



Research Article

Time Series Prediction of Wind Power using Autoregressive- Moving Average Model

Vinitha Varghese^{a*}, Divya P. S.^b, and Amala Linus^c

^{a,b,c}Department Of Mathematics, Karunya Institute of Technology and Sciences,
Karunya Nagar, Coimbatore, 641114, Tamil Nadu, India.

*Corresponding author (E-mail: vinithavarghese88@gmail.com)

Received: 15/12/2025, Accepted: 08/03/2026, Published(online): 31/05/2026

Abstract

Proper forecasting of wind is a significant requirement for an effective way of incorporating renewable energy in contemporary power systems. The nonlinear and highly changing nature of wind speed is difficult to forecast, particularly when it comes to models of the wind power curve, which is a union of wind speed and power output. This paper uses the Auto Regressive Moving Average (ARMA) model to predict the wind power curves based on the past wind speed-power records. Being a time series model, the ARMA model is used to represent the time variation in the data and offer predictions with a statistical basis. The standard error measures that were used to evaluate model performance with Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). Conclusions provided find that, using the ARMA model can give credible predictions of wind power curves; it is a simple, although effective method to forecast wind energy. The results not only emphasize the Autoregressive Moving Average Modeling as a statistical tool to integrate renewable energy, but also establish the basis to do more research on it, which may include some hybrid and ensemble approaches to represent an even better forecasting pinnacle.

Keywords: Wind power curve; ARMA; MAE; RMSE; Time series.

1 Introduction

The world energy environment is experiencing a paradigm shift, with countries all over the world investing in minimizing carbon emissions and switching to clean energy sources. Wind energy is one of the most promising renewable energy technologies whose installations have been expanding exponentially over the last ten years. Nevertheless, the nature of wind power variability and intermittency is a major challenge to the operators of power systems, especially related to grid stability and reliability of the electricity supply. Proper forecasting of wind power is important in a number of operational and economic factors. It allows grid operators to effectively meet supply and demand, it allows grid operators to meet higher penetrations of wind energy into current power systems, it allows optimal unit commitment and economic dispatch decisions, and it reduces the reserve capacity requirement, which is expensive. In spite of its significance, wind power forecasting is a difficult task owing to dynamism in the wind speed, the atmospheric dynamics, and the nonlinear correlation between wind speed and power output.

The wind power curve is the graph between the speed of wind and the electrical power of a wind generator. This is a nonlinear relationship which has three different areas: the cut-in region, in which it starts generating power; the operational region, in which the power grows roughly as a cube of the wind speed; and the rated region, where the power is restricted to something to prevent components from being destroyed by the turbines. At some point, the speed of the wind (cut-out speed) causes turbines to be stopped and shut down, and produce no power. This power curve is crucial to the forecasting of wind power generation because it needs to be understood and correctly modelled. This complexity, along with time-dependent changes in wind patterns, makes the use of advanced methods of modeling, which can handle the nonlinear nature of the results as well as the time-dependent correlation in data, a necessity.

The forecasting of wind power has been growing in importance with regard to the integration of grids and the management of energy as renewable energy gains popularity in the world's electricity markets. Short-term and medium-term prediction of wind power has received a significant amount of research interest on autoregressive-moving average (ARMA) models and variants due to their statistical rigor and computational efficiency. ARMA model is a synthesis of two basic elements in time series, namely autoregressive (AR) element, the variable by a linear equation of its previous values to the past, and the moving average (MA)

element, which explains the dependence between the observation and a residual error, which is a result of past predictions. The overall ARMA (p, q) model can be defined as a stationary stochastic process that is appropriate to consider temporal relationships in wind power generation data (Box et al. 2015). Conventional ARMA models involve stationarity, which usually involves differencing or transformation of wind power data. This limitation gave rise to the Autoregressive Integrated Moving Average (ARIMA) model, which uses differencing as an adjunct to the time series method to deal with non-stationary time series that are a common occurrence in meteorological and energy production data (Kavasseri and Seetharaman 2009). A number of scholars have reported the usefulness of ARMA-based models in the prediction of wind power over short periods of time. Research on hourly and day-ahead forecasting has revealed that ARMA models have the capability to reproduce the autocorrelation structure of wind power time series, especially on prediction horizons up to one hour to several hours. The models work best in the cases where wind turbulence is comparatively constant and it follows foreseeable trends (Erdem and Shi 2011). Studies that have been conducted to compare ARMA with persistence, and other statistical models have shown that ARMA models perform better than naive forecasting methods particularly when the time horizon of the prediction is more than one hour. This is because the model was able to observe the temporal correlation structure as well as random variation in the wind power generation (Cadenas and Rivera 2010).

One of the important things about ARMA modeling of wind power is to decide on the right model orders (p, q). In order to balance between model complexity and predictive accuracy, studies have used different information criteria such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Autocorrelation function (ACF) and partial autocorrelation function (PACF) have been very popular as diagnostic techniques in determining appropriate ARMA structures (Sideratos and Hatzigiorgiou 2007). The effects of parameter estimation methods especially maximum likelihood estimation and least squares methods have been researched in respect to the effects they create in forecast accuracy. Studies have shown that existence of accurate parameter estimation plays a major role in the predictive ability of the model, where caution should be taken to prevent overfitting particularly using limited data to train the model (Kavasseri and Seetharaman 2009).

In appreciation of the fact that pure ARMA models are not very good at non-linear relationships and complex patterns, scientists have come up with hybrid models that incorporate ARMA and other forecasting models. These encompass

ARMA-neural network hybrids, wavelet-ARMA hybrids and ensemble techniques that combine the capabilities of two or more forecasting paradigms. These hybridizations have shown to be more accurate than standalone ARMA models especially in difficult forecasting case when the volatility or irregularity are very high (Zhang 2003).

ARMA models have a number of benefits in wind power prediction, such as efficiency in computation, theoretical explainability as well as provision of probabilistic forecasts using prediction intervals. They are parametric and therefore can be easily implemented and have relatively lower computational demands than the more complex machine learning methods. Moreover, ARMA models use less training data compared to most other techniques but have a reasonable degree of accuracy at the short-term predictions (Cadenas and Rivera 2010). ARMA models have a number of limitations in applications in wind power forecasting although they are useful. Basic ARMA models assume that the data used are stationary but this is not always the case when it comes to wind power data that has trends or structural changes. The linear nature of the models restricts the capability of the models to represent non-linear relationships between the variables which may be of great importance in wind power generation. Moreover, ARMA models do not normally include exogenous forecasts like meteorological forecast, which may boost prediction efficiency (Erdem and Shi 2011).

The studies have also been conducted on how ARMA performs when the characteristics are of wind regimes. Findings suggest that the performance of models is different depending on site specific features, where ARMA models perform well in sites where wind patterns are more regular and extreme variability is less pronounced (Cadenas and Rivera 2010).

ARMA models and its forms are a classic method of wind power prediction which provides a compromise between ease and prediction ability of short-term predictions. Although there are more sophisticated techniques that have since been introduced, ARMA models are still useful especially where they are under hybrid framework or computation power is of key importance. Further investigation of adaptive applications and improved preprocessing methods will extend the applications of the ARMA-based methods to the dynamic field of wind power predictions.

In this research, the researcher examines how the Auto Regressive Moving Average (ARMA) model can be used to forecast wind power curves. The specific objectives are:

- To create an ARMA-based forecasting model of wind power curves based on

the previous wind speed and power output data.

- To measure the efficiency of the ARMA model based on the standard error measures (MAE and RMSE).
- To determine the accuracy and feasibility of the ARMA model in predicting wind energy.

The rest of this paper is organized in the following structure: The methodology and model development are described in Section 2. The results and performance evaluation are discussed in section 3. At last, Section 4 is a conclusion of the paper and a proposal of the way out, outlining the future research.

2 Developmental Framework of the Model

ARMA modeling of wind power forecasting has a number of major stages. Exploratory Data Analysis means the analysis of the statistical characteristics, distributions, and time series trends of the wind power, which is done primarily. The Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) plots are used to identify the ARMA (p, q) model order. Diagnostic plots are used for the determination of the correlation structure and for suggesting the right model orders. There were a few candidate ARMA models that were considered with varying (p, q) values. The selection was done using the computation of the criteria of information, such as Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) were used to determine the process of model selection. The model with the minimum information criteria was chosen. Also, the Maximum Likelihood Estimation (MLE), is used for model parameter estimation, when estimating the model coefficients. The Estimation process provides the best coefficients, which minimize the errors in prediction on the training data. Again, to test the model suitability, that is, validation of the model, we applied the model residual diagnostic test. Thorough diagnostic checks were done to determine whether the residuals were random, had constant variance, and a normal distribution using Ljung-Box for residual autocorrelation. The use of residuals to ensure randomness in the model is done to ascertain that all systematic patterns in the data have been projected in the model. As soon as the model had been tested for validation, it was applied to provide forecasts during the test time. One-step-ahead forecasting strategies were adopted to measure the performance of the models at various forecasting horizons.

2.1 ARMA Model Structure

Time-varying parameters of a series are the basis of the time series analysis. Time series analysis is the analysis of a series of data that are recorded at a certain time interval in order to detect patterns, trends, and dependencies. In the case of wind power forecasting, time series analysis is especially applicable since the pattern of wind and power generation has a time-dependent relationship, and accumulated historical values have an impact on the current value of wind and power generation. Among the major concepts in time series analysis are stationarity (statistical properties of a series that remain unchanged across time), autocorrelation (correlation of a series with previous values of the series), and seasonality. Time series data should usually be tested against these properties, and transformed in case they do not fit the model assumptions, before using forecasting models.

The Auto Regressive Moving Average (ARMA) Model is comprised of two basic blocks, the Auto Regressive (AR) block and the Moving Average (MA) block.

Auto Regressive (AR) Component: The current value is described in the AR (p) model as a linear combination of the a priori p past values and a random error:

$$w_t = \phi_0 + \phi_1 w_{t-1} + \phi_2 w_{t-2} + \dots + \phi_p w_{t-p} + \varepsilon_t \quad (1)$$

where:

- w_t is the wind power at time t
- ϕ_0 is the constant term
- $\phi_i, i = 1, 2, 3, \dots, p$
- p is the order of the autoregressive model
- ε_t is the error term with zero mean and constant variance

Moving Average (MA) Component: The MA (q) model represents the present value as being a linear sum of the present and q past error terms:

$$w_t = \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2)$$

where:

- w_t is the wind power at time t
- θ_0 is the mean value of the process given by, $\theta_0 = E[w_t]$
- $\theta_i, i = 1, 2, 3, \dots, q$ is moving average parameters that capture the effect of past disturbances.
- q is the order of the moving average model, i.e., the number of lagged errors considered.
- ε_t is the white noise, i.e., error term with zero mean and constant variance.

Adding both components makes us arrive at the ARMA (p, q) model:

$$w_t = \phi_0 + \phi_1 w_{t-1} + \phi_2 w_{t-2} + \dots + \phi_p w_{t-p} + \varepsilon_t + \theta_0 + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3)$$

where:

- w_t represents the explanatory wind power
- w_{t-i} represents the i^{th} predictor of wind power
- ϕ_0 denote the constant term (mean value)
- ϕ_i denotes the power coefficient,
- ε_t denotes the white noise
- ε_t temporary random shocks that cause deviations from the mean (θ_0)
- θ_0 denotes the mean value of the process
- θ_i symbolizes the error coefficients.

2.2 Data Description

The research makes use of past data on wind speed and power production assembled at a wind farm location. The dataset is taken from 01 January, 2017 to 31 December, 2021, where measurements are taken on daily basis. The dataset includes Wind speed(m/s)- wind speed_10m, wind speed_100m, Wind power output kW), temperature_2m ($^{\circ}C$), relative Humidity_2m (%), wind gusts_10m(m/s), dewpoint_2m($^{\circ}C$), wind direction($^{\circ}$)- wind direction_10m, wind direction_100m.

2.3 Data Preprocessing

There were a few preprocessing steps carried out before the development of the model:

- Missing data treatment: There are missing values that may have been created by a sensor failure, or a loss of communication, any of which were interpolated or deleted.
- Outlier Detection: The presence of extreme values in the data caused by error in measurements or extreme events was found by the IQR method.
- Data Normalization: In order to achieve good model convergence and performance, the data were normalized by Min- Max scaling (or Z-Score scaling).

2.4 Stationarity Testing

ARMA models need time series that are stationary. It was evaluated using Augmented Dickey-Fuller (ADF) test. In case non-stationarity had been identified, relevant transformations were used for differencing or detrending.

2.5 Data Splitting

The data set was divided into 79.7% and 20.3% as training and testing sets. Model parameter estimation was done in the training set, and the forecasting performance was done in the testing set.

3 Results of Analysis

The implementation of ARMA on the wind power curve forecasting implies considering the power output time series as the variable to be modelled. The AR and MA components can be used to represent the temporal dependencies in wind power generation, which is influenced by the atmospheric stability, the diurnal characteristics, and the meteorological conditions. Although the power curve itself is not linear with respect to wind speed, the time series of power output can have linear time dependencies, and ARMA can be useful to model them.

3.1 ACF and PACF Plots

The plot of the Autocorrelation Function (ACF) has a slow, gradual decay and not a steep cutoff. A tapering pattern like this normally indicates that there is

an autoregressive component as opposed to a pure moving average component. Where the ACF abruptly diminishes following a specified lag, the lag normally represents the order of the moving average. Since the observed ACF decreases slowly, as shown in Figure 1, the moving average order is presumed to be minimal. This is typical of the ARMA models, where a gradual decay in ACF is normally observed.

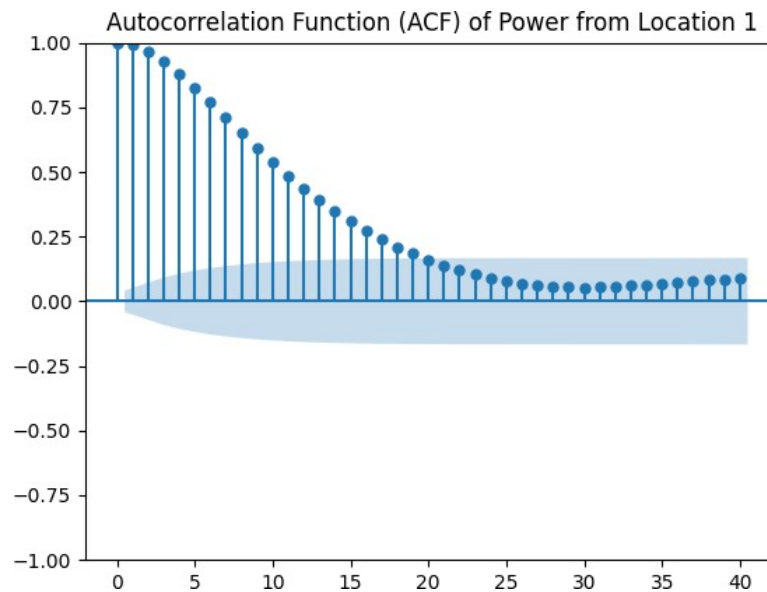


Figure 1: ACF of wind power

The plot of the Partial Autocorrelation Function (PACF) given in Figure 2 indicates great spikes at the first lags, and values decrease to nearly zero past about lag 2 or 3, with the remainder of the values confined within the confidence intervals thereafter. This trend shows that the order of autoregression, falls in the range of between 1 and 3.

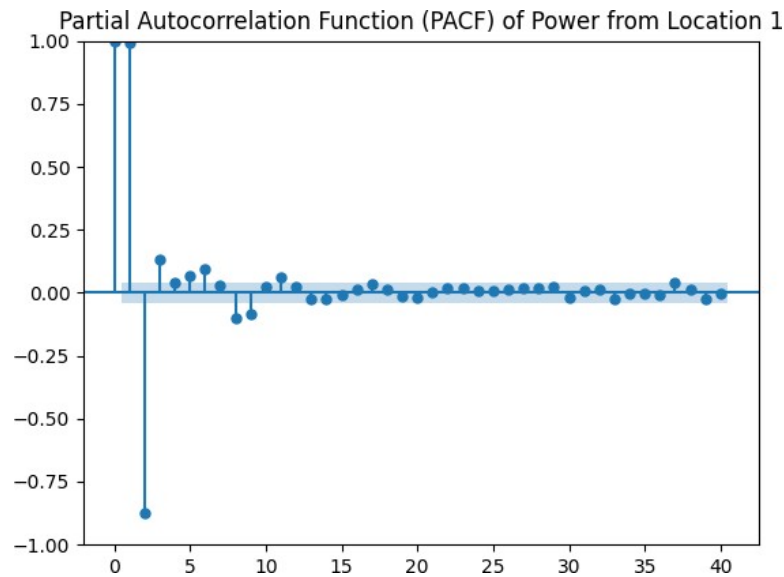


Figure 2: PACF of wind power

On the whole, the above observations indicate taking the ARMA models having an autoregressive order p of 1 to 3 and a moving average order q of either 1 or close to 0 at the beginning stage. These order selections can then be narrowed down with future modeling and information criteria comparisons such as AIC or BIC.

3.2 AIC – BIC Criteria Evaluation

The criteria of information were used to determine the process of model selection with the ARMA (p , q). Table 1 shows the AIC and BIC values for different ARMA models at different orders of models.

Table 1: Model Selection Criteria

Model	AIC	BIC
ARMA (1,1)	-13450.52173	-13432.52773
ARMA (1,2)	-13465.43851	-13441.44651
ARMA (1,3)	-13467.72493	-13437.73493
ARMA (2,1)	-13470.93664	-13442.94464
ARMA (2,2)	-13472.71942	-13442.72942
ARMA (2,3)	-13468.12649	-13432.13849
ARMA (3,1)	-13473.48404	-13443.49404
ARMA (3,2)	-13466.54731	-13430.55931
ARMA (3,3)	-13469.29416	-13427.30816

The ARMA (3,1) model possesses the least AIC value as well as the BIC value, which indicates the ARMA (3, 1) model is best suited for forecasting wind power for the data we have. Table 2 gives a statistical summary for the selected ARMA (3,1) model.

Table 2: Statistical Summary of ARMA(3,1)

Statistic	Value
Dependent Variable	power
Number of Observations	2208
Model	ARMA(3,1)
Log Likelihood	6742.742
AIC	-13473.48404
BIC	-13443.49404
HQIC	-13460.99
Covariance Type	Outer product of gradient (opg) estimator
Sample Range	2021-10-01 to 2021-12-31

3.3 Parameter Estimates

The parameters that are estimated in the chosen ARMA model are given in Table 3

Table 3: ARMA (3,1) Parameter Estimates

Parameter	Coefficient	Std Error	z-value	P> z	95% Confidence Interval
const	0.4571	0.03	15.328	0.0023	[0.399, 0.516]
ar.L1	0.482	0.103	4.679	0.000	(0.280, 0.684)
ar.L2	-0.9145	0.109	-8.388	0.0351	[-1.128, -0.701]
ar.L3	0.035	0.055	0.636	0.0502	[-0.073, 0.143]
ma.L1	0.095	0.104	0.917	0.359	(-0.109, 0.299)
σ^2	0.001	0.00003	33.22	0.000	(0.001, 0.001)

All parameters were statistically significant at the level of 5 percent, and this indicates that they are significant in the forecasting model.

3.4 Diagnostic Test Results

The diagnostic statistics for the ARMA (3, 1) Model are given in table 4. Statistic Q of Ljung- Box is 0.9 with a p-value of 0.01. This implies the existence of an autocorrelation. Jarque-Bera (JB) value and large p-value are the concluding factors that the data is normally distributed.

Table 4: Diagnostic Statistics ARMA(3,1)

Test	Statistic	p-value
Ljung-Box (L1) (Q)	0.9	0.015471
Jarque-Bera (JB)	1.4187	0.6366
Heteroskedasticity (H)	0.9	0.7345
Skewness	0.037	-
Kurtosis	2.479	-

3.5 Forecasting Performance

Root Mean Square Error and Normalized Mean Absolute Error were the error measures used in testing the performance of the forecasting during the test data set. The findings are given in Table 5, and the results indicate that the ARMA model has more accuracy in curve forecasting of wind power.

Table 5: Root Mean Square Error and Normalized Mean Absolute Error

Error Metrics	Value
RMSE	0.0146
NMAE	0.0134

3.6 ARMA Forecast

Figure 3 gives the comparison of actual power and predicted power using the ARMA model.

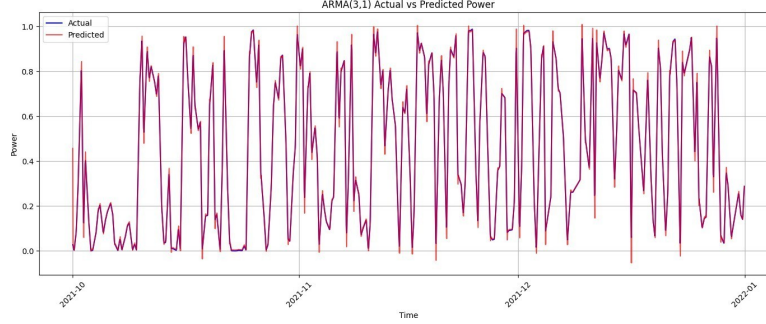


Figure 3: Actual Power vs Predicted Power using ARMA (3,1)

3.7 Interpretation and Discussion

The findings reveal that the ARMA model is a good predictor of wind power curves, and its error is acceptable. The NMAE of 0.0134 and RMSE of 0.0146 indicate that the model includes the most significant trends in terms of time in the wind power generation. The comparatively small error values indicate that, although the wind power curve is nonlinear in regard to wind speed, the temporal dynamics of power output have patterns, that can be adequately explained with linear time series models.

The ARMA model has a number of benefits in wind power prediction. ARMA model is mathematically well-posed, and parameters can be interpreted. This openness has been useful in operational uses where it is significant to know how models behave. Computational Efficiency ARMA models is another advantage, since it requires lower computational needs than complex machine learning models to train and make predictions, and therefore they are used in applications where real-time forecasts are needed. ARMA models offer a statistically rigorous model with formal hypothesis test, confidence interval as well as Model validation diagnostic tools. The outcomes suggest that the ARMA models are especially useful when it comes to short-term forecasting horizons, which tend to be the most important ones in the grid operations and energy trading.

Because of its advantages, the ARMA method has its limitations. ARMA models assume that there are linear relationships in the time domain that can be inad-

equate to represent the complex nonlinear dynamics in the generation of wind power. Data that are stationary may be expensive to preprocess and transformation may have repercussions to interpretation. Standard ARMA models do not directly use external variables like weather predictions that tend to increase the accuracy of prediction.

4 Conclusion

The proposed research was able to prove the use of ARMA modeling in predicting wind power curves. The key findings are that the ARMA model is a good representation of the wind power generation data in terms of temporal dependencies and it can give good forecasts with acceptable levels of error (NMAE: 0.0134, RMSE: 0.0146). The model is suitable especially in short-term horizons of forecasting, which is most important in operational decision-making in power systems. Although there is not a linear relationship between wind speed and power output, power generation temporal patterns can be successfully predicted by the linear ARMA framework. The ARMA models are also simple and computationally efficient and thus can be used in real-time forecasting in wind energy systems.

Forecasting of wind power has been a major problem with regard to the integration of renewable energy sources, and this paper has shown that the application of rigorous statistical techniques such as ARMA modelling, are still a useful solution. Though today there is a growing interest in complex artificial intelligence approaches to forecasting energy, this study emphasizes that simpler, well-established statistical approaches can be reliable, give interpretable and computationally low forecasts. The ARMA model is a useful instrument in operational forecasting and a strong foundation in the assessment of more advanced tools. The diversity of forecasting tools, including standard statistical and the latest machine learning, will be fundamental to ensuring that the different stakeholders in the energy ecosystem have their needs addressed, given that the renewable energy industry is steadily growing.

References

- [1] Box G E P, Jenkins G M, Reinsel G C and Ljung G M. *Time Series Analysis: Forecasting and Control*. 5th ed. Wiley, Hoboken. 2015.

- [2] Cadenas E and Rivera W. Wind speed forecasting in three different regions of Mexico, using a hybrid ARIMA-ANN model. *Renewable Energy*. 35(12) (2010) 2732–2738.
- [3] Erdem E and Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. *Applied Energy*. 88(4) (2011) 1405–1414.
- [4] Kavasseri R G and Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. *Renewable Energy*. 34(5) (2009) 1388–1393.
- [5] Sideratos G and Hatziargyriou N D. An advanced statistical method for wind power forecasting. *IEEE Transactions on Power Systems*. 22(1) (2007) 258–265.
- [6] Zhang G P. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*. 50 (2003) 159–175.
- [7] Carta J A, Ramírez P and Velázquez S. A review of wind speed probability distributions used in wind energy analysis: Case studies in the Canary Islands. *Renewable and Sustainable Energy Reviews*. 13(5) (2009) 933–955.
- [8] Jiang W and Hong T. Wind power forecasting based on AR models and improved particle swarm optimization. *Applied Energy*. 238 (2019) 494–506.
- [9] Tuncar E A. A review of short-term wind power generation forecasting methods. *Renewable Energy*. 204 (2024) 672–682.
- [10] Poultourtzidou Z and Jankūnas A. Wind offshore power forecasting: Comparative analysis between autoregressive model and machine learning approaches. M.S. thesis, Department of Energy Sciences, Lund University, Lund, Sweden. 2024.
- [11] Dai J and Fu L. A wind speed forecasting model using nonlinear autoregressive model optimized by the hybrid chaos-cloud salp swarm algorithm. *Energy*. 270 (2024) 126943.
- [12] Rajagopalan S and Santoso S. Wind power forecasting and error analysis using the autoregressive moving average modeling. In *Proceedings of the IEEE Power and Energy Society General Meeting*, Calgary, AB, Canada. 2009. pp. 1–5.

- [13] Erdem E and Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. *Applied Energy*. 88(7) (2011) 1405–1414.
- [14] Hyndman R J and Khandakar Y. Automatic time series forecasting: the forecast package for R. *Journal of Statistical Software*. 27 (2008) 1–22.
- [15] Pant A and Rajput R S. Time series analysis of gold price using R. In *Digitalization in Commerce and Its Impact on Economy*, Social Research Foundation. 1 (2019) 15–24.
- [16] Sameer M, Rajput R K S, Dibyansu, Thurkal S and Bijlwan S. Enhancing rainfall accuracy: A comparative study of ARIMA, ANN, and hybrid models using agro-climatic data. *Proceedings of the Jangjeon Mathematical Society*. 28(4) (2025) 713–725.
- [17] Blanco M I. The economics of wind energy. *Renewable and Sustainable Energy Reviews*. 13 (2009) 1372–1382.
- [18] Burton T, Sharpe D, Jenkins N and Bossanyi E. *Wind Energy Handbook*. 1st ed. John Wiley & Sons, New York. 2011.
- [19] Chang W Y. A literature review of wind forecasting methods. *Journal of Power and Energy Engineering*. 2 (2014) 161–168.
- [20] Chen G, Gan M and Chen G. Generalized exponential autoregressive models for nonlinear time series: Stationarity, estimation and applications. *Information Sciences*. 438 (2018) 46–57.