Arma Type Modeling of Certain Non-stationary Time Series in Calabar

B.O. EKPENYONG[*]

Department of Mathematics and Statistics, Cross River University of Technology, Calabar, Nigeria
*Corresponding author: budyekpes@yahoo.com

Abstract Divergently different time series are considered in this article. The monthly passengers traffic at Cross lines limited, Calabar from 1990 to 2015 and monthly incidence of tuberculosis diseases at University of Calabar Teaching Hospital based from 1990-2015. The research adopted the statistical models based on time series analysis by Box and Jenkins methodology via the autocorrelation and the partial autocorrelation functions which showed that the two series are not stationary. Logarithm transformation was used to stabilize the variances of the two series and the residual autocorrelation and the partial autocorrelation functions is made stationary. Both regular and seasonal differencing was applied to the two-transformed set of data to obtain stationary series. The study employed ARIMA model on the classes of the two series, and the parameters of the identified model were estimated by the use of SPSS. The two models so chosen were ARMA $(2,1,0)$ x $(1,1,1)_{12}$ for passengers' traffic and ARMA $(1,0,1)$ x $(1,1,2)_{12}$ for tuberculosis cases and forecasts was done for 12 months for the two series. The adequacy of the model was achieved and model fit for passengers traffic yields R-square, RMSE and MAPE of 0.876, 9.137, and 27.479 respectively and for tuberculosis cases yields R-square, RMSE and MAPE of 0.614, 6.785 and 26.522 respectively, recommendation and conclusion was made for the area of study.

Keywords: *ARMA, non-stationary time series, modeling, passenger's traffic, tuberculosis diseases*

1. Introduction

A time series is defined as a sequence of observations (measurements) ordered by time $\{xt\}$, tET. We restrict ourselves to equidistant time series, i.e the parameter set is a finite set of equidistant points of time; $T = \{1,2,3, \dots, N\}$ observation $(Y_0, Y_{-1}, Y_{-2}, Y_{-3}, Y_{-4}, \dots.)$ could have obtained earlier or later observations $(Y_{N+1}, Y_{N+2}, Y_{N+3}, \dots.)$ had the process been observed for more time. The observed sample $(Y_1 Y_2, Y_3, Y_4, \dots Y_N,)$ could then be view as a finite segment of a doubly infinite sequence, devoted.

$$\left\{Y_t,\right\}_t^{\infty} = -\infty = (\dots.Y_{-1}, Y_0, Y_1, Y_2, Y_3, \dots.Y_{N+1}, Y_{N+2}, , Y_{N+3}, \dots.).$$

Typically, a time series $\left\{Y_t,\right\}_t^{\infty} = -\infty$ is identified by describing the ith element, [1] for example, a time series is a series whose values at date t are simply the date of the observation: $Y_t, = t$.

[2] described the time series as a set of observations generated sequentially in time. The set of observations generated sequentially in time are available for passengers' traffic, population figures, health figures, prices of all kinds of goods and merchandize and other innumerable sequence based on industrial, economic and social phenomena.

Observations in these series vary due to several factors. Some can be detected easily, such as seasonal variations in prices, variation due to the state of the economy in population and major calamities in the health cases. Assuming that it is possible to separate the contributions due to these detectable causes of variation, there would remain some variation which can only be attributed to random causes as in any natural phenomenon.

Tuberculosis (TB) is a chronic, infectious disease caused by bacteria generally to as "mycobacterium tuberculosis complex". The most important source of TB is an untreated pulmonary TB (PTP) patient. When such a person coughs, spits or sneezes, tiny droplet nuclei containing the tubercle germs are released. Transmission is through inhalation of these droplet nuclei. TB is curable, provided patients are detected early and treated promptly according to the National Tuberculosis and Leprosy Control Programme (NTBLCP) guidelines. Tuberculosis diseases constitute one of the major public health problems in Nigeria. In addition, the social stigma associated with these diseases further compounds the problem.

2. Material and Methodology

2.1. Data Collection Technique

The data type and source of this paper mainly, the secondary time series data in its analysis. The data were collected as follows:

(i) Monthly passenger's traffic at Cross Lines Limited, Calabar.

(ii) Monthly incidence of tuberculosis diseases from University of Calabar Teaching Hospital.

2.1.1. Model Speciation Criteria

When more than one model is selected from the process enumerated in the study, then it becomes necessary to use criteria to select best model amongst them. The two most common criteria are:

(i) Akaike Information Criterion (AIC)

$$AIC = T.ln(\sum_{t=1}^{r}\hat{\epsilon}_t^2) + 2n \qquad (1)$$

(ii) Schwartz (or Bayesian) information criterion (SIC)

$$SIC = T.ln\left(\sum_{t=1}^{r}\hat{\epsilon}_t^2\right) + n.\ln(T) \qquad (2)$$

Where T is sample size, n is number of parameters (e.g. $P + q$), and $\hat{\epsilon}_t^2$ are estimated residuals from a given data. [3] state the choice of the best model is based on the model selection criteria, the minimum model selection criteria compared to others.

2.1.2. Box and Jenkins Methodology on ARMA

The objective of Box-Jenkins methods is to forecast univariate time series using ARMA models that are in conformity with our analysis. The methods are in steps and are listed below:

a) Transform and difference the data to achieve stationarity.

b) By inspecting the sample autocorrelation and partial autocorrelations of the data, determine a suitable ARMA (p,q) model that parsimoniously but completely captures the dynamics of the process.

c) Estimate the parameters of the chosen model.

d) Perform diagnostic test on the reseals to ensure white noise. Make adjustments and estimates if needed.

e) Compute forecast.

2.1.3. Model Identification

Identification of non-stationary models, the inability of the ACF of PACF to die out is an indication of non-stationarity. A non-stationary model can be made stationary by differencing. An example is a case where autocorrelations are large and fails to die out at higher lags. Simple differencing will need to be adopted to reduce the autocorrelations in such data. However, if a very heavy periodic component remains, then it may be evident by large correlations at lags 12, 24, 36… for seasonal data.

The mixed ARMA (p,q) model can be identified by the fact that the sample autocorrelation plot of Pth order autoregressive component and qth order moving average component is a mixture of exponential and damped vine wave after the first $q - p$ lags. However the partial autocorrelation function is dominated by a mixture of exponentials and damped sine wave after the first $p - q$.

2.1.4. Estimation of Parameters for ARMA (p,q) Process

Suppose $(Y_1, Y_2, Y_3, Y_4, \ldots Y_t)$ are generated by an ARMA(p,q) model with normal innovations, then the Y's are normal with mean and covariances depending on the parameter vectors consisting of the unknown coefficient of the lag polynomials. If the lags order p and q are known and small, Q can be estimated by maximum likelihood, however, this is computationally demanding when the sample size T is large as calculations involves inverting $T \times T$ covariance matrix.

In practice, most practitioners maximize an approximation to the likelihood based on slight change in the initial conditions. For example, in the $AR(p)$ model, if we condition on the first p values $Y_1, Y_2, Y_3, Y_4, \ldots Y_p$ and examine the process from $t = P + 1$, then maximum likelihood is equivalent to the least square regressing of the y_t on its p lagged values. Similar modification of the initial conditions in the general ARMA case leads to a non-linear least squares regression.

3. Estimation Results

3.1. Model Parameter Estimation for Traffic Flow

Technique used for estimating the parameters of the models identified is [4] and Marquardt Algorithm for linear least square estimates and these estimates for the suggested models and the fitted models for passengers' traffic in Calabar. Considering the estimates of the parameters of the models identified, we may see that the probable model to work with is $(2,1,0) * (1,1,1)_{12}$. A general multiplicative model is represented as:

$$\phi_p(B)\phi_p(B^s)\nabla^d\nabla_s^D Y_t = \theta_q(B)\theta_Q(B^s)\varepsilon_t. \qquad (3)$$

Substituting the estimates of the parameters, we have

$$= \left(1 + 0.4954B + 0.2624B^2\right)\left(1 - 0.5107B^{12}\right)\nabla\nabla_{12}Y_t$$
$$= \left(1 - 0.9934B^{12}\right)\varepsilon_t \qquad (4)$$

$$= \left(\begin{matrix}1 + 0.4954B \\ + 0.2624B^2\end{matrix}\right)\left(1 - 0.5107B^{12}\right)\left(1 - B\right)\left(1 - B^{12}\right) \qquad (5)$$
$$= \varepsilon_t - 0.9934\varepsilon_{t-12}.$$

The model can be written explicitly as:

$$= Y_t - 0.5046Y_{t-1} - 0.233Y_{t-2} - 0.2624Y_{t-3}$$
$$-0.1340Y_{t-4} - 1.5107Y_{t-12} + 0.7623Y_{t-13}$$
$$+0.4860Y_{t-14} + 0.3964Y_{t-15}$$
$$-0.5107Y_{t-24} - 0.2577Y_{t-25} \qquad (6)$$
$$+0.3870Y_{t-26} - 0.1340Y_{t-27}$$
$$= \varepsilon_t - 0.9934\varepsilon_{t-12}.$$

Appendix 2 in Figure 2 show the plot of the residuals of ACF, PACF and original series for the passenger's traffic, which both are white noise series. Since their residuals are small, this is an evidence of good fitting the

model to the series. Hence, it is reasonable to conclude that the model is adequate.

3.1.1. Forecasting

Now that we have shown that the passenger traffic data is well fitted by a model of the form in 3.1 it is convenient to values inserted, future using this model as with the parameter values inserted, future values of the series may be forecast. The minimum mean square error forecast at lead time L and origin t may be given by

$$
\begin{aligned}
\hat{y}_t\left(l\right) = {}& 0.5046 Y_{t+l} + 0.2330 Y_{t+l-2} \\
&+ 0.2624 Y_{t+l-3} + 0.1340 Y_{t+l-4} \\
&+ 1.5107 Y_{t+l-12} + 0.7623 Y_{t+l-13} \\
&+ 0.4860 Y_{t+l-14} + 0.3964 Y_{t+l-15} \\
&+ 0.5107 Y_{t+l-24} - 0.2577 Y_{t+l-25} \\
&+ 0.3870 Y_{t+l-26} - 0.1340 Y_{t+l-27} \\
&+ \varepsilon_{t+l} - 0.9934 \varepsilon_{t+l-12}.
\end{aligned}
\tag{7}
$$

Forecast data is then generated to cover the period January to December 2016, using December 2015 as the origin as shown in Table 3.

3.2. Model Parameter Estimation for Tuberculosis Cases

From estimates of the parameters identified we may now consider $(1,0,1) * (1,1,2)_{12}$ as the model to work with a general multieliative model is represented as:

$$
\phi_p\left(B\right)\phi_p\left(B^s\right)\nabla^d \nabla_s^D Y_t = \theta_q\left(B\right)\theta_Q\left(B^s\right)\varepsilon_t.
\tag{8}
$$

Substituting the estimates of the parameters, we have

$$
\begin{aligned}
&= (1 - 0.9050B)(1 + 0.4092B^{12})\nabla_{12}Y_t \\
&= (1 - 0.8B)(1 - 0.4971B^{12} - 0.494B^{24})\varepsilon_t.
\end{aligned}
\tag{9}
$$

The model can be written explicitly as

$$
\begin{aligned}
Yt - {}& 0.9050 Y_{t-1} - 0.5908 Y_{t-12} + 0.5347 Y_{t-13} \\
&- 0.4092 Y_{t-24} - 0.3703 Y_{t-25} \\
&= \varepsilon_t - 0.8 \varepsilon_{t-1} - 0.4915 \varepsilon_{t-12} + 0.3977 \varepsilon_{t-13} \\
&- 0.4940 \varepsilon_{t-24} + 0.3952 \varepsilon_{t-25}.
\end{aligned}
\tag{10}
$$

Appendix 2 in Figure 1 show the plot of the residuals of ACF,PACF and original series for the tuberculosis cases, which both are white noise series. Since their residuals are small, this is an evidence of good fitting of the model and hence adequacy of the model is achieved.

3.2.1. Foresting

From the analysis the minimum mean square error forecast at lead from L and origin is be given by

$$
\begin{aligned}
\hat{y}_t\left(l\right) = {}& 0.9050 Y_{t+l-1} + 0.5908 Y_{t+l-12} - 0.5347 Y_{t+l-13} \\
&- 0.4092 Y_{t+l-24} + 0.3703 Y_{t+l-25} + \varepsilon_{t+l} - 0.8 \varepsilon_{t+l-1} \\
&- 0.4971 \varepsilon_{t+l-12} + 0.3977 \varepsilon_{t+l-13} \\
&- 0.494 \varepsilon_{t+l-24} + 0.3952 \varepsilon_{t+l-25}.
\end{aligned}
\tag{11}
$$

This forecast models provide the forecast value for the period of January to December 2016 in Table 6.

4. Discussion

A time series model was obtained for Passengers Traffic at Cross Lines Limited, Calabar for a period of 26 years. The use of statistical packages was employed for accuracy which embraced the [4] methods implemented through an algorithm developed by [5] and Marquardt algorithm for non-linear least squares (slightly modified by G.T. Wilson). In the identification of the model, the Autocorrelation and Partial Autocorrelation Functions were used as tools in identifying the model. The Akaike Information Criterion (AIC) was used to select best model amongst those identified. The model chosen is the one with least AIC, that is the model that minimized the AIC. The auto-covariance and special characteristics for some other multiplication models can be found in [4] and in [6]. Under box and Jenkins methods, stationarity is important because non-stationarity can swamp the effect of other dynamics, [7]. The partial autocorrelation function is useful in determining the order of a pure autoregressive process. As [8] states, the partial autoregressive at lag k, denoted by $\emptyset_{KK}$ is the correlation between y_t and y_{t-k} after removing the effect of the intervening variables y_{t-k} after removing the effect of the intervening variables $y_{t-1}, y_{t-2},$ ------y_{t-kt1}. The procedure moderated by [9] postulates that before data are analyzed; sometimes they are filtered, or treated in a particular way. This calculation is particular done using the auto-covariance-generating function. An iterative procedure for deciding on transformations of a single variable is given by [10,11,12] and [13] all showed how suitable transformations of predictor variables are sometimes suggested plotting the data in various ways. The logarithmic transformation is a special case of a family of transformations called the [14] or power transformation.

The residual autocorrelation and partial autocorrelation function of the series showed that it is stationary. Logarithm transformation was used to stabilize the variance of the series. Both regular and seasonal differencing was applied to the log-transformed data to obtain a stationary series. ARMA $(2, 1, 0)$ x $(1, 1, 1)_{12}$ model was the model that was fitted to the series adequately. The model obtained was then used to make forecasts of passenger's flow for the company. The forecasts were made for a period covering 12 months.

Secondly, the monthly set of data for Tuberculosis diseases from University of Calabar Teaching Hospital was also analyzed using Time Series analysis. The data set was for a period of 25 years. The same packages as earlier mentioned were also employed during analysis for accuracy. Also in the identification of the model, the Autocorrelation and Partial Autocorrelation Function were used as tools in identifying the model. The Akaike information Criterion (AIC) was used to select best model among those identified. The model chosen is the one with least AIC, that is the model that minimizes the AIC. The model fit in Table 7 for passengers traffic yields R-square, RMSE and MAPE of 0.876, 9.137, and 27.479 respectively and in Table 8 for tuberculosis cases yields R-square, RMSE and MAPE of 0.614, 6.785 and 26.522 respectively.

The autocorrelation function and partial autocorrelation function of the series showed that it is not stationary. So it became necessary to carry out logarithm transformation,

and of course appropriate differencing. The model which fitted the TB series is the ARMA $(1, 0, 1)$ x $(1, 1, 2)_{12}$ model. The derived model was then used to make forecasts of Tuberculosis cases for the Hospital for a period coving 12 months.

5. Conclusion

There has been significant reduction in the monthly passengers' flow in Cross Lines Limited with some evidence of reduction in the number of buses plying different routes. This has been attributed to the introduction of different Urban Mass transit transportation by Federal Government, other State Governments and other allied transport companies operating in the country. Also long waiting time, distance to passengers' residence, attitude of company's workers and availability of affordable transport at any time have been identified as some of the reasons why passengers prefer patronizing the private owned transport company. They have however, remained in business and are functioning proficiently having geared up to purchasing new buses and rehabilitating the old ones.

In order to maintain the low trend in prevalence of Tuberculosis cases, the National TB control programme should identify collaboration with private health care providers as one of the future challenges facing the accelerated Expansion of the treatment strategy in the state. Implementation of treatment strategy in the state should not be limited to the Government own clinics, Hospitals, and some missionary health facilities since huge patronage in the made in the Private Health Clinics by patients. It is not enough to combat this problem in Government sector only, there is need to involve private sector in the implementation of treatment strategy to provide another front to combat this public health problem and to help in providing/scaling up the case finding activities of NTBLCP.

References

[1] Hamilton, James.D.(1994). "The series analysis new Jersey", United Kingdom Princeton University press.

[2] Bhat, U,N. G.K.Miller (2002) "Estimation of the coefficient variation for unobservable services in time series" *j. math. Sc. Society math. Sc. i.* delhi, india.

[3] Onwukwe C.E and Nwafor G. O. (2014). "A Multivariate Time Series Modeling of Major Economic Indicators in Nigeria" *American journal of applied mathematics and statistics*,vol.2, No.6 376-385.

[4] Box G.E.F. and G.M, Jenkins.(1976) "time series analysis :forecasting and control rev. ed."San Franciso holden – day.

[5] Mcleod, A.I and P.R. Holanda sales (1983) "An algorithm for approximate likelihood calculation of ARMA and seasonal ARMA models", *journal of the royal statistical society (series G)* 32,211-224.

[6] Abraham. B. and j. Ledolter (1983). "Statistical methods for forecasting." New York: wiley.

[7] Granger, C.W.J. and Paul Newbold (1974). "Spurious regressions in econometrics" *journal of econometrics* 2:111;20

[8] Myers, Raymond H. (1990) "Classical and modern regression with applications" 2nd ed. Boston:Duxbury press.

[9] Anderson Brain D.O and John B. Moore (1979) "optimal filtering. Eaglewood cliffs,"N.J. : Prentice – hall.

[10] Box G. E. P and Tidwell, A. (1962). "Transformations of independent variable" *technometrics 4*,431-550.

[11] Hoerl, E,A. (1954) "Fitting curves to data" Chemical business handbook.

[12] Dolby J. L. (1963)"A quick method of choosing transformation" *Technometrics* 5,317-325.

[13] Turkey, T.W. (1961) "Discussion, emphasizing the connection between analysis of variance and spectrum analysis" *Technometrics 3*:1-29.

[14] Box G.E.P and D.R. Cox (1964) "An analysis of transformation" *Journal of royal statistical society* B26; 211-243.

[15] Federal ministry of health (2002) "A guide for community health extension workers: providing support to TB patients and communities at TB (DOTS) treatment centre."

APPENDIX 1

Table 1. Suggested Models for Passengers' Traffic

S/NO	SUGGESTED MODELS IDENTIFIED	PARAMETER ESTIMATES	STD. ERROR	t-RATIO	APPROX. PROBABILITY	RESIDUAL VARIANCE	SES	SBC	AIC
1	$(2.1.0)$ x $(1.1.0)_{12}$	p(1)= -0.4985	0.056	-8.8929	0.00000000	0.0648	19.2487	45.4775	34.3762
		p(2)= -0.2629	0.05462	-4.6789	0.00000439				
		PS(1)=-0.2186	0.576	-3.7931	0.00018039				
2	$(2.1.0)$ x $(2.1.0)_{12}$	p(1)= -0.4955	0.0562	-8.8098	0.00000000	0.0621	18.4659	38.7615	23.9597
		p(2)= -0.2554	0.0565	-4.5229	0.00000000				
		PS(1)= -0.2802	0.0588	-4.7646	0.00000297				
		PS(2) = -0.2276	0.0635	-3.5866	0.00039208				
3	$(0.1.1)$ x $(0.1.1)_{12}$	q(1)= 0.5137	0.0498	10.3115	0.0000000	0.0626	18.7336	31.6595	24.2586
		QS(1) = 0.3690	0.0568	6.4977	0.0000000				
4	$(0.1.2)$ x $(0.1.1)_{12}$	q(1) = 0.4957	0.0579	8.5686	0.00000000	0.0627	18.6987	36.8026	25.7013
		q(2) = 0.0544	0.0583	0.9343	0.35090532				
		QS(1) = 0.3741	0.0573	6.5249	0.00000000				

American Journal of Applied Mathematics and Statistics

Table 2. Fitted Models for Passengers' Traffic

S/NO	SUGGESTED MODELS IDENTIFIED	PARAMETER ESTIMATES	STD. ERROR	t-RATIO	APPROX. PROBABILITY	RESIDUAL VARIANCE	SES	SBC	AIC
1	$(2.1.1) \times (2.1.1)_{12}$	p(1)= -0.4518	0.2018	-2.2386	0.02593558	0.0559	17.4248	33.0776	10.8749
		p(2)= -0.2506	0.0971	-2.5815	0.01032172				
		q(1)=0.0615	0.2084	0.2952	0.76814736				
		PS(1)=0.5683	0.0805	7.0644	0.00000000				
		PS(2)=-0.0414	0.0714	-0.5806	0.56195974				
		QS(1)= 0.9692	0.1518	6.3860	0.00000000				
2	$(2.1.0) \times (1.1.1)$	p(1)= -0.4954	0.0539	-9.3206	0.00000000	0.0545	17.4472	22.4958	7.6340
		p(2)= -0.2624	0.0538	-4.9940	0.00000101				
		PS(1)= 0.5107	0.0802	7.0031	0.00000000				
		QS(1) = 0.9934	0.0023	2.7429	0.00014302				
3	$(2.1.2) \times (1.1.1)_{12}$	P(1)=-0.6485	0.3541	-1.8316	0.06803238	0.0552	17.4133	33.0351	10.8325
		P(2)=-0.1696	0.1146	-1.4801	0.13991938				
		q(1)=-0.1370	0.3558	-0.3849	0.70055147				
		q(2) =0.2021	0.2305	0.8767	0.38136060				
		PS(1)=0.5632	0.0819	6.8806	0.00000000				
		QS(1)=0.9882	0.3519	2.8080	0.00532023				
4	$(1.1.0) \times (1.1.1)_{12}$	p(1) =-0.3915	0.0519	-7.5396	0.00000000	0.0597	18.7606	38.1575	27.0561
		PS(1) = 0.5054	0.0771	6.5528	0.00000000				
		QS(1) = 0.9635	0.1119	8.6074	0.00000000				

Table 3. Forecasts; Model: (2,1,0)(1,1,1) Seasonal lag: 12 (Input: TRAFFIC, Start of Origin: 1 End of origin: 312)

		Forecast	Lower 90.0000%	upper 90.0000%
1	313	37317.19	25281.17	54659.12
2	314	31786.60	20639.56	48953.95
3	315	9754.91	6090.86	15623.13
4	316	19192.72	11305.73	32581.76
5	317	9537.15	5383.55	16895.41
6	318	6018.35	3267.17	11086.19
7	319	5089.69	2656.90	9750.06
8	320	4297.99	2164.47	8534.53
9	321	5767.38	2806.34	11844.51
10	322	16608.51	7823.61	35257.75
11	323	23360.76	10665.28	51168.40
12	324	33463.31	14824.92	75534.50

Table 4. Suggested Models for Tb

S/NO	SUGGESTED MODELS IDENTIFIED	PARAMETER ESTIMATES	STD. ERROR	t-RATIO	APPROX. PROBABILITY	RESIDUAL VARIANCE	SES	SBC	AIC
1	$(1.1.0) \times (2.1.0)_{12}$	p(1)=-0.4896	0.5015	-9.5153	0.00000000	0.1328	38.5386	255.2229	244.2445
		PS(1) = - 0.6205	0.0562	-11.0509	0.00000000				
		PS(2)=-0.3515	0.562	-6.2582	0.00000000				
2	$(0.1.2) \times (0.1.1)_{12}$	q(1)=0.8351	0.0579	14.4330	0.00000000	0.0910	27.3757	157.3767	146.3983
		q(2)=0.0440	0.0580	0.7589	0.44855186				
		QS(1)=0.8412	0.0412	20.4062	0.00000000				

Table 5. Fitted Models for Tb

S/NO	SUGGESTED MODELS IDENTIFIED	PARAMETER ESTIMATES	STD. ERROR	t-RATIO	APPROX. PROBABILITY	RESIDUAL VARIANCE	SES	SBC	AIC
1	(1.1.1) x (2.1.0)$_{12}$	p(1)= 0.0791	0.0686	1.1533	0.24974216	0.1066	31.0167	198.5745	183.9366
		q(1)= 0.8782	0.0340	25.8220	0.00000000				
		PS(1)= 0.6305	0.0560	-11.2668	0.00000000				
		PS(2)= 0.3744	0.0562	-6.6604	0.00000000				
2	(1.1.0) x (2.1.1)$_{12}$	p(1)= 0.5088	0.0502	-10.1346	0.00000000	0.1151	34.1610	226.4224	211.7845
		PS(1)= 0.00162	0.0819	-0.1980	0.84315376				
		PS(2)= 0.00775	0.0733	-1.0568	0.29152107				
		QS(1)= 0.7911	0.0655	12.0861	0.00000000				
3	(1.0.1) x (1.1.2)$_{12}$	p(1)= 9050	0.0707	12.6767	0.00000000	0.0885	26.3022	156.4980	138.1832
		q(1)= 8000	0.0974	8.0224	0.00000000				
		PS(1)= 4092	0.6947	-1.9695	0.03331386				
		QS(1)= 4971	0.6722	2.1968	0.04408960				
		QS(2)= 4940	0.5494	2.0887	0.02771946				
4	(1.1.1) x (1.1.2)$_{12}$	p(1)= 0.0598	0.0662	0.9030	0.36730017	0.09131	27.3049	167.8703	149.5729
		q(1)= 0.8919	0.0317	28.1409	0.00000000				
		PS(1)= 0.6465	0.7583	-0.8526	0.39462901				
		QS(1)= 0.1658	0.7368	0.2250	0.82214006				
		QS(2)= 0.5799	0.6071	0.95653	0.34025935				

Table 6. Forecasts; Model: (1,0,1)(1,1,2) Seasonal lag: 12 (Input: TB, Start of Origin: 300)

	Forecast	Lower 90.0000%	upper 90.0000%
1	35.42759	21.86458	57.40401
2	30.09353	18.52340	48.89063
3	28.57741	17.55221	46.52797
4	11.88015	7.28392	19.37665
5	10.51909	6.44016	17.18146
6	8.11058	4.95975	13.26308
7	13.57904	8.29583	22.22687
8	16.53476	10.09363	27.08620
9	18.25258	11.13512	29.91943
10	26.47313	16.14166	43.41727
11	21.65110	13.19581	35.52416
12	23.98559	14.61349	39.36833

Table 7. Model Fit for Passengers Traffic

R-squared	.876
RMSE	31189.172
MAPE	27.479
MaxAPE	367.841
MAE	20108.413
MaxAE	151750.245
Normalized BIC	20.791

Table 8. Model Fit for Tuberculosis Cases

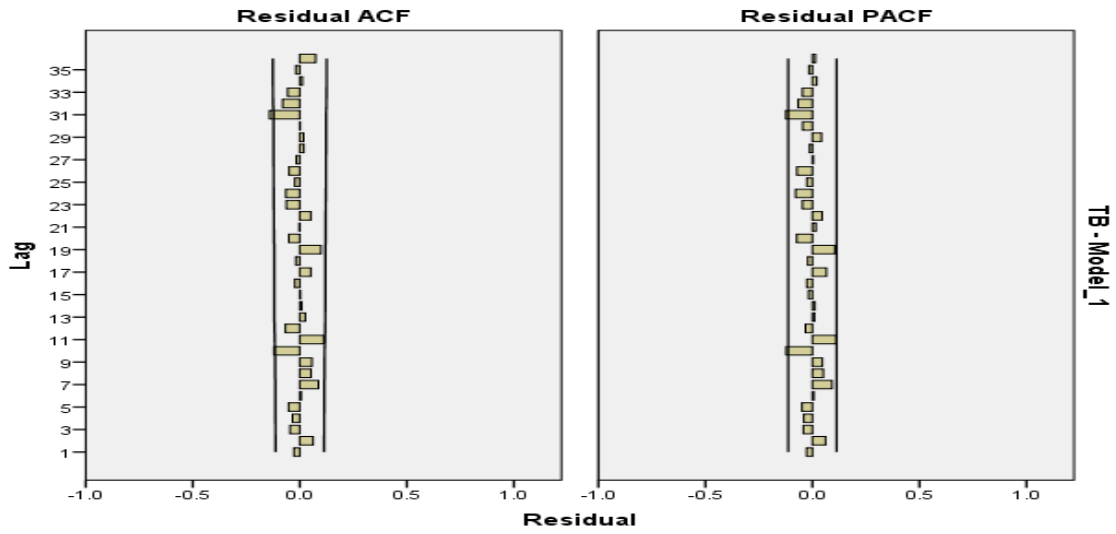|	
R-squared	0.614
RMSE	6.785
MAPE	26.522
MaxAPE	253.811
MAE	5.288
MaxAE	24.727
Normalized BIC	3.943

APPENDIX 2

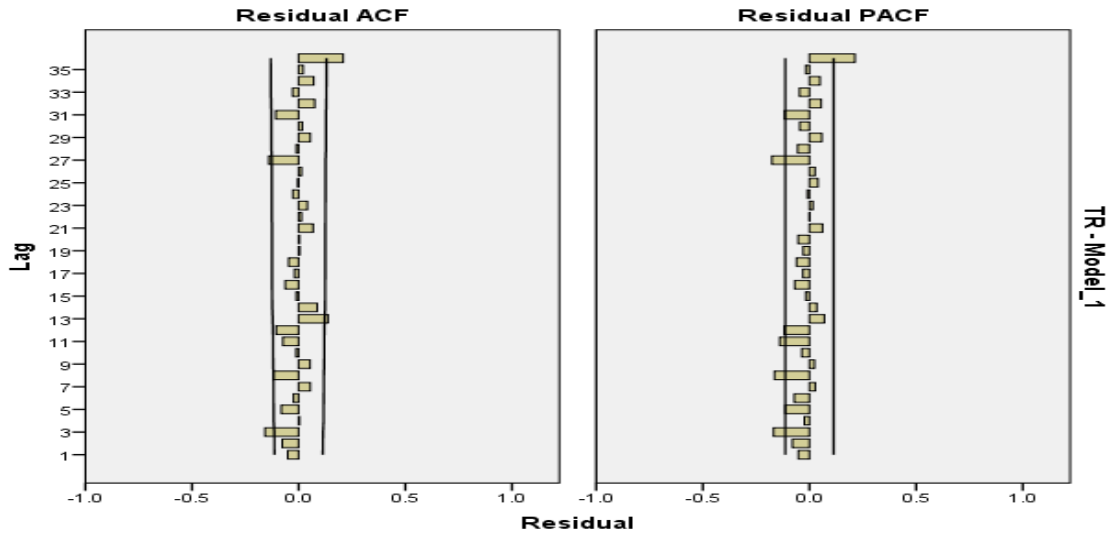**Figure 1.** The plot of the residual ACF and PACF of tuberculosis cases

Figure 2. The plot of the residual ACF and PACF of passenger's traffic
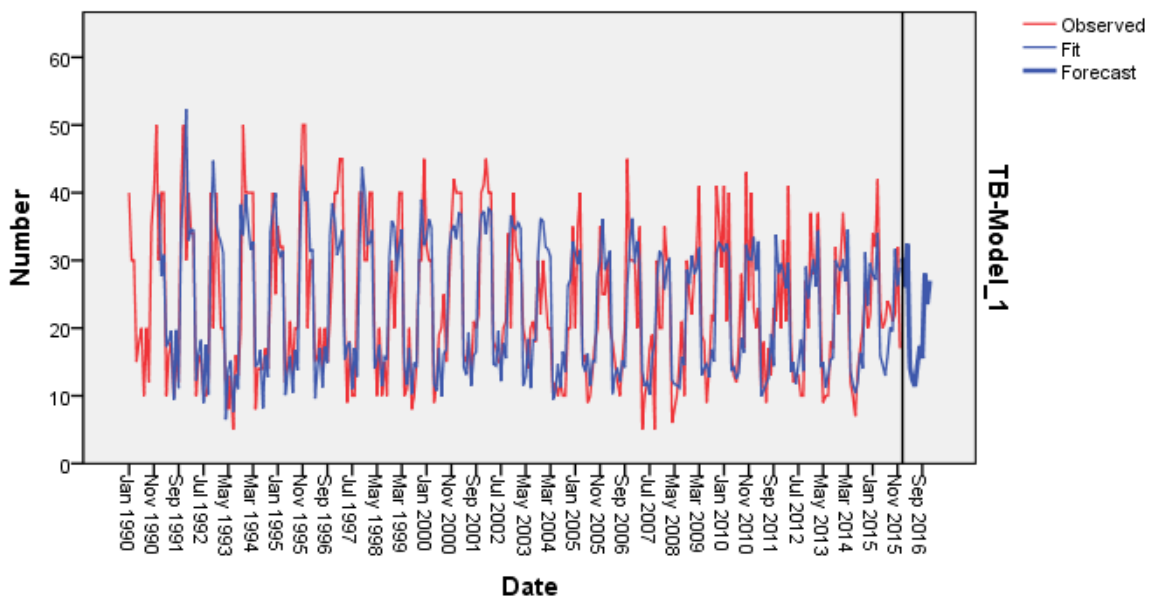

Figure 3. The plot of original series, fit and the forecast of tuberculosis cases
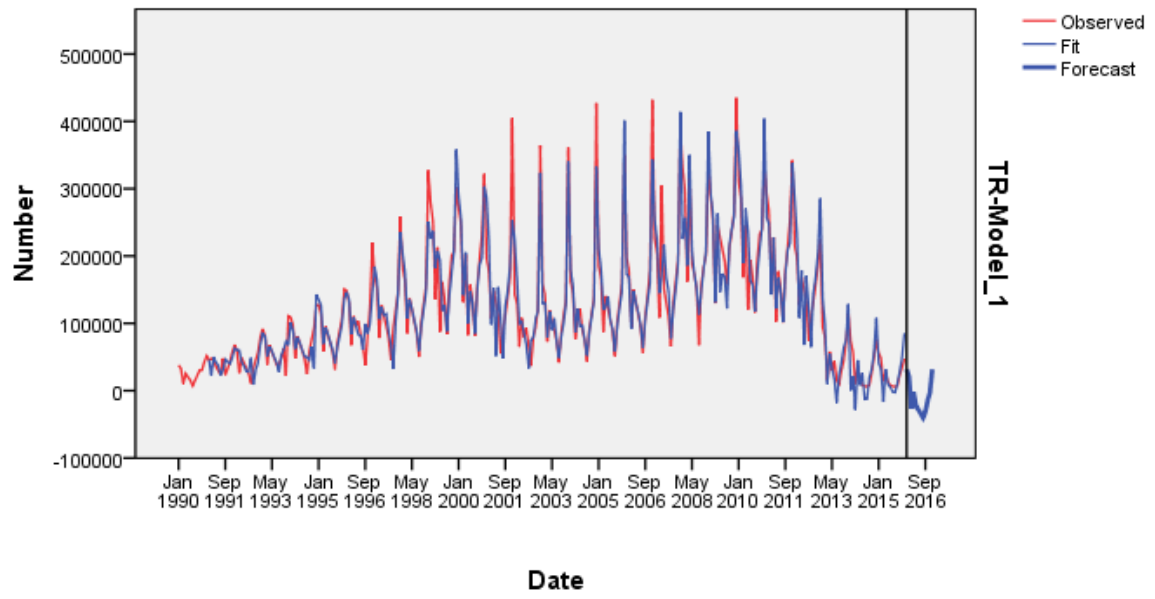
Figure 4. The plot of original series, fit and the forecast of passenger's traffic