**International Journal of Engineering, Science and Mathematics**
Vol. 5 Issue 3, September 2016,
ISSN: 2320-0294 Impact Factor: 6.238
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

# TIME SERIES DATA:

# THE APPLICATION OF ARIMA MODEL OF BUSINESS

**MONEY, UDIH (Ph.D)***

**ABSTRACT**

Auto-Regressive Integrated Moving Average (ARIMA) Models are among the most important time series models used in businesses and financial market forecasting over the past decades.

Theoretically, and empirically, findings have suggested that integration of different models can be an effective method of improving upon their predictive performance, especially when the models in the ensemble are quite different. In this paper, ARIMA models are viewed from the introduction, looking at the definitions, assumptions, general notations of ARIMA models; and also the procedure of ARIMA with respect to the three stages of ARIMA modeling 'cum' procedural statements are considered. And we demonstrate the use of ARIMA procedure statements. We demonstrate also that ARIMA model is a complex mathematical or statistical technique for forecasting business variables.

**Keywords: ARIMA, Forcasting, Moving Average, Variables, Processes.**

* **FEDERAL UNIVERSITY OF PETROLEUM RESOURCES P.M.B. 1221, EFFURUN (FUPRE), DELTA STATE, NIGERIA.**

## INTRODUCTION

Every individual, business and government, are involved in business finance; which could be sale and/or purchase of properties/items. And in doing so, data are gathered over the periods. From period to period data are gathered, and information is accumulated at regular intervals; thus we have time series data.

The individual, business and government who gathered all this information or data are in one way or the other involved in decision making, problem solving, and strategy. They forecast or predict the future for business survival. Many statistical and mathematical tools or techniques are used for forecasting or predicting the future of the business.

In real-life, nobody can look into the future, but statistical methods, econometric models and business intelligence software helped in the forecasting of the future. The ARIMA time series analyses uses lags and shifts in the historical data to uncover patterns (e.g. moving averages, seasonality) and predict the future. ARIMA is a complex statistical technique for forecasting. ARIMA methodology introduced by Box and Jenkins helps to uncover the hidden patterns in data, because patterns of data are unclear, and individual observations involve errors; and it generate forecast. ARIMA is popular in many areas and researchers confirmed that it is powerful and flexible (Hoff, 1983; Pankratz, 1983; Vandaele, 1983). It is a complex technique, very very difficult to use, experience is needed; but of benefit, it produces at all time good results.

## DEFINITION OF ARIMA PROCESSES AND ASSUMPTIONS

Autoregressive (AR) process is defined as a “Statistical forecasting model in which future values are computed only on the basis of past values of a time series data”. “It is in the Decision Making, Problem Solving, and Strategy and Statistics, Mathematics, and Analysis Subjects”. http://www.business dictionary.com

Autoregressive moving average model (ARMA) is defined “as a tool for understanding and possibly predicting future values in the time series” (Wikipedia 2006).

Both ARMA Model and ARIMA Model are referred to as Box-Jenkins (B-J) models. And Business Dictionary.com defines B-J models as a "Mathematical Models used typically for accurate short-term forecasts of 'well-behaved' data (that shows predictable repetitive cycles and patterns). It requires at least a moderately long time series (with about a hundred observations) for an effective 'fitting', and generally autoregressive, moving average, and seasonal moving average terms, and difference and seasonal moving average terms, and difference
and seasonal difference operators.

Autoregressive integrated moving average (ARIMA) Model is defined as "autoregressive moving average process (ARMA) Model of a differenced time series (one that has been rendered stationary by the elimination of 'drift' whose output needs to be anti-differenced to forecast the original series. It can represent a wide range of time series data, and are used generally in computing the probability of a future value lying between any two limits. http://www.business dictionary.com

Autoregressive integrated moving average is defined as that which "describes a stochastic process or a model of one. An autoregressive integrated moving average process is made up of sums of autoregressive and moving average components, and may not be stationary. http://economics .about.com

We defined ARIMA Model as the statistical or mathematical model use to predict or forecast, or plan by business in all areas, and cause comparison with actual values, making forecast value to be reliable. And it is ARMA Model that is differenced or an ARMA Model that requires differencing.

ARIMA is not without assumptions, the assumptions are:

* Absence of Outliers
* Shocks are randomly distributed with a mean of zero and constant
variance over time.
* Residuals exhibit homogenous variance over time, and with a mean of
zero.

* Residuals are normally distributed.
* Residuals are independent.
* It is identified with (P, d, q,).
* It is estimated and diagnosed.
* The trend is affected by random shocks.
* P stands for auto-regressive component (autocorrelation)
* d stands for integrated component.
* q stands for the moving average components, which randomized shocks.
* Model specification relies on an examination of (ACF) and (PACF) –

Autocorrelation and partial autocorrelation function respectively.

**NOTATION FOR ARIMA MODELS: AN OVERVIEW**

ARIMA Model i.e. autoregressive integrated moving average model is
generally denoted by the notation ARIMA (p, d, q), where:
P, is defined as the order of the autoregressive part.
d,is defined as the order of the differencing; and
q,is defined as the order of the moving-average process.

ARIMA (p,d,q) gives birth to ARIMA (p,q) is where there is no differencing i.e. (d = o). Thus, auto-regressive integrated moving average model becomes auto-regressive, moving average (ARMA).

Final model for ARIMA (p,d,q) is ARIMA (1,1,1) i.e. P=1, q=1, and d=1.

The notation for a pure ARIMA model is written mathematically as:

$$Wt = \mu + \frac{\Theta (B)}{\text{Ø} (B)} at$$

Where
tis the indexes time.
Wt is the response series Ytor a difference of the response series.
µis the mean term.
B is the backshift operator; that is BXt = Xt - l
Ø(B) is the autoregressive operator, represented as a polynomial in the
backshift operator:

$$Ø(B) = l - Ø, B - \ldots Ø p B^{p}$$

$\Theta$ (B) is the moving-average operator, represented as a polynomial in the backshift operator:

$$\Theta(B) = l - \Theta, B - \ldots Ø q B^{q}$$

atis the independent disturbance, known as random error.

Note that Wt is computed by the identify statement and processed by estimate statement. Thus, Wt iis either Ytor a difference of Yt specified by the differencing operators in the identify statement.

We have Wt for both non seasonal and seasonal differencing as:

$$Wt = (1\text{-}B)^{d} Yt \text{ and}$$

$$Wt = (l - B)^{d}(l - B^{s})^{D} Yt$$

respectively.

Where d and D are the degrees of non-seasonal and seasonal differencing respectively. And S is the length of the seasonal cycle.

An ARIMA (1,1,1) model is mathematically represented as:

$$(l - B)\, Yt = \mu + \frac{(l - \Theta, B)}{(l - Ø, B)} at$$

The ARIMA Model constant term can be written as:

$$Ø(B)(Wt - \mu) = \Theta(B)\, at \text{ or}$$

$$Ø(B) Wt = \text{const} + \Theta(B)\, at$$

Where const = $Ø(B)\mu = Ø1\mu - Ø2\mu - \ldots - Øp\mu$

In the model, the constant term can be $Ø(B)\mu$, provided an autoregressive operator and a mean term are included in the model.

There is the general ARIMA model with input series, and it is referred to as the ARIMAX model, written as:

Wt = μ+Σωi (B) BkiX i,i+Θ (B)at

iSi (B) Ø (B)

Where

Xi, i is the **i**th input time series or a difference of the **i**th input series at time t.

Ki is the pure time delay for t effect of the **i**th input series.

ωt (B) is the numerator polynomial of the transfer function for the **i**th input series.

Si (B) is the denominator polynomial of the transfer function for the **i**th input series.

The above model for transfer function can be compressed as:

Wt = μ+Σωi (B) Xi,I+ nI

i

Where ωi (B) is the transfer function for the **i**th input series modeled as a ration of the ω andδpolynomials:

ki

ωi(B) =(ωi (B)/δi (B)) B

Where ntis the noise series: nt = (Ø (B)/Θ (B)) at.

The other names for response series are the dependent series or output series. And other names for input series are independent series or predictor series. The response series is made up of past values of the random shocks and past values of other input series.

ARIMA models can be expressed in factored form i.e. Ø, Θ, ω, or δ polynomials are put as a products simpler polynomials, thus, ARIMA in factored form is:

Wt = μ+ Θ1 (B) Θ2 (B)at

Ø1 (B) Ø2 (B)

Where Ø1 (B) Ø2 (B) = Ø (B) and Θ1 (B) Θ2 (B) = Θ (B).

In factored form, the order of ARIMA can be expressed as the product of two factors, thus, we have:

ARIMA (p,d,q) X (P,D,Q)s.

ARIMA (p,d,q) X (P,D,Q)s is the general notation for the order of a seasonal ARIMA model with both seasonal and non-seasonal factors. This term (p,d,q) is for the non-seasonal part and (P,D,Q)s is for the seasonal put.

The S in the order of the seasonal part is the number of observations in a seasonal cycle: S can be 12 for monthly series, 4 for quarterly series and 7 for daily series.

Examples:

An ARIMA (0,1,2) X (0,1,1)4is a notation that describes a seasonal ARIMA model for quarterly datawhich can be expressed in mathematicalform as:

4 4

$(1-B)(1-B)Yt = \mu + (1-\Theta1,iB-\Theta1,2B^2)(1- \Theta2,1B)at.$

An ARIMA (0,1,2) X (0,1,1)12 is a notation that describes a seasonal ARIMA model for monthly data expressed mathematical as:

12 12 12

$(1-B)(1-B)Yt = \mu + (1-\Theta1,iB-\Theta1,2B)(1- \Theta2,1B)at.$

In general, the class of models that can be used in theory to forecast a time series which can be stationary by transformations such as differencing and logging is the ARIMA model – ARIMA (p,d,q).

We have special cases of ARIMA models, and these are: (1) Random-walk and random-trend models, (2) autoregressive models, and (3) exponential smoothing models.

ARIMA non-seasonal models are:

- Random walk i.e. ARIMA (0,1,0).
- Differenced first-order autoregressive model i.e. ARIMA (1,1,0)
- Simple exponential smoothing i.e. ARIMA (0,1,1) is without constant.

- Simple exponential smoothing (growth) i.e. ARIMA (0,1,1) is with constant.
- Linear exponential smoothing i.e. ARIMA (0,2,1) or (0,2,2) is without constant.
- "Mixed" Model – ARIMA (1,1,1) has a prediction equation of

$\hat{Y}(t) = \mu + Y(t-1) + \text{Ø}(Y(t-1)-Y(t-2)-\Theta e(t-1)$

In real life, we try to avoid over fitting of data and non-uniqueness of the co-efficient; thus, we are tied to "unmixed" models with either only autoregressive, or only moving averages.

ARIMA model which is a generalisation of ARIMA model is fitted to time series data so as to understand the data or to predict future points in the series. This model is represented as ARIMA (p,d,q), where p, d, and q are integers greater than or equal to zero. Thus, it is the integration of ARIMA (p,q) process that gives birth to ARIMA (p,d,q) process. Interestingly, it is d that controls the level of differencing, as such if d = 0ARIMA model is equal to ARMA model. It is only the autoregressive side that is always differenced i.e. AR side; such the moving average (MA) side is always I (0).

In ARIMA (p,d,q), we have the (Box-Jenkins), with p, as the order of the AR part of the model, d is the degree of preliminary differencing that is required, and q, the order of the MA part of the model.

In most cases, differencing is needed before applying the ARMA process; as such the degree of differencing will be known. Most time series processes, are viewed from ARIMA standpoint, as ARIMA models encompass AR processes, MAmodels and integration. All data may have AR component, and it may have degree of integration, of the order one, zero, or two; and it may have MA component.

Time series can be divided or decompose into three components, which enable one to determine the structure of the process that is being modeled. In all analysis of time series, its first stage is the differencing, which will help to produce a stationary series. And secondly, ARMA will then be attempted to fit into the stationary series. For example, if we have $Yt= \mu + \varepsilon t$, it is a random process. In the linear equation, Yis only dependent on the mean of the series and the error term.

The above random process is ARIMA (0,0,0), because there is no AR process, no integration, and no past error terms. That is, Y is not relying on past values of Yt, we do not have differencing of Yt and there is no dependence on past error terms. An ARIMA (0,0,0) process with a zero mean is called a WHITE NOISE PROCESS.

Considering $Yt = \alpha Yt\text{-}1 + \varepsilon t$,we see that ARIMA is written as ARIMA (1,0,0); indicating that the process depends on the immediate past values of Y with no differencing. Thus, ARIMA (1,0,0) is the same as AR (1) process.

Given that $-1 < \alpha < 1$, the $\varepsilon t$is a white – noise process.

Considering ARIMA (0,1,0) process, we can see that is a differencing or differencing would be required; and if $\alpha$ is 1 in the above, then there would be no stationary, and we would have,

$$Yt = Yt\text{-}1 + \varepsilon t$$

And an ARIMA (0,0,1) is given by

$$Yt = \varepsilon t + \Theta \varepsilon t\text{-}l$$

Since Ytis influence only by error terms. While, ARIMA (1,0,1) process would be

$$Yt = \alpha Yt\text{-}1 + \beta \varepsilon t\text{-}1 + \varepsilon t$$

In real life, in different applications, we are faced with more than one variable, and this will lead us to the idea of vector AR and vector MA processes. We may face situations like this:

$$\hat{X}t = \alpha Yt\text{-}1 + \beta Yt\text{-}1$$

$$\hat{Y}t = \gamma X\, t\text{-}1 + \delta Yt\text{-}1$$

With Xt and Yt forming a vector, the above equation can be put in matrix form:

$$\begin{bmatrix} \hat{X}t \\ \hat{Y}t \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} Xt\text{-}1 \\ Yt\text{-}1 \end{bmatrix}$$

From the above, we have a 2 X 2 matrix of $\alpha$, $\beta$,$\gamma$ and $\delta$. There is only one lag of vector variable involved, as such the process is a vector AR (1) or VAR

(1) and there is no MA processes.

Below is a process with lags in the error terms but no lags in the underlying variable, the process is VMA (1) process. We have,

$$\hat{X}t = P\varepsilon xt\text{-}1 + q\varepsilon Yt\text{-}1$$

$$\hat{Y}t = \gamma\varepsilon xt\text{-}1 + S\varepsilon Yt\text{-}1$$

In matrix form, we have,

$$\begin{bmatrix} \hat{X}t \\ \hat{Y}t \end{bmatrix} = \begin{bmatrix} p & q \\ r & S \end{bmatrix} \begin{bmatrix} \varepsilon Xt\text{-}1 \\ \varepsilon Yt\text{-}1 \end{bmatrix}$$

Combining AR and MA processes, thus, giving one lag of levels and one lag of residuals, i.e. VARMA (1,1), we have:

$$\hat{X}t = \alpha Yt\text{-}1 + \beta Yt\text{-}1 \quad + \ P\varepsilon xt\text{-}1 + q\varepsilon Yt\text{-}1$$

$$\hat{Y}t = \gamma X\, t\text{-}1 + \delta Yt\text{-}1 + \gamma\varepsilon xt\text{-}1 + S\varepsilon Yt\text{-}1$$

$$\begin{bmatrix} \hat{X}t \\ \hat{Y}t \end{bmatrix} = \begin{matrix} \\ r\ \ S \end{matrix} \begin{bmatrix} p & q \\ \varepsilon Yt\text{-}1 & \end{bmatrix} \begin{bmatrix} \varepsilon Xt\text{-}1 \\ r\ \ S \end{bmatrix} + \begin{bmatrix} p\ \ q \\ \varepsilon Yt\text{-}1 \end{bmatrix} \begin{bmatrix} \varepsilon Xt\text{-}1 \\ \ \end{bmatrix} \begin{bmatrix} \ \\ \ \end{bmatrix}$$

ARIMA processes are made up of three parts, the AR processes, the degree of integration of the series and the MA processes. In ARMA, time series data are assumed to be stationary in their mean and variance.

## PAPER OBJECTIVESS

The primary objectives of this paper are to contribute to knowledge and to

aid businesses in forecasting, predicting or planning in Nigeria. Its specific objectives are to forecast key variables in all business and to show the use of ARIMA in forecasting. The application of ARIMA models to make real-time predictions as it relates to business variables.

## SIGNIFICANCE OF PAPER

This paper is written to do essentially two things:

(1) to contribute to knowledge by evaluating ARIMA (p,d,q) as it concerns time series data of business.

(2) to assist individuals, businesses and government in forecasting or predicting or planning with their past data gathered.

## ARIMA PROCEDURE: METHODOLOGY

This section deals with a review of what the net have done in the area of autoregressive integrated moving average and its analyses and forecasts in the business, financial and more especially in the Nigerian Stock Exchange Market. It is also to examine the literature with respect to the User's Guide SAS 9.2 documentation. This section is a theoretical framework or operational framework and empirical reviews. Emphasis will be placed on ARIMA methodology or procedure, with a view to relating to organisational forecast or prediction within the Nigerian Stock Exchange Markets.

## ARIMA PROCEDURE: AN OVERVIEW

Arima procedure uses autoregressive integrated moving-average (ARIMA) or autoregressive moving-average (ARMA) for analyses and forecasts of equally spaced univariate time series data, transfer function data, and intervention data. This model predicts a value in a response time series in the form of lines combination of its own past values, shocks or innovations i.e. past errors and recent and former values of other time series.

Box and Jenkins where the first to popularise the ARIMA approach, and as such ARIMA models is sometimes called Box-Jenkins models. Box and Tiao (1975) where the first to discussed the ARIMA procedure use of the general transfer function model. ARIMA model do include other time series as input variables, where other variables are included in ARIMA model, the model is then known to be an ARIMAX model. The other name for ARIMAX model is DYNAMIC REGRESSION (Pankratz, 1991).

ARIMA procedure did not only make available a detail set of tools for univariate time series model identification, parameter estimation, and forecasting, but it also offers great flexibility in analyses of ARIMA or ARIMAX models. This procedure supports seasonal, subset and factored ARIMA models; intervention or interrupted time series models; multiple regression analysis with ARMA errors; and rational transfer function models of any state of being connected parts.

Box-Jenkins strategy for time series modeling has the features of (1) identification, (2) estimation and diagnostic checking, and (3) forecasting steps. And the PROC ARIMA design meets the above Box-Jenkins strategy for time series modeling. The triple strategies of Box-Jenkins should be known to enable one use the PROC ARIMA, and one need care and judgment before using ARIMA procedure and to use them correctly requires some expertise, because ARIMA class of time series model is complex and powerful.

**ARIMA MODELING: BOX-JENKINS 3-STAGES**

In 1976, Box and Jenkins described three stages of ARIMA modeling and
the analysis performed by PROC ARIMA is divided into these three stages, and these are:

- the identification stage,
- the estimation and diagnostic checking stage, and
- the forecasting stage.

In the identification stage, use is made of the IDENTIFY statement to
specify the response series and identify candidate ARIMA models for it. Time series are normally used in later statements, as such it is this identifies statement that reads it, and it not only reads it, it possibly computes autocorrelations, inverse autocorrelations, partial autocorrelations, and cross-correlations; and differencing the time series. To know if differencing is necessary, stationarity tests will be carried out. Identify statement out put analysis always suggests for one or more ARIMA models that could be fit. This gives room for options, which enable one to test for stationarity and tentative ARMA order identification.

In the second stage, the estimation and diagnostic checking stage, use is made of the ESTIMATE statement to specify the ARIMA model that would fit into the variable specified in the previous IDENTIFY statement and to also estimate the model parameters. The statement in the stage produces diagnostic statistics that will help one to judge whether or not the model is adequate. Series of statistical test are carried out to test parameter estimate, to compare with other models and to know if residual contains information that can be used for a more complex model. These statistical tests could range from significance tests for parameter estimates which will indicate whether some terms are unnecessary in the model; the Goodness-of-fit statistics to enable comparism with other models; and the white noise residual tests which will indicate whether the residual series have additional information that can be used in other complex model. There is the OUTLIER statement which provides another useful tool to check it the currently estimated models accounts for all the variation in the series. Diagnostic tests is used to find out if there are problems with the model, using this test, if these are problems with model then you try other model, and of course, repeat the estimation and diagnostic checking stage.

In the third stage, that is, the forecasting stage, use is made of the FORECAST statement to forecast future values of the time series and also to generate confidence intervals for these forecasts from the ARIMA model produced.

**ARIMA PROCEDURE STATEMENT: IT'S USE.**

The ARIMA procedure statements are hierachically related i.e. the three statement are related in a hierarchy – the IDENTIFY, ESTIMATE, and FORECAST statements. It is done this way, IDENTIFY statement produces a time series to be modeled; and several ESTIMATE statement can follow to estimate different ARIMA models for the series, and finally, several FORECAST can be used. So it goes this way, a FORECAST statement must be preceded at some point by an ESTIMATE statement, and so also an ESTIMATE statement must precede an IDENTIFY statement. Use can be made of additional IDEBTIFY statements so as to switch to modeling a different response series or to change the degree of differencing used.

Ordinarily PROC ARIMA can be reinvoked, if procedure statements are executed too many times but with ARIMA procedure it cannot be reinvoked; since ARIMA procedure can be used interactively and can be executed any number of times. ARIMA procedure statements can be carried out singly or in groups by a single statement or group of statements with a RUN statement. When a RUN statement is entered the output for each statement or group statements is produced. PROC ARIMA step can not be terminated by RUN statement but it only tells the procedure to execute the statements given so far. PROC ARIMA can be ended by submitting a QUIT statement, a DATA step, another PROC step, or an ENDAS statement. The example below demonstrates the Interactive use of ARIMA procedure statements. A complete PROC ARIMA program is:

```
Proc arima data = test;
            Identify rar = sales nlag = 24;
run;
            identity rar = sales (1);
            run;
estimate P=1;
            run;
            estimate P = 1. q = 1;
            run;
            Outlier;
            run;
forecast lead = 12 interval = month id = data out = result;
            run;
            quit;
```

**IDENTIFICATION STAGE**

A real sales organisation will have a SALES variable that need to be forecast. An ARIMA (1,1,1) model uses a simulated data set TEST.

The IDENTIFY statement is used first to specify the input data set in PROC ARIMA statement. Later, an IDENTIFY statement is used to read in the SALES series and its correlation properties is analyse. Firstly, IDENTIFY statement will print description statistics for the SALES series. Next following the IDENTIFY statement produces a panel of plots used for its autocorrelation and trend analysis. The panel is made up of the quad plots:

- the time series plot of the series

- the sample autocorrelation function plot (ACF)
- the sample inverse autocorrelation function plot (IACF)
- the sample partial autocorrelation function plot (PACF)

The autocorrelation function plots indicate the degree of correlation with past values of the series as a function of the number of periods in the past (that is, the lag) at which computation is made for the correlation. The IDENTIFY statement statistical test of the hypothesis that none of the autocorrelations of the series up to a given lag are significantly different from 0. If a series non-stationary, we need to transform it to a stationary one by differencing. Thus, we model the change in SALES from one period to another and not modeling the SALES series itself. Use in made of another IDENTIFY statement to difference the SALES series. It will be seen that the autocorrelation decreases rapidly, meaning the change in SALES is a stationary time series. The next level in ARIMA methodology is to study the patterns in the autocorrelation so as to choose the candidate ARMA models to the series. Partial and inverse autocorrelation functions are also use to identify the appropriate ARMA models for the series.

The heart in the identification stage in ARIMA methodology is the matching of theoretical autocorrelation functions of different ARMA models to the sample autocorrelation functions computed from the response series. And in the ARIMA approach to ARIMA modeling, the sample autocorrelation function, inverse autocorrelation function, and partial autocorrelation functions are compared with the theoretical correlation functions expected from different kinds of ARMA models. An autocorrelation model of (AR (1), might be a good model to fit the process.

## ESTIMATION AND DIAGNOSTIC CHECKING STAGE

In the identification stage, the autocorrelation for the series suggest an AR (1) model for the change in SALES. And a check on the diagnostic statistics should be made to see if AR (1) model is okay. We still have other models like MA (1) model and low-order mixed ARMA models. In estimating an AR model, one need to specify the order of the autoregressive model with the P = option in an ESTIMATE statement. This ESTIMATE statement does fits the model to the data and parameter estimates are printed as well to the data and diagnostic statistics how well the model fits the data. And next in this stage, is the use of a table of goodness-of-fit statistics, which helps to compare this model to other models. We have a "Constant Estimate" which is a function of the mean (μ) and the autoregressive (AR) parameters. And this estimate is only computed for AR or ARMA models, and strictly for MA models. The "Variance Estimate" estimates the innovation variance, and it is the variance of the residual series. "Standard Error Estimate" is the square root of the variance estimate. And a table of correlations of parameter estimates is printed by the ESTIMATE statement. On the next again, the ESTIMATE statement output is a check of the autocorrelations of the residuals. And this output has the same form as the autocorrelation check for white noise which the IDENTIFY statement prints for the response series. The chi-square ($X^2$) test statistics for the residuals series shows whether the residuals are uncorrelated (white noise) or have additional information that can be put to use in a more complex model. Finally, the ESTIMATE statement output gives a listing of the estimated model, using the backshift notation.

In estimating an ARMA (1,1) model, the previous ESTIMATE statement check of residuals shows that an AR (1) model is not sufficient. Thus, we use ARMA (1,1) model for the change in SALES. In the ARMA (1,1) model, the prediction is that the change in SALES is an average change, plus some fraction of the previous change, plus a random error, plus some fraction of the random error in the preceding period. Thus, an ARMA (1,1) model for the change in SALES is the same as an ARIMA (1,1,1) model for the level of SALES.

**FORECASTING STAGE**

In the estimation stage, ESTIMATE statement is used to decide the best
model, after which we use the FORECAST statement to forecast. If the ESTIMATE statement shows that the last model is not the best, then there is need to repeat it, to enable us get the best model, then, we can forecast by using the FORECAST statement.

A FORECAST statement like this:

forecast lead = 12 interval = month id = date out = results; run;

can be use for the SALES series that is monthly, and attempting to forecast a year ahead from the most recent available SALES figure; given the dates for the observations by a variable DATE in the input data set TEST. The FORECAST statement can print and plots forecast values.

**INTERRUPTED TIME SERIES ARIMA**

Interrupted time series ARIMA also known or called as intervention model
is a special kind of ARIMA model with input series. In this model, the input series is an indicator variable that contains discrete values that flag the occurrence of an event affecting the response series. Event like this is an intervention on or an interruption of the normal evolution of the respondtime series; where there is no interruption or intervention we4 talked about a pure ARIMA process. This model can be used to model and forecast the response series and to also analyse its impact. Ifwe are estimating the effect of the intervention, then the process is known as intervention analysis or interrupted time series analysis.

The question is, can outside event affect subsequent observations? Will the introduction of a new economic policy improve economic performance; can a new anti-crime law after subsequent crime rates? Generally, we would like to evaluate the impact of one or more discrete events on the values in the time series. We have three major types of impacts that are possible:

- permanent abrupt,
- permanent gradual, and
- abrupt temporary.

Impulse interventions can be found in a one-time event. E.g. a short-
term advertisement on the sales of a product. In this case, the value of 1 is the input variable for the period and the value 0 for other periods with no advertisement. This kind of intervention variables are called impulse or pulse functions.

We have other interventions like the continuing interventions, in this case, the input variable flags period before and after the intervention. E.g. the study of change in tax rates on some economic measure. Here the intervention variables are called or known as step functions.

With respect to interaction effects, we can multiply input variables and include the product variables as another input. Thus, you can use PROC ARIMA to fit any linear model together with an ARMA model for the error process by using input variables that corresponds to columns of the design matrix of a model that is linear.

**SUMMARY, RECOMMENDATIONS AND CONCLUSION**

**SUMMARY:**

In this paper ARIMA or ARIMA Model is the acronym for Autoregressive Integrated Moving Average (Model). It is a powerful mathematical or statistical tool used by individual, business and government in forecasting or predicting business variables say SALES. ARIMA was introduced by Box and Jenkins in 1976, and at times it is referred to as Box-Jenkins Model.

And its assumptions ranges from absence of outliers to model specification that relies on an examination of (ACF) and (PACF) – autocorrelation function and partial autocorrelation function.

ARIMA model do include other variables, and as such it is called an ARIMAX model; which is also referred to as DYNAMIC REGRESSION. The paper also indicates that ARIMA have three stages: (1) the identification stage, (2) the estimation stage, and (3) the forecasting stage. And these three stages have statements that are hierarchical.

**RECOMMENDATIONS:**

ARIMA model is a very complex mathematical or statistical technique used in forecasting or predicting business variables. Because ARIMA have been a popular method of forecasting and well-developed mathematical structure, thus making it possible to calculate various model features such as prediction intervals. And it allows forecast uncertainty to be quantified; the papers suggest that individual, business and government, should use the models to forecast or predict or plan their business variables. And this will enhance sales and reduce purchases thus, creating more profit for the entity.

This paper strongly recommends the use of this ARIMA model to forecast or predict business variables; as such there should proper education given to the individuals, business and government. There should be public enlightenment on the use of ARIMA model instead of other statistical tools that are not well – developed. Sound education of ARIMA should be given to the owners of businesses, so that they can practicalize the use and see the differences, and automation of ARIMA model should not be out of place. If proper education is given to the owners of businesses and automation is in place, this will lead to more profit or gain and per capital income will move up.

## CONCLUSION:

This paper demonstrates how ARIMA model in time series data are useful
to study and forecast of business variables e.g. sales. This paper demonstrates also how the ARIMA model in time series data can be used to construct a model of forecasting. This paper shows that ARIMA model is best for sales forecasting. This paper is not without assumptions. It shows that ARIMA model can be used in all businesses, in planning sales. ARIMA time series analysis uses lags and shifts in the past data to uncover patterns (e.g. moving averages, seasonality) and predict the future. ARIMA can be more complex to use than other statistical forecasting techniques; though if properly used is quite good.

## REFERENCES


- About.com Economics: http://economics.about.com/cs/economicsglossary/g/ar.htm (2/25/2009)
- American Psychological Association (APA): autoregressive (AR) process. BusinessDictionary.com.Retrieved February 25, 2009, from BusinessDictionary.com website: http://www.businessdictionary.com/definition/autoregressive-AR-process.html
- Box, P. and Jenkins, G.M., (1976). Time series Analysis: Forecasting andControl.Holden-day Inc., San FranciscoCA.
- Box, P. and Jenkins, G.M., Reinsel, G.C.(1994).Time series Analysis: Forecasting andcontrol. 3.E. Prentice-Hall.
- Box-Jenkins Models Definition(1994)http://www.businessdictionary.com/ Definition/Box-Jenkins-models.html
- Brockwell, P.J. and Davis, R.A. (1991).Time Series: Theory and Methods, Springer-Verlag, 2nded.
- Brockwell, P.J. and Davis, R.A. (1996) Introduction to Time Series and Forcasting.NewYork: Springer-Verlag.

- BusinessDictionaryDefinition(2009), ARIMA.http://www.businessdictionary.com/definition/autoregressive-integrated-moving-average.
- Chatfield, C. (1996). The Analysis of Time Series: An Introduction (5th edition).London: Chapman and Hall.
- Chicago Manual Style (CMS):Autoregressive (AR) process. BusinessDictionary.com. WebFinance.Inc. http://www.businessdictionary.com/definition/autoregressive-AR-process.html (accessed: February 25, 2009).
- Cochran A., Chamlin D., and Seth E. (1994).Deterrence or Brutalization? Criminology,P. 32, PP. 107-134.
- Enders, W. (2004).Applied Econometric Time Series, Wiley,2nd ed.
- Granger, C.W.J. and Paul, N. (1986). Forecasting Economic Time Series.Orlando: Academic Press.
- Harvey, A.C. (1989).Forecasting Structural Time Series ModelsAnd the KalmanFilter.Cambridge Press.
- Hamilton, J.O. (1994).Time Series Analysis. Princeton: PrincetonUniversity Press.Hodgson, B. (2006). http://en.wikipedia.org/wiki/Autoregressive
- Moving average model.Introduction to ARIMA Models: (2009) http://www.duke.ed/rnau/411 arim.htm
- Makridakis, S., Wheelwright, S.C. and Forecasting: Methods and applications, Hyndman, R.J. (1998) NewYork: John Wiley and Sons.
- McCleary, R. and Hay, R.A. Jr. (1980). Applied Time Series Analysis for the SocialSociences.Beverly Hills, Ca: Sage.
- Mills, T.C. (1990).Time Series Techniques for Economists.CambridgeUniversity Press.Modern Language Association (MLA):"autoregreeive (AR) process" BusinessDictionary.com. WebFinance, Inc. February 25, 2009 http://www.businessdictionary.com/definition/autoregressive-AR-process.html.
- NationMaster.comencyclopedia(2005).ARIMA http://www.nationMaster.com/encyclopedia/Autoregressive-integrated-moving-average.
- Pandit, S.M. and WU, S. (1983). Time Series and System Analysis with Applications.John Wiley & Sons, Inc.

- Pankratz, A. (1983).Forecasting with Univariate Box-JenkinsModels: Concepts and Cases, NewYork:
- John Wiley & Sons.Time Series Analysis (2008).Electronic Textbook Statsoft. STATISTICA Statsoft.Inchttp://www.Statsoft.com/textbook/Sttimser.
- Venables, W.N. and Ripley, B.D. (2002).Modern Applied Statistics with S. Springer-Verlag,4th ed.VisualWikipedia(2008).ARIMA-Visual Wikipediahttp://visualwikipedia.com/en/Autoregressive-integrated_moving-average.
- Watsham, T.J. and Parramore, K.(1997). Quantitative Methods in Finance, The Alden Group, Oxford, Great Britain.Wikipedia Free encyclopedia (2008).ARIMA.http://www.en.wikipedia.org/wiki/Autoregressive integrated-moving-average: