Analysis of TV White Space using a Spectrum Observatory

JPH Havenga

orcid.org 0000-0002-5028-3710

Dissertation submitted in fulfilment of the requirements for the degree *Master of Engineering* in *Computer and Electronic Engineering* at the North-West University

Supervisor: Dr M Ferreira

Graduation May 2018

Student number: 23442735

Declaration

I, Johannes Petrus Hendrik Havenga hereby declare that the dissertation entitled "Analysis of TV White Space using a Spectrum Observatory" is my own original work and has not already been submitted to any other university or institution for examination.

J.P.H. Havenga

Student number: 23442735

Signed on the 21th day of November 2017 at Potchefstroom.

Acknowledgements

I would like to acknowledge my Lord and Saviour Jesus Christ. To Him be the glory and praise forever. I am eternally grateful for this opportunity and the people He placed in my life during this time.

"*For from Him [all things originate] and through Him [all things live and exist] and to Him are all things [directed]. To Him be glory and honor forever! Amen.*" Rom 11:36 AMP

My study leader, Dr. M. Ferreira, thank you for your continuous help and motivation. Thank you for the opportunities for growth and travel you created for me. It was a privilege working with a supervisor who is involved and dedicated.

I would like to thank the following individuals and companies for giving me guidance and helping me to complete this dissertation:

- Dr. Braam Otto (SKA South Africa)
- Dr. Charles Crosby (CHPC)
- Telkom SA COE
- GEW Technologies

I would never have been able to do this without the support from my family. To my father Johan, my mother Mirrie and my brothers Okkert and Adriaan, I am extremely grateful for each one and love you. To my extended family, all the cousins, aunts and uncles, thank you for your love and support.

To my mentor and grandfather, Okkie Britz, thank you for all your prayers, continuous support and guidance.

My friends, Herman, Daneel and JJ, who went through this with me, thank you for everything!

Abstract

The concept of Dynamic Spectrum Access (DSA) allows the under-utilised spectrum to be made available for Secondary Users (SUs) on condition that such SUs do not cause harmful interference for the Primary User (PU). The first step in assigning the Radio Frequency (RF) spectrum to SUs, is to determine the PU spectrum usage, which can be done via spectrum sensing or by using a geolocation database. A geolocation database is currently the approach preferred by regulators across the world to deploy DSA in the TV bands. The aim of this study is to implement a Spectrum Observatory that also includes the building blocks for a geolocation database. The Spectrum Observatory is a research tool that contains a geolocation database, and that is used for DSA research. Central to both the geolocation database and the Spectrum Observatory is the computation node, which is responsible for performing propagation prediction over a large area.

Field strengths are computed for all PU transmitters to determine the geographic location of channels that might be available to SUs. Propagation prediction is computationally intensive and time-consuming, however, and thus it effectively becomes the limiting factor in extending the geolocation-based approach for DSA to other parts of the spectrum. A High Performance Cluster Computer (HPCC) implementation for the computation node is thus presented in this dissertation. Four propagation models are implemented and adapted for parallelisation. The performance analysis indicates computational speed-up for all propagation models on the HPCC, and factors affecting linear scalability are identified.

The success of implementing a Spectrum Observatory for DSA is directly dependent on the choice of propagation model used in the Spectrum Observatory for the propagation prediction. The above-mentioned HPCC was used to do analysis on the TV White Space (TVWS) availability, while using different propagation models. The Free-space Loss model, the Hata-Davidson model and the Irregular Terrain Model (ITM) model in both Area mode and Point-to-Point (P2P) mode were investigated.

The investigation involved comparing the amount of TVWS predicted for the country of South Africa, by using the different propagation models. The different propagation models were also verified against selected industry radio propagation software.

The TVWS research results showed that a considerable number of TVWS channels are indeed available for most of the propagation models. The terrain-aware ITM P2P model showed slightly more TVWS available for provinces with mountainous areas, thus showing its advantage over empirical models. The mean number of TVWS channels available for South Africa, according to the different propagation models, is 0.0 for the Free-space loss model, 30.4 for the Hata-Davidson model, 41.2 for the ITM Area model and 42.1 channels according to the ITM P2P model.

The favourable results from the TVWS analysis, especially when using a terrain-aware propagation model, and the speed-up obtained when using an HPCC as the computation node for the Spectrum Observatory strengthens the argument for geolocation-based DSA in the TV bands. By further improving the computation node and refining the propagation models, the concept of geolocation-based DSA in other parts of the RF spectrum can also be promoted.

Keywords: *Spectrum Management, TV White Space, Dynamic Spectrum Access, High Performance Computing*

Contents

List of Figures

List of Tables

List of Acronyms

AGL Above Ground Level

AMSL Above Mean Sea Level

CCDF Complementary Cumulative Distribution Function

CEPT European Conference of Postal and Telecommunications Administrations

CHPC Centre for High Performance Computing

C/N Carrier to Noise Ratio

COGEU COgnitive radio systems for efficient sharing of TV white spaces in EUropean context

CPU Central Processing Unit

CR Cognitive Radio

CSIR Council for Scientific and Industrial Research

CUDA Compute Unified Device Architecture

DEM Digital Elevation Model

DSA Dynamic Spectrum Access

DTT Digital Terrestrial TV

ECC Electronic Communications Committee

EU European Union

ERP Effective Radiated Power

FCC Federal Communications Commission

FDTD Finite-Difference Time-Domain

GLDB Geolocation Database

GLOBE Global Land One-km Base Elevation Project

GPGPU General Purpose Graphics Processing Unit

GPS Global Positioning System

GPU Graphics Processing Unit

HAAT Height Above Average Terrain

HPC High Performance Computing

HPCC High Performance Cluster Computer

ICASA Independent Communications Authority of South Africa

IaaS Infrastructure as a Service

ITM Irregular Terrain Model

ITS Institute for Telecommunication Sciences

ITU International Telecommunication Union

KML Keyhole Markup Language

MAE Mean Absolute Error

MCL Minimum Coupling Loss

MDTT Mobile Digital Terrestrial TV

MIMD Multiple instruction, multiple data

MPI Message Passing Interface

NASA National Aeronautics and Space Administration

NIST National Institute of Standards and Technology

NTIA National Telecommunications and Information Administration

Ofcom Office of Communication

OFDM Orthogonal Frequency-Division Multiplexing

P2P Point-to-Point

PaaS Platform as a Service

PAL Phase Alternating Line

PAWS Protocol to Access Whitespace

PMSE Program Making and Special Events

PU Primary User

RF Radio Frequency

RFI Radio Frequency Interference

RMSE Root Mean Square Error

SaaS Software as a Service

SATNAC Southern Africa Telecommunication Networks and Applications Conference

SEAMCAT Spectrum Engineering Advanced Monte Carlo Analysis Tool

SIMD Single instruction, multiple data

SKA Square Kilometer Array

SO Spectral Opportunity

SPLAT! Signal Propagation, Loss, And Terrain analysis tool

SRTM Shuttle Radar Topography Mission

SU Secondary User

TAU Tuning and Analysis Utilities

TVWS TV White Space

UHF Ultra High Frequency

UK United Kingdom

UN United Nations

USA United States of America

WS White Space

WSD White Space Device

Chapter 1

Introduction

In this chapter, the relevance of the research on TV White Space (TVWS) for South Africa is explained, and the usage of a Spectrum Observatory for research is motivated. This chapter also introduces the background to the research area, sets out the research questions and objectives, and presents the research methodology used. This chapter concludes with a discussion of the dissertation layout.

1.1 Background

In a spectrum-scarce world, every effort needs to be made to use the spectrum as efficiently as possible. One of the methods that have been proposed in [10] is to use Dynamic Spectrum Access (DSA) to enable secondary access to the spectrum. The primary concern for DSA is to protect the Primary User (PU) or licenced user. To avoid harm to the PU, two general approaches have been used, viz., firstly, Spectrum Sensing [11] and, secondly a geolocation database.

One of the first parts of the spectrum being investigated for DSA is that part of the spectrum initially assigned for terrestrial TV broadcasting. Digital Terrestrial TV (DTT) is being embraced by more and more countries as the existing analogue TV systems are being switched off systematically. One of the main reasons for switching to DTT is that DTT is more spectrum-efficient than analogue TV technologies, i.e. less bandwidth is needed to transmit the same amount of TV content, thus freeing up the spectrum. The unused spectrum in the part of the spectrum assigned to terrestrial TV is called TV White Space (TVWS). DSA by TVWS is being investigated in this study.

Spectrum sensing using only one radio for sensing is not adequate for reliable detection of PUs because of multipath effects, and thus co-operative sensing has been proposed [12]. Since reliable sensing equipment, with sufficiently high sensitivity and calibrated antennas, is expensive, it is not the preferred option for industry. This approach is also not currently preferred by regulators as a standalone DSA solution. Research is being done to enhance propagation prediction with spectrum sensing to deliver more accurate propagation prediction.

The alternative to Spectrum Sensing is an approach based on a geolocation database, and to this end the Spectrum Observatory is introduced. The Spectrum Observatory being used in this research is illustrated in Figure 1.1. It contains the components of a geolocation database and thus uses propagation prediction to make TVWS predictions. The initial purpose of the Spectrum Observatory is to create a tool that can be used by researchers to do faster TVWS research and spectrum management. The Spectrum Observatory also allows researchers to view spectrum availability, PU protection regions and other analytics.

The Spectrum Observatory consists of different parts, as seen in the Spectrum Observatory Architecture in Figure 1.1, and each part has a specific function. The coloured blocks in Figure 1.1 were developed and integrated through the course of this study. The Primary User Database unit contains the operational information of the PUs. The Propagation Models unit denotes the different propagation models that exist in the system and that can be used by the researcher.

In this research, the available propagation models (Free-Space Loss model, Hata-Davidson, Irregular Terrain Model (ITM) Area and ITM Point-to-Point (P2P)) models are implemented and investigated. The central part of the Spectrum Observatory is the Computation Node, which is responsible for the computation of the predictions; is a central focus of this study, as a High Performance Cluster Computer (HPCC) is proposed for the Computation Node. The Policies unit represents the different TVWS policies that can be used for TVWS classification; a basic TVWS policy is used in this research. The Analysis unit contains the tools created to help researchers interpret the TVWS data. The Geolocation Database (GLDB) and the Protocol to Access Whitespace (PAWS) units denote the geolocation database aspect of the Spectrum Observatory, with the PAWS unit indicating the PAWS interface that is used by TVWS devices.

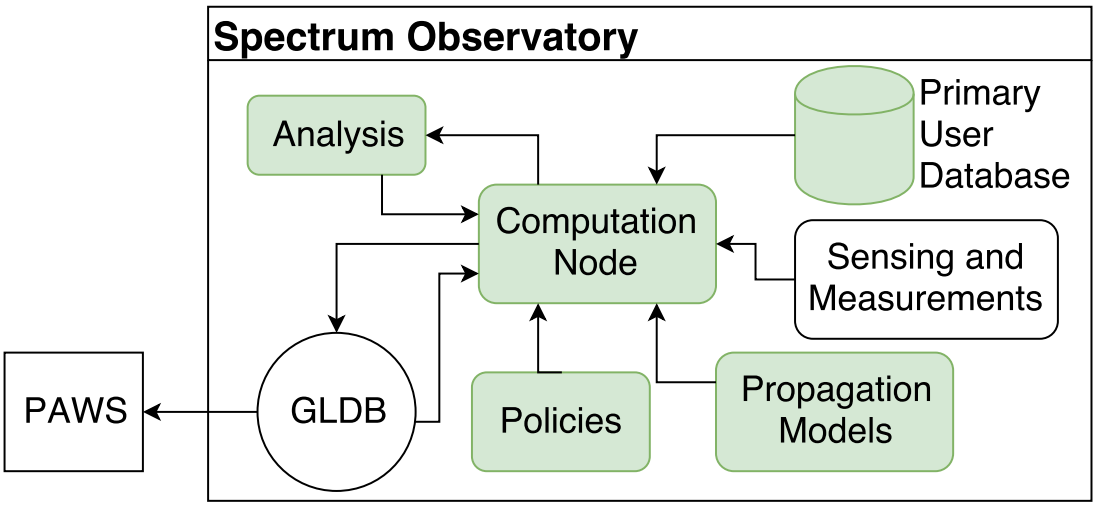


Figure 1.1: Spectrum Observatory Architecture

With a geolocation database based approach, like the Spectrum Observatory, propagation modelling is used to predict the field strengths from the PUs, and thereafter DSA decisions are made based on those predictions [1]. The choice of propagation model plays a crucial role in determining the available spectrum and in protecting the PUs. Currently, this approach has been employed in TVWS networks, in other words, DSA has been used in terrestrial TV networks.

The specifications of terrestrial TV transmitters are thus used as input for the propagation prediction and the calculated field strengths are stored in the geolocation database. The geolocation database uses these predicted propagation field strengths as a basis for determining the unused spectrum.

The core issue with propagation prediction is that it takes a long time to calculate field strengths: the Federal Communications Commission (FCC) in the United States of America (USA) requires the PUs in the geolocation database to be updated once a week [13], while the Office of Communication (Ofcom) in the United Kingdom (UK) supplies its geolocation database owners with new TVWS data every six months [14]. This is especially true for propagation models that take into account terrain data, and that is where the need for faster propagation prediction arises.

Some researchers would argue that the PU activity and hence the propagation prediction does not change significantly in the terrestrial TV part of the Radio Frequency (RF) spectrum (470 - 854 MHz), but the exploitation of TVWS for DSA is only a stepping stone for the use of DSA across all parts of the RF spectrum. The other consideration to take into account is that calculations regarding the effect of Secondary Users (SUs) on each other could be done more promptly, thus allowing detailed and fast co-existence planning. More computation power also means less downtime for the geolocation database, when PUs are being added or operational parameters change. Without more dynamic propagation prediction and secondary access determination, the idea of a geolocation-based approach to DSA is simply not feasible. SUs need rapid answers with regard to the available spectrum to avoid harm to PUs and to ensure optimum spectrum usage.

To implement DSA successfully the South African government should draw up policies that stipulate how DSA is to be used to protect the PU. In the USA, the TVWS policies have been drawn up by the FCC, and in the UK, these policies have been drawn up by Ofcom. South Africa is yet to release a clear plan on the usage of TVWS and DSA implementation. There are, however, draft regulations on the usage of TVWS for DSA in South Africa [15].

1.2 Motivation

Utilising the available spectrum more efficiently has become a necessity with the ever-growing need for more broadband connectivity. While much research is going into improving and making spectrum sensing equipment more affordable, the go-to approach for DSA, for the time-being, remains a geolocation-based one.

The propagation model used in the geolocation database to predict RF propagation plays a critical role in determining the amount of TVWS available. The primary constraint of DSA is the protection of PUs from harmful interference; the various propagation models are used to determine the signal strength of the PUs. The propagation model could be conservative in predicting the field strength, which would lead to less TVWS being available, or it could be more casual, which would make more TVWS available, but this might also cause more interference to the PUs. This study thus investigates the different propagation models used in the propagation prediction phase of the Spectrum Observatory, and evaluates the effect of the various propagation models on the amount of TVWS available.

The geolocation-based approach does require a significant amount of computing power during the propagation prediction part of TVWS determination. To reduce the time needed the use of a HPCC is proposed and the speed-up obtained from using such an HPCC as the computation node is quantified and analysed.

1.3 Research Questions

The research questions of the study are as follows:

1. What is the effect of the choice of propagation model on the amount of TVWS available?

2. How much computation-time can be saved by using an HPCC during the propagation prediction phase of the Spectrum Observatory?

1.4 Research Objectives

Objectives to be addressed and documented in the dissertation to answer the first research question are as follows:

- Create a repository for the primary transmitters as part of the Spectrum Observatory;
- Calculate the path loss and received powers using the HPCC, which is the chosen computation node of the Spectrum Observatory;
- Create a repository for the Spectrum Observatory that will contain relevant information with regard to the received powers at different frequencies and at different locations in the country, resulting from one TV transmitter, for all the propagation models;
- Create a repository for the Spectrum Observatory that will contain the maximum predicted received powers at different locations for the whole country aggregated according to Ultra High Frequency (UHF) channels, for all the propagation models;
- Create a repository for the Spectrum Observatory, which will contain the predicted received powers that are not used as the maximum signal, for the whole country also aggregated according to UHF channels for all the propagation models. This data can be used to do interference calculations in the future;
- Create a repository for the Spectrum Observatory that will contain the UHF channels available to be used as TVWS for analogue, DTT and Mobile Digital Terrestrial TV (MDTT) for the whole country and for all the propagation models;

- Analyse the TVWS availability for the different propagation models using the data generated by and stored in the Spectrum Observatory's repositories.

Objectives to be addressed to answer the second research question:

- Calculate TVWS availability using South Africa as a case study;
- Implement the propagation models for parallelisation on the HPCC platform;
- Quantify the improvement in computational speed-up by means of a computational performance analysis of the HPCC implementation relative to a quad-core desktop machine.

1.5 Research Method

In the field of research, different types of research exist. The type of research done in this study is quantitative research, because a significant amount of data was generated, from which certain conclusions were drawn. In this quantitative research study, firstly, the amount of time that can be saved when using an HPCC as the computation node for the Spectrum Observatory is quantified. Secondly, the amount of TVWS that is predicted to be available when using different propagation models is also quantified. Lastly, the research done can be broken down into various steps, viz., literature study, system model design, computational performance analysis, verification and validation and results analysis.

1.5.1 Study of Relevant Literature

The aim of this phase of the research is to review the relevant literature with regard to the particular area of research. The focus of the literature review in Chapter 2 thus starts by looking at the concept of DSA, why it is necessary, and how it is measured;

it also looks at the different policies used around the world to implement DSA. The second part of the literature study discusses the geolocation-based approach to DSA and how countries implement such a geolocation database. The geolocation database review is followed by a discussion of different RF wave propagation phenomena and a discussion of different propagation models. The literature study further includes a review of the different computing architectures that could be used as the computation node of the Spectrum Observatory. The literature study includes discussion of related work done by other researchers with regard to geolocation-based DSA, comparative TVWS studies and parallel computation solutions. The literature study in Chapter 2 is followed by a discussion of the System Model in Chapter 3.

1.5.2 System Model Design

Chapter 3 explains the System Model design that was used to conduct the comparative study with regard to the amount of TVWS available per area when using different propagation models on the Spectrum Observatory. All four of the propagation models (Free-Space, Hata-Davidson, ITM Area and ITM P2P models) that are currently in the Spectrum Observatory were parallelised to run on parallel hardware in order to decrease computation times. The path loss and received power for the analogue, DTT and MDTT transmitters in South Africa were calculated by using the four propagation models. The maximum field strengths for each terrestrial TV channel for the South Africa were then used with a basic TVWS policy to determine the TVWS for the country using the different propagation models.

1.5.3 Verification and Validation of Models

To ensure that the System Model was implemented correctly, verification was needed. Such verification was done by comparing the results from the proposed propagation model implementation against the results from existing industry software implementations. This verification is sufficient for both parts of the System Model as both the

computational, quantitative study on the speed-up gained from using an HPCC as the computation node and the comparative study on the amount of TVWS from different propagation models rely on the accuracy of the propagation models implemented. Validation involves ensuring that the proposed solution solves the problem at hand. The research done in this study relies heavily on propagation prediction using prediction models. Validation was thus done for this study by comparing the predicted field-strengths, which were then used to make TVWS predictions, against field measurements taken for defined pairs of transmitters and receivers. Verification and validation were discussed in Chapter 4.

1.5.4 Computational Performance Analysis

The computational performance analysis in Chapter 5 discusses the system that was designed to quantify the speed-up obtained by using an HPCC as the computation node for the Spectrum Observatory. The four propagation models (Free-Space, Hata-Davidson, ITM Area and ITM P2P models) were used to calculate the received power for 2 325 transmitters for the area of Gauteng. The calculations were done with a local multi-core PC and with a High Performance Cluster Computer (HPCC), with the latency of using both the architectures being recorded. The recorded latency was then used to calculate the speed-up obtained when using the HPCC. The speed-up was then analysed to determine the advantage of using the HPCC as the computation node for the Spectrum Observatory.

1.5.5 Analysis of TVWS Results

Following the Verification and the Validation of the System Model, in Chapter 4, the results generated by the System Model, discussed in Chapter 3, were analysed in Chapter 6. The analysis of the results displays and discusses the TVWS available per province for the different propagation models investigated, as well as the TVWS channels available for the country using the different propagation models. The TVWS availability per

area is analysed by using Complementary Cumulative Distribution Function (CCDF) plots as well as the mean number of TVWS channels available per area.

1.6 Spectrum Observatory Deliverables

This research work will deliver the following deliverables as part of the Spectrum Observatory:

1. A custom binary file containing the predicted received power for every PU in the PU database (.fst).
2. Custom binary files containing the maximum predicted received power per channel for the country of South Africa (.fsc).
3. A custom binary file containing information about the location of TVWS for every illumination technology (.wsc).

The specification documents for these custom binary files are shown in Appendix A.

1.7 Dissertation Overview

The rest of the dissertation will be structured in the following manner. A literature study in Chapter 2 will explain some of the concepts explored in this dissertation. This is followed in Chapter 3 by a detailed explanation of the system model, which forms the foundation of the Spectrum Observatory. The verification and validation of the implementations used in the study can be found in Chapter 4. The Computational Performance Analysis, in which the speed-up when using an HPCC as the computation node is determined, is found in Chapter 5, which is then followed by an analysis of the TVWS results in Chapter 6. The dissertation is wrapped up with a conclusion in Chapter 7.

Chapter 2

Literature Study

The Literature Study discusses the fundamental concepts used in the study. It thus contains information about TVWS, geolocation databases, propagation models and parallel computing. This chapter concludes with a related work section which discusses relevant work done by other researchers in the field.

2.1 Introduction

A thorough review of the literature is needed to comprehend the area of study and to obtain a precise definition of the different propagation models to be compared and used on the Spectrum Observatory. The literature study includes a review of DSA, and secondary access policies, as well as geolocation databases and their preferred usage by different regulators. A brief overview of common radio wave propagation phenomena is followed by detailed descriptions of the four propagation models (Free-Space loss, Hata-Davidson, ITM Area and ITM P2P) used in the Spectrum Observatory

and the TVWS analysis. The last aspect studied in the literature is the different computing architectures; this includes a comparison of both parallel hardware and software architectures. A related work review is also included in the literature review, which discusses aspects of this research that have already been investigated by other researchers.

2.2 Spectral Opportunity

Spectral Opportunity (SO) is a term that refers to the bandwidth that is free to be used by opportunistic wireless devices and that provides an opportunity for devices to connect. The term White Space (WS) can also be used in reference to SO. The exact definition of when a specific part of the spectrum can be seen as SO or WS is dependent on the policy used to classify the part of spectrum under consideration; more information on this topic is presented in Section 2.5.

As stated above, the widest known SO or WS is TVWS. The term TVWS refers to WS in that part of the RF spectrum that was initially assigned to TV broadcasting networks. The specific bands differ from country to country. According to the broadcast plan drawn up by the Independent Communications Authority of South Africa (ICASA) and published in 2013, the frequencies used for TV broadcasting in South Africa range from 470 to 854 MHz [16]. This range of frequencies translates to UHF channels 21 to 68, with a channel bandwidth of 8 MHz.

Allowing a Secondary User (SU) access to the part of spectrum that has already been assigned to a Primary User (PU) is called Dynamic Spectrum Access (DSA). The requirement of DSA is that the PU will be guaranteed protection against harmful interference from the SU. Harmful in this context implies interference levels that adversely affect the service grade of the PU [17]. The FCC and Ofcom in the USA and UK respectively already have policies in place with regard to TVWS classification and usage, while ICASA has thus far only released a draft on TVWS usage for South Africa.

South Africa is still busy with an analogue switch-off, which means that more unused spectrum will become available after the switch-over to digital. DSA is a proposed use for Digital Dividend where the Digital Dividend refers to part of the spectrum that is freed up after switching from analogue services to digital services.

2.3 Measurement of SO

To ensure that the PU does not experience harmful interference from the SUs, the SUs may not transmit with a transmitting power near the PUs in a manner that would cause notable degradation of the PUs quality of service. The fact that different primary transmitters transmit at different channels or frequencies at different times for different locations means that the SO is time, frequency, and location dependent for each RF channel [18].

The SUs can use one of two methods to determine if there is spectrum available for re-use. The first method is spectrum sensing, in which the SU can sense the spectrum and scan for activity from PUs. According to [11], the best way of sensing for PUs or incumbents is for multiple devices to sense together and share data with each other, as this will increase the probability of detecting a PU [12]. Collective sensing is, however, a costly solution since the radios used are expensive, and a large number of them are needed.

The second method is to use a geolocation database that contains information regarding the location of primary transmitters, the antenna height, transmit power, antenna polarisation and transmitting frequency, along with a propagation model that models the signal propagation from the primary transmitters. The article of [19] explains in detail the process of determining TVWS capacity using a geolocation database.

2.4 The role of Cognitive Radios

A widely accepted solution is to use Cognitive Radio (CR) for spectrum sensing and sharing. These radios are "cognitive" suggesting that they have cognitive functions and can, therefore, think for themselves [20]. The definition for CR, given by [21], states that cognitive radios are radios that understand the context in which they operate and as a result can change the communication process in line with that understanding.

In the context of DSA, the CR would be "aware" of the context in which they operate by continually being in contact with the geolocation database and thus aware which frequencies are available and at what power to transmit. The CR could also be aware of its environment by continuously sensing the spectrum to determine if there is any SO available at the current location and time. It has been shown in [12], that using multiple CRs together yields better sensing accuracy.

Using both the CR and the geolocation database would mean that the CR would be working with the geolocation database by sending the Global Positioning System (GPS) location of the CR to get information regarding the SO at the current location. The CR would then use this information to set up a connection on the available TVWS channels.

The idea of this approach is to promote agile spectrum transmission by using the CR. The CR can thus be deployed in two different scenarios. In the first scenario, the CR is used as a stand-alone device, which senses and utilises the spectrum based on the available spectrum. In the second scenario, the CR remains in contact with a geolocation database and provides the geolocation database with spectrum sensing information. This spectrum information is then used by the geolocation database to aid the prediction of WS. And lastly, the CR is able to utilise the available WS.

2.5 Secondary Access Policies

2.5.1 International Trends

The International Telecommunication Union (ITU) is an agency of the United Nations (UN) assigned to regulate information and communication technologies. The ITU is the global coordinator of spectrum usage and sets out standards, regulations and recommendations that allow the co-existence of national regulators with each other. Their regulations and recommendations, as set out for DTT, MDTT, and analogue broadcasting, are used as the basis for many TVWS policies.

2.5.2 USA

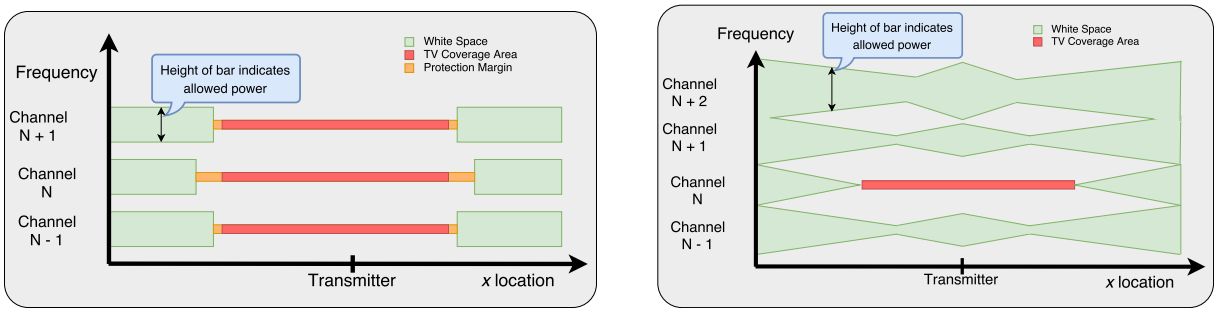
The most well-known governing spectrum regulator is the FCC, which regulates the usage of spectrum in the USA. The FCC's most recent publication regarding unlicensed operations in the television band was released in August 2015 [22]. In [22] the FCC explains the rules that fixed WS devices and portable WS devices must comply with, to ensure protection for the PU while still allowing good throughput for the WS devices. The initial rules made by the FCC were stringent, but over time, some rules were relaxed to enable more WS devices to use the spectrum, while still protecting the PUs. A graphical representation of the rules set out by the FCC can be seen in Figure 2.1. The figure shows the FCC rules with regard to SU power usage and the distance from the PU or transmitter. The allowed power for the SU is fixed, as seen in Figure 2.1. The co-channel and adjacent channel protection regions are different, but both only allow a fixed amount of power to be used by the SU when accessing TVWS.

2.5.3 Europe

There is much of activity in Europe on the European Union (EU) level via the European Conference of Postal and Telecommunications Administrations (CEPT) and the Electronic Communications Committee (ECC), but the UK regulator, Ofcom was first with its publication of regulations for using TVWS for DSA. The latest guideline on how TVWS should be implemented in the UK was published by Ofcom in February 2015 [23]. The approach by Ofcom is to have certified TVWS geolocation databases across the country that can be used by SUs to access TVWS.

However, the publications of Ofcom still have to comply with the regulations set out by CEPT. CEPT is the European Conference of Postal and Telecommunications Administrations, where Ofcom represents the UK in the European context. A graphical representation of the rules suggested by the SE43 workgroup of CEPT is seen in Figure 2.1.

It can be seen that the CEPT regulations for TVWS will allow for more TVWS usage when compared against the rules set out by the FCC. The main difference between the two approaches is that CEPT assigns variable power to the SU as a function of the distance between the SU and the PU. The FCC, in contrast, assigns a fixed allowed power to the SU, as mentioned earlier.



(a) FCC Rules

(b) CEPT Rules

Figure 2.1: FCC and CEPT Rules [1]

2.6 Geolocation Database

The geolocation database is one of the two SO measurement approaches as mentioned in Section 2.3. A geolocation database is a database containing information regarding spectrum usage and spectrum availability within a geographical area. The word geolocation is short for geographical location, and a geolocation database thus refers to a database containing information about something that is ordered according to its location on the planet. In the context of TVWS, that "something" is spectrum availability. To determine the available spectrum for the geolocation database, the strength of incumbent TV transmitters must first be determined. In the case of a geolocation database, this is done by using propagation models. More information regarding the propagation models is presented in Section 2.9.

The geolocation database becomes the central information source for the TVWS devices, which query the geolocation database to determine the available channels, and the geolocation database in turn gives them the information needed to connect without causing harmful interference to the PU.

In the Spectrum Observatory architecture seen in Figure 1.1, the geolocation database contains the PU database, the propagation models, the TVWS policies, the computation node, the results of the TVWS prediction, and the interface with which devices connect to the geolocation database. The Spectrum Observatory is at its core a geolocation database, whose purpose is to aid TVWS research by allowing comparisons of computing architectures and TVWS policies, as well as generating analytics and metrics of spectrum utilisation.

Some examples of Spectrum Observatories are the ones created by Microsoft [24] and by the University of Washington [25] in Seattle to advance research of TVWS in USA.

A South African Geolocation Spectrum Database & DSA coexistence manager was created by the Meraka Institute at the Council for Scientific and Industrial Research (CSIR) as the country's first DSA coexistence manager [26]. The proposed Spectrum Observatory being created will aid in understanding and providing analytics on the utilisation of the spectrum in South Africa.

The secondary access policies determine when the spectrum can be classified as "available" in the geolocation database. The geolocation implements the WS policy on the predicted field strengths to determine WS availability.

2.7 Existing Geolocation Database Usage and Rules

There are a couple of different regulatory bodies from different countries who have embraced the usage of TVWS for DSA using a geolocation database. The regulatory bodies have already set out requirements and recommendations on how this should be implemented. There are also different ways to implement TVWS using a geolocation database and not all regulatory bodies follow the same approach.

2.7.1 FCC

The Federal Communications Commission (FCC) has strict regulations that the commercial TVWS geo-location databases should follow. According to the FCC, there is specific operational information that needs to be included in the geo-location database. The information, found in [13], lists the broadcasting service for which protection should be sought, when spectrum is classified as TVWS, how this TVWS should be used, as well as the frequency and manner in which TVWS devices should query the geolocation database. The FCC uses a vectorised Minimum Coupling Loss (MCL)-based approach to determine the availability of TVWS channels using fixed rules for separation distance vectors [27].

2.7.2 COGEU

The COgnitive radio systems for efficient sharing of TV white spaces in EUropean context (COGEU) is a project funded by the European Commission. The purpose of the COGEU is to take advantage of the TV digital switch-over by developing cognitive radio systems that can use TVWS on a secondary access basis. The COGEU group has set out a few documents explaining their approach to different aspects of TVWS implementation. In one of their documents, they explain the information required for the geo-location database of the primary incumbents. The document explaining the requirements for the geo-location database can be seen in [28].

2.7.3 ECC

The Electronic Communications Committee (ECC) published a document in 2013 on the technical and operational requirements for the operation of White Space Device (WSD) under geo-location approach [29]. This document contains the parameters that need to be registered for Program Making and Special Events (PMSE). In South Africa the PMSE is not yet classified and therefore this information will not be discussed at this time. The ECC makes no particular requirement for the parameters needed by the geolocation database for the primary incumbents. The document does discuss the information that needs to be provided by a WSD to the geolocation database and also the information from the geolocation database to the WSD. The information from the geolocation database to the WSD can be taken into consideration when designing the geo-location database.

2.7.4 Ofcom

Office of Communication (Ofcom) has a different approach to calculating TVWS, as discussed in Section 2.5.3. Ofcom provides the database manager with the predicted signal levels, which implies that the database will not need to perform propagation

modelling [30]. The geolocation database, must, however provide the SU with the power at which the SU may operate and for what duration. To calculate the availability of TVWS channels, Ofcom uses a statistical approach that utilises a Monte Carlo simulation methodology to determine the degradation probability of a TV receiver in a pixel in a grid [27].

2.8 Radio Wave Propagation Effects

Electromagnetic signals or RF signals lose power when travelling over a distance, even in free space. The path loss is the attenuation that the transmitted RF signals undergo when travelling to the receiver. Many different phenomena can affect the RF signal propagation; the most significant effects will briefly be discussed.

2.8.1 Fading

Fading is the variation in the RF signal amplitude, which can be caused by variation in the distance between the two terminals, changes in the propagation environment, multiple signal paths and the relative motion between the two terminals. Fading can also be caused by obstacles entering the path between the transmitter and receiver. The source of fading that causes much trouble for spectrum engineers is multipath fading, also known as Raleigh Fading. This occurs because the transmitted signal can take more than one physical route to the receiver. This results in multiple signals arriving at the receiver at different times with the amplitudes differing due to different path lengths [31].

2.8.2 Reflection

The reflection of an RF signal from a surface is one of the most well known propagation phenomena [32]. The reflection coefficient of a surface is indicative of how much of

the RF signal is reflected and how much is absorbed by the surface and is used when doing reflection calculations. Finite surface reflection, when there is a gap in the surface between the terminals, is more complicated to model than infinite surface reflection, but it can be modelled as infinite surfaces if the reflection surface is big enough.

2.8.3 Scattering

RF signals that hit a rough surface do not just reflect off the surface, but they also tend to scatter. The scatter effect leads to much more attenuation than a reflection from a smooth surface, even if the reflection coefficient of the surface is relatively high [32]. The ITM model, discussed in Section 2.9.3, also accounts for scattering.

2.8.4 Diffraction

One of the effects of an obstacle between a transmitter and receiver is diffraction. Diffraction is the mechanism whereby a RF signal enters into the shadow of the obstacle [32]. To model diffraction mathematically, researchers have created a model called knife-edge diffraction, where the obstacle is modelled as having a very sharp, knife-like tip, to model the diffraction of the RF signal. Propagation models such as the ITM factor in the effects of such diffraction.

2.9 Propagation Models

Various propagation models are used for predicting field strengths resulting from the transmitter activity. There are numerous propagation models, which can be divided into two main types, namely deterministic models, like the Finite-Difference Time-Domain (FDTD), and empirical/statistical models, like the ITM [4]. A number of standard models are mentioned and classified in Table 2.1.

Propagation models used for TVWS planning are usually outdoor empirical propagation models. The FCC has developed a set of curves, which are used for propagation prediction similar to the ITU-R P.1546, while Ofcom makes use of an adapted Hata model; neither of these entirely takes the terrain between transmitter and receiver into account. Using a terrain-aware model like the ITM P2P model would lead to more accurate results.

Table 2.1: Common Propagation Models [3], [4]

Empirical/Statistical		**Deterministic**
Outdoor	**Indoor**	
Okumura	Partition Losses	FDTD
Hata	Log-Distance Path Loss	Moment-Method
COST-231	Ericsson MBM	Artificial Neural Network
Dual-Slope	Attenuation Factor	
ITM		
ITU-R P.1546		

The propagation models used in this dissertation are discussed in more detail.

2.9.1 Free-space Loss Model

The most basic propagation prediction can be made with the free-space loss formula. This was chosen because it is relatively easy to implement and because it provides a good baseline of worst-case propagation from an interference perspective. Propagation of a RF signal in 'free-space', means propagation in a vacuum. In the implemented instance, 'free-space' refers to line-of-sight through the atmosphere, meaning that there are no losses due to refraction, scattering or diffraction [32]. The free-space loss formula takes into account the frequency of transmission and the distance from the PU. Free-space loss is also defined in ITU-R P.525-3, and is thus referred to by some as the ITU-R P.525 model [33]. The formula for calculating path loss(PL_{FS}), from [31], where f is the frequency in MHz and d is the distance in km is:

$$PL_{FS}(dB) = 32.4 + 20\log(f) + 20\log(d) \tag{2.1}$$

2.9.2 Hata Model

The Hata model, also known as the Okumura-Hata model, is a set of equations derived from the Okumura outdoor propagation model, which is defined in [34]. Okumura developed a set of curves giving the median attenuation, or path loss, relative to free-space loss, based on drive measurements taken over a period in Japan [35]. The Hata model is a formulation of the curves drawn up by Okumura [5].

The Hata-Davidson model is considered a simple propagation model used by a number of researchers. The most prominent disadvantage of the original Hata model is that it neglects the terrain profile between the transmitter and the receiver when making predictions [35]. Nonetheless, the Hata model was chosen because it has been used by numerous researchers before.

The Hata model is only defined for link distances up to 100 km, as seen in Table 2.2 [36], meaning that it cannot be used for large-scale area prediction. Instead of using the Hata model as is, an extension of the model can be used. This propagation model is called the Hata-Davidson propagation model, defined in [36] and it extends the Hata model to distances of 300 km, frequencies of up to 2000 MHz and transmitter antenna heights of up to 2500m [6]. The complete list of parameters for the Hata-Davidson model can be seen in Table 2.3.

Table 2.2: Range of Parameters for Hata model [5]

Parameter	**Valid Range**
Frequency	150 - 1 500 MHz
Distance	1 - 100 km
Transmitter Antenna Height	30 - 200 m
Receiver Antenna Height	1 - 10 m

Table 2.3: Range of Parameters for Hata-Davidson model [6]

Parameter	**Valid Range**
Frequency	30 - 1 500 MHz
Distance	1 - 300 km
Transmitter Antenna Height	20 - 2 500 m
Receiver Antenna Height	1 - 10 m

The standard formula for medium path loss in an urban environment using the Hata model is

$$\begin{aligned} L_{50}(urban)(dB) = 69.55 + 26.16\log(f_{MHz}) - 13.82\log(h_{te}) - a(h_{re}) \\ + (44.9 - 6.55\log(h_{te}))\log(d_{km}) \end{aligned} \tag{2.2}$$

where $a(h_{re})$ is the mobile antenna correction factor with h_{re} being the receiver height in meter. The h_{te} is the transmitter antenna height AGL in meter. For small to medium-sized cities the mobile antenna correction factor is defined as

$$a(h_{re}) = (1.1\log(f_{MHz}) - 0.7)h_{re} - (1.56\log(f_{MHz}) - 0.8) \tag{2.3}$$

and for a large city, it is defined as

$$a(h_{re}) = 8.29(\log(1.54h_{re})^2) - 1.1 \text{ for } f_{MHz} \leq 300MHz \tag{2.4}$$

$$a(h_{re}) = 3.2(\log 11.75h_{re})^2 - 4.97 \text{ for } f_{MHz} > 300MHz \tag{2.5}$$

For suburban areas, the equation is adjusted to look as

$$L_{50}(dB) = L_{50}(urban) - 2[\log(f_{MHz}/28)]^2 - 5.4 \tag{2.6}$$

and for open or rural areas, this adjusts to

$$L_{50}(dB) = L_{50}(urban) - 4.78(\log(f_{MHz})^2) + 18.33\log(f_{MHz}) - 40.94 \tag{2.7}$$

The formulas for the Hata-Davidson as from [6] is

$$\begin{aligned} PL_{HD} = PL_{Hata} + A(h_{te}, d_{km}) - S_1(d_{km}) - S_2(h_{te}, d_{km}) \\ - S_3(f_{MHz}) - S_4(f_{MHz}, d_{km}) \end{aligned} \tag{2.8}$$

where

$$PL_{Hata} = L_{50}(urban)(dB) \tag{2.9}$$

$$A(h_{te}, d_{km}) = 0.62137(d_{km} - 20)[0.5 + 0.15\log_{10}(h_{te}/121.92)] \quad \text{for } 20 \geq d_{km} < 64.38 \tag{2.10}$$

$$S_1(d_{km}) = 0.174(d_{km} - 64.38) \text{ for } 64.38 \geq d_{km} < 300 \tag{2.11}$$

$$S_2(h_{te}, d_{km}) = 0.00784|\log_{10}(9.98/d_{km}|(h_{te} - 300) \quad \text{for } h_{te} > 300m \tag{2.12}$$

$$S_3(f_{MHz}) = f_{MHz}/250\log_{10}(1500/f_{MHz}) \tag{2.13}$$

$$S_4(f_{MHz}, d_{km}) = [0.112\log_{10}(1500/f_{MHz})] \times (d_{km} - 64.38) \text{ for } d_{km} > 64.38km \tag{2.14}$$

$$\tag{2.15}$$

For all other values, S_1, S_2, S_4 and $A = 0$

In the above equations, the parameters are:

f_{MHz}	-	Carrier frequency in MHz
d_{km}	-	Distance from transmitter in km
h_{te}	-	Transmitter effective height in m
h_{re}	-	Receiver effective height in m

Antenna Heights

The transmitter antenna height(h_{te}) is defined as AGL for the original Hata model, as the original measurements done by Okumura [34] were done for a flat urban area [37]. The Hata-Davidson extension also added the much needed Height Above Average Terrain (HAAT) factor [6].

There are mixed interpretations of what exactly is meant by effective base (or transmitter) antenna height [37] in propagation models. This is no different for the Hata model; no precise definition for the effective base antenna height for the Hata-Davidson could be found, as the original article defining the Hata-Davidson extension is not easily ac-

cessible. The Hata-Davidson has an effective base antenna height restriction, in that it needs to be between 20 and 2500 m, which leads to the assumption that HAAT is being implied. According to the authors of [6], one of the prominent correction factors that the Hata-Davidson extensions added is the HAAT factor.

Hata Area Correction Factors

The Hata equations were set out initially for an urban area over quasi-smooth terrain and added correction factors for other area types [5]. The types of area catered for in the Hata model, and consequently the Hata-Davidson model are medium-small cities (urban), large cities (urban), suburban areas and open/rural areas.

2.9.3 ITM

The ITM, also known as the Longley-Rice propagation model, is a propagation model defined in [7]. The model was refined, and the version used most often can be found in [38]. The ITM is a propagation model that takes the path geometry of the terrain into account, as well as the refractivity of the troposphere.

The Longley-Rice model was initially defined for the parameters seen in Table 2.4.

Table 2.4: Range of Parameters for ITM [7]

Parameter	**Range**
Frequency	20 - 40 000 MHz
Distance	1 - 2 000 km
Antenna Heights	0.5 - 3 000 m
Surface Refractivity	250 - 400 N-Units

The ITM consists of two modes: the first is used for point to area path loss prediction and the second is used for P2P path loss prediction. The two modes can be viewed as two different propagation models, with the point-to-area model being called the ITM Area model and the P2P mode being called the ITM P2P model.

In the ITM Area model, an irregularity parameter is used to model in the effect of the terrain between the transmitter and the receiver, whereas in the ITM P2P model, a Digital Elevation Model (DEM) is utilised.

The first of two modes available in the ITM is the area mode, where the surface refractivity, the climate of the region, ground constant and terrain irregularity parameters are all used to calculate the median transmission loss. The calculation is done at a distance from a transmitting antenna at a particular height for a receiving antenna of a certain height.

The second method is the P2P method. The P2P method does not use a terrain irregularity parameter but instead uses a path profile between the transmitting and receiving points which are extracted from a DEM.

Many parameters need to be set when using the ITM. These parameters are discussed below:

Environmental Parameters

The first couple of variables that are discussed are all variables that explain and compensate for an aspect of the physical environment in which the propagation prediction is being calculated.

Surface Refractivity The surface refractivity is the parameter that accounts for atmospheric conditions, such as climate and weather, which play an essential part in predicting the strength and fading properties of the tropospheric signals. The surface refractivity indicates the amount a RF ray is bent, or refracted, as it passes through the atmosphere [8]. Work done in [39], determined the median value for the median surface refractivity (N_0) for South Africa to be 332.9943 N-units. This value for surface refractivity is also used in this study.

Polarisation This refers to the polarisation of the transmitted electromagnetic wave or RF signal. Electromagnetic signals in the UHF band are usually either vertically or horizontally polarised. The polarisation used for each transmitter is determined by the data used from ICASA. It is assumed in the predictions that the transmitters and the receivers have the same polarisation.

Electric Ground Constants *Ground Conductivity* indicates the electrical conductivity of the ground, in Siemens per Meter, over which the signal is propagated. The first of two ground constants, the ground conductivity is the constant that has a greater effect on the signal propagation at frequencies below 50 MHz. The ground conductivity is set to 0.005, as this is the default recommended by the authors of the ITM model [8].

The *Relative Permittivity* of the ground, also known as the dielectric constant, indicates the ability of the ground to store electrical energy. The default value to be used as suggested is 15 [8]. The ground constants along with the polarisation of the signal only have an effect on the signal propagation when the signal is grazing the ground, or nearly so. The suggested values for the ground constants can be seen in Table 2.5.

Table 2.5: Suggested Values for Electrical Ground Constants [8]

	Relative Permittivity	**Conductivity (S/m)**
Average Ground	15	0.005
Poor Ground	4	0.001
Good Ground	25	0.020
Fresh Water	81	0.010
Sea Water	81	5.0

Radio Climate The radio climate is described by a set of discrete labels. The radio climate is used together with the Surface Refractivity to characterise the atmosphere and its variability in time. Different radio climates with suggested values for the surface refractivity can be seen in Table 2.6. The ITM area mode is intended for use over irregular terrain, and the preference of the authors of the ITM mode is to use the Continental Temperate climate unless there are clear indications to choose another climate [8].

Therefore, for this implementation, the radio climate is chosen as Continental Temperate.

Table 2.6: Radio Climates [8]

Radio Climate	**Ns (N_0)**
Equatorial (Congo)	360
Continental Subtropical (Sudan)	320
Maritime Subtropical (West Coast of Africa)	370
Desert (Sahara)	280
Continental Temperate	301
Maritime Temperate, over land (UK)	320
Maritime Temperate, over sea	350

Terrain Irregularity Parameter The ITM area mode does not make use of a DEM file as is the case with the P2P mode of the ITM. Instead, the ITM Area mode makes use of a Terrain Irregularity Parameter. The examination of a large number of terrain profiles of different lengths in a given area showed that median values of $\Delta h(d)$ increase with path length to an asymptotic value Δh, which is then used to characterise the terrain [8]. The world wide average for Δh is 90 m [8].

Siting Criteria This input parameter of ITM in Area mode describes the care taken at each antenna to ensure proper propagation conditions. The siting of the antenna, transmitting and/or receiving, can be classified as either **very good**, **good** or **random**. The siting criteria affect the assumed effective antenna height of that terminal within the ITM area calculations. The **very good** classification is used when the terminal system is sited on high ground and an effort was made to locate the antenna for maximum coverage. In the case of the terminal site not being on a hilltop, but where it is none-theless elevated for coverage, the siting is **good**. When the siting of the terminal is neither good nor poor, the siting is classified as **random**. In the use case of this study, the TV transmitters were placed on a **very good** siting, and the TV receivers were placed on a **random** siting. This is encouraged by the fact that the transmitters are installed by engineers, whereas the receivers are more likely to be installed by home users, who are most likely inexperienced, leading to bad installations.

Variability

The need for variability comes from the fact that even measured signals vary from observation to observation. This variation is caused by many factors, some of which are not accounted for in the propagation model. The variation in taking measurements can occur because of the equipment used and because of how, where and when the measurements were taken. In the implementation of the ITM propagation model, each mode of implementation has different variability parameters. Variability takes into account random, or somewhat unaccounted for, factors that influence the field strength resulting from the transmitters. There are different kinds of variability implemented in the models. The different variabilities for the ITM area mode are:

Time Variability The time variability indicates how the measured field strength will vary over time. In the case of hourly measurements being done over a particular time of the year, those measurements will differ. The parameter can be interpreted as "On this path for 95% of the time the attenuation did not exceed 32.6 dB".

To get a median value of the path loss, the time variability was set to 50%, meaning that the attenuation from the propagation model will not exceed the result for 50% of the time. The ITU-RRC-06 recommends 50% time variability when predicting wanted field strengths [9].

Location Variability The location variability refers to the long-term time changes from using different paths. In the point-to-point mode, a single, defined communication link is used, and a single isolated path is used, leading to the location variability being irrelevant. For a time variability of say 70% and a location variability of 60%, it can be said that this means, "For 60% of the path locations, the attenuation does not exceed 32.6 dB for 70% of the time". In the area mode, a location probability of 50% is used according to the recommendations from ITU-RRC-06 when predicting wanted field strengths [9].

Situation Variability The situation variability is analogous to using different equipment for measuring the field strength. So a situation can be seen as a particular deployment of equipment. It can be said that, for example, a situation variability of 40%, a location variability of 60% and a time variability of 70% with a resulting attenuation of 36.2 dB means that for 40% of situations at 60% of the path locations the attenuation does not exceed 36.2 dB for 70% of the time. The median situation variability was used to calculate the area mode prediction and was therefore set to 50%.

Variability Modes The four variability modes are the different ways in which these three variability parameters can be used in combination [8]. In the **Individual Mode**, the situation and the location variabilities are combined and used along with the time variability. When the location and time variables are combined, it is called the **Mobile Mode**. In a typical use case, a mobile system is employed with a single base station. All three variability parameters are combined in the **Single-Message Mode**, leading to one variable for variability. An example of this mode is a disaster warning system. In the **Broadcast Mode**, all three parameters are used separately, with the typical user being a broadcaster. For the setup of this study, the **Broadcast Mode** is used.

Point-to-point mode variability The point-to-point mode for the ITM has two variability parameters called "reliability" and "confidence". The statement, "At a range of 5 km from the transmitter, the probability is 90% that the path loss is less than 130 dB" indicates a reliability of 90%. Reliability is a measure of the variability that a radio system will observe during its operation. The reliability parameter, in the point-to-point mode, is implemented as the time variability parameter, of the area mode, and the confidence variability, in the point-to-point mode is used as the situation variability in the area mode, with the location variability set to 0, within the ITM source code. Confidence can be explained with the following example statement: "For at least 40% of the base stations, at range 5 km from the base station, the probability is 90% that the path loss is less than 135 dB". In the last mentioned example, the 40% indicates the confidence and the 90% is the reliability of the prediction [40].

To ensure a median approach, the value of 50% is chosen for reliability and a value of 50% is also chosen for confidence, as has been done in [41].

2.10 Computing Architectures

To calculate the TVWS for an area of South Africa, divided into a 1 km grid of 2 154 240 cells, with 2 194 TV transmitters, the result is 4 726 402 560 calculations for every propagation model. Such a huge number of computations need a powerful computer for the computation node in the Spectrum Observatory to do the calculations in as short a time as possible. In this section, the different architectures are considered and aspects relevant to the chosen architecture, HPCC, are discussed in detail.

The first computers used the von Neumann architecture. The von Neumann architecture consists of the main memory, a Central Processing Unit (CPU) and an interconnection between the two. This is the architecture on which the first and most serial computers were built. However, as the need for more processing power arose, the need for more CPUs rose with it. A parallel computer is a computer with multiple CPUs that work together on solving a problem [42].

There is a limit on single CPU speeds, however, and to do even faster computing than would be possible with serial computing, more than one CPU needs to be used on a computing job. These multiple CPUs are then used to work in parallel and thus exploit the advantage of parallel computing [43]. Parallel processing can be formally defined as the simultaneous processing of instruction using separate facilities [44].

2.10.1 Parallel Hardware

Parallel hardware can be divided into Single instruction, multiple data (SIMD) and Multiple instruction, multiple data (MIMD) systems. SIMD applies the same instruction on multiple data items, whereas a MIMD system is entirely parallel in the sense

that multiple instruction streams operate on multiple data sets. According to [43], MIMD systems are usually asynchronous meaning that the CPUs can operate at their own pace. Parallel systems can be divided into distributed-memory systems and shared-memory systems. In a shared-memory system, the various processors have access to the same memory, whereas in a distributed-memory system, each processor has its private memory. The different parallel hardware solutions can all be divided into these classifications. The parallel hardware solutions investigated are discussed below:

Cluster Computing

A cluster computer is a collection of computers interconnected by a network [45]. A cluster computer is a distributed-memory system. In the area of High Performance Computing (HPC), the computers used for the cluster will be high-performance computers with fast processors and memory. There are many components to a cluster computer set-up, and each plays a vital role in the performance of the cluster. The components mentioned above include the connection medium and computer node performance. The performance of different cluster computers is discussed and reviewed in [46].

The most well-known cluster-computer is the Beowulf cluster made at the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center's Center of Excellence in Space Data and Information Sciences. This project was the basis for cluster computers that were built using commercial off the shelf components. The Beowulf showed that a powerful cluster computer does not need to be expensive [47]. In [47] more detailed information about the Beowulf cluster can be found.

The demand for HPC has grown so much that there is always better computing hardware required to keep up with the demand. In [48] the author discusses the need for HPC and discusses the idea of cluster computing as a solution. The term HPCC is used to describe cluster computers used for high-performance computing. In South Africa, the CSIR has a division called the Centre for High Performance Computing (CHPC)

which built a HPCC called "Lengau" that consists of over 24,000 CPUs and boasts a peak performance of 24.9 teraflops/second. At its highest ranking, it came in 311 in the world's top 500 supercomputers and was ranked number one in all of Africa [49]. This HPCC is a distributed memory system with 24 CPUs per node.

The main advantages of using cluster computing are availability, scalability and performance [46]; when more computing resources are needed, more computing nodes can be added to the cluster. The main drawback for using cluster computing is the communications overhead that is needed to pass data between the different nodes [50]. The programs written for cluster computers are more complex than programs running on serial computers because the programs must be written to utilise this message passing.

Cloud Computing

As explained in [51], it is hard to find a formal definition of cloud computing. The article mentioned above looks at the origin of the term; according to the authors, the term cloud computing is not a word that has a single definition but rather refers to a trend in service delivery. The article goes on to discuss how the term developed, and it mentions the formal definition released by the National Institute of Standards and Technology (NIST).

A formal definition of cloud computing by NIST defines cloud computing as a paradigm for facilitating on-demand network access to a shared pool of configurable computing devices, that can be quickly provisioned and delivered with least management effort or service provider intercommunications [52]. According to this definition, a cloud computer can be seen as a distributed-memory MIMD.

A few definitions for cloud computing have been found in the literature. According to [53], it can be seen as a consumer model in which IT functions are sold as services, and it functions on a "you pay-for-what-you-use" basis. The significant factor of cloud computing is that the services are offered through the internet, and therefore an internet connection is required. A similar definition of cloud computing can be found

in [54]; according to the authors, cloud computing is "simply the delivery of applications; security and other services; storage and other infrastructures; and platforms such as those for software development to users over the Internet or a private cloud." [54]. A look at the above definitions can lead to the assumption that "cloud" refers to a network, that the term "the cloud" refers to the internet, and that cloud computing thus refers to a computing solution provided over the internet.

The architecture of cloud computing is highly dependent on the type of service that it provides. The types of services delivered in cloud computing are Infrastructure as a Service (IaaS), where the computing hardware is sourced over the cloud; Platform as a Service (PaaS), where software development environments are sourced over the internet, and Software as a Service (SaaS), where the software is not installed on a local computer but accessed and operated on the internet through a browser [55].

The advantages of cloud computing are flexibility and scalability because it is effortless to add more compute nodes to the network; it is also cost-effective because it costs less to lease a computing node than to buy one [56]. In [57] it is mentioned that current cloud computing solutions are not suitable for running HPC jobs because they do not offer Quality of Service guarantees. The significant shortcoming, however, is the increase in latency caused by communication overhead, which is worse than for HPCC, because the communication link between nodes is much slower than for HPCC. This is because the message passing in an HPCC takes place over memory interconnect busses, whereas in cloud computing, the message passing occurs over a network interface.

GPU Computing

The Graphics Processing Unit (GPU) is a specialised processing unit that was designed to process graphics [58]. GPUs are massively parallel because that have multiple cores on one unit. To program a GPU, a unique set of libraries or a specific programming language is needed.

For instance, Compute Unified Device Architecture (CUDA) is a unique tool-kit and compiler for programming GPUs.

When a GPU is used for general purpose computation, it is called a General Purpose Graphics Processing Unit (GPGPU). The GPGPU can be seen as a SIMD parallel system with the data being distributed amongst a number of cores, and these cores then execute the same instruction stream on the different datasets.

One of the few drawbacks of GPGPUs is that some of the math functions, like the trigonometric functions, are approximated when used on the GPU, leading to a slight compromise in accuracy [59]. Moreover, programs that runs on the GPGPU have to be modified to work on the GPGPU. Programs that require less memory will work better on a GPGPU.

Computation Performance Metrics

Measuring the performance of different computer architectures needs precise definitions of the metrics used. The computation metrics used depend on the computing architectures being tested and the tests being run. Arguably the most significant metric on a HPCC is the running time, also called latency. A formal definition for latency was presented as the time elapsed between starting the program and ending the program on a parallel computing architecture [60] [61].

The second metric being used is speed-up, which is defined as the ratio of the runtime of a serial solution (T_σ) to a problem to the parallel runtime (T_π) [42]. The equation for speed-up is given in (2.16),

$$S = \frac{T_\sigma}{T_\pi} \tag{2.16}$$

where T_σ is the serial solution runtime, T_π is the parallel solution runtime, both in seconds.

2.10.2 Parallel Programming

To make use of the parallel hardware discussed above, there is a need for parallel programming, in other words, to write computer programs that can exploit the multiple CPUs. Some parallel architectures work better with particular parallel programming paradigms. The two most widely known parallel programming approaches will be discussed.

MPI

The Message Passing Interface (MPI) is the most widely known message passing implementation. Message passing is used in a distributed-memory system where one process running on a processor can send information or data to another process running on another processor. One of the advantages of a MPI parallel program is that it is easily scalable to more processors. The MPI is not a programming language or a compiler. Instead, it is instead a library of functions that can be called in other programming languages, like C, C++ or Fortran. The C program is compiled with the MPI library, and this makes the message passing and other parallel functions available [43].

Message passing is often used in a distributed-memory system where one process running on a processor can send information or data to another process running on another processor. MPI is not a new programming language but rather a library that can be used with multiple programming languages and that enables message passing [42]. With MPI, the programmer has control over which CPU does which part of the code, and the programmer is also responsible for distributing data between the CPUs.

OpenMP

OpenMP is an API for shared-memory parallel programming. According to [43], the "MP" in OpenMP stands for "multiprocessing". OpenMP is intended for systems in which each CPU or processor can have access to all available memory. OpenMP is

relatively easy to implement because the system can decide at run-time and during compilation which processors should do what.

Every computing platform has its own unique software implementation challenges and approaches. The HPCC uses MPI, OpenMP or a mixture of the two to manage data between the different nodes. The CUDA platform is used in the case of GPGPU programming. The type of cloud computing available determines the programming platform used; for IaaS, for instance, the cloud computer can be used as a cluster and then MPI or OpenMP would be used.

2.11 Related Work

TVWS prediction by using propagation modelling has been done for South Africa in [39], and trials for the secondary usage of TVWS have been completted in [62] and [63]. A study was in Hull, in the UK, showed that most of the propagation models used overestimate the path loss of the signal when compared to measured data. The Egli model used, which takes into account terrain data, was the closest to the measured data, and it is thus preferred by [64] to do cognitive radio planning.

GPUs have been used for HPC to do propagation prediction in [65] and [66]. In [66] the authors used FDTD with GPUs to do radio coverage prediction, while in [65] the COST-Walfisch model was implemented on a GPU. Both of these studies only consider small urban areas for the evaluation.

Cloud computing has been used in the form of a commercial SaaS solution for propagation prediction. Examples are CloudRF [67] and LS Telcom [68]. CloudRF has been used for cellular coverage design in [69].

In [70] the authors build a special-purpose machine to do finite-difference wave propagation because the HPCs then could not yet meet the demands of the advanced algorithms used.

The authors of [71] conducted a performance analysis of cloud computing solutions. The first metric they used were time to complete calculation (to test CPU performance); they also measured bandwidth by using iPerf to test network performance, and they used file handling speed tests to test file handling, using *sysbench* from MySQL. The authors ran every test ten times to create a test case. In [72] the performance metrics used were response time, throughput and CPU utilisation. Every test case consisted of running a test three times and then using the average.

A study to draw a comparison between predictions and experimental data for different propagation models was done by the authors of [73]. In their article, the authors used the mean value and standard deviation of the difference between the measurements and the predicted values in their analysis.

The authors of [64] conducted a comparative study of the predicted field strengths from propagation models against measurements taken, with the location of the study being around Hull in the UK. The metrics used in this study are the actual difference, the average error of the propagation model, the standard deviation of the error and the Root Mean Square Error (RMSE). The propagation models used in their study are the Extended Hata model, the Hata-Davidson model and the Egli model, for frequencies ranging from 470 to 790 MHz. The authors used these empirical propagations with and without terrain data. The authors used the terrain data to calculate diffraction losses and added these to the losses calculated by the empirical models. The authors found that, when terrain data is not used, the Hata-Davidson and Extended-Hata models provided the best fitting results, while when terrain data was also implemented, the Egli model provided the best results.

In a study, by [41], the authors attempted to improve the ITM P2P model with measurements for more accurate WS detections. The authors used the ITM model with Shuttle Radar Topography Mission (SRTM)-3 terrain data, which gives terrain data for every 3 arc-seconds. The ITM mode was implemented with the default values for the dielectric constant and the conductivity of the ground parameters. The climate was set to maritime temperate, and the reliability and confidence statistical parameters were set to 50% to get a median value of the attenuation.

In [74] a quantitative study of WS in the USA was done based on the FCC rules. The concept of an erosion margin or a fading margin was introduced to protect the PU from severe fading events, and the authors used a margin of 1 dB as recommended by the FCC. The study also showed that for a fading margin of $\lambda = 0.1$ dB, more TVWS is available.

In a quantitative study of TVWS in Europe, the authors investigated the amount of TVWS available using a statistical propagation model and a deterministic propagation model [75]. The metrics used to analyse the amount of TVWS are the average amount of UHF channels available as WS weighted over a geographical area as well as the average amount of UHF channels available as WS weighted per population. In their calculation, the authors based their WS classification on the regulations set out by the FCC. The erosion or fading margin used by the authors was $\lambda = 1$ dB, as set out by the FCC.

In [19] the writers use a geolocation-based approach to determine the capacity of TVWS in south-eastern Europe. In this study, the authors used the ITM P2P model to do TVWS predictions because of the variations in the terrain of the area under investigation. The protection ratios adopted by the ITU for DTT transmitters were used as a guideline to do TVWS classification.

2.12 Conclusion

The discussion of fundamental concepts of spectrum management, DSA and TVWS in this chapter gave a solid introduction to the relevance of this study. This chapter also included information on the use of the geolocation-based approach for DSA and it considered how regulators from different countries have started to implement this approach. The different propagation models implemented on the Spectrum Observatory were also discussed. Information was presented on computation and computing architectures, and in particular parallel computation. The last section looked at the related work done by researchers in the field.

Chapter 3

System Model

This chapter explains the system model used to generate the results that were used when analysing the amount of TVWS available for the different propagation models. The Spectrum Observatory being created is presented and discussed with a focus on the parts of the Spectrum Observatory that were designed and implemented by the author.

3.1 Introduction

The PU usage must be predicted by using the different propagation models to investigate the TVWS availability for South Africa when using different propagation models. The setup for the investigation is explained in the next section. An integral part of the setup is the grid system, which is also explained in this chapter. The Spectrum Observatory, which is used for the TVWS analysis, is also described and discussed in this chapter.

3.2 TVWS Prediction Setup

To compare different large-scale RF radio propagation prediction models, the link budget used for propagation prediction, as illustrated in Figure 3.1, is employed. The propagation models are used to calculate the path loss(dB). The path loss, along with the transmitted power, is utilised to determine the received power, in dBm, as seen in Equation (3.1).

$$P_{rx}(dBm) = P_{tx}(dBm) - PL(dB) \tag{3.1}$$

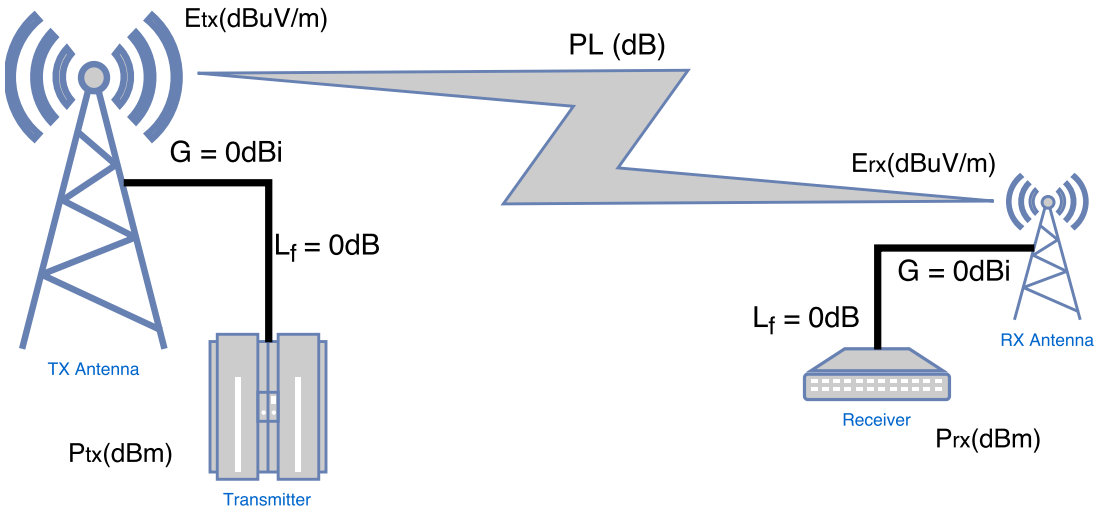


Figure 3.1: Propagation Prediction Link Budget

The symbols used in Figure 3.1 and in Equation (3.1) denote the following:

P_{tx}	-	Transmit power in dBm
P_{rx}	-	Receive power in dBm
E_{tx}	-	Electric field strength at the transmit antenna
E_{rx}	-	Electric field strength at the receiving antenna
PL	-	The predicted path loss in dB
L_f	-	Feeder loss in dB
G	-	Antenna gain relative to an isotropic radiator in dBi

Note that P_{tx} and P_{rx} are both electrical power, and are chosen because most signal generators use dBm as unit. Conversely, E_{tx} and E_{rx} are both electric field strength and are used to display the power of RF signals.

The WS policy used for predicting TVWS is applied to the field strength at the receiver. The first step is to get from the Effective Radiated Power (ERP), given by ICASA in dBW relative to a half-wave dipole, to the transmitted power(P_{tx}), relative to an isotropic radiator, in dBm, as seen in Figure 3.1. This is done by using Equation (3.2).

$$P_{tx}(dBm) = ERP(dBW) + 2.15 + 30 \tag{3.2}$$

The next step is to do the path loss (PL) prediction using the propagation model, followed by the calculation of the received power (P_{rx}) with Equation (3.1). This received power (dBm) is then converted to the field strength at the receiver relative to an isotropic radiator, which means zero feeder losses (L_f) and antenna gain (G), which is then fed to the WS policy for TVWS prediction.

3.2.1 Transmitter

ICASA is the regulatory body for South Africa, and they hold the data for the terrestrial TV broadcasters. ICASA set out a Terrestrial Broadcasting Plan in 2008 [76], which holds the information for analogue, MDTT and DTT transmitters. The Terrestrial Broadcasting Plan was used along with the ICASA TV Transmitter database, as have been done in [39], to create a list of 1800 analogue TV transmitters, 533 DTT transmitters and 73 MDTT transmitters. That list of transmitters contains the information that was used by the propagation models, which can be seen in Table 3.1. The ERP is given in dBW, relative to a half-wave dipole, so this is first converted to the power through an isotropic radiator by merely adding 2.15 dB to the power given.

Table 3.1: Transmitter Information

Station Name	Channel Number
GPS Latitude	GPS Longitude
Frequency (MHz)	ERP (dBW) (Through half-wave dipole)
Polarisation	Antenna Height (m) (AGL)

3.2.2 Path Loss

The central focus of this study is the path loss calculations. The different propagation models, which were discussed in detail in Section 2.9, are used to calculate the between every transmitter in the transmitter database and the relevant receive locations, and given in decibels (dB).

Free-space

The Free-space loss model is implemented as a worst-case scenario. The Free-space loss model does not have any restrictions on the parameters used. The only two parameters that the Free-space loss model uses are frequency and distance.

Hata-Davidson

The Hata-Davidson model was implemented with the antenna HAAT height for the effective antenna height, as described in Section 2.9.2. The area correction factor chosen for the proposed implementation was chosen as Suburban, as motivated in Section 2.9.2. The model was implemented with the parameter ranges shown in Table 2.3.

The HAAT for the same set of transmitters has been developed by [39]; the HAAT calculation approach is based on the ITU method for point-to-area predictions [77]. The average of the ground height Above Mean Sea Level (AMSL) is calculated for a distance of between 3 and 15 km from the transmitter. This average ground height AMSL is calculated at 10° intervals for 36 azimuths, starting at 0° and ending at 350°, and the maximum of these average heights is then used as the average terrain height for that transmitter site. To calculate the HAAT for every transmitter at the transmitter site, the antenna height AGL is added to the average terrain height for that transmitter site. Note that the HAAT was calculated using the 3 arc-second SRTM DEM.

In this research, the HAAT has been used as the effective antenna height for the Hata-Davidson implementation, denoted as HATAE. The antenna mast height AGL was also implemented as the effective antenna height for the second set of Hata-Davidson results, denoted as HATA, in order to observe the difference between the two implementations.

A number of transmitters had effective antenna heights that were less than 20 m, which is the lower boundary of the acceptable effective antenna height of the transmitter for the Hata-Davidson model. These transmitters are most probably transmitters that were designed to support only a small area with coverage. In the proposed implementation, the effective antenna height was forced to 20 m for these transmitters, as a worst-case scenario approach, which would lead to overprotection of the PU and the risk of missing potential TVWS. The list of transmitters with effective antenna height less than 20 m can be seen in Table 3.2, which is presented at the end of the chapter; all of these transmitters are analogue transmitters.

An investigation into how much of South Africa is rural and how much is urban revealed that, as per the 2001 census, 57.5% of South Africa is urban and the rest is rural [78]. These facts, as mentioned earlier, led the author to set the area parameter to suburban.

ITM Area

The ITM Area model was implemented with the parameters set out in Table 3.3 at the end of the chapter. The frequency, transmitter heights (as AGL) and polarisation were read from the transmitter list obtained from ICASA. The choice of values for the parameters used in the ITM Area model is discussed in Section 2.9.3.

ITM P2P

The parameters chosen for the ITM in P2P mode are listed in Table 3.4 at the end of the chapter with the discussion of these parameters in Section 2.9.3. The frequency, transmitter heights (as AGL) and polarisation were again extracted from the transmitter list provided by ICASA.

DEM Data Used Terrain-aware propagation models, like the ITM in P2P mode, apply a DEM to modulate the interaction of the RF signal with the ground. The Digital Elevation Model (DEM) holds, as its name suggests, elevation data for a number of points in an area. Most DEM data is given in meters AMSL.

One of the most well-known sources of DEM data is the SRTM data. The SRTM DEM has been made available for South Africa with resolutions of 30 arc-second and 3 arc-second. 1 Arc-second data is also available for the USA. This data has been collected using radar interferometry by various space agencies. The 3 arc-second SRTM data is used by the author of [39], to calculate the HAAT for the TV transmitters of South Africa.

Another DEM is the Global Land One-km Base Elevation Project (GLOBE) 30-arc second DEM. The GLOBE DEM is a global elevation model on a 1-kilometer grid. According to the website of the GLOBE DEM, the data came from 11 sources, through 18 combinations of image processing. This project was an international effort to create a global DEM. The ITM in P2P mode was originally designed by the Institute for Telecommunication Sciences (ITS), in the USA, to work with the GLOBE DEM. Therefore, the DEM data used in the ITM P2P implementation is the 30 arc-second or 1 km GLOBE DEM [79].

Number of elevation points The number of elevation points extracted from the DEM for each path between the transmitter and receiver can be set when using the ITM P2P. This number of points of elevation read from the DEM file also determines the distance

between each elevation point. The researchers at ITS who wrote the ITM application that is used for verification of the proposed implementation, set their number of points to a fixed number of 800, regardless of the path length. In this study, the amount of elevation points used is thus also fixed to 800 to coincide with the program created by the researchers at ITS.

3.2.3 Receiver

In this setup, the receiver is simulated as having an isotropic antenna with no cable losses or extra system gains and losses. In reality, the antenna pattern and cable losses would have an effect on the signal, but these are ignored in this study because the upper limit of the TVWS channel availability is being investigated. The receiver height AGL is chosen to be 10 meters, following recommendation set out in ITU-R BT.1386-12 [80]. The choice of 10 meters also makes logical sense, as this is about the average height of a terrestrial antenna on a household roof. This height is also above the usual suburban clutter. The choice of an isotropic antenna without any further receiver gains and losses allows the easy conversion between received power and electrical field strength.

3.2.4 Grid System

The grid system used needs to be defined before the received power is presented. This grid system makes use of geographical coordinates (GPS coordinates) and therefore no projections took place. The boundaries of the grid system are defined as the area of South Africa, stretching from 22.0° S to 35° S and 16.0° E to 33.0° E. This area is divided into a 30 arc-second grid to correspond with the 30 arc-second DEM used. The grid system can be seen in Figure 3.2, with the cell illustration in Figure 3.3 showing each cell being referenced at the centre of the cell. The cells are numbered with the left number indicating the latitudinal cell index or row and the right number indicating the longitudinal cell index or column.

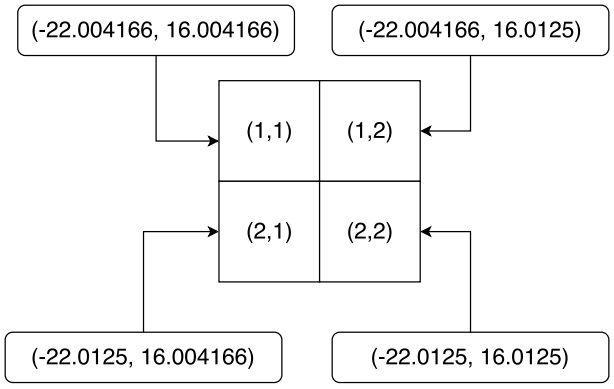


Figure 3.2: System Grid

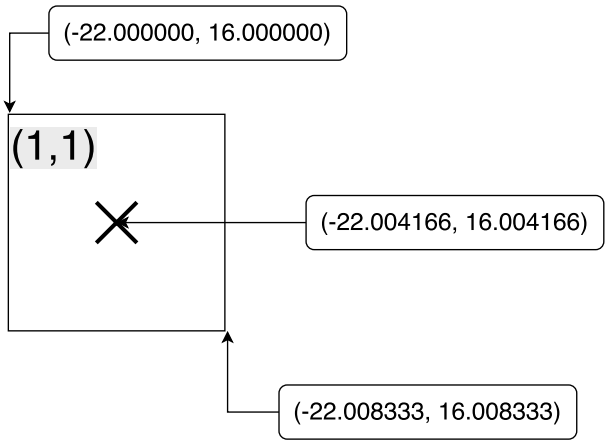


Figure 3.3: Grid Cell

The cells are referenced at the centre of the cell, as observed in Figure 3.3. Moreover, the grid is organised in such a way that the rows are the latitudes and the columns are the longitudes. The grid contains 1560 cells along the latitudinal axis and 2040 cells along the longitudinal axis, resulting in a total number of 3 182 400 cells for the area when using a 30 arc-second cell length. This step is repeated for every transmitter, but the calculations with regard to the PUs are limited by checking the thermal noise, as explained in Section 3.2.6.

3.2.5 TVWS Prediction Process

The flow diagram in Figure 3.4 displays the process that was followed to do TVWS prediction on the Spectrum Observatory.

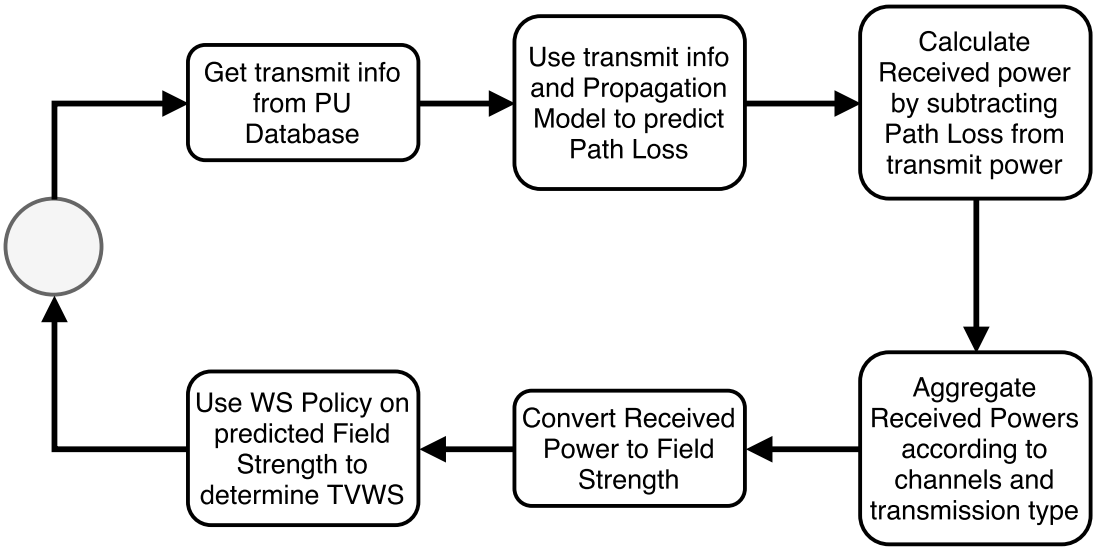


Figure 3.4: WS Prediction Flow Diagram

The first step is to obtain the relevant PU information, such as site locations, ERP, broadcast frequency and polarisation. This operational information is fed into the propagation model, which predicts the path loss for that transmit information. The path loss calculated is then subtracted from the PU transmit power to arrive at the received power. The above steps are repeated for every PU in the PU database. The received powers are then aggregated, by taking the maximum received power at every location, according to the UHF channels and the illumination technology. The maximum received powers are then converted to field strength values, which are fed into the TVWS policy to determine which locations can be used as TVWS. Note that this process is done for each cell in the grid.

3.2.6 Received Power Calculation

The first deliverable from the program is the received power (dBm) resulting from every transmitter in the list of transmitters used. This investigation was done in the South African context, and the analysis of this area was done using the grid system discussed above.

A decision was made to calculate only the received power for an area limited by received powers higher than thermal noise when using free-space loss to calculate the received powers. This was done to save computing power and to work more efficiently, because calculating the free-space loss would take much less time than calculating the propagation loss with more complex terrain-aware propagation models. The Thermal Noise at a receiver is the noise that results from the unpredictable movement of free electrons in a conductor in the receiver, caused by heat. If an RF signal has a power level at the receiver that is lower than the thermal noise, the signal to noise ratio will be too weak to get a reading [31]. The Thermal Noise can be computed as a power level in watts using Johnson's formula, where k is Boltzmann's constant, T is the room temperature in Kelvin, usually set to 290, and B is the bandwidth of the signal under consideration.

$$P_n = \frac{v_n^2}{R} = \frac{\left(\sqrt{4kTBR}\right)^2}{4R} = kTB \tag{3.3}$$

This results in a formula for the Thermal Noise in dBm being:

$$P_n(dBm) = 10\log(B) - 174 \tag{3.4}$$

Note that, in the above equation, the signal bandwidth is expressed in Hz. The maximum free-space loss can be calculated by

$$L_{FS}(dB) = T_x(dBm) - R_{Thermal}(dBm) \tag{3.5}$$

where $T_x(dBm)$ is the transmitter power in dBm.

The free-space loss equation, seen in Equation (2.1), can be rewritten to get the maximum calculation distance as shown below in Equation (3.6), where d is the distance in km and f is the frequency in MHz.

$$d = 10^{(\frac{Loss_{FS} - 20\log(f) - 32.45}{20})} \tag{3.6}$$

A buffer of 25 km is then added to this result, to account for the difference between approximating the earth as a sphere and the earth's exact shape when calculating the geodesic distance, in order to get the maximum calculation radius from the transmitting antenna for propagation prediction with all the propagation models.

This distance is then converted back to identify the cells that fall within the predicted area. A graphical presentation of the cells calculated and stored for each transmitter can be seen in Figure 3.5 The received power values calculated for each transmitter in the transmitter list are stored per transmitter as a (.fst) custom binary file. More details about this file can be found by using Appendix A.

It should be borne in mind that the only propagation model used that places a limit on the distance input parameter and that would affect calculations is the Hata-Davidson model, which is limited to a range of 300 km. In this case, if the distance calculated from the thermal noise constraint was more than 300 km, the distance was set to 300 km. In Figure 3.5, the cells that are predicted but that fall outside the cut-off distance, viz. the green cells in respect of the Hata-Davidson model, are assigned a default received power value of -500.00 dBm. This default value was chosen to allow for the classification of TVWS, while still making it possible to distinguish between default and predicted values.

The next step in the received power prediction is to aggregate the transmitters according to the UHF channels, for analogue, MDTT and DTT separately. A graphical representation of this can be seen in Figure 3.6. This part of the Spectrum Observatory generates two files, a Received Power per Channel for the country (.fsc) and an Interference Data per Channel for the country (.infc). Once again, these are custom binary files, whose specifications can be seen in Appendix A.

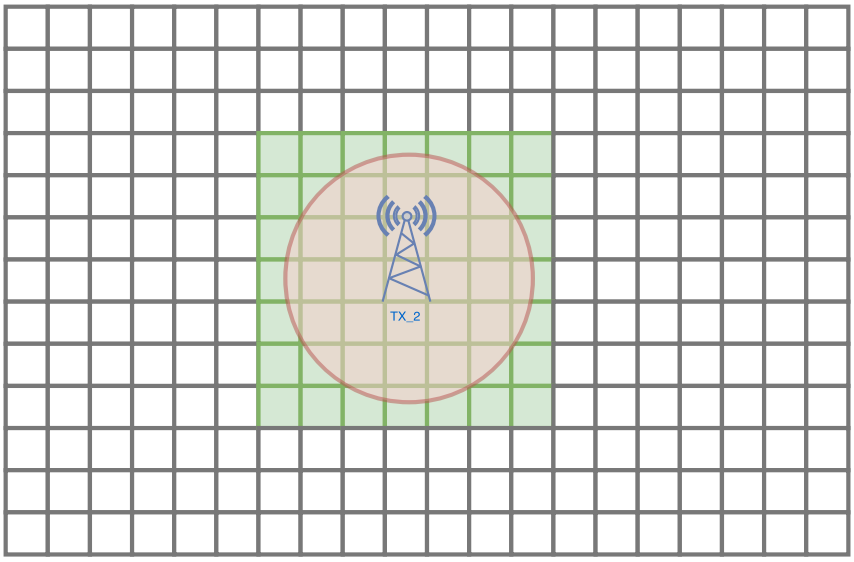


Figure 3.5: Transmitter Received Power Area

The received power per channel for the country file contains the wanted receive power for each cell, which is the maximum received power at that cell, as well as the transmitter responsible for that received power. The interference file contains all of the received powers that are present in every cell as well as the reference to the transmitters that caused the received powers.

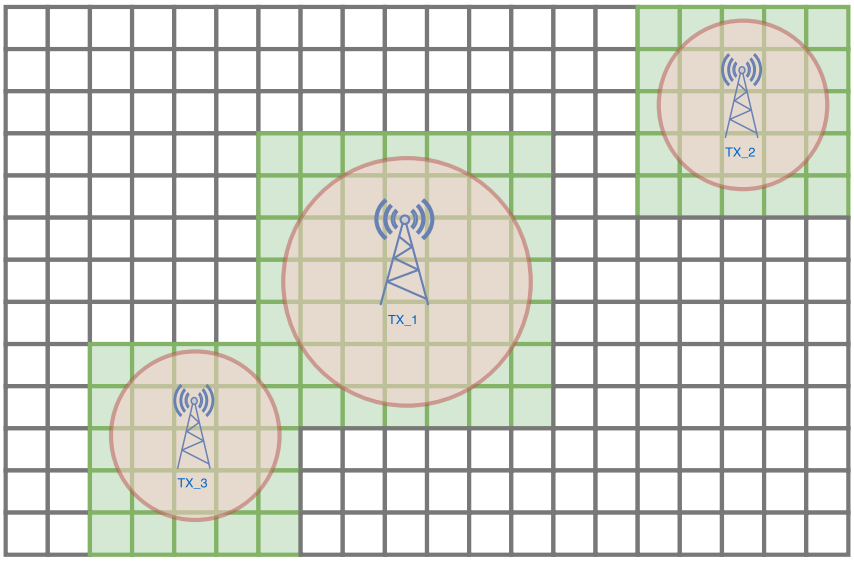


Figure 3.6: Transmitter Received Power Area per Channel

3.2.7 Received Power to Field Strength

The TVWS classification is done in the field strength unit of (dBuV/m). To perform the TVWS classification, the received power must thus first be converted to field strengths. The predicted received power in dBm can be converted back to the electric field strength in dB(uV/m) by using the assumption that the propagation prediction is done with an isotropic antenna without any equipment or feeder losses. The effective antenna aperture from an isotropic antenna can be seen in Equation (3.7), with f being the frequency in MHz [32].

$$A_{ei} = \frac{300^2}{4\pi f^2} = \frac{7160}{f^2} \tag{3.7}$$

The power density of an RF signal has both a magnetic field and an electric field component, as expressed in Equation (3.8).

$$P_d = EH \tag{3.8}$$

The transmission medium used in this study is air, and free-space properties are thus assumed for it. The assumption of free-space properties entails that the intrinsic impedance relationship between the electric field and the magnetic field, as seen in equation (3.9), can be used to obtain a relationship between only power density and electric field strength, as seen in Equation (3.10) [32].

$$E = 120\pi H \tag{3.9}$$

$$P_d = \frac{E^2}{120\pi} \tag{3.10}$$

By using the antenna aperture for an isotropic antenna along with the power density at the receiver, a relationship between the electric field strength and the received power can be seen, as in Equation (3.11) [32].

$$P_r(isotropic) = P_d A_{ei} = \frac{E^2}{120\pi}\frac{7160}{f^2} \tag{3.11}$$

Using microvolts per meter as the unit for electric field strength and milliwatts for the received power, as well as converting to logarithmic units, the above equation can be expressed as seen in Equation (3.12). This leads to Equation (3.13), showing the field strength in dBuV/m from received power in dBm with frequency f in MHz [32].

$$P_r dBm = E(dB\mu V/m) - 20\log(f) - 77.2 \tag{3.12}$$

$$E(dB\mu V/m) = P_r(dBm) + 20\log(f) + 77.2 \tag{3.13}$$

3.2.8 WS Determination

To determine whether there is spectrum available to be used as WS, the predicted field strength is compared to the minimum field strengths required for operating each of the terrestrial TV technologies (analogue, MDTT and DTT) according to the relevant ITU recommendations. This is done to protect the PU service, who has the first right of usage of the TV spectrum. At each grid location, if the predicted field strength is lower than the minimum field strength required for terrestrial TV broadcasting, then it can be concluded that the grid cell under consideration is not part of the coverage area of the PU for that channel. To get an upper bound of the available TVWS an SU transmitting at a hypothetical 0 W is used, and the effect of the SU is ignored, leading to the protection region around the PU being only equal to the minimum field strength. The WS policy used is thus a basic or self-defined TVWS policy that is based on ITU recommendations.

The TVWS policy used has the premise of protecting the PU service from harmful interference, meaning interference that would lead to a degradation in the quality of the PU service. No other protection region is added to this basic requirement of protecting the PU service, to get an upper bound on the amount of TVWS channels available for the different propagation models.

$$E_{protect} = E_{min} + \psi \tag{3.14}$$

with $\psi = 0$dB.

$E_{protect}$ denotes the PU field strength that the policy must protect, E_{min} denotes the minimum field strength from the PU, as discussed below, and ψ denotes the fading margin. The protected field strength is the minimum field strength of the PU that must be present at a location in order to imply that that location is still part of the region allocated to the PU service. The fading margin (ψ) has been set to 0 dB in the proposed implementation to give an upper boundary on the amount of TVWS available for the different propagation models. Regulators implementing the erosion margin will need to come to a consensus on the value and impact of the erosion margin, but for this study, the erosion margin is taken to be 0 dB.

Analogue

The minimum field strength, produced from the PU, for which protection may be sought in planning analogue terrestrial TV services as defined in [81] is:

$$E_{min}(dB(\mu V/m)) = 62 + 20\log(f/474) \tag{3.15}$$

Equation (3.15) gives the field strength at a height of 10 m above ground level, where f is the channel centre frequency in MHz.

DTT

The minimum field strength that may be sought when doing DTT planning is a bit more difficult and is explained in detail in [80]. The minimum field strength can be calculated using the following equations:

$$\begin{aligned} E_{min}(dB(\mu V/m)) &= \varphi_{min} + 120 + 10\log(120\pi) \\ &= \varphi_{min} + 145.8 \end{aligned} \tag{3.16}$$

$$\varphi_{min} = P_{smin} - A_a + L_f \tag{3.17}$$

$$P_{smin} = C/N + P_n \tag{3.18}$$

$$P_n = F + 10\log(kT_0B) \tag{3.19}$$

$$A_a = G + 10\log(1.64(^{\lambda}/_{4})\pi) \tag{3.20}$$

where:

P_n	-	receiver intrinsic noise power (dBW)
F	-	receiver noise figure (dB)
k	-	Boltzmann's constant($k = 1.38 \times 10^{-23}$ (J/K))
T_0	-	absolute temperature($T_0 = 290$ (K))
B	-	receiver noise bandwidth($B = 7.61 \times 10^6$ (Hz))
P_{smin}	-	minimum receiver input power (dBW)
C/N	-	RF S/N at the receiver input required by the system (dB)
A_a	-	effective antenna aperture (dBm^2)
G	-	antenna gain relative to a half dipole (dBd)
λ	-	wavelength of signal (m)
φ_{min}	-	minimum power flux density at receiver (dB(W/m^2))
L_f	-	feeder loss (dB)

The bands for which calculation were done are summarised in Table 3.5 at the end of the chapter.

The recommended values for the parameters used above are stated in [9], and can be seen in Table 3.6 at the end of the chapter. In [9] the reference planning configuration Carrier to Noise Ratio (C/N) for a fixed antenna setup is 21 dB.

MDTT

The protected field strength required for Mobile Digital Terrestrial TV (MDTT) services is calculated using the same equation as for the DTT calculations, which can be seen in Equation (3.16). Some of the parameters used for the minimum field strength calculation differ for MDTT [9]. The recommended Antenna Gain (G) is 0 dB for both band IV and band V, the Feeder Loss (L_f) is 0 dB for both bands and the Noise Figure (F) is 6 dB for both bands [9]. The C/N for a portable outdoor antenna, recommended by [9], is 19 dB.

3.2.9 WS Criteria

The WS classification is done using Equation (3.21). In Equation (3.21), E_{wanted} denotes the maximum predicted field and $E_{protect}$ denotes the protected field strength of the PU for which protection is sought, as defined in Equation (3.14). The subset of channels under investigation is $\mathcal{C} = \{21, 22, .., 67, 68\}$.

$$WS = \begin{cases} 0, & E_{max} \geq E_{protect}, c \in \mathcal{C} \\ 1, & E_{max} < E_{protect}, c \in \mathcal{C} \end{cases} \tag{3.21}$$

Each cell in the grid for the whole country will thus have a binary value of 0, if there is no TVWS available in that cell and a value of 1, if TVWS is available. The answer for all the channels, per channel, for South Africa will be stored in three custom binary TVWS files (.wsc), i.e. one for the analogue entries, one for the MDTT entries and one for the DTT entries.

The fully functional Spectrum Observatory will allow researchers to pick from a number of TVWS policies to use when conducting TVWS research. The research done in this study used the policy discussed above.

3.2.10 Exclusion Zones

The TVWS saved in the binary files mentioned above is for the entire grid defined in Section 3.2.4, but certain parts of this grid fall outside of South Africa and thus also outside of the proposed analysis. By excluding those areas from the analysis, it is implied that the excluded region is not added toward the totals for the results. Vector data of South Arica is used to manipulate the TVWS data in the grid system to exclude areas that fall outside of the landmass of the Republic of South Africa. The usage of the vector data for South Africa leads to the exclusion of Lesotho and Swaziland.

Another region that is excluded from the TVWS analysis is the core advantage area of the Square Kilometer Array (SKA). The SKA is an international radio astronomy project with a core part of the project being located in the Northern Cape in South Africa. This project is protected against Radio Frequency Interference (RFI) by the Astronomy Geographic Advantage Act [82]. The regions that are excluded from the analysis, as part of the SKA, are classified as the core and central advantage areas.

3.2.11 WS Metrics

South Africa is currently in a dual-illumination stage: the analogue switch-off is scheduled to start and some DTT and MDTT PUs have already been commissioned, but this is an ongoing process and far from complete. In order to predict TVWS without causing harmful interference to any of the illumination technologies (Analogue, DTT or MDTT), the TVWS for dual-illumination was thus determined. A cell in the grid system was classified as TVWS for the dual-illumination time, if there was TVWS for all three of the illumination technologies(Analogue, DTT or MDTT). If there was PU activity for that cell on any of the illumination technologies, the dual-illumination status for that cell was be not-TVWS.

The number of channels available per province per propagation model during dual-illumination was calculated, to quantify the amount of TVWS available when using each propagation model. To aggregate the TVWS data that was stored in the grid, as explained in Section 3.2.4, the data was weighted by area for the nine provinces of South Africa.

WS weighted by area

For every province of South Africa, the amount of TVWS UHF channels available was calculated using Equation (3.22).

$$WS_{area} = \frac{\sum_{i=1}^{N} WS_i A_i}{\sum_{i=1}^{N} A_i}, \forall \mathrm{c} \in \mathcal{C} \tag{3.22}$$

where the terms denote the following:

WS_{area}	-	WS weighted by area
i	-	Number of the grid cell
N	-	Number of grid cells in the area
WS_i	-	The WS for that cell, either a 1 or a 0
A_i	-	Area of that cell
$\mathcal{C}$	-	Set of UHF channels considered $\mathcal{C} = \{21, 22..67, 68\}$

The areas was calculated by creating a shapefile of South Africa, which includes the provincial borders. These borders, which coincides with the 2001 Census, were used together with Green's Theorem, to calculate the area of each province. The area calculated was then used in the metric discussed above to calculate the WS per area. The calculation of the areas for South Africa and the different provinces, while incorporating the exclusion zones, is part of previous work done in [39]. The exclusion zones, which were discussed in Section 3.2.10, were not included in these calculations.

3.3 Spectrum Observatory Architecture

The ultimate goal is to create an entirely automated Spectrum Observatory. The architecture of the proposed Spectrum Observatory can be seen in Figure 1.1. The Spectrum Observatory being constructed is an evolving piece of work, with multiple researchers contributing to it.

This research focuses on the following parts of the Spectrum Observatory:

1. Primary User Database
2. Propagation Models
3. Policies
4. Computation Node

The Primary User (PU) database is the transmitter list that was discussed in Section 3.2.1.

The propagation model component is being investigated in this research with the setup explained above, to determine which propagation model would work better in which scenario. The idea is to have multiple propagation models available for researchers on the Spectrum Observatory. Once the Spectrum Observatory has been completed, the researchers can use or add any RF propagation model.

The TVWS policies component of the Spectrum Observatory is there to make a list of policies available for research; in this research, though, an elementary policy is employed. The policy used is the TVWS classification discussed in Section 3.2.8. ICASA, the regulatory body of South Africa, published draft regulations for TVWS usage in South Africa in April of 2017 [15]. The final rulings on the usage of TVWS and a clear policy are yet to be released.

A central part of the spectrum observatory, which is also a focus of this research, is the computation node. The calculation of the received powers from the PUs and the calculation of the available TVWS channels will occur within this component of the Spectrum Observatory.

3.3.1 Implementations of Propagation Models

The source code for implementing the ITM propagation model and the source code for extracting the terrain data from the DEM tiles are from the ITS, which is part of the National Telecommunications and Information Administration (NTIA) in the USA. The DEM used is again the GLOBE DEM data [79], as stated above. The source code for the ITM implementation had to be ported from C++ to C. All the of the source code had to be parallelised by means of MPI routines. The source code for the DEM data extraction routines was only available in Fortran 77 and had to be ported to C. The other propagation models (Free-space and Hata-Davidson) were easier to implement and were written directly from empirical equations. The parallelisation of the propagation models is explained in more detail in Section 3.3.2.

3.3.2 Parallelisation of Models

The parallelisation of the propagation models can be graphically represented with the flow diagrams in Figure 3.7 and Figure 3.8. The ITM in P2P mode is taken as an example because it is the most complicated model of the four models chosen for the TVWS analysis. The received power prediction is an SIMD problem handled by means of MPI, which means that the same program is sent to every CPU, but every CPU only does the part written for it.

The first part of the parallelisation is where CPU 0 reads the list of PUs and divides them among the number of CPUs, as seen in Figure 3.7. MPI is a message passing interface; it is used to send the PUs that each CPU must process to the relevant CPU.

The relevant CPU receives this information using MPI, and then the programs on the different CPUs are synchronised, with the MPI_Barrier function. This function waits for all the CPUs to get to the same part of the program. In this case, this will occur when all of the CPUs have received all the PUs they should process.

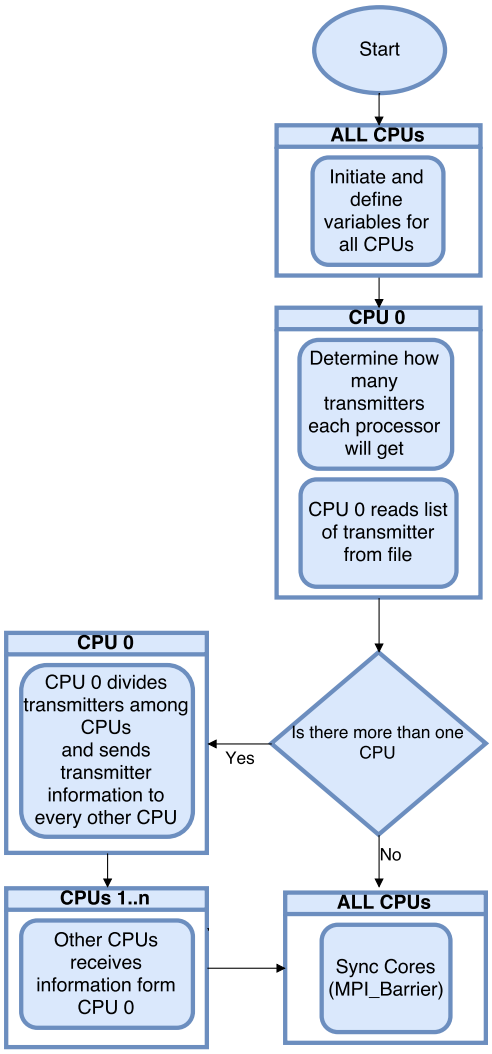


Figure 3.7: ITM P2P Work Split

The second part of the parallel program is where the path loss and the received power for each transmitter are calculated. The first step is where each CPU opens a file for the transmitter for which it is currently calculating the path loss and received powers. Then the number of longitudinal and longitudinal cells that needs to be calculated is determined. Thereafter, for each grid cell in the grid, the Haversine distance needs to be calculated from that cell to the transmitter cell.

For each of these distances and paths, an elevation profile is extracted from the DEM. The path extraction is only needed for the ITM in P2P mode; the other propagation models do not need path information, and thus this step can be skipped. This is the part of the ITM in P2P mode, where the DEM file is opened, and elevations are read for every path. After the path profile has been saved into an array, the path loss is calculated. This path loss along with the transmit power for every PU is in turn used to calculate the received power, in dBm, at each cell. The result of the path loss and the received power is then written to the transmitter result file. After this process has been followed for all points and for all the transmitters, the CPUs are synchronised with the MPI_Barrier call, and the program is ended. A flow diagram displaying the steps discussed above can be seen in Figure 3.8.

3.4 Conclusion

The system used to calculate and analyse TVWS for this study was reviewed in detail in this chapter. The link budget, grid system and the TVWS policy used were all discussed, and the details of the propagation models implemented were stated. This chapter also reviewed the Spectrum Observatory and described the parts of the Spectrum Observatory that were worked on during this research. The parallelisation of the different propagation models to be used on the HPCC was also discussed in this chapter.

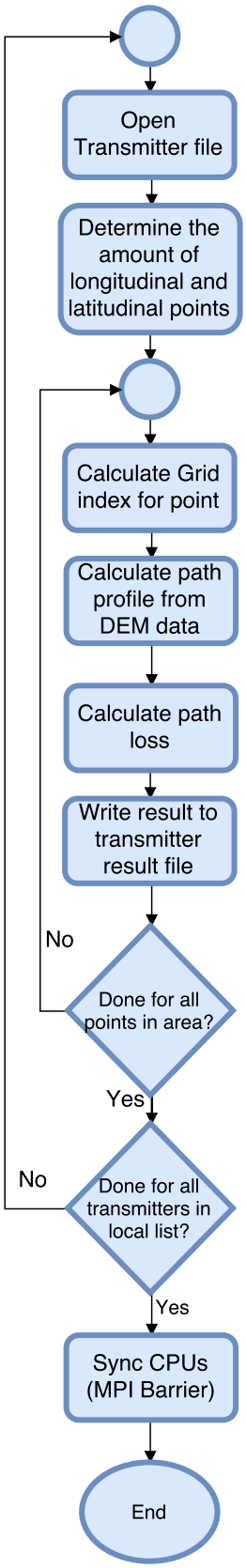


Figure 3.8: ITM P2P Path Loss

Table 3.2: List of Transmitters with Effective Antenna Height less than 20 m

Description	Channel	Effective Height (m)	Mass Height (AGL) (m)
GOODHOUSE	56	-14.17	10.00
GOODHOUSE	60	-14.17	10.00
GOODHOUSE	64	-14.17	10.00
GOODHOUSE	68	-14.17	10.00
STELLA	56	17.20	12.00
MSAULIMINE	24	-2.64	10.00
MSAULIMINE	37	-2.64	10.00
MSAULIMINE	39	-2.64	10.00
MSAULIMINE	46	-2.64	10.00
MSAULIMINE	51	-2.64	10.00
HEXRSANDHLSKANETCL	63	2.35	10.00
KAKAMASSEEKOEISTEEK	54	4.10	10.00
KKLCALTIZDORPSPA	46	10.17	0.00
PELAMISSION	38	-0.88	18.00
PELAMISSION	42	-0.88	18.00
PELAMISSION	46	-0.88	18.00
LADISMITHAMALIENSTN	31	-33.15	10.00
WARRENTON	43	16.34	10.00
LYDENBURGDOORNHOEK	40	-20.48	10.00
PORTNOLLOTH	21	16.02	10.00
PORTNOLLOTH	23	16.02	10.00
PORTNOLLOTH	27	16.02	10.00
PORTNOLLOTH	31	16.02	10.00
PORTNOLLOTH	35	16.02	10.00
FORTBEAUFORTLORR	45	1.27	0.00
PRIESKA	39	13.97	10.00
PRIESKA	43	13.97	10.00
PRIESKA	47	17.72	0.00
RIEMVASMAAKVREDESVA	53	11.34	10.00
RIEMVASMAAKVREDESVA	57	11.34	10.00
RIEMVASMAAKVREDESVA	61	11.34	10.00
RIEMVASMAAKVREDESVA	65	11.34	10.00

Table 3.3: Parameters for ITM Area Implementation

Parameter	**Value**
Surface Refractivity	322.9943
Relative Permittivity	15
Electrical Ground Conductivity	0.005
Radio Climate	Continental Temperate
Terrain Irregularity Parameter	90
Tx Siting Criteria	Very Good
Rx Siting Criteria	Random
Variability Mode	Broadcast
Time Variability	50 %
Location Variability	50 %
Situation Variability	50 %

Table 3.4: Parameters for ITM P2P Implementation

Parameter	**Value**
Surface Refractivity	322.9943
Relative Permittivity	15
Electrical Ground Conductivity	0.005
Radio Climate	Continental Temperate
Confidence	50 %
Reliability	50 %

Table 3.5: Broadcasting Band Designation [9]

Parameter	**Band IV**	**Band V**
Frequency Range (MHz)	470 - 582	582 - 862
Broadcasting Channel	21 - 34	35 - 68

Table 3.6: DTT Minimum Field Strength Parameters [9]

Parameter	**Band IV**	**Band V**
Antenna Gain (dBd)(G)	10	12
Feeder Loss (dB)(L_f)	3	5
Noise Figure (dB)(F)	7	7

Chapter 4

Verification and Validation

The verification and validation of research is important for confirming that the tools used were implemented correctly and that the research done was accurate. To ensure that the propagation models were implemented correctly, verification is needed; it is done by comparing the proposed implementations against commercial software. Validation involves ensuring that the proposed solution solves the problem at hand by comparing predicted field strengths against actual field strength measurements.

4.1 Introduction

Verification and validation are essential parts of research. in this chapter, the verification of the propagation models used is presented. How the verification was done, what metrics were used, and what the results are, are topics discussed in this chapter.

The validation of propagation prediction, which is an important part of the Spectrum Observatory, to show that predicting TVWS by using the propagation models is correct, is also included. The validation setup, metrics and the results are discussed in this chapter.

4.2 Verification

This study involves the implementation of computerised propagation models; as defined in [83], the verification of a computerised model necessitates the assurance that the computer programming and implementation of the conceptual model, or in this case, the propagation model, is correct. The verification of the different propagation models shows that the propagation models were indeed implemented correctly and that the results generated from the implementation can be trusted.

4.2.1 Verification Metrics

The metrics used to do the verification of the different propagation models are outlined below. These metrics will be used to determine whether the propagation models were implemented correctly by the author.

Minimum Percentage (MIN) The minimum difference between the results from the proposed implementation and the industry or published solution, as a percentage of the result from industry solution, is used as a metric. The industry solution uses the exact same models, implementing these as publicly available software. The minimum percentage metric is thus seen in Equation (4.1), with B denoting the path loss according to the proposed solution and C indicating the path loss according to the industry implementation.

This metric indicates the minimum difference between the proposed implementation and the published software as a percentage of the answer from the industry software.

$$MIN\left(\frac{|C-B|}{C}\times 100\right) \tag{4.1}$$

Maximum Percentage (MAX) The maximum difference between the two sets of results expressed as a percentage of the industry implementation, conversely, indicates the worst possible outcome. Equation (4.2) thus uses C as the industry solutions and B as the proposed implementation of the model.

$$MAX\left(\frac{|C-B|}{C}\times 100\right) \tag{4.2}$$

The minimum and maximum percentages give the simplest measure of variability as the range of the differences [84].

Mean Absolute Error (MAE) ($\bar{x}$) The mean difference, in dB, between the two sets of results, is expressed in Equation (4.3) [84]. The one set is the proposed implementation, whereas the other set is the standard implementation.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}|x_i| \tag{4.3}$$

where $x_i = |C - B|$ and C is the industry solution and B is the result from the proposed implementation.

Median ($\tilde{x}$) The median of the absolute of the difference set ($C - B$, where C is the proposed implementation and B is the industry implementation) is used along with the standard deviation to determine the distribution of the differences. It is important to note that the median is read from the set of differences that have been ranked from the lowest to the highest difference. It is also worth mentioning that the median was calculated from the absolute of the differences.

$$\widetilde{x} = \begin{cases} \text{single middle value} = \left(\frac{n+1}{2}\right) \text{th value on ordered list} & \text{if } n \text{ is odd,} \\ \text{average of two middle values} = \text{average of } \frac{n}{2}\text{th and } \left(\frac{n}{2}+1\right)\text{th values} & \text{if } n \text{ is even.} \end{cases} \tag{4.4}$$

Standard Deviation (σ) The standard deviation of the difference, in dB, between the two sets of results is expressed in Equation 4.5 [84].

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}} \tag{4.5}$$

where $x_i = |C - B|$ and C is the industry solution and B is the result from the proposed implementation. The number of differences is denoted by N in the above equation and $\bar{x}$ is the MAE. The standard deviation is used with the median difference to determine the variance of the verification results. The standard deviation was also calculated on the absolute difference.

4.2.2 Verification Setup

The implementation of the propagation models will be compared against other existing applications of the models using randomly generated verification test sets, to verify the implementation of the different propagation models. For instance, the verification of the various propagation models will be done by using a dataset of 50 unique propagation parameters that are randomly generated by a python script. These same parameters were previously discussed in Section 2.9. There will be test sets generated for each propagation model because the propagation models have different input parameter ranges.

Free-space Loss Model For the Free-space loss calculation a test set was generated with the following conditions:

Table 4.1: Free-space Input Parameters

Parameter	Range
Frequency(MHz)	$100 \leq f \leq 1500$
Distance (km)	$1 \leq d \leq 20$

The dataset will be put through the author's implementation as well as through software called Spectrum Engineering Advanced Monte Carlo Analysis Tool (SEAMCAT) for verification. SEAMCAT was designed to assess potential interference between different radio communication systems [85]. SEAMCAT is a collaborative project, which includes tools to test specific propagation models. One of the models available to be used is the Free-space Loss Model, which was then used to verify the author's implementation.

Hata-Davidson Model The Hata-Davidson model is an extension of the Hata model. To ensure that both the core model and the extension were implemented correctly, the test set of 50 parameters was split into two. The first 25 parameters of the test set were generated to fit the Hata model and have the conditions as seen in Table 4.2. The last 25 parameters of the test set were generated to make use of the extensions that the Hata-Davidson model adds to the base Hata model. The conditions for the parameters for this part of the test set can also be observed in Table 4.2.

Table 4.2: Hata Input Parameters

Parameter	Hata	Hata-Davidson
Frequency(MHz)	$150 \leq f \leq 1500$	$30 \leq f \leq 1500$
Distance (km)	$1 \leq d \leq 20$	$1 \leq d \leq 300$
TX Height(m)	$30 \leq tx_h \leq 200$	$20 \leq tx_h \leq 2500$
RX Height (m)	$1 \leq rx_h \leq 10$	
Environment	$1 \leq E \leq 4$	

The environmental values used are listed in Table 4.3. The parameter values and their interpretation are discussed in Section 2.9.2.

Table 4.3: Hata Enviroment Values

Value	Description
1	Urban (Small and Medium Cities)
2	Urban (Large Cities)
3	Suburban
4	Open or Rural

The dataset will be put through the proposed implementation as well as through industry created software for verification. The industry software used to verify the Hata and Hata-Davidson model is the software written by the NTIA. The software package is called Land Mobile Service but is also known as the Okumura-Hata Field Strength / Loss Calculator and calculates path loss using the Okumura-Hata Model, the Hata-Davidson Model or the COST 231 extension to the Hata model. The proposed implementation uses the Hata-Davidson model, and therefore the Hata-Davidson model of Land Mobile Service will be used. The NTIA software was written by Fredrick A. Najmy, Jr [86].

ITM Models The ITM model in P2P mode is the only model implemented that makes extensive use of terrain data via a DEM. The ITM model is discussed in detail for both operation modes in Section 2.9.3. The ITM model uses a range of parameters; some of these parameters are the same for both area and P2P modes, whereas some are not. A list of all the parameters and their boundaries can be seen in Table 4.4. The various parameters are discussed in the paragraphs below.

Area Mode. In addition to the parameters shared by both modes of the ITM model, the area mode also has a **terrain irregularity parameter**, which is set to **90 m** as default, as per the ITM recommendations [8]. The area mode has three variability parameters, viz., the time percentage, set to 50%, the location percentage, set to 50%, and the confidence percentage, also set to 50%. The area mode also has different modes of variability; the mode best suited for the proposed implementation is the broadcast mode.

The distance used in the area mode, which is the distance between the transmitter and the receiver, is calculated using the Haversine formula with the GPS coordinates generated for the P2P mode. The Haversine method has regularly been used to calculate the geodesic distance between two points on a sphere and has been shown to be accurate enough [87].

P2P Mode. The P2P mode of the ITM model has two variability parameters, the confidence, set to 50% and the reliability, also set to 50%. These values will ensure a median approach [41]. The P2P mode, inherently, uses an individual variability mode.

Table 4.4: ITM Input Parameters

Parameter	**ITM Area**	**ITM P2P**
Distance (km)	$0 \leq d \leq 1000$	N/A
TX & RX Latitude	N/A	$25.00 \leq TX_{lat} \leq 35.00$
TX & RX Longitude	N/A	$24.00 \leq TX_{lon} \leq 34.00$
Frequency(MHz)	$100 \leq f \leq 1000$	
TX Height(m)	$1 \leq tx_h \leq 300$	
RX Height (m)	$1 \leq rx_h \leq 30$	
Dielectric Constant	15.0	
Surface Refractivity (ppm)	301.0	
Conductivity of the ground (S/m)	0.005	
Radio Climate	Continental Temperate	

The verification for the ITM in both Area mode and P2P was done by using software that was written by the NTIA [88]. This NTIA ITM software provides path loss calculations for the ITM in Area mode and in P2P mode. The NTIA ITM application was developed by Dr Alakanda Paul, Paul McKenna and Fredrick Najmy from the NTIA. This software uses the GLOBE DEM, with 800 elevation points between the two terminals, when making the P2P predictions. All other parameters were set to match the proposed implementation's parameters as discussed above.

The ITM P2P mode was further verified using the Signal Propagation, Loss, And Terrain analysis tool (SPLAT!) package. SPLAT! is an RF propagational analysis tool, which provides detailed information for a RF link between two terminals [89]. SPLAT! was configured to use the same parameters as discussed above, but the SPLAT! software uses the 3 arc-second SRTM DEM.

The results for the ITM Area model are for the proposed implementation against the NTIA ITM software. The ITM P2P model verification includes results for the proposed implementation against the NTIA ITM software, the proposed implementation results against the SPLAT! software as well as the NTIA ITM software against the SPLAT! software.

4.2.3 Verification Results

According to [83], it is important to remember that some of the variations experienced while checking the accuracy of the computer programs may be caused by the computer implementation. Some of the small differences between the proposed implementations and the commercially available software are because of the different programming environments used and the variable sizes. The proposed implementation gives the result with up to six decimals, while the commercial software gives the result with only up to two decimal places.

Free-space Loss Model The results from the verification test for the Free-space loss model is given as maximum difference percentage, minimum difference percentage, mean difference ($\bar{x}$), median difference($\widetilde{x}$) and the standard deviation of the differences (σ) in Table 4.5.

Table 4.5: Free-space Verification Results

MIN	**MAX**	$\bar{x}$	$\widetilde{x}$	σ
0.0051	0.0170	0.0106	0.0116	0.0028

The difference can be ascribed to the different sizes of the variables and the rounding off done by the SEAMCAT software. The minimal differences between the proposed implementation and the commercial software prove that the Free-space loss model was implemented correctly.

Hata-Davidson The results generated from the Hata-Davidson tests can be seen in Table 4.6.

Table 4.6: Hata-Davidson Verification Results

	First 25 Entries	**Last 25 Entries**	**All 50 Entries**
MIN	0.00	0.04	0.00
MAX	0.04	2.76	2.76
$\bar{x}$	0.01	0.8	0.41
$\widetilde{x}$	0.00	0.78	0.04
σ	0.0178	0.6913	0.6261

In the results for the first 25 entries, the entries for the core Okumura-Hata model, the differences are diminutive and can again be ascribed to variable sizes and rounding off. The results for the second 25 entries, which implement the Hata-Davidson extension, show differences of up to 2.76%. This is considerably more and cannot be ascribed to rounding off only; further investigation into the manual of the NTIA software shows that the implementation of the Hata-Davidson model by the NTIA differs from the definitions and equations given in [6]. The differences are however still relatively small, and further investigation led to the discovery that the definition for the Hata-Davidson model by the NTIA, as used in their implementation, differs on two factors in the Hata-Davidson model.

The available software implementations of the Hata-Davidson model are very limited and because the author did not find another commercially available solution, the author also wrote a small excel spreadsheet for the Hata-Davidson implementation to test the proposed implementation. The result for the second test can be seen in Table 4.7.

Table 4.7: Hata-Davidson Second Test Results

MIN	**MAX**	$\bar{x}$	σ	$\widetilde{x}$
0.00	0.00	2.7E-05	1.5E-05	3.3.3E-05

The values in Table 4.7 show that the differences between implementing the Hata-Davidson by the author in C and Excel are minuscule. This gives enough evidence that the Hata-Davidson was implemented correctly.

ITM Area Mode The same test set was used for both the ITM in Area mode and in P2P mode. The results for the Area mode implementation can be seen in Table 4.8.

Table 4.8: ITM Area Mode Verification Results

MIN	MAX	$\bar{x}$	σ	$\widetilde{x}$
0.0002	0.0623	0.03	0.0164	0.02

The results show minimal differences between the author's implementation and the NTIA software. The small differences can be attributed to rounding off of the final answer, which, upon further investigation, the NTIA software does. The minuscule differences verify that the ITM model in Area model was implemented correctly.

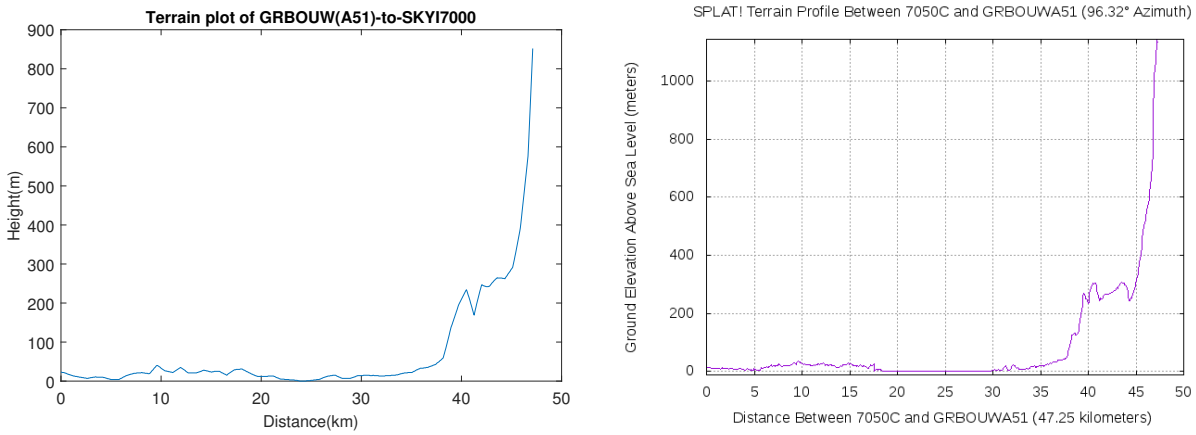
ITM P2P Mode The results of the verification of the ITM P2P Mode can be seen in Table 4.9. The results show minimal differences between the author's implementation and the NTIA implementation, which proves that the ITM was implemented correctly. The results of the author's implementation versus the SPLAT! analysis tool as well as the SPLAT! analysis tool against the NTIA software show significant differences, however. These differences are because the SPLAT! tool uses a 3 arc-second DEM, whereas the author's implementation and the NTIA software uses a 30 arc-second DEM. The more detailed DEM means that the SPLAT! tool can detect terrain peaks or lows that the author's implementation cannot detect, leading to more substantial losses being predicted in general by the SPLAT! tool.

Table 4.9: ITM P2P Mode Verification Results

	NTIA	**SPLAT!**	**NTIA vs SPLAT!**
MIN	0.0002	0.0011	0
MAX	0.0212	22.8019	22.81032
$\bar{x}$	0.0252	4.7955	4.7922
$\widetilde{x}$	0.0235	1.7122	1.735
σ	0.0128	9.6308	9.6319

To illustrate the difference between the available DEMs used, the terrain profile was drawn for the terrain between the analogue transmitter of channel 51 at the Grabouw site, and the arbitrary receiver location in Constantia, Cape Town.

The first terrain profile was drawn using the 30 arc-second GLOBE DEM that is implemented in the proposed system. The second terrain profile was drawn using the 3 arc-second SRTM DEM used in the SPLAT! software. The differences between using the 30 arc-second DEM against the 3 arc-second DEM can be seen in Figure 4.1 and shows the higher level of detail accuracy when working with a more detailed DEM.



(a) GLOBE Terrain Data (30″)

(b) SPLAT! Terrain Data (3′)

Figure 4.1: DEM Path Profile

4.3 Validation

The propagation models used to do TVWS analysis make use of empirical formulas, and some also use terrain data, but how well do the predicted field strengths compare against field measurements? For the validation of the propagation models, the four different propagation models (Free-space, Hata-Davidson, ITM Area model and the ITM P2P model) were used to predict the field strengths of PUs in the area of Cape Town at a specific location where an RF receiver is located. The predicted field strengths were then correlated against the measured field strengths from the receiver in order to draw a conclusion about the validity of using propagation prediction to determine TVWS. The receiver used in the validations is a system from GEW Technologies.

GEW Technologies gave access to their system by offering remote access to one of their computers. A site visit was also done to get a better understanding of how everything works together at their facility.

4.3.1 Validation Setup

The validation process involves comparing the predicted field strengths for a list of transmitters to a nearby fixed receiver. The predicted field strengths are then compared against the measured field strengths to draw validity conclusions. The validation will focus on the protection of the PUs and the accuracy of the models.

Transmitter Details

The first part of the validation setup is to pick the transmitters to be used in the validation study. In order to get trustworthy transmitter locations, one of the receivers from GEW Technologies, the 7050C Spectrum Analyser, was used to find the bearing of prominent TV signals that were detected in the range of 470 - 900 MHz. Three sites were identified with good certainty after consultation with the engineers from GEW Technologies, and these were used for the evaluation, which can be seen in Table 4.10. All of the transmitters used for validation are analogue transmitters, and no DTT transmitters were used. The transmitter power, in Table 4.10, is provided in dBW relative to a half-wave dipole (dBd).

Table 4.10: Validation Transmitter List

Description	A/D	Channel	Freq (MHz)	GPS location	TX Power (dBd)
Tygerberg	Analogue	22	479.25	-33.8747; 18.5961	33.00
Tygerberg	Analogue	30	543.25	-33.8747; 18.5961	30.00
Tygerberg	Analogue	34	575.25	-33.8747; 18.5961	33.00

Receiver

The receiver used to measure the field strengths is an SKY-I7000 from GEW Technologies. The location and height of the receiver setup can be seen in Table 4.11. The setup of the receiver is illustrated in Figure 4.2. The receiver used is a calibrated instrument, but it does not account for all of the cable losses and antenna gains in its current configuration. Cable losses and antenna factor corrections needed to be taken into account, therefore the best theoretical data for the cables and antenna were used to calculate the relevant losses and gains. The received power ($RP(dBm)$) read from the GEW receiver also had to be converted to the received power relative to an isotropic radiator ($RP_i(dBm)$) by using Equation (4.6).

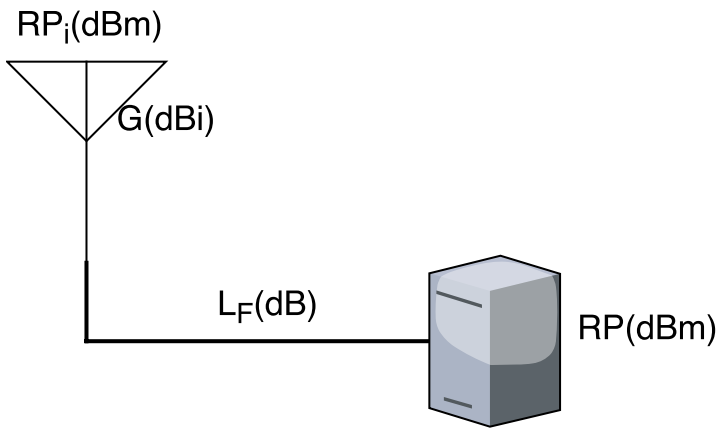


Figure 4.2: GEW Receiver

$$RP_i(dBm) = RP(dBm) + L_f(dB) + G(dBi) \tag{4.6}$$

In Equation (4.6), L_f is the cable losses, in dB, and G is the antenna gain relative to an isotropic radiator, in dBi. The calculation of the cable losses and the antenna gains is discussed below.

Table 4.11: SKY-I7000 Receiver Information

Parameter	Value
GPS Latitude	-34.0557
GPS Longitude	18.4588
Receiver Height	11 m

Antenna Gain ($G(dBi)$) The antenna used by the SKY-I7000 receiver was characterised, and the antenna factor file was received from GEW Technologies. This antenna factor file, which gave the antenna factor (dB/m) at different frequencies, was interpolated, and the antenna factors for the frequencies of the three transmitters were calculated. The antenna factor can be converted to the gain of the antenna relative to an isotropic radiator by using Equation (4.7), because this is a 50 Ω system [90].

$$G(dBi) = 10 \times log \left[2.4 \left(\frac{2\pi}{\frac{c}{f} \times 10^{Af(dB/m)/20}} \right)^2 \right] \tag{4.7}$$

In the above equation f denotes the frequency of operation, $Af(dB/m)$ is the antenna factor, and c is the speed of light. The resulting antenna factors used for the different transmitter frequencies are seen in Table 4.12.

Table 4.12: Antenna Gain Results

Transmitter	Channel	Frequency(MHz)	Gain($G(dBi)$)
Tygerberg	22	479.25	-6.0845
Tygerberg	30	543.25	-6.9885
Tygerberg	34	575.25	-5.8008

Cable Losses ($L_f(dB)$) The calculation of the cable losses for the cables between the receiver instrument and the receiving antenna was done using the length of the cables used and the cable losses per meter (dB/m) from the datasheets of the cables. The cable loss per meter in the datasheet differs according to the frequency of operation, so the information from the datasheet was once again interpolated, and the cable loss per meter for the different frequencies of operation for the three transmitters was calculated and used.

The information regarding the length of the cables and the datasheets for the different cables were provided by GEW Technologies. The total cable loss for frequencies of the transmitters used is summarized in Table 4.13.

Table 4.13: Cable Losses Results

Transmitter	**Channel**	**Frequency(MHz)**	**Cable Loss($L_f(dB)$)**
Tygerberg	22	479.25	0.8436
Tygerberg	30	543.25	0.8970
Tygerberg	34	575.25	0.9232

Field Strength Prediction

The field strengths were predicted using the propagation models under investigation in this research, namely, the Free-space model, the Hata-Davidson model, the ITM Area model and the ITM P2P model. To determine the validity of the way in which these models were implemented in the research being conducted, the models were implemented with the same parameters as discussed in Section 3.2.2.

Measurement

The analogue TV system used in South Africa is the Phase Alternating Line (PAL) system, or more accurately the PAL-I variant [91]. The PAL TV system has the video information amplitude modulated onto the vision carrier signal. The audio information is frequency modulated on the audio carrier, which is placed 6 MHz to the right of the vision carrier. When viewing the TV signal in the frequency domain, it is important to remember that the vision carrier is the carrier for the signal and that the vision carrier contains most of the transmitted power [2]. The vision carrier is found 1.25 MHz to the right of the left-most edge frequency of the 8 MHz channel. The spectrum of the PAL signal, seen in Figure 4.3, shows the different carries as well as the separation between the carriers.

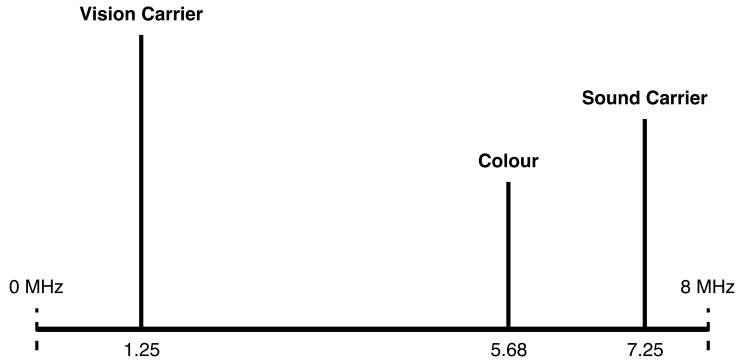


Figure 4.3: PAL TV Signal Spectrum [2]

The signal strength of an analogue signal is measured at the vision carrier. To measure the received power of the analogue signal, the vision carrier was measured by means of a narrowband receiver, taking the amplitude of the vision carrier with the marker of the spectrum analyser. The resolution bandwidth was set to 6k250/10M00 with the wideband receiver of the spectrum analyser focusing on the centre of the analogue TV signal. An example of the spectrum analyser window when measuring an analogue TV signal can be seen in Figure 4.4.

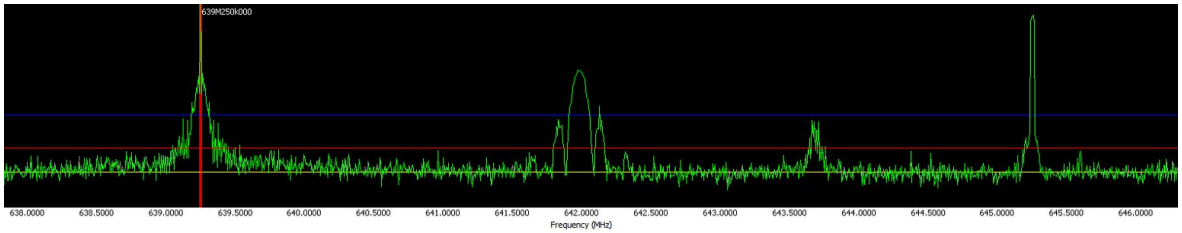


Figure 4.4: Analogue Measurement

Experimental Method

Five received power measurements were taken for every transmitter in Table 4.10; these measurements were converted to field strengths, and the averages of these values were compared against the predicted field strengths. The measurements were taken at different times of the day, at around 17:00, 20:50, 00:00, 08:30 and 12:00.

The results of these measurements in received power were then converted to received power relative to an isotropic radiator using Equation (4.6), and this in turn was then converted to field strength with the equations presented in Section 3.2.7.

4.3.2 Validation Metrics

The metrics used for the validation, where the predicted field strength is compared against the measured field strengths, are discussed in detail in this section.

False Positive Percentage

The most important condition for DSA is protection of the PUs; to ensure that the propagation models used adhere to this requirement, the field strengths predicted by the propagation models must be above the measured field strengths. The situation that must not occur is one where the propagation models predict TVWS while there is no TVWS, a.k.a they may not create a false positive, because this could lead to possible interference to the PU operation. The metric used to measure the false positives (FP) is done by assigning a value of 0 if the predicted field strength($E_{predicted}$) is at least equal or higher to the measured value ($E_{measured}$). A value of 1 is assigned if a false positive is identified, meaning that the measured field strength is higher than the predicted field strength, as seen in Equation (4.8).

$$FP = \begin{cases} 0, & E_{predicted} \geq E_{measured} \\ 1, & E_{predicted} < E_{measured} \end{cases} \tag{4.8}$$

The false positives are then presented as a percentage of the amount of transmitters for the different propagation models.

Root Mean Square Error (RMSE))

The RMSE was used by the authors in [92], during an extensive study on the accuracy of various propagation models. The RMSE gives a representation of the standard deviation of the predicted field strengths vs the measurements made using the spectrum analyser, without taking the sign of the result into account [41].

The equation for the RMSE can be seen in Equation (4.9) [93].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_i^2} \tag{4.9}$$

In Equation (4.9), e is used to show the error between the predicted field strength($E_{predicted}$) and the measured field strength($E_{measured}$) and is taken as $e = E_{measured} - E_{predicted}$ and n is taken to be the number of sites used for validation, which in this case is three.

Mean Absolute Error (MAE)

The MAE ($\bar{x}$) between the predicted field strengths and the measured field strengths is another metric that is used to determine the validity of using propagation prediction to do TVWS determination. The authors in [41] also used the mean error as a metric in their study when determining the accuracy of different ITM variants. A discussion of the equation in which the mean value is used as a statistical parameter to discuss the results is seen in Section 4.2.1.

Many researchers in the field of geosciences, use RMSE as a metric for model error evaluation; however, some avoid it and rather use MAE. In cases where a Gaussian distribution is expected, the RMSE has an advantage over the MAE [93]. In the validation of the propagation models in this study, both metrics were used.

4.3.3 Validation Results

The predicted field strengths for the three transmitters at the receiver can be seen in Table 4.14. In this table and in the following results the propagation models are denoted in the following way:

FSL - Free-space model
Hata - Hata-Davidson model with HAAT considered
ITMArea - ITM Area model
ITMP2P - ITM P2P model

Table 4.14: Predicted Field Strengths (dBuV/m)

Transmitter	**FSL**	**Hata**	**ITMArea**	**ITMP2P**
Tygerberg 22	82.4	63.7	66.5	82.4
Tygerberg 30	79.4	60.4	64.0	79.4
Tygerberg 34	82.4	63.2	67.3	82.4

The field strengths from the measurements taken with the receiver are summarized in Table 4.15, and a bar graph comparing the measured field strengths against the predicted field strengths from the different propagation models can be seen in Figure 4.5. It is evident from the bar graph as well as from the validation results that all of the propagation models over-predict the field strengths for all the transmitters, which means that there were no false positives (FP). The false positive percentage (FP) in the validation results, in Table 4.16, also shows that none of the propagation models had false positives for any of the transmitters. This means that all of the propagation models pass the requirement of not predicting TVWS or DSA when there is in fact still PU activity.

Table 4.15: Measured Field Strengths (dBuV/m)

Transmitter	**17:00**	**20:00**	**00:00**	**08:35**	**12:00**	**AVG**
Tygerberg 22	48.5	48.9	49.0	48.7	48.2	48.7
Tygerberg 30	50.4	51.3	53.7	51.2	48.2	51.4
Tygerberg 34	49.9	51.0	51.3	52.2	48.2	51.4

The other factor that can be tested by meanse of validation is the accuracy of the propagation models, which is done using the RMSE and MAE metrics.

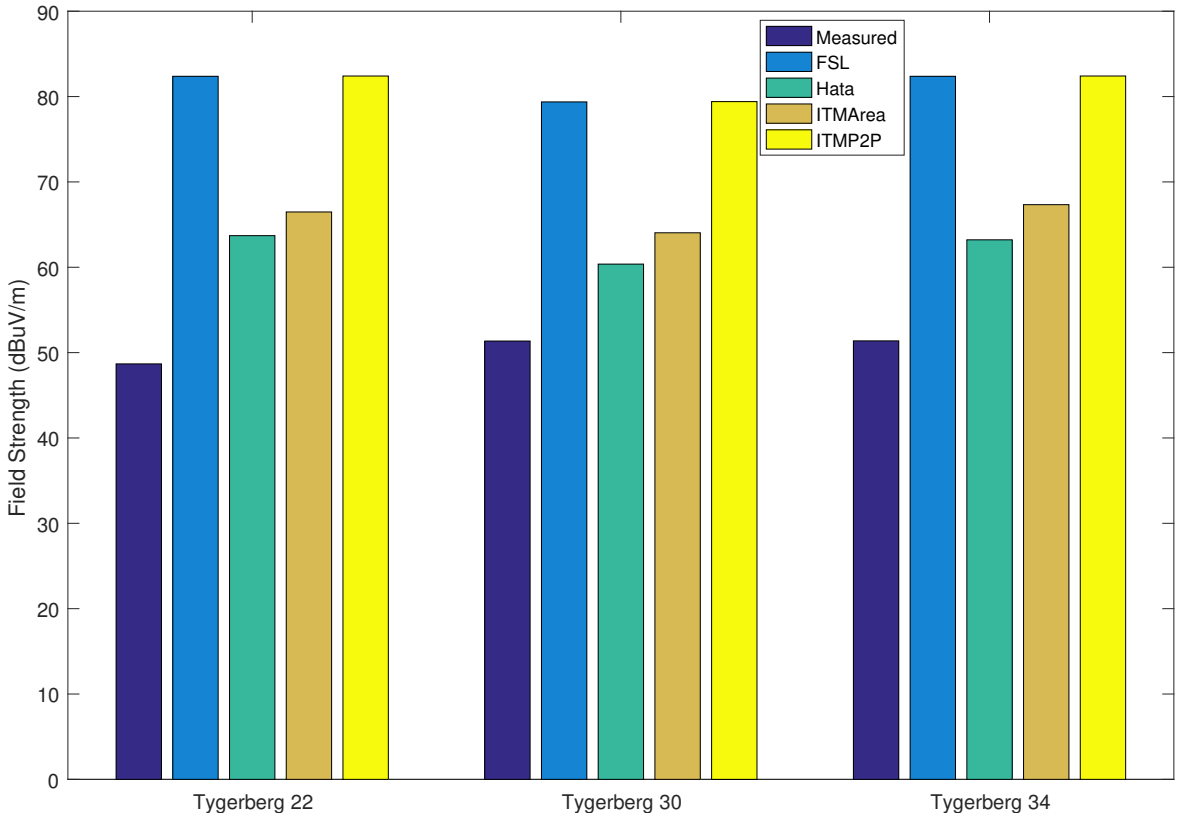


Figure 4.5: Predicted and Measured Field Strengths

Table 4.16: Validation Results

Metric	**FSL**	**Hata**	**ITMArea**	**ITMP2P**
FP (%)	0.0	0.0	0.0	0.0
MAE (dB)	30.9	12.0	15.5	30.9
RMSE(dB)	31.0	12.2	15.6	31.0

The results for these metrics for the different propagation models are presented in Table 4.16. It is encouraging to see that the Hata-Davidson model has an RMSE and an MAE of less than 15 dB.

The ITM Area model shows an MAE of 15.5 dBuV/m and an RMSE of 15.6 dBuV/m. The accuracy of the ITM Area model should improve when the terrain irregularity parameter, discussed in Section 2.9.3 and used to factor in the terrain, is chosen specifically for the type of terrain that is actually near the transmitter.

In the TVWS calculations done in this research, the irregularity parameter was fixed at the world wide average of 90, because the irregularity parameter for each location in the country was not known. The ITM P2P mode was written for the purpose of taking the specific terrain into account.

The validation results for the ITM P2P mode show an MAE and RMSE of 30.9 and 31.0 respectively. The above result is as bad as the result from the Free-space model, with regard to accuracy. Observing the predicted field strength values for the ITM P2P and the Free-space model in Table 4.14 shows that they are identical. The ITM P2P model uses a DEM to model the specific terrain. The terrain plot for the path between the Tygerberg transmitter and the receiver at the measurement site using the 30 arc-second GLOBE DEM, seen in Figure 4.6, shows that there is a line of sight link between the transmitter and the receiver, with no obstacles in between. The ITM P2P thus uses the Free-space loss model for line-of-sight links, but obstacles encountered between the transmitter and receiver are accounted for; this will be made more apparent in the Results Analysis (Chapter 6).

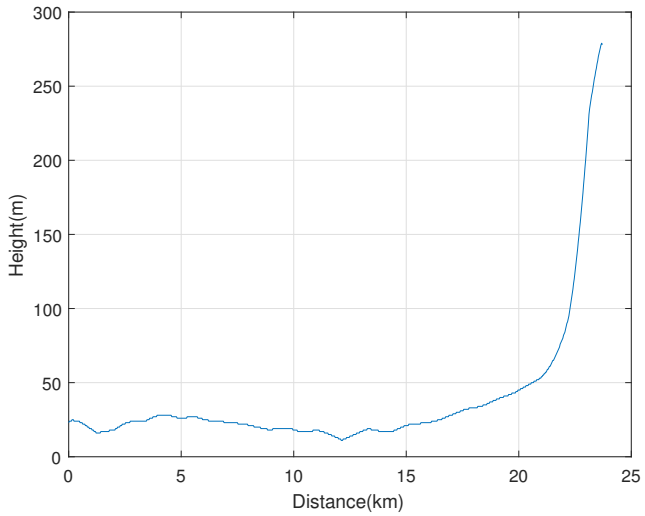


Figure 4.6: Terrain Profile between Tygerberg and Measurement Site

The ITM Area model and the ITM P2P model use statistical parameters (known as variability parameters), which were discussed in Section 2.9.3, as these have an influence on the path loss calculations and thus on the predicted field strengths.

The values for these statistical or variability parameters were set to 50% for the proposed TVWS research in order to arrive at a mean path loss and predicted field strength to do TVWS analysis with (as discussed in Section 2.9.3). However, when these values are changed to be more strict, they will give smaller field strength values. The ITM Area model and the ITM P2P model were used with all the variability parameters set to 90%, and the resulting field strengths and validation results can be seen in Table 4.17 and Table 4.18 respectively.

Table 4.17: Predicted Field Strengths for ITM models (dBuV/m)

Transmitter	**ITMArea (50,50,50)**	**ITMArea (90,90,90)**	**ITMP2P (50,50)**	**ITMP2P (90,90)**
Tygerberg 22	66.5	43.4	82.4	72.5
Tygerberg 30	64.0	40.9	79.4	69.5
Tygerberg 34	67.3	44.2	82.4	72.5

Table 4.18: Validation Results for ITM Models

	ITMArea (50,50,50)	**ITMArea (90,90,90)**	**ITMP2P (50,50)**	**ITMP2P (90,90)**
FP	0.0	100.0	0.0	0.0
MAE (dB)	15.5	7.7	30.9	21.0
RMSE (dB)	15.6	8.0	31.0	21.2

It is evident that the MAE and the RMSE for the ITM with the variability parameters set to 90% are better than they are for variability parameters set to 50% and the predicted field strengths are closer to the measured field strengths. The ITM Area model with variability parameters set to 90% did, however, predict field strengths that resulted in false positives (FP). The prediction of false positives should be avoided, even though it may mean less TVWS. The ITM P2P model with variability parameters set to 90% also resulted in field strengths nearer to the measured field strengths, but the MAE and the RMSE are still 21.0 and 21.2 dBuV/m respectively.

However, this is not such a bad result when compared with other measurement-based research studies. The authors in [73] measured the differences between different propagation models and an increasing number of obstacles between the transmitter and receiver. In this study, the average error for using the ITM P2P model with SRTM data was 15 dB, and for some of the other models, it was as high as 30 dB.

A study was done to bound the error of path loss models, and it showed an RMSE for the different propagation models of between 10 and 45 dB [92]. This shows that differences between measured and predicted field strengths can indeed be observed.

The above validation setup showed that all of the above propagation models effectively over-predicted the field strength of the transmitter, which would result in a loss of TVWS although the PUs would be protected. After all, PU protection is the primary constraint for DSA and thus any of the propagation models can be used to do TVWS prediction. It was also seen that the statistical parameters of the ITM model have an influence on the TVWS prediction and this can be a point of discussion for the regulators regarding what the statistical parameters should be set to.

It should be noted that this is only a point measurement, meaning a single receiver was used; it is possible that different results will be obtained when using more measurement points to measure the received power for the same transmitters, and that this may sketch a very different picture. It may be the case that the specific measurement point used for the validation is a local minimum or maximum of received power because of constructive and destructive interference. The ideal would be to have three or more receivers around the area measuring the same signal over a longer period, and comparing these measurements against the predicted values for every location. The extended measurement period and additional measurement sites are expected to deliver results that would be closer to the predicted median values.

4.3.4 Validation Through Literature

The validity of using propagation models to do propagation prediction for TVWS determination is further encouraged by the literature in the DSA research domain.

The authors of [19] quantified the TVWS capacity for Europe by using a geolocation-based approach, while using the ITM P2P model with a publicly available DEM for their specific region.

A study was done in the USA by researchers from the University of Washington, where the different aspects of determining TVWS were discussed, and the authors used the ITM P2P model to make TVWS predictions [94]. TVWS prediction by using propagation models was proposed as early as 2008 by the authors of [95], where the authors compared different FCC propagation curves for use in TVWS prediction.

4.4 Conclusion

In this chapter, the implementation of the different propagation models was verified by comparing the results from industry software against the results from the proposed implementations, while using a randomly generated set of transmitters and receivers. The verification tests showed that all of the models were implemented correctly.

The second part of this chapter validated the use of propagation models for TVWS prediction by doing field measurements of known transmitters. Although there were discouraging differences between the predicted field strengths and the measured field strengths, all of the propagation models over-predicted the field strengths, meaning that the PUs would not be experience harmful interference.

The usage of propagation models to do TVWS prediction was further supported by examples in literature, where propagation models were used to conduct similar studies on TVWS.

Chapter 5

Computational Performance Analysis

The computation node of the Spectrum Observatory is discussed in this chapter. The amount of speed-up obtainable by using an HPC as the computational node is investigated in this chapter. The HPC used in this Spectrum Observatory is a High Performance Cluster Computer (HPCC) from the CHPC, which is a division of the CSIR, called the Lengau cluster.

5.1 Introduction

One of the problems with geolocation based DSA is that the prediction of received power takes a considerable amount of time. The part of the prediction that takes the most time is the path loss prediction. This is especially true when a terrain-aware prediction model is being used. This part of the research thus aims to answer the second research questions, as expressed in Section 1.3, by investigating the speed-up that is possible when using an HPC for the propagation prediction.

The various options for the computational node have been discussed in Section 2.10, and the HPCC was identified as the preferred option because of its scalability and its increased performance, and the possibility of building on existing code. The scalability of the HPCC means that, if more computing resources are needed, more computing nodes can be added to the cluster.

The HPCC is also able to do the trigonometric computations needed, with less compromise in accuracy than would be experienced when using an GPGPU, and it does not have virtualisation overheads, as would be the case with cloud computing, thus leading to better performance than would be experienced with cloud computing. Usage of the HPCC as the computation node for the Spectrum Observatory is further encouraged by the fact that the CHPC, which is part of the CSIR, made its Lengau HPCC available to researchers, free of charge.

The Lengau HPCC makes use of the Lustre file system; the specifications for the Lengau HPCC can be seen in Table 5.1, along with the specifications of the local PC that was used for comparison in these experiments.

Table 5.1: Specifications for Computing Architectures

	Local PC	**Lengau HPCC**
CPU	Intel i7-6700	Intel Xeon (R) E5-2690 V3
CPU Clock	3.6 GHz	2.6 GHz
CPU Cores	4	24192
Number of Nodes	1	1008
Min CPU Cores Avail	1	1
Max CPU Cores Avail	4	240
*Rpeak	N/A	1006 TFlops
*Rmax	N/A	782.9 TFlops
Memory	8 GB	126 TB
Interconnect	N/A	FDR Infiniband Network
Storage	100 GB	4 PB

**Rmax = Max number of floating point operations*
**Rpeak = Theoretical max value of floating point operations [96]*

5.2 Experimental Setup

An experimental set-up is needed to compare the relative performances of different computing architectures, to determine the advantage of using an HPCC, which is a parallel approach, over a serial approach. The experiments entail using four different propagation models (ITM in P2P mode, ITM in Area mode, the Hata-Davidson model and the Free-space loss model) to calculate the received powers from terrestrial TV transmitters for the Gauteng area, and then to measure the computational performance of both architectures. Identical datasets were used for all the tests done.

The received powers are calculated on a logical grid with a grid spacing of 30 arc-seconds and boundaries defined by the Gauteng area, which is defined by the GPS coordinates, 25.1° S to 27° S and 27.1° E to 29.1° E. The grid spacing of 30 arc-seconds results in a grid of 240 horizontal/longitudinal cells by 228 vertical/latitudinal cells, which results in 54 720 cells that need to be calculated. This is significantly less than the 3 182 400 (2040 x 1560) cells that would be needed for the whole country, but it nonetheless serves as a useful indicator of scalability. The received powers are calculated for the centre of each cell and stored as such, as discussed in Section 3.2.4.

The number of TV transmitters used for the calculations ranges from 1 transmitter to 2325 of the transmitters in the PU database. When all the transmitters are used for the Gauteng area, it means that 127 224 000 (54 720 x 2 325) propagation predictions would need to be computed for each of the propagation models.

The propagation models, discussed in Section 2.9, were adapted for parallelisation on an HPCC. The Free-space propagation model does not have any input parameter constraints, and it was thus implemented for the whole investigation area of Gauteng.

The Hata-Davidson model is valid for distances up to 300 km for predictions and for antenna heights, which were listed in Table 2.3; however, these ranges were ignored, and the calculations were still performed for parameters falling outside of these boundaries to get an upper boundary of the time it would take to do the calculations. The antenna heights used in this computational study was the antenna height AGL.

The ITM has parameters that are implemented in both the P2P and Area mode, and these parameters can be observed in Table 5.2. The parameters that are unique to the P2P mode are the confidence level (set to 50%), the reliability (set to 50%) and the number of points used for the terrain extraction (set to 800).

Table 5.2: Parameters for ITM for Computer Experiments

Parameter	**Value**
Surface Refractivity	301
Relative Permittivity	15
Electrical Ground Conductivity	0.005
Radio Climate	Continental Temperate

The parameters used solely by the ITM in Area mode can be seen in Table 5.3 along with the values for the parameters as they were used for the computational experiments. The frequency, transmitter heights (as AGL) and polarisation are again read from the transmitter list, consisting of the 2 325 transmitters.

Table 5.3: Parameters for ITM in Area mode for Computational Experiments

Parameter	**Value**
Terrain Irregularity Parameter	90
Tx Siting Criteria	Random
Rx Siting Criteria	Random
Variability Mode	Broadcast
Time Variability	50%
Location Variability	50%
Situation Variability	50%

5.2.1 Architecture Benchmark Experiment

This benchmark experiment is created to compare the computing architecture at different stages of performance. This experiment consists of three tests as described below. The first test compares the individual CPU performance of each architecture, while the second test compares the speed-up when the maximum amount of CPUs is used on the local machine, The third test seeks to investigate the effect of internode communication between the two nodes used on the HPCC.

Serial Performance

The first test involves testing the serial performance of each computing architecture. The four different propagation models was used to compute the field strengths in the Gauteng area resulting from four transmitters. Both the local machine and the Lengau HPCC did the calculations with 1 CPU, while still utilising MPI.

Baseline MPI Performance

The second test exploits the most straightforward MPI implementation, where four CPUs are used for each computing architecture to calculate the field strengths for the defined prediction area for four transmitters. The source code is written in such a way that each CPU would be responsible for a transmitter on both the local PC and the HPCC. This benchmark shows the MPI performance of each architecture.

48 Core Parallelism

The next test seeks to highlight the advantages of parallelism by using 48 CPUs on the Lengau HPCC and the maximum of four CPUs on the local computer to calculate the received power from 48 transmitters.

As seen in Table 5.1, the Lengau HPCC has 24 CPUs per node and the minimum parallel usage, to experience inter-node communication, is when only two of the available ten nodes are used. This results in 48 CPUs; since there are 48 transmitters, this implies that one transmitter would be allocated to each CPU on the Lengau HPCC. On the local computer, the 48 transmitters will be divided among the four CPUs of the PC, giving every CPU 12 transmitters to calculate. The area of calculation is once again the defined prediction area.

5.2.2 HPCC Scalability Experiment

Different propagation models scale differently across the increase of the number of CPUs. In this experiment, the field strengths from 2325 transmitters in the database were calculated for the Gauteng area for the different propagation models using an increasing number of CPUs on the Lengau HPCC. The benchmark test was done by running the different propagation models using 48, 96, 120, 144, 192 and 240 CPUs. The results from using the HPCC were compared to the same calculation done on the local PC using only 4 CPUs.

5.2.3 Metrics Used

The metrics used for these tests are latency and speed-up. For each experiment, the latency from the local PC (T_σ) and the latency from the HPCC (T_π) will be recorded and the speed-up (S) will be calculated. Every test will be repeated three times, to build a test case, and the results averaged to get a test result, which is in-line with common performance analysis practice [72]. Running each test three times is needed because of the fact that the computers used, both the local computer and the HPCC, will not be in identically the same state for each run of a test. Care was taken to keep the computers as close to identical as possible but some factors are outside of the control of the programmer. The metrics were discussed in Section 2.10.1.

5.3 Results of the Computational Performance Analysis

Following the execution of the benchmark experiments discussed in Section 5.2.1 and Section 5.2.2 the results from these different benchmark experiments are discussed in this section. The discussion starts with the results for the Architecture Benchmark Experiments and then goes on to the results of the Scalability Experiment.

5.3.1 Architecture Benchmark Experiment

The tests were executed, and the speed-up of each test was calculated from the measured latency for each test. The speed-up for the architecture benchmark experiment can be viewed in Figure 5.1. The serial test measures the performance of a single CPU on the HPCC versus a single CPU on the local machine. Refer to Table 5.1 for the specifications of each of the computing architectures, which show that the local machine has CPUs with faster-operating frequencies than the CPUs used in the HPCC. For a single CPU, the local PC ran faster, and that is why, with all the propagation models, no speed-up is observed, $S < 1$.

The Baseline MPI test shows the speed-up that occurred when the local machine used 4 CPUs against when the HPCC computer used 4 CPUs. The scale-up is similar for the Free-space, Hata-Davidson and ITM Area mode implementations because although the local PC has CPUs with faster-operating frequencies, the HPCC was created with the purpose of exploiting parallelism. The ITM P2P mode showed no speed-up, which will be discussed later in the chapter.

The 48 Core test shows the power of parallelism versus a serial architecture. The local PC used the maximum available number of 4 CPUs while the HPCC used the maximum of 24 CPUs per computation node, with two nodes being applied, thus adding up to 48 CPUs on the Lengau HPCC. This test was done to measure the effect of inter-node communication. For all the propagation models, except for the ITM P2P model, there was notable speed-up, as seen in Figure 5.1.

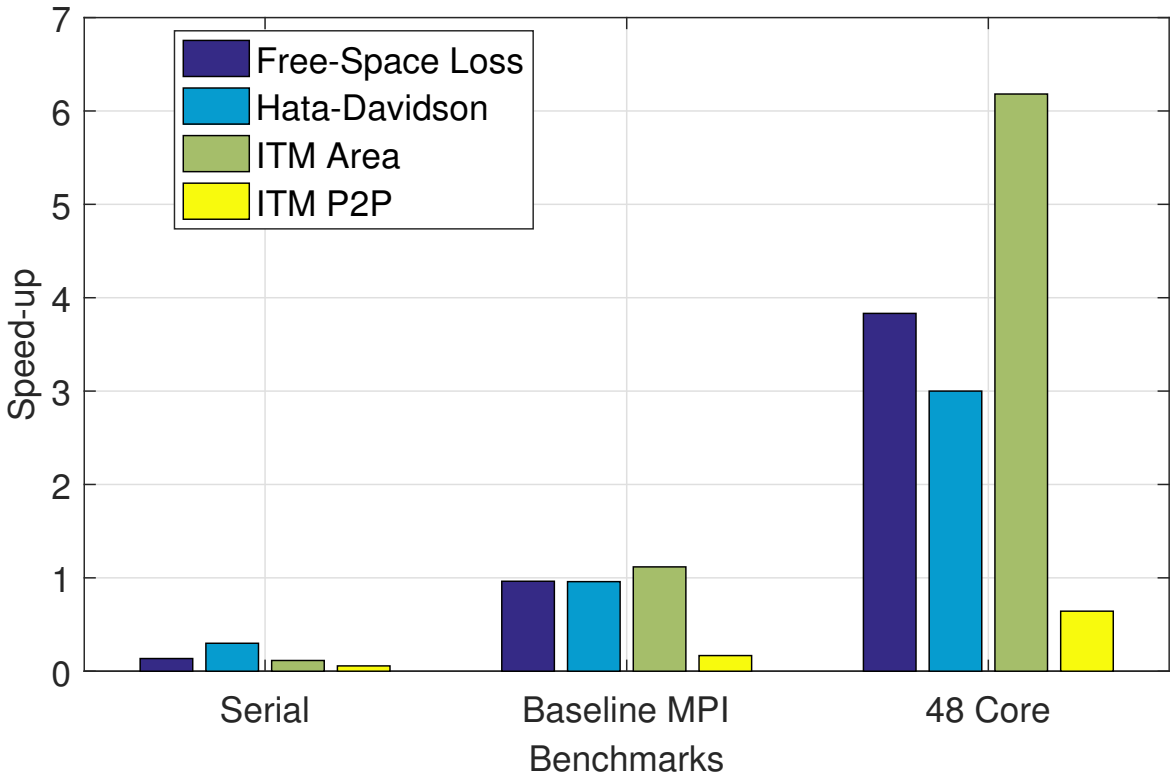


Figure 5.1: Architecture Benchmark Experiment

To understand the lack of speed-up with the ITM P2P model observed in the Baseline MPI test and the 48 Core test, a profiling tool was used to analyse the ITM P2P implementation. Tuning and Analysis Utilities (TAU) is a profiling tool that can be used to measure program performance on an HPCC. The TAU tool was used to analyse the ITM P2P for 48 transmitters with 48 CPUs. The TAU tool ran profiling on all the CPUs, and the mean results are shown in Table 5.4.

The results show that the function using the most of the execution time is the *get_GLOBE_data()* function, which is called by the *GLOBE_elevation()* function, which is part of the *get_GLOBE_pfl()* routine. The *GLOBE_pfl()* function is responsible for creating the path with elevation data between the transmitter and the point under consideration. The *get_GLOBE_data()* function is a simple function that reads the elevation data for a particular coordinate out of the DEM file.

The times in the table shows that most of the time is being lost to file I/O on the Lengau HPCC and that this has an effect on the speed-up for the ITM P2P mode. This also explains why the ITM P2P mode implementation did not show a favourable speed-up for the Baseline MPI and 48 Core tests.

Table 5.4: TAU Results for ITM P2P

Percentage of Time	**Exclusive Time (ms)**	**Total Inclusive Time (ms)**	**Name**
100.0	29	35:15.784	.TAU application
100.0	1.392	35:15.754	main()
85.7	30.684	30:13.981	get_GLOBE_pfl()
84.3	11:12.851	29:43.234	GLOBE_elevation()
52.4	18:29.601	18:29:601	get_GLOBE_data()
13.9	4:54.596	4:54.596	MPI_Barrier()

To counter the file I/O issue for the ITM P2P model, the DEM data was moved from the DEM file to the memory of the HPCC to minimise access to the DEM file. The copying of the entire DEM file is possible because HPCCs normally has large memory capacities, to such an extent that the entire DEM file would readily fit into the memory of the HPCC. This resulted in the improvements seen in Table 5.5; the time it took to do the calculations decreased from five and a half hours to a mere four and a half minutes.

Table 5.5: Path Extraction Improvements

	Before	**After**
Transmitters	2 325	2 325
CPUs	240	240
Wall Time	05:26:29	00:04:22
CPU Time	1291:06:02	17:17:00

5.3.2 HPCC Scalability Experiment

The advantage of the Lengau HPCC over the local PC is that 240 CPUs can be used on the HPCC against the maximum of 4 on the local PC. Speed-up is calculated, as described in Section 5.2.3, and observed in Figure 5.2. In Figure 5.2 the red line denotes the speed-up for the local PC, which is 1, whereas the blue line shows the speed-up for the HPCC.

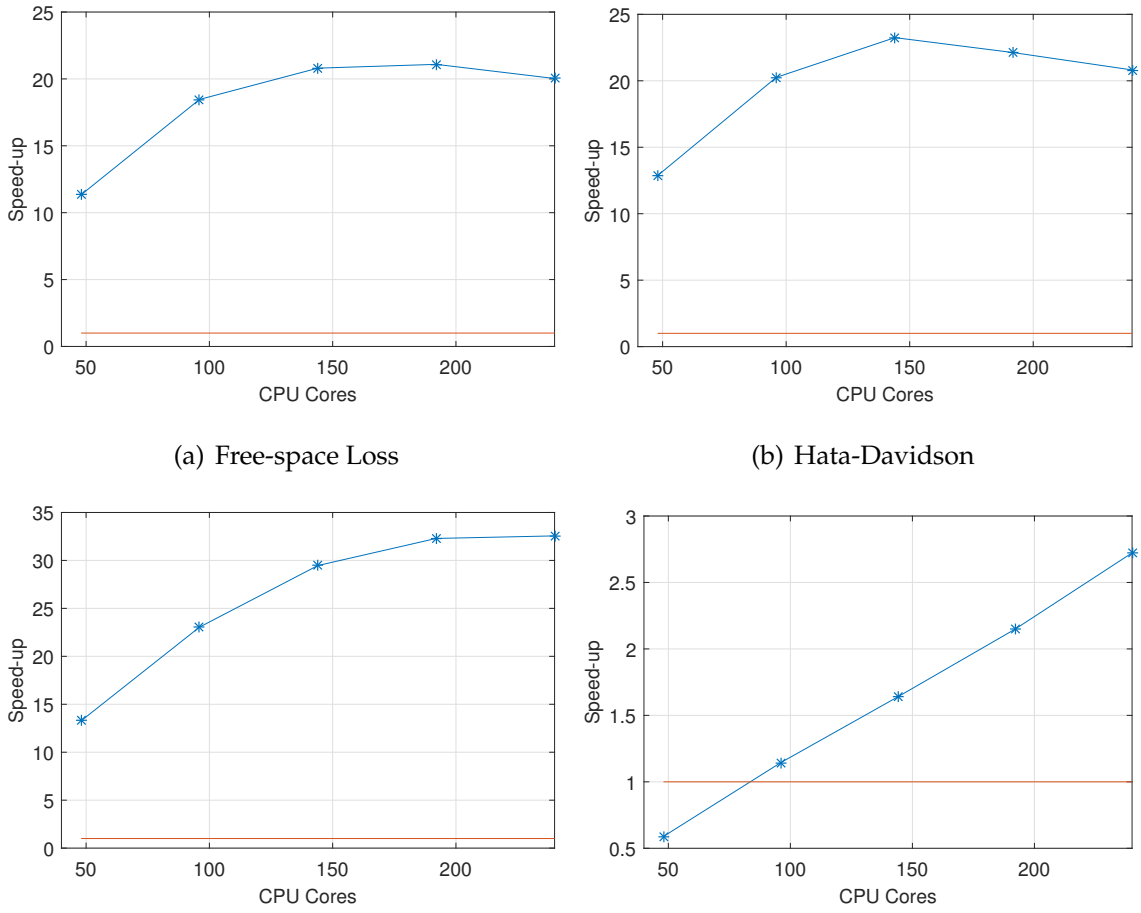
There is significant speed-up for all the propagation models, except for the ITM P2P implementation, as seen in Figure 5.2 at the end of the chapter. The results of the profiling done using TAU led to the identification of file I/O as a factor in the lack of speed-up in the ITM P2P mode implementation. Fortunately, the large amount of memory available on the HPCC enabled the movement of the DEM to the memory of the HPCC, which led to less file I/O and greater speed-up for the ITM P2P model.

Following the movement of the DEM to the memory of the HPCC, the HPCC Scalability Experiment was run again for the ITM P2P, as described in Section 5.2.2, with the file I/O improvements. Figure 5.3 at the end of the chapter shows the results from the HPCC Scalability Experiment; it is evident that exploiting the large memory capacity of the HPCC led to a huge computational performance improvement.

The scaling of the speed-up for the different propagation models also tells an intriguing story. The increase in speed-up is not completely linear for any of the propagation models. As more CPUs are employed, additional MPI protocol communication overhead is experienced, which, along with the time it takes to store the generated data on a finite number of storage places, leads to an upper boundary in the speed-up increase for every problem. However, the speed-up for the ITM P2P model is more linear than it is for the other models. This phenomenon cannot be explained yet, but it can be investigated in the future.

5.4 Conclusion

This chapter investigated the use of parallel computing on an HPCC to do large-scale propagation prediction. The study was done by performing a series of tests, where four propagation models were used, of which all were written to use MPI, and two were ported from other languages, to evaluate the performance of the HPCC. All the propagation models, including the ITM P2P mode after improvements, show favourable speed-up and scalability, which confirms that an HPCC can be used as the computation node for the Spectrum Observatory.



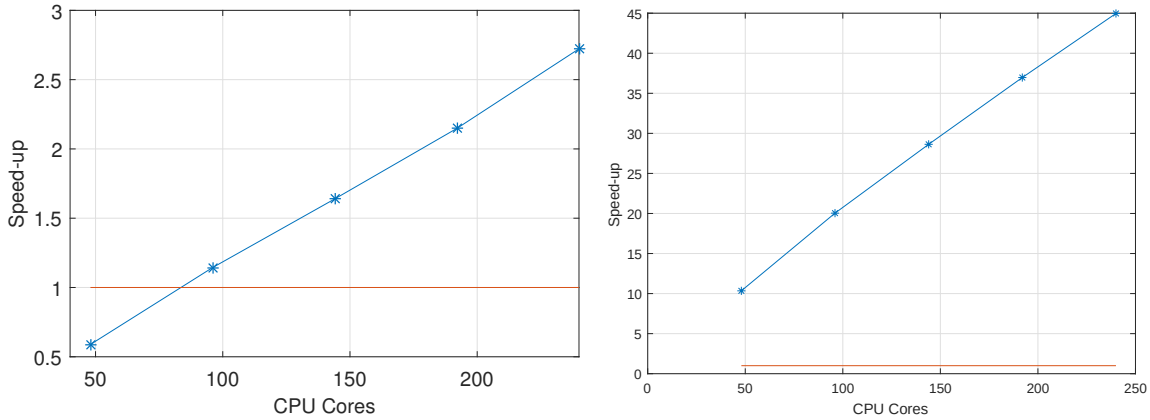
(a) Free-space Loss

(b) Hata-Davidson

(c) ITM Area

(d) ITM P2P

Figure 5.2: Lengau HPCC Scalability Experiment



(a) ITM P2P w/o Improvement

(b) ITM P2P with Improvement

Figure 5.3: HPCC Scalability Experiment after I/O Improvements

Chapter 6

Results Analysis

The calculated received power and amount of TVWS are presented and analysed in this chapter. The amount of TVWS is aggregated by area, and the results are presented as the mean number of channels available per area. CCDF plots as the TVWS channel availability per area are also presented and discussed.

6.1 Introduction

The results, obtained from calculating the received power for a transmitter, are displayed to compare the path loss calculated by means of the different propagation models. The maximum received powers are aggregated according to channel for the different propagation models (Free-space Loss, Hata, ITM Area and ITM P2P) and TV technologies (Analogue, DTT and MDTT) and an example result is displayed and discussed.

The maximum received power is converted to field strength, on which the TVWS policy is applied to classify the cell as TVWS or not for the applicable channel. The TVWS data for the different TV technologies (Analogue, DTT and MDTT) are then combined to form the TVWS data for the dual-illumination period, in which South Africa is currently operating. The CCDF and mean available channels for each province in South Africa are then calculated and used to analyse the TVWS availability for the different propagation models.

6.2 Received Power Prediction

The first observable deliverable from the Spectrum Observatory are the received powers calculated for every transmitter in the PU database, which are stored in the custom binary file (.fst). A script was written to generate an image and a Keyhole Markup Language (KML) file from the custom binary file containing the received power for every transmitter, which allows researchers to view the received powers for the relevant transmitter graphically as a face validation on Google Earth. The script is written in such a way that, the stronger the signal, the nearer to red the area would be, and the weaker the signal, the nearer to dark blue the area would appear. To show the difference between the four propagation models, the image and the KML files for the analogue TYGERBERG transmitter of channel 22 is generated from the relevant received power (.fst) files for all four propagation models. The KML files for other transmitters and channels are available as part of the Digital Artefacts in Appendix A. The resulting coverage images can be observed in Figure 6.1 at the end of the chapter, where the maximum received power is -1.69277 dBm and the minimum received power is -226.070 dBm.

Notice that the Free-space loss model predicts the least path loss. All the models, except the ITM in P2P mode, are used as point-to-area models, with no terrain path profile used, and therefore the resulting coverage plot is circular in nature.

It is evident from the contour of the coverage plot of the ITM P2P model that the ITM in P2P mode takes into account the detailed terrain when predicting the path loss.

The second deliverable of the propagation prediction using the Spectrum Observatory is a file that contains the wanted or maximum received power per channel for the whole country (.fsc). These files are aggregations of the received power for the transmitter files, and the aggregation is per channel for the different propagation models and TV technologies (Analogue, DTT and MDTT). A script was written to visualise the wanted received powers, much the same as the script that was written to visualise the .fst files. This provided a KML file that visually displays the received power on Google Earth for face validation. Again, the stronger the signal, the nearer to red the area is, and the weaker the signal is, the nearer to dark blue the area is. The wanted received powers for channel 22 of the analogue transmitters of South Africa can be seen in Figure 6.2 at the end of the chapter, where the maximum received power was 32.43 dBm, and the minimum received power was -240.00 dBm. Observe that the Free-space loss model again predicts the strongest signals overall.

6.3 TVWS Prediction Heatmaps

The next deliverable are the channels available for DSA as TVWS, after the WS policy, defined in Section 3.2.8, has been applied to the received powers for the country as a custom binary file (.wsc). Heatmaps were generated to show the amount of TVWS channels available for the different TV technologies (Analogue, DTT and MDTT) combined. For example, the heatmaps for four propagation models for dual illumination for South Africa can be seen in Figure 6.3 and Figure 6.4 at the end of the chapter, where the minimum amount of TVWS channels available for a location is 0, and the maximum amount of channels is 48. The heatmaps include a colour bar as the legend.

The amount of TVWS available when doing propagation prediction with the Free-space loss model is almost zero for most of the country, and therefore the heatmap for the Free-space loss model was omitted.

The low TVWS channel count for using the Free-space loss model is to be expected because the amount of loss it predicts on a signal is not as much as the other propagation models. It is clear to see the effect of the terrain-aware model when looking at the heatmap of the ITM P2P model in Figure 6.3 at the end of the chapter, because the contours of the ground can be seen in the predicted TVWS. It is, however, still difficult to distinguish which propagation model predicted more TVWS from the heatmaps and that is why statistical methods were used to analyse the TVWS further.

6.4 TVWS Analysis per Area

To analyse the TVWS that is identified as being available when using the different propagation models, the metrics discussed in Section 3.2.11 are used. This means that only the area that forms part of South Africa as per the 2001 census, which excludes the SKA core and the central advantage areas, is used for analysis. Note that this means that the area stated in Section 3.2.10 as part of the exclusion zones, is excluded from the analysis region.

The TVWS that were predicted for the different TV technologies (Analogue, DTT and MDTT) were combined to give the TVWS available for the area of South Africa as the dual-illumination TVWS results. This is because South Africa is still in a dual-illumination stage, because some DTT and MDTT transmitters are being installed while some analogue transmitters are still operating.

The amount of dual-illumination TVWS available is analysed by using the mean number of TVWS channels available in that particular area as explained in Section 3.2.11. The dual-illumination TVWS per area is also analysed using CCDF plots for the different provinces of South Africa, as well as for South Africa as a whole.

Mean TVWS available Quantifying the amount of TVWS available per province for the different propagation models as the mean number of channels available for that area gives insightful information about the availability of TVWS. This is one of the metrics that will be used to discuss the TVWS available for the different provinces, as well as for South Africa as a whole.

CCDF The Complementary Cumulative Distribution Function (CCDF) graph of an area provides the probability of the amount of TVWS channels available for that area. To be more specific, the CCDF gives the probability that the amount of TVWS will be a certain amount or higher. The CCDF, which is the complement of the Cumulative Distribution function (being the integral of the probability density function), is often used to show how much time an Orthogonal Frequency-Division Multiplexing (OFDM) signal spends above a certain power level, as in [97].

The CCDF plots for the TVWS channels available per area for South Africa as well as for the nine provinces in South Africa are seen in Figure 6.5 and continued in Figure 6.6 at the end of the chapter. The CCDF plots and the mean channels available per area, as set out in Table 6.1, display the results for the following propagation models:

FSL	-	Free-space loss model
HATAE	-	Hata-Davidson with HAAT used for the effective antenna height of the transmitter
ITMArea	-	ITM Area model
ITMP2P	-	ITM P2P model
HATA	-	Hata-Davidson with only height AGL used as the effective antenna height of the transmitter

The results for the mean number of channels per area as well as the number of channels available as a percentage of the maximum 48 UHF channels under consideration are presented in Table 6.1.

The Hata-Davidson model was run for the effective height of the transmitter being taken as the height of the transmitter mast added to the HAAT (HATAE) and the effective height of the transmitter being taken as only the height AGL.

Table 6.1: TVWS Available by Area during Dual-Illumination

Area	FSL	HATAE	ITMArea	ITMP2P	HATA
South Africa	0.0 (0%)	30.4 (63%)	41.2 (86%)	42.1 (88%)	39.6 (82%)
Western Cape	0.1 (0%)	29.9 (62%)	42.3 (88%)	43.9 (91%)	41.3 (86%)
Eastern Cape	0.0 (0%)	24.3 (51%)	39.8 (83%)	42.2 (88%)	38.0 (79%)
Northern Cape	0.0 (0%)	39.3 (82%)	44.8 (93%)	44.7 (93%)	43.3 (90%)
Free-State	0.0 (0%)	29.4 (61%)	39.5 (82%)	39.7 (83%)	36.7 (76%)
KwaZulu-Natal	0.0 (0%)	21.8 (46%)	37.0 (77%)	41.4 (86%)	35.9 (75%)
North-West	0.0 (0%)	29.5 (61%)	38.1 (79%)	37.7 (79%)	35.7 (74%)
Gauteng	0.0 (0%)	19.9 (41%)	33.7 (70%)	36.5 (76%)	32.2 (67%)
Mpumalanga	0.0 (0%)	23.7 (49%)	38.7 (81%)	41.7 (87%)	36.7 (76%)
Limpopo	0.0 (0%)	27.6 (57%)	41.8 (87%)	40.5 (84%)	41.5 (86%)

All of the results show that there is hardly any TVWS available when using the Free-space loss model to do propagation prediction. As stated earlier, this is due to the tendency of the Free-space loss model to over-predict the received power.

The discussion of results will focus on South Africa as a whole as well as discussing the results of the provinces that exhibit interesting phenomena.

6.4.1 South Africa

The mean number of TVWS channels available for South Africa is more than 30, which indicates that much of the spectrum is still available in South Africa to be used for DSA. This translates to 240 MHz (30 $\times$ 8 MHz) of the spectrum that could be used for DSA. The CCDF curve of TVWS per area available for South Africa, as illustrated in Figure 6.5(a), shows that the Hata model using only AGL antenna height, the ITM Area model and the ITM P2P model predict at least 27 TVWS channels with 100% probability. The difference between using the effective antenna height for the Hata-Davidson as only the antenna height AGL versus implementing the HAAT is evident with the Hata-Davidson when using the antenna height AGL as the effective antenna height having 9.2 more TVWS channels available opposed to the Hata-Davidson when using the HAAT for the effective antenna height.

To investigate the difference between the Hata-Davidson model with HAAT considered and the Hata-Davidson model without HAAT, the heatmaps for the these two approaches were plotted and depicted in Figure 6.4 at the end of the chapter. This shows how the usage of the HAAT leads the Hata-Davidson model to predict far greater received powers for the PU and less TVWS. The amount of TVWS in itself is not necessarily a measure of the usefulness of a propagation model; the accuracy of the model is more important. More about this will be in the Future Work in Section 7.2.

6.4.2 Northern Cape

The CCDF for the TVWS available per area in the Northern Cape, in Figure 6.5(d), shows that there are at least 35 TVWS channels available with a probability of 1; the mean number of TVWS channels is the highest compared to the other provinces, at a mean amount of 39 TVWS channels available. It is interesting to note that the CCDF curves for the ITM P2P model and the ITM Area models follow each other very closely, more so than for the other provinces, which is explained by the fact that the Northern Cape has a much less mountainous surface area than do the other provinces. This also leads to the Hata-Davidson with HAAT CCDF curve being closer to the other propagation models. This is also the case for the Free-State results, as well as for the results from the North-West province.

6.4.3 KwaZulu-Natal

The TVWS results of KwaZulu-Natal contrast significantly to the results from the Northern Cape. The KwaZulu-Natal CCDF curves (see Figure 6.6(d)) show a marked difference between the Hata-Davidson with HAAT and the other propagation models, with a TVWS channel count of only 9 with a probability of 1 and a mean TVWS channel availability of 21.8 (46%). The CCDF curves of the ITM P2P model also deviated from the Hata-Davidson without HAAT and from the ITM Area model.

This fact shows that terrain-aware models, like the ITM P2P model, are able to exploit the ground truth of the terrain and identify TVWS that would have been missed with other propagation models. The CCDF curves of the Hata-Davidson without HAAT and the ITM Area model follow each other closely and both of them only implement the transmitter antenna height as AGL.

6.4.4 Gauteng

The mean number of TVWS channels available for Gauteng is the lowest for all the propagation models. The most number of channels available at 100% confidence is 36.5 (76%), when using the ITM P2P model. The CCDF curves for the dual-illumination period for Gauteng, in Figure 6.5(c), show that there are fewer TVWS channels available overall than in the other provinces. The Hata-Davidson with HAAT curve (HATAE) shows considerably less TVWS than do the other models.

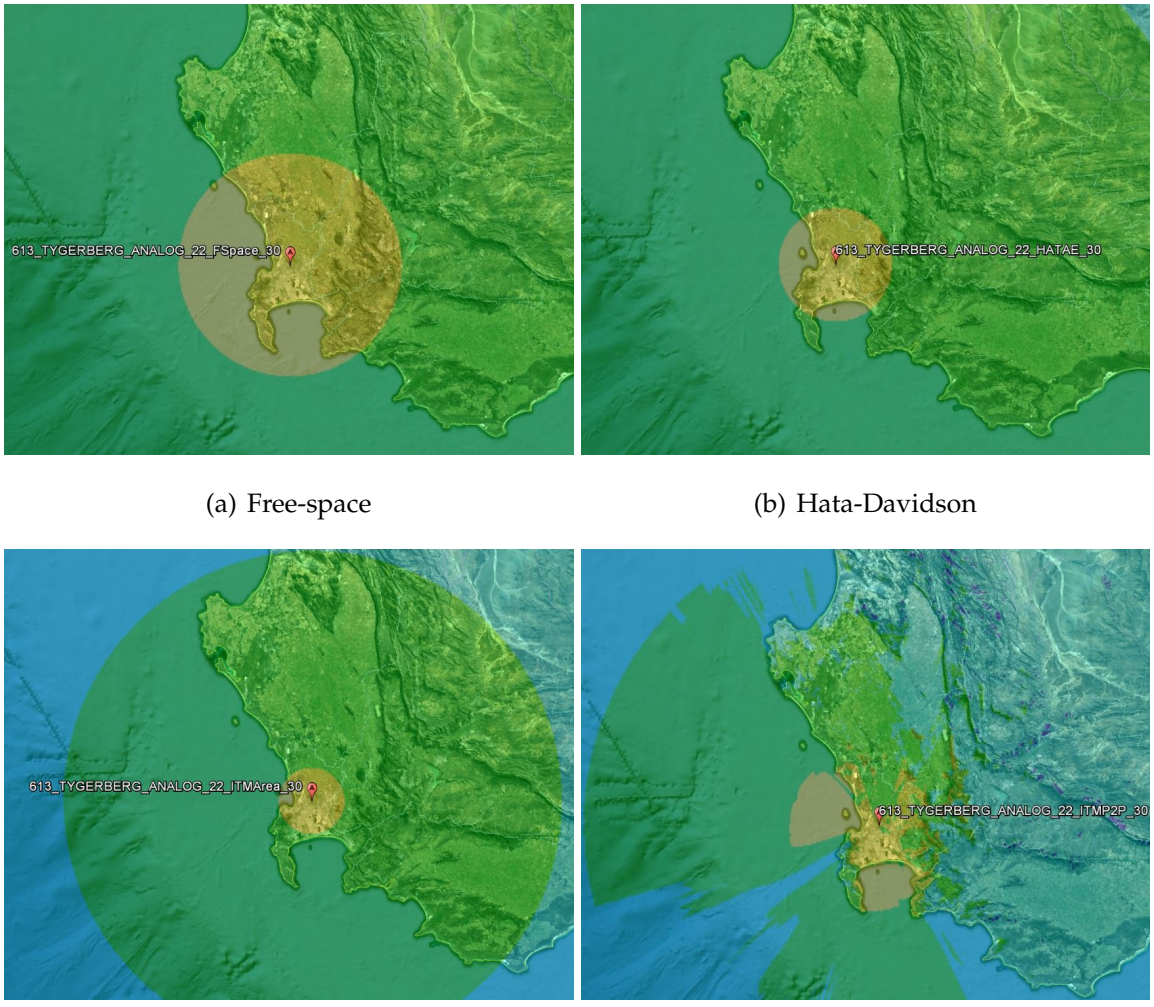
Protection Ratio Influence

The results discussed above were generated using a co-channel protection ratio or a contour of 0 dB to get a upper limit on the amount of TVWS channels available. In practice, regulators also factor in a co-channel as well as an adjacent channel protection ratio, to protect TV receivers against emissions from TVWS devices.

The amount of TVWS channels available for the ITM in P2P mode with a protection ratio of 17 dB was calculated to investigate the influence of a co-channel protection ratio on the number of TVWS channels available. The co-channel protection ratio was chosen to align with the Ofcom regulation of 17 dB, as stated in [27] and the resulting TVWS heatmap (a) and CCDF plot (b) is displayed in Figure 6.7.

6.5 Conclusion

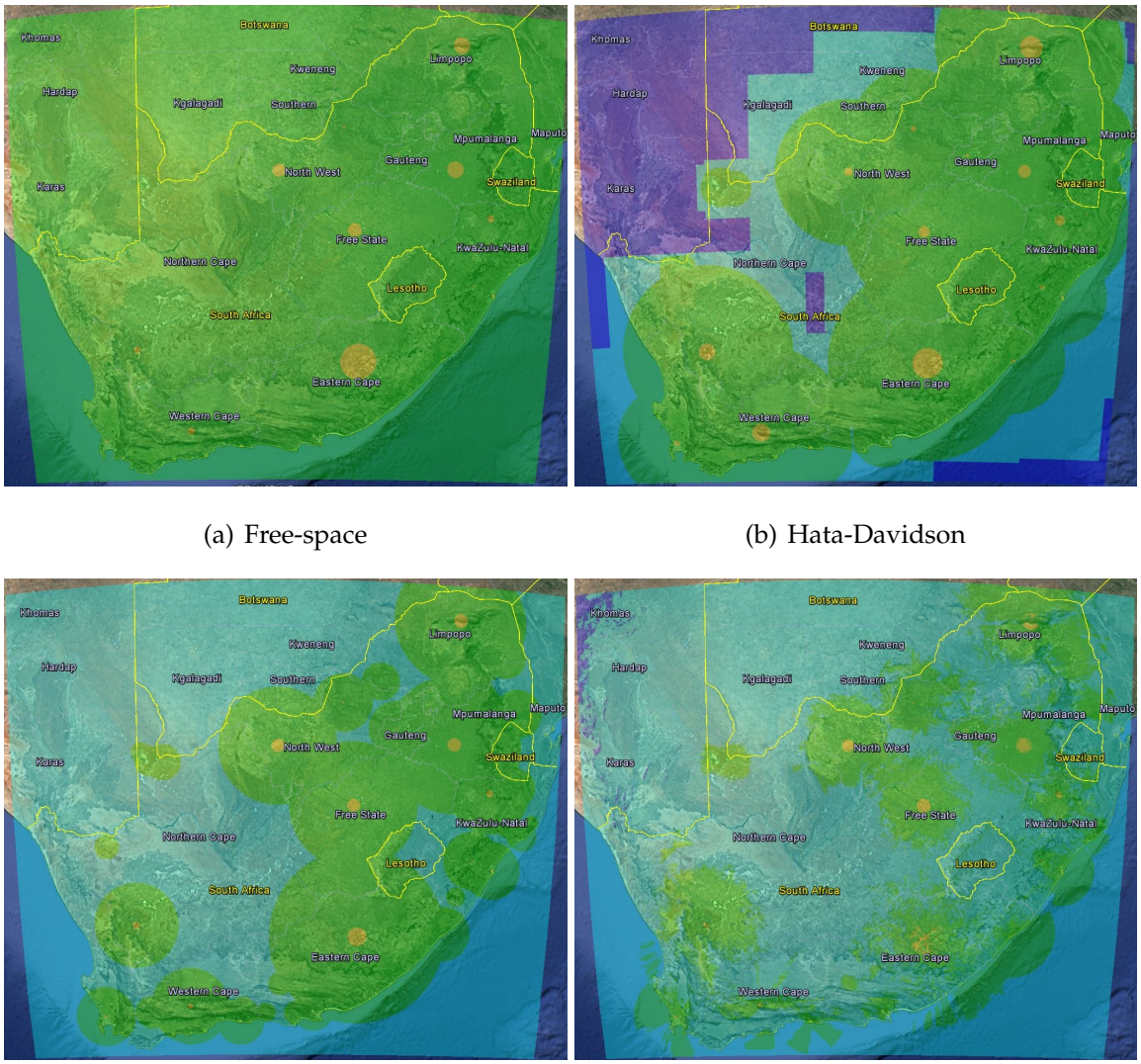
This chapter analysed the TVWS results generated using the Spectrum Observatory. The TVWS available per area was presented in the form of CCDF curves for all the provinces in South Africa as well as for the country as a whole. The results showed that, for areas with a flatter or less mountainous surface, like the Northern Cape, the terrain-aware models do not have any real advantage over the other empirical propagation models. However, for areas with an irregular terrain, like KwaZulu-Natal, there is a definite advantage for using the terrain-aware ITM P2P model over the empirical models. Nonetheless, note that, although the aggregate results show no real difference, there may still be local minima and maxima present.



(a) Free-space (b) Hata-Davidson

(c) ITM Area (d) ITM P2P

Figure 6.1: Received Power for TYGERBERG_ANALOG_22 transmitter



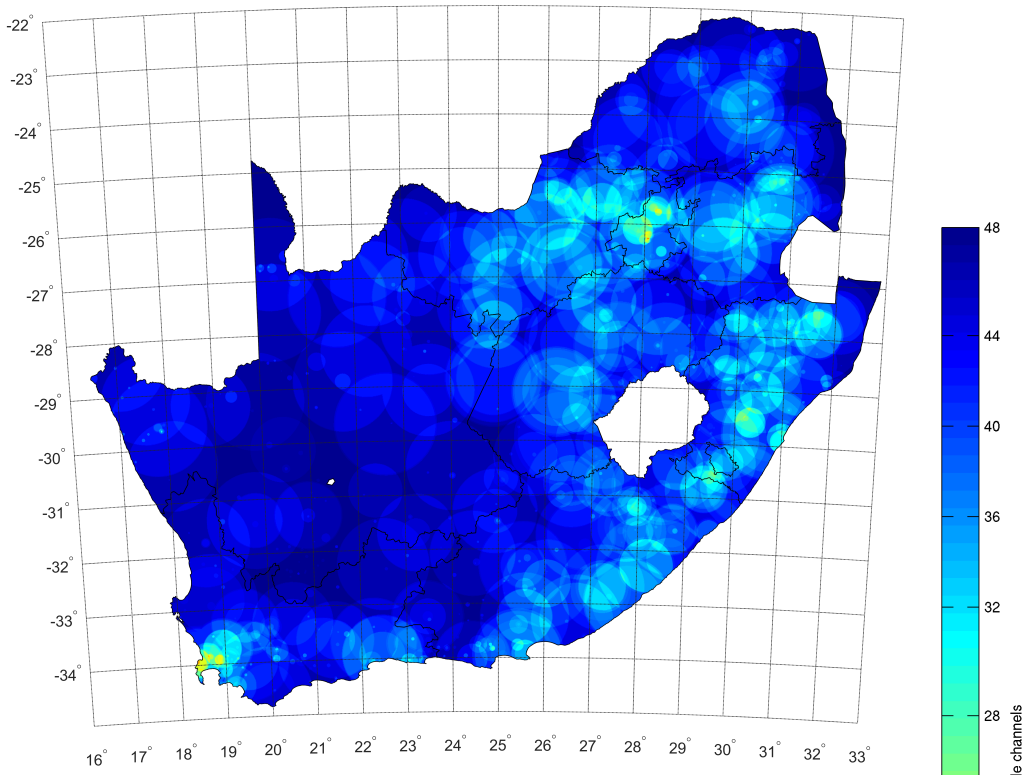
(a) Free-space

(b) Hata-Davidson

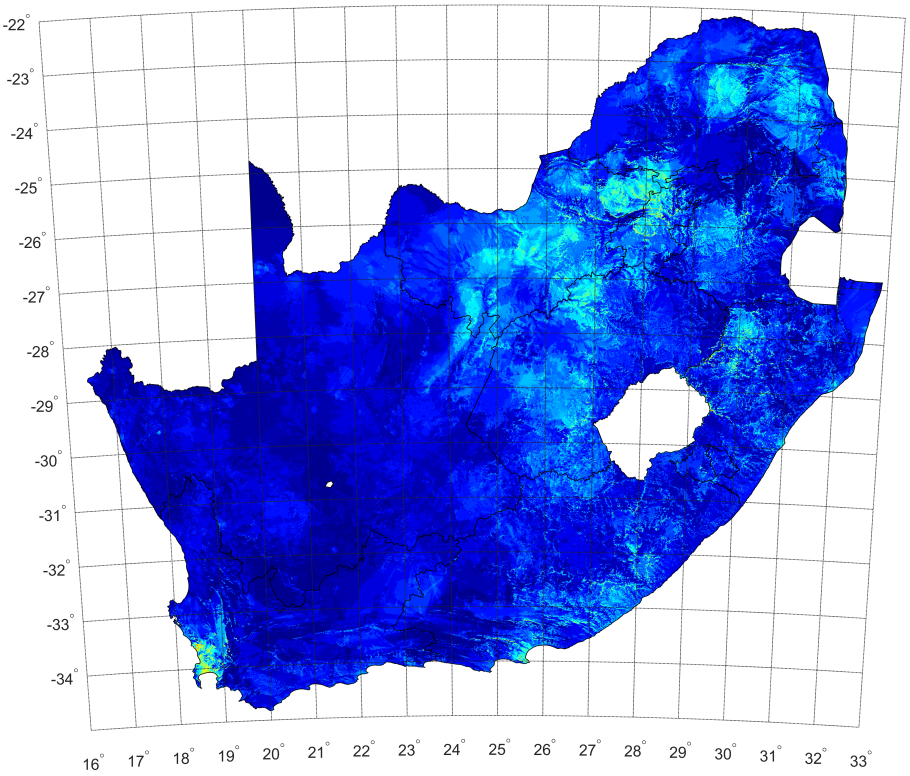
(c) ITM Area

(d) ITM P2P

Figure 6.2: Predicted Received Power for Channel 22 Analogue Transmitters across SA

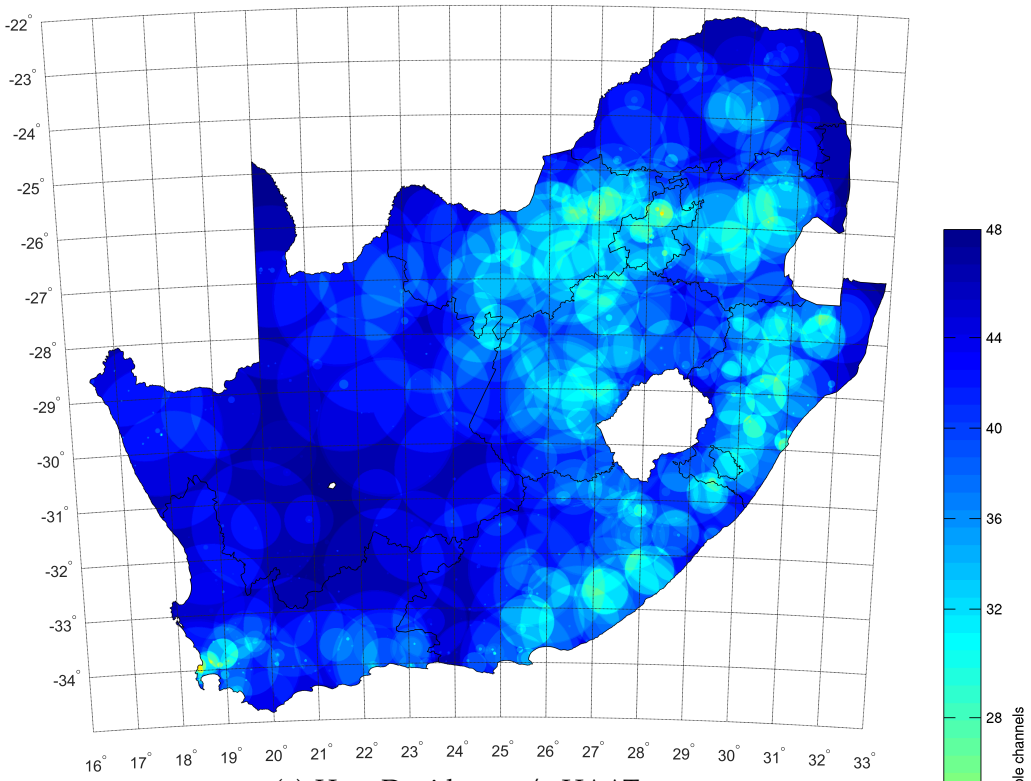


(a) ITM Area

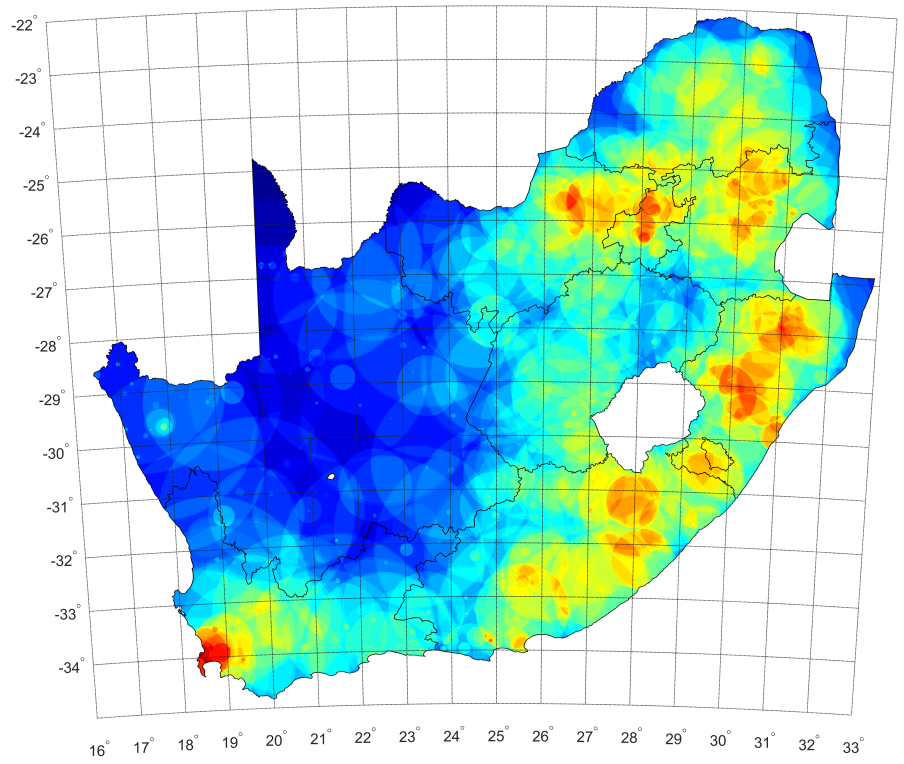


(b) ITM P2P

Figure 6.3: TVWS Heatmaps for Dual Illumination

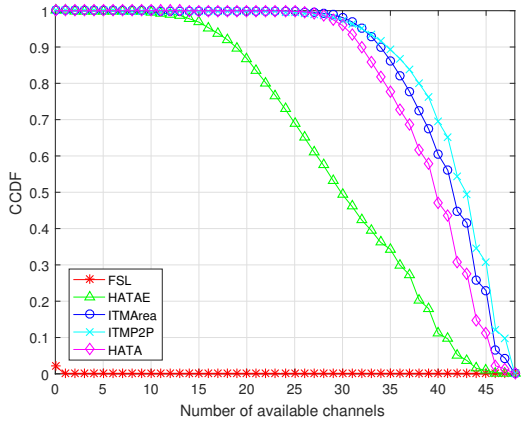


(a) Hata-Davidson w/o HAAT

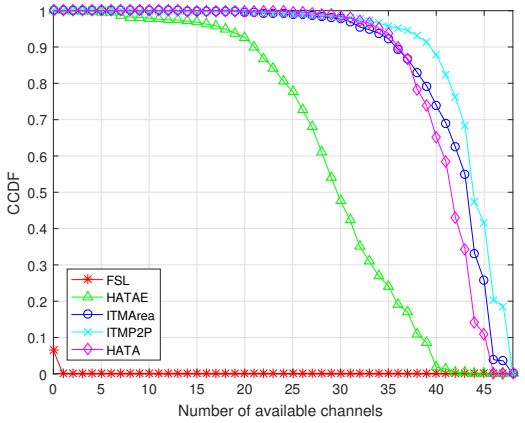


(b) Hata-Davidson with HAAT

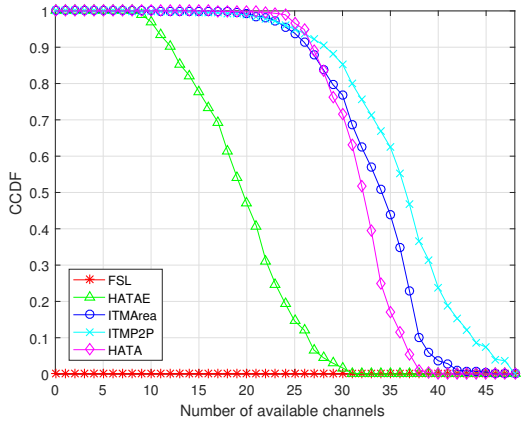
Figure 6.4: TVWS Heatmaps for Dual Illumination Cont.



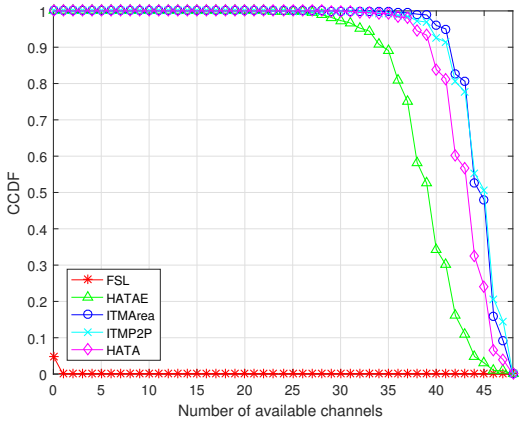
(a) South Africa



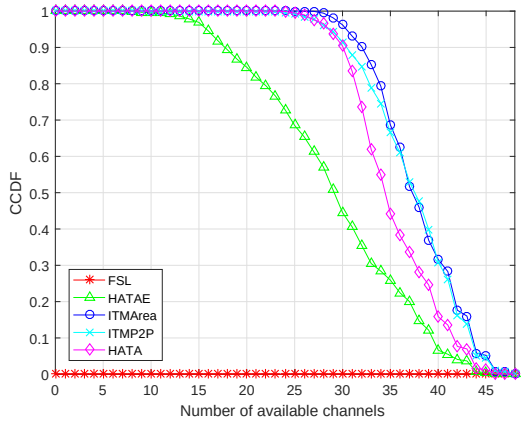
(b) Western Cape



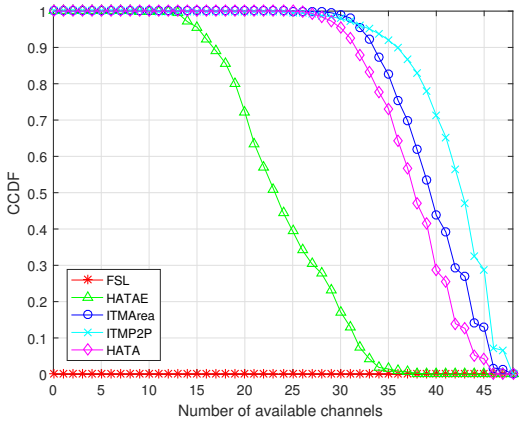
(c) Gauteng



(d) Northern Cape

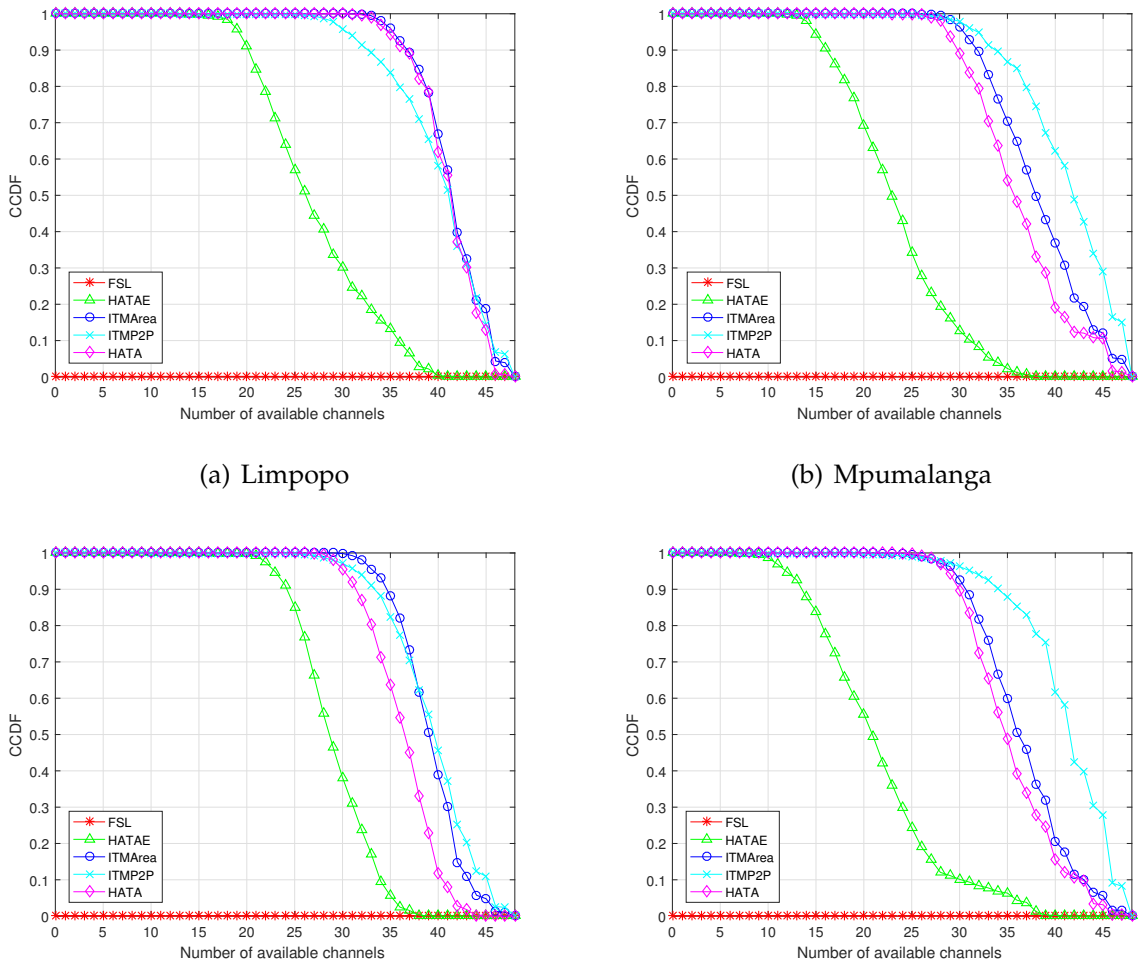


(e) North-West



(f) Eastern Cape

Figure 6.5: CCDF of TVWS per Area



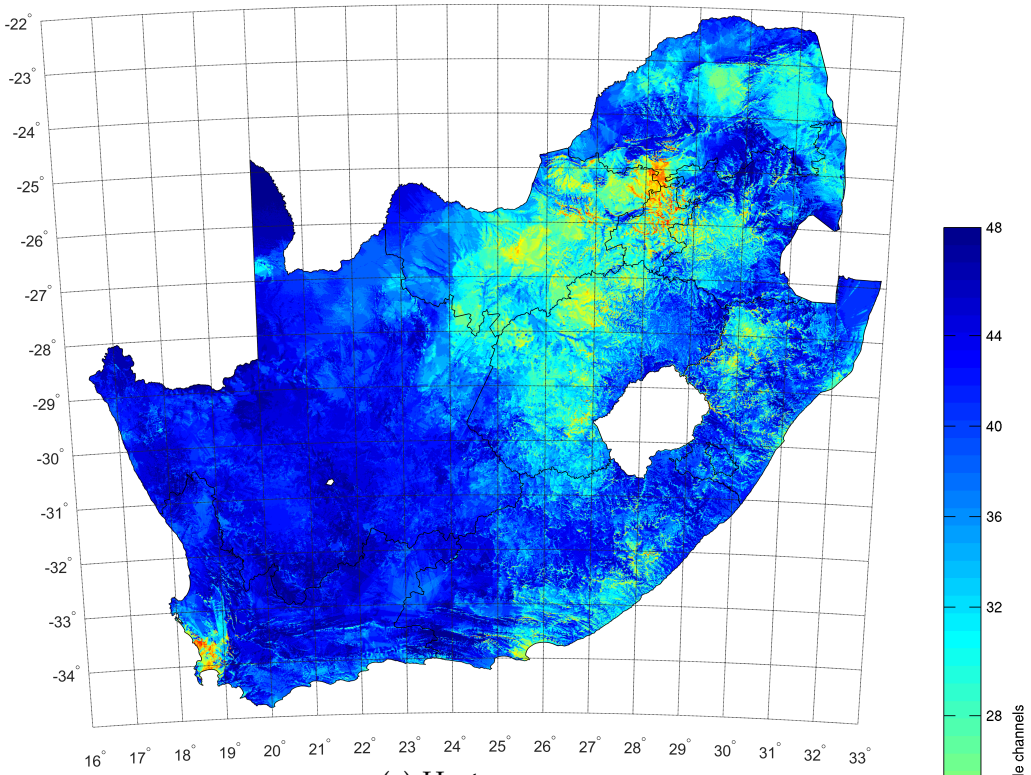
(a) Limpopo

(b) Mpumalanga

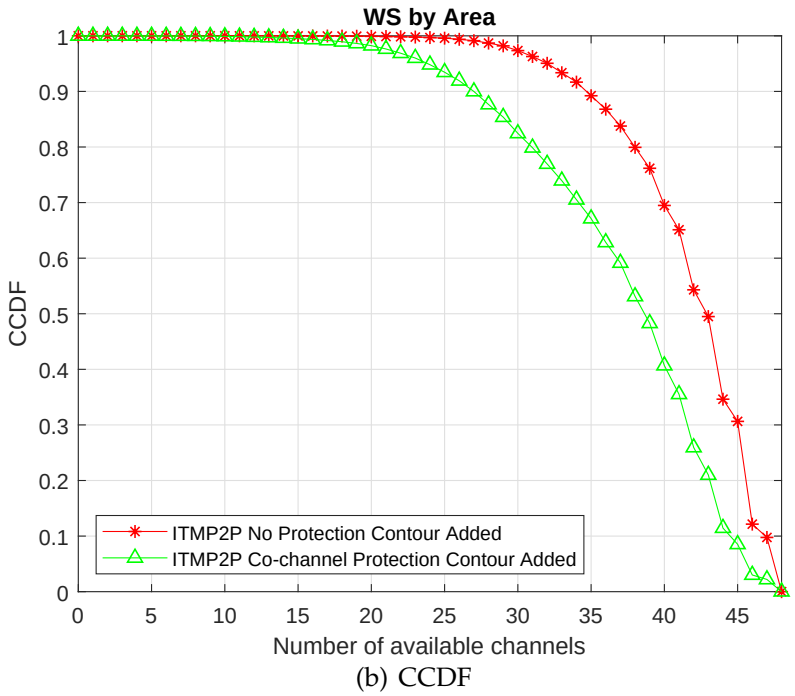
(c) Free-State

(d) KwaZulu-Natal

Figure 6.6: CCDF of TVWS per Area Cont.



(a) Heatmap



(b) CCDF

Figure 6.7: Results for ITM P2P with Co-Channel Protection Ratio

Chapter 7

Conclusion and Recommendations

An overview of the dissertation is given in this chapter and the results analysed in the dissertation are discussed. The research goals stated in Chapter 1 are revisited, and the contributions made through this research are stated. Recommendations are made for future work following this research, and the dissertation is concluded with some final remarks.

7.1 Research Overview

The aim of the research was to answer two research questions, as stated in Section 1.3. To answer the first research question regarding the effect of the propagation model on the amount of TVWS available, the Spectrum Observatory was used to calculate the number of TVWS channels available for South Africa, weighted by area.

The research process began with a detailed literature study, in Chapter 2, of relevant aspects (such as DSA, geolocation databases, TVWS, propagation models, computation architectures, parallel programming and related work done in the research field).

The literature study was followed by a discussion of the system model, in Chapter 3, implemented on the Spectrum Observatory, which was used to do TVWS research. This was followed by an in-depth verification (in Section 4.2) of the propagation models used, and validation, in Section 4.3, of using propagation models to do TVWS research. The next step was to discuss the benchmarking experiments that were used to quantify the speed-up from using HPCC as the computation node for the Spectrum Observatory, as presented in Chapter 5. This was followed by an analysis of the results, in Chapter 6, from the TVWS study, using the system model explained earlier.

South Africa is still in the process of switching from analogue TV to DTT and MDTT, with all three technologies being present at this stage, which is why the research assumed the dual-illumination period and thus sought protection for all the TV technologies (Analogue, DTT and MDTT) when predicting TVWS. The transmitter or PU information used was provided by ICASA, as discussed in Section 3.2.1.

The propagation models, parallelised, implemented and compared against each other, are the Free-space loss model, the Hata-Davidson model, the ITM Area model and the ITM P2P model. These propagation models were used to calculate the path loss between the PU and the different points on the grid system, defined in Section 3.2.4 by making use of the Spectrum Observatory.

The predicted path loss was used to calculate the received power and the received field strengths for a theoretical 10-meter high receiver antenna. The calculated field strengths were compared against the minimum field strengths required for the PU to function without harmful interference and on this basis, TVWS prediction was done. In the case of the predicted field strength being below the minimum field strength, the cell under consideration would be classified as TVWS.

The amount of TVWS channels available was weighted per area, as discussed in Section 3.2.11. The resulting amount of TVWS was described in Section 6.4 in the CCDF plots, as was the mean number of channels available in Table 6.1 for the dual-illumination period.

The results showed that, on average for South Africa, the ITM P2P model predicted the most TVWS available. This ITM P2P model was the only fully terrain-aware model used, and the TVWS predicted by this model that was missed by the other models can be ascribed to the fact that terrain was factored in.

The drawback of using a terrain-aware propagation model is that it is more computationally intensive than empirical models, which is why the second research question involved quantifying the amount of time saved when using the HPCC as the computation node.

To measure the computational improvement from using the HPCC as the computational node of the Spectrum Observatory, the experiments described in Section 5.2 were conducted. They showed speed-up for all of the propagation models when using HPCC. The initial experiments did not show not as much speed-up for the ITM P2P as it did for the other propagation models, but after I/O improvements had been made it too showed tremendous speed-up gains.

The core of the research done in this study were the propagation models used, and therefore the verification involved testing whether the propagation models had been implemented correctly. The validation involved assuring that propagation models can be used to do TVWS prediction by comparing predicted field strength values against measured field strength values. The validity of using propagation models when doing such TVWS predictions is further supported by examples of related work done in the literature.

7.1.1 Revisiting the Research Objectives

The research goals listed in Section 1.4 as part of the research question regarding the amount of TVWS available when using different propagation models were addressed in the following way:

- The PU information was extracted into a file that is easily readable by the different propagation models on the Spectrum Observatory.
- The path loss was calculated using the four propagation models (Free-space loss model, Hata-Davidson, ITM Area model and the ITM P2P model) and followed the received powers which were also calculated.
- The received powers calculated were stored in the custom binary file (.fst) for every PU in the transmitter list. The specification document is contained in Appendix A and details the various file formats described in this section.
- The predicted received powers were aggregated according to channel and stored in another custom binary file (.fsc).
- Custom binary files containing only the interference information (.infc) after the received powers had been aggregated, were also created and stored as part of the Spectrum Observatory.
- The channels available as TVWS were determined from the wanted receive powers (.fsc) by applying the TVWS policy, and they were stored as a custom binary file (.wsc).
- The TVWS channels available per area as well as the mean number of channels available per area for the different propagation models were analysed using CCDF plots.

The research objectives addressed the research question regarding the amount of speed-up that could be expected from using an HPCC as the computation node for the Spectrum Observatory. This was done by measuring the latency for the different propagation models when doing propagation predictions for the area of Gauteng by using a local PC and an HPCC. The measured latency was then used to calculate the speed-up when using the HPCC as the computation node for the Spectrum Observatory.

7.1.2 Functionality added to the Spectrum Observatory

The research done for this dissertation delivered knowledge that contributes to current research being done on DSA, and it added certain functionality to the Spectrum Observatory. The functionality added to the Spectrum Observatory being developed is listed below:

- Utilisation of the Lengau HPCC as computation node for the Spectrum Observatory.
- Parallelisation and implementation of the empirical propagation models (Free-space and Hata-Davidson) for usage on the HPCC.
- Parallelisation of the ITM models (Area and P2P mode) for usage on the HPCC.
- Addition of the four propagation models (Free-space, Hata-Davidson, ITM Area and ITM P2P) to the Spectrum Observatory.
- Addition of the TVWS policy to the Spectrum Observatory.
- Creation of the custom binary file formats for the received power per transmitter (.fst), the aggregated wanted received powers (.fsc), the interference received powers (.infc) and the TVWS channels available (.wsc). The specification for these file types is part of the digital artefacts found in Appendix A.
- Creating of tools to visualise the available TVWS channels and predicted received powers on Google Earth. Examples of this are also found as part of the digital artefacts in Appendix A.
- Quantification of the amount of TVWS channels available during the dual-illumination period for South Africa for four different propagation models.

7.1.3 Research Findings

The research done to answer the first research question regarding the amount of TVWS channels available per area for the different propagation models found that there is a large amount of spectrum available for DSA in South Africa in the terrestrial TV bands. All of the propagation models protected the PU in the validation study.

The Free-space loss model showed that there was only a small amount TVWS in the Western Cape and no TVWS for the rest of the provinces, but it should be noted that the Free-space loss model greatly over-predicts the received power and thus should not be used for DSA planning. It can be used as a first iteration, but other models give results that are closer to the ground truth. The Hata-Davidson model was implemented with the effective height of the PU incorporating the HAAT around the PU; it showed a more conservative approach by also over-predicting the received powers and thus could lead to missing out on some TVWS opportunities.

The CCDF curves for the ITM Area model and the ITM P2P model show similar amounts of TVWS available, especially for provinces with flatter terrain like the Northern Cape. The difference between the ITM Area model and the ITM P2P model is greater for areas that are more mountainous, which can be explained by the fact that the ITM P2P model takes into account the terrain by using of a DEM.

The investigation to answer the second research question posed about the speed-up that could be obtained from using an HPCC as the computation node showed that, for all the propagation models, speed-up is gained by using the HPCC. After employing improvements to the implementation of the ITM P2P model, a speed-up of up to 45 times was experienced, as was shown in Figure 5.3. The usage of an HPCC as the computation node of the Spectrum Observatory reduces the time for predicting TVWS for a large area such as South Africa from days or weeks to mere hours.

The use of more accurate, terrain-aware propagation models is encouraged, because calculations would not have to be done on a regular basis, and using the HPCC as the computation node would led to the calculations being done in a few hours.

7.2 Recommendations for Future Work

The implementations used in this research are not perfect, and there is still much that could be done to improve them. The future work identified is focused on improving the implementation of the propagation models by adding more functionality to the Spectrum Observatory and conducting more field measurements.

One of the first improvements that can be done on the ITM P2P implementation is to use a more accurate DEM, like the SRTM 1 arc-second DEM. This would improve the prediction accuracy at the cost of taking more computation time to calculate. The effect of the DEM on the TVWS available could also be a valuable research topic and is something worth investigation.

The other improvement that can be done on the ITM P2P model is to write the path extraction routines to extract the path profile for a number of points, depending on the distance of the link. This will lead to more accurate results and improve the computation time needed to do the propagation prediction.

Another worthwhile area of future study would be to investigate the effect of the interpretation of PU antenna height of the Hata-Davidson on accuracy and DSA availability, because there is little consensus on this. Some researchers use the base antenna height (AGL), whereas others use HAAT. Future work can also include investigating the impact of the PU with antenna heights that fall outside the boundaries of the antenna height parameter of the Hata-Davidson and that were forced to 20 m in the proposed implementation, seen in Table 3.2.

Part of the future work could be to add more propagation models, like the ITU-R P.1546-5 and the Irregular Terrain with Obstruction Modelling, to the Spectrum Observatory, as well as to add more TVWS policies to be used for conducting TVWS research. The TVWS per population could also be added to the analysis. The blocks not coloured in Figure 1.1 show the parts of the Spectrum Observatory that still need to be completed before this Observatory can be used as a commercial geolocation database.

The Spectrum Observatory is primarily a research tool. The parts that still need to be added are the creation of a database that can be queried, the addition of sensing capabilities and the creation of the PAWS interface.

The impact of changing the different statistical input parameters of the propagation models on the path loss predicted and the accuracy of the path loss itself can be computed as topics for future investigation. Future work on which propagation parameter values would work best for different location in South Africa would also be a worthwhile area of study. In the validation of this research, field measurements were taken from a single receiver site. Future work can be done to take field measurements at different measurement sites around an identified PU. Spectrum sensing can also be used to aid propagation prediction, thus resulting in more accurate propagation prediction.

7.3 Closure

The results generated from this study aimed to answer the research question on the amount of TVWS available for the different propagation models and also confirmed that there is considerable DSA opportunity in the UHF bands. The validation of this study also showed that all of the propagation models protected the PU, which is the primary prerequisite for allowing DSA. The results showed that propagation models that take into account the terrain, like the ITM P2P model, can result in more TVWS being identified, while still providing service protection to the PU.

The development of a Spectrum Observatory with an HPCC as the computation node will allow for faster propagation prediction, which could aid in doing TVWS research and even serve as a basis for using a geolocation-based approach for DSA in other parts of spectrum.

With the digital migration taking place in the TV broadcasting domain, even more TVWS will become available, once the analogue switch-off is complete. The continuous research taking place in propagation modelling and computing would support the notion to allow DSA in other parts of spectrum and lead to the spectrum being used more efficiently.

Bibliography

[1] M. Nekovee, T. Irnich, and J. Karlsson, "Worldwide trends in regulation of secondary access to white spaces using cognitive radio," *Wireless Communications, IEEE*, vol. 19, no. 4, pp. 32–40, 2012.

[2] H. Packard, "Cable Television System Measurements Handbook," *Santa Rosa, California: Hewlett-Packard Company*, 1994.

[3] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[4] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, "A survey of various propagation models for mobile communication," *IEEE Antennas and propagation Magazine*, vol. 45, no. 3, pp. 51–82, 2003.

[5] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980.

[6] S. Kasampalis, P. I. Lazaridis, Z. D. Zaharis, A. Bizopoulos, S. Zettas, and J. Cosmas, "Comparison of Longley-Rice, ITU-R P. 1546 and Hata-Davidson propagation models for DVB-T coverage prediction." in *BMSB*, 2014, pp. 1–4.

[7] A. G. Longley and P. L. Rice, "Prediction of tropospheric radio transmission loss over irregular terrain. a computer method-1968," DTIC Document, Tech. Rep., 1968.

[8] G. A. Hufford, A. G. Longley, W. A. Kissick *et al.*, *A guide to the use of the ITS*

irregular terrain model in the area prediction mode. US Department of Commerce, National Telecommunications and Information Administration, 1982.

[9] "Final Acts of the Regional Radiocommunication Conference for planning of the digital terrestrial broadcasting service in parts of Regions 1 and 3, in the frequency bands 174-230 MHz and 470-862 MHz (RRC-06)," 2006.

[10] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79–89, 2007.

[11] "Complementary Report to ECC Report 159: Further definition of technical and operational requirements for the operation of white space devices in the band 470-790 MHz," ECC, Report 185, Jan 2013.

[12] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004.*, vol. 1, Nov 2004, pp. 772–776 Vol.1.

[13] (2016, Jun) Electronic code of federal regulations, title 47, part 15, subpart h. U.S. Government Printing Office. [Online]. Available: http://www.ecfr.gov/

[14] (2013, Sep) TV white spaces: approach to coexistence. Ofcom. [Online]. Available: https://www.whitespacealliance.org/documents/04_Ofcom_white-spaces_consultation.pdf

[15] "Draft regulations on the use of television white spaces," in *Government Gazette*, Republic of South Africa, Apr. 2017, vol. 40772, notice 283.

[16] "National radio frequency plan 2013," in *Government Gazette*, Republic of South Africa, 28 Jun. 2013, vol. 36336, notice 354.

[17] "IEEE standard definitions and concepts for dynamic spectrum access: terminology relating to emerging wireless networks, system functionality, and spectrum management," DYSPAN P19001 Working Group and others, Tech. Rep., 2008.

[18] F. Hessar and S. Roy, "Resource allocation techniques for cellular networks in TV white space spectrum," in *Dynamic Spectrum Access Networks (DYSPAN), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 72–81.

[19] D. Makris, G. Gardikis, and A. Kourtis, "Quantifying TV White Space Capacity: A Geolocation-Based Approach," *IEEE Communications Magazine*, vol. 50, no. 9, p. 145, 2012.

[20] M. Cave, C. Doyle, and W. Webb, *Essentials of modern spectrum management*. Cambridge University Press Cambridge, 2007.

[21] J. Mitola III and G. Q. Maguire Jr, "Cognitive radio: making software radios more personal," *Personal Communications, IEEE*, vol. 6, no. 4, pp. 13–18, 1999.

[22] "Report and order on unlicenced operations in the television bands," FCC, Tech. Rep., Aug. 2015.

[23] (2015, Feb) Implementing TV White Spaces. Ofcom. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0034/68668/tvws-statement.pdf

[24] (2016, Jun) Microsoft Specrtum Observatory. Microsoft. [Online]. Available: https://observatory.microsoftspectrum.com/

[25] (2016, Jun) University of Washington Spectrum Observatory. [Online]. Available: http://specobs.ee.washington.edu/

[26] L. Mfupe, M. Mzyece, and F. Mekuria, "Geo-location white space spectrum database: Models and design of south africas first dynamic spectrum access coexistence manager," vol. 8, pp. 3810 – 3836, 11 2014.

[27] A. K. Mishra and D. L. Johnson, *White Space Communication: advances, developments and engineering challenges*. Springer, 2014.

[28] J. Lauterjung *et al.*, "COGEU D4. 1 Spectrum measurements and anti-interference spectrum database specification," vol. 1, 2010.

[29] "Technical and operational requirements for the operation of white space devices under geo-location approach," ECC, Report 186, Jan 2013.

[30] (2011, Sep) Implementing Geolocation. Ofcom. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0035/46889/statement.pdf

[31] L. Frenzel, *Principles of electronic communication systems*. McGraw-Hill, Inc., 2007.

[32] C. Haslett, *Essentials of radio wave propagation*. Cambridge University Press, 2008.

[33] Std.

[34] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Elec. Commun. Lab*, vol. 16, no. 9, pp. 825–73, 1968.

[35] P. Pathania, P. Kumar, and B. S. Rana, "Performance evaluation of different path loss models for broadcasting applications," *American Journal of Engineering Research (AJER)*, vol. 3, no. 4, pp. 335–342, 2014.

[36] M. Hata and B. Davidson, "A report on technology independent methodology for the modeling, simulation and empirical verification of wireless communications system performance in noise and interference limited systems operating on frequencies between 30 and 1500 MHz," in *Working group*, 1997.

[37] W. C. Lee, "Studies of base-station antenna height effects on mobile radio," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 2, pp. 252–260, 1980.

[38] G. A. Hufford, "The ITS irregular terrain model, version 1.2. 2 the algorithm," *Institute for Telecommunication Sciences, National Telecommunications and Information Administration, US Department of Commerce. http://flattop. its. bldrdoc. gov/itm. html*, 2002.

[39] M. Ferreira, "Spectral opportunity analysis of the terrestrial television frequency bands in South Africa," Ph.D. dissertation, North-West University, 2013.

[40] P. S. Tonkin, "A Tutorial on the Hata and ITM Propagation Models: Confidence, Reliability, and Clutter with Application to Interference Analysis," Tech. Rep., January 2017.

[41] P. Avez, P. Van Wesemael, A. Bourdoux, A. Chiumento, S. Pollin, and V. Moeyaert, "Tuning the Longley-Rice propagation model for improved TV white space detection," in *Communications and Vehicular Technology in the Benelux (SCVT), 2012 IEEE 19th Symposium on*. IEEE, 2012, pp. 1–6.

[42] P. S. Pacheco, *Parallel programming with MPI*. Morgan Kaufmann, 1997.

[43] P. Pacheco, *An introduction to parallel programming*. Elsevier, 2011.

[44] A. Geraci, F. Katki, L. McMonegal, B. Meyer, J. Lane, P. Wilson, J. Radatz, M. Yee, H. Porteous, and F. Springsteel, *IEEE standard computer dictionary: Compilation of IEEE standard computer glossaries*. IEEE Press, 1991.

[45] R. Buyya *et al.*, "High performance cluster computing: Architectures and systems (volume 1)," *Prentice Hall, Upper SaddleRiver, NJ, USA*, vol. 1, p. 999, 1999.

[46] C. S. Yeo, R. Buyya, H. Pourreza, R. Eskicioglu, P. Graham, and F. Sommers, "Cluster computing: high-performance, high-availability, and high-throughput processing on a network of computers," in *Handbook of nature-inspired and innovative computing*. Springer, 2006, pp. 521–551.

[47] T. Sterling, D. Becker, M. Warren, T. Cwik, J. Salmon, and B. Nitzberg, "An assessment of Beowulf-class computing for NASA requirements: Initial findings from the first NASA workshop on Beowulf-class clustered computing," in *Aerospace Conference, 1998 IEEE*, vol. 4. IEEE, 1998, pp. 367–381.

[48] L. P. Michelson and J. E. Kerrigan, "A High Performance Computing Primer," Academic Services and Technologies, Tech. Rep., July 2006.

[49] (2016, Dec) South African CHPC Unveils Lengau Supercomputer. HPCwire. [Online]. Available: https://www.hpcwire.com/off-the-wire/chpc-unveils-lengau-supercomputer/

[50] A. Keren and A. Barak, "Opportunity cost algorithms for reduction of I/O and interprocess communication overhead in a computing cluster," *IEEE Transactions on Parallel and Distributed Systems*, vol. 14, no. 1, pp. 39–50, 2003.

[51] N. Daylami, "The origin and construct of cloud computing." *International Journal of the Academic Business World*, vol. 9, no. 2, 2015.

[52] P. Mell and T. Grance, "The NIST definition of cloud computing," 2011.

[53] M. Yousif, "Introducing IEEE Cloud Computing: A Very Timely Magazine," *IEEE Cloud Computing*, no. 1, pp. 4–7, 2014.

[54] S. Ortiz Jr, "The problem with cloud-computing standardization," *Computer*, no. 7, pp. 13–16, 2011.

[55] G. Petri, "Shedding light on cloud computing," *The Cloud Academy*, 2010.

[56] A. Iosup, S. Ostermann, M. N. Yigitbasi, R. Prodan, T. Fahringer, and D. Epema, "Performance analysis of cloud computing services for many-tasks scientific computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 6, pp. 931–945, 2011.

[57] H. A. Duran-Limon, L. A. Silva-Bañuelos, V. H. Tellez-Valdez, N. Parlavantzas, and M. Zhao, "Using lightweight virtual machines to run high performance computing applications: the case of the weather research and forecasting model," in *Utility and Cloud Computing (UCC), 2011 Fourth IEEE International Conference on*. IEEE, 2011, pp. 146–153.

[58] F. A. Cruz, "Tutorial on GPU computing," 2009.

[59] V. Halyo, P. LeGresley, P. Lujan, V. Karpusenko, and A. Vladimirov, "First evaluation of the CPU, GPGPU and MIC architectures for real time particle tracking based on Hough transform at the LHC," *Journal of Instrumentation*, vol. 9, no. 04, p. P04005, 2014.

[60] S. Wei, L. Cai, D. Liu, Z. Liu, L. Chen, and P. Liu, "A practical test method for scalability of the cluster," in *Information Science and Service Science and Data Mining*

(ISSDM), 2012 6th International Conference on New Trends in. IEEE, 2012, pp. 414–418.

[61] X. Z. X. Zhang, Y. Y. Y. Yan, and Q. M. Q. Ma, "Measuring and analyzing parallel computing scalability," in *Parallel Processing, 1994. ICPP 1994 Volume 2. International Conference on*, vol. 2. IEEE, 1994, pp. 295–303.

[62] A. A. Lysko, M. T. Masonta, M. R. Mofolo, L. Mfupe, L. Montsi, D. L. Johnson, F. Mekuria, D. W. Ngwenya, N. S. Ntlatlapa, A. Hart *et al.*, "First large TV white spaces trial in South Africa: A brief overview," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2014 6th International Congress on*. IEEE, 2014, pp. 407–414.

[63] M. Masonta, L. Kola, A. Lysko, L. Pieterse, and M. Velempini, "Network performance analysis of the Limpopo TV white space (TVWS) trial network," in *AFRICON, 2015*. IEEE, 2015, pp. 1–5.

[64] A. Fanan, N. Riley, M. Mehdawi, and M. Ammar, "Comparison of propagation models with real measurement around Hull, UK," in *Telecommunications Forum (TELFOR), 2016 24th*. IEEE, 2016, pp. 1–4.

[65] D. Catrein, M. Reyer, and T. Rick, "Accelerating radio wave propagation predictions by implementation on graphics hardware," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*. IEEE, 2007, pp. 510–514.

[66] A. Valcarce, G. De La Roche, and J. Zhang, "A GPU approach to FDTD for radio coverage prediction," in *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on*. IEEE, 2008, pp. 1585–1590.

[67] (2017, Mar) CloudRF: Online radio planning. CloudRF. [Online]. Available: https://cloudrf.com/

[68] (2017, Mar) LSTelecom: Smart Spectrum Solutions. LSTelecom. [Online]. Available: http://www.lsofsa.co.za/

[69] S. Hossain and S. Faruque. (2017, Nov) Cellular RF Coverage Design and Optimization Using Cloudrf. [Online]. Available: http://www.academia.edu/download/32340801/Cellular_RF_Coverage_Design_and_Optimization_Using_Cloudrf_-Sehtab_hossain_saleh_faruque_ID524.docx

[70] O. Pell, J. Bower, R. Dimond, O. Mencer, and M. J. Flynn, "Finite-difference wave propagation modeling on special-purpose dataflow machines," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 5, pp. 906–915, 2013.

[71] T. Brummett, P. Sheinidashtegol, D. Sarkar, and M. Galloway, "Performance Metrics of Local Cloud Computing Architectures," in *Cyber Security and Cloud Computing (CSCloud), 2015 IEEE 2nd International Conference on*. IEEE, 2015, pp. 25–30.

[72] C.-T. Lu, C.-S. E. Yeh, Y.-C. Wang, and C.-S. Yang, "The Performance Study of a Virtualized Multicore Web System." *KSII Transactions on Internet & Information Systems*, vol. 10, no. 11, 2016.

[73] M. Liniger, M. Marghitola, M. Rohner, M. A. N. da Silva, and E. Costa, "Wave propagation models comparison of prediction results with measurements," in *Communication Technology, 2006. ICCT'06. International Conference on*. IEEE, 2006, pp. 1–5.

[74] M. Mishra and A. Sahai, "How much white space is there?" *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-3*, 2009.

[75] J. van de Beek, J. Riihijärvi, A. Achtzehn, and P. Mähönen, "UHF white space in Europe A quantitative study into the potential of the 470–790 MHz band," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 1–9.

[76] "Publication of final terrestrial broadcasting frequency plan, 2008," in *Government Gazette*, Republic of South Africa, 18 Nov. 2009, vol. 32728, notice 1538.

[77] *Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3 000 MHz*, ITU-R Recommendation P.1546-5.

[78] "Census 2001: Investigation into appropriate definitions of urban and rural areas for South Africa," Tech. Rep., Feb 2009.

[79] "GLOBE Task Team and others (Hastings, David A., Paula K. Dunbar, Gerald M. Elphingstone, Mark Bootz, Hiroshi Murakami, Hiroshi Maruyama, Hiroshi Masaharu, Peter Holland, John Payne, Nevin A. Bryant, Thomas L. Logan, J.-P. Muller, Gunter Schreier, and John S. MacDonald), eds., 1999. The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80303, U.S.A. Digital data base on the World Wide Web (URL: http://www.ngdc.noaa.gov/mgg/topo/globe.html) and CDROMs."

[80] *Planning criteria, including protection ratios, for digital terrestrial television services in the VHF/UHF bands*, ITU-R Recommendation BT.1368-12.

[81] *Minimum field strengths for which protection may be sought in planning an analogue terrestrial television service*, ITU-R Recommendation BT.417-5.

[82] "Astronomy geographic advantage act, 2007," in *Government Gazette*, Rebuplic of South Africa, 17 Jun. 2008, no. 31157, notice 27.

[83] R. G. Sargent, "Verification and validation of simulation models," in *Proceedings of the 40th Conference on Winter Simulation*, ser. WSC '08. Winter Simulation Conference, 2008, pp. 157–169.

[84] J. Devore, N. Farnum, and J. Doi, *Applied statistics for engineers and scientists*. Nelson Education, 2013.

[85] S. Lyubchenko, J.-P. Kermoal, S. Hiensch, and K. Koch, "Seamcat modeling system-level emc analysis," in *Electromagnetic Compatibility (EMC Europe), 2014 International Symposium on*. IEEE, 2014, pp. 1299–1304.

[86] Microcomputer Spectrum Analysis Models. [Online]. Available: http://ntiacsd.ntia.doc.gov/msam/

[87] H. Mahmoud and N. Akkari, "Shortest Path Calculation: A Comparative Study for Location-Based Recommender System," in *Computer Applications & Research (WSCAR), 2016 World Symposium on*. IEEE, 2016, pp. 1–5.

[88] ITM - ITS. [Online]. Available: https://www.its.bldrdoc.gov/resources/radio-propagation-software/itm/itm.aspx

[89] SPLAT! A Terrestrial RF Path Analysis Application For Linux/Unix. [Online]. Available: http://www.qsl.net/kd2bd/splat.html

[90] (2013, March) Calculation of antenna factor and gain. Micronix. [Online]. Available: http://www.micronix-jp.com/english/file-download/application/pdf/antenna_factor.pdf

[91] *Analogue television systems currently in use throughout the world*, ITU-R Recommendation BT.2043.

[92] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the error of path loss models," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 71–82.

[93] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, 2014.

[94] F. Hessar and S. Roy, "Capacity considerations for secondary networks in TV White Space," *Mobile Computing, IEEE Transactions on*, vol. 14, no. 9, pp. 1780–1793, 2015.

[95] D. Gurney, G. Buchwald, L. Ecklund, S. L. Kuffner, and J. Grosspietsch, "Geo-location database techniques for incumbent protection in the TV white space," in *New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008. 3rd IEEE Symposium on*. IEEE, 2008, pp. 1–9.

[96] D. Tomic, E. Imamagic, and L. Gjenero, "Towards new energy efficiency limits of

High Performance Clusters," in *Information Technology Interfaces (ITI), Proceedings of the ITI 2013 35th International Conference on*. IEEE, 2013, pp. 89–93.

[97] P. Kaur and R. Singh, "Complementary cumulative distribution function for performance analysis of OFDM signals," *IOSR Journal of Electronics and Communication Engineering (IOSRJECE)*, vol. 2, no. 5, pp. 05–07, 2012.

Appendix A

Digital Artefacts

The files that would allow the viewing of the received powers for transmitters, per channel for the country as well as the files that show the amount of TVWS channels on Google Earth can be seen by scanning the QR code below or going to the URL denoted below the QR code. Publications and presentations resulting from this work are also in this directory.

Figure A.1: QR Code to Digital Artefacts

http://uqr.to/9gtq

Appendix B

Research Publications

B.1 SATNAC 2016 Work-in-progress Paper

A work-in-progress paper peer reviewed and accepted as part of the Southern Africa Telecommunication Networks and Applications Conference (SATNAC) for 2016 titled "*Towards a Spectrum Observatory for TV White Space*".

B.2 SATNAC 2017 Article

Peer reviewed and presented article discussing the "*Performance Analysis of an HPC Implementation for Large-Scale Propagation Prediction*", appearing in the proceedings of SATNAC for 2017.

Towards a Spectrum Observatory for TV White Space

Johan Havenga, Melvin Ferreira
TeleNet Research Group
School of Electrical, Electronic and Computer Engineering
North-West University, Potchefstroom Campus, South Africa
Email: 23442735@g.nwu.ac.za, melvin.ferreira@nwu.ac.za

***Abstract*—The concept of Dynamic Spectrum Access (DSA) provides a solution to the artificial spectrum scarcity that has been created with current spectrum licensing regimes. The use of DSA in the TV bands, also called TV White Space (TVWS), presents interesting possibilities for mobile broadband provisioning. However, many TVWS policy and technical issues still need to be investigated. To this end, we introduce and motivate the need for our spectrum observatory architecture. The spectrum observatory will allow for rapid comparison and research of TVWS technical and policy issues.**

***Index Terms*—TV White Space, Dynamic Spectrum Access, Cloud Computing, Geolocation Database.**

I. Introduction

Wireless communication is dependent upon the amount of spectrum still available and spectrum is a limited resource. Worldwide the spectrum is licensed for exclusive use to different licensees and a large part of the spectrum is assigned to terrestrial TV service (470 - 790 MHz in RSA) [1]. This contributes to the problem of the spectrum not being utilised efficiently [2].

A solution towards more effective spectrum usage is to allow DSA. DSA allows secondary devices to utilise spectrum originally assigned to Primary Users (PUs) or incumbents without causing harmful interference to the PUs [3]. The term TVWS refers to secondary access to the part of the spectrum originally assigned to TV broadcasters.

The Secondary Users (SUs) can use one of two methods to determine if there is spectrum available for re-use. The first method is spectrum sensing in which the SU has the ability to sense the spectrum and scan for PU activity [4]; the second method is to use a geolocation database. The preferred method for TVWS is a geolocation database based approach.

Countries have different policies set up to regulate DSA. The policies determine when a part of the spectrum can be considered white space, how much power can be used by TV-band devices (TVBD), etc. The most important responsibility of the policies is to protect the incumbents from harmful interference [5].

The geolocation database uses information about the primary transmitters along with a propagation model and regulatory policies to model the protection regions of the PUs. A key issue with the geolocation based approach is the computation time needed to do fine-grained computations to determine the TVWS channels. In [6] it took the authors a month to calculate the TVWS channels available in the UK.

The Federal Communications Commission (FCC) requires the White Space Databases (WSDBs) to renew its information on the incumbents and recalculate the TVWS channels weekly [7]. Office of Communication (Ofcom) distributes updated field strengths of the Digital Terrestrial TV (DTT) incumbents every 6 months [8]. These update times are in part because of the computation time needed to calculate the field strengths of the PUs.

The first usage of the term "Spectrum Observatory" in a radio spectrum context is found in an article written about the first spectrum sensing network [9]. The spectrum observatory referred to in this paper, uses inputs from the geolocation database as well as data from sensing for research. The computation node in the spectrum observatory must also be powerful enough to (re)calculate the field strengths and TVWS available in a reasonable amount of time.

This paper proposes an architecture for a spectrum observatory that can aid researchers in answering questions about TVWS. These aspects include but are not limited to, the research on the effect of the policy on the amount and capacity of TVWS available, the impact of the accuracy of the propagation models on TVWS available and the operational parameters of TVBDs on the achievable capacity in TVWS.

This rest of this paper is organised as follows. Section II discusses related work, in Section III the proposed architecture for the spectrum observatory is discussed and the paper is concluded in Section IV.

II. Related Work

A. Commercial Uses

There are a few companies that provide a geolocation database for handling client requests from TVBDs. The FCC approved a list of WSDBs and this include, but are not limited to, Google, LStelcom and Spectrum Bridge [10]. Ofcom also has a list of certified WSDBs. This includes a database by Nominet UK, Fairspectrum and Spectrum Bridge [11].

B. Academic Work

A global spectrum observatory project has been launched by Microsoft that uses sensing nodes across the globe to

determine spectrum usage at different locations [12].

A computation and measurement based spectrum observatory was also created by the authors in [13]. Locally, a geolocation database for South Africa was created by the CSIR [14].

A recent article showed that with the current TVWS definition, which only relies on the primary transmitters, there are still some TVWS not being utilised. This article suggests that also monitoring the TV receivers for inactivity will allow the identification and exploitation of more TVWS channels. The proposed architecture uses a geolocation database in combination with sensing to determine TVWS availability [15].

III. Proposed Architecture

The spectrum observatory that will be created will have a geolocation database and a web-based Graphical User Interface (GUI) to display the data from the geolocation database. The geolocation database can also handle client requests from TVBDs. The architecture for the spectrum observatory can be seen in Fig. 1.

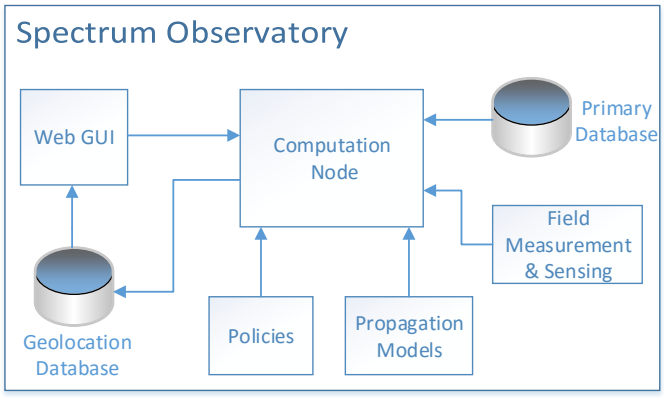


Fig. 1. Spectrum Observatory Architecture

The primary database will contain the information from the Independent Communications Authority of South Africa (ICASA) about the primary transmitters, which will be used along with a propagation model, to calculate the field strength of the PU signals. The TVWS policy, using the calculated field strengths, will determine which channels in the spectrum qualifies as TVWS. Spectrum sensing data will also be added to do TVWS analysis and allow for more accurate TVWS qualification. The GUI will allow these values to be visualised.

A central part of the spectrum observatory is the computation node. The calculation of the field strengths from the PUs and the calculation of the available TVWS channels will occur in this node. There are a few possible solutions that can be used in this regard. The first option is to use Cluster Computing, which is a collection of computers interconnected by a network [16]. The second option is to use Cloud Computing. Cloud Computing is defined as the delivery of applications, services, storage, infrastructures or platforms over the internet or a private network [17]. There are three service models used in Cloud Computing; Infrastructure as a Service (IaaS), Software as a Service (SaaS) and Platform as a Service (PaaS) [18].

The initial approach is to use an easy implementable SaaS Cloud Computing solution for the computation node. The interfaces between the different components also need to be created. A phased approach will be used in creating the spectrum observatory.

IV. Conclusion and Future Work

TVWS and spectrum usage is a pivotal research area and a spectrum observatory will propel the time it takes to answer research questions regarding TVWS, to this end our spectrum observatory was introduced.

Future work includes using the spectrum observatory to do research on the policy and regulation of TVWS in South Africa. The SaaS computation node can also be replaced with a more powerful computation solution.

References

[1] "National Radio Frequency Plan, 2013," in *Government Gazette*, Republic of South Africa, 28 Jun 2013, no. 36336, notice 354.

[2] F. S. P. T. Force, "Report of the spectrum efficiency working group," 2002.

[3] D. P. W. Group *et al.*, "IEEE standard definitions and concepts for dynamic spectrum access: terminology relating to emerging wireless networks, system functionality, and spectrum management," Technical report, IEEE, Piscataway, Tech. Rep., 2008.

[4] ECC, "Further definition to technical and operational requirements for the operation of white space devices in the band 407-790 MHz," CEPT ECC, ECC Report, January 2013.

[5] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 79–89, 2007.

[6] D. Makris, G. Gardikis, and A. Kourtis, "Quantifying TV white space capacity: A geolocation-based approach," *IEEE Communications Magazine*, vol. 50, no. 9, p. 145, 2012.

[7] (Jun 2016) Electronic Code of Federal Regulations, Title 47, Chapter I, Part 15, Subpart H. U.S. Government Publishing Office (GPO). [Online]. Available: http://www.ecfr.gov/

[8] "TV white spaces: approach to coexistence," Ofcom, Tech. Rep., September 2013.

[9] R. B. Bacchus, A. J. Fertner, C. S. Hood, and D. A. Roberson, "Long-term, wide-band spectral monitoring in support of dynamic spectrum access networks at the IIT spectrum observatory," in *New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008. 3rd IEEE Symposium on*. IEEE, 2008, pp. 1–10.

[10] *Report and Order on Unlicensed Operations in the Television Bands*, Federal Communications Commission, FCC 15-99, Aug 2015.

[11] "Implementing TV white spaces," Ofcom, Tech. Rep., February 2015.

[12] (2016) Microsoft spectrum observatory. [Online]. Available: https://observatory.microsoftspectrum.com/

[13] (2016) University of Washington spectrum observatory. [Online]. Available: http://specobs.ee.washington.edu/

[14] (2016) CSIR geolocation spectrum database. [Online]. Available: http://whitespaces.meraka.csir.co.za/index.jsp

[15] A. Saeed, M. Ibrahim, K. A. Harras, and M. Youssef, "Toward dynamic real-time geo-location databases for TV white spaces," *Network, IEEE*, vol. 29, no. 5, pp. 76–82, 2015.

[16] R. Buyya *et al.*, "High performance cluster computing: Architectures and systems (volume 1)," *Prentice Hall, Upper SaddleRiver, NJ, USA*, vol. 1, p. 999, 1999.

[17] S. Ortiz Jr, "The problem with cloud-computing standardization," *Computer*, no. 7, pp. 13–16, 2011.

[18] P. Mell and T. Grance, "The NIST definition of cloud computing," 2011.

Johan Havenga obtained his B.Eng degree in Computer and Electronic Engineering in 2015 at the North-West University and is currently pursuing a M.Eng in Computer Engineering. His interests include spectrum management, spectrum reuse, and cloud computing.

Performance Analysis of an HPC Implementation for Large-Scale Propagation Prediction

Johan Havenga, Melvin Ferreira
School of Electrical, Electronic and Computer Engineering
North-West University, Potchefstroom Campus, South Africa
[1]23442735@g.nwu.ac.za
[2]melvin.ferreira@nwu.ac.za

***Abstract*—The concept of Dynamic Spectrum Access (DSA), allows under-utilised spectrum to be made available for Secondary Users (SUs) when not utilised by the Primary User (PU). A geolocation database is currently the preferred approach by regulators to deploy DSA in the TV-bands. Central to the geolocation database is the computation node, responsible for performing propagation prediction over a large area. Field strengths are computed for all PU transmitters to determine the geographic location of channels available to SUs. Propagation prediction is computationally intensive and time-consuming, effectively becoming the limiting factor in extending the geolocation based approach for DSA to other parts of the spectrum. We present a High-Performance Cluster Computer (HPCC) implementation for the computation node. Four propagation models are implemented and adapted for parallelisation. Our performance analysis indicates computational speed-up for all propagation models on the HPCC and factors affecting linear scalability are identified.**

***Index Terms*—Dynamic Spectrum Access, TV White Space, Propagation Prediction, High Performance Computing, Parallel Computing, High Performance Cluster Computing**

I. Introduction

In a spectrum scarce world, every effort needs to be made to use spectrum as efficiently as possible. One of the methods that have been proposed is to use DSA for last-mile broadband wireless access. The requirement of DSA is that the PU will be protected against harmful interference from the SU, harmful meaning interference that affects the service grade of the PU to an unacceptable level [1]. DSA in TV-bands is referred to as TV White Space (TVWS). In order to avoid interference to the PU, two approaches are proposed. The first is Spectrum Sensing [2] and the second approach is to use a geolocation database.

The challenge with sensing is that a lot of equipment is needed to effectively cover an area. Specialised sensing equipment, with high receiver sensitivity and calibrated antennas, is required. These factors drive up cost, making sensing over a large area expensive and not viable. Research is being done to use cheaper alternatives, such as cognitive radios, to do the spectrum sensing [3].

With a geolocation database, propagation prediction is used to predict the field strengths of the PUs at a distance from each PU, to determine protected regions and to make DSA decisions based on those predictions [4]. Propagation prediction is also an essential part of broadcast and cellular network planning. In the case of TVWS, the terrestrial TV transmitters specifications are used as input for the propagation prediction and the calculated field strengths are stored in the geolocation database. The geolocation database uses this predicted field strengths as a basis to determine unused spectrum.

The core issue with propagation prediction is that it takes a long time to calculate field strengths for an area of interest. As an example, to calculate the field strengths for one transmitter for an area of Gauteng in a 30 arc-second grid, which results in 54 720 calculations, takes about 20 seconds using one Central Processing Unit (CPU). That is partially why the Federal Communications Commission (FCC) requires an update of the PUs in a TVWS geolocation database once a week [5], while Office of Communication (Ofcom) supplies its geolocation database owners with new TVWS data every 6 months [6]. This is especially true for propagation models that take terrain data into account, hence faster propagation predictions are required.

It could be argued that the user activity does not change a lot in the terrestrial TV part of the Radio Frequency (RF) spectrum and hence the propagation predictions need not be recomputed timeously. However, the exploitation of TVWS for DSA is only a stepping stone to DSA in other parts of the RF spectrum. Without more dynamic propagation prediction and White Space (WS) determination, the idea of extending a geolocation based approach for DSA to other parts of the RF spectrum is not feasible.

In this article, well-known propagation models were implemented and adapted for parallelisation on an HPCC. Various tests were run on both serial - and parallel computing architectures. This was done to analyse the increase in performance by using the parallel architecture for well-known propagation prediction models which include the Free-Space Loss model, the Hata-Davidson model and the Irregular Terrain Model (ITM) model in both Area and Point-to-Point (P2P) mode.

The rest of the article is structured as follows: background is discussed in Section II and related work in Section III. The computing architectures used is detailed in Section IV. In Section V, our methodology is discussed, followed by the results in Section VI and a conclusion in Section VII.

II. Background

A. Propagation Models

There are numerous propagation models, which can be divided into two main types, namely deterministic models, such as Finite-Difference Time-Domain (FDTD) modelling, and empirical/statistical models, such as the ITM. Some common propagation models are mentioned and classified in Table I. Propagation models used for large-scale (TVWS, cellular and broadband) planning are usually outdoor empirical propagation models. For the TVWS geolocation database, the FCC uses a set of F-curves, that bare similarity to the ITU-P.1546 model, for propagation prediction [5], while Ofcom makes use of an Extended Hata model [7], none of which fully takes the terrain path profile between transmitter and receiver into account. Using the ITM P2P model could lead to more accurate results, i.e. better protection for PUs and more WS for SUs, due to the inclusion of terrain data.

TABLE I
Common Propagation Models [8], [9]

Empirical/Statistical		Deterministic
Outdoor	**Indoor**	
Okumura	Partition Losses	FDTD
Hata	Log-Distance Path Loss	Moment-Method
COST-231	Ericsson MBM	Artificial Neural Network
Dual-Slope	Attenuation Factor	
ITM		

The propagation models used in this article will be discussed in more detail.

1) Free-space Loss: The most basic propagation prediction can be done with the free-space loss formula. The free-space loss is used as a baseline for propagation predictions.

2) Hata-Davidson Propagation Model: The Hata model is derived from the Okumura outdoor propagation model which is defined in [10]. Okumura developed a set of curves giving the median attenuation relative to free space loss. The Hata model is a formulation of these curves which define path loss drawn up by Okumura [8].

The Hata model is only defined for link distances up to 100 km, and therefore the Hata model as is, can't be used for large-scale propagation prediction. The Hata-Davidson propagation model, defined in [11], extends the Hata model to distances of 300 km, frequencies of up to 2 GHz and transmitter antenna heights of up to 2500 m [12]. This propagation model was considered because it has been used to do SU to SU interference modelling [13].

3) Irregular Terrain Model: The ITM, also known as the Longley-Rice propagation model, is defined in [14]. The model was refined and the most used version can be found in [15]. The ITM is a propagation model that takes the path geometry of the terrain into account and also takes the refractivity of the troposphere into account. The ITM has been used numerous times for large-scale propagation prediction because it supports distances of up to 2000 km.

The first of two modes available in the ITM is the area mode, where the median transmission loss at a distance from a transmitter antenna is calculated, at a particular height, for a receiving antenna of a certain height. Loss due to terrain or obstructions is done using a terrain irregularity parameter. The second method is the P2P method. The P2P method uses a path profile between the transmitting and receiving points which is extracted from a Digital Elevation Model (DEM). Using a path profile is considerably more accurate, but increases the computational burden.

B. Computing Architectures

Using multiple CPUs to work in parallel on a computation instead of one CPU leads to a performance increase and is called parallel computing [16]. The main approaches to parallel computing are discussed below.

1) GPGPU Computing: General Purpose Graphics Processing Unit (GPGPU) is when a Graphics Processing Unit (GPU) is used for general purpose computing, and researchers have successfully used GPUs as GPGPUs for numerous applications. GPUs are massively parallel because they have multiple cores on one unit. One of the drawbacks of a GPGPU is that some of the math functions, e.g. the trigonometric functions (which are used in propagation models such as the ITM model), are approximated when used on the GPU leading to a slight compromise in accuracy [17].

2) Cloud Computing: The National Institute of Standards and Technology (NIST) defines cloud computing as a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be promptly procured and made available with little management effort [18].

The advantages of cloud computing are flexibility, scalability and affordability [19]. With cloud computing, new nodes can easily be added and computing resources can be leased as and when required. Current cloud computing solutions are not suitable for running High-Performance Computing (HPC) jobs because the virtualisation overhead in cloud computing leads to a decrease in performance and because it is difficult to provide performance guarantees in a virtual environment [20].

3) Cluster Computing: A cluster computer is a collection of computers interconnected by a network [21] and can be classified as a distributed-memory Multiple Instruction, Multiple Data (MIMD) system. A cluster computer that is built for High-Performance Computing (HPC) is called an High-Performance Cluster Computer (HPCC). In South Africa, the Council for Scientific and Industrial Research (CSIR) has a division called the Centre for High Performance Computing (CHPC) which built an HPCC called "Lengau" that consists of over 24,000 CPUs and boasts a peak performance of 24.9 teraflops/second [22].

The main advantages of using cluster computing are availability, scalability and performance. If more computing resources are needed, more computing nodes can be added to the cluster. The biggest drawback for using cluster computing is the communications overhead that is needed to pass data between the different nodes [23]. Cluster computing is able to

do the trigonometric computations, with less compromise in accuracy than experienced when using GPGPU and doesn't have virtualisation overhead, as is the case with cloud computing, leading to better performance than experienced with cloud computing.

4) Computation Performance Metrics: Measuring the performance between different computer architectures needs clear definitions of the metrics used. The computation metrics used depends on the computing architectures being tested and the tests being run. Arguably the most important metric on an HPCC is the running time, also called latency. Latency is defined as the elapsed time between starting the program and ending the program on a computing architecture [24].

The second metric used is speed-up, which is defined as the ratio of the runtime of a serial solution (T_σ) to the runtime of a parallel solution (T_π) [25]. The equation for speed-up is given in (1),

$$S(n,p) = \frac{T_\sigma(n,p)}{T_\pi(n,p)} \tag{1}$$

where $T_\sigma(n)$ is the serial solution runtime and $T_\pi(n,p)$ is the parallel solution runtime using p CPUs on n computing nodes.

C. Parallel Programming

Although most computers have parallel computing capabilities, most programs are not written to fully exploit this. To write parallel programs for a parallel computer one of the following parallel programming approaches can be used.

1) OpenMP: Open Multi-Processing (OpenMP) is an Application Programming Interface (API) for shared-memory parallel programming, meaning that OpenMP is designed for systems in which each CPU or processor can have access to all available memory [16]. The OpenMP API does the distribution of tasks among the CPUs for the programmer. However, the programmer has less control over the parallelisation process.

2) MPI: The Message Passing Interface (MPI) is the most used message passing implementation. Message passing is used in a distributed-memory system where one process running on a processor can send information or data to another process running on another processor. MPI is not a new programming language but rather a library that can be used with multiple programming languages and enables message passing [25]. With MPI the programmer has control over which CPU does what part of the code and the programmer is also responsible for distributing data between the CPUs.

III. Related Work

TVWS prediction through the use of propagation prediction has been done for South Africa in [26] and trials for the secondary usage of TVWS has been done in [27] and [28]. A study was done in Hull, United Kingdom (UK), which showed that most of the propagation models used overestimate the path loss of the signal when compared to measured data. The Egli model, which takes terrain data into account, was the closest to measured data in that trial [29].

GPUs have been used for HPC to do propagation prediction in [30] and [31]. In [31] the authors used FDTD with GPUs to do radio coverage prediction, while in [30] a ray-optical approach was implemented on a GPU. Both of these studies shows the performance increase of using a GPU for predicting field strengths, but both studies only consider small urban areas for evaluation.

Cloud computing has been used in the form of a commercial Software as a Service (SaaS) solution for propagation prediction, for example, CloudRF [32]. CloudRF can do large-scale area prediction using a selection of propagation models along with 3 arc-second DEM data is a viable option. The service operates on a subscription or pay-per-use basis. CloudRF has been used for cellular coverage design in [33].

The only use of an HPCC for aiding in large-scale propagation prediction is in [34], where the authors used the HPCC as part of their Ofcom approved geolocation database [35]. The specifics of how and where the HPCC is utilised in their geolocation database, is not described. Also, the authors did not do a performance analysis of their HPCC implementation.

IV. Computing Architectures

The various options for the computational node have been discussed in Section II and because of the scalability and performance increase, the HPCC is our preferred option. This option is further encouraged by the fact that the CHPC made its Lengau HPCC available to researchers, free of charge. The Lengau HPCC makes use of the Lustre file system and the specifications for the Lengau HPCC can be seen in Table II, along with the specifications of the local PC used for comparison in this article.

TABLE II
Specifications for Computing Architectures

	Local PC	**Lengau HPCC [36]**
CPU	Intel i7-6700	Intel Xeon (R) E5-2690 V3
CPU Clock	3.6 GHz	2.6 GHz
CPU Cores	4	24192
Number of Nodes	1	1008
Min CPU Cores Avail	1	1
Max CPU Cores Avail	4	240
*Rpeak	N/A	1006 TFlops
*Rmax	N/A	782.9 TFlops
Memory	8 GB	126 TB
Interconnect	N/A	FDR Infiniband Network
Storage	100 GB	4 PB

**Rmax = Max number of floating point operations*
**Rpeak = Theoretical max value of floating point operations [37]*

Implementation

The source code for implementing the ITM propagation model and the source code for extracting the terrain data from the DEM tiles are from the Institute for Telecommunication Sciences (ITS), which is part of the National Telecommunications and Information Administration (NTIA). The DEM used is the GLOBE DEM data [38]. The source code for the ITM implementation had to be ported from C++ to C. All the of

the source code had to be parallelised with MPI. The source code for the DEM data extraction routines was only available in Fortran 77 and had to be ported to C. The other propagation models (Free-Space Loss and Hata-Davidson) was easier to implement and was written directly from empirical equations.

V. Methodology

To determine the advantage of using an HPCC, which is a parallel approach, over a serial approach, an experimental set-up is needed to compare the relative performance of different computing architectures. The experiments entail using four different propagation models (ITM in P2P mode, ITM in area mode, the Hata-Davidson model and the Free-space loss model) to calculate the field strengths from terrestrial TV transmitters for the Gauteng area and then measure the latency performance of both architectures.

The field strengths are calculated on a logical grid with grid spacing of 30 arc-seconds and boundaries defined by the Gauteng area, which is defined by the GPS coordinates, 25.1° S to 27° S and 27.1° E to 29.1° E. This results in a grid of 240 horizontal/longitudinal cells by 228 vertical/latitudinal cells, which results in 54 720 cells to be calculated. This is a lot less than the 3 182 400 (2040 x 1560) cells needed for the whole country, but still, serves as an indicator of scalability. The field strengths are calculated for the centre of each cell and stored as such.

The number of TV transmitters used for the calculations ranges from 1 transmitter to all of the 2325 transmitters in the PU database. When all the transmitters are used for the Gauteng area it leads to 1 227 224 000 (54 720 x 2 325) field strengths that need to be computed for each of the propagation models.

The Hata-Davidson model has a validity for distances up to 300 km for predictions and the ITM implementation has a validity for distances up to 1 000 km, but these ranges were ignored and the calculations were still done for parameters falling outside of these boundaries to get an upper boundary of the time it would take to do the calculations. The default parameters were used for the ITM P2P and Area mode calculations.

A. Architecture Benchmark Experiment

This experiment consists of three tests as described below.

1) Serial Performance: The first benchmark involves testing the serial performance of each computing architecture. The 4 different propagation models will be used to compute the field strengths in the Gauteng area resulting from 4 transmitters. Both the local machine and the Lengau HPCC will do the calculations with 1 CPU while still utilising MPI.

2) Baseline MPI Performance: The second benchmark exploits the simplest MPI implementation where 4 CPUs are used for each computing architecture to calculate the field strengths for the defined prediction area for 4 transmitters. The source code is written in such a way that each CPU would be responsible for a transmitter on both the local PC and the HPCC. This benchmark shows the MPI performance of each architecture.

3) 48 Core Parallelism: The next benchmark seeks to further show the advantages of parallelism by using 48 CPUs on the Lengau HPCC and the maximum of 4 CPUs on the local computer to calculate the field strengths resulting from 48 transmitters. As seen in Table II, the Lengau HPCC has 24 CPUs per node and the minimum parallel usage, to experience inter-node communication, is when only 2 of the available 10 nodes are used. This results in 48 CPUs, and since there are 48 transmitters, this implies 1 transmitter for each CPU on the Lengau HPCC. On the local computer, the 48 transmitters will be divided among the 4 CPUs of the PC, giving every CPU, 12 transmitters to calculate. The area of calculation is once again the defined prediction area.

B. HPCC Scalability Experiment

Different propagation models scale differently across the increase of the number of CPUs. For this experiment, the field strengths from all of the 2325 transmitters in the database were calculated for the Gauteng area for the different propagation models using an increasing number of CPUs on the Lengau HPCC. The benchmark test was done by running the different propagation models using 48, 96, 120, 144, 192 and 240 CPUs. This was compared to the same calculation done on the local PC using 4 CPUs.

C. Metrics Used

The metrics used for the tests are latency and speed-up. For each experiment, the latency from the local PC (T_σ) and the latency from the HPCC (T_π) will be recorded and the speed-up ($S(n,p)$) will be calculated. Every test will be repeated 3 times, to build a test case, and the results averaged to get a test result, which is in-line with common performance analysis practice [39]. The metrics were discussed in Section II-B.

VI. Results

A. Architecture Benchmark Experiment

The tests were run and the speed-up of each test was calculated from the measured latency for each test. The speed-up for the architecture benchmark experiment can be seen in Fig.1. The Serial test measures the performance of a single CPU on the HPCC versus a single CPU on the local machine. Refer to Table II for the specifications for each computing architecture, which shows that the local machine has CPUs with faster operating frequencies than the CPUs used in the HPCC. For a single CPU, the local PC ran faster and that is why, with all the propagation models, $S(1,1) < 1$.

The Baseline MPI test shows the speed-up from when the local machine used 4 CPUs against when the HPCC computer used four CPUs. The scale-up is similar for the Free-Space Loss, Hata-Davidson and ITM Area mode implementations because although the local PC has CPUs with faster operating frequencies, the HPCC was created with the purpose of exploiting parallelism. The ITM P2P mode showed a decrease in speed, which will be discussed later in the article.

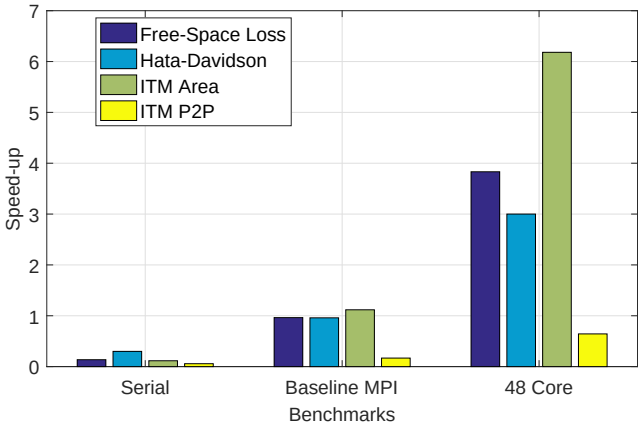
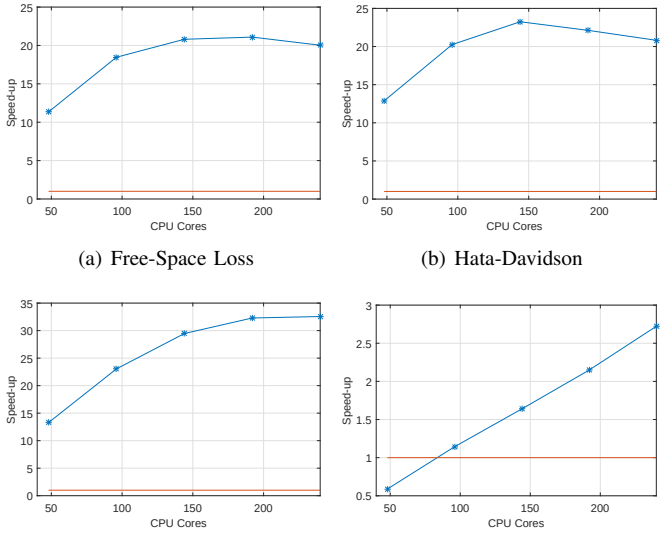


Fig. 1. Architecture Benchmark Experiment



(a) Free-Space Loss (b) Hata-Davidson

(c) ITM Area (d) ITM P2P

Fig. 2. Lengau HPCC Scalability Experiment

The 48 Core tests show the power of parallelism versus a serial architecture. The local PC used the maximum available of 4 CPUs while the HPCC used the maximum of 24 CPUs per computation node with two nodes being used, totalling to 48 CPUs on the Lengau HPCC. This was done to measure the effect of inter-node communication. For all the propagation models, except for the ITM P2P model, there was notable speed-up, as seen in Fig.1.

To understand the lack of speed-up with the ITM P2P model a profiling tool was used to analyse the ITM P2P implementation. Tuning and Analysis Utilities (TAU) is a profiling tool that can be used to measure program performance on an HPCC. The TAU tool was used to do an analysis of the ITM P2P for 48 transmitters with 48 CPUs. The TAU tool ran profiling on all the CPUs and the mean results are shown in Table III.

The results clearly show that the function using the most of the execution time is the get_GLOBE_data() function, which is called by the GLOBE_elevation() function, which is part of the get_GLOBE_pfl() routine. The GLOBE_pfl() function is responsible for creating the path with elevation data between the transmitter and the point under consideration. The get_GLOBE_data() function is a simple function that reads the elevation data for a certain coordinate out of the DEM file. This shows that most of the time is being lost to file I/O on the CHPC and this has an effect on the speed-up for the ITM P2P mode. This explains why the ITM P2P mode implementation did not show favourable speed-up for the Baseline MPI and 48 Core tests.

TABLE III
TAU RESULTS FOR ITM P2P

Time	Exclusive msec	Inclusive total msec	Name
100.0	29	35:15.784	.TAU application
100.0	1.392	35:15.754	main()
85.7	30.684	30:13.981	get_GLOBE_pfl()
84.3	11:12.851	29:43.234	GLOBE_elevation()
52.4	18:29.601	18:29:601	get_GLOBE_data()
13.9	4:54.596	4:54.596	MPI_Barrier()

B. *HPCC Scalability Experiment*

The advantage of the Lengau HPCC over the local PC is that 240 CPUs can be used on the HPCC against the maximum of 4 from the local PC. Speed-up is calculated as described in Section V-C. The effect of this can be seen for all the propagation models in the scalability experiment in Fig.2.

There is significant speed-up for all the propagation models, except for the ITM P2P implementation, as seen in Fig.2. The results from the profiling done using TAU leads us to suspect file I/O as a factor in the lack of speed-up in the ITM P2P mode implementation. The scaling of the speed-up for the different propagation models also tells an interesting story. The increase in speed-up is not linear for any of the propagation models. As more CPUs are used, additional communication overhead is experienced, which leads to an upper boundary in the speed-up increase for every problem.

VII. CONCLUSION AND FUTURE WORK

This article investigated the use of parallel computing on an HPCC to do large-scale propagation prediction. The study was done by performing tests, where four propagation models were used, of which all was written to use MPI and 2 was ported from other languages, to evaluate the performance of the HPCC.

All the propagation models, excluding the ITM P2P mode, show favourable speed-up and scalability which proves that an HPCC can be a used as the computation node for a geolocation database. The implementation of the ITM P2P model still needs to be improved in terms of parallel scalability. Future improvements include the use of a hybrid approach, using MPI and OpenMP together, and decreasing the time being spent waiting for file I/O to complete.

ACKNOWLEDGEMENT

We are thankful for the support and resources from the CHPC at the CSIR as well as for the GLOBE DEM files used for the ITM P2P implementation, that are from [38]. The source code for the original ITM implementation and the GLOBE extraction routines and continued support from the ITS, NTIA, is also gratefully acknowledged. The authors also thank the reviewers for their insightful comments and constructive feedback.

REFERENCES

[1] D. P. W. Group *et al.*, "IEEE standard definitions and concepts for dynamic spectrum access: terminology relating to emerging wireless networks, system functionality, and spectrum management," Technical report, IEEE, Piscataway, Tech. Rep., 2008.

[2] ECC, "Further definition to technical and operational requirements for the operation of white space devices in the band 407-790 MHz," CEPT ECC, ECC Report, January 2013.

[3] C.-H. Ko, D. H. Huang, and S.-H. Wu, "Cooperative spectrum sensing in TV white spaces: when cognitive radio meets cloud," in *Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on*. IEEE, 2011, pp. 672–677.

[4] M. Nekovee, T. Irnich, and J. Karlsson, "Worldwide trends in regulation of secondary access to white spaces using cognitive radio," *Wireless Communications, IEEE*, vol. 19, no. 4, pp. 32–40, 2012.

[5] (Jun 2016) Electronic Code of Federal Regulations, Title 47, Chapter I, Part 15, Subpart H. U.S. Government Publishing Office (GPO). [Online]. Available: http://www.ecfr.gov/

[6] Ofcom, "Tv white spaces: approach to coexistence," Ofcom, Tech. Rep., September 2013.

[7] ——, "Implementing TV White Spaces: Annexes 1 to 12," 2015.

[8] T. S. Rappaport *et al.*, *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, 1996, vol. 2.

[9] T. K. Sarkar, Z. Ji, K. Kim, A. Medouri, and M. Salazar-Palma, "A survey of various propagation models for mobile communication," *IEEE Antennas and propagation Magazine*, vol. 45, no. 3, pp. 51–82, 2003.

[10] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *Rev. Elec. Commun. Lab*, vol. 16, no. 9, pp. 825–73, 1968.

[11] M. Hata and B. Davidson, "A report on technology independent methodology for the modeling, simulation and empirical verification of wireless communications system performance in noise and interference limited systems operating on frequencies between 30 and 1500 MHz," in *Working group*, 1997.

[12] S. Kasampalis, P. I. Lazaridis, Z. D. Zaharis, A. Bizopoulos, S. Zettas, and J. Cosmas, "Comparison of Longley-Rice, ITU-R P. 1546 and Hata-Davidson propagation models for DVB-T coverage prediction." in *BMSB*, 2014, pp. 1–4.

[13] F. Hessar and S. Roy, "Capacity considerations for secondary networks in tv white space," *Mobile Computing, IEEE Transactions on*, vol. 14, no. 9, pp. 1780–1793, 2015.

[14] A. G. Longley and P. L. Rice, "Prediction of tropospheric radio transmission loss over irregular terrain. A computer method-1968," DTIC Document, Tech. Rep., 1968.

[15] G. A. Hufford, "The its irregular terrain model, version 1.2. 2 the algorithm," *Institute for Telecommunication Sciences, National Telecommunications and Information Administration, US Department of Commerce. http://flattop. its. bldrdoc. gov/itm. html*, 2002.

[16] P. Pacheco, *An introduction to parallel programming*. Elsevier, 2011.

[17] V. Boyer and D. El Baz, "Recent advances on GPU computing in Operations Research," in *Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW), 2013 IEEE 27th International*. IEEE, 2013, pp. 1778–1787.

[18] P. Mell and T. Grance, "The NIST definition of cloud computing," 2011.

[19] A. Iosup, S. Ostermann, M. N. Yigitbasi, R. Prodan, T. Fahringer, and D. Epema, "Performance analysis of cloud computing services for many-tasks scientific computing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 6, pp. 931–945, 2011.

[20] K. L. Kroeker, "The evolution of virtualization." *Communications of the ACM*, vol. 52, no. 3, pp. 18 – 20, 2009.

[21] R. Buyya *et al.*, "High performance cluster computing: Architectures and systems (volume 1)," *Prentice Hall, Upper SaddleRiver, NJ, USA*, vol. 1, p. 999, 1999.

[22] HPCwire. (2016) South African CHPC Unveils Lengau Supercomputer. [Online]. Available: https://www.hpcwire.com/off-the-wire/chpc-unveils-lengau-supercomputer/

[23] A. Keren and A. Barak, "Opportunity cost algorithms for reduction of i/o and interprocess communication overhead in a computing cluster," *IEEE Transactions on Parallel and Distributed Systems*, vol. 14, no. 1, pp. 39–50, 2003.

[24] X. Z. X. Zhang, Y. Y. Y. Yan, and Q. M. Q. Ma, "Measuring and analyzing parallel computing scalability," in *Parallel Processing, 1994. ICPP 1994 Volume 2. International Conference on*, vol. 2. IEEE, 1994, pp. 295–303.

[25] P. S. Pacheco, *Parallel programming with MPI*. Morgan Kaufmann, 1997.

[26] M. Ferreira, "Spectral opportunity analysis of the terrestrial television frequency bands in South Africa," Ph.D. dissertation, North-West University, 2013.

[27] A. A. Lysko, M. T. Masonta, M. R. Mofolo, L. Mfupe, L. Montsi, D. L. Johnson, F. Mekuria, D. W. Ngwenya, N. S. Ntlatlapa, A. Hart *et al.*, "First large TV white spaces trial in South Africa: A brief overview," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2014 6th International Congress on*. IEEE, 2014, pp. 407–414.

[28] M. Masonta, L. Kola, A. Lysko, L. Pieterse, and M. Velempini, "Network performance analysis of the Limpopo TV white space (TVWS) trial network," in *AFRICON, 2015*. IEEE, 2015, pp. 1–5.

[29] A. Fanan, N. Riley, M. Mehdawi, and M. Ammar, "Comparison of propagation models with real measurement around Hull, UK," in *Telecommunications Forum (TELFOR), 2016 24th*. IEEE, 2016, pp. 1–4.

[30] D. Catrein, M. Reyer, and T. Rick, "Accelerating radio wave propagation predictions by implementation on graphics hardware," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*. IEEE, 2007, pp. 510–514.

[31] A. Valcarce, G. De La Roche, and J. Zhang, "A GPU approach to FDTD for radio coverage prediction," in *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on*. IEEE, 2008, pp. 1585–1590.

[32] (2017) Cloudrf. [Online]. Available: https://cloudrf.com/

[33] S. Hossain and S. Faruque, "Cellular RF Coverage Design and Optimization Using Cloudrf: Bangladesh Perspective," *International Journal of Emerging, Technology, Science & Engineering (IJETSE) April Issue*, 2013.

[34] L. Mfupe, F. Mekuria, and M. Mzyece, "Geo-location White Space Spectrum Databases: Models and Design of South Africa's First Dynamic Spectrum Access Coexistence Manager." *TIIS*, vol. 8, no. 11, pp. 3810–3836, 2014.

[35] L. Mfupe, F. Mekuria, L. Montsi, and M. Mzyece, *Geo-location White Space Spectrum Databases: Review of Models and Design of a Dynamic Spectrum Access Coexistence Planner and Manager.* Meraka Institute-CSIR, Pretoria: Signals and Communication Technology series. New York and Heidelberg: Springer, 2015, pp. 153 – 194.

[36] (2017) Lengau cluster. [Online]. Available: http://www.chpc.ac.za/index.php/resources/lengau-cluster

[37] D. Tomic, E. Imamagic, and L. Gjenero, "Towards new energy efficiency limits of high performance clusters," in *Information Technology Interfaces (ITI), Proceedings of the ITI 2013 35th International Conference on*. IEEE, 2013, pp. 89–93.

[38] "GLOBE Task Team and others (Hastings, David A., Paula K. Dunbar, Gerald M. Elphingstone, Mark Bootz, Hiroshi Murakami, Hiroshi Maruyama, Hiroshi Masaharu, Peter Holland, John Payne, Nevin A. Bryant, Thomas L. Logan, J.-P. Muller, Gunter Schreier, and John S. MacDonald), eds., 1999. The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0. National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80303, U.S.A. Digital data base on the World Wide Web (URL: http://www.ngdc.noaa.gov/mgg/topo/globe.html) and CDROMs."

[39] C.-T. Lu, C.-S. E. Yeh, Y.-C. Wang, and C.-S. Yang, "The performance study of a virtualized multicore web system." *KSII Transactions on Internet & Information Systems*, vol. 10, no. 11, 2016.

Johan Havenga obtained his B.Eng degree in Computer and Electronic Engineering in 2015 at the North-West University and is currently pursuing an M.Eng in Computer and Electronic Engineering. His interests include spectrum management, spectrum reuse, and parallel computing.