



Comparison of Artificial Neural Network and Box- Jenkins Models to Predict the Number of Patients with Hypertension in Kalar

Layla A. Ahmed

Department of Mathematics, College of Education, University of Garmian, Kurdistan Region, Iraq
layla.aziz@garmian.edu.krd

Article history: Received 7 January 2020, Accepted 20 February 2020, Published in October 2020

Doi: [10.30526/33.4.2516](https://doi.org/10.30526/33.4.2516)

Abstract

Artificial Neural Network (ANN) is widely used in many complex applications. Artificial neural network is a statistical intelligent technique resembling the characteristic of the human neural network. The prediction of time series from the important topics in statistical sciences to assist administrations in the planning and make the accurate decisions, so the aim of this study is to analysis the monthly hypertension in Kalar for the period (January 2011- June 2018) by applying an autoregressive –integrated- moving average model (ARIMA) and artificial neural networks and choose the best and most efficient model for patients with hypertension in Kalar through the comparison between neural networks and Box- Jenkins models on a data set for predict. Comparisons between the models has been performed using Criterion indicator Akaike information Criterion, mean square of error, root mean square of error, and mean absolute percentage error, concluding that the prediction for patients with hypertension by using artificial neural networks model is the best.

Keywords: Hypertension, time series, autoregressive-integrated-moving average model, artificial neural networks.

1.Introduction

Last years, after increasing the number of patients with chronic hypertension disease, it was to be highlighted to study this disease and the use of statistical methods and artificial intelligence techniques. Hypertension is defined as the abnormal high blood pressure (more than 120/8 mm. Hg) in the arteries [1]. Uncontrolled high blood pressure makes you more likely to get heart disease, stroke, and kidney disease [2].



The time series forecasting assumes that the future values a linear combination of historical data. There are various time series forecasting models; however, the most highly frequently approach to fit such model is Box and Jenkins for fitting ARIMA model.

Box and Jenkins (1970) generalized the ARIM model to deal with seasonality [4]. Tiao and Box (1979) described a practical to ARMA modeling of multivariate time series data by three stages: identification, estimation of the parameters and model checking [5].

Artificial Neural Network is extensively used in construction industry [6], analyzing the business data stored in database [7-8], robotics industry systems, decision support systems, automated control systems, and prediction systems [9]. Artificial neural network is a calculation method resembling the characteristic of the human neural network. The important characteristics of neural network involve nonlinearity, capacity to handle large data, and generalization [10].

The main aim is to choose the best and efficient model for forecasting the number of patients with hypertension in Kalar. Through the comparison between neural networks and Box- Jenkins models.

In this paper, introduction, autocorrelation function, partial autocorrelation function, and autoregressive moving average model are introduced. Next, the neural network model is introduced and then the results of these models are compared. Finally, conclusions and recommendation for this study are given.

2. Methodology

2.1. Autocorrelation Function (ACF)

The autocorrelation function is the correlation of time series $(z_1, z_2, \dots, z_n)$ with itself. The correlation coefficient between z_t and z_{t-1} is called lag-k autocorrelation of z_t and denoted by ρ_k , which is under the assumption of weak stationary and is defined [4], [11]:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^n (z_t - \bar{z})^2}, k = 1, 2, \dots, \max k \quad (1)$$

Where

$$\bar{z} = \frac{\sum_{t=1}^n z_t}{n}$$

$$\gamma_k = \text{cov}(z_t, z_{t+k})$$

With the following Properties:

$$-1 \leq \rho_k \leq 1$$

$$\rho_k = \rho_{-k}, \quad \rho_0 = 1$$

2.2. Partial Autocorrelation Function (PACF)

The correlation coefficient between z_t and z_{t-1} after removing the impact of the intervening $z_{t-1}, z_{t-2}, \dots, z_{t-k+1}$ is called partial autocorrelation function at lag- k, denoted by ϕ_{kk} , and defined⁽¹²⁾:

$$\Phi_{kk} = \begin{cases} 1 & , k = 0 \\ \rho & , k = 1 \\ \frac{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \dots & 1 \end{vmatrix}} = 0 & , k \geq 2 \end{cases} \quad (2)$$

Calculating partial autocorrelation function of sample form:

$$\Phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \Phi_{k-1,j} \rho_{k-j}}{1 - \sum_{j=1}^{k-1} \Phi_{k-1,j} \rho_j} \quad (3)$$

Where

$$\Phi_{kj} = \Phi_{k-1,j} - \Phi_{kk} \Phi_{k-1,k-j} \quad , j = 1, 2, \dots, k-1$$

2.3. Autoregressive Moving Average Model (ARMA)

The autoregressive moving average model is denoted by $ARIMA(p, d, q)$

Where

p : The order of autoregressive.

q : The order of moving average.

d : The order of non-seasonality difference.

For stationary time series, the general form of an ARMA model can be written as:

$$\phi(B)x_t = \delta + \theta(B)a_t \quad (4)$$

A non-stationary series should be first transformed into a stationary one by considering relevant differences:

$$\nabla^d x_t = (1 - B)^d x_t = x_t - x_{t-d} \quad (5)$$

$(1 - B)^d$: The d^{th} difference.

B : The backward shift operator.

For non-stationary time series, the general form of an ARMA model can be written as [12], [13]:

$$\phi(B)(1 - B)^d x_t = \delta + \theta(B)a_t \quad (6)$$

Where

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$, is the autoregressive operator of order p .

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$, is the moving average operator of order q .

$$a_t = NID(0, \sigma_a^2) \quad (7)$$

2.4. Model Building

The Box-Jenkins iterative approach for constructing linear time series models consists of four stages [3], [12]: identification, estimation of the parameters, diagnostic checking and forecasting.

The Identification stage is the most important. It consists of the appropriate model from ARIMA models.

The major approaches of appropriate ARIMA models are nonlinear least squares and maximum likelihood estimation.

Diagnostic model checking involves testing the assumptions of the model to identify any area where the model is inadequate.

For goodness of fit test for the suggested ARIMA model such that:

$$Q_k = n(n+2) \sum_{k=1}^{\max k} (n-k)^{-1} \quad (8)$$

$$\left. \begin{array}{l} H_0 = \rho_k(a_t) = 0 \\ H_1 = \rho_k(a_t) \neq 0 \end{array} \right\} \quad (9)$$

If Q lies in the extreme 5% of the right side tail of the chi-square distribution, we reject the hypothesis where the residuals are random.

2.5 Neural Networks

2.5.1 Biological Neural

The biological neuron receives inputs from all components of body, combines the input, performs a nonlinear operation and offers the output result [14]. Human brain is highly complex, nonlinear and parallel computer [15]. The human brain includes about 100 billion neurons. The mean neuron is as complex as a small computer and has as many as 10000 physical connections with other cells [16]. A neuron contains of four parts called cell body, dendrites, axon, and synapses [14], [17].

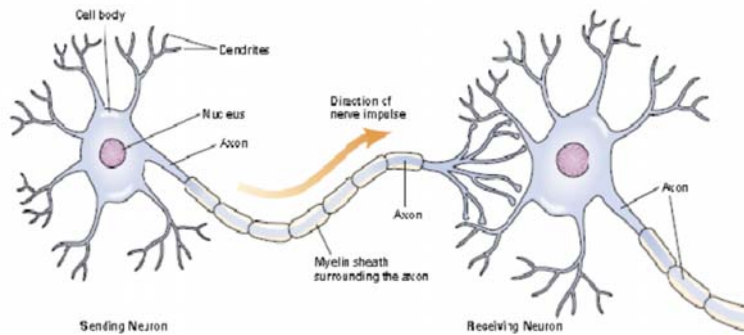


Figure 1: Biological neural network

2.5.2 Artificial Neural Network (ANN)

Artificial neural network is a computational model that functions like a human brain through biological neurons [18]. The Artificial neural networks perform helpful calculations and simulate complex modeling in the functioning of the human brain. The first artificial neuron model was proposed in 1943 by McCulloch and Pitts [19-20].

The neural networks approach is widely applied in biological, engineering, medical, financial [21]. The artificial neural network can be designed using either feed forward or fed back approach. There are three types of layers, input layer, hidden layer, and output layer in an artificial neural network [9-10], as shown in Figure 2.

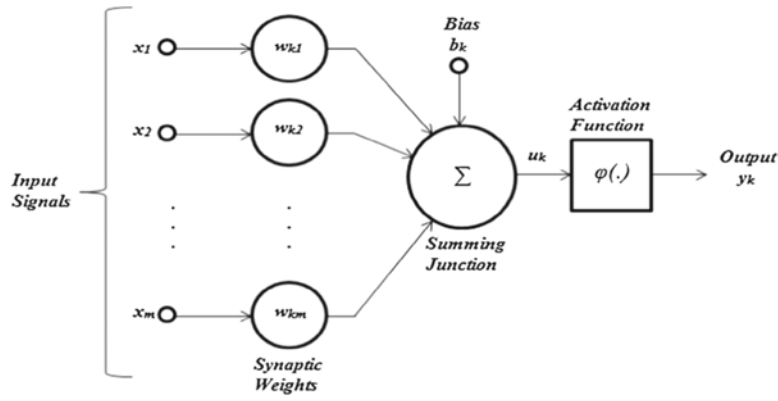


Figure 2: Artificial neural network

A statistical model is [7], [16]:

$$y_i = f(a_j) = \frac{1}{1 + \exp(-a_j)} \quad (10)$$

Where y_i is the output of the n^{th} layer; $f(a_j)$ is the activation function widely employed by the logistic sigmoid, hyperbolic tangent sigmoid and squared functions [9], [21]. The sigmoid activation function, commonly used and applied in this study, and a_j is the sum of the weight of the previous layer, which is obtained by [22]:

$$a_j = \sum_{i=1}^n w_{ij} x_i + b_j \quad (11)$$

Where w_{ij} is the linkage weight from the neuron i to neuron j ; x_i is the input data from neuron i to j ; b_j is the bias on the neuron j , and n is the total number of input neurons.

Efficiencies of human neurons and processing elements of ANN can be compared as synapses act like a weight of the arriving stimulus and inspired the weights of ANN; dendrites a cumulates the arriving weighted stimulus, inspired the summing function of ANN; cell body, that reasons conversion of collected stimulus in to a new stimulus, inspires activation function; axon, which distributes the new stimulus to the conformable neurons, inspires the output and output links; and finally, threshold value with a role of activating or inactivating increase and decrease of the stimulus, inspires the bias [17]. The hidden nodes can be in single layer or multi-layers. Usually, the multiple layers neurons are called as multiple layers perception [22].

The multiple layers perception neural network is an important architecture, and it is also one of the most widely used architectures that are applicable to different application problems [23].

The criteria of comparison between the models: *Akaike* information criterion (*AIC*) [24], [25-26], mean square of error (*MSE*) [10], root mean square of error (*RMSE*) [21], and mean absolute percentage error (*MAPE*) [27]. The formulas are expressed as below:

$$AIC = -2\text{Log}L + 2m$$

$$AIC = n(1 + \text{Log}\pi) + n\text{Log}\sigma_a^2 + 2m$$

$$AIC = n\ln\sigma_a^2 + 2m \quad (12)$$

Where:

L: The likelihood function.

σ_a^2 : The mean square of residuals.

m : Parameters of the model.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (13)$$

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right]^{\frac{1}{2}} \quad (14)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{Y_i} \times 100 \quad (15)$$

Where Y_i is the actual value and $\hat{Y}_i$ is the predicted value, and n is the size sample. The final model is used to generate predictions about the future values and then calculate the forecast errors as developments of new values watch from the time series [23].

3. Data Analysis and Results

3.1. Data Description

The data of current study were the monthly observations that included (5169) patients of hypertension in Kalar city and were obtained from the records of the general hospital- Kalar for the period (Jan. 2011- Jun. 2018) in order to reach an appropriate model to be used to forecast the monthly hypertension. SPSS 22 and Minitab17 statistical software were used.

Table 1 below shows the descriptive statistics of the series and indicates the mean of the time series that(57.43), the sample size is 90 months. From the data, we considered that the number of patients were female (60.03%) females and (39.97%) males.

Table 1: Descriptive statistics of the data

Gender	No. of patients	Mean	Max.	Min.	Std. Dev.	Percent
Male	2066	22.8889	68	8	9.7778	39.97
Female	3103	34.4778	76	9	15.83	60.03
Total	5169		144	17	23.9996	100

3.2. Results of Box- Jenkins

Figure 3, represents the series monthly hypertension and we show that data are stationary in the variance, but not stationary in the mean when we plot autocorrelation functions and partial autocorrelation function for the data.

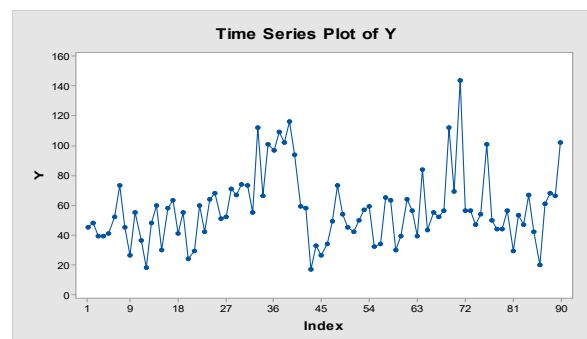


Figure 3: Plot of the data

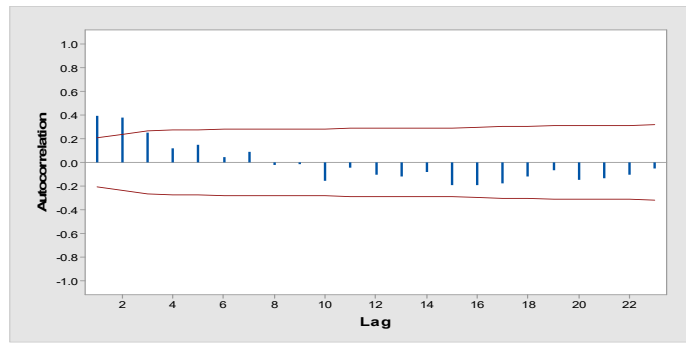


Figure 4: ACF of the data

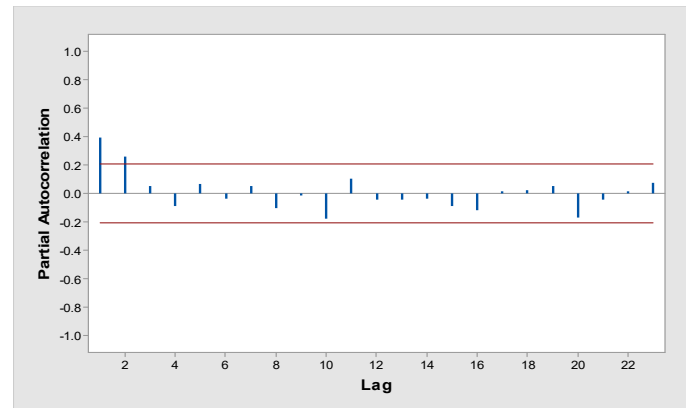


Figure 5: PACF of the data

We treat the outlier problem, then take the first difference for the data and plot ACF and PACF again for the difference time series; we show that the series become stationary in the mean and variance.

3.3. Choosing Appropriate Model

We choose the appropriate model through the use of *Akaike's* information criterion (AIC) as shown in table (2). The best model is the $ARIMA(2,1,1)$, because the values of AIC are minimum. Since the model is:

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)y_t = \delta + (1 - \theta_1 B)a_t \quad (16)$$

Table 2: Iterations of ARIMA models

Model	Res. VAR.	AIC
$ARIMA(1,1,0)$	370.6	528.446
$ARIMA(1,1,1)$	319.2	517.158
$ARIMA(0,1,1)$	328.5	517.714
$ARIMA(0,1,2)$	327.0	519.306
$ARIMA(1,1,2)$	330.9	522.362
$ARIMA(2,1,2)$	333.9	525.165

<i>ARIMA</i> (2 ,1, 1)	310.2	516.612
<i>ARIMA</i> (2 ,1, 0)	336.5	521.855
<i>ARIMA</i> (3 ,1, 0)	338.5	524.383
<i>ARIMA</i> (3 ,1, 1)	325.7	522.952
<i>ARIMA</i> (0 ,1, 3)	330.8	522.335
<i>ARIMA</i> (2 ,1, 3)	304.7	519.02
<i>ARIMA</i> (3 ,1, 3)	318.7	525.02

3.4. Estimating the Parameters

Type	Coefficient	SE. Coefficient	T	P – Value
<i>AR1</i>	0.3548	0.1122	3.16	0.002
<i>AR2</i>	0.1418	0.1129	1.26	0.212
<i>MA1</i>	1.006	0.0097	103.63	0.000
<i>Constant</i>	0.066	0.044	1.48	0.14

Modified Box – Pierce (Ljung – Box)Chi – Square Statistic				
<i>Lag</i>	12	24	36	48
<i>Chi – Square</i>	8.6	18.1	26.6	34.2
<i>DF</i>	8	20	32	44
<i>P – Value</i>	0.379	0.579	0.737	0.856

Number of observations: Original series 90, after taking the difference 89
Residuals: SS = 26371.1, MS =310.2, and d. f= 85

$$1 - 0.3548B - 0.1418B^2)(1 - B)y_t = 0.066 + (1 - 1.006B)a_t$$

3.5. Diagnostic Model Checking

After checking the stationary for the series, the autocorrelation functions and partial autocorrelation function of the residuals are plotted as shown in figures 6 and 7. It is clear that the autocorrelation functions and partial autocorrelation function of the residuals are values that fall within the confidence limits of probability 95%.

$$-1.96(1/\sqrt{n}) \leq r(\hat{e}) \leq 1.96(1/\sqrt{n})$$

The test is based on first 24 autocorrelations

P-value = 0.579 > 0.05

Accept H_0 , residuals are random

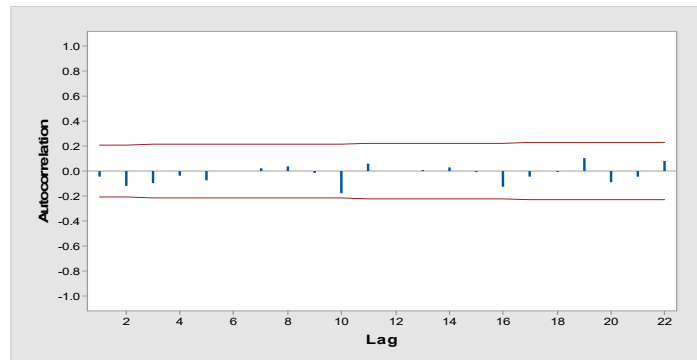


Figure 6: ACF of residual

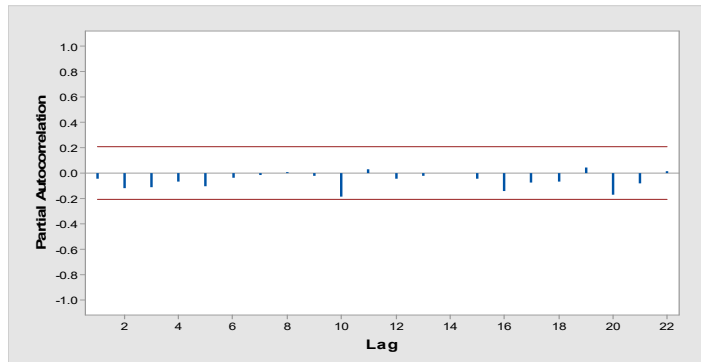


Figure 7: PACF of residual

After verification of the suitability of the model that the model used for forecasting values for monthly hypertension based on $ARIMA(2,1,1)$, **Table 2** showed that.

3.6 Results of ANN

The data were randomly divided into two independent training and testing subsets, (80%) of the data was considered for network training and (20%) of data was used for network reliability and used for testing.

Table 3 shows that the best model includes five units in the hidden layer, as well as four lags.

Table 3: Results of ANN model for the data

Number of lags	1	2	3	4	5
SSE	21.354	28.266	30.696	18.817	29.594
RMSE with 5units	0.553	0.716	0.787	0.53	0.722

Table 4 contains a comparison between of $ARIMA(1, 1, 2)$ and ANN5 (4) of the data in order to determine the best appropriate model for prediction hypertension through the use of MSE, RMSE, and MAPE. The results of this analysis show that the best model is the ANN5 (4), because the values of MSE, RMSE, and MAPE are minimum. After verification of the suitability of the model that the model used for forecasting values of the number of patients for two years in monthly (July 2014-June 2016), **Figure 8** showed that.

Table 4: Comparison of models

Year	Actual	ARIMA Forecast	ANN5(4) Forecast
Aug.2017	56	51.84	51.07
Sep.	29	54.69	38.84
Oct.	53	56.88	51.83
Nov.	47	58.13	51.83
Dec.	67	58.94	66.45
Jan. 2018	42	59.48	47.23
Feb.	20	59.85	47.21
Mar.	61	60.12	80.72
Apr.	68	60.34	80.72
May.	66	60.52	47.23
Jun	102	60.68	80.72
MSE	0	415.60	206.34
RMSE	0	20.39	14.37
MAPE	0	40.18%	29.18%

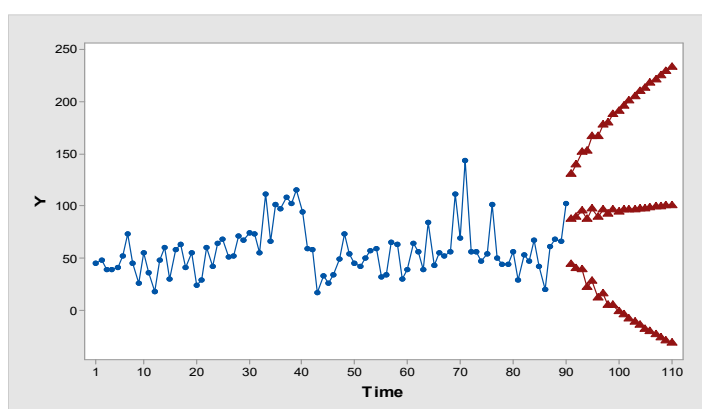


Figure 8: Plot of time series

4. Conclusions

The results of application show that (60.03) of patients with hypertension are females and (39.97) of patients with hypertension are males. The statistical tests show that the data are stationary in the variance, but not stationary in the mean. The model for monthly prediction of hypertension by using Box- Jenkins is the model of $ARIMA(2,1,1)$, and by using artificial neural network model is $ANN5(4)$. The model obtained by using artificial neural network is a suitable model for prediction, because when we use this model on data, it gives us less value of mean squares error, root mean squares error, and mean absolute percentage error. This is accuracy signs of this model.

We recommend that the health and medical practitioners should use the ANN models for the forecasting of hypertension disease cases in Kalar, because of their efficiency in prediction and their low error rates.

References

1. R., Siyad A. Hypertension, *Hygeia Journal for Drugs and Medicines*. **2011**, 3, 1, 1-16.
2. Tawfiq, LNM; Oraibi. YA. Fast Training Algorithms for Feed Forward Neural Networks. *Ibn Al-Haitham Journal for Pure and Applied Science*. **2017**, 26, 1: 1275-1280
3. Ahmed, L. A. Using ARIMA Time Series Model to Forecast Production in Spinning Factory- Kifri, 2nd Scientific Conference, University of Garmian, **2015**, 1, 125-140.
4. Chatfield, C. *The Analysis of Time Series: An Introduction*; 2nd Edition, Chapman and hall, London New York, **1980**; ISBN 0412224607.
5. Li W. K. ; McLeod, A.I. Distribution of the Residual Autocorrelations in Multivariate ARMA Time Series Models, *Royal Statistical Society, Series B (Methodological)*, **1981**, 43, 2, 231-239.
6. Golnaraghi, S. Z. et al. Application of Artificial Neural Network(s) in Predicting Formwork Labour Productivity, *Advances in Civil Engineering*, **2019**, article ID 5972620, 1-11, [http:// doi.org/10.1155/2019/5972620](http://doi.org/10.1155/2019/5972620).
7. Maliki, O. S. et al. Comparison of Regression Model and Artificial Neural Network Model for the Prediction of Electrical Power Generated in Nigeria; *Advances in Applied Science Research*, **2011**, 2, 5, 329-339.
8. Lee, K. Y.; Chung, N. ; Hwoung, S. Application of an Artificial Neural Network Model for Predicting Mosquito abundances in Urban Areas; *Ecological Informatics*, **2016**, 36, 172-180.
9. Al- Maqaleh, B. M.; Al- Mansoub, A. A ; Al- Badani, F. N. Forecasting Using Artificial Neural Network and Statistics Models, *International Journal Education and Management Engineering*, **2016**, 3, 20-32., doi:10.5815/ijeme.2016.03.03.
10. Ranganayki, V. ; Deepa, S. N. An Intelligent Ensemble Neural Network Model for Wind Speed Prediction in Renewable Energy Systems, *The Scientific World Journal*, **2016**, Article ID 9293529, 1-15, <http://dx.doi.org/10.1155/2016/9293529>.
11. Yassen, A. A. Comparative Study of Artificial Neural Network and ARIMA Models for Economic Forecasting, M.Sc. Thesis, University of Al- Azahra, Gaza, **2011**.
12. Dobre, I. and; Alexandru, A. A. Modeling Unemployment Rate Using Box-Jenkins Procedure, *Journal of Applied Quantitative Methods*, **2008**, 3, 2, 156-166.
13. John, E. E. and Patric, U. U. Short- Term Forecasting of Nigeria Model; *Science Journal of Applied Mathematics and Statistics*, **2016**, 4, 3, 101-107, doi:10.11648/j.sjams.20160403.13.
14. Singh, V. Application of Artificial Neural Networks for Predicting Generated Wind Power, *International Journal of Advanced Computer Science and Applications*, **2016**, 7, 3, 250-253.
15. Tosun, E. K., Aydin, K. ; Bilgili, M. Comparison of Linear Regression and Artificial Neural Networks Model of a Diesel Engine Fueled with Biodiesel- Alcohol Mixtures; *Alexandria Engineering Journal*, **2016**, 55, 3081-3089, doi:10.1016/j.aej.2016.08.011.
16. Balaji, S. A. ; Baskaran, K. Design and Development of Artificial Neural Networking System Using Sigmoid Activation Function to Predict Annual Rice Production in Tamilnadu; *International Journal of Computer Science, Engineering and Information Technology*, **2013**, 3, 1, 13-31, doi:10.512/ijcseit.2013.3102.

17. Guresen, E. ; Kayakutlu, G. Definition of Artificial Neural Networks with Comparison to Other Networks; *Procardia Computer Science*, **2011**, 3, 426-433, doi:10.1016/j-procs.2010.12.071.
18. Idris, M. A.; Jami, M. S. and Hammed, A. M. , Optimization process of moringa oleifera seed extract using artificial neural network (ANN); *Malaysian Journal of Fundamental and Applied Sciences*, **2019**, 15, 2, 254-259.
19. BuHamara, S.; Smaoui, N. and Gabr, M. The Box- Jenkins Analysis and Neural Networks: Prediction and Time Series Modeling, *Applied Mathematical Modeling*, **2003**, 7, 805-815, doi: 210.1016/s0307-904x (03)00079-9.
20. Zakaria, M.; Al- Shebany, M., ; Sarhan, S. Artificial Neural Network: A Brief Overview; *International Journal of Engineering Research and Applications*, **2014**, 4, 2, version.1, 7-12,
21. Alhashimi, S. A. Muttaleb. Prediction of Monthly Rainfall in Kirkuk Using Artificial Neural Network and Time Series Mode; *Journal of Engineering and Development*, **2014**, 18, 1, **2016**129-143.
22. Yusof, K. K. et al. The evaluation on artificial neural networks (ANN) and multiple linear regressions (MLR) models over particulate matter (PM10) variability during haze and non-haze episodes: A decade case study; **2019**, 15, 2, 164-172.
23. Jalham, I. S. Filtrated Artificial Neural Network Approach to Predict the Mechanical Behavior of Polystyrene Reinforced with Bakelite; *Engineering Sciences*, **2013**, 39, 1, 66-74.
24. Iwok, I. A. Seasonal Modeling of Fourier Series with Linear Trend, *International Journal of Statistics and Probability*, **2016**, 5, 6, 65-72, doi: 10.5539/ijsp.vn6p65.
25. Obubu, M. et al. Modeling and Forecasting of Armed Robbery Cases in Nigeria Using Auto Regression Integrated Moving Average (ARIMA) Models; *European Journal of Statistics and probability*, **2016**, 4, 4, 28-41.
26. Ping, Y. P.; Hamizah, N. ; Maizah, H. A. Forecasting Malaysian Gold Using GARCH Model; *Applied Mathematical Sciences*, **2013**, 7, 58, 2879-2884.
27. Yazdi, F. M. Application of Neural Networks for 24- Hour- Ahead Load Forecasting; *International Journal of Electrical and Computer Engineering*, **2009**, 3, 2. 248-251.