International Journal of Mathematics and Computational Science
Vol. 4, No. 2, 2018, pp. 35-47
http://www.aiscience.org/journal/ijmcs
ISSN: 2381-7011 (Print); ISSN: 2381-702X (Online)

A Hybrid SARIMA-NARX Nonlinear Dynamics Model for Predicting Solar Radiation in Makurdi

Emmanuel Vezua Tikyaa[1, *], Matthias Idugba Echi[2], Bernadette Chidomnso Isikwue[2], Alexander Nwabueze Amah[2]

[1]Department of Physics, Federal University Dutsin-Ma, Katsina State, Nigeria
[2]Department of Physics, University of Agriculture Makurdi, Benue State, Nigeria

Abstract

In this paper a hybrid SARIMA-NARX neural network model was successfully developed, trained using 16 years data obtained from the Nigerian Meteorological Agency (NIMET) and tested by forecasting daily solar radiation time series in Makurdi. The intrinsic parameters of the model was optimized using the predetermined nonlinear dynamics of the meteorological data in order to get the right neural network configuration, save time and ensure accurate forecasts. The results of the model testing showed that, the model performed better and faster using the Levenberg-Marquardt training function with daily solar radiation successfully forecasted using minimum temperature and maximum temperature as exogenous variables. The daily solar radiation in Makurdi for the year 2016 was successfully predicted to validate the model using the hybrid model generating a RMSE value of 1.6475, correlation coefficient of 0.8782, MAE of 1.2042 and MAPE of 5.9695%. After validation, forecasts of daily solar radiation were then made for 2016 and 2017 with quite good accuracy recorded. It was also observed that the data trends were accurately predicted as a result of the SARIMA model adopted while the NARX model generated the nonlinear part of the time series with relatively fair but acceptable RMSE values which could be as a result of the poor correlation of the meteorological variables emanating from the presence of missing data and noise in the meteorological data used.

Keywords

NARX, SARIMA, Neural Network, Chaos Theory, Forecasting

Received: March 19, 2018 / Accepted: April 3, 2018 / Published online: May 28, 2018

1. Introduction

Weather forecasting is a complex exercise as a result of the chaotic nature of the weather [1] and climate change [2]. The fact that weather is a non-linear phenomenon, traditional regression methods used in forecasting are simply not suitable for the application due to the lack of nonlinear mapping ability. Chaos theory on the other hand when applied to artificial neural network (ANN) optimal structure brings out the dynamics of the system and thus improves the accuracy of the forecast as well as the performance of the network. Meteorologists today forecast the weather using a combination of numerical computer models, observations via satellite imaging, and a fore knowledge of previous trends and patterns to obtain reasonable and precise forecasts of up to several days in advance. These techniques are not readily available to all and sundry for application when required and so one has to rely on weather broadcasts on the mass media and from the internet. That notwithstanding, these forecast are restricted to only major towns and cities and are usually on a very short term. Hence the development of local numerical weather prediction techniques (models) that can give a short-term forecast of these meteorological variables

* Corresponding author
E-mail address: etikyaa@fudutsinma.edu.ng (E. V. Tikyaa)

to an appreciable degree of accuracy will go a long way in solving this problem and enable Nigerians plan their future activities such as agriculture, transportation, renewable energy generation etc. better.

Several hybrid models have been developed using neural networks to predict and forecast weather variables in the past. Masqood *et al.* [3] developed an ensemble of neural networks (MLP, ERN, RBF and HF) for weather forecasting in Southern Saskatchewan, Canada. In their proposed approach, weights are determined dynamically from the respective certainties of the network outputs. The more certain a network seems to be of its decision, the higher the weight. The data of temperature, wind speed and relative humidity were used to train and test the different models. Empirical results indicate that HFM was relatively less accurate and RBFN was relatively more reliable for the weather forecasting while the ensemble of neural networks produced the most accurate forecasts. Diaz-Robles *et al.* [4] created a novel hybrid model combining ARIMAX and ANN to improve forecast accuracy for an area with limited air quality in Temuco, Chile, where residential wood burning is a major pollution source during cold winters, using surface meteorological data and PM_{10} (particulate matter with aerodynamic diameter 10 μm) as input measurements. Results obtained indicated that the hybrid model was able to capture 100% and 80% of alert and pre-emergency episodes respectively, and their approach illustrated the potential of hybrid neural network modeling in air quality forecasting in other cities and countries but did not consider the chaos dynamics of the systems modelled hence it failed to properly optimize the model parameters. Unsihuay-Vila *et al.* [5] proposed a new hybrid approach based on nonlinear chaotic dynamics and evolutionary strategy to forecast electricity loads and prices. A hybrid approach combined nonlinear chaotic dynamics and evolutionary strategy techniques such as ARIMA, ANN to achieve short-term load and day-ahead electricity price forecasting in New England, Alberta, Spain. Their ideology was to develop a new training or identification module to a nonlinear chaotic dynamic based predictor-PREDICT2-ES, in such a way that the time series modeling and forecasting are greatly improved. Their results show that the proposed PREDICT2-ES is capable of effectively capturing the complex dynamics of chaotic time series, since in real time series, this dynamic complex is unknown and due to its running on-line manner, the search for the optimal parameters and prediction is executed automatically thus providing a more accurate and effective forecasting than ARIMA and ANN methods.

Caiado [6] examined the daily water demand forecasting performance of combined double seasonal univariate time series models (Holt-Winters, ARIMA and GARCH) based on multi-step ahead forecast mean squared errors. The empirical results obtained from the optimal combination of forecasts can be quite useful especially for short-term forecasting. However, the forecasting performance of this approach is not consistent over the seven days of the week. Di Piazza *et al.* [7] in their work, applied Artificial Neural Networks (ANNs) to the field of wind power generation. 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. Cadenas *et al.* [8] compared two multistep ahead wind speed forecasting models. Firstly, a univariate model was developed using ARIMA and secondly, a multivariate model developed using a nonlinear autoregressive exogenous artificial neural network (NARX) was developed. The NARX model used the variables: barometric pressure, air temperature, wind direction, solar radiation, relative humidity, as well as delayed wind speed. It was observed that the NARX model produced more accurate results with adequate improvements on the ARIMA model of between 5.5% and 10.6% for the hourly database and of between 2.3% and 12.8% for the ten minute database in terms of the mean absolute error and mean squared errors respectively.

Hence, this research is aimed at developing a hybrid model for carrying out multistep ahead forecasting of meteorological variables using autoregressive integrated moving average (ARIMA) and artificial neural network (ANN) optimized by chaos theory.

2. Theoretical Concepts

2.1. Nonlinear Autoregressive Neural Network with Exogenous Inputs (NARX)

Artificial neural network (ANN) is a type of artificial intelligence technique that mimics the behavior of the human brain and resembles the brain in two respects: ANN models can recognize trends, patterns, and learn from their interactions with the environment [9]. The nonlinear autoregressive network with exogenous inputs (NARX) is a recurrent dynamic network, which has feedback connections covering several layers of the network. The model equation for the NARX model is [10]:

$$y(n+1) = f\left[y(n), y(n-1), \dots, y\left(n-d_y+1\right); u(n), u(n-1), \dots, u(n-d_u+1)\right] \quad (1)$$

Where $u(n)$ and $y(n)$ are real valued and denote respectively, the input and output regressors of the model at time step n, while $d_u \geq 1$ and $d_y \geq 1, d_u \leq d_u$, are the input and output delays, respectively. Equation (47) can be simplified as:

$$y(n + 1) = f\,[y(n);\, u(n)] \quad (2)$$

The nonlinear mapping $f[\cdot]$ is generally unknown and can be approximated, using a feed forward multilayer perceptron (MLP) network, with the next value of the dependent output signal $y(n+1)$ is regressed from the preceding values of the output signal and preceding values of an independent (exogenous) input signal [11]. NARX network can be used to predict the next value of the input signal, as a nonlinear filter, in which the target output is a noise-free version of the input signal and in the modeling of chaotic systems like weather parameters [12]. However, in terms of forecasting, NARX can only produce a one step ahead forecast, hence the desire to obtain multistep ahead forecast of time series data remains a great challenge [13].

2.2. Autoregressive Integrated Moving Average (ARIMA) Models

ARIMA is a data fitting tool developed as a systematic way of predicting and forecasting business and economic time series [14]. The acronym AR-I-MA encompasses: lags of the stationarized series called "autoregressive" *(AR: p)* terms, a series which needs to be differenced to be made stationary, i.e. an "integrated" *(I: D)* series, and lags of the forecast errors called "moving average" *(MA: q)* terms [15]. In most cases, forecasting models for time series requires that the data be 'stationarized' using transformations such as differencing, logging, and deflating techniques. A time series is stationary if all of its statistical properties—mean, variance, autocorrelations, etc. are constant in time. Thus, it has no trend, no heteroscedasticity (sub-populations with different variabilities from others) and a constant degree of wiggliness i.e. small variations [15]. The ARIMA (p,d,q) is expressed mathematically in a linear equation as [4]:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \mu + \sum_{j=1}^{q} \theta_j e_{t-i} + \varepsilon_t \quad (3)$$

where ϕ_i is the i-th autoregressive parameter, θ_j is the j-th moving average parameter, μ is the mean of the series and ϵ_t is the error term at time t. Equation (3) is usually generalized in the form:

$$Y_t = \theta_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \mu - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} + \varepsilon_t \quad (4)$$

Where p is the number of autoregressive terms, Y_t is the forecasted output, Y_{t-p} is the observation at time t_p, and $\phi_1, \phi_2, \dots, \phi_p$ is a finite set of parameters determined by linear regression. Also q is the number of the moving average terms, $\theta_1, \theta_2, \dots, \theta_q$ are the finite weights or parameters set and θ_0 is the intercept or 'constant term' while ε_t is the error associated with the regression (residuals or white noise process).

In this work a multiplicative seasonal ARIMA model (SARIMA) was adopted. This method predicts both the trend and seasonality of the input data. The SARIMA model is denoted by: $SARIMA(p,d,q) \times (P,D,Q)_S$ and can be expressed as [16]:

$$\phi_p(B)\Phi_p(B^S)\nabla^d \nabla_S^D Y_t = \theta_q(B)\Theta_Q(B^S)\varepsilon_t \quad (5)$$

Where;

$$\text{AR: } \phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad (6)$$

$$\text{SAR: } \Phi_P(B^S) = 1 - \Phi_{1,S} B^S - \Phi_{2,S} B^{2S} - \cdots - \Phi_{p,S} B^{PS} \quad (7)$$

Difference operator:

$$\nabla^d = (1 - B)^d \quad (8)$$

Seasonal difference operator:

$$\nabla_S^D = (1 - B^S)^D \quad (9)$$

MA:

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad (10)$$

SMA:

$$\Theta(B^S) = 1 - \Theta_{1,S} B^S - \Theta_{2,S} B^{2S} - \cdots - \Theta_{Q,S} B^{QS} \quad (11)$$

The parameter S is the seasonality of the time series. It was evaluated by taking the inverse of the dominant (peak) frequency of the fast Fourier transform (power spectrum) of the time series [17].

3. Methodology

3.1. The Study Area and Data Source

Makurdi is located in Benue State in the north-central region of Nigeria with geographical coordinates $7°43'50''N\ 8°32'10''E$. It is blessed with a beautiful humid Guinea Savannah vegetation type with abundant sunshine which if properly harnessed could help alleviate the power needs of the people in the state [18]. Hence there is the need for an appropriate planning and estimation of the potentials of this great free renewable energy resource so as to create awareness on the possibility of harnessing it to help solve Nigeria's power challenges. On this basis the application of a suitable neural network model to the field of solar power generation is presented. The secondary weather data for Makurdi covering sixteen years (2000-2015) was collected from the Nigerian Meteorological Agency (NIMET), Abuja. The data set includes: rainfall (mm), solar radiation ($MJm^{-2}day^{-1}$), minimum temperature (°C) and maximum temperature (°C). 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.

3.2. NARX Model Design

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

i. Pre-processing of data: This involves sorting of missing data and smoothening using cubic spline interpolation [19] and designation of input/target variables using correlation analysis.

ii. Building the network: In building the network, some key parameters required to build a suitable network were specified. 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 created using *'narxnet'* which uses the tan-sigmoid transfer function as default and linear transfer function in the output layer. The *narxnet* created has two inputs: 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 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 [20]. The hyperbolic tangent sigmoid (tan-sig) was used as the transfer function to carry data across the hidden layers via the neurons. In this work two hidden layers were used in addition to the input and output layers. This is in accordance with Takens' embedding theorem [21]. 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 ($N_{H,1} and\ N_{H,2}$) by using the expression [11]:

$$N_{H,1} = 2m + 1, N_{H,2} = round(\sqrt{N_{H,1}}) \quad (12)$$

m is the embedding dimension of the target data obtained using the method of false nearest neighbors, FNN [22]. Similarly, the Levenberg-Marquardt and Bayesian regularization training functions were both separately deployed to optimize the network training.

iii. Preparing the data for training: in training the network with tapped delay lines, it is crucial to fill delays with initial values of the inputs and outputs of the network. In MATLAB, the *'preparets'* function was used to facilitate this process. *Preparets* has three input arguments: the network, input series and target series, and returns the initial conditions needed to fill the tapped delay lines in the network and the modified input and target series. The function was implemented in MATLAB as follows:

[inputs, inputStates, layerStates, targets] = preparets (net, inputSeries,{}, targetseries) (13)

iv. Training the network: during the training process, the weights were adjusted systematically until the predicted output generated is close to the target (measured) output of the network. The comparison of the output signal with the desired response or target output consequently produced an error signal. In each step of iterative process, the error signal activates a control mechanism which applies a sequence of corrective adjustments of the weights and biases of the neuron via the weight/bias learning functions with the *learngdm* function applied. These corrective adjustments continued until the training data attained the desired mapping to obtain the target output as closely as possible. After a series of iterations (also called *training epochs*) the neural network was successfully trained and the weights were saved.

v. Testing and validation of the model performance: Performance evaluation functions 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 [7, 23]:

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	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_a(i) - y_p(i)\right)^2}$
MAE	Mean absolute error	$\frac{1}{n}\sum_{i=1}^{n}\left\|y_a(i) - y_p(i)\right\|$
MAPE	Mean absolute percentage error	$\frac{1}{n}\sum_{i=1}^{n}\left\|\frac{y_a(i)-y_p(i)}{y_a(i)}\right\| \times 100\%$

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

The root mean square error was employed in this research as it is the most preferred performance function in neural network training because it evaluates the performance of the network according to the mean squared deviations between the target and the predicted output. Other measures of accuracy used to evaluate the performance of the hybrid SARIMA-NARX model in this work are listed in Table 2 [4, 24]:

Table 2. Additional Evaluation Parameters.

Parameter	Mathematical formulation	Description
Pearson's correlation coefficient, ***R***	$\frac{n\sum xy-(\sum x)(\sum y)}{\sqrt{n(\sum x^2)-(\sum x)^2}\sqrt{n(\sum y^2)-(\sum y)^2}}$	Measures relationship between two data sets.
Schwarz-Bayesian information criterion, ***SBIC***	$n\log(SSE)+m\log n$ SSE =sum squared errors, n =number of data points, m = number of parameters in the model	Measures the goodness-of-fit of the model and also penalizes the number of model parameters.
Ljung-Box (Chi- squared) Q-test, Q	$T(T+2)\sum_{k=1}^{L}\frac{\rho(k)^2}{(T-k)}$ T=sample size, L=no. of autocorrelation lags, $\rho(k)$=sample autocorrelation at lag k	Assesses the null hypothesis that a series of residuals exhibits no autocorrelation for a fixed number of lags L, against the alternative that some autocorrelation coefficient ρ (k), k = 1, 2,..., L, is nonzero.

In order to improve the accuracy of training and avoid over fitting, multiple networks were trained to fit the given data set. This method lead to the constitution of an ensemble-based NARX framework which involves generating multiple possible realizations of models which fit the training data by training different configurations in the neighborhood of a given configuration of hidden layer nodes a given number of times and using a weighted averaging method to combine all component models in the ensemble in order to conduct an out-of-sample multi-step-ahead forecasting of the output. The weighted average output was obtained from the equation [6, 25]:

$$y(t)=\frac{1}{N}\sum_{i=1}^{N}w_i y_i(\mathrm{t}) \tag{14}$$

Where N is the number of NARX networks trained, w_i is the error weight (MSE or SSE) of each model in the ensemble and $y_i(t)$ is the output prediction of each model trained in the ensemble.

3.3. SARIMA Models Design

The Box-Jenkins ARIMA method basically involves a three-step algorithm which involves [5]:

i. Model identification: ARIMA model identification methods used in this work involves determining suitable values for parameters *p* and *q* and determining the degree of differencing, *d*, to obtain stationarity. In this work, the unit root test and the Schwarz-Bayesian information criterion (SBIC) method was used to determine the optimal values of the model parameters; *p, d, q*. The model with the least SBIC value was adopted and used. After this a $SARIMA(p,d,q)\times(P,D,Q)_S$ model is defined with the seasonality set at 365 days for daily data.

ii. Parameter estimation: After identifying the most suitable model to adopt, its parameters were estimated using the maximum likelihood and conditional least squares methods.

iii. Diagnostic checking and forecasting: After identifying the model and estimating its parameters, the forecast was carried out and diagnostic checks were used to reveal its inadequacies and suitable improvements are indicated. Residuals and their autocorrelations were also inspected. If the model is a good fit to the data, then the residuals would correspond to white noise, have very little autocorrelation and will all fall within the 95% confidence interval.

3.4. The Hybrid SARIMA-NARX Model

The hybrid model in this study is based on the hybrid model designed by Diaz-Robles *et al.* [4] to capture different patterns in the air quality data and Caiado [6] to forecast water demand. The methodology of their model consisted of an ARIMA model to forecast the Max PM_{10} ma and an ANN model was developed to describe the residuals from the ARIMA model. But in this work, a multiplicative seasonal ARIMA model was used in capturing the linear pattern (trend) and seasonality of the forecasted solar radiation time series while a NARX model was designed to captures the non-linear patterns of the time series such as noise and extreme values with the network training parameters optimized using chaos theory to save time and computational costs. That is to say in this hybrid model, the output of the ARIMA forecast is inputted into the NARX network created from the historical data with exogenous input so as to generate the nonlinear part of the final output and obtain a more accurate output and longer forecast than that obtained by previous designers. A flow chart showing details of the Hybrid model is illustrated in Figure 1:

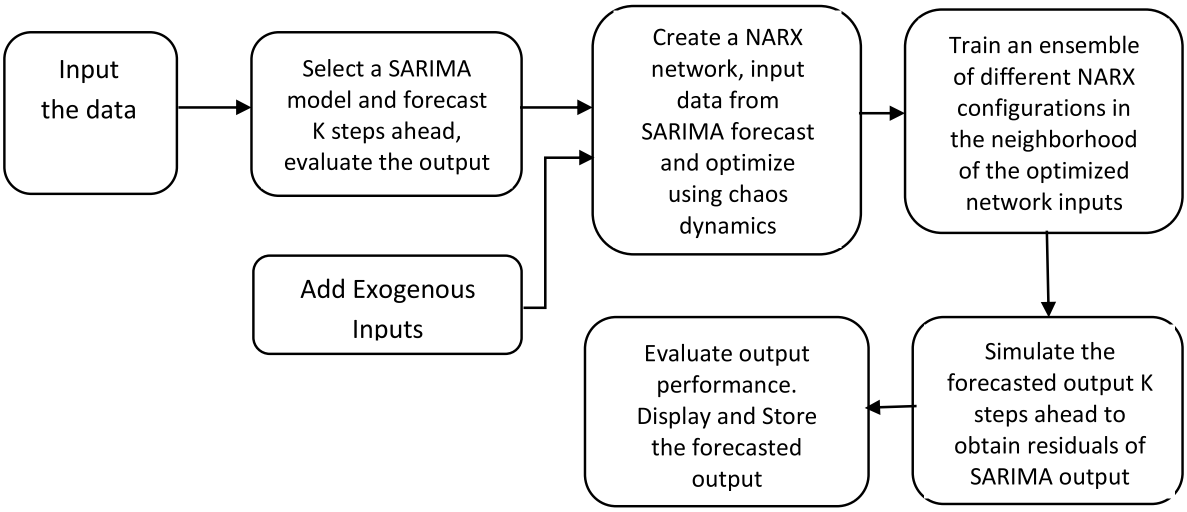


Figure 1. Flow chart of the SARIMA-NARX-Chaos model.

4. Results and Discussion

4.1. NARX Model Selection

Table 3. Correlation analysis of the meteorological variables for Makurdi from 2000 to 2015.

Parameters	Solar radiation	Rainfall	Minimum temperature	Maximum temperature
Solar radiation	1.0000	-0.2628	-0.5743	0.2942
Rainfall	-0.2628	1.0000	0.1154	-0.1108
Minimum temperature	-0.5743	0.1154	1.0000	0.5919
Maximum temperature	0.2942	-0.1108	0.5919	1.0000

Based on the correlation analysis of the different variables displayed in Table 3, the model in Figure 2 was developed for the prediction and forecasting of solar radiation in Makurdi using the hybrid model with minimum and maximum temperature selected as exogenous input parameters.

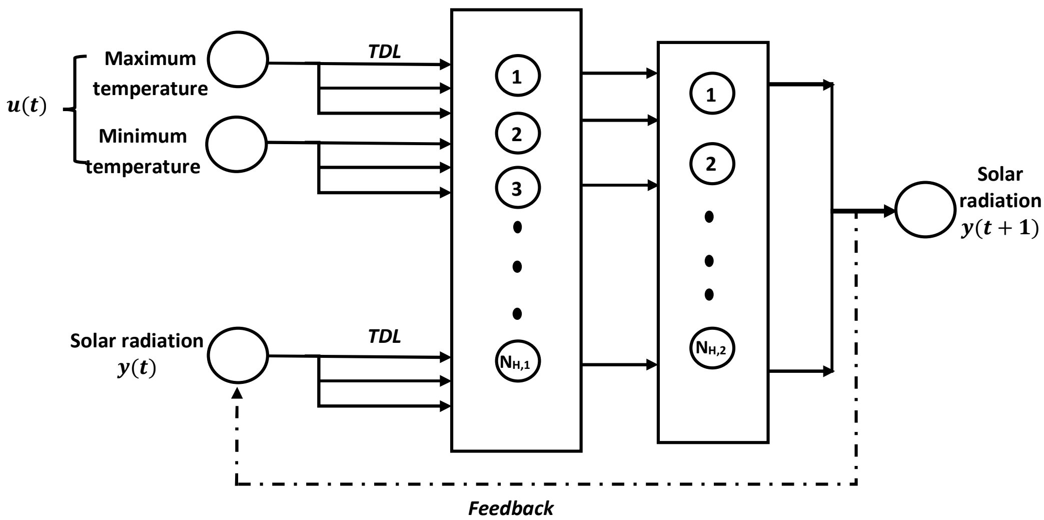
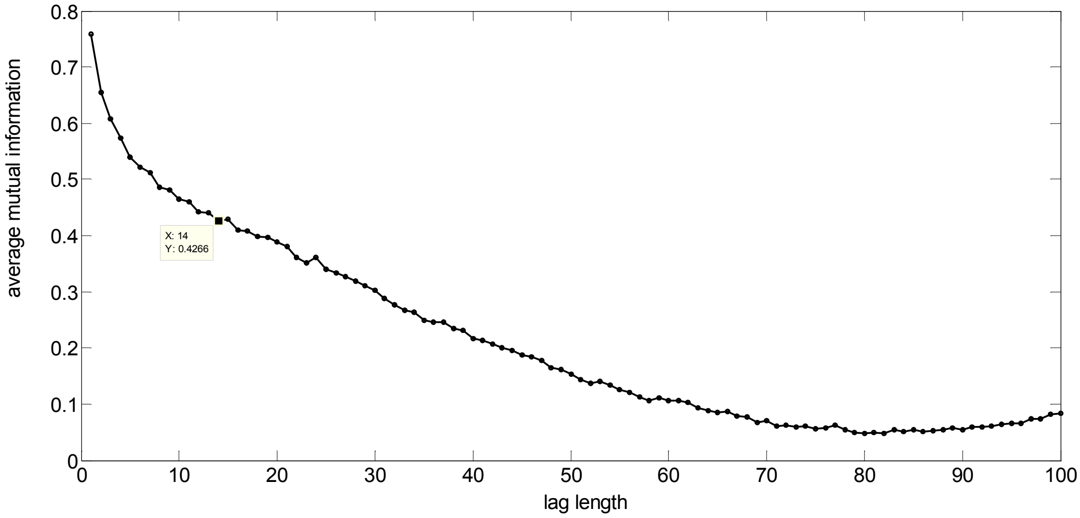
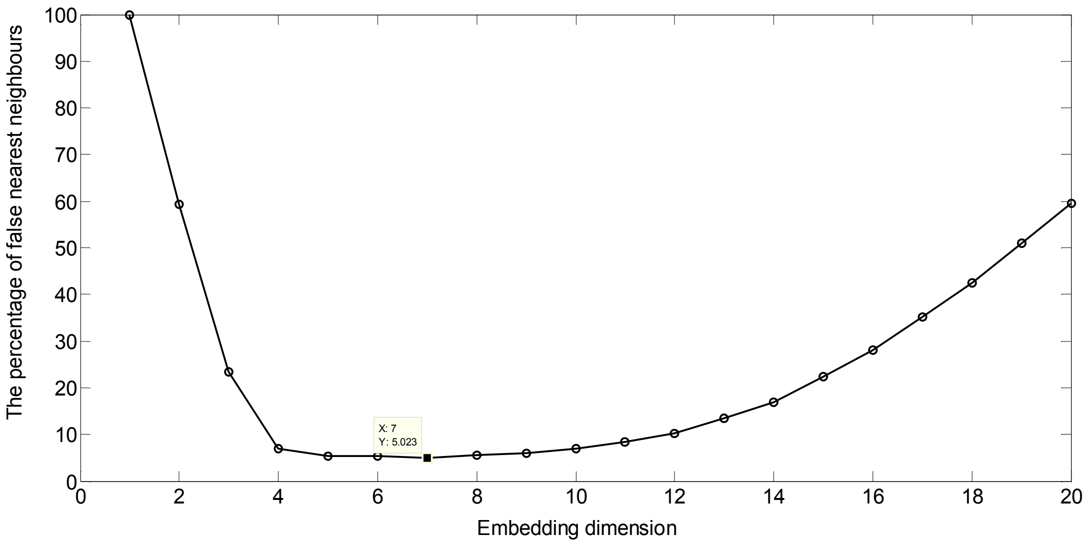


Figure 2. NARX model (series-parallel architecture) design for solar radiation prediction.

In order to optimize the number of time delay lines, hidden neurons and hidden layer sizes, the tenets of chaos theory was utilized. Under this technique, the takens embedding theory was applied as explained in equation (12). The time delay and embedding dimension for the daily solar radiation time series training data from 2000 to 2015 were estimated using the methods of AMI and FNN and the result is displayed in Figure 3.



(a)



(b)

Figure 3. Estimation of NARX network Parameters for daily solar radiation prediction; (a.) time delay, τ = 14 days, (b.) embedding dimension, m = 7. This result indicates the number of neurons in the first and second hidden layers are in the neighborhood of 15 and 4 respectively.

4.2. SARIMA Model Selection

For the different input and target parameters, the SBIC method was used to evaluate and optimize the seasonal (*AR: p* and *MA: q*) and non-seasonal model parameters (*SAR:P* and *SMA:Q*). In this case, the model fit statistic considers the goodness-of-fit and parsimony; i.e. the simplest model with the least assumptions/variables, minimum SBIC and greatest loglikelihood selected. The degree of differencing, *d* was determined using the unit root test. The seasonality of the model is set at S=365 for daily data since there are 365 days in a year. Table 4 shows a summary result of the ARIMA (p, d, q)×(P, D, Q)$_S$ models selected while Table 5 shows the SARIMA parameters estimated with their performance in forecasting of the different meteorological variables using fifteen years daily average data from 2000-2014. Figure 4 shows the result of the residual analysis for daily solar radiation in Makurdi in the year 2015 using the selected SARIMA models.

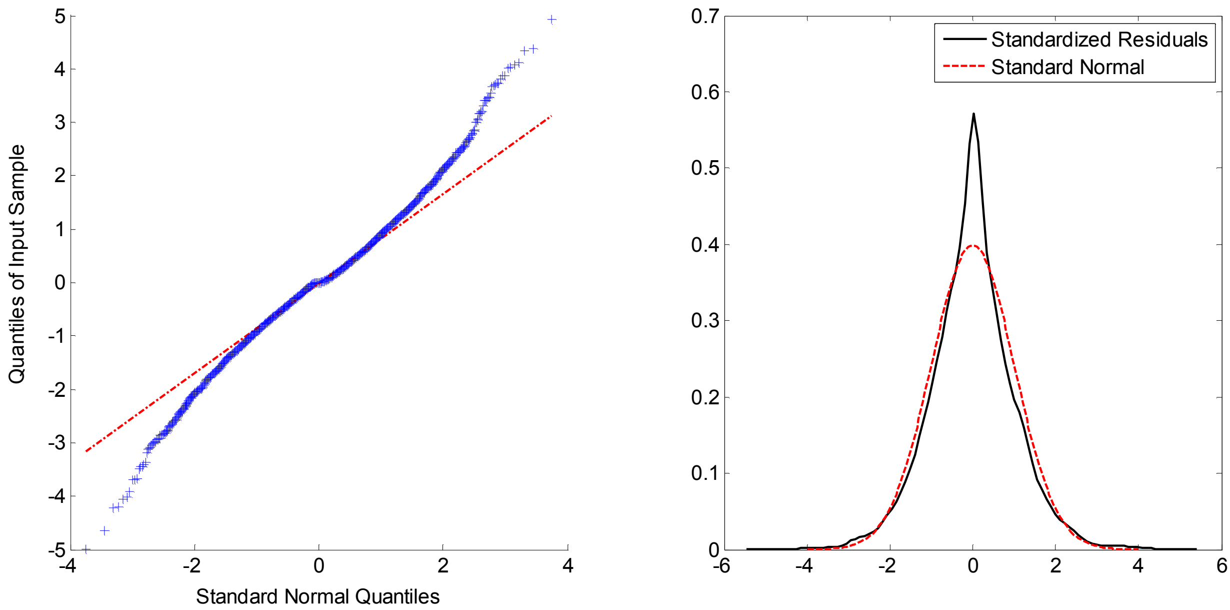
Table 4. ARIMA Model Parameters for 15 Year Data of the Meteorological Variables and their Performance.

Parameter	AR, SAR term (p, P)	Degree of differencing (d, D)	MA, SMA term (q, Q)	Seasonal ARIMA model	Constant	*MSE*	*RMSE*	*R*
Solar radiation	2	0	2	ARIMA (2,0,2)x(2,0,2)$_{365}$	0	6.3010	2.5102	0.7266
Minimum temperature	2	0	2	ARIMA (2,0,2)x(2,0,2)$_{365}$	0	7.3042	2.7026	0.8091
Maximum temperature	2	0	2	ARIMA (2,0,2)x(2,0,2)$_{365}$	0	10.1882	3.1919	0.6849

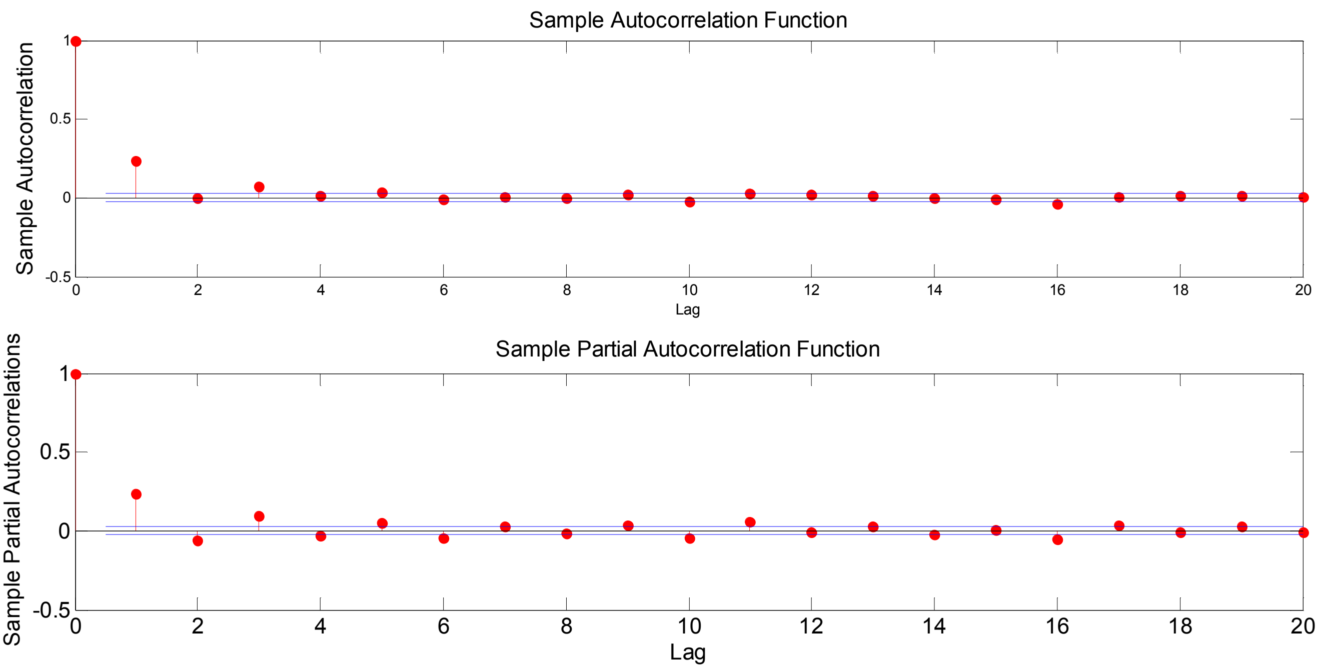
Table 5. ARIMA $(2,0,2)x(2,0,2)_{365}$ Model Parameters Seasonally Integrated with Seasonal AR(2) and MA(2); Conditional Probability Distribution: Gaussian.

Parameter	Value	Standard error	t-Statistic
Constant	0	Fixed	Fixed
AR{2}	0.969531	0.0107061	90.5587
SAR{2}	0.481775	0.0549366	8.76965
MA{2}	-0.938522	0.017753	-52.8655
SMA{2}	-0.251626	0.0538433	-4.67331
Variance	7.79296	0.113311	68.7753

$MSE = 6.30102$
$RMSE = 2.51018$
$R = 0.72657$
$D = 0.52791$



(a)



(b)

Figure 4. (a.) Quartile-Quartile plot of sample data versus standard normal for daily solar radiation SARIMA forecast, (b.) Error autocorrelation and partial autocorrelation of SARIMA forecasted daily solar radiation for 2015.

The Ljung-Box Q-test result in Figure 4 (b) shows the sample ACF and PACF results; with the null hypothesis that all autocorrelations are jointly equal to zero up to the tested lag is rejected (h = 1) for the three lags as the residuals are uncorrelated and cut off from the 5% level of significance.

4.3. Hybrid ARIMA-NARX Model Results

The results of the performance of the SARIMA-NARX hybrid model training for daily solar radiation forecast for Makurdi in 2015 using the Levenberg-Marquardt and Bayesian regularization training functions and evaluated in terms of the root-mean-square error and correlation coefficient of the target and output are displayed in Tables 6 and 7. Similarly, Figure 5 and 8 show the ARIMA output, NARX network predictions and observed data plotted on the same axis while Figure 6 and 8 show the error autocorrelations and the output regressions of the best performing configurations of the different hybrid models using the two different training functions respectively.

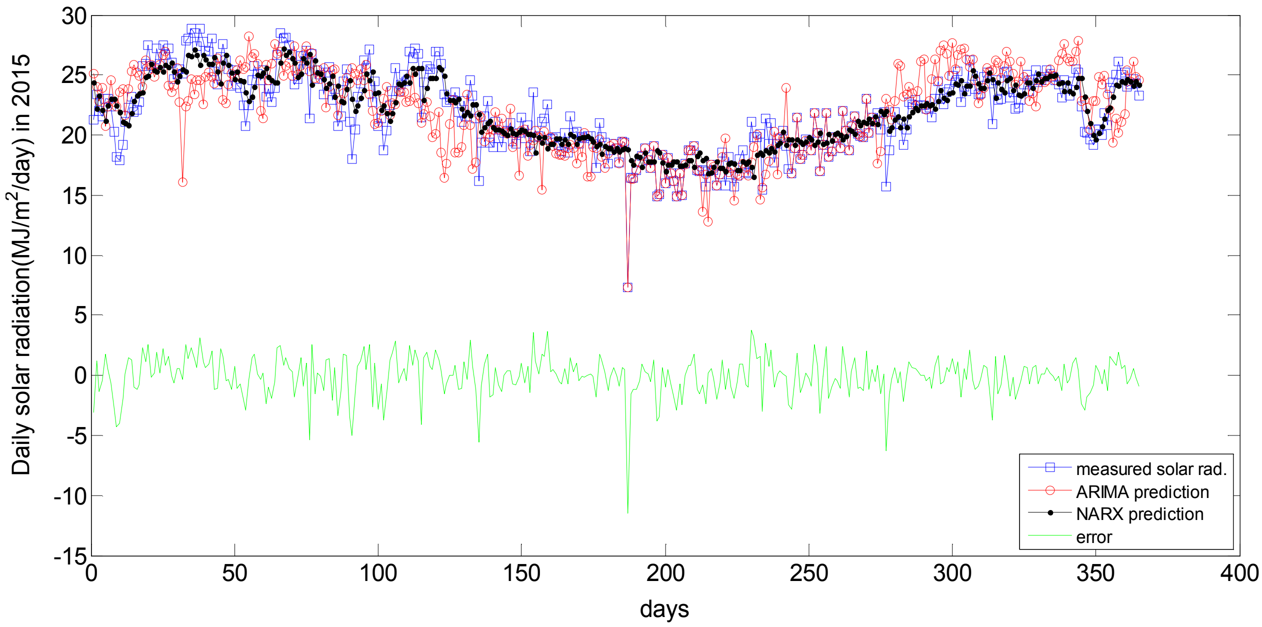
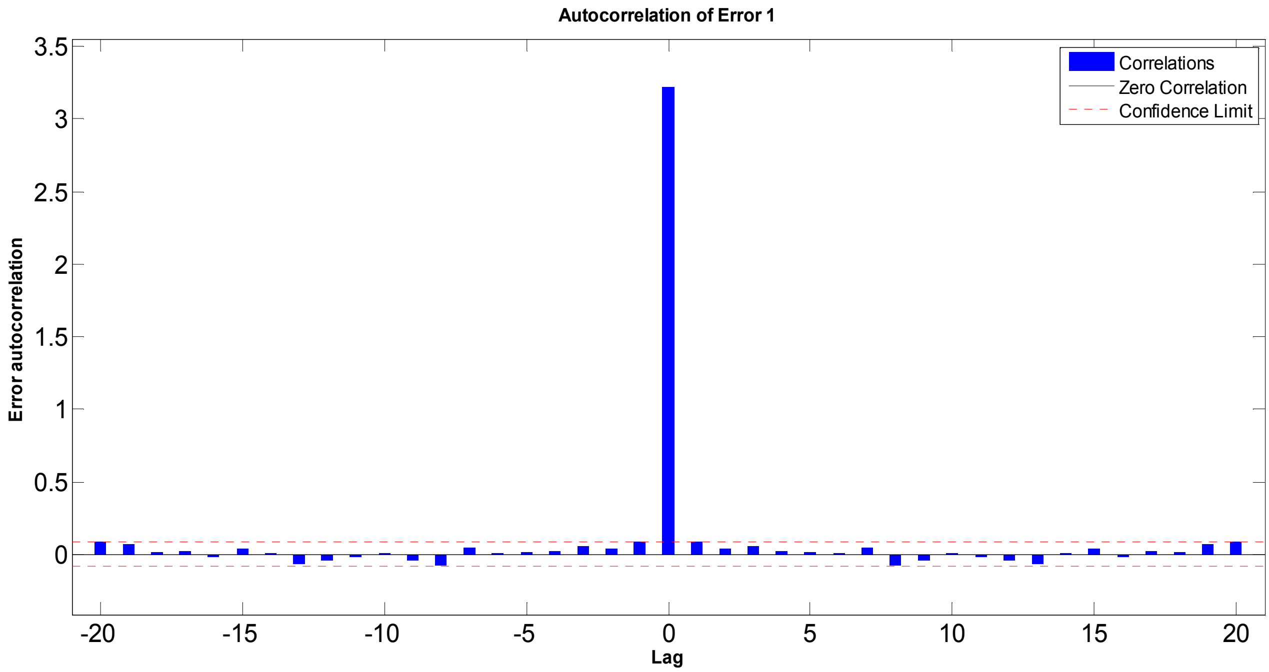
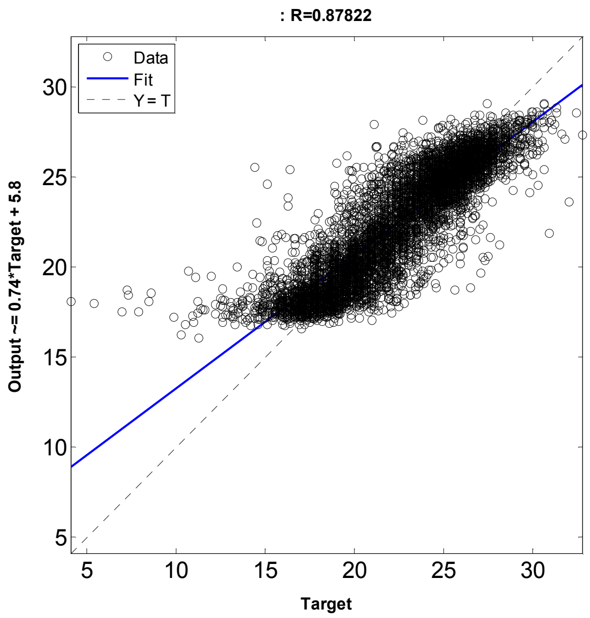


Figure 5. Forecast plot of daily solar radiation in Makurdi for 2015 using Levenberg-Marquardt training function with an ensemble of 10 NARX networks in the neighborhood of the optimized 15-4-1 neuron configuration and 13 TDL.



(a)



(b)

Figure 6. Performance evaluation of the daily solar radiation forecast for Makurdi in 2015 using Levenberg-Marquardt training function; (a.) Autocorrelation of errors (residuals), (b.) Correlation and regression of target and forecast series.

Table 6. Hybrid Arima-Narx Model Selection for Daily Solar Radiation Prediction for Makurdi using the Levenberg-Marquardt ('*Trainlm*') Training Function.

m	$N_{H,1}$	$N_{H,2}$	**No. of time delay lines (TDL)**					
			13		**14**		**15**	
			RMSE	***R***	***RMSE***	***R***	**RMSE**	***R***
6	13	4	1.6575	0.8797	1.6554	0.8761	1.6714	0.8773
7	15	4	1.6475	0.8782	1.6641	0.8787	1.6541	0.8784
8	17	4	1.6583	0.8791	1.6656	0.8807	1.6784	0.8797

Table 7. Hybrid SARIMA-NARX model selection for daily solar radiation prediction for Makurdi using the Bayesian regularization ('*trainbr*') training function.

m	$N_{H,1}$	$N_{H,2}$	**No. of time delay lines (TDL)**					
			13		**14**		**15**	
			RMSE	***R***	***RMSE***	***R***	**RMSE**	***R***
6	13	4	1.6651	0.8690	1.6628	0.8692	1.6563	0.8645
7	15	4	1.6676	0.8686	1.6644	0.8682	1.6598	0.8680
8	17	4	1.6602	0.8676	1.6618	0.8695	1.6668	0.8683

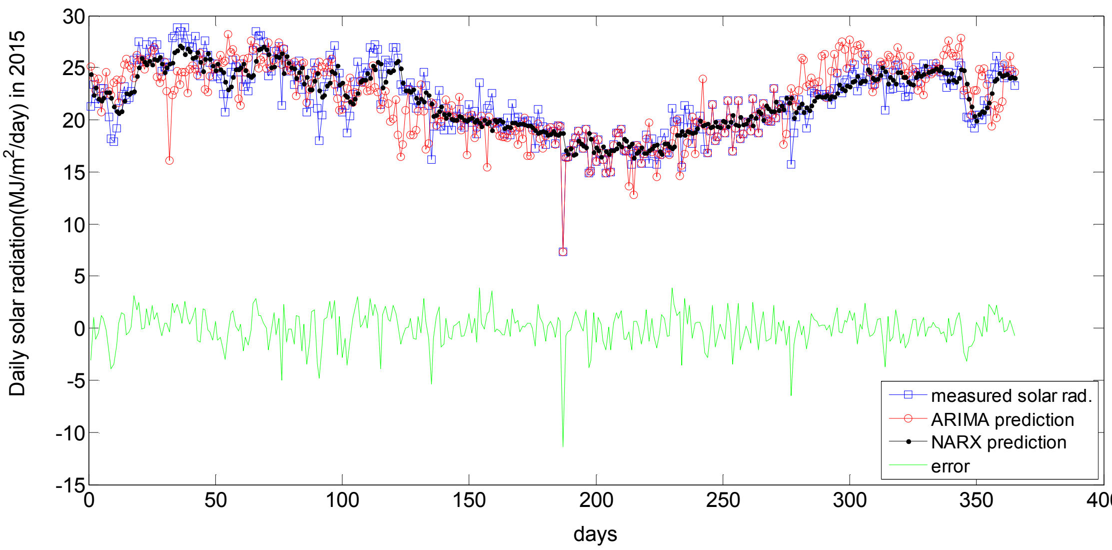
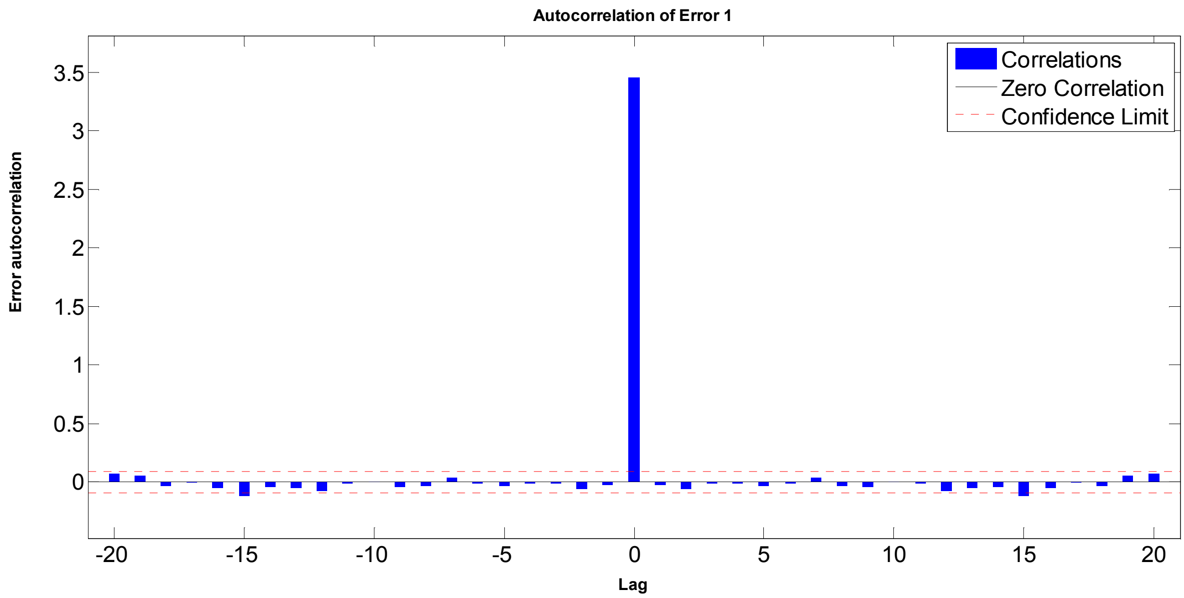
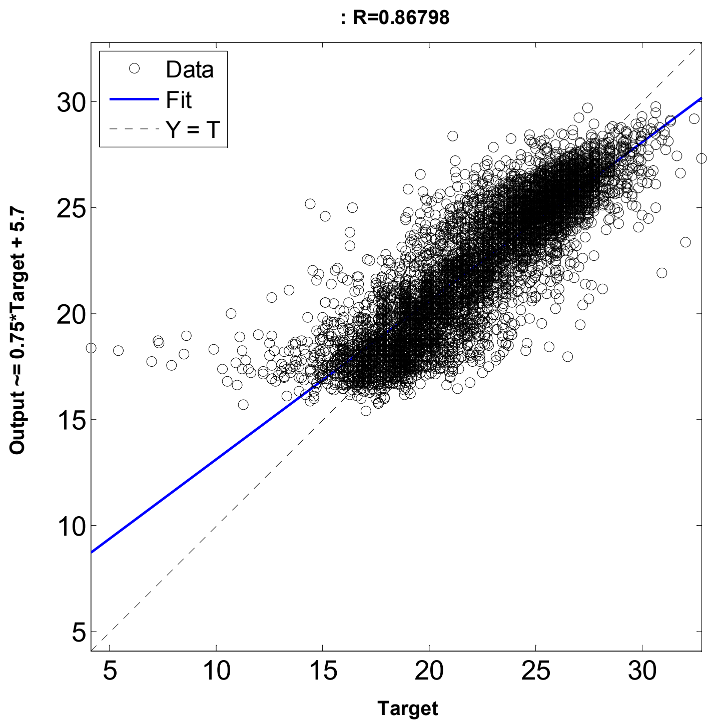


Figure 7. Forecast Plot of Daily Solar Radiation in Makurdi for 2015 using Bayesian Regularization Training Function with 15-4-1 Neuron Configuration and 15 TDL.



(a)



(b)

Figure 8. Evaluation of daily solar radiation forecast for Makurdi in 2015 using Bayesian regularization training function; (a.) autocorrelation of error, (b) correlation and regression of target and forecast series.

4.4. Summary of Forecast Results

A summary of the results of the performance of the SARIMA-NARX-Chaos model for daily solar radiation forecast in Makurdi for 2015 is displayed in Table 8.

Table 8. Summary of SARIMA-NARX forecast results.

Method (sampling rate) training function	**Evaluation parameters**				
	MSE	***RMSE***	***R***	***MAE***	***MAPE (%)***
ARIMA (daily)	6.3010	2.5102	0.7266	1.7921	8.0024
NARX (daily) $\boldsymbol{lm}$	2.7143	1.6475	0.8782	1.2042	5.9695
NARX (daily) $\boldsymbol{br}$	2.7549	1.6598	0.8680	1.2346	6.0470

Thus it is clear to see that the NARX hybridization of SARIMA forecast greatly improves the output by decreasing the $MAPE$ by about 2%. This is as a result of the non-linear nature of the multivariate NARX model which incorporates chaos dynamics in its architecture. It is also worthy to note that the Levenberg-Marquardt (lm) training function outperforms the Bayesian-regularization (br) training function in this model in terms of processing speed and accuracy of results. Hence, it is highly recommended for modeling tasks of this sort.

4.5. Daily Solar Radiation Forecasts for 2016 and 2017

The hybrid SARIMA-NARX-Chaos model was used to forecast the daily solar radiation for 2016 by inputting sixteen years training data (2000-2015) into the model and simulating the solar radiation values 365 days ahead to obtain the values for 2016. The forecasted 2016 data was added to the input data and then stored. The process was again repeated using training data from 2000-2016 and the solar radiation values for 2017 obtained. Figure 9 is a plot of the input data (2000-2015) and the forecasted values (2016 and 2017).

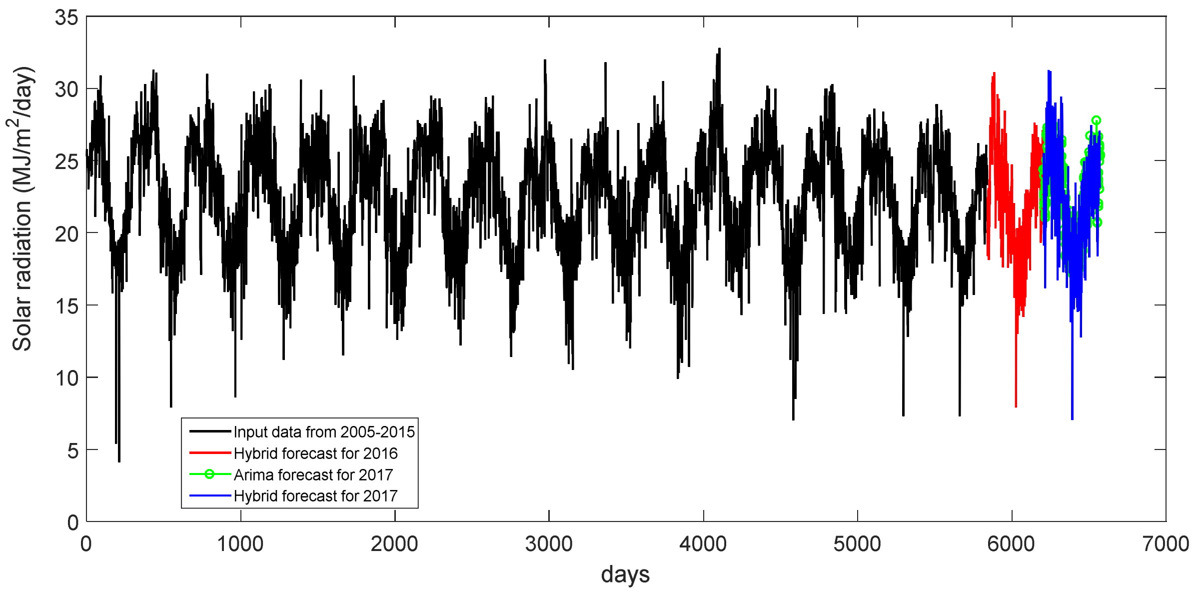


Figure 9. Daily solar radiation forecasts for 2016 and 2017 for Makurdi.

5. Conclusion

A hybrid SARIMA-NARX-Chaos neural network model was successfully developed, trained and tested for the prediction of daily solar radiation in Makurdi. The intrinsic parameters of these models were optimized using the predetermined chaos dynamics of the meteorological data. Results of the model validation shows that, the models perform better using the Levenberg-Marquardt training function predicting solar radiation successfully using minimum temperature and maximum temperature as exogenous variables within a $RMSE$ of 1.6475, correlation coefficient of 0.8782, MAE of 1.2042 and $MAPE$ of 5.9695% for the a 365 day ahead forecast in 2015. Even though the trends were accurately forecasted, the relatively fair but acceptable $RMSE$ values obtained are as a result of the poor correlation of the meteorological variables as shown in Table 4 which could be as a result of large number of missing data and white noise in the meteorological data which may have arisen from faulty equipment or human error/neglect in measurement.

Acknowledgements

We wish to acknowledge the authorities of University of Agriculture, Makurdi and Federal University Dutsin-ma, Nigeria for the support accorded us in the course of this work.

References

[1] Lorenz, E. N. (1963). Deterministic Non-periodic Flow. *Journal of the Atmospheric Sciences*, 20 (2): 130-141.

[2] IPCC (2017). Working Group I: The Physical Science Basis. Fourth Assessment. Report of the Intergovernmental Panel on Climate Change (AR4): Summary of Policymakers, Geneva.

[3] Masqood, I., Khan, R. M. and Abraham, A. (2004). An Ensemble of Neural Networks for Weather Forecasting. *Neural Computing and Application*, 13: 112-122. DOI: 10.1007/s00521-004-0413-4.

[4] Diaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G. and Moncada-Herrera, J. A. (2008). A Hybrid ARIMA and Artificial Neural Networks Model to Forecast Particulate Matter in Urban Areas: The case of Temuco, Chile. *Atmosphere environment*, 42: 8337.

[5] Unsihuay-Vila, C., Zambroni de Souza, A. C., Marangon-Lima, J. W. and Balestrassi, P. P. (2009). Electricity Demand and Spot Price Forecasting using Evolutionary Computation Combined with Chaos Nonlinear Dynamics Model. *Elsevier: Electrical Power and Energy Systems,* 32: 112.

[6] Caiado, J. (2010). Performance of combined double seasonal univariate time series models for forecasting water demand. *Journal of Hydrologic Engineering,* 15 (3): 8.

[7] 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," Proceedings of the International Conference on Renewable Energies and Power Quality, ICREPQ '14, Vol. 12, 8th-12th April, 2014.

[8] Cadenas, E., Rivera, W., Campos-Amezcua, R. and Heard, C. (2016). Wind Speed Prediction Using a Univariate ARIMA Model and a Multivariate NARX Model. *Energies,* 2016 (9), 109.

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

[10] Beale, M. H., Hagan, M. T. and Demuth, B. H. (2014). Neural Network ToolboxTM User's Guide (R2014b), MathWorks Inc., Pp. 23-34.

[11] Menezes-Jnr, J. M. and Barreto, G. A. (2007). Long-Term Time Series Prediction with the NARX Network: An Empirical Evaluation. Preprint submitted to Elsevier Science, Department of Teleinformatics Engineering, Fortaleza, Cear´a, Brazil, P. 15.

[12] Diaconescu, E. (2008). The Use of NARX Neural Networks to Predict Chaotic Time Series. *WSEAS Transactions on Computer Research*, 3 (3): 189.

[13] Menezes-Jnr, J. M. and Barreto, G. M. (2007). Multistep-Ahead Prediction of Rainfall Precipitation Using the NARX Network. Department of Teleinformatics Engineering, Fortaleza, Cear´a, Brazil, P. 12.

[14] Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). Time Series Analysis: Forecasting and Control, 3rd edition, Englewood Cliffs, New Jersey, Pp. 34-51.

[15] Nau, R. (2014). Introduction to ARIMA models. Lecture notes on forecasting, Duke University. http://people.duke.edu/~rnau/forecasting.html

[16] Iwok, A. I. (2017). Handling Seasonal Autoregressive Integrated Moving Average Model with Correlated Residuals. *American Journal of Mathematics and Statistics*, 7 (1): 1-6. DOI: 10.5923/j.ajms.20170701.01.

[17] Telgarsky, R. (2013). Dominant Frequency Extraction. *arXiv* 1 (2013): 1-12.

[18] NOAA (2016). Nigeria: Climatic Normals from 1961-1990. National Oceanic and Atmospheric Administration. http://www.ncdc.noaa.gov/wdcmet/data-access-search-viewer-tools/nigeria-climate-normals-1961-1990.

[19] Kiusalaas, J. (2010). Numerical Methods in Engineering with MATLAB, 2nd edition. Cambridge University Press, London, 2010, Pp. 116-119.

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

[21] Takens, F. (1980). Detecting Strange Attractors in Turbulence. *Lecture Notes in Mathematics,* 898: 366.

[22] 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.

[23] 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, Uppsala University, Pp: 21-35.

[24] Ljung, G. and Box, G. E. P. (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 66: 67-72.

[25] Zhongxian, M., Eugene, Y., Fue-sang, L., Zhiling, Y., and Yongquian, L. (2014). Ensemble Nonlinear Autoregressive Artificial Neural Networks for Short-term Wind Speed and Power Forecasting, *International Scholarly Research Notices*, 6: 1-16. http://dx.doi.org/10.1155/2014/972580.