Al- Azhar University– Gaza

Deanship of Postgraduate Studies & Scientific Research

Faculty of Economics and Administrative Sciences

Department of Applied Statistics

Identification methods in time series models; the case of the Palestinian banking sector

طرق التعرف على النماذج في السلاسل الزمنية(حالة قطاع البنوك الفلسطيني)

by

Ehab Mahmud Abuzuiter

A Thesis Submitted in Partial Fulfillment of Requirements for the Degree of M.Sc. of Applied Statistics

Under the Supervision

Of

Dr. Abdalla El-Habil

Associate Professor of Statistics

October 2012

بسم الله الرحمن الرحيم

نرفع درجات من نشاء وفوق كل

ذي علم عليم

صدق الله العظيم

سورة يوسف(٧٦)

Al- Azhar University – Gaza

Deanship of Graduate Studies

Identification methods in time series models; the case of the Palestinian banking sector

A Thesis submitted in partial fulfillment of requirements for the degree of Master of Applied Statistic

by

Ehab M. Abuzuiter

<u>Committee of Evaluation</u>

Title	Signature
1- Dr. Abdalla M. El-Habil Supervisor	……………………………
2- Dr. Mahmoud Okasha Internal Examiner	…..…………………………
3-Dr. Hazem E. Sheikh Ahmed External Examiner	….…………………………….

Faculty of Economics and Administrative Sciences

Department of Applied Statistics

Gaza – Palestine 2012

إهداء

إلى أول مـن نطـق بهـما لســاني والدتي ووالدي ... إلى زوجـتي و أبنـائي الذيـن صـبروا علـى وشـاركوني التعـب و العنـاء ... إلى إخـواني و أخـواتي جميعـاً... إلى أسـاتذتي المـوقرين الذين لم يبخلوا علي بالعطاء.... إلى كل أحبائي و أصدقائي......................

إلى كل شخص ساعدني في انجاز هذا العمل.

إلى كل هؤلاء اهدي هذا الجهد المتواضع،،،،،،

ACKNOWLEDGMENT

Firstly, I would like to express my deepest sense of gratitude to my supervisor Dr. ABDALLA EL- HABIL for his guidance, encouragement, understanding, and excellent advice throughout this study. I am grateful to Dr. Mahmoud Okasha, and Dr. Hazem E. Sheikh Ahmed for their effort, notes, advice and support that they willingly gave to help me. I am also thankful to all my colleagues and friends at Al-Azhar University, especially from the field of Applied Statistics for their help and support, with whom I shared pleasant times. I am deeply and forever indebted to the people in my life who touched my heart and gave me strength to move forward to something better. The people who inspire me to breathe, who encourage me to understand who I am, and who believe in me when no one else does. To my wife, two son and my two daughters. Last but not least, I would like to dedicate this work to my loving mother, father and my brothers .

Thank you all

Abstract

Model identification method is the important step in time series modeling, Box-Jenkins identification method depends on examining the patterns of the sample autocorrelation function (SACF) and partial sample autocorrelation function (PSACF), In some cases, the behavior of SACF and PACF are similar which make the identification much more difficult. In our thesis, we studied three methods [ESACF, SCAN and MINIC] of identification, and concentrated on their performance for order selection of mixed ARIMA model. The order selection of ARIMA models by using the three methods were ARIMA(3.1.0), ARIMA(1.1.0) and ARIMA(2.1.1), ARIMA(1.1.1) from MINIC, SCAN and ESACF, respectively. An application on the three methods using real data is conducted. Also their performance for order selection of mixed ARIMA model are compared by AIC, BIC, ME, RMSE, MAE and MPA criteria. We found that SCAN method gives the best order selection for the ARIMA model.

ملخص

من اهم مراحل بناء النموذج في السلاسل الزمنية هي مرحلة التعرف على رتبة النموذج.

طريقة بوكس جنكنز تختبر شكل دالة الارتباط الذاتي والارتباط الذاتي الجزئي ولكن هذه الطريقة تواجه صعوبات عندما تتشابه شكل دالة الارتباط ودالة الارتباط الجزئي وتحتاج لمزيد من الخبرة وخصوصا في النماذج المختلطة .

هذه الدراسة تهتم بطرق التعرف على نماذج التنبؤ في السلاسل الزمنية ، قمنا بالتعرف على النماذج من خلال الطرق الكلاسيكية (بوكس جنكنز) والتي لم تكن حاسمة بوضوح لتحديد رتبة النموذج ، وبعد ذلك قمنا بتحديد رتبة النماذج بثلاث طرق هي ESACF,SCAN, MINIC وقد قمنا بمقارنة كفاءة هذه الطرق في اختيار رتبة النموذج المقترح من خلالها للسلسلة الزمنية محل الدراسة والتي كانت عن تنبؤ سعر الافتتاح للسهم في قطاع البنوك الفلسطيني، وقد كانت رتب النماذج المقترحة للسلسلة محل الدراسة باستخدام هذه الطرق كالآتي ARIMA(3.1.0) من طريقة MINIC ، ARIMA(1.1.0) من طريقة SCAN و ARIMA(2.1.1) و ARIMA(1.1.1) من طريقة ESACF . وتم مقارنة الطرق من خلال النماذج التي رشحت عن كل طريقة وكذلك من خلال تقييم التنبؤ لكل نموج من خلال مقارنة Loss function و AIC و BIC .

في دراستنا وبناءا على معايير كفاءة التنبؤ كان النموذج المقترح من قبل طريقة SCAN ARIMA(1.1.0) هو الأفضل في حالة البيانات المتوفرة لدينا .

Abbreviations

.

Abbreviation	Full Word
ACF	Autocorrelation Function.
AIC	Akaike Information Criterion.
AR(p)	Autoregressive Model of order p.
ARIMA(p.d.q)	Autoregressive Integrated Moving Average Model order (p.q).
ARMA(p.q)	Autoregressive Moving Average Model order (p.q) .
BIC	Bayesian Information Criterion.
BLUE	Best Linear Unbiased Estimator.
ESACF	Extended Sample Autocorrelation Function.
HQ	Hannan Quinn Criterion.
IACF	Inverse Autocorrelation Function.
MA(q)	Moving Average Model of order q.
MAE	Mean Absolute Error.
MAPE	Mean Absolute Percentage Error.
ME	Mean Error.
MINIC	Minimum information Criterion
ML	Maximum likelihood method
MPE	Mean Percentage Error.
MSE	Mean Squared Error.
OLS	Ordinary Least Squares
PACF	Partial Autocorrelation Function.
RMSE	The Root Mean Squared Error.
SCAN	Smallest Canonical correlation method
WN	White Noise.

Identification methods in time series models; the case of the Palestinian banking sector

Table of Contents

List of tables

List of figures

List of output

Chapter 1

Introduction

1.1 Overview

In this chapter, we first motivate the thesis by explaining why we believe model identification is relevant and important. Then we present the study objectives in order to clearly define the research problem. Finally, we outline the plan of this thesis in order to review its organization.

1.2 Motivations

Model identification, in which the researcher tries to identify which model is appropriate for the data. Sometime using (Box Jenkins) Identification approach is a very difficult, complicated, and problematic task. In this study we used Box Jenkins approach and an automated methods and make comparisons between them to make the model identification process more reliable and effective.

1.3 The study importance

The importance of the study is based on identifying the best order selection of the models represented in time series, in order to brief the errors such as AIC,BIC.

1.4 Research Problem

Practically model identification using Box-Jenkins approach sometimes confusing the researchers about the order selection of the models. Therefore the research problem is can be formed by the following question:

How can we determine the best method for identification in time series models by using automated methods?

1.5 Objectives

The main objectives of the study is

1. To understand identification methods in time series models within specific criteria.
2. To help researchers to choose the methods of identifying the appropriate model to the available data.
3. To determine the best methods for identification in time series models in our data.

1.6 The study structure

The study tried to determine the best identification methods to select the appropriate order selection of the time series models.

Contents of the study:

The study consists of 5 chapters; Chapter 1 presents the study objectives, outline the plan of this thesis in order to review its organization and presents the literature review. Chapter 2 overview of time series models, discuss a different types of time series models and its properties, describes the building time series models, ARIMA forecasting and forecasting measures. Chapter 3 presents the identification methods and tools that used the identification methods structure. Chapter 4 discusses the findings of this study based on the results of implementing the proposed methods using real data. Chapter 5 contains conclusion and recommendations.

1.7 Literature Review

In this section we will mention some related studies:

Wai-Sum Chan (1999), studied identification methods using the patterns of some functions of the autocorrelations and proposed to supplement the Box Jenkins methods. His paper studies some of these proposed procedures. Their performances for order selection of a mixed ARMA process are compared with an expert system in a simulation study.

Yuhong Yang & Hui Zou(2002) proposed the use of an algorithm AFTER to convexly combine the models for a better performance of prediction. The weights are sequentially updated after each additional observation. Simulations and real data examples are used to compare performance of him approach with model selection methods. The results show advantage of combining by AFTER over selection in term of forecasting accuracy at several settings.

Hamad A. Al-Ghannam(2003), studied the analysis of time series of Stock Price Index yearly in Saudi Arabia, for the period from March 1985 until June 2002, to be recognized on the type of the pointer changes and to build a model that helps to predict the values of the index in the short term. He applied Box_ Jenkins methodology by using some statistical methods to test and examine a model such as the stationary of residual, as well as the application of Akaike & Schwarz and prediction error. He claimed that the best model appropriate for data of general index of share prices is a autoregressive model of the first order without seasonal effects in the model, and the choice was based on several criteria and diagnostic tests of several models of close .

Chen and Yang in(2004), studied the issue of evaluating forecast accuracy measures. In the theoretical direction, for comparing two forecasters, only when the errors are stochastically ordered, the ranking of the forecasts is basically independent of the form of chosen measure. They proposed well motivated Kullback– Leibler Divergence based accuracy measures. In the empirical direction, they studied the performance of several familiar accuracy measures and some new ones in two important aspects: in terms of selecting the known– to– be better forecaster and the robustness when subject to random disturbances. In addition, Their study suggested that, for cross–series comparison of forecasts, individually tailored measures may improve the performance of differentiating between good and poor forecasters.

Samreen Abu Radi (2009) aims to analyze the status of shares related to the banking sector in Amman Stock Exchange, through the use of time series analysis, access to an efficient market, through the application of the conditions existing in the market, analysis of the status of the general trend of stock prices in Amman Stock Exchange, through the turnover rate of shares over twelve months for eight years starting from 2000-2007, in order to find the variables affecting performance, identifying the most important components of the time series affecting stock prices in Amman Stock Exchange (seasonal, periodical, and random), in addition to identifying which of these components are responsible for stock price changes. Trying to determine the general tendency of the time series of stock prices for the coming period through the use of the model of basic components. The study came to a number of conclusions such as:

• Results showed that the influence of irregular variables on the turnover rate of shares related to the banking sector, listed in Amman Stock Exchange, was clear, in addition to the impact of changes related to the general trend, as well as, the seasonal and periodical changes.

• Results showed that the size of circulation plays a major role in changing the direction of prices. Thus, in the case of higher prices, increased circulation is desired, while decreased circulation will be the case for low prices.

Based on the above conclusions, the researcher presented a series of suitable recommendations for the use of analysis model of time series in analyzing the status of shares in Amman Stock Exchange.

Shittu. O.I & Asemota. M.J (2009) compared the performance of model order determination criteria in terms of selecting the correct order of an Autoregressive model in small and large samples using the simulation method. The criteria considered are the Akaike information criterion (AIC); Bayesian information criterion (BIC) and the Hannan Quinn criterion (HQ). The results shows that BIC performs best in terms of selecting the correct order of an Autoregressive model for small samples irrespective of the AR structure, HQ criteria can be said to perform best in large sample. Even though the AIC has the least performance among the criteria considered, it appears to be the best in terms of the closeness of the selected order to the true value.

Ojo J.F. & Olatayo T.O. (2009) compared subset autoregressive integrated moving average models, with full autoregressive integrated moving average models. They used residual variance, AIC and BIC, to determine the performance of the models. Results revealed that the residual variance attached to the subset autoregressive integrated moving average models is smaller than the residual variance attached to the full autoregressive integrated moving average models. Subset autoregressive integrated moving average models performed better than the full autoregressive integrated moving average models.

1.8 What feature this study from previous studies

This study agreed with some previous studies in:

1 - Time Series Analysis for the banking sector.

2 - In some comparisons between the methods of determining the order of the model in the time series.

This study characterized all previous studies in:

1 - work comprehensive comparison between the classical methods for determining the order in the time series models such as the Box Jenkins method and automatic methods.

2 - comparisons between type of automated methods .

3 - Action comparisons within type automated methods .

4 - apply the results of the comparisons that have been working on our data, and data pertaining to the analysis of the Palestinian exchange Market (banking sector).

5 - focus on combining most of the methods of identifying the order of time-series models in a way that makes it easier for researchers to know the appropriate methods of data available to them.

Chapter 2

Time Series Models

2.1 Introduction

The goal of this chapter is to give some basic ideas of time series and to show how one can go through the entire process of building the model for a time series data set, and Box-Jenkins Methodology. The time series is a set of numbers that measures the status of some activity over time. It is the historical record of some activity, with measurements taken at equally spaced intervals (daily, monthly, quarterly,...,etc). There are two types of time series data: continuous, where in observations are made at every instant of time and discrete, where in observations are made at (usually) equip spaced intervals. Adiscrete time series can be represented as [x_t: t = 1, 2, ...,N] in which the subscript t indicates the time at which the datum x_t was observed. There are different models in time series including, autoregressive (AR), moving average (MA), autoregressive moving average (ARMA) and autoregressive integrated moving average (ARIMA) . However the most commonly used model is Box-Jenkins ARIMA model that has been successfully applied in economic time series forecasting, as well as being a good tool to develop empirical model dependencies between successive time and failures time. (Box & Jenkins,1994)

The aims of time series analysis are to describe and summarize time series data, fit low-dimensional models, and make forecasts. We can summarize objectives of time series into modeling goals where included of set up a family of probabilistic models to describe the data, estimate the model parameters, check model for goodness of fit and application goals where included provide a compact description of the data, interpretation, prediction (forecasting) and hypothesis testing. Time series data provide useful information about the physical, biological, social or economic systems generating the time series, such as economics, sociology, environment,…, etc.

2.2 Basic concepts

In this section, we will discuss some basic concepts of linear time series analysis.

Stationarity

Box & Jenkins(1994) stated that the basis of time series analysis is stationarity. The time series is said to be stationary if the mean, the variance and the autocovariance (at various lags) does not change regardless of what is the point measure, it fixed over time. Moreover, the time series $\{x_t\}$ is said to be strictly stationary if the joint distribution of x_{t1} ,..., x_{tk} is identical to that of x_{t1-s} ,..., x_{tk-s} for all choices of t_1, t_2 ,…, t_k and all choices of time lag(s), where k is any positive integer and t_1, t_2 ,…, $t_k$is a collection of k positive integers. In other words, strict stationary requires that the joint distribution of x_{t1} ,..., x_{tk} is constant under time shift. A weaker version of stationarity is often assumed a time series $\{x_t\}$ is weakly stationary if both the mean of x_t and the covariance between x_t and x_{t-s} are time-invariant, where s is an arbitrary integer.

More specifically, $\{x_t\}$ is weakly stationary if :

(a) $E(x_t) = \mu$, which is a constant , for all t .

(b) $Cov(x_t, x_{t-s}) = \gamma_s$, which only depends on all time t and lag s.

However in weak stationarity, we suppose that the first two moments of x_t are finite, if x_t is strictly stationary and its first two moments are finite, then x_t is also weakly stationary, from the definitions , but the converse is not true in general.

Dependency and Autocorrelation

In time series analysis, dependence is assessed by calculating the values of the autocorrelations among the data points in the series. In contrast to a correlation coefficient, which is generally used to estimate the relationship between two different variables measured at the same time on multiple subjects, an autocorrelation estimates the relationships within one variable that is measured at regular intervals over time on only one subject. The degree of dependency in a time series is determined by the magnitude of the autocorrelations that can vary between –1.00 and 1.00, with a value of 0.00 indicating no relationship. These values can be interpreted as the strength of relationship between consecutive measurements. The accuracy of estimation improves as the number of observations increases. Generally, 50 or more observations provide reasonably accurate estimates (Ljung & Box, 1978). In practical terms, the degree of dependency indicates the extent to which an observation at any point in time is predictable from one or more preceding observations. The direction of dependency in a time series refers to whether an autocorrelation is positive or negative. The direction can be determined with a high degree of accuracy when there is strong dependency in the data. As the degree of dependency approaches zero, the direction becomes less important. With strong dependency, the direction has clear implications. When the sign of the autocorrelation is negative, a high level for the series on one occasion predicts a lower level for the series on the next occasion. When the sign is positive, a high level of the series on one occasion predicts a higher level on the next occasion. In calculating an autocorrelation, the data points of the series are paired off in a lagged manner against each other.

Autocorrelation Function(ACF)

The most important tool for study dependence is the sample autocorrelation function(SACF). The correlation coefficient between any two random variables X, Y, which measures the strength of linear dependence between X, Y, always takes values between -1 and 1. If we assuming stationary, and we want to estimate autocorrelation function ρ_k for a set of lags K = 1,2, ... The simplest way to do this is to compute the sample correlation between the pairs k units apart in time. Note that the concept of correlation expanding in the case of stationary time series to become the autocorrelation function. The correlation coefficient between $Y_t$and Y_{t-k} is called the lag k autocorrelation of Y_t and denoted by the symbol ρ_k, which under the assumption of weak stationary and defined as:

$$\rho_k = \frac{\sum_{t=K+1}^{T}(Y_t-\bar{Y})(Y_{t-k}-\bar{Y})}{\sum_{t=1}^{T}(Y_t-\bar{Y})^2} = \frac{\gamma_k}{\gamma_0} : for\ k = 1,2, ... \quad (2.1)$$

where $\gamma_k = cov\ (Y_t, Y_{t=k})$

Since ρ_k is a correlation coefficient , it has the simple properties:

a) -1 ≤ρ_k ≤ 1

b) $\rho_k = \rho_{-k}$

c) $\rho_0 = 1$

Partial Autocorrelation Function (PACF)

The correlation coefficient between two random variables Y_t and Y_{t-k} after removing the impact of the intervening Y_{t-1} , Y_{t-2} ,..., Y_{t-k+1} is called (PACF) at lag k and denoted by $\emptyset_{kk}$ One of the methods of calculation are based on the account partial regression coefficient $\emptyset_{kk}$ in the equation:

$$Y_t = \emptyset_{k1}Y_{t-1} + \emptyset_{k2}Y_{t-2} + \cdots + \emptyset_{kk}Y_{t-k} + \alpha_t \qquad (2.2)$$

$\emptyset_{kk}$ calculated from the relationship:

$$\emptyset_{00} = 1\ ;\ \emptyset_{11} = \rho_1\ ;\ \emptyset_{kk} = \frac{\rho_{k-}\sum_{j=1}^{k-1}\emptyset_{k-j}\rho_{k-j}}{1-\sum_{j=1}^{k-1}\emptyset_{k-1}\rho_j}\ , k = 2,3 \ldots \qquad (2.3)$$

$where\ \emptyset_{kj} = \emptyset_{k-j}\emptyset_j - \emptyset_{kk}\emptyset_{k-1,k-1}$; j=1,2,3…,k-1 is a linear time series model can be tentatively identified by its autocorrelation function (ACF), and partial autocorrelation function (PACF) as follows:

- If ρ_1 is non-zero, this indicates that the serial is first order serially correlated.
- Ifρ_k tails off geometrically with increasing lags, and (PACF) cut off after certain lag it means that the model is autoregressive process.
- If ρ_k cut off after a small number of lags, and (PACF) tails off geometrically with increasing lags, it means that the model is moving-average process. A plot of ρ_k versus lag K is often called a correlogram. (Box & Jenkins,(1994))

The Inverse Autocorrelation Function

Chatfield (2000) defined the inverse autocorrelation function (IACF) as the autocorrelation function associated with the reciprocal of the spectral density of a stationary time series, assuming the model is invertible, this means that if $\{z_t\}$ follows the invertible ARMA model $\emptyset(B)z_t$, = $\theta(B)a_t$ model, then the inverse autocorrelation function of this model is the autocorrelation function (ACF) of the dual model $\theta(B)z_t$= $\emptyset(B)a_t$. Note that if the original model is a pure autoregressive model, then the IACF is an ACF corresponding to a pure moving average model .Thus, the IACF has the cut-off feature and behaves similarly to the partial autocorrelation function. The confidence interval of the IACF depends heavily on the estimation method and the asymptotic variance of the IACF estimate is more complicated than that of the partial autocorrelation function (PACF) estimate, which is $\frac{1}{N}$. Therefore, there is more statistical error when estimating a confidence interval of the IACF than of the PACF.

White Noise (WN)

Cryer and Chan, (2008) said that, a very important case of a stationary process is called white noise process . If a time series { x_t } is a sequence of independently and identically distributed (iid) random variables with finite mean and variance, we shall some times, denote this process as $w_t \overset{iid}{\rightarrow} n(0\,,\sigma_w^2)$. Its importance originates from the fact that many useful processes can be constructed from white noise. If { w_t } is normally distributed with mean zero and variance σ^2 and no serial correlation , then it is said to be Gaussian white noise or more succinctly, $w_t \overset{iid}{\rightarrow} n(0\,,\sigma_w^2)$, we usually assume that the white noise process has mean zero and denote Var (x_t) by σ_w^2 . For a white noise series, all the ACFs are zero. In practice, if all SACFs are close to zero, then the series is a white noise series.

Random walk process

Box & Jenkins,(1994) said that, a random walk process is usually method used in the equity market to describe, for example, the behavior of stock prices or exchange rates. This process continually drifts from any expected value in a specific period of time. In this approach it is not considered any constant value or constant variance over time. Generally we can classify two types of random walk process:

- random walk without a drift given by :

$$y_t = y_{t-1} + w_t \tag{2.4}$$

(i.e., no constant or intercept term)

- random walk with a drift given by :

$$y_t = \alpha + y_{t-1} + w_t \quad (2.5)$$

(a constant term is present):

Where y_0 is a real number denoting the starting value of the process, the constant α is called the drift parameter. For t = 1, 2, . . ., with initial condition y_0 = 0, w_t is white noise, and when α = 0 ,in (2.5) is called a simple random walk, consider that we may rewrite (2.5) as a cumulative sum of white noise variates as:

$$y_t = \alpha_t + \sum_{j=1}^{t} w_j \quad (2.6)$$

The simple random walk process provides a good model (at least to a first approximation) for phenomena as diverse as the movement of common stock price, and the position of small particles suspended in a fluid so called Brownian motion . If a trend in a time series process is completely predictable and not variable, it is said a deterministic trend, whereas if it is not predictable, it is said a stochastic trend.

2.3 General linear processes

Yule (1927), claimed that time series can be represented as a linear combination of a sequence of uncorrelated random variables. Affirmed this, too Wild (1938), which states that every weakly stationary non deterministic stochastic process $\{r_t\}$ can be interpreted as a linear combination (or linear filter) of a sequence of uncorrelated random variables, however the linear filter representation is given by:

$$r_t = e_t + \Psi_1 e_{t-1} + \Psi_2 e_{t-2} + \ldots = \sum_{j=0}^{\infty} \Psi_j e_{t-j} \quad , \Psi_0 = 1 \quad (2.7)$$

for our purposes, it suffices to assume that

$$\sum_{i=1}^{\infty} \Psi_{i^2} < \infty \quad (2.8)$$

note the case where the Ψ 's form an exponentially decaying sequence $\Psi_j = \emptyset^j$ where ϕ is a number strictly between −1 and +1 . Then

$$r_t = e_t + \emptyset_1 e_{t-1} + \emptyset_2 e_{t-2} + \ \dots \qquad (2.9)$$

E(r_t) =0 , $\operatorname{var}(r_t) = \dfrac{\sigma_e^2}{1-\phi^2}$, $\operatorname{cov}(r, r_{t-1}) = \dfrac{\phi\sigma_e^2}{1-\phi^2}$,

$$corr(r_t, r_{t-1}) = \phi \quad , \quad corr(r_t, r_{t-k}) = \phi^k$$

for a general linear process (2.7) :

$$E(r_t) = 0, \gamma_k = cov\ (r_t, r_{t-k}) = \ \sigma_\epsilon^2 \textstyle\sum_{i=0}^{\infty} \psi_i \psi_{i+k} \ \ k \geq 0 \qquad (2.10)$$

(Cryer and Chan, (2008)).

Autoregressive process (AR)

A model written on the form

$$r_t = \ \emptyset_1 r_{t-1} + \emptyset_2 r_{t-2} + \ \dots + \emptyset_p r_{t-p} + \ \varepsilon_t \qquad (2.11)$$

called autoregressive of order p, and abbreviated AR(p), where ε_t is white noise. The current value of the series r_t is a linear combination of the p most recent past values of itself plus an “innovation” term ε_t include all that is new in this series at a time t is not interpreted by the past values. Thus, for every t, we assume that ε_t is independent of the $r_{t-1}, r_{t-2}, r_{t-3}, \dots$. In general, we say that a variable r_t is autoregressive of order p, AR(p), with p ∊ N , if it is a function of its p past values and can be represented as:

$$r_t = \sum_{i=1}^{p}{}_j \ \emptyset_i r_{t-i} + \ \varepsilon_t \qquad (2.12)$$

Box & Jenkins(1994)

Moving Average process (MA)

a series $\{r_t\}$ a moving average of order q and abbreviate the name to MA(q), if Written on the form:

$$r_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_p \varepsilon_{t-p} \qquad (2.13)$$

We consider that in the moving average process , a time series $\{r_t\}$, is obtained by applying the weights $1, -\theta_1, -\theta_2, \dots, -\theta_q$ to the variables $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ and then moving the weights and applying them to $\varepsilon_{t-1}, \varepsilon_{t-2}, \varepsilon_{t-3}, \dots, \varepsilon_{t-q+1}$ to obtain $\{r_{t+1}\}$ and so on. We can write equation (2.13) as The following form:

$$r_t = \varepsilon_t - \sum_{i=1}^{q} \theta_q \varepsilon_{t-i} \qquad (2.14)$$

Box & Jenkins(1994)

Autoregressive Moving Average process (ARMA)

Sometimes we need to reach a model for time series analysis more comprehensive, more effective and has fewer parameters. If we assume that this series is part of the autoregressive and the other part is the moving average, we obtain a quite general time series model. In general, if Suppose that $\{r_t\}$; t =... -1, 0, 1 ... is an equally spaced weakly stationary or covariance stationary, time series. There is a famous model of linear models for time series analysis in the time domain belongs to an autoregressive moving average is expressed in the form:

$$r_t = \emptyset_1 r_{t-1} + \emptyset_2 r_{t-2} + \dots + \emptyset_p r_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_p \varepsilon_{t-q} \qquad (2.15)$$

we say that $\{r_t\}$ is a mixed autoregressive moving average process of orders p and q , respectively ; and known simply as ARMA(p ,q) . For the general ARMA(p ,q) model, we state that ε_t, is independent of $r_{t-1}, r_{t-2}, r_{t-3}, \dots, r_{t-k}$ a stationary solution to Equation

(2.12) exists if and only if all the roots of the AR characteristic equation ϕ (x) = 0 are outside the unit circle . For invariability we have to assume that the roots of θ (x) = 0 are outside the unit circle. Where {ε_t} is a sequence of uncorrelated variables, also referred to as a white noise process, and $(\emptyset_1\emptyset_2 \dots \emptyset_p, \theta_1\theta_2 \dots \theta_q)$ are unknown constants or parameters. The model (2.10) is an ARMA(p,q) model or Box-Jenkins model. And we can expressed this model as:

$$(1 - \emptyset_1 B^1 - \emptyset_2 B^2 - \dots - \emptyset_p B^p) r_t = (1 + \theta_1 B^1 + \theta_2 B^2 + \dots + \theta_q B^q)\varepsilon_t \qquad (2.16)$$

where B is the backshift operator, that is $Bx_t = x_{t-1}$.

$$\emptyset(B) = 1 - \emptyset_1 B^1 - \emptyset_2 B^2 - \dots - \emptyset_p B^p$$

$$\theta\,(B) = 1 + \theta_1 B^1 + \theta_2 B^2 + \dots + \theta_q B^q$$

Any ARMA model can therefore be written as an AR or MA model. In general, an ARMA (p,q) is a combination of an AR(p), equation (2.12) , and MA(q) , equation (2.14), and can be written as:

$$r_t = \sum_{i=1}^{p} \emptyset_i r_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \qquad (2.17)$$

(Box & Jenkins(1994)

Autoregressive Integrated Moving Average process (ARIMA)

Cryer and Chan, (2008) claim that, ARIMA is one of the most traditional methods of nonstationary time series analysis known as the Box-Jenkins methodology, but it is technically known as the ARIMA methodology , this method does not focus on building models of a single equation or simultaneous equation models, but on the analysis of a probabilistic, or stochastic, the properties of economic time series itself under the principle that "let the data speak for themselves." In contrast to the regression models, the ARIMA models allows Y_t is explained from the past, or lagged values of Y itself and stochastic error terms. For these reasons, called these models sometimes theoretical models because they are not derived from any economic theory - and economic theories often this form basis of simultaneous equations models. Time series$\{r_t\}$ is classified as an integrated model Autoregressive moving average if the d^{th} difference:

$$w_t = \nabla^d r_t = (1 - B)^d r_t \tag{2.18}$$

is a stationary ARMA process , Where ∇= 1-B is the difference operator. If $\{w_t\}$ follows the ARMA (p, q) model, so we say that $\{r_t\}$ is an ARIMA (p, d, q) process. For practical purposes, we can take is usually d = 1 or 2 at most . We can write model (2.18) as :

$$\emptyset(B) w_t = \theta_0 + \theta(B)\varepsilon_t \tag{2.19}$$

where ϕ (B) is a stationary autoregressive operator, θ (B) is a stationary moving average operator, and $\{\varepsilon_t\}$ is white noise and θ_0 , is a constant usually referred to as a trend parameter. The model is called “integrated” since r_t, can be thought of as the summation (integration) of the stationary series w_t. The previous mentioned models are built on assumptions that the time series involved are (weak) stationary. But, as is well known to us that many economic time series nonstationary, that is, integrating. If the integration of a time series of order 1, I(1), it's the first differences are I(0), that is, stationary. Similarly, if a time series is I(2), it’s second difference is I(0). In general, if a time series is I(d), after differencing it d times we obtain an I(0) series. The above is clear that if we needed to the difference time series d times to make it stationary and then apply the ARMA (p, q) process, we say that the original time series is ARIMA (p, d, q) process, that is, it is an autoregressive integrated moving average time series. In the case of the pattern of seasonal time series ARIMA model is written as follows:

$$\emptyset(B)\Phi(B)\nabla^{d}\nabla_{s}^{D}r_{t} = \theta(B)\Theta(B)\varepsilon_{t} \tag{2.20}$$

Where : $W_t = \nabla^{d}\nabla_{s}^{D}r_{t}$ is a stationary series.

$\nabla^{d} = (1-B)^{d}$ represents the number of regular differences. $\nabla_{s}^{D} = (1-B^{s})^{D}$ represents the number of seasonal differences required to induce stationarity in r_t. The first step in estimating the ARIMA model is to determine (p,d,q)(P,D,Q), where p denotes the number of autoregressive terms, q denotes the number of moving average terms and d denotes the number of time series must be differenced to induce stationarity. P denotes the number of seasonal autoregressive components, Q denotes the number of seasonal moving average terms and D denotes the number of seasonal differences required to induce stationarity.

2.4 Model building

Box & Jenkins(1994)stated that, time series model building is considered as a three-stage iterative procedure based on, identification, estimation, diagnostic checking,

2.4.1 Model Identification

A preliminary Box-Jenkins analysis with a plot of the initial data should be starting with determining an appropriate model. The input data must be adjusted to form a stationary series, one whose values vary more or less uniformly about a fixed level over time. Apparent trends can be adjusted by having the model apply a technique of "regular differencing," a process of computing the difference between every two successive values, computing a differenced series which has overall trend behavior removed. If a single differencing does not achieve stationarity, it may be repeated, although rarely, if ever, are more than two regular differencing required. Where irregularities in the differenced series continue to be displayed, log or inverse functions can be specified to stabilize the series, such that the remaining residual plot displays values approaching zero and without any pattern. This is the error term, equivalent to pure, white noise. On the other hand, if the initial data series displays neither trend nor seasonality(Pure Random Serirs), and the residual plot shows essentially zero values within a 95% confidence level and these residual values display no pattern, then there is no real-world statistical problem to solve and we go on to other things. With a stationary series in place, a basic model can now be identified. Three basic models exist, AR (autoregressive), MA (moving average) and a combined ARMA in addition to the previously specified RD (regular differencing): These comprise the available tools. When regular differencing is applied, together with AR and MA, they are referred to as ARIMA, with the I indicating "integrated" and referencing the differencing procedure. Initially we can identify the pure model as we summarize in the following table(table 2.1) by using (ACF&PACF)

Model	ACF	PACF
ARIMA(p, 0, 0)	Decays slowly	= 0 after p
ARIMA(0, 0, q)	= 0 after q	Decays slowly
ARIMA(p, 0, q)	Decays slowly	Decays slowly
ARIMA(0, d, 0)	Does not decay	Does not decay

Table 2.1: using ACF & PACF to identify the time series model

2.4.2 Model Estimation

There are many method to estimate parameter in the time series model like maximum likelihood method , last square method, …..,etc, and we will talk a little bit from each.

Maximum Likelihood When p and q are known in ARMA(p, q)model, estimates of the ϕ_is and θ_js can be found when the data being observations from a Gaussian $ARMA$ model. Even if $\{X_t\}$ is not Gaussian the Gaussian likelihood still is a reasonable measure of goodness of fit of the model, so maximizing it is sensible. Also, the asymptotic distribution of maximum likelihood estimators is the same whether the white noise innovations ε_t (and so the process itself) are normal or not.

The Least Squares in Time Series The underlying idea in the fitting of a time series model by least squares, by analogy with regression, is that we should choose parameter values which minimize the sum of squared differences between the observed data and their expected values according to the model.

2.4.3 Model Diagnostic

We will consider two types of diagnostic checks. In the first we fit extra coefficients and test for their significance. In the second we examine the residuals of the fitted model to determine if they are white noise (i.e. uncorrelated). Fitting extra coefficients , suppose we have tentatively identified and estimated an ARMA(p, q) model. Consider the following ARMA(p + q, q + q) model.

$$(1 - a_1 L - \cdots - a_p L^p - \ldots - a_{p+p^*} L^{p+p^*}) X_t = (1 + b_1 L + \cdots + b_q L^q + \ldots + b_{q+q^*} L^{q+q^*}) \varepsilon_t \qquad (2.21)$$

We can calculate a Lagrange Multiplier test of the restrictions

ap+1 = ap+2 = . . . = ap+p = 0

bq+1 = bq+2 = . . . = bq+q = 0

If the hypothesis is accepted we have evidence of the validity of the original model. Tests on residuals of the estimated model, if the model is correctly specified the estimated residuals should behave as white noise (be uncorrelated).

2. 5 Forecasting ARIMA Models

The objectives of building a model for a time series including that to be able to forecast the values for that series in future times. We shall assume that the model is known, including specific values for all the parameters, and do forecasting using a known ARIMA model. Suppose we've observed a series $X_1, X_{2,} X_{3,} \dots \dots, X_t$ up to time t and wish to forecast X_{t+1}. That is, we wish to find a predictor,$\hat{X}_{t+k}$ $\left(\hat{X}_t(k)\right)$ say, of X_{t+k} based on the information in $X_1 X_2 X_3 \dots \dots X_t$. Her we will called t is the forecast origin and k the lag time. We'd like $\hat{X}_{t+k}$ to be close to X_{t+k} in some sense. One interpretation of this is that :

$$E\left[\left(\hat{X}_{t+k} - X_{t+k}\right)^2 / X_t, X_{t-1}, X_{t-2}, \dots\right] \quad (2.22)$$

should be as small as possible. The conditioning here takes account of the fact that we've observed $X_t, X_{t-1}, X_{t-2}, \dots$ and want to build this information. The $\hat{X}_{t+k}$ minimizing (2.22) is called the k-step-ahead minimum mean square error predictor. Recall that for any random variable Z with $Var\ (Z\ ) < \infty$, $E\ (Z\ - a)^2$

is minimized by : $a = E\ (Z\ )$

Hence the minimum mean square error k-step-ahead predictor above is $\hat{X}_{t+k} = X_{t+k} / X_t, X_{t-1}, X_{t-2}, \dots$

2. 5.1 Forecast Accuracy

It is important to evaluate forecast accuracy. The accuracy of forecasts can only be determined by considering how well a model performs on new data that were not used when fitting the model. When choosing models, it is common to use a portion of the available data for testing, and use the rest of the data for fitting the model. Then the testing data can be used to measure how well the model is likely to forecast on new data. The issue

of measuring the accuracy of forecasts from different methods has been the subject of much attention. We summarize some of the approaches here. A more thorough discussion is given by Hyndman & Koehler (2006). In the following discussion, $\hat{y}_t$denotes a forecast of y_t. We only consider the evaluation of point forecasts. There are also methods available for evaluating interval forecasts and density forecasts.
(Corradi & Swanson 2006)

2. 5.2 Measures of Forecast Accuracy

The crucial object in measuring forecast accuracy is the loss function, often restricted to which charts the "loss", "cost" or "disutility" associated with various pairs of forecasts and realizations. In addition to the shape of the loss function, the forecast horizon (k) is also of crucial importance. Rankings of forecast accuracy may be very different across different loss functions and different horizons, (Simonelli, 2009)

In this section we will discuss a few statistical loss functions, because they are used widely. Accuracy measures are usually defined on the forecast errors , or $e_{t+k,t} = X_{t+k} - \hat{X}_{t+k\,,t}$ or percent forecast errors, $P_{t+k,t} + \frac{(X_{t+k} - \hat{X}_{t+k\,,t})}{X_{t+k}}$.

Among a number of possible criteria that could be used, here we talk about some common of this criteria, note that in the next, T is number of periods used in the calculation.

1- **Mean Error**, ME = $\frac{1}{T}\sum_{t-1}^{T} e_{1+k,t}$, is the average of all the errors of forecast for a group of data. The ME can be very misleading. AME value of zero can mean that the method forecasted the actual values perfectly (unlikely) or that the positive and negative errors cancelled each other out. It tends to understate the error in all cases.

2- **Mean Absolute Error,** MAE $=\frac{1}{T}\sum_{t-1}^{T}\left|e_{1+k,t}\right|$, is the mean, or average of the absolute values of the errors. MAE is a way of dealing with the understatement of ME. By using the absolute values of the error, the mean gives a better indication of the model's fit.

3- **Mean Squared Error** , MSE $=\frac{1}{T}\sum_{t-1}^{T}e_{t+k,t}^{2}$ the MSE eliminates the positive-negative problem by squaring the errors. The result tends to place more emphasis on the larger errors and, therefore, gives a more conservative measure than the MAE. This approach penalizes large forecasting errors.

4- **Mean Percent Error**, MPE= $\frac{1}{T}\sum_{t-1}^{T}P_{t+k,t}$, is the average of the percentage errors of a forecast. The MPE is a relative measure of the forecasting error. It is subject to the "averaging" of the positive and negative errors. Can be used to determine if a forecasting method is biased (consistently forecasting low or high). Large positive MPE implies that the method consistently under estimates, large negative MPE implies that the method consistently over estimates, and The forecasting method is unbiased if MPE is close to (0).

5- **Root Mean Squared Error**, RASE $=\sqrt{\frac{1}{T}\sum_{t-1}^{T}e_{t+k,t}^{2}}$ The RMSE is easy for most people to interpret because of its similarity to the basic statistical concept of a standard deviation, and it is one of the most commonly used measures of forecast accuracy. We can rewrite its like RMSE = $\sqrt{\frac{1}{T}\sum_{t-1}^{T}(y_t-\hat{y}_t)^2}$.

6- **Mean Absolute Percent Error**, MAPE= $\frac{1}{T}\sum_{t-1}^{T}\left|P_{t+k,t}\right|$, is the average of the absolute values of the percentage errors of a forecast. it is a comparative measure that does not have the problem of averaging the positive and negative errors. and It is relatively easy to use to communicate a model's effectiveness. MAPE provides an indication of how large the forecast errors are in comparison to actual values of the series. Especially useful when the y_t values are large. It Can be used to compare the accuracy of the same or different methods on two different time series data. In closing this section, we note that it is sometimes informative to compare the accuracy of a forecast to that of a "naive" competitor. A simple and popular such comparison is achieved by Theil's (1961) U statistic, which is the ratio of the 1-step-ahead MSE for a given forecast relative to that of a random walk forecast $\hat{y}_{t+1,t} = y_t$, that is ,

$$U = \frac{\sum_{t=1}^{T}\left(y_{t+1}-\hat{y}_{t+1,t}\right)^2}{\sum_{t=1}^{T}(y_{t+1}-\hat{y}_t)^2} \qquad (2.22)$$

Generalization to other loss functions and other horizons is immediate. The statistical significance of the MSE comparison underlying the U statistic may be ascertained using the methods just described. One must remember, of course, that the random walk is not necessarily a naive competitor, particularly for many economic and financial variables, so that values of the U statistic near one are not necessarily bad. The use of Theil's U-statistic:

U = 1 The naive method is as good as the forecasting technique being evaluated.

U < 1 The forecasting technique being used is better than the naïve method.

U > 1 There is no point in using a formal forecasting method since using a naive method will produce better results .

2.6 Summary

In this chapter, we have discussed the basic concepts of time series models . We illustrated the (ACF, PACF , IACF) and the concepts with the basic autoregressive integral moving average (ARIMA) time series models. Also we talked about time series model building mechanism identification ,parameter estimation and model diagnostic, and talked about forecasting ARIMA models, forecast accuracy and measures of forecast accuracy .

Chapter 3
Model Identification Methods

3.1 Introduction

In this chapter we will define ARIMA parameters which needs to identify the model range, introduce identification methods like Box and Jenkins' method and the difficulties with mixed ARMA models, automated methods and make comparison between them.

Box, Jenkins, & Reinsel (1994) claims that , in time series analysis, within the ARIMA model framework, involves important steps, that model identification step, in which the researcher tries to identify which underlying mathematical model is appropriate for the data. Model identification focuses on the dependency parameters, one of the types of parameters unique to longitudinal designs. This step can sometimes be a very difficult, complicated, and problematic task specially in mixed model. Model identification can represents the primary goal of the analysis, especially if a researcher is trying to identify the basic underlying process represented in a time series data set.

3.2 Definition of ARIMA Parameters

The ARIMA model represents a family of models characterized by three parameters (p,d,q) that describe the basic properties of a specific time series model. The value of the first parameter, p, denotes the order of the autoregressive component of the model. If an observation can be influenced only by the immediately preceding observation, the model is of the first order. If an observation can be influenced by both of the two immediately preceding observations, the model is of the second order. The value of the second parameter, d, refers to the order of differencing that is necessary to stabilize a nonstationary time series. This process is described as nonstationary because values do not vary about a fixed mean level; rather, the series may first fluctuate about one level for some observations, and then rise or fall about a different level at a different point in the series. And the value of the third parameter, q, denotes the order of the moving averages component of the model. Again, the order describes how many preceding observations must be taken into account. The values of each of the parameters (p,d,q) of the model may be designated as order 0, 1, 2, or greater, with a parameter equal to zero indicating the absence of that term from the model. Higher-order models, four and above, are generally rare in the behavioral and social sciences Glass (1975).

Box, Jenkins, and Reinsel (1994) provide a more complete discussion of these parameters. The order of a time series parameter reflects how far into the past one must go to predict a present observation and thus refers to how many preceding observations must be taken into account to accurately describe the dependency present in the data series. Accuracy in determining the exact order can be quite difficult because higher-order autocorrelation terms are generally closer to zero than terms of earlier order. In effect, the higher-order terms become more likely to be included within the interval that would include an error estimate.

3.3 Model Identification Methods

Choi(1992)claimed that there are two major approaches for model identification: the traditional Box-Jenkins procedure and time series diagnostics by automated methods. The later consists in detailed inspection of empirical autocorrelation and partial autocorrelation functions (ACF and PACF) and in comparing their shape and value with theoretical ARIMA patterns. Model selection using the Box-Jenkins methodology is a sophisticated procedure requiring many data points and a great deal of researcher expertise. Velicer and Harrop (1983) demonstrated that even highly trained judges experienced considerable difficulty in identifying models present in computer-generated series, only 28% of the cases were classified correctly. Therefore, more reliable and less subjective automated methods are the focus of this research .

3.3.1 Box and Jenkins' method

As stated previously, in the model identification stage we try to determine "appropriate" orders for p, d, q, P, D and Q of the ARIMA(p, d,q) x (P, D, Q), model. possibly after the application of some suitable transformation to the data to stabilize the variance. The first step in the Box and Jenkins' method is to determine whether or not to difference the series. To achieve this, apart from the plot of the data, the sample autocorrelation function (SACF)

$$C_j = \frac{\sum_{t=1}^{N-j}(Z_t - \bar{z})(Z_{t+j} - \bar{z})}{\sum_{t=1}^{N}(z_t - \bar{z})^2}$$

where $\bar{z} = \sum_{t=1}^{N} {z_t}/{N}$, is an indispensable tool.

If a process requires some differencing to induce stationarity, it can be shown that the sample auto correlations will typically behave as a very smooth function, and thus not die out rapidly, at high lags. The failure of the SACF to die out at high lags thus indicates that differencing is required. This is so even though the first few sample autocorrelations need not necessarily be large. After having determined the stationary transformation, that is the operator $\nabla^d \nabla_s^D$ in the mode

$$\begin{cases} \emptyset(\mathrm{B})\phi(\mathrm{B^s})\nabla^{\mathrm{d}}\nabla_{\mathrm{s}}^{\mathrm{D}}\mathrm{z_t} = \mathrm{C} + \theta(\mathrm{B})\Theta(\mathrm{B^s})\mathrm{a_t} \\ \quad or\ for\ no \text{ seasonality} \\ \emptyset(B)\Phi(B)\nabla^d\nabla_s^D z_t = \theta(B)\Theta(B)\mathrm{a}_t \end{cases} \tag{3.1}$$

Where s is the degree of seasonality

the following step is to find an ARMA model for the stationary series $u_{t=\nabla^d \nabla_s^D z_t}$ .The SACF and sample partial autocorrelation function (SPACF) are the fundamental tools to obtain the orders of the autoregressive and moving average parts. The partial autocorrelation function is defined by

$$a_1 = corr\ (u_2, u_1)$$

and

$$a_1 = corr\ \left\{u_{j+1} - E\left({}^{u_{j+1}}/_{u_2}, u_3, .., u_j\right), u_1 - E\left({}^{u_1}/_{u_2}, u_3, .., u_j\right)\right\},\ \ j \geq 2$$

where E(.I.) denotes orthogonal projection and Corr stands for correlation. It can be shown that the partial autocorrelations a_j can be

estimated by successive AR(j) fittings. That is, if we fit

$$u_t + \widehat{\emptyset}_{j1} u_{t-1} + \ ... + \widehat{\emptyset}_{jj} u_{t-j}\ = \hat{C} + \ a_t$$

for j = 1,2, ... , then $\hat{a}_j = \widehat{\emptyset}_{jj}$ the sample PACF.

It can also be shown that if $\{u_t\}$ is a pure autoregressive model, i.e., q = Q = 0 in (3.1), the theoretical partial autocorrelation function dies out in an exponential or sinusoidal fashion and has the cut-off feature, which means that the partial autocorrelations a_j will obey $a_k = 0$ for all $k \geq p + P$. Also, if $\{u_t\}$ is a pure moving average model, i.e., p = P = 0 in (3.1), the theoretical autocorrelation function dies out in an exponential or sinusoidal fashion and has the cut-off feature, which means that the autocorrelations ρ_j will obey $\rho_k = 0$, for all k>q+ Q. Estimated auto correlations can have rather large variances and can be highly autocorrelated with each other. Therefore, one has to be careful when working with the SACF because detailed adherence to the theoretical autocorrelation function cannot be expected. It may be the case that moderately large estimated auto correlations occur after the theoretical autocorrelation function has damped out, or apparent ripples and trends may be present in the SACF which have no basis in the theoretical function. If $\{u_t\}$ is a pure moving average process, the elements of the sequence

$\{\sqrt{N}C_j\}$, $j > q + Q$,

are asymptotically normal with mean 0 and variance

$$\hat{\sigma}^2(C_j) = 1 + 2\sum_{k=1}^{q+Q} C_k^2 \;, j > q + Q\;.$$

Then, an approximate 100(1 - α) percent confidence interval of

$$\rho_j \; is \left(-{}^{z\alpha/2}\Big/_{\sqrt{N}}, z_{\alpha/2} \times {}^{\hat{\sigma}(C_j)}\Big/_{\sqrt{N}}\right), j > q + Q$$

where $z_{\alpha/2}$ is the z- value such that the area α /2 lies to its right in the standard normal probability density. If $\{u_t\}$ is a pure autoregressive process, then the elements of the sequence $\{\sqrt{N}\emptyset_{jj}\}$, $j > p + P$, are asymptotically independent random variables with mean 0 and variance 1. Therefore, an approximate 100(1 - α) percent confidence interval of $\emptyset_{jj}$ is

$$\left(-{}^{z\alpha/2}\Big/_{\sqrt{N}}, {}^{z\alpha/2}\Big/_{\sqrt{N}}\right), j > p + P$$

For mixed ARM A models, the identification can be more difficult if only the SACF and SPACF of the series are available. Box and Jenkins (1976) provided some information on how to determine the orders of the autoregressive and moving average parts from "reading" the SACF and SPACF of a stationary series. However, this approach is usually not very effective in practice.

The difficulties with mixed ARMA models

As we show that the sample autocorrelation, partial autocorrelation and inverse autocorrelation functions can effectively identify pure autoregressive and pure moving average models. On the other hand, when both the degrees of the autoregressive polynomial (p+P) and the moving average polynomial (q + Q) are not zero, the previous functions are much more difficult to interpret. In this case, other model identification methods, different from the classical methods, are called for. The difficulty of identifying mixed ARMA models is further increased when seasonality is also present in the time series at hand. Several major advances have been made in the last three decades to identify ARIMA models for non-seasonal time series. Among these, we can mention the extended autocorrelation function(EACF) and the smallest canonical correlation(SCAN) methods developed by Tsay and Tiao (1984, 1985). These methods are very informative in the identification of ARIMA models for non-seasonal time series, but they are less successful when they are directly applied to seasonal time series. It is to be noted that these methods can also be used with nonstationary series.

Since the early 1970s, some penalty function methods have been proposed for ARMA model identification. These methods can be used with seasonal time series and their popularity is constantly increasing. The reason for this is that they can be effective, computationally cheap and objective. However, although some results have been extended to nonstationary series, these methods are in principle only applicable to stationary series.

3.3.2 Automated method's

Numerous alternatives to the Box-Jenkins approach have been developed during the last three decades to make the model identification process more reliable and less subjective.

Choi (1992) published a survey of different procedures for model identification, where automated methods are classified into three categories:

- penalty function methods (e.g., BIC of Rissannen(1978), Schwarz(1978), AIC of Akaike (1974)).
- innovation regression methods (e.g., HR of Hannan & Rissanen(1982), KP of Koreisha & Pukkila(1990)).
- and pattern identification methods (e.g., the corner method of Beguin, Gourieroux, & Montfort(1980), ESACF and SCAN of Tsay & Tiao(1984),(1985)).

Penalty Function Methods

In the identification stage, once the differencing orders d and D in (3.1) have been obtained for the nonstationary series $\{z_t\}$, the problem remains of finding an ARMA model for the differenced series

$u_t = \nabla^d \nabla_s^D z_t$.

Since the early 1970s, some procedures to be determined the orders k and i of an ARMA(k, i) model have been proposed which minimize a function of the form

$$P(K, i) = \ln \hat{\sigma}_{k,i}^2 + (k+i)\frac{C(M)}{M}; \;\; k \leq K, i \leq I \qquad (3.2)$$

where $\hat{\sigma}_{k,i}^2$ is the maximum likelihood estimate of the variance of the

white noise variance, C(M) is some function of the number of observations M of the series, and K and I are upper bounds for the orders, usually imposed a priori. Because $\hat{\sigma}_{k,i}^2$

decreases as the orders increase, it cannot be a good criterion to choose the orders minimizing it. This is the reason why the penalty term $(k+i)\frac{C(M)}{M}$ is included.

The penalty function identification methods are regarded as objective. if $C(M)$ in (3.2) is replaced with 2, we obtain the famous AlC criterion, which stands for Akaike's Information Criterion. Other possible choices are $C(M) = \ln(M)$, which corresponds to the BIC (Bayesian Information Criterion), and $C(M) = 2\ln(\ln(M))$ which gives the HQ criterion (Hannan and Quinn). The BIC criterion imposes a greater penalty term than does AIC. One criterion for selection of AR(p) models is the FPE (Final Prediction Error) criterion, which is given by $\mathrm{FPE}(p) = \left\{1 + \left(\frac{P}{M}\right)\right\}\hat{\sigma}_p^2$.

The BIC criterion estimates the orders of an ARMA model consistently, whereas the AIC does not. However, this is not a reason to prefer BIC instead of AIC because consistency is based on the assumption that there is a "true" ARMA model for the series and this is doubtful proposition. Models are artificial constructs and probably there is no such a thing as a true model. for example, Liitkepohl (1985), that the BIC criterion works better in practice than AlC, in terms of selecting more often the original model when working with simulated series and selecting models with a better fit when working with real series. Although the penalty function methods are in principle computationally expensive, because they need maximum likelihood estimates for all possible ARMA models, there are methods, like the Hannan-Rissanen's method, which use cheaper estimates based on linear regression techniques only. Also, in the case of multiplicative seasonal ARMA models, it will be seen that it is possible to further reduce the computational burden by proceeding sequentially. That is, by iterating between selections of the regular and of the seasonal parts. The penalty function methods can also be used to identify vector ARMA models. The penalty functions to use with multivariate data are direct generalizations of the ones for the univariate case. This is a great advantage, not shared by many of the other identification methods.

Pattern Identification Methods

Since the early 1980s, some methods have been applied for determining the orders of an ARMA process which use the extended Yule Walker equations. For the following equation {the ARMA(p,q)} .

$$z_t + \emptyset_1 z_{t-1} + \ldots + \emptyset_p z_{t-p} = C + a_t + \theta_1 a_{t-1} + \ldots + \theta_q a_{t-q} \qquad (3.3)$$

these last equations are given by

$$\gamma_j = -\emptyset_1 \gamma_{j-1} - \ldots - \emptyset_p \gamma_{j-p} \quad , j = q+1, q+2, \ldots$$

Where $\gamma_j, j = 0,1, \ldots$ . ., is the covariance function of the process. These methods are often called pattern identification methods. (Choi (1992))

It is to be noted that, contrary to Choi's remark about penalty function methods being computationally exorbitant and pattern identification methods being computationally cheap. The pattern identification methods are so called because they are based on certain functions which give rise to two-way arrays with distinctive patterns. For each ARMA(p, q) model, the corresponding two-way array shows a unique pattern. Using the sample analog of this two-way array, an ARMA model is identified by looking for a theoretical pattern which is closely resembled by the sample one. Among the many pattern identification methods which have been proposed in the literature, we can mention the Corner method by Begnin, Gourieroux and Monfort (1980), the extended sample autocorrelation method by Tsay and Tiao (1984) and the smallest canonical correlation method by Tsay and Tiao (1985). These last two methods can be effective with non-seasonal time series and can also be used with nonstationary series. However, the Corner method, which can only be used with stationary series, do not seem to be very useful even for data with no seasonality.

Innovation regression method

Hannan and Rissanen's minimum information criterion (MINIC) combines the regression technique and the penalty functions AIC and BIC for the modeling of stationary and invertible ARMA (p, q) processes. The MINIC procedure consists of three steps.

1- The first is to fit a high-order AR model to the observations. The choice of autoregressive parameter is determined by the order minimizing the AIC.

2- The second step is to apply the ordinary least squares (OLS) method to the series and estimated innovations of the fitted AR model. As a result, m autoregressive and j moving-average OLS estimates are obtained $m = P_{min}, \dots, P_{max}$. is the autoregressive test order, $j = q_{min}, \dots, q_{max}$. is the moving-average test order.

3- As the last step, the BIC is computed for each of $m \times j$ ARMA models. A model with the smallest BIC is used as a MINIC recommendation.

3.4 Some selected pattern identification methods

The sample ACF and PACF provide effective tools for identifying pure AR(p) or MA(q) models. However, for a mixed ARMA model, its theoretical ACF and PACF have infinitely many nonzero values, making it difficult to identify mixed models from the sample ACF and PACF. Many methods have been proposed to make it easier to identify the ARMA orders, for example, the corner method (Becuin et al., 1980), the extended autocorrelation

(ESACF) method (Tsay and Tiao, 1984), The minimum information criterion (MINIC) method can tentatively identify the order of a stationary and invertible ARMA process. (Hannan and Rissannen (1982)), and the smallest canonical correlation (SCAN) method (Tsay and Tiao, 1985), among others. The model identification of mixed ARMA model is much more complicated than that of pure AR or MA models as we know. We shall consider four methods:

3.4.1 The Extended Autocorrelation Function

The first method to identify the order of a mixed model is the extended autocorrelation function (EACF) of Tsay and Tiao (1984). The EACF, in fact, applies to ARIMA as well as ARMA models. However, it treats an ARIMA(p, d, q) model as an ARMA(p + d, q) model. The basic idea of EACF is based on the "generalized" Yule-Walker equation. Conceptually, it involves two steps. In the first step, we attempt to obtain consistent estimates of AR coefficients. Given such estimates, we can transform the ARMA series into a pure MA process. The second step then uses the sample ACF of the transformed MA process to identify the MA order q. The best way to introduce EACF is to consider simple example:

Example: Suppose that Z_t is an ARMA(1,1) model

$$Z_t - \emptyset Z_{t-1} = a_t - \theta a_{t-1}, \qquad |\emptyset| < 1, \qquad |\theta| < 1$$

For this model, the ACF is

$$\rho_l = \begin{cases} \frac{(1-\theta\emptyset)(\emptyset-\theta)}{1+\theta^2-2\emptyset\theta} & for\ l = 1 \\ \emptyset\rho_{l-1} & for\ l > 1 \end{cases} \qquad (3.4)$$

For p = 1, the usual Yule-Walker equation is $\rho_1 = \emptyset\rho_0$

and the j-th generalized Yule-Walker equation is $\rho_{j+1} = \emptyset\rho_j$

Denote the solution of the Yule-Walker equation by $\emptyset_{1,1} = \emptyset_{1,1}^{(0)}$ and that of the j-th generalized Yule-Walker equation by $\emptyset_{1,1}^{(j)}$. Then, we have

$$\emptyset_{1,1}^{(j)} = \begin{cases} \rho_1 \neq \emptyset & for\ j = 0 \\ \emptyset & for\ j > 0 \end{cases} \qquad (3.5)$$

Thus, the solution of the usual Yule-Walker equation is not consistent with the AR coefficient $\emptyset$. However, ALL of the solutions of the j-th generalized Yule-Walker equations are consistent with the AR coefficient. In sample, these results say that the estimates of $\emptyset_{1,1}^{(j)}$ obtained by replacing the ACF by sample ACF have the property:

$$\widehat{\emptyset}_{1,1}^{(j)} = \begin{cases} \rho_1 & for\ j = 0 \\ \emptyset & for\ j > 0 \end{cases}$$

Now define the transformed series $W_{1,t}^{(j)}$ by:

$$W_{1,t}^{(j)} = Z_t - \widehat{\emptyset}_{1.1}^{(j)} Z_{t-1} \quad \text{for } j > 0 \qquad (3.6)$$

The above discussion shows that $W_{1,t}^{(j)}$ for j > 0 is asymptotically a pure MA(1) process. Consequently, by considering the ACF of the $W_{1,t}^{(j)}$ series, we can identify that the MA order is 1. (Jonathan D. Cryer • Kung,2008)

Now Model identification via EACF, to make use of the EACF for model identification , we consider the two-way table:

AR	MA(j)					
m	0	1	2	3	4	…
0	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	….
1	$\rho_{1,1}$	$\rho_{1,2}$	$\rho_{1,3}$	$\rho_{1,4}$	$\rho_{1,5}$	….
2	$\rho_{2,1}$	$\rho_{2,2}$	$\rho_{2,3}$	$\rho_{2,4}$	$\rho_{2,5}$	….
3	$\rho_{3,1}$	$\rho_{3,2}$	$\rho_{3,3}$	$\rho_{3,4}$	$\rho_{3,5}$	….
…	…..	…..	…..	….	…..	…..

Table (3.1) The EACF To identify the order of an ARMA model

In practice, the EACF in the above table(3.1) is replaced by its sample counterpart. To identify the order of an ARMA model, we need to understand the behavior of the EACF table for a given model. Before giving the theory, we shall illustrate the function of the table. Suppose that Z_t is an ARMA(1,1) model, then the corresponding EACF table is

AR	MA(j)						
m	0	1	2	3	4	5	…
0	**X**	**X**	**X**	**X**	**X**	**X**	….
1	**X**	**0***	**0**	**0**	**0**	**0**	….
2	**X**	**X**	**0**	**0**	**0**	**0**	….
3	*	*	**X**	**0**	**0**	**0**	….
4	*	*	*	**X**	**0**	**0**	
5	*	*	*	*	**X**	**0**	
…	…..	…..	…..	….	…..	……	…..

Table(3.2):autocorrelation ARMA(1,1) model, EACF method
where "X" and "O" denotes non-zero and zero quantities, respectively,
"*" represents a quantity which can assume any value between −1 & 1.

From the table, we see that there exists a triangle of "O" with vertex at (1, 1), which is the order of Z_t. In practice, the non-zero and zero terms are determined by the sample EACF and its estimated standard error via the Bartlett's formula for MA models. Of course, we cannot expect to see an exact triangle as that of the above table. However, one can often make a decision based on the pattern of the EACF table(3.2).

To understand the triangular pattern, it is best to consider a simple example such as ARMA(1,1) model of the above table. In particular, we shall discuss the reason why $\rho_{2,2}$ is different from zero for an ARMA(1,1) model. By definition, $\rho_{2,2}$ is the lag-2 ACF of the transformed series $W_{2,t}^{(2)} = Z_t - \emptyset_{2,1}^{(2)} Z_{t-1} - \emptyset_{2,2}^{(2)} Z_{t-2}$

where $\emptyset_{2,1}^{(2)}$ and $\emptyset_{2,2}^{(2)}$ are the solution of the 2nd generalized Yule-Walker equation of order 2, namely

$$\begin{bmatrix} \rho_3 \\ \rho_4 \end{bmatrix} = \begin{bmatrix} \rho_2 & \rho_1 \\ \rho_3 & \rho_2 \end{bmatrix} \begin{bmatrix} \emptyset_{2,1}^{(2)} \\ \emptyset_{2,2}^{(2)} \end{bmatrix}$$

However, for an ARMA(1,1) model, $\rho_j = \emptyset\rho_{j-1}$ for $j > 1$ so that the above Yule-Walker equation is "singular" in theory. In practice, the equation is not exactly singular, but is ill-conditioned. Therefore, the solution $\emptyset_{2,1}^{(2)}$ and $\emptyset_{2,2}^{(2)}$ can assume any real numbers. Consequently, the chance that $\emptyset_{2,i}^{(2)} = 0$ is essential zero. More importantly, this implies that the transformed series $W_{2,t}^{(2)}$ is not an MA(1) series. Therefore, $\rho_{2,2} \neq 0$. Intuitively, one can interpret this result as an over-fitting phenomenon. Since the true model is ARMA(1,1)and we are fitting an AR(2) polynomial in the construction of $W_{2,t}^{(2)}$, the non-zero $\rho_{2,2}$ is in effect a result of over fitting of the second AR coefficient.

Using exactly the same reasoning, one can deduce the triangular pattern of the EACF table. Thus, it can be said that the triangular pattern of EACF is related to the over fitting of AR polynomials in constructing the transformed series $W_{m,t}^{(j)}$.

3.4.2 The MINIC Method

Hannan and Rissanen's minimum information criterion (MINIC) combines the regression technique and the penalty functions AIC and BIC for the modeling of stationary and invertible ARMA (p, q) processes. The MINIC procedure consists of three steps. The first is to fit a high-order AR model to the observations. The choice of autoregressive parameter is determined by the order minimizing the AIC. The second step is to apply the ordinary least squares (OLS) method to the series and estimated innovations of the fitted AR model. As a result, m autoregressive and j moving-average OLS estimates are obtained $m = P_{min}, \dots, P_{max}$. is the autoregressive test order, $j = q_{min}, \dots, q_{max}$. is the moving-average test order. As the last step, the BIC is computed for each of $m \times j$ ARMA models. A model with the smallest BIC is used as a MINIC recommendation. For more illustration:

The Minimum Information Criterion (MINIC) method can tentatively identify the order of a stationary and invertible ARMA process. Given a stationary and invertible time series $\{Z_t: 1 \leq t \leq n\}$ with mean corrected form $\hat{Z}_t = Z_t - \mu_z$, with a true autoregressive order of p, and with a true moving average order of q, you can use the MINIC method to compute information criteria (or penalty functions) for various autoregressive and moving average orders. The following paragraphs provide a brief description of the algorithm. If the series is a stationary and invertible ARMA(p,q) process of the form

$$\Phi_{(p,q)}(B)\hat{Z}_t = \Theta_{(p,q)}(B)\epsilon_t \quad (3.7)$$

the error series can be approximated by a high-order AR process

$$\hat{\epsilon}_t = \widehat{\Phi}_{(p_\epsilon,q)}(B)\hat{Z}_t \approx \epsilon_t \quad (3.8)$$

where the parameter estimates $\widehat{\Phi}_{(p_\epsilon,q)}$ are obtained from the Yule-Walker estimates. The choice of the autoregressive order, p_ϵ, is determined by the order that minimizes the Akaike information criterion (AIC) in the rang $p_{\epsilon,min} \leq p_\epsilon \leq p_{\epsilon,max}$

$$AIC((p_\epsilon, 0) = \ln(\hat{\sigma}^2{}_{(p_\epsilon,0)} + 2(p_\epsilon + 0)/n \quad (3.9)$$

Where $\hat{\sigma}^2{}_{(p_\epsilon,0)} = \frac{1}{n}\sum_{t=p_\epsilon+1}^{n}\hat{\epsilon}_t^2$

Note that Hannan and Rissannen (1982) use the Bayesian information criterion (BIC) to determine the autoregressive order used to estimate the error series. Box et al (1994) and Choi (1990) recommend the AIC.

Once the error series has been estimated for autoregressive test order $m = P_{min}, \dots, P_{max}$ and for moving-average test order $j = q_{min}, \dots, q_{max}$, the OLS estimates $\widehat{\Phi}_{(m,j)}$ *and* $\widehat{\Theta}_{(m,j)}$ are computed from the regression model

$$\hat{Z}_t = \sum_{i=1}^{m}\emptyset_i^{(m,j)}\hat{Z}_{t-i} + \sum_{k=1}^{j}\theta_k^{(m,j)}\hat{\epsilon}_{t-k} + \text{error} \quad (3.10)$$

From the preceding parameter estimates, the BIC is then computed

$$\mathrm{BIC(m, j)} = \ln(\hat{\sigma}^2{}_{(\mathrm{m,j})}) + 2(\mathrm{m+j})\ln(\mathrm{n})/\mathrm{n} \qquad (3.11)$$

Where $\hat{\sigma}^2{}_{(m,j)} = \frac{1}{n}\sum_{t=t_0}^{n}(\hat{z}_t - \sum_{i=1}^{m} \emptyset_i^{(m,j)}\, \hat{z}_{t-i} + \sum_{k=1}^{j} \theta_k^{(m,j)}\, \hat{\epsilon}_{t-i})$

Where $t_0 = p_\epsilon + \max(m, j)$.

A MINIC table is then constructed using BIC (m, j) (see Table 3.3). If

$p_{max} > p_{\epsilon,min}$, the preceding regression may fail due to linear dependence on the estimated error series and the mean-corrected series. Values of BIC (m,j) that cannot be computed are set to missing. For large autoregressive and moving average test orders with relatively few observations, a nearly perfect fit can result. This condition can be identified by a large negative BIC(m, j) value.

	MA				
AR	0	1	2	3	…
0	BIC(0,0)	BIC(0,1)	BIC(0,2)	BIC(0,3)	…
1	BIC(1,0)	BIC(1,1)	BIC(1,2)	BIC(1,3)	…
2	BIC(2,0)	BIC(2,1)	BIC(2,2)	BIC(2,3)	…
3	BIC(3,0)	BIC(3,1)	BIC(2,3)	BIC(3,3)	…
..	…	…	….	…	…

Table (3.3) MINIC Table identified by a least BIC(m, j) value

3.4.3 The SCAN Method

The Smallest Canonical (SCAN) correlation method can tentatively identify the orders of a stationary or nonstationary ARMA process. Tsay and Tiao (1985) proposed the technique, and Box et al (1994) and Choi (1990) provide useful descriptions of the algorithm. Given a stationary or nonstationary time series $\{Z_t: 1 \leq t \leq n\}$ with mean corrected form $\hat{Z}_t = Z_t - \mu_z$, with a true autoregressive order of p + d, and with a true moving average order of q, you can use the SCAN method to analyze eigenvalues of the correlation matrix of the ARMA process. The following paragraphs provide a brief description of the algorithm. For autoregressive test order $m = p_{min}, \dots, p_{max}$ and for moving-average test order, $j = q_{min}, \dots, q_{max}$ perform the following steps.

1. Let$Y_{m,t} = (\hat{z}_t, \hat{z}_{t-1}, \dots, \hat{z}_{t-m})$. Compute the following $(m+1) \times (m+1)$matrix.

$$\hat{B}(m, j+1) = (\sum_t Y_{m,t-j} Y^t_{m,t-j-1})^{-1} (\sum_t Y_{m,t-j-1} Y^t_{m,t})$$

$$\hat{B}^*(m, j+1) = (\sum_t Y_{m,t} Y^t_{m,t})^{-1} (\sum_t Y_{m,t} Y^t_{m,t-j-1}) \quad (3.12)$$

$$\hat{A}^*(m, j) = \hat{B}^*(m, j+1) \hat{B}\ (m, j+1) \quad (3.13)$$

where t ranges from j + m + 2 to n.

2. Find the *smallest* eigenvalue, $\hat{\lambda}^*(m,j)\ of\ \hat{A}^*(m,j)$ and its corresponding *normalized* eigenvector , $\Phi_{m,j} = (1, -\emptyset_1^{(m,j)}, -\emptyset_2^{(m,j)}, \dots, -\emptyset_m^{(m,j)}$

The squared canonical correlation estimate is $\hat{\lambda}^*(m,j)$.

3. Using the $\Phi_{m,j}$ as AR(m) coefficients, obtain the residuals for $t = j + m + 1\ to\ n$ by following the formula:

$$w_t^{(m,j)} = \hat{Z}_t - \emptyset_1^{(m,j)} \hat{Z}_{t-1}, -\emptyset_2^{(m,j)} \hat{Z}_{t-2}, \dots, -\emptyset_m^{(m,j)} \hat{Z}_{t-m}. \quad (3.14)$$

4. From the sample autocorrelations of the residuals, $r_k(w)$, approximate the standard error of the squared canonical correlation estimate by $var(\hat{\lambda}^*(m,j)^{1/2}) \approx d(m,j)/(n-m-j)$

Where $d(m,j) = 1 + 2\sum_{i=1}^{j-1} r_k(w^{(m,j)}))$ (3.15)

The test statistic to be used as an identification criterion is

$$c(m,j) = -(n-m-j)\ln\left(1 - \frac{\hat{\lambda}^*(m,j)}{d(m,j)}\right) \quad (3.16)$$

which is asymptotically $x_1^2\ if\ n = p + d\ and\ j \geq q\ or\ if\ m \geq p + d\ and\ j = q\ , for\ m > p\ and\ j < q$ there is more than one theoretical zero canonical correlation between$Y_{m,t}$ and

$Y_{m,t-j-1}$. Since the $\hat{\lambda}^*(m,j)$ are the smallest canonical correlations for each (m, j), the percentiles of c(m, j) are less than those of a x_1^2 , therefore, Tsay and Tiao (1985) state that it is safe to assume a x_1^2, For m < p and j < q , no conclusions about the distribution of c(m, j) are made. A SCAN table is then constructed using c(m,j) to determine which of the $\hat{\lambda}^*(m,j)$ are significantly different from zero (see Table 3.5). The ARMA orders are tentatively identified by finding a (maximal) rectangular pattern in which the $\hat{\lambda}^*(m,j)$ are insignificant for all test orders$m \geq p + d \text{ and } j \geq q$. There may be more than one pair of values (p + d, q) that permit such a rectangular pattern. In this case, parsimony and the number of insignificant items in the rectangular pattern should help determine the model order. Table (3.5) depicts the theoretical pattern associated with an ARMA(2,2) series.

	MA				
AR	0	1	2	3	…
0	C(0,0)	C(0,1)	C(0,2)	C(0,3)	…
1	C(1,0)	C(1,1)	C(1,2)	C(1,3)	….
2	C(2,0)	C(2,1)	C(2,2)	C(2,3)	…
3	C(3,0)	C(3,1)	C(3,2)	C(3,3)	…
….	…	…	….	….	…

Table (3.4) SCAN Table the smallest canonical correlations for each (m,j)

	MA							
AR	0	1	2	3	4	5	6	7
0	*	X	X	X	X	X	X	X
1	*	X	X	X	X	X	X	X
2	*	X	0	0	0	0	0	0
3	*	X	0	0	0	0	0	0
4	*	X	0	0	0	0	0	0
5	*	X	0	0	0	0	0	0

Table (3.5) Theoretical SCAN Table for an ARMA(2,2) Series
X: significant terms , 0: insignificant terms , *: no pattern

3.4.4 The Corner method as mathematical pattern identification method

The Corner method produce a table whose values will show a great rectangle of zeros starting from some indexes takes place (p, q) of the table (3.6). This couple (p,q) will characterize the time series model ARMA (p, q). The table is built starting from the calculation of the following determining:

$$\Delta(r,s)\begin{vmatrix}\rho_s & \rho_{s-1} & \rho_{s-2} & \cdots & \rho_{s-r+1}\\ \rho_{s+1} & \rho_s & \rho_{s-1} & \cdots & \rho_{s-r}\\ & & \cdot & & \\ & & \cdot & & \\ \rho_{s+r-1} & \rho_{s+r-2} & \rho_{s+r-3} & \cdots & \rho_s\end{vmatrix} \qquad (3.17)$$

For r ≥ 1 and s ≥ 1. It is important to note ρ_s is the value of the Autocorrelation Function for the delay s and $\rho_s = \rho_{-s}$. In such a way that:

$$c(i,j) = \Delta(j+1, i+1) \qquad (3.18)$$

For i = 0,1,2, ..., K, being K a relatively big number.

A theoretical Corner table for a model ARMA (p,q) is presented in Table1. It is expected that the Corner table produces a large south east rectangular sub-matrix whose all elements are zeros. The coordinates of the northwest corner of this zero-rectangle are (p,q). It provides a strong clue for us to identify the order of the underlying process

AR/MA	1	…	…	q-1	q	q+1	…	k
1	X	X	X	X	X	X	X	X
…	X	X	X	X	X	X	X	X
…	X	X	X	X	X	X	X	X
p-1	X	X	X	X	X	X	X	X
P	X	X	X	X	0	0	0	0
P+1	X	X	X	X	0	0	0	0
…	X	X	X	X	0	0	0	0
k	X	X	X	X	0	0	0	0

Table (3.6) Corner Table for and ARMA (p, q). Note: X represent nonzero value

Because in the practice we generally have a finite number of observations and, therefore, the values of the Autocorrelation Functions are estimates, it will be difficult to locate the points p, q by means of a visual inspection. It becomes necessary to specify an approach to differentiate the values theoretically zero and the other ones**.** Following . we simplify the Corner table using indicator symbols. The Simplified Corner table is defined as follows:

$$c^*(i,j)\begin{cases}0\ if\ \left|\frac{\Delta(j+1,i+1)}{SE(\Delta(j+1,i+1))}\right| < 2\\ x \qquad otherwise\end{cases} \qquad (3.19)$$

for i, j = 0, 1, ..., K. ”X " represents a nonzero element. The standard error of any Δ element in the estimated Corner table is given by:

$$SE(\Delta) = \sqrt{\frac{AGA'}{n}} \qquad (3.20)$$

where A is a (1 x h) vector with elements

$$a(j) = \frac{\partial\Delta}{\partial\rho_j}$$

h is the maximal lag among all the autocorrelations in Eq (3.17), n is the sample size, G is a (h x h) matrix whose (i, j) element is:

$$\sum_{k=-\infty}^{\infty}\{\rho_k\rho_{k-i+j} + \rho_{k+j}\rho_{k-i} - 2\rho_k\rho_j\rho_{k-i} - 2\,\rho_k\rho_j\rho_{k-j} + 2\rho_i\rho_j\rho_k^2\} \qquad (3.21)$$

3.5 Comparison among identification methods

In the classical method in pure (AR) and (MA) we have no problem to identify it, since we Know that the ACF , IACF and PACF that be sufficient to identify the model but the problem is clearly in the mixed model in this case we used the automated method then we compare between it. As we Know we have Penalty Function Methods and pattern identification methods and we intrudes the innovation regression method in the pattern methods.

3.5.1 Comparison between the penalty functions and pattern methods

Penalty function method is useful with seasonal time series and their popularity is constantly increasing. It's effective, computationally, objective and it being computationally exorbitant and there is important advantage that is no shared by many of other identification methods that is use with multivariate data are direct generalizations of the ones for the univariate case. Penalty function method can also be used to identify vector ARMA models.

Pattern method so called because they are based on certain functions which give rise to two-way arrays with distinctive patterns. It is less successful when they are directly applied to seasonal time series. It's being computationally cheap, some of pattern method working with seasonally data and it's not very useful. It's effective , sample for using and if we applied it's condition we have direct identification.

3.5.2 Comparison between the penalty function methods

If we return to section 3.3.2 we find most of penalty function methods hesitate from same source but some of advantage between them appear like BIC criterion works better in practice than AlC, in terms of selecting more often the original model when working with simulated series and selecting models with a better fit when working with real series .

Hannan-Rissanen's method, which use cheaper estimates based on linear regression techniques only.

We add some approved advantage like :

1- the Bayesian information criteria (BIC) performs best in terms of selecting the correct order of an Autoregressive model for small samples irrespective of the AR structure.

2- The Hannan Quinn criteria can be said to perform best in large sample .

(Shittu. O.I & Asemota. M.J (2009))

3.5.3 Comparison between the pattern methods

Now we want to compare between four important pattern method C-Table, SCAN,MINIC and ESACF methods. The pattern identification methods are so called because they are based on certain functions which give rise to two-way arrays with distinctive patterns. The method (ESACF) and method (SCAN) can be effective with non-seasonal time series and can also be used with nonstationary series.

The Corner method (C-Table), which can only be used with stationary series.

MINIC can only be only used with the modeling of stationary and invertible series.

Tetiana stadnytsky(2008), concluded the following:

1- the Corner method works well when the sample size is large, but it works poorly in small to moderate sample sizes.

2- the ESACF method is more robust with respect to sample size and works reasonably well .

3- in the pure structures, MINIC and SCAN perform well for autoregressive models, whereas ESACF works better in moving-average cases.

4- SCAN and ESACF are superior to MINIC for mixed (p,q) models.

5- the positive effect of sample size is more pronounced for MINIC than for SCAN and ESACF.

3.6 Summary

In this chapter we discussed the definition of the ARIMA parameter and model identification methods, also introduced classical method like box and Jenkins method and automated method, we discussed class of automated method .We concentrated on the penalty functions and pattern methods and we did comparisons between them.

Chapter 4

Case Study

4.1 Introduction

The objective of the present chapter is to illustrate the model identification methods by using actual data. We will talk about the nature of the data source, parameter estimation and forecasts of future values. Furthermore, our data was obtained from Palestinian Exchange (PEX) web page (www.p-s-e.ps). This web page provides several datasets of the opening stock price of the banking sector in the Palestinian stock market, and we will take weekly data from December 2007 to September 2012 .

4.2 Data Description

The data which we have, is weekly observations of open value and its range from December 2007 to September 2012 . Thus we have about (240) observation. To take a general idea about the data, we displayed descriptive statistics of the actual data, as in the following table:

descriptive statistics	Mean	Median	Maximum	Minimum	Std. Dev	Skewness	Kurtosis
	94.58	96.10	104.50	66.56	6.993769	-1.1575	0.9332982

Table(4.1): Descriptive statistics for the opening stock price of the banking sector in Palestine

In table (4.1) above, we can see some descriptive statistics of the data. The maximum open value in the period from December 2007 to September 2012 is 104.50, while the value 66.56 is recorded as the minimum open value for the same period.

4.3 Data Preparation

First, we just need to know whether or not the data is autocorrelated. A statistical test such as the Box– Ljung test can provide the answer. By using this test, we will get the p-value. Conventionally, a p-value of less than 0.05 indicates that, the data contains a significant autocorrelations.

```
Box-Ljung test
data:  open
X-squared = 208.2588, df = 1, p-value < 0.0002
```

Output (4.1) autocorrelations test of data (opening stock price)

From the test, the p-value is near zero, which indicates that the data has a significant autocorrelations. The data describes the data set which is recorded over a period of time, so our data (weekly open price) is a time series data.

4.4 The General ARIMA Model

Model identification

Stationary of actual Palestinian banking sector Data

1. Run Sequence Plot

A plot of the actual information is given by Figure 4.1

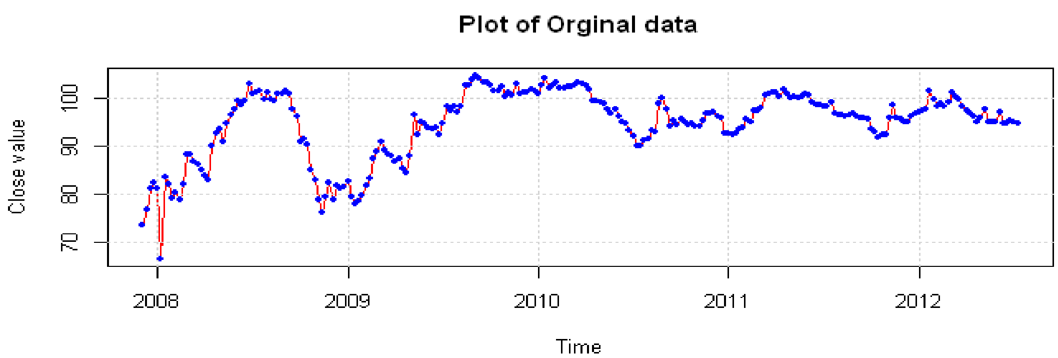

Figure 4.1 Weekly open price of Palestinian banking sector data

Figure 4.1 displays the time series of the open price Palestinin banking sector data , we might well consider a nonstationary model .

Stationary tests : PP.test , Alternative hypothesis is stationary

Phillips-Perron Unit Root Test data: open
Dickey-Fuller Z(alpha) = -17.1316, Truncation lag parameter = 4, p-value = 0.1468
alternative hypothesis: stationary

Output4.2 The first difference of open stock price Test for Level Stationary of Palestinian banking sector data

kpss.test: Alternative hypothesis is stationary

KPSS Test for Level Stationary data: open
KPSS Level = 1.511, Truncation lag parameter = 3, p-value = 0.01

Output4.3 The first difference of open stock price Test for Level Stationary of Palestinian banking sector data

Now from the plot of actual data and the stationary tests we find the data is nonstationary so, we make a first difference. After first difference we plot the data and we make stationary tests .

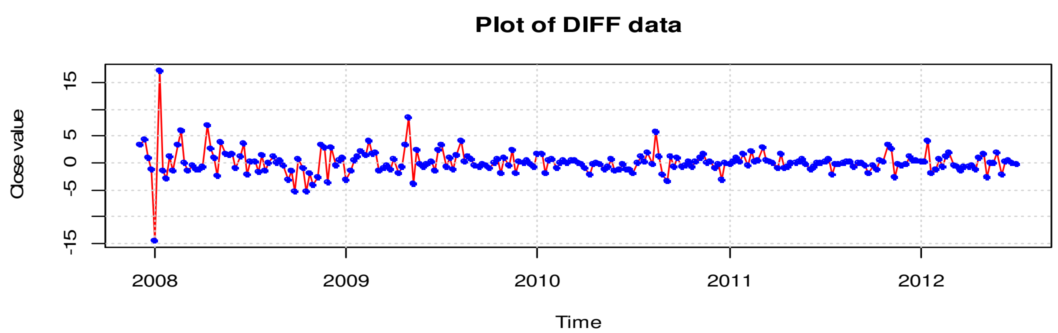

Figure 4.2 The first difference of weekly open stock price of Palestinian banking sector data

Phillips-Perron Unit Root Test data: dopen

Dickey-Fuller Z(alpha) = -261.4862, Truncation lag parameter = 4, p-value = 0.01

alternative hypothesis: stationary

Output 4.4 The first difference of weekly open stock price Test for Level Stationary by P.P.Test

KPSS Test for Level Stationary data: dopen

KPSS Level = 0.162, Truncation lag parameter = 3, p-value = 0.1

Output4.5 The first difference of open stock price Test for Level Stationary by KPSS Test

From figure 4.2 and output (4.4 & 4.5) we be sure that data stationary. Now we are ready to identify the model by using Box Jenkins methodology and automated model identification methods MINIC, SCAN, ESACF, after that we see what is the method's more effective of our data .

```
Use identify options to identify a good model
The ARIMA Procedure     Number of Observations     240
Autocorrelations
Lag  Covariance  Correlation  -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1   Std Error
 0   48.912801    1.00000  |                   |********************|   0
 1   45.280360    0.92574  |                 . |******************* |   0.064550
 2   42.701904    0.87302  |                 . |*****************   |   0.106340
 3   40.723274    0.83257  |                 . |*****************   |   0.132889
 4   38.635805    0.78989  |                 . |****************    |   0.153088
 5   36.706969    0.75046  |                 . |***************     |   0.169220
 6   33.015010    0.67498  |                 . |*************       |   0.182561
 7   30.552453    0.62463  |                 . |************        |   0.192679
 8   28.589317    0.58450  |                 . |************        |   0.200939
 9   26.289028    0.53747  |                 . |***********         |   0.207903
10   23.671988    0.48396  |                 . |**********          |   0.213614
11   20.728489    0.42378  |                 . |********.           |   0.218135
12   17.798635    0.36389  |                 . |******* .           |   0.221538
Name of Variable = dopen,  Mean of Working Series  94.58425,  Standard Deviation  6.993769
```

0utput4.6 Autocorrelations of the first difference of open stock price

```
                Inverse Autocorrelations

Lag   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

  1     -0.52291    |        **********| .                |
  2      0.13626    |                . |***               |
  3     -0.17664    |             ****| .                 |
  4      0.20444    |                . |****              |
  5     -0.37193    |          *******| .                 |
  6      0.35334    |                . |*******           |
  7     -0.14836    |              ***| .                 |
  8      0.10618    |                . |**.               |
  9     -0.13196    |              ***| .                 |
 10      0.12448    |                . |**.               |
 11     -0.21773    |             ****| .                 |
 12      0.23439    |                . |*****             |
```

output4.7 Inverse Autocorrelations the first difference of open stock price

```
                Partial Autocorrelations

Lag   Correlation   -1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1

  1      0.92574    |                . |******************* |
  2      0.11211    |                . |**.               |
  3      0.07933    |                . |**.               |
  4     -0.01014    |                . | .                |
  5      0.00769    |                . | .                |
  6     -0.27195    |            *****| .                 |
  7      0.07217    |                . |* .               |
  8      0.03732    |                . |* .               |
  9     -0.02627    |               . *| .                |
 10     -0.08397    |               .**| .                |
 11     -0.04289    |               . *| .                |
 12     -0.12498    |               .**| .                |
 13      0.01759    |                . | .                |
```

output4.8 Partial Autocorrelations the first difference of open stock price

Autocorrelation Check for White Noise									
To	Chi-		Pr >						
Lag	Square	DF	ChiSq	--------------------Autocorrelations--------------------					
6	969.95	6	<.0001	0.926	0.873	0.833	0.790	0.750	0.675
12	1363.78	12	<.0001	0.625	0.584	0.537	0.484	0.424	0.364
18	1423.84	18	<.0001	0.316	0.255	0.196	0.141	0.095	0.038
24	1435.94	24	<.0001	-0.014	-0.051	-0.076	-0.092	-0.111	-0.127

output4.9 Autocorrelation Check for White Noise

From output(4.6,4.7,4.8,4,9) we have some difficulty to determine the model order, Using the Box Jenkins method, we find it difficult to determine the order of the time series model, we need much of experience, and this is why this study is to use automatic methods, to determine model order in the time series, and this saves time and effort, especially in mixed (ARIMA) models . the order of the model.

So we go to **MINIC method**

Minimum Information Criterion						
Lags	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	3.619117	3.505762	3.41645	3.350818	3.279869	3.177738
AR 1	1.126599	1.148019	1.159875	1.137514	1.155645	1.171963
AR 2	1.130832	1.127803	1.138443	1.148141	1.162437	1.182156
AR 3	1.114247	1.129921	1.150282	1.17096	1.185272	1.20479
AR 4	1.131696	1.145868	1.168542	1.190056	1.206551	1.22079
AR 5	1.146273	1.162017	1.18401	1.200913	1.212957	1.233932
Minimum Table Value: BIC(3,0) = 1.114247						

Output4.10 MINIC method (The ARIMA Procedure) of first difference of open stock price

From output 4.10 we can essay to see that the minimum BIC(3,0) that we have ARIMA(3,1,0) or we have ARI(3,1) since BIC(3,0) = 1.114247.

ESACF method

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12
0	x	0	0	0	0	0	0	0	0	0	0	0	0
1	x	0	0	0	0	0	0	0	0	0	0	0	0
2	x	x	0	0	0	0	0	0	0	0	0	0	0
3	x	0	0	0	0	0	0	0	0	0	0	0	0
4	x	x	0	x	0	0	0	0	0	0	0	0	0
5	x	x	0	0	0	0	0	0	0	0	0	0	0
6	0	x	x	0	x	0	0	0	0	0	0	0	0

Output4.11 esacf method (The ARIMA Procedure) of of first difference of open stock price of Palestinin banking sector

From output 4.11 we can see the ESACF result recommend two models ARIMA(1,1,1)&ARIMA(2,1,1) with BIC = 1.148019 , 1.127803 respectively

SCAN method

ARMA(p+d,q) Tentative Order Selection Tests

---------SCAN--------				--------ESACF--------		
p+d	q	BIC		p+d	q	BIC
1	0	1.126599		1	1	1.148019
				2	1	1.127803
(5% Significance Level)						

Output4.12 scan method (The ARIMA Procedure) of dopen price Palestinin banking sector

From output 4.12 we can see the SCAN result recommend one model ARIMA(1,1,0) With BIC = 1.126599 . Then after using automate methods we have four models are :

Identification method	ARIMA models	BIC
MINIC	ARIMA(3,1,0)	1.114247
SCAN	ARIMA(1,1,0)	1.126599
ESACF	ARIMA(2,1,1)	1.127803
ESACF	ARIMA(1,1,1)	1.148019

Table 4.2 The proposed ARIMA models from the identification methods

Initial state of methods

From table 4.2 we see the MINIC result recommend ARIMA(3,1,0) will it have the minimum BIC(3,0) = 1.114247, but ether of result of the methods is not recommend this model and SCAN method result recommend ARIMA(1,1,0) well has BIC = 1.126599 will ESACF and MINIC methods not recommend this model and as we see too, ESACF result recommend two model exclusive are ARIMA(1,1,1) & ARIMA(2,1,1) with BIC = 1.148019 , 1.127803 respectively and both of models not appear in other methods so in our data if we used one method of former methods we have model and we don't know about many models may be it models are more efficiency, so we want know to spotlight more on four models by make diagnostic model and after that we return to chapter3 and compare our result with the comparison in section (3.5). Before comparing the suggested models, we need to test the series mean, so we have to apply the one sample T test to verify if the series mean equals zero or not, the result shows that the p-value is larger than 0.05, so the test is significant.

One Sample t-test
data: dopen
t = 0.5857, df = 238, p-value = 0.5586
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval: -0.2089339 0.3857540
sample estimates: mean of x 0.08841004

Output4.13 One Sample t-test result of first difference of open stock price

Based on that result we don't reject the null hypothesis and we find that the series mean equal zero, so we don't need to estimate the mean of the series. Now, we will examine the models and the spotlight more on four models ,we will estimate parameters for the four models using "SAS program", then We will compare some of the properties of four models.

The next is the results are an application for the previous suggested four models, respectively.

ARIMA(3,1,0)

Estimate parameter , ARIMA(3.1.0)

Model Parameter	Estimate	Std Error	T	Prob>lTl
Intercept	0.08593	0.1257	0.6837	0.4948
Autoregressive, Lag 1	-0.15069	0.0652	-2.3096	0.0218
Autoregressive, Lag 2	-0.06356	0.0658	-0.9657	0.0528
Autoregressive, Lag 3	0.01983	0.0652	0.3040	0.0614
Model Variance (sigma squared)			5.37743.	

Model Diagnostic

Now we will apply diagnostic measures to validate the ARIMA(3,1,0) model, the result of diagnostic is a set of three graphs. that shown in figure (4.3) is a plot of standardized residuals Second part is an autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics.

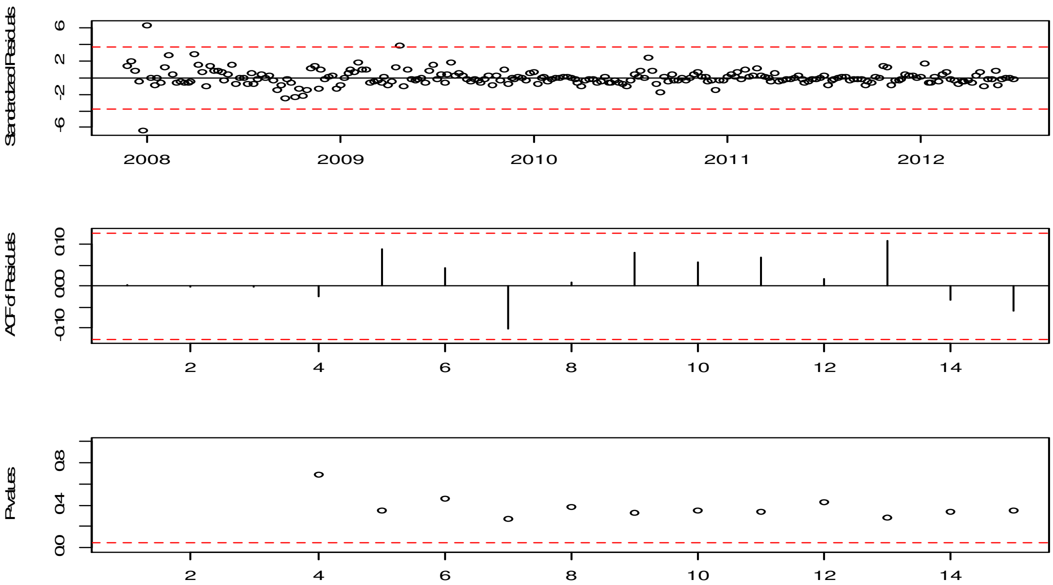

figure (4.3) A plot of standardized residuals and autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics for ARIMA(3,1,0)

The standardized residuals don't show clusters of volatility and seem to be fairly "random" with no particular patterns. Second part is an autocorrelation (ACF) plot of residuals. The plot of ACF shows that there is no significant autocorrelation between the residuals, to be sure about this we use LB test, which is a function modifies the Box.test function in the R program and it computes the Ljung-Box or Box-Pierce tests check whether or not the residuals are white noise, R program result below shows the result of LB test.

```
Box-Ljung test
data:  residuals from  a
X-squared = 8.9854, df = 8, p-value = 0.3435
```

Output4.14: Autocorrelation between the residuals ARIMA(3.1.0)

This test does not reject the randomness of the error terms based on the first eleven autocorrelations of the residuals. The p-values for the Ljung–Box statistics all are large, which indicate that the residuals are patternless, figure (4.3) shows the p-value vs. lag. Until now, these are basic diagnostics, but we need for additional check normality for the residuals, the residual figure (4.4) below abbreviated that.

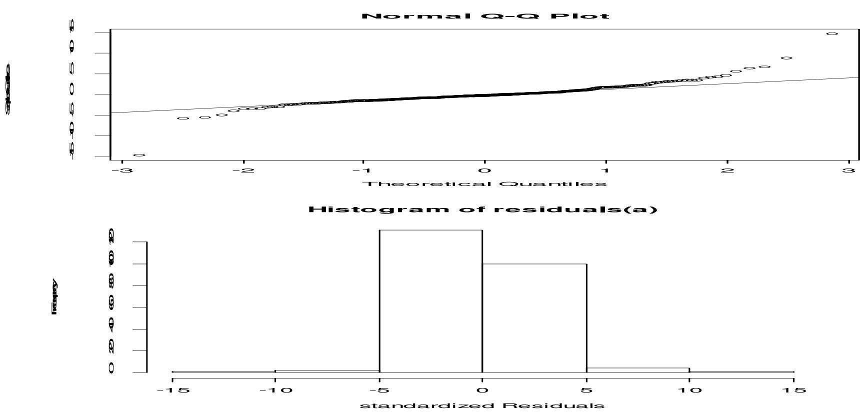

Figure (4.4) Check normality for the residuals ARIMA(3.1.0)

With a few minor exceptions in the lower tail, the Q-Q plot of the standardized residuals seems to be normal. shows that the residuals for ARIMA model are normally distributed because the values is on and near normal line. We have a large number of residuals, so by central limit theorem the residuals have normal distribution and from the figure above (normality probability plot and histogram) we can check our results. Normality By Shapiro test:

Shapiro-Wilk normality test data: residuals(a) W = 0.8303, p-value < 0.001

Output4.15: Normality autocorrelation between the residuals ARIMA(3.1.0)

Output 4.15 shows the normality of the residuals of ARIMA model is holding because p-value is less than 0.05. When we check ARIMA(3,1,0) model, we can see that its good model to forecasting .

ARIMA(1,1,0)

Estimate parameter, ARIMA(1,1,0)

Model Parameter	Estimate	Std. Error	T	Prob>ǀTǀ
Intercept	0.08691	0.1312	0.6625	0.5083
Autoregressive, Lag 1	-0.14183	0.0643	-2.2059	0.0284
	Model Variance (sigma squared)		5.35793	

Model Diagnostic

Now we will apply diagnostic measures to validate the ARIMA(1,1,0) model, the result of diagnostic is a set of three graphs. that shown in figure (4.5) is a plot of standardized residuals Second part is an autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics.

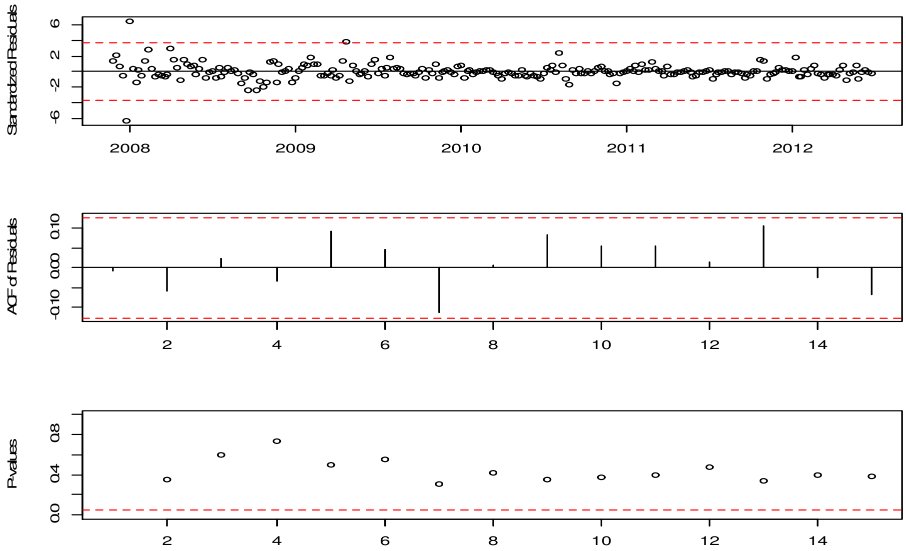

Figure (4.5) Aplot of standardized residuals and autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics

The standardized residuals don't show clusters of volatility and seem to be fairly "random" with no particular patterns. Second part is an autocorrelation (ACF) plot of residuals. The plot of ACF shows that there is no significant autocorrelation between the residuals, to be sure about this we use LB test, which is a function modifies the Box.test function in the R program and it computes the Ljung-Box or Box-Pierce tests check whether or not the residuals are white noise, R program result below shows the result of LB test.

```
Box-Ljung test
data:  residuals from  a
X-squared = 10.5477, df = 10, p-value = 0.3938
```

Output4.16: Autocorrelation between the residuals ARIMA(1.1.0)

This test does not reject the randomness of the error terms based on the first eleven autocorrelations of the residuals. The p-values for the Ljung–Box statistics all are large, which indicate that the residuals are pattern less, figure (4.5) shows the p-value vs. lag. Until now, these are basic diagnostics, but we need for additional check normality for the residuals, the residual figure (4.6) below abbreviated that.

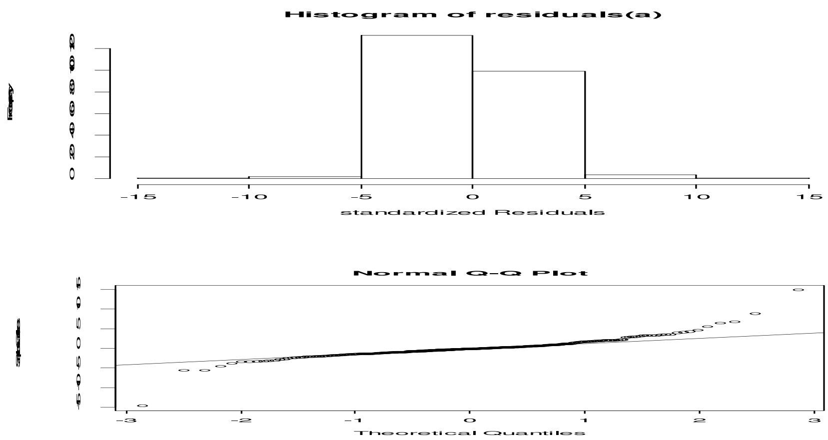

Figure (4.6) Check normality for the residuals ARIMA(1.10)

With a few minor exceptions in the lower tail, the Q-Q plot of the standardized residuals seems to be normal. shows that the residuals for ARIMA model are normally distributed because the values is on and near normal line. We have a large number of residuals, so by central limit theorem the residuals have normal distribution and from the figure above (normality probability plot and histogram) we can check our results by Shapiro test:

Shapiro-Wilk normality test data: residuals(a) W = 0.8305, p-value < 0.0019

Output4.17: Normality autocorrelation between the residuals ARIMA(1.1.0)

Output 4.17 shows the normality of the residuals of ARIMA model is holding because p-value is less than 0.05. When we check ARIMA(1,1,0) model, we can see that its good.

ARIMA(2,1,1)

Estimate parameter, ARIMA(2,1,1)

Model Parameter	Estimate	Std .Error	T	Prob>lTl
Intercept	0.08527	0.1264	0.6747	0.5005
Moving Average, Lag 1	-0.61091	0.5030	-1.2146	0.0582
Autoregressive, Lag 1	-0.75945	0.4990	-1.5218	0.0948
Autoregressive, Lag 2	-0.15198	0.0760	-1.9997	0.0467

Model Variance (sigma squared) 5.36783

Model Diagnostic

Now we will apply diagnostic measures to validate the ARIMA(2,1,1) model, the result of diagnostic is a set of three graphs. that shown in figure (4.7) is a plot of standardized residuals Second part is an autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics.

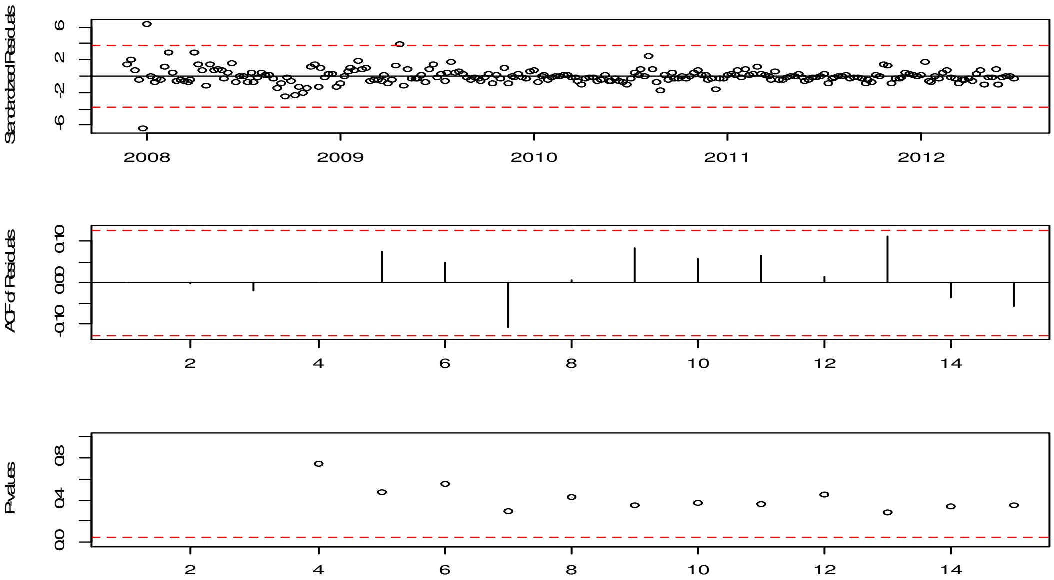

figure (4.7) A plot of standardized residuals and autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics

The standardized residuals don't show clusters of volatility and seem to be fairly "random" with no particular patterns. Second part is an autocorrelation (ACF) plot of residuals. The plot of ACF shows that there is no significant autocorrelation between the residuals, to be sure about this we use LB test, which is a function modifies the Box.test function in the R program and it computes the Ljung-Box or Box-Pierce tests check whether or not the residuals are white noise, R program result below shows the result of LB test.

```
Box-Ljung test
data:  residuals from  a
X-squared = 8.7369, df = 8, p-value = 0.365
```

Output4.18 : Autocorrelation between the residuals ARIMA(2.1.1)

This test does not reject the randomness of the error terms based on the first eleven autocorrelations of the residuals. The p-values for the Ljung–Box statistics all are large, which indicate that the residuals are pattern less, figure (4.7) shows the p-value vs. lag. Until now, these are basic diagnostics, but we need for additional check normality for the residuals, the residual figure (4.8) below abbreviated that

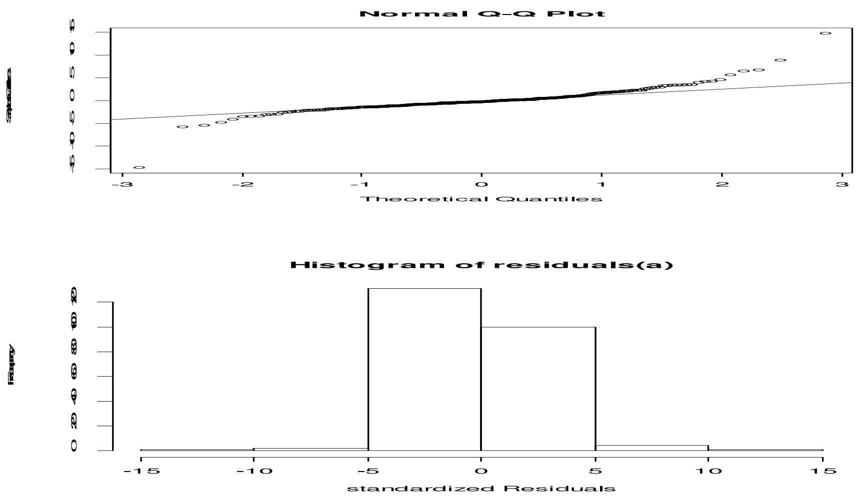

Figure (4.8) Check normality for the residuals ARIMA(2.1.1)

With a few minor exceptions in the lower tail, the Q-Q plot of the standardized residuals seems to be normal. shows that the residuals for ARIMA model are normally distributed because the values is on and near normal line. We have a large number of residuals, so by central limit theorem the residuals have normal distribution and from the figure above (normality probability plot and histogram) we can check our results by Shapiro test:

Shapiro-Wilk normality test data: residuals(a) W = 0.828, p-value < 0.0014

Output4.19: Normality autocorrelation between the residuals ARIMA(2.1.1)

Output 4.19 shows the normality of the residuals of ARIMA model is holding because p-value is less than 0.05. When we check ARIMA(2,1,1) model, we can see that its good.

ARIMA(1,1,1)

Estimate parameter , ARIMA(1,1,1)

Model Parameter	Estimate	Std .Error	T	Prob>lTl
Intercept	0.08529	0.1227	0.6952	0.4876
Moving Average, Lag1	0.31049	0.3839	0.8088	0.0514
Autoregressive, Lag1	0.15720	0.3988	0.3942	0.0693

Model Variance (sigma squared) 536291

Model Diagnostic

Now we will apply diagnostic measures to validate the ARIMA(1,1,1) model, the result of diagnostic is a set of three graphs. that shown in figure (4.9) is a plot of standardized residuals Second part is an autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics.

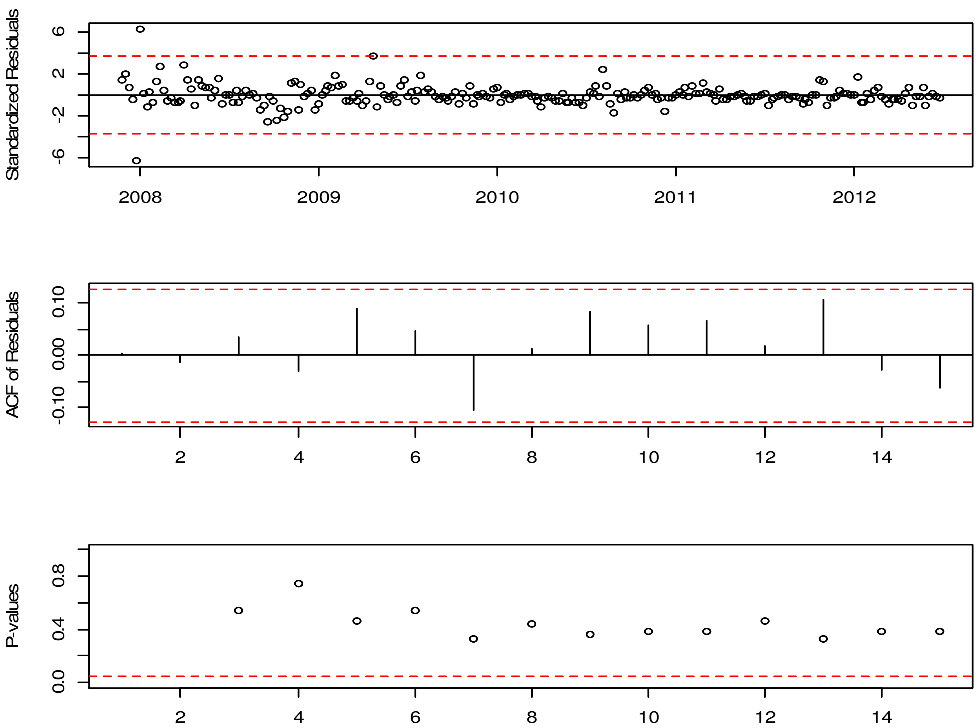

figure (4.9) A plot of standardized residuals and autocorrelation (ACF) plot of residuals, and p-values plot for the Ljung–Box Chi-squared statistics

The standardized residuals don't show clusters of volatility and seem to be fairly "random" with no particular patterns. Second part is an autocorrelation (ACF) plot of residuals. The plot of ACF shows that there is no significant autocorrelation between the residuals, to be sure about this we use LB test, which is a function modifies the Box.test function in the R program and it computes the Ljung-Box or Box-Pierce tests check whether or not the residuals are white noise, R program result below shows the result of LB test.

```
Box-Ljung test
data:  residuals from  a
X-squared = 9.6251, df = 9, p-value = 0.3817
```

Output4.20: autocorrelation between the residuals ARIMA(1.1.1)

This test does not reject the randomness of the error terms based on the first eleven autocorrelations of the residuals. The p-values for the Ljung–Box statistics all are large, which indicate that the residuals are pattern less, figure (4.9) shows the p-value vs. lag. Until now, these are basic diagnostics, but we need for additional check normality for the residuals, the residual figure (4.10) below abbreviated that

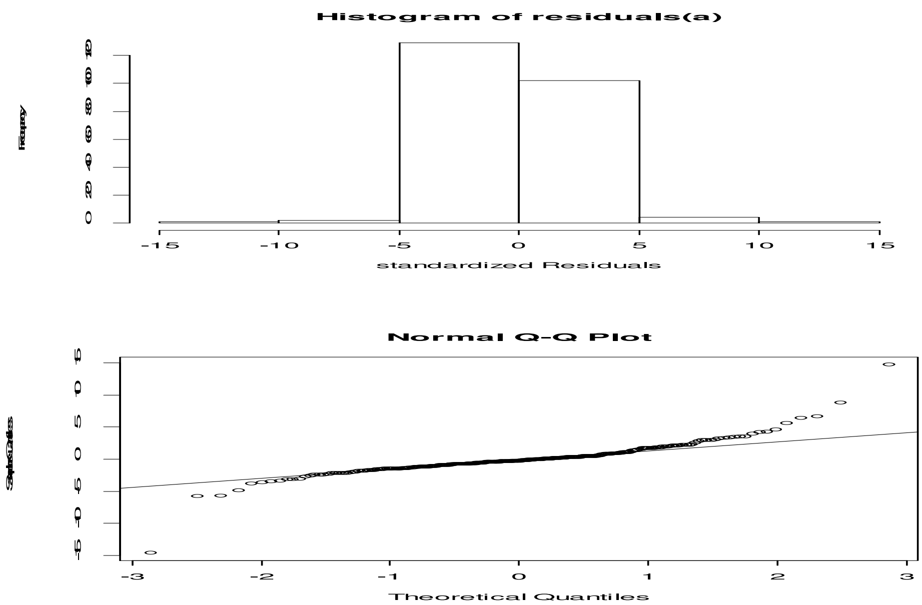

Figure (4.10) Check normality for the residuals ARIMA(1.1.1)

With a few minor exceptions in the lower tail, the Q-Q plot of the standardized residuals seems to be normal. shows that the residuals for ARIMA model are normally distributed because the values is on and near normal line. We have a large number of residuals, so by central limit theorem the residuals have normal distribution and from the figure above (normality probability plot and histogram) we can check our results by Shapiro test:

Shapiro-Wilk normality test data: residuals(a) W = 0.8335, p-value < 0.002

Output4.21 Normality autocorrelation between the residuals ARIMA(1.1.1)

Output 4.21 shows the normality of the residuals of ARIMA model is holding because p-value is higher than 0.05. When we check ARIMA(1,1,1) model, we can see that its good. When we check ARIMA(3,1,0) model, we can see that its good. The model can be written as:

ARIMA(3.1.0) $$r_t = \emptyset_1 r_{t-1} + \emptyset_2 r_{t-2} + \emptyset_3 r_{t-3} + \varepsilon_t$$

The constant term is not included in the model because the t test is not significant as we saw that in the result below .

Model.coef	Value	Test statistic	P-value
$\hat{\emptyset}_1$	-0.14894	-2.2878	0.0230
$\hat{\emptyset}_2$	-0.06194	-0.9434	0.0528
$\hat{\emptyset}_3$	0.02118	0.3255	0.0614
$\hat{\mu}$	0.08593	0.6837	0.4948

Table 4.3 ARIMA(3,1,0) model parameter estimation

When we check ARIMA(1,1,0) model, we can see that its good. The model can be written as:

ARIMA(1.1.0) $$r_t = \emptyset_1 r_{t-1} + \varepsilon_t$$

The constant term is not included in the model because the t test is not significant as we saw that in the result below.

Model.coef	Value	Test statistic	P-value
$\hat{\emptyset}_1$	-0.14183	-2.2059	0.0284
$\hat{\mu}$	0.08691	0.6625	0.5083

Table 4.4 ARIMA(1,1,0) model parameter estimation

When we check ARIMA(2,1,1) model, we can see that its good. The model can be written as:

ARIMA(2,1,1) $r_t = \emptyset_1 r_{t-1} + \emptyset_2 r_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$

The constant term is not included in the model because the t test is not significant as we saw that in the result below.

Model.coef	Value	Test statistic	P-value
$\hat{\theta}_1$	-0.61091	-1.2146	0.0582
$\hat{\emptyset}_1$	-0.75945	-1.5218	0.0948
$\hat{\emptyset}_2$	-0.15198	-1.9997	0.0467
$\hat{\mu}$	0.08527	0.6747	0.5005

Table 4.5 ARIMA(2,1,1) model parameter estimation

When we check ARIMA(1,1,1) model, we can see that its good. The model can be written as:

ARIMA(1,1,1) $r_t = \emptyset_1 r_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1}$

The constant term is not included in the model because the t test is not significant as we saw that in the result below.

Model.coef	Value	Test statistic	P-value
$\hat{\theta}_1$	0.31049	0.8088	0.0514
$\hat{\emptyset}_1$	0.15720	0.3942	0.0693
$\hat{\mu}$	0.08529	0.6952	0.4876

Table 4.6 ARIMA(1,1,1) model parameter estimation

Now we have four model are passed fitness tests we want to see the forecasting result of the models

4.5 Forecast Evaluation

In this section, we will use measures of forecast accuracy in chapter two, to compare the four models above and see which one give the best forecast, table (4.7) below contains the result for measures of forecast accuracy (ME, RMSE, MAE, MPE, MAPE) and table (4.8) contains the result of production values From these results, we can deduce that:

Methods	Models	AIC	BIC	ME	RMSE	MAE	MPE	MAPE
SCAN	ARIMA(1,1,0)	401.65623	405.13269	0.10102	2.30735	1.42456	0.07818	1.57758
ESACF	ARIMA(1,1,1)	402.92985	409.88277	0.10778	2.30385	1.4301	0.08456	1.58252
ESACF	ARIMA(2,1,1)	404.16022	414.58961	0.10481	2.30014	1.42103	0.08194	1.57172
MINIC	ARIMA(3,1,0)	404.56195	414.99134	0.10546	2.30208	1.42676	0.08246	1.57841

Table 4.7 ARIMA(p.d.q) models forecast evaluation

Methods/medal		Actual value	Predicted value	Prediction error
SCAN ARIMA(1,1,0)	1	94.5400	96.5819	-2.0419
	2	94.6500	94.8658	-0.2158
	3	95.0600	94.6346	0.4254
	4	94.9300	95.0024	-0.0724
	5	94.6000	94.9483	-0.3483
ESACF ARIMA(1,1,1)	1	94.5400	96.5708	-2.0308
	2	94.6500	94.8038	-0.1538
	3	95.0600	94.7133	0.3467
	4	94.9300	95.0171	-0.0871
	5	94.6000	94.9367	-0.3367
ESACF ARIMA(2,1,1)	1	94.5400	96.6447	-2.1047
	2	94.6500	94.7146	-0.0646
	3	95.0600	94.8774	0.1826
	4	94.9300	94.8399	0.0901
	5	94.6000	95.0251	-0.4251
MINIC ARIMA(3,1,0)	1	94.5400	96.5682	-2.0282
	2	94.6500	94.7618	-0.1118
	3	95.0600	94.8193	0.2407
	4	94.9300	94.9430	-0.0130
	5	94.6000	94.9263	-0.3263

Table 4.8 prediction values and predicted error of ARIMA(p.d.q)

From table 4.7 and table 4.8 we find :

- **From SCAN** method, model ARIMA(1.1.0)

SCAN method proposed order selection of the models was ARIMA(1.1.0) with

$$\hat{r}_t = -0.14183 r_{t-1}$$

the result for measures of forecast accuracy (AIC = 401.65623 ,BIC=405.13269, ME=0.10102 , RMSE=2.30735 , MAE = 1.42456 , MPE=0.07818 , MAPE=1.57758.

has less AIC & BIC 401.65623 & 405.13269 respectively and it has the less MPA = 0.07818 and less ME = 0.10102 but from table 4.8 SCAN method approximately has second Prediction error.

- **From the MINIC** method ARIMA (3.1.0)

  MINIC method recommends order selection of the models was ARIMA(3.1.0) with Minimum Table Value: BIC(3,0)=1.114247

$$\hat{r}_t = -0.14894\, r_{t-1} + -0.06194 r_{t-2} + 0.02118 r_{t-3}$$

  the result for measures of forecast accuracy (AIC = 404.56195, BIC=414.99134, ME= 0.10546, RMSE=2.30208 , MAE = 1.42676, MPE=0.08246, MAPE=1.57841), it has Minimum information Criterion Value: BIC(3,0) = 1.114247 but it has largest AIC & BIC 404.56195 & 414.99134 respectively and has approximately the least Prediction error.

- **ESACF method** ARIMA(1.1.1)&ARIMA(2.1.1)

ESACF method proposed order selection of the models was ARIMA(1.1.1)&ARIMA(2.1.1) with

ARIMA(1,1,1) $\hat{r}_t = 0.15720 r_{t-1} - 0.31049\, \varepsilon_{t-1}$

ARIMA(2,1,1) $\hat{r}_t = -0.75945 r_{t-1} + -0.15198 r_{t-2} + 0.61091 \varepsilon_{t-1}$

the result for measures of forecast accuracy for ARIMA(1.1.1) is (AIC = 402.92985 ,BIC=409.88277, ME=0.10778, RMSE=2.30735, MAE = 1.42456 , MPE=0.08456, MAPE=1.58252). And for ARIMA(2.1.1) is (AIC = 404.16022 ,BIC=414.58961, ME=0.10481, RMSE=2.30014, MAE = 1.42103, MPE=0.08194, MAPE=1.57172).

has good Prediction error, and has less RMSE = 2.30014 and second MPE & MAE after SCAN method.

According to previous result:

SCAN method has performs the better model with better measures of forecast accuracy.

The model with less AIC value may be not give a least Prediction error (better forecasting). It was clearly in the table(4.8), where as the ARIMA(1,1,0) has least AIC and has not least Prediction error while ARIMA (3,1,0) has big AIC and has least Prediction error.

4.6 Summary

In the present chapter, we tried to show the usefulness, effectiveness, of identification methods by using actual data and a forecasting accuracy of the models was recommended by applying them on the data of opening stock price of the banking sector in the Palestinian stock market. SCAN method has performs the better model with better measures of forecast accuracy.

Chapter five

Conclusion and Recommendations

5.1 Conclusion

Model identification method is the crucial step in time series modeling, The data which we have, is weekly observations of opening stock price of the banking sector in Palestine and its range from December 2007 to September 2012 . Thus we have about (240) observation. In this study, we discussed methods of identification models, classical method and automated method (SCAN, ESACF and MINIC), we find SCAN method has less AIC & BIC 401.65623 & 405.13269 respectively and it has the less MPA = 0.07818 and less ME = 0.10102 but from table 4.8 SCAN method approximately has approximately less Prediction error and we find MINIC method has Minimum information Criterion Value: BIC(3,0) = 1.114247 but it has largest AIC & BIC 404.56195 & 414.99134 respectively and has the least Prediction error, for the ESACF has good Prediction error, and has less RMSE = 2.30014 and second MPE & MAE after SCAN method.

From our study we concluded that:

1. Box – Jenkins methodology or methods (only used ACF,PACF) does not give a clearly model identification .
2. Automated identification methods in general, is very helpful for identifying the current model specially pattern methods .
3. In ARIMA pure structures, MINIC and SCAN perform well for autoregressive models, whereas ESACF works better in moving-average cases.
4. SCAN and ESACF are superior to MINIC for mixed (p,q) models.

For the case study, the MINIC, SCAN, and ESACF results strongly supported the recommendation of Box et al. (1994) to use automated methods as supplementary guidelines in the model selection process and not as a substitute for critical examination of the ACF, PACF and model residuals.

Recommendations:

We recommend the researchers to :

- Find new methods for identifying the models order .
- Make more comparisons among the identification methods.

References

[1] Beguin, J. M., Gourieroux, C. and Monfort, A, (1980), "Identification of a Mixed Autoregressive Moving Average Process: The Corner Method" , in Time Series, O. D . Anderson, Ed., NorthHolland, Amsterdam, 423-436.

[2] Box, G. E. P., Jenkins, G. M. and Reinsel, G. C.(1994), "Time Series Analysis, Forecasting and Control", 3rd ed., Prentice Hall, Englewood Cliffs, 1994

[3] Box, G.E.P. & Jenkins, G.M., (1970), "Time series Analysis Forecasting and Control", Holden-Day, San Francisco.

[4] Box, G.E.P. & Pierce, D.A. (1970), "Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. Journal of the American Statistical Association, Vol. 65 No. 332, 1509-26.

[5] Box, G.E.P., Jenkins, G.M., & Reinsel, G.C., (1994), "Time series Analysis Forecasting and Control", 3rd ed., Prentice Hall, Englewood Gliffs, NJ., 89- 99, 224- 247.

[6] Byoung Choi, (1992), "ARMA model identification", Springer-Verlag Berlin, 200 pp., DM 81.00, ISBN 3-540-97795-3

[7] Chan, Wai-Sum (1999), "A comparison of some of pattern identification methods for order determination of mixed ARMA models. " *Statistics & Probability Letters* vol. 42 issue 1 March 15, 1999. p. 69-79

[8] Chatfield, C. (2004), "The Analysis of Time Series: An Introduction", 6th edition, Chapman & Hall/CRC.

[9] Chatfield, C.(2000), "Time series forecasting", Chapman & Hall/CRC, Boca Raton, Florida, (2001).

[10] Chen, Z. & Yang, Y., (2004), "Assessing Forecast Accuracy Measures", Lowa Stata University, Department of Statistics(2004-10).

[11] Choi, B., (1992), "ARMA Model Identification", Springer Verlag, New York.

[12] Cochrane, J. H., (2005), "Time Series for Macroeconomics and Finance", university of Chicago Spring 1997.

[13] Corradi, V. & Swanson, N. R., (2006), "Predictive density evaluation. in 'Handbook of Economic Forecasting", North-Holland, Amsterdam, The Netherlands.

[14] Cryer, Jonathan D., Chan, Kung-Sik (2008),"Time Series Analysis with application in R ", 2nd edition , Springer Texts in Statistics (2008)

[15] Hamad A. Al-Ghannam (2003), "Time Series Analysis of Saudi Arabian Stock Market Index", Journal of King Abdulaziz University (2003)

[16] Hyndman, R. J. & Koehler, A. B., (2006), "Another look at measures of forecast Accuracy". International Journal of Forecasting, Elsevier, vol. 22(4), pages 679-688.

[17] Liitkepohl, H., (1985), "Comparison of Criteria for Estimating the Order of a Vector Autoregressive Process" , Journal of Time Series Analysis, 6, 35-52.

[18] Ljung, G. M. & Box G. E. P (1978), "On a measure of lack of fit in time series models", ' Biometrika 65(2), 297–303. Magee, L. (1985)

[19] Samreen Abu Radi (2009)," Analysis of the status of shares related to the Banking Sector in Amman Stock Exchange, through the use of Time Series Analysis Model", Middle East Technical University (2009).

[20] Shittu. O.I & Asemota. M.J (2009), "Comparison of Criteria for Estimating the Order of Autoregressive Process: A Monte Carlo Approach", European Journal of Scientific Research ISSN 1450-216X Vol.30 No.3 (2009), pp.409-416

[21] Velicer, W.F., and Harrop, J. (1983), The Reliability and Accuracy of Time Series Model Identification, Evaluation Review, 7, 551-560.

[22] Wei, W.W.S., (2006), "Time Series Analysis Univariate and Multivariate Methods", 2nd edition Pearson Education, Inc.

[23] Yuhong Yang & Hui Zou(2002), "Combining Time Series Models for Forecasting", Lowa State University (2002-15).

[24] . Chan, W. S., (1999) "A comparison of some pattern identification methods for order determination of mixed ARMA models", Statistics and Probability Letters, 42, 69–79.

[25] . Cryer, J. D. & Chan, K(2008), "Time Series Analysis With Applications in R" 2^{nd} edition, Springer Science + Business Media, LLC(2008).

[26] . Ojo, J. F. & Olatayo, T. O (2009),"On the Estimation and Performance of Subset", Autoregressive Integrated Moving Average Models", European Journal of Scientific Research ISSN 1450-216X Vol.28 No.2, pp.287-293 .