International Journal of Research in Medical and Basic Sciences
Volume 6 Issue 9, September 2020 ISSN: 2455-2569 Impact Factor: 5.433
Journal Homepage: http://mbsresearch.com, Email: mbsresearchp@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

Application of SARIMA model to forecasting monthly Diabetic patients at Al-Baha Region

Salemalzahrani

[3]Al-BahaUniversity, Faculty of Science & Arts, dep of Math, Al-Mandag, KSA
Email: salem.bb@hotmail.com

Abstract:

SARIMA were used to modelling and forecasting monthlyDiabetic patients time series. The best SARIMA models were selected based on their autocorrelation function (ACF), partial autocorrelation function (PACF). The ACF of the residuals plot shows the randomness andhomogeneity of model residuals. The performance and validation of the SARIMA model premasters wereevaluated based on various statistical measures. We found that the best model that represented the data is SARIMA (1,1,1) (2,0,0)12. Finally, the results obtained can be applied to predict the futures trend of the Diabetic in Al-Baha Region KSA.

Keywords:

Seasonal ARIMA model,Diabetic patients, Al-Baha,ACF, PACF

1. Introduction:

An increasing diabetic patient became a great challenge in The General Directorate of Health Affairs, Al-Baha– kingdom of Saudi Arabia; therefore studying of this phenomenon becomes important issue.

Diabetes is a common disease around the world, which can encourage various systemic diseases and high mortality. It is a disease, which categorizeby high sugar levels in the blood and urine. It is **usually** diagnosed by means of a glucose tolerance test (GTT). There are three kind of diabetes mellitus[1].

The first Kind of diabetes mellitus: results from the body's failure to produce sufficient insulin. It is often occurring among children young. Type 2 diabetes mellitus: results from resistance to the insulin, often initially with normal or increased levels of circulating insulin. Gestational diabetes, the third kind is Gestational diabetes and it happen when pregnant women without a previous history of diabetes develop a high blood glucose level.

Diabetes is a main health challenge. Globally, the estimated number diagnosed with Diabetes is approximately 463 million people per year and mortality is 4.2 million deaths per year [2].

Al-b\Baha Health Affairs has launched the «Diabetes Friend» Initiative, targeting 1,500 diabetics, including children, school students, and the elderly across the region. The Affairs serves about 20,000 diabetics through the Diabetes Center of King Fahad Hospital-Al-Baha and diabetes clinics of the region's hospitals, in addition to the follow-up of healthcare centers [3].

2. Literature Review:

ARIMA Models:

The Box-Jenkins methodology was described in a highly advantaged book [Gilchrist, 1976]. They effectively put together in a comprehensive manner, the relevant information required to understand and use univariate time series ARIMA models [Holt, 1957; Jonathan, 2008]. The ultimate objective is to arrive at a model that appropriately describes a time series under consideration. Their approach to construct ARIMA models consists of four phases: model identification, model estimation, model diagnostic checking and forecasting. Nemours models have been used to represent a time series depending on the underline process assumed to operate on the series.

The General Box-Jenkins model [Brown, 1956]:
The general ARIMA model of orders $(p,d,q)(P,D.Q)^s$ can be written as:

$$\emptyset(\mathrm{B})\Phi(B^s)\Delta_d\Delta_s^D x_t = \theta(\mathrm{B})\Theta(B^s)e_t \quad (1.1)$$

Where:
$(1-B)^d$ is the d[th] order difference)
$(p,d,q)$is the nonseasonal part of the model.
$(P,D.Q)^s$is the seasonal part of the model.
$S$is the seasonal Length.

3. Empirical Results:

3.1 Study area

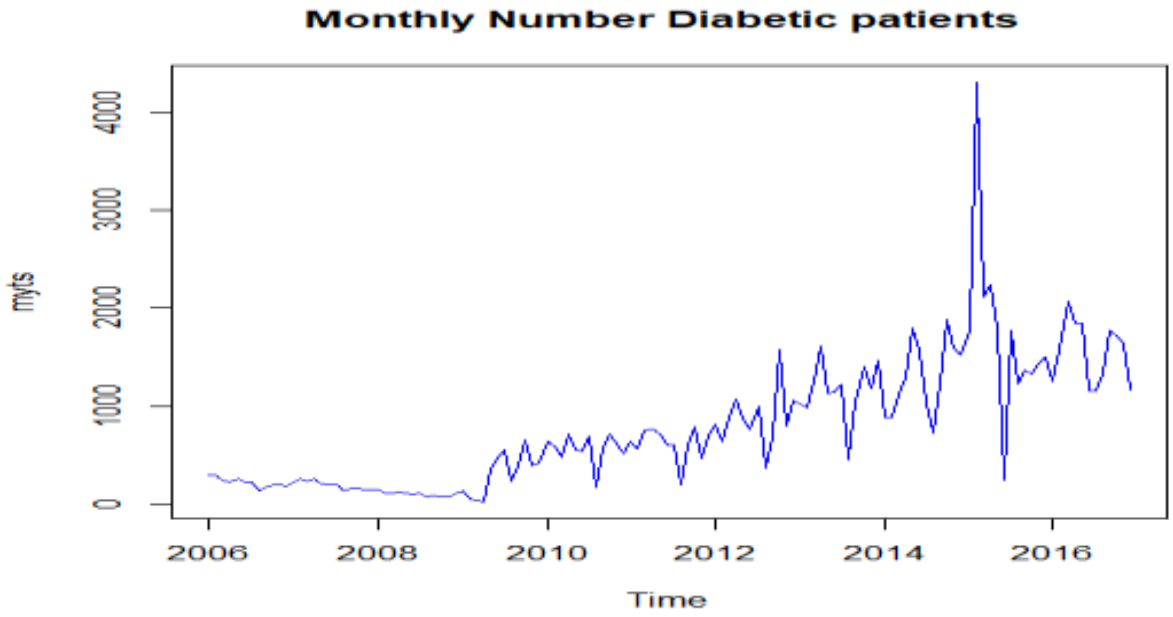
International Journal of Research in Medical and Basic Sciences
Volume 6 Issue 9, September 2020 ISSN: 2455-2569 Impact Factor: 5.433
Journal Homepage: http://mbsresearch.com, Email: mbsresearchp@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

3.2Diabeticpatients Data

This paper used monthly diabetic patients data attended Al-Baha hospitals during the period from January 2006 to November 2016.

3.3 Model Identification

The time series data cover 132 months, from January 2005 to December 2016, and depict notable seasonality and a upward trend of diabetic patients, as shown in Fig. 1



Figure(1): Sequence chart of Diabetic patients attended Al Baha Region

Table (1) ADF test results of Diabetic patients

Unit root test results / Diabetes series		
Test Type	Level	
	Test value	Prb
ADF test	−7.0928	0.01

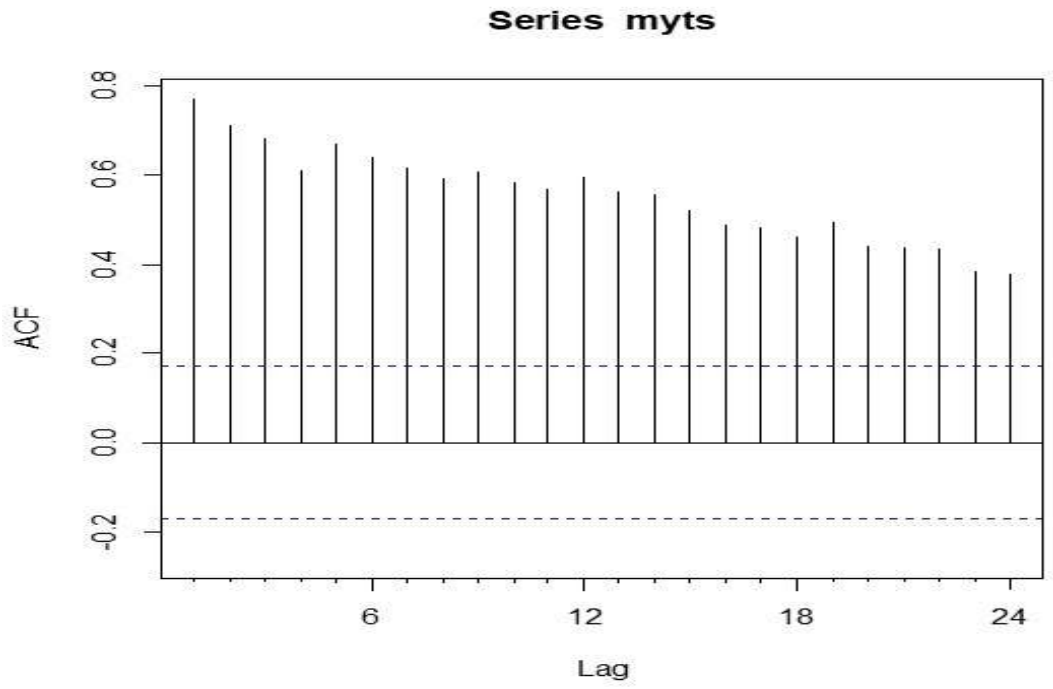
The ADF test of Diabetic patients is used to judge whether the processed series, is stationary or not. And the result of the ADF test is shown in Table 2. From the result, we can know that the value of T statistic is ( ) , which is smaller compared with the critical values at 5% confidence level. In addition, the P value is (0.01), so the original hypothesis is rejected and there is no unit root in the series . Therefore, it is a stationary series

The ACF plot of the diabetic patient's data in Fig. 2 depicts seasonality, which dies down

International Journal of Research in Medical and Basic Sciences
Volume 6 Issue 9, September 2020 ISSN: 2455-2569 Impact Factor: 5.433
Journal Homepage: http://mbsresearch.com, Email: mbsresearchp@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

slightly, while the PACF plot of the diabetic patient's data in Fig. 3 tails off after lag 1, and decays in sinewave fashion. In an initial attempt to remedy the non-stationarity of the time series, and eliminate the trend and seasonality, non-seasonal differencing was employed the non-seasonal component of the model identified by examining the ACF and PACF plots (Fig. 2 and 3) of the non-seasonal Diabetic patients. The ACF values in Fig. 2 decays exponentially in a sine-wave fashion and the PACF (Fig. 3). Decline steadily after 1 lag This suggests a moving average of order 1, resulting in an autoregressive moving average (1,1,1)12 model (i.e. p = 1, d = 1 and q = 1)

Figure (2): ACF plot of Diabetic patients attended Al Baha region



International Journal of Research in Medical and Basic Sciences
Volume 6 Issue 9, September 2020 ISSN: 2455-2569 Impact Factor: 5.433
Journal Homepage: http://mbsresearch.com, Email: mbsresearchp@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

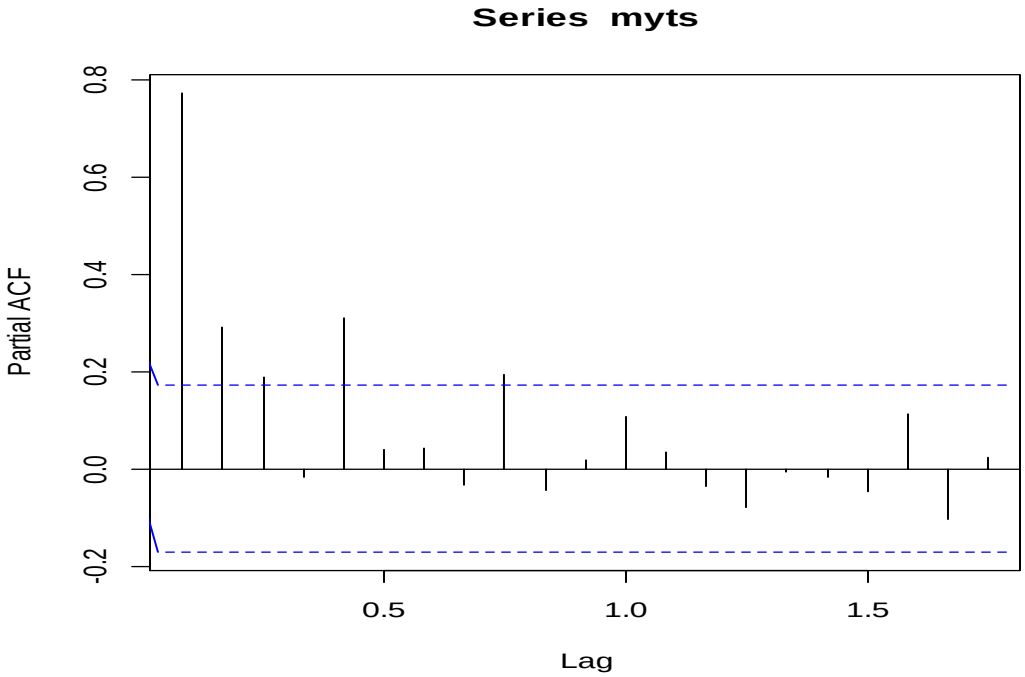


Figure (3): PACF plot of Diabetic patients attended Al Baha region

In Fig. 4 and 5 show the ACF and PACF plots of monthly Diabetic patients after non seasonal differencing transformation. In Fig. 4 , the ACF cuts off after 1 lag, which suggests a seasonal moving average MA (0) model, while the PACF plot (Fig. 4) declines after 1 lags, which suggests a seasonal autoregressive AR (1) model. Therefore, based on the non-seasonal differencing and seasonal differencing, the best model is SARIMA (1,1,1) (2,0,0)12.

International Journal of Research in Medical and Basic Sciences
Volume 6 Issue 9, September 2020 ISSN: 2455-2569 Impact Factor: 5.433
Journal Homepage: http://mbsresearch.com, Email: mbsresearchp@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

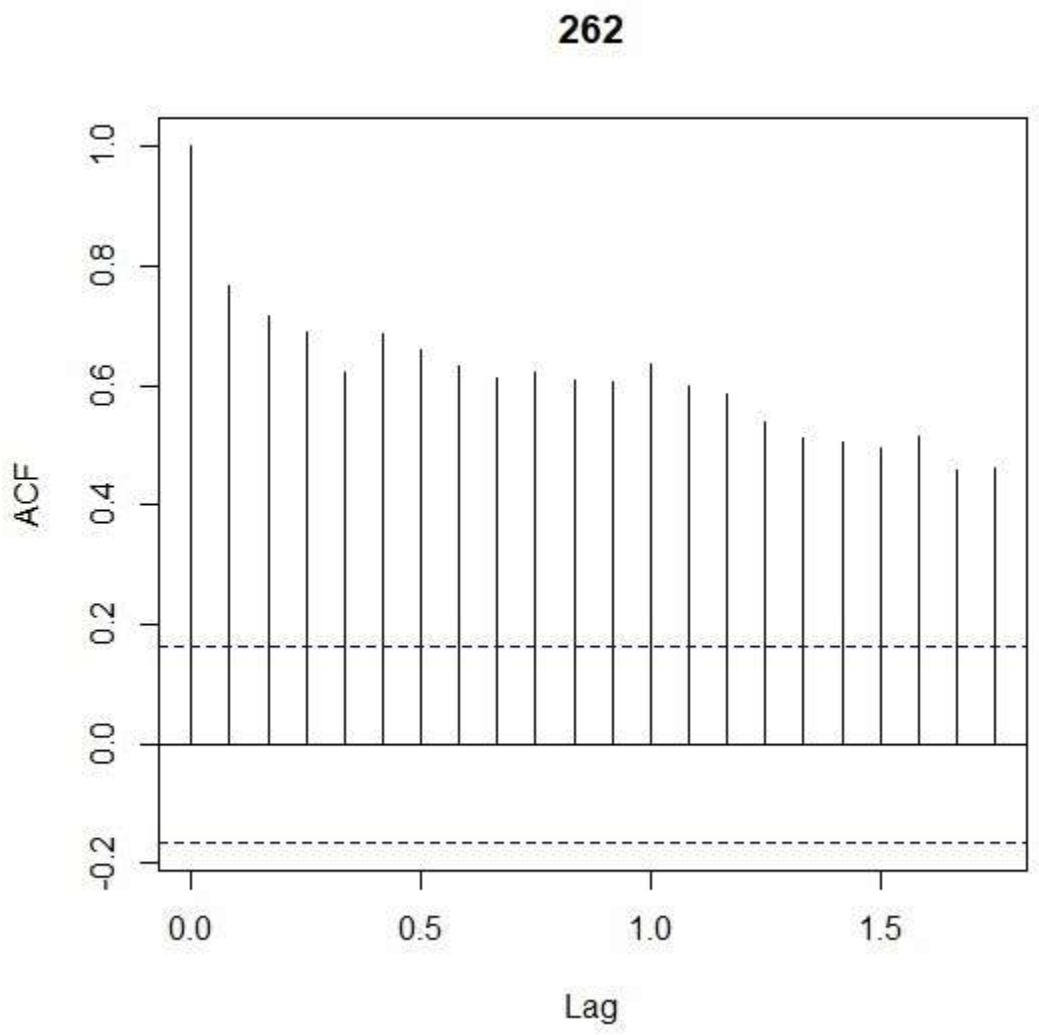


Figure **(4)** PACF plots of monthly Diabetic patients

International Journal of Research in Medical and Basic Sciences
Volume 6 Issue 9, September 2020 ISSN: 2455-2569 Impact Factor: 5.433
Journal Homepage: http://mbsresearch.com, Email: mbsresearchp@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

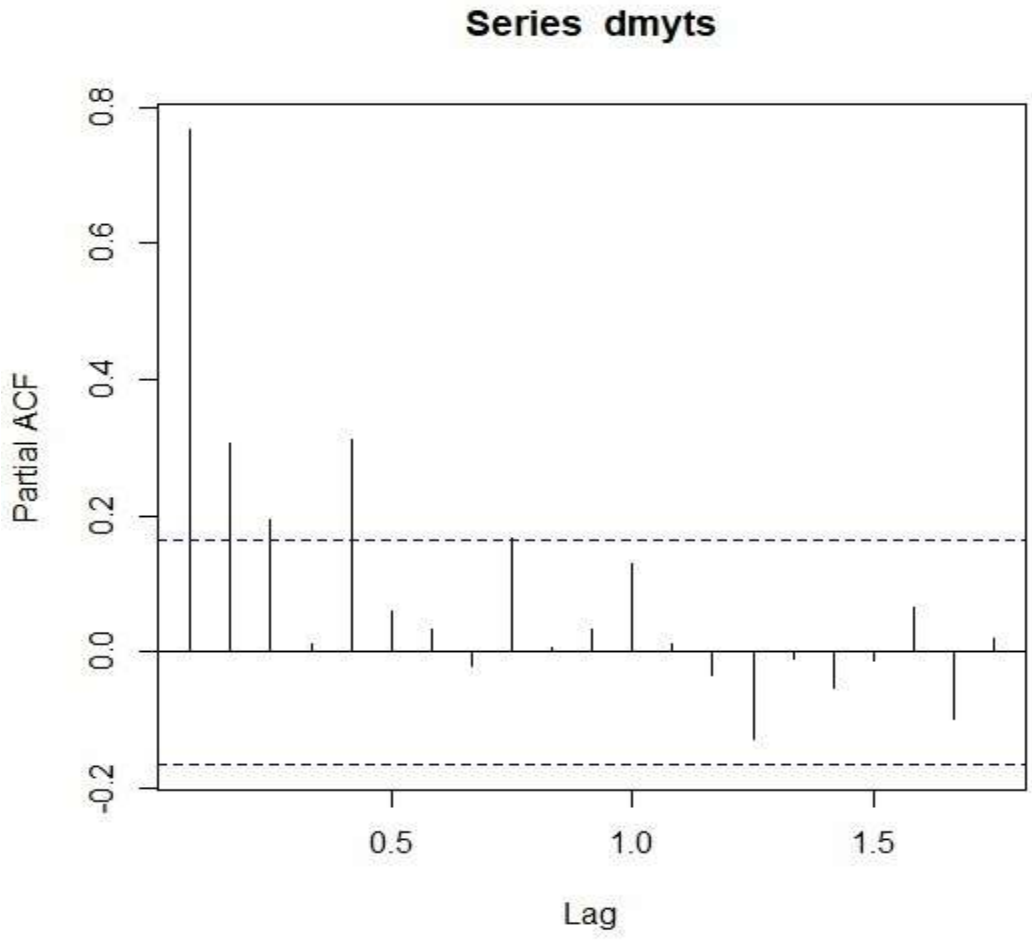


Figure (5) Partial ACF plots of monthly Diabetic patients

3.4 Model testing and parameter estimation

Table 2 shows the values of our model perimeters,and all the estimates parameters provided in Table 2 are significant

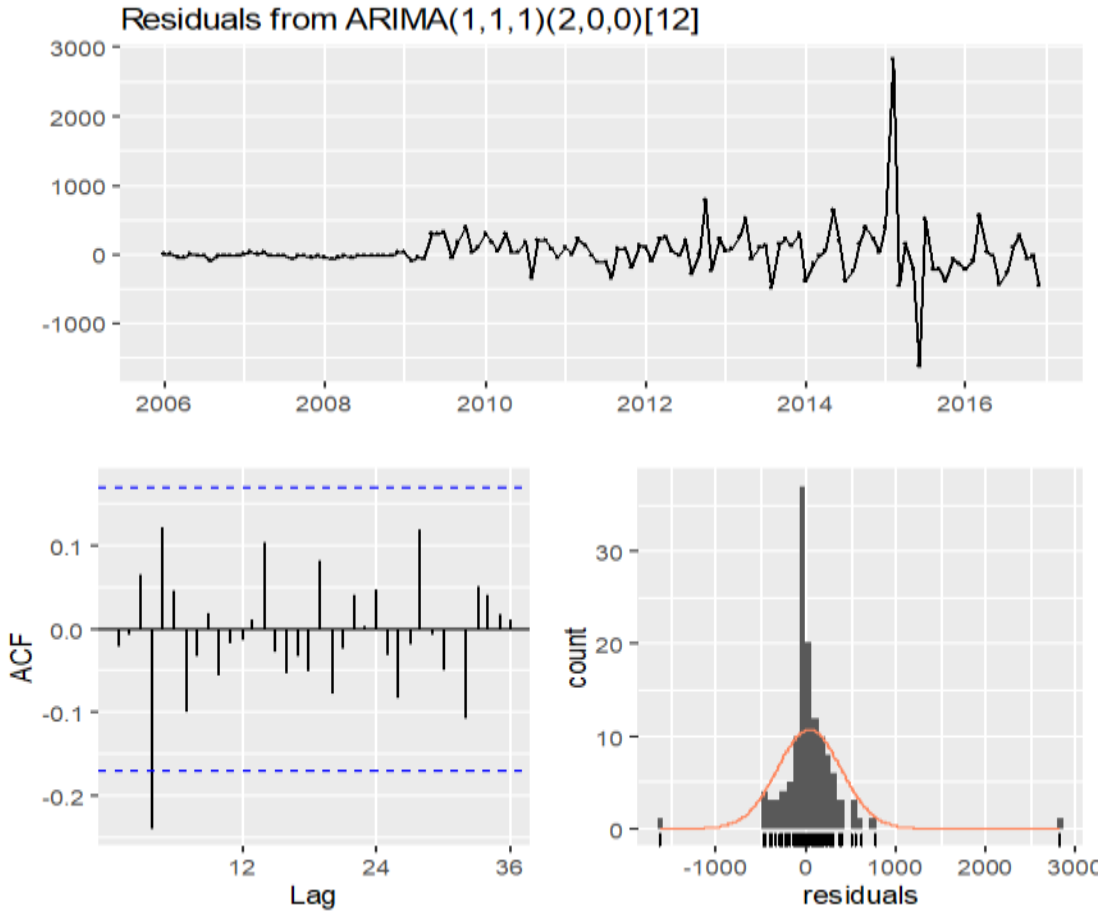
Table **(**2**):** Parameter estimation

Unit Root Test Results for Diabetes Series			
Type	Level		
	Coef	SE of coef	p-value
AR1	0.244897	0.103739	0.01824
MA1	-0.873554	0.051516	2e-16
SAR1	0.105760	0.081153	19250
SAR2	0.309625	0.117660	0.00850

International Journal of Research in Medical and Basic Sciences
Volume 6 Issue 9, September 2020 ISSN: 2455-2569 Impact Factor: 5.433
Journal Homepage: http://mbsresearch.com, Email: mbsresearchp@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal

3.5 Model validation

This was done by verifying the ACF of the residuals to check for autocorrelation, The ACF plot of residuals (Fig. 6) suggests that the residuals have a constant variance, and the autocorrelations were modelled out leaving only one significant value as indicated by the spike in lag 12. We therefore proceeded to use the SARIMA (0,1,1)(0,1,1)12 model for forecasting, since it provides a reasonable fit to the highly seasonal and non-seasonal time series data.



Figure(6) the residuals

3.6 Forecasting

The selected SARIMA (1,1,1)(2,0,0)12 model was used to predict Diabetic patients attended Al Baha region from January 2017 to December 2020 (Fig. 7).The monthly prediction plot for diabetics who attended Al-Baha area shows that the values of expected monthly cases tend to growing in the first years and then to stability

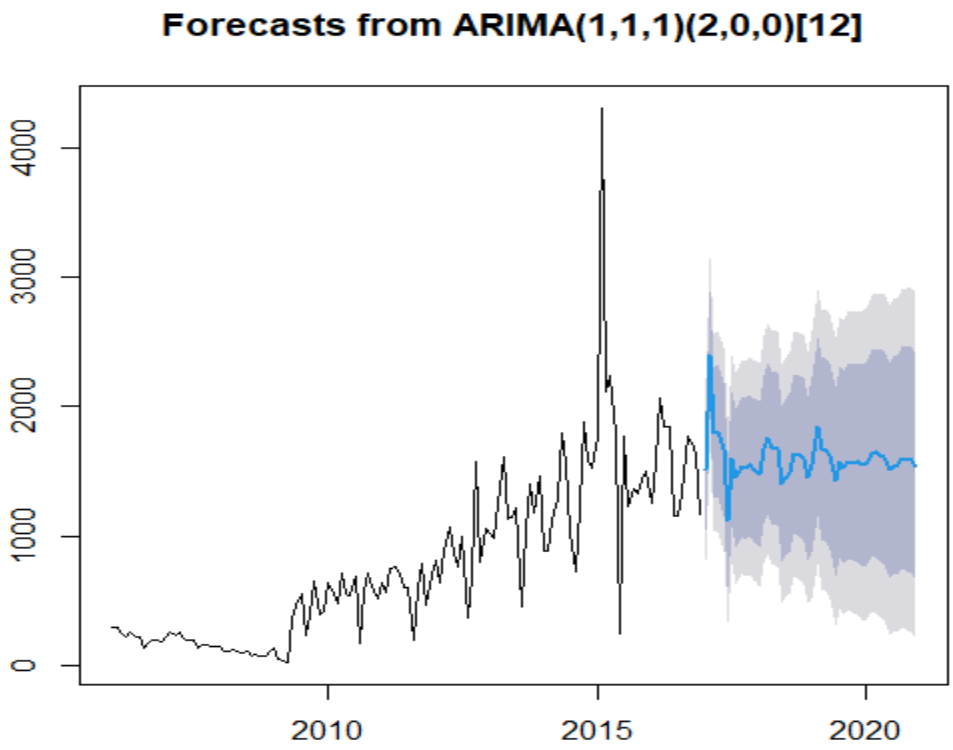


Figure (7): Observed Diabetic patients from January 2005-December 2016 and predicted Diabeticpatients from January 2017 to December 2020

4.Conclusion:

In this work the SARIMA was applied to the Diabetic patients, in order to be able to do a short run or a long run forecast for Patients. The model chosen was seasonal ARIMA (1,1,1)(2,0,0)12. The model developed showed itself statistically significant, good fit, good forecast, a low adjusted coefficient of determination, and passed diagnostic.

Declaration of competing interest

The author declare that they have no known competing financial
Interests or personal relationships that could have appeared to influence the work reported in this paper.

Conflicts of Interest

The author declare no conflicts of interest regarding the publication of this paper.

Bibliography:

[1] (CLASSIFICATION OF DIABETES MELLITUS 2019 file:///C:/Users/sshok/Downloads/9789241515702-eng.pdf)

[2] (International Diabetes Federation, 2020https://www.idf.org/our-activities/congress/idf-complications-congress-2020/abstract-submission.html?gclid=Cj0KCQjwz4z3BRCgARIsAES_OVfz2BXcgbH9U2VZ9UCzEJowzWWaBeMLOJ9vH1M8FmqTGcOZxvqt65EaAujLEALw_wcB)

[3] World Health Organization. Global report on diabetes [Internet] Available from http://www.who.int/diabetes/global-report/en/

[4]S. Al Zahrani, ,F. Al Sameeh,A. Musa,A.shokeralla , of Forecasting Diabetes Patients Attendance at Al-Baha Hospitals Using Autoregressive Fractional Integrated Moving Average (ARFIMA) Model ,January 2020 , Journal of Data Analysis and Information Processing 08(03):183-194 , DOI: 10.4236/jdaip.2020.83011

[5] Makraidkis, S. Wheelright, S. and McGee, E.,(1983)." Forecasting: Methods and Application", 2nd ed., John Wiley and sons,NewYourk, USA:

[6]Daniel Kalus, Michael Muma, Abdelhak M. Zoubir, "Distributed robust change point detection for autoregressive processes with an application to distributed voice activity detection", Acoustics Speech and Signal Processing (ICASSP) 2015 IEEE International Conference on, pp. 3906-3910, 2015.

[7]Chatfield C. The Analysis of Time Series: An Introduction. 6th ed. Florida: CRC Press, 2003.

[8] Shumway R, Stoffer D. Time Series Analysis Using the R Statistical Package. Free Dog Publishing, 2017. http://www.stat.pitt.edu/stoffer/tsa4/tsaEZ.pdf (accessed 10 November 2017

[9]Permanasari AE, Hidayah I, Bustoni IA. SARIMA (Seasonal ARIMA) implementation on time series to forecast the number of malaria incidence. International Conference on Information Technology and Electrical Engineering (ICITEE), 7 - 8 October 2013, Yogyakarta, Indonesia. Yogyakarta: IEEE, 2013:203-207. https://doi.org/10.1109/iciteed.2013.6676239

[10]ADHIKARY S.K, RAHMAN M.D.M., GUPTA A.D. 2012. A stochastic modelling technique for predicting groundwater table fluctuations with Time Series Analysis. International Journal of Applied Sciences and Engineering Research. Vol. 1. Iss. 2 p. 238–249.

[11] ADNAN R.M., YUAN X., KISI O., ANAM R. 2017. Improving accuracy of river flow forecasting using LSSVR with gravitational search algorithm. Advances in Meteorology. Vol. 2017. Art. ID 2391621 pp. 23. DOI

10.1155/ 2017/2391621.

[12]BOX G.E.P, JENKINS G.M., REINSEL G.C. 2008. Time series analysis, forecasting and control. 4th ed. New Jersey. John Wiley and Sons, Inc. ISBN 978-0470272848 pp.784.

[13]Brockwell, P. J. , and Davis, R. A. . Introduction to Time Series and Forecasting. Springer-Verlag New York, 1996.

[14]L. Zhou, P. Zhao, D. Wu, C. Cheng, H. HuangTime series model for forecasting the number of new admission inpatients
BMC Med InfDecis Making, 18 (1) (2018)

[15]Q. Xu, R. Li, Y. Liu, C. Luo, A. Xu, F. Xue, etal.Forecasting the incidence of mumps in Zibo City based on a SARIMA model
Int J Environ Res Public Health, 14 (8) (2017)

[16]Q. Mao, K. Zhang, W. Yan, C. ChengForecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model
J Infect Public Health, 11 (5) (2018), pp. 707-712

[17] J. Rob, Y.K. HyndmanAutomatic time series for forecasting: the forecast package for R
J Stat Software, 27 (3) (2008)

[18] Nobre FF, Monteiro ABS, Telles PR, Williamson GD. Dynamic linear model andSARIMA: a comparison of their forecasting performance in epidemiology. StatMed 2001;20(20):3051–69.

[19]Wangdi K, Singhasivanon P, Silawan T, Lawpoolsri S, White NJ, Kaewkungwal J. Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: A case study in endemic districts of Bhutan. Malaria J 2010;9(1):251. https://doi. org/10.1186/1475-2875-9-251

[20] Zheng L, Liu DJ, Xian XU. Application of time series analysis in forecasting theincidence of pulmonary tuberculosis. PractPrev Med 2012;6(5):375–83.

[21] Mahmood M, Narges K, Mahshid N, Abbas B. Predicting the incidence of smearpositive tuberculosis cases in Iran using time series analysis. Iran J Public Health2015;44(11):1526–34.

[22] Hosking, J. "Modelling Persistence in Hydrological Time Series using Fractional Differencing". Water Resources Research, 1984.