

# A MULTIPLICATIVE SEASONAL ARIMA MODEL FOR NIGERIAN UNEMPLOYMENT RATES

Ette Harrison Etuk

Department of Mathematics/Computer Science,  
Rivers State University of Science and Technology,  
NIGERIA

Tel: +2348136948161

**Keywords:** Unemployment Rates, Seasonal Time Series, ARIMA model, Nigeria

**Abstract:** Time series analysis of Nigerian Unemployment Rates is done. The data used is monthly from 1948 to 2008. The time plot reveals a slightly positive trend with no clear seasonality. A multiplicative seasonal model is suggestive given seasonality that typically tends to increase with time. Seasonal differencing once produced a series with no trend nor discernible stationarity. A non-seasonal differencing of the seasonal differences yielded a series with no trend but with a correlogram revealing stationarity of order 12, a nonseasonal autoregressive component of order 3 and a seasonal moving average component of order 1. A multiplicative seasonal autoregressive integrated moving average (ARIMA) model,  $(3, 1, 0) \times (0, 1, 1)_{12}$ , is fitted to the series. It has been shown to be adequate.

## 1. Introduction

A time series is defined as a set of data collected sequentially in time. It has the property that neighbouring values are correlated. This tendency is called autocorrelation. A time series is said to be stationary if it has a constant mean and variance. Moreover the autocorrelation is a function of the lag separating the correlated values and called the autocorrelation function (ACF).

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

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

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  $\alpha_i$ 's and  $\beta_j$ 's constants;

$A(L) = 1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p$  and  $B(L) = 1 + \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q$  and  $L$  the backward shift

operator defined by  $L^k X_t = X_{t-k}$ .

If  $p = 0$ , the model (1) becomes a *moving average model of order  $q$*  (designated MA( $q$ )). If, however,  $q = 0$  it becomes an *autoregressive process of order  $p$*  (designated AR( $p$ )). An AR( $p$ ) model may be defined as a model whereby a current value of the time series  $X_t$  depends on the immediate past  $p$  values:  $X_{t-1}, X_{t-2}, \dots, X_{t-p}$ . On the other hand an MA( $q$ ) model is such that the current value  $X_t$  is a linear combination of immediate past values of the white noise process:

$\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ . Apart from stationarity, invertibility is another important requirement for a time series. It refers to the property whereby the covariance structure of the series is unique (Priestley, 1981). Moreover it allows for meaningful association of current events with past history of the series (Box and Jenkins, 1976).

An AR(p) model may be more specifically written as

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

Then the sequence of the last coefficients  $\{\alpha_{ii}\}$  is called the partial autocorrelation function of  $\{X_t\}$ . The ACF of an MA(q) model cuts off after lag q whereas that of an AR(p) model is a combination of sinusoidals dying off slowly. On the other hand the PACF of an MA(q) model dies off slowly whereas that of an AR(p) model cuts off after lag p. AR and MA models are known to exhibit some duality relationships. These include:

1. A finite order AR model is equivalent to an infinite order MA model.
  2. A finite order MA model is equivalent to an infinite order AR model.
  3. The ACF of an AR model exhibits the same behaviour as the PACF of an MA model.
  4. The PACF of an AR model exhibits the same behaviour as the ACF of an MA model.
  5. An AR model is always invertible but is stationary if  $A(L) = 0$  has zeros outside the unit circle.
  6. 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 appropriate order could render a non-stationary series  $\{X_t\}$  stationary. Let degree of differencing necessary for stationarity be d. Such a series  $\{X_t\}$  may be modelled as

$$A(L)\nabla^d X_t = B(L)\epsilon_t \tag{3}$$

where  $\nabla = 1 - L$  and in which case  $A(L) \nabla^d = 0$  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 designated ARIMA(p, d, q).

## 2. Seasonal ARIMA Models

A time series is said to be seasonal of order d if there exists a tendency for the series to exhibit periodic behaviour after every time interval d. Traditional time series methods involve the identification, unscrambling and estimation of the traditional components: secular trend, seasonal component, cyclical component and the irregular movement. For forecasting purpose, 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

$$A(L)\Phi(L^s)\nabla^d \nabla_s^D X_t = B(L)\Theta(B^s)\epsilon_t \tag{4}$$

where  $\Phi$  and  $\Theta$  are polynomials of order P and Q respectively. That is,

$$\Phi(L^s) = 1 + \phi_1 L^s + \dots + \phi_P L^{sP} \tag{5}$$

$$\Theta(L^s) = 1 + \theta_1 L^s + \dots + \theta_Q L^{sQ} \tag{6}$$

where the  $N_i$  and the  $\theta_j$  are constants such that the zeros of the equations (5) and (6) are all outside the unit circle for stationarity or invertibility respectively. Equation (5) represents the autoregressive operator whereas (6) represents the moving average operator.

Existence of a seasonal nature is often evident 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 the Nigerian Unemployment Rates (NUMP).

### 3. Materials and Methods

The data for this work are Nigerian monthly Unemployment rates Prices from 1948 to 2008 obtainable from the quarterly abstracts of the Central Bank of Nigeria.

#### 3.1. Determination of the orders $d$ , $D$ , $P$ , $q$ and $Q$ :

Seasonal differencing is necessary to remove the seasonal trend. If there is secular trend non-seasonal differencing will be necessary. To avoid undue 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  $\pm 2/\sqrt{n}$  where  $n$  is the sample size are the non-significance limits for both functions.

#### 3.2. Model Estimation

The involvement of the white noise process in an ARIMA model entails a nonlinear iterative process in the estimation of the parameters. An optimization criterion like least error sum of squares, maximum likelihood or maximum entropy is used. An initial estimate is usually used. Each iteration is expected to be an improvement of the last 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 adopt linear methods to estimate ARMA models (See for example, Etuk(1987, 1998)). We shall use Eviews software which employs the least squares approach involving nonlinear iterative techniques.

#### 3.3. Diagnostic Checking

The model that is fitted to the data should be tested for goodness-of-fit. We shall do some analysis of the residuals of the model. If the model is correct, the residuals would be uncorrelated and would follow a normal distribution with mean zero and constant variance. The autocorrelations of the residuals should not be significantly different from zero.

### 4. Results and Discussion

The time plot of the original series NUMP in Figure 1 shows a slightly positive trend and no clear seasonality. However the seasonality tends to increase with time indicating a multiplicative seasonal model. Seasonal (i.e. 12-month) differencing of the series produces a series SDNUMP with a slightly positive trend but with no clear seasonality (see Figure 2). Non-seasonal differencing of SDNUMP yields a series DSDNUMP with no trend and no clear seasonality (See Figure 3). However, its ACF in Figure 4 has a negative spike at lag 12 revealing a seasonality of lag 12 and a seasonal MA component of order one to the model. The PACF has significant spikes at the early lags up to lag 3 suggesting the involvement of a nonseasonal AR component of order 3. We therefore propose the  $(3, 1, 0) \times (0, 1, 1)_{12}$  model

$$DSDNUMP_t = \alpha_1 DSDNUMP_{t-1} + \alpha_2 DSDNUMP_{t-2} + \alpha_3 DSDNUMP_{t-3} + \beta_{12} \varepsilon_{t-12} + \varepsilon_t \quad (7)$$

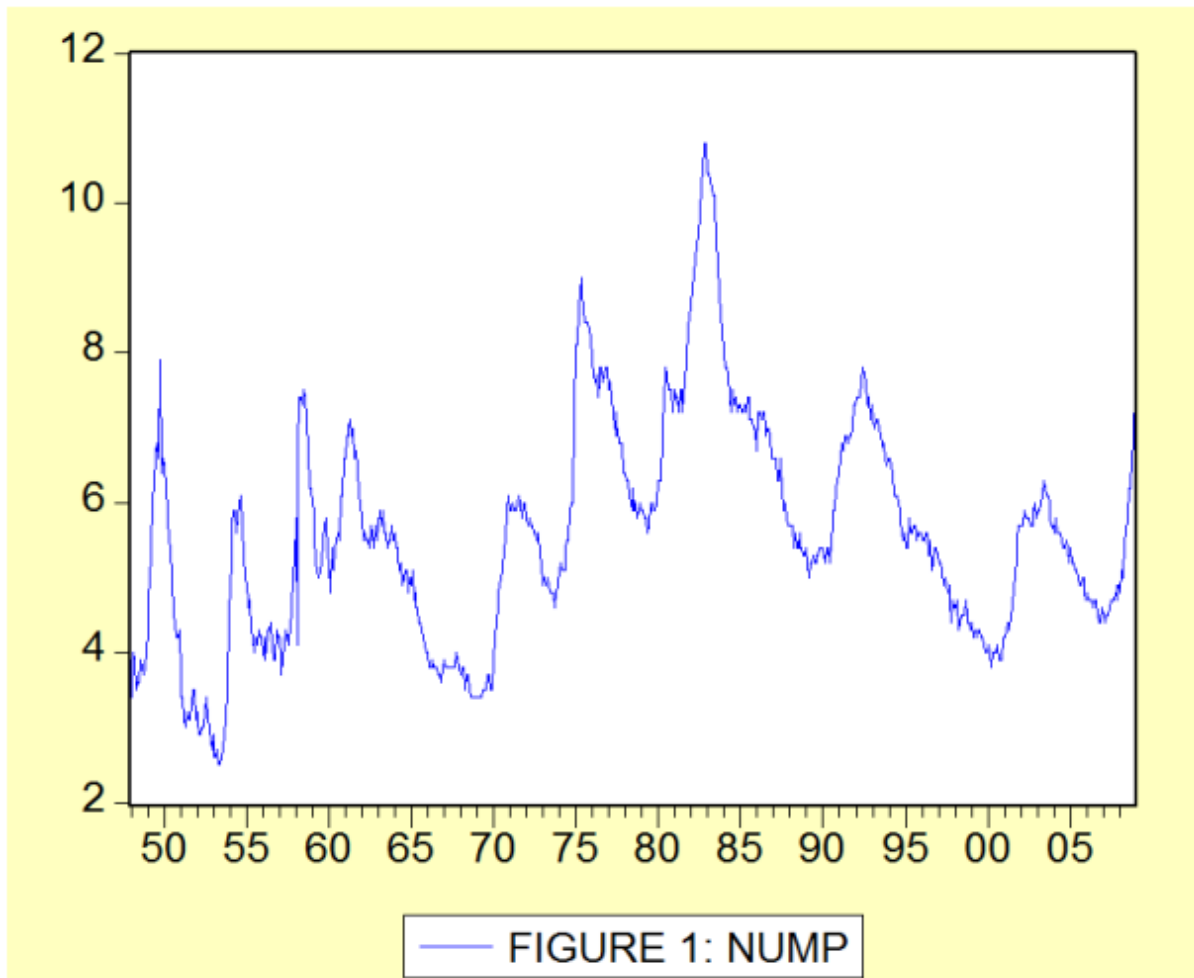
The model as summarized in Table 1 is given by

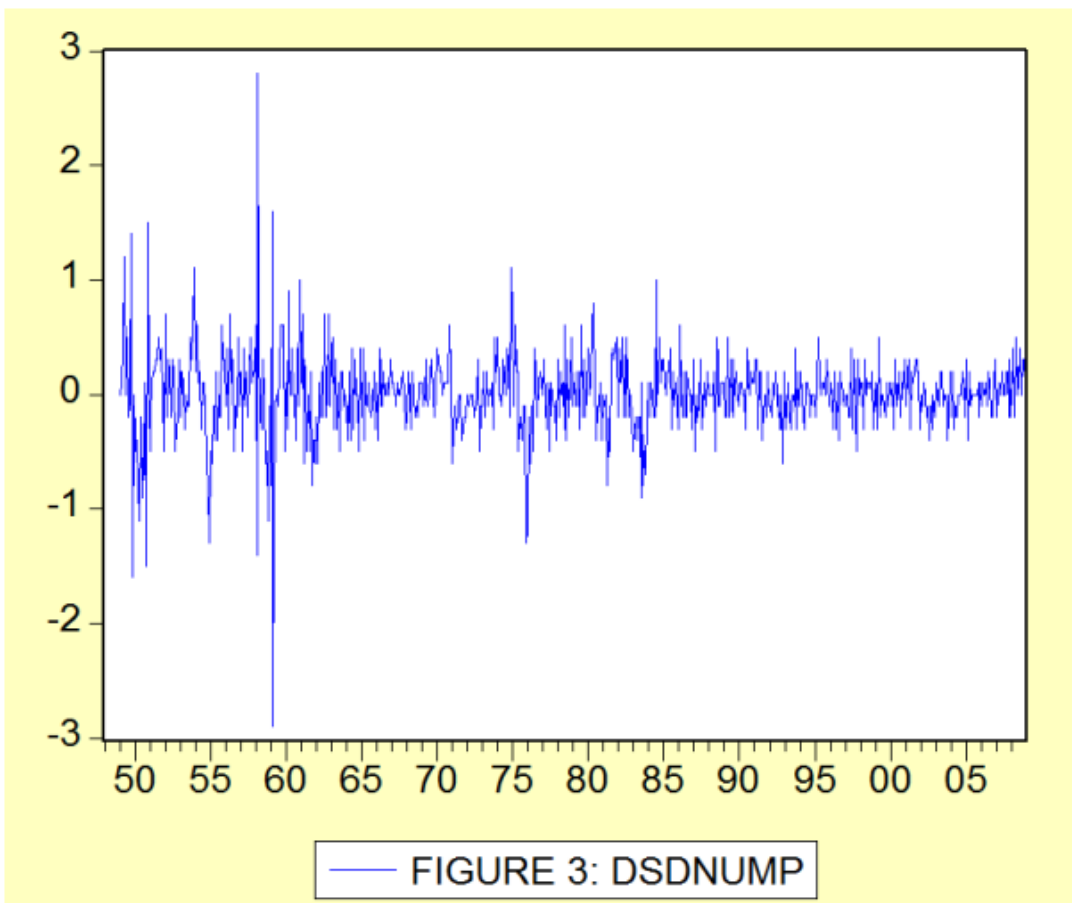
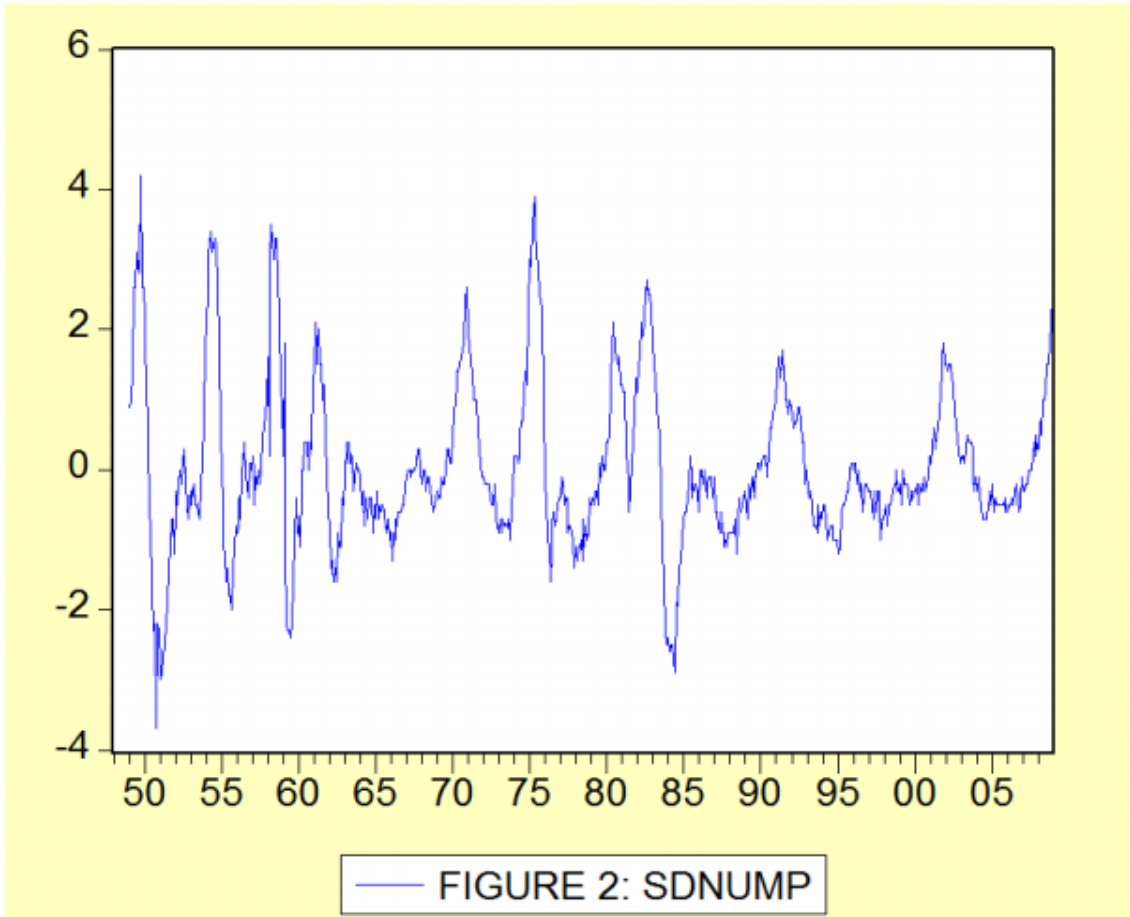
$$\text{DSDNUMP}_t = \underset{(\pm 0.0355)}{-0.0541} \text{DSDNUMP}_{t-1} + \underset{(\pm 0.0348)}{0.2002} \text{DSDNUMP}_{t-2} + \underset{(\pm 0.0356)}{0.1632} \text{DSDNUMP}_{t-3} - \underset{(\pm 0.0165)}{0.8853} \varepsilon_{t-12} + \varepsilon_t \quad (8)$$

The estimation involved 10 iterations. Only the first autoregressive coefficient is not significantly different from zero, being smaller than twice its standard error. The remaining coefficients including the MA one are statistically significant. Figure 5 shows that there is close agreement between the actual and the fitted models. Moreover the histogram of the residuals in Figure 6 shows that they are normally distributed with zero mean further indicating model adequacy.

## 5. Conclusion

We conclude that Nigerian Unemployment Rates follow an adequate  $(3, 1, 0) \times (0, 1, 1)_{12}$  model.





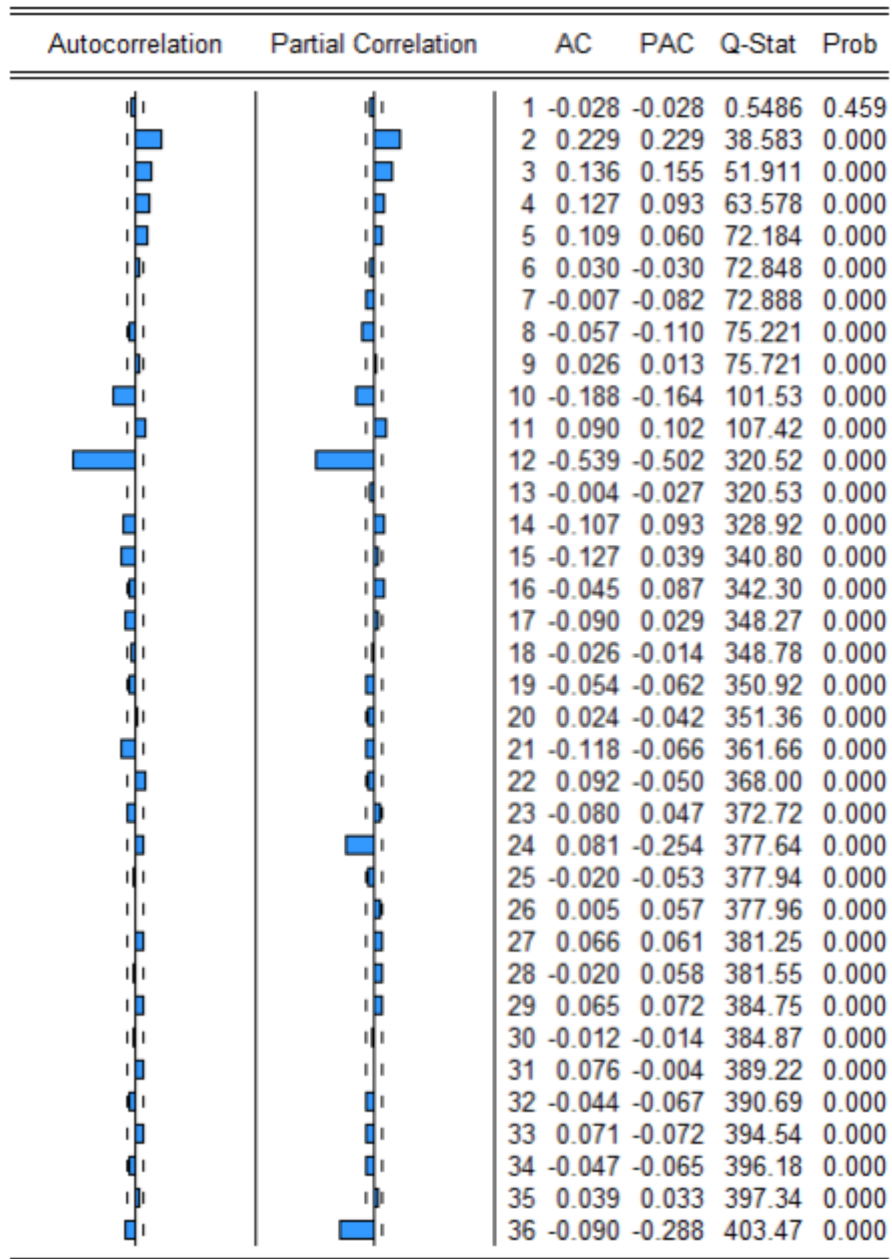
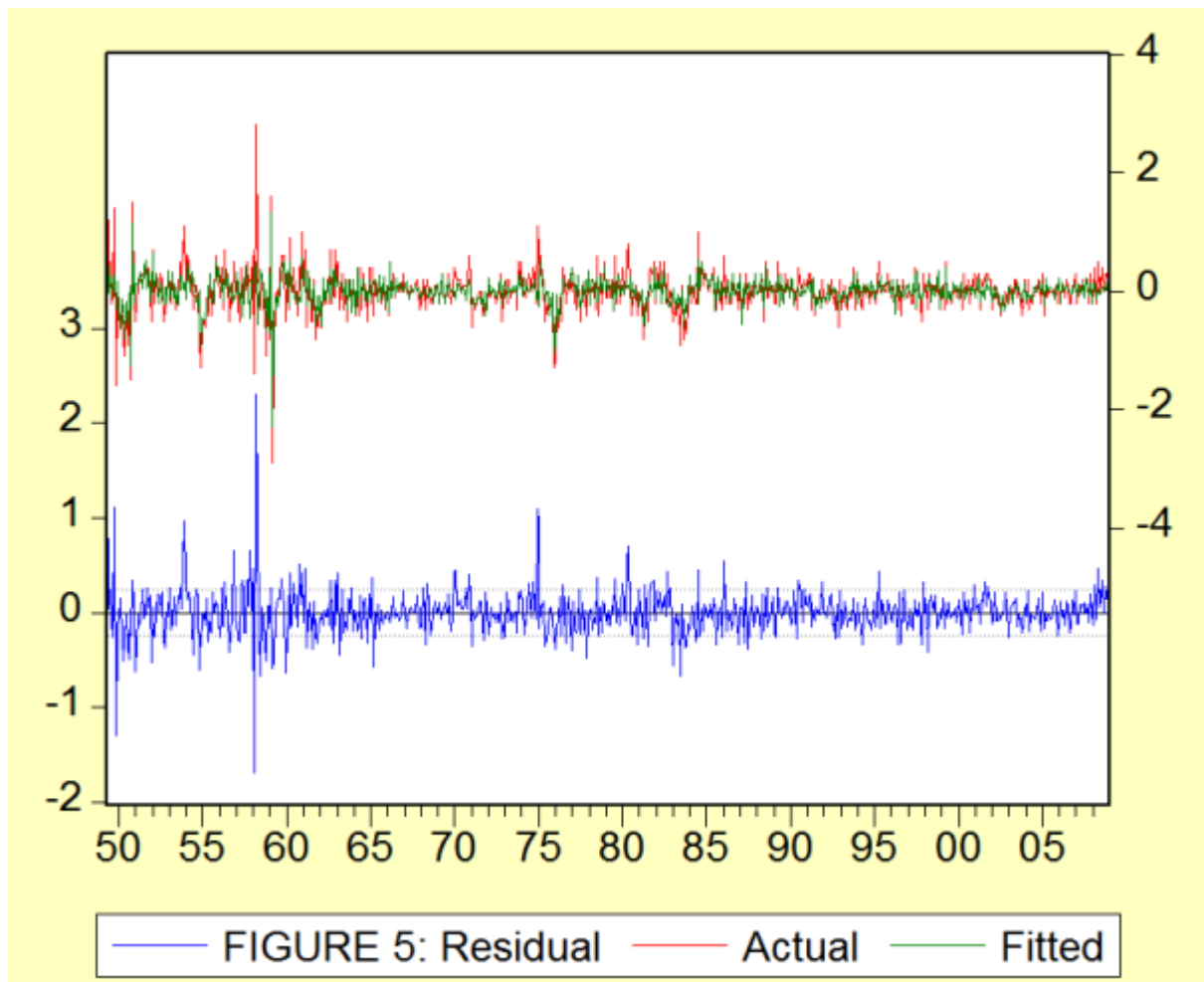


FIGURE 4: CORRELOGRAM OF DSDNUMP

TABLE 1: MODEL ESTIMATION

Dependent Variable: DSDNUMP  
 Method: Least Squares  
 Date: 08/13/12 Time: 09:41  
 Sample(adjusted): 1949:05 2008:12  
 Included observations: 716 after adjusting endpoints  
 Convergence achieved after 10 iterations  
 Backcast: 1948:05 1949:04

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.054112	0.035536	-1.522752	0.1283
AR(2)	0.200208	0.034810	5.751457	0.0000
AR(3)	0.163150	0.035625	4.579669	0.0000
MA(12)	-0.885325	0.016545	-53.51071	0.0000
R-squared	0.555877	Mean dependent var		0.001257
Adjusted R-squared	0.554005	S.D. dependent var		0.370652
S.E. of regression	0.247532	Akaike info criterion		0.051018
Sum squared resid	43.62575	Schwarz criterion		0.076569
Log likelihood	-14.26429	F-statistic		297.0526
Durbin-Watson stat	2.037684	Prob(F-statistic)		0.000000
Inverted AR Roots	.65	-.35+.36i	-.35	-.36i
Inverted MA Roots	.99	.86+.49i	.86	-.49i
	.49	-.86i	-.00	-.99i
	-.49	+.86i	-.86	-.49i
				-.99



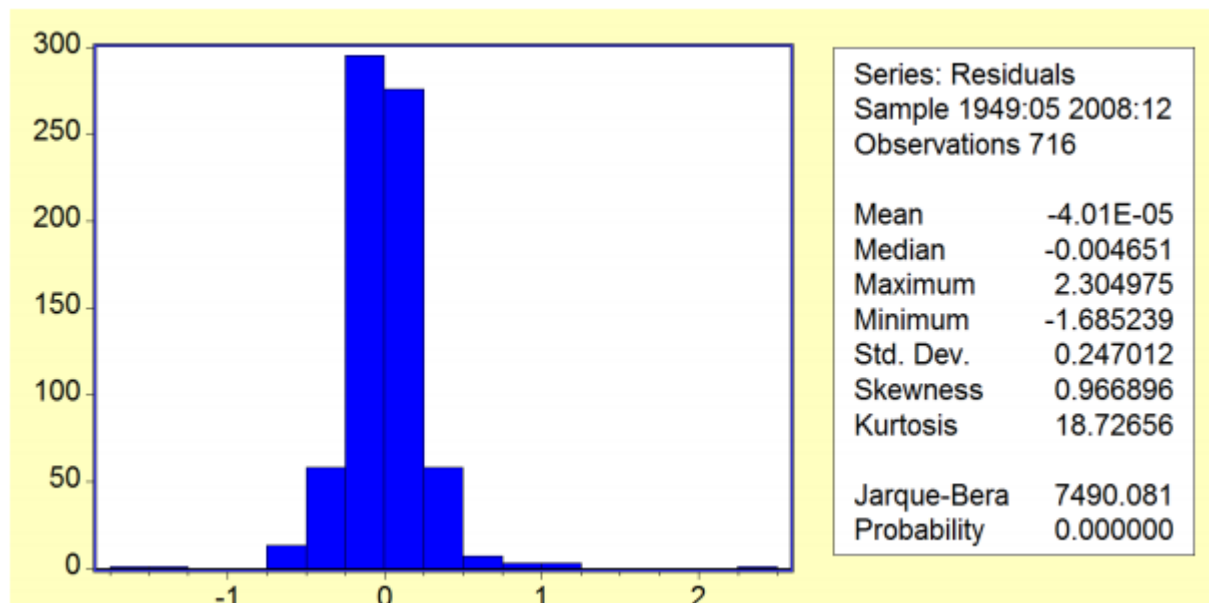


FIGURE 6: HISTOGRAM OF RESIDUALS

### References

- [1]. Box, G. E. P. and Jenkins, G. M. (1976). Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.
- [2]. Etuk, E. H. (1987). On the Selection of Autoregressive Moving Average Models. An unpublished Ph. D. Thesis, Department of Statistics, University of Ibadan, Nigeria.
- [3]. Etuk, E. H. (1998). An Autoregressive Integrated Moving Average (ARIMA) Simulation Model: A Case Study. Discovery and Innovation, Volume 10, Nos. 1 & 2: pp. 23 – 26.
- [4]. Madsen, H. (2008). Time Series Analysis, Chapman & Hall/CRC, London.
- [5]. Oyetunji, O. B. (1985). Inverse Autocorrelations and Moving Average Time Series Modelling. Journal of Official Statistics, Volume : pp. 315 – 322.
- [6]. Priestley, M. B. (1981). Spectral Analysis and Time Series. Academic Press, London.