Journal of Statistical and Econometric Methods, vol.1, no.3, 2012, 27-37
ISSN: 2241-0384 (print), 2241-0376 (online)
Scienpress Ltd, 2012

Predicting Inflation Rates Of Nigeria Using A Seasonal Box-Jenkins Model

Ette Harrison Etuk[1]

Abstract

Time series analysis of Nigerian monthly Inflation Rates (INFL) Data is done. It is observed that it is seasonal. Based on its autocorrelation structure as depicted by the correlogram, the multiplicative seasonal autoregressive integrated moving average (ARIMA) model, $(1, 1, 0)x(0, 1, 1)_{12}$, is fitted to the series. The model is shown to be adequate and the 2012 forecasts are obtained on the basis of it. These forecasts are shown to agree closely with the observations.

Keywords: Inflation Rate, Seasonal Time Series, ARIMA model, Nigeria

1 Introduction

A time series has the property that neighbouring values are correlated. This tendency is called *autocorrelation*. It is said to be stationary if it has a constant mean, constant variance and autocorrelation that is a function of the lag separating the correlated values. The autocorrelation expressed as a function of the lag is called *the autocorrelation function* (ACF).

A stationary time series $\{X_t\}$ is said to follow an *autoregressive moving average model of orders p and q* (denoted by ARMA(p,q)) if it satisfies the following difference equation

---

[1] Department of Mathematics/Computer Science, Rivers State University of Science and Technology, Nigeria, e-mail: ettetuk@yahoo.com

Article Info: *Received* : September 21, 2012. *Revised* : October 24, 2012
Published online : November 30, 2012

$$X_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + ... + \alpha_p X_{t-p} = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + ... + \beta_q \varepsilon_{t-q} \quad (1)$$

or

$$A\ (L)X_t = B\ (L)\varepsilon_t \quad (2)$$

where $\{\varepsilon_t\}$ is a sequence of uncorrelated random variables with zero mean and constant variance, called *a white noise process*, and the α_i's and β_j's constants;

$A\ (L) = 1 + \alpha_1 L + \alpha_2 L^2 + ... + \alpha_p L^p$

and

$B\ (L) = 1 + \beta_1 L + \beta_2 L^2 + ... + \beta_q L^q$

and L is the backward shift operator defined by $L^k X_t = X_{t-k}$.

If p=0, model (1) becomes *a moving average model of order q* (denoted by MA(q)). If, however, q=0 it becomes *an autoregressive process of order p* (denoted by AR(p)). An AR(p) model of order p may be defined as a model for which a current value of the time series X_t depends on the immediate past p values: $X_{t-1}, X_{t-2}, ..., X_{t-p}$. On the other hand an MA(q) model of order q is whereby the current value X_t is a linear combination of the immediate past q values of the white noise process: $\varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q}$. Besides stationarity, invertibility is another important requirement for a time series. It refers to the characteristic whereby the covariance structure of the series is unique (Priestley, 1981). In addition it makes meaningful association of current events with the past history of the series possible (Box and Jenkins, 1976).

An AR(p) model can be put more specifically:

$$X_t + \alpha_{p1} X_{t-1} + \alpha_{p2} X_{t-2} + ... + \alpha_{pp} X_{t-p} = \varepsilon_t$$

Then the sequence of the last coefficients$\{\alpha_{ii}\}$ is called *the partial autocorrelation function*(PACF) of $\{X_t\}$. The ACF of an MA(q) model cuts off after lag q whereas that of an AR(p) model is a mixture of sinusoidals tailing off slowly. On the other hand the PACF of an MA(q) model tails off slowly whereas that of an AR(p) model tails off after lag p.

AR and MA models are known to have some duality characteristics. These include:

1. A finite order of the one type is equivalent to an infinite order of the other type.
2. The ACF of the one type exhibits the same behaviour as the PACF of the other type.
3. An AR model is always invertible but is stationary if A (L) = 0 has zeros outside the unit circle.
4. An MA model is always stationary but is invertible if B (L) = 0 has zeros outside the unit circle.

Parametric parsimony consideration in model building entails preference for the mixed ARMA fit to either the pure AR or the pure MA fit. Stationarity and invertibility conditions for model (1) or (2) are that the equations A (L) = 0 and B (L) = 0 should have roots outside the unit circle respectively.

Often, in practice, a time series is non-stationary. Box and Jenkins (1976) proposed that differencing of an appropriate order could render a non-stationary series $\{X_t\}$ stationary. Suppose the degree of differencing necessary for stationarity is equal to d. Such a series $\{X_t\}$ may be modelled as

$$((1+\sum_{i=1}^{p} a_i B^i)\nabla^d X_t = B(L)\varepsilon_t \qquad (3)$$

where $\nabla = 1 - L$ and in which case $A(L) = (1+\sum_{i=1}^{p} a_i B^i)\nabla^d = 0$ shall have unit roots d times. Then differencing to degree d renders the series stationary. The model (3) is said to be an autoregressive integrated moving average model of orders p, d and q and denoted by ARIMA(p, d, q).

1.1 Seasonal ARIMA Models

A time series is said to be seasonal of order d if it has a tendency to exhibit periodic behaviour after every time interval d. Traditional time series analysis involves the identification, disintegration and estimation of the traditional components: secular trend, seasonal component, cyclical component and the irregular movement. For the purpose of making forecasts, they are reintegrated. Such techniques could be quite misleading.

The time series $\{X_t\}$ is said to follow a multiplicative (p, d, q)x(P, D, Q)$_s$ *seasonal ARIMA model* if

$$A(L)\Phi(L^s)\nabla^d\nabla_s^D X_t = B(L)\Theta(L^s)\varepsilon_t \qquad (4)$$

where Φ and Θ are polynomials of order P and Q respectively. That is,

$$\Phi(L^s) = 1 + \phi_1 L^s + \ldots + \phi_P L^{sP}, \qquad (5)$$

$$\Theta(L^s) = 1 + \theta_1 L^s + \ldots + \theta_Q L^{sQ}, \qquad (6)$$

where the ϕ_i and θ_j are constants such that the zeros of the equations (5) and (6) are all outside the unit circle for stationarity and invertibility respectively. $\Phi(L^s)$ represents the seasonal autoregressive operator whereas $\Theta(L^s)$ represents the seasonal moving average operator.

A seasonal nature is often apparent from the time plot. Moreover for a seasonal series the ACF or correlogram exhibits a spike at the seasonal lag. Box and Jenkins (1976) and Madsen (2008) are a few authors that have written extensively on such models. A knowledge of the theoretical properties of the models provides basis for their identification and estimation. The purpose of this paper is to fit a seasonal ARIMA model to Nigerian Inflation Rate (INFL) series. Earlier work on Nigerian Inflation rates includes those by Abidemi and Maliq (2010), Olatunji et al (2010) and Chiakwelu (2012). Whereas Abidemi and Maliq (2010) think that the inflation rates are stationary, Olatunji et al (2010) think otherwise. The graph of the inflation rates from 1990 to 2010 given by Chiakwelu (2012) corroborates the opinion of Olatunji et al. There is perhaps no known attempt to model Nigerian inflation rates by a seasonal ARIMA model.

2 Materials and Methods

The data for this work are inflation rates – All items (Year on Change)- from 2003 to2011 obtainable from the Data and Statistics publication of Central Bank of Nigeria retrievable from the website http://www.cenbank.org/.

2.1 Determination of the orders d, D, p, P, q and Q

Seasonal differencing is necessary to remove the seasonal trend. If there is secular trend non-seasonal differencing will be necessary. To avoid unnecessary model complexity it has been advised that orders of differencing d and D should add up to at most 2 (i.e. d + D < 3). If the ACF of the differenced series has a positive spike at the seasonal lag then a seasonal AR component is suggestive; if it has a negative spike then a seasonal MA term is suggestive.

As already mentioned above, an AR(p) model has a PACF that truncates at lag p and an MA(q)) has an ACF that truncates at lag q. In practice $\pm\frac{2}{\sqrt{n}}$, where n is the sample size are the non-significance limits for both functions.

2.2 Model Estimation

The involvement of the white noise terms in an ARIMA model necessitates a nonlinear iterative process in the model estimation. An optimization criterion like the least squares, maximum likelihood or maximum entropy is used. An initial estimate is usually used and each iteration is expected to be an improvement of the previous one until the estimate converges to an optimal one. However, for pure AR and pure MA models linear optimization techniques exist (See for example Box and Jenkins (1976), Oyetunji (1985)). There are attempts to propose linear methods to estimate ARMA models (See for example, Etuk (1987,1998)). We shall use Eviews software which employs the least squares approach to analyze the data.

2.3 Diagnostic Checking

The model that is fitted to the data should be tested for goodness-of-fit. To do this we shall do some residual analysis. In particular we shall plot the histogram of the residuals and their correlogram. If the model is correct, the residuals would be uncorrelated and would follow a normal distribution with mean zero and constant variance. Assuming an adequate model, the autocorrelations of the residuals should therefore not be significantly different from zero.

3 Results and Discussion

The time plot of the original series INFL in Figure1 shows no clear secular trend nor seasonality. Seasonal (i.e. 12-month) differencing of the series produces a series SDINFL also with no trend nor clear seasonality (see Figure 2). Non-seasonal differencing yields a series DSDINFL with no trend and no clear seasonality (see Figure 3). Its ACF in Figure 4 has spikes at lags 1 and 12 revealing a seasonality of lag 12. The spike at lag 12 is negative indicating the involvement of a seasonal MA component of order one. The PACF has a spike at lag 1 indicating a nonseasonal AR component of order one.

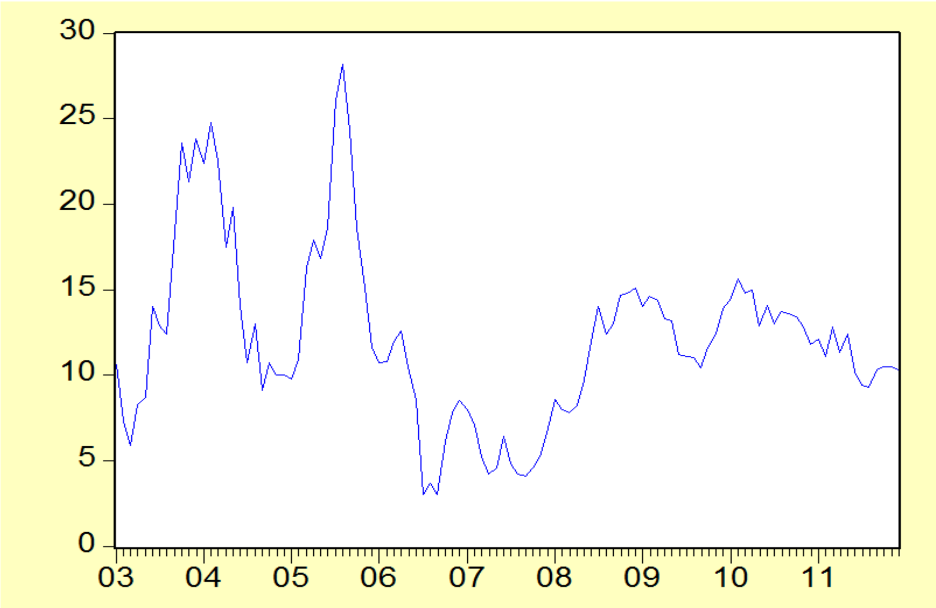


Figure 1: INFL

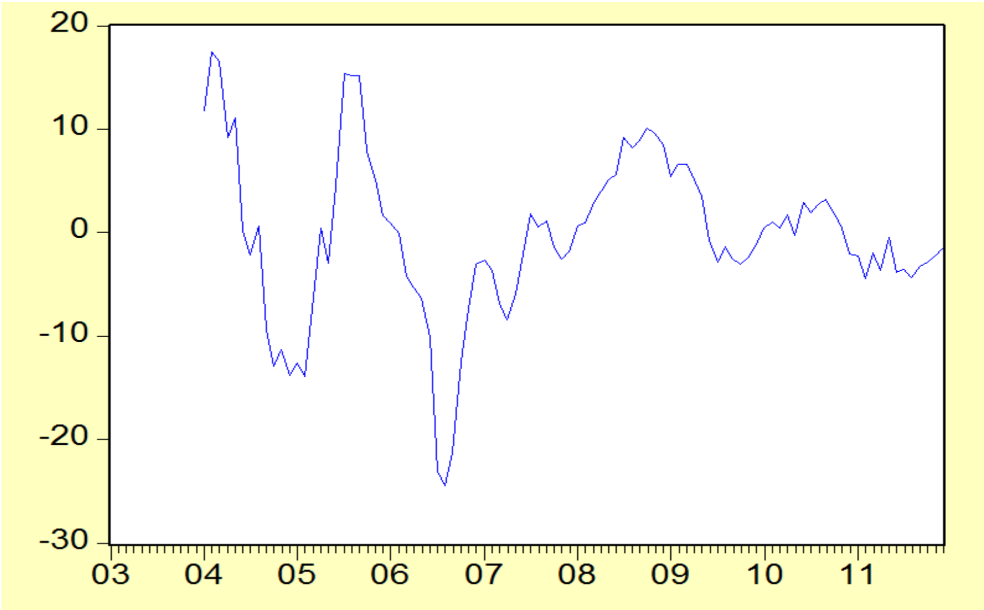


Figure 2: SDINFL

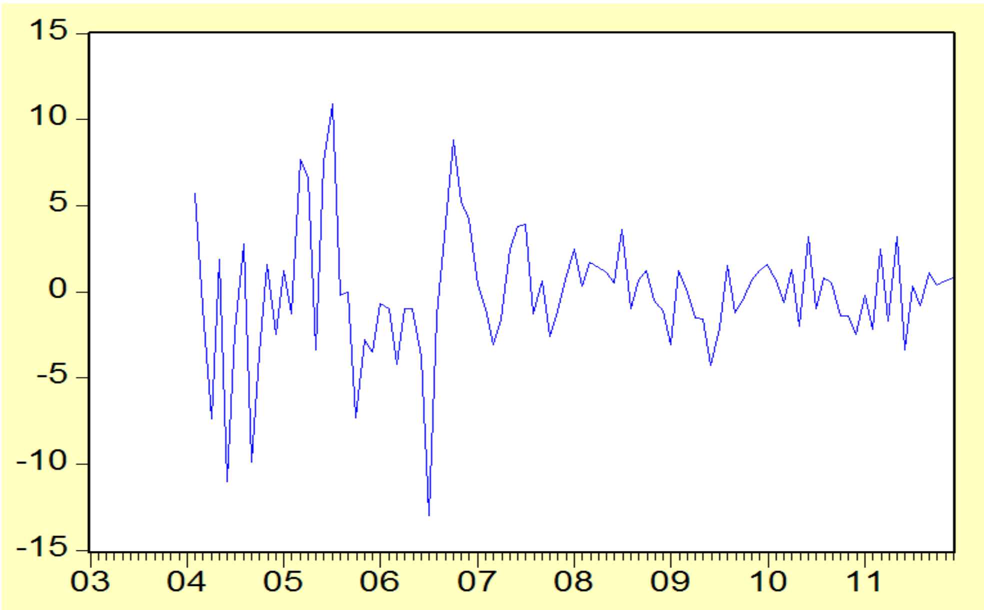
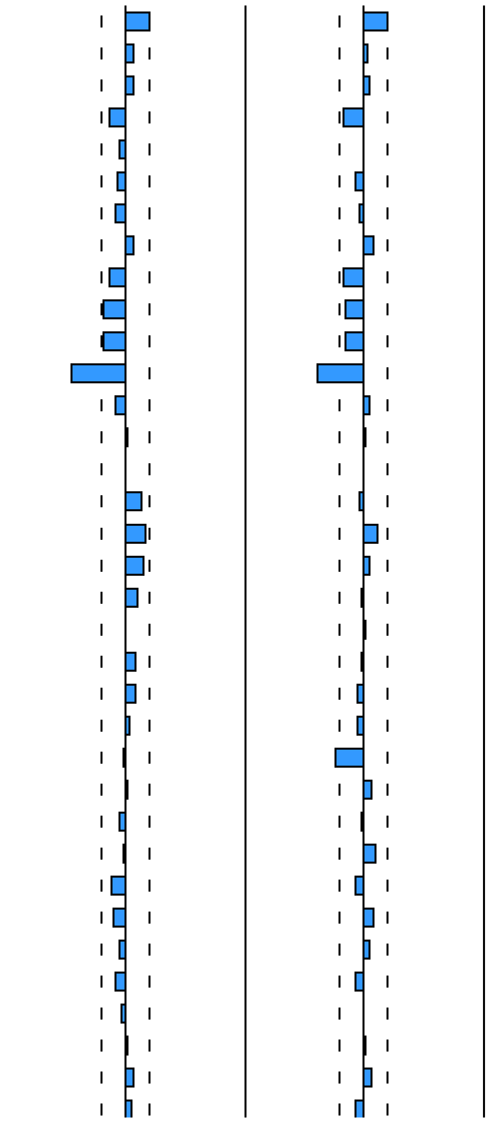


Figure 3: DSDINFL



	AC	PAC	Q-Stat	Prob
1	0.212	0.212	4.4230	0.035
2	0.081	0.037	5.0671	0.079
3	0.079	0.057	5.6852	0.128
4	-0.132	-0.171	7.4598	0.113
5	-0.050	0.005	7.7181	0.172
6	-0.080	-0.064	8.3772	0.212
7	-0.088	-0.034	9.1877	0.239
8	0.068	0.088	9.6801	0.288
9	-0.139	-0.176	11.759	0.227
10	-0.185	-0.152	15.464	0.116
11	-0.181	-0.155	19.064	0.060
12	-0.461	-0.406	42.662	0.000
13	-0.089	0.046	43.549	0.000
14	0.012	0.008	43.566	0.000
15	0.007	-0.007	43.571	0.000
16	0.137	-0.039	45.757	0.000
17	0.178	0.120	49.496	0.000
18	0.157	0.056	52.464	0.000
19	0.100	-0.015	53.677	0.000
20	0.007	0.020	53.684	0.000
21	0.091	-0.021	54.707	0.000
22	0.088	-0.042	55.683	0.000
23	0.038	-0.045	55.866	0.000
24	-0.015	-0.242	55.894	0.000
25	0.015	0.063	55.923	0.000
26	-0.046	-0.011	56.203	0.001
27	-0.012	0.106	56.224	0.001
28	-0.125	-0.077	58.377	0.001
29	-0.113	0.085	60.145	0.001
30	-0.052	0.051	60.522	0.001
31	-0.086	-0.059	61.587	0.001
32	-0.041	-0.006	61.830	0.001
33	0.011	0.009	61.847	0.002
34	0.072	0.065	62.626	0.002
35	0.052	-0.064	63.042	0.003

Figure 4: Correlogram of DSDINFL

We hereby propose the following $(1, 1, 0)x(0, 1, 1)_{12}$ model

$$DSDINFL_t = \alpha_1 DSDINFL_{t-1} + \beta_{12}\varepsilon_{t-12} + \varepsilon_t \quad (7)$$

The estimation of the model is summarized in Table 1. The model is given by

$$DSDINFL_t = \underset{(\pm 0.0672)}{0.2023} DSDINFL_{t-1} - \underset{(\pm 0.0255)}{0.8858}\varepsilon_{t-12} + \varepsilon_t \quad (8)$$

The estimation involved 20 iterations. Both coefficients are significantly different from zero, each being more than twice its standard error. As much as 65% of the variation in DSDINFL is accounted for by the model. The histogram of the residuals in Figure 5 shows that the residuals are normally distributed with zero mean indicating model adequacy; the mean of -0.16 is not significantly different from zero given a standard deviation of 2.1 by t-test. Moreover the correlogram of the residuals in Figure 6 depicts the adequacy of the model. Virtually all the residual autocorrelations are not significantly different from zero.

Table 1: Model Estimation

Dependent Variable: DSDINFL
Method: Least Squares
Date: 02/07/12 Time: 21:15
Sample(adjusted): 2004:03 2011:12
Included observations: 94 after adjusting endpoints
Convergence achieved after 20 iterations
Backcast: 2003:03 2004:02

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	0.202287	0.067199	3.010279	0.0034
MA(12)	-0.885827	0.025458	-34.79528	0.0000

R-squared	0.645616	Mean dependent var	-0.202128
Adjusted R-squared	0.641764	S.D. dependent var	3.620015
S.E. of regression	2.166682	Akaike info criterion	4.405318
Sum squared resid	431.8951	Schwarz criterion	4.459431
Log likelihood	-205.0500	F-statistic	167.6052
Durbin-Watson stat	1.888246	Prob(F-statistic)	0.000000

Inverted AR Roots	.20			
Inverted MA Roots	.99	.86+.49i	.86 -.49i	.49+.86i
	.49 -.86i	-.00 -.99i	-.00+.99i	-.49 -.86i
	-.49+.86i	-.86+.49i	-.86 -.49i	-.99

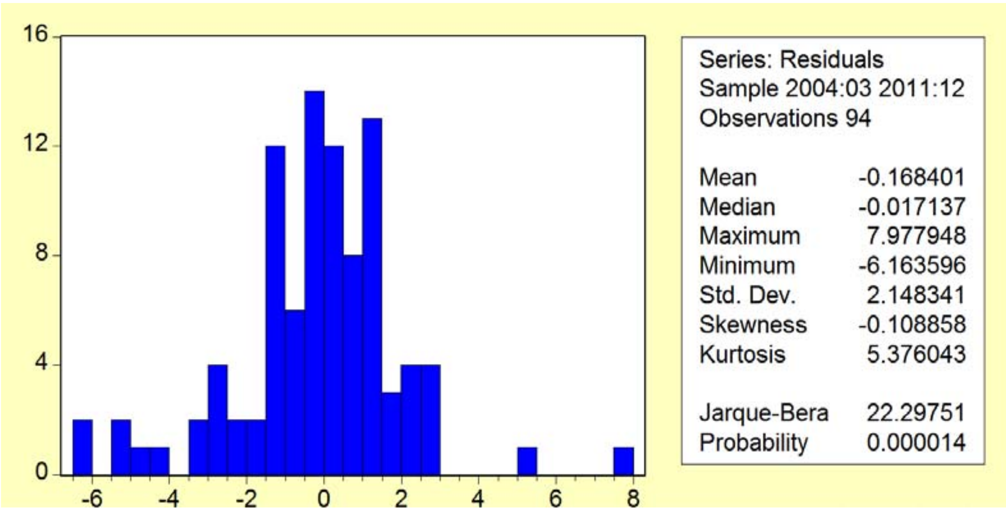


Figure 5: Histogram of residuals

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.038	0.038	0.1417	
		2	0.016	0.014	0.1666	
		3	-0.029	-0.030	0.2505	0.617
		4	-0.059	-0.057	0.5942	0.743
		5	-0.003	0.002	0.5955	0.897
		6	-0.089	-0.088	1.3988	0.844
		7	0.032	0.036	1.5055	0.912
		8	0.055	0.053	1.8253	0.935
		9	-0.031	-0.042	1.9298	0.964
		10	-0.123	-0.132	3.5513	0.895
		11	-0.088	-0.073	4.3921	0.884
		12	-0.241	-0.245	10.794	0.374
		13	-0.035	-0.032	10.927	0.449
		14	-0.019	-0.029	10.969	0.532
		15	0.060	0.029	11.377	0.579
		16	-0.044	-0.110	11.598	0.639
		17	0.095	0.096	12.654	0.629
		18	0.092	0.060	13.667	0.624
		19	0.055	0.071	14.033	0.665
		20	0.053	0.056	14.378	0.704
		21	0.002	0.000	14.378	0.761
		22	0.074	0.002	15.065	0.773
		23	0.002	-0.017	15.066	0.820
		24	-0.076	-0.136	15.816	0.825
		25	0.009	-0.007	15.826	0.863
		26	-0.045	-0.078	16.100	0.884
		27	0.083	0.121	17.039	0.880
		28	-0.063	-0.087	17.573	0.891
		29	-0.004	0.093	17.575	0.916
		30	0.032	0.072	17.722	0.933
		31	-0.132	-0.057	20.207	0.886
		32	-0.028	-0.005	20.320	0.908
		33	-0.030	0.012	20.453	0.926
		34	0.095	0.079	21.811	0.912

Figure 6: Correlogram of the residuals

4 Forecasting

For the model (7) at time t+k we have

$$X_{t+k} = \alpha_1 X_{t+k-1} + \beta_{12}\varepsilon_{t+-12} + \varepsilon_{t+k}$$

Obtaining conditional expectations given the series up to time t, we have

$$\hat{X}_t(1) = \alpha_1 X_t + \beta_{12}\varepsilon_{t-11}$$

$$\hat{X}_t(k) = \alpha_1 \hat{X}_t(k-1) + \beta_{12}\varepsilon_{t+k-12}\text{, k} \geq 2$$

where $\hat{X}_t(k)$ is the k-step ahead forecast from time t.

The 2012 forecasts are given in Table 2. With the actual observed values for the first seven months of the year being 12.6, 11.9, 12.1, 12.9, 12.7, 12.9 and 12.8 respectively, the chi-square goodness-of-fit test statistic is equal to 4.1132 which with a p-value of greater than 0.5 is not statistically significant. This further shows that the model is adequate.

TIME	RESIDUALS	DSDINFL	SDINFL	INFL
December 2010	-1.26140	-2.5	-2.1	11.8
January 2011	0.3899	-0.2	-2.3	12.1
February 2011	-1.34231	-2.2	-4.5	11.1
March 2011	1.70871	2.5	-2.0	12.8
April 2011	-1.39008	-1.7	-3.7	11.3
May 2011	1.30454	3.2	-0.5	12.4
June 2011	-2.39750	-3.4	-3.9	10.2
July 2011	0.10659	0.3	-3.6	9.4
August 2011	-0.29539	-0.8	-4.4	9.3
September 2011	1.21624	1.1	-3.3	10.3
October 2011	-0.81674	0.4	-2.9	10.5
November 2011	0.46909	0.6	-2.3	10.5
December 2011	-0.43875	0.8	-1.5	10.3
January 2012		-0.40	-1.90	10.2
February 2012		1.11	-0.79	10.3
March 2012		-1.29	-2.08	10.7
April 2012		0.97	-1.11	10.2
May 2012		-0.96	-2.07	10.3
June 2012		1.93	-0.14	10.1
July 2012		0.30	0.16	9.6
August 2012		0.32	0.48	9.8
September 2012		-1.01	-0.53	9.8
October 2012		0.52	-0.01	10.5
November 2012		-0.31	-0.32	10.2
December 2012		0.33	0.01	10.3

5 Conclusion

The INFL series has been shown to follow a $(1, 1, 0)\text{x}(0, 1, 1)_{12}$ model. This model has been shown to be adequate. On the basis of the model 2012 forecasts have been obtained. These forecasts have been shown to agree very closely with the observed values of the year.

References

[1] M.B. Priestley, *Spectral Analysis and Time Series,* Academic Press, London, 1981.

[2] G.E.P. Box and G.M. Jenkins, *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, 1976.

[3] H. Madsen, *Time Series Analysis*, Chapman and Hall/CRC, London, 2008.

[4] O.I. Abidemi and A.A.A, Maliq, Analysis of Inflation and its determinants in Nigeria, *Pakistan Journal of Social Sciences,* **7**(2), (2010), 97-100.

[5] G.B. Olatunji, O.A. Omotesho, O.E. Ayinde and K. Ayinde, Determinants of inflation in Nigeria: A co-integration approach, Contributed paper presented at the Joint 3rd African Association of Agriculture Eonomists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa, September 19 – 23, 2010. http://ageconsearch.umn.edu/bitstream/96162/2/103.

[6] E. Chiakwelu, Nigeria: Inflation spikes, Naira recoils and T-bills sell continues, *Africa Political and Economic Strategic Center*, AFRIPOL; http://www.afripol.org, (2012).

[7] O.B. Oyetunji, Inverse Autocorrelations and Moving Average Time Series Modelling, *Journal of Official Statistics*, **1**, (1985), 315-322.

[8] E.H. Etuk, *On the Selection of Autoregressive Moving Average Models*, An unpublished Ph. D. Thesis, Department of Statistics, University of Ibadan, Nigeria, 1987.

[9] E.H. Etuk, An Autoregressive Integrated Moving Average (ARIMA) Simulation Model: A Case Study, *Discovery and Innovation*, **10**(1-2), (1998), 23-26.