Transactions of the Nigerian Association of Mathematical Physics
Volume 5,(September and November, 2017), pp183 –192
© Trans. of NAMP

Application of Chaos Theory in the Optimization of NARX Neural Network Configurations for the Prediction of Solar Radiation and Wind Speed in Makurdi

[1] Tikyaa E. V., [2] Tyokyaa R. K. and [3] Anjorin F. O.

[1]Department of Physics, Federal University Dutsin-Ma, Katsina State, Nigeria
[2]Department of Mathematical Sciences and Information Technology, Federal University Dutsin-Ma, Katsina State, Nigeria
[3]Department of Physics, University of Jos, Plateau State, Nigeria.

Abstract

This paper applies chaos theory in the optimization of NARX neural network configurations for the prediction of solar radiation and wind speed in Makurdi. The meteorological data used was obtained from the Nigerian Meteorological Agency (NiMeT), Abuja-Nigeria. It comprises of records of daily averages of wind speed, solar radiation, minimum temperature, maximum temperature, maximum relative humidity and sunshine hours; with 70 % of the data used for training and 30 % used for validation and testing of the NARX models. The intrinsic parameters of the model which include the number of time delay lines (TDL), hidden neurons and hidden layers were optimized using the predetermined chaos dynamics of the meteorological data using Taken's embedding theorem in order to get the right neural network configuration, save time and ensure more accurate predictions. The number of time delay lines was optimized by evaluating the time delay of the training data using the method of average mutual information while the number of hidden neurons in the hidden layers were evaluated in relation to the embedding dimension of the training data set which was obtained from the method of false nearest neighbors. The results of the model testing showed that, the model performed better using the Bayesian Regularization training function with daily solar radiation successfully predicted using wind speed, minimum temperature and maximum relative humidity as exogenous variables within a RMSE of 2.6763 and correlation coefficient of 0.8791 in an 11-3-1 configuration with 9 TDL; while wind speed was also successfully predicted using solar radiation, maximum temperature and sunshine hours as exogenous variables within a RMSE of 1.1429 and correlation coefficient of 0.7607 in a 13-4-1 configuration with 13 TDL for the 365 day ahead predictions made. It is also worthy to note that after trying several configurations of NARX network models using sequence the trial and error method, the optimized configurations obtained which had the highest correlation between the observed and target output as well as the least mean square error were found to be almost spot on with those predetermined using the application of chaos theory vis-à-vis the chaos dynamics of the training data.

Keyword: Multi-switching, combination synchronization, high-dimensional systems, hyperchaotic systems

1. INTRODUCTION

Weather forecasting is a complex exercise as a result of the chaotic nature of the weather and climate change. The fact that weather is a non-linear phenomenon, traditional forecasting methods are simply not suitable for this application due to their lack of nonlinear mapping ability, hence the emergence of artificial intelligence packages such as artificial neural network (ANN). ANN is a biologically inspired tool which has the ability to learn a pattern from a set of input data and generate/predict the same outputs as with the learning data when presented with a new set of data which contains a similar pattern to that which the ANN was trained with [1]. ANN mimics the behavior of the human nervous system and resembles the brain in two respects: ANN models can recognize trends/patterns and learn from their interactions with the environment [2]. Over the years, so many researchers have applied ANN in the modelling and prediction of different systems, due to its versatility and ability to learn new patterns and have observed that the NARX seems to produce the best results for nonlinear dynamical systems. Diaconescu [3] investigated the performance of the prediction for different time series using the Nonlinear Auto Regressive with eXogenous inputs (NARX) dynamic recurrent neural network.

Correspondence Author: Tikyaa E.V., E-mail: etikyaa@fudutsinma.edu.ng, +2347032096960

Comparative analysis with real and artificial chaotic time series from diverse domains of physical systems were made and the results obtained inferred that NARXrecurrent neural network has the potential to capture the dynamics of nonlinear dynamical system such as the Mackey Glass system with different delays. This affirmation was based on the fact that the correlation coefficient R estimated for the original and generated time series is close to 1 in many cases, and the prediction can be considered of real interest or significance if R> 0.98.

Di Piazza *et al.* [4]carried out a prediction and forecast of climate time series, for better planning and management of the power grid from wind power generation using different neural networks. Two dynamic recurrent ANNs; the focused time-delay neural network (FTDNN) and the nonlinear autoregressive network with exogenous inputs (NARX), were used to develop a model for the estimate and forecast of daily wind speed. The daily wind speed and the daily maximum and minimum temperature in the period between 2010 and 2012 registered on Palermo weather station, in the northeast of Sicily were used as dataset to train the ANNs. The ANNs-based models were experimentally validated and they both showed good performance since reliable and precise representations of daily wind speed were obtained. Rocha Vaz [5] addressed the variability of PV electric generation based on the premise that the performance and management of small electric networks can be improved when solar power forecast information is used. A neural network architecture system for the Nonlinear Autoregressive with eXogenous inputs (NARX) was implemented using not only local meteorological data but also measurements of neighboring PV systems. The input configuration with the best network performance was selected by optimization and forecasts of up to several hours in advance were tested to verify the model forecasting accuracy for different short-term time horizons and this was compared with the persistence model. The NARX model clearly outperformed the persistence model and yielded a 3.7% and a 4.5% *RMSE* for the anticipation of the 5min and 2hr 30min forecasts, respectively.

Ahmad and Anderson [6]undertook a study using NARX to predict global solar radiation across New Zealand. Data for nine hourly weather variables recorded across New Zealand from January 2006 to December 2012 were used to create, train and test Artificial Neural Network (ANN) models using the Levenberg−Marquardt (LM) training algorithm, with global solar radiation as the target parameter. ANN models with different numbers of neurons (from 5 to 250) in the hidden layer as well as different numbers of delays were experimented with, and their effect on prediction accuracy were analyzed. The predicted values of hourly global solar radiation were compared with the measured values, and it was found that the mean squared error (MSE) and regression (R) values showed close correlation. As such, the study illustrated and affirmed the reliability of the model for solar radiation forecasting at a later time, thus demonstrating the generalization capability of the approach over unseen data.Sandhya and Kavitha [7] proposed a system used to predict solar radiation based on a non-linear autoregressive exogenous input model and found the suitability of crops for cultivation based on the atmospheric conditions. The daily average solar radiation for Tiruvallur region was correctly predicted using input climatic parameters such as minimum air temperature, maximum air temperature, air pressure and humidity that influence the crop yield for Tiruvallur region. A non-linear autoregressive exogenous input (NARX) was built to predict the solar radiation and the performance of the network was analyzed by calculating the mean squared error and Regression analysis to determine the accuracy between the actual data and predicted data. Based on the accuracy achieved from the solar radiation prediction along with other climatic parameters such as the amount of rainfall, soil, duration and suitable months the crops that were suitable for cultivation around Tiruvallur region were identified and suggestions of the suitable planting times were made to farmers.

All these works are a testimony of the efficacy of NARX in modelling and prediction of different nonlinear and chaotic systems. But in all, the problem of optimization of neural network configuration remains a challenge that needs to be properly addressed. Hence in this paper, chaos theory is applied to NARX models to help select the optimal neural network configuration for wind speed and solar radiation prediction in Makurdi.

1. Theoretical Framework

1.1 Takens' embedding theorem and phase space reconstruction

Takens [8]extended Whitney's embedding theorem by proving that for any system, an embedding can be obtained from a single time series of infinite precision by using $(2D+1)$delay coordinates, where $D$is the fractal dimension. Hence phase space reconstruction is usually done so as to draw out a multi-dimensional description of system in an embedded space called state space. This can be simply achieved using the method of delay (MOD) [8, 9]. For a time series $\{x_1, x_1,\ldots,x_N\}$, with Nbeing the number of data points, the attractor (collection of closely packed trajectories) can be reconstructed into an m-dimensional phase space of delay coordinates by forming the following vectors:

$$X_n = [x_n, x_{n+\tau}, x_{n+2\tau}, \ldots, x_{n+(m-1)\tau}] \quad (1)$$

where τ is the delay time, and the integer m is the embedding dimension. The phase space vector $X_n$is an$M \times m$matrix, and the constants m, M, τ, and N are related by:

$$M = N - (m - 1)\tau \quad (2)$$

Mbeing the number of reconstructed phase space points. Furthermore, [8] postulated that:

$$m \geq 2D + 1 \quad (3)$$

Equation 3 is the focal point of our NARX configuration optimization using Chaos theory.

1.2 NARX neural network

Nonlinear autoregressive network with exogenous inputs (NARX) is a recurrent dynamic network, which has feedback connections that cover several layers of the network. The model equation for the NARX model is given in[10]:

$y(n+1) = f\,[y(n), \ldots, y(n-d_y+1); u(n), u(n-1), \ldots, u(n-d_u+1)]$ (4)

$u(n)$and$y(n)$ are real valued functions and denote the input and output regressors of the model at time step n respectively. The parameters$d_u \geq 1, d_y \geq 1,\ and\ d_u \leq d_u$, are the input and output delays respectively while the nonlinear function$f[\cdot]$ is unknown and is usually approximated using a feed forward multilayer perceptron (MLP) network [11]. The predicted value of the dependent output signal $y(n+1)$ is often regressed from the preceding values of the output signal and preceding values of an independent (exogenous) input signal. In this work, the NARX network created was used to train and predict the target signal and carry out a step-ahead prediction of the output signal, hence it could act as a nonlinear filter, in which the target output is a noise-free version of the input signal especially when dealing with chaotic nonlinear dynamic systems [3].

2. Materials and Methods

2.1 The Data source

The secondary weather data for Makurdi covering eleven years (2006-2016) was collected from the Nigerian Meteorological Agency (NiMeT), Abuja. The data set includes: wind speed (km/hr), solar radiation ($MJm^{-2}day^{-1}$), minimum temperature (℃) and maximum temperature(℃), minimum relative humidity (%), maximum relative humidity (%) and sunshine hours (hr).The data was divided into three sets, the training set corresponding to 70% of the data, the target set, corresponding to 15% of the data and the testing/validation set, corresponding to 15% of the data.

2.2 NARX model design and Optimization using Chaos Theory

Designing NARX models requires that one must follow a number of systemic procedures. The six basics steps followed in predicting with NARXin this work include:

i. Pre-processing of data by sorting out the missing data and designating the input/target variables using correlation analysis.

ii. Creating the network: here some key parameters required to create a suitable network are assigned. Some of these include: choosing the network name, the number of hidden layers, neurons in each layer, the number of time delay lines, transfer function in each layer, training function, weight and bias learning function, data division function and performance function. A feedforward network was createdusing*'narxnet'*which uses the tan-sigmoid transfer function as default and linear transfer function in the output layer.

iii. Choice of the number of time delay lines, hidden layers and neurons: This is key in optimizing the performance of the network. The *narxnet*created has two types of input: the external input $x(t)$ and the feedback connection from the network output. Each of these inputs has a tapped delay line to store previous values.

The number of time delay lines (TDL) in the NARX network was chosen as the time delay of the data set which was obtained by applying the method of average mutual information, AMI [12]:

$$I(\tau) = -\sum_{ij}^{N} P_{ij}(\tau) ln \frac{P_{ij}(\tau)}{P_i(\tau)P_j(\tau)} \quad (5)$$

$P_i(\tau)$and$P_j(\tau)$ are the probabilities of finding the values$x_i$in the ith and $x_{i+\tau}$in the jth intervals, and $P_{ij}(\tau)$ is their joint probability. By plot$I(\tau)$ against the lag length τ, the first local minimum of the curve corresponds to the delay time τ[13].The hyperbolic tangent sigmoid (tan-sig) was used as the transfer function to carry data across the hidden layers via the neurons.

In this model, two hidden layers were used in addition to the input and output layerswhile in determining the number of neurons in each hidden layer, chaos theory was used to optimize the number of hidden neurons in the 1st and 2nd hidden layers ($N_{H1} and\ N_{H2}$) in accordance with Takens' embedding theorem (equation 2).This is applied and summarized in the expression [14]:

$$N_{H1} = 2m+1, N_{H2} = round(\sqrt{N_{H1}}) \quad (6)$$

The parameter$m$ is the embedding dimension ofthe target data obtained using the method of false nearest neighbors, FNN. This criterion for designating a false nearest neighbor is given by [15]:

$$\left[\frac{R_{m+1}^2(n,r)-R_m^2(n,r)}{R_m^2(n,r)}\right]^{\frac{1}{2}} = \frac{|x(n+m\tau)-x^{(r)}(n+m\tau)|}{R_m^2(n,r)} > R_{tol} \quad (7)$$

Where;

$$R_m^2(\text{n,r}) = \sum_{k=0}^{m-1}\left[x(n+k\tau)-x^{(r)}(n+k\tau)\right]^2 \quad (8)$$

and

$$R_{m+1}^2(\text{n,r}) = R_m^2(\text{n,r}) + \left[x(n+k\tau)-x^{(r)}(n+k\tau)\right]^2 \quad (9)$$

$x^{(r)}(n)$ is the rth nearest neighbor of x(n). R_m denotes the Euclidean distance (norm) in phase space between nearest neighbors with embedding dimension m,τ is the time delay and R_{tol} is the tolerance threshold. The graph of the percentage of FNN against increasing embedding dimension and has a monotonic decreasing graph and the minimum embedding dimension is usually evaluated at the point where the percentage of FNN drops to almost zero or a minimum value [15].

iv. Data preparation for training: in training the NARX network which has tapped delay lines, it is mandatory to fill delays with initial values of the inputs and outputs of the network. This is implemented in MATLAB using the *'preparets'* function.

v. Training the network: the training process involves a systematic adjustment of the weights until the predicted output generated is very close in value and trend to the target (measured) output of the network. The comparison of the output signal with the target output produces an error signal which activates a control mechanism during the iterative process and then applies a sequence of corrective adjustments of the weights and biases of the neuron using the weight/bias learning functions. These corrective adjustments will continue until the training data attains the desired mapping to obtain the target output as closely as possible. After a series of *training epochs* the neural network was successfully trained and the weights were saved.

vi. Testing and validation of the model performance using some performance evaluation functions. These are statistical tools used to test the accuracy of the network training by testing the variation between the actual and predicted output data sets. The functions used in the NARX training in this research are listed in Table 1 [16]:

Table 1. Neural Network Performance Functions

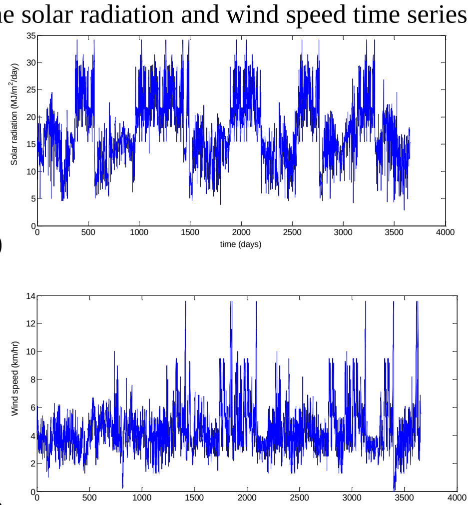
Function	Name	Algorithm
MSE	Mean squared error	$\frac{1}{n}\sum_{i=1}^{n}\left(y_a(i)-y_p(i)\right)^2$
RMSE	Root mean squared error	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_a(i)-y_p(i)\right)^2}$
MAPE	Mean absolute percentage error	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_a(i)-y_p(i)}{y_a(i)}\right\rvert \times 100\%$
R	Pearson's correlation coefficient	$\frac{n\sum y_a y_p-(\sum y_a)(\sum y_p)}{\sqrt{n(\sum {y_p}^2)-(\sum y_a)^2}\sqrt{n(\sum {y_p}^2)-(\sum y_a)^2}}$

y_a is the observed output and y_p is the predicted output.

3. Results and Discussion

3.1 Solar Radiation and Wind speed Time Series

The solar radiation and wind speed time series for Makurdi used in the network training are displayed in Figure 1.



(a)

(b)

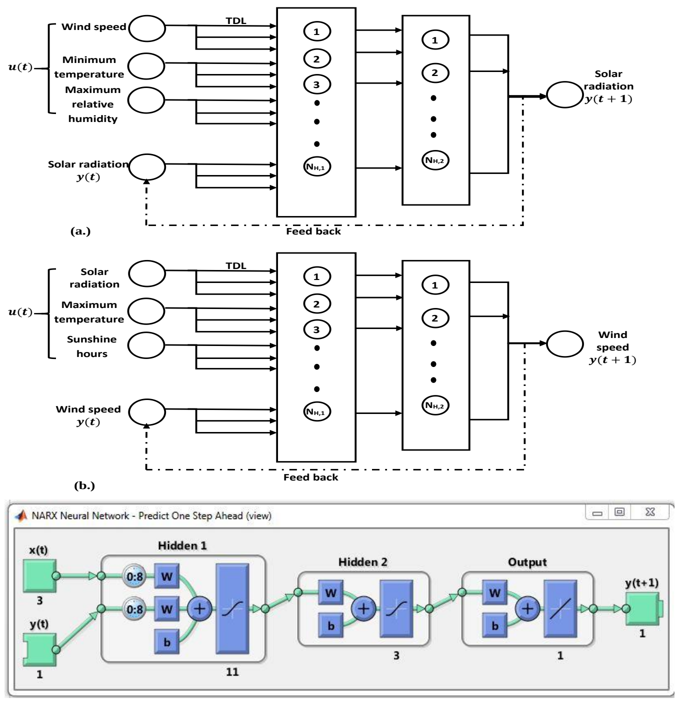
Figure 1. (a)Solar radiation and (b) wind speed time series for Makurdi from 2006-2015

3.2 NARX model selection

Based on the correlation coefficients of the different variables earlier computed and displayed in Table 2, the models in Figure 2 were developed for the prediction of solar radiation and wind speed.

Table 2: Correlation coefficients of the meteorological parameters for Makurdi from 2006 to 2015

Parameter	Wind speed	Solar radiation	Min. temperature	Max. temperature	Min. relative humidity	Max. relative humidity	Sunshine hours
Wind speed	1.0000	-0.0857	-0.0227	-0.0665	0.0027	-0.0460	-0.1077
Solar radiation	-0.0857	1.0000	0.0930	0.0128	0.0509	0.1071	0.0037



(c.)

Figure 2. NARX model (parallel architecture) designs for: (a.) solar radiation, (b.) wind speed prediction, (c.) Simulink view of the NARX neural network for prediction.

In order to optimize the time delay, hidden neurons and hidden layer sizes, the tenets of chaos theory was utilized as explained earlier in section 3.2(iii).The time delay and embedding dimension for solar radiation and wind speed time series test data from 2006 to 2015 were estimated and the results are presented in Figures 3 and 4 respectively.

(a) (b)

Figure 3. Estimation of NARX network parameters for solar radiation prediction: (a.) time delay, τ=9days (i.e. 9 TDL); (b.) embedding dimension, m=5 (11-3-1 configuration)

(a) (b)

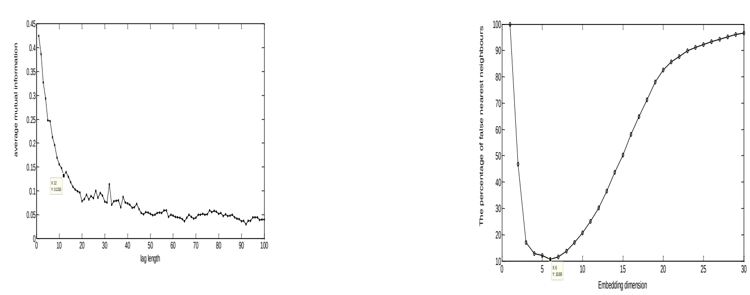


Figure 4. Estimation of NARX network parameters for wind speed prediction: (a.) time delay, τ=12days (i.e. 12 TDL); (b.) embedding dimension, m=6 (13-4-1)

3.3 Results of the NARX Models Training and Prediction

In applying the selected NARX networks for the prediction of solar radiation and wind speed in Makurdi, ten years data (2006-2015) of both the input and target series was used as training data i.e. 3650 points with the extra day in the leap years (February 26th) removed, while one year (2016) was used as target data i.e. points 3650-4015, for validation of the network created. The results of the performance of the NARX network models' training for solar radiation and wind speed prediction for Makurdi in 2016 using the Levenberg-Marquardt and Bayesian regularization training functions which was validatedusing the root-mean-square error and correlation coefficient of the target and output are displayed in Tables 3-6. Figures 5-8 show a display of the network predictions and actual data for the best performing configuration of the different NARX models.

Table 3: NARX model selection for solar radiation prediction in Makurdi using the Levenberg-Marquardt ('*trainlm*') training function

			No. of time delay lines (TDL)					
			8		9		10	
m	$N_{H,1} = 2m + 1$	$N_{H,2} = round(\sqrt{N_{H,1}})$	$RMSE$	R	$RMSE$	R	RMSE	R
5	**11**	**3**	2.7732	0.8893	3.4191	0.8253	3.5587	0.8112
6	13	4	3.4182	0.8296	3.5477	0.8158	3.1814	0.8512
7	15	4	3.4234	0.8244	3.1952	0.8519	3.1816	0.8512
10	21	5	3.4372	0.8304	3.1001	0.8591	3.3690	0.8401
20	41	6	3.0595	0.8629	4.7211	0.6255	3.1220	0.8572

From the results in Table 3, the best model is 11-3-1 with 8 TDL, as this has the highest correlation coefficient (R) and the least root-mean-square error ($RMSE$). 11-3-1 means there are 11 neurons, 3 neurons and 1 neuron in the first hidden, second hiddenand output layers respectively. Figure 5 shows a plot of the network output and original outputof the model.

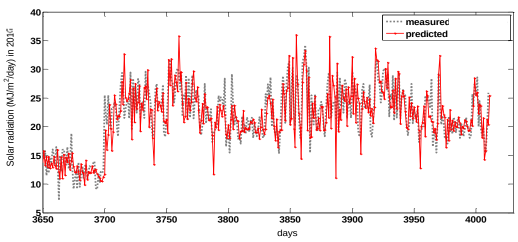


Figure 5. Prediction plot of daily solar radiation in Makurdi for 2016 using Levenberg-Marquardt training function

Table 4: NARX model selection for solar radiation prediction in Makurdi using the Bayesian regularization ('*trainbr*') training function

			No. of time delay lines (TDL)					
			8		9		10	
m	$N_{H,1} = 2m + 1$	$N_{H,2} = round(\sqrt{N_{H,1}})$	$RMSE$	R	$RMSE$	R	RMSE	R
5	**11**	**3**	2.7480	0.8911	2.6763	0.8971	3.0058	0.8679
6	13	4	2.9522	0.8728	3.0013	0.8682	3.0679	0.8619
7	15	4	3.1113	0.8576	2.9946	0.8689	3.0969	0.8593
10	21	5	3.0103	0.8690	3.1000	0.8588	2.8295	0.8840
20	41	6	3.1149	0.8572	3.1216	0.8566	2.8363	0.8835

The results in Table 4 indicate that the best model is 11-3-1 with 9 TDL; this has the highest correlation coefficient (R) and the least root-mean-square error ($RMSE$) and is the exact configuration obtained using the time delay and embedding dimension of the target data. Figure 6 shows a plot of the network output, original output and the network error of the model.

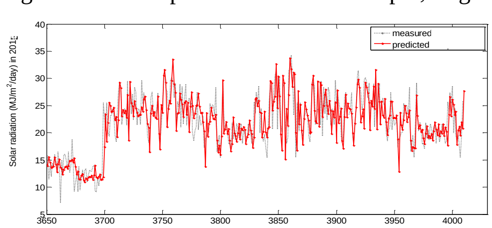
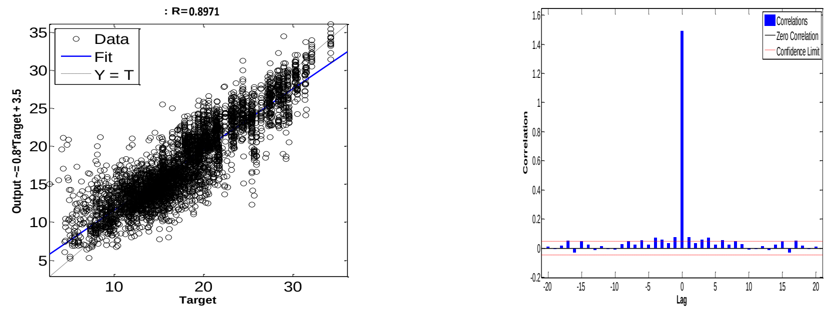


Figure 6. Prediction plot of daily solar radiation in Makurdi for 2016 using Bayesian regularization training function.

Figure 7 further shows both the output regression of the best performing configuration of the NARX model (11-3-1; 9 TDL) obtained using the Bayesian regularization training function and the autocorrelation of errors between the observed and predicted values of daily solar radiation. It is observed that most of the errors fall within the 95% confidence interval which makes it a good prediction.



(a) (b)

Figure 7. Performance validation of daily solar radiation prediction for Makurdi in 2016 using Bayesian Regularization Training Function; (a) Correlation of observed and predicted Series, (b) Autocorrelation of Errors.

Table 5: NARX model selection for wind speed prediction in Makurdi using the Levenberg-Marquardt ('*trainlm*') training function

			No. of time delay lines (TDL)					
			11		12		13	
m	$N_{H,1} = 2m+1$	$N_{H,2} = round(\sqrt{N_{H,1}})$	$RMSE$	R	$RMSE$	R	$RMSE$	R
5	11	3	1.2796	0.7179	1.1870	0.7432	1.2576	0.7007
6	**13**	**4**	**1.1670**	**0.7487**	1.2260	0.7275	1.2748	0.7017
7	15	4	1.2193	0.7259	1.2117	0.7317	1.2045	0.7313
10	21	5	1.1958	0.7342	1.1704	0.7530	1.2022	0.7316
20	41	6	1.2052	0.7366	1.3333	0.6801	1.4424	0.5735

The results in Table 5 indicate that the best model is 13-4-1 with 11 TDL; as this has the highest correlation coefficient (R) and the least root-mean-square error ($RMSE$). Figure 8 shows a plot of the network output and original output for the 13-4-1 model.

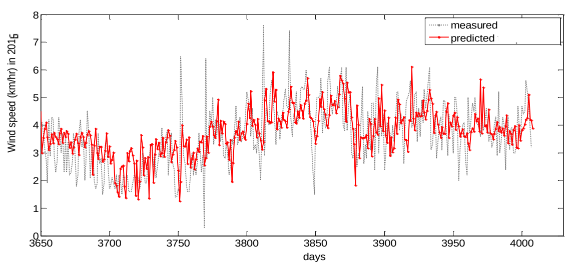


Figure 8. Prediction plot of daily wind speed in Makurdi for 2016 using Levenberg-Marquardt training function

Table 6: NARX model selection for wind speed prediction in Makurdi using the Bayesian regularization ('*trainbr*') training function

			No. of time delay lines (TDL)					
			11		12		13	
m	$N_{H,1} = 2m+1$	$N_{H,2} = round(\sqrt{N_{H,1}})$	$RMSE$	R	$RMSE$	R	RMSE	R
5	11	3	1.2025	0.7305	1.2058	0.7292	1.2274	0.7254
6	**13**	**4**	1.2142	0.7246	1.2014	0.7312	**1.1429**	**0.7607**
7	15	4	1.2038	0.7300	1.2083	0.7280	1.2140	0.7255
10	21	5	1.2049	0.7293	1.1902	0.7373	1.1643	0.7502
20	41	6	1.1896	0.7372	1.1996	0.7323	1.2062	0.7301

The results in Table 6 shows that the best model is 13-4-1 with 13 TDL; as this has the greatest correlation coefficient (R) and the least root-mean-square error ($RMSE$). Figure 9 shows a plot of the network output and original output of this model.

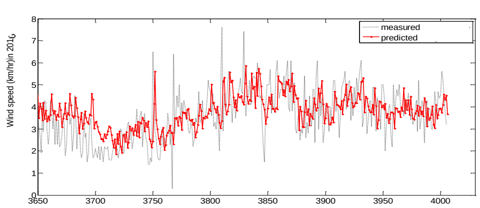
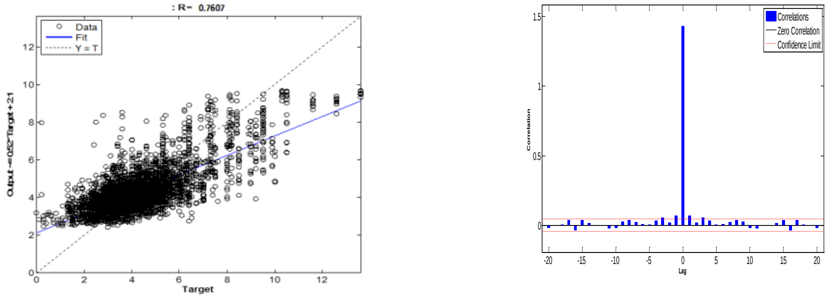


Figure 9. Prediction plot of daily wind speed in Makurdi for 2016 using Bayesian regularization training function

Figure 10shows both the output regressions of the best performing configurations of the NARX model (13-4-1; 13 TDL) using the Bayesian regularization training function and the autocorrelation of errors between the observed and predicted values of daily wind speed.Here also most of the errors fall within the 95% confidence interval indicating an acceptable prediction. Even though the NARX configuration was properly optimized, the averagely fair value of correlation coefficient of the best performing model is as a result of the noisy nature of the data used which resulted in weak correlation of input and target variables as shown in Table 2.



(a) (b)

Figure 10. Performance validation of daily wind speed prediction for Makurdi in 2016 using Bayesian Regularization Training Function; (a) Correlation of observed and predicted Series, (b) Autocorrelation of Errors.

From the results in Tables 3, 4, 5 and 6, it is observed that the Bayesian regularization training function outperforms the Levenberg-Marquardt training function. The is because it is more robust,are difficult to overtrain, since evidence procedures provide an objective Bayesian criterion for stopping training, are also difficult to overfit, because the Bayesian regularization training function calculates and trains on a number of effective network parameters or weights, effectively turning off those that are not relevant; as this effective number is usually considerably smaller than the number of weights in a Levenberg-Marquardt back-propagation neural network [17]. Hence it performs better when applied to chaotic time series, even though Levenberg-Marquardt is generally faster[10].

4. Conclusion

ANARX neural network model was successfully developed and its configuration optimized for the prediction of daily solar radiation and wind speed in Makurdi. The intrinsic parameters of the models were optimized using the predetermined chaos dynamics of the meteorological data in accordance with Takens' embedding theorem. Results of the model validation shows that, the models perform better using the Bayesian regularization training function predicting solar radiation successfully using wind speed, minimum temperature and maximum relative humidity as exogenous variables within a $RMSE$ of 2.6763 and correlation coefficient of 0.8791; while wind speed was also successfully predicted using solar radiation, maximum temperature and sunshine hours as exogenous variables within a $RMSE$ of 1.1429 and correlation coefficient of 0.7607 for the 365 day ahead forecast in 2016. Even though the trends were accurately predicted, the relatively fair but acceptable $RMSE$ values obtained was as a result of the noisy nature of the meteorological variables used which could be as a result of some missing data which may have arisen from faulty equipment or human error/neglect of duties.

References

[1] Amakali, S. (2008). Development of Models for Short-term Load Forecasting using Artificial Neural Networks. M.Sc. Dissertation, Department of Electrical Engineering Cape Peninsula University of Technology, Cape town, South Africa, Paper 32, pp. 44-53.

[2] Haykin, S. (2009). Neural Networks and Learning Machines, 3rd edition. New Jersey, Pearson Education, Inc., p. 37.

[3] Diaconescu, E. (2008). The Use of NARX Neural Networks to Predict Chaotic Time Series. *WSEAS Transactions on Computer Research*, 3(3): 182-191.

[4] Di Piazza, A., Di Piazza, M. C. and Vitale, G. (2014). Estimation and Forecast of Wind Power Generation by FTDNN and NARX-net based Models for Energy Management Purpose in Smart Grids. International Conference on Renewable Energies and Power Quality (ICREPQ '14), 12, 8th-12th April, 2014.

[5] Rocha Vaz, A. G. (2014). Photovoltaic Forecasting with Artificial Neural Networks. M.Sc. Dissertation, Department of Engineering and Geography, Universidade de Lisboa, p. 57.

[6] Ahmad, A. and Anderson, T. (2014) Global Solar Radiation Prediction using Artificial Neural Network Models for New Zealand. Solar2014: The 52nd Annual Conference of the Australian Solar Council, School of Engineering, Auckland University of Technology, Auckland, 1010, New Zealand.

[7] Sandhya, T. and Kavitha, V. R. (2015). A Neural Network Based Approach to Predict Solar Radiation and Analyze the Suitability of Crops Based on Atmospheric Conditions. *International Journal of Advanced Research in Computer Engineering and Technology*, 4(5): 124-133.

[8] Takens, F. (1981). Detecting strange attractors in turbulence. In Rand, D. A. and Young, L. S.(eds.), Dynamical Systems and Turbulence.*Lecture Notes in Mathematics*, **898**: 366–381. Springer.

[9] Packard, N., Crutchfield, J., Farmer, J. and Shaw, R. (1980). Geometry from a Time Series. *Phys. Rev. Lett.*, **45**: 712.

[10] Beale, M. H., Hagan, M.T. and Demuth, B.H. (2014). Neural Network Toolbox™. User's Guide (R2014b). USA, TheMathWorks, Inc., pp. 23-34.

[11] Menezes -Jnr, J. M. and Barreto, G. A. (2007a). Long-Term Time Series Prediction with the NARX Network: An Empirical Evaluation. Preprint submitted to Elsevier Science, Department of Teleinformatics Engineering, Av. Mister Hull, S/N, CP 6005, CEP 60455-760, Fortaleza, Cear´a, Brazil, 15 pp.

[12] Cellucci, C., Albano, A. and Rapp, P. (2003). Comparative Study of Embedding Methods. *Physical Review E.*, 67(6): 1-13.

[13] Echi, M.I., Tikyaa, E.V. and Isikwue, B.C. (2015). Dynamics of Rainfall and Temperature in Makurdi, *International Journal of Science and Research*, **4** (7): 493-499.

[14] Menezes-Jnr, J. M. and Barreto, G. M. (2007b). Multistep-Ahead Prediction of Rainfall Precipitation Using the NARX Network. Department of Teleinformatics Engineering, Av. Mister Hull, S/N, CP 6005, CEP 60455-760, Fortaleza, Cear´a, Brazil, 12 pp.

[15] Kennel, M. B., Brown, R. and Abarbanel, H.D. (1992). Determining Embedding Dimension for Phase-Space Reconstruction using a Geometrical Construction. *Physical Review A*, 45(6): 3403-3411.

[16] Baghirli, O. (2015). Comparison of Lavenberg-Marquardt, Scaled Conjugate Gradient and Bayesian Regularization Backpropagation Algorithms for Multistep Ahead Wind Speed Forecasting Using Multilayer Perceptron Feedforward Neural Network. Published M.Sc. Thesis in Energy Technology With Focus on Wind Power, Dept. of Earth Sciences, Campus Gotland, Uppsala University, pp. 25-31.

[17] Burden, F. and Winkler, D. (2008). Bayesian Regularization of Neural Networks. In: Livingstone, D.J. (eds.),Artificial Neural Networks. *Methods in Molecular Biology™*, **458.** Humana Press, Springer science.