

A comparison between the ARIMA model and neural networks average death in Iraq for the period (1980-2018)

Ashraf Mohamed Shareefa^a, Sarmad Jaafar Nasera^{a,*}, Ahmad M. Hashim^b

^aFaculty of Nursing, National University of Science and Technology, Thi-Qar, Iraq

^bMazaya University College, Nasiriyah, Iraq

(Communicated by Madjid Eshaghi Gordji)

Abstract

In this research, the (Box-Jenkins) methodology and artificial neural networks in predicting theoretical and practical levels were identified and clarified by constructing time series models and artificial networks to predict the mortality rate of Iraq and the data represented by the mortality rate for the time period (1980-2028). It was obtained from the Central Statistical Organization where the data were analyzed using time series according to the Box-Jenkins method and artificial neural networks using the program (Eviws.v9, SPSS, Zaitun.TS) and the most important conclusions and recommendations were reached, the most important of which proved the time series model using residues and values. The ARMA model has its advantage over the neural network model for predicting Iraqi mortality. So we recommend using this form.

Keywords: Box-Jenkins, Neural Networks, Prediction, ARIMA model, Death in Iraq
2020 MSC: 60G25, 62M45

1 Introduction

Statistical tools are one of the most important methods and methods used in scientific studies and research on the basis of identifying changes in phenomena. And prediction methods in time series are one of the important statistical tools through which changes in the values of the phenomenon can be measured over time, as well as predicting future changes depending on the previous values of the phenomenon [2, 5].

1.1 Research problem

The research problem lies in reaching the most efficient statistical methods that are used in forecasting and that give the best reliable results.

1.2 Research objective

The research aims to build a model of neural networks and time series models, predicting mortality, comparing estimated models and choosing the ideal model.

*Corresponding author

Email addresses: ashraf.m.shareef@nust.edu.iq (Ashraf Mohamed Shareefa), sarmad.j.Naser@nust.edu.iq (Sarmad Jaafar Nasera), ahmed.hashem@mpu.edu.iq (Ahmad M. Hashim)

1.3 Search tools

In order to achieve the goal of the research, we relied on the Box Jenkins method and the method of artificial neural networks, using Eviws.v9, SPSS, and Zaitun.TS.

1.4 Research importance

The importance of research lies in studying which method is more efficient.

1.5 Research assumes

1. There is no difference between the Box Jenkins method and the method of artificial neural networks in constructing nonlinear models.
2. There is no difference between the Box Jenkins method and the style of artificial neural networks in terms of accuracy and ease in predictions.

2 The theoretical aspect [2]

2.1 Time series and prediction

2.1.1 Definitions

Time series: A series of measurements recorded for one or more variables arranged on the basis of their occurrence in time and give values of a specific phenomenon, or it is a number of statistical observations that describe the change of the phenomenon under study.

The stability of the time series: The time series is stable from the statistical side if it does not contain a trend up or down in time, that is, it contains a variance and a fixed average, then if y_t represents the value of the phenomenon studied in time t then it is considered stable if the following conditions are met:

$$E(Y_t) = M \quad \text{for } t \in Z \quad (2.1)$$

$$\text{cov}(Y_t, Y_{t+k}) = \gamma_0 = \text{Fixed for } k = 0, \quad \gamma_k; k = 0, 1, 2, \dots \quad (2.2)$$

Such that, k represents the deceleration, which is the period that separates the previous values used in the model $k = n/4$.

Auto-correlation function (ACF): This function evaluates the correlations between the values shown for different periods i.e. measures the relationship between the series of residues, (i.e. the correlations between the errors of the series of values) and is mathematically defined in [1] as follows

$$\gamma_K = \frac{\text{cov}(k)}{\text{cov}(0)} = \frac{\gamma_k}{\gamma_0} \quad (2.3)$$

$$\gamma_k = \frac{\sum(y_t - \bar{y})}{n - k}; \quad \gamma_0 = \frac{\sum(y_t - \bar{y})^2}{n}, \quad \gamma_k = \gamma_{-k}; -1 \leq \gamma_k \leq 1. \quad (2.4)$$

That's where n is the sample size. The Auto-correlation function (ACF) can be relied upon to know whether the chain is stable or not if it is stable if the value of $\gamma_K = 0$ or tends to zero at the slowdown K . In general, the stable time series are distinguished from the unstable by the values of the self-correlation coefficients, as their value approaches the difference after the second period in relation to the static series, whereas the unstable have significant differences close to zero after the seventh and eighth period.

$$\phi_{cc} = \frac{\text{cov}[X_{(t)} - \hat{X}_{(t)}, X_{(t+c)} - \hat{X}_{(t+c)}]}{\sqrt{\text{var}[X_{(t)} - \hat{X}_{(t)}]} \sqrt{\text{var}[X_{(t+c)} - \hat{X}_{(t+c)}]}}. \quad (2.5)$$

Prediction: It is known that predicting what will happen in the future appears to be the dependence of the apparent in the past by using one of the known prediction models. Or knowing the behavior of the phenomenon in the future from its behavior in the past period.

2.2 Box and Jenkins models

Box and Jenkins models are considered one of the important statistical techniques in analyzing time series, as they are considered one of the most important methods used to build models in time series analysis and predicting their values in the future. They consider the extension of the regression method.

y_t : indicates the values of the visible time series $t = 1, 2, \dots, m$.

ϕ : Indicates the parameters of the Auto-regression model.

θ : Indicates the parameters of the mobile media model.

α : indicates the constant of the model.

e_t : indicates the random error in the time period

As can be illustrated types of model box Genghis time series as follows

2.2.1 Autoregressive model

A auto-regression (AR) model can be defined mathematically as follows

$$X_t = m + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \dots + \theta_p X_{t-p} \quad (2.6)$$

whereas, (P) represents the rank of the model for Auto-regression, and it is a expression of the previous values used in the model.

2.2.2 Moving Average

The Moving Media (MA) model can be defined mathematically as follows

$$X_t = Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \dots - \phi_q Z_{t-q}, \quad (2.7)$$

whereas, q represents the rank of the moving media model, which is the past values used in the model or in other form, representing the period of slowness of the model, and is symbolized by the symbol MA (q)

2.2.3 Mixed model

Sometimes the phenomenon can be represented by a mixed model of the auto-regression of the rank (p) and the moving media of the rank (q) . To obtain a model called the mixed model, which is symbolized by the symbol (p,q) ARMA can be represented by the following mathematical model.

$$X_t = m + \theta_1 X_{t-1} + \theta_2 X_{t-2} + \dots + \theta_p X_{t-p} + Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \dots - \phi_q Z_{t-q} \quad (2.8)$$

whereas (p,q) represents the rank of the model and is denoted by the symbol ARMA (p,q) . If the time series is unstable, it can be converted into a stable chain by taking the differences. For example, the first difference is given by the following formula.

$$D_t = X_t - X_{t-1} \quad (2.9)$$

where it represents the same previous model, but only an integrated word is added to the name of the model to indicate that this model was used to represent an unstable time series.

2.3 Box and Jenkins methodology stages [1]

We can summarize the construction of the Box-Genghis model in four main stages: identification, assessment, diagnosis, and prediction.

The first stage of the Box and Jenkins methodology in constructing time series models is one of the most difficult stage, in which at this stage several models are obtained to represent the time series, where this stage examines the stability of the time series as the stability condition is a prerequisite. There are several statistical tools It can be inferred from the stability of the time series, including the Q statistic, the time series propagation, unit root tests (simple Dickey Filler test and extended Dickey Filler test) and the self-correlation function fee as the chain is stable if all or most of the correlations fall within the confidence interval limit. Line The subsequent phase of the first stage after confirming the stability of the time series (or processing the time series to make it stable) is by determining the degree

of self-regression (p) and the rank of the moving media (q), which is through drawing the functions of self-correlation (ACF and partial self-correlation (PACF) The partial self-correlation function determines the self-regression rank, as it represents the number of decelerations outside the boundary of the confidence interval (i.e., not equal to zero). About the confidence period. At the end of the first stage, more than one ARIMA model will be nominated, the differentiation between them will be made, and the optimal model will be chosen based on several statistical criteria, including the AIC and Ashuratz (SIC) standards. They can be defined as follows:

1. Akaike standard (AIC): It is calculated from the following equation

$$AIC(C) = n \ln(\sigma_u^2) + 2c \quad (2.10)$$

The model with the lowest AIC value is chosen.

2. Schwartz Standard (SIC): It is calculated from the following formula:

$$SIC = n \ln(\sigma_u^2) + k \ln(n) \quad (2.11)$$

Since m is the number of model parameters and n is the number of time series observations, mse is the average error squared of the model. It is the best model that has the lowest amount of two criteria.

After specifying the ARIMA model from the first stage, the stage of estimating the model emerges as a second stage and then moving to the third stage of diagnosing the model to ensure its validity before prediction, through testing the remainder of the model to ensure that it has the property of white noise and its independence from each other. And finally after making sure Model validity in the third stage The model is approved for forecasting future time periods.

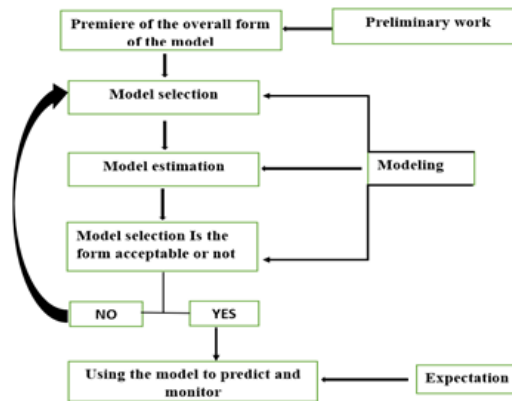


Figure 1: Box and Jenkins implementation stages in time series analysis

2.4 The concept of artificial intelligence [4]

One of the most important ways of artificial intelligence is ANN (Artificial Neural Networks). Its idea revolves around simulating the ability of the human mind to recognize patterns and distinguish things using a computer through following the self-learning process that occurs in the mind, which is It is to benefit from previous experiences in order to reach the best results in the future.

Where each operating unit has one or more input path whose mission is limited to transfer information from the outside world to the operating unit, which in turn performs a simple collection process, then converts the information by an activating function known as the conversion function, after which the information is transferred as output through the output path. Mathematically, each neuron receives a number of input signals (X_1, X_2, \dots, X_n) Input Signals, which are similar to Dendrites, then all inputs are multiplied by weights (W_1, W_2, \dots, W_n) which is the primary means of long-term artificial neural network memory and expresses the degree of importance of the inputs, The summation is obtained using the $F(Y)$ Transfer function to obtain the Output Signals (Y).

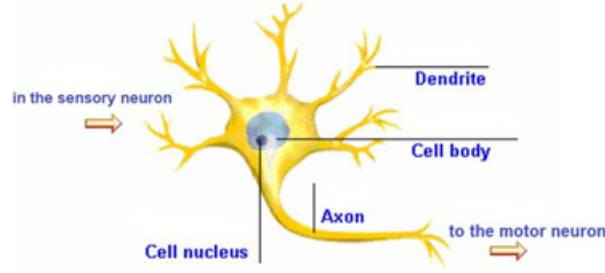


Figure 2: drawing the neural network (Introduction to Artificial Neural Networks)

3 General structure of the network Typical Architectures [3]

The process of organizing neurons in layers and how they communicate between these cells to form a network is called the structure of the architectures. In general, it is possible to divide the structure of the artificial neural network into three main types:

1. Single - Layer Feed forward Network.
2. Multi - Layer Feed forward Network
3. Multi - Layer Recurrent Network We will explain each type as follows

3.1 Front layer single layer mesh

It is the simplest structure of artificial neural networks, and the most common type of it is called Perceptron, which is the simplest form of artificial neural networks. The outputs are direct, as they are learned through supervisory education and through that network. The learning process is accomplished by the processing units taking several main steps: The first step: random selection of the primary values of weights W_1, W_2, \dots, W_n and the threshold value θ is in the range $[0.5, -0.5]$. The second step: the activation process: after each unit of Neuron receives many input signals, and calculating the weighted sum of these inputs used, which is usually done using the following Summation Function.

$$X = \sum_{i=1}^n w_i x_i, \quad (3.1)$$

where, X is the net weighted input of the neuron, w is the relative weights of the connection nodes between layers, x_i is input value, n is the number of neuron inputs, and (X_1, X_2, \dots, X_n) Perceptron is activated by applying input.

The desired outputs $T(K)$, and the calculation of the actual outputs on repeat k are as follows:

$$y(k) = \text{step} \left[\sum_{i=1}^n x_i(k) w_i(k) - \theta \right]. \quad (3.2)$$

This type of conversion (activation) function is called a step function. It should be noted that, besides the step function, there are many activation functions, but it was found that only a few of them have practical applications and the most important of these are the following

1. Sign Function
2. Step Function
3. Linear Function
4. Function Sigmoid and Usually Sign and Step Function of Neurons in Decision Tasks Classification and Pattern Recognition

The sigmoid function is the most commonly used conversion state because of its simplicity of discrimination and ease in calculating the slope. The linear function can be used in some applications of time series, and in most cases the conversion function is combined from a linear combination.

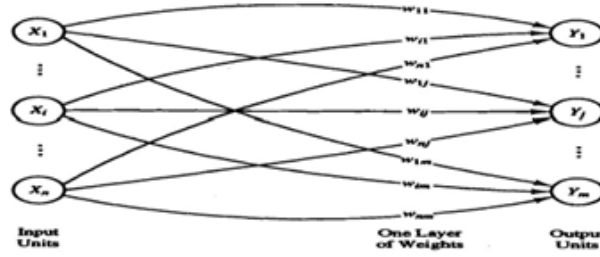


Figure 3: shows the shape of the single layer

3.2 Artificial neural network architecture

It is divided into two parts

1. A single-layer network

It contains a single layer of processors or an input layer. It only passes the inputs and it consists of only one link.

2. Multilayer network

It contains more than one layer of processors, and some layers may be hidden, and the function of these layers is to increase the strength of the neural network and improve its performance, as the layer may be composed of only one neuron, and the hidden layer increases the network's ability to process data, but it is faulty that it slows down the process Treated.

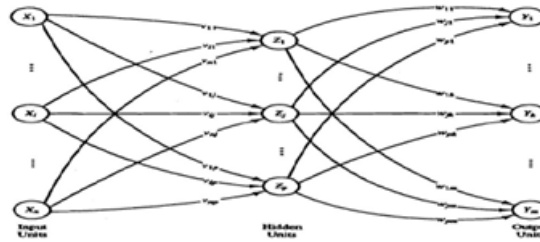


Figure 4: shows a multi-layer shape

3.3 Front Layer Mesh Multilayer [3, 4]

A multi-layered network consists of one or more layers (or levels) of nodes called Hidden Layers found between the Input Layers and Output layers, and in front-feed networks the input signals are spread in the forward direction from layer to layer, nor Going back is allowed, and this type of network can solve many complex problems that a single-layer network cannot solve, but training may take longer.

The mathematical elements in the hidden layer perform the weighted collection process and send an output signal (alert patterns) to the output layer, which also performs the weighted collection process and specifies the output patterns for the whole network. Also, by increasing the number of processing units inside the hidden layer or adding more than one hidden layer in the network, the network is allowed to deal with many complex functions and the precise approximation of continuous functions, and this is what is stipulated by the theory (Cybenko) As for the education process, it continues in the same way in the single-layer network, where weights occur using Back propagation Learning, and before this update there are two consecutive steps, the first being Forward and Backward, in the front step the network outputs are calculated from the data entered, and a comparison of these Outputs with target outputs by error calculation. In the regression step, the network adjusts the relative weights in order to reduce the error, and the process that includes the previous two phases is called the epoch, and this cycle is repeated until we reach the least sum of squares of the error.

3.4 Multilayer Network with Feedback [4]

It is the second type of multi-layer neural networks, and this type contains at least one background loop (closed loop of feedback loop), whereas unlike the multi-layer network with front feeding, the error between the network outputs and the real outputs is calculated but is fed The error is in the reverse direction (opposite the input direction) of the network to adjust the weights and thus the error is reduced, and this procedure is repeated until the error reaches the lowest possible value.

3.5 Use of neural networks in the prediction process [3, 4]

This method is usually used when there is a time series containing previous observations and we want to predict future values. Here, the aggregation function itself is used in addition to the logistic transfer function until we obtain new outputs that are compared with the desired values (to calculate the error $E = X_i - Y_I$) through the difference between them. Then the weight is corrected and adjusted through the learning process that is done on the network and is done automatically. Often, through the following equation, where alpha is the amount of learning, and it is imposed initially at a low level, whereas beta expresses the difference between calculated and desired values.

$$W_i(Final) = W_i + \alpha\beta x_i \quad (E = X_i - Y_I).$$

Then all of the above is repeated until we reach the desired goal, the acceptable error rate, or the end of the specified number of attempts.

4 Results and discussion

The data that we are analyzing and estimating the ARIMA model appropriate for predicting mortality with a time series of mortality in Iraq for the period 1980-2018, which represents the mortality per thousand people shown in Table 1 below. In order to study the characteristics of the time series and estimate the ARIMA model appropriate for prediction according to For the box and Jenkins method, my statistical software, EvIEWS v.9, was summarized as follows:

4.1 Time Series stationary

Before defining the model, we must demonstrate the stability of the time series of the research sample as follows:

4.1.1 Spread the time series

Figure 5 shows the graph of the time series representing the mortality rate in Iraq for the period 1980-2018. We note that the time series of data in 1980 was increasing due to wars until the world of 1990 and then gradually decreased until 2018.

Table 1: Mortality in Iraq for the period 1980-2018

year	death rate	year	death rate	year	death rate	year	death rate
1980	9.86	1990	7.06	2000	5.63	2010	5.51
1981	9.96	1991	6.79	2001	5.62	2011	5.4
1982	9.97	1992	6.56	2002	5.63	2012	5.28
1983	9.85	1993	6.36	2003	5.65	2013	5.16
1984	9.59	1994	6.18	2004	5.69	2014	5.06
1985	9.21	1995	6.03	2005	5.72	2015	4.97
1986	8.75	1996	5.9	2006	5.73	2016	4.89
1987	8.26	1997	5.8	2007	5.72	2017	4.83
1988	7.8	1998	5.72	2008	5.68	2018	4.78
1989	7.4	1999	5.66	2009	5.61		

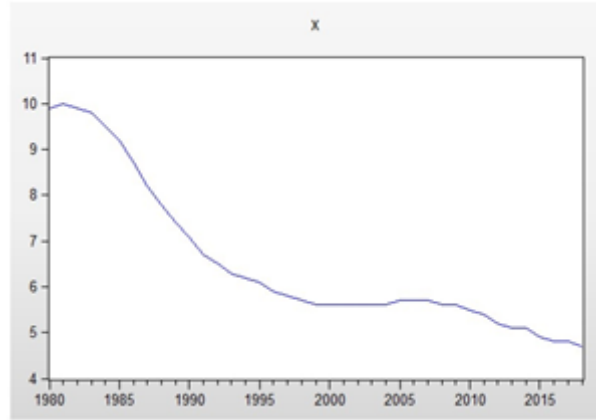


Figure 5: graphical representation of mortality for the period 1980-2018 in Iraq

Table 2: Extended Dickey-Filler Test

Model	t -static	p
Intercept	-4.1006	0.0029
Trend and Intercept	-4.2696	0.0092
N0ne	-2.0608	0.0392

4.1.2 Augmented Dikey-Fuller

The unit root of the extended cockpipe test was tested and the test results were as in Table 2

We note that the extensive Dickey Fuller test in Table 2 confirms that the time series of mortality are stable in mean and variance as We note that its test statistic $t = -4.1006$ has a significant level $p = 0.00029$ which is less than $\alpha = 0.05$ in Intercept and this indicates acceptance of the alternative hypothesis.

The test statistic $t = -4.2696$ has a significant level $p = 0.0092$ which is less than $\alpha = 0.05$ in Intercept Trend and this indicates acceptance of the alternative hypothesis, We also note the test statistic $t = -2.0608$ with a significant level $p = 0.0392$ which is less than 0.05 in None and this indicates acceptance of the alternative hypothesis, and this indicates that the time series is stable.

As we observe from Figure 6 that the time series is stable for the original data. We also note that Figure 5 confirms the stability of the time series, as we find all the factors of the auto-correlation of the series fall within the confidence period. Before starting the study of the stable chain behavior, the probability distribution of the chain must be studied, and this is what we explain in Figure 6, which indicates that the stable chain is subject to a normal

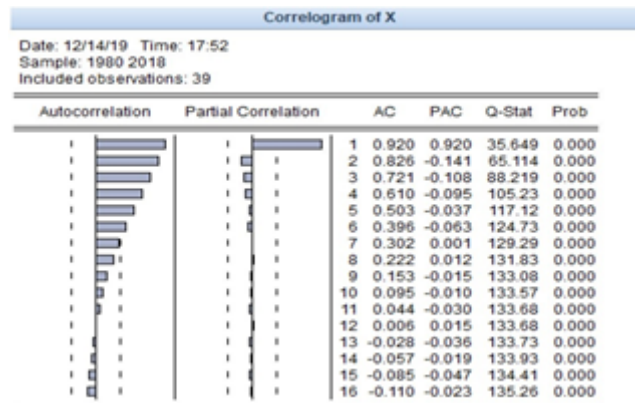


Figure 6: partial self-link coefficients of the stable chain

Distribution, as the amount of statistics (Jarque-Bera)=7.641636 with a significant level $p = 0.021910$ which is less than $\alpha = 0.05$. The value of the flatness coefficient (Kurtosis) (2.764314) is less than (3) indicating the symmetry of the probability distribution.

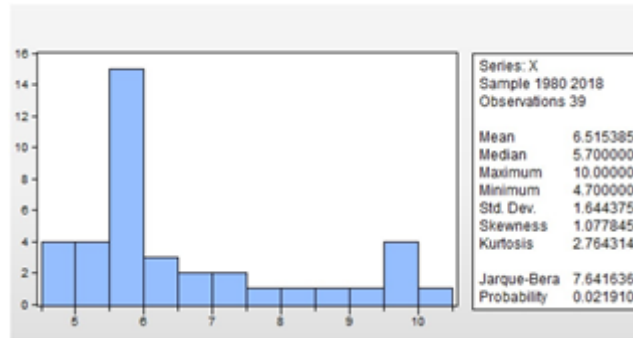


Figure 7: Time series histogram

4.2 ARMA Model Identification and Estimation

Returning to Figure 5, which represents the self-correlation coefficient and the partial self-correlation of the stable chain, we notice that the coefficient of self-correlation at the first decelerations differs significantly from zero (falls outside the confidence interval) and thus we infer that the rank of the moving media ($p = 2$). Partial self-correlation coefficients differ significantly from zero after the first deceleration (located outside the confidence interval), when we find that the self-regression rank is ($p = 2,1$).

Now we propose a set of ARMA models (p,q) and prefer them among them to choose the best one based on the AIC self-information standard and the BIC standard to determine the appropriate model as shown in Table 3.

Table 3: Estimation of the proposed ARMA models for x series data

Rank Model		Parameter Model				Year
AR(1)	AR(2)	MA(1)	MA(2)	AIC	BIC	
0.991877 p=0.0000		0.697385 p=0.0000		-0.646787	-0.476166	
1.945885 p=0.0000		-0.289991 p=0.2516		-1.518362	-1.305085	
1.937771	-0.955139	-0.448432 p=0.014	0.281292	-1.542350	-1.286418	
0.993510 p=0.0000		0.756910 p=0.0000	0.685153 p=0.0001	-1.007068	-0.793791 p=0.097	

We notice from Table 3 that the ARMA model (1,1) has the lowest AIC and BIC standards -0.646787 and -0.476166 respectively, and therefore the appropriate model for representing time series X data in which the AR parameter (1) was significant which Its value reached (0.94877) and its level of significance ($P = 0.000$) which is less than the level of significance specified for the test $\alpha = 0.05$ and the parameter MA (1) had a value that reached (0.697385) and its level of significance ($P = 0.000$) It is less than the significance level specified for the test $\alpha = 0.05$. The fixed limit was significant because the corresponding level of significance was (0.0097) which is less than $\alpha = 0.05$, meaning that its inclusion in the model is important.

From Table 4 shows the training group (79.5%) of the total sample size as well as the number of training group views of 31 views, while the test group (20.5% of the sample size and number of observations 8), as for excluded observations There is no view excluded for the current network.

From Table 5 in the first section in it for the Input Layer, which consists of a unit, one independent variable (Covariates), either in the second section of the table, which is related to the hidden layer (Hidden Layer) and consists of two units and the activation function (Hyperbolic tangent), and the last section of the table that shows the Output

Table 4: represents the total training

		N	Percent
Sample	Training	31	79.5%
	Testing	8	20.5%
Valid		39	100.0%
Excluded		2	
Total		41	

Table 5: Neural Network Information

	Covariates	1	x
Input Layer	Number of Units	(Excluding the bias unit)	1
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers	1	
	Number of Units in Hidden Layer	1 (Excluding the bias unit)	2
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	y^*
	Number of Units		2
	Activation Function		Softmax
	Error Function		Cross-entropy

Layer information and which consists of a Dependent Variables one y^* with a number of units and an activation function.

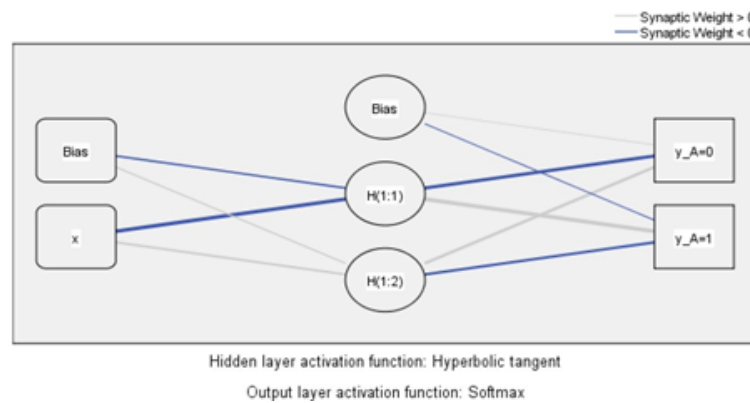


Figure 8: Structure of the Artificial Neural Network

Table 6: Network Summary (Dependent Variable: y^*)

Training	Cross Entropy Error	0.017
	Percent Incorrect Predictions	0.0%
	Stopping Rule Used	Training error ratio criterion (0.001) achieved
		0.0000
Testing		8.170E-8
	Percent Incorrect Predictions	0.0%

From Table 6, we notice in its first section the results related to the Training Group, which had an error value of 0.017 and a false prediction rate of 0.0%, while the other section of the table related to the results of the test sample, which was with an error of 8.170E-8 and a false prediction ratio of 0.0.

Table 7: Estimation of model parameters

Predictor		Predicted			
		Hidden Layer 1		Output Layer	
		H(1:1)	H(1:2)	$[y_A = 0]$	$[y_A = 1]$
Input Layer x	Bias	-4.282	3.232		
		-5.684	4.526		
Hidden Layer 1	Bias			1.314	-1.603
	H(1:1)			-5.419	5.717
	H(1:2)			4.997	-4.925

4.3 Predicted

From Table 7 the parameters of the model or the so-called model weights in its first section represent the values of the weights between the Input Layer and the Hidden Layer, and in the other division the values of the weights between the Hidden Layer and the Output Layer.

Table 8: Estimated values

year	ANN	ARMA(1,1)	year	ANN	ARMA(1,1)
1980	-	-	2000	5.6607	9.6368061
1981	-	9.8927061	2001	5.5182	9.624017
1982	-	9.878606	2002	5.4184	9.6112935
1983	-	9.8645782	2003	5.3353	9.5986352
1984	-	9.8506223	2004	5.2985	9.5860417
1985	-	9.8367378	2005	5.2462	9.5735127
1986	-	9.8229245	2006	5.2257	9.5610479
1987	-	9.8091819	2007	5.1961	9.5486469
1988	-	9.7955098	2008	5.1632	9.5363095
1989	-	9.7819077	2009	5.1326	9.5240353
1990	-	9.7683753	2010	5.09	9.511824
1991	-	9.7549122	2011	5.0366	9.4996753
1992	8.0472	9.7415181	2012	5.0231	9.4875887
1993	7.6414	9.7281926	2013	5.0005	9.4755641
1994	7.2022	9.7149353	2014	4.9866	9.4636011
1995	6.9443	9.701746	2015	4.9638	9.4516994
1996	6.6986	9.6886243	2016	4.9648	9.4398587
1997	6.4755	9.6755698	2017	4.9915	9.4280786
1998	6.1781	9.6625821	2018	4.9899	9.4163589
1999	5.8738	9.649661			

From Table 8 it is shown that the estimated prediction values using the artificial neural networks (ANN) method are better than the estimated predictive values of the ARMA(1,1) model.

Table 9: Comparison of Box Genghis Methodological Predictions with Neural Networks

year	Values to predict a model ARMA	Values to predict a model ANN
2019	4.7334	5.2457
2020	4.6922	5.2275
2021	4.6562	5.3609
2022	4.6252	5.3514
2023	4.5992	5.4039
2024	4.5778	5.3042
2025	4.5610	5.2486
2026	4.5486	5.2727
2027	4.5405	5.1625
2028	4.5365	5.1721

From Table 9 the predictive values show the following years for both models for a period of ten years

Table 10: Comparison of Box Genghis Methodological Predictions and Neural Networks

	ANN	ARMA(1.0,1)
MAE	0.21978	3.22765
MSE	0.06517	3.520694
ERROR	0.86702	1.644375

We note from Table 10 that the criteria for the Box-Jenkins methodology are high and this shows that the predictions of artificial neural networks are the best.

5 Conclusions and recommendations

1. The Neural Networks model proved its preference from the time series model using the residuals and the estimated values. ARMA model for predicting Iraqi mortality numbers. So we recommend using this form.
2. The research concluded that the best model for predicting the numbers of Iraqi deaths is the model of artificial neural networks.
3. The research recommends studying the most important variables that affect the mortality rate while trying to estimate other multivariate models that take this into consideration.
4. Neural networks inability to predict the years (1981-1991).

References

- [1] K.K.S. Al-Satori and B.M.A. Al-Hiti, *Using ARIMA models to predict the money supply for Qatar*, Anbar Univ. J. Econ. Administrat. Sci. **35** (2010), 58–83.
- [2] M. Hajji, *International trade in technology*, J. Econ. **5** (1975), no. 57.
- [3] H.B.A.-A. Mazouzi and A. Al-Mu'tar, *Prediction of the use of artificial neural networks*, Doctoral diss. Ahmed Deraya-Adrar University, 2018.
- [4] S.M.A. Mustafa, *Using ARIMA models and artificial neural networks in predicting the Egyptian stock exchange index EGX30*, J. Financ. Commercial Res. **18** (2017), no. 1, 392–416.
- [5] S.A.-K. Tumo, *Using time series download for predicting people with malignant diseases in Anbar governorate*, Anbar Univ. J. Econ. Sci. (2012), no. 8.