Original Research Article

Time Series Modelling and Forecasting of Consumer Price Index in Ghana

Abstract

The knowledge of economic and financial indicators is the bedrock of making a thorough investment assessment and sound judgment of the various investment portfolios of an institution. Such important indicators include the consumer price index, which measures the change in the prices paid by households for goods and services consumed. A trigger in the consumer price in Ghana triggers inflation which affects the purchasing power of its citizens. Knowledge of the trend of the CPI is crucial in economic planning. The study sought to determine the appropriate time model for the CPI and use the model to predict the next nine months CPI. The study further sought to determine the type of trend model that characterizes the CPI. The Box-Jenkins methodology was adopted. The results of analysis showed SARIMA(2, 1, 1)(1, 0, 0)$_{12}$ as most fitted time series model and was used to predict the consumer price index for the next nine months. The S-model was also found to be the appropriate trend for the CPI. The SARIMA (2, 1, 1)(1, 0, 0)$_{12}$ is recommended for forecasting consumer price index in Ghana.

Keywords: Time series analysis, trend analysis and consumer price index.

1.0 Introduction

Economic and financial indicators are the bedrock of making a thorough investment assessment as well as making sound judgment concerning the various areas of investments [15]. Economic and financial indicators are also a source of ignition to the growth of every economy. Economic indicators can have a huge impact on the market. Therefore, knowing how to interpret and analyze them is crucial for all investors. They help us to understand the past and current economic situations so as to predict an expectation of the future for our investments.

One of the major economic indicators is the consumer price index (CPI). From the resolution concerning consumer price indices released in 2003 by the Seventeenth International Conference of Labour Statisticians convened by the International Labour Organization (ILO), the consumer price index (CPI) is a current social and economic indicator that is constructed to measure changes over time in the general level of prices of consumer goods and services that households acquire, use or pay for consumption. The index aims to measure the change in consumer prices over time. This may be done by measuring the cost of purchasing a fixed basket of consumer goods and services of constant quality and similar characteristics, with the products in the basket being selected to be representative of households' expenditure during a year or other specified period.

In Ghana, the consumption or shopping basket (CPI) consists of the following major commodity groups: food and non-alcoholic beverages, alcohol and tobacco, clothing and footwear, fuel & light, housing and utilities, household goods, operations and services, medical care and health expenses, transport and communications, recreation, entertainment, education and cultural services, and miscellaneous goods and services studied by Ghana Economy Watch [7].

Rising CPI also often leads to the central bank to raise interest rates, tightening money supply, reduce the money supply and other measures to tighten monetary policy, which flow into reduction of capital stock funds for greater returns, often accompanied by high inflation and therefore the stock market decline according to Mei [14]. Mei [14] revealed that CPI is one of the most important economic indicators in any country and it is used for measurement of the rate of inflation, for indexation of public pension benefits and for many policy purposes.

Inflation is directly calculated from CPI and it is considered to be a major problem in transition economies and thus fighting inflation and maintaining stable prices is the main objective of monetary authorities like Central Bank by Habimana et al. [8]. The consumer price index, CPI, which is probably the best known index number, is an economic indicator of the rate of inflation according to Gordor and Howard [10]. Changes in CPI are used to assess price changes associated with the cost of living. A rising rate means inflation, but economists usually predict an acceptable change of between 1% to 2% by Nketiah and Obeng-Aboagye [15]. Price change beyond 2% means there has been significant inflation and as a result, consumers' purchasing power decreases.

Investors need to know the trend of the CPI since the knowledge of it will help boost the earnings and the future prospects of their investments. Government, management of institutions and consumers equally need to know the projections of CPI to make useful decisions. This calls for a study to be conducted to formulate time series model for the CPI using data from January 1998 to December 2017 and then forecast the monthly and yearly CPI of Ghana using the resulted model. The study will also determine the trend of the CPI using data from March 2013 to November 2018. The regional CPI will also be compared using the analysis of variance to check if the differences in CPI among the regions are significant or not.

2.0 Methods

A quantitative research method was used to determine the trend of the CPI and specifically, the Box-Jenkins model methodology was adopted in formulating the appropriate model for the CPI. Minitab 16 software was used in analyzing the data.

2.1 Trend Analysis

Four trend models were considered. The trend models considered for the CPI were linear, quadratic, exponential growth model and S-curve model.

2.1.1 Linear trend model

The linear trend model is given by $y = b_0 + b_1 t + \varepsilon_t$, where b_0 is the value of y when x is zero or the y intercept and b_1 is the change in y for a unit change x.

2.1.2 Quadratic trend model

The quadratic trend model is given by $y = b_0 + b_1 t + b_2 t^2 + \varepsilon_t$. The quadratic trend model accounts for simple curvature in the data.

2.1.3 Exponential growth

The exponential growth model is given by $y = b_0 b_1^t \varepsilon_t$. Exponential growth trend model accounts for exponential growth or decay.

2.1.4 S-curve model

The S-model is given by $y = \dfrac{10^3}{b_0 + b_1 b_2^t}$. The S-curve model fits the Pearl-Reed logistic trend model. It accounts for the case where the series follows an S-shaped curve.

In the fitted trend equation, the letters represent the following:

- y_t is the variable
- b_0 is the constant
- b_1 and b_2 are the coefficients
- t is the value of the time unit

2.2 Seasonal Autoregressive integrated moving average models (SARIMA*)*

Box and Jenkins [4] generalized the autoregressive integrated moving average (ARIMA) model known as Seasonal ARIMA (SARIMA) to deal with seasonality. Seasonal ARIMA (SARIMA) is used when the time series exhibits a seasonal variation. According to Permanasari et al. [16], a seasonal autoregressive notation (*P*) and a seasonal moving average notation (*Q*) will form the multiplicative process of SARIMA as $(p, d, q)(P, D, Q)^s$. The subscripted letter 's' shows the length of seasonal period. For example, in a monthly data $s = 12$. The SARIMA as $(p, d, q)(P, D, Q)^s$ in terms of lag polynomials is given below [13]:

$$\Phi_P\left(L^s\right)\varphi_p(L)(1-L)^d(1-L^s)^D y_t = \Theta_Q(L^s)\theta_q(L)\varepsilon_t$$

2.2.1 *Data*

In this study, past monthly national consumer price indices from March 2013 to November 2018 of Ghana ("Monthly Statistical bulletin of Ghana"). A time series model would be formulated for each period. Based on this data, time series forecasting will be conducted to forecast CPI up to May 2019.

2.2.2 Box Jenkins Methodology

Box and Jenkins developed a systematic methodology for identifying and estimating models that could incorporate both approaches and this makes Box-Jenkins models a powerful class of models by Dobre et al. [6]. The Box-Jenkins methodology includes model data preparation, model selection, parameter estimation, model diagnosis and forecasting.

2.2.3 Data Preparation

Data preparation involves transformations and differencing. The first step in data preparation is to check whether the series is stationary or not. If the series is non-stationary, transforming the series by using square root, logarithm, and so on can help stabilize the variance. The data can also be differenced to achieve stationarity. The Augmented-Dickey Fuller (ADF) test is performed to confirm the stationarity or otherwise of the series. If the calculated value is greater than the t-value at α level of significance, then the CPI has a unit root problem. Or for a smaller p-value, the null hypothesis is rejected at α level of significance, meaning that the CPI has no unit root problem.

2.2.4. Model selection

Model selection is based on the use of graphs based on the transformed or differenced data to identify potentially ARIMA processes. The major tools used in the identification phase are plots of the series, correlograms of auto correlation (ACF), and partial autocorrelation (PACF). Arriving at the right model is not always straightforward but also a good deal of experimentation with alternative models. Among the competing models, the one with the least Bayesian information criterion (BIC) is selected. Minitab does not display BIC value but rather the mean square error (MS) of the residuals. The mean square error is also a measure of the accuracy of the fitted model. Smaller values of the mean square error usually indicate a better fitting model.

2.2.5 Parameter estimation

Parameter estimation means finding the values of the model coefficient that gives the best fit to the data. Minitab version 16.0 is used to estimate the values of the coefficients. To determine whether the association between the response variable (CPI) and each term in the model is statistically significant, compare the p-value for the term to the significance level to assess the null hypothesis. The null hypothesis is that the term is not significantly different from 0, which indicates that no association exists between the term and the response. A significance level of α =0.01is used. If p-value ≤ α: The term is statistically significant. If the p-value is less than or equal to the significance level, you can conclude that the coefficient is statistically significant.

But if the p-value > α: the term is not statistically significant. If the p-value is greater than the significance level, you cannot conclude that the coefficient is statistically significant. You may want to refit the model without the term.

2.2.6 Model diagnosis

Diagnosis of the model is a complementary step to selecting the best model from class of competing models by Alonso and Garcia-Martos [2]. Model diagnosis involves testing to find out if there is no violation of the assumptions of the developed model. If there is violation in any of the assumptions, then it calls for selecting a new model. The adequacy of the model is checked by analyzing the residuals. If the residuals are white noise we accept the model, else we go to the first step again and start over. The ACF and PACF together with modified Box-Pierce would be used to determine the validity of the model. For a well fitted model, the residuals should be uncorrelated (white noise).

2.3 Forecasting with the ARIMA and SARIMA Models

The developed model should be able to forecasting the time series with minimal forecasting errors. According to Guerrero, 2003 a model fitting is defined as the sum of the residuals squares divided by the sample size. Its object is to measure the model's capacity to reproduce the sample data (i.e. to verify how similar the modeled series and the actual series are).

3.0 Results and Discussion

The data for consumer price indices from March 2013 to November 2018 from Ghana Statistics Service (GSS) was used. The data was input into Minitab 16 to generate the required statistics, models and charts.

3.1 Descriptive Statistic

The table below is the descriptive statistic of the CPI of Ghana.

Table 1: Descriptive Statistics of CPI

Variable	N	Mean	SE Mean	StDev	Minimum	Maximum	Q1	Median	Q3	Range
CPI	69	165.5	4.48	37.24	108.00	224.20	130.60	168.00	200.00	116.20

From table 1 above, the mean of the sixty-nine (69) data point was 165.5 and the standard error of the mean (SE Mean) was 4.48. The standard deviation (StDev) was 37.24 with minimum and maximum values as 108.00 and 224.20 respectively. The range of the CPI was 116.20. The table further revealed that 25% of the CPI were below 130.60 and above 200.00, 50% were below and above 168.00.

3.2 Trend of CPI

The four trend models which had already been discussed were linear, quadratic, growth curve model and S-curve. Table 2 below had the four trend models as well as their coefficients and measures of accuracies.

Table 2: Trend models

	Linear model	Quadratic model	Exponential model	S-model
b_0	100.505	101.20	107.638	2.857
b_1	1.851	1.793	1.012	6.916
b_2	*	8.35×10^{-4}	*	0.978
MAPE	1.486	1.454	2.011	1.274
MAD	2.252	2.227	3.522	1.976
MSD	7.402	7.314	21.329	5.748

From table 2 above, among the competing trend models, the S-model seemed appropriate as it had the lowest measures of accuracy. The maximum absolute percentage error (MAPE) was 1.274, the mean absolute deviation (MAD) was 1.976 and the mean square deviation (MSD) was also 5.748.

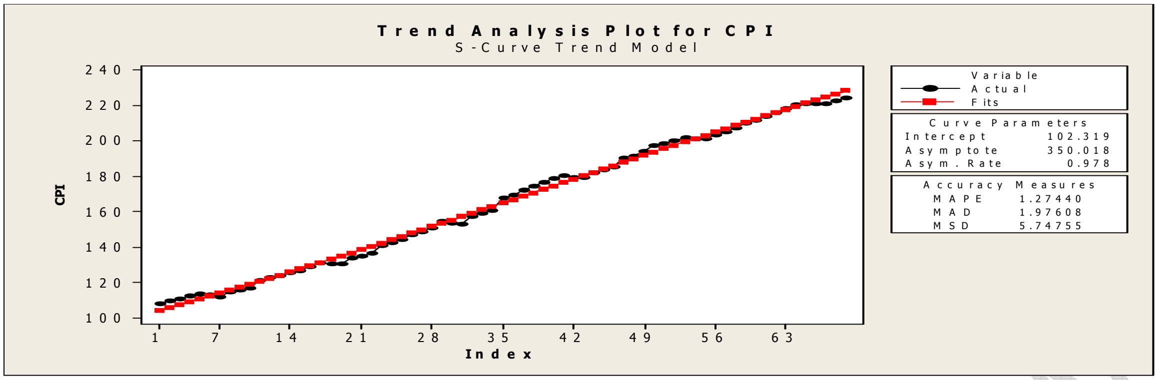

Figure 1: The Trend analysis plot for CPI

Figure 1 is the trend analysis plot for CPI. The plot portrayed the s-model with intercept 102.319, asymptote 350.018 and asymptotic rate 0.978. The trend equation established using the coefficients of the s-model in table 2 above was

$$y = \frac{10^3}{2.857 + 6.916(0.978)^t}$$

3.3 Test for Stationarity

The graph in figure 1 above, clearly indicates the CPI has no constant mean and variance meaning that it is non-stationary. To test for the presence of an intercept and a trend, the ADF-test is used. The result of the ADF-test is shown in the table 3 below.

Table 3: Result of ADF-test on the CPI

	t-statistics	P-value
ADF-test statistics	-3.057	0.125
Test critical values: 1%	-4.108	
5%	-3.482	
10%	-3.169	

From table 3 above, the ADF-test statistic value is -3.057. The test critical values at 1%, 5% and 10% are -4.108-3.482 and -3.169 respectively. Since the test statistic value (-3.057) is greater than the critical values at 1%, 5% and 10% respectively, the test confirms the presence of a trend and thus, the CPI is non-stationary. So the CPI needs to be differenced.

Table 4: Result of ADF-test on differenced CPI

	t-statistics	P-value
ADF-test statistics	-4.382	0.004
Test critical values: 1%	-4.116	
5%	-3.485	
10%	-3.171	

After differencing the CPI, a further test indicates that the CPI has no unit root problem and the result is shown in table 4 above.

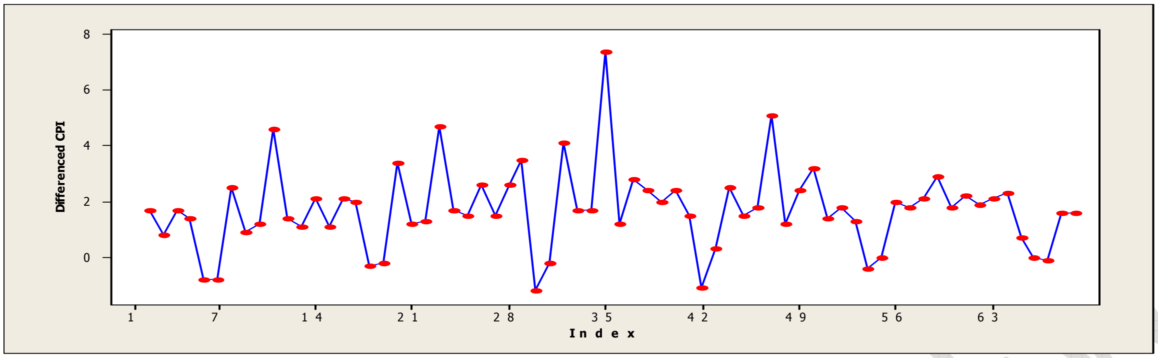

Figure 2: Plot of the differenced CPI

The graph in figure 2 above is the plot of the differenced CPI when appears to have constant mean and variance as confirmed by the ADF-test above.

3.3 ARIMA/SARIMA Modeling

The Minitab software was used to formulate three tentative SARIMA models. The results are shown in table 5 below.

Table 5: Tentative SARIMA models

SARIMA Structure	***Type***	***Coefficients of parameters***	***P-value***	***MSE***
$(1,1,1)(1,0,0)_{12}$	AR 1	0.980	0.000	0.549
	SAR 12	0.922	0.000	
	MA 1	0.898	0.000	
$(2,1,1)(1,0,0)_{12}$	AR 1	0.6507	0.001	0.546
	AR 2	0.2582	0.054	
	SAR 12	0.9715	0.000	
	MA	0.7584	0.000	
$(1,1,1)(1,0,1)_{12}$	AR 1	0.9810	0.000	0.552
	SAR 12	0.8902	0.000	
	MA 1	0.8999	0.000	
	SMA 12	0.0041	0.980	

From table 5 above, the three tentative SARIMA models formulated were $(1,1,1)(1,0,0)_{12}$, $(2,1,1)(1,0,0)_{12}$ and $(1,1,1)(1,0,1)_{12}$. The model with the least mean square error is $(2,1,1)(1,0,0)_{12}$ and was deemed appropriate for forecasting the CPI of Ghana for the next eight (8) months. The coefficients of the SARIMA $(2,1,1)(1,0,0)_{12}$ are significantly different from zero at 1% level of significance.

Table 6: Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	16.2	31.2	42.9	44.4
DF	8	20	32	44
P-Value	0.040	0.053	0.094	0.454

Table 6 showed the results of the Ljung-Box statistics. The Ljung-Box statistics at lags 12, 24, 36 and 48 were 16.2 31.2, 42.9 and 44.4 respectively. The Ljung-Box statistics give non-significant p-values indicating that the residuals appeared to be uncorrelated at 1% level of significance.

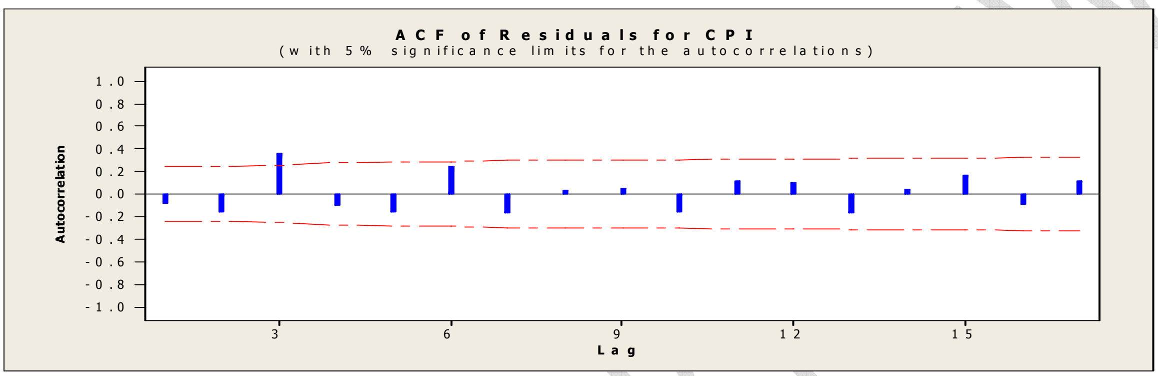

Figure 3: ACF of residuals for CPI

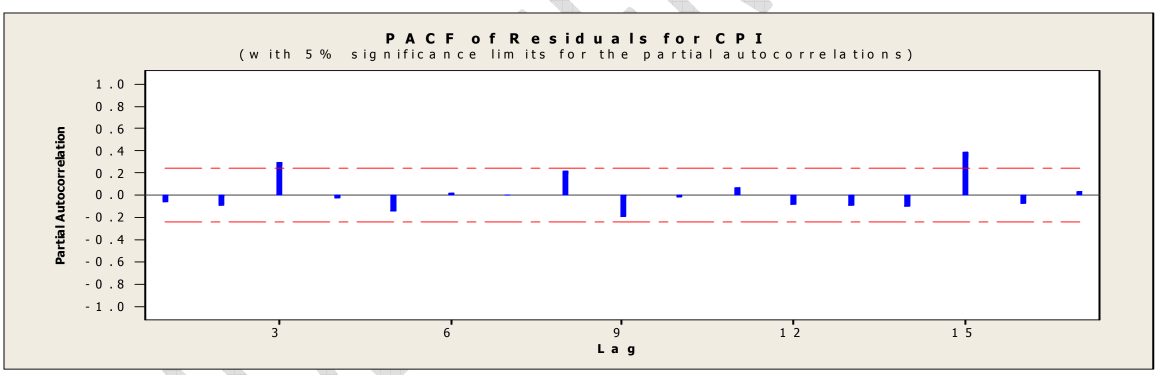

Figure 4: PACF of residuals for CPI

The residual autocorrelations and partial autocorrelations supported the view that the residuals are uncorrelated at 1% and the spikes at lag 3 of the ACF in figure 3 and lags 3 and lags 15 in figure 4 are as a result of randomness in the CPI.

3.4 Forecasting with SARIMA Model

ARIMA models are basically developed to forecast the corresponding variable. There are two types of forecasts: sample period forecasts and post sample period forecasts. The former is used to develop confidence interval in the model and the latter to generate genuine forecasts for planning and other purposes.

3.4.1 In-sampling Forecast

The sample period forecasts were obtained by simply plugging the actual values of the explanatory variables in the formulated model $(2,1,1)(1,0,0)_{12}$. The forecast values together with the lower confidence limits (LCL) and the upper confidence limits (UCL) constructed at 95% had been displayed in table 7 below.

Table 7: In-sampling forecast of the CPI

Period	Forecast	LCL	UCL	Actual
18-Mar	213.825	212.376	215.274	214.100
18-Apr	216.833	214.891	218.775	216.000
18-May	218.023	215.528	220.518	218.100
18-Jun	219.635	216.614	222.655	220.400
18-Jul	220.765	217.216	224.313	221.100
18-Aug	220.254	216.180	224.328	221.100
18-Sep	220.140	215.542	224.739	221.000
18-Oct	221.978	216.856	227.100	222.600
18-Nov	223.629	217.986	229.272	224.200

From table 7 above, the forecast errors which are the differences between the actual and the forecast values were very minimal, that is, the predicted values are very close to the actual values and this indicates the reliability of the model.

3.4.2 Post Sample Forecast

The main objective of developing a SARIMA model for a variable is to generate post sample period forecast.

Table 8: Forecasts of the CPI

Period	*Forecast*	*LCL*	*UCL*
18-Dec	226.184	224.734	227.633
19-Jan	228.926	226.984	230.868
19-Feb	230.611	228.116	233.105
19-Mar	232.687	229.667	235.708
19-Apr	234.477	230.929	238.025
19-May	236.465	232.391	240.539
19-Jun	238.651	234.052	243.250
19-Jul	239.286	234.164	244.408
19-Aug	239.245	233.602	244.888

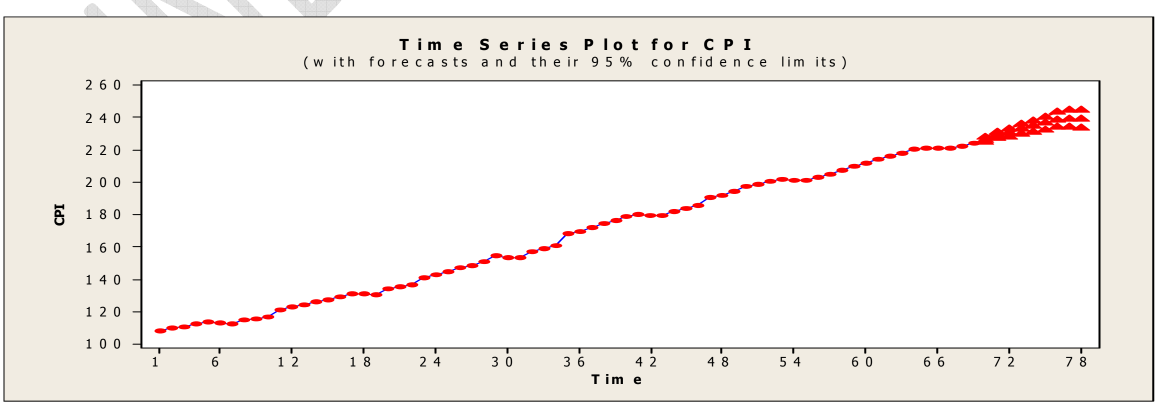

Figure 5: Plot of CP with forecast and their 95% confidence limits

From table 8above, the forecast of the CPI of Ghana for the next nine months from December 2018 to August 2019 provided there are no policy interventions or shocks in the economy. The forecast values are increasing in magnitude. Figure 5 above is the plot of the CPI with the forecasts and their 95% confidence limits.

3.5 Conclusion

The most appropriate SARIMA model for forecasting the CPI using data from March 2013 to November 2018 is the SARIMA $(2,1,1)(1,0,0)_{12}$.This model has been shown to adequately explain the variation in the monthly CPI. The model is also used to forecast the CPI for the next nine months. The trend of the CPI has an S-model.

Reference

[1] Adams, O. S, Awujola, A and Alumgudu, I.A, (2014).Modeling Nigeria's Consumer Price Index Using Arima Model. International Journal of Development and Economic Sustainability, Vol.2, No. 2, pp. 37-47, June 2014.

[2] Alonso, M. A. and Garcia-Martos, C. (2012). Time Series Analysis: Model diagnosis and prediction. U*C3M-UPM*

[3] Anim-Ansah, A.V, (2010).Time Series Analysis of the Consumer Price Index in Ghana (1964 – 2009). Published thesis

[4] Box, G. E. P., Jenkins, G. M., and Reinsel, G. C, (1994). Time series analysis: Forecasting and control, 3rd ed. Englewood Cliffs, N.J.: Prentice Hall.

[5] Dielman, T. (2006).Choosing Smoothing Parameters for Exponential Smoothing: Minimizing Sums Of Squared Versus Sums Of Absolute Errors. Journal of Modern Applied Statistical Methods, May, 2006, Vol. 5, No. 1, 118-129

[6] Dobre, I., Alexandru, A. A., (2008). Quantitative Methods Inquires, Journal of Applied Quantitative Methods vol. 40, pp.165-166.

[7] Ghana Economy Watch, Vo.l 6 January-June 2010.

[8] Habimana, N, Wanjoya, A and Waititu, A, (2016). Modeling and Forecasting Consumer Price Index (Case of Rwanda).American Journal of Theoretical and Applied Statistics2016; 5(3): 101-107, Science Publishing Group.

[9] Ghana Statistical Service (2016). Consumer Price Index August 2016. Statistical Newsletter, No. B12-2003, Accra.

[10] Gordor,K. B, Howard, K. N, (2000). Elements of Statistical Analysis. City Printers, Accra- Ghana.

[11] Guerrero, V. M., (2003). Análisisestadístico de series de tiempoeconómicas, Segundaedición, México.

[12] ILO (2003). The Seventeenth International Conference of Labour Statisticians: Resolution concerning consumer price indices. Geneva.

[13] Ljung, G. M. and Box, G. E. P. (1978). A Measure of Lack of Fit in Time Series Models.Biometrika.65, 297-3030.

[14] Mei, T. W. L (2011).The Relationship between Consumer Price Index (CPI) and Producer Price Index (PPI) in Malaysia. Faculty of Economics and Business University of Malaysia Sarawak

[15] Nketiah, K. N. and Obeng-Aboagye, S. (2016). Economic and Financial Indicators - A Tool for Investment Decisions.Omega capital, Accra.

[16] Permanasari, E. A., Hidaya, I. and Bustoni, A. I. (2013).*SARIMA (*Seasonal ARIMA) implementation on time series to forecast the number of Malaria incidence. https://www.researchgate.net/publication/261307350. Accessed on 30th April 2018.

[17] TIBCO Software Inc. (2018). How to identify patterns in time series data: Time series analysis. http://www.statsoft.com/Textbook/Time-Series-Analysis**.** Accessed on 25th March 2018.

[18] Zhang X., Zhang T., Young A. A., Li, X. (2014). Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance *Data.*PLoS ONE 9(2): e88075. doi:10.1371/journal.pone.0088075

APPENDIX

Table 1:DATA USED IN THE ANALYSIS

Month	CPI	Month	CPI
13-Mar	108.0	16-May	176.4
13-Apr	109.7	16-Jun	178.8
13-May	110.5	16-Jul	180.3
13-Jun	112.2	16-Aug	179.2
13-Jul	113.6	16-Sep	179.5
13-Aug	112.8	16-Oct	182.0
13-Sep	112.0	16-Nov	183.5
13-Oct	114.5	16-Dec	185.3
13-Nov	115.4	17-Jan	190.4
13-Dec	116.6	17-Feb	191.6
14-Jan	121.2	17-Mar	194.0
14-Feb	122.6	17-Apr	197.2
14-Mar	123.7	17-May	198.6
14-Apr	125.8	17-Jun	200.4
14-May	126.9	17-Jul	201.7
14-Jun	129.0	17-Aug	201.3
14-Jul	131.0	17-Sep	201.3
14-Aug	130.7	17-Oct	203.3
14-Sep	130.5	17-Nov	205.1
14-Oct	133.9	17-Dec	207.2
14-Nov	135.1	18-Jan	210.1
14-Dec	136.4	18-Feb	211.9

15-Jan	141.1	18-Mar	214.1
15-Feb	142.8	18-Apr	216
15-Mar	144.3	18-May	218.1
15-Apr	146.9	18-Jun	220.4
15-May	148.4	18-Jul	221.1
15-Jun	151.0	18-Aug	221.1
15-Jul	154.5	18-Sep	221
15-Aug	153.3	18-Oct	222.6
15-Sep	153.1	18-Nov	224.2
15-Oct	157.2		
15-Nov	158.9		
15-Dec	160.6		
16-Jan	168.0		
16-Feb	169.2		
16-Mar	172.0		
16-Apr	174.4		