TIME SERIES MODEL ON FEMALE BIRTH RECORDS AND THE PROPENSITY OF FEMALE BIRTHS IN ABAK LOCAL GOVERNMENT: A STUDY OF GENERAL HOSPITAL, UKPOM ABAK

www.akwapolyjournal.org
P-ISSN: 2536-6733 | E-ISSN: 2616-0773

Supported by: tetfund TERTIARY EDUCATION TRUST FUND
TETFund/DESS / POLY / IKOT-OSURUA/ ARJ /1-2)

Essien, Eduma & Udoh, Grace
Department of Statistics,
Akwa Ibom State Polytechnic, Ikot Osurua

Abstract
This research study utilizes the Time Series Model analytical approach to investigate and predict the monthly records of female births at the General Hospital in Ukpom Abak, estimating the expected pattern of female births in the Abak Metropolis. Currently, there is a significant neglect of the documentation of female births, leading to a need for more reliable statistics on female births. This lack of data adversely affects the accuracy of predicting the number of daughters born to a female as she progresses through her reproductive age. Furthermore, with rapid population growth, government planners require assistance planning for various aspects such as workforce, education, healthcare facilities, and other essential amenities. Despite the efforts of government agencies such as the Bureau of Statistics to ensure proper birth records, parents have shown reluctance to provide the necessary information. The researcher suggests using time series analysis, which involves collecting well-defined data items obtained through repeated measurements over time, to gain reliable insights into the trends and patterns of female births. The analysis's findings could prove invaluable to governmental planners and other stakeholders managing the population.

Keywords: *Sarima, Aic ,Bic,Time Series,Health*

Introduction
Compulsory birth registration with governmental agencies originated in the UK in 1853. Now, most countries have laws regulating the registration of births, which is the responsibility of the mother's physician, midwife, hospital administrator, or the child's parents. The original purpose of vital statistics was tax purposes and determining the

available military workforce. The practice of registering births with churches continued into the 19th century.

Government agencies retain the actual birth record and provide certified copies or representations of the original birth record upon request. These records are used to apply for government benefits, including passports. Birth predictions are essential for labour force planning, social welfare, education, health, housing, transportation, and public infrastructure. Planners and decision-makers in Commonwealth agencies, governments, and private enterprises rely on information on future births to create development plans.

Female births are of particular interest since they can be used to obtain the gross reproduction rate (GRR) and net reproduction rate (NRR), which measure the average number of daughters born to a female if she passes through her lifetime. The case study focused on Ukpom General Hospital's recorded female births between 2006 and 2016. The hospital is vital since it serves women from diverse locations within the local government area who come to deliver their babies.

Many studies have analyzed time series models, focusing on the autoregressive integrated moving average (ARIMA) models. These models have been utilized in various fields, including but not limited to population growth, birth rates, malaria infections, and consumer price index (CPI) data. For instance, Wangdi et al. (2010) developed a temporal model utilizing ARIMAX analysis to forecast malaria infections in Bhutan. Akhter (2013) created a model for Bangladesh's monthly CPI from 2000 to 2012 using seasonal ARIMA models. Akpanta and Okorie (2015) used SARIMA models to analyze Nigeria's CPI data from 1996 to 2013.

Additionally, time series models have been employed to analyze seasonal and non-seasonal variations in monthly birth data for regions such as Norway, Ontario, and London. Lee et al. (2006) performed a systematic review of countries where the seasonal pattern of preterm birth has been reported and analyzed the seasonal variability of preterm birth in a London-based cohort. Their results demonstrated that the seasonality in the London-based cohort differed from other developed countries that previously reported the seasonal pattern of preterm birth. Therefore, it is crucial to model female birth using SARIMA models, particularly since preterm birth has a seasonal pattern. This research aims to determine the trend of female births, assess the series' stationarity using ACF and PACF, and fit an appropriate time series model to the data using the SARIMA model. Birth forecasts are vital to any developmental plan, especially in developing countries where female births are not monitored from reproductive age to menopause. Predicting the trend of female births can help estimate the population of a country or state at a given time, taking into account the average number of children a woman is expected to have and the probability of their survival.

Method of Analysis

On generalizing data collected for time series analysis, two Seasonal Autoregressive Integrated Moving Average Models (SARIMA) are identified. They are SARIMA (1,1,1)(0,1,1)12 and SARIMA (1,1,1)(1,1,1)12 for female births. Jain et al., 1985 believe that the Autoregressive Integrated Moving Average (ARIMA) Model is adequate for analyzing births. The result shows that the SARIMA (1,1,1)(0,1,1)12 model is appropriate for female births predictions.

Seasonal Autoregressive Integrated Moving Average Model (SARIMA)

1. Represent the series $\{Y_t\}$ in Buys-Ballot table which contains
   $Y_{t-2,}\ Y_{t-1,}\ Y_{t,}\ Y_{t+1,}\ Y_{t+2,\ldots\ldots,}$ the within periods, and
   $Y_{t-2s,}\ Y_{t-s,}\ Y_{t,}\ Y_{t+s,}\ Y_{t+2s,\ldots\ldots}$ the between periods relationships.
2. Fit a non-seasonal ARIMA model for the series.
   $\emptyset_p(B)(1-B)^d Y_t = \theta_q(B)a_t$ then a_t will not be white noise if

, $\dot{p}_{j(s)} = E\left[(a_{t_{-js}} - \mu_a)(a_t - \mu_a)\right] / \sigma_a^2$ = 1, 2, 3… represents the

unexplained between periods relationship.

3. The between periods can be represented as an ARIMA model.
   $\varphi_P(B^S)(1-B^S)^D Y_t = \vartheta_Q(B^s)a_t$
   $where,$
   $\varphi_P(B^S) = 1 - \varphi_1(B^S) - \varphi_2(B^S) - \cdots \varphi_P(B^{PS}).$
   $and \vartheta_Q(B^s)a_t = 1 - \vartheta_1(B^S)a_t - \vartheta_2(B^{2S})a_t - \cdots \vartheta_Q(B^{QS})a_t$
   are polynomials in B^S with common roots. The root of these polynomials liesoutside the unit circle, a_t and is zero mean white noise process.

4. combining steps 2 and 3 we get the multiplicative seasonal ARIMA model

$$\varphi_P(B^S)\emptyset_p(B)(1-B)^d(1-B^S)^D Y_t = \theta_q(B)\vartheta_Q(B^s)a_t$$

where, $\emptyset_p(B)$ and $\theta_q(B)$ are regular autoregressive and moving average factors (polynomials) and $\vartheta_Q(B^s)$ and $\varphi_P(B^S)$ are the seasonal autoregressive andmoving average factors (or polynomials) respectively.

5. The model is denoted as ARIMA (p,d,q)*(P,D,Q)$_s$ where the sub index S refers to the seasonal period.

The forecast can be obtained using
$$\varphi_P(B^S)\emptyset_p(B)(1-B)^d(1-B^S)^D Y_t = \theta_q(B)\vartheta_Q(B^s)a_t.$$

Diagnostic Check and Model Selection

Model Selection

The Akaike Information Criterion (AIC) and Normalized Bayesian Information Criterion (BIC) are used to select the most suitable model.

Diagnostic Checking

Diagnostic checking is basically used to know if the identified model is adequate for the data. According to Box and Jenkins (1970), there are several methods that can be used for diagnostic checking, but in this study, the diagnostic plots were used to check if the residuals of each model are uncorrelated at various lag.

The forecast was obtained using
$$\varphi_P(B^S)\emptyset_p(B)(1-B)^d(1-B^S)^D Y_t = \theta_q(B)\vartheta_Q(B^s)a_t$$

Data Analysis, Results and Interpretation

The data collected as shown in Table 1 is analyzed and presented

Data Analysis

Before model identification, parameter estimation and validation (diagnostic checking), some preliminary analysis such as time plot of the original series, monthly mean and yearly mean plot were obtained.

Time plot of the original Series

The plot of the original series against time is shown in Figure 1. The graph of the seriesshows that the series is nonstationary.

Monthly mean Plot

The monthly mean of the data set were computed and plotted against time as are shownin Table 3 and Figure 2. The highest monthly mean birth was recorded in the month of May with value 93, while August and November recorded the lowest yearly mean birth with value 55 see Table 2.

Model Identification

The Figure representing the original series displays characteristics of a nonstationary series with peaks occurring at equal intervals, indicating the presence of a seasonal trend. The original series' autocorrelation function (ACF) and partial autocorrelation function (PACF) in Figures 3 and 4 confirmed these characteristics. The inability of the autocorrelation function to rapidly decay at higher lags, such as a lag of 30, indicates that the series is nonstationary. To achieve stationarity, a different technique for series transformation was implemented. The stationary nature of the differenced series is apparent from the graph in Figure 5. The ACF and PACF plots of the differenced series in Figures 6 and 7 confirmed that the series was stationary. Examining the seasonal lags (for K=12, 24) in Figures 8 and 9, it is observed that the ACF and PACF have significant spikes, which indicates that a seasonal AR or MA term is present. Since the ACF and PACF of the differenced series cut off at significant lags, suggesting a mixed process, a seasonal ARIMA model may be appropriate.

The following seasonal ARIMA models were tentatively proposed based on the abovementioned features.
SARIMA $(1, 1, 1)(1, 1, 1)_{12}$
SARIMA $(1, 1, 1)(0, 1, 1)_{12}$

Estimation of Parameters

Based on estimation of parameters using statistical softwares (SPSS and GRETL), theresults showed that:

(i) SARIMA $(1, 1, 1)(1, 1, 1)_{12}$:

$\emptyset_1$	3= 0.33	SE = 0.136	t – value = 9.701
θ_1	= 0.847	SE = 0.087	t - value = 9.701
$\varphi_{12,1}$	= 0.04	SE = 0.231	t – value = 0.171
$\vartheta_{12,1}$	= 0.757	SE = 0.241	t – value = 3.138

ii) SARIMA (1,1,1 (0,1,1) 12:

$\emptyset_1 = 0.339$	SE = 0.135	t – value = 2.509
$\vartheta_{12,1} = 0.850$	SE = 0.086	t – value = 9.893
$\theta = 0.729$	SE = 0.126	t – value = 5.80

Diagnostic Checking

Having identified the model, and the parameters estimated, we now apply diagnostic check to test the adequacy of the models. Based on the Akaike Information Criterion (AIC) and the Normalized Bayesian Information Criterion (BIC), we compare the adequacy of the models.

SARIMA (1,1,1) $(1,1,1)_{12}$ BIC = 6.481 AIC = 1103.047
SARIMA (1,1,1) $(0,1,1)_{12}$ BIC = 6.434 AIC = 1101.102

The SARIMA (1,1,1) $(0,1,1)_{12}$ model has the least BIC value of 6.434 and AIC value of 1101. Since, the parameters are statistically significant the model is appropriate and adequate and could be used for prediction.

Interpretation of Results

From the result of the analysis, SARIMA (1, 1, 1) $(0, 1, 1)_{12}$ model for female births has the smallest BIC and AIC values, this indicates that it is a better model than the other estimated model. The fitted models can be written using backward shift operator.

$\Rightarrow$ SARIMA (2,1,1) $(1,1,1)_{12}$:

$$(1 - \emptyset_1 B - \emptyset_2 B^2)(1 - \varphi_{12,1} B_{12})\nabla\nabla_{12} Y_t = (1 + \theta_1 B)(1 + \vartheta_{12,1} B^{12})a_t$$

$$(1 - 0.3068B + 0.203B_2)(1 + 0.146B^{12})\nabla\nabla_{12} Y_t = (1 + 0768B)(1 + 0.710B^{12})a_t$$

$$1 = 0.306B + 0.203 B^{12})(1 + 0.146B^{12})\nabla\nabla_{12} Y_t = (1 + 0.768B)(1 + 0.710B^{12})a_t$$

Where

$$W_t = \nabla\nabla_{12} Y_t,$$

$\Rightarrow$ SARIMA (1,1,1) $(0,1,1)_{12}$

$$(1 - \emptyset_1 B) W_t = (1 + \theta_1 B)(1 + \vartheta_{12,1} B^{12})a_t$$

$$(1 - \emptyset_1 B) W_t = (1 + \theta_1 B)(1 + (\vartheta_{12,1} B^{12})a_t$$

$$(1 - 0.339B)W_t = (1 + 0.850B)(1 + 0.729B_12)a_t$$

Conclusion

In figure10, upon examining the residual plot of the ACF and the PACF for the SARIMA (1,1,1) (0,1,1)12 model, it was observed that the residuals exhibit white noise and that the P-values of the Ljung-Box statistic for the SARIMA (1,1,1) (0,1,1)12 model are 0.137, all of which are greater than or equal to 0.05. Thus, the model fits the data.

Given that the prediction of births is a crucial factor in any developmental plan, particularly in a developing country where female births are not tracked from their reproductive age to menopause, resulting in the inability to obtain estimates like the GRR and NRR, this female birth model, based on the available data, can be utilized in predicting the trend of female births, which in turn can be used to estimate the population of the state as of 2065, taking into account the average number of children that are expected to be born to a woman and the probability of her survival throughout her reproductive period.

Table 1: Female Birth

M/Y	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total	Mean
Jan	19	38	30	35	56	65	67	48	73	92	73	596	54.18
Feb	12	48	44	38	66	58	59	47	66	97	86	721	56.45
Mar	17	54	47	62	72	78	79	83	71	108	107	778	70.72
Apr	19	28	38	72	83	83	91	96	98	103	104	815	74.09
May	20	46	39	70	92	72	90	121	105	154	102	911	82.81
Jun	09	32	54	50	65	78	63	97	109	121	121	799	72.63
Jul	10	38	36	60	72	55	24	95	98	102	110	700	52.81
Aug	20	26	43	50	56	71	79	109	117	68	63	702	67.18
Sep	25	29	33	49	62	65	70	97	79	91	81	681	61.90
Oct	24	15	50	64	81	59	77	86	105	9	73	643	58.45
Nov	10	20	33	62	69	46	53	86	92	9	15	637	57.9
Dec	13	18	29	52	84	66	77	57	133	59	52	640	58.18
Total	198	392	476	664	848	796	829	1022	1148	1133	997	9456	
Mean	16,5	32.66	39.66	55.33	71.5	66.33	69.08	85.2	95.6	84.4	83.08		

Source: General Hospital Ukpom birth record

Table 2: The Yearly Mean of Births

S/N	Year	Yearly Total	Yearly Mean
1	2008	198	16.50
2	2009	392	32.66
3	2010	476	39.66
4	2011	664	55.33
5	2012	848	71.50
6	2013	796	66.33
7	2014	829	69.08
8	2015	1022	85.20
9	2016	1148	95.60
10	2017	1133	84.41
11	2018	997	83.08

Table 3: The Monthly Mean of Births

S/N	Month	Monthly Total	Month Mean
1	JAN	596	54.18
2	FEB	721	56.45
3	MAR	778	70.72
4	APR	815	74.09
5	MAY	911	82.81
6	JUN	799	72.63
7	JUL	700	52.81
8	AUG	702	67.18
9	SEP	681	61.90
10	OCT	643	58.45
11	NOV	637	57.90
12	DEC	640	58.18

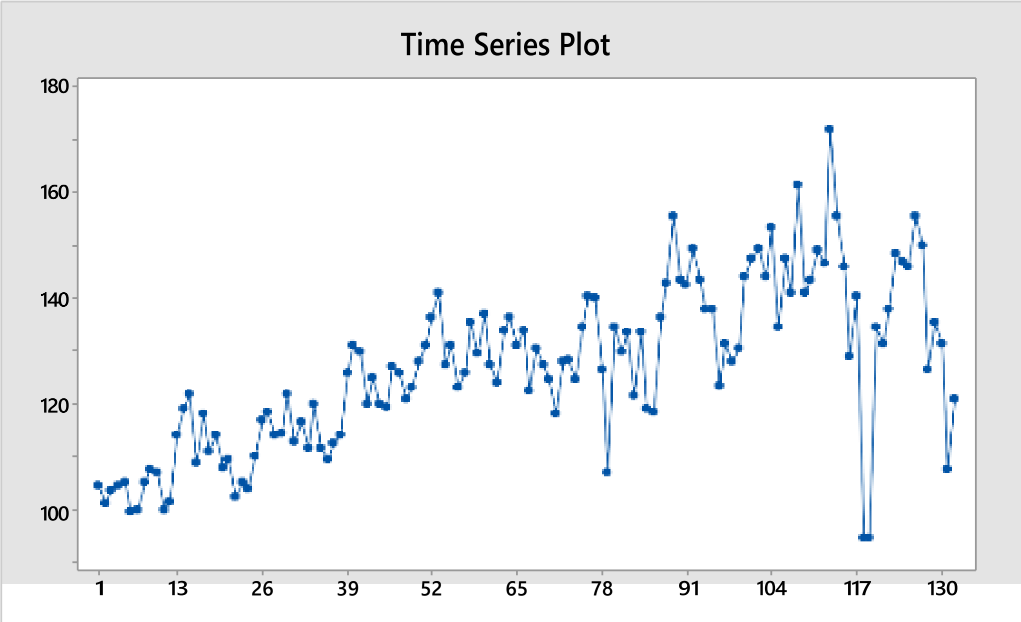


Fig.1: Time Plot of the Original Series for Female Births

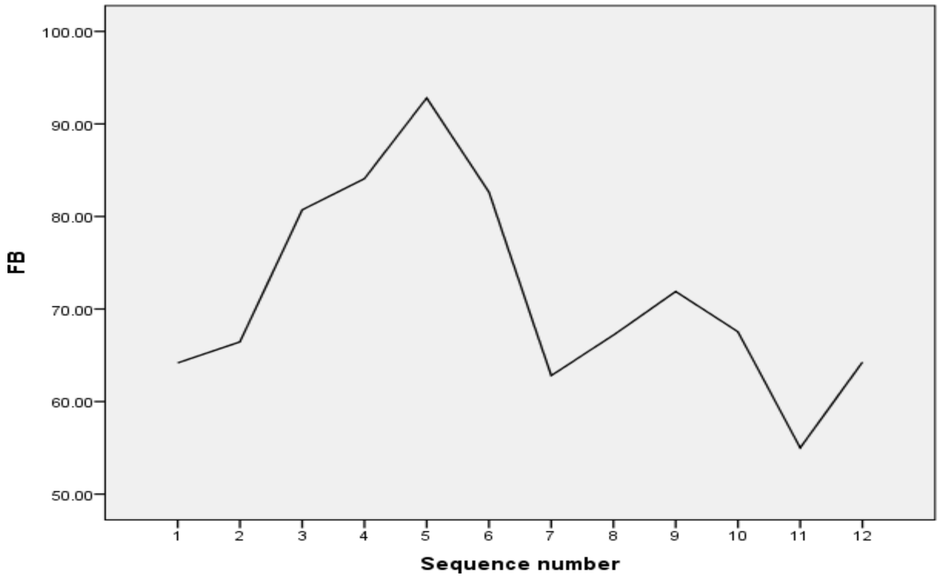


Fig. 2: Plot of Monthly Mean against Time of female births

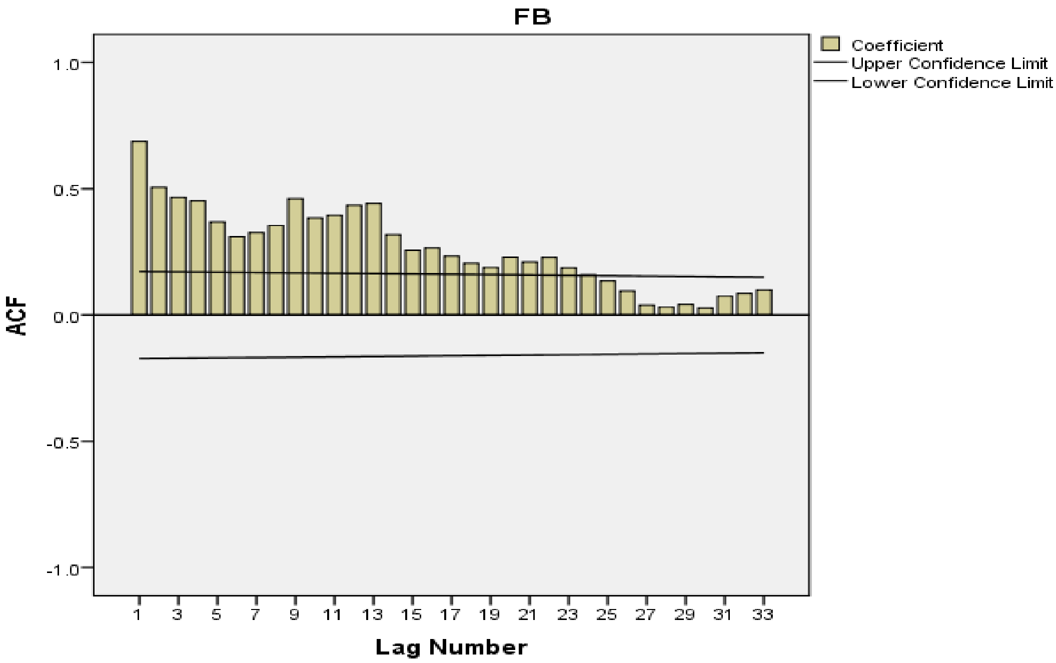


Fig. 3: Plot of ACF of the Original Series of female births

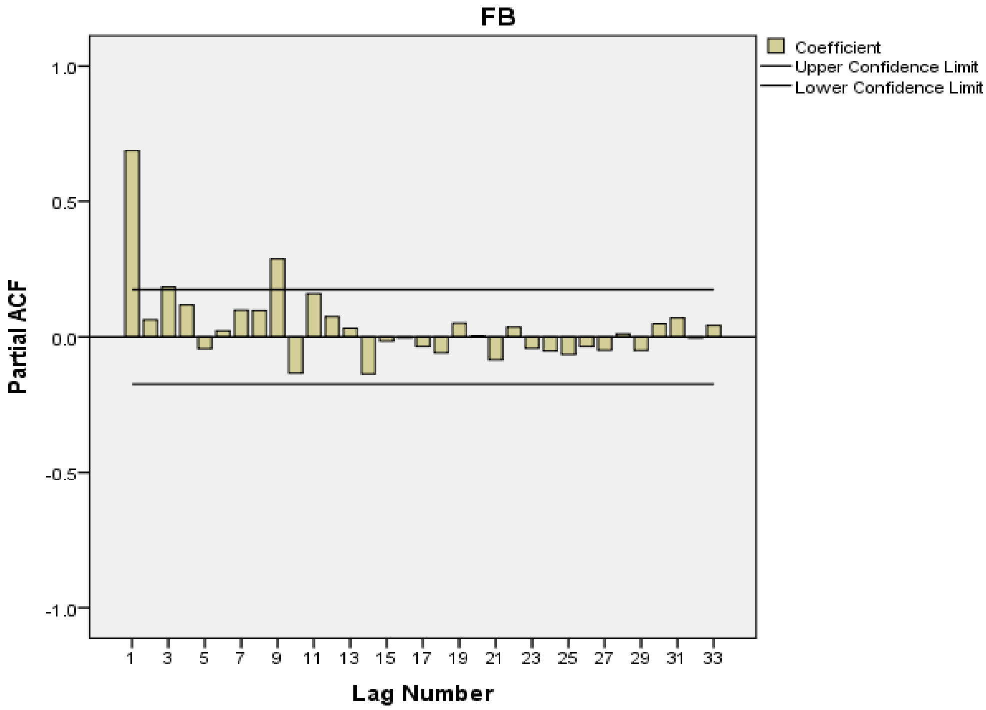


Fig.4: Plot of PACF of the Original Series of female births

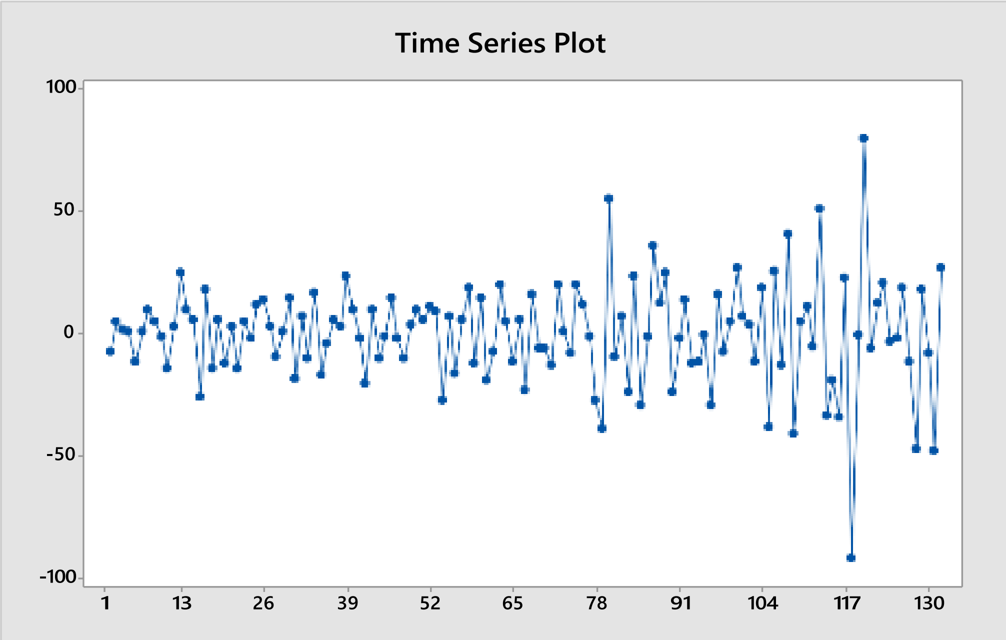


Fig.5: Time Plot of First Regular Difference of the Original Series of female births

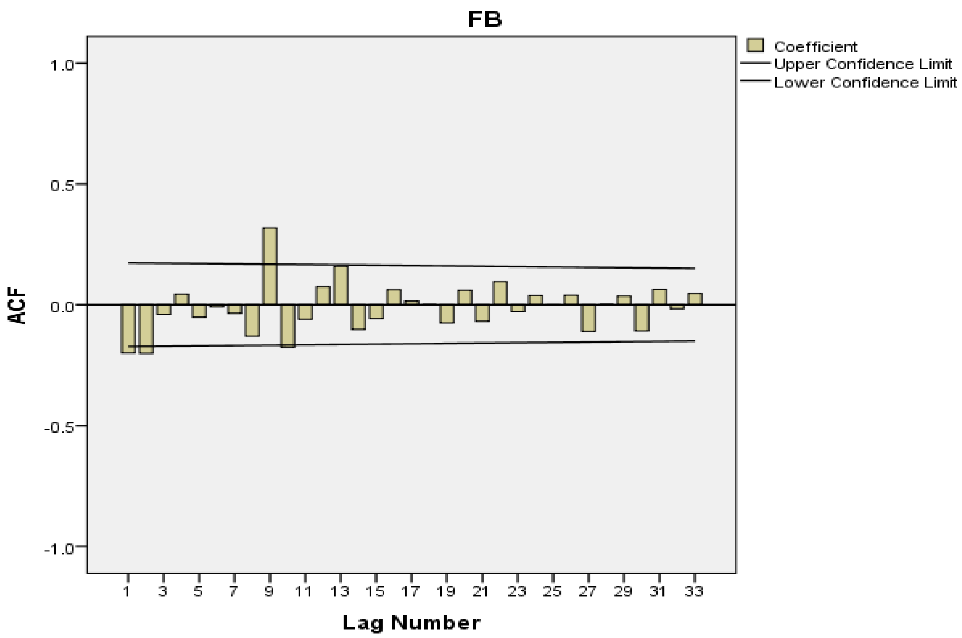


Fig.6: Plot of ACF of First Regular Difference of the Original Series of female births

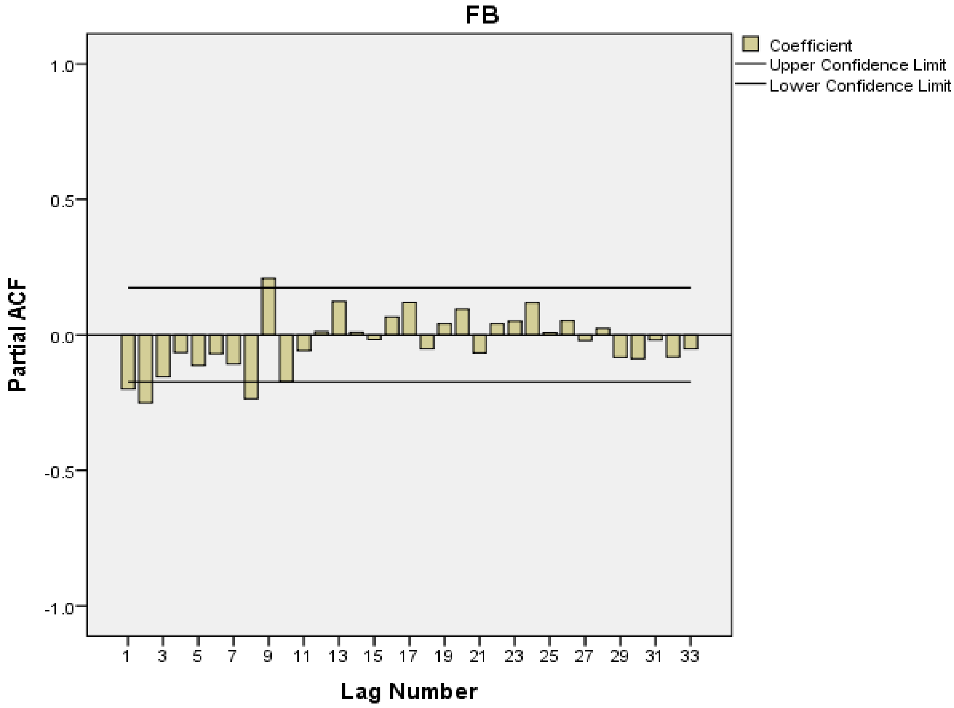


Fig.7: Plot of PACF of First Regular Difference of the Original Series of female births

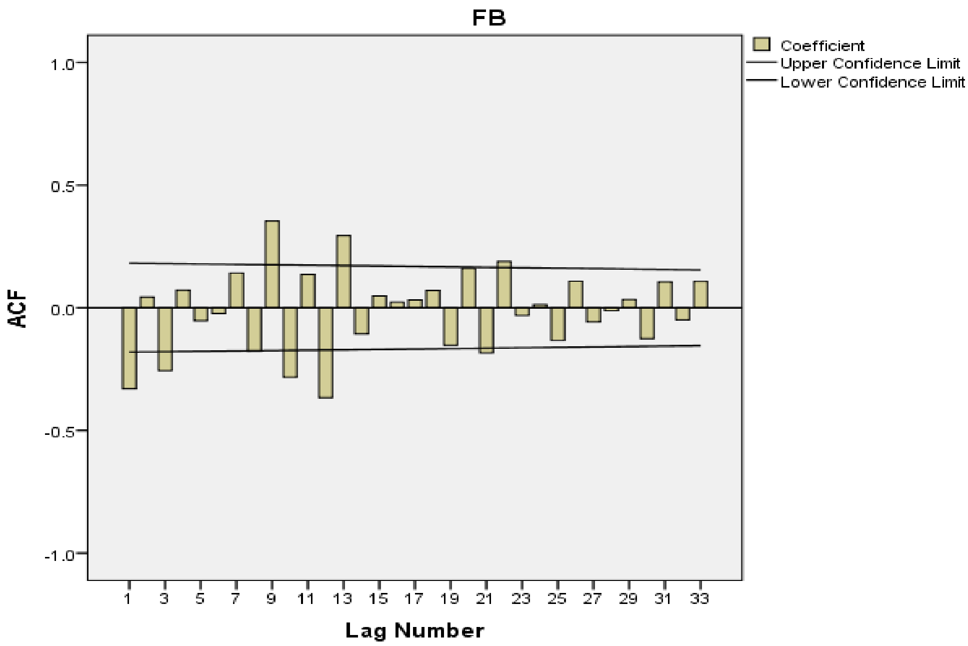
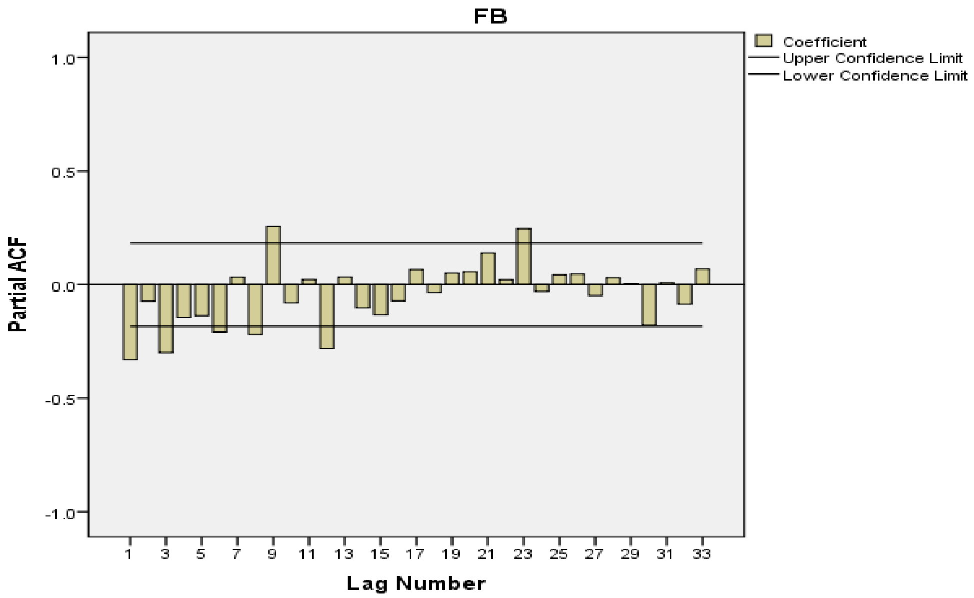


Fig.8: Plot of ACF of Seasonal Difference of the First Regular Difference of the Original Series of female births



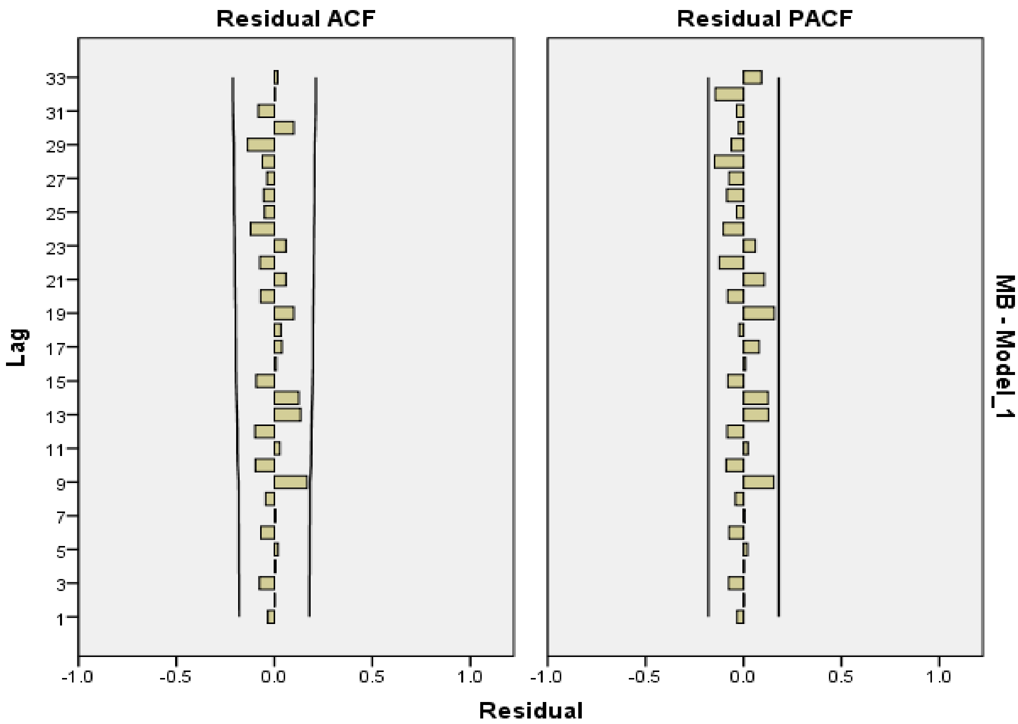


Fig. 9: Plot of PACF of Seasonal Difference of the First Regular Difference of theOriginal Series of female births

REFERENCES

Adams, S. O., Awujola, A. and Alumgudu A. I. (2014). Modeling Nigerian Consumer Price Index Using ARIMA. *International Journal of Development and Economic Sustainability*, 2(2), 37-47.

Akhter, T. (2013). Short-Term Forecasting of Inflation in Bangladesh with Seasonal ARIMA Process. *Munich Personal RepEc Archive Paper*. NO 43739 http://Mpra.Ub.Uni-Muenchen.de/43729.

Akpanta, A. C. and Okorie I. E. (2015). On the Time Series Analysis of Consumer Price Index data of Nigeria -1996 to 2013. *American Journal of Economics* 5(3): 363-369.

Box, G.E.P. and Jenkins, G.M. (1970). *Time Series Analysis, Forecasting and Control.*Holden Day, San Francisco

Jain, R.K., Sharma, R.D. and Jain, S. (1985). Application of ARIMA model in adjustment of seasonal and non-seasonal variations in births of Ontario. *Genus*, 41,127-133

Jondeau E., Le Biham H. and Sedillot F. (1999). Modeling and Forecasting the French Consumer Price Index Components (In French) Banque de France Working Paper No. 68, Available at *Social Science Research Network* (SSRN): *https://ssrn.com/abstract=1734313* or *http://dx.doi.org/10.2139/ssrn.1734313*

Lee, S. J, Steer P. J. and V. Filippi (2006). Seasonal patterns and preterm birth: a systematic review of the Literature and an analysis in a London based Cohort. *BOLOG: An international Journal of obstetrics and Gynecology*, 113(11), 1280- 1288. *https://doi.org/10.1111/j.1471-0528.2006.01055.x.*

Leung C. C. S. (1995). Time Series Modelling of Birth Data. A sub-thesis submitted forthe degree of Master of Statistics of the Australian National University.

McDonald, J. (1980). Births time series models and structural interpretations. *Journalof American Statistics Association*, 75, 39-41.

Olajide, J. T., Ayansola O. A., Odusina, M. T. and Oyenuga, I. F. (2012). Forecasting the Inflation Rate in Nigeria: Box Jenkins Approach. *IOSR Journal of Mathematics (IOSR-JM)*, 3(5):15-19. http://www.iosrjournals.org/iosr-jm/papers/Vol3- issue5/D0351519.pdf

Saboia, J.L.M. (1977). Autoregressive integrated moving average (ARIMA) models forbirth forecasting. *Journal of American Statistics Association*, 72, 264-270.

Wangdi K., Singhasivanum P., Silawan T., Lawpoilsri S.,White J. and Wkungwal J. (2010). 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 Journal* 9(1), 1-9. https://doi.org/10.1186/1475-2875-9-251.

Wilkie A.D., Watson R. and sons (2012). A Stochastic Model for Consumer Price Indices and Exchange Rates in Several Countries. Watson House.
Wikipedia (Nd). Birth Certificate. https://en.wikipedia.org ›wiki› Birth_certificate