AAiT Addis Ababa Institute of Technology
Addis Ababa University

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

Seasonal Auto-Regression Integrated Moving Average-based Data Traffic Forecasting: The Case of UMTS Network in Addis Ababa, Ethiopia

By

Amel Salem

Advisor

Dr. - Ing. Dereje Hailemariam

A Thesis Submitted to the School of Graduate Studies of Addis Ababa

University in Partial Fulfillment of the Requirements for the Degree of

Masters of Science in Electrical Engineering

February, 2016

Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

Seasonal Auto-Regression Integrated Moving Average-based Data Traffic Forecasting: The Case of UMTS Network in Addis Ababa, Ethiopia

Amel Salem

Approval by Board of Examiners

__________________ ______________

Chairman, School Graduate Committee Signature

Committee

Dr. -Ing. Dereje Hailemariam ______________

Advisor Signature

__________________ ______________

Internal Examiner Signature

__________________ ______________

External Examiner Signature

AAiT Addis Ababa Institute of Technology
Addis Ababa University

Declaration

I, the undersigned, declare that this thesis is my original work, has not been presented for a degree in this or any other university, and all sources of materials used for the thesis have been fully acknowledged.

Amel Salem ____________________

Name Signature

Place: Addis Ababa

Date of Submission: ________________

This thesis has been submitted for examination with my approval as a university advisor.

Dr.-Ing. Dereje Hailemariam ______________

Advisor's Name Signature

Abstract

In planning, operating and developing mobile data networks, one crucial input is the telecommunication demand that includes number of subscribers and their required service data rates. These numbers should be predicted accurately for optimal planning and to capture the needs of the subscriber thereby creating customer satisfaction.

Ethio-telecom, the sole telecom service provider in Ethiopia, has recently introduced different service charging systems that include flat rate and package-based data services. This has increased the number of subscribers who are using these services, which in turn has led to a substantial increase in data traffic, and hence, a burden on the existing infrastructure. Such increases in demand should be considered in planning phases, where proper forecasting of the data demand growth is one integral input for the planning. Based on the available information, the current data growth forecast practice being employed by Ethio-telecom is mainly based on marketing information.

This thesis presents Seasonal Auto-Regression Integrated Moving Average (SARIMA) model as an alternative way of forecasting Universal Mobile Telecommunication System (UMTS) mobile data traffic taking the city of Addis Ababa as a case study. The approach in this thesis involves investigating the past UMTS data traffic load collected from the core network to find an appropriate model which describes the inherent structure of the UMTS data-traffic and forecast the future data traffic load. With this forecasting model, it is observed that the expected monthly data traffic per user for smart phones can reach up to 7GB as compared to the current 1GB cap. As the first practice (to the best of our knowledge) for data forecasting using available data in Ethio-telecom, it is hoped that the approach shown here will be useful for subsequent infrastructure expansion planning in a way that guarantees better customer satisfaction.

Key Words: ACF, forecasting, PACF, SARIMA, Seasonality, Trend, UMTS, UMTS Data-traffic.

Acknowledgment

I am grateful to Allah for the good health and well-being that were necessary to complete this thesis.

I would like to thank my advisor Dr. - Ing. Dereje Hailemariam for his patience, motivation and immense knowledge, his door was always open whenever I ran into a trouble spot or had a question.

I would like to acknowledge Ethio telecom staff members, especially Ato Fedlu Mohammed and Ato Fistum Mergia as they were very cooperative in providing the necessary data for this work and as they were very encouraging throughout the whole process.

Finally, I must express my very profound gratitude to my parents and to my friends for providing me with unfailing support and continuous encouragement throughout my year of study and through the process of researching and writing the thesis. This accomplishment would not have been possible without you. Thank you.

Table of Contents

List of Tables

List of Figures

Abbreviations

3G	Third Generations
3GPP	3rd Generation Partnership Project
ACF	Auto correlation Function
AR	Auto-Regressive model
ARMA	Auto-Regressive Moving Average
ARIMA	Auto-Regressive Integrated Moving Average
BSS	Base Station Subsystem
CN	Core Network
DL	Downlink
EDGE	Enhanced Data for GSM Evolution
ETA	Ethiopian Telecommunication Agency
ETSI	European Telecommunications Standards Institute
GPRS	General Packet Radio System
GSN	GPRS Support Node
HLR	Home Location Register
IMT	International Mobile Telecommunication
ITU	International Telecommunication Union
MA	Moving Average
MGW	Media Gateway

MS	Mobile Station
MSC	Mobile Switching Center
Node B	Radio Base Station
PACF	Partial Auto-Correlation Function
QoS	Quality of Service
RAN	Radio Access Network
S-GW	Serving Gateway
SARIMA	Seasonal Auto-Regression Moving-Average model
SGSN	Serving GPRS Support Node
SINR	Signal-to-Interference plus Noise Ratio
SISO	Single Input Single Output
UE	User Equipment
UL	Uplink
UMTS	Universal Mobile Telecommunication System
VLR	Visitor Location Register

1. Introduction

Mobile networks have evolved through four generations, starting with the analogue or first-generation (1G) networks deployed in the early 1980s, and moving on to the digital second-generation (2G) networks deployed in the early 1990s. Third generation (3G) networks were first deployed in 2001; and 3.5G networks around 2005; and 4G systems were first rolled out in the 2010 [1].

3G network, which is the core of this thesis, is based on a set of standards used for mobile devices, services and networks that comply with the International Mobile Telecommunications-2000 (IMT-2000) specifications set by the International Telecommunication Union (ITU). 3G is developed to support the effective delivery of multimedia services. In addition, it provides a more efficient system for the over-the-air transmission of existing services, such as voice, text and data [2].

The Universal Mobile Telecommunication System (UMTS) is a 3G mobile cellular system for networks that evolved from the 2G Global System for Mobile Communication (GSM) standard and is a component of the IMT-2000 standard set. UMTS offers a higher spectral efficiency and available bandwidth to mobile network operators and this is attributed to its use of Wideband Code Division Multiple Access (W-CDMA) radio access technology [1].

Ethio-telecom, the sole provider of telecommunication services in Ethiopia, is in charge of installing and operating network infrastructures to provide the needed services for its customers. 3G wireless service was introduced in 2013 in Addis Ababa and then extended to the other major cities of Ethiopia. 3G services are becoming more popular and the traffic over the network is growing rapidly. Continuing with its massive expansion plans, Ethio-telecom expects to have 53 million subscribers in a year time from now. According to Ethiopian Government's second Growth and Transformation Plan (GTP), this number is expected to rise to 103 million by the year 2020. This massive expansion plan will require efficient network planning, deployment and operations.

With the rapid growth of subscribers, wireless service providers, like Ethio-telecom, face the challenge of satisfying their customers' demands. Unlike voice service, which has a well understood traffic characteristics, data services vary greatly in their behavior and demand on network resource such as volume of data, signaling load and radio frequency (RF) air time. Data volume is likely to burst due to the increase of customer and degrade the ability of the network to support the data traffic. The ability to accurately predict the data traffic is critical to managing the data bursting [8].

In planning, developing and operating telecommunication networks, one crucial input is the telecommunication demand that includes the number of subscribers and their requested data rates. The traffic load must be predicted (forecasted) to properly design the future network infrastructure. Forecasting involves the process of making predictions of the future based on the past and the present analysis of data trends [10].

Over the last few years, forecasting has become a dynamic study area which has attracted attentions of research communities. The main aim of time series forecasting is to gather the past observations (data) to develop the best model that accurately fits the data and then generate the future data. Time series forecasting thus can be termed as the act of predicting the future by understanding the past time series data [12].

The current trend of forecasting network traffic being followed by Ethio-telecom, based on observations, is in marketing level. This method uses parameters such as subscriber growth, age of customers, technology being used, GTP targets and population growth rate to do the forecast. The major drawback of such method is that is does not exploit the information (e.g., trend in subscriber growth as well as data demand and usage) which is inherent in available previous data.

This thesis first intends to find the model that fits the UMTS data traffic of Addis Ababa city. For that, two alternative methods of forecasting, namely Auto-Regressive Moving Average (ARMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA) were considered. ARMA is a hybrid of Auto-Regressive model and Moving Average model which are used to forecast stationary time series, while SARIMA is a variant of the ARMA model that is used for non-stationary time series with seasonal component. Secondly, based on the results, SARIMA is selected to be the model which has a better accuracy for our data, and hence, used for subsequent forecasting.

1.1 Statement of the Problem

As we are embracing the age of technology, the electronic devices we use are becoming more interconnected and communicating in ways that were only fictitious previously. In this age, our smart phones are at the centre of this network of things. The recent trend in the smart phone world shows devices that are being shipped with many new applications like Viber and Facebook that consume the data bandwidth and can cause an unprecedented behaviour like delayed response and data bursting in a specific network [8]. Because the network demand is expanding at a quick pace, both in size and in the types of applications, especially with the growth of the Internet, the performance and dimensioning issue of networks are becoming more critical.

Nowadays in Addis Ababa, cellular data is being used heavily and experiencing a significant increase in data volume which might result in slower upload and download speeds. In the near future, this network infrastructure could face a significant increase in the number of subscribers with fancy smart phones, which could increase the Internet demand and cause a serious problem by congesting the existing network infrastructure.

The current forecasting method being employed by Ethio-telecom is focused on input from the marketing department, such as the number of subscribers. This can cause a serious deviation in forecast as it does not consider the information (e.g., trend in subscriber growth as well as data demand and usage) which is inherent in available previous data.

1.2 Objective

1.2.1 General Objective

The main objective of this thesis is to select model from ARMA and SARIMA that better fits the measured UMTS data traffic and use the selected model to forecast the future UMTS mobile data-traffic in Addis Ababa.

1.2.2 Specific Objectives

Specifically, the thesis aims to:

- Understand the basics of UMTS mobile network;

- Understand forecasting, types of forecasting and the different models of forecasting;
- Choose the model that fits to the UMTS data-traffic sample data;
- Implement the ARMA and SARIMA forecasting model using MATLAB tool,
- Forecast using the two models and choose the one with minimum prediction error,
- Calculate the traffic Usage in GB/month/user,
- Assist Ethio-telecom in planning their network infrastructure growth to accommodate future needs.

1.3 Literature Review

In the Ethiopian context, there is no prior work, to the best of our knowledge, which applies known forecasting techniques to understand the country's data growth. Hence, some related studies that were conducted abroad are surveyed.

Svoboda, etal. analysed different methods for long term forecasts of packet switched traffic from live 3G networks [6]. The dataset consists of over 400 values, each representing the peak load for a separate day. Four different methods were applied to forecast the increase in traffic, two simple (linear and exponential regression) and two more sophisticated (ARMA and Dynamic Harmonic Regression (DHR)) methods. They showed the cases in which the sophisticated models deliver better performance and also discussed whether the added benefit is significant enough to justify the increased complexity in the models. They presented different methods to predict the load increase in 3G live network. This thesis drew compelling evidence from [6]on the advantages of using the aforementioned sophisticated forecasting methods.

Stefan Valentin, etal. studied how combining econometric forecasting methods with methods from optimization and experimental design increased the accuracy of demand forecasts by orders of magnitude [8]. They investigated the economic, social, and technical factors behind a trend and points to interesting fundamental research to improve once understanding of telecommunication markets by using a “crystal ball” forecasting method. After surveying various methods (e.g., using diffusion or choice modelling), the authors highlighted that, unlike other fields, telecommunication completely lacks of forecasting methods that have been rigorously evaluated on their effectiveness

S. Makridakis, etal. examines that Box-Jenkins methodology to Auto-Regression Integrated Moving Average (ARIMA) models and determines the why in empirical test it is found that the post-sample forecasting accuracy of the model is worse than much simpler time series methods [9]. They concluded that the major problem is the way of making the data sample series stationary in its mean by differencing. From here it’s gained the idea of differencing in mean to make the data stationary when we faced the same challenge.

Haviluddinet, et al. presented a new approach for a network characterization by using ARIMA techniques [7]. The data sample that they used was taken from the Internet network traffic. They obtained result using Box-Jenkins Methodology and selected the best model and order that fit the sample. Some of the approaches from [7] were taken to fit our sample data and select the best one.

1.4 Methodology

Review on previous related studies is used for the indication of the importance and significance of forecasting models in Addis Ababa. The first phase of the research includes obtaining data for UMTS mobile data traffic from Ethio-telecom. The collected sample data was checked whether it was a time series or not, to make the forecasting in the next step. Then, the data was analyzed using the two different forecasting models and the better fitted model was simulated and forecasted in MATLAB. Finally, by using the forecasted data, the future Traffic usage in GB per month per user was calculated.

1.5 Scope and Limitation

1.5.1 Scope of the Thesis

This thesis is a case study addressing the importance of forecasting in network infrastructure capacity terms. The scope of this thesis is to forecast the sampled mobile UMTS data traffic using the best fitted model (SARIMA) comparing two different forecasting methods, ARMA and SARIMA and select the best one with minimum prediction error.

1.5.2 Limitation of the Thesis

This thesis has some limitation mainly due to inability to find related works in the Ethiopian context in forecasting area and the data concerning the two years data traffic for UMTS. In addition to this, the design of the core network or the existing network plan had to be calculated in other way because of confidentiality reasons.

1.6 Contribution

As indicated in the literature survey, no related work could be found that takes the case of forecasting the network demand of Ethio telecom's subscribers. Hence, to the best of our knowledge, this work is the first to analyse the forecasting of UMTS data traffic in more scientific ways and create awareness. This thesis selects from the available methodologies in forecasting UMTS data traffic to give a contribution in the area. Above all, it can be taken as a benchmark in the future related works.

1.7 Thesis Layout

The structure of the thesis is as follows. In Chapter one, as seen before an introductory overview, objectives, methodologies and scope of the thesis is given. This provides the necessary background and terminology for presenting the subsequent literature review. Chapter two presents brief description about UMTS network. Chapter three, the review chapter in turn describes the fundamentals of stochastic process and indicated the need of the moments and test of stationary. Chapter four explains about the different models that help forecasting time series and detail mathematical review about them. Chapter five deal with capacity planning of UMTS for existing 3G mobile networks in Addis Ababa. Chapter six gives summary of results, analysis, conclusion and recommendation for future work.

2. Universal Mobile Telecommunication System

2.1 Introduction

Mobile telecommunication system was first introduced as a commercial cellular system for the public in the late 20th century. As indicated in Chapter 1, according to their stages of development, cellular networks are classified into different generations, namely, 1G, 2G, 3G and 4G [20].

From the early 1G analog mobile to the last implemented 4G the system has been changing both in terms of services provided (from simple voice to data and video) and technological complexity. In the latest generation of networks communication is ubiquitous and they provide users with a new set of services. The growth of the number of mobile subscribers over the last years has also contributed to an evolution of mobile network. Table 2-1 summarizes the various generations along with the provided services in each generations.

3G radio networks are based on the CDMA (Code Division Multiple Access) technology and has deployed all over the world [20]. The aim of the technology is to fulfill the user requirement for innovative services such as multimedia messaging through high-speed data channels.

Types of mobile generations	Features	Data-rate (bps)	Service	Standard
1G	Analogue	2.4K, max 22K	Voice	AMPS
2G and 2.5	Digital	14.4-64K	Voice	GSM, GPRS, EDGE
3G	Digital	200M-1G (UMTS- 200-384K)	Integrated higher quality of voice, data and video	UMTS
4G	Digital	LTE :>100Mbps with adequate spectrum (15 or 20MHZ)	higher quality of voice, data and video	LTE
5G	Digital			

Table 2-1: Comparison of different mobile technologies [23].

CDMA is a digital cellular technology that uses spread-spectrum techniques. It does not assign a specific frequency to each user. Rather, every channel uses the full available spectrum. Individual conversations are encoded with a pseudo-random digital sequence [24].

UMTS is the European vision of 3G mobile communication systems. One of the key functionalities of UMTS is the ability to provide services anywhere and anytime. In UMTS the mobile equipment will be used for any possible purpose such as communication, entertainment, business and all kinds of services. With the increase in the number of smart phones having different applications setup, the data traffic demand is likely to be higher which pushes researchers and network operators to develop a new plan for the existing infrastructure of the network.

In the case of the city of Addis Ababa, Ethiopia, 3G and 4G systems have already been deployed to provide data services in addition to the 2G system which is predominantly used for voice services. The richness of features and functionalities with high quality of service in 3G is bringing people closer to the rest of world.

The following sections review the UMTS systems as a 3G platform for mobile communications and services. The sections provide a comprehensive overview of UMTS systems from evolution, architecture, protocols and service capabilities perspectives.

2.2 UMTS System Architecture

UMTS can, in many aspects, be looked upon as an extension to Global System for Mobile communication (GSM) and General Packet Radio Service (GPRS). The greatest changes are related to the access part of the network. The UMTS system consists of three main parts: the User Equipment (UE), the UMTS Terrestrial Radio Network (UTRAN) and the core network (CN).

2.2.1 User Equipment

User Equipment (UE) or terminal in UMTS corresponds to MS in GSM and is responsible for the communication functions needed on the radio interface.

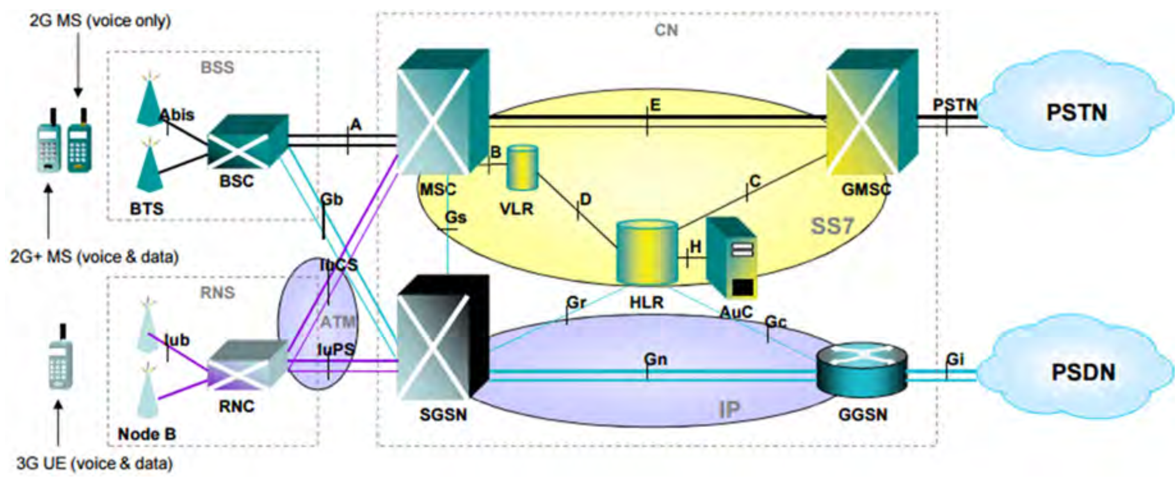


Figure 2-1: General UMTS Architecture [26].

The UE consists of two parts:

- The Mobile Equipment (ME): is the radio used for radio communication;
- The UMTS Subscriber Identity Module (USIM): is smartcard that holds the subscriber identity, performs authentication algorithms, and stores authentication and encryption keys and some subscription information that is needed at the terminal [19].

2.2.2 Core Network

The basic Core Network (CN) architecture for UMTS is based on GSM network with GPRS [21]. The main function of the CN is to provide:

- Switching and routing functionalities;
- Transit for user traffic;
- Contains the database and the network management functions.

The CN is divided into circuit switched and packet switched domains. Some of the circuit switched elements are Mobile Switching Center (MSC), Visitor Location Register

(VLR) and Gateway MSC. On the other hand, packet switched elements are Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). Some network elements like Home Location Register (HLR) and VLR are shared by both domains. Below is a brief description of the CN elements.

- **MSC:** is a central station that makes all the required functions of signaling and commutation for CS services from/to all the located MSs in a geographic area. It connects with UTRAN via the Iu-CS interface, with external networks such as Public Switched Telephone Network (PSTN) and Integrated Service Digital Network (ISDN) via the PSTN/ISDN interface, with HLR via the C interface, with MSC/VLR, GMSC or Short Message Center (SMC) via the E interface, with Serving GPRS Support Node (SGSN) via the Gs interface [26].
- **GMSC:** is the gateway node between the circuit domain of the mobile network and the external network, it's an optional functional node. It connects with external networks (PSTN, ISDN and other PLMN) through the PSTN/ISDN interface, connects with HLR through the C interface [25].
- **VLR:** The HLR has a data base and his function is to manage the mobile subscribers. The HLR stores the information of subscriptions and location data that allow to the appraisal and direction of call and message to the MSC/SGSN where the MS has been registered [27].
- **SGSN:** is responsible for the delivery of packets to/from the MSs within its service area and communications with the GGSN. It also keeps track of the mobiles within its service area.
- **GGSN:** acts as a logical interface to external packet data networks and maintains routing information used to tunnel Protocol Data unit (PDU) to the SSGN that is currently serving the MS.

- **HLR:** is a functional node shared by the CS and PS domains in the WCDMA core network. It connects with MSC/VLR or GMSC through the C interface, with SGSN through the Gr interface, and with GGSN through the Gc interface. And its main functions are to store subscription information for subscribers, support new services and provide the enhanced authentication function [25].

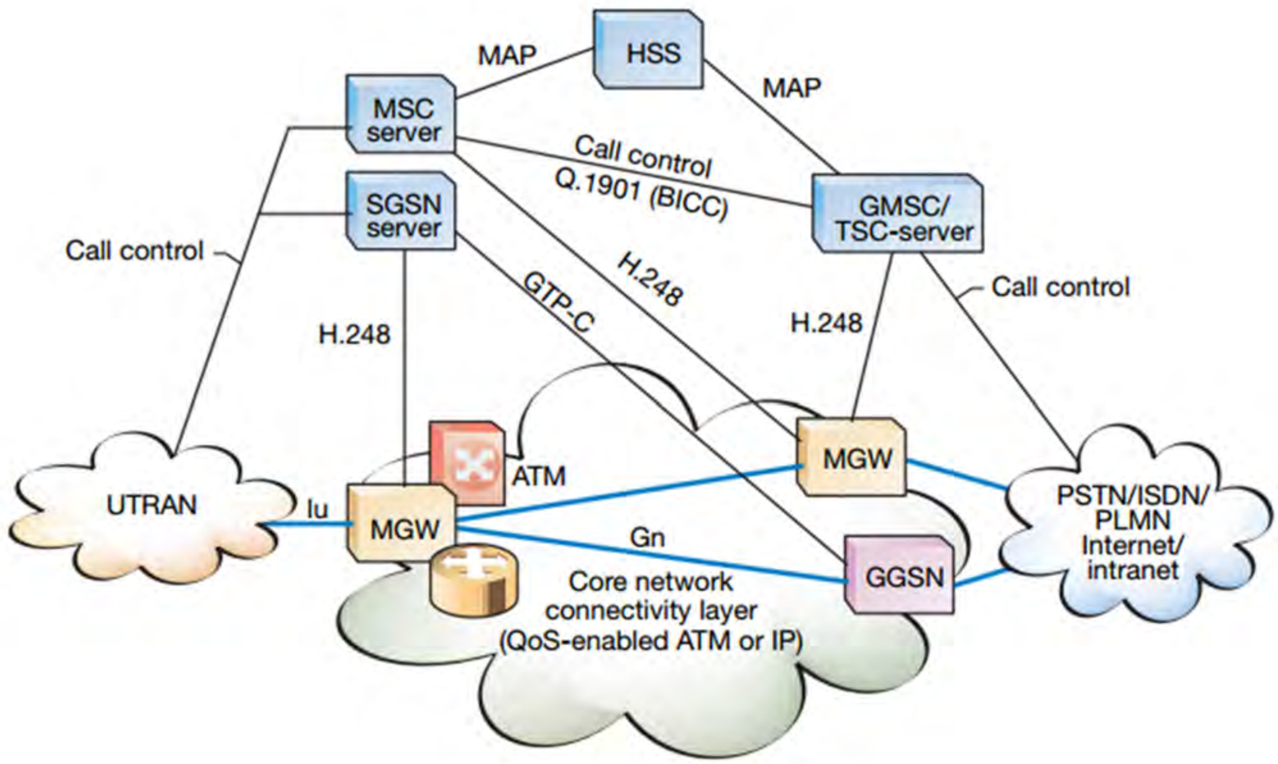


Figure 2-2: Core network architecture [25].

2.2.3 Basic Structure of UMTS's UTRAN

The access network, also called UTRAN, provides the air interface access method for user equipment [18]. Figure 2-2 below shows the structure of UTRAN.

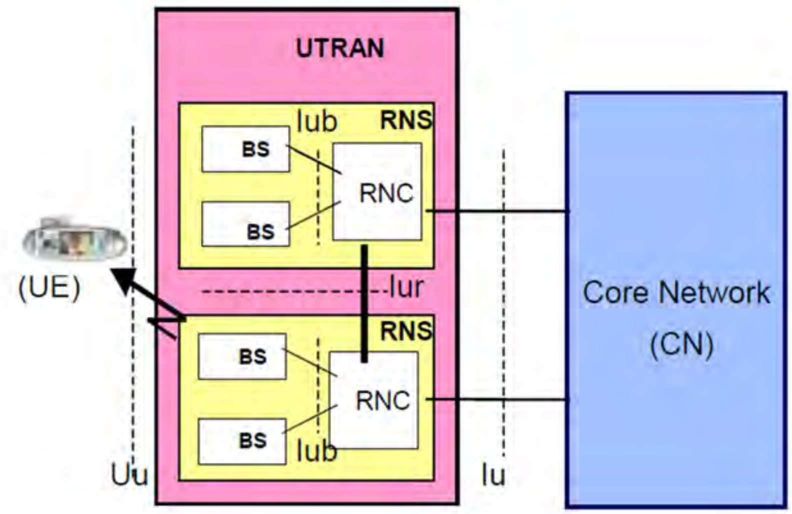


Figure 2-3: UTRAN Architecture [24].

UTRAN consists of base stations and base stations controllers.

- ✓ The base stations are called Node B. The functions of Node-B are:
  - Air interface transmission/reception;
  - Modulation/demodulation;
  - CDMA and physical channel coding;
  - Error Handing;
  - Closed loop power control.
- ✓ Several base stations are managed by a Radio Network Controller (RNC).The RNC is responsible for the Handover decisions that require signaling to the UE. The function of RNC are:
  - Radio Resource control;
  - Admission Control;
  - Channel Allocation;
  - Power Control setting;

- Handover control;
- Macro Diversity;
- Segmentation/Reassembly.

2.3 Basic Interfaces of UTRAN

The UMTS standards are structured so that internal functionality of the network elements is not specified in detail. Instead, the interfaces between the logical network elements have been defined. The following main open interfaces are specified:

- Iub interface: connects between an RNC and a Node B.
- Iur interface: is the logical interface between two RNCs.
- Iur represents a point-to point link between RNCs; however the physical realization may not be a point-to-point link.
- Iu interface: is the interconnection point between an RNC and the 3G Core Network (UTRAN).
- Cu interface: is the link between the USIM smartcard and the ME.
- Uu interface: is the interface through which the UE accesses the fixed part of the system.

2.4 Basic Protocol Structure of UTRAN Interfaces

UTRAN interface consists of a set of two layers. One is the separation of generic transport aspect from UMTS specific mobile networking aspect by horizontal layers, the other is the separation of network control aspect from the user data transfer aspect by vertical planes as shown in Figure 2-5.The UTRAN requirements are addressed in the horizontal radio network layer across different types of control and user planes.

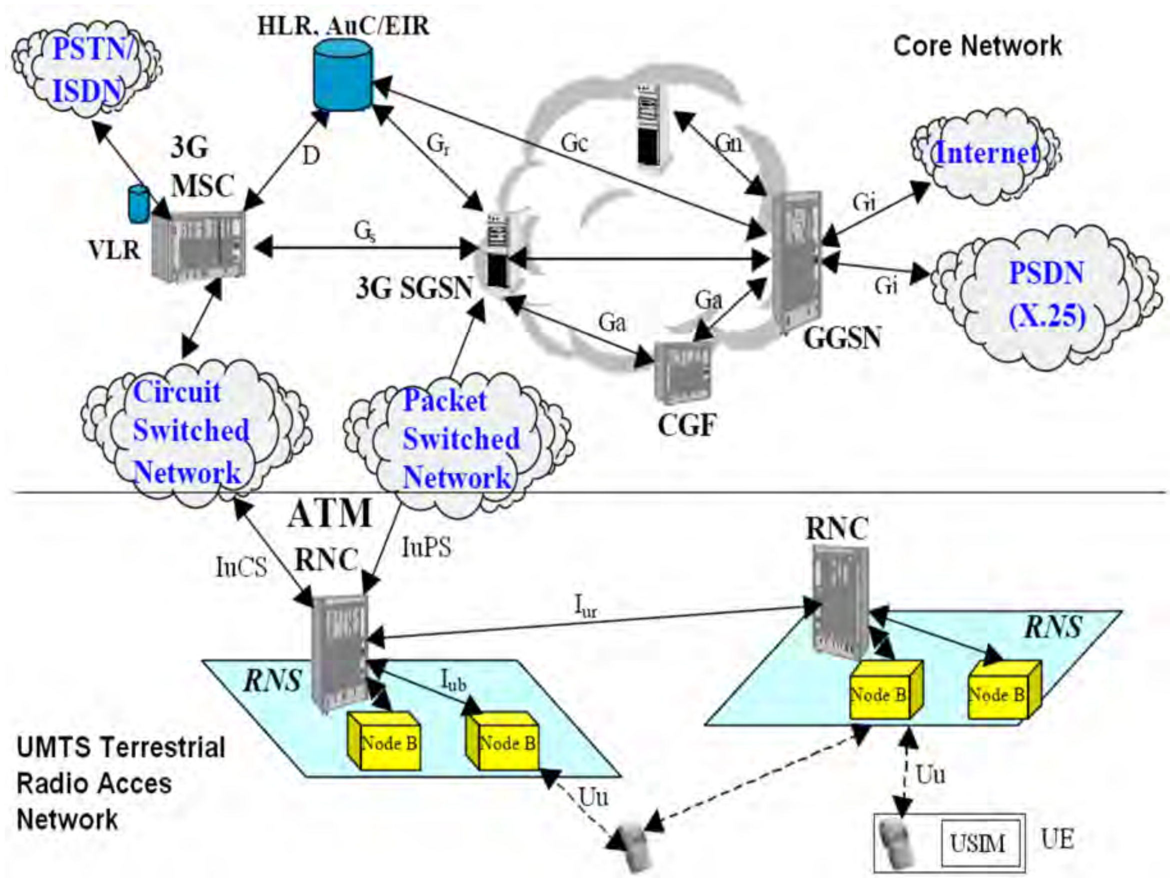


Figure 2-4: UTRAN interfaces [24].

Control Planes are used to control a link or a connection; User Planes are used to transparently transmit user data from the higher layer [24].

2.5 UMTS Radio Interface Technology

Both the UE and the UTRAN are composed of different layers. The four lowest layers are: The Physicallayer (PHY), the Medium Access Control (MAC), the RadioLink Control (RLC) and Radio Resource Control(RRC).

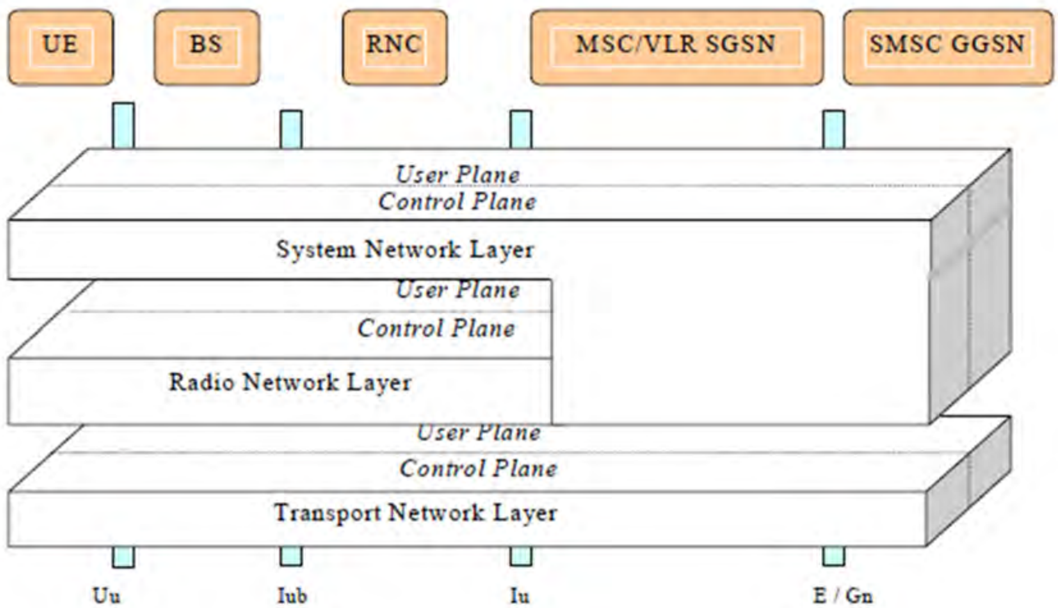


Figure 2-5: UMTS protocol architecture [24].

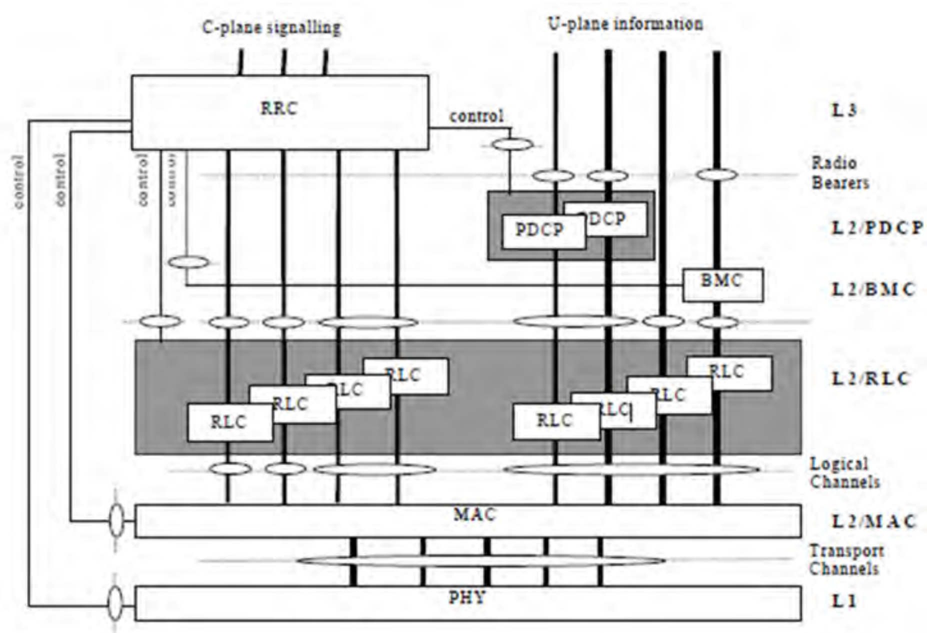


Figure 2-6: Basic radio interface protocol architecture [24].

- **The Physical layer**: takes care of coding, interleaving and the adding of Cyclic Redundancy Check (CRC) to the packets. Some of the features of the physical layers are:
  - ✓ Error detection on transport channels and indication to higher layers.
  - ✓ Encoding/decoding of transport channels.
  - ✓ Modulation and spreading/demodulation and dispreading of physical channels.
  - ✓ Frequency and time (chip, bit, slot, frame) synchronization.
  - ✓ Radio characteristics measurements and indication to higher layers.
  - ✓ Inner - loop power control.
  - ✓ Radio frequency processing.

  The Physical layer administrates all radio communication. It handles power control, modulation and measurements.

- **The MAC layer**: is responsible for the handling of the logic channels and most of the priority and multiplexing issues. The functions of MAC include:
  - ✓ Mapping between logical channels and transport channels.
  - ✓ Selection of appropriate Transport Format for each Transport Channel.
  - ✓ Priority handling between data flows of one UE.
  - ✓ Multiplexing/de-multiplexing of upper layer PDUs: into/from transport blocks delivered to/from the physical layer on common transport channels.
  - ✓ Traffic volume measurement.
  - ✓ Transport Channel type switching.
  - ✓ Ciphering for transparent mode RLC.

The MAC layer handles the timing of the packet releases and the adding of transport entity addresses on the outgoing traffic. The received traffic is sent to the corresponding transport entity via the MAC layer, which reads the address and removes it.

- **The RLC layer**: is the layer below the RRC in the protocol stack and it is focused on the actual data transfers. Below follows a sample of the functions:
  - ✓ Segmentation and reassembly,
  - ✓ Padding,
  - ✓ Error correction,
  - ✓ In-sequence delivery of upper layer Packet Data Units (PDUs),
  - ✓ Duplicate detection,
  - ✓ Flow control,
  - ✓ Sequence number check,
  - ✓ Protocol error detection and recovery,
  - ✓ Ciphering.

The RLC is responsible for retransmission, segmentation and reassembly. This layer contains the transport entities, which are created and deleted dynamically in pairs as services are established or released. One transport entity handles the incoming traffic and the other handles the outgoing traffic.

- **The RRC layer**: is the highest layer in the protocol stack and it handles most of the decisions and supervisory functions. Below follows a sample of the functions:
  - ✓ Broadcast of information.
  - ✓ Establishment, maintenance and release of an RRC connection between the UE and UTRAN.
  - ✓ Establishment, reconfiguration and release of Radio Bearers.

- ✓ Assignment, reconfiguration and release of radio resources for the RRC connection.
- ✓ RRC connection mobility functions.
- ✓ Control of requested Quality of Service.
- ✓ UE measurement reporting and control of the reporting.
- ✓ Outer loop power control.
- ✓ Control of ciphering.
- ✓ Paging,
- ✓ Initial cell selection and cell re-selection.
- ✓ RRC message integrity protection.

The RRC layer dynamically establishes and releases logical communication channels (transport entities), which is used by the various services in the UMTS network. It controls the parameters available, for example: bit rate, level of retransmission and coding scheme.

3. Fundamentals of Stochastic Process

This chapter presents an introduction to the branch of statistics known as time series analysis as well as fundamental concepts in the area of stochastic processes that are relevant for this thesis. Often the sample data we gather is collected sequentially over time–this type of data is known as *time series* data. Collecting data sequentially over time induces a correlation between measurements because observations near each other in time will tend to be more similar. However, the usual assumption is that observations are independent and stationary. Certain tests are applied to verify whether the collected sample data is stationary process or not.

The following sections present an introduction to time series models which include time series and their characteristics as well as stochastic process. Moreover, autocorrelation function (ACF), the partial autocorrelation function (PACF), Gaussian and white noise process and stationary process are defined so as to lay the foundation for the next chapter which is focused on forecasting.

3.1 Time series

The term time series refers to a set of t ordered data $x_1, x_{2,\dots,} x_t$, equally separated in time. A time series is a discrete set of numbers that may arise from instantaneous sampling of continuous process or when the original source is discrete in nature. The successive values of the time series must be separated by equal time intervals also called lags.

In the study of time series, a suitable model is fitted to the known data values (time series) and the corresponding model parameters are estimated from the data. This procedure of fitting a time series to a proper model is termed as *Time Series analysis*.

3.1.1 Types of time series

Time series can be classified based on two different ways; namely, based on modeling methods of time series and the other is by observing the phenomena of the time series.

Based on modeling methods time series can be grouped as:

a) **Univariate modeling method**: This is a type of modeling which generally uses only time as an input variable with no other outside explanatory variables. Some of the few employed methods are exponential smoothing, ARIMA and SARIMA [14].

b) **Multivariate modeling method:** When two or more independent variables are considered for the analysis. Variables may or may not be dependent on each other [14].

Based on the dynamic phenomenon that is observed a time series can be grouped into two classes:

a) The first are those that take stable values in time around a constant level, without showing a long term increasing or decreasing trend. These processes are called *stationary*.

b) A second class of processes is the *non-stationary processes*, which are those that can show trend, seasonality and other evolutionary effects over time [12].

In practice, the classification of a time series as stationary or non-stationary depends on the period of observation, since the series can be stationary in a short period and non-stationary in a longer one. Stationary process will be covered later in this chapter.

3.1.2 Components of time series

Time series are often examined to discover a historical pattern that can be exploited in the preparation of a forecast. In order to identify this pattern, it is often suitable to think of a time series as consisting of several components such as trend, seasonal variation and cyclic changes [10].

- A trend is the long-term change in the mean level and often thought of as the underlying growth or decline component in the series.
- The seasonal component is concerned with the periodic fluctuation in the series within each year. Seasonal fluctuations are most often attributed to social customs or weather changes.
- Cyclic changes within a time series are similar to the seasonal component in that it is revealed by a wavelike pattern. Cyclic changes can be thought of as variation across a fixed period due to some physical cause other than seasonal effects. Cycles are normally confined to a particular fixed period and can be a behavior that takes place over a period of years.

Once the trend and cyclic variations have been accounted for, the remaining movement is attributed to irregular fluctuations and the resulting data is a series of residuals. This set of residuals is not always random, so this series of residuals is also analyzed to determine if all the cyclic variation has truly been removed.

3.2 Stochastic Processes

A time series is non-deterministic in nature, but it can be assumed to follow certain probability model which describes the joint distribution of the random variable. The mathematical expression describing the probability structure of a time series is termed as *stochastic process* [28]. Thus, the sequence of observation of the series is actually a sample realization of the stochastic process that produced it.

Time series fall into the general field of stochastic processes which can be described as statistical phenomenon that evolves over time. Stochastic (random) process is a collection of random variables representing the evolution of some system of random values over time. In the simple case of discrete time, as opposed to continuous time, a stochastic process is a sequence of random variables.

The time series should be visualized as one possible realization of the stochastic process. One approach to visualize time series may be to model the random variables as random functions of one or several deterministic arguments (in most cases, the time parameter). Although the random values of a stochastic process at different times may be independent random variables, in most commonly considered situations they exhibit complicated statistical dependence [13].

The value of an observed time series are considered to be the outputs $x_n$of the model process whose input x_{n-k} is the past data also called the independent random shocks. The description is represented schematically in Figure3-1.

Figure 3-1: Time Series in Random shocks [17].

The random shocks or the inputs are regarded as disturbance of randomly varying magnitude, each independent of the preceding one, which enter the filter and transformed and combined into observed output values. The input or the sampled data occur at equally spaced intervals with the same index t (time) as the outputs.

For most statistical purpose, time series must be assumed that the sampled data (random shocks) are uncorrelated and normally distributed with zero mean and constant variance. A sequence possessing these properties will be referred to here as white *noise*, although white noise need not be normally distributed. A box or the filter transform the noise into an observed series. Checking the adequacy of the model used to describe the stochastic process consists of verifying that the model's residuals (observed minus model-calculated values) are white noise [10].

Mathematically:

$$\mathbf{x[n] - x[n]_e = e[n]} \qquad (3.1)$$

Where,

- $x(n)_e$-is the exact values,
- x(n)– is the forecasted data,
- e(n)-is the White noise.

Some important terms and behavior of stochastic process is mentioned in the below section.

3.3 First and Second-order Moments of Stochastic Process

An important part of the analysis of a time series is the selection of a suitable probability model (or class of models) for the data to allow for the possibly unpredictable nature of future observations. It is natural to suppose that each observation is a realized value of a certain random variable. A complete probabilistic time series model for the sequence of random variables would specify all of the joint distributions of the random vectors [11]. Such a specification is rarely used in time series analysis (unless the data are generated by some well-understood simple mechanism), since in general it will contain far too many parameters to be estimated from the sample data. Instead we specify only the first- order moments of the joint distributions, i.e., mean, variance and covariance.

3.3.1 Mean

Stochastic process can be defined via its first-order moments, i.e., $E(X_t)$, $E(X_t X_s)$ etc., where E(.) is the expectation operation. The sampled time series mean is an estimator of the mean from all the set from which the sample was taken [11].Mathematically, the mean function of the data traffic load process is given by:

$$\mathbf{E[x[n]] = \mu_x = \frac{1}{N}\sum_{i=1}^{N} x(n)} \qquad (3.2)$$

Where,

- x[n]-is the sample data,
- N-is the number of the sample data,
- n-is the number of times that value occurred (lags).

3.3.2 Variance

The other first–order parameter in probability theory is the variance of time series. It is the expectation of the squared deviation of a random variable from its mean, and it informally measures how far a set of (random) numbers are spread out from their mean [10].

Mathematically, the variance of the time series given by [10]:

$$\mathbf{Var[X[n]] = E[\boldsymbol{x}[\boldsymbol{n}]^2] = \boldsymbol{\alpha}_x^2 = x[n] - \mu_x{}^2} \qquad (3.3)$$

Where,

- x[n] - is the mean of the sample data,
- μ_x- is the mean of the sample data,
- n - is the number of times that value occurred (lags).

3.3.3 Covariance

The covariance of a time series is a measure of how much two random variables change together. If the greater values of one variable mainly correspond with the greater values of the other variable, and the same holds for the lesser values, i.e., the variables tend to show similar behavior, the covariance is positive. In the opposite case, when the greater values of one variable mainly correspond to the lesser values of the other, i.e., the variables tend to show opposite behavior, the covariance is negative. The sign of the covariance therefore shows the relationship between the variables. The magnitude of the covariance is not easy to interpret. The normalized version of the covariance, the correlation coefficient, however, shows by its magnitude the strength of the linear relation [11].Mathematically, the covariance function between the time series$X_s$and , X_k is given by:

$$\mathbf{Cov}(\mathbf{X}[\mathbf{n_s}],\mathbf{X}[\mathbf{n_k}]) = \mathbf{E}[(\mathbf{X}[\mathbf{n_s}] - \mu_\mathbf{s})(\mathbf{X}[\mathbf{n_k}] - \mu_\mathbf{k})] \qquad (3.4)$$

Where,

- $X[n_s]$ −is the sampled dataat s interval,
- $X[n_k]$ −is the sampled data at k interval,
- μ_s-is the mean of data at s interval,
- μ_k-is the mean of data at k interval.

The other relevant stochastic process properties are described in the below sections.

3.4 Correlation Functions

3.4.1 Autocorrelation Functions

One of the primary interests in studying time series is the extent to which successive terms in the series are correlated. In this thesis UMTS data traffic data set is considered and it seems reasonable to expect that the data traffic of next month will be correlated with the current month. However, will the data-traffic of five months from now depend in any way with the current months? In order to answer questions of this sort, we need to define autocorrelation function.

Autocorrelation of the time series is the correlation of a data (or time series) with itself at different points in time. It measures the similarity (correlation) between observations (information) at different times, as a function of the two times or of the time lag.

Autocorrelation is used to find repeating patterns, such as the presence of a periodic signal obscured by noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is often used in signal processing for analyzing functions or series of values, such as time domain signals [14].

Let X be time series repeatable process, and s be point in the interval (lag) after the start of that process. (It may be an integer for a discrete time process or a real number for a continuous time process).

Mathematically, the correlation function between X_s, X_k is given by;

$$\boldsymbol{\rho}(\mathbf{X_s}, \mathbf{X_k}) = \frac{\mathbf{Cov} \quad _{s},\mathbf{X_k})}{\sqrt{\alpha_s^2}\sqrt{\alpha_k^2}} \qquad (3.5)$$

Where,

- x_s −is the sampled at s interval,
- x_k −is the sampled at s interval,
- $\mathrm{Cov}(X_s, X_k)$-is the covariance of x_s and x_k ,
- α_s-isvariance of the UMTS data-traffic dataset at t time,
- α_k- is variance of the UMTS data-traffic dataset at k time.

As functions of k-lags, $\boldsymbol{\rho}(k)$ is called the autocorrelation function (ACF). They represent the covariance and correlation between $x(n)$ and $x(n + k)$ from the same process, separated only by k intervals.

3.4.2 Partial Autocorrelation Function

In time series analysis, the partial autocorrelation function (PACF) gives the partial correlation of a time series with its own lagged values, controlling for the values of the time series at all shorter lags. It contrasts with the autocorrelation function, which does not control for other intervals. Mathematically, the conditional for partial autocorrelation is given as:

$$\mathbf{Cor}\ (\boldsymbol{x_n}, \boldsymbol{x_{n+k}} \backslash \boldsymbol{x_{n+1}}, \boldsymbol{x_{n+k-1}}) \qquad (3.6)$$

Where "\" represents "without".

This function plays an important role in data analysis aimed at identifying the extent of the lag in an autoregressive model (will be discussed in the next chapter).

3.5 Joint Probability Distribution

Given at least two random variables x_t,x_k that are defined on a probability space, the joint probability distribution for them is a probability distribution that gives the probability that each of x_t, x_k falls in any particular range or discrete set of values specified for that variable

If X and Y are discrete, this distribution can be described with a joint probability mass function. If X and Y are continuous, this distribution can be described with a joint probability density function [34].

In general, if X_s and X_K are two random variables, the probability distribution that defines their simultaneous behavior is called a joint probability distribution.

3.6 Stationary Processes

The main goal of a time series analysis may be to understand seasonal changes and/or trends over time. However, another goal that is often of primary importance is to understand and model the correlational structure in the time series. This type of analysis is generally done on stationary processes [10].

This section covers stationary time series which are those whose statistical properties remain constant over time i.e., autocorrelation, power spectrum etc.

A random variable that is a time series is stationary if its statistical properties are all constant over time. A stationary series has no trend, its variations around its mean have constant amplitude, and it wiggles in a consistent fashion, i.e., its short-term random time patterns always look the same in a statistical sense. The latter condition means that its autocorrelations remain constant over time, or equivalently, that its power spectrum remains constant over time. The signal could be a pattern of fast or slow mean reversion, or sinusoidal oscillation, or rapid alternation in sign, and it could also have a seasonal component [7].

If $\{X_t\}$ is a strictly stationary process and $E(X_t^2) < \infty$, then the mean function is a constant and the variance function is also a constant.

Moreover, for a strictly stationary process with finite mean and variance, the covariance function, and the correlation function depend only on the time differences.

3.6.1 Test of stationary

The following are tests that can be utilized to check whether a time series is stationary.

- **Auto-correlation function plot**

Autocorrelation plots are a commonly used tool for checking randomness in a data set. This randomness is ascertained by computing autocorrelations for data values at varying time lags. If random, such autocorrelations should be near zero for any and all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero. The ACF plot is useful for identifying non-stationary time series. For a stationary time series, the ACF will drop to zero relatively quickly, while the ACF Of non-stationary data decreases slowly.

- **Unit root test**

A unit root test tests whether a time series variable is non-stationary and possesses a unit root. The null hypothesis is that the variable contains a unit root, and the alternative is that the variable and the alternative hypothesis is either; stationary, trend-stationary or explosive root depending on the test used.

After using the above different tests of stationary process, if the sample time series data set is not stationary there are different mechanisms to make it stationary to apply it into different forecasting model described in the next chapter. Generally, most of the time series are non-stationary and it must be removed if there is any non-stationary sources of variation i.e., trend or seasonality effects, and then fit a stationary model to the time series data. The below section shows how to make non- stationary process stationary.

3.6.2 Removing the Trends and Periodic Effects in a Time Series

In order to study a time series in greater detail, it is helpful to remove any trends and seasonal components from the sample data first [11]. There are a variety of ways this can be done. Consider the model:

$$\mathbf{x_t = \mu_t + e_t} \qquad (3.9)$$

Where,

- x_t -is the time series sample data set and;

- μ_t-is the trend and;
- e_t – is error.

Eliminating a Trend when there is no seasonality:

A time series that exhibits a trend is a non-stationary time series. Modeling and Forecasting of time series is greatly simplified if the trend can be eliminated. There are two methods of removing a trend from a data that doesn't have seasonality component [33, 34].

a) Smoothing by moving average: this process converts a time series into another time series by linear operation.

$$y_t = \frac{1}{2q+1}\sum_{i=0}^{q} x_{t+i} \qquad (3.10)$$

Where q-is the choice of the analyst for smoothing.

b) Differencing: another way of removing a trend in data is by differencing; that is applying the difference operator to the original time series to obtain a new time series [34]. The first difference operator ∇ defined by:

$$\nabla x_t = x_t - x_{t-1} \qquad (3.11)$$

Differencing has two advantages relative to fitting a trend model to the data. First it doesn't require estimation of any parameter. And second, model fitting assumes that trends is fixed throughout the time series history and will remain so in the future. In other words, the trend component, once estimated is assumed to be deterministic.

Differencing can allow the trend components to change through time. The first differencing is for a trend that impacts the change in the mean of the time series [34].

- **Eliminating a seasonal or periodic effect:**

The differencing method described above can be used to eliminate a seasonal effect in a time series as well. The seasonality components of a given time series can be eliminated by using a seasonal differencing such as ∇_s.

$$\nabla x_t = x_t - x_{t-s} \tag{3.12}$$

Where s is the seasonality components of the time series.

When both trend and seasonal components are simultaneously present, sequentially differencing can be used to eliminate these effects. Then, difference one or more using the regular difference operator to remove then seasonally difference to remove the seasonal components [34].

3.7 Gaussian Process

The above sections describe how time series is uncorrelated observation or not. If this time series observations have a constant variance, it is white noise (discussed in the next section).In addition, if the observation in this time series are normally distributed, the time series is Gaussian White Noise. Ideally forecast errors are Gaussian white noise. In a Gaussian process, every point in some continuous input space is associated with a distributed random. Moreover, every finite collection of those random variables has a univariate normal distribution. The distribution of a Gaussian process is the joint distribution of all those (infinitely many) random variables, and as such, it is a distribution over functions with a continuous domain (time).

A Gaussian process is stationary because the normal distribution is uniquely characterized by its first two moments (the mean and variance).

3.8 White Noise Processes

The simplest type of time series model is the white noise process, which is defined as a sequence of independent, identically distributed random variables. A time series x_t is said to be a white noise process w_t with variance σ^2 [5]. A process w_t is called white noise process, if the distribution of the sample autocorrelation coefficient at lag k in large sample is approximately normal with mean zero and constant variance of σ^2.

In discrete time, white noise is a discrete signal whose samples are regarded as a sequence of uncorrelated random variables with zero mean and finite variance; a single realization of white noise is a random shock. In particular, if each sample has a normal distribution with zero mean, the signal is said to be Gaussian white noise. The white noise is used as an input in the moving average model discussed in the next chapter.

4. FORECASTING MODELS for TIME SERIES

The parameters that are described in the previous chapters are the basis to understand modeling and forecasting time series data. Different techniques of modeling and forecasting are presented in this chapter.

4.1 Introduction to Forecasting

Forecasting is the process of making predictions of the future based on the past and the present and analysis of data trends. Forecasts empower people because their use implies that we can modify variables now to alter (or be prepared for) the future. Scholars have proposed different ways to categorize forecasting methodologies [15]. Three of them are presented below:

- **Genius Forecasting**: This method is based on a combination of intuition, insight, and luck. Psychics and crystal ball readers are the most extreme case of genius forecasting. Their forecasts are based exclusively on intuition. Science fiction writers have sometimes described new technologies with uncanny accuracy.
- **Trend extrapolation**: These methods examine trends and cycles in historical data, and then use mathematical techniques to extrapolate to the future. The assumption of all these techniques is that the forces

responsible for creating the past will continue to operate in the future. This is often a valid assumption when forecasting short term horizons, but it falls short when creating medium and long term forecasts. The further out we attempt to forecast, the less certain we become of the forecast.

- **Simulation methods** - Simulation methods involve using analogies to model complex systems. These analogies can take on several forms. It might be equation/formulas to predict the future and it is a mathematical analogy. Mathematical analogies are of particular importance to future research. They have been extremely successful in many forecasting applications. Nowadays prediction is done using these forecasting methods.

Forecasting starts with certain assumptions based on experience, knowledge, and judgment. These estimates are projected into the coming months or years using one or more techniques such as ARIMA models, moving averages, regression analysis, and trend projection (discussed in the following section).

In the statistical analysis of time series, ARMA model provide a parsimonious description of a (daily) stationary stochastic process in terms of two polynomials, one for the auto-regression and the second for the moving average. The detail description of this model will be seen in the following sections.

4.2 Auto-Regressive Moving Average Model

4.2.1 Introduction

ARMA model is a hybrid of two models, Auto Regression (AR) and Moving Average (MA).

4.2.2 Auto-Regression Model

In auto-regression (AR) model forecasting the future variable is predicted using linear combinations of past values. An AR model expresses a stochastic process of the data-traffic past values. The common notation is AR (p), where p is the order of the model.

The term "auto-regression" indicates that it is a regression of the variable itself. The order of the AR model tells how many lagged past values are included. Auto-regressive models are remarkably flexible at handling a wide range of different time series patterns [10].

4.2.2.1 Mathematical description

In the AR model, the value of the current output (observation)$\mathbf{x_n}$ (stochastic process) depends solely on "p" prior outputs and the current input (random shock) at $\mathbf{x_{n-k}}$.Mathematically the model is written as:

$$\mathbf{x(n) = \sum_{k=1}^{p} a_k x(n-k) + w(n)} \qquad (4.1)$$

Where,

- $x(n-k)$-is the previous samples,
- a_k- auto-regressive coefficients at Lag-k,

- w(n)- White noise with zero mean.

The name is appropriate, as the model involves regression a variable on previous values of itself, plus an error or random term. The error or the noise goes by various other names: the error, the random-shock, and the residual. The residuals are assumed to be random in time (not auto correlated), and normally distributed [8, 11, 34].

The simplest AR model is the first-order autoregressive, or AR (1) and it given by;

$$\mathbf{x(n) = \sum_{k=1}^{p=1} a_k x(n-k) + w(n) = a_1 x(n-1) + w(n)} \qquad (4.2)$$

Where

- $x(n-1)$-is the previous first sample,
- a_1-is theLag-1 autoregressive coefficients,
- w(n)- White noise with zero mean.

The first order AR (1) model has the form of a regression model in which x(n)is regressed on its Previous value. In this form, a_1 is analogous to the negative of the regression coefficient, andw(n) to the regression residuals.

Figure below show how to find the coefficients the AR model using the auto-correlation sequence r_k, when the value of b_k is zero.

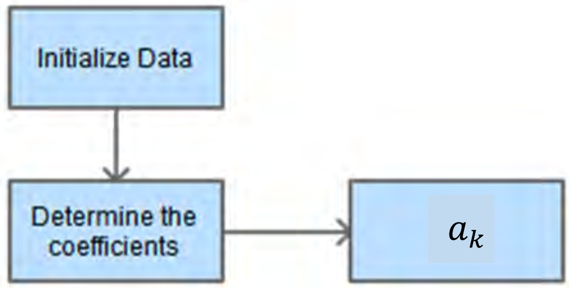


Figure 4-1: Flowchart to determine the Auto-regressivecoefficients (a_k).

4.2.3 Moving Average Model

Rather than using past values of the forecast variable in a regression, a moving average model uses past forecast errors in a regression-like model. The moving average (MA) model is model in which the sample process is regarded as a moving average (unevenly weighted) of a random sample process. The letter ‘’q” is used for the order of the moving average model.

4.2.3.1 Mathematical description

When the current output depends solely on the current input and "q" prior inputs, the model is given as:

$$\mathbf{x(n) = \sum_{k=0}^{q} b_k w(n-k) = x(n) = \sum_{k=1}^{q} b_k w(n-k) + b_o w(n)} \qquad (4.3)$$

Where,

- $w(n-k)$=white noise sampled at K interval,
- b_k-is the Moving average coefficients,
- w(n)- is the White noise with zero mean.

The above equation is called the Moving Average model of order q. The notation is MA (q).

- The first-order moving average, or MA (1) model is given by:

$$\mathbf{x(n)} = \sum_{\mathbf{k=0}}^{\mathbf{1}} \mathbf{b_k w(n-K)} = \mathbf{b_o w(n)} + \mathbf{b_1 w(n-1)} \qquad (4.4)$$

Where,

- $w(n-1)$-is the white noise sampled at 1,
- b_1 -is the Moving average coefficients,
- $w(n)$- is the White noise with zero mean.

Equation 4.4 is sampled process at time n and n-1, and is the first-order moving average coefficient.

The term "Moving Average" is misleading, since the model is not the familiar it's consisting of the arithmetic mean of past observations. However the term has become traditional.

The idea that an observation could be modeled as a linear weighted sum of random numbers is unfamiliar to most non statisticians. It is important to realize that a series composed of such linear sums of white noise elements is not itself white noise, but rather has a definite autocorrelation structure.

Figure 4-2, shows how to determine the coefficients of the Moving Average model using the auto-correlation sequence r_k and Substituting a_k to determine b_k.

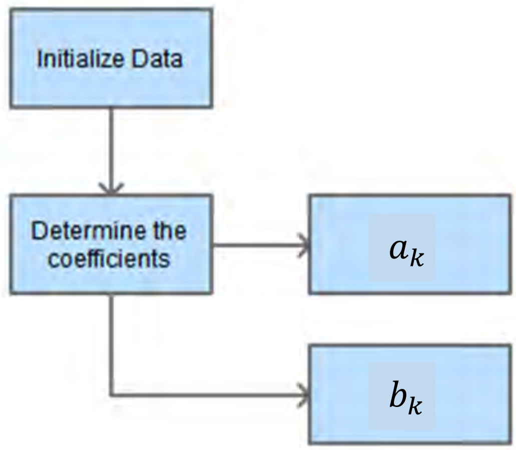


Figure 4-2: Flowchart to determine the Moving-Average coefficients (b_k).

ARMA is a well-known method used for forecasting time series. In this paper, the model is used to forecast the future time series based on past sample of 3G data traffic.

ARMA models are mathematical models of persistence, or autocorrelation, in a time series. It was introduced by Box and Jenkins (1976). ARMA models allow us not only to uncover the hidden patterns in the data but also to generate forecasts and predict a variable's future values from its past values [8].

ARMA models can also be used to predict behavior of a time series from past values alone. Such a prediction can be used as a baseline to evaluate the possible importance of other variables to the system. ARMA models are widely used for prediction of economic and industrial time series.

The autoregressive model includes lagged terms on the sample process itself, and that the moving average model includes lagged terms on the white noise.

By including both types of lagged terms, we arrive at what are called autoregressive-moving-average, or ARMA, models. The order of the ARMA model is included in parentheses as $\mathrm{ARMA}\,(p, q)$, where p is the autoregressive order and q the moving-average order.

Another use of ARMA models is simulation, in which synthetic series with the same persistence structure as an observed series can be generated. Simulations can be especially useful for established confidence intervals for statistics and estimated time series quantities.

4.2.4 Mathematical description

ARMA models are mathematical models of the persistence, or autocorrelation, in a time series. There are several possible reasons for fitting ARMA models to data, i.e. for interpolation purpose and to forecast.

The general $\mathrm{ARMA}(p, q)$equation is given by;

$$\mathbf{x(n) = -\sum_{k=1}^{p} a_k x(n-k) + \sum_{k=0}^{q} b_k w(n-k)} \quad (4.5)$$

Where,

- $x(n-k)$ -is the previous sample at k,
- a_k-is the Lag-k autoregressive coefficients p-order of auto-regressive model,
- q-is the order of Moving-Average model,
- $w(n-k)$-is the white noise sampled at K,
- b_k-is the Moving average coefficients.

From (4.5), the output undergoes regression over p of its previous values and at the same time a moving average on $W(n), W(n+1), \dots, W(n-q)$ of the input over $(q+1)$values is added to it, thus generating an ARMA(p,q) process $x(n)$.

The first-order autoregressive and first moving average orARMA (1,1) ;

$$\mathbf{x(n) = -\sum_{k=1}^{1} a_k x(n-k) + \sum_{k=0}^{1} b_k w(n-k)}$$

$$\mathbf{= a_k x(n-1) + b_o w(n) + b_1 w(n-1)} \quad (4.6)$$

Where,

- $x(n-1)$ –is the previous first sample,
- a_1-is the Lag-1 autoregressive coefficients,
- $w(n-1)$-is the white noise sampled at t=1,
- b_1-is the Moving average coefficients,
- $w(n)$- White noise with zero mean.

In processing ARMA model, the autoregressive and moving average coefficients must be determined to apply the model. In order to do that using the z-transform of the samples time series will help to get to values of the coefficients.

The z-transform of the sampled time series process $x(n)$ will be given as:

$$\mathbf{x(n) \xrightarrow{z-transfor} x(z)} \quad (4.7)$$

$$\mathbf{x(z) = -\sum_{k=1}^{p} a_k z^{-k} x(z) + \sum_{k=0}^{q} b_k z^{-k} w(z)} \quad (4.8)$$

Rearranging (4.8) as,

$$\mathbf{x(z) + \sum_{k=1}^{p} a_k z^{-k} x(z) = \sum_{k=0}^{q} b_k z^{-k} w(z)} \quad (4.9)$$

$$\mathbf{x(z)\left[1 + \sum_{k=1}^{p} a_k z^{-k}\right] = \sum_{k=0}^{q} b_k z^{-k} w(z)} \quad (4.10)$$

$$\mathbf{x(z)} = \frac{\sum_{k=0}^{q} b_k z^{-k} w(z)}{\left[1+\sum_{k=1}^{p} a_k z^{-k}\right]} = \left[\frac{b_{o+}\sum_{k=1}^{q} b_k z^{-k}}{\left[1+\sum_{k=1}^{p} a_k z^{-k}\right]}\right] \mathbf{w(z)} \qquad (4.11)$$

The inputs are regarded as disturbances of randomly varying magnitude (lag), each data-traffic load is independent of the preceding one, which enters the box and are transformed and combined into observed output values.

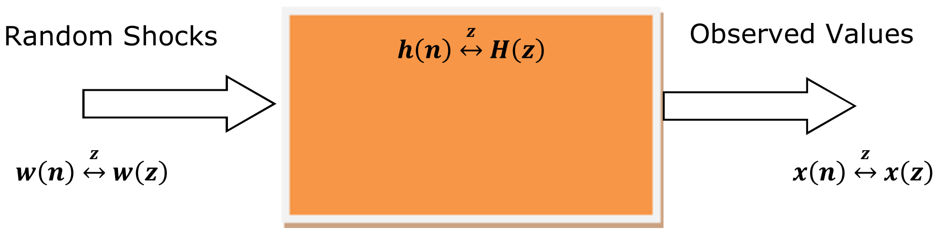


Figure 4-3: UMTS Data-traffic Time Series in white noise [15].

Figure 4-4 is given by:

$$\mathbf{x(z) = H(z)W(z)} \qquad (4.12)$$

The function is similar with the LTI (Linear Time invariant) function,

$$\mathbf{H}(z) = \frac{w(z)}{x(z)} = \frac{\sum_{k=0}^{q} b_k z^{-k} w(z)}{\left[1+\sum_{k=1}^{p} a_k z^{-k}\right]} = \left[\frac{b_{o+}\sum_{k=1}^{q} b_k z^{-k}}{\left[1+\sum_{k=1}^{p} a_k z^{-k}\right]}\right] \qquad (4.13)$$

$$\mathbf{H(z)} = \sum_{k=0}^{\infty} \mathbf{h(k)} z^{-k} = \frac{b_0+b_1+z^{-1}+\cdots+b_q z^{-q}}{1+a_1 z^{-1}+\cdots+a_p z^{-p}} \qquad (4.14)$$

Where: $\mathbf{a_0 = 1}$

Notice that the transfer function $\mathbf{H(z)}$ in (4.13) is rational with p poles and q zeros that determine the model order of the underlying system. The number of the poles should be greater than the number of zeros for the transfer functions $\mathbf{H(z)}$ to be stable.

The number of poles and zeros for (4.14) is p and q respectively. (4.13) can be written as;

$$\mathbf{H(z)} = \frac{\mathbf{w(z)}}{\mathbf{x(z)}} = \frac{\mathbf{b_o + b_1 + z^{-1} + \cdots + b_q z^{-q}}}{\prod_{\mathbf{k=1}}^{\mathbf{p}}(\mathbf{1-\alpha_k z^{-k}})} \qquad (4.15)$$

$$\mathbf{H(z)} = \frac{\mathbf{w(z)}}{\mathbf{x(z)}} = \sum_{\mathbf{k=1}}^{\mathbf{p}} \frac{\mathbf{A_k}}{\mathbf{1-\alpha_k z^{-k}}} \qquad (4.16)$$

Taking the inverse transform of (4.16);

$$\mathbf{h(n)} = \begin{cases} \sum_{\mathbf{n=0}}^{\mathbf{p}} \mathbf{A_k \alpha_k^n} & \mathbf{, n \geq 0} \\ \mathbf{0} & \mathbf{, n \leq 0} \end{cases} \qquad (4.17)$$

The above equation is LTI Impulse response.

The Cross-correlation for the sample process is given by;

$$\mathbf{R_{wx}(k) = R_{ww}(k) * h^*(-k)} \qquad (4.18)$$

The auto correlation of $\mathbf{x(n)}$ is given by;

$$\mathbf{R_{xx}(k) = R_{wx}(k) * h(k)} \qquad (4.19)$$

Substituting Equation 4.18 in Equation 4.19, The Auto correlation will be;

$$\mathbf{R_{xx}(k) = R_{ww}(k) * h(k) * h^*(-k)} \qquad (4.20)$$

The power spectrum of the output x (k) is be given by;

$$\mathbf{S_{xx}(w) = S_{ww}(w) * |H(jw)|^2} \qquad (4.21)$$

Where; $\mathbf{z = e^{jw}}$

The above formulas generally represent**ARMA** $\mathbf{(p, q)}$ process.

From the above equations there are two cases that needed to be considered;

- **AUTOREGRESSIVE MODEL**: When q is zero, x (n) ~AR(P)Case, the output ARMA(p, 0) is given by;

$$\mathbf{x(n)} = \sum_{\mathbf{k=1}}^{\mathbf{p}} \mathbf{a_k x(n-k) + w(n)} \qquad (4.22)$$

- The transfer function of AR (p) model is given by,

$$\mathbf{H(z)} = \frac{\mathbf{w(z)}}{\mathbf{x(z)}} = \frac{\mathbf{b_o}}{\mathbf{1+\sum_{k=1}^{p} a_k x(n-k)}} \qquad (4.23)$$

$$\mathbf{H_{AR}(z)} = \frac{\mathbf{w(z)}}{\mathbf{x(z)}} = \frac{\mathbf{b_o}}{\mathbf{1+\sum_{k=1}^{p} a_k z^{-k}}} \qquad (4.24)$$

~$\mathbf{AR(p)}$ All pole model.

Therefore, the block diagram of a $\mathbf{AR(p)}$ process (generated by white noise $\mathbf{w(n)}$with an all-pole Coefficients), is as follows:

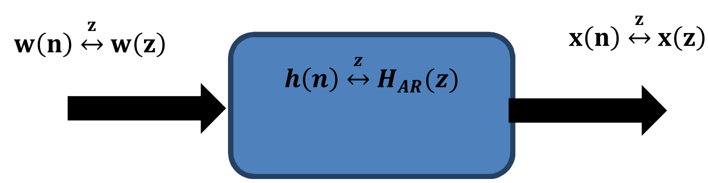


Figure 4-4: Block diagram of AR (p) model.

- **MOVING AVERAGE MODEL**: Whenp is zero, x (n) ~MA(q) Case, the output ARMA(0, q) is given by;

$$\mathbf{x(n)} = -\sum_{k=0}^{q} \mathbf{b_k w(n-k)} \qquad (4.25)$$

- The transfer function of moving average is given by;

$$\mathbf{H_{MA}(z)} = \frac{\mathbf{w(z)}}{\mathbf{x(z)}} = \sum_{k=0}^{q} \mathbf{b_k z^{-k}} \qquad (4.26)$$

$$\mathbf{H(z)} = \frac{\mathbf{w(z)}}{\mathbf{x(z)}} = \sum_{k=0}^{q} \mathbf{b_k z^{-k}} = \mathbf{b_o} + \sum_{k=1}^{q} \mathbf{b_k z^{-k}} \qquad (4.27)$$

~ $\mathbf{MA(q)}$ All zero model.

Therefore, the block diagram of $\mathbf{MA(q)}$ process (generated by white noise $\mathbf{w(n)}$with an all-zero coefficients), is as follows:

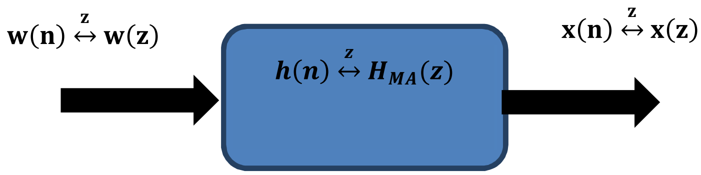


Figure 4-5: Block diagram of MA (q) model.

After determining the coefficients for the model the Autocorrelation of the time series is determined.

The Auto correction of $\mathbf{x(n)}$ is given by;

$$\mathbf{r_k = E[x(n+k)x^*(n)]} \tag{4.28}$$

$$\mathbf{r_k = E\left[-\sum_{m=1}^{p} a_k x(n+k-m) + w(n+k)\right] x^*(n)} \tag{4.29}$$

$$\mathbf{r_k = -\sum_{m=1}^{p} a_k\, r_{k-m} + E[w(n+k)x^*(n)]} \tag{4.30}$$

The white noise $\mathbf{w(n)}$has an effect only on the current value of $x(\mathbf{n})$ it doesn't affect the next values of $x(\mathbf{n+K})$.

When k≥1, for (4.30), the present value of x (n) will not be affected by the next $(\mathbf{n+K})$ values then the expectation$\mathbf{E[w(n+k)x^*(n)] = 0}$.

- When k>1, $\mathbf{r_k}$is given by;

$$\mathbf{r_k = -\sum_{m=1}^{p} a_m\, r_{k-m}} \tag{4.31}$$

- When k=o, $\mathbf{r_k}$is given by;

$$\mathbf{r_k = -\sum_{m=1}^{p} a_k\, r_{k-m} + \alpha^2 \delta(k)} \tag{4.32}$$

Generally the input $\mathbf{W(n)}$ represents a sequence of uncorrelated random variables of zero mean and constant variance $\mathbf{\alpha_w^2}$ so that,

$$\mathbf{r_{ww}(k) = \alpha_w^2 \delta(k)} \tag{4.33}$$

In addition, if $\mathbf{W(n)}$ is normally distributed then the output $\mathbf{x(n)}$ also represent a strict-sense stationary normal process.

From (4.4), an ARMA $(\mathbf{p,q})$the process system has only $\mathbf{p+q+1}$independent coefficients,

$(\mathbf{a_k, k = 1 \to p, b_i, i = 0 \to q})$, Hence its impulse response sequence $\{\mathbf{h_k}\}$ also must exhibit a similar dependence among them.

The theory of Kronecker[11] states that the necessary and sufficient condition for $\mathbf{H(z)} = \sum_{\mathbf{k=0}}^{\infty} \mathbf{h(k)z^{-k}}$ to represent a rational system (ARMA) is that;

$$\mathbf{detH_n = 0}, \qquad \mathbf{n \geq N} \text{ (For all sufficiently large n)},$$

$$\text{Where} \quad \mathbf{H_n} \triangleq \begin{bmatrix} \mathbf{h_0} & \mathbf{h_1} & \mathbf{h_2} & \cdots & \mathbf{h_n} \\ \mathbf{h_1} & \mathbf{h_2} & \mathbf{h_3} & \cdots & \mathbf{h_{n+1}} \\ \vdots & & & \cdots & \\ \mathbf{h_n} & \mathbf{h_{n+1}} & \mathbf{h_{n+2}} & \cdots & \mathbf{h_{2n}} \end{bmatrix} \qquad (4.34)$$

i.e., In the case of the rational system for all sufficiently large n, the above matrices (square matrix with….) $\mathbf{H_n}$ in (4.34) all have the same rank.

Finally, to find the coefficients, using;

$$\mathbf{H(z)} = \sum_{\mathbf{k=0}}^{\infty} \mathbf{h(k)z^{-k}} = \frac{\mathbf{b_o + b_1 + z^{-1} + \cdots + b_q z^{-q}}}{\mathbf{1 + a_1 z^{-1} + \cdots + a_p z^{-p}}}, \qquad (4.35)$$

The necessary part easily follows from (4.16) by crossing multiplying and equating coefficients of power of$\mathbf{z^{-k}}$, $\quad \mathbf{k = 0, 1, 2, \cdots}$.

This gives;

$$\mathbf{b_0 = h_0} \tag{4.36}$$

$$\mathbf{b_1 = h_0 a_1 + h_1}$$

$$\vdots$$

$$\mathbf{b_q = h_0 a_q + h_1 a_{q-1} + \cdots + h_m}$$

$$\mathbf{0 = h_0 a_{q+i} + h_1 a_{q+i-1} + \cdots + h_{q+i-1} a_1 + h_{q+i}} \quad , \mathbf{i \geq 1} \tag{4.37}$$

Now let's assume that,

$$\mathbf{i = p - q}\,, \qquad \mathbf{forsystemq \leq p -} \tag{4.38}$$

Substituting (4.38) in (4.37), it will be;

$$\mathbf{h_0 a_p + h_1 a_{p-1} + \cdots + h_{p-1} a_1 + h}_\mathrm{p} = 0 \vdots$$

$$\mathbf{h_p a_p + h_{p+1} a_{p-1} + \cdots + h_{2p-1} a_1 + h_{2p} = 0} \tag{4.39}$$

Writing (4.39) in matrices form results in

$$\begin{bmatrix} \mathbf{h_0} & \mathbf{h_1} & \mathbf{h_2} & \cdots & \mathbf{h_p} \\ \mathbf{h_1} & \mathbf{h_2} & \mathbf{h_3} & \cdots & \mathbf{h_{n+1}} \\ \vdots & & & \cdots & \\ \mathbf{h_p} & \mathbf{h_{p+1}} & \mathbf{h_{p+2}} & \cdots & \mathbf{h_{2p}} \end{bmatrix} \begin{bmatrix} \mathbf{a_p} \\ \mathbf{a_{p-1}} \\ \vdots \\ \mathbf{1} \end{bmatrix} = [\mathbf{0}] \tag{4.40}$$

It's possible to obtain similar determinantial conditions for ARMA system in terms of the above matrices generated from its output autocorrelation sequence.

Referring to (4.9), the input white noise process $\mathbf{W(n)}$ there is uncorrelated with its own past sample values as well as the past values of the system output. This gives,

$$\mathbf{E\{w(n)w^*(n-k)\} = 0}\,, \qquad \mathbf{k \geq 1} \tag{4.41}$$

$$\mathbf{E\{w(n)x^*(n-k)\} = 0}\,, \qquad \mathbf{k \geq 1} \tag{4.42}$$

Together with (4.11), below equation is obtained;

$$\mathbf{r_i = E\{x(n)x^*(n-i)\}}$$

$$\mathbf{= -\sum_{k=1}^{p} a_k r_{i-k} + \sum_{k=0}^{q} b_k \{w(n-k)w^*(n-i)\}} \tag{4.43}$$

Hence in general the,

$$\mathbf{\sum_{k=1}^{p} a_k r_{i-k} + r_i \neq 0} \qquad \mathbf{, i \leq q} \tag{4.44}$$

And

$$\mathbf{\sum_{k=1}^{p} a_k r_{i-k} + r_i = 0} \qquad \mathbf{, i \leq q} \tag{4.45}$$

Notice that (4.34) is the same as (4.45) replacing $\{\mathbf{h_k}\}$ by $\{\mathbf{r_k}\}$ and hence the Kronecker conditions for rational systems can be expressed in terms of its output autocorrelation as well.

Thus if $x(n) \sim ARMA(p, q)$ represents a wide sense stationary stochastic process, then its output autocorrelation sequence $\{r_k\}$ is given by,

$$\mathbf{D} \triangleq \begin{bmatrix} \mathbf{r_0} & \mathbf{r_1} & \mathbf{r_2} & \cdots & \mathbf{r_k} \\ \mathbf{r_1} & \mathbf{r_2} & \mathbf{r_3} & \cdots & \mathbf{r_{k+1}} \\ \vdots & & & \cdots & \\ \mathbf{r_k} & \mathbf{r_{k+1}} & \mathbf{r_{k+2}} & \cdots & r_{2k} \end{bmatrix} \tag{4.46}$$

Once the coefficients are known the order of the models is figured out.

4.3 Auto-Regressive Integrated Moving Average Model

4.3.1 Introduction

One of the most popular and frequently used stochastic time series models is the ARIMA. The integrated ARMA (ARIMA) is a broadening of the class of ARMA that includes differencing. ARIMA is an acronym for Autoregressive Integrated Moving Average model ("integration" in this context is the reverse of differencing). ARIMA model is denoted as ARIMA (p,d,q) where d represent the order of differencing.
The ARIMA procedure analyzes and forecasts equally spaced univariate time series data, transfer function data, and intervention data using the ARIMA model. An ARIMA model predicts a value in a response time series as a linear combination of its own past values, past errors (also called shocks or innovations), and current and past values of other time series.

ARIMA model is the most general class of models for forecasting a time series which can be made to be "stationary" by differencing (if necessary), perhaps in conjunction with nonlinear transformations such as logging or deflating (if necessary). An ARIMA model can be viewed as a "filter" that tries to separate the signal from the noise, and the signal is then extrapolated into the future to obtain forecasts.

The ARMA procedure provides a comprehensive set of tools for univariate time series model identification, parameter estimation, and forecasting, and it offers great flexibility in the kinds of ARIMA models that can be analyzed. The ARIMA procedure supports seasonal, subset, and factored ARIMA models; intervention or interrupted time series models; multiple regression analysis with ARMA errors; and rational transfer function models of any complexity [9, 10].

4.3.2 Mathematical description

Combining differencing with auto regression and a moving average model, it is obtained a non-seasonal ARIMA model. The full model can be written as:

$$\mathbf{x}'(\mathbf{n}) = -\sum_{\mathbf{k}=\mathbf{1}}^{\mathbf{p}} \mathbf{a}_{\mathbf{k}} \mathbf{x}'(\mathbf{n}-\mathbf{k}) + \sum_{\mathbf{k}=\mathbf{0}}^{\mathbf{q}} \mathbf{b}_{\mathbf{k}} \mathbf{w}(\mathbf{n}-\mathbf{k}) \qquad (4.47)$$

Where,

- $\mathbf{x}'(\mathrm{n}-\mathrm{k})$-is the differenced series sampled at k,

  $\mathbf{x}'(\mathbf{n})$-is the differenced series in n samples.

The above equation is called An ARIMA model, and notated as ARIMA (p, d,q).

Most of the models already discussed above are special cases of the ARIMA model as shown in Table 4-1 below.

The same stationary and invertibility conditions that are used for autoregressive and moving average models apply to this ARIMA model.

Models	ARIMA (p,d,q)
White Noise	ARIMA (0,0,0)
Random Walk	ARIMA (0,1,0) with no constant
Random Walk with drift	ARIMA (0,1,0) with constant
Auto-Regression	ARIMA (p,0,0)
Moving-Average	ARIMA (0,0,q)

Table 4-1: Types of ARIMA model and their orders.

The basic tools for model identification are the graphs of the sample ACF and sample PACF obtained from the series. The ACF (sometimes called the correlogram) indicates the degree of correlation within the series for lags and it checks the stationary of the data sample. In a similar fashion, the PACF indicates the degree of correlation at a given lag after accounting for the correlation from the intervening lags.

Some behavior of ACF and PACF to identify the right model is shown in the below table.

Process	ACF	PACF
AR(P)	Infinite: damps out	Finite: Cuts off after lag p
MA(q)	Finite: Cuts off after lag q	Infinite :damps out
ARMA (p,q)	Infinite: damps out	Infinite: damps out

Table 4-2: Summary of Correlation Patterns.

4.4 Seasonal Auto-Regressive Integrated Moving Average (SARIMA) Model

4.4.1 Introduction of SARIMA

The above models are restricted in non-seasonal data and non-seasonal ARIMA models. However, ARIMA models are also capable of modeling a wide range of seasonal data. A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA models.

SARIMA model takes into account the seasonal characters of the time series. The model is used in the analysis of stochastic but not stationary time series and complements ARIMA models [10, 11, 16].

4.4.2 Mathematical description

Mathematically ARIMA is given as:

$$\mathbf{x'(n)} + \sum_{\mathbf{k=1}}^{\mathbf{p}} \mathbf{a_k x'(n-k)} + \mathbf{a_k x'(n-k^s)} = \sum_{\mathbf{k=0}}^{\mathbf{q}} \mathbf{b_k w(n-k)} + \mathbf{b_k w(n-k^s)} \qquad (4.48)$$

Where,

- $\mathbf{x'(n-k)}$ -is the differenced series in n samples,
- $\mathbf{x'(n)}$ -is the differenced series in n samples,
- a_k-is the Lag-k autoregressive coefficients,
- p-is the order of auto-regressive model,
- q-is the order of Moving-Average model,
- $w(n-k)$-is the white noise sampled at K,
- b_k-is the Moving average coefficients,
- s −is the Seasonal Components, number of periods per season.

The above equation is known as a Seasonal ARIMA of order (p,d,q) or ARIMA$(p, d, q)x(P, D, Q)_s$ with p=non-seasonal AR order, d= non-seasonal differencing, q= non-seasonal MA order, P= Seasonal AR order, D = Seasonal differencing, Q=Seasonal MA order and S=time span of repeating seasonal pattern.

The seasonal part of the model consists of terms that are very similar to the non-seasonal components of the model, but they involve backshifts of the seasonal period.

From a practical perspective, fitted seasonal ARIMA models provide linear state transition equations that can be applied recursively to produce single and multiple interval forecasts.

4.5 Order Selection Criterion

Model selection criterion allows the best model to fit the data by striking a balance and finding a model that neither under-fits nor over-fits the data. In model selection, the idea is to find the smallest set of variables which provides an adequate description of the data.

The following are the motivations for model selection criteria:

- **Maximum likelihood estimation (MLE):** This technique finds the values of the parameters which maximize the probability of obtaining the data that we have observed. For ARIMA models, MLE is very similar to the least squares estimates that would be obtained by minimizing.

  $$\mathbf{MLE} = \sum_{t=1}^{T} e_t^2 \qquad (4.49)$$

  Where; e_t-is theerror between the exact and the estimated data.

- **Akaike's Information Criterion (AIC):** The intent of AIC is to measure the mathematical distance between the true data sample and the fitted model. AIC is a widely used measure of a statistical model. It basically quantifies ;

  1) The goodness of fit, and

  2) The simplicity/parsimony, of the model into a single statistic.

For the above mentioned reasons, this thesis uses the AIC parameter to compare and select from the forecasting methods. Now, let us apply this powerful tool in comparing various ARIMA models, often used to model time series.

Mathematically, it can be written as,

$$\mathbf{AIC = -2log(L) + 2(p + q + k + 2)} \tag{4. 50}$$

Where,

- $\boldsymbol{L}$-is the likelihood of the data,
- K $is\ constant$ $==\begin{cases} 1, & constant \neq 0 \\ 0, & constant \quad 0 \end{cases}$

For ARIMA models, the AIC can be written as:

$$\mathbf{AIC_c = AIC + \frac{2(p+q+k+1)(p+q+k+2)}{T-p-q-k-2}} \tag{4. 51}$$

Where,

- K=$\begin{cases} 1, & constant \quad 0 \\ 0, & constant = 0' \end{cases}$
- T –is the number of data points in the sampled data.

- **Bayesian Information Criterion (BIC):** is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred. It is based, in part, on the likelihood function and it is closely related to AIC.

    - Mathematically it is given as:

$$\boldsymbol{BIC = -2lnL + kln(T) = AIC + (log(T) - 2)(p + q + k + 1)} \quad (4.51)$$

Where;

- $\boldsymbol{L}$ −is the likelihood of the data,
- K= $\begin{cases} 1, & constant \neq 0 \\ 0, & constant = 0' \end{cases}$

In a model selection application, the optimal fitted model is identified by the minimum value of BIC. However, as with the application of any model selection criterion, the criterion values are important; models with similar values should receive the same "ranking" in assessing criterion preferences [17].

BIC versus AIC

AIC and BIC share the same goodness-of-fit term, but the penalty term of BIC ($\boldsymbol{kln(n)}$) is potentially much more stringent than the penalty term of AIC ($\boldsymbol{2K}$). Thus, BIC tends to choose fitted models that are more parsimonious than those favored by AIC. The differences in selected models may be especially pronounced in large sample settings.

5. UMTS Capacity Planning

The previous chapters gave some insight on the time series analysis and the different techniques of modeling time series. UMTS capacity planning is the process of estimating the number of nodes needed to support a given number of data demand or subscribers. One input parameter for capacity planning is the demanded data throughput capacity, which is the result of proper forecasting. This chapter presents the capacity planning procedures first and then explains the process of calculating the data volume per user per month and the busy hour traffic per active user.

5.1 Capacity planning procedures

Most of the 3G network license holders have an existing 2G license and will face great planning challenges when upgrading a network from 2G to 3G.2G network capacity planning depends on the amount of transceivers, but 3G capacity planning depends on the amount of customer phone usage and other carried traffic.

The number of installed transceivers limits the mobile network theoretical capacity. In UMTS systems interference, accepted and planned quality and grade of service will determine the system capacity. UMTS systems also have soft capacity, which complicates the network area capacity estimations.

The capacity estimation is an important part of the scale estimation. The purpose of capacity estimation is to estimate the approximate BS number needed by the capacity according to the service model and service traffic demand of the network planning. It should be performed from the uplink and downlink.

For the UMTS system capacity, the interference is limited in the uplink direction and the BS power is limited in the downlink direction. However, in the UMTS network, the data service proportion is obviously increased and the network uplink and downlink traffic becomes asymmetric generally, and even the downlink capacity may be limited. Therefore, the UMTS capacity estimation should be performed from the uplink and downlink respectively [29].

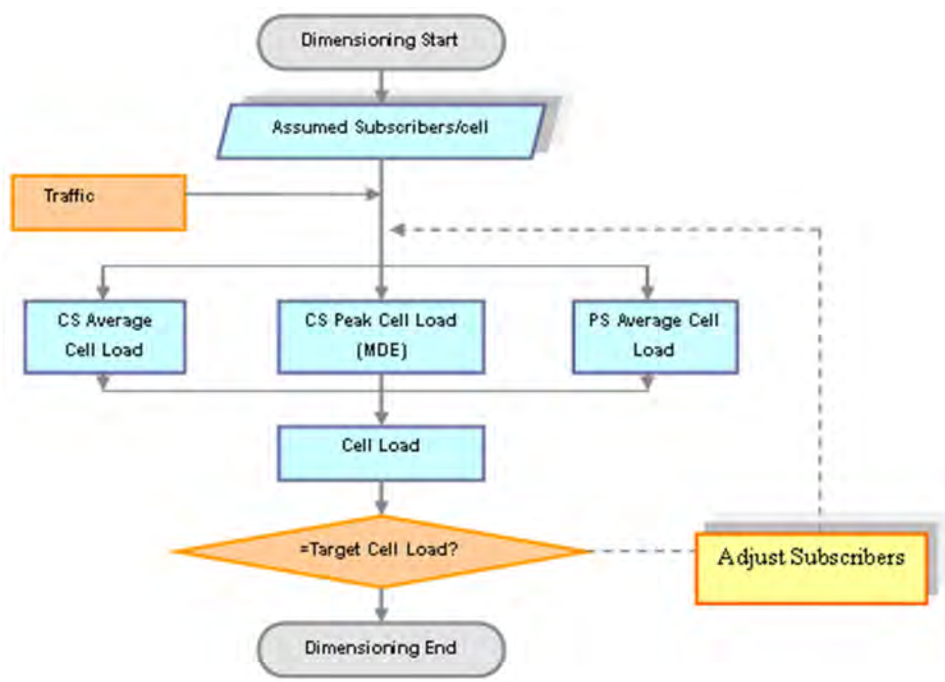


Figure 5-1: UMTS Capacity Dimension Procedure [31].

The aim of UMTS capacity planning in this thesis is to provide an input to the service provider to help obtain the number of subscribers supported per cell based on the traffic model. According to the traffic model mentioned in this planning, load per connection for different service can be calculated. According to load per connection for different service and cell load, the subscribers per cell per different service can be obtained and finally get the minimum one of different service subscribers per cell. Finally, the largest one among coverage and capacity result will be the final value.

Capacity planning is divided into two parts:

- The first thing is to estimate a single transceiver and site capacity. Calculations how the noise raises as the cell load increases is out of the scope of this page, but in-cell noise, Eb/No requirements, planned data rates, coverage probability, air resources usage activity factor, target interference margin and processing gains are needed to approximate the transceiver and site capacity. Depending on the parameter values, planned transceiver capacity is typically from400kbits/s to 700kbits/s per transceiver [30].
- The second part of the process is to estimate how many mobile users each cell can serve. Once the cell capacity and subscriber traffic profiles are known, network area base station requirements can be calculated. Estimations can be done in Erlangs per subscriber or kilobits per subscriber.

  In this thesis we try to find the number of kilobits per subscriber by using the best fitted forecasting model for the future. This number can be used in the second part of the capacity planning which is in the estimation of the number of subscribers and the usage traffic profile in kilobits per subscriber. The next section shows the traffic model of UMTS in Addis Ababa and considers the traffic model of Ethio-Telecom.

5.2 UMTS Traffic Model

The monthly data allowance is specified based on the operator marketing analysis [29].Accordingly, 10Gigabyte per month is set for data card users and 1Gigabyte per month is set for smartphone (SP) users. The percentage share of data card compared to smartphone user is 37% and this information is taken from the operator design document. The data rate per user at a busy hour is calculated based [29]:

$$\mathbf{DVBH[kbps]} = \frac{\mathbf{DV*BHR*1024*1024}}{\mathbf{30\ days*3600\ sec}} \qquad (5.1)$$

Where;

- DVBH – is the data volume per month per user at busy hour given in kbps
- DV – is the data volume per month per user in GByte collected from the result of the selected forecasting model,
- BHR – is the busy hour ratio (%), which is 10% obtained from Ethio telecom.

Table 5-1 below shows the input of UMTS Traffic model for Ethio-telecom

Traffic Usage in GB/Month/User	UMTS Data-Traffic model	
	Dongle	Smart Phone
Distribution percentage	37%of subscribers	63% of subscribers
Traffic per user per month in GB	10	1

Table 5-1: input UMTS Traffic Model from Ethio-telecom [30].

Type	Data Dongle	Data SP
Data per month(GB)	10	1
Proportion (%)	37%	63%
Data volume per month per user	3.7	0.63
Data Volume per month per user(GB)	4.33	
Busy hour ratio (%)	8.3%	
Data volume per month per user @ BH(kbps)	33.88	
Active User ratio (%)	70%	
Active user Data volume per user @BH(kbps)	33.88	

Table 5-2:Detail input UMTS Traffic Model from Ethio-telecom [29].

According to current Ethio-telecom sells model as shown in the table 5-2, the provided traffic model is an economical and acceptable traffic model.

In recent studies [27, 29] and researches the planning phase of network infrastructure is done using the operator traffic model provided as an input and currently, the capacity calculation is done as follows:

Step1: Data volume per month per user calculation

10GB*37%+1GB*63% =4.33GB.

Step2: HSPA throughput per user calculation

Per day per user: "Step1" *8*1024*1024/30/3600 =322.34Kbps.

Busy hour is 10%;

Busy hour per user =322.34*10% =32.237Kbps

Step3: Busy hour per active user calculation

Active user ratio is assumed to be 70%;

Busy hour per active user: 32.237*70%=22.5659Kbps.

6. Results and Discussions

The ARIMA analysis consisted of three stages; identification, estimation and forecast (SeeFigure-6.1).

- Identification; is the stage at which a tentative model for the series is selected from the large family of candidate ARIMA (p,d,q) models. Clearly there are many possible combinations of the orders p,d, and q. Thus, the identification stage consists of specifying the: AR, I, and MA orders (p,d,q).The stage consists of specifying the appropriate model (AR, MA, ARMA, or ARIMA) and order of model. The basic tools for identifying the model are the graphs of Sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF) obtained from the data-traffic time series. The ACF (sometimes called the correlogram) indicates the degree of correlation within the series for lags 1, 2, 3 . . . etc. In a similar fashion, the PACF indicates the degree of correlation at a given lag after accounting for the correlation from the intervening lags.

  Sometimes it is done by an auto fit procedure – fitting many different possible model structures and orders and using a goodness-of-fit statistic to select the best model.

- In Estimation and Diagnostic checking (second stage), here the stage is used to specify the ARIMA model to fit to the variable specified in the previous stage, and to estimate the parameters of the model. This stage also produces diagnostic statics to evaluate the adequacy of the model.

- In the last stage, Matlab command is used to forecast future values of the data-traffic time series.

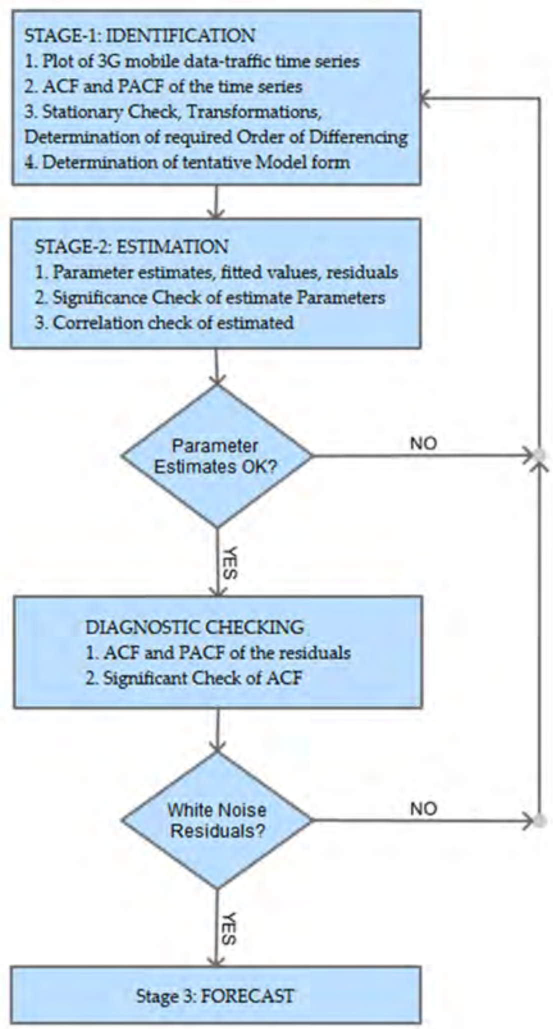


Figure 6-1: The Three stages of Forecasting.

These three stages are explained further in the following two sections. The first section describes the process of determining alternative model orders to choose from for the forecasting and the second section describes how the final model order selection is accomplished and shows the forecasting results.

6.1 Section 1: Identification and Estimation Results

Step 1.Plot the Peak Traffic load for UMTS data network.

The first step in analyzing time series data is to plot the data against time – this is called a time plot. The plotted time series is used to obtain simple descriptive measures of the main properties of the series. This plot can immediately reveal features such as trend, seasonal variation, discontinuities, and outliers that may be present in the data.

Figure 6.2 shows the time plot of the peak load of the Addis Ababa UMTS data traffic per day from October 16, 2015 to July 4, 2016 obtained from Ethio telecom.

To determine the order of differencing d, the time series must be checked for non-stationary. If non-stationary is indicated, differencing or other transformations must be performed prior to further analysis.

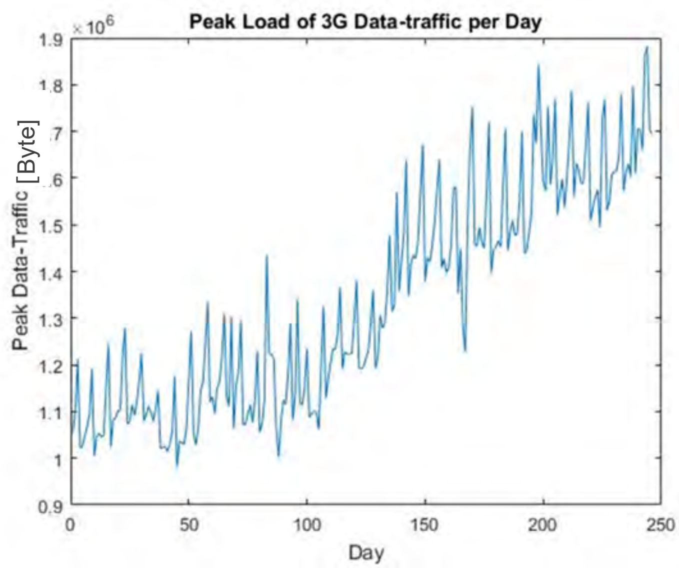
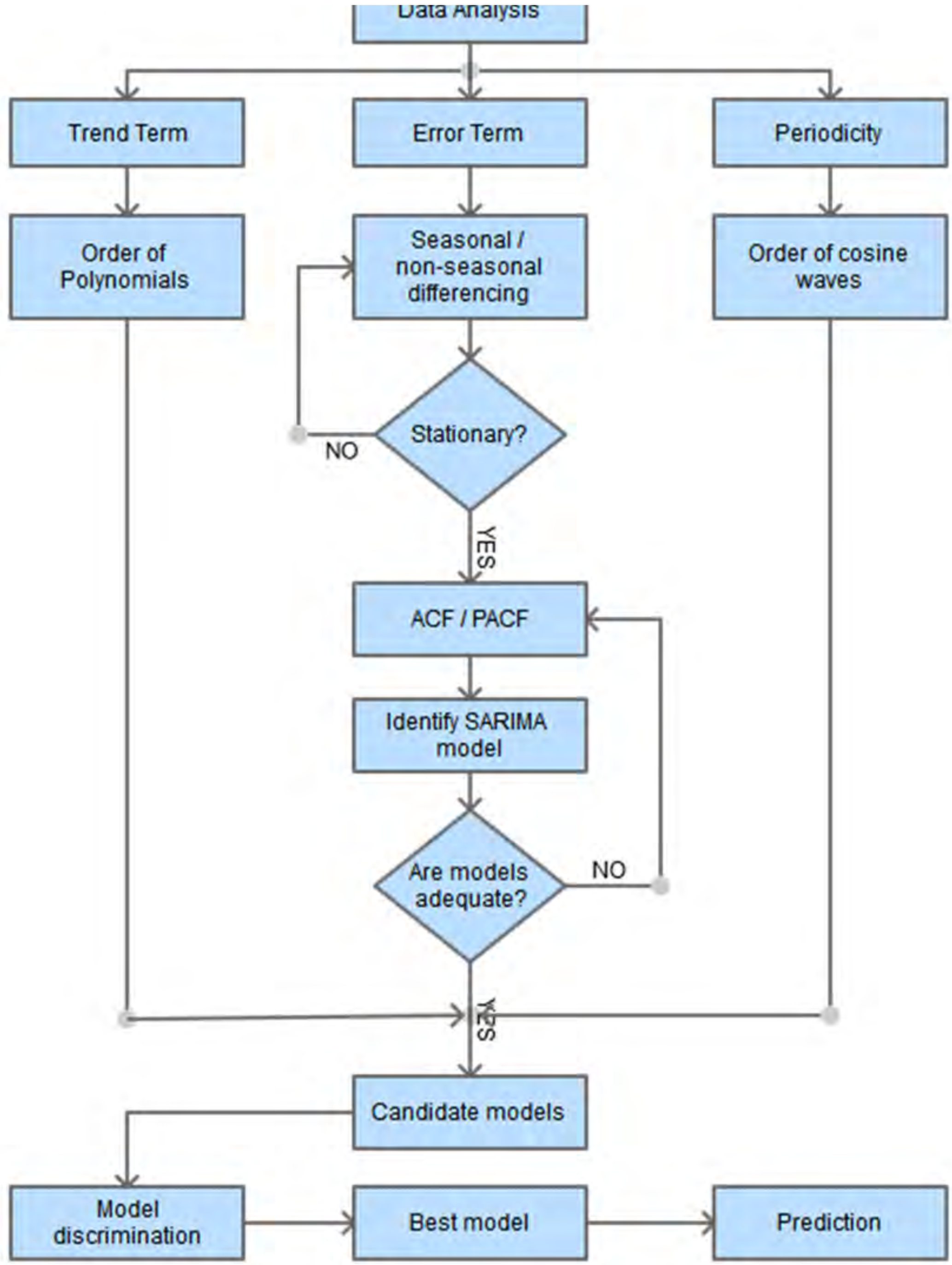


Figure 6-2: UMTS Data-Traffic Peak Load of Addis-Ababa.

Figure 6-3: Flowchart of General ARIMA



Step 2. Difference the sample data to make it stationary on mean (remove trend).Using System Identification tools in MATLAB, the mean is removed as shown in the figure below.

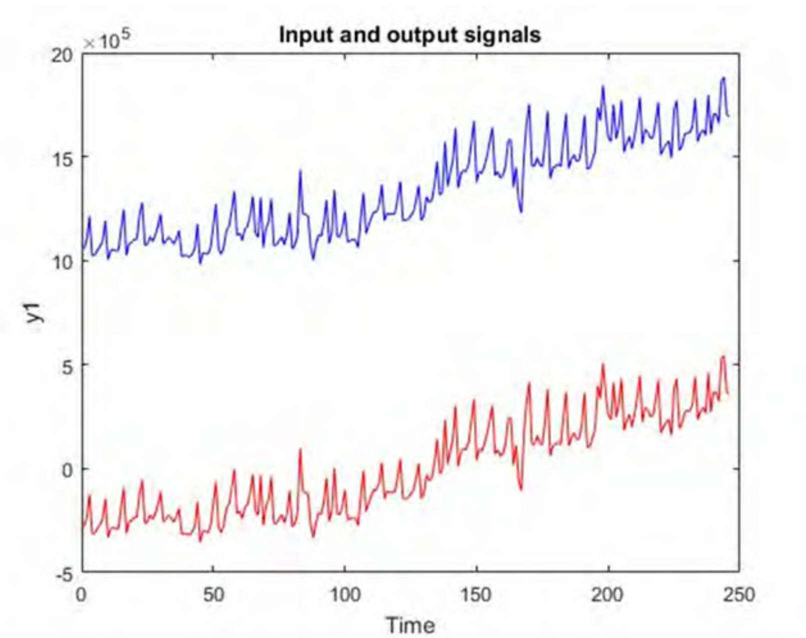


Figure 6-4: Removed Mean Peak Traffic Load

In the above figure, the vertical axis labeled y1 indicates the peak traffic load in GBytes and the horizontal axis indicates the corresponding number of days.

The blue line is the main peak data-traffic load and the red one is the data after the mean is removed. Still the data is not stationary and that's because the data is not stationary in variance. The data need to be stationary on variance to produce reliable forecast through ARMA model.

Step 3. The Log transform of the data to make it stationary variance.

To make a series stationary on variance the best way is through transforming the original Peak load.

Going back to the original peak traffic load data and log transform it, makes it stationary in variance.

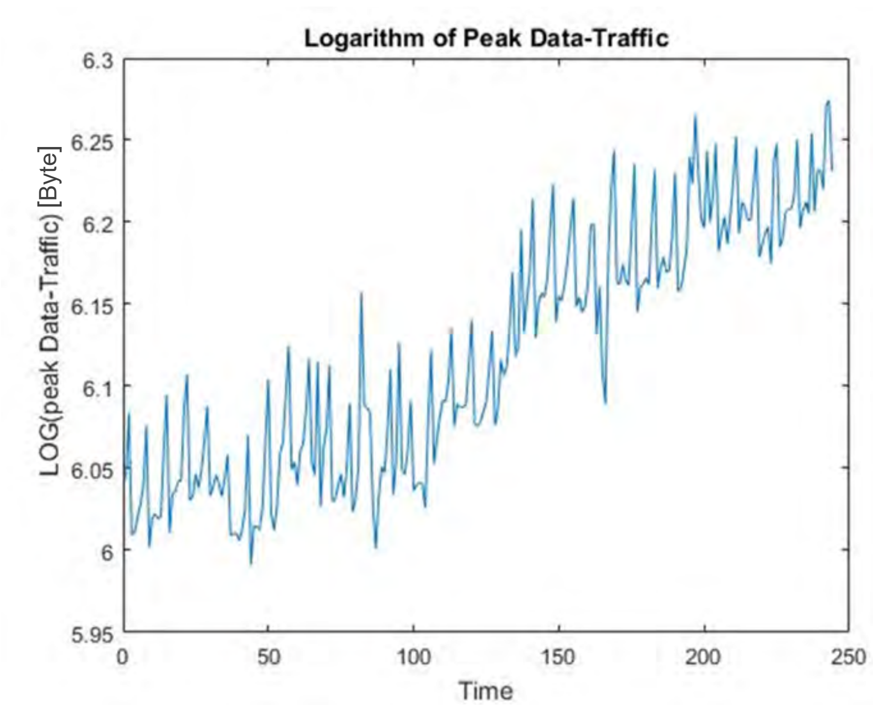


Figure 6-5: Log of Peak Load Data-Traffic.

Still, the series is not stationary on mean since we are using the original data.

Step 4.the difference plot for log transformed data series will make the Peak Traffic data stationary both in mean and variance.

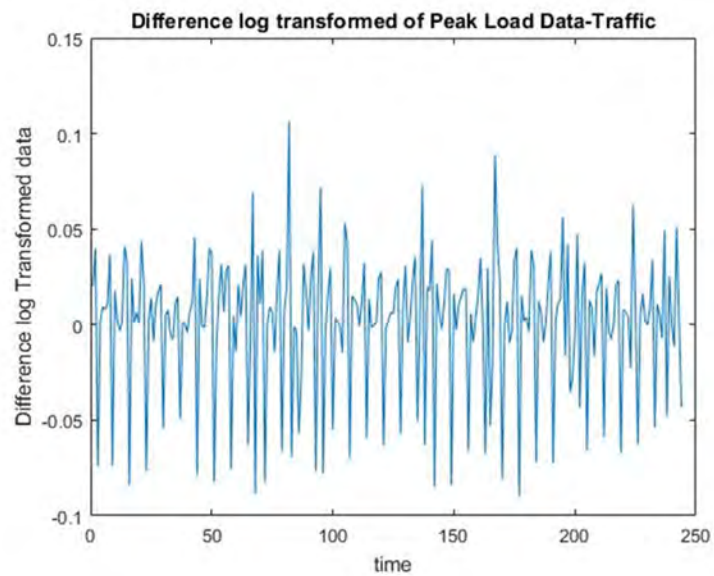


Figure 6-6: Difference log. Transformed Peak Load Data-Traffic.

Step 5: Plot ACF and PACF to identify potential AR and MA model Creating ACF and PACF plots helps to identify pattern in the above data which is stationary both in mean and variance. The aim is to identify presence of AR and MA components in the residuals.

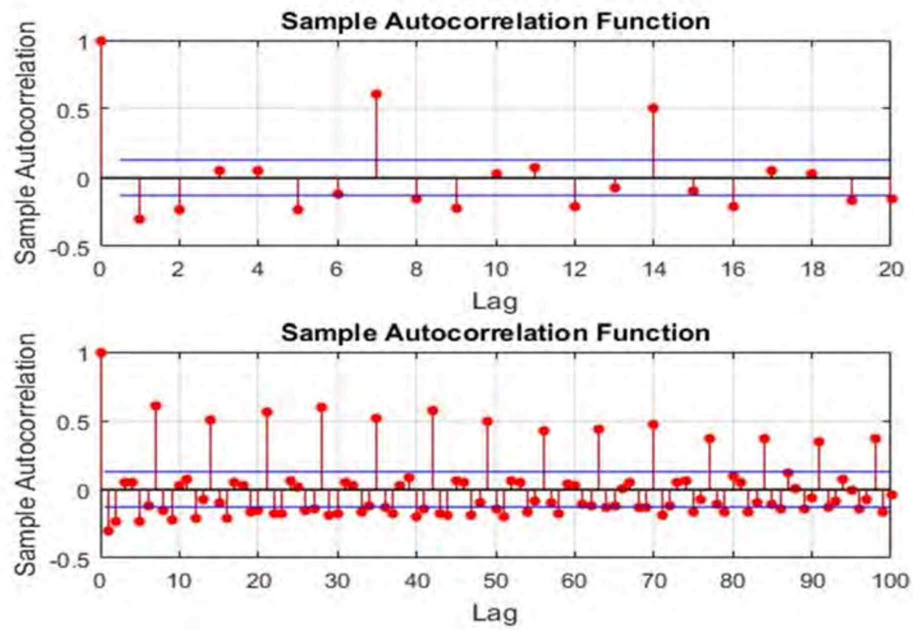


Figure 6-7: Autocorrelation Function for 20 and 100 lags.

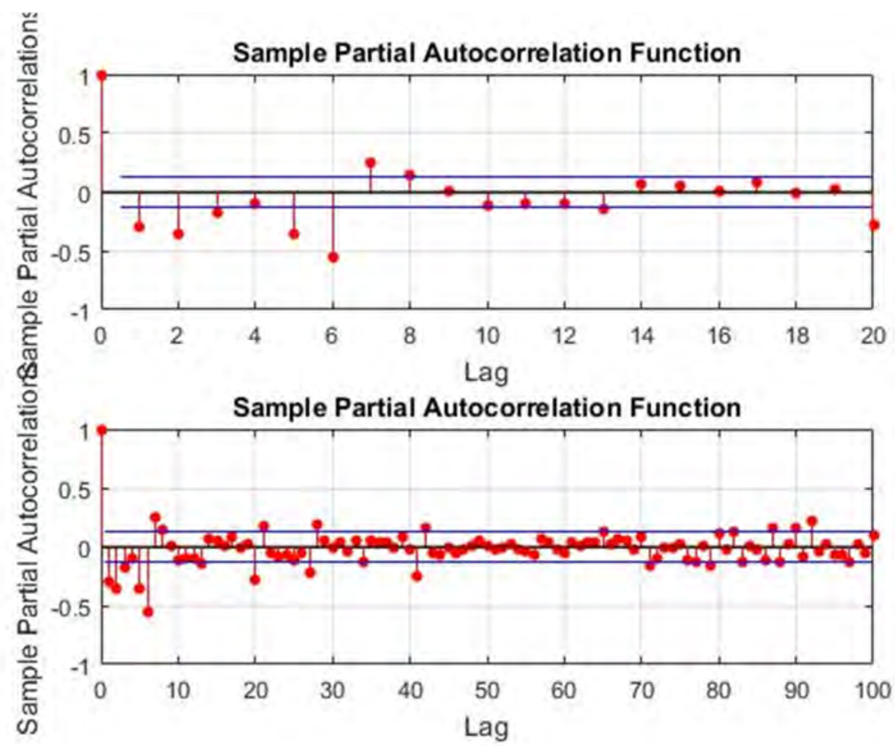


Figure 6-8: Partial autocorrelation Function for 20 and 100 lags.

At the First stage of differencing, the data is stationary. The plots of sample ACF and the sample partial autocorrelation function PACF have "cuts off" the bounds ±1.96 / n. Since there are enough spikes in the plots outside the insignificant zone (dotted horizontal lines).It can be concluded that the residuals are not random. This implies that there is information available in the residuals to be extracted by AR and MA models.

If the ACF trails off and the PACF shows spikes, then an autoregressive (AR) model with order q equal to the number of significant PACF spikes is considered the "best" model. If the PACF trails off and the ACF shows spikes, the moving average (MA) model with order q equal to the number of significant ACF spikes is the best model. If both the ACF and the PACF trail off then an autoregressive moving average (ARMA) model is used with p and q equal to one. If the data had to be differenced for it to become stationary, then the ARIMA model is used.

The next step is to examine the patterns in the autocorrelation plot to choose candidate ARMA model to the series. The partial autocorrelation function plots are useful aids in identifying appropriate ARMA models for the data-traffic time series.

Taking the logarithm and differencing by one, the data is stationary. The sample ACF decays more slowly than the sample PACF .The plot of ACF and PACF gradually tail off. Hence, it's suggested that an ARMA model with order of p and q of no more than six.

Also there is seasonal component available in the residuals at the lag 7, 14, 21…, 7n where n is 1, 2, 3…, and the sample ACF has periodicity, the series has seasonality. This means there is high peak load data traffic every Sunday which yield the data-traffic sample data to be modeled in Seasonal ARIMA process.

As stated earlier, it has been found that adequate models for many observed data sets require only non-seasonal and seasonal orders p, d, q, P, D, and Q less than or equal to about six with seasonal components in 7^{th} ,14^{th} ,.. days.

- The second stage is to estimate the order of the model. At this stage, the coefficients are estimated, so that the sum of squared residuals is minimized.

The residuals, i.e. the differences between the observed time series values and the model calculate for "fitted" values, are also obtained at this stage.

In this thesis, we will use the BIC and best fit criterion methods to chosen the candidates orders of the model. Final decisions between models are made based on the above criterion also by using MSE (minimum square error).

Therefore, the best model is the model adequately describes data and has fewest parameters.

6.2 Section 2: Best Fit Model Identification and Forecast Results

Step 6: Identification of best fit ARMA model

To identify the best combination of p and q lags BIC and Best fit criterion methods are used in this thesis. The best model is chosen with the minimum AIC and BIC values. The BIC suggested that p=5 and q=1 are used to model the peak-load of the data traffic. Using Matlab Codes, the candidates and the best orders are identified.

The below table shows the result of the BIC for all orders less than (6, 6).

(p , q)	p=1	p=2	p=3	p=4	p=5	p=6
q=1	**-0.9198**	-0.9709	**-0.9654**	-0.9663	**-0.9393**	-0.9625
q=2	**-0.9655**	-1.0054	-1.0000	-1.0076	-1.0022	-1.0054
q=3	-0.9688	-0.9999	-0.9945	-1.0028	-0.9976	-0.9936
q=4	-0.9655	-0.9967	-0.9917	-1.0030	-1.0070	-1.0120
q=5	-0.9949	-1.0221	-1.0438	-1.0384	-1.0434	-1.0380
q=6	-1.0762	-1.0778	-1.0883	-1.0853	-1.0823	-1.0804

Table 6-1: BIC values for the order less than (6, 6) (in 10^3).

The smallest BIC value is -0.9198,-0.9393,-0.9654 and -0.9655 respectively. The figure below shows the order and their BIC value.

Smallest BIC Value	Order of p	Order of q
-0.9198	1	1
-0.9393	5	1
-0.9654	3	1
-0.9655	1	2

Table 6-2: Selected order of the models in BIC method.

After estimating the orders of the ARIMA model, diagnostic checking is implemented to check whether the model fit properly or not. One of the important elements in this stage is to make sure that the residuals of the candidate model are random and normally distributed. And the other one is to ensure that the estimated parameters are statistically significant. The fitting process is usually guided by the principle of parsimony, by which the best model is the one which has fewest parameters among all models that fit the data. This step is also called diagnostic checking, or verification.

Two important elements of checking are to ensure that the residuals of the model are random, and to ensure that the estimated parameters are statistically significant. Usually the fitting process is guided by the principal of parsimony, by which the best model is the simplest possible model – the model with the fewest parameters - that adequately describes the data.

Checking the Goodness of Fit in Matlab, it is necessary to plot the normal distribution of the residuals of the orders.

The QQ-plot, sample autocorrelation function and the partial autocorrelation function for order (1, 1), (5, 1), (3, 1) are shown in the figure below. The residuals are reasonably normally distributed and uncorrelated.

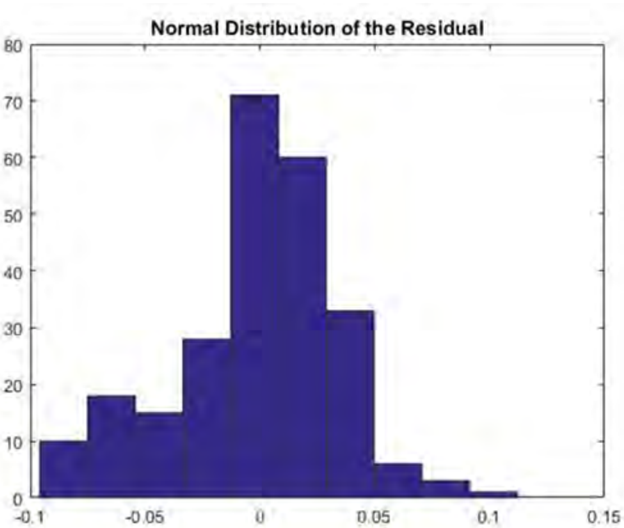


Figure 6-9: Normal distribution of the residual of order ARIMA(1,1,1)X(0,1,1)7.

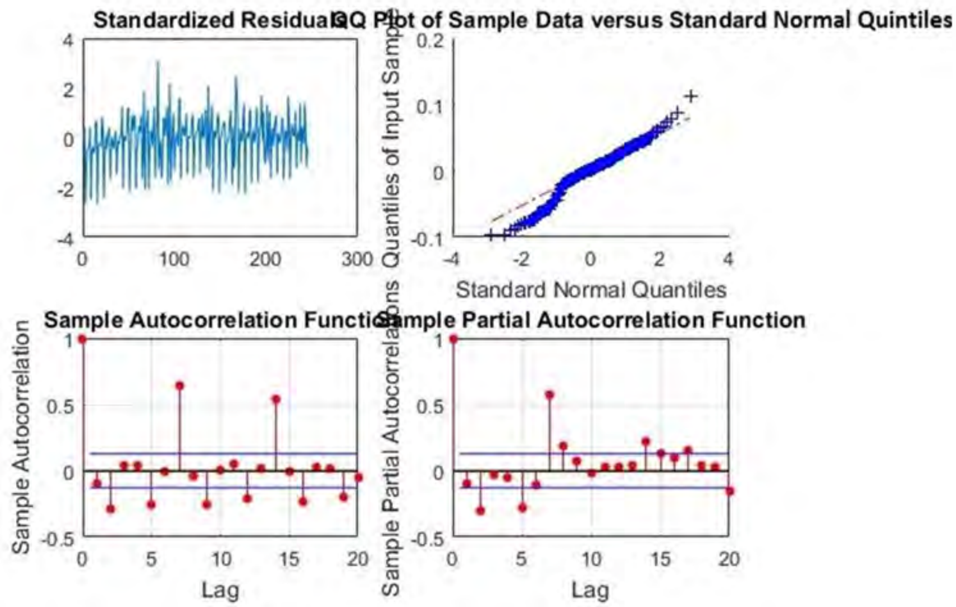


Figure 6-10: Standardized Residuals, Standard Normal Quintiles Sample ACF andPACF of SARIMA (1, 1, 1) X (0, 1, 1)$_7$

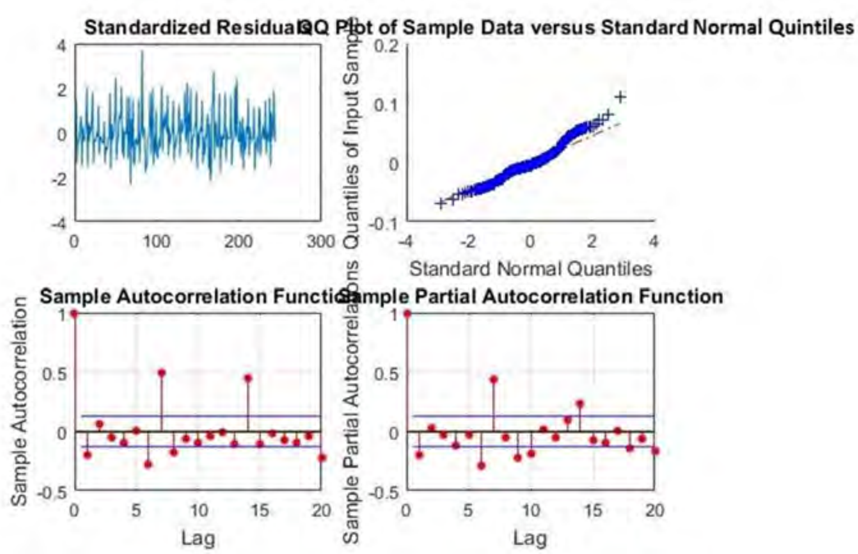


Figure 6-11: Standardized Residuals, Standard Normal Quintiles Sample ACF and PACF of SARIMA (5, 1, 1) X (0, 1, 1)$_7$

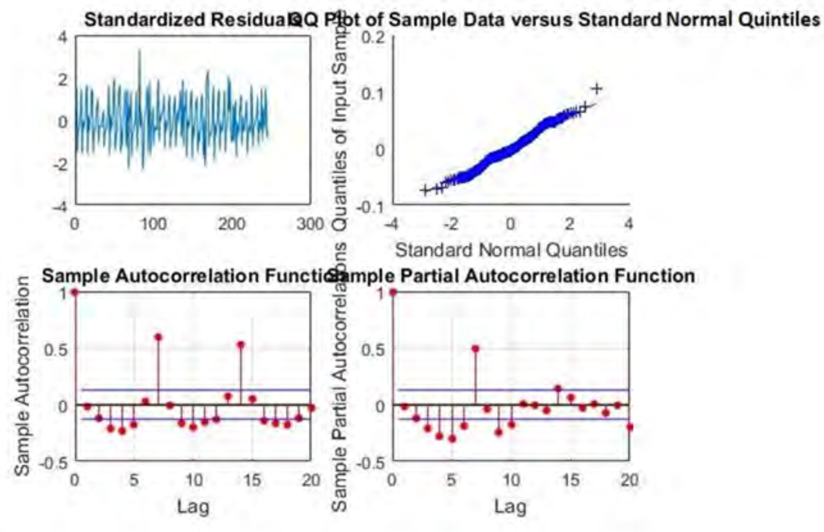


Figure 6-12: Standardized Residuals, Standard Normal Quintiles Sample ACF and PACF of SARIMA (3, 1, 1) X $(0, 1, 1)_7$

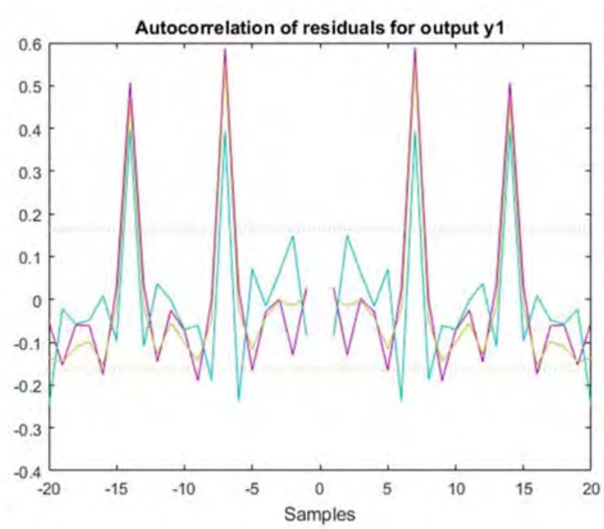


Figure 6-13: ACF of residuals for the selected models.

- The Green line=the residues of SARMA (5, 1, 1) x $(0, 1, 1)_7$
- The Purple line= the residues of SARMA (1, 1, 1) x $(0, 1, 1)_7$
- The Yellow line= the residues of SARMA (3, 1, 1) x $(0, 1, 1)_7$

By using the minimum square error and the BIC values the models are selected as shown in the table below.

Model ARIMA(p,1,q)	**BIC(10^3)**	**Fit to Estimation in %**
(1,1,1)	-0.9198	58
(5,1,1)	-0.9393	60
(3,1,1)	-0.9654	58
(5,0,0)	-0.9655	58
(1,0,0)	-	45
(0,0,1)	-	57
(3,0,0)	-	55

Table 6-3: Selected models in minimum BIC and Fit to Estimation.

After determining the orders, the coefficients must be checked whether they are stable or not and invertible or not. To check stability and Invertible the coefficients of AR must be inside the unit circle and the MA coefficients must be different from 0 respectively.

- The coefficients for (1,1,1):

  $A(z) = 1 - 0.2151z^{-1}$

  $B(z) = 1 - 0.86232z^{-1}$

- The coefficients for (5,1,1):

$$A(z) = 1 + 0.1654z^{-1} + 0.3819z^{-2} + 0.2544z^{-3} + 0.2416z^{-4} + 0.385z^{-5}$$

$$B(z) = 1 - 0.4820z^{-1}$$

- The coefficients for (3,1,1):

$$A(z) = 1 - 0.2086z^{-1} + 0.1704z^{-2} + 0.007494z^{-3}$$

$$B(z) = 1 - 0.8128z^{-1}$$

The figures below show the selected model type plots versus the real data, to compare the fitting process. In these figures, the plots colored in black represent the real data.

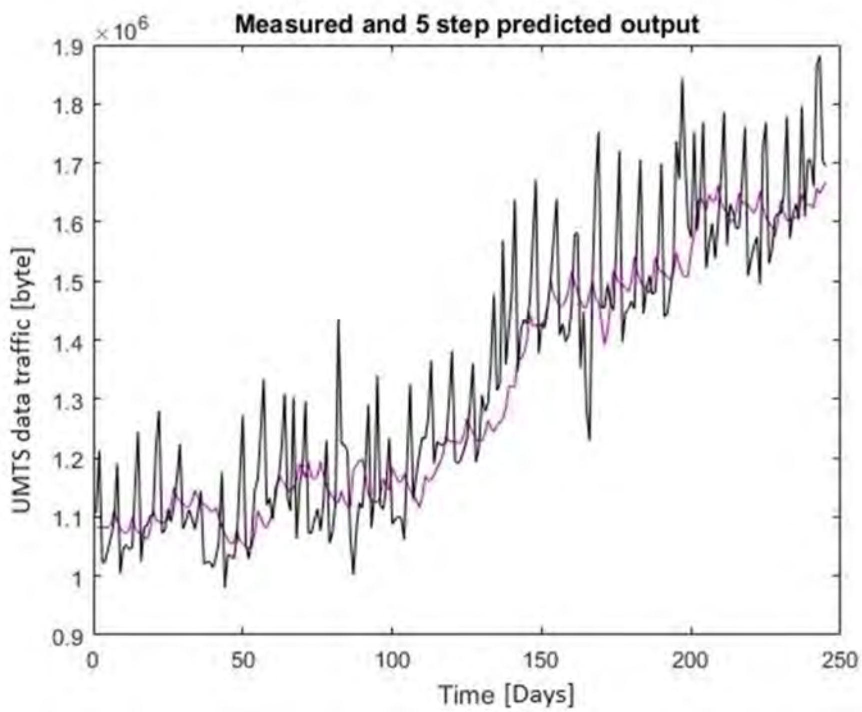


Figure 6-14: Real Data vs. Forecasted data in ARIMA (1, 1, 1) x (0, 1, 1)7.

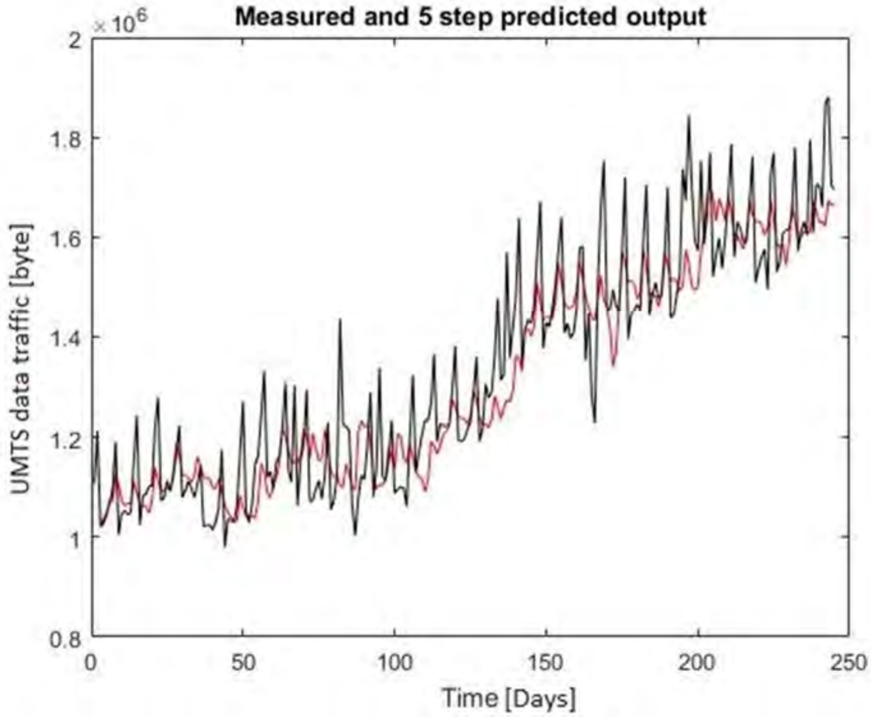


Figure 6-15: Real Data vs. Forecasted data in ARIMA (5, 1, 1) x $(0, 1, 1)_7$.

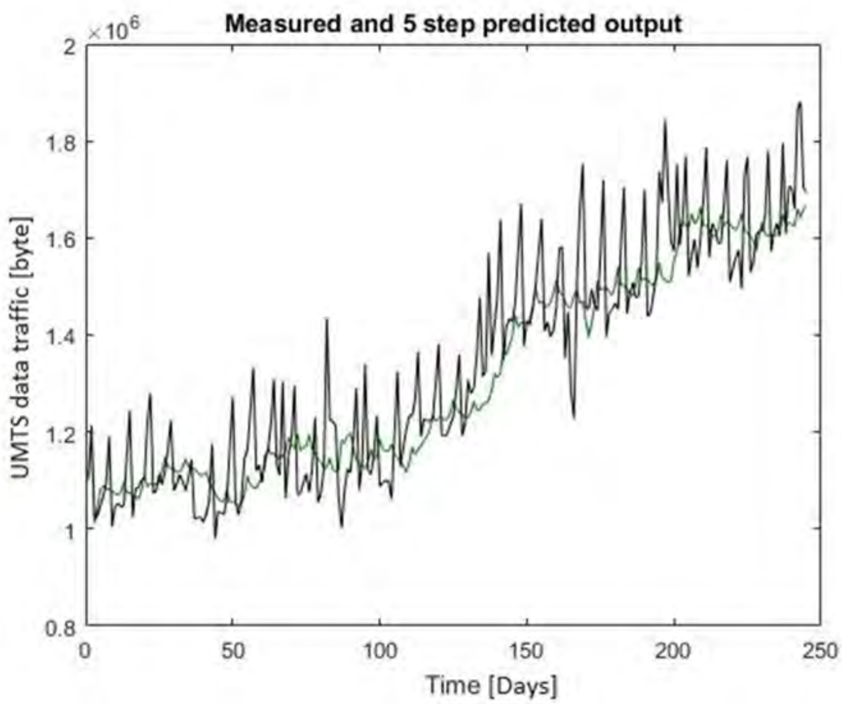


Figure 6-16: Real Data vs. Forecasted data in ARIMA (3, 1, 1) x $(0, 1, 1)_7$.

- The last stage is **forecasting stage**, the forecast statement to forecast future values of the time series and to generate confidence intervals for these forecasts from the ARMA model produced by the previous stage.

Step 7: Forecast the time series for short term of one month and long term of one year.

In the figure below, it is observed that by using SARIMA (1,1,1)x(0,1,1)$_7$, the forecasted data have a periodical pattern of7 days with an increasing trend and the traffic tends to reach the peak value on the second day of the week.

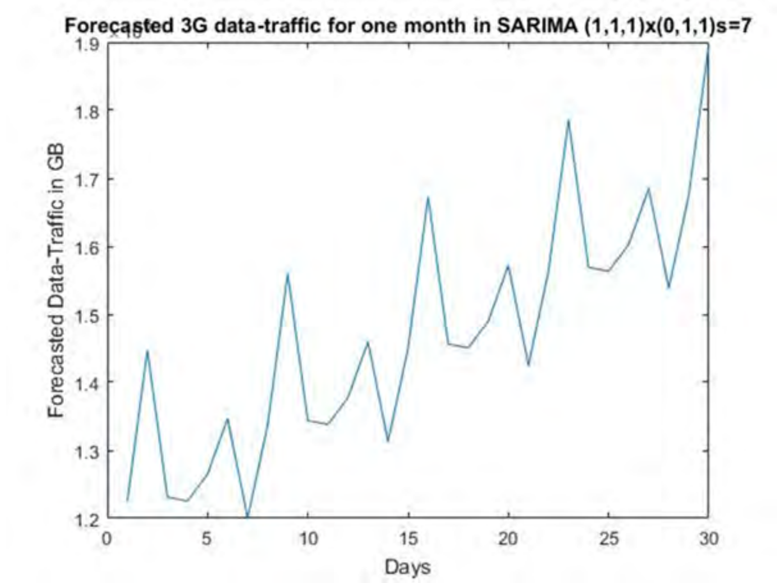


Figure 6-17: UMTS Data-Traffic Forecasted data for one month in

SARIMA (1, 1, 1) x (0, 1, 1)$_7$.

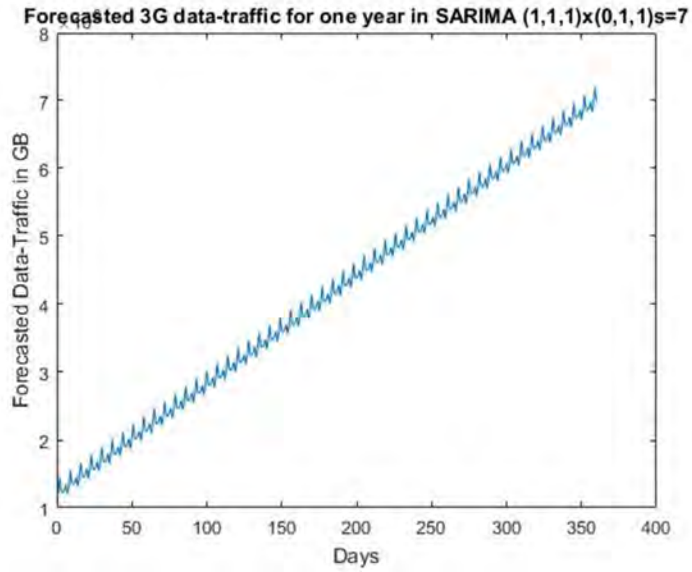


Figure 6-18: UMTS Data-Traffic Forecasted data for one year in

SARIMA (1, 1, 1) x (0, 1, 1)$_7$

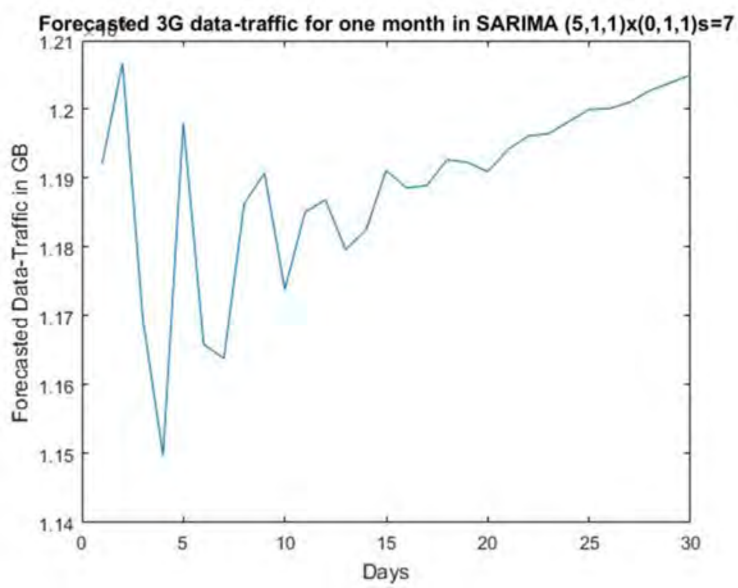


Figure 6-19: UMTS Data-Traffic Forecasted data for one month in

SARIMA (5, 1, 1) x (0, 1, 1)$_7$

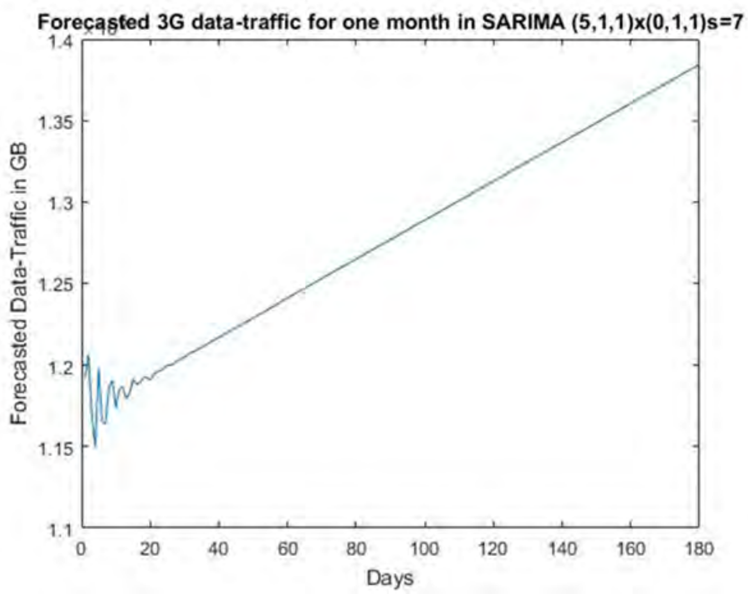


Figure 6-20: UMTS Data-Traffic Forecasted data for one year in

SARIMA (5, 1, 1) x (0, 1, 1)$_7$

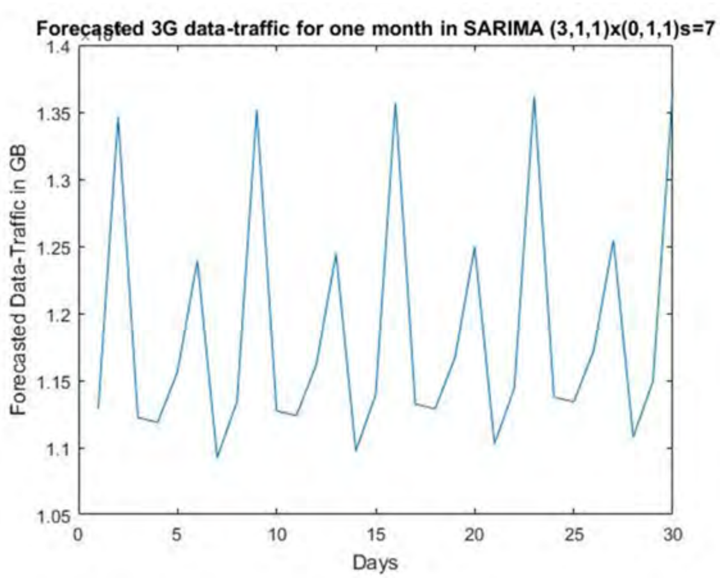


Figure 6-21: UMTS Data-Traffic Forecasted data for one month in

SARIMA (3, 1, 1) x (0, 1, 1)$_7$

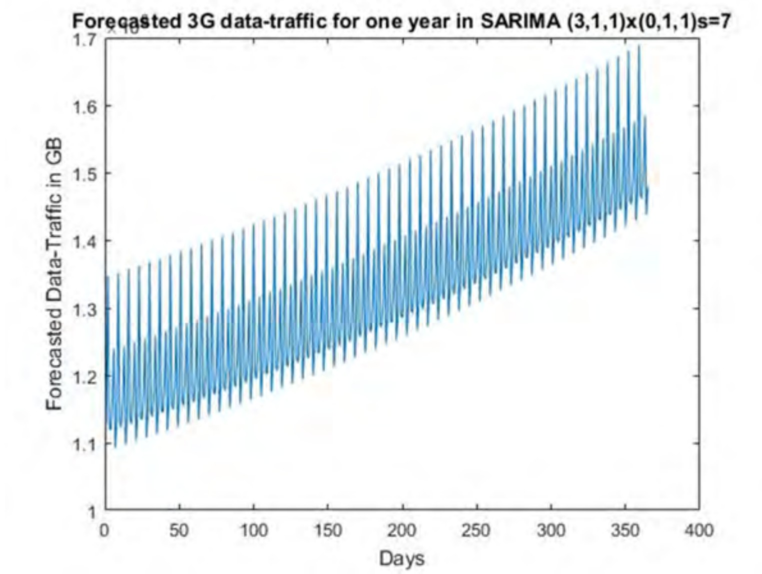


Figure 6-22: 3G Data-Traffic Forecasted data for one year in

SARIMA (3, 1, 1) x (0, 1, 1)$_7$

After predicting the UMTS data-traffic load for eachuser per month in SARIMA (1, 1, 1) (0, 1, 1)$_7$,the last step is to consider the traffic model table and arrange the input for planning the future infrastructure of the existing 3G network or to implement and widen the next generation of the network(4G-LTE).

Considered traffic model:

Traffic Usage in GB/Month/User	**UMTS Data-Traffic model**	
	Dongle	**Smart Phone**
Distribution percentage	37%of subscribers	63% of subscribers
Traffic per user per month in GB	10	2

Table 6-4: Input of UMTS Traffic Model from one month forecasted data.

Taking the operator traffic Model [27] as an input, the capacity calculation is done as follows:

Step 1: Data volume per month per user calculation

10GB*37%+2GB*63% =4.96GB

Step2: HSPA throughput per user calculation

Per day per user: "Step1" *8*1024*1024/30/3600 =477.69Kbps

Busy hour is 10%;

Busy hour per user =477.69*10% =47.77Kbps

Step3: Busy hour per active user calculation

In Ethio-telecom active user ratio is assumed to be 70%;

Busy hour per active user 47.77*70%=33.74Kbps.

To design the maximum traffic usage, the other alternative is way to use 100% of smart phone instead of using dongle. It helps customers use the unlimited data-traffic rate and anyone can access the Internet services .Therefore the traffic model table will be as shown below.

Traffic Usage in GB/Month/User	Data-traffic model
Distribution percentage	100% of subscribers
Traffic per user per month in GB	7

Table 6-5: Input of UMTS Traffic Model from one month forecasted data

Taking the operator traffic Model as an input, the capacity calculation is done as follows:

Step1: Data volume per month per user calculation

7GB*100%=7GB

Step2: HSPA throughput per user calculation

Per day per user: “Step1” *8*1024*1024/30/3600 =652.44Kbps

Busy hour is 10%;

Busy hour per user =652.44*10% =65.24Kbps

Step3: Busy hour per active user calculation

Active user ratio is assumed to be 70%;

Busy hour per active user 65.24*70%=45.67Kbps.

7. CONCLUSION and RECOMMENDATION

7.1 Conclusion

The aim of this study was to evaluate two Univariate time series methods, namely SARIMA and ARIMA, from underlying time series analysis method to predict the UMTS data traffic. The intention was to find a model that fits well to the data and could forecast the UMTS data-traffic in contrast to the forecasting performed today.

This was performed by first estimating differencing to remove both the seasonal and trend components and ARIMA models to fit UMTS data-traffic time series. When performing the time series analysis procedure the mean and variance are subtracted to make the time series stationary. Exponential models as well as ARIMA models were fit in the time series analysis and the best performing models were selected by looking at error terms and an information criterion. The best performing models were used to the forecast the UMTS data traffic. Adding the seasonal component and the trend helped the time series to be estimated with better accuracy.

With the results of the modeling performed for the study, the following conclusions can be drawn:

- The forecasting results in general revealed an increasing pattern over the forecasted period (one month and a year).In light of the forecasted results, Ethio-telecom could gain insight into the expected data traffic over the coming months and plan a more appropriate network infrastructure.

- Compared to ARIMA, the SARIMA model (ARIMA with seasonality) performed well for every forecasting period, according to the pattern. The plotted UMTS forecasted data-traffic seemed to exhibit a seasonal pattern of every week but ACF indicated a possible seasonal component on the 7th day of the week. When fitting the SARIMA (1, 1, 1) x $(0, 1, 1)_7$ models with this seasonality, it fit the UMTS data-traffic with low error terms, and a lower value of BIC.
- Seasonal component exists in data-traffic flow, which can be removed by one-step difference of the original data. The differenced UMTS data-traffic flow data are one-step correlated. In other word, the increase or decrease of data-traffic flow can influence the change of data in the next time step. The prediction of UMTS data-traffic can be made simpler by studying the differenced original data-traffic.
- Based on minimum BIC, the best-fitter SARIMA models tends to be SARIMA(1,1,1)$(0,1,1)_7$ and (5,1,1)$(0,1,1)_7$ model. After the estimation of the parameters of selected model, a series of diagnostic and forecast accuracy test were performed. Having satisfied all the model assumptions, ARIMA (1, 1, 1) $(0, 1, 1)_7$ model was judge to be the best model for forecasting UMTS data-traffic.
- When fitting SARIMA models to the UMTS data-traffic peak load, the SARIMA(1,1,1)X$(0,1,1)_7$ had a better fit than the others with minimum error , minimum BIC value and best fit, but also the SARIMA(5,1,1)X$(0,1,1)_7$ model had a good fit.

To evaluate the model for use in forecasting UMTS time series, it is necessary to consider quantitative measures of forecasting accuracy. In this thesis, a traditional error criterion such as the mean square error is found to be useful to assess the forecasting performance of both ARIMA and SARIMA models. The proper choice of performance criterion will depend on the particular application for which the model is being considered.

The attractiveness of the SARIMA model is mainly due to its flexibility to represent several varieties of time series. Having such a model greatly enhances the process of working towards a better customer satisfaction and carefully planned customer expansion and use of investment.

7.2 Recommendation for Future Works

This study addresses the UMTS data-Traffic flow forecasting model, accuracy and the errors. Different methods such as the AR model, MA model, ARMA model, ARIMA model and SARIMA model are proposed here to forecast UMTS data-traffic flow but mainly the last two models were evaluated.

It is a goal for Ethio-telecom to have their system operates more efficiently, securely and economically. To meet this goal, the behavior of the UMTS mobile data-traffic load must be well understood.

The UMTS mobile data-traffic Peak load forecasting can be used to assist in planning how much the goals of the traffic demand is managed. If the difference between the peak load data-traffic forecast and the desired maximum peak load is known, this information could be used to schedule for the future.

The seasonal ARIMA model presented in this thesis can be represented in time series analysis form and can therefore be implemented using Matlab to acquire and update the parameter estimates. Follow-on research is needed to develop and test the validity and efficiency of such a seasonal ARIMA-based on different approach.

At minimum, future research on alternate univaritate forecast approaches should include comparisons to seasonal ARIMA and heuristic forecasting performance. The ARIMA (1, 1, 1) (0, 1, 1) $_s$ model provides a simple three-parameter linear recursive estimator.

Therefore, an appropriately applied seasonal ARIMA model should be considered the parametric model benchmark for univariate traffic condition forecasting. Likewise, the deviation from the historical average prediction method described in this thesis provides good forecasts and should be considered a key heuristic forecast benchmark.

Time series forecasting is a fast growing area of research and as such provides many scope for future works. One of them is the Combining Approach, i.e., to combine a number of different and dissimilar methods to improve forecast accuracy.

Appendix

The results for selected model are shown in the table below:

Forecasted Day	Data Traffic in GB	Error in MSE
1	1.2247	0.0013
2	1.4474	0.0068
3	1.2311	0.0136
4	1.2254	0.0208
5	1.2647	0.0279
6	1.3468	0.0351
7	1.2002	0.0423
8	1.3374	0.0569
9	1.56	0.0823
10	1.3437	0.1103
11	1.3381	0.139
12	1.3773	0.1678
13	1.4594	0.1966
14	1.3128	0.2254
15	1.45	0.2676
16	1.6726	0.3272
17	1.4564	0.3909
18	1.4507	0.4555
19	1.4899	0.5203

20	1.5721	0.5852
21	1.4255	0.65
22	1.5626	0.7344
23	1.7853	0.8426
24	1.569	0.9564
25	1.5634	1.0713
26	1.6026	1.1865
27	1.6847	1.3018
28	1.5381	1.4171
29	1.6753	1.558
30	1.8979	1.7293

Reference

1. **H. Holma** and **A.Toskala**, "UMTS services and applications," in WCDMA for UMTS', Third Ed. England: John Wiley and Sons, 2004, PP.11-47.
2. **Guangdong**, "WCDMA System overview", Huawei Tech.co, Shenzhen, 2004, pp. 1–12.
3. **Guangdong**,"Radio network planning for WCDMA", Huawei Tech.co, Shenzhen, 2004, pp. 1–37.
4. **M.Farooq, M.Ahmed** and **U.Mohammed**,"Future Generations of Mobile Communication Networks", Vol.2, No.1, Jan 2013.
5. **P.Chen**,"ARIMA Based Time Series Model of Stochastic Wind Power Generation", IEEE Transaction on Power System, Vol.25, No.2, June, 2010.
6. **P.Svoboda, M.Buerger and M.Rupp**, "Forecasting of traffic load in a live 3G packet Switched Core Network'', Institute of Communication and Radio-Frequency Engineering, ,University of Technology, Vienna ,Austrila,2004.
7. **H. Alfred**, 'Forecasting Activates Using ARIMA model ", Journal of Advances in Computer Network, Vol.2, No.3, September 2014.
8. **S.Valentin, C. Liu** and **M.Tangemann**,'' Predicting the number of Mobile Subscribers: An accurate Forecasting System and its Application' 'in Proc.IEEE WCNC, April, 2012.
9. **S.Makridakis** and **M.Hibon**,''ARMA model and the Box-Jenkins Methodology' 'Journal offorecasting.Vol.16, No.3, January 1997, pp.147-263.

10. **K.Reagan**, ''An Evaluation of ARIMA (BOX-Jenkins) models for Forecasting Wastewater Treatment Process Variables", M.Sc thesis, University of California, Los Angeles, 1984.
11. **P.Brockwell and R.Davis**, ''Introduction to time series and forecasting'', Second Edition, 2002, Springer-Verlag New York, Inc, New York, pp-45-57.
12. **H.M.Taylor** and **S.Karlin**,v''An introduction to Stochastic Modeling'', San Diego, California, Third edition, 1998, pp-1-53.
13. **A.Papoulis** and **S.Pillai**, "Probability, Random Variables, and Stochastic Process'', Fourth Edition, 2002.
14. **S.Green**,"Time series Analysis of Stock Prices Using the Box-Jenkins Approach", MSc. thesis, Georgia Southern University, 2011.
15. **D.S**.**Walonick,Ph.D**, "An Overview of Forecasting Methodology", 1993.
16. **Wangdong Fu**, "ARIMA model for forecasting Poisson data: applications for long term earthquake predictions", University of Nevada, Las Vegas, August 2010.
17. **G .Schwarz**. ''Estimating the dimension of a model''. Annals of Statistics 6, 1978, pp-461–464.
18. **J.Wiley**,"UMTS Radio Access Network" in "UMTS Networks-Architecture, Mobility and services", 2001.PP-61-91.
19. **T.Balderas** and**R. A.Cumplido**,"Security Architecture in UMTS Third Generation Cellular Network", 2004.
20. **J.Wiley**,"UMTS Core Network" in"UMTS Networks-Architecture, Mobility and services", 2001.PP-103-126.
21. **J.Wiley**,"UMTS Terminal" in"UMTS Networks-Architecture, Mobility and services", 2001.PP-129-139.
22. **A.Mitschele**,"UMTS system Architecture and Protocol Architecture", 2002.

23. **L-F Pau**,Prof, "Summary introduction to wireless LTE 4G architecture and key business implications ",2011.
24. **Ye Ouyang** and **M.Hosien**,"Performance Analysis for UMTS packet switched network based on multivariate KPIs", International Journal of Next Generation Network, vol.2, March 2010.
25. **V.Britvic**,"Steps in UMTS network Design," 12th IEEE Mediterranean Electro technical Conference.Glasgow.UK. 2004.
26. **M.Fyro,K.Heikkinen**, **L.Petersen**, and **P.Wiss**,"Media getaway for mobile network", Ericsson,2000.
27. **R.Martinez**,"3G system -WCDMA (UMTS) and CDMA 2000." February 28, 2005.
28. **L.Bramer**,"Methods for Modeling and Forecasting Wind Characteristics ", MSc. Thesis, Iowa University, 2013.
29. **F.Mergia**,"UMTS Radio Network Planning for the case of Addis Ababa City", MSc. Thesis, Addis Ababa University, 2016.
30. **T. Raicharoen**, "Application of critical support vector machine to time series prediction", Circuits and Systems, 2003. ISCAS '03.Proceedings of the 2003 International Symposium on Volume 5, 25-28 May, 2003, pages: V-741-V-744.
31. **L.Mitikie**,"UMTS Coverage and Capacity Planning for the case of Bole Sub-City in Addis Ababa", MSc. Thesis, Addis Ababa University, 2016.
32. **R.Adhikari**,"An introductory study on time series Modeling and Forecasting", Msc.thesis, 2002.
33. **Chatfield C**.,"The Analysis of Time Series", Seventh edition 2, CRC press LLC Florida, 2004.
34. **J.Wiley**, ''Wiles series in Probability and Statistics: Introduction to time series analysis and forecasting", John Wiley & Sons.Inc. Hoboken, New Jersey 2008.