Global Journal of Computer and Engineering Technology (GJCET)

Vol. 1, No. 1, January, 2021 ISSN: 2767-1933

A COMPARATIVE STUDY OF SHORT-TERM WIND POWER FORECASTING ON WIND DATA

Minaxi, D.C.R.U.S.T, India (minaxichalia@gmail.com)
Rohtash Dhiman, D.C.R.U.S.T, India (rohtash.k@gmail.com)
J.S. Saini, D.C.R.U.S.T, India (jssain@rediffmail.com)

ABSTRACT

The paper presents a novel scheme for short-term wind power forecasting based upon minimization of Mean Square Error (MSE) by employing Autoregressive Integrated Moving Average (ARIMA) and Non-linear Autoregressive with Exogenous (NARX) input. Numerous schemes have been reported in the literature, which states that the minimization of MSE is a major issue. In this paper, the wind power forecasting results using ARIMA and NARX are presented and compared.

Keywords: Power Forecast, ARIMA model, NARX model, and MATLAB Software.

1. INTRODUCTION

Last few decades, generation of wind power is a rapidly developing energy system, generally due to concerns on global warming, financial back-up from government, and patterns development for power electronic devices and their processes. With growing involvement in advanced power system, accurate and reliable prediction method of wind power is necessary for system promoters, to entail wind generation interested in Economic Scheduling (ES), UC (Unit Commitment), and Reserve Assignments problems. The prediction also improves bidding in electricity markets. An accurate device for wind power prediction is applicable to decrease or overcome undesirable outcomes in developed wind power. In pool form electrical energy markets, a short-term wind power prediction is used for wind power manufacturers to share within day-ahead cooperative data and provide balancing in market. For multiple wind energy approaches, it is important to predict the energy information from accessible meteorological information.

2. LITERATURE REVIEW

In literature review different sources related to the topic should be presented. The literature should be related with either domain or method/technique/algorithm used in the correspondence research. The previous researchworks should be described in the form of title, problem statement, objectives like all the cited literature must be written in references.

2.1 Physical Model

Numerical Weather Prediction (NWP) using data like temperature, pressure, surface roughness, etc. This method is used for long term prediction. The NWP system frequently provides wind power forecasts for a grid of adjacent point about the wind generator. According to the type of NWP system, those forecasts are given with a spatial declaration. Numerous researchers have used those models for wind power/speed prediction; on the other hand, collecting information on environmental conditions is one of the main difficulties in the implementation.

2.2 Statistical Models

Statistical methods aim to find the relationship of on-line considered power data. The statistical models are simple in modeling, cheaper toward developing, and good in comparison to other models, for short-term prediction. Numerous researchers have used some of these models for wind power/speed prediction, e.g., Support Vector Regression

[(Fang & Chiang, 2016)]; Spatio-temporal statistic (Xie et al., 2013), Gaussian process (Lee et al., 2013), Particle Swarm Optimization (Quan et al., 2014), Adaptive filtering network (Khalid & Savkin, 2012),etc.

2.3 Hybrid Models

The main aim of these models is to get benefit and obtain a globally optimal forecasting performance. Since by using individual forecasting technique information contained is incomplete, hybrid method be able to maximize offered information, incorporate individual model information and create better use of multiple forecasting methods therefore prediction accuracy improves (Bracale et al., 2013)(Boutsika et al., 2012), many classical models have been proposed in the literature for improving the accuracy and efficiency of wind prediction. But, intelligent models like artificial neural networks (ANNs) (Paixão et al., 2017) can have higher features above conventional methods in forecasting since they perform well even with incomplete data due to their capability of generalization (Huang et al., 2017), (Wu, 2016), (Shi et al., 2014), (Catalão et al., 2011). In this work, an artificial neural network is trained for forecasting wind power by using Levenberg-Marquardt optimization (generally the quickest back-propagation learning algorithm in the MATLAB toolbox). To evaluate prediction accuracy, performance measures, viz. MSE (Mean Square Error) has been used.

3. EXPERIENTIAL WORK

This model has included two techniques ARIMA and NARX:

3.1 ARIMA

This technique was popularized by Martyr Box and Gwilym Jenkins in 1976. It is used to introduce a general category of time series models known as ARMIA models. The block diagram of ARIMA model is revealed in figure 1. This modeling is done with help of Box- Jenkins approach(B-J).

To develop a time series model emerge ARIMA it is necessary to recognize p, d, q parameters (those are the parameters of AR, MA, ARMA models) and to learn the time series. ARIMA method is an iterative, exploratory process intended to best–fit in time series observation it involves the following five steps:

1. Ensuring stationary: To find out the adequate values of d.
2. Identification: To find out the adequate values of p & q via PACF, ACF as well as unit root test.
3. Estimation: To estimate an ARIMA model by use of these values (p, q, and d).
4. Diagnostic checking: Check residual of expected ARMIA model(s) check if value is white noise or not; then choose best model among well-behaved residuals.
5. Forecasting: To develop unit forecasts.

3.2 Training of TIME SERIES models and Predictive Performances

To validate a forecasting model, split data into training and testing sets. For model building firstly choose training data, then choose test data for a trial run and analyze that data. After that check how well a model predicts future unknown data values.

Then fitted model use for forecasting toward assesses predictive ability. But models to facilitate fitted healthy in sample, not guaranteed toward forecast healthy values. For example, good sample be able to obtain from over fitting, but poor predictive performance obtain while under fitting.

When checking prediction performance, in a set not use data twice i.e., data employ to fit model ought to be different than data used to evaluate forecasts. Apply cross-validation in the direction of estimate out-of-sample forecasting capability; overall process can be summarized as follows:

1. Time series split into 2 parts: a) Set of Training data and b) Set of Validation data.
2. Use Fit model for training data.
3. Forecast fitted model over validation period.

Compare forecasting to validation process, observations compare using plots and numerical summaries (like Mean Square Error).

3.3 NARX model:

In real time applications, an essential correlation between modeled time series as well as additional external data is necessary. The model NARX was proposed to predict series y (t), specified p past values of series y and a further external series x (t), which can be any dimension (one and multi-dimensional). Given equation 1 shows NARX model for time series prediction.

$$Y(t)=h(x(t-1),x(t-2),\ldots\ldots x(t-k),y(t-1),y(t-2),\ldots\ldots y(t-p)+e(t)) \qquad (1)$$

NA RX is nonlinear type of model, based on last output and external data it estimates future values of time series. NARX model is shown in figure.

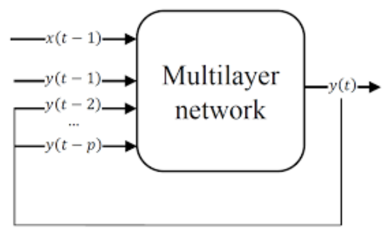

Figure 1: Non-linear Autoregressive with Exogenous (External) inputs (NARX) network.

In this NARX model, there is one input time series by time (t-1), y (t-1), and other with exogenous data at time (t-1), x (t-1). The model is analogous to NAR network, but, difference in inputs.

3.4 Architecture

A multilayer feed-forward propagation network through one layer of z- hidden units is revealed into figure 2. Output unit has W_{ok} and V_{ok} as bias. Output and Hidden units contain a bias. Bias acts similar to weights or connection as of units whose output is 1 all the time. From figure 2 it is understandable that network has an input layer, hidden layer, and output layer also. Input layer is linked with hidden layer and a hidden layer linked with output layer via interconnection weights.

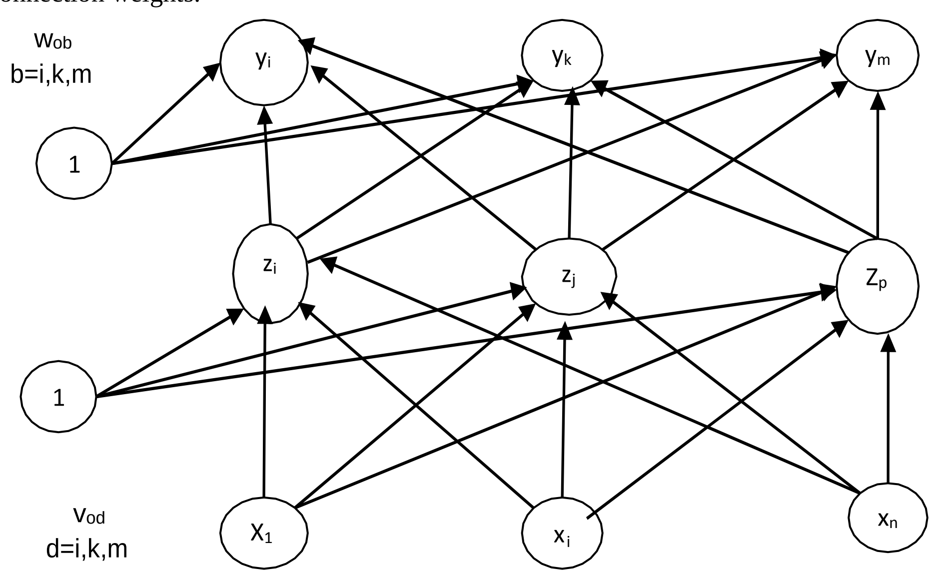

Figure 2: Architecture of Back Propagation network

3.5 Forecast Accuracy

Evaluation of forecasting methods, start values, and smoothing parameters it is essential to determine differences between possible alternatives. On the other hand, selection method is difficult for measurement of forecast errors because different measures have their advocates. To reduce forecasting errors only a measure does not appear achievable. It impossible to reduce them because different types of forecast errors have different dimensions. That's possible by considering only one dimension without sacrificing information(Zhao et al., 2011).

MSE (Mean Square Error)

The MSE defines closeness of a regression line towards set of data. This error is calculated by taking distance from the data points to regression line then squaring them as defined in equation 2; squaring is necessary to remove any negative sign.

General steps of MSE from a set of X and Y values:

i. Find regression line.
ii. Insert X values into a linear regression equation to find new Y values.
iii. Subtract new Y value from original value, to get error value.
iv. Square errors values.
v. Add up errors values.
vi. Find mean values.

$$\text{MSE} = {}^{1}/_{N} \sum_{i=1}^{N} (\text{fi} - \text{yi})\text{ ^2} \quad (2)$$

Where N is the number of data points, fi the value returned by the model and yi the actual value for data points i.

3.6 Objective

i. To forecast the wind power and minimize MSE (Mean Square Error).
ii. The process utilized two steps for the execution
   a) Firstly, train the wind farm power data.
   b) Then; compare the predicted data to the actual wind series data for the final forecasting.

3.7 Wind power data train on ARIMA models

For validation of a forecasting model, split data into training and testing data. For model building need training data and then use test data as an out samples for trial run and also for to analyze how well predicts the future unknown data values.

Step: 1 wind power data series

The wind power data has been collected from the Suzlon Pvt. Ltd wind farm. Thus, the hourly Data series consists of 510 samples.

Step: 2 Control the random numbers generation

In which random numbers are controlled across wind speed, temperature, and direction of the wind.

Step: 3 Create regression model

From the above data, series create the regression model for training and construct the ARIMA model using regression model.

Step: 4 Training the ARIMA model

From the above wind power data series, 400 samples are trained by the ARIMA model.

Step: 5 Compares the predicted signal to the wind series

In this compare the predicted signal to the actual wind series for the final forecasting.

Step: 6 Error Calculations

The measures of error used here are Mean Squared Error (MSE).

3.8 Implementation of NARX for wind power prediction

The implementation of NARX has been done by using the following steps:-
Here, step 1 to step 3 are performed in the same manner as explained in the previous section 3.2, further detail of this scheme from step 4 is as follows:

Step: 4 NARX model trained by the neural network

This is one form of a neural network used for training where one sample of outputs is fed with time stamping. 10 numbers of neurons are taken in a single hidden layer with one output layer. This network is to trainthe neural network for the specific wind datasets.

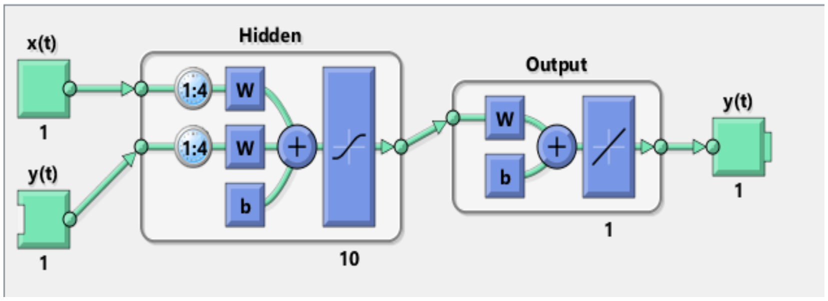

Figure 3: Train the NARX model

This network is used to get the neural network train for the forecasting of wind power. To achieve that, feedback of the outputs is served as the input to predict the value.

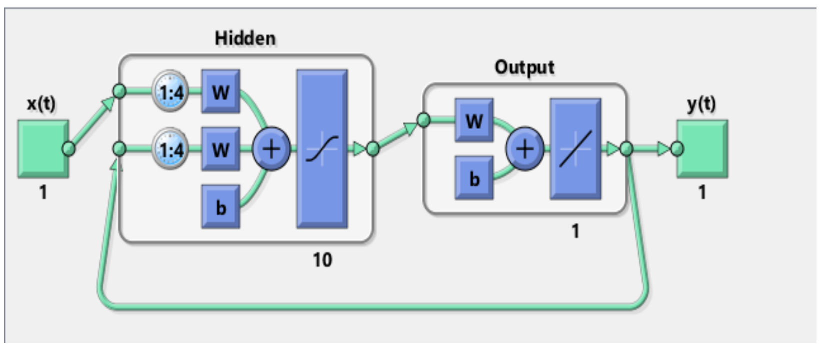

Figure 4: Performance of NARX feedback model

Step: 5 NARX model testing by the neural network
This is the final testing network after the feedback system. The NARX network gives the final wind forecasted values.

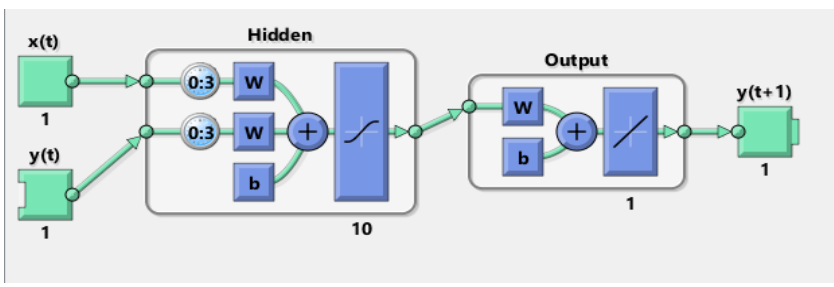

Figure 5: Performance of the testing model

Step: 6 NARX models training validation and testing.
The graphs show the training, validation, and testing of the NARX model.

3.9 Flow chart of NARX neural network

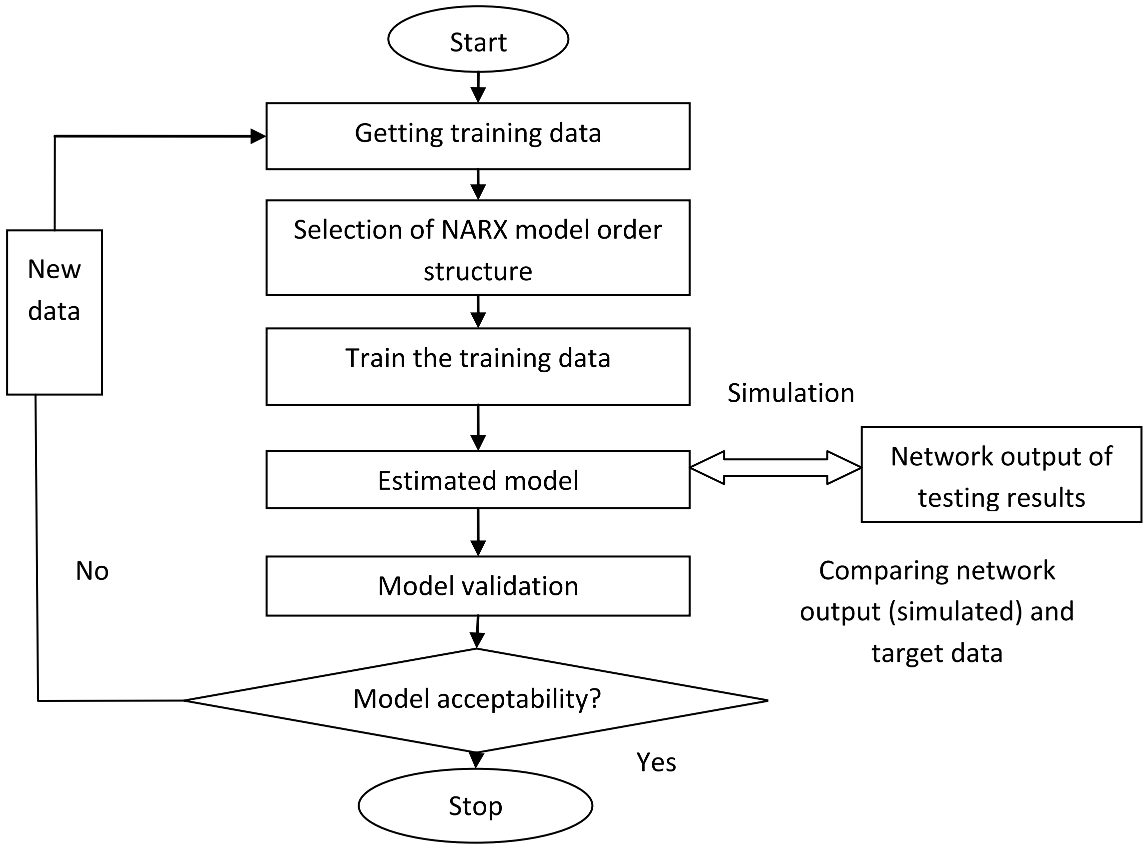

Figure 6: Flow chart of NARX neural network

4. METHODOLOGY

Forecasting of wind power has implemented in this paper using ARIMA and NARX networks. Firstly, the network is trained by using ARIMA. To further improve the parameters, a NARX network is used. Usage of these techniques and corresponding forecasting results has been presented in this paper with details. Using these techniques, all simulations have been performed in MATLAB 2015. This paper presents the results from all such simulations.

mean (yMSE) = 1.7020e+03 i.e. 34.185584

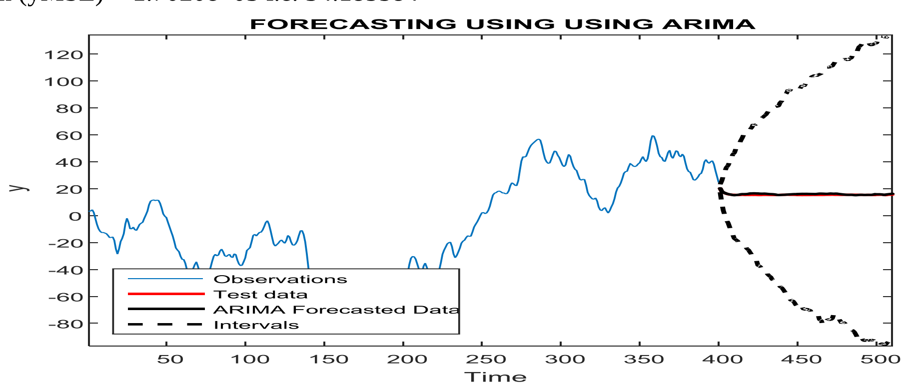

Figure 7: ARIMA model performance

4.2 Forecasting Results from the NARX model

For prediction (1 hour ahead) of these three variables wind speed, wind direction, and temperature of a wind farm, neural networks are used.

Step 1: Samples of wind data trained by neural network

Training data for this network is generated by using the first 400 samples of data. Training of this network is on MATLAB 2015. The network is trained by using LM back propagation learning algorithm. Mean Squared Error is chosen to be the performance criteria. Other graphs are described as follows:

4.2.1 Performance

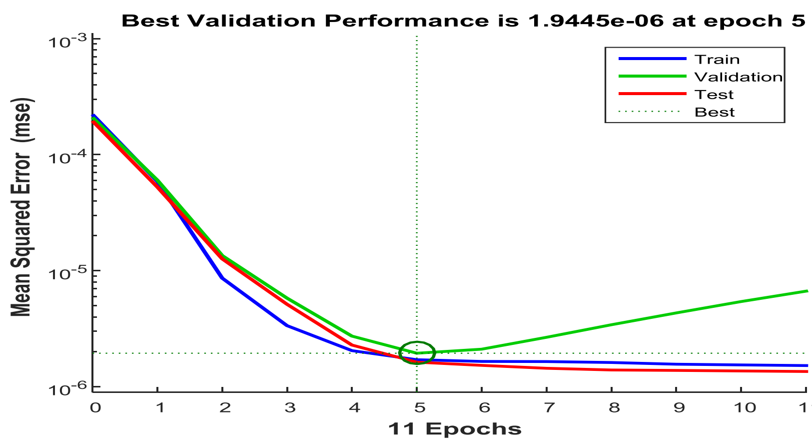

Figure 8: Training performance of neural network

Figure 8 shows the performance of the network i.e. train, test, validation, and the best value of the performance. The best validation performance is 1.9445e-06 i.e. 0.00481993 at epoch 5.

4.3.2 Training state

Figure 9 shows gradient coefficient variation in several epochs. Final value of gradient coefficient obtain at epoch 11 is 0.00276604 i.e. approximate to 1. Training and testing will better for minimum value of gradient coefficient of network. Validation value at epoch11 is 6.

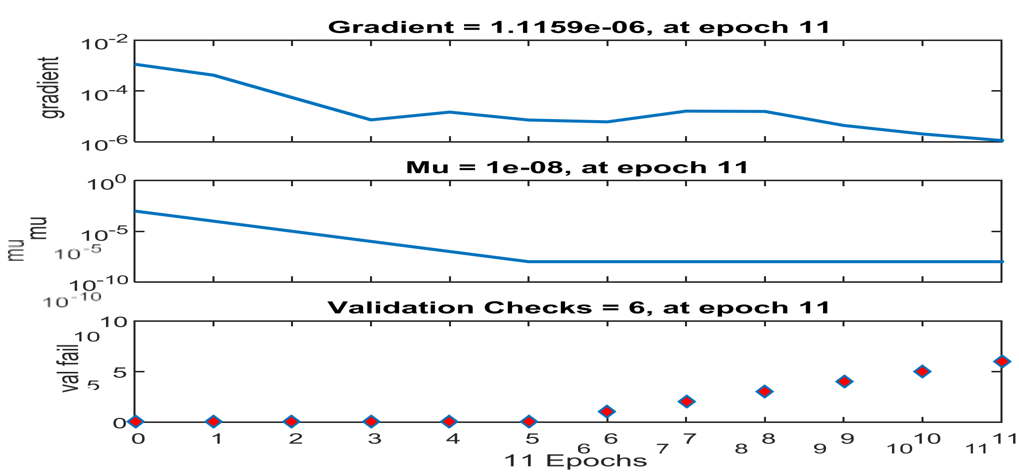

Figure 9: Neural network training states

4.2.3 Error Histogram

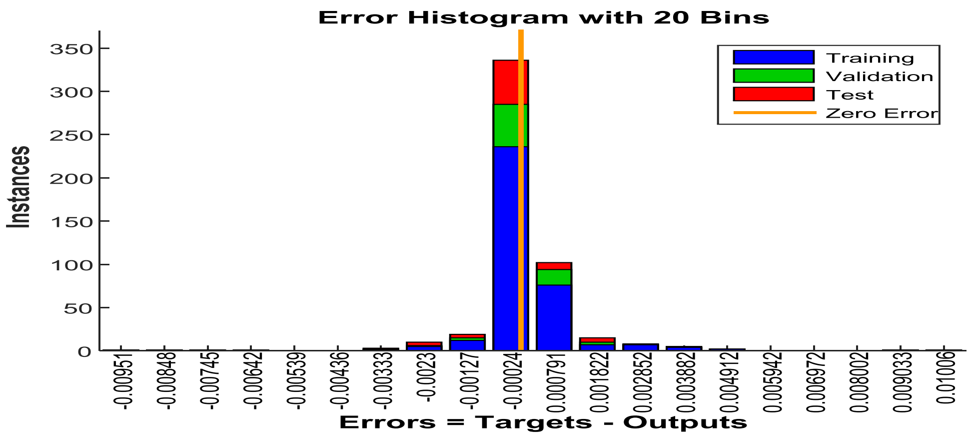

Figure 10: Neural network training error histogram

This means that the data fitting has been quite precise. Figure 10 also shows the error histogram of the trained NN for the training, validation and testing steps.

4.2.4 Regression

Regression is a statistical measure that used to determines the strength of the relationship between one dependent variable and a series of other changing variables. The figure 11 shows the output graph of training, validation, and testing and also shows the overall output values graphs.

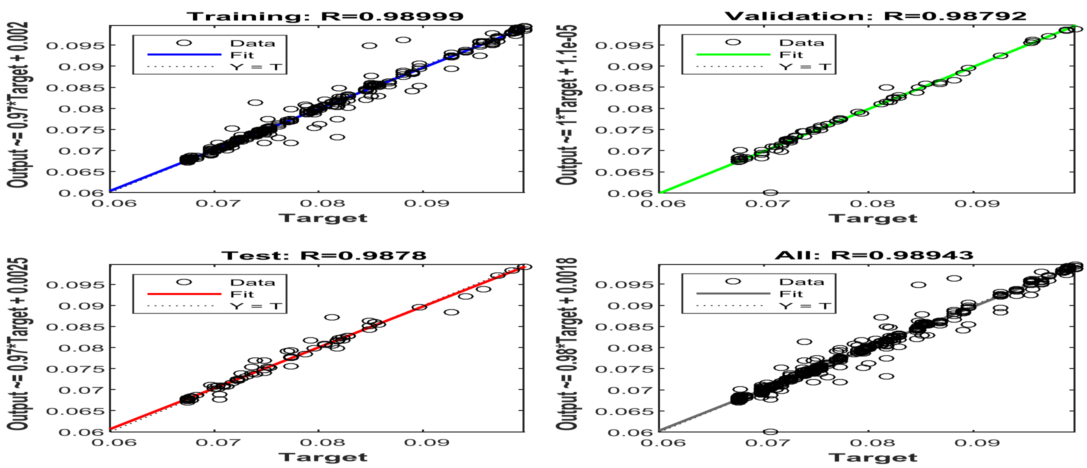

Figure 11: Neural network training regression

4.2.5 Time series response

Figure 12 shows the time series responses of targets and outputs values. The error graph gives the difference between the outputs and the values of targets.

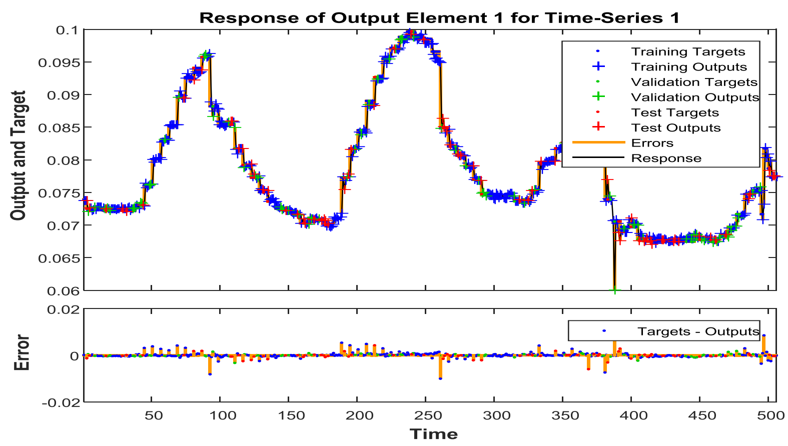

Figure 12: Neural network training Time-series response

4.2.6 Error auto correlation

Autocorrelation is also known as series correlation. The error follows a pattern that showing something is wrong with the regression model. Autocorrelation is present if assumption is violated and error value is also correlated. Autocorrelation is a general problem in time series regression. Figure 13 shows a graph of Autocorrelation.

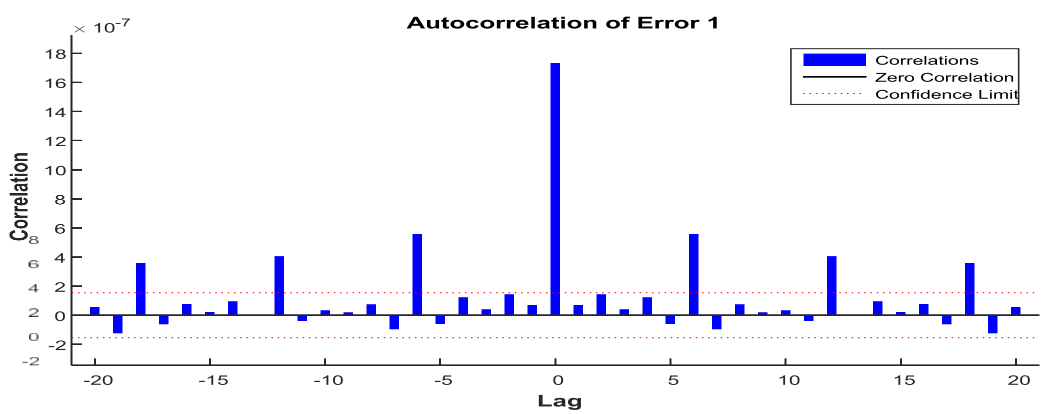

Figure 13: Neural network training error autocorrelation

4.2.7 Input-Error-cross-correlation

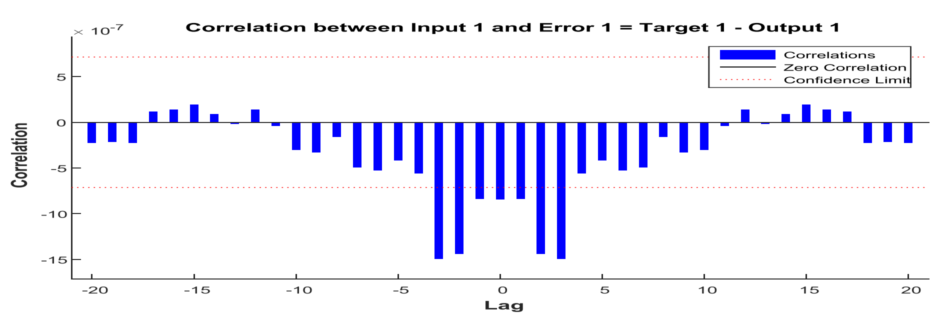

Figure 14: Neural network training input-error-cross-correlation

Figure 14 shows the NN training of input-error-cross-correlation in this correlation, an input time series and error time series are taken and plot the cross-correlation. The correlation plots between the inputs and error across varying lags for finding the input element which is being correlated and plotted. Which error element is being correlated and plotted 1 by default.

4.3 Final comparison graph and Table

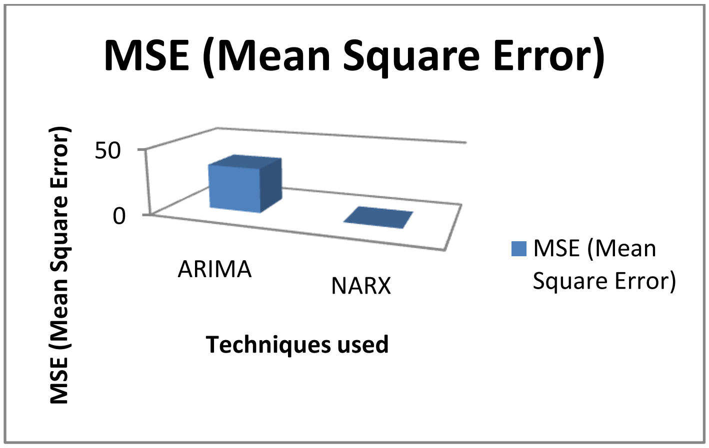

Figure 15: Comparison graph of ARIMA and NARX techniques

The graph and compression table shows that the result of forecasting data is better in case of NARX as compared to the ARIMA model. The MSE (Mean Square Error) is less in NARX i.e 0.0042959254 or case of ARIMA; MSE error is 1.7020e+03, i.e 34.18558, so accuracy is more in case of NARX instead of ARIMA model.

Table 1: Comparison Table of ARIMA and NARX technique

Techniques used	MSE (Mean Square Error)
ARIMA	34.18558
NARX	0.0042959254

Performance=1.7331e-06i.e
0.0042959254
Closed loop NARX performance = 0.000057
Prediction of a one day ahead NARX performance = 0.000002
Next predicted value = 0.018477

5. CONCLUSION

In this paper, two intelligent models have been used for the prediction of nonlinear wind power series. The performance of the models is tested for the prediction of wind power. The real data of wind is collected from Suzlon Pvt. Ltd. wind farm. The time is taken in an hour format prediction. As shown in Table 4.9 improvement of prediction performance by using the NARX model technique. These input variables are wind speed, wind direction, and temperature and output is power. To improve the prediction accuracy the MSE is reduced. ARIMA technique gives the MSE error value 34.18558 and NARX gives the MSE error value 0.0042959254 that shows the NARX model gives a better result as compare to ARIMA network.

The following studies based on the presented research work can be extended for consideration in the future. The prediction of wind power can be done with the help of other intelligent techniques like filtering, ANFIS, different optimization techniques, etc.

REFERENCES

Boutsika, T., Santoso, S., & Member, S. (2012). *Quantifying Short-Term Wind Power Variability Using the Conditional Range Metric*. *3*(3), 369–378.

Bracale, A., Caramia, P., Carpinelli, G., Rita, A., Fazio, D., & Varilone, P. (2013). *A Bayesian-Based Approach for a Short-Term Steady-State Forecast of a Smart Grid*. *4*(4), 1760–1771.

Catalão, J. P. S., Pousinho, H. M. I., Member, A., & Mendes, V. M. F. (2011). *Hybrid Wavelet-PSO-ANFIS Approach for Short-Term Wind Power Forecasting in Portugal*.

Fang, S., & Chiang, H. D. (2016). Improving supervised wind power forecasting models using extended numerical weather variables and unlabelled data. *IET Renewable Power Generation*, *10*(10), 1616–1624. https://doi.org/10.1049/iet-rpg.2016.0339

Huang, C.-M., Huang, Y.-C., Huang, K.-Y., Chen, S.-J., & Yang, S.-P. (2017). Deterministic and probabilistic wind power forecasting using a hybrid method. *2017 IEEE International Conference on Industrial Technology (ICIT)*, 400–405. https://doi.org/10.1109/ICIT.2017.7913264

Khalid, M., & Savkin, A. V. (2012). *A Method for Short-Term Wind Power Prediction With Multiple Observation Points*. *27*(2), 579–586.

Lee, D., Member, S., & Baldick, R. (2013). *Short-Term Wind Power Ensemble Prediction Based on Gaussian Processes and Neural Networks*. 1–10.

Paixão, J. L., Rigodanzo, J., Sausen, J. P., Hammarstron, J. R., Abaide, A. R., Canha, L. N., & Santos, M. M. (2017). Wind generation forecasting of short and very short duration using Neuro-Fuzzy Networks: A case study. *Proceedings - 2017 International Conference on Modern Power Systems, MPS 2017*, *Mps*. https://doi.org/10.1109/MPS.2017.7974446

Quan, H., Srinivasan, D., & Khosravi, A. (2014). Short-term load and wind power forecasting using neural network-based prediction intervals. *IEEE Transactions on Neural Networks and Learning Systems*, *25*(2), 303–315. https://doi.org/10.1109/TNNLS.2013.2276053

Shi, J., Ding, Z., Member, S., Lee, W., Yang, Y., Liu, Y., & Analysis, A. G. R. (2014). *Hybrid Forecasting Model for Very-Short Term Wind Power Forecasting Based on Grey Relational Analysis and Wind Speed Distribution Features*. *5*(1), 521–526.

Wu, Y. (2016). *Stratification-Based Wind Power Forecasting in a High Penetration Wind Power System using a Hybrid Model Po-En Su Jing-Shan Hong*. *9994*(c). https://doi.org/10.1109/TIA.2016.2524439

Xie, L., Gu, Y., Member, S., Zhu, X., & Genton, M. G. (2013). *Short-Term Spatio-Temporal Wind Power Forecast in Robust Look-ahead Power System Dispatch*. 1–10.