Nigerian Journal of Mathematics and Applications
Volume **25**, (2016), 205 − 216. *Printed by Unilorin press*
©*Nig. J. Math. Appl.* *http://www.njmaman.com*

MODELLING AND EVALUATION PERFORMANCES WITH NEURAL NETWORK USING CLIMATIC TIME SERIES DATA

T. O. Olatayo* and A. I. Taiwo

Abstract

Neural network model is an alternative powerful data modeling tool capable of capturing and representing non-linear relationship between variables.
The methodology used in this research article were Artificial Neural network (ANN), Autoregressive Integrated Moving Average (ARIMA) and fuzzy time series (FTS) models.
The study revealed that, since ANN forecast performance evaluation has the lowest value of Sum of square error (SSE), Mean square error (MSE) and Root mean square error (RMSE), then ANN model outperforms ARIMA and FTS models. The ANN model was found to be more efficient with minimum parameters and capable of handling the non-linearity that characterized climatic data series.

1. Introduction

Artificial Neural Networks (ANN) has recently been used as a research tools in several fields. Investigators have been attracted by ANN's freedom from restrictive assumptions such as linearity and non-stationarity that are often needed to make the traditional mathematical models tractable. Artificial Neural Networks (ANN) is a flexible structure that can be applied to a wide range of forecasting

Received: 15/05/2016, Accepted: 24/07/2016, Revised: 10/08/2016.
2015 *Mathematics Subject Classification.* 49Nxx & 00Axx. * Corresponding author.
Key words and phrases. Artificial Neural Network, Autoregressive integrated Moving Average, Forecasting evaluation, Fuzzy time series and Climate change
Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye.;
Emails: otimtoy@yahoo.com & ab.com1982@yahoo.com

problems with a high degree of accuracy. However, the neural networks need a large amount of historical data to reach the highest level of accuracy of the results and have as the main advantage the ability of modeling nonlinear systems (Khashei et al., 2008).
It can also be defined as a computer system containing many simple nonlinear units or nodes interconnected by links. In this research we wish to consider the predictive power and accuracy of Artificial Neural network models in the context of non-linearity, as ANN has the ability of modeling nonlinear systems. To assert and validate the predictive performance of ANN, we will compare ANN with an autoregressive integrated moving average model (ARIMA) that is used to model non-stationary. While we will as well compare it with another interested and powerful free distribution or non-parametric method and non-linear, that is, fuzzy time series model (FTS) which has the advantage of reducing the calculation time and simplifying the calculation processes. For application purpose, the three methods will be applied to time series climatic data as the research will be used to discuss climate change, since it's seems to be the foremost global challenge facing humans at the moment. Particular attention will be given to rainfall, since climate change with respect to rainfall has a strong in?uence on traffics, sewer systems, and other human activities.
Nevertheless, rainfall is one of the most complex and difficult elements of the hydrology cycle to understand and to model due to the complexity of the atmospheric processes that generate rainfall and the tremendous range of variation over a wide range of scales both in space and time (French et al., 1992). Guoqiang et al. (1998) discussed artificial neural networks (ANNs) for forecasting and he discovered that ANNs provide a great deal of promise and also embody much uncertainty. In his paper, he revealed ANN as a state-of-the-art method for forecasting and the future research directions. He further stated that ANN can be applied to almost all the fields, since real-life situations are mostly nonlinear in nature. Following this, Fabio (2010) proposed a neural network model for forecasting the production time series of eleven different industries in Brazil. Firstly, he studied different networks topologies that have been implemented in the literature in recent years, such as perceptron, linear networks, multi-layer perceptron (MLP), probabilistic network, Hopfield model, Kohonen model, time delay neural network (TDNN), Elman and Jordan Network, in addition to the back-propagation and Levenberg-Marquadt algorithms. He concluded that the neural network model proposed was effective for forecasting production time series in these industries.
Adewole e tal (2011), proposed an artificial neural network foreign exchange rate forecasting model (AFERFM) for foreign exchange rate forecasting. The proposed model was divided into two phases, namely: training and forecasting. In the training phase, back propagation algorithm was used to train the foreign

exchange rates and learn how to approximate input. Sigmoid Activation Function (SAF) was used to transform the input into a standard range [0, 1]. The learning weights were randomly assigned in the range [-0.1, 0.1] to obtain the output consistent with the training. SAF was depicted using a hyperbolic tangent in order to increase the learning rate and make learning efficient. Feed forward Network was used to improve the efficiency of the back propagation. Multilayer Perceptron Network was designed for forecasting.

Gianluigi (2002) used NN's model to forecast 30 time series, ranging on several fields, from economy to ecology. The statistical approach to artificial neural networks modeling developed in his paper was compared with linear modeling and to other three well-known neural network modeling procedures: Information Criterion Pruning (ICP), Cross-Validation Pruning (CVP) and Bayesian Regularization Pruning (BRP). The findings shows that the linear models outperform the artificial neural network models and albeit selecting and estimating much more parsimonious models, the statistical approach stands up well in comparison to other more sophisticated ANN models.

Nevertheless, Saeed et. al. (2000), also compared the performance of Back-Propagation Artificial Neural Network (BPN) models with the traditional econometric approaches to forecasting the inflation rate. Of the traditional econometric models, he used a structural reduced-form model, an ARIMA model, a vector autoregressive model, and a Bayesian vector autoregression model. He compared each econometric model with an hybrid BPN model which uses the same set of variables. Dynamic forecasts are compared for three different horizons: one, three and twelve months ahead. Root mean squared errors and mean absolute errors are used to compare quality of forecasts. The results show the hybrid BPN models are able to forecast as well as all the traditional econometric methods, and it outperformed them in some cases.

Hung (2009), used artificial Neural Network technique to improve rainfall forecast performance. A real world case study was set up in Bangkok; 4 years of hourly data from 75 rain gauge stations in the area were used to develop the ANN model. The developed ANN model was applied to real time rainfall forecasting and flood management in Bangkok, Thailand.

Toth et al. (2000) compared short-time rainfall prediction models for real-time ?ood forecasting. Different structures of autoregressive moving average (ARMA) models, arti?cial neural networks and nearest-neighbors approaches were applied for forecasting storm rainfall occurring in the Sieve River basin, Italy, in the period 1992-1996 with lead times varying from 1 to 6h. The ANN adaptive calibration application proved to be stable for lead times longer than 3 hours, but inadequate for reproducing low rainfall events. In addition, ANN could learn and generalize from examples to produce a meaningful solution even when the input data contain errors or is incomplete (Luk et al., 2000).

Mahsin et al. (2012) use Box-Jenkins methodology to build seasonal ARIMA model for monthly rainfall data taken for Dhaka station, Bangladesh, for the period from 1981-2010. In their paper, $ARIMA(0,0,1)^*(0,1,1)_{12}$ model was found adequate and the model was used for forecasting the monthly rainfall. Seyed et al.,(2011) use time series method to model weather parameter in Iran at Abadeh Station and recommended $ARIMA(0,0,1)^*(0,1,1)_{12}$ as the best fit for monthly rainfall data and $ARIMA(2,1,0)^*(2,1,0)_{12}$ for monthly average temperature for Abadeh station.

Al-Ansari et. al. (2006) dealt with the statistical analysis of the rainfall measurements for three meteorological stations in Jordan: Amman Airport (central Jordan), Irbid (northern Jordan) and Mafraq (eastern Jordan). Normal statistical and power spectrum analyses as well as ARIMA model were performed on the long-term annual rainfall measurements at the three stations. Further result indicated that there is decreasing trend for forecasted rainfall results in all stations.

Amha (2010) studied the monthly rainfall in Tigray region based on Mekelle station. He employed univariate Box-Jenkins method to analyze rainfall in the region and found that SARIMA model was suitable for forecasting future value of monthly rainfall data and used this model to forecast 12-month rainfall pattern in the study area. Further he concluded that there is no tendency of decreasing or increasing pattern of monthly rainfall over the forecast period from January 2010 to September 2011.

Adejuwon (2010) studied annual rainfall in Nigeria using power spectral analysis based on Benin, Sapele, Warri and Forcados Synoptic station in Edo and Delta States (formerly Mid-Western Nigeria) over 67 years and found that Benin synoptic station shows significant spectral peaks at 6.7, 4.6 and 3.7 years periodicities. The most pronounced peak at the station was found to be 3.7 years periodicity. In Sapele, the most pronounced periodicity of 5 years was observed. Although, the spectral peaks were significant at 4.6 and 3.7 years, respectively, at Warri, the most pronounced of these peaks was found to be 3.7 years. However, in the case of Forcados, a single significant spectral peak of 3.6 years cycle was prominent and it was then concluded that periodicities were evident with significant cycles of between 3 and 6 years.

In order to evaluate the prediction efficiency, Olatayo et al., (2014) made use of 31 years of annual rainfall data from year 1982 to 2012 of Ibadan South West, Nigeria. The fuzzy time series model has it universe of discourse divided into 13 intervals and the interval with the largest number of rainfall data was divided into 4 sub-intervals of equal length. Three rules were used to determine if the forecast value under FTS is upward 0.75–point, middle or downward 0.25-point. ARIMA (1, 2, 1) was used to derive the weights and the regression coefficients, while the theil's regression was used to fit a linear model. The performance of the model was evaluated using mean squared forecast error (MAE), root mean square forecast error (RMSE) and Coefficient of determination (R^2) . The study reveals that FTS model can be used as an appropriate forecasting tool to predict the rainfall, since it outperforms the ARIMA and Theil's models. Therefore, going through some of the available literatures on modeling climatic data series, most of the authors compared non- linear model with linear and stationary models, fitted their models with few data and based the data on raining periods. In this article, attempt was made to compare models which are capable of capturing the non-linearity of the series, since most of the climatic data series are non-linear in nature.

Most of the existing ANN models applied in rainfall forecasting are event based; the models were fed in input with screened/generated data that contained only rainy periods (i.e., rainfall events, rainy days or monthly rainfall data). By using only the data from the rainy periods as training data, ANN models could easily identify the patterns characterizing the rainfall. How- ever, on the other hand, any features or characteristics not included within the training data will not be learned by ANN. This implies the conventional ANN models could only issue accurate forecasting when rain occurred already and they can estimate how long the rain would last, but they are unable to predict whether it would rain or not if there is no rain at the time of issuing forecast. Hence, most of the previous studies of ANN on rainfall forecast are not fully suitable for the application in real time forecasting. But in this article, we extended ANN model to capture both rainy and non-rainy periods of the years. The extended ANN model was designed to run a real time task, in this situation, the input to the model should not be event based data but consecutive data including non-rainy periods to acquire a fully representation of both rain and no rain conditions.

2. Materials and Methods

An artificial neural network (ANN) is an interconnected group of arti?cial neurons that has a natural property for storing experiential knowledge and making it available for use. This original perceptron model contained only one layer, inputs are fed directly to the output unit via the weighted connections. Although the perceptron initially seemed promising, it was eventually proved that perceptron could not be trained to recognize many classes of patterns. After that, multilayer perceptron (MLP) model was derived and gradually became one of the most widely implemented neural network topologies. Multilayer perceptron means a feed-forward network with one or more layers of nodes between the input and output nodes. The MLP overcomes many limitations of the single layer perceptron, their capabilities stem from the non-linear relationships among the nodes. In their study of nonlinear dynamics, Lapedes and Farber (1987) have pointed out the important that the MLP is capable of approximating arbitrary functions.

Two important characteristics of the MLP are: its nonlinear processing elements (PEs) which have a nonlinearity that must be smooth (the logistic function and the hyperbolic tangent are the most widely used); and their massive interconnectivity (i.e. any element of a given layer feeds all the elements of the next layer).

The first or the lowest layer is an input layer where external information is received. The last or the highest layer is an output layer where the problem solution is obtained. The input layer and output layer are separated by one or more intermediate layers called the hidden layers. The nodes in adjacent layers are usually fully connected by a cyclic arcs from a lower layer to a higher layer.

The functional relationship estimated by the ANN can be written as $y = f(x_1, x_2, \ldots, x_p)$ where $x_1, x_2, \ldots, x_p$ are p independent variables and y is a dependent variable. In this sense, the neural network is functionally equivalent to a nonlinear regression model. On the other hand, for an extrapolative or time series forecasting problem, the inputs are typically the past observations of the data series and the output is the future value. The ANN performs the following function mapping $y_{t+1} = (y_t, y_{t-1}, \ldots, x_{t-p})$ where y_t is the observation at time t. Thus the ANN is equivalent to the nonlinear autoregressive model for time series forecasting problems. It is also easy to incorporate both prediction variables and time lagged observation into one ANN model, which amount to the general transfer function model. Before an ANN can be used to perform any desired task, it must be trained to do so.

Basically, training is the process of determining the arc weights which are the key elements of an ANN. The training input data is in the form of vectors of input variables or training patterns.

The total available data is usually divided into a training set (in-sample data) and a test set (out-of-sample or hold-out sample) and a training pattern consists of a fixed number of lagged observations of the series. Suppose we have N observations $y_1, y_2, \ldots, y_N$ in the training set and we need 1-step-ahead forecasting, then using an ANN with n input nodes, we have $n - N$ *training* patterns. The first training pattern will be composed of $y_1, y_2, \ldots, y_{n+1}$ as inputs and y_{n+1} as the target output. The second training pattern will contain $y_{N-n}, y_{N-n+1}, \ldots, y_{N-1}$ as inputs and y_{n+2} as the desired output. Finally, the last training pattern will be $y_1, y_2, \ldots, y_N$ for inputs and y_N for the target output. If a multilayer perceptron has a linear activation function in all neurons, that is, a linear function that maps the weighted inputs to the output of each neuron, then it is easily proved with linear algebra that any number of layers can be reduced to the standard two-layer input-output. What makes a multilayer perceptron different is that some neurons use a nonlinear activation function which was developed to model the frequency of action potentials, or firing, of biological neurons in the brain.

The two main activation functions used in current applications are both sigmoids and are described by

$$y\left(v_i\right) = tanh\left(v_i\right) \tag{1}$$

and

$$y\left(v_i\right) = \left(1 + e^{-v_i}\right)^{-1} \tag{2}$$

where y is the output and v_i is the level of training of the i^th node (Neuron).

The former function is hyperbolic tangent which ranges from -1 to 1, and the latter, the logistic function, is similar in shape but ranges from 0 to 1. Here y_i is the output of the i^{th} node (neuron) and v_i is the weighted sum of the input synapses. The multilayer perceptron consists of three or more layers (an input and an output layer with one or more hidden layers) of nonlinearly-activating nodes and is thus considered a deep neural network. Each node in one layer connects with a certain weight w_{ij} to every node in the following layer. Some people do not include the input layer when counting the number of layers and there is disagreement about whether w_{ij} should be interpreted as the weight from $i\ to\ j$ or the other way round.

Learning occurs in the perceptron by changing connection weights after each piece of data is processed, based on the amount of error in the output compared to the expected result. This is an example of supervised learning, and is carried out through back-propagation, a generalization of the least mean squares algorithm in the linear perceptron.

We represent the error in output node j in the n^{th} data point (training example) by

$$e_j\left(n\right) = d_j\left(n\right) - y_j\left(n\right) \tag{3}$$

where d is the target value and y is the value produced by the perceptron. We then make corrections to the weights of the nodes based on those corrections which minimize the error in the entire output, given by

$$\varepsilon\left(n\right) = \frac{1}{2}\sum_{j=1}^{n} e_j^2\left(n\right) \tag{4}$$

2.1 Multilayer Perceptron Model

Based on the work of Gardnera and Dorlinga (1998), a Multilayer perceptron model is proposed and this is discussed below. The proposed extended ANN model is given thus Let denote the weight of the link from the i^{th} neuron in the l^{th} layer to the j^{th} neuron in the $l + i^{th}$ layer $w_{li,\ (l+1)j}$, then the i^th neuron in the l^{th} layer is given by

$$y_{il} = f_{li}\left(z_{li}\right):\ z_{li} = \sum_{j=1}^{n_t-1} w_{(l-1)}j,\ l_i y_{(l-1)j} + b_{li} \tag{5}$$

where y_{il}, f_{li} *and* b_{li} are respectively its output, activation function and bias.

2.2 Basic Concept of Fuzzy Time series

Song et al (1993a, 1994b and 2003) proposed the definition of fuzzy time series based on fuzzy sets as follows: Let T be the universe of discourse, $T = \{t_1, t_2, \ldots, t_n\}$ and let Z be a fuzzy set in the universe of discourse U defined as follows:

$$Z = f_z\,(t_1)/t_1 + \; f_z\,(t_2)/t_2 + \cdots + f_z\,(t_n)/t_n \tag{6}$$

where f_z is the membership function of $Z.f_Z : T \rightarrow [0, 1]\,, f_z(t_i)$ indicates the grade of membership of t_i in the fuzzy set Z, $f_z\,(t_i) \in [0, 1]$ and $1 \leq i \leq n$. For more details on fuzzy time-series model, see Olatayo et al (2014).

2.3 Autoregressive Integrated Moving Averages (ARIMA) Models

Autoregressive integrated moving average (ARIMA) models are specific subset of univariate modeling, in which a time series is expressed in terms of past values of itself (the autoregressive component) plus current and lagged values of a white noise error term (the moving average component). ARIMA models are univariate models that consist of an autoregressive polynomial, an order of integration (d), and a moving average polynomial, Olatayo et al., (2014).

A process (x_t) is said to be an autoregressive integrated moving average process, denoted by $ARMA\,(p,\; d, q)$if it can be written as:

$$\emptyset\,(B)\,\nabla^d x_t = \theta\,(B)\,w_t \tag{7}$$

where $\nabla^d = (1 - B)^d$ with $\nabla^d x_t$and d^{th}consistitutes differencing.

2.4 Forecasting

There are two kinds of forecast and these are sample period forecasts and post-sample period forecasts. The former will be used to develop confidence in the model and the latter will be used to generate genuine desired forecasts using neural network, ARIMA and fuzzy time series models obtained. In forecasting, the goal is to predict future values of a time series, $x_{t+m},\; m = 1, 2, \ldots$ based on the data collected, $x = \{x_t, x_{t-1}, \ldots, x_1\}$. (Olatayo and Alabi 2011) and (Taiwo and Olatayo 2013).

2.4.1 Forecasting Accuracy Measures

Once forecasts are made, they can be evaluated if the actual values of the series to be forecasted are observed. The measurement parameters for the accuracy of forecasts to be used in this research are sum of square error, mean square error and root mean square error respectively. This is given by

$$SSE = \frac{1}{h} \sum_{t=s}^{h+s} \left(X_t - \overline{X}_t\right)^2 \tag{8}$$

$$MSE = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{X}_t - X_t\right)^2 \tag{9}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\hat{X}_t - X_t\right)^2} \tag{10}$$

The actual and predicted values for corresponding t values are denoted by $\hat{X}_t \; and \; X_t$ respectively. The smaller the values of *RMSE SSE* and M*SE*, the better the forecasting performance of the model. (Olatayo and Taiwo 2012), (Olatayo and Adeboye 2013) and (Olatayo et al 2014).
Therefore, the model with minimum parameters of measurements is the most efficient model for predictive or forecasting purpose.

3. RESULTS AND DISCUSSIONS

The ANN analysis was based on multilayer perceptron and rainfall series as the dependent variable while temperature and evaporation were used as factors and covariates respectively. The data used in this research was obtained from the Nigerian Meteorological Agency, Lagos, Lagos State. It consists of yearly annual rainfall from 1982 to 2014.

3.1 Data Analysis with Multilayer Perceptron

Since the climatic variables are not obtained based on the same measurement then the covariate will be rescaled by using the method of standardization by subtracting y_i from the mean and divide by the standard deviation, that is

$$Standardized = \frac{(y_i - mean)}{S} \tag{11}$$

The first stage in multilayer perceptron is to partition the dataset and randomly assign cases based on relative number of cases and specify relative number (ratio) of cases to each sample that is, training, testing, and holdout. Here we specify 7, 3, and 0 as the relative numbers for training, testing, and holdout samples corresponding to 70%, 30%, and 0%.
In the architecture stage, we considered two hidden layers and the activation function will be based on sigmoid given as; $y\left(v_i\right) = \left(1 + e^{-v_i}\right)^{-1}$ and it takes real-valued arguments and transforms them to the range (0, 1).
The training stage was used to determine how the network processes the records. We make used of batch training method since it can update the synaptic weights only after passing all training data records; that is, batch training uses information from all records in the training dataset and it is preferred because it directly minimizes the total error and the optimization algorithm, Olatayo (2015).

Table 1. Estimated Parameter Values.

Model	SSE	MSE	RMSE
ARIMA	110.23	3.4447	1.8560
Fuzzy Time Series	85.45	2.6703	1.6341
Neural Network	0.051	0.0016	0.04

3.2 Interpretation and Discussion

This work comparatively evaluates the forecast performance of three model which are ANN, ARIMA and fuzzy time series. With respect to climate change, it was evidence from the time plots of rainfall data series that the series contain cyclical behavior and this is due to seasonal changes in the series. Augmented dickey test showed stationarity of the series and Autoregressive integrated moving average model obtained was

$$y_t = 4.37 + u_t$$
$$(1 - 0.39L)\, u_t = (1 - 0.99L)\varepsilon_t$$

with the white noise variance $\widehat{\sigma}_{\varepsilon}^2$ estimated as 17452. However, the forecast of the yearly rainfall from 1982 to 2014 deviated slightly from the original data, sum of square error (SSE) = 110.23, MSE = 3.4447 and RMSE = 1.8560.

Under fuzzy time series, we made use of the visual Basic Version 6.0 on a Pentium 4 PC. From the summary of the forecasting results of fuzzy time series method, Olatayo and Taiwo (2015), where the universe of discourse is divided into 13 intervals and the interval with the largest number of rainfall data is divided into 4 sub-intervals of equal length. The fuzzy time series forecast of the yearly rainfall data from 1982 to 2014 did not deviated much from the original data, the measurement of performance with respect to SSE = 85.45, MSE = 2.6703 and RMSE = 1.6341.

Under ANN, we made use of SPSS Version 20.0. The ANN result shows that 23 observation are trained and 1 observation is tested and this implies that 24 observations are valid. This shows that 95.8% of the observations inputted is trained and 4.2% is tested. The rainfall prediction from the ANN model with respect to multilayer perceptron shows a slight perfect match to the original data. The forecast performance measurement shows that SSE = 0.051, MSE = 0.0016 and RMSE = 0.04.

Based on the forecasting parameters performance measurements of ARIMA, FTS and ANN models, there are indication that the ANN (Multilayer perceptron) model outperforms both the ARIMA and FTS model. While the FTS model is better than the ARIMA model. Therefore, our study establishes that ANN method is more efficient, better and an appropriate forecasting tool to model and predict rainfall in order to combat complexity in climatic change.

4. Conclusion

Complexity of the nature of rainfall as a factor for climatic change with evident of non-linearity has been studied using ANN, ARIMA and FTS techniques. The study reveals that ANN model can be used as an appropriate forecasting tool to predict the rainfall series, which is more efficient with minimum parameters and performs better than the ARIMA and FST models. Therefore, our study establishes that ANN method is better and an appropriate forecasting tool to model and predict rainfall in order to combat climatic change. The ANN model was found to be more efficient with minimum parameters and capable of handling the non-linearity that are typical, in the case of climatic change data series.

References

[1] Adewole Adetunji P, Akinwale Adio T and Akintomide Ayo B. (2011). Artificial Neural Network Model for Forecasting Foreign Exchange Rate. ISSN: 2221-0741. *World of Computer Science and Information Technology Journal (WCSIT)*, **1** No: 3 559-570.

[2] Adejuwon, J. O. (2010). A spectral analysis of rainfall in Edo and Delta States, Nigeria (formerly Mid- Western Region). *International Journal of Climatology*, **31**, 2365-2370.

[3] Al-Ansari, A., Al-Shamali B.and Shatnawi A, (2006). Statistical Analysis of rainfall Records at three Major Meteorological Stations in Jordan, Al-Mararah University special publications, **12**.

[4] Amha, G., (2010): Modelling and forecasting monthly rainfall in Tigray region. Case Mekelle station. (M.Sc. Thesis), Department of Statistics, Addis Ababa University, Ethiopia.

[5] Fabio Augusto Mollick, (2010): A Neural Network Model for Forecasting Production time series in Brazilian Industries, POMS 21^{st} Annual conference, Canada.

[6] French, M. N., Krajewski, W. F., and Cuykendall, R. R (1992): Rainfall forecasting in space and time using neural network, J. Hydrol., **137**, 1-31.

[7] Gianluigi Reck, (2002): Forecasting with Artificial Neural network models, SSE/EFI Working Paper series in Economic and Finance, No 491.

[8] Guoqiang Zhang, B. Eddy Patuwo and Michael Y. Hu,(1998): Forecasting with Artificial Neural Network: the state of the act, International Journal of Forecasting, **14**, 35-62.

[9] Hung N. Q, Babel M. S, Weesakul S and Tripathi N. K., (2009): An Artificial Neural network model for Rainfall forecasting in Bangkok, Thailand, Hydol. Earth Syst. Sci., 1413-1425.

[10] Khashei, M, hejazi, S and Bijari, M, (2008): A new hybrid artificial neural networks and fuzzy regression model for time series forecasting, fuzzy sets and systems, 769-786.

[11] Lapedes, A. S. and Farber, R. M, (1987): Nonlinear signal processing using neural networks: Prediction and system modeling, Los Alamos Report LA-UR 87-2662.

[12] Luk, K. C., Ball, J. E., and Sharma, A (2000): A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting, J. Hydrol., **227**, 56-65.

[13] Mahsin, M.D. Yesmin, A. and Monira, B.,(2012). Modeling Rainfall in Dhaka Division of Bangladesh Using Time SeriesAnalysis. Journal of Mathematical Modelling and Application, **1**, No. 5, 67-73.

[14] Olatayo T. O and Alabi O. O, (2011): Forecasting Modelling in Stochastic Time Series Process Journal of Mathematical Sciences, **22** (2), 135-142.

[15] Olatayo T. O and Taiwo A. I, (2012). Measuring the Structural Relationship of Some Macro-Economic Time Series Data (A Vector Autoregressive Approach). African Journal of Pure and Applied (ASPS). **2** (1), 63-69.

[16] Olatayo T. O and Adeboye N.O, (2013): Predicting Population Growth through Births and Death Rate in Nigeria, Mathematical Theory and Modelling, **3** (1), 96-101.

[17] Olatayo T. O, Taiwo A. I and Afolayan R. B. (2014), Statistical Modelling and Prediction of Time Series Data. Journal of the Nigerian Association of Mathematical Physics, **27**, 201-208.

[18] Saeed Moshiri and Norman Cameron, (2000): Neural Network versus Econometric models in Forecasting Inflation, Journal of Forecasting, **19**, 201-217.

[19] Song, Q. (2003). A note on fuzzy time series model selection with sample Autocorrelation functions. Cybernetics and Systems: An International Journal, **34**, 93-107.

[20] Song, Q. and Chissom, B. S. (1993a). Fuzzy time series and its models, Fuzzy Sets and Systems, **54**, 269-277.

[21] Song, Q. and Chissom, B. S. (1993). Forecasting enrollments with fuzzy time series –Part I. Fuzzy Sets and Systems, **54**, 1-9.

[22] Taiwo A. I. and Olatayo T. O., (2013): Measuring Forecasting Performance of Vector Autoregressive and Time Series Regression Analysis. American Journal of Scientific and Industrial Research, **4** (1), 49-58.