Improvement of Neural Networks Artificial Output

Dr. Marwan Abdul Hameed Ashour[1], Dr. Nada Ismaeel Jabbouri[2]

[1,2]College of Administration and Economics, University of Baghdad, Iraq

Abstract: *The purpose of this research is to discuss the improvement Artificial Neural Network (ANN) outputs by increase the number of nodes in the neural network layer inputs of quantitative and qualitative terms to reach the desired results. The stage of determining the inputs of the network is the most important stage in the neural network to enable the compatibility of the model to reach of the desired results, activate the property of learning, and self-adaptation with any model owned by artificial neural networks. In some times, the neural network is not able to reach the desired goal not because of the lack of methodology of artificial neural networks, but because the input is not properly defined. Therefore, the neural network cannot identify the correct model. Simulation was adopted as an experimental method of research. As well as, the practical side was used to demonstrate the quality and efficiency of neural networks. The empirical and practical results were proving that the systematic increase in the number of nodes of the artificial neural network (ANN) inputs. leads to the improvement of the neural networks outputs and the obtain of outputs corresponding to inputs. Then, reach to the desired results.*

Keywords: Back Propagation Network Error, improved outputs, Box-Jenkins models, Time series of seasonality

1. Introduction

The main objective of building the "modeling" model is to analyze the components of the current system's behavior to identify the influencing factors and the circumstances surrounding it in order to improve its performance. In addition, the aims to build the model and determine the ideal formula for the future system.The process of building a model for any problem or data series is one of the most difficult important problem faced by researchers, this is because quantitative methods are unable to solve or estimate the parameters of any model or problem unless the process of model construction represents a precise representation of the main features of the problem in order to achieve the desired results.Research in the area of artificial neural networks (ANN) is ongoing to identify its quality and efficiency in data processing and analysis.Research in the area of artificial neural networks (ANN) is ongoing to identify its quality and efficiency in data processing and analysis.The research is seeking to study the effect of modeling in improving the output of (ANN) by identifying inputs that describe the components of the system behavior of the problem and the benefit of the learning process and training of the neural network with any model owned by ANN.Most researchers rely in the formulation of time series data on the Box-Jenkins tools in the process of diagnosing the model and determining the network input.Most researchers rely in the formulation of time series data on the Box-Jenkins tools in the process of diagnosing the model and determining the network input leading to results similar to the results of the Box-Jenkins model. Thus, maximal utilization of the characteristics of artificial neural networks is not done, especially if the time series have irregular, while the Box-Jenkins models assume the regularity of the time series.

In this research, the researchers adopted on the process of modeling the neural network input, that is according to the " modifyed method" based on the cause-effect relationship to describe the behavior of the time series.The reverse spread network for errors has been taken as a model for artificial neural networks (ANN) because it is one of the most common networks in data analysis.The main idea of this research came from the accumulated experience of a studies group related to artificial neural networks.On this basis, the research hypothesis was formulated, which states (that the increase of the number and quality of network inputs will improve the outputs of the neural network, and then achieve the desired results.To prove the hypothesis, the simulation was used in this study as a way to generate a time series.In addition, the applied side has been used to electrical power data sold to the city of Baghdad to know the researcher about previous field research behavior for this series, being irregular in the length of time period for each season because of the nature of the work causing the overlap or interconnection between the month.

The first method is represented by the adoption of the process of modeling "network input selection" according to the modelsBox – Jenkins, the most common method. The second method is the adoption of a time series model, "Determining the network input of the time series" according to the concept of causal models "the proposed method".The results of methods adopted in this study were compared based on statistical criteria of distinctive to selection of the most efficient method.Furthermore, the neural networks to predict future values for the time series trace phenomenon that are irregular in the length of time for each season have been employed.Whereas, the prediction methods for regular seasonal models assume the regularity of the length of each season for each of these chains.It should be noted, that the background of this study is not specific in certain sources, where sources and research are rare in this field for employment of the conversion function for financial forecasting or cash flow.The sources mentioned in this research are those related to the quantitative methods were used in the research.On this basis, the researchers faced great difficulty of following the subject in published literature. The background of this research is depends on the knowledge of institution and the experience gained by the researchers.

2. Method

1) Artificial Neural Networks (ANN)

Artificial neural networks are defined as an information processing system based on simple mathematical models that have specific performance characteristics in a way that simulates the neural network. Artificial neural networks are computational techniques designed to simulate the way the human brain to performs a task, Through a massive processing distributed on parallel, consisting of simple processing units. These units are only computational elements called Nodes or Neurons, which have a neurological characteristic.They store practical knowledge and experimental information to make them available to the user by adjusting the weights. These units are work to store practical knowledge and experimental information to make them available to the user by adjusting the weights (Nasser,A & Obeidi,M. AH. (2003); Alexander, I & Galushkin, G. (2012); and Allende, H & Salas, R.(2012).
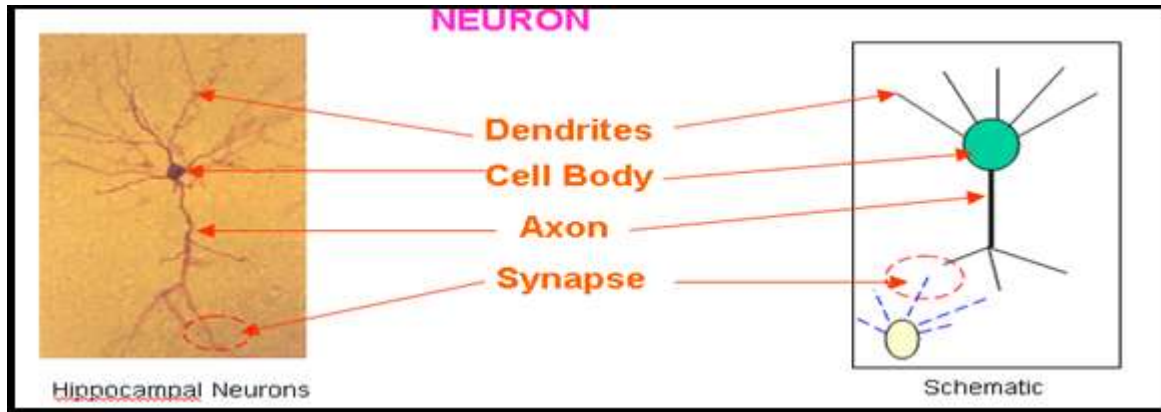
The following diagram shows the biological neuron diagram.

Figure 1: Represents the biological neuron diagram.
Source: Nasser, Abdul Majeed Hamza. Obeidi, Marwan Abdel Hamid (2003). Use of neural networks to predict future time series values that are irregular in seasonal length. Iraqi Journal of Statistical Sciences. Volume 3., 6 issue

Whereas:
- Dendrites: Represents receiving "data" signals
- Cell Body: Represents "data" signals operations
- Axon: Represents transferred information manufactured to other neurons
- Synapse: Represents the junction between the end of Axon and Dendrites for other neurons.

Artificial neural networks are nonlinear models. The artificial neural network has suggested an actress of sensual perception that represents a simulation of the Biological neuron in 1958.Artificial neural networks are nonlinear models. The artificial neural network has suggested an actress of sensory perception that represents a simulation of the biological neuron in 1958. The sensory perception is consists of neurons, nodes or units that representing the input, simulating incoming signals into the biological neuron. The signals between the nodes pass through the connecting lines and attach each linked line to a certain weight,the signals entering in the node are multiplied by these weights, and then the balanced inputs are assembled into the nodes or neurons. The outputs of each node are then processed by a non-linear function with a threshold called the Activation Function (Allende,H&Salas,R.(2012); Palit, A& Popovic,D.(2005);and Qing,Cao & Thompson, Mark.(2012).

Attention to neural networks were delayed until the 1970 due to the limited use of single-layer neural networks, but the discovery of the idea of reverse propagation of the error of multilayered network training by many researchers in the late 1980, played a key role in re-emerging neural networks as a tool to solve many problems on a large scale.Although single-layer networks are very limited in "corresponding" processes, but they can learn. As well as, the multi-layer neural networks with a "hidden layer or more" can learn any continuous corresponding process due an optional accuracy.The neural network defined as a Gradient Descent method to find the minimum value of the total error box for the outputs value calculated by the neural network. The fault-prone neural network is based on the concept of network training based on error size. The weights between the layers occur as much as the weight contribution in the composition of the error and the network continues to update the weights until the reach of optimal weights that achieve the best fit of the model. The following figure illustrates the general methodology of artificial neural networks ANN.
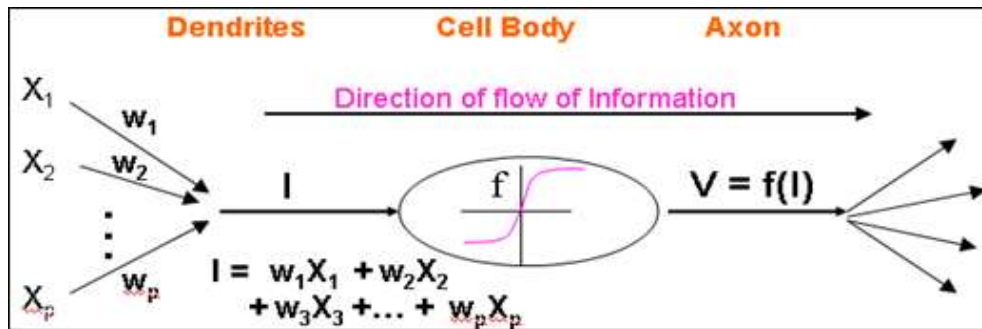
Figure 2: Shows the general methodology of artificial neural networks
Source: Allende, H and Salas, R.(2012). Artificial Neural Networks in Time Series Forecasting.
International Journal of Forecasting 8:338-358

Whereas:
Delivers Inputs X1, X2 …. Xp of neurons or other environment.
Inputs fed during the links to 'weights'
Inputs feed by related to 'weights'
Total contribution = total input weight of all sources
The "activation function" converts the input to outputs.
The output goes to the neurons or other environment

2) Engineering Structure of Neural Networks:

The neural network is consists of at least three levels and each level includes a number of nodes. The first level represents inputs, the second represents the outputs and the other level represents the hidden layer.This layer is exists between the input level and the output level, each level contains a number of nodes. The nodes are related at the three levels with the connecting lines and each line connecting to a given weight. The following figure illustrates the engineering structure of artificial neural networks (Phili,K& Deligiorgi,D.(2012); Ratnadip,RK.(2011); andNasser, A. M. &Obeidi, M. AH. (2003).
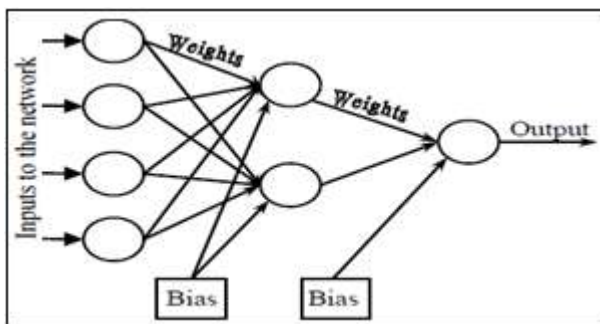

Figure 3: Illustrates the engineering structure of artificial neural networks.
Source: Palit,A & Popovic,D.(2005).*Computational Intellgence in Time Series Forecasting*
Springer-Verlag London UK., PP.972.

3) Back Propagation Network
The training of the neural network using back propagation involves three stages as following
(Alexander,I&Galushkin,G.(2012); and
Allende,HandSalas,R.(2012).
a) Forward propagation phase of error.
b) Back propagation phase of error.
c) Phase of network weights generating

Forward propagation phase of error: in this phase, the inputs signal is propagated in each of node of the hidden layer nodes, and then activation value for each node of the hidden layer nodes for this signal, and then these nodes send their signals to each node of the output layer nodes. Calculate of activation value for each node of the outputs layer nodes to form the network response for the inputs sample. In the training phase, each node of the output layer by comparing their calculated activations with the actual output value to determine the error value of that node.Depending on the error value, the error correction coefficient is calculated δK, The error correction coefficient is used δK to distribute the error on the output layer nodes to be returned to each node in the previous layer. In addition, the coefficient of updating weights between the output layer and the hidden layer was used. Similarly, the error correction coefficient is calculated δJ for each node of the hidden layer nodes, the coefficient of updating weights ‹between the input layer and the hidden layer was used. After determining all the factors to correct the error δ the weights for all layers are synthesized at the same time (Ratnadip,RK.(2011);Riahy,GH& Abedi,M., (2008); and Sana,Hamra&Gabr,Mahmoud. (2003).

4) Neural network algorithm with back propagation error
The algorithm or methodology of this network can be summarized at the following steps:
- Generate initial values of weights "from one of the statistical distributions."
- Each node in the inputs layer is receiving its input signal, and then sends it to all nodes of the hidden layer (Nasser, A. M. &Obeidi, M. Hamid (2003); Phili,K& Deligiorgi,D.(2012); and Ratnadip,RK.(2011); and Stetter,Olav.(2012).
- Each node in the hidden layer collects the values and its input signals according to the following equation:

$$h_j = f(\Sigma w_{ij}x_i - \theta_j) \qquad -------- \quad (1)$$

- An application continues activation to estimate the output of the hidden layer, and sends the activation values to all nodes in the output.
- Each node in the output layer collects its balanced input signals under the following equation.

$$_K = f(\Sigma w_{jk}h_j - \theta_k)Y \text{ -------- (2)}$$

A continued activation application to estimate the outputs of the output layer.

- Comparison of the neural network outputs with the real values to estimate the error according to the following equation:,

$$\delta_K = (t_K - y_K).f \ (v) \qquad -------- \quad (3)$$

- Calculate the change in error size Δ WJK and under the following equation:

$$\Delta \ W_{JK} = \alpha.\delta_K.H_j \ldots\ldots\ldots(4)$$

- Each node in the hidden layer has collected the inputs signals weighted δ according to the following equation:

$$\Delta_J = \Sigma \ \delta_K W_{JK} -------(1-5)$$

Then, this value is multiplied in the activation sequence of the account δ j. Hence, the change in error size is calculated Δv_{ij} through the following equation:

$$\Delta v_{ij} = \alpha.\delta_j.x_i \ ----------- \ (2\text{-} 5)$$

- Update weights of eachnodo in the outputs ayer according of the following equation:

$$W_{JK}(new) = W_{JK}(old) + \Delta \ W_{JK} \ \text{\tiny ....}(6)$$

- Update the weights for each node in the hidden layer, according to the following equation

$$Vij(new) = Vij(old) + \Delta vij \ -------- \ (7)$$

The network is continues to update the weights of the "learning process and training" until the ideal weights are obtained. Thus, the obtaining the desired outputs and finding the best fit for the model in question (Qing,Cao & Thompson, Mark.(2012); and Allende, H & Salas, R.(2012).

5) Symbols used in the neural network

X_i represents the input to the network
Wrepresents weights between levels
Θrepresents the Threshold)
f represent is Activation Function, and calculate the activation function according to the following equation:

$$f(v) = 1/ \exp(-v) \ ....(8)$$

The value of the function f (v) is between (-1,1), and v_k represents the linear synthesis output
Y_KRepresent network outputs
tRepresent actual or desired outputs
αLearning rate

6) The time series and Box-Jenkins models of the seasonal mixed:

The time series is defined as a sequence of values observation of a specific random phenomenon associated with time. There are some statistical criteria that are used to describe the quality of the time series, thus facilitating their modeling: these criteria are (Sana,H. &Gabr,M.(2003); and Wei,William.(1990).

- Self-correlation function (ACF): Is indicates the relationship or correlation of the values of the contiguous series and the value of the self-correlation coefficient between (-1،1).
- Partial Self-Correlation Function (PACF): The partial self-correlation function measures the partial effect of adding the delayed values of variable.

PACF: The partial self-correlation function measures the partial effect of adding the delayed values of a variable. Partial self-correlation function (PACF) coefficients can be obtained from the self-series regression equation the research topic. The best predictor of the value of Xt in time t+L which follows the seasonal mixed model:

$$\Phi p(Bs) \ \Phi p \ (B)(1\text{-}B)d \ (1\text{-}Bs)D \ Xt = \Theta Q \ (Bs) \ \theta q(B) \ at \ldots\ldots (9)$$

$$Xt \ (L) = E \ [Xt_+ LXt \ , \ Xt+1 \ , \ \ldots\ldots.] \quad \ldots\ldots(10)$$

When model (9) is stable, it can be rewritten in terms of random errors and as follows:

$$X_t = \Psi \ (B) \ a_t = a_t + \Psi_1 a_{t-1} + \Psi_2 \ a_{t-2} +$$
$$= \Sigma \ \Psi_J \ a_{t-J} \ , \ \Psi_0 = 1 \ \ldots\ldots \ (11)$$

Therefore, the future error prediction can be written as:

$$e_t(L) = X_{t+L} - X^\wedge_t(L) =$$

Where S the length of the seasonal period is "conditional expectation" where:

$$\text{And be:} \quad E[e_t(L)] = 011 \Psi_J \ a_{t+L-J} \ \ldots\ldots\ldots \ (3.4)$$

The rate of error squared for future prediction is:

$$\text{FMSE} \ [X^\wedge_t(L)] = E[X^\wedge_t(L) - \ X^\wedge_{t+L}]^2 = E[e^2_t(L) \quad = \sigma^2_a$$
$$\Sigma \Psi^2_J \ldots\ldots \quad (12)$$

7) Statistical Standards of differentiation:

For the purposes of comparing the results of approved methods, statistical criteria can be adopted (Phili,K&Deligiorgi,D.(2012); andStetter,Olav.(2012):

1. Mean absolute error:

$$\text{As:MAE} = \frac{\Sigma |et|}{n} \ldots\ldots\ldots \ (13)$$

2. Mean square error:

$$\text{As:MSE} = \frac{\Sigma \ et^2}{n} \ldots\ldots..\ldots \ (14)$$

3. Mean absolute percentage:

$$\text{As:} \ \ \text{MASPE} = \frac{\Sigma \left|\frac{et}{Xt}\right|}{n} *100 \ \ldots\ldots.. \ (15)$$

4. Root mean squared error:

$$\text{As} \ \ \ \text{RMSE} = \sqrt[2]{\frac{\Sigma \ et}{n}} \qquad \ldots\ldots\ldots \ (16)$$

As:Xt: represents real values
X ^ t: represent the predicted values
Et:: The error represents "the difference between real and predicted values"
n: represents the sample size

5. Coefficient of determination

$$\text{As:} \quad R^2 = SSR/SST \ \ldots\ldots\ldots(17)$$

Whereas:
SST = Total sum of square = $\Sigma \ (y(i) - \mu(y))^2$;for i=1,2, ..., n
SSR = Regression sum of square = $\Sigma \ (\hat{Y}(i) - \mu(y))^2$;for i=1,2, ..., n

The method that is gives a smallest value to statistical standards in equations (13-16) and the largest value of the statistical standard in equation (17) is the best method.

Artificial Neural Networks and Time Series:
Neural networks are becoming increasingly important in processing and analyzing time series for its ability to self-learning and training. All hypotheses of the researchers are based on one or more of the style of the removal of the series to determine the input of the artificial neural network of the time series designed to be solved using artificial neural networks.

Practical part:
Simulation is one of the important practical methods used to determine the behavior of a particular system. Simulation is involves designing a particular model or models based on real models. In addition, simulation is using in the absence of real experiments or lacks of mathematical formulas that represent the problem to be solve, and it is a method adopted in the fields of multiple sciences.

In this aspect, the application of the model has been through the simulation method to monitor and detect errors, processing and gaining experience. In many cases, simulations are using to observe the model before it is applied in practice. Simulations help with the following operations:

- Simulation helps to observe changes in the formulation of the problem if implemented in practice, leading to the development of a system model that meets the required purpose.

- Simulation is helps to study the system and see the results clearly, which facilitates the development of the system.
- Simulation helps to gain experience when designing a simulation model in an electronic computer. The experience gained is during the process of design and study of simulation that is leads to important suggestions and conclusions to change the system, where these changes can be tested by using the simulation method before their practical application.
- Simulations help to obtain information and conclusions for future situations that do not know their nature through repeated experiences for those situations.
- Simulation is using to know the change of system results when system changes occur.

The generation of a time series is following the second-order self-regression model AR (2) with a sample size of 200 and the value of a model parameter equal to($\Phi 2=0.25$‹$1=0.7\Phi$). The variation is equal to $\sigma\_e^2=15$ during the generation of random data by the Rand function that follows the regular distribution U(0, 1). The generate a random variable to represent the white noise e_t depending on the (Box and Muller) method, which is follows normal distribution at zero and variation ($\sigma = 0.15$).Figure (4) shows the generated time series.
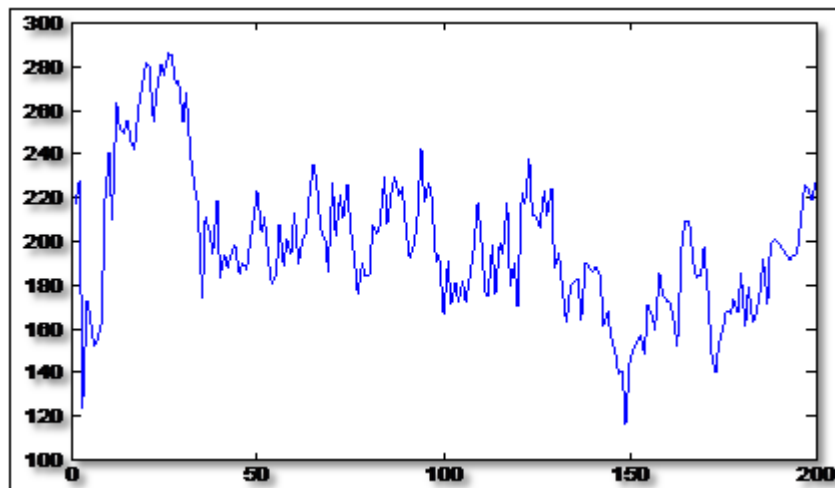

Figure 4: Represents the self-generated regression time series

Artificial Neural Networks:
The artificial neural network inputs have dependent on one variable by the displacement of the series one degree or dependent on two variables through the displacement of the series two degree "Lag1, Lag2".

Network type:
A reverse spread network for error BP multilayer is one of the most important types of artificial neural networks that have been adopting in time series processing, prediction, analysis and data classification, which are the most types of immunity networks.

Network architecture:
The selected network is composed of three layers (inputs, hidden, outputs) interconnected by links or weights, as the

trade-off between them is determined by the Means Square Error (MSE).

Back propagation "BP":
The BP network architecture for this series is as follows:

Input layer:
The number of nodes in this layer is one node of the lag of the time series generating one degree and another two nodes by lag the series two degrees.

Network Training:
In the network training phase, the logistic function and the exponential function were adopted (Sigmoid) reaches a maximum of 1500 repetitions or achieves one of the training conditions the following criteria are adopted (Regression by extracting the regression value R to evaluate the best

performance of the network and training status) in determining the neural network priority.

Hidden layer: The most appropriate number of nodes in this layer is 25 nodes, and 1layer.

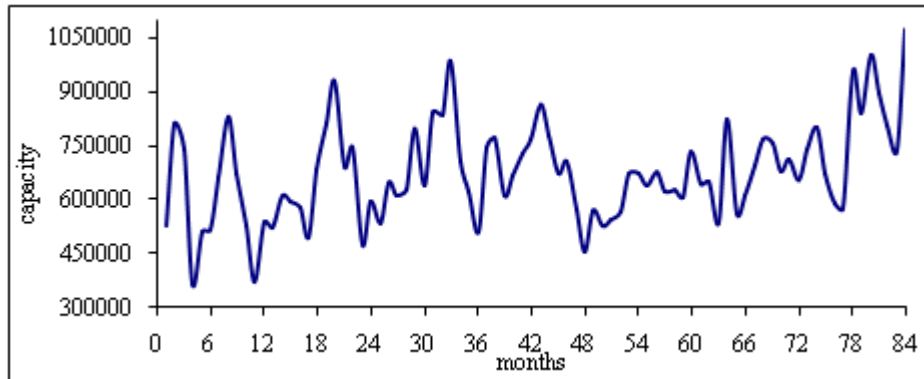**Output layer:** The number of nodes in this layer is one node represented by the variable X_t.

Figure 6: Time series of electricity sold to Baghdad for the period 1995-2001

The results of evaluation of the back propagation "BP " fortime series under study and based on the above data and the algorithm (1), it was as follows:

Table 1: Represents the results of the neural network error criteria for the generated time series

| R^2	MAPE	RMSE	MSE	Error criteria / Number of network input
0.80	13.15	16.84	283.58	One node
0.93	9.73	13.29	176.67	Two nodes

The results of the error criteria shown in Table (1), the best model is the artificial neural network when the number of inputs of two nodes and the amount of ratio of network improvement when the number of inputs two nodes of one node 35%this is shows the efficiency of the network model when two nodes in the input. This is indicates of the hypothesis significance of the search of any improvement of the outputs of the neural network when increasing the number of inputs nodes of that network.

Application part:
On the practical side, the electrical energy data sold to Baghdad City were selected for the researcher's knowledge of the nature of the work of this institution through a previous research group conducted by the researchers. Figure (5) shows the total electric energy data sold for the city of Baghdad for the period 1995-2001, the most regular period of electrical energy. In this respect, the methodologies used to model artificial neural networks were exposed. It is as follows:

1) Traditional Method: This method is based on the tools and principle of time series.
2) The proposed method: This method is based on the principle of causal models.
3) Box-Jenkins methodology in string processing was used.

Figure (7) shows the electrical power sold to the city of Baghdad for the period 1995-2001,and illustrates the behavior of the self-correlation and partial self-correlation of the selected series in this study.
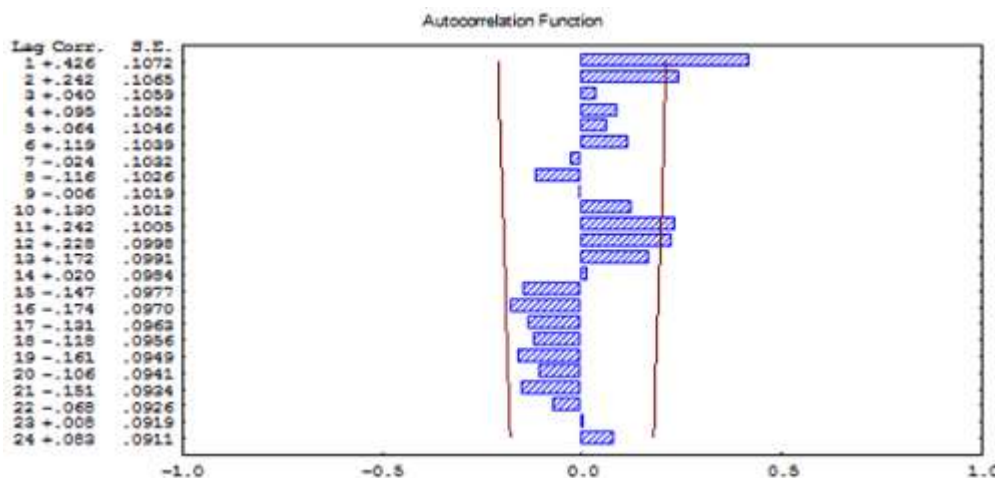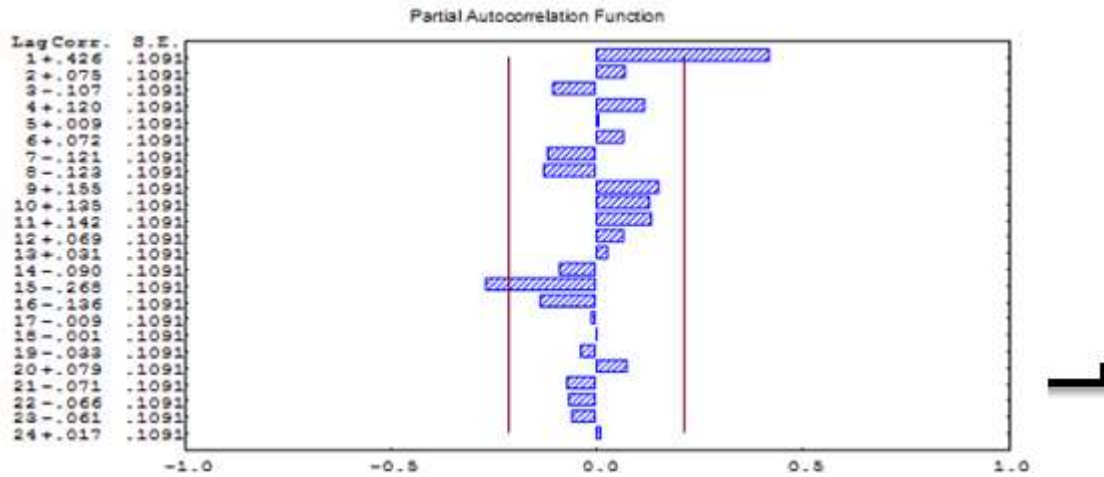
Figure 7: The behavior of autocorrelation and patial autocorrelation functions for time series

The previous form of behavior of autocorrelation and partial autocorrelation is illustrating that the irregular series in the seasonal length due to the nature of the action causative to the overlap "interconnection" between the months, this is leading to irregular seasonal length.

Forecasting Models
At this stage, the Box-Jenkins models and the neural networks with reverse error propagation were used to predict future electrical power data sold to Baghdad. The traditional and proposed method was used in the modeling of artificial neural networks. The results were as follows:

Predicting the use of the Box-Jenkins models:
The appropriate model that has been diagnosed in the light of available data is the multiplier seasonal modelARIMAS $(1,0,1)(0,1,1)_{12}$. The parameters of model were estimated using the ready program Statistics and as follows:

Table 2: Represents the estimation of model parameters and Bucks Jenkins

Parameters	Estimation	S.E
Φ	0.95453	0.06593
Θ	0.73079	0.11356
Θ	0.55965	0.16122

The following figure 8, illustrates the behavior of the autocorrelation function and the partial self-correlation of errors.
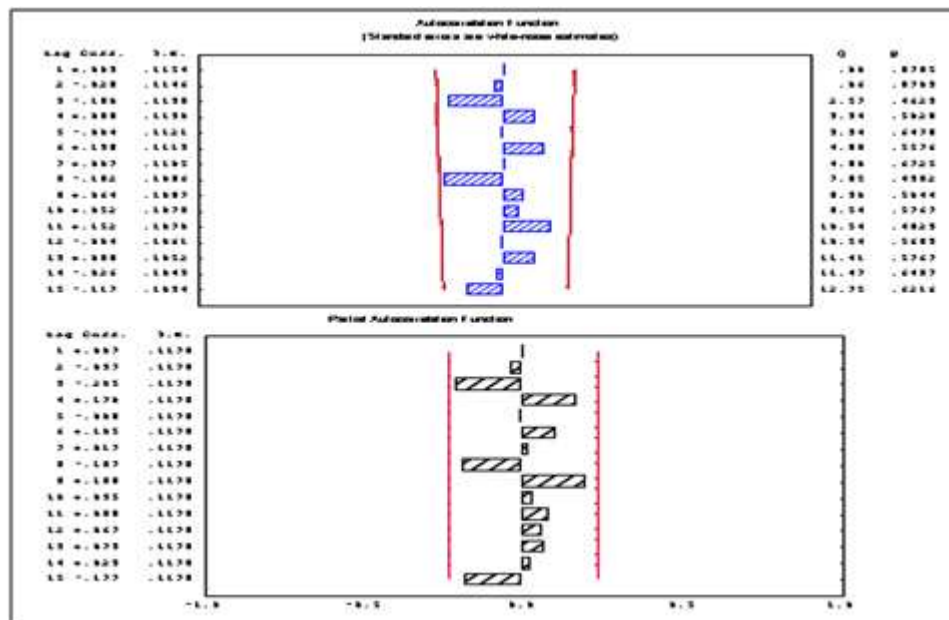

Figure 8: The behavior of autocorrelation and partial autocorrelation functions for residual

The behavior of the autocorrelation function, the partial autocorrelation function of the errors and Q value, that the errors are random behavior and the model is appropriate. The following figure (9) shows the conciliation orientation of the series and Table (3) shows the results of the statistical parameters of the Box-Jenkins model. (The program Statistica extracted the results).
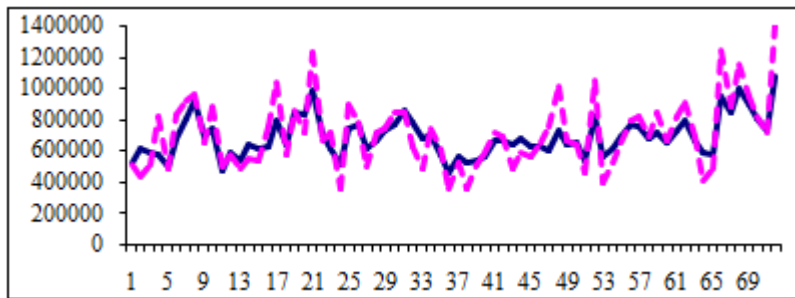
Figure 9: Curve fitting for Box-Jenkins method

Traditional neural network:

Most researchers are rely on the Box-Jenkins methodology in the modeling of neural network input when the data are of the time series type (time series instruments) Self and partial correlation functions. Therefore, the model can be diagnosed based on the behavior of the self and partial correlation functions Figure 5 shows the seasonal multiplier modelARIMA$(1,0,1)(0,1,1)_{12}$. Therefore, the input of the neural network based on this method is as follows:

X_t-1,Xt-2Xt-12,Xt-24

Where, the two variables(X_{t-1},X_{t-2})refer to the direction in the series. While, the two variables $(X_{t-12},X_{t-24})$are refer to the seasonal in the time series.The optimal model of the neural network that is reachedcan be described as follows:

Network type:
Back propagation "BP" multilayer is one of the most important and best types of artificial neural networks. Therefore, back propagation "BP" was adopted in time series processing ،predicting, analyzing, and classifying data in this study.

Network architecture
The selected network consists of three layers are (inputs, hidden, outputs) interconnected with links or weights, as the trade-off between them is determined by Mean Square Error (MSE).

Network of Back propagation "BP"of Error:
Back Propagation "BP" of network architecture for this series is as follows:
Input layer based on the Box-Jenkins model, the number of input nodes are four node Xt-1,Xt-2Xt-12,Xt-24, where the two variables (Xt-1, Xt-2) is refer reflects the direction in the series, while the variables (Xt-12, Xt-24) reflect seasonal in the time series.

Network Training: In the network-training phase, the logistic function and the exponential function (Sigmoid) are used.The maximum limits of repetition reaches to 1500 or achieve one of the training conditions. The following criteria (regression by extracting the regression value "R" to assess the best performance of the network and training status) were adopted.in determining the neural network priority.

Hidden layer: The number of nodes in this layer is 25 knots and one layer.

Output Layer: The number of nodes in this layer is one node represented by the Xt variable.
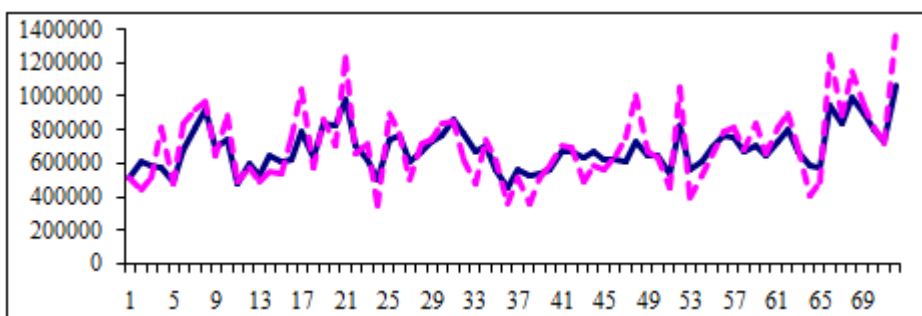The results of the Back Propagation "BP" evaluation for the time series under this study were as follows:

Figure 10: Curve fitting for traditional method

Neural Network Modeling in Proposed Method:
The idea of modeling neural network input by the proposed method is based on the idea of modeling causal models where search for a relationship is a cause and a result of the series involved. As a result of prior knowledge of the nature of the electrical system through a set of previous field research of the researcher. The energy sold can be formulated depending on the electrical power received and based on the relationship (17) and the electrical energy sold.

$$X_t = .25XX_{t-2} + .5XX_{t-1} + .25XX_t + \frac{\xi}{1-\phi B}\ e_t \ \ldots\ldots(17)$$

Whereas:
Xt: represents the energy sold
XXt: represents the energy received
It is worth mentioning that the relationship (17) has depended on the operations research and information unit in the Ministry of Electricity.

Therefore, the input of the neural network based on this method is as follows:
Xt-1, Xt-2Xt-12, Xt-24, XXt-2, XXt-1, XXt.

The two variables (Xt-1, Xt-2) reflect the direction in the series. The variables (Xt-12, Xt-24) reflect seasonal in the time series, while the variables are represented, XXt-2 XXt-1,XXt energy received on energy sold. The optimal model of the neural network reached can be described as follows: (The program Mat lab extracted the results).

Network of Back propagation "BP" of Error:
Back Propagation "BP" Network architecture for this series is as follows:
Inputs layer: based on the Box-Jenkins model, the number of input nodes are of seven-node input nodes are Xt-1,Xt-2, Xt-12,Xt-24, XXt-2, XXt-1, XXt.

Where the two variables are reflect (Xt-1,Xt-2) the direction in the time series, and the other two variables are reflect (Xt-12, Xt-24) the seasonal in the time series. While, evaluate of variables (XXt-2 ،XXt-1,XXt) are represented the energy received that effect on the energy sold.
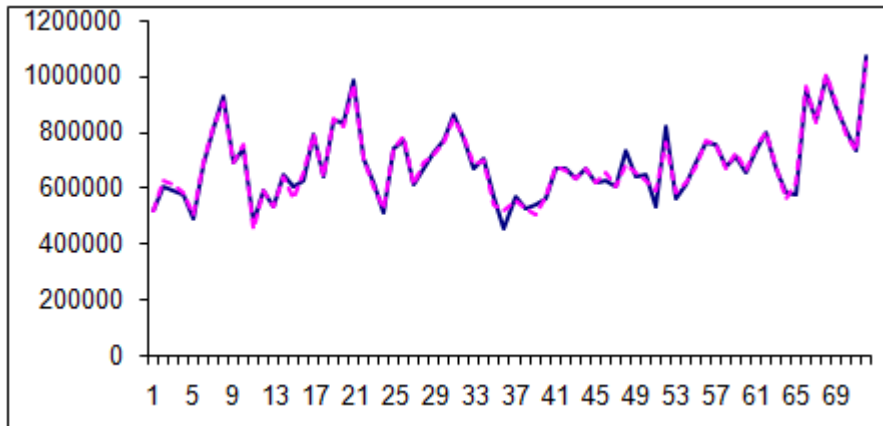

Figure 11: Curve fitting for proposed method

3. Analysis of the Results

Table (3) shows the results of the statistical standards of the artificial neural network ANN which the Back Propagation "BP" of network error of that series.For the purposes of comparison and analysis among the methods used in this research, statistical standards were adopted (Mean Average Error "MAE", Mean Square Error "MSE", Rate Mean Square Error "RMSE", MAPE ", and R2) the results were as follows:

Table 3: shows the statistical criteria of the methods used

Criteria / Ways	ME	MSE	RMSE	MAPE	R^2
Box-Jenkins	18.372	1457.901	38.18247	13.422	0.702
tradicional ANN	18.1872	1457.804	38.1812	13.302	0.704
(ANN) **proposed method**	0.1177	407.203	20.17927	2.1938	0.901

In contrast, the proposed method can model these series accurately and more robustly to take advantage of the learning, training and self-adaptability of any model owned by neural networks. Based on the causal models and building a mathematical model, there is a similarity is evident in the results of the standards with the direction of the Box-Jenkins method and the traditional way of modeling artificial neural networks (ANN).

4. Conclusion

The most important results reached in this research can be summarized as follows:
1) The results of the empirical and applied side show that the systematic increase in the number of nodes of artificial neural network inputs (ANN) leads to improving the outputs of neural networks, and to obtain outputs corresponding to inputs and then reach the desired results. The amount of improvement in the results of the outputs in the empirical side was 35% and in the applied side 51%..
2) The results of the statistical criteria show in Table (3) that the imposition of the Box-Jenkins method in diagnosing the model to determine the input of the artificial neural network (ANN) leads to results similar to the results of the Box - Jenkins method, especially if the behavior of the data is irregular, as well as the lack of full use of the property of learning, training, and self-adaptation with any form of the models possessed by artificial neural networks. While the Box-Jenkins methodology assumes the "traditional method of modeling" is the regularity of data behavior in general and the regularity of the seasonal length of the series, although the behavior of that data is irregular or that the series of behavior is iregular. This is demonstrated by the results of the statistical standards, where it is assumed that the difference between the methodology results (1, 2) and methodology (3) is significant, while the results of methods (1) and (2) are very similar. The difference between the results of the methods that is used in the research does not appear to be relatively consistent in data with regular behavior, While, occuring in data or the series with irregular behavior which is more common in applied reality, Due to the lack of research and studies in this area, and the difficulty of the theoretical treatment of these types of data or time series.
3) Artificial neural networks were used to predict time series with irregular behavior in seasonal length as an alternative to traditional methods adopted. This is

because these networks are able to learn, train and adapt themselves to any model while the previous models are assume seasonal regularity although they are irregular.

4) The methodology of causal models has been adopted in the modeling process to determine input and output only for neural networks in case of time series, Where there is no need to build the mathematical model, that sometimes is requires, especially in nonlinear relationships, to elaborate mathematical formulation or describe the effect of some variables on a particular phenomenon. While, artificialneural networks do not require the process of formulating a mathematical relationship and describing the effect of variables concerned on a specific phenomenon or hypotheses but to identify inputs and outputs to take advantage of the methodology of causal models in the modeling process to identify inputs and outputs only..

References

[1] Nasser, A. M. and Obeidi, M. Hamid (2003). Use of neural networks to predict future time series values that are irregular in seasonal length. Iraqi Journal of Statistical Sciences. Volume 3. 6 issue.

[2] Alexander,I and Galushkin,G.(2012).*Neural Networks Theory*. Springer-Verlag London UK PP884.

[3] Allende,H and Salas,R.(2012).Artificial Neural Networks in Time Series Forecasting. *International Journal of Forecasting* 8:338-358.

[4] Palit,A and Popovic,D.(2005).*Computational Intellgence in Time Series Forecasting*. Springer-Verlag London UK PP972.

[5] Peralta,J and Sanchwas,A.(2012).Time Series Forecasting By Evolving Artificial Neural Networks Using Genetic Algorithms And Differential Evolution. *International Journal of Forecasting* 9:496-521.

[6] Phili,K and Deligiorgi,D.(2012).Application Of Artificial Neural Networks For The Spatial Estimation Of Wind Speed In A Coastal Region With Complex Topography. *Journal of Computer Applications* 38:75-82.

[7] Qing,Cao and Thompson,Mark.(2012). Forecasting Wind Speed With Recurrent Neural Networks. *European Journal Of Operational Research* 21:148–154.

[8] Ratnadip,RK.(2011).A Homogeneous Ensemble of Artificial Neural Networks for Time Series Forecasting. *Internacional Journal of Computer Applications* 32: 887-975.

[9] Riahy,GH and Abedi,M.(2008). Short Term Wind Speed Forcasting For Wind Turbine Applications Using Linear Prediction Method. *European Journal of Operational Research* 33:35-41.

[10] Sana,Hamra and Gabr,Mahmoud.(2003). The Box–Jenkins Analyswas And Neural Networks: Prediction And Time Series Modelling. *Journal of Applied Mathematical Modelling* 27: 805–815.

[11] Sandy,D and Keith,Ord.(2000). Automatic Neural Network Modeling For Univariate Time Series. *International Journal of Forecasting* 16: 509–515.

[12] Stetter,Olav.(2012). Model-Free Reconstruction Of Neuronal Network Connectivity From Calcium Imaging Signals.*International Journal of Forecasting* 39: 759–771.

[13] Wei,William.(1990). *Time Series Analyswas: Univariate and Multivariate Methods*. Addwason-Wesley Publwashing Company Inc NewYork USA PP968.