Very Short term Wind Power Prediction Using Hybrid Univariate ARIMA-GARCH Model

Archee Gupta[1], Kailash Chand Sharma[2], Archita Vijayvargia[1] and Rohit Bhakar[1]
[1]Centre for Energy and Environment, Malaviya National Institute of Technology, Jaipur, India
[2]Dr. B R Ambedkar National Institute of Technology, Jalandhar, India
2018PCV5364@mnit.ac.in, kailashsharma3889@gmail.com, 2017PCV5168@mnit.ac.in, rbhakar.ee@mnit.ac.in

Abstract— **The integration of wind power generation with the grid reflects many challenges to the utility and market operators. For this purpose, very short-term Wind Power Prediction (WPP) has become an indispensable requirement for efficient power systems operations. Typically, statistical time series models like Autoregressive Integrated Moving Average (ARIMA) is widely used for WPP. Although ARIMA is capable of capturing conditional mean appropriately. However, it does not cover the time-varying volatility present in wind power. Hence, for more precise parameter estimation and forecasting, this paper utilizes the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. To fortify the forecasting model, this paper proposes a Hybrid ARIMA-GARCH model, incorporating the assets of ARIMA and GARCH model, for forecasting of wind power. The proposed model combines the ARIMA based conditional mean forecast and GARCH based conditional variance forecast. The propounded model is implemented on three wind farms located in Australia. The simulation results obtained shows that the proposed model works better than conventional ARIMA Model in term of Mean Absolute Error (MAE) and Root Mean Square Error (RMSE).**

Keywords—**ARIMA, GARCH, Hybrid ARIMA-GARCH, Statistical Forecasting, Time series, Wind Power Prediction.**

I. INTRODUCTION

The penetration of Renewable Energy Sources (RES) into the existing power grid is gaining much focus. The main reason being that the effective penetration of these RES fortifies grid sustainability and reduces the pressure on existing conventional energy sources. Wind energy being one of the RES, has the potential to fulfill about 35% of the global energy demands [1]. However, a high level of wind energy penetration to the grid imparts a factor of uncertainty on the grid. It further introduces severe challenges concerning load following, unit commitment or scheduling of generators or storage, market mechanisms, *etc.* to power system planners and utility operators [2].

The deep-rooted cause of these power system operational issues is the chaotic nature of wind energy. Hence, to mitigate these challenges, one of the possible and technically viable solutions is the accurate prediction of the wind power [6]. Several WPP models are available corresponding to the type of power system planning or operation application. In this regard, the WPP models are classified as Very-short term forecasting (minutes-1 hr), Short-term forecasting (1hr-1day ahead), Medium-term forecasting (an hours-1 week ahead), and Long-term forecasting (1 week-year ahead) [3], [12].

Further, on the grounds of the methodology adopted, WPP models are classified into statistical models and physical models. Physical prediction models are based on deterministic methods and use numerical weather prediction model data such as surface roughness, pressure, temperature, scaling of the wind speed within wind farms area [7]. However, these models generally take comparatively longer prediction time, hence does not apt the short-term prediction intervals.

Statistical prediction methods use online measurements of historical wind resource data. Therefore, these models are relatively less complicated, have low computational time, and are economical [13]. A brief classification of statistical models includes a time series approach and an artificial intelligence approach. For accounting the non-linearity of time series data various deep learning and ANN approach has been developed. A deep learning time series forecasting based on LSTMs, SVRM and EO is introduced [16]. A hybrid forecasting system comprises of three models (a data pre-processing module, forecasting module, optimization module) is propounded to enhance the wind speed prediction accuracy [17]. The five different ANN models were developed for predicting wind speed [18]. A hybrid approach is also developed for wind power forecasting using wavelet transform, particle swarm optimization and adaptive network based fuzzy inference system [19]. In [20], wind speed forecasting in 1h, in advance horizon is done using non-linear autoregressive exogenous model. AI methods are competent in dealing with complex and non-linear data but don't give much insight into the structure of model. So, in this paper statistical time series approach is used for accounting the non-linearity present in data.

Some of the time series models are Moving average (MA), Autoregressive (AR), Autoregressive Moving Average (ARMA), ARIMA, and Kalman filter models [10]. Out of these, ARIMA is one of the most influential models to predict wind power. ARIMA model is advantageous in dealing with non-stationary data like wind power. It requires a lesser number of parameters for estimation and is competent in dealing with linearity present in data [11]. Despite having these advantages, the ARIMA model is not able to capture the volatility in data. Volatility in data means the presence of varying conditional standard deviation of wind power time series data. In other words, the ARIMA model counts only

This work is supported by the DST grant for UKICERI project, DST/RCUK/JVCCE/2015/02

conditional mean in time series dataset with an assumption of having constant conditional variance [4].

The study depicts that wind power data shows time-varying volatility in its nature [5]. Therefore, incorporating such volatility into the ARIMA model leads to more precise parameter estimation and WPP. To overcome this issue, the GARCH model and its variants are used for capturing time-varying volatility. GARCH model assumes constant conditional mean and utilizes the significance of non-constant conditional standard deviation for WPP [5]. However, the GARCH model is applicable when residuals do not show significant serial autocorrelation, but the square of residuals shows correlation [4]. To highlight the strengths of ARIMA and GARCH model, a hybrid ARIMA-GARCH model can be developed to forecast wind power more accurately.

In the above context, this paper presents a hybrid ARIMA-GARCH model for short-term WPP. The proposed model combines conditional mean forecast by the ARIMA model and conditional variance forecast by GARCH model. The propounded model is implemented on three wind farms located in Australia. The results obtained reflect the improvement in WPP from ARIMA to ARIMA-GARCH model. Thus, the proposed model seems to be better to the ARIMA model with minimum MAE and RMSE.

The rest of the paper is formulated as follows: Section II discusses the proposed hybrid ARIMA-GARCH model in detail. Section III presents the case study for the propounded model implemented on three wind farms. Finally, conclusions are drawn in section IV based on the performance analysis of various models.

II. PROPOSED HYBRID ARIMA-GARCH MODEL

A brief introduction of ARIMA and GARCH model is also presented showing the significance of incorporating conditional variance for the WPP. Finally, the proposed Hybrid ARIMA-GARCH model is discussed.

A. ARIMA Model

The AR (p) model forecasts the future values using the linear combination of past values of a variable. The MA (q) model forecasts the future error of wind data using a linear combination of past forecasted errors. ARMA (p, q) model, a combination of MA (q) and AR (p) model and is applicable to univariate and stationary time series. It measures the linear relationship between the historical wind speed data and previous forecast error of wind speed data series [14]. Hence, to address non-stationary processes like wind power, the ARIMA model comes into forecasting scenario.

The equation of ARIMA (p,d,q) model is defined as:

$$y_t^d = c + \sum_{i=1}^{p} \phi_i y_{t-i}^d + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{1}$$

ARIMA (p,d,q) model is a combination of AR(p) model, and MA (q) model with finite differencing to make data stationary, *i.e.,* this model can be used with non-stationery data. Here, y_t^d is the observed time series, ϕ_i is the parameter of the autoregressive part and θ_j is the parameter of moving average part. ε_t is a moving average term. ARIMA Model captures the linearities present in wind power time series data. However, study exhibits the non-linearity and volatility present in wind power time series data [9].

B. GARCH Model

GARCH model is used for the time series having its elements that do not conform to a linear pattern such as wind power dataset. GARCH model estimates the conditional volatility of the time series data. It measures the non-linear relationship, and it can best fit data with heavier-tailed error distribution, *i.e*., high standard deviation [7].

GARCH model is an extension of Autoregressive Conditional heteroskedastic (ARCH) model, which incorporates MA term along with the AR component. ARCH model presents the conditional variance as a function of past residual errors only. However, in GARCH variance depends on its past value also. The GARCH model is successful in predicting and estimating volatility changes, *i.e.,* it considers the conditional variance as a function of the square of past error and its own value. Equation (2) and (3) shows the mean and residual error mathematical formulations respectively.

$$y_t = u_t + \varepsilon_t \tag{2}$$

$$\varepsilon_t = \sigma_t v_t \tag{3}$$

Where v_t is independent and identically distributed random number with (0,1). Equation (4) shows the mathematical formulation of the time-varying variance.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha p \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{4}$$

Equation (4) is subjected to the following constraints as mentioned in (5) to (7) to make the y_t stationary.

$$\alpha_0 > 0 \tag{5}$$

$$\beta_j \geq 0 \tag{6}$$

$$\sum_{i=1}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 < 1 \tag{7}$$

where, $\sum_{i=1}^{p} \alpha_t \varepsilon_{t-1}^2$ is a moving average term, sum of the p previous lag of each squared error term, $\sum_{j=1}^{q} \beta_j \sigma_{t-j}^2$ autoregressive term, square of the q previous lag of each variance term.

C. Hybrid ARIMA-GARCH Model

The proposed hybrid ARIMA-GARCH model incorporates the qualities of ARIMA and GARCH for better time series prediction purposes. First, the linearities of wind power time series data are modelled using ARIMA Model and then the non-linear part of the ARIMA model, *i.e.,* residuals, are modelled using GARCH Model. Finally, this hybrid model formed by the combination of univariate ARIMA and GARCH model is used to forecast the wind power time series dataset.

Combining the ARIMA and GARCH models is an effective way overcomes the disadvantages of an individual model [21], [22].

The methodology used to forecast wind power using the proposed hybrid ARIMA-GARCH model is discussed below:

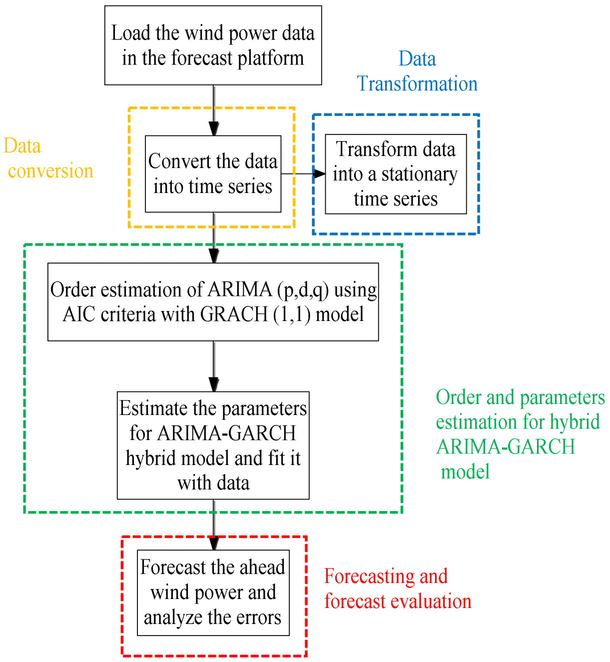

Fig. 1. Proposed hybrid ARIMA-GARCH forecasting model.

Step 1: Data conversion:
The given wind power data is converted into time-series. Then converted data is plotted to identify any non-stationarity.

Step 2: Data transformation:
If any non-stationarity is present in data, then data should be transformed through logs and differencing to make data stationary.

Step 3: Order and parameter estimation:
Order is estimated using Autocorrelation function and partial autocorrelation function plot and AIC criteria. In this paper, the order is estimated through Akaike information criteria and parameter of ARIMA-GARCH model is estimated through Maximum Likelihood Estimation Method.

Step 4: Forecasting:
Estimated ARIMA-GARCH model is fitted to model and forecasting is obtained for 6-step ahead values.

Step 5: Forecast evaluation:
In hybrid ARIMA-GARCH model, dynamic ahead forecast is used. The dynamic forecast means more than the one-step-ahead forecast. Forecasting efficiency is evaluated through two static measures, Root mean square error (MSE) and Mean absolute error(MAPE).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - y)^2}{n}} \qquad (8)$$

$$MAE = \frac{\sum_{i=1}^{n}|y_i - y|}{n} \qquad (9)$$

Where, n = number of observations, y_i is the predicted value and y is the observed value.

A concise review of the methodology used for the proposed model is presented in fig. 1.

III. CASE STUDY

A. Dataset

The forecast models of the ARIMA, hybrid ARIMA-GARCH are evaluated by using data obtained from Australia AEMO-22 wind farms, at an interval of 5 minutes. The model is tested on three wind farms, namely Lake Bonny (Lkbonny), Mount Millar (Mitmillar), Starfish Hill (Starhlwf).

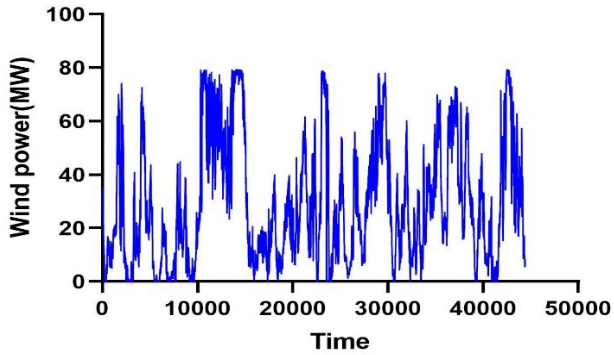

Fig. 2. Wind power time series dataset for the LKBONNEY1 wind farm.

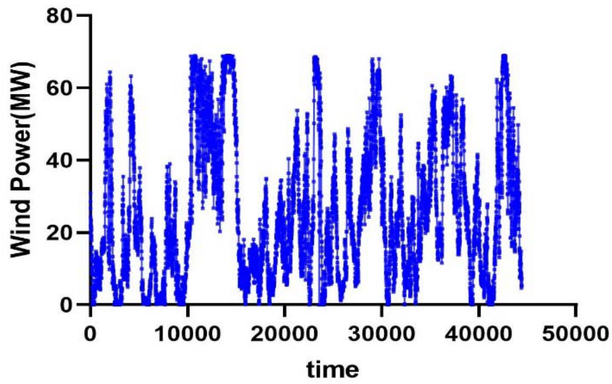

Fig. 3. Wind power time series dataset for the Mount Millar wind farm.

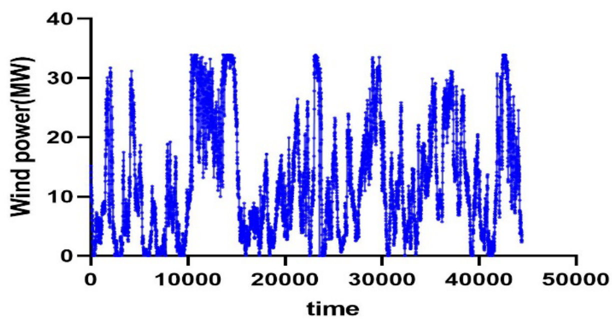

Fig. 4. Wind power time series dataset for the Starfish Hill wind farm.

TABLE I
Estimated parameter of ARIMA-GARCH model for wind farms

Parameter	LKBONNY wind farm			MITMILLAR wind farm			STARFISH HILL windfarm		
	Estimate	Std. error	p-statistics	Estimate	Std. error	p-statistics	Estimate	Std. error	p-statistics
ϕ_1	.982029	.011250	0	.982035	.011250	0	.982043	.011251	0
ϕ_2	-.016601	.017124	.332320	-.016605	.017124	.33220	-.016607	.017124	.332143
ϕ_3	-.023346	.011811	.048088	-.023344	.011811	.048102	-.023344	.011811	.048101
θ_1	-.924191	.005920	0	-.924195	.005919	0	-.924202	.005917	0
α_0	.061132	.005966	0	-.046224	.004511	0	.011228	.001096	0
α_1	.059542	.007023	0	-.175126	.007023	0	.175125	.007023	0
β_1	.174780	.006646	0	.823874	.006646	0	.823875	.006646	0

The registered capacity of each plant is as 80.55 MW, 70 MW, and 34.5 MW. The data set contains 8879 samples out of which 8870 samples are used as training purpose for model fitting and parameter estimation. The next 6 observations are used for validation of the forecasting model. The algorithms are simulated in R studio.

From figs. 2, 3, and 4; it can be seen that wind power data of three of the wind farms are non-stationary, so finite differencing is required to make them stationary. In fig. 5 volatility clustering can be easily seen after taking up the first difference of data. Volatility clustering means large changes come after large changes and small changes come after small changes. So, it ensures the applicability of the GARCH model.

B. *Results and Discussions*

Table I represents the parameter estimation of ARIMA-GARCH model. Parameter estimation is done through maximum likelihood equation. For testing coefficient stability, Nyblom criteria test is performed. Nyblom test assures that parameters are constant with time against alternative that they are time varying [10]. Results show the stability of parameter as an individual statistic of all parameters is less than joint statistics (10.881). Order estimation for ARIMA model is done through AIC criteria, and it is typical to predict the order of the GARCH model, so the standard GARCH model with order (1,1) is used.

The diagnostic checking is done to test the adequacy of the model. For this, Weighted Ljung test is performed on residuals and its squares. This test checks the fitness of the time series model. The null hypothesis is that residuals shows no serial correlation. Result statistics shows that residuals and its squares show no serial correlation after fitting the model. From the obtained parameter estimation, it is observed that all parameters satisfy the constraints present in (5), (6), and (7).

Time series model works with stationary data. However, the advantage of ARIMA Model is that it resolves the issue of stationary data in the model itself. It requires only the no. of differencing required to make data stationary. So Augmented-Dicky Fuller (ADF) test is used to check the stationarity of data [15]. After having one differencing data become stationary as indicated by the ADF test.

TABLE II
Order identification of ARIMA Model

ARIMA(p,d,q)	AIC
ARIMA(3,1,0)	46479.3
ARIMA(3,1,1)	46455.2
ARIMA(3,1,2)	46456.77

From table II, the model with the least AIC values is selected as an appropriate model. So ARIMA (3,1,1) is selected for forecasting. After that, the parameter of the ARIMA model is evaluated using the maximum likelihood estimation method, as shown in Table III.

TABLE III
ARIMA Model and parameters for wind farms

Parameters/ Model	LKBONNY	MITMILLAR	STARFISH HILL
	ARIMA (3,1,1)	ARIMA (3,1,1)	ARIMA (3,1,1)
ϕ_1	.7465	.7465	.7496
ϕ_2	-.0303	-.0303	-.0303
ϕ_3	-.0406	-.0406	-.0406
θ_1	-.7300	-.7300	-.7331
σ	3.3181	2.8858	1.4216

The results obtained from the two methods are presented in the table III. It shows the parameter estimated from the ARIMA model for three wind farms.

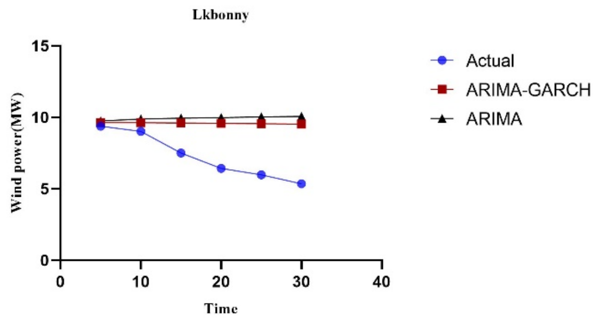

Fig. 5. 6-step ahead forecast for Lake Bonney wind farm.

TABLE IV
Error comparison of ARIMA+GARCH and ARIMA model

Model	Lake Bonney		Mount Millar		Starfish Hill	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
ARIMA	3.2443	2.880	2.5810	2.210	1.2727	1.08936
ARIMA+GARCH	2.7379	2.310	2.3807	2.009	1.17339	.9902515

TABLE V
6- step ahead forecast for wind farms

Time	LKBONNY			MITMILLAR			STARFISH HILL		
	Actual	ARIMA+GARCH	ARIMA	Actual	ARIMA+GARCH	ARIMA	Actual	ARIMA+GARCH	ARIMA
5	9.377483	9.628543	9.749835	8.154333	8.372652	8.461099	4.018921	4.126523	4.217669
10	9.020246	9.632090	9.885987	7.843692	8.375735	8.543941	3.86582	4.128042	4.210942
25	7.501987	9.605230	9.939459	6.523467	8.352378	8.556247	3.215137	4.116531	4.217007
20	6.430274	9.579838	9.98273	5.591543	8.330299	8.559927	2.755832	4.105649	4.218821
25	5.983727	9.555917	10.02938	5.203241	8.309498	8.558938	2.564454	4.095397	4.218334
30	5.358562	9.534126	10.06779	4.659619	8.290549	8.557588	2.296526	4.086058	4.217669

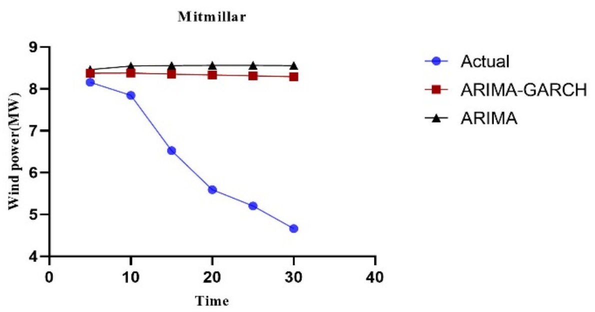

Fig. 6. 6-step ahead forecast for Mount Millar wind farm.

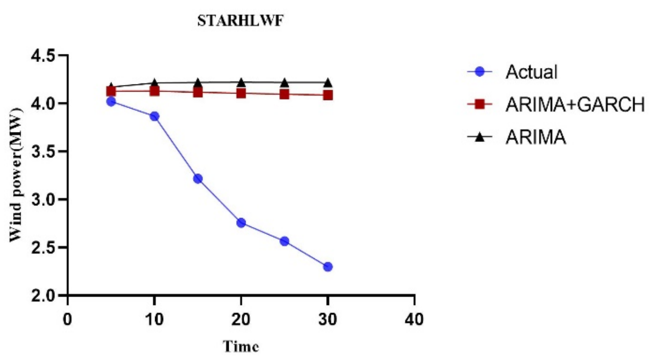

Fig. 7. 6-step ahead forecast for Starfish Hill wind farm.

Results show that the model works well up to 3-step prediction. Graphical comparison of the two models is also shown in figs. 5, 6, and 7.

Table IV shows the error analysis of two models using RMSE and MAE. Table V represents the actual value of three wind farms and the predicted value obtained from the ARIMA and ARIMA+GARCH model.

IV. CONCLUSIONS

Wind power forecasting is beneficial for resolving the grid integration issues that arises due to the integration of intermittent wind power into the grid. In this paper, two WPP models are implemented on 3 wind farms, *i.e.,* ARIMA and hybrid ARIMA+GARCH. This work reveals that the hybrid ARIMA-GARCH performs better than ARIMA Model in terms of accuracy, *i.e.,* ARIMA-GARCH model works well up to the three-step ahead forecast. Determining the prediction interval would be the future scope of the work.

REFERENCES

[1] IRENA, "Future of wind: Deployment, investment, technology, grid integration and socio-economic aspects (A Global Energy Transformation paper)," International Renewable Energy Agency, Abu Dhabi,2019.

[2] T. R. Ayodele, A. A. Jimoh, J. L. Munda, and J. T. Agee, "Challenges of grid integration of wind power on power system grid integrity: A review," *Int. J. Ren. Energy Res*., vol. 2, no. 4, pp. 618–626, 2012.

[3] X. Wang, P. Guoc, X. Huanga, "A Review of Wind Power Forecasting Models," *Energy Pro*., no. 12, pp. 770 – 778, 2011.

[4] C. Francq, J.M. Zakoian, GARCH Models Structure, Statistical Inference and Financial Applications, 2nd ed. Hoboken , NJ: John Wiley and sons, 2019.

[5] P.S. Georgilakis, "Technical challenges associated with the integration of wind power into power systems," *Ren. Sust. Energy Rev*., vol. 12, no .3, pp. 852–863, 2008.

[6] J.Xu, Z.Zhang, L.Zhao and D.Ai, "The application review of GARCH model," *Proc. 2011 Int. Conf. on Multimedia Tech., Hangzhou,* pp. 2658-2662, 2011.

[7] S. M. Lawan, W. A. W. Z. Abidin, W. Y. Chai, A. Baharun and T. Masri, "Different models of wind speed prediction; a comprehensive review," *Int. J. of Scientific & Eng. Res.,* vol. 5, no. 1, 2014.

[8] S. S. Soman, H. Zareipour, O. Malik and P. Mandal, "A review of wind power and wind speed forecasting methods with different time horizons*," North American Power Symposium 2010, Arlington, TX, 2, pp. 1-8, 2010.*

[9] P.L and L.Yue, "Short-term wind speed forecasting based on non-stationary time series analysis and ARCH model," *2011 Int. Conf. Multimedia Tech., Hangzhou*, 2011

[10] J. Nyblom, "Testing for the Constancy of Parameters Over Time*," J. of the American Stat. Asso.*, vol. 84, no. 405, pp. 223–230, 1989.

[11] K. Bhaskar and S.N. Singh, "AWNN-Assisted Wind Power Forecasting Using Feed-Forward Neural Network", *IEEE Tran. Sust. Energy*, vol. 3, no. 2, 2012.

[12] J. Dowell and P. Pinson, "Very-Short-Term Probabilistic Wind Power Forecasts by Sparse Vector Autoregression," *IEEE Tran. Smart Grid*, vol. 7, no. 2, pp. 763-770, 2016

[13] G. J. Osório, J. C. O. Matias and J. P. S. Catalão, "A review of short-term wind power forecasting approaches,"*2nd IET Ren. Power Gen. Conf. (RPG 2013)*, *Beijing*, 2013.

[14] R.J Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 2nd ed. Melbourne, Australia: Otexts, 2018.

[15] W. Enders and Y.Lin, "Inflation Revisited: New evidence from modified unit root tests," *Int. J. of App. Eco*., vol. 11, no. 2, pp. 1-9, 2014 .

[16] J. Chen, G.Q. Zengb, W.Zhoua., W.Dua and K.D. Lua, "Wind speed forecasting using nonlinear-learning ensemble of deep learning time series prediction and extremal optimization," *Energy Conv. & Mgmt.,* vol. 165, pp. 681–695, 2018.

[17] E.B. Iversen, J. M. Morales, J.K. Moller and H. Madsen, "Short-term probabilistic forecasting of wind speed using stochastic differential equations," *Int. J. of Forecasting*, vol. 32, no. 3, pp. 981–990, 2016.

[18] T. Kaur, S. Kumar and R. Segal, “Application of artificial neural network for short term wind speed forecasting,” *2016 Biennial Int. Conf. on Power and Energy Sys.: Towards Sustainable Energy (PESTSE), Bangalore,* pp. 1-5, 2016.
[19] J. P. S. Catalão, H. M. I. Pousinho and V. M. F. Mendes, “Hybrid Wavelet-PSO-ANFIS Approach for Short-Term Wind Power Forecasting in Portugal,” *IEEE Trans. on Sust. Energy*, vol. 2, no. 1, pp. 50-59, 2011.
[20] E. Cadenas, W. Rivera, Campos-Amezcua, R. Cadenas, “Wind speed forecasting using the NARX model, case: La Mata, Oaxaca ,Mexico,” *Neural Comp. & App.* ,vol. 27, no. 8, pp. 2417-2428, 2016.
[21] C. Dritsaki , “The Performance of Hybrid ARIMA-GARCH Modeling and Forecasting Oil Price,” *Int. J. of Energy Eco. and Policy*, vol. 8, 2018.
[22] C. Chen, J. Hu, Q. Meng and Y. Zhang, "Short-time traffic flow prediction with ARIMA-GARCH model," *2011 IEEE Intelligent Vehicles Symposium (IV)*, Baden-Baden, pp. 607-612, 2011.