Available online at www.ijpe-online.com

International Journal of Performability Engineering

vol. 15, no. 7, July 2019, pp. 1772-1782
DOI: 10.23940/ijpe.19.07.p3.17721782

Short-Term Wind Power Forecasting using Wavelet-based Hybrid Recurrent Dynamic Neural Networks

Pavan Kumar Singh*, Nitin Singh, and Richa Negi

Motilal Nehru National Institute of Technology Allahabad, Prayagraj, 211004, India

Abstract

In the recent past, the integration of wind energy generation into smart grids has gained lot of momentum because of its availability. The major hurdle in the integration of wind power in smart electric grids, at present time is the irregularity and unpredictability of wind power. Therefore, in order to deal with these challenges, the superior forecasting tool plays an important role in the planning and execution of the wind energy integration. In the expanding power system, because of increasing wind power penetration, a precise wind power forecasting technique is greatly needed to help system operators and consider wind power production in economic scheduling, unit commitment, and allocation trouble reservation. In this paper, two hybrid recurrent dynamic neural networks have employed hybridizing wavelet transform (WT) for short-term prediction of wind power. The proposed approach consists of wavelet decomposition of wind power and wind speed time series, and NAR and NARX recurrent dynamic neural networks are employed to regress upon each decomposed sub-series. Thereafter, the individual outputs of sub-series are aggregated to achieve final prediction of wind power, with up to 24 hours of forecast horizon. The performance of the proposed method is obtained in terms of MAE, MSE, and MAPE values and compared to the results of the persistence method. The forecast results reveal that WT-NARX model is better in terms of the selected performance criteria as compared to the WT-NAR and persistence models respectively.

Keywords: wavelet transform; Smart Grids; Recurrent Neural Network; Dynamic Neural Network; Nonlinear Autoregressive Network (NAR); Nonlinear Autoregressive Network with exogenous input network (NARX).

(Submitted on June 24, 2019; Revised on June 30, 2019; Accepted on July 26, 2019)

1. Introduction

Wind power production was the most rapidly increasing system of energy conversion during the last three decades. This is essentially because of the development related to economic encouragement from governments, evolution in design, as well as the manufacturing of power electronic and global warming. Because of its greater correlation with stochastic and nonstationary behavior of wind speed, in the current power system, the irregular behavior of wind power creates the maximum difficulty for wind power integration. Wind power integration brings about many problems for system operators regarding power system operation, such as preserving power balance, power quality, system frequency, and voltage support; planning and economy, such as wind power's insecurity to unit commitment, economic load scheduling, and spinning reserve computations; and more. A perfect tool is needed for wind power prediction to diminish the unwanted effects in the increasing scenario of wind power penetration. For wind power producers to take part in day-ahead adjustment and balancing markets, in the pool-based electricity markets, a short-term wind power prediction tool is required [1]. In modern times, the value of renewable energy sources has grown. In reality, they are sufficient, free of pollution, and liberally available in our environment, while traditional energy sources are gradually coming to an end. Accurate wind speed prediction can play a very important role in our current and upcoming wind power market. For appropriate grid operations with the wind power integration into the power system, precise wind power prediction has been gaining great significance [2].

Short-term forecasting of wind power is an extremely essential area of research in the energy sector because system operators must manage a significant quantity of variable power due to its growing installed capacity. The short-term forecasting in time scales is in the range of a few days, and from minutes to hours for the forecast time-steps. In literature, a

* Corresponding author.
E-mail address: impksingh@gmail.com

number of techniques have been reported to predict wind power, specifically physical methods in addition to statistical methods. It is beneficial to use physical methods for long-term forecasting, whereas statistical methods perform better in short-term forecasting. The traditional statistical models are similar to direct stochastic time-series models incorporating auto regressive (AR) and auto regressive integrated moving average (ARIMA) models. In spite of being the simplest, persistence time-series models can outperform numerous others in very short-term forecasting. They have been utilized extensively and practically regardless of their unstable forecasting effectiveness. It has been confirmed that the persistence method is a valuable first approximation in short-term wind power forecasting. Therefore, persistence in addition to ARIMA approaches offer a vital benchmark against which other optional techniques can be evaluated. Some novel methods, such as data mining, evolutionary algorithms, fuzzy logic (FL), artificial neural networks (ANNs), and hybrid methods, have been attracting the interest of researchers in recent years. Comparing all the methods accurately is somewhat complicated because they depend on various circumstances and the collection of data is a difficult task. Nevertheless, in short-term forecasting, artificial-based models have been reported to be better than other methods. If there is sufficient training data, a sufficient choice of input-output examples, a proper number of hidden neurons, and the existence of adequate computational resources, the simple, powerful, and flexible prediction tools are provided by neural networks (NNs) [3].

Large-scale wind power integration accompanies large influences on power systems in addition to huge challenges to grid dispatching operation because of its high sensitivity and low predictability. Enhancing the accuracy of wind power forecasts is of great importance to ensure stable power system operation. Among the various factors affecting the accuracy of wind power prediction, the accurate forecasting of wind speed is a key one. Wind speed prediction has crucial effects on wind power forecasting. While forecasting the wind power of a wind farm, a general and efficient method is to first forecast the wind speed of the wind farm and then utilize the wind power curve to determine the value of wind power prediction. The methods for forecasting wind speed can be classified into two kinds. i.e., physical and statistical methods. The first type necessitates establishing physical models of wind farms accompanied by the nearby surroundings in addition to the roughness, terrain, and other related information of the wind farms. All of these factors are taken into account while performing forecasts with physical equations. In statistical methods, generally historical (measured) data of wind farms is used to perform statistical analyses and achieve wind speed forecasting. Inbuilt laws identify mapping relationships of system inputs and wind speed. In short-term and very short-term forecasting, the statistical models have added appliances because of their automatic matching to the location as well as surroundings of the wind farms. Techniques such as time series, Kalman filtering, SVM, WT, and ANNs have been largely applied in research on statistical methods. For wind speed prediction on the basis of statistical models, many researchers generally process data initially to enhance the accuracy of the forecast efficiently and to assist in finding the laws of wind speed. To build powerful forecasting tools, artificial neural network (ANN) and multiresolution analysis (MRA) by wavelet transform (WT) techniques have been applied to predict the wind speed as well as wind power production [4-5].

At every instant, the total amount of electricity provided in a power system must be equal to the changing load from consumers of electricity. This can be accomplished in an economical way by scheduling the power plants well in advance in accordance with increasing marginal costs of operation. The inconsistent wind power pattern of generation alters the scheduling of the other generating plants as well as the utilization of transmission capability among different regions. In case the wind variations are not being predicted or are incorrectly predicted, it is compulsory to utilize additional reserves. These are usually costly backup generators, operating in the conditions of low efficiency and used to rapidly balance between generation and load, hence ensuring the system reliability. At present, it is obvious that an added growth of this kind of renewable energy is strongly dependent on the capability for accurate wind speed and wind power forecasting [6-8]. Generally, the input of wind power is greatly reliant on various meteorological parameters. In this work, hybrid models employing wavelet decomposition and recurrent dynamic neural networks (NAR and NARX) have been developed [9].

The rest of the sections in this paper have been arranged as follows. In Section 2, the wavelet decomposition process and two widely-used recurrent dynamic neural networks, i.e., NAR and NARX, are described. Section 3 gives details of the proposed hybrid WT-NAR and WT-NARX methodology. Section 4 discusses the simulation results of the proposed hybrid models, and the conclusion is given in Section 5.

2. Forecasting Techniques

2.1. Wavelet Decomposition

Pre-processing the input data of wind speed and wind power is necessary because the collection of observations from such locations are extremely uncertain and nonstationary. The wavelet transform (WT), an efficient signal processing tool used for time frequency analysis of signals, is capable of handling time-varying random signals. WT permits the wind time series to decompose into a set of consecutive data patterns (sub-series) that possess better statistical properties than the raw input data. The chief concern of the WT is to gather the significant information with removal of noise and irregularities from the original signal. The Daubechies wavelet has been found to perform smoothening of the signal appropriately. With the application of wavelet transform, the decomposition of an original signal into a number of wavelet functions at various time and frequency levels can be obtained with the help of expansion and translation scaling of a mother wavelet function. In comparison to Fourier transform, the wavelet transform possesses a merit of time and frequency resolution with automatic adjustment. WT offers superior frequency resolution in addition to improved time resolution for low-frequency and high frequency components of a signal, respectively. By using wavelet transform efficiently, the extremely nonlinear wind speed and wind power signals can be analyzed effectively, due to its adaptive adjustment capability of time and frequency resolution.

The continuous wavelet transform (CWT) and discrete wavelet transform (DWT) are two kinds of wavelet transform. Regarding a mother wavelet function $w(t)$, a continuous wavelet transform $f(t)$ can be given by Equation (1):

$$CWT_f\ (a,b) = \left(f(t), \psi_{a,b}(t)\right) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{|a|}}\psi^*\left(\frac{t-b}{a}\right)dt \tag{1}$$

Where *, a, and b denote the complex conjugate, scale coefficient, and translation coefficient, respectively.

A digital equivalent of continuous wavelet transform is discrete wavelet transform, which provides sufficient information and highly decreased computational effort. Equation (2) gives the transformation used to achieve the DWT:

$$\begin{cases} a = 2^j \\ b = k2^j \end{cases} \tag{2}$$

Where j and k have only integer values and denote the scale and translation coefficients, respectively.

The Mallat algorithm is a fast DWT that builds up the foundation of multi-resolution analysis (MRA) theory. The algorithm decomposes a signal into various levels of resolution during the decomposition and reconstruction processes. The algorithm is utilized in multi-resolution computation to make simpler calculation processes that employ filters and signifies an effective way for implementation of the DWT. In the Mallat algorithm, two signals, A1 (approximation) and D1 (details), are achieved by passing the actual wind speed and power signals through high and low pass filters, respectively. The coefficients for the A1 (low-frequency) and D1 (high-frequency) components are obtained respectively by using low pass and high pass filters. Coefficients A1 and D1 describe the approximate and detail information of the signal S, respectively. After the decomposition process, the coefficients have half the length compared to the those of the actual signal, making it necessary to recover the coefficients length during the reconstruction process. Hence, by passing the coefficients through reconstruction filters, a low frequency series A1 and a high-frequency series D1 that have same length as those of the actual signal S are obtained.

In case the level of the decomposition of the original signal S is greater than one, the process of signal decomposition must carry on further after the two coefficients A1 and D1. Only the approximate coefficient requires further decomposition in the succeeding levels. Therefore, the approximate coefficient A1 is decomposed into a further detail coefficient D2 and approximate coefficient A2. In continuance of this process, A2 is again decomposed into A3 and D3 coefficients, and so on. The original signal S is decomposed J times if the decomposition level is J, obtaining D1, D2, ⋯, DJ and AJ as high-frequency coefficients and the low-frequency coefficient, respectively. The relation of decomposed sub-series to the original signal S can be represented as given in Equation (3):

$$S = D1 + D2 + \cdots + DJ + AJ \tag{3}$$

It can be understood that the signal decomposition process using wavelet is the division of the actual signal into groups of approximations step by step. In Matlab, "wavedec" and "wrcoef" functions are applied to DWT decomposition by selecting a mother wavelet. The process of signal decomposition of up to three levels of resolution is shown in Figure 1 [10-11].

2.2. Dynamic Neural Network (*DNN*)

The common forecasting methods of time series are compatible for linear data, whereas the nonlinear methods are required for forecasting non-linear data like wind power and wind speed. One such nonlinear technique uses neural networks, where input data is not required to be linearly varying. The time series data is given as input to train neurons of the neural network. After proper training, the neural network becomes ready for prediction of future values of the time series. In this work, two of the important neural networks, non-linear auto regressive (NAR) and non-linear auto regressive with exogenous input (NARX), are employed as univariate and multivariate forecast models, respectively. Only one variable is involved in univariate modeling, utilizing the past values of wind power to forecast the future values. Two or more input variables are given to the neural network in multivariate modeling as well as prediction. The topologies of NAR and NARX are shown in Figures 2 and 3, respectively. The Levenberg-Marquardt (LM) algorithm (trainlm) is utilized to train both models. The numbers of input lags are selected using ACF and PACF [12]. The cross-correlation functions (CCFs) between the residuals of output and each input time series are specified by a number of lags at which maximum communication occurs. After selecting the maximum lag k_{max}, the values at lags 1 to k_{max} are utilized as model inputs. The need of forecasting steps (periods) determines the number and outputs lags [13]. In this paper, two recurrent dynamic neural network architectures are selected to be utilized as prediction models for wind power. The NARX model enhances the performance of time series forecast due to external (exogenous) information fed as input. Both dynamic neural network models have been developed using the neural network toolbox in MATLAB [14].

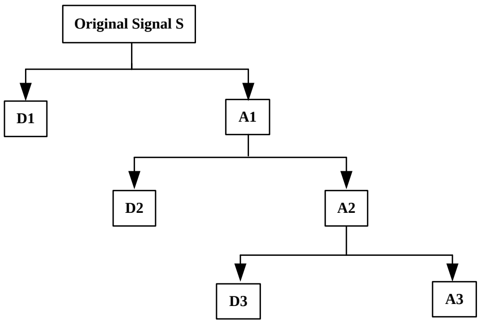


Figure 1. Multilevel (three-level) decomposition process

2.2.1. NAR Neural Network

The NAR, a recurrent type neural network, is suitable for assessing the future values of the input variable using delays of a univariate time series. The future value of forecasting is enabled by a NAR network based on past background of a time series. For a new prediction, the predicted value serves as an input in a re-feeding mechanism. The actual target values are used as feedback to build and train the network in open loop mode, thus ensuring superior accuracy in training. The open loop architecture, similar to a three-layer feedforward structure, is used to train the NAR network. This network depicts a discrete, nonlinear, and autoregressive model when applied to time series forecasting. The network is changed to a closed loop after training, and the predicted values are used as new feedback inputs to the network. This architecture is an arrangement of multilayer perceptron and nonlinear filtering. The prediction of the NAR network is termed as a function of preceding values of observation, as given in Equation (4):

$$y(t) = f(y(t\text{-}1), y(t\text{-}2), \cdots, y(t\text{-}p)) \tag{4}$$

Where $y(t)$ and p are the value of the series at time t and the number of feedback delays, respectively. Equation (4) illustrates how the value $y(t)$ of a data series y at time t is predicted by a NAR network utilizing p previous values of the time series. The function $f(.)$ is not identified earlier and approximated by neural network training with the optimization of weights and bias of neurons in a neural network.

The p features of time series $y(t\text{-}1)$, $y(t\text{-}2)$, $\cdots$, $y(t\text{-}p)$, are known as feedback delays and shown in Figure 2. The functions of the hidden layer and number of neurons are extremely significant from several booming appliances of neural networks. The feature detection and capture of the existing pattern in the data and the complex input-output nonlinear mapping of variables are performed by the neurons in hidden layer. Theoretically, it is shown that one hidden layer is adequate for neural networks to approximate any nonlinear function with any desired accuracy. The optimizations for the entirely flexible numbers of hidden layers and neurons per layer are achieved using the trial-and-error method. However, it is vital to know that excess neurons make the system more complicated, whereas deficient neurons reduce the generalizing capability and computational power of the neural networks. The Levenberg-Marquardt backpropagation (LMBP) is the most general learning rule applied to the NAR networks because it is usually the fastest backpropagation-type algorithm. Levenberg Marquardt Back-Propagation (LMBP) algorithms do not require computing the Hessian matrix for approximation of the second-order derivative, which increases the training speed [15-16].

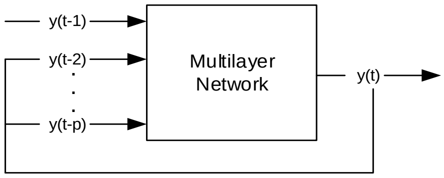


Figure 2. Topology of a NAR network

2.2.2. NARX Neural Network

There is a significant correlation that exists between the modeled time series and added exogenous data in the numerous real applications. It has been shown that the high dependency of the wind power plant's production on the weather environment is a common characteristic of renewable energy. Therefore, to provide an accurate forecast, the knowledge or data integration concerning the weather could be beneficial to the time series modeling process, in comparison to a single approach having one single value related to the wind power. The NARX architecture can be used if the model has more associated variables. With the given p past values of the time series $y(t)$ in addition to an added exogenous time series $x(t)$, which may be single or multidimensional, NARX models predict the wind power series $y(t)$. For the time series prediction, the behavior of the NARX neural network model can be given as Equation (5):

$$y(t) = h(x(t\text{-}1), x(t\text{-}2), \cdots, x(t\text{-}k), y(t\text{-}1), y(t\text{-}2), \cdots, y(\text{t-p})) \quad (5)$$

The assessment of the future values of a time series by the NARX model is dependent on its previous outputs and the exogenous input. The NARX model utilizes past values of the wind power time series $y(t)$ as one input and the wind speed data at $x(t)$ as other exogenous inputs in this study, and it gives the output $y(t)$ one step ahead concerning the value of wind power. The NARX architecture, as shown in Figure 3, is similar to the NAR network in that they both have dissimilarity in their inputs. Hence, external data is taken into account by output $y(t)$, as shown in Equation (5). For training of the NARX model, LMBP is used as the learning rule.

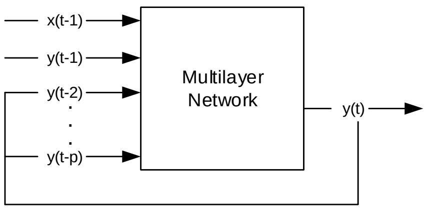


Figure 3. Topology of a NARX network

In neural networks, the selection of input variables is a necessary element because the chosen features highly influence the prediction accuracy. Traditionally, the number of the input lags of the ANN is chosen on the basis of the trial and error method. This means that several neural networks should be created and checked with various numbers of input lags, and thereafter, the networks having superior performance would be chosen as the forecasting model. This time-consuming trial and error method can be avoided by applying two statistical measures, namely ACF and PACF for selecting input lags of variables. These statistical tools are utilized for analyzing time series. The analysis of ACF and PACF is performed for sub-series of signals after decomposition. It has been examined from ACF and PACF plots that the number of lags significantly correlates to the future values of wind power and wind speed, and they can be chosen as the inputs of the neural networks. The sample autocorrelation function (ACF) of a series indicates the correlations of the series with its lagged values. The optimal lag for the NARX is chosen from subsets of positive peaks higher than the confidence limit lines. For large datasets, a complex error minimization algorithm needs to be formed in order to determine optimal lag, as there may be numerous options for better efficiency [17-19].

3. Proposed Hybrid Model

In this study, wind power (P) time series data has been decomposed as the output (target) for the NAR, wind speed (WS) time series (as exogenous input), and NARX forecast models. Both the signals are decomposed using wavelet transform, and each sub-series of wind power and wind speed, obtained after wavelet decomposition, are respectively used as the target and exogenous variable for the NAR and NARX models.

The network consists of inputs corresponding to the wind power time series as well as the future output values of the time series prediction. After data pre-processing using the mapminmax function and analysis stage, the number of delays is determined experimentally with the trial and error method. The steps to model wind power and wind speed time series, employing the proposed hybrid methodology with the NAR and NARX neural networks, are given below:

(1) The original wind power and wind speed time series are decomposed into a set of more stationary sub-series, using db4 mother wavelet with two levels of decomposition.

(2) The wind power and wind speed time series data are normalized using mapminmax before training of the neural networks.

(3) For each sub-series after decomposition, NAR and NARX networks are created with appropriate training. The number of input lags (input as well as feedback delays) of the models are chosen by means of the ACF and PACF time series plots.

(4) The built NAR and NARX networks are employed to perform multi-steps ahead predictions for the related decomposed sub-series.

(5) After obtaining multi-steps ahead predictions for each decomposed sub-series with NAR and NARX models, they are aggregated (reconstructed) to achieve the final predictions of the wind power time series (target).

The steps of the proposed WT-NAR and WT-NARX modeling and forecasting process are given in Figure 4. The NAR and NARX neural networks employ "tanh" and "linear" functions for the hidden and output layers as activation functions, respectively. In the training process, weights and biases are computed to optimize the task performance. The LMBP algorithm is used to train both univariate and multivariate models. The trial and error method, with the minimum value of the mean square error, has been applied for the selection of optimal values of input, feedback delays, and the appropriate number of hidden layers. An epoch is a learning step during the training process, and in each epoch, all the training examples are passed through the learning algorithm simultaneously before the modification of weights in the training process. The weight updating is performed so as to reduce the mean square error after each epoch. To attain multi-steps ahead prediction, the network is changed to a closed loop configuration. In the training process, example sets of inputs and corresponding outputs of historical data are given to the network. The success of training largely depends on the choice of a sufficient number of inputs presented to the neural network. A neural network maps the input and output relationship in the learning process by adjusting the weights and biases to minimize deviation between the produced output and desired output at each iteration. The iterations are repeated until the results converge to the global minima..

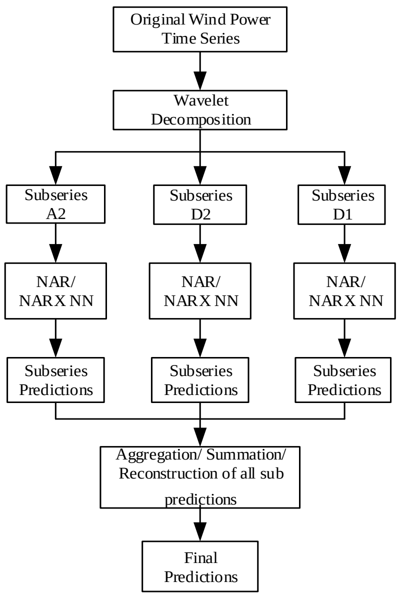


Figure 4. Steps for the proposed WT-NAR and WT-NARX modeling and forecasting process

For the estimation of performance of the models, the mean absolute error (MAE), mean square error (MSE), and mean absolute percentage error (MAPE) are utilized as three error evaluation criteria in this paper. The formulas for calculating the three different measures of performance are respectively given in Equations (6)-(8).

$$MAE = \frac{1}{N}\sum_{i=1}^{n}\left|P_i^{true} - P_i^{forecast}\right| \tag{6}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{n}\left(P_i^{true} - P_i^{forecast}\right)^2 \tag{7}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}\frac{\left(\left|P_i^{true} - P_i^{forecast}\right|\right)}{P_i^{true}} \times 100\ \% \tag{8}$$

4. Simulation Results and Discussion

In this paper, the wind power and wind speed observation data with five-minute resolution has been taken for the year 2012 from a wind farm located in Aagar, South Dakota (Longitude: -100.067^0, Latitude: 44.92891^0) in the USA. Averaging every twelve data points, the averaged hourly data has been obtained from January 1 to April 30, 2012. The wind power time series is shown in Figure 5. The daubechies "db4" has been used as the mother wavelet, and the decomposition level is taken as 2. The A2 approximation sub-series of each time series explains the original time series in a better way and results in the smallest amount of error. Because of elimination of high frequency outliers at this level, a smoother and simpler signal is achieved for forecasting. All the approximations and detail sub-series have been trained with NAR and NARX neural network models. The decomposed wind power time series with "db4" mother wavelet and level 2 is given in Figure 6.

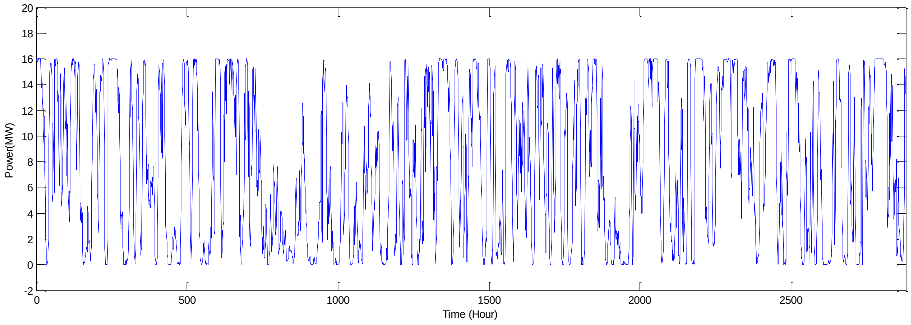


Figure 5. Original wind power time series

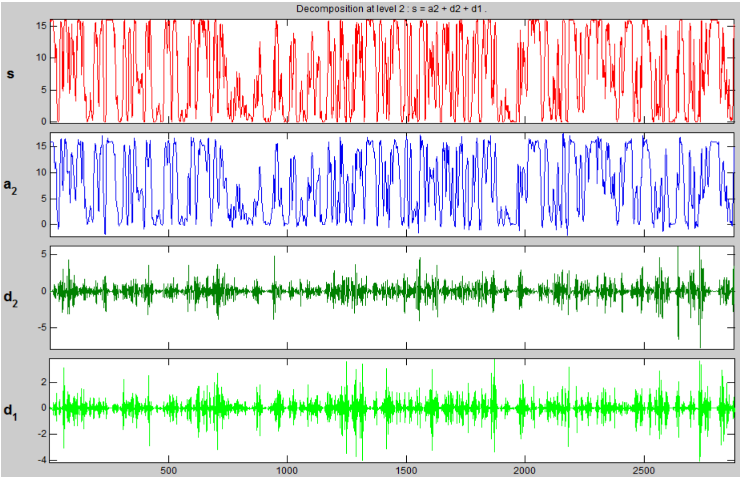
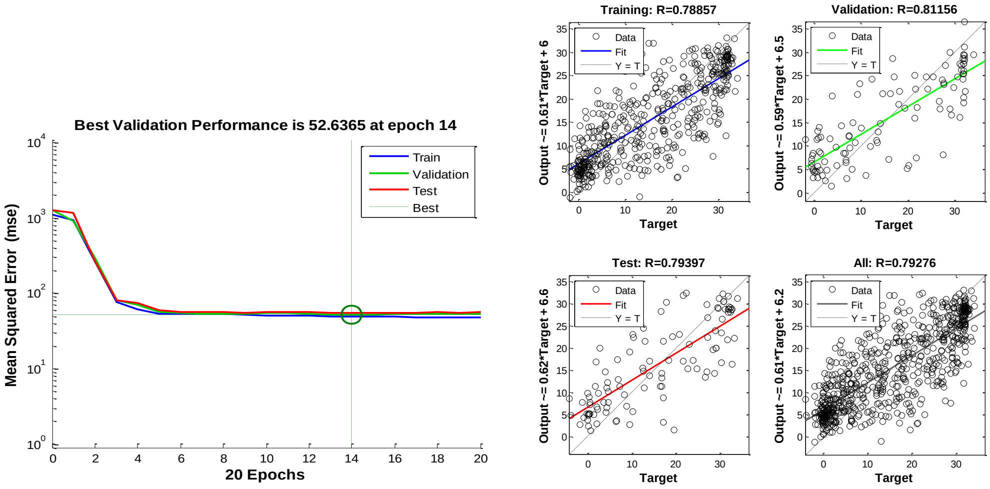


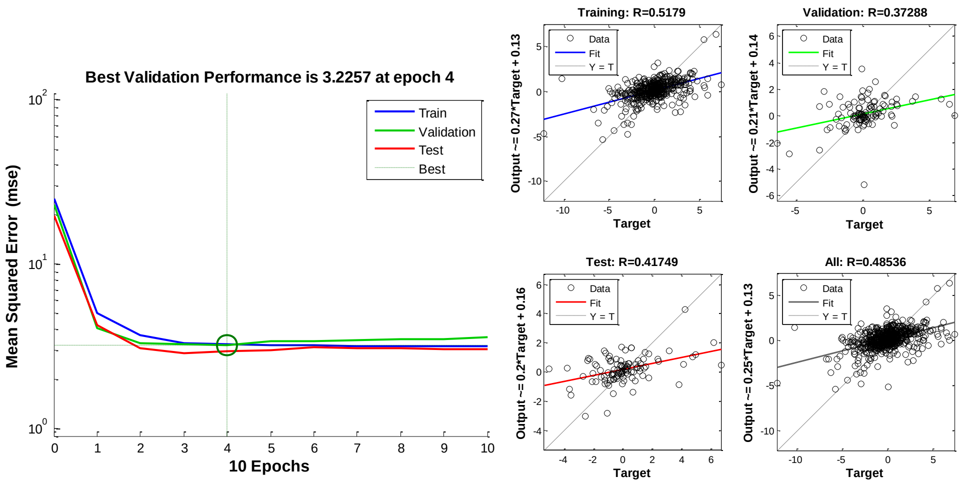
Figure 6. Decomposed wind power time series with "db4" mother wavelet and level 2

Each of the NAR and NARX models decomposed the sub-series of wind power, and the wind speed time series utilised 70%, 15%, and 15% data for training, validation, and testing, respectively. In each NAR network modeling, the number of feedback delays and hidden layer neurons are respectively selected as 4 and 12, 3 and 10, and 2 and 6, using the trial and error method. For each model, ten simulation runs are carried out with a feedback delay variation of 1:6 and hidden layer neurons of 5:20. The Training performances and regressions plots of three NAR models for each sub-series are shown in Figures 7-9, respectively.



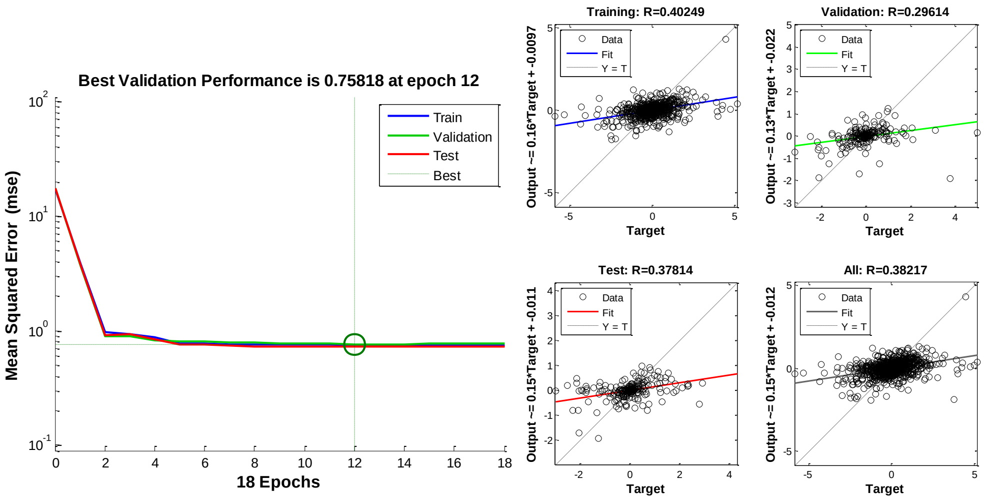
(a) (b)

Figure 7. (a) Training performance; (b) Regression of NAR1 model for sub-series cA2_P



(a) (b)

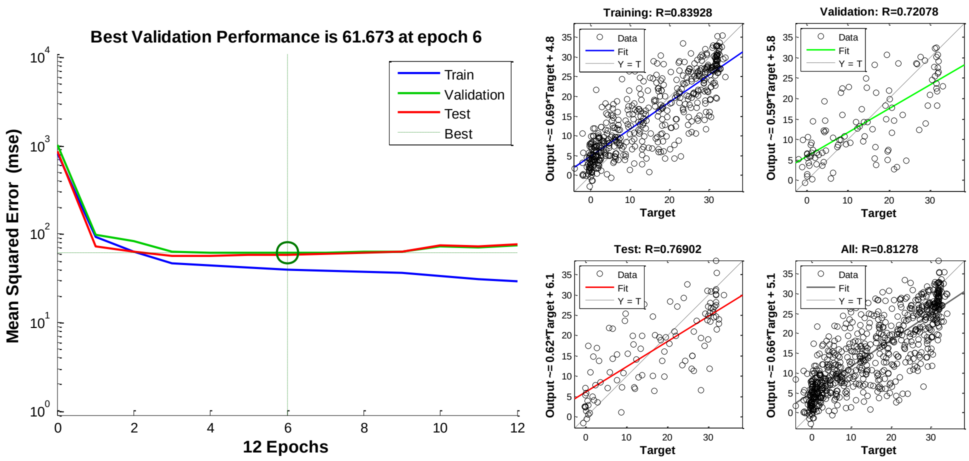
Figure 8. (a) Training performance; (b) Regression of NAR2 model for sub-series cD2_P



(a) (b)

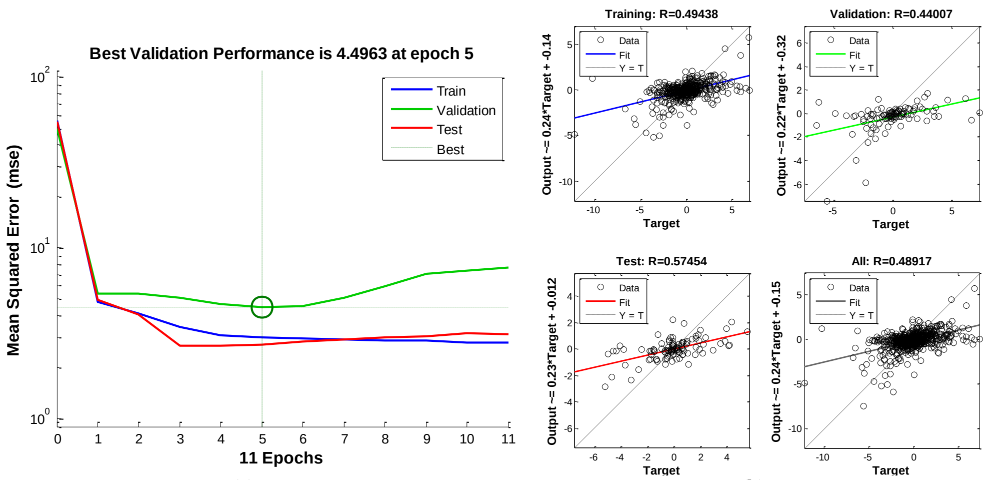
Figure 9. (a) Training performance; (b) Regression of NAR3 model for sub-series cD1_P

In the NARX model, the sub-series of wind speed data, attained from the wind farm, has been utilized as exogenous input to the neural networks. Due to the existence of higher cross-correlation of wind power with wind speed, wind speed has been considered as exogenous input. Along with wind speed as the exogenous input, NARX uses a feedback of output (wind power) as an input. For NARX modeling, input and feedback delays and hidden layer neurons for three models are respectively selected as 7 and 12, 2 and 10, and 2 and 10, using trial and error method. For each model, ten simulation runs are carried out with a feedback delay variation of 1:7 and hidden layer neurons of 5:20. The network performances and regressions of the three NARX models for each sub-series are respectively shown in Figures 10-12.



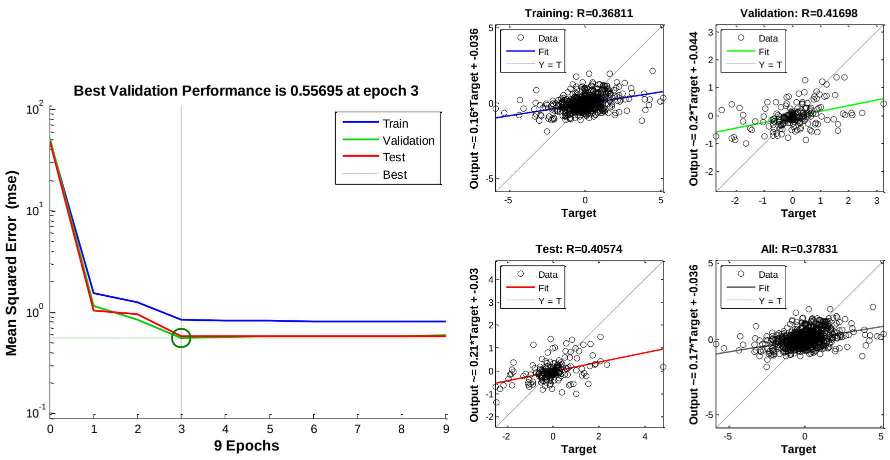
(a) (b)

Figure 10. (a) Training performance; (b) Regression for NARX1 model for sub-series cA2_P, cA2_WS



(a) (b)

Figure 11. (a) Performance; (b) Regression for NARX2 model for sub-series cD2_P, cD2_WS



(a) (b)

Figure 12. (a) Performance; (b) Regression for NARX3 model for sub-series cD1_P, cD1_WS

The 24 hours ahead forecast outputs of each sub-series, with the NAR and NARX models, have been aggregated separately to achieve the final forecast value using wavelet reconstruction in each case. The performances of the two models are compared to that of the persistence model and evaluated in terms of the MAE, MSE, and MAPE criteria. Figure 13 shows the actual observation and 24 hours ahead forecast with the persistence, WT-NAR, and WT-NARX models. The actual and forecast values of each model are given in Table 1 and shown in Figure 13. The performances of the persistence, WT-NAR, and WT-NARX forecast models in terms of MAE, MSE, and MAPE are given in Table 2. From Table 2, the superiority of WT-NARX model performance is revealed in terms of the three evaluation criteria.

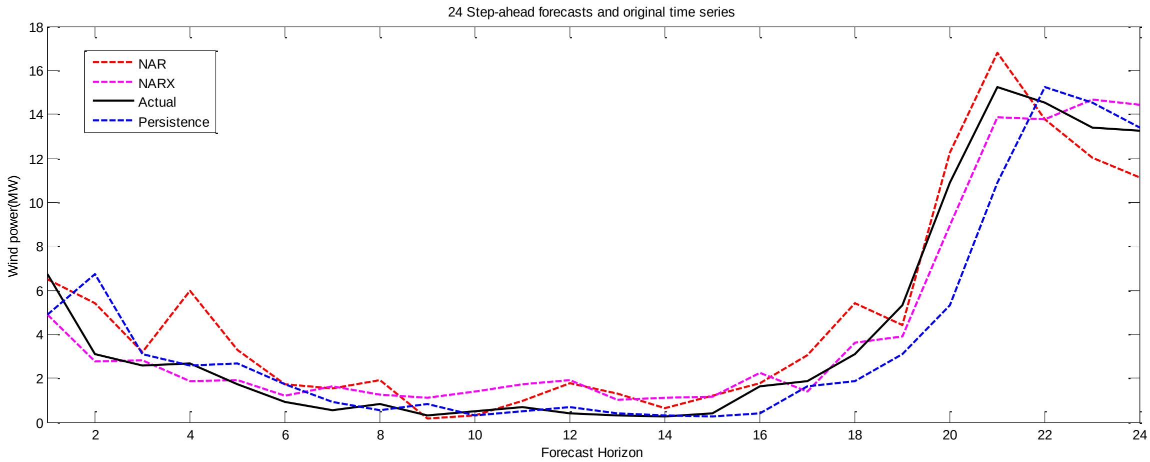


Figure 13. Actual and 24 hours ahead forecast with persistence, WT-NAR and WT-NARX models

Table 1. Actual and 24 hours ahead forecast values

Forecast Horizon	Actual	WT-NAR		WT-NARX		Persistence	
		Forecast	Error	Forecast	Error	Forecast	Error
1	6.721333	6.49284	0.22849	4.89563	1.8257	4.894083	1.82725
2	3.10325	5.41194	-2.3087	2.78347	0.31978	6.721333	-3.61808
3	2.554667	3.1995	-0.6448	2.81948	-0.2648	3.10325	-0.54858
4	2.664167	5.9734	-3.3092	1.86157	0.80259	2.554667	0.1095
5	1.704417	3.29923	-1.5948	1.93059	-0.2262	2.664167	-0.95975
6	0.91825	1.7364	-0.8182	1.18812	-0.2699	1.704417	-0.78617
7	0.559833	1.52847	-0.9686	1.62552	-1.0657	0.91825	-0.35842
8	0.80475	1.92406	-1.1193	1.22917	-0.4244	0.559833	0.244917
9	0.292667	0.15695	0.13572	1.10987	-0.8172	0.80475	-0.51208
10	0.485083	0.31038	0.17471	1.38818	-0.9031	0.292667	0.192417
11	0.6615	0.98641	-0.3249	1.70629	-1.0448	0.485083	0.176417
12	0.395083	1.75585	-1.3608	1.90841	-1.5133	0.6615	-0.26642
13	0.29575	1.31026	-1.0145	1.00967	-0.7139	0.395083	-0.09933
14	0.271333	0.63494	-0.3636	1.1072	-0.8359	0.29575	-0.02442
15	0.39125	1.21972	-0.8285	1.17498	-0.7837	0.271333	0.119917
16	1.630083	1.77433	-0.1442	2.26492	-0.6348	0.39125	1.238833
17	1.862917	3.0679	-1.205	1.39721	0.46571	1.630083	0.232833
18	3.072417	5.40069	-2.3283	3.59165	-0.5192	1.862917	1.2095
19	5.334667	4.40435	0.93032	3.8769	1.45776	3.072417	2.26225
20	10.90792	12.2543	-1.3464	8.96527	1.94265	5.334667	5.57325
21	15.23975	16.8158	-1.576	13.8517	1.38805	10.90792	4.331833
22	14.51	13.7589	0.75109	13.7672	0.74281	15.23975	-0.72975
23	13.39175	12.0428	1.34891	14.6549	-1.2632	14.51	-1.11825
24	13.26033	11.1126	2.14777	14.43	-1.1697	13.39175	-0.13142

Table 2. Performances of persistence, WT-NAR, and WT-NARX forecast models

Model	*MAE*	*MSE*	*MAPE*
WT-NAR	1.1239	1.8775	40.1058
WT-NARX	0.8915	1.0212	37.2980
Persistence	1.1313	3.2788	49.5338

5. Conclusion

In this paper, two NAR and NARX recurrent dynamic hybrid models based on wavelet decomposition have been used for short-term prediction of wind power. The wavelet based decomposition of wind power and wind speed data are carried out using Daubechies (db4) as the mother wavelet up to level 2, and the decomposed sub-series is fed as input to neural networks. In the NARX model, the wind speed time series is considered as exogenous input due to its higher cross-correlation with the wind power in comparison to wind direction and temperature. The model inputs and hidden layer size are determined using the trial and error method with minimum mean square error. The Levenberg-Marquardt (LM) algorithm has been utilized to train both models. The proposed hybrid models based on wavelet transform with univariate (NAR) and multivariate (NARX) neural networks are built up and compared to the persistence model. It is revealed from the simulation results that the WT-NARX network gives the minimum MAE, MSE, and MAPE values of 0.89, 1.02, and 37.30, respectively. In comparison, the MAE, MSE, and MAPE values are 1.12, 1.88, and 40.11 for WT-NAR and 1.13, 3.28, and 49.53 for persistence for 24 hours ahead prediction. The results validate the superiority and usefulness of the proposed approach for multi-steps ahead wind power forecasting.

References

1. B. Kanna and S. N. Singh, "AWNN-Assisted Wind Power Forecasting using Feed-Forward Neural Network," *IEEE Transactions on Sustainable Energy*, Vol. 3, No. 2, pp. 306-315, 2012
2. D. R. Chandra, M. S. Kumari, M. Sydulu, F. Grimaccia, and M. Mussetta, "Adaptive Wavelet Neural Network based Wind Speed Forecasting Studies," *Jo*urnal *of Electrical Engg & Technology*, Vol. 9, No. 6, pp. 1812-1821, 2014
3. J. P. S. Catalão, H. M. I. Pousinho, and V. M. F. Mendes, "Short-Term Wind Power Forecasting in Portugal by Neural Networks and Wavelet Transform," *Renewable Energy*, Vol. 36, No. 4, pp. 1245-1251, 2011
4. R. Ye, Z. Z. Guo, R. Y. Liu, and J. N. Liu, "Short-Term Wind Speed Forecasting Method based on Wavelet Packet Decomposition and Improved Elman Neural Network," in *Proceedings of IEEE 2016 International Conference on Probabilistic Methods Applied to Power Systems* (*PMAPS*), DOI:10.1109/PMAPS.2016.7764077, 2016
5. R. Aquino, M. M. S. Lira, J. B. de Oliveira, M. A. Carvalho, O. N. Neto, and G. Almeida, "Application of Wavelet and Neural Network Models for Wind Speed and Power Generation Forecasting in a Brazilian Experimental Wind Park," in *Proceedings of International Joint Conference on Neural Networks*, pp. 172-178, USA, 2009
6. D. L. Faria, R. Castro, C. Philippart, and A. Gusmão, "Wavelets Pre-Filtering in Wind Speed Prediction," in *Proceedings of IEEE 2009 International Conference on Power Engineering*, *Energy and Electrical Drives*, pp. 168-173, 2009
7. M. Madhiarasan and S. N. Deepa, "Performance Investigation of Six Artificial Neural Networks for Different Time Scale Wind Speed Forecasting in Three Wind Farms of Coimbatore Region," *International Journal of Innovation and Scientific Research*, Vol. 23, No. 2, pp. 380-411, 2016
8. R. Abhinav and N. M. Pindoriya, "Short-Term Wind Power Forecasting using Wavelet-based Neural Network," *Energy Procidia*, Vol. 142, pp. 455-460, 2017
9. A. Prasetyowati, H. Sudibyo, and D. Sudiana, "Wind Power Prediction by using Wavelet Decomposition Mode based NARX-Neural Network," in *Proceedings of the 2017 International Conference on Computer Science and Artificial Intelligence*, *Jakarta, Indonesia* (*CSAI 2017*), pp. 275-278, 2017
10. S. Saroha and S. K. Aggarwal, "Wind Power Forecasting using Wavelet Transforms and Neural Networks with Tapped Delay," *CSEE Journal of Power and Energy Systems*, Vol. 4, No. 2, pp. 197-209, 2018
11. W. Wei, G. Wu, M. H. Yang, Y. W. Zhang, S. X. Qiu, and A. Sun, "Short-Term Forecasting for Wind Speed based on Wavelet Decomposition and LMBP Neural Network," in *Proceedings of IEEE 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies* (*DRPT*), pp. 1126-1131, 2011
12. H. Liu, H. Q. Tian, C. Chen, and Y. F. Li, "An Experimental Investigation of Two Wavelet-MLP Hybrid Frameworks for Wind Speed Prediction using GA and PSO Optimization," *International Journal of Electrical Power and Energy Systems*, Vol. 52, pp. 161-173, 2013
13. H. R. Maier and G. C. Dandy, "Determining Inputs for Neural Network Models of Multivariate Time Series," *Computer-Aided Civil and Infrastructure Engineering*, Vol. 12, No. 5, pp. 353-368, 1997
14. V. Prema, K. U. Rao, B. S. Jnaneswar, C. A. Badarish, P. S. Ashok, and S. Agarwal, "Application of Hybrid Neuro-Wavelet Models for Effective Prediction of Wind Speed," *Intelligent Systems Technologies and Applications*, Vol. 384, pp. 345-354, Springer International Publishing Switzerland, DOI: 10.1007/978-3-319-23036-8_30, 2016
15. L. G. B. Ruiz, M. P. Cuéllar, M. D. Calvo-Flores, and M. D. C. P. Jiménez, "An Application of Non-Linear Autoregressive Neural Networks to Predict Energy Consumption in Public Buildings," *Energies*, Vol. 9, No. 684, pp. 1-21, 2016
16. F. H. Pereira, F. E. Bezerra, S. Junior, J. Santos, I. Chabu, G. Souza, et al., "Nonlinear Autoregressive Neural Network Models for Prediction of Transformer Oil-Dissolved Gas Concentrations," *Energies*, Vol. 11, No. 1691, pp. 1-12, 2018
17. A. S. Shekhawat, "Wind Power Forecasting using Artificial Neural Networks," *International Journal of Engineering & Technology* (*IJERT*), Vol. 3, No. 4, pp. 993-998, 2014
18. C. Erasmo, R. Wilfrido, C. A. Rafael, and C. Roberto, "Wind Speed Forecasting using the NARX Model, Case: La Mata, Oaxaca, México," *Springer Neural Computing and Applications*, Vol. 27, No. 8, pp 2417-2428, 2015
19. C. Erasmo, R. Wilfrido, C. A. Rafael, and H. Christopher, "Wind Speed Prediction using a Univariate ARIMA Model and a Multivariate NARX Model," *Energies*, Vol. 9, No. 109, pp. 1-15, 2016