporal data: A survey and outlook. arXiv:2310.10196
38. Wang Y, Yu Y, Cao S, Zhang X, Gao S (2020) A review of applications of artificial intelligent algorithms in wind farms. Artif Intell Rev 53:3447–3500
39. Hanifi S, Liu X, Lin Z, Lotfian S (2020) A critical review of wind power forecasting methods-past, present and future. Energies 13(15):3764
40. Lu P, Ye L, Zhao Y, Dai B, Pei M, Tang Y (2021) Review of meta-heuristic algorithms for wind power prediction: Methodologies, applications and challenges. Appl Energy 301:117446
41. Qian Z, Pei Y, Zareipour H, Chen N (2019) A review and discussion of decomposition-based hybrid models for wind energy forecasting applications. Appl Energy 235:939–953
42. Bokde N, Feijóo A, Villanueva D, Kulat K (2019) A review on hybrid empirical mode decomposition models for wind speed and wind power prediction. Energies 12(2):254
43. Wang Y, Zou R, Liu F, Zhang L, Liu Q (2021) A review of wind speed and wind power forecasting with deep neural networks. Appl Energy 304:117766
44. Tc Fu (2011) A review on time series data mining. Eng Appl Artif Intell 24(1):164–181
45. Ouyang T, Huang H, He Y, Tang Z (2020) Chaotic wind power time series prediction via switching data-driven modes. Renew Energy 145:270–281
46. Wang Y, Xiong W, Liu Q, Yang N, Fu P, Gong K et al (2022) Wind power prediction based on a hybrid granular chaotic time series model. Front Energy Res 9:1023
47. Bhaskar K, Singh SN (2012) AWNN-assisted wind power forecasting using feed-forward neural network. IEEE Trans Sustain Energy 3(2):306–315
48. Osório G, Matias J, 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
49. Wang H, Lei Z, Liu Y, Peng J, Liu J (2019) Echo state network based ensemble approach for wind power forecasting. Energy Convers Manage 201:112188
50. Wang C, Zhang H, Fan W, Ma P (2017) A new chaotic time series hybrid prediction method of wind power based on EEMD-SE and full-parameters continued fraction. Energy 138:977–990
51. Li Z, Luo X, Liu M, Cao X, Du S, Sun H (2022) Wind power prediction based on EEMD-Tent-SSA-LS-SVM. Energy Rep 8:3234–3243
52. Zhang C, Ding M, Wang W, Bi R, Miao L, Yu H et al (2019) An improved ELM model based on CEEMD-LZC and manifold learning for short-term wind power prediction. IEEE Access 7:121472–121481
53. Naik J, Bisoi R, 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
54. Ding J, Chen G, Yuan K (2020) Short-term wind power prediction based on improved grey wolf optimization algorithm for extreme learning machine. Processes 8(1):109
55. Wu J, Ji T, Li M, Wu P, Wu Q (2015) Multistep wind power forecast using mean trend detector and mathematical morphology-based local predictor. IEEE Trans Sustain Energy 6(4):1216–1223
56. Lu P, Ye L, Sun B, Zhang C, Zhao Y, Teng J (2018) A new hybrid prediction method of ultra-short-term wind power forecasting based on EEMD-PE and LSSVM optimized by the GSA. Energies 11(4):697
57. Zhao Y, Ye L, Li Z, Song X, Lang Y, Su J (2016) A novel bidirectional mechanism based on time series model for wind power forecasting. Appl Energy 177:793–803
58. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297
59. Zendehboudi A, Baseer MA, Saidur R (2018) Application of support vector machine models for forecasting solar and wind energy resources: A review. J Clean Prod 199:272–285
60. Wang Y, Wang D, Tang Y (2020) Clustered hybrid wind power prediction model based on ARMA, PSO-SVM, and clustering methods. IEEE Access 8:17071–17079

61. Lu P, Ye L, Tang Y, Zhao Y, Zhong W, Qu Y et al (2021) Ultra-short-term combined prediction approach based on kernel function switch mechanism. Renew Energy 164:842–866
62. Zhang Y, Wang P, Ni T, Cheng P, Lei S (2017) Wind power prediction based on LS-SVM model with error correction. Adv Electr Comput Eng 17(1):3–8
63. Lin WH, Wang P, Chao KM, Lin HC, Yang ZY, Lai YH (2021) Wind power forecasting with deep learning networks: Time-series forecasting. Appl Sci 11(21):10335
64. Yu R, Liu Z, Li X, Lu W, Ma D, Yu M et al (2019) Scene learning: deep convolutional networks for wind power prediction by embedding turbines into grid space. Appl Energy 238:249–257
65. Shao H, Deng X, Jiang Y (2018) A novel deep learning approach for short-term wind power forecasting based on infinite feature selection and recurrent neural network. J Renew Sustain Energy 10(4):043303
66. Bengio Y, Simard P, Frasconi P (1994) Learning long-term dependencies with gradient descent is difficult. IEEE Trans Neural Netw 5(2):157–166
67. Ewees AA, Al-qaness MA, Abualigah L, Abd Elaziz M (2022) HBO-LSTM: Optimized long short term memory with heap-based optimizer for wind power forecasting. Energy Convers Manage 268:116022
68. Ko MS, Lee K, Kim JK, Hong CW, Dong ZY, Hur K (2020) Deep concatenated residual network with bidirectional LSTM for one-hour-ahead wind power forecasting. IEEE Trans Sustain Energy 12(2):1321–1335
69. López E, Valle C, Allende H, Gil E, Madsen H (2018) Wind power forecasting based on echo state networks and long short-term memory. Energies 11(3):526
70. Yu R, Gao J, Yu M, Lu W, Xu T, Zhao M et al (2019) LSTM-EFG for wind power forecasting based on sequential correlation features. Futur Gener Comput Syst 93:33–42
71. Devi AS, Maragatham G, Boopathi K, Rangaraj A (2020) Hourly day-ahead wind power forecasting with the EEMD-CSO-LSTM-EFG deep learning technique. Soft Comput 24(16):12391–12411
72. Neshat M, Nezhad MM, Abbasnejad E, Mirjalili S, Groppi D, Heydari A et al (2021) Wind turbine power output prediction using a new hybrid neuro-evolutionary method. Energy 229:120617
73. Liu Y, Guan L, Hou C, Han H, Liu Z, Sun Y et al (2019) Wind power short-term prediction based on LSTM and discrete wavelet transform. Appl Sci 9(6):1108
74. Li S, Wang P, Goel L (2015) Wind power forecasting using neural network ensembles with feature selection. IEEE Trans Sustain Energy 6(4):1447–1456
75. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp 770–778
76. Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. IEEE Trans Signal Process 45(11):2673–2681
77. Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw 18(5–6):602–610
78. Liu W, Liu Y, Fu L, Yang M, Hu R, Kang Y (2022) Wind power forecasting method based on bidirectional long short-term memory neural network and error correction. Electr Power Compon Syst 49(13–14):1169–1180
79. Jiao R, Huang X, Ma X, Han L, Tian W (2018) A model combining stacked auto encoder and back propagation algorithm for short-term wind power forecasting. IEEE Access 6:17851–17858
80. Yin H, Ou Z, Huang S, Meng A (2019) A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition. Energy 189:116316
81. Li H (2022) Short-term wind power prediction via spatial temporal analysis and deep residual networks. Front Energy Res 10:920407
82. Duan J, Wang P, Ma W, Fang S, 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
83. Cleveland RB, Cleveland WS (1990) STL: A seasonal-trend decomposition procedure based on loess. J Off Stat 6(1):3–73
84. Wen Q, Zhou T, Zhang C, Chen W, Ma Z, Yan J, et al (2022) Transformers in time series: A survey. arXiv:2202.07125
85. Wu H, Xu J, Wang J, Long M (2021) Autoformer: decomposition transformers with auto-correlation for long-term series forecasting. Adv Neural Inf Process Syst 34:22419–22430
86. Zhou T, Ma Z, Wen Q, Wang X, Sun L, Jin R (2022) Fedformer: frequency enhanced decomposed transformer for long-term series forecasting. In: International Conference on Machine Learning, PMLR, pp 27268–27286
87. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al (2017) Attention is all you need. Adv Neural Inform Process Syst 30
88. Zheng H, Hu Z, Wang X, Ni J, Cui M (2023) VMD-CAT: a hybrid model for short-term wind power prediction. Energy Rep 9:199–211
89. Wang L, He Y, Li L, Liu X, Zhao Y (2022) A novel approach to ultra-short-term multi-step wind power predictions based on encoder-decoder architecture in natural language processing. J Clean Prod 354:131723
90. Xiao H, He X, Li C (2023) Probability density forecasting of wind power based on transformer network with expectile regression and kernel density estimation. Electronics 12(5):1187
91. Huang S, Yan C, Qu Y (2023) Deep learning model-transformer based wind power forecasting approach. Front Energy Res 10:1055683
92. Zhang Y, Yan J (2022) Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting. In: The 11th International Conference on Learning Representations
93. Chen W, Wang W, Peng B, Wen Q, Zhou T, Sun L (2022) Learning to rotate: quaternion transformer for complicated periodical time series forecasting. In: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pp 146–156
94. Zhou T, Ma Z, Wen Q, Wang X, Sun L, Jin R. Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting. In: International Conference on Machine Learning. PMLR; 2022. p. 27268–27286
95. Drouin A, Marcotte É, Chapados N (2022) Tactis: Transformer-attentional copulas for time series. In: International Conference on Machine Learning, PMLR. pp 5447–5493
96. Shen L, Wang Y (2022) TCCT: Tightly-coupled convolutional transformer on time series forecasting. Neurocomputing 480:131–145
97. Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H, et al (2021) Informer: Beyond efficient transformer for long sequence time-series forecasting. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol 35, pp 11106–11115
98. Liu S, Yu H, Liao C, Li J, Lin W, Liu AX, et al (2021) Pyraformer: Low-complexity pyramidal attention for long-range time series modeling and forecasting. In: International Conference on Learning Representations

99. Lim B, Arık SÖ, Loeff N, Pfister T (2021) Temporal fusion transformers for interpretable multi-horizon time series forecasting. Int J Forecast 37(4):1748–1764
100. Tang B, Matteson DS (2021) Probabilistic transformer for time series analysis. Adv Neural Inf Process Syst 34:23592–23608
101. Boroojeni KG, Amini MH, Bahrami S, Iyengar S, Sarwat AI, Karabasoglu O (2017) A novel multi-time-scale modeling for electric power demand forecasting: From short-term to medium-term horizon. Electr Power Syst Res 142:58–73
102. Peng X, Wang H, Lang J, Li W, Xu Q, Zhang Z et al (2021) EALSTM-QR: Interval wind-power prediction model based on numerical weather prediction and deep learning. Energy 220:119692
103. Jency W, Judith J et al (2022) Homogenized point mutual information and deep quantum reinforced wind power prediction. Int Trans Electr Energy Syst. https://doi.org/10.1155/2022/3686786
104. Ouyang T, Zha X, Qin L (2017) A combined multivariate model for wind power prediction. Energy Convers Manage 144:361–373
105. Vargas L, Paredes G, Bustos G (2010) Data mining techniques for very short term prediction of wind power. In: (2010) IREP Symposium Bulk Power System Dynamics and Control-VIII (IREP), IEEE, pp 1–7
106. Kusiak A, Zhang Z (2010) Short-horizon prediction of wind power: A data-driven approach. IEEE Trans Energy Convers 25(4):1112–1122
107. Li S, Wunsch DC, O'Hair E, Giesselmann MG (2001) Comparative analysis of regression and artificial neural network models for wind turbine power curve estimation. J Sol Energy Eng 123(4):327–332
108. Xia J, Zhao P, Dai Y (2010) Neuro-fuzzy networks for short-term wind power forecasting. In: 2010 International Conference on Power System Technology. IEEE. pp 1–5
109. Kusiak A, Zheng H, Song Z (2009) Wind farm power prediction: A data-mining approach. Wind Energy Int J Progr Appl Wind Power Conv Technol 12(3):275–293
110. Barbounis TG, Theocharis JB, Alexiadis MC, Dokopoulos PS (2006) Long-term wind speed and power forecasting using local recurrent neural network models. IEEE Trans Energy Convers 21(1):273–284
111. Devlin J, Chang MW, Lee K, Toutanova K (2019) Bert: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of NAACL-HLT, pp 4171–4186
112. Raffel C, Shazeer N, Roberts A, Lee K, Narang S, Matena M et al (2020) Exploring the limits of transfer learning with a unified text-to-text transformer. J Mach Learn Res 21(1):5485–5551
113. Cao D, Jia F, Arik SO, Pfister T, Zheng Y, Ye W, et al (2023) Tempo: Prompt-based generative pre-trained transformer for time series forecasting. In: The 12th International Conference on Learning Representations
114. Jin M, Wang S, Ma L, Chu Z, Zhang JY, Shi X, et al (2023) Time-llm: Time series forecasting by reprogramming large language models. arXiv:2310.01728
115. Zhou T, Niu P, Sun L, Jin R et al (2024) One fits all: Power general time series analysis by pretrained LM. Adv Neural Inf Process Syst 36
116. Yin H, Dong Z, Chen Y, Ge J, Lai LL, Vaccaro A et al (2017) An effective secondary decomposition approach for wind power forecasting using extreme learning machine trained by crisscross optimization. Energy Convers Manage 150:108–121
117. Chen C, Liu H (2020) Medium-term wind power forecasting based on multi-resolution multi-learner ensemble and adaptive model selection. Energy Convers Manage 206:112492
118. Yan H, Wu Z (2020) A hybrid short-term wind power prediction model combining data processing, multiple parameters optimization and multi-intelligent models apportion strategy. IEEE Access 8:227126–227140
119. Yu M, Zhang Z, Li X, Yu J, Gao J, Liu Z et al (2020) Superposition graph neural network for offshore wind power prediction. Futur Gener Comput Syst 113:145–157
120. Zhen H, Niu D, Yu M, Wang K, Liang Y, Xu X (2020) A hybrid deep learning model and comparison for wind power forecasting considering temporal-spatial feature extraction. Sustainability 12(22):9490
121. An G, Jiang Z, Chen L, Cao X, Li Z, Zhao Y et al (2021) Ultra short-term wind power forecasting based on sparrow search algorithm optimization deep extreme learning machine. Sustainability 13(18):10453
122. Naik J, Satapathy P, Dash P (2018) Short-term wind speed and wind power prediction using hybrid empirical mode decomposition and kernel ridge regression. Appl Soft Comput 70:1167–1188
123. Ribeiro MHDM, da Silva RG, Moreno SR, Mariani VC, dos Santos Coelho L (2022) Efficient bootstrap stacking ensemble learning model applied to wind power generation forecasting. Int J Electr Power Energy Syst 136:107712
124. Abdoos AA (2016) A new intelligent method based on combination of VMD and ELM for short term wind power forecasting. Neurocomputing 203:111–120
125. Xing Z, Qu B, Liu Y, Chen Z (2022) Comparative study of reformed neural network based short-term wind power forecasting models. IET Renew Power Gener 16(5):885–899
126. Jiang P, Li R, Li H (2019) Multi-objective algorithm for the design of prediction intervals for wind power forecasting model. Appl Math Model 67:101–122
127. Khorramdel B, Chung C, Safari N, Price G (2018) A fuzzy adaptive probabilistic wind power prediction framework using diffusion kernel density estimators. IEEE Trans Power Syst 33(6):7109–7121
128. Naik J, Dash PK, Dhar S (2019) A multi-objective wind speed and wind power prediction interval forecasting using variational modes decomposition based multi-kernel robust ridge regression. Renew Energy 136:701–731

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.