Prediction of Wind Speed using Random Forest Learning Model considering multi-meteorological Variables

Jayati Vaish*, [1], Sandhya Shrivastava*, Amit Kumar*, and Bindeshwar Singh^

www.RERICProceedings.ait.ac.th

Abstract – *Wind power and other renewable energy sources are crucial for minimizing emission of greenhouse gas and preventing weather variation. The prediction of wind speed accurately is essential for maximizing production of wind energy and successfully integrating it into the electrical grid. In this paper, the real time data of wind speed for 24 hours of Bengaluru city (Karnataka, India) from 2011-2021 i.e., 11 years is considered to train Ensemble Learning based Random Forest (RF) learning model to forecast wind speed. The meteorological variable quantities such as dry bulb temperature, relative humidity and dew point are correspondingly taken as input parameters to train RF model. The short-term prediction of the wind speed is estimated for the day ahead and one hour. The statistical measures such as mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE) and normalized root mean square error (NRMSE) are evaluated to assess the performance of the results obtained from RF model. The outcomes determined from the proposed RF model is also equated with the conventional and advanced method to evaluate its efficacy and accuracy. The analysis demonstrate that the Random Forest model provides improved outcome than existing methods, providing accurate and reliable wind speed forecasts by handling non-linearities in the data.*

Keywords – Ensemble learning, Random Forest, Renewable energy, Meteorological variables, Prediction of wind speed.

1. INTRODUCTION

The main strategy in addressing the worldwide issue of climate change and moving towards a future of sustainable environment is obtained from the quick development of renewable energy resources, especially wind power. For efficient planning, operation, and integration of the wind energy into the energy grid, accurate wind speed forecast is essential. Prediction of the wind speed helps policymakers, energy producers, and power grid operators to make well-informed decisions, maximize energy production, and maintain grid stability. [1].

The depletion of the fossil fuels and the expanding demand for the ecological fortification have propelled the rapid development of the alternative energy and new energy industries in recent years. Wind energy has the advantages of being clean and renewable. As a result, there are several financial and community benefits to the widespread growth and usage of wind energy. Therefore, wind energy may be a very beneficial resource globally. However, the dominant and extensive development in the wind system is severely constrained by the inherent randomness and intermittency of wind speed, which presents a significant challenge to power grid operation and management, especially when taking alternative energy integration into account [2]. Hence, for accurate forecasting of wind speed, it is divided into various groups: Very short-term predictions range from a few hours to a day, short-term predictions range from more than a few days to a week. Long-term predictions are those that span more than a year, whereas medium-term predictions typically range from one week to a year. The short-term forecast takes into account load statistics, holidays, festivals, social events, and season data. In medium and long-term prediction, past wind speed and seasonal data are taken into account. Also, the total number of consumers in all markets and sectors is also considered in long term forecasting.

For wind turbines to be managed efficiently and machinery to operate as intended, the short-term wind speed forecast provides certain characteristics which are relevant to scientific power system control and comparatively more accurate. The numerical weather prediction model is mostly used for medium-term forecasts, which are used to schedule equipment maintenance and troubleshooting. The selection of wind farm locations is based on long-term projections that forecast the possible advantages.

For wind energy installations, a number of methods have been developed and put into use to forecast wind speed. Persistence method, physical method, statistical approaches, spatial correlation technique, fuzzy logic, support vector machines (SVM), artificial intelligence based approaches and hybrid methods are the six basic categories into which wind speed prediction techniques are divided.

The Persistence Method makes forecasts based on the simple assumption that wind power and speed will not change over time. For ultra-short-term prediction, the persistence model outperforms other wind speed prediction methods. The most economical and useful

*Department of Electrical Engineering, Shri Ramswaroop Memorial College of Engineering and Management, Ayodhya Road, Lucknow, 226028, Uttar Pradesh, India.

^Department of Electrical Engineering, Kamla Nehru Institute of Technology, Sultanpur, 228118, Uttar Pradesh, India.

[1] Corresponding author;
Tel: 8840011977
E-mail: itayaj26@gmail.com

technique for prediction of the wind speed is the persistence method. For estimates provided in the near future, the electrical business uses the persistence strategy. It is essential to compare a newly developed prediction technique to the conventional measure for the persistence approach in order to ascertain the extent to which it may enhance predictions produced by persistence. However, the persistence approach's accuracy quickly decreases as the anticipated time scale increases. This approach is only appropriate when the weather is stable or when there is minimal change in the synoptic conditions [3].

The physical approach uses numerical weather forecast information from the lower atmosphere, including surface roughness, temperature, and pressure. Meteorologists use the Numerical Weather Prediction (NWP) approach to forecast weather on a large scale. To improve the NWP model's real accuracy and produce reliable weather forecasts, physical models are employed. Supercomputers are used to apply physical techniques since they need a lot of calculations. For the lower atmosphere, the physical approach depends on numerical weather forecast data [4].

Statistical methods are employed to ascertain the correlation between measured power and online data. A statistical model can be constructed using Wind Energy Conversion Systems (WECS) records. Statistical models are less expensive and easier to develop than other models. The statistical technique includes methods such as autoregressive (AR), auto-regressive moving averages (ARMA), auto-regressive integrated moving averages (ARIMA). This method is taken for short period of time. The drawback of this approach is that the error of prediction error increases with the increase in time. The application of arithmetical approaches is applied where large data and its observation is required such as engineering, economics and natural sciences [5]-[6].

The geographic relation between wind speed at various sites is taken into account by spatial correlation models. This is crucial for applications in the field of environmental impact research, wind energy and air quality evaluation. The approaches considered for predicting wind speed using geographic correlation models are gaussian process regression (GPR), Kriging models, gaussian weighted regression (GWR), geographic autoregressive (GAR) models and spatial interpolation approaches. The need and characteristics of the wind speed data will determine whether the above models are utilized separately or in combination. The selected approach or model is validated using proper validation methodologies, when working with time series data for forecasting [7]

The technique employed using artificial intelligence (AI), also known as Artificial neural network (ANN) is the technique created to imitate the operation of the human brain and its functions. The ANN method is a perfect method for managing complex problems and dynamics of the wind speed variations using the approach of interconnected neurons seen in human brain and thus, can operate complex relationship of the data. They can also learn and improve their performance to enhance its efficacy and stability. There are two varieties of ANN topologies known as feedforward and feedback. A feedforward ANN network is structured into layers without a feedback path. The feedback ANN network has feedback channel between the layers into the network. The application of ANN covers all fields including meteorology and hydrology for developing prediction models. Therefore, early warning systems (EWS) are established in impacted areas through the artificial neural networks (ANNs) to construct accurate prediction models. Early in the 1990s, ANNs were first observed to be used in meteorology and hydrology. In recent years, predictive models have gone a revolution due to ANNs which allow us to solve complex and challenging issues of wind speed predictions with accuracy and precision [8]-[9].

Hybrid models apply the best aspects of the multiple forecasting approaches to increase the wind speed estimation, precision, robustness and flexibility. The wind speed outputs of these hybrid models are combined to determine accurate and efficient outcomes. These models are especially useful in scenarios where individual single models can have limits. The hybrid models can be updated to include the recent information when weather patterns change or new data available. By combining the advantages of the machine learning models and domain specific models, hybrid models can maximize efficiency of both the techniques and the kind of data being examined. The use of hybrid models is still being studied and efforts are being made to improve these model performances [10].

Fuzzy logic is a strong and adaptable method for handling imprecision and uncertainty in forecasts and decision making. It offers a solution for dealing with the uncertainty and unpredictability present in meteorological data for wind speed forecasting. The linguistic variables which characterize the circumstances and elements, affecting the accuracy of forecasting, are defined in a fuzzy logic based wind speed prediction system. A fuzzy logic based wind speed prediction system successfully necessitates the careful design of fuzzy sets and validation against historical data. Thus, fuzzy logic models have become alternative to traditional prediction methods. However, these models require accurate modeling for forecasting, which require more time, causing a drawback for this method. It is crucial to strike a balance between simplicity (to guarantee interpretability) and complexity (to capture correct predictions) [11].

Support vector machine (SVM) is a useful procedure for dealing with learning problems. It is based on the Statistical Learning Theory of the 1990s. It was initially applied to classification and pattern recognition problems, but it has now been modified for use in regression problems. Time-series forecasting applications can be emphasized, and the SVM technique has been considered highly competitive. Support Vector Machines (SVM) are used to predict short-term wind.

An illustration of a non-parametric learning process is this tactic. However, performance deterioration may be directed by parametric and structural factors. The impact of the model parameters on the prediction model is also studied while employing support vector machines [12].

In recent years, several research have employed machine learning-based ensemble techniques to generate more precise wind speed estimates. To achieve a more accurate outcome, the Ensemble Learning technique integrates predictions from multiple machine learning algorithms. Among the most popular advanced methods based on supervised ensemble learning algorithms are random forest (RF), bagging, and boosting. Many hybrid approaches based on Ensemble Learning techniques are currently being used to offer more consistent and acceptable results than using a single model [13].

The novel feature of this study is that it makes accurate predictions of wind speed by taking meteorological conditions into consideration. The inconsistency of wind speed is greatly affected by meteorological parameters, which also affect the forecast accuracy of wind speed. Consequently, diminutive research is done on the topic of estimating wind speed using meteorological parameters. The key contribution of this research work is:

1. Analysis of the real-time wind speed data and meteorological dataset from Bengaluru city, Karnataka, India, is collected from 1st January 2011, to 31st December 2021, over a period of 11 years.
2. The wind speed prediction is evaluated employing the ensemble learning algorithm based Random Forest model. The statistical metrics like MAE, MAPE, RMSE, and NRMSE are used to do performance analysis on the data.
3. A comparison of the outcomes from the RF learning model with those from more conventional and innovative methods like Support Vector Regression (SVR), Exponential Smoothing (ES), Auto Regression Moving Averages (ARIMA), and Artificial Neural Networks (ANN).

The layout of the paper has been split into six sections. Section I provides an introduction and a summary of the development of the various prediction algorithms. Section II discusses the Random Forest Model and the Ensemble Learning Method. In section III, the projected method based on machine learning for this study is described. Section IV covers the data analysis used in this investigation. Section V discusses the findings and comparison of numerous techniques employed for prediction of wind speed. The result of the study is discussed in Section VI.

2. METHODS

Machine learning algorithms have lately established their potential to improve wind speed estimation. Among these methods, ensemble learning based Random Forest (RF) model stands out as an appropriate and successful technique for solving complex data for forecasting. The proposed RF method implied for forecasting is explained in next section.

2.1 Ensemble Learning

Ensemble learning is an effective and popular machine learning approach that combines predictions of several separate models to increase its robustness and performance of the predictive models. Ensemble learning was first created in 1990 to solve the high variation and low accuracy problems in machine learning systems. This model is effective in enhancing the functionality of the algorithm used for decision making and forecasting. The fundamental principle of the ensemble learning is that the collective outcome of the numerous models can produce superior results as compared to anyone model alone. The accuracy of ensemble learning models is greatly increased by building several classifiers or base learners and combines them for final outcome. The ensemble learning can solve various problems including error correction, estimation, adaptive learning and missing features. It has the potential to successfully enhance poor learners by incorporating a foundational learning algorithm on learning data [14].

The various domains where ensemble learning models are implied are: speech recognition, computer vision, intrusion detection, processing of natural languages etc. The ensemble learning algorithms which are currently used in various fields are classified as: Bagging (Bootstrap Aggregation), Boosting, Adaboost, Gradient Boosting, XGBoost, Random Forest and Stacking [15]. In this study, the algorithm implied for the wind speed forecasting is Random Forest (RF) model.

2.2 Random Forest Learning Model

The Random Forest (RF) method is a algorithm based on machine learning technique which falls within the group of ensemble learning approach. The RF learning model is first created by Leo Breiman and Adele Cutler in 2001 which has emerged as a key method in machine learning field. The RF method This method is adaptable and powerful in handling the regression and classification problems, which are common in contemporary machine learning algorithms. For RF techniques, three primary hyperparameters need to be set before training. They include the number of trees used, the size of the knots, and the range of splits. Each decision tree in the ensemble that comprises the random forest technique uses a bootstrap sample, which is a data sample selected from a training set with replacement. Fig. 1 shows the Random Forest model used to forecast wind speed by implementing Breiman's theory for generating and predicting RF [16]- [17].

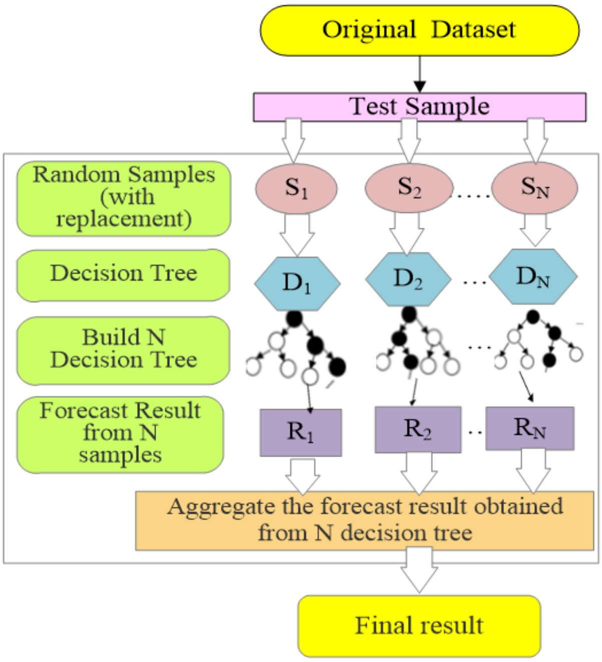


Fig 1. Random Forest Learning Model

In order to diversify the dataset and break its connections to earlier decision trees, point bagging is utilized to fit yet another instance of randomness. Each decision tree will be equalized for a maturity vote and a retrogression job. Finally, the sample is tested by cross-validation and outcome. The Random Forest offers excellent generalization and is unaffected by outlier data. The regression tree is measured using the least necessary residuals, which are shown below.

$$M_r(t) = \frac{1}{n_t}\sum_{i=1}^{N} D_N(c_N) \tag{1}$$

where the average value of individual node item is represented as c_N and n_t is the total number of values at node *t*. Bootstrapping samples S_N is used to create the random set of training data T_k along with the associated decision tree D_N. The value of an individual decision tree taken for prediction is provided using:

$$D_i = \sum_{i=1}^{N} \Theta_i \Pi_{pi} \tag{2}$$

where predicted value is denoted by Π_{pi} and Θ_i is the weight of the value of the node. The average value of a decision tree is calculated as the final expected output using-

$$\psi_N = \frac{1}{N}\sum_{i=1}^{N} \Pi_{pi}(D_i) \tag{3}$$

where ψ_N represents the average or forecasted value.

The steps that the Random Forest (RF) approach use to estimate wind speed are explained as below:

Step 1: Create test samples using the original data set, where N represents the number of samples.

Step 2: From the test sample, random samples are created i.e., $S_1, S_2 .....S_N$ with replacement (Bagging Method).

Step 3: From each sample decision trees $D_1, D_2 .....D_N$ are constructed.

Step 4: The optimum split is determined at individual node of the decision tree by considering only a arbitrary set of features.

Step 5: Every decision tree produces a unique prediction. $R_1, R_2 .....R_N$.

Step 6: The final outcome is determined by taking the average of (in regression) the predictions from decision trees $D_1, D_2 .....D_N$ as stated in equation (3).

Step 7: The final outcome is the predicted value.

3. PROPOSED METHODOLOGY

The schematic layout of the RF learning model utilized in this investigation is displayed in Fig. 2. This study makes use of the original wind speed data and the number of meteorological factors, including relative humidity (assumed to be 70%), dry bulb temperature (DBT), and dew point. The procedures used to determine the predicted wind speed are explained below:

Step 1: Take the original input data of wind speed of Bengaluru city (Karnataka, India) for 11 years from 2011 to 2021.

Step 2: The various meteorological factors, such as relative humidity, dry bulb temperature, and dew point, which influence wind speed are also considered.

Step 3: In this step, the data is pre-processed and applied to the RF model.

Step 5: Following Step 3, the data is fragmented into these datasets: training data and testing data.

Step 5: The training data is divided into training sample data with replacement. By these samples, decision trees are formed randomly.

Step 6: The most significant result obtained from these decision trees are aggregated to obtain the predicted result using Equation 3.

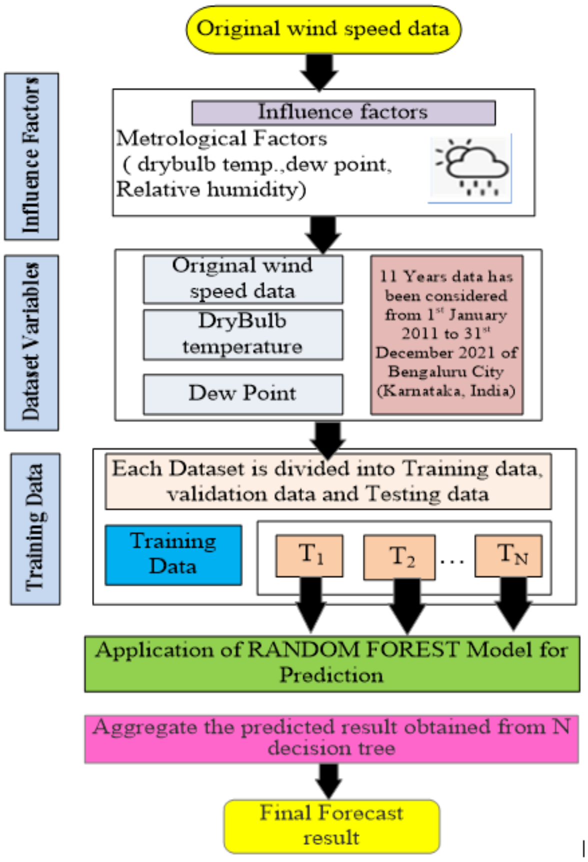


Fig 2. Proposed Methodology of Wind Speed Prediction using RF Learning Model

4. DATA ANALYSIS

4.1 Selection of data

A dataset is explored and examined during data analysis for prediction in order to find trends, connections, and patterns that may be included into predictive models. In this case, the wind speed data of Bengaluru City (Karnataka, India) during an 11-year period from January 2011 to December 2021 is taken into consideration for prediction of wind speed. In this case, wind speed predictions for short term are estimated for an hour and a day.

Figure 3 shows the wind speed data of 11 years for 572 weeks (i.e., 52 weeks*9 years = 468 weeks and 53 weeks*2 years = 106 weeks, therefore total weeks 574 weeks). For one-hour prediction, the total data considered from 1st January 2011 to 31st December 2021 is 96408 hours (i.e., 365 days*9 years*24 hours = 78,840 hours and 366*2*24 = 17,568). For one-day prediction, the total data considered from 1st January 2011 to 31st December 2021 is 4017 days (i.e., 365 days*9 years = 3285 days and 366*2 = 732 days, therefore total days 4017). Relative humidity is considered 70% for the assessment of the data. The air or atmospheric temperature in the city is assumed to be similar to the dry bulb temperature.

The variability and central tendency of the data are described by the statistical analysis. It identifies departures from normality and specifies the structure of the distribution. This analysis is significant for precise and accurate estimations, particularly when extrapolating sample data of the larger dataset. By understanding these statistical metrics, researchers, scientists, and analysts can get important insights into the underlying patterns and features of the data. Decision-making and further research can also be guided by these insights.

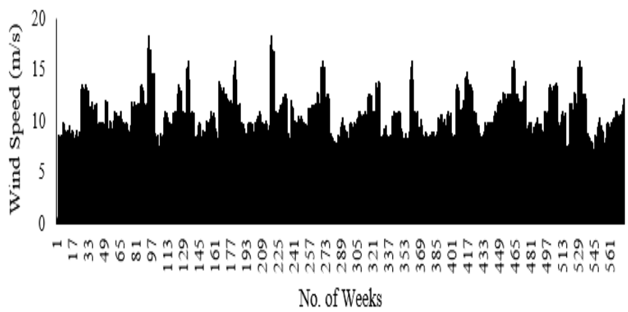


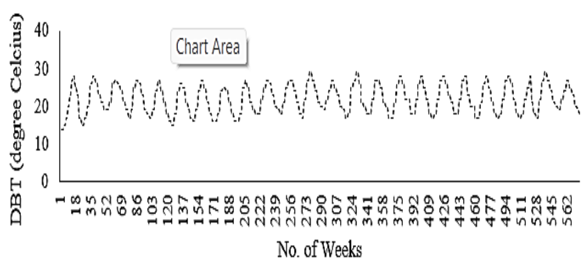
Fig. 3. Wind Speed data for 11 years from 1st January 2011 to 31st December 2021

Table 1. Descriptive Statistics of Actual data of wind speed

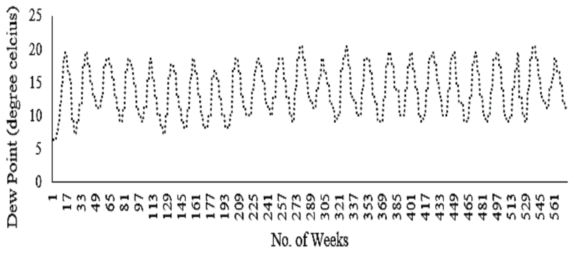
S. No.	Parameters	Values
1	Mean	10.69
2	Median	10.36
3	Mode	9.85
4	Standard Deviation	1.99
5	Kurtosis	0.82
6	Skewness	0.91
7	Min. wind speed	7.24
8	Max. wind speed	18.26

4.2 Meteorological Data Analysis

Meteorological variation is the primary cause of fluctuations in wind speed. This is because many utilities, including space heating, air conditioning, and agricultural irrigation, depend on wind components that are susceptible to weather. One of the weather variables that most affects wind speed is temperature. If the temperature of the place deviates from the actual value on any particular day, the wind speed may change significantly.



(a) Dry Bulb Temperature



(b) Dew Point

Fig. 4. Meteorological variables weekly data for ten years from January 2011 to December 2021

The numerous meteorological factors that affect wind patterns and its forecasting are: temperature, humidity, atmospheric pressure. Although dew point and humidity do not directly affect wind speed, they can have an impact on weather patterns. Weather patterns in turn can have an impact on wind speed due to pressure gradients and other meteorological phenomena.

Table 2. Descriptive Statistics of Meteorological Variables

S. No.	Parameters	Values	
		Dry bulb Temp.	Dew Point
1	Mean	21.878	13.757
2	Standard Error	0.169	0.158
3	Median	22	13.9
4	Minimum	14	6.4
5	Maximum	29	20.4
6	Sample Variance	14.854	13.093
7	Skewness	0.055	0.057
8	Kurtosis	-1.253	-1.251

The dry bulb temperature (DBT) is a type of temperature that measures the air's temperature using a thermometer that is completely exposed to the atmosphere but shielded from moisture and radiation. At the extremes, DBT typically runs from around -20°C to 46°C, rising to about on a really hot day in the desert. Fig. 4(a) displays the DBT statistics for Bengaluru City (Karnataka, India) from January 2011 to December 2021. In a similar vein, Fig. 4(b) displays the dewpoint data for Bengaluru City (Karnataka, India) from January 2011 to December 2021. Table 2 presents the descriptive statistical data analysis.

4.3 *Statistical Evaluation Parameters*

The following statistical parameters are used to assess and analyze the usefulness of the RF learning model output based on:

a) Mean Absolute Percent Error (MAPE):

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{\theta_{Ai}-\theta_{Pi}}{\theta_{Ai}}\right) \tag{4}$$

b) Mean Absolute Error (MAE):

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left(\theta_{Ai}-\theta_{Pi}\right) \tag{5}$$

c) Root Mean Square Error (RMSE):

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\theta_{Ai}-\theta_{Pi}\right)} \tag{6}$$

d) Normalized Root Mean Square Error (NRMSE):

$$NRMSE = \frac{RMSE}{\theta_{Ai}} \tag{7}$$

Where, θ_{Ai} is actual value, θ_{Pi} is predicted value and N is total number of observations.

5. RESULT ANALYSIS AND DISCUSSION

5.1 *Model Setting*

The Random Forest model is implemented on the actual wind speed data with meteorological variables as shown in Fig. 3 and Fig. 4(a) and 4(b). The RF model is constructed on the parameters given in Table 3.

Table 3. Model Setting for Random Forest Model

Model	Parameters
Random Forest	No. of Trees N_{tree} = 100
	No. of variables in each split = 3
	Minimum Node size = 5

5.2 *Pre-processing of Data*

Each input data set of wind speed and meteorological variables is converted into a unified MATLAB structure file. Table 4 displays the data that was taken into account for the structure file. This data features are applied for prediction of the wind speed for day ahead and one day.

Table 4. Input Data for Wind Speed Prediction for RF Learning Model

Input	Parameters
1	Average of weekly wind speed data of Bengaluru city from January 2011 to December 2021 i.e., for 10 years (52 weeks*10 years=52 weeks data)
2	Average of weekly Dew Point statistics of Bengaluru city from January 2011 to December 2021 i.e., for 10 years (52 weeks*10 years=52 weeks data)
3	Average of weekly Dry Bulb Temperature statistics of Bengaluru city from January

	2011 to December 2021 i.e., for 10 years (52 weeks*10 years=52 weeks data)
4	Week dates from January 2011 to December 2021.

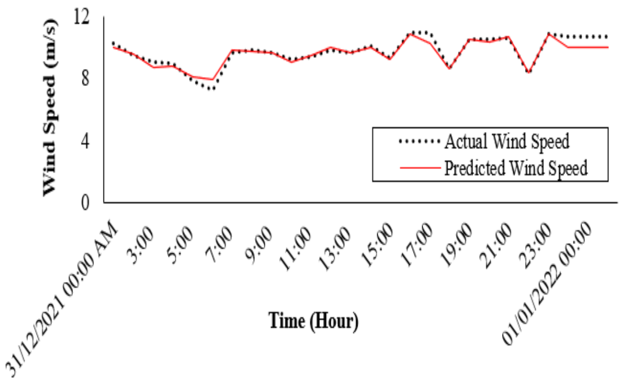
Table 5. Dataset considered for Wind Speed Prediction for RF Learning Model

S. No.	Prediction	Training Dataset	Validation Dataset	Testing Dataset
1	One hour	1st January 2011 to 30th December 2021 (i.e., 96384 hours)	31st December 2021 (i.e., 24 hours)	31st December 2021 (i.e., 24 hours)
2	One day	1st January 2011 to 24th November 2021 (i.e., 4010 days)	25th December 2021 to 31st December 2021 (i.e., 7 days)	25th December 2021 to 31st December 2021 (i.e., 7 days)

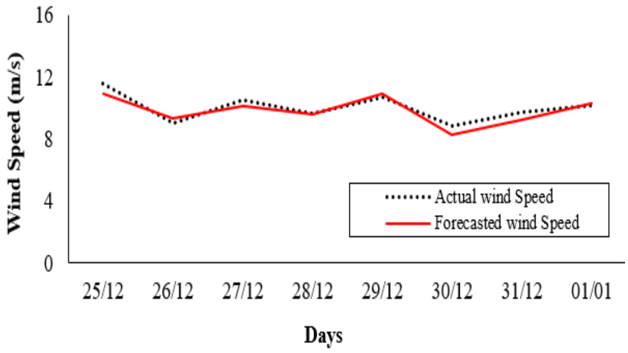
The above dataset considering various meteorological variables are separated into two datasets: training dataset and testing dataset. Table 5 displays the data division. The training data for one-hour prediction is collected from 1 January 2011 to 30 December 2021, or 96384 hours, while the validation and testing data is taken from 31 December 2021, or 24 hours. For one-day prediction, training data are obtained from 1 January 2011 to 24 December 2021, or 4010 days, and validation and testing data are gathered from 7 days. The training dataset of the RF model are utilized to generate a number of trained forecasters and define the optimum forecasted parameters.

5.3 Model Forecasting Results

The prediction of wind speed for short-term using the RF Model for 1 hour and 1 day is depicted in Fig. 5(a)-5(b). The pictorial representation displays the comparison between real time wind speed data with forecasted wind speed data.



(a) Wind Speed Prediction for one hour



(b) Wind Speed Prediction for one day

Fig. 5. Wind Speed Prediction using RF Learning Model

Table 6. Statistical Analysis of Wind Speed Prediction using RF Model

Wind Speed Prediction	MAE	MAPE	RMSE	NRMSE
For one hour	0.19	2.02%	0.39	0.04
For one day	0.36	2.40%	0.57	0.06

The statistical characteristics obtained from the RF model for predicting wind speed for one hour and one day indicates that MAPE is 2.02% and 2.42%, RMSE is 0.39 and 0.57, MAE is 0.19 and 0.36 and NRMSE is 0.04 and 0.06 respectively as shown in Table 6.

5.4 Comparative Analysis of various prediction methods

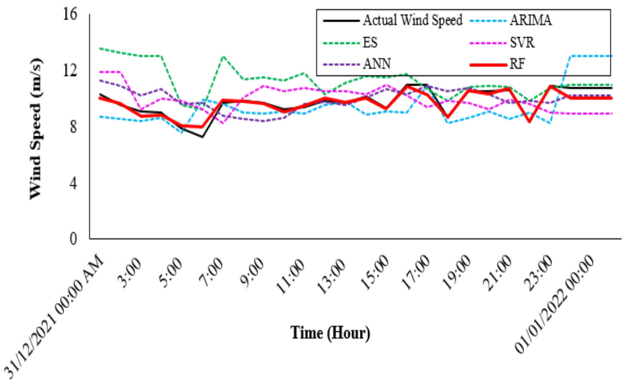
The wind speed prediction of one hour and one day obtained from the RF model is compared with other old and innovative predicting models to determine the effectiveness of the proposed RF method. The traditional methods consist of: Auto Regressive Moving Averages (ARIMA) and Exponential Smoothing (ES). The advanced methods considered for comparison are SVR and ANN. Each method is applied to the data considered for prediction of the wind speed. The comparison between actual wind speed and predicted wind speed obtained from each method is shown in Fig. 6(a)-6(b).

The results shows that wind speed forecasted for one day delivers maximum MAPE for ES method 14.35% in comparison to other techniques such as ARIMA gives 11.24%, SVR gives 12.29%, ANN gives 9.61% and RF model gives 2.02% of MAPE. The Exponential smoothing works best with data that has no trend or seasonal components.

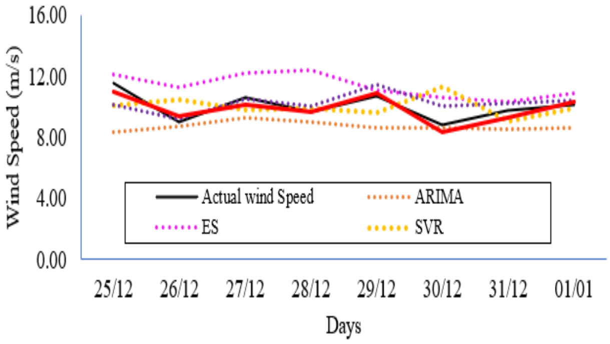
Since the data exhibits trends or seasonality and also the simple form of exponential smoothing is considered. Therefore, the MAPE is higher in ES than other techniques for one day wind speed forecasting. The MAPE obtained for week ahead forecasting by implementing various methods such as ARIMA is 13.52%, ES is 11.41%, SVR is 10.44%, ANN is 6.98% and RF is 2.40%.

The analysis shows that in comparison to one hour, the day ahead wind forecasting has maximum MAPE values except proposed model i.e., RF. This is due to that hourly data typically shows more short-term fluctuations and volatility than daily data. Also, the events or anomalies (e.g., sudden spikes in demand, temperature changes, or traffic) are more common and pronounced in hourly time frames.

Table 7 tabulates the statistical evaluation of the outcomes derived from different approaches. According to the comparison analysis, the results from more conventional approaches, such ES and ARIMA, yield higher MAPE than those from advanced approaches. Additionally, modern approaches can gather meteorological data for analysis and prediction, whereas older methods only use time series data. As a result, advanced techniques outperform traditional ones, with the RF method performing the best.



(a) Wind Speed Prediction for one hour



(b) Wind Speed Prediction for one day

Fig. 6. Comparison between Actual wind speed and Predicted wind speed using various techniques

Table 7. Comparison between wind speed prediction using various techniques.

Techniques	MAE		MAPE		RMSE		NRMSE	
	one hour	one day	one hour	one day	one hour	one day	one hour	one day
ARIMA	1.02	1.20	11.24%	13.52%	0.91	1.03	0.09	0.10
ES	1.68	1.31	14.35%	11.41%	1.19	1.08	0.13	0.11
SVR	1.22	1.07	12.29%	10.44%	1.06	0.97	0.11	0.10
ANN	0.94	0.74	9.61%	6.98%	0.90	0.78	0.10	0.08
RF	0.19	0.36	2.02%	2.40%	0.39	0.57	0.04	0.06

6. CONCLUSION

The RF learning model has shown itself to be an effective and successful method for wind speed prediction when taking into account a variety of meteorological factors. This model utilizes the benefits of ensemble learning and addresses the complexity present in meteorological data with remarkable accuracy. The model's predictive ability is enhanced by a number of meteorological factors, including drybulb temperature, relative humidity and wind direction. Together, these components offer a more comprehensive understanding of the factors influencing wind speed. Due to its superior ability to capture non-linear correlations and interactions between different parameters, the Random Forest model produces predictions that are correct.

The ongoing advancements in machine learning and the accessibility of more varied and high-resolution meteorological variables portend well for the possibility of further improvements in wind speed prediction. Researchers and practitioners should continue exploring new approaches and look into integrating many data sources to increase the precision and consistency of wind speed forecasts.

REFERENCES

[1] J. Jenis et. al., "Engineering applications of Artificial Intelligence in mechanical design and optimization, volume 11, no. 6, pp 577, 2023.

[2] S Giftson Samuel et. al., "Improved Prediction of Wind Speed Using Machine Learning", 1964, 2021, 052005, IOP Publishing.

[3] Wen-Yeau Chang "A Literature Review of Wind Forecasting Methods" Journal of Power and Energy Engineering, 2, 161-168, April 2014.

[4] X. Qin, C. Jiang and J. Wang, "Online clustering for wind speed forecasting based on combination of RBF neural network and persistence method," Chinese Control and Decision Conference (CCDC), Mianyang, China, 2011, pp. 2798-2802.

[5] Daniele Lagomarsino Oneto et. al, "Physics informed machine learning for wind speed prediction", Energy, Volume 268, 2023, 26628, ISSN 0360-5442.

[6] Haoyin Ye et. al, "Wind Speed and Power Prediction Approaches: Classifications, Methodologies, and Comments", Front. Energy Res., Sec. Process and Energy Systems Engineering, Volume 10, 28 April 2022.

[7] Long Wang et al., "Short-Term Wind Speed Forecasting Based on Spatial-Temporal Graph T/F networks", Energy, vol. 253,2022,124095, ISSN 0360-5442.

[8] Madhiarasan, M., "Accurate prediction of different forecast horizons wind speed using a recursive radial basis function neural network", Prot Control Mod Power Syst, vol 5, pages 22, 2020.

[9] J. Vaish, et. al., "short term wind speed forecasting using ANN and ensemble models considering solar irradiance," 2020 ICE3, Gorakhpur, India, 2020, Page. 44-48.

[10] Vaish, J, et al. (2022). "Empirical mode decomposition with random forest model based short term load forecasting. distributed generation &Amp; Alternative Energy Journal, 37(4), 1159–1190.

[11] H. Bevrani et al., "A Fuzzy Inference Model for Short-Term Wind Speed Forecasting," 2012 2nd Iranian Conference on Renewable Energy and Distributed Generation, Tehran, Iran, 2012, Pages. 39-44.

[12] Chauhan M., Gupta S., and Sandhu M., "Short term electric wind speed forecasting using support vector machines (SVM), ECS Transactions, Volume 107, No. 1, Pages 9731-9737, 2022.

[13] Kim J., Afzal A., Kim HG et. al., "Wind power forecasting based on hourly wind speed data in South Korea using Machine Learning Algorithm", Journal of Mechanical Science & Technology, pages 6107-6133, 2022.

[14] Liu, Zongxu, Hui Guo et. al., "A comprehensive review of wind power prediction based on machine learning: models, applications and challenges", Energies, Volume 18, No. 2, 350.

[15] I. D. Mienve and Y. Sun, "A survey of ensemble learning: concepts, algorithms, applications and prospects", IEEE access, volume 10, pages 99129-99149, 2022.

[16] Mohd. Savargiv et. al., "A new Random Forest algorithm based on learning automata", Computational intelligence and neuroscience, pages 19, 2021.

[17] Leo Breiman, "Random Forests", Machine Learning, 45, pages 5-32, 2001.