AUTOREGRESSIVE INTEGRATED MOVING AVERAGE-BASED PREDICTIVE MODEL FOR BASE STATION AVAILABILITY OF TELECOMMUNICATION NETWORKS IN MINNA

BY

OHIHOIN, Emmanuel Esebanme
MEng/SEET/2018/8522

DEPARTMENT OF TELECOMMUNICATION ENGINEERING
FEDERAL UNIVERSITY OF TECHNOLOGY MINNA

NOVEMBER, 2021

AUTOREGRESSIVE INTEGRATED MOVING AVERAGE-BASED PREDICTIVE MODEL FOR BASE STATION AVAILABILITY OF TELECOMMUNICATION NETWORKS IN MINNA

BY

OHIHOIN, Emmanuel Esebanme
MEng/SEET/2018/8522

A THESIS SUBMITTED TO THE POSTGRADUATE SCHOOL, FEDERAL UNIVERSITY OF TECHNOLOGY, MINNA, NIGERIA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF ENGINEERING (M.ENG) IN COMMUNICATION ENGINEERING.

NOVEMBER, 2021

ABSTRACT

There is a standard of 99.999% (five 'nines') availability for telecommunication hardware and software. This is to guarantee the high level of service required by the Mobile Network Operator (MNO) for service delivery. MNOs in Nigeria and most sub-Saharan Africa countries are, however, not being able to meet up with the expected base station availability mainly due to high restoration time after the outage. In this thesis, the historical Base Transceiver Station (BTS) Availability reports of a thousand data points each for four MNOs were used. The MNOs (MNO W, MNO X, MNO Y and MNO Z in Minna) data were acquired from 1st of January 2018 to 26th September 2020. The first 73% of the data was partitioned into the Training period and the remaining 27% was set for Validation. The data is in the form of Time Series (TS) and was modelled using Autoregressive Integrated Moving Average (ARIMA) prediction. Correlation plots of the data were done and the ARIMA (p,d,q) parameters were got with the aid of the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF). The ARIMA-Based models for the MNOs are ARIMA (0,1,3), ARIMA (1,0,1), ARIMA (2,0,4) and ARIMA (0,1,1) for MNO W, MNO X, MNO Y and MNO Z, respectively. The predictive models were used to predict BTS Availability for the MNOs from 27th September 2020 to 20th December 2020. The performance of the models was evaluated with data in the validation period for Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The MAEs for the respective MNOs are: 1.3959, 0.6602, 1.5666 and 0.6177; while their MAPE are: 0.0150, 0.0068, 0.0176 and 0. 0063. The long short-term (LSTM) model was used for comparison with the ARIMA model for the same MNOs and their MAE and MAPE are 2.8397, 0.8894, 2.8223, and 1.1245; 0.0322, 0.0092, 0.0349 and 0.0118 for MNO W, MNO X, MNO Y and MNO Z respectively. From the results, it is observed that the LSTM models have higher MAE values than the ARIMA models by 51%, 26%, 44% and 45% for MNO W, MNO X, MNO Y and MNO Z respectively. Similarly, for MAPE, the LSTM models have 53%, 26%, 50% and 47% higher values than the ARIMA models for the respective MNOs. These indicate that the ARIMA models have performed better than the LSTM models in all the MNOs. The values of the MAE and MAPE for the predictive models are very low which implies that the predicted Availability data is close to the actual values and can be used for proper planning and decision-making. MNOs can proactively schedule Predictive Maintenance (PdM) with the PdM algorithm developed in this work. Using the 95% availability threshold of this algorithm, MNO W and MNO Y have no savings in maintenance count, while MNO X and MNO Z have savings of 33 and 32 respectively.

TABLE OF CONTENTS

Cover Page	i
Title Page	ii
DECLARATION	**ii**
CERTIFICATION	**iii**
ACKNOWLEDGEMENT	**iv**
ABSTRACT	**v**
TABLE OF CONTENTS	**vi**
LIST OF TABLES	**ix**
LIST OF FIGURES	**x**
LIST OF PLATES	**xi**
ABBREVIATIONS	**xii**
CHAPTER ONE	**1**
1.0 INTRODUCTION	**1**
1.1 Overview	**1**
1.2 Statement of the Research Problem	**4**
1.3 Aim and Objectives of the Study	**5**
1.4 Scope of the Study	**5**
1.5 Justification for the Study	**6**
1.6 Thesis Outline	**7**

CHAPTER TWO **8**

2.0 LITERATURE REVIEW **8**

2.1 Overview of Literature review **8**

2.2 Theoretical background **17**

2.3 Research Gap **26**

CHAPTER THREE **27**

3.0 RESEARCH METHODOLOGY **27**

3.1 Data Acquisition and Processing **27**

3.2 Data Preparation **29**

3.3 Modelling Using the TS Data **29**

3.4 PdM Scheduling Algorithm **37**

CHAPTER FOUR **40**

4.0 RESULTS AND DISCUSSION **40**

4.1 Plot of BTS Availability for MNO W, X, Y and Z **40**

4.2 Creating the Predictive Models for the MNOs **48**

4.3 Results of the LSTM Models **54**

4.4 Prediction Accuracy **56**

4.5 Result of Scheduling Algorithm **60**

CHAPTER FIVE **63**

5.0 CONCLUSION AND RECOMMENDATIONS **63**

5.1 Conclusion **63**

5.2 Recommendations **64**

5.3 Contributions to Knowledge **65**

REFERENCES **66**

APPENDIX A **70**

APPENDIX B **76**

APPENDIX C **78**

LIST OF TABLES

Table	Title	Page
2.1	Summary of Literature Review	13
2.2	Availability and Corresponding Downtime	19
3.1	Base Station Site Count for the MNOs under Consideration	27
4.1	Model Parameters for the MNOs	48
4.2	Predicted BTS Availability for MNO	51
4.3	Performance Metrics for the ARIMA Model	57
4.4	Performance Metrics Comparison for the ARIMA and LSTM Models	58
4.5	Scheduling Algorithm of the MNOs Using Thresholds	61

LIST OF FIGURES

Figure	Title	Page
1.1 A Typical Telecommunication Network		3
1.2 Evolution of Mobile Network Technology		4
2.2 Availability of Systems (a) Series (b) parallel		19
2.3 Long Short-Term Memory Block.		23
3.1 Schematic of Base Stations and their Serving Network Elements for MNO X		28
3.2 Flowchart of the ARIMA-Based Predictive Model		30
3.3 Flowchart of the LSTM Model		36
3.4 PdM Scheduling Algorithm		39
4.1 Plots of Base Station Availability in Groups of 200 Days for MNO		43
4.2 Percentage of Base Station Availability Over 95% Considering Groups of 50 Days Over the 1000 days		43
4.3 Percentage of Base Station Availability Over 98% Considering Groups of 50 Days Over the 1000 days		44
4.4 (a) Autocorrelation Plot for BTS Availability for MNO W		45
4.5 Plot of the transformed TS for MNO W with a First Order Differencing		46
4.6 ACF Plot for the Transformed TS Data for MNO W		47
4.7 Plot of predicted availability for MNO W (Validation period)		49
4.8 Plot of predicted availability for MNO X (Validation period)		49
4.9 Plot of predicted availability for MNO Y (Validation period)		50
4.10 Plot of predicted availability for MNO Z (Validation period)		50
4.13 Predicted BTS Availability for MNO Y Using ARIMA (2,0,4) Model		53
4.14 Predicted BTS Availability for MNO Z Using ARIMA (0,1,1) Model		54
4.15 LSTM Predicted BTS Availability for (a) MNO W, (b) MNO X, (c) MNO Yand (d) MNO Z		56
4.16 MAE against MNO		57
4.17 MAPE against MNO		58
4.18 Performance Chart for the MNO		58
4.19 MAE Comparison for ARIMA and LSTM		59
4.20 MAPE Comparison for ARIMA and LSTM		59

LIST OF PLATES

Plate	Title	Page
I	BTS3900	76
II	BSC6900	76
III	UMG8900	77
IV	OptiX900	77
V	Tower	77
VI	A Rack of Microwave Radio Indoor Units for NEs	77

ABBREVIATIONS

Acronyms	**Meaning**
ACF	Autocorrelation Function
ARIMA	Autoregressive Integrated Moving Average
BCF	Base Station Control Function
BSS	Base Station Subsystem
BSC	Base Station Controller
BTS	Base Transceiver Station
DaaS	Desktop as a Service
EDGE	Enhanced Data rates for GSM Evolution
EMTS	Emerging Market for Telecommunication Service
GPRS	General Packet Radio Service
GSM	Global System for Mobile communication
HW	Hardware
ICS	Integrated Communication System
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MNO	Mobile Network Operator
MS	Mobile Station
MSS	Mobile Switching Server
MTBF	Mean Time Between Failure
MTTR	Mean Time To Repair

NA	Network Availability
NE	Network Element
NFV	Network Function Virtualisation
O&M	Operation and Maintenance
PACF	Partial Autocorrelation Function
PdM	Predictive Maintenance
PPM	Planned Preventive Maintenance
RAN	Radio Access Network
RBD	Reliability Block Diagram
RFU	Radio Frequency Unit
RMSE	Root Mean Square Error
SDH	Synchronous Digital Hierarchy
SDN	Software Defined Networking
SMS	Short Message Service
TAS	Telecommunication Application Server
TDM	Time Division Multiplexing
TRX	Transmit/Receive
TS	Time Series
TSA	Time Series Analysis
UAS	Unavailable Second
UE	User Equipment
Um	U-mobile interface for the communication between a BTS and MS
UMG	Universal Media Gateway

UMTS	Universal Mobile Telecommunication System
VIM	Virtualised Infrastructure Management

CHAPTER ONE

1.0 INTRODUCTION

1.1 Overview

The massive influx of cellular mobile technology has made virtually all aspects of human activities dependent on the use of telecommunications services. For instance, in Nigeria, according to the Nigerian Communication Commission (NCC), the number of active lines in the Global System for Mobile Communications (GSM) rose from 148,774,015 in April 2017 to 198,961,361 in July 2020. A few decades ago, the delivery of telecommunications services like voice calls, SMS and the internet were just emerging and users of the services were not too keen about the extent of the availability of the service. Today, the narrative has changed tremendously; high network availability is usually requested from the Mobile Network Operators (MNO).

Subscribers are anticipating a significant degree of service, from which availability is viewed as the major evaluation of quality (Mahdi *et al*., 2018). Availability measurement might be utilised as a contribution to attract customers, but perhaps more importantly, the operators may profit from its use in evaluating the overall quality of the network (Thulin, 2004).

This research is on the Prediction of Base Station Availability for Telecommunication Operators in Minna. This work can be of use to the MNO in the planning process to focus on the aspects of the Network Elements (NE) maintenance. The areas in the NE that are prone to failure could be provided with redundancy. The major MNOs in Nigeria are 9mobile or EMTS (formerly known as Etisalat), Globacom Mobile, Airtel Nigeria and MTN Nigeria.

1.1.1 The Base Transceiver Station (BTS) and the Mobile Cellular Network

The BTS is the cell phone's admission point to the network (Qing, 2017). It is accountable for carrying out radio communications between the network and the mobile phone. It executes speech encoding, encryption, multiplexing (Time Division Multiplexing), and modulation/demodulation of the radio signals.

The U-mobile (Um) interface is a sort of radio interface liable for the communication between the mobile station and the BTS. It makes available the link joining the Mobile Station (MS) and the GSM system. Its physical linking is accomplished through the radio waves. The Um interface is the main interface amongst all the interfaces in the GSM framework.

A BTS is supervised by a Base Station Controller (BSC) using the Base Station Control Function (BCF). The BCF is designed as a discrete unit or even combined in a Transceiver (TRX) in compact base stations. The BCF offers an Operations and Maintenance (O&M) connection to the Network Management System (NMS) and accomplishes the operational conditions of each Transmit/Receive (TRX), as well as software management and alarm gathering (Qing, 2017). The basic structure and roles of the BTS remain the same irrespective of the wireless technologies.

According to Mahdi *et al.* (2018), a mobile cellular network is described as a communication infrastructure comprising Network Elements (NEs) that allow Mobile Stations or User Equipment (UE) to access network services through radio channels. Mobile cellular networks usually span large geographical service area which is subdivided into smaller service areas known as cells. Each cell has a fixed access point called a BTS (NodeB for 3G, eNodeB for

4G and gNodeB for 5G) which resides within the cell for wireless communication with UE. Figure 1.1 shows a typical communications network.

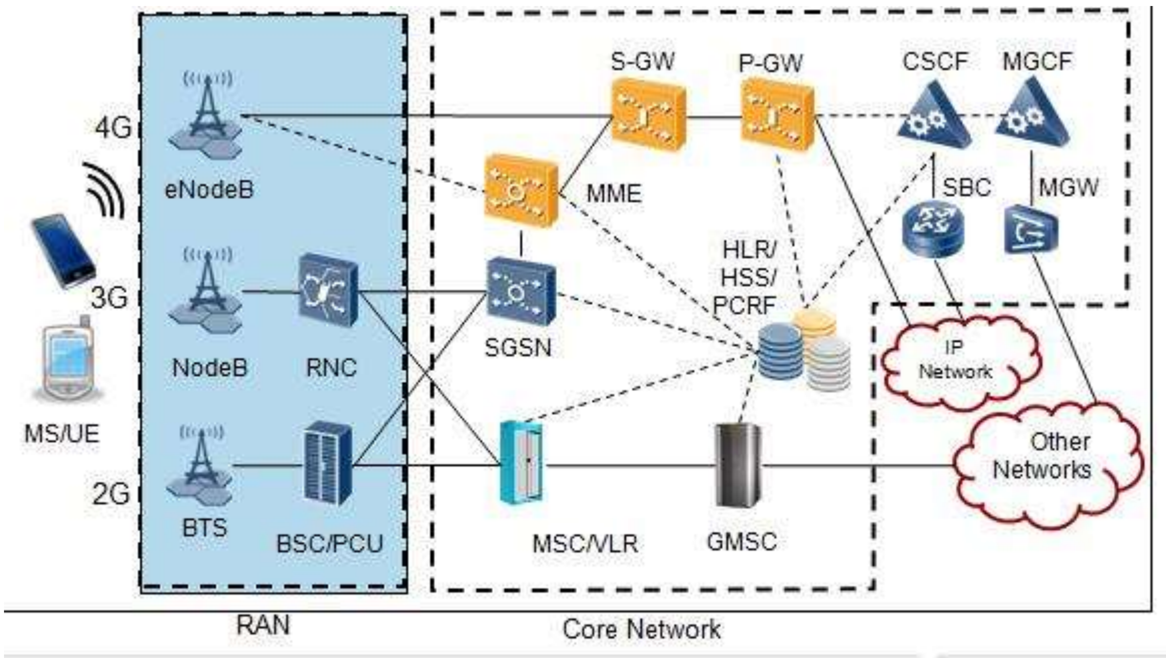


Figure 1.1 A Typical Telecommunication Network. Source: (Qing, 2017)

For illustrative purposes, the author has shown from fieldwork, some of the equipment that make up the telecommunications network as depicted in plates I to III in Appendix B.

1.1.2 Brief History and Evolution of Mobile Telecommunication in Nigeria

In the year, 2001, Global System for Mobile Communication (GSM) was rolled out in Nigeria with the deployment of the second-generation technology (2G). It came with both voice and SMS services and the data rate was 9.6kbps. The 3G was first launched in Japan using the Wide Code Division Multiple Access (WCDMA) technology in 2001 (Meraj and Kumar, 2015). 3G was deployed in Nigeria much later. In 2008, 4G was launched with the

Long-Term Evolution (LTE) technology and it has a data rate of 150Mbps. Figure 1.2 illustrates the evolution of mobile technology from 2G in 1980 to 4G in 2008.

As the technology evolved, the data rate increased as well. There was an improvement of the data transmission rate of 117.2Kbps of the General Packet Radio Service (GPRS) to 384Kbps of the Enhanced Data rates for GSM Evolution (EDGE). With further evolution as depicted in Figure 1.2, a transmission data rate of 2Mbps of the Universal Mobile Telecommunication System (UMTS) was obtained in 3G technology. The 4G LTE has a data transmission rate of 150Mbps.

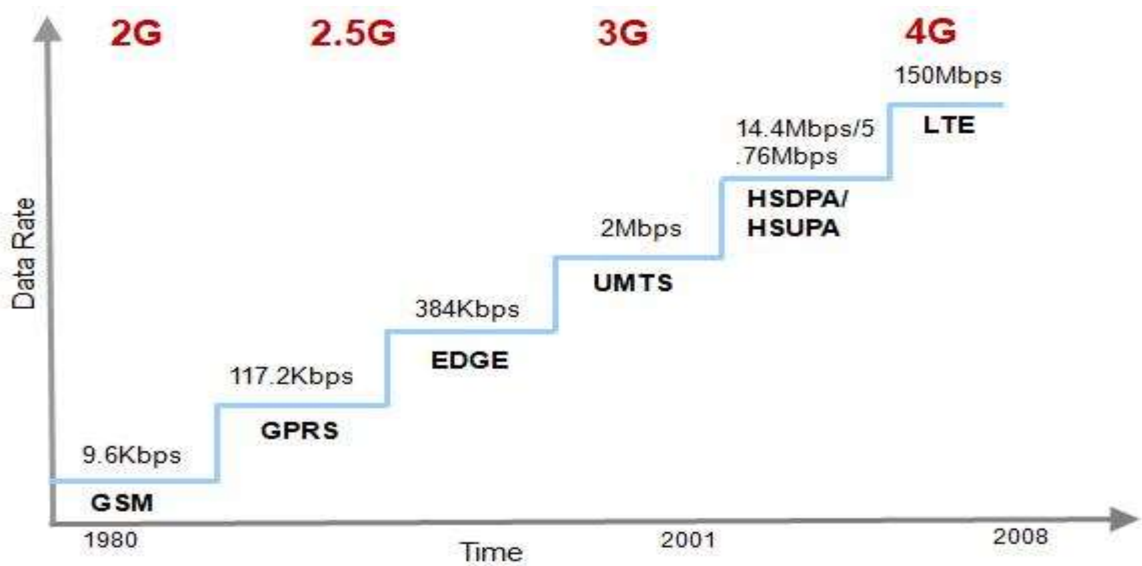


Figure 1.2 Evolution of Mobile Network Technology. Source: (Qing, 2017)

1.2 Statement of the Research Problem

Telecommunication hardware (HW) and software(SW) are specifically intended to support the Availability requirement of five or six ‘nines’ (Hilt, 2019; Netes, 2018; Akinsanmi and Adebusuyi, 2016; Thulin, 2004). Five or six ‘nines’ mean 99.999% or 99.9999% respectively. Availability of 99.999% for instance corresponds to an outage duration of 5 minutes, 15 seconds in a year. MNOs in Nigeria and most Sub-Saharan Africa countries are struggling to attain BTS Availability of two 'nines' (99.0%). This is very far from expectations. There is a need for MNOs to be equipped with the means to proactively schedule predictive maintenance (PdM) or planned preventive maintenance (PPM) on NEs and BTS. This will mitigate or appreciably reduce outage probability and improve BTS availability.

1.3 Aim and Objectives of the Study

This research aims at predicting the availability of Base Stations for Telecommunication Networks in Minna.

To achieve the aim stated above, the following are the objectives:

I. To acquire and process the Base Station Availability data of four MNOs: MNO W, MNO X, MNO Y and MNO Z in Minna from 1st January 2018 to 26th September 2020.

II. To use the Data in Objective I to develop Autoregressive Integrated Moving Average-Based (ARIMA-Based) Predictive Models of Base Station Availability for the stated MNOs.

III. To evaluate the performance of the Models using MAE and MAPE, and to compare with the Long Short-Term Memory (LSTM) Model.

1.4 Scope of the Study

This study is limited to focus on BTS Availability for telecommunications network operators in Minna. The population of Availability data (measured in percentage) for the research is a thousand each for four MNOs named as MNO W, MNO X, MNO Y and MNO Z in Minna metropolis spanning the period from 1st January 2018 – 26th September 2020. The work discusses Availability as a Time Series (TS) data and as such, Time Series Analysis (TSA) shall be discussed. Also, in the study, the concepts of the following: Availability, BTS, predictive models like the Autoregressive Integrated Moving Average (ARIMA) and the LSTM are discussed. Factors affecting BTS Availability are also highlighted in this study.

1.5 Justification for the Study

MNOs use BS Availability for attracting customers. It is used as a measure of quality (Mahdi *et al*. 2018; Thulin, 2004). Hilt *et al*. (2016) studied the Availability prediction of telecommunication application servers on cloud. The result showed that legacy telecommunication availability could be attained on cloud-based core networks. This work could have been enhanced if real time TS data was used. Both Hilt *et al*. (2016) and Mahdi *et al*. (2018) used Reliability Block Diagram (RBD) method. Mahdi *et al*. (2018) focussed on Availability measurement and the work lacks predictive ability. Fan *et al*. (2016) improved the base station Availability by improving the maintenance of the back-up battery group. This research however did not take a complete view of other factors that could impair BS Availability. MNOs in Minna and most Sub-Saharan Africa do not provide the required standard of Availability of 99.999% (5 minutes, 15 seconds of downtime within a year) to their subscribers. In the light of the drawbacks and merits of the previous research, this work leveraged on the amenability of the ARIMA model on TS data to design a predictive model for PdM. This predictive model will be utilised by the policy makers of the MNO to schedule PPM and PdM for a proactive maintenance. This improves the Base Station Availability and reduces operation expenses.

1.6 Thesis Outline

This thesis consists of five chapters as follows: Chapter one is the Introduction. Chapter Two gives the Literature Review and introduces various theoretical background. Chapter Three is the Research Methodology. Chapter Four presents the Results and Discussions. Chapter Five contains the Conclusion, Recommendation and Contributions of the Research to Knowledge.

CHAPTER TWO

2.0 LITERATURE REVIEW

2.1 Overview of Literature review

Previous research work on base station availability has centred on optimum resource utilization, power-saving and improved radio network planning and maintenance. It can be categorised based on the technique used, measurement, prediction and application areas.

Thulin (2004) considered the Network Availability (NA) of Synchronous Digital Hierarchy (SDH) link in their company, Song Network AB. A background investigation of NA was done and subsequently, a method for the measurement of NA was designed for use in their SDH network. Availability parameters were got from ITU-T Standard G.826 and a log file from the network surveillance system was stored in a database. A parser program was employed to acquire, investigate and present the data in the database and create NA reports. The study, however, did not consider the prediction of NA for the network.

Netes (2018) worked on the End-to-End Availability of Cloud Services. Service Availability was explained and quantitative estimation of end-to-end availability of cloud services was carried out for distinctive conditions bearing in mind all components of the service structure. It shows that realising high availability needs redundancy for such components as network connections and data centres. General reflection was detailed for the cloud service “Desktop as a Service” (DaaS). End-to-End availability for a single Data centre and two Data centres were each calculated; the use of only one data centre is not desirable keeping in mind to ensure survivability, that is, the ability to continue to function during and after a natural or man-made disorder. All these led to the idea of redundancy for data centres with geographical separation of their locations (geo-redundancy). However, this will involve an additional cost.

Netes (2018) and Hilt *et al*. (2016) researched availability relating to cloud services. The latter, focussed on ‘cloud-based Mobile Switching and Telecommunication Application Servers (MSS and TAS). The paper demonstrates the possible redundancy principles and a simulation method to forecast availability for mobile communication NEs on the cloud. Nokia AS on ‘Telco-cloud’ is offered, which combines several redundancy principles such as full protection (2N), standby and load sharing. The study showed that the general availability of telecommunication networks depends on NE, interconnection and NW level redundancy methods. Simulated results predict that ‘Telco-grade’ availability can be attained on cloud-based core network elements. It was also shown that full protection and load sharing are more efficient than other methods of redundancy especially with increasing number of parallel nodes or units.

Fan *et al.* (2016) looked at boosting service availability of base stations by consideration of profiling of power backup to sustain both the base stations and the backhaul infrastructure used for the aggregation of the BSs in the network. The paper conducted a systematic analysis on a real-world dataset collected from the battery groups installed on the base stations of China Mobile Limited company and proposed an event-driven battery profiling approach that precisely extracts the features causing the degradation of the battery group. Dataset was collated and analysed by cloud computing platform with Hadoop 1.2.1 and Hive 1.2.1. Based on the data analysis, an event-driven battery profiling approach that can be used to schedule preventive maintenance and replacement of battery was obtained, to reduce power outages and guarantee the high service availability for the cellular networks which is usually the main target of MNOs. Although their work showed an 18.09% increase in service availability, the consideration is only limited to the battery group.

The Network Function Virtualization (NFV) initiative aimed at allowing network operators to deploy a set of services in a reliable, secure and fast way has been considered as it is a milestone towards the deployment of Telco-cloud Di Mauro *et al.* (2017) and Hilt (2019). Considering that one of the main benefits of an NFV infrastructure is its high availability, Hilt (2019) focused on the analysis of the Virtualized Infrastructure Manager (VIM), a core element that is implemented through the OpenStack platform and is aimed at managing the whole NFV architecture. The system availability was evaluated on the VIM by performing both a steady-state and transient analysis of Stochastic Reward Nets (SRNs) to obtain the best system configuration with the "five nines" availability requirement.

Ibrahimovic and Bajgoric (2016) presented a modified Bayesian Belief Network model for predicting information system availability. The work adapted the model built by (Franke *et al.*, 2012). The model made use of parameters by using the probability elicitation process. The model validation was performed using the Monte-Carlo simulation. Work by other authors that used Bayesian Network is presented by (Beuzen *et al.*, 2018; Franke *et al.*, 2012; Usman *et al.*, 2017).

Nencioni *et al.* (2017) investigated the Availability of Software Defined Networking, SDN. The impact of the deployment of SDN controllers indicated that using a three or four-homed controller would not improve the availability performance. According to the report, using a two-homed SDN controller gave an availability value that is comparable to that of the legacy network. One SDN controller is not sufficient to assure an acceptable availability. Since network operators' target is to make sure that availability does not decrease as compared to the traditional network, the two SDN controllers would be the best choice.

Säe and Lempiäinen (2016) considered saving energy to be used by eNodeB sites in a critical disaster scenario. A portion of the entire site is intended to be provided with backup coverage during a period of a temporary electricity power cut. A kind of sleep mode concept in cellular networks is the main approach utilized in the paper for the functionality of eNodeB sites operating only with backup power (that is, during a disturbance scenario). Commercially available radio network planning software called Integrated Communication System (ICS) Designer was used in the implementation of simulation on a link-level basis. Though a significant availability was achieved during the backup period, the study however does not include algorithms in how to select the usable eNodeB sites.

Bikcora *et al*. (2016) investigated day-ahead probabilistic forecasting of the charging rate and the availability of plug-in electric vehicles at a charging station. Two forecast scenarios were used in the paper. Since the availability forecasting was considered as a binomial problem, forecasting was done to address the availability as such, while the charging rate was forecasted using an ordered logistic model having categorized the feasible range of values. The result of their findings indicated that the predictive model was essential for the optimum performance of the charging station at some charging points, whereas, it was not so at some others. The study may not be applicable at some other locations as it was localized in Netherland. Other places with different environmental factors may experience different outcomes.

Garrido *et al*. (2016) presented a context-aware architecture that provides an adaptable and reusable avenue for the availability of cloud-based mobile services using an election algorithm. A service replication scheme together with a self-configuration method was adopted for the activation/hibernation of the replicas of the service depending on user context information from the mobile system.

In the research, ‘Survivability Analysis of GSM Network Systems’, Mahdi *et al*. (2018) posit that Availability is the major measure of quality. This is in agreement with the claim of (Thulin, 2004) that MNOs use availability statistics in assessing the overall quality of the network. Mahdi *et al*. (2018) used Reliability Block Diagram (RBD) to model the system and their work showed that the components with lower Availability have higher Mean Time To Repair (MTTR) while the ones with higher Availability have lower MTTR which agrees with the theoretical definition of Availability. MTTR is the average time taken to repair a repairable componet or system after its failure.

Mahdy *et al.* (2020) used a set of datasets for Statistical, Machine Learning and Deep Learning models in the prediction of traffic load on Base Stations. In their work, they clustered base stations that possess similar load behaviour before using clustered data for the predictive model. The results showed that the Root Mean Square Error (RMSE) for the clustered scenario performed better than in the un-clustered scenario. Table 2.1 gives a summary of the reviewed literature for this research.

Table 2.1 Summary of Reviewed Literature

S/N	Author(s)	Title	Category	Strength	Weakness
1	Dicholkar and Dongree (2018)	Cost-Effective Adaptive Modulation for Microwave Link Availability Improvement in Plain, Hilly Terrain, Water Bodies	Availability improvement	Adaptive modulation was used to achieve desired availability at a low cost.	Sensitivity to measurement error and delay.
2	Netes (2018)	End-to-End Availability of Cloud Services.	Availability in Cloud computing	Gave a quantitative estimation of the end-to-end availability of cloud services	High cost of redundancy for higher availability.
3	Thulin (2004)	Measuring Availability in Telecommunications Networks.	Availability Measurement	Developed a Method for measuring Network Availability.	Lacks predictive ability.
4	Hilt *et al*. (2016)	Availability Prediction of Telecommunication Application Servers Deployed on Cloud	Availability Prediction in cloud computing	Simulated outcomes predict that 'Telco-grade' availability can be attained on cloud-based core network elements.	Technical complexity issues
5	Fan *et al*. (2016)	Boosting Service Availability for Base Stations of Cellular Networks by Event-driven Battery Profiling.	Availability improvement through PdM on batteries	An improvement in service availability for cellular networks	The model may not adapt in other locations.

6	Nencioni *et al.* (2017)	Impact of SDN Controller Deployment on Network Availability	Availability Improvement in SDN	Using a two-homed SDN controller gave an availability value that is comparable to that of legacy networks.	High technical complexity.
7	Di Mauro *et al.* (2017)	Availability Evaluation of the Virtualized Infrastructure Manager in Network Function Virtualization Environments	Availability in NFV	Their results showed that the Availability value for the redundancy level of 4 gave a value of 5 ‘’nines’ which is better than for the redundancy levels of 3 and 2.	The work did not give a vital performance metric like response delay which would have been used to assess the SDN Controller.
8	Säe and Lempiäinen (2016)	Maintaining Mobile Network Coverage Availability in Disturbance Scenario	Application to Availability	Significant availability was achieved during the backup period.	No algorithms for the selection of the usable eNodeB sites
9	Ibrahimovic and Bajgoric, (2016)	Modelling Information System Availability by Using Bayesian Belief Network Approach	Availability Prediction using BN	Offered a modified Bayesian Belief Network model for predicting information system availability.	The hypothesis built into this model is the independence of the variables that enabled the application of the Leaky Noisy-OR approach.
10	Bikcora *et al*. (2016)	Prediction of Availability and Charging Rate at Charging Stations for Electric Vehicles.	Availability Prediction for PdM	Findings indicated that the predictive model was essential for the optimum performance of the charging station at	The study may not be applicable at some other location as it was localized, other places with different

				some charging points, whereas, it was not so at some others.	environmental factors may experience a different outcome.
11	Bakar and Rosbi (2017)	Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in High Volatility Environment: A New Insight of Bitcoin Transaction	Technique Utilized	The error analysis between forecasting value and actual data was performed and the MAPE of the forecasting is 5.36%.	Cryptocurrency exchanges are often very volatile. Full Reliance on the model may fail.
12	Özs (2020)	Monitoring unstable slopes in an open-pit lignite mine using ARIMA.	Technique Utilized	Prediction results from the ARIMA method were compared with results from regression methods and were shown to be more successful.	The assumption of mathematical modelling may not agree with real-life situations, this is affected by various factors such as floods and landslides.
13	Mahdy *et al*. (2020)	A Clustering-Driven Approach to Predict the Traffic Load of Mobile Networks for the Analysis of Base Stations Deployment.	Technique Utilized	Results showed that the RMSE for the clustered scenario performed better than in the un-clustered scenario.	A very limited dataset was used in the work.

14	Mahdi *et al*. (2018)	Survivability Analysis of GSM Network Systems.	Availability Measurement	Components with lower Availability have higher MTTR while the ones with higher Availability have lower MTTR.	Lacks predictive ability
15	Beuzen *et al*. (2018)	Bayesian Networks in Coastal Engineering: Distinguishing Descriptive and Predictive Applications	Application	Results show that two BNs: one with high predictive skill and one with an optimized descriptive skill could be developed from the given dataset	There is a challenge of selecting an appropriate input to include in the model
16	Carvalho *et al*. (2019)	A systematic literature review of machine learning methods applied to predictive maintenance		A systematic literature review of ML methods applied to PdM	Work does not have a quantitative finding
17	Gandomi and Roke (2015)	Valuation of artificial neural network and genetic programming as predictive tools	Technique/Application	A case study in punching shear prediction of RC slabs is modelled using a hybrid ANN and GP. The result shows the relevance of parametric studies in model acceptance criteria	Technical complexity
18	Mahmood and Munir, (2020)	Predictive and Preventive Maintenance using IoT and Big Data in the Telecom Sector	Technique/Application	The paper gave a qualitative approach in enabling PdM for RAN	The framework does not have a quantitative evaluation.

2.2 Theoretical background

To understand more clearly the reviewed work in this research, this section discusses the theoretical principles of some relevant terms in the researched papers.

2.2.1 Availability Parameters in ITU-T Standard G.826

The International Telecommunication Union (ITU) defined a set of standards providing procedures, objectives and limits for links in Synchronous Digital Hierarchy (SDH) Networks. The standard relating to performance evaluation and measurements of Availability is G.826 (ITU-T G.826, 2002; Akinsanmi and Adebusuyi, 2016; Säe and Lempiäinen, 2016; Thulin, 2004). The Unavailable Seconds (UAS) which is indicative of downtime is deduced from the G.826 standard as shown in figure 2.1.

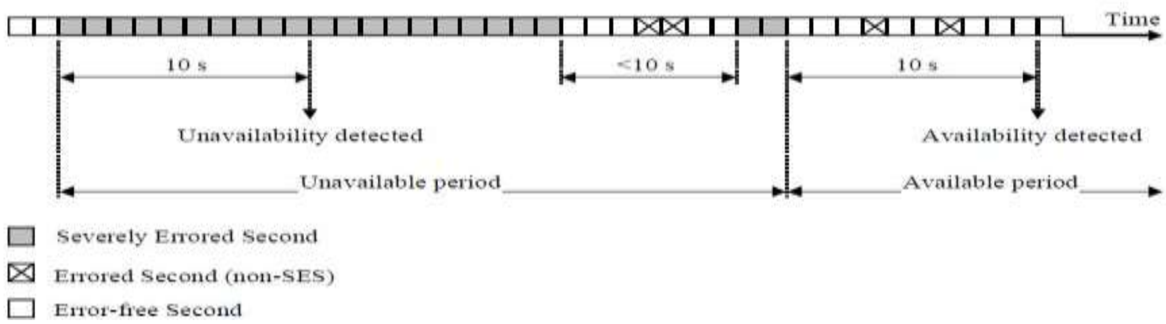


Figure 2.1 Determination of UAS. (Source**:** ITU-T G.826 Standard.)

A period of unavailable time begins at the onset of ten consecutive SES events (these ten seconds is part of the unavailable time). A new period of available time commences at the onset of ten consecutive non-SES events. These ten seconds is part of the available time.

$$= \frac{(\quad - \quad) * 100}{\quad} \qquad (2.1)$$

Akinsanmi & Adebuyi (2016)

Where A is the Availability.

a. Definition of Availability

Availability is defined by the International Standard Organisation (ISO) as the capability of a useful unit to be in a state to accomplish a vital role under specified circumstances at a given instant of time or over a given time interval, assuming that the required external resources are provided (ITU-T E.802, 2007; ISO/IEC, 2015).

Service Availability is defined in some ITU-T Recommendations, like (X.140) as the ratio of the total time during which acceptable or tolerable service is, or could be, provided to the total observation period. According to E.860, it is the percentage of time during which the contracted service is functioning at the respective service access points. In this case, the access point is the BTS.

Unavailability,

$$= 1 - \qquad (2.2)$$

Down Time (D_T),

$$=(1-)* \qquad (2.3)$$

b. BTS Availability

In line with the definition of service availability, BTS Availability is the percentage of time duration during which the intended service expectation is achieved. It is presented in nines. That is five or six nines (99.999, 99.9999 respectively) according to standard. However, these values are hardly met in practice by MNOs.

c. Availability Targets in Telecommunication Networks

Availability values are usually presented in terms of nines. For many telecommunication networks and equipment, BTS inclusive, there is a strict standard of five or six nines. Telecommunications HW (and SW) is specially designed to support these very strict requirements (Hilt, 2019). Table 2.2 illustrates availability values.

Table 2.2 Availability and Corresponding Downtime

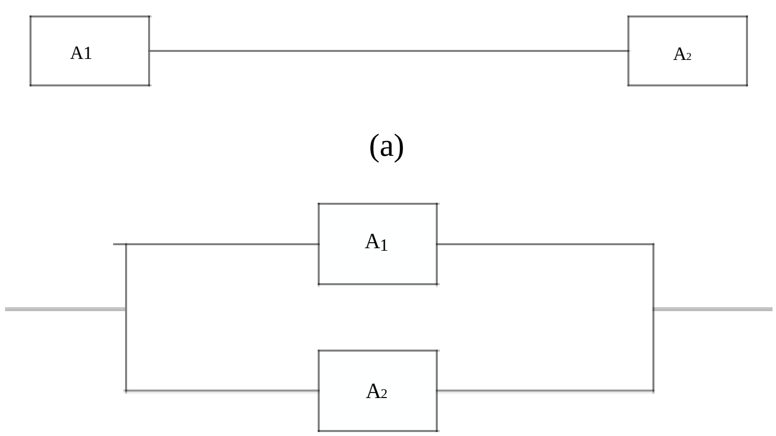
Availability	Downtime per year
99.9999%	32s
99.999%	5min 15s
99.99%	52min 32s
99.9%	8hr 46min
99%	3days 15hr 40min

Source: (Hilt *et al*. 2016; Netes 2018; Akinsanmi and Adebusuyi 2016; Thulin 2004)

The values in Table 2.2 are computed from equation 2.3.

d. Availability and Redundancy

If a system is made up of useful units in series, the failure of any of the functional units leads to a total system failure (Netes 2018; Thulin 2004). A series system does not have redundancy. It is represented in Figure 2.2a.



(b)

Figure 2.2 Availability of Systems (a) Series (b) parallel

The Availability (A) for the system in Figure 2.2 (a) is:

$$=_{1}*_{2} \tag{2.4}$$

Generally, the availability for a system in series can be computed as:

$$A = \prod_{=1} \tag{2.5}$$

For all $= 1,2,3, \ldots , .$

Where A_i is the Availability of the *i*th element and n is the number of elements that make up the system. The overall Availability (A) for this kind of system is less than the least Availability of an individual element of the system.

The overall availability of a system that has its elements in series reduces and is even less than that of the least element. The provision of redundancy increases Availability (Ahmed *et al*. 2017; Hilt *et al*. 2016; Krajnović 2017; Netes 2018; Nagy et al., 2016). The Availability for a system having redundancy as in Figure 2.2b is more guaranteed since the system will remain available so long as either of the element is available, and the system only fails when all the elements fail. The Availability is computed as:

$$=_{1}+_{2}-_{1}*_{2} \tag{2.6}$$

$$A=1-\prod_{=1}(1-) \tag{2.7}$$

2.2.2 Factors that Affect BTS Availability

The Mean Time Between Failure (MTBF) of the components and subsystems of the BTS and NEs in the network affect the overall Availability. A high MTBF and low MTTR improve Availability. It is pertinent to note that the MTTR in case of an outage is a major barrier to Availability. According to ‘the Availability Digest’, the focus has been on the improvement in system MTBF which has boosted the Reliability and Availability of both hardware and software used in the systems. However, systems do not have infinite MTBF. Failures do occur due to application software bugs or even as a result of operator error. This implies that further improvement on MTBF is not likely going to yield a better Availability. The Availability barrier becomes the MTTR. The system may fail but it should be fixed fast. When the duration of time for fixing or restoration is so fast (that is, MTTR tending to zero), users may not be aware of the failure. So, it is much beneficial to consider ways of reducing MTTR to break/reduce the barrier to Availability. Some of the factors that affect BTS Availability include failure or degradation in components (BTS and NE components), software, microwave link, transmission cable, infrastructure, accessibility and force majeure.

2.2.3 Predictive Analytics

Predictive analytics have been used in diverse fields such as Statistics, Marketing, Natural Sciences and Engineering (Mahmood and Munir 2020; Almeida 2002; Beuzen et al. 2018; Bikcora *et al*. 2016; Carvalho *et al*. 2019; Chodorek 2005; Fan *et al.* 2016; Gandomi and Roke 2015; Boulos and Niraula 2016; Odom, 1990; Ercsey-Ravasz *et al.*, 2013). Prediction is useful in the improvement of operations. Companies and other establishments like MNOs use it for adequate planning and management of resources for an optimal realization of set goals. For this work, the emphasis is on BTS Availability.

Predictive modelling is a process of using known data and statistics of a variable to forecast or predict the future outcome of the variable with data models (Daren and Paul, 2019). The Predictive Modelling tool used in this work is based on ARIMA (Autoregressive Integrated Moving Average). The Long Short-Term Memory (LSTM) algorithm, a type of Recurrent Neural Network (RNN) was used for prediction and performance comparison.

Other predictive modelling tools in the researched papers are the ANN (Artificial Neural Networks), Python, R, Apache Hadoop, Windchill and Bayesian Networks.

a. The ARIMA Model

This is a very popular linear model that is very amenable to modelling with Time Series (TS) data. ARIMA model is very flexible and could be used on different TS. It has good statistical properties and can be used for exponential smoothing models Zhang (2003). A limitation of the ARIMA model is the assumption that the predicted value of a variable is a linear function of the previous value of the variable. This model is discussed in more detail in chapter three of this work.

b. The LSTM

LSTM is a type of RNN and was first presented by Sepp Hochreiter and Juergen Schmidhuber as a means of resolving the vanishing gradient trouble in training some networks (Hochreiter and Schmidhuber 1997). This issue is because of an enormous input being constrained into a little input space. When this happens, there is a loss of memory. Most traditional RNNs are faced with this problem of vanishing gradient as the input become large.

An RNN takes input information each in turn and keeps up with data from past inputs, it takes data from past input in making new calculations. LSTMs can choose similitudes in successions of information like TS data. LSTMs are utilized for prediction or forecast Laptev *et al*. (2017). LSTMs utilize the condition of the last neuron from the last time value as a format to make an output.

A significant element of the LSTM is a memory cell known as a memory block which can safeguard its state throughout an extensive period. According to (Greff *et al*. 2017), an LSTM cell contains gates that control the stream of information into and out of the cell. LSTM has the following gates: Forget gate, Input and Output gates as shown in Figure 2.3.

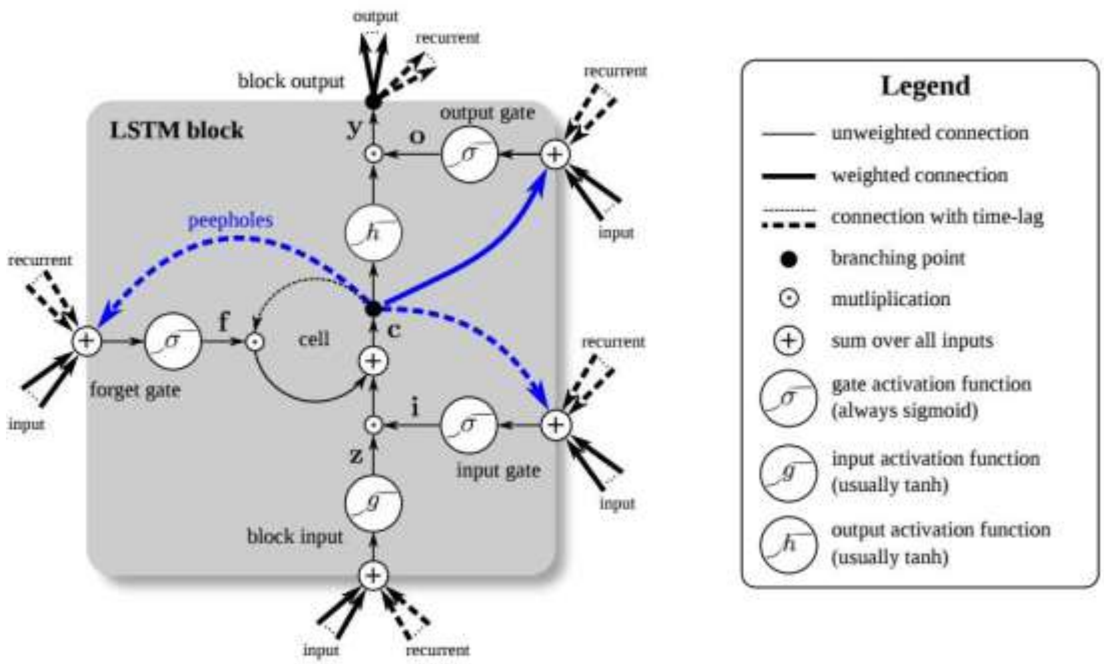


Figure 2.3 Long Short-Term Memory Block. Source: Greff et al. (2017).

2.2.4 Time Series Analysis

A time series is a set of time-ordered observations of a process where the intervals between observations remain fixed as in daily, weekly, monthly or yearly intervals over a period (Daren & Paul, 2019; Jebb et al., 2015). A good example is the daily BTS Availability report of a Mobile Network Operator taken for several months or the weekly report for a few years or the monthly report for several years. According to Jebb *et al.* (2015), the length of time series do vary, but should at least be more than twenty observations long. Some models do require more than fifty observations. It is generally required that more data that capture the phenomena of interest is used for optimum model accuracy. The most important reason for carrying out TS analysis is for forecasting or predicting future values of the series.

a. Components of a Time Series

The change in the pattern of a time series is characterized by four components which are: trend, seasonality, cyclical and irregularity (Shmueli and Lichtendahl, 2016)(Jebb *et al.*, 2015).

- **i.** The trend shows the increase or decrease in the series over a long period. An example is the population growth over the years.
- **ii.** Seasonality indicates a regular pattern of fluctuation in the TS. It is a short-term variation occurring as an impact of the seasonal effect. An instance is the effect of fog and stormy rainfall on network availability during the harmattan and rainy season respectively.
- **iii.** The cyclical component in TS is conceptually like the seasonality component. It is a pattern of fluctuations that repeat at irregular periods. The patterns are not caused by any fixed period.
- **iv.** The irregularity (Random) component in TS refers to variations that occur at unpredictable factors and do not have a particular repetitive pattern. This component represents the noise and can be termed as the error component in most statistical models.

b. Decomposition of Time Series

For the TS depicted in Figure 2.4, the trend and cyclical components are treated as the same. The additive decomposition of the TS is represented in equation 2.8.

$$= + + \tag{2.8}$$

The multiplicative decomposition is given as:

$$= * * \tag{2.9}$$

Where T_t, S_t and E_t are the trend, cyclical and the random components respectively. Equation 2.8 is used when trend, cyclical and seasonal components are constant throughout the series, while equation 2.10 is used otherwise (Jebb *et al*. 2015).

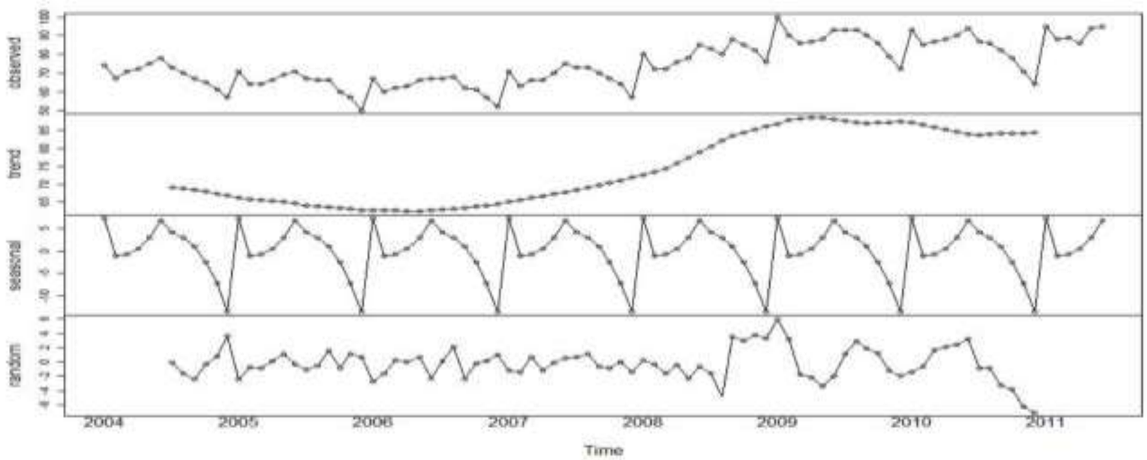


Figure 2.4 Time Series Decomposed into its Trend, Seasonal, and Irregular Components.

Source: Jebb *et al.* (2015)

c. Stationarity for a Time Series and Transformation of Non-Stationary Series

It is a necessity that for predictive purposes, the time series should be stationary. A stationary TS has a constant mean and variance or autocorrelation in the TS data. When the series is stationary, any section of the data can be used in the model irrespective of when the variable was observed (Jebb *et al.* 2015). A stationary series is very good for prediction and as such, a non-stationary series must be transformed to a stationary one before modelling.

For ARIMA and some other forecasting models, a non-stationary series can be transformed into a stationary series by a process called differencing (Hyndman and Athanasopoulos 2014). Differencing removes the trend in the series. For custom ARIMA models, the time series to be modelled are made stationary by transforming a non-stationary series into a stationary one by a ‘difference’ transformation.

2.3 Research Gap

Hilt *et al*. (2016) predicted that the legacy telecommunications Availability can be obtained in the cloud scenario. The work would have been more authentic if real field TS Availability data was used. Like Hilt *et al*. (2016), Mahdi *et al.* (2018) used RBD approach in the analysis of NA without prediction of the Availability. Fan *et al*. (2016) focussed on the improvement of Base Station Availability by considering the optimum maintenance of the Base Station backup battery groups. This work does not have a wholistic view of Base Station Availability. MNOs in Minna and most Sub-Saharan Africa have a challenge of attaining a Base Station Availability value of 99.999%, to improve the Base Station Availability, the TS data of the MNOs must be used for predictive modelling. These predictive models will enable the policy makers of the MNOs to proactively Schedule Maintenance. The ARIMA model is very suitable for TS modelling and has been successfully used by (Bakar and Rosbi, 2017; Özs, H., 2020).

CHAPTER THREE

3.0 RESEARCH METHODOLOGY

3.1 Data Acquisition and Processing

The technical department of the various Mobile Network Operators: MNO W, MNO X, MNO Y and MNO Z were approached to acquire the base station daily availability data. This data is used in developing a predictive model that will assist in the proactive planning of preventive maintenance in a bid to reduce BTS downtime and subsequent improvement of BTS Availability. Consequently, the historic BTS availability report from 1st January 2018 to 26th September 2020 was used as the data processed and used for building the predictive model. The report for each base station was retrieved from the Network Operation Centre (NOC) archive. Table 3.1 shows the Base Station site count for the MNOs in Minna.

Table 3.1 Base Station Site Count for the MNOs under Consideration

MNO W	MNO X	MNO Y	MNO Z
29	43	38	65

Source. Author's Field Work

3.1.1 Data Source

The availability report was retrieved from the Network Operation Centre's (NOC's) archive and from the cloud surveillance application (Mateline) which contains the Radio Access Network (RAN) report. The daily BTS Availability is the percentage uptime within a day which is obtainable from equation 2.1. For this research, a daily BTS Availability for a period of one thousand days was taken from1st January 2018 – 26th September 2020. From the BTS Availability data, other useful technical information is an idea of the corresponding Unavailability (U) and Downtime (D_T) as in equations 2.3 and 2.4 respectively. Information about the Uptime can also be gotten from the report. Both Uptime and Downtime are measured in a unit of time. Figure 3.1 illustrates the network of base stations for one of the Mobile Network Operators (MNO X). Plates IV to VI in Appendix B indicate OptiX for terminating optic fibre; the tower and the radio rack respectively.

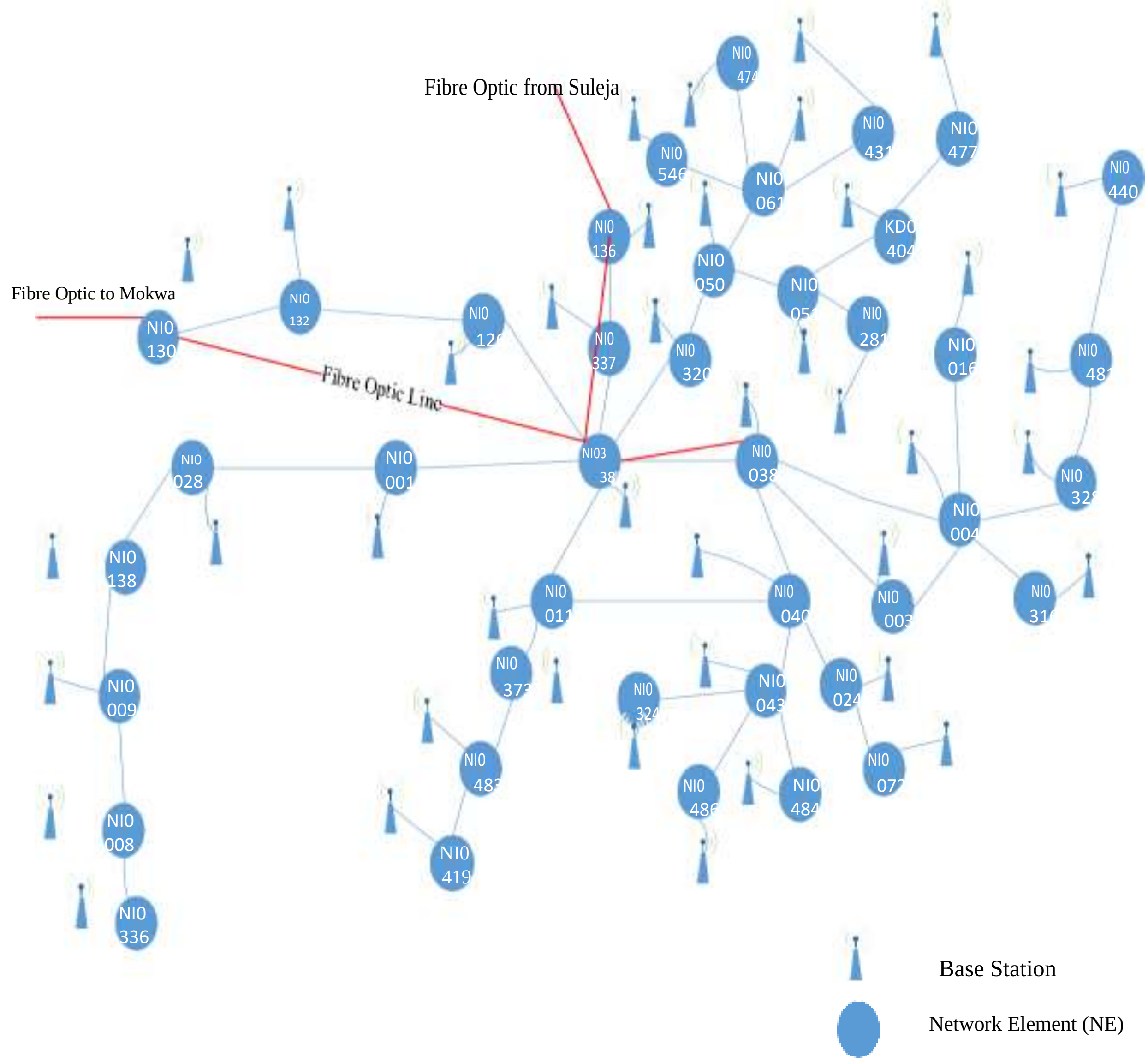


Figure 3.1 Schematic of Base Stations and their Serving Network Elements for MNO X.

The Base Stations are backhauled through the SDH network elements to the backbone and the BSC.

3.2 Data Preparation

This section discusses the preparation of the raw data got from the field. The Availability report was gotten for each Base Station and the mean daily Availability for each MNO was computed.

$$= \frac{1}{\ } \Sigma \qquad (\ ) \tag{3.1}$$

Where N is the number of base stations, A_i is the ith Availability for each base station for $= 1,2,3, \ldots ,$

3.2.1 Expectation-Maximization (EM)

For a good model to be developed, one of the most crucial requirements is data devoid of errors. Errors occur in the form of missing data and outliers. Outliers are data records that do not fit well in the given set of data and are excluded from the list. Missing data are estimated by taking the arithmetic mean or median of the neighbouring data points. When the volume of data is much, it becomes very difficult to manually check for missing data values and outliers. In this case, it becomes necessary to implement an automated process for data validation. The Data Preparation add-on module in the IBM SPSS Statistical tool allows the identification of unusual cases and invalid cases, variables, and data values in the active dataset, and prepares data by a process of finding maximum likelihood estimation for modelling.

3.3 Modelling Using the TS Data

This thesis work is aimed at predicting the Availability of a Base Station for Mobile Network Operators (MNO) in Minna. In the preceding sections and subsections of this chapter, it was shown how the BTS Availability data of the four MNOs were acquired from their base stations. Each MNO number of base station sites is indicated in Table 3.1. The data is observed to have the nature of Time Series (TS) data. The data was prepared so that it will be adequate for producing a good predictive model. Figure 3.2 is the flowchart for the ARIMA-Based model.

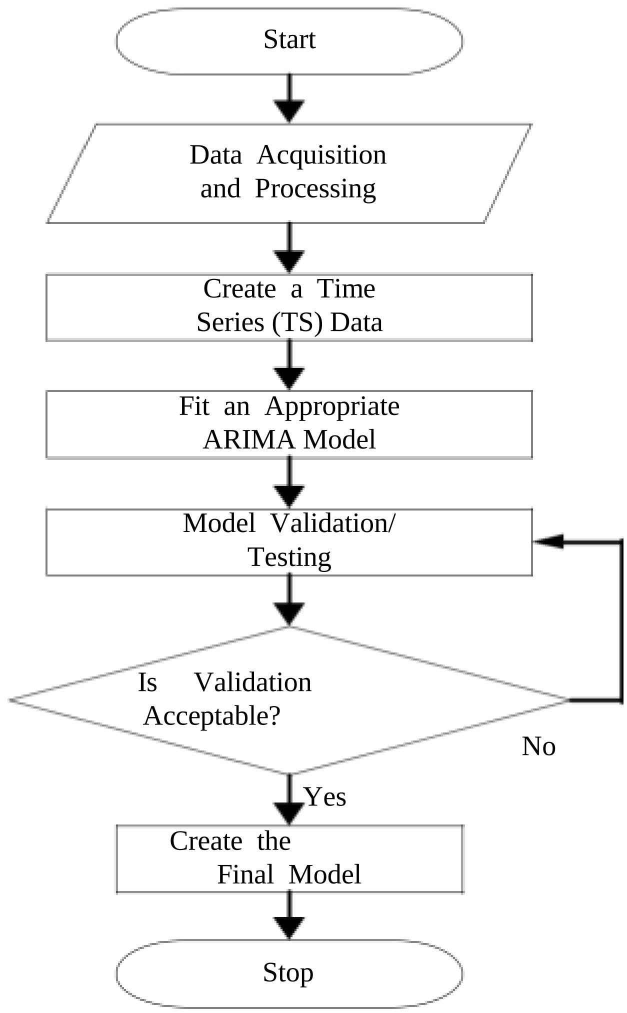


Figure 3.2 Flowchart of the ARIMA-Based Predictive Model

3.3.1 The ARIMA Model

The ARIMA model is very suitable for Time Series analytics. It is considered in more detail here. The BTS Availability A_t is a sequence of the form;

$$= {}_1, {}_2, {}_3, {}_4, \ldots, \quad (3.2)$$

The Autoregressive (AR) (p) Integrated (I) (d) Moving Average (MA) (q) is a hybrid of models that exploits the functions of the Regressive, the differencing factor (Integrating) and the Moving Average in the models. The **p, d and q** are their respective orders.

For the Autoregressive (AR) model, the output variable is linearly dependent on the previous values and a set of stochastic terms. The AR component of the Base Station Availability having an order (p) is generally written as AR (p) and is mathematically defined as:

$$= {}_1 {}_{-1} + {}_2 {}_{-2} + \cdots + {}_{-} + \varepsilon \qquad (3.3)$$

Equation 3.3 can be written as:

$$= \sum_{=1} {}_{-} + \varepsilon \qquad (3.4)$$

Where ,…, are the model's parameters and ε is the random error. Similarly, the output variable ( ) for the MA depends linearly on the present value and past values of stochastic terms. For MA of order (q) written as MA(q), ( ) is defined as:

$$= + {}_1\varepsilon_{-1} + {}_2\varepsilon_{-2} + \cdots + \varepsilon_{-} + \varepsilon \qquad (3.5)$$

$$= + \sum_{=1} \varepsilon_{-} + \varepsilon \qquad (3.6)$$

Where is the mean of the series, ,…, q are the parameters for the MA model, ε and ε_{-} are the white noise errors.

From equations 3.4 and 3.6 give equation 3.7.

$$= + \sum_{=1} {}_{-} + \varepsilon + \sum_{=1} \varepsilon_{-} \qquad (3.7)$$

Equation 3.7 defines the ARMA (p, q) model.

ε_t is assumed to be an independent identically distributed (iid) variable obtained from a normal distribution whose mean is zero. In statistics and probability theory, a group of variables is iid if each random variable has the same probability distribution as the others and are all mutually independent.

$\varepsilon \approx \ (0,\ ^2)$. Mean is zero, $\ ^2$ is the variance.

From equation 3.7,

$$= \sum_{=1} \ _{-} + \varepsilon + \sum_{=1} \ \varepsilon_{-} \qquad (3.8)$$

According to Bakar and Rosbi (2017) and Özs (2020), a backshift or lag operator L operates on an element in a time series to give the previous element. Using the lag operator L,

$$= \ _{-1} \qquad (3.9)$$

Generally, for the lag operator with any arbitrary power, i,

$$= \ _{-} \qquad (3.10)$$

Re-arranging and applying the lag operator on equation 3.8,

$$- \sum_{=1} = \varepsilon + \sum_{=1} \ \varepsilon \qquad (3.11)$$

$$(1 - \sum_{=1} \ ) \ = (1 + \sum_{=1} \ ) \varepsilon \qquad (3.12)$$

Equation 3.12 represents the ARMA model with the lag operator.

For an ARIMA model, the differencing operation is paramount as pointed out earlier in this write-up. The reason is not far-fetched; the differencing operation helps in removing the trend component and is considered subsequently.

The differencing operator is defined in Özs (2020) as:

$$= - {}_{-1} \Rightarrow = (1 -) \text{ . So,}$$

$$=(1-) \tag{3.13}$$

If the polynomial of equation 3.9 has a solution of a single root $(1 - \ )$, then, applying equation 3.13 on equation 3.12 gives:

$$(1-\sum_{=1})(1-) \qquad =(1+\sum_{=1})\varepsilon \tag{3.14}$$

Equation 3.14 represents the ARIMA (p, d, q) model where p, d and q are the respective orders for the AR, Integrating (differencing) and the MA models which are the parameters to be determined and used for the model.

3.3.2 The IBM SPSS Statistics Tool (SPSS)

The IBM SPSS Statistics version 23 is used in creating the predictive model in this thesis. IBM SPSS Statistics is just one of the world's major and most effective statistical software companies. The software package was originally known as SPSS (Statistical Package for the Social Sciences) and was created in the 1960s by three Stanford graduates (Daren & Paul 2019). With the vast usage of the software package by researchers in the field of sciences and Engineering, SPSS was renamed "Statistical Product and Service Solutions." The package was bought by IBM in 1990 and was named IBM SPSS Statistics. The package shall simply be termed SPSS in this work.

3.3.3 Creating Time Series Data using SPSS

As stated earlier in this chapter, the network availability data for the MNOs were processed and stored in Microsoft Excel. The data was imported into the data and variable editor of the Time Series Modeller of the SPSS and the periodicity of the data was set on weekly periodicity.

3.3.4 Fitting an Appropriate ARIMA Model

For an ARIMA model, the parameters p, d and q need to be determined. The standard way to do this is to make a correlation plot of the Time Series data to ascertain the stationarity of the data, and the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) examined in terms of lags.

3.3.5 Model Training, Validation and Testing

In this predictive model and in most models where data is used, the concept of data partitioning is very important. The data is normally partitioned into the **training** and the **validation** periods. The training period is the period from the beginning of observation of the data up to about 60 – 75% of the entire period; while the validation period is the remaining period after the end of the training period up to the end of the entire observation period which may be 25 – 40% of the entire period of observation) (Shmueli & Lichtendahl, 2016).

To forecast the future data, for instance, the future BTS Availability for MNOs, the BTS Availability data is partitioned into three periods: the Training, Validation and Future periods. However, there is no data for the future at the initial instance but it is intended to be forecasted. The available data is thus partitioned into the Training and Validation periods. The model is fitted to the data in the Training period and the model forecasts data into the

Validation period, and assumed that there is no data for the Validation period. Two sets of data in the Validation period are the actual data, (A_t) and the forecasted data (Ft) using only the Training data. The essence of this is to ascertain how well the model can predict the data which it did not see. In essence, the performance of the model is measured or assessed in the Validation period.

The data of both the Training and Validation periods are recombined and the model is rerun on the entire data (Training plus Validation data) to give the intended forecast. This gives the model greater credibility as more data is utilised; also, the validation data has the most recent data which is often the most useful data in forecasting future data. The model is then ready to be used for forecasting.

3.3.6 The LSTM Model

The Software (SW) used for this model is the Python programming language which has some important libraries such as Keras, NumPy, Pandas and matplotlib.

i. The Keras library is used for creating good networks.

ii. The Numpy library is used for numerical analysis.

iii. The Pandas library is used for the robust structuring of data and feature extraction.

iv. Matplotlib is for graphical plotting.

As seen in the Flowchart of Figure 3.3, the libraries are imported after which the training dataset is imported. The data is processed to make it suitable for use. The data is re- shaped to be used by the Numpy library. The LSTM model is created and the test dataset is imported to test the LSTM model. The matplotlib library is finally used to plot the graph.

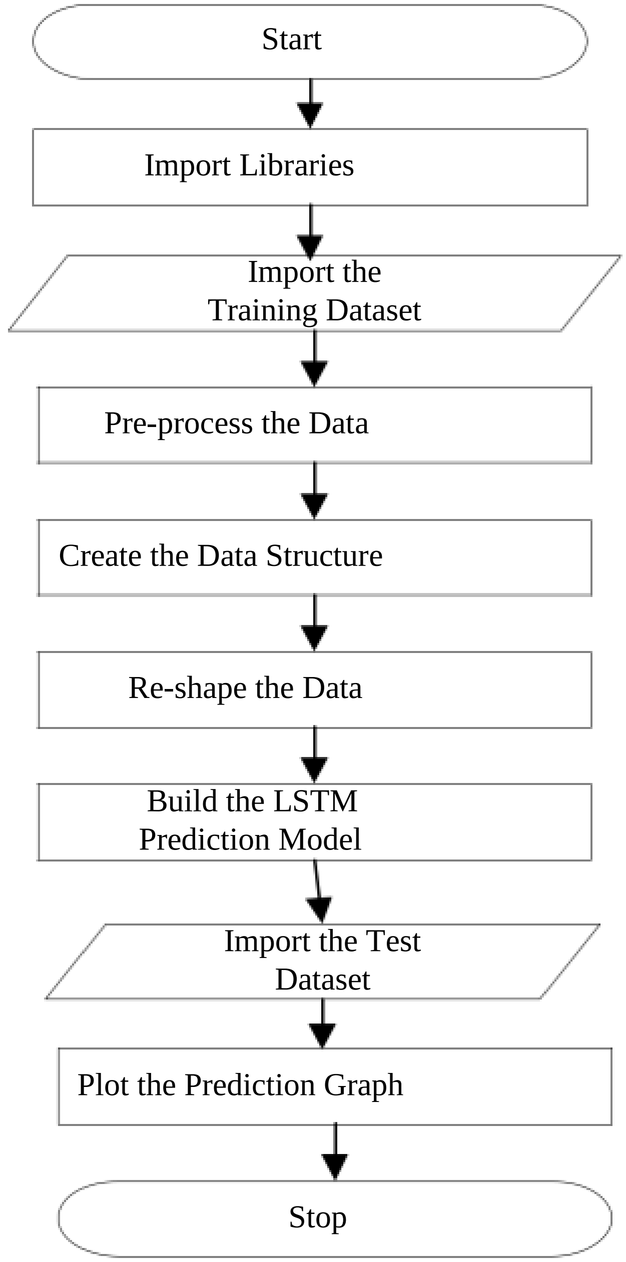


Figure 3.3 Flowchart of the LSTM Model.

The MAE and MAPE are used to evaluate the performance of the LSTM model using equations 3.15 and 3.16, respectively.

3.3.7 Performance Metrics

The reason for evaluating the performance of a predictive model is to determine how well the model performs when it is eventually deployed. The relevant performance metrics of the model are the Mean Absolute Error or deviation (MAE) and the Mean Absolute Percentage Error (MAPE) shown in equations 3.15 and 3.16 for assessing the prediction accuracy of the model.

a. Mean Absolute Error

The MAE is the ratio of the summation of the difference between the observed value (A_t) and the forecasted value (F_t) to the total number of observations (n).

$$= \frac{\Sigma \quad | - |}{} \qquad (3.15)$$

b. Mean Absolute Percentage Error

The mean absolute percentage error is a measure of the error as a percentage of actual value.

It is the average of the absolute errors divided by the actual observed values.

$$= \frac{\Sigma_{=1} | \frac{-}{} |}{} \qquad (3.16)$$

3.4 PdM Scheduling Algorithm

Maintenance impact which is quantified as the change in component availability leading to system availability is key to optimum maintenance scheduling (Rahdar *et al.,* 2020). The main reasons for PdM are to avoid unnecessary maintenance tasks for the reduction of operational cost and to improve system availability.

Fixed interval PM is often adopted by maintenance/facility managers. To improve availability, they shorten the intervals to increase the PM count. Though, this may improve the availability but at increased operating expenditure (OPEX). On the other hand, increasing the interval length to reduce the frequency of PM count will improve OPEX at the expense of system availability.

In this work, the predictive models of Base Station Availability can be used for the PdM schedule as seen in the algorithm of Figure 3.4. Diagnostic tests are conducted on various sub-systems and systems such as the Radio Frequency Units (RFU), transmission (TX) equipment, Base Station Subsystem (BSS) equipment and power equipment. After the diagnostic tests and alarm analysis that are made possible by modules in the equipment, the right decision for scheduling maintenance is triggered using the predictive tool.

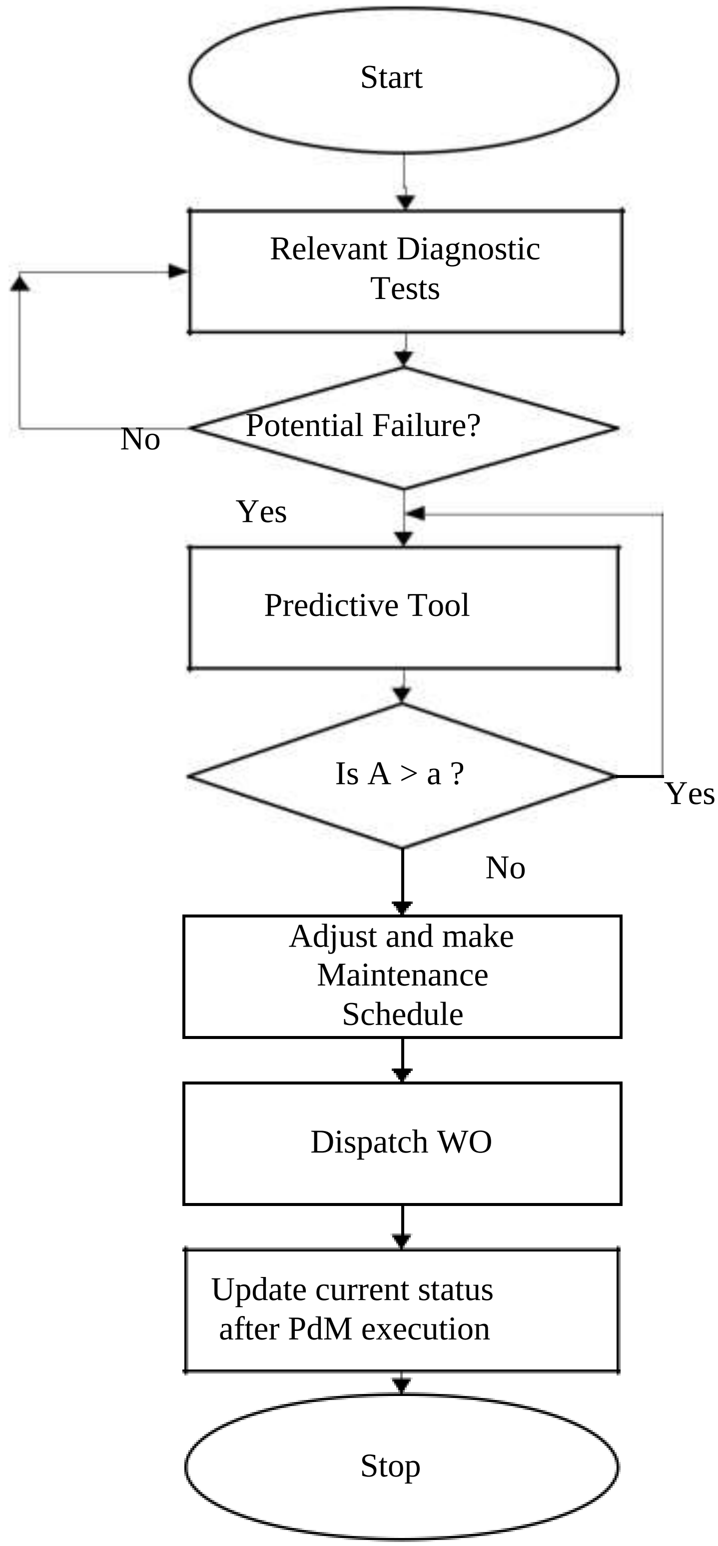


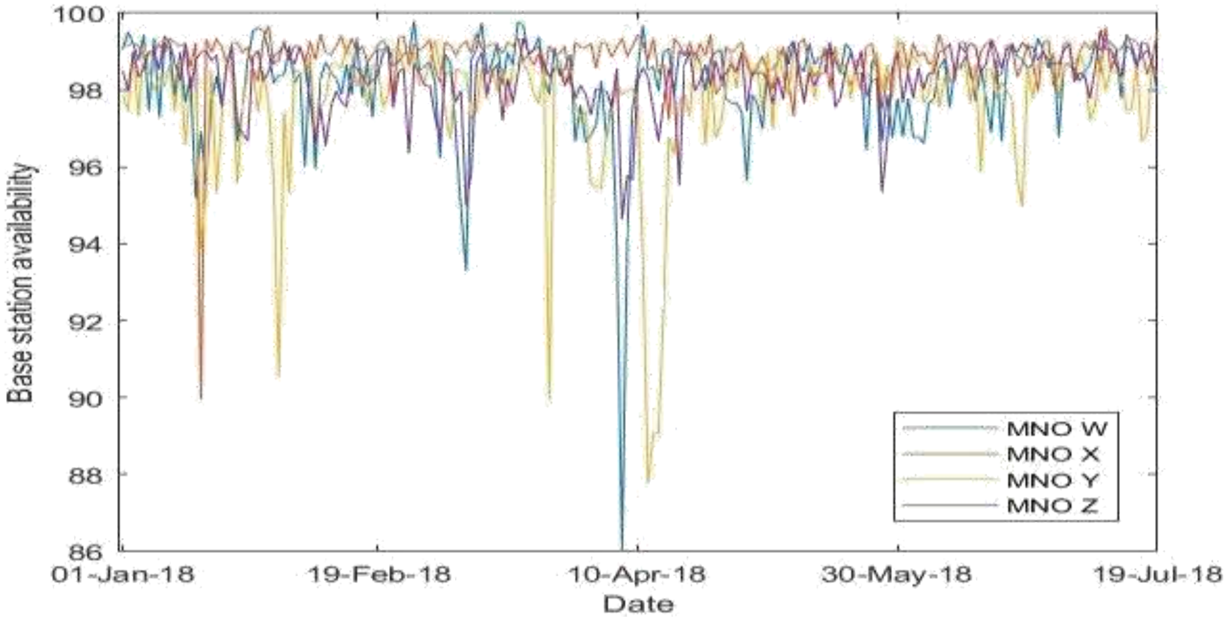
Figure 3.4 PdM Scheduling Algorithm

CHAPTER FOUR

4.0 RESULTS AND DISCUSSION

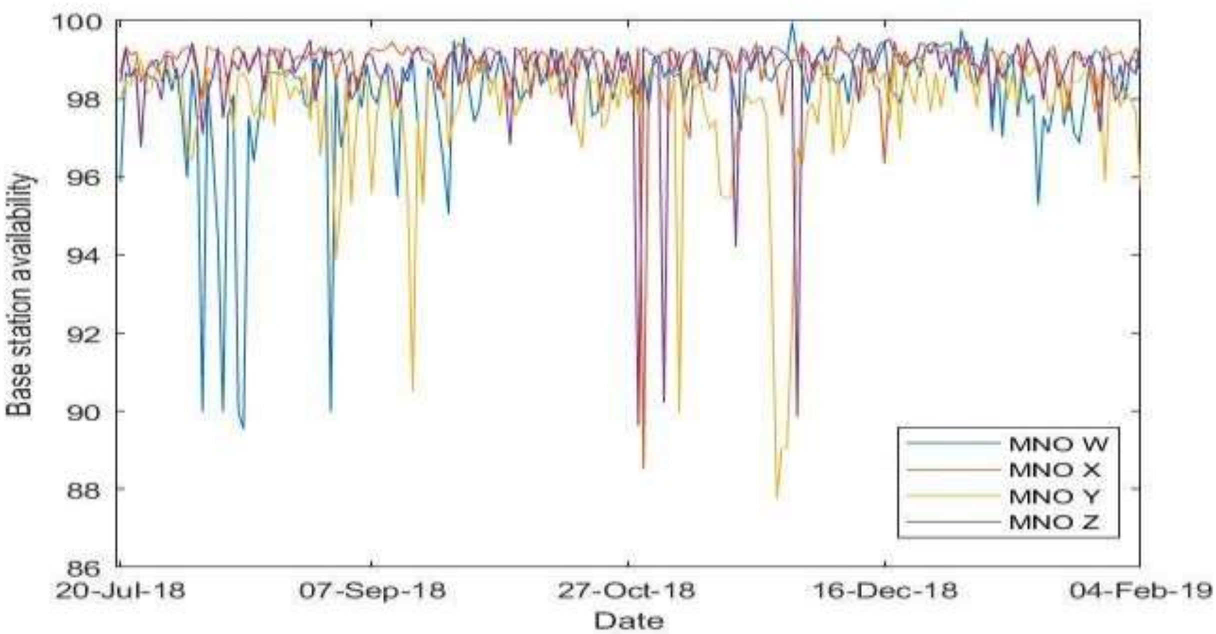
4.1 Plot of BTS Availability for MNO W, X, Y and Z

The TS data for the four MNOs were taken from 1st January 2018 to 26th September 2020 making a thousand data points. The TS plots for the four MNOs are shown in groups of 200 days as shown in Figures 4.1 (a) to (e). The observed availability of the MNOs is seen to dip in the availability values. For instance, in Figure 4.1 (a), MNO W had an Availability of 86% on the 7th of April 2018. MNO Y had similar dips from 11th to 15th April 2018, MNO X dipped in availability value in the middle of January 2018 having an Availability of 90%. In Figure 4.1 (e), MNO Y was observed to have a very poor availability of around 65% in July 2020. These dips in availability values are often caused by factors like weather, equipment failures and inadequate maintenance of passive facilities. These flash periods are crucial for further analysis for proper planning of maintenance (Predictive, Preventive).



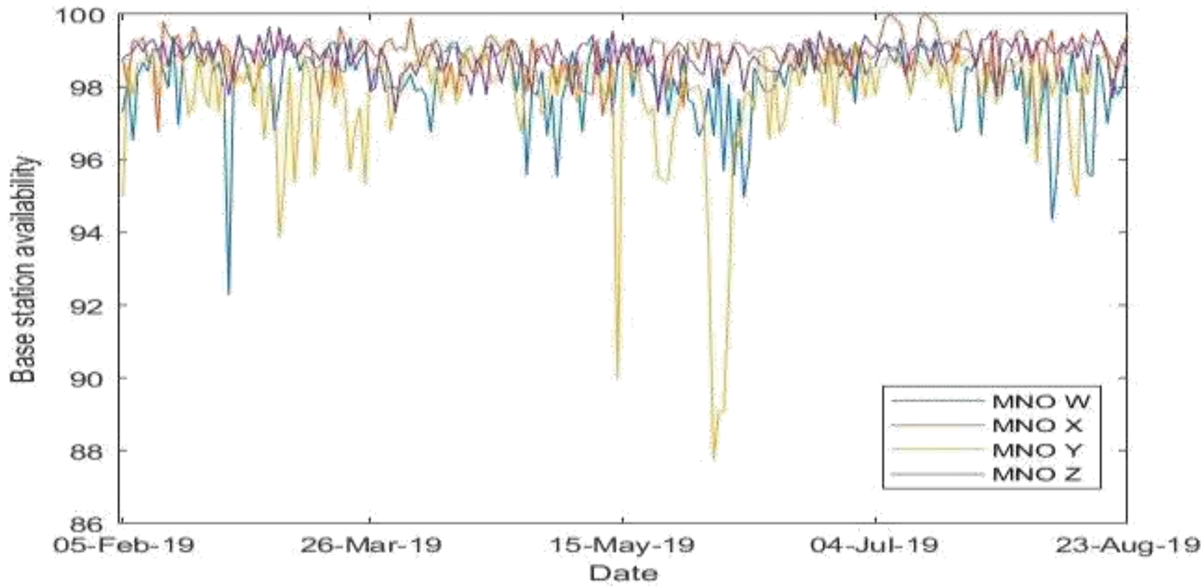
100
98
96
94
92
90
88
86
Base station availability
01-Jan-18
19-Feb-18
10-Apr-18
30-May-18
19-Jul-18
Date
MNO W
MNO X
MNO Y
MNO Z

(a)

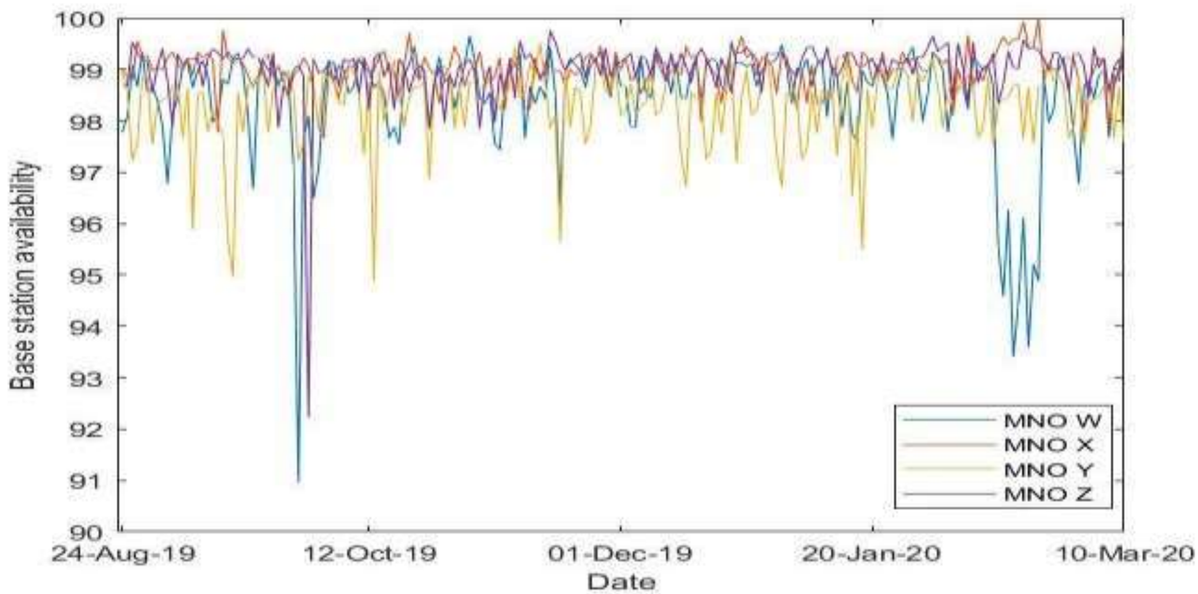


100
98
96
94
92
90
88
86
Base station availability
20-Jul-18
07-Sep-18
27-Oct-18
16-Dec-18
04-Feb-19
Date
MNO W
MNO X
MNO Y
MNO Z

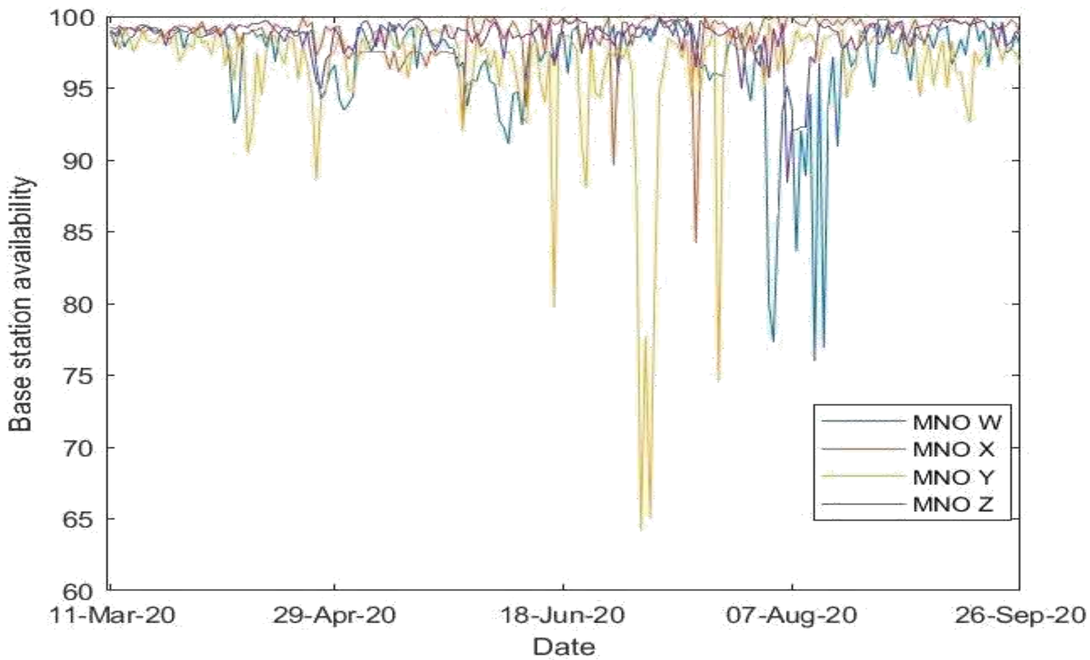
(b)


100
98
96
94
92
90
88
86
Base station availability
05-Feb-19
26-Mar-19
15-May-19
04-Jul-19
23-Aug-19
Date
MNO W
MNO X
MNO Y
MNO Z

(c)


100
99
98
97
96
95
94
93
92
91
90
Base station availability
24-Aug-19
12-Oct-19
01-Dec-19
20-Jan-20
10-Mar-20
Date
MNO W
MNO X
MNO Y
MNO Z

(d)



(e)

Figure 4.1 (a – e) Plots of Base Station Availability in Groups of 200 Days for MNO

For a threshold of 95%, Figure 4.2 shows the percentage number of days having availability values above 95% for the 4 MNOs considering groups of 50days over 1000 days.

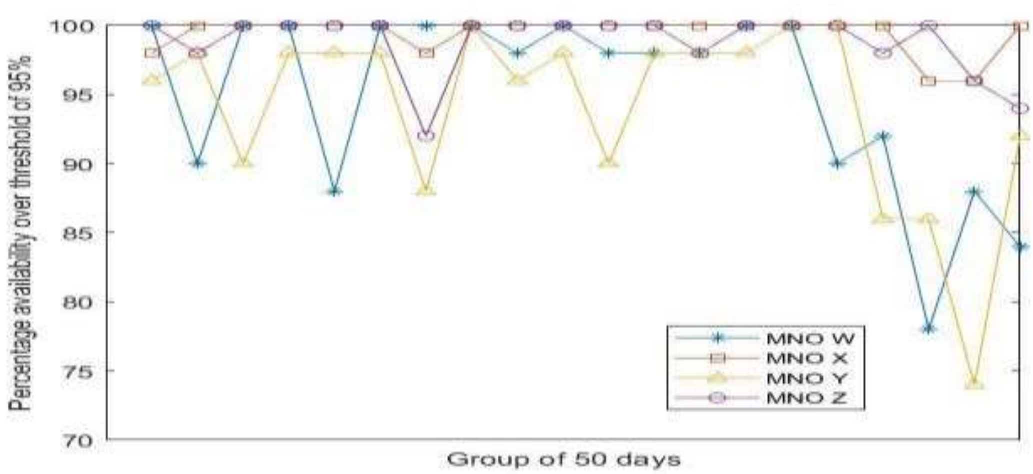


Figure 4.2 Percentage of Base Station Availability Over 95% Considering Groups of 50 Days Over the 1000 days.

Similarly, Figure 4.3 shows the percentage of days in which the MNOs have availability values above the 98% threshold.

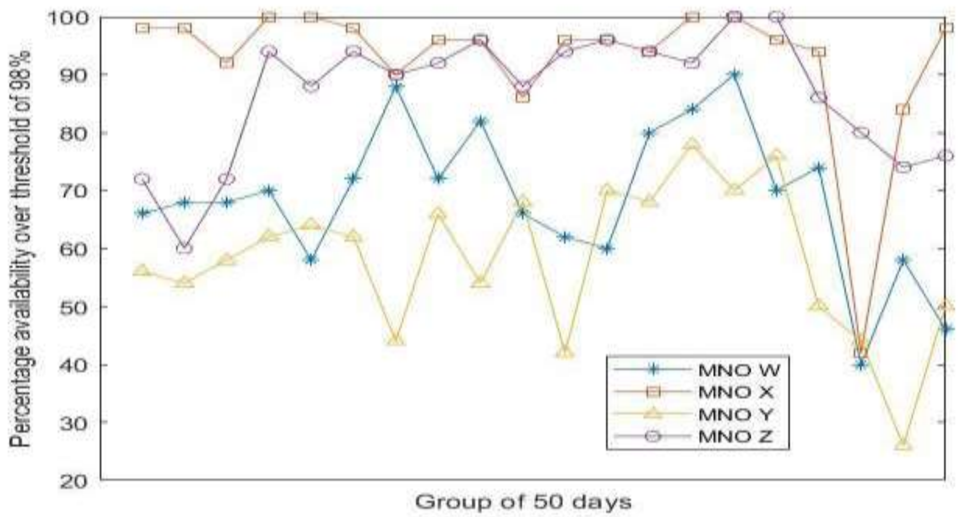


Figure 4.3 Percentage of Base Station Availability Over 98% Considering Groups of 50 Days Over the 1000 days

The ARIMA p, d and q parameters were obtained by using autocorrelation and partial autocorrelation plots are for the systematic determination of the model's parameters. Using SPSS, the plots for the MNOs were taken and illustrated subsequently. Figures 4.4 (a) and (b) are the autocorrelation and partial autocorrelation plots respectively for MNO W.

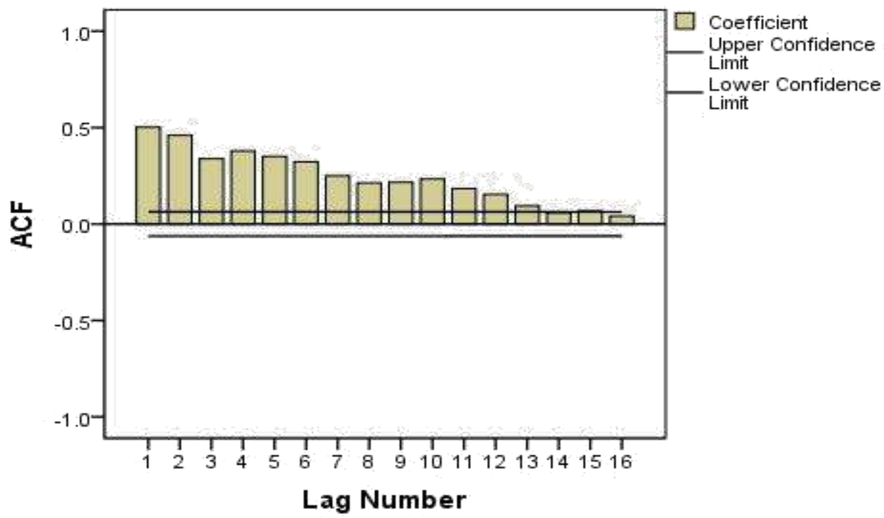


Figure 4.4 (a) Autocorrelation Plot for BTS Availability for MNO W

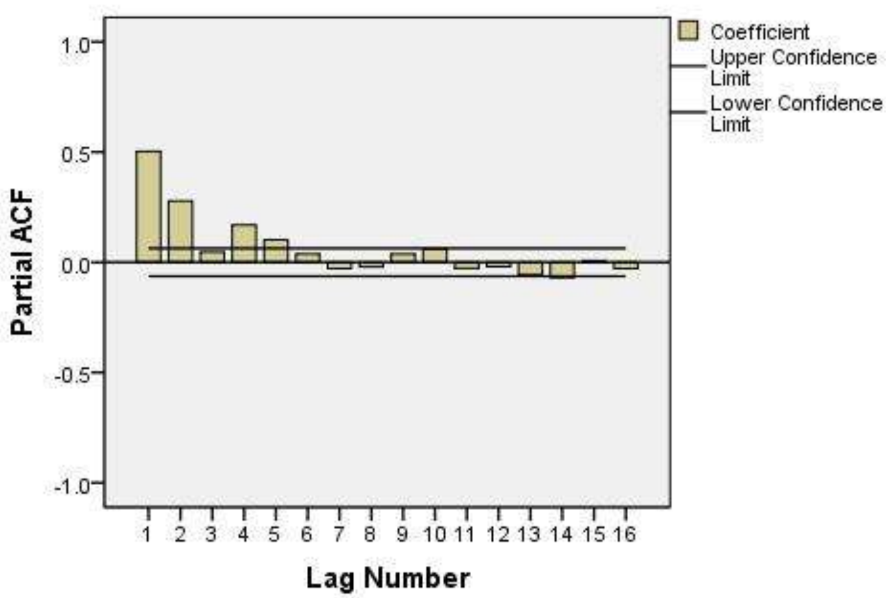


Figure 4.4 (b) Partial Autocorrelation Plot for BTS Availability for MNO W

From Figure 4.4(a), it is observed that the Autocorrelation Function (ACF) is strongly autocorrelated across all the lags (from lag 1 to lag 16 have significant ACF) and this is a characteristic of non-stationary data. For predictive modelling, the data must be stationary, so, the non-stationary data is transformed to a stationary one by a first-order differencing, which means that the value of ‘d’ in the ARIMA model will be equal to 1. The process gives a differenced TS plot for the Availability of MNO W as in Figure 4.5.

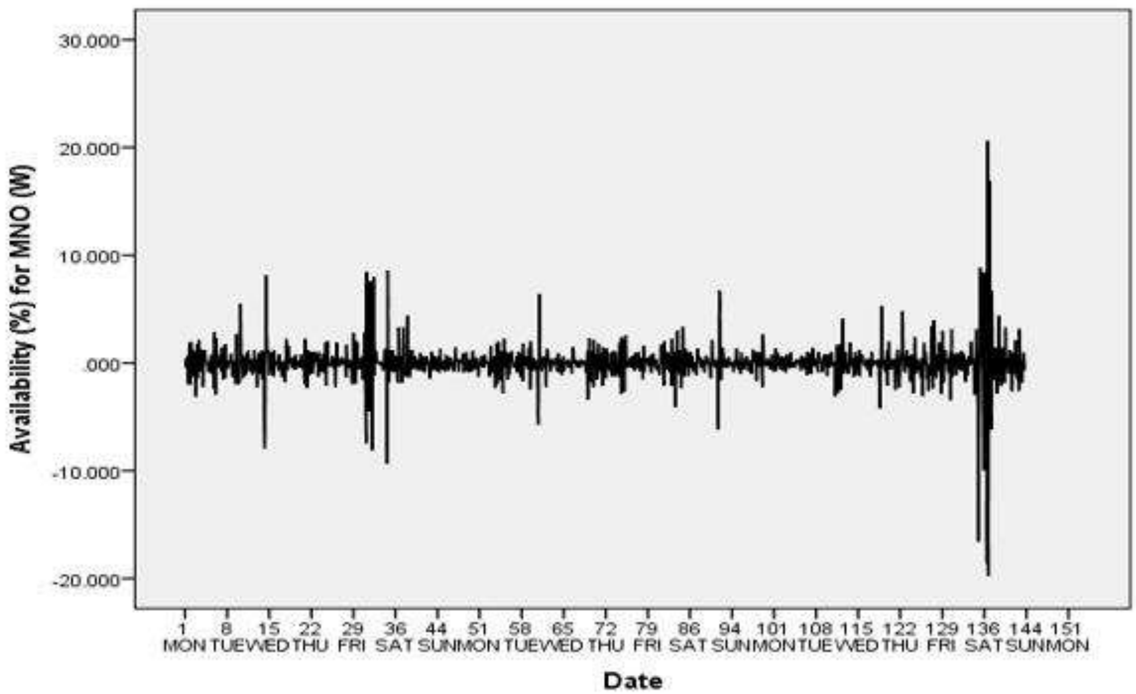


Figure 4.5 Plot of the transformed TS for MNO W with a First Order Differencing.

The plot shows a variation around a constant mean (around 0.0) and the variance around the mean is within a reasonable bound which is expected for stationary data. The ACF and PACF plots done on this transformed TS gave the outcomes depicted in Figures 4.6 (a) and (b). The earlier (starting from lag 1 to lag 16) significant lags in the ACF and PACF plots indicate the value for the ARIMA parameters q and p respectively. A close examination of Figures 4.6 (a) and (b) show that q is equal to three while p is equal to zero. This is to say that the ARIMA (p, d, q) parameters for MNO W are (0,1,3).

ARIMA (p, d, q)$_{\text{MNO W}}$ = (0,1,3).

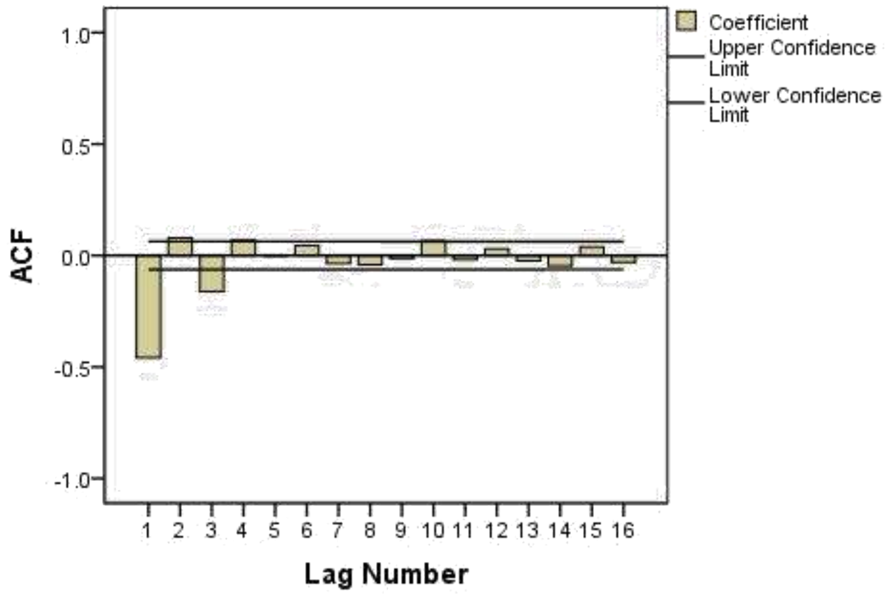


Figure 4.6 (a) ACF Plot for the Transformed TS Data for MNO W.

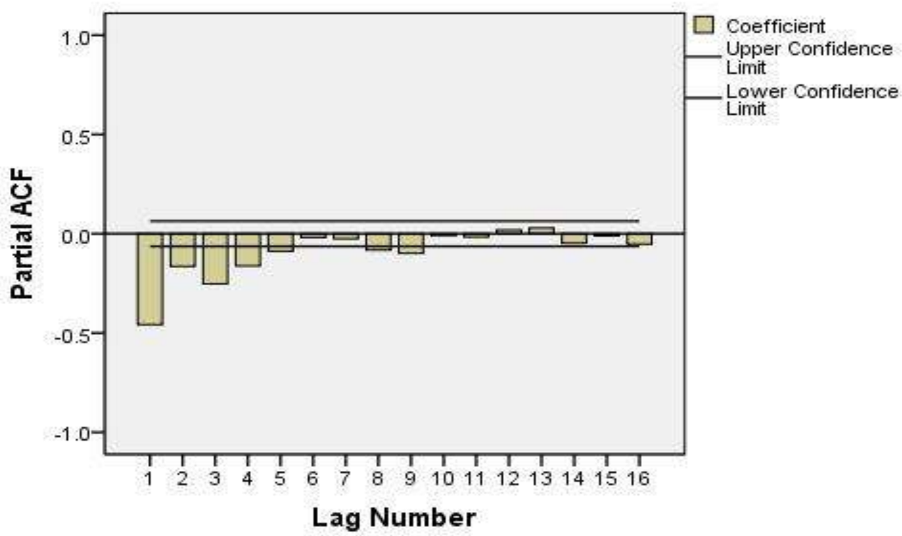


Figure 4.6 (b) PACF Plot for the Transformed TS Data for MNO W.

A similar procedure gave the ARIMA parameters for the other MNOs and the parameters are presented in Table 4.1.

4.2 Creating the Predictive Models for the MNOs

The ARIMA model parameters for each of the MNOs were used in the creation of the predictive model for the BTS availability for the MNOs. The parameters are arranged in Table 4.1.

Table 4.1 Model Parameters for the MNOs

MNO	ARIMA MODEL
W	ARIMA (0,1,3)
X	ARIMA (1,0,1)
Y	ARIMA (2,0,4)
Z	ARIMA (0,1,1)

Source: Author's Field Work.

The BTS availability data as a TS data was taken for a period of 143 weeks (from 1st January 2018 – 26th September 2020) which is from week 1 to week 143; while the model, is to forecast the availability data up to week 155. The plots of the BTS availability for MNO using the ARIMA models are presented in Figures 4.11, 4.12, 4.13 and 4.14.

The observed Availability data spanned from 1st January 2018 to 26th September 2020 corresponding to 1000 data points. The result of the prediction from 27th September 2020 to 20th December 2020 is displayed in Table 4.2.

From the prediction, MNO would have a working tool for proactive rather than reactive planning for scheduling operations such as PPM and PdM on BTS and even network nodes for the improvement of Availability. Since the prediction will beam more light on the general health status of Base Station and other NEs, appropriate decision-making processes that enhance efficiency are made possible and nodes with repetitive failures could be furnished with redundancies.

Figures 4.7 to 4.10 are the plots of the ARIMA prediction during the validation period.

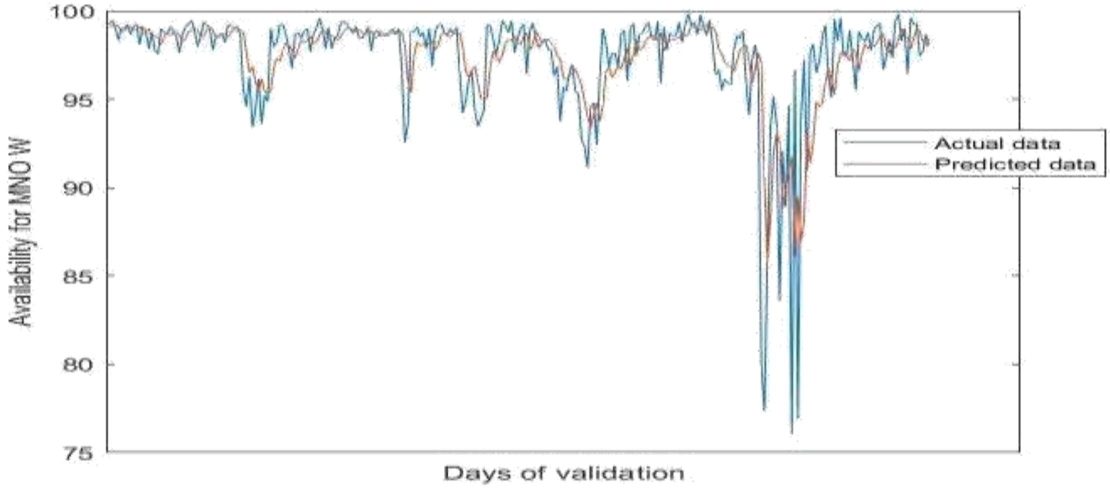


Figure 4.7 Plot of predicted availability for MNO W (Validation period)

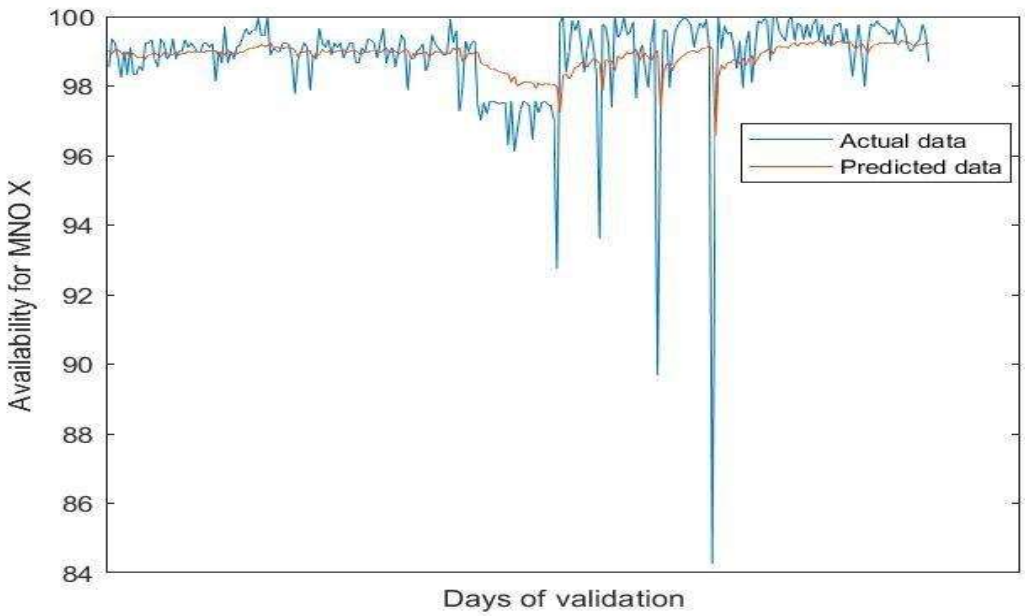


Figure 4.8 Plot of predicted availability for MNO X (Validation period)

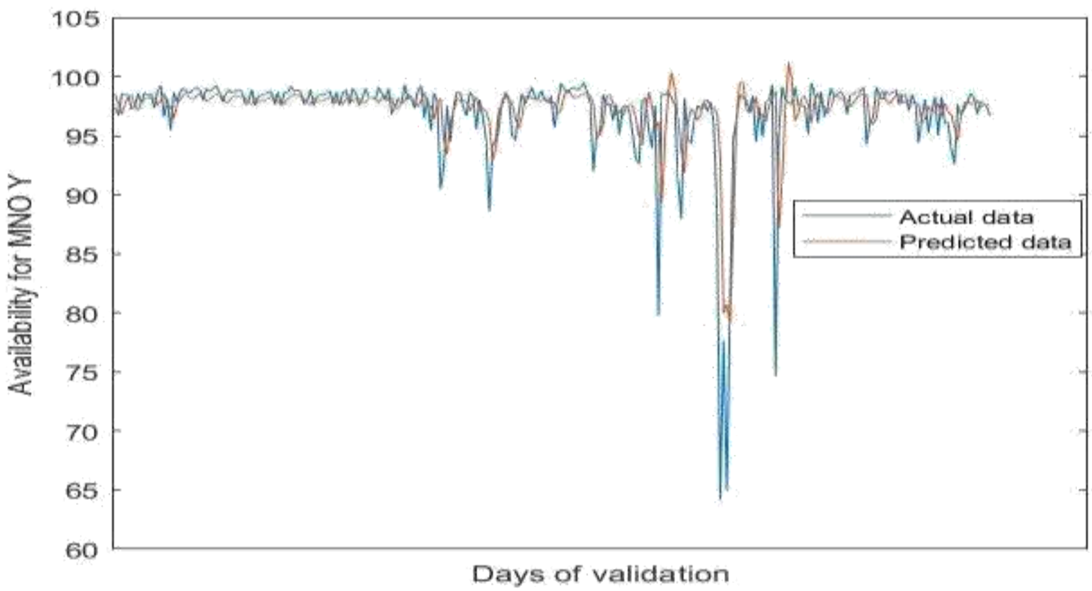


Figure 4.9 Plot of predicted availability for MNO Y (Validation period)

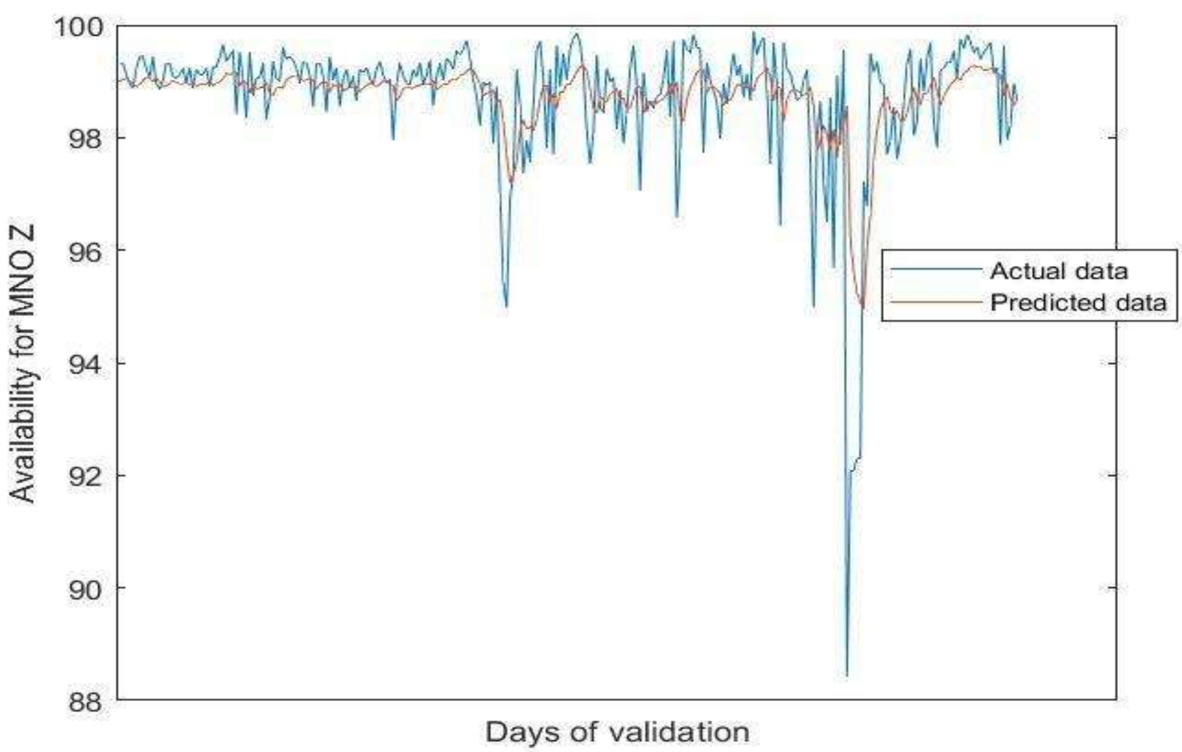


Figure 4.10 Plot of predicted availability for MNO Z (Validation period)

Table 4.2 Predicted BTS Availability for MNO (ARIMA)

Date	MNO W	MNO X	MNO Y	MNO Z	Date	MNO W	MNO X	MNO Y	MNO Z
27-Sep-20	98.34	99.13	97.26	98.81	09-Nov-20	98.29	98.91	97.63	98.82
28-Sep-20	98.25	99.12	97.18	98.81	10-Nov-20	98.28	98.91	97.63	98.82
29-Sep-20	98.31	99.10	97.32	98.81	11-Nov-20	98.28	98.91	97.63	98.82
30-Sep-20	98.31	99.09	97.53	98.81	12-Nov-20	98.28	98.91	97.63	98.82
01-Oct-20	98.30	99.08	97.65	98.81	13-Nov-20	98.28	98.91	97.63	98.82
02-Oct-20	98.30	99.07	97.62	98.81	14-Nov-20	98.28	98.91	97.63	98.82
03-Oct-20	98.30	99.06	97.63	98.81	15-Nov-20	98.28	98.90	97.63	98.82
04-Oct-20	98.30	99.05	97.63	98.81	16-Nov-20	98.28	98.90	97.63	98.82
05-Oct-20	98.30	99.04	97.63	98.81	17-Nov-20	98.28	98.90	97.63	98.82
06-Oct-20	98.30	99.03	97.63	98.81	18-Nov-20	98.28	98.90	97.63	98.82
07-Oct-20	98.30	99.02	97.63	98.81	19-Nov-20	98.28	98.90	97.63	98.82
08-Oct-20	98.30	99.01	97.63	98.81	20-Nov-20	98.28	98.90	97.63	98.82
09-Oct-20	98.30	99.00	97.63	98.81	21-Nov-20	98.28	98.90	97.63	98.82
10-Oct-20	98.30	99.00	97.63	98.81	22-Nov-20	98.28	98.90	97.63	98.82
11-Oct-20	98.30	98.99	97.63	98.81	23-Nov-20	98.28	98.90	97.63	98.82
12-Oct-20	98.30	98.98	97.63	98.81	24-Nov-20	98.28	98.90	97.63	98.82
13-Oct-20	98.30	98.98	97.63	98.81	25-Nov-20	98.28	98.90	97.63	98.82
14-Oct-20	98.30	98.97	97.63	98.81	26-Nov-20	98.28	98.90	97.63	98.82
15-Oct-20	98.30	98.97	97.63	98.81	27-Nov-20	98.28	98.90	97.63	98.82
16-Oct-20	98.30	98.96	97.63	98.81	28-Nov-20	98.28	98.90	97.63	98.82
17-Oct-20	98.30	98.96	97.63	98.81	29-Nov-20	98.28	98.90	97.63	98.82
18-Oct-20	98.30	98.96	97.63	98.81	30-Nov-20	98.27	98.90	97.63	98.82
19-Oct-20	98.30	98.95	97.63	98.81	01-Dec-20	98.27	98.90	97.63	98.82
20-Oct-20	98.30	98.95	97.63	98.81	02-Dec-20	98.27	98.90	97.63	98.82
21-Oct-20	98.29	98.94	97.63	98.81	03-Dec-20	98.27	98.90	97.63	98.82
22-Oct-20	98.29	98.94	97.63	98.81	04-Dec-20	98.27	98.90	97.63	98.82
23-Oct-20	98.29	98.94	97.63	98.82	05-Dec-20	98.27	98.90	97.63	98.82
24-Oct-20	98.29	98.94	97.63	98.82	06-Dec-20	98.27	98.90	97.63	98.83
25-Oct-20	98.29	98.93	97.63	98.82	07-Dec-20	98.27	98.90	97.63	98.83
26-Oct-20	98.29	98.93	97.63	98.82	08-Dec-20	98.27	98.90	97.63	98.83
27-Oct-20	98.29	98.93	97.63	98.82	09-Dec-20	98.27	98.90	97.63	98.83
28-Oct-20	98.29	98.93	97.63	98.82	10-Dec-20	98.27	98.90	97.63	98.83
29-Oct-20	98.29	98.92	97.63	98.82	11-Dec-20	98.27	98.90	97.63	98.83
30-Oct-20	98.29	98.92	97.63	98.82	12-Dec-20	98.27	98.90	97.63	98.83
31-Oct-20	98.29	98.92	97.63	98.82	13-Dec-20	98.27	98.90	97.63	98.83
01-Nov-20	98.29	98.92	97.63	98.82	14-Dec-20	98.27	98.90	97.63	98.83
02-Nov-20	98.29	98.92	97.63	98.82	15-Dec-20	98.27	98.90	97.63	98.83
03-Nov-20	98.29	98.92	97.63	98.82	16-Dec-20	98.27	98.90	97.63	98.83
04-Nov-20	98.29	98.91	97.63	98.82	17-Dec-20	98.27	98.90	97.63	98.83

05-Nov-20	98.29	98.91	97.63	98.82	18-Dec-20	98.27	98.90	97.63	98.83
06-Nov-20	98.29	98.91	97.63	98.82	19-Dec-20	98.27	98.90	97.63	98.83
07-Nov-20	98.29	98.91	97.63	98.82	20-Dec-20	98.26	98.90	97.63	98.83
08-Nov-20	98.29	98.91	97.63	98.82					

Source: Author's Field Work.

Figures 4.11, 4.12, 4.13 and 4.14 show the predicted BTS availability graph for MNO W, MNO X, MNO Y and MNO Z, respectively. The BTS Availability data were analysed using their ACF and PACF plots. MNO W and MNO Z were observed not to be stationary, a requirement for TSA is that the data should be stationary. After the process of differencing, the p, d, q values for the ARIMA models were got. MNO X and MNO Y did not require the process of differencing and their ARIMA model parameters were obtained. The ARIMA models for all the MNOs are: (0,1,3), (1,0,1), (2,0,4) and (0,1,1) for MNO W, MNO X, MNO Y and MNO Z, respectively as contained in Table 4.1. The parameters were used for their respective data and the predictive models were used for prediction and the performance of the models were evaluated. The MAE and MAPE results are shown in Table 4.3.

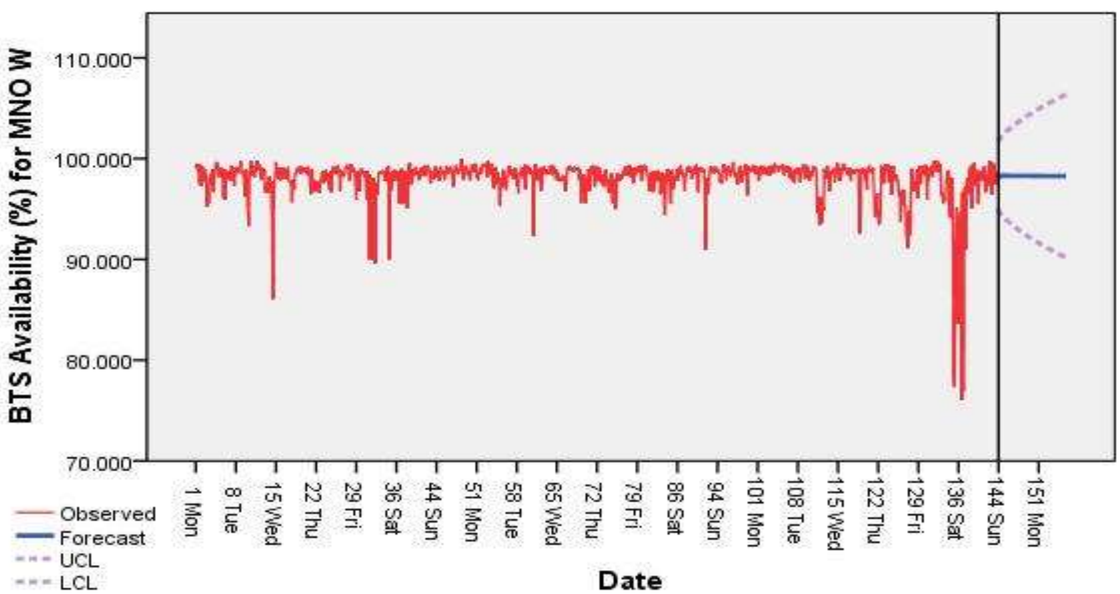


Figure 4.11 Predicted BTS Availability for MNO W Using ARIMA (0,1,3) Model.

Also, the ARIMA (1,0,1) model was used for the prediction of BTS Availability for MNO X and is shown in Figure 4.12.

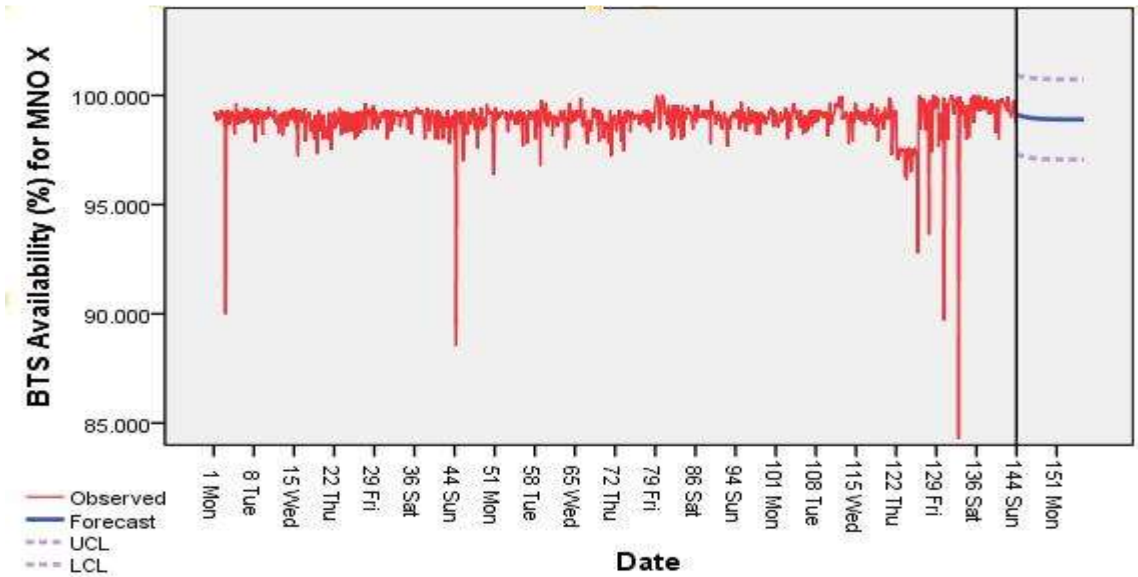


Figure 4.12 Predicted BTS Availability for MNO X Using ARIMA (1,0,1) Model.

Similarly, the predicted BTS availability are presented in figures 4.12 and 4.13 using ARIMA (2,0,4) and ARIMA (0,1,1) respectively.

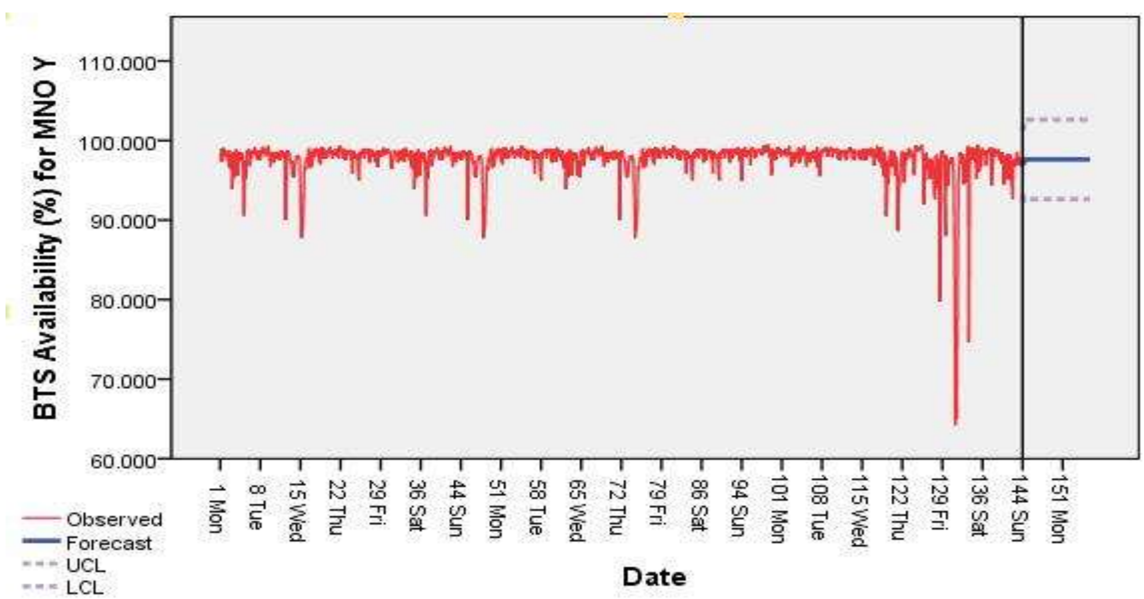


Figure 4.13 Predicted BTS Availability for MNO Y Using ARIMA (2,0,4) Model

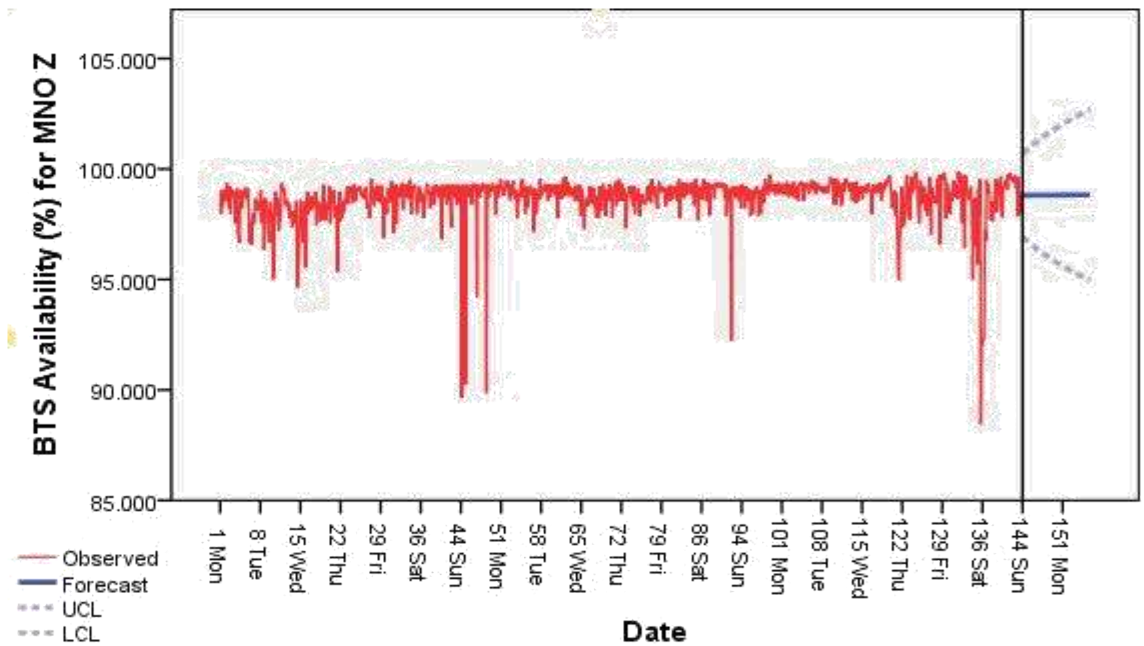
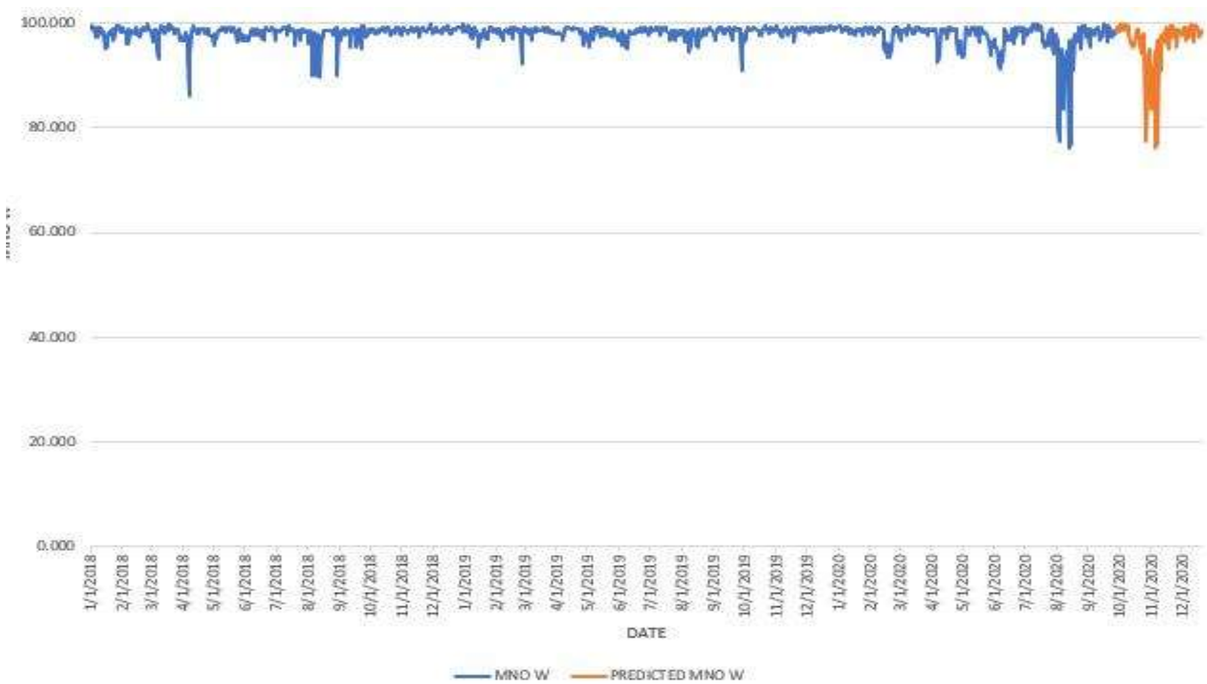
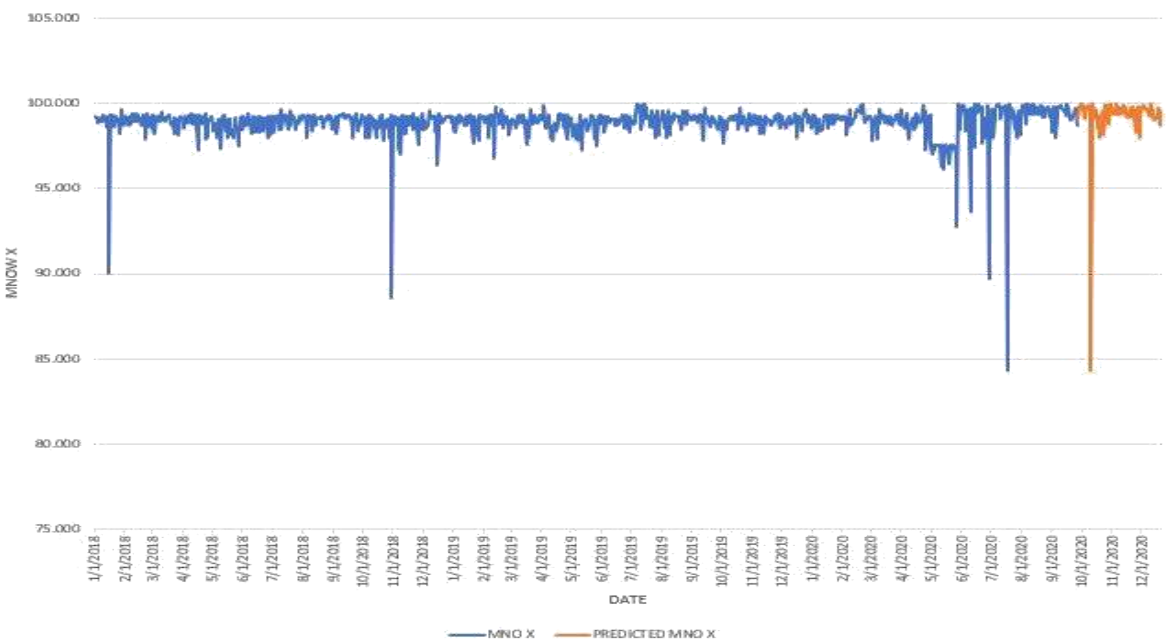


Figure 4.14 Predicted BTS Availability for MNO Z Using ARIMA (0,1,1) Model.

4.3 Results of the LSTM Models

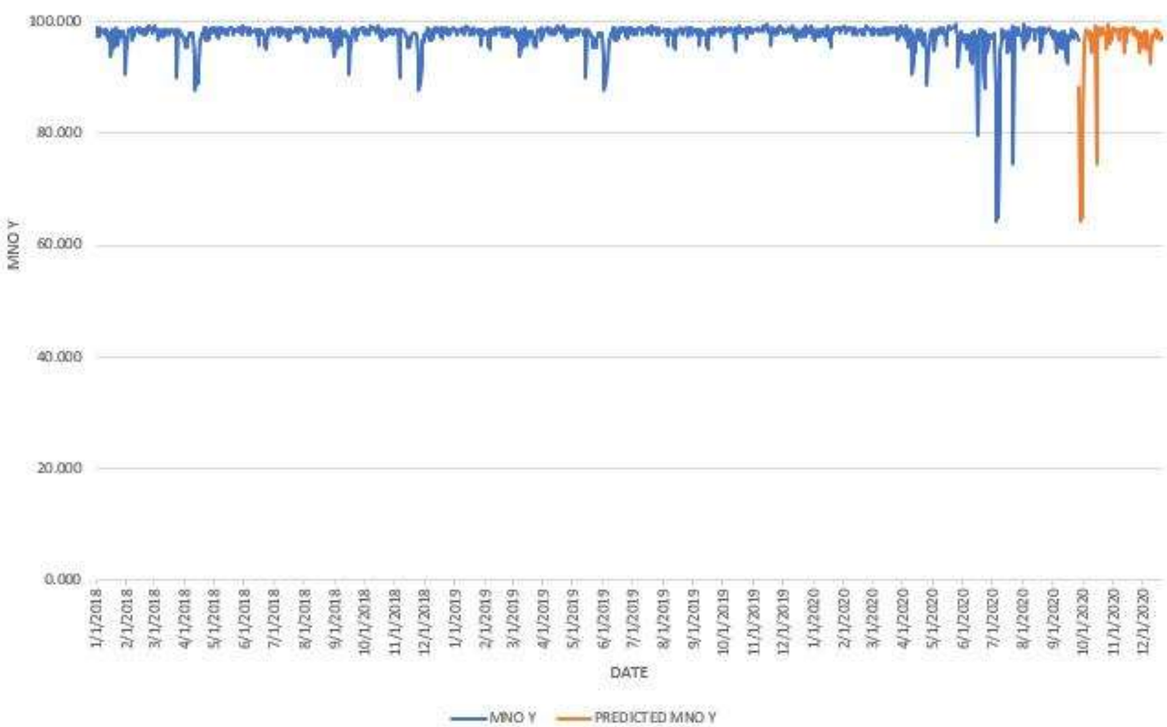


(a)



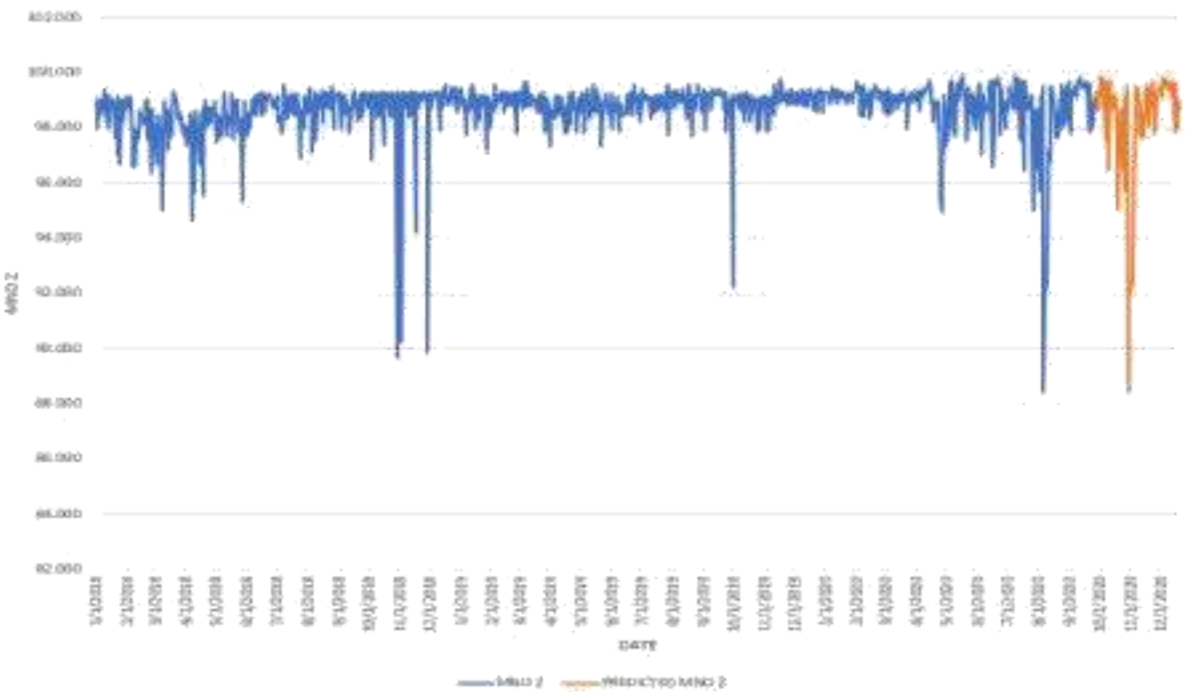
105.000
100.000
95.000
90.000
85.000
80.000
75.000
MNOW X
1/1/2018
2/1/2018
3/1/2018
4/1/2018
5/1/2018
6/1/2018
7/1/2018
8/1/2018
9/1/2018
10/1/2018
11/1/2018
12/1/2018
1/1/2019
2/1/2019
3/1/2019
4/1/2019
5/1/2019
6/1/2019
7/1/2019
8/1/2019
9/1/2019
10/1/2019
11/1/2019
12/1/2019
1/1/2020
2/1/2020
3/1/2020
4/1/2020
5/1/2020
6/1/2020
7/1/2020
8/1/2020
9/1/2020
10/1/2020
11/1/2020
12/1/2020
DATE
MNO X
PREDICTED MNO X

(b)



100.000
80.000
60.000
40.000
20.000
0.000
MNO Y
1/1/2018
2/1/2018
3/1/2018
4/1/2018
5/1/2018
6/1/2018
7/1/2018
8/1/2018
9/1/2018
10/1/2018
11/1/2018
12/1/2018
1/1/2019
2/1/2019
3/1/2019
4/1/2019
5/1/2019
6/1/2019
7/1/2019
8/1/2019
9/1/2019
10/1/2019
11/1/2019
12/1/2019
1/1/2020
2/1/2020
3/1/2020
4/1/2020
5/1/2020
6/1/2020
7/1/2020
8/1/2020
9/1/2020
10/1/2020
11/1/2020
12/1/2020
DATE
MNO Y
PREDICTED MNO Y

(c)



(d)

Figure 4.15 LSTM Predicted BTS Availability for (a) MNO W, (b) MNO X, (c) MNO Yand (d) MNO Z

4.4 Prediction Accuracy

The data used in the ARIMA model was shared into the training (from 1st January 2018 – 31st December 2019) data which is about 73% and validation (1st January 2020 – 26th September 2020) represented about 27% of the entire data. The performance metrics (MAE and MAPE) were used to evaluate data in the validation period. Table 4.3 shows the MAE and MAPE for the ARIMA models.

The performance metrics from Table 4.3 indicates that the predictive model for MNO Z has the best accuracy, followed by that of MNO X, MNO W and MNO Y in the order in terms of both MAE and MAPE.

Table 4.3 Performance Metrics for the ARIMA Models

MNO	MAE	MAPE
MNO W	1.40	0.015
MNO X	0.66	0.0068
MNO Y	1.57	0.018
MNO Z	0.62	0.0063

Source: Authors Field Work

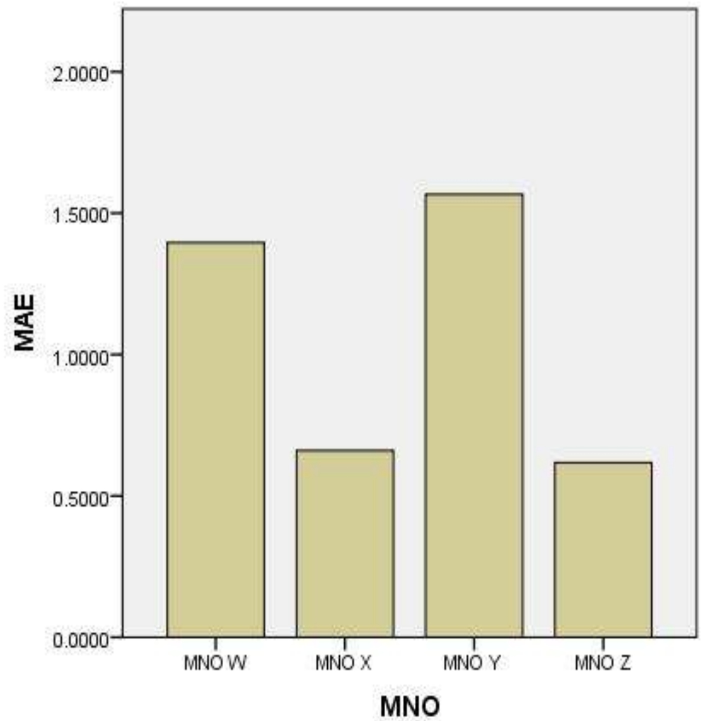


Figure 4.16 MAE against MNO

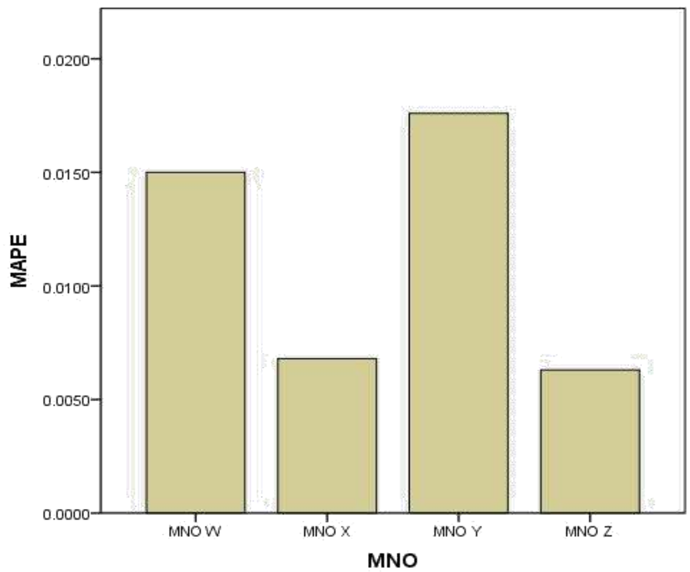


Figure 4.17 MAPE against MNO

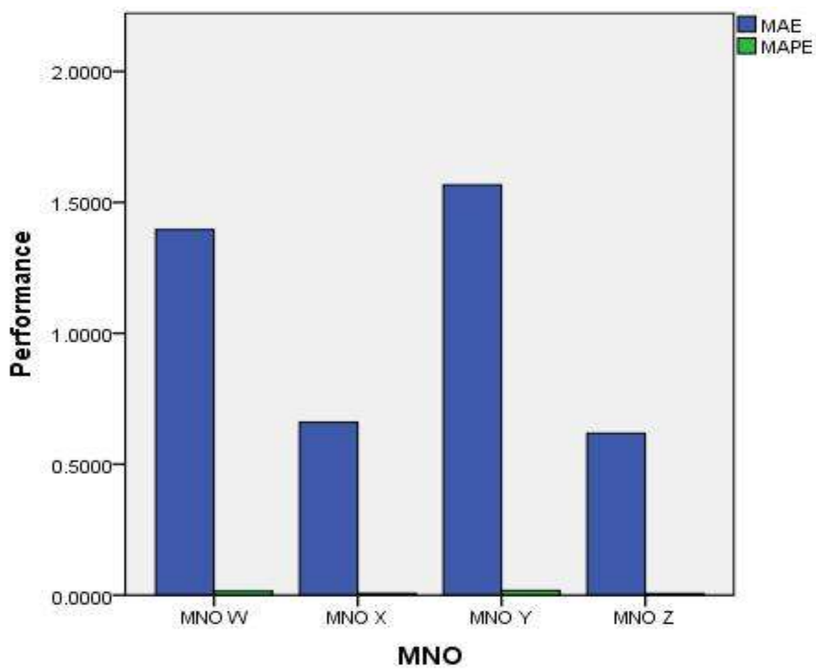


Figure 4.18 Performance Chart for the MNO.

Table 4.4 Performance Metrics Comparison for the Models (MAE)

MNO	ARIMA	LSTM
MNO W	1.40	2.84
MNO X	0.66	0.89
MNO Y	1.57	2.82
MNO Z	0.62	1.12

Table 4.5 Performance Metrics Comparison for the Model (MAPE)

MNO	ARIMA	LSTM
MNO W	0.015	0.032
MNO X	0.0068	0.0092
MNO Y	0.0176	0.035
MNO Z	0.0063	0.012

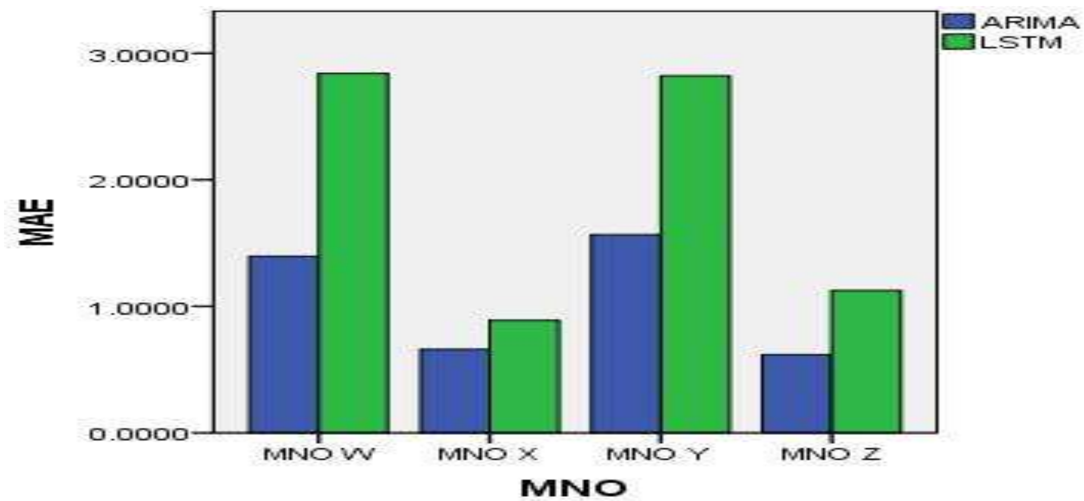


Figure 4.19 MAE Comparison for ARIMA and LSTM

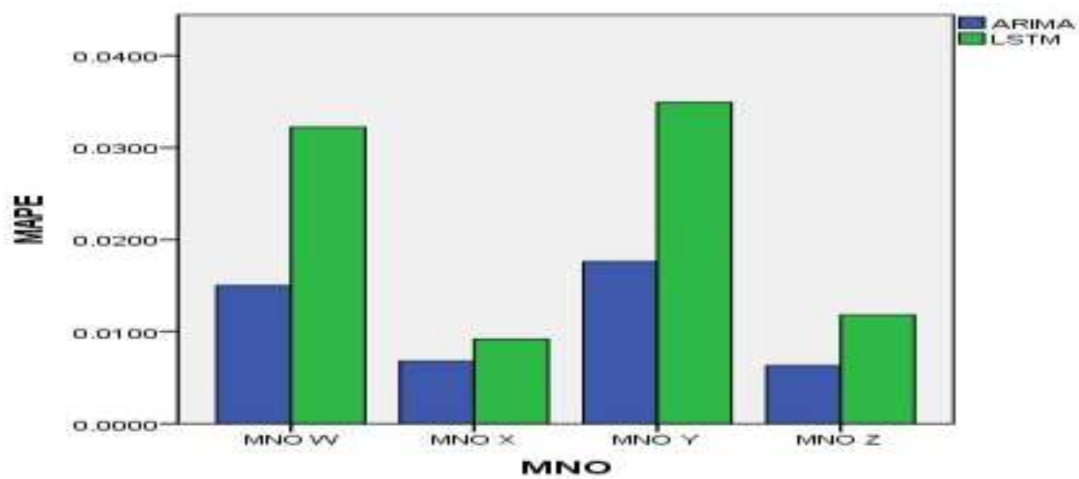


Figure 4.20 MAPE Comparison for ARIMA and LSTM.

From the results, it is observed that the LSTM models have higher MAE values than the ARIMA models by 51%, 26%, 44% and 45% for MNO W, MNO X, MNO Y and MNO Z respectively. Similarly, for MAPE, the LSTM models have 53%, 26%, 50% and 47% higher values than the ARIMA models for the respective MNOs.

4.5 Result of Scheduling Algorithm

The result of the PdM scheduling algorithm is contained in Table 4.6. The maintenance counts for the MNOs based on the ARIMA prediction were estimated using the thresholds of 95% and 90%.

Table 4.6 Scheduling Algorithm for the MNOs Using Thresholds

MNO W		MNO X		MNO Y		MNO Z	
Maintenance count based on predicted availability < 95	Number of saved maintenances	Maintenance count based on predicted availability < 95	Number of saved maintenances	Maintenance count based on predicted availability < 95	Number of saved maintenances	Maintenance count based on predicted availability < 95	Number of saved maintenances
34	0	0	33	36	0	1	32
Maintenance count based on predicted availability < 90	Number of saved maintenances	Maintenance count based on predicted availability < 90	Number of saved maintenances	Maintenance count based on predicted availability < 90	Number of saved maintenances	Maintenance count based on predicted availability < 90	Number of saved maintenances
8	25	0	33	6	27	0	33

Based on the algorithm that determines when maintenance should be carried out at the Base Station based on predicted low availabilities, the number of maintenances was calculated for each MNO using two results; 95% and 90% availabilities as thresholds for doing maintenance.

These were compared with a fixed maintenance schedule of once-monthly which results in 33 maintenances within the period under investigation.

MNO X and MNO Z have higher counts of Base Station availability for both thresholds. MNO Y has the least number of Availability counts of high Availability above the thresholds. The MNOs in order of ranking based on Base Station Availability in descending order are MNO Z, MNO X, MNO W and MNO Y.

Based on the Predictive models, MNO Z has the best accuracy, followed by MNO X, and MNO W. MNO Y is the worst in accuracy.

CHAPTER FIVE

5.0 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

There is a great dependence on communication services in modern times and the demand for very high Availability is now taking an important concern as users of the services must get their devices connected irrespective of the status of the channels. This quest for very high Availability is even expected to increase with the envisaged full deployment of 5G technology. Of what use is it going to be anyway if, after a well-deployed technology, service is not available when it is needed most? Applications like the IoT, telemedicine, Forex and Cryptocurrency trading will demand very high BTS availability and MNOs would do much to retain their subscribers otherwise, they will lose out.

Conscious efforts have been put by telecommunications vendors to improve the reliability and the MTBF of components and equipment. Such components and devices have strict Availabilities of five or six ‘nines.’ This has not yielded the desired BTS Availability for the MNOs in Minna as studied in this research due to the major Availability barrier, namely: MTTR. Even if the system fails, it should be restored fast, such that the period of the outage becomes minimal. Therefore, it is necessary to come up with a means of reducing MTTR to boost Availability.

In this research, the historic BTS Availability data for a thousand data points were used for designing ARIMA-based predictive models for four MNOs in Minna from 1st January 2018 to 26th September 2020. 73% of the data was used for model training, the remaining 27% was used for model validation and a prediction was made by the model from 27th September

2020 to 20th December 2020. The model parameters (p,d,q) used after a careful ACF and PACF analysis are: ARIMA (0,1,3), ARIMA (1,0,1), ARIMA (2,0,4) and ARIMA (0,1,1) for MNO W, MNO X, MNO Y and MNO Z respectively. The performance metrics used for the models are the MAE and MAPE. The MAE and MAPE for MNO W, MNO X, MNO Y and MNO Z are 1.40 and 0.015; 0.66 and 0.0068; 1.57 and 0.018; and 0.62 and 0.0063 respectively. Based on the results of the performance metrics, the ranking of the MNOs in terms of descending order of model performance are MNO Z, MNO X, MNO W and MNO Y, respectively. It is pertinent to note that low values of MAE and MAPE across all the MNOs indicate that the prediction values are close to the actual values and are therefore valid for good planning and decision making for proper PdM and PPM geared towards a substantial reduction of MTTR, which invariably improves the BTS Availability of the MNOs.

The LSTM models were compared with the ARIMA models. Their MAE and MAPE are: 2.84, 0.89, 2.82, and 1.12; 0.032, 0.0092, 0.035 and 0.012 for MNO W, MNO X, MNO Y and MNO Z respectively. From the results, it is observed that the ARIMA models performed better than the LSTM models in all the MNOs.

5.2 Recommendations

It would be recommended that MNOs should pay more attention to perceived interruptions and their causes by careful study of the trends in the BTS Availability to dedicate their resources to units and areas that show lower Availability. This helps in curtailing hidden future failures and reduces operational costs.

In this work, the concept for availability focused on the MNOs. It is recommended that the concept could be investigated from the customer or subscribers' perspective.

5.3 Contributions to Knowledge

This work has contributed to knowledge by developing:

i. ARIMA-based predictive models of Base Station Availability for MNOs in Minna.

ii. PdM scheduling algorithm for proactive rather than reactive planning of BTS and network nodes maintenance for the improvement of availability and OPEX.

iii. Provision of Base Station Availability data for MNOs in Minna.

.

REFERENCES

Ahmed, W., Hasan, O., Pervez, U., & Qadir, J. (2017). Reliability modelling and analysis of communication networks. *Journal of Network and Computer Applications, 78,* 191–215. https://doi.org/10.1016/j.jnca.2016.11.008

Akinsanmi, O., & Adebuyi, K. (2016). *Assessment of the Reliability of Global Services Mobile Communication Networks on Quality of Services (QoS): A Comparative Study of Four Major Giants*. *5*(6), 552–562.

Almeida, J. S. (2002). Predictive non-linear modelling of complex data by artificial neural networks. In *Current Opinion in Biotechnology* (Vol. 13, Issue 1, pp. 72–76). https://doi.org/10.1016/S0958-1669(02)00288-4

Bakar, N. A., & Rosbi, S. (2017). Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Cryptocurrency Exchange Rate in High Volatility Environment: A New Insight of Bitcoin Transaction. *International Journal of Advanced Engineering Research and Science*, *4*(11), 130–137. https://doi.org/10.22161/ijaers.4.11.20

Beuzen, T., Splinter, K. D., Marshall, L. A., Turner, I. L., Harley, M. D., & Palmsten, M. L. (2018). Bayesian Networks in coastal engineering: Distinguishing descriptive and predictive applications. *Coastal Engineering*, *135*(July 2017), 16–30. https://doi.org/10.1016/j.coastaleng.2018.01.005

Bikcora, C., Refa, N., Verheijen, L., & Weiland, S. (2016). *Prediction of Availability and Charging Rate at Charging Stations for Electric Vehicles*. 1–6.

Boulos, P. F., & Niraula, A. (2016). Optimize Operations Using Real-Time Data and Predictive Tools. *Opflow*, *42*(4), 22–24. https://doi.org/10.5991/opf.2016.42.0022

Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., & Francisco, P. (2019). *Computers & Industrial Engineering A systematic literature review of machine learning methods applied to predictive maintenance*. *137*(September).

Chodorek, R. (2005). QoS measurement and evaluation of telecommunications quality of service [Book Review]. In *IEEE Communications Magazine* (Vol. 40, Issue 2). https://doi.org/10.1109/mcom.2002.983906

Daren, G., & Paul, M. (2019). *IBM SPSS Statistics 25 Step by Step: A Simple Guide and Reference.* (15th ed.). Taylor & Francis Group. http://taylorandfrancis.com

Di Mauro, M., Longo, M., Postiglione, F., Restaino, R., & Tambasco, M. (2017). Availability Evaluation of the Virtualized Infrastructure Manager in Network Function Virtualization Environments. *Risk, Reliability and Safety: Innovating Theory and Practice - Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016, January 2020*, 423. https://doi.org/10.1201/9781315374987-393

Ercsey-Ravasz, M., Markov, N. T., Lamy, C., VanEssen, D. C., Knoblauch, K., Toroczkai, Z.,

& Kennedy, H. (2013). A Predictive Network Model of Cerebral Cortical Connectivity Based on a Distance Rule. *Neuron*, *80*(1), 184–197. https://doi.org/10.1016/j.neuron.2013.07.036

Fan, X., Wang, F., & Liu, J. (2016). Boosting Service Availability for Base Stations of Cellular Networks by Event-driven Battery Profiling. *Performance Evaluation Review*, *44*(2), 88–93. https://doi.org/10.1145/3003977.3004002

Franke, U., Johnson, P., König, J., & Marcks von Würtemberg, L. (2012). Availability of Enterprise IT Systems: An Expert-based Bayesian Framework. *Software Quality Journal*, *20*(2), 369–394. https://doi.org/10.1007/s11219-011-9141-z

Gandomi, A. H., & Roke, D. A. (2015). *Advances in Engineering Software Assessment of artificial neural network and genetic programming as predictive tools*. *88*, 63–72. https://doi.org/10.1016/j.advengsoft.2015.05.007

Garrido, L., Balderas-d, S., Guerrero-contreras, G., & Rodr, C. (2016). *A Context-Aware Architecture Supporting Service Availability in Mobile Cloud Computing*. *13*(9). https://doi.org/10.1109/TSC.2016.2540629

Greff, K., Srivastava, R. K., Koutn, J., & Steunebrink, B. R. (2017). *LSTM : A Search Space Odyssey*. 1–12. https://doi.org/10.1109/TNNLS.2016.2582924

Hilt, A. (2019). *Evolution towards Telco-Cloud : Reflections on Dimensioning, Availability and Operability*. 1–8.

Hilt, A., Járó, G., & Bakos, I. (2016). Availability Prediction of Telecommunication Application Servers Deployed on Cloud. *Periodica Polytechnica Electrical Engineering and Computer Science*, *60*(1), 72–81. https://doi.org/10.3311/PPee.9051

Hochreiter,S., and Schmidhuber, J. (1997). Long Short-Term Memory. *NeuralComputation*, 9(8), 1735 - 1730.

Hyndman, R., & Athanasopoulos, G. (2014). *Forecasting: Principles and Practice - Forecasting data and methods*. *1*, 1–5.

Ibrahimovic, S., & Bajgoric, N. (2016). Modelling Information System Availability by using the Bayesian Belief Network Approach. *Interdisciplinary Description of Complex Systems*, *14*(2), 125–138. https://doi.org/10.7906/indecs.14.2.2

International Telecommunication Union (ITU-T). (2002). *Standard G.826: End-to-end error performance parameters and objectives for international, constant bit-rate digital paths and connections*.

ISO/IEC. (2015). INTERNATIONAL STANDARD ISO / IEC Information technology — Vocabulary. *27009*, *2015*.

Jebb, A. T., Tay, L., Wang, W., Huang, Q., & Croudace, T. J. (2015). *Time series analysis for psychological research : examining and forecasting change*. *6*(June), 1–24.

https://doi.org/10.3389/fpsyg.2015.00727

Krajnović, N. (2017). Network topologies for achieving maximal all-terminal network availability. *24th Telecommunications Forum, TELFOR 2016*. https://doi.org/10.1109/TELFOR.2016.7818723

Laptev, N., Yosinski, J., Erran, L., & Slawek, L. (2017). *Time-series Extreme Event Forecasting with Neural Networks at Uber*.

Mahdi, Q. S., Hamarash, I. I., & Hassan, J. (2018). Survivability Analysis of GSM Network Systems. *Eurasian Journal of Science & Engineering*, *3*(3). https://doi.org/10.23918/eajse.v3i3p113

Mahdy, B., Abbas, H., Hassanein, H. S., Noureldin, A., & Abou-Zeid, H. (2020). A clustering-driven approach to predict the traffic load of mobile networks for the analysis of base stations deployment. *Journal of Sensor and Actuator Networks*, *9*(4). https://doi.org/10.3390/jsan9040053

Mahmood, T., & Munir, K. (2020). Enabling Predictive and Preventive Maintenance using IoT and Big Data in the Telecom Sector. *Proceedings of the 5th International Conference on Internet of Things, Big Data and Security*, 169–176. https://doi.org/10.5220/0009325201690176

Meraj, M., and Kumar, S. (2015). Evolution of Mobile Wireless Technology from 0G to 5G. *International Journal of Computer Science and Technologies*, 6(3), 2545 - 2551.

Nagy, L., Hilt, A., Jaro, G., Olah, D., & Taliyigas, Z. (2016). *Comparison of Availability Figures of Distributed Systems Using Multiple Redundancy Methods*. *June*.

Nencioni, G., Helvik, B. E., Gonzalez, A. J., Heegaard, P. E., & Kamisinski, A. (2017). *Impact of SDN Controllers Deployment on Network Availability*. http://arxiv.org/abs/1703.05595

Netes, V. (2018). End-to-end availability of cloud services. *Conference of Open Innovation Association, FRUCT*, *2018-May*, 198–203. https://doi.org/10.23919/FRUCT.2018.8468272

Odom, M. D. (1990). *A neural network model for bankruptcy prediction - Neural Networks, 1990., 1990 IJCNN International Joint Conference on*. *January 2015*. https://doi.org/10.1109/IJCNN.1990.137710

Özs, H. (2020). Monitoring Unstable Slopes in an Open Pit Lignite Mine Using ARIMA. *Journal of the Southern African Institute of Mining and Metallurgy*, *120*(3), 173-180. https://dx.doi.org/10.17159/2411-9717/665/2020

Qing, W. Q. (2017). *Introduction to Wireless Network Technology-Huawei Technologies Limited Training Manual*.

Recommendation ITU-T E.802 (2007). *Framework and methodologies for the determination* *and application of QoS parameters*.

Säe, J., & Lempiäinen, J. (2016). Maintaining Mobile Network Coverage Availability in Disturbance Scenarios. *Mobile Information Systems*, *2016*. https://doi.org/10.1155/2016/4816325

Shmueli, G., & Lichtendahl Jr, K. C. (2016). *Practical time series forecasting with r: A hands-on guide*. Axelrod Schnall Publishers.

Thulin, M. (2004). *Measuring Availability in Telecommunications Networks*. 50. https://people.kth.se/~e98_thu/thesis/Availability_Thesis_final.pdf

Usman, M. R., Usman, M. A., & Shin, S. Y. (2017). Channel blocking analysis and availability prediction in cognitive radio networks. *2017 International Conference on Computing, Networking and Communications, ICNC 2017*, 963–968. https://doi.org/10.1109/ICCNC.2017.7876264

Zhang, G. P. (2003). *Time series forecasting using a hybrid ARIMA and neural network model*. *50*, 159–175.

APPENDIX A

SPSS CODE FOR THE ARIMA – BASED MODEL

The following new variables are being created:

```
Name    Label
WEEK_   WEEK, not periodic
DAY_    DAY, period 7
DATE_   Date.  Format:  "WWWW DDD"
```

```
* Sequence Charts.
TSPLOT VARIABLES=MNOW MNOX MNOY
  MNOZ /ID=DATE
  /NOLOG.


ACF VARIABLES=MNOW
  /NOLOG
  /MXAUTO 16
  /SERROR=IND
  /PACF.


* Sequence Charts.
TSPLOT VARIABLES=MNOW
  /ID=DATE
  /NOLOG
  /DIFF=1
```

```
  /FORMAT NOFILL NOREFERENCE.

ACF VARIABLES=MNOW

  /NOLOG

  /DIFF=1

  /MXAUTO 16

  /SERROR=IND

  /PACF.


PREDICT THRU END.

* Time Series Modeler.

TSMODEL

   /MODELSUMMARY PRINT=[MODELFIT RESIDACF RESIDPACF]

   /MODELSTATISTICS DISPLAY=YES MODELFIT=[ SRSQUARE]

   /MODELDETAILS PRINT=[ PARAMETERS RESIDACF RESIDPACF] PLOT=[
RESIDACF RESIDPACF]

   /SERIESPLOT OBSERVED FORECAST FORECASTCI

   /OUTPUTFILTER DISPLAY=ALLMODELS

   /SAVE PREDICTED(Predicted)

   /AUXILIARY CILEVEL=95 MAXACFLAGS=24

   /MISSING USERMISSING=EXCLUDE

   /MODEL DEPENDENT=MNOW INDEPENDENT=DATE

      PREFIX='Model'
```

```
  /ARIMA AR=[0] DIFF=1 MA=[3,2,1] ARSEASONAL=[0] DIFFSEASONAL=0
MASEASONAL=[0]
     TRANSFORM=NONE CONSTANT=YES
  /AUTOOUTLIER DETECT=OFF.

ACF VARIABLES=MNOX
 /NOLOG
 /MXAUTO 16
 /SERROR=IND
 /PACF.

PREDICT THRU END.
* Time Series Modeler.
TSMODEL
  /MODELSUMMARY PRINT=[MODELFIT RESIDACF RESIDPACF]
  /MODELSTATISTICS DISPLAY=YES MODELFIT=[ SRSQUARE]
  /MODELDETAILS PRINT=[ PARAMETERS RESIDACF RESIDPACF FORECASTS]
PLOT=[ RESIDACF RESIDPACF]
  /SERIESPLOT OBSERVED FORECAST FORECASTCI
  /OUTPUTFILTER DISPLAY=ALLMODELS
  /SAVE PREDICTED(Predicted)
  /AUXILIARY CILEVEL=95 MAXACFLAGS=24
  /MISSING USERMISSING=EXCLUDE
  /MODEL DEPENDENT=MNOX INDEPENDENT=DATE
```

```
    PREFIX='Model'
  /ARIMA AR=[1] DIFF=0 MA=[1] ARSEASONAL=[0] DIFFSEASONAL=0
MASEASONAL=[0]
    TRANSFORM=NONE CONSTANT=YES
  /AUTOOUTLIER DETECT=OFF.

ACF VARIABLES=MNOY
 /NOLOG
 /MXAUTO 16
 /SERROR=IND
 /PACF.

PREDICT THRU END.
* Time Series Modeler.
TSMODEL
  /MODELSUMMARY PRINT=[MODELFIT RESIDACF RESIDPACF]
  /MODELSTATISTICS DISPLAY=YES MODELFIT=[ SRSQUARE]
  /MODELDETAILS PRINT=[ PARAMETERS RESIDACF RESIDPACF FORECASTS]
PLOT=[ RESIDACF RESIDPACF]
  /SERIESPLOT OBSERVED FORECAST FORECASTCI
  /OUTPUTFILTER DISPLAY=ALLMODELS
  /SAVE PREDICTED(Predicted)
  /AUXILIARY CILEVEL=95 MAXACFLAGS=24
  /MISSING USERMISSING=EXCLUDE
```

```
  /MODEL DEPENDENT=MNOY
     INDEPENDENT=DATE PREFIX='Model'
  /ARIMA AR=[2,1] DIFF=0 MA=[4,3,2,1] ARSEASONAL=[0] DIFFSEASONAL=0
MASEASONAL=[0]
     TRANSFORM=NONE CONSTANT=YES
  /AUTOOUTLIER DETECT=OFF.

ACF VARIABLES=MNOZ
 /NOLOG
 /MXAUTO 16
 /SERROR=IND
 /PACF.

* Sequence Charts.
TSPLOT VARIABLES=MNOZ
 /ID=DATE
 /NOLOG
 /DIFF=1
 /FORMAT NOFILL NOREFERENCE.

ACF VARIABLES=MNOZ
 /NOLOG
 /DIFF=1
 /MXAUTO 16
```

```
  /SERROR=IND

  /PACF.

PREDICT THRU END.

* Time Series Modeler.

TSMODEL

   /MODELSUMMARY PRINT=[MODELFIT RESIDACF RESIDPACF]

   /MODELSTATISTICS DISPLAY=YES MODELFIT=[ SRSQUARE]

   /MODELDETAILS PRINT=[ PARAMETERS RESIDACF RESIDPACF FORECASTS]

PLOT=[ RESIDACF RESIDPACF]

   /SERIESPLOT OBSERVED FORECAST FORECASTCI

   /OUTPUTFILTER DISPLAY=ALLMODELS

   /SAVE PREDICTED(Predicted)

   /AUXILIARY CILEVEL=95 MAXACFLAGS=24

   /MISSING USERMISSING=EXCLUDE

   /MODEL DEPENDENT=MNOZ INDEPENDENT=DATE

      PREFIX='Model'

   /ARIMA AR=[0] DIFF=1 MA=[1] ARSEASONAL=[0] DIFFSEASONAL=0

MASEASONAL=[0]

      TRANSFORM=NONE CONSTANT=YES

   /AUTOOUTLIER DETECT=OFF.


MEANS TABLES=MAE MAPE BY MNO
  /CELLS=MEAN COUNT STDDEV.
```

APPENDIX B

PLATES

Plate I (a) BTS3900 (An Open Indoor BTS, Vendor: Huawei)

Plate I (b) BTS3900 (Closed Indoor BTS, Vendor: Huawei)

Plate II   BSC6900 (Vendor: Huawei)

Plate III UMG8900 (Universal Media Gateway, Vendor: Huawei)

Plate IV OptiX900 (Fibre Termination Equipment, Vendor: Huawei)

Plate V Tower (For Mounting RF, Microwave Antennas and Terminating Cables)

Plate VI A Rack of Microwave Radio Indoor Units for NEs.

APPENDIX C

PUBLICATION

Ohihoin, E., Ohize, H., David, M., and Alenoghena, C. (2021, May). Base Station Availability and Telecommunication Network Quality of Service–A Review. *Proceedings of the 2020 IEEE 2nd International Conference on Cyberspace (Cyber Nigeria)*. http://repository.futminna.edu.ng:8080/jspui/handle/123456789/906