A comparative Study of Forecasting the Electrical Demand in Basra city using Box-Jenkins and Modern Intelligent Techniques

Khadeega Abd Al-zahra
Dept. of Geographic
college of Arts
University of Basrah
Basrah, Iraq
kadeejha2007@yahoo.com

Khulood Moosa
Dept. of Electrical
Engineering
University of Basrah
Basrah, Iraq
khuloodmo@yahoo.com

Basil H. Jasim
Dept. of Electrical
Engineering
University of Basrah
Basrah, Iraq
hanbas632@gmail.com

Abstract The electrical consumption in Basra is extremely nonlinear; so forecasting the monthly required of electrical consumption in this city is very useful and critical issue. In this Article an intelligent techniques have been proposed to predict the demand of electrical consumption of Basra city. Intelligent techniques including ANN and Neuro-fuzzy structured trained. The result obtained had been compared with conventional Box-Jenkins models (ARIMA models) as a statistical method used in time series analysis. ARIMA (Autoregressive integrated moving average) is one of the statistical models that utilized in time series prediction during the last several decades. Neuro-Fuzzy Modeling was used to build the prediction system, which give effective in improving the predict operation efficiency. To train the prediction system, a historical data were used. The data representing the monthly electric consumption in Basra city during the period from (Jan 2005 to Dec 2011). The data utilized to compare the proposed model and the forecasting of demand for the subsequent two years (Jan 2012-Dec 2013). The results give the efficiency of proposed methodology and show the good performance of the proposed Neuro-fuzzy method compared with the traditional ARIMA method.

Index Terms- **ARIMA models, Artificial neural networks, Box-Jenkins, Neuro-fuzzy.**

I. INTRODUCTION

The development in human population, buildings and technology applications result in fast increase in the electric demand. So that, predicting electric demand is necessary in decision-making for the efficient management of electric energy and important in effective energy saving [1]. Prediction of electric load is very important issue for modern power system engineers and a good month ahead prediction of electric load is required for efficient performance of various Energy Management System functions such as unit commitment, economic dispatch, fuel scheduling, and unit maintenance . Prediction of an event requires vague, imperfect and uncertain knowledge [2].

Time series forecasting is an important issue of prediction in which past values of the same variable are given and analyzed to simulate a system model describing the specify relationship. The model is then used to predict the time series into the future [3]. This model approach is needed when little knowledge is available on the available data values of the specified process. Earlier, forecasting systems models were built using rules given manually. When the number of inputs increased the rules of the model became complicated and forecast operation will be tedious. In the last years, many Articles gave efforts to develop the time series forecasting models[1-5].The prediction using neural networks are modern methods which received wide

attention in multiple areas, including predict prices currencies and financial resources, weather conditions[6], the flow of water in rivers[7], and consumption of electrical energy [1-3]. And have been widely used because they do not need to be strict and precise conditions for the purpose of predicting the data of non-linear behavior.

The Engineering models help to find and build a forecasting system model that could adapt the increasing number of inputs and generate rules accordingly. The speed and accuracy obtained is superior to manual prediction models. Forecasting a system is done by learning from the past data values for which historical data is obtained and analyzed to give the data values results in the event [5].

The main purpose of this Article is to build a prediction system to forecast the future occurrence of the electric energy demand in Basra city. The theoretical part gives the principles of the intelligent techniques including neural network , Neuro -Fuzzy model and the statistical ARIMA model. A fuzzy based approach for month ahead prediction of electric demand using modern technologies such as neural network and fuzzy logic are utilized in this work.

The practical part deals with data of electric demand in Basra city. Actual data are taken from monthly records of the Basra electric distribution management during the period from (Jan 2005 to Dec 2011). Actual and predicted monthly electric energy demand data of the corresponding months are used in the time series analysis . The data were used to compare the proposed model of data model and the forecasting with demand of electricity for two years ahead (Jan 2012-Dec 2013).The predicted results are compared with actual demand data values to measure the performance of the proposed methods.

II.A FUZZY INFERENCE SYSTEM (FIS)

The FIS has two parts a knowledge base a processing part [6]. The structure of the

prediction model based on fuzzy logic is given in Fig.1. As shown in that figure , In first stage, the crisp numerical variables were the input of the FIS system. These variables were changed to linguistic variables by passing through a fuzzification part which became the fuzzy input for the inference engine. The fuzzy input was converted to fuzzy output by using the rules of the inference engine [6].
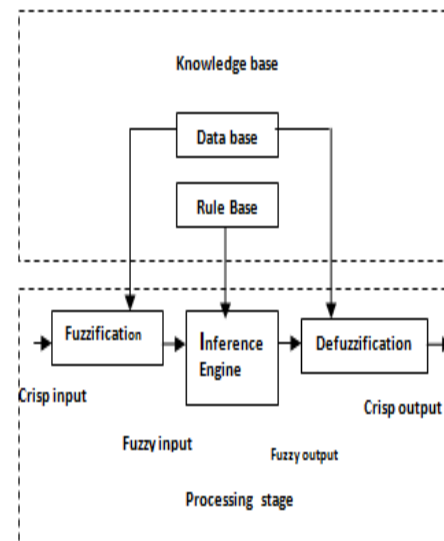

Fig.1:Fuzzy inference system

The part of knowledge base gives the membership functions (MFs) and fuzzy rules which are needed for the process[8]. The linguistic values are then transformed by a defuzzification part into numerical data that became the output of the FIS system. Several models have been utilized to test the data in the complicated systems to get the IF- THEN rules to fuzzy system [8]. To implement a fuzzy system, the experts of human often used to generate the IF-THEN rules to express their knowledge. ANFIS system model collects the neural networks learning abilities with the linguistic rule of fuzzy logic systems [9]. To establish the optimal definition of the premise and consequent part of fuzzy IF-THEN rules the ANFIS system performed the knowledge base rules by utilizing the capability of training for neural networks, and then using the error function of the system. The ANFIS model is utilized in this article to solve the problem of electricity demand forecasting in Basra city.

III. ARTIFICIAL NEURAL NETWORK

There are many types of artificial neural networks (ANN). It can use different topologies and learning algorithms. The neural network models is a computational approach to simulate the biological neural network information and simulate the processing that occurs in the brain. The main purpose of ANN is its ability to learn. The neural network is able to find the relationships between input and output values in the training process [2]. The ability of neural networks to predict comes directly from their ability to generalize and find the hidden relationships between input and output data[3]. After training, the network is able to forecast the future values of a specified sequence on the basis of several history data and any current factors. Neural networks are used for simulating a range of nonlinear problems [10]. A feed forward network with single hidden layer is the most popular used . The ANN models can approximate many class of functions with a high degree of accuracy using parallel processing of the data . The following equation indicates the relationship between the output (yt) and the inputs (y_{t-1}; y_{t-2}; …;y_{t-p}) [12,17]:

$$y_t = w_o + \sum_{j=1}^{q} w_j \, f \left(d_{oj} + \sum_{i=1}^{p} d_{ij} y_{t-i} \right) + \varepsilon_t \quad …(1)$$

where the variables p and q represent the number of input nodes and the number of hidden nodes respectively . w_j and d_{ij} are the weights of model , with j = 0, 1, 2… , q;   i = 0, 1, 2, … , p; In a feed forward neural network model, the data values moves forward in one direction , from the input nodes, through the hidden nodes and to the output nodes. ANN is trained using a simple learning algorithm called the delta rule which calculates the error values between actual output and desired output data value, and utilize this to make an adjustment to the weights, thus implementing a form of gradient descent[10].This type of networks implements using multiple layers of computational units, interconnected in a feed-forward way. Each neuron in one layer has directed connections to the neurons of the subsequent layer. In many applications the units of these networks apply a Gaussians function as an activation function[13,15]. Multi-layers network use several techniques of learning, such as back propagation. The back propagation has one input layer and one output layer as well as one or more hidden layer [12]. To calculate the error values the difference between the actual output with the values of target is calculated. Then the error is fed back in the network to adjust the weights of each connection in the network therefore the error values will be reduced using the algorithm[13] . A non-linear optimization method that is called gradient descent is utilized . The derivative of the error function with respect to the network weights is determined, and the weights are then changed such that the error decreases [12]. The back propagation network implement supervised training with a finite number of pattern pairs. Each one of these pairs consist of an input pattern and a target pattern. The training process consists of two steps[14]:

1.Feed forward: each one of the pattern pairs are enters and the actual output is computed.

2.Feed backward: in this step all the weights are adjusted, which presented in the network, depending on the difference error between the desired and actual output. This operation is repeated until the error between desired and actual outputs reaches the specified minimum value [14].The least mean square error function (LMS) is the most widely used in the back propagation algorithm to determine the error of the network.

IV. ADAPTIVE NEURO FUZZY INFERENCE SYSTEM(ANFIS):

A neuro-fuzzy network is a fuzzy inference system in the body of an artificial neural network. Depending on the FIS type, there are several layers that simulate the processes involved in a fuzzy inference like fuzzification, inference, aggregation and defuzzification.

Embedding an FIS in a general structure of an ANN has the benefit of using available ANN training methods to find the parameters of a fuzzy system. The Neuro fuzzy hybrid system is a learning mechanism that utilizes the training and learning neural networks to find parameters of a fuzzy system based on the symptoms created by the mathematical model[10].

The neuro-fuzzy computing provides the system identification and interpret ability of fuzzy models and adaptive learning capability of neural networks in a single system. In the last decade, various neuro-fuzzy systems have been developed [13]. Among them, adaptive neuro-fuzzy inference system (ANFIS) provides a systematic and directed approach for model building and gives the best possible design parameters in minimum time. They have got wide acceptance for modeling many real world problems.

V. THE MATHEMATICAL MODEL OF FORECASTING PROBLEM USING ANFIS

Let us have n discrete samples $x_1, x_2, x_3, \ldots x_n$ at successive time points t_1, t_2, t_3,**....** t_n Then the problem of forecasting consists of predicting the value x $_{n+k}$ at some future point of time t $_{n+k}$ where k is the duration of the forecast:

$$x_{n+k} = F (x_1, x_2, x_3, \ldots x_n) \qquad \ldots\ldots(2)$$

where F is some unknown function. The ANFIS model components is shown in Fig. 2 [21]. The input vector $xi(k)$ (i=1 …n) is entered to the network model input at time k with output values of the model are $u(k)$. n is the number of nodes in the input layer. In first layer the input signals transmitted directly to the next layer without any computation. The number of nodes is equal to the sum of inputs. The outputs of this layer can be expressed as

$$y_i^{(1)} = x_i^{(1)} \quad , \quad i = 1, 2, \ldots, n.$$

In second layer each node in this layer performs the calculation of Gaussian membership function MF, so that a small or large represents one linguistic term. Gaussian MF s is utilized   as shown below [2,18]:

$$u^{(2)}_j(x_i^{(1)}) = \mathbf{e}\exp\left(\frac{-(xi - aij)^2}{2bij^2}\right) \quad i=1,\ldots n, \; j=1\ldots J, \;..(3)$$

where $u^{(2)}_j$ is the output signal with respect to the i th input pattern variable $x_i^{(1)}$ in the j th fuzzy rule , a_{ij} and b_{ij} are center point and spreading of activation function respectively and referred to as premise parameters.
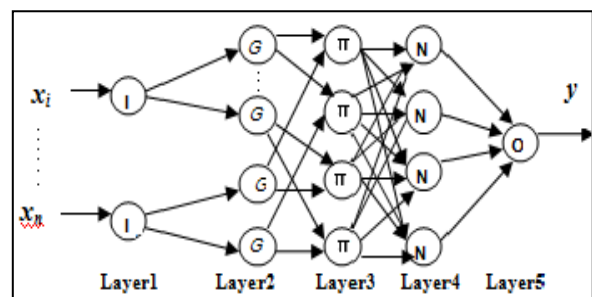

Fig. 2 The structure of neuro-fuzzy inference system where G is referred Gaussian membership function

In third layer the number of rules equal to the number the of nodes. The min AND operator is utilized as a fuzzy T-norm operation as shown below:

$$y_j^{(3)} = \prod_i u_j^{(2)}(x_i^{(1)}) \qquad \ldots(4) \quad , \quad i = 1,2,\ldots 4, j = 1,2,\ldots,16$$

where Π represents the min AND operator. The output of this layer represents the firing strength of antecedent part for each rule.

The fourth layer is a consequent layer. The sum of the output signals $y_j^{(3)}$ is calculated. The firing strength was normalized through dividing its value by the summation of all the firing strengths of all rules as:

$$y_j^{(4)} = \frac{y_j^{(3)}}{\sum_{l=1}^n y_l^{(3)}} , \qquad \ldots(5) \quad , j=1,2, \ldots,16$$

The ANFIS proposed system used a Fuzzy Sugeno model, with 4 inputs $x_i, x_{i-1}, x_{i-2}, x_{i-3}$ and each input variable was denoted with two membership MFs small and large. The parameters of the system are optimized by using ANN training, and structure is calculated by expert knowledge [2,19].   To get the forecasting operation sixteen fuzzy IF-THEN rules are used, which are the same rules as that in [2,11] as follows: Rule j

IF $(x_{i-3}$ is $B_1{}^j$) AND $(x_{i-2}$ is $B_2{}^j)$ AND

$(x_{i-1}$ is $B_3{}^j)$ AND $(x_i$ is $B_4{}^j)$THEN

$y^j = n_1{}^j x_i + n_2{}^j x_{i-1} + n_3{}^j x_{i-2} +$

$n_4{}^j x_{i-3} + n_5{}^j$, …..(6)   j = 1,2,…16

where y^j is the prediction result according to the jth fuzzy rule, B_i^j is the fuzzy set associated with the i_{th} input variable in the jth fuzzy rule and n_k^j is the parameter with i=1,2,..,4 and k=1,2,..,5. In layer five , after a linear combination of the input signals, the output of the ANFIS was calculated by[17],:

$$y = \sum_j y_j^{(4)} \left(n_1^j x_i + n_2^j x_{i-1} + n_3^j x_{i-2} + n_4^j x_{i-3} + n_5^j \right) ..(7)$$

Where j=1,2,….,16 and n_1, n_2, n_3, n_4, n_5 are set of unknown parameters called consequent parameters. The adjusted parameters in the neuro-fuzzy network can be divided into two categories based on the IF- part(premise part) and THEN-part (consequent part) parameters of the fuzzy rules [12]. In the premise part the center a_{ij} and the width b_{ij} of the Gaussian membership functions are to be fine tuned ,where as in the consequent part the adjusted parameters are the consequent weight c_i. A gradient descent back propagation algorithm is used least mean square error function to adjust the parameters of the neuro-fuzzy network by using the training patterns[12,14]. The main goal of supervised learning algorithm is to minimize this error function which has the formula[13]:

$$E^P(Z) = \frac{1}{2}(y^p - y^d)^2 \text{ and }\quad E(Z) = \sum E^P(Z)$$

Where E is the total error of the ANFIS ,E^P is the error in the pattern y^d is the desired output in the pattern p , y^p is the actual output in the pattern p ,p is the patterns number ,Z is the parameter vector(a_{ij}, b_{ij}, c_{ij}), i=1,..n. In order to minimize the objective function E(Z) ,a neuro fuzzy learning algorithm for tuning the parameters a_{ij} ,b_{ij} ,c_{ij} which are described as followes [11,13]:

$$a_{ij}(t+1) = a_{ij}(t) - k_a \frac{\partial E}{\partial aij(t)} \quad …..(8)$$

$$= a_{ij}(t) - k_a [y(t) - (y_d(t)) * (c_i(t) - y(t))$$

$$* \frac{y_j}{\sum_{l=1}^n y_l} * \frac{(x_j - a_{ij}(t))}{(b_{ij}(t)^2)}] …..(9)$$

$$b_{ij}(t+1) = b_{ij}(t) - k_b \frac{\partial E}{\partial bij(t)} \quad ….(10)$$

$$= b_{ij}(t) - k_b [y(t) - (y_d(t)) * (c_i(t) - y(t))$$

$$* \frac{y_j}{\sum_{l=1}^n y_l} * \frac{(x_j - a_{ij}(t))}{(b_{ij}(t)^2)}] …(11)$$

$$c_{ij}(t+1) = c_i(t) - k_c \frac{\partial E}{\partial ci(t)} \quad ….(12)$$

$$= c_i(t) - k_c [y(t) - (y_d(t)) * \frac{y_j}{\sum_{l=1}^n y_l}]….(13)$$

Where k_a , k_b , k_c are the learning rates parameters take values between [0,…1].

VI. BOX- JENKINS BACKGROUND

A-Time Series

The time series can be defined as a variety of observations related with each other, and made sequentially in time to represent the manner of specific phenomena in past period of time, so if (Y_t) is a symbol represents the series then [19]:

$Y_t = y_1, y_2, y_3, ……..,y_n,$

where :t: time , n: integer number represent the length of the time series ,Yi: observations,(i =1, 2, ……, n).

B-Stationary

Stationary concept consider an essential and basic issue in study and analysis any time series, as well as it forms basic depended in building a prediction models. So, the time series classify depending on the style of the available data (its fluctuation)[20,23].

1)Stationary Time Series

the time series be stationary if there is constancy in its mean and its variance, even not to be oblique to growth or to decrease, and the mean could calculate as in the following equation [21,24]:

$$\bar{Y} = \frac{1}{N}\sum_{t=1}^{N} y_t \dots (14)$$

where : t: time, N: length of the time series ,

Y_t: observation in t time. While the variance calculate in the following equation[24]:

$$\bar{\delta} = \frac{1}{N}\sum_{t=1}^{N}(y_t - \bar{Y})^2 \dots (15)$$

where $\bar{Y}$:the mean

2) Non Stationary Time Series

They are series, which its data be in growth or decline , or in another words, data of the series has (trends). If the time series is non-stationary, (the Method of Differences) is used to removing the non- stationary by taking (d) from the suitable differences to the original data. And differences take in different degrees to warranty stationary, and almost it turned on to stationary after taking the first difference, or the second, or the third[20]. If we assumed, the series of observations (Y_t) are non-stationary, it could be removing after taking the first difference[24]:

$$\nabla^1 w_t = (1 - B)w_t = w_t - w_{t-1} \dots (16)$$

Where   *B :* Backshift operator

∇: Back Ward Differences Operator

Or could be removing after taking the second difference[21]:

$$\nabla^2 y_t = (1 - B) \; y_t = (1 - 2B - B^2)y_t ..(17)$$

In general, removing non-stationary will be done after the difference (d) as in the following equation[20,23]:

$$\nabla^d w_t = y_t ...(18)$$

To distinguish if the Time Series is stationary or not by elicitation the value of the Auto Correlation function , the seasonality could determine by drawing the time series and test Auto Correlation values for the time periods, if it found that Correlations have significant differences at constant time periods(represent season length), so, the time series is seasonality.

C- Auto Correlation Function

the Auto Correlation function consider an important criterion to association power between the adjacent points in the time series (Yt),through which the degree of related between the same variable values is determined at periods of time and with a different displacements which represent the time lags, which define as: it is the value of the displacement betweentwo series that are wanted to compared between each other[20]. The values of the Auto Correlation function is equal (1) when the displacement (K= 0) lag, that means there is fully resemblance between the measured amount and itself. In the case of series is non- stationary autocorrelation function shows a deviation from right to left, which means the sequential values have correlation to each other, and it goes close to (zero) in case of stationary, and auto correlation function for maximum displacement expressed with the following equation[21],[22]:

$$P_k = \frac{\text{cov}(y_t, y_{t+k})}{\delta_y^2} = \frac{\gamma_k}{\gamma_0} \dots (19)$$

where γ_k: variable variation value, γ_0: variable dissimilarity value. The Auto correlation Function is considered an important criterion within time series analysis because it benefit the researcher at several things as[21]:

1-   Give an important information about the basic style of the time series.

2-   The stationary of the series could be determined through its values.

3-   It is used to determine if the chosen symbol is suitable to predicate or not, and that by testing the random of the predicative errors .

4-   Predicative(Forecasting).

VII. BOX - JENKINS METHOD

ARIMA   (Auto Regressive Integrated Moving Average) considers one of the best efficient ways for alteration with different types of the time series that contain of different styles in the same time whether it stationary or not, seasonality or not[21]. Box – Jenkins method depends non-specific group of symbols and with different degrees, and that according to the characteristic of time series (trends, seasonality , random, periodical).   The work with Box – Jenkins method includes four stages which are[22]:

1. Model Identification .

2. Parameters Estimation.

3. Diagnostic Checking.

4. Forecasting(Predictive).

After testing the Box-Jenkins's symbols of the study data appear that the symbol which gave a close result to the original data style is Multiplicative Seasonal Model, and that because the study data contain both properties(trends, seasonality) furthermore, exist some random fluctuation [19]. This model uses for the time series have autocorrelation values   non – zero,

after taking the difference (d) for time period (12) represent seasonal value. The model writes as the following formula[20,21]:

$$\emptyset_P(B)\phi_p(B^S)\nabla^d\nabla^D s\, Y_t = \Theta_q(B)\Theta_Q(B^s)a_t .. (20)$$

And writes as the following formula[21]:

$$\emptyset_P(B)\phi_p(B^S)w_t = \Theta_q(B)\Theta_Q(B^s)a_t .. (21)$$

And that model symbolizes with ARIMA (p , d , q) x(P , D , Q)$_s$,Where p: represent order of the non-seasonal linear regressive part, q: represent order of the non-seasonal moving average part, d: represent order of the non-seasonal first difference involved part, $\varphi_p(B)$: represent the non-seasonal autoregressive coefficient. $\phi_p(B^s)$: represent the seasonal autoregressive coefficient, $\Theta_q(B)$: represent the non-seasonal moving averages coefficient, $\Theta_Q(B^s)$: represent the seasonal moving averages coefficient, P: represent order of the seasonal linear regressive part, Q: represent order of the seasonal moving averages part, D: represent order of the seasonal difference part.

VIII. THE DEPENDABLE STATISTICAL STANDARDS TO TEST FORECASTING MODEL

A. Mean Square Error(MSE)

The estimated model is tested in which sum of squares error   (SSE) , or average of sum of squares error(MSE) which has the less possible quantity . And it can be calculated in the following equation[24,25]:

$$MSE = \frac{\sum_{t=1}^{n} e_t^2}{n-k} \dots\dots (22)$$

Where: t: time, k: number of estimated models, e: error value , n: number of observations.

B. The Test of Coefficient Limits

Coefficient Limits test is one of the statistical tools which utilized to test the estimated model, by using the autocorrelation of errors(residual)

followed by the normal distribution with mean equal to zero and variance (1/n) and independent of each other.

On this basis, the upper and the lower limits of the autocorrelation is estimated   with the possibility of 95% as the following[22]:

- 1.96 $(1/\sqrt{n})$ < (P_k e_t) < +1.96 96 $(1/\sqrt{n})$..(23)

Where: P_k: auto correlation function coefficient. e_t : the errors(residual) ,n: number of observations. So, if the values of P_k e_t are within trust limits[22] with possibility 95%, that means the errors(residual) are random and thus the stipulated model is suitable and efficient.

C. Test Q for (Box & Pierce)

This criterion uses to measure random value of residuals of the estimated model and with free degrees (k-p-q), and at significant level (5%), so , if the calculated value for Q less than classified value, that point torandom errors   . And Q calculated by the following equation[23]:

$$Q = (N-d)\sum_{j=1}^{k} p_j e_t \approx X_\alpha \ (k-p-q-d)...(24)$$

Where the variables *p*, e_t denoted the auto correlation coefficients for errors of estimated model, p : order of the auto regressive, q : order of the moving averages, n : number of observation for time series under study, d : degree of differences , k = 1,2,3,…n/4 it represents the maximum displacement.

D. T- Test

This criterion uses to knowing the range of the random errors value of the estimated model, and that through test range of differences significance between the real and predictive values of the time series [23].

If T value is not significant at the confidence level(5%) that means the estimated model can be able to omit all existing styles of data, and what actually remains is random errors and vice versa.

And   T   calculate   as   in   the   following equation[23,24]:

$$T = \frac{x-\mu}{\delta^2/\sqrt{n}}......(25)$$

where: x: represent mean of the symbol, μ : mean of society, δ^2 : variance, n : volume symbol,

E. The coefficient of determination: R^2

One of the most important statistical criteria which reliance in test the estimated model is R2, and through which was known the identification range of the estimated model with the original data. And it calculate   as in the following equation[24]:

Where: n: number of observations, t: time,

$$R^2 = \frac{\sum_{t=1}^{n} \hat{Y}^2}{\sum_{t=1}^{n} Y^2}...(26)$$

$Y^{\hat{}}$: the predicative values, Y: the actual values.

IX. SIMULATION AND EXPERIMENTAL RESULTS

In the last years a variety of intelligent techniques have being used to forecast the future demand   accurately   for   electric   energy consumption. In this Article three different predicting models is utilized such as ARIMA statistical method and an intelligent techniques such as ANN and ANFIS to build a forecasting system model for electric energy in Basra city. The statistical data were obtained from electrical distribution management in Basra city for the period of 2005-2013. The neuro-fuzzy structure and its learning algorithm is used to build a forecast model. In the Article the 3 input history data points are used x(t-3) x(t-2) x(t-1) x(t) for forecast monthly model. It includes 4 input and one output neuron. The 12 neurons are used in the hidden layer . A Gaussian membership functions

is utilized in the second layer of system for each input signal.

A.ARIMA   forecasting   model   results   :
The work of Time Series Analysis including four stages for ARIMA analysis to predict the electric demand which are as follows:

1) First Stage: Time Series Analysis and Diagnosis the suitable   symbol
Plot the time series of electrical consumption values for the period 2005-2011 as shown in Fig. (3) . The Fig. indicates that the electricity demand has high order nonlinearity and the time series has trends clearly, it turns to be high and that means the series is non-stationary, so the difference should be taken and the value of (d) should determine according to equations (19-24) , and the series be stationary after the first difference (when d=1). As appears in Fig. (4) which represents autocorrelation functions for study data , also it appears the series is non stationary.
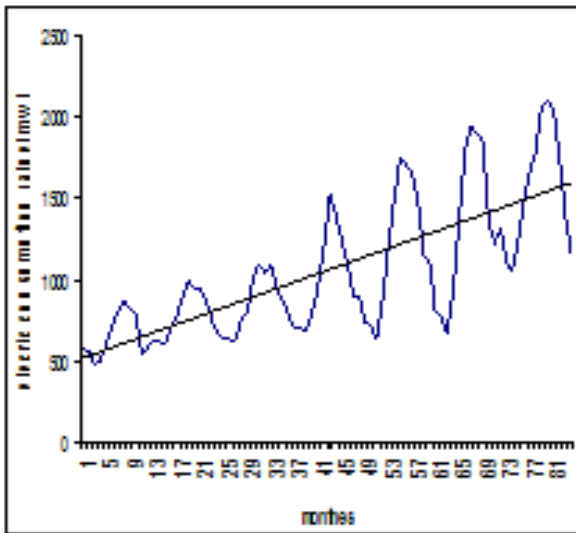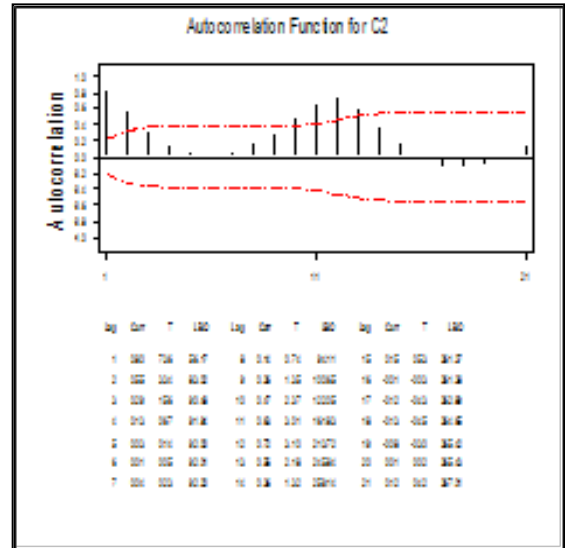

Figure (4)Autocorrelation function of monthly time series(2005-2011) in Basra governorate

2)Second stage: Estimate the Parameters and determine the order of the symbol:

1. calculate the value of the autocorrelation functions of the stationary time series to determine the degree of each autoregressive(p) and moving average(q) which its values close to zero, which means, the predicative model that proposed is ARIMA type. Numbers of parameters of autocorrelation sometimes are inadequate to get an accurate estimation, so that requires number of attempts(for trial) to select the degree of each p &q.
The repeated experience has shown that the appropriate model to predicate is multiplicative seasonal model of degree $ARIMA(2,1,2)(2,1,2)_{12}$.

3)Third stage: Examining the Estimated Model :

1-   Dependence a group of statistical criteria to identify best order for the estimated model which are(MSE, T – test, coefficient limits test, Test Q, The coefficient of determination R2).

Figure (3) monthly Time series for the average of electrical consumption for period (2005-2011)   in Basra governorate

schedule below shows the estimated model's degree with the statistical criteria dependable for test.

Table. I Statistical criteria. dependence

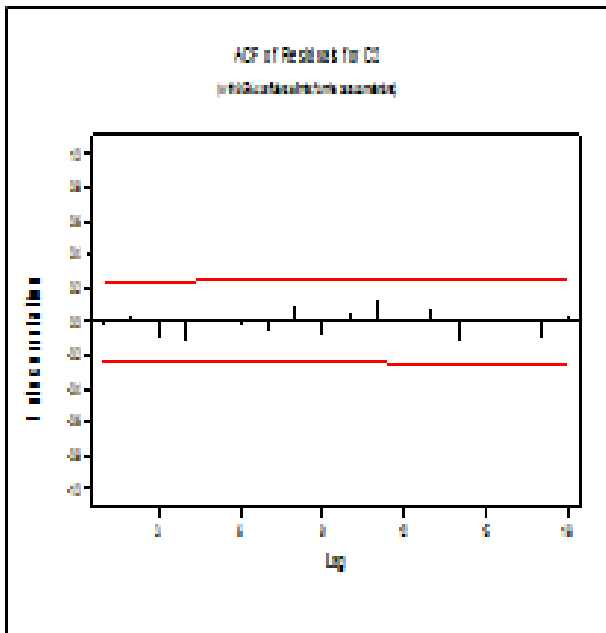| statistical criteria dependence		
Q	T- test	R^2
5	0.47	0.822
suggested model order		
ARIMA$(2,1,2)(2,1,2)_{12}$		

Fig. (5) Autocorrelation functions for residual of the estimated model of electrical consumption quantity for the period(2005-2011)

As for the coefficient limits test to residual of the estimated model, in the Fig.(5) that appears all the residual are not exceed the highest and the lowest coefficient limits test, and that means the residual contained in the model is random residual has no obvious effect on the predictive values of the suggested model to predict.

4) Fourth Stage :Predictive operation

This stage represents the last step, as the estimated model uses after its test   to find predictive values where is calculated according to the estimated model as the best model. As it has been calculated the first predictive value , i.e. for year 2012, then it is used to predict the following value ,i.e. for year 2013. After achieved the four stages to test the estimated model to forecast and generate the required forecasting values, it was found that there is a general trend towards the rising of the actual and the predictive values with a magnitude of change equal to 162.0833.Table. I shows the statistical criteria dependence such as R^2 ,T-TEST and Q for the estimated model . Fig. (6) shows the predictive and actual values according to the suggested ARIMA model.

b. Neural network experiment results:

The neural networks have a flexible nonlinear modeling capability. ANN is used in building the predicting   model for electric consumption of Basra city. A two layers of feed forward neural network have been used . Hidden layer with 12 neuron have been selected . the inputs to ANN are the previous three months electrical consumption. Data from 2005 to 2011 have been used to train our network, while the available data for 2012 and 2013 was used to test the network. Data have been scaled to unity.
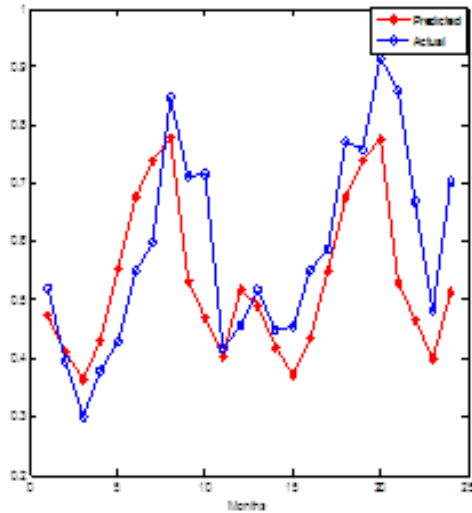
Fig.(6) Actual and predictive values for electrical consumption for period (2012-2013) in Basra governorate using ARIMA forecasting model

Fig. (7) Actual and predictive values for electrical consumption using ANN model

The data-driven approach is suitable model for a different data sets. The mean squared error (MSE) , the root mean square error (RMSE) and mean absolute error MAE are selected to be the forecasting accuracy measures and used as an indications for network efficiency[25]:

$$MSE = \sum_{t=1}^{n} e_t^2 \Big/ n.....(27)$$

$$RMSE = \sqrt{\sum_{t=1}^{n} e_t^2 / n}...(28)$$

$$MAE = \sum_{t=1}^{n} |e_t| / n.....(29)$$

The error $e = C_i - N_i$, where, C_i and N_i are the actual and the predicted values respectively. Fig.(7) shows the actual and predicted data for the test data (2012 and 2013) using ANN model .The number of epochs is 8000 epoch.
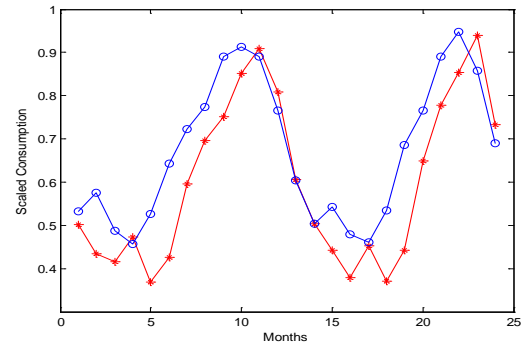
c. The ANFIS experiments results:

ANFIS is utilized for building the   forecasting system model depends on the history data values The ANFIS system is used to reduce the searching space and time for achieving optimal solution, the combination of neural networks with fuzzy knowledge base is implemented for electricity consumption prediction. A Back Propagation, gradient descent learning algorithms are used in this article as a training algorithms to get a suitable architecture of neural network for training the ANFIS structure and to adjust the parameters of learning algorithm from a given set of input and output data, Gaussian MFs have been used where two MFs for each input was selected. The network has two rules. Here, also data from 2005 to 2011 have been used to train the network , while data of 2012 and 2013 was used to test the efficiency of the trained network. Before training, the train set of data has been scaled by multiplying it by 1.5 to generate a network which can deal with inputs having values greater than the maximum vales of the training set.

Fig.8 shows the actual and the predicted results. The results showed that ANFIS provides better training error results than other two methods and less number of epoches. The training is carried out for 2000 epochs. In the result of training the values of parameters of neuro-fuzzy system were found. The result of simulation of neuro-fuzzy prediction model is compared with result of simulation of prediction model based on neural network. To get the performance of neural and neuro-fuzzy forecasting systems the MSE ,RMSE and MAE values of errors between forecast and actual output is given in table II. As shown from the table the error performance of neuro-fuzzy prediction is better compared with nural model and ARIMA model . The result obtained from the simulation satisfies the efficiency of application of neuro-fuzzy technology in building forecasting model of electric energy demand as shown in Fig.8.
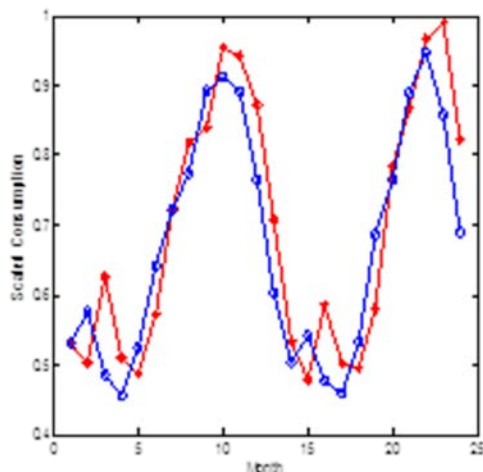

Fig.(8) Actual and predictive values for electrical consumption using ANFIS model

Conclusions:

Energy strategy is necessary for countries development. The energy demand increased when the economy of these countries is grown. Energy management   should improve new strategies for responding this growing of energy consumption. The accurate forecasting of the future energy demand are useful for developing such strategies. In this work a comparative study is implemented between traditional statistical model such as ARIMA model and modern intelligent techniques such as ANN and ANFIS models in which predictions of the future values are depends on past history data. The research deals with using seasonal time series models to study and analysis the monthly data on consumption of electricity in Basra city for the period (2005-2011) , these models are different with high accuracy and flexible in analysis of the time series .

The results of application indicated that the suitable and efficiency model for representing time series data are the seasonal model of order ARIMA$(2,1,2)(2,1,2)_{12}$. According to forecast results of this model we got

the forecasting to monthly consumption of electric energy for future data of two years a head from the period January 2012 to December 2013 , these results showed a harmonic direction with the same actual data values of the time series. The Arima model is implemented via Minitab system program while ANFIS and neural network models are built using Mat lap program tools for training the proposed system model. The one-step-ahead forecasting is considered.

The mean squared error (MSE) , the root mean square error (RMSE) and mean absolute error MAE are selected to be the forecasting accuracy measures. Table II gives the error values of forecasting results. The different approaches for modeling the time series of electrical demand give good results and the models approach made a good forecasting values followed the behavior of the original time series for the months of the years 2012 and 2013, the prediction values of these models converge to the values of the original series .

Table. II A comparative error results

Statistical parameters of error	Box-Jenkins model	ANN model	ANFIS model
$RMSE = \sqrt{\sum_{t=1}^{n} e_t^2 / n}$	1.5801	1.5056	1.2497
$MSE = \sum_{t=1}^{n} e_t^2 / n$	2.497	2. 2668	1.562
$MAE = \sum_{t=1}^{n} |e_t| / n$	0.31604	0.301	0.2491

The degree of changes in the time series and especially random changes, directly affects the results obtained using the methods under study, The rapidly changes in the time series models reduces the efficiency of Box - Jenkins compared with artificial intelligent techniques.
Neural network and ANFIS models are directly affected by the size of the available data and the length of the time series. whenever a sufficient data size is available so that each change in the time series leads to increase the learning degree of the intelligent techniques ,and then increased the efficiency of the network models to predict. And whenever forecasting period increased in the future, the results of the networks is more accurate than the results of Box - Jenkins, through the prediction results obtained from these models. We discussed the results of a comparative study of the performance of neuro fuzzy,   neural networks and traditional   ARIMA method in forecasting time series.
Our work demonstrates that for time series used ANFIS gets a good error performance compared with   neural networks and the Box-Jenkins model. The obtained result from the simulation satisfies the efficiency of application of neuro-fuzzy technology in constructing prediction model of electricity consumption.

REFERENCES

[1] Ahmad, M.Y., Hassan ,M.P. , Abdullah, H.A. Rahman, F., " A review on applications of ANN and SVM for building electrical energy consumption forecasting ",Renewable and Sustainable Energy Reviews, Malaysia, vol. 33, pp. 268-279,May 2014.

[2] Rahib Abiyev, Vasif H. Abiyev, Cemal   Ardil," Electricity Consumption Prediction Model using Neuro-Fuzzy System ",World Academy of Science, Engineering and Technology vol,8 ,pp. 620-624,2007.

[3] G. Peter Zhang, "Time series forecasting using a hybrid ARIMA and neural network model", Neuro computing ,no. 50, pp.159 – 175, 2003.

[4] Gamze Oğcu, Omer F. Demirel, Selim Zaim," Forecasting Electricity Consumption with Neural Networks and Support Vector Regression" ,Procedia - Social and Behavioral Sciences, vol. 58, no. 12, pp. 1576–1585,October, 2012.

[5] Sineglazov, Victor; Chumachenko, Elena; Gorbatyuk, Vladyslav, "An algorithm for solving the problem of forecasting", Aviation , Academic Magn., Vilnius Gediminas technical University, Mar 1, 2013.

[6] Pankaj Kumar, "Minimum Weekly Temperature Forecasting using ANFIS", Computer Engineering and Intelligent Systems , VCSG College of Horticulture, GBPUA&T, Pantnagar, Uttarak hand ,vol 3, No.5, 2012.

[7] Kumar, D. Negesh , Raju , K. Srinivasa and Sathish , T. "River Flow Forecasting using Recurrent neural Networks", Water Resources Management 18,pp.1431-161,2004.

[8] Vaidehi .V ,Monica .S , Mohamed Sheik Safeer .S, Deepika .M, Sangeetha .S, "A prediction system based on Fuzzy Logic", Proceedings of the World Congress on Engineering and Computer Science, 2008.

[9] Yue Wu ,Biaobiao Zhang ,Jiabin Lu Yue Wu "Fuzzy Logic and Neuro-fuzzy Systems: A Systematic Introduction", Faculty of Electromechanical Engineering, International Journal of Artificial Intelligence and Expert Systems (IJAE), pp47-80, vol. 2 : Issue (2) , 2011.

[10] Imianvan Anthony Agboizebeta , Obi Jonathan Chukwuyeni ,"Cognitive Neuro-Fuzzy Expert System for Hypotension Control", Computer Engineering and Intelligent Systems , ,vol 3, No.6, pp21-31,2012.

[11] Pradeep Kumar, Sudhir Kumar Sharma ,Sidheshwar Prasad, "Detection of Fetal Electrocardiogram from Multivariate Abdominal Recordings by using Wavelets and Neuro-Fuzzy systems", Computer Engineering and Intelligent Systems , vol.5, No.4, 2014.

[12] C. Loganathan ,K.V.Girija, "Hybrid Learning for Adaptive Neuro Fuzzy Inference System " ,International Journal Of Engineering And Science vol.2, Issue 11 , pp. 06-13, April 2013.

[13] R. S. Khaudeyer, "Hybrid Approaches: Neuro-Fuzzy and Geno-Neuro-Fuzzy Hybrid Systems for Solving Some Classification and Functions Approximation Problems", Ph.D. Thesis, Department of Computer Science ,University of Basrah, 2003.

[14] E.S.Jabber, "Neuro Fuzzy System for Function Approximation", M.Sc. Thesis, Department of computer science ,University of Basrah, 2004.

[15] Mohsen, H.; Yazdan, S. ,"Artificial Neural Network Approach for short term Load Forecasting for Llam Region", Engineering and Technology vol.28,pp. 280-284 , October 2008.

[16]. A. Abraham and Baikunth Nath, " Hybrid Intelligent Systems: A Review of a decade of Research", School of Computing and Information Technology, Faculty of Information Technology, Monash University, Australia, Technical Report Series, pp. 1-55, 2000.

[17] Yue Wu, Biaobiao Zhang, Jiabin Lu & K. L. Du," Fuzzy Logic and Neuro-fuzzy Systems: A Systematic Introduction", International Journal of Artificial Intelligence and Expert Systems (IJAE), vol 2, Issue 2,pp.47-80, 2011.

[18] Pankaj Kumar, "Minimum Weekly Temperature Forecasting using ANFIS", Computer Engineering and Intelligent Systems , vol 3, No.5, pp.1-6, 2012.

[19] W.T. Dent and J.A. Swanson, ''Forecasting with Limited Information : ARIMA Models of the Trailer on flactor Transportation Market'', Jasa, Journal of the America Statistical Association , Jun , vol 73. No. 326, pp. 295-298, June, 1978.

[20] Box, G.E.P. and Jenkins, G.M., "Time Series analysis Forecasting and Control'' . San Francisco, Holden Day , pp. 300-320, 1976 .

[21] A.C. Harrvy, ''Time series models'', Philip Allan, Britain, pp. 1-15, 1981.

[22] K,H,Abd Al-Zahra, "Computerized System For Forecasting of Temperature & Rain in Iraq", A thesis in Msc. Degree , College of Science , University of Basra , pp. 29-31,. 2003

[23] Neville Davies and Joseph D. Petruccelli , ''On the use of the general Partial Auto correlation function for order Determination in ARIMA (p,q) Processes'' , Jasa journal of the America Statistical Association , vol.79. No.386 , pp. 374-377,June,1984.

[24] Robert S. Pindyck, D.L. Rubinfeld , ''Econometric Models and Economic Forecasts '', U.S.A McGraw-Hill, 1976.

[25] Liu , Lon-Mu," Time Series Analysis and Forecasting "University of Illionis at Chicago ,Copyright Scientific Computing Associates Corp. ,U.S.A ,2006.