Nouf J. Al Qahtani
Control & Instrumentation Engineering Department
King Fahd University of Petroleum & Minerals, Dhahran, 31261
Saudi Arabia

Mohamed S. Abolouz
Electrical Engineering Department
King Fahd University of Petroleum & Minerals, Dhahran, 31261
Saudi Arabia

Mohamed A. Mohandes
Electrical Engineering Department
Interdisciplinary Research Center for Sustainable Energy Systems
King Fahd University of Petroleum & Minerals, Dhahran, 31261
Saudi Arabia

Shafiqur Rehman
Interdisciplinary Research Center for Sustainable Energy Systems
King Fahd University of Petroleum & Minerals, Dhahran, 31261
Saudi Arabia

Comparative Evaluation of Recurrent and Attention-Based Deep Learning Models for Wind Speed Forecasting

Wind speed forecasting is critical for enhancing wind energy output and improving efficiency. Although traditional statistical and physical models are commonly used but fail to capture the complex, non-linear, and time-varying nature of meteorological data. Therefore, this study explores four deep learning architectures: Recurrent Neural Network (RNN), Gated Recurrent Unit (GRU), Long Short-Term Memory (LSTM), and transformer. These models are analyzed using multivariate and hourly meteorological data from the Al-Ahsa region over 24 years (2001-2024). The dataset includes five features: wind speed at 10 meters above ground (used as a reference), air temperature at 2 meters, specific humidity at 2 meters, wind direction at 10 meters, and surface pressure. The models were trained on 80% of the dataset, validated on 10%, and tested on the remaining 10%. The GRU model achieved the best performance with RMSE = 0.3126 m/s and R² = 0.9759.

Keywords:*Deep learning architectures; wind speed, short-term prediction; Recurrent Neural Network (RNN); Gated Recurrent Unit (GRU); Long Short-Term Memory (LSTM); and Transformer*

1. INTRODUCTION

Traditionally, energy generation has relied heavily on electricity produced from fossil fuels or other conventional sources, where most systems require either fuel or grid-based electricity to operate. However, with rising energy consumption and the associated environmental impacts, there is a growing need to adopt more sustainable, low-emission energy sources for electricity and energy production [1-4]. As a response to the growing environmental concerns linked to fossil fuel combustion, many countries are transitioning toward renewable and sustainable energy sources. Kingdom of Saudi Arabia, in particular, has outlined clear national targets to annually reduce carbon emissions by 278 million tons and increase the renewables share by 50% of total electricity generation by 2030. These efforts align with the country's objective of achieving net-zero carbon emissions by 2060 [5].

Wind energy is a renewable and environmentally acceptable source of power that is an alternative to conventional fossil fuel-based energy systems. The main advantage of wind energy is its ability to generate electricity with minimal greenhouse gas emissions or other pollutants[6-10]. Thus, it is an eco-friendly system and an essential component of global decarbonization efforts to mitigate climate change. However, it has some drawbacks as well, including intermittency, noise pol–lution, and the environmental costs of turbine produc–tion and disposal, all of which limit its widespread adoption.

The working principle of wind energy generation involves converting wind kinetic energy into electrical using wind turbines. The wind flow causes blades to rotate and generate power through gear box and the generator. Since, wind is an inexhaustible and widely available resource, it enables countries to diversify their energy portfolio toenhance energy security and reduce dependence on fossil fuels [8]. However, wind power deployment faces challenges due to its unpredictable and variable characteristics, which can undermine system reliability and power quality while increasing wind penetration levels.

The electrical power generated by a wind turbine is directly dependent on the cube of wind speed, making it a critical in the production process. It indicates that small changes in wind speed may lead to significant variations in output power. Thus, accurate prediction of wind speed is crucial for optimizing turbine operation, scheduling energy supply, managing grid stability, and integrating wind power effectively into the electricity system [11]. These predictions can significantly reduce wind curtailments, helping to lower operational costs while increasing revenue in electricity market operations. However, accurate forecasting is affected by the inherent unpredictability and variability of the wind, making it difficult at times. Accordingly, many studies have attempted to develop and refine advanced forecasting techniques to better address the complexities of wind speed and power prediction [11-18].

Based on the existing literature, a wide spectrum of prediction methods has been used. These methods range from simple to complex approaches, each differing in characteristics and performance. In general, there are two types of forecasting: short-term and long-term, each with its own advantages. For example, short-term forecasts play a critical role in power system operations, particularly in day-ahead predictions, as they are

Received: August 2025, Accepted: October 2025
Correspondence to: Dr Shafiqur Rehman, IRC-SES, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran-31261, Saudi Arabia
E-mail: srehman@kfupm.edu.sa
doi: 10.5937/fme2504693A

essential for unit commitment, load balancing, and scheduling [19-22]. On the other hand, long-term fore–casts support strategic planning by anticipating seasonal trends and optimizing resource allocation; thus, redu–cing reliance on reserve capacities [23]. In the recent times, machine learning (ML), artificial neural networks (ANNs) have emerged as some of the effective models for wind speed prediction. This is attributed to their non-linear learning capabilities, which can reduce wind power grid costs and support the selection of optimal reserve capacities at various time intervals [24-26].

Recurrent neural networks (RNNs) are used for sequential and time-series data as they retain a hidden state that evolves at each time step to capture temporal dependencies. Thus, RNNs are efficient for time-series wind speed predictions. However, standard RNNs often struggle to preserve information over long sequences due to the vanishing gradient problem [21]. To address this, gated variants, namely gated recurrent units (GRU) and long-short-term memory (LSTM) networks, incor–porate learnable gating mechanisms that selectively retain and update relevant context over much longer horizons [22-26]. As a result, GRUs and LSTMs are particularly well suited for wind speed forecasting, where current values are heavily influenced by past behavior. In the present study, we implement and com–pare three architectures: simple RNN, GRU, and LSTM to evaluate their effectiveness in predicting 24-hour wind speeds.

Being able to retain hidden state of the sequential and time-series data, RNNs are used to capture temporal dependencies. Thus, RNNs are efficient for time-series wind speed predictions. However, standard RNNs often struggle to preserve information over long sequences due to vanishing gradient problem [27]. To address it, gated recurrent units (GRU) and long-short-term memory (LSTM) networks are used to retain and update relevant context over much longer horizons [28-32]. As a result, GRUs and LSTMs are particularly well suited for wind speed forecasting, where current values are heavily influenced by past behavior. In the present study, we implement and compare three architectures: simple RNN, GRU, and LSTM to evaluate their effectiveness in predicting 24-hour wind speeds.

Recent advancements in wind speed forecasting have focused on addressing key challenges related to the nonlinearity, uncertainty, and volatility inherent in wind data [33]. Lv et al. [16] proposed a dynamic adaptive interval prediction model that integrates fuzzy infor–mation granulation with a dynamically adjusted time window and least squares support vector machine. This method was designed to enhance the accuracy of pre–diction point and reliability of intervals. By effectively capturing the non-stationary behavior of wind speed time series, their approach demonstrated superior performance across various interval evaluation metrics at confidence levels of 0.90 and 0.95. Both Maruthi et al. [20] and Amirteimoury et al. [24] investigated the challenge of wind speed variability with advanced hyb–rid and ensemble techniques designed to enhance fore–casting accuracy and grid integration. While Amirtei–moury et al. developed a hybrid model combining Discrete Wavelet Transform (DWT) for signal smoot–hing, Mutual Information (MI) for time-series compo–nents extraction, the Coot Optimization Algorithm (COOT) for optimal feature selection, and a Bidi–rectional LSTM network for temporal modeling.Maruthi et al. [20] introduced a multi-model integration frame–work (MIDF) based on the ensemble of DeepAR and Temporal Fusion Transformer (TFT). These studies attained coefficient of determination above 0.89.

Unlike the prior studies, Xiao et al. [25] prioritized data quality and noise reduction as the foundation for accurate wind speed forecasting. Thus, they developed a hybrid approach combining Weighted Principal Com–ponent Analysis (WPCA) and a Particle Swarm Optimi–zation-tuned Gated Recurrent Unit (PSO-GRU) net–work. WPCA was used to extract the most informative features while mitigating data noise, and PSO was employed to optimize the GRU model's hyperpa–rameters. This hybrid architecture achieved significant reductions in MAE and RMSE by 5.3% to 16% and improved R^2 scores by 2.1% to 3.1% compared to conventional models. Similarly, Du et al. [26]proposed framework integrates Variational Mode Decomposition (VMD) for effective signal denoising, Runge–Kutta Optimization (RUN) for fine-tuning decomposition parameters, and a Sequence-to-Sequence model with an Attention mechanism (Seq2Seq-Attention) for multi-step forecasting. The study demonstrated substantial improvements in forecast accuracy, with correlation coefficients exceeding 0.9 across 1-to-12-hour predic–tion horizons and up to a 21% increase in predictive performance.

Despite significant advancements, considerable cha–llenges remain in pre-processing noisy wind speed data, selecting the most informative features, and balancing prediction accuracy with computational efficiency. Mo–reover, current models often lack the capacity to dyna–mically adapt to changing atmospheric conditions or to seamlessly integrate domain knowledge. Recent lite–rature emphasizes that continued progress in wind speed forecasting depends on the hybridization of advanced deep learning architectures with robust data processing techniques and adaptive learning capabilities. In res–ponse to these gaps, the main objective of the present study is to systematically evaluate and compare the per–formance of four prominent models, which are RNN, GRU, LSTM, and transformer, in forecasting wind spe–ed using real-world meteorological datasets, focusing on their accuracy, robustness, and practical applicability.

2. METHODOLOGY

This study employed a detailed analysis to investigate the application of neural network architectures in wind prediction. The models are selected to predict wind speed using a dataset obtained from the NASA POWER database (https://power.larc.nasa.gov/). NASA POWER provides high-quality meteorological data from satellite observations and weather models, with global coverage that supports diverse applications. Additionally, the da–taset spans a wide range of climatic conditions, inclu–ding remote and extreme environments, which can int–roduce complexities and affect the accuracy of wind pre–dictions in specific regions. Once the dataset was obta–

ined, it was pre-processed, and features were extracted. These features were then used in the neural network models to predict wind speed. Figure 1 summarizes the overall methodological framework for this study.

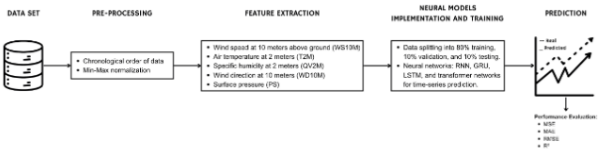

Figure 1. Methodological framework: Acquired data are pre-processed, features are extracted, and then used for prediction

2.1 Dataset

Al-Ahsa region, located in the eastern part of the Kingdom of Saudi Arabia, is widely recognized for its desert climate observing high temperatures. During the summer months, the area experiences average daily maximum temperatures that typically range between 44°C and 46°C, making it one of the hottest regions in the country [34]. These climatic conditions pose significant challenges for various sectors, including agriculture, energy, and urban planning, which rely heavily on accurate meteorological data for decision-making and resource management. The dataset used in this investigation contains meteorological measurements collected hourly from this region over 24 years (from 2001 to 2024), resulting in 210,360 data points.

2.2 Pre-processing of the Data

Pre-processing is a crucial step to ensure data quality and improve model performance. In Python software (version 3.11), the raw data underwent several pre-processing steps, like handling missing values, normalizing it, and selecting relevant features for wind speed prediction. After the dataset was loaded, the time-related columns (Year, MO, DY, and HR) were renamed as year, month, day, and hour, respectively. Then, a new timestamp column was created by combining these time-related columns into a single datetime column to ensure chronological order. This was essential for time-series analysis to sort the data based on the newly created timestamp. Afterward, the data was sorted based on the “TIMESTAMP” to maintain the correct sequence of observations. A key aspect of this pre-processing was handling missing values (if exists), though in this dataset, no explicit missing values were found. If any gaps were detected, interpolation techniques could have been used to fill missing time-series data based on the surrounding points. The dataset was then filtered to include only the relevant features. One of the most significant pre-processing steps involved normalizing the features using a Min-Max scaler.

The min-max scaler was applied to range between 0 and 1. This step is important for improving the performance of machine learning models, which are sensitive to the magnitude of input values. Without this scaling, features with larger ranges could dominate the learning process, leading to suboptimal model performance. By normalizing the features, the data was made more uniform and ready for model training. Finally, the scaled features were combined with the TIMESTAMP column.

2.3 Feature Extraction

Once the data was pre-processed, the features in the dataset were used to facilitate the prediction process. The features were selected based on their established significance in meteorological modeling and their direct influence on the accuracy of wind speed prediction. The features are illustrated in Table 1. The wind speed (WS10M) and the wind direction (WD10M) are measured at 10 meters above the ground (used as a reference). The air temperature (T2M) and the specific humidity (QV2M) are measured at 2 meters above the ground. The wind direction at 10 meters indicates the direction from which the wind is blowing at a height of 10 meters above the ground. To account for its circular nature, the WD10M was transformed using its sine and cosine components.

Wind speed and other meteorological variables can vary significantly at different altitudes due to factors such as atmospheric stability. By using data from multiple heights, a more comprehensive representation of the atmospheric conditions influencing wind patterns at various levels can be obtained. Moreover, the dataset does not offer measurements at all possible heights or a flexible selection of heights. Other combinations of features could have been explored. For example, features such as relative humidity, solar radiation, or cloud cover could be included to capture other atmospheric variables that influence wind patterns and speed. Another alternative would be using forecasted weather data or seasonal variations to improve model accuracy, especially in regions with strong seasonal wind variations.

Table 1. Relevant features used for accurate wind speed prediction

Feature	Description	Units
WS10M	Wind speed at 10 meters above ground	meters/second (m/s)
T2M	Air temperature at 2 meters	degrees Celsius (°C)
QV2M	Specific humidity at 2 meters	grams/kg (g/kg)
WD10M	Wind direction at 10 meters	degrees (°)
PS	Surface pressure	kilopascals (kPa)

2.4 Neural Models Implementation and Training

The pre-processed data was divided into three segments: training (80%), validation (10%), and testing (10%). Based on the data split strategy, the training set covers a total of 168,290 hourly samples, while the validation and test sets contain 21,034 and 21,036 hourly samples, respectively. The 80/10/10 configuration offered the best balance between training data and model evaluation. Following such a temporal split ensures that the model is tested on unseen data, which is important to simulate a real-world forecasting scenario. Also, it helps to eliminate data leakage that occurs in random splits.

Four neural models were used in this study, which are RNN, GRU, LSTM, and Transformer networks.

These models were implemented and trained on the pre-processed data to predict wind speed. All models followed a sequence-to-one regression design, where a 24-hour sliding window of past data was used to predict wind speed for the following hour. This structure reflects practical forecasting applications and enables the models to learn from recent temporal patterns. For the initial setup, each model began with a single recurrent or encoder layer, followed by a dense layer and a final output node for regression tasks. Once the baseline architecture was established, various parameters, such as the number of units, the addition of a second recurrent layer, dropout rates, and learning rate, were fine-tuned to improve performance. All models were trained for 30 epochs, providing enough iterations for convergence without overfitting. This number was chosen after considering the trade-off between adequate learning and model generalization, though techniques like early stopping could further refine training in future experiments.

The optimization parameters were determined through a process of hyperparameter tuning, where multiple combinations of model parameters were tested to identify the best performing configuration. The tuning process involved adjusting key parameters such as dropout rate and learning rate, among others. The goal was to maximize model performance, ensuring that it learned from the data effectively while avoiding overfitting. For example, the RNN model was optimized with 192 units in the first recurrent layer, and a dropout rate of 0.2, which helped prevent overfitting by randomly dropping some neurons during training. The GRU was configured with 128 units in its first layer and similarly had a 0.2 dropout rate. The LSTM model, which is known for its ability to handle long-range dependencies, was optimized with 160 units in the first recurrent layer and a 0.1 dropout rate, emphasizing its ability to learn from complex sequences while minimizing overfitting. For the Transformer model, a more advanced architecture, the hyperparameters were adjusted to incorporate 3 attention heads in the encoder layers, which helps the model focus on different aspects of the input sequence at once. The learnable positional encoding was included to help the Transformer model understand the order of sequence elements, a feature crucial for handling time-series data. Each model was optimized using the Adam optimizer, which is widely used for training deep learning models due to its efficient gradient descent approach. A summary of the optimization parameters can be found in Table 2.

Evaluating the neural network models is a crucial step for performance of the trained models. Specifically, the models were assessed based on their ability to accurately forecast wind speed values using the test dataset. Four key statistical measures Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2) were employed. To illustrate, MSE and RMSE are more sensitive to large prediction deviations, making them useful for evaluating models. MAE, on the other hand, provides average absolute difference between predicted and actual values. The R^2 score measures how well the predicted and true values match each other. By analyzing these metrics, the effectiveness and reliability of the trained models were systematically evaluated, allowing for a robust comparison between different architectures and configurations.

Table 2. Comparison of optimized hyperparameters across all evaluated models

Parameter	RNN	GRU	LSTM	Transformer
Recurrent Layer 1	192 units	128 units	160 units	3 heads, 32-dim (Attention)
Recurrent Layer 2	32 units	128 units	96 units	3 heads, 32-dim (Attention)
Dense Layer Units	96	64	96	64
Dropout Rate	0.2	0.2	0.1	0.3
Learning Rate	1.888×10^{-4}	4.692×10^{-4}	1.661×10^{-4}	11.99×10^{-4}
Total Trainable Parameters	48,481	159,233	214,337	122,625
Sequence Length	24	24	24	24
Architecture Depth	2layers	2 layers	2layers	2 encoder blocks
Optimizer	Adam	Adam	Adam	Adam
Positional Encoding	-	-	-	Learnable embeddings

2.5 SHAP and Ablation Analyses

SHAP (SHapley Additive exPlanations) and ablation analyses are widely used to interpret machine learning models and assess feature importance. SHAP technique measures the contribution of each input parameter to demonstrate how individual features influence the output. SHAP technique measures the contribution of each input feature to the model's prediction to demonstrate how individual features influence the output, while the ablation technique is the removal of individual features or groups of features from the trained model and observing the resulting change in performance of the model. These methods are highly useful in analyzing model behavior, diagnosing potential biases, and selecting relative features. Therefore, SHAP analysis was conducted on all features across the RNN, GRU, LSTM, and transformer models to evaluate their relative importance. SHAP is reliable because it ensures fair, consistent explanations by considering all possible feature combinations, providing both local and global interpretability. The SHAP library is a widely used Python tool for efficiently computing Shapley values and offering visualizations like summary plots, force plots, and dependence plots. Similarly, in the ablation analysis, each feature (T2M, QV2M, WD10M, and PS) was individually removed from the input set to assess its specific impact on model accuracy. For each ablation experiment, the corresponding feature was excluded, the models were retrained using the remaining features, and performance metrics were recorded. This process was repeated for each feature to evaluate its contribution to the overall model performance.

3. RESULTS

3.1 Dataset Pre-processing

The pre-processed dataset was thoroughly analyzed, where the main features used in this study were T2M, QV2M, WD10M, and PS. The WS10M was used as a reference to compare the predicted wind speed with true values. The boxplot of the normalized features is pre–sented in Figure 2 to analyze distribution of features. The black dots in Figure 2 represent outliers for WS10M and QV2M, indicating data points significantly higher or lower than the rest. They are shown only for these two variables because the other parameters (T2M, WD10M, and PS) do not have extreme values outside the defined outlier range.

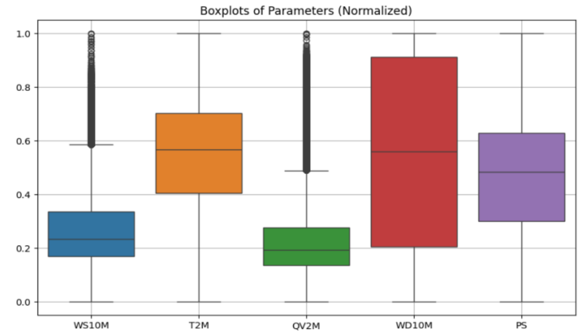

Figure 2. Boxplots of normalized meteorological parame–ters: WS10M, T2M, QV2M, WD10M, and PS, showing the distribution, variability, and presence of outliers in the dataset

3.2 Feature Extraction

The features used in this investigation were chosen depending on the relevance to the prediction task. Table 3 presents summary statistics for each feature. Among all, T2M shows the highest mean value of 0.561 and a notably high median of 0.568, indicating consistently significant values.

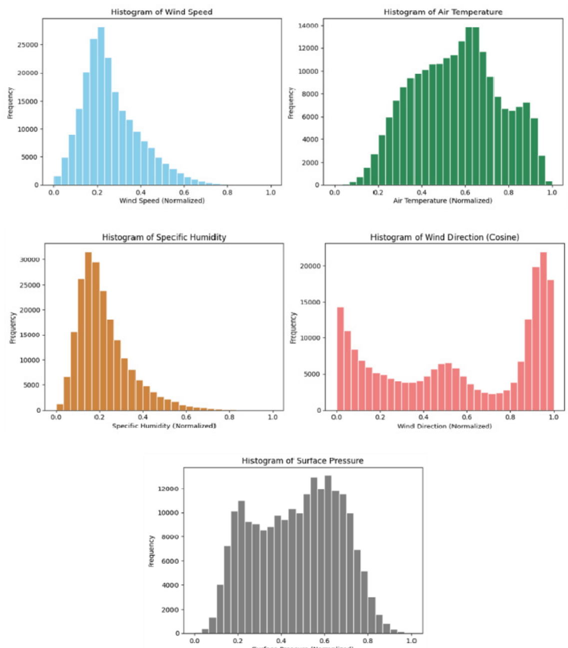

Figure 3. Histogram distributions of meteorological parameters illustrating the frequency and variability of observed values

WD10M also exhibits a high mean of 0.555, but with a much larger standard deviation of 0.350, reflecting substantial variability. In contrast, WS10M and QV2M show lower mean values of 0.262 and 0.222, respectively. PS has a moderate mean of 0.470 and low standard deviation of 0.746, suggesting relatively stable values across the dataset. Moreover, the histogram distribution of each feature, shown in Figure 3, shows the frequency and spread of each feature's values to identify patterns, skewness, and potential outliers in the data. In Table 3, all the normalized values are reported.

Table 3. Statistics for each feature, including mean, median, and standard deviation

Feature	Mean	Median	Standard Deviation
WS10M	0.262	0.235	0.133
T2M	0.561	0.568	0.199
QV2M	0.222	0.193	0.125
WD10M	0.555	0.561	0.350
PS	0.470	0.484	0.197

3.3 Neural Models Implementation and Training

3.3.1 RNN Model

The performance of the RNN model was evaluated using the metrics, provided earlier. The performance matricsrefers to the specific performance indicators are used to evaluate the effectiveness of the various models employed for wind prediction.

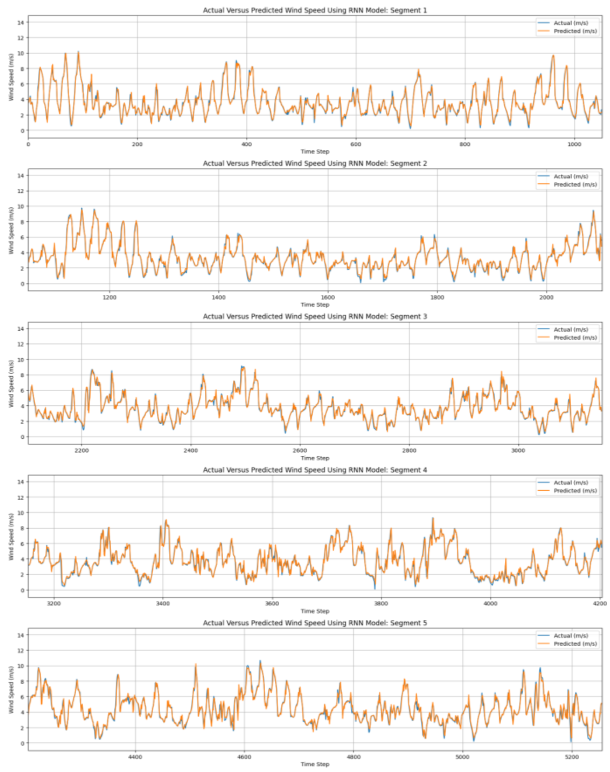

Figure 4. Comparison between the actual and predicted values for multiple time segments. The plots illustrate the performance of RNN model in capturing temporal variati–ons and short-term fluctuations in wind speed, highlighting the accuracy and potential discrepancies between observed and predicted values

These metrics are Mean Squared Error (MSE), Mean Absolute Error (MAE), Root Mean Squared Error

(RMSE), and the Coefficient of Determination (R^2). It provides a quantitative measure of how well the model predicts the wind speed or direction compared to actual observed values.Figure 4 shows a set of segments of actual versus predicted WS10M. It is noticeable that the predicted values have the same trend as the actual va–lues. Based on Table 4, MSE value is 0.1191m/s bet–ween predicted and actual values. The MAE is 0.2191m/s, reflecting the average magnitude of predic–tion errors regardless of direction. The RMSE, which provides an error measure in the original units, is 0.3452m/s, highlighting the typical prediction deviation. Finally, the R^2 is 0.9707, demonstrating that the model explains approximately 97.07% of the variance in the observed data, indicating a strong predictive capability.

Table 4. Performance metrics of the RNN model for wind speed prediction using a random seed of 42

Metrics of RNN Model			
MSE (m/s)	MAE (m/s)	RMSE (m/s)	R^2
0.1191	0.2191	0.3452	0.9707

3.3.2 GRU Model

The proposed GRU model was evaluated using the error metrics explained above. The predicted values are almost identical to actual values of WS10M as shown in Figure 5.

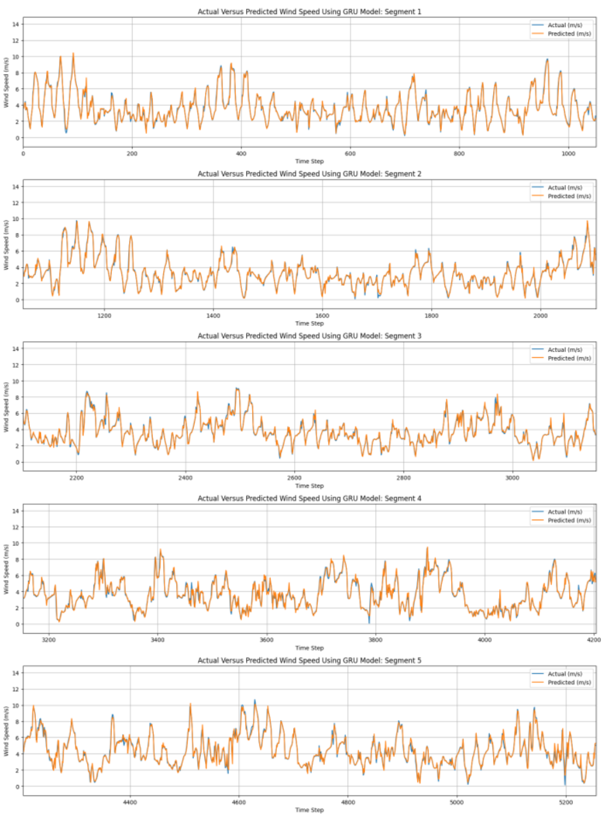

Figure 5. Comparison between the actual and predicted values across multiple time segments. The plots illustrate the performance of GRU model in capturing temporal variations and short-term fluctuations in wind speed

Multiple examples of comparison between the pre–dicted and the actual values of the wind speed over different time pans show an excellent trend matching, Figure 5.The R^2 value of 0.9759 showed an excellent match between the predicted and the actual values, Table 5. Lower values of RMSE, MAE, and MSE of 0.3126 m/s, 0.1928 m/s, and 0.0977; further strengthen the good performance of the GRU model.

Table 5. Performance metrics of the GRU model for wind speed prediction using a random seed of 42

Metrics of GRU Model			
MSE (m/s)	MAE (m/s)	RMSE (m/s)	R^2
0.0977	0.1928	0.3126	0.9759

3.3.3 LSTM Model

The LSTM model showed further better performance where the difference between predicted and actual values is concerned, as illustrated in Figure 6. This model achieved an MSE of 0.1120m/s, MAE of 0.2062m/s, and RMSE of 0.3347m/s, as summarized in Table 6.The model's R^2 value suggests that appro–ximately 97.24% of the variance in the observed data is explained by the model.

3.3.4 Transformer Model

The transformer model also showed a good perfo–rmance, where predicted values were close to actual values, as shown in Figure 7. The R^2 value of 0.9402 indicates a strong correlation between the predicted and observed data, Table 7.

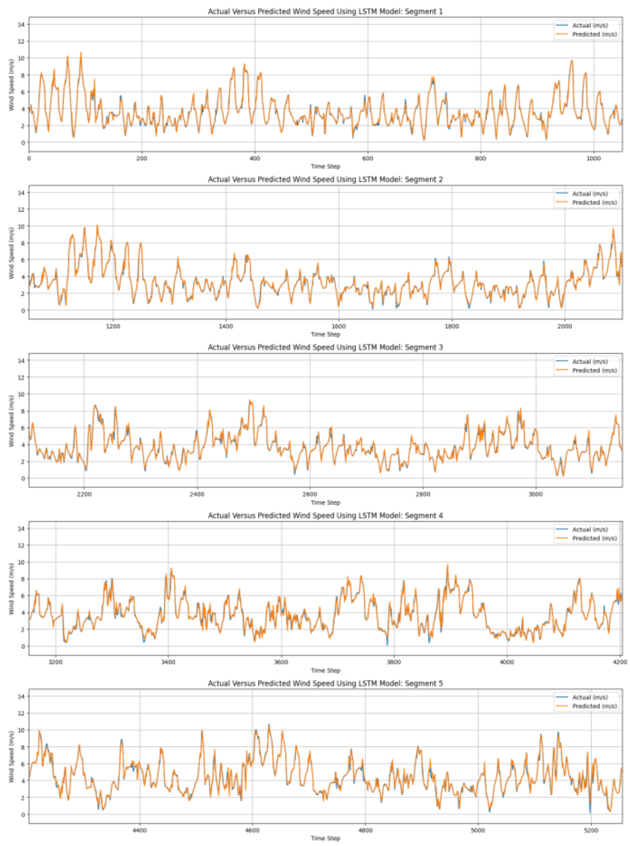

Figure 6. Comparison of actual versus predicted wind speed across multiple time segments. The plots illustrate the performance of LSTM model in capturing temporal variations and short-term fluctuations in wind speed, highlighting the accuracy and potential discrepancies between observed and predicted values

Table 6. Error values for LSTM model for wind speed prediction using a random seed of 42

Metrics of LSTM Model			
MSE (m/s)	MAE (m/s)	RMSE (m/s)	R^2
0.1120	0.2062	0.3347	0.9724

3.3.5 Comparison Across Models

To compare the performance of different predictive models, a zoomed-in segment is plotted for visual inspection, where both actual and predicted wind speeds are compared for each model. The zoomed-in segment helps identify discrepancies, trends, and overall model reliability, as shown in Figure 8. The plots show the comparison of actual versus predicted wind speed for the first segment across four different models: RNN, GRU, LSTM, and Transformer. All models display relatively close predictions to the actual values, with the Transformer model appearing to capture the overall trend more smoothly, particularly in handling fluctuations. The GRU and LSTM models also show strong performance but with slight discrepancies in higher and lower peaks. The RNN model exhibits more noticeable variations between the predicted and actual values, especially in the latter part of the segment.

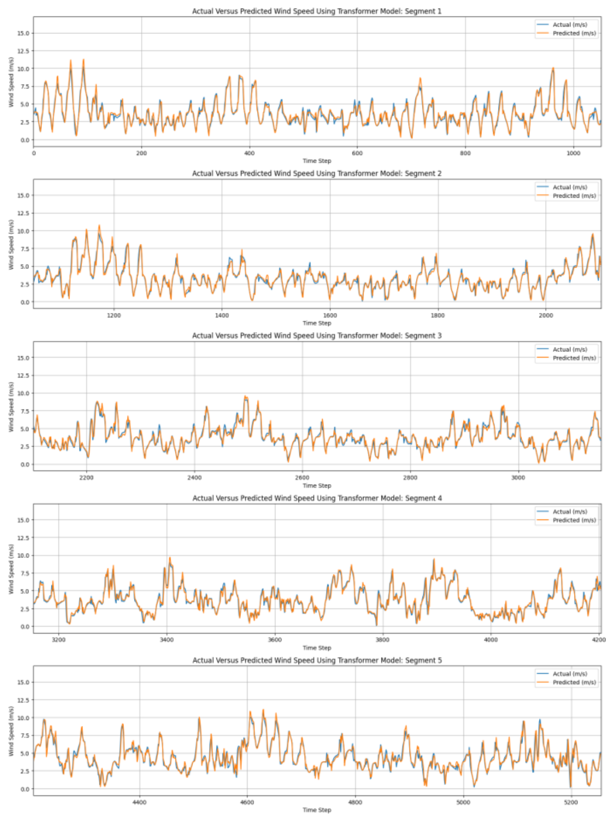

Figure 7. Comparison of actual versus predicted wind speed across multiple time segments. The plots illustrate the performance of transformer model in capturing temporal variations and short-term fluctuations in wind speed, highlighting the accuracy and potential discrepancies between observed and predicted values

Table 7. Error values for transformer model for wind speed prediction using a random seed of 42

Metrics of Transformer Model			
MSE (m/s)	MAE (m/s)	RMSE (m/s)	R^2
0.2430	0.3439	0.4930	0.9402

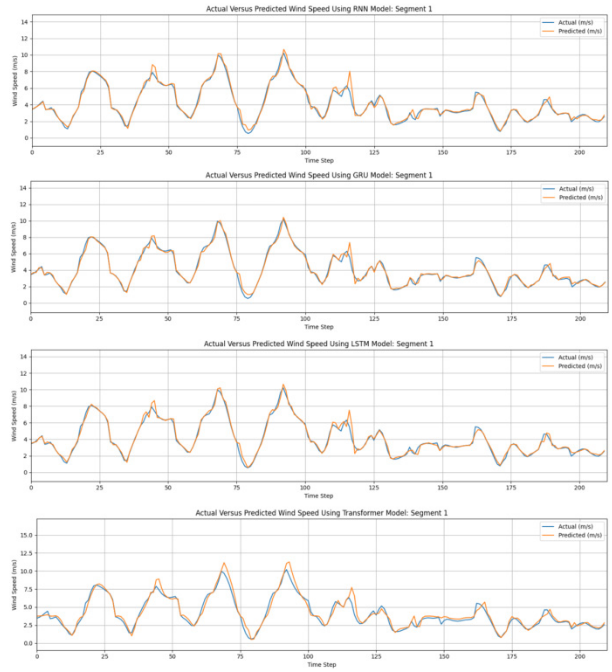

Figure 8. Zoomed-in view of actual vs. predicted wind speed for the first segment across different models.

3.3.6 Multiple Runs of Models

Different seed numbers in machine learning algorithms affect the accuracy of tested models because they control random processes like weight initialization, data shuffling, and stochastic operations such as dropout in neural networks. A different seed leads to a different initialization of model weights, which can cause the model to converge to a different local minimum during training. This variability in the training process can impact the final performance and accuracy of the model. Thus, the use of different seeds typically results in slight variations in model accuracy because the model might learn slightly different or make different decisions during training. This randomness, especially in complex models, can lead to changes in accuracy between runs. Running the model multiple times with different seeds and reporting the average accuracy can provide a more stable and reliable evaluation of its performance.

Table 8. Performance evaluation of all models across five different random seeds (7, 42, 88, 123, 2024)

Model	Seed	MSE (m/s)	MAE (m/s)	RMSE (m/s)	R^2
RNN	7	0.2151	0.3103	0.4637	0.9471
	42	0.1191	0.2191	0.3452	0.9707
	88	0.1795	0.2853	0.4237	0.9558
	123	0.1840	0.2875	0.4289	0.9547
	2024	0.2211	0.3490	0.4702	0.9456
GRU	7	0.1569	0.2593	0.3962	0.9614
	42	0.0977	0.1928	0.3126	0.9759
	88	0.1471	0.2467	0.3836	0.9638
	123	0.1473	0.2485	0.3838	0.9637
	2024	0.1549	0.2565	0.3936	0.9619
LSTM	7	0.2179	0.3190	0.4668	0.9464
	42	0.1120	0.2062	0.3347	0.9724
	88	0.1998	0.3031	0.4469	0.9508

	123	0.1989	0.3006	0.4460	0.9510
	2024	0.2061	0.3023	0.4540	0.9493
Transformer	7	0.2347	0.3502	0.4844	0.9422
	42	0.2430	0.3439	0.4930	0.9402
	88	0.2933	0.4156	0.5415	0.9278
	123	0.2106	0.3101	0.4589	0.9481
	2024	0.2161	0.3112	0.4648	0.9468

The predictability of all models was evaluated across five different random seeds (7, 42, 88, 123, 2024), see Table 8. The GRU consistently achieved the best per–formance, with the lowest MSE values ranging from 0.0977 to 0.1569 and the highest R^2 values between 0.9614 and 0.9759, indicating strong predictive accu–racy. The LSTM and RNN models showed moderate predictability, with MSE values between 0.1120 and 0.2211, and R^2 values around 0.9456 to 0.9724. The tra–nsformer model exhibited slightly higher MSE values (0.2106 to 0.2933) and lower R^2 scores (0.9278 to 0.9481), indicating relatively less accurate predictions under the tested configurations.

3.4 Neural Models Implementation and Training

3.4.1 SHAP Analysis

SHAP analysis was conducted on the dataset to analyze feature contributions in each model prediction. In RNN model, the model predicted a value of approximately 0.20, slightly below the SHAP base value, as shown in Figure 9(a). The most influential factor in lowering the prediction was WS10M. T2M also contributed nega–tively, while QV2M, WD10M, and PS had positive contributions, increasing the prediction.

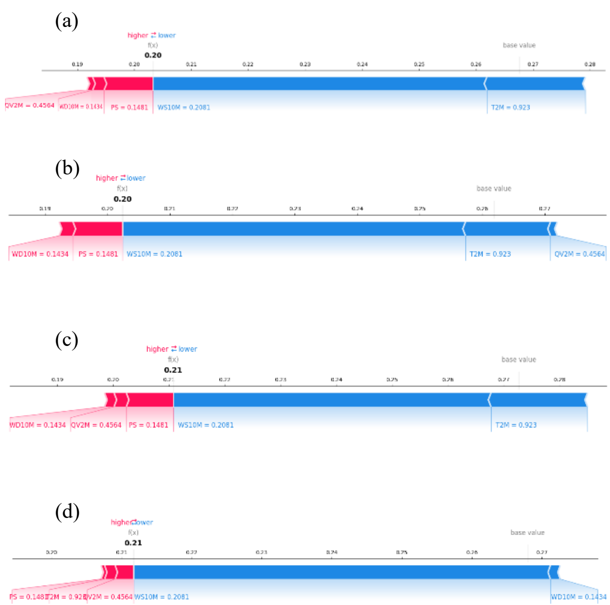

Figure 9. SHAP analysis of feature contributions to the (a) RNN, (b) GRU, (c) LSTM, and (d) Transformer model predictions

Similar interpretations were made across other models. In the GRU model, only WD10M and PS inc–reased the prediction, as shown in Figure 9(b). In the LSTM model, WD10M, QV2M, and PS showed positive contributions, as shown in Figure 9(c). For the Transformer model, PS, T2M, and QV2M increased the prediction, as shown in Figure 9(d). While these fin–dings demonstrate model behavior, they should be interpreted cautiously.

3.4.2 Ablation Analysis

For each feature, models were retrained after feature removal, and performance metrics were recorded, as presented in Table 9. Across all models, the removal of T2M resulted in a noticeable decline in performance, especially for the GRU and LSTM models. These findings confirm the positive contribution of T2M fea–ture in increasing model predictive accuracy. For ins–tance, the GRU model's RMSE increased from its typical baseline (=0.3126m/s) to 0.3257m/s when T2M was excluded. Similarly, removing QV2M led to moderate degradation, particularly in LSTM and RNN, while GRU remained relatively robust (RMSE: 0.3169m/s). Interestingly, removing WD10M had a smaller impact, with performance metrics showing only slight increases in error across all models, suggesting its influence may be more context dependent. The exc–lusion of PS also caused minor degradation in all models, though GRU again demonstrated the highest resilience. Notably, the transformer model showed consistently higher error values regardless of which feature was removed, indicating its relative sensitivity and potentially lower adaptability in this task setting.

Table 9. Impact of individual feature removal on model per–formance. Each model was retrained without one feature (T2M, QV2M, WD10M, or PS), and performance metrics (MSE, MAE, RMSE, R^2) recorded to assess the contribution of each to wind speed prediction

Features	Model	MSE (m/s)	MAE (m/s)	RMSE (m/s)	R^2
Without T2M	RNN	0.1382	0.2342	0.3717	0.9660
	GRU	0.1061	0.2007	0.3257	0.9739
	LSTM	0.1255	0.2189	0.3542	0.9691
	Transformer	0.2203	0.3181	0.4694	0.9458
Without QV2M	RNN	0.1374	0.2380	0.3707	0.9662
	GRU	0.1004	0.1953	0.3169	0.9753
	LSTM	0.1167	0.2117	0.3417	0.9713
	Transformer	0.2058	0.3001	0.4536	0.9493
Without WD10M	RNN	0.1361	0.2336	0.3689	0.9665
	GRU	0.1182	0.2061	0.3437	0.9709
	LSTM	0.1351	0.2261	0.3675	0.9667
	Transformer	0.2081	0.3186	0.4562	0.9488
Without PS	RNN	0.1261	0.2258	0.3551	0.9689
	GRU	0.1002	0.1926	0.3165	0.9753
	LSTM	0.1149	0.2133	0.3389	0.9717
	Transformer	0.1977	0.2909	0.4447	0.9513

4. DISCUSSION

The study presented a comprehensive approach to deve–lop and evaluate four deep learning models for wind speed forecasting using high-resolution hourly meteo–rological data obtained forAl-Ahsa region over a 24-year period. To achieve the objective of this study, a rigorous methodological approach was employed, in–volving comprehensive data pre-processing, sequential data splitting, features selection, and models architecture design.

The four architectures were constructed to predict wind speed as intended. The RNN model demonstrated strong predictive capability and successfully captured temporal dependencies in the input features. Visual inspection revealed a close alignment between actual and predicted wind speeds, as shown in Figure 4. Although RNN model is the simplest among the recurrent models, it performed reasonably well, with an R^2 of 0.9707 and error values moderately higher than GRU and LSTM, as illustrated in Table 4. The GRU model consistently outperformed its counterparts across all evaluation metrics, where predicted values perfectly overlapped actual values, as shown in Figure 5. It achieved the lowest error values (MSE: 0.0977m/s, MAE: 0.1928 m/s, RMSE:0.3126 m/s) and the highest coefficient of determination ($R^2 = 0.9759$), as illustrated in Table 5.

The LSTM model also demonstrated strong performance, with an R^2 of 0.9724, as illustrated in Figure 6 and Table 6. Although its error metrics were higher than those of GRU, it still outperformed both the standard RNN and transformer models. It indicates that while LSTM is capable of handling complex temporal structures, the GRU may achieve similar or better performance with fewer parameters and faster training. Lastly, the transformer, despite its success in many sequential tasks, showed the weakest performance in this study. As illustrated in Table 7, the transformer lagged behind all three recurrent models with the highest MSE of 0.2430 m/s and lowest R^2 of 0.9402. This could be attributed to the model's architecture and data size, which may favor recurrence-based models over self-attention mechanisms. The transformer may require more extensive data and tuning to match the performance of recurrent models in this context. Nonetheless, its statistical metrics remain well within acceptable limits, indicating reliable model performance. Overall, all models achieved high coefficients of determination ranging from approximately 0.94 to 0.97, demonstrating their effectiveness in wind speed forecasting. Since the coefficient of determination ranges from 0 to 1, values closer to 1 indicate a stronger agreement between the predicted and observed data.

The consistent superiority of the GRU model performance across all random seeds, as shown in Table 8, suggests that its architecture is particularly well-suited for the temporal dynamics present in meteorological data. Its gating mechanism appears effective in preserving relevant temporal dependencies while minimizing overfitting and noise sensitivity. The moderate performance of the transformer model, despite its growing popularity in time-series forecasting, may reflect a mismatch between its self-attention-based design and the relatively smooth, low-frequency patterns typical of meteorological time series. The relatively high variability in R^2 across models, particularly in the transformer, also points to differences in stability under random initialization. This sensitivity may have implications for real-world deployment, where models need to generalize well despite slight changes in training conditions. As the performance of the GRU model remained nearly consistent across different seeds, with only small variations in results, it can be considered more reliable for forecasting.

It is worth mentioning that the development of RNN, GRU, LSTM, and Transformer models for wind prediction varies significantly in terms of complexity and computational resource requirements. RNNs are the simplest and least computationally demanding but struggle with long-term dependencies due to vanishing gradient problems. GRUs and LSTMs address this issue by introducing gates to manage memory, making them more complex and computationally expensive, with LSTMs requiring more resources due to their more intricate structure. Transformers, while powerful for capturing long-range dependencies, have the highest complexity and require substantial computational resources, especially for training on large datasets, due to their attention mechanism and parallelization needs. In terms of temporal requirements, training time increases progressively from RNNs to Transformers, with transformers often needing the longest training times due to their high parameter count and large memory usage.

Based on Figures 9(a), (b), (c), and (d) and Table 9, the SHAP and ablation analyses together reveal that input features do not contribute equally across models. T2M and QV2M showed low SHAP values in some predictions, which may indicate a limited impact on model performance. However, ablation results revealed that removing these features significantly reduced model performance, indicating they play a more important role overall than SHAP alone suggests.The minor performance degradation caused by removing PS and WD10M may indicate redundancy or collinearity among meteorological features. However, the fact that GRU model still performed relatively well even when these features were excluded highlights its robustness in handling reduced or partially missing data, which is an important practical advantage in real-world meteorological applications where sensor failures or missing data are common.

When compared to existing literature, the present study shows highly competitive performance in wind speed forecasting, particularly with the GRU model. As illustrated in Table 10, the GRU architecture achieved an R^2 value of 0.9759 and an RMSE of 0.3126m/s, outperforming or matching the R^2 values of several advanced hybrid models applied in diverse geographical settings and with varying temporal resolutions. Notably, while some studies reported higher R^2 scores, such as Amirteimoury et al. [24] reported coefficient of determination as 0.999. However, such results are often obtained under shorter forecast horizons (e.g., 1-hour) and potentially more favorable climatic conditions. In contrast, the models developed in this study were designed for a 24-hour forecast horizon and tested under the challenging meteorological conditions of Al-Ahsa, Saudi Arabia, a region characterized by extreme desert climate variability. The strong performance observed, particularly for GRU and LSTM models, underscores the effectiveness of the methodological framework adopted in this work, including the choice of input features, the sequential data segmentation strategy, and the design of robust recurrent architectures. Significantly, this perfor–

mance was achieved without reliance on computationally intensive hybrid techniques, complex optimization algorithms, or elaborate data pre-processing. This simp–licity enhances the practical applicability of the models, making them well-suited for deployment in operational environments with limited computational resources.

Table 10. Comparison of the performance of models in the current study with other approaches reported in different studies and regions

Ref.	Method	MSE (m/s)	MAE (m/s)	RMSE (m/s)	R^2
[16]	FIG-MSS-LSSVM	-	-	-	0.936
[26]	RUN-VMD-Seq2Seq-Attention	3.873	-	-	0.952
[35]	C-LSTM	0.437	0.127	-	0.913
[20]	DeepAR	0.00494	0.04609	0.07026	0.780
	TFT	0.00294	0.0595	0.05422	0.850
	MIDF	0.0035	0.01739	0.01913	0.890
[24]	DWT-MI-BiLSTM-COOT	0.009	0.069	-	0.999
[25]	WPCA-PSO-GRU	-	1.43	2.22	0.917
[36]	EEMD-BA-RGRU-CSO	-	0.181	0.225	0.988
Current Study	RNN	0.1191	0.2191	0.3452	0.9707
	GRU	0.0977	0.1928	0.3126	0.9759
	LSTM	0.1120	0.2062	0.3347	0.9724
	Transformer	0.2430	0.3439	0.4930	0.9402

5. CONCLUSIONS

This study provides a comprehensive evaluation of multiple deep learning architectures for short-term wind speed forecasting, utilizing an extensive meteorological dataset collected from the Al-Ahsa region. The primary objective of this study was to develop, optimize, and analyze multiple deep learning architectures for accurate short-term wind speed forecasting using long-term meteorological data. The study aimed to evaluate these models' predictive performance, robustness, and gene–ralization ability under realistic operational conditions, while also investigating feature importance through interpretability (SHAP) and sensitivity (ablation).

In this study, four sequence-based models were deve–loped, optimized, and rigorously compared through real-world evaluation scenarios. These models are RNN, GRU, LSTM, and transformer. Firstly, the dataset was pre-processed, and segmented, where 80% of dataset were used for training, 10% for validation and 10% for testing. Comprehensive hyperparameter tuning was conducted to optimize model architectures, focusing on key parameters such as the number of units, dropout rates, and learning rates. The transformer model was augmented with posi–tional encoding to effectively preserve temporal depen–dencies within the input sequences. Furthermore, all models were evaluated across multiple random seeds to ensure the reliability and robustness of results.

Across all evaluation metrics, the GRU model attained the highest coefficient of determination of 0.9759 and the lowest error metrics (MSE, RMSE, and MAE). The results confirmed that the GRU model de–monstrated the best performance among all models, especially in terms of mean absolute error and root mean squared error. Although the LSTM model exhi–bited competitive performance, the other models sho–wed relatively lower accuracy. While the GRU model exhibited superior predictive performance across all evaluation metrics, the comparatively moderate results of the transformer architecture do not diminish its theoretical potential. Rather, they underscore opportu–nities for methodological advancement and refinement. The inherent capacity of the transformer's attention me–chanism to model long-range dependencies and capture intricate temporal relationships remains underexploited in its current configuration.

To enhance interpretability, SHAP analysis was applied to investigate feature contributions. This ana–lysis, performed on a representative subset of the data to reduce computational demands, revealed that while some features showed variable importance across individual predictions, others consistently influenced the models' outputs. Complementing this, ablation analysis provided a global sensitivity perspective. These ana–lyses confirmed the critical roles of certain meteoro–logical variables such as temperature and humidity, which, despite sometimes low local SHAP values, were essential for maintaining overall predictive accuracy.

Given the potential for changes in environmental conditions and the availability of new data when these models are used in real applications. It is worth men–tioning that, retraining the models should be done periodically, perhaps annually or quarterly, to adapt to evolving environmental conditions and ensure accuracy. The used models differ in how they process sequential data, with LSTMs and Transformers excelling at cap–turing long-term dependencies, while Transformers require more computational resources. Adding more features like wind speed at different altitudes, solar radiation, or atmospheric pressure gradients could en–hance model accuracy by providing a more compre–hensive understanding of the factors influencing wind patterns.

Future research directions include the incor–poration of external environmental variables (e.g., solar radiation, terrain elevation, and seasonal indices), integ–rating real-time data, the exploration of probabilistic forecasting techniques to better quantify prediction un–certainty, and the development of adaptive learning strategies such as online learning or transfer learning to improve model adaptability across varying climatic conditions and geographical locations.

AUTHOR CONTRIBUTIONS

Conceptualization, M.M., and S.R.; literature search, N.A. and M.A; methodology, N.A.; resources, N.A., M.A., and M.M.; writing-original draft preparation, N.A.; writing-review and editing, N.A., M.M., S.R.; proofreading, N.A., M.M., and S.R.; supervision, M.M. and S.R. All authors have read and agreed to the published version of the manuscript.

FUNDING

The authors acknowledge the support of King Fahd University of Petroleum & Minerals.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article are available at https://power.larc.nasa.gov/.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

[1] M. A. Hanif, F. Nadeem, R. Tariq, U. Rashid, "Chapter 1 - Energy resources and utilization," *Re–newable and Alternative Energy Resources*, M. A. Hanif, F. Nadeem, R. Tariq and U. Rashid, eds., pp. 1-30: Academic Press, 2022.

[2] O. Ellabban, H. Abu-Rub, F. Blaabjerg, "Rene–wable energy resources: Current status, future pros–pects and their enabling technology," *Renewable and Sustainable Energy Reviews,* vol. 39, pp. 748-764, 2014/11/01/, 2014.

[3] B. Rašuo, M. Dinulović, A. Veg, A. Grbović, A. Bengin, "Harmonization of new wind turbine rotor blades development process: A review," *Renewable and Sustainable Energy Reviews,* vol. 39, pp. 874-882, 2014/11/01/, 2014.

[4] B. Rašuo, A. Bengin, A. Veg, "On Aerodynamic Optimization of Wind Farm Layout," *PAMM,* vol. 10, no. 1, pp. 539-540, 2010/12/01, 2010.

[5] [5] A. M. Ghanem, Y. A. Alamri, "The impact of the green Middle East initiative on sustainable development in the Kingdom of Saudi Arabia," *Journal of the Saudi Society of Agricultural Scie–nces,* vol. 22, no. 1, pp. 35-46, 2023.

[6] S. Roga, S. Bardhan, Y. Kumar, S. K. Dubey, "Re–cent technology and challenges of wind energy generation: A review," *Sustainable Energy Techno–logies and Assessments,* vol. 52, pp. 102239, 2022/08/01/, 2022.

[7] F. H. Ouerghi et al.: "Feasibility evaluation of wind energy as a sustainable energy resource," *Alexandria Engineering Journal,* vol. 106, pp. 227-239, 2024/11/01/, 2024.

[8] S. Mathew, *Wind energy: fundamentals, resource analysis and economics*: Springer, 2006.

[9] B. P. Rašuo, A. Č. Bengin, "Optimization of wind farm layout," *FME transactions,* vol. 38, no. 3, pp. 107-114, 2010.

[10] B. P. Rasuo, A. D. Veg, "Design, fabrication and verification testing of the wind turbine rotor blades from composite materials," *Proceedings of the ICCM-16, Japan*, pp. 1-4, 2007.

[11] A. Tascikaraoglu, M. Uzunoglu, "A review of co–mbined approaches for prediction of short-term wind speed and power," *Renewable and Sustai–nable Energy Reviews,* vol. 34, pp. 243-254, 2014.

[12] J. Lerner, M. Grundmeyer, M. Garvert, "The im–portance of wind forecasting," *Renewable Energy Focus,* vol. 10, no. 2, pp. 64-66, 2009/03/01/, 2009.

[13] D. Wahdany, C. Schmitt, J. L. Cremer, "More than accuracy: end-to-end wind power forecasting that optimises the energy system," *Electric Power Sys–tems Research,* vol. 221, pp. 109384, 2023/08/01/, 2023.

[14] A. M. Foley, P. G. Leahy, A. Marvuglia, E. J. McKeogh, "Current methods and advances in fore–casting of wind power generation," *Renewable Energy,* vol. 37, no. 1, pp. 1-8, 2012/01/01/, 2012.

[15] O. Karakuş, E. E. Kuruoğlu, M. A. Altınkaya, "One -day ahead wind speed/power prediction based on polynomial autoregressive model," IET Renewable Power Generation, vol. 11, no. 11, pp. 1430-1439, 2017.

[16] G. Lv, L. Chen, Z. Dou, Z. An, "Dynamic self-ada–ptive segmentation of wind speed interval predic–tion based on fuzzy information granulation." pp. 422-428.

[17] [S. Rehman, S. A. Khan, L. M. Alhems, "The Ef–fect of Acceleration Coefficients in Particle Swarm Optimization Algorithm with Application to Wind Farm Layout Design," *FME Transactions,* vol. 48, no. 4, 2020.

[18] S. A. Khan, "Adaptation of the Simulated Evolution Algorithm for Wind Farm Layout Optimization," *FME Transactions,* vol. 50, no. 4, 2022.

[19] J. Cai, C. Peng, L. Xu, B. Wu, X. Bian, "A multi-stage multivariable offshore wind speed assessment system incorporating statistical analysis, decompo–sition, feature selection, bias correction, and atten–tion-based hybrid prediction," *Wind Engineering*, pp. 0309524X251314204, 2024.

[20] M. Maruthi, B. Kim, S. Sujeen, J. An, Z. Chen, "Multi-model integration for dynamic forecasting (MIDF): a framework for wind speed and direction prediction," *Artificial Intelligence Review,* vol. 58, no. 6, pp. 164, 2025.

[21] Y.-Y. Hong, C. L. P. P. Rioflorido, "A hybrid deep learning-based neural network for 24-h ahead wind power forecasting," *Applied Energy,* vol. 250, pp. 530-539, 2019.

[22] S.-X. Wang, M. Li, L. Zhao, C. Jin, "Short-term wind power prediction based on improved small-world neural network," *Neural Computing and Applications,* vol. 31, no. 7, pp. 3173-3185, 2019.

[23] S. Han, Y.-h. Qiao, J. Yan, Y.-q. Liu, L. Li, Z. Wang, "Mid-to-long term wind and photovoltaic power generation prediction based on copula func–tion and long short term memory network," *Applied energy,* vol. 239, pp. 181-191, 2019.

[24] F. Amirteimoury, F. Keynia, E. Amirteimoury, G. Memarzadeh, H. Shabanian, "A novel wind speed prediction model based on neural networks, wavelet transformation, mutual information, and coot opti–mization algorithm," *Scientific Reports,* vol. 15, no. 1, pp. 10860, 2025.

[25] Y. Xiao, C. Zou, H. Chi, R. Fang, “Boosted GRU model for short-term forecasting of wind power with feature-weighted principal component ana–lysis,” *Energy,* vol. 267, pp. 126503, 2023.

[26] M. Du, Z. Zhang, C. Ji, “Prediction for Coastal Wind Speed Based on Improved Variational Mode Decomposition and Recurrent Neural Network,” *Energies,* vol. 18, no. 3, pp. 542, 2025.

[27] S. Rehman, H. Nuha, M. Islam, F. Schulze, “Accu–racy of wind speed predictability with heights using Recurrent Neural networks,” *FME Transactions,* vol. 49, no. 4, pp. 909, 2021.

[28] D. Wei, Z. Tian, “A comprehensive multivariate wind speed forecasting model utilizing deep lear–ning neural networks,” *Arabian Journal for Science and Engineering,* vol. 49, no. 12, pp. 16809-16828, 2024.

[29] M. E. Karare eta l.: “A New Gated Recurrent Unit Network-Based Wind Turbine Power Prediction,” *Arabian Journal for Science and Engineering*, pp. 1-10, 2024.

[30] M. A. Benmahdjoub, A. Mezouar, M. Ibrahim, L. Boumediene, Y. Saidi, M. Atallah, “Accurate esti–mation of effective wind speed for wind turbine control using linear and nonlinear Kalman filters,” *Arabian Journal for Science and Engineering,* vol. 48, no. 5, pp. 6765-6781, 2023.

[31] M. Elsaraiti, A. Merabet, “Application of long-short-term-memory recurrent neural networks to forecast wind speed,” *Applied Sciences,* vol. 11, no. 5, pp. 2387, 2021.

[32] D. Gupta, V. Kumar, I. Ayus, M. Vasudevan, N. Natarajan, “Short-Term Prediction of Wind Power Density Using Convolutional LSTM Network,” *FME Transactions,* vol. 49, no. 3, 2021.

[33] S. Rehman, U. T. Salman, M. A. Mohandes, F. A. Al-Sulaiman, S. Adetona, L. M. Alhems, M. A. Baseer, “Wind Speed Prediction Based on Long–Short Term Memory using Nonlinear Autoreg–ressive Neural Networks,” *FME Transactions,* vol. 50, no. 2, 2022.

[34] M. Al-Bouwarthan, M. M. Quinn, D. Kriebel, D. H. Wegman, “Assessment of heat stress exposure among construction workers in the hot desert cli–mate of Saudi Arabia,” *Annals of work exposures and health,* vol. 63, no. 5, pp. 505-520, 2019.

[35] D. Wang, M. Xu, Z. Guangming, F. Luo, J. Gao, Y. Chen, “Enhancing wind power forecasting accuracy through LSTM with adaptive wind speed calib–ration (C-LSTM),” *Scientific Reports,* vol. 15, no. 1, pp. 5352, 2025.

[36] A. Meng, S. Chen, Z. Ou, W. Ding, H. Zhou, J. Fan, H. Yin, “A hybrid deep learning architecture for wind power prediction based on bi-attention mechanism and crisscross optimization,” *Energy,* vol. 238, pp. 121795, 2022.

КОМПАРАТИВНА ЕВАЛУАЦИЈА РЕКУРЕНТНИХ И НА ПАЖЊИ ЗАСНОВАНИХ МОДЕЛА ДУБОКОГ УЧЕЊА ЗА ПРОГНОЗИРАЊЕ БРЗИНЕ ВЕТРА

Н.Џ. Ал-Кахтани, М.С. Аболуз, М.А. Мохендес С. Рехман

Прогнозирање брзине ветра је кључно за повећање производње енергије ветра и побољшање ефикас–ности. Иако се традиционални статистички и физич–ки модели често користе, не успевају да обухвате сложену, нелинеарну и временски промењиву при–роду метеоролошких података. Стога, ова студија истражује четири архитектуре дубоког учења: реку–рентну неуронску мрежу (RNN), затворену реку–рентну јединицу (GRU), дугорочну краткорочну меморију (LSTM) и трансформатор. Ови модели су анализирани коришћењем мултиваријантних и сат–них метеоролошких података из региона Ал-Ахса током 24 године (2001-2024). Скуп података укљу–чује пет карактеристика: брзину ветра на 10 метара изнад тла (користи се као референца), температуру ваздуха на 2 метра, специфичну влажност на 2 метра, смер ветра на 10 метара и површински при–тисак. Модели су обучени на 80% скупа података, валидирани на 10% и тестирани на преосталих 10%. GRU модел је постигао најбоље перформансе са RMSE = 0,3126 m/s и R^2 = 0,9759.