DURBAN UNIVERSITY OF TECHNOLOGY
INYUVESI YASETHEKWINI YEZOBUCHWEPHESHE

Predicting the Impact of IoT Devices on Radio Frequency Noise in South African Environments Using Machine Learning.

A Thesis submitted to the Department of Electronic and Computer Engineering in the Faculty of Engineering and the Built Environment at the Durban University of Technology in fulfilment of the academic requirements for the Doctoral Degree.

Doctor of Engineering

By

Dominique Guelord Kumamputu Ingala

20929809

[redacted]

[redacted]

Dr. N. Pillay, Supervisor — 14/03/2024 Date

[redacted]

Dr. S. Reddy, Co-Supervisor — 14/03/2024 Date

PLAGIARISM DECLARATION

I, with this, declare that the contents of this thesis are a true reflection of my work and that this thesis has not been submitted, in whole or part, for a degree to any other University or Institution.

MOTIVATIONAL THOUGHTS

You are an engineer, an analytical person. While conducting your engineering activities, you will undoubtedly meet a plethora of culminating, correlating, confusing, and contradicting thoughts. However, never forget the issue you are solving. That will lead you to practical solutions.

Dominique G.K. Ingala

In God, we trust. All others must bring data.

W. Edwards Deming

DEDICATION

To my parents, Claude Ingala and Philomene Anger.

No matter how harsh life seemed, they ensured their four children never went uneducated.

ACKNOWLEDGEMENTS

I am grateful to many people and institutions for their various contributions towards completing this research. My acknowledgement list will never be exhaustive. However, let me first recognise Mr. Stuart MacPherson, the first person with whom I discussed this research concept. His thoughts were supportive and built optimism in me to go ahead.

Since my early stage, I realized that it does not only take an excellent student to complete scientific research; it also requires a good supervisor. I thank Dr. N. Pillay for guiding me to deliver quality outcomes and to submit my work timeously. Similarly, I appreciate Dr. S. Reddy for guiding me to build a solid foundation for this research.

I recognise Mr. Peter Morris for allowing me overtime access to the state-of-the-art RF and Microwave Laboratory at the Durban University of Technology. His availability contributed as the research demanded to conduct vital measurements to evaluate, verify, validate, and characterize my RF system.

This study necessitated designing, developing, and assembling electronic hardware for IoT Radio Noise Generators. While 2019, 2020, and 2021 suffered a catastrophic worldwide shortage in the manufacturing of electronic components, Texas Instruments in South Africa supplied me with 100 samples of the much-needed radio chip CC1352 free of charge. In the same way, I am ever grateful to Arrow Altech in Durban for providing all the components required in the bill of material, also free of charge. Their kind actions raised my enthusiasm to pursue this study, as I did not want their considerate interventions to go in vain.

Data collection was another crucial step in this research. The survey campaign required me to have access to private properties where I could work safely. My notes of sincere appreciation go to Mr. Dieudonné Kabongo and Mr. Kabulo Loji for allowing me to operate my equipment at their residential premises. I must thank Inzalo Utility Systems and IDS, where I collected data for industrial candidate sites.

My parents, Claude Ingala, and Philomene Anger, and my uncles, Antoine Bulu and Lally Anger, encouraged this journey. Their regular messages and phone calls made me realize how much they believed I could achieve and complete this research. Although they could not provide engineering solutions to my problems, their experience, and constant follow-up helped me manage my tasks conveniently.

I can never omit my home-ground fans: my sweethearts, wife Benita, and daughter Nyota. They were my true backbones during this cycle. I conducted most of this research from home, especially

during the COVID-19 pandemic when the world witnessed unprecedented abnormalities. My family allowed space and time for me to focus on this journey. As they played their supporting part very well, it gave me a deep sense of responsibility that I did not wish to take their accompanying efforts for granted. When desperations tested me, their “go! go! daddy, go! go! daddy” energized me to carry on.

ABSTRACT

The increasing number of Radio Frequency (RF) devices proliferating in our environment inspired this research. Forecasts suggest that the Internet of Things (IoT) industry will populate society with an ever-increasing number of RF-operated applications such as smart homes, remote surveillance, intelligent vehicles, tracking, smart grid, remote metering, innovative health, and smart cities. This research aimed to investigate whether the presence of IoT-like radiations could influence the levels of ambient RF noise. With that in mind, the study required surveying by collecting real-world ambient radio noise data in target urban, suburban, and industrial environments over the Industrial Scientific Medical (ISM) bands, such as 433 MHz, 868 MHz, and 2.4 GHz. At the time of this survey, the IoT industry was still in the infancy stage in South Africa. Therefore, the exercise necessitated two series of survey campaigns. The first set of measurements had as its primary mission to assess the existing levels of ambient RF noise in selected candidate sites considering their early IoT development phase. Subsequently, this phase helped to verify and validate that the research deployed appropriate equipment, hardware, and software for collecting environmental radio noise data. This study designed a Radio Noise Surveying System (RNSS) using software-defined radio techniques with the Universal Software Radio Peripheral (USRP) and the GNU Radio platforms as part of the equipment. The simulation and test results agreed that the RNSS performed adequately, and that all system was suitable for radio noise surveying. This first phase also helped to confirm that the post-processing methods of importing and transforming raw data into clean data and the applied calibration techniques were correct. Exploratory data analysis with these baseline measurements revealed ambient radio noise data characteristics, for example, their extensive data volume. One of the remarkable findings was that, out of six candidate sites, the Steeve Biko Campus showed the highest levels of ambient radio noise compared to the rest, irrespective of frequency bands.

The second measurement trial, over five candidate sites, envisaged assessing the direct contribution of IoT operations. Therefore, the exercise necessitated environments populated with IoT devices. Hence, this research created IoT radio noise generators (ING) to produce intentional RF emissions in the ISM bands to imitate the presence of IoT devices in selected environments. The research underwent a complete product design cycle covering conceptualisation, component selection, schematic and PCB design, board assembly, firmware development, and functional testing. Survey campaigns deployed forty-five ING units, of which fifteen covered each of the three frequencies of interest. Data analysis exploited the elements of descriptive statistics to understand the characteristic nature of data emanating from ambient RF measurements. Concerning the central

question of this research, exploratory results revealed that 80% of analysed cases show an increase in environmental radio noise levels, with a conclusion that ambient radio noise levels were directly proportional to radio activities in given environments. This finding forewarns that the proliferating presence of IoT products will directly influence ambient radio noise levels.

Finally, this study applied Machine Learning techniques to develop linear regression models to predict the levels of ambient RF noise. The research developed a computer application as Radio Noise Predictor (RNP) software installable in Windows PC. Based on models produced in this research, the RNP application allows interested users to estimate the radio noise levels from a selected environment and frequency.

DECLARATION OF RESEARCH WORK PUBLICATION

At the time of this writing, this research produced two publications as follows:

Journal paper

- Paper title : An Assessment of Environmental RF Noise Due to IoT Deployment
- Authors : Dominique G. K. Ingala; Nelendran Pillay; Aritha Pillay
- Journal : Sensors
- Volume : 23
- Issue : 18
- Year : 2023
- DOI : 10.3390/s23187899
- URL : https://doi.org/10.3390/s23187899
- Publisher : MDPI

Conference paper

- Paper title : Baseline Measurement of the Environmental RF Noise Floor in the ISM Bands
- Authors : Dominique G. K. Ingala; Serendra Reddy
- Conference name : Information Communication Technology and Society (ICTAS)
- Organiser : Durban University of Technology
- Location : Virtual
- Year : 2022
- Website : http://www.ictas.org/ictas2022.html
- Publisher link : https://ieeexplore.ieee.org/xpl/conhome/9744636/proceeding

TABLE OF CONTENTS

LIST OF ACRONYMS

ACR Average Cross Rating

ADC Analogue-to-Digital Converter

AI Artificial Intelligence

APD Amplitude Probability Distribution

BER Bit Error Rate

BOM Bill of Material

BPF Band Pass Filter

CCIR Comité Consultatif International des Radiocommunications
(International Radio Consultative Committee)

CDF Cumulative Distribution Function

CISPR Comité International Spécial des Perturbations Radioélectriques
(International Special Committee on Radio Interference)

CPU Central Processing Unit

DAC Digital-to-Analog Converter

DDC Digital Downconverter

DOC Department of Commerce

DOD Department of Defence

DOT Department of Transport

DSP Digital Signal Processing

DUT Durban University of Technology

EMC Electromagnetic Compliance

EMI Electromagnetic Interference

EN Equipment Noise

FM Frequency Modulation

FPGA Field Programming Gate Array

GPL General Public License

GPRS General Packet Radio Services

GPS Global Positioning System

GRC GNU Radio Companion

GSM Global System for Mobile Communications

GUI Graphical User Interface

HDD Hard Disk Drive

ICASA Independent Communications Authority of South Africa

ICT Information and Communications Technologies

IDE Integrated Development Environment

IEEE Institute of Electrical and Electronics Engineers

IF Intermediate Frequency

IMD Intermodulation Distortion

ING IoT Noise Generator

IoT Internet of Things

IP Internet Protocol

IQR Interquartile Range

ISM Industrial Scientific and Medical

IT Information Technologies

ITS Institute for Telecommunication Sciences

ITU International Telecommunication Union

KDE Kernel Density Estimate

LED Light-Emitting Diode

LNA Low Noise Amplifier

LPF Low Pass Filter

LTE Long-Term Evolution

LTRS Lower Transmission Rate Side

M2M Machine-to-Machine

MAC Media Access Control

MIMO Multiple Inputs Multiple Outputs

ML Machine Learning

MSE Mean Squared Error

NAD Numerical Amplitude Distribution

NASA National Aeronautics and Space Administration

NATO North Atlantic Treaty Organization

NF Noise Figure

NOAA National Oceanic and Atmospheric Administration

NVMe Non-Volatile Memory Express

OLS Ordinary Least Squares

PAPD Pulse Amplitude Probability Distribution

PCB Printed Circuit Board

PDD Pulse Duration Distribution

PDF Probability Density Function

PSD Pulse Spacing Distribution

QPM Quasi-Peak Measurements

RAM Random Access Memory

RF Radio Frequency

RFI Radio Frequency Interference

RFID Radio Frequency Identification

RMS Root Mean Square

RNP Radio Noise Predictor

RNSS Radio Noise Surveying Systems

RX Receiver

SARL South African Radio League

SDR Software-Defined Radio

SGD Stochastic Gradient Descent

SKA Square Kilometre Array

SMA Surface Mount Assembly

SMD Surface Mount Device

SMT Surface Mount Technology

SOP Standard Operating Procedure

SSD Solid-State Device

TI Texas Instruments

UHD USRP Hardware Driver

UPS Uninterruptible Power Supply

USA United States of America

USB Universal Serial Bus

USRP Universal Software Radio Peripheral

UTRS Upper Transmission Rate Side

VHF Very High Frequency

VNA Vector Network Analyser

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

1.1. Research Problem

This research questioned whether the number of radio transmissions generated by the ubiquitous presence of Internet of Things (IoT) devices would directly influence the level of radio frequency noise in the environments. Literature shows that artificial sources of radio noise continue to multiply, for example, broadcast equipment, wireless communication, computers, engine ignitions, electrical motors, and power lines.

The fast growth of the IoT industry could pose significant risks to the industry itself when looking at radio noise. According to Spaulding [1], natural and artificial radio noise are among the factors affecting the performance of radio systems. Seeing the proliferation of RF devices and the radio spectrum becoming increasingly crowded is worrisome. Other related challenges that could affect IoT include, among others, difficulties in radio spectrum management, risk of inefficient communications in ISM bands, tight design considerations owing to interference, and unsustainable IoT business risk. According to Enge et al., the growth of unlicensed uses may not be sustainable due to the degradation of the noise floor environment [2].

IoT applications are innovative and beneficial concepts. However, this research raises awareness of the probable consequences that could be coming from the influx of radio noise associated with IoT. Pertinent questions must be asked if everything is IoT-enabled: what will the environment radio noise become with RF interference everywhere? Will IoT communication be efficient with the increase of ambient RF noise? Will the IoT applications follow the directives on radio communication acts? The outcomes of this research should interest many sectors, including academia, industry, regulatory authority, and society.

1.2. Research Aim

This research investigated the relationship between the ambient RF noise floor level versus the quantity of RF emissions generated by the IoT apparatuses at 433 MHz, 868 MHz, and 2.4 GHz. The study considered three categories of environments, including industrial, urban, and suburban. The interest in these three frequencies and selected sites was because these were the most exploited IoT scenarios. The intention was to conduct environment radio noise measurements over five selected candidate sites in Durban (South Africa) presented in Table 1.1.

In the first surveying campaign [3], the research aimed to validate the logic behind the method and verify the functionality of the hardware and software equipment selected for the experiment. The goal was to characterise the quantitative figure of radio noise in these environments at the time of

this experiment and to get insights into the nature of data produced by the measurement system. It was necessary to confirm that the system could acquire reliable data.

Table 1.1: Candidate sites in the KwaZulu Natal (KZN) region.

Candidate Sites	Environments
New Germany (NGM)	Industrial
Westmead (WSM)	Industrial
Berea Steve Biko Campus (DUT)	Urban
Morningside (MNS)	Urban
Montclair (MNC)	Suburban
Glenmore (GLM)	Suburban

In the second data collection campaign, the study aimed to emulate the IoT radio operations in selected environments to generate IoT-like radio noise at ten different transmit rates. Collected data formed the most essential asset for predictive machine learning modelling.

1.3. Research Methodology

This research planned four stages: measurement, generation, data analysis, and predictive modelling. The study developed a low-cost measurement platform called Radio Noise Surveying System (RNSS) to perform digital data acquisition for ambient radio noise. The requirement led to designing the RNSS as an RF receiver using modular components in the analogue RF front-end and the Universal Software Radio Peripheral (USRP) in the digital back-end. The GNU Radio software performed as a spectrum analyser and controller for digital data acquisition.

At the time of these experiments, the technology of IoT was still in the infant stage, specifically in South Africa. The environments were not yet exposed to the proliferation of IoT devices as predicted by the forecasts. This research had to innovate ways of generating and injecting IoT-like radio noise into the environments. It was essential to develop the IoT Noise Generators (ING) units. One of the requirements for the design of ING was compliance with the Electronic Communications Act of South Africa. The intention was to plant these ING units in the selected environments during surveying campaigns.

This study proceeded with descriptive statistical analysis to get informative insights about the recorded data. Finally, this thesis investigated linear regression algorithms to develop predictive data modelling.

1.4. Research Delimitation

Any environment contains radio noise. This research limited its scope to the industrial, urban, and suburban environments because these sites are target destinations for deploying IoT products. Insights obtained with selected candidate sites can extend to places with similar geographical and agglomerative characteristics. The official definitions of industrial and urban environments may

change with dynamics in towns, landmarks, and populations. This research adopted the nomenclature by ITU [4] and adapted it to the relevance of this thesis:

- The industrial environment was an agglomeration occupied by workshops, factories, and machinery. Specific areas like airports fall into this category.
- The urban environment was a populated residential area with many businesses, offices, and commercial activities. Central business districts (CDB) fall into this category.
- Suburban environment: populated residential area with less or no business activities and absence of CBDs.

Radio noise can appear at any frequency in the spectrum. However, this study investigated the unlicensed ISM frequency bands where most IoT applications operate: 433 MHz, 868 MHz, and 2.4 GHz. The three frequencies of interest complied with the ISM frequency ranges (433.05 – 433.79 MHz, 868.0 – 868.6 MHz, 2400.0 – 2483.5 MHz) as established by the Independent Communication Authority of South Africa (ICASA) [5].

Literature categorises radio noise as internal or external to the equipment. Radio noise emanates from natural or artificial sources. The scope of this research did not explore internal noise and naturally occurring noise. The study had an interest in external noise generated by manufactured gadgets.

While IoT could carry different meanings according to the applications, the functions, and the use of the device, this research did pay attention to the semantic definitions of IoT. The engagement was solely on the radio noise these apparatuses generated, irrespective of what the IoT device was in nature.

As many algorithms exist for data mining, artificial intelligence, machine learning, and deep learning, this research study limited itself to descriptive statistical analysis. Further data analysis applied OLS and Ridge regressions to develop predictive machine learning models.

1.5. Gaps and Key Findings

The literature review brought limited resources concerning the impact of radio noise generated by IoT devices. This finding may make sense as the IoT was still in the infancy stage in South Africa. As predicted, the full deployment of IoT could be visible in the upcoming years.

Specialised agencies usually sponsor radio noise survey campaigns to provide reputable radio noise measurement systems. Among these institutions are the National Aeronautics and Space Administration (NASA), ITU, National Oceanic and Atmospheric Administration (NOAA), Institute for Telecommunication Sciences (ITS), Department of Transport (DOT), Department of

Commerce (DOC), Department of Defence (DOD), and other high-level international institutions. Many of these measurements cover broad frequency bands for different uses. Studying the topic as academic research with low-cost and simplified tools was undoubtedly a reasonable effort.

The ITU recommendation on radio noise (ITU–R P.372) provides an artificial radio noise model based on legacy radio noise data. As one of the shortcomings of the ITU model, the statistical data originated from US measurements conducted in the early 1970s. The nature and number of possible artificial sources contributing to radio noise have probably changed considerably during these 40 years. Hence, the literature questioned the validity of the ITU–R P.372 as a referral tool for predictions of radio noise and expressed concerns about relying on too old data [6-8]. These remarks must have provoked the researcher's enthusiasm to acquire new data. Leck [9] measured ambient and noise floors in the USA between 2005 – 2006. Bradshaw performed a radio noise survey in Australia [10]. In 2016, Skeie did radio noise measurement in Norway [6]. This thesis's literature review section shows more radio noise measurement campaigns.

Concrete data on radio noise for South African environments was a scarce resource. The South African Radio League (SARL) has shown interest in monitoring radio noise. According to Groenendaal, the SARL technical committee's strategic business plan had to include the increasing level of the RF noise floor [11]. Radio amateurs suspected that the noise floor in the radio spectrum would increase given that the production volume of equipment that emits radio energy grew. However, scientific data was insufficient or unavailable to support the hypothesis. Radio astronomy drove more Radio Frequency Interference (RFI) studies, especially with candidate sites for the Square Kilometre Array (SKA) project. Dunn discussed a typical RFI measurement system for the South African SKA site [12]. It provided considerable technical details and measurement methods. However, the experiment limited their work to remote and reserved areas in compliance with the SKA requirements. Radio noise collected from remote sites may not be comparable to populated agglomerations.

With all these gaps, launching new data collection campaigns was necessary, especially in areas never exposed to radio noise measurements. It is relevant when these new campaigns focus on IoT radio noise in industrial, urban, and suburban environments.

1.6. Contributions and novelties

To the author’s knowledge, no previous research has investigated the direct influence of IoT radio noise in South African industrial, urban, and suburban environments.

The precedent projects on surveying radio noise environments utilised high-end laboratory or field instrumentation. To the author’s knowledge, this research was the first attempt to employ software-

defined radio technology to examine ambient noise levels. The method of this research's RNSS, built with SDR technology, is repeatable and replicable to alleviate the challenges brought by laboratory-grade instruments.

At this time, the IoT Industry was still going through shy development rather than predictions. As a critical contribution, the research innovated with the design from scratch, development, and production of IoT noise generators (ING) to mimic IoT devices in radio transmissions.

The measurement campaigns allowed the recording of heavy data size of about 4.95 TB. This research provides a method for data reduction by segmentation instead of sophisticated parallel routing methods, Hadoop, MapReduce, and Spark.

Another contribution was the development of the Radio Noise Predictor (RNP) software installable on Windows computers. This graphical user interface (GUI) utilises the Machine Learning models delivered in this research as background data. The user can select one of the three target environments and frequency bands, and the RNP application displays the regression lines. The models are exportable for third-party applications. The software also returns the predicted ambient radio noise levels after the user has entered a given TX Rate value. RNP is available[1] to any interested users.

1.7. Research Relevance

This research was pertinent as current data would be helpful for many sectors: science, industry, and government, to name a few.

1.7.1. Relevance to Science

NASA and DOT sponsored research to study the noise floor in selected frequency bands to understand better how it is being affected today and how it may be affected tomorrow. Their study examined whether observed noise floor results were consistent with the applicable regulations governing those bands. NASA was mainly concerned with the noise floor in bands used for crucial safety and scientific purposes [2].

NOAA runs various very sensitive and space-borne meteorological observing and remote sensing systems. The potential deployment of new systems and services within the frequency bands currently used by NOAA has generated the need to measure and collect baseline noise floor data. Collected data served to quantify future deviations in current noise floor levels [9].

It is common knowledge that radio astronomical observations are vulnerable to artificial radio RFI. For this reason, choosing a site suitable to host a new radio astronomical facility requires much

[1] Download RNP software here: https://tinyurl.com/mrxnn44f

effort. A site that today looks good in terms of RFI for radio astronomical observations may not have the same characteristics in the future. The growth and the proliferation of the commercial use of the radio spectrum are indeed unpredictable on a long-term scale. Reserved, protected, and quiet radio locations are among the characteristics of modern and sensitive radio telescopes [13].

The free-space path loss equation[2] does not account for any impairments. Because of the vagaries of wireless communications due to atmospheric and multipath deviations, a fade margin ensures link reliability during unanticipated irregularities. Since free-space path loss is not the only loss encountered, real-life path loss should consider other impairments. Attenuation from rainfall, wet snow, dense fog, water vapour, and the atmosphere alters path losses [14]. This link budget equation does not include the influence of radio noise in the environment. As discussed earlier, the quality of RF communication degrades as the background noise increases. The concept of "free space" cannot keep its definition in a noisy environment. There must be a model to characterise RF noise in the environment. This research study finds relevance as it can contribute to estimating the levels of ambient radio noise.

1.7.2. Relevance to Industry

Forecasts predict billions of operational IoT devices in the upcoming years. In many instances, new wireless technologies support applications in unlicensed bands. The use of unlicensed spectrum has fostered many innovative applications recently [2]. Depending on the region, most IoT technology operates in the ISM bands (433 MHz, 868 MHz, 915 MHz, and 2.4 GHz). Increasing ISM device production volume could raise these bands' RF background noise level.

Consequently, the operation may not satisfy good functionality amid high RFI. The IoT industry must not ignore the environmental radio noise because the future of the whole project behind the IoT concept depends on it. It would not only affect the recent technology of IoT but there could be negative impacts on existing devices sharing the same ISM bands. From an engineering perspective, the design would no longer be about operability only, but design for susceptibility would become a requirement to pass.

The radio link range, as the operational physical distance between RF devices (transmitter and receiver), has often been one of the critical specifications describing the performance of RF devices in M2M or IoT networks. In a noisy radio environment, communication over a long range becomes uncertain. As more devices radiate RF signals, the range between the devices may become compromised.

[2] $L_p = 32.4 + 20\log(f) + 20\log(d)$
L_p is the free-space path loss in dB, f is the frequency in MHz, and d is the distance in km.

1.7.3. Relevance to Government

This topic will be of interest to regulatory bodies on RF communication. The regulators must be adequately equipped with personnel and technology to deal with the challenges imposed by the rapid development of IoT devices. The civil authority must regulate the RF industry so that the products implanted within the communities comply with the relevant standards. A lot of M2M and IoT products exploit the unlicensed ISM bands. New design and development should be mindful of the integrity of environmental radio noise. Failing to regulate these bands will have detrimental consequences.

In town planning, urbanisation, and developing new cities, the relevant authorities must study the levels of ambient RF noise that IoT infrastructure can generate. Bradshaw [10] conducted a radiation survey to determine the impact on the Minto urban development in Australia. The purpose of the survey was to ensure that the levels of RF noise were within the terms of the City Council Development Control Plan.

1.8. Thesis Structure

Activities in this study followed the organisational structure of tasks shown in Figure 1.1.

The Introduction in Chapter 1 offers the research problem, aim, methodology, delimitation, gaps, and this research relevancy. The Literature Review in Chapter 2 covers the theoretical background of radio noise, including radio noise sources, mathematical and statistical considerations for analysing radio noise, and measurement systems for the environmental or ambient radio noise. This chapter also briefly explains the IoT and its future trends. Chapter 3 describes the development process of RNSS. This chapter discusses the equipment selection, simulation, and component performance tests. Chapter 4 relates to the development of ING units, including the conceptualisation, design, hardware construction, firmware coding, and performance validation steps. This thesis exploited survey campaign data in Chapter 5 for descriptive statistical analysis. Chapter 6 implements machine learning predictive models using linear regression algorithms. The Conclusion in Chapter 7 summarises the entire thesis by highlighting the findings and suggesting prospects.

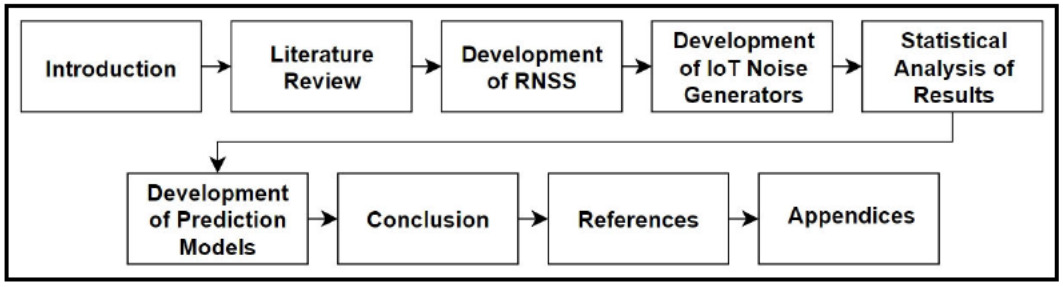

Figure 1.1: Thesis structure

CHAPTER 2: LITERATURE REVIEW

2.1. Introduction

This literature review presents an in-depth theoretical overview of radio noise, its sources, categorisation in frequency ranges according to the ITU, mathematical and statistical elements in assessing radio power levels, risks of exposure to electromagnetic fields, and measurement techniques for radio noise data collection. Concerning the IoT topic, this chapter elaborates more on predictions projected for the future of this industry.

2.2. Definition of Radio Noise

Radiation is a phenomenon by which matter transmits energy in space through waves. Electromagnetic radiation and particle radiation are the two primary forms of radiation. Electromagnetic radiation is a specific instance of radio emission using an electric field and magnetic field oscillating perpendicularly to one another. The process induces electron and proton particle acceleration to form radiations [15-17]. Radio noise, as electromagnetic radiation, is present in all terrestrial vicinities from natural and artificial sources [18].

“RF noise floor refers to the power spectral density of thermal noise plus the ambient signal background noise within a frequency band” [9]. This definition suggests that radio noise combines two variables: 1- internal thermal noise and 2- external environmental or ambient noise. As measured with receiving equipment, radio noise power level can be described in terms of noise figure as the amount of dB above thermal noise [4]. The definition shows that the study of radio noise shall relate to frequency as a parameter of interest. A given environment can exhibit different radio noise figures depending on the measurement frequency. In 2016, the International Telecommunication Union (ITU) published a referral report (ITU-R P 372-13) [19] on the background levels of radio noise at specific frequencies 0.1 Hz to 100 GHz offers another aspect of radio noise as a time-varying signal conveying no significant information and may interfere with the intentional signal adversely affecting the reception and accurate reproduction of wanted signals. Radio noise may communicate characteristic details about their sources and origins, exploitable in specific applications. In channel cohabitation situations, several wanted signals may aggregate as radio noise if they cannot be distinguishable [19, 20].

Pioneer studies showed interest in investigating environmental radio noise about 80 years ago. In his report on natural and artificial noise, Spaulding warned that overlooking the notion of radio background noise may have detrimental consequences on the performance of communication systems. Designing radio systems demands a good understanding of the "real-world electromagnetic environment” [1] in terms of parameters such as radio noise sources, noise factors,

radiation frequency, and amplitude [19]. In the same way, Leferink et al. suggested that assessing radio noise and interference is a critical step toward achieving successful wireless systems [21].

2.3. Natural Radio Noise Sources

Natural and anthropogenic or artificial sources generate radio noises. Natural noise sources form two categories: 1- terrestrial sources (confined within the Earth's ionospheric and magnetospheric cavities), and 2- extra-terrestrial sources (from cosmic and galactic spaces) [18].

2.3.1. Terrestrial Sources

Understanding Earth's geophysics is necessary in the study of natural radio noise.

2.3.1.1. From Earth's Surface to Inner Core

The study of natural radio noise sources necessitates a basic account of understanding the internal geology and geospace of the terrestrial environment. Figure 2.1 shows planet Earth's internal structure and spatial composition.

Earth, or terrestrial planet, is a natural magnetic field source, generating electric currents flowing from its inner layers [22]. Of spheric shape with an equatorial radius of 6378 km, the Earth's internal structure counts three main layers:

- The crust (40 km depth) is the outermost layer, consisting of sedimentary rocks, limestone, sandstone, granite, and basalt. Earthquakes happen within the crust. This crust includes two layers called the continental crust and oceanic crust. [23, 24];
- The mantle (2900 km depth) contains magnesium, iron, silicate rocks, and peridotite. Seismic waves carry most of the scientific details about the mantle layer. The lower layer of the mantle is called the mesosphere [23, 24].
- The core layer consists of the outer and inner cores. The outer core (from 2900 km to 5150 km depth, about 2250 km thick) includes liquid. This layer is the "seat of the Earth's magnetic field" due to the liquid iron and nickel movements. The inner core (1228 km thick) is solid and rich in iron and nickel. The Earth's interior heat drives mechanical powers responsible for earthquakes and volcanic eruptions. The temperature at the Earth's core may increase beyond 4000 °C. The natural constituent rocks of the Earth integrate radioactivity, causing heat to maintain these temperatures [23, 24].

The lithosphere and asthenosphere layers are related to volcanic activities. The lithosphere extends from Earth's surface to 280 km depth. Multiple geological plate tectonics form parts of the lithosphere. The asthenosphere is a weak layer able to move while retaining its solid state. The

lithosphere and asthenosphere layers are responsible for continental drift and seafloor spreading when convection occurs inside the asthenosphere [23, 24].

2.3.1.2. From Earth's Surface to Geospace

Geospace, which consists of the Earth's atmosphere and magnetosphere, is a solar-terrestrial region closest to Earth's planet. Solar energies affect the geospace substances which originate from Earth. Interactive phenomena occur within the geospace between terrestrial material and solar radiation, between terrestrial and solar magnetic fields, and between Earth magnetic fields and charged particles. Geospace reacts to filter or stop solar radiations like solar flares, known as localised outbursts of short duration and high intensity, streams of magnetised matter called solar wind directed towards Earth, photons, and the interplanetary magnetic field. These natural dynamics create radio waves [22].

a. Atmospheric and ionospheric sources

The atmosphere region takes off from the ground and mainly contains molecules of nitrogen, oxygen, carbon dioxide, water, hydrogen, and gases. The geo characteristics change with altitudes [22]. Variations in temperature and increase in altitude divide the atmosphere in five layers [25]:

- Troposphere: takes off from Earth's surface to 8 km in altitude. It is related to clouds and weather science. Infrared radiation and convection from the ground contribute to troposphere heat. The temperature reaches -55 °C. The upper troposphere develops an intermediate layer called tropopause, where the temperature falls to -80 °C approximately.
- Stratosphere: 50 km altitude, the temperature increases to 0 °C maximum. The ozone layer consists of heated particles after the absorption of solar ultraviolet energy. The stratopause is the medium at the upper part of the stratosphere.
- Mesosphere: 85 km in altitude, falls below -90 °C in temperature. The mesopause layer develops as the upper layer of the mesosphere.
- Thermosphere: from 150 km in altitude can reach 500 °C, and 1300 °C at 500 km altitude due to solar ultraviolet radiations.
- Exosphere: extends beyond 500 km to form the boundary medium with the lower magnetosphere layer.

The atmospheric iris window, also exploited in radio astronomy, allows radio wave penetration (10 MHz to 300 GHz) without significant attenuation or reflection [23].

Thermodynamic activities form the thermosphere due to solar emissions (ultra-violet and X-ray) penetrating the atmosphere and initiating ionisation. The ultraviolet and X-ray emissions are parts

of the Sun's radiations which influence the ionisation levels. The Institute of Radio Engineers formally adopted the term ionosphere in 1950. The definition suggests the ionosphere as an electrically conducting medium in the upper atmosphere where the quantity of ionised particles is significant enough to interfere with the propagation of radio waves. The electrical conductivity depends on the content of charged particles in the ionosphere. Interactions of ionised particles generate various radio waves [22, 25, 26]. The ionosphere is a vital medium in radio science and essential to the quality of radio communications [27].

The ionosphere consists of the D, E, and F layers, which experience several ionisation sources. The lower ionospheric region is the D layer, which exists due to Lyman ionising the nitric oxygen. Extending 80 km altitude, region D reflects and attenuates radio waves during the day. It disappears at nighttime. Formed by X-ray and far ultra-violets ionisation, Region E extends to 110 km altitude and reflects radio waves at night and daytime. Formed by extreme solar ultraviolet ionising atomic oxygen, the F region splits into F_1 (150 km) and F_2 (300 km) layers during the day and merges into a single layer at night. The ionosphere affects low-frequency radio waves by attenuation, reflection, and refraction. The level of solar activity, season, and time of day determines the degree of ionisation and seasonal variations. The geographical position also contributes to the ionosphere medium's state because solar flux varies with geographical latitude [25, 26].

The Earth's atmosphere experiences electrical disturbances causing atmospheric noise. The ionosphere is the source of most of the atmospheric noise. Lightning discharges are the most visible source of static. Lightning's random nature spreads spurious radio signals across the extremely low frequency (ELF) to very high frequency (VHF) ranges. At the higher frequencies, the ionosphere is more transparent and more penetrable. Ionospheric waves are more powerful at the very low frequency (VLF) and High frequency (HV) ranges. Worldwide, Earth's surface receives about 100 lightning strikes every second, with discharge capable of reaching ten kA in less than a second, equivalent to the energy of tenths of GJ and 10 GW of power. The short-duration strikes indicate the impulsive characteristic of atmospheric noise. The atmosphere contains an electric field directed radially as the Earth carries negative charges. In fair weather, the magnitude of the atmospheric electric field can reach 100 V/m. Depending on climate, altitude, and time of the day, it may vary from 50 V/m to 500 V/m and about 40 kV/m during precipitation [18, 28].

b. Magnetic sources

From the Earth's internal core, geomagnetic fields propagate a thousand kilometres outwards to interact with magnetised solar wind-charged particles and develop a magnetic cavity called

magnetosphere in the interplanetary geospace. The magnetosphere is an obstacle against emissions from the Sun because of its magnetic properties, which oppose solar wind flow. The size of the magnetosphere continuously varies with the changes by the solar wind. The Earth is comparable to a giant magnetic dipole at the central core. The very high temperature, thousands of degrees, at the Earth's inner core does not allow the existence of a magnet. The heat convection dynamics of iron-rich metallic fluid create the Earth's magnetic field in the outer core. The generated magnetic fields multiply with induced electrical current due to Earth's rotation. 10 Terra-Watts (10^{13} W) is required to produce Earth's magnetic field [22, 26, 29].

The magnetosphere is the medium above the ionosphere dominated by Earth's magnetic fields, which interact with the movement of ionised particles. Modifying electric current after ionised components' motion generates radio waves [22]. The natural magnetic noise develops within the magnetosphere from the ultra-low frequency (ULF) to extremely low frequency (ELF) ranges. Their amplitudes vary from tenths of nT to hundreds and thousands of nT (in the case of magnetic storms). Magnetic noise sources include geomagnetic pulsations, chorus emissions, auroral hiss, and whistlers. A low-frequency receiver with audio features can detect and listen to magnetic chirps and hiss-like tones since their emissions coincide with the audible frequencies [18]. Natural magnetic noise is exploited in the navigation systems [29]. The Earth's planet experiences several celestial bombardments, and disturbances characterised as magnetic and ionospheric storms. The magnetosphere is vital to shield the planet from harmful perturbations in the geospace environment [26].

2.3.2. Extra-terrestrial Sources

The Sun and all other stars in the universe are responsible for space noise. A solar flare is a violent storm on the Sun's surface. The amount of radio noise emitted by the Sun is called solar radio noise flux, or simply solar flux, an indicator of general solar activity. The solar flux varies with frequency. However, the solar flux level increases abruptly at any frequency when a solar flare occurs. A sudden increase in the solar flux indicates that ionospheric propagation conditions will deteriorate within a few hours.

Celestial radio sources or distant stars radiate noise as the Sun does. The noise generated by the stars is termed cosmic noise. Although the stars are more distant from the Earth than our Sun, their vast numbers and additive effect contribute to space noise about equal to that of the Sun. Cosmic noise occurs at all wavelengths, from the very low RF band to the X-ray band and above [19, 20]. The study in [30] explored more extra-galactic radio sources.

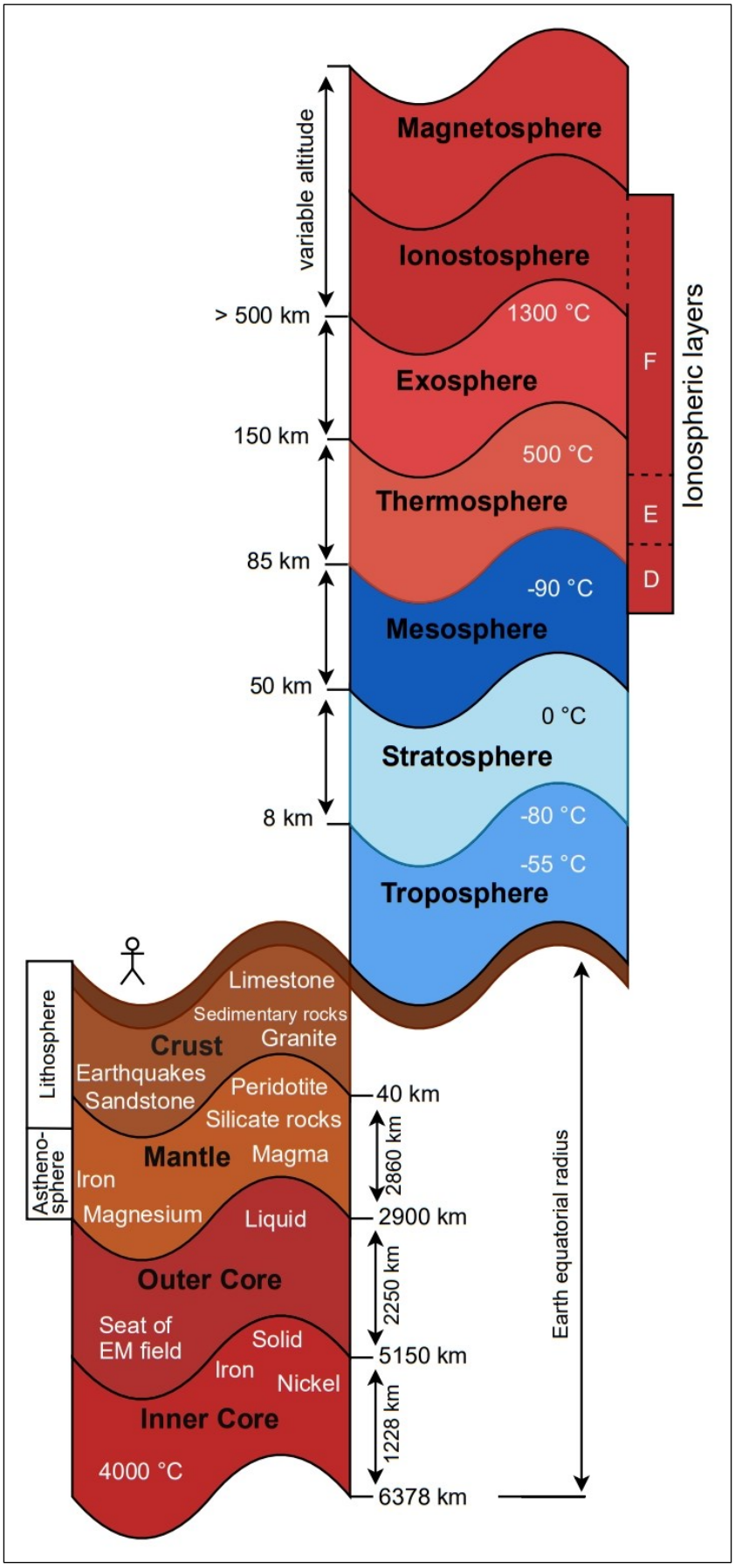

Figure 2.1: Unscaled representation of Earth geology and geospace

2.4. Man-made Radio Noise

A North Atlantic Treaty Organization (NATO) report indicated that RF noise may occur naturally and artificially. The former includes, among others, lightning, atmospheric disturbances, galactic sources, and solar flares; these are all highly random phenomena. The latter includes, among others, broadcast, communication, radar and navigation transmitters, electrical motors and generators, fluorescent and neon lighting, arc welders and RF induction heaters [31]. The recommendations from ITU give more details about radio noise due to lightning, artificial sources, the galaxy and the temperature of the lower atmosphere [19].

2.4.1. Man-made Ambient Noise

Artificial noise is any form of electromagnetic interference originating from any non-natural causes. The primary sources of artificial noise refer to interference, such as ignition and impulse noise originating from combustion engines, automotive ignitions, generation facilities, industrial equipment, electrical appliances, power transmission lines, lighting systems, consumer products, and medical equipment. It has become an increasing problem in radio communications because of the proliferation of household appliances that generate electromagnetic energy [1, 20].

Electrical and mechanical machinery, highly reflective industrial facilities and co-existing communication systems are the significant sources of disturbances in industrial and factory environments. Electromechanical machinery is the primary source of impulsive interference for wireless systems. This type of interference makes it difficult to predict the emission pattern without measurements. A measurement system can capture the ambient signals emitted by broadcast radio and TV stations, cellular telephone and other wireless systems, and from unintentional emitters such as inverters, electric motors, welders, vehicles, leaking antenna cables, robots, microwave ovens, cranes, arcs, freight trains, elevators and gasoline generators. Modern transport mopeds and four-wheeled motorcycles in the industrial environment introduce disturbances in the existing wireless communication systems. Repairing work in an industrial environment with welding or similar processes also generates disturbances [32]. Typical sources of substantial radiated electromagnetic interference in industrial environments are electrical engines, power converters, charging devices for battery-driven equipment, frequency converters, welding processes using pulsed power, other wireless systems, and personal computers [33].

2.4.2. IoT Radio Noise

2.4.2.1. Background about IoT

There are several ways of describing the concept of IoT. Atzori et al. describe IoT semantically as a worldwide network of interconnected objects uniquely addressable, based on standard communication protocols [34]. Another definition provided by Qin et al. is that in contrast to the

traditional internet, where the main focus is on computers, IoT aims to connect everyday objects such as coats, shoes, watches, home appliances, cars, plants, animals and other things to the internet to enable interaction between these objects [35]. Hatton et al. define all connected devices as machine-to-machine (M2M); however, IoT is a broader concept incorporating application development, data sharing, analytics, and a whole ecosystem. Although connected devices are M2M devices, they form parts of IoT [36]. Although these references provided holistic meanings of IoT, this research ignored the payload and focused only on radio noise generated by IoT devices.

A study on artificial noise floors has shown that using unlicensed spectrum has fostered many innovative applications in recent years [2]. New wireless technologies in the IoT industry often support applications in unlicensed bands. In South Africa, the Independent Communications Authority of South Africa (ICASA) has reserved, among others, licence exemptions for the 433 MHz, 868 MHz and 2.4 GHz frequency bands [5]; these are the three most common global ISM bands [33].

The short-range IoT primarily consists of devices connected by unlicensed radio technologies, with a typical range of up to 100 meters, such as Wi-Fi, Bluetooth, and ZigBee. The wide-area IoT consists of devices using cellular connections and unlicensed low-power technologies, such as Sigfox and LoRa [37].

Years ago, equipment operating in the ISM bands did not have significant harmonics in the frequency up to 1 GHz [1]. However, the current environment has seen much development in technology. The spectrum will be even more saturated when IoT sees full capacity deployment. It motivated this research to collect new data.

2.4.2.2. IoT Forecast

Several studies perceived that the number of IoT and wireless connections will significantly increase. The Ericsson Mobility Reports have published insights that clearly illustrate the tremendous evolution of mobile technology and the increase in mobile handhelds, base stations, mobile subscriptions, and data traffic. Heuveldop, in the 2017 Ericsson Mobility Report, predicted that the number of connected devices will reach 29 billion by 2022, of which 18 billion will be related to IoT (connected cars, machines, meters, sensors, point-of-sales terminals, consumer electronics, Smart TVs, and wearables) as shown in Figure 2.2. Figure 2.3 shows global critical figures by the end of 2022: five billion LTE subscriptions and more than 500 million 5G subscriptions [37]. As shown in Figure 2.4, the 2019 Ericsson Mobility Report predicts a massive development in cellular IoT connections in the 4G, LTE, NB-IoT, Cat-M, and 5G bands: 5.4 billion

wide-area IoT, 5.0 billion Cellular IoT, and 19.5 billion short-range IoT connections. 2.6 billion 5G subscriptions are expected by the end of 2025 [38].

Another research by Hatton et al. predicted that by 2024, the total number of connected devices will have grown to 39 billion, of which the majority (69%) will be in the category of new devices, which exclude handsets, tablets and PCs [36]. Frost and Sullivan also foresee that there will be about ten connected devices for every human on the planet [39].

These predictions expected the industry of IoT to grow exponentially in the next few years. How crowded will the spectrum become? What will the impacts of this spectrum saturation be? These thought-provoking questions raised the problem motivating this research.

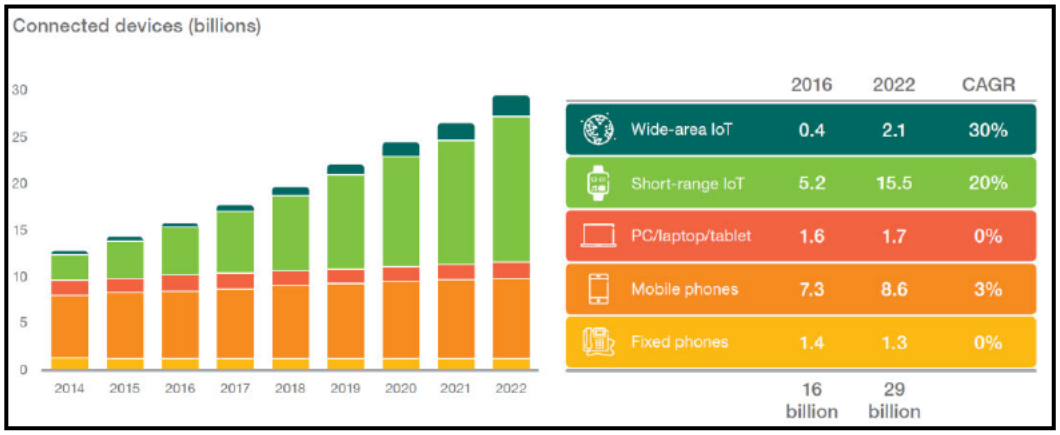

Figure 2.2: Number of connected devices by the end of 2022 [37].

	2015	2016	2022 forecast	CAGR** 2016–2022	Unit
Mobile subscriptions					
Worldwide mobile subscriptions	7,260	7,520	8,980	3%	million
> Smartphone subscriptions	3,280	3,860	6,830	10%	million
> Mobile PC, tablet and mobile router subscriptions	240	240	320	5%	million
> Mobile broadband subscriptions	3,530	4,390	8,280	11%	million
> Mobile subscriptions, GSM/EDGE-only	3,600	3,050	670	-22%	million
> Mobile subscriptions, WCDMA/HSPA	2,080	2,280	2,780	3%	million
> Mobile subscriptions, LTE	1,090	1,860	4,960	18%	million
> Mobile subscriptions, 5G			530		million
Mobile traffic*					
> Data traffic per smartphone	1.4	2.1	12	33%	GB/month
> Data traffic per mobile PC	5.8	7.7	23	20%	GB/month
> Data traffic per tablet	2.5	3.6	11	20%	GB/month
Total traffic					
Total mobile data traffic	5.3	8.8	71	42%	EB/month
> Smartphones	4.1	7.2	66	45%	EB/month
> Mobile PCs	0.4	0.5	1.3	17%	EB/month
> Tablets	0.2	0.3	1.5	30%	EB/month
> Total fixed data traffic	60	70	170	15%	EB/month

Figure 2.3: Global key figures by the end of 2022 [37].

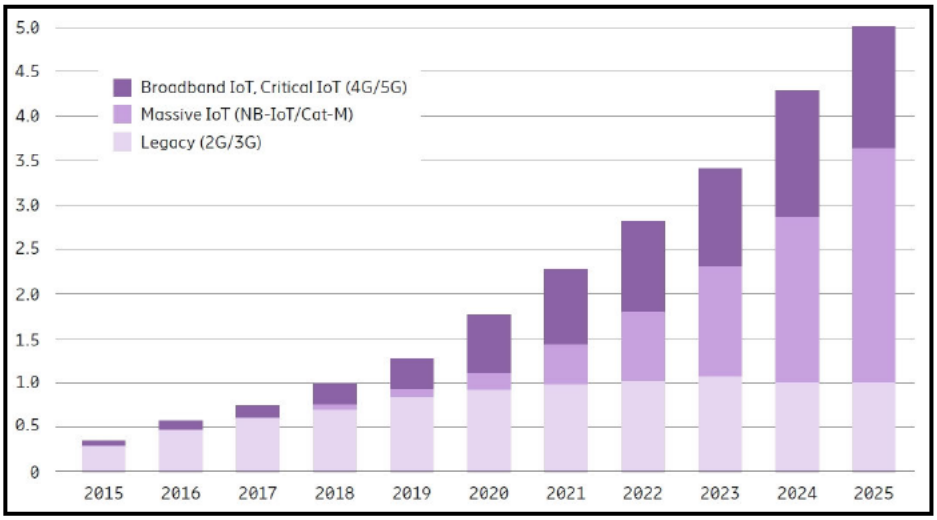

Figure 2.4: Billions of IoT Connections by the end of 2025 [38].

2.4.3. Equipment Internal Noise

Random noises created by the passive or active devices inside the equipment are distributed equally over the entire RF spectrum. The power of this random noise, termed internal noise, is proportional to the bandwidth. Internal noise has two categories: thermal and shot noise [20]. Phase noise has become essential in designing and developing systems for communications [40].

2.4.3.1. Thermal Noise

Any resistance or the resistive component generates thermal noise. Noise, termed thermal, thermal agitation, white, or Johnson noise, emanates from the rapid and random motion of the atoms and electrons in a resistor. The power of this random noise is related directly to bandwidth and the absolute temperature of the resistance. Noise power can be calculated in Equation (2.1) [20].

$$P_n = kTB \tag{2.1}$$

Where P_n is expressed in Watt.

k: Boltzmann's constant: 1.38×10^{-23} J/°K

T: absolute resistance temperature in degrees Kelvin

B: bandwidth in Hertz.

Even when a resistor is not in a circuit with a voltage source, the internal movement of electrons generates a random RMS voltage across the resistor. Equation (2.2) expresses the value of the RMS noise voltage.

$$e_n = \sqrt{4kTBR} \tag{2.2}$$

With e_n in Volt and R, the resistor generating noise in Ohm.

2.4.3.2. Shot Noise

All active devices provoke shot noise. It is present in all amplifying devices due to the shot effect caused by random variations of most current carriers (electrons or holes) at the output terminal of an amplifying device. The term shot noise relates to the excessive randomly varying noise currents.

Both shot and thermal noise are present in amplifiers and are, therefore, additive. However, since shot noise is a current and thermal noise is a voltage, it is impossible to add them together in calculations. Hence, equations for shot noise are usually approximations. For simplicity, the manufacturer's datasheets specify an equivalent input noise resistance (R_{eq}), which has a resistance value that will produce a noise level equivalent to that produced by the shot noise of an amplifying device. In calculations, R_{eq} is treated like any other noise-generating resistance, operating at the same temperature and electrically connected in series with the input terminal of the amplifying device [20].

2.4.3.3. Phase Noise

Phase noise is a topic of theoretical and practical interest in oscillators used in RF and digital electronic systems that require a clock signal for time reference. Phase Noise is of significant concern in oscillators because introducing even a little noise into an oscillator leads to dramatic changes in its frequency spectrum and timing properties. This phenomenon, peculiar to oscillators, is known as phase noise or timing jitter [41]. In the frequency domain phase, noise presents itself as noise sidebands. In the time domain, phase noise is called a jitter on a signal. As an effect, phase noise reduces RF receiver sensitivity in the presence of strong signals. It reduces the full-scale dynamic range in the ADC clock with a jitter. It increases the bit error rates (BER) in phase-modulated systems in digital communications [42].

A perfect frequency source generates only one output signal without instability in its output frequency or level. Both the output level and frequency are measurable to great precision. In reality, however, all signal sources exhibit instability in their instantaneous output frequency and level [40].

Without suitable, low-phase noise oscillators in systems, the interference levels experienced by users can become unacceptable. Improvements in these systems are often contingent upon improved oscillators, which can only be quantified through good measurements of phase noise [40].

Phase noise is the most generic method of expressing frequency instability. The carrier frequency instability is expressed by deriving the average carrier frequency and measuring the power at various offsets from the carrier frequency in a defined bandwidth. The result is a logarithmic ratio

compared to the total carrier power. Normalising the power ratio to be the equivalent signal power present in a measurement bandwidth of 1 Hz is essential. Phase noise measurements are inherently ratio-metric. In principle, the absolute level of the carrier signal to be measured is irrelevant. The measured performance of the source should not change with the level it is measured at unless the source operating conditions are changed (the addition of an attenuator pad, for example, should not fundamentally change the performance) [40].

Several different factors are responsible for phase noise on a source [40]:

- Noise in the semiconductor devices that generate the signal
- Noise from power supplies converted to a phase deviation by non-linear processes
- Varactor diodes internal noise (or other frequency tuning devices)
- Noise on the tuning lines to frequency tuning elements
- Additive noise from amplifier systems, including white noise
- Phase-locked loops in synthesised sources

Phase noise can be measured using one of these five measurement techniques:

- Spectrum Analysers
- Delay Line Discriminators
- Quadrature Technique
- FM Discriminators
- Direct Digital Measurement

2.5. Electromagnetic Spectrum

Electromagnetic waves are composed of electric fields and magnetic fields oscillating perpendicularly to each other. Frequency measured in Hertz (Hz), defined as the number of oscillations per second, is an essential electromagnetic parameter. Equation (2.3) describes the relation between frequency and wavelength λ in free space. Measured in meters (m), the wavelength is the distance separating two oscillation peaks [17].

$$\lambda = \frac{c}{f} \tag{2.3}$$

with $c \approx 3\times10^8$ m/s being the speed of light.

The electromagnetic spectrum comprises ionising and non-ionising radiations. The former is energetically strong, > 100 electron-Volts (eV), and can break electrons from atomic structure. The ionising radiations include the X-ray and gamma (γ) ray wavelengths. Ionising radiation can lead to electronic excitation, transition, vibration, and rotation and activate the ionisation process in

biological materials. Consequently, the electron is more reactive as it is less bound to the nucleus. Equation (2.4), the Planck – Einstein relation, is the mathematical expression describing photon energy (E) as the electromagnetic energy radiated by hot bodies. Due to high energy, the ionisation process causes the electron to detach from the coulomb attraction of the nucleus. It is critical to monitor exposure to ionising radiation. In contrast, radiations from the non-ionising portion do not extract electrons due to their weak energy. They experience negligible effects on the stability of the attraction between the nucleus and electron. The non-ionising process can provoke molecular excitation, but the energy cannot expel the electron from the atomic nucleus. The non-ionising radiations (NIR) include electromagnetic waves (up to 300 GHz) and optical spectrum (300 GHz to 3 PHz). The optical spectrum counts the infrared, visible light, and ultraviolet waves. Acoustic sound (20 Hz to 20 kHz) and ultrasound waves (> 20 kHz) are also parts of NIR. Medical low-resolution diagnostic sonography or ultrasound echography machines operate between 2 MHz and 12 MHz range and offer deeper skin penetration. The high-resolution machines have poor skin penetration but extend about 15 to 20 MHz. The mutual dependence between the electrical field (E) and magnetic field (H) ceases to be below the extremely low frequency (ELF). 50 Hz and 60 Hz are frequencies utilised for electricity distribution [15, 17, 43-47].

$$E = hf \tag{2.4}$$

Where h = 6.626x10^{-34} J.Hz^{-1} is the Planck's constant; f is the radiation frequency in Hertz (Hz); and photon energy E measured in Joule (J) or electron-volt (eV), with 1 J = 6.244x10^{18} eV.

With this wealth of knowledge, this thesis has established in Figure 2.5 a more inclusive electromagnetic spectrum from ultra-low frequency (ULF) to cosmic waves. Natural magnetospheric and atmospheric radiations [18, 28]. The RF electromagnetic spectrum, used for radiocommunications, extends from 3 kHz to 300 GHz and is subdivided by the International Telecommunications Union (ITU) as in Table 2.1. The subset from 300 MHz to 300 GHz is called microwave frequencies [17, 48]. As Sanad [47] explained, the actual frequency or wavelength of electromagnetic radiations may vary depending on the physical parameters of particles (nucleus, atom or molecule) such as the mass of electron, nucleon or molecule in kg, atomic or nuclear or molecular distance in m.

Table 2.1: Radiofrequency electromagnetic spectrum

Band	**Frequency**	**Wavelength (m)**	**Characteristics**
Very low frequency (VLF)	3 kHz – 30 kHz	$10^5 – 10^4$	
Low frequency (LF)	30 kHz – 300 kHz	$10^4 – 10^3$	Kilometric waves
Medium frequency (MF)	300 kHz – 3 MHz	$10^3 – 10^2$	Hectometric waves
High frequency (HF)	3 MHz – 30 MHz	$10^2 – 10^1$	Decametric waves
Very high frequency (VHF)	30 MHz – 300 MHz	10 – 1	Metric waves
Ultra-high frequency (UHF)	300 MHz – 3 GHz	$1 – 10^{-1}$	Centimetric waves
Super high frequency (SHF)	3 GHz – 30 GHz	$10^{-1} – 10^{-2}$	Decimetric waves
Extremely high frequency (EHF)	30 GHz – 300 GHz	$10^{-2} – 10^{-3}$	Millimetric waves

Table 2.2 represents a dedicated naming convention employed exclusively for satellite frequency bands [48].

Table 2.2: Satellite band names

Band name	**Frequency**	**Application**
L	1 GHz – 2 GHz	Global Positioning System (GPS) , Satellite mobile phones (Iridium, Inmarsat), WorldSpace satellite radio
S	2 GHz – 4 GHz	Weather radar, surface ship radar, NASA communication with International Space Station (ISS), Space Shuttle.
C	4 GHz – 8 GHz	Satellite TV network, Telstar satellite
X	8 GHz – 12 GHz	Military, civil, government services. Defense tracking, vehicle speed detection, air traffic control, weather monitoring, maritime vessel traffic control
Ku	12 GHz – 18 GHz	Direct broadcast services, Astra
K	18 GHz – 26 GHz	Unused
Ka	26 GHz – 40 GHz	Military aircraft, close range targeting radar.

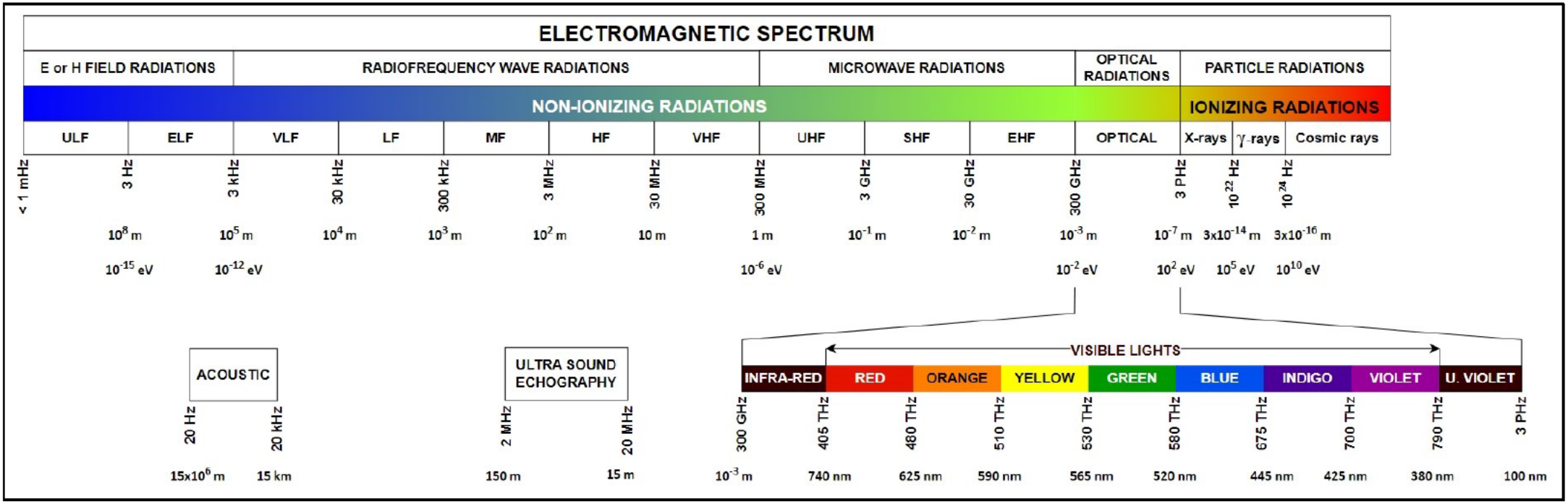

Figure 2.5: Electromagnetic spectrum

2.6. Exposimetry and Dosimetry

Physical materials respond when exposed to electromagnetic radiation due to interactions between electromagnetic fields and the atoms and molecules. Power transferred from radiation source to media, measured in Watt (W), is equivalent to energy expressed in Joule per second (J/s). The frequency of radiation determines how each matter reacts to EMF exposure. In free space and air, electromagnetic waves propagate at 3×10^8 m/s (speed of light) but slow down in dielectric materials and biological tissues. Reflection and absorption of EMF energy occur when biological bodies respond to EMF impacts. The interaction induces currents within the tissues, and the intensity of internal currents depends on the frequency (wavelength) of EMF, material properties and the body size [16, 17]. Dosimetric mechanisms operate in three frequency ranges: 30 kHz to 10 MHz for body size larger than the wavelength, 10 MHz to 10 GHz for body size comparable to the wavelength, and 10 GHz to 300 GHz for body size much larger than the wavelength. EMF consists of electrical and magnetic components. The electrical field is the main component that affects the exposed body [17].

Tissue heating is one of the confirmed adverse effects of radio exposure. The ICNIRP considers for thresholds temperature rise of 1 °C average over the whole body, five °C on localised Type-1 regions (upper arm, forearm, hand, thigh, leg, foot, pinna, cornea, anterior chamber, eye iris, epidermal, dermal, fat, muscle, and bone), and two °C on Type-2 regions (head, eye, abdomen, back, thorax, and pelvis). The specific energy absorption rate (SAR), defined as the variation in time of energy consumption by heat, represents a vital dosimetry quantity. Equation (2.5), which relates the temperature change to SAR, only applies to brief exposure cases of insignificant heat loss [16].

$$SAR = C\frac{dT}{dt} \tag{2.5}$$

With SAR expressed in W/kg, C is tissue heat capacity expressed in Joule per kilogram and degree Celsius (J kg^{-1} $°C^{-1}$), T is temperature (°C), and t is exposure duration in seconds.

Equation (2.6) describes SAR as the quantity of EMF power absorbed dW by biological bodies per unit of mass dm, contained in a volume dV of given density φ [16].

$$SAR = \frac{d}{dt}\left(\frac{dW}{dm}\right) = \frac{d}{dt}\left(\frac{dW}{\varphi dV}\right) \tag{2.6}$$

Biological tissues are dielectric lossy matter transparent to the magnetic field (relative magnetic permeability, $\mu_r = 1$). Equation (2.7) describes SAR as proportional to the rms value of the electrical field E [16].

$$SAR = \frac{\sigma|E|^2}{\varphi} \quad (2.7)$$

Where:

σ is the conductivity of the tissue (S/m)

φ is the mass density of the tissue (kg/m^3)

E is rms electric field strength in tissue (V/m)

2.7. Risks of Exposure to Radio Noise

Radiofrequency science and technology produced countless advantageous applications. However, scientists have been investing in the risks associated with exposure to electromagnetic fields (EMF) or radio noise to plant and animal ecosystems and human health. These three components interact essentially to maintain a healthy ecological system.

2.7.1. Impacts on Plants

Most studies suggest that RF exposure affects plant development in metabolism, stress-alleviation, photosynthesis, and physiology. The study results relied on factors including frequency, duration of exposure, plant type, and test greenhouse conditions. Studies suspected abnormalities in numerous plants such as tomato, mung beans, onion, celery, basil, sunflowers, and lettuce. Duckweed developed slow growth and morphological abnormalities due to RF exposure within 156 – 162 MHz. Other plants, such as mung beans, bush rose, and radish, experienced growth reduction. A minority of studies found no correlation between RF exposure and plant ecosystems [49].

The methods employed by certain previous studies may have resulted in inconclusive assessments: laboratory-controlled conditions instead of real-world environments, short-term evaluations often when the species were still in the seed germination phase, ignoring the responses during plant maturity and lifecycle, and minimal sample size under assessment. On four-month experimentation, Czerwinski et al. concluded that only specific plants exhibit responses to RF exposure in the 866 – 868 MHz range. Six species showed no observable modifications out of ten types of herbaceous grasses, forbs and legumes. Four species varied in seedling emergence rate, height, and leaf discolouration. The Trifolium Arvense (Hare's-Foot Clover) was the unique plant to develop the most significant long-term effects over RF exposure, displaying discoloured patterns on the leaves and modified leaf orientation [50]. Certain levels of RF exposure introduce dielectric heating in plants, causing effects resulting in plant mortality. Low power levels provided inconclusive results [51].

2.7.2. Impacts on Animals

Several studies delivered contradictory results, making it hard to formulate a generalised conclusion on the impacts of RF exposure on biological organisms. Using standards may allow researchers to achieve repetitions of observations [52].

The European Parliamentary Research Services (EPRS) requested to investigate the effects of exposure to radio-frequency electromagnetic fields emitted within the 5G network frequencies on biological and ecological species. The hypothesis affirmed that RF radiations could penetrate biological tissues and influence dielectric heating, genetic processes, and behavioural responses. The results indicated a thermal effects relationship. All categories demonstrated dielectric heating due to RF-EMF exposure. For instance, an increase in internal temperature in organisms may cause modifications of thermoregulation responses [51]. Table 2.3 summarises the other observations.

Table 2.3: EPRS summary of observations

	450 MHz – 6 GHz		6 GHz – 300 GHz	
Animal	In vivo	In vitro	In vivo	In vitro
Vertebrate	Contradictory results among different studies. Literature unconclusive.	Uncorrelated results due to inadequate evidence or unconclusive analysis.	Observation of eye corneal lesions and cataract. Rodent male fertility was affected.	Neural effects demonstrated. Evidence that a measured dose of high frequency exposure could induce anti-inflammatory response.
Invertebrate	Inconclusive results.	Increased neural effects demonstrated.		Neural effects demonstrated.

2.7.3. Impacts on Human Health

In 2011, the International Agency for Research on Cancer (IARC) evaluated that RF Electromagnetic fields were "possibly carcinogenic to humans", categorised as Group 2B. The same report underlines that there is "limited evidence" in humans and animals for the carcinogenicity of RF radiation. Studies on reported association cases presented severe lacunas: mixed epidemiological evidence, uninformative studies, exposure misclassification, inconsistencies between results, and limited evidence. In conclusion, the current evidence in humans was "inadequate" [17].

In 2020, the International Commission on Non-Ionizing Radiation Protection (ICNIRP) published updated guidelines [16] for limiting human exposure to EMF up to 300 GHz. This report emphasised that high RF EMF power levels beyond the recommended limits can adversely affect

human health. Observing the recommended limits and exposure duration is essential to ensure short-term and long-term safety. Not all people have common body reactions to EMF. For example, metallic implants, as a conductive material, may influence the body's response to EMF. The ICNIRP relied on recent available scientific knowledge, including independent reviews by the World Health Organization (WHO), the Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR), the Swedish Radiation Safety Authority (SSM), to assess a wide range of health-related risks: foetal development, brain physiology, audition, vestibular and ocular systems, endocrine, cardiovascular, immune system, haematology, fertility, reproduction, childhood development, and cancer. The ICNIRP concluded that "the only substantiated adverse health effects were nerve stimulation, modification in the permeability of cell membranes, and effects due to temperature elevation. There was no evidence of adverse health effects at exposure levels below the ICNIRP restriction levels, and no evidence of an interaction mechanism that would predict that adverse health effects could occur due to radiofrequency EMF exposure below those restriction levels" [16].

Human life is exposed to EMF regularly. High-voltage powerlines, power substations, train and tram powerlines, welding equipment, induction heaters, and MRI machines are some sources of extremely low frequency (ELF, 1 Hz – 100 kHz) radiations. Exposure to this category of EMF can cause short-term side effects, including nausea, dizziness, muscle contractions, body heating, nerve stimulation, and injury hazards from flying metallic projectiles. There have been reports of health-related risks such as leukaemia issues for children living near powerlines, Parkinson's as long-term effects on people exposed to workplace ELF radiations, indications of headache, painful muscles, and dizziness in people at residential places. However, no causal effect has been proven [53].

2.8. Noise Measurement

Ambient radio noise measurement is feasible. The ITU report (ITU-R SM.1753-2) [4] provides guidelines for radio noise surveying. Several studies appeared to have referred to these ITU design recommendations. Bradshaw conducted an electromagnetic radiation survey for a city renewal project in Australia [10]. Wepman et al. performed a comprehensive radio noise survey describing the technical parameters, artificial noise models, and data analysis [7]. In Norway, Skeie measured external artificial radio noise in city, residential, and rural areas [6]. Instead of a fixed location survey, Haedrich et al. embarked on a walk-by mobile measurement campaign to produce street-scale radio noise mapping in the USA (Boston and Massachusetts). Measurement equipment was transported in a backpack [54]. Breton et al. performed a similar study of mobile surveys. It was instead a cycling exercise where the RF instrumentation was towed by a bicycle [55].

2.8.1. Equipment requirements

A radio noise measurement system should be a transportable receiver ready for use in many locations [4]. Most radio noise survey systems use an antenna, RF filter, low-noise amplifier (LNA), and radio receiver. Additional components, such as RF switches, can be optional depending on the project requirements.

2.8.1.1. Antenna

In the report (ITU-R P.372) [19], their study employed a grounded short vertical monopole reference antenna with omnidirectional and lossless characteristics. The same antenna is required to achieve comparable results with the ITU artificial noise models. However, this is not always applicable in all cases. Different types of antennas can be exploited [6]. Radio noise arrives at the receive antenna from different directions. The receiving antenna with omnidirectional property is appropriate for radio noise measurement. A quarter-wave short monopole with a ground plane can be selected. A single wideband antenna is useful when targeting multiple observatory frequencies as it makes it easier for result analysis. The typical value of the voltage standing wave ratio (VSWR) obtained from antenna frequency response can be 1.4 maximum [7]. It is required to electrically match the monopole antenna to a standard 50 Ω impedance [4]. A low VSWR indicates that the antenna meets the requirement for impedance matching. When streams of radio noise originate from all angles, it is sufficient to use directional antennas rather than omnidirectional antennas [4]. Without a reference antenna, measurement results are relative to radio noise power levels.

2.8.1.2. RF Filter

The survey system may incorporate low-pass and pre-selector filters (LPF). The LPF is first used to reject components higher than the frequencies of interest. The pre-selector filters, such as band-pass filters (BPF), manually tune the system according to the measurement frequency [6]. It is advisable to insert pre-selector filters before the amplifier to prevent the occurrence of intermodulation distortion (IMD) by attenuating in-band and out-of-band interfering signals. The bandwidth of such filters may be a trade-off between achieving narrow band performance and insertion loss, noise figure, and group delay [6, 14].

2.8.1.3. LNA

A pre-amplifier is usually the first RF component. The objective is to increase the weak incoming signal, with minimum RF noise, to a sufficient level for further RF processing in the signal chain. Any circuitry or component noise introduced in the RF front end affects the receiver's ability to detect information. The LNA provides the initial system gain [56]. The survey system shall operate with low noise performance to detect external power. For this reason, LNA is essential, and it

should perform within the linearity range to prevent intermodulation products and spurious emissions [6]. High gain and low noise are essential characteristics for the LNA to keep measured power about 10 dB above the internal noise of the equipment [4]. In cascaded receiver systems, the gain of the first stage is the most important quantity to guarantee optimum performance for the whole system. High gain and low noise figures are the preferred options. Subsequent components have negligible impacts on the system noise figure, as the overall performance in this regard depends on the quality of the first stage [57, 58].

2.8.1.4. Radio Receiver

Most of the previous studies used standard spectrum analysers as radio receivers. The experience in this thesis highlights that one of the advantages of using standard spectrum analysers is that the transformation process for signal analysis is accurate, and post-processing tasks are less demanding. For example, the instrument already translates the received power level in dBm. Other researchers applied SDR tools to architecture radio receivers in software and selected hardware. Post-processing activities are not negligible, time-consuming, and can quickly introduce computing mistakes.

Standard spectrum analysers sweep over a span related to the observatory frequency. Each sweep consists of N points, translated to N power values. Equation (2.8) translates the relationship between the spectrum analyser's span and resolution bandwidth (RBW) and N. Equation (2.9) introduces sweep time (T_s) if $Span$ and RBW are known. The k factor must be at least 2 or 3 to allow enough time for the filters for the intermediate frequency (IF) to settle [7].

$$Span = (N - 1) \cdot RBW \tag{2.8}$$

$$T_s = \frac{k \cdot Span}{RBW^2} \tag{2.9}$$

The relationship between RBW, T_s, and $Span$ is essential to maintain in a spectrum analyser. To measure shallow signal levels, the RBW must be improved. Efforts to optimise the RBW may slow the sweep time, consequently delaying the acquisition time [59].

Standard benchtop spectrum analysers offer dynamic range and sensitivity advantages. However, they are unsuitable for field measurements due to their heavy weight, larger size, high energy consumption, and expensive cost. Portable USB spectrum analysers proved to be effective alternative options. They are more attractive if they are real-time spectrum analysers, as they provide the ability to detect impulsive radio noise [60]. Real-time spectrum analysers incorporate advanced and high-speed technology to process the entire bandwidth in real time. In contrast to sweeping instruments, the real-time analysers allow total bandwidth measurements continuously

and instantaneously. The acquisition time is much faster. Radio receivers implemented in SDR can be inefficient in capturing low-power signals. However, some SDR hardware, including the USRP, has been tested and demonstrated comparable performances with classic spectrum analysers [61].

2.8.2. Measurement Variables

Artificial radio noise is remarkable because its levels depend on human-made activities or processes. For this reason, the artificial radio noise level will vary with location and time. The external noise field strength is a statistical variable over space and time. Data analysis must express the noise figure by statistical parameters relevant to a given frequency, geographic location, and time [6].

2.8.2.1. Temporal Variable

It is essential to characterise the noise environment as a function of day, week, and season. The recommendation is to take measurements in the mid-morning, rush hour, and evening. An extended measurement period, especially using a wide bandwidth, requires a large storage capacity to store raw data [26]. Skeie and Solberg collected data for 10 minute measurement period [6].

2.8.2.2. Spectral Variable

IoT consists of two categories in terms of range: short-range and wide-area. The short-range segment primarily consists of devices connected by unlicensed radio technologies, with a typical range of up to 100 meters, such as Wi-Fi, Bluetooth, and ZigBee. This category also includes devices connected over fixed-line local area networks and powerline technologies. The wide-area segment comprises devices using cellular connections and unlicensed low-power technologies, such as Sigfox, LoRa and RPMA [37].

2.8.2.3. Environmental Variable

Even on one frequency, the radio noise level, mainly when dominated by artificial noise, varies depending on the time and location. The recommendation is to obtain multiple measurements at separate locations selected according to their categories. However, for the benefit of a more detailed evaluation, it is recommended to classify noise measurements into the following categories [4]:

- Remote rural: No apparent civilisation, no buildings, no traffic, no electrical installations within 5 km.
- Rural: Open countryside with essentially agricultural activity, building density < 1/ha, no major roads, no electrified railways.

- Residential: Villages and purely residential areas with no commercial or industrial activities. No electrified railways, major roads, and high voltage overhead lines or facilities within 1 km.
- Urban: Dense residential buildings, including minor commercial or industrial activities and shops. No electrified railways, major roads, high voltage overhead lines, or facilities are within 500 m.
- City: Dense commercial or industrial buildings and offices. Major roads and railways can be nearby but should not be dominating.
- Industrial area: Areas with dense factory sites and heavy industry
- Railway: Locations with dominant electrified major railways
- Road: Locations with dominant road traffic, e.g. highway

All the above measurement locations should be outdoors. Measurement results should be evaluated separately for each location category [4].

Experience shows, however, that indoor noise levels tend to be even higher than those measured outdoors due to the domination of a few single noise sources inside the building. Indoor measurements always measure the sum of noise and interference from single sources. In most cases, emissions from single sources inside the house will dominate. The locations for indoor measurement are in the following categories [4]:

- Domestic: Single house or flat with typical electrical and electronic appliances for private use.
- Office: Electrical and electronic appliances for business use, IT and telecommunication equipment, namely computers, printers, and local area networks.
- Shopping centre: Locations with shops and supermarkets.
- Railway station: Major railway stations inside roofed platform area.
- Airport terminal: Major airports, inside the terminal building.
- Factory: Inside factory buildings dominated by electrical machinery.
- Hospital: Locations dominated by medical appliances

2.9. Mathematical Foundations for Radio Receivers

2.9.1. Noise Voltage

Planck's blackbody radiation law defines noise voltage in Equation (2.10) as random fluctuating quantities caused by random motions of electrons in a resistor exposed at a given temperature [62].

$$V_n = \sqrt{\frac{4hfBR}{e^{(hf/kT)} - 1}} \tag{2.10}$$

h: Planck's constant = 6.626x10^{-34} J.sec or J.Hz^{-1},

k: Boltzmann's constant = 1.38x 10^{-23} J/K,

T: temperature in degrees Kelvin (K),

B: bandwidth of the system in Hz,

f: centre frequency of the bandwidth in Hz,

R: resistance in Ω.

Assuming f=100 GHz and T=100 K at microwave frequencies, the expression hf becomes too small that $e^{(hf/kT)} - 1 \approx hf/kT$. Equation (2.11) reduces the noise voltage as:

$$V_n = \sqrt{4kTBR} \tag{2.11}$$

Figure 2.6 represents the Thevenin equivalent circuit where the load resistor R is in a circuit with noise source V_n incorporated with internal noise resistor R_n. These two resistors must be equal to achieve a maximum power transfer from source to load. Noise power P_n, in Equation (2.12), can be derived from the root mean square of V_n (with $V_n = \sqrt{2} \cdot V_{n_{rms}}$).

$$P_n = \frac{\left(V_{n_{rms}}\right)^2}{2R} = \frac{\left(\frac{V_n}{\sqrt{2}}\right)^2}{2R} = \frac{\left(\frac{\sqrt{4kTBR}}{\sqrt{2}}\right)^2}{2R} = KTB \tag{2.12}$$

This equation implies that noise power is proportional to the temperature and bandwidth. As a rule of thumb, the cooler the device, the narrower the bandwidth, and the lesser the noise power.

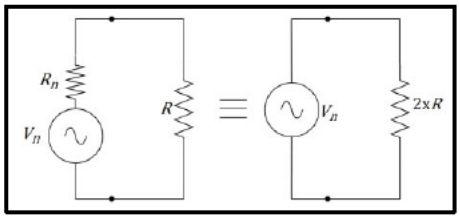

Figure 2.6: Voltage noise source equivalent circuit

At the baseband level, the time-varying voltage is a complex quantity derived from Equation (2.13) in cartesian and polar forms [63, 64].

$$\hat{v}(t) = x(t) + jy(t) = \sqrt{x(t)^2 + y(t)^2} \times e^{j \arctan\left(\frac{y(t)}{x(t)}\right)} \tag{2.13}$$

$x(t)$ and $y(t)$ are real and imaginary components, respectively, and are retrievable in Equation (2.14). R_e denotes the real part. The instantaneous noise voltage $v(t)$ is a passband signal at the carrier frequency.

$$v(t) = R_e\{\hat{v}(t)e^{j2\pi f_c t}\} \tag{2.14}$$

Equation (2.15) defines the instantaneous noise power level by squaring the noise voltage quantity.

$$w = |\hat{v}(t)|^2 = x(t)^2 + y(t)^2 \tag{2.15}$$

2.9.2. Noise Factor

The RF power received by the antenna determines the power of radio environment noise or external radio noise. The antenna's external noise figure indicates the level of radio ambient noise measured by the receiver [7]. Equation (2.16) invokes a combination of parameters at the receiving terminal that characterises the receiver noise factor (f). The external noise factor, in Equation (2.17), is included among these parameters. The antenna feed-point is the reference point that demarks the receiver's internal and external noise. Conventionally, the lowercase variables denote the linear value quantities, while the uppercases denote the transformed values in dB [19, 54].

$$f = f_a + (f_c - 1) + l_c (f_t - 1) + l_c l_t (f_r - 1) \tag{2.16}$$

f_a: external noise factor,

l_c: passive circuit loss,

f_c: noise factor associated with l_c,

l_t: transmission line loss,

f_t: noise factor associated with l_t,

f_r: internal noise factor of the receiver.

$$f_a = \frac{p_n}{k\, T_0\, b} \tag{2.17}$$

p_n: available noise power from an equivalent lossless antenna

k : Boltzmann's constant = 1.38x 10^{-23} J/K

T_0: reference temperature (K) taken as 290 K

b: noise power bandwidth of the receiving system (Hz)

External noise factor f_a is most effective for describing ambient electromagnetic noise power [65]. Equation (2.18) resolves F_a is the external noise expressed in dB as:

$$F_a = 10\log\left(\frac{p_n}{k\,T_0\,b}\right)$$

The logarithmic transformation describes the following parameters:

$$P_n = 10\log(p_n);\ B = 10\log(b);\ 10\log(kT_0) = -204\ \text{dB}$$

$$F_a = P_n - B + 204 \tag{2.18}$$

Equation (2.19) is the noise factor (f_c) associated with passive circuit loss (l_c) [19].

$$f_c = 1 + (l_c - 1)\,\frac{T_c}{T_0} \tag{2.19}$$

T_c: actual temperature (K) of the antenna and nearby ground.

In the event of known noise power, p_{n_r}, referenced to the antenna output or receiver input, the internal noise factor of the receiver is obtained as [65]:

$$f_r = \frac{p_{n_r}}{kT_0 b} \tag{2.20}$$

Loss due to passive circuit (l_c) and thermal noise power (p_{nc}) radiated into the transmission line circuitry affect the noise power p_{n0}, where:

$$l_c = \frac{p_n}{p_{n_r}} \tag{2.21}$$

$$p_{nc} = kt_c b\left(1 + \frac{1}{l_c}\right) \tag{2.22}$$

Leading to the expression of noise power p_{n_r}, as:

$$p_{n_r} = \frac{p_n}{l_c} + p_{nc} \tag{2.23}$$

Applying Equations (2.17), Equation (2.16), and Equation (2.23), Equation (2.24) is the radio noise power at the receiver input:

$$p_{n_r} = \frac{f_a \times kT_0 b}{l_c} + kt_c b\left(1 + \frac{1}{l_c}\right) \tag{2.24}$$

Equation (2.25) is the noise factor (f_t) associated with transmission line loss (l_t) [19].

$$f_t = 1 + (l_t - 1)\,\frac{T_t}{T_0} \tag{2.25}$$

T_t is the actual temperature (K) of the transmission line.

Equation (2.26) is the external noise figure of the antenna, expressed in dB.

$$F_a = 10 \log f_a \tag{2.26}$$

Equation (2.27) is the internal noise figure of the receiver, expressed in dB.

$$F_r = 10 \log f_r \tag{2.27}$$

Equation (2.28) is the external noise factor f_a in terms of noise temperature.

$$f_a = \frac{T_a}{T_0} \tag{2.28}$$

T_a is the effective antenna temperature due to external noise.

2.9.3. Noise Power

Equation (2.29) called, radio link budget equation, is useful to evaluate the relationship between the transmitted power and the received power. It can serve to characterize a radio noise measurement system [54].

$$P_o + L_t + G_t = R + L_b(d_o) + G_r + L_r + F_n + B + 10 \log(kT) \tag{2.29}$$

P_o = the transmission power of the signal (dBm);

L_t : insertion loss of the transmitting system components, such as cables (dB),

G_t : transmitting antenna gain (dBi),

R : minimum signal-to-noise ratio (SNR) for satisfactory reception (dB),

$L_b(d_o)$: transmission or path loss over range d_o (dB),

G_r : receiving antenna gain (dBi),

L_r : insertion loss of the receiving system components (dB),

F_n : total system noise figure measured in decibels relative to thermal noise (dB),

B : receiver noise bandwidth (dB),

k : Boltzmann's constant, 1.38×10^{-23} JK^{-1},

T : receiver temperature, typically assumed to be 290 K.

Assumptions must be made when using the RF link budget with measurement of radio noise environment:

- intentional transmitter is unknown. Variables P_o, L_t, and G_t are ignore ,
- receiver the path loss and insertion loss $L_b(d_o)$ and L_r are nullified.
- receiver antenna had isotropic properties and the gain G_r was equal to zero.
- Without intentional transmitted signal, the receiver signal-to-noise ratio SNR is irrelevant. R is 1 in linear units as the ratio by the received noise over received noise.

Equation (2.30) simplifies the RF link budget applicable to ambient radio noise measurements, as noise power expressed in dBW.

$$P_n = F_n + B + 10\log(kT) \tag{2.30}$$

2.10. Statistical Characteristics of Radio Noise Environment

2.10.1. Random Variable and Random Process

Environmental radio noise (man-made noise, atmospheric noise, interfering signals, etc.) is a random process. This is true for noise from a lone source, like an automobile, as well as the more general case in which the interfering noise is from a collection of many individual sources. The fact that environmental radio noise is a random process means that the noise can be described only in probabilistic or statistical terms [1]. To analyse random phenomena, a probabilistic approach must be used. The concepts and theory of probability and estimation provide a fundamental mathematical framework for the techniques of analysing random signals [66]. A variable is random if its value cannot be precisely predicted in advance [67]. A random variable, subject to uncertainty, is a numerical variable whose value depends on the outcome of a chance experiment. A random variable is discrete if its set of possible values is a collection of isolated points along the number line. A random variable is continuous if its set of possible values includes an entire interval on the number line [68]. Signals and data that possess random characteristics arise in a broad array of technical, scientific, and economic fields. There are many waveforms of continuous random variables, but there are also many applications that concern discrete random variables. These are mostly counting processes; for example, the number of particles produced; the number of calls in a telephone switching; the frequency of heartbeats. In practice, a discrete random variable arises in connection with counting (for example, the number of items purchased, the number of gas pumps in use, or the number of broken eggs in a carton). A continuous random variable is one whose value is typically obtained by measurement (temperature in a freezer compartment, weight of a pineapple, amount of time spent in the store) [66, 68].

Continuous and discrete random variables obey the same general laws. The range of all possible values of random variables comprises its sample space. All values in a sample space are associated with probabilities of occurrence and obey the axioms of probability. Each value in a discrete sample space and any set of continuous or discrete sample spaces are associated with a probability [66].

A random process is a random variable with an additional dimension: time. For each measurement or outcome of an experiment, there exists a time function instead of a single number. This is also true for all signal and time series measurements with random components or properties. The laws of probability and statistics are applied by describing the behaviour of all the processes at a specific time as a random variable. The description of the properties of random processes is accomplished by extending the probabilistic description of random variables to include the additional independent variable: time [66].

2.10.2. Gaussian Characteristic

Gaussian radio noise has properties to be uncorrelated electromagnetic vectors. Bandwidth equal to or greater than receiver bandwidth. Spectral power level increases linearly with bandwidth. Gaussian noise originates from computers, power line communication networks, wired computer networks, and cosmic noise [4].

2.10.3. Non-Deterministic Characteristic

Many situations occur that involve non-deterministic or random phenomena. Similarly, there are many signal and time series measurements with random characteristics. Their behaviour is not predictable with certainty because either it is too complex to model, or knowledge is incomplete, or as with some noise processes, it is essentially indeterminate [66].

Radio noise is random such that the instantaneous voltage cannot be predicted for arbitrary points in time. Thus, radio noise is non-deterministic. Noise cannot be characterised in the time domain by simple parameters amplitude and phase since the voltage at any point in time is a random function. Noise can be characterised in the time domain with a statistical approach [69].

The environmental radio noise is a random process that cannot be represented by a deterministic waveform or any collection of deterministic waveforms [1].

2.10.4. Non-Stationary Characteristic

Man-made noise power statistics are non-stationary in that they have been found to vary over time and location [63]. The behaviour of any random process is not necessarily the same at various

times. The process is stationary if random variables at separate times are equal. A useful form of stationarity is when the mean and covariance do not change with time [66].

Environmental radio noise is nonstationary and, therefore, great care must be exercised in planning, measuring, and interpreting the results. One must measure long enough to obtain a good estimate of the required parameter, but it necessitates that the noise remains "stationary enough" during this period. This is no small point and is frequently overlooked in the design of measurement experiments. It is assumed that the random noise process is stationary enough over some period to obtain the required statistics [1].

2.10.5. Non-Gaussian Characteristic

The best-known kind of noise is a continuously present one called white Gaussian noise [70], which exhibits an instantaneous level that varies following a normal random process of mean zero and variance equal to its mean power. It originates from a mixture of multiple sources of which thermal phenomena are predominant. Samples from white Gaussian noise are independent, and they have a plain power spectrum with the same contributions in all the frequencies. For digital communication systems, Gaussian noise does not represent a major problem as long as the mean power of the desired received signal is high enough [70].

Commonly, the probability density function (PDF) of a particular waveform happens to have a Gaussian shape, but other shapes are possible [69]. Many naturally occurring random vibrations have the well-known “bell-shaped” probability distribution. The random variable has a normal or Gaussian probability distribution, and this is extensively used in random vibration theory to approximate the characteristics of random excitation [67].

Many researchers rely on the assumption that background noise is Gaussian in character. This assumption is not generally valid for HF communications. HF noise consists of galactic noise, atmospheric noise, and man-made noise. Galactic noise is produced by cosmic radiators such as the sun and other stars. Atmospheric noise changes with the time of year and is largely influenced by lightning activity; while man-made noise is produced by electromagnetic radiating devices such as electric machinery, power lines and transformers, and computers [71].

2.10.6. Impulsive Characteristic

A receiving system with a conventional true RMS measurement method gives the power sum of both components, Gaussian, and impulsive noise. The impulse noise component is, in most situations, present in a small fraction of the time. However, it may exhibit high instantaneous power levels during its time of presence. For this reason, the presence and the influence of impulse

noise can be detected by collecting enough data to estimate the amplitude probability distribution (APD) of the received noise signal [6].

Most of the time, the received signal is background noise (Gaussian). Some noises appear unexpectedly as pulses of high amplitude. These are called impulsive noises, and they are not easily tractable. For digital communications, the impulsive noise is harmful because each pulse may cause bursts of bit errors and loss of synchronization. As the occurrence of noise pulses and their amplitude is so unpredictable, it makes it difficult to build digital communication systems that can avoid the effects of impulsive noise [70].

Impulsive noise can be divided into three classes according to its bandwidth:

• Class A: Impulsive noise with a bandwidth smaller than that of the receiver.

• Class B: Impulsive noise with a bandwidth larger than that of the receiver.

• Class C: A case that comprises both Class A and Class B.

In practice, most of the impulsive emissions are Class B. To measure impulsive noise, several functions were proposed: Amplitude Probability Distribution (APD), Numerical Amplitude Distribution (NAD), Average Cross Rating (ACR), and Quasi-Peak Measurements (QPM). There are also other methods to statistically determine some parameters of the received pulses, such as peak amplitude, duration, and spacing. These are Pulse Amplitude Probability Distribution (PAPD), Pulse Duration Distribution (PDD) and Pulse Spacing Distribution (PSD). When measuring impulsive noise, it is customary to set a threshold level to separate it from the Gaussian noise. If surpassed, the signal would be considered impulsive noise. Thus, impulsive noise lower than the threshold would not be included in the analysis. If the threshold is lowered to overcome this drawback, then too much Gaussian noise will be registered instead of impulsive noise [70].

Impulsive noise arises from natural sources such as electrical storms and from many different artificial sources including car ignition systems, power transmission lines, industrial plants, and domestic equipment. Such noise can harm radio communication systems. It is therefore desirable to have a statistical framework for characterising man-made noise. The essential characteristics of impulsive noise are that the individual events have a very short temporal duration, and the noise voltage is random during this brief interval.

2.11. Statistical Function

2.11.1. Measures of Central Tendency

Mean, median, and peak statistical functions are commonly used to characterize noise power. Median external radio noise power is also a preferred statistic for describing radio noise power

levels because, unlike the mean, it will not be impacted by the presence of infrequent but powerful impulsive noises [54, 63]. The measure of central tendency is a value that represents a typical or central entry of a dataset. The most common measures of central tendency are mean (average), median and mode. The mean is a reliable measure because it takes into account every entry of a dataset [72].

2.11.1.1. Mean

The mean or the arithmetic average describes an entire sample with a single number that represents the centre of the data. You calculate the mean by adding up all the observations and then dividing the total by the number of observations. The mean is sensitive to skewed data and extreme values. For datasets with these properties, the mean gets pulled away from the centre of the data. In these cases, the mean can be misleading because the most common values in the distribution might not be near the mean [73].

In Equation (2.31), the mean value of a random variable x, denoted by μ_x, describes where the probability distribution of x is centred. Although the mean is computed differently for discrete and continuous random variables, the interpretation is the same in both cases. The mean value of a discrete random variable x, denoted by μ_x, is computed by first multiplying each possible x value by the probability of observing that value and then adding the resulting quantities. Each value of x is multiplied by its corresponding probability and the products are added [68, 72].

$$\mu_x = \sum x \cdot p(x) \tag{2.31}$$

As described by Witte [69], the waveform will have an average or mean value given by

$$\bar{x} = E[x] = \int_{-\infty}^{+\infty} x \cdot p(x) \cdot dx \tag{2.32}$$

$E[x] =$ Expected value of x; and $p(x) =$ probability density function of $x(t)$

The mean of a data set is computed as $\bar{x} = \frac{\sum x}{n}$ where n is the total number of entries in the data set [72].

2.11.1.2. Median

The median of a dataset is the value that lies in the middle of the data when the dataset is numerically ordered. It measures the centre of an ordered dataset by dividing it into two equal parts. If the dataset has an odd number of entries, the median is the middle data entry. If the dataset had an even number of entries, the median is the mean of the two middle data entries [72].

2.11.1.3. Mode

The mode is the data entry that occurs with the greatest frequency. A dataset can have one mode, more than one mode, or no mode. If no entry is repeated the dataset has no mode. If two entries occur with the same greatest frequency, each entry is a mode (bimodal). Ordering the data set helps to find the mode [72].

2.11.2. Measures of Variation

The measures of variation include range, variance and standard deviation [72].

2.11.2.1. Range

The range of a dataset is the difference between the maximum and minimum data entries in the set [72].

$$Range = (Max\ Data\ Entry) - (Min\ Data\ Entry) \tag{2.33}$$

2.11.2.2. Variance

The variance, σ^2, is a measure of how the instantaneous value of x strays from the mean value of x [69]. Thus, given by

$$\sigma^2 = E[(x - \bar{x})^2] \tag{2.34}$$

Closely related to the variance is the standard deviation, σ. Since the square of the standard deviation is equal to the variance, the two quantities are redundant [69].

2.11.2.3. Standard Deviation

Standard deviation is the square root of variance. The standard deviation of a random variable x, denoted by σ_x, describes variability in the probability distribution. When the value of σ_x is small, observed values of x will tend to be close to the mean value (little variability). When the value of σ_x is large, there will be more variability in observed x values. Although the standard deviation is computed differently for discrete and continuous random variables, the interpretation is the same in both cases [68].

2.11.3. Density and Distributions Functions

The probability distribution for a random variable is a model that describes the long-run behaviour of the variable [68].

2.11.3.1. Shape of Distributions

The three main shapes of distributions are [72, 73]:

- Symmetrical distributions: data are equally distributed on both sides of the graph (distribution). Both halves of the distribution are approximately mirrored images. The mean and median have the same value at the centre of the distribution.
- Uniform or rectangular distributions: all entries in the distribution have equal frequencies. Uniform distributions are therefore symmetrical. The mean and median have the same value at the centre of the distribution.
- Skewed distributions: data are not equally distributed on both sides of the graph. Histograms easily show where the data are skewed. Distributions can be referred to as skewed right or skewed left distributions.
    - Skewed left distribution: negatively skewed. most of the data are on the right side of the graph. The mean is to the left of the median. The tail of the graph elongates more to the left.
    - Skewed right distribution: positively skewed. Most of the data are on the left side of the graph. The mean is to the right of the median. The tail of the graph elongates more to the right.

2.11.3.2. Probability Distributions for Continuous Random Variables

A random statistical process can be characterized in terms of a Probability Density Function (PDF) which can take many different forms according to the type of process being modelled. In the case of Additive White Gaussian Noise (AWGN), by definition, the voltage is modelled by a Gaussian distribution with zero mean and a uniform Power Spectral Density (PSD). Assuming a Gaussian mean power level of w_{0_g} and a random noise variable, w_{RV}, Equation (2.35) is the power PDF for AWGN [63, 74]:

$$p_g(w_{RV}) = \frac{1}{w_{0_g}} e^{-\frac{w_{RV}}{w_{0_g}}} \tag{2.35}$$

A probability distribution for a continuous random variable x is specified by a curve called a density curve. The function that defines this curve is denoted by $f(x)$ and it is called the density function. The probability that x falls in any interval is the area under the density curve and above the interval. The following are properties of all continuous probability distributions [68]:

- $f(x) \geq 0$: so that the curve cannot dip below the horizontal axis.
- The total area under the density curve is equal to 1.

When the density is constant over an interval (resulting in a horizontal density curve), the probability distribution is called a uniform distribution [68].

Probabilities for continuous random variables are often calculated using cumulative areas. A cumulative area is all of the areas under the density curve to the left of a particular value [68].

2.11.3.3. Probability Distributions for Discrete Random Variables

Three characteristics of impulsive electromagnetic noise are important: the amplitude of the pulse, the duration of the pulse, and the interval between the two adjacent pulses. The six most suitable statistical metrics for impulsive radio noise are [65, 75]:

- Cumulative Distribution Function (CDF)
- Amplitude Probability Distribution (APD)
- Noise Amplitude Distribution (NAD)
- Average Crossing (ACR)
- Pulse Duration Distribution (PDD)
- Pulse-Spacing Distribution (PSD)

Common ways to display a probability distribution for a discrete random variable are a table, a probability histogram, or a formula. A probability distribution table for a discrete variable shows the possible x values and $p(x)$ for each possible x value. Because $p(x)$ is a probability, it must be a number between 0 and 1, and because the probability distribution lists all possible x values, the sum of all the $p(x)$ values must equal 1. A pictorial representation of a discrete probability distribution is called a probability histogram. The picture has a rectangle centred above each possible value of x, and the area of each rectangle is proportional to the probability of the corresponding value [68].

2.11.3.4. Probability Density Function and Cumulative Distribution Function

The frequency of occurrence of values in a sample space is described by a pair of complementarity functions called the Probability Density Function (PDF) and Cumulative Distribution Function (CDF) [66].

Noise can be described by tabulating how often a certain voltage appears. In a continuous form, this results in the Probability Density Function (PDF). It tells the probability of certain voltage occurring [69].

We need some way of probabilistically describing random variables that works equally well for discrete and continuous random variables. One way of accomplishing this is through the CDF. Continuous random variables have continuous CDF. Discrete random variables have step-discontinuous CDF. The CDF is a complete and useful description for the computation of probabilities. However, for purposes of computing the statistical averages, the Probability Density

Function of a random variable is more convenient. The PDF of a random variable is defined to be the derivative of the CDF. Since the CDF of a discrete random variable is discontinuous, its PDF, mathematically speaking, does not exist at the points of discontinuity [76].

The CDF of instantaneous noise power describes the probability that the noise power will not exceed a value. CDF uniquely defines the probability that the random variable X is less than or equal to some allowed value L. The distribution function is defined in Equation (2.36) where W is the noise power random variable, w is the noise power independent variable, and $P(w)$ is the probability density function (PDF) of the noise power random variable [63, 65, 74]:

$$CDF(L) = Probability(X \leq L) = P(X \leq L) = P(W \leq w) = \int_0^w p\,(w_{RV}) dw_{RV} \qquad (2.36)$$

2.11.3.5. Amplitude Probability Distribution

The Amplitude Probability Distribution (APD) is defined as the percentage of time when the measured envelope of an interfering signal exceeds a threshold level [77, 78]. It is necessary to set a threshold level so that any sample that exceeds it is taken into account [70]. The EMC engineers are more concerned about the noise envelope exceeding a certain level; thus the APD is preferable which is complementary to the cumulative distribution function (CDF) [78].

APD is a function used in noise measurement to describe signal amplitude statistics. The APD technique is useful to identify noise emissions from various sources of electrical appliances and can then be used for noise modelling and simulation. The APD provides important information to characterize the noise. As an example, the APD approach can be used to characterize the amplitude statistics of non-Gaussian noise produced from man-made electrical or electronic devices. The APD is an important statistical tool to evaluate the characteristics of impulsive noise. The APD has been proposed within CISPR for the measurement of electromagnetic noise emission [79].

In general, radio disturbance is measured by using a measuring receiver as specified in the CISPR standard. Several detectors in the receiver provide statistics regarding the disturbance amplitude, such as the peak values, average values, and RMS values. What value one should use depends on the purpose of the measurement. Among them, APD provides promising values if you want to evaluate the performance degradation in digital wireless communication systems affected by a disturbance – it has been reported that under some conditions the APD of disturbance correlates well with the BER degradation of wireless systems affected by interference. APD measurement method is therefore useful for measuring the effects of disturbance to protect wireless communication systems and has become a means of addressing the industry's concerns regarding

interference problems in wireless systems [80]. APD is one of the most represented statistical distributions; it has been used for impulsive noise measurements, analysis, and theoretical characterization [70]. There is a need to determine the probability that the noise power will exceed a value. Equation (2.37) show the APD defined as the complement of the CDF.

$$APD(L) = A(w) = Probability(X \geq L) = P(X \geq L) = 1 - CDF(L) = \int_{w}^{\infty} p(w_{RV}) dw_{RV} \quad (2.37)$$

The APD is dependent on variables such as the bandwidth of the measurement, bandwidth of the noise, and time intervals between noise pulses. The APD of zero-mean, white Gaussian noise is completely described by its mean power; therefore the performance of a radio receiver in white noise can be predicted from mean noise power alone [63]. In radio engineering, the APD function is used rather than the CDF, as it indicates the probability that the power exceeds a given level. Equation (2.38) is the APD if the noise is WGN distributed [74]:

$$A_g(w) = e^{-\frac{w}{w_{0g}}} \quad (2.38)$$

2.11.3.6. Noise Amplitude Distribution (NAD)

Equation (2.39) shows the NAD as the total sum of the number of pulses, which exceeds a given level divided by the total length of the measuring time.

$$Num(X > L) = \frac{\sum_{k=0}^{n} p_k}{T} \quad (2.39)$$

2.11.3.7. Average Crossing (ACR)

Equation (2.40) shows the ACR as the average number of crossings in positive direction above given level divided by the total length of the measuring time.

$$Num(X = L) = \frac{\sum_{k=0}^{n} A_k}{T} \quad (2.40)$$

2.11.3.8. Pulse Duration Distribution (PDD)

Equation (2.41) defines PDD as the fraction of time that the pulse duration at level exceeds time τ.

$$PDD(\tau, L) = \int_{\tau}^{\infty} p_d\left(\frac{t}{L}\right) dt = CDF\left(\frac{\tau}{L}\right) \quad (2.41)$$

Where $p_d\left(\frac{t}{L}\right)$ is the density function of the pulse duration time about a threshold L.

2.11.3.9. Pulse-Spacing Distribution (PSD)

Equation (2.42) defines PSD as the fraction of time that the pulse-spacing duration at level exceeds time s.

$$PDD(s, L) = \int_{s}^{\infty} p_s\left(\frac{t}{L}\right) dt = CDF\left(\frac{s}{L}\right) \tag{2.42}$$

Where $p_s\left(\frac{t}{L}\right)$ is the density function of the pulse spacing time about a threshold L.

2.12. IoT Challenges as a Contributor to Big Data

According to Alam et al., the number of objects connected to IoT is expected to give rise to an enormous amount of valuable data which will be used to understand and control complex environments around us, enabling better decision-making, greater automation, higher efficiencies, productivity, accuracy, and wealth generation [81].

A major challenge in these settings is the timely analyses of large amounts of data to produce highly reliable and accurate insights and decisions so that IoT could live up to its promise. Data mining and other machine learning methods would play a critical role in creating smarter IoTs. Machine learning is among the top methods to gain hidden insights from IoT data [81].

Romero et al. suggested that the collection of radio frequency data can overwhelm even the largest data storage capacities very quickly due to high sampling frequencies [82]. Data rates can exceed 200 GB per second, and it is prohibitively expensive to store large samples in real-time. Adding to the problem is the time required to verify the desired signals were recorded in the data collection and properly annotate the data for convenient retrieval at subsequent times. Also noteworthy is the problem created by both expected and unexpected sources of radio frequency signals that can diminish the value of the data collected.

Human error may also lead to the incorrect annotation of data whose consequences can be difficult to mitigate. These issues can be solved, in large part, through an automatic classification of datasets. A process has been described here that uses machine learning algorithms to classify datasets composed of RF signals. Mahdavinejad et al. reinforce that intelligent processing and analysis of this big data is the key to developing smart IoT applications [83]. Their research assessed the different machine learning methods that deal with the challenges in IoT data by considering smart cities as the main use case. One key contribution of this study is the presentation of a taxonomy of machine learning algorithms explaining how different techniques are applied to the data to extract higher-level information. According to Mary et al., spatial and temporal correlation is the noteworthy and exclusive characteristic of IoT data that can be exploited to

considerably enrich the accuracy of analysis [84]. Existing pre-processing techniques don't consider the heterogeneous nature of IoT data and so the deployment of these techniques in the IoT environment to perform data preparation tasks produces biased, inaccurate, or misleading results. More fundamentally, the nature of the errors in IoT data cannot be easily corrected by traditional pre-processing solutions which don't consider the temporal or spatial aspects of data. In their research, a unified pre-processing framework is proposed that comprises three pre-processing models to pre-process IoT data by carefully considering the spatial, temporal, attribute, and time-lagged correlations, to enhance the accuracy of subsequent analysis. Another method to consider is the use of statistical analysis to assist with the probabilistic nature of radio noise [85].

2.13. Summary

This chapter provided in-depth literature coverage on the definition of electromagnetic radiation, sources of radio noise, radiation phenomena, electromagnetic spectrum, risks of exposure to electromagnetic radiations, radio noise measurement techniques, and mathematical and statistical foundations of radio noise. The chapter also reviewed IoT as a contributor to big data.

Radio noise is everywhere, from natural and artificial sources. This chapter explained the pros and cons of electromagnetic radiation. Although classified as non-ionizing radiations, stakeholders in the RF industry shall remain attentive to regulatory compliance. Failure to comply with electromagnetic norms may not only harm the technological performance of the equipment but may lead to ecosystem and health hazards. Although discrepancies in study results do not allow scientists to establish common conclusions on health risks, exposure to RF requires caution. The adoption of IoT in homes, offices, farming and agriculture must be mindful as exposure to RF may affect plant development, food production, and pharmaceutical industries.

CHAPTER 3: DEVELOPMENT OF RADIO NOISE SURVEYING SYSTEM

3.1. Introduction

This chapter describes the equipment setup for the noise measurement system employed in this study. It also discusses the technical implementation of the Radio Noise Surveying System (RNSS) and data acquisition. RNSS is a radio receiver for external noise measurement.

Typical systems for ambient radio noise measurements are expensive and cumbersome. Many of these projects necessitate vehicles (truck or van) as housing for the payload, masts as mounting structures for the external antenna, diesel generators for onsite power supply, battery chargers or UPS if mains are unavailable onsite. Benchtop laboratory-type instruments are heavy and space-taking. Where it was impossible to use vehicles, it was helpful to use a backpack (about 15 kg of payload) or tow away by bicycle [60, 74]. These characteristics showed that the RF noise survey was demanding in terms of budget, logistics, equipment, and workforce. Hence, these campaigns often need support from well-established organizations [3]. Adding to these challenges, many passive components (cables, switches, filters) can deteriorate the system noise figure. A system like this requires care to optimize the system's dynamic range and noise figure. For data acquisition, if the radio receiver was a standard spectrum analyzer, it must record measurements in real-time, considering that recording an immense volume of data demands adequate memory. If third-party software was essential to control the instrument, the spectrum analyzer should feature an interface port for computer interactions. Not all spectrum analyzers incorporate these features.

A low-cost, simple, and agile system was proposed in this research to provide:

- Lightweight antennas eliminate the need for a mast or tripod.
- Connectable components bring no need for RF cables in between modules.
- A reduced number of passive RF components does not require many amplification stages.
- Software-Defined Radio System. No need for a state-of-the-art spectrum analyzer.
- Low-power consumption devices do not need for massive power generation equipment.
- Easily transportable equipment brings no need for delicate logistic transportation.
- Low-cost components. No need for fund-raising campaigns.
- Repeatable setup. Swift system configuration after transferring to different locations.

3.2. Design Specifications

With the body of knowledge consolidated in the literature review, this thesis proposed Figure 3.1 as the schematic for the RNSS. The analogue RF front end included an antenna to receive external RF waves, an LNA to optimize the noise figure by amplifying the incoming signals and a Band-Pass

Filter (BPF) to reject out-of-band frequencies. Each front-end component had to feature SMA termination to enable smooth modular assembly.

The digital back end, controllable from an external laptop, was developed using the Universal Software Radio Peripheral (USRP) hardware.

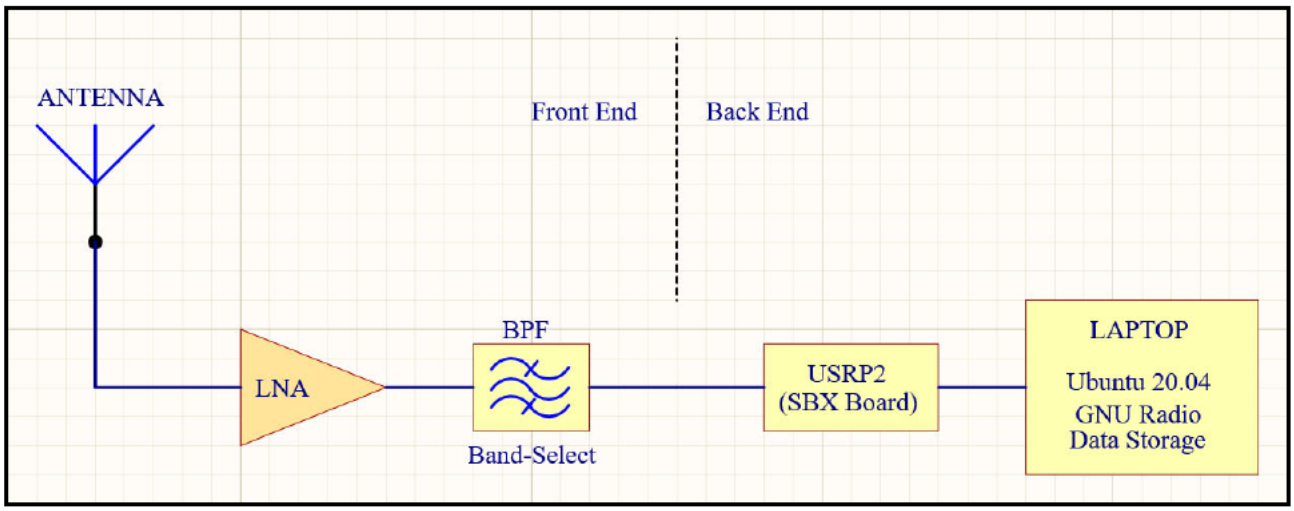

Figure 3.1: Block Diagram of the Noise Measurement System

3.2.1. Antenna

The antennas were lightweight terminal monopole antennas. Ideally, a single broadband antenna was the optimal choice to cover the three frequency bands of interest: 433 MHz, 868 MHz, and 2.4 GHz. However, such an antenna was scarce to procure to meet the RNSS strategies on cost, size, ergonomics, and weight. This research selected three passive terminal monopole antennas available on the market shelves.

Input impedance and radiation pattern are critical characteristics in antenna tests and measurements. The return loss and the voltage standing wave ratio (VSWR) specify the input impedance, and their measurements are relatively straightforward. However, radiation pattern measurements require a lot more effort [86].

Because this research worked with already-manufactured antennas, testing every detail was unnecessary in this thesis. A Vector Network Analyser (VNA), model E5062A by Keysight, was utilized to verify and validate the performing characteristics of each antenna. The low-loss and ultra-flexible RF cable, model ULC-1.5FT-SMSM+ by Mini-circuits, was used to connect the VNA to the device under test. During these measurements, parameters such as RF cable loss, connector loss, adaptor loss, and antenna factors were negligible. Figures 3.2, 3.3, and 3.4, respectively, present the return loss measured values of -13.501 dB (433 MHz), -13.401 dB (868 MHz), and -7.242 dB (2.4 GHz). Figure 3.4 estimates a return loss of -15 dB at 2.442 GHz.

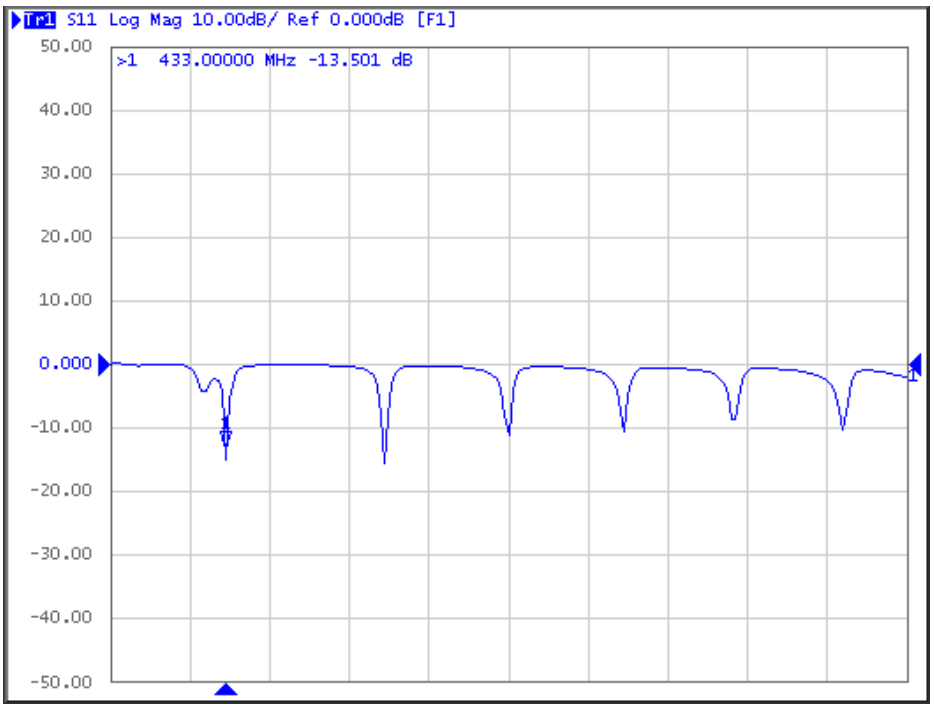

Figure 3.2: Return loss of the 433 MHz antenna (Asian Creation AC-Q433-L20)

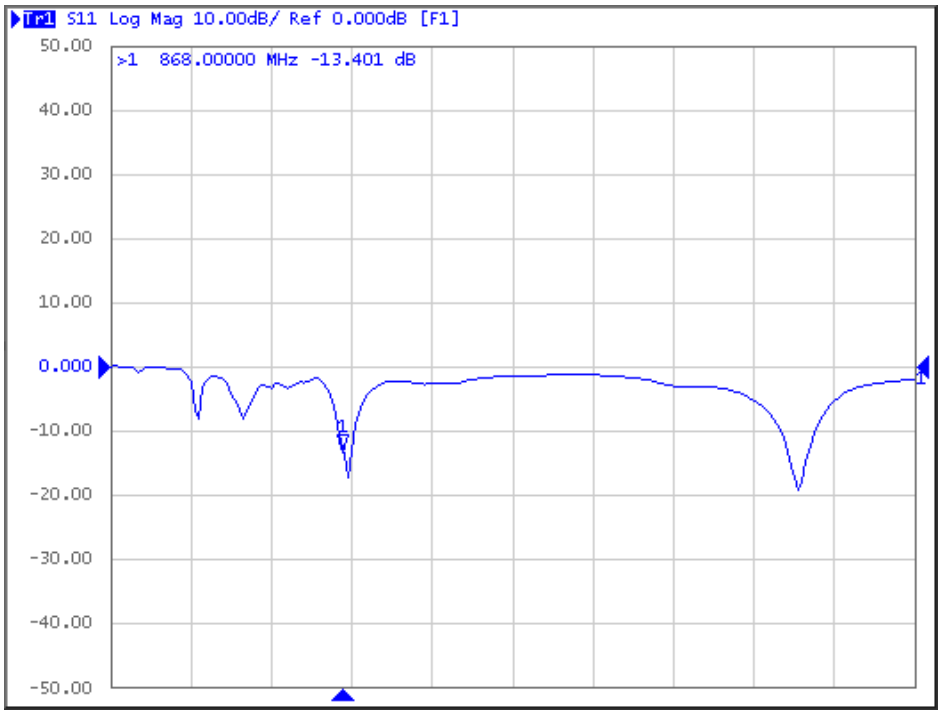

Figure 3.3: Return loss of the 868 MHz antenna (Taoglas TI.85.2113)

Equation (3.1) calculates the voltage standing wave ratio (VSWR) and verifies the antenna impedance matching a 50 Ω system impedance [87].

$$VSWR = \frac{1 + |\Gamma|}{1 - |\Gamma|} \tag{3.1}$$

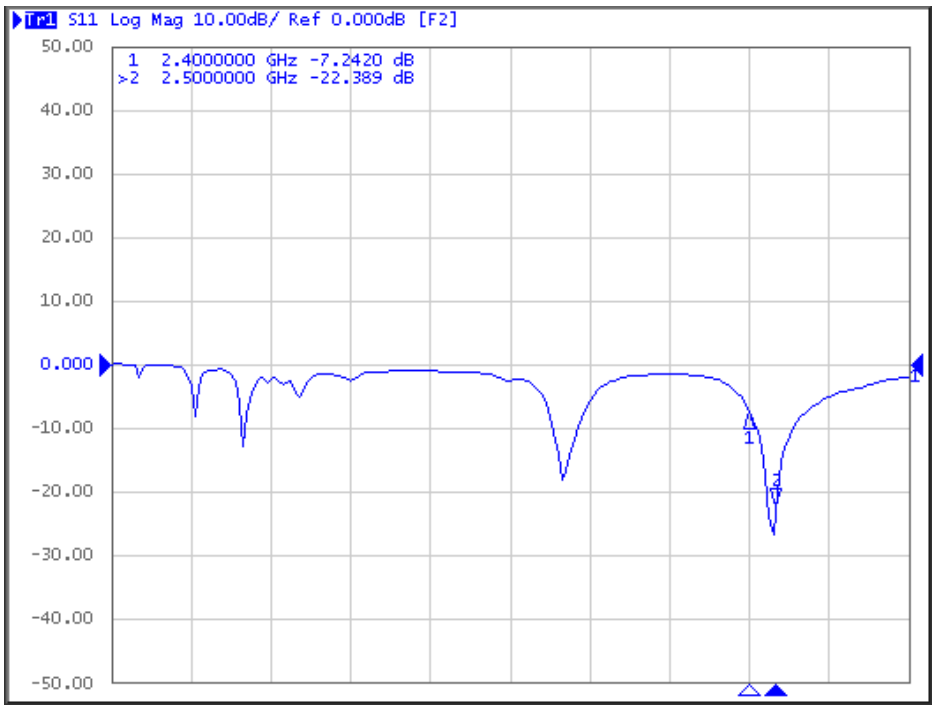

Figure 3.4: Return loss of the 2.4 GHz antenna (Taoglas GW.15.2113)

With |Γ| being the antenna reflection coefficient absolute value. Equation (3.2) is the relation between the return loss RL in dB and reflection coefficient Γ. In Equation (3.3), the reflection coefficient defines the ratio of reflected over incident voltages. An optimal antenna impedance matching requires at least 0.316 of the reflection coefficient. This criterion translates to a -10 dB return loss, ensuring that 90 % of the incident signal could reach the radiating element. In theory, the VSWR ranges from 1 to infinite. However, $1 < \text{VSWR} \leq 2$ achieves realistic impedance matching [87].

$$RL = 20\log(|\Gamma|) \tag{3.2}$$

$$\Gamma = \frac{V_{reflected}}{V_{incident}} \tag{3.3}$$

Table 3.1 shows the measured return loss, the calculated VSWR values and antennas selected for the RNSS. Measured and calculated results met the design criteria and endorsed the suitability of the three chosen antennas.

Table 3.1: Selected antennas

Band	Manufacturer	Part Number	Type	Return Loss	VSWR
433 MHz	Asian Creation	AC-Q433-L20	Monopole	-13.5 dB	1.536
868 MHz	Taoglas	TI.85.2113	Monopole	-13.4 dB	1.543
2.4 GHz	Taoglas	GW.15.2113	Monopole	-15 dB	1.433

3.2.2. Low-Noise Amplifier

The RF front-end in this thesis incorporated an ultra-low noise figure LNA, model ZX60-P33ULN+ by Mini-Circuits. The most attractive features of this LNA were:

- Frequency range: 400 MHz - 3 GHz, covering the three frequencies of interest in this thesis,
- Noise figure: 0.38 – 0.9 dB, guaranteeing outstanding RF performance for the front end,
- Gain: 8.8 – 24.5 dB, more than sufficient for this RNSS,
- Input power max: 22 dBm, robust to handle strong external RF interference, if any,
- IP3: 38 dBm: high IP3, ensures high linearity operation,
- Output power 1 dB compression: 17 dBm
- Input and output impedances: 50 Ω, allowing direct impedance matching in the RF chain.
- DC power: 5.5 V, 0.5 W max, operable from low-power DC sources.
- Size: 19 x 18.8 mm, adequate with the RNSS.

There was no requirement for more front-end amplification because the RF daughter board in the USRP included further RF processing (filtering, amplifying, mixing, and down-converting). Figure 3.5 shows the LNA gain measurement with markers placed at the frequencies of interest: 23.546 dB (433.92 MHz), 18.847 dB (868.30 MHz), and 10.298 dB (2.442 GHz). The VNA port power delivered -30 dBm, which is low enough to avoid overdriving the LNA. The measured results, in good correlation with the datasheet, validated the suitability of the selected LNA.

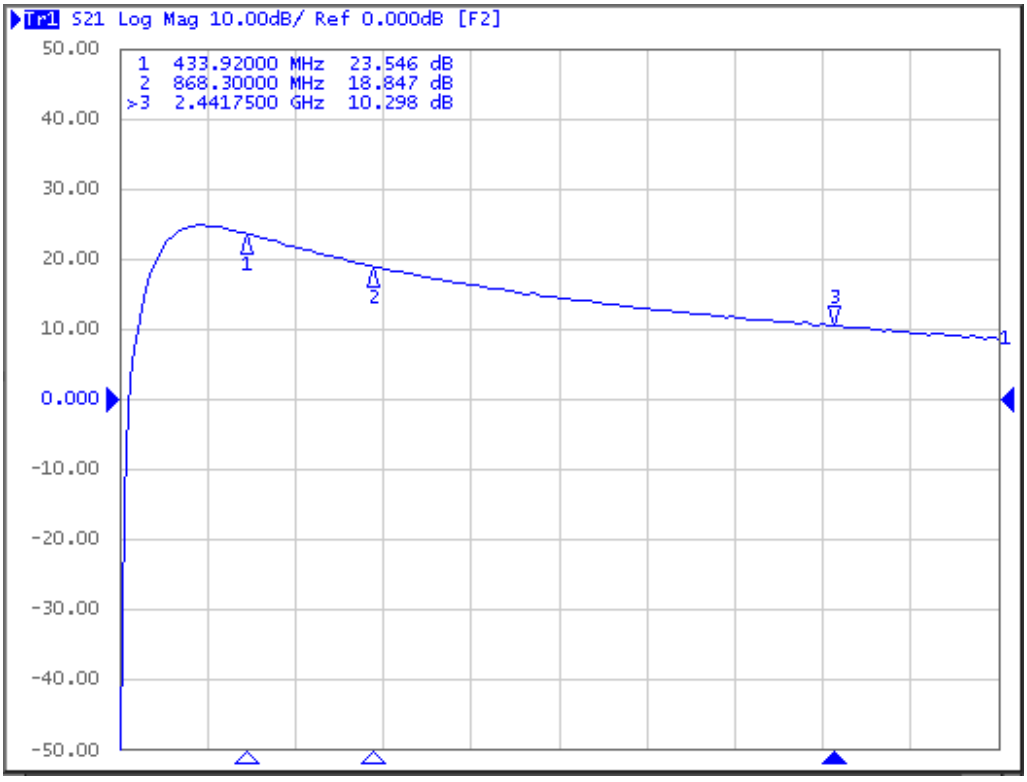

Figure 3.5: The gain curve of the LNA (ZX60-P33ULN+)

The LNA gain was also measured for individual frequency of interest using an RF signal generator and a spectrum analyzer. Figure 3.6 shows this scenario's block diagram, which includes the RF Signal Generator, model N9310A by Keysight, and the Spectrum Analyser, model N9010A by Keysight. The datasheet warns that the maximum input power to this LNA is +14 dBm and 17 dBm for output power at 1 dB compression. With these conditions, the output of the Signal Generator was set to -10 dBm to ensure the LNA operates safely in linearity. The Spectrum Analyser was connected to the output of the LNA to measure the power level due to the LNA amplification. Figure 3.7, Figure 3.8, and Figure 3.9, respectively, show the measured results for LNA's output power in this test scenario: 13.55 dBm (433.9 MHz), 8.55 dBm (868 MHz), and -0.92 dBm (2.4 GHz).

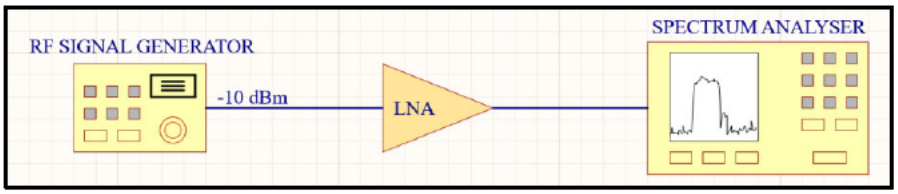

Figure 3.6: LNA Test using RF Signal Generator and Spectrum Analyser

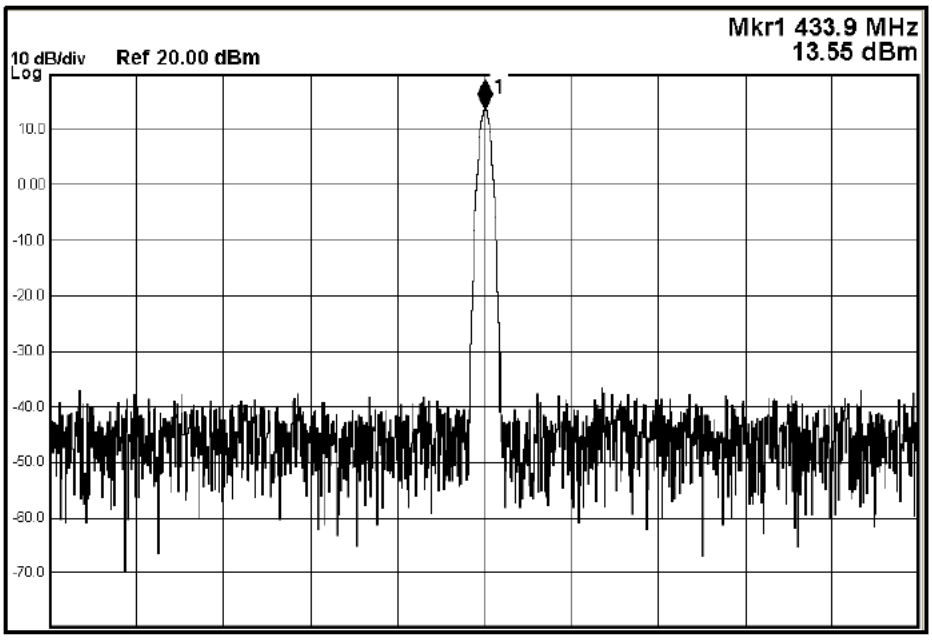

Figure 3.7: LNA output power level at 433 MHz

Matlab RF Budget Analyzer examined LNA power performance at each frequency by simulation. The device datasheet provided LNA properties such as gain, noise figure, and output intercept point (OIP). For the 433 MHz band, Figure 3.10 shows the simulation setup with a signal bandwidth of 25 MHz (the simulation worst-case scenario) and an input power of -10 dBm (same as the measurement). Harmonic Balance (HB) analysis computed the characteristics.

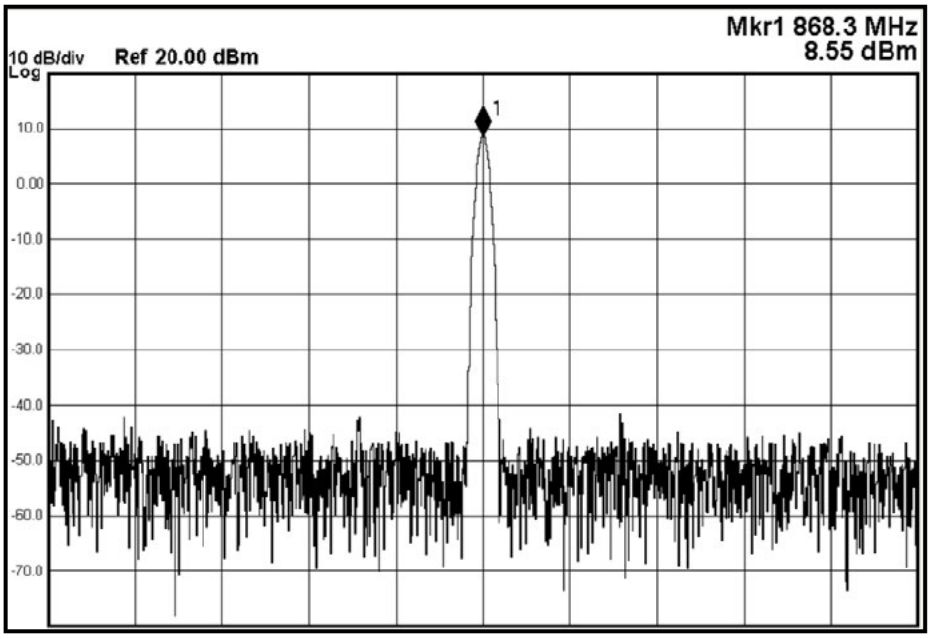

Figure 3.8: LNA output power level at 868 MHz

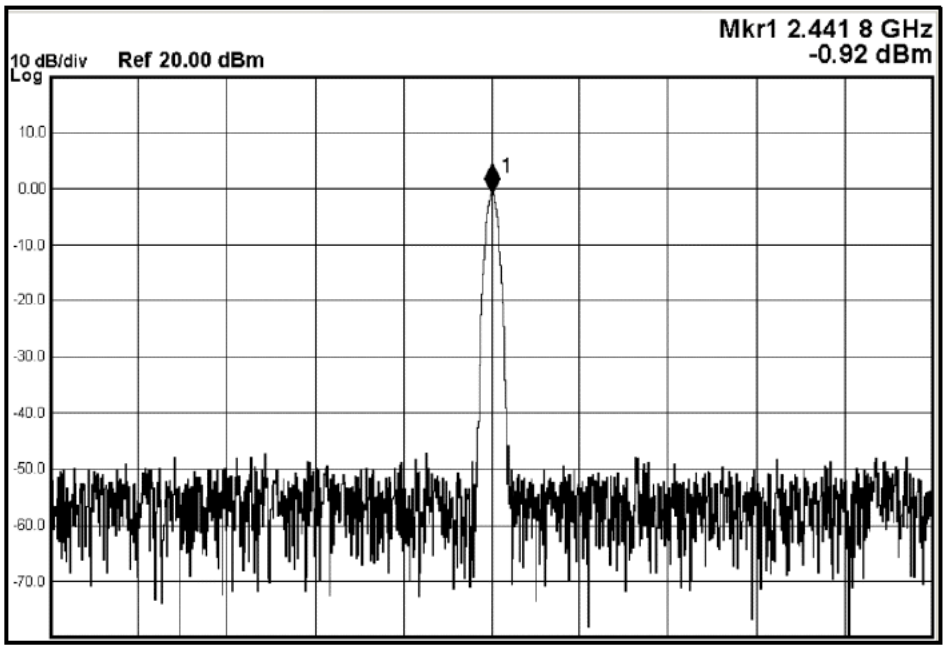

Figure 3.9: LNA output power level at 2.4 GHz

Figure 3.11 shows the simulation results for 433 MHz. Figure 3.12 and Figure 3.13, respectively, show the simulation setup and results for 868 MHz. Figure 3.14 and Figure 3.15, respectively, show the simulation setup and results for 2.4 GHz. The results clarify the usually time-consuming parameters to test in a laboratory. Among these parameters are those related to the amplifier linearity performance, such as the 1 dB compression point, the input and output saturation values, the third-order intercept point (IP3), and the signal-to-noise ratio (SNR). Additionally, the simulation differentiates between the ideal and actual saturation powers. The results prove that the -10 dBm level correctly kept the LNA within its linear region during measurement.

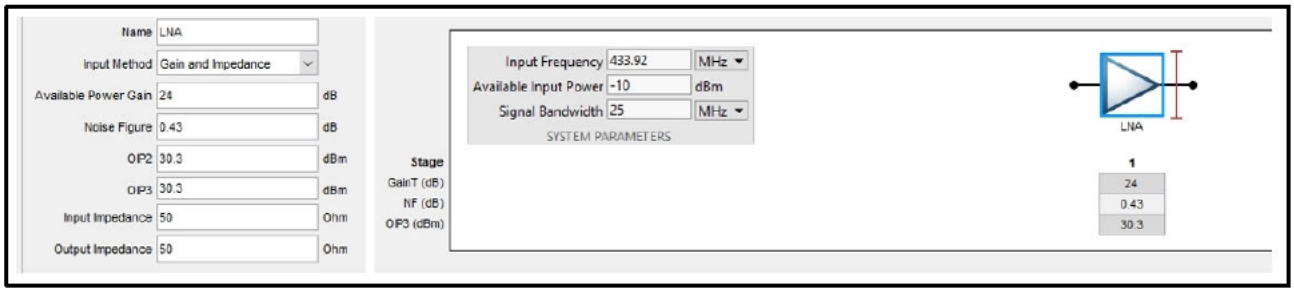

Figure 3.10: LNA simulation setup, 433 MHz

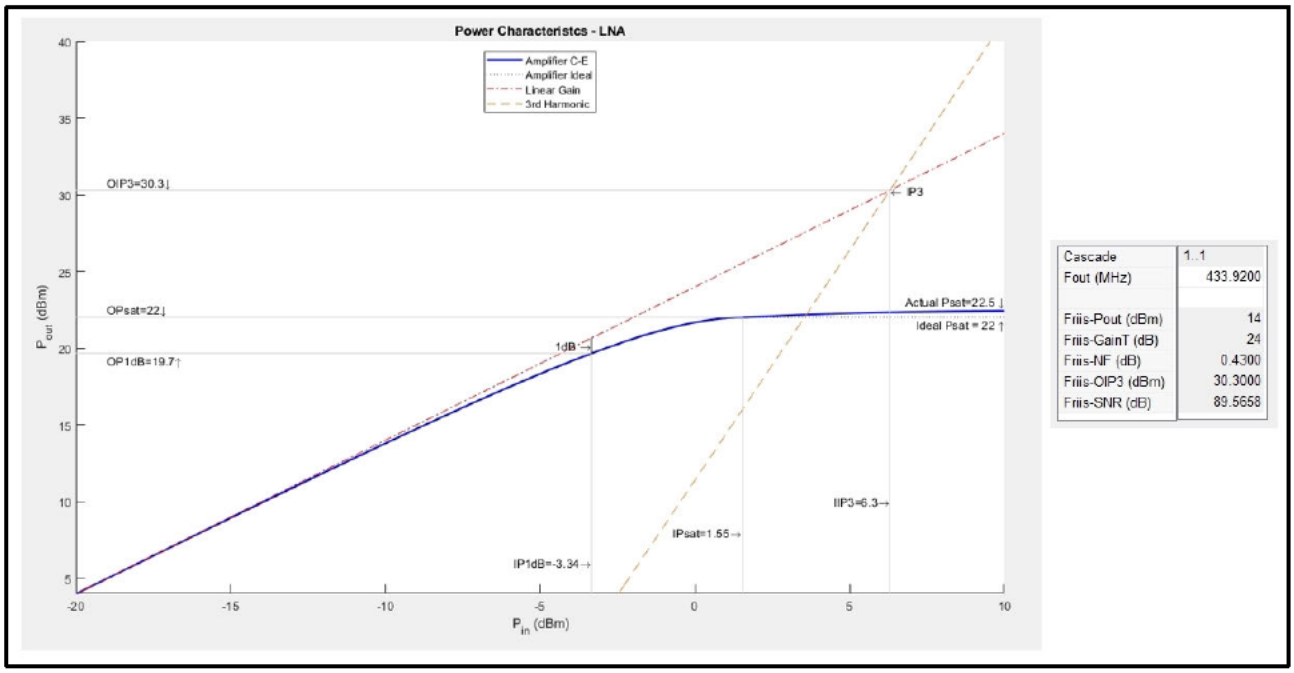

Figure 3.11: LNA simulation results, 433 MHz

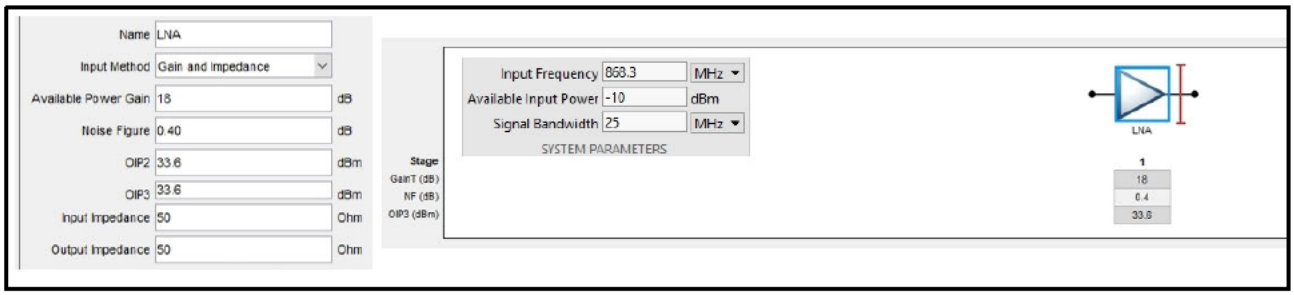

Figure 3.12: LNA simulation setup, 868 MHz

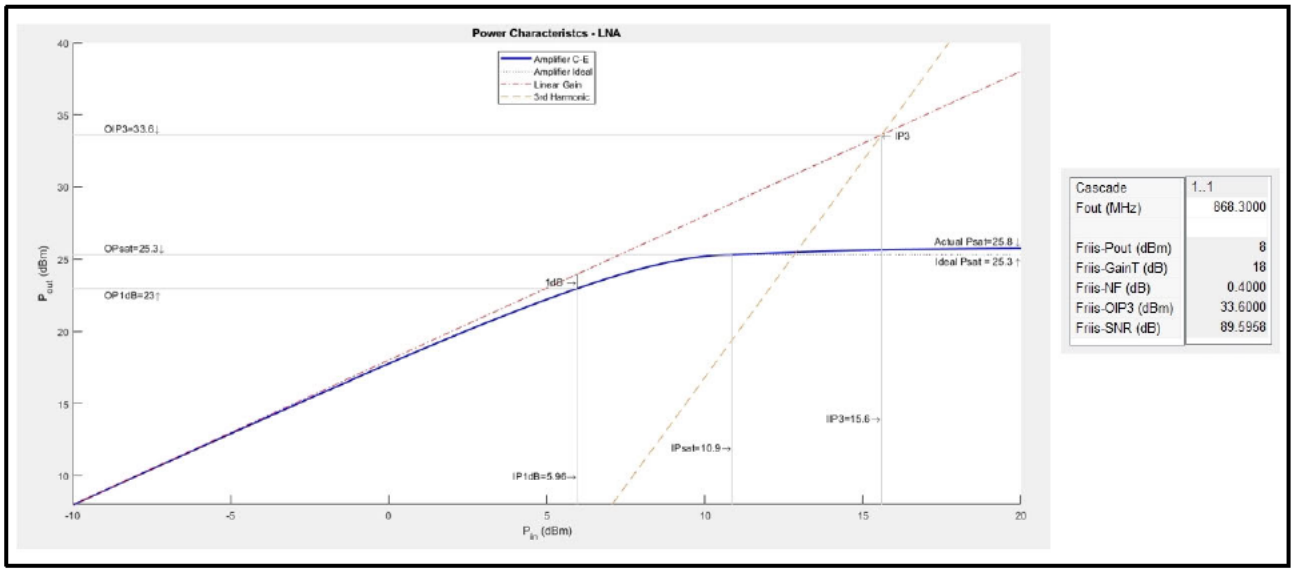

Figure 3.13: LNA simulation results, 868 MHz

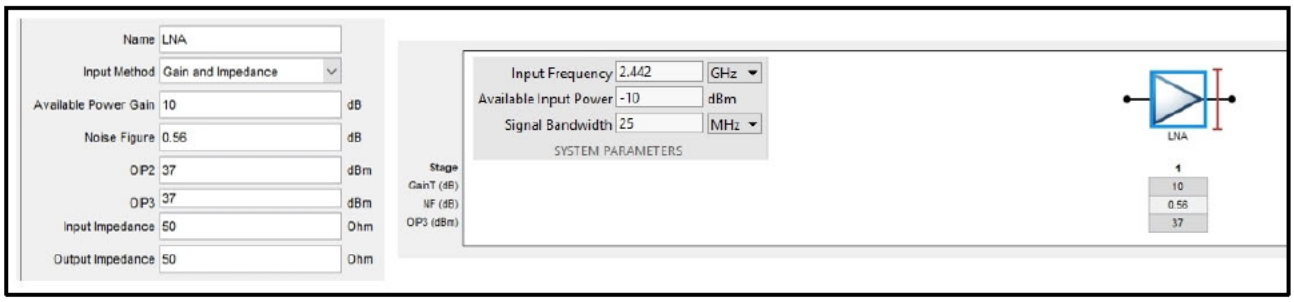

Figure 3.14: LNA simulation setup, 2400 MHz

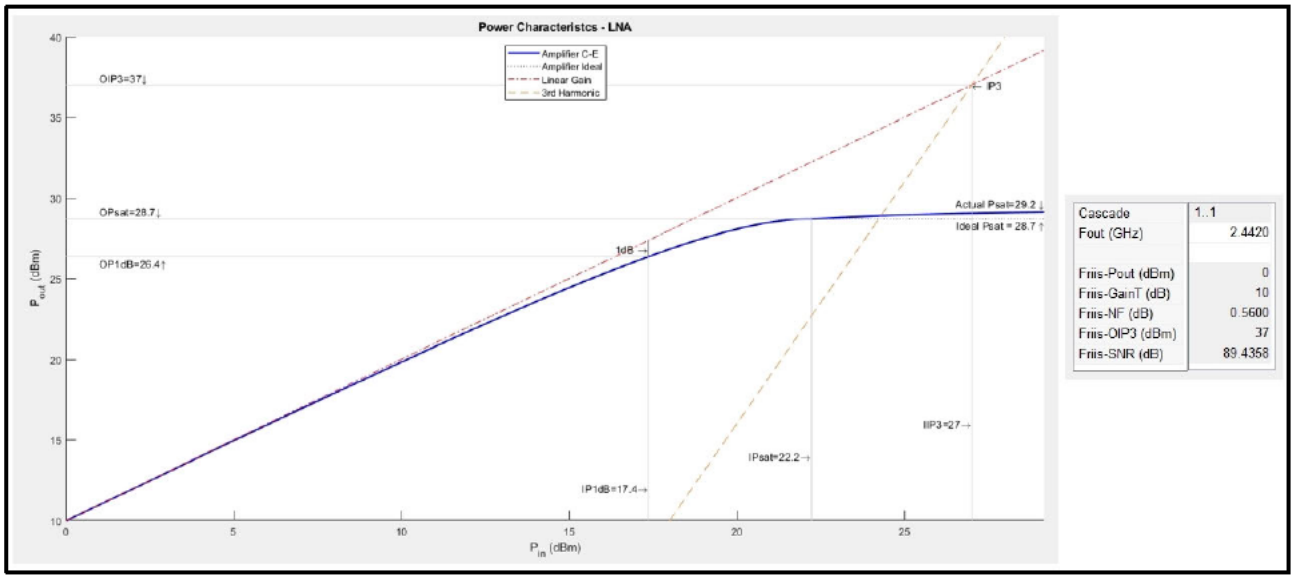

Figure 3.15: LNA simulation results, 2400 MHz

Table 3.2 summarizes the LNA test results comparing measured (Meas.), simulated (Sim.), and calculated (Calc.) data. Comparable values adhere to the expectations and validate this LNA.

Table 3.2: LNA gain - Calculated versus Measured

Frequency (MHz)	Input (dBm)	Output (dBm)		Gain (dB)			SNR (dB)	IPsat (dBm)	OPsat (dBm)	IP1dB (dBm)	OP1dB (dBm)
		Meas.	Sim.	Calc.	Meas.	Sim.	Simulated only				
433	-10	13.55	14	23.55	23.546	24	89.57	-3.34	22.5	-3.34	19.7
868	-10	8.55	8	18.55	18.847	18	89.60	10.9	25.8	5.96	23
2400	-10	-0.92	0	9.08	10.298	10	89.44	22.2	29.2	17.4	26.4

3.2.3. Band-Pass Filters

Three band-pass filters (BPF) incorporated the following features:

- Technology: Surface Acoustic Wave (SAW) on 433 MHz and 868 MHz bands. Ceramic type for 2.4 GHz band. Both types provide excellent out-of-band rejection,
- Input and output impedances: 50 Ω, allowing direct impedance matching in the RF chain.
- Minimum out-of-band attenuation: 28 dB on 433 and 868 MHz bands. 15 dB on 2.4 GHz band.
- Maximum insertion loss: 2 dB (433 MHz band), 2.2 dB (868 MHz band), 1.3 dB (2.4 GHz band),
- Maximum in-band ripple: 0.3 dB (433 MHz band), 0.7 dB (868 MHz band), 0.8 dB (2.4 GHz band)
- Pass bandwidth: 10 MHz (433 MHz band), 25 MHz (868 MHz band), 100 MHz (2.4 GHz)

VNA measurements verified the properties and performance of each filter. Figure 3.16, Figure 3.17, and Figure 3.18 present their insertion loss, respectively. Markers indicate the peak insertion loss at the designated frequency and the frequencies of the 3 dB bandwidth. Table 3.3 shows the part numbers and summarises the performance of each filter. Test results validated that the selected components agreed with the specifications.

The simulation setups and frequency responses of the BPF are shown in Figure 3.19, Figure 3.20, and Figure 3.21 for each band, respectively. The filter simulation's purpose was to observe the insertion loss curves, as the filters influence the frequency response of the overall front-end. The simulated results generated a sharp filter response, insinuating highly selective BPF. Matlab RF Toolbox (version 4.5, R2023a) did not include a SAW filter model. Instead, the simulation exploited the Chebyshev filter. The out-of-band rejection shape was sufficient to indicate the expected filtering properties.

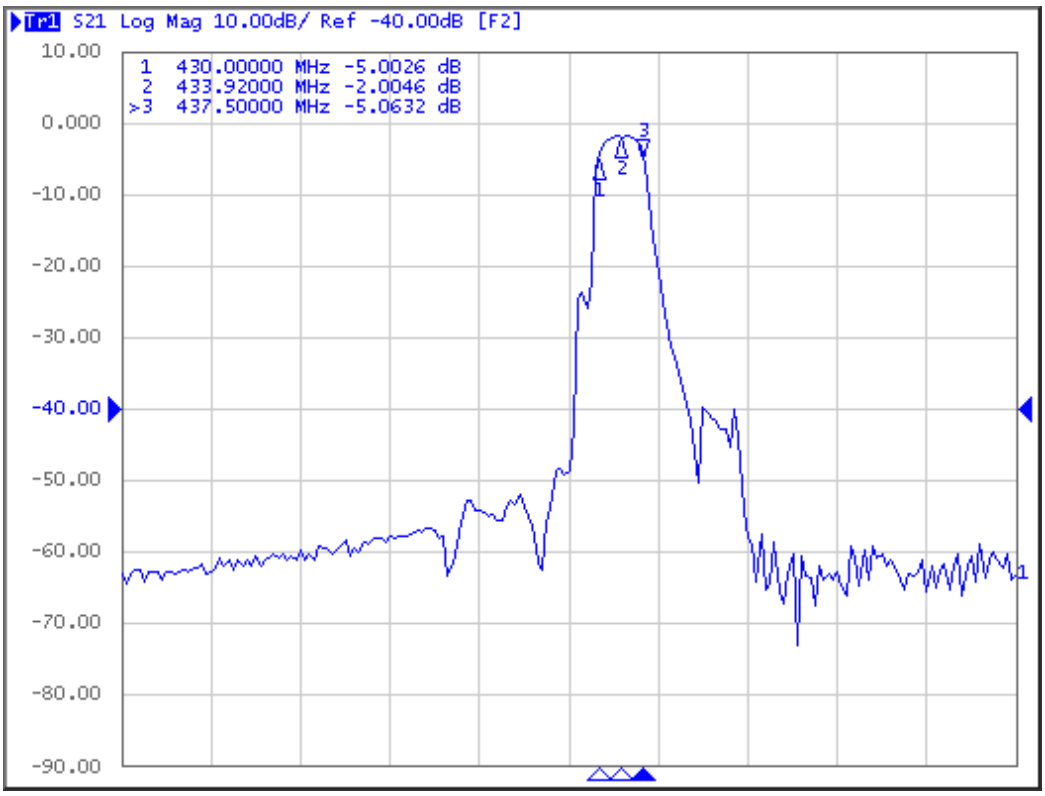

Figure 3.16: Insertion loss of the 433 MHz band-pass filter (Crystek CBPFS-0433)

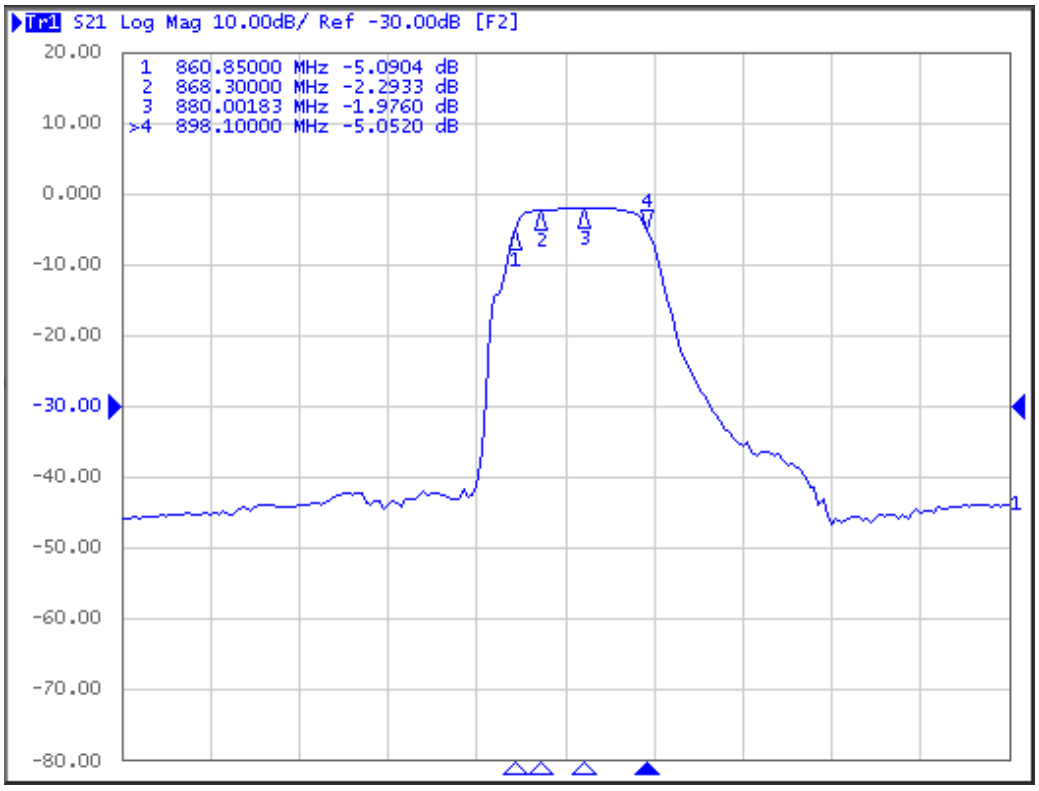

Figure 3.17: Insertion loss of the 868 MHz band-pass filter (Crystek CBPFS-0866)

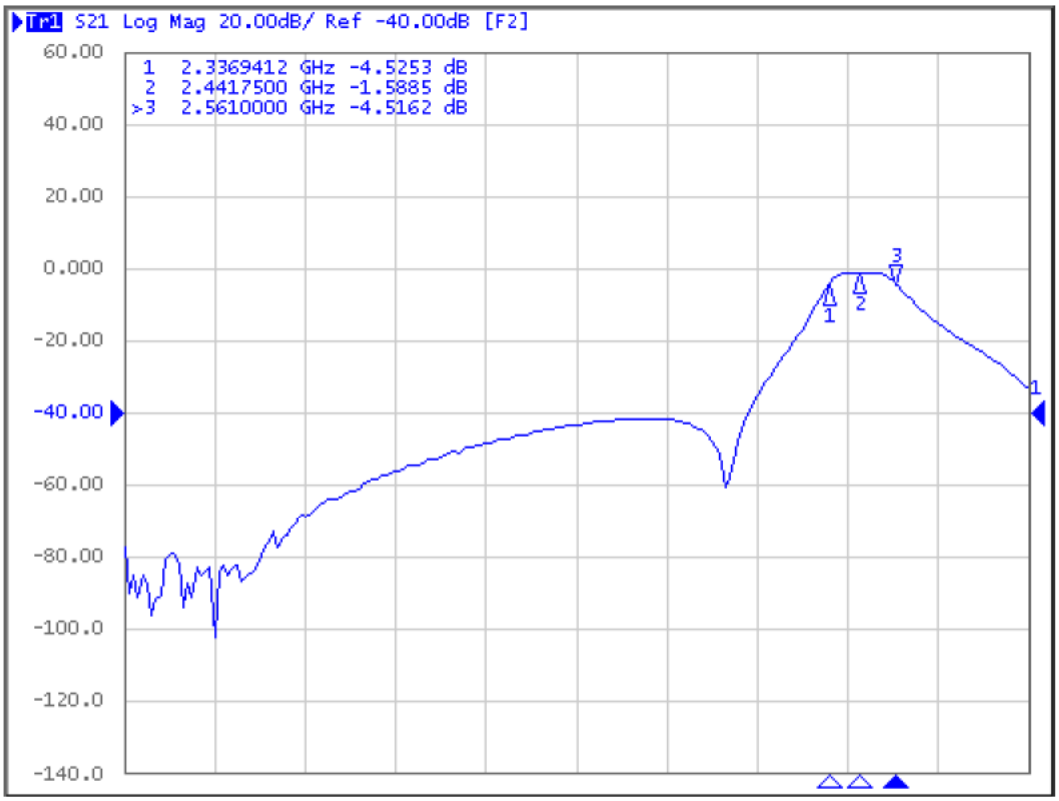

Figure 3.18: Insertion loss of the 2.4 GHz band-pass filter (Taoglas BPF.24.01)

Table 3.3: Performance and test results of the band-pass filter

Manufacturer	Model	Type	Test Frequency	Insertion Loss	3dB Bandwidth
Crystek	CBPFS-0433	SAW	433.92 MHz	2 dB	7.5 MHz (430.0 - 437.5 MHz)
Crystek	CBPFS-0866	SAW	868.30 MHz	2.29 dB	37.4 MHz (860.8 - 898.2 MHz)
Taoglas	BPF.24.01	Ceramic	2.44175 GHz	1.58 dB	0.224 GHz (2.337 – 2.561 GHz)

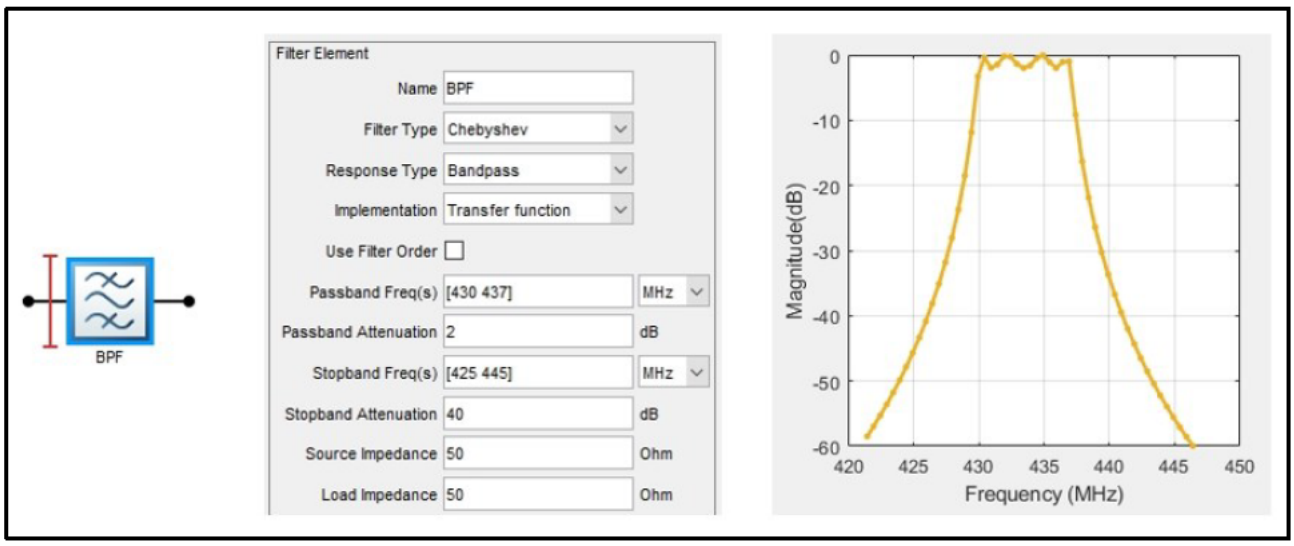

Figure 3.19: BPF simulation setup and frequency responses, 433 MHz

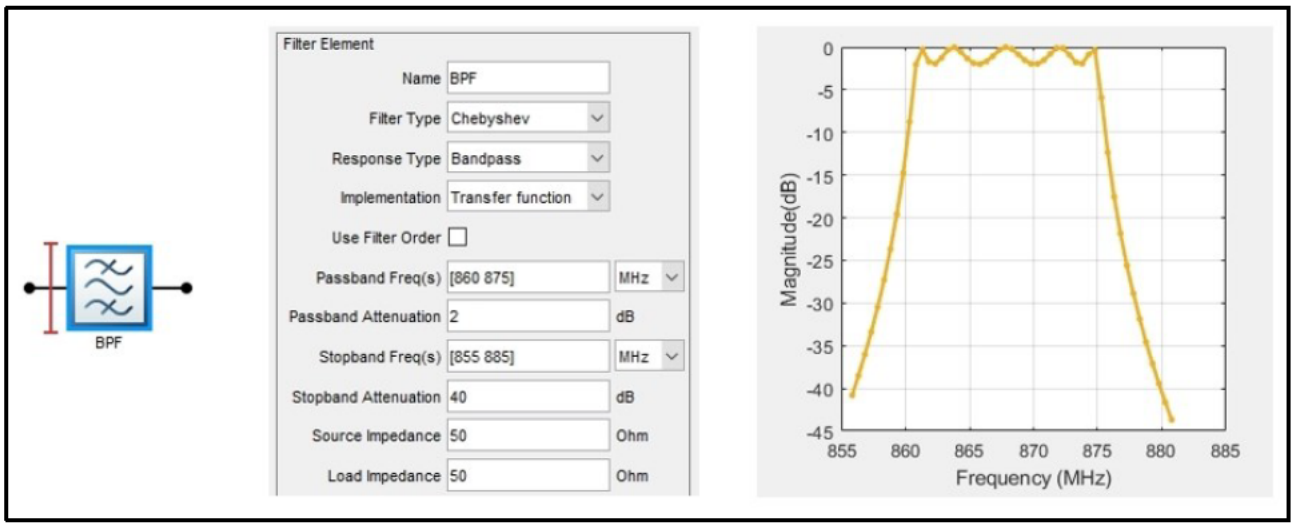

Figure 3.20: BPF simulation setup and frequency responses, 868 MHz

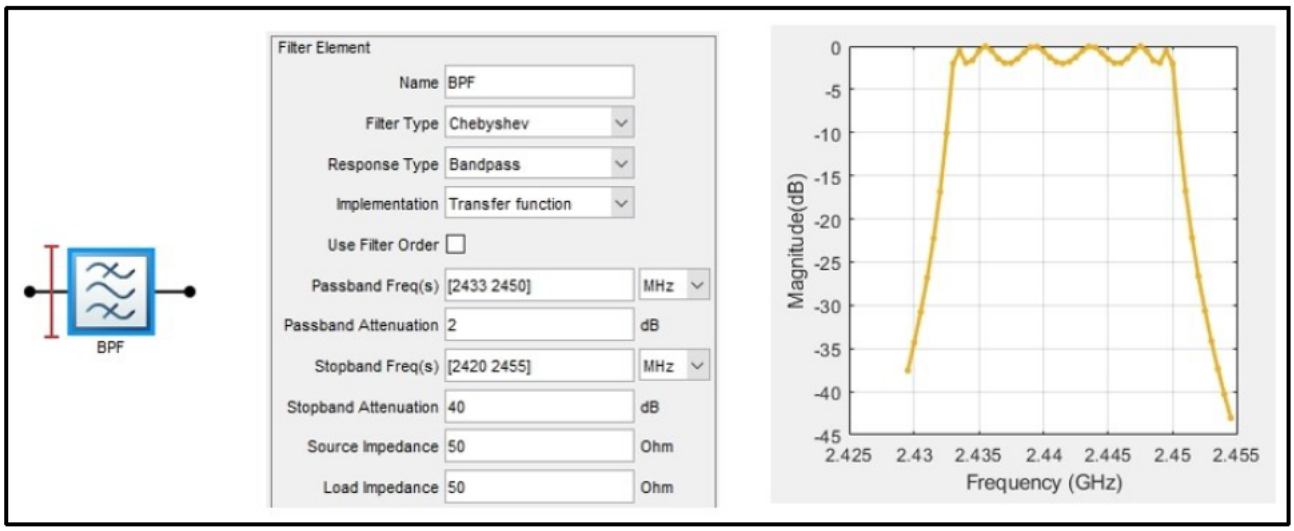

Figure 3.21: BPF simulation setup and frequency responses, 2.4 GHz

3.2.4. System Level Characterisation

The exercise was to determine the front-end system gain, noise figure and the minimum detectable signal (MDS). In Figure 3.1, the LNA was the first stage component; therefore, the LNA input was the reference point for system-level characterization.

3.2.4.1. RNSS Front-end System Gain

Figure 3.22 outlines the LNA, and the BPF cascaded in series to measure the total gain and estimate the front-end system noise figure. Figure 3.23 is the photograph of the actual front-end assembly.

Testing for each frequency required replacing the BPF. VNA port power was adjusted to -30 dBm to prevent the VNA input from overdriving. Figure 3.24, Figure 3.25, and Figure 3.26 illustrate the

measured results of the total front-end gain for 433 MHz, 868 MHz, and 2.4 GHz, respectively. These results demonstrate each component's contribution. The BPF dictates the system selectivity. The system frequency responses follow that of the BPF (discussed earlier). Measurements also demonstrate that the front-end amplification is proportional to the LNA gain associated with the filter's insertion losses. Table 3.4 shows comparable front-end data between calculations, measurements, and simulations.

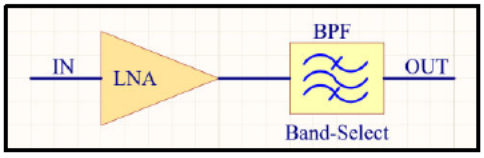

Figure 3.22: Cascaded System LNA and BPF

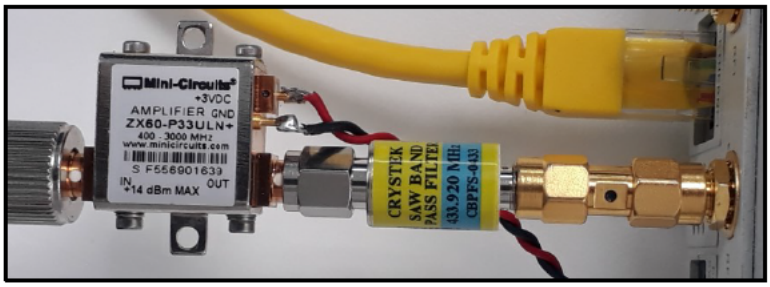

Figure 3.23: Fron-end system assembly

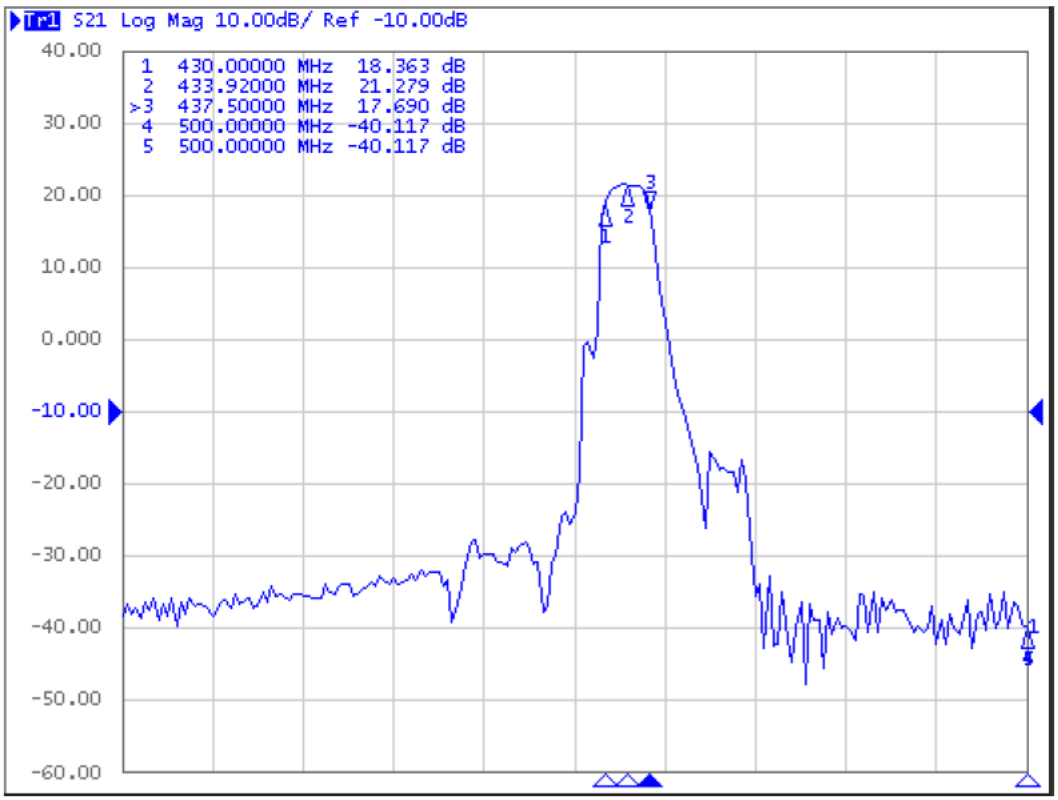

Figure 3.24: Cascaded system total gain 433 MHz

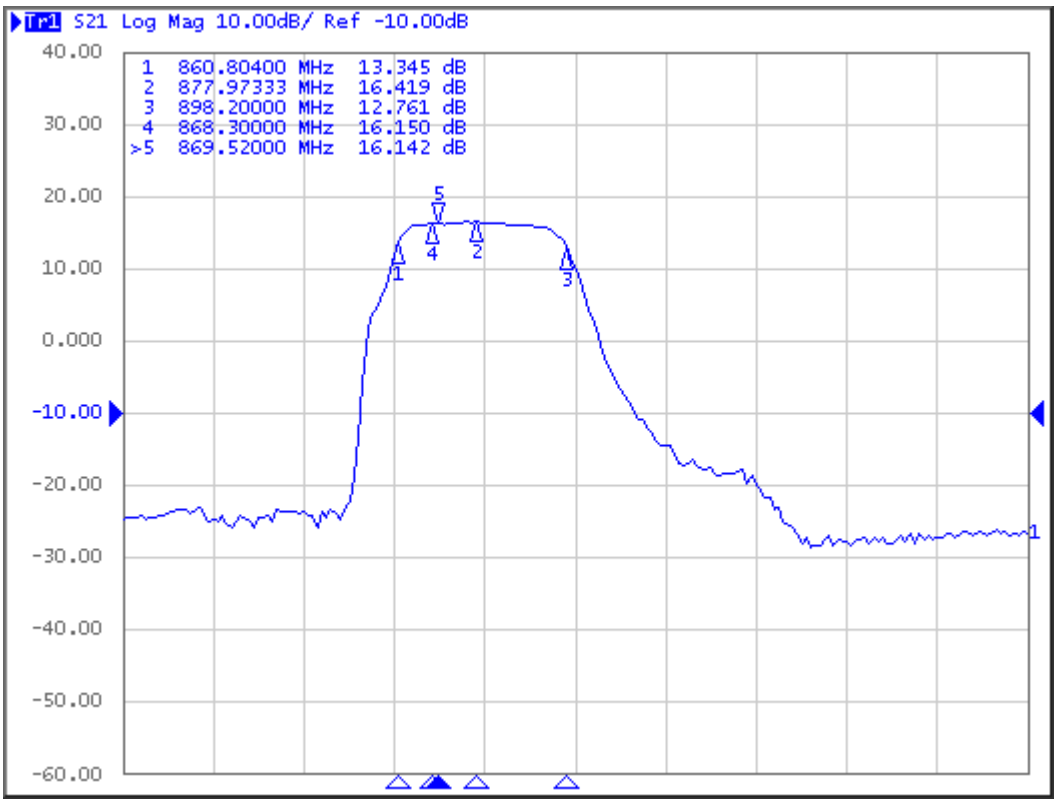

Figure 3.25: Cascaded system total gain 868 MHz

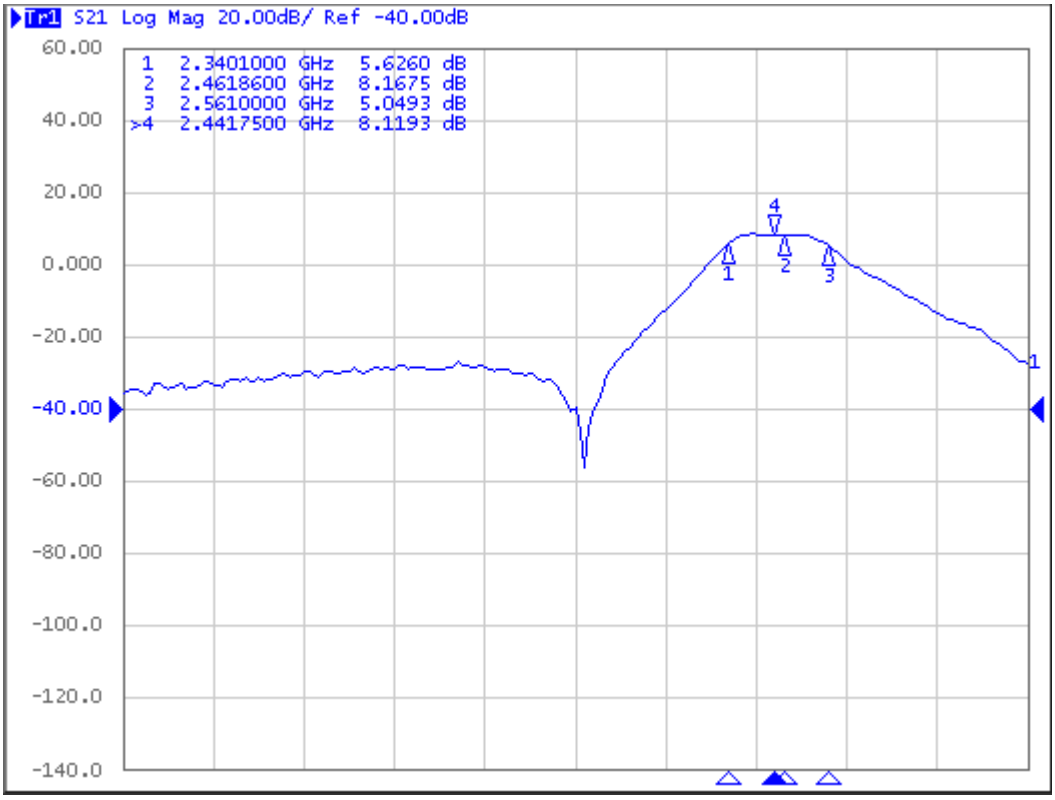

Figure 3.26: Cascaded system total gain 2.4 GHz

The RF Signal Generator and Spectrum Analyser verified the front-end total gain. Figure 3.27 illustrates the test setup for this verification exercise. Figure 3.28, Figure 3.29, and Figure 3.30, respectively, show the measured results on the spectrum: 11.48 dBm (433 MHz), 6.24 dBm (868 MHz), and -1.76 dBm (2.4 GHz). With known input power and measured output levels, computation makes the system gain simple.

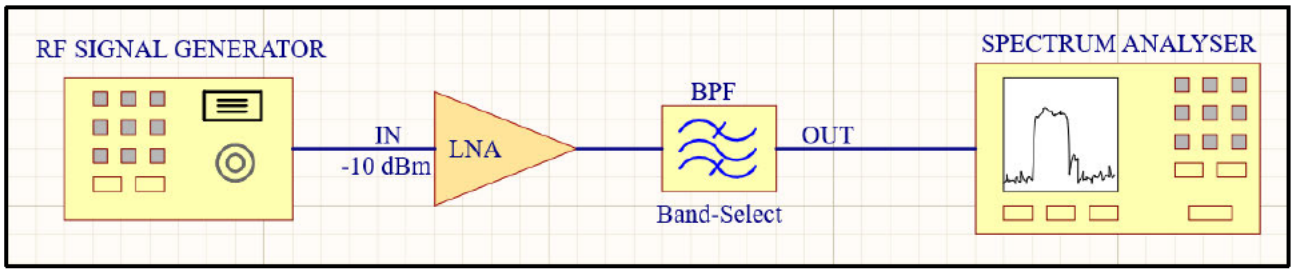

Figure 3.27: Total gain test using RF Signal Generator and Spectrum Analyser

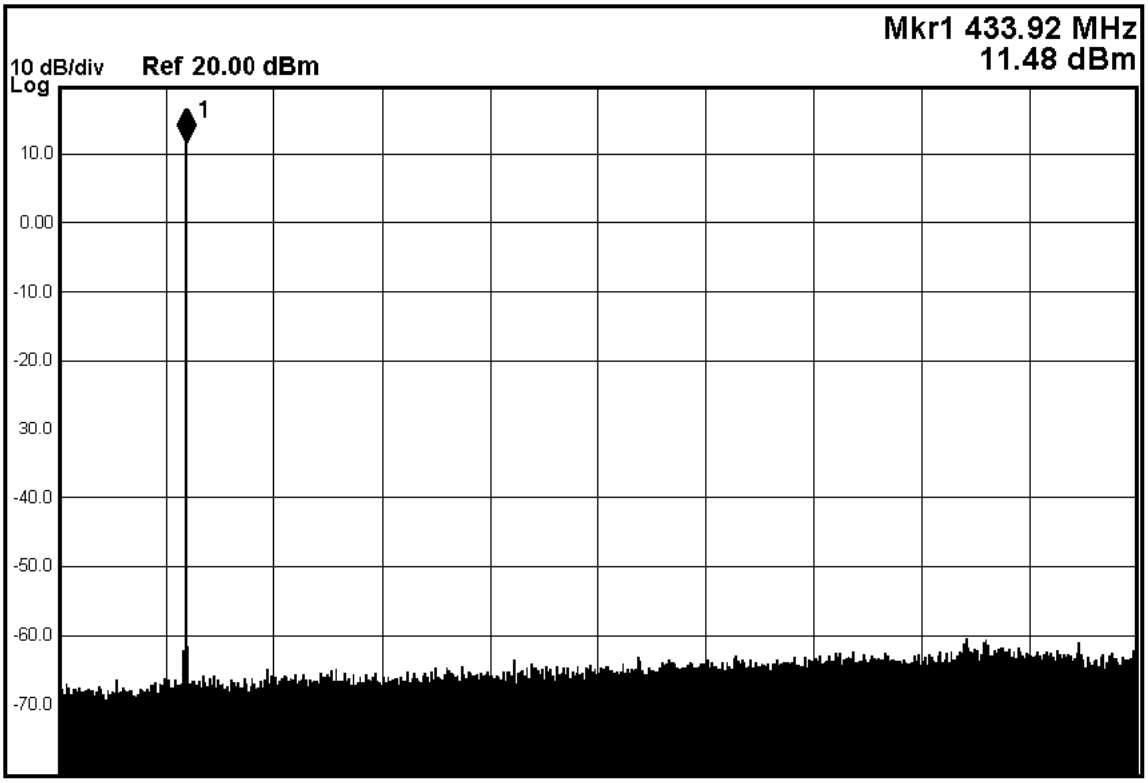

Figure 3.28: System output power level at 433 MHz

Simulation setups integrating the cascaded LNA and BPF are shown in Figure 3.31, Figure 3.32, and Figure 3.33 for each frequency, respectively. Settings applied from individual simulations were unchanged. As observed with the measured results, the simulation results corroborate each component's influence: the LNA contributed by amplifying the input stimulus (-10 dBm) by the gain value. The BPF dictated the shape of the frequency response and the out-of-band rejection.

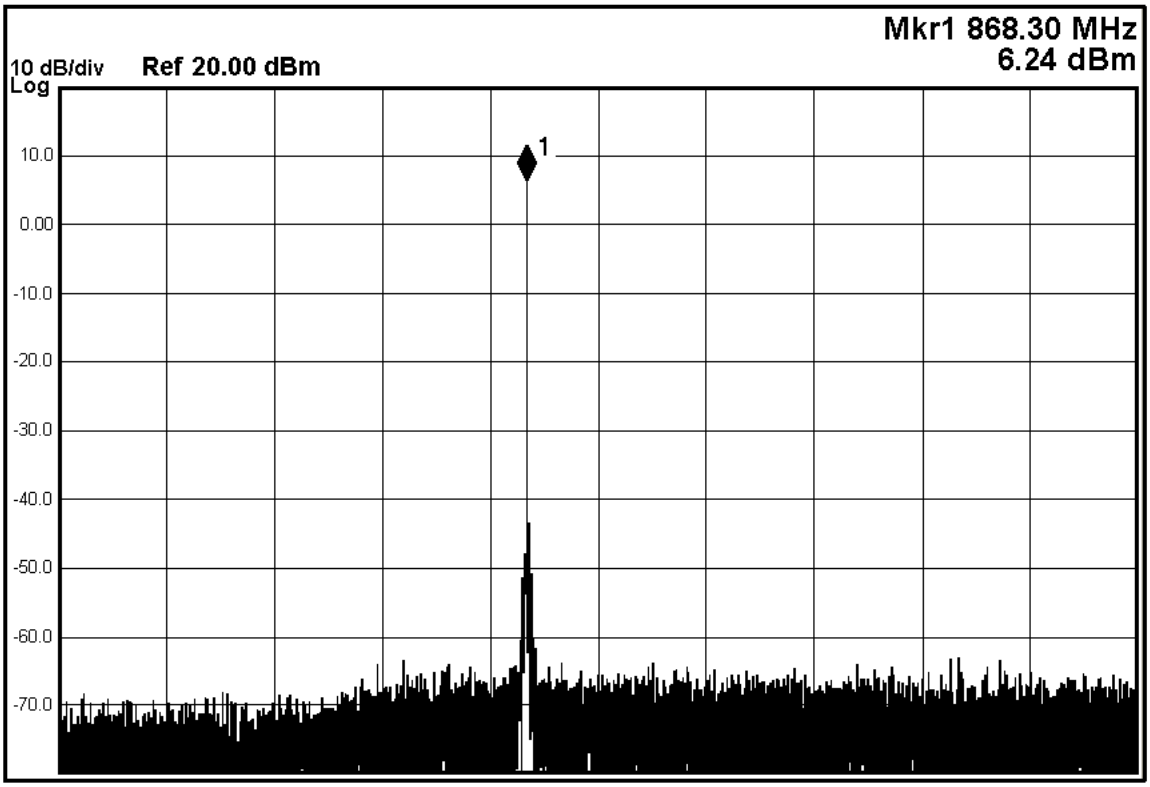

Figure 3.29: System output power level at 868 MHz

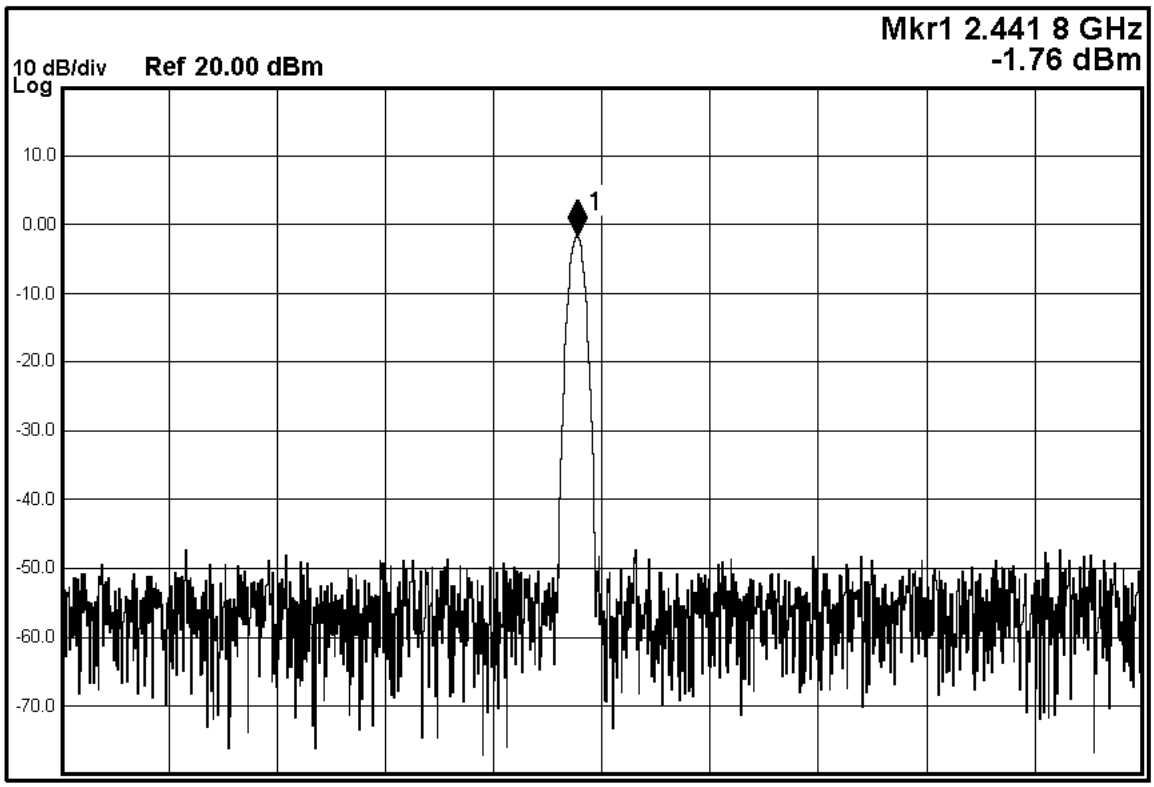

Figure 3.30: System output power level at 2.4 GHz

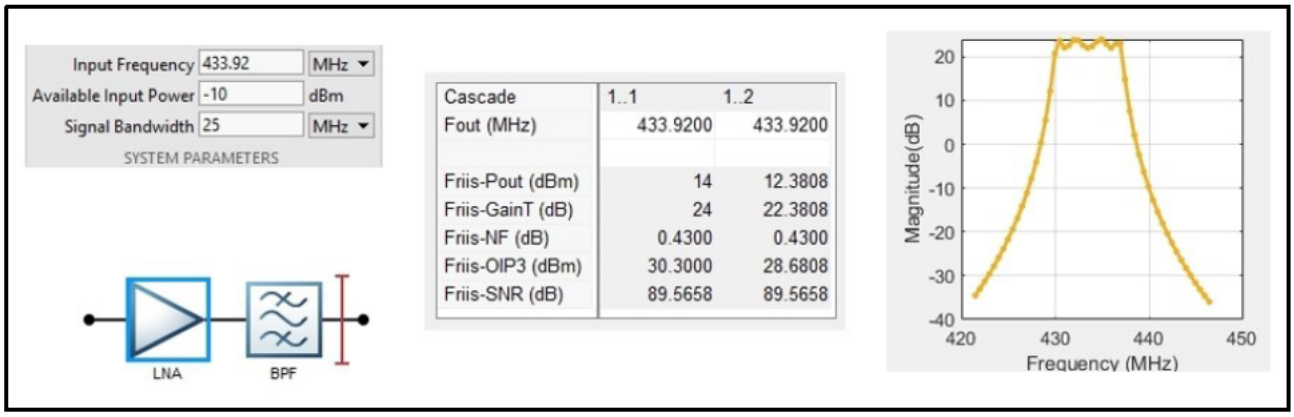

Figure 3.31: Front-end system level simulation, 433 MHz

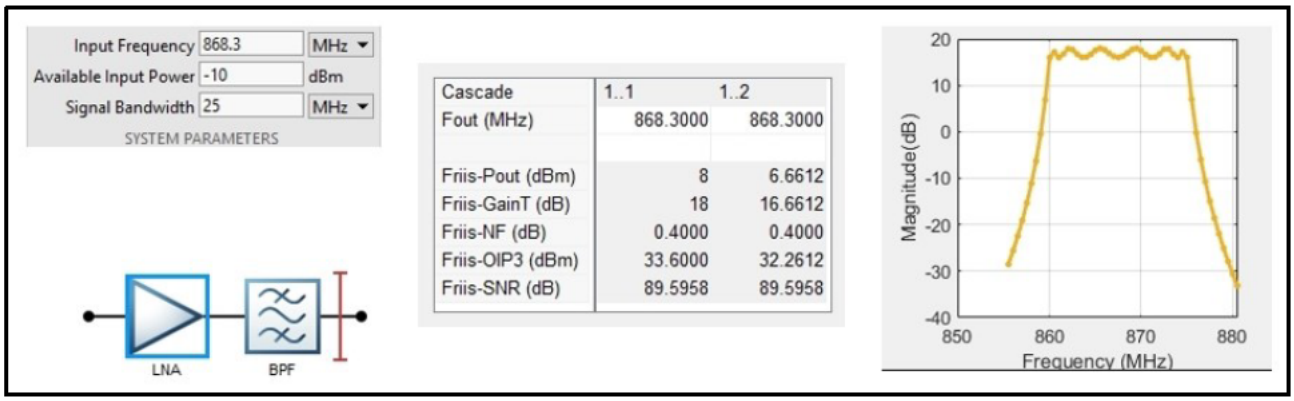

Figure 3.32: Front-end system level simulation, 868 MHz

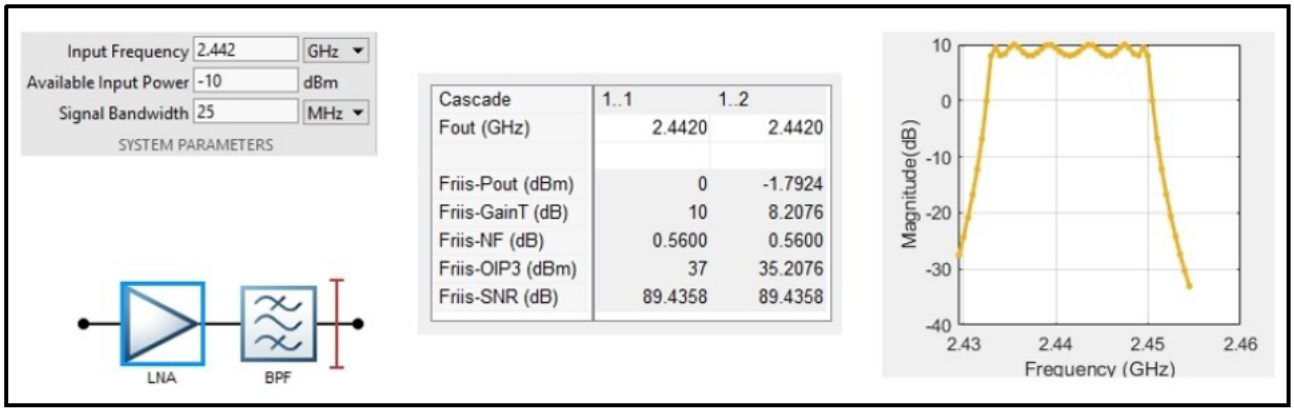

Figure 3.33: Front-end system level simulation, 2.4 GHz

Table 3.4: Cascaded system total gain - Calculated versus Measured

Frequency (MHz)	Input (dBm)	Output (dBm)		Gain (dB)			SNR (dB)	OIP3 (dBm)
		Meas.	Sim.	Calc.	Meas.	Sim.	Simulated only	
433	-10	11.48	12.38	21.48	21.279	22.38	89.57	28.68
868	-10	6.24	6.66	16.24	16.150	16.66	89.56	32.26
2400	-10	-1.76	-1.79	8.24	8.119	8.2	89.43	35.21

3.2.4.2. RNSS Front-end System Noise Figure

Equation (3.4) is the Friis formula for a two-stage cascade system applicable to this RF front-end.

$$F = F_1 + \frac{F_2 - 1}{G_1} \tag{3.4}$$

F_1 and G_1 are the first stage (LNA) noise factor and gain, respectively. F_2 is the second stage (BPF) noise factor.

As explained by Steer [88], for passive devices such as filters, the insertion loss determines the noise figure. Table 3.5 shows the calculated noise figures for each frequency considering measured values for LNA gains and BPF insertion losses. The LNA datasheet provides information on the noise figure.

The total system front-end exhibited an excellent noise figure of less than 1 dB across the three frequencies of interest. This result confirmed the theory that in a cascade system, the system noise figure only depended on the first element, the LNA. As a follower stage, the effect of BPF on noise figures was negligible.

Table 3.5: Total system analysis on noise figure and gain

Frequency	Stage 1 (LNA)	Stage 2 (BPF)	Total	
	Calculated	Calculated	Calculated	Simulated
433 MHz	G_1 = 23.546 dB NF_1 = 0.43 dB	G_2 = -2 dB NF_2 = 2 dB	G_T = 21.55 dB NF_T = 0.44 dB	G_T = 22.38 dB NF_T = 0.43 dB
868 MHz	G_1 = 18.847 dB NF_1 = 0.38 dB	G_2 = -2.29 dB NF_2 = 2.29 dB	G_T = 16.56 dB NF_T = 0.42 dB	G_T = 16.66 dB NF_T = 0.40 dB
2.4 GHz	G_1 = 10.298 dB NF_1 = 0.55 dB	G_2 = -1.58 dB NF_2 = 1.58	G_T = 8.72 dB NF_T = 0.70 dB	G_T = 8.20 dB NF_T = 0.56 dB

3.2.4.3. RNSS Front-end Minimum Detectable Signal (MDS)

Equation (3.5) defines the receiver mds as a set margin m above the internal noise power (see Equation (2.20)). Equation (3.6) and Equation (3.7) demonstrate the logarithmic transformation for MDS (expressed in dBm) [89].

$$mds = p_{n_r} \cdot m = kT_0 b f_r \cdot m \tag{3.5}$$

$$MDS = 10\log\left(\frac{kT_0}{10^{-3}}\right) + 10\log(B) + F_r + m \tag{3.6}$$

$$MDS = -174 + 10\log(B) + F_r + m \tag{3.7}$$

p_{n_r}: receiver noise power; k : Boltzmann's constant (1.38×10^{-23} J/K); T_0: reference temperature (290 K); b: noise power bandwidth of the receiving system (Hz); f_r: receiver noise factor.

With B = 25 MHz (worst case scenario), noise figures from Table 3.5, and let m = 3 dB, Table 3.6 shows the calculated values for MDS.

For simulation, the models developed using Matlab RF Budget Analyzer were exported to Simulink Testbench in RF Blockset in Figure 3.34. As demonstrated in Equation 2.12, noise power is independent of frequency. Since the noise factor of LNA is frequency-related, the simulation stage had to evaluate the MDS for each band. Figure 3.35, Figure 3.36, and Figure 3.37 show the simulated results on each band, respectively. The MDS was the average from peak noise floor levels shown in markers P2 and P3. Marker P1 measures the peak power at the centre frequency (433 MHz, 868 MHz, or 2.4 GHz). Table 3.6 compares the calculated versus the simulated MDS and the power levels at the centre frequencies. The results correlate with the expectations.

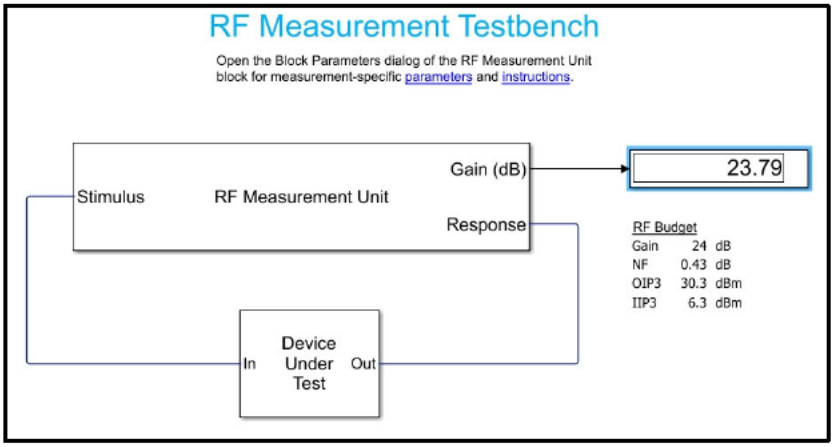

Figure 3.34: Front-end system - MDS simulation on Simulink Testbench

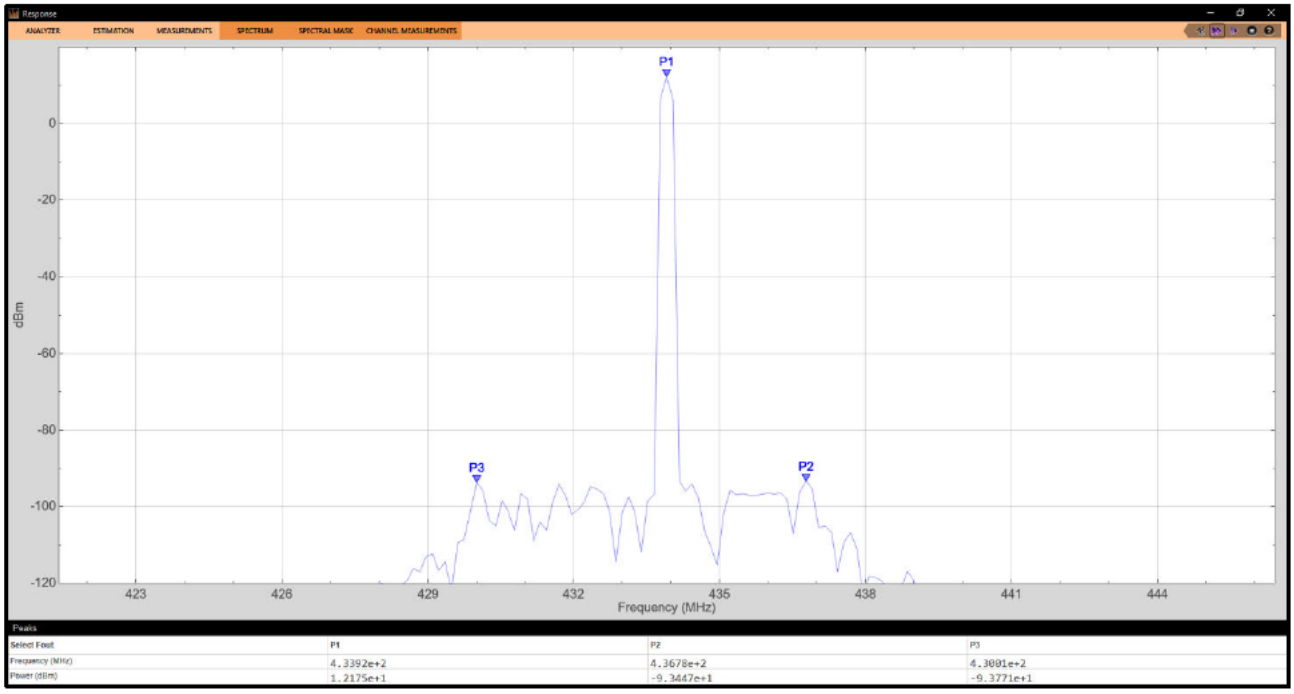

Figure 3.35: Front-end system - MDS simulation, 433 MHz

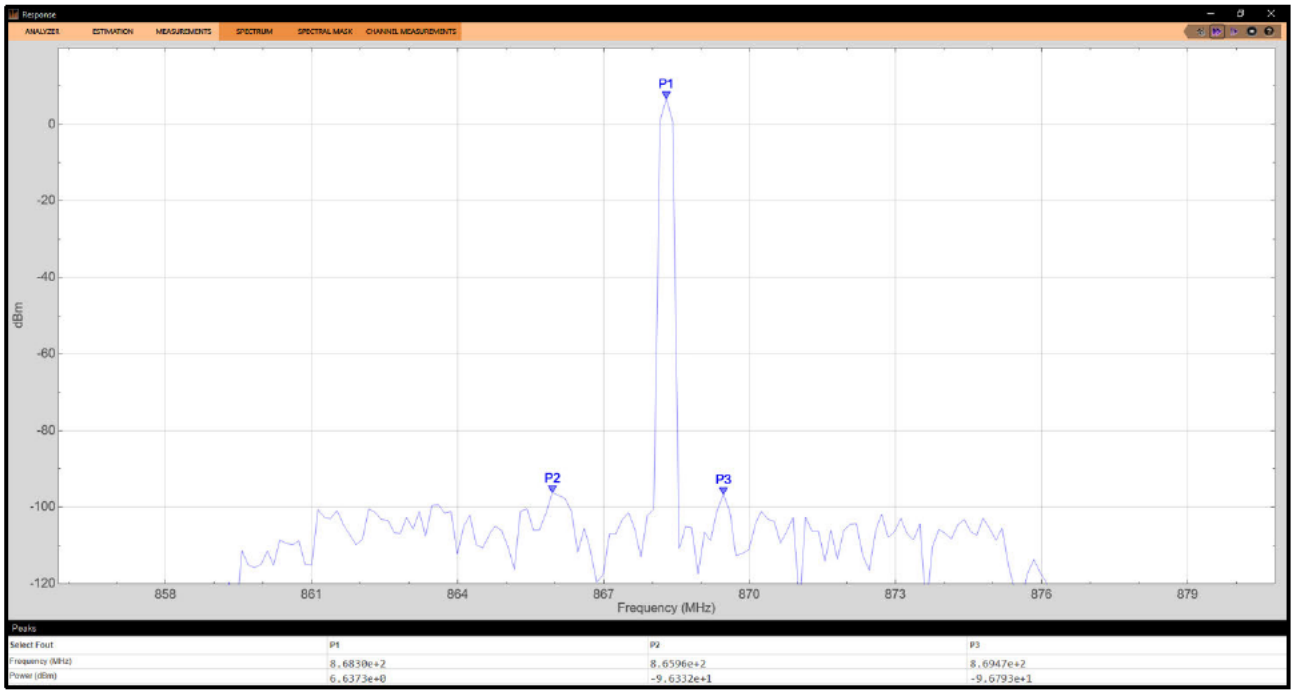

Figure 3.36: Front-end system - MDS simulation, 868 MHz

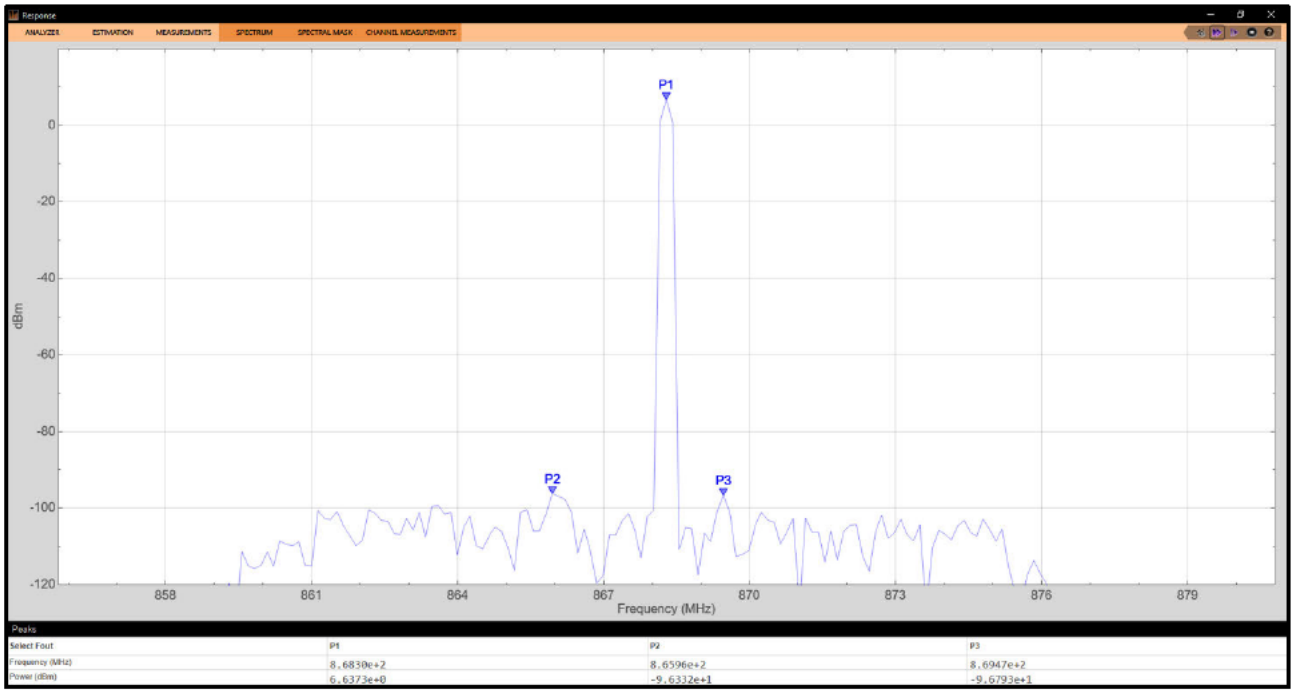

Figure 3.37: Front-end system - MDS simulation, 2.4 GHz

Table 3.6: System MDS

Frequency	MDS		Power	
	Calculated	Simulated	Calculated	Simulated
433 MHz	-96.59 dBm	-93.61 dBm	11.48 dBm	12.17 dBm
868 MHz	-96.62 dBm	-96.03 dBm	6.24 dBm	6.64 dBm
2.4 GHz	-96.46 dBm	-99.10 dBm	-1.76 dBm	-1.75 dBm

3.3. Software-Defined Radio (SDR)

The digital back-end system in this research comprised Software-Defined Radio (SDR) technology. SDR implements the radio function by software using digital signal processing technology on a programmable general-purpose hardware platform [90]. SDR is a radio system in which some or all the physical layer functions are software-defined, and this implies that the architecture is flexible so that the radio may be configurable. SDR is a multifunctional, programmable, and easy-to-upgrade radio that can support various services and standards while simultaneously providing a low-cost, power-efficient solution. An SDR Radio Receiver includes an RF Tuner, Analogue-to-Digital Converter (ADC), Digital Down Converter (DDC), and Digital Signal Processing (DSP) functions. The complete form of an SDR system requires the combination of software-based signal processing and the enabling hardware components [91]. Rather than using traditional chipsets, SDR has become increasingly popular due to its cost-effectiveness, capability and versatility [92].

There are several platforms developed for SDR based on their capabilities. The user will mainly be interested in the following features: programmability, flexibility, energy consumption and computing power. Choosing a computing platform for a given application is a trade-off between these cost functions. Some of these platforms are USRP, Quicksilver, SORA, KUAR, Texas Instruments C64+, Imec ADRES, Hiveflex, NXP EVP16, Infineon MuSIC, Sandblaster, ARDBEG, Tomahawk, Fujitsu SDR LSI, Imec BEAR, CEA Magali, XiSystem, WARP, WINC2R, Lyrtech, Picochip, AsAP, CEA Genepsy, eFalcon, COBRA, CORAL, BEE2. Different SDR software packages are available. Four of these are GNU Radio, IRIS, ASGARD, and OSSIE [93, 94]. The combination of GNU Radio and USRP is among the most popular SDR platforms of choice, especially among radio designers and researchers [95].

3.3.1. GNU Radio Software

GNU Radio software operates under the GNU General Public Licence (GPL). It is a free software development toolkit that provides signal processing blocks. It can be used with readily available low-cost external RF hardware to create software-defined radios or without hardware in a simulation-type environment. GNU Radio has a complete transmitter and receiver, a software-defined spectrum analyzer and oscilloscope, a multichannel receiver and a collection of modulators and demodulators. It covers applications such as mobile phones, radar, radio astronomy, RFID detectors, broadcast TV, radio receivers, and satellite navigation [91]. In this development, GNU Radio functioned alongside Ubuntu 20.0. The installation launches from the Ubuntu terminal [96].

3.3.2. USRP Hardware

Each USRP device comprises two main sub-devices: a motherboard and one among various available daughterboards. The primary objective of USRP motherboards is to convert analogue

Intermediate Frequency (IF) signals to digital baseband signals and vice versa. The daughterboards implement RF-to-IF conversion [97]. Ettus Research™ have developed different series of USRP hardware and different daughterboard kits. This study used the USRP2 motherboard and the SBX daughterboard. The USRP2 hardware features [98]:

- 50 MHz of RF bandwidth with 8-bit samples
- 25 MHz of RF bandwidth with 16-bit samples
- Gigabit Ethernet connectivity
- MIMO capable
- FPGA processing: Xilinx Spartan XC3S2000
- ADCs: 14-bits 100 MS/s
- DACs: 16-bits 400 MS/s
- Ability to lock to external 5 or 10 MHz clock reference.
- 1 daughterboard slot (1 RX + 1 TX connectors)

The SBX is a wide bandwidth transceiver that provides [99]:

- Frequency Range: 400 MHz - 4.4 GHz
- RF Output Power: 100 mW (20 dBm)
- RF Input Power: -15 dBm (max)
- Two antenna ports

Figure 3.38 outlines the front panel of the USRP2 device. Figure 3.39 shows the SBX daughterboard.

Figure 3.38: USRP2 hardware front panel

The installation of GNU Radio includes the USRP Hardware Driver (UHD), which supports application development for all USRP SDR products. UHD also offers cross-platform support for multiple industry-standard development environments and frameworks, including RFNoC, GNU Radio, LabVIEW, MATLAB and Simulink. It is available on Linux, Windows, and Mac OS [100].

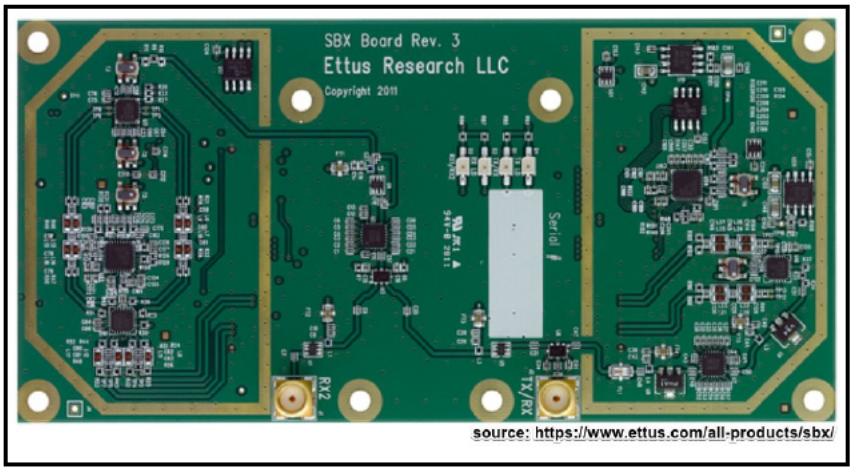

Figure 3.39: SBX daughterboard

The firmware for USRP2 boots from a memory card. A standard 2GB SD card is sufficient. Experiments demonstrated that High Capacity (SDHC), Class 4, and Class 10 memory cards may not work with the USRP2. Shell Terminal script[3], available for download, explains the Gnu Radio installation steps, the USRP2 SD card firmware burning process, the network settings, and testing the USRP2 configuration.

3.3.3. GNU-Radio Companion

The GNU Radio Companion (GRC), a built-in application of the GNU Radio software, was used to build the signal-processing components of the RNSS back end. Figure 3.40 illustrates the GRC flowgraph applied in this RNNS. The USRP Source block configures the network settings, RF options, sampling rate, and data type. The gain was set to 14 dB to improve the resolution on the graphical display. The output data type was a 32-bit complex floating-point number. The wire format was an 8-bit complex integer. The system sample rate was 25 Msps. The receiver bandwidth was 25 MHz. The QT GUI Sink block represented a spectrum analyzer used as a monitor to display the RF activities. The Complex-to-Real converter block extracts the real part of the complex quantity. The imaginary part was negligible. As explained by Ingala et al., only the real part is prevalent when evaluating the noise power of complex signals. The File Sink block writes binary data streams into files storable on PC hard drives [3].

3.4. Exploratory Data Analysis for Radio Noise

As an output of this research, a published conference article [3] provided insights for a baseline measurement of environmental radio noise. The study necessitated at this stage to experiment with real-world radio noise data to confirm that the system could detect external radio noise. This data collection campaign aimed to measure environmental radio noise in selected candidate sites listed in Table 3.7.

[3] https://tinyurl.com/595s7x5m

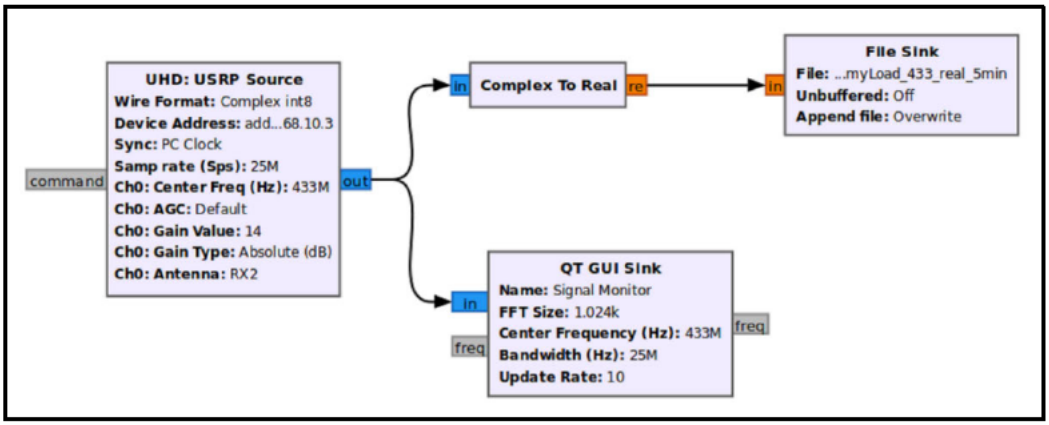

Figure 3.40: Flowchart of the RNSS

These data formed baseline resources to test and validate the post-processing methods through exploratory data analysis. The initial computational operation concerned reading the raw data stored as unformatted binary files. This chapter presented a brief theory about big data and the alternative approaches employed in this research to deal with big data.

Table 3.7: Surveyed candidate sites

Environment	Location	Situated	Date
Industrial	New Germany (NG)	Indoor	10/09/2021
	Westmead (WM)	Indoor	10/09/2021
Urban	Berea – DUT/Steve Biko (DUT)	Outdoor	10/09/2021
	Morningside (MS)	Outdoor	28/09/2021
Sub-urban	Montclair (MC)	Outdoor	22/03/2021
	Glenmore (GM)	Outdoor	21/08/2021

3.4.1. Facing Big Data

Several researchers and practitioners defined big data as a term used to refer to the increased volume of data that is difficult to store, process, and analyze through traditional database technologies. Big data implies data beyond technology's capability to store, manage, and process efficiently [101]. Big or large data means it is too big, fast, or hard for existing tools to process [102].

The continuous increase in the volume and detail of data captured by organizations, such as the rise of social media, the Internet of Things (IoT), and multimedia, has produced an overwhelming flow of data in structured or unstructured formats [101]. Scientific measurements are one of the generators of big data. They are often made at a high resolution and start to get huge when they involve two or three dimensions of space [103].

Addressing big data is a challenging and time-demanding task requiring extensive computational infrastructure to ensure successful data processing and analysis [101]. Computing applications

often experience limitations on the data size they can handle. Some software reads the entire datasets into memory and computes them in given memory locations. Constraints come when software must manipulate a massive chunk of data more prominent than the RAM. On most hardware platforms, there is a much harder limit on memory expansion than disk expansion: the motherboard has only so many slots to fill. Even under modern 64-bit operating systems, many applications today (for example, R under Windows) have only 32-bit executables and are limited to 4GB address spaces—this often translates into a 2- or 3GB working set limitation. Finally, even where a 64-bit binary is available—removing the absolute address space limitation—all too often, relics from the age of 32-bit code still pervade software, particularly in the use of 32-bit integers to index array elements. Thus, for example, 64-bit versions of R (available for Linux and Mac) use signed 32-bit integers to represent lengths, limiting data frames to at most 231–1, or about two billion rows. Even on a 64-bit system with sufficient RAM to hold the data, a 6.75-billion-row dataset is too big for R to manage. Any given computer has absolute and practical limits: memory size, disk size, and processor speed [103].

Cloud computing is one of the most significant shifts in modern ICT and service for enterprise applications and has become a robust architecture to perform large-scale and complex computing. The advantages of cloud computing include virtualized resources, parallel processing, security, and data service integration with scalable data storage. Cloud computing can not only minimize the cost and restriction for automation and computerization by individuals and enterprises. However, it can also provide reduced infrastructure maintenance costs, efficient management, and user access [101].

Distributed computing is the most successful strategy known for analysing massive datasets. This technique distributes the dataset to multiple machines to gain an advantage in parallel processing power [103].

3.4.2. Exploratory Data Analysis

3.4.2.1.Data Wrangling

With a five-minute observation, RNSS produced 30 GB of raw data containing about 10 billion rows of numerical linear values representing the amplitude of the ambient radio noise. Processing this amount of data added enormous strain on general-purpose computers. One of the issues was the computer suffering from out-of-memory. Cloud computing, multiple parallel, or distributed computing are costly, take time, and require effort. Following these techniques would defeat this research strategy aimed to be cost-effective and user-friendly. Alternative methods for data wrangling were necessary to prepare data for computing.

Exploratory data analysis employed Python with Numpy, Pandas, Scipy, Sklearn, Matplotlib, Seaborn libraries and JupyterLab IDE. MathWorks® explained several techniques to use memory efficiently [104, 105]. This study applied a data reduction approach to optimize memory usage with the following steps [3, 106]:

- The laptop upgraded to 48 GB RAM and 1 TB NVMe SSD
- Removing all the zeros because only non-zero signal levels were necessary for exploration.
- As shown in Figure 3.41, the last step was averaging by segmentation. The task required splitting the original raw data to create 10 million segments and averaging the sample values in each segment. The last step combines the mean values to form a new dataset for further processing.

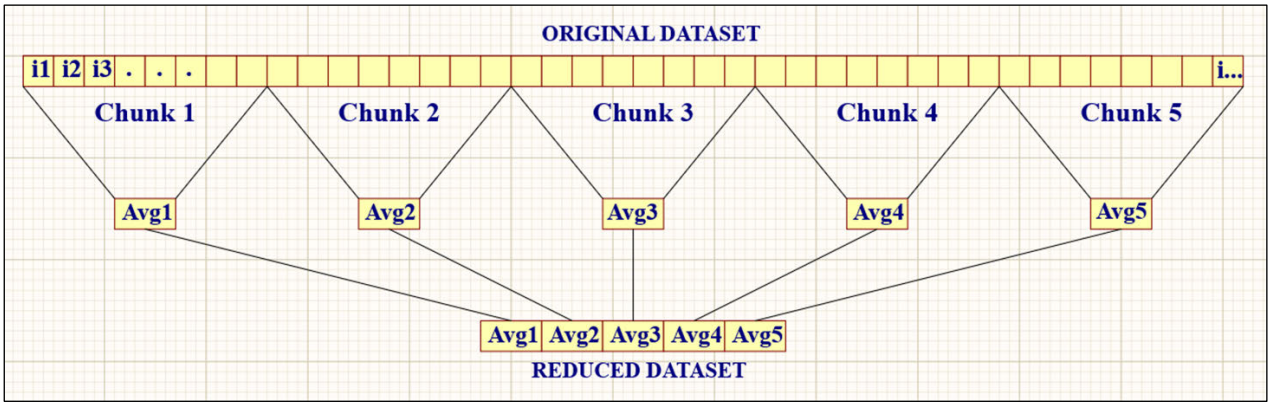

Figure 3.41: Data reduction method

This data reduction method transformed from raw data containing 10 billion rows (30 GB) to 10 million rows (30 MB), friendly for further computing.

3.4.2.2. Calibration

According to Skeie and Solberg [6], a noise measurement system needs to distinguish between the system's internal and external noise. The external noise from the antenna adds to the internal noise generated by the receiving system. The samples of the capture file will reflect the sum of these two contributions. The best measure of the level of the external noise component requires subtracting an estimate of the internal noise level from the total noise recorded. The calibration process must use a 50-ohm termination instead of the antenna to measure the internal noise level [6].

This study extracted the internal or equipment noise (EN) after replacing the antenna in the RF front end with a 50 Ω dummy load. Five minutes of recording was sufficient to collect a reasonable amount of data for equipment noise. A data reduction technique was also necessary. Equipment noise (EN), measured with the dummy load, was compared with the external ambient noise measured with the antenna per frequency band and location. Analysis for visualization involved the rolling average method [3].

Figure 3.42, Figure 3.43, and Figure 3.44 show that the amplitude of EN was lower compared to that of external noise in any frequency band. These results verified and validated that the RNSS could measure environmental radio noise and was suitable for this study.

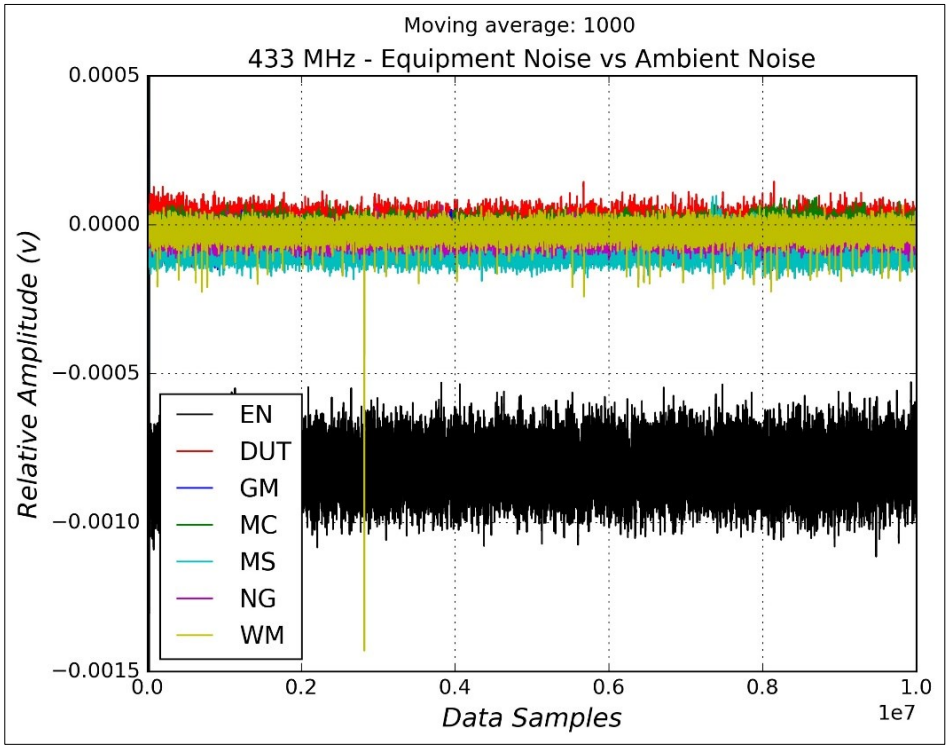

Figure 3.42: Equipment noise vs external noise – Band 433 MHz

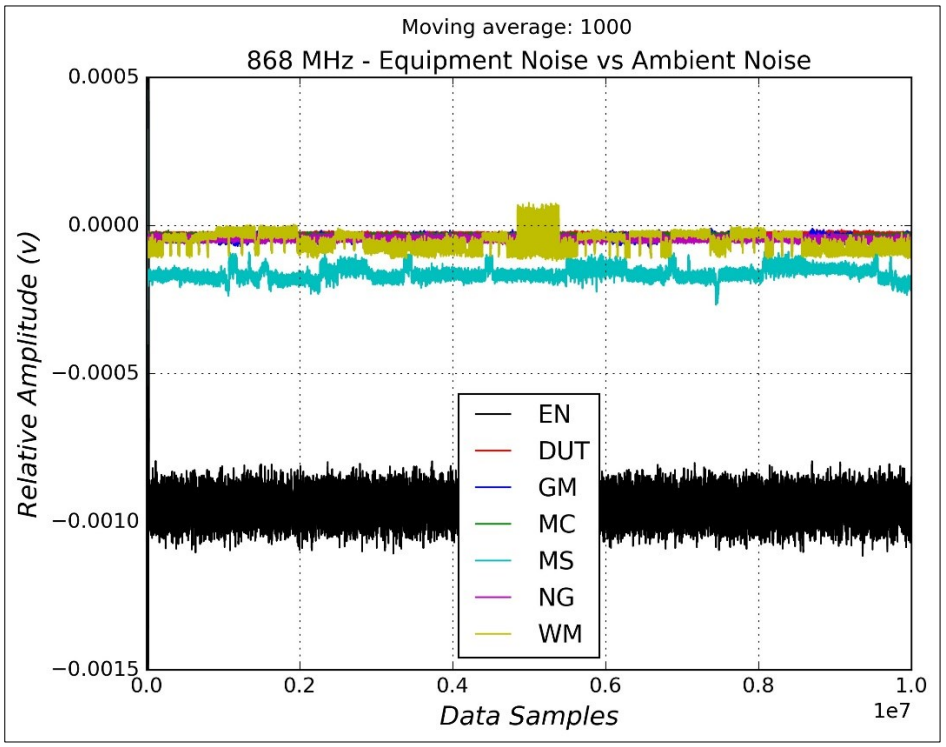

Figure 3.43: Equipment noise vs external noise – Band 868 MHz

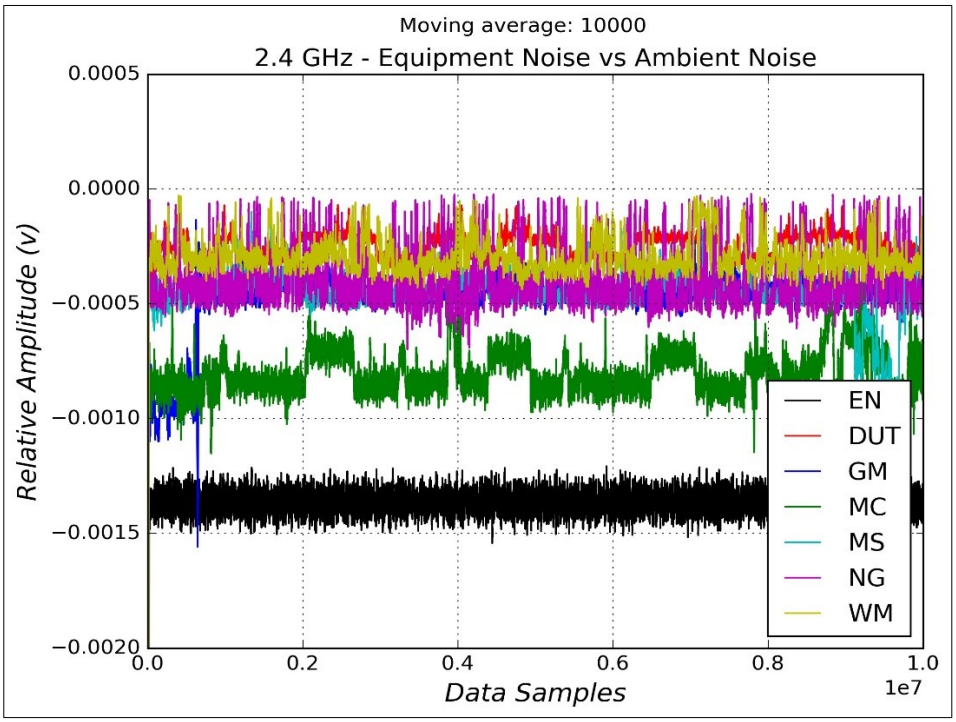

Figure 3.44: Equipment noise vs external noise – Band 2.4 GHz

3.4.2.3. Amplitude Correction

Amplitude correction factors included the LNA gain, the BPF insertion loss in the analogue front end, and the digital gain (14 dB) in the digital back end (USRP Source block). Table 3.8 displays the values of the amplitude factors computed as the difference between the front-end and back-end gains. The next post-processing operation shall subtract the amplitude correction values from the recorded data.

Table 3.8: Amplitude correction factors

Band	Gain Measured	USRP Source Gain	Amplitude Correction
433 MHz	21.3 dB	14 dB	7.3 dB
868 MHz	16.2 dB	14 dB	2.2 dB
2.4 GHz	8.1 dB	14 dB	-5.9 dB

3.4.3. Baseline Measurement of Ambient RF Noise

The analysis needed to stratify the collected raw data per frequency bands and locations. Data analysis considered data reduction method, calibration for equipment noise, and amplitude correction factors. Figure 3.45, Figure 3.46, and Figure 3.47, respectively, display the relative amplitude of environmental radio noise grouped per candidate sites and frequency bands. The results clearly inform the candidate site carrying the lowest ambient noise levels: Morningside at 433 MHz and 868 MHz and Montclair at 2.4 GHz. However, which site has the highest noise level

is not apparent. Hence, Table 3.9 shows the numerical averages for the relative amplitudes expressed in volts, revealing that DUT carried the highest levels of ambient radio noise [3].

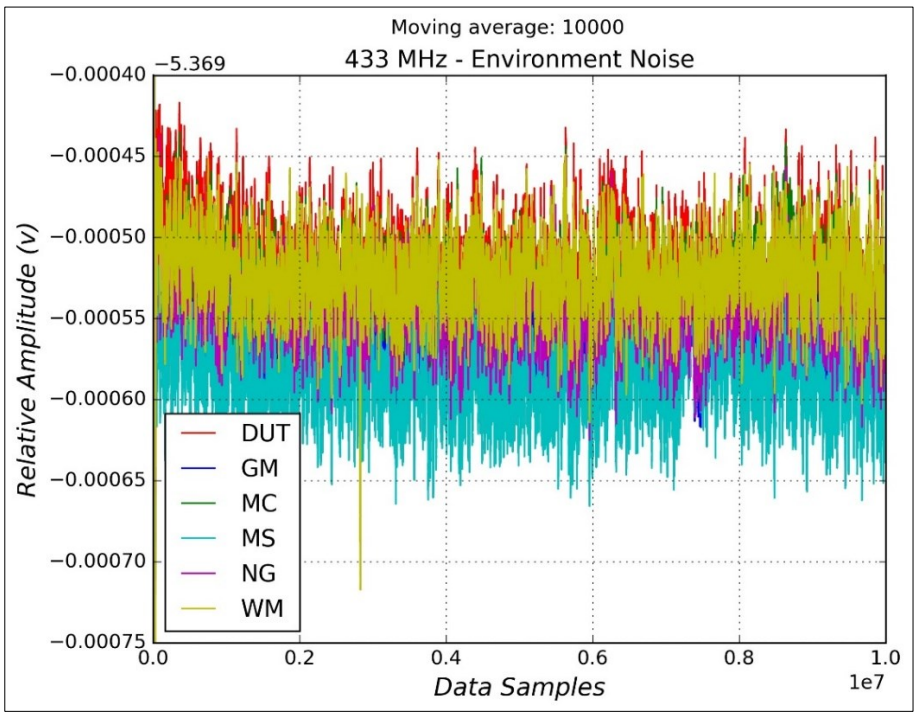

Figure 3.45: Radio noise environment – 433 MHz

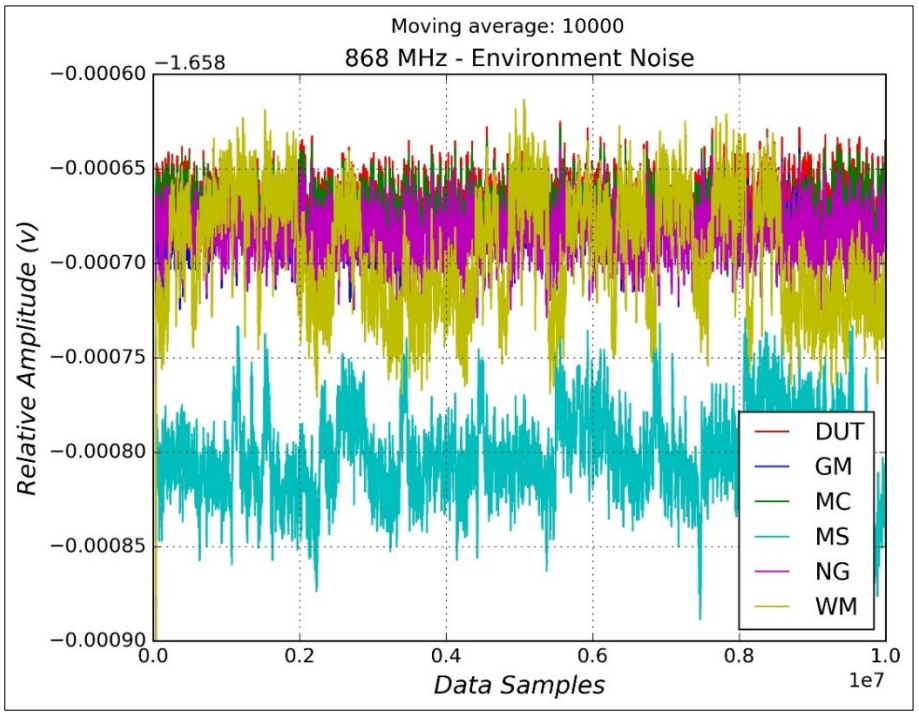

Figure 3.46: Radio noise environment – 868 MHz

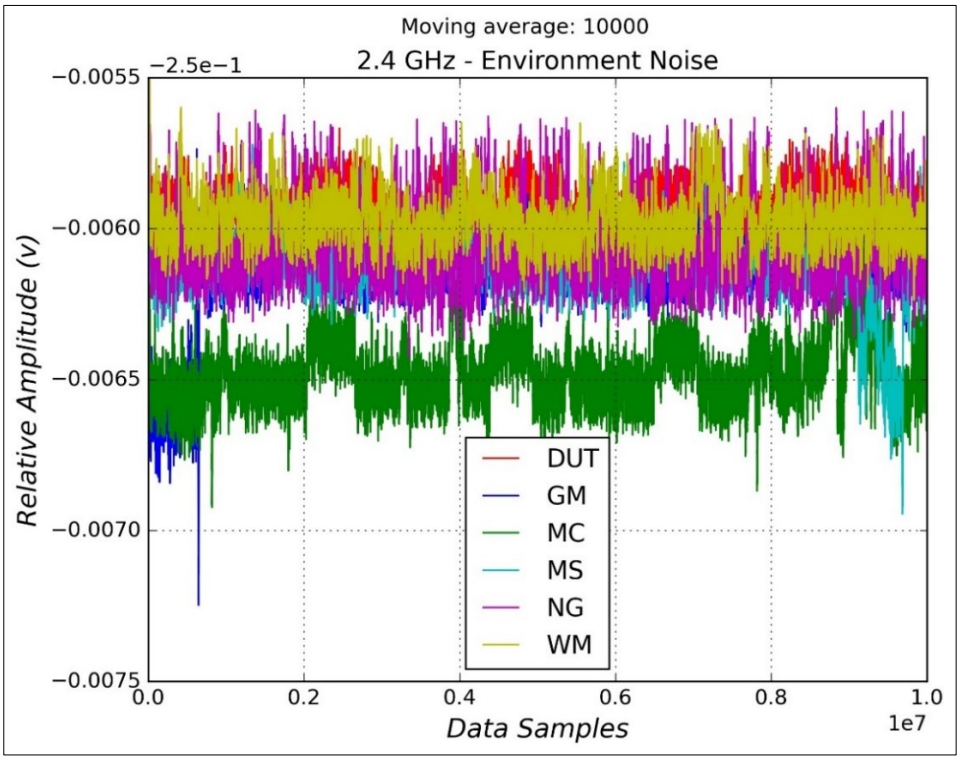

Figure 3.47: Radio noise environment – 2.4 GHz

Table 3.9: Amplitude Average values in volt for the environmental radio noise

Band	DUT	GM	MC	MS	NG	WM
433 MHz	-5.369517	-5.369559	-5.369533	-5.369589	-5.369541	-5.369527
868 MHz	-1.658671	-1.658685	-1.658677	-1.658804	-1.658686	-1.658697
2.4 GHz	-0.255924	-0.256146	-0.256475	-0.256129	-0.256056	-0.255966

3.5. Summary

Engineering activities in this chapter conducted laboratory RF testing as part of the component acceptance and validation processes. Results affirmed that all selected parts agreed with the specifications.

A noteworthy outcome found the level of equipment noise below the collected noise by the antenna. This result proved that the RNSS was the correct instrument for radio noise measurement. This chapter was a guideline for initial exploratory data analysis involving environmental radio noise data.

CHAPTER 4: DEVELOPMENT OF IOT RADIO NOISE GENERATORS

4.1. Introduction

This chapter introduces the IoT Noise Generators (ING) hardware and firmware development. In this thesis, the intention in designing ING units was to emulate the RF transmission behaviour for real-world IoT devices. Typically, IoT devices perform proprietary tasks following known specifications. Multiple IoT products may exist, but their standard function is RF operation. Hence, they all radiate radio noise in the environment. From an operational point of view, each IoT device transmits RF data at regular time intervals or on event-driven responses. From a noise surveying point of view, their RF transmissions appear randomly because each device transmits independently. The RF spectrum is, therefore, always occupied by random RF traffic. As a result, an environment surrounded by IoT gadgets would observe radio noise on aggregate to have continuous characteristics rather than impulsive [106].

4.2. Hardware

The following requirements guided the design of the ING units:

- generate random RF transmissions over time
- operate under three frequencies (433 MHz, 868 MHz and 2.4 GHz)
- comply with the ICASA regulations on unlicensed ISM bands
- be cost-effective
- be a low-power in current consumption

With these requirements in mind, the hardware[4] development followed the design strategies such as:

- Production of 3 PCB hardware variants:
  - Variant 1: 433 MHz
  - Variant 2: 868 MHz
  - Variant 3: 2.4 GHz
- Dedicated antenna on each variant
- Impedance matching circuits for each antenna
- Board variants to be distinguishable using labels on silkscreen
- Battery-operated device with replaceable CR2032 3V/225mAh coin-cell
- Board to include battery holder/retainer for CR2032 size.

[4] Link to download ING hardware files: schematics, board overlays, Gerbers, BOM : https://tinyurl.com/yr36kap7

- 3 Volts power supply and provision for the linear voltage regulator. The default bill of material (BOM) shall not count as this regulator.
- Texas Instruments CC1352R was the ideal processor. It consists of a Cortex M4 processor. The most intriguing feature was that it incorporates dual-band sub-1 GHz and 2.5 GHz RF modules, making it advantageous to cover all frequencies.
- The same component layout for 433 MHz and 868 MHz as the CC1352R has the same pinout for the sub-1GHz section. The 2.4 GHz section required a dedicated topology.
- Antennas to be custom-made as onboard PCB trace antennas
- Provision of two LEDs for battery power indicator and RF transmission feedback. The default BOM shall not include the power indicator LED to reduce power consumption
- Provision for push buttons to manipulate the device through different transmit rate modes.

Figure 4.1 shows the functional block diagram with compartments included in the design of ING units. Figure 4.2 shows the three designed boards after production assembly.

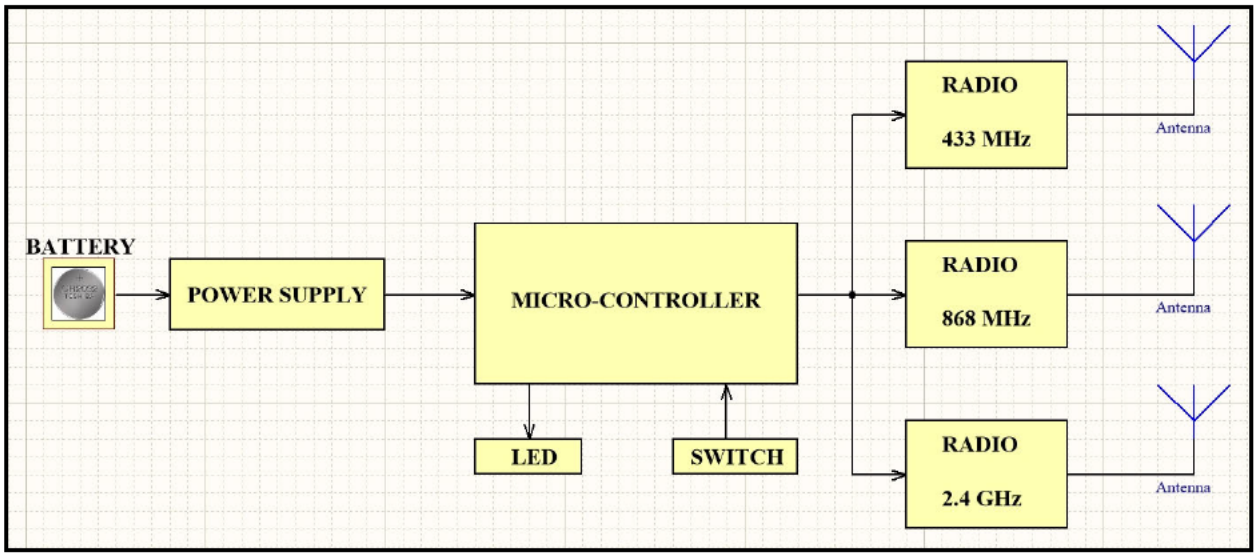

Figure 4.1: Functional Block Diagram of the IoT Noise Generator

4.3. Firmware

Development tools such as Code Composer Studio v8, Simplelink (built-in with the Code Composer Studio), and SmartRF Studio 7, all by Texas Instruments, were used to program the microcontroller for the following routines:

- The device was to start running on power-up after inserting the battery.
- The device continuously broadcasts on a set transmit rate interval.
- The transmitted signal carries random data with no user-specific payload.
- Transmitted burst to last 100 ms as airtime every cycle.

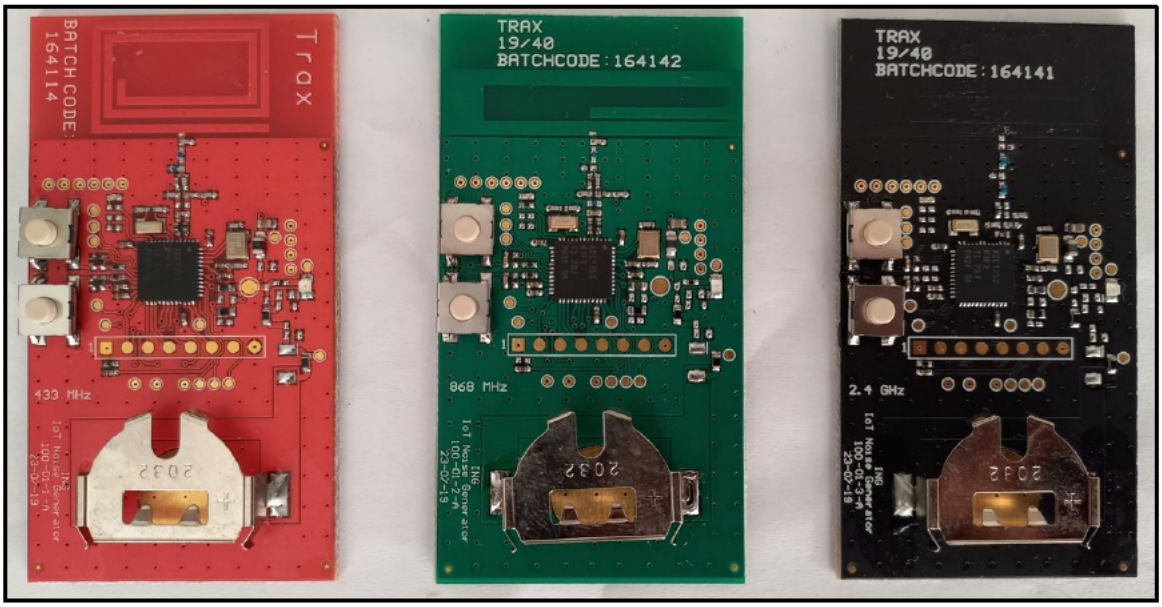

Figure 4.2: Assembled boards for ING units (Red: 433 MHz; Green: 868 MHz; Black: 2.4 GHz)

Figure 4.3 shows the development flowchart applied in the firmware[5] design. Each board variant was programmed using dedicated firmware based on the frequency band. Regarding functionality, the ING units must transmit modulated random RF signals at a controllable time interval called Transmit Rate. On power up, the module moved to radio silence mode of no radioactivity. The LED used to indicate the Transmit Rate remained off. On button press, the microcontroller increments the transmit rate counter by one, starting from transmit rate Tx0. The highest counter was 10 for the Transmit Rate Tx10. The transmit rate LED indicator had to flash with the number of flashes corresponding to the transmit rate counter. For example, in transmit rate Tx1, the LED would have blinked only once. After Tx10, the counter would overflow and reset to Tx0 (radio silence state). Table 4.1 displays the transmit time interval for individual Transmit Rates. The higher the transmit rate, the shorter the transmit interval and the more radio noise could propagate into free space. The process produced the ING capable of producing radio noise in the environment [106].

Table 4.1: Transmit rate and corresponding time interval

Transmit Rate	Transmit interval
Tx0	Idle
Tx1	5 sec
Tx2	4 sec
Tx3	3.5 sec
Tx4	3 sec
Tx5	2.5 sec
Tx6	2 sec
Tx7	1.5 sec
Tx8	1 sec
Tx9	0.5 sec
Tx10	0.1 sec

[5] Text file with firmware code is available here: https://tinyurl.com/3dvvdv8s

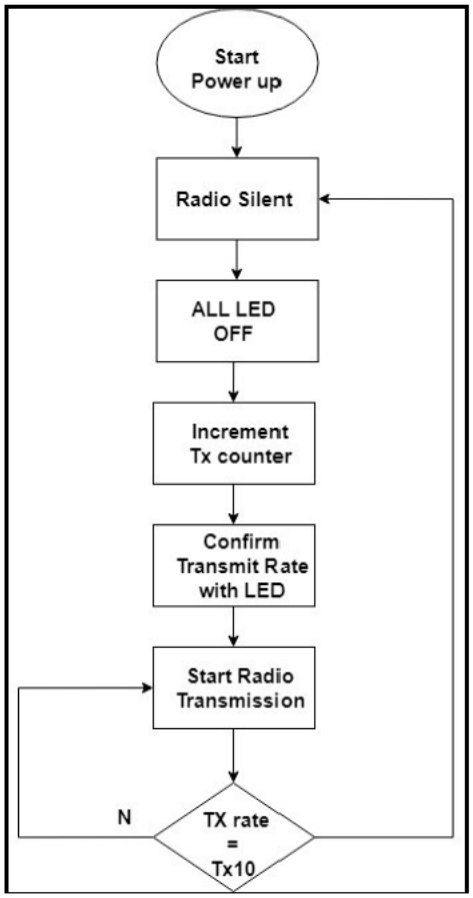

Figure 4.3: Firmware development flowchart

4.4. Environmental IoT Radio Noise Measurement Campaign

Successful manufacturing of ING units enhanced the equipment readiness for the in-situ data collection campaign. Radio transmissions from the ING devices mimicked the RF behaviours of IoT products. Data analysis demonstrated IoT radio activities' influence on the noise environment levels. As a proceeding of this research, a published journal article [106] highlighted the process in this measurement campaign.

4.5. Candidate Sites

This survey took place over five candidate sites representing the three targeted environment types. Table 4.2 shows the full names and abbreviations this thesis applied. This experiment occurred randomly at various times of the day.

Table 4.2: Candidate sites for IoT noise surveying campaign

Site	Description	Date
New Germany (NGM)	Industrial	30/06 - 04/07/2022
Berea Steve Biko Campus (DUT)	Urban	06/07 - 08/07/2022
Morningside (MNS)	Urban	19/08/2022
Montclair (MNC)	Suburban	16/07/2022
Glenmore (GLM)	Suburban	10/07/2022

A Google Earth map highlights in Figure 4.4 the actual locations for candidate sites, all from the KwaZulu Natal province in South Africa. Pointers (white-circled stars) indicate the GPS locations for each site (DUT, GLM, MNC, MNS, NGM). Figure 4.5 shows the equipment arrangement on-site during the campaign scenario.

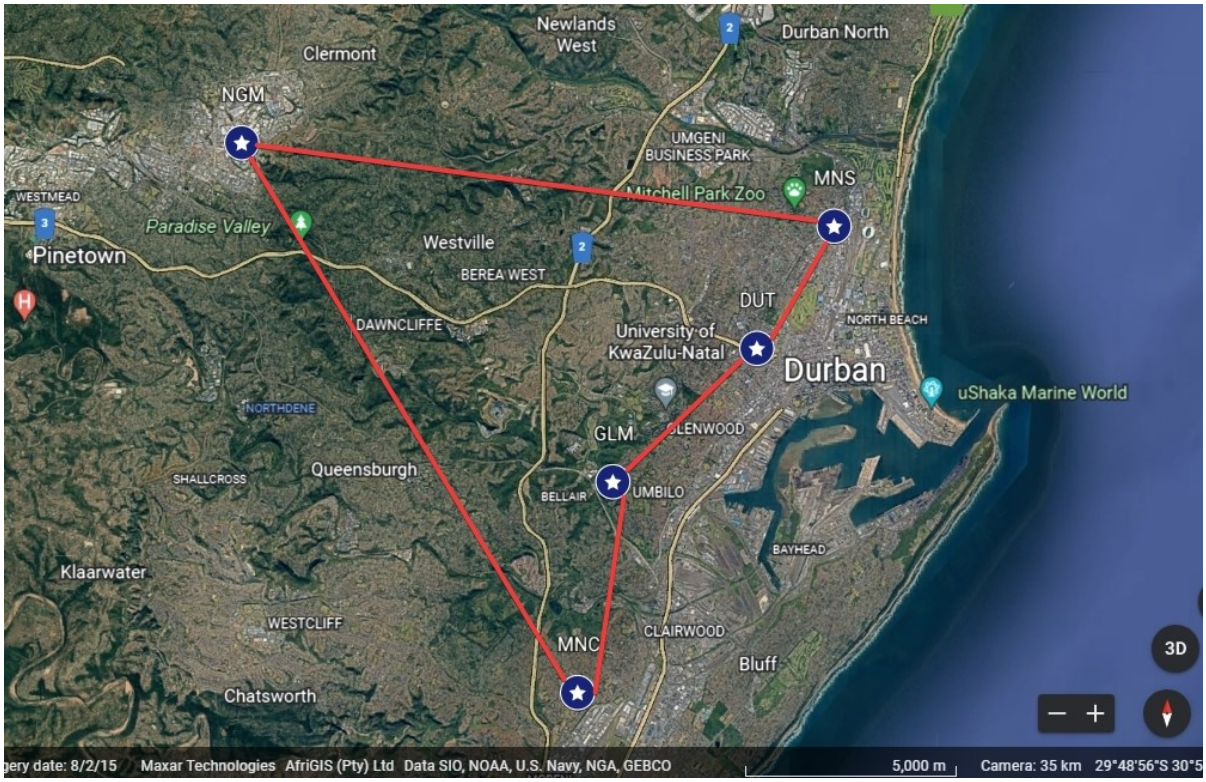

Figure 4.4: Candidate Sites on the map

Figure 4.5: Equipment setup on the fields

4.6. Operating Procedure on Site

Handling equipment led to repetitive tasks prone to introducing human errors. The research established standard operating procedures (SOP) to prevent unforeseen deviations. As in Figure 4.6, the SOP started by preparing the USRP hardware, including the USRP2 unit, power supply, and ethernet cable. The next stage performed data collection on each frequency band. Inserting the coin-cell battery was the only task to turn on the ING units. Some candidate sites did not have mains power (220V AC) available. Fully charged Uninterruptible Power Supply (UPS) secured up to 2 hours of power autonomy. Finally, computer work consisted of operating the GRC to change the frequency, composing the file name for recorded data (convention: *Site-name_Frequency_Tx-rate),* starting measurements, and stopping the recording after 5 minutes of observation. The routine needed to repeat itself 11 times to record raw data from transmit rates Tx0 to Tx10.

The following characteristics summarize the data requirements from this survey:

- 11 TX rates (Tx0 – Tx10)
- 33 raw datasets (over three frequency bands on every site)
- 165 raw datasets (obtained by five candidate sites)
- One raw dataset weighted about 30 GB in a 5-minute recording
- Required storage: 4.95 TB
- Recorded data type: unformatted binary file
- Estimated operation time: 3 hours spent on every site

4.7. Summary

This research developed, by hardware and firmware design, proprietary radio noise generators which successfully reproduced the IoT radio activities on fields. The manufacturing and production process availed 45 units of ING from 15 units per board variant. By modulating the transmit rates from Tx0 to Tx10, the ING units delivered different quantities of radio emissions into environments. The understanding can hypothesize that slower Transmit Rates imply a reduced environment radio occupancy in terms of power density. The faster transmit rates should induce an increased environment radio occupancy. Therefore, these concepts can help establish the relationship between the IoT industry's growth and ambient radio noise levels.

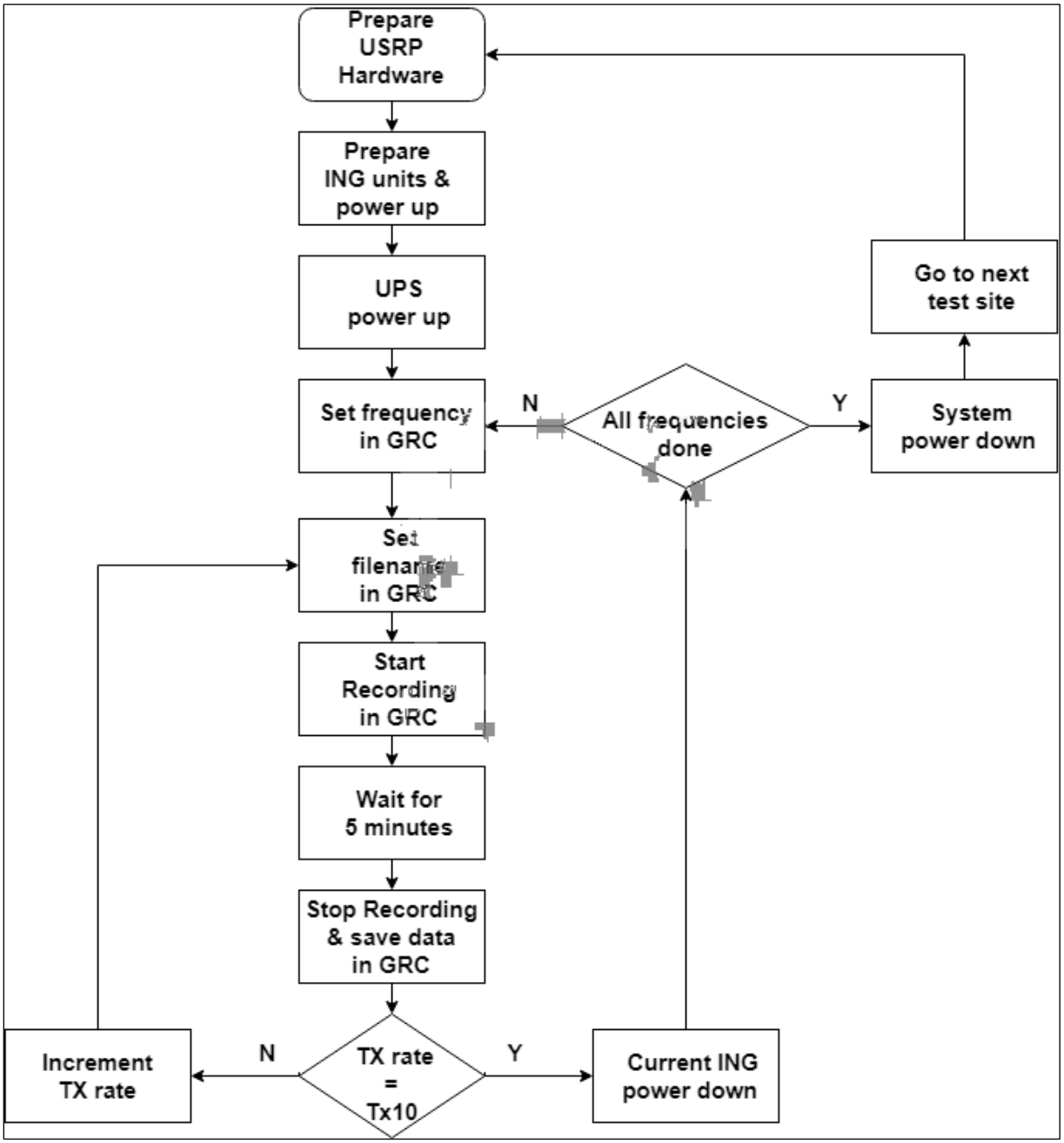

Figure 4.6: SOP followed at the candidate sites.

CHAPTER 5: STATISTICAL ANALYSIS OF RESULTS

5.1. Introduction

This analysis exploited data collected from the measurement campaign in Chapter 4. Wepman et al. [7] calculated the maximum, mean and median to analyse measurement data in their experiment. Based on descriptive statistical analysis, the investigation evaluated the mean, median, variance, standard deviation, kurtosis, skewness, and elements of boxplots (Q1 at 25th, Q2 at 50th, Q3 at 75th percentiles, interquartile range IQR, minimum and maximum fences/scores). This research applied the Kernel Density Estimate (KDE) function to evaluate the density distributions. Regression lines demonstrated the trend directions with the mean values. This study evaluated these metrics and grouped the results per candidate sites, frequency bands, and transmit rates.

5.2. Data Preparation

As in Chapter 3, the post-processing phase applied the same techniques: data reduction by data segmentation, calibration by removing equipment noise and subtracting any external gain values introduced in the RNSS front-end and back-end. The amplitude-corrected dataset contained positive as well as negative values. For convenience and a more straightforward result interpretation, it was necessary to shift all data values to the positive number range without affecting the relational proportions within datasets. This Chapter introduces amplitude scaling to give all samples positive values [106]. Literature in [106, 107] recommends adding a negligible value to avoid conflict with zero in mathematics. Post-processing implemented three steps to complete the amplitude scaling: 1- Identify the minimum value in the dataset and compute its absolute value denoted as $|X_{min}|$. 2- Calculate the scaling factor SF in Equation (5.1) to prevent mathematical errors with zeros. 3- In Equation (5.2), add the scaling factor to every sample in the dataset to make all quantities positive. The final dataset is relative power, P_r, expressed in milliwatts (mW).

$$SF = |X_{min}| + 0.000001 \tag{5.1}$$

$$P_r = (X_n + SF)^2 \times 10^3 \tag{5.2}$$

5.3. Statistical Analysis

This thesis applied descriptive statistical analysis to examine the concept of central tendency, variation, and shapes of distributions. This analysis explored the central tendency using the mean and the median because they define the centre concept. The study also examined the concept of variability, drift, or dispersion using variance and standard deviation [106, 108]. Finally, the investigation assessed the shapes of distributions using Fisher's measure of skewness and kurtosis

or the coefficient of excess. Skewness values = 0 indicate a symmetrical shape; positive values mean the curve skewed to the right (right-tail), and negative values suggest skewing to the left (left-tail). The kurtosis coefficient determines the measure of peakedness and flatness. Normal distributions have kurtosis = 0. Positive values characterise peaked distributions, and negative kurtosis values indicate flatter distributions [106, 109]. Statistical practices recommend detecting and removing the outliers as they may induce detrimental influences such as skewed distributions [110].

As in Figure 5.1, the boxplot is helpful to visualise all the statistics associated with the outlier detection. The interquartile range (IQR), in Equation (5.3), is valuable when evaluating the outliers. Potential outliers are values measured below the lower fence or above the upper fence [111].

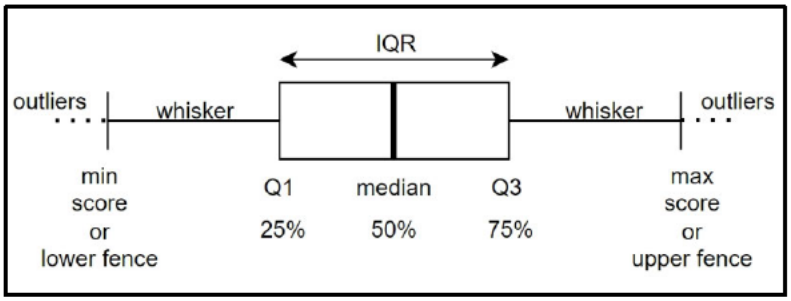

Figure 5.1: Elements of boxplot

$$IQR = Q3 - Q1 \tag{5.3}$$

5.3.1. DUT 433 MHz

Table 5.1 shows values for statistical metrics for the DUT 433 MHz case. From transmit rates Tx0 to Tx10, the small values of standard deviations and variances, approximately zero, expected dense distributions around the mean values. The values for skewness, approximately zero, indicated symmetrical distributions. The positive values for kurtosis, with a minimum of about 19.93 mW at Tx7 and a maximum of about 44.96 mW at Tx4, developed a tendency for peaked distributions. The mean and median are almost equal. The number of outliers combines those below the lower fence and those above the upper fence. The outlier rates indicate the ratio of outliers over the entire data. Variable Tx6 contains the highest score of outliers at 1.39 %. The entire observation shows a low rate of outliers. Besides some high values for kurtosis, data for the DUT 433 MHz case satisfy conditions for statistical normality. Figure 5.2 confirms these observations, showing a high concentration of samples around the mean, symmetrically shaped and slightly peaked distributions. The boxplot in Figure 5.3 shows outliers below and above the fences. Figure 5.4 adjusted the limits of the ordinate axis to visualise zoomed versions of the boxplots. The result reinforces the necessity to remove the outliers. The ring marks in the boxplots highlight the mean values.

Table 5.1: Statistical characteristics of IoT Radio Noise at DUT 433 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
DUT	433 MHz	Mean	1.054105e-02	1.054027e-02	1.054078e-02	1.053312e-02	1.053516e-02	1.053477e-02	1.053855e-02	1.054385e-02	1.054275e-02	1.054372e-02	1.054426e-02
		Std	3.534312e-04	3.473585e-04	3.350036e-04	3.096071e-04	3.283400e-04	3.280084e-04	3.444774e-04	3.437997e-04	3.388229e-04	3.392542e-04	3.418226e-04
		Var	1.249136e-07	1.206579e-07	1.122274e-07	9.585655e-08	1.078072e-07	1.075895e-07	1.186646e-07	1.181983e-07	1.148009e-07	1.150934e-07	1.168427e-07
		Kurt	2.331253e+01	2.341741 e+01	2.476091e+01	3.738505e+01	4.496607e+01	4.186730e+01	2.367878e+01	1.935876e+01	2.005544e+01	1.987552e+01	1.943532e+01
		Skew	1.470215e-01	1.258298e-01	1.199231e-01	1.567761e-01	3.411671e-01	3.316216e-01	1.366726e-01	6.994978e-02	6.292748e-02	6.189534e-02	7.017456e-02
		Min Score	9.652000e-03	9.667000e-03	9.706000e-03	9.777000e-03	9.727000e-03	9.723000e-03	9.695000e-03	9.679000e-03	9.697000e-03	9.693000e-03	9.688000e-03
		Q1	1.031600e-02	1.031900e-02	1.032800e-02	1.034100e-02	1.033000e-02	1.032900e-02	1.032700e-02	1.032500e-02	1.032800e-02	1.032800e-02	1.032800e-02
		Median	1.053500e-02	1.053300e-02	1.053200e-02	1.052500e-02	1.052600e-02	1.052600e-02	1.052800e-02	1.053600e-02	1.053300e-02	1.053500e-02	1.053500e-02
		Q3	1.075800e-02	1.075400e-02	1.074300e-02	1.071600e-02	1.073200e-02	1.073200e-02	1.074800e-02	1.075500e-02	1.074800e-02	1.075200e-02	1.075400e-02
		IQR	4.420000e-04	4.350000e-04	4.150000e-04	3.750000e-04	4.020000e-04	4.040000e-04	4.210000e-04	4.300000e-04	4.200000e-04	4.240000e-04	4.260000e-04
		Max Score	1.142200e-02	1.140600e-02	1.136500e-02	1.127900e-02	1.133500e-02	1.133800e-02	1.138000e-02	1.140000e-02	1.137900e-02	1.138700e-02	1.139300e-02
		Outlier samples	1.080440e+05	1.077720e+05	1.162550e+05	1.234200e+05	1.167750e+05	1.136250e+05	1.386280e+05	1.105610e+05	1.161550e+05	1.100850e+05	1.125570e+05
		Outlier rates (%)	1.080000e+00	1.080000e+00	1.160000e+00	1.230000e+00	1.170000e+00	1.140000e+00	1.390000e+00	1.110000e+00	1.160000e+00	1.100000e+00	1.130000e+00

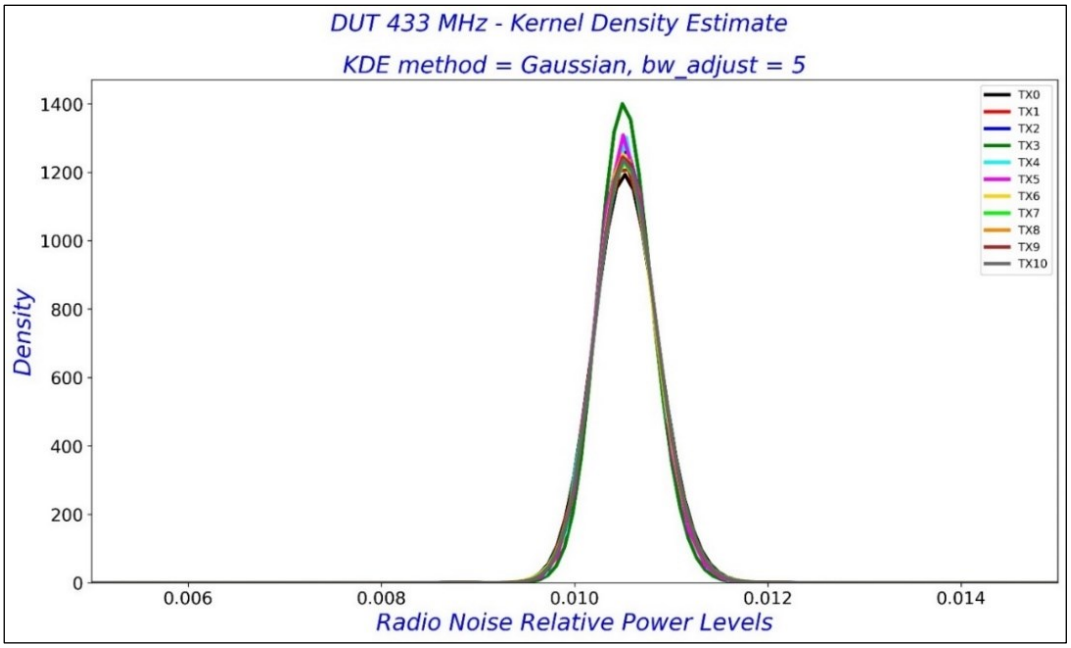

Figure 5.2: Density Distributions of IoT Radio Noise at DUT 433 MHz

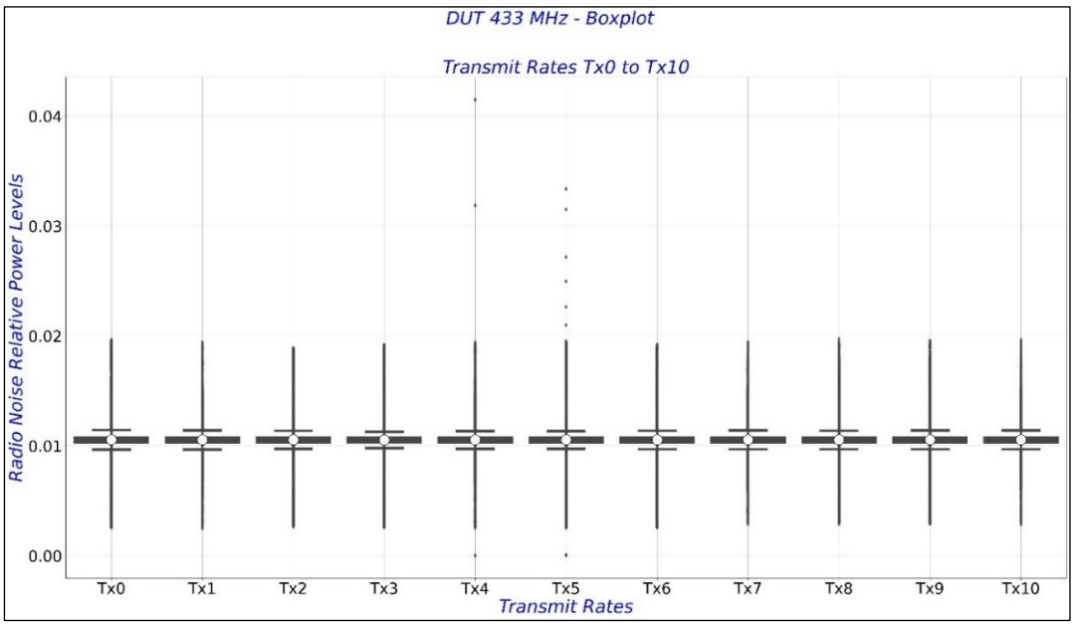

Figure 5.3: Full boxplots for IoT Radio Noise at DUT 433 MHz

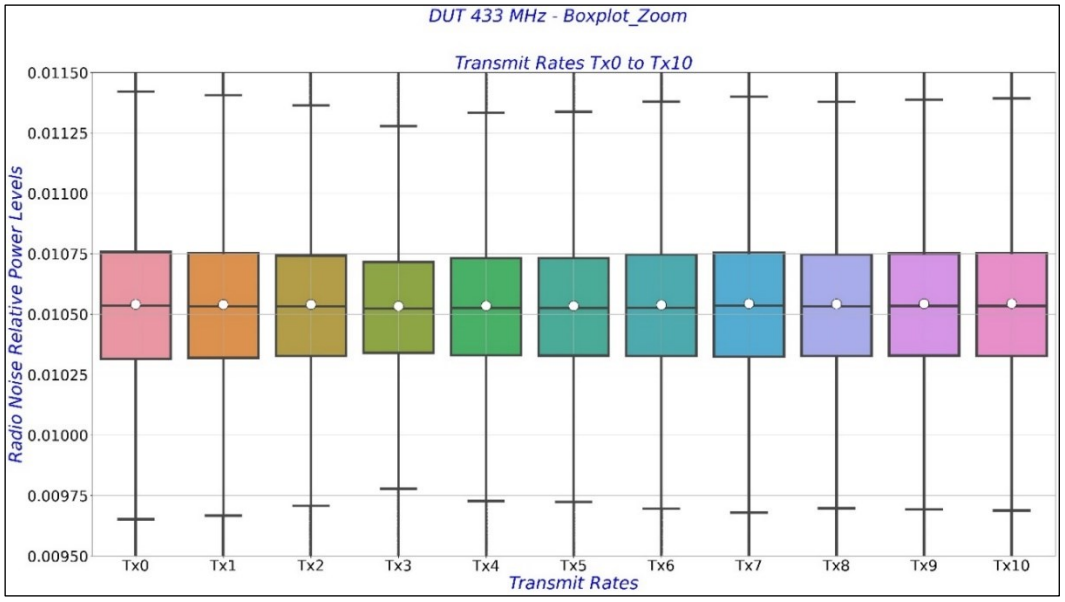

Figure 5.4: Boxplots for IoT Radio Noise at DUT 433 MHz – adjusted y-axis

5.3.2. DUT 868 MHz

Table 5.2 shows the statistical metrics for the DUT 868 MHz case. These data expect no dispersion due to insignificant standard deviation and variance values. The distributions show most of the data confined around the mean values. Data expect abnormality due to high values for kurtosis, for example, about 723 mW at Tx8. Skewness values slightly different from zero did not cause remarkably skewed distributions for most data. Tx8, with 12.24 mW of skewness, may not be enough to show a seriously skewed distribution to the right. Tx8 produced the highest outlier at 6%. Though the rest of the data showed reduced outlier rates, there was a necessity to remove them as every dataset counted a considerable number of outliers. The fact that median and mean values were approximately identical signified probable normality if further processing could reduce the outliers. The small values of IQR corroborate that the distributions would see most of the data squeezed within a small range. Figure 5.5 shows the density distributions and confirms the observations discussed on centrality aspects. Figure 5.6 expects to show boxplots. However, the presence of the outliers obstructed the boxplot visibility. To palliate, Figure 5.7 improves the resolution after adjusting the y-axis. The results point out the necessity to remove the outliers.

Table 5.2: Statistical characteristics of IoT Radio Noise at DUT - 868 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
DUT	868 MHz	Mean	1.003552e+00	1.003672e+00	1.003720e+00	1.003729e+00	1.003729e+00	1.003794e+00	1.003816e+00	1.003862e+00	1.003876e+00	1.003666e+00	1.003982e+00
		Std	2.934000e-03	2.227800e-02	2.619300e-02	2.692200e-02	2.692200e-02	3.151000e-02	3.246800e-02	3.508600e-02	3.628600e-02	2.088500e-02	4.072000e-02
		Var	9.000000e-06	4.960000e-04	6.860000e-04	7.250000e-04	7.250000e-04	9.930000e-04	1.054000e-03	1.231000e-03	1.317000e-03	4.360000e-04	1.658000e-03
		Kurt	7.934167e+01	5.593848e+02	4.470676e+02	4.118026e+02	4.118026e+02	3.174897e+02	3.093061e+02	2.482331e+02	7.239953e+02	6.498772e+01	2.364371e+01
		Skew	-3.346260e-01	8.526816e+00	7.928197e+00	7.491961e+00	7.491961e+00	6.697823e+00	6.645309e+00	5.893685e+00	1.224578e+01	1.035623e+00	7.573410e-01
		Min Score	9.972020e-01	9.969070e-01	9.966050e-01	9.966700e-01	9.966700e-01	9.966640e-01	9.965440e-01	9.964020e-01	9.964080e-01	9.953600e-01	9.885320e-01
		Q1	1.001887e+00	1.001822e+00	1.001755e+00	1.001775e+00	1.001775e+00	1.001807e+00	1.001755e+00	1.001755e+00	1.001755e+00	1.001515e+00	9.999700e-01
		Median	1.003269e+00	1.003331e+00	1.003331e+00	1.003301e+00	1.003301e+00	1.003275e+00	1.003271e+00	1.003282e+00	1.003289e+00	1.003319e+00	1.003350e+00
		Q3	1.005011e+00	1.005099e+00	1.005189e+00	1.005178e+00	1.005178e+00	1.005235e+00	1.005230e+00	1.005324e+00	1.005320e+00	1.005619e+00	1.007595e+00
		IQR	3.124000e-03	3.277000e-03	3.434000e-03	3.403000e-03	3.403000e-03	3.428000e-03	3.474000e-03	3.569000e-03	3.565000e-03	4.103000e-03	7.625000e-03
		Max Score	1.009697e+00	1.010014e+00	1.010340e+00	1.010282e+00	1.010282e+00	1.010377e+00	1.010441e+00	1.010678e+00	1.010668e+00	1.011774e+00	1.019032e+00
		Outlier samples	2.515510e+05	3.697520e+05	3.699450e+05	3.867290e+05	3.867290e+05	4.077770e+05	4.685520e+05	5.194580e+05	6.036170e+05	1.018303e+06	2.465600e+06
		Outlier rates (%)	2.520000e+00	3.700000e+00	3.700000e+00	3.870000e+00	3.870000e+00	4.080000e+00	4.690000e+00	5.190000e+00	6.040000e+00	1.018000e+01	2.466000e+01

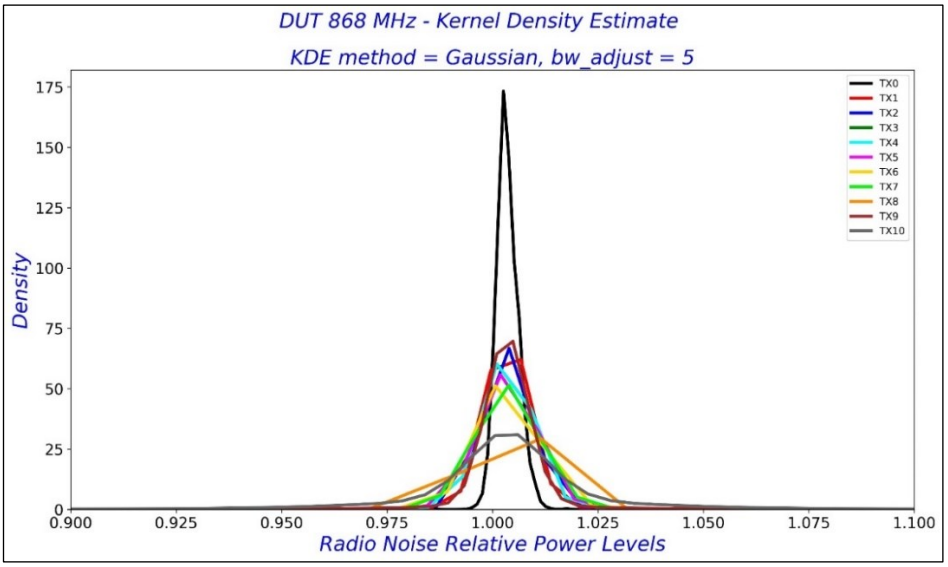

Figure 5.5: Density Distributions of IoT Radio Noise at DUT 868 MHz

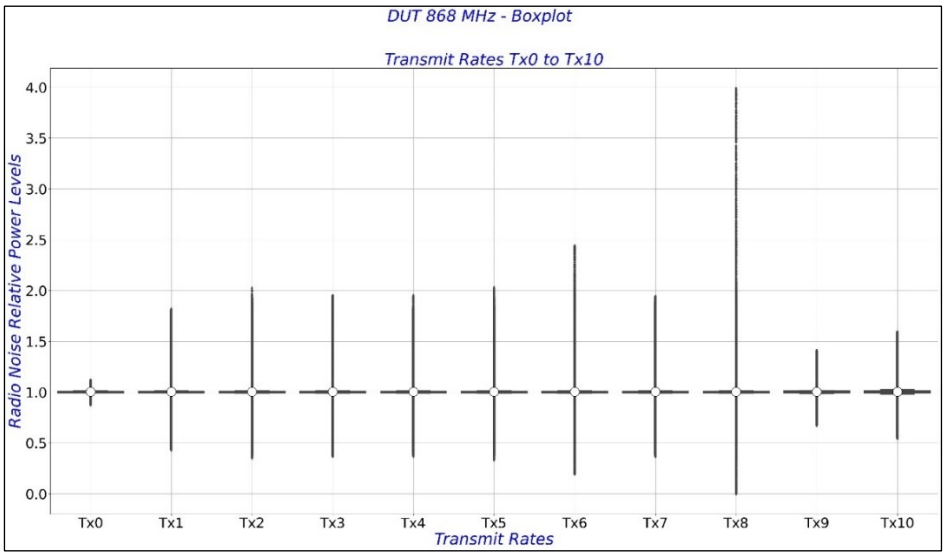

Figure 5.6: Full boxplots for IoT Radio Noise at DUT 868 MHz

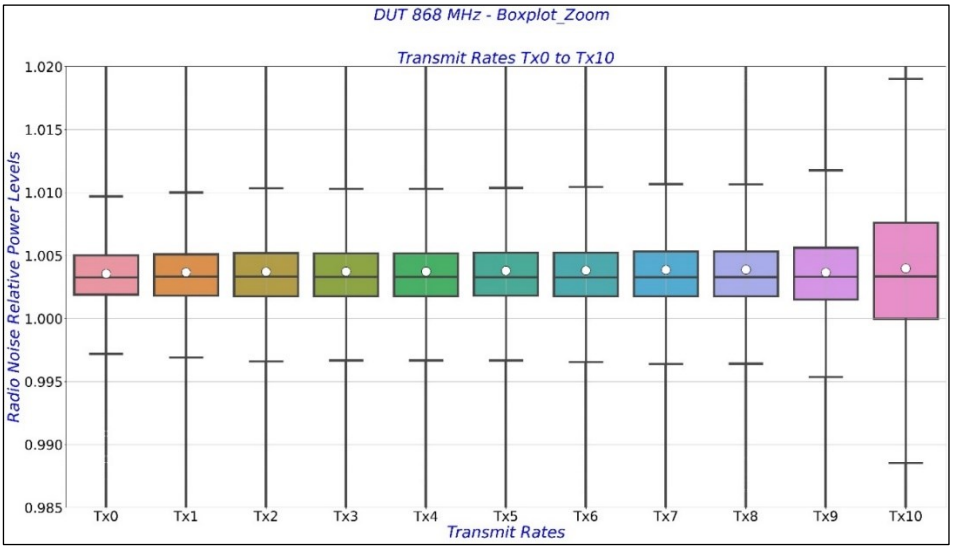

Figure 5.7: Boxplots for IoT Radio Noise at DUT 868 MHz – adjusted y-axis

5.3.3. DUT 2400 MHz

Table 5.3 shows statistics for the DUT 2400 MHz case. There was a possibility of normality as the mean and median values were very close. Slight standard deviation and variance values suggested that the observation should see most data confined around the mean. However, significant kurtosis values, up to 559.2 mW at Tx8, insinuate peaked distributions. With the highest skewness values of 9.9 mW at Tx3, these data should not expect remarkable skewed distributions. Even though the outlier rates were low, for example, 2.98% as a maximum at Tx2, removing them will only improve the distributions. Figure 5.8 corroborates these observations. Figure 5.9 shows the presence of outliers, making it challenging to interpret the graph. In Figure 5.10, the result of eliminating the outliers made informative boxplots.

Table 5.3: Statistical characteristics of IoT Radio Noise at DUT - 2400 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
DUT	2400 MHz	Mean	3.899210e-02	3.899797e-02	3.899700e-02	3.899561e-02	3.900300e-02	3.900231e-02	3.900400e-02	3.901432e-02	3.901700e-02	3.901300e-02	3.898289e-02
		Std	9.426682e-04	9.848377e-04	1.112000e-03	9.495766e-04	1.160000e-03	9.912186e-04	1.173000e-03	9.875899e-04	1.142000e-03	1.024000e-03	9.299148e-04
		Var	8.886232e-07	9.699053e-07	1.000000e-06	9.016956e-07	1.000000e-06	9.825142e-07	1.000000e-06	9.753337e-07	1.000000e-06	1.000000e-06	8.647415e-07
		Kurt	6.263878e+01	3.889665e+01	2.195567e+02	1.544737e+01	3.123230e+02	3.933557e+01	4.201304e+02	2 .579819e+02	5.592899e+02	1.461365e+02	1.551277e+02
		Skew	1.663775e+00	1.302008e+00	4.048037e+00	9.902095e-01	5.492811e+00	1.356131e+00	6.867987e+00	4 .059387e+00	7.753762e+00	2.856238e+00	2.813828e+00
		Min Score	3.692500e-02	3.690900e-02	3.691900e-02	3.691900e-02	3.689400e-02	3.689400e-02	3.689400e-02	3.692700e-02	3.694200e-02	3.692700e-02	3.667100e-02
		Q1	3.848700e-02	3.845200e-02	3.846300e-02	3.846300e-02	3.845200e-02	3.845200e-02	3.845200e-02	3.847100e-02	3.848700e-02	3.847100e-02	3.835400e-02
		Median	3.901300e-02	3.896500e-02	3.897500e-02	3.897500e-02	3.896500e-02	3.896500e-02	3.896500e-02	3.898400e-02	3.900000e-02	3.898400e-02	3.891300e-02
		Q3	3.952900e-02	3.948200e-02	3.949200e-02	3.949200e-02	3.949200e-02	3.949200e-02	3.949200e-02	3.950100e-02	3.951600e-02	3.950100e-02	3.947600e-02
		IQR	1.042000e-03	1.029000e-03	1.029000e-03	1.029000e-03	1.039000e-03	1.039000e-03	1.039000e-03	1.029000e-03	1.030000e-03	1.029000e-03	1.122000e-03
		Max Score	4.109200e-02	4.102500e-02	4.103600e-02	4.103600e-02	4.105100e-02	4.105100e-02	4.105100e-02	4.104500e-02	4.106100e-02	4.104500e-02	4.115900e-02
		Outlier samples	2.206710e+05	2.894010e+05	2983 14.00	2.767720e+05	2.873740e+05	2.761650e+05	2.896760e+05	2.556450e+05	2.640230e+05	2.569760e+05	1.290110e+05
		Outlier rates (%)	2.210000e+00	2.890000e+00	2.980000e+00	2.770000e+00	2.870000e+00	2.760000e+00	2.900000e+00	2.560000e+00	2.640000e+00	2.570000e+00	1.290000e+00

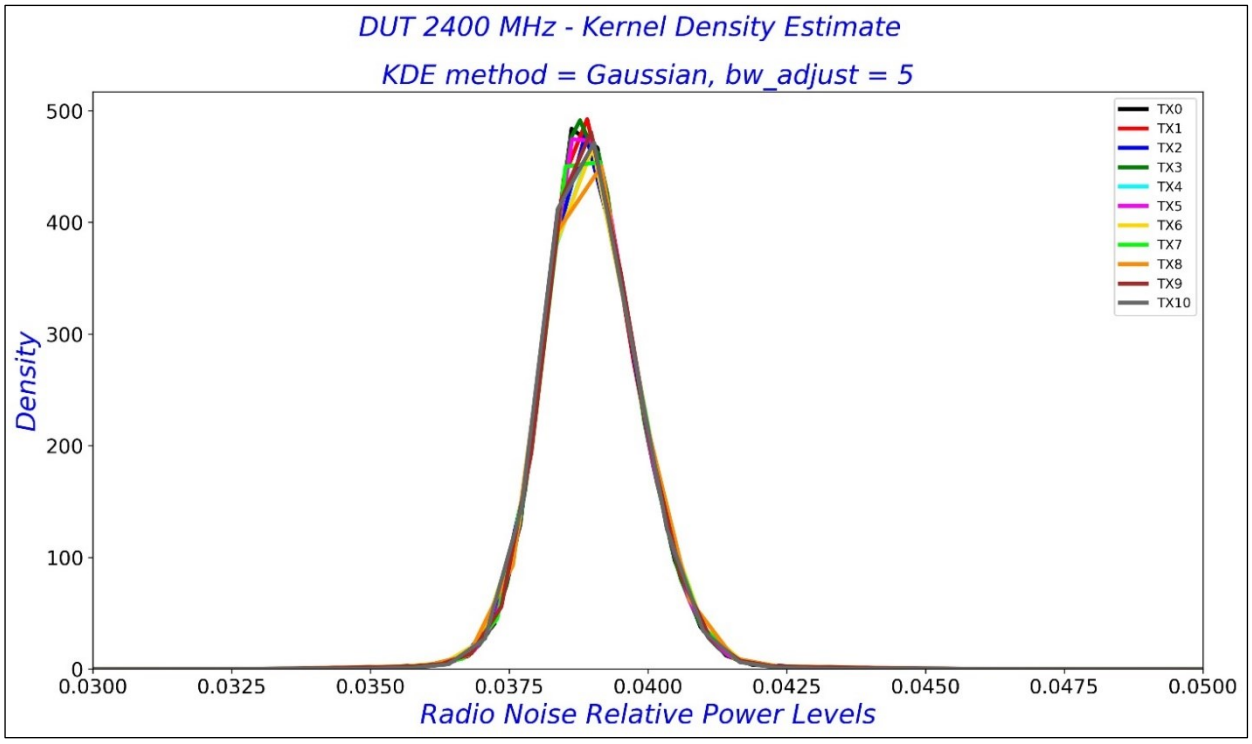

Figure 5.8: Density Distributions of IoT Radio Noise at DUT - 2400 MHz

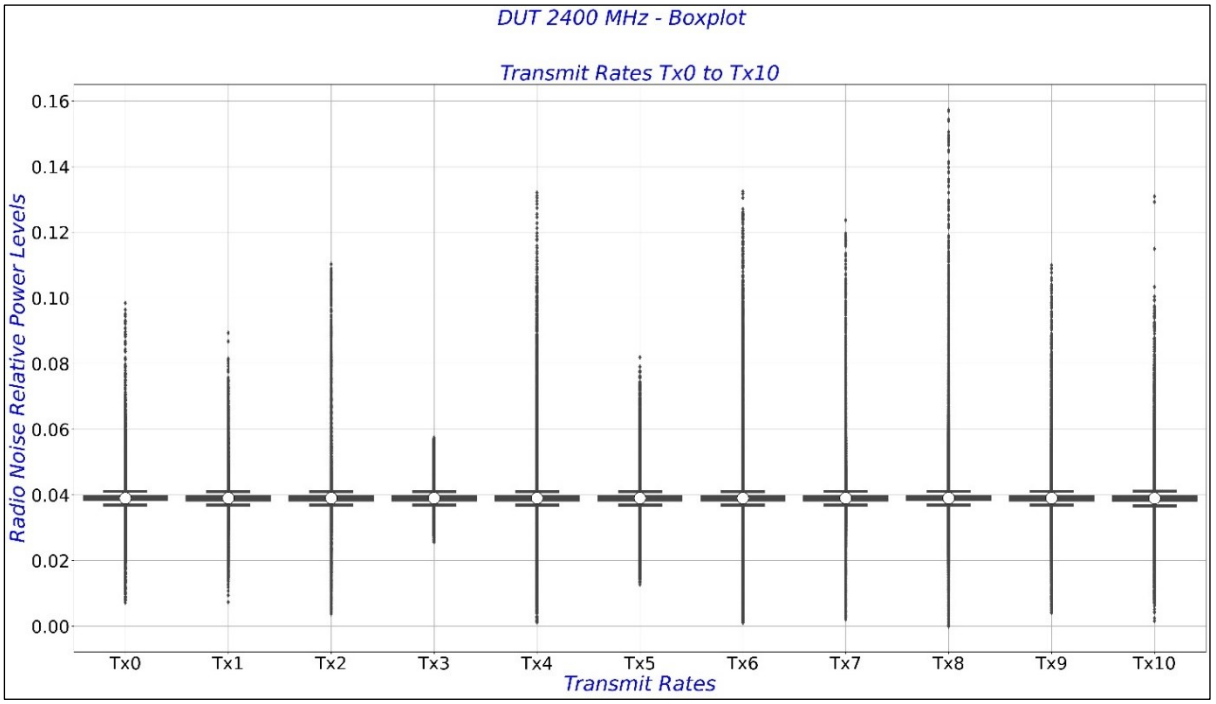

Figure 5.9: Full boxplots for IoT Radio Noise at DUT - 2400 MHz

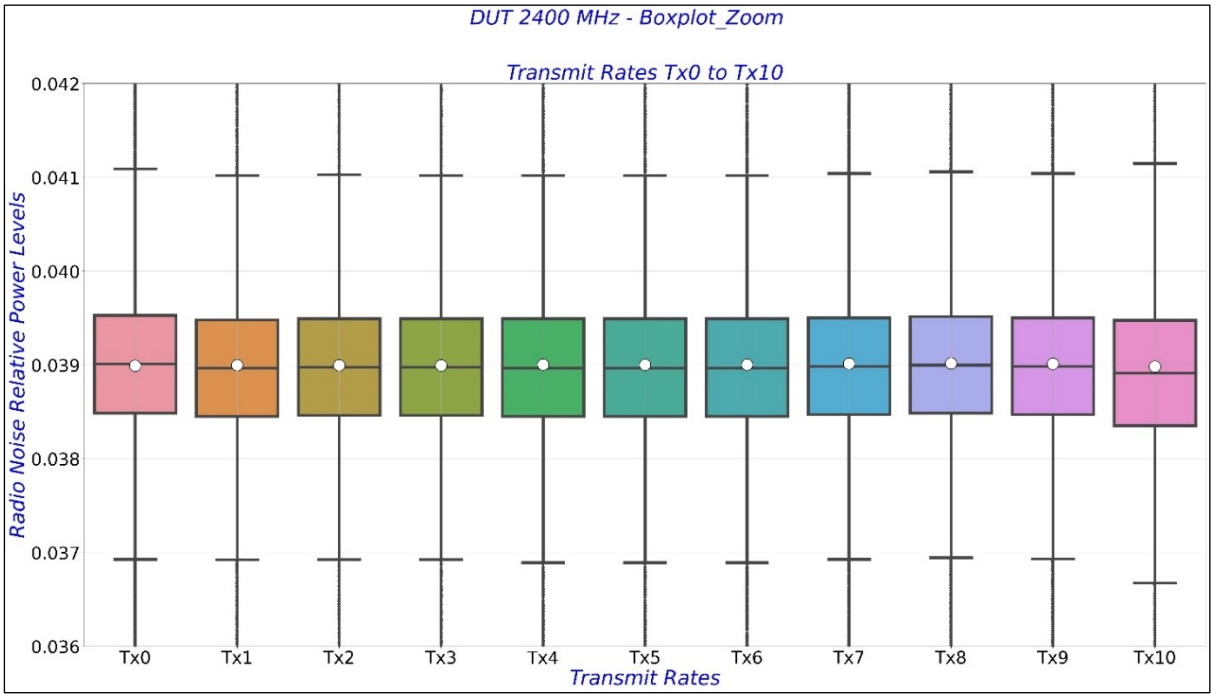

Figure 5.10: Boxplots for IoT Radio Noise at DUT - 2400 MHz – no outliers

5.3.4. GLM 433 MHz

Table 5.4 shows that the GLM 433 MHz case experienced minor standard deviations and variance values. Data expect centralised distributions about the mean values. From Tx0 to Tx11, the values for skewness, approximately zero, should not cause skewed distributions. The kurtosis values between 24.94 mW (minimum) at Tx1 and 29.35 mW (maximum) at Tx9 are not too high to cause

peaked distributions. The mean and median are comparable and approximately zero mW. These responses may satisfy the conditions for a normal distribution. However, the rates of outliers are low, with the highest rate of 1.30 % at Tx1. Analytic practices recommend reducing the number of outlier samples. Figure 5.11 confirms these observations as the distributions form an approximately symmetrical bell shape. The boxplot in Figure 5.12 already visualised all the elements of the boxplots to confirm the low rates of outliers. The modified minimum and maximum limits of the ordinate axis in Figure 5.13 demonstrated the improvement one can expect, provided the the distributions obtained a reduced count outlier.

Table 5.4: Statistical characteristics of IoT Radio Noise at GLM - 433 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
GLM	433 MHz	Mean	2.847475e-03	2.847503e-03	2.847257e-03	2.847404e-03	2.847095e-03	2.847066e-03	2.846959e-03	2.847322e-03	2.847462e-03	2.847011e-03	2.847463e-03
		Std	1.690073e-04	1.758847e-04	1.631939e-04	1.732418e-04	1.719088e-04	1.714887e-04	1.683389e-04	1.714087e-04	1.707211e-04	1.709894e-04	1.712372e-04
		Var	2.856347e-08	3.093542e-08	2.663226e-08	3.001273e-08	2.955263e-08	2.940837e-08	2.833797e-08	2.938094e-08	2.914570e-08	2.923737e-08	2.932220e-08
		Kurt	2.494326e+01	2.668491e+01	2.764921e+01	2.801175e+01	2.919053e+01	2.791289e+01	2.895373e+01	2.903117e+01	2.859289e+01	2.935260e+01	2.903279e+01
		Skew	7.066023e-01	7.982482e-01	7.441999e-01	8.175024e-01	8.386396e-01	8.107824e-01	8.184314e-01	8.333889e-01	8.220846e-01	8.339812e-01	8.333563e-01
		Min Score	2.424000e-03	2.411000e-03	2.445000e-03	2.419000e-03	2.419000e-03	2.424000e-03	2.430000e-03	2.421000e-03	2.418000e-03	2.422000e-03	2.420000e-03
		Q1	2.739000e-03	2.736000e-03	2.745000e-03	2.738000e-03	2.738000e-03	2.739000e-03	2.741000e-03	2.739000e-03	2.738000e-03	2.739000e-03	2.739000e-03
		Median	2.843000e-03	2.842000e-03	2.842000e-03	2.842000e-03	2.842000e-03	2.842000e-03	2.842000e-03	2.842000e-03	2.841000e-03	2.841000e-03	2.842000e-03
		Q3	2.950000e-03	2.952000e-03	2.944000e-03	2.951000e-03	2.951000e-03	2.950000e-03	2.948000e-03	2.950000e-03	2.951000e-03	2.951000e-03	2.951000e-03
		IQR	2.100000e-04	2.170000e-04	2.000000e-04	2.130000e-04	2.130000e-04	2.100000e-04	2.070000e-04	2.120000e-04	2.130000e-04	2.120000e-04	2.120000e-04
		Max Score	3.265000e-03	3.277000e-03	3.244000e-03	3.271000e-03	3.270000e-03	3.266000e-03	3.258000e-03	3.268000e-03	3.271000e-03	3.269000e-03	3.269000e-03
		Outlier samples	1.187090e+05	1.295580e+05	1.273540e+05	1.274110e+05	1.205790e+05	1.255370e+05	121635 .00	1.213080e+05	1.154600e+05	1.182720e+05	1.194410e+05
		Outlier rates (%)	1.190000e+00	1.300000e+00	1.270000e+00	1.270000e+00	1.210000e+00	1.260000e+00	1.220000e+00	1.210000e+00	1.150000e+00	1.180000e+00	1.190000e+00

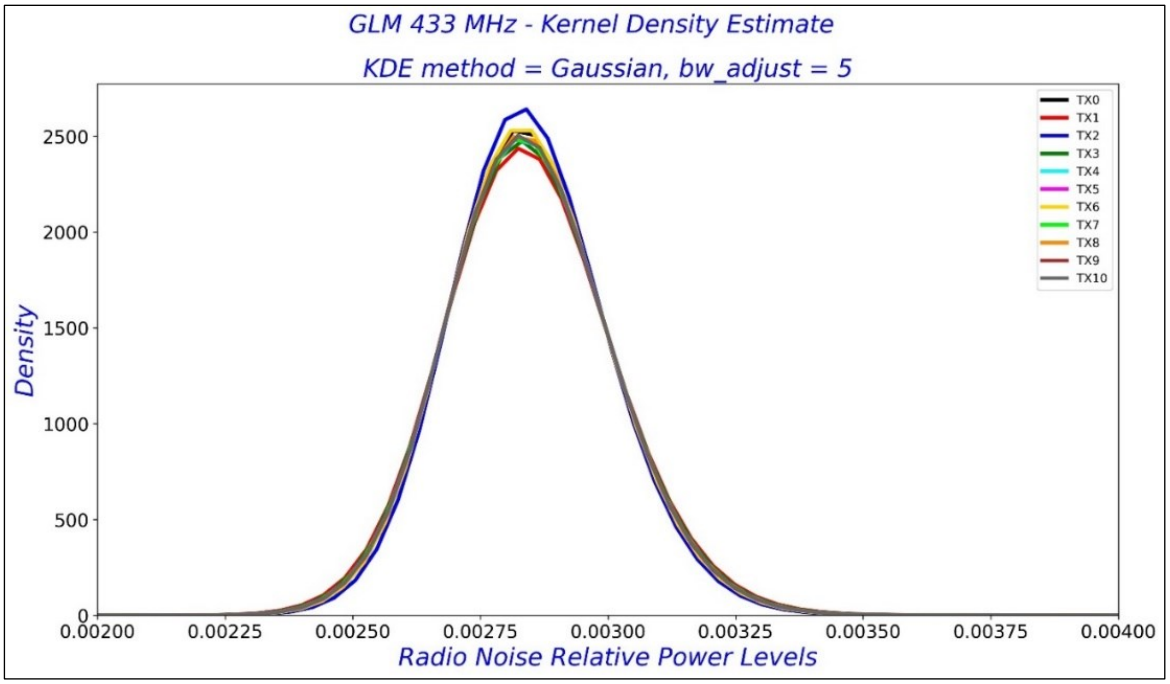

Figure 5.11: Density Distributions of IoT Radio Noise at GLM - 433 MHz

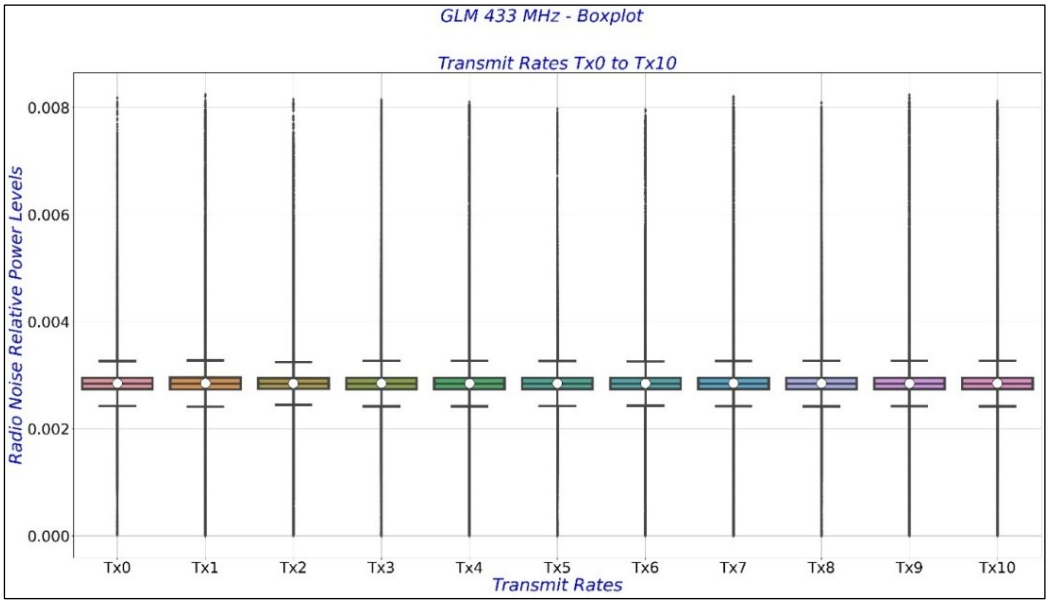

Figure 5.12: Full boxplots for IoT Radio Noise at GLM - 433 MHz

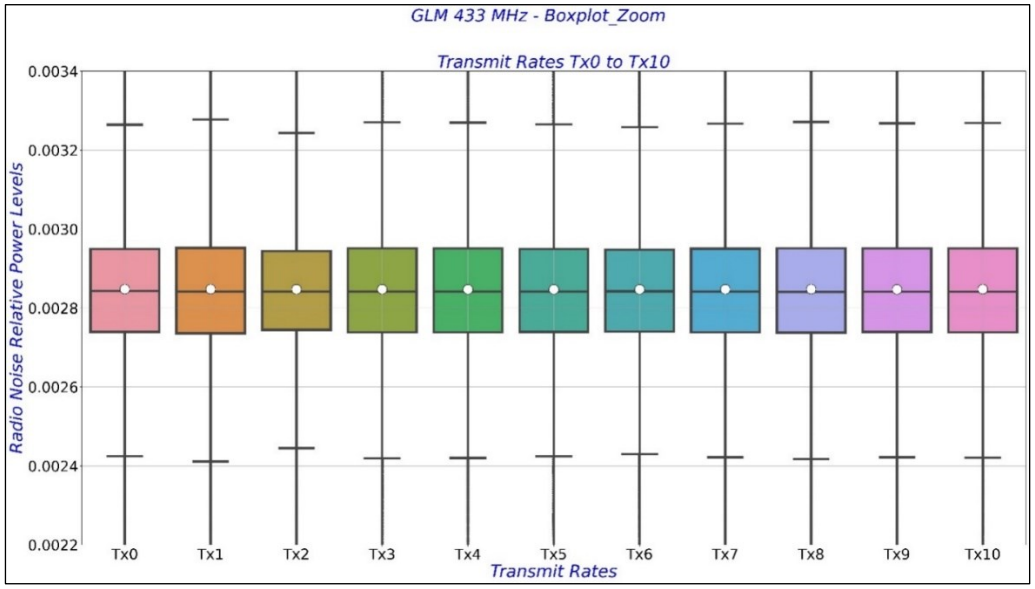

Figure 5.13: Boxplots for IoT Radio Noise at GLM - 433 MHz – no outliers

5.3.5. GLM 868 MHz

Table 5.5 shows the statistical metrics for the GLM 868 MHz case. The distributions expect no dispersion due to insignificant standard deviation and variance values. The distributions see most of the samples confined around the mean values. The distributions expect a symmetrical shape due to low values for skewness. However, Tx1 to Tx6 have high kurtosis values, for example, 327.57 mW at Tx4, capable of influencing the peak or smoothness of the distributions. Tx10 recorded 22.29 mW as the highest rate of outliers in this series. The fact that the median and mean showed approximate and comparable values indicated the possibility for normal distributions, provided further processing removed the outliers. The small IQR values confirmed that the distributions had

most data centred within a small range. Figure 5.14 corroborates the observations with peaked density distributions. Figure 5.15 shows the influence of the outliers obstructing the boxplot visibility. Figure 5.16 achieves an improved resolution of visualisations after tuning the y-axis and reinforces the necessity to remove the outliers.

Table 5.5: Statistical characteristics of IoT Radio Noise at GLM - 868 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
GLM	868 MHz	Mean	8.290513e-02	8.292300e-02	8.293400e-02	8.293400e-02	8.293500e-02	8.293200e-02	8.293800e-02	8.294000e-02	8.299000e-02	8.305400e-02	8.325400e-02
		Std	8.495090e-04	2.579000e-03	3.145000e-03	3.172000e-03	3.144000e-03	2.887000e-03	3.216000e-03	3.282000e-03	5.274000e-03	6.973000e-03	1.065000e-02
		Var	7.216655e-07	7.000000e-06	1.000000e-05	1.000000e-05	1.000000e-05	8.000000e-06	1.000000e-05	1.100000e-05	2.800000e-05	4.900000e-05	1.130000e-04
		Kurt	8.731086e+01	2.698557e+02	2.737227e+02	2.769290e+02	3.275723e+02	2.399274e+02	1.362141e+02	8.843579e+01	9.344413e+01	5.882927e+01	2.074641e+01
		Skew	2.216498e-01	5.680360e+00	6.726043e+00	6.654069e+00	7.391822e+00	5.469801e+00	3.671437e+00	2.518281e+00	3.743000e+00	2.958566e+00	1.761425e+00
		Min Score	8.106400e-02	8.102100e-02	8.092400e-02	8.091400e-02	8.091400e-02	8.087100e-02	8.088000e-02	8.083500e-02	8.068500e-02	8.039400e-02	7.835300e-02
		Q1	8.243300e-02	8.242300e-02	8.239800e-02	8.239800e-02	8.239800e-02	8.239100e-02	8.239800e-02	8.239800e-02	8.236600e-02	8.228800e-02	8.182300e-02
		Median	8.283900e-02	8.284500e-02	8.283500e-02	8.282400e-02	8.283100e-02	8.281300e-02	8.284600e-02	8.283700e-02	8.281900e-02	8.285000e-02	8.286300e-02
		Q3	8.334600e-02	8.335800e-02	8.338000e-02	8.338700e-02	8.338700e-02	8.340400e-02	8.341000e-02	8.343900e-02	8.348700e-02	8.355100e-02	8.413600e-02
		IQR	9.130000e-04	9.350000e-04	9.820000e-04	9.890000e-04	9.890000e-04	1.013000e-03	1.012000e-03	1.042000e-03	1.121000e-03	1.263000e-03	2.313000e-03
		Max Score	8.471600e-02	8.476100e-02	8.485400e-02	8.487100e-02	8.487100e-02	8.492400e-02	8.492700e-02	8.500200e-02	8.516900e-02	8.544600e-02	8.760500e-02
		Outlier samples	224632 .00	396464 .00	396381 .00	420758 .00	4.517580e+05	460801 .00	5.382920e+05	684929 .00	8.650650e+05	1.402268e+06	2229328 .00
		Outlier rates (%)	2.250000e+00	3.960000e+00	3.960000e+00	4.210000e+00	4 .52	4 .61	5.380000e+00	6.850000e+00	8.650000e+00	1.402000e+01	2.229000e+01

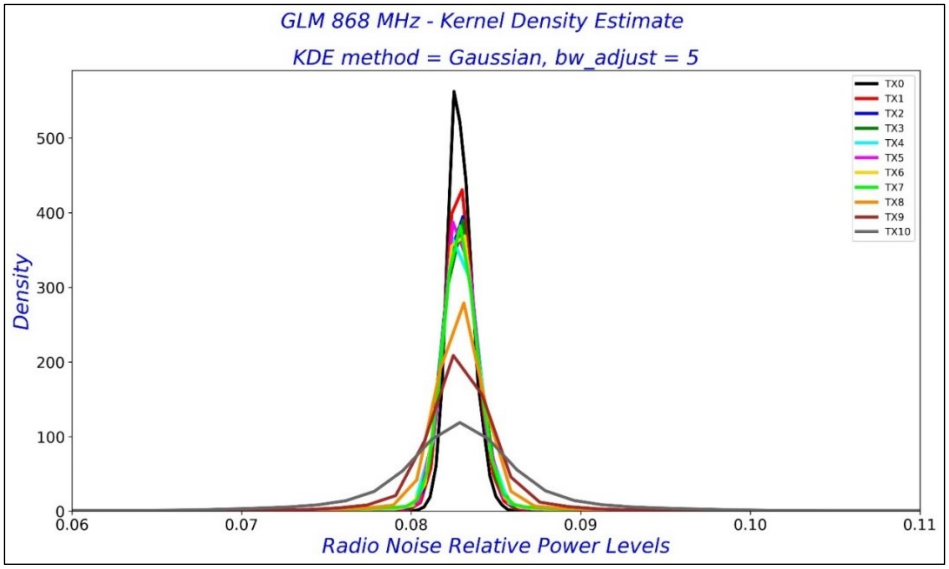

Figure 5.14: Density Distributions of IoT Radio Noise at GLM - 868 MHz

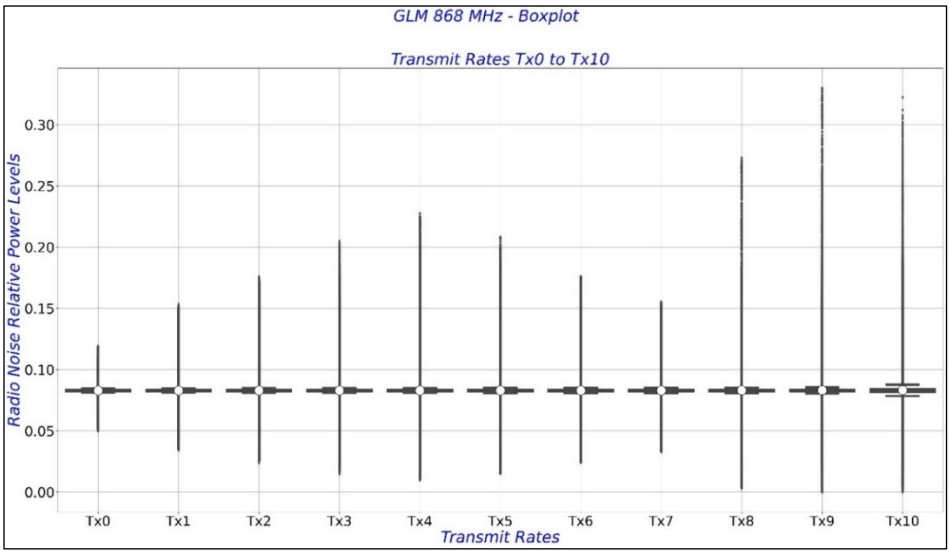

Figure 5.15: Full boxplots for IoT Radio Noise at GLM - 868 MHz

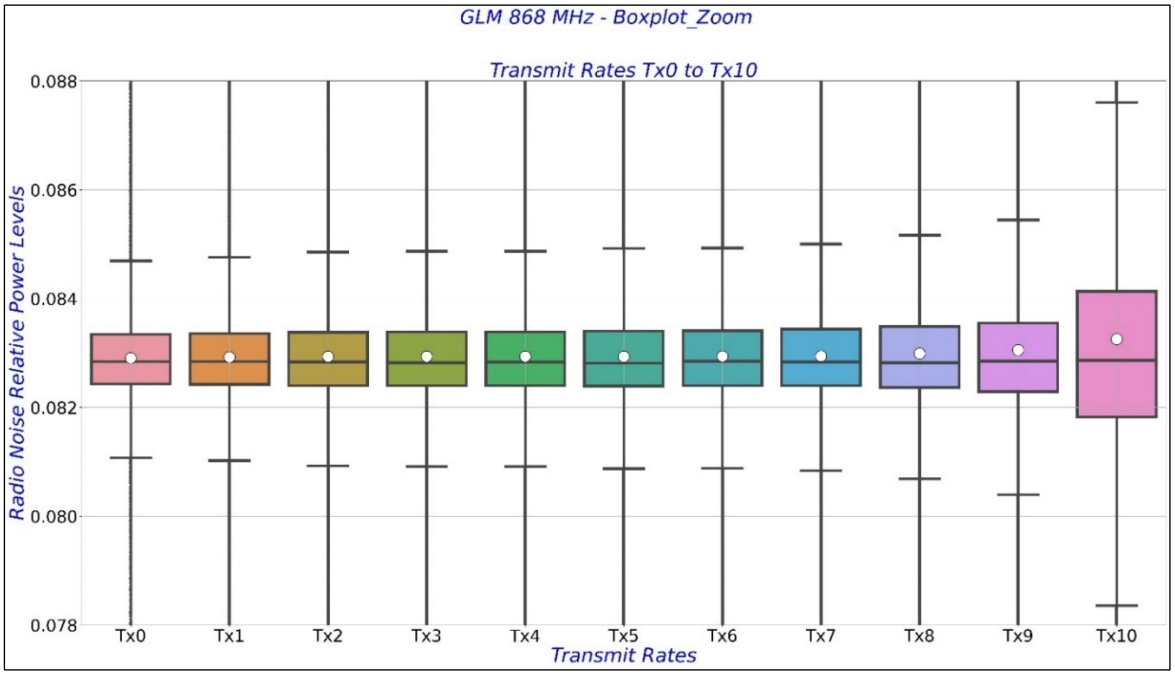

Figure 5.16: Boxplots for IoT Radio Noise at GLM - 868 MHz – no outliers

5.3.6. GLM 2400 MHz

Table 5.6 shows statistics for the GLM 2400 MHz case. Normal distribution was possible as the mean and median values were close, and both were approximately zero mW. Weak standard deviations and variances suggested that the observation should see most data confined around the mean. These data series carry significant kurtosis values above 68.76 mW at Tx0 up to 502.74 mW at Tx9 and may influence distributions' peakedness and smoothness. The weak skewness factors, for example, 10.75 mW as the highest at Tx9, should not lead to skewed distributions. Even though the outlier rates were low, for example, 2.64% as a maximum at Tx7, removing them would only improve the distributions. Figure 5.17 corroborates these observations as the distributions show unsmooth responses. In Figure 5.18, the high number of outliers prevents obtaining meaningful visualisation for boxplots. In Figure 5.19, as a result of simulating extraction of the outliers, the visualisation resolution allows clear visibility for all elements of boxplots. All the datasets show to have a comparable IQR range.

Table 5.6: Statistical characteristics of IoT Radio Noise at GLM - 2400 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
GLM	2400 MHz	Mean	5.372400e-02	5.373900e-02	5.372600e-02	5.372000e-02	5.374400e-02	5.374200e-02	5.374500e-02	5.376100e-02	5.374500e-02	5.376300e-02	5.375300e-02
		Std	1.359000e-03	1.847000e-03	1.273000e-03	1.300000e-03	1.657000e-03	1.594000e-03	1.457000e-03	1.913000e-03	1.291000e-03	2.222000e-03	1.531000e-03
		Var	2.000000e-06	3.000000e-06	2.000000e-06	2.000000e-06	3.000000e-06	3.000000e-06	2.000000e-06	4.000000e-06	2.000000e-06	5.000000e-06	2.000000e-06
		Kurt	6.876356e+01	3.083926e+02	9.516235e+01	7.505375e+01	3.740635e+02	1.457575e+02	2.084527e+02	4.045508e+02	1.931022e+02	5.027450e+02	1.223456e+02
		Skew	1.838075e+00	6.335283e+00	2.076680e+00	1.889881e+00	6.610726e+00	3.366846e+00	4.173798e+00	7.736165e+00	3.297862e+00	1.075965e+01	2.901883e+00
		Min Score	5.097200e-02	5.082300e-02	5.108200e-02	5.097200e-02	5.108200e-02	5.108200e-02	5.108200e-02	5.121900e-02	5.108200e-02	5.108200e-02	5.108200e-02
		Q1	5.304900e-02	5.304900e-02	5.306000e-02	5.304900e-02	5.306000e-02	5.306000e-02	5.306000e-02	5.311400e-02	5.306000e-02	5.306000e-02	5.306000e-02
		Median	5.371700e-02	5.370600e-02	5.371700e-02	5.371700e-02	5.371700e-02	5.371700e-02	5.371700e-02	5.371700e-02	5.371700e-02	5.371700e-02	5.371700e-02
		Q3	5.443300e-02	5.453300e-02	5.437800e-02	5.443300e-02	5.437800e-02	5.437800e-02	5.437800e-02	5.437800e-02	5.437800e-02	5.437800e-02	5.437800e-02
		IQR	1.384000e-03	1.484000e-03	1.318000e-03	1.384000e-03	1.318000e-03	1.318000e-03	1.318000e-03	1.264000e-03	1.318000e-03	1.318000e-03	1.318000e-03
		Max Score	5.651000e-02	5.675900e-02	5.635500e-02	5.651000e-02	5.635500e-02	5.635500e-02	5.635500e-02	5.627300e-02	5.635500e-02	5.635500e-02	5.635500e-02
		Outlier samples	1.995970e+05	1.756020e+05	1.842920e+05	174637 .00	228449 .00	224995 .00	2.159060e+05	263626 .00	1.949010e+05	256766 .00	2.531480e+05
		Outliers rates (%)	2.000000e+00	1.760000e+00	1.840000e+00	1.750000e+00	2.280000e+00	2.250000e+00	2.160000e+00	2.640000e+00	1.950000e+00	2.570000e+00	2.530000e+00

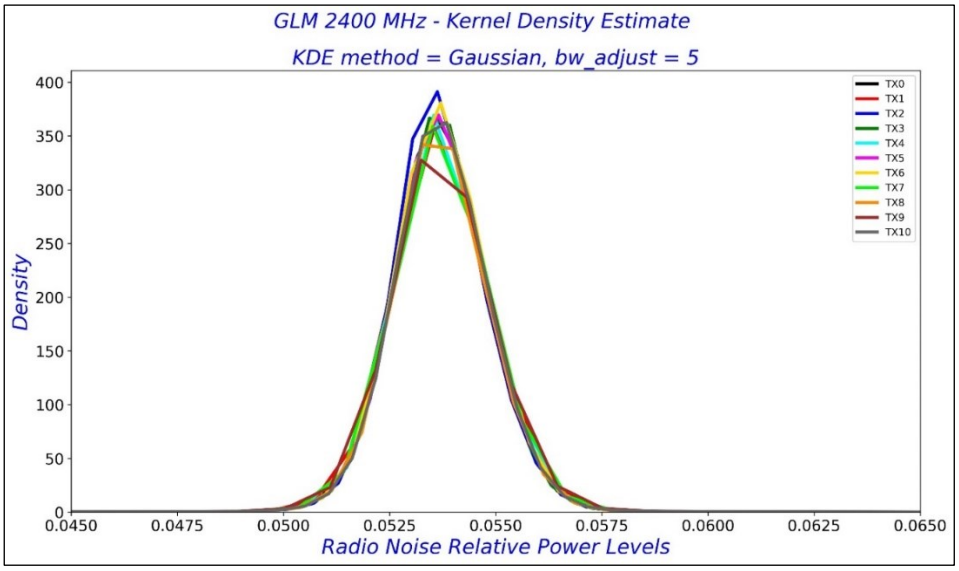

Figure 5.17: Density Distributions of IoT Radio Noise at GLM - 2400 MHz

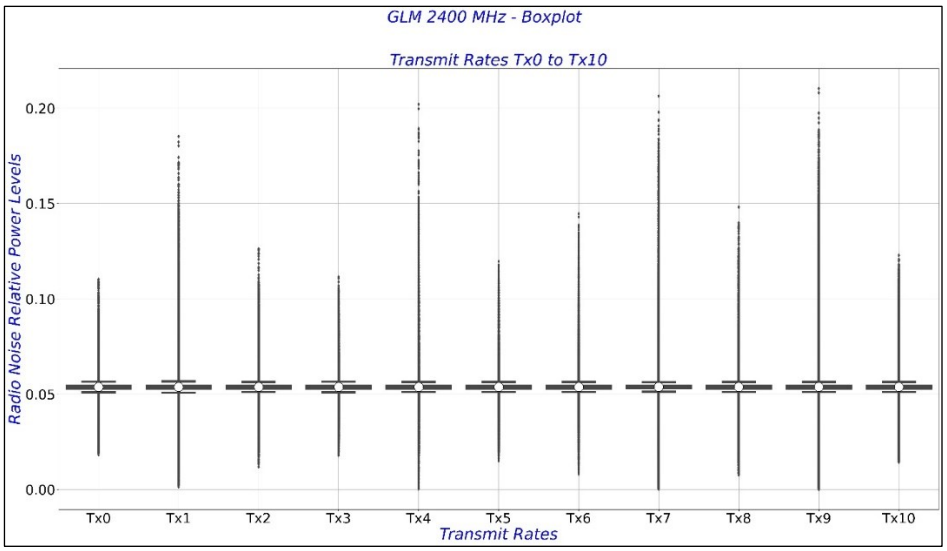

Figure 5.18: Full boxplots for IoT Radio Noise at GLM - 2400 MHz

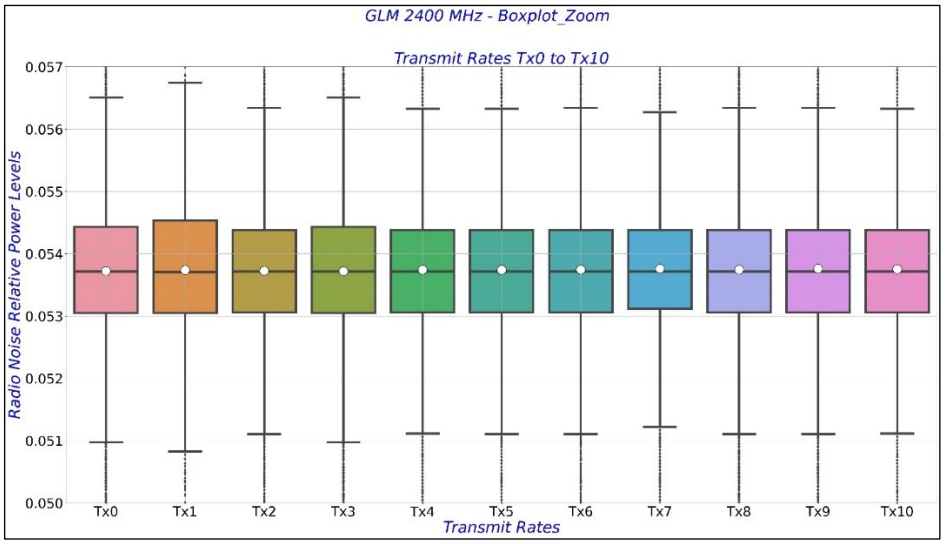

Figure 5.19: Boxplots for IoT Radio Noise at GLM - 2400 MHz – no outliers

5.3.7. MNC 433 MHz

Table 5.7 shows that MNC 433 MHz case expected density distributions with most of the samples confined about the means due to approximately zero mW levels of standard deviations and variances. Data should not expect skewed distributions with small values for skewness. The kurtosis of 36.86 mW (maximum) at Tx1 is not high enough to influence the shape of distributions. The mean and median are comparable and approximately zero mW. These data expect responses similar to those of normal distributions. Tx1 shows the highest rate of outliers at 1.26 %. Figure 5.20 agrees with these observations as it shows the symmetrical bell shape of the distributions. Even though the outlier rates were low, Figure 5.21 could not visualise all the boxplot elements. The result was palliated in Figure 5.22 by tuning the y-axis scale. The results encourage removing the outliers.

Table 5.7: Statistical characteristics of IoT Radio Noise at MNC - 433 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
MNC	433 MHz	Mean	2.807315e-03	2.809429e-03	2.809776e-03	2.808835e-03	2.808604e-03	2.809356e-03	2.808765e-03	2.809694e-03	2.809670e-03	2.809392e-03	2.809136e-03
		Std	1.635268e-04	1.672270e-04	1.665232e-04	1.657745e-04	1.656697e-04	1.685060e-04	1.676935e-04	1.692995e-04	1.688092e-04	1.657857e-04	1.683287e-04
		Var	2.674102e-08	2.796486e-08	2.772997e-08	2.748117e-08	2.744644e-08	2.839429e-08	2.812111e-08	2.866232e-08	2.849653e-08	2.748491 e-08	2.833454e-08
		Kurt	3.683873e+01	3.308213e+01	3.316484e+01	3.297052e+01	3.329025e+01	3.517111 e+01	3.324725e+01	3.220476e+01	3.316810e+01	3.253427e+01	3.314183e+01
		Skew	9.953848e-01	9.302457e-01	9.250232e-01	9.135042e-01	9.209347e-01	9.745232e-01	9.32151oe-01	9.172996e-01	9.399292e-01	9.025599e-01	9.348742e-01
		Min Score	2.405000e-03	2.397000e-03	2.398000e-03	2.394000e-03	2.396000e-03	2.391000e-03	2.392000e-03	2.389000e-03	2.391000e-03	2.399000e-03	2.394000e-03
		Q1	2.704000e-03	2.705000e-03	2.704000e-03	2.703000e-03	2.702000e-03	2.702000e-03	2.702000e-03	2.702000e-03	2.702000e-03	2.704000e-03	2.703000e-03
		Median	2.802000e-03	2.803000e-03	2.804000e-03	2.804000e-03	2.804000e-03	2.802000e-03	2.803000e-03	2.803000e-03	2.802000e-03	2.803000e-03	2.802000e-03
		Q3	2.903000e-03	2.911000e-03	2.908000e-03	2.908000e-03	2.906000e-03	2.910000e-03	2.909000e-03	2.911000e-03	2.910000e-03	2.908000e-03	2.909000e-03
		IQR	1.990000e-04	2.050000e-04	2.040000e-04	2.050000e-04	2.040000e-04	2.080000e-04	2.070000e-04	2.090000e-04	2.080000e-04	2.040000e-04	2.060000e-04
		Max Score	3.202000e-03	3.219000e-03	3.214000e-03	3.216000e-03	3.212000e-03	3.222000e-03	3.219000e-03	3.223000e-03	3.222000e-03	3.214000e-03	3.219000e-03
		Outlier samples	1.262690e+05	1.184490e+05	1.221020e+05	1.143220e+05	1.204010e+05	1.186740e+05	1.171610e+05	118400 .00	1.201600e+05	1.201370e+05	1.209380e+05
		Outlier Rates (%)	1.260000e+00	1.180000e+00	1.220000e+00	1.140000e+00	1.200000e+00	1.190000e+00	1.170000e+00	1.180000e+00	1.200000e+00	1.200000e+00	1.210000e+00

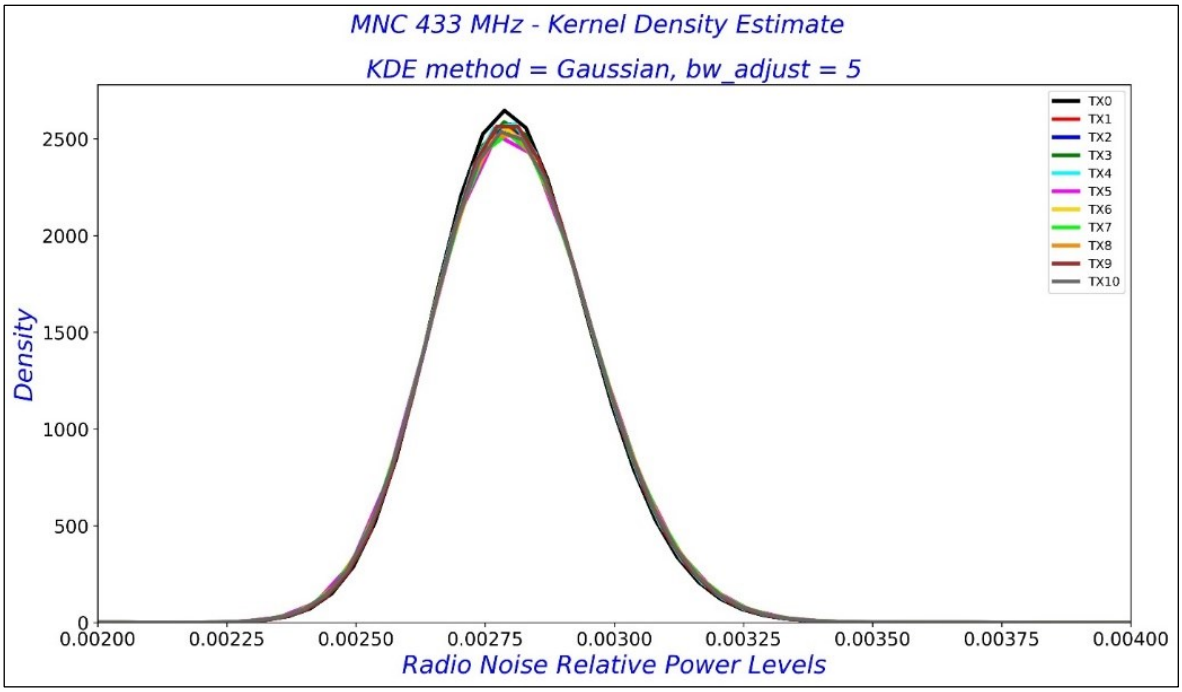

Figure 5.20: Density Distributions of IoT Radio Noise at MNC - 433 MHz

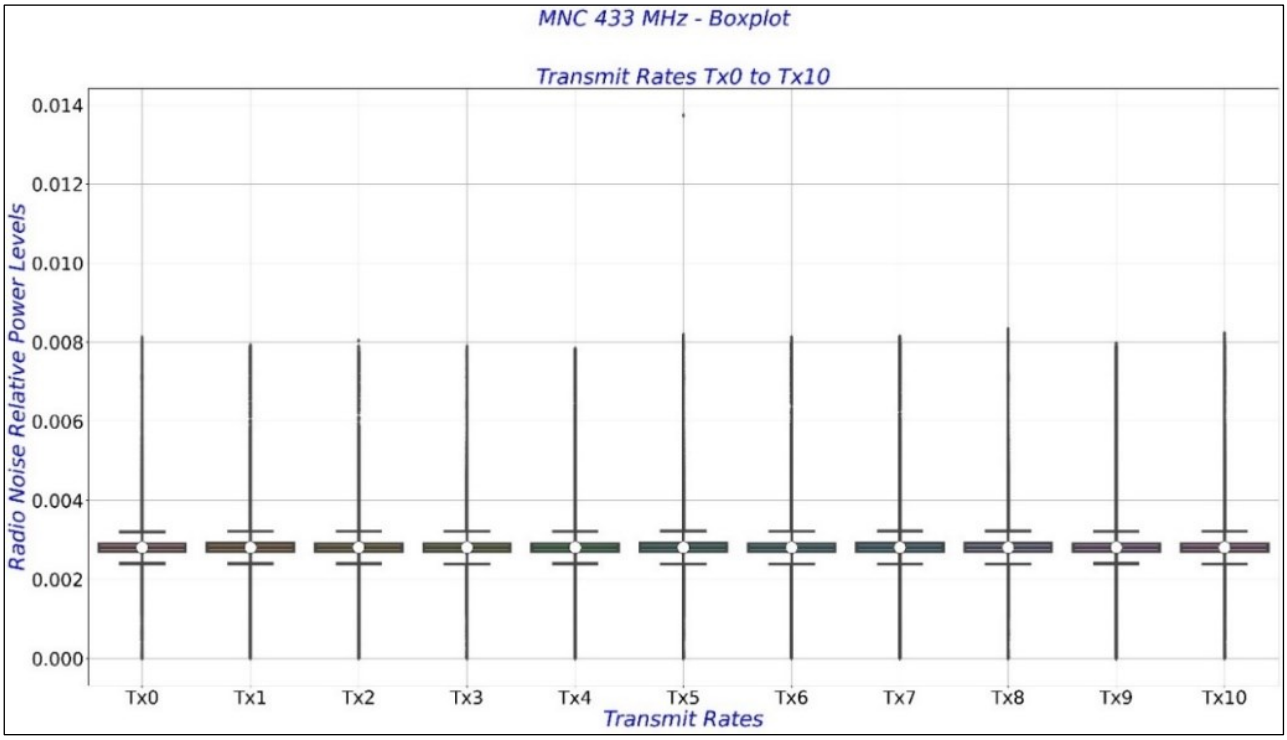

Figure 5.21: Full boxplots for IoT Radio Noise at MNC - 433 MHz

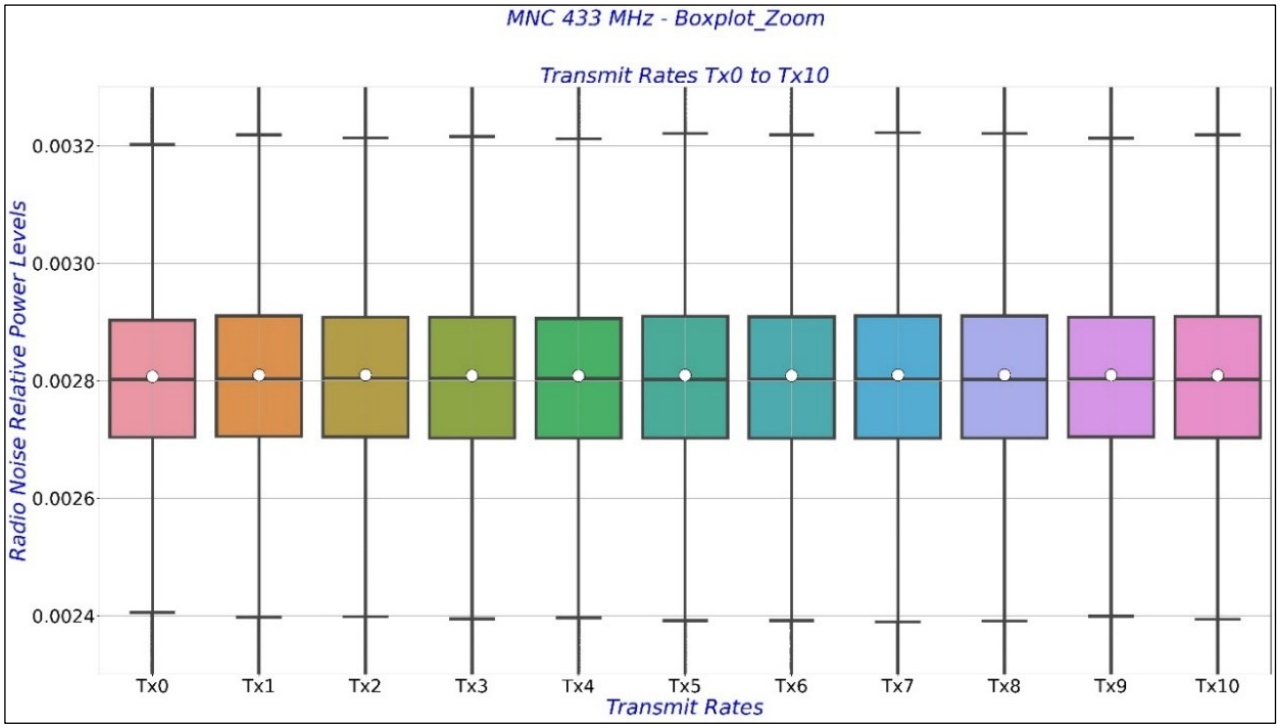

Figure 5.22: Boxplots for IoT Radio Noise at MNC - 433 MHz – no outliers

5.3.8. MNC 868 MHz

Table 5.8 shows the statistical metrics for the MNC 868 MHz case. The distributions expect no dispersion due to insignificant standard deviations and variance values. Most of the samples were, therefore, confined around the mean. From Tx0 to Tx10, the skewness values were unlikely to cause skewed distributions. Except for Tx10, which showed a minimum kurtosis of 26.65 mW, the

rest of the datasets exhibited higher kurtosis values above 100 mW, with the highest value of 352.30 mW at Tx1. These values may influence the peakedness or the smoothness of the distributions. Data experienced many outliers. Expect Tx0 and Tx1, which counted 1.8 % and 7.71 % of outlier rates; the rest of the dataset counted higher outlier rates, all above 10%. For example, Tx9 recorded 29.23 % as the highest outlier rate in this series. The median and mean showed approximate and comparable values. With slight standard deviations and variances, data showed characteristics of normal distributions, provided that further processing removed the outliers. The small IQR values confirmed that the distributions should see most data confined within a small range. Figure 5.23 corroborates the observations with peaked density distribution on Tx0 data. Tx1 to Tx10 exhibit flatter distributions. In Figure 5.24, the influence of the outliers obstructs the visibility of boxplots. Figure 5.25 achieves an improved resolution of visualisations after adjusting the y-axis. This result reinforces the necessity to remove the remove the outliers. Tx10 data exhibit a wider IQR compared to the rest of the data.

Table 5.8: Statistical characteristics of IoT Radio Noise at MNC - 868 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
MNC	868 MHz	Mean	1.031534e+00	1.033859e+00	1.034373e+00	1.034823e+00	1.036095e+00	1.037659e+00	1.039316e+00	1.041388e+00	1.044504e+00	1.051661e+00	1.062488e+00
		Std	3.053000e-03	1.037090e-01	1.154360e-01	1.239800e-01	1.475040e-01	1.714780e-01	1.937610e-01	2.177360e-01	2.499130e-01	3.103450e-01	3.805890e-01
		Var	9.000000e-06	1.075600e-02	1.332600e-02	1.537100e-02	2.175700e-02	2.940500e-02	3.754300e-02	4.740900e-02	6.245600e-02	9.631400e-02	1.448480e-01
		Kurt	1.031614e+02	3.523080e+02	2.810884e+02	269 .953500	2.313532e+02	1.825999e+02	1.432514e+02	1.117852e+02	8.353467e+01	5.038023e+01	26 .652107
		Skew	-2.665870e-01	1.453684e+01	1.300280e+01	1.299954e+01	1.239278e+01	1.118384e+01	9.929911e+00	8.773448e+00	7.593405e+00	5.877315e+00	4.108453e+00
		Min Score	1.024788e+00	1.024173e+00	1.023513e+00	1.023495e+00	1.023418e+00	1.023053e+00	1.022676e+00	1.022161e+00	1.020640e+00	1.014800e+00	9.704840e-01
		Q1	1.029820e+00	1.029672e+00	1.029470e+00	1.029452e+00	1.029462e+00	1.029395e+00	1.029354e+00	1.029258e+00	1.028867e+00	1.027586e+00	1.016448e+00
		Median	1.031244e+00	1.031352e+00	1.031320e+00	1.031347e+00	1.031330e+00	1.031263e+00	1.031296e+00	1.031288e+00	1.031254e+00	1.031320e+00	1.031485e+00
		Q3	1.033175e+00	1.033338e+00	1.033441e+00	1.033423e+00	1.033491e+00	1.033624e+00	1.033805e+00	1.033989e+00	1.034352e+00	1.036110e+00	1.047091e+00
		IQR	3.355000e-03	3.666000e-03	3.971000e-03	3.971000e-03	4.029000e-03	4.228000e-03	4.452000e-03	4.731000e-03	5.485000e-03	8.524000e-03	3.064300e-02
		Max Score	1.038206e+00	1.038837e+00	1.039398e+00	1.039380e+00	1.039535e+00	1.039966e+00	1.040483e+00	1.041087e+00	1.042578e+00	1.048896e+00	1.093055e+00
		Outlier samples	1.798160e+05	771340 .00	1.043045e+06	1087226 .00	1.110785e+06	1236250 .00	1.585319e+06	2000978 .00	2608905 .00	2923064 .00	1.557386e+06
		Outlier rates (%)	1.800000e+00	7.710000e+00	1.043000e+01	1.087000e+01	1.111000e+01	1.236000e+01	1.585000e+01	20 .01	2.609000e+01	29 .23	1.557000e+01

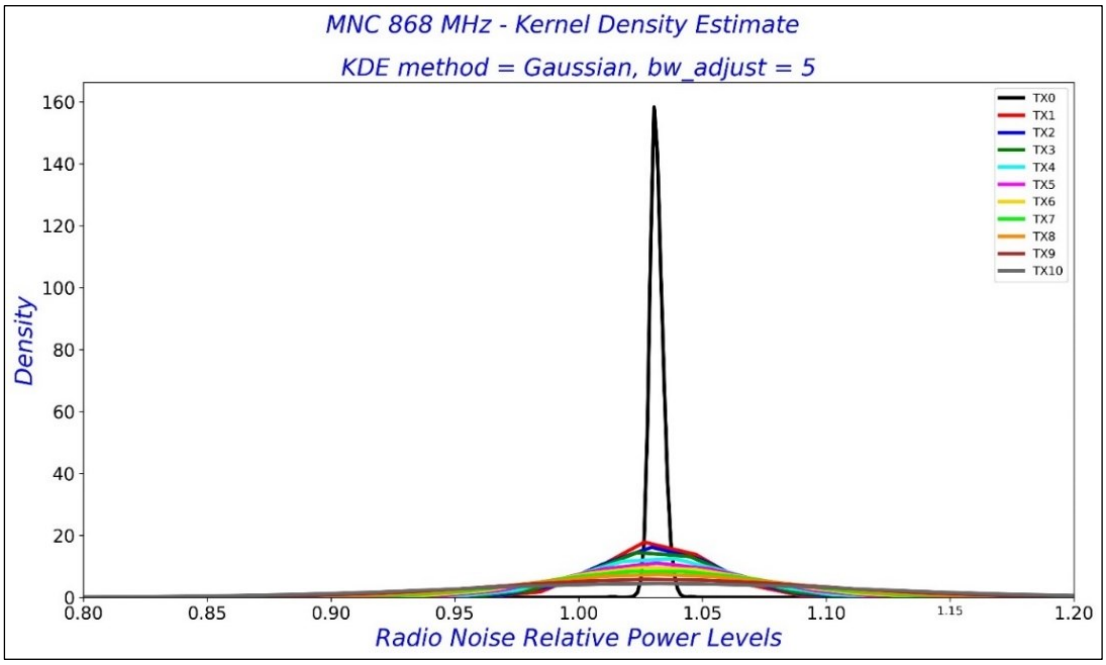

Figure 5.23: Density Distributions of IoT Radio Noise at MNC - 868 MHz

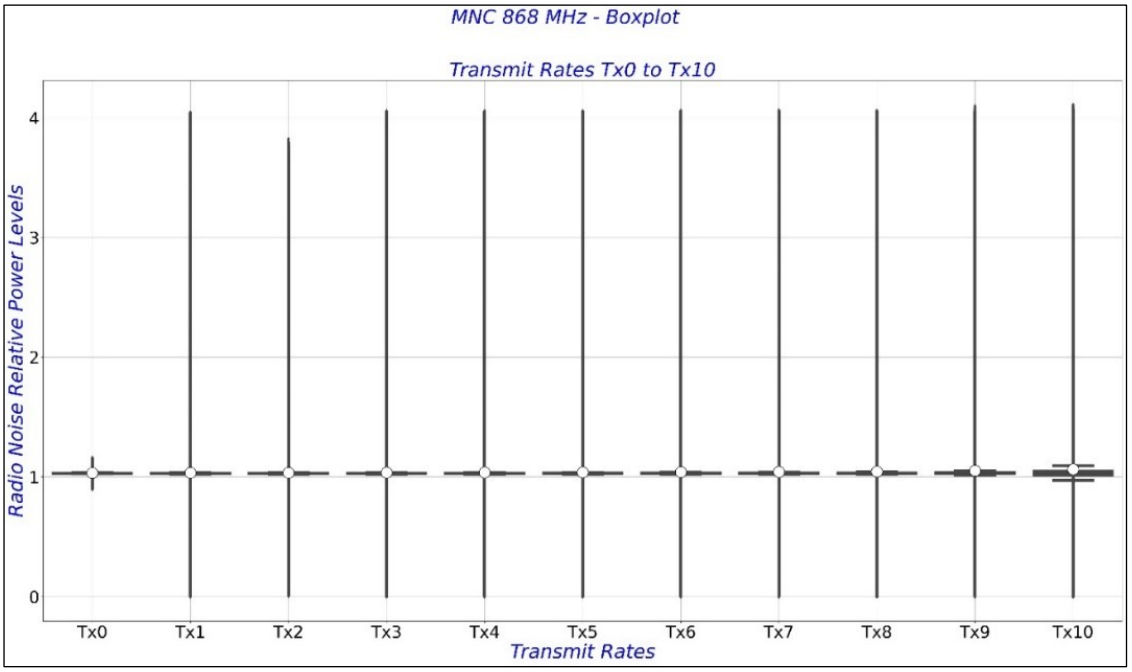

Figure 5.24: Full boxplots for IoT Radio Noise at MNC - 868 MHz

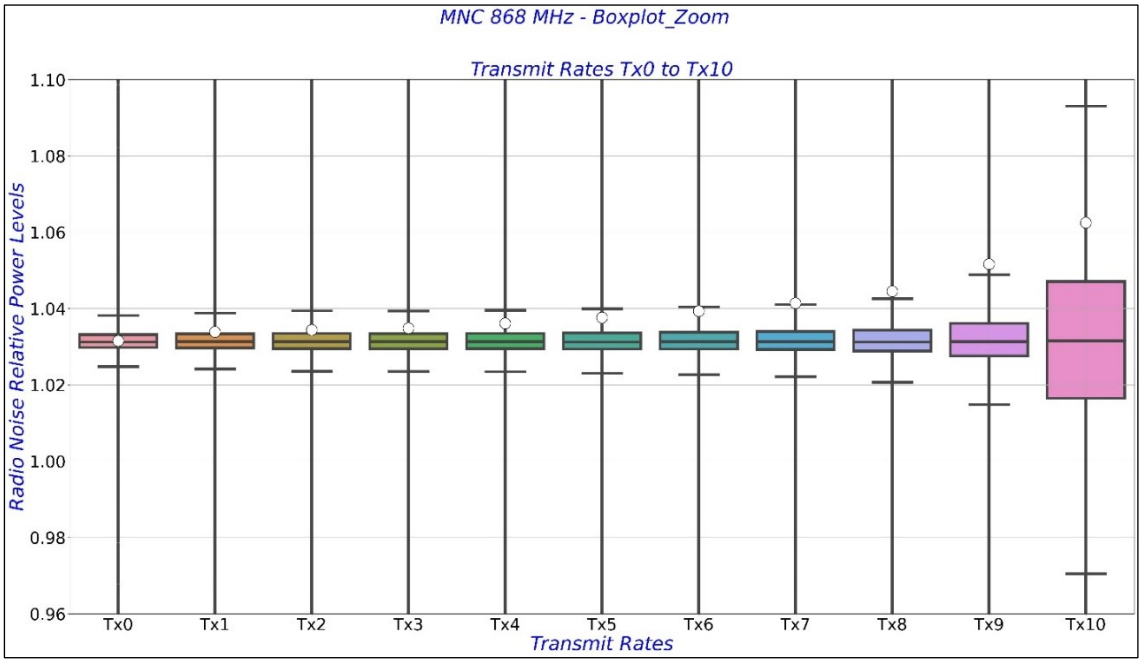

Figure 5.25: Boxplots for IoT Radio Noise at MNC - 868 MHz – no outliers

5.3.9. MNC 2400 MHz

Table 5.9 shows statistics for the MNC 2400 MHz case. They provide comparable mean and median values, approximately zero mW, insinuating the possibility of normal distributions. Weak standard deviations and variances suggested that the distributions have most samples squeezed around the mean. These data series carry significant kurtosis values above 66.61 mW at Tx0 up to 531.22 mW at Tx9. These values may be able to influence the peakedness and smoothness of distributions. Tx9 also produced the highest skewness factor of 7.21 mW. With that, these data

series should not lead to skewed distributions. Even though the outlier rates were low, for example, 3.19% as a maximum at Tx8, removing them would only improve the quality of distributions. Figure 5.26 corroborates these observations as the distributions show unsmooth responses. Figure 5.27 shows the influence of the outliers preventing from producing meaningful visualisation for the boxplots. Figure 5.28 simulates the extraction of the outliers by adjusting the y-axis. As a result, the resolution of visualisation improved with clear visibility of all elements of boxplots. All the datasets show to have comparable IQR.

Table 5.9: Statistical characteristics of IoT Radio Noise at MNC - 2400 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
MNC	2400 MHz	Mean	4.836000e-02	4.835500e-02	4.834800e-02	4.834300e-02	4.835500e-02	4.835300e-02	4.836300e-02	4.834500e-02	4.835400e-02	4.834500e-02	4.834500e-02
		Std	1.047000e-03	1.176000e-03	1.123000e-03	1.244000e-03	1.245000e-03	1.230000e-03	1.145000e-03	1.272000e-03	1.388000e-03	1.366000e-03	1.366000e-03
		Var	1.000000e-06	1.000000e-06	1.000000e-06	2.000000e-06	2.000000e-06	2.000000e-06	1.000000e-06	2.000000e-06	2.000000e-06	2.000000e-06	2.000000e-06
		Kurt	6.661487e+01	1.150995e+02	7.306716e+01	1.922035e+02	2.054518e+02	2.320643e+02	1.972265e+02	1.408181e+02	421 .256062	531 .223320	5.312233e+02
		Skew	1.577002e+00	2.190663e+00	1.754429e+00	3.501756e+00	3.574481e+00	3.864910e+00	3.309804e+00	2.711260e+00	6.107103e+00	7.214686e+00	7.214686e+00
		Min Score	4.604100e-02	4.595800e-02	4.577800e-02	4.592900e-02	4.600600e-02	4.546300e-02	4.597100e-02	4.563200e-02	4.583000e-02	4.585400e-02	4.585400e-02
		Q1	4.778000e-02	4.773400e-02	4.766900e-02	4.772300e-02	4.776000e-02	4.761200e-02	4.776000e-02	4.765200e-02	4.772300e-02	4.771900e-02	4.771900e-02
		Median	4.831200e-02	4.831900e-02	4.824000e-02	4.831900e-02	4.833100e-02	4.831900e-02	4.833100e-02	4.830500e-02	4.831900e-02	4.830500e-02	4.830500e-02
		Q3	4.894000e-02	4.891900e-02	4.892900e-02	4.891900e-02	4.892900e-02	4.904400e-02	4.895200e-02	4.899900e-02	4.898500e-02	4.896200e-02	4.896200e-02
		IQR	1.160000e-03	1.185000e-03	1.260000e-03	1.196000e-03	1.169000e-03	1.433000e-03	1.192000e-03	1.347000e-03	1.262000e-03	1.243000e-03	1.243000e-03
		Max Score	5.067900e-02	5.069600e-02	5.082000e-02	5.071300e-02	5.068300e-02	5.119300e-02	5.074100e-02	5.102000e-02	5.087800e-02	5.082700e-02	5.082700e-02
		Outlier samples	1.740940e+05	277245 .00	1.414470e+05	270591 .00	238426 .00	1.208920e+05	209789 .0	1.964760e+05	3.190650e+05	2.237710e+05	2.237710e+05
		Outlier rates (%)	1.740000e+00	2.770000e+00	1.410000e+00	2.710000e+00	2.380000e+00	1.210000e+00	2.100000e+00	1.960000e+00	3.190000e+00	2.240000e+00	2.240000e+00

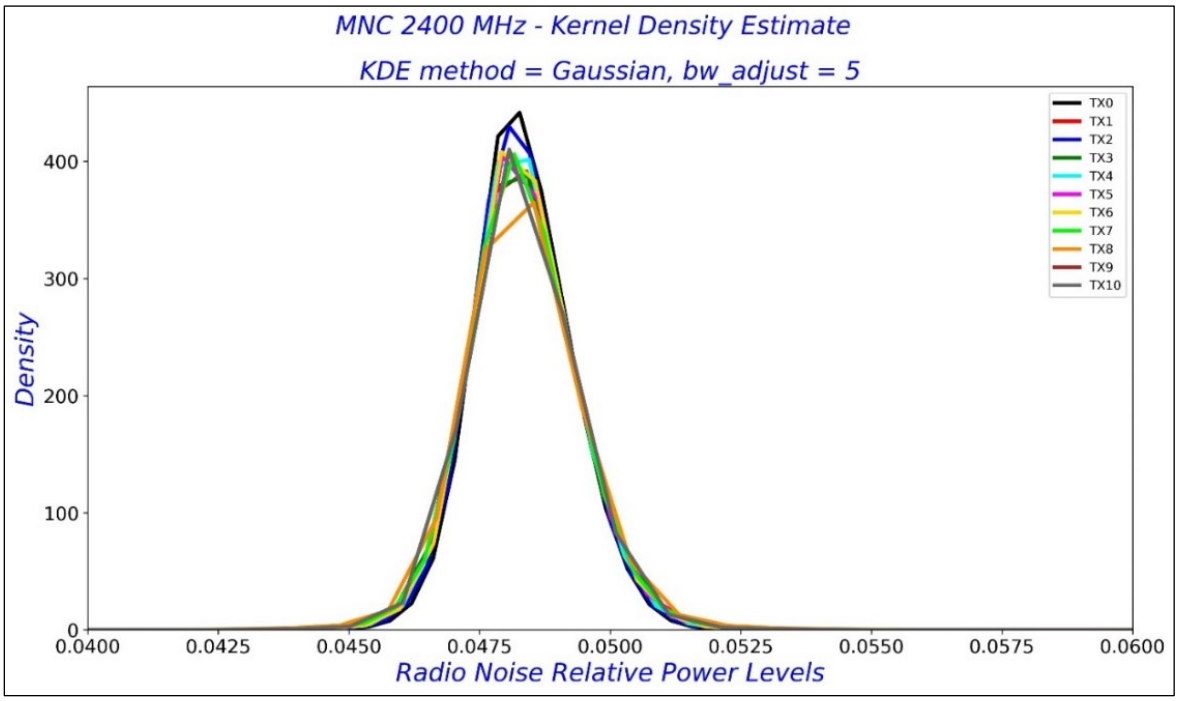

Figure 5.26: Density Distributions of IoT Radio Noise at MNC - 2400 MHz

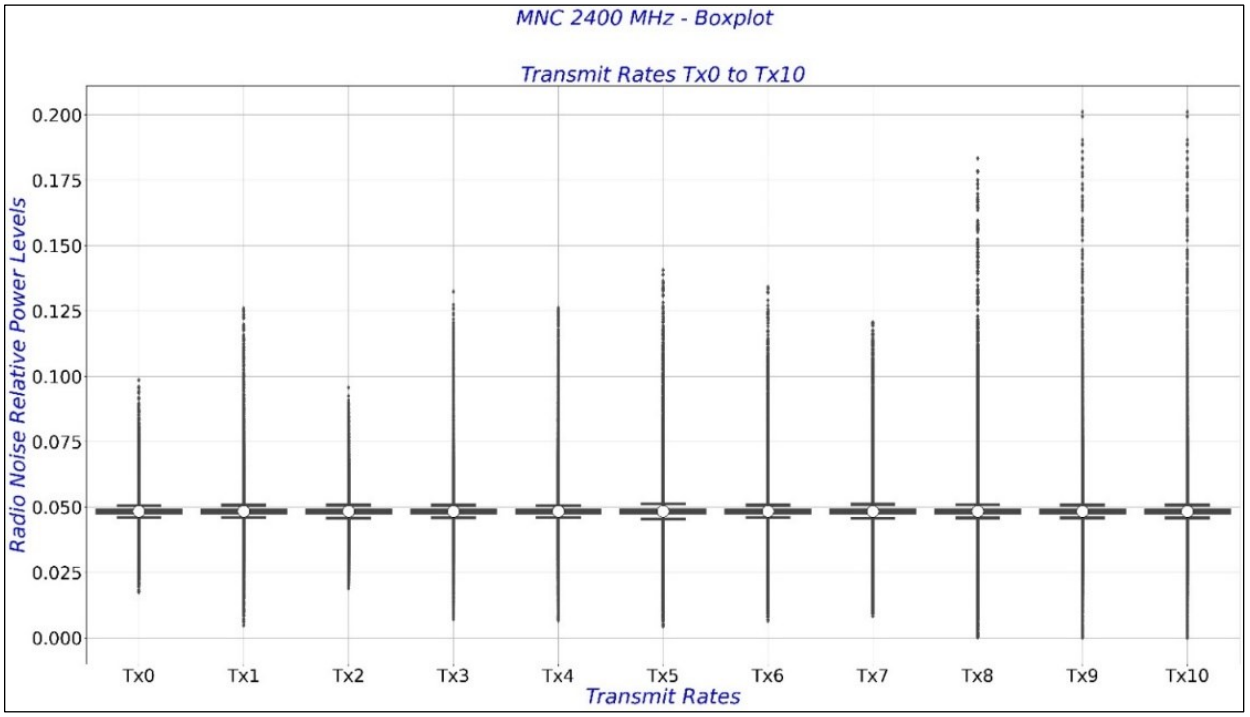

Figure 5.27: Full boxplots for IoT Radio Noise at MNC - 2400 MHz

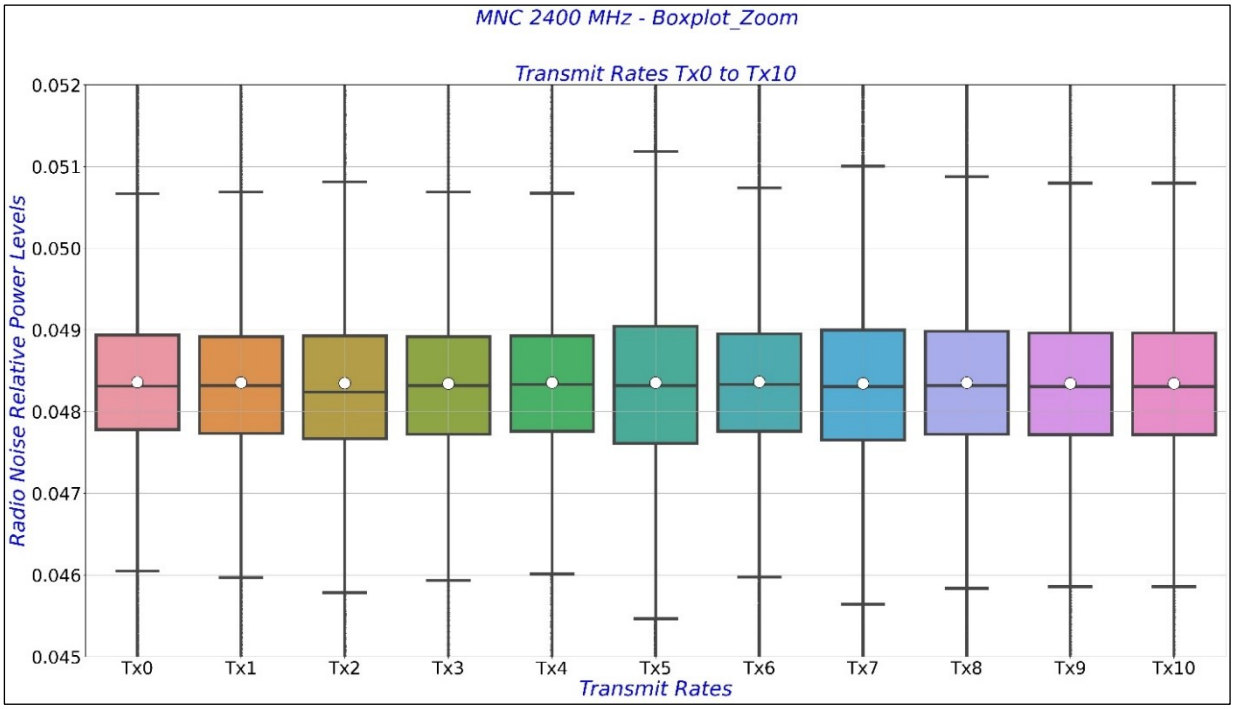

Figure 5.28: Boxplots for IoT Radio Noise at MNC - 2400 MHz – no outliers

5.3.10. MNS 433 MHz

Table 5.10 shows values for statistical metrics for the MNS 433 MHz case carried approximately zero mW for standard deviations and variances. These data did not produce dispersed distributions. The insignificant skewness values could only obtain symmetrical distributions between 0.96 mW (minimum) at Tx8 and 4.57 mW at Tx7 (maximum). The kurtosis was not negligible in this case.

With a minimum of 37.38 mW at Tx8 and a maximum of 143.57 mW at Tx7, these kurtosis may influence the shape of distributions. The mean and median showed remarkable differences, especially at Tx1, Tx3, Tx4 and Tx7. Though the outlier rates were low, their quantity was considerable. Figure 5.29 showed almost symmetrical distributions due to low skewness. The kurtosis must have impacted the smoothness of the distributions. The figure shows that Tx8, Tx9, and Tx10 carried the highest mean values compared to the rest of the transmit rate data. Boxplots in Figure 5.30 confirm the presence of the outliers. Figure 5.31 improves the resolution of the boxplots due to adjusting the scale of the y-axis.

Table 5.10: Statistical characteristics of IoT Radio Noise at MNS - 433 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
MNS	433 MHz	Mean	2.849332e-03	2.850177e-03	2.850606e-03	2.850249e-03	2.850099e-03	2.849311e-03	2.850197e-03	2.851003e-03	2.926515e-03	2.926636e-03	2.925853e-03
		Std	1.973871e-04	1.970946e-04	2.046453e-04	2.108025e-04	2.122811e-04	1.934846e-04	2.042927e-04	2.102520e-04	1.632141e-04	1.732183e-04	1.722029e-04
		Var	3.896166e-08	3.884629e-08	4. 187971e-08	4 .443770e-08	4.506327e-08	3.743630e-08	4. 173549e-08	4 .420589e-08	2.663883e-08	3.000458e-08	2.965383e-08
		Kurt	1.055790e+02	1.039017e+02	1.035650e+02	1.400179e+02	1.371945e+02	1.118699e+02	1.047022e+02	1.435793e+02	3.738170e+01	4.111549e+01	4.061442e+01
		Skew	3.567531e+00	3.523264e+00	3.210525e+00	4.485429e+00	4.410713e+00	3.551090e+00	3.237364e+00	4.579399e+00	9.607908e-01	1.123458e+00	1.112751e+00
		Min Score	2.393000e-03	2.393000e-03	2.377000e-03	2.372000e-03	2.368000e-03	2.380000e-03	2.377000e-03	2.372000e-03	2.526000e-03	2.500000e-03	2.501000e-03
		Q1	2.736000e-03	2.736000e-03	2.729000e-03	2.729000e-03	2.726000e-03	2.726000e-03	2.729000e-03	2.729000e-03	2.824000e-03	2.818000e-03	2.817000e-03
		Median	2.844000e-03	2.844000e-03	2.851000e-03	2.844000e-03	2.844000e-03	2.844000e-03	2.851000e-03	2.844000e-03	2.921000e-03	2.920000e-03	2.919000e-03
		Q3	2.964000e-03	2.964000e-03	2.964000e-03	2.968000e-03	2.964000e-03	2.956000e-03	2.964000e-03	2.968000e-03	3.023000e-03	3.030000e-03	3.028000e-03
		IQR	2.280000e-04	2.280000e-04	2.350000e-04	2.380000e-04	2.380000e-04	2.300000e-04	2.350000e-04	2.380000e-04	1.990000e-04	2.120000e-04	2.110000e-04
		Max Score	3.307000e-03	3.307000e-03	3.316000e-03	3.325000e-03	3.322000e-03	3.301000e-03	3.316000e-03	3.325000e-03	3.322000e-03	3.348000e-03	3.344000e-03
		Outlier samples	1.103610e+05	1.102920e+05	1.259660e+05	1.301510e+05	1.289160e+05	1.010900e+05	1.237820e+05	127640 .00	1.243640e+05	1.220260e+05	1.236360e+05
		Outlier rates (%)	1.100000e+00	1.100000e+00	1.260000e+00	1.300000e+00	1.290000e+00	1.010000e+00	1.240000e+00	1.280000e+00	1.240000e+00	1.220000e+00	1.240000e+00

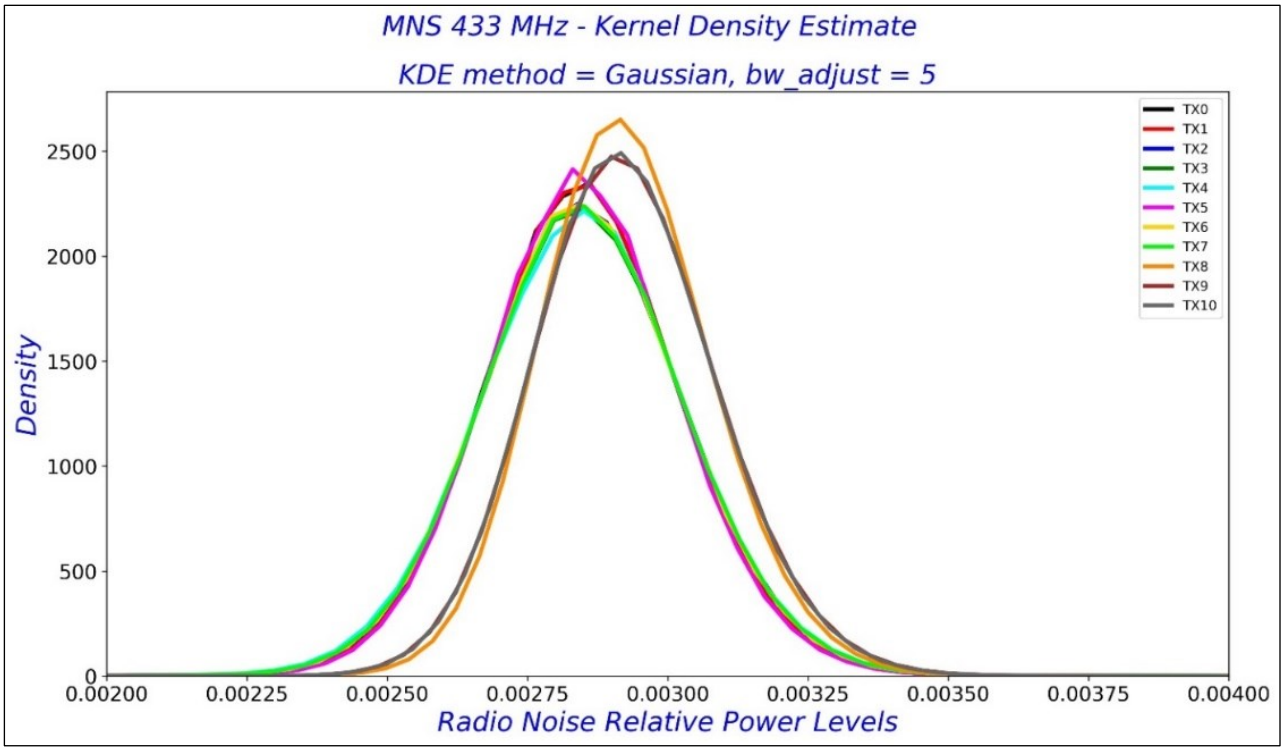

Figure 5.29: Density Distributions of IoT Radio Noise at MNS - 433 MHz

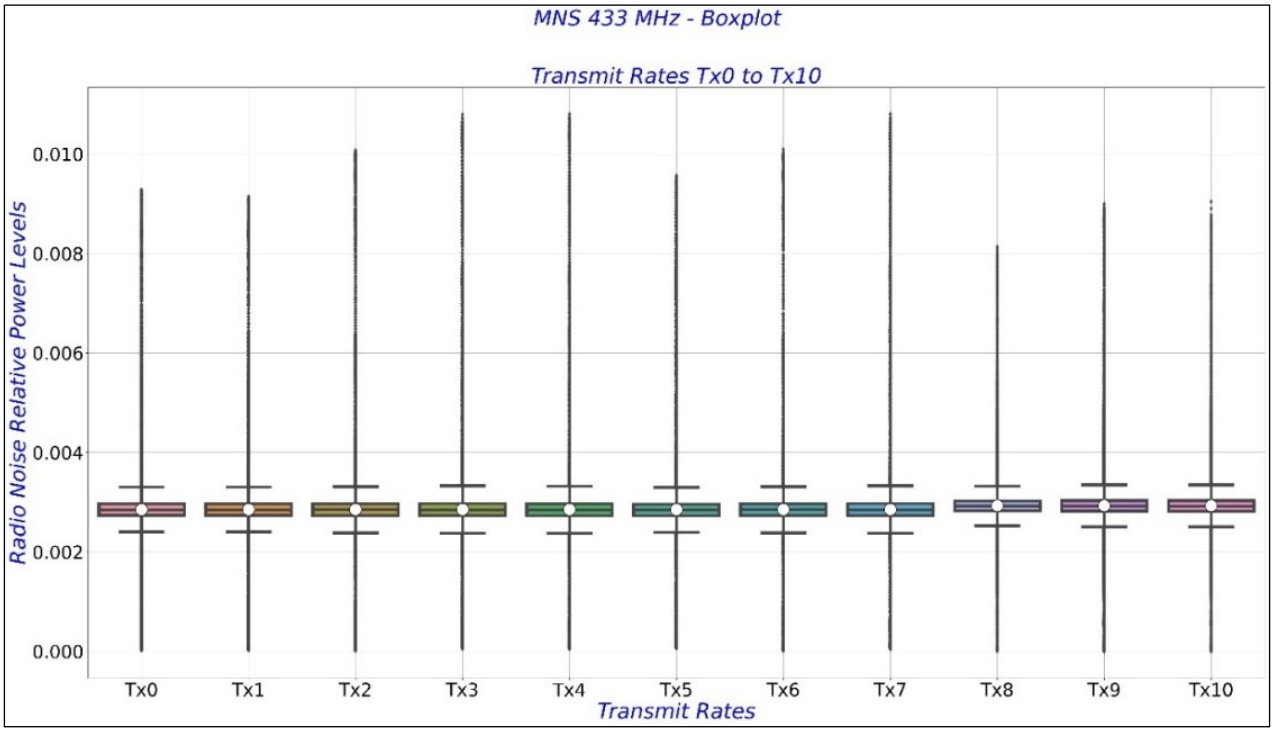

Figure 5.30: Full boxplots for IoT Radio Noise at MNS - 433 MHz

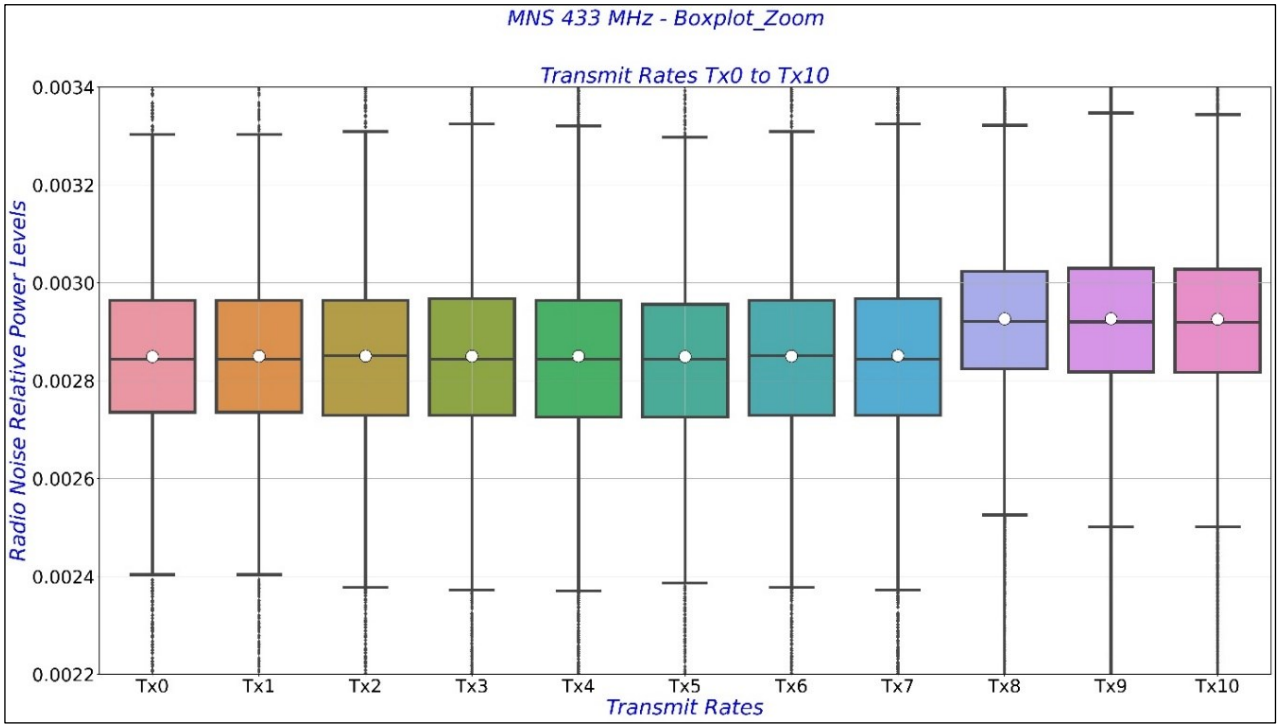

Figure 5.31: Boxplots for IoT Radio Noise at MNS - 433 MHz – no outliers

5.3.11. MNS 868 MHz

Table 5.11 shows the statistical metrics for the MNS 868 MHz case. The distributions expect no dispersion due to insignificant standard deviation and variance values. The distributions show most of the samples confined to the mean values. The skewness values, for example, 14.29 mW (highest) at Tx1, may not be significant enough to cause skew distributions. These data series

exhibited strong kurtosis as high as 506.39 mW at Tx1. The weakest kurtosis was 25.58 mW at Tx10. These datasets may influence the peakedness or the smoothness of the distributions. 24.19 % at Tx9 was the highest rate of outliers in this series. The fact that the median and mean showed approximate and comparable values envisioned the possibility for normal distributions, provided further processing removed the outliers. Figure 5.32 confirms the observations as some data introduced unsmooth and peaked density distributions. In Figure 5.33, the boxplots could not be visible due to high rates of kurtosis. To palliate, Figure 5.34 achieves improved resolution of boxplot visualisations after adjusting the y-axis. The results reinforce that it is necessary to remove the outliers.

Table 5.11: Statistical characteristics of IoT Radio Noise at MNS - 868 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
MNS	868 MHz	Mean	3.253370e-01	3.255740e-01	3.256310e-01	3.256810e-01	3.256980e-01	3.257790e-01	3.266450e-01	3.264200e-01	3.274160e-01	3.278000e-01	3.308350e-01
		Std	1.695000e-03	1.820000e-02	1.994500e-02	2.152600e-02	2.204500e-02	2.475400e-02	4.332000e-02	3.971500e-02	5.483700e-02	5.877600e-02	8.769100e-02
		Var	3.000000e-06	3.310000e-04	3.980000e-04	4.630000e-04	4.860000e-04	6.130000e-04	1.877000e-03	1.577000e-03	3.007000e-03	3.455000e-03	7.690000e-03
		Kurt	8.509975e+01	5.063961e+02	382 .710793	3.615660e+02	3.039141e+02	254 .955311	1.588359e+02	2.018629e+02	9.776824e+01	5.855614e+01	2.558241e+01
		Skew	-1.349230e-01	1.429238e+01	1.223604e+01	1.231678e+01	1.087094e+01	1.011463e+01	9.321148e+00	1.064712e+01	7.358712e+00	5.219355e+00	3.505551e+00
		Min Score	3.215240e-01	3.213180e-01	3.211910e-01	3.211440e-01	3.211090e-01	3.210870e-01	3.208380e-01	3.205130e-01	3.199140e-01	3.188910e-01	3.027510e-01
		Q1	3.243620e-01	3.243160e-01	3.242830e-01	3.242760e-01	3.242700e-01	3.242610e-01	3.242030e-01	3.241110e-01	3.239660e-01	3.237660e-01	3.197480e-01
		Median	3.252170e-01	3.251640e-01	3.252030e-01	3.252070e-01	3.252160e-01	3.252460e-01	3.252520e-01	3.252250e-01	3.252290e-01	3.252020e-01	3.252410e-01
		Q3	3.262540e-01	3.263150e-01	3.263440e-01	3.263630e-01	3.263770e-01	3.263780e-01	3.264470e-01	3.265100e-01	3.266670e-01	3.270150e-01	3.310790e-01
		IQR	1.892000e-03	1.999000e-03	2.061000e-03	2.088000e-03	2.107000e-03	2.116000e-03	2.244000e-03	2.399000e-03	2.701000e-03	3.250000e-03	1.133100e-02
		Max Score	3.290910e-01	3.293130e-01	3.294360e-01	3.294950e-01	3.295370e-01	3.295520e-01	3.298120e-01	3.301090e-01	3.307190e-01	3.318900e-01	3.480750e-01
		Outlier samples	1.689870e+05	5.127230e+05	5.848190e+05	593883 .00	7.041130e+05	8.015340e+05	1.133130e+06	1.387172e+06	1.932979e+06	2.418663e+06	1.262368e+06
		Outlier rates (%)	1.690000e+00	5.130000e+00	5.850000e+00	5.940000e+00	7.040000e+00	8.020000e+00	1.133000e+01	1.387000e+01	1.933000e+01	2.419000e+01	1.262000e+01

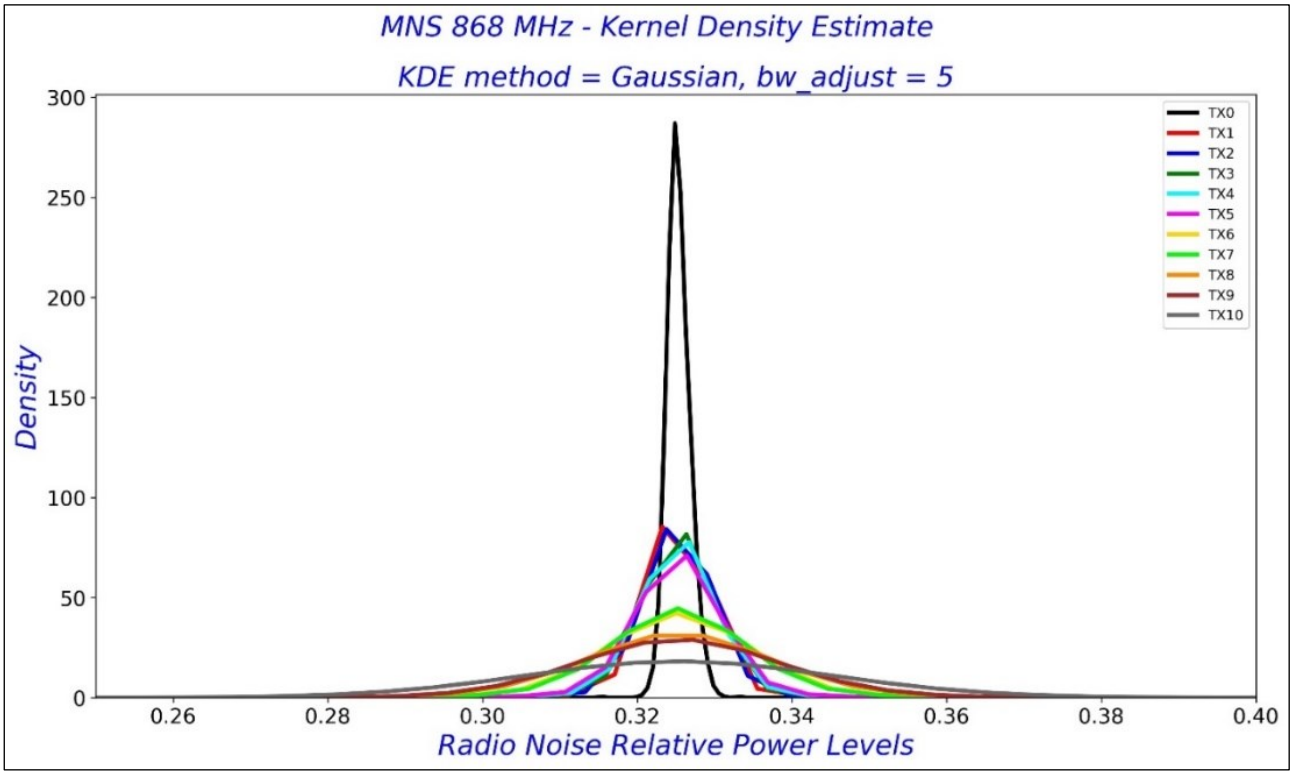

Figure 5.32: Density Distributions of IoT Radio Noise at MNS - 868 MHz

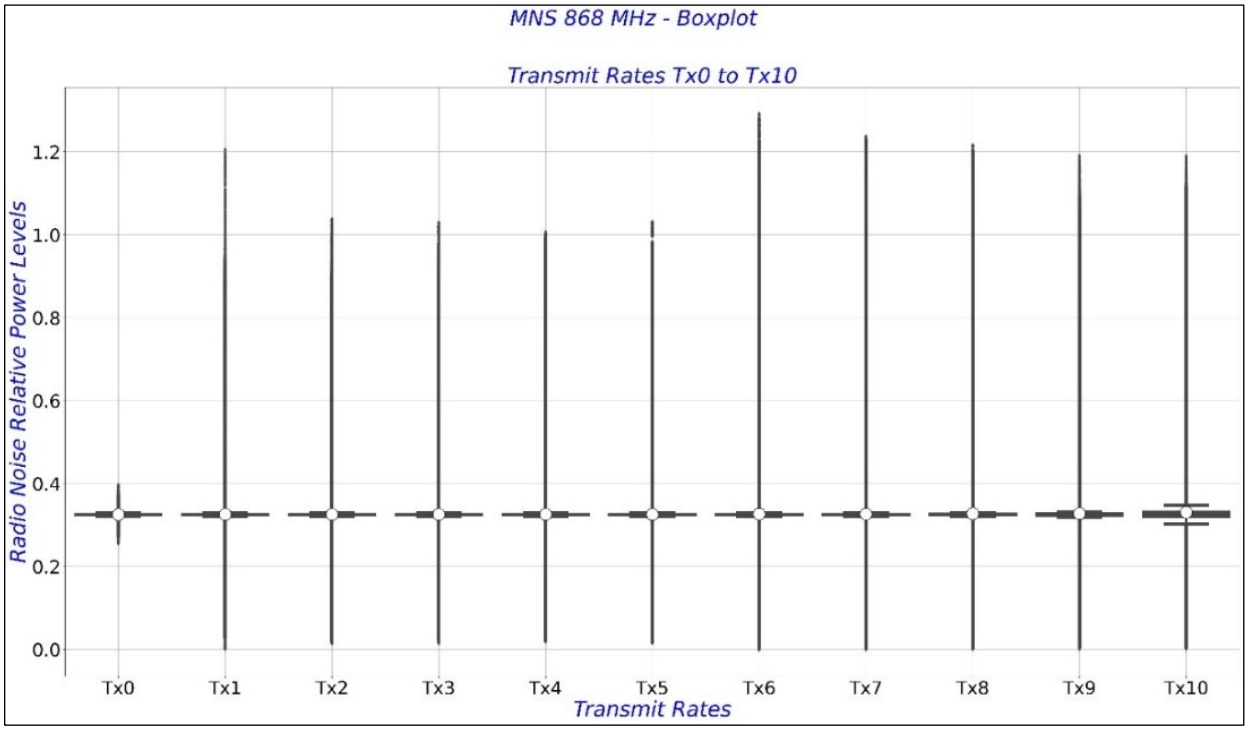

Figure 5.33: Full boxplots for IoT Radio Noise at MNS - 868 MHz

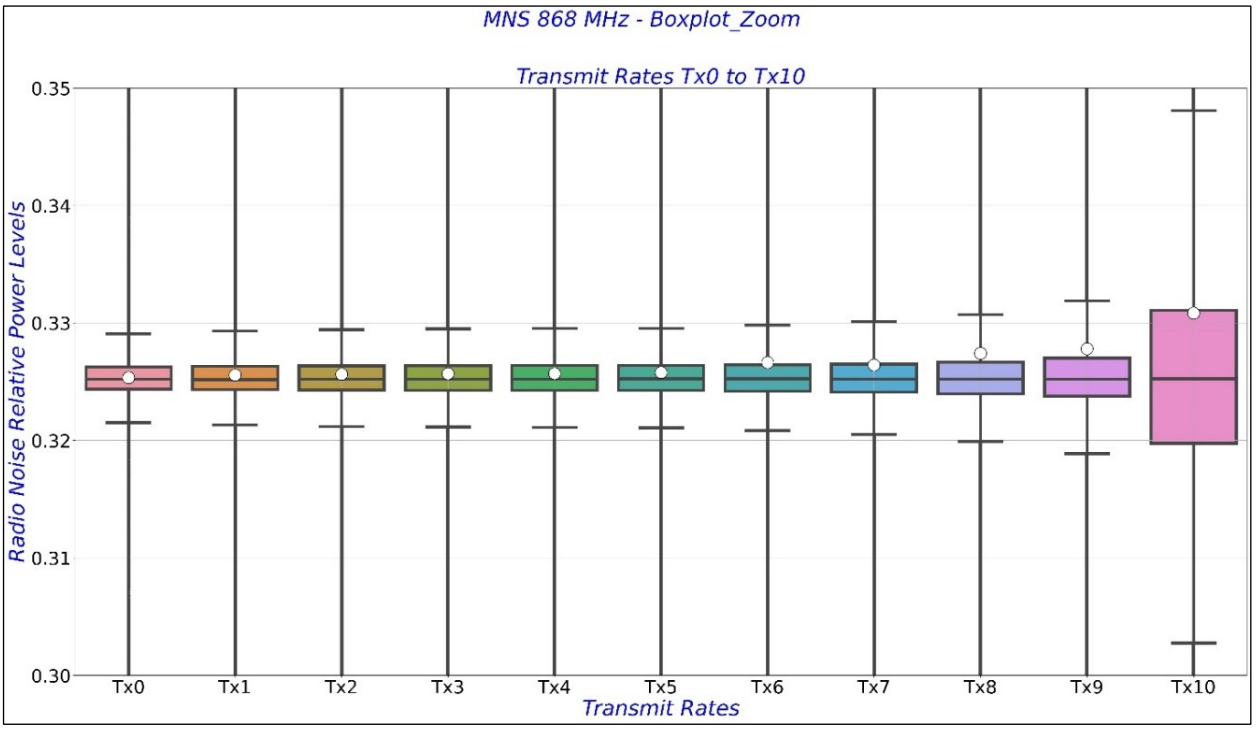

Figure 5.34: Boxplots for IoT Radio Noise at MNS - 868 MHz – no outliers

5.3.12. MNS 2400 MHz

Table 5.12 shows statistics for the MNS 2400 MHz case. The mean and median values are comparable, and both are approximately zero mW, insinuating the possibility of normal distributions. Weak standard deviations and variances suggested that most samples would fit around the mean. These data series carry varied kurtosis values with a minimum of 22.25 mW at

Tx4 and a maximum of 1148 mW at Tx9. These values seem high enough to influence the peakedness and smoothness of distributions. Tx8 produced the highest skewness factor of 10.73 mW, which should not lead to skewed distributions. Tx9 produced 2.1% as the maximum outlier rate, meaning the presence of the outliers was weak. Figure 5.35 corroborates these observations as the distributions show unsmooth responses. Figure 5.36 shows the influence of the outliers, as poor visualisation for the boxplots prevents them from getting insights from these responses. Figure 5.37 simulates the extraction of the outliers by adjusting the y-axis. As a result, the resolution of visualisation improved with clear visibility of all elements of boxplots.

Table 5.12: Statistical characteristics of IoT Radio Noise at MNS - 2400 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
MNS	2400 MHz	Mean	7.023000e-02	7.022700e-02	7.021700e-02	7.021600e-02	7.022500e-02	7.022500e-02	7.022800e-02	7.022600e-02	7.021600e-02	7.021600e-02	7.023600e-02
		Std	1.332000e-03	1.212000e-03	1.300000e-03	1.218000e-03	1.203000e-03	1.203000e-03	1.228000e-03	1.259000e-03	1.465000e-03	1.307000e-03	1.340000e-03
		Var	2.000000e-06	1.000000e-06	2.000000e-06	1.000000e-06	1.000000e-06	1.000000e-06	2.000000e-06	2.000000e-06	2.000000e-06	2.000000e-06	2.000000e-06
		Kurt	5.436486e+02	2.446267e+01	1.089930e+02	3.146853e+01	2.225518e+01	2.225518e+01	3.177114e+01	5.677470e+01	9.757539e+02	1.148007e+03	9.250974e+01
		Skew	5.842174e+00	1.057966e+00	2.055575e+00	1.163357e+00	1.066508e+00	1.066508e+00	1.157354e+00	1.449646e+00	1.073525e+01	1.004095e+01	1.834259e+00
		Min Score	6.727200e-02	6.743500e-02	6.707500e-02	6.707500e-02	6.743500e-02	6.743500e-02	6.743500e-02	6.743500e-02	6.724700e-02	6.707500e-02	6.743500e-02
		Q1	6.943700e-02	6.950700e-02	6.940600e-02	6.940600e-02	6.950700e-02	6.950700e-02	6.950700e-02	6.950700e-02	6.947500e-02	6.940600e-02	6.950700e-02
		Median	7.012600e-02	7.019600e-02	7.012600e-02	7.012600e-02	7.019600e-02	7.019600e-02	7.019600e-02	7.019600e-02	7.016300e-02	7.012600e-02	7.019600e-02
		Q3	7.088100e-02	7.088800e-02	7.096000e-02	7.096000e-02	7.088800e-02	7.088800e-02	7.088800e-02	7.088800e-02	7.096000e-02	7.096000e-02	7.088800e-02
		IQR	1.444000e-03	1.381000e-03	1.554000e-03	1.554000e-03	1.381000e-03	1.381000e-03	1.381000e-03	1.381000e-03	1.485000e-03	1.554000e-03	1.381000e-03
		Max Score	7.304600e-02	7.296000e-02	7.329000e-02	7.329000e-02	7.296000e-02	7.296000e-02	7.296000e-02	7.296000e-02	7.318700e-02	7.329000e-02	7.296000e-02
		Outlier samples	1.480870e+05	1.917600e+05	1.094760e+05	86504 .00	1.858640e+05	185864 .00	1.975480e+05	1.958830e+05	1.500630e+05	8.567300e+04	2.099580e+05
		Outliers rates (%)	1.480000e+00	1.920000e+00	1.090000e+00	8.700000e-01	1.860000e+00	1.860000e+00	1.980000e+00	1.960000e+00	1.500000e+00	8.600000e-01	2.100000e+00

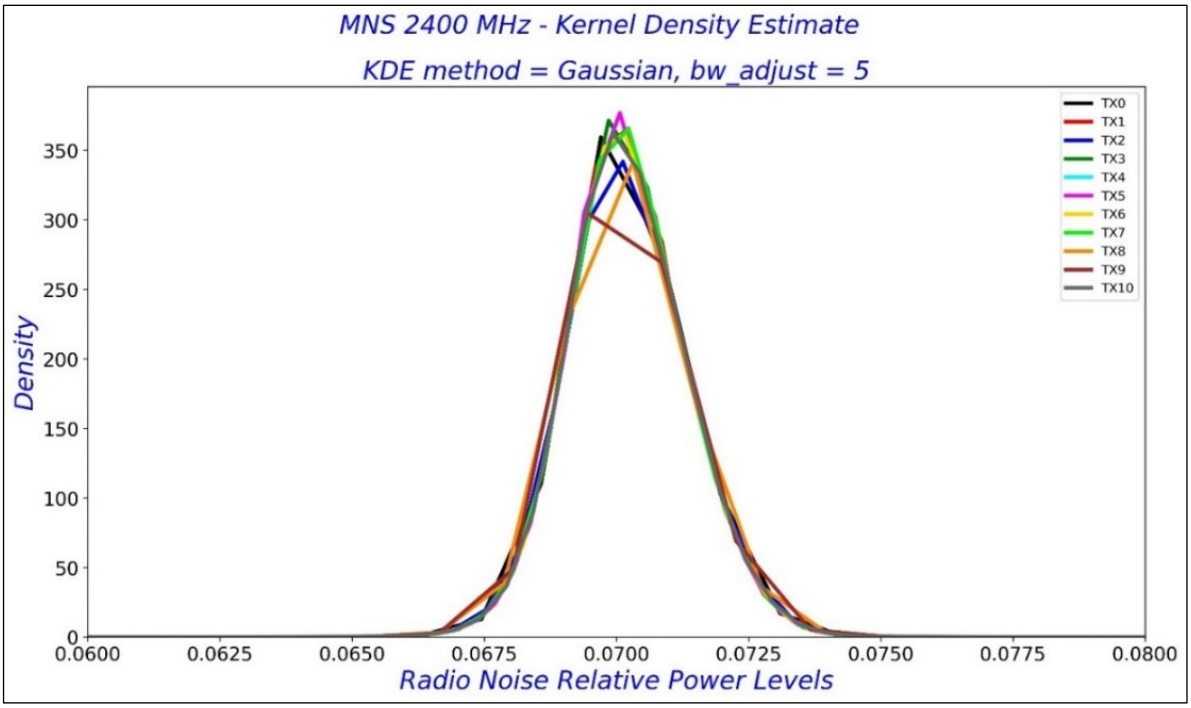

Figure 5.35: Density Distributions of IoT Radio Noise at MNS - 2400 MHz

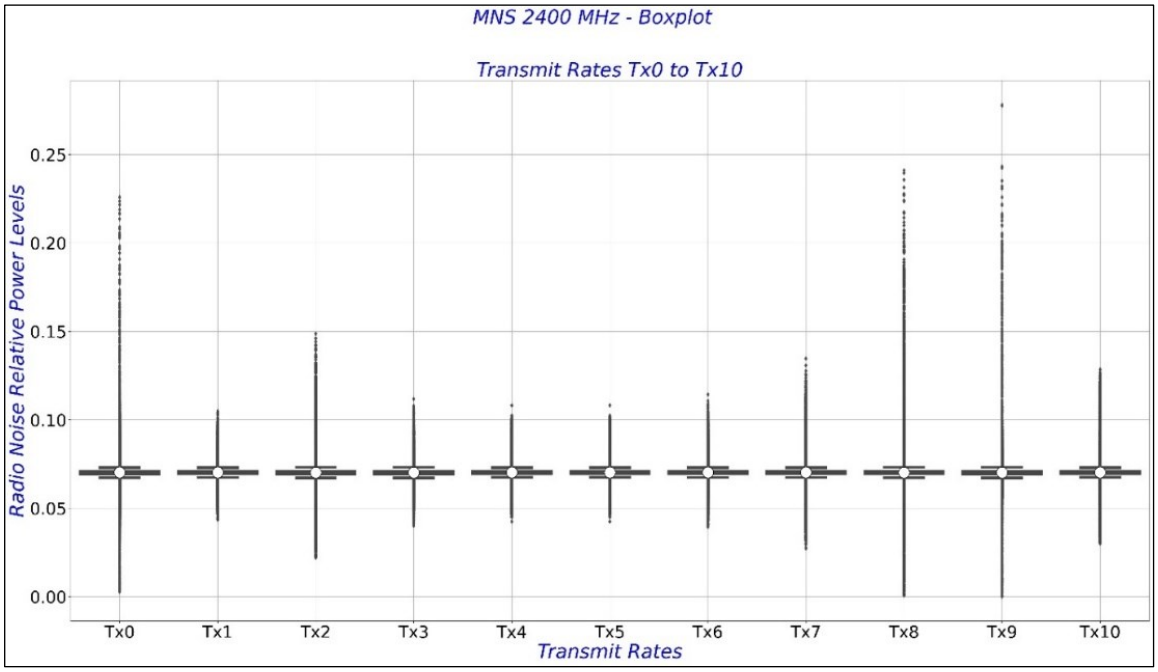

Figure 5.36: Full boxplots for IoT Radio Noise at MNS - 2400 MHz

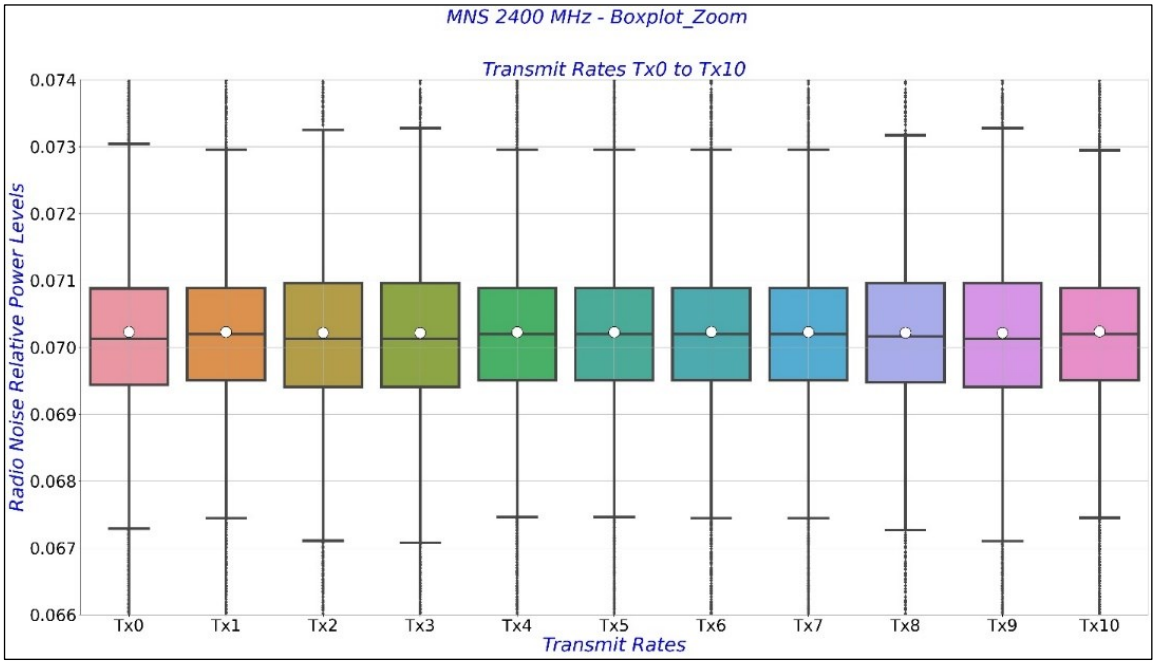

Figure 5.37: Boxplots for IoT Radio Noise at MNS - 2400 MHz – no outliers

5.3.13. NGM 433 MHz

Table 5.13 shows that the NGM 433 MHz case also expects confined density distributions around the means due to weak standard deviations and variance values. The slight skewness could not impact the distribution skew. The insignificant kurtosis values could not influence the shape of distributions. The mean and median are comparable and approximately zero mW. These data expect responses similar to those of normal distributions. Data obtained 1.18 % as the maximum outlier rate. Figure 5.38 shows symmetrical bell shape distributions with comparable mean values

from Tx0 to Tx10. However, some datasets did not show smooth curves. Figure 5.39 visualises all the elements of the boxplots as well as the outlier portions. Figure 5.40 simulates the removal of outliers by tuning the scale of the y-axis. The results improve the resolution of boxplots.

Table 5.13: Statistical characteristics of IoT Radio Noise at NGM - 433 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
NGM	433 MHz	Mean	2.593588e-03	2.594416e-03	2.594282e-03	2.594514e-03	2.595042e-03	2.595042e-03	2.594640e-03	2.595098e-03	2.594873e-03	2.595200e-03	2.594825e-03
		Std	1.570943e-04	1.650605e-04	1.609160e-04	1.650200e-04	1.680573e-04	1.680573e-04	1.673369e-04	1.688190e-04	1.690994e-04	1.704589e-04	1.628456e-04
		Var	2.467861e-08	2.724496e-08	2.589395e-08	2.723161e-08	2.824326e-08	2.824326e-08	2.800164e-08	2.849987e-08	2.859461e-08	2.905622e-08	2.651869e-08
		Kurt	2.567341e+01	2.341687e+01	2.379263e+01	2.157618e+01	2.023490e+01	2.023490e+01	2.120373e+01	2.058646e+01	2.078468e+01	1.982651e+01	2.110424e+01
		Skew	7.426763e-01	7.322625e-01	7.185692e-01	6.767359e-01	6.537486e-01	6.537486e-01	6.755379e-01	6.702899e-01	6.774145e-01	6.532230e-01	6.498476e-01
		Min Score	2.198000e-03	2.183000e-03	2.194000e-03	2.181000e-03	2.172000e-03	2.172000e-03	2.174000e-03	2.172000e-03	2.171000e-03	2.165000e-03	2.189000e-03
		Q1	2.493000e-03	2.490000e-03	2.492000e-03	2.489000e-03	2.487000e-03	2.487000e-03	2.488000e-03	2.487000e-03	2.486000e-03	2.485000e-03	2.492000e-03
		Median	2.588000e-03	2.590000e-03	2.589000e-03	2.589000e-03	2.590000e-03	2.590000e-03	2.588000e-03	2.590000e-03	2.589000e-03	2.589000e-03	2.590000e-03
		Q3	2.689000e-03	2.694000e-03	2.691000e-03	2.694000e-03	2.697000e-03	2.697000e-03	2.697000e-03	2.697000e-03	2.697000e-03	2.699000e-03	2.693000e-03
		IQR	1.960000e-04	2.040000e-04	1.990000e-04	2.050000e-04	2.100000e-04	2.100000e-04	2.090000e-04	2.100000e-04	2.100000e-04	2.140000e-04	2.010000e-04
		Max Score	2.984000e-03	3.001000e-03	2.989000e-03	3.002000e-03	3.011000e-03	3.011000e-03	3.012000e-03	3.013000e-03	3.012000e-03	3.019000e-03	2.995000e-03
		Outlier samples	1.125170e+05	1.180620e+05	1.221670e+05	1.177000e+05	1.177920e+05	1.177920e+05	1.122440e+05	1.175380e+05	1.196580e+05	1.121150e+05	1.198980e+05
		Outlier rates (%)	1.130000e+00	1.180000e+00	1.220000e+00	1.180000e+00	1.180000e+00	1.180000e+00	1.120000e+00	1.180000e+00	1.200000e+00	1.120000e+00	1.200000e+00

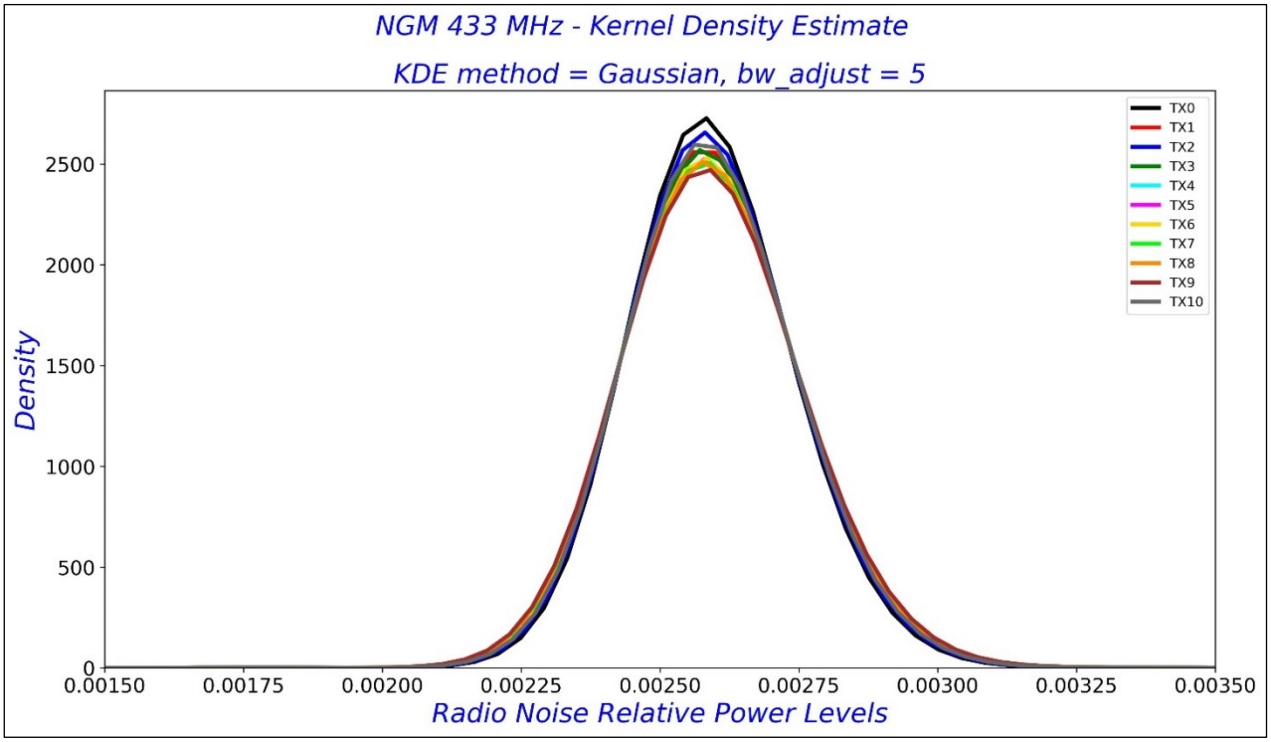

Figure 5.38: Density Distributions of IoT Radio Noise at NGM - 433 MHz

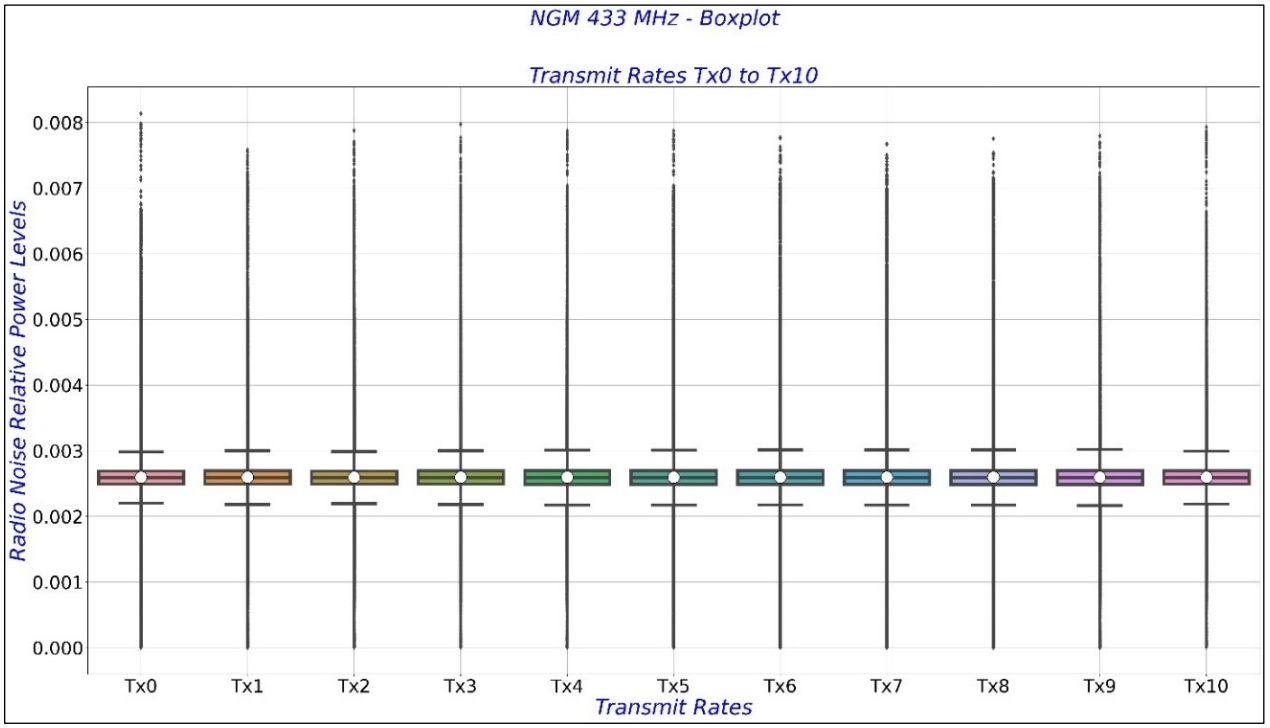

Figure 5.39: Full boxplots for IoT Radio Noise at NGM - 433 MHz

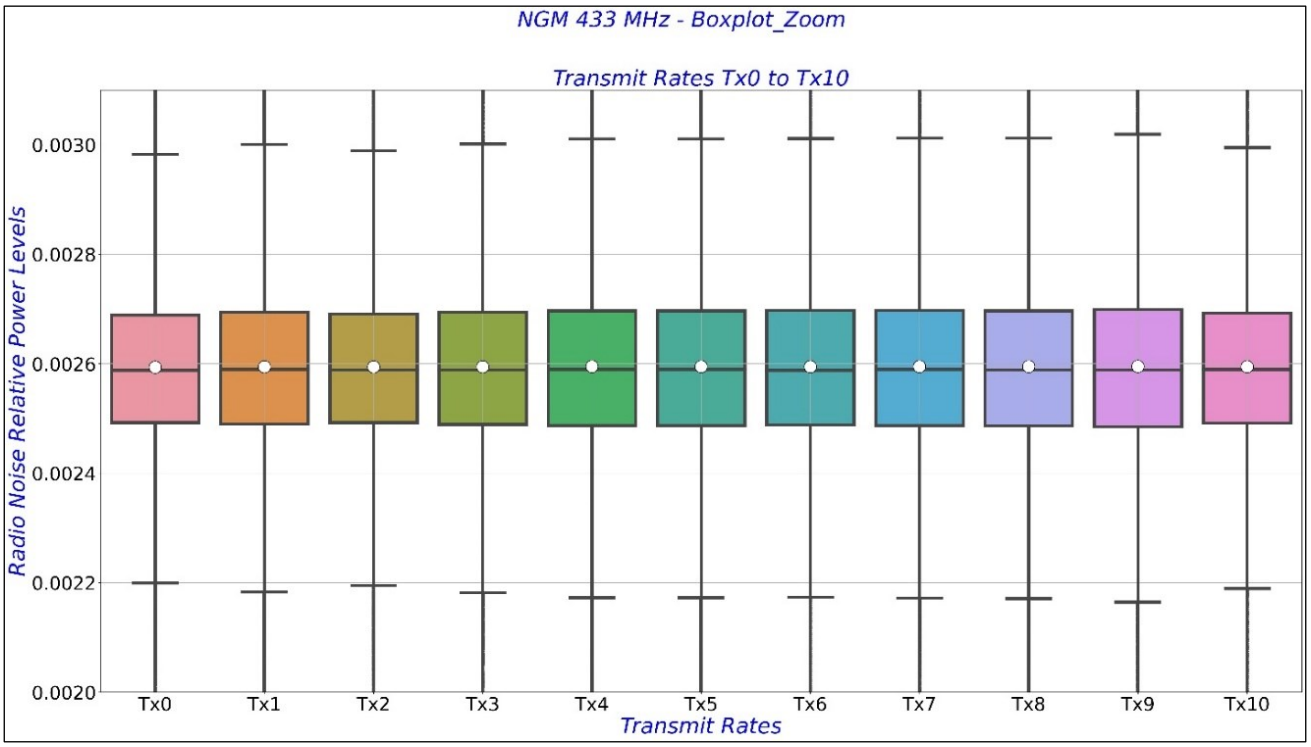

Figure 5.40: Boxplots for IoT Radio Noise at NGM - 433 MHz – no outliers

5.3.14. NGM 868 MHz

Table 5.14 shows the statistical metrics for the NGM 868 MHz case. The distributions expect no dispersion due to insignificant values for standard deviations and variances of approximately zero mW. Most of the samples were, therefore, confined around the mean. From Tx0 to Tx10, the

skewness values were unlikely to cause skewed distributions, as the highest value was 14.22 mW at Tx5. Except for Tx0, the weakest kurtosis was 56.08 mW; the rest of the datasets exhibited higher kurtosis values above 142 mW, with the highest value of 856.25 mW at Tx5. These values may influence the peakedness or the smoothness of the distributions. Data in this series did not carry many outliers as the rates varied between 2.71 % (minimum) at Tx0 and 9.7% (maximum) at Tx7, Tx8, Tx9, and Tx10. The median and mean showed comparable values of approximately zero mW. With minor standard deviations and variances, data showed characteristics of normal distributions, provided that further processing removed the outliers. The small values of IQR confirmed that the distributions see the majority of data confined within a small range. Figure 5.41 corroborates the observations with unsmooth and peaked density distributions. In Figure 5.42, the outliers' influence obstructs the boxplot visibility. Figure 5.43 achieves an improved resolution of visualisations after adjusting the y-axis. This result confirms that it is necessary to remove the outliers. Datasets in this series exhibit comparable IQR range.

Table 5.14: Statistical characteristics of IoT Radio Noise at NGM - 868 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
NGM	868 MHz	Mean	5.891150e-01	5.892470e-01	5.893070e-01	5.893450e-01	5.893760e-01	5.893270e-01	5.895510e-01	5.892540e-01	5.892540e-01	5.892540e-01	5.892540e-01
		Std	2.647000e-03	1.783500e-02	2.136500e-02	2.312600e-02	2.487900e-02	2.250900e-02	3.249500e-02	1.790000e-02	1.790000e-02	1.790000e-02	1.790000e-02
		Var	7.000000e-06	3.180000e-04	4.560000e-04	5.350000e-04	6.190000e-04	5.070000e-04	1.056000e-03	3.200000e-04	3.200000e-04	3.200000e-04	3.200000e-04
		Kurt	5.608361e+01	3.554966e+02	2.657154e+02	2.741077e+02	2.224324e+02	8.562556e+02	1.727174e+02	1.420760e+02	1.420760e+02	1.420760e+02	1.420760e+02
		Skew	-0 .109029	7.155354e+00	6.436797e+00	7.073440e+00	6.198063e+00	1.422149e+01	5.956075e+00	3.099585e+00	3.099585e+00	3.099585e+00	3.099585e+00
		Min Score	5.838200e-01	5.835040e-01	5.834660e-01	5.833150e-01	5.831640e-01	5.833500e-01	5.831580e-01	5.831230e-01	5.831230e-01	5.831230e-01	5.831230e-01
		Q1	5.877520e-01	5.876690e-01	5.876880e-01	5.876450e-01	5.876060e-01	5.876730e-01	5.876020e-01	5.876010e-01	5.876010e-01	5.876010e-01	5.876010e-01
		Median	5.888950e-01	5.889440e-01	5.889520e-01	5.889080e-01	5.889150e-01	5.889570e-01	5.889230e-01	5.888890e-01	5.888890e-01	5.888890e-01	5.888890e-01
		Q3	5.903740e-01	5.904460e-01	5.905030e-01	5.905320e-01	5.905680e-01	5.905550e-01	5.905640e-01	5.905870e-01	5.905870e-01	5.905870e-01	5.905870e-01
		IQR	2.622000e-03	2.777000e-03	2.815000e-03	2.887000e-03	2.962000e-03	2.882000e-03	2.963000e-03	2.986000e-03	2.986000e-03	2.986000e-03	2.986000e-03
		Max Score	5.943060e-01	5.946110e-01	5.947260e-01	5.948610e-01	5.950110e-01	5.948780e-01	5.950080e-01	5.950660e-01	5.950660e-01	5.950660e-01	5.950660e-01
		Outlier samples	271295 .00	5.953000e+05	6.490920e+05	720026 .0	8.040430e+05	8.782650e+05	8.592320e+05	9.766230e+05	9.766230e+05	9.766230e+05	9.766230e+05
		Outlier rates (%)	2.710000e+00	5.950000e+00	6.490000e+00	7.200000e+00	8.040000e+00	8.780000e+00	8.590000e+00	9.770000e+00	9.770000e+00	9.770000e+00	9.770000e+00

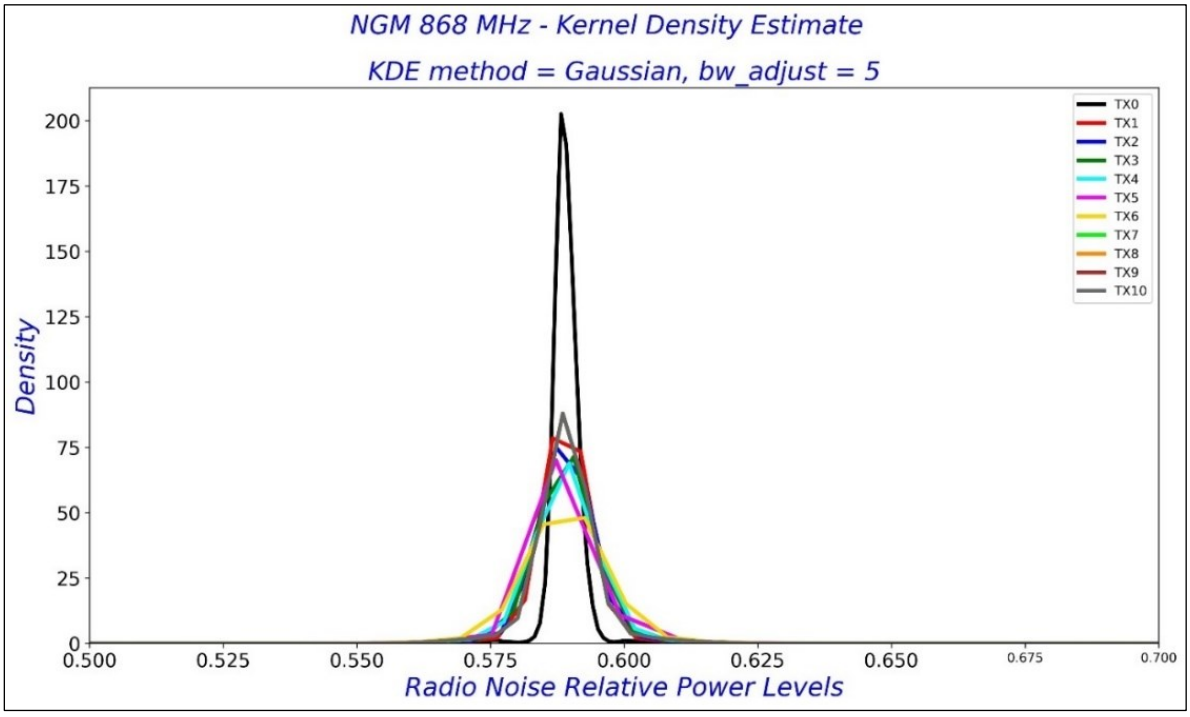

Figure 5.41: Density Distributions of IoT Radio Noise at NGM - 868 MHz

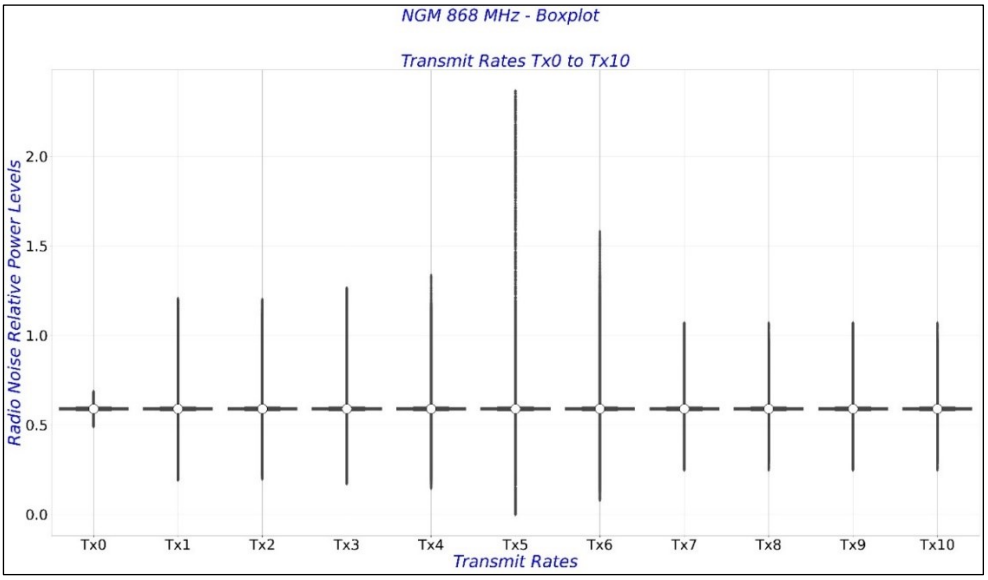

Figure 5.42: Full boxplots for IoT Radio Noise at NGM - 868 MHz

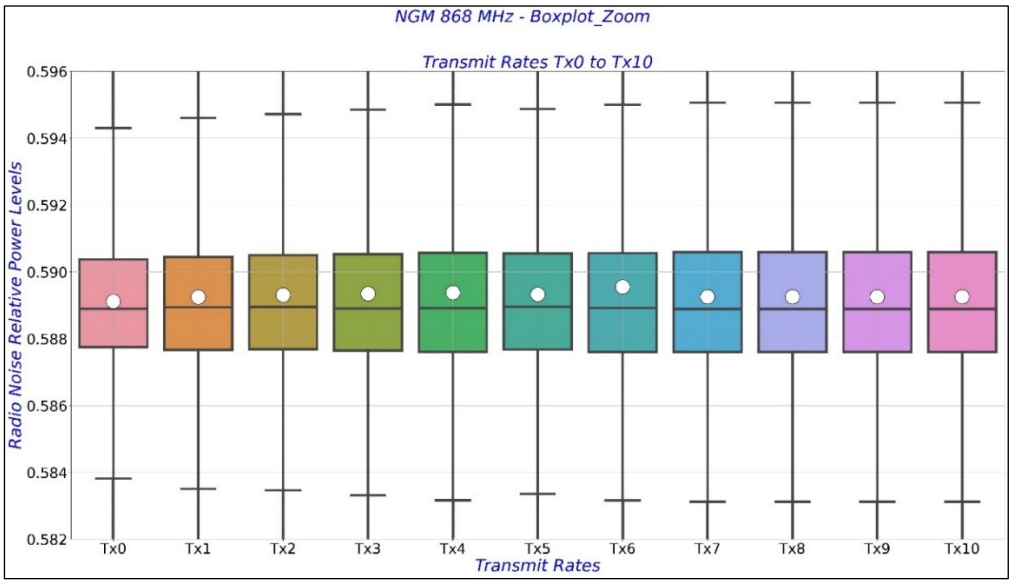

Figure 5.43: Boxplots for IoT Radio Noise at NGM - 868 MHz – no outliers

5.3.15. NGM 2400 MHz

Table 5.15 shows statistics for the NGM 2400 MHz case. Comparable mean and median values, and both approximately zero mW, informed of the possibility of normal distributions. Weak standard deviations and variances suggested that most samples would fit around the mean. These data series also carry varied kurtosis values with a minimum of 74.55 mW at Tx7 and a maximum of 1523.22 mW at Tx2. These values are high enough to influence the peakedness and smoothness of distributions. Tx2 produced the highest skewness factor of 15.49 mW, which should not lead to skewed distributions. Tx10 produced 3.17% as the maximum outlier rate, which means the presence of the outliers was weak. Figure 5.44 corroborates these observations as the distributions show unsmooth and peaked responses. Figure 5.45 shows the influence of the outliers with poor

visualisation for the boxplots, preventing them from obtaining meaningful insights from these responses. Figure 5.46 simulates the extraction of the outliers by adjusting the ordinate axis. As a result, the resolution of visualisation improved with clear visibility of all elements of boxplots.

Table 5.15: Statistical characteristics of IoT Radio Noise at NGM - 2400 MHz (values in mW)

Site	Band	Stat.	Tx0	Tx1	Tx2	Tx3	Tx4	Tx5	Tx6	Tx7	Tx8	Tx9	Tx10
NGM	2400 MHz	Mean	1.866010e-01	1.865170e-01	1.865550e-01	1.865830e-01	1.866520e-01	1.865060e-01	1.864890e-01	1.864610e-01	1.864420e-01	1.864460e-01	1.865650e-01
		Std	1.988000e-03	2.765000e-03	3.771000e-03	2.649000e-03	2.621000e-03	3.597000e-03	3.746000e-03	2.360000e-03	2.808000e-03	2.922000e-03	2.614000e-03
		Var	4.000000e-06	8.000000e-06	1.400000e-05	7.000000e-06	7.000000e-06	1.300000e-05	1.400000e-05	6.000000e-06	8.000000e-06	9.000000e-06	7.000000e-06
		Kurt	8.018528e+01	293 .970901	1.523224e+03	2.997975e+02	3.118775e+02	3.881705e+02	3.618448e+02	7.485119e+01	1.857818e+02	2.714661e+02	1.945709e+02
		Skew	1.343330e+00	3.586582e+00	1.549393e+01	3.474021e+00	3.314868e+00	5.364937e+00	5.241255e+00	1.322751e+00	2.556718e+00	3.399649e+00	2.456446e+00
		Min Score	1.815500e-01	1.819510e-01	1.813580e-01	1.820200e-01	1.820640e-01	1.819510e-01	1.817860e-01	1.796600e-01	1.814490e-01	1.812950e-01	1.817040e-01
		Q1	1.852340e-01	1.853280e-01	1.852340e-01	1.853970e-01	1.854720e-01	1.853280e-01	1.853280e-01	1.847240e-01	1.851320e-01	1.851320e-01	1.852340e-01
		Median	1.864600e-01	1.864520e-01	1.864090e-01	1.865210e-01	1.865960e-01	1.864520e-01	1.864520e-01	1.864600e-01	1.863580e-01	1.863580e-01	1.863580e-01
		Q3	1.876900e-01	1.875790e-01	1.878180e-01	1.876490e-01	1.877440e-01	1.875790e-01	1.876900e-01	1.881010e-01	1.875870e-01	1.876900e-01	1.875870e-01
		IQR	2.456000e-03	2.251000e-03	2.584000e-03	2.252000e-03	2.272000e-03	2.251000e-03	2.362000e-03	3.376000e-03	2.455000e-03	2.558000e-03	2.353000e-03
		Max Score	1.913740e-01	1.909560e-01	1.916950e-01	1.910260e-01	1.911520e-01	1.909560e-01	1.912330e-01	1.931660e-01	1.912700e-01	1.915270e-01	1.911170e-01
		Outlier samples	1.251780e+05	244036 .00	1.476510e+05	2.377630e+05	295790 .00	2.855290e+05	2.993510e+05	66209 .00	2.477080e+05	1.710950e+05	3.169940e+05
		Outlier rates (%)	1.250000e+00	2.440000e+00	1.480000e+00	2.380000e+00	2.960000e+00	2.860000e+00	2.990000e+00	6.600000e-01	2.480000e+00	1.710000e+00	3.170000e+00

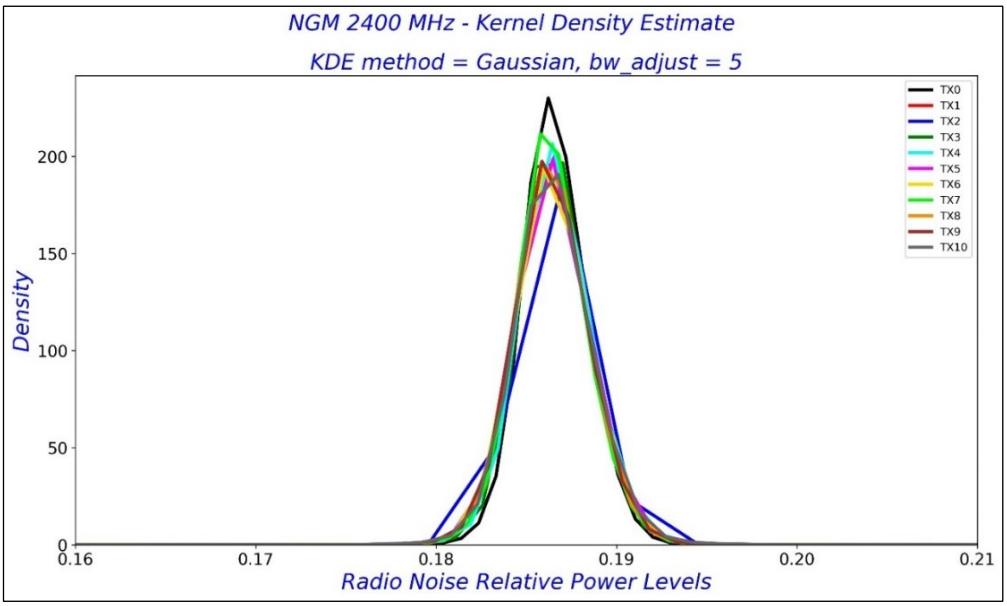

Figure 5.44: Density Distributions of IoT Radio Noise at NGM - 2400 MHz

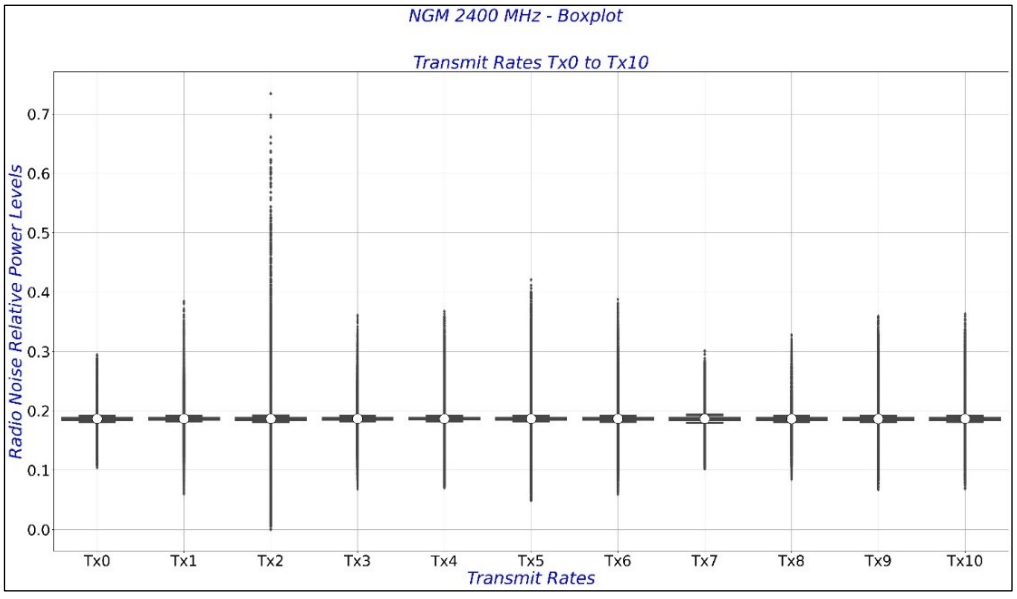

Figure 5.45: Full boxplots for IoT Radio Noise at NGM - 2400 MHz

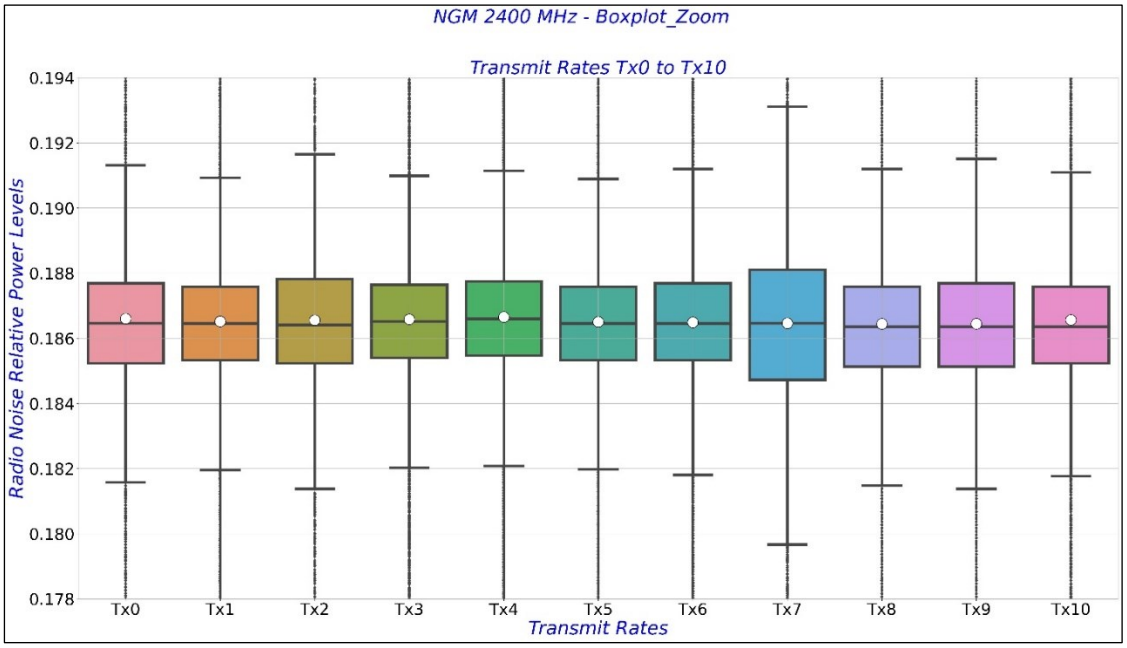

Figure 5.46: Boxplots for IoT Radio Noise at NGM - 2400 MHz – no outliers

5.4. Removing the Outliers

Statistical analysis in the previous section revealed potential outliers in the datasets. The outlier elimination proceeded by rejecting data samples below the lower and those above the upper fences of boxplots.

Figure 5.47 visualises boxplots in a subplot format for data containing the outliers from all fifteen cases (five sites and three frequency bands). As a common observation, every dataset carries several outliers, which are detrimental to boxplot analysis. With the removal of outliers, Figure 5.48 shows interpretable boxplots. The results exhibit comparable IQR ranges in each dataset from Tx0 to Tx9, except for MNC 868 MHz and MNS 868 MHz, where their Tx10 deviates and presents wider IQR compared to datasets in the same cases.

Figure 5.49 combines the density plots in a subplot format for data with the outliers removed. All data show symmetrical and smooth responses. Most cases developed normal distribution with a bell shape. A few datasets show noisy ripples of negligible impacts. MNC 868 MHz and MNS 868 MHz developed peaked distributions on datasets Tx0 to Tx9, reflecting their very narrow IQR ranges; their Tx10 data exhibit no normal distributions.

It was necessary to interrogate how outlier removal would affect the proportional relationship within the datasets. This analysis exploited scatter and line plots to trace the mean values. Figures 5.50 and 5.51, for data with and without the outliers, respectively, obtained comparable trace evolutions before and after removing the outliers. The average values should change, modifying the total number of samples due to removing the outliers. The results of the outlier removal show no

significant difference in proportions within the data. With that, there were no risks to proceed with this analysis using the outlier-less data.

5.5. Regression Lines

A critical question in this thesis was to demonstrate the influence of IoT radio activities on ambient radio noise levels. The analysis applied the regression technique to visualise the direction of trends as the transmit rates increased from the radio silence state (Tx0) to the busiest radio activities (Tx10).

As discussed in [106], Figure 5.52 is a subplot representing regression lines for all fifteen cases. The slopes of the regression lines indicate the direction of trends. Out of 15 assessments, only 3 cases (20 %) showed a decreasing trend. Twelve cases (80 %) indicated an augmenting trend. The levels of radio noise power as linear quantities show mere differences in the third, fourth or fifth decimal places. However, these variations may, though small linearly, have significant merit in the context of environmental radio noise. Table 5.16 summarises the observations.

Table 5.16: Conclusion of analysis from the regression lines

Site	Band (MHz)	Mean min (mW)	Mean max (mW)	Direction of trend
DUT	433	10.529603	10.541081	Augmenting
	868	1003.463161	1003.604141	Augmenting
	2400	38.965016	38.995504	Augmenting
GLM	433	2.845122	2.844545	Diminishing
	868	82.883005	82.922968	Augmenting
	2400	53.710115	53.740330	Augmenting
MNC	433	2.804772	2.807212	Augmenting
	868	1031.475526	1031.693036	Augmenting
	2400	53.710115	53.740330	Augmenting
MNS	433	2.846615	2.924101	Augmenting
	868	325.305783	325.350955	Augmenting
	2400	70.211370	70.200025	Diminishing
NGM	433	2.591371	2.592906	Augmenting
	868	589.076361	589.092486	Augmenting
	2400	186.602416	186.409612	Diminishing

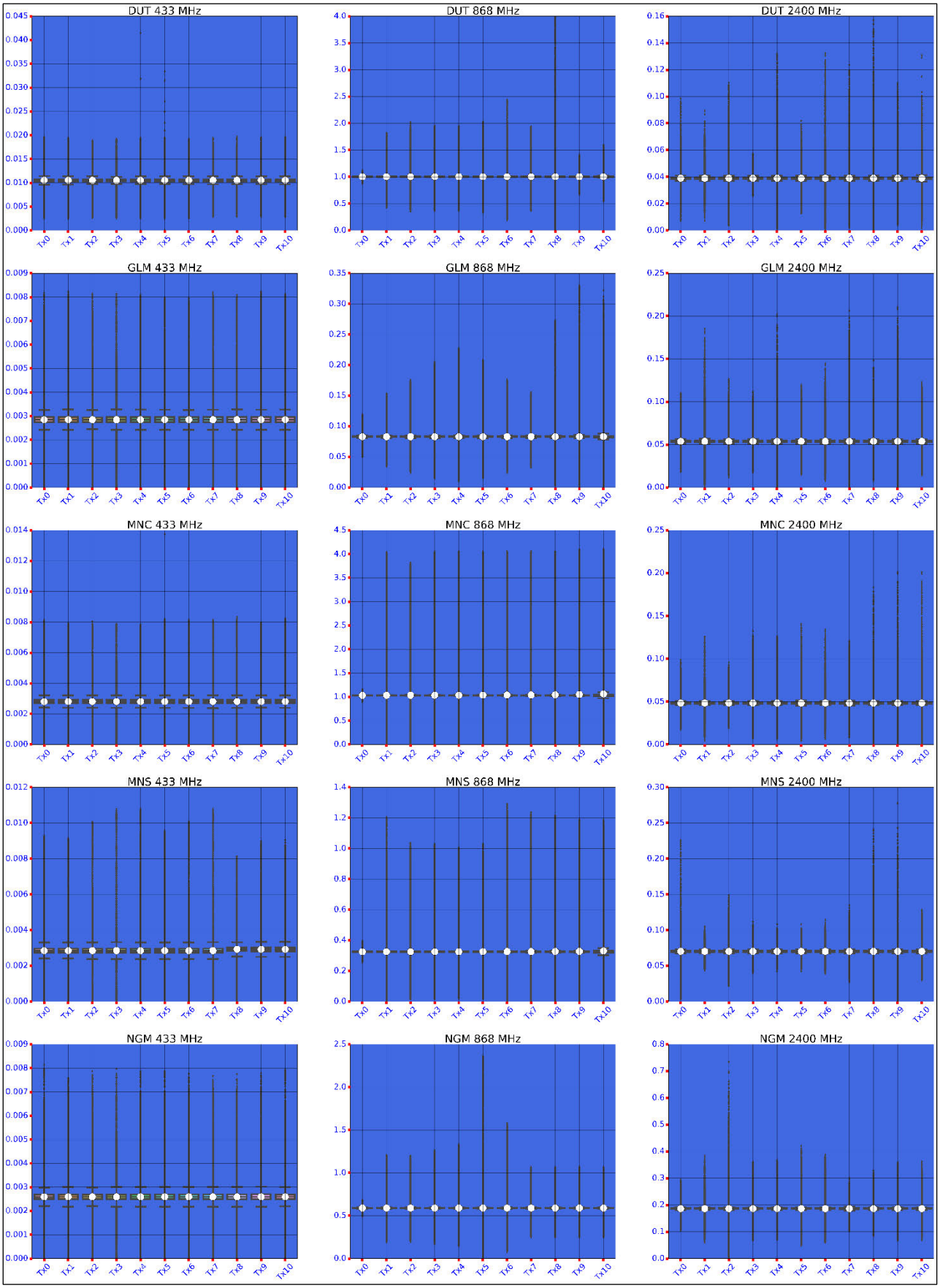

Figure 5.47: Data with the outliers

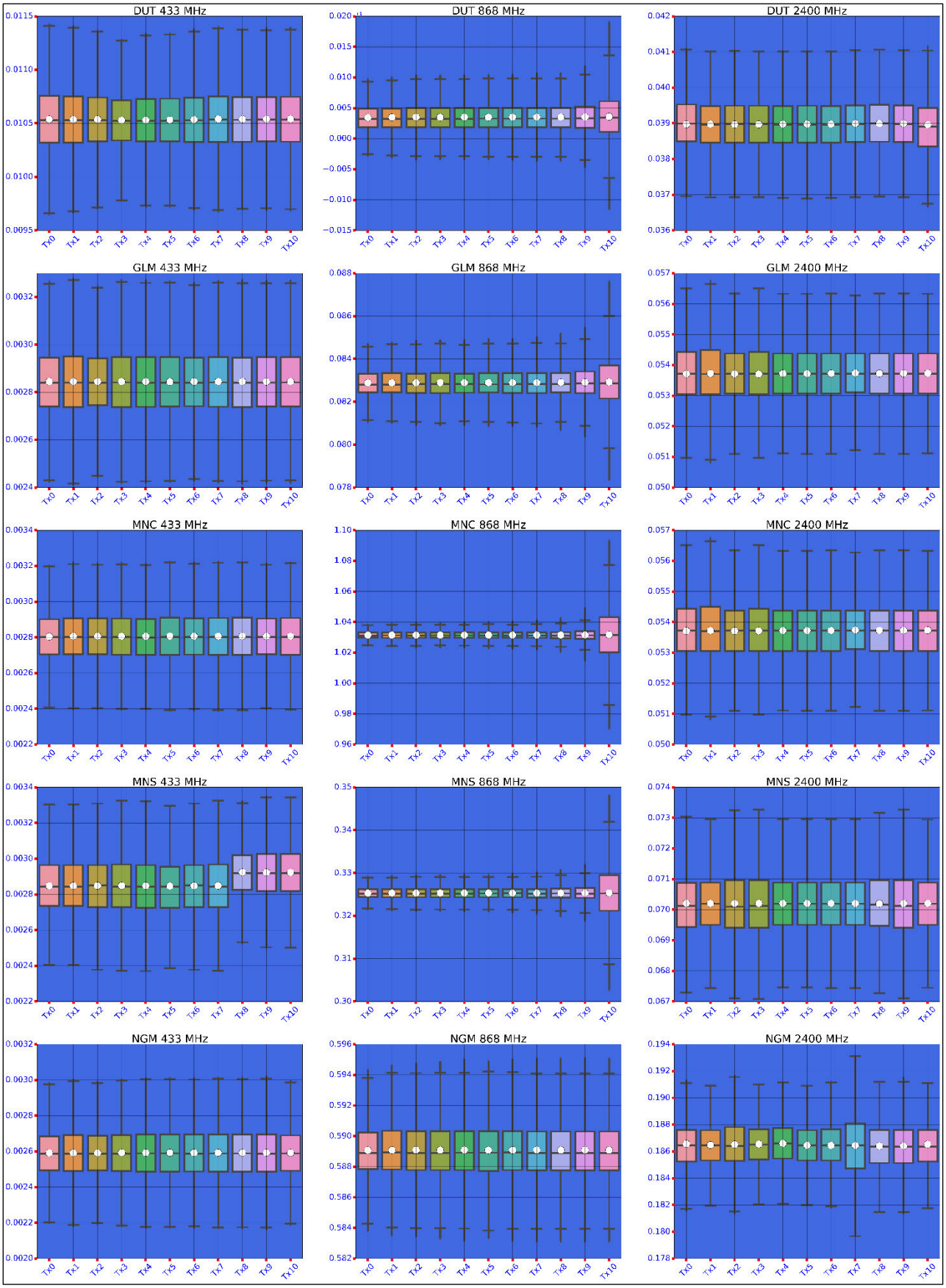

Figure 5.48: Data with the outliers removed.

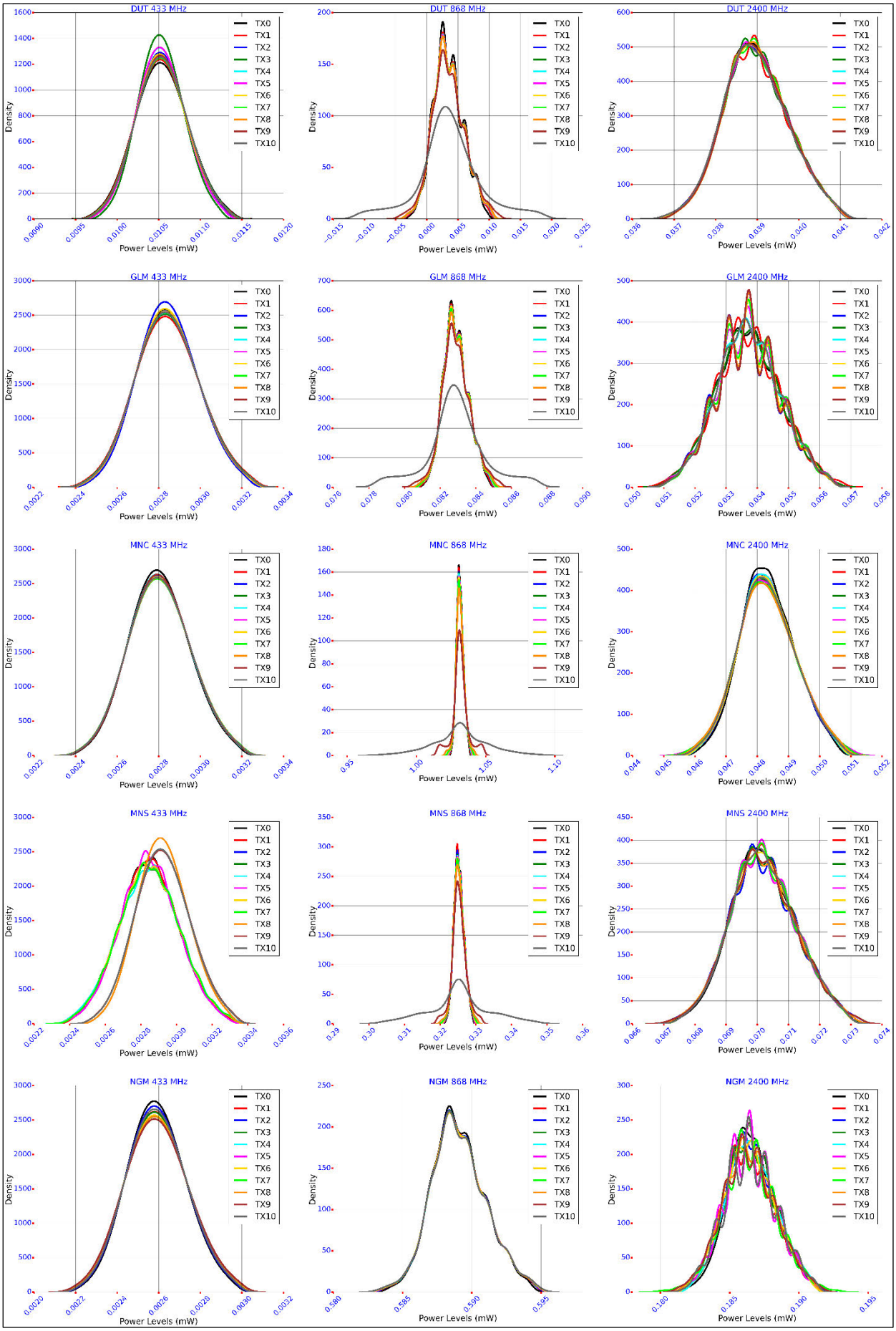

Figure 5.49: Density plots for data with the outliers removed.

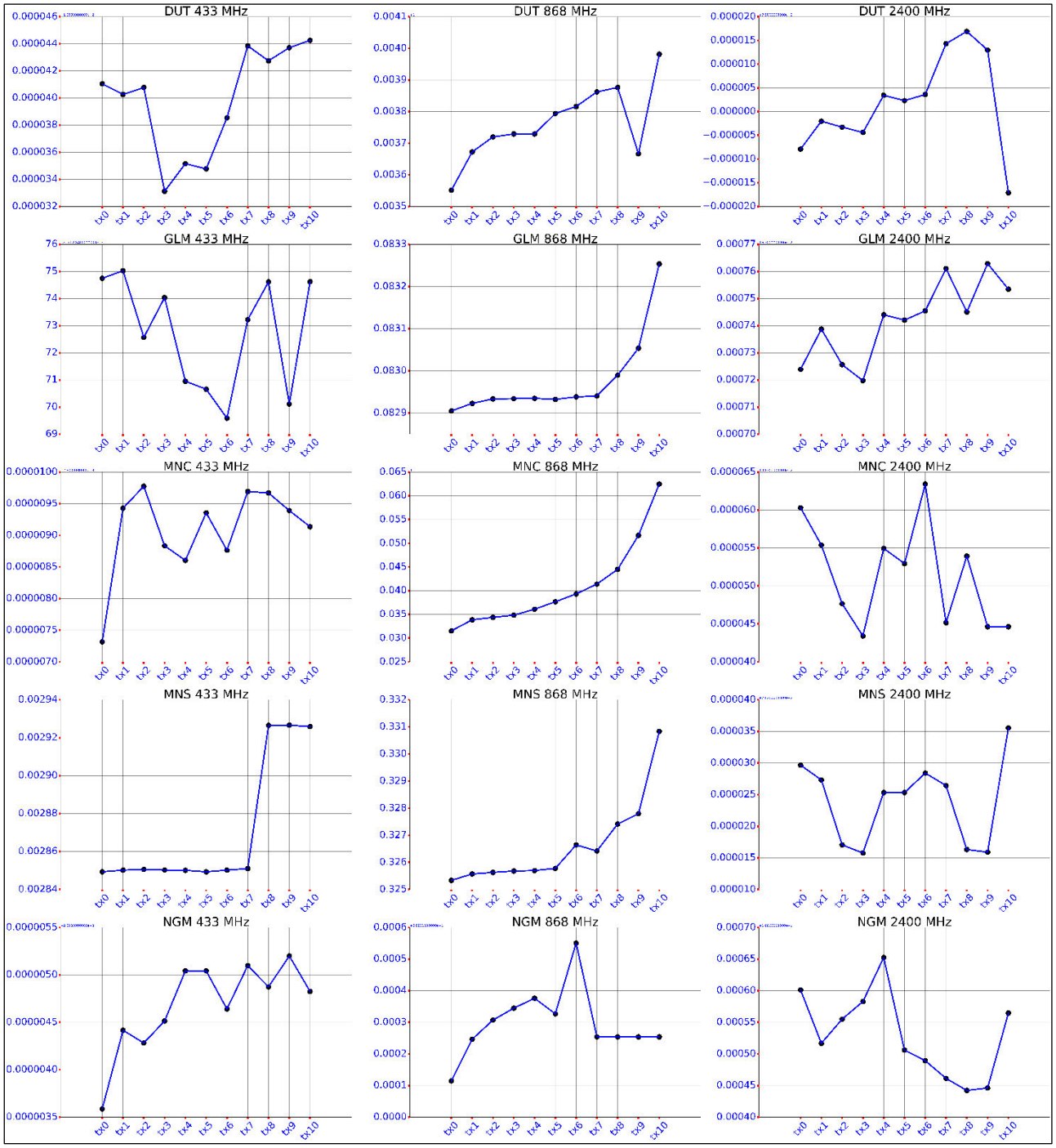

Figure 5.50: Average values with the outliers

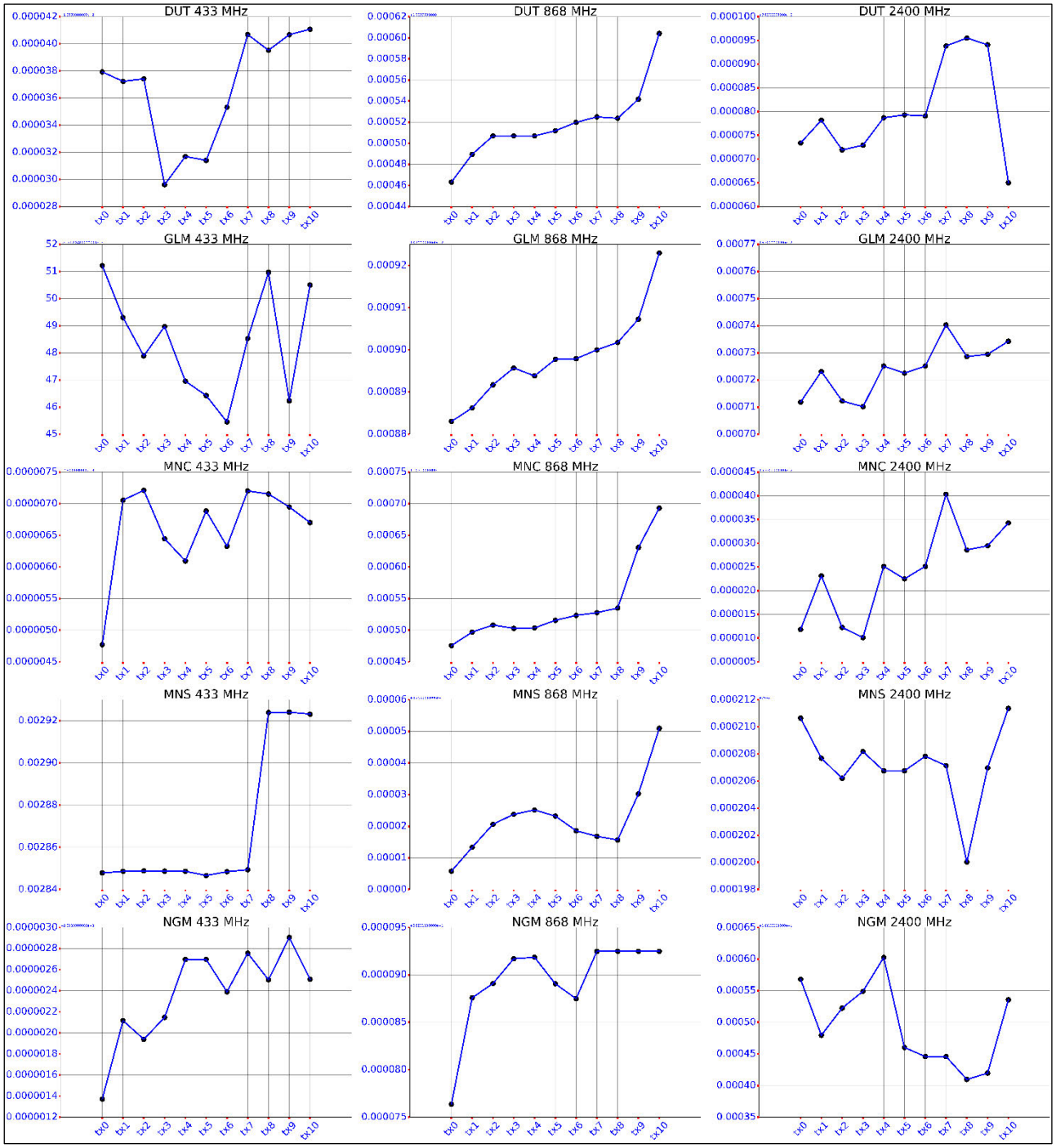

Figure 5.51: Average values with the outliers removed.

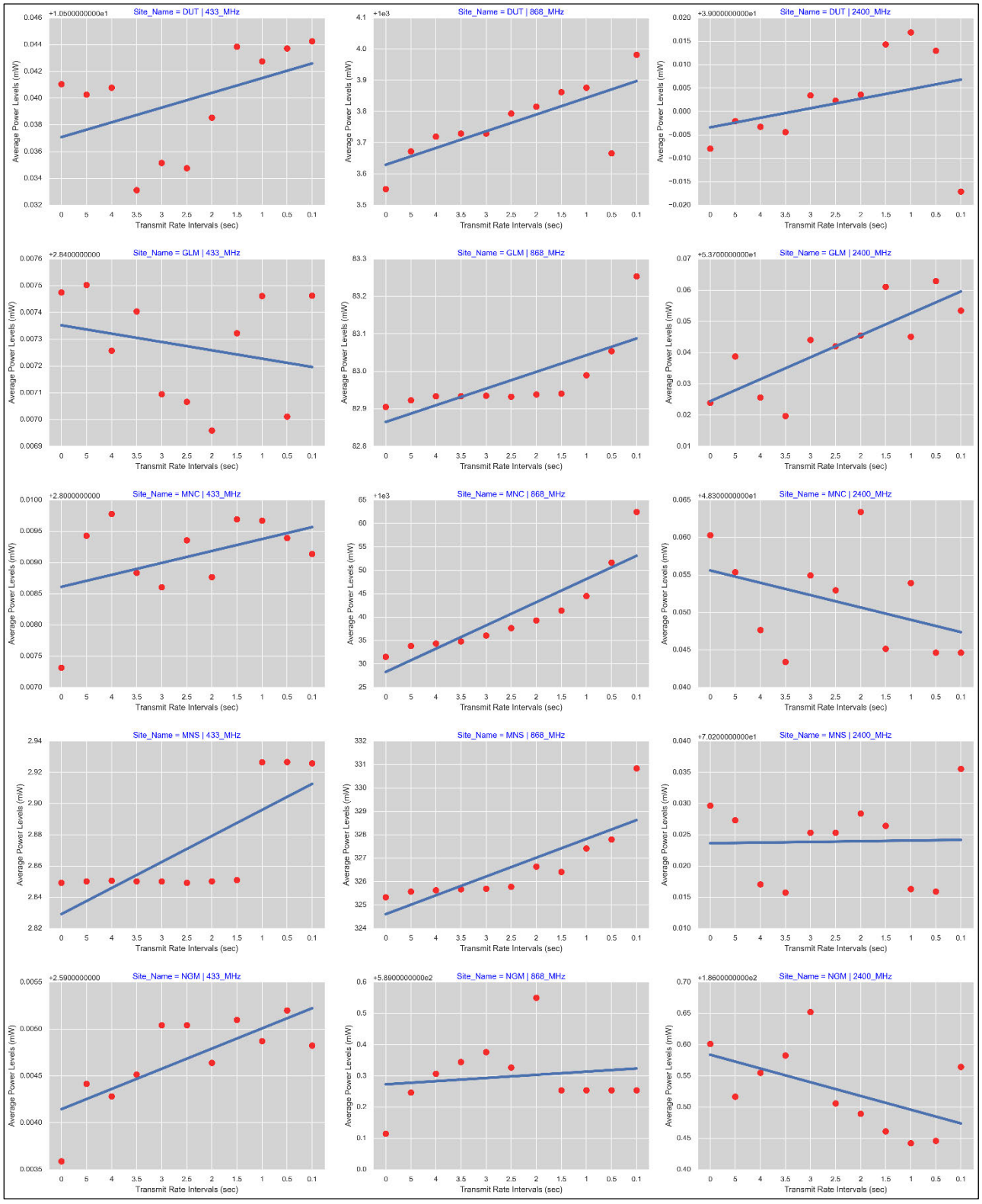

Figure 5.52: Regression lines and scatter plots

5.6. Empirical Analysis

As discussed in [106], data processing could identify the features or transmit rate columns containing the highest average value per case (site and frequency). This empirical and heuristic approach required to split these features into two categories:

- Lower Transmission Rate Side (LTRS): associating Transmit Rates Tx0 to Tx5
- Upper Transmission Rate Side (UTRS): associating Transmit Rates Tx6 to Tx10

Table 5.17 represents the transmit rates of maximum noise and counts how many belonged in the LTRS or UTRS categories. The analysis demonstrated that the UTRS category dominated and therefore confirmed the theory that the quantity of ambient radio noise is directly proportional to the volume of RF activities, even in this case of low-power IoT operations, and irrespective of frequency bands and type of environment (industrial, urban, or suburban).

Table 5.17: Conclusion of the empirical analysis.

Band	**DUT**	**GLM**	**MNC**	**MNS**	**NGM**	**LTRS**	**UTRS**
433 MHz	Tx10	Tx0	Tx2	Tx9	Tx9	2	3
868 MHz	Tx10	Tx10	Tx10	Tx10	Tx7	0	5
2400 MHz	Tx8	Tx7	Tx7	Tx10	Tx4	1	4

5.7. Summary

An essential contribution from this Chapter was observing the influence of IoT radio activities on environmental radio noise levels. This Chapter transformed the original raw data from amplitude in volts to relative power expressed in mW. The first analysis demonstrated the presence of outliers in every dataset. Removing the outliers contributed to improving the resolution for data visualisation. Data showed similarities in statistics, such as standard deviations and variances. All data cases showed approximately zero mW levels of standard deviations and variances.

The linear values for the radio noise levels are too small that the statistical tables and density distribution plots could not easily demonstrate the differences in power levels as the Transmit Rates increased from Tx0 to TX10. The differences appear in the third or fourth decimal place. However, regression lines and scatterplots resolved this difficulty as it was visibly easy to interpret the direction of trends.

This Chapter addressed the most critical question of this thesis, affirming that the 80% of data developed increase in radio noise levels directly proportional to the quantity of IoT radio activities.

CHAPTER 6: DEVELOPMENT OF RADIO NOISE PREDICTION MODELS

6.1. Introduction

Machine learning can reliably retrieve information from IoT data [81]. As a definition, *machine learning* refers to the automated detection of meaningful patterns in data. Nowadays, it is a popular tool required to extract insights from data. It englobes notions of probability, statistics, linear algebra, analysis, and algorithms. Machine Learning also signifies "automated learning", where devices train from data delivered [112].

There may be different approaches to solving Machine Learning problems. This thesis exploited the supervised learning method, which fits well with this research's data. The supervised learning maps an input to an output based on sample input-output pairs. It applies a labelled data training mechanism. Supervised learning is suitable for determining a relationship between the input and output variables. "Supervised" means the outcome or output variable guides the learning process. Supervised learning is appropriate for applications requiring predicting a particular outcome from a given input [113-115].

6.2. Machine Learning techniques

6.2.1. Machine Learning Algorithms

Figure 6.1 describes the methods or algorithms used in the supervised problems. Machine Learning expects quantitative or qualitative (categorical or discrete) variables. On prediction tasks, data types imply two algorithm methods for supervised learning: 1- regression and 2- classification. The regression method predicts continuous quantitative outputs. In linear regression, prediction is a line for a single feature, a plane for two features, or a hyperplane in higher dimensions. On the other end, predictions of qualitative outputs require classification methods. Classification intends to predict a choice predefined from a list of possibilities [114, 115].

Standard classification algorithms are Naive Bayes, Linear Discriminant, Logistic Regression, K-Nearest Neighbours, Support vector machine, Decision Tree, Random Forest, Adaptive Boosting, Extreme Gradient Boosting, Stochastic Gradient Descent, Rule-based Classification [113].

Popular regression algorithms are linear, polynomial, Lasso, and Ridge regressions. Simple and Multiple Linear Regression is the most popular Machine Learning modelling technique, well-known as regression techniques. In regression, the dependent variable or output is continuous, and the independent variable(s) or input can be continuous or discrete. Instead of a linear response, a polynomial regression produces an outcome where a relationship between the independent and the dependent variables is the polynomial of n^{th} degree. LASSO and Ridge regressions are powerful

regression techniques due to their capability to prevent over-fitting and reduce the complexity of the model. They are usually advantageous for building learning models with many features [113].

6.2.2. Machine Learning Mechanism

Literature interchanges the terms for the input variables with features, predictors, or independent variables. Likewise, the output variables are synonyms to responses, dependent variables, or targets. Machine Learning mechanisms explore the inputs to predict the output [114].

Modelling is an iterative process. Clean data and a good understanding of data contents set the prerequisite for Machine Learning modelling to make efficient predictions. The modelling process consists of the following three main steps: 1- selection of algorithm based on data type, 2- execution of the model, and 3- diagnosis and model comparison, classifying objects, or gaining an understanding of the system. Instead of coding models from scratch using traditional statistical approaches, importing pre-built libraries available in programming languages is advisable. For example, *StatsModels* or *Scikit-learn* are popular Machine Learning libraries in Python [116].

Supervised learning usually splits the labelled data into two parts. The first part is the training data used to build a model. The second part is the test data reserved to evaluate model performance. It serves to measure whether an algorithm will perform well on new data. A model generalises from the training set to the test set when it can make accurate predictions. As a rule of thumb, simple models generalise better to new data. Complex models are prone to overfitting, which happens when a model works well on the training set however, unable to generalise to new data. On the contrary, too simple models might not be able to capture all the aspects of variability in the data due to underfitting, which happens when a model cannot work well on training data [115].

6.2.3. Linear Regression

Different linear regression models exist based on how the algorithm trains the model parameters slope (also called weights or coefficients) and intercept (also called offset) from the training data and how the algorithm manages complexities in the model. Linear regression, or ordinary least squares (OLS), is the simplest and most classic linear method for regression. Linear regression finds the parameters slope and intercept that minimise the mean squared error between predictions and the true regression targets on the training set. The mean squared error is the sum of the squared differences between the predictions and the true values. The intercept or the offset attribute is always a single float number, while the slope or coefficient attribute is an array with one entry per input feature. Ridge regression, as a linear model for regression, uses ordinary least squares to make predictions. The Ridge model makes a trade-off between the simplicity of the model (near-zero coefficients) and its performance on the training set [115].

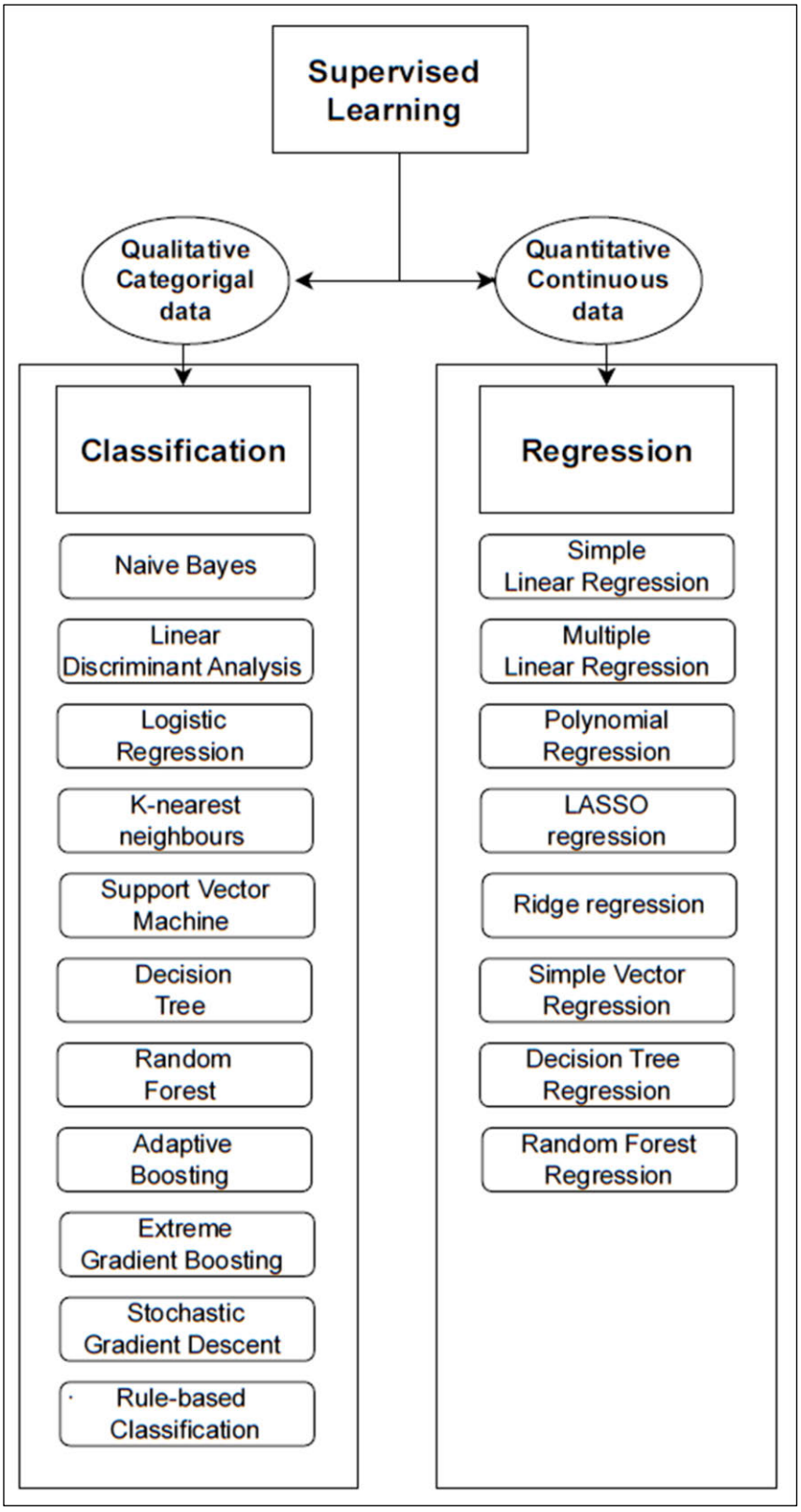

Figure 6.1: Algorithms for Supervised Learning

6.3. Data Interpolation

A dataset case, for example DUT 433 MHz, contained only eleven data points corresponding to transmit rates Tx0 to Tx10. Using machine learning on small-size datasets presents a problem because, in general, the power of machine learning in recognising patterns is proportional to the

size of the dataset; the smaller the dataset, the less powerful and less accurate the machine learning algorithms [117]. Standard training algorithms assume that datasets are large enough to train massive models successfully. In practice, and mainly when a dataset is not large enough, it is a frequent practice to artificially generate additional samples, widely known as "data augmentation". The artificial generation of training samples for machine learning has been a common practice for many years. The paper documented some sophisticated ways of dealing with small data, such as Subspace Approximation, Sparse Decomposition, Empirical Mode Decomposition, Tensor Decomposition, Class preserving transforms, Statistical imputation, Probabilistic modelling, Low-rank matrix completion [118].

This research applied data interpolation to generate artificial samples to fill the gaps alongside eleven known values. There are several general facilities available in SciPy for interpolation and smoothing for data in 1, 2, and higher dimensions. The interp1d, CubicSpline, and PchipInterpolator functions perform data interpolation. The choice of a specific interpolation routine depends on the data: whether it is one-dimensional, is given on a structured grid, or is unstructured [119]. This research implemented data interpolation using the CubicSpline to generate additional samples alongside the eleven known values. Experiments showed later that the total number of interpolated samples influenced the model accuracy. The individual respective plot shows the Data size indicating the number of interpolated samples.

Figure 6.2 through Figure 6.16 show the results of data augmentation for each case grouped in candidate site and band. The black trace represents all interpolated values. The red marks indicate eleven known values. The interpolation operation did not change the original values of the eleven known points. With that, the responses from the models with interpolated data still reflect that of the original data.

6.4. Predictive Modelling

The numerical and continuous data nature of the target or dependant or output variable dictated the application of supervised and linear regression methods. This research applied two linear regression algorithms for predictive modelling: OLS and Ridge. The univariate Train-Test learning applied the transmit rate (TX Rate) and noise levels (Level) data from the interpolated data. Experiments demonstrated that sample splitting size influenced the model performance. 67 % on the train and 33 % on the test subsets showed satisfactory accuracy trade-offs among all data cases. With datasets grouped per candidate site and frequency band, this modelling exercise produced 15 models for each OLS and the Ridge linear regressions.

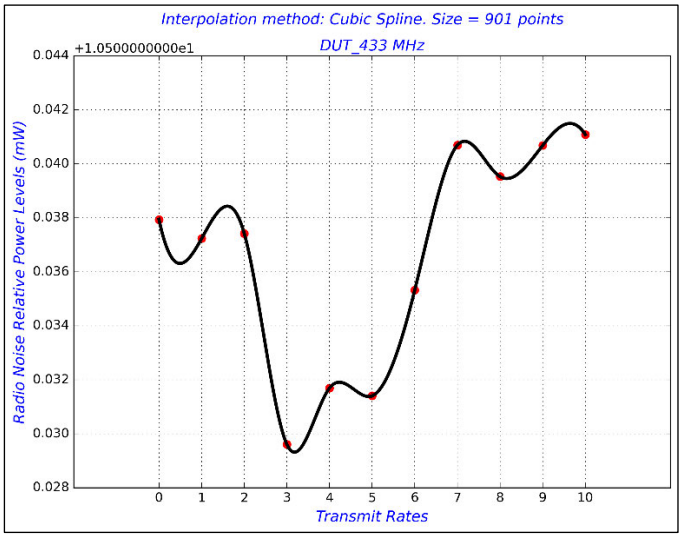

Figure 6.2: Interpolated data for DUT 433 MHz

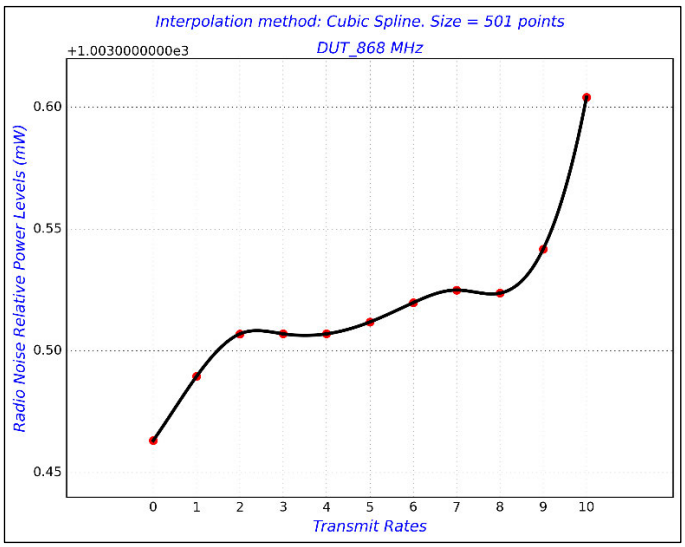

Figure 6.3: Interpolated data for DUT 868 MHz

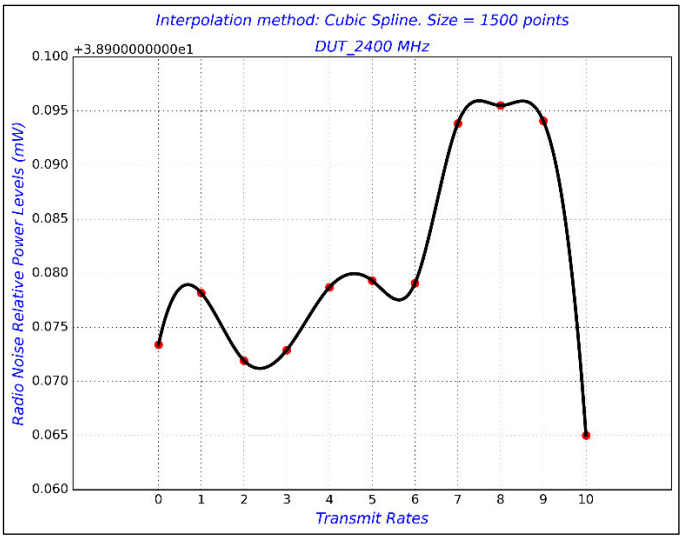

Figure 6.4: Interpolated data for DUT 2400 MHz

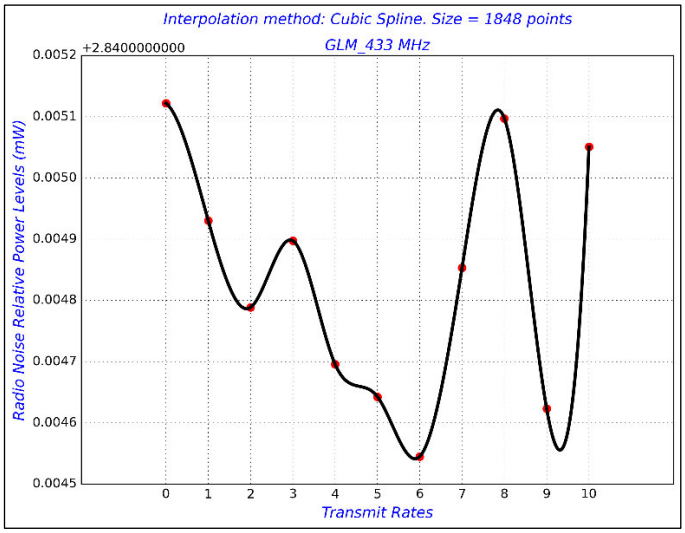

Figure 6.5: Interpolated data for GLM 433 MHz

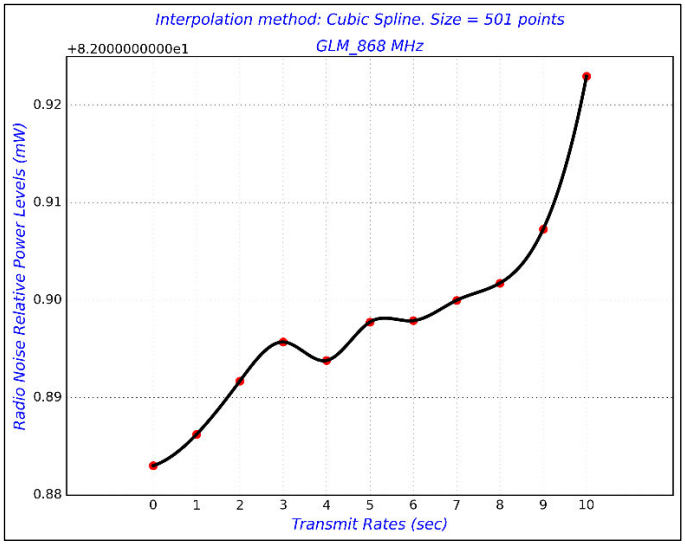

Figure 6.6: Interpolated data for GLM 868 MHz

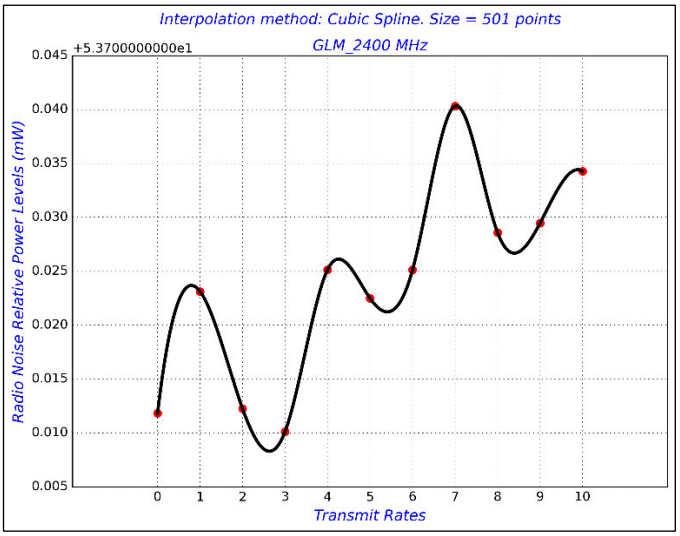

Figure 6.7: Interpolated data for GLM 2400 MHz

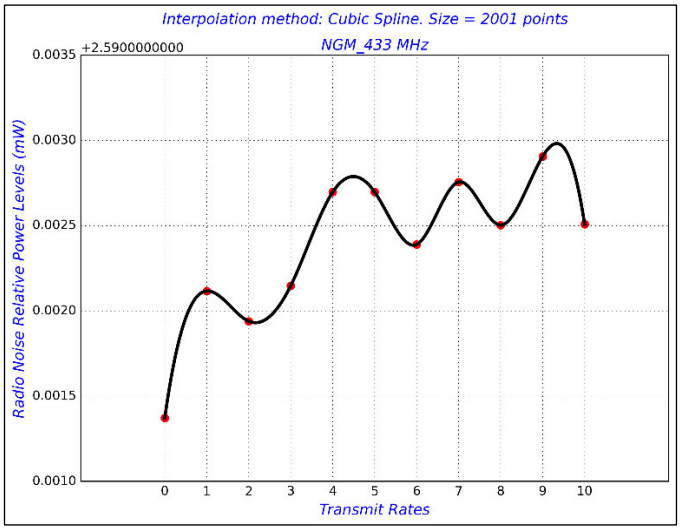

Figure 6.8: Interpolated data for MNC 433 MHz

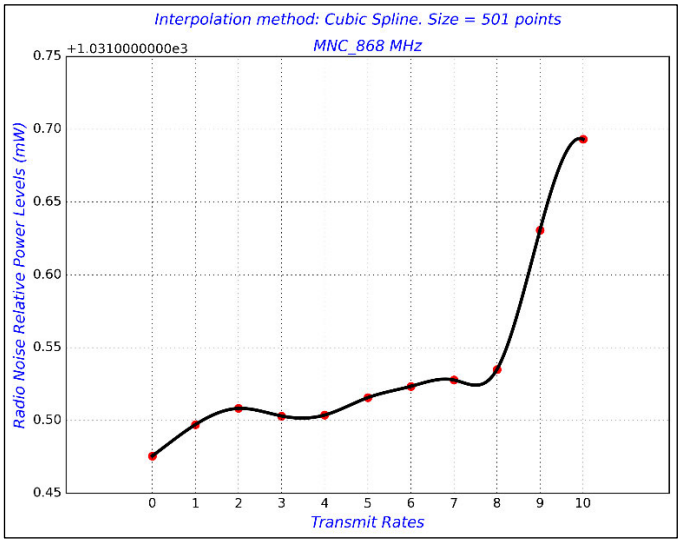

Figure 6.9: Interpolated data for MNC 868 MHz

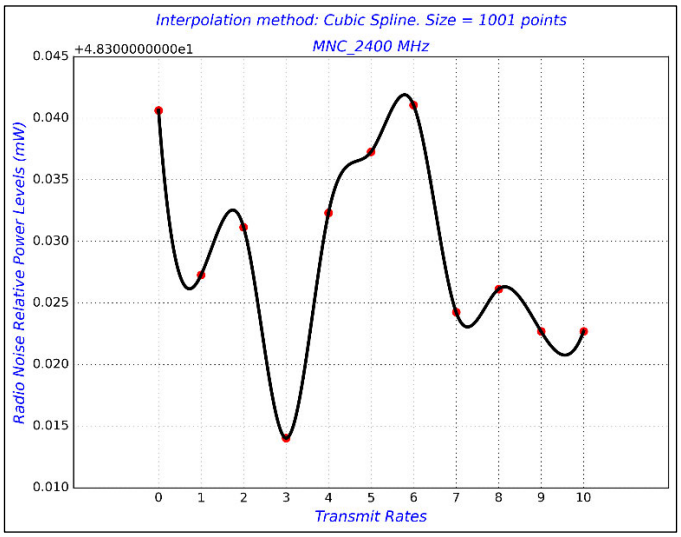

Figure 6.10: Interpolated data for MNC 2400 MHz

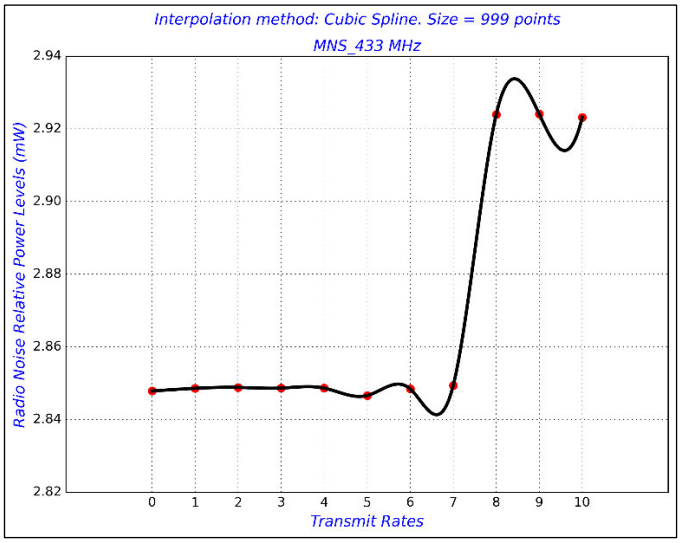

Figure 6.11: Interpolated data for MNS 433 MHz

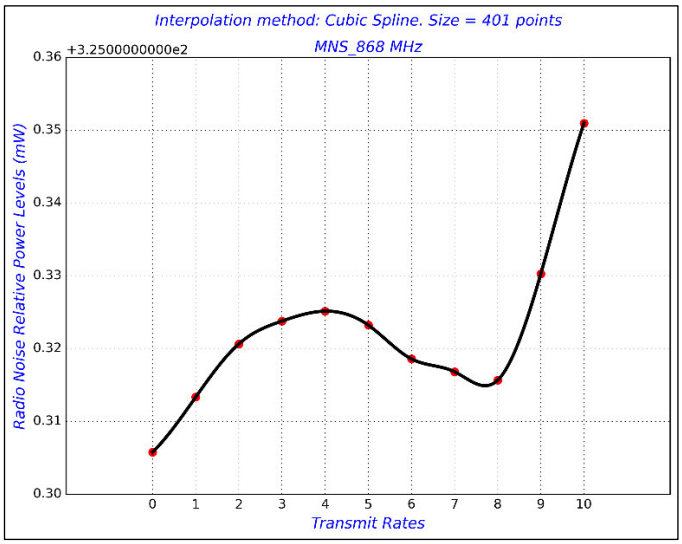

Figure 6.12: Interpolated data for MNS 868 MHz

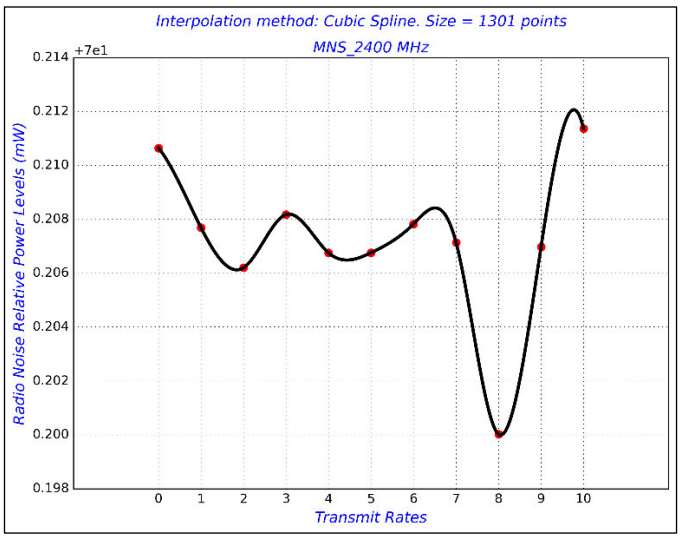

Figure 6.13: Interpolated data for MNS 2400 MHz

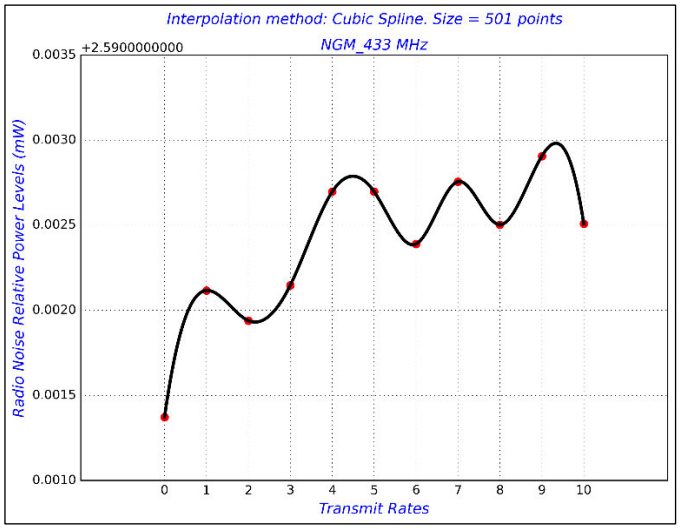

Figure 6.14: Interpolated data for NGM 433 MHz

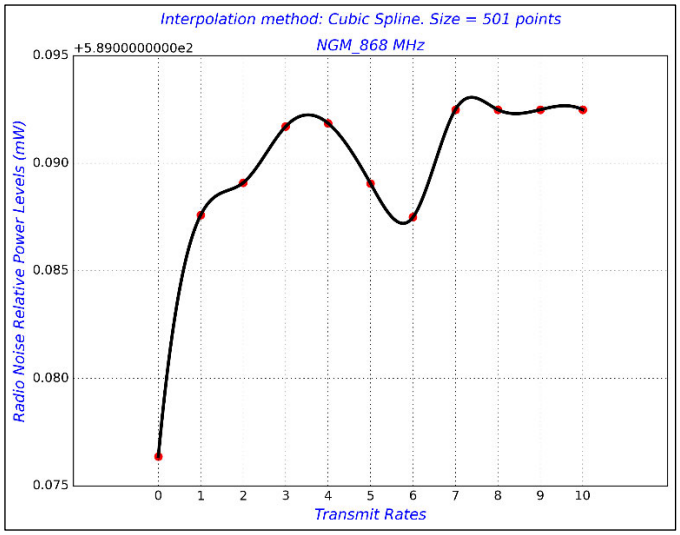

Figure 6.15: Interpolated data for NGM 868 MHz

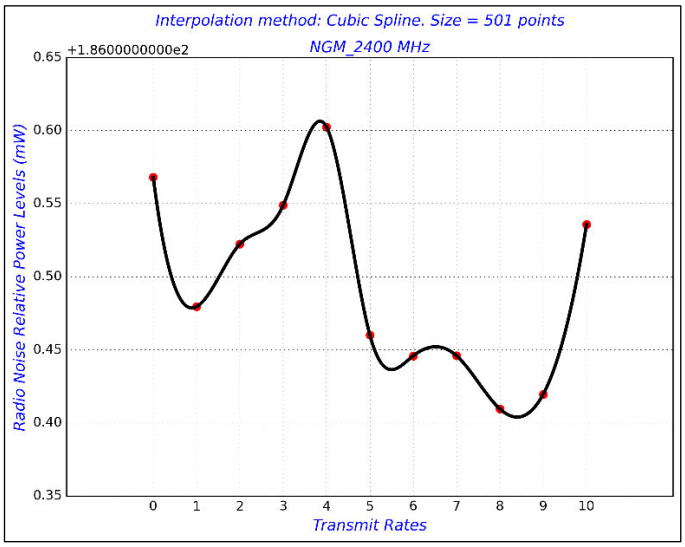

Figure 6.16: Interpolated data for NGM 2400 MHz

6.4.1. OLS models

Figure 6.17 through Figure 6.31 combined the train and test data for the OLS method and their corresponding regression lines. Analysis grouped the results per candidate sites and frequency bands. The figure legends show the proportion between the train and test data. The regression lines obtained with interpolated data produced the same slope and intercept values as those before the interpolation process. The train and test subsets followed the same trends because they originated from the same dataset.

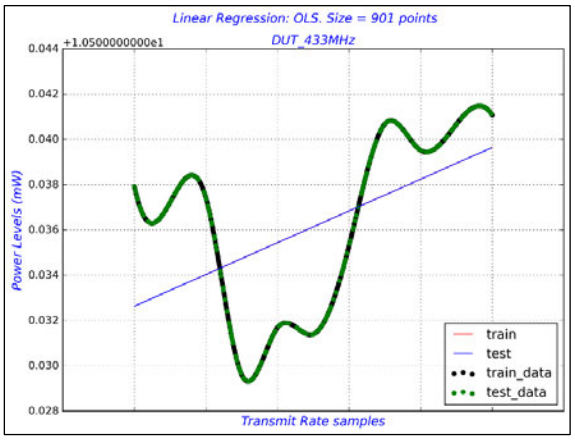

Figure 6.17: OLS regression for DUT 433 MHz

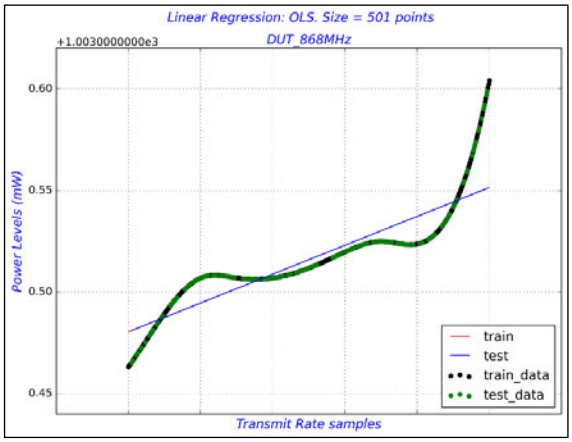

Figure 6.18: OLS regression for DUT 868 MHz

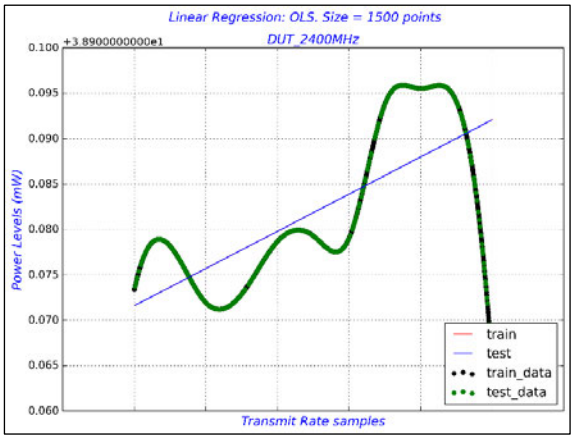

Figure 6.19: OLS regression for DUT 2400 MHz

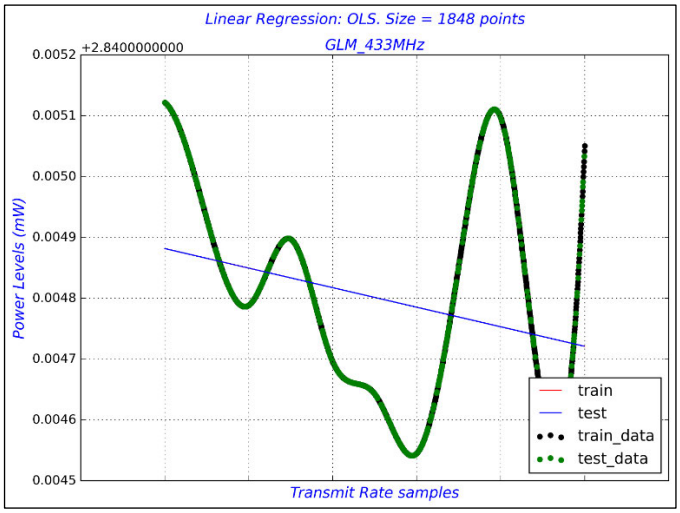

Figure 6.20: OLS regression for GLM 433 MHz

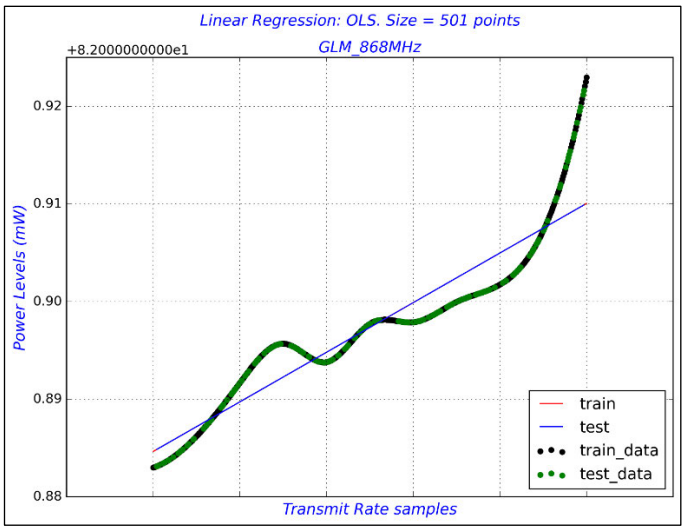

Figure 6.21: OLS regression for GLM 868 MHz

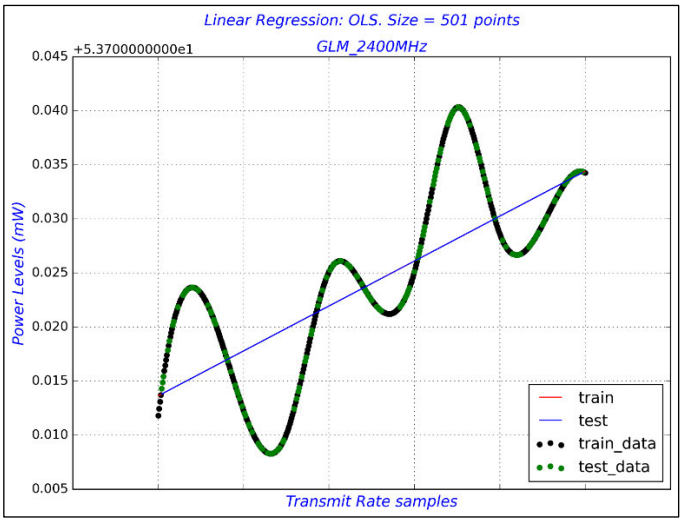

Figure 6.22: OLS regression for GLM 2400 MHz

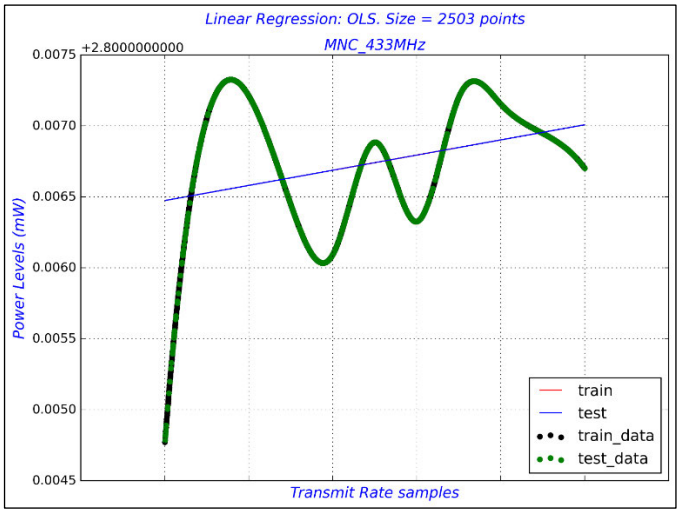

Figure 6.23: OLS regression for MNC 433 MHz

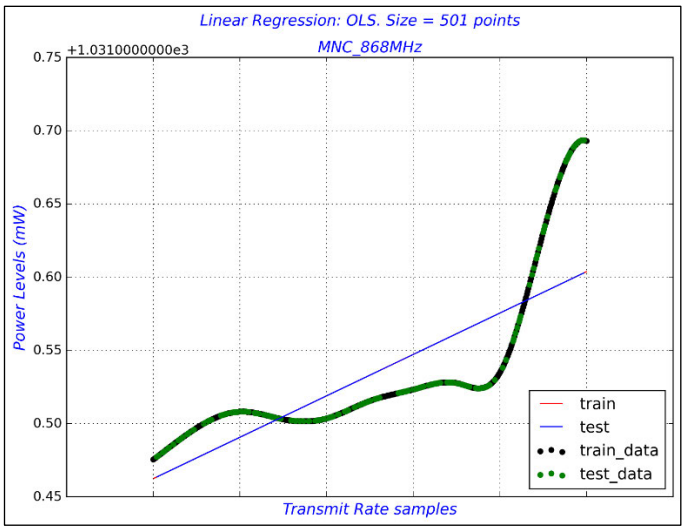

Figure 6.24: OLS regression for MNC 868 MHz

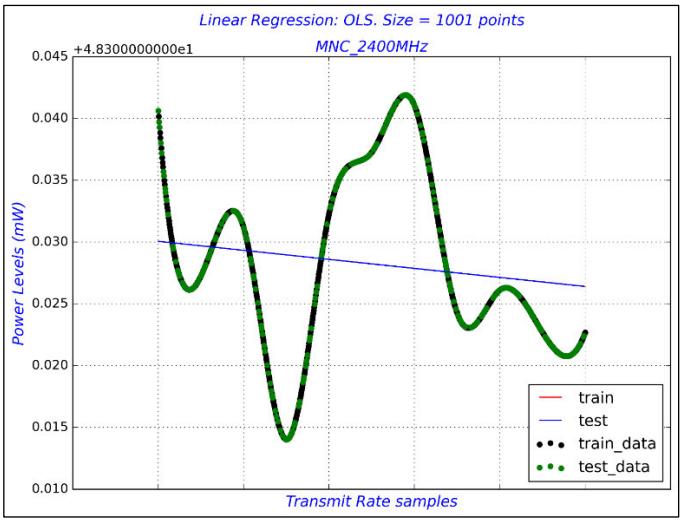

Figure 6.25: OLS regression for MNC 2400 MHz

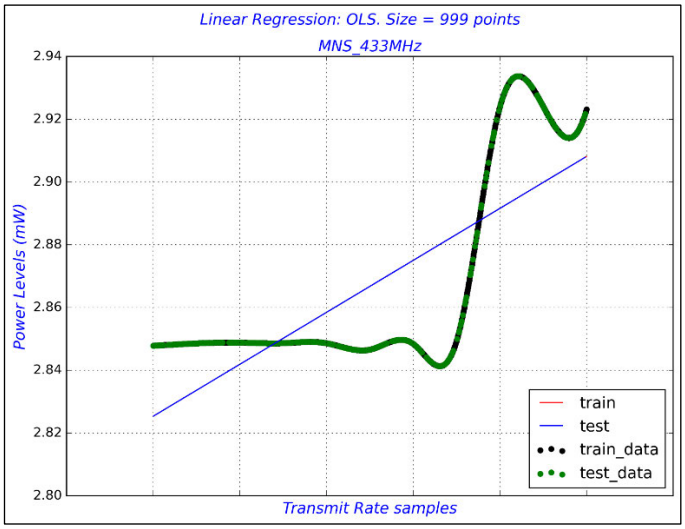

Figure 6.26: OLS regression for MNS 433 MHz

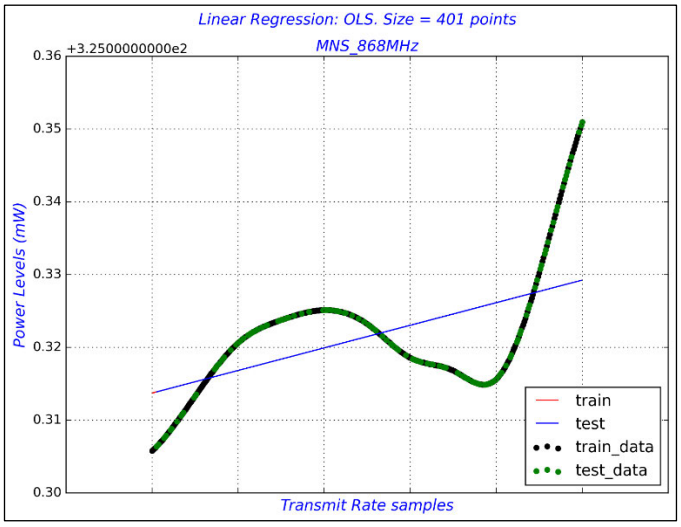

Figure 6.27: OLS regression for MNS 868 MHz

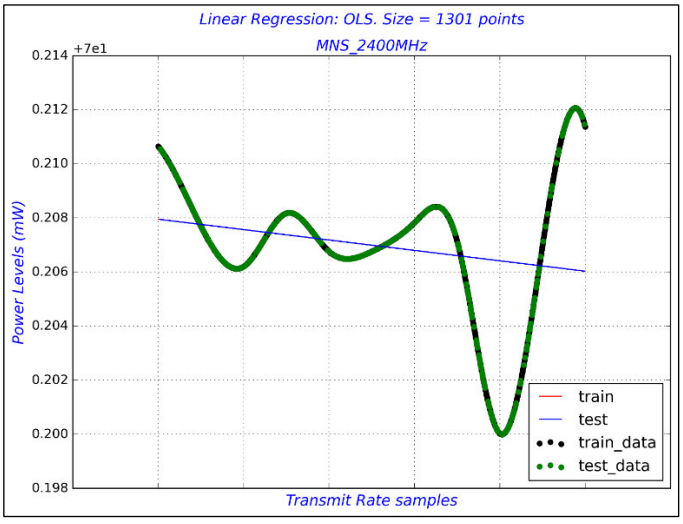

Figure 6.28: OLS regression for MNS 2400 MHz

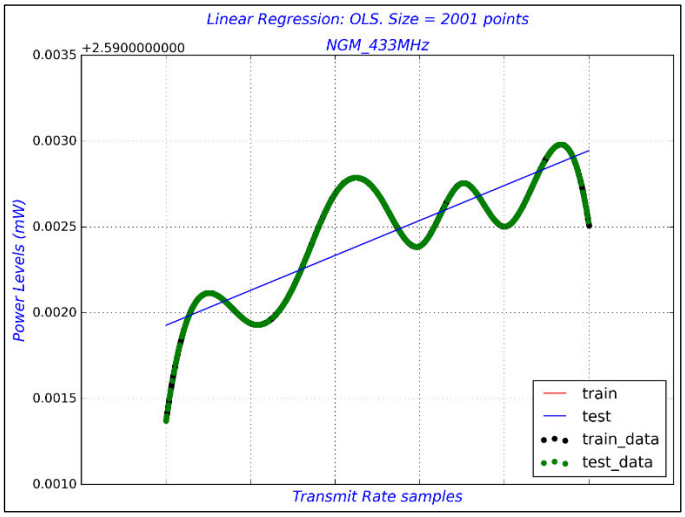

Figure 6.29: OLS regression for NGM 433 MHz

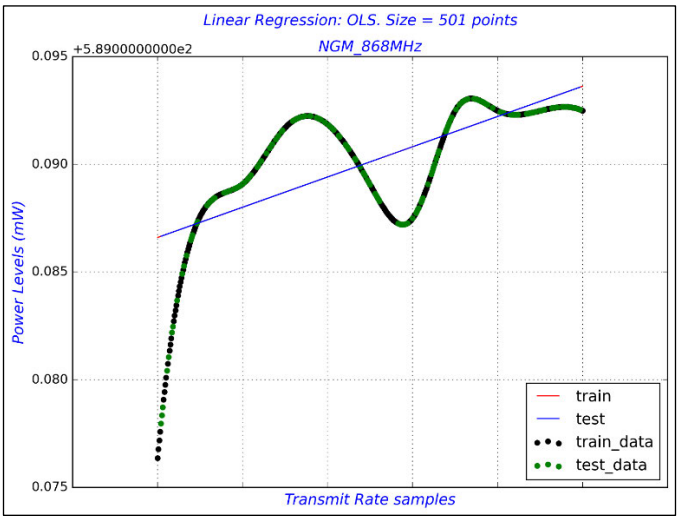

Figure 6.30: OLS regression for NGM 868 MHz

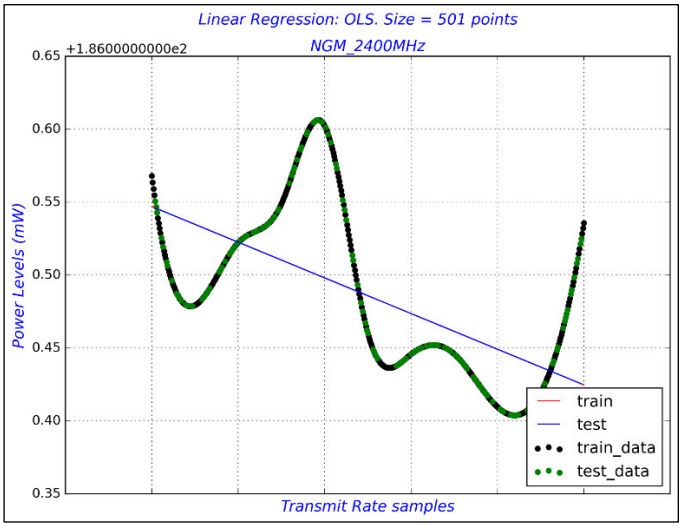

Figure 6.31: OLS regression for NGM 2400 MHz

6.4.2. Ridge models

Figure 6.32 through Figure 6.46 combined the train and test data for the Ridge method and their corresponding regression lines. The results show an optimum performance after tuning the alpha parameter to 0.01. The analysis presented the results per candidate sites and frequency bands.

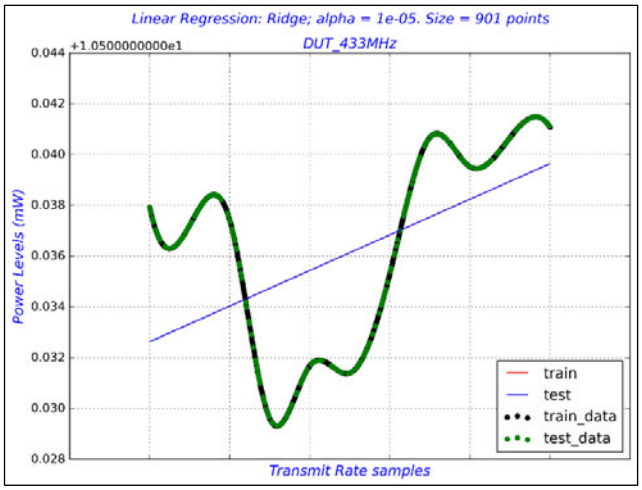

Figure 6.32: Ridge regression for DUT 433 MHz

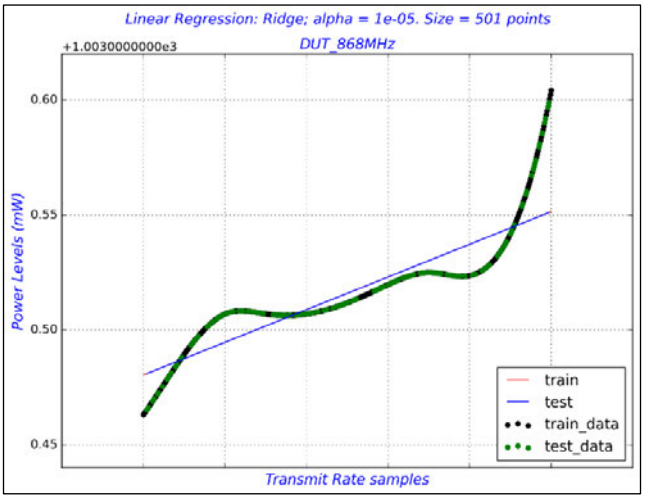

Figure 6.33: Ridge regression for DUT 868 MHz

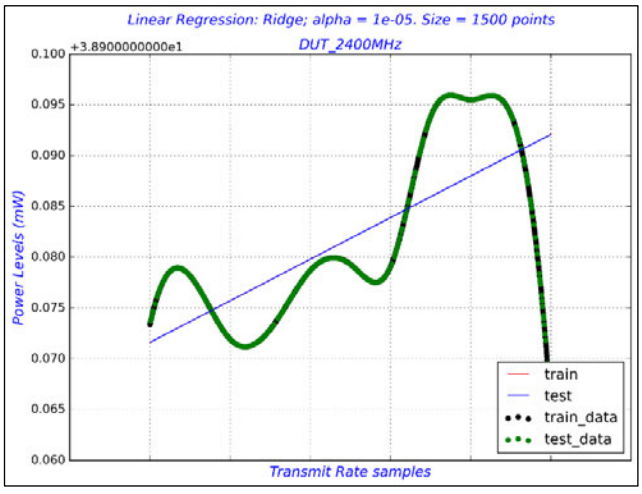

Figure 6.34: Ridge regression for DUT 2400 MHz

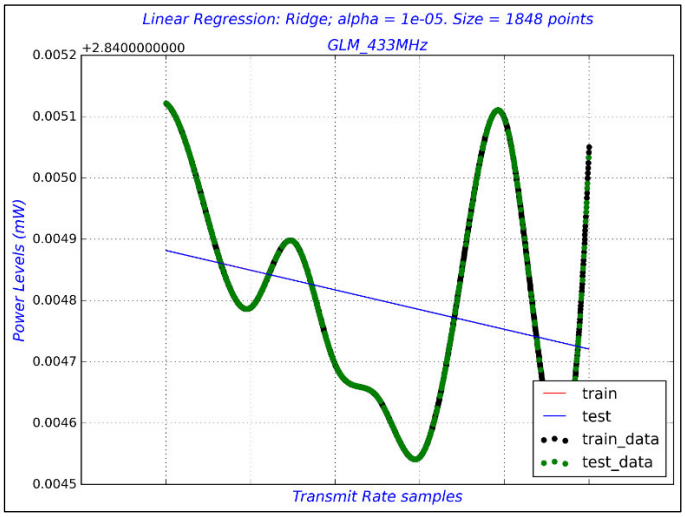

Figure 6.35: Ridge regression for GLM 433 MHz

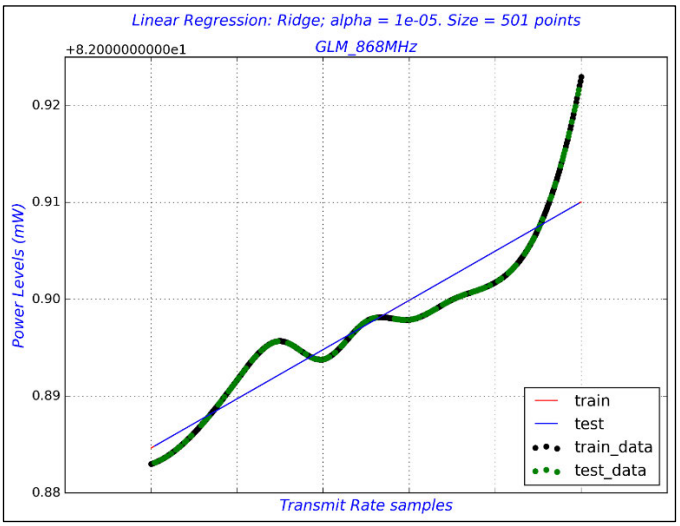

Figure 6.36: Ridge regression for GLM 868 MHz

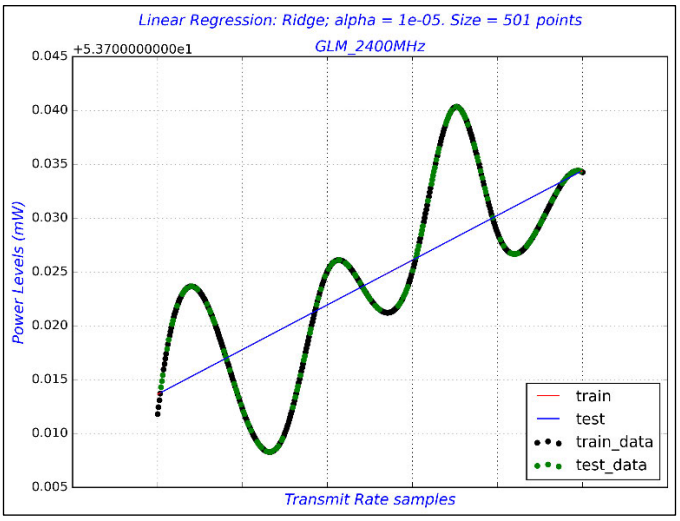

Figure 6.37: Ridge regression for GLM 2400 MHz

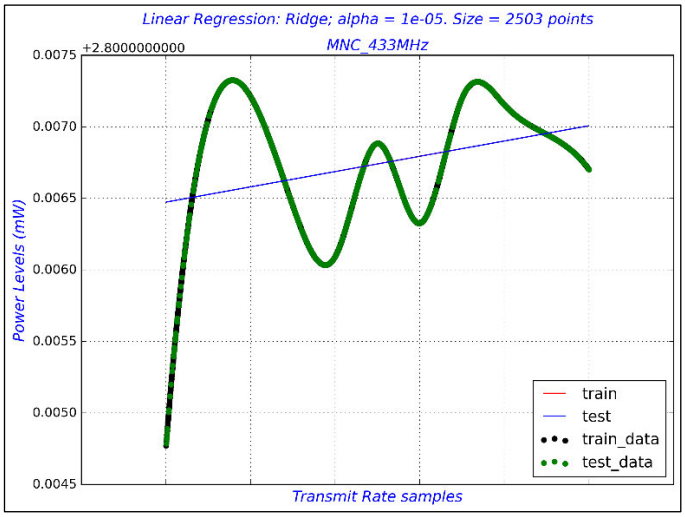

Figure 6.38: Ridge regression for MNC 433 MHz

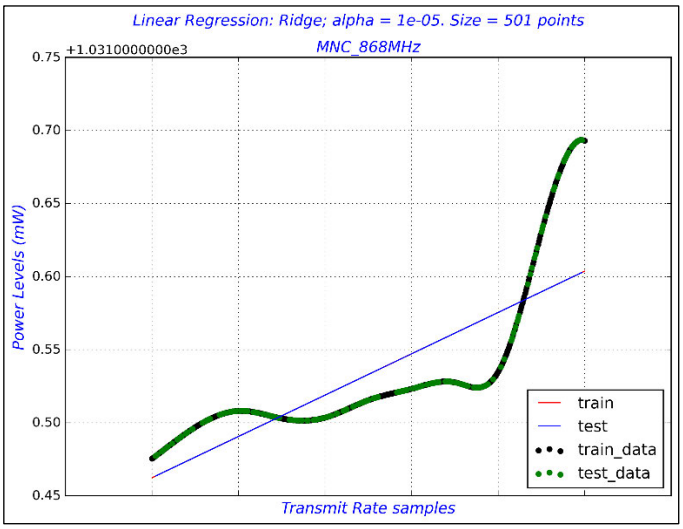

Figure 6.39: Ridge regression for MNC 868 MHz

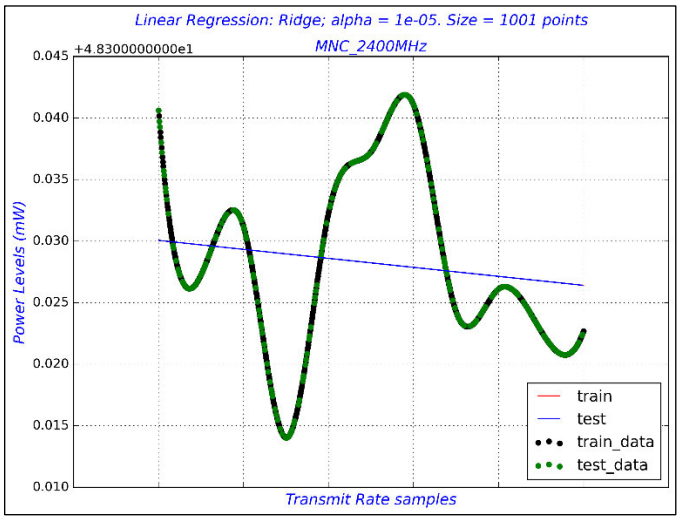

Figure 6.40: Ridge regression for MNC 2400 MHz

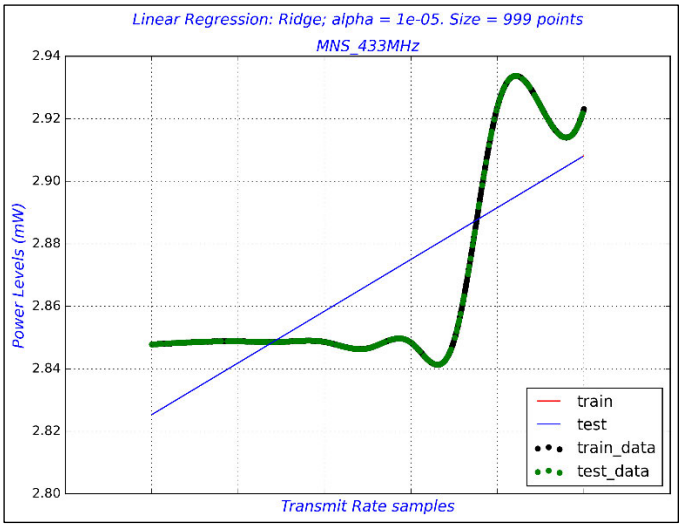

Figure 6.41: Ridge regression for MNS 433 MHz

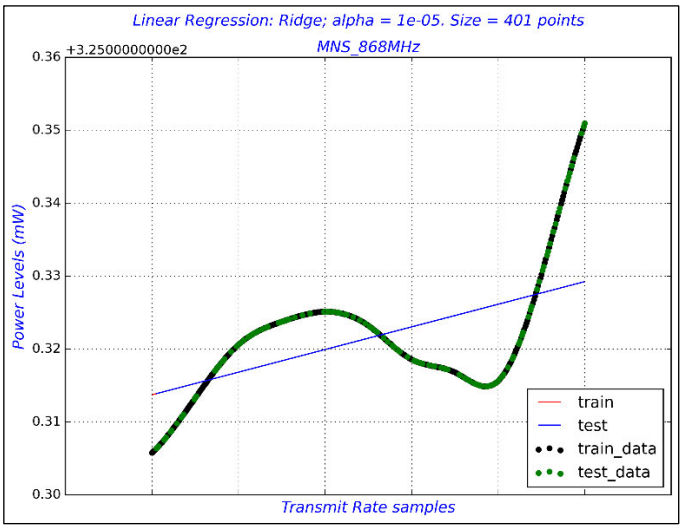

Figure 6.42: Ridge regression for MNS 868 MHz

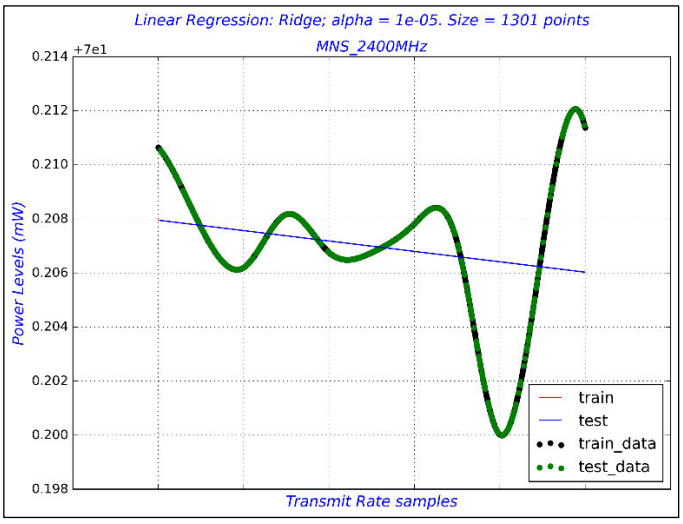

Figure 6.43: Ridge regression for MNS 2400 MHz

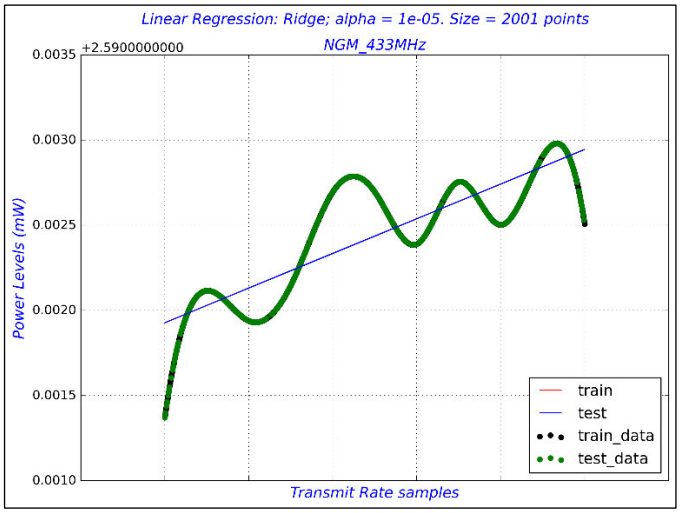

Figure 6.44: Ridge regression for NGM 433 MHz

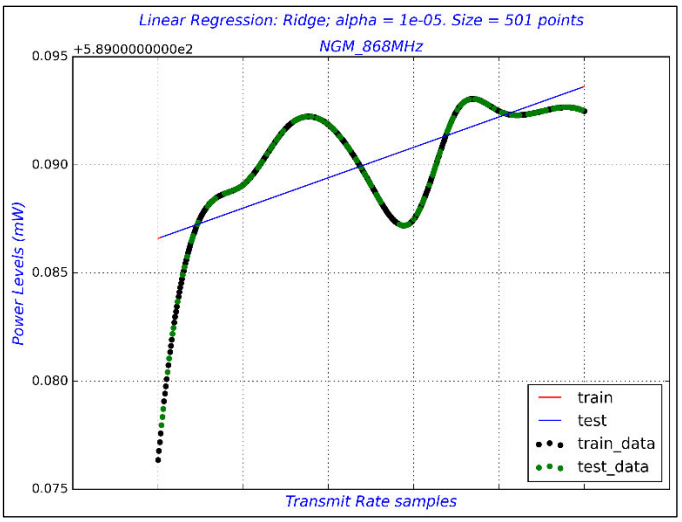

Figure 6.45: Ridge regression for NGM 868 MHz

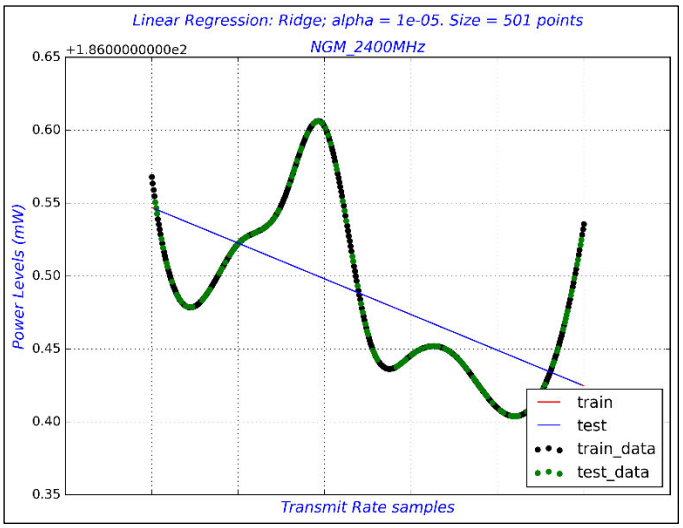

Figure 6.46: Ridge regression for NGM 2400 MHz

6.5. Model Evaluation

In regression modelling, performance metrics summarise the predictive skill of a model. Evaluating these metrics involves calculating the error as the difference between the target and predicted values. Mean squared error (MSE) and R-squared score (R2) are among the commonly-exploited evaluation metrics. MSE is the average of the squares of the errors of all the data points. R2 refers to the coefficient of determination and tells how well a model will predict the unknown samples. R2 returns values in the range [0, 1], in other words, between 0 and 100%. Keeping track of every metric can get tedious. A good practice is ensuring MSE is low. The higher the R2 value, the more valuable the linear model is [120, 121].

Results for model evaluation in this study showed that the produced models obtained very low MSE. Careful adjustments of the sample size of the interpolated data were the techniques applied for tuning the performance of these models. The intention was to obtain comparable R2 values between train and test datasets.

Figures 6.47 through Figure 6.61 present the absolute errors from the OLS regression method for each case grouped per candidate site and frequency band. Figure 6.62 through Figure 6.76 present the absolute errors from the Ridge method. As a general observation, these errors were minuscule.

6.5.1. OLS Errors

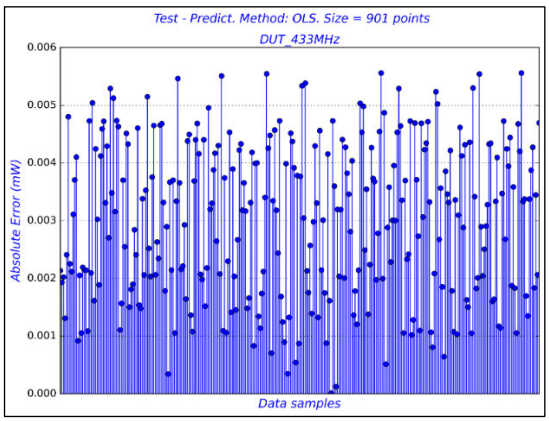

Figure 6.47: OLS absolute error for DUT 433 MHz

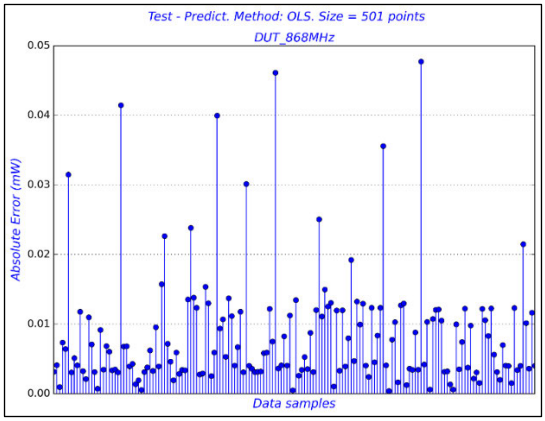

Figure 6.48: OLS absolute error for DUT 868 MHz

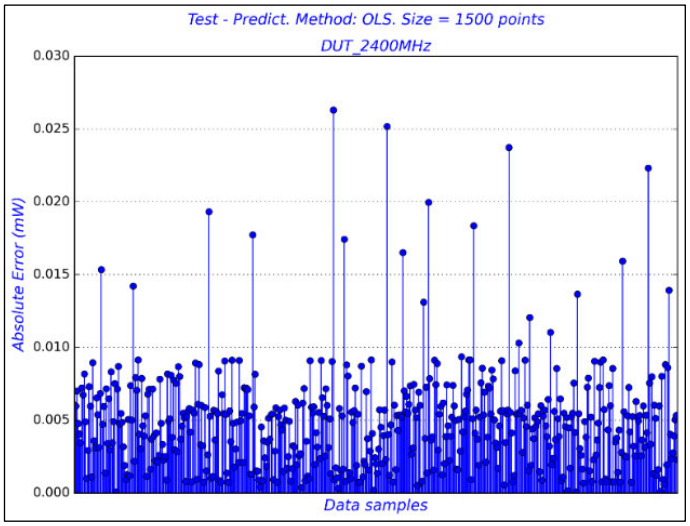

Figure 6.49: OLS absolute error for DUT 2400 MHz

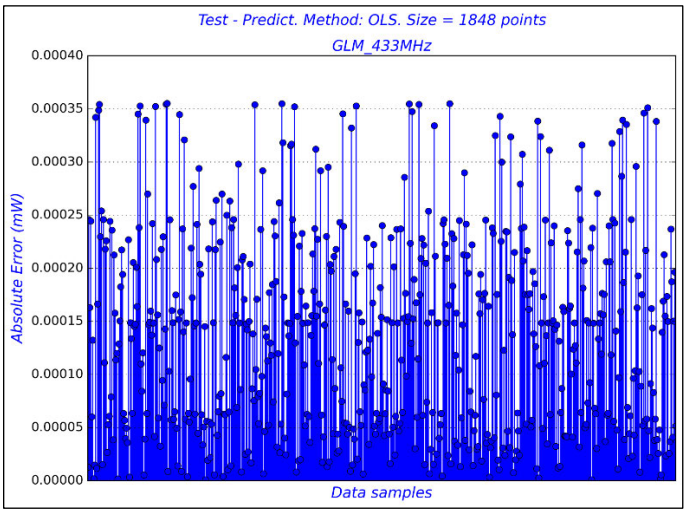

Figure 6.50: OLS absolute error for GLM 433 MHz

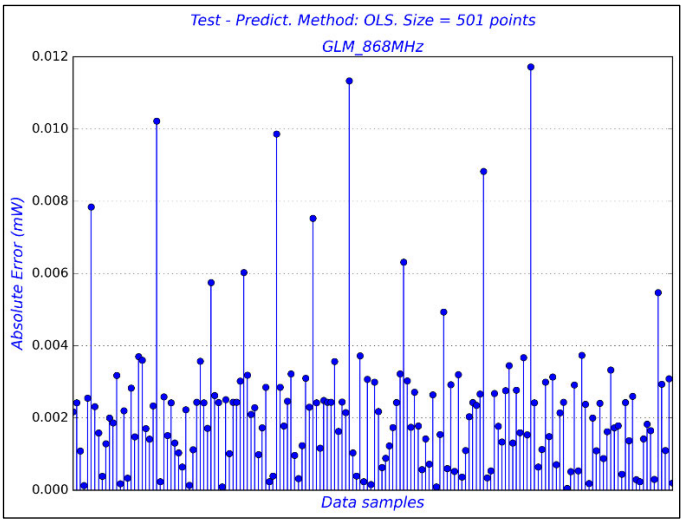

Figure 6.51: OLS absolute error for GLM 868 MHz

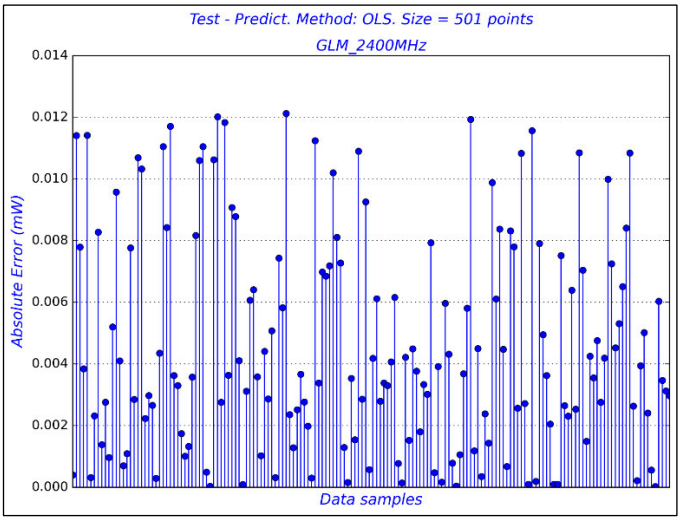

Figure 6.52: OLS absolute error for GLM 2400 MHz

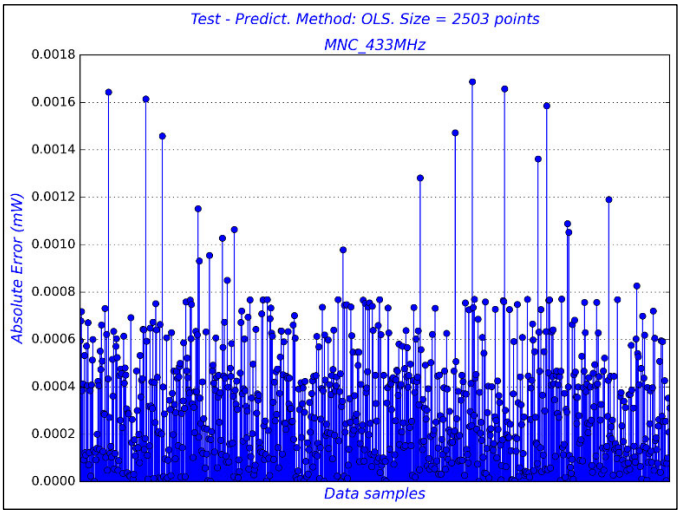

Figure 6.53: OLS absolute error for MNC 433 MHz

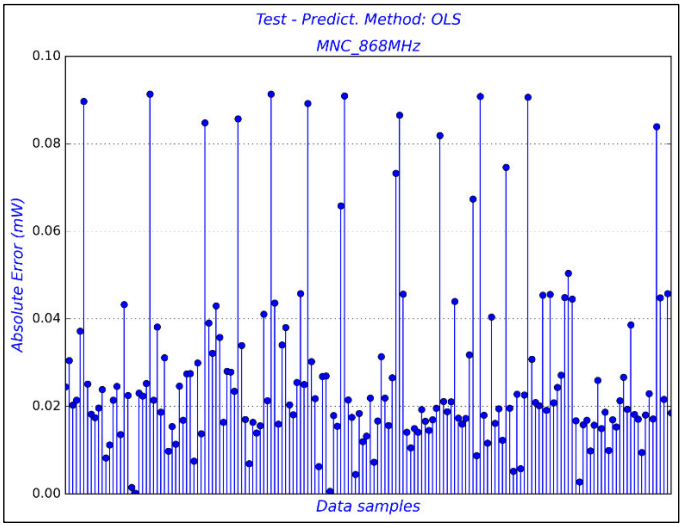

Figure 6.54: OLS absolute error for MNC 868 MHz

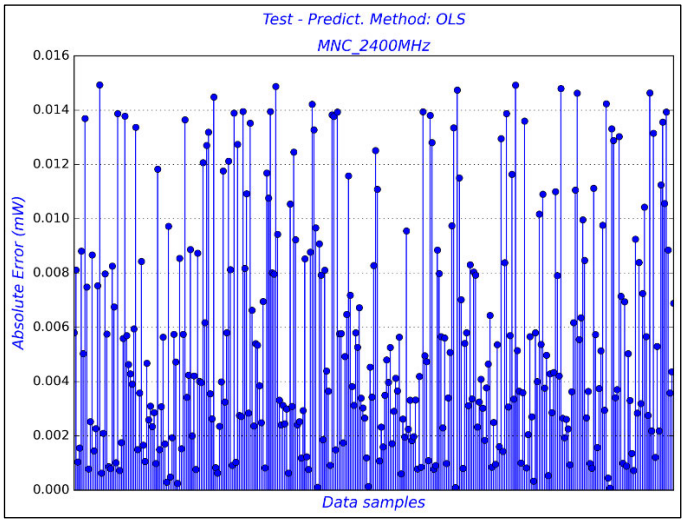

Figure 6.55: OLS absolute error for MNC 2400 MHz

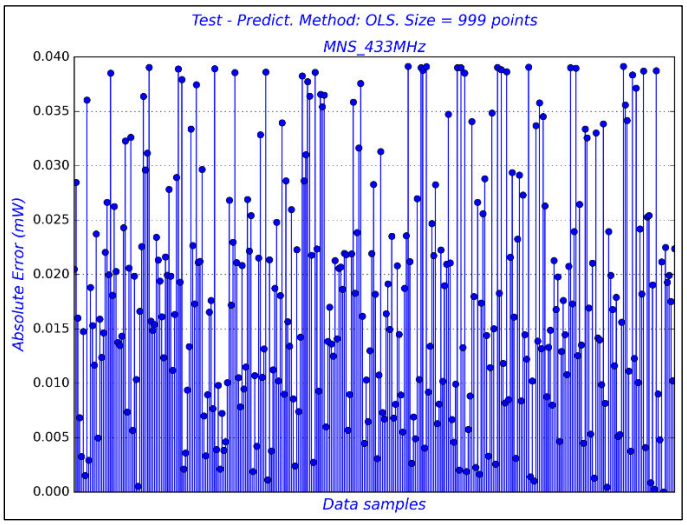

Figure 6.56: OLS absolute error for MNS 433 MHz

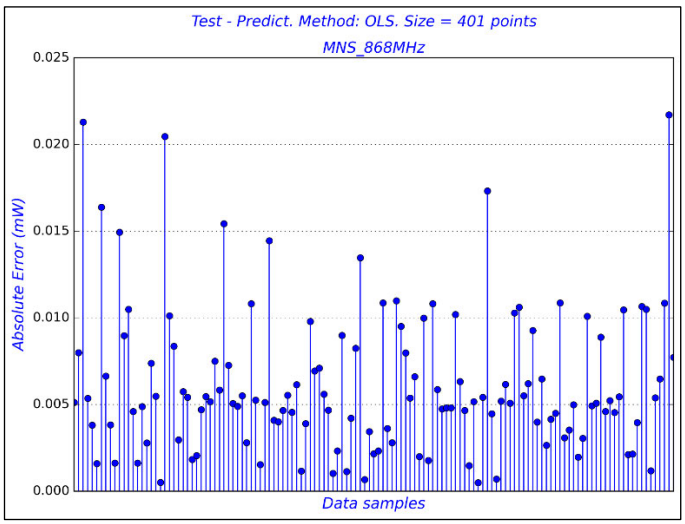

Figure 6.57: OLS absolute error for MNS 868 MHz

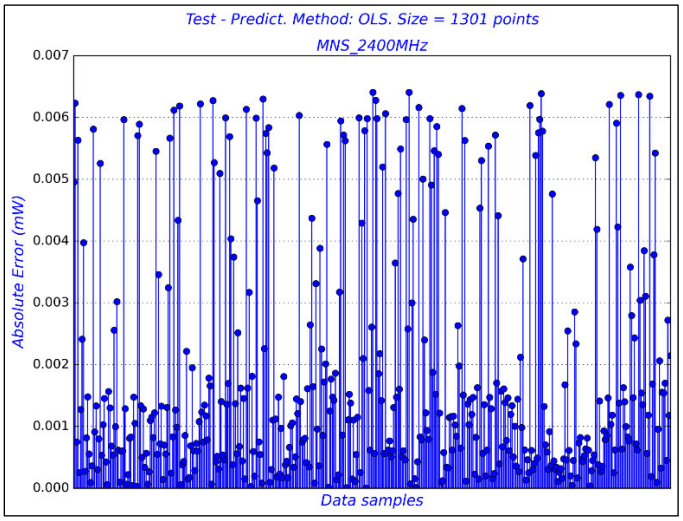

Figure 6.58: OLS absolute error for MNS 2400 MHz

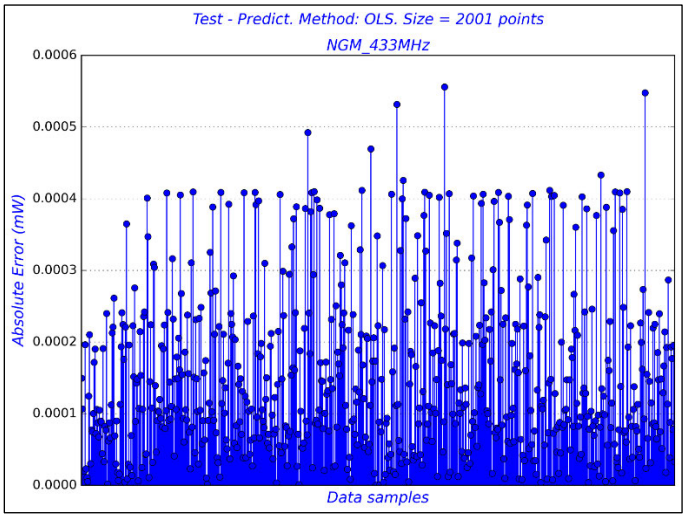

Figure 6.59: OLS absolute error for NGM 433 MHz

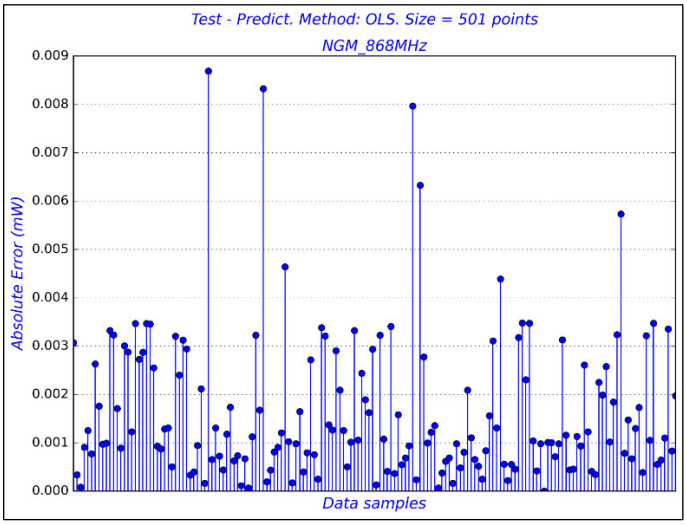

Figure 6.60: OLS absolute error for NGM 868 MHz

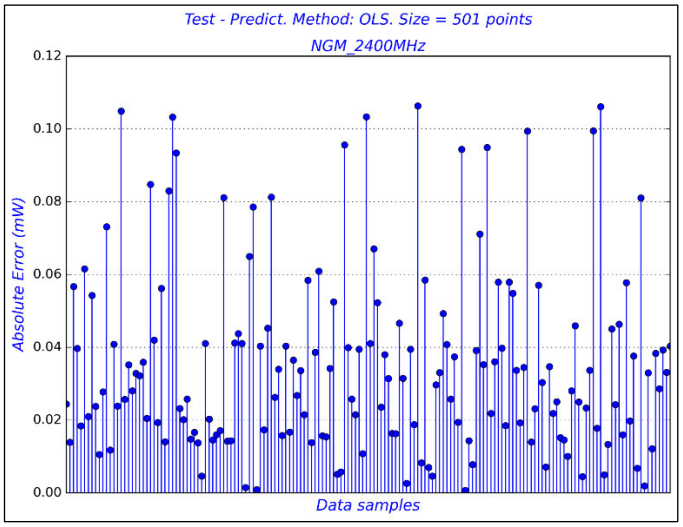

Figure 6.61: OLS absolute error for NGM 2400 MHz

6.5.2. Ridge Errors

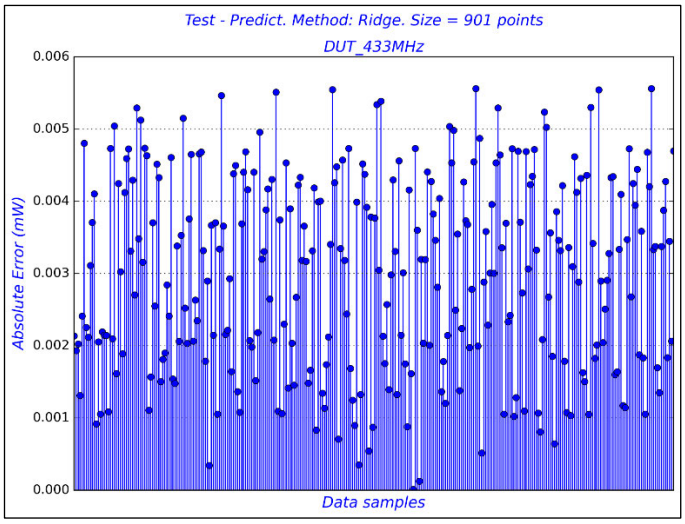

Figure 6.62: Ridge absolute error for DUT 433 MHz

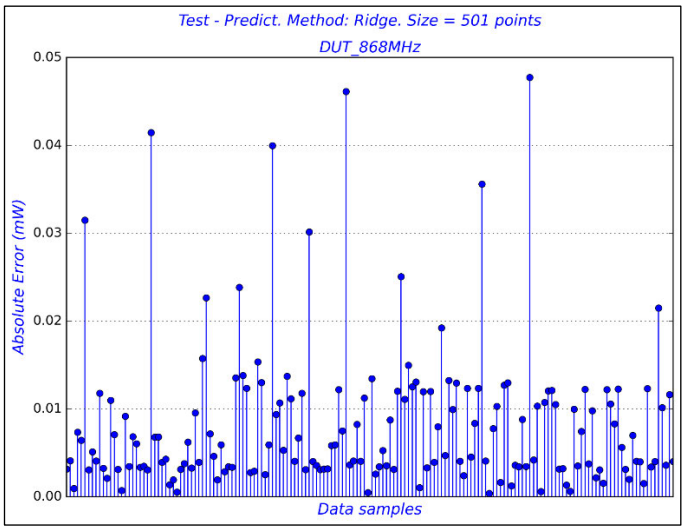

Figure 6.63: Ridge absolute error for DUT 868 MHz

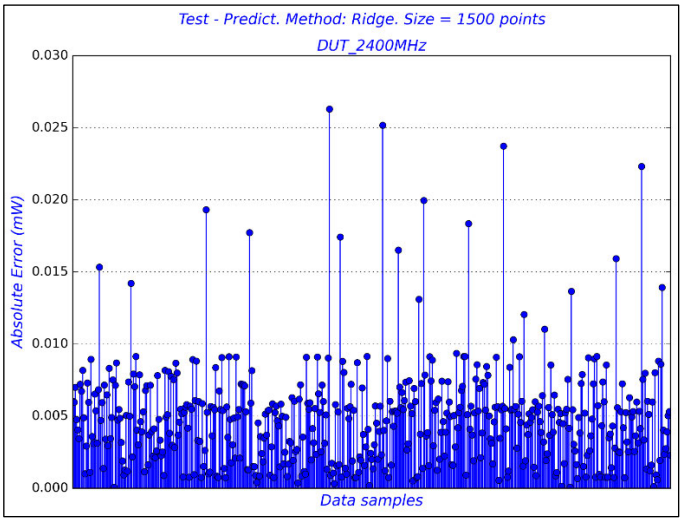

Figure 6.64: Ridge absolute error for DUT 2400 MHz

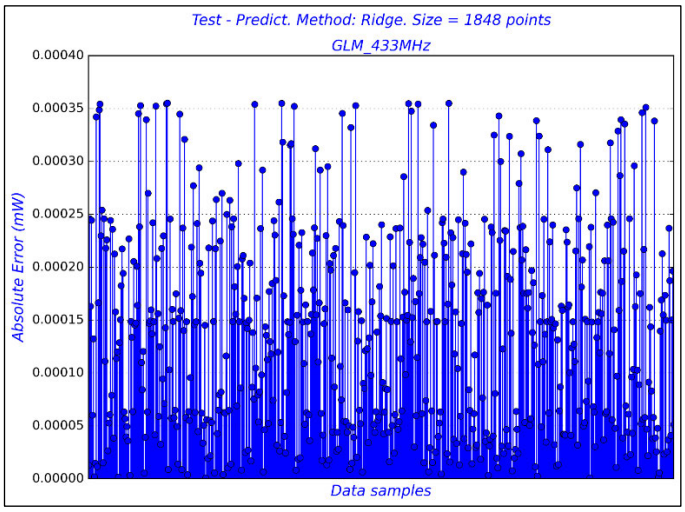

Figure 6.65: Ridge absolute error for GLM 433 MHz

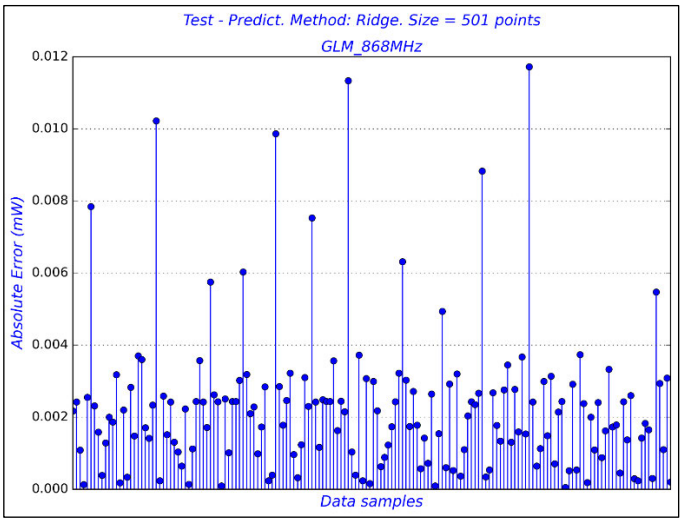

Figure 6.66: Ridge absolute error for GLM 868 MHz

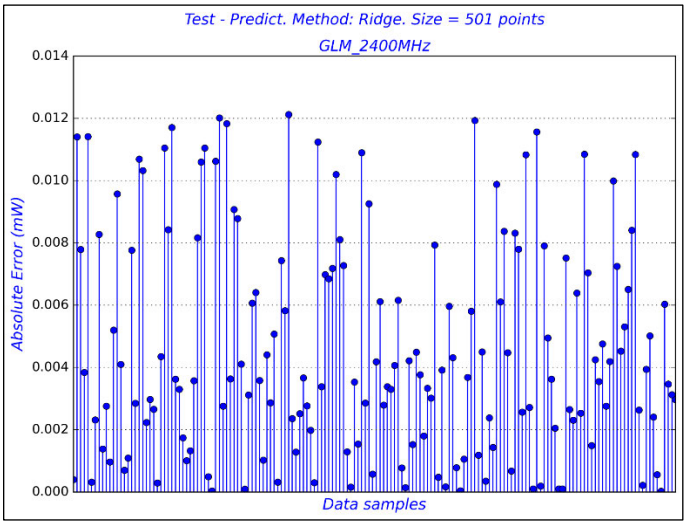

Figure 6.67: Ridge absolute error for GLM 2400 MHz

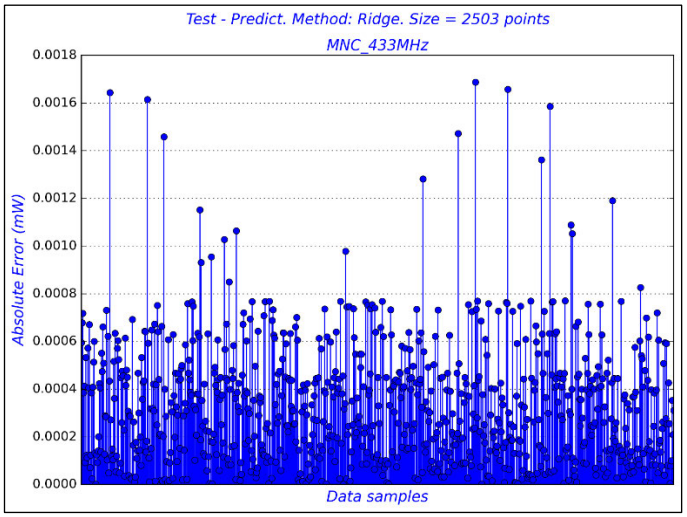

Figure 6.68: Ridge absolute error for MNC 433 MHz

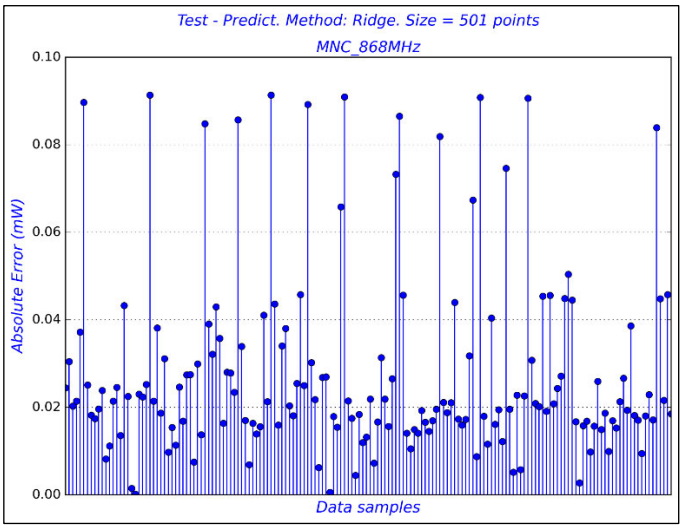

Figure 6.69: Ridge absolute error for MNC 868 MHz

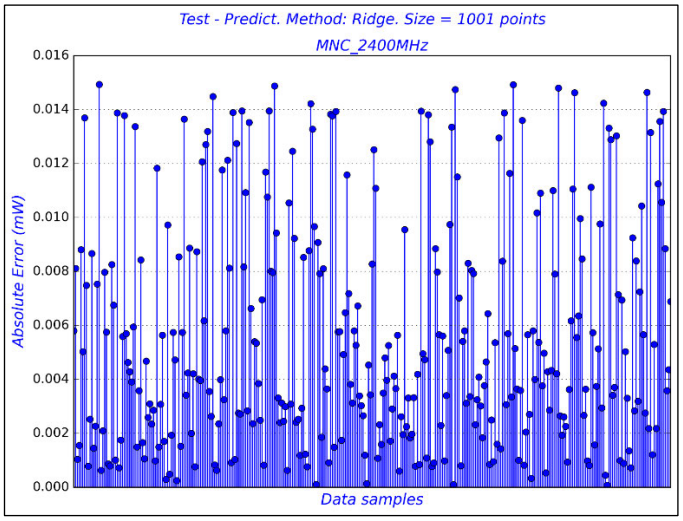

Figure 6.70: Ridge absolute error for MNC 2400 MHz

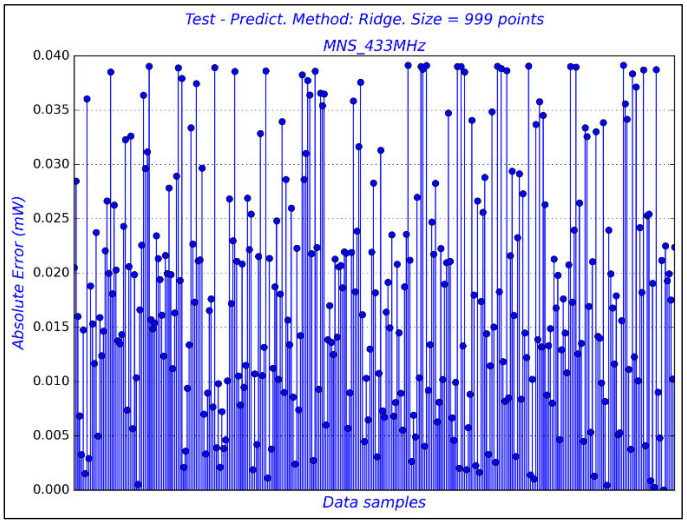

Figure 6.71: Ridge absolute error for MNS 433 MHz

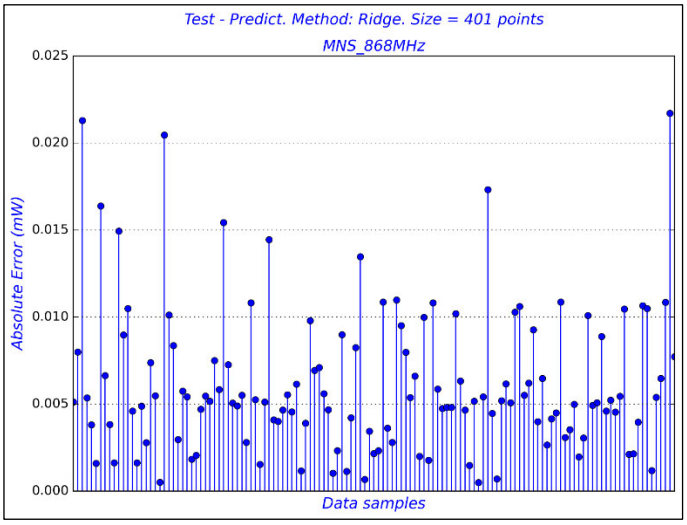

Figure 6.72: Ridge absolute error for MNS 868 MHz

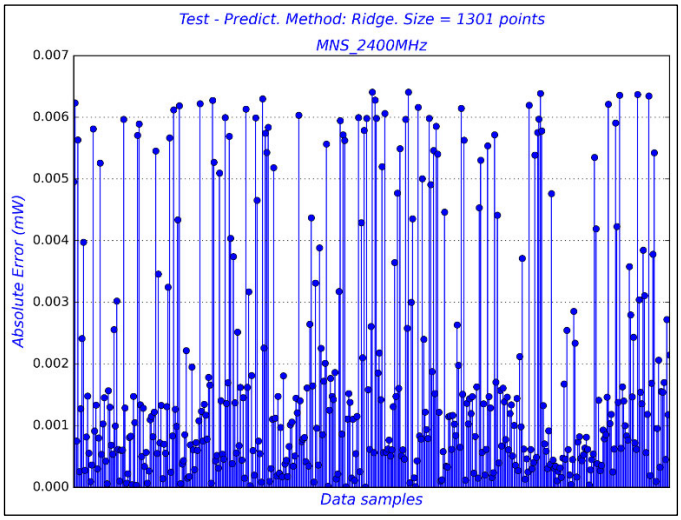

Figure 6.73: Ridge absolute error for MNS 2400 MHz

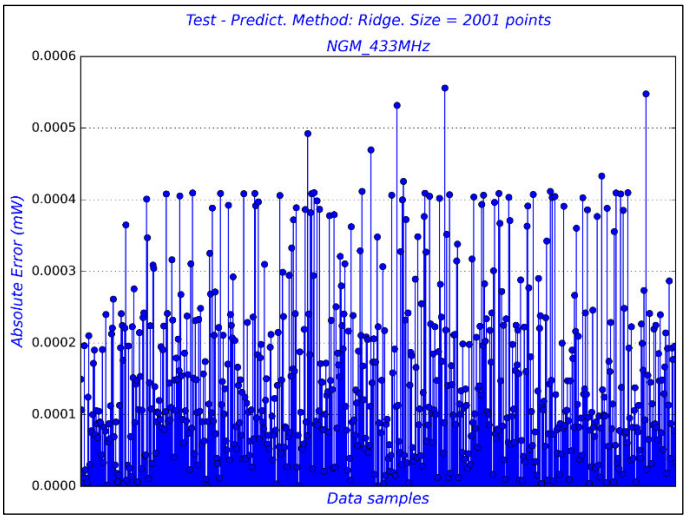

Figure 6.74: Ridge absolute error for NGM 433 MHz

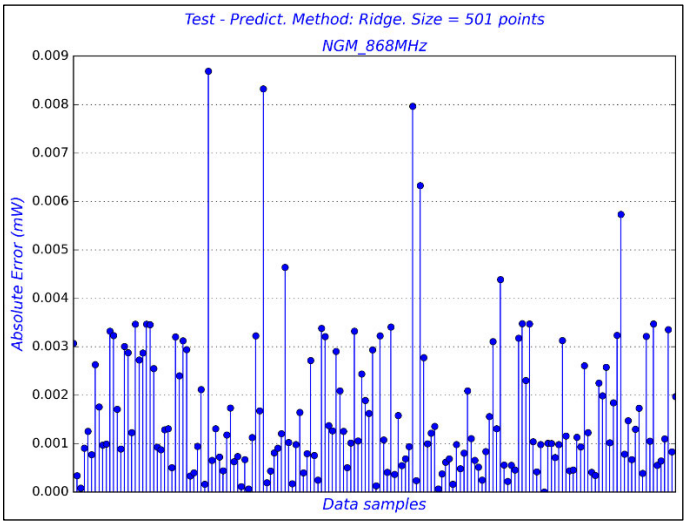

Figure 6.75: Ridge absolute error for NGM 868 MHz

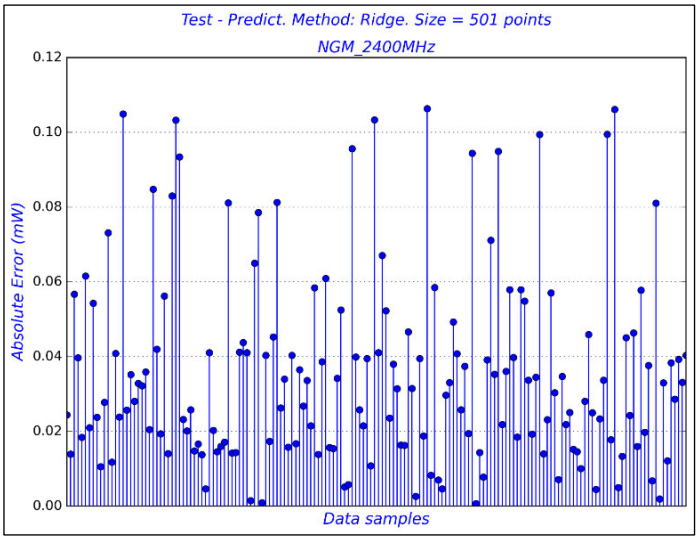

Figure 6.76: Ridge absolute error for NGM 2400 MHz

6.5.3. Evaluation Metric Summary Tables

It was convenient to summarise the performance metrics in Table 6.1 through Table 6.5, which group MSE and R2 per candidate site. The results show the performance of the OLS and Ridge models considering the train and test data. Generally, the metric values are pretty much comparable in all data cases. As discussed earlier, the Ridge regression is helpful for complex models with multiple input variables. With univariate models like this research, the Ridge and OLS performed similarly.

Table 6.1: OLS and Ridge– MSE and R2 metrics for DUT

Data	Method	Metric	Train	Test
DUT 433	OLS	MSE	0.000011	0.000011
	Ridge	MSE	0.000011	0.000011
	OLS	R2	25.861670	25.688542
	Ridge	R2	25.861670	25.688548
DUT 868	OLS	MSE	0.000119	0.000135
	Ridge	MSE	0.000119	0.000135
	OLS	R2	77.913630	77.624246
	Ridge	R2	77.913630	77.624216
DUT 2400	OLS	MSE	0.000037	0.000036
	Ridge	MSE	0.000037	0.000036
	OLS	R2	48.827838	48 .970293
	Ridge	R2	48.827838	48 .970291

Table 6.2: OLS and Ridge – MSE and R2 metrics for GLM

Data	Method	Metric	Train	Test
GLM 433	OLS	MSE	2.629315e-08	2.779452e-08
	Ridge	MSE	2.629315e-08	2.779452e-08
	OLS	R2	7.570041e+00	7.362440e+00
	Ridge	R2	7.570041e+00	7.362440e+00
GLM 868	OLS	MSE	0.000008	0.000009
	Ridge	MSE	0.000008	0.000009
	OLS	R2	87.290382	86.54766
	Ridge	R2	87.290382	86.547639
GLM 2400	OLS	MSE	0.000033	0.000033
	Ridge	MSE	0.000033	0.000033
	OLS	R2	52.128608	53.964257
	Ridge	R2	52.128608	53.964233

Table 6.3: OLS and Ridge – MSE and R2 metrics for MNC

Data	Method	Metric	Train	Test
MNC 433	OLS	MSE	1.954045e-07	1.775974e-07
	Ridge	MSE	1.954045e-07	1.775974e-07
	OLS	R2	1.119996e+01	1.083085e+01
	Ridge	R2	1.119996e+01	1.083085e+01
MNC 868	OLS	MSE	0.001012	0.00122
	Ridge	MSE	0.001012	0.00122
	OLS	R2	62.228984	61.216427
	Ridge	R2	62.228984	61.216397
MNC 2400	OLS	MSE	0.00005	0.000049
	Ridge	MSE	0.00005	0.000049
	OLS	R2	2.147546	3.267597
	Ridge	R2	2.147546	3.267595

Table 6.4: OLS and Ridge – MSE and R2 metrics for MNS

Data	Method	Metric	Train	Test
MNS 433	OLS	MSE	0.000427	0.000455
	Ridge	MSE	0.000427	0.000455
	OLS	R2	56.821334	56.67619
	Ridge	R2	56.821334	56.67619
MNS 868	OLS	MSE	0.000044	0.000055
	Ridge	MSE	0.000044	0.000055
	OLS	R2	30.6988	30.963924
	Ridge	R2	30.6988	30.963902
MNS 2400	OLS	MSE	0.000006	0.000007
	Ridge	MSE	0.000006	0.000007
	OLS	R2	4.726941	4.947832
	Ridge	R2	4.726941	4.947829

Table 6.5: OLS and Ridge – MSE and R2 metrics for NGM

Data	Method	Metric	Train	Test
NGM 433	OLS	MSE	3.643304e-08	3.655356e-08
	Ridge	MSE	3.643304e-08	3.655356e-08
	OLS	R2	7.023056e+01	7.003340e+01
	Ridge	R2	7.023056e+01	7.003340e+01
NGM 868	OLS	MSE	0.000005	0.000005
	Ridge	MSE	0.000005	0.000005
	OLS	R2	43.704054	44.604593
	Ridge	R2	43.704054	44.604601
NGM 2400	OLS	MSE	0.002179	0.001897
	Ridge	MSE	0.002179	0.001897
	OLS	R2	36.390629	35.979657
	Ridge	R2	36.390629	35.979676

Table 6.6 summarises the performance of all models in terms of their coefficients of determination R2. This Table combines the results for the OLS and Ridge models. Out of the thirty models, twenty of them show good and high performance, and so, are deployable with no concern. Ten models produced low performance and would require attention before deploying them. Finally, the Python Pickle module is one of the Python libraries exploited to save and open processed data as a Pickle object[6]. The "dump" method exports models as Pickle objects to make them available in the

[6] https://docs.python.org/3/library/pickle.html#module-pickle

future without needing to re-train the same models from scratch. The Pickle "load" method imports existing models into Pickle objects.

Table 6.6: Performance overview for models produced in this research.

Performance	R2 score	Cases	Count
High	≥ 60 %	DUT 868 GLM 868 MNC 868 NGM 433	8
Good	≥ 30 % & ≤ 60 %	DUT 2400 GLM 2400 MNS 433 MNS 868 NGM 868 NGM 2400	12
Low	≥ 10 % & ≤ 30 %	DUT 433 MNC 433	4
Undesired	≤ 10 %	GLM 433 MNC 2400 MNS 2400	6

6.6. Radio Noise Predictor Software

Figure 6.77 shows the graphical user interface (GUI) of the RNP software[7] that this research developed using the Matlab App Designer tool, version 2023a. This software features two dropdown buttons. The first dropdown button enables users to select an environment among the urban, suburban, and industrial environments. The second dropdown allows users to select one of the exploratory frequencies (433 MHz, 868 MHz, or 2.4 GHz). A click on the RUN button performs three operations:

1. Preloads radio noise data hard-coded in the software. For simplicity and based on the quality of recorded data from the surveying campaigns, the RNP considers only the following cases in Table 6.7.

Table 6.7: Selected data for the RNP development

	433 MHz	868 MHz	2.4 GHz
Urban	DUT	DUT	DUT
Suburban	MNC	GLM	GLM
Industrial	NGM	NGM	NGM

[7] Download RNP software source code here: https://tinyurl.com/2yfbv9uz

2. Performs the interpolation using the cubic spline method.
3. Computes the linear regression model (OLS method) and assesses the evaluation metrics. This routine plots graphs for the original data and the regression line.

Finally, as an essential requirement for this software, the RNP allows users to enter a given value of the Transmit Rate from which the software predicts the radio noise levels in a selected environment and frequency. Two more buttons are available to export the graphs as high resolution (300 dpi) PNG files and save the model in a MAT file.

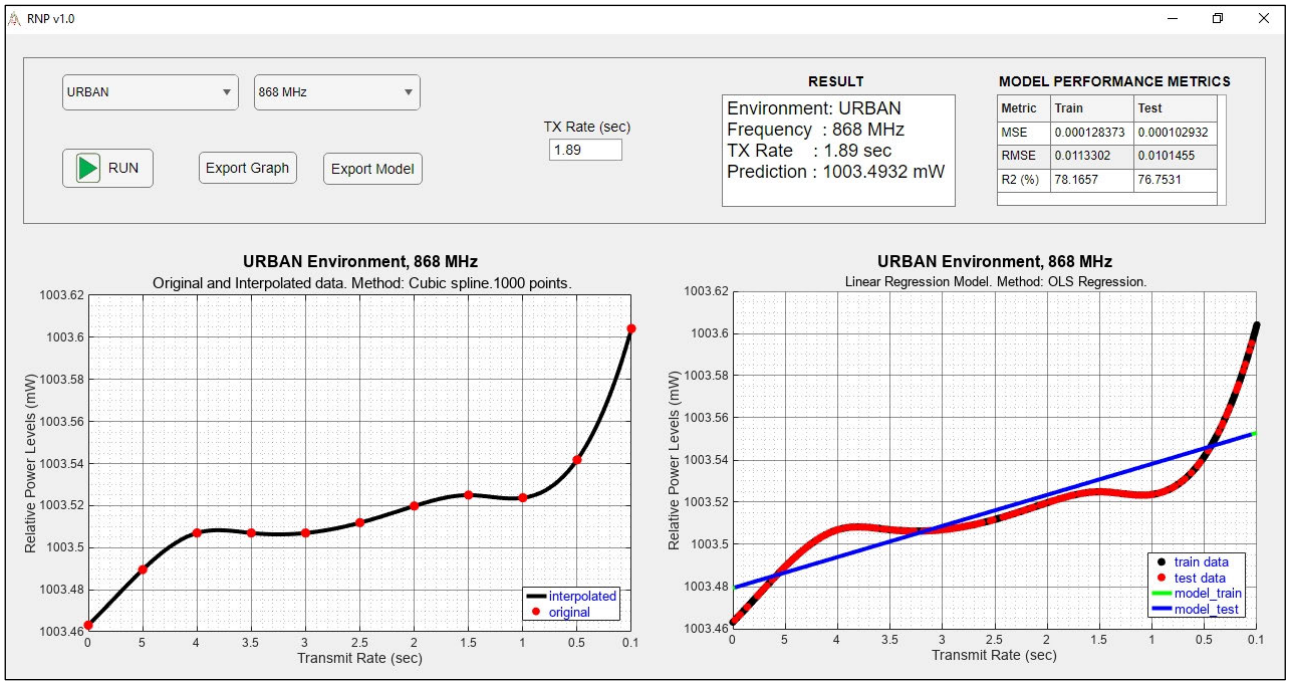

Figure 6.77: RNP GUI

6.7. Summary

It was necessary to start this chapter with a brief theoretical background for Machine Learning. Without diving into minute details, the knowledge gathered provided a secure guideline for identifying and selecting the two algorithms applied in this thesis.

This research did not intend to explore all the existing machine learning mechanisms. The priority was reliably transforming raw radio noise data into a learnable format ready for any machine learning approach. The OLS and Ridge linear regressions trained successful and exploitable models. The RNP linear models developed in Matlab performed comparably to those developed in Python.

CHAPTER 7: CONCLUSION

This thesis, on *"Predicting the Impact of IoT Devices on Radio Frequency Noise in South African Environments"* completes with a satisfactory impression as it resolved the problematic and aim elaborated in Introduction in Chapter 1. Chapter 2 extensively reviewed the literature on the scientific notions and the existing state-of-the-art technologies on environmental radio noise. It was essential to this study to consolidate a sound theoretical understanding of definitions and sources of radio noise, electromagnetic spectrum, exposimetry, risks from RF exposure, noise measurement techniques, mathematical and statistical foundations for radio noise, and the general background of IoT. The acquired knowledge allowed this study to advance confidently.

Building a radio noise measurement system requires human resources, equipment, and budget efforts. Often, well-established institutions support and fund this kind of initiative. However, as documented in Chapter 3, this research had to innovate by developing a simplified and low-cost Radio Noise Surveying System (RNSS). It was a commendable effort to aim to produce reliable scientific results out of an ordinary computer, open-source tools, SDR hardware, and a few modular RF parts. Simulation of the RF front end and in-depth performance testing process validated the choice of components such as LNA, BPF, and antennas. Measurement results demonstrated satisfactory performance. It expanded the research in discovering and innovating solutions in areas of big data, from IoT radio noise processing, data preparation, calibration, amplitude correction, and data science programming. The first survey exposed this RNSS to real-world radio noise data. The encouraging results instilled momentum to progress with this study.

This thesis initially assumed that IoT devices would saturate the environment. However, at the time of this conclusion, the environments, including in South Africa, were still in the infancy development stage as opposed to the predictions. The forecasts certainly suffered as the years 2019, 2020, and 2021 suffered a worldwide shortage in the manufacturing and supply of electronic components, and the year 2020 experienced the ever-unexpected COVID-19 pandemic. Creative work in Chapter 4 designed and constructed IoT radio noise generators, without which there was no way to gather the required quantitative data. The thesis documented the entire design cycle of these units, from formalizing the requirements to sourcing out adequate components to conceptualizing the schematics to laying out the PCB to assembling the boards to programming them, to testing and optimizing their performance. The PCB for the ING devices included three frequency-based variants, each covering 433 MHz, 868 MHz, and 2.4 GHz bands.

It was a remarkable endeavour to drive around many places in Durban, daytime and night times, on and off weekends, spending hours on different candidate sites for collecting data as part of the

surveying campaign. Although the intention was to have many more candidate sites, social factors such as safety, security, and crime rate limited the selection of locations for the campaign. As documented in Chapter 4, a standard operating procedure was prepared to follow repeatable work instructions during measurement.

The statistical analysis in Chapter 5 allowed this study better to discern the mathematical nature of the collected data. Visualization approaches exposed the particularities of each case grouped in candidate site and frequency band. Statistical and empirical analysis revealed that although operating as low-power applications and transmitting RF signals in a disciplined manner to comply with regulatory directives, the levels of ambient radio noise were indeed proportional to the IoT RF transmit rates. The more radio activities by the IoT industry, the more the environmental radio noise increases, irrespective of the type of environment and operating frequency.

This research would not have achieved the intended goals without developing predictive machine learning models. Chapter 6 provided a brief theory of machine learning as guidance towards predictive modelling. The nature of collected data, as quantitative and continuous, dictated exploiting linear regressions. This thesis was in favour of the OLS and Ridge linear regression models. Five candidate sites and three frequency bands produced forty-five predictive models. As evaluation metrics, all models showed low root mean squared (MSE). The coefficient of determination R2 measured good percentages on thirty cases with combined OLS and Ridge's models. Six models showed low performance, and nine models were undesired. From these models, the Radio Noise Predictor (RNP) software allows any interested users to assess ambient radio noise for their sites.

Besides this thesis, which has elaborated intellectual and engineering activities, this research delivered two outputs: 1- a peer-reviewed journal article [106] published by MDPI, and 2- a conference paper [3] presented at the ICTAS 2022.

Executing this research navigated through many programs like systems engineering, general and embedded electronics, RF and microwaves, and data science, earning new expertise and contributing as a valuable resource to the knowledge domain. Although approached within the delimitations discussed in the Introduction, this study revealed interesting aspects that future works may interrogate. Another study can explore radio noise outside the ISM bands. A project like mapping radio noise for a city or a town requires surveying more than five candidate sites. While this research allowed randomness concerning the survey time, other research can focus on specific times of day, week, and seasons. Machine learning is a forever unending process. While this research applied linear regression algorithms, another study can investigate diverse algorithms.

REFERENCES

[1] A. D. Spaulding, "The Natural and Man-made Noise Environment in Personal Communications Services Bands.," U.S. Department of Commerce, Report TR-96-330, 1997. Accessed: 24 January 2018. [Online]. Available: https://www.its.bldrdoc.gov/publications/download/96-330.pd

[2] P. Enge, D. Akos, J. Do, J. Simoneau, W. Pearson, and V. Seetharam, "Measurements of Man-Made Spectrum Noise Floor," NASA, , CR-2004-213551, 2004. [Online]. Available: https://ntrs.nasa.gov/search.jsp?R=20050041714

[3] D. G. K. Ingala and S. Reddy, "Baseline Measurement of the Environmental RF Noise Floor in the ISM Bands," presented at the ICTAS 2022, Durban, 2022.

[4] ITU, "Methods for measurementsof radio noise: Recommendation ITU-R SM.1753-2," in "Spectrum management," ITU, , 2012. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/sm/R-REC-SM.1753-2-201209-I!!PDF-E.pdf

[5] ICASA, "Electronic Communications Act, N°36 of 2005: Regulations in respect of Licence Exemptions ", N°31290 ed. Pretoria, South Africa: Government Gazette, 2008.

[6] B. Skeie and B. Solberg, "External man-made radio noise measurements," Norwegian Defence Research Establishment, Norway, , FFI-RAPPORT 16/00869, 2016.

[7] J. A. Wepman and G. A. Sanders, "Wideband Man-Made Radio Noise Measurements in the VHF and Low UHF Bands," U.S. Department of Commerce, USA,, TR-11-478, 2011.

[8] A. Shukla, "Feasibility Study into the Measurement of Man-Made Noise," DERA, United Kingdom, , AY 3952, 2001. [Online]. Available: http://educypedia.karadimov.info/library/mmnoise.pdf

[9] R. Leck, "Results of Ambient RF Environment and Noise Floor Measurements Taken in the U.S in 2004 and 2005," World Meteorological Organization, Geneva, , CBS/SG-RFC 2005/Doc. 5(1), 2006. Accessed: 10 January 2018. [Online]. Available: https://www.wmo.int/pages/prog/www/TEM/SG-RFC06/Ambient-RF-noise.pdf

[10] C. Bradshaw, "Electromagnetic Radiation Survey of the Minto Urban Renewal Project," Epping NSW, Australia, No 50816, 2005.

[11] H. v. d. Groenendaal, "Rising RF noise floor calls for urgent action." [Online]. Available: http://www.ee.co.za/article/rising-rf-noise-floor-calls-urgent-action.html

[12] S. B. Dunn, "A study of the SA SKA RFI Measurement Systems," Master of Science, Department of Electrical Engineering, University of Cape Town, 2008. [Online]. Available: http://www.rrsg.uct.ac.za/theses/msc_theses/sdunn_thesis.pdf

[13] Pietro Bolli, Francesco Gaudiomonte, Filippo Messina, Roberto Ambrosini, Claudio Bortolotti, and M. Roma, "The RFI monitoring systems for the Medicina and the Sardinia Radio Telescopes," *Proceedings of Science,* 2010. [Online]. Available: https://www.researchgate.net/publication/239593072_The_RFI_monitoring_systems_for_the_Medicina_and_the_Sardinia_Radio_Telescopes.

[14] C. W. Sayre, *Complete Wireless Design*. New-York: McGraw-Hill, 2001.

[15] R. J. Tuieng, S. H. Cartmell, C. C. Kirwan, and M. J. Sherratt, "The Effects of Ionising and Non-Ionising Electromagnetic Radiation on Extracellular Matrix Proteins," *Cells,* vol. 10, no. 3041, 2021, doi: 10.3390/cells10113041.

[16] ICNIRP, "Guidelines For Limiting Exposure To Electromagnetic Fields (100 KHz To 300 GHz)," *Health Physics,* vol. 118, no. 5, pp. 483–524, 2020. [Online]. Available: https://www.icnirp.org/cms/upload/publications/ICNIRPrfgdl2020.pdf.

[17] IARC, "Non-ionizing Radiation, Part 2: Radiofrequency Electromagnetic Fields - Monographs on the Evaluation of Carcinogenic Risks to Humans," IARC, 2013, vol. Volume 102. [Online]. Available: https://publications.iarc.fr/Book-And-Report-Series/Iarc-Monographs-On-The-Identification-Of-Carcinogenic-Hazards-To-Humans/Non-ionizing-Radiation-Part-2-Radiofrequency-Electromagnetic-Fields-2013

[18] C. Bianchi and A. Meloni, "Natural and man-made terrestrial electromagnetic noise: an outlook," *Annals of Geophysics,* vol. 50, no. 3, 2007.

[19] ITU, "Radio Noise: Recommendation ITU-R P.372-13," in "Radiowave Propagation," ITU, Geneva, , 2016. [Online]. Available: https://www.itu.int/dms_pubrec/itu-r/rec/p/R-REC-P.372-13-201609-I!!PDF-E.pdf

[20] F. R. Dungan, *Electronic Communications Systems*, 2nd ed. New York: Delmar, 1993.

[21] F. S. Frank Leferink, J. Catrysse, S. Batterman, V. Beauvois, A. Roc'h, "Man-Made Noise in Our Living Environments," *The Radio Science Bulleting,* no. 334, p. 86, 2010. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7911080.
[22] J. K. Hargreaves, *The solar-terrestrial environment: An introduction to geospace - the science of the terrestrial upper atmosphere, ionosphere and magnetosphere.* (Cambridge atmospheric and space science). Cambridge University Press, 1992.
[23] J. Daintith and W. Gould, *The Facts on File Dictionary of Astronomy*, 5 ed. Facts On File, 2006.
[24] R. A. Dastrup, *Physical Geography and Natural Disasters*. SLCC, 2020.
[25] E. Britannica, " Ionosphere and Magnetosphere," ed, 2023.
[26] J. D. Keyser, S. Stankov, and T. Verhulst, "Earth's magnetosphere and ionosphere," *Revue E tijdschrift,* vol. 130, no. 2, 2014. [Online]. Available: http://ionosphere.meteo.be/sites/ionosphere.meteo.be/files/documentation/publications/Team-Journals/Stankov_2014_KBVE_Earths%20magnetosphere%20and%20ionosphere.pdf.
[27] A. Petrukovich *et al.*, "Monitoring of Physical Processes in Upper Atmosphere, Ionosphere and Magnetosphere in Ionosphere Space Missions.," presented at the XII International Conference "Solar-Terrestrial Relations and Physics of Earthquake Precursors (STRPEP)", 2021. [Online]. Available: https://www.epj-conferences.org/articles/epjconf/pdf/2021/08/epjconf_strpep2021_02010.pdf.
[28] IARC, "Non-ionizing Radiation, Part 1: Static and Extremely Low-frequency (ELF) Electric and Magnetic Fields - Monographs on the Evaluation of Carcinogenic Risks to Humans," IARC, Volume 80, 2002. [Online]. Available: https://publications.iarc.fr/Book-And-Report-Series/Iarc-Monographs-On-The-Identification-Of-Carcinogenic-Hazards-To-Humans/Non-ionizing-Radiation-Part-1-Static-And-Extremely-Low-frequency-ELF-Electric-And-Magnetic-Fields-2002
[29] T. Verhulst, S. Stankov, and J. Rasson, "The Geomagnetic Field: an Actively Changing Global Phenomenon," *Revue E tijdschrift,* vol. 130, no. 2, 2014. [Online]. Available: http://ionosphere.meteo.be/sites/ionosphere.meteo.be/files/documentation/publications/Team-Journals/Verhulst_2014_KBVE_The%20geomagnetic%20field%20-%20an%20actively%20changing%20global%20phenomenon.pdf.
[30] P. Dabhade, D. J. Saikia, and M. Mahato, "Decoding the giant extragalactic radio sources," *Journal of Astrophysics and Astronomy* vol. 44, no. 13, 2023, doi: https://doi.org/10.1007/s12036-022-09898-5.
[31] NATO, "Radio Frequency Ambient Environments," *Electrical and Electromagnetic Environmental Conditions,* vol. 250, no. 1, pp. 16-26, 2009. [Online]. Available: https://www.dau.mil/cop/e3/DAU%20Sponsored%20Documents/aectp%20250%201.pdf.
[32] P. Ängskog, C. Karlsson, J. F. Coll, J. Chilo, and P. Stenumgaard, "Sources of Disturbances on Wireless Communication in Industrial and Factory Environments," in *2010 Asia-Pacific International Symposium on Electromagnetic Compatibility*, Beijing, April 12 - 16, 2010 2010.
[33] J. C. Peter Stenumgaard, Javier Ferrer-Coll, Per Ängskog, "Challenges and Conditions for Wireless Machine-to-Machine Communications in Industrial Environments," *IEEE Communications Magazine,* vol. 51, no. 6, 2013. [Online]. Available: http://pgembeddedsystems.com/securelogin/upload/project/IEEE/33/PGEMB0020/06525614.pdf.
[34] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Computer Networks,* vol. 54, pp. 2787-2805, 2010. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.719.9916&rep=rep1&type=pdf.
[35] Y. Qin, Q. Z. Sheng, N. J. G. Falkner, S. Dustdar, H. Wang, and A. V. Vasilakos, "When things matter: A survey on data-centric internet of things," *Network and Computer Applications,* vol. 64, pp. 137-153, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1084804516000606.
[36] M. Hatton and G. Chua, "Service Provider Opportunities & Strategies in the Internet of Things," Machina Research, , 2015. Accessed: 20 January 2017. [Online]. Available: http://cdn.iotwf.com/resources/106/SPs_in_IoT_-_State_of_the_Industry_2015_12_03.pdf
[37] N. Heuveldop, "Ericsson Mobility Report - June 2017," Ericsson, , Stockholm, Sweden, 2017. Accessed: 17 January 2018. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports
[38] F. Jejdling, "Ericsson Mobility Report - November 2019," Ericsson, Stockholm, Sweden, 2019. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports

[39] F. Sullivan, "Growing Industry Applications of LPWAN Technologies," Frost & Sullivan, , 2018. [Online]. Available: https://go.murata.com/rs/382-MEZ-125/images/Ke5NTPU7.pdf
[40] D. Owen, "Good Practice Guide to Phase Noise Measurement," in "Measurement Good Practice Guide," National Physical Laboratory, United Kingdom, 2004, vol. 68. Accessed: 27 August 2019. [Online]. Available: https://www.npl.co.uk/special-pages/guides/gpg68_noise
[41] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase Noise in Oscillators: A Unifying Theory and Numerical Methods for Characterisation," *Design Automation Conference,* vol. 35, 1998. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.31.8589&rep=rep1&type=pdf.
[42] G. Veale, C. Nel, and G. Ewation, "Phase noise in RF synthesizers," in *Aardvark AOC Little Crow Conference*, CSIR, A. AOC, Ed., 10 November 2010 2010: Aardvark AOC. [Online]. Available: http://www.aardvarkaoc.co.za/wp-content/Proceedings/201011%20Aardvark%20AOC/Gert.pdf. [Online]. Available: http://www.aardvarkaoc.co.za/wp-content/Proceedings/201011%20Aardvark%20AOC/Gert.pdf
[43] H. Omer, "Radiobiological effects and medical applications of non-ionizing radiation," *Saudi Journal of Biological Sciences,* vol. 28, no. 2021, pp. 5582-5592, 2021.
[44] R. Tinker *et al.*, "A coherent framework for non-ionising radiation protection," *Journal of Radiological Protection,* vol. 42, 2022, doi: 10.1088/1361-6498/ac3bc3.
[45] V. S. Lucas, R. S. Burk, S. Creehan, and M. J. Grap, "Utility of High-Frequency Ultrasound: Moving Beyond the Surface to Detect Changes in Skin Integrity," *Plastic Surgical Nursing,* vol. 34, no. 1, pp. 34-38, 2014, doi: 10.1097/PSN.0000000000000031.
[46] UiB. "The Electromagnetic spectrum." University of Bergen. https://www.uib.no/en/hms-portalen/75292/electromagnetic-spectrum (accessed 23/08/2023, 2023).
[47] M. R. Sanad, "Determination of Photon Energy, Frequency and Wavelength with Main Physical Parameters of Particles," *SSRN,* 2023, doi: 10.2139/ssrn.4461441.
[48] ESA. "Satellite frequency bands." European Space Agency (ESA). https://www.esa.int/Applications/Connectivity_and_Secure_Communications/Satellite_frequency_bands (accessed 22-08-2023, 2023).
[49] N. T. Tran, L. Jokic, J. Keller, J. U. Geier, and R. Kaldenhoff, "Impacts of Radio-Frequency Electromagnetic Field (RF-EMF) on Lettuce (Lactuca sativa) - Evidence for RF-EMF Interference with Plant Stress Responses," *Plants,* vol. 12, no. 1082, 2023, doi: https://doi.org/10.3390/plants12051082.
[50] M. Czerwinski, A. Vian, B. A. Woodcock, P. Golinski, L. R. Virto, and L. Januszkiewicz, "Do electromagnetic fields used in telecommunications affect wild plant species? A control impact study conducted in the field," *Ecological Indicators,* vol. 150, no. 110267, 2023, doi: https://doi.org/10.1016/j.ecolind.2023.110267.
[51] A. Thielens, "Environmental impacts of 5G: A literature review of effects of radio-frequency electromagnetic field exposure of non-human vertebrates, invertebrates and plants.," EPRS, PE 690.021, 2021.
[52] S. Cucurachi, W. L. M. Tamis, M. G. Vijver, W. J. G. M. Peijnenburg, J. F. B. Bolte, and G. R. d. Snoo, "A review of the ecological effects of radiofrequency electromagnetic fields (RF-EMF)," *Environment International,* vol. 51, pp. 116-140, 2013, doi: http://dx.doi.org/10.1016/j.envint.2012.10.009.
[53] T. A. M. Stege, J. F. B. Bolte, L. Claassen, and D. R. M. Timmermans, "Electromagnetic field exposure in power plants: a qualitative assessment of work safety perceptions among employees," *Risk Research,* vol. 23, no. 12, 2020, doi: 10.1080/13669877.2020.1750459.
[54] C. E. Haedrich and D. J. Breton, "Measuring Very High Frequency and Ultrahigh Frequency Radio Noise in Urban Environments: A Mobile Measurement System for Radio-Frequency Noise," U.S. Army Engineer Research and Development Center, ERDC/CRREL TR-19-8, 2019.
[55] D. J. Breton and A. C. Meyer, "Land Use and the Character of Urban Radio-frequency Noise in a Small City," presented at the USNC-URSI NRSM, Boulder, CO, USA, 2023. [Online]. Available: https://ieeexplore.ieee.org/document/10043164.
[56] W. Schweber, *Electronic Communication Systems: A complete Course*, 2 ed. New Jersey: Prentice-Hall, 1996.
[57] D. Cabric, D. Taubenheim, G. Cafaro, and R. Farrell, "Cognitive radio platforms and testbeds," *Cognitive Radio Communications and Networks,* pp. 539-585, 2010, doi: https://doi.org/10.1016/B978-0-12-374715-0.00019-8.

[58] P. Poshala, R. KK, and R. Gupta, "Signal Chain Noise Figure Analysis," Texas Instruments, 2014. [Online]. Available: https://www.ti.com/lit/an/slaa652/slaa652.pdf?ts=1616188452534&ref_url=https%253A%252F%252Fwww.google.com%252F
[59] A. PALAIOS, V. M. MITEVA, and P. MÄHÖNEN, "Contemporary Study of Radio Noise Characteristics in Diverse Environments," *IEEE Access,* vol. 6, 2018, doi: 10.1109/ACCESS.2017.2654064.
[60] D. J. Breton, C. E. Haedrich, M. J. Kamrath, and D. K. Wilson, "Street-scale mapping of urban radio frequency noise at very high frequency and ultra high frequency.," *Radio Science,* vol. 54, no. 11, pp. 934-948, 2019, doi: 10.1029/2019RS006893.
[61] M. Sorecau, E. Sorecau, A. Sarbu, and P. Bechet, "Real-Time Statistical Measurement of Wideband Signals Based on Software Defined Radio Technology," *Electronics,* vol. 12, no. 13, 2023, doi: https://doi.org/10.3390/electronics12132920.
[62] D. M. Pozar, *Microwave Engineering*, 4th ed. Wiley, 2012.
[63] R. J. Achatz and R. A. Dalke, "Man-Made Noise Power Measurements at VHF and UHF Frequencies," in "Report 02-390," U.S. DEPARTMENT OF COMMERCE, 2001.
[64] A. J. Wagstaff and N. Merricks, "Man-Made Noise Measurement Programme," MAAS Consultants, Cambridgeshire, UK, 2003.
[65] M.-H. Chang and K.-H. Lin, "A Comparative Investigation on Urban Radio Noise at Several Specific Measured Areas and Its Applications for Communications.," *IEEE Transactions on Broadcasting,* vol. 50, no. 3, 2004.
[66] R. Shiavi, *Introduction to Applied Statistical Signal Analysis: Guide to Biomedical and Electrical Engineering Applications*. Elsevier, 2007.
[67] D. E. Newland, *An Introduction to Random Variations, Spectral and Wavelet Analysis*. Longman.
[68] Roxy Peck, Chris Olsen, and J. L. Devore, *Introduction to Statistics and Data Analysis*, 4 ed. Richard Stratton, 2012.
[69] R. A. Witte, *Spectrum and Network Measurements*. Printice-Hall, 2001.
[70] P. Torio and M. G. Sanchez, "Novel Procedure to Determine Statistical Functions of Impulsive Noise," *IEEE Transactions on Electromagnetic Compatibility,* vol. 47, no. 3, 2005, doi: 10.1109/TEMC.2005.852219.
[71] J. E. Giesbrecht, D. Abbott, and R. Clarke, "An empirical study of the probability density function of HF noise," *Fluctuation and Noise Letters,* vol. 6, no. 2, p. 8, 2006, doi: 10.1142/S0219477506003203.
[72] J. M. Martinez, "Elementary Statistics: Measures of Central Tendency," Compton College, 2012 2012. [Online]. Available: http://www.compton.edu/facultystaff/jmmartinez/docs/Math-150-Spring-2015/Math-150-Ch3-LectureSlides.pdf
[73] J. Frost, "Statistics by Jim: Glossary," vol. 22019, ed, 2019.
[74] A. J. Wagstaff and N. Merricks, "Autonomas Interference Monitoring System," MAAS Consulting, 2007.
[75] H. Li, L. Liu, Y. Li, Z. Yuan, and K. Zhang, "Measurement and Characterization of Electromagnetic Noise in Edge Computing Networks for the Industrial Internet of Things.," *Sensors,* vol. 19, no. 3104, 2019, doi: 10.3390/s19143104.
[76] Rodger E. Ziemer and W. H. Tranter, *Priciples of Communications Systems, Modulation, and Noise*. John Wiley and Sons, 2010.
[77] K. Wiklundh, "Relation between the amplitude probability distribution of an interfering signal and its impact on digital radio receivers," *IEEE Transactions on Electromagnetic Compatibility,* vol. 48, no. 3, p. 7, 2006, doi: 10.1109/TEMC.2006.877782.
[78] A. N. Z. Abidin, W. R. W. Abdullah, M. Z. M. Jenu, and A.Ramli, "The Development of Statistical Parameter Measuring Analyser Using the APD Technique for DSL Communication," in *IEEE International RF and Microwave*, Kuala Lumpur, Malaysia, IEEE, Ed., 2008: IEEE.
[79] A. N. Z. Abidin, W. R. W. Abdullah, T. C. Chuah, M. Z. M. Jenu, and A.Ramli, "APD-based Measurement Technique to Estimate the Impact of Noise Emission from Electrical Appliances on Digital Communication Systems.," presented at the Electrical Design of Advanced Packaging and Systems Symposium, Seoul, South Korea, 2008.
[80] K. Gotoh, Y. Matsumoto, S. Ishigami, T. Shinozuka, and M. Uchino, "Development and evaluation of a prototype multichannel APD measuring receiver," presented at the International Symposium on Electromagnetic Compatibility, Honolulu, HI, USA, 2007.

[81] F. Alam, R. Mehmoodb, I. Katiba, and A. Albeshria, "Analysis of Eight Data Mining Algorithms for Smarter Internet of Things," *Procedia Computer Science,* vol. 98, pp. 437-442, 2016. [Online]. Available: https://www.sciencedirect.com/search?pub=Procedia%20Computer%20Science&cid=280203&volume=98&page=437-442&show=25&sortBy=relevance.

[82] P. Romero and K. Dighe, "Fast and Unsupervised Classification of Radio Frequency Data Sets Utilizing Machine Learning Algorithms," presented at the The Fourth International Conference on Data Analytics, Nice, France, 2015. [Online]. Available: https://www.thinkmind.org/download.php?articleid=data_analytics_2015_8_30_60114.

[83] M. S. Mahdavinejad, M. Rezvan, Mohammadamin, Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine Learning for Internet of Things Data Analysis: A Survey," *Digital Communications and Networks,* vol. 4, no. 3, pp. 161-175, 2017, doi: 10.1016/j.dcan.2017.10.002.

[84] I. P. S. Mary and L. Arockiam, "Aro-Nims: A New-Fangled Framework for IoT Data Pre-processing," *International Journal of Research in Advent Technology,* vol. 6, no. 5, 2018. [Online]. Available: http://www.ijrat.org/downloads/Vol-6/may-2018/paper%20ID-652018117.pdf.

[85] S. Faruque, *Radio Frequency Propagation Made Easy*. Switzerland: Springer, 2015.

[86] Y. Huang and K. Boyle, *Antennas: from Theory to Practice*. United Kingdom: Wiley, 2008.

[87] J. P. Uwiringiyimana, U. Khayam, Suwarno, and G. C. Montanari, "Design and Implementation of Ultra-Wide Band Antenna for Partial Discharge Detection in High Voltage Power Equipment," *IEEE Access,* vol. 10, 2022, doi: 10.1109/ACCESS.2022.3144416.

[88] M. Steer, *Microwave and RF Design Modules*, Third ed. 2019.

[89] Mini-Circuits, "Understanding noise parameter measurements," 2009. [Online]. Available: https://www.minicircuits.com/app/AN60-040.pdf

[90] X. Liu, G. Qiu, and F. Yu, "Experimental study of cooperative communication utilizing GNU Radio and USRP2," in *CECNet*, Yichang, China, 2012: IEEE, doi: 10.1109/CECNet.2012.6201638.

[91] D. G. K. Ingala, "Development of a Multi-frequency Interferometer Telescope for Radio Astronomy (MITRA)," Master of Engineering, Electronic Engineering, Durban University of Technology, Durban, 2015. [Online]. Available: https://openscholar.dut.ac.za/handle/10321/1383

[92] T. J. Mathews and Z. Han, "USRP2 Implementation of Compressive Sensing Based Channel Estimation in OFDM," Beirut, Lebanon, 2013: IEEE, doi: 10.1109/ICCITechnology.2013.6579527.

[93] M. Dardaillon, K. Marquet, T. Risset, and A. Scherrer, "Software Defined Radio Architecture Survey for Cognitive Testbeds," presented at the IWCMC, Limassol, Cyprus, 2012. [Online]. Available: https://ieeexplore.ieee.org/document/6314201.

[94] M. T. Masonta, M. Mzyece, and F. Mekuria, "A comparative study of cognitive radio platforms," *MEDES,* pp. 145-149, 2012, doi: 10.1145/2457276.2457301.

[95] J. Muslimin, A. L. Asnawi, A. F. Ismail, and A. Z. Jusoh, "SDR-Based Transceiver of Digital Communication System Using USRP and GNU Radio," in *ICCCE 2016*, Kuala Lumpur, Malaysia, 2016, vol. 1: IEEE, pp. 49-453, doi: 10.1109/ICCCE.2016.100. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7808358

[96] GNU_Radio. "InstallingGR." https://wiki.gnuradio.org/index.php/InstallingGR (accessed 27 Nov 2020, 2020).

[97] L. C. Tran, D. T. Nguyen, F. Safaei, and P. J. Vial, "An Experimental Study of OFDM in Software Defined Radio Systems Using GNU Platform and USRP2 Devices," in *ATC'14*, Hanoi, Vietnam, 2014: IEEE, doi: 10.1109/ATC.2014.7043470. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7043470

[98] Ettus-Research™. "USRP2 Knowledge Base." https://kb.ettus.com/USRP2 (accessed 26-Nov-2020, 2020).

[99] Ettus-Research™. "SBX Daughterboard Knowledge Base." https://kb.ettus.com/SBX (accessed 26-Nov-2020, 2020).

[100] Ettus-ResearchTM. "USRP Hardware Driver." https://www.ettus.com/sdr-software/uhd-usrp-hardware-driver/ (accessed 27 Nov 2020, 2020).

[101] I. A. T. Hashem, IbrarYaqoob, N. BadrulAnuar, SalimahMokhtar, b. Gania, and S. U. Khanb, "The rise of "big data" on cloud computing: Review and open research issues," *Information Systems,* vol. 47, pp. 98-115, January 2015 2015. [Online]. Available: https://doi.org/10.1016/j.is.2014.07.006.

[102] S. Madden, "From Databases to Big Data," *IEEE Internet Computing,* vol. 16, no. 3, pp. 4-6, May-June 2012 2012, doi: 10.1109/MIC.2012.50.

[103] A. Jacobs. (2009, August 2009) The pathologies of big data. *Communications of the ACM*. 36-44. Available: https://dl.acm.org/doi/pdf/10.1145/1536616.1536632
[104] MathWorks®. "Strategies for Efficient Use of Memory." https://www.mathworks.com/help/matlab/matlab_prog/strategies-for-efficient-use-of-memory.html (accessed 10 August 2021, 2021).
[105] MathWorks®. "Resolve "Out of Memory" Errors." https://www.mathworks.com/help/matlab/matlab_prog/resolving-out-of-memory-errors.html (accessed 10 August 2021, 2021).
[106] D. G. K. Ingala, N. Pillay, and A. Pillay, "An Assessment of Environmental RF Noise Due to IoT Deployment," *Sensors,* vol. 23, no. 18, 2023, doi: 10.3390/s23187899.
[107] S. W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing*, 2nd ed. New York, USA: California Technical Publishing, 1999.
[108] L. J. Cardinal, "Central tendency and variability in biological systems," *JCHIMP,* vol. 5, no. 3, 2015, doi: http://dx.doi.org/10.3402/jchimp.v5.27930.
[109] M. J. Blanca, J. Arnau, D. López-Montiel, R. Bono, and R. Bendayan, "Skewness and kurtosis in real data samples," *Methodology,* vol. 9, no. 2, pp. 78-84, 2013, doi: https://doi.org/10.1027/1614-2241/a000057.
[110] C. S. K. Dash, A. K. Behera, S. Dehuri, and A. Ghosh, "An outliers detection and elimination framework in classification task of data mining," *Decision Analytics,* vol. 6, no. 2023, 2023, doi: 10.1016/j.dajour.2023.100164.
[111] B. ILLOWSKY and S. DEAN, *Introductory Statistics*. Houston, USA: OpenStax, 2018.
[112] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014, p. 449.
[113] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Scienc,* vol. 2, no. 160, 2021, doi: https://doi.org/10.1007/s42979-021-00592-x.
[114] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2 ed. Springer, 2016.
[115] A. C. Müller and S. Guido, *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly, 2016.
[116] N. Ceder, *Exploring Data with Python*. Manning, 2018.
[117] P. Kokol, M. Kokol, and S. Zagoranski, "Machine learning on small size samples: A synthetic knowledge synthesis," *Science Progress,* vol. 105, no. 1, 2022, doi: https://doi.org/10.1177/00368504211029777.
[118] C. F. Caiafa, J. Solé-Casals, S. Zhe, and T. Tanaka, "Decomposition Methods for Machine Learning with Small, Incomplete or Noisy Datasets," *Applied Sciences,* vol. 10, no. 23, 2020, doi: https://doi.org/10.3390/app10238481.
[119] SciPy, "User Guide - Interpolation," SciPy, 2023. [Online]. Available: https://docs.scipy.org/doc/scipy/tutorial/interpolate.html
[120] P. Joshi, J. Hearty, B. Sjardin, L. Massaron, and A. Boschetti, *Python: Real World Machine Learning*. Birmingham - UK: Packt Publishing, 2016.
[121] V. Plevris, G. Solorzano, N. Bakas, and M. E. A. B. Seghier, "Investigation of performance metrics in regression analysis and machine learning-based prediction models," presented at the ECCOMAS, Oslo, Norway, 2022.