International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 03, 2020
ISSN: 1475-7192

Box- Jenkins Methodology Used in Predict The Numbers of Patients with Diabetes in DiyalaProvince

Dr.HaithamYaqubYousif[1]

Abstract:

The sugar disease hadknown since ancient times, where a Greek physician Aritaos observed in early 200 BC that some patients develop symptoms of frequent urination and extreme thirst has been named this phenomenon "policyholders" or "Aldiabats" which is a Latin word that means going to the bathroom chair, have discovered the scientists Joseph von Mereynj Oscar Menkoski 1889 the role of the pancreas in diabetes when they removed the pancreas completely dogs, where they showed signs and symptoms of diabetes led to their deaths after a brief period, and in 1910 discovered the world Sir Edward drinkers Scheffer said diabetic patients suffer a shortage of a chemical one produced by the pancreas called insulin. The diabetes at present from more diseases prevalent in the entire world advanced him and developing, as the number of people infected with the disease increasing, and because of the seriousness of this disease and how to predict the numbers of infected and build a standard form to him, it is here came the idea of this research, which was intended to use one of the advanced methods of time series and of methodology Box- Jenkins analysis to predict the numbers of people with diabetes in Diyala province, through the examination of a number of models and choose the best model among these models to represent monthly from 2009-2012 series and thus to predict until the end of 2014.

Keywords: *Stationary, Partial autocorrelation Function PACF, Estimation, Forecasting*

I. Introduction:

Diabetes is defined as an imbalance in the process of sugar metabolism, which leads to an abnormally high level of sugar (glucose) in the blood for various reasons that may be psychological or organic, or due to excessive intake of sugars or due to genetic factors, and it occurs as a result of an imbalance in the secretion of insulin From the pancreas, the amount of insulin that is secreted may be less than what is required, or there is a complete stop of its production, and this condition is called "insulin insufficiency", or the amount secreted is large in some cases, such as individuals with obesity, but there is resistance from tissues and cells in the body that exceeds the function of insulin This condition is called "insulin resistance". In both cases, glucose is unable to enter the cells, which leads

[1] *College of Administration and Economics, University of of Diyala*

to its accumulation in the blood and the possibility of its appearance in the urine. Over time and with the increase in the accumulation of sugar in the blood instead of entering the cells of the body, it may lead to chronic complications. Some parts of the body, such as the tiny blood vessels in the retina, kidney sacs, and those that supply nerves.

For the purpose of shedding more light on the incidence of diabetes - and to the extent this research allows - the following points were addressed:

First: the types of diabetes mellitus

There are two types of diabetes: (National reference for diabetes education, Kingdom of Saudi Arabia, Ministry of Health, 2011, p. 23)

1. Diabetes mellitus type 1: It is diabetes that depends on its treatment on insulin, and usually affects young children and young people under 20 years old, and this type is characterized by the inability of the pancreas to secrete insulin, and this type needs:

A- Insulin

B- Diet

C- Sports

It is necessary for patients to understand that this type does not respond to treatment with pills.

2. Diabetes mellitus type 2: It is the most common type and constitutes 90% of patients with diabetes. It is a non-insulin dependent disease, and it is the most prevalent type in adults over forty years old or overweight and sometimes affects children over ten, and it occurs as a result of inability The body makes enough insulin hormone, or there is enough insulin, but it is ineffective, which results in high blood sugar. This type needs:

A- Diet

B- Sports

C- Beans

Second: Symptoms of diabetes

Symptoms are varied among people with diabetes, some of which are general (Rewiha, Dr. Amin, 1973, p.13), the most important of which are:

Fatigue and weakness (weakness of forces)

Shivering of the limbs

Lack of desire to work

• Anxiety, mental disorder, anxiety, impaired memory, and stuttering

Vertigo

Nausea

Hunger and low body weight

• Thirst, itching, and tongue pain

• Frequent urination with feeling or feeling the need to urinate, despite the emptiness of the bladder, and irregular bowel movements.

The disorder of the metabolism in diabetes mellitus bad effects on all types of organs and tissues in the body, and this is not surprising .. because these all cause nutritional disorders due to the abnormal deviation in blood sugar.

Research Value:

The importance of the research lies in developing an appropriate model used to predict the numbers of people expected to have diabetes in Diyala Governorate, using the time-series analysis method and relying on the Box-Jenkins methodology, and then predicting these numbers until December 2014.

II. Research Methodology:

The descriptive and analytical method was used in the completion of this research, by reviewing many sources that dealt with the Box-Jenkins methodology in analyzing time series. It has advanced capabilities in such analysis.

Research Aims:

It can be said that the main goal of the research is to provide a sound and accurate statistical tool for predicting the number of people with diabetes. We summarize the following:

1. The possibility of applying the Box-Jenkins method in predicting the number of people expected to have diabetes.

2. Developing a standard model to predict the number of people expected to have diabetes.

3. Predicting the number of people expected to have diabetes until December 2016

Among these main objectives above, sub-goals can be added to benefit the health sector and decision-makers in that sector, namely:

A- Supporting the efforts of decision-makers in the health sector when they plan to control the health system on a monthly basis.

B- Helping to control the accuracy of one of the most important scientific indicators in the health sector, which is the number of people with diabetes.

C- Determining carefully the monthly needs of medicines and medical supplies in the health sector of Diyala Governorate.

D- Presenting a study for the Arab Library, which includes both the theoretical and practical framework in a simplified manner for the Box-Jenkins methodology in time series analysis, especially since the Arab library suffers a relatively shortage in that.

Research Problem:

The research problem lies in answering the following questions:

1. Does the diabetes time series include the general trend vehicle?

2. Does the time series for the number of people with diabetes include the seasonal component?

3. Does the time series of the number of people with diabetes suffer from the problem of heterogeneity of variance?

4. Does the time series for the number of people with diabetes in Diyala governorate have a normal distribution?

5. Is the appropriate model for the monthly series of numbers of people with diabetes subject to one of the self-regressive models and moving averages?

6. Is the appropriate model for the monthly series of numbers of people with diabetes subject to one of the self-regressive models and integrative moving averages?

7. How to determine the best model among the Box-Jenks models to represent the monthly series of numbers of people with diabetes.

Research Hypothesis:

To answer the research questions, the hypotheses were formulated as follows:

The first hypothesis:The time series of numbers of people with diabetes does not have a problem of heterogeneity

The second hypothesis:The monthly series of numbers of people with diabetes does not include the general trend vehicle

The third hypothesis:The time series for the number of people with diabetes in Diyala governorate has a normal distribution

Fourth hypothesis:The possibility of applying self-regression models and moving averages in the representation of the monthly series of numbers of people with diabetes.

Fifth hypothesis:The possibility of applying self-regression models and integrative moving averages in the representation of the monthly series of numbers of people with diabetes

Research Limits:

The variables for the numbers of people with diabetes and according to the age group greater than 15 years, which were obtained from the records of Baquba General Hospital for the period 2009-2012, were used.As shown in the following table:

Table (1)

The number of people with diabetes, according to the months

	2009	2010	2011	2012
1	144	197	207	186
2	148	242	196	166
3	150	117	150	146
4	142	254	183	178
5	150	239	176	168
6	160	180	102	101
7	168	239	163	82
8	81	234	175	90
9	42	235	155	100
10	315	164	169	90
11	145	186	191	168
12	229	211	157	107

Search Plan:

This research was divided into four sections, which included the following:

Topic I (current): the general framework of the research.

The second topic: The standard theoretical framework: It included a brief presentation of the Box-Jenkins methodology in analyzing time series, and the most important tools used in it.

The third topic: Presentation, analysis and discussion of results using EVIEWS 10.0 program and graphic interfaces.

The fourth topic: Presenting the most important conclusions reached in the body of the research, and its recommendations.

III. The theoretical standard framework

Box-Jenkins' methodology means that the methodology that George Box and Gwilyn Jenkins applied to time series in 1970, that this methodology gives a strong strategy to solve many problems of the time series, and gives accurate predictions for time series. What is sometimes called ARIMA models is an organized method for constructing and analyzing models in order to find the optimal model from among the models based on time-series data, as the optimal model is obtained with minimum errors, and it is considered an optimal model if all the information in it is statistically significant and errors in The form distributed independently.

For the purpose of dealing with the Box-Jenkins methodology in time series analysis, this topic has been divided into the following:

Considerations for applying the Box-Jenkins models

Before applying the Box-Jenkins methodology in time series analysis, the following concepts must be studied because of their close relationship with this methodology:

First: Stationary

All economic applications assume that time series have the characteristic of stability or stationary, and that the first step in applying the Box-Jenkins method is to ensure that the time series is stable. , 2004, p. 648)

It is possible to identify whether the time series is stable or unstable by viewing the graph of the studied phenomenon, or by viewing the Auto correlation function or the Partial auto correlation function, as its values do not approach zero after the second or third displacement, but rather Its values remain large for a number of shifts, and the instability of the time series is due to one of the following reasons:

1. There is a general trend in the time series

2. The presence of seasonal fluctuations in the time series

3. The instability of variance and the arithmetic mean of the series in question.

The following is a brief presentation of the concepts of autocorrelation and partial autocorrelation and how to use it to detect instability of the time series.

Second: Auto correlation function

The autocorrelation function plays a major role in testing the stability of the time series through the following:

1. Confidence intervals

The autocorrelation coefficient falls between -1,1). The stability of the chain here requires that the estimated value of the autocorrelation coefficient be equal to zero or not fundamentally different from it for any time gap (Attia, Abdul-Qadir Muhammad, 2004, p.650). In the event that the series data enjoys stability, the autocorrelation coefficients of the sample often have a normal distribution, the arithmetic mean of which is equal to zero and the variance of (1 / n). Hence, the confidence limits for a large sample are**:**

$$\pm 1.96\sqrt{1/n}$$

And if the estimated value of the autocorrelation coefficient falls within these limits, that is:

$$-1.96\big/\sqrt{n} \;\le\; \rho_k \le\; 1.96\big/\sqrt{n}$$

And below the level of significance (95%), the confidence limits are as follows:

$$\Pr\left\{-1.96\big/\sqrt{n} \;\le\; \rho_k \le 1.96\big/\sqrt{n}\right\} = 0.95$$

Therefore, we accept the null hypothesis that this parameter is equal to zero and therefore the time series is considered stable.

2. Box test, Pierce

A test can be performed to reveal the significance of the autocorrelation coefficient as a group, using the Box, Pierce test statistic, which takes the following formula (Attia, Abdul Qadir Muhammad, 2004, p. 653):

$$\mathbf{Q=n}\sum_{k=1}^{m} \hat{P}_k^{\,2}$$

So that:

m: represents the number of time slots, k: represents the time slot number

For large samples, the Q statistic has a chi-square distribution with a degree of freedom equal to m. If the calculated Q value is greater than the tabular value, we reject the null assumption that all autocorrelation coefficients are equal to zero and the series is unstable.

3. The Ljung-Box Test

There is another alternative statistic used to perform the same Box, Pierce test, called the Ljung-Box statistic, which takes the following formula (Attia, Abdul Qadir Muhammad, 2004, p. 654):

$$\mathrm{LB} = n(n+2)\sum_{k=1}^{m} \frac{\hat{P}^2{}_k}{n-k}$$

+

Which has a chi-square distribution with a degree of freedom equal to m and gives better results than Q in the case of small-sized samples, although it is suitable for large-sized samples.

Third: PACF Partial Autocorrelation Function

The partial autocorrelation function PACF represents the relationship between successive values of a variable during two different time periods, assuming the stability of other periods, and the partial autocorrelation function is symbolized by the symbol, so the partial autocorrelation function between Yt, Yt-k indicates the correlation between them, excluding other Yt values that fall Between the two periods (t, tk) (Berri, Dr. Adnan Majid Abdel Rahman 2002, p.11), the mathematical equation for the partial autocorrelation coefficient can be formulated from the autocorrelation equation as follows:

$$r_{kk} = \begin{cases} 1, & k=0 \\ r_1, & k=1 \\ \dfrac{\begin{vmatrix} 1 & r_1 & \cdots & r_{k-2} & r_1 \\ r_1 & 1 & \cdots & r_{k-3} & r_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ r_{k-1} & r_{k-2} & \cdots & r_1 & r_k \end{vmatrix}}{\begin{vmatrix} 1 & r_1 & \cdots & r_{k-2} & r_{k-1} \\ r_1 & 1 & \cdots & r_{k-3} & r_{k-2} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ r_{k-1} & r_{k-2} & \cdots & r_1 & 1 \end{vmatrix}}, & k=2,3,\ldots \end{cases}$$

To calculate the partial autocorrelation function for the sample repeatedly, this is done through the following**:**

$$r_{00} = 1, \quad \text{by definition}$$

$$r_{11} = r_1$$

$$r_{kk} = \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j}, \quad k = 2,3,\ldots$$

where:

$$r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-1}, \quad j = 1,2,...,k-1$$

$$V\left(r_{kk}\right) \cong \frac{1}{n}, \quad k > 0$$ 1.

2.For large values, they have almost a normal distribution, and therefore we can perform the following test:

$$H_0 : \phi_{kk} = 0$$
$$H_1 : \phi_{kk} \neq 0$$

And by using the statistics:

$$\frac{\left|r_{kk}\right|}{n^{-\frac{1}{2}}} = \sqrt{n}\left|r_{kk}\right|$$

$\sqrt{n}\left|r_{kk}\right| > 1.96$ if it was H_0 and refusesAnd that's at a significant level $\alpha = 0.05$

3.Under the hypothesis, $H_0 : \phi_{kk} = 0, \forall k \ corr\left(\phi_{kk}, \phi_{k-s,k-s}\right) \cong 0, s \neq 0$ where s represents the number of the other time gap.

4.The variances of the autocorrelation function of the sample are estimated as follows:

$$\hat{V}\left(r_{kk}\right) \cong \frac{1}{n}, \quad k > 0$$

Formulation of Box-Jenkins models

Box-Jenkins's method relies on a set of (Stochastic models), including: - (Berri, Dr. Adnan Majid Abdel-Rahman 2002, p. 22)

First: the p-degree autoregressive model

A self-regression model of p-degree can be written according to the following formula:

$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + e_t$

whereas:

P: score of the autoregressive model

Yt: a variable representing the time series

Yt-p: A variable representing the slowed time series p of scores

$\phi_0, ..., \phi_p$: Represents the parameters of the model to be assessed.

e_t : Represents the random error component, which is assumed to have a normal distribution with mean of zero and constant variance equal to σ^2

Second: The MA (q) moving average model: Moving Average Model

It is said that yt is a moving averages process of finite order q if it can be expressed as**:**

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - ..... - \theta_q \varepsilon_{t-q} ; t = 0, \pm 1, ...$$

Where: q: represents the degree of moving media

ε_t Process of quiet disturbances

$\theta_1, ..., \theta_q$ Constants representing the parameters or parameters of the model

These models are referred to as MA (q) and they are always "static operations, because the order of the form q is limited. The models MA (q) are inverted if the roots of the characteristic equation are".

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - .... - \theta_q B^q = 0$$

They are all outside the unit circle.

Third: ARMA (p, q): Autoregressive-Moving Average Model

The self-regression model of a moving average of degree can be written as follows**:**

$$z_t = \delta + \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \cdots - \theta_q a_{t-q}$$

whereas:

$a_t \sim WN\left(0, \sigma^2\right)$ White noise series

$-\infty < \delta < \infty$: Fixed parameter representing the plane

$\phi_1, \phi_2, \ldots, \phi_p$: Autoregressive Parameters

$\theta_1, \theta_2, \ldots, \theta_q$:Moving AverageOperators : are the parameters of the moving average

Stages of applying the Box-Jenkins methodology in time series analysis

This topic has been devoted to presenting the Box-Jenkins methodology and the basic stages of applying this method in time series analysis by addressing the following points: - (Shaarawi, Dr. Samir Mustafa, 2005, p.257)

First: Identification

The first stage of the modern analysis of time series using the Box-Jenkins methodology is the identification of the appropriate initial model for the observed time series data, and by identifying the model, the selection of the three model ranks (p, d, q) means that:

d: represents the rank or degree of differences necessary to accommodate the time series

P: Represents the number of previous observation boundaries that should be included in the appropriate prototype

Q: represents the number of quiet disturbance variables that should be included in the appropriate model. The recognition tools are represented by three main tools:

A- Autocorrelation function

B- Partial Autocorrelation function

T- The form of the correlation between the coefficient of each previous function and the length of the Correlogram gap.

The recognition stage is one of the most difficult and important stages of analysis, as it requires skill, experience, practical practice and a measure of personal judgment for the researcher, and the following is an explanation of the ranks of the model (p, d, q)

1. Determine the differences

Most of the time series that arise in the different fields of application usually show a certain type of non-rest, whether in its mean or in its variation, or in both together. In fact, non-staticness may occur in more than one way. We have mentioned that the judgment on the static or non-static processes comes through examining the roots of the characteristic equation, so if all the roots of this equation are clearly outside the unit circle, then this means that the process or series is static and in this The case the autocorrelation function rapidly fades as the time gaps increase. But if some or all of the roots are located within the unit circle, then this means that the process or chain is not static, but if one of the roots is located on the unit circle, this means that the process or series is not static but is homogeneous. This kind of non-static is characteristic of most of the actual strings that arise in practical applications, and such strings can be transformed into static strings by using mathematical transformations. After settling the time series, the value of d is determined according to the following steps: -

The first step: drawing the time curve of the original series and drawing the autocorrelation function for it. If the drawing does not show any kind of non-staticity in the mean and the variance, and the autocorrelation function of the sample [r (k)] fades rapidly with the increase in the time gap k, then it is not taken Any differences of the series and d = 0, and we proceed directly to looking for the values of p, q.

Second step: If the drawing shows non-staticity with respect to the mean and the autocorrelation function of the sample is slowly fading, then the first differences of the series must be taken, and then we draw the time curve and the sample autocorrelation function for the first series of differences. If the graph shows static in the series and the sample's autocorrelation function is rapidly fading with the increase in the time gap, then no other difference is taken and p = 1 and we proceed directly to searching for q values.

The third step: If it is observed from the time curve of the series of differences, that the series of differences still suffers from the lack of inactivity of its basic characteristics and that the self-correlation function of this series slowly fades away, so the second differences of the original series must be taken. Then we draw the time curve and the autocorrelation function for the second series of differences. If static in the time curve and rapid decay in the autocorrelation function is observed, then we stop taking the differences and form d = 2 and proceed directly to searching for the values of p, q.

Usually the value of d is small (zero, one or two), and we would like to emphasize the risk of taking unnecessary differences. Although the differences of any static series give a static series, taking unnecessary differences leads to a model that contains unnecessary features and a pattern More complex self-association. In addition, taking unnecessary differences usually results in large series variance.

2. Determining the self-regression and moving averages ranks

After determining the rank of the differences necessary for chain housing, the rank of the autoregressive segment p and the order of the moving averages segment q must be determined. The autocorrelation function and the partial autocorrelation function are one of the effective keys in distinguishing between AR (p) models, MA (q) models, and ARMA (p, q) mixed models and the ranking of each. Before the explanation and elaboration, it may be necessary to recall the basic properties of these two functions and the type of each of them for the most important models that belong to these three families.

Second: Estimation

After completing the stage of identifying the prototype suitable for the available data, the parameters of this model must be estimated using one of the known methods in statistical theory, the most important of which is the least squares method, which can be derived as follows:

For ARMA (p, q) models, they are written as:

$$\phi_p(B)(z_t - \mu) = \theta_q(B)a_t, \quad a_t \sim N(0, \sigma^2)$$

Where $\phi_p(B)$ there are no common roots between them and the roots of the equation $\theta_q(B) = 0$ are all outside the unit circle (the inversion condition). And the error model a_t is:

$$a_t = \frac{\phi_p(B)}{\theta_q(B)}(z_t - \mu)$$

Whereas, the right-hand side can be viewed as a function of features $\boldsymbol{\varphi} = \{\phi_1, \phi_2, \ldots, \phi_p\}$ $\boldsymbol{\theta} = \{\theta_1, \theta_2, \ldots, \theta_q\}$ and μ written as:

$$a_t(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu) = \frac{\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right)}{\left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^p\right)}(z_t - \mu)$$

The conditional minimum squares method and given observations $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$ depend on miniaturizing the function

$$\min_{\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu} S_c(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu) = \sum_{t=p+1}^{n} a_t^2(\boldsymbol{\varphi}, \boldsymbol{\theta}, \mu | \mathbf{z})$$

Third: Diagnostic checking

This stage is one of the most important and most dangerous stages of the analysis wherein it is assured that the initial model is appropriate and thus can be used in forecasting, or in which this model is modified based on the results of the examinations and tests that take place at this stage, and in this case the modified form must be subjected to all the examinations and tests Which we will talk about in detail here. That is, the diagnostic phase is in its essence a problem of improving or developing the prototype in order to be more suitable for the available data, and it is a complex problem with multiple dimensions and aspects. The time series analyst, taking into account all these dimensions and aspects, must end with a better model For available data.

The diagnosis of the form in general depends on the conduct of many examinations and tests, the most important of which are:

1. Static analysis

The model is considered static if the roots of the characteristic equation are all outside the unit circle, in other words if the absolute value of each of these roots is greater than the integer one, then this indicates the staticity of the random process that gave birth to the observed series.

2. Reflection analysis

The model that represents the time series is characterized by the character of reflection if the roots of the equation are outside the unit circle, that is, if the absolute value of each of these roots is greater than the integer one, then this indicates the reflection of the original model.

3. Residual analysis

The steps involved in analyzing residues are:

A- Draw remainder

The first and important step in analyzing residues is the graphic signature of these values as a time series, and this drawing is a necessary step that cannot be dispensed with by conducting statistical checks and tests. The drawing of the residues shows the basic features of the remainders, such as the general trend, dispersion and anomalous data in a way that the statistical tests may not be able to show. And if the prototype was good, this means that it was able to absorb all the patterns and regular movements in the data, leaving the remainder free of such patterns and movements, and then the remainder on the graphic signature sheet must oscillate with a constant dispersion around zero as a midline parallel to the time axis, as well as The figure should appear random, devoid of any information that can be used in predicting the time series under study.

B- Examine the residual self-correlation function

Here each sample autocorrelation coefficient is examined separately, and then the sampling distributions of these factors must be examined. The study by Anderson (1942) showed that if the model is appropriate - that is, if the errors represent purely random variations - then the autocorrelation coefficients for the medium and large samples are non-correlated and follow a normal distribution with a standard deviation of n-0.5 and thus the autocorrelation coefficient of the remainder And at a certain time slot, which falls outside the interval (), the corresponding theoretical correlation coefficient is significant to zero.

Fourth: Prediction

Prediction is the last stage of the Box-Jenkins methodology, and it is usually the final goal of time series analysis, and it is not possible to move to this stage until after the prototype passes all the diagnostic tests and tests that were previously exposed, where the following are tested:

1. The teams' significance test

This criterion relies on "Ex-Post Forecast" in testing the model's ability to predict. If the expected value is equal to the actual value of the predicted variable, or the difference between them is not substantial, then the model's ability to predict is high, but if the difference between them is substantial, then This indicates a lack of the model's ability to predict.

2. The unequal coefficient of methyl

The higher the Thiel factor value than the correct one, the lower the model's predictability.

3. The Jans coefficient

This parameter measures the model's ability to predict during the sample period and during the post-sample period, and its value ranges between zero and infinity. The higher the value of this parameter, the more this indicates the weakness of the model's ability to predict, and when it is equal to the correct one, this means that the model's ability to predict in the past is equal with it in the future.

1. Average square error

This scale is used to compare more than one model on predictive power, and the best model is the one with the lowest mean of squares of error.

IV. Presentation, analysis and discussion of the results

In this study, the results of the time series for the number of people with diabetes were presented, analyzed and discussed, depending on the EVIEWS 10.0 program, as follows:

First: the hypothesis testing phase:

Before starting the application of the Box-Jenkins method, it is necessary to test the hypotheses that were discussed in the general framework of the research, as follows:

The first hypothesis: the homogeneity test of the time series variance of the numbers of people with diabetes in Diyala governorate for the period 2009: 1-2012: 12.

For the purpose of detecting the homogeneity of the series variance in relation to the numbers of people with diabetes, the time series has been drawn and is shown as follows:

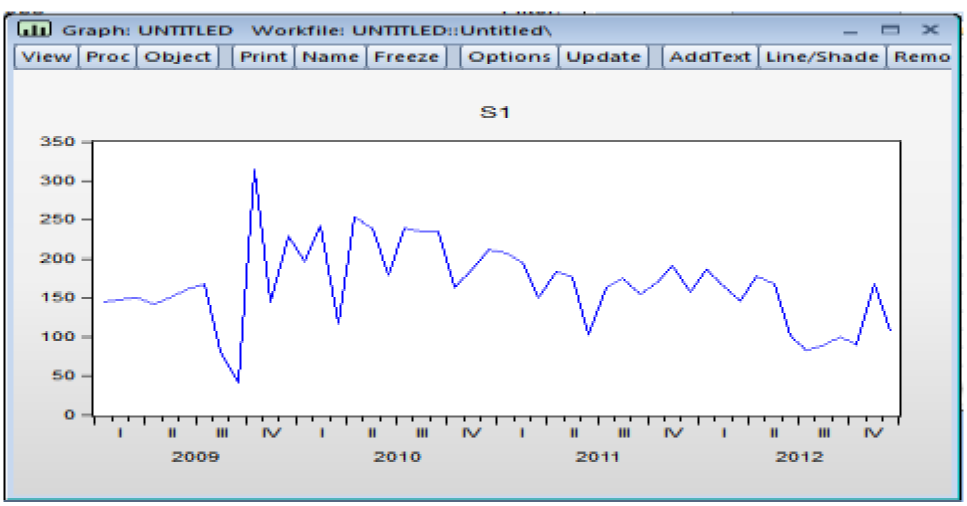


Figure (1)

Diabetes chart of the number of people with diabetes

It is noticed from the chart of the number of diabetic patient series that there is no difference in the homogeneity of the series variance, and to make sure of this, the tests for the homogeneity of the series variance test including each one who tested Bartlett, Levene, Brown-Forsythe's test, to test the null hypothesis that variance is homogeneous in the series, where the test result was as follows:

Test for Equality of Variances of S1
Categorized by values of S1
Date: 02/26/16 Time: 21:48
Sample: 2009M01 2012M12
Included observations: 48

Method	df	Value	Probability
Bartlett	3	3.716543	0.2937
Levene	(3, 44)	1.578157	0.2081
Brown-Forsythe	(3, 44)	1.462566	0.2378

Category Statistics

S1	Count	Std. Dev.	Mean Abs. Mean Diff.	Mean Abs. Median Diff.
[0, 100)	5	20.02498	14.00000	11.40000
[100, 200)	33	27.02517	20.85950	20.51515
[200, 300)	9	14.87261	11.03704	10.33333
[300, 400)	1	NA	0.000000	0.000000
All	48	52.70672	17.86869	17.22917

Bartlett weighted standard deviation: 24.65446

Figure (2)

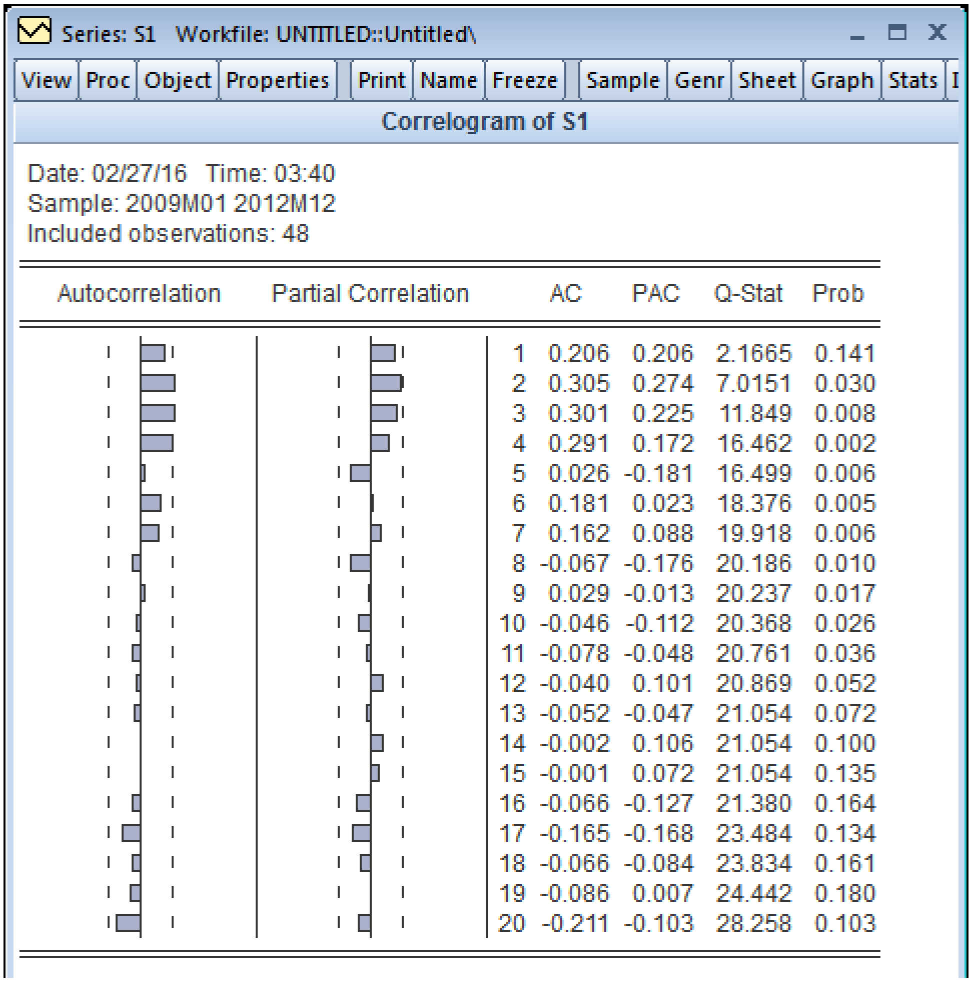
Test results for the homogeneity of time series variance of the numbers of people with diabetes

It is noted from the results of the test for the uniformity of variance of the time series for the numbers of people with diabetes, that all the tests confirm the acceptance of the null hypothesis that the series is homogeneous in the variance, the fact that the probability values for the tests are equal to (0.2937) for Bartlett's test and (0.2081) for the Levene test and (0.2378) for Brown-Forsythe test is greater than the level of significance (5%).

The second hypothesis: a test of the general trend of the time series of numbers

People with diabetes in Diyala governorate

From the diagram for the number of people with diabetes, shown in figure (1), it is noticed that the series does not suffer from the general trend, and this is an indication of the stability of the time series, and to make sure of this, the values of the autocorrelation function and the partial self-correlation were found with the diagram for them as shown in the following figure:



Series: S1 Workfile: UNTITLED::Untitled\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Graph | Stats

Correlogram of S1

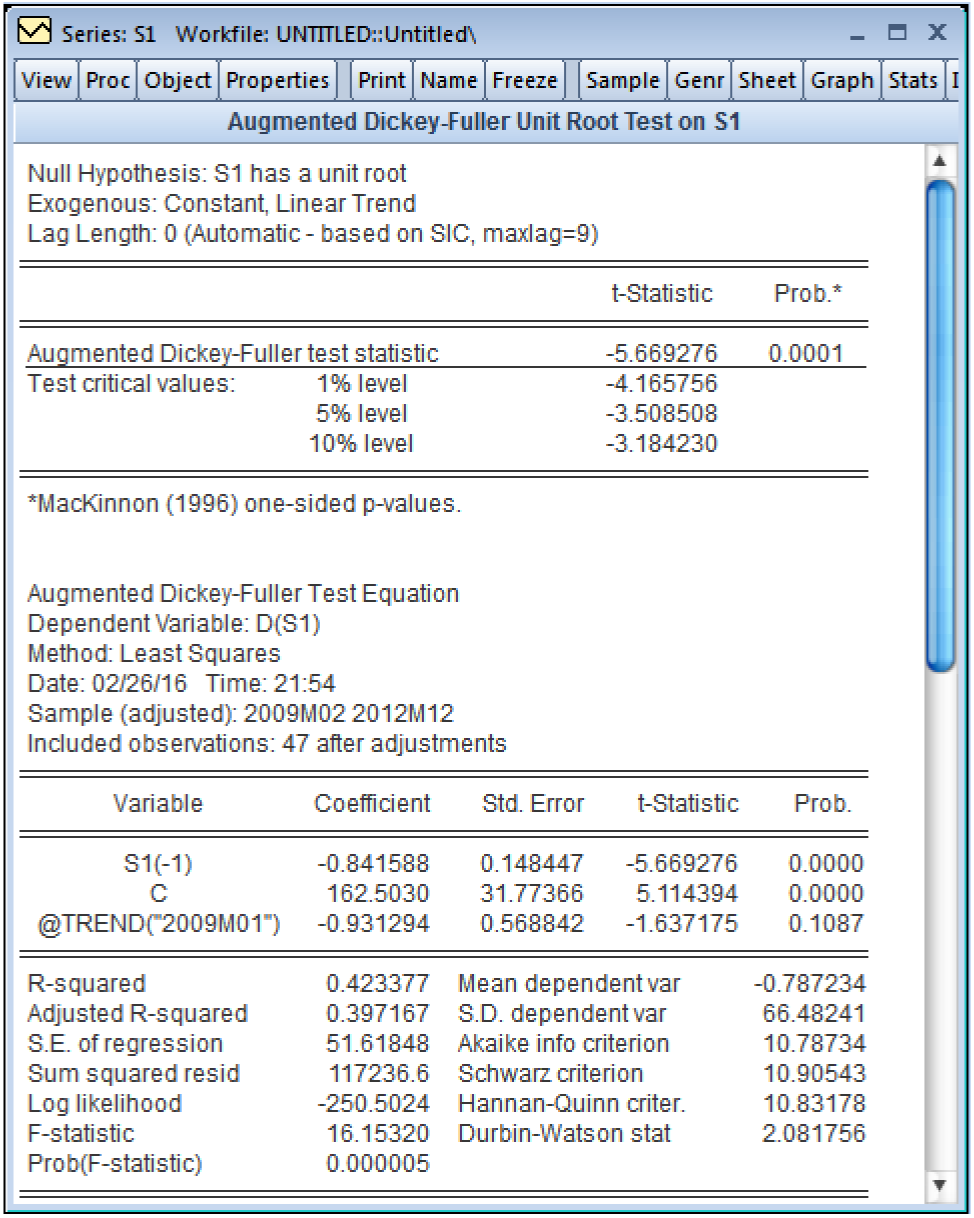
Date: 02/27/16 Time: 03:40
Sample: 2009M01 2012M12
Included observations: 48

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.206	0.206	2.1665	0.141
		2	0.305	0.274	7.0151	0.030
		3	0.301	0.225	11.849	0.008
		4	0.291	0.172	16.462	0.002
		5	0.026	-0.181	16.499	0.006
		6	0.181	0.023	18.376	0.005
		7	0.162	0.088	19.918	0.006
		8	-0.067	-0.176	20.186	0.010
		9	0.029	-0.013	20.237	0.017
		10	-0.046	-0.112	20.368	0.026
		11	-0.078	-0.048	20.761	0.036
		12	-0.040	0.101	20.869	0.052
		13	-0.052	-0.047	21.054	0.072
		14	-0.002	0.106	21.054	0.100
		15	-0.001	0.072	21.054	0.135
		16	-0.066	-0.127	21.380	0.164
		17	-0.165	-0.168	23.484	0.134
		18	-0.066	-0.084	23.834	0.161
		19	-0.086	0.007	24.442	0.180
		20	-0.211	-0.103	28.258	0.103

Figure (3)

The interface for the values of each of the autocorrelating functions and partial autocorrelation with the chart for them

It is noticed from Figure (3) that most of the values of the autocorrelation function fall within the limits of confidence, in addition to that the test statistic for Box-Pierce of (28,258) is not significant due to the fact that the probability value of the test of (0.103) is greater than the level of significance (5%). It means that the chain is free from the general direction component, and for the purpose of making sure that the first series of differences does not include the general direction component, the extended Dickie-Fuller test was used as follows:



Series: S1 Workfile: UNTITLED::Untitled\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Graph | Stats | I

Augmented Dickey-Fuller Unit Root Test on S1

Null Hypothesis: S1 has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.669276	0.0001
Test critical values:	1% level	-4.165756	
	5% level	-3.508508	
	10% level	-3.184230	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(S1)
Method: Least Squares
Date: 02/26/16 Time: 21:54
Sample (adjusted): 2009M02 2012M12
Included observations: 47 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
S1(-1)	-0.841588	0.148447	-5.669276	0.0000
C	162.5030	31.77366	5.114394	0.0000
@TREND("2009M01")	-0.931294	0.568842	-1.637175	0.1087

R-squared	0.423377	Mean dependent var	-0.787234
Adjusted R-squared	0.397167	S.D. dependent var	66.48241
S.E. of regression	51.61848	Akaike info criterion	10.78734
Sum squared resid	117236.6	Schwarz criterion	10.90543
Log likelihood	-250.5024	Hannan-Quinn criter.	10.83178
F-statistic	16.15320	Durbin-Watson stat	2.081756
Prob(F-statistic)	0.000005		

Figure (4)

Extended Dickey-Vollier test results for chain-infected subjectsDiabetes in Diyala governorate

It is noted from the results of the extended Dickey - Fuller test that the probability value of the test (0.0001) is less than the level of significance (5%). This means rejecting the null hypothesis that the series includes a unit root, meaning that the chain is stable and does not contain the general trend component.

The third hypothesis: the normal distribution test for the number of people with diabetes in Diyala Governorate

For the purpose of testing whether the first series of differences in the number of people with diabetes have a normal distribution, the tests available in the EVIEWS 10.0 program were used where the results were as follows:

Series: S1 Workfile: UNTITLED::Untitled\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Graph | Stats | I

Empirical Distribution Test for S1
Hypothesis: Normal
Date: 02/28/16 Time: 18:54
Sample: 2009M01 2012M12
Included observations: 48

Method	Value	Adj. Value	Probability
Lilliefors (D)	0.114675	NA	> 0.1
Cramer-von Mises (W2)	0.089076	0.090004	0.1539
Watson (U2)	0.088922	0.089848	0.1284
Anderson-Darling (A2)	0.501283	0.509605	0.1976

Method: Maximum Likelihood - d.f. corrected (Exact Solution)

Parameter	Value	Std. Error	z-Statistic	Prob.
MU	166.2083	7.607560	21.84779	0.0000
SIGMA	52.70672	5.436283	9.695360	0.0000

Log likelihood	-257.9167	Mean dependent var.	166.2083
No. of Coefficients	2	S.D. dependent var.	52.70672

Figure (5)

The results of the normal distribution test for the data of the number of people with diabetes in Diyala Governorate

It is noticed from Figure (5) that the probability values for testing each of (Lillefors, Cramer-von mises, Watson, Anderson-Darling) are greater than the level of significance (5%). This means accepting the null hypothesis that the time series data for the numbers of people with diabetes in the province Diyala follows the normal distribution, and that there are no anomalies in the data that may affect the results of the estimate and thus the prediction process.

The second stage: the application of the Box-Jenkins methodology in the chain analysis

The time period for the numbers of people with diabetes in Diyala Governorate

After passing the hypothesis testing phase, we now apply the Box-Jenkins methodology in analyzing the time series for the numbers of people with high blood pressure in Diyala Governorate, as follows:

First: the model recognition stage:

The initial model for applying the Box-Jenkins methodology to the time series of numbers of people with diabetes can be identified through the graphical representation of each of the self-correlation and partial self-correlation functions shown in Figure (6), as it is noticed that the graph of the autocorrelation function appears to be discontinued after the time gap Fourth, in addition to the fact that the partial autocorrelation function gradually approaches zero in a reluctant way in the sign, which in turn leads to the weighting of the fourth-degree moving average model, which is symbolized by the symbol Ma (4).

Second: the assessment stage

After completing the initial recognition phase that is appropriate for the available data, the parameters of this model should be estimated using the least squares method, as follows:

Equation: UNTITLED Workfile: UNTITLED::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: S1
Method: Least Squares
Date: 02/29/16 Time: 08:18
Sample: 2009M01 2012M12
Included observations: 48
Convergence achieved after 19 iterations
MA Backcast: 2008M09 2008M12

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	165.8178	14.28189	11.61036	0.0000
MA(1)	0.172107	0.145071	1.186361	0.2420
MA(2)	0.278358	0.143170	1.944250	0.0584
MA(3)	0.261688	0.141311	1.851858	0.0709
MA(4)	0.297276	0.145848	2.038258	0.0477

R-squared	0.183082	Mean dependent var	166.2083
Adjusted R-squared	0.107090	S.D. dependent var	52.70672
S.E. of regression	49.80464	Akaike info criterion	10.75243
Sum squared resid	106661.6	Schwarz criterion	10.94734
Log likelihood	-253.0582	Hannan-Quinn criter.	10.82609
F-statistic	2.409222	Durbin-Watson stat	2.104291
Prob(F-statistic)	0.063860		

Inverted MA Roots	.41+.69i	.41-.69i	-.50-.46i	-.50+.46i

Figure (6)

Estimating the fourth-order moving average model for the time series

With regard to the numbersof people with diabetes in Diyala governorate

It is noted from the results of Figure (6) that the estimated model for the first-order moving averages can be written as follows:

$$y_t = \delta + \theta(B)\varepsilon_t \Rightarrow y_t = \delta + (1 - \theta B - \theta_2 B^2 - \theta_3 B^3 - \theta_4 B^4)\varepsilon_t$$

$$\Rightarrow y_t = 165.8178 + (1 - 0.172107B - 0.278358B^2 - 0.261688B^3 - 0.297276B^4)\varepsilon$$

Third: the diagnostic examination

Determining the best model for the time series for numbers of people with diabetes, which was estimated in the second step, depends on an important set of theoretical assumptions about the random process that generated the data, the general shape of the model, and the random changes, and for the purpose of determining whether the MA (4) model is a good model for representing the series of sufferers. Diabetes has been tested for the following:

1. Check the reversal for the fourth order moving average model

For the purpose of making sure that the first-degree moving average model possesses the characteristic of reflection, the unit circle was found and as shown in Figure (7) where it is noticed that the inverse root of the estimated model is located on the boundaries of the unit circle.

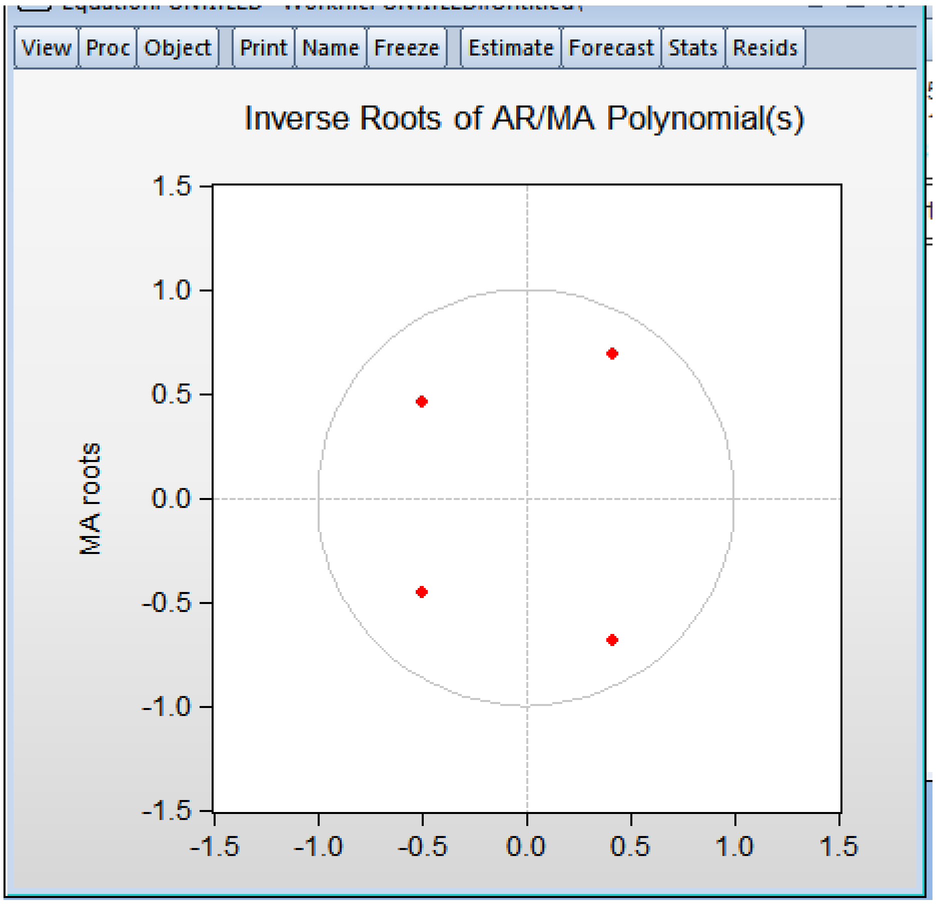


Figure (7)

Unit circle with respect to the inverse root of the first-order moving average model

1. Draw a diffusive remainder chart for the first-order moving averages model

Figure (8) below shows the remainder drawing, which seems devoid of all patterns and regular moves that can be used to improve the model, as the data oscillates randomly around the zero line.

International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 03, 2020
ISSN: 1475-7192

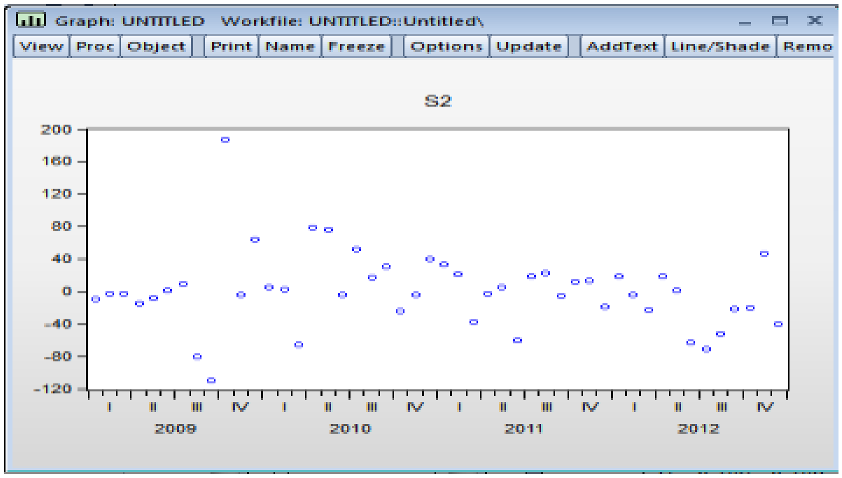
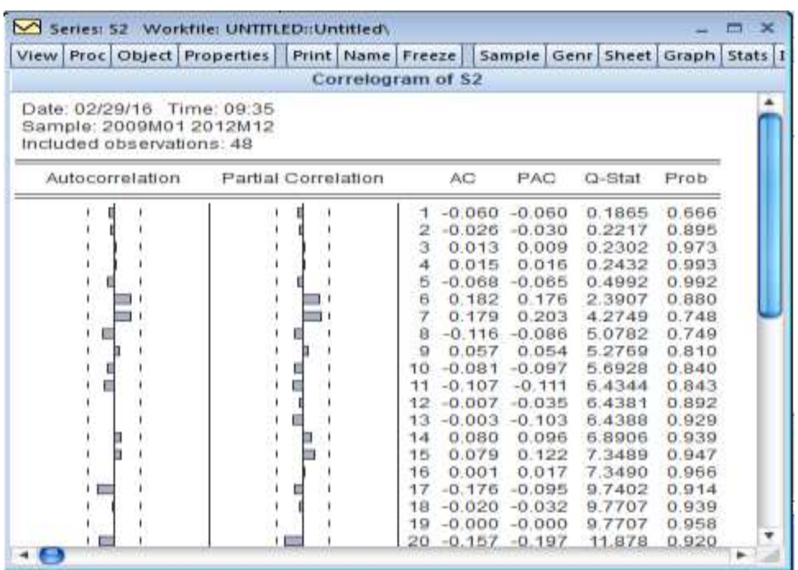


Figure (8)

Plot the spread points with respect to the rest of the fourth order moving average model

1. Examine the residual autocorrelation function for the fourth degree moving average model

Figure (9) shows a drawing of the autocorrelation function for the remainder for the model estimated for the fourth-degree moving average, where it is noticed that the values of the autocorrelation function fall within the limits of confidence and for a large sample of gaps, and this is an indication that the form of the autocorrelation function is devoid of protrusions, and this is an indicator Another good thing is that errors are purely random changes.



Series: S2 Workfile: UNTITLED::Untitled\

Correlogram of S2

Date: 02/29/16 Time: 09:35
Sample: 2009M01 2012M12
Included observations: 48

	AC	PAC	Q-Stat	Prob
1	-0.060	-0.060	0.1865	0.666
2	-0.026	-0.030	0.2217	0.895
3	0.013	0.009	0.2302	0.973
4	0.015	0.016	0.2432	0.993
5	-0.068	-0.065	0.4992	0.992
6	0.182	0.176	2.3907	0.880
7	0.179	0.203	4.2749	0.748
8	-0.116	-0.086	5.0782	0.749
9	0.057	0.054	5.2769	0.810
10	-0.081	-0.097	5.6928	0.840
11	-0.107	-0.111	6.4344	0.843
12	-0.007	-0.035	6.4381	0.892
13	-0.003	-0.103	6.4388	0.929
14	0.080	0.096	6.8906	0.939
15	0.079	0.122	7.3489	0.947
16	0.001	0.017	7.3490	0.966
17	-0.176	-0.095	9.7402	0.914
18	-0.020	-0.032	9.7707	0.939
19	-0.000	-0.000	9.7707	0.958
20	-0.157	-0.197	11.878	0.920

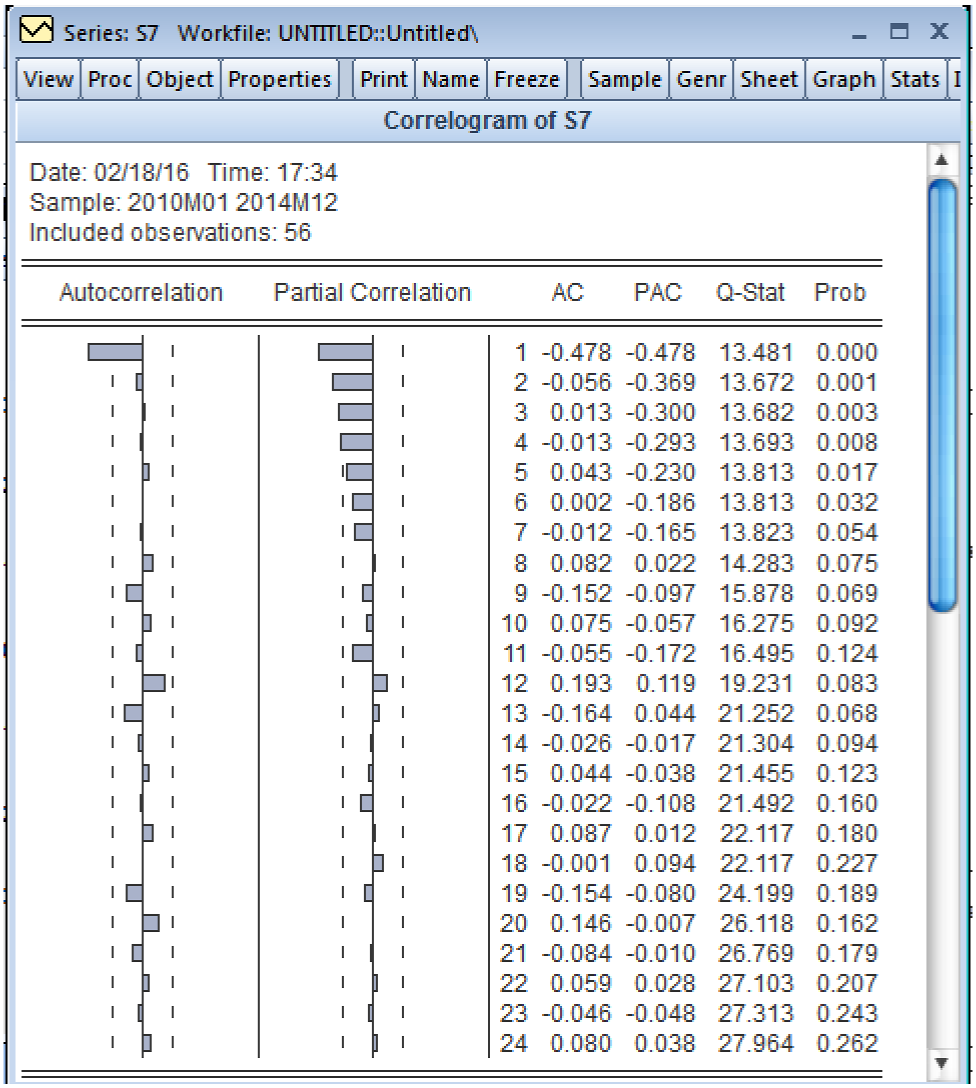
Figure (9)

Plot the autocorrelation function with respect to the remainder of the fourth order autoregressive model

It can also be noted from Figure (9) that the values of the Box-Pierce counts are not significant, and this is an indication that the residues are random, and this confirms the suitability of the estimated model.

1. Examine the first differences model of the remainder for the fourth degree moving average model

For the purpose of increasing the assurance that the residuals follow random changes, the residual first differences model must follow a first-order moving average model, and on this basis the self-correlation function was extracted with respect to the first differences of the remainders as follows:



	AC	PAC	Q-Stat	Prob
1	-0.478	-0.478	13.481	0.000
2	-0.056	-0.369	13.672	0.001
3	0.013	-0.300	13.682	0.003
4	-0.013	-0.293	13.693	0.008
5	0.043	-0.230	13.813	0.017
6	0.002	-0.186	13.813	0.032
7	-0.012	-0.165	13.823	0.054
8	0.082	0.022	14.283	0.075
9	-0.152	-0.097	15.878	0.069
10	0.075	-0.057	16.275	0.092
11	-0.055	-0.172	16.495	0.124
12	0.193	0.119	19.231	0.083
13	-0.164	0.044	21.252	0.068
14	-0.026	-0.017	21.304	0.094
15	0.044	-0.038	21.455	0.123
16	-0.022	-0.108	21.492	0.160
17	0.087	0.012	22.117	0.180
18	-0.001	0.094	22.117	0.227
19	-0.154	-0.080	24.199	0.189
20	0.146	-0.007	26.118	0.162
21	-0.084	-0.010	26.769	0.179
22	0.059	0.028	27.103	0.207
23	-0.046	-0.048	27.313	0.243
24	0.080	0.038	27.964	0.262

Figure (10)

Values of the autocorrelation function and the partial autocorrelation for the first residual differences

It is noticed from Figure (10) that the autocorrelation function is suddenly disconnected after the first time gap, and the partial autocorrelation function is reluctantly close to zero in the signal, and this is an indication that the first series of differences for the remainder follows a first-degree moving average model, and to make sure that The moving average parameter is significant (i.e. it does not differ from the correct one). The following regression was performed:

Equation: UNTITLED Workfile: UNTITLED::Untitled\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: D(S2)
Method: Least Squares
Date: 02/29/16 Time: 09:48
Sample (adjusted): 2009M02 2012M12
Included observations: 47 after adjustments
Convergence achieved after 18 iterations
MA Backcast: 2009M01

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.654470	0.595439	-1.099137	0.2776
MA(1)	-0.970516	0.034407	-28.20727	0.0000

R-squared	0.523472	Mean dependent var	-0.647294
Adjusted R-squared	0.512883	S.D. dependent var	69.84881
S.E. of regression	48.75013	Akaike info criterion	10.65291
Sum squared resid	106945.9	Schwarz criterion	10.73164
Log likelihood	-248.3435	Hannan-Quinn criter.	10.68254
F-statistic	49.43310	Durbin-Watson stat	2.160774
Prob(F-statistic)	0.000000		

Inverted MA Roots	.97

Figure (11)

Estimation of the first order moving average model for the first residual differences

It is noticed from the estimated model that the moving average parameter is significant, since the probability value of the test of (0.000) is less than the level of significance (5%), and this confirms the randomness of the remainder, in addition to that the estimated moving average model is characterized by the character of reflection.

Fourth: Prediction

It is noticed from the previous results of the diagnostic examination that the time series model for the numbers of people with diabetes in Diyala governorate is represented by the fourth-degree moving average model, especially as it has the lowest index relative to the ACEK and Schwarz data criteria, and on this basis the numbers of people with diabetes were predicted according to the proposed model above From 2009-2013, the estimates by months were as follows:

Table (2)

Predictive values using the fourth degree moving average model

	2009	2010	2011	2012	2013
1	153.76	180.33	218.06	211.73	211.17
2	153.11	183.57	212.43	214.56	213.03
3	155.57	185.84	210.65	216.63	217.82
4	160.61	190.01	213.89	219.81	220.40
5	165.82	200.47	216.87	221.62	233.18
6	165.88	200.71	200.33	222.73	235.27
7	167.91	204.72	217.44	197.84	240.32
8	170.33	212.83	221.02	200.36	242.78
9	180.91	220.07	228.91	192.44	245.13
10	200.45	220.31	230.41	195.17	241.51
11	240.78	217.44	233.58	200.40	244.33
12	244.69	233.96	229.84	203.71	260.71

Figure (12) shows the graphical curve for each of the real and estimated values of the numbers of people with diabetes in Diyala Governorate, where a great convergence is observed between the two curves and this indicates the quality of the model in forecasting, in addition to the indicators for each of the Thiel coefficient, and the average of the error squares were The value for them is small and this is also an indication that the MA (4) model is good in the forecasting process.

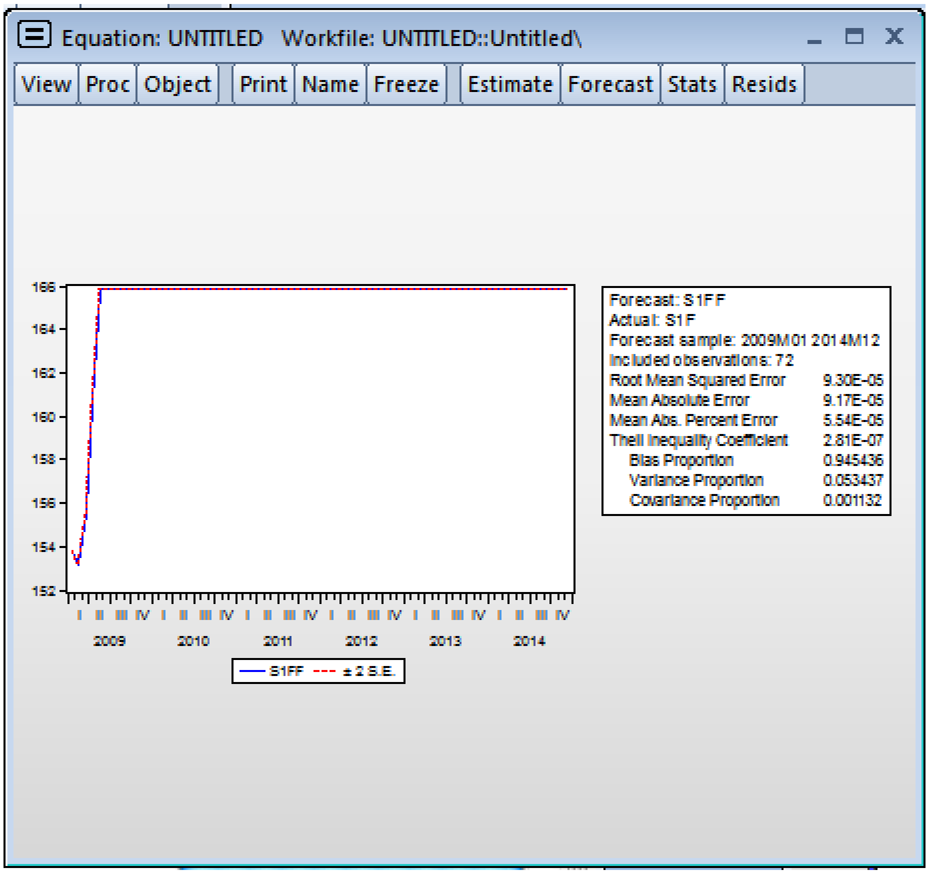


Figure (12)

Curve graph for both real and estimated numbers of people with diabetes In Diyala governorate

V. Conclusions

In light of the results contained in the body of the research, the following points were reached:

1- The time series of the number of people with diabetes for the age group greater than (15) years in Diyala governorate does not suffer from the problem of heterogeneity of variance.

2- The time series for the numbers of people with diabetes does not include the general trend component, meaning that the series is stable on average.

3- The time series for the numbers of people with diabetes in Diyala governorate does not include anomalous values and it follows the normal distribution.

4- By examining both the self-correlation and partial self-correlation functions for the stable series, it was found that the fourth-degree moving averages model MA (4) is the one that fits the series of numbers of people with diabetes in Diyala governorate.

5- The number of people with diabetes started to increase and then stabilized in the number during the 2013 forecast period.

VI. Recommendations

Through what has been presented in terms of the definition, causes and types of diabetes infection in the general framework of the research, and the approval of the sources included in the research, as well as the results that have been reached on the practical side and the conclusions that have been reached, we recommend the following:

1- Health institutions, the media, and the health policy adopted by the government should spread health awareness and hold seminars and conferences on introducing this disease and its severity among citizens.

2- Paying attention to diet and exercise

3- Women should pay attention to breastfeeding their children, because breastfeeding improves glucose and insulin levels.

4- Using sugar-lowering tablets, in consultation with a specialist doctor

5- Relying on the results of this research and the model that was reached for the purpose of predicting the number of people infected with this disease.

6 -Conducting similar studies for people with diabetes in the rest of the governorates of Iraq.

References

1- Ahmed, Ben Ahmed “Standard Modeling of National Electricity Consumption in Algeria during the Period (1988: 10-03-2007)” The People's Democratic Republic of Algeria - Ministry of Higher Education and Scientific Research - University of Algeria - Faculty of Economic Sciences and Management Sciences, a note presented within Requirements for obtaining a Master’s Degree in Economic Sciences - Quantitative Economics Branch - 2008.

2- Al-Hamid, Muhammad bin Saad, “Diabetes, its causes, complications and treatment” Kingdom of Saudi Arabia - King Abdulaziz City for Science and Development 2008

3- Al-Jubouri, Abeer Hassan Ali, “Predicting Iraqi Oil Prices for the Year 2010 using Time Series”, University of Babylon Journal, Human Sciences, Volume 18, Issue 1, 2010

4- Al-Sous, Muhammad Fayek Mahmoud, “The Use of ARFIMA Models in Predicting the Food and Agriculture Organization (FAO) Indicators” Al-Azhar University - Gaza, Deanship of Postgraduate Studies, Faculty of Economics and Administrative Sciences, Master of Statistics Program, Master Thesis in Statistics, 2014.

5- Al-Ghannam, A.M. Hamad bin Abdullah “Time series analysis of the stock price index in the Kingdom of Saudi Arabia using the Box-Jenkins methodology, Saudi Arabia-Riyadh-King Saud University, King Abdulaziz University Journal, Economics and Management, Volume 17, Issue 2, 2003.

6- "The National Reference for Diabetes Education", Kingdom of Saudi Arabia, Ministry of Health - Assistant Agency for Preventive Medicine - General Administration of Non-Communicable Diseases - First Edition 2011.

7- Insaf, Alaa Hussein Omran, “Using the ARIMA Model to Predict the Number of Tuberculosis Cases in Wasit Governorate” Al-Technical Magazine - Volume Twenty Seven - Issue Five, 2014

8- Iman, Tali Muammar, “An Analytical Standard Study of Household Electricity Consumption - Case Study of Sonelgaz Unit of Bouira - During the period 2008: 1 - 2013: 12” Ministry of Higher Education and Scientific Research - AkliMohandOulhaj University - Bouira - Faculty of Economic, Commercial and Management Sciences, 2014

9- Berry, Dr. Adnan Majed Abdel-Rahman, “Methods of Statistical Prediction - Part One -” King Saud University, Department of Statistics and Operations Research, 2002.

9- Report on High Blood Pressure, King Faisal Specialist Hospital and Research Center, 2007

10- Rahim, Ibrahim “A Standard Study of Family Demand for Electricity in Algeria for the Period 1969-2008” Ministry of Higher Education and Scientific Research - University of Ouargla - Algeria - Faculty of Economic, Business and Management Sciences - Department of Economic Sciences, Memorandum submitted for completing the requirements of the Master’s degree in Economic Sciences 2012.

11- Rockaby, Dr. Mukhlis Abdullah, the servant, Prof. Muhammad “Using the Box-Jenkins Methodology in Predicting the Number of Umrah Performers from Abroad Per Month” Umm Al-Qura University - The Custodian of the Two Holy Mosques Institute for Hajj and Umrah Research - 2014.

12- Rouihah, Dr. Amin, "Diabetes, its causes - its symptoms", Dar Al-Qalam - Beirut - Lebanon 1973.

13- Sharawi, Dr. Samir Mustafa, “Introduction to Modern Time Series Analysis”, Kingdom of Saudi Arabia - King Abdulaziz University - College of Science - Scientific Publishing Center, 2005.

14- Attia, Abd al-Qadir Muhammad Abd al-Qadir, “The Hadith in Econometrics between Theory and Practice”, Kingdom of Saudi Arabia, Makkah Al-Mukarramah, 2004

15- Alloush, M.D. JaafarBaqer Mahmoud, Muhammad, and Noor Hazem “Measuring risks and forecasting the indexes of the Iraq Stock Exchange using the Box-Jenkins methodology” Wasit Journal for the Humanities, Volume 11, Issue 30, 2015.

16- Fatih, Luqi, “The Quality of SARIMA Mixed Seasonal Time Series Models in Sales Forecasting - A Case Study of the Jadei Mills Corporation in the People's Democratic Republic of Algeria - Ministry of Higher Education and Scientific Research - Mohamed Khidir University - Biskra - Faculty of Economic, Business and Management Sciences - Department of Sciences Management, dissertation submitted as part of the requirements for a master's degree in management science, 2014.

17- College of Pharmacy, "an educational brochure on high blood pressure", Kingdom of Saudi Arabia, Ministry of Higher Education and Scientific Research, University of Hail.

18- Muhammad, Ibrahim Humaidan Ahmed “Application of integrated self-regression models and moving averages to crude oil production in Sudan for the period (2005-2012)” Sudan - Gezira University, part of the requirements for obtaining a master's degree in statistics, 2013

19- Makhramash, Abla “Estimating a Model for Forecasting Sales Using Time Series - Box Models - Jenkins - Case Study of the National Electricity and Gas Company in Ouargla Region” University of QasdiMurbah - Ouargla - College of Law and Economic Sciences, a note submitted for a master’s degree in economic sciences, 2006

20- World Health Organization "Global Briefing Note on High Blood Pressure" 2013

21 -Naqar, Dr. Othman. Al-Awwad, Dr. Munther, “The Box-Jenkins Methodology in Time Series Analysis and Prediction - An Applied Study on the Number of First Grade Students of Basic Education in Syria” Damascus University Journal of Economic and Legal Sciences - Volume 27 - Third Issue - 2011.