International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Review on Deep Learning in Wind Speed Forecasting: Techniques and Challenges

Ari Shawkat Tahir[1*], Adnan Mohsin Abdulazzeez[2], Ismail Ali Ali[3]

Submitted:14/03/2024 **Revised**: 29/04/2024 **Accepted**: 06/05/2024

Abstract: Wind speed forecasting is crucial for optimizing wind energy systems, enhancing turbine efficiency, ensuring grid stability, and planning energy production. This review critically examines advancements in wind speed forecasting through deep learning algorithms, which surpass traditional statistical methods in handling complex, non-linear, and non-stationary wind speed data. It focuses on various deep learning models, including CNNs, RNNs, LSTMs, and GRUs, and their ability to capture spatial and temporal dependencies. Essential data preprocessing techniques and evaluation metrics like RMSE, MAE, R, R2, and MAPE, are discussed to assess model performance. The review also synthesizes recent case studies demonstrating practical applications. Despite progress, challenges such as data quality, computational demands, overfitting, and model interpretability remain. Future research directions include improving data collection, developing efficient model architectures, enhancing interpretability, and mitigating overfitting. This review provides a concise overview of the current state of deep learning in wind speed forecasting, highlighting key methodologies, challenges, and future research opportunities.

Keywords: *Wind Speed Forecasting, Artificial Intelligence, Deep Learning, Wind Energy*

1. Introduction

Wind speed forecasting is a crucial element in the management and optimization of wind energy systems, playing a significant role in enhancing the efficiency of wind turbines, ensuring grid stability, and planning energy production schedules. The inherently intermittent and variable nature of wind poses considerable challenges for accurate forecasting, necessitating sophisticated modeling approaches that can capture its complex behavior[1]. Over the past three decades, the global wind energy sector has undergone substantial development, marked by significant advancements in both theoretical and applied research domains. This rapid progress has paved the way for a promising future in wind energy technology. As reported by the Global Wind Energy Council in the Global Wind Report 2023, approximately 77.6 GW of new wind power capacity was integrated into power grids in 2022, elevating the total installed wind capacity to 906 GW—an increase of 9% compared to 2021 (as shown in Fig. 1b). The report highlights that the top ten markets for new installations in 2022 include notable contributions from China, which accounted for 16% of the new capacity, followed by the United States with 11% (as shown in Fig. 1a). Additionally, the report forecasts that by 2024, the global installed capacity of onshore wind power is expected to surpass 100 GW for the first time, with offshore wind power capacity projected to reach unprecedented levels by 2025[2]. In recent years, deep learning algorithms have emerged as powerful tools capable of addressing these challenges, providing robust frameworks for analyzing and predicting wind speed patterns[3].

Deep learning (DL), a subset of artificial intelligence (AI), leverages neural networks with multiple layers to model intricate relationships within large datasets. These models have demonstrated superior performance over traditional statistical methods, particularly in handling the non-linear and non-stationary characteristics typical of wind speed data[4]. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), are among the most commonly employed deep learning architectures in this domain. CNNs are adept at extracting spatial features from data, making them useful for analyzing meteorological patterns, while RNNs excel at capturing temporal dependencies, crucial for understanding time-series data.

Several performance metrics are employed to evaluate and compare these models, such as accuracy, computing time, decomposition techniques, and statistical testing. The comparisons consider potential influences from datasets of varying sizes, locations, resolutions, weather conditions, and periods. This paper aims to identify key factors while summarizing pertinent information.

This review and synthesis of methodologies utilized in wind speed highlight the evolution of DL approaches and explore new applications within these domains while identifying

[1] *Computer Science dept.. University of Zakho, Zakho – 42002, Iraq*
ORCID ID : 0009-0006-8380-5323
[2] *Energy Department, Duhok Polytechnic University., Duhok – 42001, Iraq*
ORCID ID : 0000-0002-4357-7331
[3] *Computer Science dept.. University of Zakho, Zakho – 42002, Iraq*
ORCID ID : 0000-0002-4800-345X
** Corresponding Author Email: ari.tahir@uoz.edu.krd*

opportunities for extending current trends in research and application. Systematically reorganizing and analyzing existing research outcomes allows for identifying superior methodologies and valuable experiences, avoiding redundant efforts and errors, and innovating novel combinations of techniques for broader application. Additionally, this review can provide valuable insights and stimulate further research directions. comprehensive literature review not only offers a critical evaluation of existing datasets from prior case studies but also serves as an indispensable resource for research in the future. Such an evaluation is crucial for exploring the effectiveness and efficiency of potential applications in various contexts.

Nomenclature

SCADA	Supervisory Control and Data Acquisition	SRD	Signal-to-Noise Ratio Decomposition
DWT	Discrete Wavelet Transform	EWT	Empirical Wavelet Transform
DE	Differential Evolution	WA	Weighted Average
HELM	Hysteretic Extreme Learning Machine	NSCE	Nonlinear State Estimation
CWB	Central Weather Bureau	EFD	Empirical Fourier Decomposition
OS	Optimal Selection	NWP	Numerical Weather Prediction
NCL	Nonlinear Control Law	EWT	Empirical Wavelet Transform
RELM	Regional Extreme Learning Machine	IWOA	Improved Whale Optimization Algorithm
SNAP	Sentinel Application Platform	FWA	Firework Algorithm
GPR	Gaussian Process Regression	CEEMD	Complete Ensemble Empirical Mode Decomposition
LASSO	Least Absolute Shrinkage and Selection Operator	CSNN	Convolutional Spiking Neural Network
BPNN	Backpropagation Neural Network	SEEMD	SEEMD: Subsampled Ensemble Empirical Mode Decomposition
STSR	Spatio-Temporal State Representation	PSO	Particle Swarm Optimization
WCT	Wavelet Coherence Transform	IPSO	Improved Particle Swarm Optimization
ELM	Extreme Learning Machine	STL	Seasonal and Trend decomposition using Loess
CEEM	Complementary Ensemble Empirical Mode Decomposition	SA	Simulated Annealing
WD	Wind Direction	ASD	Amplitude Spectral Density
RTCN	Recurrent Temporal Convolutional Network	QS	query selection
SSA	Singular Spectrum Analysis	CEEMDAN	Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
ORELM	Online Sequential Extreme Learning Machine	PE	Prediction Error
SVMD	Sparse Variational Mode Decomposition	FCGRU	Fully Connected Gated Recurrent Unit
TF	Transformer	NW	Numerical Weather
		KF	Kalman Filter

The potential for applications of DL in renewable energy to advance significantly and become increasingly mature in both technique and cost management is evident, given the rapid development of supportive technologies. This review aims to explore the advancements in wind speed forecasting using deep learning algorithms. It will discuss the various DL models employed, the preprocessing techniques critical for enhancing model performance, and the metrics used to evaluate forecasting accuracy. Additionally, it will highlight real-world case studies and applications, addressing the current challenges and outlining future research directions to further improve the efficacy of DL models in wind speed forecasting. By examining the current state of research and identifying areas for improvement, this review seeks to contribute to the ongoing development and refinement of DL models in the field of wind speed forecasting.

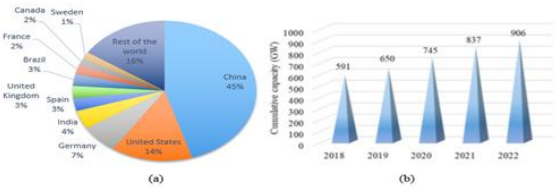

Fig. 1. (a) Top 10 markets for installations in 2023; (b)wind power global cumulative capacity, 2018–2022;

2. Deep Learning Applications of Wind speed

Wind energy has become an essential source of clean, renewable energy. The adoption of DL methods in the renewable energy sector has proven to be highly effective. These methods provide a feasible approach for modeling and predicting linear correlations and complex nonlinear dynamic processes. As the exploration and utilization of marine energy resources continue to grow[5]. The vast potential of wind energy brings numerous possibilities for advancement, including energy harvesting, control mechanisms, behavioral analysis, movement tracking, stability enhancement, and generation efficiency. These areas of focus provide ample opportunities for growth and innovation within the field of renewable energy[6]. DL methodologies, in particular, offer significant potential for optimizing these processes, thereby driving forward the development and deployment of wind energy systems. By addressing these various aspects, researchers and engineers can contribute to the broader goal of creating sustainable and efficient renewable energy systems, ultimately supporting the global transition to cleaner energy sources [7].

To gradually harness wind energy and expand its commercial scale, one of the options is to explore wind energy by analyzing its behavior and correlations. Another option is summarizing and analyzing various wind energy applications across different models. This comprehensive analysis can encourage the application of these models in optimizing wind energy, enhancing visualization techniques, and improving forecasting accuracy. By doing so, it can leverage the strengths of different models to advance wind energy technology and achieve more efficient and reliable energy production.

Wind speed applications were selected for review and discussion in this paper based on recent publications from the past few years. A variety of scientific search engines were utilized to gather relevant literature Science Direct, IEEE, Springer, Google Scholar, and ReseachGate. The search keywords included wind energy, wind speed, decomposition algorithm, Machine Learning (ML), AI, DL, ANN, CNN, RNN, LSTM, GRU, and others. Articles were chosen for review based on their publication date (starting from 2018) and their relevance to the topic.

2.1. Forecasting of Wind Energy

While wind conditions can be forecasted, accurately predicting the actual wind power output to meet demand remains challenging. This difficulty arises because power output is affected by various factors beyond just wind conditions. The application of DL structures to prediction provides an efficient method for processing large amounts of historical data for predictive analysis. Forecasting methods have been developed using various approaches, such as physical models, traditional statistical methods, AI techniques, and hybrid structures. The Classification of wind speed forecasting shown in Fig. 2.

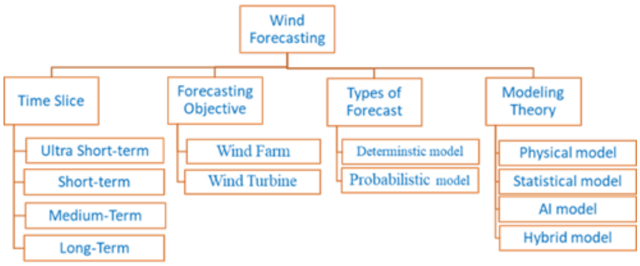

Fig.2. Classification of Wind Speed forecasting

Forecasting wind speed through physical models involves using mathematical representations of the atmosphere and its processes to predict future wind conditions. These models are built on fundamental principles of physics, such as fluid dynamics and thermodynamics, and incorporate a range of atmospheric variables. The primary objective of physical models is to simulate the behavior of the atmosphere under different conditions to forecast wind speed with a high degree of accuracy. A highly regarded physical model for wind speed forecasting is the Weather Research and Forecasting (WRF) model. This advanced numerical weather prediction system is tailored for both atmospheric research and operational forecasting. The WRF model incorporates various initial conditions, boundary conditions, and physical parameterizations to accurately simulate atmospheric phenomena across different scales. By resolving fine-scale atmospheric features, the WRF model can deliver high-resolution wind forecasts, making it particularly valuable for applications in renewable energy, including wind power generation[8]. Another example of a physical model used for wind speed forecasting is the Global Forecast System (GFS) developed by the National

Oceanic and Atmospheric Administration (NOAA). The GFS is a global numerical weather prediction model that runs four times daily, providing forecasts up to 16 days in advance. It incorporates data from various sources, including satellite observations and ground-based measurements, to produce comprehensive wind speed forecasts. The GFS model's ability to integrate a wide range of observational data helps improve the accuracy of its wind speed predictions[9]. The European Centre for Medium-Range Weather Forecasts (ECMWF) also employs a sophisticated physical model known as the Integrated Forecast System (IFS). The IFS uses a combination of deterministic and ensemble prediction systems to provide high-quality weather forecasts, including wind speed. The model integrates data assimilation techniques, which continuously incorporate observational data to update and improve the initial conditions of the forecast. This approach helps to reduce uncertainties and enhance the reliability of wind speed forecasts [10]. Cassola et al [11]. Proposed a Kalman filtering method to improve wind speed forecasts from a Numerical Weather Prediction (NWP) model in complex terrain. The Kalman filter adjusted the model's forecast using wind speed measurements from two anemometric stations. The results showed that this approach significantly enhanced the accuracy of short-term wind speed predictions, with low errors compared to actual measurements. This suggests that Kalman filtering is an effective tool for improving wind energy forecasting in challenging terrains, benefiting applications like short-term power prediction in wind farms. Ambach, et al [12]. Proposed a high-dimensional time series approach that integrates a multivariate seasonal time-varying threshold autoregressive model with interactions (TVARX) and a threshold seasonal autoregressive conditional heteroscedastic (TARCHX) model. This approach allows for the inclusion of periodicity, conditional heteroscedasticity, variable interactions, and a complex autoregressive structure with nonlinear impacts. The model's methodology lies in its use of a high-dimensional shrinkage technique and the iteratively re-weighted least absolute shrinkage and selection operator (LASSO) method. The model improves the accuracy of individual variable forecasts and accounts for their interdependencies, providing a comprehensive framework that reflects the dynamic nature of meteorological processes.

Statistical methods are crucial in wind speed forecasting by using historical data to predict future conditions. Techniques like time series analysis, regression models, and machine learning algorithms identify patterns and improve prediction accuracy. These methods enhance the reliability of forecasts, which is essential for optimizing wind turbine operations, planning energy production, and integrating wind energy into the power grid. Thus, statistical approaches are vital for the efficient use of wind energy resources. Statistical methods are most commonly used for forecasting within a 6-hour timeframe, which can significantly aid in wind turbine control and monitoring.

Cadenas et al [13]. Used wind speed and direction data collected by the Instituto de Investigaciones Eléctricas (IIE) from 2004 to 2005. Measurements at 10 meters above ground were taken with high-accuracy sensors, recording data at 1 Hz and averaging wind speeds every 10 minutes. Statistical analysis was conducted on the time series data, and forecasts were made using the single exponential smoothing (SES) method. The SES method, especially with an α value of 0.9, proved to be effective and accurate, outperforming the ANN method.

Gendeel et all[14] . Proposed a Variational Mode Decomposition (VMD) and weighted Least Squares Support Vector Machine (LS-SVM) to enhance wind power forecasting. VMD addressed non-stationarity by decomposing the wind power series, while weighted LS-SVM improved model robustness. Dividing the data into training, validation, and testing sets, and using a learning rate of 0.6, the results showed superior forecasting performance with an 80% confidence level. This combined approach effectively handled uncertainties, offering more reliable and accurate predictions for wind farm operations and power system management.

AI-based models, including back propagation, support vector machines (SVMs), fuzzy logic methods, and ANNs, have been widely used in various forecasting fields. As their use grows rapidly, many different structures of DL networks have been developed for different applications. The crucial role of wind power forecasting in both electricity grids and the energy market has placed a premium on accuracy. This has led to a significant shift towards using intelligent forecasting models. These models excel at capturing the complex relationships between different factors, making them superior to traditional statistical or physical methods.

Recent years have seen a surge in the development and application of various deep learning models, including CNN, RNN, LSTM, deep belief networks, stacked auto-encoders, and deep neural networks in general. These models with gated recurrent units and the ability to combine different architectures (deep hybrid models) have proven to outperform traditional statistical and physical models in prior research.

Ömer Ali Karaman[15] proposed a series of multiobjective predictive models utilized in wind power prediction involving developing multi-objective predictive models through advanced machine learning techniques, specifically CNN and LSTM networks, to enhance accuracy. Additional input parameters, such as air temperature, precipitation, and air density, were integrated with wind speed, wind direction, active power, and theoretical power data from the SCADA system to improve predictive capabilities. The study rigorously analyzed the optimal parameters using input-output correlation matrices to determine the influence of independent variables on the dependent variable. The study found that the LSTM model is more successful in estimating wind power. Chen et al. [16] employed a combination auto-encoder of CNN and LSTM to perform 2-D wind plane prediction. The dataset used in the case study comprised meteorological data from the Wind Integration National Dataset by NREL, collected in Indiana, US, from 2010 to 2012, within a 10 by 10 wind array. The raw data, with a resolution of 5-minute time series, was modified to a 2-hour time interval, resulting in 1,314,000 data points over the three-year period. The data was split into a 4:1 ratio for training and testing to evaluate the CNN-LSTM model.

Yao Liu et al. [17] proposed a DWT_LSTM prediction method that employs a divide-and-conquer strategy. The DWT decomposes the original wind power data into sub-signals, isolating key information. Separate LSTM networks are then used to model the temporal dynamic behaviors of each sub-signal independently. Three-time series wind power datasets from different wind farms were selected for experiments in this study. Each time series spanned 12 months to ensure comprehensive analysis. The data was recorded at 15-minute intervals, resulting in a total of 35,000 data points.

Ya-Lan et al. [18] proposed a nonlinear hybrid LSTM-DE-HELM model to achieve highly accurate and stable wind speed forecasting results. This model combines an LSTM network, HELM, DE, and an LSTM-based nonlinear mechanism. In the LSTM-DE-HELM model, three types of LSTM networks optimized by the DE algorithm and three types of HELM with varying numbers of neurons in each hidden layer are used to learn wind speed time series data. To demonstrate the effectiveness of this nonlinear hybrid model, wind speed data collected at ten-minute and one-hour intervals from a wind farm in Inner Mongolia, China, were used in two case studies.

Zhongda et al.[19] The proposed wind speed forecasting model utilizes VMD and an Improved Whale Optimization Algorithm (IWOA)-optimized Echo State Network (ESN). VMD decomposes the original wind speed data into several stationary components with different frequencies, simplifying the modeling process. Each stationary component is forecasted using an ESN optimized by IWOA. The final forecast is obtained by summing the predictions for each component. Case studies were conducted using actual ultra-short-term wind speed data with a 15-minute sampling period and short-term wind speed data with a 1-hour sampling period. Zhuoyi et al. [20]Proposed a hybrid wind speed forecasting system is developed based on the data area division (DAD) method and a deep learning neural network the model proposed for short-term wind forecasting is based only on the wind speed history data from the forecast location and the surrounding locations. The system consists of three modules: extraction module, data preprocessing module, and forecasting module. a large amount of valid historical data is extracted, filtered, and classified and used a complementary ensemble empirical mode decomposition for preprocessing while an LSTM network optimized by a genetic algorithm is used to forecast the decomposed wind speed data and integrate them into the final forecast result. Chih-Chiang et al. [21] focused on predicting real-time wind and wave changes in coastal waters during typhoon periods to prevent damage to infrastructure in international ports. The data sources included ground station data, buoy data, and hourly radar reflectivity images from CWB ground stations. The predictive models combined RNN-based GRUs and CNNs to forecast wind speeds. These wind speed predictions were then used to model wave heights. The dataset included information from 21 typhoons that affected Taiwan between 2013 and 2019. The study predicted and analyzed wind speeds and significant wave heights in the coastal waters of Keelung and Kaohsiung Ports over lead times of 1 to 6 hours.

Yuansheng et al. [22] proposed a novel model incorporating the ensemble empirical mode decomposition (EEMD) method along with a combination forecasting approach using Gaussian process regression (GPR) and the LSTM neural network based on the variance-covariance to enhance the accuracy of wind prediction. It collects wind speed data from a wind farm of Zhangjiakou, North China, two forecasting cases are represented in the case study. The data in Dataset A are recorded every 5 min and from 1 January 2014 to 4 January 2014 while the data in Dataset B are recorded every 60 min and from 1 July 2014 to 25 July 2014. The proposed model is expected to provide a useful reference for the power sector to forecast the short-term wind speed. Tian et al. [23] proposed a negative correlation learning-based regularized extreme learning machine ensemble model (NCL-RELM) integrated with optimal variational mode decomposition (OVMD) and sample entropy (SampEn) for multi-step ahead wind speed forecasting. The wind speed data was recorded at an intervals of 10 min and wind speed data for the whole year of 2018. Four data sets of 1008 continuous points (one week) are chosen for experiments. The RMSE values of the proposed OS-NCL-RELM model were 0.102, 0.125, 0.057, and 0.177 m/s for the four wind speed datasets, respectively.

2.1.1. Variations in Datasets Used for Wind Forecasting

The datasets employed in wind forecasting research are notably diverse, underscoring the complexity of this field. Typically, forecasting objectives encompass predicting wind speed and wind power, frequently utilizing DL methodologies. The sources of data for these studies are multifaceted, including historical meteorological records, remote sensing data, and geographic information. Meteorological datasets furnish essential variables such as wind speed, wind direction, temperature, atmospheric pressure, and humidity, which are critical for understanding atmospheric conditions influencing wind behavior. Remote sensing data, derived from satellites or aerial sensors, provide spatial and temporal insights into environmental conditions, capturing variables like sea surface salinity and ocean depth, particularly relevant for coastal and offshore wind forecasting. Geographic data supply information on terrain features, such as orography, which significantly impact wind patterns.

In addition to these environmental datasets, data from wind energy devices, including turbine-specific parameters like hub height and pitch angle, are integrated to enhance forecast precision. This amalgamation of diverse datasets facilitates a comprehensive analysis of the factors affecting wind energy generation. Moreover, datasets may encompass time series data, essential for capturing temporal dynamics, and spatial data, which offer context regarding location-specific characteristics that influence wind patterns.

When historical meteorological data are utilized, additional computations are often required to convert these data into accurate wind power predictions. Key parameters such as wind speed, turbine hub height, and blade pitch angle must be incorporated to model energy output effectively. This comprehensive approach to data integration and analysis is vital in enhancing the accuracy and reliability of wind forecasting models, particularly those employing advanced DL techniques. Consequently, the selection and combination of datasets are crucial, reflecting the intricate interplay of various environmental and technical factors in wind energy forecasting.

Xiaoyu et al. [24] utilized continuous wavelet transforms (CWT) to examine the spatial and temporal correlations of wind speed, centering on an SC-LSTM network. The analysis was based on data from the Buck City wind farm in Washington State, USA. The efficacy of the SC-LSTM network was evaluated against a conventional back-propagation model and an SVM through the parameters RMSE, MAE, and MAPE. The dataset, derived from the wind turbines at the wind farm, encompassed the year 2010 with a 5-minute temporal resolution, totaling 10,656 data points. The data were segmented into training, validation, and testing sets in a ratio of 30:2:5 days, respectively. Chen et al.[25] Employed a multiperiod-ahead stacked denoising autoencoder model, which operates in an unsupervised manner and utilizes unlabeled reconstructed data for wind speed forecasting. The model leverages these reconstructed data inputs to improve prediction accuracy. The training dataset consisted of 20,000 samples per wind speed series, with intervals spanning from 15 minutes to 24 hours, over a six-month period. Jianming et al. [26] proposed a novel hybrid model comprising a Quantile Regression Neural Network (QRNN) integrated with the Multi-Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (MCEEMDAN) and Grasshopper Optimization Algorithm (GOA) to achieve precise wind speed forecasting. One-hour and two-hour wind speed data collected from Yumen, Gansu Province, China, were used as case studies. The proposed models demonstrated lower RMSE and MAPE compared to similar models, indicating that these hybrid models are highly competitive for short-term wind speed forecasting.

Danxiang et al. [27] proposed a hybrid dual temporal information wind speed forecasting system combining GRUs and CNNs. The research utilized four wind speed datasets with a 10-minute resolution, collected from two wind plants: Donggang (42.5°N, 122.5°E) and Dandong (40°N, 125°E). Each dataset comprised 1,000 observations post-preprocessing and was divided into training (800 observations), validation (100 observations), and testing (100 observations) sets. The proposed system was evaluated against benchmark models, including AutoRegressive Integrated Moving Average (ARIMA), Support Vector Regression (SVR), and CNNs, to assess its ability to extract temporal information and enhance forecasting accuracy. The experimental results indicated significant improvements in MAPE compared to traditional models.

Huanling et al. [28] present a hybrid model, VMD-DE-ESN, which combines VMD, DE, and ESN for accurate wind speed forecasting. It combines VMD for data decomposition, DE for parameter optimization, and an ESN for forecasting the decomposed subseries of wind speed data. The dataset comprises wind speed data collected from the Sotavento wind farm in Galicia, Spain. The data is preprocessed and then divided into training and testing sets for model evaluation. The accuracy and stability of the VMD-DE-ESN model are validated using this dataset through a series of experiments and comparisons with other forecasting models.

Duan et al. [29] proposed an innovative hybrid forecasting system that incorporates effective data decomposition techniques, RNN prediction algorithms, and error correction methods. The system first applies a novel decomposition approach to break down the original wind speed series into subseries. It then predicts the wind speed using a recurrent neural network and subsequently decomposes the error to correct the initial predictions. The wind speed data, collected from a wind farm in the Ningxia Hui Autonomous Region of China, consists of four series with 1,000 samples each. These series were divided into training and testing sets, with the first 700 samples used for training and the remaining 300 for testing. The wind speed was measured at a height of 70 meters with a 15-minute sampling interval.

Majidi et al. [30] presented a novel wind speed forecasting model, integrating Sentinel satellite imagery analysis and ML methods. The model operates in two phases using multi-sensor satellite data. Initially, wind speed and bathymetry are analyzed using images from Sentinel-1 (S-1) and Sentinel-2 (S-2) satellites, respectively.

Ref.	***Model Used***	***Pre-Processing***	***Location***	***Time Step***	***Type***
[15]	RNN, LSTM, ANN, CNN	Normalization	Onshore	1 h	-
[16]	CNN + LSTM	ELM	Onshore	2 h	1–3 h
[17]	LSTM	DWT	Onshore	15 min	Short-Term
[18]	LSTMDE-HELM	DE	Onshore	10-min/1 h	Short term
[19]	VMD+ IWOA-ESN	IWOA	Onshore	15 min/ 1 h	ultra-short-term
[20]	GA+LSTM	CEEM	Onshore	10 min, 30 min, 60 min	short-term
[21]	CNN + GRU	-	Onshore/coast	hourly	1–6 h
[22]	EEMD -GPR-LSTM	EEMD	Onshore	5 min/ 60 min	short-term
[31]	CNN	-	Onshore	24-h	short-term
[23]	Negative correlation learning	OVMD	Offshore	10-min	Short term
[24]	SC-LSTM	WCT	Onshore	5-min	37 days
[32]	LASSO-QRNN	WD	Onshore	10-min	Short-term
[26]	QRNN	SRD	Onshore	1h-2h	Short-term
[27]	GRU+ CSNNs	EWT	Onshore/offshore	10-min/ 30-min/1 h	Short-term
[28]	ESN	VMD-DE	Onshore	10-min	Short-term
[29]	BPNN-RNN	ICEEMDAN	Onshore	15-min	Short-term
[33]	STSR-LSTM	-	Onshore/offshore	monthly/seasonal/ annual	Long-term
[30]	GRNN	WOA	Offshore	-	short-term
[34]	GRU	WA	Onshore	-	short-term
[35]	DBN	IPSO-HHT	Onshore	1-min	Ultra-short-term

Table 1. Summary of wind speed forecasting models based on deep learning

Subsequently, a hybrid forecasting model is proposed to assess and predict wind speed, utilizing a combination of a Generalized Regression Neural Network (GRNN) and Whale Optimization Algorithm (WOA). Additionally, the model evaluates offshore areas with potential for wind farm installation by considering wind speed, bathymetry, and the distance of high-scoring (HS) points from the shoreline. The method employs S-1 (SAR) and S-2 (optical) satellite images, processed using SNAP software and the Radar Wind Data (RWD) tool to extrapolate nearshore and offshore wind speed and bathymetry around the area of interest. The model was tested using S-1 images from January 1, 2015, to December 29, 2018, with a case study conducted on Favignana Island. Additionally, Zhenhao et al. [36] Proposed a Temporal Convolutional Network (TCN) model for interval prediction of wind speed. The dataset used in the study comprised wind speed data from San Francisco in 2012, the data facilitated wind speed interval prediction experiments with time horizons varying from 15 to 90 minutes. Ceyhun et al. [37] Introduced a novel two-step DL method, VMD-CNN, for wind power forecasting. Wind power data from a wind farm in southeastern Turkey was used to evaluate the performance of this method. The dataset, collected between January 1, 2018, and December 31, 2018, was sampled at hourly intervals, resulting in a total of 8,760 data points. It includes sequences of wind power, wind speed, and wind direction.

Li et al. [38] proposed a novel hybrid model integrating a genetic algorithm (GA), VMD, an improved dung beetle optimization algorithm (IDBO), and a Bidirectional Long Short-Term Memory network based on an attention mechanism (BiLSTM-A) is proposed to achieve superior forecasting performance. First, GA is used to select the optimal parameters for VMD, effectively extracting intrinsic patterns from historical wind speed data. Next, an attention mechanism is incorporated into the BiLSTM model to mitigate the loss of important information typically associated with long time series. Additionally, IDBO, enhanced by three strategies, is employed to optimize the BiLSTM parameters. Hourly wind speed data from the California Irrigation Management Information System were used, covering four seasons: spring (February 2021 to April 2021), summer (May 2021 to July 2021), autumn (August 2021 to October 2021), and winter (November 2020 to January 2021). Furthermore, Fantini et al. [34] A model based on GRU was utilized, incorporating a methodology that integrates wavelet transforms (WT) with RNNs. The study also examined potential errors resulting from improper partitioning and processing of training and validation data. Wind speed time series data from NASA POWER for a specific location in Brazil was analyzed. The results indicate that using WT as a preprocessing step for GRU input data does not produce significant improvements to warrant its application. Liu et al.[35] Introduced a Deep Belief Networks (DBN)-Elman hybrid forecasting model utilizing Improved Particle Swarm Optimization (IPSO) - Hilbert-Huang Transform (HHT), designed to manage the nonlinear and nonstationary characteristics of wind-speed data. This model integrates DBN with the Elman neural network, employing the IPSO algorithm and HHT for preprocessing the wind-speed data. The experimental dataset comprised wind-speed time series data with a one-minute temporal resolution, collected from a wind turbine in a wind farm in Jingbian, Shanxi, China. The dataset included 7200 data points, with 7140 used for training and 60 for testing the model's forecasting accuracy.

Different parameters significantly impact prediction models, and the choice of dataset can lead to varying forecasting results. The diverse weights assigned to each variable influence feature extraction and the overall model performance. In DL models, variables with smaller weights have a lesser impact on predictions. Table 1 summarizes the wind forecasting models reviewed in this paper.

2.1.2. Preprocessing

Effective pre-processing is crucial for improving the accuracy and reliability of wind speed forecasting models. Raw data from meteorological stations and wind farms often contains noise, missing values, and inconsistencies, which can negatively impact forecasting algorithms. Therefore, strong pre-processing techniques are needed to prepare the data for accurate and efficient model training and prediction.

- Data Cleaning and Noise Reduction

The initial step in pre-processing is data cleaning, which focuses on removing anomalies and filling in missing values. Techniques like interpolation and imputation are commonly used to address gaps in the dataset. For noise reduction, methods such as moving average smoothing, Gaussian filters, and advanced techniques like WT and empirical mode decomposition (EMD) can be applied. These approaches help isolate the true signal from the noise, ensuring that forecasting models are trained on high-quality data [39]–[44].

- Feature Engineering

Feature engineering is another vital aspect of pre-processing for wind speed forecasting. It involves creating new features or modifying existing ones to better capture the underlying patterns in the data. Common features derived from wind speed data include wind direction, temperature, pressure, and humidity. Additionally, temporal features such as time of day, day of the week, and seasonal indicators can be incorporated to account for periodic variations in wind speed. Feature selection techniques, such as correlation analysis and principal component analysis (PCA), help identify the most relevant features, reducing dimensionality and improving model performance [45]–[48].

- Normalization and Scaling

Normalization and scaling are critical pre-processing steps that prepare data for ML, and DL algorithms. Wind speed data often shows significant variability in magnitude, which can lead to biased model training. Techniques such as min-max scaling[49]–[51], z-score normalization [52], [53], and logarithmic transformations [44] are employed to standardize the data. This process ensures that each feature contributes equally to the model, preventing any single feature from disproportionately influencing the predictions.

- Decomposition Techniques

To enhance the predictive power of models, decomposition techniques can be applied to wind speed time series data. Methods such as EMD[54], [55], [56], CEEMDAN[57], [58], HHT[35], EEMD[59][60], WT[34], [61], [62], and VMD [14], [41], [57], [63] decompose the original time series into constituent components. These components, including trend, seasonality, and residuals, can be modeled separately, allowing for a more detailed analysis and improved forecasting accuracy.

- Handling Non-Stationarity

Wind speed data often exhibits non-stationarity, characterized by changing trends and seasonal patterns over time. Addressing non-stationarity is crucial for accurate forecasting. Techniques such as differencing, detrending, and seasonal adjustment are employed to transform the data into a stationary series. Additionally, more advanced methods, such as ARIMA[13], [29], [47] models and their extensions, can effectively manage non-stationarity in the data.

In summary, pre-processing is a vital step in the wind speed forecasting pipeline, encompassing data cleaning, noise reduction, feature engineering, normalization, decomposition, and addressing non-stationarity. These steps ensure that the data used in forecasting models is of high quality, leading to more accurate and reliable predictions. Effective pre-processing not only enhances model performance but also contributes to the overall robustness and efficacy of wind speed forecasting systems.

2.1.3. Evaluation and Comparison Methods

CNNs leverage convolutional layers to automatically extract features from input data, reducing the need for extensive manual feature engineering. By employing layers of convolutional filters, CNNs can identify and learn important patterns and correlations in wind speed data, enhancing predictive accuracy. The structure of CNNs enhances their ability to extract features from target variables. Typically, CNNs consist of multiple layers and are often referred to as "black boxes" due to their complex input and output processes. These layers include a padding layer, a convolution layer, a pooling layer, a fully connected layer (flatten layer), and dropout and activation functions (ReLU/Sigmoid). Numerous studies have integrated CNNs with other models due to their superior feature extraction capabilities, which can also help reduce computing costs to some extent. The CNN module functions as a feature extractor, translating raw data into intrinsic deep features. Through this process, spatial features are extracted using the CNN structure. Meanwhile, time series features can be captured using models such as RNN, GRU, and LSTM, which are equipped with memory gates to record sequential data [16].

The structure of CNNs can be classified into one-dimensional, two-dimensional, or three-dimensional formats based on the input dataset's feature extraction needs. In the existing literature, CNNs are often employed as feature extractors because of their exceptional capability to extract deep and detailed information, making them an ideal choice for applications requiring thorough feature analysis. Xinyu et al. [64] employed a one-dimensional CNN on an input matrix and convolutional filters, which are typically effective at identifying simple patterns and require only a few samples in each channel. The input wind speed data can be transformed into a two-dimensional format, making it adaptable for use in a multistep wind speed and turbulent standard deviation combination dataset. Zhu et al. [65] predicted wind power using a CNN model that utilized historical wind farm data as input. This study exemplifies the application of a CNN with two-dimensional matrix data. Fig. 3 explains how the CNN structure works to extract features.

Nazemi et al. [66] proposed a short-term DL-based wind speed forecasting approach utilizing a one-dimensional (1D)-CNN. This method aggregates weather information from the past hour to accurately predict wind speed for the next hour. Experimental results demonstrate that this 1D-CNN-based technique offers precise wind speed predictions, confirming its effectiveness for short-term forecasting. Abdulrahman et al. [67] proposed a 1D-CNN model for wind speed prediction at various heights above ground level (AGL). The study demonstrates that using wind speed data captured at an 18-meter height for training is sufficient for accurately predicting wind speed at higher elevations. Nazemi et al. [68] proposed an innovative two-dimensional (2D)-CNN based technique for hour-ahead wind speed prediction. This 2D-CNN model demonstrates exceptional performance in overcoming regression and prediction challenges, offering a significant advancement in wind speed forecasting. Zhu et al. [69] introduced a prediction method utilizing coupled feature analysis of wind speed behavior. This approach harnesses the advanced behavior recognition capabilities of three-dimensional (3D)-CNNs for wind speed forecasting. By employing 3D convolutions, the model extracts spatiotemporal features of wind speed behavior, enhancing prediction accuracy through a comprehensive analysis of the temporal and spatial characteristics of wind patterns within a wind farm.

Trebing and Mehrkanoon [70] present a new model utilizing CNNs for wind speed prediction. This model applies convolutions across multiple dimensions, enabling it to capture a wider range of data modalities. Unlike traditional CNN-based models, the proposed approach excels in characterizing the spatiotemporal evolution of wind data by learning the intricate input-output relationships from various dimensions of the input data. Utilizing CNNs allows researchers to improve the accuracy and reliability of wind speed forecasts, thereby facilitating more efficient wind energy management and integration into the power grid. The ongoing development and application of CNN-based models offer significant potential for advancing the field of wind speed forecasting.

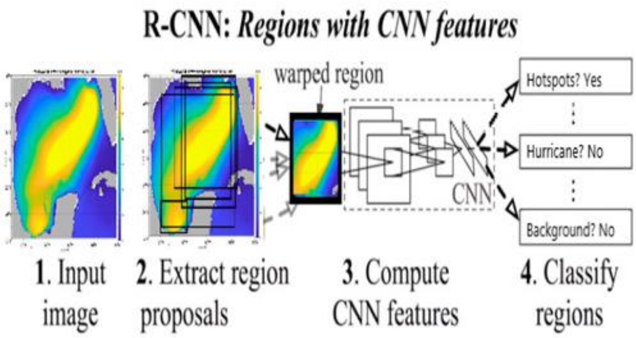

Fig. 3. Sample structure of a CNN

Recurrent neural networks (RNNs) are extensively applied in time-series-sensitive domains such as speech prediction, wind speed/power forecasting, audio recognition, and weather forecasting. These networks can feed their output data back into the model as input for subsequent steps. Training is accomplished using backpropagation, evolving from a feed-forward network with adjusted weights. Thus, RNNs are well-suited for handling temporal dynamics in sequential datasets. An RNN's current input not only takes into account the present time series data but also integrates features learned from previous steps, embodying the concept of "memory" which is distinct from traditional feed-forward networks[71]. RNNs are essential in wind speed forecasting due to their capability to model temporal sequences and capture the evolving nature of wind patterns over time. Unlike traditional feedforward neural networks, RNNs have internal memory states that allow them to retain information from previous time steps, making them particularly effective for time series analysis. This feature enables RNNs to identify patterns and dependencies in wind speed data, which is crucial for accurate forecasting[72].

The basic components of RNNs include inputs, outputs, weights, and biases as shown in Fig. 4. By forwarding previous inputs to subsequent hidden layers, RNNs can retain and "memorize" the data[72]. This recurrent mechanism allows temporal dependencies to be incorporated at various stages without losing previously learned weights. Backpropagation aids in updating these weights during training, enhancing the network's learning capabilities. Combining RNNs with CNNs can further improve the efficiency of feature extraction. Shivani et al. [73] presented a comparative study of the time series model ARIMA and the DL model RNN it aims to minimize the use of conventional power plants through unit commitment and optimize plant operations via economic dispatch. A novel framework proposed by Chuanjin et al [74] to improve the forecasting accuracy for wind speed utilized the RNN to extract the deeper features and involved in suitable machine learning methods for predicting.

However, RNNs encounter challenges such as the vanishing gradient problem, where the gradient diminishes, and impairing effective learning and memory retention from earlier steps. This limitation often results in RNNs retaining only short-term memory. The gradient represents the slope of the function, with a steeper slope facilitating faster learning. Conversely, if the slope approaches zero, learning stagnates. Addressing the vanishing gradient issue can significantly enhance RNN performance and preserve its memory capabilities. To address these challenges, LSTM networks were developed, providing improved convergence and performance in RNN architectures.

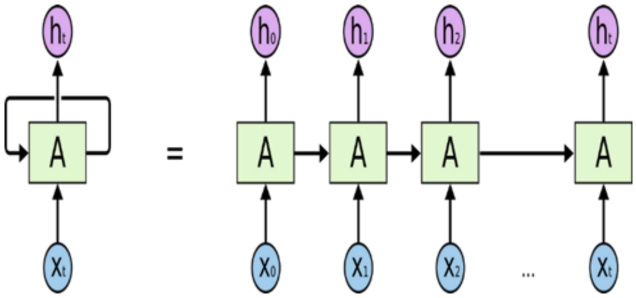

Fig.4. Simple Structure of RNN

Long Short-Term Memory (LSTM) networks significantly enhance wind speed forecasting by overcoming the limitations of traditional RNNs. LSTMs maintain information over extended periods, capturing long-term dependencies in sequential data. This capability is crucial for accurate wind speed predictions, as it enables the recognition of patterns and trends over time. Utilizing gates to control information flow, LSTMs selectively remember or forget data, improving forecast accuracy and reliability. Consequently, LSTM-based models outperform conventional methods, contributing to more efficient and reliable renewable energy management. Numerous studies have demonstrated the effectiveness of LSTM-based models in enhancing forecast accuracy and reliability compared to conventional statistical methods and basic neural networks. Numerous studies have demonstrated the effectiveness of LSTM-based models in enhancing forecast accuracy and reliability compared to conventional statistical methods and basic neural networks.

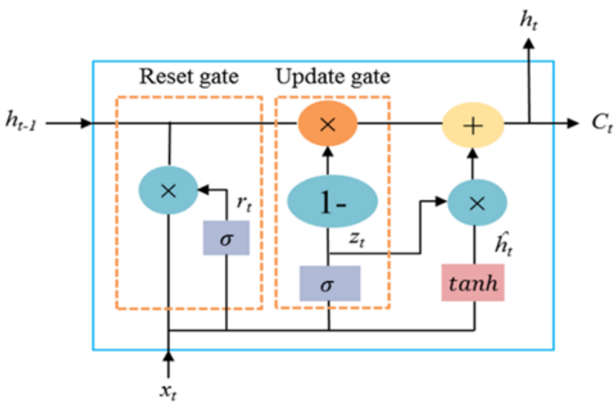

Fig. 6. The Architecture of GRU cell

Shao et al.[75] Developed a wind speed prediction model using a nonparametric LSTM neural network. They optimized the hyperparameters of the LSTM model with the firework algorithm, resulting in a reduced RMSE compared to traditional empirical parameter estimation methods. Yao et al. [76] proposed a model that combines the LSTM network with deep learning capabilities and fuzzy-rough set theory for short-term wind speed prediction. The LSTM neural network facilitates wind speed prediction by learning and processing various parameters. Aytaç et al. [53] The study also employed an LSTM model combined with a decomposition method and an optimizer for wind speed forecasting. The model's weights were estimated and optimized using Grey Wolf Optimizer (GWO), while the data were processed with a Weighted Moving Average (WMA) before being input into the model. To address missing data, Kalman filters were used for reconstruction, and interpolation was applied to prevent accuracy offsets in the system. Although the LSTM model addresses overfitting and benefits from longer memory retention, it has some limitations. Specifically, it requires substantial training time and large datasets, and it is sensitive to different weight initializations. Prabha et al.[77] Proposed an LSTM network for one-hour-ahead wind speed prediction, employing a clustering approach to segment the time series data into windy and non-windy months. This model is recommended for optimizing the scheduling of wind power to ensure stability within the power system. Fig. 5 shows the simple structure of the LSTM Cell.

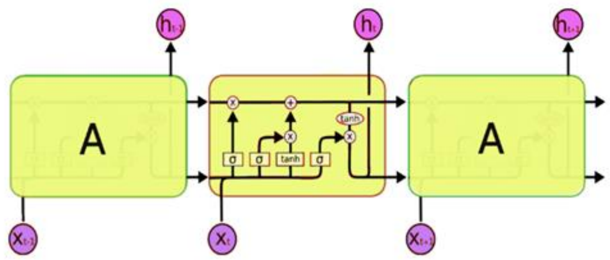

Fig. 5. Simple Structure of LSTM

A Gated Recurrent Unit (GRU) is a recurrent neural network architecture that addresses the limitations of traditional RNNs, such as the vanishing gradient problem [78]. GRUs use update and reset gates to control information flow, enabling them to capture long-term dependencies effectively. The simple structure of GRU is shown in Fig.6. In wind speed forecasting, GRUs excel at modeling temporal dependencies in sequential data. They accurately predict wind speeds by learning complex patterns from historical data. GRUs' efficiency and ability to handle long-term dependencies make them suitable for both short-term and long-term wind speed forecasting. Recent studies, like Liu et al.[79], have demonstrated GRUs' robust performance in this field, enhancing wind energy management and grid integration. Ding et al. [78] introduced a bi-dimensional GRU model designed to improve wind power prediction. This model addresses wind speed error by using it as a weighting factor, incorporating it into the time series as an input to correct wind speed forecasts. The performance of the forecasting model was evaluated using RMSE and MAE metrics. The bidirectional GRU model features two gates: one for integrating new input and another for managing previous memory. Sun et al [80]. Proposed a CNN-GRU model leveraging meteorological data from 2019 to 2021. This model combines the strengths of CNN and GRU to create a deep convolutional gated recurrent unit network. The GRU component is utilized to establish temporal relationships between historical time points, while the final wind speed predictions are derived based on spatio-temporal correlation analysis.

The hybrid model trend aims to enhance forecasting accuracy by leveraging the strengths of various model combinations. Integrating feature extraction methods with forecasting models enables effective handling of spatio-temporal data, thereby improving both efficiency and computational costs for forecasting applications. By combining different techniques, hybrid models can better capture the complexities of data patterns, leading to more reliable and precise forecasts.

Udeship et al.[81] Proposed a hybrid model by combining ARIMA and RNN-LSTM for estimating the wind speed. In [82], Tonglin Fu proposed a hybrid wind speed forecasting model known as VMD-NSCE-LSTM. This model utilizes VMD to decompose wind speed time series data into Intrinsic Mode Functions (IMFs). The NSCE method is then employed to assign weights to the forecasting results of each IMF. Finally, this pre-processed data is fed into an LSTM network for prediction. The accuracy of short-term wind speed forecasting in Huan County, Loess Plateau, China was improved by combining these techniques. Moreover, San et al. [83] proposed a wavelet decomposition-based hybrid DL model (CNN and LSTM) for one-step-ahead wind speed prediction. Historical data on wind speed, temperature, and relative humidity from Mandalay and Meiktila, Myanmar, is first filtered using wavelet techniques. These filtered features are then used in a CNN-LSTM model to forecast wind speed. The proposed hybrid model demonstrates superior performance compared to other benchmark models, highlighting its effectiveness in wind speed prediction.

Mohapatra et al. [84] proposed a new hybrid model by combining ARIMA, Kalman filter and LSTM for estimating wind speed. Kumar et al.[85] Introduced a sophisticated hybrid model incorporating Empirical Fourier Decomposition (EFD), LSTM, and GWO. The wind speed time series is decomposed via EFD into multiple sub-series and a residual component, which is stationary and thus suitable for modeling with an RNN. Each sub-series and residual is forecasted using LSTM, while GWO optimizes the predicted outputs. This methodology significantly enhances prediction accuracy and stability in wind speed forecasting. Additionally, In [18], Hu and Ya-Lan introduced a novel nonlinear hybrid model, LSTM-DE-HELM, designed to enhance wind speed prediction. This model integrates a LSTM, HELM, DE, and a nonlinear combined mechanism, providing an innovative approach to wind speed forecasting. Moreover, in [86], Nguyen introduced a novel hybrid model EEMD 4with a CNN and Bi-LSTM network, optimized through a GA. The model leverages GA for hyperparameter optimization to enhance forecasting accuracy, while EEMD aids in data decomposition to improve model performance. CNN-Bi-LSTM networks are utilized for feature extraction and capturing both historical and future wind speed data. Yin et al. [87] and Yang et al. [88] implemented Q-learning to achieve model integration in their respective studies. Meanwhile, the approaches taken by authors in [89] and [90] involved combining multiple models through techniques such as weighting and data-adaptive censoring strategies. Additionally, Xing et al. [91] and Duan et al. [59] focused on data preprocessing by decomposing the raw data into its components, predicting each component separately, and subsequently integrating these predictions through residual correction methods.

This hybrid approach outperforms single models, demonstrating superior forecasting accuracy. Table 2 provides an overview of the DL models pertinent to wind speed forecasting, highlighting their advantages, disadvantages, and potential applications.

Table 2. Summary of different deep learning models related to Wind Speed Forecasting.

Model	***Application***	***Pros.***	***Cons.***
CNN	• Pre-processing and feature extraction in wind speed forecasting. • Spatial data analysis from meteorological sensors.	• Efficient Feature Extraction: • Handling Multidimensional Data. • Robustness to Noise	• Inability to Capture Temporal Dependencies • High Computational Requirements • Complex Hyperparameter Tuning • Overfitting
RNN	• Short-term wind speed forecasting. • Sequential data analysis in meteorological time series.	• Temporal Data Processing: • Captures Short-Term Dependencies	• Limited Long-Term Performance • Complex Training:
LSTM	• Long-term wind speed forecasting. • Modeling and predicting extended temporal patterns in wind speed data.	• Long-Term Dependency Handling. • Memory Retention.	• Complex Architecture • Long Training Times
GRU	• Time series forecasting for wind speed. • Error modeling and correction in wind speed predictions.	• Simpler and Faster than LSTMs. • Handles Long-Term Dependencies: • Mitigates Vanishing Gradient Problem	• Lower Learning Efficiency • Slower Convergence:
Hybrid	• Comprehensive wind speed forecasting.	• High Accuracy. • Versatility.	• Computational Complexity

	• Integration of spatial and temporal data for enhanced predictions. • Power system optimization and energy management.	• Flexibility.	• Difficult to Tune.

2.1.4. Performance Evolution Metrics

The classical metrics used to assess the performance of wind prediction, their definitions, equations, and evaluation criteria are detailed in Table 3. MAE assesses the average absolute differences between two continuous variables, providing a straightforward measure of prediction accuracy by calculating the average magnitude of the errors in a set of predictions, without considering their direction. MAPE, on the other hand, evaluates the accuracy of the prediction by considering both the error and the ratio of the measured values to the predicted values. Meanwhile, RMSE quantifies the standard deviation between the predicted outputs and the actual observation values, offering insight into the prediction model's accuracy by measuring the average magnitude of the errors. A lower MAPE value signifies better performance, as it indicates a smaller average percentage error between the predicted and actual values, thereby reflecting the model's precision in relative terms. Additionally, the coefficient of determination (R^2) measures how well the predicted values align with the actual values, indicating the proportion of the variance in the dependent variable that is predictable from the independent variable. An R^2 value closer to 1 denotes a model with higher explanatory power and better fit to the data. Lastly, the correlation coefficient (R) evaluates the strength and direction of the linear relationship between observed and predicted values, with an R value closer to 1 indicating a strong positive correlation and better model performance.

The efficiency and performance of each model are typically assessed using a series of indicators, the results of which are summarized in Table 4.

Table 3. Performance evolution metrics.

Index	***Definition***	***Equation***
MAE	Mean absolute error	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert X_i - X_i\rvert$
MAPE	Mean absolute percentage error	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{X_i - X_i}{X_i}\right\rvert$
MAPE	Mean absolute percentage error	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{X_i - X_i}{X_i}\right\rvert \times 100\%$
RMSE	Root mean square error	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(X_i - X_i)^2}$
R	correlation coefficient	$R = \frac{\sum_{i=1}^{N}(X_i - X_i)\,(P_i - P_i)}{\sqrt{\sum_{i=1}^{N}(X_i - X_i)^2\sum_{i=1}^{N}(P_i - P_i)^2}}$
R^2	Coefficient of determination (R2)	$R^2 = 1 - \frac{\sum_{i=1}^{N}(X_i - X_i)^2}{\sum_{i=1}^{N}(X_i - X_i)^2}$

Table 4. Summary of wind speed forecasting models based on deep learning.

Ref.	*Models*	*MAE*	*MAPE*	*RMSE*	*R*	R^2	*Lead Time*
[64]	CNN-LSTM	0.3	0.4285	0.3925			
[55]	NWP-EMD-LSTM	0.1447		0.1603			50-min
[82]	VMD-NSCE-LSTM	0.0720	2.52%	0.0940			
[86]	EEMD-GA-CNN-Bi-LSTM	0.039	2.235	0.028			
[18]	LSTM-DE-HELM	1.133	1.000	1.100	0.93031		
[92]	EWT-Q-LSTM-DBN-ESN	0.716	0.960	0.9327			
[19]	VMD + IWOA-ESN	0.7452	10.2424%	0.8544		0.9287	
[75]	FWA-LSTM	0.46	30.05%	0.64			
[37]	VMD-CNN	0.0376		0.0499	0.9744		1h
[20]	DAD- *CEEMD-GA-LSMT*	0.4585	6.1188%	0.6233			
[25]	single SDAE-ELM	0.68		0.89			15-min
[69]	3D-CNN	0.0745		0.681			
[28]	VMD-DE-ESN	0.1065	2.016%	0.1384			1h
[27]	GRU-CSNN		4.80%	0.2297	99.40%		1h
[1]	SEEMD-LSTM	0.0504		0..0656		0.9353	1-min
[93]	EEMD-VMD-GRU-PSO	0.116		0.148		0.966	1-min
[81]	ARIMA-RNN-LSTM	0.097		0.124			
[35]	IPSO-HHT-DBNElman	0.3864	0.0693	0.4416		0.88	
[94]	STL-VMD-CNN-LSTM-SA	0.52566	0.13084	0.68537	0.97335		6 h
[85]	EFD-LSTM-GWO	0.465	15.0%	0.617		0.837	
[95]	ASD+RTCN	0.2257	3.77%	0.2862			1h
[96]	SSA-VMD-LSTM-ORELM	0.2385	2.3388%	0.3079			
[97]	SVMD-TF-QS	0.5598		0.7645		0.9217	1h
[98]	CEEMDAN-CNN-LSTM	0.3452	13.4191%	0.4046			
[99]	VMD-PE-FCGRU	0.199	2.45%	0.030	0.996		
[100]	NW-LSTM		3.322	1.0215			
[84]	ARIMA-KF-LSTM	1.68		2.09			
[101]	SSA-VMD-TCN-GWO	0.1062		0.1363		0.9935	10-min

3. Challenges and Future Directions

Forecasting wind speed and wind power faces major challenges because of the inherently stochastic and non-linear nature of wind patterns. The difficulty of precisely forecasting these trends is further complicated by the impact of several environmental and geographical factors. While deep learning (DL) models are capable of capturing these complex relationships, their performance can be inconsistent, particularly when dealing with diverse datasets and rapidly changing weather conditions.

The quality of training model data is crucial. Errors, noise, and missing values in meteorological and wind farm data may decrease forecasting model accuracy. Despite modern preprocessing algorithms like WT, EMD, and VMD, data cleaning and preparation are challenging to ensure reliable predictions. The interpretability of deep learning models, especially hybrid models that combine different algorithms such as LSTM, RNN, CNN, and ANN, decreases as their complexity increases. Additionally, these complex models often require extensive computational resources, which can be a barrier to their widespread adoption. Overfitting is a common issue in deep learning models, characterized by a model's ability to perform well on training data but unable to generalize to unknown data. This issue is especially significant in predicting wind speed, since models may excessively adapt to particular situations or datasets. Ensuring the generality of models across various wind farms and geographic locations remains a challenging task. Due to the decreasing correlation among data point in wind speed data points over extended periods, making it harder for models to capture long-term dependencies and trends accurately, so achieving high accuracy in long-term wind speed predictions remains difficult, unlike short-term forecasting.

Future research might prioritize the seamless integration of heterogeneous information, such as a combination of meteorological data with remote sensing and turbine-specific metrics. This integration would enable the development of more holistic models that consider a wider array of elements which affect wind speed and the generation of power. Developing hybrid models that combine deep learning and statistical approaches. These models can leverage the strengths of different approaches. These models can use LSTM networks' temporal pattern recognition and CNNs' spatial analysis to increase prediction accuracy. In the future, one of the most important areas of study will be the development of algorithms that will make deep learning models more interpretable. Among these options are the development of tools or frameworks that offer insights into the process by which models create predictions, as well as the creation of models that manage a balance between complexity and transparency. Regularization, cross-validation, and the use of larger, more diverse datasets can reduce deep learning models overfitting. Additionally, exploring ensemble methods that combine multiple models could improve generalization and reduce the likelihood of overfitting. Future research should prioritize improving long-term wind speed forecasting by developing models that better capture long-term dependencies. As the use of deep learning models in wind speed forecasting becomes more widespread, there will be a rising need to guarantee that these models are scalable and can be implemented in real-time environments. To make these models more widely applicable in the energy sector, it will be necessary to do research on how to reduce the amount of processing that these models require without compromising their accuracy.

4. Conclusion

In this study, a comprehensive summary and organization of research on DL-based wind speed prediction that was carried out between the years 2018 and 2024 is presented. It focuse on important steps like the preparation of data, the extraction of features, the learning of relationships, and the optimization of parameters. Provided a comprehensive analysis as well as a discussion of the most important technologies and models that are involved in the essential process.

Data preprocessing is essential for handling anomalies and identifying patterns in wind speed data. This process includes detecting outliers and using decomposition methods to manage abnormal data points and reveal trends. Identifying the main features and eliminating unnecessary features can be done through Feature extraction, whether by traditional statistical techniques or neural networks. Relationship learning which connecting input features to wind speed or power output, using models like nonlinear regression, tree-based methods, deep learning (DL), or hybrid approaches that combine multiple methods. By fine-tuning hyper-parameters, parameter optimization, which is often accomplished via intelligent algorithms, enhances the model's performance. In large-scale or real-time applications, computational cost and efficiency must be balanced. Metrics such as MAE, MAPE, RMSE, R, and R^2 are crucial for assessing model accuracy and reliability.

This review paper focuses on the recent challenges in DL-based wind prediction and suggests future directions for increasing accuracy to improve the performance of DL models in predicting wind speed. The review helps wind speed forecasting experts enhance DL technology by studying and analyzing these challenges and trends.

References

[1] M. Bahri, S. Vahidnia, L. Ghignone, and M. Mohtashamkhani, “Wind Speed Prediction Using a SEEMD-LSTM Model,” 9th Int. Conf. Eng. Emerg. Technol. ICEET 2023, no. October, pp. 1–6, 2023, doi:

10.1109/ICEET60227.2023.10525919.

[2] G. W. Energy, “Global Wind Energy Council,” Bonn,Germany, 2023.

[3] P. P. Prabha, V. Vanitha, and R. Resmi, “Wind Speed Forecasting using Long Short Term Memory Networks,” 2019 2nd Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2019, no. August, pp. 1310–1314, 2019, doi: 10.1109/ICICICT46008.2019.8993279.

[4] Z. Zhang, L. Lin, S. Gao, J. Wang, and H. Zhao, “Wind speed prediction in China with fully-convolutional deep neural network,” Renew. Sustain. Energy Rev., vol. 201, no. June, p. 114623, 2024, doi: 10.1016/j.rser.2024.114623.

[5] U. Yuzgec, E. Dokur, and M. Balci, “A novel hybrid model based on Empirical Mode Decomposition and Echo State Network for wind power forecasting,” Energy, vol. 300, no. May, p. 131546, 2024, doi: 10.1016/j.energy.2024.131546.

[6] M.-R. Chen, G.-Q. Zeng, K.-D. Lu, and J. Weng, “A Two-Layer Nonlinear Combination Method for Short-Term Wind Speed Prediction Based on ELM, ENN, and LSTM,” IEEE Internet Things J., vol. 6, no. 4, pp. 6997–7010, 2019, doi: 10.1109/JIOT.2019.2913176.

[7] Y. Y. Hong, C. L. P. P. Rioflorido, and W. Zhang, “Hybrid deep learning and quantum-inspired neural network for day-ahead spatiotemporal wind speed forecasting,” Expert Syst. Appl., vol. 241, no. November 2023, p. 122645, 2024, doi: 10.1016/j.eswa.2023.122645.

[8] J. D. Adams, “Reinterpreting evaluation classics in the modern age,” J. Contin. High. Educ., vol. 49, no. 2, pp. 14–22, 2001, doi: 10.1080/07377366.2001.10400427.

[9] N. C. for E. P. (U.S.), Ed., “The GFS atmospheric model.” 2004, [Online]. Available: https://repository.library.noaa.gov/view/noaa/11406.

[10] N. Procedures, “IFS DOCUMENTATION – Cy47r1 Operational implementation 30 June 2020 PART III : DYNAMICS AND NUMERICAL PROCEDURES,” no. June, pp. 1–31, 2020.

[11] F. Cassola and M. Burlando, “Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output,” Appl. Energy, vol. 99, pp. 154–166, 2012, doi: https://doi.org/10.1016/j.apenergy.2012.03.054.

[12] D. Ambach and W. Schmid, “A new high-dimensional time series approach for wind speed, wind direction and air pressure forecasting,” Energy, vol. 135, pp. 833–850, 2017, doi: https://doi.org/10.1016/j.energy.2017.06.137.

[13] E. Cadenas, O. A. Jaramillo, and W. Rivera, “Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method,” Renew. Energy, vol. 35, no. 5, pp. 925–930, 2010, doi: 10.1016/j.renene.2009.10.037.

[14] M. Gendeel, Y. Zhang, X. Qian, and Z. Xing, “Deterministic and probabilistic interval prediction for wind farm based on VMD and weighted LS-SVM,” Energy Sources, Part A Recover. Util. Environ. Eff., vol. 43, no. 7, pp. 800–814, 2021, doi: 10.1080/15567036.2019.1632980.

[15] Ö. A. Karaman, “Prediction of Wind Power with Machine Learning Models,” Appl. Sci., vol. 13, no. 20, p. 11455, 2023, doi: 10.3390/app132011455.

[16] Y. Chen et al., “2-D regional short-term wind speed forecast based on CNN-LSTM deep learning model,” Energy Convers. Manag., vol. 244, no. March, p. 114451, 2021, doi: 10.1016/j.enconman.2021.114451.

[17] Y. Liu et al., “Wind power short-term prediction based on LSTM and discrete wavelet transform,” Appl. Sci., vol. 9, no. 6, 2019, doi: 10.3390/app9061108.

[18] Y. L. Hu and L. Chen, “A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm,” Energy Convers. Manag., vol. 173, no. July, pp. 123–142, 2018, doi: 10.1016/j.enconman.2018.07.070.

[19] Z. Tian, H. Li, and F. Li, “A combination forecasting model of wind speed based on decomposition,” Energy Reports, vol. 7, pp. 1217–1233, 2021, doi: 10.1016/j.egyr.2021.02.002.

[20] Z. Liu, R. Hara, and H. Kita, “Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting,” Energy Convers. Manag., vol. 238, no. January, p. 114136, 2021, doi: 10.1016/j.enconman.2021.114136.

[21] C. C. Wei and H. C. Chang, “Forecasting of typhoon-induced wind-wave by using convolutional deep learning on fused data of remote sensing and ground measurements,” Sensors, vol. 21, no. 15, 2021, doi: 10.3390/s21155234.

[22] Y. Huang, S. Liu, and L. Yang, “Wind speed forecasting method using EEMD and the combination forecasting method based on GPR and LSTM,” Sustain., vol. 10, no. 10, 2018, doi: 10.3390/su10103693.

[23] T. Peng, C. Zhang, J. Zhou, and M. S. Nazir, “Negative correlation learning-based RELM ensemble model integrated with OVMD for multi-step ahead wind speed forecasting,” Renew. Energy, vol. 156, pp. 804–819, 2020, doi: 10.1016/j.renene.2020.03.168.

[24] X. Shi et al., “Deep-learning-based Wind Speed Forecasting Considering Spatial–temporal Correlations with Adjacent Wind Turbines,” J. Coast. Res., vol. 93, no. sp1, pp. 623–632, 2019, doi: 10.2112/SI93-084.1.

[25] d Y. Z. C Lei, Zhijun Li, “Multiperiod-Ahead Wind Speed Forecasting Using Deep Neural.pdf.” .

[26] J. Hu, J. Heng, J. Wen, and W. Zhao, “Deterministic and probabilistic wind speed forecasting with de-noising-reconstruction strategy and quantile regression based algorithm,” Renew. Energy, vol. 162, pp. 1208–1226, 2020, doi: 10.1016/j.renene.2020.08.077.

[27] [27]D. Wei, J. Wang, X. Niu, and Z. Li, “Wind speed forecasting system based on gated recurrent units and convolutional spiking neural networks,” Appl. Energy, vol. 292, no. December 2020, p. 116842, 2021, doi: 10.1016/j.apenergy.2021.116842.

[28] H. Hu, L. Wang, and R. Tao, “Wind speed forecasting based on variational mode decomposition and improved echo state network,” Renew. Energy, vol. 164, pp. 729–751, 2021, doi: 10.1016/j.renene.2020.09.109.

[29] J. Duan, H. Zuo, Y. Bai, J. Duan, M. Chang, and B. Chen, “Short-term wind speed forecasting using recurrent neural networks with error correction,” Energy, vol. 217, p. 119397, 2021, doi: 10.1016/j.energy.2020.119397.

[30] M. Majidi Nezhad, A. Heydari, E. Pirshayan, D. Groppi, and D. Astiaso Garcia, “A novel forecasting model for wind speed assessment using sentinel family satellites images and machine learning method,” Renew. Energy, vol. 179, pp. 2198–2211, 2021, doi: 10.1016/j.renene.2021.08.013.

[31] A. J. H. Goulart and R. de Camargo, “On data selection for training wind forecasting neural networks,” Comput. Geosci., vol. 155, no. December 2020, 2021, doi: 10.1016/j.cageo.2021.104825.

[32] L. Yaodong, X. Zidong, W. Hao, W. Yawei, M. Jianxiao, and Z. Yiming, “Probabilistic short-term wind speed forecasting using a novel ensemble QRNN,” Structures, vol. 57, p. 105286, 2023, doi: https://doi.org/10.1016/j.istruc.2023.105286.

[33] T. Ahmad and D. Zhang, “A data-driven deep sequence-to-sequence long-short memory method along with a gated recurrent neural network for wind power forecasting,” Energy, vol. 239, p. 122109, 2022, doi: https://doi.org/10.1016/j.energy.2021.122109.

[34] D. G. Fantini, R. N. Silva, M. B. B. Siqueira, M. S. S. Pinto, M. Guimarães, and A. C. P. Brasil, “Wind speed short-term prediction using recurrent neural network GRU model and stationary wavelet transform GRU hybrid model,” Energy Convers. Manag., vol. 308, p. 118333, 2024, doi: https://doi.org/10.1016/j.enconman.2024.118333.

[35] W. Liu et al., “Wind-speed forecasting model based on DBN-Elman combined with improved PSO-HHT,” Glob. Energy Interconnect., vol. 6, no. 5, pp. 530–541, 2023, doi: 10.1016/j.gloei.2023.10.002.

[36] Z. Gan, C. Li, J. Zhou, and G. Tang, “Temporal convolutional networks interval prediction model for wind speed forecasting,” Electr. Power Syst. Res., vol. 191, no. October 2020, p. 106865, 2021, doi: 10.1016/j.epsr.2020.106865.

[37] C. Yildiz, H. Acikgoz, D. Korkmaz, and U. Budak, “An improved residual-based convolutional neural network for very short-term wind power forecasting,” Energy Convers. Manag., vol. 228, no. December 2020, p. 113731, 2021, doi: 10.1016/j.enconman.2020.113731.

[38] Y. Li, K. Sun, Q. Yao, and L. Wang, “A dual-optimization wind speed forecasting model based on deep learning and improved dung beetle optimization algorithm,” Energy, vol. 286, p. 129604, 2024, doi: https://doi.org/10.1016/j.energy.2023.129604.

[39] X. Xia, X. Wang, and L. Guo, “A Novel Hybrid Model for Short-Term Wind Speed Forecasting Based on Twice Decomposition, PSR, and IMVO-ELM,” Complex., vol. 2022, Jan. 2022, doi: 10.1155/2022/4014048.

[40] W. Xinxin, S. Xiaopan, A. Xueyi, and L. Shijia, “Short-term wind speed forecasting based on a hybrid model of ICEEMDAN, MFE, LSTM and informer.,” PLoS One, vol. 18, no. 9, p. e0289161, 2023, doi: 10.1371/journal.pone.0289161.

[41] Z. Zhao et al., “Hybrid VMD-CNN-GRU-based model for short-term forecasting of wind power considering spatio-temporal features,” Eng. Appl. Artif. Intell., vol. 121, p. 105982, 2023, doi: https://doi.org/10.1016/j.engappai.2023.105982.

[42] N. Sun, J. Zhou, L. Chen, B. Jia, M. Tayyab, and T. Peng, “An adaptive dynamic short-term wind speed forecasting model using secondary decomposition and an improved regularized extreme learning machine,” Energy, vol. 165, pp. 939–957, 2018, doi: https://doi.org/10.1016/j.energy.2018.09.180.

[43] A. Meng, J. Ge, H. Yin, and S. Chen, “Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm,” Energy Convers. Manag., vol. 114, pp. 75–88, 2016, doi:

https://doi.org/10.1016/j.enconman.2016.02.013.

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

[45] H. Wei, W. Wang, and X. Kao, "A novel approach to ultra-short-term wind power prediction based on feature engineering and informer," Energy Reports, vol. 9, pp. 1236–1250, 2023, doi: https://doi.org/10.1016/j.egyr.2022.12.062.

[46] X. Liu, H. Zhang, X. Kong, and K. Y. Lee, "Wind speed forecasting using deep neural network with feature selection," Neurocomputing, vol. 397, pp. 393–403, 2020, doi: https://doi.org/10.1016/j.neucom.2019.08.108.

[47] M. U. Yousuf, I. Al-Bahadly, and E. Avci, "Short-Term Wind Speed Forecasting Based on Hybrid MODWT-ARIMA-Markov Model," IEEE Access, vol. 9, pp. 79695–79711, 2021, doi: 10.1109/ACCESS.2021.3084536.

[48] W. Zhang, J. Wang, J. Wang, Z. Zhao, and M. Tian, "Short-term wind speed forecasting based on a hybrid model," Appl. Soft Comput., vol. 13, no. 7, pp. 3225–3233, 2013, doi: https://doi.org/10.1016/j.asoc.2013.02.016.

[49] C. Patel, A. Pandey, R. Wadhvani, and D. Patil, "Forecasting Nonstationary Wind Data Using Adaptive Min-Max Normalization," in 2022 1st International Conference on Sustainable Technology for Power and Energy Systems (STPES), 2022, pp. 1–6, doi: 10.1109/STPES54845.2022.10006473.

[50] V. B. Kumar, V. M. Nookesh, B. S. Saketh, S. Syama, and J. Ramprabhakar, "Wind Speed Prediction Using Deep Learning-LSTM and GRU," in 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), 2021, pp. 602–607, doi: 10.1109/ICOSEC51865.2021.9591886.

[51] S. S. Subbiah, S. K. Paramasivan, K. Arockiasamy, S. Senthivel, and M. Thangavel, "Deep Learning for Wind Speed Forecasting Using Bi-LSTM with Selected Features," Intell. Autom. Soft Comput., vol. 35, no. 3, pp. 3829–3844, 2023, doi: 10.32604/iasc.2023.030480.

[52] K. U. Jaseena and B. C. Kovoor, "Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks," Energy Convers. Manag., vol. 234, p. 113944, 2021, doi: https://doi.org/10.1016/j.enconman.2021.113944.

[53] A. Altan, S. Karasu, and E. Zio, "A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer," Appl. Soft Comput., vol. 100, p. 106996, 2021, doi: https://doi.org/10.1016/j.asoc.2020.106996.

[54] J. Hu, J. Heng, J. Wen, and W. Zhao, "Deterministic and probabilistic wind speed forecasting with de-noising-reconstruction strategy and quantile regression based algorithm," Renew. Energy, vol. 162, pp. 1208–1226, 2020, doi: https://doi.org/10.1016/j.renene.2020.08.077.

[55] F. Zhang, Z. Guo, X. Sun, and J. Xi, "Short-term wind power prediction based on EMD-LSTM combined model," IOP Conf. Ser. Earth Environ. Sci., vol. 514, no. 4, pp. 0–10, 2020, doi: 10.1088/1755-1315/514/4/042003.

[56] H. Pan, Y. Yang, X. Li, J. Zheng, and J. Cheng, "Symplectic geometry mode decomposition and its application to rotating machinery compound fault diagnosis," Mech. Syst. Signal Process., vol. 114, pp. 189–211, 2019, doi: https://doi.org/10.1016/j.ymssp.2018.05.019.

[57] Y. Dong, C. Li, H. Shi, and P. Zhou, "Short-Term Probabilistic Wind Speed Predictions Integrating Multivariate Linear Regression and Generative Adversarial Network Methods," Atmosphere (Basel)., vol. 15, no. 3, 2024, doi: 10.3390/atmos15030294.

[58] T. Peng, J. Zhou, C. Zhang, and Y. Zheng, "Multi-step ahead wind speed forecasting using a hybrid model based on two-stage decomposition technique and AdaBoost-extreme learning machine," Energy Convers. Manag., vol. 153, pp. 589–602, 2017, doi: https://doi.org/10.1016/j.enconman.2017.10.021.

[59] X. Ai, S. Li, and H. Xu, "Wind speed prediction model using ensemble empirical mode decomposition, least squares support vector machine and long short-term memory," Front. Energy Res., vol. 10, 2023, doi: 10.3389/fenrg.2022.1043867.

[60] Y. Yan, X. Wang, F. Ren, Z. Shao, and C. Tian, "Wind speed prediction using a hybrid model of EEMD and LSTM considering seasonal features," Energy Reports, vol. 8, pp. 8965–8980, 2022, doi: https://doi.org/10.1016/j.egyr.2022.07.007.

[61] P. Mandal, H. Zareipour, and W. D. Rosehart, "Forecasting aggregated wind power production of multiple wind farms using hybrid wavelet-PSO-NNs," Int. J. Energy Res., vol. 38, no. 13, pp. 1654–1666, 2014, doi: https://doi.org/10.1002/er.3171.

[62] D. Liu, D. Niu, H. Wang, and L. Fan, "Short-term wind speed forecasting using wavelet transform and support

vector machines optimized by genetic algorithm," Renew. Energy, vol. 62, pp. 592–597, 2014, doi: https://doi.org/10.1016/j.renene.2013.08.011.

[63] Y. Zhang, Y. Zhao, C. Kong, and B. Chen, "A new prediction method based on VMD-PRBF-ARMA-E model considering wind speed characteristic," Energy Convers. Manag., vol. 203, p. 112254, 2020, doi: https://doi.org/10.1016/j.enconman.2019.112254.

[64] X. Zhao, N. Jiang, J. Liu, D. Yu, and J. Chang, "Short-term average wind speed and turbulent standard deviation forecasts based on one-dimensional convolutional neural network and the integrate method for probabilistic framework," Energy Convers. Manag., vol. 203, p. 112239, 2020, doi: https://doi.org/10.1016/j.enconman.2019.112239.

[65] A. Zhu, X. Li, Z. Mo, and R. Wu, "Wind power prediction based on a convolutional neural network," 2017 Int. Conf. Circuits, Devices Syst., pp. 131–135, 2017, [Online]. Available: https://api.semanticscholar.org/CorpusID:715312.

[66] M. Nazemi, S. Chowdhury, A. H. Khan, and X. Liang, "A Hour-Ahead Wind Speed Forecasting Using One-Dimensional Convolutional Neural Network," in 2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2023, pp. 25–30, doi: 10.1109/CCECE58730.2023.10288848.

[67] A. Katranji, M. Shafiullah, and S. Rehman, "Short-Term Wind Speed Prediction for Saudi Arabia via 1D-CNN," in 2023 IEEE 13th International Conference on System Engineering and Technology (ICSET), 2023, pp. 153–158, doi: 10.1109/ICSET59111.2023.10295102.

[68] M. Nazemi, S. Chowdhury, and X. Liang, "A Novel Two-Dimensional Convolutional Neural Network-Based an Hour-Ahead Wind Speed Prediction Method," IEEE Access, vol. 11, pp. 118878–118889, 2023, doi: 10.1109/ACCESS.2023.3327135.

[69] X. Zhu, R. Liu, Y. Chen, X. Gao, Y. Wang, and Z. Xu, "Wind speed behaviors feather analysis and its utilization on wind speed prediction using 3D-CNN," Energy, vol. 236, p. 121523, 2021, doi: 10.1016/j.energy.2021.121523.

[70] K. Trebing and S. Mehrkanoon, "Wind speed prediction using multidimensional convolutional neural networks," 2020 IEEE Symp. Ser. Comput. Intell. SSCI 2020, pp. 713–720, 2020, doi: 10.1109/SSCI47803.2020.9308323.

[71] C. Gu and H. Li, "Review on Deep Learning Research and Applications in Wind and Wave Energy," Energies, vol. 15, no. 4, 2022, doi: 10.3390/en15041510.

[72] Q. Cao, B. T. Ewing, and M. A. Thompson, "Forecasting wind speed with recurrent neural networks," Eur. J. Oper. Res., vol. 221, no. 1, pp. 148–154, 2012, doi: 10.1016/j.ejor.2012.02.042.

[73] Shivani, K. S. Sandhu, and A. Ramachandran Nair, "A Comparative Study of ARIMA and RNN for Short Term Wind Speed Forecasting," in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019, pp. 1–7, doi: 10.1109/ICCCNT45670.2019.8944466.

[74] C. Yu, Y. Li, Y. Bao, H. Tang, and G. Zhai, "A novel framework for wind speed prediction based on recurrent neural networks and support vector machine," Energy Convers. Manag., vol. 178, no. September, pp. 137–145, 2018, doi: 10.1016/j.enconman.2018.10.008.

[75] B. Shao, D. Song, G. Bian, and Y. Zhao, "Wind Speed Forecast Based on the LSTM Neural Network Optimized by the Firework Algorithm," Adv. Mater. Sci. Eng., vol. 2021, 2021, doi: 10.1155/2021/4874757.

[76] W. Yao, P. Huang, and Z. Jia, "Multidimensional LSTM Networks to Predict Wind Speed," in 2018 37th Chinese Control Conference (CCC), 2018, pp. 7493–7497, doi: 10.23919/ChiCC.2018.8484017.

[77] P. P. Prabha, V. Vanitha, and R. Resmi, "Wind Speed Forecasting using Long Short Term Memory Networks," 2019 2nd Int. Conf. Intell. Comput. Instrum. Control Technol. ICICICT 2019, pp. 1310–1314, 2019, doi: 10.1109/ICICICT46008.2019.8993279.

[78] M. Ding, H. Zhou, H. Xie, M. Wu, Y. Nakanishi, and R. Yokoyama, "A gated recurrent unit neural networks based wind speed error correction model for short-term wind power forecasting," Neurocomputing, vol. 365, pp. 54–61, 2019, doi: 10.1016/j.neucom.2019.07.058.

[79] S. Lv, L. Wang, and S. Wang, "A Hybrid Neural Network Model for Short-Term Wind Speed Forecasting," Energies, vol. 16, no. 4, 2023, doi: 10.3390/en16041841.

[80] Y. Sun, J. Du, and S. Wang, "Spatio-Temporal Wind Speed Prediction Based on CNN-GRU," in 2023 IEEE/ACIS 23rd International Conference on Computer and Information Science (ICIS), 2023, pp. 86–90, doi: 10.1109/ICIS57766.2023.10210249.

[81] D. M. Udeshi, S. G. L. Divakarla, N. C. Rajdev, and A. M, "Wind Speed Forecasting using Hybrid Model," in 2022 IEEE 7th International conference for Convergence in Technology (I2CT), 2022, pp. 1–5,

doi: 10.1109/I2CT54291.2022.9823995.

[82] T. Fu, "A hybrid prediction of wind speed based on variational mode decomposition method and long short-term memory," Proc. - 2020 Int. Conf. Comput. Eng. Appl. ICCEA 2020, pp. 408–412, 2020, doi: 10.1109/ICCEA50009.2020.00094.

[83] K. M. San, J. G. Singh, and K. Prakash N., "Wind Speed Forecasting using Hybrid Model of CNN and LSTM with Wavelets," in 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT), 2023, pp. 297–301, doi: 10.1109/APSIT58554.2023.10201713.

[84] M. R. Mohapatra, R. Radhakrishnan, and R. M. Shukla, "A Hybrid Approach using ARIMA, Kalman Filter and LSTM for Accurate Wind Speed Forecasting," in 2023 IEEE International Symposium on Smart Electronic Systems (iSES), 2023, pp. 425–428, doi: 10.1109/iSES58672.2023.00095.

[85] B. Kumar, N. Yadav, and Sunil, "A novel hybrid algorithm based on Empirical Fourier decomposition and deep learning for wind speed forecasting," Energy Convers. Manag., vol. 300, p. 117891, 2024, doi: https://doi.org/10.1016/j.enconman.2023.117891.

[86] T. H. T. Nguyen and Q. B. Phan, "Hourly day ahead wind speed forecasting based on a hybrid model of EEMD, CNN-Bi-LSTM embedded with GA optimization," Energy Reports, vol. 8, pp. 53–60, 2022, doi: 10.1016/j.egyr.2022.05.110.

[87] [87]Y. Zhang, X. Kong, J. Wang, S. Wang, Z. Zhao, and F. Wang, "A comprehensive wind speed prediction system based on intelligent optimized deep neural network and error analysis," Eng. Appl. Artif. Intell., vol. 128, p. 107479, 2024, doi: https://doi.org/10.1016/j.engappai.2023.107479.

[88] [88] F. Xing, X. Song, Y. Wang, and C. Qin, "A New Combined Prediction Model for Ultra-Short-Term Wind Power Based on Variational Mode Decomposition and Gradient Boosting Regression Tree," Sustainability, vol. 15, no. 14, 2023, doi: 10.3390/su151411026.

[89] [89] H. Wu, K. Meng, D. Fan, Z. Zhang, and Q. Liu, "Multistep short-term wind speed forecasting using transformer," Energy, vol. 261, p. 125231, 2022, doi: https://doi.org/10.1016/j.energy.2022.125231.

[90] [90]J. Wang, D. Gao, Z. Zhuang, and J. Wu, "An optimized complementary prediction method based on data feature extraction for wind speed forecasting," Sustain. Energy Technol. Assessments, vol. 52, p. 102068, 2022, doi: https://doi.org/10.1016/j.seta.2022.102068.

[91] [91]T. Jiang and Y. Liu, "A short-term wind power prediction approach based on ensemble empirical mode decomposition and improved long short-term memory," Comput. Electr. Eng., vol. 110, p. 108830, 2023, doi: https://doi.org/10.1016/j.compeleceng.2023.108830.

[92] H. Liu, C. Yu, H. Wu, Z. Duan, and G. Yan, "A new hybrid ensemble deep reinforcement learning model for wind speed short term forecasting," Energy, vol. 202, p. 117794, 2020, doi: 10.1016/j.energy.2020.117794.

[93] Y. Sun and J. Wang, "Wind Speed Prediction Based on Gate Recurrent Unit," 2022 IEEE Conf. Telecommun. Opt. Comput. Sci., pp. 1047–1052, 2023, doi: 10.1109/tocs56154.2022.10015928.

[94] J. Pang and S. Dong, "A novel ensemble system for short-term wind speed forecasting based on hybrid decomposition approach and artificial intelligence models optimized by self-attention mechanism," Energy Convers. Manag., vol. 307, no. March, p. 118343, 2024, doi: 10.1016/j.enconman.2024.118343.

[95] G. Zhang, Y. Zhang, H. Wang, D. Liu, R. Cheng, and D. Yang, "Short-term wind speed forecasting based on adaptive secondary decomposition and robust temporal convolutional network," Energy, vol. 288, no. October 2023, p. 129618, 2024, doi: 10.1016/j.energy.2023.129618.

[96] J. Wang, X. Niu, L. Zhang, Z. Liu, and X. Huang, "A wind speed forecasting system for the construction of a smart grid with two-stage data processing based on improved ELM and deep learning strategies," Expert Syst. Appl., vol. 241, no. March 2023, p. 122487, 2024, doi: 10.1016/j.eswa.2023.122487.

[97] S. Parri and K. Teeparthi, "SVMD-TF-QS: An efficient and novel hybrid methodology for the wind speed prediction," Expert Syst. Appl., vol. 249, no. PA, p. 123516, 2024, doi: 10.1016/j.eswa.2024.123516.

[98] Q. Li, G. Wang, X. Wu, Z. Gao, and B. Dan, "Arctic short-term wind speed forecasting based on CNN-LSTM model with CEEMDAN," Energy, vol. 299, no. March, p. 131448, 2024, doi: 10.1016/j.energy.2024.131448.

[99] Z. Gong, A. Wan, Y. Ji, K. AL-Bukhaiti, and Z. Yao, "Improving short-term offshore wind speed forecast accuracy using a VMD-PE-FCGRU hybrid model," Energy, vol. 295, no. March, p. 131016, 2024, doi: 10.1016/j.energy.2024.131016.

[100] H. H. Aly, "A Proposed Hybrid Deep Learning Model for Wind Power Forecasting," 2024 ASU Int. Conf. Emerg. Technol. Sustain. Intell. Syst. ICETSIS

2024, pp. 1011–1014, 2024, doi: 10.1109/ICETSIS61505.2024.10459633.

[101] [101] S. Fan, "Short-term wind speed prediction based on temporal convolutional networks," IEEE Jt. Int. Inf. Technol. Artif. Intell. Conf., vol. 11, pp. 165–169, 2023, doi: 10.1109/ITAIC58329.2023.10409006.