



**Bruno Miguel Machado** **Técnicas de Previsão: Aplicação a Redes Celulares**  
**Carvalho**

**Forecasting Techniques: Application to Cellular  
Networks**



**Bruno Miguel Machado Técnicas de Previsão: Aplicação a Redes Celulares  
Carvalho**

## **Forecasting Techniques: Application to Cellular Networks**

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Eletrónica e Telecomunicações (Mestrado Integrado), realizada sob a orientação científica do Professor Dr. Aníbal Manuel de Oliveira Duarte do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro.

## **O Júri**

### **Presidente**

Professor Doutor José Rodrigues Ferreira da Rocha

Professor Catedrático, Universidade de Aveiro

### **Vogal – Arguente externo**

Professora Doutora Isabel Maria Simões Pereira

Professora Auxiliar da Universidade de Aveiro – Departamento de Matemática

### **Vogal - Orientador**

Professor Doutor Aníbal Manuel de Oliveira Duarte

Professor Catedrático da Universidade de Aveiro

## **Agradecimentos**

Quero, desde já, agradecer ao professor Manuel de Oliveira Duarte pela disponibilidade, orientação e conhecimento que me foi transmitido ao longo deste trabalho, fundamentais para a sua conclusão.

Agradecer à minha família, principalmente aos meus pais, pelo apoio, carinho e paciência que sempre tiveram.

Por fim, um agradecimento especial à Raquel pelo apoio incondicional em todos os momentos.

**palavras-chave**

UMTS, Arquitetura, Previsão, Otimização, KPIs, Monitorização.

**resumo**

Devido ao aumento da competitividade e agressividade do mercado, cada vez mais a estratégia dos operadores de redes móveis passa pelo melhoramento das infraestruturas e otimização dos recursos já existentes, de modo a proporcionar a melhor experiência aos seus utilizadores. Para isto, recorrem à análise de indicadores chave de desempenho (KPIs) e ao uso de métodos de previsão para prever e planear alterações a realizar na sua rede.

Tendo isto como base, esta dissertação foca-se no estudo e análise de diferentes métodos de previsão e sua implementação em Python, de maneira a obter previsões do futuro comportamento da rede em tempo real e de forma automatizada.

**keywords**

UMTS, Architecture, Forecasting, Optimization, KPIs, Monitoring.

**abstract**

Due to the growing competitiveness and aggressiveness of the market, the network operators' strategy is increasingly based on improving infrastructures and optimizing existing resources, in a way that provides the best experience to the user. To do this, the operator analyzes the Key Performance Indicators (KPIs) and uses forecasting methods to predict and plan the modifications needed in the network.

With this as basis, this work focuses on the study and analysis of different forecasting methods and their implementation in Python, so that the operator can obtain automate real-time predictions of the future behavior of his network.

# Index

Figures Index.....	ix
Tables Index.....	x
Acronyms.....	xi
1. Introduction.....	1
1.1 Motivation.....	2
1.2 Objectives.....	2
1.3 Organization of the Dissertation.....	2
2. Mobile Network.....	3
2.1 Mobile Network Evolution overview.....	3
2.2 GSM Architecture.....	5
2.3 UMTS Architecture.....	8
2.4 LTE Architecture.....	18
2.5 Network Management.....	25
2.6 Key Performance Indicators.....	27
3. Forecasting.....	28
3.1 Telecommunication Forecasting.....	28
3.2 Forecasting Categories.....	29
3.3 Time Series Analysis.....	31
3.4 Forecasting Models.....	34
3.5 Auto Regressive Models.....	40
3.6 Moving Average Models.....	41
3.7 Auto Regressive Moving Average Models.....	43
3.8 Forecast Validation.....	46
3.9 Forecast Evaluation Methods.....	48
3.10 Modelling ARIMA procedure.....	51
3.11 ACF and PACF.....	52
3.10 Fourier Transform.....	57
4. Analysis and application of Forecasting in Python.....	61
4.1 Weekly data analysis and Forecast - CSFF.....	61

4.2 Weekly data analysis and Forecast - PSFF.....	68
5. Conclusion.....	75
6. Future Work.....	75
References.....	76



## Figures Index

Figure 1 - Global mobile traffic from "On Pulse of the Networked Society.....	1
Figure 2 - Mobile Network evolution.....	3
Figure 3 - Multiple Access Techniques.....	5
Figure 4 - GSM Architecture.....	5
Figure 5 - UMTS Architecture.....	9
Figure 6 - UTRAN components and interfaces.....	11
Figure 7 - 3GPP network implementation on R99.....	13
Figure 8 - 3GPP network implementation on R4.....	14
Figure 9 - 3GPP R5 implementation .....	15
Figure 10 - 3GPP IMS Architecture.....	15
Figure 11 - 3GPP R5 implementation.....	18
Figure 12 - 3GPP LTE implementation (Release 8).....	20
Figure 13 - MME logical connections and functions.....	21
Figure 14 - S-GW logical connections and functions.....	22
Figure 15 - PDN-GW logical connections and functions .....	23
Figure 16 - PCRF logical connections and functions.....	23
Figure 17 - Total number of airline passengers over a period of time.....	31
Figure 18 - Trend component of the time series of airline passengers .....	32
Figure 19 - Seasonal component of the time series of airline passengers .....	32
Figure 20 - Irregular component of the time series of airline passengers .....	32
Figure 21 - Scatter plot example.....	33
Figure 22 - Application of basic methods .....	36
Figure 23 - Australian beer production.....	38
Figure 24 - Total number of airline passengers over a period of time.....	39
Figure 25 - p-order IIR filter implementing an AR model.....	41
Figure 26 - q-order FIR implementing MA model.....	42
Figure 27 - ARMA (p,q) model implemented by a FIR/IIR.....	44
Figure 28 - Autocorrelation of square finite pulse.....	53
Figure 29 - Simulated MA(1).....	54
Figure 30 - ACF and PACF of the MA(1).....	55
Figure 31 - Simulated AR(1).....	55
Figure 32 - ACF and PACF of AR(1).....	56

Figure 33 - Illustration of Fourier Transform.....	57
Figure 34 - Simulated time series.....	58
Figure 35 - FFT amplitude.....	59
Figure 36 - FFT phase.....	59
Figure 37 - Plots of weekly CSFF data.....	62
Figure 38 - ACF plot of weekly aggregated CS fill factor.....	62
Figure 39 - Weekly CSFF after first difference.....	63
Figure 40 - ACF of first difference of weekly CSFF.....	64
Figure 41 - PACF of first difference of weekly CSFF.....	65
Figure 42 - RNC1 weekly CSFF real value (blue line) vs forecast (orange line).....	66
Figure 43 - RNC2 weekly CSFF real value (blue line) vs forecast (orange line).....	66
Figure 44 - RNC3 weekly CSFF real value (blue line) vs forecast (orange line).....	67
Figure 45 - RNC4 weekly CSFF real value (blue line) vs forecast (orange line).....	68
Figure 46 - Weekly PSFF of four RNCs.....	68
Figure 47 - ACF plot of weekly aggregated PS fill factor.....	69
Figure 48 - PS fill factor after first difference.....	70
Figure 49 - ACF of first differences of weekly PSFF.....	71
Figure 50 - PACF of the first difference of weekly PSFF.....	72
Figure 51 - RNC1 weekly PSFF real value (blue line) vs forecast (orange line).....	73
Figure 52 - RNC3 weekly PSFF real value (blue line) vs forecast (orange line).....	73
Figure 53 - RNC4 weekly PSFF real value (blue line) vs forecast (orange line).....	73
Figure 54 - RNC2 weekly PSFF real value (blue line) vs forecast (orange line).....	74

## Tables Index

Table 1.....	9
Table 2.....	17
Table 3.....	18
Table 4.....	19
Table 5.....	25
Table 6.....	27
Table 7.....	30
Table 8.....	56
Table 9.....	63
Table 10.....	64
Table 11.....	65
Table 12.....	66
Table 13.....	67
Table 14.....	67
Table 15.....	70
Table 16.....	72
Table 17.....	74

## Acronyms

1G	1st Generation
2G	2nd Generation
3G	3rd Generation
3GPP	3rd Generation Partnership Project
ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
AIC	Akaike's Information Criteria
AICc	Corrected Akaike's Information Criteria
AR	Auto-Regressive
ARIMA	AutoRegressive Integrated Moving Average
ARMA	Auto Regressive Moving Average
ARPU	Average revenue per user
AuC	Authentication Center
BS	Base Station
BSC	Base Station Controller
BSS	Base Station Subsystem
BTS	Base Transceiver Station
CAPEX	Capital Expenditure
CDMA	Code Division Multiple Access
CN	Core Network
CS	Circuit Switch
CSFF	Circuit Switched Fill Factor
E-DCH	Enhanced Dedicated Channel
EDGE	Enhanced GPRS
EIR	Equipment Identity Register
eTom	Enhanced Telecom Operations Map
E-UTRAN	Evolved UMTS Terrestrial Radio Access Network
FIR	Finite Response Filter
FM	Frequency Modulation
GERAN	GSM Edge Radio Access Network
GGSN	Gateway GPRS Support Node
GMSC	Gateway Mobile Switching Center

GRAN	GSM Radio Access Network
GSM	Global System for Mobile communications
H-ARQ	Hybrid Automatic Repeat Request
HLR	Home Location Register
HSDPA	High Speed Downlink Packet Access
HSS	Home Subscriber Server
HSUPA	High Speed Uplink Packet Access
I-CSCF	Interrogating-Call Session Control Function
IIR	Infinite Response Filter
IMEI	International Mobile Equipment Identity
IMS	IP Multimedia Subsystem
KPI	Key Performance Indicator
LTE	Long Term Evolution
MA	Moving Average
MAE	Mean of Absolute errors
MAPE	Mean absolute percentage errors
ME	Mobile Equipment
MIMO	Multiple Input Multiple Output
MRFC	Media Resource Function Controller
MRFP	Media Resource Function Processor
MS	Mobile Station
MSC	Mobile Switching Center
MSE	Mean of Squared errors
NE	Network Element
NPV	Net Present Value
OFDMA	Orthogonal Frequency Division Multiple Access
OPEX	Operation Expenditure
PACF	Partial Autocorrelation Function
P-CSCF	Proxy-Call Session Control Function
PLMN	Public Land Mobile Network
PS	Packet Switch
PSFF	Packet Switched Fill Factor
RAN	Radio Access Network
RNC	Radio Network Controller
S-CSCF	Serving-Call Session control function

SIM	Subscriber Identity Module
SIP	Session Initiated Protocol
SLC	Service Life-cycle
SGSN	Service GPRS Support Node
SLF	Subscription Locator Function
SMA	Simple Moving Average
TDMA	Time Division Multiple Access
TMN	Telecommunication Management Network
UE	User Equipment
UMTS	Universal Mobile Telecommunication System
UTRAN	UMTS Terrestrial Access Network
VLR	Visitor Location Register
W-CDMA	Wideband Code Division Multiple Accessed
WLAN	Wireless Local Area Network

## 1. Introduction

The world population is increasingly getting more dependent on technology. Nowadays, the smartphones and tablets are taking over the market, meaning that we do not only need voice oriented networks, but also a variety of data services. The Universal Mobile Telecommunication System (UMTS) is what brought the operators the possibility of providing this kind of data services, and the traffic trend analysis show that the data services usage will grow exponentially, making this subject the focus of network operators.

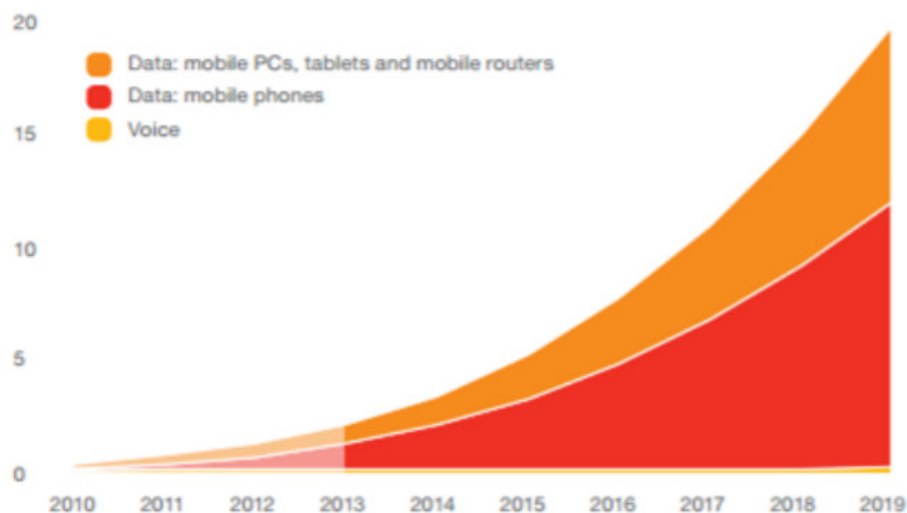


Figure 1 - Global mobile traffic from "On Pulse of the Networked Society"

The increasing diversity of services with multiple requirements leads mobile operators to have difficulties in managing the performance of their networks

Facing this scenario, the network operators strategy to deal with this growth in data variety and services usage is to invest in the analysis and forecasting of the network performance, maximizing the capacity of the current resources.

## **1.1 Motivation**

The motivation is to study and apply forecasting mechanisms to data from a real communication service provider by developing a forecasting tool using programming language Python that forecasts the network capacity performance, allowing the mobile service provider to prevent future network issues.

## **1.2 Objectives**

The main objective of this work is to develop a forecasting tool in Python, so that networks operators can predict and prevent network capacity issues. Using Python opens the possibility to improve the automatization of the forecasting process, reducing the need of human interaction (e.g. to identify the model orders) and the mobile network provider response to a network problem. As this work is focused in mobile networks, starts by getting acquainted with the different mobile network architecture, with enhanced focus on UMTS and his KPIs, since the data provided comes from a RNC. The forecasting methods are based in mathematical methods, so it takes to study the different methods and analyze which one should be used in each situation.

## **1.3 Organization of the Dissertation**

Besides the introduction chapter, this dissertation is composed by for main 4 chapters:

- Chapter 2 introduces the mobile networks architectures. After this, it presents the KPIs and discusses its importance in network managing;
- Chapter 3 introduces the concept of forecasting, starting by some basic concepts for better understanding the methods discussed later in chapter;
- Chapter 4 is about the programming language Python and how to use it to forecast network capacity;



## 2. Mobile Network

In this chapter is presented an overview of mobile networks, its features and evolution, with enhanced focus in the UMTS and LTE architectures. Furthermore, the concepts of network management and KPIs are approached.

### 2.1 Mobile Network Evolution overview

The evolution of mobile network has been categorized in different generations, term used to refer fundamental technological implementations and transformations that occur in the type and characteristics of a specific service, non-backwards compatible with technology and new frequency bands of a particular network system [31]. In Figure 2 are shown some of the important changes between generations.

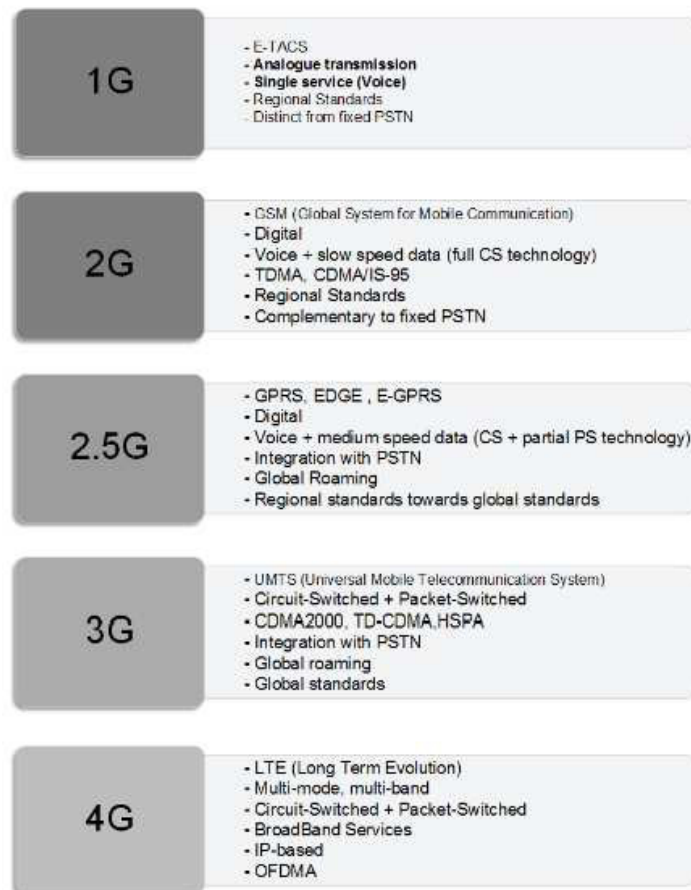


Figure 2: Mobile Network evolution

### 2.1.1 Radio Access Network

The RAN is one of the most important components of a mobile telecommunication network, being responsible for providing a connection between the different users devices and the network through radio connections. Given this, the evolution of RAN marked each mobile network generation:

- GERAN: when EDGE is not present it is called GRAN but they are conceptually identical. This is the GSM (2G) RAN and is composed by the BS and BSC;
- UTRAN: appears with UMTS (3G) to work with the W-CDMA technique. Similar to GERAN, it also has 2 components but it the BS and BSC are substituted by the RNC and Node B;
- E-UTRAN: present in LTE, this RAN comes satisfy the need for higher data rates and even more efficient spectrum usage. Here, the only component of the RAN is the eNodeB.

Since the UMTS started as a GSM-based architecture, the GERAN and the UTRAN can coexist and interwork, allowing the 2G networks radio interface to use the 3G core network. This implied an optimization in both radio technologies and reduced the need of additional instalations which lowered the cost of the 'transition' from 2G to 3G.

### 2.1.2 Multiple Access Techniques

Multiple access techniques allows the RAN to make a more efficient division of the spectrum available among the multiple users. Depending on the type of channel, the techniques used are [18]:

- Time-slot within Frequency Bands: TDMA is the technique used in 2G which allows every user to transmit using a common frequency band during a certain amount of time, allocating a specific timeslot for each call;
- Distinct Codes: CDMA identifies every user by its assigned code, allowing all user to transmit at the same time in the same frequency band improving the efficiency of radio resources usage and providing higher data rates. The CDMA (W-CDMA) is used in 3G.
- Frequency Channels: consists of dividing the spectrum into small frequency channels that are assigned to different users. One example is the FM radio, where each radio station has

is exclusive frequency. In 4G, the OFDMA is used for downlink and the SC-FDMA for uplink.

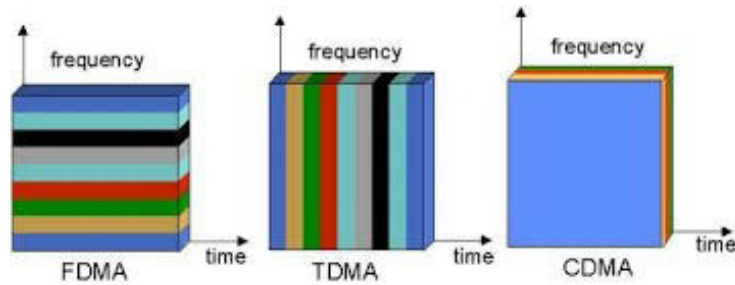


Figure 3: Multiple Access Techniques

Figure 3 is an illustration of the way the different techniques use the spectrum to allocate multiple users.

## 2.2 GSM Architecture

The mobile networks exploded in the beginning of the 90s with the GSM (2G) networks, reaching their cruise speed during the second half of the 90s and start becoming a major daily need by the end of the 90s [32]. Figure 3 depicts a general GSM architecture.

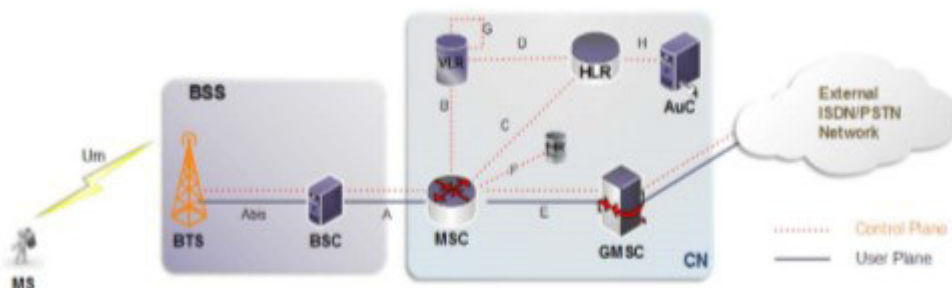


Figure 4: GSM Architecture

The network system can be divided into three main elements:

- Mobile Station (MS);

- Base Station Subsystem (BSS);
- Core Network (CN);

### **A. Mobile Station (MS)**

Generally, the MS is the device used by a subscriber to communicate with the GSM network. It represents the mobile part of the network and is divided into Subscriber Identity Module (SIM) and Mobile Equipment (ME): SIM is used to allow mobility and subscription of services to the user and ME consists on a device used to connect to the network. Some functions of the MS are:

- Speech to digital conversion to the BTS and digital to speech conversion from the BTS;
- Radio Signal encryption;
- Monitoring the signal quality from BTS and report back.

### **B. Base Station Subsystem (BSS)**

The BSS contains all necessary equipment, hardware and software to connect the MS to the CN and is responsible for all radio path control. It is composed by two elements:

- **Base Transceiver Station (BTS):** radio equipment (transceivers and antennas) necessary to provide network coverage to a designated area and allows the MS to connect and communicate with the network through the radio interface. The BTS functions include [28]:
  - Uplink measurements and advanced timing calculations;
  - Channel encryption and decryption and speech processing;
  - Broadcast and channels control.
- **Base Station Controller (BSC):** the BSC controls all the functionality of the BTS attached to it. The quantity of BTS connected to a BSC depends on the network configuration. Some of the main functions are [28]:
  - Manage of radio resources of the BTS, including frequency allocation;
  - Sustain radio connections to the MS and terrestrial connections to the CN;
  - Power control;
  - Intercell handovers.

## **C. Core Network**

The Core Network consists in a group of elements responsible for calls management, allowing mobile devices to communicate with each other and with telephones.

### **Mobile Switching Center (MSC)**

The MSC is the central element of the CN when referring to control the services provided to the subscribers. The MSCs are used to interface GSM networks to fixed networks and bridge the core network to the access network. Some of its functions are:

- Mobility management of the subscribers;
- Manage connections between subscribers;
- Call control and billing information gathering;
- BSS control functions;
- Handover procedures (e.g. change to other MSC)

### **Home Location Register (HLR)**

The HLR is a database that stores information about the subscribers within its network, holding the following information:

- User identification, addressing and security information for authentication and authorization;
- User location at inter-system level;
- Handle the user registration and stores inter-system location information;
- User profile.

### **Visitor Location Register (VLR)**

The VLR is a database that contains information about the current MSs in its area. Unlike the HLR, the information in the VLR is dynamic. When a MS enters a MSC area and tries to connect with the network, the MSC is informed of its arrival and its associated VLR is updated. After this, a

message is sent to the HLR informing that the VLR contains the location of the UE. The VLR stores data such as:

- Subscriber location;
- Identification numbers of the mobile subscribers;
- Services that subscribers are allowed to use;
- Security information for authentication purposes.

### **Authentication Center (AuC)**

The AuC cooperates closely with the HLR database, and is responsible for the identification of the MS and the network, encryption keys generation and regulation of integrity of the transmission.

### **Equipment Identity Register (EIR)**

The EIR is a database with the International Mobile Equipment Identity (IMEI) of a given MS for security purposes.

## **2.3 UMTS Architecture**

The UMTS architecture has several features implemented based on GSM core network and its radio access network, allowing 2G to be used for UMTS. However, while the GSM uses TDMA transmissions, UMTS uses CDMA which has its physical layer referred as wideband CDMA (W-CDMA). The W-CDMA as a wider band compared to the common CDMA, with the advantage of efficient utilization of radio spectrum giving a higher maximum data rate [33]. In the next table are presented the specifications of the initial format.

### 3G UMTS Specification

Parameter	Specification
Maximum data rate	2048kbps low range
RF channel bandwidth	5MHz
Multiple access scheme	CDMA
Duplex schemes	FDD and also TDD

The UMTS network system can be divided into three logical components:

- Core Network;
- Radio Access Network;
- Mobile Equipment.

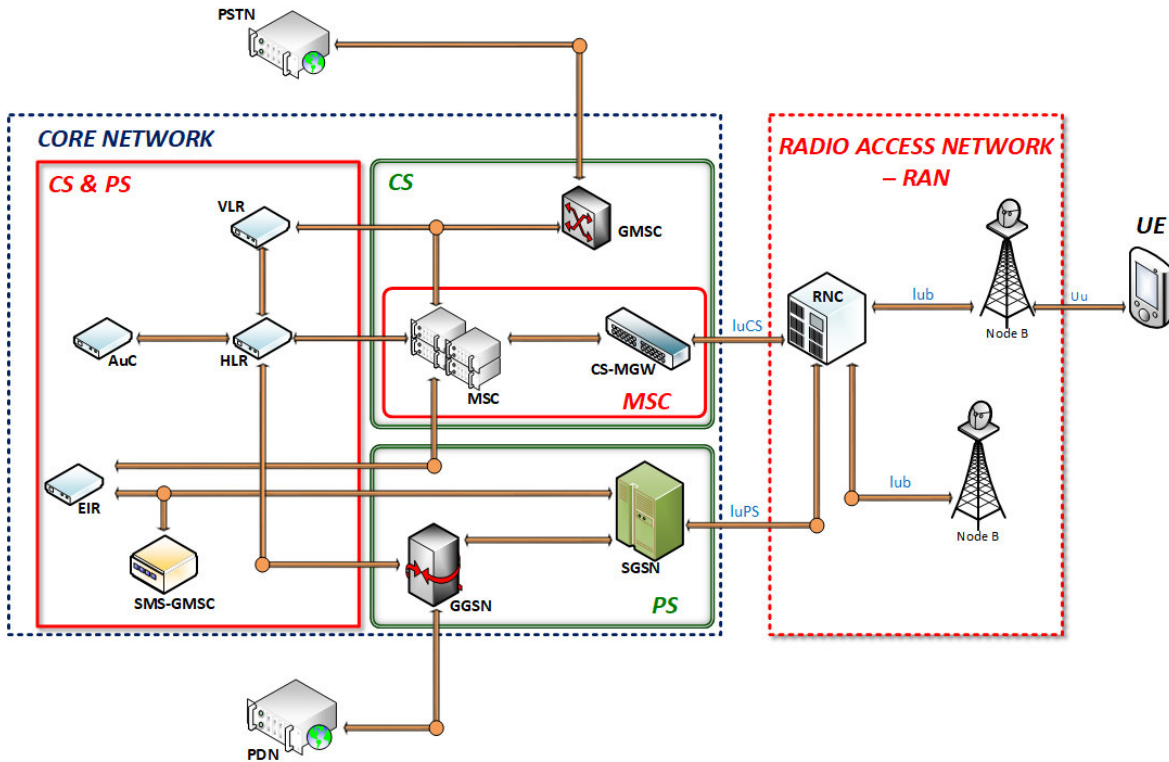


Figure 5: UMTS Architecture

### **2.3.1 Core Network**

The core network of UMTS has its base on the GSM core network, where elements are overlaid to allow additional functions. It can be divided into two groups of elements:

- Circuit-switched elements: used to carry data by applying CS and contains the MSC and the GMSC;
- Packet-switched elements: used to carry packed data, allowing an improved network usage since it carries data as packets which can be routed directly to their destination. The PS contains the SGSN and the GGSN.

There is another group of elements that are shared by both the previous groups, which are:

- HLR and VLR;
- EIR;
- AuC.

### **2.3.2 UMTS Terrestrial Radio Access Network**

UTRAN is the main type of RAN in UMTS and is equivalent to BSS in GSM, providing and managing all air interfaces within the network. The UTRAN system can be divided into:

- Node B or Base Station;
- Radio Network Controller.



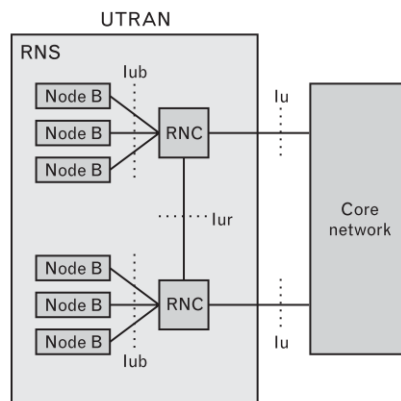


Figure 6: UTRAN components and interfaces

Externally, the Iu interface connects the UTRAN to the CN and the uu interface connects to the UE. Internally, the Iub interface connects the RNCs to the Node Bs and the Iur interface connects the RNCs to each other. This is illustrated in Figure 5.

The UTRAN uses two techniques to ensure concurrent transmissions (uplink and downlink):

- FDD: uplink and downlink occur using different frequencies, which implies that the bandwidth has to be doubled and a filter is needed to avoid signal interfering;
- TDD: a single frequency carrier is used for uplink and downlink with the transmissions occurring at different time intervals.

### A. Node B

In general, Node B are used to refer logical concepts and to physically refer an entity the term used BS is used, the UMTS equivalent to a BTS in the GSM. The Node B is controlled by the RNC and provides connections between the MS and the fixed part of the network, managing the data over the air interface. It can perform the following functions [28]:

- Map logical resources onto hardware resources;
- Transmit RNC system information messages;
- Report uplink interferences and downlink power information;
- Synchronize time and frequency.

## **B. Radio Network Controller**

The RNC is used to control the Node B and can perform the following functions [28]:

- RAN resources management;
- Allocation of traffic channels in the lu and lub interfaces;
- System information management and scheduling of system information;
- Traffic and reporting management.

### **2.3.3 Mobile Equipment**

In the UMTS there is a terminal equipment, which has into it a Mobile Equipment and an USIM. The USIM consists on an improved SIM card, including a list of preferred and prohibited PLMN as well as information about language to be displayed when roaming.

### **2.3.4 Release 99**

The first UMTS release was the Release 99. This release has a strong GSM presence since it had to be compatible and interoperate with the existing GSM networks and the improvements focus mostly on the radio access network with the introduction of UTRAN and software improvements to support the new lu interface between the UTRAN and the MSC. Since the W-CDMA was not compatible with the GSM network, the UTRAN consists on a translation of the GERAN where the BTS and BSC components are substituted by the BS and RNC [34].

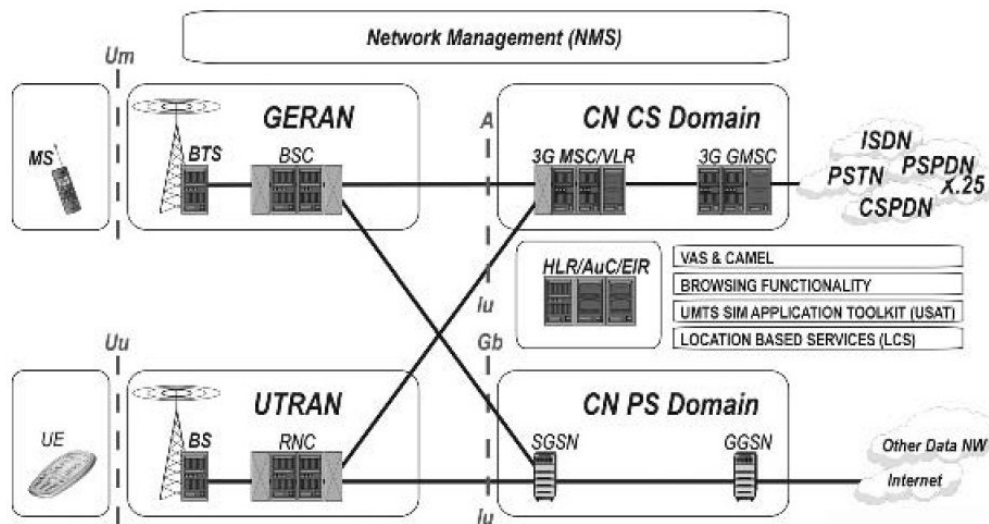


Figure 7: 3GPP network implementation on R99

To simplify, the 3GPP R99 implementation GSM-based, GSM-evolved network with two access networks and providing both CS and PS traffic with variable speeds [34].

### 2.3.5 Release 4

Release 4 improved the core CS voice and data services by introducing the bearer-independent CS CN and IP transport for CN protocols: it offers an option to convert protocol stacks in such a way that the transport protocols becomes IP-based and bearer-independent CS CN brings scalability to do system [34].

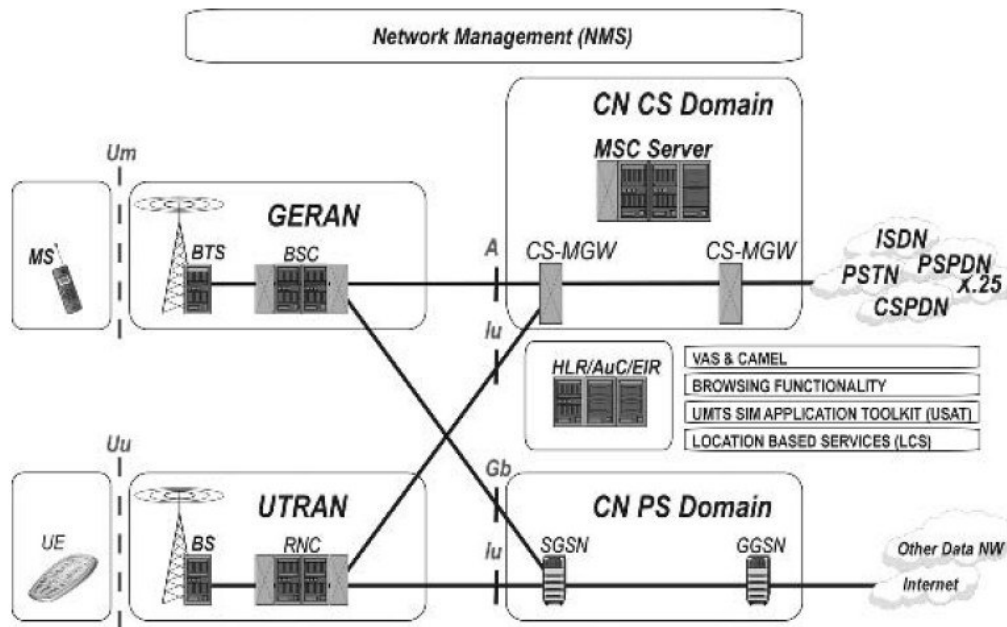


Figure 8: 3GPP network implementation on R4

The Circuit Switched Media Gateway (CS-MGW) is the node that maintains CS connection capacity and is responsible for all physical connection set-up. The MSC server maintains connection control capacity and one MSC server can control various CS-MGW. This separation facilitates the allocation of more efficient bearer services with different characteristics.

### 2.3.6 Release 5

Release 5 brought the introduction of the IP Multimedia Subsystem (IMS) to the CN side, allowing a wide use of various multimedia.

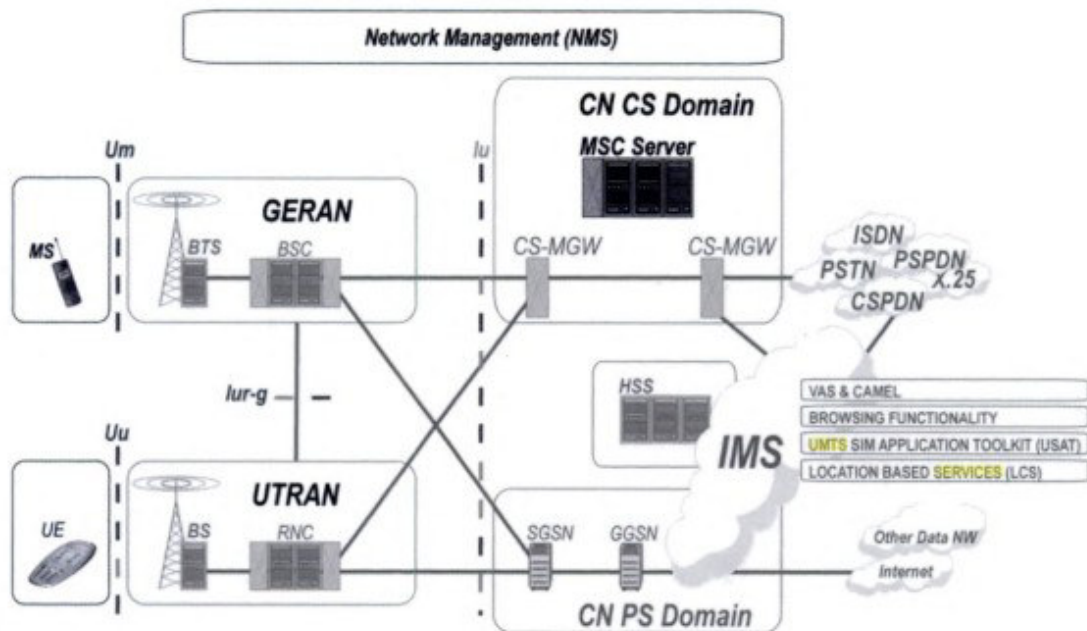


Figure 9: 3GPP R5 implementation

The IMS is a separate system solution that is able to utilise various network itself, allowing the user to use sophisticated multimedia and messaging services [34]. IMS provides the possibility to set up multimedia sessions by means of the Session Initiated Protocol (SIP) and has the following functions and domains [28]:

- Separation of control and data paths;
- Implements All-IP network;
- Handle voice and data in a similar way;

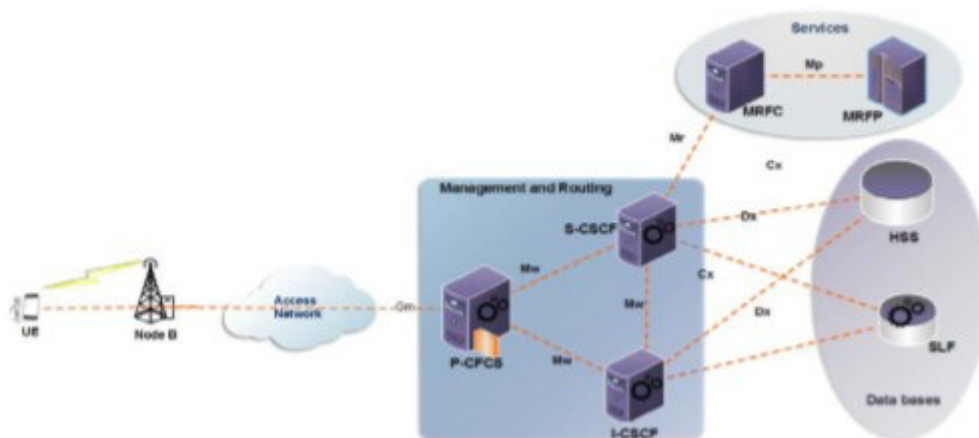


Figure 10: 3GPP IMS Architecture

Figure 10 shows an overview of IMS architecture with the most relevant entities, functions and interfaces.

The Proxy-Call Session Control Function (P-CSCF) is the first point of contact between the UE and the IMS network; Serving-Call Session Control Function (S-CSCF) is responsible for registration process, routing decision and service profile storing; Interrogating-Call Session Control Function (I-CSCF) is the contact point between the subscriber and the service provider. The Databases are the Home Subscriber Server (HSS), which is a merge of HLR and AuC, and the Subscription Locator Function (SLF), that allows I-CSCF and S-CSCF to find the address of the HSS that holds the information of a certain subscriber. The Services functions provides the home network with a source of media and the ability to mix media streams, media analysis, etc, and are divided into a signaling plane (MRFC) and a media plane (MRFP).

Appart from the IMS, this release also introduced the High Speed Downlink Packet Access (HSDPA) to the radio access side. This is a packet based data service which includes the following functional entities [34]:

- Hybrid Automatic Repeat Request (H-ARQ):
- Adaptive Modulation and Coding;
- Fast Packet Scheduling

HSDPA allows to achieve higher peak rates, reduce delay and increase throughput by employing a time-multiplexing approach to transfer data packets through a single shared channel and using the 16-QAM to enhance the air interface.

### **2.3.7 Release 6**

Release 6 is the second phase of the 3GPP High data packet specifications by introducing an uplink dedicated specification, the High Speed Uplink Packet Access (HSUPA). The HSUPA improves the bandwidth available in the uplink direction, increasing the uplink speed, which has not been improved since Release 99. The new uplink transport channel, the Enhanced Dedicated Channel (E-DCH) support fast Node B based scheduling, fast physical H-ARQ and shorter transmission

time interval. Despite having the same features as the HS-DSCH for HSDPA, the E-DCH is not a shared channel and each user has its own dedicated and independent E-DCH [36].

Release 6 also brought improvements to the IMS, the capability of interworking with WLANs and to the HSDPA, introducing MIMO antennas that allow to reach higher data rates. The next table presents a comparison between HSDPA and HSUPA.

<b>Feature</b>	<b>HSDPA</b>	<b>HSUPA</b>
Peak data rate	14.4 Mbps	5.6 Mbps
Modulation scheme(s)	QPSK, 16 QAM	QPSK
TTI	2 ms	2 ms (optional)/ 10 ms
Transport channel type	Shared	Dedicated
Adaptive Modulation and Coding (AMC)	Yes	No
HARQ	HARQ with incremental redundancy; Feed Back in HS-DPCCH	HARQ with incremental redundancy; Feedback in dedicated physical channel (E-HICH)
Packet scheduling	Downlink scheduling (for capacity allocation)	Uplink scheduling (for power control)
Soft handover support(U-Plane)	No (in the downlink)	Yes

### 2.3.8 Release 7

Release 7 is marked by a set of new substantial enhancements to the end-user performance, network capacity and radio technology improvements but compared to the previous releases, this improvements mark an evolution smoothing. The High Speed Packet Access (HSPA) and its evolution, the HSPA+, were introduced. The HSPA is the result of the combination of HSDPA with HSUPA and is compatible with existing Releases 5 and 6, putting UMTS closer to LTE; the HSPA+ is considered the last UMTS improvement and adds new functions such as higher order modulation and MIMO, till then only used in downlink [35].

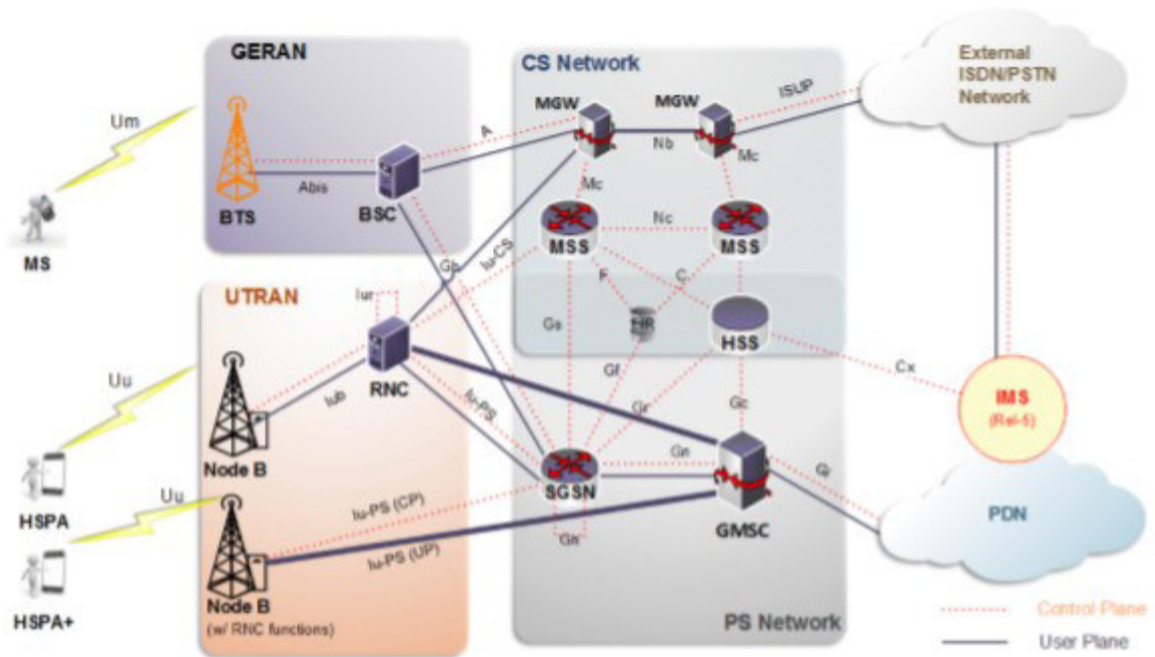


Figure 11: 3GPP Release 7 implementation

To summarize, the next table show the functionalities present in 3GPP releases upon the UMTS network:

	Release 99	Release 4	Release 5	Release 6	Release 7
Soft Handover	✓	✓	✓	✓	✓
Hard Handover	✓	✓	✓	✓	✓
HSDPA			✓	✓	✓
HSUPA				✓	✓
MBMS				✓	✓

## 2.4 LTE Architecture

Developed by the 3GPP, the LTE is the successor of the UMTS. The LTE comes from the need to ensure competitiveness of the 3G system for the future, the demand of higher data rates, QoS and lower latency, as well as the continuous desire to reduce CAPEX and the OPEX. Some of the main feature of LTE are [18]:



- OFDMA based;
- FDD and TDD support and interworking;
- Flexible spectrum support;
- Low latency and high data rates.

While UMTS used W-CDMA for both uplink and downlink, the LTE adopted OFDMA for downlink and SC-FDMA for uplink. The table below present a comparison between the LET specifications and the previous releases.

	<b>WCDMA(UMTS)</b>	<b>HSPA HSDPA/HSUPA</b>	<b>HSPA+</b>	<b>LTE</b>
Max downlink speed bps	384 k	14 M	28 M	100 M
Max uplink speed bps	128 k	5.7 M	11 M	50 M
Latency round trip time approx.	150 ms	100 ms	50 ms(max)	~10ms
3GPP releases	Rel 99/4	Rel 5/6 2005/6	Rel 7	Rel 8
Approx. years of initial roll out	2003/4	HSDPA 2007 /8 HSUPA	2008 /9	2009/1'
Access methodology	CDMA	CDMA	CDMA	OFDMA /SC-FDMA

It is noticeable the large improvements in the max download and max upload speeds as well as significant decrease in latency time.

LTE architecture elements and interfaces can be grouped into three main components:

- User Equipment;
- Evolved Packet Core;
- Evolved UMTS Terrestrial Radio Access Network (E-UTRAN).

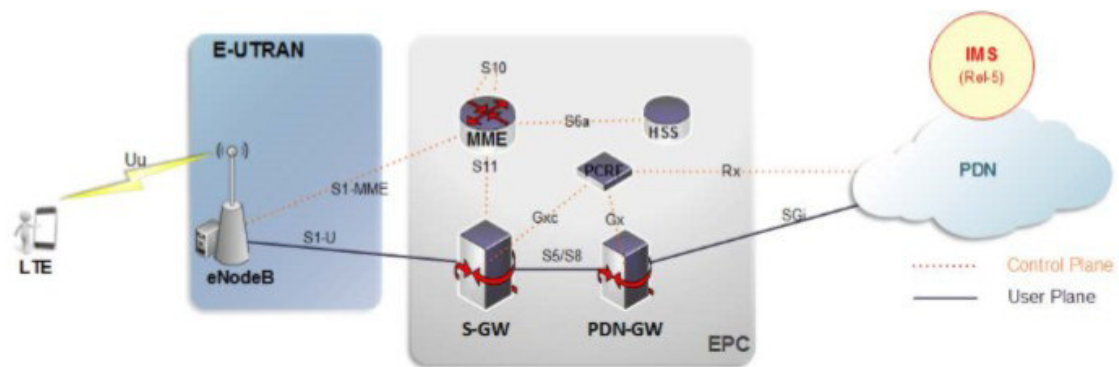


Figure 12: 3GPP LTE implementation (Release 8)

The UE internal architecture in LTE is similar to UMTS and GSM ME. EPC and E-UTRAN are presented with more detail in the following subsections.

### 2.4.1 E-UTRAN

In the radio access side, where used to be BTS and BSC in GSM and RNC with NodeB in UMTS, in LTE there is only one element: eNodeB. This new RAN is based on OFDM that combined with spatial multiplexing through multiple antennas, modulation up to 64QAM and larger bandwidth, allows to achieve higher data rates.

The E-UTRAN is a network of interconnected eNBs with no centralized controllers and that are connected to the EPC over S1 interface. The S1 interface is divided into S1-U and S1-MME, where the first is essentially used for user data plane and S1-MME is used for control plane. The eNodeB is responsible for the following control plane functions [37]:

- Radio bearer management;
- Radio interface transmission and reception;
- UL and DL Dynamic RRM and data scheduling;
- Mobility management;
- User data IP header compression and encryption.

## 2.4.2 EPC

The Evolved Packet Core, also known as LTE core network, is all-IP based meaning both real time and data services will be carried by IP protocol, making the traffic performance and cost more efficient. In LTE is introduced the separation of user data, user plane and the control plane, making it scale independently which brings improvements in dimensioning and adapting the network system.

As Figure 12 shows, the EPC contains the following components:

### A Mobility Management Entity (MME)

The MME is responsible for control plane functions related to the user and the user access to the network, establishing a direct path between itself and the user equipment. Also handles security related to network access by the user.

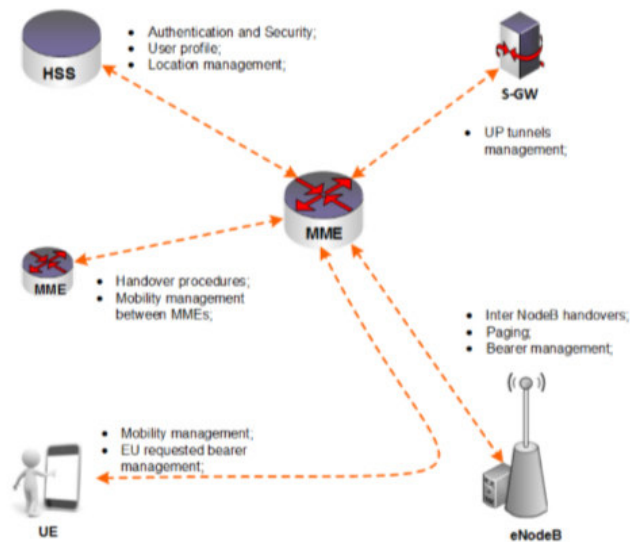


Figure 13: MME logical connections and functions

## B Serving Gateway (S-GW)

S-GW is more concerned with the user plane tunnel management and can use GPRS Tunneling Protocol (GTP) or Proxy Mobile IPv6 (PMIP) tunnels for data flow depending on the data bearer setup.

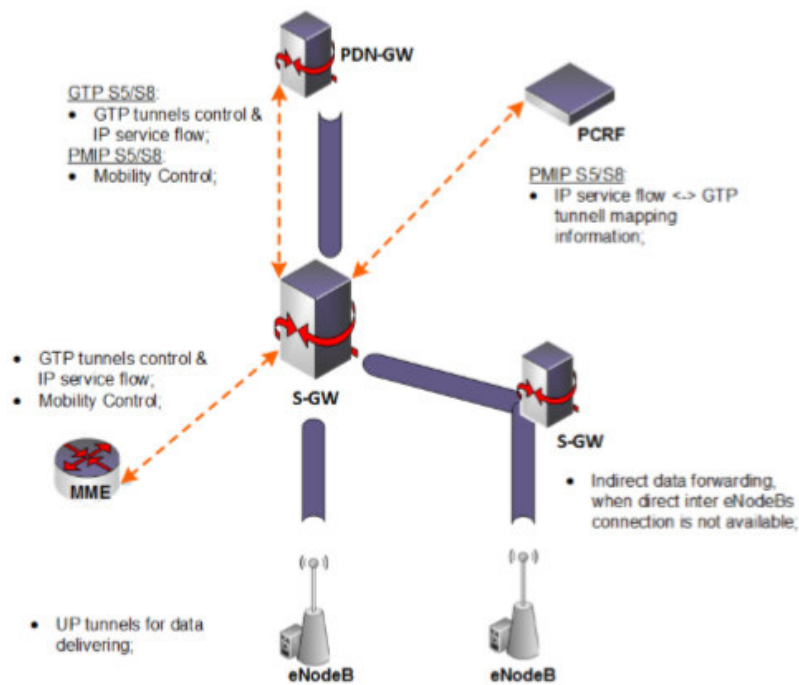


Figure 14: S-GW logical connections and functions

The S-GW only controls its own resources and works based on request from PDN-GW, MME or PCRF.

## C Packet Data Network Gateway (PDN-GW)

PDN-GW is a gateway to the Packet Data Network (PDN), assigning an IP address to the UE which is used to communicate with external networks. It also performs the Dynamic Host Configuration Protocol (DHCP) [37].

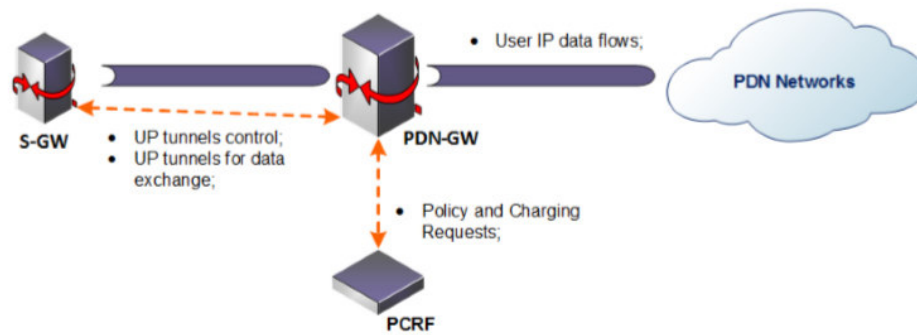


Figure 15: PDN-GW logical connections and functions

#### D Policy and Charging Rules Function (PCRF)

PCRF is the policy, decision making and charging control element of service data flows and IP bearer services. Based on QoS, the PCRF decides on how to handle the services, then sends information to the PDN-GW and, if applicable, the S-GW sets up the appropriate bearer and policing.

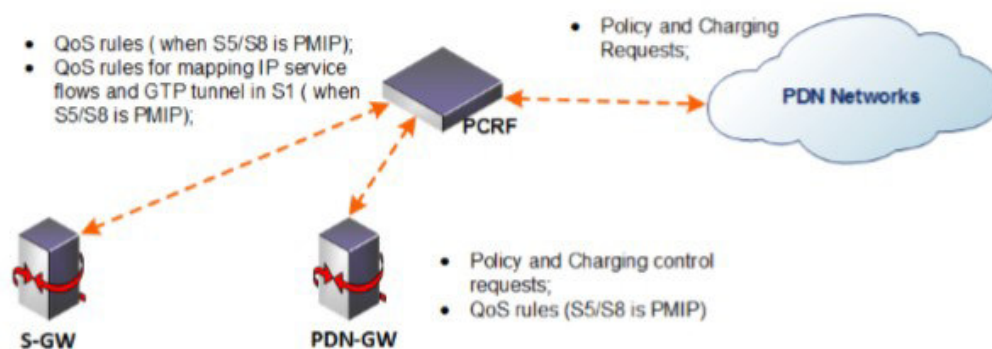


Figure 16: PCRF logical connections and functions

## **E Home Subscriber Service (HSS)**

The HSS contains information about the MME upon which the user is registered and the PDNs that the user can connect to. In comparison to GSM and UMTS, the HSS contains the HLR and AuC present in those architectures.

### **2.4.3 Releases 8 and 9**

The LTE architecture and specifications previously present correspond to the 3GPP Release 8. The consequent releases are focused on enhancing the technology and improving the data rates. In Release 9 are included improvements that were not in Release 8 such as [38]:

- Public Warning System (PWS): include different types of alerts sent to the user in case of emergencies;
- Femto Cell: implementations of small cells inside homes, offices or public spaces which are connected to networks providers using a landline broadband connection.
- Self-Organizing Networks (SON): reduces manual works by means of self-installations, consequently causing a cost reduction.

### **2.4.4 Releases 10 and 11**

In these releases is introduced the LTE-Advanced, increasing the data capability of LTE networks. The improvements included in Releases 10 and 11 are the following [38]:

- Carrier aggregation improvements: the UE uses multiple carriers in downlink and uplink;
- Evolved MIMO operation: support up to eight transmitter and eight receiver antennas for downlink;
- Heterogeneous network operation: allows the combination of large macro cells with small cells networks;
- Enhanced Uplink multiple access: SC-FDMA in uplink;
- SON improvements: improvement of SON features from Release 8 and 9.

The next table presents a comparison between LTE and LTE-Advanced, exposing the significant improvements accomplished with the last.

<b>Parameter</b>		<b>LTE</b>	<b>LTE-Advanced</b>
Scalable bandwidths		1.4-20 MHz	20-100 MHz
Peak data rate	DL	300 Mbps	1Gbps
	UL	75 Mbps	500 Mbps
Transmission bandwidth	DL	20 MHz	100 MHz
	UL	20 MHz	40 MHz
Peak Spectrum Efficiency [bps/Hz]	DL	15	30
	UL	3.75	15

The current state of mobile telecommunications networks is characterized by the cohabitation between GSM (2G), UMTS (3G) and LTE (4G) technologies.

## 2.5 Network Management

In Telecommunication context, Network Management consists on a set of capabilities of a given system and elements to allow an exchange and processing of information, assisting the service providers in their business. It is concerned with four aspects:

- Operation: keeping the network and its services running with minimum impact on the user experience;
- Maintenance: concerned with the state of the NEs (need for upgrade or repair);
- Provision: related to facilitate service requirements by means of network configuration;
- Administration: deals with network control.

The constant network developments allied with introduction of new technologies and services require improvements regarding to network management. Some of the existing network management models include:

- Telecommunication Management Network (TMN);
- Information Technology Infrastructure Library;
- Cisco Lifecycle Services;
- Enhanced Telecom Operations Map (eTOM).

## **2.5.1 Network Management in UMTS**

The UMTS network management has its base on TMN and contains definitions for general management frameworks and concepts. Some of its main management functions are [28]:

- Software management;
- Quality of Service (QoS) management;
- Performance Management;
- User equipment management.

### **A. Software management**

Software management is divided into two processes: software-management process used to manage or correct new software releases and software-fault-management process used to handle with software malfunctions identified when monitoring the network.

### **B. QoS Management**

Includes QoS policy provisioning, used to configure and maintain network elements with the QoS policies created based in network performance, and QoS monitoring, responsible for collecting all statistical data related to the performance which will be used when applying change to the network

### **C. Performance Management**

This function is used to access the current configuration of the network, allowing to collect the following types of data:

- User data and control signalling traffic levels in the network and the usage of resources by network nodes;
- Network configuration and verification;
- Resource-access measurements and availability;
- QoS measurements experienced by the user.



## D. User Equipment Management

This allows the network service provider to trace a specific subscriber activity and report to network management system.

## 2.6 Key Performance Indicators

The Key Performance Indicators (KPI) are the most important indicators expressing the state of the network and can be used not only to detect problems but also for a variety of subjects such as performance trends. KPI are measurements of parameters used for monitoring and there many parameters that can be used. Since there is no criterion to choose the right parameter, it is up to communication service provider to choose the one that best its network behaviour.

There is a wide variety of KPI, divided into the following categories [29][30].

<b>KPI</b>	<b>Counter</b>
<b>Accessibility</b>	<ul style="list-style-type: none"><li>• Radio access success ratio;</li><li>• IU paging access ratio;</li><li>• AMR RAB setup success;</li></ul>
<b>Mobility</b>	<ul style="list-style-type: none"><li>• Soft handover success ratio;</li><li>• Inter/intra frequency hard handover success ration;</li></ul>
<b>Availability</b>	<ul style="list-style-type: none"><li>• Worts cell ratio;</li><li>• Congested cell ratio;</li><li>• Average CPU load;</li></ul>
<b>Traffic</b>	<ul style="list-style-type: none"><li>• CS equivalent traffic;</li><li>• PS UL throughput;</li><li>• PS DL throughput;</li></ul>

In the RNC, the counters will take performance measures using Traffic KPI, which are essentially used to evaluate all CS equivalent traffic [29]. In this dissertation, forecast models will be applied to traffic KPI data set provided by a communication service provider, since KPI are able to be estimated, helping the service provider to plan its network management and focus on problems. This is an important aspect since it allows the service provider to provide a better experience to the end user.

### 3. Forecasting

Forecasting can be used in different areas such as economic (stock exchange forecast), political (outcome of elections forecast), telecommunication (network performance forecast), etc. This section is focused on forecasting the mobile network behaviour. The already existing literature in the field can be divided in two approaches [1][2] :

- Historic data (times series) based forecasting.
- Diffusion of innovation forecasting.

In this dissertation will be considered the historic data approach. This consists on predicting, as accurately as possible, the future values of an event given the information already available. According to [3], the predictability of an event depends on several factors including:

- How well we understand the factors that contribute to it;
- How much data are available;
- Whether the forecast can affect the thing we are trying to forecast.

Forecasting has been in constant evolution as it can be an important support for efficient and effective planning [2][4]. There is already a vast amount of methodologies that can be applied in different situations (e.g. finances, weather), but there are still areas where the forecasting technologies have not been fully explored. One of those is forecasting and planning in mobile networks.

#### 3.1 Telecommunication forecasting

During its life-cycle, every product or service passes through the following phases: introduction, growth, saturation and decline. Therefore, the understanding and forecasting of each segment of Service life-cycle (SLC) have become more and more important in a competitive market [9].

The use of forecasting techniques may contribute for the mobile operators to be prepared to solve and/or prevent network problems such as reduced QoS, queuing delays, traffic service almost reaching capacity limit and packet loss. Telecommunication forecasting is usually base on indicators such as [9]:

- User growth;
- Market share;
- Volume – Pricing;
- Average revenue per user (ARPU).

In this dissertation, the use of forecast will be applied to help communication service providers in the management of the network degradation of Quality of Service (QoS). The existing difficulty in knowing when a NE has its capacity reached is causing communication service providers revenue losses. The application of forecasting techniques allows the service provider to estimate when a capacity upgrade is needed.

### 3.1.1 Forecasting Methods in Telecommunication

According to [9], there are several software tools and methods available for telecommunication forecasting. The following methods are the most often used:

- New telecommunications service penetration forecasting by using growth models;
- Forecasting models based on seasonal variations elimination and autoregression;
- Cross-section models for the forecasting based on the relations between different services or the relations between equal services in different markets;
- Scenario methods;
- Monte Carlo – for revenue, costs and net present value (NPV) forecasting.

In this dissertation will be studied different quantitative statistical forecasting models.

## 3.2 Forecasting Categories

The forecasting methods can be divided in two basic categories which are described as follows:

- **Qualitative Methods:** rely exclusively on the intuition of the forecaster which can lead to subjective and nonmathematical human opinion. Some of the methods that can be cited are [5][6]: Judgmental method, Delphi method and Scenario method.

- **Quantitative Methods:** based on analytical and statistical models of an observed event. It uses consistent mathematical modelling to generate an objective forecast, taking in consideration a lot of information at one time. The most important methods are [6][9]: Times series methods and Causal methods

In general, quantitative forecasting uses either time series data or cross-sectional data collected at a single point in time [3].

	<b>Qualitative Methods</b>	<b>Quantitative Methods</b>
<b>Strengths</b>	May incorporate Latest changes both in the environment and inside information	Consistent and objective which allow to consider and data at one time
<b>Weaknesses</b>	Reduces forecast accuracy	Depends on quantifiable data

The object of work in this dissertation are quantitative methods, the qualitative methods of forecasting will not be addressed, for further information please refer to [8].

### 3.2.1 Quantitative Methods

As the methods can be grouped in two categories [6][10]:

- **Time-series models:** consists of a set of observations taken at regular intervals over a specific period of time. Only the history of the variable is considered in the forecast.
- **Causal models:** identifies the casual relationship between the forecast variable and a set of predictors (causal factors). Uses regression models and various techniques for the evaluation of their applicability, as well as the reliability of forecasting results.

Both of the categories are mathematical but the way they generate forecasts is different: time series models have all the necessary information in the data and generate a forecast based on patterns on the data while casual models require a more complex process and the variable to forecast is related to other variables[5].

### 3.3 Times Series Analysis

A time series is a sequence of several quantitative data observations taken over different periods of time at equally space intervals [13][14]. The figure shows an example of a time series of the total number of airline passengers over a period of time [15].

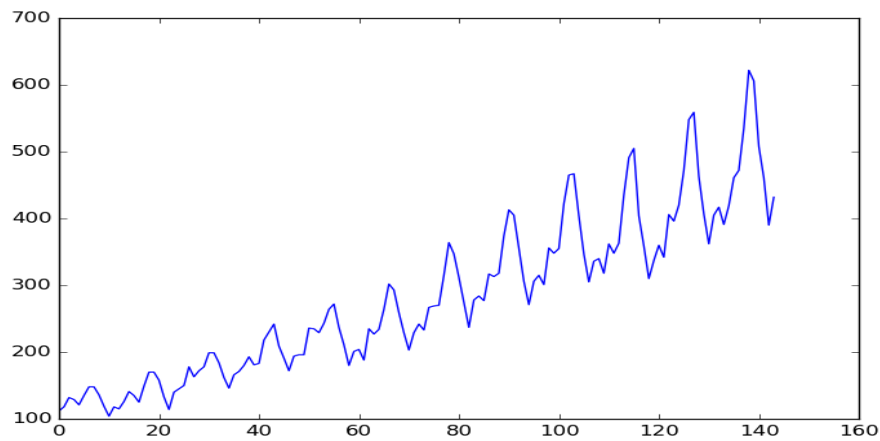


Figure 17: Total number of airline passengers over a period of time [15]

#### 3.3.1 Times Series Components

A times series can have a variety of patterns and behaviors and it can be divided into the following components[8]:

- **Trend (T)** – it is the long term pattern; It can be positive or negative depending on whether the time series show a long term increase or decrease in the data. This component may be associated with the cycle component resulting in trend-cycle component [3].

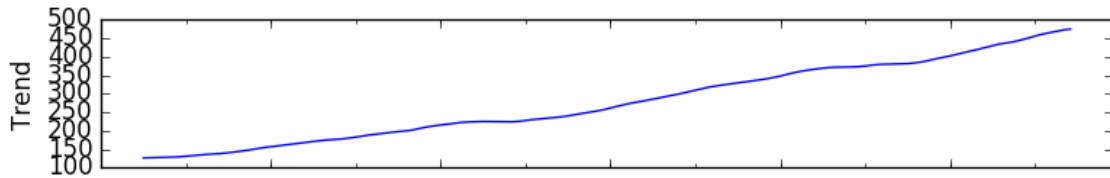


Figure 18: Trend component of the time series of airline passengers [15]

- **Seasonal (S)** – when the time series exhibits regular fluctuations during the same time period (e.g. during the same week, every month). This type of series is often called periodic despite not repeating over each time period [9].

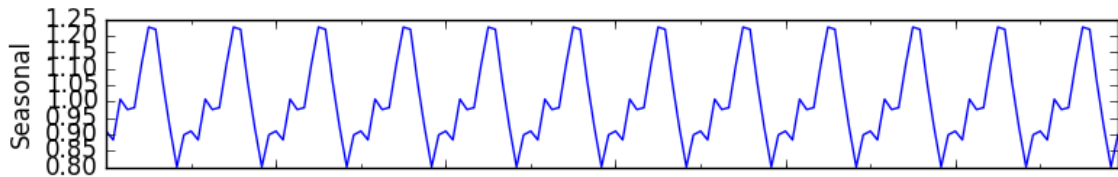


Figure 19: Seasonal component of the time series of airline passengers [15]

- **Irregular (random) (I)** – it is the random component and what is left when all the others components are extracted, consisting of unexplained variation that may occur and cannot be predicted.

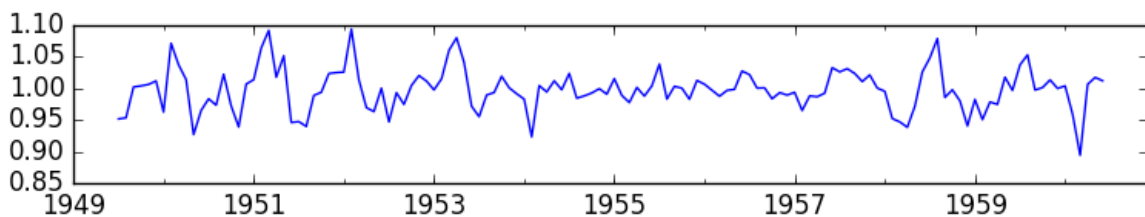


Figure 20: Irregular component of the time series of airline passengers [15]

- **Cycle (C)** – rises and falls around a given trend with variable duration. Generally, when a cycle component is present, is also possible to find seasonality, thus some literature do not separate the trend and cycle for long-term behaviour times series due to the difficulty of separating them, originating the trend-cycle component [13] [14].

Although seasonality and cycle seem similar concepts, when referring to seasonality it has fixed known length and magnitude while the cyclic component varies, making it very difficult to forecast.

After decomposing a time series to obtain the components above mentioned, the original series can be re-composed by adding or multiplying all the components.

### 3.3.2 Graphics

One of the most important steps when working with time series is to graphically represent the data. The analysis of the graphics allows and helps to identify some characteristics of the time series such as patterns or unusual observations (outliers).

There are two common types of graph frequently applied to represent time series data:

- **Time plots:** the observations are plotted as function of the consecutive steps and linked by a straight line to enhance visualization. An example of this graphic is in Figure 17.
- **Scatter plots:** a graphical representation using Cartesian coordinates to display two variables of a data set. This is very important to study the relationship between variables. An example of a scatter plot is shown below:

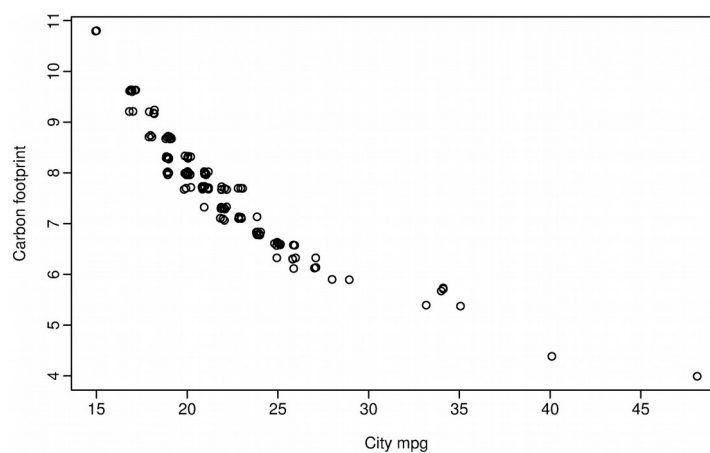


Figure 21: Scatter plot example

Such as the method of forecasting depends on the type of data, so the graphic to be used depends on the type of data to be analysed [3]. Considering that scatter plot purpose is to plot more than one

variable, in this dissertation only time plots will be used since the data being used only has one variable.

### 3.4 Forecasting models

Several mathematical methods of forecasting have been developed to be applied to various situations such as economics, technological forecast and so on. Each method has its own behaviour therefore it is necessary to study and test different methods to select the most suitable for the case in study. According to [18], the recommended ones are:

- Basic methods;
- Smoothing models;
- Decomposition;
- ARMA models

The fact that the methods above use historical data and perform better in comparison to more statistically sophisticated and complex methods is another reason for these methods to be chosen [17].

The description of basic methods, smoothing models and time series decomposition will be presented in this section, leaving the ARMA models to the next chapter, since it is the main forecasting method of this dissertation.

#### 3.4.1 Basic methods

This section is focused on basic forecasting methods and their application on time series data.

##### A. Naive Method

This method relies on the assumption that the forecast period value is equal to the last observation. In other words, if we have daily data until today and we are forecasting the value for tomorrow by naive method, the forecasted value will be equal to today value.

Mathematically, it is represented as follows:

$$F_{t+h} = x_t \quad \text{Equation 1}$$



Where:

$F_{t+h}$  is the forecast for t+h.

$y_{t+h}^o$  is the observed value.

$y_{t+h}$  is the forecast horizon.

## B. Seasonal Naive Method

The seasonal naive method is very similar to the previous method, with the difference that forecasts values for a particular season are equal to the last observation from that same season [3]. For example, given daily seasonal data, the forecast value for the next Monday will be equal the observation in previous Monday.

The mathematical expression for this method is represented as:

$$F_{T+h} = x_{T+h-m(k+1)} \quad \text{Equation 2}$$

Where:

- $m$  is the seasonal period
- $k$  is the integer part of  $(h-1)/m$

Another variation of the naive method is the naive method with drift, which allows the forecast to change (increase or decrease) over time, where the amount of change is the average change in the observations. For further information, please refer to [10][11].

## C. Average Method

The average method or simple mean states that the forecast of  $h$  periods are obtained by the average of all data available.

$$F_{t+h} = (x_1 + \dots + x_N) / N \quad \text{Equation 3}$$

Where:

$F_{T+h}$  is forecast for T+h.

$N$  is the total number of data to be averaged.

$h$  is the forecast horizon (quantity of periods ahead to be forecasted).

In Figure 8 is possible to see the results from the application of the previous methods to the quarterly beer production in Australia from 1992 to 2006.

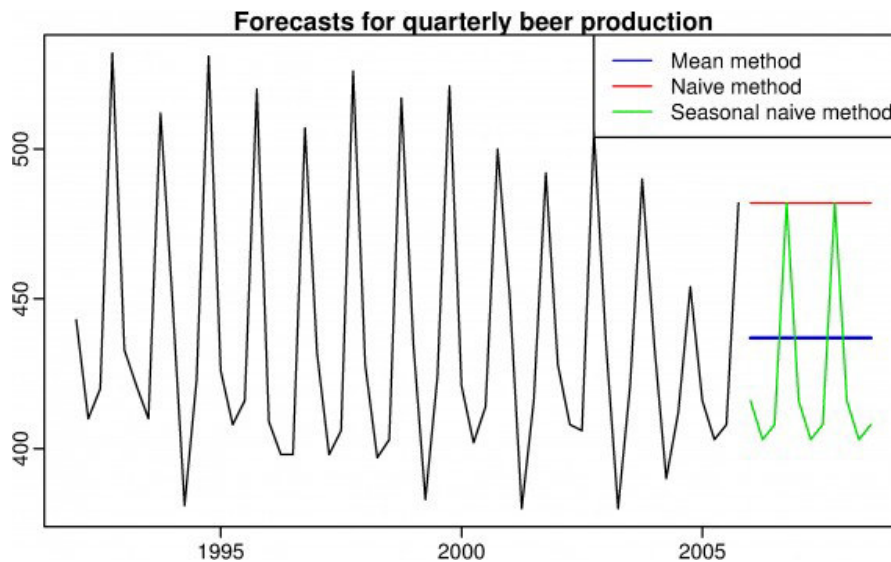


Figure 22: Application of basic methods

From observing the graphic is possible to conclude that these methods may not serve to forecast with this data. However, results obtained with this simple methods can be used to compare with forecasts obtained with more complex methods, i.e., we can discard complex methods with worse results than this simple methods.

### 3.4.2 Smoothing models

The smoothing models are another group of models that can be used to forecast. The moving average concept is the best known smoothing model because as new data become available, old data is disposed, causing the average to move across time. This models require that the time series is locally stationary and have a slow varying mean [19]. The difference between the moving averages comes from the weight assigned to the values. A description of the simple moving average (SMA) and the weighted moving average will be given.

### 3.4.2.1 Simple Moving Average

This method uses a constant number of observations to forecast, meaning that as new data is incoming, old data is discarded. Simplifying, is an average that moves along time [8].

$$F_{t+h} = \frac{X_t + X_{t-1} + X_{t-2} + \dots + X_{t-n+h}}{n} \quad \text{Equation 4}$$

Where:

- $n$  is the number of periods to be averaged;
- $h$  is the forecast horizon;
- $F_{(t+h)}$  is the forecast based on previous data.

Both simple moving average and average method use the same mathematical expression, the difference is that the SMA divisor  $n$  only takes a fraction of the data. However, for high  $n$  values, this method become equivalent to average method but the data to forecast will also be more lagged [20].

### 3.4.2.2 Weighted Moving Average

The weighted moving average is the simple moving average with weights assigned to each value of the data.

$$F_{t+h} = \frac{W_1 X_t + W_2 X_{t-1} + W_3 X_{t-2} + \dots + W_n X_{t-n+h}}{n} \quad \text{Equation 5}$$

Where:

- $n$  is the number of periods to be averaged;
- $h$  is the forecast horizons;
- $W_k$  is the weight assigned. The sum of all weights equal to 1.

Usually, is given more weight to the most recent data and less weight to older data [9] and requires a lot of historical data.

### 3.4.3 Classical decomposition

Time series decomposition help to identify the different patterns a time series can present, such as trend or seasonality. The classical decomposition method identifies the series components (trend (T), seasonality (S), cycle (C) and irregular (I)) which provides an easier forecast since the underlying regular patterns of the time series are already known. There are two decomposition models: the additive and the multiplicative [3][7].

#### 3.4.3.1 Additive Model

The additive model assumes that the time series is the sum of its components and can be applied when the seasonal variation and its amplitude are constant over time. In this model, the seasonality is independent from the trend.

$$Y = T + S + C + I$$

Equation 6

In case the data does not have one component, it is considered zero. The next figure shows an example of an additive model: the trend and the seasonal variation remain constant over time.

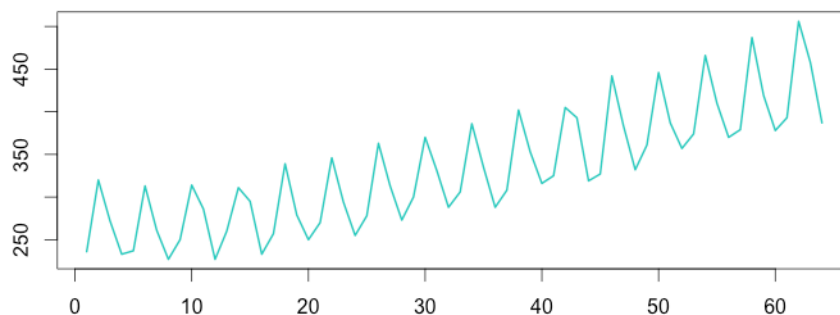


Figure 23: Australian beer production [15]

The series components can be obtained by following the steps:

1. Use the moving average smoothing to calculate the trend-cycle component.

2. Calculate the detrended series by subtracting the trend-cycle component to the data (Y).
3. Obtain the seasonality (S) by averaging the detrended series values.
4. The irregularity (I) is the result of subtracting all the components to the data (Y).

### 3.4.3.2 Multiplicative model

In this model, the data is considered the result of the product of the series components. This is used when the data presents an increasing variation in the seasonal amplitude.

$$Y = T * S * C * I$$

Equation 7

An example of a multiplicative model is represented in Figure 10, where is possible to observe a seasonal variation, with amplitude increasing over time.

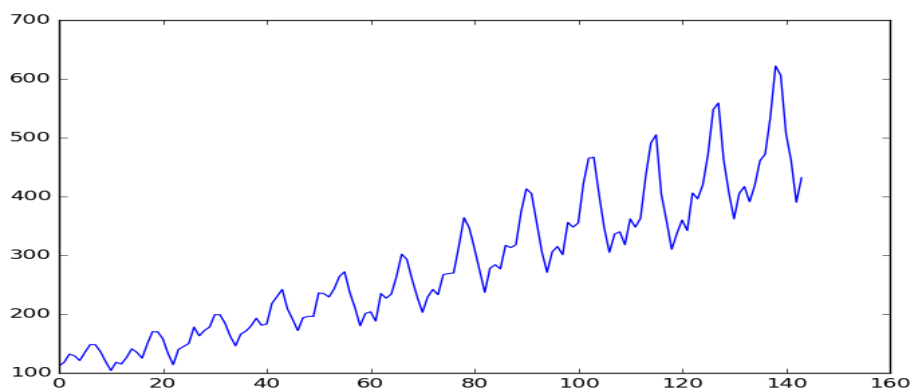


Figure 24: Total number of airline passengers over a period of time [15]

The process to get the series components is similar to the additive model, with some operation changes:

1. Use the moving average smoothing to calculate the trend-cycle component.
2. Calculate the detrended series by dividing the data (Y) by the trend-cycle component.

3. Obtain the seasonality (S) by averaging the detrended series values.
4. The irregularity (I) is the result of dividing the data (Y) by all the components.

The multiplicative model can be override by transforming the data (through differencing, for example) to stabilize the variation in the time series and then use the additive model [3].

### 3.5 Auto-Regressive (AR) Models

A model is considered auto-regressive when a value from the time series can be expressed as a linear combination of previous equally spaced observations [3] and is denoted as AR(p), where  $p$  is the model order. The forecast is obtained through a combination of the last  $p$  known times series values and is expressed by the following expression:

$$F_t = c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t \quad \text{Equation 8}$$

or

$$F_t = c + \sum_{p=1}^N \phi_k x_{t-p} + \varepsilon_t \quad \text{Equation 9}$$

Where:

- $c$  is the constant;
- $\varepsilon_t$  is white noise at time  $t$ ;
- $\phi_k$  are the auto-regressive coefficients ( $k=1, \dots, p$ );
- $p$  autoregressive order.

The variance of  $\varepsilon_t$  and  $c$  will change the scale of the time series but not its patterns.

The AR model can be compared to an infinite response filter (IIR) with  $p$ -order.

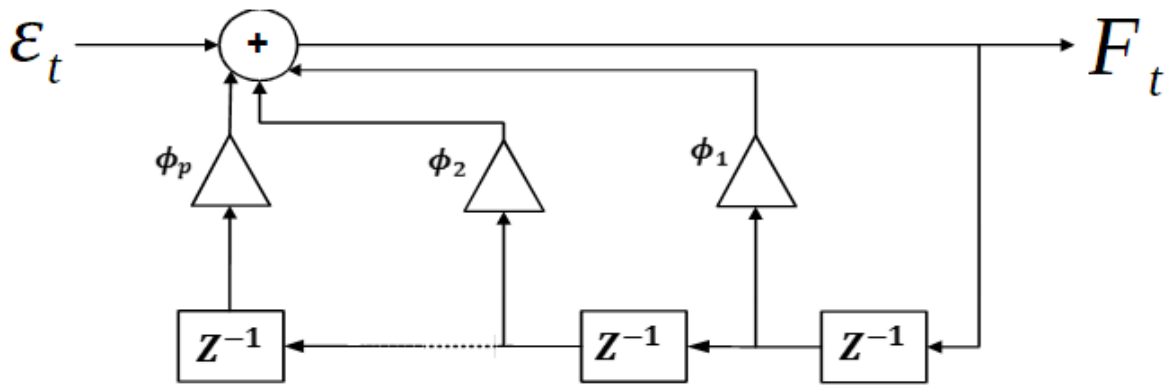


Figure 25: p-order IIR filter implementing an AR model

### 3.5.1 First order auto-regression

The first order auto-regression ( $p=1$ ) is given by the following expression:

$$F_t = c + \phi_1 x_{t-1} + \varepsilon_t \quad \text{Equation 10}$$

In order to predict the current value is necessary to know to value immediately before. The  $\phi_1$  is the coefficient that will be used to minimize the error of the model and get better forecasts. There some particular cases when in presence of AR(1) model [3]:

- $\phi_1=0$             the result is white noise;
- $\phi_1=1 \wedge c=0$     the result is random walk;
- $\phi_1=1 \wedge c \neq 0$    the result is random walk with drift.

### 3.6 Moving Average (MA) Models

The Moving Average model uses the residuals to compute a forecast ( $F_t$ ) and the MA model of order  $q$ , MA( $q$ ), can be denoted as:

$$F_t = c + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_p \varepsilon_{t-p} + \varepsilon_t \quad \text{Equation 11}$$

or

$$F_t = c + \sum_{p=1}^N \theta_k \varepsilon_{t-q} + \varepsilon_t$$

Equation 12

Where:

- $c$  is the constant;
- $\varepsilon_t$  is white noise at time  $t$ ;
- $\theta_k$  are the moving average coefficients ( $k=1, \dots, q$ );
- $q$  is the moving average order.

Although this appears to be a regression, it is not because the  $\varepsilon_t$  values are not observed values. This method can be represented by a block diagram of finite response filter (FIR)

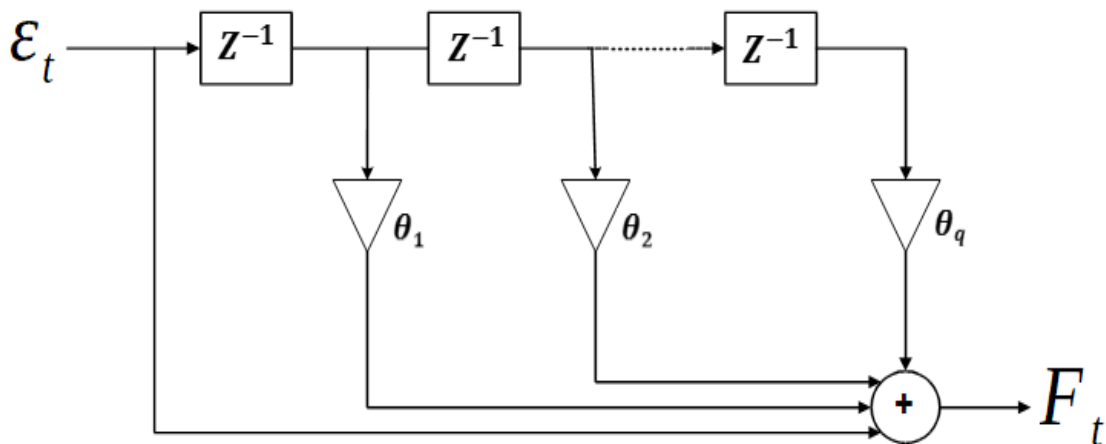


Figure 26: q-order FIR implementing MA(q) model

The moving average model can be seen as a weighted moving average of residuals, considering the MA parameters as the weights. Given the similarities in the nominations, it is important to not confuse MA models with MA smoothing: while this model is used to forecast values, the MA smoothing is applied to obtain the trend-cycle estimation of past values [3].



### 3.7 Auto Regressive Moving Average (ARMA) Models

The Auto Regression Moving Average results from the combination of the Auto Regressive (AR) and the Moving Average (MA) models: the AR is responsible for the memory of past data and the MA treats the forecast errors. These models require the time series to be stationary. A time series is stationary when its probability distribution is constant over time [12]. The ARMA models are divided into two types [22]:

- Non-seasonal ARMA;
- Seasonal ARMA.

#### 3.7.1 Non-seasonal ARMA

The combination of Equation 8 with Equation 11 gives the mathematical expression of an ARMA(p,q), allowing to express the forecast value in function of past values and/or past errors.

$$F_t = c + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad \text{Equation 13}$$

Where:

- $\phi_k$  are the AR coefficients;
- $\theta_k$  are the MA coefficients;

Analogously, the block diagram representation of this method is given by combination of a FIR (MA part) and a IIR (AR part).

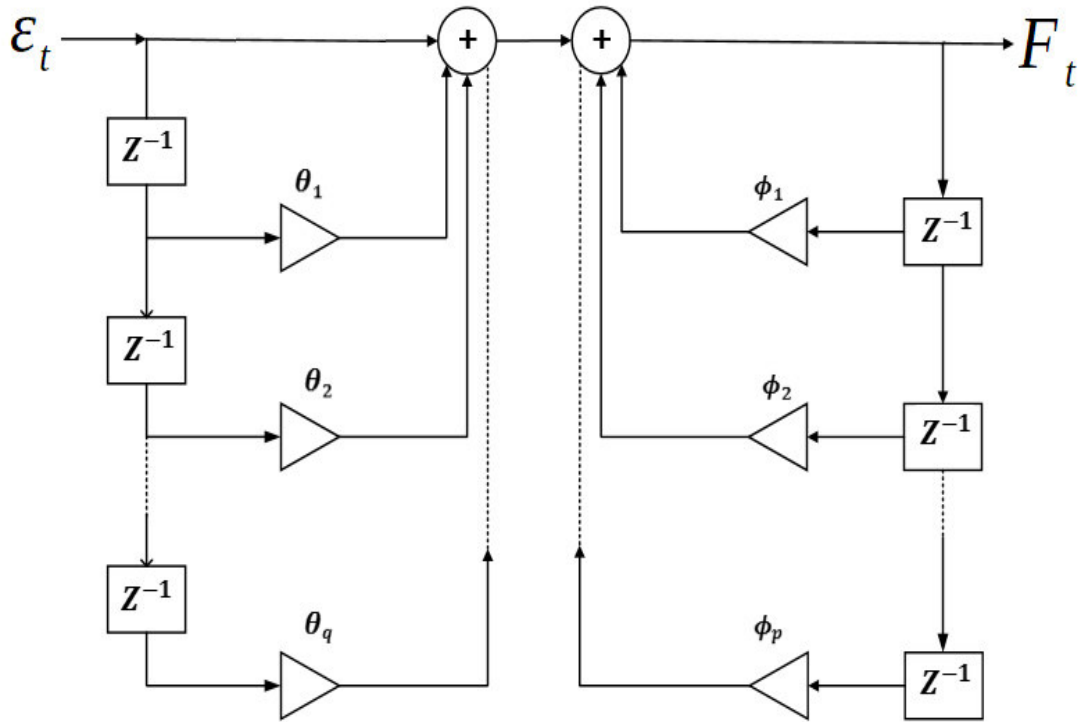


Figure 27: ARMA(p,q) implemented by a FIR/IIR

### 3.7.2 Seasonal ARMA

When data presents seasonality, this model can be applied to present the series in terms of past values at lag equal to length of the seasonal period [23]. This models is denoted as ARMA(P,Q)<sub>s</sub> and the forecast can be computed using the following expression:

$$F_t = \Phi_1 x_{t-1s} + \Phi_2 x_{t-2s} + \dots + \Phi_p x_{t-ps} + \Theta_1 \varepsilon_{t-1s} + \Theta_2 \varepsilon_{t-2s} + \dots + \Theta_Q \varepsilon_{t-Qs} + \varepsilon_t \quad \text{Equation 14}$$

Where:

- $\Phi_p$  are the seasonal AR coefficients;
- $\Theta_Q$  are the seasonal MA coefficients;
- $s$  is the length of the seasonal period.

As non-seasonal ARMA, seasonal ARMA can also be represented by a block diagram similar to the Figure 12 but with delay blocks of  $s$ .

### 3.7.3 Mixed Seasonal ARMA

The combination of seasonal and non-seasonal ARMA result in Mixed Seasonal ARMA, denoted as ARMA(p,q)x(P,Q)<sub>s</sub>. Since the expression of this model is complex, the backward shift operator  $B$  is used to simplify:

$$B^n x_t = x_{t-n} \quad \text{Equation 15}$$

The ARMA(p,q)x(P,Q)<sub>s</sub> model is expressed by:

$$\begin{aligned} (1 - \Phi_1 B^s - \dots - \Phi_p B^{ps})(1 - \phi_1 B^1 - \dots - \phi_p B^p) x_t = \\ c + (1 + \Theta_1 B^s + \dots + \Theta_Q B^{Qs})(1 + \theta_1 B^1 + \dots + \theta_q B^q) \varepsilon_t \end{aligned} \quad \text{Equation 16}$$

### 3.7.4 ARIMA models

The Auto-Regressive Integrated Moving Average model is the combination of the ARMA model with differencing, i.e., computing the differences between consecutive observations, converting a non-stationary time series into a stationary one. Considering  $x_t^d$  as the differenced series of order  $d$

$$x_t^d = (1 - B)^d x_t \quad \text{Equation 17}$$

Replacing in the non-seasonal ARMA model presented in Equation 13 leads to representation of non-seasonal ARIMA(p,d,q):

$$F_t^d = c + \phi_1 x_{(t-1)}^d + \phi_2 x_{(t-2)}^d + \dots + \phi_p x_{(t-p)}^d + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad \text{Equation 18}$$

Where:

- $\theta_q$  are MA coefficients;
- $\phi_p$  are AR coefficients;
- $d$  is the differencing order.

Given that the forecast value  $F_t^d$  is obtained from the differenced time series, in order to obtain the ‘real’ forecast  $F_t$  is necessary to sum this value to the last time series value  $x_{t-1}$ .

When the data presents seasonal nonstationarity, seasonal difference can be performed in order to obtain seasonal stationarity and is expressed by:

$$x_t^D = (1 - B^s)^D x_t \quad \text{Equation 19}$$

This leads to seasonal ARIMA model denoted as ARIMA(p,d,q) x (P,D,Q)<sub>s</sub>.

$$\begin{aligned} & (1 - \Phi_1 B^s - \dots - \Phi_p B^{ps}) (1 - \phi_1 B^1 - \dots - \phi_p B^p) (1 - B^s)^D (1 - B)^d x_t = \\ & c + (1 + \Theta_1 B^s + \dots + \Theta_Q B^{Qs}) (1 + \theta_1 B^1 + \dots + \theta_q B^q) \varepsilon_t \end{aligned} \quad \text{Equation 20}$$

Where:

- $(1 - B^s)^D$  is the seasonal difference of order D;
- $(1 - B)^d$  is the non-seasonal difference of order d

### 3.8 Forecast validation

The main goal of this dissertation is to produce forecasts, which demands a necessary validation of the forecasted value. Forecast validation is divided into three main groups [26]:

- In sample forecasting;
- Out-of-sample forecasting;
- Pseudo out-of-sample forecasting.

In this dissertation, the out-of-sample forecasting will be tested for both expanding and moving window. For further information on the other methods, please refer to [19][26].

The out-of-sample method consists on using historical data available to fit a forecasting model and then forecast the next value ( $F_{t+1}$ ). Once this is done, stores the forecast error and when there is a new input, re-estimates the model including the new value and computes a new forecast ( $F_{t+2}$ ). When applying out-of-sample method, the data is divided into two groups:

- **Estimation:** also known as training set, is used to fit the forecast model and obtain forecasted values;
- **Validation:** also known as testing set, is used to test the forecast values and measure the accuracy of the model used.

There two methods that can be used to obtain out-of-sample forecasts:

- **Expanding Window:** Starts with a determined number of observations in the training set to estimate the model and produce one forecast. Then the training set size is incremented by one (equivalent to a new data input), the model is re-estimated including the new value and computes a new forecast. For example, the training data set has three values that will be used to obtain the first forecast value; when a new input arrives, it will be included in the training data set (increasing its size to four) and used to produce the second forecast value.
- **Moving Window:** In this method, the size of the training set is fixed. When a new input arrives and is added to the training set, the oldest value is discarded, maintaining the fixed size of the window. Using the simple example above, lets assume that the initial values of the training set are  $x_1$ ,  $x_2$  and  $x_3$  and that the window has size three. The first forecast is obtained. Considering  $x_4$  as a new data input, the second forecast is computed using the training set with  $x_2$ ,  $x_3$  and  $x_4$ , discarding the old value  $x_1$ .

Summing up, the expanding window uses all available data to the current time while the moving window uses a fixed number of the most recent data.

### 3.9 Forecasting evaluation methods

The forecasting evaluation methods allow to evaluate the forecasting model in terms of accuracy and goodness of fit. There are different methods based on knowing the forecast error  $e_t$ , which is given by the difference between the real value  $x_t$  and the forecasting value  $F_t$ :

$$e_t = x_t - F_t \quad \text{Equation 21}$$

The most commonly used forecast measures of accuracy are [24]:

- Mean of absolute errors – MAE;
- Mean of squared errors – MSE;
- Mean absolute percentage errors – MAPE.

In this dissertation, this three measures will be determined and the forecasting model with the lowest MAPE is chosen to forecast.

#### 3.9.1 Mean of Absolute Errors (MAE)

The mean absolute error consists in computing the mean of the absolute value of the forecast errors, represented in the following equation:

$$MAE = \frac{1}{n} \sum_{t=0}^{n-1} |e_t| \quad \text{Equation 22}$$

This measure is sensitive to small deviations near zero and can be interpreted as the expected value for the forecast error

#### 3.9.2 Mean of Squared Errors (MSE)

The mean of squared errors is a non-negative estimator to measure the average of the squared errors: the closer this value is to zero, the better a certain model fits the data.

$$MSE = \frac{1}{n} \sum_{t=0}^{n-1} e_t^2$$

Equation 23

The fact it uses squared errors makes them very sensitive to larger deviations (higher individual forecast errors).

### 3.9.3 Mean of Absolute Percentage Errors (MAPE)

The MAPE is defined as the average of percentage errors and expressed in Equation 22.

$$MAPE = \frac{1}{n} \sum_{t=0}^{n-1} \left( \left| \frac{e_t}{x_t} \right| * 100 \right)$$

Equation 24

According to different literatures [9][11][25], this is one the most used forecast accuracy measure and the best forecasting model is the one that minimizes the MAPE (null MAPE means perfect fit). However, besides being simple and scale-independent, the use of MAPE has two practical disadvantages:

- Division by zero if there are data equal to zero;
- There is no upper limit for MAPE values.

### 3.9.4 Akaike's Information Criteria (AIC)

This criterion is used to measure the relative quality of a model, telling, for example, if an ARMA(1,1) fits best the data then an ARMA(2,2). However, this criterion does not provide information about the quality of the comparing models by themselves (if the models to be compared fit poorly the data).

The AIC calculates a trade-off between goodness of fit and complexity of the model, providing means to model selection.

$$AIC = \ln(\hat{\sigma}_\varepsilon^2) + \frac{2k}{T}$$

Equation 25

- $\hat{\sigma}_\varepsilon^2$  is the residuals variance;
- $k$  is the number of the estimated ARMA parameters ( $p + q$ );
- $T$  is total number of observations.

The model that best fits the time series is the one with the lower AIC. Though, it is possible that the model that presents the best AIC is not the model with the best forecasting results.

### 3.9.5 Corrected Akaike's Information Criteria (AICc)

The AICc is defined as an AIC corrected for finite size samples which penalizes the use of extra parameters. Such as for AIC, the model with the lowest AICc is preferred.

$$AICc = AIC + \frac{2(k+2)(k+3)}{T-k}$$

Equation 26



### 3.10 Modelling ARIMA procedure

There are several procedures that can be used to model ARIMA models. In this dissertation, the steps taken until obtain a forecast are:

1. Plot the times series to check for identifiable patterns;
2. Analyse ACF of the times series and perform ADF test to check stationarity;
3. If non stationary, take differences until stationarity is obtained;
4. Estimate the  $p$  and  $q$  order by analysing the ACF and PACF of the differenced series;
5. Test identified models and choose the one with the lowest MAPE;
6. Forecast the data.

The *ARMA* function from Python's *statsmodels* package is used to fit the model, the differencing of the times series in order to obtain stationarity is performed separately such as the reverse process to produce the forecast. In addition to observing the times series ACF, the Augmented Dickey-Fuller (ADF) test is performed, which is a unit root statistical hypothesis test to check for stationarity: if the p-value is greater than 0.05, the time series is non stationary and is necessary to apply differencing; if p-value is smaller than 0.05, the series is stationary [12]. To perform this test it is used the function *adfuller* from *statsmodels* package.

The most complex part of modelling an ARIMA is the identification of the  $p$  and  $q$  orders. This is done recurring to the ACF and PACF plots of the stationary time series, produced using the *acf\_plot* and *pacf\_plot* of the *statsmodels* package. Given the importance of this step, a more detailed study of the ACF and PACF is presented in Section 3.11.

Once the identification of the orders is done, the *ARMA.fit* function is used to fit the model to the time series and the function *forecast* to obtain the forecast value.

### 3.11 ACF and PACF

The Autocorrelation Function and Partial Autocorrelation Function are used when analyzing a time series in order to help identify the  $p$  and  $q$  order of the AR and MA models. In this section is presented the definition of both functions as well as some examples to better understand the goal of using these functions.

#### A. Autocorrelation Function (ACF)

The autocorrelation is the correlation of a function with a delayed copy of itself. Considering the function  $f(t)$  where  $t$  is continuous finite time, the autocorrelation  $\rho_{ff}(\tau)$  is defined as:

$$\rho_{ff}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T [f(t) - \bar{f}] \cdot [f(t+\tau) - \bar{f}] dt \quad \text{Equation 27}$$

When  $f(t)$  is a function with finite duration, the autocorrelation is given by:

$$\rho_{ff}(\tau) = \int_{t_1}^{t_2} f(t) \cdot f(t+\tau) dt \quad \text{Equation 28}$$

Where:

- $\tau$  is the delay;
- $f(t+\tau)$  is the delayed function

In order to get a better understanding of this concept, the next figure presents the graphical development of the autocorrelation function of a square pulse of amplitude  $a$  and duration  $T$ .

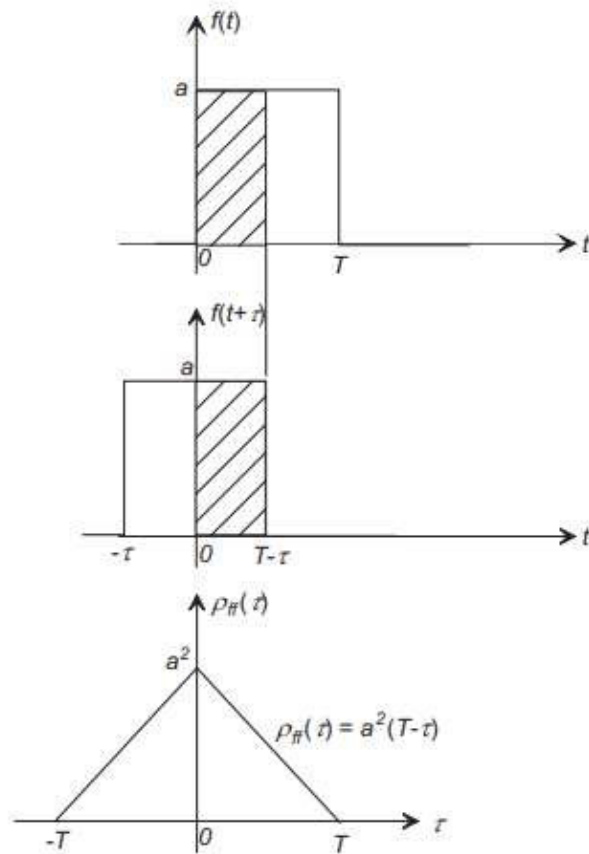


Figure 28: Autocorrelation of square finite pulse

Observing Figure 28, it is possible to interpret that the autocorrelation function computes the area of the section where the delayed function overlaps the original function.

In case  $f(t)$  is a periodic function (such as a sinusoid), the ACF presents the highest values when  $\tau = T$  and the lowest when  $\tau = T/2$ .

As the main focus of this work is time series analysis, where  $t$  is discrete with equally spaced data, the autocorrelation is given by:

$$ACF(h) = \frac{\text{Covariance}(x_t, x_{t-h})}{\text{StdDev}(x_t) \text{StdDev}(x_{t-h})} \quad \text{Equation 29}$$

## B. Partial Autocorrelation Function (PACF)

If  $x_t$  is correlated with  $x_{t-1}$  (autocorrelation at lag 1) and  $x_{t-1}$  is correlated with  $x_{t-2}$ , it is also expected to find correlation between  $x_t$  and  $x_{t-2}$  (autocorrelation at lag 2) and so on. Thus, the result from ACF does not exclude the propagation factor.

The partial autocorrelation (PACF) between  $x_t$  and  $x_{t-h}$  is defined as the conditional correlation between  $x_t$  and  $x_{t-h}$  conditional on the observations between  $t$  and  $t-h$ :

$$PACF(h) = \frac{\text{Covariance}(x_t, x_{t-h} \vee x_{t-1}, x_{t-2}, \dots, x_{t-h+1})}{\sqrt{(\text{Variance}(x_t \vee x_{t-1}, x_{t-2}, \dots, x_{t-h+1})) \text{Variance}(x_{t-h} \vee x_{t-1}, x_{t-2}, \dots, x_{t-h+1})}} \quad \text{Equation 30}$$

For example, the partial autocorrelation at lag 2 is the difference between the autocorrelation at lag 2 and the expected correlation due to the propagation of autocorrelation at lag 1.

Considering the MA(1) given by [39]:

$$x_t = 0,9 \varepsilon_{t-1} + \varepsilon_t \quad \text{Equation 31}$$

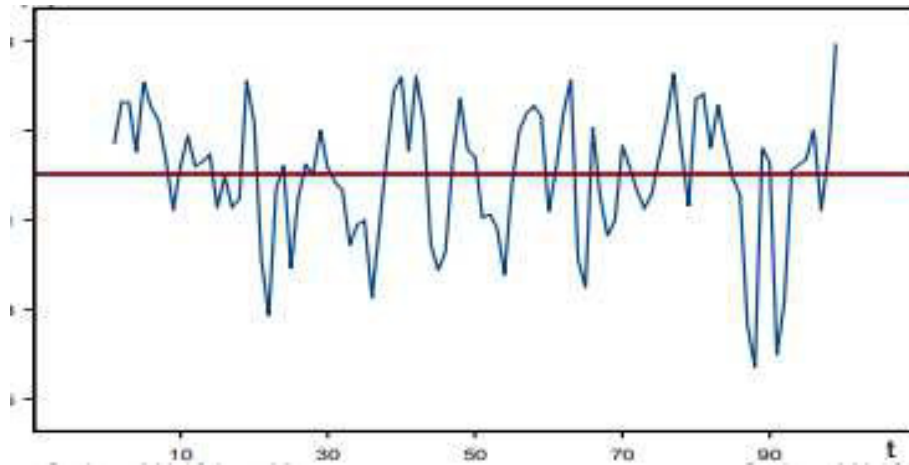


Figure 29: Simulated MA(1)

Computing the ACF and PACF, the results obtained are shown in the following figure [39]:

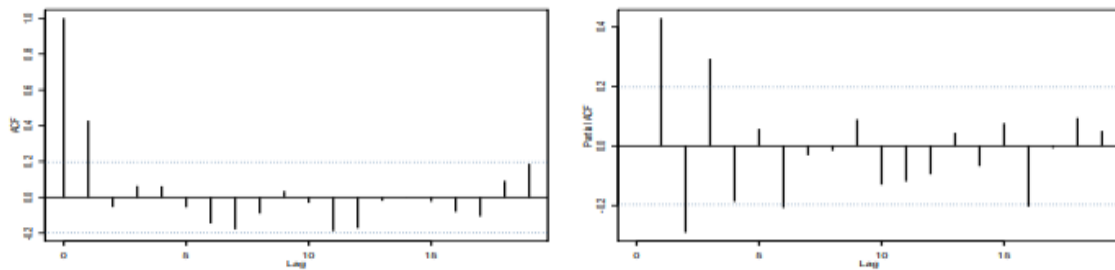


Figure 30: ACF and PACF of the MA(1)

The ACF of the MA(1) process cuts off after lag 1 and the PACF presents a decaying pattern. This can be extended to general case of MA( $q$ ) where the ACF cuts off after the lag  $q$  and the PACF shows a similar pattern. Then, the ACF is used to estimate the  $q$  of the MA process: considering a 95% confidence interval between which the ACF value is considered zero it is possible to identify the first lag  $n$  where the ACF value is within that interval and determine the MA order given by  $q = n - 1$ .

Considering now the AR(1) model expressed as follows [39]:

$$x_t = 0,9x_{t-1} + \varepsilon_t$$

Equation 32

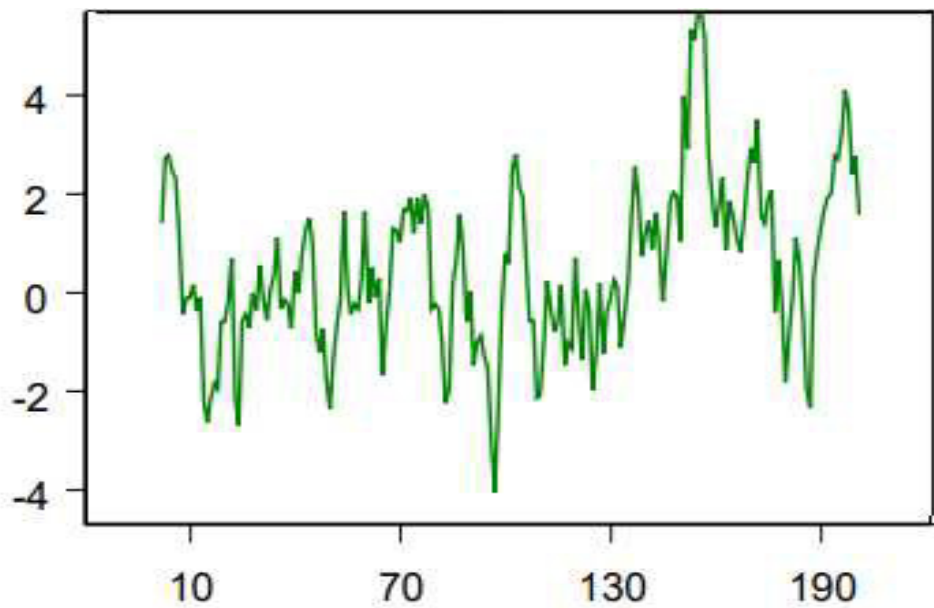


Figure 31: Simulated AR(1)

The ACF and PACF graphics for this model are presented in the next figure:

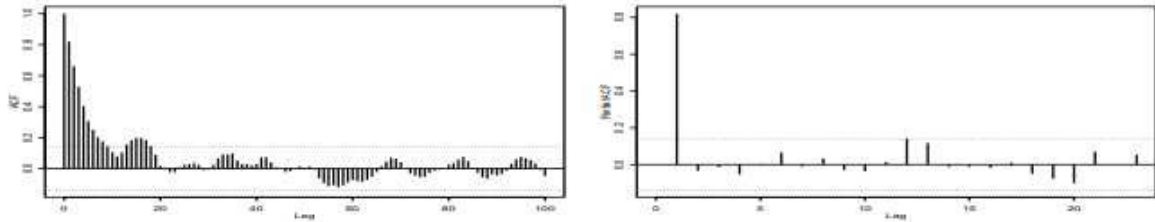


Figure 32: ACF and PACF of AR(1)

Contrarily to the MA(1), in this case it is the PACF that cuts off after lag 1 while the ACF shows a slow decaying pattern. Extending to the general case of AR(p), the PACF cuts off after lag p and the ACF will slowly decrease with the lag. Concluding, the PACF allows to estimate the order  $p$  of AR model by identifying the lag  $n$  in which the value of the function is below the confidence interval (red lines in previous figures) and then  $p=n-1$ .

	AR(p)	MA(q)
ACF	Tails off	Cuts off after lag q
PACF	Cuts off after lag p	Tails off

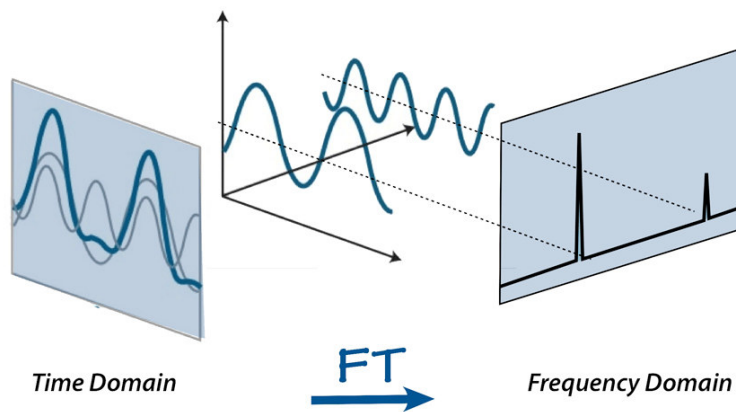
### 3.12 Fourier Transform

The work of Jean-Baptiste Joseph Fourier shows that any signal  $f(t)$  can be written as a sum of sine waves with different amplitudes, frequencies and phases:

$$f(t) = a_0 + a_1 \sin(\omega_1 t + \phi_1) + a_2 \sin(2\omega_2 t + \phi_2) + a_3 \sin(3\omega_3 t + \phi_3) + \dots \quad \text{Equation 33}$$

The Fourier transformation is a mathematical procedure that allows to obtain a frequency domain representation of a time domain signal.:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad \text{Equation 34}$$



33: Illustration of Fourier Transform

Figure

The Discrete Fourier Transform (DFT) is the equivalent of the continuous Fourier Transform for a discrete signal, known only at  $N$  instants separated by sample time  $T$ . Let  $x(k)$  be the  $N$  samples of the continuous signal  $f(t)$ . Each sample of  $x(k)$  can be regarded as an *impulse* with area equal to  $x(k)$ . Then:

$$F(\omega) = \int_{t=0}^{(N-1)T} f(t) e^{-i\omega t} dt \quad \text{Equation 35}$$

$$= x(0) e^{-i0} + x(1) e^{-j\omega T} + \dots + x(k) e^{-j\omega kT} + \dots + x(N-1) e^{-j\omega(N-1)T} \quad \text{Equation 36}$$

Which leads to:

$$X(\omega) = \sum_{k=0}^{N-1} x(k) e^{-i\omega kT} \quad \text{Equation 37}$$

The Fourier Transform can be used to identify patterns in a time series. Consider the following time series:

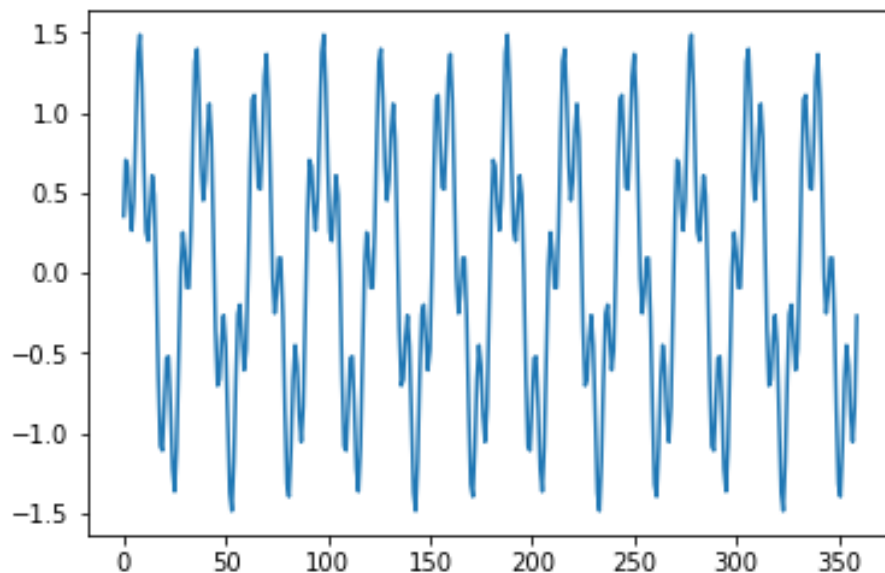
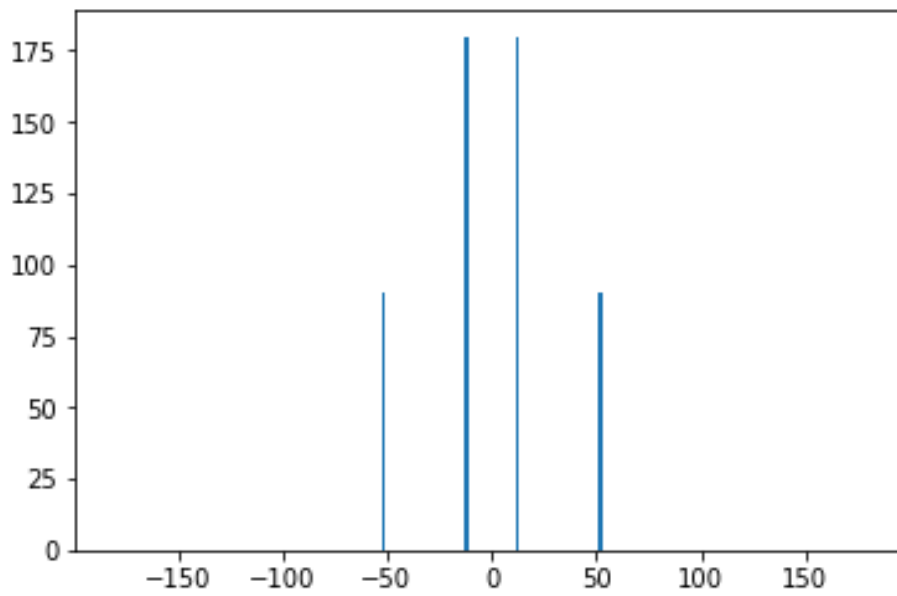


Figure 34: Simulated time series

The discrete Fourier Transform is then computed using Fast Fourier Transform. In Python, this is done through means of the `numpy.fft` function. The results of the FFT are complex numbers, containing information about the amplitude and phase offset of a certain frequency: amplitude is obtained by taking its absolute value and phase by computing its angle. The following two figures show the result obtained.





35: FFT amplitude

Figure

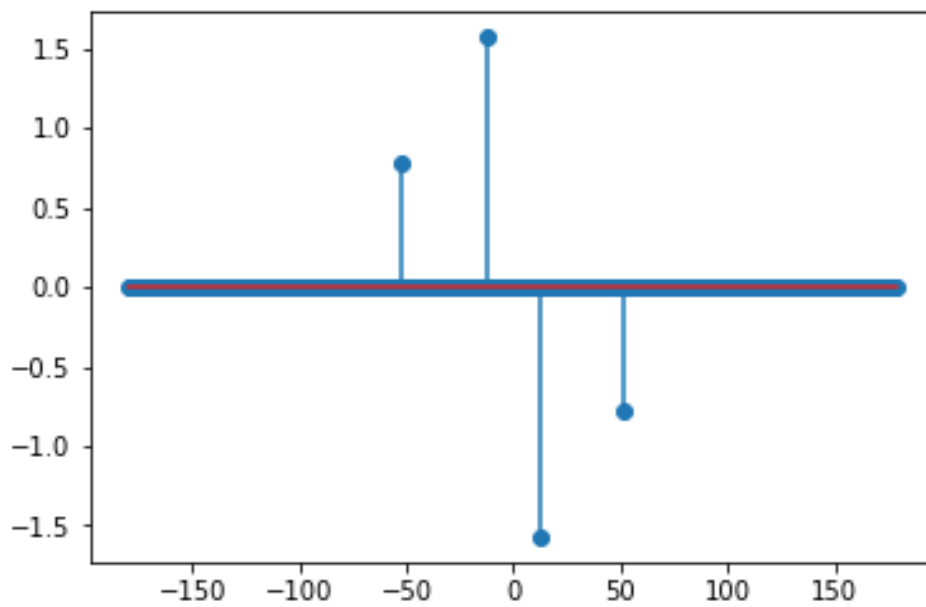


Figure 36: FFT phase

In Figure 34 it is observable that there the only positive frequencies with energy are 12Hz and 52Hz and Figure 35 shows that their phase shifts are  $-\left(\frac{\pi}{2}\right)$  and  $-\left(\frac{\pi}{4}\right)$ , respectively. This is expected since the data series is sampled from a created signal  $f(t)$  that consists in the sum of two cosine waves with those frequencies and phase shifts:

$$f(t) = \cos\left(2\pi * 12t - \frac{\pi}{2}\right) + \frac{1}{2} \cos\left(2\pi * 52t - \frac{\pi}{4}\right)$$

Equation 38

Despite, in case this was real data series representing daily data over one year, the FFT analysis allowed to identify periodicity in the data: one monthly ( $N/12=30.5$ ) and another one weekly ( $N/52=7.04$ ).

## 4. Analysis and applications of Forecasting in Python – Case Study

Python is a high-level programming language that is processed at runtime by the interpreter, running one line at a time and excluding the need of compiling the program before execution. There is the possibility to use an editor (in this work is Spyder) to enter multiple lines of codes at one, however the interpreter will run them individually until the last line (or appear an error).

When comparing to other programming languages, like Java and C++, Python is generally expected to run slower, but the developing time is much lesser: where Python code is typically 3-5 times shorter than equivalent Java code, it is often 5-10 times shorter than equivalent C++ code [XX].

In this section, the forecasting model ARMA of the package *statsmodels* from Python is used to forecast data in telecommunications context. As this model requires data to be stationary, will be made an initial analysis of the data to check for stationarity (differencing if necessary) followed by the application of the forecast model and analysis of the errors.

The data sets used were provided by a communication service provider. This type of data is obtained from the OSS performance management software and contains information regarding a set of RNCs belonging to a 3G network. The data set describes a KPI traffic retrieved by 4 RNCs, where Circuit Switch and Packet Switch fill factors are measured. These measures are commonly used to provide a based proportion related to the RNC licensed capacity being used for voice and data traffic at a specific moment.

### 4.1 Weekly data analysis and forecast – Circuit-switch fill factor

The weekly data (one sample per week) provided corresponding to the each of the 4 RNCs has 146 observations, equivalent to almost three years. Of this, 100 observations are used for modelling and the remaining 46 are used to evaluate the model. The ARMA models are determined for different combinations of  $p$  and  $q$  and evaluated using both expanding and moving window. In Figure 37, there is presented the CS fill factor observations of the four RNCs.

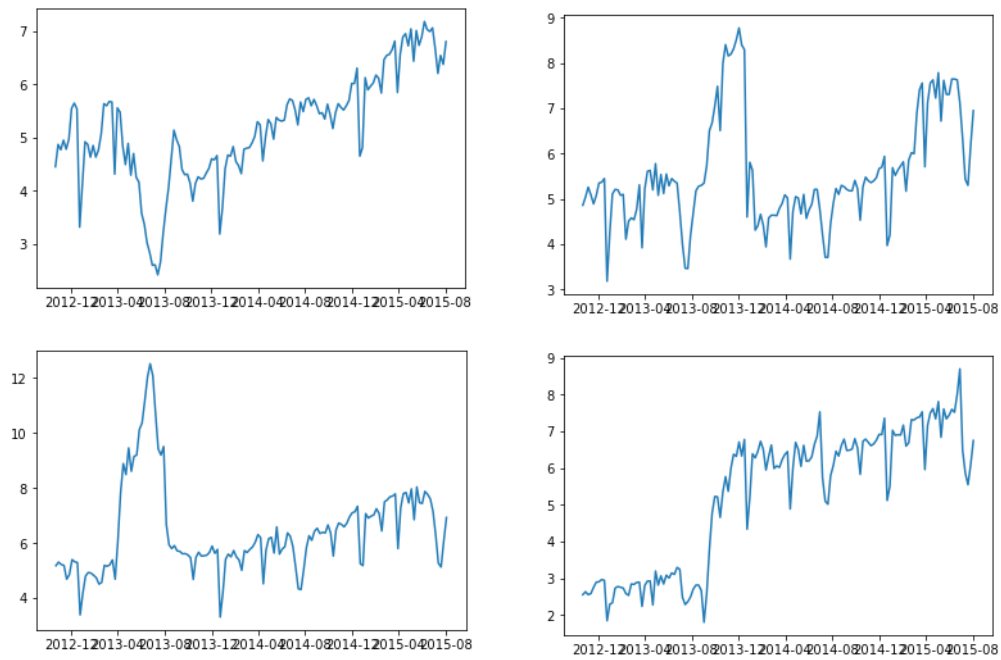


Figure 37: Plots of CSFF data

The first procedure is to determine if the data sets are stationary. Through the analysis of the ACF plots, presented in Figure 38, an exponential decrease in lag and not an abrupt drop to zero is observed. This behaviour in the ACF plot is an indicator of non-stationarity.

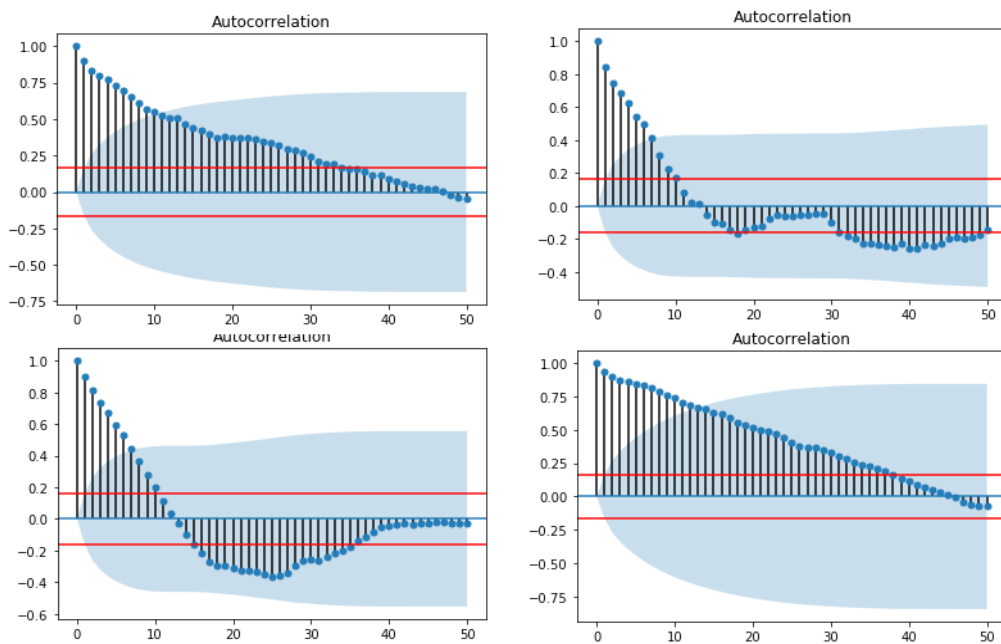


Figure 38: ACF plot of weekly aggregated CS fill factor

To further verify the data stationarity, the Augmented-Dickey-Fuller-test was resorted. The results are presented in next table.

Circuit-switched fill factors	
RNC	ADF p-value
1	0.5204
2	0.0562
3	0.0671
4	0.6492

The p-values are all greater than 0.05 which implies that the null hypothesis is accepted and that the times series are not stationary. Since the ARMA model requires that data is stationary, a first difference is applied to original data sets, shown in Figure 39.

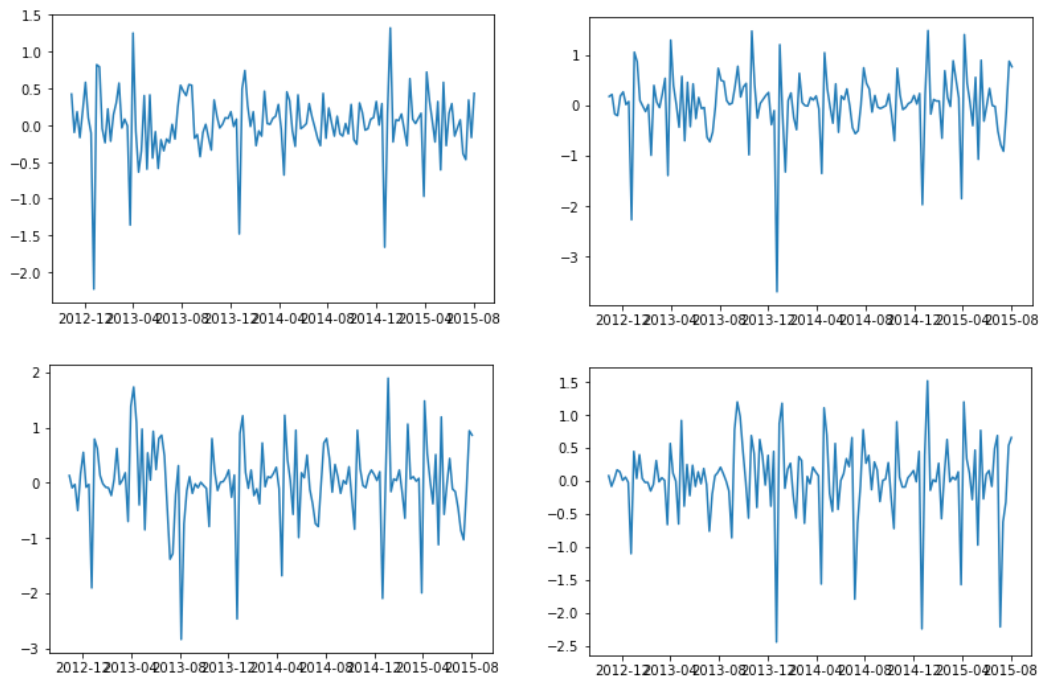


Figure 39: CSFF after first difference

In following table are presented the p-values for the first difference data, confirming its stationarity.

1 <sup>st</sup> difference Circuit-switched fill factors	
RNC	ADF p-value
1	2.36e-20
2	5.59e-27
3	9.52e-23
4	3.26e-13

The ACF of the first differences is shown in Figure 40.

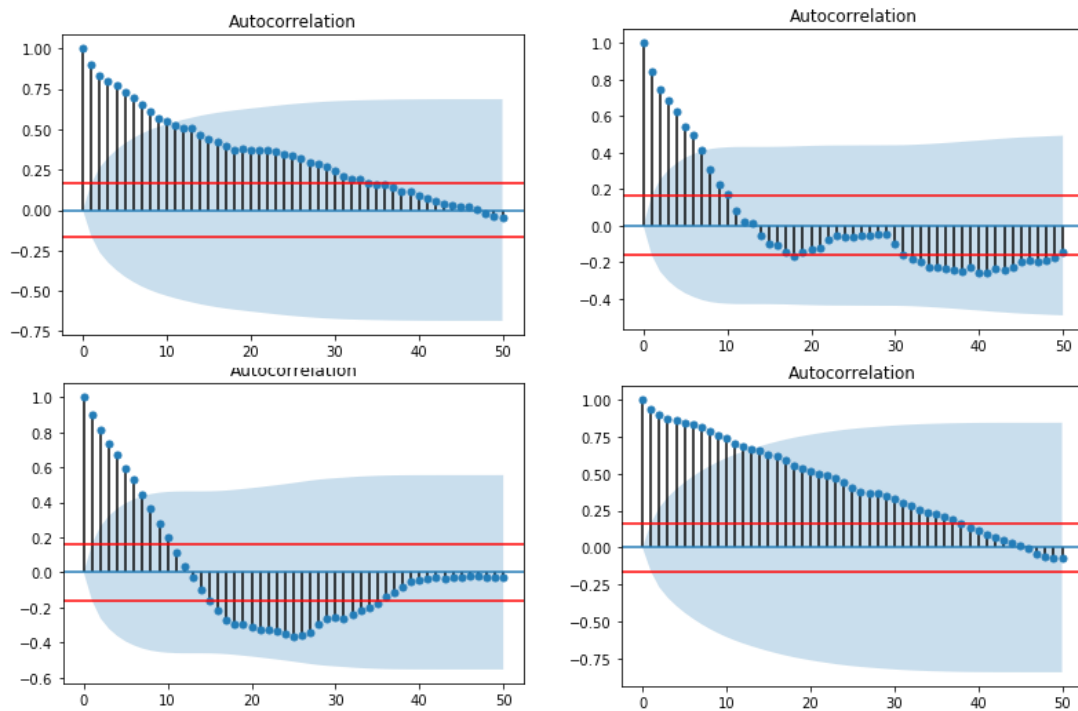


Figure 40: ACF of first difference of weekly CSFF

Based on the ACFs above, there is no seasonal component since there are no identifiable cycles of spikes. The same ACF plots are used to identify possible candidates for the order  $q$ ; the candidates for order  $p$  are recognizable by observing the PACF of the time series.

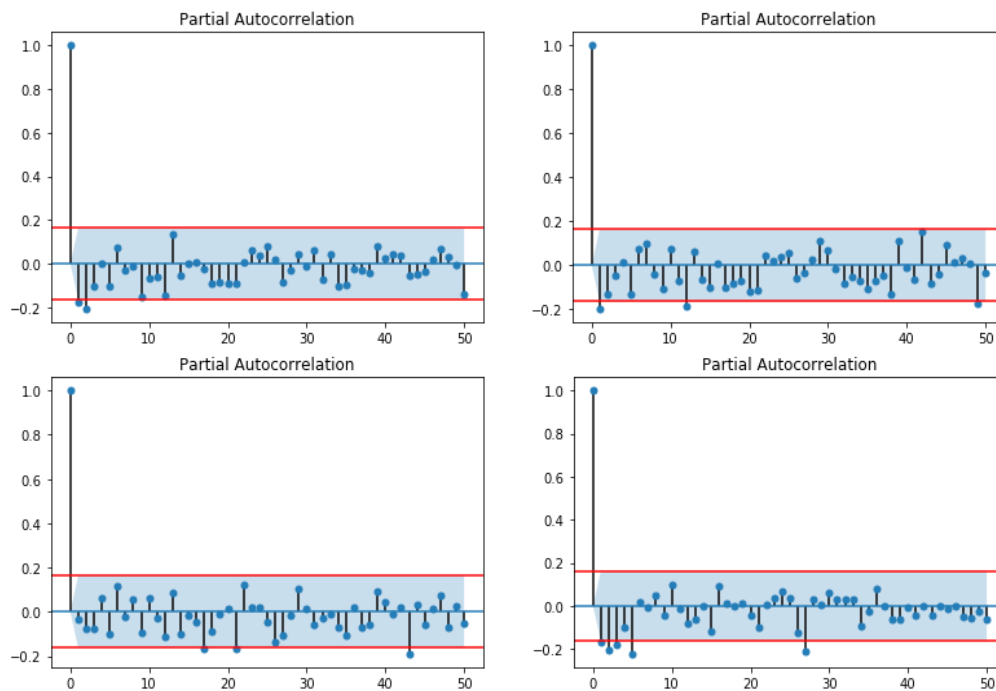


Figure 41: PACF of first difference of weekly CSFF

For the RNC1, the possible candidates are  $q \leq 2$  and  $p \leq 2$ . Different combinations of this orders are used to model ARMA and then the evaluation of the forecasts is made by computing MSE, MAPE and MAE.

	Expanding Window			Moving Window		
	MSE	MAPE	MAE	MSE	MAPE	MAE
<b>ARMA(0,1)</b>	0,1762	4,7149	0,2816	0,1875	4,7202	0,2823
<b>ARMA(0,2)</b>	0,1681	4,7415	0,2847	0,1857	4,7866	0,2869
<b>ARMA(1,0)</b>	0,1850	4,7476	0,2845	0,1889	4,7597	0,2853
<b>ARMA(2,0)</b>	0,1734	4,7253	0,2838	0,1882	4,8035	0,2888
<b>ARMA(2,1)</b>	0,1790	4,8311	0,2902	0,1862	4,8048	0,2885
<b>ARMA(0,0) Rnd Walk</b>	0,1969	4,9887	0,2990	0,1971	4,9874	0,2989

For both expanding window and moving window, the chosen model is ARMA(0,1) which has the lowest MAPE value. In Figure 42,

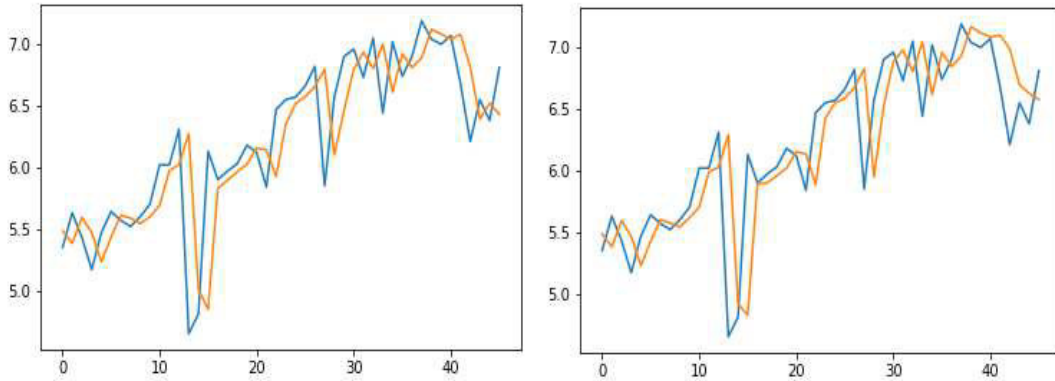


Figure 42: RNC1 weekly CSFF real value (blue line) vs forecast (orange line)

For RNC2, the candidates are  $p \leq 1$  e  $q \leq 1$ . The following table presents the evaluation results.

	Expanding Window			Moving Window		
	MSE	MAPE	MAE	MSE	MAPE	MAE
<b>ARMA(0,1)</b>	0,4355	8,1089	0,4745	0,4442	8,0692	0,4729
<b>ARMA(1,0)</b>	0,4462	8,0629	0,4724	0,4519	8,0638	0,4722
<b>ARMA(1,1)</b>	0,4296	8,0631	0,4722	0,4621	8,2117	0,4836
<b>ARMA(0,0) Rnd Walk</b>	0,4559	7,9493	0,4651	0,4568	7,9569	0,4659

In this case, the lowest MAPE value is for  $p=0$  and  $q=0$ . Notice that performing an ARMA(0,0) is equivalent to Random Walk.

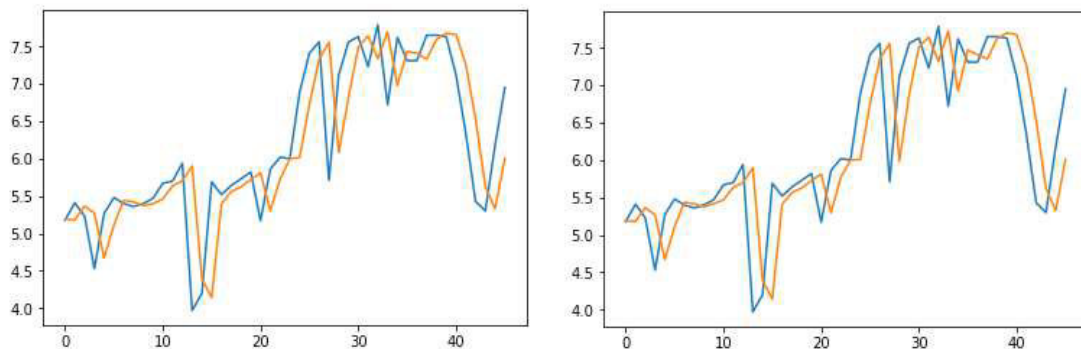


Figure 43: RNC2 weekly CSFF real value (blue line) vs forecast (orange line)



For the RNC3, ACF and PACF do not show significant spikes, so the chosen values for the orders are  $p=0$  and  $q=0$ .

	Expanding Window			Moving Window		
	MSE	MAPE	MAE	MSE	MAPE	MAE
<b>ARMA(0,0)</b>	0,5543	7,6664	0,5008	0,5529	7,6418	0,4998

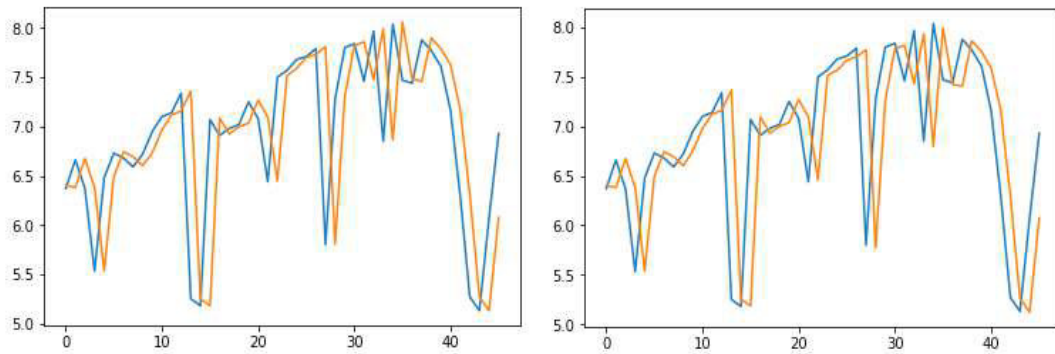


Figure 44: RNC3 weekly CSFF real value (blue line) vs forecast (orange line)

At last, for the RNC4 the values are  $p \leq 3$ ,  $p=5$  and  $q \leq 2$ .

	Expanding Window			Moving Window		
	MSE	MAPE	MAE	MSE	MAPE	MAE
<b>ARMA(0,1)</b>	0,4741	7,0312	0,4541	0,4803	7,0604	0,4457
<b>ARMA(0,2)</b>	0,4262	6,5219	0,4211	0,4348	6,5303	0,4209
<b>ARMA(1,0)</b>	0,4872	6,8630	0,4466	0,4899	6,8540	0,4458
<b>ARMA(2,0)</b>	0,4644	6,9791	0,4529	0,4712	6,9957	0,4535
<b>ARMA(2,1)</b>	0,4265	6,5103	0,4208	0,4328	6,4632	0,4199
<b>ARMA(3,0)</b>	0,4478	6,6795	0,4323	0,4562	6,7069	0,4335
<b>ARMA(3,1)</b>	0,4282	6,5506	0,4235	0,4354	6,5085	0,4199
<b>ARMA(5,0)</b>	0,4167	6,6430	0,4302	0,4252	6,6929	0,4328
<b>ARMA(5,1)</b>	0,4189	6,6739	0,4325	0,4298	6,7311	0,4355
<b>ARMA(0,0) Rnd Walk</b>	0,5016	6,9918	0,4557	0,5022	6,9802	0,4546

The ARMA(2,1) has the lowest MAPE and the forecast are in the following figure.

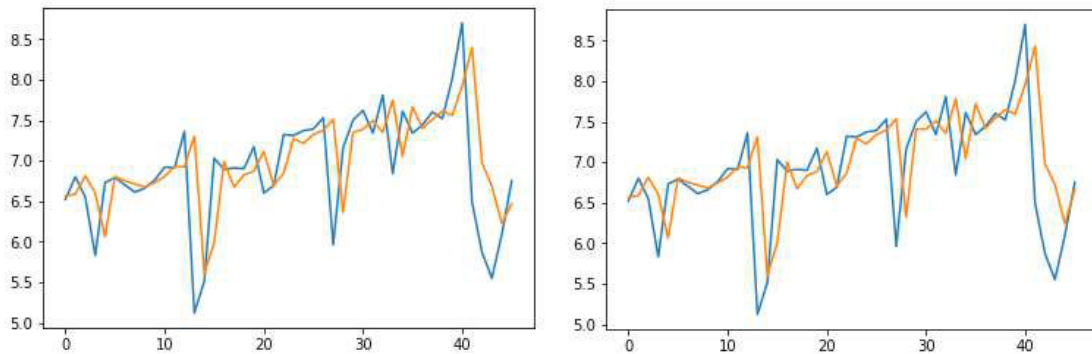


Figure 45: RNC4 weekly CSFF real value (blue line) vs forecast (orange line)

#### 4.2 Weekly data analysis and forecast – PS fill factor

The time series representation Packet-Switch fill factor have the same size of the CS fill factor, 146 observations. Like in CS case, 100 observations are used for modelling and the remaining 46 are used to evaluate the model using expanding and moving windows. The data is shown in Figure 46.

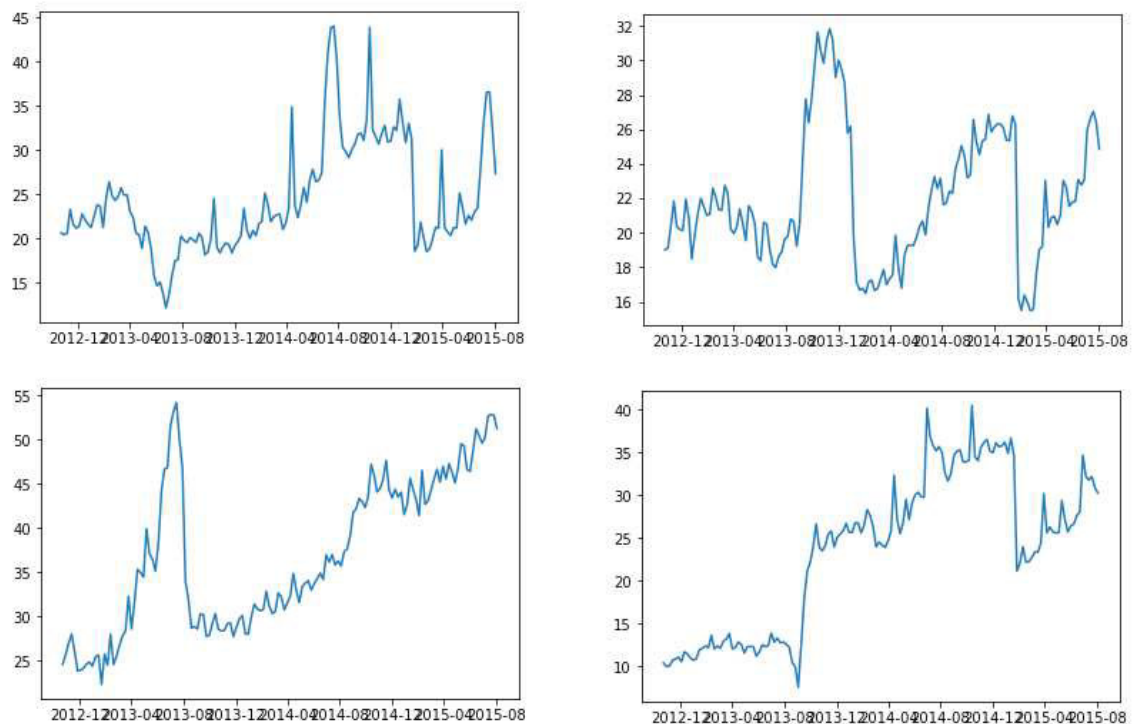


Figure 46: Weekly PSFF of four RNCs

Observing the plots, RNC1 has an increasing tendency after 2013-08; RNC2 seem to not have trend in the beginning until the big spike, showing a trend after; RNC3 and RNC4 the trend is identifiable along the series.

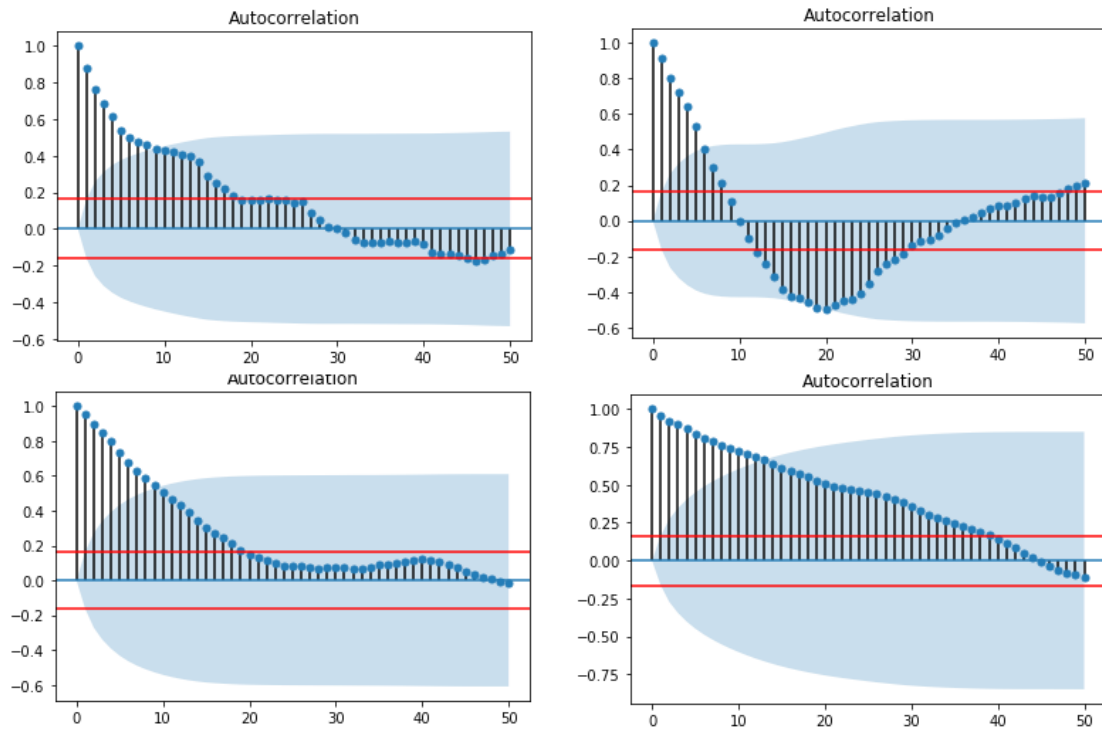


Figure 47: ACF plot of weekly aggregated PS fill factor

To remove the trend in the data and obtain stationarity, the first difference is applied to the data provided.

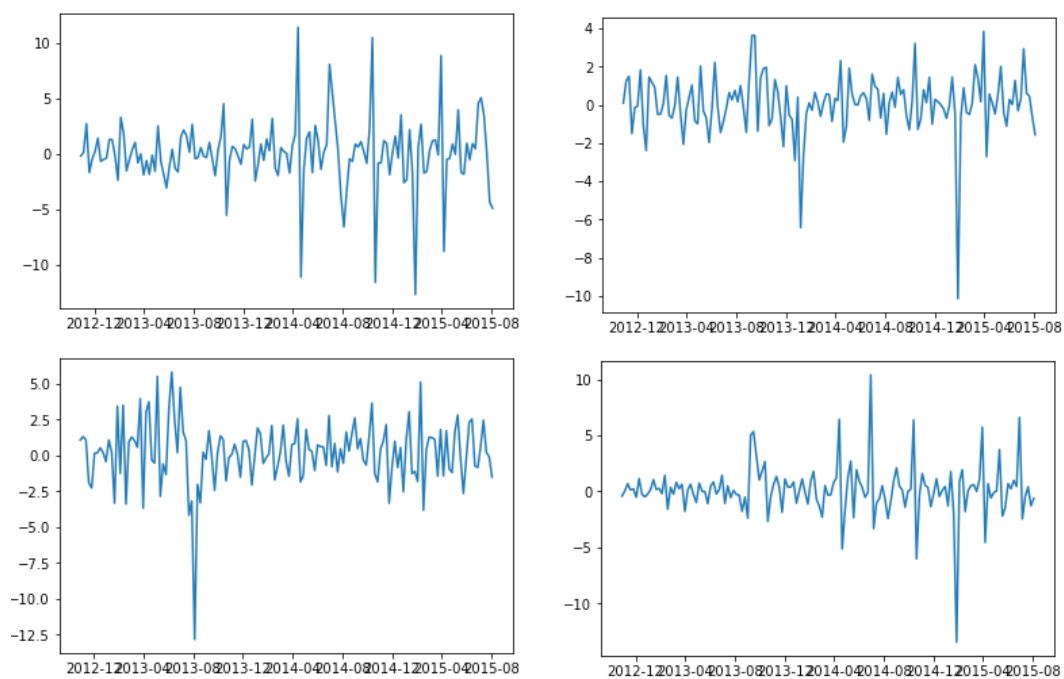


Figure 48: PS fill factor after first difference

To ensure the stationarity is performed the Augmented-Dickey-Fuller-test.

<b>1<sup>st</sup> difference Packet-switched fill factors</b>	
<b>RNC</b>	<b>ADF p-value</b>
1	2.36e-20
2	5.59e-27
3	9.52e-23
4	3.26e-13

Now that the data is stationary through first differencing, the ACF and PACF are computed to identify the possible values for  $p$  and  $q$ .

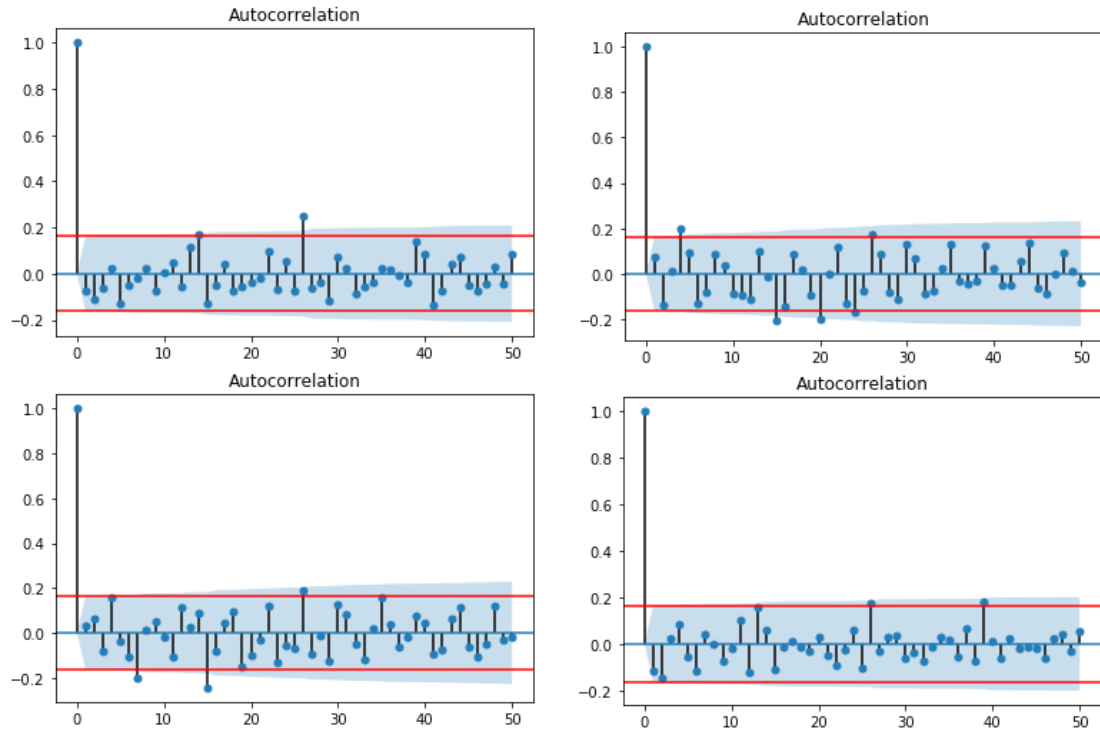


Figure 49: ACF of first differences of weekly PSFF

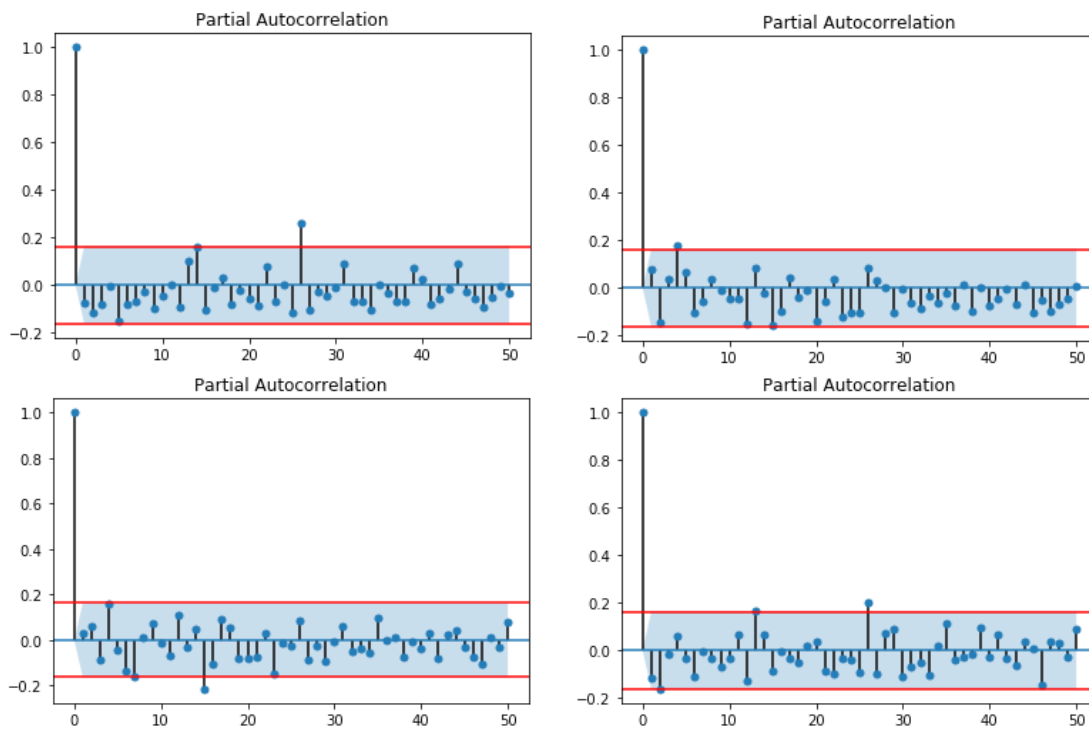


Figure 50: PACF of the first difference of weekly PSFF

For all RNCs except the RNC2, the candidates for  $p$  and  $q$  are equal:  $p=0$  and  $q=0$ . RNC2 has two more candidates in  $p=4$  e  $q=4$ . Since the  $p$  and  $q$  for RNC1, RNC3 and RNC4 are equal and there is only combination possible, the results for the three are presented in the same table , leaving the RNC2 for last.

		Expanding Window			Moving Window		
		MSE	MAPE	MAE	MSE	MAPE	MAE
<b>RNC 1</b>	<b>ARMA(0,0)</b>	16,5089	9,6699	2,6166	16,5178	9,6891	2,6215
<b>RNC 3</b>	<b>ARMA(0,0)</b>	3,4943	3,3535	1,5041	3,5418	3,3531	1,5404
<b>RNC 4</b>	<b>ARMA(0,0)</b>	9,1874	6,0747	1,7392	9,2110	6,0937	1,7447

The forecast results for the three RNCs are presented in the following figures.

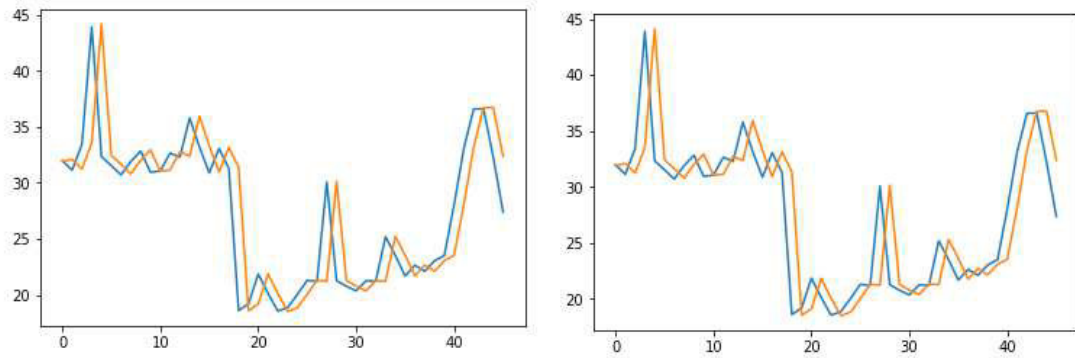


Figure 51: RNC1 weekly PSFF real value (blue line) vs forecast (orange line)

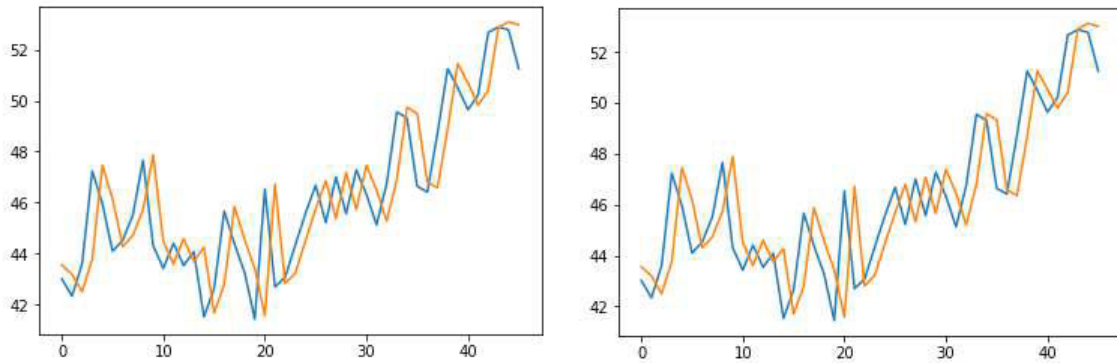


Figure 52: RNC3 weekly PSFF real value (blue line) vs forecast (orange line)

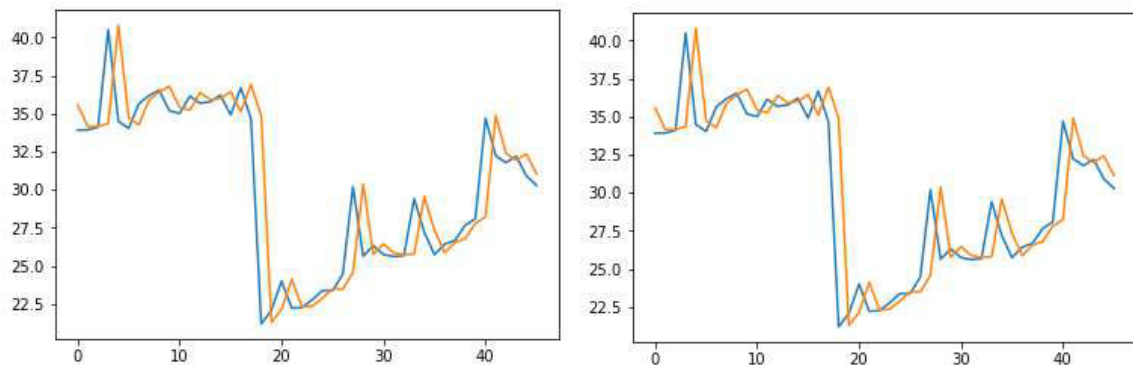


Figure 53: RNC4 weekly PSFF real value (blue line) vs forecast (orange line)

For the RNC2, besides  $p=0$  and  $q=0$  like the other RNCs, the ACF and PACF also suggests  $p=4$  and  $q=4$  as candidates.

	Expanding Window			Moving Window		
	MSE	MAPE	MAE	MSE	MAPE	MAE
<b>ARMA(0,0)</b>	3,8262	5,2522	1,1013	3,8304	5,2625	1,1039
<b>ARMA(4,0)</b>	4,1251	6,1496	1,2751	4,0947	6,0375	1,2563
<b>ARMA(0,4)</b>	4,0618	5,8150	1,2221	4,1068	5,9077	1,2396

Despite having more  $p$  and  $q$  candidates, and choosing the model with the lower MAPE, the best model is ARMA(0,0) just as for the other three RNCs. The forecast results are shown in Figure 53.

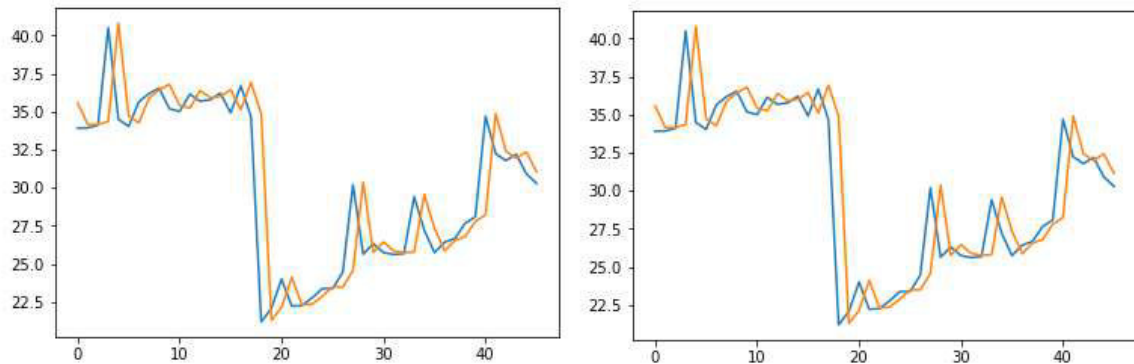


Figure 54: RNC2 weekly PSFF real value (blue line) vs forecast (orange line)



## **5. Conclusion**

The constant evolution and development of the communication technology is leading to a point where telecommunication service providers may have some difficulties to satisfy the needs of the user. Hence, forecasting is an important tool for the service provider to be able to predict the behaviour of the network, with the goal of helping to better plan network modifications and prevent problems. However, the task of obtaining a good forecast is difficult.

The case study in chapter 4 is a practical application of a forecasting model to data from a real case of weekly aggregated CS and PS. This data was analyzed, stationarized by differencing and applied to an ARMA model. Then the accuracy measure were taken and the model with the lowest MAPE was chosen. Although it was used both expanding and moving window, the lowest MAPE for both methods do not present significant differences, leading to a conclusion that both methods can be used in real cases.

More than the case study results, this work has shown that is possible use historical to apply a statistical model to historical data and forecast future data, providing useful information to the communication service provider.

## **6. Future Work**

The number of forecasting methods is vast, so it would be interesting to explore other forecasting methods that were not covered by this work. Related with the KPIs, this is also a vast area that can be explored by forecasting different KPI in order to forecast other parts of the network.

Another important aspect in this area is the automation of the modelling and forecasting procedure in order to include it in the monitoring component of the network.

## References

- [1] X. Hu and J. Wu, "Traffic Forecasting Based on Chaos Analysis in GSM Communication Network," in *2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007)*, 2007.
- [2] A. M. Álvarez-Socarrás, A. Berrones, G. J. Moreno, J. A. Rodríguez-Sarasty, and M. Cabrera-Ríos, "Practice Summary: Enhancing Forecasting and Capacity Planning Capabilities in a Telecommunications Company," *Interfaces (Providence)*, vol. 43, no. 4, Aug. 2013.
- [3] R. J. Hyndman and G. Athanasopoulos, "Forecasting : principles and practice," 2013. [Online]. Available: <https://www.otexts.org/fpp/8/1>. [Accessed: 1-Oct-2018].
- [4] C. N. Konstantinopoulou, K. A. Koutsopoulos, G. L. Lyberopoulos, and M. E. Theologou, "Core network planning, optimization and forecasting in GSM/GPRS networks," in *IEEE Benelux Chapter on Vehicular Technology and Communications. Symposium on Communications and Vehicular Technology. SCVT-2000. Proceedings (Cat. No.00EX465)*.
- [5] J. S. Armstrong, *Principles of forecasting : a handbook for researchers and practitioners*. Kluwer Academic, 2001.
- [6] M. Sokele, "Analytical Method For Forecasting Of Telecommunications Service Life-Cycle Quantitative Factors Analitički Postupak Predviđanja Kvantitativnih Čimbenika Životnog Vijeka Telekomunikacijske Usluge," 2009.
- [7] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications*, vol. 102. Springer New York, 2006.
- [8] R. D. Reid and N. R. Sanders, *Operations management : an integrated approach*. John Wiley, 2010.
- [9] S. Makridakis, R. J. Hyndman, and S. C. Wheelwright, *Forecasting : methods and applications*. Wiley, 1998.
- [10] E. H. Hassani, "Causal Method and Time Series Forecasting model based on Artificial Neural Network," *Int. J. Comput. Appl.*, vol. 75, no. 7, 2013.
- [11] D. C. Montgomery, C. L. Jennings, and M. Kulahci, *Introduction to Time Series Analysis and Forecasting*. Wiley Series in Probability and Statistics, 2008.
- [12] J. and E. W. S. M. P.F. Miller, "Introduction to Time Series Regression and Forecasting," *University of Pennsylvania*. [Online]. Available: [http://www.ssc.upenn.edu/~fdiebold/Teaching104/Ch14\\_slides.pdf](http://www.ssc.upenn.edu/~fdiebold/Teaching104/Ch14_slides.pdf). [Accessed: 29-Aug-2018].

- [13] J. Caiado, *Métodos de previsão em gestão : com aplicações em excel*, 2nd ed. Lisboa: Sílabo, 2016.
- [14] Statistics Canada, “Seasonal adjustment and trend-cycle estimation,” 2015. [Online]. Available: <http://www.statcan.gc.ca/pub/12-539-x/2009001/seasonal-saisonnal-eng.htm>. [Accessed: 20-Jun-2018].
- [15] “Extract Seasonal & Trend: using decomposition in R - Anomaly,” *anomaly*, 2015. [Online]. Available: <https://anomaly.io/seasonal-trend-decomposition-in-r/>. [Accessed: 05-Mar-2018].
- [16] G. van de Ven, “Removal of Trend & Seasonality,” 2010. [Online]. Available: <https://www.stat.berkeley.edu/~gido/Removal of Trend and Seasonality.pdf>. [Accessed: 08-Apr-2018].
- [17] S. Makridakis and M. Hibon, “The M3-Competition: results, conclusions and implications,” *Int. J. Forecast.*, vol. 16, 2000.
- [18] M. D. O. Duarte, D. Teixeira, B. D. O. Cruz, and J. P. V. Dias, “Cellular Telecommunications network: planning, dimensioning and economics,” Aveiro, 2015.
- [19] R. F. Nau, “Statistical forecasting: notes on regression and time series analysis.” [Online]. Available: <https://people.duke.edu/~rnau/411avg.htm#SMA>. [Accessed: 23-Jun-2018].
- [20] W. J. Stevenson, *Operations management*, 11th ed. Rochester Institute of Technology, 2012.
- [21] R. Upadhyay, “Forecasting & Time Series Analysis – Manufacturing Case Study,” 2015. [Online]. Available: <http://ucanalytics.com/blogs/forecasting-time-series-analysis-manufacturing-case-study-example-part-1/>. [Accessed: 19-Mar-2018].
- [22] The Pennsylvania State University, “14.1 - Autoregressive Models | STAT 501,” 2016. [Online]. Available: <https://onlinecourses.science.psu.edu/stat501/node/358>.
- [23] “Seasonal Arma” [Online]. Available: [http://www.maths.qmul.ac.uk/~bb/TS\\_Chapter7\\_2.pdf](http://www.maths.qmul.ac.uk/~bb/TS_Chapter7_2.pdf) [Accessed: 17-Sep-2018]
- [24] R. M. Kunst, “Evaluating predictive accuracy,” vol. 1, no. 1, Department of Economics of the University of Vienna, 2004.
- [25] International Telecommunication Union, “Models for forecasting international traffic,” *Blue Book*, vol. E.507. ITU, 1993.
- [26] J. H. Stock and M. W. Watson, *Introduction to econometrics*, Third edit. Addison -Wesley.
- [27] N. S. Nalawade and M. M. Pawar, “Forecasting telecommunications data with Autoregressive Integrated Moving Average models,” in *2015 2nd International Conference on Recent Advances in Engineering & Computational Sciences (RAECS)*, 2015

- [28] J. Korhonen, *Introduction to 3G mobile communications*, Second edi. London: Artech House, 2003.
- [29] Huawei technologies, “Huawei Ran KPI for performance Management,” unknow, 2006.
- [30] I. Ray Khastur, LTE Optimization Consultant at P.I.Works, “KPI in LTE Radio Network (Huawei Based),” 2015. [Online]. Available: <http://www.slideshare.net/RayKhastur/kpi-in-lte-radio-network-huawei-based>. [Accessed: 09-Jan-2017].
- [31] P. Sharma, “Evolution of Mobile Wireless Communication Networks-1G to 5G as well as Future Prospective of Next Generation Communication Network,” *Int. J. Comput. Sci. Mob. Comput. - not index*, vol. 2, no. August
- [32] José Silva, “Proposta 3GPP de Indicadores de Desempenho de Rede – R4 CS Core Network”, Universidade de Aveiro, 2011
- [33] V. Pereira and T. Sousa, “Evolution of Mobile Communications: from 1G to 4G,” *Dep. Informatics Eng. Univ. Coimbra*, p. 7, 2004.
- [34] H. Kaaranen, “UMTS Networks : Architecture, Mobility and Services,” *John Wiley Sons*, 2005.
- [35] Harri Holma and Antti Toskala, “HSDPA/HSUPA for UMTS: High Speed Radio Access for Mobile Communications”, John Wiley & Sons, 2006.
- [36] Gonzalo Camarillo, Miguel A. García-Martín, “The 3G IP Multimedia Subsystem (IMS): Merging the Internet and the Cellular Worlds”, John Wiley & Sons, 2006.
- [37] Dhruv Sunil Shah, “A tutorial on LTE Evolved UTRAN (EUTRAN) and LTE Self Organizing Networks”, The University of Texas at Arlington, December 2010.
- [38] J. Reunanen, J. Salo, and R. Luostari, “LTE Key Performance Indicator Optimization,” in *LTE Small Cell Optimization*, Chichester, UK: John Wiley & Sons Ltd, 2015.
- [39] “ARMA Models – ACF and PACF” [Online] Availabe: [http://www.maths.qmul.ac.uk/~bb/TimeSeries/TS\\_Chapter6\\_2\\_2.pdf](http://www.maths.qmul.ac.uk/~bb/TimeSeries/TS_Chapter6_2_2.pdf)