Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

A Decomposition-Based Deep Learning Approach for Wind Speed Forecasting Using CEEMDAN and LSTM

G. Rakesh Yadav[1], Dr. V. Ramu[2]

[1]Research Scholar Dept. of EEE, Kakatiya University, Warangal, India.
[2]Associate Professor, Dept. of EEE University College of Engineering, Kakatiya University, Kothagudem.

Abstract

Accurate prediction of wind speed is crucial for the reliable operation of wind energy systems and the stability of modern power grids. Nevertheless, wind speed data are highly nonlinear, non-stationary, and influenced by various environmental factors, making precise forecasting a challenging task. To address these challenges, this study proposes a hybrid deep learning framework that combines Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and a Long Short-Term Memory (LSTM) neural network for short-term wind speed prediction. In the proposed approach, the CEEMDAN technique is first employed to decompose the original wind speed time series into a finite number of intrinsic mode functions (IMFs), each representing different frequency components and hidden patterns within the data. This decomposition reduces noise and enhances the interpretability of complex temporal structures. Subsequently, each IMF sub-series is individually modelled using LSTM networks, which are well-suited for capturing long-term dependencies and nonlinear relationships in time-series data. The outputs of these individual models are then aggregated to produce the final wind speed forecast. The effectiveness of the proposed hybrid model is validated using real-time wind speed data obtained from the National Institute of Wind Energy (NIWE), India. Performance evaluation is carried out using standard statistical error metrics, demonstrating that the proposed CEEMDAN–LSTM model significantly outperforms conventional and standalone models in terms of prediction accuracy. The results confirm that the developed approach provides a robust and efficient solution for short-term wind speed forecasting in practical renewable energy applications.

Keywords: Renewable energy systems, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN), Deep neural network, short-term forecasting, Wind speed prediction

1. Introduction

The rapid growth in global energy demand and environmental concerns has accelerated the adoption of renewable energy sources, particularly wind energy. Wind power is a sustainable and clean energy solution as presented in figure 1. However, its integration into modern power systems presents significant challenges due to the intermittent and unpredictable nature of wind speed. Accurate short-term wind speed forecasting is therefore essential for enhancing grid stability, improving energy management, and optimizing wind farm operations. Wind speed time series are inherently nonlinear, non-stationary, and highly volatile, making accurate predictions a complex task. Traditional statistical methods such as autoregressive integrated moving average (ARIMA) have been widely applied for wind speed forecasting but often fail to handle nonlinear patterns effectively [1]. Persistence models, although simple, are limited in capturing dynamic variations in wind behavior [2]. Regression-based approaches have also been explored, but their performance remains constrained under fluctuating environmental conditions [3].

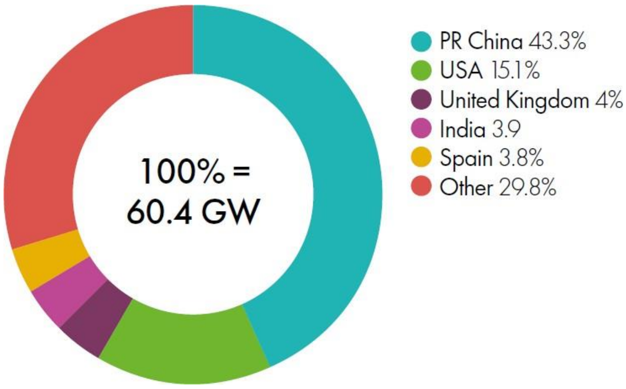


Figure 1. capacity in 2019 and share of top five markets (%)

To address these limitations, machine learning techniques have been introduced for wind speed prediction. Support vector machines (SVM) have demonstrated improved performance in modelling nonlinear relationships [4]. Artificial neural networks (ANN) have also been widely used due to their capability to approximate complex functions [5]. Random forest models have shown robustness in handling large datasets and reducing overfitting [6]. However, these models often lack the ability to capture temporal dependencies effectively. Deep learning approaches have recently gained prominence in time-series forecasting. Recurrent neural networks (RNN) are capable of modelling sequential data but suffer from vanishing gradient problems [7]. Long Short-Term Memory (LSTM) networks overcome this limitation by incorporating memory cells and gating mechanisms [8]. Several studies have reported improved forecasting performance using LSTM-based models for wind speed prediction [9]. Gated recurrent units (GRU) have also been explored as a computationally efficient alternative to LSTM [10]. Hybrid deep learning architectures combining convolutional neural networks (CNN) with LSTM have further enhanced feature extraction and prediction accuracy [11].

Despite these advancements, standalone deep learning models may struggle with highly non-stationary signals. To address this, signal decomposition techniques have been integrated into forecasting frameworks. Empirical mode decomposition (EMD) decomposes signals into intrinsic mode functions (IMFs) for better analysis [12]. Ensemble empirical mode decomposition (EEMD) improves decomposition by reducing mode mixing through noise addition [13]. Variational mode decomposition (VMD) provides better frequency separation and stability [14]. Wavelet transform has also been used for multi-resolution analysis of wind speed signals [15]. Among these techniques, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) has emerged as a robust approach that effectively eliminates mode mixing and enhances decomposition quality [16]. CEEMDAN-based hybrid models have shown superior performance in handling noisy and complex signals [17]. The

combination of decomposition techniques with machine learning models has been widely studied to improve forecasting accuracy.
Hybrid models such as EMD-SVM have demonstrated improved prediction capability by decomposing signals prior to modelling [18]. EEMD-ANN frameworks have shown enhanced performance in capturing nonlinear characteristics [19]. EEMD-LSTM models have further improved forecasting accuracy by leveraging deep learning for each decomposed component [20]. VMD-based hybrid models combined with deep neural networks have achieved superior results in short-term prediction tasks [21]. Recent studies have focused on integrating advanced decomposition methods with deep learning architectures for improved performance. CEEMDAN-LSTM models have demonstrated high accuracy in wind speed forecasting by effectively capturing multi-scale temporal patterns [22]. Optimization-based hybrid models incorporating metaheuristic algorithms have also been explored to enhance model parameters [23]. Attention-based LSTM models have been proposed to improve feature weighting and prediction accuracy [24]. Additionally, ensemble learning approaches have been used to combine multiple forecasting models for robust performance [25]. Motivated by these developments, this study proposes a hybrid CEEMDAN–LSTM model for short-term wind speed forecasting. The CEEMDAN technique is employed to decompose the wind speed signal into multiple IMFs, each representing distinct temporal characteristics. Subsequently, LSTM networks are used to model each IMF independently, enabling effective learning of nonlinear and temporal dependencies. The outputs of individual models are aggregated to produce the final forecast. The proposed model is validated using real-time data from the National Institute of Wind Energy (NIWE), India, and results demonstrate its superiority over conventional and existing hybrid approaches.

2. Background of Deep Learning Approach

This section presents a comprehensive description of the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) technique, the Long Short-Term Memory (LSTM) neural network, and the proposed hybrid CEEMDAN–LSTM framework for accurate short-term wind speed forecasting. The integration of signal decomposition with deep learning provides a powerful mechanism to handle the nonlinear, non-stationary, and stochastic nature of wind speed data.

2.1 Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)

Empirical Mode Decomposition (EMD) is a data-driven signal processing technique introduced as part of the Hilbert–Huang Transform (HHT). It decomposes a complex signal into a set of intrinsic mode functions (IMFs) along with a residual component. Each IMF represents oscillatory modes embedded in the signal, capturing local temporal characteristics at different frequency scales. This adaptive decomposition makes EMD particularly suitable for analyzing nonlinear and non-stationary signals such as wind speed time series. The decomposition process in EMD is based on an iterative procedure called sifting. In this process, local extrema are identified, upper and lower envelopes are constructed using spline interpolation, and their mean is subtracted from the original signal. This procedure is repeated until the extracted component satisfies the conditions of an IMF: (i) the number of zero crossings and extrema differ at most by one, and (ii) the mean of the envelopes is zero. The signal can therefore be expressed as the sum of IMFs and a residual trend component.

Despite its adaptability, EMD suffers from a critical limitation known as mode mixing, where a single

Advanced International Journal for Research (AIJFR)
E-ISSN: 3048-7641 ● Website: www.aijfr.com ● Email: editor@aijfr.com

IMF may contain signals of widely disparate scales or similar frequency components may appear in different IMFs. This phenomenon reduces interpretability and negatively affects forecasting accuracy. To address this issue, Ensemble Empirical Mode Decomposition (EEMD) was introduced as a noise-assisted variant of EMD. In EEMD, white noise is added to the original signal multiple times, and EMD is applied to each noisy realization. The final IMFs are obtained by averaging the corresponding components across all trials. This process reduces mode mixing by distributing the signal energy across different scales. However, EEMD introduces residual noise in the reconstructed signal and requires a large number of ensemble trials, resulting in high computational complexity.

To overcome these drawbacks, Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) was developed as an advanced extension of EEMD. CEEMDAN improves decomposition accuracy by introducing adaptive noise at each stage of the decomposition process, ensuring that the resulting IMFs are more physically meaningful and free from residual noise. Unlike EEMD, CEEMDAN computes each IMF sequentially, using the residual obtained from the previous stage, which enhances both computational efficiency and decomposition precision. As presented in Figure 2 The CEEMDAN algorithm operates as follows. Initially, white noise is added to the original signal, and EMD is applied to generate the first IMF. The mean of these IMFs across multiple realizations is taken as the final first IMF. Subsequently, the residual signal is computed by subtracting the obtained IMF from the original signal. At each subsequent stage, adaptive noise is added to the residual, and the decomposition process is repeated to extract higher-order IMFs. This iterative procedure continues until the residual becomes a monotonic function or contains no further oscillatory modes.

Mathematically, the wind speed signal $x(t)$can be decomposed as:

$$x(t) = \sum_{i=1}^{n} IM\,F_i(t) + r_n(t)$$

where $IMF_i(t)$ represents the i^{th} intrinsic mode function and $r_n(t)$ denotes the final residual component.

The advantages of CEEMDAN include effective elimination of mode mixing, reduced reconstruction error, improved separation of frequency components, enhanced robustness to noise and lower computational burden compared to EEMD. In the context of wind speed forecasting, CEEMDAN plays a crucial role in transforming a highly complex signal into multiple simpler sub-series. Each IMF captures specific temporal dynamics such as high-frequency fluctuations, medium-scale variations, and long-term trends. This decomposition significantly simplifies the forecasting task, enabling machine learning models to learn more effectively from each component.

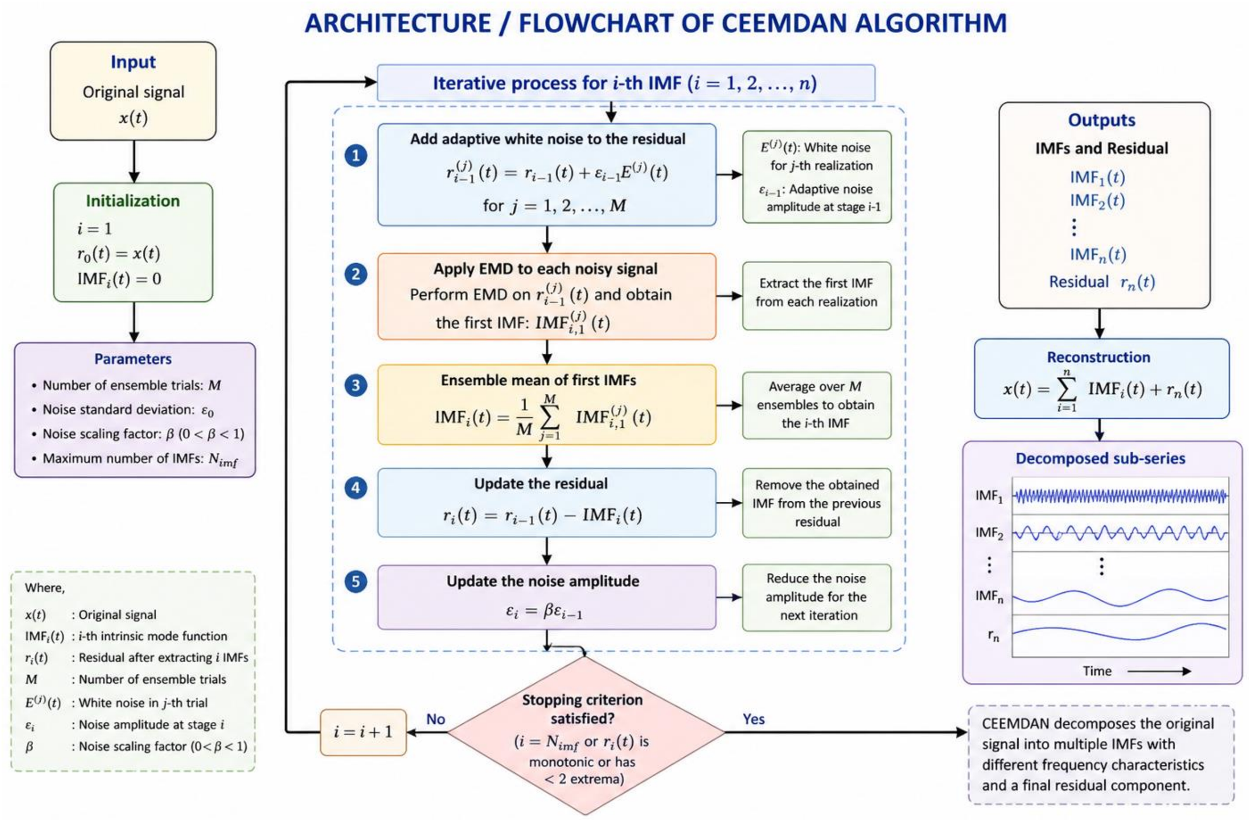


Figure 2. Detailed flowchart of CEEMDAN algorithm

2.2 Long Short-Term Memory (LSTM) Network

Recurrent Neural Networks (RNNs) are widely used for sequential data modeling due to their ability to retain information from previous time steps. However, conventional RNNs suffer from vanishing and exploding gradient problems, which limit their capability to capture long-term dependencies. Long Short-Term Memory (LSTM) networks were developed to address these limitations by introducing a specialized memory architecture. An LSTM network consists of memory cells that maintain information over time, along with three key gating mechanisms: the input gate, forget gate, and output gate. These gates regulate the flow of information, allowing the network to selectively retain or discard information as needed. The input gate determines how much new information should be stored in the cell state. The forget gate decides which information from the previous cell state should be discarded. The output gate controls the information that is passed to the next hidden state and the final output. Architecture of LSTM is presented in Figure 2.

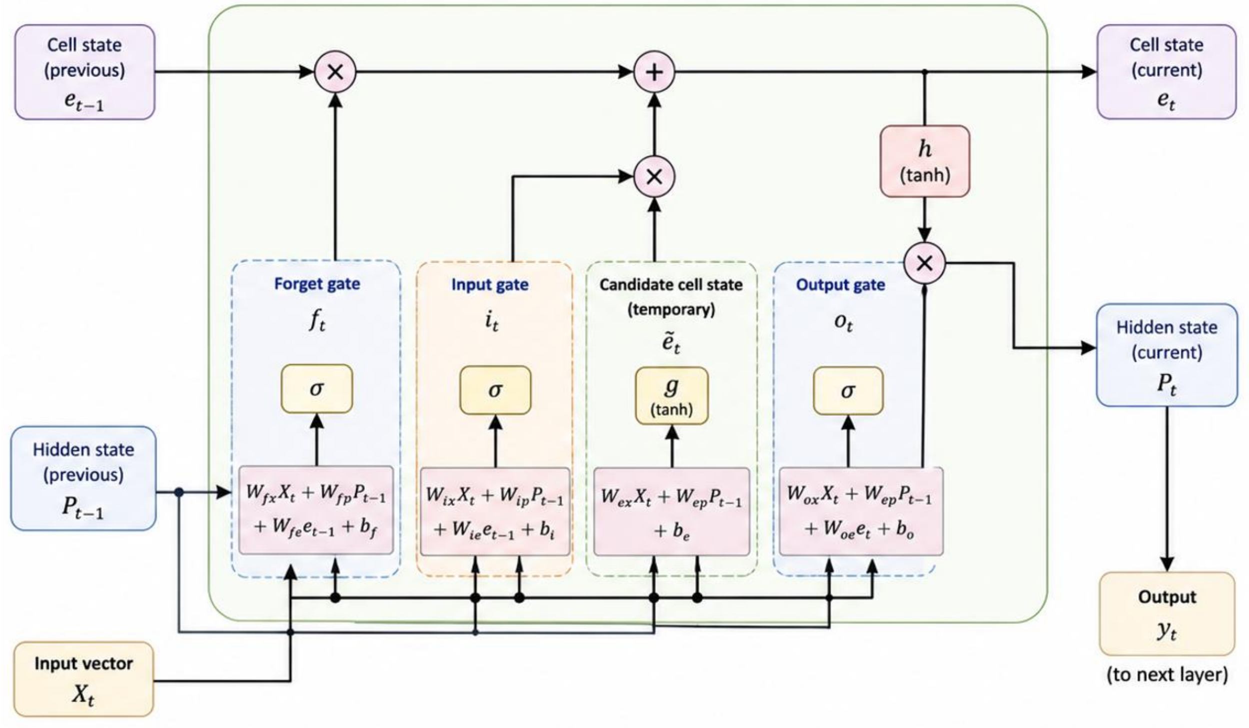


Fig. 2. Architecture of LSTM

The mathematical operations governing an LSTM cell are as follows:

Input gate:

$$i_t = \sigma(W_{ix}X_t + W_{ip}P_{t-1} + W_{ie}e_{t-1} + b_i)$$

Forget gate:

$$f_t = \sigma(W_{fx}X_t + W_{fp}P_{t-1} + W_{fe}e_{t-1} + b_f)$$

Output gate:

$$o_t = \sigma(W_{ox}X_t + W_{op}P_{t-1} + W_{oe}e_t + b_o)$$

Candidate cell state:

$$\tilde{e}_t = g(W_{ex}X_t + W_{ep}P_{t-1} + b_e)$$

Cell state update:

$$e_t = i_t \odot \tilde{e}_t + f_t \odot e_{t-1}$$

Hidden state (output of memory block):

$$P_t = o_t \odot h(e_t)$$

Final output layer:

$$y_t = \phi(W_{yp}P_t + b_y)$$

Where X_t is the input vector at time t, y_t is the output vector, i_t, f_t, o_t are gate activations, e_t is the cell state, P_t is the hidden state, σ denotes the sigmoid activation function, $g(\cdot)$ and $h(\cdot)$ represent activation, $\odot$ denotes element-wise multiplication, W and b are weights and biases. The key strength of LSTM lies in its ability to learn long-term dependencies and nonlinear relationships in sequential data. This makes it highly suitable for wind speed forecasting, where temporal correlations and dynamic patterns play a critical role.

2.3 Proposed Hybrid CEEMDAN–LSTM Framework

The proposed hybrid framework integrates CEEMDAN and LSTM to leverage the strengths of both signal decomposition and deep learning. The core idea is to decompose the original wind speed time series into multiple simpler components and then model each component individually using LSTM networks.

The overall workflow of the proposed model is as follows:

1. **Data Preprocessing:** The raw wind speed data obtained from measurement stations are first normalized to ensure numerical stabilitya and improved learning performance.
2. **Signal Decomposition using CEEMDAN:** The preprocessed wind speed signal is decomposed into multiple IMFs and a residual component. Each IMF represents a specific frequency band, capturing different temporal characteristics of the original signal.
3. **Individual Modeling using LSTM:** Each IMF sub-series is fed into an independent LSTM model. Since each component is simpler and more stationary than the original signal, the LSTM can effectively learn its underlying patterns and temporal dependencies.
4. **Prediction of Sub-series:** The trained LSTM models generate predictions for each IMF independently.
5. **Reconstruction of Final Output:** The predicted IMFs are aggregated (summed) along with the residual component to obtain the final wind speed forecast.

By combining CEEMDAN's powerful decomposition capability with LSTM's sequence learning strength, the proposed framework effectively addresses the challenges associated with wind speed forecasting. The decomposition step isolates different signal behaviours, while the LSTM models capture both short-term fluctuations and long-term dependencies within each component. The CEEMDAN–LSTM hybrid model provides a robust, accurate, and scalable solution for short-term wind speed prediction, making it highly suitable for real-world renewable energy applications and smart grid integration. Overall architecture of CEEMDAN–LSTM hybrid model is presented in Figure 3. The decomposition results for one month of training data are shown in Figure 4.

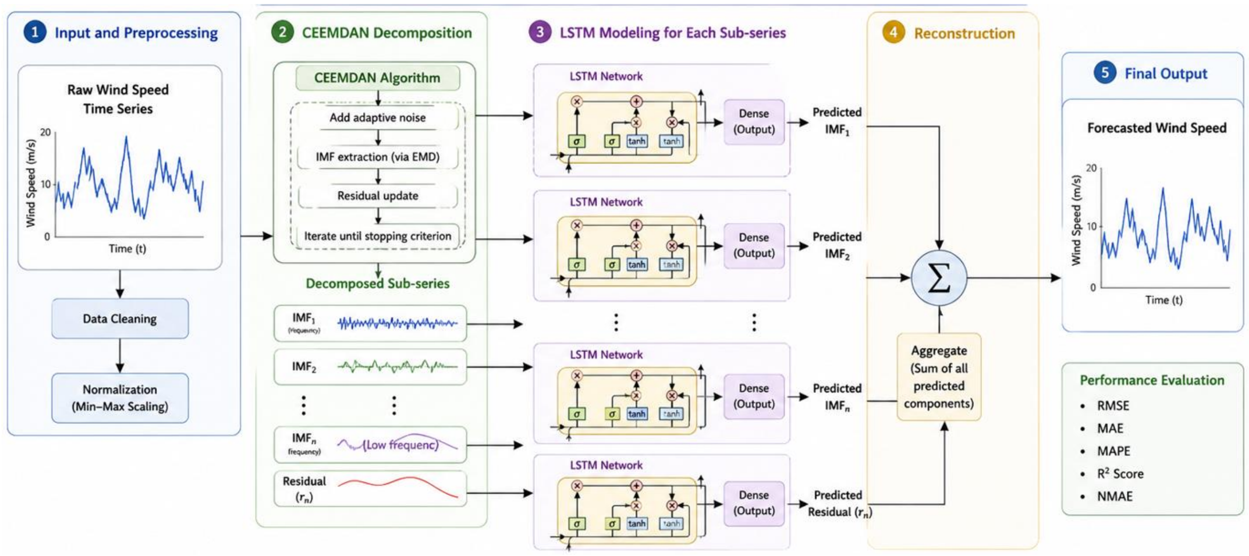


Figure 3. Overall architecture of CEEMDAN–LSTM hybrid model

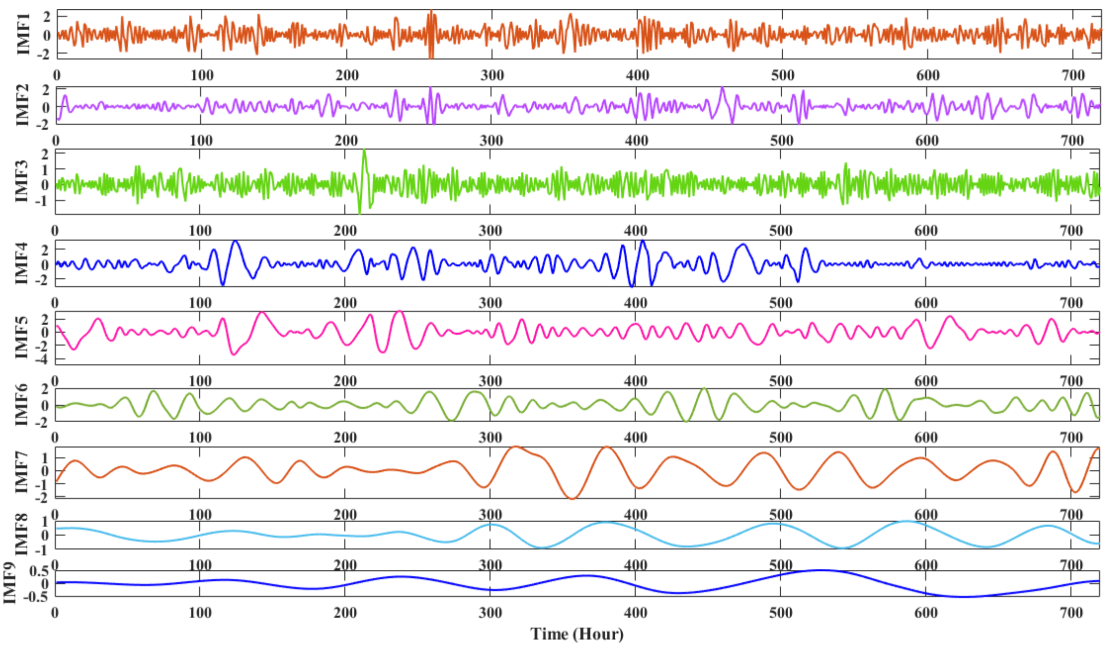


Figure 4. Decomposed sub-series using CEEMDAN technique

3. Performance Evaluation of the Developed Model

The performance of the proposed hybrid CEEMDAN–LSTM model is evaluated using real-time wind speed data obtained from the National Institute of Wind Energy (NIWE), India [27]. The dataset consists of hourly sampled wind speed measurements, which capture the stochastic and dynamic characteristics of wind behavior. Such high-resolution temporal data is well-suited for validating short-term forecasting models in practical wind energy applications. To ensure a robust and unbiased evaluation, the dataset is divided into two subsets: 75% of the data is used for training the models, while the remaining 25% is reserved for testing. A one-step-ahead forecasting approach is employed, where the model predicts the wind speed at time $t+1$based on previous observations. This setup reflects real-world operational requirements in wind energy systems, where accurate short-term predictions are essential for grid stability and energy management. The predictive performance of the proposed model is assessed using standard statistical error metrics, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics provide complementary insights into forecasting accuracy, considering both absolute and relative deviations between predicted and actual values. Here, Ndenotes the total number of samples, $i$represents the time index, and $X_{\text{predicted},i}$and $X_{\text{actual},i}$correspond to the predicted and observed wind speed values, respectively.

Table 1. Statistical Error Measures Used for Performance Evaluation

Error Metric	**Mathematical Expression**
Root Mean Square Error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(X_{\text{predicted},i} - X_{\text{actual},i}\right)^2}$
Mean Absolute Error (MAE)	$MAE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_{predicte,i} - X_{actual,i}}{N}\right)$
Mean Absolute Percentage Error (MAPE)	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{X_{actual,i} - X_{predicte,i}}{X_{actual,i}} * 100\right)$

The proposed hybrid CEEMDAN–LSTM model is benchmarked against six comparative approaches: the persistence method, feed-forward backpropagation neural network (FFBP), Elman neural network (ENN), standalone LSTM, CEEMDAN-based FFBP, and CEEMDAN-based ENN models. These baseline models represent a range of traditional statistical, machine learning, and hybrid techniques, providing a comprehensive framework for comparison. The effectiveness of incorporating CEEMDAN decomposition is illustrated in Fig. 7, where forecasting results obtained with and without CEEMDAN are compared. It is evident that models utilizing CEEMDAN preprocessing exhibit significantly improved prediction accuracy. This improvement is attributed to the decomposition of the original wind speed signal into multiple intrinsic mode functions (IMFs), which simplifies the learning process and reduces the impact of non-stationarity.

Further validation of the proposed model is presented in Fig. 8, which compares the predicted wind speed values with the actual test data. The close agreement between the predicted and observed values indicates that the hybrid model effectively captures both high-frequency fluctuations and long-term temporal dependencies. This demonstrates the capability of the CEEMDAN–LSTM framework to model complex wind speed dynamics with high precision. A quantitative comparison of all models is provided in Fig. 9, where bar charts of RMSE, MAE, and MAPE values are presented. Lower error values correspond to higher forecasting accuracy. The results clearly show that the proposed CEEMDAN–LSTM model achieves the lowest error metrics among all considered approaches. While the standalone LSTM model performs better than conventional neural networks such as FFBP and ENN, its performance is further enhanced when combined with CEEMDAN decomposition. Similarly, CEEMDAN-based FFBP and

ENN models demonstrate improved accuracy compared to their standalone counterparts but still do not match the performance of the proposed hybrid model. The results confirm that integrating CEEMDAN with LSTM significantly enhances forecasting performance by effectively addressing the challenges of nonlinearity and non-stationarity in wind speed data. The proposed hybrid model provides a reliable and accurate solution for short-term wind speed forecasting, making it highly suitable for real-world wind energy applications.

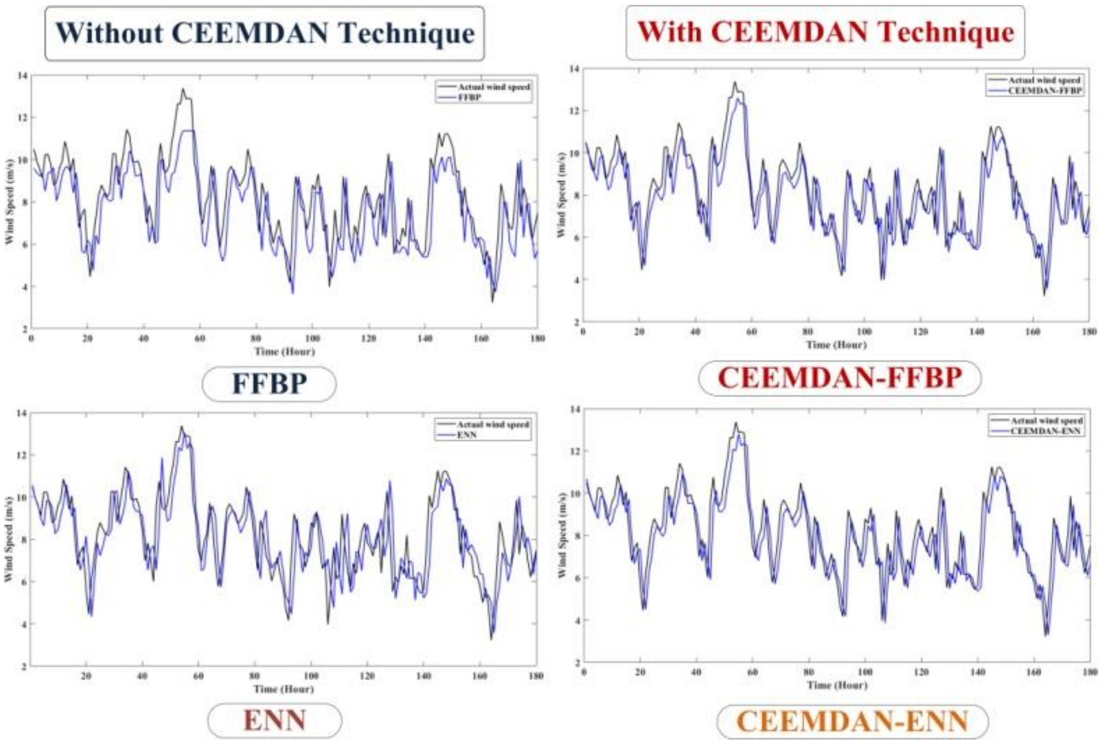


Fig. 7. Effect of CEEMDAN technique in prediction results

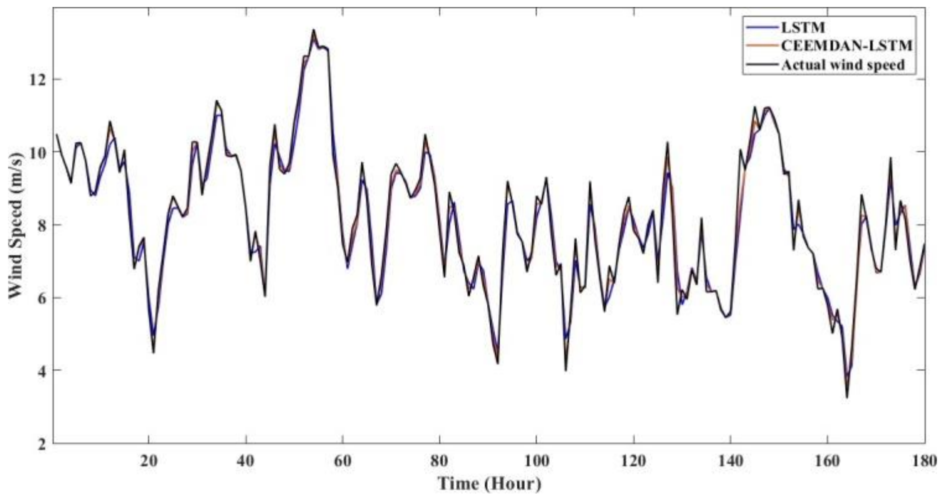
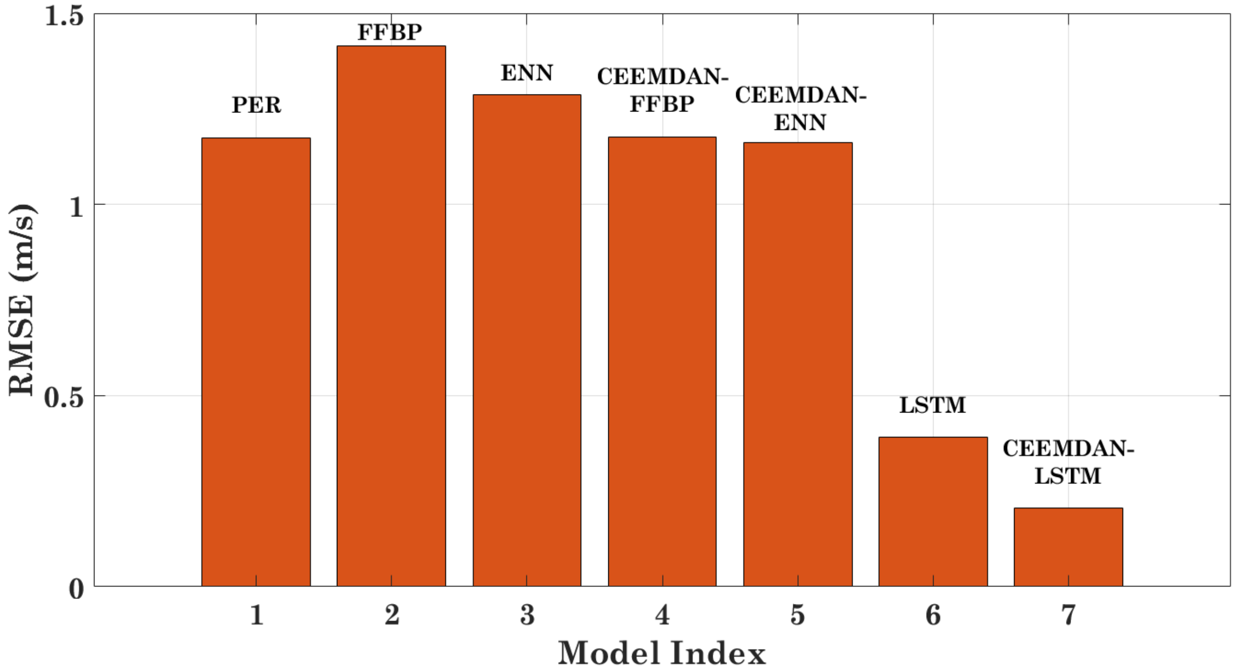
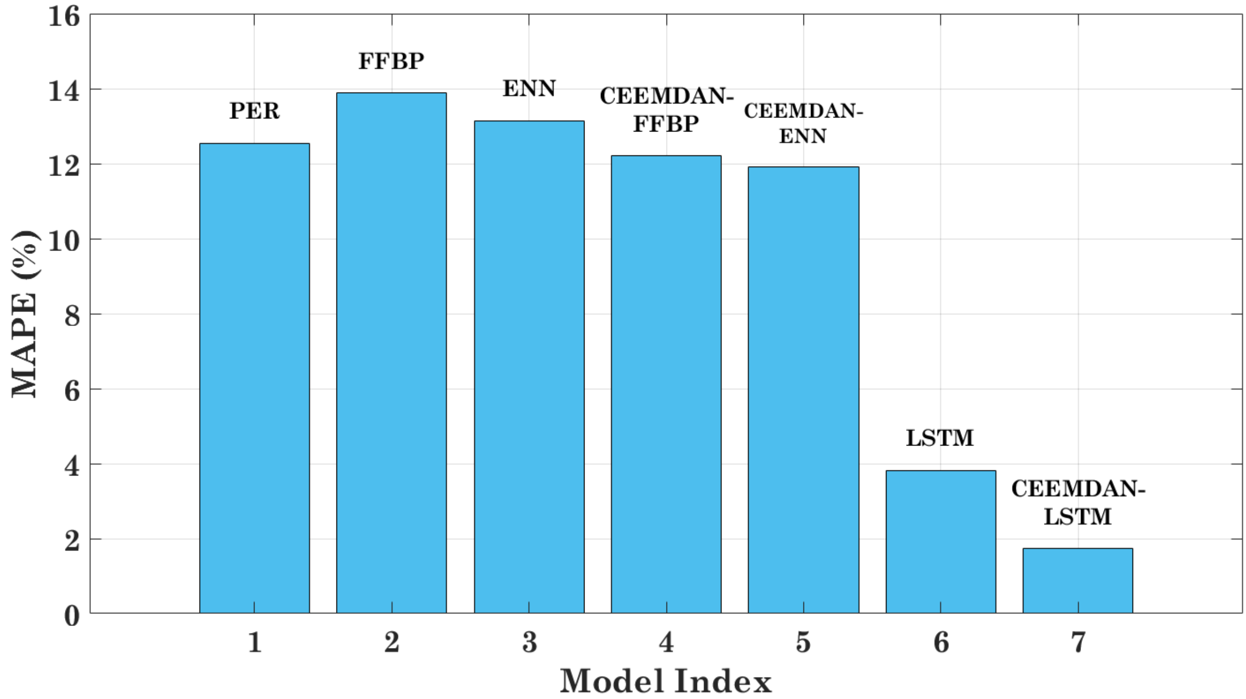


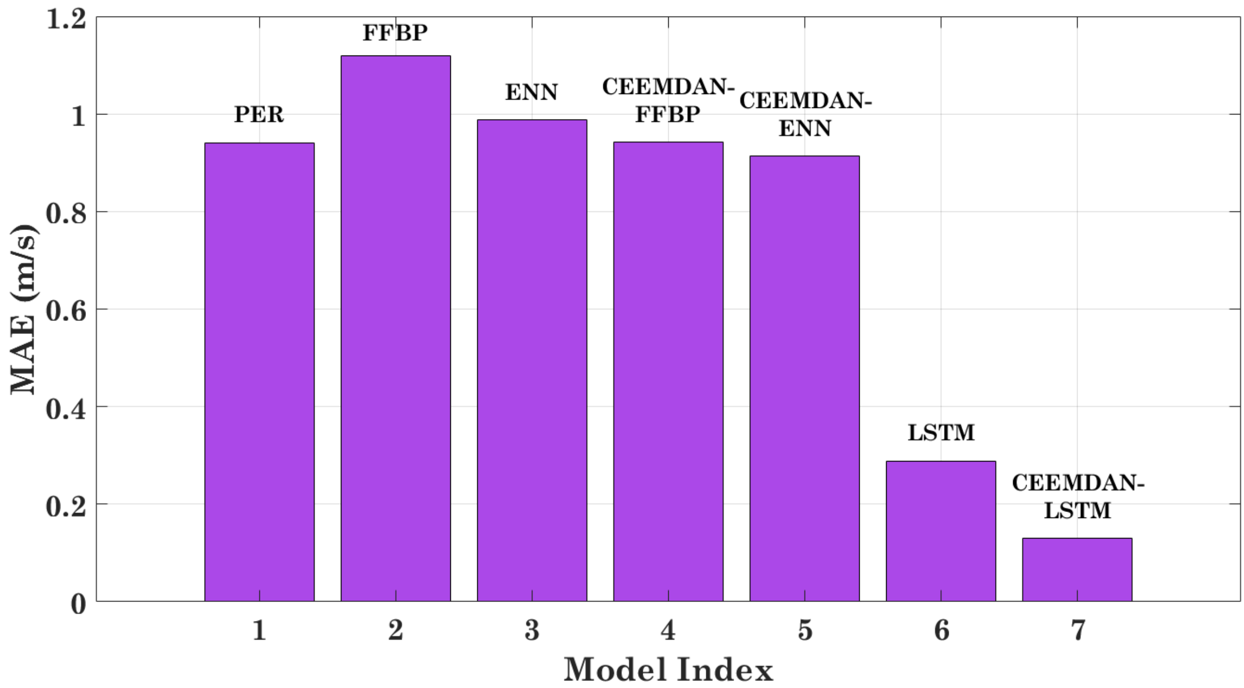
Fig. 8. Comparison of actual wind speed, LSTM model and developed CEEMDAN-LSTM model forecast results


1.5
1
0.5
0
RMSE (m/s)
PER
FFBP
ENN
CEEMDAN-
FFBP
CEEMDAN-
ENN
LSTM
CEEMDAN-
LSTM
1
2
3
4
5
6
7
Model Index

(a)


16
14
12
10
8
6
4
2
0
MAPE (%)
PER
FFBP
ENN
CEEMDAN-
FFBP
CEEMDAN-
ENN
LSTM
CEEMDAN-
LSTM
1
2
3
4
5
6
7
Model Index

(b)



(c)

Fig. 9. Bar charts of statistical error values (a) RMSE (b) MAPE (c) MAE

3.1 Comparative Performance Analysis Using Benchmark Models

Lower values of statistical error metrics indicate higher forecasting accuracy and better predictive efficiency. The comparative performance of the benchmark models and the proposed hybrid CEEMDAN–LSTM model is presented in Table 2. The evaluation is based on three widely accepted metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE).

Table 2. RMSE, MAE, and MAPE Errors for Benchmark Models and Proposed CEEMDAN–LSTM Model

Approach Used	**RMSE (m/s)**	**MAE (m/s)**	**MAPE (%)**
Persistence (PER)	1.1748	0.9409	12.5446
FFBP	1.4142	1.1200	13.8863
ENN	1.2873	0.9880	13.1402
CEEMDAN–FFBP	1.1769	0.9426	12.2124
CEEMDAN–ENN	1.1606	0.9133	11.9165
LSTM	0.3906	0.2890	3.8234
CEEMDAN–LSTM	**0.2064**	**0.1298**	**1.7298**

From Table 2, it is evident that the persistence model exhibits relatively high error values, indicating its limited capability in capturing the nonlinear and dynamic characteristics of wind speed. Among the conventional neural network models, the feed-forward backpropagation (FFBP) and Elman neural network (ENN) models show moderate improvements; however, their performance remains constrained due to their inability to effectively capture long-term temporal dependencies. The incorporation of CEEMDAN as a preprocessing technique significantly enhances the performance of these models. For instance, the RMSE of the FFBP model reduces from 1.4142 to 1.1769 after applying CEEMDAN, while the MAPE decreases from 13.8863% to 12.2124%. Similarly, the ENN model shows improvements in all error metrics when combined with CEEMDAN. This improvement can be attributed to the ability of CEEMDAN to decompose the wind speed signal into simpler sub-series, thereby reducing noise and improving model learning. Among all individual models, the LSTM network demonstrates superior performance due to its ability to capture long-term dependencies and nonlinear temporal patterns. The RMSE and MAPE values of the standalone LSTM model are significantly lower than those of traditional models, highlighting its effectiveness for time-series forecasting. The proposed hybrid CEEMDAN–LSTM model achieves the best performance across all metrics, with RMSE = 0.2064, MAE = 0.1298, and MAPE = 1.7298%. These results clearly indicate that the combination of CEEMDAN and LSTM provides a powerful framework for accurate wind speed prediction by addressing both non-stationarity and temporal complexity.

3.2 Performance Improvement Analysis

To quantitatively assess the improvement achieved by the proposed model, percentage improvements in RMSE, MAE, and MAPE are calculated using the following expressions:

$$P_{RMSE} = \frac{RMSE_1 - RMSE_2}{RMSE_1} \times 100$$

$$P_{MAE} = \frac{MAE_1 - MAE_2}{MAE_1} \times 100$$

$$P_{MAPE} = \frac{MAPE_1 - MAPE_2}{MAPE_1} \times 100$$

where $RMSE_1, MAE_1, MAPE_1$ represent the error values of the benchmark model, and $RMSE_2, MAE_2, MAPE_2$ correspond to the error values of the proposed CEEMDAN–LSTM model.

Table 3. Performance Improvements Achieved by the Proposed Model**

Comparison	**PRMSE (%)**	**PMAE (%)**	**PMAPE (%)**
CEEMDAN–LSTM vs PER	82.4310	86.2047	86.2108
CEEMDAN–LSTM vs FFBP	85.4052	88.4107	87.5431
CEEMDAN–LSTM vs ENN	83.9664	86.8623	86.8358
CEEMDAN–LSTM vs CEEMDAN–FFBP	82.4624	86.2296	85.8357
CEEMDAN–LSTM vs CEEMDAN ENN	82.2161	85.7878	85.4839
CEEMDAN–LSTM vs LSTM	47.1582	55.0865	54.7575

Table 3 clearly highlights the substantial performance gains achieved by the proposed hybrid model. The CEEMDAN–LSTM model shows an improvement of over 82% in RMSE and approximately 86% in MAE and MAPE when compared with the persistence model. Similar significant improvements are observed when compared with FFBP and ENN models, demonstrating the superiority of the proposed approach. Even when compared with CEEMDAN-based hybrid models such as CEEMDAN–FFBP and CEEMDAN–ENN, the proposed model achieves notable improvements exceeding 80% across all metrics. This indicates that while CEEMDAN effectively enhances model performance, the integration with LSTM further amplifies forecasting accuracy due to its advanced temporal learning capability. When compared with the standalone LSTM model, the proposed hybrid approach still achieves considerable improvements of 47.16% in RMSE, 55.08% in MAE, and 54.75% in MAPE. This demonstrates that the inclusion of CEEMDAN preprocessing significantly enhances the predictive capability of LSTM by simplifying the input signal and enabling more efficient learning. The combined analysis of Fig. 7, Fig. 8, Fig. 9, Table 2, and Table 3 confirms that the proposed CEEMDAN–LSTM model consistently outperforms all benchmark models. The superior performance can be attributed to two key factors. First, CEEMDAN effectively decomposes the wind speed signal into multiple intrinsic mode functions, reducing noise and isolating meaningful patterns. Second, the LSTM network efficiently captures both short-term fluctuations and long-term dependencies within each decomposed component. The hybrid CEEMDAN–LSTM framework provides a highly accurate and robust solution for short-term wind speed forecasting, making it well-suited for real-world applications in wind energy systems and smart grid operations.

4. Conclusion

This study presented a hybrid deep learning framework for short-term wind speed forecasting by integrating Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) and a Long Short-Term Memory (LSTM) network. The primary objective was to address the inherent challenges of wind speed prediction arising from its nonlinear, non-stationary, and highly stochastic nature. By combining signal decomposition with advanced sequence learning, the proposed model effectively enhances forecasting accuracy and reliability. The CEEMDAN technique was employed to decompose the original wind speed time series into multiple intrinsic mode functions (IMFs) and a residual component, thereby isolating different frequency characteristics and reducing noise. This decomposition simplifies the complexity of the input data and enables more effective learning. Subsequently, individual LSTM networks were used to model each decomposed sub-series, capturing both short-term fluctuations and long-term temporal dependencies. The final forecast was obtained by aggregating the outputs of all sub-models. The performance of the proposed hybrid CEEMDAN–LSTM model was evaluated using real-time wind speed data obtained from the National Institute of Wind Energy (NIWE), India. Comparative analysis against multiple benchmark models, including Persistence, FFBP, ENN, standalone LSTM, and CEEMDAN-based hybrid models, demonstrated the superiority of the proposed approach. The model achieved significantly lower error values, with RMSE, MAE, and MAPE of 0.2064, 0.1298, and 1.7298%, respectively. Furthermore, performance improvement analysis revealed substantial gains across all evaluation metrics, confirming the effectiveness of the hybrid framework. The results clearly indicate that the integration of CEEMDAN and LSTM provides a robust and efficient solution for wind speed forecasting. The decomposition process enhances data representation, while the LSTM network effectively models temporal dependencies and uncertainties. As a result, the proposed model outperforms conventional and standalone approaches in terms of accuracy and stability. The CEEMDAN–LSTM hybrid model offers a promising tool for short-term wind speed prediction in renewable energy systems. Its ability to handle complex time-series characteristics makes it suitable for practical applications such as wind farm management, power system scheduling, and grid integration. Future work may focus on extending the framework to multi-step forecasting, incorporating additional meteorological variables, and integrating optimization techniques to further enhance model performance.

References

1. Liu, Ming-De, Lin Ding, and Yu-Long Bai. "Application of hybrid model based on empirical mode decomposition, novel recurrent neural networks and the ARIMA to wind speed prediction." *Energy Conversion and Management* 233 (2021): 113917.
2. Cancino-Solórzano, Yoreley, Antonio J. Gutiérrez-Trashorras, and Jorge Xiberta-Bernat. "Analytical methods for wind persistence: Their application in assessing the best site for a wind farm in the State of Veracruz, Mexico." *Renewable Energy* 35, no. 12 (2010): 2844-2852.
3. Roberts, Callum, David Garcia Cava, and Luis D. Avendaño-Valencia. "Addressing practicalities in multivariate nonlinear regression for mitigating environmental and operational variations." *Structural Health Monitoring* 22, no. 2 (2023): 1237-1255.
4. Yu, Chuanjin, Yongle Li, Yulong Bao, Haojun Tang, and Guanghao Zhai. "A novel framework for wind speed prediction based on recurrent neural networks and support vector machine." *Energy Conversion and Management* 178 (2018): 137-145.
5. Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, *22*(5), 717-727.
6. Shi, Kunpeng, Ying Qiao, Wei Zhao, Qing Wang, Menghua Liu, and Zongxiang Lu. "An improved random forest model of short-term wind-power forecasting to enhance accuracy, efficiency, and robustness." *Wind energy* 21, no. 12 (2018): 1383-1394.
7. Torres, José F., Dalil Hadjout, Abderrazak Sebaa, Francisco Martínez-Álvarez, and Alicia Troncoso. "Deep learning for time series forecasting: a survey." *Big data* 9, no. 1 (2021): 3-21.
8. Ghojogh, Benyamin, and Ali Ghodsi. "Recurrent neural networks and long short-term memory networks: Tutorial and survey." *arXiv preprint arXiv:2304.11461* (2023).
9. Zheng, Lang, Weisheng Lu, and Qianyun Zhou. "Weather image-based short-term dense wind speed forecast with a ConvLSTM-LSTM deep learning model." *Building and Environment* 239 (2023): 110446.
10. Nosouhian, Shiva, Fereshteh Nosouhian, and Abbas Kazemi Khoshouei. "A review of recurrent neural network architecture for sequence learning: Comparison between LSTM and GRU." (2021).
11. Agga, Ali, Ahmed Abbou, Moussa Labbadi, Yassine El Houm, and Imane Hammou Ou Ali. "CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production." *Electric Power Systems Research* 208 (2022): 107908.
12. Chen, Youwei, Pengwei Zhao, Zhen Zhang, Juncheng Bai, and Yuqi Guo. "A stock price forecasting model integrating complementary ensemble empirical mode decomposition and independent component analysis." *International Journal of Computational Intelligence Systems* 15, no. 1 (2022): 75.

13. Pan, Zhen, Biao Xu, Wenjia Chen, Dian Fan, Xianghan Meng, Mengfan Peng, and Ciming Zhou. "Improved complete ensemble empirical mode decomposition with adaptive noise and its application in noise reduction for fiber optic sensing." *Journal of Lightwave Technology* 43, no. 5 (2024): 2466-2474.
14. Jia, Hao, Pengfei Cao, Tong Liang, Cesar F. Caiafa, Zhe Sun, Yasuhiro Kushihashi, Antoni Grau, Yolanda Bolea, Feng Duan, and Jordi Solé-Casals. "Short-time variational mode decomposition." *Signal Processing* (2025): 110203.
15. Hu, Rui, Weihao Hu, Nuri Gökmen, Pengfei Li, Qi Huang, and Zhe Chen. "High resolution wind speed forecasting based on wavelet decomposed phase space reconstruction and self-organizing map." *Renewable energy* 140 (2019): 17-31.
16. Elouaham, S., A. Dliou, R. Latif, and M. Laaboubi. "Filtering of biomedical signals by using complete ensemble empirical mode decomposition with adaptive noise." *International Journal of Computer Application* 149, no. 7 (2016).
17. Xu, Yang, Junyan Hu, Cheng Yao, Xiaorui Hu, Hexin Peng, and Hongyu Long. "Integrating CEEMDAN-EWT and Hybrid Neural Network for Traffic Flow Forecasting." *Engineering Letters* 33, no. 8 (2025).
18. Chen, Ning, Hongxin Sun, Qi Zhang, and Shouke Li. "A short-term wind speed forecasting model based on EMD/CEEMD and ARIMA-SVM algorithms." *Applied Sciences* 12, no. 12 (2022): 6085.
19. Liao, Shengli, Huan Wang, Benxi Liu, Xiangyu Ma, Binbin Zhou, and Huaying Su. "Runoff forecast model based on an EEMD-ANN and meteorological factors using a multicore parallel algorithm." *Water Resources Management* 37, no. 4 (2023): 1539-1555.
20. Wu, Sen, Wei Wang, Yanan Song, and Shuaiqi Liu. "An EEMD-LSTM, SVR, and BP decomposition ensemble model for steel future prices forecasting." *Expert Systems* 41, no. 11 (2024): e13672.
21. Niu, Hongli, Kunliang Xu, and Weiqing Wang. "A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network: A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network." *Applied Intelligence* 50, no. 12 (2020): 4296-4309.
22. Song, Yuping, Jiefei Huang, Yang Xu, Jinrui Ruan, and Min Zhu. "Multi-decomposition in deep learning models for futures price prediction." *Expert Systems with Applications* 246 (2024): 123171.
23. Pérez-Rodríguez, Sergio A., José M. Álvarez-Alvarado, Julio-Alejandro Romero-González, Marcos Aviles, America Eileen Mendoza-Rojas, Carlos Fuentes-Silva, and Juvenal Rodríguez-Reséndiz. "Metaheuristic algorithms for solar radiation prediction: A systematic analysis." *IEEE Access* 12 (2024): 100134-100151.
24. Kang, Qing, Elton J. Chen, Zhong-Chao Li, Han-Bin Luo, and Yong Liu. "Attention-based LSTM predictive model for the attitude and position of shield machine in tunneling." *Underground Space* 13 (2023): 335-350.
25. Didona, Diego, Francesco Quaglia, Paolo Romano, and Ennio Torre. "Enhancing performance prediction robustness by combining analytical modeling and machine learning." In *Proceedings of the 6th ACM/SPEC international conference on performance engineering*, pp. 145-156. 2015.