IJE TRANSACTIONS B: Applications Vol. 39, No. 02, (February 2026) 309-320

International Journal of Engineering

Journal Homepage: www.ije.ir

A Novel Approach for Accurate Wind Speed Time Series Forecasting Using ICEEMDAN Decomposition and Sample Entropy through Integration of Deep Learning Models

H. Mezaache*[a], H. Bouzgou[b], C. Raymond[c], N. Zemouri[a]

[a] *Department of Electronics, Faculty of Technology, University of M'sila, Lab. G.E. University Pole, Algeria*
[b] *Department of Industrial Engineering, Faculty of Technology, Laboratory of Automation and Manufacturing, University of Batna 2 (Mostefa Ben Boulaïd), Batna, Algeria*
[c] *INSA Rennes, INRIA/ IRISA Beaulieu Campus 35042 Rennes, France*

PAPER INFO

Paper history:
Received 26 January 2025
Received in revised form 23 March 2025
Accepted 29 March 2025

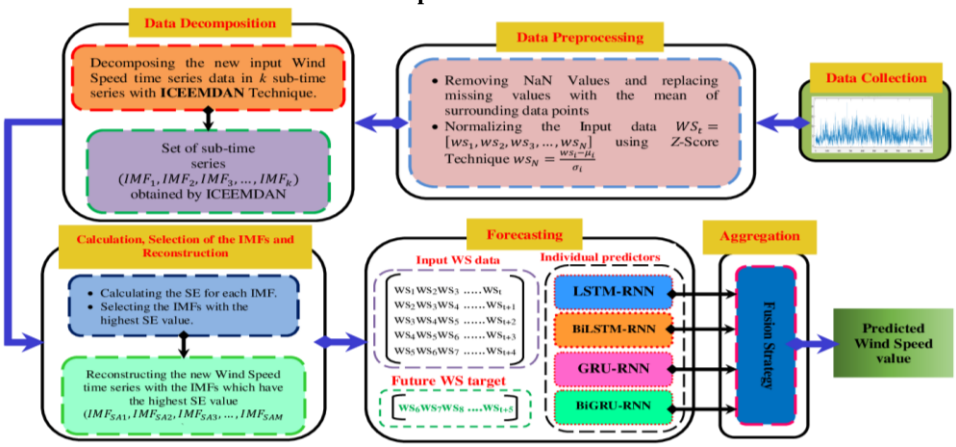
Keywords:
Wind Speed
Times Series
Deep Neural Networks
Sample Entropy
Decomposition
Fusion

ABSTRACT

This study proposes a novel hybrid model for wind speed forecasting (WSF) based on a three-stage framework comprising decomposition, feature selection, and forecasting. The proposed approach employs Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) to decompose wind speed time series into Intrinsic Mode Functions (IMFs). A distinctive contribution of this study is the use of sample entropy as a feature selection mechanism to identify the most relevant Intrinsic Mode Functions (IMFs). The selected IMFs are then integrated through a classification-based fusion technique, significantly enhancing forecasting accuracy and distinguishing this approach from conventional methods. Two distinct forecasting approaches are evaluated using multiple performance metrics, including RMSE, MAE, MAPE, and R^2. The first approach applies the fusion technique directly to the original wind speed time series, while the second incorporates ICEEMDAN to decompose the time series. Experimental validation using two real-world datasets from Algeria demonstrates the superiority of the proposed hybrid model over individual forecasting models, yielding significant improvements in prediction accuracy, robustness, and stability. These findings underscore the effectiveness of the three-stage framework, offering a reliable and efficient solution for short-term wind speed forecasting, with potential applications in renewable energy management and grid optimization.

***doi**: 10.5829/ije.2026.39.02b.03*

Graphical Abstract



*Corresponding Author Email: hatem.mezaache@univ-msila.dz (H. Mezaache)

Please cite this article as: Mezaache H, Bouzgou H, Raymond C, N. Zemouri. A Novel Approach for Accurate Wind Speed Time Series Forecasting Using ICEEMDAN Decomposition and Sample Entropy through Integration of Deep Learning Models. International Journal of Engineering, Transactions B: Applications. 2026;39(02):309-20.

NOMENCLATURE

BiLSTM-RNN	Bidirectional Long Short-Term Memory Recurrent Neural Networks	IMF	Intrinsic Mode Functions
BiGRU-NN	Bidirectional Gated Recurrent Unit Neural Networks	ICEEMDAN	Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise
EMD	Empirical Mode Decomposition	MAPE	Mean Absolute Percentage Error
EEMD	Ensemble Empirical Mode Decomposition	MABE	Mean Absolute Bias Error
GRU-NN	Gated Recurrent Unit Neural Networks	RMSE	Root Mean Square Error
SE	Sample Entropy	R^2	Coefficient of Determination
LSTM-RNN	Long Short-Term Memory Recurrent Neural Networks	WSF	Wind Speed Forecasting

1. INTRODUCTION

Renewable energies play a pivotal role in transforming modern lifestyles. By decreasing dependence on fossil fuels, they contribute considerably to mitigating climate change through the reduction of greenhouse gas emissions. The integration of clean energy sources such as solar, wind, and hydropower fosters the diversification of energy supplies, thus enhancing the energy security of nations. Economically, renewable energy systems drive job creation within advanced and emerging sectors, supporting sustainable economic growth. Moreover, they enable households and businesses to lower electricity costs through self-consumption and on-site clean energy generation. Access to sustainable energy also enhances living standards, principally in remote and underserved areas where conventional electrification is either expensive or logistically challenging. In summary, renewable energies are key enablers of a cleaner, more secure, and socially equitable future (1).

Wind speed is essential factor in harnessing wind energy, as it directly impacts the effectiveness and power output of wind turbines. Higher wind speeds generate more kinetic energy, which turbines convert into electricity. Forecasting wind power is important for improving energy production, grid stability, and operational organization. Precise predictions help also to integrate wind energy into power grids, reducing dependence on fossil fuels and enhancing energy security. It also enables efficient scheduling of maintenance and resource allocation, ultimately lowering operational costs and supporting the shift to a more sustainable energy system (2).

Several studies in the literature have addressed the problem of wind speed forecasting (3-6), an EMD-MGM-based network architecture was introduced by Xiong et al. (7) as the core framework. The method extends three-point regression loss functions to enable wind speed interval regression, ensuring more robust predictive capabilities. Additionally, a novel quantile regression mechanism, termed SLF, is proposed to enhance the precision of probabilistic forecasts. To further improve forecasting accuracy, a new technique called NS is developed, allowing for the superposition of subsequence prediction results, thereby capturing more comprehensive temporal patterns in wind speed data. Ran et al. (8) introduced a novel prediction model for long-term forecasting of wind power generation in China. Based on the forecasting results, the study proposes targeted policy recommendations aimed at supporting the sustainable growth and development of China's wind power industry. The study by Parri et al. (9) proposed a hybrid approach for wind speed forecasting (WSF) based on Variational Mode Decomposition (VMD), CoST, and Support Vector Regression (SVR) models. The VMD algorithm is employed to denoise the wind speed data, while the CoST-SVR framework predicts wind speed from the denoised signals. The approach is evaluated using two sets of WSF models, demonstrating superior performance for short- to medium-term forecasts at 5, 10, 15, 30 minutes, as well as 1-hour and 2-hour ahead prediction intervals. In another work by Parri et al. (10), a hybrid approach, named VMD-Ts2Vec-SVR, is proposed for accurate wind speed forecasting (WSF). The method employs the VMD algorithm to decompose wind speed data and introduces contextual feature representation for the first time in WSF. The Ts2Vec-SVR model is then used to predict wind speed with high accuracy. Experimental results demonstrate the superior performance of the proposed approach in comparison to other methods. Li et al. (11) proposed an approach to improve local feature extraction by incorporating causal convolution, which enhances the model's ability to capture temporal patterns. The Empirical Mode Decomposition (EMD) algorithm generates features across different time scales, enriching the input data. A Transformer model that integrates EMD with causal convolution demonstrates significant improvements in forecasting accuracy (12), addressed the nonlinearity problem of the original sequence by employing multiple decomposition methods. Two decomposition frameworks are proposed to uncover the underlying relationships within the sequence. A novel hybrid model, CNN-ECA-BIGRUa (with attention mechanism), is used for the prediction phase. Additionally, the forecasting is further enhanced using gated Temporal Convolutional Network (TCN) combinatorial sub-models, resulting in improved prediction performance. In the studies discussed above, the methodologies primarily relied on a combination of decomposition techniques and forecasting models. However, the present work introduces an additional critical component: feature

selection. This innovative block is designed to identify and retain the most relevant features, thereby optimizing the forecasting performance. By incorporating feature selection into the framework, the proposed approach aims to enhance the accuracy and reliability of wind speed predictions, setting it apart from existing methods in the literature.

The remainder of this paper is structured as follows: In Section 2, we present the proposed methodology, providing an in-depth explanation of the various algorithms introduced for the study. This section elaborates on the approach's design, the rationale behind each algorithm, and how they contribute to the overall framework. Section 3 is dedicated to the experimental setup, detailing the data used for testing the proposed models, as well as presenting and analyzing the results obtained from the different experiments conducted. Finally, Section 4 concludes the paper by summarizing the key findings and discussing the implications of the results.

2. PROPOSED METHODOLOGY

The proposed methodology for wind speed forecasting is a comprehensive hybrid approach that begins with the collection of wind speed time series data, ensuring a reliable and accurate dataset for analysis. Next, the data undergoes preprocessing, including the removal of missing values and normalization using the *Z*-score method to ensure consistency and comparability. The normalized data is then decomposed using the Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN), resulting in several Intrinsic Mode Functions (IMFs), each representing specific frequency components. To isolate the most relevant components, Sample Entropy (SE) is calculated for each IMF, and only those with the highest entropy values, indicating their importance, are selected. These selected IMFs are then reconstructed to form a refined dataset, capturing the essential dynamics of the wind speed time series. This refined dataset is then used as input for several deep learning models, including LSTM-NN, BiLSTM-NN, GRU-NN, and BiGRU-NN, designed to analyze complex temporal patterns and dependencies in sequential data. The outputs of these models are integrated using a classification-based aggregation technique, which combines the strengths of each model to produce an optimized forecast. The complete workflow of this methodology is illustrated in the graphical abstract, providing a clear overview of the system architecture and its successive steps.

2. 1. Z-SCORE The *Z*-score normalization technique transforms the data into a standardized format by centering it around a mean of zero and scaling it to a unit standard deviation. It is a widely used data preprocessing technique, particularly in time-series forecasting tasks. For each original data value x_i, the *Z*- score normalization z_i formula is given by:

$$z_i = \frac{x_i - \beta}{\alpha} \quad (1)$$

where: β represents the mean and α denotes the standard deviation of the dataset. This process ensures that the dataset becomes dimensionless, enabling fair comparison and eliminating biases caused by differences in scale among variables (13, 14).

2. 2. Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (ICEEMDAN) The ICEEMDAN is an algorithm that has been developed by Colominas et al. (15). It is an enhanced version of the ensemble empirical mode decomposition method (EEMD). This method effectively addresses the primary obstacles encountered in EEMD method. This decomposition technique makes it possible to decode each time series signal into a finite number of components called IMFs (16). The algorithm is designed to handle adaptive noise and also to correct frequency aliasing faults. The ICEEMDAN method enhances the quality of IMFs by incorporating white noise mode into the original signal, effectively diminishing the remaining noise (17). The detailed steps of ICEEMDAN decomposition technique are as follows:

Step 1:

To obtain specific noise, add Empirical Mode Decomposition (EMD) to the original signal as shown in the following equation:

$$y^i = y + \beta_0 E_1[\omega^i] \quad (2)$$

This equation represents the process of adding special noise $E_1[\omega^i]$ to the original signal y. The variable y^i represents the i^{th} added white noise, with i ranging from *1* to *N*. The expected signal-to-noise ratio of the first decomposed signal is represented by β_0.

Step 2:

To compute the IMF value, the EMD algorithm is utilized to acquire the local mean value of the reconstructed signal, and the first residual is obtained by taking their mean value.

$$R_1 = \frac{1}{N}\sum_{i=1}^{N} M[y^i] \quad (3)$$

$$\tilde{r}_1 = y - R_1 \quad (4)$$

This equation utilizes the residual of the first decomposition denoted as R_1, along with the operator $M[\blacksquare]$ that calculates the local mean, and $\tilde{r}_1$ represents the value of the first IMF.

Step3:

Calculate the second IMF value using the following equation:

$$R_2 = \frac{1}{N}\sum_{i=1}^{N} M\left[R_1 + \beta_1 E_2[\omega^i]\right] \quad (5)$$

$$\tilde{r}_2 = R_1 - R_2 \quad (6)$$

Where R_2 and $\tilde{r}_2$ represent the residual and the value of the second IMF, respectively.

Step4:

Further apply the above computational steps to determine the k^{th}IMF value:

$$R_k = \frac{1}{N}\sum_{i=1}^{N} M\left[R_{k-1} + \beta_{k-1} E_k[\omega^i]\right] \quad (7)$$

$$\tilde{r}_k = R_{k-1} - R_k \quad (8)$$

where $k = 1,2,3,\dots,N$. The iterative calculation of this equation can accurately decompose the IMF of the original signal.

2. 3. Sample Entropy (SE) Sample Entropy (SE) is a widely used method for assessing the complexity and regularity of time series data, introduced by Richman et al. (18) as an improvement over Approximate Entropy (ApEn). Unlike Approximate Entropy, which can produce inconsistent results, especially with short or noisy data, SE offers a more reliable and precise measure of the unpredictability of a time series. The core concept of SE is to evaluate the similarity of sequences within the time series, with lower values of SE indicating higher similarity (or regularity) and higher values indicating more complexity or unpredictability. The calculation of SE involves three primary parameters: the length of the time series, the tolerance for similarity (often defined as a fraction of the standard deviation of the series), and the embedding dimension, typically set to 2. The process of calculating sample entropy (SE) is based on the following steps (19, 20):

Step 1: The original data consisting of N points is represented as:

$$X_i = [x_i, x_{i+1}, \dots, x_{i+m-1}] \quad (9)$$

Step 2: The dimension between the vector X_i and X_j is defined as:

$$D_m(X_i, X_j) = \max_{0\sim m-1} |x_{i+1} - x_{j+k}| \quad (10)$$

Step 3: The mean value of the vectors is then calculated as:

$$B^m(r) = \frac{1}{N-M+1}\sum_{i=1}^{N-M+1} B_i^m \quad (11)$$

Step 4: Steps 1 through 3 are repeated to compute the final value of the sample entropy (SE), which is expressed as follows:

$$SE(m,r,N) = \lim_{N\to\infty}\left[-\ln\frac{B^{m+1}(r)}{B^m(r)}\right] \quad (12)$$

Where m indicates the embedding dimension, and r represents the conditional threshold.

2. 4. Individual Deep Learning Forecasting Models

2. 4. 1. Long Short-Term Memory Recurrent Neural Networks (LSTM-RNN) Long Short-Term Memory (LSTM) network is a special kind of recurrent networks (RNN), for the first time, LSTM-RNN is proposed by Hochreiter and Schmidhuber (21) to learn long-term dependence information. Next, Gers et al. (22) improved the LSTM-RNN by introducing the technique of the forget gate, which is relevant to the prediction of continuity (23). The LSTM-RNN has been very successful and widely used in different fields such as image processing, image and text recognition, video data recognition. Currently, LSTM-RNN is widely used with success in time series forecasting problems. The classical neural networks contain neurons, while LSTM recurrent networks are composed of blocks of memory interconnected through successive layers. To control the flow of historical information in cell LSTM-RNN, its architecture contain three types of gates which are the Forget Gate, the Input Gate, and the Output Gate (24, 25), using these gates an LSTM-RNN processes, stores and transfer the pertinent information for a long period and reduce the rate of data loss (26). Figure 1 shows the internal architectures of an LSTM-RNN memory cell.

The following expressions can represent the internal calculation process for LSTM-RNN cell:

$$f_t = \delta(X_f \times [h_{t-1}, I_t] + \beta_f) \quad (13)$$

where f_t is the forget gate; δ represent the sigmoid activation function; X_f is the weight matrix between forget gate and input gate; h_{t-1} represents the previous hidden state; I_t is the input value of the LSTM-RNN at the current time and β_f represents the bias term of forgetting gate.

$$i_t = \delta(X_i \times [h_{t-1}, I_t] + \beta_i) \quad (14)$$

$$\widetilde{C_t} = \mathcal{G}(X_C \times [h_{t-1}, I_t] + \beta_C) \quad (15)$$

$$C_t = f_t \times C_{t-1} + i_t \times \widetilde{C_t} \quad (16)$$

where: i_t is the input gate; $\widetilde{C_t}$ is the candidate cell state; C_t represent update cell state; C_{t-1} is the cell state from previous time step; $\mathcal{G}$ represent hyperbolic tangent (tangh) activation function; X_i, β_i represents weight

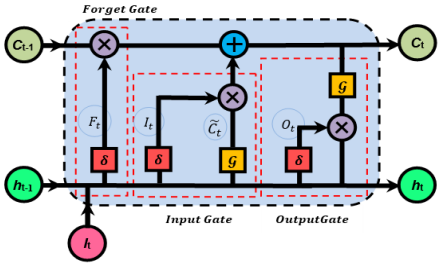


Figure 1. Basic Structure of LSTM-RNN model

matrix and bias term of input gate respectively; X_C, β_C represents the weight matrix and bias term for generating candidate values respectively.

$$h_t = o_t \times \mathcal{G}(C_t) \quad (17)$$

$$o_t = \delta(X_o \times [h_{t-1}, I_t] + \beta_o) \quad (18)$$

where h_t is hidden state for the current time step; o_t represent the output gate which controls how much information from C_t will be included in the h_t; X_o, β_o are the weights matrix and bias term of the output gate.

The sigmoid and hyperbolic tangent (tangh) are a nonlinear activation functions defined by:

$$\delta(\mathcal{Z}) = \frac{1}{1-e^{-\mathcal{Z}}} \quad (19)$$

$$\mathcal{G}(\mathcal{Z}) = tanh(\mathcal{Z}) = \frac{e^{\mathcal{Z}}-e^{-\mathcal{Z}}}{e^{\mathcal{Z}}+e^{-\mathcal{Z}}} \quad (20)$$

It converts the output value to a value between - 1 and 1.

2. 4. 2. Bidirectional Long Short-Term Memory Recurrent Neural Networks (BiLSTM-RNN) Bidirectional Long Short-Term Memory (BiLSTM) is a recurrent neural network (RNN), introduced by Schuster and Paliwal (27), it is an advanced variant of the LSTM model developed to enhance sequence modeling by capturing bidirectional contextual dependencies. Unlike conventional LSTMs, which process data in a single forward direction, BiLSTM networks leverage two LSTM layers, one operating in the chronological order (past to future) and the other in reverse order (future to past) (28). This dual-layer structure allows BiLSTM to capture comprehensive temporal patterns, making it particularly effective in applications where future information influences present predictions, such as in wind speed forecasting and complex time-series analyses.

The representation of the BiLSTM-RNN model is shown in Figure 2, where both forward and backward LSTM layers process the input sequence in parallel, combining their outputs to form the final hidden state at each time step.

At each time step t, the hidden states for the forward and backward layers are computed independently. The forward hidden state $h_t^{\rightarrow}$ and the backward hidden state $h_t^{\leftarrow}$ are calculated using the following equations:

$$h_t^{\rightarrow} = \delta(W^{\rightarrow}.[h_{t-1}^{\rightarrow}, x_t] + b^{\rightarrow}) \quad (21)$$

$$h_t^{\leftarrow} = \delta(W^{\leftarrow}.[h_{t+1}^{\leftarrow}, x_t] + b^{\leftarrow}) \quad (22)$$

where: δ denotes the sigmoid activation function, x_t is the input vector at time step t, $h_{t-1}^{\rightarrow}$ and $h_{t+1}^{\leftarrow}$ represent the hidden states of the previous and next time steps for the forward and backward LSTM layers, respectively, and $W^{\rightarrow}$, $W^{\leftarrow}$ are the weight matrices for the forward and backward directions. The vectors $h_t^{\rightarrow}$ and $h_t^{\leftarrow}$ are combined, typically by concatenation, to form a bidirectional representation of the sequence at each time step:

$$h_t = [h_t^{\rightarrow} \quad h_t^{\leftarrow}] \quad (23)$$

This concatenation ensures that the network benefits from both the past and future context, improving its performance. By capturing long-range dependencies in both directions, BiLSTM-RNN is particularly effective in situations where the full context of the sequence is necessary for making accurate forecasting.

2. 4. 3. Gated Recurrent Unit Neural Networks (GRU-NN) The Gated Recurrent Units (GRU) are a variant of Recurrent Neural Networks (RNN), address the vanishing and exploding gradient problems, while simplifying the complexity of Long Short-Term Memory (LSTM) networks. GRUs are particularly effective for modeling sequential data and have demonstrated their efficiency in various applications, such as wind speed forecasting and speech recognition (29, 30). The GRU's architecture is built around two main gates: the update gate, which controls how much information from the past is retained, and the reset gate, which determines what portions of the previous state should be forgotten (31). Together, these gates dynamically regulate the flow of information across time steps, ensuring adaptive memory retention. Figure 3 presents a schematic diagram of the GRU architecture, illustrating the flow of information through the update and reset gates.

The update gate z_t is defined as:

$$z_t = \delta(W_z.[h_{t-1}, x_t] + b_z) \quad (24)$$

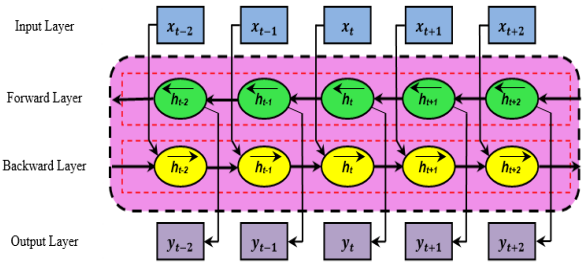


Figure 2. Architecture of BiLSTM-RNN model

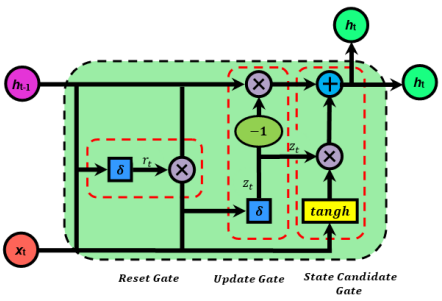


Figure 3. Architecture of GRU-NN model

where: δ is the sigmoid activation function, W_Z is the weight matrix, h_{t-1} is the previous hidden state, x_t is the current input, and b_z is the bias term.
The reset gate r_t is computed as follows:

$$r_t = \delta(W_r.[h_{t-1}, x_t] + b_r) \quad (25)$$

where W_r and b_r are the corresponding weight matrix and bias term for the reset gate.

Together, these gates dynamically regulate the flow of information across time steps, ensuring adaptive memory retention. Using the reset gate, the GRU calculates a candidate hidden state $\tilde{h}_t$ as follows:

$$\tilde{h}_t = tanh(W.[r_t \odot h_{t-1}, x_t] + b) \quad (26)$$

where: $tanh$ is the hyperbolic tangent activation function, $\odot$ represents element-wise multiplication, and W and b are the weight matrix and bias term.

Finally, the GRU computes the updated hidden state h_t as a combination of the previous hidden state h_{t-1} and the candidate hidden state $\tilde{h}_t$, weighted by the update gate:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (27)$$

This formulation allows the GRU to integrate the past hidden state and current input, enabling a balance between retaining historical information and incorporating new data (32).

2. 4. 4. Bidirectional Gated Recurrent Unit Neural Networks (BiGRU-NN) The BiGRU model (Bidirectional Gated Recurrent Unit) represents an enhancement of traditional recurrent networks, designed to capture temporal dependencies in both past and future directions. Unlike a standard GRU, which processes sequences in a single direction, the BiGRU combines two GRU layers: one processes the data sequentially, while the other processes it in reverse. This provides better contextualization for complex temporal sequences. This structure is particularly effective in applications such as wind energy forecasting, where bidirectional temporal correlations play a crucial role (33, 34). The BiGRU unit can be described by the following equations:

$$h_t^f = GRU(x_t, h_{t-1}^f) \quad (28)$$

$$h_t^b = GRU(x_t, h_{t+1}^b) \quad (29)$$

where h_t^f and h_t^b represent the hidden states generated in the forward and backward directions, respectively, and x_t is the input at time t. The final output is obtained by combining these two states.

$$h_t = [h_t^f, h_t^b] \quad (30)$$

Figure 4 illustrates the architecture of the BiGRU-NN, highlighting its two parallel processing flows (forward and backward) and their fusion to produce a globally enriched temporal context output.

2. 5. Classification Based Technique The classification-based strategy is founded on the premise that predictive models exhibit varying levels of performance across different regions within the input variable space of time series data. To leverage this variability, the strategy partitions the input space into distinct regions, each of which is assigned to the predictive model that achieves the lowest forecasting error within that region. During the training phase, the optimal model for each segment is determined through a comprehensive analysis of the training set. This segmentation process ensures that each region is paired with the most accurate predictor, thereby enhancing overall forecasting precision (35-37).

The optimal predictor $\hat{P}(Z)$ for a given series S is selected using:

$$\hat{P}(Z) = \underset{i=1,2,\ldots..N}{\arg\min} |f_i(S_j) - y_j| \quad (31)$$

where: $f_i(S_j)$ represents the prediction from model i for series S_j and y_j is the actual value.
Once the segmentation is established, any new time series data is classified into a region, and the associated optimal model is applied for forecasting. The output forecast $\hat{F}(Z)$ is expressed as:

$$\hat{F}(Z) = f_{\hat{P}(Z)}(Z) \quad (32)$$

This strategy's effectiveness lies in its ability to dynamically leverage the strengths of different models for varying data characteristics, thereby improving the overall forecasting accuracy. As illustrated in Figure 5, this method involves two main phases: training, where the variable space is partitioned, and testing, where the optimal model for each incoming series is selected.

3. CASE STUDY

3. 1. Description of Data Used To evaluate the accuracy and effectiveness of the proposed hybrid approach, simulation experiments were conducted using historical wind speed time series data from two sites in southern Algeria. The first site represents a recorded data

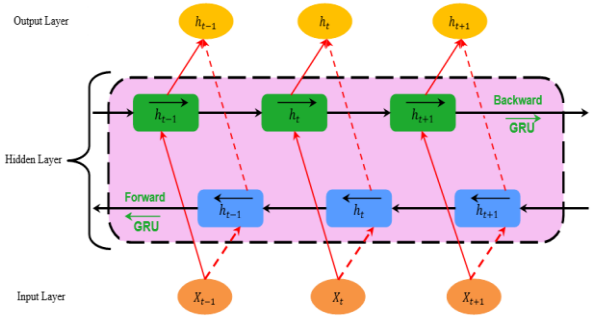


Figure 4. Architecture of BiGRU-NN model

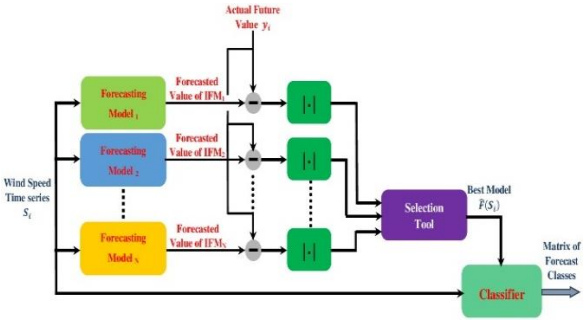


Figure 5. Flowchart of the classification-based technique

every 30 minutes over a one-year period from January 1st to December 31, 2016, and the second site recorded data every 15 minutes over the same period from January 1st to December 31, 2019. These sites were chosen for their representation of distinct wind regimes. Table 1 describes the geographical characteristics of the two sites, including latitude, longitude, altitude, and the data size associated with each site. Fifty values (50) of the time series are used as input data, while the 51th value is utilized as the target (future value).

Table 2 summarizes the statistical characteristics of the data, including the mean, standard deviation, and extremes of wind speed.

The data employed in this study to evaluate the various methodologies consist of the wind speed time series recorded at two locations in Algeria, which represent the two sub-climates, as illustrated in Figures 6 and 7 presents the normalized samples for each site, emphasizing the normalization process.

3. 2. Performance Metrics In this study, four statistical metrics are employed to evaluate the proposed approach:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (33)$$

$$MABE = \frac{1}{N}\sum_{i=1}^{N}|\hat{y}_i - y_i| \quad (34)$$

TABLE 1. Geographical characteristics of the two test sites

Dataset	Geographical position	Altitude (m)	Latitude (°)	Longitude (°)	Sample size
Site# 1	South-west	263	27.88	-0.28	10172
Site# 2	South-east	450	36.667	5.667	34943

TABLE 2. Statistical characteristics of the two test sites

Dataset	Resolution	Statistical values (m/s)				
		Mean	Std	Max	Med	Min
Site # 1	30 min	5.39	2.89	21.80	4.9	0
Site # 2	15 min	4.69	3.45	23.30	4.0	0

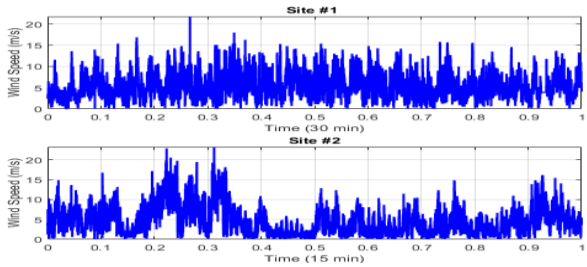


Figure 6. A small window of original time series of wind speed from the two test sites

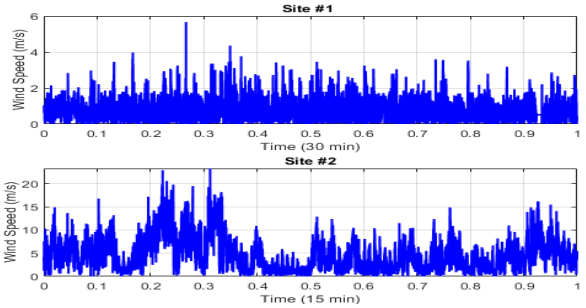


Figure 7. A small window of wind speed time series from the two test sites after normalization

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right| * 100 \quad (35)$$

$$R^2 = 1 - \frac{var(y_i - \hat{y}_i)}{var(y_i)} \quad (36)$$

where $\hat{y}_i$ and y_i represent the forecasted and actual wind speed values, respectively, $var(y_i)$ denotes the variance of the measured values, and N is the total number of time series.

3. 3. Forecasting Experiments The original wind speed data from the two test sites, as outlined in the previous section, are employed to validate and showcase the effectiveness of the proposed approach. To achieve this, two distinct sets of hybrid models are developed to accurately forecast short-term wind speed for each location.

3. 4. Results of the Different Models and Discussion This study uses two approaches to forecast wind speed by joining the advantages of the decomposition method (ICEEMDAN), Deep Learning Neural Networks (LSTM, BiLSTM, GRU, BiGRU) and combination technique (Classification). The first approach uses only Deep Learning Neural Networks as individual models, applied to original wind speed time series and the classification combination to aggregate the different outcomes issued from the DNNs. The second approach incorporates the ICEEMDAN decomposition method to decompose the original wind speed signals and then employs the Sample Entropy (SE) method to select

the best IMFs. The best models are obtained through fine-tuning and cross-validation, with the final prediction derived by aggregating the results from all the deep neural network (DNN) models. The statistical performances of the two approaches for both locations are listed in Table 3, where the best results are marked in bold.

3. 4. 1. Results Obtained with First Approach

This study evaluates the performance of individual models (LSTM, BiLSTM, GRU, BiGRU) and their combination through classification. The results as shown in Table 3. The BiGRU models stand out due to their lightweight and efficient architecture, offering better accuracy and reduced complexity, surpassing the BiLSTM models. While BiLSTM effectively captures complex long-term relationships, BiGRU proves more suitable for sequential predictions thanks to its faster learning capability. Figures 8 and 9 illustrate the predictions of both individual and combined models, demonstrating that the fusion significantly enhances accuracy and reduces errors, thereby confirming the superiority of combined approaches.

3. 4. 2. Results Obtained with Second Approach

In this approach, a new processing phase based on the ICEEMDAN technique is inserted to enhance the accuracy of wind speed forecasts. This time-space analysis method decomposes non-stationary and non-linear time series into Intrinsic Mode Functions (IMFs) while optimizing the stability and accuracy of the decomposition. The addition of adaptive noise mitigates the effects of over-decomposition and improves the reliability of the extracted components, providing a better representation of complex signals, particularly natural signals. The components extracted from the two test sites are illustrated in Figures 10 and 11.

The optimal Intrinsic Mode Functions (IMFs) are selected using the Sample Entropy (SA) method, as presented in Table 4. Following selection, the signal is

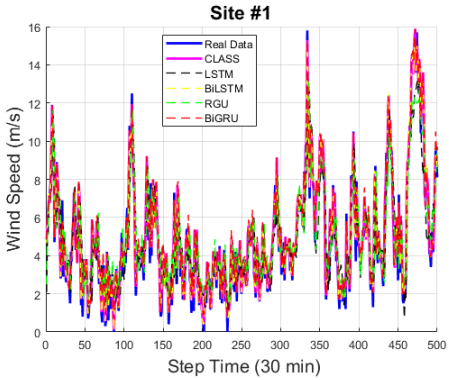


Figure 8. Measured and forecasted non-decomposed wind speed series of Site #1 using deep learning models with combination

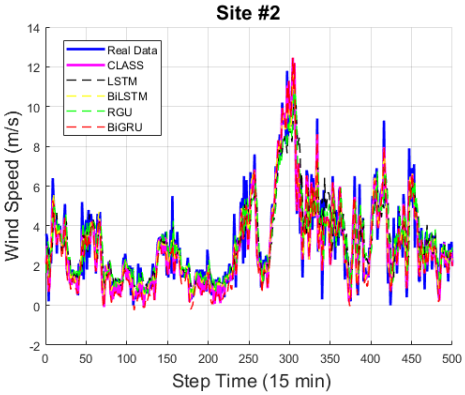


Figure 9. Measured and forecasted non-decomposed wind speed series of Site #2 using deep learning models with combination

TABLE 3. Details of the model performance evaluation metrics without decomposition for the two test sites. The best results are in bold

Test sites/ Methods	Without Decomposition				
	Individual Deep Learning Model				Combination
	LSTM -NN	BiLST M-NN	GRU -NN	BiGRU -NN	Class
Site# 1					
RMSE	1.45	0.39	1.40	0.43	**0.30**
MAPE	21.77	5.93	20.98	6.92	**4.29**
MABE	0.05	0.11	0.09	0.27	0.21
R^2	0.69	0.97	0.71	0.98	**0.98**
Site# 2					
RMSE	1.50	0.83	1.48	0.64	**0.53**
MAPE	26.25	14.69	26.05	12.08	**8.72**
MABE	0.009	0.12	0.031	0.175	0.36
R^2	0.71	0.91	0.71	0.95	**0.96**

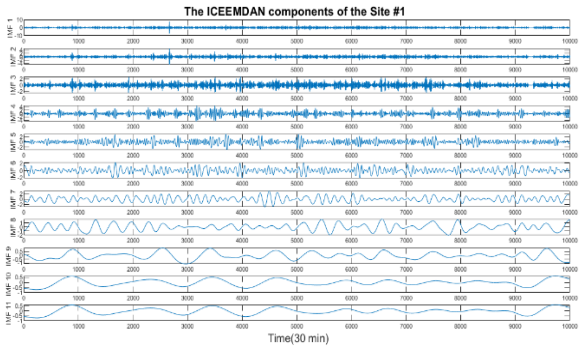


Figure 10. The waveforms of ICEEMDAN components of wind speed time series of Site# 1

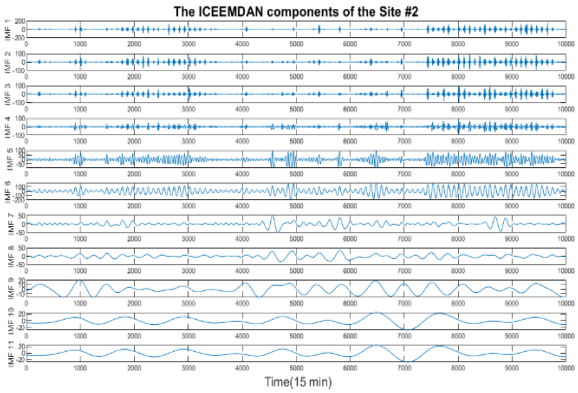


Figure 11. The waveforms of ICEEMDAN components of wind speed time series of Site# 2

reconstructed and subjected to predictions by individual predictors. To further improve the performance of the individual Deep Neural Network (DNN) models, an additive block incorporating a "Classification" combination strategy is introduced, which aggregates the outcomes of the different predictors.

The performances of the hybrid models, including ICEEMDAN-LSTM, ICEEMDAN-BiLSTM, ICEEMDAN-GRU, and ICEEMDAN-BiGRU, as well as their combination using the "Class" strategy, are illustrated in Figures 12 and 13 for both test sites. The statistical results presented in Table 5 show that the combined hybrid model (ICEEMDAN-DNN-Class) outperforms the individual models across all error metrics.

TABLE 4. SE of different IMFs for the two sites

	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8	IMF9	IMF10	IMF11
Site 1	1.63	1.73	0.95	0.61	0.25	0.11	0.05	0.02	0.01	0.007	0.001
Site 2	1.50	1.64	1.07	0.53	0.24	0.11	0.04	0.01	0.009	0.004	0.001

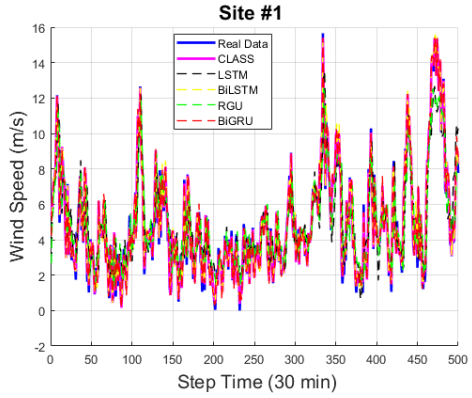


Figure 12. Measured and forecasted decomposed wind speed series of Site #1 using deep learning models with class-based aggregation

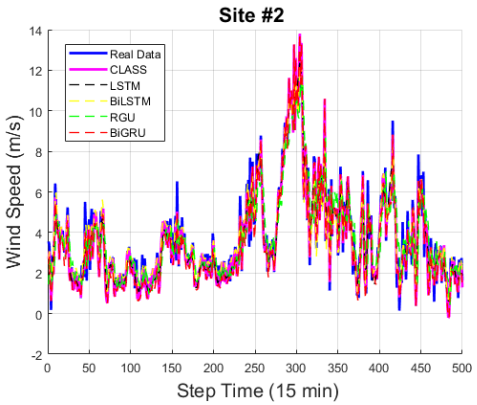


Figure 13. Measured and forecasted decomposed wind speed series of Site #2 using deep learning models with class-based aggregation

The results depicted in Figures 14 and 15, illustrated through histograms of the RMSE metric for test sites #1 and #2, clearly indicate that the "Class" combination method is the most effective strategy for reducing prediction errors. For site #1, this method achieves an RMSE of 0.30 without decomposition, which is further reduced to 0.21 when paired with the ICEEMAN decomposition method. Similarly, for site #2, the RMSE decreases from 0.53 without decomposition to 0.41 with ICEEMAN. These results show the significant impact of incorporating decomposition techniques, particularly

TABLE 5. Details of model performance evaluation metrics with decomposition for the two test sites (best results in bold)

Sites/ Methods	Decomposition With ICEEMDAN				
	Individual Deep Learning Model				Combination
	LSTM -NN	BiLSTM-NN	GRU -NN	BiGRU -NN	Class
Site# 1					
RMSE	1.44	0.33	1.39	0.32	**0.21**
MAPE	21.92	5.00	20.97	5.00	**3.08**
MABE	0.08	0.08	0.08	0.08	**0.15**
R^2	0.69	0.98	0.71	0.98	**0.99**
Site# 2					
RMSE	1.47	0.74	1.48	0.54	**0.41**
MAPE	25.67	13.22	25.63	10.22	**6.83**
MABE	0.019	0.026	0.027	0.246	**0.289**
R^2	0.712	0.927	0.718	0.968	**0.979**

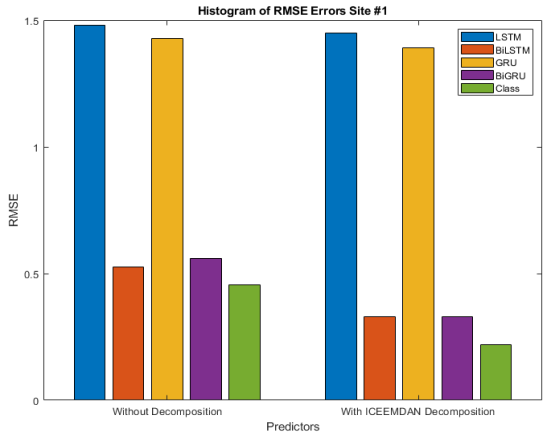


Figure 14. Histogram of RMSE errors for different approaches at Site #1

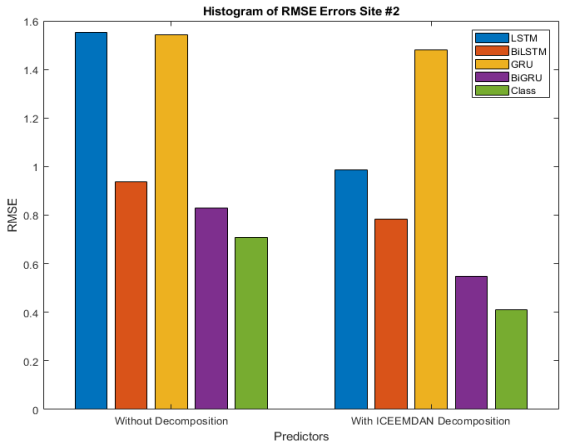


Figure 15. Histogram of RMSE errors for different approaches at Site #2

ICEEMAN, in enhancing predictive performance. Bidirectional models such as BiGRU and BiLSTM achieve lower RMSE values compared to unidirectional models, but they are still outperformed by the "Class" method, especially when combined with ICEEMAN. These findings highlight the ability of combined approaches, strengthened by advanced decomposition techniques, to significantly improve the accuracy and stability of forecasts, confirming their effectiveness for reliable wind speed prediction across the two test sites.

3. 4. 3. Comparison with Existing Methods in the Literature The proposed approach is evaluated in comparison with several state-of-the-art methods from the literature to demonstrate its effectiveness and superiority. Table 6 below provides a comprehensive comparison of the proposed method with some existing techniques, highlighting key performance metrics such as RMSE, MAE, MSE and MAPE. The comparison encompasses only advanced deep learning models.

TABLE 6. Comparison with Some Existing Methods in the Literature

Ref.	Resolution	Model	Forecasting performance
(38)	15-min	SSA-CNNGRU-SVR	MAPE=0.97, MAE=1.72
(39)	10-min	SD-BiGRU	MAE=0.1456 MAPE 4.78 RMSE 0.1871
(40)	10-min	NESN-MP	MAE 0.3 RMSE 0.43 MSE 0.19
(41)	1-h	SDAE-ELM	RMSE 0.21 MAE 0.28
(42)	1-h	WSTDGRU	RMSE 0.3757
Our approch	**Site 1**: 30min	ICEEMDAN-SA-DL-Class	RMSE 0.21 MAPE 3.08 MABE 0.15
	Site 2: 15min		RMSE 0.41 MAPE 6.83 MABE 0.289

The table compares the forecasting performance of various models across different time resolutions. For instance, SSA-CNNGRU-SVR achieves a MAPE of 0.97 and MAE of 1.72 at a 15-minute resolution, while SD-BiGRU demonstrates strong performance with MAE=0.1456 and RMSE=0.1871 at a 10-minute resolution. The proposed ICEEMDAN-SA-DL-Class approach shows competitive results, with RMSE=0.21, MAPE=3.08, and MABE=0.15 for Site 1 (30-minute resolution) and RMSE=0.41, MAPE=6.83, and MABE=0.289 for Site 2 (15-minute resolution). These results highlight the effectiveness of the proposed method in achieving high forecasting accuracy across different time scales.

4. CONCLUSION

Accurate wind speed forecasting (WSF) is essential for optimizing wind energy integration, ensuring grid stability, and enhancing the efficiency of renewable energy systems. This paper proposed a novel three-stage framework for WSF, comprising decomposition, model selection, and forecasting. The decomposition stage leverages advanced techniques to extract meaningful temporal features from the wind speed data using ICEEMDAN decomposition technique, while the features selection phase identifies the most suitable IMFs via Sample Entropy. The forecasting stage employs a fusion of multiple deep learning models to capture complex patterns and temporal dependencies.

The proposed classification-based fusion approach, combined with decomposition, has demonstrated superior predictive accuracy compared to individual models. Experimental results reveal significant improvements in error reduction and forecast stability,

highlighting the robustness and adaptability of the framework. By effectively exploiting the strengths of decomposition, selection, and hybrid forecasting, this method provides a more precise, stable, and reliable solution for wind speed prediction, thereby contributing to the advancement of sustainable wind energy management. Future developments will involve applying more advanced decomposition and fusion techniques to a broader range of time horizons and diverse wind profiles.

5. REFERENCES

1. Kaygusuz K. Renewable energy: Power for a sustainable future. Energy exploration & exploitation. 2001;19(6):603-26. https://doi.org/10.1260/0144598011492723
2. Valdivia-Bautista SM, Domínguez-Navarro JA, Pérez-Cisneros M, Vega-Gómez CJ, Castillo-Téllez B. Artificial intelligence in wind speed forecasting: A review. Energies. 2023;16(5):2457. https://doi.org/10.3390/en16052457
3. Chinforoush N, Latif SG. A novel method for forecasting surface wind speed using wind-direction based on hierarchical Markov model. International Journal of Engineering, Transactions B: Applications. 2021;34 (2):414-26. https://doi.org/10.5829/IJE.2021.34.02B.13
4. Wang Y, Zou R, Liu F, Zhang L, Liu Q. A review of wind speed and wind power forecasting with deep neural networks. Applied energy. 2021;304:117766. https://doi.org/10.1016/j.apenergy.2021.117766
5. Bosch JL, Zheng Y, Kleissl J. Deriving cloud velocity from an array of solar radiation measurements. Solar Energy. 2013;87:196-203. https://doi.org/10.1016/j.solener.2012.10.020
6. Lydia M, Edwin Prem Kumar G, Akash R. Wind speed and wind power forecasting models. Energy & Environment. 2024:0958305X241228515. https://doi.org/10.1177/0958305X241228515
7. Xiong Z, Yao J, Huang Y, Yu Z, Liu Y. A wind speed forecasting method based on EMD-MGM with switching QR loss function and novel subsequence superposition. Applied Energy. 2024;353:122248. https://doi.org/10.1016/j.apenergy.2023.122248
8. Ran M, Huang J, Qian W, Zou T, Ji C. EMD-based gray combined forecasting model-Application to long-term forecasting of wind power generation. Heliyon. 2023;9(7). https://doi.org/10.1016/j.heliyon.2023.e18053
9. Parri S, Teeparthi K, Kosana V. A hybrid methodology using VMD and disentangled features for wind speed forecasting. Energy. 2024;288:129824. https://doi.org/10.1016/j.energy.2023.129824
10. Parri S, Teeparthi K, Kosana V. A hybrid VMD based contextual feature representation approach for wind speed forecasting. Renewable Energy. 2023;219:119391. https://doi.org/10.1016/j.renene.2023.119391
11. Li N, Dong J, Liu L, Li H, Yan J. A novel EMD and causal convolutional network integrated with Transformer for ultra short-term wind power forecasting. International Journal of Electrical Power & Energy Systems. 2023;154:109470. https://doi.org/10.1016/j.ijepes.2023.109470
12. Xia H, Zheng J, Chen Y, Jia H, Gao C. Short-term wind speed combined forecasting model based on multi-decomposition algorithms and frameworks. Electric Power Systems Research. 2024;227:109890. https://doi.org/10.1016/j.epsr.2023.109890
13. Wu H, Meng K, Fan D, Zhang Z, Liu Q. Multistep short-term wind speed forecasting using transformer. Energy. 2022;261:125231. https://doi.org/10.1016/j.energy.2022.125231
14. Fei N, Gao Y, Lu Z, Xiang T, editors. Z-score normalization, hubness, and few-shot learning. Proceedings of the IEEE/CVF International Conference on Computer Vision; 2021.
15. Colominas MA, Schlotthauer G, Torres ME. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. Biomedical Signal Processing and Control. 2014;14:19-29. https://doi.org/10.1016/j.bspc.2014.06.009
16. Ghimire S, Deo RC, Casillas-Perez D, Salcedo-Sanz S. Improved complete ensemble empirical mode decomposition with adaptive noise deep residual model for short-term multi-step solar radiation prediction. Renewable Energy. 2022;190:408-24. https://doi.org/10.1016/j.renene.2022.03.120
17. Xia X, Wang X. A Novel Hybrid Model for Short-Term Wind Speed Forecasting Based on Twice Decomposition, PSR, and IMVO-ELM. Complexity. 2022;2022(1):4014048. https://doi.org/10.1155/2022/4014048
18. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. American journal of physiology-heart and circulatory physiology. 2000;278(6):H2039-H49.
19. Yao P, Xue J, Zhou K, Wang X. Sample Entropy-Based Approach to Evaluate the Stability of Double-Wire Pulsed MIG Welding. Mathematical Problems in Engineering. 2014;2014(1):869631. https://doi.org/10.1155/2014/869631
20. Che J, Ye Y, Wang H, Huang W. A Sample Entropy Parsimonious Model Using Decomposition-ensemble with SSA and CEEMDAN for Short-term Wind Speed Prediction. Engineering Letters. 2023;31(1).
21. Hochreiter S, Schmidhuber J. Long short-term memory. Neural computation. 1997;9(8):1735-80. https://doi.org/10.1162/neco.1997.9.8.1735
22. Gers FA, Schmidhuber J, Cummins F. Learning to forget: Continual prediction with LSTM. Neural computation. 2000;12(10):2451-71. https://doi.org/10.1162/089976600300015015
23. Qadeer K, Rehman WU, Sheri AM, Park I, Kim HK, Jeon M. A long short-term memory (LSTM) network for hourly estimation of PM2. 5 concentration in two cities of South Korea. Applied Sciences. 2020;10(11):3984. https://doi.org/10.3390/app10113984
24. Wang K, Qi X, Liu H. Photovoltaic power forecasting based LSTM-Convolutional Network. Energy. 2019;189:116225. https://doi.org/10.1016/j.energy.2019.116225
25. Ai X, Li S, Xu H. Short-term wind speed forecasting based on two-stage preprocessing method, sparrow search algorithm and long short-term memory neural network. Energy Reports. 2022;8:14997-5010. https://doi.org/10.1016/j.egyr.2022.11.051
26. Sedai A, Dhakal R, Gautam S, Dhamala A, Bilbao A, Wang Q, et al. Performance analysis of statistical, machine learning and deep learning models in long-term forecasting of solar power production. forecasting. 2023;5(1):256-84. https://doi.org/10.3390/forecast5010014
27. Schuster M, Paliwal KK. Bidirectional recurrent neural networks. IEEE transactions on Signal Processing. 1997;45(11):2673-81. https://doi.org/10.1109/78.650093
28. Li J, Song Z, Wang X, Wang Y, Jia Y. A novel offshore wind farm typhoon wind speed prediction model based on PSO–Bi-LSTM improved by VMD. Energy. 2022;251:123848. https://doi.org/10.1016/j.energy.2022.123848
29. Liu X, Lin Z, Feng Z. Short-term offshore wind speed forecast by seasonal ARIMA-A comparison against GRU and LSTM.

Energy. 2021;227:120492. https://doi.org/10.1016/j.energy.2021.120492

30. Zhang G, Liu D. Causal convolutional gated recurrent unit network with multiple decomposition methods for short-term wind speed forecasting. Energy Conversion and Management. 2020;226:113500. https://doi.org/10.1016/j.enconman.2020.113500
31. Haque E, Tabassum S, Hossain E. A comparative analysis of deep neural networks for hourly temperature forecasting. IEEE Access. 2021;9:160646-60. https://doi.org/10.1109/ACCESS.2021.3131533
32. Bibi I, Akhunzada A, Malik J, Iqbal J, Musaddiq A, Kim S. A dynamic DL-driven architecture to combat sophisticated Android malware. IEEE Access. 2020;8:129600-12. https://doi.org/10.1109/ACCESS.2020.3009819
33. Zhang Y, Zhang L, Sun D, Jin K, Gu Y. Short-term wind power forecasting based on VMD and a hybrid SSA-TCN-BiGRU network. Applied Sciences. 2023;13(17):9888. https://doi.org/10.3390/app13179888
34. Tao C, Tao T, He S, Bai X, Liu Y. Wind turbine blade icing diagnosis using B-SMOTE-Bi-GRU and RFE combined with icing mechanism. Renewable Energy. 2024;221:119741. https://doi.org/10.1016/j.renene.2023.119741
35. Zemouri N, Bouzgou H, Gueymard CA. Multimodel ensemble approach for hourly global solar irradiation forecasting. The European Physical Journal Plus. 2019;134(12):594. https://doi.org/10.1140/epjp/i2019-12966-5
36. Bouzgou H, Benoudjit N. Multiple architecture system for wind speed prediction. Applied Energy. 2011;88(7):2463-71. https://doi.org/10.1016/j.apenergy.2011.01.037
37. Bruzzone L, Melgani F. Robust multiple estimator systems for the analysis of biophysical parameters from remotely sensed data. IEEE Transactions on Geoscience and Remote Sensing. 2005;43(1):159-74. https://doi.org/10.1109/TGRS.2004.839818
38. Liu H, Mi X, Li Y, Duan Z, Xu Y. Smart wind speed deep learning based multi-step forecasting model using singular spectrum analysis, convolutional Gated Recurrent Unit network and Support Vector Regression. Renewable energy. 2019;143:842-54. https://doi.org/10.1016/j.renene.2019.05.039
39. Xiang L, Li J, Hu A, Zhang Y. Deterministic and probabilistic multi-step forecasting for short-term wind speed based on secondary decomposition and a deep learning method. Energy Conversion and Management. 2020;220:113098. https://doi.org/10.1016/j.enconman.2020.113098
40. Chitsazan MA, Fadali MS, Trzynadlowski AM. Wind speed and wind direction forecasting using echo state network with nonlinear functions. Renewable energy. 2019;131:879-89. https://doi.org/10.1016/j.renene.2018.07.060
41. Chen L, Li Z, Zhang Y. Multiperiod-ahead wind speed forecasting using deep neural architecture and ensemble learning. Mathematical Problems in Engineering. 2019;2019(1):9240317. https://doi.org/10.1155/2019/9240317
42. Peng Z, Peng S, Fu L, Lu B, Tang J, Wang K, et al. A novel deep learning ensemble model with data denoising for short-term wind speed forecasting. Energy Conversion and Management. 2020;207:112524. https://doi.org/10.1016/j.enconman.2020.112524

Persian Abstract

چکیده

این مطالعه یک مدل ترکیبی جدید برای پیش‌بینی سرعت باد (WSF) بر اساس یک چارچوب سه مرحله‌ای شامل تجزیه، انتخاب ویژگی و پیش‌بینی پیشنهاد می‌کند. رویکرد پیشنهادی از تجزیه حالت تجربی کامل مجموعه بهبود یافته با نویز تطبیقی (ICEEMDAN) برای تجزیه سری های زمانی سرعت باد به توابع حالت ذاتی (IMFs) استفاده می کند. سهم متمایز این مطالعه استفاده از آنتروپی نمونه به عنوان مکانیزم انتخاب ویژگی برای شناسایی مرتبط‌ترین توابع حالت ذاتی (IMF) است. سپس IMF های انتخاب شده از طریق یک تکنیک همجوشی مبتنی بر طبقه بندی ادغام می شوند و دقت پیش بینی را به طور قابل توجهی افزایش می دهند و این رویکرد را از روش های مرسوم متمایز می کنند. دو رویکرد پیش‌بینی مجزا با استفاده از معیارهای عملکرد چندگانه، از جمله RMSE، MAE، MAPE و R^2 ارزیابی می‌شوند. روش اول تکنیک همجوشی را مستقیماً در سری زمانی اصلی سرعت باد اعمال می‌کند، در حالی که روش دوم ICEEMDAN را برای تجزیه سری‌های زمانی ترکیب می‌کند. اعتبار سنجی تجربی با استفاده از دو مجموعه داده دنیای واقعی از الجزایر، برتری مدل ترکیبی پیشنهادی را نسبت به مدل‌های پیش‌بینی فردی نشان می‌دهد و پیشرفت‌های قابل‌توجهی در دقت پیش‌بینی، استحکام و پایداری دارد. این یافته‌ها بر اثربخشی چارچوب سه‌مرحله‌ای تأکید می‌کند، که راه‌حلی قابل اعتماد و کارآمد برای پیش‌بینی کوتاه‌مدت سرعت باد، با کاربردهای بالقوه در مدیریت انرژی‌های تجدیدپذیر و بهینه‌سازی شبکه ارائه می‌دهد.