JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Developing an efficient statistical model to forecast the Eye surgeries at the Mecca Eye complex hospital

Salma Abass Basher Abass[1], Altaiyb Omer Ahmed Mohmmed[2]
[1,2]Department of statistics
[1,2]Sudan University of science and technology, college of science
[1]corresponding author: email: nazar9090215@gmail.com

Abstract
A major disease prevalence rate in Sudan is eye disease, which has been spreading rapidly during the last decade. Mecca eye complex hospital treats about (22,150) patients per month for all frequent eye disease patients, in this study, the number of cases in all patients visiting the complex and the six main eye disease surgeries (Cataract - Phaco - Retina - Glaucoma - Cornea and Strabismus) considered for the period from January 2014 up to April 2022 (Series size 100). It was found that the time series analysis showed that it is non-stationary, so the inverse was taken to change it to stationary without requiring different assumptions; in addition, a model was developed to forecast each surgery of each disease. Thus, the study result showed that using the proposed models to forecast the number of cases for each disease in the future was an effective method to forecast future cases by updating the data every period of time and rebuilding the models. In addition, the models clarify that the number of cases is increasing for all disease in the coming few months.

Key Words: Eye disease, ACF, surgeries clinic, Stationary, Forecasting

1. Introduction
Mecca eye complex hospital is one of the international eyes hospitals, they have experts in all fields related to eye problems including cosmetics, and the prices of their Services are very affordable (albasar.wikipedia.org). They have higher degree of success in all fields with 60 beds capacity.
Mecca Eye Complex includes a number of specialized ophthalmic clinics to contribute in the domestication of treatment in Sudan, these clinics include specialists in Ophthalmology, nursing staff and technicians, and equipped with the modern instruments for screening and diagnosis of eye diseases (charity<medadcenter.org).
The Operation Theater in Mecca Eye Complex is provided with hi-tech surgical equipment, many types of surgeries are performed in the theater, such as: Cataract surgery with Phaco machine/ small incision / surgical; Glaucoma surgery; Corneal Graft surgery; Squint surgery; Retina surgery; Orbital surgery; Oculoplastics surgery; DCR. (www.gleauty.com) About 100 surgical operations are done per day. Laser procedures are also performed (Argon laser / Yag laser). Other services: One optical shop and a pharmacy are Mecca available to serve eye patients. Free Camps & Clinical Days: A number of Free Eye Camps and clinical days were conducted inside the Republic of Sudan (www.albasar-sudan.org).
As for general frequently and the disease of the Mecca eye complex, they are not considered as main of the prevalence and will discuss in this paper. In addition, the main aim of this study is to forecast patient's size for these surgeries clinic conducted in hospital by building mixed models for

time series analysis, autoregressive and moving average (ARMA). (Alzawbaee et al 2020 and He & Tao 2018).

Where the paper includes, additionally to the introduction, an explanation and presentation of the data used, and then testing the stationary of the series, after which the models are built. Finally, these models are used in forecasting after diagnostic checking step. (Alzawbaee et al 2013 and Gujarati 2009 and Hyndman 2018).

2. Data description

The data used in this paper is represent as monthly surgeries in Mecca eye complex hospital (Sudan) for main six eye disease of the period from January 2014 up to April 2022, which has been collected from planning department – Mecca eye complex hospital- Riyadh region -Sudan, and the figure (1) represents the time series graph for each surgeries of the six disease studied. (Elliott et al 1996 and Ljung & Box 1978).

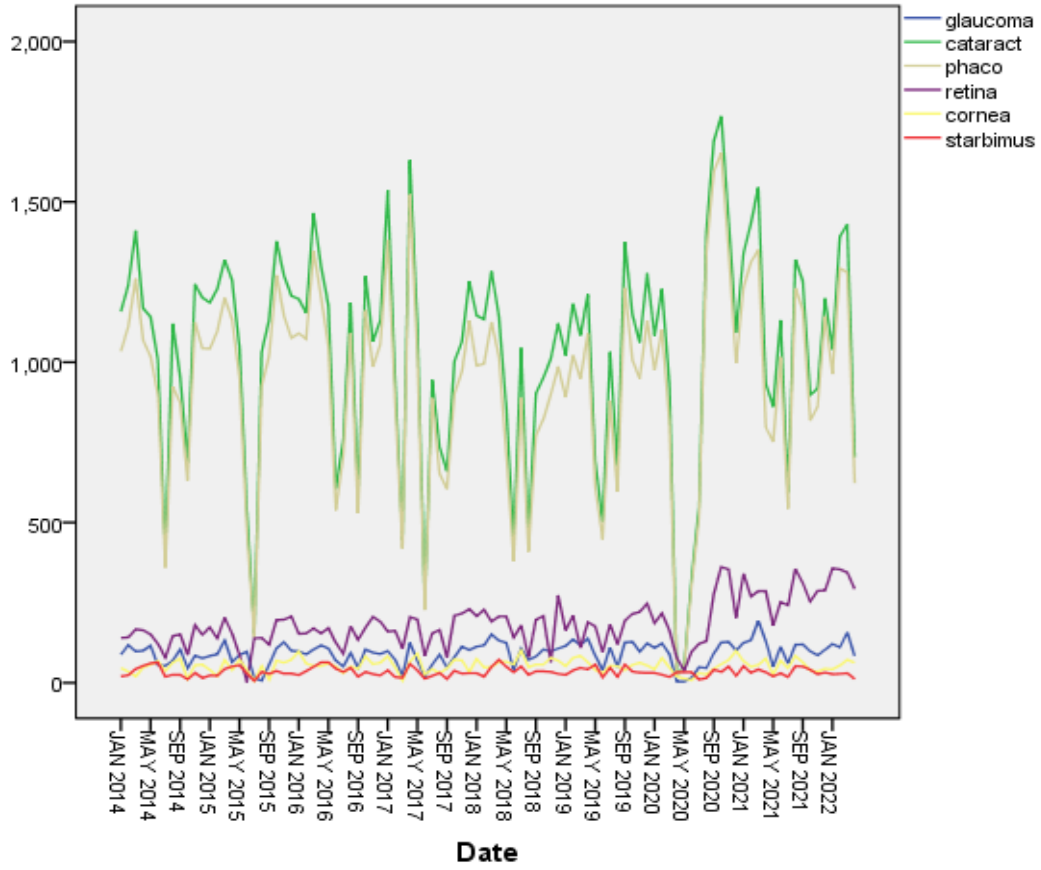

Figure (1): series of six surgeries eye done in Mecca eye complex hospital

From the figure (1) it is clear that some series have a zero frequently because it proportion to the conditions the Sudan was going through during that period specially, like revaluation that affected in operations in a hospital, the descriptive analysis of the data indicates that some series are non-stationary (containing a decreasing trend) clearly, while other series are not clear their stationary and will be confirmed by using transfer test to obtained a stationary such as Dickey-Fuller test (Dickey & Fuller 1981and Xinghua & Gao 2012).

3. Methodology and material study

There are some studies that tried to forecast the final size of surgeries eye disease cases through using many different methods over the past few years. Whereas, the main steps of time series analysis are multiple and are in order; Identification which includes knowing the type of the model and determining its rank, Estimation to estimate the model parameters, Diagnostic checking, and the last step is Forecasting (Brokwell & Davis 2002). However, all of these steps must be preceded by a series stationary test and the transformation of the non-stationary series into a stationary using many statistical packages such as SPSS, MINTAB and STATA addition to EXCEL to obtain the graphs and tables. In this paper a new technique which is determine the model has been done

ARMA (p, q) where P, q = 0, 1, 2 and there is no difference taken on this series because it stationary without it (Bisgaard & Kulahci 2011)

Meanwhile, the best model will be chosen after finding the rank selection criteria such as: Akaike's Information Criterion, (AIC) and Bayesian Information Criterion, (BIC) as well as, find Root Mean Square Error, (RMSE) and Mean Absolute Error, (MAE), and take the model that matches to the minimum of these criteria, then its parameters are estimated and diagnostic checking and then used for forecasting (Box et al 1994 and Cryer & Chan 2008).

The time series for the six main surgeries above are drawn as shown in Figure (1), which shows that most of the series are non-stationary and for the purpose of examining the stationary of the series, Dickey-Fuller and Phillips-Perron for the unit root (ERS) tests were used, as it shown in table (1). (Hyndman 2018 and Dickey & Fuller 1981and Engle 1982).

Table (1): ADF and PP test for the unit root test

Unit root test (ADF)						
	glaucoma	cataract	phaco	retina	cornea	strabismus
Constant	7.460 (0.000)	8.220 (0.000)	9.160 (0.000)	9.310 (0.000)	6.940 (0.000)	7.170 (0.000)
Trend	-0.270 (0.787)	-1.260 (0.212)	1.220 (0.226)	0.420 (0.677)	0.050 (0.959)	-0.560 (0.576)
L1	-9.370 (0.000)	-9.640 (0.000)	-9.870 (0.000)	-12.111 (0.000)	-8.700 (0.000)	-9.050 (0.000)
Take inverse						
	glaucoma	cataract	phaco	retina	cornea	strabismus
Constant	1.220 (0.227)	0.680 (0.496)	0.690 (0.492)	2.240 (0.027)	6.940 (0.000)	6.080 (0.000)
Trend	0.200 (0.840)	0.670 (0.503)	0.671 (0.504)	-1.320 (0.189)	0.050 (0.959)	-0.330 (0.745)
L1	-5.420 (0.000)	-6.784 (0.000)	-6.759 (0.000)	-9.943 (0.000)	-8.700 (0.000)	-9.274 (0.000)
Unit root test (PP)						
	glaucoma	cataract	phaco	retina	cornea	strabismus
Constant	7.460 (0.000)	8.220 (0.000)	8.350 (0.000)	9.310 (0.000)	6.940 (0.000)	7.170 (0.000)
Trend	-0.270 (0.787)	-1.260 (0.212)	1.220 (0.226)	0.420 (0.677)	-0.050 (0.959)	-0.560 (0.576)
L1	0.290 (0.770)	-0.020 (0.987)	-0.230 (0.820)	-2.120 (0.037)	1.060 (0.290)	0.680 (0.498)
Take inverse						
	glaucoma	cataract	phaco	retina	cornea	strabismus
Constant	1.220 (0.227)	0.680 (0.496)	0.690 (0.492)	2.240 (0.027)	6.940 (0.000)	6.080 (0.000)
Trend	0.200 (0.840)	0.670 (0.503)	0.671 (0.504)	-1.320 (0.189)	-0.050 (0.959)	-0.330 (0.745)
L1	6.140 (0.000)	3.680 (0.000)	3.720 (0.000)	-0.140 (0.889)	1.060 (0.290)	0.280 (0.782)

Table (1) shows that all the time series of surgeries clinic done in hospital at the period time are non-stationary except the cornea is stationary, and other clinic become stationary when take the inverse of origin series.

Figure (2) shows that when take inverse the all series become stationary for the series that non-stationary.

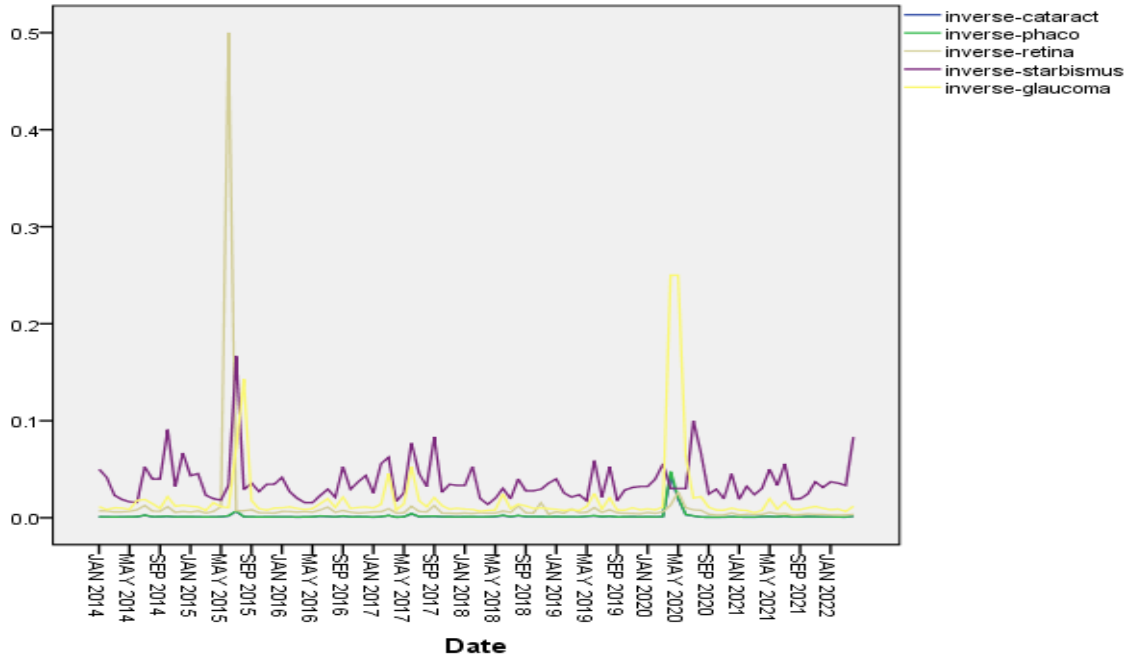

Figure (2): the surgeries clinic after taking the inverse

Hence, series is stationary after taking the inverse of original series without any difference then the model that represent this series is ARMA model with (p,d), for the possible values 0,1,2 for each, whenever the models are :

ARMA (0,1) (0,2) (1,0) (1,1) (1,2) (2,0) (2,1) (2,2)

To choose the best model many criterion models were used to fit best one, the most popular criteria are Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC), also Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) was found as criterion. The minimum values for all criterions model used it represent the appropriate model (Alzawbaee et al 2020 and Cryer & Chan 2008).

Table (2): all possible model ARMA (p,d) and are criterion measure for cornea surgeries

ARMA (p,d)	MAE	RMSE	AIC	BIC
ARMA (0,1)	15.972	20.587	610.93	6.188
ARMA (0,2)	15.985	20.683	611.86	6.243
ARMA (1,0)	15.987	20.569	610.76	6.186
ARMA (1,1)	15.987	20.655	611.59	6.241
ARMA (1,2)	15.987	20.733	612.35	6.298
ARMA (2,0)	15.990	20.688	611.91	6.241
ARMA (2,1)	**15.670**	**20.545**	**610.52**	**6.176**
ARMA (2,2)	15.828	20.719	612.21	6.338

The table (2) shows the results of criterion of optimal parameter model to fit the data of cornea clinic surgeries it was found that ARMA (2, 1) is the best one among the other models, because had a minimum values of (MAE, RMSE, AIC, BIC), as shown in above table, and for the rest parameter model of surgeries clinic criterion test table (3) represent it.

Table (3): best criterion for the all models

Surgeries clinic	ARMA (p,d)	MAE	RMSE	AIC	BIC
Cornea	ARMA (2,1)	15.670	20.545	610.52	6.176
Cataract	ARMA (1,0)	250.816	327.313	1159.18	11.720
Phaco	ARMA (0,1)	229.792	295.481	1138.72	11.496
Retina	ARMA (1,1)	43.971	57.677	812.97	8.248
Strabismus	ARMA (2,2)	10.458	13.463	527.99	5.430
Glaucoma	ARMA (2,2)	24.364	32.368	703.43	7.185

Above table (3) determine the order and quality of the model for each surgeries clinic done, and the model parameter were estimate as shown in below equations.

Table (4): the models equation

Surgeries clinic	ARMA (p,q)	Equation
Cornea	**ARMA (2,1)**	$Z_t = 53.589 + 1.149Z_{t-1} - 0.206Z_{t-2} + 0.998a_{t-1}$
Cataract	**ARMA (1,0)**	$Z_t = 1031.811 + 0.364Z_{t-1}$
Phaco	**ARMA (0,1)**	$Z_t = 929.878 - 0.457a_{t-1}$
Retina	**ARMA (1,1)**	$Z_t = 189.859 + 0.865Z_{t-1} + 0.420a_{t-1}$
Strabismus	**ARMA (2,2)**	$Z_t = 33.693 + 0.664Z_{t-1} - 0.823Z_{t-2} + 0.493a_{t-1} - 0.880a_{t-2}$
Glaucoma	**ARMA (2,2)**	$Z_t = 90.984 + 1.648Z_{t-1} - 0.918Z_{t-2} + 1.468a_{t-1} - 0.786a_{t-2}$

After we determined the model for each surgeries clinic and estimated its parameters, the next step is diagnostic checking; now we calculate the residual of ACF and PACF for each surgeries clinic and then used Ljug-Box test for the efficient model has been used. And its formula is: (Ljung & Box 1978 and Box et al 1994)

$$Q = n(n+2)\sum_{k=1}^{n}\frac{\hat{\rho}_k^2}{n-k} \cong \chi^2_{(k-p-q)}$$

Table (5): Ljung-Box test for all surgeries studied

Model	Ljung-Box Q(18)		
	Statistics	df	Sig
Cornea- model 1	10.306	15	0.800
Cataract-model 2	2.225	17	1.000
Phaco – model 3	0.977	17	1.000
Retina – model 4	25.764	16	0.057
Strabismus – model -5	16.325	14	0.294
Glaucoma –model 6	20.549	14	0.114

Table (5) shows the results of the Ljung-Box test of residual it is clear that all the models are efficient. Also Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) for the estimated residuals were drawn, and its shows that there is no significant.

The final step which is, forecasting step by using all the models, one by one, where it forecast for the period of (8) month for each surgeries clinic as its shown in the table (6) (Box et al 1994 and Alzawbaee et al 2013 and Bisgaard & Kulahci 2011).

Table (6): forecasting values of surgeries eye clinic in next months

Model		May 2022	Jun 2022	Jul 2022	Aug 2022	Sep 2022	Oct 2022	Nov 2022	Dec 2022
Cornea – model 1	Forecast	55	53	53	53	53	53	53	53
	UCL	95	94	94	94	94	94	94	94
	LCL	15	12	12	12	12	12	12	12
Cataract – model 2	Forecast	912	988	1016	1026	1030	1031	1032	1032
	UCL	1559	1676	1709	1720	1724	1725	1726	1726
	LCL	266	300	323	332	336	337	338	338
Phaco – model 3	Forecast	754	930	930	930	930	930	930	930
	UCL	1340	1575	1575	1575	1575	1575	1575	1575
	LCL	167	285	285	285	285	285	285	285

Retina – model 4	Forecast	285	272	261	251	243	236	230	224
	UCL	399	397	394	389	385	380	376	372
	LCL	171	147	129	114	101	91	83	76
Strabismus – model 5	Forecast	24	28	38	41	35	29	29	35
	UCL	51	55	65	68	63	56	57	63
	LCL	2	1	11	13	7	1	1	7
Glaucoma – model 6	91	86	82	81	82	86	91	96	91
	156	151	148	147	149	153	158	163	156
	27	20	15	14	16	20	24	28	27

In addition, the original and forecasting series for each clinic were drown, as shown in below figure (3). By looking at the figures and observing the original series behavior that forecasted for each clinic under the search, can notice the strength and efficiency of the models that were built on the one side, and the continuing number of cases increasing relatively for the next period.

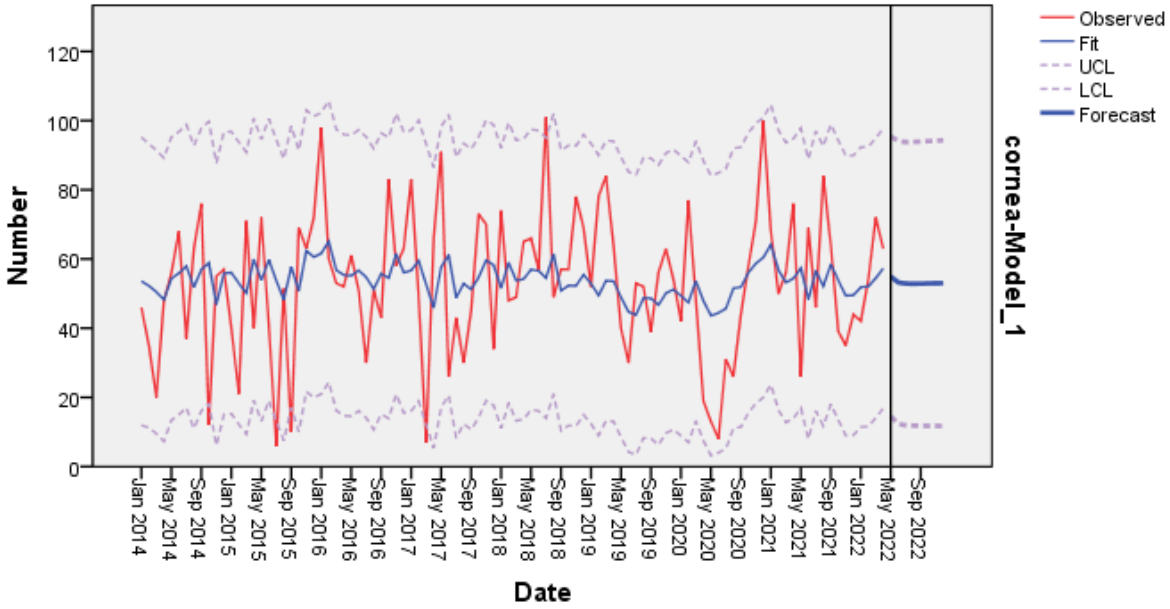

Figure (3): the forecast series of cornea model

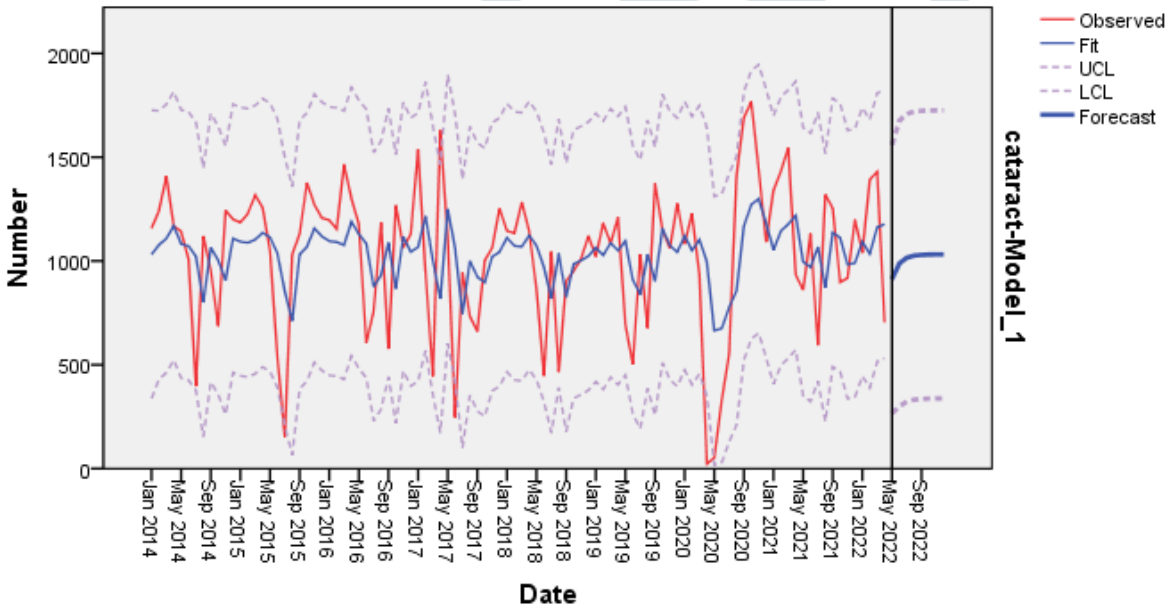

Figure (4): the forecast series of cataract model

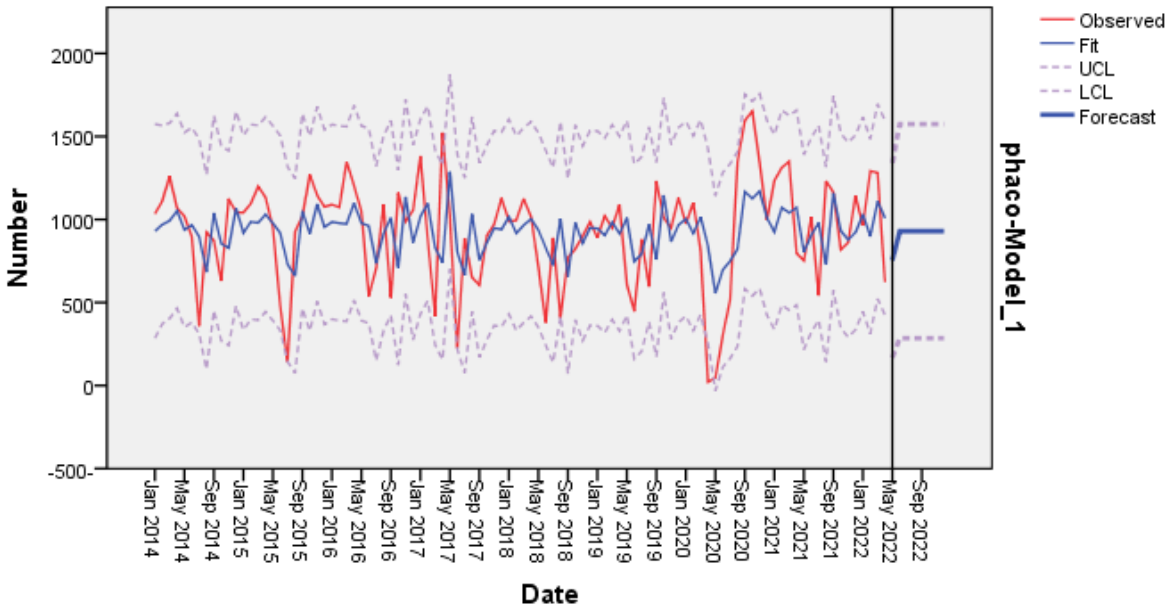

Figure (5): the forecast series of phaco model

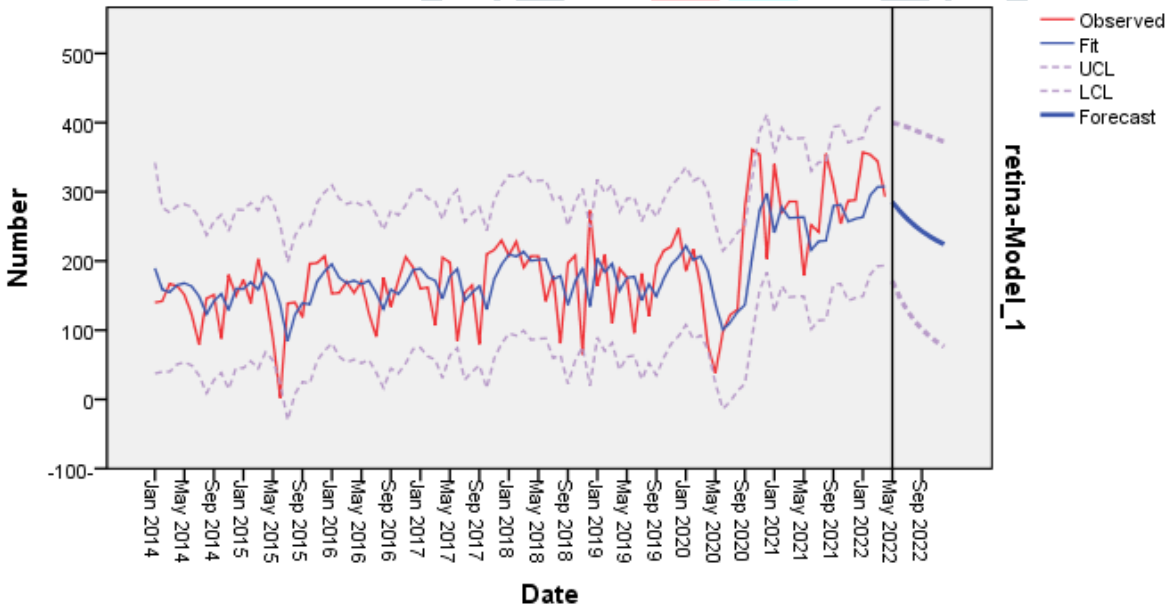

Figure (6): the forecast series of retina model

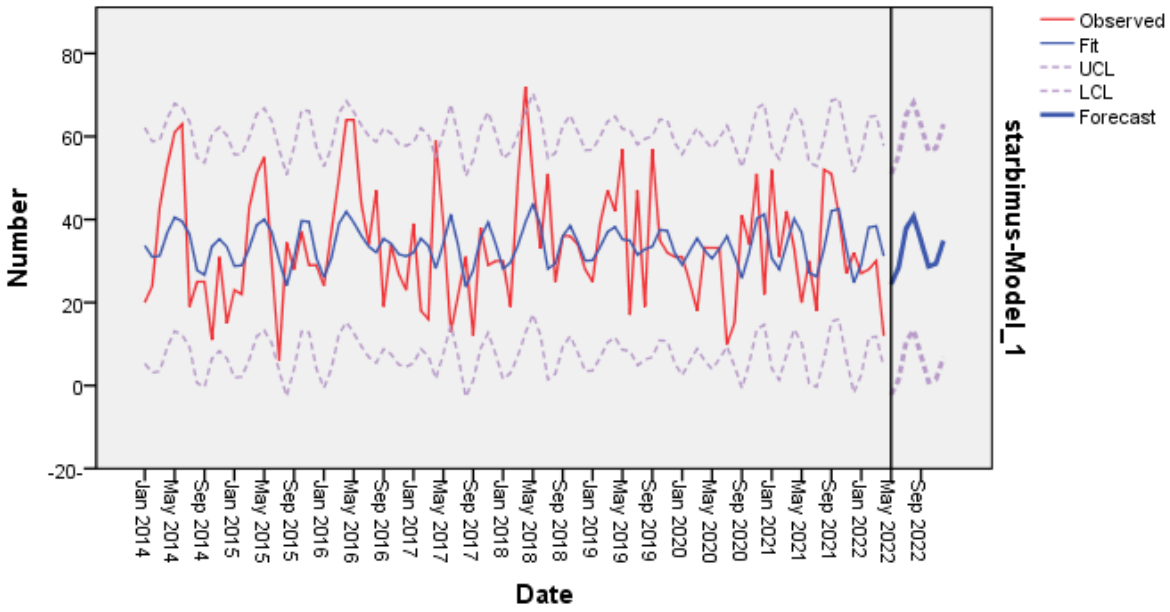

Figure (7): the forecast series of strabismus model

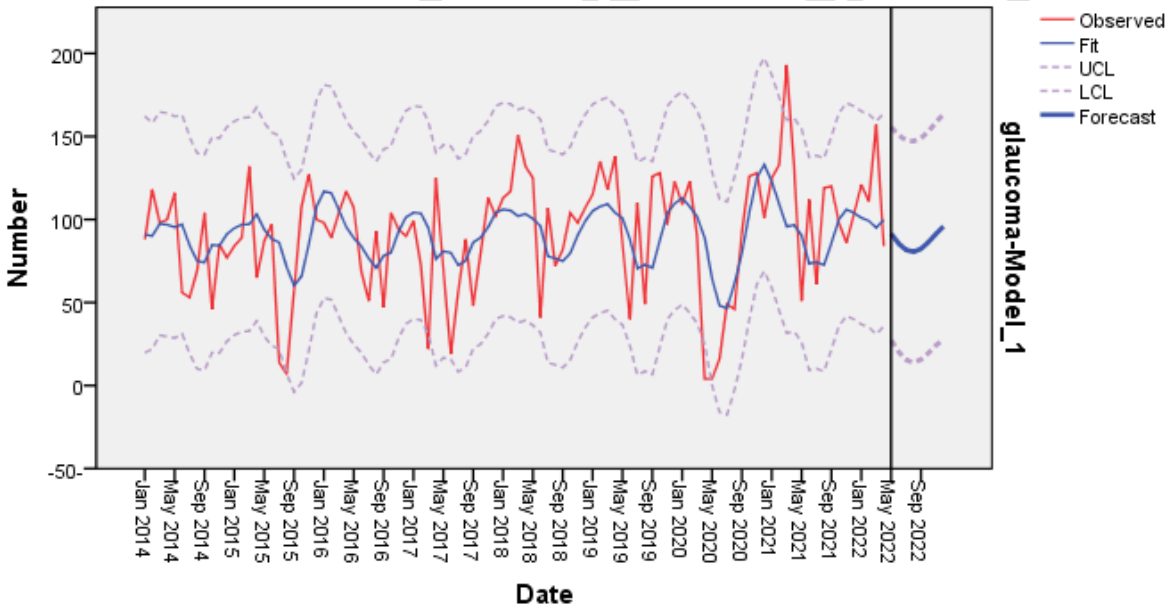

Figure (8): the forecast series of retina model

4. Conclusions

This study aimed to forecast the surgeries clinic eye disease in Mecca complex hospital using times series model technique in other to establish the one that is most suitable model for the analysis of surgeries eye patients' data in clinical research. ARMA model procedure was used to estimate the surgeries size done in complex clinic monthly of eye disease patients' whom length of stay in hospital upon admission, all the time series for all surgeries are non-stationary series (except the cornea series was stationary), and were converted to stationary series by taking the inverse.

The appropriate model for each surgeries was chosen according to the model criterion measures, it corresponds the minimum values of AIC, BIC, MAE, RMSE, for forecasting issue the tests have proved were efficiency of the selected models, the models that have been built are effective by adopting to forecast the future cases, the possibility of benefiting from the proposed models to forecast the numbers of cases at each new surgeries by updating the data for each period of time

and rebuilding the models in the same manner. And from the models, it is clear that the number of cases is increasing for all clinic in the coming few months, it is possible by updating the models to reach inflection point, to know the expected end of the cases.

References:

[1] albasar.wikipedia.org
[2] Alzawbaee, O. M. M., Hamd A. A., and Mahmood. A. O. (2020), using time series analysis to forecast the prevalence of the COVID-19 pandemic (IRAQ and gulf corporation council countries. Journal of critical reviews, 7(15).
[3] Alzawbaee, O. M. M., AL-Takriti, M. S. (2013), A New Suggested Technique to Determine the Best Models of Time Series. Damascus University Journal for Basic Sciences, 29 (1).
[4] Bisgaard, Søren and Kulahci,Murat , (2011) time series analysis and forecasting by example, A John Wiley & Sons, Inc., publication.
[5] Brokwell, Peter J and Davis, Richard A, (2002) Introduction to Time Series and Forecasting (second edition), Springer Texts in Statistics, USA.
[6] Box, Gerorge E.P , Jenkins, Gwilym M and Reinsel, Gregory C, (1994)Time Series and Forecasting and Control (3th edition).
[7] charity<medadcenter.org
[8] Cryer, Jonathan D and Chan, Kung-Sik, (2008) Time series Analysis with Applications in R, Springer Texts in Statistics, USA.
[9] Dickey D. A., Fuller W. A. (1981), Likelihood ratio statistics for autoregressive time series with a unit root, Econometrica, 49 (4), pp. 1057-1072. DOI: 10.2307/1912517.
[10] Elliott G., Rothenberg T., Stock J. (1996), Efficient Tests for and Autoregressive Unit root, Econometrica, 64 (4), pp. 813-836. DOI: 10.2307/2171846.
[11] Engle R. F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica: Journal of the Econometric Society, 50 (4), pp. 987-1007. DOI: 10.2307/1912773.
[12] Gujarati D. N., Porter D. C. (2009), Basic Econometrics, 5th Edition, McGraw Hill Inc., New York.
[13] He Z., Tao H. (2018), Epidemiology and ARIMA model of positive-rate of influenza viruses among children in Wuhan, China: A nine-year retrospective study, International Journal of Infectious Diseases, 74, pp. 61-70. doi: 10.1016/j.ijid.2018.07.003.
[14] Hyndman R. J., Athanasopoulos G. (2018), Forecasting: principles and practice, OTexts, Melbourne.
[15] Ljung G. M., Box G. E. (1978), On a measure of lack of fit in time series models, Biometrika, 65 (2), pp. 297-303. DOI: 10.2307/2335207.
[16] www.albasar-sudan.org
[17] www.gleauty.com
[18] Xinghua,Chang and Meng,Gao, (2012) seasonal autoregressive integrated moving average model for precipitation time series, Journal of Mathematics and Statistics 8 (4): 500-505.