# CHAPTER ONE

# INTRODUCTION

## 1.1 Background of the study

According to statistical literature for example Chatfield (2003), time series is defined as a set of observations of a random process made sequentially over time. It is usually based on an underlying real-valued discrete-time stochastic process $X_t,\ t \in \mathbb{Z}$ and the available data $\{X_t\}$, t = 1, 2,…, n which is a subset of all possible series from which $X_t$ could have been selected. These observations however, may be affected by four different forms of variation usually referred to as the components of time series namely the Trend ($T_t$), the seasonal ($S_t$), the Cyclical $(C_t)$ and the Irregular $(I_t)$ variations. Hence, the entire methods and techniques employed in formulating and estimating models for the purpose of studying these variations and various statistical analyses of the observed data are generally referred to as time series analysis. Usually the process begins with a preliminary data-analysis such as graphical illustration of the series, generally referred to as the 'time plot' and then the inspection of the autocorrelation function (ACF) and the partial autocorrelation function (PACF) plots. After which analysis of the data can be done using either the frequency domain method or time-domain method.

According to Subba Rao and Gabr (1984), the study of time series analysis and modeling has been well developed for the past three decades. The frequency domain as well as time domain techniques have been examined and applied in the analysis and modeling linear time series and a great number of studies have examined the analysis of linear time series models usually designed to model the covariance structure of the series. These are either the Autoregressive (AR) or the Moving Average (MA) models, which can be combined to have the Autoregressive Moving Average (ARMA) models. The series is said to be an autoregressive process (AR) of order $p$ if it satisfies the relation:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + e_t \tag{1.1}$$

where { $e_t$ } are independent and identically distributed ($iid$) random variables with mean zero and variance $\sigma^2$.

This implies that $E(X_t \mid X_{t-1},...,X_{t-p}) = \sum_{i=1}^{p} \phi_i X_{t-i}$. So, the past values of $X_t$ would have a linear influence on the conditional mean of $X_t$. Then the parameters are estimated using the conventional least squares method.

On the other hand, $X_t$ is said to follow a moving average of order $q$, if it satisfies the relation:

$$X_t = \sum_{j=1}^{q} \theta_j e_{t-j} \tag{1.2}$$

where { $e_t$ } are $iid$ random variables with zero mean and variance 1.

In general, a time series is said to be linear and have a moving average of order $q$ representation if:

$$X_t = \sum_{j=-\infty}^{\infty} \theta_j e_{t-j} \tag{1.3}$$

where { $e_t$ } are independent and identically distributed random variables with finite variance.

Similarly, the general Autoregressive Moving Average process, ARMA $(p,q)$ is a linear stochastic model given by the relation:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + e_t + \sum_{j=1}^{\infty} \theta_j e_{t-j} \tag{1.4}$$

where $X_t$ is being expressed in terms of its own past values and a disturbance.

The ARMA model assumes that $X_t$ is the result of a linear filter that transforms the past innovations $e_{t-i}, i = 0,1,...,\infty$. That is, $X_t$ is linear, which is an assumption based on the Wold's decomposition theorem (Wold; 1938).

In most cases, the assumption of linearity in time series is found dubious due to the fact that a lot of physical real-life phenomena would not be adequately explained by linear models. As a result, there is the need for models that would take the nonlinear nature of such phenomena into consideration and models with higher order spectra. Volterra (1930) and Wiener (1958) presented presented a good background study that led to the development of nonlinear models. However, the Wiener's representation looks too complex and the statistical analysis procedures involved were cumbersome. As a result, some other authors like Ozaki and Oda (1977), Jones (1978), Haggan and Ozaki (1980) and Tong and Lim (1980) presented some more specific types of non-linear models. Also, Campbell and Walker (1977) made important contributions which led to further development of nonlinear time series models. They made us to realize nonlinear models can be categorized into two namely, models which are nonlinear in mean and also models which are non-linear in variance (heteroskedastic). For example, the following are three nonlinear time series models:

(i) *Exponential Autoregressive Models*:

The exponential autoregressive models were proposed by Ozaki (1980) and Haggan and Ozaki (1981). They examined the problem of numerical evaluation and simulations of the parameters by minimizing the sum of squares of the penalty function;

$$Q_n(\beta) = \sum_{t=m+1}^{n} \{X_t - X_{t|t-1}(\beta)\}^2 \tag{1.5}$$

The first order Exponential autoregressive model is defined as:

$$X_t - \{\psi + \pi \exp(-\gamma X_{t-1}^2)\} X_{t-1} = e_t \tag{1.6}$$

for all $t \geq 2$ and $X_t$, being an initial variable.

*(ii) Threshold autoregressive (TAR) processes*

These models were first examined by Tong with the idea of a piecewise linearized autoregressive model obtained by introduction of a local threshold dependence on the amplitude $X_t$.

Tong and Lim (1980) considered the estimation of the parameters of the threshold model using the maximum likelihood method of estimation. The first order threshold autoregressive is;

$$X_t - \sum_{j=1}^{m} a^j H_{t-1} H_t(X_{t-1}) = e_t, \qquad \forall t \geq 2 \tag{1.7}$$

where $X_t$, is an initial variable and $H_j(X_{t-1}) = 1(X_{t-1} \in F_j), 1(\cdot)$ is an indicator function and $F_1, ...., F_m$ are disjoint regions.

*(iii) Random coefficient autoregressive (RCA) models*

This nonlinear model is studied by Aase (1984). He defined the model by allowing random additive perturbations of the coefficients of the AR models. Hence, a $d$ - dimensional RCA model of order p is given by:

$$X_t - \sum_{i=1}^{p} (a_i + b_{ti}) X_{t-i} = e_t \qquad \forall -\infty < t < \infty \tag{1.8}$$

where $a_i, i = 1, ..., p$ are deterministic $d \times d$ matrices and $\{b_t(p)\} - \{[b_{t1}, ..., b_{tp}]\}$ is a $d \times pd$ zero matrix process where $b_i(p)'s$ independent and identically distributed and also independent of $e_t$.

It should be noted that nonlinear time series are obtained through nonlinear dynamic equations. Hence, they display features and attributes that one cannot express as linear processes for example time-changing variance, higher-moment structures, asymmetric cycles, breaks and thresholds.

Moreover, in most cases many time series processes usually characterized by a seasonal component that occurs at every '$s$' observations. That is, at every length of season. For example, the amount of rainfall and temperature in the months of a year usually vary from one month to another with similar pattern over the years. So, a farmer for example would need to understand the seasonality of these weather conditions for proper planning in planting and harvesting of his farm produce. Similarly, companies and organizations that sell items like sunscreen, Christmas light and decorative items, and so on would see sales jump up at some period of the year and drastic fall at other times. Hence, there is need for them to understand the seasonality behaviour of their business for proper planning, staffing and other decision making in order to maximize profit as well as minimize cost of production. So, fitting a non-seasonal time series model in such situations would result in a poor and bias estimate of the parameters of such series. Therefore, it is very important to remember the effects of seasonality when modeling and analyzing data from such series.

In addition, quite a number of studies in time past have considered seasonal time series models and these examined the performance of the series only at the peak of seasons. However, less attention has been given to a seasonal nonlinear time series model that would study the performance of a time series before the peaks are reached, at the peaks and after the peaks of seasons.

## 1.2 Statement of the problem

The concept of seasonality has been widely discussed and examined in relation to linear time series models such as Beaulieu and Miron (1993), Ghysels *et al*. (1994), Harvey and Scott (1994), Hylleberg (1992), Hylleberg *et al*. (1993), Osborn and Rodrigues (1998), Xinghua *et al*., (2012), Eleazar *et al.* (2016). These include the linear Seasonal Autoregressive (SAR) and the pure Seasonal Autoregressive Integrated Moving Average (SARIMA) models. However, little or no attention has been given to the development of a seasonal model particularly the mixed seasonal autoregressive integrated moving average bilinear model (a nonlinear type) when it is applied to real seasonal time series data. It is our belief that when a mixed seasonal nonlinear model is fitted, it would help us to examine the behaviour of the series before the peaks are reached, at the peaks as well as after the peaks of the season and would equally perform better than a linear model. Therefore in this study, our intention is to fit mixed nonlinear seasonal time series model to a real series with a view to tracking the total behavior of the seasonal time series for adequate planning.

## 1.3 Aim and Objectives of the study

The main aim of this study is to develop a seasonal nonlinear time series model (Mixed seasonal autoregressive integrated moving average one-dimensional bilinear (MSARIMAODBL)) that would study the performance of a series before, at and after the peak of seasons.

The specific objectives are to:

i. specify the MSARIMAODBL model with its associated subsets namely; Pure SARIODBL, Mixed SARIODBL and Pure SARIMAODBL.
ii. derive the stationarity condition required for each specified model.
iii. estimate the parameters of the MSARIMAODBL models.
iv. validate the models using real life data as well as simulated data.
v. compare the performance of the specified models and also with the existing linear seasonal models.
vi. forecast based on the optimal model.

## 1.4 Justification of the study

Bearing it in mind that many time series processes exhibit both the seasonal and non-seasonal behaviour coupled with the fact that they are nonlinear in nature, it is expedient to obtain a model that would be able to capture and explain each of these behavior in the series. Since a pure seasonal time series model studies the series only at the peak of seasons and a nonseasonal time series model studies the series only along a trend line, none of these is suitable to model the series effectively. Hence, a mixed seasonal time series model that would study the series before the peaks, at and after the peaks is inevitable in such situations. So, this study proposes a nonlinear seasonal model known as the mixed seasonal autoregressive integrated moving average one-dimensional bilinear time series model that is capable of attaining stationarity for nonlinear seasonal time series. One major advantage of this model is that it would avail us the opportunity to track the total behaviour of the series in and out of season.

## 1.5 Some useful definitions

### 1.5.1 Convergence absolutely almost surely

Given a sequence $\{X_n, n > 1\}$ of random vectors, all of the same order $p \times 1$ defined on some probability space, then $\sum_{n\geq 1} X_n$ is said to converge absolutely almost surely if:

$$\sum_{n\geq 1} |(X_n)_i| < \infty \qquad \forall \text{ i} = 1,2,\ldots\text{p} \tag{1.9}$$

### 1.5.2 Convergence in the mean

Given a sequence $\{X_n, n > 1\}$ of random vectors of the same order $p \times 1$ defined on some probability space, then $\sum_{n\geq 1} X_n$ converges in the mean if;

$$\lim_{m\to\infty} E\left|\sum_{n=1}^{m} (X_n)_i - (X)_i\right| = 0 \tag{1.10}$$

for every $i = 1,2,\ldots p$ and any there exists a random vector $X$, $E$ denotes the expectation value.

### 1.5.3 Cauchy-Schwartz inequality:

The Cauchy-Schwartz inequality is one of the most important inequalities in Mathematics. The inequality form for sums was published by Augustin-Louis Cauchy in 1821, while its corresponding integral form was first proved by Viktor Bunyakovsky in 1859. The modern proof of the integral form of the inequality was given by Hermann Amandus Schwarz in 1888. The inequality can be expressed in different ways. For example,

$$Cov^2(X,Y) \leq \sigma_X^2 \sigma_Y^2$$

Similarly, the inequality states the relationship between any two vectors X and Y of an inner product space as:

$$\left|\langle X,Y\rangle\right|^2 \le \langle X,X\rangle \bullet \langle Y,Y\rangle$$

where $\langle \bullet , \bullet \rangle$ is the inner product such as the dot product.

If the two vectors are linearly dependent, squaring both sides and referring to the norms of the vectors, then inequality is given by:

$$\left|\langle X,Y\rangle\right| \le \|X\|\|Y\| \tag{1.11}$$

### 1.5.4 The Kronecker product

In Mathematics, matrix functions are of two categories namely the Kronecker matrix products and the non-Kronecker or ordinary matrix products. Neudeker (1969). According to Huamin and Feng (2013) the Kronecker product is named after Leopold Kronecker (1823–1891), a German mathematician. This product (usually represented by $\otimes$) is an operation on any two matrices of arbitrary sizes which result in a block matrix. So, given any matrices $A$ and $B$, the Kronecker product of $A$ and $B$ is given by:

$$A_{mXn} \otimes B_{pXq} = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

$$\Rightarrow (A \otimes B)_{p(r-1)+v,q(s-1)+w} = a_{rs}b_{vw}$$

Given a matrix F, if we represent the $(i,j)^{th}$ element of F by $(F)_{ij}$ or $(F)_{i,j}$, if the element of F are not indicated specifically otherwise and a column vector, C. The $i^{th}$ component of C written $(C)_i$ or $(C)_{i1}$ as in the case where the elements of C are implicitly represented. With respect to this given notation, then;

$$\Rightarrow (A \otimes B)_{ij,uv} = a_{ij}b_{uv} = (A \otimes B)_{(i-1)r+u,(j-1)s+v} \le k\lambda^{r-1} \tag{1.12}$$

Some of the properties of the kronecker product include:

(i) Bilinearity and associativity

$A \otimes (B+C) = A \otimes B + A \otimes C$

$(A+B) \otimes C = A \otimes C + B \otimes C$

$$(kA)\otimes B = A\otimes(kB) = k(A\otimes B)$$

$$(A\otimes B)\otimes C = A\otimes(B\otimes C) \quad (1.13)$$

(ii) Non-commutativity

Generally speaking, $A\otimes B$ and $B\otimes A$ are not the same, but are permutation equivalent. That is;

$$A\otimes B = P(B\otimes A)Q \quad (1.14)$$

(iii) The mixed-product property:

Given matrices A, B, C and D such that AC and BD exist, then;

$$(A\otimes B)(C\otimes D) = (AC)\otimes(BD) \quad (1.15)$$

This is referred to as the mixed-property because it mixes the conventional matrix product with the Kronecker product.

(iv) The inverse property: if each of A and B are invertible, the inverse of A $\otimes$ B also exists. Hence;

$$(A\otimes B)^{-1} = A^{-1}\otimes B^{-1} \quad (1.16)$$

# CHAPTER TWO

# LITERATURE REVIEW

## 2.1 Introduction

The theory of time series analysis began early with stochastic processes, which refers to the evolution of some system over time usually represented by a variable which is subject to irregular variation. The initial actual application of autoregressive models to time series data can be attributed to the study carried out by Yule and Walker in the 1920s and 1930s. They also presented the moving average process to deduce periodic fluctuations in a series, for example fluctuations due to seasonality. Herman Wold then later introduced the Autoregressive Moving Average (ARMA) models for stationary time series. However, he was unable to obtain a likelihood function to derive maximum likelihood (ML) estimation of the parameters. This challenge was later overcame by Box and Jenkins (1970) in their book titled "Time series Analysis" which contains the specification, estimation, diagnostics and forecasting for individual series. Moreover the conventional and commonly used Box-Jenkins models for forecasting and seasonal adjustment can be traced back to these models.

According to Hyndman (2010), we have a seasonal fluctuation when a series is affected by seasonal factor which is usually of a fixed and known period, which can be quarterly, monthly, weekly or daily. Hence, sometimes we refer to seasonal series as periodic time series. Seasonality can be detected using some graphical techniques such as a run sequence plot and seasonal subseries plot. We can also use the Multiple box plots as an alternative to the seasonal subseries plot to detect seasonality. The autocorrelation function plot can also help in identifying seasonality.

A pure seasonal ARMA model, usually denoted by $ARMA(P,Q)s$, is given by:

$$X_t = \Phi_1 X_{t-s} + \Phi_2 X_{t-2s} + \ldots \Phi_P X_{t-Ps} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + \ldots + \Theta_Q e_{t-Qs} + e_t \qquad (2.1)$$

This can be written in operator form:

$$\Phi_P(B^s)X_t = \Theta_Q(B^s)e_t \tag{2.2}$$

where $\Phi_P(B^s)$ and $\Theta_Q(B^s)$ are seasonal autoregressive and moving average operators respectively.

## 2.2 Stochastic process

A stochastic or random process is defined as a collection of some random variables that would take values in a set *S*, usually referred to as the state space. This group of random variables is usually denoted by another set *T*, known as the index set. In general, it can be described by an n-dimensional probability distribution $p(\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_\mathrm{n})$. The two most common index sets are the set of natural numbers $\mathbb{N} = \{0,1,2,\ldots\}$ which usually represent discrete time and the positive real numbers $\mathbb{R}^+ = [0, \infty)$ which represents continuous time. Usually, the first set are random variables $\{X_0, X_1, X_2, \ldots\}$ while the second is is such that $\{X_t, t \geq 0\}$.

If $\mathbb{R}^+ = (0,1, \ldots)$ then, $X_t$ is said to be a discrete time stochastic process, but if $\mathbb{R}^+ = [0, \infty)$, then $X_t$ is a continuous time process. The joint distribution is stable with respect to time and is given by:

$$p\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\} = p\{X_{t_1+k}, X_{t_2+k}, \ldots, X_{t_n+k}\}$$

The autocovariances are given by:

$$\gamma_k = Cov(X_t, X_{t-k}) = E(X_t - \mu)(X_{t-k} - \mu) \tag{2.3}$$

and the autocorrelations:

$$\rho_k = \frac{Cov(X_t, X_{t-k})}{[Var(X_t) \cdot Var(X_{t-k})]^{\frac{1}{2}}} = \frac{\gamma_k}{\gamma_0} \tag{2.4}$$

while the sample autocorrelation function (SACF) is given by:

$$r_k = \frac{\sum_{t=k+1}^{n} (X_t - \overline{X})(X_{t-k} - \overline{X})}{\sum_{t=1}^{n} (X_t - \overline{X})^2}, \quad k = 0,1,2,\ldots \tag{2.5}$$

But, for uncorrelated observations, the variance of $r_k$ is given by:

$$\text{var}(r_k) \approx \frac{1}{n} \tag{2.6}$$

We also refer to a discrete parameter random variable as a Markov process. Hence, a time series $X_t$ is referred to as a stochastic process if it is generated in time with respect to some probabilistic laws and the future values are only partially predicted by the past values. However, in situations where predictions can be done exactly, we refer to it as deterministic.

## 2.3 Stationary process

Stationarity in time series is a concept which describes a situation in which statistical properties such as mean, autocorrelation, variance, and so on remain the same over a long period of time. So, when the statistical properties of a process do not change over time, it is said to be. It may be difficult to tell if a model is stationary or not by mere looking at its graph. However, a hypothesis test can be enployed. This includes:

i. Test of unit root tests.
ii. KPSS test (a complement to the unit root tests).
iii. A run sequence plot,
iv. The Priestley-Subba Rao (PSR) Test or Wavelet-Based Test.

A stationary process with mean reverting property fluctuates around the mean, which acts as an attractor and crosses the mean line in a large number of ways. Stationarity can be of two types namely Strict stationarity or Weak stationarity.

### 2.3.1 Strict stationarity

A stochastic process $X_t$ satisfies the strict stationarity condition if for any positive integer $n$ at any points $t_1, t_2, \ldots, t_n$ and $n \in \mathbb{Z}$, the joint distribution of $\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\}$ is equivalent to the joint probability distribution of $\{X_{t_1+k}, X_{t_2+k}, ..., X_{t_n+k}\}$. The process is referred to as a stationary process of order $m$, if $\forall\ n \in \mathbb{Z}^+$ and any points $t_1, t_2, ..., t_n$ and $n \in \mathbb{Z}$, then the joint moments (up to order $m$) of $\{X_{t_1}, X_{t_2}, ..., X_{t_n}\}$ is equivalent to the joint moments of $\{X_{t_1+k}, X_{t_2+k}, ..., X_{t_n+k}\}$ which implies that:

$$E\{(X_{t_1})^{k_1}(X_{t_2})^{k_2}...(X_{t_n})^{k_n}\} = E\{(X_{t_1+h})^{k_1}(X_{t_2+h})^{k_2}...(X_{t_n+h})^{k_n}\}\ \forall k_1, k_2, ..., k_n \in \mathbb{Z}^+$$

such that $k_1 + k_2 + ... + k_n \leq m$.

One of the major assumptions upon which most forecasting techniques in statistics are based is that the series can be approximated to attain stationarity (i.e., "stationarized") after mathematical transformations. If a series is not stationary, we can render it stationary via differencing. In otherwords, given the series $X_t$, we can obtain a new set of series:

$$Y_t = X_t - X_{t-1} \tag{2.7}$$

The differenced series $Y_t$ then contains one less point than the original one. It should be noted that this can be done more than once.

### 2.3.2 Weak stationarity

If the first and second moments of a series do not change over time, then it is said to be weakly stationary. That is:

$$E(X_t) = E(X_{t-1}) = \mu\ \forall T$$

$$Var(X_t) = \gamma_0 < \infty$$

$$Cov(X_t, X_{t-k}) = \gamma_k$$

This type of stationarity is also referred to as covariance stationary

## 2.4 Autocovariance and autocorrelation function

In Statistics, the autcovariance function expresses the covariance of a stochastic process with itself at pairs of time intervals. Mathematically;

$$\gamma_k = \text{cov}(X_t, X_{t+k}) \tag{2.8}$$

$$= E\{(X_t - \mu)(X_{t+k} - \mu)\} \qquad k = 0,1,2,...$$

where $\mu$ is the expected value of $X_t$ for all t.

Autocorrelation or serial correlation is defined as the dependence or association in any statistical relationship that relays information about the behaviour of a phenomenon with a delayed copy of itself as a function of delay. It helps in capturing re-occuring patterns or detecting missing fundamental frequency in a signal due to its harmonic frequencies. It also helps in showing relationship between the values of time series process at two different times in relation to the times or the time lag. Given the set of autocorrelation coefficients described by following autocovariance matrix:

$$\gamma_k = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-2} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_0 \end{pmatrix} \qquad (2.9)$$

The autocorrelation function (ACF) matrix is given as:

$$\rho_k = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{pmatrix} \quad where\ (-1 \le \rho_i \le +1) \quad i = 1,2,...,k-1 \qquad (2.10)$$

The partial autocorrelation function (PACF) is defined as the degree of relationship between the random variables of a time series process with its own lagged values. It is different from the autocorrelation function in that the autocorrelation function does not control for other lags. Given two series $Y_t$ and $Y_{t-k}$, we refer to the value $Y_t$ , $Y_{t-k}$ as the partial autocorrelation function (PACF).

We define the Toeplits matrix as:

$$\rho_k = \begin{pmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{pmatrix} \qquad (2.11)$$

Then, the partial autocorrelation $\phi_{kk}$ is given by:

$$\phi_{kk} = \frac{\rho_k^*}{\rho_k} \tag{2.12}$$

where $\rho_k^*$ is $\rho_k$ with the last column substituted by $(\rho_1, \rho_2, ..., \rho_k)$

## 2.5 Methods of estimating autocovariance function

We can estimate the covariance functions using any of the following methods.

### 2.5.1 Maximum likelihood estimator.

The estimate of the autocovariance $\gamma_k$ using the maximum likelihood method is;

$$C_k = \frac{1}{N}\sum_{t=1}^{N-k}(X_t - \bar{X})(X_{t+k} - \bar{X}) \qquad k = 0,1,2,...N-1 \tag{2.13}$$

For the mean deleted data;

$$C_k = \frac{1}{N}\sum_{t=1}^{N-k} X_t^* X_{t+k}^* .$$

In real analysis, $k \le \frac{N}{4}$ and $C_k$ give a good estimate of $\gamma_k$ since $C_k$ is an unbiased estimate of $\gamma_k$ as $N \to \infty$, $E(C_k) = (1 - \frac{k}{N})[\gamma_k + \text{var}(\bar{x})]$, $C_k$ is positive definite.

<u>Alternative Estimator</u>

Another method of estimating the ACF is given by:

$$\bar{C}_k = \frac{1}{N-k}\sum_{t=1}^{N-k}(X_t - \bar{X})(X_{t+k} - \bar{X}) \tag{2.14}$$

This has a smaller bias than $C_k$ but in general, $C_k$ has a smaller mean square error. However, $\bar{C}_k$ is not often used in modeling as it is semi-positive definite.

### 2.5.2 Jackknife Estimator

The Jackknife estimator, originally proposed by Quenouille (1949) is a procedure for correcting bias. It was later redefined and given its current name by John Tukey in 1958 and described its use in constructing confidence limits for a large class of estimators. To estimate $\gamma_k$ , the series is split into two equal parts and the sample autocovariance function (SACF) estimated for each part as well as for the entire series. Then,

$$\breve{C}_k = 2C_k - \frac{1}{2}(C_{k_1} + C_{k_2}) \tag{2.15}$$

where, $\tilde{C}_k$ is the SACF for the entire series

$C_{k_1}$ is the SACF for the first part of the series

$C_{k_2}$ is the SACF for the second half series.

## 2.6 The general linear process (GLP)

This is a class of linear models known as "linear time series models" usually designed to model the dynamic behavior of a series. They include the Autoregressive (AR), Moving Average (MA) and the Autoregressive Moving Average (ARMA) models. Several authors such as Hannan (1980), Box and Jenkins (1970), and so on have considered them in their study according to literature. A time series $\{X_t\}$ is said to be a general linear process (g.l.p) if;

$$X_t = \sum_{j=0}^{\infty} \Psi_j \in_{t-j} \ (\Psi_0 = 1) \tag{2.16}$$

where $\{\in_t\}$ is a white noise process with $E(\in_t) = 0$, $E(\in_t^2) = \sigma^2 \ \forall t$ and $E(\in_t \in_j) = 0$ $\forall t \neq j$ . (2.16) can be written as:

$$\in_t = X_t - \sum_{j=1}^{\infty} \Psi_j \in_{t-j} \tag{2.17}$$

$$\Rightarrow \in_{t-1} = X_{t-1} - \Psi_1 \in_{t-1} - \Psi_2 \in_{t-2} - \Psi_3 \in_{t-3} - \ldots \tag{2.18}$$

From (2.16) and (2.17), we have;

$$\in_t = X_t - \psi_1 \in_{t-1} - (\psi_2 - \psi_1^2) \in_{t-2} - \xi \in_{t-3} - \ldots$$

$$\in_t = X_t + \pi_1 X_{t-1} + \pi_2 X_{t-2} + \pi_3 X_{t-3} + \ldots$$

Hence, we can write the g.l.p as:

$$X_t + \pi_1 X_{t-1} + \pi_2 X_{t-2} + \pi_3 X_{t-3} + \ldots = \in_t$$

$$\Rightarrow \in_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} \qquad (2.19)$$

Using the backward shift operator; $B^k X_t = X_{t-k}$, equation (2.16) becomes;

$$X_t = \sum_{j=0}^{\infty} \psi_j B^j \in_t = (\sum_{j=0}^{\infty} \psi_j B^j) \in_t$$

$$= \Psi(B) \in_t \qquad (2.20)$$

From (2.19), $\in_t = (\sum_{j=0}^{\infty} \pi_j B^j) X_t \Rightarrow \in_t = \pi(B) X_t$

Multiplying (2.20) by $\pi(B)$, we have;

$$\pi(B) X_t = \pi(B)\Psi(B) \in_t$$

$$\in_t = \pi(B)\Psi(B) \in_t$$

$$\Rightarrow \pi(B)\Psi(B) \in_t = 1 \qquad (2.21)$$

$\Rightarrow \pi(B) = \dfrac{1}{\Psi(B)}$ and this has established relationship between $\Psi(B)$ and $\pi(B)$.

## 2.7 Autoregressive process AR($p$)

A stochastic process $X_t$ is said to be an autoregressive model of order $p$, AR($p$) where specific lagged values of $X_t$ are used as predictor variables if it satisfies the relation:

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \ldots + \phi_p X_{t-p} + e_t \qquad (2.22)$$

where $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$ are past values (lags) and $e_t$ is a white noise.

$$E\{X_t\} = \mu$$

Using the backward shift operator: $B^k X_t = X_{t-k}$, equ ation 2.22 is written as:

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p) X_t = e_t \qquad (2.23)$$

$$\Rightarrow \phi(B) X_t = e_t$$

where:

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

and $\phi(B) = 0$ is the characteristic equation.

The Autoregressive model predicts future behaviour based on past behaviour**.** It can be used for forecasting when there exists some relationship between values in a series and the preceeding and succeeding values.

### 2.7.1 Estimation of Autoregressive model AR ($p$).

Given the autoregressive process of order 1, AR (1):

$$X_t = \varphi X_{t-1} + e_t \tag{2.24}$$

In lag-operator form, we have:

$$(1 - \varphi B) X_t = e_t \tag{2.25}$$

and the characteristic polynomial;

$$\Phi(B) = (1 - \varphi B)$$

If $\Phi(B) = (1 - \varphi B) = 0$, the only characteristic root is:

$$\beta = \frac{1}{\varphi}, \quad \varphi \neq 0 \tag{2.26}$$

The AR(1) process is regarded as a stationary process if and only if $|\varphi| < 1$ or $-1 < \varphi < 1$. When $\varphi = 1$, this corresponds to a non-stationary explosive random walk process having a zero drift, $X_t = X_{t-1} + e_t$. If AR (1) equation recursively applied, the random walk process becomes:

$$X_t = e_t + e_{t-1} + e_{t-2} + \ldots \tag{2.27}$$

Then,

$$\mathrm{Var}(X_t) = \sum_{t=0}^{\infty} \sigma^2 = \infty \tag{2.28}$$

For AR (2) process $X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + e_t$, the characteristic polynomial is given by:

$$\Phi(B) = (1 - \varphi_1 B - \varphi_2 B^2)$$

Then the solutions to $\Phi(B) = 0$ are:

$$\alpha_1 = \frac{\varphi_1 + \sqrt{\varphi_1^2 + 4\varphi_2}}{2} \text{ and } \alpha_2 = \frac{\varphi_1 - \sqrt{\varphi_1^2 + 4\varphi_2}}{2}$$

The AR(2) is stationary if and only if $\|\alpha_1\| < 1$ and $\|\alpha_2\| < 1$

$$\Rightarrow \|\alpha_1\alpha_2\| = |\varphi_2| < 1 \text{ and } \|\alpha_1 + \alpha_2\| = |\varphi_1| < 2$$

$$\Rightarrow -1 < \varphi_2 < 1 \text{ and } -2 < \varphi_1 < 2.$$

For real values of $\alpha_1$ and $\alpha_2$, $\varphi_1^2 + 4\varphi_2 \geq 0$

$$\Rightarrow -1 < \alpha_2 \leq \alpha_1 < 1 \text{ and } \varphi_1 + \varphi_2 < 1;\ \varphi_2 - \varphi_1 < 1.$$

In general, for autoregressive model of order $p$,

$$\Phi(B)X_t = e_t$$

where $\Phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p$

$$\Rightarrow X_t = \varphi^{-1}(B)e_t = \psi(B)e_t.$$

Then, by algebraic factorization, we can factorise;

$$\Phi(B) = (1 - G_1 B)(1 - G_2 B)\ldots(1 - G_p B)$$

Such that;

$$\psi(B) = \Phi^{-1}(B) = (1 - G_1 B)^{-1}(1 - G_2 B)^{-1}\ldots(1 - G_p B)^{-1}$$

By partial fractions we have;

$$\psi(B) = \frac{\sum_{i=1}^{p} k_i}{(1 - G_i B)} \tag{2.29}$$

$$= k_1(1 - G_1 B)^{-1} + k_2(1 - G_2 B)^{-1} + \ldots + k_p(1 - G_p B)^{-1}$$

For convergence of $\psi(B) = \Phi^{-1}(B)$, we must have $|B| \leq 1$ and for stationarity $\Phi(B) = 0$ must lie outside the unit circle.

## 2.8 The Moving Average MA ($q$) process

A stochastic process $X_t$ is said to be a Moving Average process of order $q$, denoted by MA ($q$), if it satisfies the difference equation:

$$X_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} + e_t \text{, for every } t \in \mathbb{Z} \tag{2.30}$$

where: $\theta_1, \theta_2 \ldots, \theta_q$ are constants and $e_t$ is a purely random process with mean zero and variance $\sigma^2$.

Similarly, by the backward shift operator, then;

$$X_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q) e_t \tag{2.31}$$

$$\Rightarrow X_t = \theta(B) e_t$$

where: $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$.

The characteristic equation of (2.30) is $\theta(B) = 0$. It can be shown that for any values of $X_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} + e_t$, MA(q) is stationary. Therefore, no stationarity condition is required. But for the expression $e_t = \theta^{-1}(B) X_t$, we say that;

$$\theta^{-1}(B) = 1 - \theta_1 B - \ldots - \theta_q B^q)^{-1} = (1 - P_1 B)^{-1} (1 - P_2 B)^{-1} \ldots (1 - P_q B)^{-1}$$

and the partial fraction;

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

$$= \sum_{i=1}^{q} T_i (1 - P_i B)$$

Hence, for the convergence of $K(\theta) = \theta^{-1}(B)$, we must have $|B| \leq 1$.

$$\Rightarrow |P_i| < 1 \quad \forall \; i = 1, 2, \ldots, q .$$

where $P_i^{-1} (i = 1, 2, \ldots, q)$ are the roots of the characteristic equation.

Hence, for the invertibility condition of the MA($q$), all the roots of $\theta(B)$ must lie outside the unit circle.

## 2.9 The Autoregressive Moving Average (ARMA) process

Independent Autoregressive AR ($p$) and the Moving Average MA($q$) models are sometimes unrealistic by themselves. They can however be combined to form the extremely useful ARMA $(p,q)$ models. Therefore, the ARMA $(p,q)$ is a stochastic process satisfying the relation:

$$X_t = [\phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p}] + [e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + ... + \theta_q e_{t-q}] \tag{2.32}$$

Alternatively, we can write (2.32) as:

$$\phi(B) X_t = \theta(B) e_t \tag{2.33}$$

This implies that $\{X_t\}$ is the sum of an autoregressive and the moving average processes. Given $e_t$ and the $p$ starting values of $\{X_t\}$, the whole series can be iteratively formed and the series will be stationary given that the autoregressive part is stationary.

Another practical importance of the mixed ARMA $(p,q)$ models is the fact that when two independent ARMA processes are summed up, the resulting process is also ARMA. That is, if $X_t$ and $Y_t$ are two independent ARMA series given by:

$$X_t = ARMA(p_1, q_1) \quad , \quad Y_t = ARMA(p_2, q_2)$$

And their sum:

$$Z_t = X_t + Y_t = ARMA(p_3, q_3)$$

$$\text{where: } p_3 = p_1 + p_2 \text{ and } q_3 = \max(p_1 + q_2,\ p_2 + q_1)$$

This implies that mixed models are much more applicable to real-world series than the individual AR and MA models.

Forecasting for the ARMA models

For one-step forecast situation, we have:

$$X_{n+1} = \phi_1 X_n + \phi_2 X_{n-1} + ... + \phi_p X_{n-p+1} + e_{n+1} + \theta_1 e_n + \theta_2 e_{n-1} + ... + \theta_q e_{n-q+1} \tag{2.34}$$

The optimal forecast is then given by:

$$f_{n,1} = \phi_1 X_n + \phi_2 X_{n-1} + ... + \phi_p X_{n-p+1} + e_{n+1} + \theta_1 e_n + \theta_2 e_{n-1} + ... + \theta_q e_{n-q+1} \tag{2.35}$$

A step forecast error given by:

$$\varepsilon_{n,1} = X_{n+1} - f_{n,1} = e_{n+1} \tag{2.36}$$

Then to obtain the optimal forecast, we use $\{e_t\}_{t=n-q+1}^{n}$ which can be estimated by starting with $f_{0,1} = 0$ and then forming the $e_t$ recursively by:

$$e_t = X_t - f_{t-1,1} \qquad t = 1, 2, ..., n$$

### 2.9.1 Estimation of Autoregressive Moving Average ARMA ($p, q$) process

We would recall that the AR process admits an MA$(\infty)$ structure, however they do impose restrictions on the decay structures of the coefficients $\varphi_i$. Similarly, the MA requires some finite terms, but never impose any conditions to restrict the coefficients. According to autocorrelation pattern, the AR processes do allow many non-zero coefficients, but usually with a fixed decay pattern, the MA permits a few coefficients appart from zero with arbitrary values. According to Anderson (1977), the ARMA processes do combine the properties, then make it possible to denote them in a reduced form with fewer number of parameters.

The simplest ARMA process, the ARMA(1,1) is written as:

$$X_t = \varphi_1 X_{t-1} + e_t - \theta_1 e_{t-1}$$

Writing this in the backward operator form, we have:

$$(1 - \varphi_1 B) X_t = (1 - \theta_1 B) e_t \tag{2.37}$$

where $|\varphi_1| < 1$ for the process to be stationary and $|\theta_1| < 1$ for it to be invertible.

Also, we assume that $\varphi_1 \neq \theta_1$. If the two parameters are identical, multiplying (2.37) by $(1 - \varphi_1 B)^{-1}$, we have a white noise process;

$$X_t = e_t$$

Furthermore, the Autocorrelation Function (ACF) of the ARMA (1,1) can be obtained equation (2.37) is multiplied by $X_{t-k}$ and expectations are taken on both sides. Then;

$$\gamma_k = \varphi_1 \gamma_{k-1} + E(e_t X_{t-k}) - \theta_1 E(e_{t-1} X_{t-k}) \tag{2.38}$$

If $k$ >1, $e_t$ becomes uncorrelated with the history of the series. Hence,

$$\gamma_k = \varphi_1 \gamma_{k-1} \tag{2.39}$$

If $k = 0$, $E[e_t X_t] = \sigma^2$

and $E[e_{t-1} X_t] = E[e_{t-1}(\varphi_1 X_{t-1} + e_t - \theta_1 e_{t-1})] = \sigma^2(\varphi_1 - \theta_1)$

Substituting these into (2.37), for $k = 0$, then;

$$\gamma_0 = \varphi_1 \gamma_1 + \sigma^2 - \theta_1 \sigma^2 (\varphi_1 - \theta_1) \tag{2.40}$$

Taking $k = 1$ in (2.38);

$$E[e_t X_{t-1}] = 0,\ E(e_{t-1} X_{t-1}) = \sigma^2$$

$$\text{and } \gamma_1 = \varphi_1 \gamma_0 - \theta_1 \sigma^2 \tag{2.41}$$

solving (2.40) and (2.41) simultaneously, we have;

$$\gamma_0 = \sigma^2 \frac{1 - 2\varphi_1\theta_1 + \theta_1^2}{1 - \varphi_1^2} \tag{2.42}$$

Dividing (2.41) by the expression above, we have the first autocorrelation coefficient;

$$\rho_1 = \frac{(\varphi_1 - \theta_1)(1 - \varphi_1\theta_1)}{1 - 2\varphi_1\theta_1 + \theta_1^2}$$

It can be recalled that if $\varphi_1 = \theta_1$, the autocorrelation becomes zero, then the operators $(1 - \varphi_1 B)$ and $(1 - \theta_1 B)$ cancel out and this results in a white noise process.

The rest of the autocorrelation coefficients is obtained by dividing (2.39) by $\gamma_0$ which gives:

$$\rho_k = \varphi_1 \rho_{k-1}, \quad k > 1 \tag{2.43}$$

This means that from the first coefficient, the ACF of an ARMA (1,1) decays exponentially, depending on the value of $\varphi_1$ of the AR part.

## 2.10 The Autoregressive Integrated Moving Average, ARIMA ($p, d, q$) process

The generalized ARMA model also known as the Autoregressive Integrated Moving Average (ARIMA) model is often employed in some cases of non-stationarity of the data. Both are usually fitted in time series analysis either for better understanding of the series or in order to forecast future values. The Autoregressive (AR) part of ARIMA indicates that there is regression of the variable of interest on its own lagged, prior values while the Moving Average (MA) part shows that the regression error is a linear combination of error terms whose values occurred simultaneously with some other events in the past. The integrated part (denoted by “I") shows that the original values are substituted for by the difference between present and previous values. This is sometimes done more than once before stationarity is acchieved.

Non-seasonal ARIMA models are usually denoted by ARIMA (*p*,*d*,*q*), where *p* is the order of the autoregressive model, *d* is the number of times differencing is carried out before stationarity is achived and *q* is the moving average process order. By general definition we have;

$$\psi(B)X_t = \phi(B)\nabla^d X_t = \theta(B)e_t \qquad (2.44)$$

where:

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$ is the autoregressive operator.

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q$ is the moving average operator.

$\psi(B) = \nabla^d \phi(B)$ is the genearalised autoregressive non-stationary operator.

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

$$\Rightarrow \psi(B) = \phi(B)(1-B)^d = 1 - \psi_1 B - \psi_2 B^2 - \psi_{p+d} B^{p+d}$$

Then (2.44) may be written as:

$$X_t = \psi_1 X_{t-1} + \ldots + \psi_{p+d} X_{t-p-d} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q} + e_t \qquad (2.45)$$

## 2.11 Model Identification.

After examining the stationarity and seasonality status of a model, then we consider the identification of the order of *p* and *q*. This is usually done in time domain model using the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) plots. The sample autocorrelation function plots as well as the sample partial autocorrelation plots are then compared with the theoretical pattern of the ACF and PACF plots after identifying the order.

The ACF plot of the moving average process usually cut off after a particular lag q, while the PACF of the autoregressive model cuts off at a particular lag p and its ACF decays exponentially to zero. However, for the ARMA $(p, q)$ model, neither the ACF nor the PACF cuts off at $p$ and $q$. So, to determine the cut-off points, we compare the coefficients of the ACF and the PACF with the value $\pm\frac{2}{\sqrt{n}}$ such that any coefficient greater or less than this critical value is significantly greater than zero.

Usually, since the sample autocorrelation and PACF are random variables which do not give similar picture as that of the theoretical function, model identification usually becomes more difficult especially in mixed models. Hence, in recent years information-based techniques which would help in automating the model identification process are used. These include the Final Prediction Error (FPE), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and others.

From prediction error point of view, the model becomes better as the order increases since degrees of freedom of the model increases except for the fact that there will be need for more computation time and memory for higher orders. The parsimony principle is to choose the model with the smallest degree of freedom or number of parameters, given that all the models are true representation of the data. Therefore, to examine the order of a model, we would not depend on the prediction error alone but also incorporate a penalty any time the order increases. We can also obtain the plot of the prediction error as a function of the dimension of the model and then obtain the minimum in the curve visually or carry out an F-test to obtain an appropriate value of the model order.

Final Prediction Error (FPE) Criterion

According to Akaike (1970) The FPE is defined as the mean squared prediction error when a model is fitted to a given set of data and is applied to another independent observation, to make a one-step ahead prediction. Based on the final prediction error, the parameters in each model will be estimated for the minimum final prediction error to be attained for the model and then a model with the minimum final prediction error is selected.

For example, given a T-realization $\{Y_1, ...Y_T\}$ of an AR(p) process satisfying the relation:

$$Y_t = \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + v_t \qquad (2.46)$$

and another AR(p) process $X_t$ which is independent of $Y_t$ but having the same statistical structure such that:

$$X_t = \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + u_t \tag{2.47}$$

where the white noise processes $u_t$ and $v_t$ have the same distribution but are independent of each other. Then a one-step predictor of $X_{T+1}$ is given by:

$$\widehat{X}_{T+1} = \widehat{\phi}_{p,1} X_T + \ldots + \widehat{\phi}_{p,p} X_{T-p+1} \tag{2.48}$$

where $\widehat{\phi}_{p,1},\ldots,\widehat{\phi}_{p,p}$ are the estimates obtained by fitting the AR(p) model to the observations $\{Y_1,\ldots Y_T\}$. Since $X_t$ and $\{\widehat{\phi}_{p,1},\ldots,\widehat{\phi}_{p,p}\}$ are independent, Akaike (1969) has shown that, for large T, the mean square prediction error is:

$$E(\widehat{X}_{T+1} - X_{T+1})^2 \simeq \sigma^2 (1 + \frac{p}{T}) \tag{2.49}$$

$$\Rightarrow E(\widehat{\sigma}_p^2) \simeq (1 - \frac{p}{T})\sigma^2 \tag{2.50}$$

for large T.

Hence, $(1 - \frac{p}{T})^{-1} \widehat{\sigma}_p^2$ is a less biased estimate of $\sigma^2$ than $\widehat{\sigma}_p^2$.

The corresponding estimate of the asymptotic mean square error of $\widehat{X}_{T+1}$ is given by:

$$(1 + \frac{p}{T})(1 - \frac{p}{T})^{-1} \widehat{\sigma}_p^2$$

which is asymptotically equivalent to;

$$(1 + 2\frac{p}{T})\widehat{\sigma}_p^2 .$$

Hence, the final prediction error (FPE) is given by:

$$FPE(p) = (1 + 2\frac{p}{T})\widehat{\sigma}_p^2 \tag{2.51}$$

Which is an unbiased and consistent estimator of the mean square error of the one-step ahead predictor; $\widehat{X}_{T+1}$.

If the mean of the process is unknown, accoding to Akaike (1970a), the final prediction error (FPE) becomes:

$$FPE(p) = \hat{\sigma}_p^2 (1 + 2\frac{p+1}{T}) \tag{2.52}$$

## 2.12 Diagnostic checking

This is a procedure by which a specified model is examined and verified in order to ascertain its correctness through the nature of its residuals $(e_t)$. According to Box and Jenkins (1976), this can be done by observing the autocorrelation function plots of the residuals so as to detect whether large correlation values exist. If all the autocorrelations and partial autocorrelations values are small, then the model is regarded adequate and can then be used for forecasting. If there are some large autocorrelation values, $p$ and/or $q$ would be adjusted and the model re-estimated. Therefore, every necessary information from the data would be obtained through the model. In otherwords, the residuals must be small, leaving no systematic or predictable patterns in them.

If the residuals of the specified model resemble a white noise process, then it is correctly specified. In general, we expect that the residuals would be almost uncorrelated with each other for their sample autocorrelation function to be $\rho_k(e)$ and distributed approximately normally about zero and variance $N^{-1}$.

Box and Jenkins proposed a diagonistic method such that the first $k$ autocorrelations $\rho_e(k), k \in \mathbb{Z}^+$ are observed under the null hypothesis ($H_0$):

$$\rho_{\hat{e}}(1) = \rho_{\hat{e}}(2) = ... = \rho_{\hat{e}}(K) = 0$$

with test statistic;

$$Q = N \sum_{i=1}^{K} \hat{\rho}_{\hat{e}}^2(k) \quad \sim \chi^2(k - p - d)$$

where $\hat{\rho}_{\hat{e}}^{2}(k)$ is the k$^{th}$ order sample autocorrelation of the estimated residuals, N is the sample size, and K is chosen such that it is sufficiently large. Then we reject $H_0$ if the Q is greater than the tabulated critical value.

## 2.13 Forecast in probability models

Forecasting is a technique in time series analysis for predicting future events through a sequence of time by analyzing the trends of the past based on historical data and assumming that future trends will follow similar pattern. The forecast error is defined as the absolute difference in the actual value and the forecast value for the corresponding period.

Smoothing methods may be employed when a series shows no significant effects of trend, cyclical, or seasonal components. In this situation, the aim is to smooth out the irregular component of the series by applying an averaging technique such as the moving averages method. It is one of the most commoly used smoothing techniques. It is such that the forecast is the average of the last "x" number of observations, where "x" is some suitable number.

Recently in literature, Artificial Neural Networks (ANNs) have also been employed in time series forecasting. Although they can be used to model linear and nonlinear phenomena, they could not model both structures equally well. Hence, the hybrid methodology integrating ARIMA and ANN models were introduced by Cagdas *et al.*, (2009). They proposed a different approach combining Elman's Recurrent Neural Networks (ERNN) with the ARIMA models. This hybrid approach was then applied to the Canadian Lynx data and was found to have optimum forecasting accuracy.

Ahmed *et al.* (2017) demonstrated how the inefficiency of the conventional neural networks in identifying the behavior and pattern of nonlinear or dynamic series with moving average components, thereby resulting in low forecasting capability. According to them, this gingered the idea that leads to the development of new models such as the Deep Learning neural networks with or without hybrid methodologies as in Fuzzy Logic.

Benkachcha *et al.* (2015) proposed two other methods of investigating variability in seasonal time series using ANN namely; a multilayer perceptron model and a causal method based on artificial neural networks using the components of decomposed time series as input variables. They discovered that ANNs yielded almost the same accuracy whether the original time series is decomposed or not.

However in most cases, prediction errors are inevitable and almost all forecasting methods have errors in predicted results. One of the commonly used forecasting methods is the least square method. It performs calculations with time series data and has a seasonal trend, with time series calculated data error possibility. Prediction errors will often occur (Ismail and Shabri 2014), these prediction errors can be computed using the Mean Absolute Deviation method and Mean Absolute Percentage Error.

## 2.14 Measures of performance of models

Durig the process of building a proper time series model, one has to consider the principle of model parsimony which says the model with smallest possible number of parameters is always to be selected in order to provide an adequate representation of the underlying time series data (Chatfield 1996). Hence, having a set of suitable models, the simplest one shall be considered, still maintaining an accurate description of inherent properties of the time series. The idea of model parsimony is similar to the famous Occam's razor principle, discussed by Hipel and McLeod (1994). An important aspect of this principle is that when one is faced with a number of competing andadequate explanations, pick the most simple one. It also provides considerable inherent informations when applied to logical analysis. More importantly, the more complicated the model is, the more possibilities will arise for departure from the actual model assumptions. So, the more the increase in the number of model parameters, the more the risk of over-fitting. An over fitted time series model may describe the training data very well, but it may not be suitable for future forecasting. As potential overfitting affects the ability of a model to forecast well, parsimony is often used as a guiding principle to overcome this issue. Hence, it can be said that, while making time series forecasts, keen attention should be given to select the most parsimonious model among all others.

Moreover, we examine the criteria needed to measure the performance of the forecast values in relation to the existing models. These include:

**Mean Absolute Error (MAE)**

The MAE measures the difference between two continuous variables *X* and *Y* explaining similar phenomenon. For example, comparisons of predicted against observed or a particular technique of measurement against another. In a scatter plot of *n* points, with coordinates ($x_i$, $y_i$), the MAE is the average vertical distance between each point and the identity line or the average horizontal distance between each point and the identity line.

The MAE is mathematically written as:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n} \tag{2.53}$$

It can also be expressed as the sum of the Quantity Disagreement (QD) and the Allocation Disagreement (AD). Quantity Disagreement is the absolute value of the Mean Error while Allocation Disagreement is the difference between the MAE and the Quantity Disagreement. The Mean Error is therefore defined as:

$$ME = \frac{\sum_{i=1}^{n} y_i - x_i}{n} \tag{2.54}$$

**Mean Percentage Error (MPE)**

The mean percentage error (MPE) is a measure of the average of percentage errors by which forecasts of a model differ from actual values of the quantity being forecast.

Mathematically,

$$MPE = \frac{100\%}{n}\sum_{i=1}^{n}\frac{A_t - F_t}{A_t}, \quad A_t \neq 0 \tag{2.55}$$

where $A_t$ is the actual value of the quantity, $F_t$ is the forecast, and $n$ is the number of times the variable is forecast.

Due to the fact that actual values are used in the formula, positive and negative forecast errors can cancel each other. Hence, equation (2.55) can be used to measure bias in the forecasts.

**Mean Absolute Percentage Error (MAPE)**

This method of prediction is also referred to as as the mean absolute percentage deviation (MAPD). It is a measure of prediction accuracy of a forecasting method. It expresses the accuracy as a percentage, and is given by:

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|, \quad A_t \neq 0 \tag{2.56}$$

where $A_t$ is the actual value and $F_t$ is the forecast value.

As shown above, the difference between $A_t$ and $F_t$ is divided by the Actual value $A_t$ and the absolute value is summed for every forecasted point with respect to time and then divided by the number of fitted points $n$.

## 2.15 Autoregressive bilinear models

The strictly stationary stochastic process $X_t$ is said to be an autoregressive bilinear time series model if:

$$X_t = \sum_{i=1}^{p}\varphi_i X_{t-i} + \sum_{i=1}^{m}\sum_{j=1}^{n} b_{ij} X_{t-i} e_{t-j} \tag{2.57}$$

where: $\varphi_i$ are the parameters of the linear autoregressive components

$b_{ij}$ are the nonlinear bilinear components.

## 2.16 Autoregressive Moving Average Bilinear models

A strictly stationary stochastic process $X_t$ is said to be an autoregressive moving average bilinear time series model if:

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{j=1}^{q} \theta_j e_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} X_{t-i} e_{t-j} \tag{2.58}$$

where: $\varphi_i$ are the parameters of the linear autoregressive components

$\theta_j$ are the parameters of the linear moving average components

$b_{ij}$ are the nonlinear bilinear components.

## 2.17 Autoregressive Integrated bilinear models

The strictly stationary stochastic process $X_t$ is referred to as an autoregressive integrated bilinear time series model if:

$$\Psi(B) X_t = \psi_p(B) \nabla^d X_t + \sum_{i=1}^{m} \sum_{j=1}^{n} b_{ij} X_{t-i} e_{t-j} \tag{2.59}$$

where:

$\Psi(B) = \psi(B)(1-B)^d = 1 - \psi_1 B - \psi_2 B^2 - .... - \psi_{p+d} B^{p+d}$

The expanded form of (2.59) is written as:

$$X_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + .... + \psi_{p+d} X_{t-p-d} + b_{11} X_{t-1} e_{t-1} + .... + b_{mn} X_{t-m} e_{t-n} + e_t \tag{2.60}$$

where: $\psi_i$ are the parameters of the linear autoregressive integrated components

$b_{ij}$ are the nonlinear bilinear components.

## 2.18 Autoregressive integrated moving average bilinear models

The strictly stationary stochastic process $X_t$ is said to be an autoregressive integrated moving average bilinear time series process if it satisfies the relation:

$$X_t = \psi_1 X_{t-1} + .... + \psi_{p+d} X_{t-p-d} + \theta_1 e_{t-1} + ... + \theta_q e_{t-q} + b_{11} X_{t-1} e_{t-1} + .... + b_{mn} X_{t-m} e_{t-n} + e_t$$

where: $\psi_i$ are the parameters of the linear autoregressive integrated components.

$\theta_j$ are the parameters of the linear moving average components.

$b_{ij}$ are the nonlinear bilinear components.

## 2.19 Related literature on Bilinear Time Series Models

Many researchers for example, Hannan (1962,1970), Box and Jenkins (1970) and many others have actually used the linear time series models to explain the covariance structure in the time series and apply them to different fields of life. In fact, many real-life stationary phenomena can be explained by stationary linear time series models, for example the ARMA $(p, q)$ models (Priestely, 1978, 1980). However, many nonlinear phenomena like frequency modulations (Tong, 1990), limit cycles as well as animal population cycles (Oster and Ipaktchi, 1978) cannot be adequately examined by linear models, and as a result second-order properties, such as covariances and spectra, can no longer adequately characterize the properties of the series . Hence, second-order properties, such as covariances and spectra, can no longer adequately characterize the properties of the series. So, we would need to consider non-linear models in order to obtain a better outcome because for any nonlinear time series, any statistical inferences relating to analysis for nonlinear time series model, would reveal the nonlinear characteristics of the data more excellently than those employing linear approximation.
Usually, nonlinear dynamic equations are used to generate nonlinear time series. They possess features such as time-changing variance, asymmetric cycles, higher-moment structures, thresholds and breaks that linear processes will not be able to examine.

Bilinear models is one of the major class of nonlinear time series models which have been widely discussed by different authors in literature (Mohler, 1973, Ruberti, Isidori and d'Alessandro, 1974 and Subba, 1992). It is a class of time series model, which assumes both linear and non-linear components of autoregressive moving average processes. Its linear part of is the aggregation of autoregressive and moving average processes, while the non-linear part is the product of the two processes. One interesting thing about the bilinear model is that despite the fact that it is nonlinear, it has a structural theory which is similar to those obtained in linear systems (Ruberti, Isidori and d'Alessandro, 1974). Following Granger and Anderson (1978a), Subba (1981), Subba and Gabr (1984), a time series is said to be a bilinear process usually denoted by BL $(p, q, m, k)$ if:

$$X_t + \sum_{i=1}^{p} a_i X_{t-i} = \sum_{j=1}^{q} c_j e_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{k} b_{ij} X_{t-j} e_{t-k} + e_t \tag{2.61}$$

where: $\{\varepsilon_t\}$ is a set of *i.i.d* random variables, usually but not always having zero mean and variance $\sigma^2$.

$\phi_i$ are the autoregressive components of order $p$ $\{i = 1,2,..p\}$

$\theta_j$ are the moving average components of order $q$ $\{i = 1,2,..q\}$

$b_{ij}$ are the bilinear components.

Poskitt and Tremayne, (1986) in a paper titled 'The selection and use of linear and bilinear time series models' considered problems usually encountered when linear and bilinear processes are used to model time series phenomena. The paper gives some explanation on reasons why bilinear models may prove to be useful for modelling non-Gaussian time series. Explanation was also given on the effects of adopting a Bayesian stance in the use of model selection criteria when addressing the issue of determining suitable parameterizations. They illustrated the idea using a familiar data set and their ramifications for forecasting are explored. It showed clearly that the one-step ahead forecasts with the minimum mean squared error are generated by combining the two classes of models under consideration.

Pham (1985) introduced an extension of the linear Markovian repsentation called the bilinear Markovian representation, and is shown to provide representations of all-diagonal bilinear time series models. Some of its properties are also given.

Jian Liu and Peter Brockwell (1988) derived the necessary condition for the general bilinear time series equations to be strictly stationary. This was shown to reduce to the conditions given by Pham and Tran (1981) and Bhaskara Rao *et al.* (1983) in the special cases which they considered. Under this specified condition, they constructed a solution which is proven to be causal, stationary and ergodic. It is moreover the unique causal solution and the unique stationary solution of the defining equations. In the special case when the defining equations contain no non-linear terms, our condition reduces to the well-known necessary and sufficient condition for existence of a causal stationary solution.

John and Todd (1991) in a paper titled 'parameter estimation for a bilinear time series model' presented a direct approach to estimating the parameters of a bilinear time series model. They made us realize that the approach depends on the expressions for certain higher-order statistics of the signals that satisfy the bilinear model. According to them, the expressions are usually linear in most of the model parameters and can be evaluated from an over-determined set of equations. Results of an experiment that employs their technique and demonstrates its good properties are also included in the paper.

Cathy (1992) propose a Bayesian approach to bilinear time series analysis as an extension of the study of Broemeling and Shaaraw (1988) on linear time series. They derived both the predictive and marginal posterior distributions of the bilinear parameters, by which they made inferences about the parameters and for a future observation using the conjugate prior for parameters of the model. They illustrated this approach using the Wolf sunspot numbers from Box and Jenkins (1976) and made comparison with a linear time series.

Jian Liu (1992) obtained the proof of the inequality:

$$\rho(A\otimes A+B\otimes B)\geq\rho(A\otimes A),$$

where $\rho(M)$ is the spectral radius of any given square matrix $M$, i.e. max{|eigenvalues| of $M$}, and $M\otimes N$ is the Kronecker product of any two matrices $M$ and $N$. he then used the inequality to show that stationarity of the bilinear model:

$$X_t = Z_t + \sum_{i=1}^{l} \phi_i X_{t-i} + \sum_{j=i}^{r} \theta_j Z_{t-j} + \sum_{j=1}^{q} \sum_{i=1}^{m} b_{ij} X_{t-i} Z_{t-j} \tag{2.62}$$

and this implies the stationarity of the linear ARMA model;

$$Y_t = e_t + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=i}^{r} \theta_j e_{t-j} \text{ for } r = 1 \text{ and } q = 1.$$

In addition, it shows that stationarity of the bilinear model with $b_{ij} = 0$ for $i < j$, also implies stationarity of its linear part, given that the stationarity condition stated by Bhaskara Rao *et.al* is met. The conclusion reached was that stationarity of the subdiagonal bilinear models, implies that the linear component models cannot be extended to the general non-subdiagonal bilinear models. Finally, they demonstrated the last observation using an example where $p = m = 1$, $r = 0$ and $q = 2$.

Subba Rao *et al.* (1992) in a paper titled 'identification of bilinear time series' showed how the Yule-Walker difference equations for higher order moments and cumulants derived for certain types of bilinear time series models BL$(p, 0, p, 1)$ could be used for tentative identification of the model. They used the canonical correlation analysis carried out between the observations and the linear combination of higher powers of the observations for idfentification. They further illustrated the method using real and simulated data.

Abdelouahab Bibi (2006) considered some bilinear time series models having time varying coefficients. Under the study, he aim to examine the structure of usual time series analysis tools, particularly the sample autocovariance function which was developed for analyzing stationary linear time series. He used appropriately defined Markovian representations to obtain a necessary and sufficient condition for the existence of a unique solution with Bounded First and Second order Moments (BFSM). Moreover, he provided a more explicit and sufficient condition for the existence of a BFSM solution and then obtained expression for the autocovariance function. He equally showed the existence of a weak time-varying ARMA representation of the bilinear model with time varying coefficients and that of higher order moments. Several sub-groups of the model were discovered to be quasi-stationary and finally obtained the asymptotic distributions of the sample mean and sample covariances under the assumption.

Sarma and Singh (2007) expressed a bilinear time series model in the form of Akaike's Markovian representation in order to use the Kalman recursive estimation approach. They showed the equivalence of the Akaike's Markovian representation of autoregressive moving average models of order $p$ and $q$ (ARMA($p$,$q$)) and bilinear models. This facilitates the use of the maximum likelihood method of estimation to obtain the parameters bilinear models, which otherwise is an unwieldy problem. They further said that the approach can easily be extended to take into consideration the case of missing observations.

Christian Francq (2007) stated that in general, any purely non-deterministic stationary process $(X_t)$ with finite variance can be represented as an infinite moving average in terms of its innovation process. A property usually employed in the estimation and prediction of linear time series but may give poor results when the innovations are not independent especially in practical situations. The class of models is usually enlarged so as to ensure improvement in the quality of fit and forecasts. In this study he aimed to address the problem using ARMA models with bilinear white noise. He consider the probabilistic and the statistical properties (namely the stationarity, invertibility, moments, prediction, identification, estimation, tests) of this class of models. Then, applicability is examined through simulations and sunspot numbers.

Chikezie, (2007) considered the estimation of the parameters of the nonlinear bilinear seasonal ARIMA time series model which was followed by some numerical illustrations. In this study, it was discovered that the closeness of the estimated values confirmed that the entire procedure of minimizing error (the nonlinear least square method) as well as the Newton-Raphson iterative procedure employed are adequate.

Ojo, (2011) developed the Generalized and subset bilinear time series models which are adequate in achieving stationarity for all non-linear series. These were developed to admit error process of $2^q - 1$ possible subsets with the goal of achieving a better generalized bilinear model. Using the robust nonlinear least-square method and Newton-Raphson iterative method, the parameters of the models are estimated. He then proposed an algorithm for fitting Error Process Included (EPI) generalized and subset generalized bilinear models. He adopted the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in order to obtain the optimum model order. According to him, the EPI made the generalized bilinear model a better model using the residual variance and then fitting models for all possible subsets listed to have the best model is not necessary. Finally, he justified the concept using real life data.

Ojo, (2012) compare new time-series methods for evaluating the death rate of an emerging and re-emerging disease based on One-Dimensional Integrated Autoregressive Bilinear Time Series Model and Generalized Integrated Autoregressive Bilinear Time Series Model. He estimated the parameters of the models using Newton-Raphson iterative method and investigated the statistical properties of the derived estimates. He then proposed an algorithm for fitting the two models. The Akaike Information Criterion (AIC) was used in model order determination while the Residual variance was used to see which model performed better. He then illustrated the approach with real life data and concluded that One-Dimensional Integrated Autoregressive Bilinear Time Series Model performed better than the Generalized Integrated Autoregressive Bilinear Time Series Model in the estimation of death rate of a disease.

Omekara, (2016) studied in details the invertibility, stationarity and covariance nature of pure diagonal bilinear time series models. In his study, he transformed the pure diagonal bilinear time series model into the vector form and then examined the conditions for process to be stationary and invertibile. He also obtained the covariance structure of the pure diagonal bilinear time series model and proved that for every pure diagonal bilinear process there is an ARMA process having identical covariance structure.

Mohammed and Wissam (2016) studied the periodic integer-valued bilinear time series model and dealt with the study of some probabilistic and statistical properties of a periodic integer-valued diagonal bilinear model. They showed the existence of a periodically strict stationary integer-valued process, stated the sufficient conditions for periodical stationarity in the first and second orders, as well as the closed-form of the periodic autocovariance function. They also obtained the closed-forms of the mean and the second moment as well as the Yule-Walker estimations of the underlying parameters and then provided a simulation study.

Anthony (2018) examined two bilinear time series models namely the Bilinear Autoregressive (BAR) and Bilinear Moving Average (BMA) models under certain conditions obtained from the general bilinear autoregressive moving average time series model and then showed that the models exist for $q = Q = 0$, $j = 0$ and $p = P = 0$, $i = 0$, respectively.

## 2.20 Related literature on Seasonal time series models

Seasonality as related to time series is defined as a regular movement or changes that re-occurs at every "$s$" time period, where "$s$" denotes the number of time periods until the pattern is repeated again. For instance, when $s = 12$, we have monthly data, when $s = 4$ we have quarterly data and so on.

Pure seasonal ARIMA time series models on the other hand, are usually denoted by ARIMA $(P, D, Q)_s$, where $P$,$D$ and $Q$ are the seasonal autoregressive, differencing, and moving average terms respectively. The general seasonal ARIMA model, usually denoted by SARIMA $(p, d, q) \times (P, D, Q)_s$, incorporates both non-seasonal and seasonal factors in a multiplicative model. Where;

$p$ is the order of the non-seasonal AR,

$d$ is the non-seasonal differencing,

$q$ is the order of the non-seasonal MA,

$P$ is the order of the seasonal AR,

$D$ is the seasonal differencing,

$Q$ is the order of the seasonal MA,

The model could be written more formally as:

$$\Theta(B^s)\varphi(B)X_t = \Theta(B^s)\theta(B)e_t \quad (2.63)$$

where:

$\varphi(B) = 1 - \varphi_1 B - ... - \varphi_p B^p$ are the non-seasonal AR components

$\theta(B) = 1 + \theta_1 B + ... + \theta_q B^q$ are the non-seasonal MA components.

$\Phi(B^s) = 1 - \Phi_1 B^s - ... - \Phi_P B^{Ps}$ are the seasonal AR components

$\Theta(B^s) = 1 + \Theta_1 B^s + ... + \Theta_Q B^{Qs}$ are the seasonal MA components.

<u>Seasonal Differencing</u>

When $s = 12$ for example, a seasonal difference is given by:

$$(1\text{-}B^{12})X_t = X_t - X_{t-12} \quad (2.64)$$

The differences from the previous year could be as much as that for each month resulting in a stationary series.

Similarly when $s = 4$, for quarterly data, a seasonal difference is given by:

$$(1\text{-}B^4)X_t = X_t - X_{t-4} \quad (2.65)$$

This procedure usually eliminates seasonal trend and seasonal random walk type of nonstationarity.

George *et al.* (1979) revealed that, as a result of the practice-theory iteration extended over many decades and examined by many different investigators, a class of stochastic models capable of representing nonstationary and seasonal time series has evolved. They further investigated and showed that when these models are applied to forecasting and intervention analysis, they have worked well. They equally showed how a fitted model can then determine appropriate techniques for smoothing and seasonal adjustment of a particular time series. Using the shipments-order-inventories series, they illustrated how multivariate models allow the analysis of complex relationships, more accurate forecasts and have the potential for improving smoothing and seasonal adjustment methods.

Billy and Lester, (2003) examined the theory behind the basis for modeling univariate traffic condition data streams as seasonal autoregressive integrated moving average processes. This was based on the Wold decomposition theorem and the assertion that a one-week lagged first seasonal difference applied to discrete interval traffic condition data would result in a weakly stationary transformation. In addition, they presented empirical results using actual intelligent transportation system data and found them to be consistent with the theoretical hypothesis. They concluded by stating the effects of their assertions and findings in relation to the present intelligent transportation systems research, deployment and operations.

Xinghua Chang *et al.* (2012) applied statistical approaches from time series analysis to investigate precipitation trend by presenting a seasonal ARIMA time series model incorporating seasonal characteristics using the monthly precipitation data in Yantai, China for the period of 1961-2011. They built a Seasonal ARIMA $(1, 0, 1)(0, 1, 1)_{12}$ and found that the model is a true representation of the data and modeling of the stochastic seasonal fluctuation was successfully carried out except for some extreme values which were affected by environmental factors like temperature, geographical location and changes in climate. Forecast values based on this model indicated that there would be a decline in precipitation for the next three years.

Wiredu Sampson *et al.* (2013) opined that fluctuations in rainfall pattern would directly or indirectly affect various sectors like agricultural, insurance and other allied fields that play major roles in the development of any economy. In their study, they applied the Seasonal Autoregressive Integrated Moving Average (SARIMA) time series model to rainfall pattern of Navrongo municipality. They used the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion (AICc) and the Bayesian Information Criterion (BIC) in selecting the optimum model and was considered as the best. Seasonal ARIMA $(1, 0, 0)(0, 1, 1)_{12}$ and $(0, 0, 1)(0, 1, 1)_{12}$ were the top two competing models with the least values of AIC, AICc and BIC. The duo was again compared based on their in-sample and out-of-sample forecast performances. ARIMA $(0, 0, 1)(0, 1, 1)_{12}$ performed better than ARIMA $(1, 0, 0)(0, 1, 1)_{12}$ in the forecast performances. Hence, it was identified to be the optimum model for the rainfall data. A global suitability check for the model with the Ljung-Box procedure showed that the model was suitable for forecasting the rainfall data.

Akintunde *et al.* (2017) considered forecast comparison of two independent time series models namely the Seasonal Autoregressive Integrated Moving Average (SARIMA) and the Self Exciting Threshold Autoregressive (SETAR) models. The study was based on the established nonlinearity of the financial and economic time series data as it applies to the rate of inflation in Nigeria. The in-samples and out sample forecast performances for SETAR were compared with that of SARIMA model and SETAR was found to perform better.

Adubisi *et al.* (2017) examined the pattern and growth of money in circulation in Nigeria using the Box-Jenkins procedure involving three modelling stages. They considered the appropriate model which fits the monthly record of Nigeria money in circulation for the period of January 2000 to December 2016 using figures provided by Central Bank of Nigeria. The results showed that the seasonal ARIMA $(2, 1, 0)(0, 1, 1)_{12}$ model was the most suitable model for the series with the smallest information criteria.

Mohammad Valipour (2015) examined the ability of the SARIMA and ARIMA models for long-term runoff forecasting in the United States using the Akaike Information Criterion (AIC) to check the fitted SARIMA and ARIMA models. The MINITAB software was employed to create all necessary stages for determining the parameters involved. He forecasted the amount of runoff for 2011 in each US state using the data from 1901 to 2010 with respect to the average value of all stations in individual state in the first stage. The outcome of the study showed that the SARIMA model was more accurate than the ARIMA model. He also forecasted the amount of runoff for 2001 to 2011 in the second stage using the average annual runoff data from 1901 to 2000. SARIMA model with periodic term equal to 20, $R^2$=0.91, and Mean Bias Error (MBE) = 1.29 mm emerged the optimum model in this stage. Moreover, a trend was observed between annual runoff data in the United States every 20 years or almost a quarter century.

Xiaosheng Li, *et al.* (2013) carried out a research on applications of SARIMA model in forecasting outpatient amount. A time series model forecasting the monthly outpatient amount was built in order to understand the trend of 2002–2012 outpatient amounts and to provide a scientific statistical foundation for the optimization of medical treatment resource allocation. The yearly outpatient data from January 2002 to December 2012 was obtained and SPSS18.0 software was used to verify the outpatient amount in July 2012 to December 2012 in relation to the seasonal autoregressive integrated moving average (SARIMA) model. Results showed that the SARIMA model perfectly fits the variation trend of the outpatient amounts and devoid of external intervention factor, the outpatient amount in the hospital will be on the increase.

## 2.21 Summary

Critical scrutiny of the existing literatures in the previous sections on both linear and nonlinear seasonal time series models reveals that a lot has been done in the examination of the performance of a time series at different length of seasons ($s$). However, the seasonal linear models only reveal the performance of the models at regions before and after the peaks of season while the seasonal nonlinear models do that only at the peaks of season. But, little or no attention has been paid to consideration of a seasonal model that would examine the performance of the series before the peaks are reached, at the peaks and after the peaks at the same time. This poses a course for concern. So, this study considers a nonlinear seasonal model that would enable us to track the behaviour of a seasonal time series in and out of season.

CHAPTER THREE

# METHODOLOGY

## 3.1 Introduction:

Having identified a challenge as a result of the perusal of literatures in chapter two, the development of statistical techniques involved in addressing the challenge is considered a necessity. So, in this chapter we consider the formulation of models and statistical procedures for estimating their parameters as well as testing their validity.

## 3.2 Seasonal Autoregressive (SAR) model.

A stochastic process $X_t$ is regarded as a Seasonal Autoregressive (SAR) model of order $P$, if the following relation is satisfied:

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_P X_{t-Ps} + e_t \qquad (3.1)$$

where:

$\Psi(B) = \Phi(B^s) = 1 - \Psi_1 B - \Psi_2 B^2 .... - \Psi_P B^P$ is the seasonal autoregressive operator.

## 3.3 Seasonal Autoregressive Moving Average (SARMA) model.

A stochastic process $X_t$ is said to be a Seasonal Autoregressive Moving Average model of order $P, Q$ is given by:

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_P X_{t-Ps} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} .... + \Theta_Q X_{t-Qs} + e_t \qquad (3.2)$$

where:

$\Psi(B) = \Phi(B^s) = 1 - \Psi_1 B - \Psi_2 B^2 - .... - \Psi_P B^P$ is the seasonal autoregressive operator.

$\Theta(B) = \phi(B^s) = 1 - \theta_1 B - \theta_2 B^2 .... - \theta_Q B^Q$ is the seasonal Moving-Average operator.

## 3.4 Estimation of SAR and SARMA models

Estimation of seasonal autoregressive (SAR) model

Estimation of the parameters of autoregressive processes can be compared in many ways to the sampling theory approach in univariate regression models. According to Grahn (1995), different methods based on the least squares procedure are commonly used to estimate the parameters of the autoregressive models. With additional assumption of normality, Whittle (1951) also proved that the maximum likelihood estimator is consistent and asymptotically normal. In the absence of normality, the least squares estimators usually have equivalent asymptotic distribution as the maximum likelihood estimators. Given the first-order seasonal autoregressive process of period k which satisfies the stochastic difference equation:

$$Y_t = \sum_{i=0}^{k-1} \delta_{it}\theta_i + \rho Y_{t-k} + e_t, t = k+1, k+2, ..., \quad (3.3)$$

where, $\delta_{it} \begin{array}{l} =1 \quad if \quad i = (t-1) \bmod k \\ =0 \quad otherwise, \end{array}$

and $Y_1, Y_2, ...., Y_k$ are initial conditions and $\{e_t\}$ is a set of independent normal (0, $\sigma^2$) random variables, Walker (1964). The absolute value of parameter $\rho$ is assumed to be strictly less than one. The parameters $\theta_i$ , $i = 0,1,\dots,k-1$, are the seasonal intercepts associated with the various periods. The first-order stationary autoregressive process with nonzero mean is a special case of (3.3) obtained when k = 1.

$\{Y_t\}$ can be expressed in another more useful form by making use of the double subscripts notation $Y_{ij}$, where $t = (i+1) + k(j-l)$ . The variable $Y_{ij}$ represents the sampled value for period $i$ of the $j^{th}$ cycle and satisfies the stochastic difference equation:

$$Y_{ij} = \theta_i + \rho Y_{i,j-1} + e_{ij} \quad (3.4)$$

$j = 2,3,4,\dots$, and $= 0,1,2,\dots,k-1$ . The observations for $ith$ period form a realization from a first-order autoregressive process with a nonzero intercept and parameter $p$ . Consider a realization of $nk$ observations from (3.4), the least squares estimator of $p$ is such that:

$$\hat{\rho}_{OLS} = \frac{[(n-1)k]^{-1}\sum_{i=0}^{k-1}\sum_{j=2}^{n}(Y_{ij}-\bar{Y}_{i0})(Y_{ij,j-1}-\bar{Y}_{i1})}{[(n-1)k]^{-1}\sum_{i=0}^{k-1}\sum_{j=2}^{n}(Y_{i,j-1}-\bar{Y}_{i1})^2} \tag{3.5}$$

where $\bar{Y}_{i0} = (n-1)^{-1}\sum_{j=2}^{n} Y_{ij}$

and $\bar{Y}_{i1} = (n-1)^{-1}\sum_{j=2}^{n} Y_{i,j-1}$, $i = 0,1,2,\ldots,$

Conditioning on the initial k observations, the maximum likelihood estimator of $\rho$ is the same as the least squares estimator of $p$ in (3.5). The estimator $\hat{\rho}_{OLS}$ can lie outside the stationary region $(-1,1)$ even when the observed series is stationary.

Sawa (1978) gave a method for calculating the exact moments of the least squares estimator of the parameter of a stationary autoregressive series of order one but no closed form expressions are given. However, the mean of $\hat{\rho}_{OLS}$ can be obtained exactly if the series is random, i.e., $\rho = 0$ . The result is as shown by the following lemma.

Lemma 3.1. Given $Y_{ij}$, a sequence of independent ($\mu_i, \sigma^2$) random variables, the expectation of $\hat{\rho}_{OLS}$ in (3.5) is;

$$E(\hat{\rho}_{OLS}) = -(n-1)^{-1} \tag{3.6}$$

The least squares estimator has a downward bias in the case of independent observations with unknown means. If we have a random series with known means, the least squares estimator of $\rho$ is given as:

$$\rho^{*}_{OLS} = \frac{\sum_{i=0}^{k-1}\sum_{j=2}^{n}(Y_{ij}-\mu_i)(Y_{i,j-1}-\mu_i)}{\sum_{i=0}^{k-1}\sum_{j=2}^{n}(Y_{i,j-1}-\mu_i)^2} \tag{3.7}$$

The method of obtaining the mean of $\hat{\rho}_{OLS}$ may not be effective in obtaining the mean of $\rho^{*}_{OLS}$, however Sawa's method can be used to achieve that.

Estimation of the seasonal Autoregressive moving average (SARMA) model

The seasonal ARMA model is usually denoted by ARMA $(P,Q)_s$ and of the form:

$$\Phi(B^s)X_t = \Theta(B^s)e_t \tag{3.8}$$

where $\Phi(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \Phi_P B^{Ps}$ ,

and $\Theta(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \ldots + \Theta_Q B^{Qs}$ are the seasonal autoregressive operator and the seasonal moving average operators respectively, with seasonal period of length $s$ . It should be noted that similarly to ARMA $(p, q)$, the ARMA $(P,Q)_s$ model is causal only when the roots of $\Phi(z^s)$ lie outside the unit circle, and satisfies the invertibility condition when the roots of $\Theta(z^s)$ are outside the unit circle. For example, the simplest case of seasonal ARMA$(1, 1)_{12}$, is;

$$(1 - \Phi B^{12})X_t = (1 + \Theta B^{12})e_t,$$

$$\Rightarrow X_t - \Phi X_{t-12} = e_t + \Theta e_{t-12}$$

$$\Rightarrow X_t = \Phi X_{t-12} + e_t + \Theta e_{t-12} \tag{3.9}$$

Comparing (3.9) with ARMA (1,1):

$$X_t = \varphi X_{t-1} + e_t + \theta e_{t-1}$$

implies that the seasonal ARMA presents the series as a form of its past values at lag equal to the length of the period ($s = 12$), while the non-seasonal ARMA presents the series in terms of its past values at lag 1. Just as in the non-seasonal ARMA case, $|\Phi| < 1$ and $|\Theta| < 1$ are the respective required conditions for series to be stationary and invertible.

When $P = 0$, $Q = 1$, $s = 12$ then;

$$X_t = e_t + \Theta_1 e_{t-12}$$

$$\gamma(0) = (1 + \Theta_1^2)\sigma_e^2$$

$$\gamma(12) = \Theta_1 \sigma_e^2,$$

$$\gamma(s) = 0 \text{ for } s = 1, 2, \ldots, 11, 13, 14, \ldots.$$

When $P = 1$, $Q = 0$, $s = 12$, then;

$$X_t = \varphi_1 X_{t-12} + e_t$$

$$\gamma(0) = \frac{\sigma_e^2}{(1-\Phi_1^2)}$$

$$\gamma(12i) = \frac{\sigma_e^2 \Phi_1^i}{1-\Phi_1^2},$$

$$\gamma(s) = 0 \text{ for other } s.$$

## 3.5 Seasonal Autoregressive Integrated (SARI) model

In general, a pure seasonal autoregressive integrated (SARI) model of order $P$ is defined as;

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} .... + \Psi_{P+D} X_{t-Ps-D} + e_t \ ,$$

where $\Psi(B) = \Phi(B^s) = 1 - \Psi_1 B - \Psi_2 B^2 .... - \Psi_P B^P$ is the seasonal autoregressive operator.

## 3.6 Seasonal Autoregressive Integrated Moving Average (SARIMA) model

The seasonal autoregressive integrated moving average (SARIMA) model of order $P$ and $Q$ is generally given by:

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} .... + \Psi_{P+D} X_{t-Ps-D} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + ... + \Theta_Q e_{t-Qs} + e_t$$

where: $\Psi_i$ ,i=1,2,...,P are the seasonal autoregressive components.

$\Theta_i$ ,i=1,2,...,Q are the seasonal moving average components.

## 3.7 Behaviour of ACF and PACF of seasonal linear Time Series Model

Given a multiplicative seasonal ARMA process:

$$\Phi_P(B^s)\varphi_p(B)X_t = \Theta_Q(B^s)\theta_q(B)e_t \tag{3.10}$$

The autocorrelation function (ACF) of the process is a combination of the ACF which corresponds to the nonseasonal and seasonal components. If we represent by $r_i$ the ACF coefficients of the nonseasonal ARMA ( $p,q$ ) process:

$$\varphi_p(B)X_t = \theta_q(B)e_t \tag{3.11}$$

and $R_{si}$, the ACF coefficients in the lags s, 2s, 3s, ... of the seasonal ARMA $(P,Q)$ process:

$$\Phi_P(B^s)X_t = \Theta_Q(B^s)e_t \tag{3.12}$$

and $\rho_j$ the AC coefficients of the combined process, then we have;

$$\rho_j = \frac{r_j + \sum_{i=1}^{\infty} R_{si}(r_{si+j} + r_{si-j})}{1 + 2\sum_{i=1}^{\infty} r_{si}R_{si}} \tag{3.13}$$

Suppose $s = 12$ and $r_j \simeq 0$ for large lag values (such as when j ≥ 8), the denominator of (3.13) gives the unit and the autocorrelation function such that:

(i) In small lags (e.g j = 1, ..., 6) only the regular part is observed, this implies:

$$\rho_j \simeq r_j$$

(ii) In seasonal lags basically the seasonal part is observed. Hence;

$$\rho_{12i} \simeq R_{12i}(r_{24i} + r_0) + R_{24}(r_{36i} + r_{12i}).$$

Also assuming $r_{12i} \simeq 0$ that for i ≥ 1, with $r_0$ = 1 the expression becomes:

$$\rho_{12i} \simeq R_{12i}(i = 1, 2, ...).$$

(iii) Interaction between the regular and seasonal part is observed around the seasonal lags. This is reflected in the re-occurence of the regular part of the autocorrelation function on the two sides of each seasonal lag. In particular, if the regular nonseasonal part is a moving average of order $q$, on both sides of each non-null seasonal lag there will be $q$ coefficients not equal to zero. On the otherhand, if the regular part is an autoregressive process of order $p$, a decrease will be observed as a result of the structure of the AR on the two sides of the seasonal lags.

The partial autocorrelation function of a multiplicative seasonal process is a bit ambigous due to the fact that it is a function of the partial autocorrelation functions of the regular or nonseasonal and seasonal parts (3.11) and (3.12) and also on the sample autocorrelation of the regular part as follows.

(i) In the first lags the PACF of the nonseasonal structure is observed while in the seasonal lags, the PACF of the seasonal structure is seen.

(ii) At the right hand side of each seasonal coefficient (lags $js+1, js+2,...$) the PACF of the regular part will appear. If the seasonal coefficient is positive the regular PACF is inverted. However, if it is negative the PACF takes its own sign.

(iii) On the left hand side of the seasonal coefficients (lags $js-1, js-2$), we notice the ACF of the nonseasonal part.

## 3.8 Nonlinearity Test

Of great importance is the examination of the nonlinearity behaviour of a set of data before fttting a nonlinear time series model. In time domain, different methods have been employed in detecting nonlinearity. For example Subba Rao *et al*. (1980) and Hinnich (1982) used the bispectrum test which is based on the principle that the squared modulus of the normalized bispectrum is constant when the series is linear. The hypothesis uses the non-centrality paremeter $\lambda_i$ of the marginal distributions $\sigma_N^2\chi_2^2(\lambda_i)$ of the squared moduli and also Hjellvik *et al.* (1998).

Granger and Newbold (1976) proved that for a linear and normal time series $X_t$;

$$\rho_k(X_t^2) = \{\rho_k(X_t)\}^2 \tag{3.14}$$

where $\rho_k(.)$ is the lag $k$ autocorrelation. Non-satisfaction of this condition would indicate nonlinearity of some degree, as shown by Granger and Andersen (1978b).

Yasumasa (1997) defined two types of nonlinearity and proposed statistics to detect them as follows. Given { $X_t$ }, a strictly stationary process with zero mean and finite variance. If the spectral distribution function of $X_t$ is absolutely continuous with the density function

$$f \text{ and } \int_{-\pi}^{\pi} \ln f(\lambda)d\lambda > -\infty,$$

then by the Wold decomposition theorem,

$$X_t = \sum_{i=0}^{\infty}\phi_i Z_{t-i} \tag{3.15}$$

where $\left|\phi_i^2\right| < \infty$ and { $Z_t$ } is a set of mutually orthogonal random variables. Moreover;

$$Z_t = \sum_{i=0}^{\infty} \pi_i X_{t-i} \tag{3.16}$$

If $\{ Z_t \}$ is independent and identically distributed random variables, then $\{ X_t \}$ is a linear process.

According to Yajima, (1994), if $X_t$ is a linear process, the following properties are satisfied.

$$E(X_t \big| \zeta_{t-1}) = \sum_{i=1}^{\infty} -\pi_i X_{t-i} \tag{3.17}$$

$$E(Z_t^2 \big| \zeta_{t-1}) = EZ_t^2 = \sigma^2 \tag{3.18}$$

where $\zeta_t$ is a σ-field generated by $\{X_s, s < t\}\}$.

Hence, for the two types of nonlinearity, if (3.17) is not satisfied and the conditional mean of $X_t$ given $X_{t-1}$, $X_{t-2}$,… is some nonlinear function of $X_{t-1}$, $X_{t-2}$,…, then $X_t$ is called a process with nonlinear conditional mean (NCM). If (3.18) is not satisfied and the conditional variance of the innovation process $Z_t$ given $X_{t-1}$, $X_{t-2}$,… depends on the values of $X_{t-1}$, $X_{t-2}$,…. Then, we refer to $X_t$ as a process with heteroscedastic conditional variance (HCV).

Keenan's Nonlinearity Test

This is based on Tukey's (1949) test for non-aditivity to generate a time-domain statistic. It is a test based on the relationship of Volterra expansions with polynomials. That is, a time series $X_t$, $t = 1, 2, \ldots, n$ can be explained by a second-order Volterra expansion of the form:

$$X_t = \mu + \sum_{i=-\infty}^{p} c_i e_{t-i} + \sum_{i,j=-\infty}^{\infty} c_{ij} e_{t-i} e_{t-j} \tag{3.19}$$

which will be linear if the last term on the right-hand side of (3.19) is zero. We should note that the general bilinear time series model is a special case of (3.19). Keenan's test usually regress $X_t$ on $\{1, X_{t-1},..., X_{t-M}\}$, then calculate the fitted values $\{\hat{X}_t\}$ and the residuals $\{\hat{e}_t\}$, $t = M+1, \dots, n$ and its sum of squares $\{\hat{e},\hat{e}\} = \sum \hat{e}_t^2$. It further regresses $\hat{X}_t^2$ on $\{1, X_{t-1},..., X_{t-M}\}$ and calculate the residuals, $\{\hat{\xi}_t\}$ for $t = M+1, \dots, n$ and lastly regresses $\hat{e} = (\hat{e}_{M+1},..., \hat{e}_n)$ on $\hat{\xi} = (\hat{\xi}_{M+1},..., \hat{\xi}_n)$ to obtain $\hat{\eta}$ and $\hat{F}$ via;

$$\hat{\eta} = \hat{\eta}_0 \left( \sum_{t=M+1}^{n} \hat{\xi}_t^2 \right) \quad \text{where } \hat{\eta}_0 \text{ is the regression coefficient} \tag{3.20}$$

$$\text{and } \hat{F}_k = \frac{\hat{\eta}(n-2M-2)}{\{\hat{e},\hat{e}\} - \hat{\eta}^2} \tag{3.21}$$

The F-test

Tsay (1986) later modified Keenan's test to obtain the F-test by substituting the aggregated quantity $\hat{X}_t^2$ with the disaggregated variables $\hat{X}_{t-i}\hat{X}_{t-j}$, $i,j = 1,\dots,M$. The F-test regresses $X_t$ on $\{1, X_{t-1},..., X_{t-M}\}$, and calculates fitted values $\{\hat{X}_t\}$ and the residuals $\{\hat{e}_t\}$, $t = M+1, \dots, n$ with the regression model;

$$X_t = W_t\Phi + e_t \tag{3.22}$$

where $W_t = (1, X_{t-1},..., X_{t-M})$ and $\Phi = (\Phi_0, \Phi_1,..., \Phi_M)^T$

It further regresses the vector $Z_t$ on $W_t$, where $Z_t$ is the multivariate regression model;

$$Z_t = W_t H + \Lambda_t \tag{3.23}$$

where $Z_t$ is an $m = \frac{1}{2}M(M+1)$ dimensional vector given by $Z_t^T = \text{vech } (U_t^T U_t)$ with $U_t = (X_{t-1},..., X_{t-M})$.

Nonlinear least square Method

Consider a group of $m$ data values; $(x_1, y_1), (x_2, y_2), \dots (x_m, y_m)$, and a nonlinear model $y = f(x, \alpha)$, such that $x$ also depends on $n$ parameters, $\alpha = (\alpha_1, \dots, \alpha_n)$, $m \geq n$. The Nonlinear least squares is a type of least squares analysis used to fit some the set of $m$ observations with nonlinear model in $n$ unknown parameters. In order to find the vector $\alpha$ of parameters such that the nonlinear model fits best given the data in the least square sense, that is, the sum of square:

$$S = \sum_{i=1}^{m} e_i^2 \tag{3.24}$$

is minimized, where the in-sample prediction errors $e_i$ are given by:

$$e_i = y_i - f(x_i, \alpha) \qquad i = 1,2,...m$$

Then, the minimum value of S ($S_{\min}$) occurs when the gradient is zero. Since the model contains $n$ parameters, then there exist $n$ gradient equations:

$$\frac{\partial S}{\partial \alpha_j} = 2\sum_i e_i \frac{\partial e_i}{\partial \alpha_j} = 0 \qquad (j = 1, 2, \ldots n) \tag{3.25}$$

where $\frac{\partial e_i}{\partial \alpha_j}$ are functions of $x$ and $\alpha$.

### 3.9 Seasonality test

Several methods have been used in order to detect seasonality in time series analysis. These include:

(i) A run sequence plot.

(ii) A seasonal subseries.

(iii) Multiple box plots which can be used as an alternative method to the seasonal subseries plot.

(iv) The autocorrelation function plot.

However we can categirise them into three major groups, namely the Chi-Square ( $\chi^2$ ) Goodness-of-Fit test and the Kolmogorov-Smirnov type statistic. The Harmonic analyses based on the Edwards' type statistic (Edwards, 1961), and the Nonparametric Tests. The $\chi^2$ goodness-of-fit test is relatively popular due to the fact that its mathematical theoryis simple. The test is based on whether the empirical data can be a sample of a certain distribution with sampling error as the only source of variability (McLaren *et al*. 1994). It uses a sample from a population with an unknown distribution function F(x) and a certain theoretical distribution function $F_0(x)$. Although no restriction is set on the underlying distribution, the hypothetical distribution is usually a uniform distribution. For seasonality studies, the frequency $O_i$, i = 1, 2,…, k is the observed value at the $i^{th}$ season, while the frequency $E_i$, i = 1, 2,…, k is the expected cell frequency at the $i^{th}$ season. Under the null hypothesis that there is no seasonal effect (i.e., $F_0(x)$ is a uniform distribution), then $E_1 = E_2 = \ldots = E_K$ and the statistic:

$$T = \sum_{i=1}^{k} \left[ \frac{(O_i - E_i)^2}{E_i} \right] \tag{3.26}$$

is asymptotically distributed as $\chi^2$ with $\nu = k - 1$ degrees of freedom.

<u>The Hegy Tests</u>

HEGY (1990) presented a factorization of the seasonal differencing polynomial $\Delta_4 \equiv (1 - B^4)$, then provided a procedure for testing seasonal unit roots, which consists the procedures in estimating the following regression via OLS:

$$\Delta_4 X_t = \pi_1 y_{1,t-1} + \pi_2 y_{2,t-1} + \pi_3 y_{3,t-2} + \pi_4 y_{3,t-1} + \varepsilon_t \tag{3.27}$$

where $y_{1t} = (1 + B + B^2 + B^3)X_t$, $y_{2t} = -(1 - B - B^2 - B^3)X_t$ and $y_{3t} = -(1 - B^2)X_t$.

If $X_t$ is a univariate stochastic process given by;

$$X_t = \alpha X_{t-4} + u_t \tag{3.28}$$

where $u_t$, is a stationary process with zero mean and constant variance.

If $\alpha$ = 1 in (3.28), then $y_{1t}, y_{2t}$ and $y_{3t}$ have unit roots only at $\theta$ = 0, $\pi$, and $\frac{\pi}{2}$, respectively. Therefore, the unit root found at $\theta$ = 0 in $X_t$ which implies acceptance of the null hypothesis;

$$H_0: \pi_1 = 0.$$

Similarly, when $\pi_2 = 0$, this implies the existence of a unit root at $\theta = \pi$. When both $\pi_3 = \pi_4 = 0$, these result in complex unit roots at $\theta = \frac{\pi}{2}$.

### 3.10 Pure seasonal autoregressive integrated one-dimensional Bilinear Time Series (PSARIODBL) Model.

Recall that a pure seasonal autoregressive integrated model of order P is generally given by:

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} .... + \Psi_{P+D} X_{t-Ps-D} + e_t \text{ ,}$$

where $\Psi(B) = \Phi(B^s) = 1 - \Psi_1 B - \Psi_2 B^2 .... - \Psi_P B^P$ is the seasonal autoregressive operator.

Hence, we then define the pure seasonal autoregressive integrated one-dimensional bilinear (PSARIODBL) time series model as:

$$\Psi(B) X_t = \Phi_P (B^s) \nabla_s^D \nabla^d X_t + \sum_{i=1}^{m} b_{i1} X_{t-i} e_{t-1} \tag{3.29}$$

where $\Psi(B) = \Phi(B^s)(1 - B^s)^D = 1 - \Psi_1 B - \Psi_2 B^2 .... - \Psi_{P+D} B^{P+D}$ is the seasonal autoregressive integrated operator.

$b_{i1}$ are the nonlinear one-dimensional bilinear components.

The expanded form of (3.29) is written as:

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_{P+D} X_{t-Ps-D} + b_{11} X_{t-1} e_{t-1} + .... + b_{m1} X_{t-m} e_{t-1} + e_t \tag{3.30}$$

#### 3.10.1 Vector form of the pure seasonal autoregressive integrated one-dimensional bilinear time series (PSARIODBL) model specified model:

According to Akaike (1974), the properties of a process are more easily examined when the model represented in the vector form due to the Markovian nature of the model.

Consider:

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_{P+D} X_{t-Ps-D} + b_{11} X_{t-1} e_{t-1} + .... + b_{m1} X_{t-m} e_{t-1} + e_t \tag{3.31}$$

$$\Rightarrow X_{t-1} = \Psi_1 X_{t-s-1} + \Psi_2 X_{t-2s-1} + .... + \Psi_{P+D} X_{t-Ps-D-1} + b_{11} X_{t-2} e_{t-2} + .... + b_{m1} X_{t-m-1} e_{t-2} + e_{t-1}$$

$$X_{t-2} = \Psi_1 X_{t-s-2} + \Psi_2 X_{t-2s-2} + .... + \Psi_{P+D} X_{t-Ps-D-2} + b_{11} X_{t-3} e_{t-3} + .... + b_{m1} X_{t-m-1} e_{t-3} + e_{t-2}$$

$$\vdots$$

$$\begin{aligned} X_{t-p-1} = {} & \Psi_1 X_{t-s-p-1} + \Psi_2 X_{t-2s-p-1} + .... + \Psi_{P+D-p-1} X_{t-Ps-D-p-1} + b_{11} X_{t-p-2} e_{t-p-1} + .... \\ & + b_{m1} X_{t-i-m-1} e_{t-p-2} + e_{t-p-1} \end{aligned}$$

Let: $\Psi = \begin{pmatrix} \Psi_1 & \Psi_2 & \Psi_3 & .... & \Psi_{P+D} & \Psi_{P+D+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix}$, $B_{m\times m} = \begin{pmatrix} b_{11} & b_{21} & b_{31} & .... & b_{m1} \\ 1 & 0 & 0 & .... & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & .... & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$

and the vectors:

$\underset{1\times p}{\mathbf{X}'} = (X_t \quad X_{t-1} \quad X_{t-2} \quad .... \quad X_{t-p-1})$, $\underset{1\times p}{H'} = (1 \quad 0 \quad 0 \quad .... \quad 0)$ and $C' = (1 \quad 0 \quad 0 \quad .... \quad 0)$

where $H'$ represents the transpose of a matrix H, $t = \cdots - 1, 0, 1, ....$ .

Hence, (3.31) in vector form is given by:

$$\mathbf{X}_t = +\Psi\mathbf{X}_{t-s} + \mathrm{B}\mathbf{X}_{\mathrm{t-1}}\mathbf{e}_{\mathrm{t-1}} + \mathrm{C}\mathbf{e}_{\mathrm{t}} \qquad (3.32)$$

$$\Rightarrow X_t = H' \, \mathbf{X}_t \qquad (3.33)$$

### 3.10.2 Stationarity and Convergence of the PSARIODBL

Following Rao *et al.* (1983) and Sangodoyin *et al.* (2010), a sufficient condition necessary for the existence of strictly stationary process and convergence satisfying the pure seasonal one-dimensional bilinear model (3.33) can be achieved through the following theorem.

**Theorem:**

Let $\{e_t, t \in \mathbb{Z}\}$ be a set of independent and identically distributed random variables defined on the probability space $\{\Omega, F, P\}$ such that; $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$. Let $\Psi$ *and* $\mathrm{B}$ be matrices as defined above such that;

$$\rho[(\Psi \otimes \Psi + \sigma^2 \mathrm{B} \otimes \mathrm{B}] = \delta < 1 \qquad (3.34)$$

And **C** be any column vector with components $c_1, c_2, ... c_p$.Then the series of random vectors;

$$\sum_{u\geq 1} \prod_{i=1}^{u} (\Psi + \mathrm{B}e_{t-i}) C e_{t-u}$$

converges absolutely almost surely and also in the mean $\forall \; t \in \mathbb{Z}$. On the other hand, if:

$$X_t = Ce_t + \sum_{u\geq 1}\prod_{i=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}, \quad t \in \mathbb{Z} \tag{3.35}$$

Then for every $t \in \mathbb{Z}, X_t$ is a strictly stationary process satisfying the pure seasonal bilinear model:

$$\mathbf{X}_t = \Psi\mathbf{X}_{t-s} + \mathbf{B}\mathbf{X}_{t\text{-}1}\mathbf{e}_{t\text{-}1} + \mathbf{C}\mathbf{e}_t$$

On the otherhand, if $\{X_t, t \in \mathbb{Z}\}$ is a stationary process satisfying the pure seasonal bilinear relation;

$$\mathbf{X}_t = \Psi\mathbf{X}_{t-s} + \mathrm{B}\mathbf{X}_{t\text{-}1}\mathrm{e}_{t\text{-}1} + \mathrm{C}\mathrm{e}_t$$

$\forall\ t \in \mathbb{Z}$ , for some sequences $\{e_t, t \in \mathbb{Z}\}$ of independent and identically distributed random variables with $E(e_t) = 0$ *and* $E(e_t^2) = \sigma^2 < \infty$ and for the matrices $\Psi, \mathrm{B}$ and $C$ of respective orders $P \times P,\ m \times m$ and $p \times 1,$ such that $\rho[(\Psi \otimes \Psi + \sigma^2\mathrm{B} \otimes \mathrm{B}] = \delta < 1$

Then;

$$X_t = Ce_t + \sum_{u\geq 1}\prod_{i=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}, \ \forall t \in Z \tag{3.36}$$

**Proof:**

The proof shall be established via;

<u>Step 1</u>: For almost sure convergence, we prove that;

$$\sum_{u\geq 1} E\left|\left(\prod_{j=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}\right)_j\right| < \infty \quad \forall \quad i = 1, 2, ..., P \tag{3.37}$$

$\Rightarrow \sum_{u\geq 1}\prod_{i=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}$ is absolutely convergent almost surely and also in the mean.

<u>Step 2:</u>

We establish (3.37) for i=1 and it should be noted that

$\forall\ t \in Z, u \geq 1\ \ \textit{and}\ \ n = 1, 2, ...P$

$$E\left|((\Psi + \mathrm{B}e_{t-i})Ce_{t-u})_n\right| = E\left|\sum_{j=1}^{P}(\Psi)_{nj}C_j e_{t-u} + \sum_{j=1}^{m}(B)_{nj}c_j e_{t-u}^2\right| \leq k_0$$

where $k_0$ is a constant depending on $\Psi, \mathrm{B}, C$ and $\sigma^2$ only.

Step 3:

Similarly if $u \geq 2$, this implies that there exists some $k_1 > 0$ such that:

$$E\left|\left(\prod_{i=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}\right)_1\right| \leq k_1 P\delta^{(\frac{u-1}{2})} \tag{3.38}$$

However, it should be noted that:

$$E\left|\left(\prod_{i=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}\right)_1\right| = E\left|\left[(\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i})(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}\right]_1\right|$$

$$= E\left|\sum_{n=1}^{p}(\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i}))_{1n}((\Psi + \mathrm{B}e_{t-i})Ce_{t-u})_n\right|$$

$$\leq \sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\Psi + \mathrm{B}e_{t-i})Ce_u)_n\right|\right)$$

In the above step, we have considered the fact that $\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i})$ and $(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}$ are independently distributed. Hence by step 2 and application of the Cauchy-Schwartz inequality,

$\sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\Psi + \mathrm{B}e_{t-i})Ce_{t-u})_n\right|\right)$ is not greater than

$$k_0\sum_{n=1}^{p}\left[E\left(\left((\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i})\right)_{1n}\right)^2\right]^{\frac{1}{2}}$$

Now for any n=1,2,…,P

$$\left(\left((\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i})\right)_{1n}\right)^2 = ((\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i})) \otimes (\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i})))_{1n;1n}$$

$$= (\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i}) \otimes (\Psi + \mathrm{B}e_{t-i}))_{1n;1n} \tag{3.39}$$

Hence;

$$E\left(\left(\prod_{i=1}^{u-1}(\Psi + \mathrm{B}e_{t-i})\right)_{1n}\right)^2 = \prod_{i=1}^{u-1}(E(\Psi + \mathrm{B}e_{t-i}) \otimes (\Psi + \mathrm{B}e_{t-i}))_{1n;1n}$$

$$= ((E[(\Psi + \mathrm{B}e_t) \otimes (\Psi + \mathrm{B}e_t)])^{u-1})_{1n;1n}$$

$$= ((E[(\Psi \otimes \Psi + 2\Psi \otimes \mathrm{B}e_t + \mathrm{B} \otimes \mathrm{B}e_t^2)])^{u-1})_{1n;1n}$$

$$= ((\Psi \otimes \Psi + \sigma^2 \mathrm{B} \otimes \mathrm{B})^{u-1})_{1n;1n} \leq k\delta^{u-1} \quad (3.40)$$

Therefore; $E\left|\left(\prod_{i=i}^{u}(\Psi + \mathrm{B}e_{t-j})Ce_{t-i}\right)_1\right| \leq k_1 P\delta^{(\frac{u-1}{2})}$ for any suitable choice of $k_1$.

Step 4:

Since $\delta < 1$,

$$\Rightarrow \sum_{u \geq 1} E\left|\left(\prod_{i=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}\right)_i\right| < \infty \qquad \forall \;\; i = 1, 2, \ldots P$$

Thus (3.37) is satisfied.

It is then obvious that the vector-valued stochastic process $\{X_t, t \in \mathbb{Z}\}$ given by:

$$X_t = Ce_t + \sum_{u \geq 1} \prod_{i=1}^{u}(\Psi + \mathrm{B}e_{t-i})Ce_{t-u}, \;\; t \in \mathbb{Z} \text{ is strictly stationary.}$$

Hence;

$$\rho(\Psi \otimes \Psi + \sigma^2 \mathrm{B} \otimes \mathrm{B}) = \delta < 1$$

is a sufficient condition for the pure seasonal one-dimensional bilinear model (3.36) to be stationary.

So, if we desire a real-valued process $X_t$ conforming to the bilinear model:

$$X_t = \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + \ldots. + \Psi_{P+D} X_{t-Ps-D} + b_{11} X_{t-1} e_{t-1} + \ldots. + b_{m1} X_{t-m} e_{t-1} + e_t$$

$\forall t \in \mathbb{Z}$ under the assumptions on $e_t$, a sufficient condition for its existence is given by:

$$\Psi^2 + \sigma^2 \mathrm{B}^2 = \delta < 1 \quad (3.41)$$

Thus, the proof is established.

## 3.11 Mixed Seasonal Autoregressive Integrated One-dimensional Bilinear (MSARIODBL) time series model.

We then define the mixed seasonal autoregressive integrated One-dimensional Bilinear Time Series Model (MSARIODBL) as follows:

$$\psi(B)X_t = \varphi_p(B)X_t \Phi_P(B^s)\nabla_s^D \nabla^d + \sum_{i=1}^{m} b_{i1} X_{t-i} e_{t-1} \quad (3.42)$$

where: $\psi(B)=\varphi(B)(1-B)^d=1-\psi_1 B-\psi_2 B^2-....-\psi_{p+d}B^{p+d}$, is the non-seasonal autoregressive integrated operator.

$\Psi(B)=\Phi(B^s)(1-B^s)^D=1-\Psi_1 B-\Psi_2 B^2....-\Psi_{P+D}B^{P+D}$ is the seasonal autoregressive integrated operator.

$b_{i1}$ are the nonlinear one-dimensional bilinear components.

The expanded form of equation (3.42) is written as:

$$X_t=\psi_1 X_{t-1}+\psi_2 X_{t-2}+....+\psi_{p+d}X_{t-p-d}+\Psi_1 X_{t-s}+\Psi_2 X_{t-2s}+....+\Psi_{P+D}X_{t-Ps-D}$$
$$+b_{11}X_{t-1}e_{t-1}+....+b_{m1}X_{t-m}e_{t-1}+e_t$$

(3.43)

### 3.11.1 Vector form of the mixed seasonal autoregressive integrated one-dimensional bilinear time series model:

In a similar manner, we shall express mixed seasonal autoregressive integrated one-dimensional bilinear time series in the state space form as in Akaike (1974).

Given:

$$X_t=\psi_1 X_{t-1}+\psi_2 X_{t-2}+....+\psi_{p+d}X_{t-p-d}+\Psi_1 X_{t-s}+\Psi_2 X_{t-2s}+....+\Psi_{P+D}X_{t-Ps-D}$$
$$+b_{11}X_{t-1}e_{t-1}+....+b_{m1}X_{t-m}e_{t-1}+e_t$$

(3.44)

$$\Rightarrow X_{t-1}=\psi_1 X_{t-2}+\psi_2 X_{t-3}+....+\psi_{p+d}X_{t-p-d-1}+\Psi_1 X_{t-s-1}+\Psi_2 X_{t-2s-1}+....$$
$$+\Psi_{P+D}X_{t-Ps-D-1}+b_{11}X_{t-2}e_{t-2}+....+b_{m1}X_{t-m-1}e_{t-2}+e_{t-1}$$

$$X_{t-2}=\psi_1 X_{t-3}+\psi_2 X_{t-4}+....+\psi_{p+d}X_{t-p-d-2}+\Psi_1 X_{t-s-2}+\Psi_2 X_{t-2s-2}+....$$
$$+\Psi_{P+D}X_{t-Ps-D-2}+b_{11}X_{t-3}e_{t-3}+....+b_{m1}X_{t-m-1}e_{t-3}+e_{t-2}$$
$$\vdots$$
$$X_{t-p-1}=\psi_1 X_{t-p-2}+\psi_2 X_{t-p-3}+....+\psi_{p+d}X_{t-2p-d-1}+\Psi_1 X_{t-s-p-1}+\Psi_2 X_{t-2s-p-1}+....$$
$$+\Psi_{P+D-p-1}X_{t-Ps-D-p-1}+b_{11}X_{t-p-2}e_{t-p-1}+....+b_{m1}X_{t-i-m-1}e_{t-p-2}+e_{t-p-1}$$

Let us define the matrices:

$$\psi = \begin{pmatrix} \psi_1 & \psi_2 & \psi_3 & .... & \psi_p & \psi_{p+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix}, \quad \Psi = \begin{pmatrix} \Psi_1 & \Psi_2 & \Psi_3 & .... & \Psi_{P+D} & \Psi_{P+D+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix},$$

and $$B_{m\times m} = \begin{pmatrix} b_{11} & b_{21} & b_{31} & .... & b_{m1} \\ 1 & 0 & 0 & .... & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & .... & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

and the vectors:

$\underset{1\times p}{\mathbf{X}'} = (X_t \quad X_{t-1} \quad X_{t-2} \quad .... \quad X_{t-p-1}),\ \underset{1\times p}{H'} = (1 \quad 0 \quad 0 \quad .... \quad 0)$ and $C' = (1 \quad 0 \quad 0 \quad .... \quad 0)$

where $H'$ stands for the transpose of a matrix H,

With these notations, (3.44) can be written in the vector form as:

$$\mathbf{X}_t = \psi \mathbf{X}_{t-1} + \Psi \mathbf{X}_{t-s} + \mathrm{B}\mathbf{X}_{\mathrm{t-1}}\mathbf{e}_{\mathrm{t-1}} + \mathrm{C}\mathbf{e}_{\mathrm{t}} \tag{3.45}$$

Hence;

$$X_t = H' \, \mathbf{X}_t \tag{3.46}$$

### 3.11.2 Stationarity and convergence of the mixed seasonal autoregressive integrated one-dimensional bilinear time series model.

In the same vein, following Rao et.al (1983) and Sangodoyin *et al.* (2010), a sufficient condition necessary for process satisfying the mixed seasonal one-dimensional bilinear model (3.45) to be stationary and convergent can be achieved via the proceeding theorem.

**Theorem:**

Given $\{e_t, t \in \mathbb{Z}\}$, a set of independent and identically distributed random variables defined on the probability space $\{\Omega, F, P\}$ such that; $E(e_t) = 0 \; and \; E(e_t^2) = \sigma^2 < \infty$.

Let $\psi, \Psi, and \; \mathrm{B}$ be matrices as defined above such that;

$$\rho[(\psi \otimes \psi + 2\psi \otimes \Psi + \Psi \otimes \Psi + \sigma^2 \mathrm{B} \otimes \mathrm{B}] = \alpha < 1$$

And **C** be any vector with components $c_1, c_2, \ldots c_p$. Therefore, the series of random vectors

$$\sum_{v \geq 1} \prod_{i=1}^{v} (\psi + \Psi + \mathrm{B} e_{t-i}) C e_{t-v}$$

converges absolutely almost surely and also in the mean for all fixed $t \in \mathbb{Z}$.

On the other hand, if:

$$X_t = Ce_t + \sum_{v \geq 1} \prod_{i=1}^{v} (\psi + \Psi + \mathrm{B} e_{t-i}) C e_{t-v}, \quad t \in \mathbb{Z}$$

Then for every $t \in \mathbb{Z}, X_t$ is a strictly stationary process conforming to the mixed seasonal bilinear model:

$$\mathbf{X}_t = \psi \mathbf{X}_{t-1} + \Psi \mathbf{X}_{t-s} + \mathrm{B} \mathbf{X}_{\mathrm{t-1}} \mathrm{e}_{\mathrm{t-1}} + \mathrm{Ce}_{\mathrm{t}}$$

Conversely;

If $\{X_t, t \in \mathbb{Z}\}$ is a stationary process conforming to the mixed seasonal bilinear model;

$$\mathbf{X}_t = \psi \mathbf{X}_{t-1} + \Psi \mathbf{X}_{t-s} + \mathrm{B} \mathbf{X}_{\mathrm{t-1}} \mathrm{e}_{\mathrm{t-1}} + \mathrm{Ce}_{\mathrm{t}}$$

$\forall \; t \in \mathbb{Z}$ , for some sequences $\{e_t, t \in \mathbb{Z}\}$ of independent and identically distributed random variables with $E(e_t) = 0$ , $E(e_t^2) = \sigma^2 < \infty$ and for the matrices $\psi, \Psi, \mathrm{B}$ and $C$ of respective orders $p \times p, P \times P, \; m \times m$ and $p \times 1$ with:

$$\rho[(\psi \otimes \psi + 2\psi \otimes \Psi + \Psi \otimes \Psi + \sigma^2 \mathrm{B} \otimes \mathrm{B}] = \alpha < 1$$

Then,

$$X_t = Ce_t + \sum_{v \geq 1} \prod_{i=1}^{v} (\psi + \Psi + \mathrm{B} e_{t-i}) C e_{t-v}, \qquad \forall t \in \mathbb{Z} \tag{3.47}$$

**Proof:**

Similar to the procedures employed in the previous proof, we shall establish the proof as follows:

<u>Step 1</u>: For almost sure convergence;

$$\sum_{v\geq 1} E\left|\left(\prod_{j=1}^{v}(\psi+\Psi+\ \mathrm{B}e_{i-i})Ce_{t-v}\right)_j\right| < \infty \quad \forall \quad i=1,2,...,p \tag{3.48}$$

$\Rightarrow \sum_{v\geq 1}\prod_{i=1}^{v}(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v}$ absolutely converges almost surely and in the mean.

<u>Step 2:</u>

We establish (3.48) for i=1 and it should be noted that for all $t\in\mathbb{Z}, v\geq 1$ and $i=1,2,\dots p$

$$E\left|((\psi+\Psi+\ \mathrm{B}e_{t-i})Ce_{t-v})_n\right|$$

$$= E\left|\sum_{j=1}^{p}(\psi)_{nj}C_j e_{t-v} + \sum_{j=1}^{P}(\Psi)_{nj}C_j e_{t-v} + \sum_{j=1}^{m}(B)_{nj}c_j e_{t-v}^2\right| \leq k_0$$

<u>Step 3:</u>

Similarly if $v \geq 2$, hence there exists some $k_1 > 0$ such that:

$$E\left|\left(\prod_{i=1}^{v}(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v}\right)_1\right| \leq k_1 p\alpha^{(\frac{v-1}{2})} \tag{3.49}$$

However, it should be noted that:

$$E\left|\left(\prod_{i=1}^{v}(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v}\right)_1\right| = E\left|\left[(\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i})(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v}\right]_1\right|$$

$$= E\left|\sum_{n=1}^{p}(\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i}))_{1n}((\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v})_n\right|$$

$$\leq \sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v})_n\right|\right)$$

In the above step, we have used the fact that $\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i})$ and $(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v}$ are independently distributed. By step 2 and application of the Cauchy-Schwartz inequality, we have;

$$\sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v})_n\right|\right)$$

$$\leq k_0\sum_{n=1}^{p}\left[E\left(\left((\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i})\right)_{1n}\right)^2\right]^{\frac{1}{2}}$$

Now for any n=1,2,…,p

$$\left(\left((\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i})\right)_{1n}\right)^2=((\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i}))\otimes(\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i})))_{1n;1n}$$

$$=(\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i})\otimes(\psi+\Psi+\mathrm{B}e_{t-i}))_{1n;1n} \qquad (3.50)$$

Hence;

$$E\left(\left(\prod_{i=1}^{v-1}(\psi+\Psi+\mathrm{B}e_{t-i})\right)_{1n}\right)^2=\prod_{i=1}^{v-1}(E(\psi+\Psi+\mathrm{B}e_{t-i})\otimes(\psi+\Psi+\mathrm{B}e_{t-i}))_{1n;1n}$$

$$=((E[(\psi+\Psi+\mathrm{B}e_t)\otimes(\psi+\Psi+\mathrm{B}e_t)])^{v-1})_{1n;1n}$$

$$=((E[(\psi\otimes\psi+2\psi\otimes\Psi+2\psi\otimes\mathrm{B}e_t+\Psi\otimes\Psi+2\Psi\otimes\mathrm{B}e_t+\mathrm{B}\otimes\mathrm{B}e_t^2)])^{v-1})_{1n;1n}$$

$$=((\psi\otimes\psi+2\psi\otimes\Psi+\Psi\otimes\Psi+\sigma^2\mathrm{B}\otimes\mathrm{B})^{v-1})_{1n;1n}\le k\alpha^{v-1} \qquad (3.51)$$

Therefore;

$E\left|\left(\prod_{i=1}^{v}(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v}\right)_1\right|\le k_1 p\alpha^{(\frac{v-1}{2})}$ for any suitable choice of $k_1$.

Step 4:

Since $\alpha<1$,

$$\Rightarrow\sum_{v\ge1}E\left|\left(\prod_{i=1}^{v}(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v}\right)_i\right|<\infty \qquad \forall\ i=1,2,\ldots p$$

Thus (3.48) is proven.

Then, it is evident that the vector-valued process $\{X_t,t\in\mathbb{Z}\}$ described as:

$X_t=Ce_t+\sum_{v\ge1}\prod_{i=1}^{v}(\psi+\Psi+\mathrm{B}e_{t-i})Ce_{t-v},\ \ t\in\mathbb{Z}$ is strictly stationary.

Hence;

$$\rho(\psi\otimes\psi+2\psi\otimes\Psi+\Psi\otimes\Psi+\sigma^2\mathrm{B}\otimes\mathrm{B})=\alpha<1$$

is the sufficient condition that the model is strictly stationary.

This implies that if we desire a real-valued process $X_t$ satisfying the mixed seasonal autoregressive integrated one-dimensional bilinear model:

$$X_t=\psi_1X_{t-1}+\psi_2X_{t-2}....+\psi_{p+d}X_{t-p-d}+\Psi_1X_{t-s}+\Psi_2X_{t-2s}....+\Psi_{P+D}X_{t-Ps-D}$$
$$+b_{11}X_{t-1}e_{t-1}+....+b_{m1}X_{t-m}e_{t-1}+e_t$$

$\forall t \in \mathbb{Z}$ under the assumptions on $e_t$, a sufficient condition for its existence is;

$$\rho(\psi \otimes \psi + 2\psi \otimes \Psi + \Psi \otimes \Psi + \sigma^2 \mathrm{B} \otimes \mathrm{B}) = \alpha < 1 \tag{3.52}$$

Thus, the proof is established.

## 3.12 Pure seasonal autoregressive Integrated Moving Average one-dimensional Bilinear Time Series (PSARIMAODBL) Model.

We then define the pure seasonal autoregressive integrated moving average one-dimensional Bilinear Time Series Model (PSARIMAODBL) as follows:

$$\begin{aligned} X_t &= \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + \dots + \Psi_{P+D} X_{t-Ps-D} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + \dots \\ &\quad + \Theta_q e_{t-Qs} + b_{11} X_{t-1} e_{t-1} + \dots + b_{i1} X_{t-i} e_{t-1} + e_t \end{aligned} \tag{3.53}$$

### 3.12.1 Vector form of the specified Pure Seasonal Autoregressive Integrated Moving Average One-Dimensional Bilinear Time Series (SARIMABL) Model.

Similarly, we shall consider the vector form of the process due to the Markovian nature of the model. Akaike (1974). Therefore, given the process;

$$\begin{aligned} X_t &= \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + \dots + \Psi_{P+D} X_{t-Ps-D} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + \dots + \Theta_q e_{t-Qs} \\ &\quad + b_{11} X_{t-1} e_{t-1} + \dots + b_{m1} X_{t-i} e_{t-1} + e_t \end{aligned}$$

$$\begin{aligned} \Rightarrow X_{t-1} &= \Psi_1 X_{t-s-1} + \Psi_2 X_{t-2s-1} + \dots + \Psi_{P+D} X_{t-Ps-D-1} + \Theta_1 e_{t-s-1} + \Theta_2 e_{t-2s-1} + \dots \\ &\quad + \Theta_Q e_{t-Qs-1} + b_{11} X_{t-2} e_{t-2} + \dots + b_{m1} X_{t-i-1} e_{t-2} + e_{t-1} \\ X_{t-2} &= \Psi_1 X_{t-s-2} + \Psi_2 X_{t-2s-2} + \dots + \Psi_{P+D} X_{t-Ps-D-2} + \Theta_1 e_{t-s-2} + \Theta_2 e_{t-2s-2} + \dots \\ &\quad + \Theta_q e_{t-Qs-2} + b_{11} X_{t-3} e_{t-3} + \dots + b_{m1} X_{t-i-1} e_{t-3} + e_{t-2} \\ &\vdots \\ X_{t-p-1} &= \Psi_1 X_{t-s-p-1} + \Psi_2 X_{t-2s-p-1} + \dots + \Psi_{P+D-p-1} X_{t-Ps-D-p-1} + \Theta_1 e_{t-s-p-1} + \Theta_2 e_{t-2s-p-1} + \dots \\ &\quad + \Theta_q e_{t-Qs-p-1} + b_{11} X_{t-p-2} e_{t-p-1} + \dots + b_{m1} X_{t-i-p-1} e_{t-p-2} + e_{t-p-1} \end{aligned}$$

Let:

$$\Psi = \begin{pmatrix} \Psi_1 & \Psi_2 & \Psi_3 & .... & \Psi_{P+D} & \Psi_{P+D+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix}, \qquad \Theta = \begin{pmatrix} \Theta_1 & \Theta_2 & \Theta_3 & .... & \Theta_Q & \Theta_{Qs+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix},$$

and
$$B_{m\times m} = \begin{pmatrix} b_{11} & b_{21} & b_{31} & .... & b_{m1} \\ 1 & 0 & 0 & .... & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & .... & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

And the vectors: $\underset{1\times p}{\mathbf{X}'} = (X_t \quad X_{t-1} \quad X_{t-2} \quad .... \quad X_{t-p-1})$ , $\underset{1\times p}{H'} = (1 \quad 0 \quad 0 \quad .... \quad 0)$ and $C' = (1 \quad 0 \quad 0 \quad .... \quad 0)$.

where $H'$ is the transpose of $H$ and $t = \cdots -1, 0, 1, \ldots$. With these notations, (3.53) can be expressed in the vector form as:

$$\mathbf{X}_t = \Psi\mathbf{X}_{t-s} + \Theta\mathrm{e}_{\mathrm{t-s}} + \mathrm{B}\mathbf{X}_{\mathrm{t-1}}\mathrm{e}_{\mathrm{t-1}} + \mathrm{Ce}_{\mathrm{t}} \qquad (3.54) \qquad 3.54)$$

$$\Rightarrow X_t = H'\,\mathbf{X}_t \qquad (3.55)$$

### 3.12.2 Stationarity and convergence of the Pure Seasonal Autoregressive Integrated Moving Average One-Dimensional Bilinear Time Series (PSARIMAODBL)

Similarly, following Rao *et al.* (1983) and Sangodoyin *et al.* (2010), a sufficient condition necessary for the stochastic process conforming to the pure seasonal autoregressive integrated moving average one-dimensional bilinear model (3.53) can be achieved through the following theorem.

**Theorem:**

Consider a sequence of independent and identically distributed random variables $\{e_t, t \in \mathbb{Z}\}$ defined on the probability space $\{\Omega, F, P\}$ such that; $E(e_t) = 0$ *and* $E(e_t^2) = \sigma^2 < \infty$.

Let $\Psi, \Theta$ *and* $\mathrm{B}$ be matrices as defined above such that;

$$\rho[(\Psi \otimes \Psi + \sigma^2(\Theta \otimes \Theta + 2\Theta \otimes \mathrm{B} + \mathrm{B} \otimes \mathrm{B})] = \beta < 1$$

And **C** be any column vector with components $c_1, c_2, \dots c_p$. Hence;

$$\sum_{w \geq 1} \prod_{i=1}^{w} (\Psi + \Theta e_{t-is} + \mathrm{B} e_{t-i}) C e_{t-w}$$

converges absolutely almost surely and in the mean for every fixed t in $\mathbb{Z}$.

Moreover, if:

$$X_t = Ce_t + \sum_{w \geq 1} \prod_{i=1}^{w} (\Psi + \Theta e_{t-is} + \mathrm{B} e_{t-i}) C e_{t-w}, \quad t \in \mathbb{Z}$$

Then for all $t \in Z$, there exists a strictly stationary process $X_t$ satisfying the pure seasonal autoregressive integrated moving average one-dimensional bilinear model:

$$\mathbf{X}_t = \Psi \mathbf{X}_{t-s} + \Theta \mathrm{e}_{\text{t-s}} + \mathrm{B} \mathbf{X}_{\text{t-1}} \mathrm{e}_{\text{t-1}} + \mathrm{C} \mathrm{e}_{\text{t}} \tag{3.54}$$

Conversely;

If $\{X_t, t \in \mathbb{Z}\}$ is a stationary process conforming to the mixed seasonal bilinear model

$$\mathbf{X}_t = \Psi \mathbf{X}_{t-s} + \Theta \mathbf{e}_{\text{t-s}} + \mathbf{B} \mathbf{X}_{\text{t-1}} \mathbf{e}_{\text{t-1}} + \mathbf{C} \mathbf{e}_{\text{t}}$$

$\forall t \in Z$, for some sequences $\{e_t, t \in \mathbb{Z}\}$ of independent and identically distributed random variables with $E(e_t) = 0$ *and* $E(e_t^2) = \sigma^2 < \infty$ and for matrices $\Psi, \Theta, \mathrm{B}$ and $C$ of orders $P \times P,\ Q \times Q,\ m \times m$ and $p \times 1$ respectively with:

$$\rho[(\Psi \otimes \Psi + \sigma^2(\Theta \otimes \Theta + 2\Theta \otimes \mathrm{B} + \mathrm{B} \otimes \mathrm{B})] = \beta < 1$$

Then;

$$X_t = Ce_t + \sum_{w \geq 1} \prod_{i=1}^{w} (\Psi + \Theta e_{t-is} + \mathrm{B} e_{t-i}) C e_{t-w}, \quad \text{for every } \ t \in \mathbb{Z} \tag{3.55}$$

**Proof:**

Similar to the procedures employed in the previous proof, we shall establish the proof as follows:

Step 1: To establish almost sure convergence condition, we prove that;

$$\sum_{w\geq 1} E\left|\left(\prod_{j=1}^{w}(\Psi+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-w}\right)_j\right|<\infty,\quad \forall \quad i=1,2,\ldots,p \qquad (3.56)$$

$\Rightarrow \sum_{w\geq 1}\prod_{i=1}^{w}(\Psi+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-w}$ is absolutely convergent almost surely as well as in the mean.

Step 2:

We establish (3.56) for $i$=1 and it should be noted that $\forall\ t\in\mathbb{Z}, w=1$ and $i=1,2,\ldots p$

$$E\left|((\Psi+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-w})_n\right|\leq k_0$$

where $k_0$ depends on $\Psi,\Theta,\mathrm{B},C$ and $\sigma^2$

Step 3:

Similarly if $r\geq 2$, then there exists some $k_1>0$ such that:

$$E\left|\left(\prod_{i=1}^{w}(\Psi+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-w}\right)_1\right|\leq k_1\beta^{(\frac{w-1}{2})}$$

However, it should be noted that:

$$E\left|\left(\prod_{i=1}^{w}(\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w}\right)_1\right|$$

$$=E\left|\left[(\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i})(\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w}\right]_1\right|$$

$$=E\left|\sum_{n=1}^{p}(\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i}))_{1n}((\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w})_n\right|$$

$$\leq\sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w})_n\right|\right)$$

In the above step, we have used the fact that $\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i})$ and $(\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w}$ are independently distributed. So, by step 2 and application of the Cauchy-Schwartz inequality,

$\sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w})_n\right|\right)$ is not greater than

$$k_0\sum_{n=1}^{p}\left[E\left(\left((\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i})\right)_{1n}\right)^2\right]^{\frac{1}{2}}$$

Now for any n=1,2,…,p

$$\left(\left((\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i})\right)_{1n}\right)^2=((\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i}))\otimes(\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i})))_{1n;1n}$$

$$=(\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i})\otimes(\Psi+\Theta+\mathrm{B}e_{t-i}))_{1n;1n} \qquad (3.57)$$

Hence;

$$E\left(\left(\prod_{i=1}^{w-1}(\Psi+\Theta+\mathrm{B}e_{t-i})\right)_{1n}\right)^2=\prod_{i=1}^{w-1}(E(\Psi+\Theta+\mathrm{B}e_{t-i})\otimes(\Psi+\Theta+\mathrm{B}e_{t-i}))_{1n;1n}$$

$$=((E[(\Psi+\Theta+\mathrm{B}e_t)\otimes(\Psi+\Theta+\mathrm{B}e_t)])^{w-1})_{1n;1n}$$

$$=((E[(\Psi\otimes\Psi+2\Psi\otimes\Theta+2\Psi\otimes\mathrm{B}e_t+\Theta\otimes\Theta+2\;\Theta\otimes\mathrm{B}e_t+\mathrm{B}\otimes\mathrm{B}e_t^2)])^{w-1})_{1n;1n}$$

$$=((\Psi\otimes\Psi+2\Psi\otimes\Theta+\Theta\otimes\Theta+\sigma^2\mathrm{B}\otimes\mathrm{B})^{w-1})_{1n;1n}\le k\beta^{m-1} \qquad (3.58)$$

Therefore;

$E\left|\left(\prod_{i=1}^{w}(\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w}\right)_1\right|\le k_1P\beta^{(\frac{w-1}{2})}$ for any suitable choice of $k_1$.

Step 4:

Since $\beta<1$,

$$\Rightarrow\sum_{w\ge1}E\left|\left(\prod_{i=1}^{w}(\Psi+\Theta+\mathrm{B}e_{t-i})Ce_{t-w}\right)_i\right|<\infty\;\forall\;\; i=1,2,\dots p$$

Thus (3.56) is established.

It is then obvious that the vector-valued stochastic process;

$$X_t = Ce_t + \sum_{w\geq 1}\prod_{i=1}^{w}(\Psi + \Theta + \mathrm{B}e_{t-i})Ce_{t-w},\ \ t\in\mathbb{Z} \text{ is strictly stationary.}$$

Hence;

$$\rho(\Psi\otimes\Psi + 2\Psi\otimes\Theta + \Theta\otimes\Theta + \sigma^2\mathrm{B}\otimes\mathrm{B}) = \beta < 1$$

is a sufficient condition for strict stationarity of the model.

Therefore, if we desire a real-valued process $X_t$ conforming to the mixed seasonal autoregressive integrated one-dimensional bilinear model:

$$\begin{aligned} X_t &= \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_{P+D} X_{t-Ps-D} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + .... \\ &\quad + \Theta_q e_{t-Qs} + b_{11} X_{t-1} e_{t-1} + .... + b_{m1} X_{t-i} e_{t-1} + e_t \end{aligned}$$

$\forall t\in\mathbb{Z}$ under the stated assumptions on $e_t$, a necessary and sufficient condition for its existence is given by:

$$\rho(\Psi\otimes\Psi + 2\Psi\otimes\Theta + \Theta\otimes\Theta + \sigma^2\mathrm{B}\otimes\mathrm{B}) = \beta < 1 \qquad (3.57)$$

Thus, the proof is established.

## 3.13 Mixed seasonal Autoregressive Integrated Moving Average One-Dimensional Bilinear (MSARIMAODBL) time series model.

We define the Mixed Eeasonal Autoregressive Integrated Moving Average One-Dimensional Bilinear Time Series Model (MSARIMAODBL) as follows:

$$\begin{aligned} X_t &= \psi_1 X_{t-1} + .... + \psi_{p+d} X_{t-p-d} + \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_{P+D} X_{t-Ps-D} + \theta_1 e_{t-1} \\ &\quad + .... + \theta_q e_{t-q} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + .... + \Theta_q e_{t-Qs} + b_{11} X_{t-1} e_{t-1} + .... + b_{m1} X_{t-i} e_{t-1} + e_t \end{aligned} \qquad (3.58)$$

### 3.13.1 Vector form of the MSARIMAODBL time series model.

Similarly, we shall consider the state space form because of the Markovian nature of the model Akaike (1974).

Given:

$$X_t = \psi_1 X_{t-1} + .... + \psi_{p+d} X_{t-p-d} + \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_{P+D} X_{t-Ps-D} + \theta_1 e_{t-1}$$
$$+ .... + \theta_q e_{t-q} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + .... + \Theta_q e_{t-Qs} + b_{11} X_{t-1} e_{t-1} + .... + b_{m1} X_{t-i} e_{t-1} + e_t$$

$$\Rightarrow X_{t-1} = \psi_1 X_{t-2} + \psi_2 X_{t-3} + .... + \psi_{p+d} X_{t-p-d-1} + \Psi_1 X_{t-s-1} + \Psi_2 X_{t-2s-1} + ....$$
$$+ \Psi_{P+D} X_{t-Ps-D-1} + \theta_1 e_{t-2} + \theta_2 e_{t-3} + .... + \theta_q e_{t-q-1} + \Theta_1 e_{t-s-1} + \Theta_2 e_{t-2s-1} + ....$$
$$+ \Theta_Q e_{t-Qs-1} + b_{11} X_{t-2} e_{t-2} + .... + b_{m1} X_{t-i-1} e_{t-2} + e_{t-1}$$

$$X_{t-2} = \psi_1 X_{t-3} + \psi_2 X_{t-4} + .... + \psi_{p+d} X_{t-p-d-2} + \Psi_1 X_{t-s-2} + \Psi_2 X_{t-2s-2} + ....$$
$$+ \Psi_{P+D} X_{t-Ps-D-2} + \theta_1 e_{t-3} + \theta_2 e_{t-4} + .... + \theta_q e_{t-q-2} + \Theta_1 e_{t-s-2} + \Theta_2 e_{t-2s-2} + ....$$
$$+ \Theta_q e_{t-Qs-2} + b_{11} X_{t-3} e_{t-3} + .... + b_{m1} X_{t-i-1} e_{t-3} + e_{t-2}$$

$$\vdots$$

$$X_{t-p-1} = \psi_1 X_{t-p-2} + \psi_2 X_{t-p-3} + .... + \psi_{p+d} X_{t-2p-d-1} + \Psi_1 X_{t-s-p-1} + \Psi_2 X_{t-2s-p-1} + ....$$
$$+ \Psi_{P+D-p-1} X_{t-Ps-D-p-1} + \theta_1 e_{t-p-2} + \theta_2 e_{t-p-3} + .... + \theta_q e_{t-q-p-1} + \Theta_1 e_{t-s-p-1}$$
$$+ \Theta_2 e_{t-2s-p-1} + .... + \Theta_q e_{t-Qs-p-1} + b_{11} X_{t-p-2} e_{t-p-1} + .... + b_{m1} X_{t-i-p-1} e_{t-p-2} + e_{t-p-1}$$

Let the matrices:

$$\psi = \begin{pmatrix} \psi_1 & \psi_2 & \psi_3 & .... & \psi_p & \psi_{p+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix}, \quad \Psi = \begin{pmatrix} \Psi_1 & \Psi_2 & \Psi_3 & .... & \Psi_{P+D} & \Psi_{P+D+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix},$$

$$B_{m\times m} = \begin{pmatrix} b_{11} & b_{21} & b_{31} & .... & b_{m1} \\ 1 & 0 & 0 & .... & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & .... & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$\theta = \begin{pmatrix} \theta_1 & \theta_2 & \theta_3 & .... & \theta_q & \theta_{q+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix} \quad \text{and} \quad \Theta = \begin{pmatrix} \Theta_1 & \Theta_2 & \Theta_3 & .... & \Theta_Q & \Theta_{Qs+1} \\ 1 & 0 & 0 & .... & 0 & 0 \\ 0 & 1 & 0 & .... & 0 & 0 \\ \vdots & 0 & 1 & .... & 0 & 0 \\ 0 & 0 & 0 & .... & 1 & 0 \end{pmatrix}$$

And the vectors:

$\underset{1\times p}{\mathbf{X}'}=(X_t \quad X_{t-1} \quad X_{t-2} \quad .... \quad X_{t-p-1}), \quad \underset{1\times p}{H'}=(1 \quad 0 \quad 0 \quad .... \quad 0)$ and

$C'=(1 \quad 0 \quad 0 \quad .... \quad 0)$ where $H'$ denotes for the transpose of a matrix H, $t = \cdots -1, 0, 1, \dots.$.

Therefore, (3.58) in the state space form is:

$$\mathbf{X}_t = \psi\mathbf{X}_{t-1} + \Psi\mathbf{X}_{t-s} + \theta\mathrm{e}_{t\text{-}1} + \Theta\mathrm{e}_{t\text{-}s} + \mathrm{B}\mathbf{X}_{t\text{-}1}\mathrm{e}_{t\text{-}1} + \mathrm{C}\mathrm{e}_{t} \tag{3.59}$$

Hence;

$$X_t = H'\,\mathbf{X}_t \tag{3.60}$$

### 3.13.2 Stationarity and convergence of the mixed seasonal autoregressive Integrated Moving Average one-dimensional Bilinear Time Series model.

As in the previous sections, following Rao *et al.* (1983) and Sangodoyin *et al.* (2010), a sufficient condition necessary for the existence of strictly stationary process and convergence conforming to the mixed seasonal autoregressive integrated moving average one-dimensional bilinear time series model can be achieved through the following theorem.

**Theorem:**

Let $\{e_t, t \in \mathbb{Z}\}$ be a sequence of independent and identically distributed random variables defined on the probability space $\{\Omega, F, P\}$ such that; $E(e_t) = 0$ and $E(e_t^2) = \sigma^2 < \infty$

Let $\psi, \Psi, \theta, \Theta$ and $\mathrm{B}$ be matrices as defined above such that;

$$\rho[(\psi \otimes \psi + 2\psi \otimes \Psi + \Psi \otimes \Psi + 2\sigma^2(\theta \otimes \Theta + \theta \otimes \mathrm{B} + \frac{1}{2}\Theta \otimes \Theta + \Theta \otimes \mathrm{B}$$
$$+\frac{1}{2}\mathrm{B} \otimes \mathrm{B})] = \tau < 1$$

And $\mathbf{C}$, any column vector with components $c_1, c_2, \dots c_p$. Then, the vectors;

$$\sum_{r\geq 1}\prod_{i=1}^{r}(\psi + \Psi + \theta e_{t-i} + \Theta e_{t-is} + \mathrm{B}e_{t-i})Ce_{t-r}$$

would converge absolutely almost surely and in the mean for all fixed t in $\mathbb{Z}$.

Moreover, suppose:

$$X_t = Ce_t + \sum_{r\geq 1}\prod_{\mathrm{i}=1}^{\mathrm{r}}(\psi + \Psi + \theta e_{t-i} + \Theta e_{t-is} + \mathrm{B}e_{t-i})Ce_{t-r}, \quad t \in \mathbb{Z}$$

Then for every $t \in \mathbb{Z}, X_t$ is a strictly stationary process conforming to the mixed seasonal bilinear model:

$$\mathbf{X}_t = \psi\mathbf{X}_{t-1} + \Psi\mathbf{X}_{t-s} + \theta \mathrm{e}_{\mathrm{t\text{-}1}} + \Theta \mathrm{e}_{\mathrm{t\text{-}s}} + \mathrm{B}\mathbf{X}_{\mathrm{t\text{-}1}}\mathrm{e}_{\mathrm{t\text{-}1}} + \mathrm{Ce}_{\mathrm{t}}$$

Conversely;

If $\{X_t, t \in Z\}$ is a stationary process conforming to the mixed seasonal bilinear model

$$\mathbf{X}_t = \psi\mathbf{X}_{t-1} + \Psi\mathbf{X}_{t-s} + \theta \mathrm{e}_{\mathrm{t\text{-}1}} + \Theta \mathrm{e}_{\mathrm{t\text{-}s}} + \mathrm{B}\mathbf{X}_{\mathrm{t\text{-}1}}\mathrm{e}_{\mathrm{t\text{-}1}} + \mathrm{Ce}_{\mathrm{t}}$$

$\forall\, t \in Z$ , for some sequences $\{e_t, t \in \mathbb{Z}\}$ of independent and identically distributed random variables with $E(e_t) = 0$ *and* $E(e_t^2) = \sigma^2 < \infty$ and for some matrices $\psi, \Psi, \theta, \Theta,$ B and $C$ of orders $p \times p, P \times P, q \times q, Q \times Q,\ m \times m$ and $p \times 1$ respectively with

$$\rho[(\psi \otimes \psi + 2\psi \otimes \Psi + \Psi \otimes \Psi + 2\sigma^2(\theta \otimes \Theta + \theta \otimes \mathrm{B} + \frac{1}{2}\Theta \otimes \Theta + \Theta \otimes \mathrm{B}$$
$$+ \frac{1}{2}\mathrm{B} \otimes \mathrm{B})] = \tau < 1 \tag{3.61}$$

**Proof:**

Similar to the procedures employed in the previous proof, we shall establish the proof via the succeeding steps:

Step 1: For almost sure convergence, we prove that;

$$\sum_{x\geq 1} E\left|\left(\prod_{i=1}^{x}(\psi + \Psi + \theta e_{t-i} + \Theta e_{t-is} + \mathrm{B}e_{t-i})Ce_{t-x}\right)_j\right| < \infty, \ \forall \quad i = 1, 2, ..., p \tag{3.62}$$

$\Rightarrow \sum_{x\geq 1}\prod_{i=1}^{x}(\psi + \Psi + \theta e_{t-i} + \Theta e_{t-is} + \mathrm{B}e_{t-i})Ce_{t-x}$ is absolutely convergent almost surely as well as in the mean.

Step 2:

We establish (3.62) for $i$=1 and $\forall\ t \in \mathbb{Z}, x=1$ and $i=1,2,\ldots p$, we have;

$$E\left|((\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x})_n\right| \le k_0$$

where $k_0$ is a function of $\psi, \Psi, \Theta, \mathrm{B}, C$ and $\sigma^2$

Step 3:

Similarly if $x \geq 2$, then there exists some $k_1 > 0$ such that:

$$E\left|\left(\prod_{i=1}^{x}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x}\right)_1\right| \le k_1\tau^{(\frac{x-1}{2})}$$

However, it should be noted that:

$$E\left|\left(\prod_{i=1}^{x}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x}\right)_1\right|$$

$$=E\left|\left[(\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x}\right]_1\right|$$

$$=E\left|\sum_{n=1}^{p}(\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i}))_{1n}((\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x})_n\right|$$

$$\le\sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x})_n\right|\right)$$

In the above step, we have used the fact that; $\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})$ and

$(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x}$ are independently distributed.

Therefore by step 2 and the Cauchy-Schwartz inequality,

$\sum_{n=1}^{p}\left(E\left|(\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i}))_{1n}\right|\right)\left(E\left|((\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x})_n\right|\right)$ is not

greater than;

$$k_0\sum_{n=1}^{p}\left[E\left(\left((\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})\right)_{1n}\right)^2\right]^{\frac{1}{2}}$$

Now for any n=1,2,…,p

$$\left(\left((\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})\right)_{1n}\right)^2$$

$$=((\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i}))\otimes(\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})))_{1n;1n}$$

$$=(\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})\otimes(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i}))_{1n;1n} \tag{3.63}$$

Hence;

$$E\left(\left(\prod_{i=1}^{x-1}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})\right)_{1n}\right)^2$$

$$=\prod_{i=1}^{x-1}(E(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})\otimes(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i}))_{1n;1n}$$

$$=((E[(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t})\otimes(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t})])^{x-1})_{1n;1n}$$

$$=((E[(\psi\otimes\psi+2\psi\otimes\Psi+2\psi\otimes\theta e_t+2\psi\otimes\Theta e_t+2\,\psi\otimes\mathrm{B}e_t+\Psi\otimes\Psi$$
$$+2\Psi\otimes\theta e_t+2\Psi\otimes\Theta e_t+2\Psi\otimes\mathrm{B}e_t+2\theta\otimes\Theta e_t^2+2\theta\otimes\mathrm{B}e_t^2+\Theta\otimes\Theta e_t^2$$
$$+2\Theta\otimes\mathrm{B}e_t^2+\mathrm{B}\otimes\mathrm{B}e_t^2)])^{x-1})_{1n;1n}$$

$$=((\psi\otimes\psi+2\psi\otimes\Psi+\Psi\otimes\Psi+2\sigma^2(\theta\otimes\Theta+\theta\otimes\mathrm{B}+\frac{1}{2}\Theta\otimes\Theta+\Theta\otimes\mathrm{B}$$
$$+\frac{1}{2}\mathrm{B}\otimes\mathrm{B}))^{x-1})_{1n;1n}\le k\tau^{x-1}$$

Therefore;

$$E\left|\left(\prod_{i=1}^{x}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x}\right)_1\right|\le k_1 p\tau^{(\frac{x-1}{2})}$$

for any suitable choice of $k_1$.

Step 4:

Since $\tau<1$,

$$\Rightarrow\sum_{x\ge1}E\left|\left(\prod_{i=1}^{x}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x}\right)_i\right|<\infty\ \forall\ \ i=1,2,\dots p$$

Thus (3.62) is established.

It is then obvious that the vector-valued stochastic process $\{X_t, t\in\mathbb{Z}\}$ defined by:

$$X_t=Ce_t+\sum_{x\ge1}\prod_{i=1}^{x}(\psi+\Psi+\theta e_{t-i}+\Theta e_{t-is}+\mathrm{B}e_{t-i})Ce_{t-x},\ \ t\in\mathbb{Z}$$

is strictly stationary.

Hence;

$$\rho((\psi \otimes \psi + 2\psi \otimes \Psi + \Psi \otimes \Psi + 2\sigma^2(\theta \otimes \Theta + \theta \otimes \mathrm{B} + \frac{1}{2}\Theta \otimes \Theta + \Theta \otimes \mathrm{B}$$
$$+ \frac{1}{2}\mathrm{B} \otimes \mathrm{B})) = \tau < 1$$

is a sufficient and necessary condition for strict stationarity of the mixed SARIMA bilinear model.

So, if we desire a real-valued process $X_t$ conforming to the mixed seasonal autoregressive integrated one-dimensional bilinear model:

$$X_t = \psi_1 X_{t-1} + .... + \psi_{p+d} X_{t-p-d} + \Psi_1 X_{t-s} + \Psi_2 X_{t-2s} + .... + \Psi_{P+D} X_{t-Ps-D} + \theta_1 e_{t-1}$$
$$+ .... + \theta_q e_{t-q} + \Theta_1 e_{t-s} + \Theta_2 e_{t-2s} + .... + \Theta_q e_{t-Qs} + b_{11} X_{t-1} e_{t-1} + .... + b_{m1} X_{t-i} e_{t-1} + e_t$$

$\forall t \in \mathbb{Z}$ under the assumptions on $e_t$, a sufficient condition for its existence is given by:

$$\rho((\psi \otimes \psi + 2\psi \otimes \Psi + \Psi \otimes \Psi + 2\sigma^2(\theta \otimes \Theta + \theta \otimes \mathrm{B} + \frac{1}{2}\Theta \otimes \Theta + \Theta \otimes \mathrm{B}$$
$$+ \frac{1}{2}\mathrm{B} \otimes \mathrm{B})) = \tau < 1 \qquad (3.64)$$

Thus, the proof is established.

## 3.14 Estimation of parameters of the mixed seasonal autoregressive integrated moving average one-dimensional bilinear time series model.

Due to the fact that the process of parameter estimation of the models is similar to that of the full model, we shall report the estimation of the parameters of the one-dimensional case for the mixed seasonal autoregressive integrated moving average model. Suppose that $X_t$ are generated by equation (3.58), the sequence of random deviates $\{e_t\}$ could be determined from the relation:

$$e_t = X_t - \psi_1 X_{t-1} - \psi_2 X_{t-2} .... - \psi_{p+1} X_{t-p-1} - \Psi_1 X_{t-s} - \Psi_2 X_{t-2s} .... - \Psi_{P+D} X_{t-Ps-D}$$
$$- \theta_1 e_{t-1} - \theta_1 e_{t-2} - ... - \theta_q e_{t-q} - \Theta_1 e_{t-s} - \Theta_2 X_{t-2s} .... - \Theta_Q e_{t-Qs} - b_{11} X_{t-1} e_{t-1} - ....$$
$$- b_{m1} X_{t-i} e_{t-1} \tag{3.65}$$

To estimate the unknown parameters we shall minimize the error by obtaining the first and second order partial derivatives of (3.65) with respect to the individual parameters; $\psi_1, \psi_2 .... \psi_{p+1};\ \Psi_1,\ \Psi_2 .... \Psi_{P+D};\ \ \theta_1,\ \theta_2 .... \theta_q;\ \ \ \Theta_1, \Theta_2 .... \Theta_Q$ and $b_{i1}$ as follows:

$$\frac{\partial e_t}{\partial \psi_i} = -X_{t-i} - \sum_{i=1}^{q} \theta_i \frac{\partial e_{t-i}}{\partial \psi_i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial e_{t-is}}{\partial \psi_i} - \sum_{i=1}^{m} b_{i1} X_{t-i} \frac{\partial e_{t-1}}{\partial \psi_i} \tag{3.66}$$

$$\frac{\partial e_t}{\partial \Psi_i} = -X_{t-is} - \sum_{i=1}^{q} \theta_i \frac{\partial e_{t-i}}{\partial \Psi_i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial e_{t-s}}{\partial \Psi_i} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial e_{t-1}}{\partial \Psi_i} \tag{3.67}$$

$$\frac{\partial e_t}{\partial \theta_i} = -e_{t-i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial e_{t-is}}{\partial \theta_i} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial e_{t-1}}{\partial \theta_i} \tag{3.68}$$

$$\frac{\partial e_t}{\partial \Theta_i} = -e_{t-is} - \sum_{i=1}^{q} \theta_i \frac{\partial e_{t-i}}{\partial \Theta_i} + \sum_{i=1}^{m} b_{i1} X_{t-i} \frac{\partial e_{t-1}}{\partial \Theta_i} \tag{3.69}$$

$$\frac{\partial e_t}{\partial b_{i1}} = -X_{t-i} e_{t-1} - \sum_{i=1}^{q} \theta_i \frac{\partial e_{t-i}}{\partial b_{i1}} - \sum_{i=1}^{Q} \Theta_i \frac{\partial e_{t-s}}{\partial b_{i1}} - b_{i1} X_{t-i} \frac{\partial e_{t-1}}{\partial b_{i1}} \tag{3.70}$$

And the second derivatives;

$$\frac{\partial^2 e_t}{\partial \psi_i^2} = -\sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \psi_i^2} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-s}}{\partial \psi_i^2} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \psi_i^2} \tag{3.71}$$

$$\frac{\partial^2 e_t}{\partial \Psi_i^2} = -\sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \Psi_i^2} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-s}}{\partial \Psi_i^2} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \Psi_i^2} \tag{3.72}$$

$$\frac{\partial^2 e_t}{\partial \theta_i^2} = \frac{\partial e_{t-i}}{\partial \theta_i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-is}}{\partial \theta_i^2} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \theta_i^2} \tag{3.73}$$

$$\frac{\partial^2 e_t}{\partial \Theta_i^2} = \frac{\partial e_{t-is}}{\partial \Theta_i} - \sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \Theta_i^2} + \sum_{i=1}^{m} b_{i1} X_{t-i} \frac{\partial^2 e_{t-1}}{\partial \Theta_i^2} \tag{3.74}$$

$$\frac{\partial^2 e_t}{\partial b_{i1}^2} = -X_{t-i} \frac{\partial e_{t-1}}{\partial b_{i1}} - \sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial b_{i1}^2} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-s}}{\partial b_{i1}^2} - b_{i1} X_{t-i} \frac{\partial^2 e_{t-1}}{\partial b_{i1}^2} \tag{3.75}$$

$$\frac{\partial^2 e_t}{\partial \psi_i \Psi_i} = -\sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \psi_i \Psi_i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-is}}{\partial \psi_i \Psi_i} - \sum_{i=1}^{m} b_{i1} X_{t-i} \frac{\partial^2 e_{t-1}}{\partial \psi_i \Psi_i} \tag{3.76}$$

$$\frac{\partial^2 e_t}{\partial \psi_i \theta_i} = -\sum_{i=1}^{q} \frac{\partial^2 e_{t-i}}{\partial \psi_i \theta_i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-is}}{\partial \psi_i \theta_i} - \sum_{i=1}^{m} b_{i1} X_{t-i} \frac{\partial^2 e_{t-1}}{\partial \psi_i \theta_i} \tag{3.77}$$

$$\frac{\partial^2 e_t}{\partial \psi_i \Theta_i} = -\sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \psi_i \Theta_i} - \sum_{i=1}^{Q} \frac{\partial^2 e_{t-is}}{\partial \psi_i \Theta_i} - \sum_{i=1}^{m} b_{i1} X_{t-i} \frac{\partial^2 e_{t-1}}{\partial \psi_i \Theta_i} \tag{3.78}$$

$$\frac{\partial^2 e_t}{\partial \psi_i b_{i1}} = -\sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \psi_i b_{i1}} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-is}}{\partial \psi_i b_{i1}} - \sum_{i=1}^{m} X_{t-i} \frac{\partial^2 e_{t-1}}{\partial \psi_i b_{i1}} \tag{3.79}$$

$$\frac{\partial^2 e_t}{\partial \Psi_i \partial \theta_i} = -\sum_{i=1}^{q} \frac{\partial^2 e_{t-i}}{\partial \Psi_i \partial \theta_i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-s}}{\partial \Psi_i \partial \theta_i} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \Psi_i \partial \theta_i} \tag{3.80}$$

$$\frac{\partial^2 e_t}{\partial \Psi_i \partial \Theta_i} = -\sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \Psi_i \partial \Theta_i} - \sum_{i=1}^{Q} \frac{\partial^2 e_{t-s}}{\partial \Psi_i \partial \Theta_i} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \Psi \partial \Theta_{ii}} \tag{3.81}$$

$$\frac{\partial^2 e_t}{\partial \Psi_i \partial b_{i1}} = -\sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \Psi_i} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-s}}{\partial \Psi_i} - \sum_{i=1}^{m} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \Psi_i b_{i1}} \tag{3.82}$$

$$\frac{\partial^2 e_t}{\partial \theta_i \partial \Theta_i} = -\frac{\partial e_{t-i}}{\partial \Theta_i} - \sum_{i=1}^{Q} \frac{\partial^2 e_{t-is}}{\partial \theta_i \partial \Theta_i} - \sum_{i=1}^{m} b_{i1} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \theta_i \partial \Theta_i} \tag{3.83}$$

$$\frac{\partial^2 e_t}{\partial \theta_i b_{i1}} = -\frac{\partial e_{t-i}}{\partial b_{i1}} - \sum_{i=1}^{Q} \Theta_i \frac{\partial^2 e_{t-is}}{\partial \theta_i b_{i1}} - \sum_{i=1}^{m} X_{t-1} \frac{\partial^2 e_{t-1}}{\partial \theta_i b_{i1}} \tag{3.84}$$

$$\frac{\partial e_t}{\partial \Theta_i b_{i1}} = -\frac{\partial e_{t-is}}{\partial b_{i1}} - \sum_{i=1}^{q} \theta_i \frac{\partial^2 e_{t-i}}{\partial \Theta_i b_{i1}} + \sum_{i=1}^{m} X_{t-i} \frac{\partial^2 e_{t-1}}{\partial \Theta_i b_{i1}} \tag{3.85}$$

Proceeding as in Subba Rao (1981), we can note that maximizing the likelihood function of $(X_{n_0,}X_{n_0+1,} \dots, X_n)$ is the same as minimizing the function;

$$V = V(\lambda) = \sum_{t=1}^{n} e_t^2$$

The first and second-order derivatives of $V(\lambda)$ are solved to obtain the components of:

$$G(\lambda) = \begin{pmatrix} \dfrac{\partial V}{\partial \psi} \\[2ex] \dfrac{\partial V}{\partial \Psi} \\[2ex] \dfrac{\partial V}{\partial \theta} \\[2ex] \dfrac{\partial V}{\partial \psi} \\[2ex] \dfrac{\partial V}{\partial b_{m1}} \end{pmatrix} \quad \text{and} \quad \lambda = \begin{pmatrix} \psi \\ \Psi \\ \theta \\ \Theta \\ b_{m1} \end{pmatrix} \tag{3.86}$$

$$\text{and } H(\lambda) = \begin{pmatrix} \dfrac{\partial^2 V}{\partial \psi^2} & \dfrac{\partial^2 V}{\partial \psi \partial \Psi} & \dfrac{\partial^2 V}{\partial \psi \partial \theta} & \dfrac{\partial^2 V}{\partial \psi \partial \Theta} & \dfrac{\partial^2 V}{\partial \psi \partial b_{m1}} \\ & \dfrac{\partial^2 V}{\partial \Psi^2} & \dfrac{\partial^2 V}{\partial \Psi \partial \theta} & \dfrac{\partial^2 V}{\partial \Psi \partial \theta} & \dfrac{\partial^2 V}{\partial \Psi \partial b_{m1}} \\ & & \dfrac{\partial^2 V}{\partial \theta^2} & \dfrac{\partial^2 V}{\partial \theta \partial \Theta} & \dfrac{\partial^2 V}{\partial \theta \partial b_{m1}} \\ & & & \dfrac{\partial^2 V}{\partial \Theta^2} & \dfrac{\partial^2 V}{\partial \Theta \partial b_{m1}} \\ & & & & \dfrac{\partial^2 V}{\partial b_{m1}{}^2} \end{pmatrix} \tag{3.87}$$

When minimising V($\lambda$) with respect to $\lambda$, the normal equations are non-linear in $\lambda$. The solutions of these equations require the application of Newton Raphson algorithm which iterative equation is given as follows:

$$\lambda_{k+1} = \lambda_k - H^{-1}(\lambda_k)G(\lambda_k) \qquad (3.88)$$

And can be adopted to obtain the (k+1)$^{th}$ iteration $(\hat{\lambda}_{k+1})$ of the estimates from the k$^{th}$ estimate $(\hat{\lambda}_k)$.

By the iterative equations (3.88) which usually converge, the estimates shall be obtained with a good set of initial values of the parameters. This can be achieved by fitting the best full seasonal autoregressive moving average bilinear model.

### 3.15 Implication of estimation technique

Bearing in mind that a nonseasonal bilinear time series model is generally denoted by $(p, d, q, m, n)$ while the pure seasonal form is denoted by $(P, D, Q, m, n)_S$. These have been examined in existing literature. For example, Chikezie (2007) worked on the pure seasonal $(1, 0, 1, 1, 1)_S$ at $s$ =1, 2, 3, 4, 6 and 12 with simulated data illustration. This study however, has extended this to $(P, D, Q, m, 1)_S$ using both real-life and simulated data and further to the mixed seasonal one-dimensional bilinear $(p, 0, q)(P, D, Q, m, 1)_S$ time series model with numerical application using both real-life and simulated data.

### 3.16 Criteria for selecting optimum models

Residual variance

In statistical analysis, the residual of an observed value is defined as the difference between the observed value and the *estimated* value of the quantity of interest (for example, a sample mean). Residual variance on the otherhand, also known as unexplained variance is that part of the variation which cannot be attributed to specific causes. It can generally be grouped into two. That is, the part related to random, everyday, free-will differences in a population or sample and the part that emanates from some un-identified systematic conditions. The latter gives rise to a bias and if not detected would result in a false conclusion.

Akaike's Information Criterion (AIC)

Given several statistical models, the Akaike's information criterion (AIC) compares the quality of a set of models with each other and ranks them from best to worst. The "best" or "optimum" model will be the one that neither under-fits nor over-fits the model. So, it best represents the model. It is a weighted estimation error based on the unexplained variation of a given time series with a penalty term when the optimal number of parameters to represent the system is exceeded. It was developed by Hirotsugu Akaike (1970) and proposed in Akaike (1974). In general;

$$AIC = 2k - 2\ln(L) \tag{3.89}$$

where k is number of parameters and L represents the likelihood function.

The errors are usually assumed to be normally and independently distributed. If n is the number of observations and RSS is the residual sum of squares, then the AIC is given by:

$$AIC = 2k + n\ln(\frac{RSS}{n}) \tag{3.90}$$

Increasing the number of free parameters usually results in better goodness of fit irrespective of the number of free parameters in the data generating process. Therefore, it does not encourage overfitting of parameters and it is often recommended as criterion for comparing models in time series forecasting.

The optimum model is the one with the minimum value of AIC. However, AIC won't say anything about *absolute* quality. That is, if all of our models are poor, it will choose the best of the bad bunch. Therefore in model selection, we consider carrying out a hypothesis test to identify the relationship between the variables of the model and the outcome of interest.

The Bayesian Information Criterion

The Bayesian Information Criterion (BIC) or Schwartz criterion proposed by Schwartz (1978) is another criterion used in selecting an optimum model among a finite set of models. It is defined by the formula:

$$BIC = (\frac{-2\ln L + k\ln(n)}{n}) \qquad (3.91)$$

where $n =$ the number of observations

$k =$ number of free parameters to be estimated

$L =$ maximized the likelihood function for the estimated model.

When the model errors are assumed to be normally distributed, the BIC becomes:

$$BIC = \ln(\frac{RSS}{n}) + k(\frac{\ln n}{n}) \qquad (3.92)$$

Just as it was said concerning the AIC, a model with a lower BIC is preferred when comparing two different models. However, the BIC penalizes free parameters more strongly than the AIC.

## 3.17 Monte Carlo simulation technique

The Monte Carlo approach is a process of deducing conclusions from simulated data. It involves generating free data sets with the aid of computerized mathematical technique, and then stating the values for all independent variables and the parameters. Values are then generated for the error terms for specified sample sizes. With these values, estimates are then calculated for the dependent variable at each sample point. The experiments are then repeated as many times as desired. This is referred to as replications, thereby the experimenter obtains a large number of estimates from the replication. According to Kennedy (1998), there are four stages in a Monte Carlo study namely;

(i) Constructing a model of data generating process

(ii) Creation of sets of data

(iii) Using estimator with the artificial data sets in estimating the model parameters.

(v) Analysis of results.

# CHAPTER FOUR

# RESULTS AND DISCUSSION

## 4.1 Exploratory Data Analysis (EDA):

Monthly Meteorological Rainfall data from 1984 to 2016 recorded in Lagos were obtained from the Nigerian Meteorological Agency (NMA). The data were subjected to Exploratory Data Analysis (EDA) via time plot, ACF plot, PACF plot, Density curve plot, Seasonal trend plot and the Normal QQ plot. For the purpose of In-sample forecast, we used the monthly data from 1984 to 2014 (amounting to 372 monthly data points) which also serves as Out-sample forecast if 2015 and 2016 data are not available and we have considered the specified models at different length/period of seasons (s=1,2,3,4,6 and 12). The results are as follows.

### 4.1.1 Time plot of the monthly rainfall between 1984-2014

It can be inferred from figure 4.1 that there is spurious differences among the monthly rainfall recorded within the thirty-one years due to non-predictable amount of downpour in the so called seasons. Also, it is to be noted that higher, constant and predictable downpour were experienced in Junes of the last 10 years (i.e, Junes of 2005 to 2014) while it varies from 1984 to 2004 but recorded the highest measured in June 1988 followed by June 1990. It can also be deduced visually that there exists non-constant, non-predictable measure of rainfall in each month and the same month succeeding year.

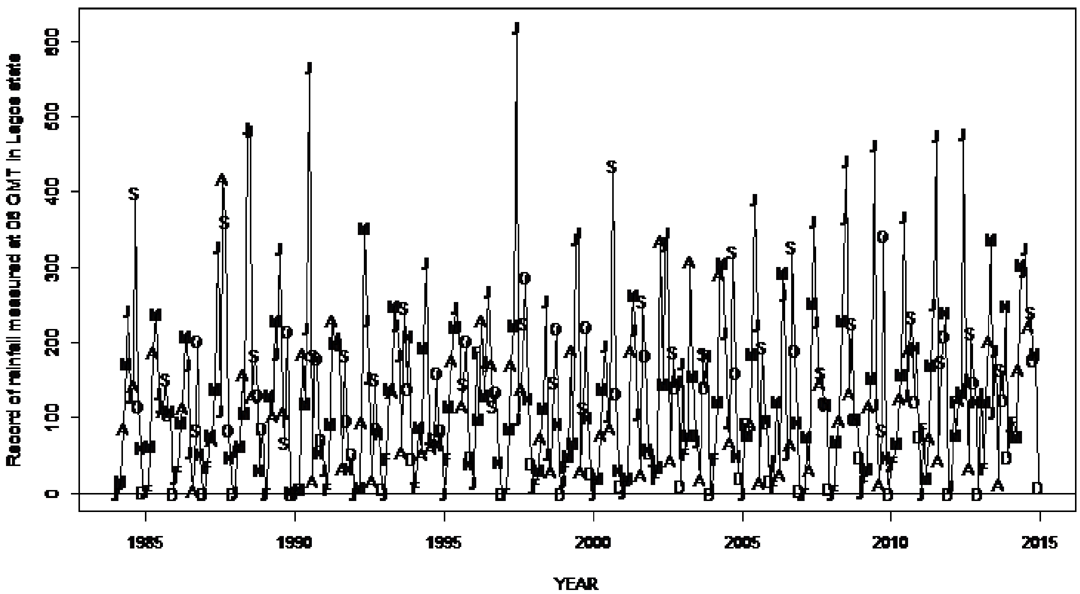


Figure 4.1: Time plot of the monthly rainfall measured in Lagos between 1984-2014.

### 4.1.2 Descriptive Statistics

From table 4.1, it is shown clearly that the Skewness of the measured rainfall data is positively skewed and less than three (<3) which suggested that it is within the scope of normality. That is, the area under the normality curve should be approximately equal 1. Furthermore, the kurtosis of positive 1.640177 suggested it is not heavily tailed from the normal curve as indicated in the modals Density curve in Figure 4.2.

Table 4.1: Major descriptive statistics.

| Minimum | 1st Quarter | Median | Mean | 3rd Quarter | Maximum |
|---|---|---|---|---|---|
| 0.00 | 40.05 | 106.80 | 128.40 | 188.40 | 619.50 |

Skewness= 1.202361

kurtosis= 1.640177

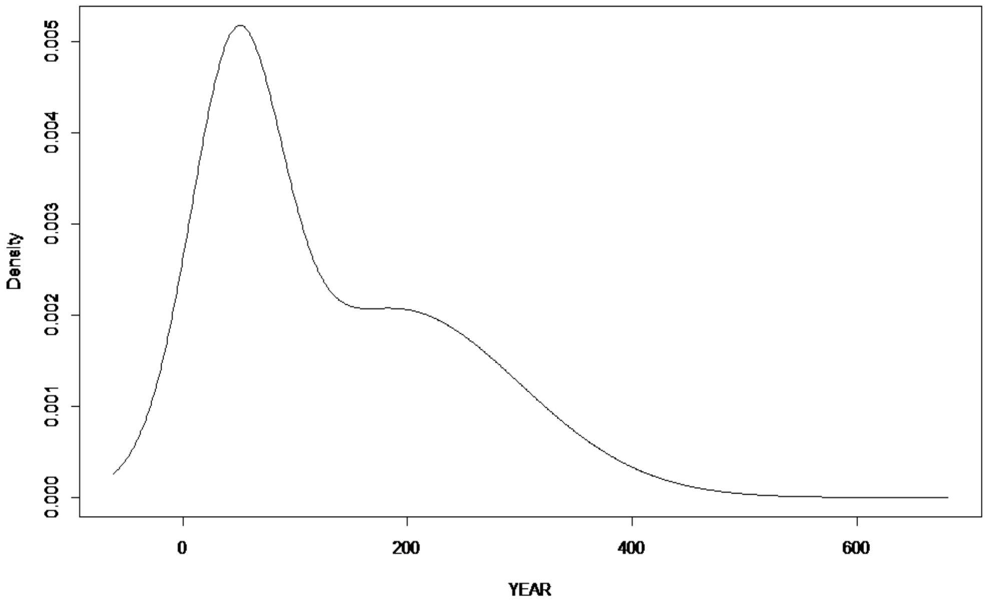


Figure 4.2: Modals Density curve

### 4.1.3 Autocorrelation function (ACF) and Partial Autocorrelation function (PACF)

The ACF and PACF of the original data $\{X_t\}$, t =1, 2, ... , 372, are as shown in Figure 4.3 and Figure4.4. They both show a seasonal fluctuation. Concentrating on the ACF of the original data, we note a slow decreasing trend in the ACF and it peaks at seasonal lags, h = 1s, 2s, 3s, 4s, where s = 12. This indicates a nonstationary behaviour and suggests a seasonal difference (Wang, 2008; Momani and Naill, 2009).

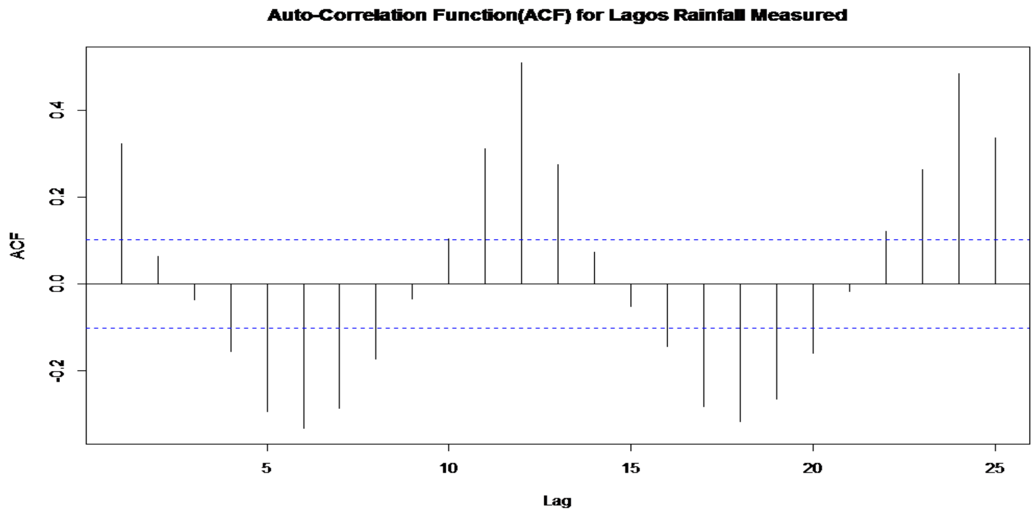


Figure 4.3 ACF plot for the rainfall data

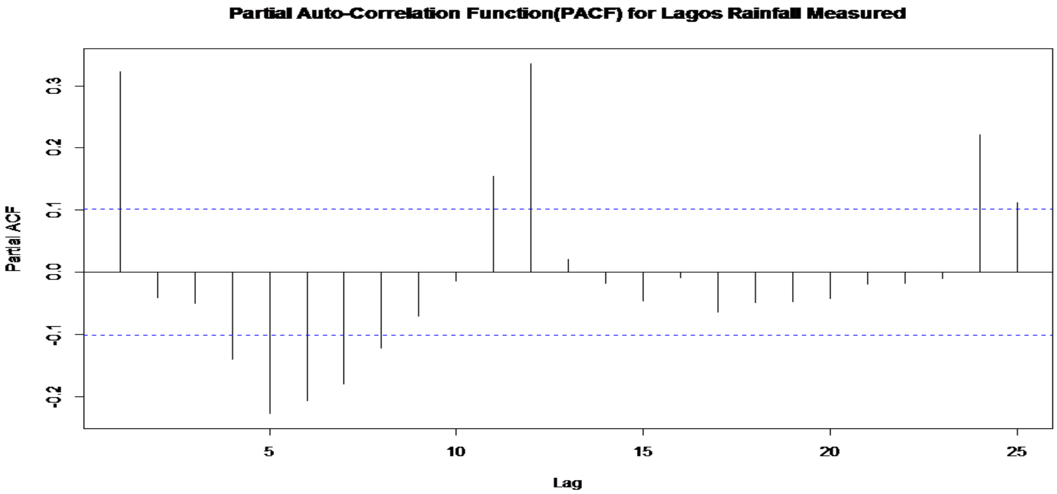


Figure 4.4: PACF for the rainfall data

.

### 4.1.4: The normal QQ plot

Looking at the Normal QQ plot in figure 4.5, the points fall along a line in the middle of the graph, but curve off in the extremities. This shows that the data have more extreme values than would be expected and hence truly came from a Normal distribution.

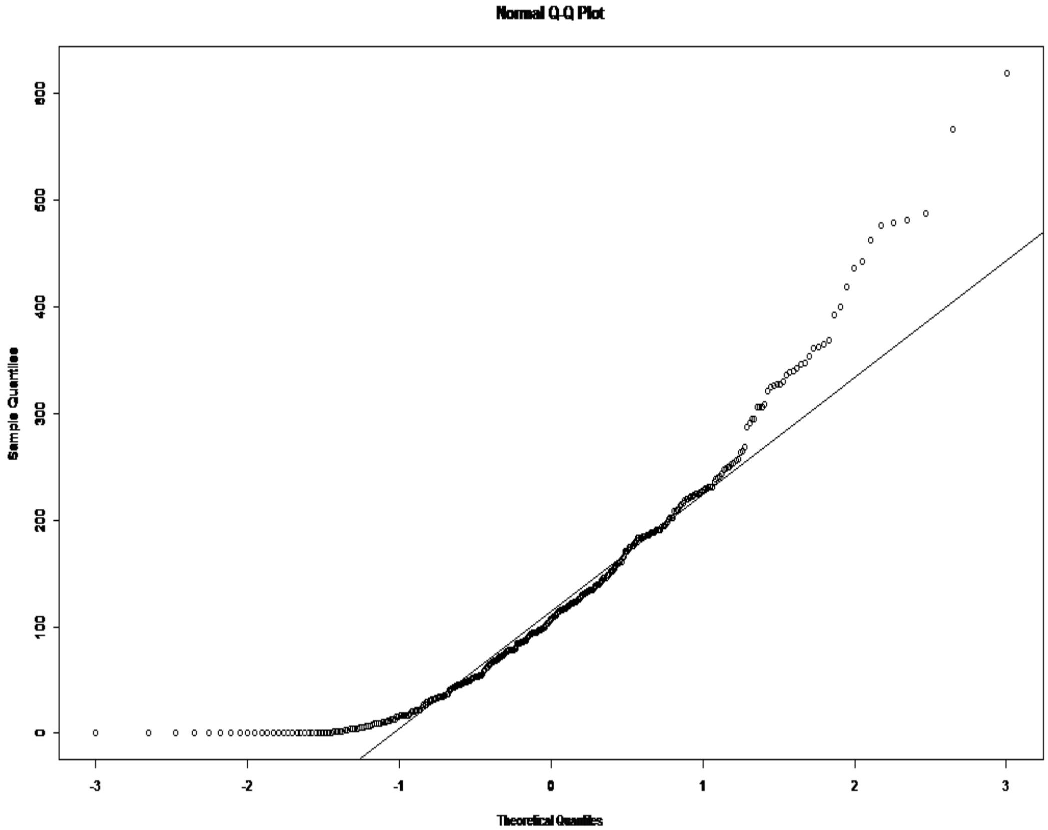


Figure 4.5: Normal QQ plot

## 4.2 Seasonality test (Hegy Test)

The famous HEGY test is carried out to test the seasonality effect because of the aforementioned seasonal pattern noticed in Figure 4.3.

Hypothesis:

$H_0$: $\pi_i = 0$

$H_1$: $\pi_i \neq 0$

Level of significance, $\alpha = 0.05$

Test statistic: The testing procedure for seasonal unit roots involves estimating via OLS the following regression:

$$\Delta^s X_t = \alpha + \beta t + \sum_{j=2}^{s} b_j Q_{jt} + \sum_{i=1}^{s} \pi_i W_{it-1} + \sum_{\ell=1}^{k} \gamma_\ell \Delta^s Y_{t-\ell} + \varepsilon_t$$

Decision: Since the calculated value 0.0105 < 0.05, we reject the null hypothesis Conclusion: the monthly seasonal effect reveals that seasonal effect is present. Hence, seasonal effect is present in the meteorological rainfall measured for the years considered.

## 4.3 Keenan's Nonlinearity test.

The Keenan's (1985) Nonlinearity test was carried out and the result is as follows.

Hypotheses:

$$H_0: \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \theta_{uv} e_{t-u} e_{t-v} = 0$$

$$H_1: \sum_{u=-\infty}^{\infty} \sum_{v=-\infty}^{\infty} \theta_{uv} e_{t-u} e_{t-v} \neq 0$$

Level of significance; $\propto = 0.05$

Test statistic:

$$\hat{F} = \frac{(n-2p-2)\hat{\eta}^2}{(SSR-\hat{\eta}^2)}$$

Test Results:

Test statistic = 9.0113

Critical value = 2.80615

Decision: since the test statistic value (9.0113) is > the critical value (2.80615), we do not accept the null hypothesis.

Conclusion: the series is nonlinear.

## 4.4 Fitted pure seasonal autoregressive integrated models

Here, we first fitted different models with different orders of $P$, for the linear Pure seasonal Autoregressive integrated time series at the different length of seasons (s=1,2,3,4,6 and 12). Optimum models are selected based on the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), as well as the residual variances. Model with minimum AIC, BIC and residual variance is the optimum model in each case. They are $(0,0,0)(3,2,0)_{1}$, $(0,0,0)(3,2,0)_{2}$, $(0,0,0)(3,2,0)_{3}$, $(0,0,0)(3,2,0)_{4}$, $(0,0,0)(3,2,0)_{6}$ and $(0,0,0)(3,2,0)_{12}$ as shown in table 4.2 and table 4.3. This shows that optimum models are obtained when the seasonal order $P$ =3 in each case.

Table 4.2: Fitted pure seasonal autoregressive integrated models.

| s/n | MODEL | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(0,0,0)(1,2,0)_1$ | 11.30858 | 10.31911 | 25.4513 |
| 2 | $(0,0,0)(2,2,0)_1$ | 11.08261 | 10.10368 | 25.1056 |
| **3** | **$(0,0,0)(3,2,0)_1$** | **10.94843** | **9.980037** | **25.0034** |
| 4 | $(0,0,0)(4,2,0)_1$ | 11.87543 | 10.56843 | 25.67432 |
| 1 | $(0,0,0)(1,2,0)_2$ | 11.58023 | 10.59077 | 25.6724 |
| 2 | $(0,0,0)(2,2,0)_2$ | 11.45904 | 10.48011 | 24.8903 |
| **3** | **$(0,0,0)(3,2,0)_2$** | **11.36895** | **10.40056** | **24.3702** |
| 4 | $(0,0,0)(4,2,0)_2$ | 11.88665 | 10.76540 | 25.7543 |
| 1 | $(0,0,0)(1,2,0)_3$ | 11.85232 | 10.86285 | 25.9856 |
| 2 | $(0,0,0)(2,2,0)_3$ | 11.65839 | 10.67946 | 25.7805 |
| **3** | **$(0,0,0)(3,2,0)_3$** | **10.850004** | **9.881645** | **24.6045** |
| 4 | $(0,0,0)(4,2,0)_3$ | 11.92346 | 10.99854 | 26.0852 |
| 1 | $(0,0,0)(1,2,0)_4$ | 12.00422 | 11.01476 | 26.29220 |
| 2 | $(0,0,0)(2,2,0)_4$ | 10.92725 | 9.94832 | 24.2516 |
| 3 | **$(0,0,0)(3,2,0)_4$** | **10.70965** | **9.741258** | **24.1784** |
| 4 | $(0,0,0)(4,2,0)_4$ | 12.3462 | 11.7632 | 26.6378 |
| 1 | $(0,0,0)(1,2,0)_6$ | 11.04643 | 10.05696 | 24.8044 |
| 2 | $(0,0,0)(2,2,0)_6$ | 10.93671 | 9.957778 | 24.1034 |
| **3** | **$(0,0,0)(3,2,0)_6$** | **10.56727** | **9.598878** | **24.0087** |
| 4 | $(0,0,0)(4,2,0)_6$ | 11.65832 | 10.37559 | 24.9348 |
| 1 | $(0,0,0)(1,2,0)_{12}$ | 10.95411 | 9.964642 | 24.3171 |
| 2 | $(0,0,0)(2,2,0)_{12}$ | 10.52465 | 9.54572 | 23.9995 |
| **3** | **$(0,0,0)(3,2,0)_{12}$** | **10.38454** | **9.416143** | **23.6734** |
| 4 | $(0,0,0)(4,2,0)_{12}$ | 10.87652 | 10.21784 | 24.6372 |

The parameters of each of these optimum models are then estimated as follows.

Fitted Pure Seasonal autoregressive Integrated model at s=1**:**

$$\hat{X}_t = \Psi_1 X_{t-1} + \Psi_2 X_{t-2} + \Psi_3 X_{t-3}$$

where: $\Psi_1 = -0.9973$, $\Psi_2 = -0.7530$, $\Psi_3 = -0.3588$

AIC=10.94843,

BIC=9.980037

Residual Variance = 25.0034

Fitted Pure Seasonal autoregressive Integrated model at s=2:

$$\hat{X}_t = \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Psi_3 X_{t-6}$$

$\Psi_1 = -0.9503$ , $\Psi_2 = -0.5984$ , $\Psi_3 = -0.3000$

AIC=11.36895

BIC=10.40056

Residual Variance = 24.3702

Fitted Pure Seasonal autoregressive Integrated model at s=3:

$$\hat{X}_t = \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Psi_3 X_{t-9}$$

$\Psi_1 = -0.0654$ , $\Psi_2 = -0.9871$ , $\Psi_3 = -0.7440$

AIC=10.85004

BIC=9.881645

Residual Variance = 24.6045

Fitted Pure Seasonal autoregressive Integrated model at s=4:

$$\hat{X}_t = \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Psi_3 X_{t-12}$$

$\Psi_1 = -0.3819$, $\Psi_2 = -0.2667$ , $\Psi_3 = -0.4453$

AIC= 10.70965

BIC= 9.741258

Residual Variance = 24.1784

Fitted Pure Seasonal autoregressive Integrated model at s=6:

$$\hat{X}_t = \Psi_1 X_{t-6} + \Psi_2 X_{t-12} + \Psi_3 X_{t-18}$$

$\Psi_1 = -0.3818$ , $\Psi_2 = -0.9926$ , $\Psi_3 = -0.5500$

AIC= 10.56727

BIC= 9.598878

Residual Variance = 24.0087

Fitted Pure Seasonal autoregressive Integrated model at s=12:

$$\hat{X}_t = \Psi_1 X_{t-12} + \Psi_2 X_{t-24} + \Psi_3 X_{t-36}$$

$\Psi_1 = -0.2423$, $\Psi_2 = -0.9811$ , $\Psi_3 = -0.3797$

AIC= 10.38454

BIC= 9.416143

Residual Variance = 23.6734

Table 4.3: Fitted optimum pure seasonal autoregressive integrated models

| S/N | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(0,0,0)(3,2,0)_1$ | 10.94843 | 9.980037 | 25.0034 |
| 2 | $(0,0,0)(3,2,0)_2$ | 11.36895 | 10.40056 | 24.3702 |
| 3 | $(0,0,0)(3,2,0)_3$ | 10.850004 | 9.881645 | 24.6045 |
| 4 | $(0,0,0)(3,2,0)_4$ | 10.70965 | 9.741258 | 24.1784 |
| 5 | $(0,0,0)(3,2,0)_6$ | 10.56727 | 9.598878 | 24.0087 |
| 6 | $(0,0,0)(3,2,0)_{12}$ | 10.38454 | 9.416143 | 23.6734 |

## 4.5 Fitted Pure Seasonal Autoregressive Integrated One-Dimensional Bilinear Models (Order one)

The nonlinear parameters of the model are estimated as follows.

Fitted pure Seasonal autoregressive integrated bilinear model of order one at s=1.

$$\hat{X}_t = \Psi_1 X_{t-1} + \Psi_2 X_{t-2} + \Psi_3 X_{t-3} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\Psi_1 = -0.9973,\ \Psi_2 = -0.7530,\ \Psi_3 = -0.3588,\ b_{11} = 0.1007$

AIC= 9.7823,

BIC= 8.5624

Residual Variance = 0.00947

Fitted Pure Seasonal autoregressive integrated bilinear model of order one at s=2:

$$\hat{X}_t = \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Psi_3 X_{t-6} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\Psi_1 = -0.9503$ , $\Psi_2 = -0.5984$ , $\Psi_3 = -0.3000,\ b_{11} = 0.7386$

AIC= 10.8023,

BIC= 9.9025,

Residual Variance = 0.01047

Fitted pure Seasonal autoregressive integrated bilinear model of order one at s=3:

$$\hat{X}_t = \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Psi_3 X_{t-9} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\Psi_1 = -0.0654$ , $\Psi_2 = -0.9871$ , $\Psi_3 = -0.7440,\ b_{11} = -0.8021$

AIC= 9.8296**,**

BIC= 8.9032,

Residual Variance = 0.06241

Fitted Seasonal autoregressive integrated bilinear model of order one at s=4:

$$\hat{X}_t = \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Psi_3 X_{t-12} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\Psi_1 = -0.3819,$ $\Psi_2 = -0.2667$ , $\Psi_3 = -0.4453,\ b_{11} = -0.9030$

AIC= 10.3892**,**

BIC= 9.5369,

Residual Variance = 0.089241

Fitted Pure Seasonal autoregressive integrated bilinear model of order one at s=6:

$$\hat{X}_t = \Psi_1 X_{t-1} + \Psi_2 X_{t-12} + \Psi_3 X_{t-18} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\Psi_1 = -0.3818$ , $\Psi_2 = -0.9926$ , $\Psi_3 = -0.5500$, $b_{11} = 0.4840$

AIC= 9.9578**,**

BIC= 8.8034

Residual Variance = 0.097241

Fitted Seasonal autoregressive integrated bilinear model of order one at s=12:

$$\hat{X}_t = \Psi_1 X_{t-12} + \Psi_2 X_{t-24} + \Psi_3 X_{t-36} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\Psi_1 = -0.2423$, $\Psi_2 = -0.9811$ , $\Psi_3 = -0.3797$, $b_{11} = -0.7943$

AIC= 9.99478,

BIC= 8.9652

Residual Variance = 0.097241

From table 4.4, it is observed that the least AIC, BIC and residual variance with respective values 9.7823, 8.5624 and 0.00947 are obtained when the length of season $s = 1$ (in bold print). This means that the model, $(0,0,0)(3,2,0,1,1)_1$ performs best while the model $(0,0,0)(3,2,0,1,1)_2$ with the highest AIC value of 10.8023, BIC value of 9.9025 and residual variance 0.01047 performs most poorly among the estimated models.

Table 4.4: Fitted optimum pure seasonal autoregressive integrated one-dimensional bilinear models (Order one)

| S/N | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| **1** | **$(0,0,0)(3,2,0,1,1)_1$** | **9.7823** | **8.5624** | **0.00947** |
| 2 | $(0,0,0)(3,2,0,1,1)_2$ | 10.8023 | 9.9025 | 0.01047 |
| 3 | $(0,0,0)(3,2,0,1,1)_3$ | 9.8296 | 8.9032 | 0.06241 |
| 4 | $(0,0,0)(3,2,0,1,1)_4$ | 10.3892 | 9.5369 | 0.089241 |
| 5 | $(0,0,0)(3,2,0,1,1)_6$ | 9.9578 | 8.8034 | 0.097241 |
| 6 | $(0,0,0)(3,2,0,1,1)_{12}$ | 9.99478 | 8.9652 | 0.127241 |

## 4.6 Fitted MIXED Seasonal Autoregressive Integrated models

Similarly, the linear Mixed Seasonal Autoregressive Integrated models are fitted at different orders of the non-seasonal parameter, *p* and the seasonal parameter *P* at the different length of seasons (s = 1, 2, 3, 4, 6 and 12). Optimum models are selected based on the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), as well as the residual variances. Model with minimum AIC, BIC and residual variance is the optimum model in each case. They are $(2,0,0)(3,2,0)_1$, $(3,0,0)(3,2,0)_{2,}$ $(3,0,0)(3,2,0)_{3,}$ $(1,0,0)(3,2,0)_{4,}$ $(3,0,0)(3,2,0)_6$ and $(2,0,0)(2,2,0)_{12}$ as printed in bold figures in table 4.5 and presented in table 4.6.

Table 4.5: Mixed seasonal autoregressive integrated models are fitted at different orders *p* and *P*.

| s/n | MODEL | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(1,0,0)(1,2,0)_1$ | 11.21465 | 10.23572 | 23.023 |
| 2 | $(2,0,0)(1,2,0)_1$ | 10.94843 | 9.980038 | 22.2064 |
| 3 | $(3,0,0)(1,2,0)_1$ | 10.9167 | 9.95884 | 22.2001 |
| 4 | $(1,0,0)(2,2,0)_1$ | 10.94843 | 9.980038 | 22.2202 |
| 5 | $(2,0,0)(2,2,0)_1$ | 10.97537 | 10.01751 | 22.2223 |
| 6 | $(3,0,0)(2,2,0)_1$ | 10.88026 | 9.932929 | 22.2209 |
| 7 | $(1,0,0)(3,2,0)_1$ | 10.9167 | 9.95884 | 22.20008 |
| **8** | $\mathbf{(2,0,0)(3,2,0)_1}$ | **10.88026** | **9.932929** | **22.0145** |
| 9 | $(3,0,0)(3,2,0)_1$ | 10.89343 | 9.956636 | 22.0195 |
| 1 | $(1,0,0)(1,2,0)_2$ | 11.52478 | 10.54585 | 25.0891 |
| 2 | $(2,0,0)(1,2,0)_2$ | 11.42977 | 10.46137 | 25.894 |
| 3 | $(3,0,0)(1,2,0)_2$ | 11.42062 | 10.46276 | 25.0035 |
| 4 | $(1,0,0)(2,2,0)_2$ | 11.3443 | 10.37591 | 24.8927 |
| 5 | $(2,0,0)(2,2,0)_2$ | 11.30734 | 10.34948 | 24.6729 |
| 6 | $(3,0,0)(2,2,0)_2$ | 11.21947 | 10.27215 | 24.6060 |
| 7 | $(1,0,0)(3,2,0)_2$ | 11.21602 | 10.25816 | 24.2461 |
| 8 | $(2,0,0)(3,2,0)_2$ | 11.20385 | 10.25652 | 24.2441 |
| **9** | $\mathbf{(3,0,0)(3,2,0)_2}$ | **11.19589** | **10.2591** | **24.1074** |
| 10 | $(4,0,0)(3,2,0)_2$ | 11.75678 | 10.3623 | 24.3212 |
| 1 | $(1,0,0)(1,2,0)_3$ | 11.21465 | 10.23572 | 26.5789 |
| 2 | $(2,0,0)(1,2,0)_3$ | 11.66298 | 10.69459 | 27.36048 |
| 3 | $(3,0,0)(1,2,0)_3$ | 11.57961 | 10.62175 | 26.7803 |
| 4 | $(1,0,0)(2,2,0)_3$ | 11.48644 | 10.51805 | 26.5389 |
| 5 | $(2,0,0)(2,2,0)_3$ | 11.48316 | 10.5253 | 26.5050 |
| 6 | $(3,0,0)(2,2,0)_3$ | 10.85683 | 9.909505 | 23.070 |
| 7 | $(1,0,0)(3,2,0)_3$ | 10.83393 | 9.87607 | 23.0450 |
| 8 | $(2,0,0)(3,2,0)_3$ | 10.83579 | 9.888459 | 23.0089 |
| **9** | $\mathbf{(3,0,0)(3,2,0)_3}$ | **10.62372** | **9.68693** | **22.8999** |

| | | | | |
|---|---|---|---|---|
| 10 | $(4,0,0)(3,2,0)_3$ | 10.99674 | 9.87652 | 23.0075 |
| 1 | $(1,0,0)(1,2,0)_4$ | 11.77003 | 10.7911 | 26.6292 |
| 2 | $(2,0,0)(1,2,0)_4$ | 11.77457 | 10.80617 | 26.7805 |
| 3 | $(3,0,0)(1,2,0)_4$ | 11.75334 | 10.79548 | 26.6111 |
| 4 | $(1,0,0)(2,2,0)_4$ | 10.92394 | 9.95554 | 25.3901 |
| 5 | $(2,0,0)(2,2,0)_4$ | 10.92913 | 9.971267 | 25.3930 |
| 6 | $(3,0,0)(2,2,0)_4$ | 10.9345 | 9.987172 | 25.3938 |
| **7** | $\mathbf{(1,0,0)(3,2,0)_4}$ | **10.70556** | **9.747698** | **24.1110** |
| 8 | $(2,0,0)(3,2,0)_4$ | 10.70973 | 9.762399 | 24.1779 |
| 9 | $(3,0,0)(3,2,0)_4$ | 10.71502 | 9.778229 | 24.1787 |
| 1 | $(1,0,0)(1,2,0)_6$ | 11.04854 | 10.0696 | 25.0780 |
| 2 | $(2,0,0)(1,2,0)_6$ | 11.05081 | 10.08241 | 25.0698 |
| 3 | $(3,0,0)(1,2,0)_6$ | 11.05591 | 10.09805 | 25.0728 |
| 4 | $(1,0,0)(2,2,0)_6$ | 10.93478 | 9.966379 | 24.7834 |
| 5 | $(2,0,0)(2,2,0)_6$ | 10.93799 | 9.980128 | 24.7888 |
| 6 | $(3,0,0)(2,2,0)_6$ | 10.94314 | 9.995814 | 24.7890 |
| 7 | $(1,0,0)(3,2,0)_6$ | 10.65797 | 9.610109 | 25.8008 |
| 8 | $(2,0,0)(3,2,0)_6$ | 10.76969 | 9.622367 | 25.8008 |
| **9** | $\mathbf{(3,0,0)(3,2,0)_6}$ | **10.5957** | **9.637975** | **25.8019** |
| 10 | $(4,0,0)(3,2,0)_6$ | 10.7890 | 9.837626 | 25.9452 |
| 1 | $(1,0,0)(1,2,0)_{12}$ | 10.95872 | 9.97979 | 25.8013 |
| 2 | $(2,0,0)(1,2,0)_{12}$ | 10.95562 | 9.987229 | 25.5723 |
| 3 | $(3,0,0)(1,2,0)_{12}$ | 10.96085 | 10.00299 | 25.5745 |
| 4 | $(1,0,0)(2,2,0)_{12}$ | 10.93478 | 9.966379 | 24.7834 |
| **5** | $\mathbf{(2,0,0)(2,2,0)_{12}}$ | **10.52747** | **9.569606** | **25.1490** |
| 6 | $(3,0,0)(2,2,0)_{12}$ | 10.52802 | 9.580692 | 25.1495 |
| 7 | $(1,0,0)(3,2,0)_{12}$ | 11.39911 | 10.441252 | 25.1089 |
| 8 | $(2,0,0)(3,2,0)_{12}$ | 11.39293 | 10.445602 | 25.1066 |
| 9 | $(3,0,0)(3,2,0)_{12}$ | 11.39531 | 10.458516 | 25.1202 |

The parameters of the optimum models at the different length of seasons are estimated as follows.

Fitted Mixed Seasonal autoregressive integrated model at s=1:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \Psi_1 X_{t-1} + \Psi_2 X_{t-2} + \Psi_3 X_{t-3}$$

$\psi_1 = -0.7581$, $\psi_2 = -0.4532$, $\Psi_1 = -0.3574$, $\Psi_2 = -0.2936$,

$\Psi_3 = -0.3473$

AIC=10.88026**,**

BIC=9.932929

Residual Variance = 22.0145

Fitted Mixed Seasonal autoregressive integrated model at s=2:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_3 X_{t-3} + \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Psi_3 X_{t-6}$$

$\psi_1 = 0.4113$, $\psi_2 = -0.5273$, $\psi_3 = 0.2328$, $\Psi_1 = -0.6599$, $\Psi_2 = -0.5495$,

$\Psi_3 = -0.2356$

AIC=11.19589,

BIC=10.2591

Residual Variance = 24.1074

Fitted Mixed Seasonal autoregressive integrated model at s=3:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_3 X_{t-3} + \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Psi_3 X_{t-9}$$

$\psi_1 = 0.0512$, $\psi_2 = -0.0451$, $\psi_3 = -0.6521$, $\Psi_1 = -0.7097$,

$\Psi_2 = -0.9082$,

$\Psi_3 = -0.5847$

AIC=10.62372,

BIC=9.68693

Residual Variance = 22.8999

Fitted Mixed Seasonal autoregressive integrated model at s=4:

$$\hat{X}_t = \psi_1 X_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Psi_3 X_{t-12}$$

$\psi_1 = 0.1128$, $\Psi_1 = -0.3717$, $\Psi_2 = -0.2467$, $\Psi_3 = -0.4452$

AIC=10.70556,

BIC=9.740698,

Residual Variance = 24.1110

Fitted Mixed Seasonal autoregressive integrated model at s=6:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_3 X_{t-3} + \Psi_1 X_{t-6} + \Psi_2 X_{t-12} + \Psi_3 X_{t-18}$$

$\psi_1 = 0.0802$, $\psi_2 = -0.0559$, $\psi_3 = -0.0167$, $\Psi_1 = -0.3805$,

$\Psi_2 = -0.9945$, $\Psi_3 = -0.5498$

AIC=10.5957,

BIC=9.637975,

Residual Variance = 25.8019

Fitted Mixed Seasonal autoregressive integrated model at s=12:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \Psi_1 X_{t-12} + \Psi_2 X_{t-24}$$

$\psi_1 = 0.0268$, $\psi_2 = -0.0844$, $\Psi_1 = -0.0247$, $\Psi_2 = -0.602$

AIC=10.52747,

BIC=9.569606

Residual Variance = 25.1490

Table 4.6: Optimum mixed seasonal autoregressive integrated models

| s/n | MODEL | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(2,0,0)(3,2,0)_1$ | 10.88026 | 9.932929 | 22.0145 |
| 2 | $(3,0,0)(3,2,0)_2$ | 11.19589 | 10.2591 | 24.1074 |
| 3 | $(3,0,0)(3,2,0)_3$ | 10.62372 | 9.68693 | 22.8999 |
| 4 | $(1,0,0)(3,2,0)_4$ | 10.70556 | 9.740698 | 24.1110 |
| 5 | $(3,0,0)(3,2,0)_6$ | 10.5957 | 9.637975 | 25.8019 |
| 6 | $(2,0,0)(2,2,0)_{12}$ | 10.52747 | 9.569606 | 25.1490 |

### 4.7 Fitted Mixed Seasonal Autoregressive Integrated One-dimensional Bilinear Models (Order one)

The initial parameters obtained from the linear mixed seasonal autoregressive integrated models above are then used as initial parameters for fitting the nonlinear bilinear models in table 4.7 and the parameters of the fitted model are as presented as follows.

The parameters of the mixed SARI one-dimensional bilinear models are estimated as follows:

Fitted Mixed Seasonal autoregressive integrated bilinear model of order one at s=1:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \Psi_1 X_{t-1} + \Psi_2 X_{t-2} + \Psi_3 X_{t-3} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\psi_1 = -0.7581$, $\psi_2 = -0.4532$, $\Psi_1 = -0.3574$, $\Psi_2 = -0.2936$,

$\Psi_3 = -0.3473$, $b_{11} = 0.1007$

AIC= 7.0064,

BIC= 6.6015

Residual Variance = 0.000893

Fitted Mixed Seasonal autoregressive integrated bilinear model of order one at s=2:

$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_3 X_{t-3} + \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Psi_3 X_{t-6} + b_{11} X_{t-1}\varepsilon_{t-1}$ $\psi_1 =$ 0.4113, $\psi_2 = -0.5273$, $\psi_3 = 0.2328$, $\Psi_1 = -0.6599$, $\Psi_2 = -0.5495$,

$\Psi_3 = -0.2356$, $b_{11} = -0.7916$

AIC= 9.2019,

BIC=8.8052

Residual Variance = 0.00563

Fitted Mixed Seasonal autoregressive integrated bilinear model of order one at s=3:

$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_3 X_{t-3} + \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Psi_3 X_{t-9} + b_{11} X_{t-1}\varepsilon_{t-1}$ $\psi_1 =$ 0.0512, $\psi_2 = -0.0451$, $\psi_3 = -0.6521$, $\Psi_1 = -0.7097$, $\Psi_2 = -0.9082$,

$\Psi_3 = -0.5847$, $b_{11} = -0.8261$

AIC= 8.6201**,**

BIC=7.8062

Residual Variance = 0.02289

Fitted Mixed Seasonal autoregressive integrated bilinear model of order one at s=4:

$$\hat{X}_t = \psi_1 X_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Psi_3 X_{t-12} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\psi_1 = 0.1128$, $\Psi_1 = -0.3717$, $\Psi_2 = -0.2467$, $\Psi_3 = -0.4452$,

$b_{11} = 0.7941$

AIC= 7.9876

BIC=6.9011

Residual Variance = 0.000998

Fitted Mixed Seasonal autoregressive integrated bilinear model of order one at s=6:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_3 X_{t-3} + \Psi_1 X_{t-6} + \Psi_2 X_{t-12} + \Psi_3 X_{t-18} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\psi_1 = 0.0802$, $\psi_2 = -0.0559$, $\psi_3 = -0.0167$, $\Psi_1 = -0.3805$, $\Psi_2 = -0.9945$,

$\Psi_3 = -0.5498$, $b_{11} = 0.9012$

AIC= 8.7843

BIC=9.0990

Residual Variance = 0.005967

Fitted Mixed Seasonal autoregressive integrated bilinear model of order one at s=12:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \Psi_1 X_{t-12} + \Psi_2 X_{t-24} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\psi_1 = 0.0268$, $\psi_2 = -0.0844$, $\Psi_1 = -0.0247$, $\Psi_2 = -0.602$,

$b_{11} = -0.3270$

AIC= 8.6902

BIC=7.6598

Residual Variance = 0.00837

Table 4.7: Fitted mixed seasonal autoregressive integrated one-dimensional Bilinear models

| s/n | MODEL | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(2,0,0)(3,2,0,1,1)_1$ | 7.0064 | 6.6015 | 0.000893 |
| 2 | $(3,0,0)(3,2,0,1,1)_2$ | 9.2019 | 8.8052 | 0.00563 |
| 3 | $(3,0,0)(3,2,0,1,1)_3$ | 8.6201 | 7.8062 | 0.02289 |
| 4 | $(1,0,0)(3,2,0,1,1)_4$ | 7.9876 | 6.9011 | 0.000998 |
| 5 | $(3,0,0)(3,2,0,1,1)_6$ | 8.7843 | 9.0990 | 0.005967 |
| 6 | $(2,0,0)(2,2,0,1,1)_{12}$ | 8.6902 | 7.6598 | 0.00837 |

## 4.8 Fitted pure seasonal autoregressive integrated moving average models

In this section, we introduced the Moving average component. In a similar manner, linear pure seasonal autoregressive integrated moving average models of different orders of seasonal autoregressive parameter *P* and the seasonal moving average, *Q* at the different length of seasons (s=1,2,3,4,6 and 12) are obtained as shown in table 4.8. Optimum models (in bold prints in table 4.8) are selected based on the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), as well as the residual variances. These optimum models are $(0,0,0)(1,2,2)_1$, $(0,0,0)(3,2,2)_{2,}$ $(0,0,0)(3,2,3)_{3,}$ $(0,0,0)(3,2,3)_{4,}$ $(0,0,0)(1,2,3)_6$ and $(0,0,0)(1,2,3)_{12.}$ Their estimated parameters are then used as initial parameters for fitting the nonlinear, pure seasonal autoregressive integrated moving average one-dimensional bilinear time series models. The optimum models are presented in tables 4.9.

Table 4.8: Fitted Pure Seasonal Autoregressive Integrated Moving Average (SARIMA ) models.

| s/n | MODEL | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(0,0,0)(1,2,1)_1$ | 10.65373 | 9.674799 | 24.6728 |
| **2** | $\mathbf{(0,0,0)(1,2,2)_1}$ | **10.35153** | **9.383136** | **23.0089** |
| 3 | $(0,0,0)(1,2,3)_1$ | 10.35532 | 9.397454 | 24.6783 |
| 4 | $(0,0,0)(2,2,1)_1$ | 10.60387 | 9.635475 | 25.8245 |
| 5 | $(0,0,0)(2,2,2)_1$ | 10.35856 | 9.400698 | 22.8747 |
| 6 | $(0,0,0)(2,2,3)_1$ | 10.35964 | 9.41231 | 24.6109 |
| 7 | $(0,0,0)(3,2,1)_1$ | 10.59500 | 9.637143 | 26.1208 |
| 8 | $(0,0,0)(3,2,2)_1$ | 10.35793 | 9.410607 | 24.7731 |
| 9 | $(0,0,0)(3,2,3)_1$ | 10.3636 | 9.426809 | 25.7054 |
| 1 | $(0,0,0)(1,2,1)_2$ | 10.92497 | 9.946036 | 24.7719 |
| 2 | $(0,0,0)(1,2,2)_2$ | 10.46178 | 9.493385 | 23.7429 |
| 3 | $(0,0,0)(1,2,3)_2$ | 10.62359 | 9.413852 | 26.0128 |
| 4 | $(0,0,0)(2,2,1)_2$ | 10.89146 | 9.923061 | 25.8418 |
| 5 | $(0,0,0)(2,2,2)_2$ | 10.43718 | 9.479317 | 22.9153 |
| 6 | $(0,0,0)(2,2,3)_2$ | 10.38495 | 9.437624 | 24.8610 |
| 7 | $(0,0,0)(32,1,)_2$ | 10.78385 | 9.825993 | 26.2398 |
| **8** | $\mathbf{(0,0,0)(3,2,2)_2}$ | **10.33298** | **9.385657** | **24.5390** |
| 9 | $(0,0,0)(3,2,3)_2$ | 10.35229 | 9.415497 | 25.6823 |
| 1 | $(0,0,0)(1,2,1)_3$ | 11.05652 | 10.07759 | 25.0049 |
| 2 | $(0,0,0)(1,2,2)_3$ | 10.4536 | 9.485208 | 23.7002 |
| 3 | $(0,0,0)(1,2,3)_3$ | 10.45927 | 9.50141 | 24.5634 |
| 4 | $(0,0,0)(2,2,1)_3$ | 10.80178 | 9.833388 | 25.6180 |
| 5 | $(0,0,0)(2,2,2)_3$ | 10.31755 | 9.359688 | 22.8141 |

| | | | | |
|---|---|---|---|---|
| 6 | $(0,0,0)(2,2,3)_3$ | 10.32517 | 9.377848 | 24.5310 |
| 7 | $(0,0,0)(3,2,1)_3$ | 10.22412 | 9.266262 | 25.6510 |
| 8 | $(0,0,0)(3,2,2)_3$ | 10.22299 | 9.275668 | 24.4169 |
| **9** | $\mathbf{(0,0,0)(3,2,3)_3}$ | **10.19405** | **9.257262** | **24.9835** |
| 1 | $(0,0,0)(1,2,1)_4$ | 11.02538 | 10.04645 | 25.01008 |
| 2 | $(0,0,0)(1,2,2)_4$ | 10.35003 | 9.381636 | 23.6211 |
| 3 | $(0,0,0)(1,2,3)_4$ | 10.39528 | 9.437422 | 24.5341 |
| 4 | $(0,0,0)(2,2,1)_4$ | 10.26358 | 9.295181 | 24.6714 |
| 5 | $(0,0,0)(2,2,2)_4$ | 10.25819 | 9.300327 | 22.7858 |
| 6 | $(0,0,0)(2,2,3)_4$ | 10.17048 | 9.223154 | 24.2812 |
| 7 | $(0,0,0)(3,2,1)_4$ | 10.26265 | 9.304785 | 25.8075 |
| 8 | $(0,0,0)(3,2,2)_4$ | 10.26147 | 9.314143 | 24.3989 |
| **9** | $\mathbf{(0,0,0)(3,2,3)_4}$ | **10.07154** | **9.134746** | **23.3284** |
| 1 | $(0,0,0)(1,2,1)_6$ | 10.31826 | 9.339328 | 24.1721 |
| 2 | $(0,0,0)(1,2,2)_6$ | 10.20121 | 9.232814 | 23.3004 |
| 3 | $\mathbf{(0,0,0)(1,2,3)_6}$ | **9.82435** | **8.866489** | **23.8799** |
| 4 | $(0,0,0)(2,2,1)_6$ | 10.31223 | 9.343837 | 24.8451 |
| 5 | $(0,0,0)(2,2,2)_6$ | 10.06034 | 9.102483 | 22.3298 |
| 6 | $(0,0,0)(2,2,3)_6$ | 10.41788 | 9.238844 | 25.8838 |
| 7 | $(0,0,0)(3,2,1)_6$ | 10.11734 | 9.159476 | 25.2785 |
| 8 | $(0,0,0)(3,2,2)_6$ | 10.04366 | 9.096329 | 24.1006 |
| 9 | $(0,0,0)(3,2,3)_6$ | 10.21244 | 9.189762 | 25.1457 |
| 1 | $(0,0,0)(1,2,1)_{12}$ | 10.16622 | 9.187286 | 23.8995 |
| 2 | $(0,0,0)(1,2,2)_{12}$ | 9.778136 | 8.80974 | 23.0021 |
| 3 | $\mathbf{(0,0,0)(1,2,3)_{12}}$ | **9.732754** | **8.774892** | **23.4271** |
| 4 | $(0,0,0)(2,2,1)_{12}$ | 10.0044 | 9.036002 | 23.9845 |
| 5 | $(0,0,0)(2,2,2)_{12}$ | 9.795466 | 8.837605 | 22.2783 |

| | | | | |
|---|---|---|---|---|
| 6 | $(0,0,0)(2,2,3)_{12}$ | 10.745778 | 9.798451 | 25.2091 |
| 7 | $(0,0,0)(3,2,1)_{12}$ | 10.00265 | 9.044785 | 25.1066 |
| 8 | $(0,0,0)(3,2,2)_{12}$ | 10.01142 | 9.064091 | 24.1001 |
| 9 | $(0,0,0)(3,2,3)_{12}$ | 11.817427 | 10.880635 | 25.4252 |

The estimates of the parameters of the optimum models are presented as follows:

Fitted pure Seasonal autoregressive Integrated Moving Average model at s=1:

$$\hat{X}_t = \Psi_1 X_{t-1} + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2}$$

$\Psi_1 = 0.3374$, $\Theta_1 = -0.9980$, $\Theta_2 = 0.9987$

AIC=10.35153,

BIC=9.383136

Residual Variance = 23.0089

Fitted pure Seasonal autoregressive Integrated Moving Average model at s=2:

$$\hat{X}_t = \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Psi_3 X_{t-6} + \Theta_1 \varepsilon_{t-2} + \Theta_2 \varepsilon_{t-4}$$

$\Psi_1 = 0.0103$, $\Psi_2 = -0.1498$, $\Psi_3 = -0.3383$, $\Theta_1 = -0.9679$, $\Theta_2 = 0.9681$

AIC=10.33298,

BIC=9.385657

Residual Variance = 24.5390

Fitted pure Seasonal autoregressive Integrated Moving Average model at s=3:

$$\hat{X}_t = \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Psi_3 X_{t-9} + \Theta_1 \varepsilon_{t-3} + \Theta_2 \varepsilon_{t-6} + \Theta_3 \varepsilon_{t-9}$$

$\Psi_1 = -0.9273$, $\Psi_2 = -0.5955$, $\Psi_3 = -0.5746$, $\Theta_1 = -0.0352$,

$\Theta_2 = -0.5691$,

$\Theta_3 = 0.6043$

AIC=10.19405,

BIC=9.257262

Residual Variance = 24.9835

Fitted pure Seasonal autoregressive Integrated Moving Average model at s=4:

$$\hat{X}_t = \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Psi_3 X_{t-12} + \Theta_1 \varepsilon_{t-4} + \Theta_2 \varepsilon_{t-8} + \Theta_3 \varepsilon_{t-12}$$

$\Psi_1 = -0.2781$, $\Psi_2 = -0.2108$, $\Psi_3 = 0.3931$, $\Theta_1 = -0.8450$,

$\Theta_2 = 0.6942$, $\Theta_3 = 0.1522$

AIC= 10.07154,

BIC= 9.134746

Residual Variance = 23.3284

Fitted pure Seasonal autoregressive integrated Moving Average model at s=6:

$$\hat{X}_t = \Psi_1 X_{t-6} + \Theta_1 \varepsilon_{t-6} + \Theta_2 \varepsilon_{t-12} + \Theta_3 \varepsilon_{t-18}$$

$\Psi_1 = -0.9998$, $\Theta_1 = -0.9545$, $\Theta_2 = 0.9604$, $\Theta_3 = 0.9165$

AIC= 9.82435,

BIC= 8.866489

Residual Variance = 23.8799

Fitted pure Seasonal autoregressive integrated Moving Average model at s=12:

$$\hat{X}_t = \Psi_1 X_{t-12} + \Theta_1 \varepsilon_{t-12} + \Theta_2 \varepsilon_{t-24} + \Theta_3 \varepsilon_{t-36}$$

$\Psi_1 = -0.9215$, $\Theta_1 = -0.9919$, $\Theta_2 = -0.9978$, $\Theta_3 = 0.9937$

AIC= 9.732754,

BIC= 8.774892

Residual Variance = 23.4271

Table 4.9: Fitted optimum pure SARIMA models

| s/n | MODEL | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(0,0,0)(1,2,2)_1$ | 10.35153 | 9.383136 | 23.0089 |
| 2 | $(0,0,0)(3,2,2)_2$ | 10.33298 | 9.385657 | 24.5390 |
| 3 | $(0,0,0)(3,2,3)_3$ | 10.19405 | 9.257262 | 24.9835 |
| 4 | $(0,0,0)(3,2,3)_4$ | 10.07154 | 9.134746 | 23.3284 |
| 5 | $(0,0,0)(1,2,3)_6$ | 9.82435 | 8.866489 | 23.8799 |
| 6 | $(0,0,0)(1,2,3)_{12}$ | 9.732754 | 8.774892 | 23.4271 |

## 4.9 Fitted Pure Seasonal Autoregressive Integrated Moving Average One-Dimensional Bilinear (PSARIMAODBL) models (Order one).

Using the obtained parameters in the previous section as initial parameters for the nonlinear Mixed SARIMA one-dimensional bilinear part, we have the following results presented in table 4.10.

Fitted Pure Seasonal autoregressive integrated moving average one-dimensional bilinear model at s=1:

$$\hat{X}_t = \Psi_1 X_{t-1} + \Theta_1 \varepsilon_{t-1} + \Theta_2 \varepsilon_{t-2} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\Psi_1 = 0.3374$, $\Theta_1 = -0.9980$, $\Theta_2 = 0.9987$, $b_{11} = -0.6029$

AIC= 8.1082

BIC=7.8602

Residual Variance = 0.000893

Fitted Pure Seasonal autoregressive integrated moving average one-dimensional bilinear model at s=2:

$$\hat{X}_t = \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Psi_3 X_{t-6} + \Theta_1 \varepsilon_{t-2} + \Theta_2 \varepsilon_{t-4} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\Psi_1 = 0.0103$, $\Psi_2 = -0.1498$, $\Psi_3 = -0.3383$, $\Theta_1 = -0.9679$, $\Theta_2 = 0.9681$, $b_{11} = 0.5046$

AIC= 8.2908

BIC=7.5690

Residual Variance = 0.0061982

Fitted Pure Seasonal autoregressive integrated moving average one-dimensional bilinear model at s=3:

$$\hat{X}_t = \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Psi_3 X_{t-9} + \Theta_1 \varepsilon_{t-3} + \Theta_2 \varepsilon_{t-6} + \Theta_3 \varepsilon_{t-9} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\Psi_1 = -0.9273$, $\Psi_2 = -0.5955$, $\Psi_3 = -0.5746$, $\Theta_1 = -0.0352$, $\Theta_2 = -0.5691$,

$\Theta_3 = 0.6043$, $b_{11} = -0.3905$

AIC= 8.6448

BIC=7.8743

Residual Variance = 0.0063002

Fitted Pure Seasonal autoregressive integrated moving average one-dimensional bilinear model at s=4.

$$\hat{X}_t = \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Psi_3 X_{t-12} + \Theta_1 \varepsilon_{t-4} + \Theta_2 \varepsilon_{t-8} + \Theta_3 \varepsilon_{t-12} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\Psi_1 = -0.2781$, $\Psi_2 = -0.2108$, $\Psi_3 = 0.3931$, $\Theta_1 = -0.8450$, $\Theta_2 = 0.6942$, $\Theta_3 = 0.1522$, $b_{11} = 0.7171$

AIC= 8.1645

BIC=7.5409

Residual Variance = 0.0057496

Fitted Pure Seasonal autoregressive integrated bilinear model of order one at s=6:

$$\hat{X}_t = \Psi_1 X_{t-6} + \Theta_1 \varepsilon_{t-6} + \Theta_2 \varepsilon_{t-12} + \Theta_3 \varepsilon_{t-18} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\Psi_1 = -0.9998$, $\Theta_1 = -0.9545$, $\Theta_2 = 0.9604$, $\Theta_3 = 0.9165$, $b_{11} = 0.9741$

AIC= 7.7450

BIC=7.2803

Residual Variance = 0.0053793

Fitted Pure Seasonal autoregressive integrated bilinear model of order one at s=12:

$$\hat{X}_t = \Psi_1 X_{t-12} + \Theta_1 \varepsilon_{t-12} + \Theta_2 \varepsilon_{t-24} + \Theta_3 \varepsilon_{t-36} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\Psi_1 = -0.9215$, $\Theta_1 = -0.9919$, $\Theta_2 = -0.9978$, $\Theta_3 = 0.9937$,

$b_{11} = 0.5765$

AIC= 7.5026

BIC=6.7859

Residual Variance = 0.0053004

From table 4.10, the AIC values of the nonlinear pure SARIMA one-dimensional models vary between 7.60 and 8.70, the BIC values vary between 6.60 and 7.90 while the residual variances vary between 0.000893 and 0.0063. These show significant smaller differences from those obtained for the linear models in table 4.9.

Table 4.10: Fitted pure SARIMAODBL models (Order one)

| s/n | Model | AIC | BIC | Residual Variance |
|---|---|---|---|---|
| 1 | $(0,0,0)(1,2,2,1,1)_{1}$ | 8.1082 | 7.8602 | 0.000893 |
| 2 | $(0,0,0)(3,2,2,1,1)_{2}$ | 8.2908 | 7.5690 | 0.0061982 |
| 3 | $(0,0,0)(3,2,3,1,1)_{3}$ | 8.6448 | 7.8743 | 0.0063002 |
| 4 | $(0,0,0)(3,2,3,1,1)_{4}$ | 8.1645 | 7.5409 | 0.0057496 |
| 5 | $(0,0,0)(1,2,3,1,1)_{6}$ | 7.7450 | 7.2803 | 0.0053793 |
| 6 | $(0,0,0)(1,2,3,1,1)_{12}$ | 7.5026 | 6.7859 | 0.0053004 |

## 4.10 Fitted Mixed Seasonal Autoregressive Integrated Moving Average models

Here, we considered the Mixed Seasonal Autoregressive Integrated Moving Average Models. The estimates of the parameters of the linear Mixed Seasonal Autoregressive integrated Moving Average time series models of different orders of the nonseasonal Autoregressive parameters $p$, seasonal Autoregressive parameter $P$ , nonseasonal Moving average parameters, $q$ and the seasonal Moving Average parameters, $Q$ at the different length of seasons (s=1,2,3,4,6 and 12) are obtained. Optimum models are selected considering the Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), as well as the residual variances. Their estimates are then used as initial parameters for the nonlinear, Mixed Seasonal Autoregressive integrated Moving Average One-Dimensional Bilinear time series models. The various linear mixed SARIMA models and the optimum models are presented in tables 4.11 and 4.12

Table 4.11: Fitted mixed SARIMA models

| s/n | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(1,0,1)(1,2,0)_1$ | 10.64689 | 9.678491 | 22.781 |
| 2 | $(0,0,1)(1,2,0)_1$ | 10.65373 | 9.674799 | 22.983 |
| 3 | $(1,0,0)(0,2,1)_1$ | 10.65373 | 9.674799 | 23.007 |
| 4 | $(1,0,1)(0,2,1)_1$ | 10.34685 | 9.378455 | 22.010 |
| 5 | $(0,0,1)(0,2,1)_1$ | 10.34685 | 9.378455 | 22.014 |
| 6 | $(1,0,0)(1,2,1)_1$ | 10.64689 | 9.678491 | 22.089 |
| **7** | $\mathbf{(1,0,1)(1,2,1)_1}$ | **9.35039** | **8.39253** | **21.873** |
| 8 | $(0,0,1)(1,2,1)_1$ | 10.34685 | 9.378455 | 22.014 |
| 9 | $(1,0,0)(2,2,1)_1$ | 10.59501 | 9.63715 | 22.574 |
| 10 | $(1,0,1)(2,2,1)_1$ | 10.35243 | 9.405105 | 22.010 |
| 11 | $(0,0,1)(2,2,1)_1$ | 10.35005 | 9.392184 | 22.000 |
| 12 | $(1,0,0)(2,2,2)_1$ | 10.35824 | 9.410909 | 23.067 |
| 13 | $(1,0,1)(2,2,2)_1$ | 12.0891 | 11.329212 | 24.8971 |
| 14 | $(0,0,1)(2,2,2)_1$ | 10.35648 | 9.409152 | 23.0921 |
| 15 | $(1,0,2)(2,2,2)_1$ | 10.36123 | 9.43497 | 23.8897 |
| 16 | $(2,0,1)(2,2,2)_1$ | 10.29732 | 9.37106 | 25.6802 |
| 17 | $(2,0,2)(2,2,2)_1$ | 11.30613 | 10.390404 | 26.7026 |
| 1 | $(1,0,1)(1,2,0)_2$ | 11.1252 | 10.15681 | 26.2967 |
| 2 | $(0,0,1)(1,2,0)_2$ | 11.18342 | 10.20449 | 26.5605 |
| 3 | $(1,0,0)(0,2,1)_2$ | 11.04304 | 10.06411 | 26.0088 |
| 4 | $(1,0,1)(0,2,1)_2$ | 10.66491 | 9.696517 | 25.4509 |
| 5 | $(0,0,1)(0,2,1)_2$ | 10.73338 | 9.764982 | 25.6121 |
| 6 | $(1,0,0)(1,2,1)_2$ | 10.80936 | 9.840961 | 25.6595 |
| **7** | $(1,0,1)(1,2,1)_2$ | 10.60457 | 9.646711 | 25.5006 |
| 8 | $(0,0,1)(1,2,1)_2$ | 10.73338 | 9.764982 | 24.6236 |
| 9 | $(1,0,0)(2,2,1)_2$ | 10.7506 | 9.792738 | 25.6700 |
| 10 | $(1,0,1)(2,2,1)_2$ | 10.60103 | 9.653704 | 25.5002 |

| | | | | |
|---|---|---|---|---|
| 11 | $(0,0,1)(2,2,1)_2$ | 10.71173 | 9.753871 | 25.6010 |
| 12 | $(1,0,0)(2,2,2)_2$ | 10.3681 | 9.420772 | 25.0011 |
| 13 | $(1,0,1)(2,2,2)_2$ | 10.55414 | 9.617348 | 25.3496 |
| 14 | $(0,0,1)(2,2,2)_2$ | 10.36412 | 9.416794 | 25.3002 |
| 15 | **$(1,0,2)(2,2,2)_2$** | **10.31248** | **9.386221** | **23.8001** |
| 16 | $(2,0,1)(2,2,2)_2$ | 10.63753 | 9.711268 | 26.6721 |
| 17 | $(2,0,2)(2,2,2)_2$ | 11.35065 | 10.434922 | 26.7026 |
| 1 | $(1,0,1)(1,2,0)_3$ | 10.64689 | 9.678491 | 25.3891 |
| 2 | $(0,0,1)(1,2,0)_3$ | 11.70064 | 10.72171 | 26.6782 |
| 3 | $(1,0,0)(0,2,1)_3$ | 11.05494 | 10.07601 | 26.0171 |
| 4 | $(1,0,1)(0,2,1)_3$ | 11.05805 | 10.08966 | 26.0440 |
| 5 | $(0,0,1)(0,2,1)_3$ | 10.90368 | 9.935285 | 25.2410 |
| 6 | $(1,0,0)(1,2,1)_3$ | 10.87284 | 9.904448 | 25.2203 |
| **7** | $(1,0,1)(1,2,1)_3$ | 10.87633 | 9.918464 | 25.2410 |
| 8 | $(0,0,1)(1,2,1)_3$ | 10.90368 | 9.935285 | 24.8745 |
| 9 | $(1,0,0)(2,2,1)_3$ | 10.69443 | 9.736569 | 25.4623 |
| 10 | $(1,0,1)(2,2,1)_3$ | 10.6968 | 9.749472 | 25.0009 |
| 11 | $(0,0,1)(2,2,1)_3$ | 10.70592 | 9.345226 | 25.0021 |
| 12 | $(1,0,0)(2,2,2)_3$ | 10.28252 | 9.335192 | 24.3431 |
| 13 | $(1,0,1)(2,2,2)_3$ | 10.28686 | 9.35007 | 24.3896 |
| **14** | **$(0,0,1)(2,2,2)_3$** | **10.27991** | **9.332585** | **24.3001** |
| 15 | $(1,0,2)(2,2,2)_3$ | 10.79085 | 9.764596 | 25.6301 |
| 16 | $(2,0,1)(2,2,2)_3$ | 10.6873 | 9.91077 | 25.8945 |
| 17 | $(2,0,2)(2,2,2)_3$ | 11.27194 | 10 .356214 | 26.8934 |
| 1 | $(1,0,1)(1,2,0)_4$ | 11.77489 | 10.8065 | 26.6456 |
| 2 | $(0,0,1)(1,2,0)_4$ | 11.79615 | 10.81722 | 26.5682 |
| 3 | $(1,0,0)(0,2,1)_4$ | 11.10639 | 10.12746 | 25.7802 |
| 4 | $(1,0,1)(0,2,1)_4$ | 11.11177 | 10.14337 | 25.7854 |
| 5 | $(0,0,1)(0,2,1)_4$ | 10.86325 | 9.894854 | 24.6492 |

| | | | | |
|---|---|---|---|---|
| 6 | $(1,0,0)(1,2,1)_4$ | 10.84038 | 9.871989 | 24.4814 |
| **7** | $(1,0,1)(1,2,1)_4$ | 10.84559 | 9.887726 | 24.4879 |
| 8 | $(0,0,1)(1,2,1)_4$ | 10.86325 | 9.894854 | 24.7129 |
| 9 | $(1,0,0)(2,2,1)_4$ | 10.25934 | 9.301478 | 23.9021 |
| 10 | $(1,0,1)(2,2,1)_4$ | 10.26262 | 9.315295 | 23.9785 |
| 11 | $(0,0,1)(2,2,1)_4$ | 10.25859 | 9.300724 | 23.7022 |
| 12 | $(1,0,0)(2,2,2)_4$ | 10.24646 | 9.299138 | 23.7001 |
| **13** | $\mathbf{(1,0,1)(2,2,2)_4}$ | **10.23915** | **9.302358** | **23.4056** |
| 14 | $(0,0,1)(2,2,2)_4$ | 10.24516 | 9.297831 | 23.7000 |
| 15 | $(1,0,2)(2.2.2)_4$ | 10.65603 | 9.52977 | 25.2018 |
| 16 | $(2,0,1)(2,2,2)_4$ | 10.8895 | 9.7452 | 26.8945 |
| 17 | $(2,0,2)(2,2,2)_4$ | 11.25281 | 10.337086 | 26.6073 |
| 1 | $(1,0,1)(1,2,0)_6$ | 10.95002 | 9.981624 | 24.8562 |
| 2 | $(0,0,1)(1,2,0)_6$ | 11.0481 | 10.06917 | 25.0783 |
| 3 | $(1,0,0)(0,2,1)_6$ | 11.21908 | 10.24015 | 25.2891 |
| 4 | $(1,0,1)(0,2,1)_6$ | 11.22244 | 10.25405 | 25.29751 |
| 5 | $(0,0,1)(0,2,1)_6$ | 10.31712 | 9.348723 | 23.7132 |
| 6 | $(1,0,0)(1,2,1)_6$ | 10.3178 | 9.349402 | 23.7157 |
| **7** | $(1,0,1)(1,2,1)_6$ | 10.32036 | 9.362498 | 23.7159 |
| 8 | $(0,0,1)(1,2,1)_6$ | 10.31712 | 9.348723 | 23.8745 |
| 9 | $(1,0,0)(2,2,1)_6$ | 10.31016 | 9.352302 | 23.7102 |
| 10 | $(1,0,1)(2,2,1)_6$ | 10.31306 | 9.365736 | 23.7115 |
| 11 | $(0,0,1)(2,2,1)_6$ | 10.30936 | 9.351504 | 23.7100 |
| 12 | $(1,0,0)(2,2,2)_6$ | 10.06382 | 9.11649 | 23.2710 |
| 13 | $(1,0,1)(2,2,2)_6$ | 10.23915 | 9.302358 | 23.5343 |
| **14** | $\mathbf{(0,0,1)(2,2,2)_6}$ | **10.05624** | **9.108912** | **23.2303** |
| 15 | $(1,0,2)(2,2,2)_6$ | 10.8685 | 9.342329 | 25.1903 |
| 16 | $(2,0,1)(2,2,2)_6$ | 11.8936 | 10.1455 | 26.9834 |
| 17 | $(2,0,2)(2,2,2)_6$ | 11.02962 | 10.113897 | 26.0045 |
| 1 | $(1,0,1)(1,2,0)_{12}$ | 10.91086 | 9.942465 | 25.7943 |
| 2 | $(0,0,1)(1,2,0)_{12}$ | 10.95855 | 9.979618 | 25.8122 |

| | | | | |
|---|---|---|---|---|
| 3 | $(1,0,0)(0,2,1)_{12}$ | 10.42466 | 9.445733 | 23.7154 |
| 4 | $(1,0,1)(0,2,1)_{12}$ | 10.38919 | 9.42079 | 23.7178 |
| 5 | $(0,0,1)(0,2,1)_{12}$ | 10.17115 | 9.202758 | 23.3451 |
| 6 | $(1,0,0)(1,2,1)_{12}$ | 10.17122 | 9.202826 | 23.3432 |
| **7** | $(1,0,1)(1,2,1)_{12}$ | 10.32036 | 9.362498 | 23.7021 |
| 8 | $(0,0,1)(1,2,1)_{12}$ | 10.17115 | 9.202758 | 23.3067 |
| 9 | $(1,0,0)(2,2,1)_{12}$ | 10.31016 | 9.352302 | 23.7002 |
| 10 | $(1,0,1)(2,2,1)_{12}$ | 9.991618 | 9.044292 | 23.0043 |
| 11 | $(0,0,1)(2,2,1)_{12}$ | 10.00973 | 9.051868 | 23.2891 |
| 12 | $(1,0,0)(2,2,2)_{12.}$ | 10.06382 | 9.11649 | 23.4814 |
| 13 | $(1,0,1)(2,2,2)_{12}$ | 10.23915 | 9.302358 | 23.4569 |
| **14** | $\mathbf{(0,0,1)(2,2,2)_{12}}$ | **9.801701** | **8.854374** | **22.7016** |
| 15 | $(1,0,2)(2,2,2)_{12}$ | 10.820449 | 9.894191 | 25.2792 |
| 16 | $(2,0,1)(2,2,2)_{12}$ | 11.8107 | 10.884507 | 26.8891 |
| 17 | $(2,0,2)(2,2,2)_{12}$ | 11.766747 | 10.851024 | 26.7899 |

The estimates of the parameters of these optimum models are as follows:

Fitted Mixed Seasonal autoregressive Integrated Moving Average model at s=1:

$$\hat{X}_t = \psi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-1} + \Theta_1 \varepsilon_{t-1}$$

$\psi_1 = 0.1692$, $\theta_1 = -0.0467$, $\Psi_1 = 0.1692$, $\Theta_1 = 0.0984$

AIC= 9.35039,

BIC= 8.39253

Residual Variance = 21.873

Fitted Mixed Seasonal autoregressive Integrated Moving Average model at s=2:

$\hat{X}_t = \psi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Theta_1 \varepsilon_{t-2} + \Theta_2 \varepsilon_{t-4}$ $\psi_1 = 0.3453$, $\theta_1 = -0.019$, $\theta_2 = -0.981$ $\Psi_1 = -0.9636$, $\Psi_2 = 0.036$,

$\Theta_1 = -0.0075$, $\Theta_2 = -0.9924$.

AIC= 10.31248,

BIC= 9.386221

Residual Variance = 23.8001

Fitted Mixed Seasonal autoregressive Integrated Moving Average model at s=3:

$$\hat{X}_t = \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Theta_1 \varepsilon_{t-3} + \Theta_2 \varepsilon_{t-6}$$

$\theta_1 = 0.2317$, $\Psi_1 = -0.0819$, $\Psi_2 = -0.3042$, $\Theta_1 = -0.9110$,

$\Theta_2 = 0.9512$

AIC= 10.27991,

BIC= 9.332585

Residual Variance = 24.3001

Fitted Mixed Seasonal autoregressive Integrated Moving Average model at s=4:

$$\hat{X}_t = \psi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Theta_1 \varepsilon_{t-4} + \Theta_2 \varepsilon_{t-8}$$

$\psi_1 = 0.0835$, $\theta_1 = 0.1062$, $\Psi_1 = -0.4812$, $\Psi_2 = -0.4703$,

$\Theta_1 = -0.6159$,

$\Theta_2 = 0.6159$

AIC= 10.23915,

BIC= 9.302358

Residual Variance = 23.4056

Fitted Mixed Seasonal autoregressive Integrated Moving Average model at s=6:

$$\hat{X}_t = \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-6} + \Psi_2 X_{t-12} + \Theta_1 \varepsilon_{t-6} + \Theta_2 \varepsilon_{t-12}$$

$\theta_1 = 0.0914$, $\Psi_1 = -0.2297$, $\Psi_2 = 0.3893$, $\Theta_1 = -0.9440$, $\Theta_2 = 0.9988$

AIC= 10.05624,

BIC= 9.108912

Residual Variance = 23.2303

Fitted Mixed Seasonal autoregressive Integrated Moving Average model at s=12:

$$\hat{X}_t = \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-12} + \Psi_2 X_{t-24} + \Theta_1 \varepsilon_{t-12} + \Theta_2 \varepsilon_{t-24}$$

$\theta_1 = 0.0142$, $\Psi_1 = 0.0851$, $\Psi_2 = 0.0281$, $\Theta_1 = -0.9191$, $\Theta_2 = 0.9991$

AIC= 9.801701,

BIC= 8.854374

Residual Variance = 22.7016

From the summary presented in table 4.12, it is observed that the AIC values vary between 9.30 and 10.40, the BIC values vary between 8.30 and 9.40 while the residual variances vary between 21.5 and 24.4. These we shall compare with the values of their nonlinear counterparts in the next section.

Table 4.12: Fitted optimum Mixed SARIMA models

| s/n | Model | AIC | BIC | Residual Variance |
|---|---|---|---|---|
| 1 | $(1,0,1)(1,2,1)_1$ | 9.35039 | 8.39253 | 21.873 |
| 2 | $(1,0,2)(2,2,2)_2$ | 10.31248 | 9.386221 | 23.8001 |
| 3 | $(0,0,1)(2,2,2)_3$ | 10.27991 | 9.332585 | 24.3001 |
| 4 | $(1,0,1)(2,2,2)_4$ | 10.23915 | 9.302358 | 23.4056 |
| 5 | $(0,0,1)(2,2,2)_6$ | 10.05624 | 9.108912 | 23.2303 |
| 6 | $(0,0,1)(2,2,2)_{12}$ | 9.801701 | 8.854374 | 22.7016 |

## 4.11 Fitted Mixed Seasonal Autoregressive Integrated Moving Average One-Dimensional Bilinear (MSARIMAODBL) models (Order one).

The parameters of the optimum linear models obtained in the previous section are used as initial parameters to fit the nonlinear mixed SARIMAODBL models and presented in table 4.13 with their corresponding AIC, BIC as well as residual variances. Then we obtained the nonlinear parameters as follows.

Fitted Mixed Seasonal Autoregressive Integrated Moving Average Bilinear model of order one at s=1:

$$\hat{X}_t = \psi_1 X_{t-1} + \theta_1\varepsilon_{t-1} + \Psi_1 X_{t-1} + \Theta_1\varepsilon_{t-1} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\psi_1 = 0.1692$, $\theta_1 = -0.0467$, $\Psi_1 = 0.1692$, $\Theta_1 = 0.0984$,

$b_{11} = 0.0562$

AIC= 7.3076

BIC = 6.7629

Residual Variance = 0.0000858

Fitted Mixed Seasonal Autoregressive Integrated Moving Average Bilinear model of order one at s=2:

$\hat{X}_t = \psi_1 X_{t-1} + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \Psi_1 X_{t-2} + \Psi_2 X_{t-4} + \Theta_1\varepsilon_{t-2} + \Theta_2\varepsilon_{t-4} + b_{11} X_{t-1}\varepsilon_{t-1}$ where:

$\psi_1 = 0.3453$, $\theta_1 = -0.019$, $\theta_2 = -0.981$ $\Psi_1 = -0.9636$, $\Psi_2 = 0.036$,

$\Theta_1 = -0.0075$,

$\Theta_2 = -0.9924$, $b_{11} = 0.8023$

AIC= 8.0985

BIC=7.6786

Residual Variance = 0.005967

Fitted Mixed Seasonal Autoregressive Integrated Moving Average Bilinear model of order one at s=3:

$$\hat{X}_t = \theta_1\varepsilon_{t-1} + \Psi_1 X_{t-3} + \Psi_2 X_{t-6} + \Theta_1\varepsilon_{t-3} + \Theta_2\varepsilon_{t-6} + b_{11} X_{t-1}\varepsilon_{t-1}$$

$\theta_1 = 0.2317, \quad \Psi_1 = -0.0819, \Psi_2 = -0.3042, \quad \Theta_1 = -0.9110, \quad \Theta_2 = 0.9512,$

$b_{11} = 0.0089$

AIC= 9.5623

BIC= 8.7009

Residual Variance = 0.00936

Fitted Mixed Seasonal Autoregressive Integrated Moving Average Bilinear model of order one at s=4:

$$\hat{X}_t = \psi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Theta_1 \varepsilon_{t-4} + \Theta_2 \varepsilon_{t-8} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\psi_1 = 0.0835, \quad \theta_1 = 0.1062, \quad \Psi_1 = -0.4812, \Psi_2 = -0.4703,$

$\Theta_1 = -0.6159, \quad \Theta_2 = 0.6159, \ b_{11} = 0.6813$

AIC= 6.5623

BIC=5.7009

Residual Variance = 0.000000936

Fitted Mixed Seasonal Autoregressive Integrated Moving Average Bilinear model of order one at s=6:

$$\hat{X}_t = \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-6} + \Psi_2 X_{t-12} + \Theta_1 \varepsilon_{t-6} + \Theta_2 \varepsilon_{t-12} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\theta_1 = 0.0914, \quad \Psi_1 = -0.2297, \Psi_2 = 0.3893, \quad \Theta_1 = -0.9440, \quad \Theta_2 = 0.9988,$

$b_{11} = 0.8065$

AIC= 9.4002

BIC= 8.0843

Residual Variance = 0.04892

Fitted Mixed Seasonal Autoregressive Integrated Moving Average Bilinear model of order one at s=12:

$$\hat{X}_t = \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-12} + \Psi_2 X_{t-24} + \Theta_1 \varepsilon_{t-12} + \Theta_2 \varepsilon_{t-24} + b_{11} X_{t-1} \varepsilon_{t-1}$$

$\theta_1 = 0.0142, \quad \Psi_1 = 0.0851, \Psi_2 = 0.0281, \quad \Theta_1 = -0.9191,$

$\Theta_2 = 0.9991$, $b_{11} = 0.6539$

AIC= 8.5892

BIC= 7.3021

Residual Variance = 0.056290

Table 4.13: Fitted mixed SARIMAODBL models (Order One)

| s/n | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(1,0,1)(1,2,1,1,1)_1$ | 7.3076 | 6.7629 | 0.0000858 |
| 2 | $(1,0,2)(2,2,2,1,1)_2$ | 8.0985 | 7.6786 | 0.005967 |
| 3 | $(0,0,1)(2,2,2,1,1)_3$ | 9.5623 | 8.7009 | 0.00936 |
| **4** | $\mathbf{(1,0,1)(2,2,2,1,1)_4}$ | **6.5623** | **5.7009** | **0.000000936** |
| 5 | $(0,0,1)(2,2,2,1,1)_6$ | 9.4002 | 8.0843 | 0.04892 |
| 6 | $(0,0,1)(2,2,2,1,1)_{12}$ | 8.5892 | 7.3021 | 0.056290 |

## 4.12 Fitted optimum one-dimensional bilinear models of order $m$

So far, we have considered and obtained the best model for PARIODBL, MARIODBL, PARIMAODBL and MARIMAODBL at order one when $s = 1, 2, 3, 4, 6$ and $12$ as shown in table 4.14. There is the need to increase this order which will further reveal if there exists a more robust model using the models in table 4.14.

Table 4.14: Optimal one-dimensional bilinear models

| S/N | Model | AIC | BIC | Residual Variance |
|---|---|---|---|---|
| 1 | PSARIODBL $(0,0,0)(3,2,0,1,1)_1$ | 9.7823 | 8.5624 | 0.00947 |
| 2 | MSARIODBL $(2,0,0)(3,2,0,1,1)_1$ | 7.0064 | 6.6015 | 0.000893 |
| 3 | PSARIMAODBL $(0,0,0)(1,2,2,1,1)_1$ | 8.1082 | 7.8602 | 0.000893 |
| 4 | MSARIMAODBL $(1,0,1)(2,2,2,1,1)_4$ | 6.5623 | 5.7009 | 0.000000936 |

### 4.12.1 Pure Seasonal Autoregressive Integrated one-dimensional Bilinear models of order $m$.

For the Pure Seasonal Autoregressive Integrated Bilinear models shown in table 4.15, it is discovered as shown in table 4.14, that the AIC and the BIC increase as the parameter $m$ increases from 1 to 2. However, there is a decrease as the parameter increases to 3, after which it increases again. Hence, the maximum order of $m$ that can be fitted is $m = 4$ and the optimum model is given by $(0,0,0)(3,2,0,3,1)_{1}$.

Hence the fitted model $(0,0,0)(3,2,0,3,1)_{1}$ is:

$\hat{X}_t = \Psi_1 X_{t-1} + \Psi_2 X_{t-2} + \Psi_2 X_{t-3} + b_{11} X_{t-1}\varepsilon_{t-1} + b_{21} X_{t-1}\varepsilon_{t-1} + b_{31} X_{t-1}\varepsilon_{t-1}$ where:

$\Psi_1 = -0.9973$, $\Psi_2 = -0.7530$, $\Psi_3 = -0.3588$, $b_{11} = 0.4638$,

$b_{21} = 0.6903$, $b_{31} = 0.4027$

Table 4.15: Pure SARIODBL models of order $m$

| S/N | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(0,0,0)(3,2,0,1,1)_1$ | 9.7823 | 8.5624 | 0.00947 |
| 2 | $(0,0,0)(3,2,0,2,1)_1$ | 11.3892 | 10.5903 | 0.0678 |
| **3** | $\mathbf{(0,0,0)(3,2,0,3,1)_1}$ | **8.4672** | **7.9091** | **0.000043920** |
| 4 | $(0,0,0)(3,2,0,4,1)_1$ | 11.7865 | 11.1079 | 0.096934 |

### 4.12.2 Mixed Seasonal Autoregressive Integrated one-dimensional Bilinear (MSARIODBL) models of order $m$

Similarly, for the Mixed Seasonal Autoregressive Integrated one-dimensional Bilinear models, it is also discovered as shown in table 4.16, that the AIC and the BIC increase as the parameter $m$ increases from 1 to 2. However, there is a decrease as the parameters increase to 3, after which it increases again. Hence, the maximum order of $m$ that can be fitted is $m = 4$ and the optimum model is given by $(2,0,0)(3,2,0,3,1)_1$. Hence, the fitted model $(2,0,0)(3,2,0,3,1)_1$ is:

$$\hat{X}_t = \psi_1 X_{t-1} + \psi_2 X_{t-2} + \Psi_1 X_{t-1} + \Psi_2 X_{t-2} + \Psi_3 X_{t-3} + b_{11} X_{t-1}\varepsilon_{t-1} + b_{21} X_{t-1}\varepsilon_{t-1} + b_{31} X_{t-1}\varepsilon_{t-1}$$

where: $\psi_1 = -0.7581, \psi_2 = -0.4532, \Psi_1 = -0.3574, \Psi_2 = -0.2936, \Psi_3 = -0.3473, b_{11} = -0.9024, b_{21} = 0.3902, b_{31} = 0.5032$

Table 4.16: Fitted MSARIODBL models of order $m$

| S/N | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(2,0,0)(3,2,0,1,1)_1$ | 7.0064 | 6.6015 | 0.000893 |
| 2 | $(2,0,0)( 3,2,0,2,1)_1$ | 10.7082 | 9.4606 | 0.007390 |
| **3** | $\mathbf{(2,0,0)( 3,2,0,3,1)_1}$ | **6.8025** | **6.3803** | **0.00001936** |
| 4 | $(2,0,0)( 3,2,0,4,1)_1$ | 12.7984 | 11.9968 | 0.09894534 |

### 4.12.3 Pure Seasonal Autoregressive Integrated Moving Average one-dimensional Bilinear (PSARIMAODBL) models of order $m$

For the Pure Seasonal Autoregressive Integrated Moving Average one-dimensional Bilinear models, the results are as shown below, it is observed that the AIC and the BIC increase as the parameter $m$ increases from 1 to 3 (table 4.17). However, there is a decrease as the number of parameters increases to 4, after which it increases again. Hence, the maximum order of $m$ that can be fitted is $m = 5$ and the optimum model is given by $(0,0,0)(1,2,2,4,1)_{1.}$ Therefore the fitted optimum model is:

$$X_t = \Psi_1 X_{t-1} + \Theta_1 e_{t-1} + \Theta_2 e_{t-1} + b_{11} X_{t-1} e_{t-1} + b_{21} X_{t-2} e_{t-1} + b_{31} X_{t-3} e_{t-1} + b_{41} X_{t-4} e_{t-1} + e_t$$

where:

$\Psi_1 = 0.3374, \Theta_1 = -0.9980, \Theta_2 = 0.9987, b_{11} = 0.3904, b_{21} = -0.6920, b_{31} = 0.8974,$
$b_{41} = 0.6702$

Table 4.17: Fitted PSARIMAODBL models of order $m$.

| S/N | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(0,0,0)(1,2,2,1,1)_1$ | 8.1082 | 0.6054 | 0.000893 |
| 2 | $(0,0,0)(1,2,2,2,1)_1$ | 10.7922 | 9.9685 | 0.08329 |
| 3 | $(0,0,0)(1,2,2,3,1)_1$ | 11.2187 | 10.4045 | 0.04796 |
| **4** | $\mathbf{(0,0,0)(1,2,2,4,1)_1}$ | **7.5007** | **6.8945** | **0.000214** |
| 5 | $(0,0,0)(1,2,2,5,1)_1$ | 11.7891 | 10.3905 | 0.073028 |

### 4.12.4 Mixed Seasonal Autoregressive Integrated Moving Average one-dimensional Bilinear (MSARIMAODBL) models of order $m$.

The results for the Mixed Seasonal Autoregressive Integrated Moving Average Bilinear models of order $m$ is as shown in table 4.18. It is observed that the AIC and the BIC increase as the number of parameter $m$ increases from 1 to 3. However, there is a decrease as the parameter increases to 4 to 5, after which it increases again at $m = 6$. However, the minimum AIC and BIC still occur at $m$=1. Hence, the optimum model is given by $(1,0,1)(2,2,2,1,1)_{4.}$

Since the minimum is obtained at order 1 then, the fitted model is given by:

$$\hat{X}_t = \psi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Theta_1 \varepsilon_{t-4} + \Theta_2 \varepsilon_{t-8} + b_{11} X_{t-1} \varepsilon_{t-1}$$

where:

$\psi_1 = 0.0835$, $\theta_1 = 0.1062$, $\Psi_1 = -0.4812$, $\Psi_2 = -0.4703$,

$\Theta_1 = -0.6159$, $\Theta_2 = 0.6159$, $b_{11} = 0.6813$.

Table 4.18: Fitted MSARIMAODBL models of order $m$.

| S/N | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| **1** | **$(1,0,1)(1,2,2,1,1)_4$** | **6.5623** | **5.7009** | **0.0000000936** |
| 2 | $(1,0,1)(1,2,2,2,1)_4$ | 10.7193 | 9.6982 | 0.0456 |
| 3 | $(1,0,1)(1,2,2,3,1)_4$ | 10.45003 | 9.5891 | 0.044332 |
| 4 | $(1,0,1)(1,2,2,4,1)_4$ | 8.1901 | 6.5891 | 0.63016 |
| 5 | $(1,0,1)(1,2,2,5,1)_4$ | 6.9043 | 5.8963 | 0.000921 |
| 6 | $(1,0,1)(1,2,2,6,1)_4$ | 11.8790 | 11.8790 | 0.06740 |

From the previous section, the optimum one-dimensional models of order $m$ are as presented in table 4.19. This reveals the fact that the mixed seasonal autoregressive integrated moving average one-dimensional bilinear (MSARIMAODBL) model of order one $(1,0,1)(2,2,2,1,1)_4$ has the least AIC, BIC and residual variance values 6.5623, 5.7009 and 0.000000936 respectively, followed by the mixed seasonal autoregressive integrated one-dimensional bilinear (MSARIODBL) model $(2,0,0)(3,2,0,3,1)_1$, the pure seasonal autoregressive integrated one-dimensional bilinear (PSARIODBL) model

$(0,0,0)(3,2,0,3,1)_1$and the pure seasonal autoregressive integrated moving average one-dimensional bilinear (PSARIMAODBL) model $(0,0,0)(1,2,2,4,1)_1$ in that order.

Table 4.19: Fitted Optimum one-dimensional bilinear models of order $m$.

| s/n | Model | AIC | BIC | Residual variance |
|---|---|---|---|---|
| 1 | $(1,0,1)(1,2,2,1,1)_4$ | 6.5623 | 5.7009 | 0.000000936 |
| 2 | $(2,0,0)(3,2,0,3,1)_1$ | 6.8025 | 6.3803 | 0.00001936 |
| 3 | $(0,0,0)(3,2,0,3,1)_1$ | 8.4672 | 7.9091 | 0.000043920 |
| 4 | $(0,0,0)(1,2,2,4,1)_1$ | 7.5007 | 6.8945 | 0.000214 |

### 4.13 Simulation Study

In order to further establish the performance of our optimal model, simulated data of sample sizes n=250, 500 and 1000 are generated using the Monte Carlo technique. The results are as follows.

Given the fitted optimal mixed seasonal one-dimensional bilinear model:

$$\hat{X}_t = \psi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Theta_1 \varepsilon_{t-4} + \Theta_2 \varepsilon_{t-8} + b_{11} X_{t-1} \varepsilon_{t-1}$$

where:

$\psi_1 = 0.0835$, $\theta_1 = 0.1062$, $\Psi_1 = -0.4812$, $\Psi_2 = -0.4703$,

$\Theta_1 = -0.6159$, $\Theta_2 = 0.6159$, $b_{11} = 0.6813$.

**The procedure:**

A data set of sample size, n = 250 values are generated for $X_t$ and $\varepsilon_t$. The data set is subjected to the optimum model above and new set of parameters are estimated using the nonlinear least square method. The procedure is repeated 50 times (replicates) with different sets of $X_t$ and $\varepsilon_t$. Averages are then obtained for each estimate and compared with the test values (hypothesized values). The entire procedure is then repeated for sample sizes n=500 and n=1000. The results are as presented in tables 4.20, 4.21 and 4.22. From the tables, we discover that:

(i) The parameters of the simulated data compared favourably with the parameters of our mixed seasonal bilinear model.

(ii) As the sample size increases (from n=250 to n=1000), the model further shows better performance with non-significance in difference for all the parameters.

### 4.13.1 Simulation results when n=250 and r=50

Table 4.20: Results of simulated data when n=250, replication=50

| S/n | Parameter | Hypothesized value | Mean (simulated value) | Standard error | Prob. | Remark |
|---|---|---|---|---|---|---|
| 1 | $\psi_1$ | 0.0835 | 0.0823 | 0.6989 | 0.1820 | Not significant |
| 2 | $\theta_1$ | 0.1062 | 0.1065 | 0.3087 | 0.0942 | Not significant |
| 3 | $\Psi_1$ | −0.4812 | −0.4699 | 0.9310 | 0.4302 | Not Significant |
| 4 | $\Psi_2$ | −0.4703 | −0.4629 | 0.0568 | 0.0214 | Significant |
| 5 | $\Theta_1$ | −0.6159 | −0.6002 | 2.2013 | 0.0319 | Significant |
| 6 | $\Theta_2$ | 0.6159 | 0.5946 | 1.1090 | 0.0118 | Significant |
| 7 | $b_{11}$ | 0.6813 | 0.6824 | 0.4801 | 0.6090 | Not Significant |

### 4.13.2 Simulation results when n=500 and r=50

Table 4.21: Results of simulated data when n=500, replication=50

| S/n | Parameter | Hypothesized value | Mean (simulated value) | Standard error | Prob. | Remark |
|---|---|---|---|---|---|---|
| 1 | $\psi_1$ | 0.0835 | 0.08310 | 0.6456 | 0.3892 | Not significant |
| 2 | $\theta_1$ | 0.1062 | 0.1068 | 0.7045 | 0.3190 | Not Significant |
| 3 | $\Psi_1$ | −0.4812 | −0.4798 | 1.3018 | 0.0597 | Not significant |
| 4 | $\Psi_2$ | −0.4703 | −0.4686 | 0.0566 | 0.6508 | Not significant |
| 5 | $\Theta_1$ | −0.6159 | −0.6023 | 0.8766 | 0.5400 | Not significant |
| 6 | $\Theta_2$ | 0.6159 | 0.6009 | 0.6210 | 0.8334 | Not significant |
| 7 | $b_{11}$ | 0.6813 | 0.6798 | 0.0045 | 0.8234 | Not significant |

### 4.13.3 Simulation results when n=1000 and r=50

Table 4.22: Results of simulated data when n=1000, replication=50

| s/n | Parameter | Hypothesized value | Mean (simulated value) | Standard error | Prob. | Remark |
|---|---|---|---|---|---|---|
| 1 | $\psi_1$ | 0.0835 | 0.08341 | 0.0378 | 0.3920 | Not significant |
| 2 | $\theta_1$ | 0.1062 | 0.1063 | 0.0256 | 0.4902 | Not significant |
| 3 | $\Psi_1$ | -0.4812 | −0.4810 | 0.0741 | 0.2901 | Not significant |
| 4 | $\Psi_2$ | -0.4703 | −0.4695 | 0.0034 | 0.0892 | Not significant |
| 5 | $\Theta_1$ | -0.6159 | −0.6057 | 0.3802 | 0.9048 | Not significant |
| 6 | $\Theta_2$ | 0.6159 | 0.6084 | 0.7201 | 0.6930 | Not significant |
| 7 | $b_{11}$ | 0.6813 | 0.6808 | 0.0894 | 0.0928 | Not significant |

### 4.14 In-sample Forecast

The Ikeja rainfall data from the Nigerian Meterological Agency (NMA) used in obtaining the results above ranges from (1984-2004). Furthermore, rainfall data for the same location for the year 2015 and 2016 are obtained as shown in columns (i) and (iii) in the table 4.23. Then, our optimum model with the minimum residual variance given by $(1,0,1)(2,2,2,1,1)_4$ :

$$\hat{X}_t = \psi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Theta_1 \varepsilon_{t-4} + \Theta_2 \varepsilon_{t-8} + b_{11} X_{t-1} \varepsilon_{t-1}$$

where:

$\psi_1 = 0.0835$, $\theta_1 = 0.1062$, $\Psi_1 = -0.4812$, $\Psi_2 = -0.4703$, $\Theta_1 = -0.6159$, $\Theta_2 = 0.6159$, $b_{11} = 0.6813$.

is the used to obtain the forecast of the year 2015 and 2016. The results are as shown in columns (ii) and (iv) in table 4.23.

Critical look at this shows that the forecast values for the year 2015 are very close estimates of the actual values at every month of each of year. Similarly, forecast values for the year 2016 give very close estimates of the actual values at each corresponding month of each of year.

Table 4.23: In-sample Forecast of 2015 and 2016

| Month | (i)<br>2015 (Actual) | **(ii)**<br>**2015 (Forecast)** | (iii)<br>2016 (Actual) | **(iv)**<br>**2016 (Forecast)** |
|---|---|---|---|---|
| Jan | 2.1 | **0.32** | 0 | **0.92** |
| Feb | 141.3 | **139.82** | 0 | **0.47** |
| Mar | 113.9 | **119.87** | 71.0 | **68.53** |
| Apr | 67.5 | **65.42** | 91.3 | **88.67** |
| May | 72.4 | **75.03** | 234.8 | **222.69** |
| Jun | 268.6 | **273.22** | 124.9 | **121.40** |
| Jul | 74.4 | **73.72** | 115.6 | **113.49** |
| Aug | 58.9 | **60.31** | 81.5 | **79.47** |
| Sep | 84.4 | **76.02** | 301.1 | **298.32** |
| Oct | 265.4 | **249.24** | 343.9 | **335.17** |
| Nov | 31.6 | **30.99** | 48.5 | **30.32** |
| Dec | 2.1 | **1.782** | 14.2 | **10.56** |

Table 4.24 shows the summary statsistics of forecast and the original data for the years 2015 and 2016. The approximate average values for the 24 original and forecast data points are respectively 109 and 106 with respective standard error of mean as 21 and 20. The observations are also at a close range of 335-355 with a variance of 10189 for the original data and 9867 for the forecast.

Table 4.24: Summary of forecast and original data

| | Original (approximate value) | Forecast (approximate value) |
|---|---|---|
| Mean | 109 | 106 |
| Size (N) | 24 | 24 |
| Standrad Deviation | 101 | 99 |
| Standard error of mean | 21 | 20 |
| Sum | 2609 | 2535 |
| Minimum | 00 | 0.32 |
| Maximum | 344 | 335 |
| Range | 344 | 335 |
| Variance | 10189 | 9867 |

## 4.15 Graphical illustrations of the AIC and Residual variances

The following figures show the plots of the AIC and the residual variances of the models in order to further establish the performances of the fitted linear models in comparison with one another as well as in comparison with the fitted nonlinear models. Also, AIC and the residual variances of the fitted nonlinear models are equally examined graphically as follows.

Figure 4.6 reveals that the nonlinear pure SARIODBL model have lower AIC values than the linear pure SARI models at all length of seasons represented on the horizontal axis by ‘value’.

Figure 4.7 also reveals that the fitted nonlinear mixed one-dimensional SARIODBL models have lower AIC values that the existing linear mixed SARI models at every length of season.

It is obvious from Fig.8 that the fitted nonlinear pure one-dimensional SARIMAODBL models have lower AIC values that the existing linear pure SARIMA models at every length of season.

Figure 4.9 also reveals that the fitted nonlinear mixed one-dimensional SARIMAODBL models have lower AIC values that the existing linear mixed SARIMA models at every length of season.

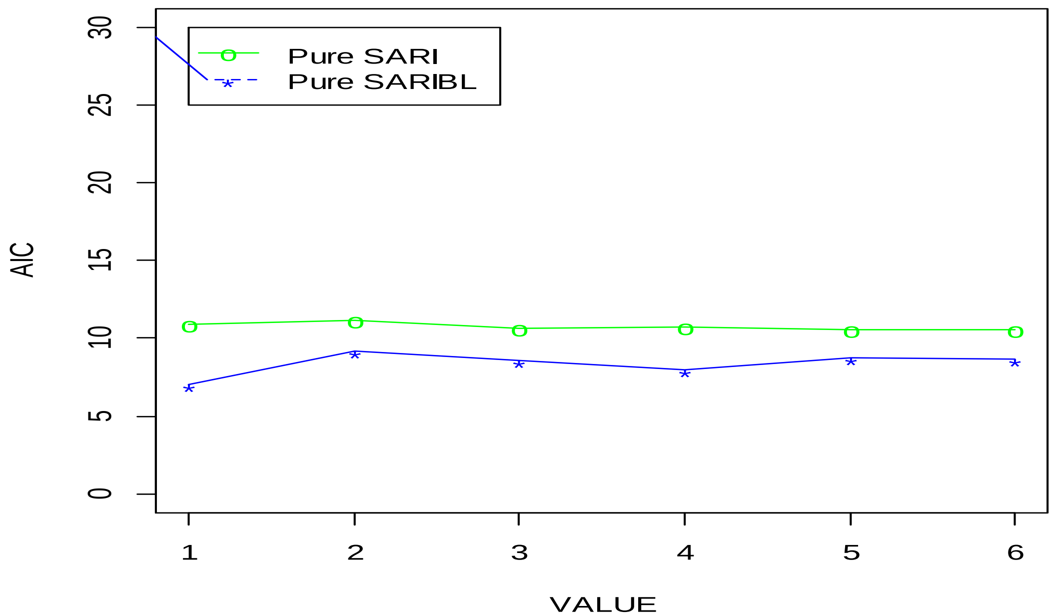


Figure 4.6: AIC values of PURE SARI and pure SARIBL

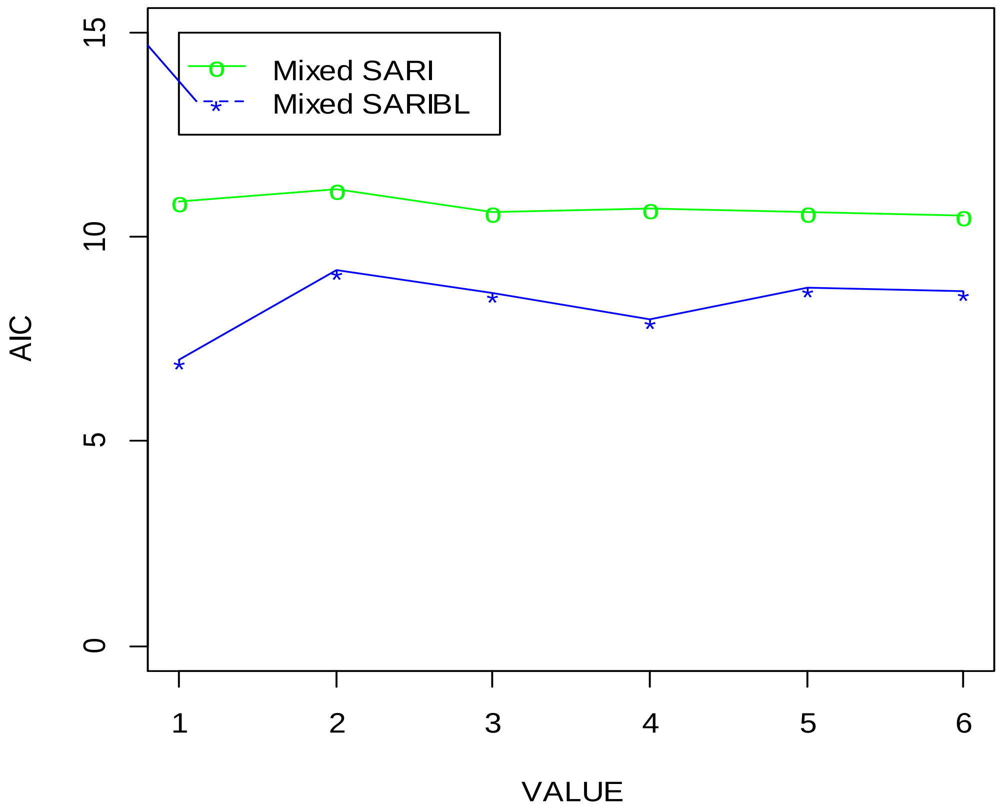


Figure 4.7 AIC values of Mixed SARI and Mixed SARIBL

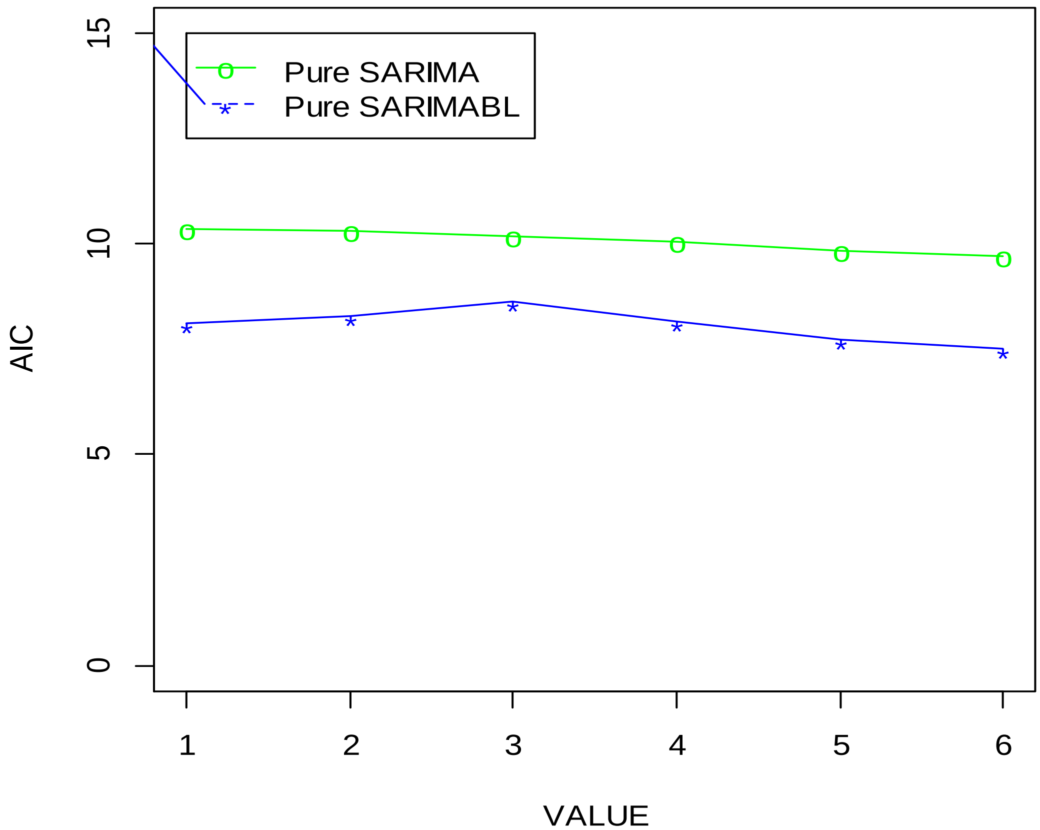


Figure 4.8: AIC values of Pure SARIMA and Pure SARIMABL

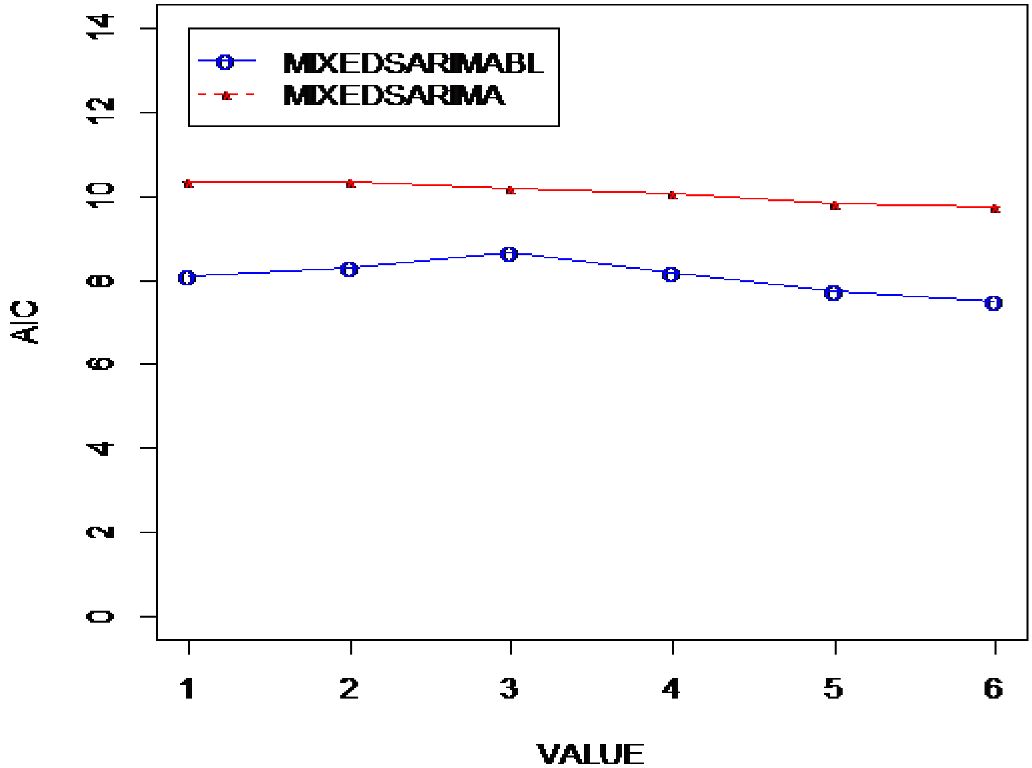


Figure 4.9: AIC values of Mixed SARIMA and Mixed SARIMABL

Comparing the performances of the fitted nonlinear pure SARIODBL models and the fitted nonlinear mixed SARIODBL models in figure 4.10, it shows that the fitted nonlinear mixed one-dimensional SARIODBL models have lower AIC values that the fitted nonlinear pure SARIODBL models at every length of season.

As observed earlier, all the nonlinear One-dimensional Bilinear models perform better than their linear counterparts. This is clearly indicated in the resisual variance plots shown figures 4.11 to 4.18. Moreover, they reveal that the nonlinear Mixed Seasonal Autoregressive Integrated One-Dimensional Bilinear Model shows a better performance than the nonlinear Pure Seasonal Autoregressive Integrated One-Dimensional Bilinear model.

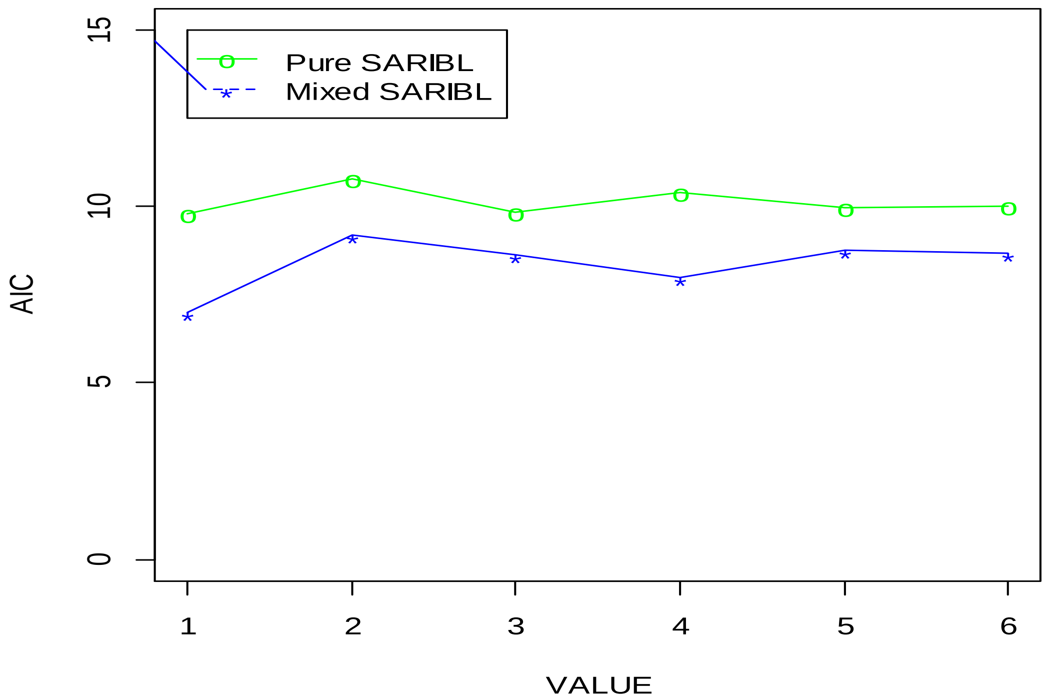


Figure 4.10: AIC values of Pure SARIBL and Mixed SARIBL

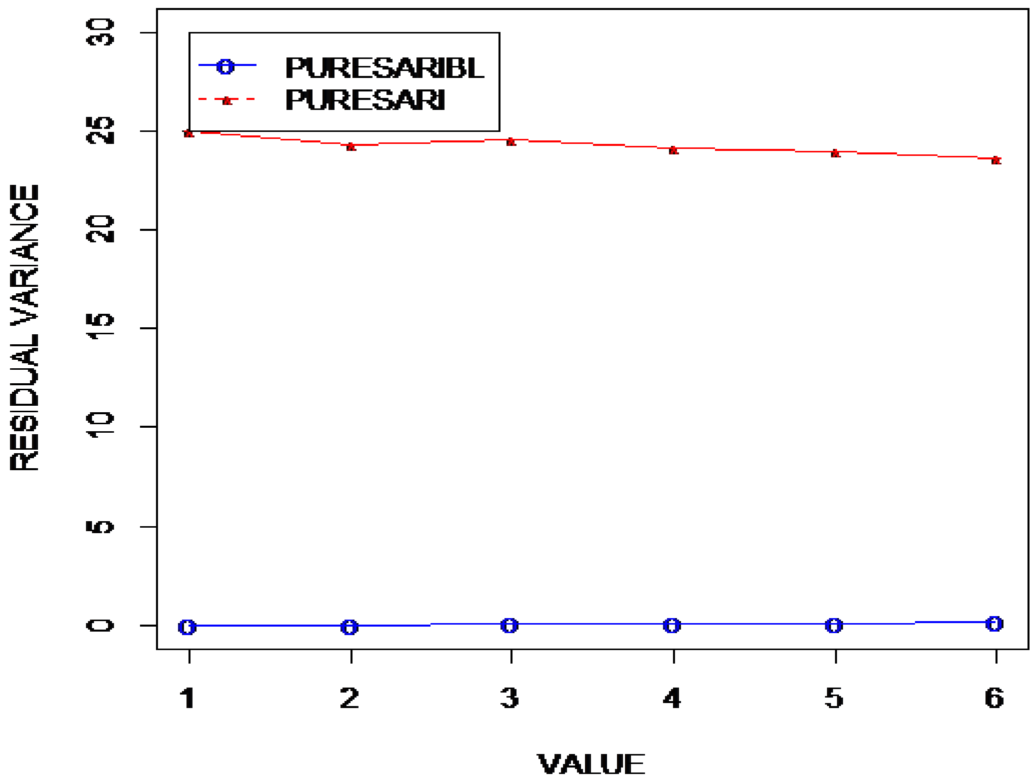


Figure 4.11: Residual Variance of Pure SARI and Pure SARIBL

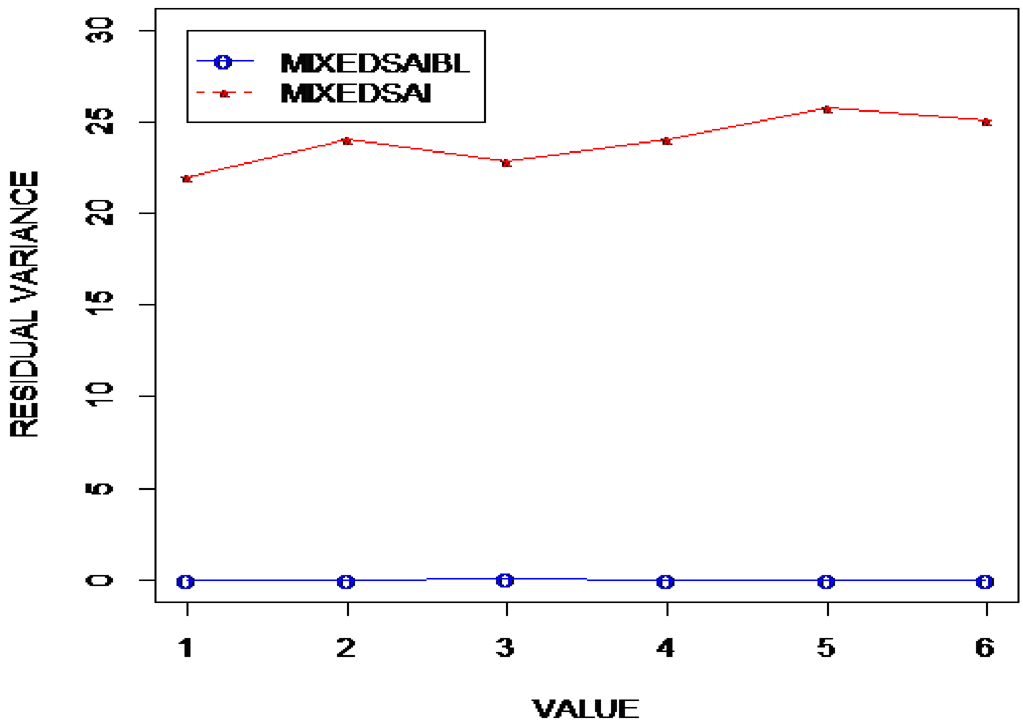


Figure 4.13: Residual Variance of Mixed SARI and Mixed SARIBL

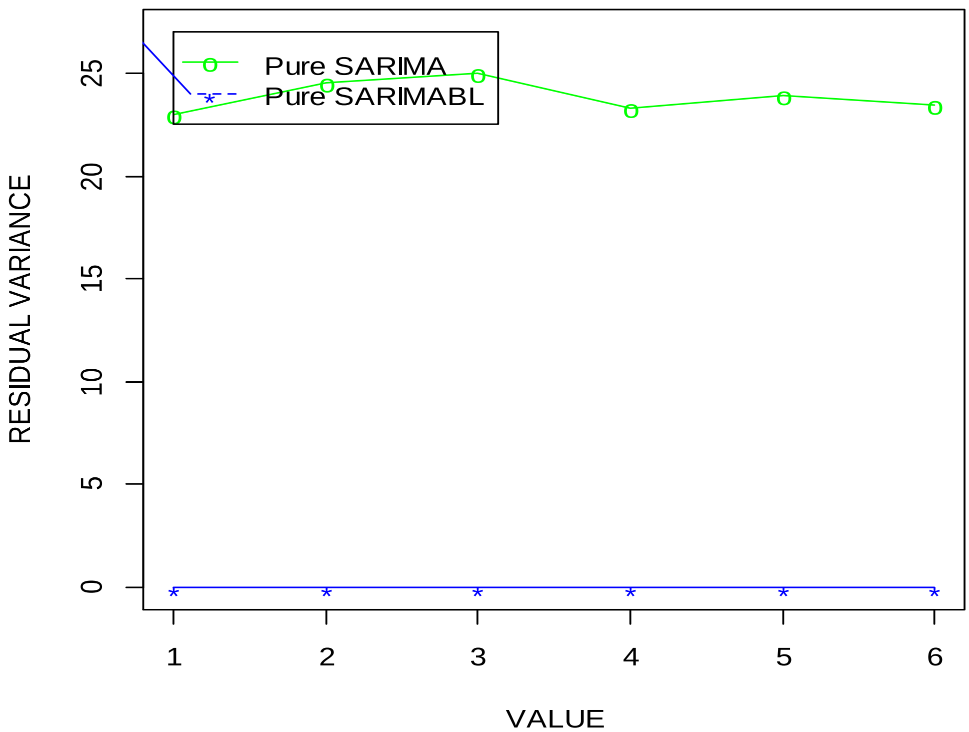


Figure 4.14: Residual Variance of Pure SARIMA and Pure SARIMABL

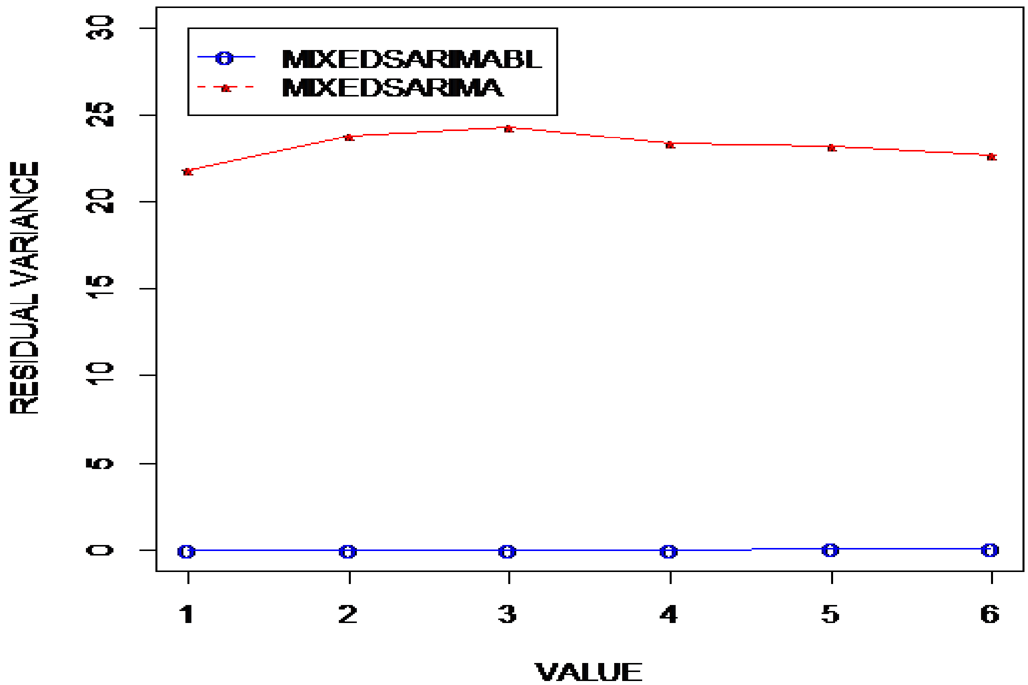


Figure 4.16: Residual Variance of Mixed SARIMA and Mixed SARIMABL

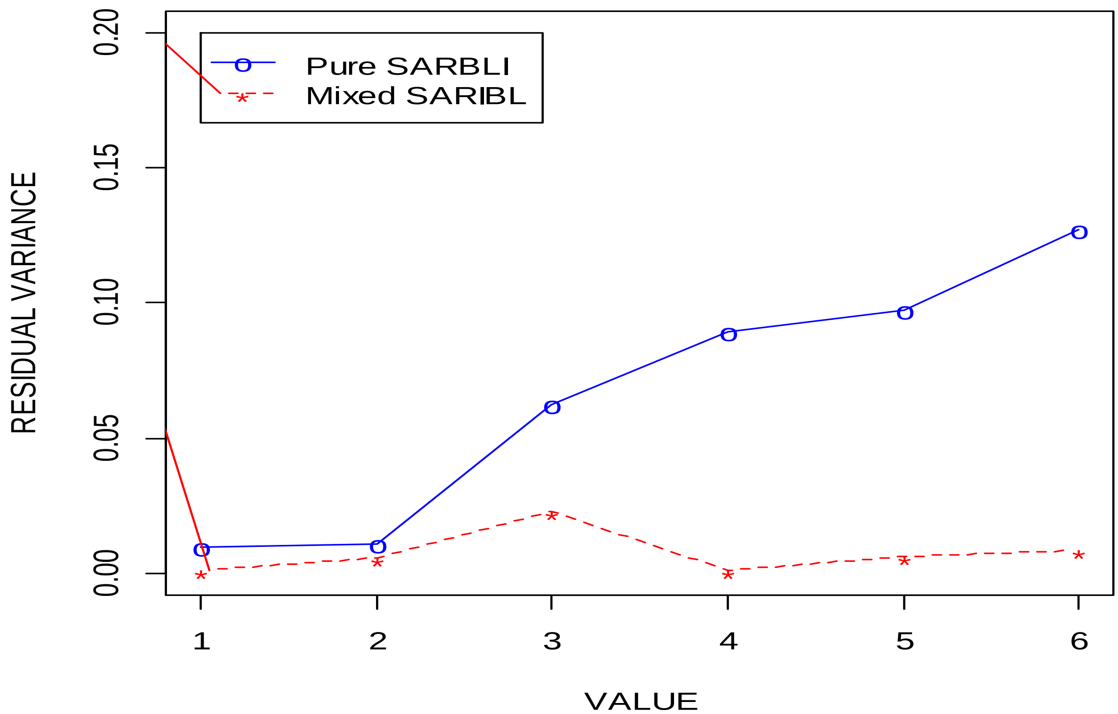


Fig. 4.12: Residual Variance of Pure SARIBL and Mixed SARIBL

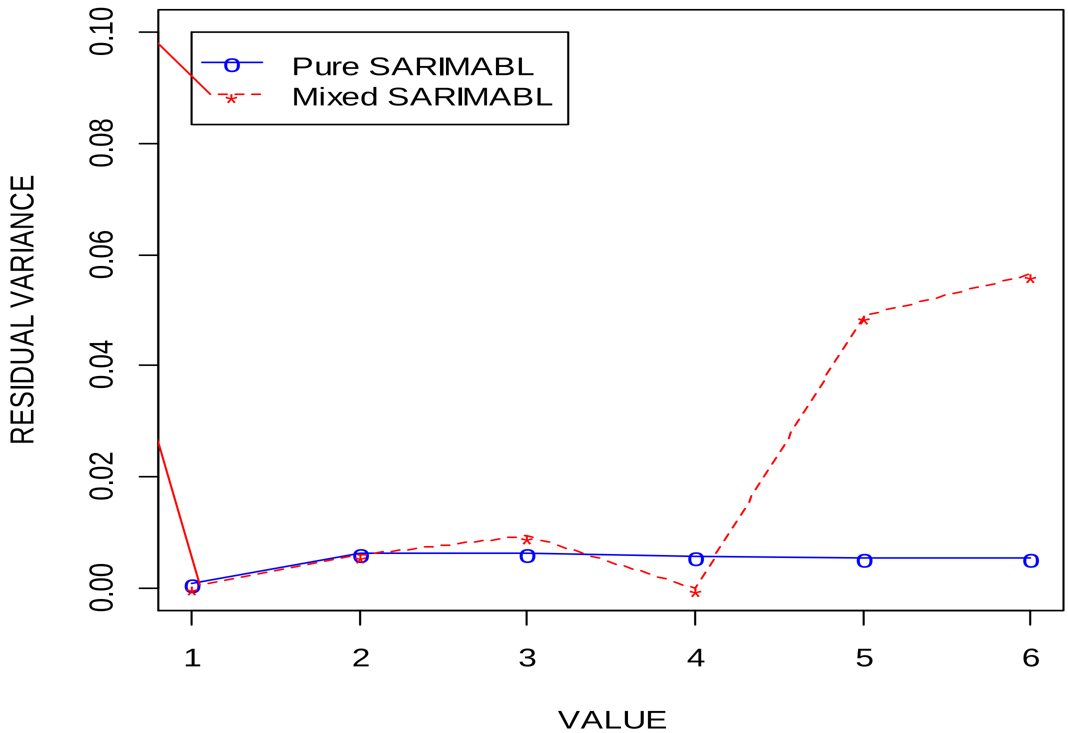


Figure 4.15: Residual Variance of Pure SARIMABL and Mixed SARIMABL

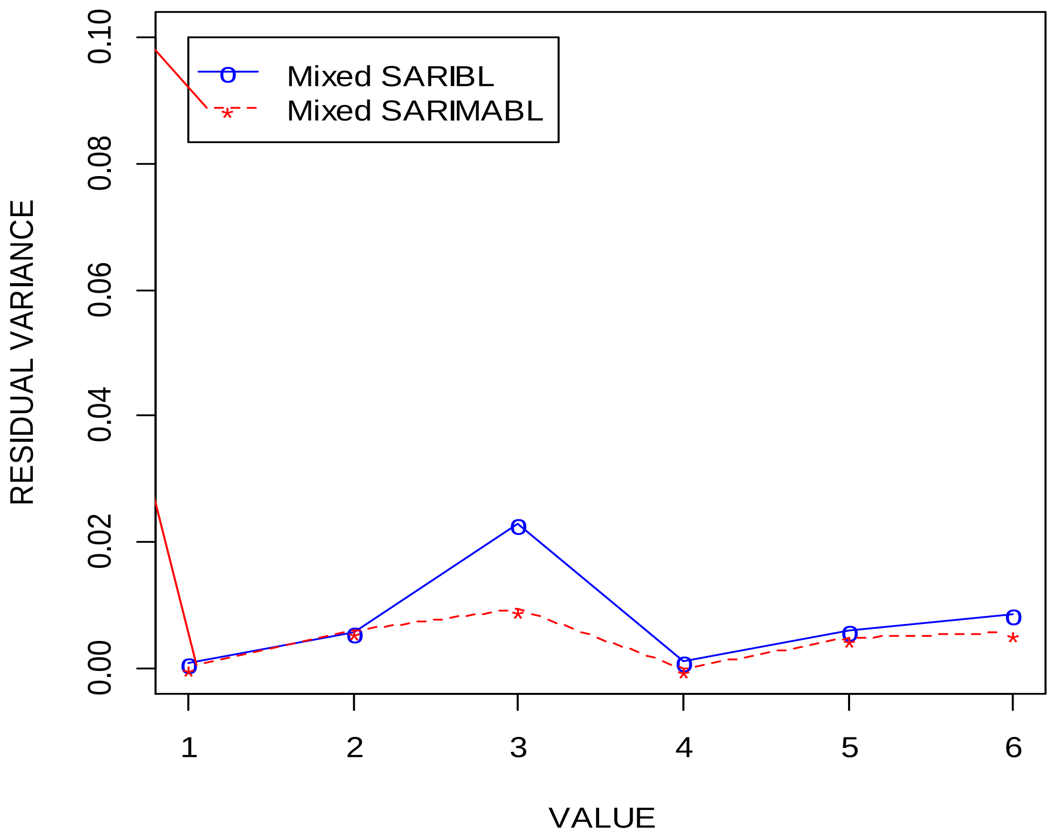


Figure 4.18: Residual Variance of Mixed SARIBL and Mixed SARIMABL

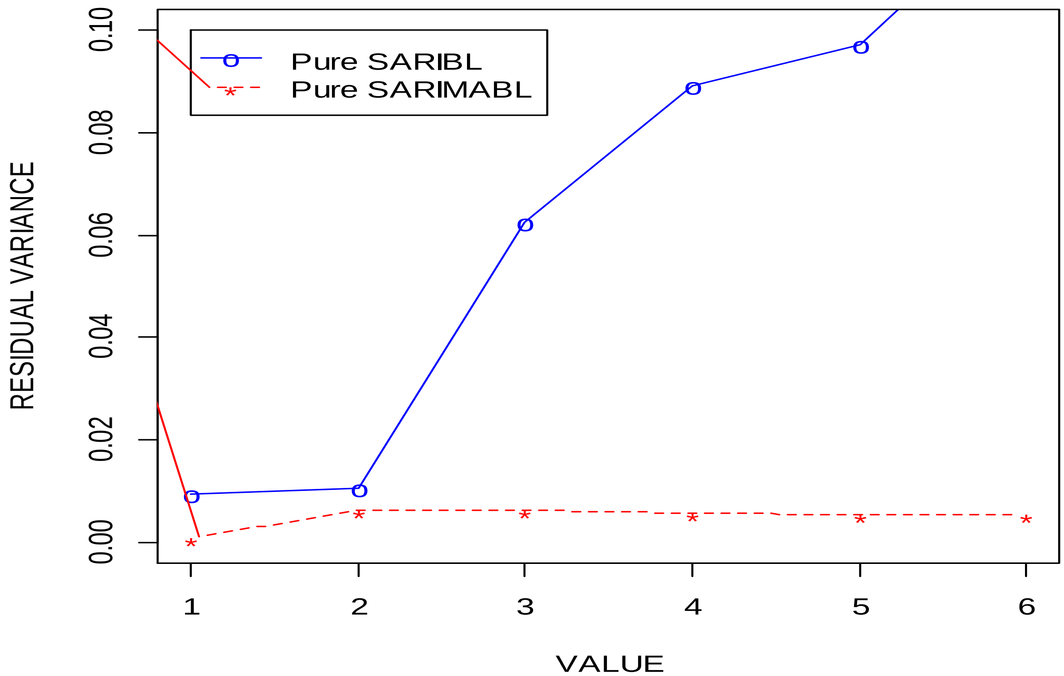


Figure 4.17: Residual Variance of Pure SARIBL and Pure SARIMABL

The performances of the four fitted nonlinear models are examined using the plot of their residual variances as shown in fig. 4.19. The mixed seasonal one-dimensional bilinear time series (MSARIMAODBL) models have the least residual variance, followed by the Mixed SARIODBL then the pure SARIODBL and the pure SARIMAODBL in that order.

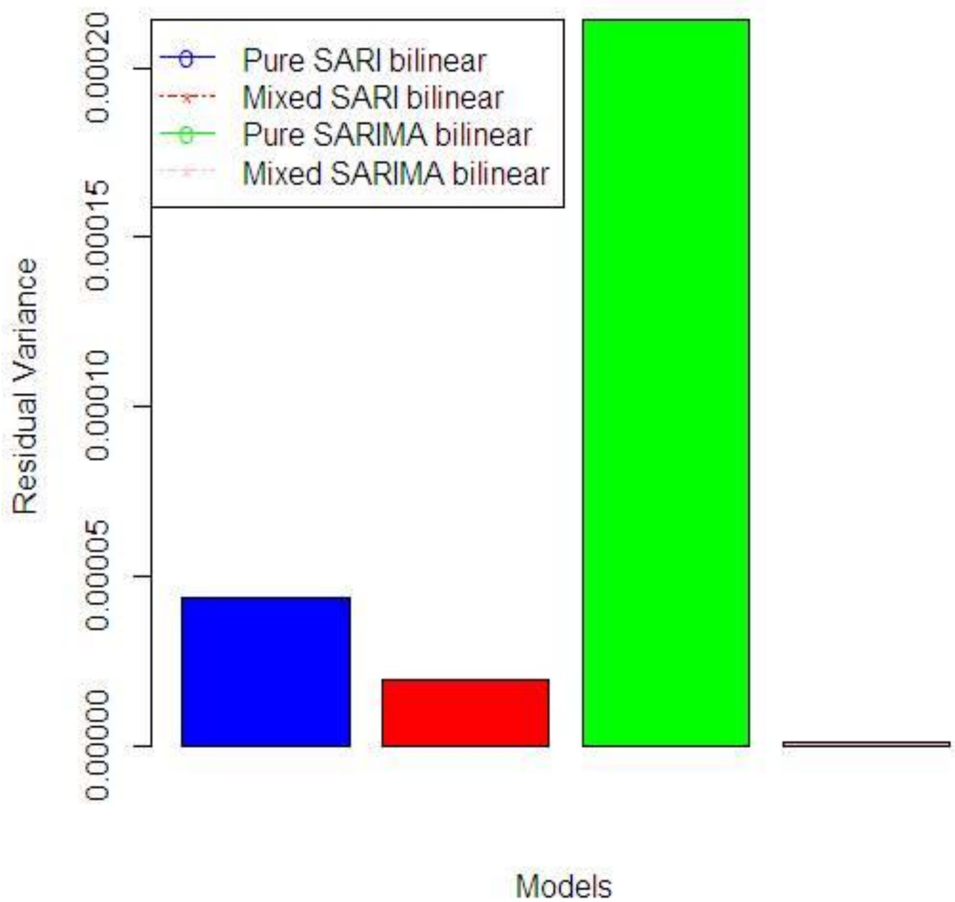


Figure 4.19: A bar chart showing the residual variances of the four fitted models.

# CHAPTER 5

# SUMMARY, CONCLUSION AND SUGGESTIONS FOR FURTHER RESEARCH

## 5.1 Summary

In summary, perusal of existing linear existing literatures on linear seasonal time series particularly the seasonal autoregressive integrated moving average (SARIMA) model has revealed the performance of a time series before and after the peaks of season. Also, the existing nonlinear pure SARIMA first order bilinear studied a series only at the peak of seasons. However, we realize that little or no attention has been given to the development of a seasonal time series model that would track the behaviour a time series before, at and after the peaks of season. Hence, this study developed the mixed SARIMA one-dimensional bilinear (MSARIMAODBL) time series model that is capable of studying the performance of a series before, at and after the peaks of season. The developed model gave rise to three other one-dimensional bilinear time series models as subsets as follows.

(i) Pure SARI One-Dimensional Bilinear Time series Model,

(ii) Mixed SARI One-Dimensional Bilinear Time series Model,

(iii) Pure SARIMA One-Dimensional Bilinear Time series Model,

The stationarity conditions and the estimation of their parameters are considered. Then, both real life and simulated data are used to justify the performance of these models at different lengths of season.

The results show clearly that at the different lengths of season‘s’:

- The nonlinear Pure seasonal autoregressive integrated one-dimensional bilinear models give better performance than their linear Pure seasonal counterparts.
- The Mixed Seasonal Autoregressive Integrated models also performed better than their pure seasonal counterparts.

- Also, the nonlinear Mixed seasonal autoregressive integrated one-dimensional bilinear models give better performance than their linear Mixed seasonal counterparts.
  Similarly;
- The mixed seasonal autoregressive integrated moving average models performed better than the pure seasonal autoregressive integrated moving average models.
- The nonlinear pure seasonal autoregressive integrated one-dimensional bilinear models performed better than the nonlinear pure seasonal autoregressive integrated moving average one-dimensional bilinear models.
- The nonlinear mixed seasonal autoregressive integrated moving average models one-dimensional bilinear models performed better than the nonlinear mixed seasonal autoregressive integrated one-dimensional bilinear models.
- This optimum model is given by the mixed SARIMAODBL model:

$$X_t = \psi_1 X_{t-1} + \theta_1 e_{t-1} + \Psi_1 X_{t-4} + \Psi_2 X_{t-8} + \Theta_1 e_{t-4} + \Theta_2 e_{t-8} + b_{11} X_{t-1} e_{t-1} + e_t$$

## 5.2 Conclusion

Based on the results above, we conclude that the fitted nonlinear seasonal models perform excellently better than their linear counterparts at every length of season under consideration. Similarly the fitted mixed seasonal models perform far better than the pure seasonal models at all length of seasons. Moreover, the mixed seasonal autoregressive integrated moving average one-dimensional bilinear time series model (MSARIMAODBL) captures seasonality better and performs best in terms of estimation along a trend line and at the peak of seasons when compared with other nonlinear seasonal bilinear models at different length of seasons.

Moreover, this study has made us to realize that since a mixed seasonal time series model would help us to study the performance of seasonal time series both along a trend line and at the peak of seasons, then it is of great importance for everyone or organization who engages in businesses or occupations that can easily be affected by seasonal fluctuation. For example farmers that produce crops which can only be available for harvest at a particular period of the year or companies that produce and sell items used during festive seasons or rainy seasons like umbrella, air-conditioner, Christmas light, and so on. This would help in making logical decisions when those items can be made available to end users and when they are out of season.

Furthermore, the fitted models are better substitutes for existing seasonal time series models in order to obtain better outcome in seasonal time series analysis.

## 5.3 Contribution to knowledge.

Due to better performances of the specified models than their existing counterparts in literature, they are therefore better substitutes and recommended for better analysis of seasonal time series.

Furthermore, since a mixed seasonal time series model would help us to study the performance of seasonal time series both along a trend line and at the peak of seasons, the fitted mixed seasonal models are better substitutes for the existing pure seasonal models in seasonal time series analysis.

## 5.4 Suggestions for further studies.

The following are therefore suggested for further studies:

(i) Extension of the mixed seasonal ARIMA One-dimensional Bilinear models to the mixed Seasonal Generalised bilinear models.

(ii) Examination of the Spatial Mixed Seasonal Bilinear time series models.

## REFERENCES


Abdelouahab Bibi. 2006. On the Covariance Structure of Time Varying Bilinear Models. *Journal of Stochastic Analysis and Applications*. Vol. 21. doi.org/10.1081/SAP-120017531

Adubisi, O. D., Eleke, C.C., Mom, T.T. and Adubisi, C.E. 2017. Application of SARIMA to Modelling and Forecasting Money Circulation in Nigeria. *Research Gate*. doi: 10.9734/ARJOM/2017/35555.

Ahmed, T., Hesham H. and Amr, B. 2017. Forecasting of nonlinear time series using ANN. *Science Direct*. 1. 39- 47.

Akaike, H. 1970. Statistical predictor identification. *Annals of the Institute of Statistical Mathematics.* 22. 203–17

Akaike, H. 1974. A new look at the statistical model identification. *IEEE transactions on automatic control*. 6. 716-723.

Akaike, H. 1974. Markovian Representation of Stochastic processes and its application to the analysis of Autoregressive Moving Average processes. *Ann. Ins. Statisti. Math.* 26. 363-383.

Akintunde, M.O., Olalude, G.A. and Oseghale, O.I. 2017. Forecast Comparison of Seasonal Autoregressive Integrated Moving Average (SARIMA) and Self Exciting Threshold Autoregressive (SETAR) Models. *American Journal of Theoretical and Applied Statistics*. 6. 278-283. doi:10.11648/j.ajtas.20170606.13

Anderson, T.W. 1971. The statistical Analysis of Time series. *New York and London Wiley*.

Anderson, T.W. 1977. Estimation for Autoregressive Moving Average models in the time and Frequency domains. *The Ann of Statistics.* 5. 842-865.

Anthony, E.U. 2018. Modelling of Nigeria gross domestic product using seasonal and bilinear autoregressive integrated moving average models. *Journal of Statistical and Econometric Methods*. 2. 1-21

Benkachcha. S. B. Benhra, J. and El Hassani H. 2015. Seasonal Time Series

Forecasting Models based on Artificial Neural Network. *International Journal of Computer Applications*. 20. 0975-8887.

Beaulieu, J. J., Miron, J. A. 1993. Seasonal unit roots in aggregate U.S. data. *Journal of Econometrics* 55. 305-328.

Bhaskara Rao, M., Subba Rao, T. and Walker, A.M. 1983. On the existence of Bilinear Time Series Models. *Journal of Time Series Analysis.* 2. 95110. doi.org/10.1111/j.1467-9892.1983.tb00362.x

Box, G.E.P and Jenkins, G.M. 1976. Time Series Analysis: Forecasting and Control, Revised Edition. *San Francisco: Holden Day*.

Box, G.E.P and Jenkins, G.M. 1970. Time series Analysis, Forecasting and Control. *San Francisco: Holden Day*.

Billy, M. W. and Lester, A. H. 2003. Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results. *Journal of transportation Engineering.*

Broemeling, L.D. and Shaarawy, S.1988. Time Series: A Bayesian Analysis in the Time Domain. In: Bayesian Analysis of Time Series and Dynamic Models, J.C. Spall. 2nd ed. 1-22, New York. Marcel Dekker.

Cagdas, H.A., Erol, E. and Cem, K. 2009. Forecasting nonlinear time series with a hybrid methodology. *Science Direct*. 9. 1467-1470. doi.org/10.1016/j.aml.2009.02.006

Campbell, M.J. and Walker, A.M. 1977. A survey of statistical work on the McKenzie River series of annual Canadian lynx trappings for the years 1821-1934, and a new analysis. *J.R. Statist. Soc. A* 140. 411-431.

Cathy, W. S. C. 1992. Bayesian inferences and forecasting in bilinear time series models. *Journal of Communications in Statistics - Theory and Methods*. Vol. 21.

Cauchy, A. L. 1821. Cours d’Analyse de L’Ecole Royale Polytechnique. Premi`ere Partie. Analyse alg´ebrique (Paris: Imprim´erie Royale)

Chatfield, C. 1996. Model uncertainty and forecast accuracy. *J. Forecasting* 15. 495-508.

Chatfield, C. 2003. *The Analysis of Time series: An introduction.* 6th ed.

Chikezie, D.C. 2007. Estimation of the parameters of the Bilinear Seasonal ARIMA Time series model. *Global journal of Mathematical Sciences*. 1. 1-4.

Christian, F. 2007. ARMA models with bilinear innovations. *Journal of communications in Statistics. Stochastic Models.* Vol.15.

Edwards, J. H. 1961. The recognition and estimation of cyclic trends. *Annals of Human Genetics*. 1. 83-86. doi: 10.1111/j.1469- 1809.1961.tb01501.x

Eleazar, C.N., Iheanyi, S.I. and Valentine, U.N. (2016). Some Tests for Seasonality in Time Series Data. *Journal of Modern Applied Statistical Methods.* 2. 382-399. doi: 10.22237/jmasm/1478002920.

George, E.P. B., Steven, G.H. and George C.T. 1979. Analysis and Modeling of seasonal time series. *National Bureau of Economic Research. (NBER)*. 309 - 346.

Ghysels, E., Lee, H.S. and Noh, J. 1994. Testing for unit roots in seasonal time series: some theoretical extensions and a Monte Carlo investigation. *Journal of Econometrics.* 62. 415-442.

Grahn, T. 1995. A conditional least squares approach to bilinear time series estimation. *Journal of Time Series Analysis* 16. 5 doi.org/10.1111/j.1467-9892.1995.tb00251.x

Granger, C.W.J. and Anderson, A.P. 1978a. Introduction to Bilinear time series Models. *Vandenhoeck and Puprecht*.

Granger, C.W.J. and Anderson, A.P. 1978b. Nonlinear Time series Modelling in applied Time series Ananlysis. D.F Findely, ed. *New York: academic press*.

Granger, C.W.J. and Newbold P. 1976. Forecasting transformed data. *J.R.S.S. Series B.* 38. 184-203.

Haggan, V. and Ozaki, T. 1980. Amplitude-dependent exponential AR model fitting for non-linear random vibrations. *In pro. 1979 Nottingham International Time Series Meeting, ed*. O.D. Anderson, North-Holland, Amsterdam.

Haggan, V. and Ozaki, T. 1981. Modelling nonlinear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika* 68. 189-196.

Hannan, E. J. 1962. Time Series Analysis. *London: Methuen Monographs.*

Hannan, E.J. 1970. Multiple Time Series. *New York: John Wiley.*

Hannan, E.J. 1980. A note on bilinear time series models. Stochastic Processes and their Applications 22. 1-224. *North-Holland Publishing Company*.

Harvey, A. C., and Scott, A. 1994. Seasonality in dynamic regression models. *Economic Journal*. 104. 1324-45.

Hinnich, M. J. 1982. Testing for Gaussianity and Linearity of a Stationary Time Series. *Journal of Time Series Analysis*. 3. 169-176.

Hipel, K.W. and McLeod, A.I. 1994. Time Series Modelling of Water Resources and Environmental Systems. *Amsterdam, Elsevier*.

Hjellvik, V., Yao, Q. and Tjostheim, D. 1998. Linearity testing using local polynominal approximation. *Journal of statistical planning and inference.* 2. 295-321. doi:10.1016/S0378-3758(97)00146-8

Huamin, Z. and Feng D. 2013. On the Kronecker Products and Their Applications. *Journal of Applied Mathematics.* Article ID 296185, doi.org/10.1155/2013/296185

Hylleberg, S., Engle, R.F. Granger, C.W.J. and B.S. Yoo; (HEGY). 1990. Seasonal Integration and Cointegration. *Journal of Econ.*, 44. 215-238.

Hylleberg, S. 1992. Modelling Seasonality. *Oxford University Press, Oxford*.

Hylleberg, S., Jorgensen, C. and Sorensen, N.K. 1993. Seasonality in macro-economic time series. *Empirical Economics*, 18, 321-335. doi.org/10.1007/BF01205406

Hyndman, R. J. 2010. Forecasting functions for time series. R version 2.09. 560 http://robjhyndman.com/software/forecast/

Ismail, S. and Shabri, A. 2014. Time series forecasting using least square support vector machine for canadian lynx data. *Journal of Physics:* Conference Series. 5. 11–15.

Jian Liu. 1992. Spectral Radius, Kronecker Products and Stationarity. *Journal of Time Series Analysis.* Vol. 13.

Jian Liu and Peter, J. B.1988. On the general bilinear time series model. *Journal of Applied Probability*. 3. 553-564.

John, M. V. and Todd, K. Moon. 1991. Parameter estimation for a bilinear time series model. *IEEE Xplore Digital Library.*

Jones, D.A. 1978. Nonlinear Autoregressive Processes. *Proc. Royal Soc. London* (A). 360; 71-95.

Keenan, D. 1985. A tukey nonadditivity-type test for time series. *Biometrika* 72. 39–44.

Kennedy, P. E. 1998. Using Monte Carlo studies for teaching econometrics. In W. E. Becker and M. Watts, eds., Teaching undergraduate economics: Alternatives to chalk and talk. *Cheltenham, U. K.: Edward Elgar*, 141-59.

Liu, J. 1992. On stationary and Asymptotic Inference of bilinear time series models. *Statistica Sinica.* 2. 479-494.

McLaren, C.E., Legler, J.M. and Brittenham, G.M. 1994. The generalized goodness-of-fit test. *Royal Statistical Society*. 2. 247-258. doi:10.2307/2348342

Mohammed, B. and Wissam, B. 2016. Periodic integer-valued bilinear time series model. *Journal of communications in Statistics- Theory and Methods***.** doi.org/10.1080/03610926.2015.1014107

Mohler, R. R. 1973. Bilinear Control Processes. *New York; Academic Press*.

Momani, M. and Naill, P.E. 2009. Time series analysis model for rainfall data in Jordan: Case study for using time series analysis. *American Journal of Environ. Sci.,* 5. 599-604. doi:10.3844/ajessp.2009.599.604

Neudeker, H. 1969. Some Theorems on Matrix differentiation with special reference to Kronecker Matrix Products. *Journal of American Statistics Association.* 327. 953-963.

Ojo, J.F. 2010. On the Estimation and Performance of One-dimensional Autoregressive Integrated Moving Average Bilinear Time Series Models. *Asian Journal of Mathematics & Statistics,* 3. 225-236.

Ojo, J.F. 2011. Generalized and subset integrated autoregressive bilinear time series models with applications. ISSN: 2153-649X. doi: 10.5251/ajsir.2011.2.3.401.408

Ojo, J.F. 2012. Bilinear time series model for estimating a disease death rate. *JORIND.* 10(3). ISSN 1596-8308. www.ajol.info/journals/jorind.

Omekara, O.C. 2016. The Properties of Pure Diagonal Bilinear Models. *American Journal of Mathematics and Statistics.* 4. 139-144.

Osborn, D. R., Rodrigues, P. M. M. 1998. The asymptotic distributions of seasonal unit root tests: a unifying approach. Discussion Paper 9811, *School of Economic Studies, University of Manchester*.

Oster, G. and Ipaktchi, A. 1978. `Population cycles', Theor. Chem. Periodicity. *New York: Academic Press*, 111-132.

Ozaki, T. and Oda, H. 1977. Nonlinear time series model identification by Akaike's information criterion. *In information and Systems, ed. B. Dubuisson. Pergamon Press, Oxford.*

Ozaki, T. 1980. Nonlinear time series model for non-linear random variables. *Journal of Applied probability*. 17. 84-93

Pham Dinh, T. 1985. Bilinear Markovian representation and Bilinear Models. *Science Direct*. 2. 295-306**.**

Pham Dinh, T. and Tat Tran, L. 1981. On the first order Bilinear Time Series. *Journal of Applied Probability.* 3. 617-627.

Poskitt, D.S and Tremayne A. R. 1986. The selection and use of linear and bilinear time series models. *Science Direct*. 1. 101-114. doi.org/10.1016/0169-2070(86)90033-6

Priestely, M.B. 1978. Nonlinear Models in Time series Analysis. *The Statistician.* 27. 159-176.

Priestely, M.B. 1980. State-Dependent Models; General Approach to Non- linear Time series Analysis. *Journal of Time series Analysis.* 1, 47-71.

Ruberti, A., Isidori, A., and D'Alessandro, P. 1974. Realization and Structure Theory of Bilinear Dynamical Systems. *SIAM Journal on Control.* 3. 517-535. doi.org/10.1137/0312040

Sawa, T. 1978. The exact moments of the least squares estimator for the Autoregressive Model. *Journal of Econometrics*. 8. 159-172.

Schwartz, G.E. 1978. Estimating the dimension of a model. *Ann. Stat*. 6. 461–464

Sangodoyin, D. K., Ojo, J.F. and Kozak, M. 2010. Subsetting and Identification of Optimal Models in One-Dimensional Bilinear Time Series Modelling. *International Journal of Management Science and Engineering Management.* 4. 252-260.

Sarma, Y. and Singh, N. 2007. Markovian representation of a bilinear time series model and maximum likelihood estimation of the parameters. *Journal of Stochastic Analysis and Applications.* 2. 309-321. doi.org/10.1080/07362990008809670.

Subba Rao, T., and M. M. Gabr. 1980. A Test for Linearity of Stationary Time Series. *Journal of Time Series Analysis*. 1. 145-158.

Subba Rao, T. 1981. On the Theory of Bilinear Time Series Models. *Journal of the Royal Statistical Society*. Series B. 2. 244-255.

Subba Rao, T. and Gabr, M. M. 1984. An introduction to Bispectral analysis and Bilinear Time Series Models. *Springer-Verlag New York*.

Subba Rao, T. and M. Eduarda A da Silva. 1992. Identification of Bilinear Time series model. *University of Manchester institute of Science and Technology, U.K*.

Tong, H and Lim J. 1980. A view on non-linear time series model building. In proc.

1979 *Nottingham international tims series Meeting, ed. O.D. Anderson.* North-Holland, Amsterdam.

Tong, H. 1990. Non-Linear Time Series: A Dynamical System Approach. *Oxford University Press*.

Tsay, R. S. 1986. Nonlinearity Test for Time Series. *Biometrika*. 73(2): 461-466.

Tukey, J. W. 1949. One Degree of Freedom for Non-additivity. *Biometrics*. 5: 232-242.

Tukey, J. W. 1958. Bias and confidence in not quite large samples (abstract). *Annals of Mathematical Statistics* 29, 614.

Valipour, M. 2015. Long-term runoff study using SARIMA and ARIMA models in the United States. doi.org/10.1002/met.1491

Volterra, V. 1930. Theory of Functionals and of Integral and Integro-Differential Equations, *Blackie and Sons, London*.

Wang, Y. 2008. Applied Time Series Analysis. 1st Edn., *China Renmin University Press, Beijing.*

Whittle, P. 1951. Hypothesis Testing in Time Series Analysis. *University of Uppsala Publication.*

Wiener, N. 1958. Nonlinear problems in random theory. *New York: Wiley*

Wiredu, S., Nasiru, S. and Asamoah, Y.G. 2013. Proposed Seasonal Autoregressive Integrated Moving Average Model for Forecasting Rainfall Pattern in the Navrongo Municipality, Ghana. *Journal of Environment and Earth Science*. www.iiste.org ISSN 2224-3216.

Wold, H. 1938. A Study in the Analysis of Stationary Time Series. *Almqvist and Wiksell, Sweden.*

Xiaosheng Li, Chunliu Ma, Haike Lei and Haixia Li. 2013. Applications of SARIMA Model in Forecasting Outpatient Amount. 3. 124-128. doi.org/10.3109/23256176.2013.798456

Xinghua Chang, Meng, G., Yan Wang and Xiyong, H. 2012. Seasonal autoregressive integrated moving average model for precipitation time series. *Journal of Mathematics and Statistics.* 4. 500-505. doi:10.3844/jmssp.2012.500.505

Yajima, Y. 1994. New Models in Time Series Analysis in Theory and Applications of Mathematical Statistics. 175-194, *University of Tokyo Press*.

Yasumasa, M. 1997. A nonparametric test for Nonlinearity by the weighted least squares method. *J. Japan Statist. Soc.* 2. 141-156.

# APPENDIX A

## R Code for Rainfall analysis

```
x<-ts(LAG, start=c(1984,1), end=c(2014,12), frequency =12)

  win.graph(width=4.875, height=2.5,pointsize=8)

  ts.plot(x)

  ts.plot(x,ylabel)

ts.plot(x,ylab="Record of rainfall measured at 06 GMT",xlab="YEAR")

> ts.plot(x,ylab="Record of rainfall measured at 06 GMT in Lagos state",xlab="YEAR")

>  Month=c('J','F','M','A','M','J','J','A','S','O','N','D')

>     points(window(x,start=c(1984,1)),main="LAGOS     STATE     RAINFALL
MEASURED",pch=Month,abline(h=0))

qqline(x)

> qqnorm(x)

> qqline(x)

> qqplot(x)

acf(x)

> acf(x,main=" Auto-Correlation Function(ACF) for Lagos Rainfall Measured")

summary(x)

skewness(x)

kurtosis(x)
```

```
adf.test(x)

PP.test(x)

kpss.test(x, "Level")

white.test(x)

Box.test(LAG)

stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D,  Q), period = S),
include.mean = !no.constant, optim.control = list(trace = trc,

   REPORT = 1, reltol = tol))

stats::ari(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
include.mean = !no.constant, optim.control = list(trace = trc,

   REPORT = 1, reltol = tol))

stats::sarima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
include.mean = !no.constant, optim.control = list(trace = trc,

   REPORT = 1, reltol = tol))

stats::mai(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S),
include.mean = !no.constant, optim.control = list(trace = trc,

   REPORT = 1, reltol = tol))

R> Par <- c(1, 0, 1, 0, 0, 0)

R> fit <- arima(data, order = c(Par[1], Par[2], Par[3]),

+ seasonal = list(order = c(Par[4], Par[5], Par[6])))
phi<-function(par,x){

phi1_AR=c(par[1]i,par[1]ii,par[1]iii)

Bigphi_AR=c(par[2]i,par[2]ii,par[2]iii)
```

```
theta_MA=par[3]

Bigtheta_MA=par[4]

d= c(0,0,0,0)

D=c(1,1,1,1)

z=diff(x, differences=2)

s=(QTR1:QTR4)

residual(AR)=e


model<-par[1]i*(z-z[1])+par[1]ii*(z-z[1]-z[2])+par[1]iii*(z-z[1]-z[2]-z[3]-d)

+par[2]i*(z-z[1]-length(s))+par[2]ii*(z-z[1]*2length(s))+par[2]iii*(z-z[1]*3length(s))

+par[3]*(e-e[1])+par[4]*(e-e[1]) + nls( mm ~z + e, data =dat$rainLagos,start = c( xmid =
0, scal = 1 ),alg = "bilinear", trace = TRUE )

BILNEAR<BI(dat,arma=(1,2)*varying, family = gaussian(), trial.size = 1, lambda,

start = NULL, cov.groups = NULL,hybrid.est = FALSE, offset = NULL, intercept =
TRUE, save.data = FALSE,

control = list(tol = 1e-4, maxit = 100, trace = FALSE, restarts = 5, seed = NULL)

m1 <- cov(dat, PURESARI, h=2,z=NULL,lambda0=0.8, lambda1=0.15, lambda3=1,
lambda4=0.25,

 lambda5=1, mu5=0, mu6=0, qm=12,alpha.prior=c(100,30)*diag(2) +

            matrix(12, 2, 2), prior=0, max.iter=40,

 initialize.opt=NULL)
```

```
m2 <- cov(dat, MSAI, h=2,z=NULL,lambda0=0.8, lambda1=0.15, lambda3=1, lambda4=0.25,
 lambda5=1, mu5=0, mu6=0, qm=12,alpha.prior=c(100,30)*diag(2) +
            matrix(12, 2, 2), prior=0, max.iter=40,
 initialize.opt=NULL)

m3 <- cov(dat, PUREMSAI, h=2,z=NULL,lambda0=0.8, lambda1=0.15, lambda3=1, lambda4=0.25,
 lambda5=1, mu5=0, mu6=0, qm=12,alpha.prior=c(100,30)*diag(2) +
            matrix(12, 2, 2), prior=0, max.iter=40,
 initialize.opt=NULL)

m4 <- cov(dat, PUREMSARIMA, h=2,z=NULL,lambda0=0.8, lambda1=0.15, lambda3=1, lambda4=0.25,
 lambda5=1, mu5=0, mu6=0, qm=12,alpha.prior=c(100,30)*diag(2) +
            matrix(12, 2, 2), prior=0, max.iter=40,
 initialize.opt=NULL)
m5<- cov(dat, MSAIMA, h=2,z=NULL,lambda0=0.8, lambda1=0.15, lambda3=1, lambda4=0.25,
 lambda5=1, mu5=0, mu6=0, qm=12,alpha.prior=c(100,30)*diag(2) +
            matrix(12, 2, 2), prior=0, max.iter=40,
 initialize.opt=NULL)
```

```
regressors <- cbind(x,linearTrend = seq(along = timeseries)/12)

norm <- tsglm(x, link = "log",model = list(past_obs = c(1, 12)), xreg = regressors, distr =
"normal")


marcal(c(SARI,MSAI,PURESARI,MSAIMA,PUREMSAIMA),main="
",type="o",col="blue")

lines(marcal(c(MSE,AIC,BiC,RES,), plot = FALSE),lty=2, color= "dashed")

legend("bottomright", legend = c(SARI,MSAI,PURESARI,MSAIMA,PUREMSAIMA),
pch=c("o","*"),lty=c(1,2),ncol=1,lwd = 1,col= c("blue", "dashed"))

 legend("bottomright", legend = c(SARI,MSAI,PURESARI,MSAIMA,PUREMSAIMA),
lwd = 1,lty = c("green", "dashed"....))

function (x, residual(ar),y tau = 0.5, lambda = NULL, weights = NULL,

   intercept = TRUE, nfolds = 10,

   foldid = NULL, nlambda = 100, eps = 1e-04, init.lambda = 1)

{p <- dim(x)[2]

   if (is.null(order)) { lag <- 1:s:p:q:b:P:Q}

   p_range <- lag + intercept

   n <- dim(x)[1]

   lag_func <- switch(which(c(SARI,MSAI,PURESARI,MSAIMA,PUREMSAIMA),

      if (is.null(lambda)) {

       sample_q <- quantile(residual(ar), tau)
```

```
inter_only_rho <- sum(check(residual(ar) - sample_q, tau))

lambda_star <- init.lambda

searching <- TRUE

while (searching) {

if (lag == "SARI") { init_fit <- fit(x, s, tau, lambda = lambda_star, weights, intercept, Vars)}

else (lag == "MSAI"){

else (lag == "PURESARI"){

else (lag == "MSAIMA"){

else (lag == "PUREMSAIMA"){}}}

if (sum(init_fit$coefficients[p_range]) == 0) {searching <- FALSE }

else {lambda_star

<- inter_only_rho/sum(sapply(init_fit$coefficients[p_range],

pen_func, 1))}}

lambda_min <- eps * lambda_star

lambda <- exp(seq(log(max(lambda_min)), log(max(lambda_star)),

length.out = nlambda))}

models <- list()

fit_models <- TRUE

lam_pos <- 1

if (fit_models == c(SARI,MSARI,PURESARI,MSARIMA,PUREMSARIMA)) {while (fit_models) { if (fit_models)}
```

```
        if (sum(abs(coefficients(models[[lam_pos]])[p_range])) ==
          0 || lam_pos == length(lambda)) {
          fit_models <- FALSE
          lambda <- lambda[1:lam_pos]}
        lam_pos <- lam_pos + 1}
  if (sum(abs(coefficients(models[[lam_pos]])[p_range])) == 0 || lam_pos ==
length(lambda)) {
          fit_models <- FALSE
          lambda <- lambda[1:lam_pos] }
        lam_pos <- lam_pos + 1} }
  cv_results <- NULL
  if (criteria == "c(BIC,AIC)") {
   if (is.null(foldid)) {foldid <- randomly_assign(n, nfolds)}
    for (i in 1:nfolds) {
    train_x <- x[foldid != i, ]
    train_y <- y[foldid != i]
    test_x <- x[foldid == i, ]
    test_y <- y[foldid == i]
    train_weights <- weights[foldid != i]
    if (lag == c(SARI,MSAI,PURESARI,MSAIMA,PUREMSAIMA)) {
    cv_models <- lapplyinit_fit <- fit(x, s, tau, lambda = lambda_star,
```

```
weights, intercept, Vars) }}
  if (cvFunc == "check") {cv_results <- cbind(cv_results, sapply(cv_models, model_eval, test_x,
  test_y, tau = tau)) }
  else {cv_results <- cbind(cv_results, sapply(cv_models, model_eval, test_x, test_y, func = cvFunc))}}
 cv_results <- apply(cv_results, 1, mean) }
 if (criteria == "BIC") {cv_results <- sapply(models, bic)}
 if (criteria == "AIC") {cv_results <- sapply(models, bic, largeP = TRUE) }
 lambda.min <- lambda[which.min(cv_results)]
 return_val <- NULL
 return_val$models <- models
 return_val$cv <- data.frame(lambda = lambda, cve = cv_results)
 colnames(return_val$cv)[2] <- criteria
 return_val$lambda.min <- lambda.min}
```

# APPENDIX B

## R Code for Simulation analysis

X=c(sim(mean),sim(sigma))= ( 0.7662, 191.8901)

e=c(sim(mean),sim(sigma))= ( 0.3995, 117.371)

x<-rnorm(0

phi<-function(par,x){

phi1_AR=c(par[1]i,par[1]ii,par[1]iii)

Bigphi_AR=c(par[2]i,par[2]ii,par[2]iii)

theta_MA=par[3]

Bigtheta_MA=par[4]

d= c(0,0,0,0)

D=c(1,1,1,1)

z=diff(x, differences=2)

s=(QTR1:QTR4)

residual(AR)=e

par[1]i= 0.0835, e[1] = 0.1062, par[2]i= -0.4812, par[2]ii= -0.4703, par[4] = -0.6159, = 0.6159,

mm ~z=0.6813.

model<-par[1]i*(z-z[1])+par[1]ii*(z-z[1]-z[2])+par[1]iii*(z-z[1]-z[2]-z[3]-d)

+par[2]i*(z-z[1]-length(s))+par[2]ii*(z-z[1]*2length(s))+par[2]iii*(z-z[1]*3length(s))

```
+par[3]*(e-e[1])+par[4]*(e-e[1]) + nls( mm ~z + e, data =dat$rainLagos,start = c( xmid =
0, scal = 1 ),alg = "bilinear", trace = TRUE )

residual<- model

e<-rnorm(250, mean(residual),sigma(residual))

sim<-rnorm(250,0, residual(initialmodel))

X=c(sim(mean),sim(sigma))

x<-rep(X,0,50)

repeat for

e<-rnorm(500, mean(residual),sigma(residual))

sim<-rnorm(500,0, residual(initialmodel))

and

e<-rnorm(1000, mean(residual),sigma(residual))

sim<-rnorm(1000,0, residual(initialmodel))
```

# APPENDIX C

Rainfall data presentation

**Source: Department of Metrological Services**

Register of Rainfall in 19…..

Station: IKEJA   Station Number:   State: LAGOS

Latitude: 06°35$^{IN}$   Longitude: 03°20$^{IE}$   Altitude: 128.55m

Record of rainfall measured at 06 GMT in Nigeria and entered against day preceding that on which read, in accordance with the Nigerian instructions in form met. 111/2.

| Date | Jan | Feb | Mar | Apr | May | June | July | Aug | Sept | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | mm | Mm | mm | Mm | Mm | Mm | Mm | mm | mm | mm | mm | mm |
| **1984** | 0 | 0 | 17 | 87.3 | 174.3 | 243.1 | 129.9 | 143.7 | 399.6 | 116.5 | 61.9 | 2.2 |
| **1985** | 4.6 | 4 | 64.3 | 188.2 | 238.2 | 134.1 | 116.2 | 110.9 | 152.6 | 105 | 110.9 | 0 |
| **1986** | 21.2 | 30.4 | 95.6 | 114.2 | 210.2 | 172.3 | 54.3 | 3.9 | 84.2 | 202.1 | 52.5 | 0 |
| **1987** | 0 | 35.8 | 78.7 | 71.7 | 139.1 | 327.3 | 109.2 | 419.1 | 362.1 | 84.3 | 48.5 | 0 |
| **1988** | 3.1 | 58 | 64.5 | 158.1 | 108.9 | 487.4 | 481.7 | 129.9 | 183.4 | 132.1 | 32.2 | 87.3 |
| **1989** | 0 | 0 | 132.1 | 102.7 | 229.6 | 186.9 | 325.5 | 107.2 | 67.2 | 216.3 | 1.2 | 0 |
| **1990** | 0 | 6.5 | 5.8 | 186 | 119.7 | 219.2 | 567 | 16.8 | 183.2 | 180.6 | 54.7 | 71.7 |
| **1991** | 31.9 | 6.3 | 92.9 | 230.6 | 201.4 | 208.3 | 198.2 | 34.8 | 184.9 | 96.6 | 33 | 52.8 |
| **1992** | 0 | 0 | 8.6 | 95 | 353.8 | 231.3 | 154.6 | 16.4 | 152.7 | 86.2 | 79.8 | 5.5 |
| **1993** | 0 | 45.7 | 138.9 | 134.5 | 249.4 | 224.2 | 183.8 | 55.2 | 248.5 | 139.7 | 209 | 46.7 |
| **1994** | 21.3 | 9.1 | 89.1 | 53 | 195.1 | 306.2 | 73.8 | 61.4 | 72.1 | 161.2 | 66.4 | 85 |
| **1995** | 0 | 46.3 | 115.8 | 178.6 | 222.6 | 247.2 | 223.1 | 117.3 | 146.4 | 202.5 | 41 | 48.8 |
| **1996** | 15.3 | 188.3 | 98.7 | 230.4 | 131.2 | 175.5 | 268.4 | 170.4 | 117.4 | 135.3 | 43.3 | 0 |
| **1997** | 0 | 0 | 86.2 | 171.3 | 224.1 | 619.5 | 99 | 139.1 | 225.5 | 287.9 | 126.2 | 41.1 |
| **1998** | 8.8 | 13.6 | 30.8 | 73.2 | 114.8 | 256.9 | 52.2 | 29.6 | 147.7 | 220.2 | 92.1 | 0 |
| **1999** | 17.4 | 36.8 | 51.2 | 189.4 | 67.9 | 338.4 | 347.9 | 29.6 | 113.3 | 221.7 | 102.2 | 26.9 |
| **2000** | 0.3 | 12.5 | 21 | 77.9 | 138.8 | 196.2 | 103.8 | 86.3 | 436.6 | 133.6 | 32.8 | 11.3 |
| **2001** | 1.6 | 10 | 22.2 | 190.4 | 265 | 218 | 106.5 | 24.7 | 255.7 | 183.8 | 59.5 | 54.7 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **2002** | 44.9 | 21.5 | 35.1 | 336.3 | 145.9 | 329.5 | 346 | 45.2 | 188.6 | 142.7 | 149.2 | 9.7 |
| **2003** | 174.4 | 53.4 | 79.1 | 308.1 | 157.4 | 78.1 | 69.5 | 18.5 | 185.2 | 141 | 184.8 | 0 |
| **2004** | 6.4 | 45.7 | 122.8 | 291.1 | 306.1 | 213.9 | 94.5 | 68.5 | 321.2 | 160.9 | 50 | 20.5 |
| **2005** | 0 | 93.1 | 78.2 | 94.1 | 185.3 | 392.3 | 225.3 | 15 | 194.2 | 94.8 | 96.4 | 16.2 |
| **2006** | 44.2 | 10.7 | 121.8 | 26.4 | 294.3 | 264 | 52.8 | 65.7 | 327.6 | 191.3 | 95.3 | 4.6 |
| **2007** | 0 | 0 | 76.1 | 31.6 | 253.7 | 361.7 | 228 | 146.1 | 160.1 | 120.3 | 118.3 | 5.4 |
| **2008** | 0.8 | 3.3 | 69.6 | 96.8 | 230 | 365 | 442.7 | 134.3 | 226.8 | 98.8 | 98.9 | 49 |
| **2009** | 1.6 | 16.3 | 33.9 | 115.5 | 154.2 | 463.2 | 119 | 12 | 84.1 | 342.7 | 48.7 | 0 |
| **2010** | 37.2 | 42.4 | 68 | 126.9 | 159.3 | 368.7 | 130.8 | 190.6 | 235.7 | 122.8 | 193.8 | 76.7 |
| **2011** | 0 | 87.2 | 21.6 | 74.7 | 170.6 | 251.9 | 476.9 | 43.7 | 175.3 | 209.3 | 240.5 | 0 |
| **2012** | 10.5 | 122.2 | 78.1 | 124.7 | 134.9 | 478.8 | 152.1 | 34.3 | 214.1 | 148.9 | 123.2 | 0 |
| **2013** | 133.7 | 34.7 | 121.8 | 202.3 | 339.4 | 108 | 190.8 | 12.6 | 165.3 | 125.6 | 249.9 | 47.9 |
| **2014** | 90.7 | 94.9 | 76.9 | 164.8 | 305.7 | 295.4 | 326.2 | 221.5 | 240.2 | 178.2 | 186.4 | 8.9 |