Int. J. Agricult. Stat. Sci. Vol. 16, Supplement 1, pp. 1611-1615, 2020 www.connectjournals.com/ijass
DocID: https://connectjournals.com/03899.2020.16.1611 ISSN : 0973-1903, e-ISSN : 0976-3392

ORIGINAL ARTICLE

THE USE OF GENERAL TREND MODELS TO PREDICT NUMBERS OF PATIENT WITH ASTHMA IN WASIT GOVERNORATE

Mahdi A. Abdulhussein*, Hadiya H. Matrood and Suheir K. Romani

Department of Business Administration, Al-Kut University College, Wasit, Iraq.
E-mail: Ali@alkutcollege.edu.iq

Abstract: The study aims to analyze the time series using general trend models to find the best model for predicting the number of people with asthma in Wasit Governorate based on the monthly data recorded in Zahraa Teaching Hospital for the period 2015-2018. The results showed that the appropriate model for predicting the number of infected people Asthma is a quadratic model. Based on this model, the number of patients with the disease was predicted for a period of two months and for two years where the predicted values were consistent with the original values and this indicates the efficiency of the model.

***Key words*:** Time Series, General trend model, Prediction model, Quadratic model, Asthma.

Cite this article

Mahdi A. Abdulhussein, Hadiya H. Matrood and Suheir K. Romani (2020). The Use of General Trend Models to Predict Numbers of Patient with Asthma in Wasit Governorate. *International Journal of Agricultural and Statistical Sciences.* DocID: https://connectjournals.com/03899.2020.16.1611

1. Introduction

Asthma is a chronic respiratory disease that affects the airways of the human lungs, as it varies in severity, causes, and response to treatment. There are several factors that lead to an asthma attack such as smoking, where inhalation of cigarette smoke causes the airways of people with asthma to be stimulated and cigarette smoke is responsible for more than half of Asthma attacks in people aged 40 years and over, associated medical conditions that occur in some other health conditions in people with asthma, such as esophageal reflux, nasal sinusitis, obstructive sleep apnea, genetic factors, family history is a risk factor for asthma, and there are many influencing genes that affect some people initially with stable asthma which lasts for weeks or months, then suddenly develops into a seizure of severe asthma, and the reactions of individuals differ due to some different factors and other factors. The study relied on monthly data for the numbers of asthma patients in Wasit Governorate for the period (2015-2018) as a time series for the purpose of its analysis to arrive at the best model for predicting the numbers of people with this disease for later periods in order to take the necessary measures to reduce this disease in the future.

Research Objective

Study of general trend models is to determine the best model of general trend models to predict the numbers of asthma patients in Wasit Governorate

Research Hypothesis

The study assumes that the number of people with asthma is growing in an increasing number. Predicting the numbers of people with this disease in Wasit Governorate is to forecast the numbers of people with asthma in provinces that face the same conditions.

Tools used in the Research

The researcher relied on his research on the theoretical side, where he dealt with general trend models in analyzing time series and supporting the theoretical side of the application by relying on realistic data regarding the number of patients with asthma in Wasit Governorate in order to reach the best model for predicting the numbers of people with asthma for later periods and then included the part. The last is the most important conclusions and recommendations that were

***Author for correspondence Received** September 14, 2020 **Revised** November 04, 2020 **Accepted** November 30, 2020*

reached, then the sources. As for the program used in data analysis, the MINITAB program was used.

2. The Theoretical Side

2.1 Time Series

A time series is defined as a set of observations or data recorded digitally, which represents a specific phenomenon and for consecutive periods of time, and often these data are equal in terms of their occurrence, if they are equally expressed as (S_{t1}, S_{t2},…, S_{tn}) at time periods (t_1, t_2,…, t_n), where n represents the values of observations [Kneher (2005), Al-Juburi and Abdullah (2012)].

3. General Trend Measurement Models

The general trend is one of the most important vehicles in the time series [Al-Sudani (2018)], which represents the long-term change in the chain and its measurement helps us in

- Knowing how the phenomenon will develop in the long run.
- Predict what will be the case for future values.
- Delete the effect of the general direction of the chain, and then other changes can be studied better.

The traditional methods of measuring the general trend depend on the reconciliation of what is known as the general trend curve. There are several methods that can be used to reconcile the general trend curve, including primitive depends on consideration, others depend on the media to get rid of these irregular changes, and the third depends on regression analysis, which is related to statistical theory, which is our study's subject to this research.

3.1 Regression Analysis

A regression analysis method is an important traditional method for estimating the general trend. It depends on determining a non-random mathematical equation g(t) to determine the general trend, and then it is assumed that the model is suitable for studying the evolution of the phenomenon, it can be written in the following form

$$y_t = g(t) + \varepsilon_t \tag{1}$$

where, $g(t)$ is a function of time, which represents the general trend vehicle.

The ε_t represents random errors and assumes that these variables are unrelated and has a fixed expectation that is equal to zero as well as a constant variance σ^2.

3.1.1 Linear Trend

Most of the phenomena that arise in the fields of dysfunctional knowledge show a general linear trend over time in the long term [Shaarawi (2005)], and in such phenomena the appropriate form can be written as follows

$$z_t = b_0 + b_1 t + \varepsilon_t \tag{2}$$

where, z_t is the value of the phenomenon at the time period t, b_0, b_1 are the two form parameters

The model is appropriate if the level of the chain is changing by a constant amount with the change of time, one period of time, the estimate of the general trend may be limited to the estimation of the two parameters of the linear model b_0, b_1 . This reminds us of the simple linear regression model t the role of the independent variable, we can write the model (1) in the form of agency matrices

$$Z = Xb + \varepsilon \tag{3}$$

whereas, Z is vertical vector of the rank n contains string observations, in other words, that

$$Z = \left[z_1 z_2 ... z_n \right]'$$

As for the matrix X, it is of the rank n*2, where the elements of its first column are equal to the correct one. As for the other column, elements represent different values of time units and always equal to the numbers 1, 2… n that is

$$X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ . & . \\ . & . \\ . & . \\ 1 & n \end{bmatrix}$$

The vertical vector b is of the order 2 and contains the two parameters of the model meaning that

$$b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}$$

and the vertical vector ε is of the order n, which

contains the variables, ε_i that meaning

$$\varepsilon_i = \varepsilon_1, \varepsilon_2, ..., \varepsilon_n$$

Assuming that model (3) is correct, the least squares estimate of a parameter vector is

$$\hat{b} = \left(X'X \right)^{-1} X'Z \qquad (4)$$

where

$$X'X = \begin{bmatrix} n & \sum_{i=1}^{n} t \\ \sum_{i=1}^{n} t & \sum_{i=1}^{n} t^2 \end{bmatrix}$$

The vector of the estimated values is as follows

$$\hat{Z} = X\hat{b}$$

As for residual vector is

$$e = \hat{\varepsilon} = Z - \hat{Z}$$

and sum of error squares is

$$SSE = e'e = \sum_{i=1}^{n} e^2$$

The estimate of the variance of errors is

$$s^2 = \sigma^2 = \frac{SSE}{n-2} = \frac{\sum_{i=1}^{n} \left(z - \hat{z} \right)^2}{n-2}$$

We can convert the above model from exponential to linear formula using logarithms according to the following

$$\ln z_t = \ln c + rt + \varepsilon_t$$

To obtain the constant estimation lnc, r, we apply special rules for simple linear regression on the data converted $\ln z_t$ by performing a regression of $\ln z_t$ over time, Then we can get the constant estimate c as follows

$$\hat{c} = e^{\ln c}$$

Based on the foregoing, the estimate of the exponential trend curve is according to the following

$$\hat{z}_t = \hat{c}e^{rt}$$

3.1.2 Application Accuracy Measures

Mean Square Error

The Mean Square Error one of the most used statistical measures that are calculated according to the following law

$$MSE = \frac{\sum_{t=1}^{n} \left(z_t - \hat{z}_t \right)^2}{n}$$

Mean Absolute Percentage Error

We can write Mean Absolute Percentage Error with the following formula

$$MAPE = \frac{\sum_{t=1}^{n} \left| \frac{z_t - \hat{z}_t}{z_t} \right|}{n} \times 100 \ , z \neq 0$$

Mean Absolute Deviation

Calculates mean absolute deviation according to the following formula

$$MAD = \frac{\sum_{t=1}^{n} \left| z_t - \hat{z}_t \right|}{n}$$

3.1.3 Prediction

Prediction is an important statistical technique in decision-making management, it is not possible to deny the urgent need for accurate, reliable and administratively reliable prediction systems. The main role of prediction is to reduce the size of risk indecision-making in the future. Prediction methods differed recently where contributed to that major development in statistical packages [Kneher (2005), Ahmed and Adel (2012)].

3.1.4 Applied side

Data was collected for the numbers of people with asthma and taken from the special records in the Statistics Division of Al-Zahra Teaching Hospital consisting of (48) views that belong to the period (2018-

Table 1: Number of people infected with asthma in Wasit Governorate.

Year/Month	2015	2016	2017	2018
1	187	185	159	201
2	174	184	159	200
3	85	179	123	199
4	141	177	127	197
5	197	181	115	209
6	114	176	172	213
7	170	169	182	204
8	155	188	186	198
9	112	191	177	202
10	123	190	91	217
11	154	198	100	200
12	176	147	129	194

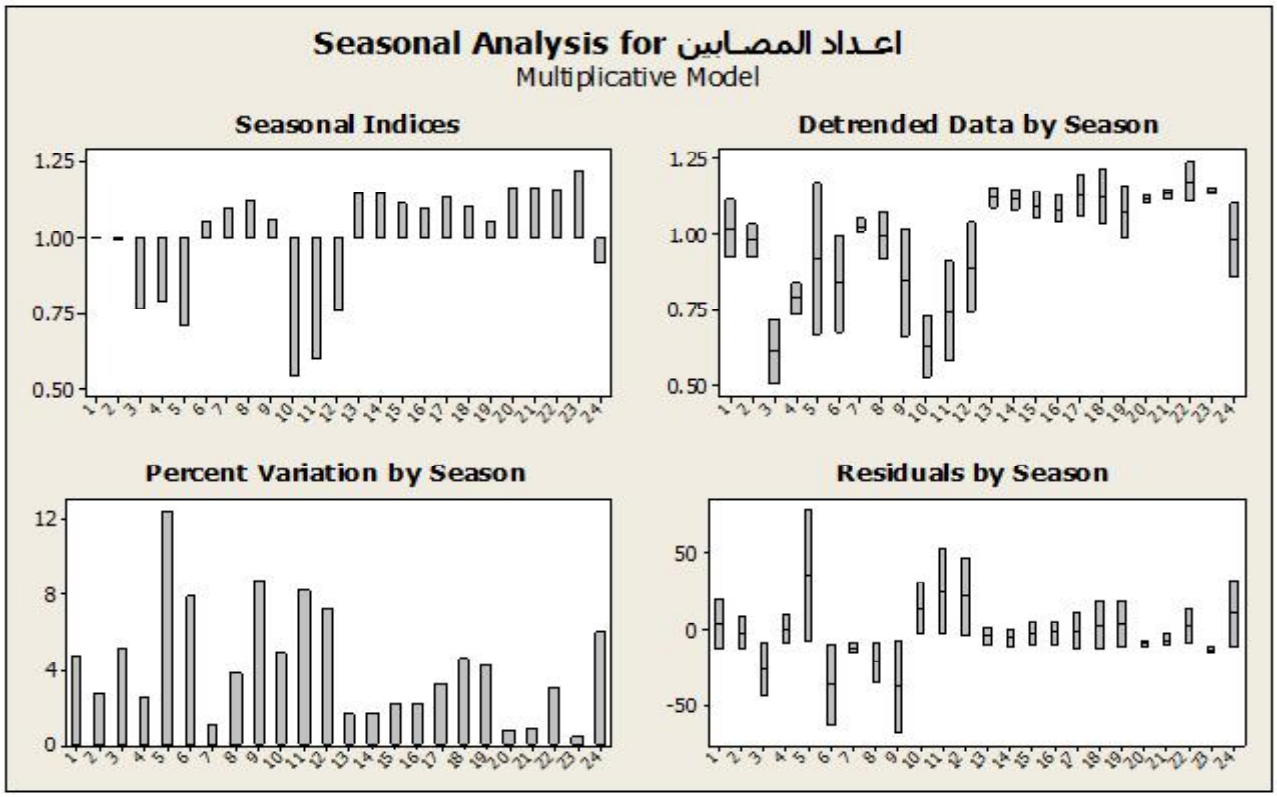

Fig. 1: Seasonal indicators

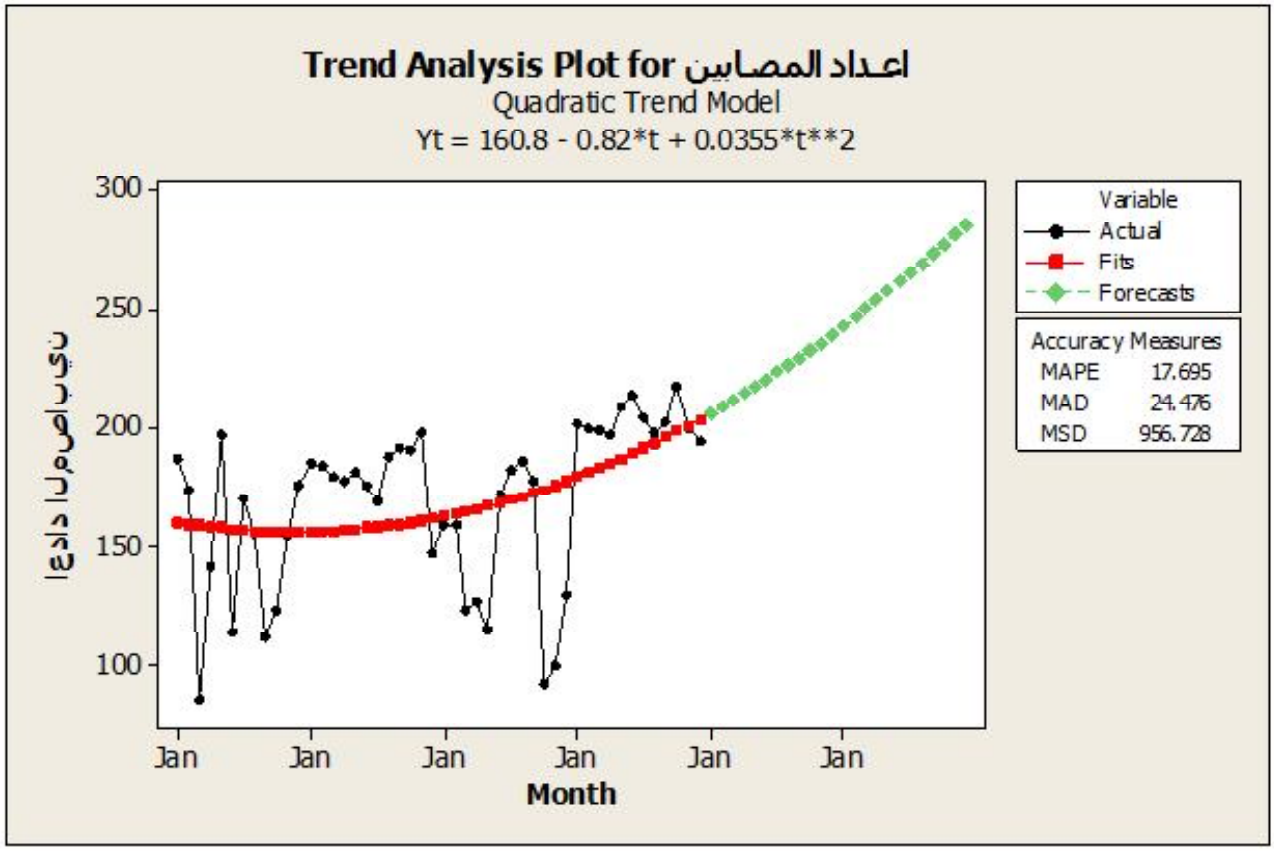

Fig. 2: Prediction values for the twenty-four future months

Table 2: Values of the criteria and the general trend equation for the three models.

Model	Equation of the general trend	MAPE	MAD	MSE
Linear Model	Yt = 146.35 + 0.920*t	17.970	25.030	993.811
Quadratic Model	Yt = 160.8 - 0.82*t + 0.0355*t**2	17.695	24.476	956.728
Exponential model	Yt = 144.052 * (1.00552**t)	18.31	26.29	1001.37

Table 3: Prediction of future values over a period of (24).

Period	Forecast	Forecast
JAN	205.938	242.966
FEB	208.633	246.513
MAR	211.399	250.131
APR	214.236	253.821
MAY	217.145	257.581
JUN	220.124	261.412
JUL	223.174	265.314
AUG	226.295	269.287
SEP	229.487	273.331
OCT	232.751	277.447
NOV	236.085	281.633
DEC	239.49	285.89

2015), then study the pattern of this series and determine the most suitable model and then use it to predict the values of the phenomenon for a period of 24 months (Table 1). The Minitab statistical program was used to obtain the best results (Figs. 1 & 2).

From Table 2, which represents the values of the models accuracy criteria, it was found that the quadratic model has recorded a noticeable preference over the rest of the models. The results for this model were as follows

- Mean Square Error (MSE) = 956.728
- Mean Absolute Percentage Error (MAPE) = 17.695
- Mean Absolute Deviation (MAD) = 24.476

They are the lowest values in relation to other models, and accordingly, the prediction of future values over a period of 24 months will be calculated according to the quadratic model as shown in Table 3.

4. Conclusions

Based on the results of the practical application, it was reached that

1. The quadratic model clearly outperformed on other models as it will be the lowest values for the three metrics.
2. The linear model came order second, with a slight relative difference from the quadratic model.
3. We have noticed from the results that all models have registered a noticeable increase in the mean values of the error squares, which makes the use of these models unwanted mostly.

Recommendations

Based on the findings of the Conclusions, we recommend the following

1. We recommend using the quadratic model due to its superiority over the other models.
2. Using the Box-Genghis model in building time series forecasting models.
3. Providing an integrated database in all health institutions so that researchers can prepare accurate and clear studies related to this aspect.

References

Ahmed, A.S. and M.Y. Adel (2012). Using time series to predict the productivity of gum Arabic in the white crops market for the period (1960-2012). *Journal of Scientific Research for Sciences and Arts*, **10 (15)**, 211-238.

Al-Juburi, N.S. and H.H. Abdullah (2012). The time series model for monthly temperature averages in the city of Samarra. *Tikrit Journal of Administrative and Economic Sciences*, **8 (25)**, 143-158.

Al-Sudani, M.M. (2018). Comparison in the Use of General Trend Equations in Predicting Wheat Production in Iraq, University of Baghdad - Ibn Al-Rushd College of Humanities. *Journal of Arts*, **3**, 13-29.

Kneher, A.L. (2005). Building a time series model to predict the numbers of patients attending Al-Zahraa Teaching Hospital in Wasit Governorate. *Ph.D Thesis, Al-Kut Journal of Economic and Administrative Sciences*, Wasit University, Iraq.

Shaarawi, S.M. (2005). *An Introduction to Modern Time Series Analysis*. El- Malk Abdulaziz University, Saudi Arabia.