



Point and interval wind speed forecasting of multivariate time series based on dual-layer LSTM

Haipeng Zhang ^a, Jianzhou Wang ^{b,*}, Yuansheng Qian ^a, Qiwei Li ^c

^a School of Statistics, Dongbei University of Finance and Economics, Dalian, 116025, PR China

^b Faculty of Innovation Engineering, Macau University of Science and Technology, PR China

^c Dalian Ocean University, Dalian, 116025, PR China

ARTICLE INFO

Handling Editor: Jesse L. The

Keywords:

Interval forecasting
Point forecasting
The dual-layer LSTM
Wind speed

ABSTRACT

Accurate prediction of wind speed is significant importance in various applications such as renewable energy management, weather prediction, and aviation safety. However, more papers only focus on the time series of wind speed, ignoring the impact of other factors on wind speed, which will reduce the effectiveness of wind speed prediction. Therefore, an innovative methodology is suggested for both point and interval forecasts of multivariate wind speed time series using a Dual-layer Long Short-Term Memory (NVDL) in this paper. The proposed multivariate forecasting model takes into account the dependencies and correlations among different variables, which are essential for capturing the complex dynamics of wind speed variations. The first layer of the Dual-layer LSTM is responsible for capturing the temporal dependencies within each variable individually, while the second layer captures the interdependencies among the variables. By incorporating this dual-layer architecture, our model effectively captures the complex spatiotemporal patterns present for multivariate wind speed information. The results obtained from both interval and point prediction demonstrate that the proposed methodology outperforms all comparative models in the precision and stability of wind speed forecasting. Therefore, the proposed forecasting methodology, characterized by minimal prediction errors and exceptional generalization ability, can be a reliable tool for smart grid programming.

1. Introduction

Precise wind speed prediction holds significance across various domains such as renewable energy management, weather forecasting, and aviation safety. The ability to accurately predict wind speeds enables optimal utilization of wind resources, efficient planning of energy generation systems, and improved decision-making for operational and safety purposes.

Accurate prediction of wind speed holds importance in decision-making for controlling and operating wind energy systems, given its significance as a cornerstone of wind energy utilization. However, the inherent characteristics of wind speed, including low density, instability, and regional variations, present significant challenges to achieving accurate predictions (J. [1]). To enhance prediction accuracy, extensive efforts and research have been devoted to this field. These research endeavors can be broadly categorized into four classes for wind speed forecasting (C. [2]): (a) physical modeling, (b) statistical modeling, (c) hybrid prediction methods, and (d) multi-dimensional

time series prediction approaches. Table 1 lists the advantages, disadvantages, and related models of these four types of methods (The abbreviations that appear in this article are listed in Table 2).

Broadly, the physical approach involves utilizing numerical weather simulations and atmospheric data to forecast wind speed. Yang introduces a new approach, referred to as the ultra-short-term prediction of wind speed, which incorporates wind speed correction using Numerical Weather Prediction (NWP) techniques and the division of transitional weather processes [39]. Yakoub et al. [3] present a wind power forecasting system specifically designed for the Nordic energy market that cater to both short-term and medium-term predictions. These systems are developed by integrating wind speed condition inputs from three different numerical weather prediction sources.

In contrast to the physical approach, the statistical approach offers unique benefits in effectively addressing short period forecasting challenges by using the history data. Among the statistical models, Gray Forecast Model (GM), Markov chain forecasting technique [4], Autoregressive Integrated Moving Average Model (ARIMA), and other

* Corresponding author.

E-mail address: jzwang2022@163.com (J. Wang).

commonly utilized models demonstrate consistent and reliable prediction performance [5]. Yousuf et al. [6] puts forth an improved gray prediction model for wind speed forecasting that addresses the limitations of traditional methods through the incorporation of L'Hospital rule and the remnant model. H. Liu et al. [7] presents a novel short period wind speed forecasting model that adopts the EMD and the improved Recursive ARIMA model. The study conducted by H. Liu et al. [8] assesses the efficacy of various ARMA-GARCH methodologies in simulating the volatility and value of wind speed. The evaluation is carried out using a dataset spanning seven years of hourly data collected from an observation site in Colorado, USA, at four different heights.

Apart from the previously mentioned models, artificial intelligence techniques, specifically neural networks, have garnered considerable interest due to their capacity for adaptability, robustness, and fault tolerance in wind speed forecasting. Presently, neural networks are commonly integrated with other intelligent algorithms to enhance their performance in wind speed prediction. Hybrid models, which amalgamate the advantages of multiple individual models, offer superior accuracy and prediction performance compared to utilizing a single model in isolation [9]. Dong et al. [10] demonstrates denoising achieved via data decomposition, improved adaptability through optimal sub model selection, and effective optimization of reduction coefficients via a multi-objective bald eagle search algorithm to achieve better wind speed prediction results. Hu et al. [11] introduce a novel combined wind speed forecasting methods, which entails a two-phase data processing strategy. The first phase involves employing variable mode decomposition-sparse autoencoder for extracting the original wind speed features. In the second phase, batch gradient descent optimization algorithm and high-order fuzzy cognitive mapping neural network method are utilized to address the limitations of the model.

However, many existing methods solely focus on predicting a single wind speed variable without considering the influence of other factors on wind speed. This limitation poses challenges in meeting the requirements and risk management of wind power systems. As a result, researchers have increasingly shifted their attention towards developing multidimensional wind speed prediction models[12]. introduce a novel temporal attention encoder-decoder model as a solution to the problem of forecasting multivariate time series. This model is designed as an end-to-end deep learning framework that incorporates a traditional encode context vector, enabling effective integration of temporal information for accurate predictions. To address the issue of collinearity and achieve sparsity in the solution (M. [13]), propose a novel method called Adaptive Elastic Net, which use the adaptive ESN to compute uncharted weights, thereby combining the advantages of quadratic regularization and adaptively weighted lasso shrinkage methods. Wang et al. [14] present a comprehensive model for wind speed forecasting

Table 2
The abbreviations that appear in this article.

Nomenclature			
LSTM	Long Short-Term Memory	NWP	Numerical Weather Prediction
GM	Gray Forecast Model	ARIMA	Autoregressive Integrated Moving Average Model
EMD	Empirical Mode Decomposition	GARCH	Generalized Autoregressive Conditional Heteroskedastic
ESN	Elastic Net	CNN	Convolutional Neural Network
RNN	Recurrent Netural Network	VMD	Varational Mode Decomposition
PSO	Particle Swarm Optimization	NNIA	Nondominated Neighbor Immune Algorithm
VECM	Vector Error Correction Model	MAPE	The average absolute percentage error
MSE	The average squared error	MAE	The average absolute error
DC	Trend prediction accuracy	RMSE	The square root of the average of the error squares
FICP	Interval coverage	FINAW	Forecast interval normalized average width
MPI	The average predictions interval width	AWD	Accumulated width deviation of testing dataset
BPNN	Back Propagation Neural Network	ELMAN	Elman Neural Network
SVM	Support Vector Machine	FTDNN	Focused Time-delay Neural Network
NAR	Nonlinear autoregressive neural network	CEEMD	Complementary Ensemble Empirical Mode Decomposition

This table shows the full names of all the abbreviations that appear in this article.

that integrates 2 novel feature-selection models, 6 individual prediction models built upon CNN and RNN, and chameleon swarm optimization algorithm. It enables effective selection and combination of features and models, leading to enhanced accuracy in short-term wind speed predictions. Multivariate wind speed prediction offers more accurate and comprehensive forecasts by considering multiple variables, capturing intricate relationships, integrating additional influencing factors, and detecting extreme events.

Based on the analysis above, all methods have its intrinsic limitations because of its specific attributes. As a solution, this paper proposes a new approach for predicting point and interval wind speeds in multivariate time series using a Dual-layer LSTM (Long-Short-Term-Memory) model and an adaptive data processing model. Specifically, in contrast to prior personal evaluations for the amount of denoising sequence, in this study, artificial intelligence algorithms were employed to ascertain the optimal number of denoising layers, thereby limiting the need for subjectivity in the evaluation process. Then, the first layer of the Dual-layer LSTM model aims to capture the temporal dependencies within each variable

Table 1
Advantages, disadvantages, and representative literature of different wind speed prediction methods.

Categories	Advantages	Disadvantages	Models	Representative literature
Physical modeling	The ability to incorporate domain-specific knowledge and physical principles into the model, such as atmospheric dynamics, terrain features, and meteorological parameters.	They often require extensive computational resources and a vast amount of parameter tuning to achieve optimal performance.	Fluid mechanics; NWP and other models	(H. [21]) [22] [23]
Statistical modeling	Their flexibility and simplicity, which enables them to capture complex relationships between input variables and wind speed outputs efficiently.	The success of statistical modeling depends on the quality and quantity of available data, the complexity of the problem.	ARIMA; SARIMA; GM(1,1) and other models	(X. [24]) [25] [26] [27]
Hybrid prediction methods	Combine the strengths of different types models, has higher robustness and prediction accuracy	They require extensive parameter tuning to achieve optimal performance. They also require a significant amount of data to train and test the model adequately.	EMD; VMD; PSO; NNIA and other models	[28] [29] [30] (W. [20])
Multivariable time series prediction models approaches	the ability of multi-dimensional models can capture complex temporal dependencies and relationships among multiple variables simultaneously, leading to more robust predictions.	the selection and integration of dimensions in these models require careful consideration, as irrelevant or redundant variables can introduce noise and impact the accuracy of predictions.	VECM and other models	[31] (D. [32]) [33]

individually, while the second layer focuses on the interdependencies among variables. This approach leads to minimal prediction errors and improved overall performance.

Novel contributions of this paper in addressing prior research limitations:

- (1) Diverging from the conventional processing approach[15,16], this research introduces a more efficient self-adaption data preprocessing framework to manage the impact of noise-induced interference. Based on the theory of decomposition and reconfiguration, the acquired wind speed time series are disintegrated into multiple series with distinct frequencies, followed by their reconstruction using optimized combined weight determined by optimization algorithm. This preprocessing procedure effectively mitigates the randomness and volatility observed in the wind speed sequence.
- (2) This article proposes an enhanced version of the traditional LSTM model[17,18] by introducing a bilateral LSTM structure, whose novel architecture allows for the direct processing of multidimensional data. The first layer is designed to capture the temporal dependencies within each individual variable, while the second layer specifically targets the interdependencies among variables. This approach effectively reduces prediction errors and enhances the overall performance of the model.

- (3) Unlike traditional single variable wind speed prediction (Z. [19]; S. [20]), this article uses wind speed and the related three variables to simultaneously predict multi-dimensional wind speed points and intervals, improving the accuracy of wind speed prediction. Based on the wind farm data set collected from Türkiye, the experimental findings indicate that the proposed system has higher prediction accuracy than the benchmark models.

The framework of this paper is presented as follows. The chapter 2 elucidates the complete modeling framework and the relevant methods of the multivariable wind speed forecasting model. Chapter 3 introduces the dataset, evaluation indicators, and model parameters setting. Chapter 4 introduces the content and results of three experiments. The last chapter outlines principal conclusions.

2. Development of a multivariable wind speed prediction model

This section introduces an overview of the implement steps and the involved models deployed for implementing the multivariate wind speed prediction methodology proposed in this paper.

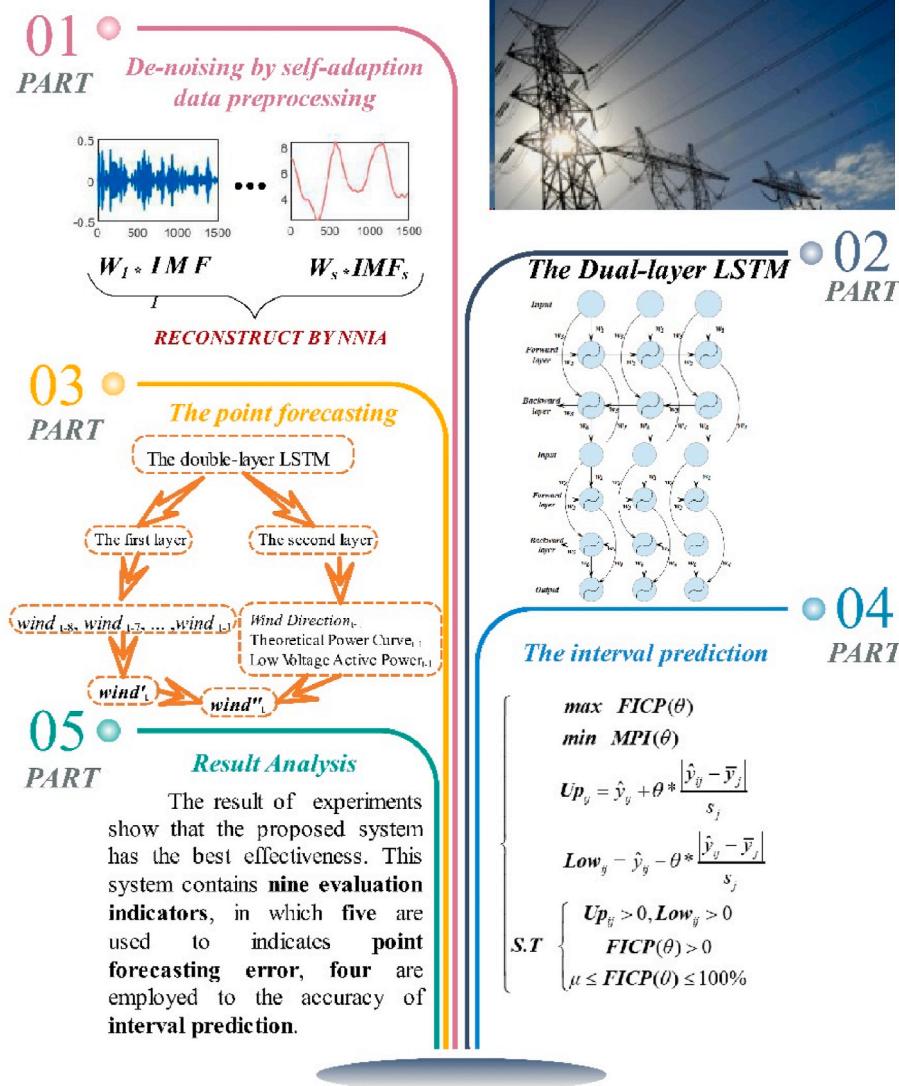


Fig. 1. The flowchart of the proposed forecasting system.

2.1. The implement steps of the proposed multivariable wind speed forecasting model

The proposed multivariate wind speed prediction model mainly includes four main steps: (1) data preprocessing based on the self-adaption data preprocessing; (2) Constructing the Dual-layer LSTM; (3) Point forecasting of wind speed time series; (4) Interval forecasting based on the result of point forecasting. The entire process of constructing the proposed multivariable wind speed prediction model is presented in this article is depicted in Fig. 1.

Step 1. data preprocessing built upon the self-adaption wind speed preprocessing

The wind speed data is initially disintegrated into multiple filters using the Variational Mode Decomposition (VMD) technique, each filter representing a different frequency component. Afterwards, the Non-Dominated Neighbor Immune Algorithm (NNIA), an intelligent optimization model, is employed to determine the corresponding weights assigned to each filter, aiming to enhance the denoising performance. By adaptively adjusting these weights, the denoising ideology effectively eliminates the negative influences caused by high-frequency noise. The Linear programming functions of the self-adaption wind speed preprocessing model can be defined:

$$\left\{ \begin{array}{l} \text{Objective : } \min(\text{MAE}) \\ \text{function : } \text{Data}_i = \omega_1 * \text{IMF}_{i1} + \omega_2 * \text{IMF}_{i2} + \dots + \omega_n * \text{IMF}_{in} \\ \text{MAE} = \frac{1}{m} * \sum_{m=1}^i |\text{y}_i - \hat{\text{y}}_i| \end{array} \right. \quad (1)$$

Where **MAE** is Mean Absolute Error, **IMF_{in}** is the n-th IMF of the i-th data. y_i is the actual data, \hat{y}_i is the forecasting value of BPNN after denoising.

Step 2. Constructing the Dual-layer LSTM

The double-layer LSTM is a neural network structure consisting of two LSTM layers, each containing multiple LSTM units. Below is a typical example of the double-layer LSTM structure:

$$\text{Input} \rightarrow [\text{LSTM}_1] \rightarrow \text{Hidden}_1 \rightarrow [\text{LSTM}_2] \rightarrow \text{Hidden}_2 \rightarrow \text{Output} \quad (2)$$

In this structure, the first LSTM layer, LSTM1, can be seen as a feature extractor that captures global features from the input sequence. The second LSTM layer, LSTM2, further extracts and transforms the outputs from the first layer, enabling integration and prediction of multi-dimensional time sequences.

To implement this structure, the multi-dimensional wind speed time series data is transformed into a suitable sequence format for the LSTM model. During the training process, a loss function such as mean squared error is utilized to optimize the model parameters, gradually fitting the training data. Afterwards, the model's hyperparameters (e.g., learning rate, batch size) are modified based on the predictions from the validation set. Finally, the trained double-layer LSTM model is used for the forecasting of multivariable wind speed data.

Step 3. Point forecasting of multivariable wind speed data

The wind speed data undergoes preprocessing techniques to ensure data integrity and eliminate outliers. The length of the dataset is 2000, and top 1600 data points in 2000 data are utilized for training the structure of the model, while the last 400 data points serve as the testing set for testing the accuracy of the model. Furthermore, through multiple experiments, the optimal ratio between the input and output is determined to be 8:1. The double-layer LSTM architecture is employed to obtain the final point forecasting results. The output of the first layer represents the preliminary predicted value \hat{y} of the wind speed. In the second layer, the input includes the preliminary predicted value \hat{y} from the first layer, as well as additional features at the 8th time point such as Wind Direction, Theoretical Power Curve (KWh), and Low Voltage

Active Power.

Step 4. Interval forecasting built upon the result of point forecasting

The multivariate wind speed prediction results are utilized as the basis for employing fuzzy c-means clustering, which classifies the predicted values into distinct groups based on their numerical values. Consequently, each group exhibits unique statistical characteristics, including varying mean and variance. To establish the desired interval width for each group, a parameter " θ " is introduced. Subsequently, an optimization algorithm is employed to adjust parameter " θ ", aiming to expand the coverage of data points within the prediction interval while simultaneously reducing its width. Importantly, these aforementioned steps rely solely on the predicted values, enabling the direct application of optimized fuzzy grouping and parameters to data beyond the sample. The specific objective function is as follows:

$$\left\{ \begin{array}{l} \max \text{ FICP}(\theta) \\ \min \text{ MPI}(\theta) \\ \text{Up}_{ij} = \hat{y}_{ij} + \theta * \frac{|\hat{y}_{ij} - \bar{y}_j|}{s_j} \\ \text{Low}_{ij} = \hat{y}_{ij} - \theta * \frac{|\hat{y}_{ij} - \bar{y}_j|}{s_j} \\ \text{S.T} \quad \left\{ \begin{array}{l} \text{Up}_{ij} > 0, \text{Low}_{ij} > 0 \\ \text{FICP}(\theta) > 0 \\ \mu \leq \text{FICP}(\theta) \leq 100\% \end{array} \right. \end{array} \right. \quad (3)$$

Where **FICP**(θ) is the coverage of the forecasting interval under the parameter. θ ; **MPI**(θ) is the average width of prediction interval under the parameter. θ ; θ is the parameter that needs to be optimized; Up_{ij} is the upper and Low_{ij} is the lower bounds of the prediction interval, respectively; \hat{y}_{ij} is the predicted value of the point of i-th data in j-th group obtained by FCM; \bar{y}_j is the average of j-th group; s_j is the standard deviation of j-th group.

2.2. The methodologies related to the proposed multivariable model

Section 2.2 provides a comprehensive overview of the algorithmic implementation processes detailed in this article.

2.2.1. Variational mode decomposition (VMD)

VMD is a signal processing method that decomposes highly nonlinear or nonstationary signals into a set of IMFs [34]. Each IMFs exhibits strict bandwidth and time-frequency localization characteristics.

$$\mathbf{r}(t) = \sum_{i=1}^n \mu_i(t) + \mathbf{r}_n(t) \quad (4)$$

The original signal is denoted as $\mathbf{r}(t)$, each individual IMF is represented by $\mu_i(t)$, and the residual signal is given by $\mathbf{r}_n(t)$. The total number of IMFs is denoted as n. IMFs exhibit favorable local time-frequency characteristics, making them suitable for signal analysis, denoising, feature extraction, and other applications.

Firstly, the analytic signal is obtained through the Hilbert transform. Subsequently, the signal is translated to baseband, and the H1 Gaussian smoothness is employed to estimate the signal bandwidth. This process is represented by Equation (6).

$$\left\{ \begin{array}{l} \text{Min} \left\{ \sum_K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * \mu_k(t) \right] e^{-j\varphi_k t} \right\|_2^2 \right\} \\ \text{s.t.} \sum_K \mu_k = f_0 \end{array} \right. \quad (5)$$

where $\{\mu_k\} = \{\mu_1, \dots, \mu_K\}$ is pattern sets, and $\{\varphi_k\} = \{\varphi_1, \dots, \varphi_K\}$ is center frequencies. To solve this issue, the above constrained optimization problem can be transformed into an unconstrained optimization

problem through augmented Lagrangian function, as follows:

$$\begin{aligned} L(\{\mu_k\}, \{\varphi_k\}, \lambda) = & \alpha \sum_k \left\| \partial_t \left\{ \left[\delta(t) + \frac{j}{\pi t} \right] * \mu_k \right\} e^{-j\varphi_k} \right\|^2 \\ & + \left\| f(t) - \sum_k \mu_k(t) \right\|^2 + \langle \lambda(t), f(t) - \sum_k \mu_k(t) \rangle \end{aligned} \quad (6)$$

The constrained optimization problem can be solved using the alternating direction method of multipliers that is based on the idea of fixing two variables and updating one of them, as follows:

$$\hat{\mu}_k^{n+1}(\varphi) = \frac{f(\varphi) - \sum_{i < k} \hat{\mu}_k^{n+1}(\varphi) - \sum_{i > k} \hat{\mu}_k^{n+1}(\varphi) + \frac{\hat{\lambda}^n(\varphi)}{2}}{1 + 2\alpha(\varphi - \varphi_k^n)^2} \quad (7)$$

$$\varphi_k^{n+1} = \frac{\int_0^\infty \varphi |\hat{\mu}_k^{n+1}(\varphi)|^2 d\varphi}{\int_0^\infty |\hat{\mu}_k^{n+1}(\varphi)|^2 d\varphi} \quad (8)$$

$$\hat{\lambda}^{n+1}(\varphi) = \hat{\lambda}^n(\varphi) + \tau \left(\hat{f}(\varphi) - \sum_k |\hat{\mu}_k^{n+1}(\varphi)| \right) \quad (9)$$

Ultimately, employing the concept of sequence reconstruction, the transformed sequence is formulated as follows [35].:

$$S_{Re} = \sum_k \omega_k * \mu_k = \omega_1 * \mu_1 + \omega_2 * \mu_2 + \dots + \omega_k * \mu_k \quad (10)$$

where S_{Re} is the reconstructed data ω_k is the weight of each mode, obtained by optimization algorithm.

2.2.2. Nondominated Neighbor Immune Algorithm (NNIA)

The weights of each IMF component in the data preprocessing module as well as the interval parameters θ in the interval prediction module are optimized using the NNIA [36].

To establish a comprehensive forecasting model, an innovative multiobjective optimization algorithm, NNIA, is introduced. The algorithm is employed for optimizing the weights in data preprocessing module and interval prediction module. The fundamental steps of NNIA can be outlined as follows.

Step 1: Initialization:

The algorithm begins by generating an initial population of candidate solutions: size of dominant population, clonal scale C_s , N_m size of active population N_a , and maximum number of iterations G_{max} , $N_a = 15$, $C_s = 200$, $N_m = 200$, $G_{max} = 300$, and evaluating the fitness of each solution using the objective function.

Step 2: antibodies clone operation.

In this step, the solutions are ranked based on their fitness values $F_{fitness}$ using non-dominated sorting.

$$\begin{cases} \max & F_{fitness}(x) = (f_1(x), f_2(x), \dots, f_k(x))^T \\ s.t. & x \in \Omega \end{cases} \quad (11)$$

The top solutions $B = \{b_1, b_2, \dots, b_n\}$ are selected according to a pre-defined clone size.

$$\begin{aligned} \forall p = 1, 2, \dots, k & \quad f_p(e^{-1}(b_j)) \geq f_p(e^{-1}(b_i)) \wedge \\ \exists q = 1, 2, \dots, k & \quad f_p(e^{-1}(b_j)) \geq f_p(e^{-1}(b_i)) \end{aligned} \quad (12)$$

Step 3: Mutation operation.

Mutation operators are applied to the selected solutions to generate diverse offspring solutions C_t and C'_t . The fitness of the mutated offspring is then updated.

Step 4: Suppression operation.

Dominant solutions D_t are identified from the offspring and inferior solutions in the population are replaced. The mutated offspring solutions are incorporated in the population.

Step 5: Immune memory update.

The non-dominated solutions are kept track of in the memory repository. The repository is updated by integrating the newly found non-

dominated solutions. If A_t is better than D_t , let $D_t = A_t$, otherwise, $D_t = D_t$.

Step 6: Termination criteria.

The algorithm checks if the termination conditions are met, such as a maximum number of iterations $t \geq G_{max}$. If not met, the algorithm returns to step 2, otherwise, it proceeds to the next step $t = t + 1$.

Step 7: Output.

The final set of non-dominated solutions D_{Gmax+1} is extracted as the algorithm's output.

2.2.3. The dual-layer LSTM

LSTM possesses the capability to capture long-term dependencies in sequences. Hochreiter and Schmid Huber initially proposed the LSTM network in 1997 [40], which has since undergone refinement and gained popularity through various contributions from researchers. The main equations of the two-layer LSTM is as follows:

$$\begin{cases} Z^t = g(\theta_z X^t + \zeta_z Y^{t-1} + b_z) \\ I^t = \sigma(\theta_i X^t + \zeta_i Y^{t-1} + \psi_i C^{t-1} + b_i) \\ F^t = \sigma(\theta_f X^t + \zeta_f Y^{t-1} + \psi_f C^{t-1} + b_f) \\ C^t = Z^t \cdot I^t + C^{t-1} \cdot F^t \\ o^t = \sigma(\theta_o X^t + \zeta_o Y^{t-1} + \psi_o C^{t-1} + b_o) \\ X_2^t = o^t \\ Z_2^t = g(\theta_z X_2^t + \zeta_z Y_2^{t-1} + b_{z2}) \\ I_2^t = \sigma(\theta_i X_2^t + \zeta_i Y_2^{t-1} + \psi_i C_2^{t-1} + b_{i2}) \\ F_2^t = \sigma(\theta_f X_2^t + \zeta_f Y_2^{t-1} + \psi_f C_2^{t-1} + b_{f2}) \\ C_2^t = Z^t \cdot I^t + C_2^{t-1} \cdot F_2^t \\ o_2^t = \sigma(\theta_o X_2^t + \zeta_o Y_2^{t-1} + \psi_o C_2^{t-1} + b_{o2}) \end{cases} \quad (13)$$

where Z^t is input block, I^t is input gate and F^t is forget gate's activation function. Y^t is the output of the cell at time t and C^t is the cell state; o^t is the activation of the output gate; $h(x)g(x)$, and $\sigma(x)$ are activation functions.

3. Dataset, evaluation indicators, and model parameters setting

This section focuses on the dataset used, evaluation indicators employed, and the setting of model parameters.

3.1. Dataset and training/test set settings

The four wind speed series are the data of Turkey in March (**dataset-Spring**), June (**dataset-Summer**), September (**dataset-Autumn**) and December (**dataset-Winter**) 2018, which include Wind Speed (**m/s**), Wind Direction, Theoretical Power Curve (**KWh**) and Low Voltage Active Power (**kW**). What's more the time interval is 10 min. The wind speed data is partitioned into two subsets: the train-set and the test-set. After conducting several experiments, it has been determined that the optimal ratio is 8:1 ($N_{input} : N_{output} = 8 : 1$). Therefore, in a dataset consisting of a total length of 2000 data points ($N_{data} = 2000$), the first 1600 data points ($N_{train} = 1600$) are utilized for training the model's architecture, while the remaining 400 data points ($N_{test} = 400$) serve as the test-set. The input of the first layer in the double-layer LSTM consists of the wind speed values at the first 8 time points, which are used to predict the wind speed at the 9th point. The output of the first layer represents the preliminary predicted value y' of the wind speed. In the second layer, the input includes the preliminary predicted value y' from the first layer, as well as additional features at the 8th time point such as Wind Direction, Theoretical Power Curve (KWh), and Low Voltage Active Power. In addition, this article provides a multi-step prediction of wind speed. The objective of multi-step prediction is to provide accurate forecasts for a sequence of future steps beyond the next immediate value. In the 2-step prediction, the predicted value obtained in the first step is used as the true value to predict the next wind speed value, while keeping everything else unchanged. Table 3 lists the data structure of wind speed of four seasons.

Table 3

The data structure of wind speed of four seasons.

Sites	Train test	Sample size	Data structure			
			Max	Min	Mean	Std.
Spring	All	2000	16.8438	0.2905	7.0370	3.8908
	Train	1600	16.2937	0.2905	6.6317	3.6419
	Test	400	16.8438	0.7286	8.6583	4.4042
Summary	All	2000	16.5546	0.3162	5.6839	3.9624
	Train	1600	16.5546	0.3162	5.6254	4.0650
	Test	400	13.4771	0.3405	5.9182	3.5171
Autumn	All	2000	16.5231	1.7548	10.3103	2.5057
	Train	1600	16.5231	4.4605	10.4120	2.4875
	Test	400	14.2299	1.7548	9.9036	2.5400
Winter	All	2000	20.4633	0.3667	8.6969	3.0960
	Train	1600	15.0079	0.3667	8.5098	2.6402
	Test	595	20.4633	3.4768	9.4452	4.4028

3.2. Performance indicators

Upon acquiring the forecasting wind speed data, it's important to build an explicit and standardized evaluation framework to assess the predictive ability of the model. Consequently, this research proposes a systematic assessment framework that encompasses seven indices, enabling a holistic assessment of the model's prediction performance in terms of both point prediction and interval precision. Specifically, MAE, MAPE, RMSE, and DACC are selected to measure prediction errors, while FICP, AWD, FINAW, and MPI measure the model's interval forecasting accuracy. Detailed information regarding these indices can be found in Table 4.

3.3. The relevant parameters of the multivariable forecasting system

This section delves into the pertinent parameters for every utilized models in this study, aiming to optimize the forecasting performance and attain superior forecasting efficacy.

(1) Variational mode decomposition

The variational model governing signal decomposition can be reformulated as an unconstrained optimization problem by introducing

a quadratic penalty term and a Lagrange multiplier (L). Through iterative alternating direction multiplier method updates process, all the modes of signal decomposition can be obtained. These modes include both the ones containing the main signal and the ones containing noise. By reconstructing the modes that contain the main signal, one can achieve a denoising effect. Six parameters are set in VMD to realize better effectiveness of denoising. α is set to 1900 ($\alpha = 1900$), which is the data fidelity constraint; the slack coefficient is set to zero ($sc = 0$); there are ten modes ($M = 10$); $DC = 0$; Initial value = 0; the tolerance criterion tol = $\widehat{10}(-8)$.

(2) Nondominated Neighbor Immune Algorithm

NNIA is a multi-objective optimization algorithm that incorporates principles from the immune system and nondomination. It generates an initial population, evaluates solution candidates, performs non-dominated sorting, stores top-ranking non-dominated solutions in immune memory, clones a subset of high-ranking solutions and introduces diversity by applying hypermutation, selects nondominated neighbors among the mutated clones, replaces a portion of the current population with new solutions, and iteratively repeats these steps until a termination criterion is met. NNIA needs to adjust several parameters according to different problems to achieve better optimization results, including the size of Frequency maximum ($F_{max} = 1$), Frequency minimum ($F_{min} = 0$), Polynomial variation parameter ($Pv = 0.5$), Analog binary cross ($Abc = 0.25$), Population ($P_{size} = 10$), and iterations ($It = 200$). While minor modifications to one of the parameters may limited impact on accuracy, it can significantly impact the program's runtime. Therefore, considering both factors, default parameters have been established for the NNIA model.

(3) The Dual-layer Long Short-Term Memory

The Dual-layer LSTM deals with the vanishing gradient problem by introducing gating mechanism for controlling the flow of information. The Dual-layer LSTM consists of memory cells that are connected through three distinct gates: input gate, forget gate, and output gate. The role of the input gate is to regulate the addition of information to the memory cell. The forget gate determines which data should be dis-

Table 4

Detailed information of the systematic assessment framework.

Indicators	Meaning	Equations
MAPE	The average absolute percentage error	$MAPE = \sum_{i=1}^N (y_i - \hat{y}_i) / y_i \times 100\% / N$
MAE	The average absolute error	$MAE = \sum_{i=1}^N (y_i - \hat{y}_i) / N$
RMSE	The square root of the average of the error squares	$RMSE = \sqrt{\sum_{i=1}^N (y_i - \hat{y}_i)^2} / N$
MSE	The average squared error	$MSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 / N$
R2	R-squared	$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$
DC	Trend prediction accuracy	$DC = \frac{100}{N-1} \sum_{t=1}^{N-1} a_t,$ $a_t = \begin{cases} 0, & \text{otherwise} \\ 1, & \text{if } ((y(t+1) - y(t))(\hat{y}(t+1) - \hat{y}(t))) > 0 \end{cases}$
FICP	Interval coverage	$FICP = \sum_{i=1}^N c_i \times 100\% / N$
MPI	The average predictions interval width	$MPI = \sum_{i=1}^N (U_i - L_i) / N$
FINAW	Forecast interval normalized average width	$FINAW = \sum_{i=1}^N MPI / (Max(y) - Min(y)) \times 100\% / N$
AWD	Accumulated width deviation of testing dataset	$AWD = \sum_{i=1}^N AWD_i / N$ $AWD_i = \begin{cases} (\hat{y}_i - U_i) / (U_i - L_i), & \hat{y}_i \geq U_i \\ 0, & L_i \leq \hat{y}_i \leq U_i \\ (L_i - \hat{y}_i) / (U_i - L_i), & \hat{y}_i \leq L_i \end{cases}$

Note: U_i is the upper bound of the interval, L_i is the lower bound of the interval, \hat{y}_i is the forecasting value, y_i is the raw data, \bar{y}_i is the average value N is the amount of samples.

carded, while the output gate controls the information that should be output from the memory cell. The Dual-layer LSTM achieves multivariate wind speed prediction by overlaying two layers of LSTM, with specific parameters as follows: The optimizer d is ADAM; the maximum Epochs ($Epochs_{max} = 200$); the Gradient Threshold ($GT = 1$); the Learn Rate Drop Factor ($LRDF = 0.18$); the Learn Rate Drop Period ($LRDP = 125$); the Learn Rate Schedule ($LRS = piecewise$); the Initial Learn Rate ($ILR = 0.006$); the Verbose is 0; the amount of input nodes for the first layer of LSTM is 8 ($N_{features}^1 = 8$); the amount of input nodes for the second layer of LSTM is 4 ($N_{features}^2 = 4$); the number of hidden layers for both layers of LSTM is 200 ($N_{hidden}^1 = N_{hidden}^2 = 200$), and the number of output nodes is 1 ($N_{responses}^1 = N_{responses}^2 = 1$).

4. Multivariate wind speed prediction experiments

This section outlines three experiments that were carried out for verifying predictive proficiency of the proposed forecasting model. These experiments aim at analyzing and comparing the effectiveness of our forecasting system with alternative models using the acquired wind speed time series as input. Through these experiments, we provide a detailed comparison of the multivariate wind speed forecasting model's effectiveness in predicting wind speed.

4.1. Experiment I: comparative analysis of point forecasting accuracy against traditional models

Experiment I aimed to compare the point forecasting accuracy of the proposed model with benchmark models. The experiment utilized four datasets (**dataset-Spring**, **dataset-Summer**, **dataset-Autumn** and

dataset-Winter) and evaluated each model's performance based on Mean Absolute Percentage Error (ϵ_{MAPE}), Mean Absolute Error (ϵ_{MAE}), Root Mean Squared Error (ϵ_{RMSE}), Mean Squared Error (ϵ_{MSE}), R-squared (R^2) and Directional Change (ϵ_{DC}). The objective was to determine whether the proposed model exhibited superior predictive accuracy and provide insights into its effectiveness compared to traditional approaches. The parameter settings of the Dual-layer LSTM are as follows: $Epochs_{max} = 200$; $GT = 1$; $LRDP = 125$; $LRDF = 0.18$; $LRS = piecewise$; $ILR = 0.006$; $Verbose = 0$. The results of Experiment I, listed in Table 5 and Fig. 2, revealed several key findings. In Set-Spring, both in one-step and two-step prediction, the developed model demonstrates remarkable performance. For one-step, the proposed multivariable forecasting model has $\epsilon_{MAPE}^{Proposed} = 4.1777$, which is significantly lower than the ϵ_{MAPE}^{Others} of the benchmark models (ranging from **9.5023 to 13.5358**), and ARIMA is the best one, whose parameters are *autoregressive coefficient* ($p = 4$), *Number of differences* ($d = 1$) and *Moving average coefficient* ($q = 3$). The proposed model demonstrates $\epsilon_{MAE}^{Proposed} = 0.2703$, which is smaller than the ϵ_{MAE} of the benchmark models (ranging from **0.5958 to 0.7982**), which demonstrates that the proposed model exhibits the smallest average percentage prediction error in comparison.

The lower ϵ_{MAPE} and ϵ_{MAE} values indicate that the proposed multivariable forecasting model exhibits a higher accuracy and precision in predicting wind speeds compared to the other models. It accurately captures the patterns and trends in the data, leading to smaller prediction errors. The proposed model demonstrated $\epsilon_{DC}^{Proposed} = 88.7218$, significantly higher than the ϵ_{DC}^{Others} of other models (ranging from **48.3709 to 56.2814**). The higher ϵ_{DC} signifies that the proposed model exhibits greater accuracy and precision in predicting the direction of wind speed trends compared to other models. Table 5 provides detailed

Table 5
Comparative analysis of point forecasting accuracy against traditional models.

Data	Models	Step 1						Step 2					
		MAPE	MAE	RMSE	MSE	R ²	DC	MAPE	MAE	RMSE	MSE	R ²	DC
Spring	ARIMA	9.5023	0.5958	0.8048	0.6477	0.9665	56.2814	12.5375	0.7270	0.9769	0.9544	0.9506	55.2764
	BPNN	10.2185	0.6521	0.9078	0.8241	0.9574	51.3784	14.0787	0.8908	1.2372	1.5308	0.9208	56.0302
	ELMAN	13.5358	0.7982	1.0911	1.1904	0.9384	56.2814	13.5358	0.7982	1.0911	1.1904	0.9384	56.2814
	SVM	9.7153	0.6256	0.8521	0.7260	0.9625	48.3709	12.8230	0.7393	1.0009	1.0018	0.9482	55.2764
	FTDNN	10.2480	0.6881	0.9782	0.9570	0.9505	49.1228	11.0626	0.7594	1.3265	1.7595	0.9090	52.0101
	NAR	10.3445	0.6584	0.9013	0.8123	0.9580	51.3784	14.4243	0.9609	1.3417	1.8001	0.9068	56.5327
	LSTM	10.0394	0.6246	0.8545	0.7302	0.9623	48.3709	13.0558	0.8212	1.1177	1.2492	0.9354	55.0251
Summer	NVDL	4.1777	0.2703	0.3667	0.1344	0.9931	88.7218	6.1623	0.4449	0.6203	0.3848	0.9802	74.9373
	ARIMA	12.1139	0.4653	0.6557	0.4300	0.9652	58.8972	17.0438	0.6632	0.9343	0.8730	0.9293	53.5176
	BPNN	14.1330	0.5158	0.7117	0.5065	0.9590	52.8822	20.5040	0.7310	1.0040	1.0079	0.9184	50.2513
	ELMAN	17.4915	0.5784	0.7836	0.6141	0.9502	51.8797	21.3906	0.7924	1.1249	1.2654	0.8976	52.2613
	SVM	12.4822	0.4776	0.6686	0.4470	0.9638	53.6341	18.1275	0.6748	0.9433	0.8899	0.9280	51.2563
	FTDNN	15.6893	0.6735	1.3948	1.9455	0.8423	53.6341	21.6708	0.8433	1.1261	1.2680	0.8974	55.2764
	NAR	13.0033	0.5168	0.7154	0.5118	0.9585	50.1253	26.1221	0.8515	1.1651	1.3574	0.8901	53.7688
Autumn	LSTM	12.3620	0.4886	0.6860	0.4706	0.9619	52.3810	20.5477	0.7095	0.9821	0.9645	0.9219	51.5075
	NVDL	3.0479	0.1058	0.1412	0.0199	0.9984	92.2306	5.5631	0.2138	0.2896	0.0839	0.9932	82.4561
	ARIMA	5.6083	0.4730	0.6182	0.3822	0.9406	51.6291	7.4978	0.6358	0.8194	0.6714	0.8955	49.4975
	BPNN	9.0239	0.5665	0.7926	0.6282	0.9024	57.3935	11.2993	0.7264	0.9926	0.9853	0.8466	53.0151
	ELMAN	6.0758	0.4992	0.6523	0.4255	0.9339	50.1253	10.8046	0.8632	1.1002	1.2104	0.8115	54.5226
	SVM	5.6230	0.4740	0.6187	0.3828	0.9405	52.3810	7.5411	0.6390	0.8191	0.6709	0.8955	50.0000
	FTDNN	6.1812	0.5003	0.6615	0.4375	0.9320	56.1404	6.3737	0.5143	0.7164	0.5132	0.9201	58.7940
Winter	NAR	7.6809	0.5266	0.7083	0.5017	0.9220	57.8947	11.6601	0.7527	1.0331	1.0674	0.8338	51.5075
	LSTM	5.9140	0.4851	0.6273	0.3935	0.9388	53.8847	8.3936	0.6939	0.8697	0.7563	0.8822	51.5075
	NVDL	2.2920	0.1973	0.2616	0.0685	0.9894	87.7193	5.2206	0.4768	0.6273	0.3935	0.9406	58.1454
	ARIMA	7.3280	0.5755	0.7607	0.5787	0.9701	52.1303	7.3307	0.5756	0.7612	0.5795	0.9701	52.2613
	BPNN	9.1383	0.8783	1.2181	1.4837	0.9233	52.8822	12.6879	1.2031	1.6323	2.6645	0.8625	52.5126
	ELMAN	9.9507	0.7785	1.0114	1.0229	0.9471	53.1328	14.8702	1.5555	2.2708	5.1567	0.7339	51.7588
	SVM	7.3860	0.5778	0.7582	0.5748	0.9703	51.6291	10.6504	0.8324	1.0878	1.1833	0.9389	53.2663
Note:	FTDNN	10.3568	1.0004	1.4279	2.0390	0.8945	52.3810	15.3517	1.4615	2.0134	4.0537	0.7908	50.0000
	NAR	8.2562	0.6838	0.9389	0.8815	0.9544	54.3860	16.8061	1.8688	3.1119	9.6837	0.5004	51.2563
	LSTM	8.4119	0.7446	0.9817	0.9637	0.9502	52.1303	13.6880	1.1913	1.5599	2.4332	0.8745	54.7739
	NVDL	3.2389	0.2481	0.3217	0.1035	0.9946	86.2155	8.5306	0.6597	0.8535	0.7284	0.9627	57.8947

Note: The bold font is an experimental result obtained from the proposed model in this article. NVDL is the proposed model. MAPE is the average of absolute percentage error. MAE is the mean absolute error of forecasting results. RMSE is the square root of the average of the error squares. MSE is mean Squared Error. R² is R-square. DC is directional Change.

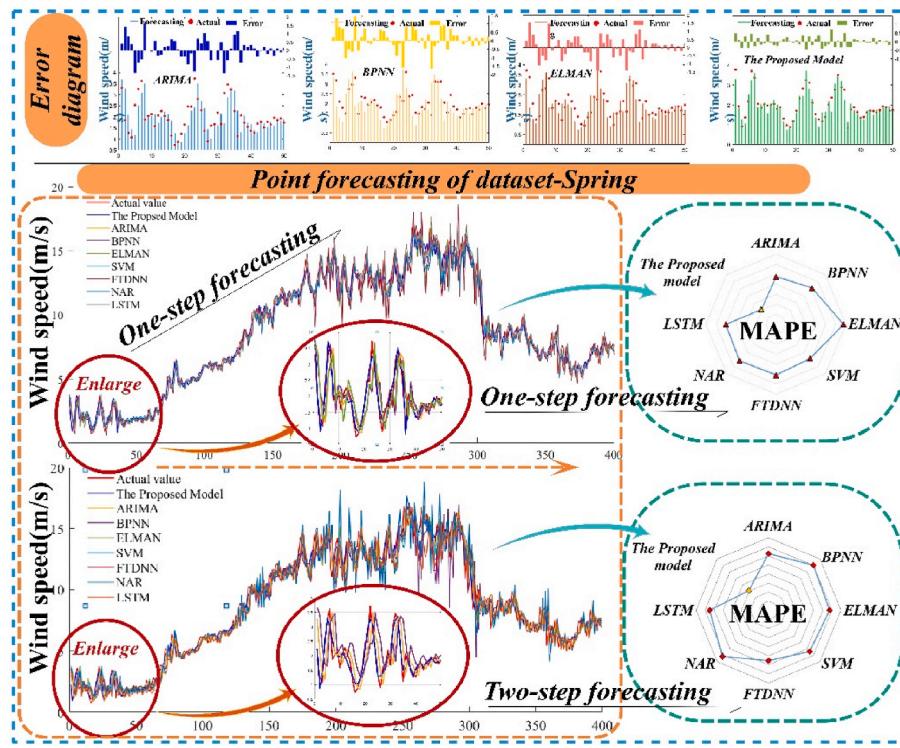


Fig. 2. Comparative analysis of point forecasting accuracy against traditional models.

evidences that the proposed multivariable forecasting model, characterized by the lowest $\epsilon_{MAPE} = 4.1777\%$ (**one-step**), 3.0479% (**one-step**), 2.2920% (**one-step**) and 3.2389% (**one-step**) exhibits optimal prediction performance compared to the hybrid models. Additionally, other assessment indicators also demonstrate favorable values, further substantiating the effectiveness of the proposed model for multivariable wind speed forecasting.

Fig. 2 presents a comparative analysis of wind speed predictions of **dataset-Spring** using different algorithms. The observed trend shows that the predicted values got by the proposed algorithm closely align

with the real values, demonstrating lower errors. Notably, in both one and two step prediction part, the proposed model consistently achieves the highest level of prediction accuracy. Furthermore, the evaluation indicators consistently highlight the preeminence of the proposed algorithm. In summary, Experiment I compared the point forecasting accuracy of the proposed model with traditional models and found that the proposed model exhibited great effectiveness in terms of lower ϵ_{MAPE} and ϵ_{MAE} values, as well as higher ϵ_{DC} . The results suggest that the proposed model has the potential to improve forecasting accuracy and contribute to advancements in point forecasting methodology for

Table 6
Comparative analysis of point forecasting accuracy against other denoising models.

Data	Models	step 1						step 2					
		MAPE	MAE	RMSE	MSE	R2	DC	MAPE	MAE	RMSE	MSE	R2	DC
Spring	EMD-LSTM	8.1847	0.5154	0.6851	0.4694	0.9757	73.1830	13.0133	0.7084	0.9029	0.8152	0.9578	60.8040
	CEEMD-LSTM	6.4397	0.3999	0.5339	0.2851	0.9853	75.6892	10.4292	0.6598	0.8658	0.7495	0.9612	58.5427
	SSA-LSTM	10.0230	0.5446	0.7131	0.5085	0.9737	68.6717	11.6804	0.6039	0.8033	0.6452	0.9670	69.3267
	VMD-LSTM	7.7717	0.3998	0.5310	0.2819	0.9854	77.4436	8.0674	0.4915	0.6525	0.4258	0.9780	68.3417
	NVDL	4.1777	0.2703	0.3666	0.1344	0.9931	88.7218	6.1623	0.4449	0.6203	0.3848	0.9802	74.9373
Summer	EMD-LSTM	10.2397	0.3658	0.4939	0.2439	0.9802	74.1855	15.2190	0.5665	0.7705	0.5936	0.9520	55.2764
	CEEMD-LSTM	14.3041	0.3794	0.4924	0.2425	0.9803	69.4236	19.3281	0.5629	0.7512	0.5643	0.9543	56.7839
	SSA-LSTM	13.1448	0.5133	0.6807	0.4634	0.9624	69.1729	22.7141	0.5956	0.7596	0.5770	0.9533	62.0603
	VMD-LSTM	11.5869	0.3278	0.4353	0.1894	0.9846	74.4361	16.3598	0.4687	0.5979	0.3574	0.9711	65.3266
	NVDL	3.0479	0.1058	0.1412	0.0199	0.9984	92.2306	5.5631	0.2138	0.2896	0.0839	0.9932	82.4561
Autumn	EMD-LSTM	3.3954	0.2885	0.3682	0.1356	0.9789	79.9499	6.0921	0.5103	0.6723	0.4519	0.9296	60.0503
	CEEMD-LSTM	3.5529	0.2878	0.3761	0.1415	0.9780	81.4536	5.4448	0.4683	0.6063	0.3676	0.9428	59.2965
	SSA-LSTM	5.6091	0.4830	0.6299	0.3968	0.9383	66.9173	7.7074	0.5295	0.6895	0.4754	0.9260	65.0754
	VMD-LSTM	4.9203	0.4628	0.5486	0.3009	0.9532	62.6566	6.0730	0.4704	0.6016	0.3619	0.9436	59.0452
	NVDL	2.2920	0.1973	0.2616	0.0685	0.9894	87.7193	5.2206	0.4768	0.6273	0.3935	0.9406	58.1454
Winter	EMD-LSTM	5.9251	0.5091	0.7187	0.5165	0.9733	76.9424	7.8494	0.7135	0.9363	0.8767	0.9548	66.5829
	CEEMD-LSTM	4.5346	0.3686	0.4781	0.2286	0.9882	76.6917	9.2190	0.8181	1.0550	1.1131	0.9426	57.0352
	SSA-LSTM	7.0186	0.5396	0.7240	0.5242	0.9729	67.4185	7.6711	0.6230	0.8025	0.6440	0.9668	61.8090
	VMD-LSTM	6.3552	0.5932	0.7722	0.5964	0.9696	64.9254	9.6954	0.8423	1.0864	1.1803	0.9391	53.5176
	NVDL	3.2389	0.2481	0.3217	0.1035	0.9946	86.2155	8.5306	0.6597	0.8535	0.7284	0.9627	57.8947

Note: The bold font is an experimental result obtained from the proposed model in this article. **MAPE** is the average of absolute percentage error. **MAE** is the mean absolute error of forecasting results. **RMSE** is the square root of the average of the error squares. **MSE** is mean Squared Error. **R²** is R-square. **DC** is directional Change. **NVDL** is the proposed model.

practical applications.

4.2. Experiment II: comparative analysis of point forecasting accuracy against other hybrid models

Experiment II aims to elucidate the comparative analysis of point forecasting accuracy of the proposed multivariable forecasting model against other preprocessing models, namely M_{EMD} , M_{CEEMD} , M_{SSA} , M_{VMD} , and the adaptive preprocessing model $M_{Adaptive-VDM}$, whose parameters are $\alpha = 1900$, $sc = 0$, $M = 10$, $DC = 0$; Initial value = 0; $tol = \widehat{10}(-8)$. The number of IMFs for M_{EMD} , M_{CEEMD} , M_{VMD} , and the adaptive preprocessing model $M_{Adaptive-VDM}$ is 10, and the highest frequency IMF is removed. The Window length of M_{SSA} is [1, 50]. The evaluation indicators are provided in Table 6 and Table 7 while the comparison chart of experimental results is visualized in Fig. 3. Table 6 compares the point forecasting accuracy against other denoising models, and Table 7 compares other hybrid models. For Set-Spring, in terms of $\epsilon_{MAPE}^{Proposed}$, the proposed model value has an improvement of approximately 34.77%. Similarly, the $\epsilon_{MAE}^{Proposed}$ improvement is about 33.17%, the $\epsilon_{RMSE}^{Proposed}$ improvement is around 30.61%, the $\epsilon_{MSE}^{Proposed}$ improvement is approximately 51.88%, the R^2 improvement is about 0.75%, and the $\epsilon_{DC}^{Proposed}$ improvement is approximately 13.55%. These calculations show that the proposed model value exhibits significant improvements across all metrics when compared to EMD-ARIMA. The proposed model value gets lower error and greater accuracy, indicating a notable advancement in wind speed forecasting performance (see Table 8).

When comparing the effectiveness of the proposed model value and the NNIA-VMD-LSTM value in terms of DC metric, specific values and improvement percentages are considered. The proposed model has $DC = 88.7218$, while the DC for the NNIA-VMD-LSTM is 81.7043. This yields an improvement percentage of approximately 8.59% [$(88.7218 - 81.7043)/81.7043 * 100$], when comparing the proposed model to the NNIA-VMD-LSTM. This suggests that the proposed model outperforms the NNIA-VMD-LSTM in the context of wind speed prediction. In contrast to the subjective determination of denoising layer number in recombination, our approach assigns weights to each sub-sequence obtained through optimization algorithm. The findings from Experiment II consistently demonstrate the superior effectiveness of the proposed multivariable forecasting model over other preprocessing techniques, highlighting the potential of the adaptive combination reconstruction approach. Fig. 3 illustrates the comparison of prediction outcomes from various algorithms for summer and autumn wind speed data points. The line graph clearly demonstrates that the predicted values generated by the method proposed in this study closely align with the actual wind speed data. Moreover, the bar chart and radar chart provide additional

evidence supporting the superior forecasting effectiveness of the proposed method. Furthermore, the scatter plots depicting the relationship between real and predicted values highlight that the proposed algorithm exhibits the lowest degree of scatter, thereby reaffirming its exceptional performance.

4.3. Experiment III: comparative analysis of interval forecasting accuracy among different models

The present study aims to compare and analyze the interval forecasting accuracy of different models for wind speed prediction. Interval forecasting is a crucial technique that provides information on the future range of wind speed, and various models may yield diverse interval prediction results, which necessitates a precise evaluation and comparison of predictive accuracy. To this end, the study evaluates the performance of each model using several metrics including γ_{FICP} , γ_{MPI} , γ_{AWD} , γ_{FINAW} . Specifically, evaluations were conducted on the EMD-ARIMA, EMD-GRNN, CEEMD-FTDNN, SSA-NAR, VMD-LSTM, NNIA-VMD-SVM, NNIA-VMD-LSTM models, and the proposed model. The results of Experiment III are shown in Table 8. Regarding the NNIA model, there are the several parameters: $F_{max} = 1$, $F_{min} = 0$, $Abc = 0.25$, $Pv = 0.5$, $P_{size} = 10$, and $It = 200$. The Proposed model achieved $\gamma_{FICP}^{Proposed} = 97.2500$, $\gamma_{FINAW}^{Proposed} = 0.1147$, $\gamma_{AWD}^{Proposed} = 0.0054$, and $\gamma_{MPI}^{Proposed} = 1.8488$. Comparing these metrics against the other models, The Proposed model exhibits a notable improvement in precision. It achieves an increase in γ_{FICP} of approximately 3.25% when compared to the highest performing model, NNIA-VMD-LSTM. The Proposed model also outperforms all other models in terms of γ_{MPI} , γ_{AWD} , and γ_{FINAW} by achieving lower values. Overall, The Proposed model shows significant improvements in accuracy compared to the other models. It achieves higher γ_{FICP} , lower γ_{MPI} , γ_{AWD} , and γ_{FINAW} values, indicating enhanced precision and reliability in wind speed interval predictions. According to Fig. 4, The majority of wind speed data falls within the predicted range, with only a limited number of points falling outside. To achieve the same level of prediction coverage as one-step prediction, the interval width of two-step interval prediction is wider. Comparing the performance indicators, it is evident from the chart that the algorithm proposed in this study achieves a high prediction coverage while maintaining a narrow interval width. Therefore, the method presented in this article demonstrates superior performance.

5. Conclusions

Wind energy has emerged as a crucial component in addressing the global need for clean and sustainable energy sources. Precise prediction of wind velocity holds significant importance in maximizing the efficient

Table 7
Comparative analysis of point forecasting accuracy against other hybrid models.

Data	Models	step 1						step 2					
		MAPE	MAE	RMSE	MSE	R2	DC	MAPE	MAE	RMSE	MSE	R2	DC
Spring	NVS	4.7767	0.3086	0.4247	0.1803	0.9907	86.2155	8.5175	0.5391	0.7257	0.5266	0.9727	66.0804
	NVL	6.0421	0.3746	0.5270	0.2778	0.9856	81.7043	10.0937	0.6066	0.8110	0.6578	0.9660	58.2915
	NVDL	4.1777	0.2703	0.3666	0.1344	0.9931	88.7218	6.1623	0.4449	0.6203	0.3848	0.9802	74.9373
Summer	NVS	5.9872	0.2225	0.3111	0.0968	0.9922	91.7293	11.8314	0.4538	0.6448	0.4157	0.9664	59.7990
	NVL	10.7647	0.4018	0.5338	0.2850	0.9769	61.6541	18.6779	0.5976	0.7826	0.6125	0.9504	53.2663
	NVDL	3.0479	0.1058	0.1412	0.0199	0.9984	92.2306	5.5631	0.2138	0.2896	0.0839	0.9932	82.4561
Autumn	NVS	2.8985	0.2415	0.3155	0.0996	0.9845	84.9624	5.3849	0.4573	0.5913	0.3496	0.9456	58.2915
	NVL	3.1719	0.2694	0.3518	0.1238	0.9808	83.4586	5.4406	0.4772	0.6258	0.3916	0.9390	56.7839
	NVDL	2.2920	0.1973	0.2616	0.0685	0.9894	87.7193	5.2206	0.4768	0.6273	0.3935	0.9406	58.1454
Winter	NVS	3.9326	0.3026	0.3923	0.1539	0.9920	83.2080	7.5469	0.5882	0.7516	0.5649	0.9709	56.0302
	NVL	4.7365	0.3663	0.4712	0.2220	0.9887	77.1144	9.9484	0.8802	1.4632	2.1410	0.8895	56.5327
	NVDL	3.2389	0.2481	0.3217	0.1035	0.9946	86.2155	8.5306	0.6597	0.8535	0.7284	0.9627	57.8947

Note: The bold font is an experimental result obtained from the proposed model in this article. MAPE is the average of absolute percentage error. MAE is the mean absolute error of forecasting results. RMSE is the square root of the average of the error squares. MSE is mean Squared Error. R² is R-square. DC is directional Change. NVDL is the proposed model. NVS is NNIA-VMD-SVM. NVL is NNIA-VMD-LSTM.

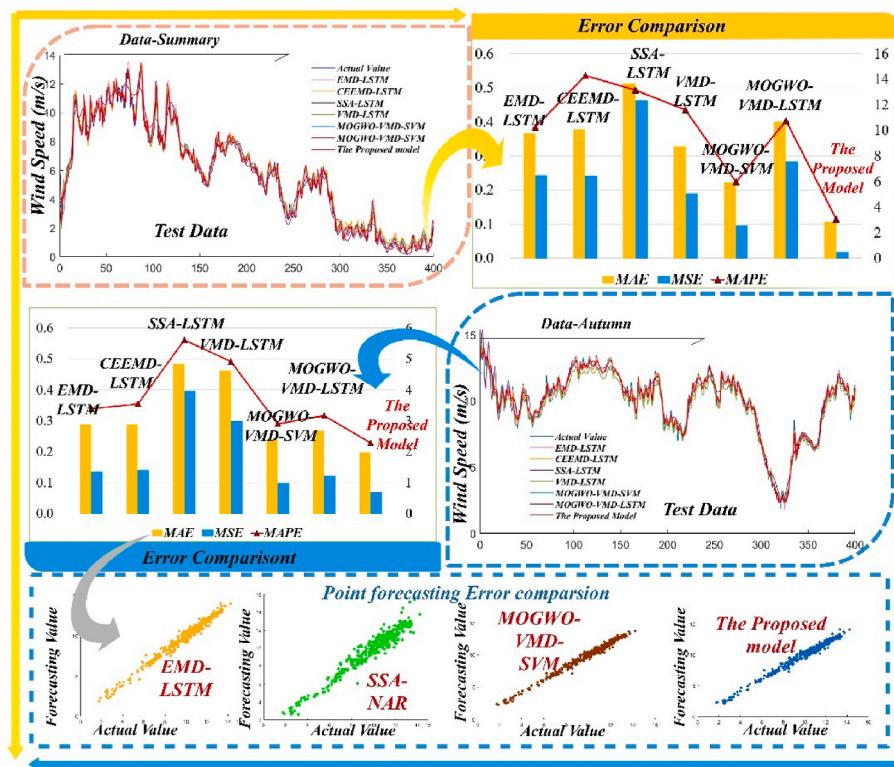


Fig. 3. Experiment II: Comparison with models employing diverse data pretreatment techniques.

Table 8
Comparative analysis of interval forecasting accuracy among different models.

Data	Models	Step 1				Step 2			
		FICP	FINAW	AWD	MPI	FICP	FINAW	AWD	MPI
Spring	EMD-LSTM	95.0372	0.1672	0.0086	2.7374	96.0199	0.2583	0.0057	4.1682
	CEEMD-LSTM	96.7742	0.1609	0.0050	2.6647	95.5224	0.2442	0.0064	3.9522
	SSA-LSTM	92.0596	0.1808	0.0157	2.9532	96.5174	0.2616	0.0025	4.2466
	VMD-LSTM	93.7500	0.1393	0.0109	2.2445	96.7419	0.2121	0.0044	3.4180
	NVS	96.5000	0.1158	0.0095	1.8664	97.2431	0.2246	0.0040	3.6190
	NVL	95.2500	0.1381	0.0083	2.2249	96.2406	0.2203	0.0060	3.5498
Summer	NVDL	97.2500	0.1147	0.0054	1.8488	97.5000	0.2075	0.0040	3.4023
	EMD-LSTM	94.7891	0.1500	0.0069	1.9850	96.0199	0.2232	0.0079	3.3216
	CEEMD-LSTM	95.5335	0.1578	0.0053	2.2191	94.5274	0.2065	0.0110	3.9200
	SSA-LSTM	94.7891	0.1895	0.0079	2.4896	96.0199	0.2040	0.0064	4.5653
	VMD-LSTM	96.0000	0.1596	0.0064	2.0967	94.9875	0.1945	0.0076	2.5554
	NVS	95.7500	0.0959	0.0068	1.2604	96.2406	0.2056	0.0076	2.7010
Autumn	NVL	93.7500	0.1599	0.0109	2.1007	95.7393	0.2345	0.0072	3.0811
	NVDL	96.7500	0.0479	0.0047	0.6296	96.5000	0.0915	0.0047	1.2134
	EMD-LSTM	93.5484	0.1615	0.0128	2.2766	95.7711	0.2862	0.0048	3.9434
	CEEMD-LSTM	93.5484	0.1783	0.0092	2.5123	95.2736	0.2788	0.0069	3.8648
	SSA-LSTM	90.5707	0.3059	0.0188	4.2310	94.0299	0.2846	0.0073	3.9147
	VMD-LSTM	96.0000	0.2666	0.0055	3.3259	96.2406	0.3173	0.0038	3.9588
Winter	NVS	93.7500	0.1441	0.0094	1.7979	96.9925	0.3226	0.0044	4.0246
	NVL	94.5000	0.1726	0.0089	2.1530	96.9925	0.3430	0.0028	4.2792
	NVDL	96.0000	0.1348	0.0066	1.6811	95.2500	0.3458	0.0054	4.7559
	EMD-LSTM	94.0447	0.1613	0.0116	3.2188	96.7662	0.2714	0.0046	5.3598
	CEEMD-LSTM	95.2854	0.1428	0.0093	2.8902	93.7811	0.2631	0.0108	5.1687
	SSA-LSTM	94.0447	0.1985	0.0133	3.9573	96.5174	0.2628	0.0057	5.2148
NVS	VMD-LSTM	92.8040	0.1812	0.0100	3.5991	94.9875	0.2827	0.0085	4.8022
	NVDL	95.5000	0.1355	0.0076	2.3013	96.9925	0.3213	0.0044	5.4581
	EMD-LSTM	94.0447	0.1405	0.0121	2.8430	96.9925	0.4700	0.0148	7.9839
	CEEMD-LSTM	95.5000	0.1148	0.0071	1.9505	96.0000	0.3148	0.0077	6.2340

Note: The bold font is an experimental result obtained from the proposed model in this article. FICP is forecast interval coverage probability. MPI is the average predictions interval width. FINAW is forecast interval normalized average width. AWD is accumulated width deviation of testing dataset. NVDL is the proposed model. NVS is NNIA-VMD-SVM. NVL is NNIA-VMD-LSTM.

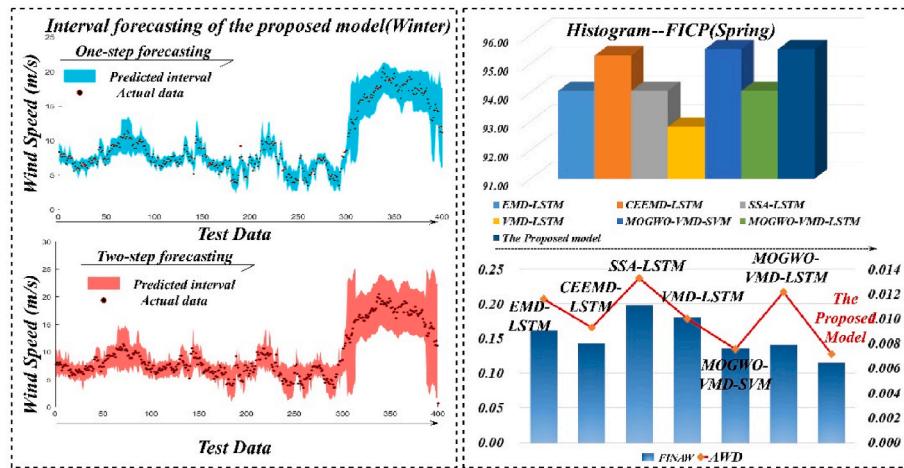


Fig. 4. Experiment III: Comparison with classic individual models.

utilization of wind power resources and facilitating the transition towards a low-carbon energy landscape. Continued advancements in wind speed prediction techniques are essential for enhancing the efficiency, reliability, and economic viability of wind energy systems. In this context, this paper presents an advanced multi-dimensional time series prediction system that integrates data preprocessing techniques, multi-objective optimization theory, and a deep-learning model. The proposed model introduces a Dual-layer LSTM model specifically designed for wind speed forecasting in multivariate time series data. The model demonstrates exceptional accuracy, stability, and generalization ability. One of the key strengths of the Dual-layer LSTM model lies in its capability to capture complex spatio-temporal patterns by considering interdependencies among different variables. This feature makes the model a valuable tool for various applications, including renewable energy management, and weather prediction. This study provides empirical evidence that the proposed system outperforms benchmark models in terms of both stability and accuracy. Its superior performance establishes it as an effective tool for applications such as smart grid programming and other related fields. Furthermore, the continued development and refinement of wind speed forecasting techniques will contribute to the optimization of wind energy resource utilization. It is helpful for increasing operational efficiency, reducing reliance on fossil fuels, and mitigating environmental impacts, thereby conducive to the global shift towards a cleaner and more sustainable energy future.

CRediT authorship contribution statement

Haipeng Zhang: Writing – review & editing, Writing – original draft,

Appendix

The comparison of prediction results of different models for the entire sample is list in [Table 9](#).

Table 9

Comparison of prediction results of different models for the entire sample.

Models	MAPE	MAE	RMSE	MSE	R2	DC	FICP	FINAW	AWD	MPI
ARIMA	8.87	0.58	0.80	0.63	0.95	51.75	87.2	0.160	0.030	3.209
BPNN	9.31	0.62	0.85	0.72	0.94	51.05	85.3	0.146	0.037	2.937
SVM	8.85	0.58	0.80	0.64	0.95	52.25	87.7	0.160	0.030	3.224
FTDNN	9.38	0.66	0.98	0.95	0.93	51.05	86.7	0.168	0.036	3.368
NAR	9.15	0.61	0.83	0.68	0.95	50.55	87.3	0.145	0.031	2.923
LSTM	9.07	0.60	0.83	0.70	0.95	52.05	89.1	0.168	0.027	3.370
EMD-LSTM	6.67	0.41	0.55	0.31	0.98	74.47	95.5	0.156	0.007	3.144
CEEMD-LSTM	7.45	0.51	0.66	0.43	0.97	69.17	93.8	0.169	0.009	3.397

(continued on next page)

Table 9 (continued)

Models	MAPE	MAE	RMSE	MSE	R2	DC	FICP	FINAW	AWD	MPI
SSA-LSTM	8.97	0.56	0.75	0.56	0.96	65.07	86.6	0.147	0.035	2.963
VMD-LSTM	6.02	0.39	0.54	0.29	0.98	78.98	96.8	0.164	0.005	3.303
NVS	6.72	0.41	0.56	0.32	0.98	76.48	96.10	0.16	0.01	3.26
NVL	6.80	0.40	0.53	0.28	0.98	75.08	95.9	0.149	0.007	2.995
NVDL	5.08	0.32	0.43	0.19	0.99	84.78	97.6	0.156	0.004	3.127

References

- [1] Wang J, Zhang H, Li Q, Ji A. Design and research of hybrid forecasting system for wind speed point forecasting and fuzzy interval forecasting. *Expert Syst Appl* 2022;209. <https://doi.org/10.1016/j.eswa.2022.118384>.
- [2] Liu C, Zhang X, Mei S, Zhen Z, Jia M, Li Z, Tang H. Numerical weather prediction enhanced wind power forecasting: rank ensemble and probabilistic fluctuation awareness. *Appl Energy* 2022;313. <https://doi.org/10.1016/j.apenergy.2022.118769>.
- [3] Yakoub G, Mathew S, Leal J. Intelligent estimation of wind farm performance with direct and indirect ‘point’ forecasting approaches integrating several NWP models. *Energy* 2023;263. <https://doi.org/10.1016/j.energy.2022.125893>.
- [4] Al-Duaies FS, Al-Sharqi RS. A unique Markov chain Monte Carlo method for forecasting wind power utilizing time series model. *Alex Eng J* 2023;74. <https://doi.org/10.1016/j.aej.2023.05.019>.
- [5] Aasim, Singh SN, Mohapatra A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. *Renew Energy* 2019;136. <https://doi.org/10.1016/j.renene.2019.01.031>.
- [6] Yousuf MU, Al-Bahadly I, Avci E. A modified GM(1,1) model to accurately predict wind speed. *Sustain Energy Technol Assessments* 2021;43. <https://doi.org/10.1016/j.seta.2020.100905>.
- [7] Liu H, Tian H qi, Li Y fei. An EMD-recursive ARIMA method to predict wind speed for railway strong wind warning system. *J Wind Eng Ind Aerod* 2015;141. <https://doi.org/10.1016/j.jweia.2015.02.004>.
- [8] Liu H, Erdem E, Shi J. Comprehensive evaluation of ARMA-GARCH(-M) approaches for modeling the mean and volatility of wind speed. *Appl Energy* 2011;88(3). <https://doi.org/10.1016/j.apenergy.2010.09.028>.
- [9] Li J, Wang J, Zhang H, Li Z. An innovative combined model based on multi-objective optimization approach for forecasting short-term wind speed: a case study in China. *Renewable Energy*; 2022. p. 201. <https://doi.org/10.1016/j.renene.2022.10.123>.
- [10] Dong Y, Li J, Liu Z, Niu X, Wang J. Ensemble wind speed forecasting system based on optimal model adaptive selection strategy: case study in China. *Sustain Energy Technol Assessments* 2022;53. <https://doi.org/10.1016/j.seta.2022.102535>.
- [11] Hu Y, Guo Y, Fu R. A novel wind speed forecasting combined model using variational mode decomposition, sparse auto-encoder and optimized fuzzy cognitive mapping network. *Energy* 2023;278. <https://doi.org/10.1016/j.energy.2023.127926>.
- [12] Du S, Li T, Yang Y, Horng SJ. Multivariate time series forecasting via attention-based encoder-decoder framework. *Neurocomputing* 2020;388. <https://doi.org/10.1016/j.neucom.2019.12.118>.
- [13] Xu M, Han M. Adaptive elastic echo state network for multivariate time series prediction. *IEEE Trans Cybern* 2016;46(10). <https://doi.org/10.1109/TCYB.2015.2467167>.
- [14] Wang J, Lv M, Li Z, Zeng B. Multivariate selection-combination short-term wind speed forecasting system based on convolution-recurrent network and multi-objective chameleon swarm algorithm. *Expert Syst Appl* 2023;214. <https://doi.org/10.1016/j.eswa.2022.119129>.
- [15] Cai C, Li Y, Su Z, Zhu T, He Y. Short-term electrical load forecasting based on VMD and GRU-TCN hybrid network. *Appl Sci* 2022;12(13). <https://doi.org/10.3390/app12136647>.
- [16] Fang T, Zheng C, Wang D. Forecasting the crude oil prices with an EMD-ISBM-FNN model. *Energy* 2023;263. <https://doi.org/10.1016/j.energy.2022.125407>.
- [17] Lin Y, Yan Y, Xu J, Liao Y, Ma F. Forecasting stock index price using the CEEMDAN-LSTM model. *N Am J Econ Finance* 2021;57. <https://doi.org/10.1016/j.najef.2021.101421>.
- [18] Yan K, Li W, Ji Z, Qi M, Du Y. A hybrid LSTM neural network for energy consumption forecasting of individual households. *IEEE Access* 2019;7. <https://doi.org/10.1109/ACCESS.2019.2949065>.
- [19] Liu Z, Jiang P, Zhang L, Niu X. A combined forecasting model for time series: application to short-term wind speed forecasting. *Appl Energy* 2020;259. <https://doi.org/10.1016/j.apenergy.2019.114137>.
- [20] Zhang S, Wang C, Liao P, Xiao L, Fu T. Wind speed forecasting based on model selection, fuzzy cluster, and multi-objective algorithm and wind energy simulation by Betz’s theory. *Expert Syst Appl* 2022;193. <https://doi.org/10.1016/j.eswa.2022.116509>.
- [21] Wang H, Han S, Liu Y, Yan J, Li L. Sequence transfer correction algorithm for numerical weather prediction wind speed and its application in a wind power forecasting system. *Appl Energy* 2019;237. <https://doi.org/10.1016/j.apenergy.2018.12.076>.
- [22] Miao C, Wang X, Li H, Han L, Wen C. Day-ahead prediction of wind power based on NWP wind speed error correction. *Dianwang Jishu/Power System Technology* 2022;46(9). <https://doi.org/10.13335/j.1000-3673.pst.2022.0834>.
- [23] Al-Yahayi S, Charabi Y, Al-Badi A, Gastli A. Nested ensemble NWP approach for wind energy assessment. *Renew Energy* 2012;37(1). <https://doi.org/10.1016/j.renene.2011.06.014>.
- [24] Liu X, Lin Z, Feng Z. Short-term offshore wind speed forecast by seasonal ARIMA - a comparison against GRU and LSTM. *Energy* 2021;227. <https://doi.org/10.1016/j.energy.2021.120492>.
- [25] Husain NH, Yusof F, Jamaludin AR, Norrulashikin SM. Forecasting wind speed in peninsular Malaysia: an application of Arima and Arima-Garch models. *Pertanika Journal of Science and Technology* 2021;29(1). <https://doi.org/10.4783/pjst.29.1.02>.
- [26] Tena García JL, Cadenas Calderón E, González Ávalos G, Rangel Heras E, Mbikayi Tshikala A. Forecast of daily output energy of wind turbine using sARIMA and nonlinear autoregressive models. *Adv Mech Eng* 2019;11(2). <https://doi.org/10.1177/1687814018813464>.
- [27] Qian W, Wang J. An improved seasonal GM(1,1) model based on the HP filter for forecasting wind power generation in China. *Energy* 2020;209. <https://doi.org/10.1016/j.energy.2020.118499>.
- [28] Wu C, Wang J, Chen X, Du P, Yang W. A novel hybrid system based on multi-objective optimization for wind speed forecasting. *Renew Energy* 2020;146. <https://doi.org/10.1016/j.renene.2019.04.157>.
- [29] Wei D, Wang J, Li Z, Wang R. Wind power Curve modeling with hybrid copula and grey wolf optimization. *IEEE Trans Sustain Energy* 2022;13(1). <https://doi.org/10.1109/TSTE.2021.3109044>.
- [30] Gao Y, Wang J, Yang H. A multi-component hybrid system based on predictability recognition and modified multi-objective optimization for ultra-short-term onshore wind speed forecasting. *Renew Energy* 2022;188. <https://doi.org/10.1016/j.renene.2022.02.005>.
- [31] Guo H, Wang J, Li Z, Jin Y. A multivariable hybrid prediction system of wind power based on outlier test and innovative multi-objective optimization. *Energy* 2022;239. <https://doi.org/10.1016/j.energy.2021.122333>.
- [32] Xu D, Shao H, Deng X, Wang X. The hidden-layers topology analysis of deep learning models in survey for forecasting and generation of the wind power and photovoltaic energy. *CMES - Computer Modeling in Engineering and Sciences* 2022;131(2). <https://doi.org/10.32604/cmes.2022.019245>.
- [33] Salazar AA, Che Y, Zheng J, Xiao F. Multivariable neural network to postprocess short-term, hub-height wind forecasts. *Energy Sci Eng* 2022;10(7). <https://doi.org/10.1002/ese.3928>.
- [34] Dragomireski K, Zosso D. Variational mode decomposition. *IEEE Trans Signal Process* 2014;62(3). <https://doi.org/10.1109/TSP.2013.2288675>.
- [35] Niu X, Wang J, Zhang L. Carbon price forecasting system based on error correction and divide-conquer strategies. *Appl Soft Comput* 2022;118. <https://doi.org/10.1016/j.asoc.2021.107935>.
- [36] Gong M, Jiao L, Du H, Bu L. Multiobjective immune algorithm with nondominated neighbor-based selection. *Evol Comput* 2008;16(2). <https://doi.org/10.1162/evo.2008.16.2.225>.
- [39] Yang M, Guo Y, Huang Y. Wind power ultra-short-term prediction method based on NWP wind speed correction and double clustering division of transitional weather process. *Energy* 2023;282:128947. ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2023.128947>.
- [40] Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput* 1997;9(8):1735–80.