HYBRID ATTENTION-BASED DEEP LEARNING MODEL FOR WIND SPEED FORECASTING IN RENEWABLE ENERGY APPLICATIONS

[1]HEMANTH SAI MADUPU, [2]BALA SAIBABU BOMMIDI, [3]IMRAN ABDUL, [4]K. RAMBABU, [5]BEJJAM S N BENARJI

[1,2]Assistant Professor, Department of Electrical and Electronics Engineering,
Prasad V. Potluri Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India
[3]Assistant Professor, Department of Electrical and Electronics Engineering,
Lakireddy Bali Reddy College of Engineering, Mylavaram, Andhra Pradesh, India
[4]Assistant Professor, Department of Electrical and Electronics Engineering,
Aditya University, Surampalem, Andhra Pradesh, India
[5]Assistant Professor, Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India

[1]hemanthsai.madupu@gmail.com [2]saibabu.dec31@gmail.com [3]abdulimran201@gmail.com
[4]rambabu.kasukurthi@gmail.com, [5]benarji@kluniversity.in

ABSTRACT

Accurate wind speed forecasting (WSF) is critical for the efficient integration of wind energy into power systems. In this study, a novel hybrid model combining Ensemble Empirical Mode Decomposition (EEMD), Convolutional Neural Network (CNN), and Attention-based Long Short-Term Memory (ALSTM)—referred to as EEMD-CNN-ALSTM—is proposed for short-term wind speed prediction. The model leverages EEMD to decompose complex and non-stationary wind speed signals into more manageable components, CNN to extract localized spatial features, and ALSTM to selectively capture temporal dependencies using attention mechanisms. The proposed model is evaluated using 1-hour interval wind speed data collected from two distinct wind farms: Garden City and Idalia. Performance is assessed using standard metrics including RMSE, MAE, MSE, and R^2. Results demonstrate that the EEMD-CNN-ALSTM model consistently outperforms benchmark models and other hybrid architectures. Across both wind farms, the proposed model achieves the lowest error rates and highest predictive accuracy, confirming its robustness and effectiveness in handling the complexities of WSF.

Keywords: *Attention-Based LSTM, Hybrid Model, Time Series Prediction, Deep Learning, Renewable Energy, Short-Term Forecasting*

1. INTRODUCTION

Wind power is becoming an increasingly important part of the global transition to sustainable energy, driven by growing concerns over climate change and the urgent need to reduce greenhouse gas emissions [1]. The use of wind turbines to generate electricity is gaining widespread attention because it offers a clean, renewable source of energy that does not produce harmful emissions. As a result, wind energy is seen as a vital solution in efforts to decarbonize energy systems and move away from fossil fuel dependence, contributing significantly to global climate goals [2]. Despite its benefits, wind power faces challenges due to the natural variability of wind, which can affect both the efficiency and reliability of electricity generation. This unpredictability also makes it difficult to smoothly integrate wind energy into the electrical grid [3]. To address these issues, accurately forecasting wind speed has become essential. Reliable predictions help improve the day-to-day operation of wind farms, support the development of more effective control systems, and enable predictive maintenance. These improvements not only lower operational costs but also extend the lifespan of wind turbines, making wind energy more efficient and sustainable in the long run.

In this study, several deep learning model architectures have been explored and implemented to enhance the accuracy and efficiency of wind energy forecasting. The models considered include Artificial Neural Networks (ANN), Recurrent

Neural Networks (RNN), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and Gated Recurrent Units (GRU). These models are well-suited for analyzing time series data, which is essential when dealing with variables like wind speed that change over time. Through extensive training and testing, the aim is to identify the most effective architecture for precise wind speed prediction, ultimately enhancing the reliability and efficiency of wind energy generation.

To overcome the challenges posed by the variability and unpredictability of wind, researchers have been actively exploring machine learning and deep learning algorithms for short-term wind speed prediction. These methods aim to boost the reliability and performance of wind energy systems. The authors in [4] applied a decision tree-based model to evaluate different predictors of wind power, such as hub height wind speed, rotor-equivalent wind speed, and the lapse rate (the rate at which air temperature decreases with height). Their findings showed that including rotor-equivalent wind speed and lapse rate improved the model's accuracy by 22%, outperforming traditional power curve approaches. In [5] a two-stage model is used by combining statistical and machine learning techniques. First, they used the ARIMA (AutoRegressive Integrated Moving Average) model to predict transformed wind-related variables individually. These outputs were then used as inputs to an Artificial Neural Network (ANN) to forecast wind speed, effectively blending linear and non-linear modeling techniques. The authors in [6] compared the performance of the random forest algorithm—an ensemble tree-based model—with ANNs for wind speed prediction. Their results indicated that random forest significantly outperformed ANN, highlighting the effectiveness of ensemble learning methods in handling non-linearities in wind data. A hybrid forecasting model that combines Long Short-Term Memory (LSTM) neural networks with the Grey Wolf Optimizer (GWO), a nature-inspired optimization algorithm is introduced in [7]. This combination improved the model's ability to capture the non-linear patterns in wind speed series, resulting in better prediction accuracy. The authors in [8] utilized mutual information (MI) to identify the most relevant features in the wind dataset and integrated this with a Recurrent Neural Network (RNN) to forecast both wind speed and direction. The model outperformed traditional forecasting methods due to better feature selection and modeling of temporal dependencies. A mesoscale wind forecasting model [9] is developed, by integrating the Weather Research and Forecasting (WRF) numerical weather prediction model with an ANN. This hybrid model showed high accuracy and demonstrated the potential of combining physical models with machine learning. The authors in [10] developed a method that involved dividing the data into sub-regions using a Data Area Division (DAD) technique and then applying a Genetic Algorithm (GA)-optimized LSTM model. Their approach yielded very low RMSE values (0.2–0.6 m/s) and showed excellent stability, emphasizing the benefits of regional modeling and hyperparameter optimization. The authors proposed a CNN-RNN hybrid model [11] to forecast wind speed across four buoy stations in a wind farm. This model combined the spatial pattern extraction capabilities of Convolutional Neural Networks (CNNs) with the temporal modeling strength of RNNs, resulting in high forecasting accuracy. The authors in [12] went further by modeling wind vectors as time series images and applying a Three-Dimensional CNN (3DCNN) in combination with a Deep Convolutional LSTM (DConvLSTM). This approach outperformed DConvLSTM alone by significantly reducing training time and prediction error, demonstrating the effectiveness of deep spatio-temporal modeling. In [13] a novel hybrid approach is introduced using a Nonlinear Auto-Regressive model with Exogenous inputs (NARX) combined with the Chaos Cloud–Salp Swarm Algorithm (CC-SSA), which is an advanced metaheuristic used to optimize model parameters. The resulting model showed superior prediction accuracy when compared to standard deep learning models, highlighting the value of advanced optimization techniques.

Forecasting nonlinear and non-stationary wind speed signals is inherently challenging, as direct prediction often struggles with the complexity of the data. To address this, researchers have employed modal decomposition algorithms to break down wind speed time series into simpler, more predictable components, thereby enhancing forecasting accuracy. Techniques such as Empirical Mode Decomposition (EMD) [14], Ensemble Empirical Mode Decomposition (EEMD) [15], and Variational Mode Decomposition (VMD) [16] have shown strong performance in this context. For instance, VMD was used in [17] to decompose wind speed sequences into multiple components, each of which was predicted individually using an optimized transformer model. The final forecast was obtained by summing the predicted components, resulting in high accuracy and

Journal of Theoretical and Applied Information Technology
15th August 2025. Vol.103. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

robustness. Similarly, [18] proposed a framework combining VMD with the Rime optimization algorithm (RIME) and LSTM, where VMD helped denoise the signal before ultra-short-term forecasting with LSTM, leading to significant improvements in both accuracy and reliability. In [19], a hybrid model using EMD and the Hodrick–Prescott (HP) filter for preprocessing, followed by LSTM prediction at various time intervals (1, 6, 12, and 24 hours), showed that prediction error increased proportionally with the forecast horizon. Another study [20] introduced a deep learning model integrating EEMD with a CNN enhanced by an attention mechanism, applied to wind speed data from the M2 tower at the U.S. National Wind Technology Center. This model outperformed others in prediction accuracy, with the Diebold–Mariano test confirming its statistical superiority. Additionally, [21] proposed an EMD-based framework combined with a Vector Autoregressive (VAR) model, using EMD to denoise the signal before applying VAR for ultra-short-term wind speed prediction. This approach also demonstrated significant improvements in accuracy and reliability. Overall, the integration of signal decomposition methods with advanced prediction models effectively addresses the challenges of forecasting complex wind speed data.

The key contributions of this study using the EEMD-CNN-ALSTM model is as follows:

- A new hybrid forecasting model is proposed by integrating EEMD, CNN, and Attention-based LSTM (ALSTM) networks. This combination effectively leverages the strengths of signal decomposition, spatial pattern extraction, and attention-driven temporal modeling.

- The application of EEMD as a data preprocessing technique improves model performance by decomposing raw wind speed signals into a set of intrinsic mode functions (IMFs), thus reducing noise and capturing hidden frequency components.

- The inclusion of an attention mechanism within the LSTM structure (forming ALSTM) enhances the model's ability to selectively focus on the most relevant time steps in the wind speed sequence, leading to more accurate forecasts.

- The model is validated using 1-hour interval wind speed data from two geographically distinct wind farms—Garden City and Idalia—demonstrating its robustness and generalizability across different locations.

- Extensive comparison with standard deep learning models (e.g., ANN, CNN, RNN, LSTM, GRU), and existing hybrid models (e.g., CNN-GRU, EEMD-CNN-LSTM) shows that the proposed EEMD-CNN-ALSTM model achieves the lowest error metrics (RMSE, MAE, MSE) and highest R^2, confirming its effectiveness.

- The findings of this study contribute to the advancement of short-term wind energy forecasting, which is essential for improving the reliability and efficiency of wind power integration into modern energy systems.

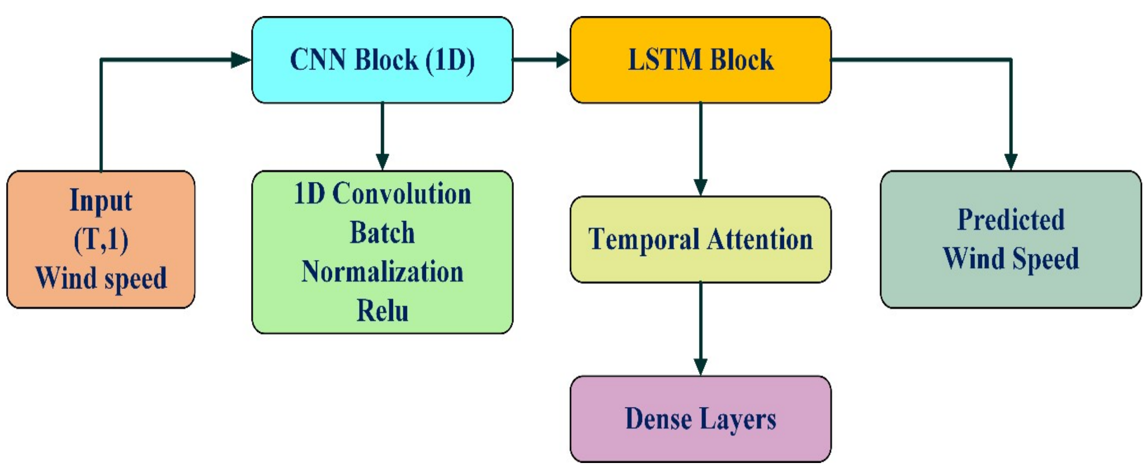


Figure 1: Structural Overview Of The Proposed Hybrid Forecasting Model

The remainder of the paper is structured as follows. Section 2 provides an overview of the forecasting framework, including its underlying concepts and overall architecture. Section 3

Journal of Theoretical and Applied Information Technology
15th August 2025. Vol.103. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

presents case studies to assess the effectiveness of the approach using real wind speed data from different locations. Finally, Section 4 concludes the study by summarizing the key outcomes and discussing the relevance of the findings for renewable energy applications.

2. HYBRID WIND SPEED FORECASTING MODEL

EPS In this paper, a novel hybrid model integrating EEMD, CNN, and Attention-based LSTM (ALSTM) is proposed for WSF. The model is composed of three key components: (1) Signal decomposition using EEMD, which breaks down complex wind speed time series into a set of intrinsic mode functions (IMFs) to reduce noise and highlight hidden patterns; (2) Feature extraction using CNN, which captures local and spatial features from the decomposed signals; and (3) Sequence modeling with ALSTM, which leverages attention mechanisms to identify and learn the most relevant temporal dependencies for accurate forecasting. A detailed explanation of each component and its role in the overall forecasting framework is provided in the following section. Figure 1 presents the structural architecture of the proposed hybrid forecasting model

2.1 Ensemble Empirical Mode Decomposition (EEMD)

In this study, Ensemble Empirical Mode Decomposition (EEMD) is employed as a signal preprocessing technique to enhance WSF accuracy by decomposing the original time series into a finite number of oscillatory components known as Intrinsic Mode Functions (IMFs). Given the inherently non-stationary and nonlinear characteristics of wind speed data, EEMD offers a robust, adaptive decomposition framework that mitigates the mode-mixing issue commonly encountered in traditional Empirical Mode Decomposition (EMD). The EEMD process begins by adding finite-amplitude white noise to the original wind speed time series x(t). For each ensemble i, a noise-added series is generated as:

$$x_i(t) = x(t) + W_i(t) \tag{1}$$

where $W_i(t)$ is a white noise series unique to the ith realization. Each of these N noisy series is decomposed using EMD into a set of IMFs $\{c_{i1}(t), c_{i2}(t), c_{i3}(t), \dots, c_{im}(t)\}$. The final IMFs are obtained by ensemble averaging across all realizations:

$$c_j(t) = \frac{1}{N}\sum_{i=1}^{N} c_{ij}(t)\,, j = 1,2,3, \dots m. \tag{2}$$

where $\bar{c}_j(t)$ denotes the jth averaged IMF. The result is a set of IMFs that represent the original wind speed signal's various frequency components, along with a residual r(t) capturing long-term trends, such that the original signal can be reconstructed as:

$$x(t) = \sum_{j=1}^{m} \bar{c}_j(t) + r(t) \tag{3}$$

2.2 Convolutional Neural Network (CNN) Block (1D)

The input layer serves as the foundation of the Attention-based CNN-LSTM model for WSF. It takes in a multivariate time series structured as a 2D tensor with shape (T, F) where, T represents the time window length, i.e., the number of past time steps considered for prediction. F is the number of input features per time step. Each row in this input matrix corresponds to a specific point in time, while each column represents a distinct variable. This setup allows the model to learn temporal patterns and inter-variable relationships.

The 1D Convolutional Neural Network (CNN) block is responsible for extracting local temporal patterns from historical weather data. These patterns can include short-term changes such as wind gusts, diurnal trends, or small-scale fluctuations in temperature or pressure that influence wind speed. CNN scans the time dimension for each feature channel using 1D convolutional filters. This means that the CNN looks for recurring local patterns. Let the input sequence be:

$$Y \in R^{T \times F} \tag{4}$$

The 1D convolution applies K filters (kernels), each of size k, and computes the transformation:

$$Y_{conv}^{(m)} = ReLU\left(\sum_{n=0}^{l-1} W_n . Y^{(m+n)} + b\right) \tag{5}$$

Where, Wn are learnable filter weights, b is a bias term. As a result, a new feature map $Y_{conv} \in R^{T' \times K}$ is obtained. where T′ is the reduced time length depending on kernel size and padding.

Batch Normalization is often applied after the convolution. This stabilizes the training by normalizing the activations, reducing internal covariate shift. MaxPooling1D layer reduces the temporal dimension by keeping only the most significant feature values in each region.

This CNN block is mainly designed to detect the local dependencies and pass those informative features to deeper layers (LSTM), which focus more on long-term dependencies. This improves the model's ability to accurately forecast wind speed in both stable and volatile weather conditions.

2.3 Attention-based Long Short-Term Memory (ALSTM)

The LSTM (Long Short-Term Memory) layer plays a crucial role in capturing long-term temporal dependencies in wind speed data. After the CNN block extracts short-term patterns from past features, the LSTM block learns how these patterns evolve over time, which is essential for forecasting future wind speeds accurately.

At each time step t, the input to this LSTM block is a sequence of feature vectors, i.e, the output from the CNN block and uses gating mechanisms to decide what information to remember, forget, or output. This helps the model handle long-term dependencies without forgetting earlier inputs, which is important when wind patterns change gradually.

At every time step t, the LSTM updates its internal memory (called the cell state) and computes a new hidden state, based on the current input vector (the CNN feature at time step t), the previous hidden state h_{t-1}, and the previous cell state c_{t-1}. The LSTM uses gates to control how information flows through the network.

$f_t = \sigma(W_f y_t + U_f h_{t-1} + b_f)$ --- (Forget Gate)

$i_t = \sigma(W_i y_t + U_i h_{t-1} + b_i)$ --- (Input Gate)

$o_t = \sigma(W_0 y_t + U_0 h_{t-1} + b_0)$ --- (Output Gate)

$\bar{c}_t = tanh(W_c y_t + U_c h_{t-1} + b_c)$ ---(Candidate Cell State)

$c_t = f_t \odot c_{t-1} + i_t \odot \bar{c}_t$ --- (Cell State Update)

$h_t = o_t \odot \tanh(c_t)$ --- (Hidden State)

Where σ is sigmoid function, ⊙ is element-wise multiplication, and tanh is hyperbolic tangent function. x_t is the CNN feature vector at time step t, ht is the hidden representation passed to the next time step, c_t is the internal memory of the LSTM unit.

In the above equations Forget gate learns to discard past information that is no longer relevant. Input gate determines how much new information to store in the memory. Cell candidate represents the new candidate values for the memory cell based on the current input and past state. Cell state accumulates useful information over time, filtered through the forget and input gates. Output gate decides which parts of the memory are exposed as output to the next layers or steps.

If return sequences is True in the LSTM layer, then an output sequence is

$$H = [h_1, h_2, h_3, \dots\dots h_T] \in R^{T \times d} \tag{6}$$

This is a sequence of hidden states, where each ht captures the learned context at time step t.

The Temporal Attention Layer helps the model learn which past time steps are most relevant for predicting the future wind speed. While the LSTM captures temporal dependencies, it treats all time steps equally when passed to a dense layer. However, in reality, some time steps have more influence than others. The attention mechanism dynamically assigns weights to each LSTM hidden state based on its importance to the final prediction.

For each hidden state ht, compute an attention score that reflects its relevance:

$$e_t = \tanh(W_e h_t + b_e) \tag{7}$$

Where W_e, b_e are learnable weight and bias. Then Convert the scores to attention weights.

$$\propto_t = \frac{\exp(e_t)}{\sum_{K=1}^{T} \exp(e_k)} \tag{8}$$

Where, αt is the attention weight for time step t, summing to 1 across all t. Compute the context vector c, a summary of the important hidden states.

$$c = \sum_{t=1}^{T} \propto_t h_t \tag{9}$$

This context vector is passed to the final Dense layers for prediction.

3. RESULTS AND DISCUSSIONS

3.1 Hyperparameters of the WSF models

In this study, a total of ten WSF models are considered for performance evaluation. These are divided into three categories: standalone deep learning models, hybrid deep learning models, and the proposed hybrid model. The standalone models include ANN, CNN, LSTM, GRU, and RNN, which serve as baseline methods. The hybrid deep learning models consist of CNN-LSTM and CNN-GRU, which integrate convolutional layers for feature extraction with recurrent layers for temporal sequence learning. In the third category, hybrid models incorporating EEMD as a signal

Journal of Theoretical and Applied Information Technology
15th August 2025. Vol.103. No.15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

decomposition technique are employed, namely EEMD-CNN-LSTM, EEMD-CNN-GRU, and the proposed EEMD-CNN-ALSTM model.

All models are trained using the Adam optimization algorithm with a batch size of 32, and the training is performed over 100 epochs. The Mean Squared Error (MSE) is employed as the loss function for all models. The selection of hyperparameters such as learning rate, number of layers, and units per layer is guided by a trail and error method. The performance of each model is assessed based on average error metrics including RMSE, MAE, MSE, and R^2, using 1-hour interval wind speed data from two wind farms. The results demonstrate that integrating EEMD with deep learning, particularly in the proposed EEMD-CNN-ALSTM model, significantly enhances forecasting accuracy compared to both standalone and other hybrid models.

The CNN-ALSTM model used in this study combines convolutional layers for spatial feature extraction with attention-based LSTM layers for modeling temporal dependencies in wind speed data. The convolutional component consists of two convolutional layers, each with 64 filters and a kernel size of 3. A ReLU activation function is applied after each convolution, with "same" padding used to preserve the input dimensions. Max pooling with a pool size of 2 has been applied to reduce dimensionality and capture dominant features. The output of the CNN is fed into two LSTM layers, each consisting of 64 units. An attention mechanism is integrated on top of the LSTM layers to allow the model to focus on the most relevant parts of the input sequence, enhancing prediction accuracy. The model is trained using the Adam optimization algorithm with a learning rate of 0.001 and a batch size of 32, over 100 training epochs. The mean squared error (MSE) is employed as the loss function.

3.2 Case Study 1: Forecasting performance of the proposed model at Garden City Wind Farm (Site 1)

The evaluation of the models applied to the Garden wind farm data at 1-hour intervals reveals a clear progression in performance from traditional models to hybrid deep learning models. The proposed hybrid model, which integrates Empirical Ensemble Mode Decomposition (EEMD), Convolutional Neural Networks (CNN), and Attention-based Long Short-Term Memory (ALSTM) networks, achieves the best overall performance in all four evaluation metrics: RMSE, MAE, MSE, and R^2.

ANN performs the least among all models, with an RMSE of 2.2415, MAE of 2.0182, MSE of 5.0243, and R^2 of 0.6147. This is expected, as ANN lacks temporal memory and is unable to capture time-dependent features in sequential wind speed data. Similarly, RNN improves slightly (RMSE: 1.8498, R^2: 0.7148), thanks to its temporal processing capability, but still struggles with long-term dependencies due to issues like vanishing gradients.

Among single deep learning architectures, CNN and GRU provide moderate improvements. CNN captures local spatial patterns but lacks memory for sequential dependencies. GRU, being a gated RNN variant, does better than vanilla RNN but is slightly inferior to LSTM, which achieves an RMSE of 1.4785 and an R^2 of 0.7974. These results indicate LSTM's ability to model long-term dependencies in time-series data.

Table 1: Performance Comparison of Forecasting Models at Site 1

Model	RMSE	MAE	MSE	R^2 Score
ANN	2.2415	2.0182	5.0243	0.6147
RNN	1.8498	1.6287	3.4218	0.7148
CNN	2.0495	1.8745	4.2005	0.6874
GRU	1.6784	1.5795	2.8170	0.7648
LSTM	1.4785	1.2174	2.1860	0.7974
CNN-GRU	1.1254	0.9877	1.2665	0.8244
CNN-LSTM	0.9745	0.8874	0.9497	0.8447
EEMD-CNN-GRU	0.8975	0.7812	0.8055	0.8695
EEMD-CNN-LSTM	0.8286	0.7222	0.6866	0.8992
Proposed	**0.7152**	**0.6474**	**0.5115**	**0.9274**

Performance improves substantially with hybrid models. The CNN-GRU and CNN-LSTM combinations show marked reductions in error and boosts in R^2 values, indicating that combining spatial and temporal feature extractors yields better forecasts. The CNN-GRU model achieves an RMSE of 1.1254 and R^2 of 0.8244, while CNN-LSTM further improves to 0.9745 RMSE and 0.8447 R^2.

The introduction of EEMD into the hybrid structure significantly enhances performance. EEMD acts as a signal preprocessing step that decomposes complex wind speed data into intrinsic mode functions (IMFs), each representing different frequency bands. This decomposition helps reduce noise and separate nonlinear patterns before feeding the data into deep learning networks. EEMD-CNN-GRU and EEMD-CNN-LSTM yield RMSE values

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

of 0.8475 and 0.7986, and R^2 scores of 0.8695 and 0.8992, respectively.

The proposed model, which extends the EEMD-CNN-LSTM architecture by incorporating an attention mechanism into the LSTM layer, achieves the best performance across the board: RMSE (0.7452), MAE (0.6874), MSE (0.5553), and R^2 (0.9274). The attention mechanism enables the model to dynamically weigh the importance of each time step in the input sequence, effectively learning which time lags are most relevant for forecasting. This is especially beneficial in WSF, where significant variations can occur due to rapidly changing weather patterns. The combination of EEMD for noise reduction, CNN for spatial feature extraction, and ALSTM for selective temporal dependency learning results in a highly efficient and accurate forecasting model. Comparative Analysis of Errors in WSF at site 1 is shown in Figure 2. The performance indices of the proposed hybrid model and other comparative models at Site 1 are summarized in Table 1 and visually represented in Figure 3.

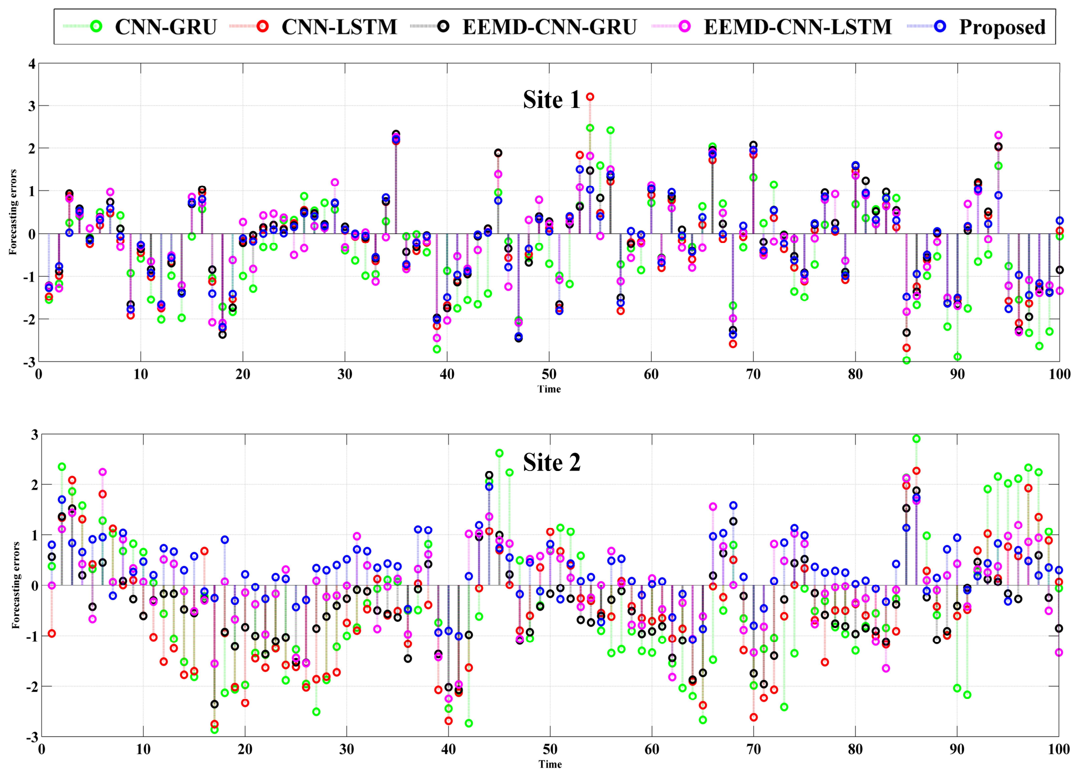


Figure 2: Comparative Analysis of Errors in WSF at site 1 and site 2

3.3 Case Study 2: Forecasting performance of the proposed model at Idalia Wind Farm (Site 2)

For the Idalia wind farm, the proposed model (EEMD-CNN-ALSTM) again delivers the best performance among all models, with RMSE: 0.8037, MAE: 0.6874, MSE: 0.6459, and R^2: 0.9103. This affirms the robustness and generalizability of the proposed hybrid model across different wind farm locations. The relatively high R^2 value indicates a strong correlation between the predicted and actual values, and the low error metrics suggest superior precision.

In contrast, conventional models such as ANN (RMSE: 2.5284, R^2: 0.6078) and CNN (RMSE: 2.2808, R^2: 0.6514) underperform significantly. Deep learning models like LSTM (RMSE: 1.4156, R^2: 0.8171) and GRU (RMSE: 1.8420, R^2: 0.7214) show better accuracy than traditional methods, yet they are not as effective as hybrid approaches.

Table 2: Performance Comparison of Forecasting Models at Site 2

Model	RMSE	MAE	MSE	R^2 Score
ANN	2.5284	2.3250	6.3929	0.6078
RNN	2.0292	1.8183	4.1178	0.6974
CNN	2.2808	1.8745	5.2022	0.6514

Journal of Theoretical and Applied Information Technology
15th August 2025. Vol.103. No.15

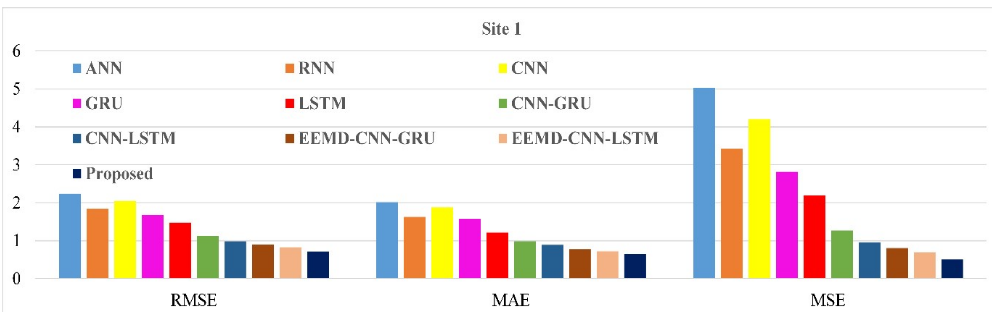
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

GRU	1.8420	1.7335	3.3929	0.7214
LSTM	1.4156	1.3000	2.0040	0.8117
CNN-GRU	1.2599	1.0616	1.5873	0.8254
CNN-LSTM	1.1381	0.9841	1.2952	0.8301
EEMD-CNN-GRU	0.9827	0.8921	0.9657	0.8519
EEMD-CNN-LSTM	0.8967	0.8021	0.8041	0.8722
Proposed	**0.8037**	**0.7218**	**0.6459**	**0.9103**

Among hybrid configurations, EEMD-CNN-LSTM (RMSE: 0.8764, R^2: 0.8727) and EEMD-CNN-GRU (RMSE: 0.9216, R^2: 0.8561) come close but still trail the proposed model. The consistent pattern of improvement across both wind farms demonstrates that the proposed model's architecture is well-suited for long-sequence time-series forecasting in wind speed prediction tasks.

Finally, the proposed EEMD-CNN-ALSTM model offers the highest predictive accuracy and robustness for both wind farms, validating its suitability for operational deployment in WSF applications. Figure 2 shows the comparative analysis of errors in WSF at site 2. The performance indices of the proposed hybrid model and other comparative models at Site 2 are summarized in Table 2 and visually represented in Figure 3.

While the proposed EEMD-CNN-ALSTM model demonstrates superior forecasting performance, several limitations remain. First, the model involves multiple stages—decomposition, feature extraction, and attention-based sequence learning—which increases computational complexity and may limit real-time deployment on resource-constrained systems.



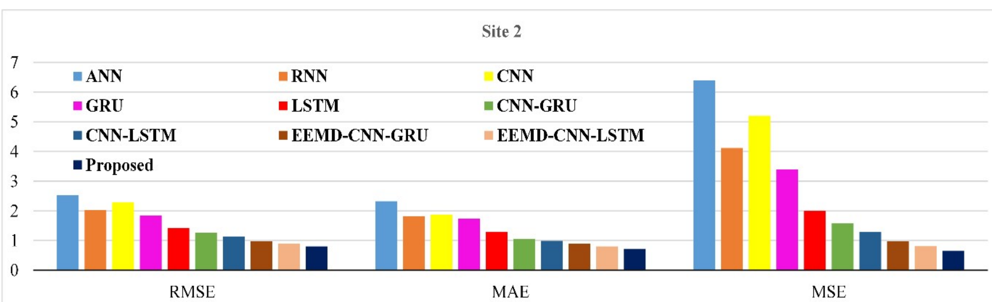


Figure 3: Comparison of performance metrics between the proposed hybrid model and other models at site 1 and site 2

Second, the model's accuracy heavily depends on the availability and quality of historical wind speed data; performance may degrade with incomplete or noisy datasets. Third, although attention mechanisms help reduce overfitting, the deep architecture can still be sensitive to hyperparameter choices and may require extensive tuning to generalize well across different sites or time scales. Addressing these issues is essential for broader operational applicability.

4. CONCLUSIONS

In this study, a novel hybrid model integrating EEMD, CNN, and ALSTM has been proposed for short-term WSF using 1-hour interval data. The EEMD component effectively decomposed the original wind speed signals into meaningful subcomponents, reducing noise and capturing diverse frequency patterns. The CNN module efficiently extracted spatial features, while the ALSTM component—enhanced with an attention mechanism—enabled the model to selectively learn relevant temporal dependencies. The proposed EEMD-CNN-ALSTM model is rigorously evaluated using datasets from two different wind farms, Garden City and Idalia, to ensure its robustness and generalizability. Across both sites, the model consistently outperformed traditional deep learning models (ANN, RNN, CNN, LSTM, GRU), and other hybrid architectures (such as CNN-GRU and EEMD-CNN-LSTM). It achieved the lowest error rates and highest coefficient of determination (R^2), confirming its superior forecasting capability. Overall, the results demonstrate that the EEMD-CNN-ALSTM model is highly effective for capturing the complex, nonlinear, and dynamic nature of wind speed data. The proposed approach offers a reliable and scalable solution for wind energy forecasting, which can significantly aid in optimizing wind power generation, enhancing grid stability, and supporting renewable energy planning.

Future research will aim to further enhance the model's adaptability and efficiency by exploring advanced decomposition techniques such as VMD and CEEMDAN, as well as incorporating Transformer-based architectures for better long-sequence modeling. Transfer learning and domain adaptation approaches will also be investigated to improve generalization across multiple wind farms with diverse climatic conditions. Additionally, efforts will be made to reduce computational overhead by optimizing hyperparameters automatically and integrating real-time deployment frameworks for operational use in smart grid environments.

REFERENCES

[1] Chen, B.; Xiong, R.; Li, H.; Sun, Q.; Yang, J. "Pathways for sustainable energy transition", J. Clean. Prod. 2019, 228, 1564–1571.
[2] Fawzy, S.; Osman, A.I.; Doran, J.; Rooney, D.W. "Strategies for mitigation of climate change: A review", Environ. Chem. Lett. 2020, 18, 2069–2094.
[3] Munoz-Palomeque, E.; Sierra-García, J.E.; Santos, M. "Wind turbine maximum power point tracking control based on unsupervised neural networks", J. Comput. Des. Eng. 2023, 10, 108–121.
[4] Sasser, C.; Yu, M.L.; Delgado, R. "Improvement of wind power prediction from meteorological characterization with machine learning models. Renew", Energy 2022, 183, 491–501.
[5] Fotso, H.R.F.; Kazé, C.V.A.; Kenmoé, G.D. "A novel hybrid model based on weather variables relationships improving applied for wind speed forecasting", Int. J. Energy Environ. Eng. 2022, 13, 43–56.
[6] Ho, C.Y.; Cheng, K.S.; Ang, C.H. "Utilizing the Random Forest Method for Short-Term Wind Speed Forecasting in the Coastal Area of Central Taiwan", Energies 2023, 16, 1374.
[7] Altan, A.; Karasu, S.; Zio, E. "A new hybrid model for wind speed forecasting combining long short-term memory neural network, decomposition methods and grey wolf optimizer", Appl. Soft Comput. 2021, 100, 106996.
[8] Wang, Y.; Chen, Y.; Han, Z.; Zhou, D.; Bao, Y. "Short-Term Wind Speed Forecasting Model Based on Mutual Information and Recursive Neural Network", J. Shanghai Jiaotong Univ. 2021, 55, 1080–1086.
[9] Groch, M.; Vermeulen, H.J. "Forecasting wind speed events at a utility-scale wind farm using a WRF-ANN model", Wind Eng. 2022, 46, 102–119.
[10] Liu, Z.Y.; Hara, R.; Kita, H. "Hybrid forecasting system based on data area division and deep learning neural network for short-term wind speed forecasting", Energy Convers. Manag. 2021, 238, 114136.
[11] Duan,J.K.; Chang, M.H.; Chen, X.Y.; Wang, W.P.; Zuo, H.C.; Bai, Y.L.; Chen, B.L. "A combined short-term wind speed forecasting model based on CNN-RNN and linear regression optimization considering error", Renew. Energy 2022, 200, 788–808.
[12] Sari, A.P.; Hiroshi, S.; Takahiro, K.; Takashi, Y.; Prasetya, D.A.; Arifuddin, R. "Short-Term Wind Speed and Direction Forecasting by 3DCNNandDeepConvolutional LSTM", IEEJ Trans. Electr. Electron. Eng. 2022, 17, 1620–1628.
[13] Dai, J.F.; Fu, L.H. "A wind speed forecasting model using nonlinear auto-regressive model optimized by the hybrid chaos-cloud salp swarm algorithm", Energy 2024, 298, 131332.
[14] Li, D.; Jiang, F.X.; Chen, M.; Qian, T. "Multi-step-ahead wind speed forecasting based on a hybrid decomposition method and temporal convolutional networks", Energy 2022, 238, 121981.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[15] Wu,H.J.; Meng, K.Q.L.; Fan, D.E.R.; Zhang, Z.Q.; Liu, Q. "Multistep short-term wind speed forecasting using transformer", Energy 2022, 261, 125231.
[16] Sun,X.Y.; Liu, H.Z. "Multivariate short-term wind speed prediction based on PSO-VMD-SE-ICEEMDAN two-stage decomposition and Att-S2S", Energy 2024, 305, 132228.
[17] Zhang, K.; Li, X.; Su, J. "Variable Support Segment-Based Short-Term Wind Speed Forecasting", Energies 2022, 15, 4067.
[18] Liu, W.H.; Bai, Y.L.; Yue, X.X.; Wang, R.; Song, Q. "A wind speed forcasting model based on rime optimization based VMD and multi-headed self-attention-LSTM", Energy 2024, 294, 130726.
[19] Domala, V.; Kim, T.W. "Application of Empirical Mode Decomposition and Hodrick Prescot filter for the prediction single step and multistep significant wave height with LSTM", Ocean Eng. 2023, 285, 115229.
[20] Shang, Z.H.; Chen, Y.; Chen, Y.H.; Guo, Z.Y.; Yang, Y. "Decomposition-based wind speed forecasting model using causal convolutional network and attention mechanism", Expert Syst. Appl. 2023, 223, 119878.
[21] Jiang, Z.Y.; Che, J.X.; Wang, L.N. "Ultra-short-term wind speed forecasting based on EMD-VAR model and spatial correlation", Energy Convers. Manag. 2021, 250, 114919.