International Journal of Modern Mathematical Sciences, 2021, 19(1): 79-89

International Journal of Modern Mathematical Sciences ISSN: 2166-286X

Journal homepage: www.ModernScientificPress.com/Journals/ijmms.aspx Florida, USA

Review

Review of Family of Autoregressive Integrated Moving Average Models in the Comportment of Autocorrelation Function for Non-Seasonal Time Series Data

Johnson Funminiyi Ojo, Rasaki Olawale Olanrewaju*

Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria

* Author to whom correspondence should be addressed; E-Mail: rasakiolawale@gmail.com

Article history: Received 14 April 2021, Revised 25 May 2021, Accepted 26 May 2021, Published 11 June 2021.

Abstract: A general linear-type time series imitation namely Autoregressive Integrated Moving Average (ARIMA) models have played indispensable theoretical and practical role in the representation and analysis of time series data. These theoretical and practical representations are usually denoted by ARIMA (p, d, q). When *d*, the differencing operator (integrating parameter) is zero the resulting models are also a generalize type called Autoregressive Moving Average Models (ARMA). It is of great importance to review members of the family of ARIMA because seeing these members at a glance will bring a better understanding and correct application of these models to time series data. Thirteen members of the family of ARIMA were reviewed in the presence of autocorrelation function.

Keywords: ARMA; ARIMA; Autocorrelation function; differencing operator

1. Introduction

It is sometimes a difficult task to see at glance different types of linear time series models in any given write-up. Most of the time, the user of such linear time series models will have to consult different write-up and gather them to be able to use the right linear-type time series representation for uniformly time series data in question. This review attempts to solve this problem, as in when a researcher lays hands on this write-up, the researcher will be able to see at a glance the different linear-types time series representation that will be appropriate for the selected study the at hand. The thirteen members of family of ARIMA considered in this review are as follows: Moving Average (MA) models, Autoregressive (AR) models, ARMA, Subset Autoregressive (SAR) models, Subset Autoregressive Moving Average (SARMA) models, Autoregressive Fractional Incorporated (ARFI) model, Autoregressive Fractional Incorporated Moving Average (ARFIMA) models, Subset Autoregressive Fractional (SARFI) models, Subset Autoregressive Fractional Incorporated Moving Average (SARFIMA) models, Integrated Autoregressive (IAR) models, ARIMA, Subset Integrated Autoregressive (SIAR) models, Subset Autoregressive Incorporated Moving Average (SARIMA) models.

2. Mathematical Formulation

2.1. Auto-Correlation Function

The assessment of the mean value of the product Y_t and Y_{t+k} from their deviations from their respective means at time interval "k" unit is called the auto-covariance of lag "k" and is usually denoted by $\gamma_k = \mathrm{cov}(Y_t, Y_{t+k})$

$$Autocorrelation(Y_{t+k}, Y_t) = \frac{Cov(Y_t, Y_{t+k})}{\sigma(Y_t)\sigma(Y_{t+k})} \quad (1)$$

$$= \frac{\gamma_k}{\gamma_0} \quad = \rho_k$$

The Autocorrelation Function usually denoted by (ACF) do furnish with guide to choose the optimal lag and traits associated to a time series. The ACF quantities evaluate the association between uniformly recorded time series at distinct distances "k" apart. The function aids to depict the different stages in which a time series pass through. It entails set of continuous values usually denoted by ρ_k that ranges from (-1, +1).

For preliminary model designation, the ACF function and its partial related function called Partial Autocorrelation Function are useful tools, (Shangodoyin and Ojo 2002). A typical example is found in the figure 1 below:

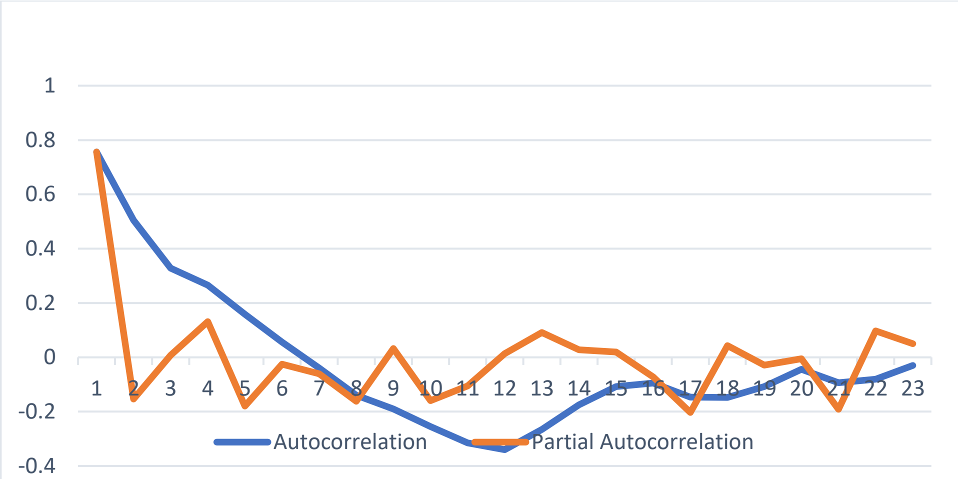

Figure 1. The graph of ACF and PACF

2.2. Autoregressive (AR) Models

This is correspondent to the form of multiple linear regression generalization of the form:

$$Z_i = \sum_{j=0}^{\infty} b_j Y_{t-j} + \varepsilon_t; \quad i = 1, \cdots, n \quad b\varepsilon\square$$

Such that the error term $\{\varepsilon_t\}$ is a purely white noise process with mean zero and variance σ_e^2, then the uniformly time varying series $\{Y_t\}$ is said to follow an autoregressive process of order "p" if it satisfies the difference equation:

$$Y_t - \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t \qquad (2)$$

where ε_t is the random noise Gaussian

$\phi_1, \phi_2, ..., \phi_p$ is a finite valued coefficient

$E\{Y_t\} = \mu$

The abbreviation for an autoregressive process of lag "p" is AR (p), which gave forth from the fact that Y_t lies solely on previous or immediate past values of Y and not as explanatory variables.

In general, the autoregressive process of lag "p" in equation (2) could be expressed as:

$$\Phi_1(B) Y_t = \varepsilon_t$$

where:

$$\Phi(B) = 1 - \phi_1 B^1 - \phi_2 B^2 - ... - \phi_p B^p$$

$$\Rightarrow Y_t = \phi_1^{-1}(B)\varepsilon_t = \psi(B)\varepsilon_t$$

We can then factorize

$$\Phi\left(B\right)=\left[1-G_1B\right]\left[1-G_2B\right],...,\left[1-G_pB\right].$$

(factoring out via Algebraic expression), we have

$$\Phi^{-1}(B)=\psi(B)=\left[1-G_1B\right]^{-1}\left[1-G_1B\right]^{-1},...,\left[1-G_1B\right]^{-1}$$

By using partial fractions, we have

$$\psi(B)=\frac{\sum k_i}{(1-G_iB)}=k_1\left[1-G_1B\right]^{-1}+\cdots+k_p\left[1-G_1B\right]^{-1} \qquad \text{where } i=1,\cdots,p$$

Thus for convergence of $\psi(B)=\Phi^{-1}(B)$, we must have $|B|\le 1$. Therefore, for stationary all the roots of the characteristics equation $\Phi(B)=0$ must lie outside the unit circle.

2.2.1. Condition for identifying and fitting of autoregressive process

The partial autocorrelation graph cut-off at dawdle "p" and the ACF function graph at order "k" decay exponentially to zero. (See the figure 1 above)

2.3. Moving Average (MA) Models

An MA model of lag "q" can be express as:

$$Y_t=\varepsilon_t-\theta_1\varepsilon_{t-1}-....-\theta_q\varepsilon_{t-q} \tag{3}$$

where $\{\varepsilon_t\}$ connote the random noise process with mean zero and variance σ^2. Then

$$Y_t=(1-\theta_1B-\theta_2B^2-\cdots-\theta_qB^q)\varepsilon_t$$

that is, $Y_t=\theta(B)\varepsilon_t$ where $\theta(B)=1-\theta_1B-\theta_2B^2-\cdots-\theta_qB^q$

The characteristic equation of (3) is $\theta(B)=0$. It can be shown that for any values of $\theta_1,\theta_2,...,\theta_q$, MA(q) is stationary, so no stationary condition is required. But in the expression

$$\varepsilon_t=\theta^{-1}(B)Y_t$$

we say that

$$\theta^{-1}(B)=(1-\theta_1B-\theta_2B^2-\cdots-\theta_qB^q)^{-1}$$

$$=(1-P_1B)^{-1}(1-P_2B)^{-1}....(1-P_qB^q)^{-1}$$

and partial fraction $K(\theta)=\theta^{-1}(B)=\sum_{i=1}^{q}T_i(1-P_iB),$

thus for convergence of $K(\theta)=\theta^{-1}(B)$, we must have $|B|\le 1$, then it implies that

$|P_i| < 1 \; \forall i = 1, 2, 3, ..., q$, where $P_i^{-1}; \; \forall \, i = 1, 2, 3, ..., q$ connote the characteristic equation of the roots. Hence, for invertibility condition to be satisfied, the characteristic equation of the roots must lie outside the unit circle, that is, $|P_i| < 1$.

2.3.1. Condition for identifying and fitting moving average process

The autocorrelation graph of order "k" cut-off after a particular lag say q whereas the partial autocorrelation graph decay exponentially to zero. Figure 1 above is applicable provided the ACF function replaced by the PACF in the table and vice versa.

2.4. Autoregressive Moving Average (ARMA) Model

A more logical extension in representing a linear process is to merge Autoregressive (AR) and Moving Average (MA) processes together. Apparently, stationary and non-stationary time series processes can be well generalized by ARMA models with parameters within the range of (-1, +1) so as to curb the problem parsimony.

A time series $\{Y_t\}$ is said to follow an autoregressive moving average model of order (p, q) (ARMA (p, q)) if it satisfies

$$Y_t - \phi_1 Y_{t-1} - \cdots - \phi_p Y_{t-p} = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \tag{4}$$

$$\Phi(B) Y_t = \theta(B) \varepsilon_t$$

Such that,

$$\Phi(B) = 1 - \phi_1(B) - \cdots - \phi_p B^p$$

$$\Theta(B) = 1 - \theta_1(B) - \cdots - \theta_q B^q$$

Both stationarity and invertibility condition must be met for the expression in (4).

2.4.1. Condition for identifying and fitting of autoregressive moving average process

To identify this model neither the ACF nor PACF cuts-off after a particular lag. AR, MA or ARMA models were broaden examined and analyzed some authors like (Chatfield, 1980; Walker, 1952; Shangodoyin and Ojo, 2002).

2.5. Subset Autoregressive (SAR) Models

A stochastic process $\{Y_t\}$ with zero-mean stationarity generated from an AR process with lag "k", denoted by AR (p), if it satisfies the difference equation

$$Y_t + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_k Y_{t-p} = e_t \tag{5}$$

Such that $\{e_t\}$ connote the Gaussian noise process with variance σ_ε^2. The $\{e_t\}$ will follow a Gaussian process, though it can follow any distributional form. Whenever the differential equation of equation (5) is fitted to uniformly time varying series $\{Y_t\}$, which incorporate full terms of the past values $\{Y_{t-i}; \forall\, i = 1,...,k\}$, it is referred to as full autoregressive model. Full autoregressive model is always characterizing by numerous parameterization, which always give rise to over-parameterization. Parable, some of these coefficients are sometimes eliminated from the system of equation when they are close to zero since they contribute insignificantly to the model. When these parameters are removed the resulting model is subset autoregressive (SAR). An appropriate algorithm is used to eliminate lags that are close to zero (Ojo *et al.*, 2008).

2.6. Subset Autoregressive Moving Average (SARMA) Models

This is similar to subset autoregressive (SAR) model described in section 5 above. The difference is in the inclusion of error term in the subset autoregressive model. The algorithm is also similar to SAR model. Given a full autoregressive moving average model which is optimal at order four for instance, the various subset autoregressive moving average having considered the algorithm is given as follows:

1, 2, 3, 4, 12, 13, 14, 23, 24, 34, 123, 124, 134, 234, 1234.

The equations for the results from the algorithm are:

$$Y_t = a_1 y_{t-1} - b_1 e_{t-1} + e_t$$

$$Y_t = a_2 y_{t-2} - b_1 e_{t-1} + e_t$$

$$Y_t = a_3 y_{t-3} - b_1 e_{t-1} + e_t \qquad (6)$$

.

.

.

$$Y_t = a_1 y_{t-1} - a_2 y_{t-2} + a_3 y_{t-3} + a_4 y_{t-4} - b_1 e_{t-1} + e_t \qquad (7)$$

Whenever 1=a_1, 2=a_2, 3=a_3, 4=a_4, minima model performance of AIC is achieved in (7) and such model is called Subset Time Series model (See Ojo, 2007).

2.7. Autoregressive Fractional Integrated and Autoregressive Fractional Integrated Moving Average Models (ARFI/ARFIMA) Models

According to Box-Jenkins ARIMA (p, d, q) examination, if the uniformly time-varying series of interest is non-stationary, the first order differencing (either by logarithm or $\Delta_{y_t} = Y_{t-1} - Y_t$ differencing) will behave well in trend provided there is no contamination of seasonal effect. However, it is believed that Δ_{y_t} will speedily experience a decomposing autocorrelations with free-like trend traits, so as to meet up the necessary condition of stationary-invertible process. In some highly contaminated time-varying series, the first order differencing might not be able to fit because stationarity is not ascertained, that is, slow decomposition of the autocorrelation to zero is experienced. In cases where the slow decomposition is experienced and non-stationarity was not attained, we used the fractional form of d $\in$ (-.5,.5) and say the stationary uniformly time series $\{Y_t\}$ has long memory.

The process of the Autoregressive Fractionally Incorporated Moving Average (AFRIMA) of lag order (p, d, q), is normally captured as ARFIMA (p, d, q). Alternatively, it can be rewritten using what is called "d"-operator notation as

$$\Phi(B)(1-B)^d Y_t = \Theta(B)\varepsilon_t$$

Such that "B" denotes the Backward shift operator; "d" connote the fractional integration index Assuming a uniformly time series observation Y_t with integer "t" as the index, such that Y_t are real numbers from the real line, then the ARFIMA (*p*, d, *q*) model could be generalized via

$$\Psi(B)Y_t = \Phi(B)\nabla^d Y_t + \Theta(B)\varepsilon_t$$

Where

$$1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p=\Phi(B);$$

$$1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q=\Theta(B);$$

$$y_t=\psi_1 y_{t-1}+\cdots+\psi_{p+d}\, y_{t-p-d}+\varepsilon_t-\theta_1\varepsilon_{t-1}-\cdots-\theta_q\varepsilon_{t-q}$$

ψ_i are the coefficients of the autoregressive process of the ARFIMA (*p*, d, *q*) model; θ_j are the moving average process coefficients associated with the error terms (ε_t). These error terms are usually subjected to $\varepsilon_t \sim IID\left(0,\sigma^2\right)$, that is, independent, identically distributed with mean zero and variance σ^2 . When q=0 in ARFIMA (p, d, q) the resulting equation is resulted to autoregressive fractional integrated model (Ojo, 2016).

For necessity condition of stationary-invertible process of the ARFIMA model, the interval d $\in$ (-.5,.5) catered for achieving suitably number of differencing. In a nut shell, the ARFIMA (p, d, q) is classified among the long-memory models with slow dying out decomposition of autocorrelations of the Box-Jenkins ARMA (p, q) models. It is to be noted that the additional index "d" accommodates and captures long-run traits in the observational series, while the Box-Jenkins ARMA (p, q) models capture short-run traits. When $d \in \left(0, 0.5\right)$, it is assumed that the ARFIMA model has a long-memory such that autocorrelation function decomposes to zero at a rate of hyperbolic; when $d \in \left(-0.5, 0\right)$, the ARFIMA process has a long-memory such that autocorrelation function decomposes to zero at a rate of hyperbolic. According to Bos, *et al*. (2008), an ARFIMA processes with $d \in (-0.5, 0)$ possessed intermediate-memory, which could also be referred to as long memory processes. A typical example is shown below

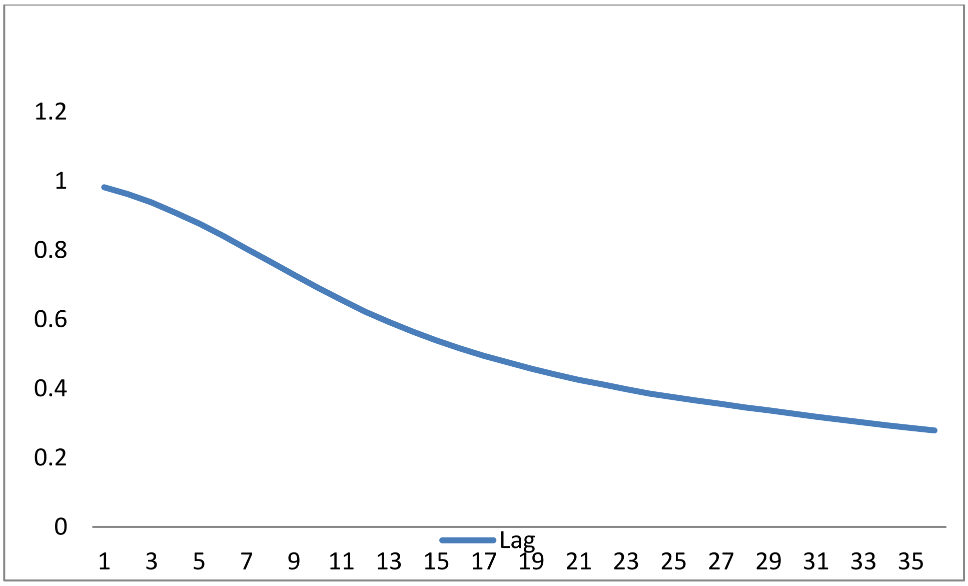

Figure 2. Graph of autocorrelation function showing slow decay

2.8. Subset Autoregressive Fractional Integrated and Subset Autoregressive Fractional Incorporated Moving Average Models (SARFI/SARFIMA) Models

ARFIMA model of selected orders can be fitted over an optimal chosen model performance of Akaike Information (AIC) that is minima in a pool of AIC. Assuming the order of an ARFIMA model is p+d+q with a generalization of $\Phi(B)\nabla^d Y_t + \Theta(B)\varepsilon_t = \Psi(B)Y_t$. If the average sum of squares for the residual of the model is $\hat{\sigma}_e^{2(1)}$ such that the minima AIC for the full model equals AIC(1). Having fitted the initial full model, a linear combination of best subset model can be achieved via $2^k - 1$ subset via choosing the minima AIC (Sangodoyin and Ojo, 2003). Let the subset ARFIMA model be

$$y_t = \psi_{n1} y_{t-n1} + \cdots + \psi_{nl+d} y_{t-nl-d} + \varepsilon_t - \theta_{k1}\varepsilon_{t-k1} - \cdots - \theta_{kq}\varepsilon_{t-kq}$$

Where

$n_1, n_2, ..., n_{l+d}$; $k_1, \ldots k_q$ are subsets of the integers (1, 2, …, p+d+q).

Assuming the average sum of squares of residuals is $\hat{\sigma}_e^{2(2)}$, the $\hat{\sigma}_e^{2(2)}$ valued AIC at AIC(2) is AIC(1)>AIC(2). The generalization with AIC (2) is the fraction part of the ARFIMA model. When "q" is zero in the above model, the resulting equation is subset autoregressive fractional integrated model. (Ojo and Rufai, 2016).

2.9. Integrated Autoregressive Model

Integrated autoregressive models (p, d, 0,) is given as

$\psi(B)Y_t = \phi(B)\nabla^d Y_t$

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$

$$Y_t = \psi_1 Y_{t-1} + \cdots + \psi_{p+d} Y_{t-p-d} + \varepsilon_t \tag{8}$$

$\phi(B)$ is the operator function for the autoregressive process that is assumed to have attained stationarity, such that its roots lies outside unit circle, that is, $\phi(B) = 0$. $\nabla^d \phi(B) = \psi(B)$ will be termed as the operator of the generalized autoregressive process that is assumed not to have attained stationarity. ψ_i are the coefficients of the incorporated autoregressive process part with ε_t error terms. These error terms ε_t are usually subjected to $\varepsilon_t \sim IID\left(0, \sigma^2\right)$, that is, independent, identically distributed with mean zero and variance σ^2

2.10. Autoregressive Incorporated Moving Average Model

In continuous-type time series generalization, ARIMA model is known to be a full sketch generalization of ARMA model. The model is normally fitted to time variant series observations either via from the model performance view or forecast point of view. It is usually via the framework of ARIMA (*p*, *d*, *q*), where p, d, and p are non-zero positive integers referring to order of autoregressive, integrated and moving average processes respectively.

Assuming a uniformly time-varying observations Y_t, where the integer "*t"* connote the time index such that Y_t are real numbers from the real number line, then the ARIMA (*p*, d, *q*) model could be written as

$\Phi(B)\nabla^d Y_t + \Theta(B)\varepsilon_t = \Psi(B)Y_t$

Where

$1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p = \Phi(B)$;

$1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q = \Theta(B)$;

$$y_t = \psi_1 y_{t-1} + \cdots + \psi_{p+d} y_{t-p-d} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \tag{9}$$

ψ_i are the coefficients of the autoregressive process of the ARFIMA (*p*, d, *q*) model; θ_j are the moving average process coefficients associated with the error terms (ε_t). These error terms are usually subjected to $\varepsilon_t \sim IID\left(0,\sigma^2\right)$, that is, independent, identically distributed with mean zero and variance σ^2. Autoregressive Integrated (ARI) and ARIMA were extensively studied by (Box and Jenkins, 1970; Anderson, 1971, 1977).

It is to be note that if the autocorrelations fail to tail-off, but rather the autocorrelation values are proximity approaches one over many lags, then time series is non-stationary and differencing might be needed. An extremely slow decaying ACF is signal for non-stationarity which can be formally verified by unit root tests.

2.11. Subset Integrated Autoregressive and Subset Autoregressive Integrated Moving Average Models

Fitting autoregressive processes and ARIMA models to optimal orders at minima AIC. Assuming fully integrated autoregressive model with p+d and autoregressive integrated moving average. With p+d+q such that models

$$Y_t = \psi Y_{t-1} + \cdots + \psi_{p+d} Y_{t-p-d} + \varepsilon_t$$

denoted by IA (p, d) and

$$y_t = \psi_1 Y_{t-1} + \cdots + \psi_{p+d} Y_{t-p-d} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q}$$

Assuming the average sum of squares of residuals is $\hat{\sigma}_e^{2(2)}$, the $\hat{\sigma}_e^{2(2)}$ valued AIC at AIC(2) is AIC(1)>AIC(2). The generalization with AIC (2) is the fraction part of the ARFIMA model. When q is zero in the above model, the resulting equation is subset autoregressive fractional integrated model.
Let the best fractional part of the Integrated Autoregressive model be

$$Y_t = \psi_{m1} Y_{t-m1} + \cdots + \psi_{ml+d} Y_{t-ml-d} + \varepsilon_t$$

where $m_1, m_2, ..., m_{l+d}$ are subsets of the integers (1, 2,…, p+d). AIC(1)>AIC(2) is usually meant for average sum of squares residual for $\hat{\sigma}_e^{2(2)}$. This is the subset integrated autoregressive. Additionally, let the best subset autoregressive integrated moving average model be

$$Y_t = \psi_{n1} Y_{t-n1} + \cdots + \psi_{nl+d} Y_{t-nl-d} + \varepsilon_t - \theta_{k1} \varepsilon_{t-k1} - \cdots - \theta_{kq} \varepsilon_{t-kq}$$

where $n_1, n_2, ..., n_{l+d}$; $k_1,….k_q$ are fractional part of the model with integers (1, 2,..., p+d+q). $AIC(22) \le AIC(11)$ is usually meant for average sum of squares residual for $\hat{\sigma}_e^{2(21)}$ (Ojo, 2009).

3. Conclusions

In this review, we have been able to note down the members of family of ARIMA model. The condition to be fulfilled for us to identify these models have been enumerated. Not only this, when we have data and when we want to fit these models, how the autocorrelation function will behave for us to know the appropriate one to fit have been discussed. In a nutshell, seeing these models at a glance and the condition to be fulfilled for us to fit these models to the data under study have been achieved.

References

[1] Anderson TW. Estimation for Autoregressive Moving Average Models in the Time and Frequency Domains. *The Ann. of Stat., 5(5)* (**1977**): 842-865.

[2] Anderson TW. *The Statistical Analysis of Time Series,* New York and London Wiley, **1971**.

[3] Bos CS, Franses PHBF, Ooms M. Long memory and level shifts: Re-analyzing inflation rates. *Empirical Economics*, *24*(**2008**): 427-449.

[4] Box GEP, Jenkins GM. *Time Series, Forecasting and Control*. Holden-Day: San Francisco, **1970**.

[5] Chatfield C. *The Analysis of Time Series: An Introduction*. Chapman and Hall, **1980**. Fsb.muohio.edu/lij14/690_ss.pdf.

[6] Ojo JF. On the Estimation and Performance of Subset Autoregressive Moving Average Models.

[7] *European Journal of Scientific Research, 18 (4)*(**2007**): 700-706.

[8] Ojo JF. On the estimation and performance of subset autoregressive integrated moving average models. *European Journal of Scientific Research,* 28(2)(**2009**): 287-293.

[9] Ojo JF. On fractionally Integrated Bilinear Time Series Models. *Science Focus, 21(1)*(**2016**): 68 - 76.

[10] Ojo JF, Olatayo TO, Alabi OO. Forecasting in Subsets Autoregressive Models and auto-projectives Models. *Asian Journal of Scientific Research, 1(5)*(**2008**): 481-491.

[11] Ojo JF, Rufai O. On Subset Autoregressive Fractionally Integrated Moving Average Models.

[12] *International Journal of Physical Science, 11(1)* (**2016**): 15-20.

[13] Online.stat.psu.edu/stat510/book/export/htm/665.

[14] Shangodoyin DK, Ojo JF. Elements of Time series Analysis. *Rashmed Publication, Ibadan* (**2002**).

[15] Shangodoyin DK, Ojo JF. On the Determination and Estimation of Subset Stationary Bilinear Time Series Autoregressive models. *Journal of Science, Engineering and technology, 10(3)*(**2003**): 5080-5091.

[16] Walker AM. Some properties of Asymptotic Power Functions of Goodness-of-fit Tests for Linear Autoregressive Schemes. J. *Roy. Stat. Soc., Series, 14*(**1952**): 117-134.