

Augmenting large scale propagation models using field measurements

NJ Buys

 orcid.org/0000-0002-6496-4963

Dissertation accepted 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: June 2021

Student number: 24300535

Acknowledgements

Thank you to God Almighty, Jesus Christ, for giving me the strength and ability to complete this study. All the glory to You. Thank you for endless love and grace.

To my wife, Franciska. Thank you for your love, patience, support, encouragement and also sacrifices you made so that I could complete this study. Thank you for always believing in me. You are the little voice in my head shouting “You can do it!”.

To my family. Thank you for your continuous support throughout this study. You always believed in me, loved me and cared about me. Your encouragement and motivating words are appreciated.

To my in-laws. Thank you taking an interest in my work and also for your encouraging words.

To my best friend, Anton. Thank you for your encouragement and support throughout this study. Having you there, only a few offices away, certainly made this challenge more enjoyable. Our tea and fingerboard games will always be remembered.

To my supervisor, Dr. Melvin. Thank you for your input, guidance and patience. I appreciate your approach of always being positive about the study and motivating me with words of encouragement.

Thank you for the financial support received from the Telkom CoE at the NWU. It gave me the opportunity to further my studies.

Thank you for Prof. Roy and the University of Washington, Seattle for hosting me and allowing me access to your measurement equipment.

I appreciate everyone’s involvement in this study. Ultimately, this was a team effort.

“If you want to go fast, go alone. If you want to go far, go together.” African Proverb

Abstract

There is a need to use the available radio spectrum more efficiently because of the current spectrum scarcity and the ongoing increase in demand for usable bands of the radio spectrum. One way to use the radio spectrum more efficiently is to share the licensed band when and where it is not used by the licensed user. Spectrum sharing approaches like Dynamic Spectrum Access (DSA) typically makes use of a Geolocation Database (GLDB) to determine where secondary users can reuse the licensed spectrum band without interfering with the licensed primary users. Television White Space (TVWS) networks operate on this principle. Radio wave propagation models are fundamental to geolocation databases and are used to predict the coverage areas of transmitters, and so also, determine where the spectrum can be shared.

Modelling the attenuation of the signal power of a radio wave in a wireless channel is no trivial task and all propagation models have a degree of inaccuracy. Large scale radio propagation models suffer from inaccuracies because there are numerous parameters with variability to account for. This research study aims to develop and implement a method to augment large scale radio propagation models with field measurements in order to reduce the mean error between propagation predictions and field strength measurements.

Field measurements includes field strength measurements, or signal power, and land cover data. Transmitters in the radio band (88 - 108 MHz) were available and selected for this study. A measurement study spanning a period of three months, was conducted to determine the average received power from the various transmitters at a single stationary receiver. The effective heights of the land cover data were varied and added to the Digital Elevation Model (DEM), hence affecting the path obstruction-loss calculations. The heights of the land cover data modelled the effect of the land cover data on the attenuation of the signal and not the physical height of the objects associated with the land cover type. The gradient descent method was used to adjust the heights associated with the different land cover classes, by minimising the Root Mean

Square Error (RMSE) between the propagation predictions and the field strength measurements. The standard deviation and Pearson correlation coefficient were also considered when evaluating the augmented propagation model. The Longley-Rice irregular terrain model (ITM) and the irregular terrain model with obstructions (ITWOM) were chosen for this study.

The results of the study showed that it is possible to augment large scale radio propagation models with field measurements and reduce the RMSE between predictions and field strength measurements by more than 7 dBm. The results of the augmented ITM follows. The RMSE was reduced from 12.16 dBm to 4.44 dBm. That is reduction of 7.72 dBm. The standard deviation increased from 5.87 dBm to 8 dBm. That is an increase of 2.13 dBm. The Pearson correlation coefficient reduced from 0.88 to 0.78 which is a reduction of 0.1. The results of the augmented ITWOM follows. The RMSE was reduced from 10.95 dBm to 3.12 dBm, which is a reduction of 7.83 dBm. The standard deviation increased from 5.87 dBm to 6.1 dBm. That is only an increase of 0.23 dB. The Pearson correlation coefficient increased from 0.88 to 0.9, which is an increase of 0.02.

The augmented large scale propagation models require initial training but provide more accurate propagation predictions. This approach can be used to improve on the propagation models used in network planning, spectrum monitoring and spectrum management. A better informed decision regarding the use of spectrum can be made, to ultimately use the available spectrum more efficiently.

Keywords: *radio wave propagation, radio propagation model tuning, land cover, field strength measurements*

Contents

List of Figures	xi
List of Tables	xv
List of Acronyms	xvii
1 Introduction	1
1.1 Introduction	1
1.1.1 Radio spectrum	3
1.1.2 Spectrum management	3
1.1.3 Spectrum sharing	7
1.1.4 Propagation modelling	9
1.2 Proposed research	11
1.3 Importance of study	11
1.4 Research question/statement	12
1.5 Research aims and objectives	13

1.6	Research methodology	14
1.6.1	Literature study	14
1.6.2	Research model design, exploratory data analysis and implemen- tation	14
1.6.3	Results of augmenting large scale propagation models	15
1.6.4	Verification and validation	15
1.7	Dissertation overview	16
2	Literature study	17
2.1	RF propagation	17
2.1.1	Inverse square law	18
2.1.2	Free space loss	18
2.1.3	Gain	20
2.1.4	Effective radiated power	20
2.2	Multipath	21
2.2.1	Scattering	22
2.2.2	Diffraction	22
2.2.3	Reflection	26
2.2.4	Refraction	27
2.2.5	Fresnel zone	27
2.3	Radio wave propagation models	28

2.3.1	Empirical models	29
2.3.2	Deterministic models	30
2.3.3	Prediction errors	38
2.3.4	Propagation prediction software	38
2.3.5	Digital elevation model	41
2.4	Field measurements	42
2.4.1	Land cover	42
2.4.2	Field strength measurements	49
2.4.3	Power Spectral Density (PSD)	50
2.4.4	Considerations for measurement campaign	51
2.5	Model tuning and evaluation	52
2.5.1	Gradient descent method	52
2.5.2	Statistical evaluation metrics	54
2.6	Related work on augmenting large scale propagation models	57
2.6.1	Related work discussion	60
2.7	Conclusion	61
3	Research model and implementation	62
3.1	Introduction	62
3.2	Measurement study	64
3.2.1	Receiver selection and setup	65

3.2.2	Transmitter selection	71
3.2.3	Measurement approach	75
3.2.4	Processing of measurements/data processing workflow	75
3.2.5	Removing outliers	80
3.3	Processing land cover data	80
3.3.1	Projection of Coordinate Reference System (CRS)	82
3.3.2	Grid system	83
3.4	Reducing the error	87
3.4.1	Modelling rationale	87
3.4.2	Cross-validation	89
3.4.3	Gradient descent method	90
3.4.4	Logical flow of the method	91
3.4.5	Initial parameters and assumptions	97
3.4.6	Implementation	100
3.5	Conclusion	103
4	Exploratory data analysis and results	105
4.1	Measurement study	105
4.2	Land cover	117
4.3	Reducing the error	120
4.3.1	ITM	121

4.3.2	ITWOM	125
4.4	Conclusion	129
5	Verification and validation	131
5.1	Verification	131
5.1.1	Introduction	131
5.1.2	Measurements	132
5.1.3	Land cover	133
5.1.4	Gradient descent method	137
5.2	Validation	139
5.2.1	Introduction	139
5.2.2	Evaluation set	140
5.2.3	Conclusion	145
6	Conclusion	147
6.1	Dissertation overview	147
6.2	Findings	149
6.3	Recommendations for future work	152
6.4	Closure	154
	Bibliography	156
	Appendices	

List of Figures

1.1	ITU regions adapted from [1]	5
1.2	South Africa 3kHz - 300GHz regulated frequency allocations chart [2] . .	6
2.1	Multipath effects due to diffraction, reflection and scattering	22
2.2	Double knife-edge diffraction using the Bullington method adapted from [3,4]	24
2.3	Double knife-edge diffraction using the Epstein-Peterson method adapted from [5,6]	25
2.4	Double knife-edge diffraction using the Deygout method adapted from [3,5]	26
2.5	Fresnel Zone	27
2.6	Calculation of terrain irregularity parameter, (Δh)	34
2.7	Median reference attenuation vs distance plot [7,8]	36
2.8	Geographic coordinates	48
2.9	Effect of initial parameters in finding local minima	54
3.1	High-level overview of the methodology	64

3.2	Measurement study setup	65
3.3	CityScape sensor layout adapted from [9]	67
3.4	Data acquisition cycle adapted from [9]	70
3.5	Effect of cable loss on calculated average power	71
3.6	Flowchart of data processing workflow	75
3.7	File structure after data is extracted	76
3.8	Example of plots illustrating the importance of units when computing the PSD.	78
3.9	Structure of a 3-dimensional matrix containing average received power measurements	79
3.10	Bounds highlighted for land cover and SRTM data	82
3.11	MATLAB grid system	84
3.12	SDF grid system	84
3.13	SRTM grid system	85
3.14	Direction with degrees	86
3.15	4-fold CV example	90
3.16	Flow diagram of the gradient descent method for reducing error	92
3.17	Proportion of land cover over all paths	98
3.18	Interaction of software in the implementation of the method	100
4.1	Average received power for transmitter KLSY for 1 day	106

4.2	Average received power for transmitter KBCS for 1 day	107
4.3	Geographic locations of transmitters and receiver	108
4.4	Average received power: 20km cluster for 8 days	109
4.5	Average received power: 30km cluster for 8 days	109
4.6	Average received power: 100km cluster for 8 days	110
4.7	Received power (on day 2) from two of the transmitters with the greatest variance in signal strength.	111
4.8	Average received power: 4km cluster for 93 days	112
4.9	Average received power: 20km cluster for 93 days	112
4.10	Average received power: 30km cluster-1 for 93 days	113
4.11	Average received power: 30km cluster-2 for 93 days	113
4.12	Average received power: 40km cluster for 93 days	114
4.13	Average received power: 100km cluster for 93 days	114
4.14	Box plot of averaged received power for all transmitters	115
4.15	Example of a land cover tile over a region of interest	118
4.16	Proportion of land cover classes in all paths for UW station	118
4.17	Proportion of land cover subclasses for unique paths to the UW station .	119
4.18	Proportion of land cover classes in all paths for UW station	120
4.19	Comparison of RMSE between measured and predicted signal power when varying two land cover classes' heights.	127

5.1	PSD plot from CityScape website	132
5.2	PSD plot generated in MATLAB	133
5.3	10 000 samples (blue dots) used for verification of land cover extraction process overlaid on the land cover dataset in QGIS	135
5.4	SDF file modification process	136
5.5	Google Maps screen capture	137
5.6	SDF grid output	137
5.7	Proportion of Land Cover	141

List of Tables

2.1	ITM valid input parameter ranges [10]	31
2.2	ITM point-to-point general input parameters [8]	32
2.3	Radio climates and suggested values for N_s [8,11]	35
2.4	Typical values for ground constants [8,11]	35
2.5	NLCD classes [12]	44
2.6	Correlation coefficient interpretation [13]	57
3.1	UW receiver station parameters	66
3.2	Microsoft receiver station parameters	66
3.3	Tunable parameters of the UW spectrum observatory	69
3.4	Transmitter information of interest obtainable from the FCC database . .	72
3.5	FM transmitter set used for study	74
4.1	Measurement results from UW station	116
4.2	Land cover tile bounds	117
4.3	Initial parameters before the method is implemented for ITM	121

4.4	Final parameters after the method is implemented for ITM	124
4.5	Evaluation set results for ITM	125
4.6	Final heights per land cover class for ITM	125
4.7	Initial parameters before the method is implemented for ITWOM	126
4.8	Final parameters after the method is implemented for ITWOM	128
4.9	Evaluation set results for ITWOM	128
4.10	Final heights per land cover class for ITWOM	128
5.1	Step-wise gradient descent example	138
5.2	Evaluation set results for ITM	142
5.3	Evaluation set results after the PSD is modified for ITM	143
5.4	Evaluation set results for ITWOM	144
5.5	Evaluation set results after the PSD is modified for ITWOM	145
A.1	Results for first random initial height set using ITM	165
A.2	Results for second random initial height set using ITM	165
A.3	Results for third random initial height set using ITM	166

List of Acronyms

AGL Above Ground Level

AMSL Above Mean Sea Level

CBRS Citizen Broadband Radio Service

CR Cognitive Radio

CRS Coordinate Reference System

DEM Digital Elevation Model

DMS Degrees Minutes Seconds

DSA Dynamic Spectrum Access

DTV Digital Television

EIRP Effective Isotropic Radiated Power

ERP Effective Radiated Power

FCC Federal Communication Commission

FEM Finite Element Method

FFT Fast Fourier Transform

GCS Geographic Coordinate System

GLDB Geolocation Database

GRS80 Geodetic Reference System 80

ICASA Independent Communications Authority of South Africa

IoT Internet of Things

ITM Irregular Terrain Model

ITU International Telecommunication Union

ITWOM Irregular Terrain Model with Obstructions

MAE Mean Absolute Error

NAD83 North American Datum of 1983

NLCD National Land Cover Database

PSD Power Spectral Density

RMSE Root Mean Square Error

SDR Software Defined Radio

SEM Spectral Emission Mask

SRTM Space Shuttle Radar Topography mission

TVWS Television White Space

USRP Universal Software Radio Peripheral

WGS84 World Geodetic System of 1984

WRC World Radiocommunications Conferences

Chapter 1

Introduction

In this chapter, background is given on the radio spectrum and spectrum management. The artificial scarcity of radio spectrum is explained. The role that radio propagation models can play in addressing the spectrum scarcity problem is discussed, and the importance of the study is motivated. The research problem, aims and objectives and the research methodology are set out. The chapter is concluded with an overview of the dissertation.

1.1 Introduction

The radio spectrum is used for various applications including television and radio broadcasting, internet access and public safety. Every day the number of wireless devices that uses the spectrum increases. The radio spectrum is regulated by country-specific regulators and the International Telecommunication Union (ITU) through the means of licensing radio bands, as will be discussed in section 1.1.2 under current spectrum management mechanisms. With the current spectrum management system particular frequency bands are designated for specific use. The spectrum is basically divided into blocks and now every block is filled. There is thus a spectrum scarcity

and the demand keeps growing. It should now be noted that although the spectrum is fully set out for specific uses, all of the frequency bands are not used efficiently. For example, in the television broadcasting band not all channels are used at all geographic locations all the time. There exist geographical areas where a lot of the radio spectrum is unused.

The problem is that there is a spectrum scarcity which can be relieved if the available radio spectrum is used more efficiently. A solution to use the radio spectrum more efficiently is to share the spectrum between primary licensed users and secondary unlicensed users at the unoccupied geographic locations known as white spaces. In order to identify the unoccupied white spaces, the coverage areas of the primary users must first be known so that they are protected from harmful interference. This can be achieved by making use of cognitive radios that can sense the presence of primary users and identify white spaces as will be discussed in section 1.1.3. Alternatively, this can also be achieved by using a Geolocation Database (GLDB) that contains details about all the licensed transmitters, together with the use of radio propagation models to predict the field strength and determine the coverage area of the transmitters. This is currently the preferred option because regulators trust the GLDB and commercial off-the-shelf sensing equipment is not yet reliable enough in detecting white spaces.

Radio propagation models are at the core of spectrum sharing using a GLDB and are used to determine the coverage areas of transmitters. The shortcoming of the radio propagation models is that they usually underestimate the path loss or overestimate the coverage area. The advantage is that the primary user is better protected against harmful interference, but the disadvantage is that opportunities are wasted to share the spectrum and use it more efficiently. Radio propagation models are not only used for spectrum co-ordination and sharing but also in the planning of networks. Network operators could gain from more accurate predictions since they can better plan the infrastructure of their network and possibly save money. They can possibly provide a better service to their clients by better knowing how much users are in the coverage area and where the coverage area ends. This would be especially helpful on the edge of the coverage area. This highlights the importance of accurate propagation models.

The research focuses on augmenting the radio propagation prediction in order to make better decisions regarding the use of the radio spectrum so that it can be used more efficiently for the economic and social benefit of society.

1.1.1 Radio spectrum

The term radio spectrum refers to the full electromagnetic frequency range from 10 kHz to 300 GHz that may be used for wireless communication [14]. Like land, gas and minerals, the radio spectrum is a limited resource and is usually controlled by the government. The ITU Constitution fully recognizes the sovereign right of each State to regulate its telecommunication [15]. Radio spectrum is, however, a reusable resource. Like all limited resources radio spectrum needs to be managed in order for it to be utilised fairly and effectively for both economic growth and social benefit.

1.1.2 Spectrum management

Spectrum management is the process of regulating the use of radio frequencies to promote efficient use and to gain social benefit [14]. The purpose of spectrum management is to maximise the benefit of the usable radio spectrum. Spectrum management is a dynamic field that should be able to adjust to new and emerging technologies that can be used to optimise spectrum use.

Importance of spectrum management

A large and growing part of the world's output relies on telecommunication and therefore the use of spectrum. Spectrum is used commercially, notably for mobile and cellular applications and for broadcasting. It is also worth mentioning the Internet of Things (IoT) for commercial use. Spectrum is also used by public sector bodies for national defence, public safety (maritime, air, police) and other emergency services as

well as for transport. All the aforementioned bodies and uses require effective communication for their respective operations. The spectrum needs to be managed in order to avoid excessive interference between the different communicators [14]. Interference can render systems unusable. If the systems are unusable, the services cannot be provided as effectively, if at all to the public. The same applies to the commercial users, and the suppliers of these services are then losing money because of it. The lack of spectrum management may result in excessive interference in the radio spectrum causing degradation in services (and stifles productivity) and a monetary loss, both haltering social and economic development.

Current spectrum management mechanisms

The radio spectrum is currently managed and monitored by regulators. Different countries have different national communication regulatory bodies, and each of them can decide how to utilise the radio spectrum in their country. The Independent Communications Authority of South Africa (ICASA) is one example of a regulatory body. The ITU is an example of an international communication regulatory body. The ITU is a specialised agency of the United Nations (UN) whose purpose is to coordinate telecommunication operations and services throughout the world commonly known as spectrum harmonisation. Spectrum harmonisation refers to the uniform allocation of radio frequencies across an entire region which is not limited to a single country [16]. The ITU divided the world into three regions with each region having its own set of frequency allocations to co-ordinate. This can be seen in Figure 1.1. The distance that the electromagnetic radio waves can propagate is determined by factors including the transmitted power, channel, and frequency etc. and not by geographical borders. International coordination is essential in some cases because interference can extend beyond national boundaries and in some cases the users are inherently international for example aviation and maritime. The advantages of spectrum harmonisation are the reduction in interference along borders, and it helps with international roaming, reducing the cost of telecommunication equipment including the cost of handheld de-

vices [14]. Usually, international bodies set out high-level guidance and regulations which national bodies adhere to but usually have a more detailed policy within that bounds [14]. The regulatory bodies are free to manage the radio spectrum in their country as long as it adheres to the guidelines and regulations set out by the ITU.

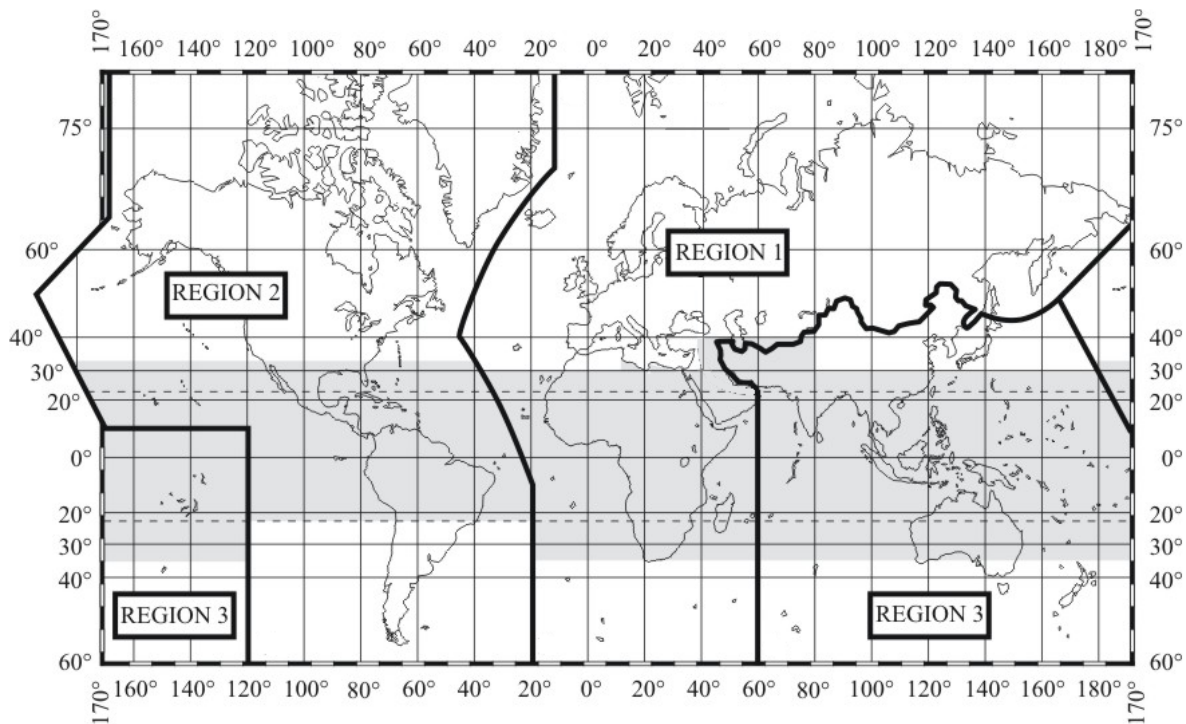


Figure 1.1: ITU regions adapted from [1]

Regulators often decide on the use of a band i.e. the purpose of the band or the service provided and the technical details that include the maximum radiated power and the Spectral Emission Mask (SEM) for example. They also often decide who may legally use the band through right of use licensing regimes. Figure 1.2 is an example of how the radio spectrum is divided into blocks or bands of frequency that is intended for specified uses. Regulators keep tight control over the use of the spectrum making it easier to model the interaction between neighbouring services and to ensure that there aren't excessive interference between users. It also makes spectrum planning easier. The spectrum can be planned to achieve a higher efficiency of spectrum use by packing a large number of users into the spectrum. This can be done because the regulator can pack alike services in neighbouring bands closer together. Alike services tend to

interfere less than unlike services [14] due to technical requirements. Examples of such technical requirements include the spectral emission mask and guard bands. It can increase the capacity of the spectrum. Increasing the capacity of the spectrum doesn't necessarily mean that the benefit that society gain from the spectrum is maximised because the spectrum can be completely used by a low-value application. A low-value application is one that doesn't have monetary or social value. Increasing the capacity of the spectrum that is used by a high-value application will increase the benefit that society gain from the spectrum and will contribute to effective spectrum management. Thus there is an opportunity to improve spectral efficiency by repurposing the low-value and unused spectrum by means of different spectrum sharing schemes discussed later.

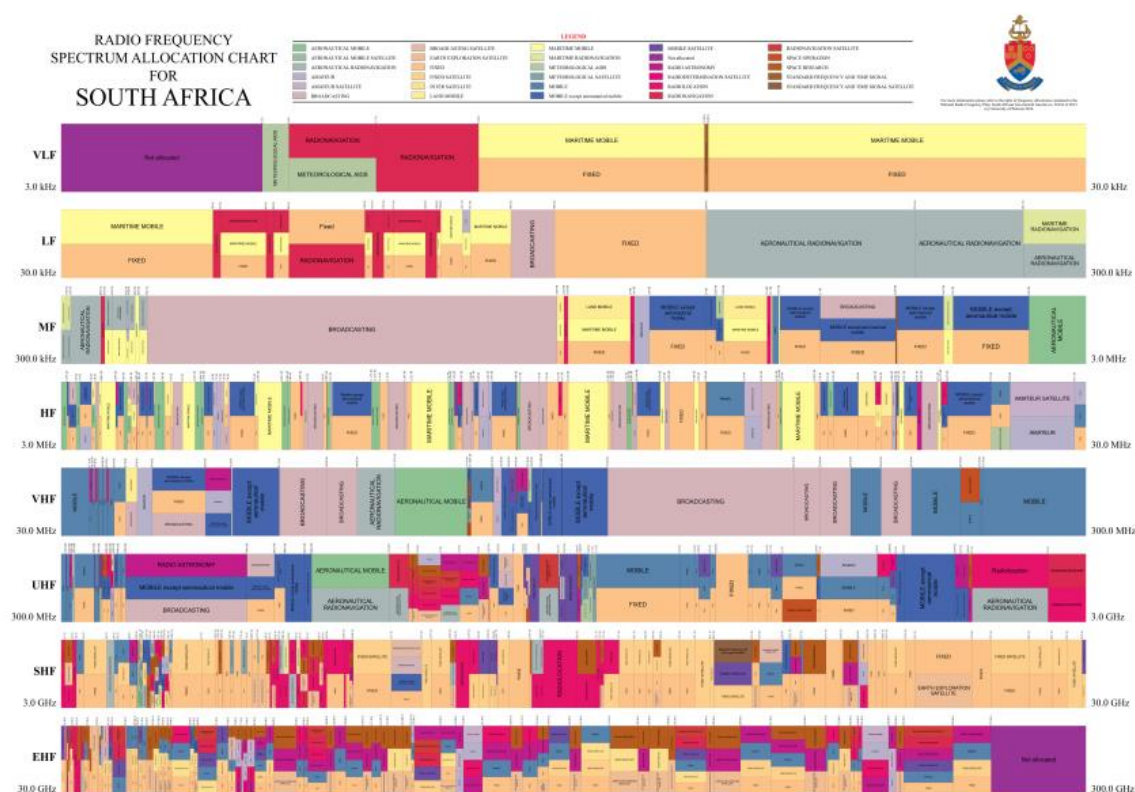


Figure 1.2: South Africa 3kHz - 300GHz regulated frequency allocations chart [2]

Current management shortcomings

The current method of spectrum management allows for the use of spectrum in a stable and predictable environment but does not necessarily achieve the goal of spectrum management, which is to maximise the benefit that society gains from the use of spectrum whether it is socially or economically [14]. Acquiring spectrum usage rights is slow and expensive. It may slow down development because spectrum can't be acquired to test the new technologies or protocols. ITU World Radiocommunications Conferences (WRC) are held every five years. The responsibility of the WRC is to review and revise regulations. Some developments are dependent on changes in regulation and are delayed by this slow process.

Another regulatory shortcoming is that larger companies can buy a temporary licence for a certain frequency band but not use it, just in order to prohibit their competitor or other smaller companies from utilising the band. The company may benefit financially from that decision but not society.

1.1.3 Spectrum sharing

An effective way to use the spectrum more efficiently is to share it. Spectrum sharing is the simultaneous usage of a specific radio frequency band in a specific geographical area by a number of independent entities, leveraged through mechanisms other than traditional multiple- and random-access techniques [17]. The spectrum can be shared in time, frequency and space, and there are a few techniques to accomplish this. Spectrum can be shared in the dedicated unlicensed bands like the 2.4 GHz ISM band. An example of spectrum sharing in the unlicensed band is Wi-Fi. Spectrum can also be shared in the licensed bands like the television broadcasting band (470 to 694 MHz in South Africa). An example of sharing spectrum in the licensed band is Television White Space (TVWS). One technique to share is with dynamic spectrum access and cognitive radio which makes use of sensing and understanding the environment the device is operating in. Another technique is to make use of a geolocation database that

knows where primary users transmit and can tell a secondary user where and how it can utilise a certain frequency band.

Dynamic spectrum access and cognitive radio

Dynamic spectrum access is a spectrum sharing technique that allows secondary users to access the ample spectrum holes or white spaces in certain areas of the licensed spectrum bands. In order to access and use the spectrum holes or white spaces, the secondary user needs to determine where these holes and spaces are, if any. Not only should the devices find the holes but also adapt to operate in the various channels. A Cognitive Radio (CR) is a smart radio that can understand the environment or context it is operating in and configure itself to communicate in that environment or context [18]. A cognitive radio will, therefore, be able to sense the spectrum holes or white spaces and adapt to communicate in the empty channels. True cognitive radios are very expensive and fairly (physically) large equipment. It is not yet a viable option for, example, mobile use. The other challenge with cognitive radios and dynamic spectrum access is to minimise the time it takes to find an unused channel and configure itself to use that channel. Another approach is to make use of a centralised point to manage the dynamic spectrum of all the unlicensed users. This approach comes with a set of its own challenges. The centralised station will still rely on sensing the spectrum. The use of smaller sensors scattered in the area could collect information about the spectrum occupancy and relay the information to the central station. The main challenge would be to minimise the time the process takes in order to have enough time left to transmit especially since the occupancy as a result of primary users may change dynamically [19].

Use of geolocation database

A geolocation database with regards to TVWS is a database containing information of primary transmitters that should be protected. The information includes the ge-

ographical location of all the transmitters, antenna height, polarisation, transmitting power and frequency amongst others. From this information and assuming a certain threshold, the coverage area of the transmitter can be predicted. Assuming there exist a receiver on the very edge of the coverage area, a secondary user can be permitted to transmit at a certain geographical location at a specified frequency and power so that the primary user is still protected from harmful interference.

Geolocation databases are used in the application of TVWS in the United States of America, Canada and the United Kingdom [20–22]. South Africa also published regulations on the use of TVWS in which it will make use of a geolocation database [23]. The US also has a service called Citizen Broadband Radio Service (CBRS), which makes use of a geolocation database [24]. Thus the method of using a geolocation database for spectrum sharing is commonly proposed.

The present geolocation database is conservative in the sense that a large range of protection contour is defined in order to ensure no interference to the primary users although the primary users might be far away. Adding to the protection range is the inaccuracy of field strength predictions used to calculate the contour. Usually the propagation models used for the prediction underestimates the path loss thus overestimating the coverage area of the transmitter. This results in many otherwise vacant channels being left unused. This also presents another opportunity to enhance the efficient use of spectrum.

1.1.4 Propagation modelling

A radio wave propagation model is a mathematical formulation of the characteristics of radio wave propagation as a function of, most notably, frequency and distance. Propagation models are used to predict the attenuation of signal power in a wireless channel. Propagation models are therefore commonly used in the planning of networks. It is faster and less expensive than taking field strength measurements to determine the coverage area of a network. The physical network also doesn't have to be in

place before the coverage area can be determined.

Propagation models are used to predict the coverage areas of networks or transmitters when using a geolocation database for TVWS. The accuracy of the propagation model used is, therefore, important. If the propagation model is inaccurate the coverage area of a network may be overestimated, reducing the amount of spectrum to be reused. The coverage area could also be underestimated, and that would amount to harmful interference to the primary licensed spectrum user when a secondary user is allowed to transmit in that area.

There are many propagation models that one can use. Propagation models are usually suitable for specific use cases. For example there are small scale and large scale propagation models, where small scale models are typically used indoors and large scale models are used over longer distances like radio broadcasting networks. Some models are specifically made for urban use, while some are more general. Propagation models differ in complexity. Empirical models are less complex than ray-tracing models and are computationally less expensive but usually less accurate. There are thus many factors to consider when choosing a propagation model to be used together with a geolocation database in order to improve the efficiency of spectrum usage.

Modelling the attenuation of signal power in a wireless channel is no trivial task, and all propagation models have a degree of inaccuracy. Large scale radio propagation predictions suffer from inaccuracies because there are so many parameters with variability to account for. It is not always feasible or possible to take all parameters into consideration. Small scale radio propagation predictions like the Finite Element Method (FEM) are more accurate but much more computationally intensive. It is not feasible to apply it to a large area. A prediction error of 8-15 dB can be expected for a priori path loss models [25].

1.2 Proposed research

A hybrid approach, where propagation prediction and spectrum measurements are combined, has been the topic of numerous research endeavours. The goal with all these approaches, in general, is to improve the mean error between propagation prediction and reality (spectrum measurement).

The focus of this research is to develop a method to augment large scale propagation predictions by using field measurements and to quantify the reduction in the mean error achievable using said method. In this context, field measurements are high resolution terrain Digital Elevation Model (DEM), land cover information and field strength measurements.

Therefore a long term measurement study will be done in order to collect field strength measurements. This measurements, together with other field measurements like land cover data, will be used to augment a large scale propagation model. The error between field strength measurements and predictions will be quantified and compared to that of the baseline large scale propagation model. Comparing the results of the baseline model and the augmented model should lead to a conclusion regarding the extent to which the error can be reduced. The research should also give an indication of what is needed for a general method to augment large scale propagation models with.

1.3 Importance of study

An improvement in propagation predictions can have both social and economic benefits. In the implementation of TVWS, a reduced error of a few decibels can potentially free up tens of square kilometres of area for secondary usage depending on the radiated power of the transmitter. The monetary value of the extra available spectrum can be significant.

The augmented propagation prediction may be the difference between a remote settlement wrongly included in the coverage area of the transmitter or not. The implication if the area is wrongly included is that this area won't benefit from the primary service in that area and will also not be able to make use of the possible secondary service. Best case is that the augmented radio propagation prediction puts the remote settlement outside the coverage area of the primary transmitter, which it actually is, and that this community can have broadband internet access through the secondary network. This will result in a social benefit to the community and create opportunities for economic benefit. The broadband penetration of South Africa could be increased by utilising TVWS, and the impact could be far greater if the underlying large scale propagation model used in the GLDB is more accurate. The increase in broadband penetration could lead to more wealth generation and job creation for the country.

Broadcasters could also benefit from more accurate large scale propagation models. Planning of a network and the needed infrastructure could be better with improved propagation predictions, which in turn can save money on infrastructure and maintenance. The network can also be planned in such a way to improve the quality of service since broadcasters and network operators might have a better understanding of the coverage area and the population therein.

It is too expensive, time-consuming and sometimes impossible to determine the coverage area or available white spaces by only taking field measurements. Contrarily propagation models aren't accurate; hence, an in-the-middle hybrid approach may be a good compromise.

1.4 Research question/statement

The research problem can be stated as follows:

Develop a method to augment large scale propagation predictions by using field measurements and quantify the reduction in mean error achievable using said method.

The research aims to answer the following questions:

- Can large scale propagation models be augmented with field measurements in order to reduce the mean error between measurements and predictions, and if so, by how much can the error be reduced?
- What field measurements are needed to augment large scale propagation models?

1.5 Research aims and objectives

The aim of the research is to create a method to augment large scale propagation models in order to reduce the mean error between field strength measurements and predictions made by the models.

The following objectives are identified:

- study and understand large scale propagation models;
- determine significant field measurements to augment large scale propagation models with;
- identify transmitters and receivers sites for the case study;
- conduct a long-term field strength measurement study;
- establish a baseline for the error between field strength measurements and predictions;
- implement a method to reduce the mean error between field strength measurements and predictions;
- quantify the reduction in error achievable by using the implemented method;
- and finally, draw a conclusion of the input parameters needed for the method of augmenting large scale propagation models with field measurements.

1.6 Research methodology

The research is approached in four phases, where each phase is arranged into three important aspects of the method for augmenting large scale propagation models. The three aspects are a) field strength measurements, b) other field measurements like land cover data, and c) a method to fit the propagation model to the measured data in order to reduce the mean error between measurements and predictions. Each phase will address one or more of the objectives that were identified. A brief description of each phase follows.

1.6.1 Literature study

In this phase, research is done on RF propagation, field measurements and a method for reducing the error between measurements and predictions. Related work in this area of research is also studied to get an understanding of the possible solutions to the problem and give an indication of significant field measurements to use. The purpose of this phase is to get a thorough understanding on how RF propagation models work and where the shortcomings are, what should be considered when taking field strength measurements and how the error between the output of the model and real measurements can be evaluated and reduced.

1.6.2 Research model design, exploratory data analysis and implementation

From the knowledge of the literature study, propagation models are identified that will be augmented. The field measurements to augment the models with are also established. The method for fitting the propagation model to field strength measurements and reducing the error is also established. The transmitter- and receiver-sites for the case study are identified, the measurement study setup is defined and executed. The

data from the measurement study is analysed and processed so that it can be used in the implementation of the method. The other field measurements, i.e. land cover data, is also analysed and processed to be used in the implementation of the method. The tools needed to implement the method for augmenting large scale propagation models are identified, and the toolchain is setup.

1.6.3 Results of augmenting large scale propagation models

In this phase, a baseline for the mean error between measurements and predictions is first established. The results from the method to augment large scale propagation models with field measurements are analysed to gain insight into the method as well as the achievable reduction in the mean error. The output of the method is compared to the baseline, and the reduction in error is quantified.

1.6.4 Verification and validation

Finally, the three important aspects of this study are verified, and the augmented propagation models are validated. The processing of the field strength measurements is verified so that one can have confidence that the average received power used as the ground truth is the actual average received power. The processing of the land cover data is verified to ensure that the land cover data is accurate and that the model is correctly augmented with the data. The method of minimising the error is also verified to ensure that it is implemented correctly and works as expected.

The augmented propagation model is then validated against a set of transmitter-receiver pairs that weren't used in the training of the model. From the validation results, conclusions can be made about the method of augmenting large scale propagation models. Conclusions can be made about the generalisation of the method, whether the reduction in error is valid, and whether the measurement study was sufficient.

1.7 Dissertation overview

The remainder of the dissertation is structured as follows. A literature study in chapter 2 will discuss the concepts of RF propagation, field measurements and model tuning and will provide the background required for this study. The research model and implementation will be discussed in chapter 3. The measurement study and processing of both field strength measurements and land cover data will be explained. The implementation of the method to reduce the error between measurements and predictions will also be explained. The exploratory data analysis and results will be discussed in chapter 4. First, the results of processing the land cover data and field strength measurements will be discussed. That is the results of the exploratory data analysis. Thereafter the results of implementing the method to augment large scale radio propagation models, will be given and discussed. In chapter 5, verification and validation will be done. The processing of the field strength measurements from the measurement study, as well as the processing of the land cover data will be verified. The implementation of the method to reduce the error will also be verified. The augmented propagation models will then be validated using an independent set of transmitter-receiver pairs. Finally, the dissertation will be concluded in chapter 6 with a discussion on the findings of the research and recommendations for future work.

Chapter 2

Literature study

In this chapter, background on radio wave propagation and large scale propagation models are discussed. Attention is given to the Longley-Rice Irregular Terrain Model in section 2.3.2. Field measurements are discussed in section 2.4.2 and model tuning, and evaluation is discussed in section 2.5. The chapter is closed with a review of related work in which the authors improved the accuracy of large scale propagation models using field measurements.

2.1 RF propagation

Assuming a vacuum or perfect scenario, radio wave propagation or electromagnetic wave propagation is described by Maxwell's equations, which state that a changing magnetic field produces an electric field and a changing electric field produces a magnetic field [26]. Thus a changing magnetic field produces a changing electric field which in turn produces a changing magnetic field and so on causing a self-propagating wave. As the wave propagates outwards and away from the source, energy is dispersed until the electric and magnetic fields are so weak that the wave can no longer propagate. If the wave is not propagating in a vacuum there may be additional losses.

The wave can propagate through obstructions and undergo diffraction, reflection and absorption. All of which contribute to additional loss of energy and reduces the range of the propagating wave.

2.1.1 Inverse square law

To calculate the power density produced by a radiating antenna consider an isotropic antenna in the centre of a sphere. The isotropic antenna radiates power equally in all directions thus the power density must be equal on all parts of the surface of the sphere. Since the antenna is in the sphere all of the power must pass through the entire surface of the sphere. Assuming that the isotropic antenna is a perfect transmitter and all of the power entering from the feeder cable is radiated into free space, and knowing that the area of a sphere of radius r is equal to $4\pi r^2$, the inverse square law follows in equation 2.1 [18]:

$$P_d = \frac{P_t}{4\pi r^2} \quad (2.1)$$

,where P_d is the power density produced by the antenna measured in watts per meter, P_t is the power entering the feed of the antenna measured in watts and r is the distance from the antenna measured in metres. It can be seen that the power density reduces with the square of the distance. This is true in free space and used as the basis for propagation models.

2.1.2 Free space loss

Radio waves travel from a source or transmitter to the surrounding area at approximately 3×10^8 m/s, the speed of light, in free space. Free space is actually a vacuum but clear air is a good approximation to this [6]. There are different types of antennas and in order to compare them a reference antenna needs to be defined. The isotropic

antenna, a fictional antenna in which the transmitted power is radiated equally in all directions [27] is defined as the reference antenna. As seen in section 2.1.1, some of the transmitted power is lost along the way when electromagnetic waves propagate through free space. The free-space loss is defined as the loss between a transmitter and receiver in free space, expressed as a power ratio [27].

If both antennas are isotropic radiators the free-space loss, FSL , in dB is given by equation 2.2:

$$FSL = 32.4 + 20 \log d + 20 \log f \quad (2.2)$$

, where d indicates the distance between the antennas in kilometres and f indicates the frequency in MHz [6]. The difference between the received power and the transmitted power is called the path loss. The path loss indicates how much power is lost along the path of propagation. Equation 2.2 can be derived from the Friis transmission equation given by equation 2.3 [28]:

$$P_r = \frac{P_t G_t G_r \lambda^2}{(4\pi R)^2} \quad (2.3)$$

, where P_r is the power of the receiving antenna, P_t is the output power of the transmitter, G_r and G_t are the gains of the receiving and transmitting antennas, λ is the wavelength in metres, and R is the distance between the transmitter and receiver in metres. Free space loss is the ratio of the received power to the transmitted power. For isotropic antennas the gains are set to one.

The link loss is the free-space loss less antenna gains plus any miscellaneous losses. The link loss is given by equation 2.4:

$$link\ loss = 32.4 + 20 \log d + 20 \log f - G_t - G_r + L_m \quad (2.4)$$

, where G_t and G_r are the gains of the transmitting and receiving antennas respectively,

in dBi and L_m represents any miscellaneous losses in the system in dB [6]. f is the frequency in MHz and d is the distance between the transmitter and receiver in kilometers. Any antenna gain must be quoted relative to a reference antenna, in this case an isotropic radiator, and therefore the antenna gains are given in dBi, decibels relative to the isotropic antenna. Similarly an antenna gain can also be expressed relative to a half wave dipole, that is in dBd. In subsection 2.1.4 the relation between dBi and dBd is discussed.

2.1.3 Gain

Antenna gain is not amplification of signal power but rather how well the power is directed. It can therefore be seen as passive gain. The term antenna gain describes how much power is radiated in the direction of peak radiation to that of an isotropic antenna [29]. An isotropic antenna radiates power equally in every direction. Directional antennas take the equivalent power that an isotropic antenna radiates and concentrates it in a specific direction. It therefore appears to be amplified or have gain.

Usually the gain is considered in a single plane of interest for example the horizontal or vertical plane.

2.1.4 Effective radiated power

Effective Radiated Power (ERP) is the electromagnetic power that is radiated from the antenna into space. That includes the output power of the transmitter, plus the gain of the antenna, minus all the attenuation and losses from cables and connectors between the transmitter and antenna. Otherwise stated, the ERP is the product of the input power and gain of the antenna relative to a dipole [30]. The ERP is always given with respect to a certain direction. For antennas with a non-uniform gain pattern, the direction is that of the maximum gain.

The ERP is usually related to a half wave dipole and is given in dBd, decibels relative to

a dipole antenna. The Effective Isotropic Radiated Power (EIRP) is the ERP relative to an isotropic radiator. The EIRP allows making comparisons between different emitters regardless of type, size or form.

It can be shown mathematically that the relation between the EIRP and the ERP (for a half wave dipole) is as given in equation 2.5. The EIRP will be given in dBi. This is a useful relation since the ERP is mostly reported in information regarding transmitters. From equation 2.5 it can be inferred that an ideal half wave dipole antenna has a gain of 2.15 dBi.

$$EIRP(dBi) = ERP(dBd) + 2.15 dB \quad (2.5)$$

2.2 Multipath

In wireless telecommunications a radio wave can follow multiple paths from a transmitter to a receiver because of scattering, diffraction, refraction and reflection. Figure 2.1 illustrates these effects. This is known as multipath propagation. The signals arrive at the receiver with different phases due to the different paths and lengths thereof travelled. The signals can either be in phase and constructive interference could take place, increasing the signal strength or it can be out of phase and destructive interference could take place, reducing the signal strength. Perfect constructive interference causes peaks in the received power and perfect deconstructive interference causes nulls in the received power. In addition the transmitter, receiver or obstacle that reflects the signal can move and it can cause a varying received field strength due to the change in relative path lengths. This is known as multipath fading.

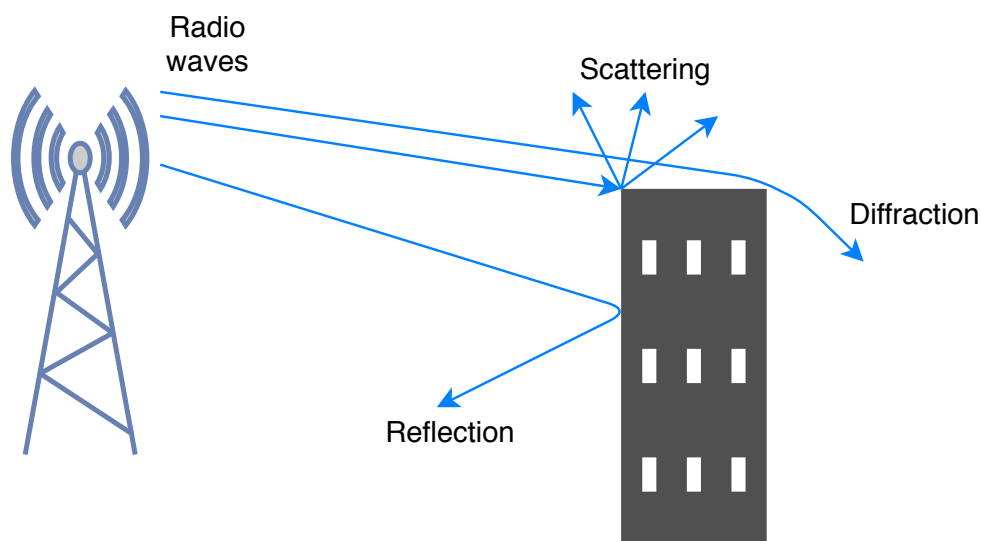


Figure 2.1: Multipath effects due to diffraction, reflection and scattering

2.2.1 Scattering

When an RF wave comes in contact with a rough area, the wave is scattered off the jagged edges into multiple waves. This causes a reduction in the main signal strength. A rough area is an area that has sharp or jagged edges that is much smaller than the wavelength of the incoming electromagnetic wave. Objects like these typically include street lights, signs, small trees and leaves.

2.2.2 Diffraction

Diffraction is the bending of waves around obstacles and openings. The amount of diffraction is dependant on the wavelength and the size of the object. When the wavelength is larger than the obstructing object it will diffract or bend around the object. The larger the wavelength the more it will diffract.

Huygens' principle states that every point on a wave-front is a source of wavelets, which spread forward at the same speed of the original wave [31]. This principle can be used to explain why diffraction is possible [32].

Knife-edge diffraction is the phenomenon in which an electromagnetic wave diffracts as it hits the sharp edge of an obstructing obstacle and changes the direction of propagation. Sharp edges of obstacles make diffraction more eminent. The diffraction loss can be approximated in propagation models for a single edge or for multiple edges.

If the direct line-of-sight is obstructed by one knife-edge type obstacle with height h_m , the diffraction or Fresnel parameter v is defined as follows [6,33].

$$v = h_m \left(\sqrt{\frac{2}{\lambda} \left(\frac{1}{d_T} + \frac{1}{d_R} \right)} \right) \quad (2.6)$$

,where d_T and d_R are the respective distances from the obstacle to the transmitter and the receiver, in meters. λ is the wavelength in meters. The diffraction loss can then be calculated from the parameter v . Given v , a good approximation for the diffraction loss, G_{diff} in dB, is given as follows [33].

$$G_{diff} = \left\{ \begin{array}{ll} 0 & \text{for } v < -1 \\ 20\log(0.5 - 0.62v) & \text{for } -1 \leq v \leq 0 \\ 20\log(0.5 \exp(-0.95v)) & \text{for } 0 \leq v \leq 1 \\ 20\log(0.4 - \sqrt{0.1184 - (0.38 - 0.1v)^2}) & \text{for } 1 \leq v \leq 2.4 \\ 20\log\left(\frac{0.225}{v}\right) & \text{for } v > 2.4 \end{array} \right\} \quad (2.7)$$

There are three common models to compute the diffraction loss over multiple knife edges. For the purpose of explanation, two knife edges will be considered.

Bullington

The first model proposed by Bullington [34], defines a new effective obstacle at the point where the line-of-sight from the two antennas cross [3,6] as seen in Figure 2.2. The two triangles with solid lines indicate the two obstacles between the transmitter and receiver. The triangle in the centre indicates the new hypothetical obstacle that

will replace the others obstacles when the diffraction loss is calculated. The dotted line indicates the line-of-sight path between the transmitter and receiver.

The formula for single knife-edge diffraction is then applied but using the new effective obstacle.

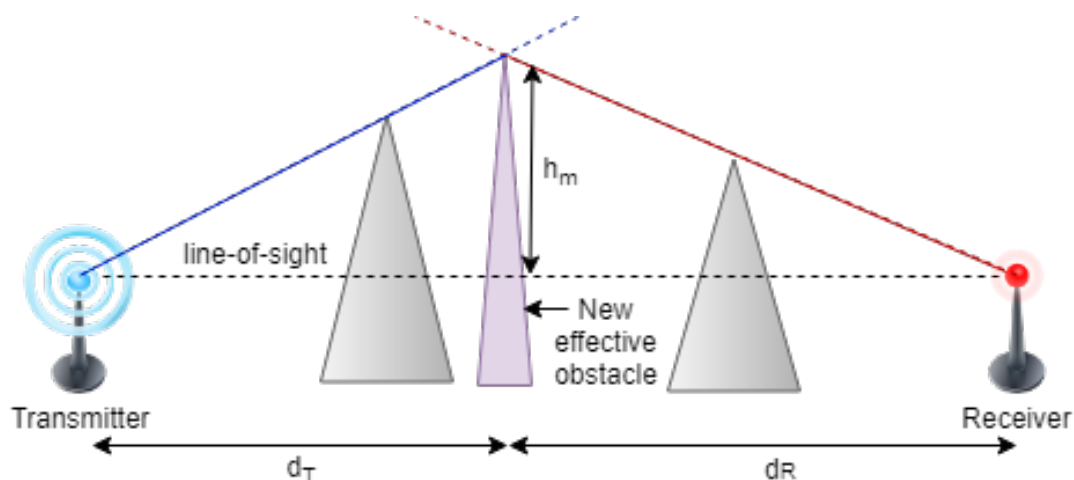


Figure 2.2: Double knife-edge diffraction using the Bullington method adapted from [3,4]

Epstein and Peterson

The second model is proposed by Epstein and Peterson [35]. In this model the link is divided into two parts or hops, each involving only one diffracting edge as illustrated in Figure 2.3. The first hop is from the transmitter to the second obstacle, with the first obstacle in the path. The second hop is from the first obstacle to the receiver, with the second obstacle in the path. The diffraction loss for each obstacle is calculated and added together for the calculation of the final diffraction loss [6]. The height used in the calculation of the Fresnel parameter is the height of the obstacle protruding above the line of sight path between the points of a hop. The dotted lines in Figure 2.3 depict the line-of-sight paths between the transmitter and receiver and also the various hops.

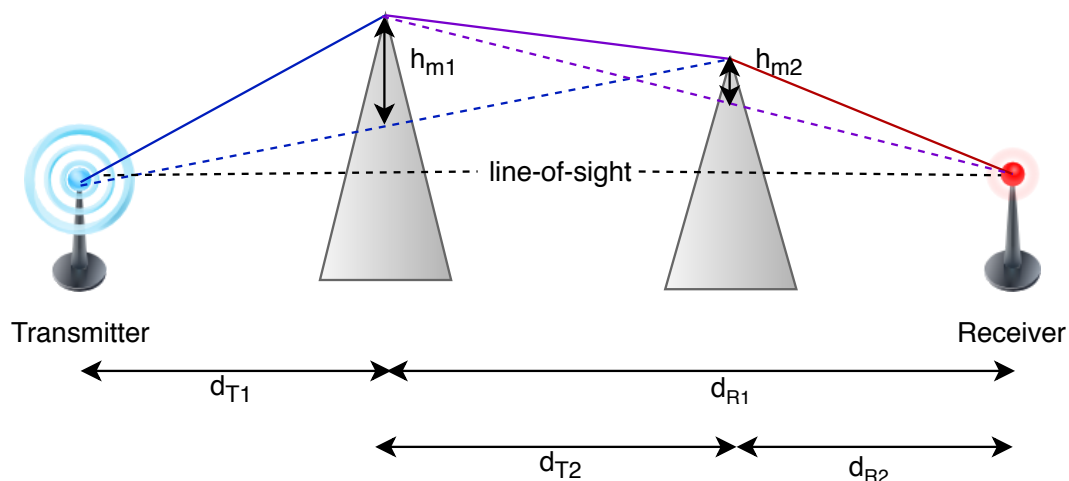


Figure 2.3: Double knife-edge diffraction using the Epstein-Peterson method adapted from [5,6]

Deygout

The third model is proposed by Deygout [36] where the main obstacle is used to calculate the diffraction loss [5]. The main obstacle is calculated and called the principle edge. The main obstacle is determined by calculating the Fresnel parameter for every edge as if it were the only obstacle. The edge with the highest Fresnel parameter will then be the principle edge [6]. Working from the principle edge it is treated as if there is a new path between the transmitter and the principle edge and that path is the reference plane for other obstacles. The loss for the intermediate obstacle is calculated with the height of the obstacle adjusted according to the reference plane, which is the path from the principle edge to the receiver. That is, the height of the obstacle is the height protruding above the line-of-sight path between the points, in this case the first obstacle and receiver as illustrated in Figure 2.4. The reference plane will be the purple dotted line and the height is denoted by h_{m2} . In short the method searches for the main obstacle and calculates the diffraction loss. Diffraction losses over secondary obstacles are calculated and added to the main diffraction loss. This method ignores edges within the first Fresnel zone. Figure 2.4 illustrates double knife-edge diffraction calculations with the Deygout method. Assume the first obstacle is the principle edge.

The dotted line indicates the line-of-sight paths between the points. The heights are indicated as the height of the obstacle above the line-of-sight plane.

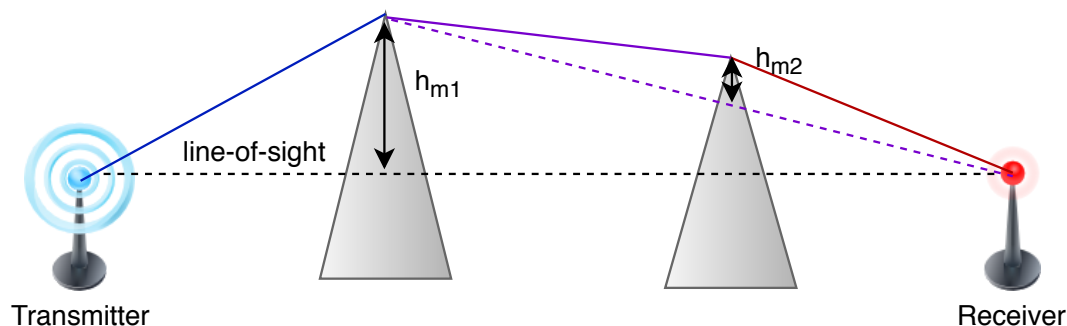


Figure 2.4: Double knife-edge diffraction using the Deygout method adapted from [3,5]

All the models are approximations and have the potential for error, especially if the edges are close to one another.

2.2.3 Reflection

Reflection is when electromagnetic waves bounce off a surface. There is normally loss accompanied by the reflection of electromagnetic waves due to either absorption or some of the signal passing through the medium. A variety of surfaces can reflect radio waves. For long distance communication the sea provides a very good reflective surface, the desert a poor reflective surface and other land areas are in between [37]. The reflectivity is related to the conductivity of a medium or surface [37]. A good conductor will reflect waves well. Propagation models will typically include a conductivity parameter that is used in the calculations of wave propagation. Table 2.4 gives typical values for the ground constants used in propagation models for various surfaces and from this it can be seen that the sea provides a very good reflective surface because of the high conductivity parameter. For short range communications buildings can be very good reflectors depending on size and structure. The smoothness of the surface also has an effect on the reflection of radio waves.

2.2.4 Refraction

Refraction is the phenomenon of electromagnetic waves being deflected when passing through the interface between two mediums with varying density and refractive index. For radio signals the change in direction due to refraction is gradual and causes the signal to bend rather than abruptly change direction.

2.2.5 Fresnel zone

A Fresnel zone is a cylindrical ellipse drawn between a transmitter and receiver as shown in figure 2.5. The size of the ellipse is determined by the frequency of the radio wave and the distance between the transmitter and receiver [38]. It used to calculate reflections and diffraction loss between a transmitter and receiver. When a signal reflects off a surface its phase reverses by 180° and due to the slightly longer path it takes from the transmitter to the receiver it also shifted further in phase relative to the original signal. These phase shifts have an influence on the signal strength at the receiver. The Fresnel zones describe how the phase shift would influence the signal at the receiver. There is an infinite number of Fresnel zones but only the first three are significant for calculating reflection and diffraction loss using a propagation model.

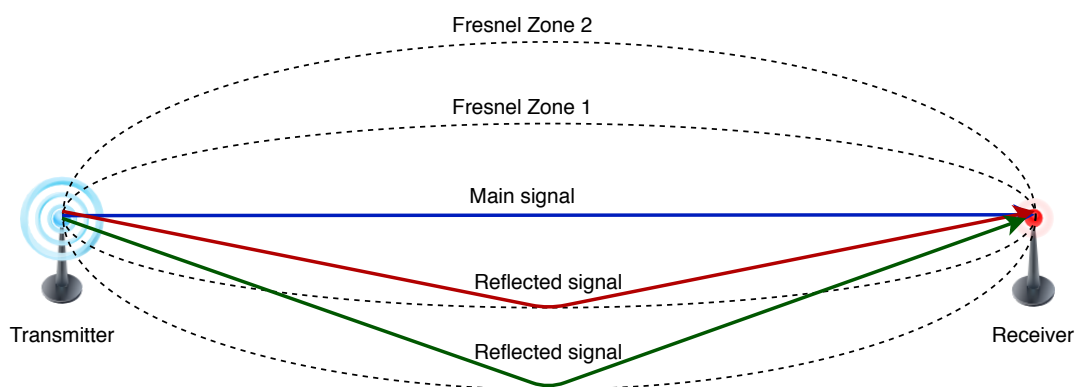


Figure 2.5: Fresnel Zone

The first Fresnel zone radius is calculated so that the path length difference between

the main and reflected signal results in a 180° phase shift [38]. This phase shift added to the phase shift of the reflection causes a 360° phase shift and therefore has no negative effect. The second Fresnel zone radius is calculated so that the path length difference results in a 360° phase shift and the third zone in a 540° shift and so on. The second Fresnel zone describes a total phase shift of 180° of the reflected signal at the receiver which will have a negative impact on the received signal.

The odd numbered Fresnel zones result in a 360° shift and have no negative effect on the received signal but the even numbered Fresnel zones result in a 180° shift and have a negative phase cancelling effect. However, the rule of thumb to maintain a good connection is that the first Fresnel zone should be 60% clear of obstructions measured from the centreline to the edge boundary of the Fresnel zone.

2.3 Radio wave propagation models

A radio wave propagation model is a mathematical formulation of the characteristics of radio wave propagation as a function of, most notably, frequency and distance. Propagation models are used to predict the attenuation of signal power in a wireless channel.

There are various types of propagation models that predict the attenuation loss in a wireless channel. Most models are either empirical or deterministic while some are a combination of the two called hybrid models. Propagation models are developed for specific scenarios or use cases where the transmitter operates in a specific frequency range and environment. Some models are better for indoor propagation prediction while others are better for outdoor propagation prediction and longer distances. Choosing an applicable model for the application is important.

Propagation models are used to determine the coverage areas of transmitters in a network and are therefore useful when planning a network.

2.3.1 Empirical models

Empirical propagation models describe radio wave propagation based on previous field strength measurements and other measured data. It is linked to the environment and the parameters of the measurement campaign. The advantages of empirical models are their speed of execution because of their simplicity and their limited reliance on detailed knowledge of the terrain [39]. The models do not rely on a detailed knowledge about the terrain because the mathematical equations are fitted to empirical data and the equations are implemented in the propagation program. Because these models are not site-specific they can have a greater error in the predicted field strength value when compared to other site-specific and more complex models. This is a disadvantage of empirical models.

The following common empirical propagation models are considered for this study.

Hata models

The Okumura-Hata model [40] is a radio propagation model for predicting path loss in urban areas where diffraction, scattering and reflection of city structures play an important role [28]. It is based on the empirical data of the Okumura model which was collected in the city of Tokyo, Japan. There are three different formulas used for urban, suburban and open areas.

The model can be used for a frequency range of 150 MHz to 1500 MHz, distances of 1km to 20km, and effective antenna heights of 30m to 200m for base stations, and 1m to 10m for the mobile station. The model is usually used for predicting propagation loss in cellular systems [28].

The shortcoming of the Okumura-Hata model is that it is applicable for short distances between a transmitter and receiver. The Hata-Davidson model adds correction terms to the Hata model and extends the distance, d , to 300km [41].

ITU-R P.1546

The ITU-R P.1546 [42] is a method for point-to-area radio propagation predictions for terrestrial services in the frequency range 30 MHz to 3 GHz [6, 42]. It is intended for use on radio links over land, sea and land-and-sea combined paths for distances up to 1000 km. The method is based on interpolation/extrapolation from empirically derived field-strength curves as functions of distance, antenna height, frequency and time variability. The calculation procedure also includes corrections to the results obtained from this interpolation/extrapolation to account for terrain clearance and other obstructions [42]. The ITU-R model can also be adjusted to produce more accurate results. It can be adjusted for different climactic regions, paths and other parameters. The model also takes average terrain into account but it is not path specific. The model is not path specific because it uses averaged terrain heights in point-to-area prediction mode but the exact path between two specific points is not considered.

2.3.2 Deterministic models

Deterministic models estimate the field strength directly from the path profile (profile of the terrain between the transmitter and the receiver) and other given input parameters. It is not based on field strength measurements like empirical models. These methods adjust the terrain elevation to take the earth's curvature into account. In addition to free space losses, these models also take into account losses due to diffraction in cases where there is insufficient clearance between the radio path and the terrain (or structures on the terrain). Accordingly, deterministic models require detailed knowledge of the terrain.

Longley-Rice irregular terrain model (point-to-point mode)

The Longley-Rice irregular terrain model, hereafter referred to as the Irregular Terrain Model (ITM), is a radio propagation model used to predict signal attenuation over

irregular terrain for a frequency range of 20 MHz to 40 GHz and path lengths of 1 km to 2000 km. It is based on electromagnetic theory and Maxwell's equations and it also includes statistics that account for time, location and situation variability of the radio signal. The model is not concerned with channel characterisation. The ranges for which the model is intended to be used are given in Table 2.1.

Table 2.1: ITM valid input parameter ranges [10]

Parameter	Minimum	Maximum
Frequency	20 MHz	40 000 MHz
Distance	1 km	2000 km
Antenna heights	0.5 m	3000 m
Surface refractivity	250 N-units	400 N-units

The model can be used in point-to-point and point-to-area mode. The difference between point-to-point and point-to-area mode is that the area mode does not require detailed terrain information or even a specific path. The area mode will thus be less complex but also less accurate than the point-to-point mode. The point-to-point model is usually implemented in a computer program because it requires detailed terrain information and is computationally complex. The NTIA developed the ITM software and made the FORTRAN code available. Various other software has since been developed that implements the ITM algorithm to predict the path loss in a wireless channel. The Federal Communication Commission (FCC) in the U.S. uses the Longley-Rice ITM as the core software in its Longley-Rice Digital Television (DTV) coverage and reception analysis implementations [43].

General input parameters to the ITM in point-to-point mode are given in Table 2.2. The time, location and situation variability can also be set by the user.

Table 2.2: ITM point-to-point general input parameters [8]

Parameter	Unit	Description
d	km	Distance between two terminals.
f	MHz	Frequency of radio wave.
h_{e1}, h_{e2}	m	Antenna effective heights.
d_{L1}, d_{L2}	km	Distance from each terminal to its radio horizon.
θ_{e1}, θ_{e2}	rad	Elevation angles of the horizons from each terminal at the height of the antennas.
Polarisation		Polarisation of antenna (horizontal or vertical).
Δh	m	Terrain irregularity parameter.
N_s	N-units	Surface refractivity.
σ, ϵ	Siemens/m, None	Electrical ground constants: Earth conductivity, Earth Dielectric Constant.
Radio climate		Expressed qualitatively as one of a number of discrete climate typess.

Time variability

Time variability accounts for variations of median values of attenuation due to slow changes in the environment for example changes in atmospheric refraction [44]. The computed field strength value is a median value per unit of time. The actual field strength at the receiver is expected to be equal or higher than the predicted value for the fraction of time specified. If the fraction of time is specified as 50% the actual value will be greater or equal to the predicted value for half the time and lower for the other half.

Situation variability

Situation variability accounts for the variations that occur in systems with the same parameters and environmental conditions. That is the link might appear to be exactly

the same but the actual field strength values differ. This accounts for “hidden” variables or variables chosen not to be controlled since the effect of these variables are not fully understood or characterisable in the modelling process. Systems would behave identical if not for these uncontrolled or hidden variables. The higher the situation variability the lower the predicted field strength. However one would then have more confidence that the actual field strength value will be equal or greater than the computed value [45].

Typical variability values for analogue broadcasting are 50% time and 50% situation variability. So for 50% of the time at 50% of the situations the actual field strength is expected to be equal or greater than the predicted value. Typical values for DTV is 50% time and 90% situation variability [11]. The average predicted power would be calculated for a time and situation variability of 50%.

Location variability

Location variability accounts for variations in signal strength that would occur if the receiver would be moved around the exact intended location. For example if a receiver is moved around in a 10 meter radius of the exact intended location, the received signal power will vary a little. For most paths the slight variation in location will not make a discernible difference in the path profile or the obstructions that are included in the path profile. The location variability factor can compensate for uncertainty in the exact topographic location of a receiver. With high resolution topographic maps, databases and exact GPS coordinates, location variability is less important since the location of a receiver can be determined with high accuracy [45].

Terrain irregularity parameter

The terrain irregularity, Δh , is determined from the path profile in the point-to-point mode [7]. Figure 2.6 illustrates a terrain profile and indicates how Δh is determined. The path between 10% and 90% of the distance between the transmitter and receiver is considered. From this section of the path the elevation between 10% and 90% of the highest elevation are used to calculate the terrain irregularity parameter.

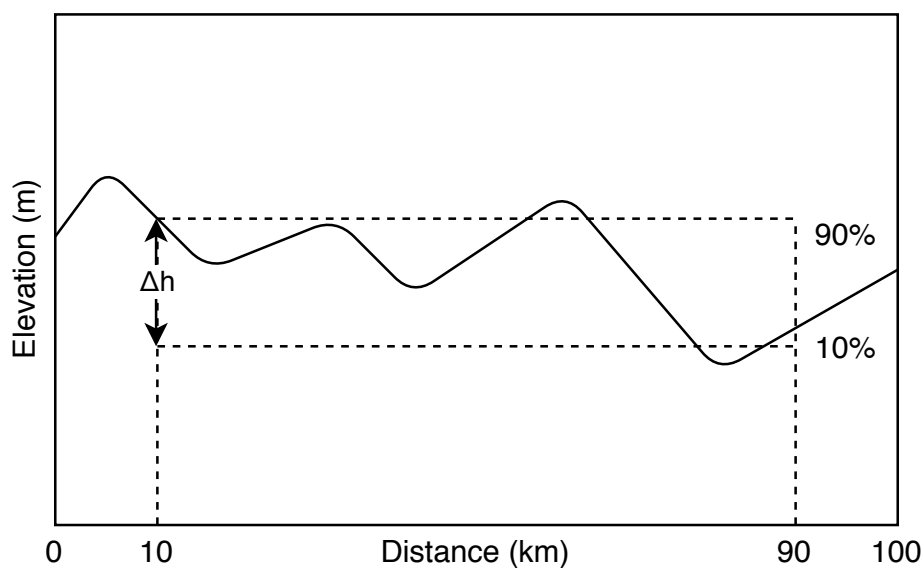


Figure 2.6: Calculation of terrain irregularity parameter, (Δh)

The antenna effective heights (h_{e1}, h_{e2}), the radio horizon distances (d_{L1}, d_{L2}) and the elevation angles (θ_{e1}, θ_{e2}) are also determined from the path profile when using the ITM in point-to-point mode.

The surface refractivity parameter, N_s , is commonly set as 301 N-units. This is commonly used with a radio climate of continental temperate. For short distances (less than 100 km) the model is not sensitive to changes in the value of surface refractivity [8] and there is little difference between Continental and Maritime temperate radio climates [11]. For average atmospheric conditions a Continental temperate radio climate and $N_s = 301$ N-units can be used. Table 2.3 gives radio climates and suggested values for N_s .

Ground constants are used in the calculation of line-of-sight and diffraction attenuation. The ground constant is the Earth's conductivity (σ) and dielectric constant (ϵ). Typical values for the ground constants are given in Table 2.4. The constants for average ground is suitable for most cases [8].

Table 2.3: Radio climates and suggested values for N_s [8,11]

Radio climate	N_s (N-units)
Equatorial (Congo)	360
Continental subtropical (Sudan)	320
Maritime subtropical (West coast of Africa)	370
Desert (Sahara)	280
Continental temperate	301
Maritime temperate, over land (United Kingdom and continental west coasts)	320
Maritime temperate, over sea	350

Table 2.4: Typical values for ground constants [8,11]

Type of ground	Dielectric Constant (ϵ)	Conductivity (σ) - Siemens per meter
Poor ground	4	0.001
Average ground	15	0.005
Good ground	25	0.020
Farmland, forest	15	0.005
Marshy land	12	0.007
Mountain, sand	13	0.002
City	5	0.001
Fresh water	80	0.010
Salt water	80	5.000

The basic transmission loss is determined using equation 2.8 [7]. L_b is the basic transmission loss, L_{bf} is the free-space loss as determined by equation 2.2, and A_{ref} is the median reference attenuation relative to free-space.

$$L_b = L_{bf} + A_{ref} \quad (2.8)$$

The median reference attenuation, A_{ref} , is computed first. A_{ref} is calculated differently depending on the propagation distance. A_{ref} is given by a piecewise, continuous function of distance. Continuity is forced at each endpoint. Figure 2.7 illustrates the median reference attenuation as a function of distance. In the line-of-sight region, the two-ray optics formula is used to calculate the median reference attenuation [7,46]. For the diffraction region, the double knife-edge diffraction is calculated with the Epstein-Peterson model [7]. For greater distances forward scattering becomes the dominant factor affecting path loss. The free space loss is then added to calculate the basic transmission loss. Statistics of the variability of the basic transmission loss are computed and added to the median basic transmission loss, L_b [7].

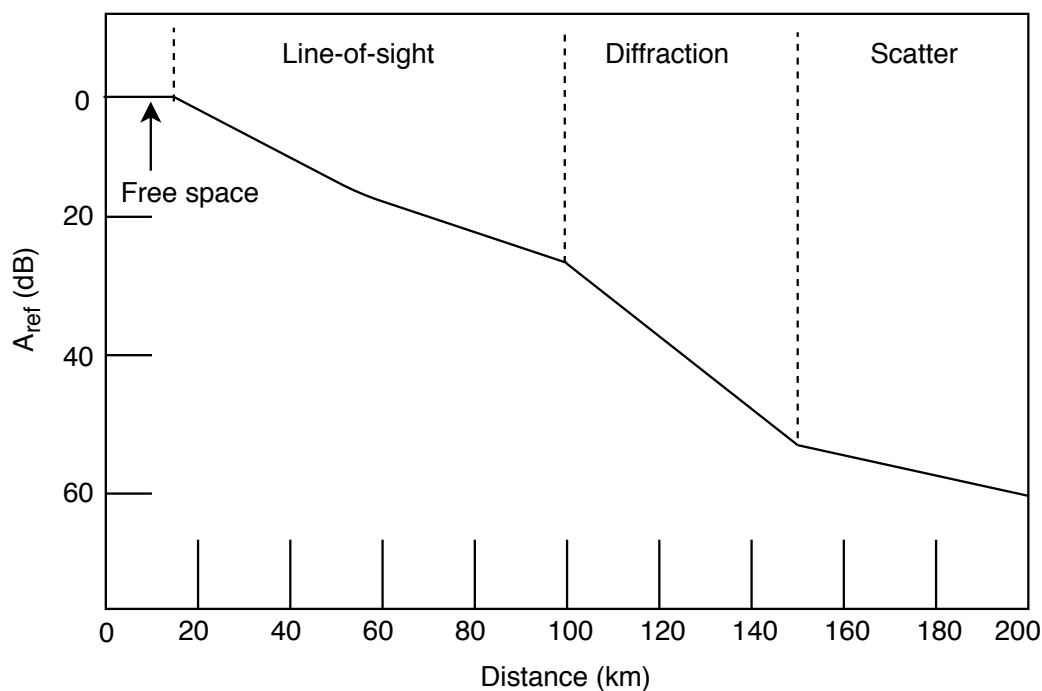


Figure 2.7: Median reference attenuation vs distance plot [7,8]

If the input parameters, or calculated parameters based thereon, are out of range the ITM provides error messages. The software implementations of this method communicates these error messages to the user. Detailed information about these error messages can be found in [8,46].

Irregular Terrain Model with Obstructions (ITWOM)

The ITWOM is an enhanced model of the Longley-Rice ITM [43]. Enhancements have been made by Sid Shumate and the model is also implemented in SPLAT! [47]. Shumate used the ITU-R P.1546 along with Snell's Law and Beer's Law to derive deterministic approximations for the software implementation. The model can be classified as a hybrid model since it uses deterministic approximations but also makes use of empirical data from empirical models.

Shumate identified some shortcomings with the Longley-Rice ITM [43]. According to Shumate the ITM works best with 30-arc-second terrain data and using 3-arc-second or 1-arc-second data do not provide better results. The 30-arc-second terrain data refers to the surface area sample size. It gives an average height for a block of land approximately 1km by 1km. In short, it means that the ITM does not benefit from higher resolution terrain data. The ITM only considers two obstructions in the radial path. It also only directly considers obstructions after the radio horizon and then only consider the highest obstruction seen from the receiver and transmitter respectively. Shumate also identified some mathematical errors and approximations in the ITM implementation. Shumate mostly identified shortcomings in the implementation of the model and not the model itself. He did however add to the model when he implemented the ITWOM.

The corrections made by Shumate are discussed in a series of articles which are published in the IEEE Broadcast Technology Society (BTS) Newsletter [48–59]. Notably he altered the switch point between diffraction and scatter ranges to the ranges proposed by Longley and Rice rather than that implemented in the NTIA ITM software imple-

mentation. The use of effective antenna heights and an average terrain height is also changed to use exact heights from the digital elevation model.

2.3.3 Prediction errors

Research shows that the (RMSE) prediction errors of propagation models compared with measurement campaigns typically range from 8-15 dB [25]. In this study 30 propagation models were described and analysed together with a large set of measurements collected on production networks. The networks used operated in the 2.4 GHz, 5.8 GHz and 900 MHz bands. The authors used various metrics to evaluate the propagation models including the RMSE. It was found that for urban scenarios the RMSE can be expected to be around 8-9 dB, even for fitted models such as a straight line fit of the relationship between path loss and distance. It was also found that for rural environments the RMSE can be expected to be around 15 dB. The authors therefore concluded that for the models evaluated in the study, which includes the ITM, a priori path loss modelling will achieve root mean square errors of between 8-15 dB.

The authors in [60] found that the average prediction error of the ITM can be up to 15 dB depending on the situation. The standard deviation can also be large, in one case it was 16.2 dB. From [61] the average error for the ITM used to predict FM radio signal strength, is found to be as high as 6.1 dB with a standard deviation of 5.9 dB. Even though the average error is found to be relatively low, the absolute errors are still large. This can be deduced from the high standard deviation values. For the ITM average errors of 10 dB can easily be expected.

2.3.4 Propagation prediction software

There are various radio wave propagation prediction tools, both commercially and publicly available. Free tools include SPLAT! for Linux, SPLAT! for Windows and Radio Mobile. Q-Rap is another free, open source radio planning tool.

Q-Rap [62]

The core propagation prediction model used in Q-Rap is based on free-space loss calculations and multiple knife-edge diffraction losses using the Deygout method. Q-Rap makes use of terrain elevation data and allows the use of almost any DEM. Ballot [63] used land cover data to tune the propagation model in Q-Rap in order to improve the prediction accuracy.

Radio Mobile [64]

Radio Mobile is a free RF propagation tool developed by Roger Coudé. The ITM model is used in cases where the path is obstructed and in the case of line-of-sight, 2-ray calculations are used. Land cover data can also be applied to apply a clutter correction factor.

SPLAT! for Windows [47]

SPLAT! for Windows is a free RF propagation tool for performing Longley-Rice path loss using the ITM. It is a Windows version of the Linux-based SPLAT! by John Magliancane.

There is a difference between the predicted field strength values when calculated by SPLAT! for Windows and Radio Mobile for the same exact same case using the same propagation model according to a case study in [60]. The difference is noted, but not explained. It may be because Radio Mobile doesn't use the ITM in case of line-of-sight or perhaps the settings or input parameters of the two programs weren't the same. It could also be that land cover data was used in Radio Mobile and not in the Windows implementation of SPLAT! since it is not an ordinary input to the model. This can also account for the differences in the two implementations.

Since SPLAT! for Windows is a port of the original implementation, SPLAT! for Linux, the tool and its functionality will be described in the next section.

SPLAT! for Linux [65]

SPLAT! (for Linux) is an RF Signal Propagation, Loss, And Terrain analysis tool, hence the acronym SPLAT. SPLAT! is free software and may be redistributed and/or modified under the terms of the GNU General Public License Version 2. It is used for radio and digital TV broadcasting, wireless network design, amateur radio communications and more.

The tool implements the Longley-Rice ITM as well as the enhanced model by Shumate, the ITWOM. It can be used in either area or point-to-point mode. Although the intended frequency range for the ITM is 20 MHz to 40 GHz, the tool allows for a frequency range of 20 MHz to 20 GHz. SPLAT! uses the U.S. Geological Survey and Space Shuttle Radar Topography mission (SRTM) elevation data as a DEM. It can work with SRTM1 and SRTM3 data. SPLAT! also uses coordinates based on World Geodetic System of 1984 (WGS84) or North American Datum of 1983 (NAD83) datum [65].

SPLAT! is command-line driven. In order for SPLAT! to implement the ITM, three input files are required in addition to the elevation data, which for SPLAT! are called SDF files from here on. Two files are site location (referred to as QTH) files and the other is an ITM parameter file, referred to as an LRP file from the extension *.lrp*. A QTH file contains information regarding the receiver or transmitter's name, geographical coordinates, and its height above ground level Above Ground Level (AGL). The LRP file contains input parameters for the ITM. If no LRP file is found, default values are used instead. An optional file to include is an antenna pattern file. This file indicates the gain with regards to specific bearings. Although the gain doesn't affect the path loss calculations it does affect the actual link loss since the gain "reduces" the radiated power in that bearing. Examples of the LRP and QTH file formats are given in chapter 3.

SPLAT! can generate multiple output files. An important output file for the point-to-point mode is the path analysis report given in the form of a plain-text file. This file includes information regarding the calculated attenuation, distance and bearing between the receiver and transmitter, ITM error numbers and obstructions detected. Another useful file is a PNG image file that illustrates the point-to-point path profile

generated by SPLAT!. It also shows the line-of-sight path, the first Fresnel zone and 60% clearance area of the first Fresnel zone.

2.3.5 Digital elevation model

A DEM is a 3-D representation of a terrain's surface created from the terrain's elevation data. Considering the terrain profile on the propagation path can improve the propagation prediction's accuracy since the terrain can diffract, reflect or scatter the electromagnetic waves.

SRTM1 and SRTM3 data are examples of digital elevation models. SRTM1 data has a resolution of 1-arc-second, or about 30m where as SRTM3 data has a resolution of 3-arc-seconds, or about 90m. In short, an arc-second is an angular measure of degrees between two longitudes or latitudes. The distance is the distance on the surface of the earth between two points denoted by coordinates. See section 2.4.1 for more details regarding longitudes, latitudes and arc-seconds. The SRTM3 model is just the SRTM1 model that is upsampled. Nine blocks of the SRTM1 data are averaged to form one block of the SRTM3 DEM [66]. For the SRTM3 digital elevation model, the terrain is represented as a grid of 90m by 90m blocks, with each block representing the average height of the terrain Above Mean Sea Level (AMSL) for that area. The SRTM data can be found at [67].

SRTM data was obtained with radar which uses a wavelength of 5.6cm [68]. The wavelength did not always penetrate vegetation very well. That means that moderate-to-heavy vegetation was mapped near the canopy top as ground level. Where vegetation was sparse, or had no leaves, the wave sometimes penetrated to the ground [69]. Therefore the DEM can sometimes represent canopy heights of forests and not the ground level.

The SRTM data is in a geographic projection, the datum is WGS84 and the units are in metres.

2.4 Field measurements

Field measurements are considered to be something that can be physically measured. Field strength measurements and land cover can physically be measured. Field strength can be measured with an antenna connected to a spectrum analyser and land cover using remote sensing techniques such as radar.

2.4.1 Land cover

Land cover indicates the physical land type, such as forest, water, buildings and so forth, that is present at a location. Land cover data is useful for a variety of applications including environmental planning and protection, development planning, economic planning and enforcement, and strategic decision making [70]. For example, land cover can be used to estimate and monitor the water level of dams during droughts and make strategic decisions based on that. Another creative way that land cover is used is by adding it to propagation models in order to improve the prediction accuracy [63,71].

There exist multiple land cover databases. The databases are for different geographic regions and they can also differ in the classification of the land cover. Some databases have 72 land cover classes [72] while others only have 16 [73]. Since different regions have different physical land cover types, it can be expected that different land cover classes will be present in the various databases.

National Land Cover Database (NLCD)

The 2014 version of the Multi-Resolution Land Characteristics Consortium National Land Cover Database (NLCD2011) is the most recent database containing land cover class information for the entire USA with a spatial resolution of 30 meters [73]. The land cover is classified into 16 classes using a decision-tree classification scheme [73].

The classification scheme is modified from the Anderson Land Cover Classification System. Table 2.5 on the end of this chapter describes the 16 land cover classes.

Table 2.5: NLCD classes [12]

Class number	Name	Description
11	Open water	Areas of open water, generally with less than 25% cover of vegetation or soil.
12	Perennial ice/snow	Areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover.
21	Developed - Open space	Areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20% of the total cover. These areas most commonly include large-lot single-family housing units, parks, golf courses, and vegetation planted in developed settings for recreation, erosion control, or aesthetic purposes.
22	Developed - Low intensity	Areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20% to 49% percent of total cover. These areas most commonly include single-family housing units.
23	Developed - Medium intensity	Areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50% to 79% of the total cover. These areas most commonly include single-family housing units.
24	Developed - High intensity	Highly developed areas where people reside or work in high numbers. Examples include apartment complexes, row houses and commercial/industrial. Impervious surfaces account for 80% to 100% of the total cover.

31	Barren land (Rock/Sand/Clay)	Areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits and other accumulations of earthen material. Generally, vegetation accounts for less than 15% of the total cover.
41	Forest - Deciduous	Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75% of the tree species shed foliage simultaneously in response to seasonal change.
42	Forest - Evergreen	Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. More than 75% of the tree species maintain their leaves all year. Canopy is never without green foliage.
43	Forest - Mixed	Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover. Neither deciduous nor evergreen species are greater than 75% of total tree cover.
52	Shrub/Scrub	Areas dominated by shrubs; less than 5 meters tall with shrub canopy typically greater than 20% of total vegetation. This class includes true shrubs, young trees in an early successional stage or trees stunted from environmental conditions.
71	Grassland/ Herbaceous	Areas dominated by graminoid or herbaceous vegetation, generally greater than 80% of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.

81	Pasture/hay	Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20% of total vegetation.
82	Cultivated crops	Areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20% of total vegetation. This class also includes all land being actively tilled.
90	Wetlands - Woody	Areas where forest or shrubland vegetation accounts for greater than 20% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.
95	Wetlands - Emergent herbaceous	Areas where perennial herbaceous vegetation accounts for greater than 80% of vegetative cover and the soil or substrate is periodically saturated with or covered with water.

The map projection of the dataset used is the Albers Conical Equal Area projection. The horizontal datum used is NAD83 and the ellipsoid used is Geodetic Reference System 80 (GRS80). The planar coordinates are specified in meters [74].

Coordinate reference system

A Geographic Coordinate System (GCS) uses a three-dimensional spherical surface to define locations on the surface of the earth [75]. A GCS includes an angular unit of measure, a prime meridian and a datum which is based on a reference ellipsoid. A

prime meridian is a line of longitude in a GCS at which the longitude is defined as 0° . An example of such a line of longitude is the Greenwich meridian.

A reference ellipsoid defines the size and the shape of the earth and the datum defines the position of the ellipsoid relative to the centre of the earth [76]. The datum therefore provides a frame of reference for measuring locations on the earth by defining the origin and orientation of longitude and latitude lines.

The shape of the surface of the earth is irregular. For most calculations the shape of the earth can be approximated to the reference ellipsoid. The ellipsoid is defined by a semi-major axis (longer radius of an ellipse) and a semi-minor axis (shorter radius of an ellipse).

GRS80 is a geodetic reference system consisting of a reference ellipsoid and datum. The datum used is NAD83. WGS84 is also a geodetic reference system with a reference ellipsoid and datum. The datum is also called WGS84. The GRS80 and WGS84 ellipsoids are almost identical. A GCS is commonly called by the name of the datum since it gives a reference point based on the accompanying ellipsoid. There is almost no difference in the geographic coordinates when the NAD83 and WGS84 datums are respectively used as reference ellipsoids.

A geographic location on the surface of the earth is described in terms of latitude and longitude. Longitude and latitudes values are commonly given in decimal degree or Degrees Minutes Seconds (DMS) format. There are 60 seconds in a minute, 60 minutes in a degree and the degrees range from 90° South to 90° North and from 180° East to 180° West. The latitude is referenced to the equator which is the reference line of latitude. An arc-second is an angular measurement of $1/3600$ of a degree where an arc-minute will be $1/60$ of a degree. At the equator an arc-second, is approximately equal to 30m. The longitude is referenced to the Greenwich meridian for the case of the WGS84 GCS.

Figure 2.8 shows how the geographic coordinate is determined. The longitude, denoted as ϕ for this example, is given by the angle between the reference meridian and

the meridian passing through the point on the surface of the earth. The latitude, denoted as ω for this example, is given by the angle between the reference line of longitude, the equator, and the parallel or line of longitude that passes through the point on the surface of the earth.

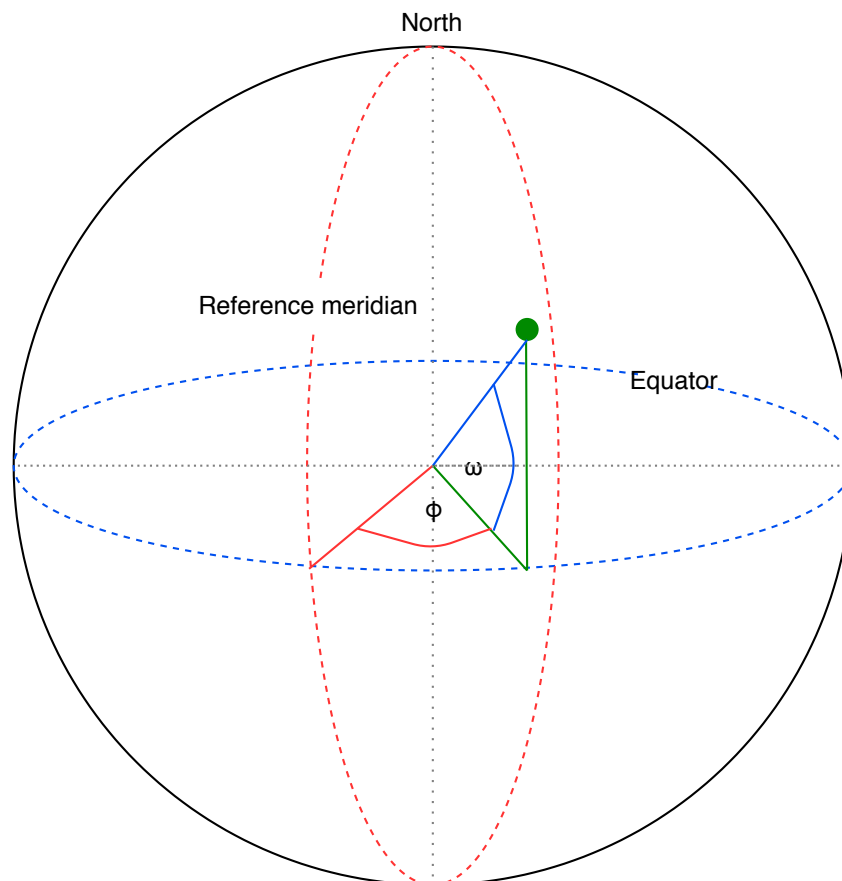


Figure 2.8: Geographic coordinates

Projection

A projected coordinate system is defined on a flat, two-dimensional surface. It is always based on a GCS that is based on a reference ellipsoid. When mapping a sphere or ellipsoid to a 2-dimensional surface, some features are distorted. Different projections respectively preserve area, distances, directions and shapes. While one feature is preserved, some of the others are distorted during projection.

2.4.2 Field strength measurements

RF field strength measurements have numerous applications. It is used to monitor the radio spectrum and enforce compliance. It is also used to verify the coverage area of a broadcast transmitter by means of a drive test. It is used to determine spectrum occupancy. It is also used to verify propagation models. Empirical models are based on field strength measurements.

Field strength is commonly measured with a spectrum analyser or a Software Defined Radio (SDR). Universal Software Radio Peripheral (USRP) is a range of software defined radios which are commonly used by research labs, universities and hobbyist because it is a comparatively inexpensive hardware platform for software radios. Spectrum analysers are generally more expensive but has superior performance when compared to inexpensive SDRs.

An antenna is connected to the spectrum analyser or SDR to intercept the electromagnetic wave as it propagates through space.

Antenna types

Antennas can be categorised in three broad categories namely omni-directional, semi-directional and directional. The antennas are categorised depending on the extent to which the radiated power is concentrated in a direction.

Omni-directional antenna

Omni-directional antennas radiate radio wave power uniformly in all directions in a single plane, usually the horizontal plane [77].

Dipole antennas are the most commonly used omni-directional antenna. These antennas are constructed to resonate at half or quarter wavelength of the applied frequency. The half wave dipole is the most widely used of the aforementioned.

The radiation pattern will look like a doughnut if viewed in three dimensions. The gain of an ideal half wave dipole relative to an isotropic radiator is 2.15 dBi as discussed in section 2.1.4.

Directional & semi-directional antenna

Directional and semi-directional antennas focus radiated power in a direction, adding significant gain in that direction [77]. Therefore the radiated power will be less in some directions and more in others.

Examples of commonly used directional antennas are the Yagi and parabolic or dish antennas. Both of these antennas focus the radiated power into a narrow beam for a significant gain in a specific direction. The radiation patterns vary but a common and important factor is the beam width.

Patch and sector antennas are examples of semi-directional antennas. Sector antennas provide a pie shape radiation pattern in the horizontal plane. That is when the radiation pattern is viewed from above it looks like a slice of pie. These antennas are usually mounted in a sectorized array to provide omni-directional coverage and are extensively used in cellular networks.

2.4.3 Power Spectral Density (PSD)

The PSD is a measure of the amount of power present in a signal as a function of frequency. It is commonly expressed in dBm/Hz. The PSD can be computed by taking the Fast Fourier Transform (FFT) of the measured time signal. The PSD can then be expressed in a power scale per FFT bin width. The amount of power over a small frequency range, a bin, is determined and that value is used as a point on the PSD plot.

The PSD is integrated over a specified frequency range to determine the average power of a signal for a given bandwidth.

2.4.4 Considerations for measurement campaign

A spectrum observatory is a dedicated station that continuously takes field strength measurements and stores the data. The data can be then analysed for various applications like spectrum monitoring, TVWS analysis and more.

A spectrum observatory consists out of a measurement device like a spectrum analyser or USRP which is connected via feeder cables to an antenna. The USRP or spectrum analyser is then also connected to a computer where the measurement data can be processed and stored.

Things to consider when measuring RF signals are the frequency range, instantaneous and resolution bandwidths, sensitivity and selectivity of the receiver, noise and other losses.

The frequency range is dependant on the RF instrumentation and the antenna. Using an omni-directional antenna is useful for a spectrum observatory since the RF signal can come from any direction.

The instantaneous bandwidth is the maximum bandwidth that the instrument can acquire without re-tuning the local oscillator [78]. The resolution bandwidth is the minimum bandwidth over which two signals can be separated and distinguished. The resolution bandwidth has an effect on the noise floor.

The noise floor is the measure of all the unwanted signals in a system [6]. This typically includes thermal noise and interference from unwanted signals. Thermal noise determines the minimum noise floor of the receiver. At room temperature the minimum thermal noise power in a bandwidth of one Hertz is -174 dBm. The noise floor determines the smallest signal that can be measured with the instrument.

2.5 Model tuning and evaluation

In the previous sections propagation models and field measurements were discussed. First propagation models are used to predict what the received power will be at a specific geographic location given the transmitter's and receiver's parameters. Then field strength measurements are taken to establish what the true received power is at the receiver's geographical location. This is usually done for many transmitter-receiver pairs. The propagation model can then be evaluated to determine how well and accurate the received power was predicted. The model can even be augmented and tuned using the true received power measurements.

The following sections will discuss a method that can be used to fit the model with field measurements, both field strength and other measurements, in order to reduce the error between the predicted and true measured received power. In other words, to improve the accuracy of the propagation model. Different metrics that can be used to evaluate a propagation model are also discussed.

2.5.1 Gradient descent method

The gradient descent method is a first-order iterative optimisation algorithm for finding the minimum of a function [63,79]. To find a local minimum, one moves proportionally in the negative direction of the gradient of the function. The approximate gradient of the multi-variable function is the vector of partial derivatives of the function, hence it being a first-order algorithm. The function $f(\mathbf{a})$ must therefore be differentiable in the neighbourhood of $\mathbf{a}$.

Consider a multi-variable function $f(\mathbf{x})$, with $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_n]^T$. Let $f(\mathbf{x})$ be differentiable in the neighbourhood of $\mathbf{x}$. The gradient vector, $\nabla F(\mathbf{x})$ will be given as

$$\nabla F(\mathbf{x}) = \left[\frac{\partial f}{\partial x_1} \ \frac{\partial f}{\partial x_2} \ \dots \ \frac{\partial f}{\partial x_n} \right]^T \quad (2.9)$$

If $\mathbf{x}_{n+1} = \mathbf{x}_n - \gamma \nabla F(\mathbf{x}_n)$ for γ , the step size, small enough, then $f(\mathbf{x}_{n+1}) \leq f(\mathbf{x}_n)$. Thus by moving against the gradient, the output of the function is reduced.

The value of the step size is allowed to change every iteration. It has an influence on the convergence rate of the algorithm. Determining the optimal step size can be time-consuming and using a fixed small step size can lead to poor convergence. Depending on the function the step size can also determine whether the algorithm converges or not.

There can be multiple stopping conditions for the algorithm. One can choose to stop it after a predetermined number of iterations. The algorithm can also be stopped when the error is small enough. The algorithm can also be stopped if the norm of the gradient vector is small enough. It is the same as stopping when the change in variables is small enough. In short the algorithm stops when a local minimum is reached or when the algorithm doesn't make meaningful changes towards the actual minimum of the function.

When the function is convex the local minimum will also be the global minimum. Otherwise it cannot be guaranteed that a global minimum will be found. The initial values of the variables will determine which minimum could be reached. Figure 2.9 illustrates how different initial values can lead to different local minimums being found. The blue and red dots indicate the initial values for $\mathbf{x}$.

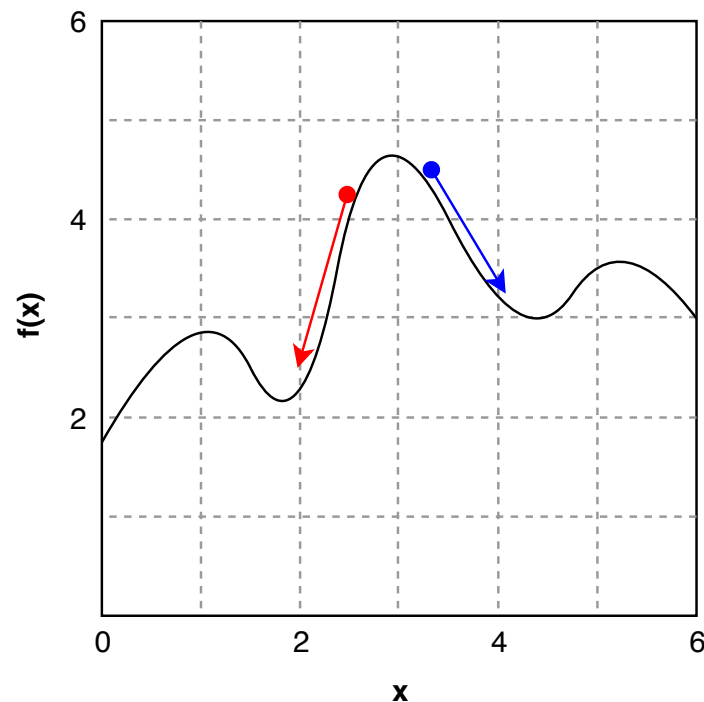


Figure 2.9: Effect of initial parameters in finding local minima

There are techniques to get out of a minimum and continuing the search of the function space. One such technique is called gradient ascent. It does the exact opposite of gradient descent and moves with the gradient to find a maxima.

The gradient descent method is considered slow when compared to other algorithms, meaning it takes longer to converge to a local minimum. It does however work for a function space with any number of dimensions, or a function with any number of variables.

2.5.2 Statistical evaluation metrics

A model is a representation and/or approximation of the real thing. There are thus differences between the model and the thing or aspect being modelled. The difference is referred to as the error.

From here on the error term will be defined as the value of the measured/real data minus the value of the predicted/modelled data as shown in equation 2.10.

$$\epsilon = x_{real} - x_{model} \quad (2.10)$$

The metrics discussed next can be used to evaluate the performance of large scale propagation models.

Mean

The arithmetic mean is calculated by dividing the sum of all the values by the number of values [13] as in equation 2.11. It reports on the central tendency of data. It is sensitive to outliers. In the case where a lot of outliers are present the median can be a better indication of the central tendency of the data. The mean error is just the average of the error between the model and the real measurement.

$$\mu = \frac{1}{n} \sum_{i=0}^n x_i \quad (2.11)$$

Standard deviation

The standard deviation gives an indication of the variation in the data around the mean [13]. A low standard deviation indicates that the data points tend to be close to the mean. The standard deviation is given in equation 2.12.

$$\sigma = \sqrt{\frac{\sum_{i=0}^n (x_i - \mu)^2}{n}} \quad (2.12)$$

Mean Absolute Error (MAE)

The mean absolute error is the sum of the absolute of the error divided by the number of samples. The mean absolute error gives a better indication of the magnitude of errors than the mean error does, because negative and positive errors can't cancel out. The mean absolute error is given in equation 2.13

$$MAE = \frac{1}{n} \sum_{i=0}^n |x_m - x_p| \quad (2.13)$$

, where x_m is the measured sample and x_p is the predicted sample.

Root Mean Square Error (RMSE)

To calculate the RMSE, the errors are first squared and then summed. The average is then taken and then the square root. The RMSE gives an indication of the average model prediction error. The RMSE is always positive and a lower value is better. The RMSE is given in equation 2.14

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=0}^n (x_m - x_p)^2} \quad (2.14)$$

, where x_m is the measured sample and x_p is the predicted sample.

Squaring the errors before taking the average gives a larger weight to large errors. The RMSE should therefore be more useful when large errors are undesirable.

Correlation

Correlation between sets of data is a measure of how well the sets are related. Pearson correlation is most commonly used. It shows the linear relationship between two sets of data. The Pearson correlation coefficient is given in equation 2.15

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}} \quad (2.15)$$

, where n is the sample size, x_i, y_i are the individual sample points indexed with i and $\bar{x}, \bar{y}$ are the sample means. [13].

The size of the correlation coefficient, r , indicates the strength of the correlation. The correlation coefficient can be interpreted as shown in Table 2.6. The correlation coefficient is always between -1 and 1. For example if $r = 0.9$, one can say that there exists a strong positive correlation.

Table 2.6: Correlation coefficient interpretation [13]

Interval	Correlation
$0.0 = r $	None
$0.0 < r < 0.2$	Weak
$0.5 \leq r < 0.8$	Moderate
$0.8 \leq r < 1.0$	Strong
$1.0 = r $	Perfect

2.6 Related work on augmenting large scale propagation models

In this section related working on augmenting large scale propagation models will be discussed. Some studies used only field strength measurements while others used a combination of field strength measurements and land cover data in order to tune large scale propagation models to reduce the error between predicted and measured field strengths.

In previous work done on tuning the Longley-Rice propagation model using measure-

ments [80] it was shown that this model usually underestimates path loss, thus overestimating the coverage area of a transmitter. The result is that an opportunity is missed for spectrum sharing at a specific location. In this study, the Longley-Rice model was tuned via the confidence level parameter, which is, according to the author, the only parameter that has a non-negligible effect. Field strength measurements were used to tune this parameter, and it was shown that only a few samples (5 to 10) of measurements were necessary. These samples were taken on a line starting at the transmitter and moving further away from where at any point a receiver can be placed. The result is, however, path specific and may not be valid for the whole coverage area due to terrain that may vary on other paths. Another path may result in another value for the confidence level parameter. The study also showed that cheaper self-made measuring equipment could be used in the measurement campaign but it is not so successful when the noise floor of the device is as low as the signal being measured. The advantage of more expensive measurement instruments is that the noise floor of the device is usually lower, and the resolution better, yielding more accurate measurements.

In a study on integrating a sensing and database architecture for white space networking [81] it was shown that measurements perform much better than predictions. It also showed that using kriging can reduce the number of samples needed for fitting a model, but the method still required many samples for a good fit. Ultimately the findings of the study was that measurements improved the coverage prediction in such away that a lot more opportunity is created for secondary spectrum use, which improves efficiency of spectrum use, but it also led to more false positives where white spaces were identified that is not white spaces, and allowing a secondary user to transmit in those regions can cause unwanted interference. There is thus a trade-off between more efficient use of the radio spectrum and unwanted harmful interference.

In another study on improving the accuracy for TVWS geo-location databases from measurement driven estimation approaches [82], the ITM was marginally improved by subtracting the mean of the measured path loss from the predicted path loss. The improvement accounted for fixed power offsets due to antenna gains and cable losses which weren't taken into account by the model. The study showed that drive testing,

which is expensive and time consuming, can be minimised if sophisticated interpolation techniques are applied.

The author of [63] used field strength measurements and land cover data to tune a propagation model based on free space loss and diffraction using the Deygout method. The model was tuned by adding an effective land cover height to the DEM and maximising the correlation coefficient between measurements and predictions. The study made use of a lot of field strength measurements from various transmitters at various receiver locations. The correlation coefficient was maximised by means of a gradient descent method. The model was tuned for a wide range of frequencies and it was found that adding an effective clutter height according to land cover, affected the higher frequencies the most. The result of tuning the model with these measurements and land cover data is a better correlation between measurements and predictions together with a lower standard deviation. The mean was not used as a metric of the tuned model's performance since it is an arbitrary task to make the mean zero by adding an offset.

In another study, the authors of [71] used field strength measurements and land cover data to tune a deterministic model, called Fresnel/Deygout, with reflections from the ground taken into account, and taking into account all obstacles between transmitter and receiver sites. The model is tuned by adding an effective clutter height for a forest land cover type. The height was increased and decreased as necessary in order to reduce the error between field strength measurements and predictions. Additionally, they tuned another term, which is an additional loss per distance unit, and added that to the loss predicted by the propagation model.

In another study, the authors of [83] improved on the Okumura-Hata model in the 800MHz band using a Newton second order method to tune the coefficients of the model. The RMSE was used as a metric to optimise and evaluate the performance of the model.

2.6.1 Related work discussion

Thus from the related work discussed it is clear that large scale radio propagation models can be augmented with field strength measurements and other field measurements for example land cover information. Field strength measurements are critical in model tuning since the true value is required to know the prediction error. Although measurements are very important, predictions are still an important aspect when determining the coverage area of a transmitter. Propagation predictions are far less time consuming and less expensive than drive tests to take field strength measurements. Different paths should be considered when augmenting a propagation model with field measurements otherwise the augmented model may produce large errors on other paths. Depending on the transmitter, a few dB difference in error can amount to an error of tens of kilometres for the predicted coverage area boundary. It means that there are either more opportunities to reuse the radio spectrum that are being missed or there is a lot of unwanted harmful interference.

Also seen from the related work, there are three important aspects when augmenting large scale propagation models with field measurements. First is to obtain field strength measurements in order to determine what the ground truth is in terms of the received signal power. Secondly, the additional field measurements that are used to augment the model with, should be taken and processed. Lastly the error between the measured field strength and predicted field strength should be reduced.

It is important to continue research in this area in order to find the balance in taking field measurements for augmenting large scale propagation models so that the reduction in error or increased accuracy of the propagation model is worth the cost of taking measurements.

2.7 Conclusion

In this chapter, background was given on radio wave propagation and large scale propagation models were discussed. Attention was given to the Longley-Rice model. Field measurements were discussed, followed by tuning and evaluation of large scale propagation models. Lastly, related work on tuning large scale propagation models with field measurements, was discussed.

In the next chapter, the research model and implementation will be discussed. It will consist out of a measurement study, processing land cover data and tuning and evaluating a large scale propagation model.

Chapter 3

Research model and implementation

In this chapter, the research model and implementation will be discussed. First, the measurement study and processing of the measurements will be discussed, followed the processing of the land cover data. Finally the method of reducing the error between field strength measurements and propagation predictions will be discussed.

3.1 Introduction

As seen from the related work in the previous chapter, there are three important aspects when augmenting large scale propagation models with field measurements. First is to obtain field strength measurements in order to determine what the ground truth is in terms of the received signal power. Secondly, the additional field measurements should be taken and processed. In this case it is the land cover data. The height of the DEM, which is an input to the ITM, will be adjusted when the model is fitted to the measurements. Lastly the error between the measured field strength and predicted field strength should be reduced. For this study the gradient descent method was chosen to tune the ITM and ITWOM with land cover data by minimising or reducing

the RMSE between the measured and predicted field strength. As discussed in section 2.5.1, the gradient descent method is suitable for a function space with any number of dimensions, or a function with any number of variables. Since there may be many land cover classes to consider, the dimensionality of the search space may become large. As discussed in section 2.5.2, the RMSE penalises larger errors more. Therefore the RMSE between predictions and measurements is chosen to be minimised using the gradient descent method.

Figure 3.1 is a high-level overview of the methodology implemented in this chapter. The field strength measurements from the measurement study are initially processed in MATLAB to remove outliers and calculate an average received power for every transmitting frequency. The land cover data is also preprocessed in MATLAB so that the resolution and geographical locations of the land cover are compatible with the DEM data used in the propagation tool. This way the DEM can easily be modified with the land cover data. The DEM is an input to the ITM or another propagation model that makes use of it. The output of the propagation model is an input to the gradient descent method. This method is used to make changes to the DEM by means of changing the heights for the different land cover classes and then adding it to the original DEM. The changes are made so that the error between the measurements and predicted received powers, or output of the ITM, is reduced. The process is repeated until the error cannot be reduced further by the gradient descent method. The output is an augmented DEM or height associated with each land cover class that can be added to a DEM. Using the augmented DEM as input to the propagation model should then result in a smaller error between measurements and predictions.

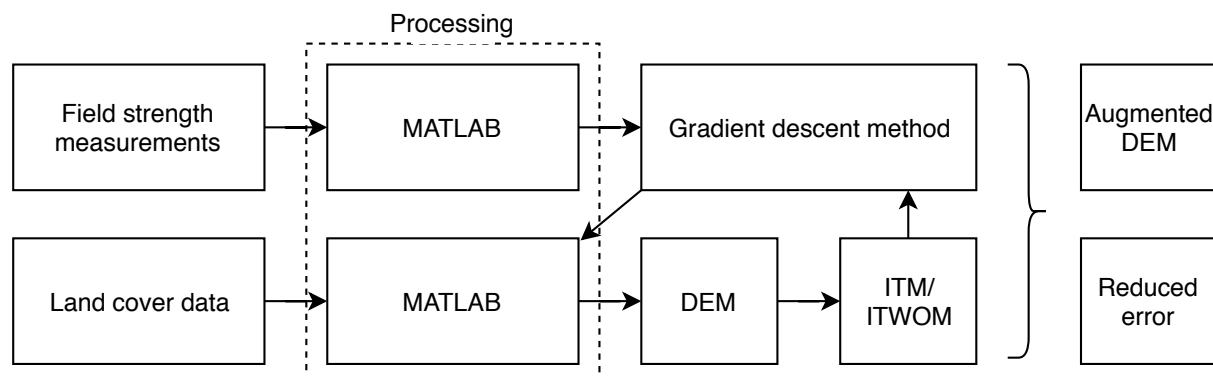


Figure 3.1: High-level overview of the methodology

3.2 Measurement study

Field strength or received power measurements are needed to establish the ground truth. This provides a baseline error for the difference between the predicted and measured received power. It is also used in the gradient descent method to drive down the error so that the predicted power is closer to the measured power i.e. reducing the error.

Field strength measurements are typically obtained via a drive test. In such a setup there are one or more transmitters and one mobile receiver. The receiver follows a pre-determined route and takes measurements along the way. The receiver also captures a time stamped GPS location for every measurement. The big advantage of a drive test is that a large number of measurements can be taken. The disadvantages are that it can be expensive and time consuming. Measurements are also taken for a short period and the variation in the signal strength over time may not be captured.

Another option is to use one stationary receiver that measures the field strength from multiple stationary transmitters like the case for a spectrum observatory. As mentioned in chapter 2, a spectrum observatory is a dedicated station that continuously takes field strength measurements and stores the data. Multiple stationary receivers at different locations can also be used to increase the number of unique paths from a

transmitter to a receiver and also increase the number of field strength measurements. The advantages of such a setup are that it is less expensive than a drive test and measurements can be taken continuously for longer periods revealing the variation in the signal strength.

For this study a single stationary receiver is used that continuously measured the field strength from thirty transmitters at various distances, heights and bearings. Figure 3.2 gives an example of such a setup. The stationary receiver is on a roof and the stationary transmitters are various distances and bearings from the receiver. Some are on roofs of buildings while others are on a hill or mountain. The paths differ in terms of the terrain elevation profile and the type of land cover thereon.

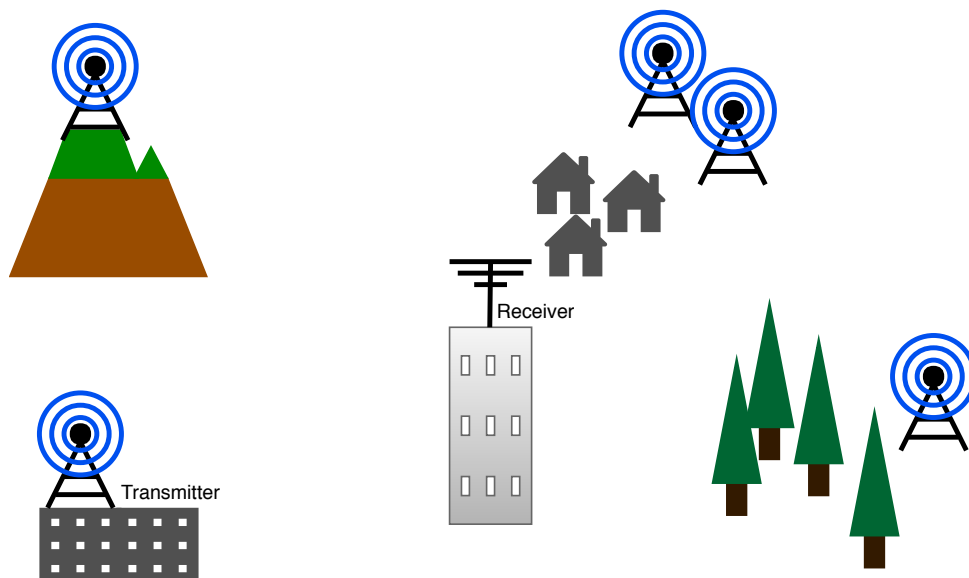


Figure 3.2: Measurement study setup

3.2.1 Receiver selection and setup

The receiver chosen for the study is the spectrum observatory station at the University of Washington (UW) campus, Seattle. It is part of the CityScape observatories. CityScape is a collaborative effort between the University of Washington, Shared Spectrum Company and Daintree Technologies to create a metro-scale spectrum observatory [9]. The period for which it was in operation is from 4 April 2017 until the end of

2017, when looking at the times of available data and the latest data uploads. The Microsoft Station in Redmond, which is nearby the UW station, is also part of CityScape metro-scale observatory and is used for validation in section 5.2.2. Historic raw data is available via the CityScape observatory website. The stations are offline and new data is not added to the website. The latest PSD data for the Microsoft station is 13 November 2017 and 5 January 2018 for the UW station. However direct access was granted to the UW station and the newly obtained data could be stored locally for further processing. The scanner parameters could also be configured as needed. Table 3.1 shows the UW receiver parameters and Table 3.2 shows the Microsoft receiver parameters.

Table 3.1: UW receiver station parameters

Parameter	Value
Latitude (N)	47.654839°
Longitude (W)	122.306728°
Height (AGL)	15m

Table 3.2: Microsoft receiver station parameters

Parameter	Value
Latitude (N)	47.640503°
Longitude (W)	122.141472°
Height (AGL)	6m

Hardware setup

The UW spectrum observatory consists out of wideband antenna (MultiPolarised Super-M Ultra Base Station (25MHz to 6GHz)) mounted on a roof connected to an SDR sensor (Ettus Research USRP N200 + UBX40 daughterboard combination) that is attached to a PC. The PC runs SDR management software and is remotely accessible as a web client

that can be used to configure the scanner process parameters for data acquisition [9]. The hardware and software subsystems are setup and configured as stated in [9]. The station setup for this study differs slightly from the original setup as given in [9]. The data is stored locally and not uploaded on the public Azure cloud since the subscription ended. The antenna is no longer connected to a Mini Circuits Splitter ZX10R-14-S before connecting to the SDR but is directly connected to the SDR with a coax cable connector. The cable connecting the antenna to the SDR is still a 10m LMR-400 cable. Figure 3.3 is adapted from [9] and shows the CityScape sensor hardware layout.

The USRP is calibrated in two steps. The first step is to calibrate the USRP with Ettus calibration tools in order to suppress the I-Q imbalance of the sensor. The second step is to calibrate the USRP using the CityScape tools in order to obtain measurements from the collected I-Q data in dBm. This is required since the I-Q data from the USRP is on an arbitrary scale. The second step of calibration is well documented in [9]. Importantly, the hardware is calibrated and the resulting periodogram from the station can be trusted. It should be noted that the cable loss is not taken into account in the calibration process and should be added to the resulting periodogram.

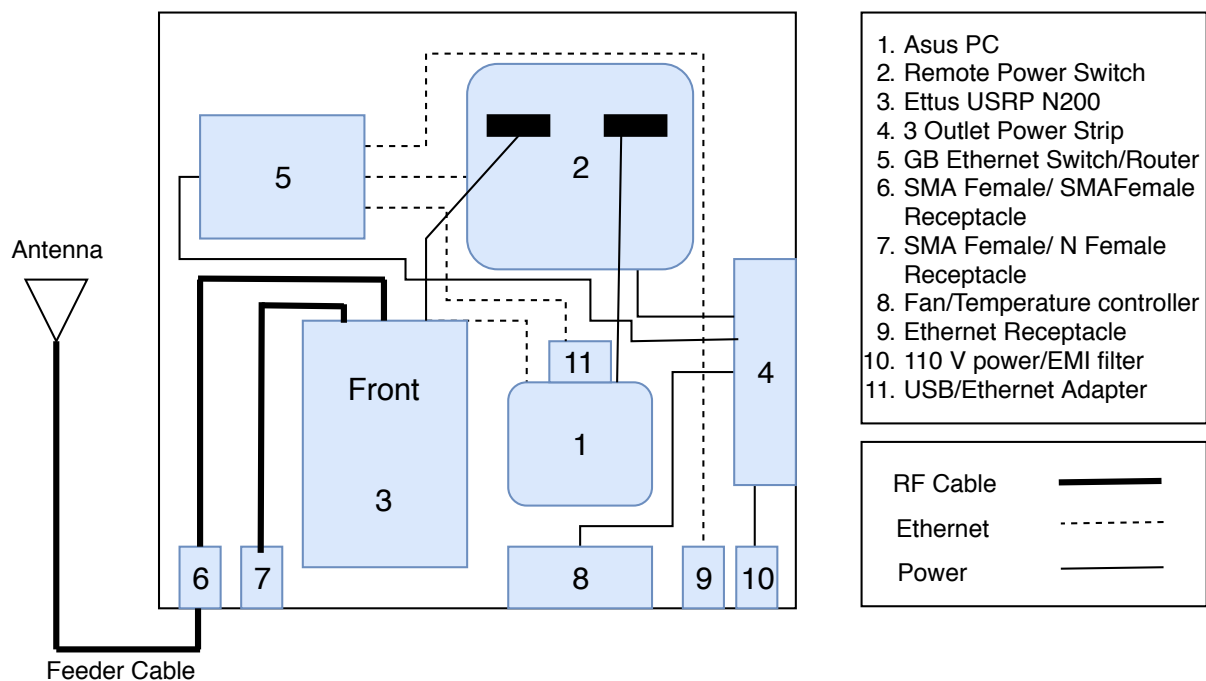


Figure 3.3: CityScape sensor layout adapted from [9]

Software setup

The spectrum observatory has two data acquisition modes namely raw I-Q and time averaged spectrogram (PSD). For the PSD data the device is programmed to compute average, minimum and maximum Power Spectral Density (PSD) values over the specified frequency range time-averaged over one-minute intervals. The operation is repeated for a 60-minute duration and a per-minute time-averaged PSD is stored in a file for upload [9]. The device would therefore capture the PSD data for one second for a specific frequency range. It will move on to the next frequency range and capture the PSD data for one second. It will continue this until the entire frequency range of importance is scanned and then start at the beginning. In the course of one minute it would have scanned the entire frequency band multiple times but note that every frequency is not taken every second. After one minute all the PSD data for the entire spectrum of importance is averaged and the results saved. This data is collected for a period of 60 minutes and the results are saved every minute. At the end of the hour there are 60 averaged PSD data for every frequency saved. This data is then written to a file and uploaded for storage.

The PSD data is of importance for this study since the received power can be calculated from it. Table 1 in [9] shows the major software components for management and control of the spectrum observatory. These were utilised and adjusted as necessary for this study. Table 3.3 shows the available settings and their values for the UW station as used in this measurement study.

Table 3.3: Tunable parameters of the UW spectrum observatory

Setting	Value	Description
Start/Stop Frequency (MHz)	75/700 MHz	Frequency, inclusive at which device should start/stop collecting data.
Effective sampling rate	25 MHz	Determines the effective sampling rate (rate after decimation) of the scanner.
Rx Gain	0 dB	Adjusts the receiver gain (dB) of the USRP.
Scan Pattern	DCSpikeAdaptiveScan	DCSpikeAdaptiveScan tries to remove DC spikes observed from the PSD data obtained and calculated from homodyne receivers (like USRP) and is used for PSD measurements.
Number of snapshots per channel visit	1	Indicates the number of snapshots taken per channel visit.
Samples per snapshot	1024	Indicates the number of samples to capture for each snapshot.
Cycle time	1000 ms	Defines the length of the cycle duration.
On time	100 ms	Defines how long Raw I-Q data should be captured within the cycle time.
PLL flag poll delay	0 ms	Period of time between each PLL lock flag polling which occurs after each frequency retune.
Additional tune delay	30 ms	Duration after each frequency retune in addition to the PLL flag poll delay.
Seconds of data per sample	60 s	Period of time-averaging interval for PSD report. A maximum, minimum and average PSD will be calculated for this time period.
Minutes of data per upload	60 min	Interval (in minutes) for upload of PSD ScanFile (.dsox) to the local server.

Figure 3.4 shows that the PSD data is captured in the off-time during a cycle while the raw I-Q data is captured during the on-time. For this study the cycle time is 1000 ms, the on time is 100 ms, leaving the off time, for PSD data capture, to be 900 ms. The cycle is repeated for one minute and a time averaged PSD is calculated for the minimum, average and maximum. The cycle also includes time to tune to a new centre frequency and any additional tune delay that allows the local oscillator to settle before acquisition begins. After an hour of continuous scanning a file containing time averaged PSD data (minimum, average and maximum) for each minute of the hour is created and stored locally for further processing.

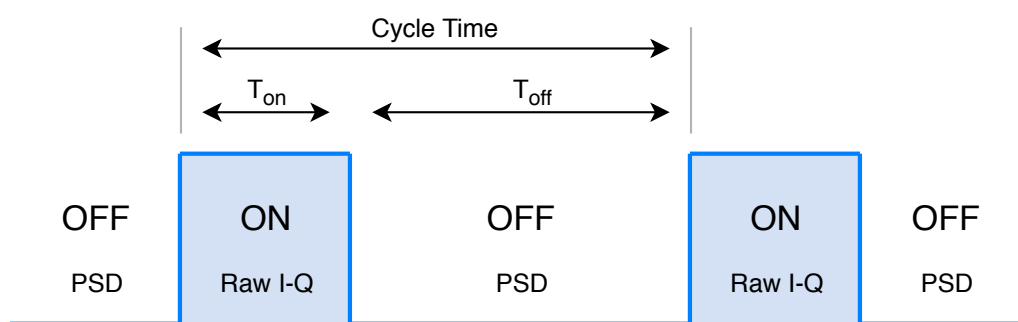


Figure 3.4: Data acquisition cycle adapted from [9]

The USRP is not capable of taking an absolute signal power measurement in dBm, and all it can provide the host PC with is ADC readings, which is amplitude in an arbitrary scale [9]. According to [9] and [84] the station is amplitude calibrated thus PSD estimate data are in dBm/FFT bin size, an absolute scale. The measured data (at the ADC of the USRP) excludes cable and connector losses. The power levels shown by the PSD plots are lower than the true received power at that location. To compare it with the propagation predictions the cable and connector losses should be considered and the value of these losses should be added to power levels of the PSD. This is illustrated in Figure 3.5. The cable losses are around 1.3dB/100 feet [85] and with the cable length being 40 ft the additional power to be added is 0.4 dB. Connector loss is insignificant as high quality interconnects were used.

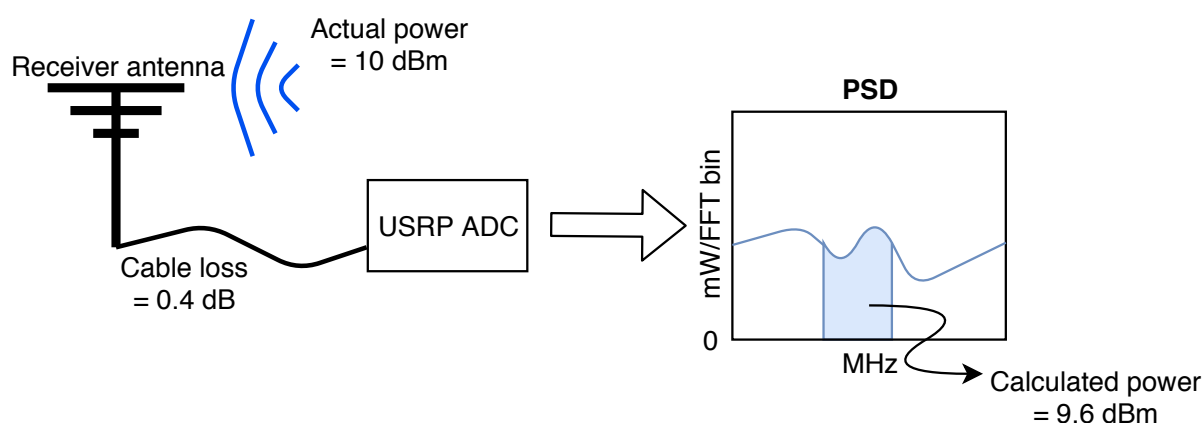


Figure 3.5: Effect of cable loss on calculated average power

3.2.2 Transmitter selection

Transmitters were selected based on the availability of transmitter information and whether the signal is measurable at the receiver location. The FCC provides publicly accessible information on licensed transmitters in the radio (88 MHz – 108 MHz) and television (174 MHz – 698 MHz) bands that are used for broadcast services. Table 3.4 shows transmitter information that could be obtained from the FCC database. The FM query search function on the FCC database [86] is used to search for transmitters. From the information from the FCC database, available transmitters in the area could be identified. The final set of transmitters to use then depended on whether the signal from a transmitter could be identified with a degree of certainty on the PSD plot by searching for the center frequency.

Table 3.4: Transmitter information of interest obtainable from the FCC database

Parameter	Description
Longitude and latitude	Geographical location of the transmitter (NAD83).
Antenna height above ground level (AGL)	Height of transmitter antenna above ground level.
Effective radiated power (ERP)	Power that the transmitter radiates relative to a dipole (dBd).
Frequency/Channel number	Centre frequency at which transmitter radiates power in MHz.
Antenna pattern	Describes the gain of the antenna in all directions in the radiation plane.
Polarisations	Orientation of the electric field (horizontal or vertical).

Considering the location of the UW spectrum observatory and the distance that radio signals in the VHF range can propagate, an initial search for all the licensed radio transmitters in Washington State was done. The initial search yielded 219 records with transmitters broadcasting at a frequency from 88.1 MHz to 107.9 MHz. In order to sort the applicable transmitters from the results of the initial search a quick measurement (1 hour) was taken. From the generated PSD plots, the frequencies at which signals were measurable became clear. The identified frequencies were used to narrow the search result. There were still multiple possible transmitters for every determined frequency. The search results included the city in which the transmitter is located. That together with the ERP usually indicated which transmitter's signal was received. In the cases where it was unclear each transmitter information is further examined. When further examining a transmitter the FCC website provides an option to plot the transmitter's coverage area. This additional information yielded the correct transmitter. The cov-

erage area plot was also used for all the other transmitters to verify that the correct transmitter was indeed chosen. Table 3.5 shows the list of 30 transmitters identified for this study. Note that all the transmitters broadcast at unique frequencies. The distance is the distance between the transmitter and receiver and the azimuth is the angle from the receiver to the transmitter. Due North would be 0° and due East would be 90° .

Table 3.5: FM transmitter set used for study

Call sign	Latitude (°N)	Longitude (°W)	Antenna height	Frequency	Directional			
			AGL (m)	(MHz)	ERP (kW)	antenna	Distance (km)	Azimuth (°)
KNKX	47.503611	121.975833	41	88.5	64	Yes	29.97	123.98
KNHC	47.542889	122.108167	64	89.5	8.5	Yes	19.4	129.8
KEXP-FM	47.615833	122.308889	122	90.3	4.7	Yes	4.33	182.2
KBCS	47.543472	122.108694	79	91.3	1.8	Yes	19.32	129.73
KYFQ	47.547778	122.807500	175	91.7	4.3	Yes	39.39	252.61
KQMV	47.504639	121.968833	66.7	92.5	56.8	No	30.34	123.24
KUBE	47.544167	122.108333	70	93.3	98	No	19.3	129.5
KLSY	47.312667	123.372083	89	93.7	28	Yes	88.64	244.98
KSWD	47.504639	121.968833	65	94.1	69	No	30.34	123.24
KUOW-FM	47.615833	122.308889	137	94.9	69	No	4.33	182.2
KJR-FM	47.544167	122.108333	70	95.7	98	No	19.3	129.5
KXXO	46.972889	123.139472	111	96.1	37	Yes	98.44	219.93
KJAQ	47.504639	121.968833	65	96.5	49	No	30.34	123.24
KIRO-FM	47.503611	121.975833	57	97.3	52	No	29.97	123.98
KOMO-FM	47.312667	123.372083	72	97.7	69	Yes	88.64	244.98
KING-FM	47.503611	121.975833	41	98.1	66	Yes	29.97	123.98
KNBQ	46.975167	123.139167	78	98.5	1.6	Yes	98.23	220.02
KNUC	47.504639	121.968833	66.7	98.9	63.9	No	30.34	123.24
KDDS-FM	47.312667	123.372083	113	99.3	64	Yes	88.64	244.98
KISW	47.503611	121.975833	41	99.9	67	Yes	29.97	123.98
KKWF	47.503611	121.975833	41	100.7	67	Yes	29.97	123.98
KPLZ-FM	47.544361	122.108889	57.2	101.5	99	No	19.25	129.53
KZOK-FM	47.504639	121.968833	65	102.5	68	No	30.34	123.24
KFNY	46.975167	123.139167	60	102.9	70	No	98.23	220.02
KHTP	47.503611	121.975833	41	103.7	67	Yes	29.97	123.98
KLSW	47.542889	122.108167	91	104.5	6.7	No	19.4	129.8
KCMS	47.544167	122.108333	69	105.3	547	No	19.3	129.5
KBKS-FM	47.504639	121.968833	65	106.1	68	No	30.34	123.24
KRWM	47.543472	122.108694	87	106.9	49	Yes	19.32	129.73
KNDD	47.503611	121.975833	41	107.7	67	Yes	29.97	123.98

3.2.3 Measurement approach

Due to the available transmitter information, the RF band in the frequency range of 75 MHz to 700 MHz were continuously measured for a period of three months (20 June 2018 – 21 September 2018). With the long term measurements the variation in signal strength could be observed. The period spans mainly one season, summer, and seasonal variation is not observed. Ultimately all the PSD data (data points for generating PSD plots) were stored locally for further processing.

3.2.4 Processing of measurements/data processing workflow

Google’s Protocol Buffer, or Protobuf, data format is used to store the generated data and corresponding metadata [9]. Protobuf is a language-neutral mechanism to store data. Then lossless compression is applied. In order to use the data, the files need to be decompressed and then Protobuf parser software needs to be used [9,87]. There are Python scripts available that can decompress and parse a single PSD ScanFile (.dsox) [84].

The data processing workflow is shown in Figure 3.6. The raw data can be downloaded from the CityScape Spectrum Observatory web client, or in this case it was obtained directly from the local computer storing the data. The data is zipped when downloaded and should be unzipped and extracted before it can be used. Python is then used to extract the power measurements into matrices so that each matrix contain 24 hours of power measurements per frequency observed. MATLAB is then used to calculate the average received power of the measurements and analyse the data.

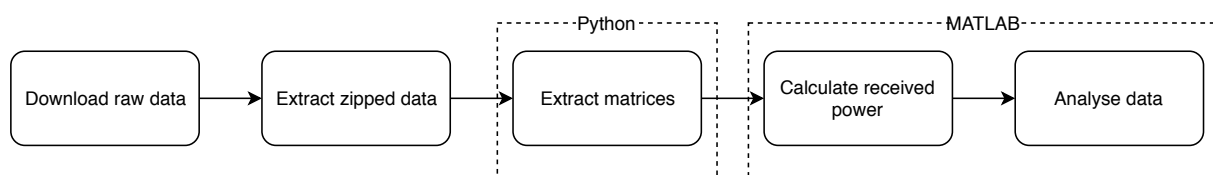


Figure 3.6: Flowchart of data processing workflow

The downloaded data is structured in folders with each folder containing the 24 unzipped PSD ScanFiles for that day, one for every hour. Each of these files needs to be decompressed and then parsed. The result of parsing a decompressed ScanFile is 180 MATLAB matrices per file. That is, 3 matrices for every minute sampled, containing the minimum, maximum and average PSD data respectively. Figure 3.7 illustrates the file structure after the raw downloaded data has been parsed and extracted.

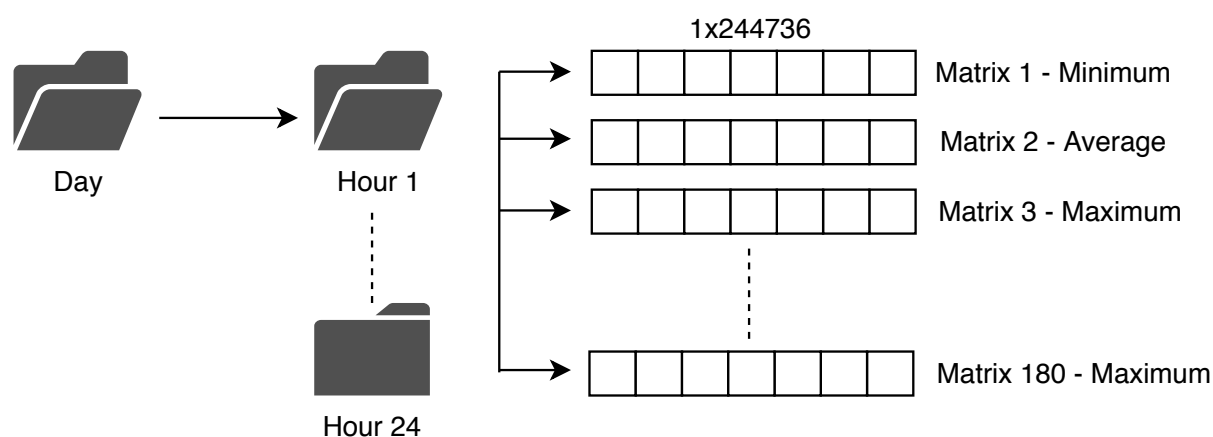


Figure 3.7: File structure after data is extracted

For the UW station, the extracted matrices have dimensions of 1×244736 , with each entry in the matrix representing a frequency bin with the size of 24 414.0625 Hz. Every matrix represents PSD data over a frequency range of 50MHz to 6 GHz, for one minute. The station was configured to only take measurements from 75 MHz to 700 MHz. All the values outside of this frequency range are given as NaN (not a number). Every value in the matrix is thus a point on the generated PSD plot and gives the power for the specific frequency bin.

The computation time to decompress and parse the data with the Python scripts, for one day's samples or PSD's, is approximately one and a half hours using a computer with 16 GB RAM and an Intel Core i7-6700 CPU @ 3.4GHz, running Windows 10.

Python scripts were created that automated the processing of data once it was downloaded and extracted into the (day) folders. This is done by looping through all the files and calling the downloaded Python scripts for decompressing and parsing. One script

decompresses all the ScanFiles and the other parse and process the data. The script that decompresses the ScanFile creates a folder for every hour in the day and moves the decompressed file to that folder. The reason for creating subfolders is to organise the generated MATLAB matrices when the file is parsed. The second script parses each file and stores hourly averaged PSD data matrix with dimensions of 1×244736 . This is done by averaging the sixty minute-average PSD matrices. When an entire day's worth of PSD data is parsed and processed, i.e. all 24 files, the hourly-averaged PSD data is combined and stored in a matrix with dimensions of 24×244736 , one row for every hour. Each day's PSD data is then further processed in MATLAB.

Using MATLAB, the average received power is calculated from the PSD data. As discussed in chapter 2, the PSD plot is the plot of the average power per frequency bin. By adding all those power for every bin in the frequency range of the signal, the average power for the signal can be calculated. Because FM radio signals have a bandwidth of 200 kHz, the PSD is integrated over a bandwidth of 200 kHz, 100 kHz lower and higher than the centre frequency of the transmitter. The PSD is given in dBm/FFT bin size, with the bin size for the UW station measurements being 24414.0625 Hz, versus the frequency in MHz. Therefore for a FM radio signal of 200 kHz, the value of eight frequency bins are summed to determine the averaged power.

Integrating a function is the same as determining the area under the plotted function. The y-axis scale of the PSD plot should first be converted from a logarithmic to a linear scale. Consider the PSD plot after the conversion is done. The amplitude (y-axis) is given in mW/24414.0625 Hz, and the frequency range (x-axis) is given in MHz.

To understand the importance of the units, consider the following intuitive example. A man is travelling in his car at 1 km/minute . If he drives for one hour (60 minutes) he would have travelled a distance of 60 km. Now consider the two plots in Figure 3.8. Both show the same speed at which the car is travelling over a period of one hour. The only difference is the unit of time of the x-axis. The area under the dotted line is coloured in. Ignoring the axis, the area in green is 60 and the area in red is one. For the first plot the y-axis is in km/minute and the x-axis is in *minutes*. The unit of the area

is therefore *km*. For the second plot the y-axis is also in *km/minute* but the x-axis is in hours. The unit of the area is therefore given as

$$\frac{km}{minute} \times \frac{hour}{1} = \frac{km \times hour}{minute} \quad (3.1)$$

When substituting an hour for 60 minutes the unit then also becomes *km*. By integrating both plots and considering the units, the distance travelled both equal 60 *km*.

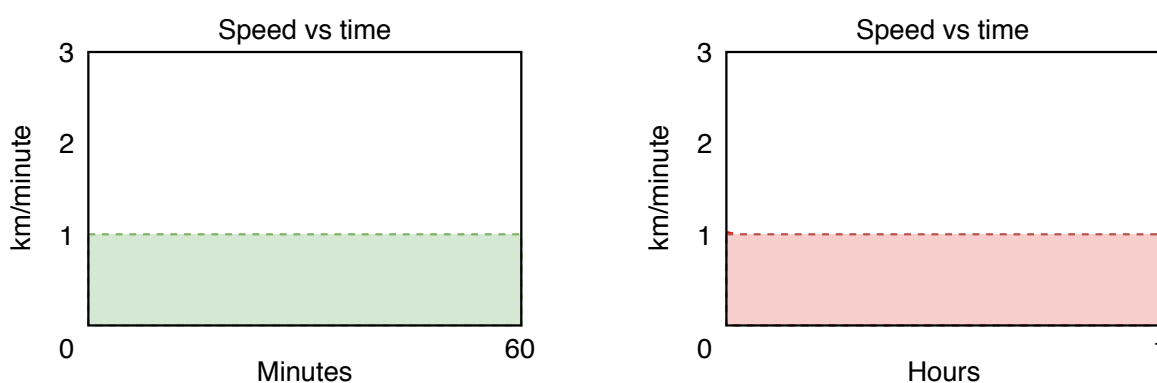


Figure 3.8: Example of plots illustrating the importance of units when computing the PSD.

Similarly, the units must be considered when computing the average power from the PSD plot. Now consider the PSD plot again, with y-axis in *mW/24414.0625 Hz*, and x-axis in *MHz*. Substituting *MHz* for a constant times 24414.0625 *Hz*, the unit of the calculated area is *mW*. For example, when integrating over a range of 200kHz with the y-axis given in *mW/24414.0625Hz*, the limits of the integral will not be 0 and 200 000 but rather 0 and 8.192 (200 000 ÷ 24414.0625). The resulting unit of the area will then be *mW*.

Alternatively one can calculate the integral over the range of 200 kHz and just remember to take the scale of the y-axis axis into account and divide the result by FFT bin size. The advantage of using the first approach is that the raw PSD data points are already in the correct scale and one can be sure of the input to the MATLAB function to compute the integral.

The averaged received power calculated in milliwatt is then converted back to dBm using equation 3.2. The power of each signal is distributed over the complete 200 kHz and the result of integrating over this band should be compared against the published ERP.

$$P_{dBm} = 10 \log_{10} \left(\frac{P_{mW}}{1mW} \right) \quad (3.2)$$

This is done for every hour of the day and a new matrix is stored, containing the average received power per hour for the 30 identified transmitters. The dimensions of the matrix are thus 24x30. The process is repeated for every day's measurements and a final matrix is generated and stored, containing the averaged received power per hour (in dBm) for the 30 transmitters for 93 days. The dimensions of the multidimensional matrix are thus [24x30x93]. Figure 3.9 shows the structure of the final matrix.

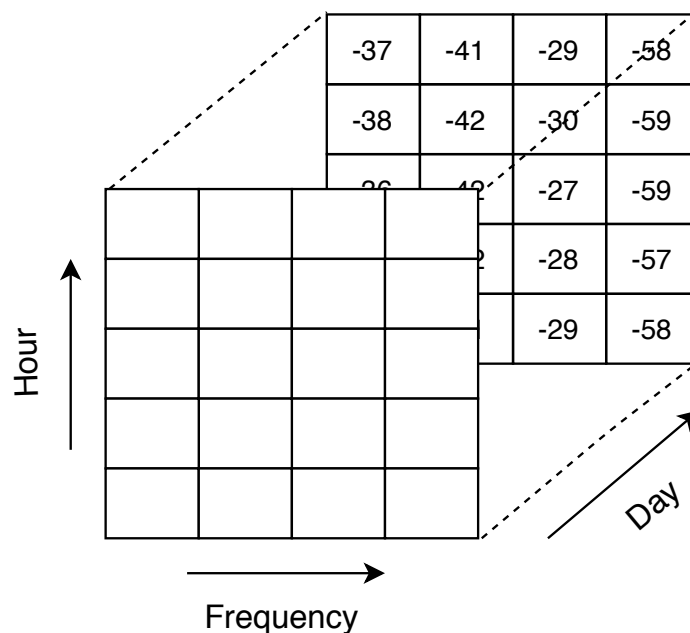


Figure 3.9: Structure of a 3-dimensional matrix containing average received power measurements

The average received power per transmitter is taken over a period of 93 days and is the measured received power that is taken as the ground truth. These values are used to

establish the base line error between the ground truth and the propagation predictions without augmented data as input. It is then also used to reduce the error between measured and predicted received power.

3.2.5 Removing outliers

By analysing the measured data, which is documented in chapter 4, it can be seen that there is intra-day and inter-day variation. Some variation is expected but it is clear that some variation is out of the norm. This is true for large variations within a day and also between days. Without being able to explain the extraordinary variation it can still be seen that for some days the average received power is outliers in the data and needs to be removed. Outliers can be identified by looking at the line plots, for example figure 4.13, and by looking at the box plot in figure 4.14. The box plot clearly indicates outliers. A measurement is considered an outlier if it is more than three scaled median absolute deviations from the median for that transmitter of the period of 93 days. The MATLAB function, *isoutlier*, is used to determine outliers in the data. The median absolute deviation is a robust measure of the variability [88] meaning it has good performance for data drawn from a wide range of probability distributions.

The data revealed that the measurement period could have a significant effect on the measurement study. It could happen that measurements are taken on a day that would be considered an outlier with respect to a longer period and the results could be misleading if that is the only time measurements are taken.

3.3 Processing land cover data

The 2014 version of the Multi-Resolution Land Characteristics Consortium National Land Cover Database (NLCD2011) is used for land cover only. The data used for the DEM is SRTM version 2 data where the voids are filled. The database is discussed in chapter 2. Only 13 of the 16 classes are present in the paths. It is determined by

extracting the unique class values from all the paths combined. The classes present are discussed in chapter 4.

The land cover file is edited in QGIS, a free and open-source geographic information system application, and manipulated to be used in MATLAB. The most important considerations were the file size and format. The transmitters and receivers were plotted on a geographical map in order to establish and confirm the geographical bounds of the area of interest. That is, the area where the SRTM and land cover data are needed to determine the path profile over which the signal propagates. Since the SRTM data is available in 1-degree blocks the bounds were established by the bounds of the SRTM data blocks. The bounds are given as 48°N, 46°N, 121°W and 124°W with the exception being the two unused blocks, influencing the latitude bounds. The blocks represent the SRTM files that need to be stored and also the corresponding land cover files that need to be created. Figure 3.10 is a screenshot of the receiver and transmitters' locations on a map, generated in MATLAB. The bounds of the blocks where the transmitters are in, are highlighted. The locations of the transmitters and receiver have been highlighted with bigger dots. One such dot can represent multiple transmitters in close vicinity to each other.

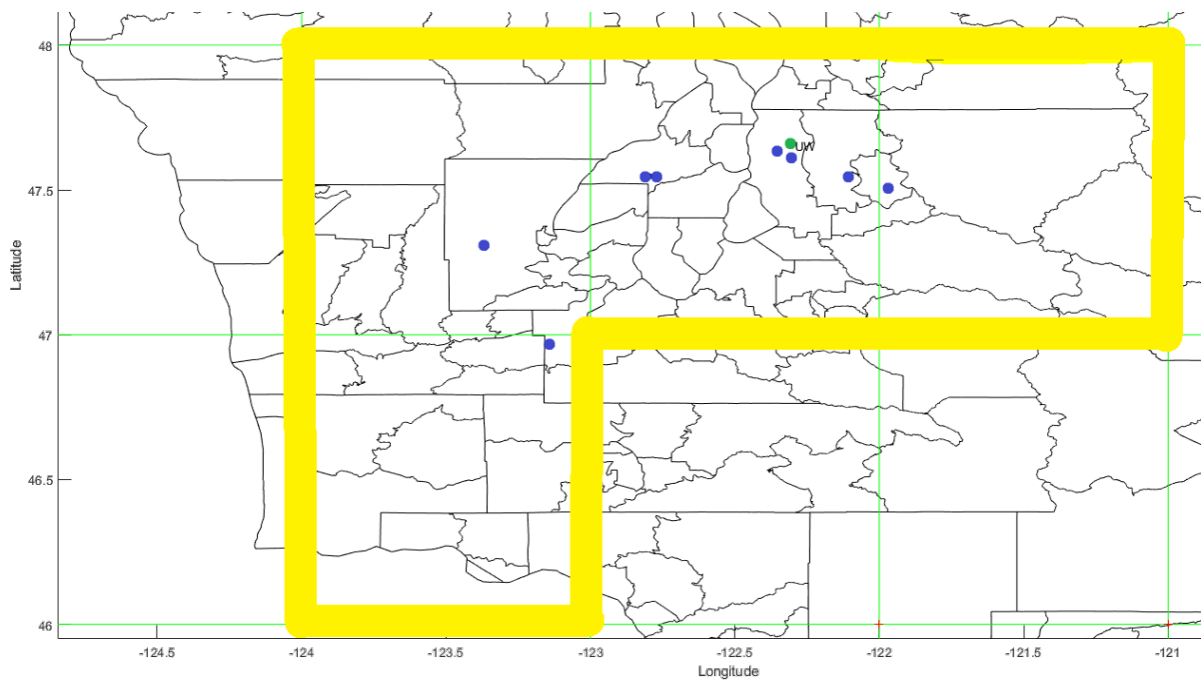


Figure 3.10: Bounds highlighted for land cover and SRTM data

3.3.1 Projection of Coordinate Reference System (CRS)

The land cover data is originally presented in the NAD83 datum with the Conus Albers projection. The EPSG code is EPSG:5070. The EPSG code identifies the CRS and projection of the dataset. The CRS of the SRTM data, which the program SPLAT! uses as the DEM, is WGS84 [68] and the land cover data is thus converted to have the same CRS. The EPSG code is EPSG:4326. The process of projecting, clipping and converting the original land cover dataset in QGIS, to make it useable in MATLAB, follows. The original land cover dataset (in NAD83 / Conus Albers) is clipped or reduced to a slightly larger area than needed, similar to the area of interest as identified by the bounds. This is done to reduce the data size and to reduce the computation time of further manipulation of the dataset. The land cover dataset is then projected to the WGS84 CRS. The CRS of the application that displays the map is changed to WGS84 in order to view the land cover dataset with GPS coordinates. Before that, location information was given in metres East and North and the bounds weren't as intuitively

recognised. The projected dataset is clipped again to reduce the dataset size in order to represent an area just larger than that of the bounds. The dataset is then converted and saved as a TIFF file that can be used in MATLAB.

3.3.2 Grid system

The SRTM and land cover data are from different sources. The coordinates for the land cover data needs to be aligned with the coordinates for the SRTM data so that the associated height of the land cover can be added to the height of the DEM at the correct locations.

The grid system of MATLAB and that of the SRTM and SDF files (Splat's processed DEM files) differ. The goal is to create a grid for the land cover dataset that is similar to the SDF grid, in order to be able to add the data contained in both grids. The MATLAB grid system is shown in Figure 3.11, the SDF grid system is shown in Figure 3.12 and the SRTM grid system is shown in Figure 3.13. The MATLAB and SDF grids represent a 1-degree tile and each tile consists out of 1200x1200 pixels, each pixel representing a 3-arc second by 3-arc second area. The SRTM grid only differs by having 1201x1201 pixels. For the MATLAB grid system the reference vector or reference point is given as the north-west corner of the grid and the north-west corner of the pixel. The reference vector for the SRTM grid is given as the south-west corner of the grid and the centre of the pixel. The SDF grid has the same system as the SRTM grid with the only difference being the the number of columns and rows. The SDF grid is a processed SRTM grid in SPLAT! and the additional overlapping rows of the SRTM data is already removed. The voids and outliers in the SRTM data is also replaced when the SDF data is generated. Only the SDF grid will be discussed further since it is similar to the SRTM grid and will be augmented with the land cover data. The SRTM grid will only be used as verification of the generation of the SDF and land cover grids. Both the SDF and MATLAB grids columns are ordered from left to right and the rows from bottom to top. This means that the matrix containing the grid data will have data in the first row, that when plotted on a map, will be displayed at the bottom as depicted in the figures.

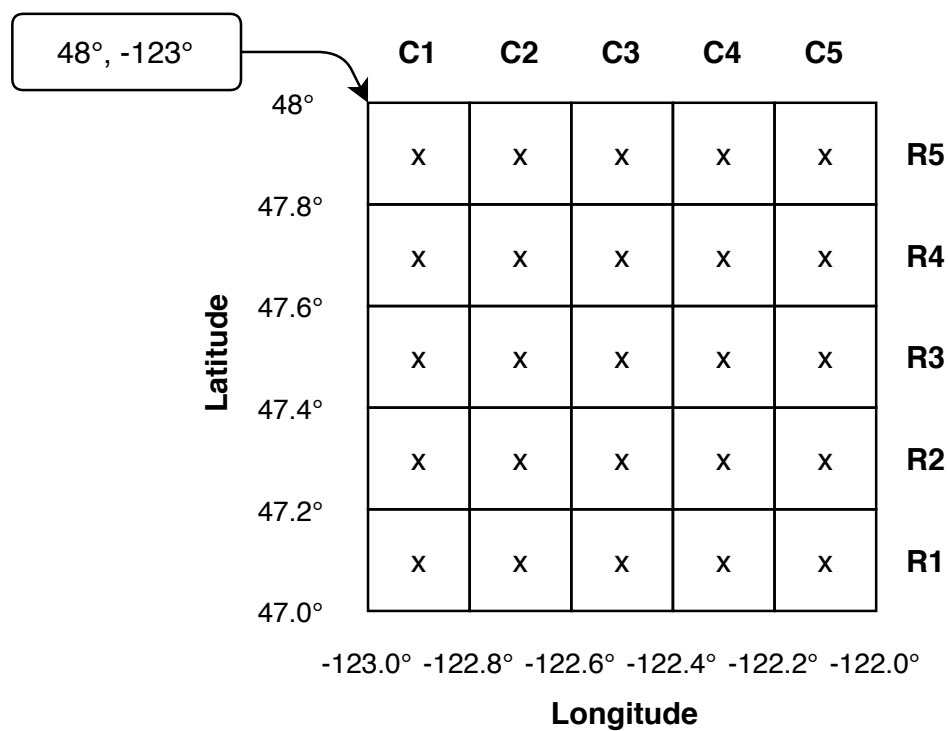


Figure 3.11: MATLAB grid system

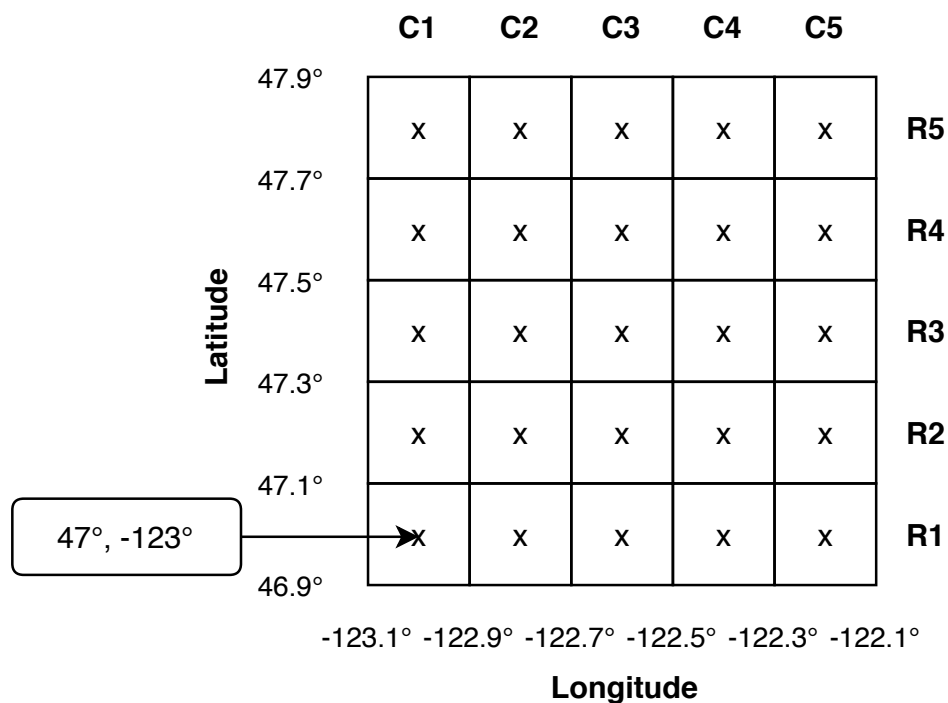


Figure 3.12: SDF grid system

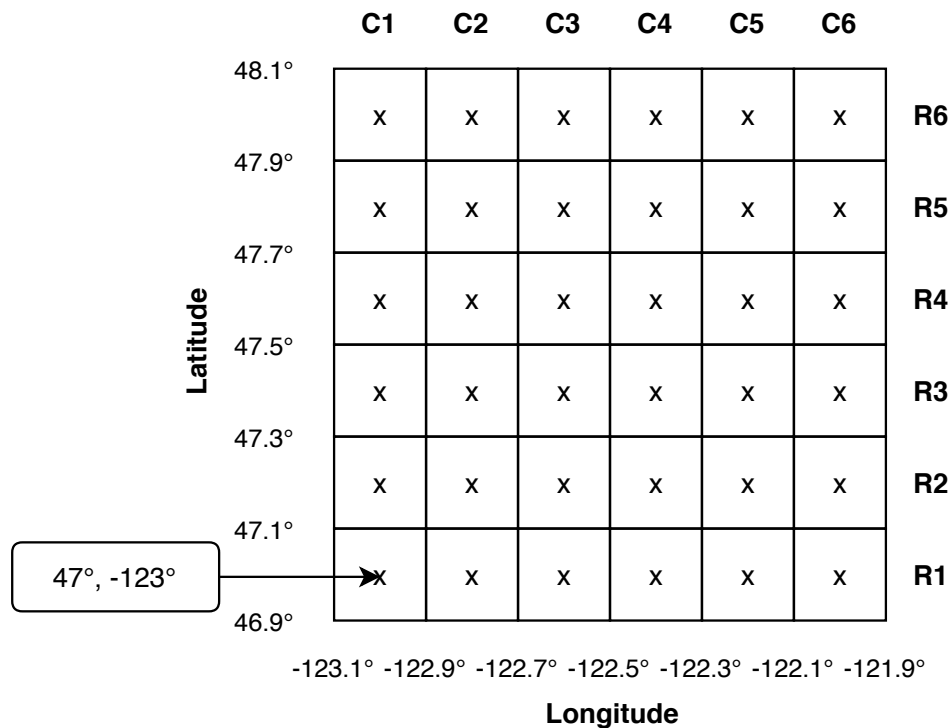


Figure 3.13: SRTM grid system

An empty SDF grid is created to store the land cover data. The grid has 1200 rows and 1200 columns for 3-arc second (90m resolution) data. The processed land cover dataset is read with the MATLAB function, *geotiffread*. The function also creates a geographical cell reference object containing the metadata, and which is used to reference the grid in order to extract a class value at a geographical location given in GPS coordinates. The MATLAB function, *ltn2val*, is used to populate the grid. The function takes a single latitude and longitude value and returns the value in the land cover grid at that geographical location. The land cover's reference cell object is used as a reference.

The latitude and longitude values are determined by examining the SDF file name. The SDF file name convention, e.g. 47:48:-122-123.sdf, gives the bounds of the grid. For this example 47 and -123 in the name would point to the minimum degrees latitude and longitude and 48 and -122 would point to the maximum degrees latitude and longitude. Figure 3.14 illustrates the degrees associated with the directions. North and East are positive and South and West are negative. The coordinates are therefore

referenced to North and East. For example $47^\circ, -123^\circ$ would be 47 degrees North and 123 degrees West. It should be noted that SPLAT! uses West instead of East as positive. The longitude of the transmitter and receiver should be given accordingly.

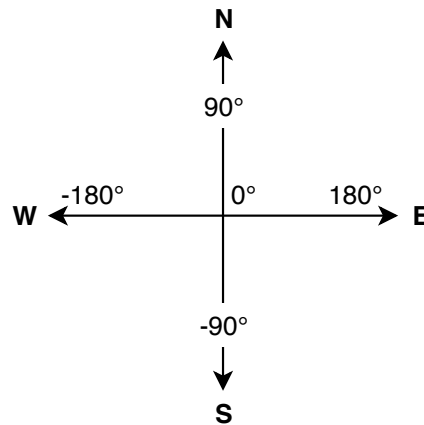


Figure 3.14: Direction with degrees

As per the format of the SDF grid, the reference vector is defined as the minimum latitude and longitude. When extracting the land cover data, this latitude and longitude are used as the first coordinate. The matrix is populated by increasing the latitude and longitude with 3-arc seconds until all 1200 rows and columns are populated. The last coordinate will be the maximum latitude and longitude less 3-arc seconds.

The exact coordinates, as needed for the SDF grid system, are used and the SDF grid and land cover grid can be aligned exactly. Comparing the original land cover dataset with the SDF grid, the value of the land cover is also in the centre of the pixel. The land cover dataset has a resolution of 30m. When populating the empty grid the resolution is changed to 90m by using the class value nearest to the geographic location in the original land cover dataset.

To display the grid for visual conformation, the reference vector is changed to fit the format of the MATLAB grid since the difference between the grid systems is in the reference point. Essentially the data is only a matrix and useful to the tool if the correct reference point is applied. The reference vector is therefore changed to the maximum latitude and minimum longitude less half of 3-arc seconds respectively. That is because

in the MATLAB grid system the values are in the centre of the pixel with the reference point on the corner and in the SDF reference system the value and reference point is in the centre of the pixel.

3.4 Reducing the error

The aim of the study is to ultimately reduce the average error between measured and predicted received power, as mentioned in chapter 1. Additional information in the form of land cover data is added to the DEM in order to achieve the reduction in error. A height is associated with each land cover class that is added to the height of the DEM.

3.4.1 Modelling rationale

The rationale is that the land cover, which is not included in the DEM, is an additional obstruction to the electromagnetic wave and causes additional loss, be it due to absorption, refraction, reflection or scattering. The difference in measured and predicted path loss is attributed to this. Different land cover classes can have different effects on the electromagnetic wave due to physical and electrical properties of the type of land cover. The assumption is made that land cover class has the same effect on the frequencies measured since the wavelength of each transmitter in this band (88-108 MHz) is around 3m long. The land cover is also defined for an area of 90mx90m and the assumption is made that the land cover is homogeneous in that area. The original land cover dataset has a resolution of 30m and a nearest-neighbour method is used for upsampling it to a 90m resolution. When considering the nine 30m blocks which makes up one 90m block, it is possible that the land cover chosen for the 90m block is not the majority of the land cover of the nine 30m blocks. Even if the majority of land cover is chosen the assumption may still not be entirely true. As the assumption could be made for the original dataset, the assumption for the upsampled set is necessary

and not an unreasonable simplification.

A gradient descent method is used to determine the heights associated with each land cover type in order to get the most reduction in error. There is thus a correlation between the effect of the land cover class on the path loss and the additional heights added to the DEM used in the propagation prediction. As mentioned in chapter 2, although the gradient descent method is slow, it works for a function with any number of variables. Since there are many land cover classes, and considering that by adding a height to the DEM according to the land cover class, the predicted path loss function has many variables that can be tuned. The gradient descent method is chosen because the function to be fitted has many variables affecting the path loss, and it is also relatively easy to implement. The author of [63] also successfully used the gradient descent method to augment a large scale radio propagation model with land cover data.

The added height due to the land cover class does not however represent the physical height of what is on that land, e.g. trees are 50m tall so the height of that class is set to 50m. It rather indicates the effect of the land cover class on the attenuation due to wave propagation through or over obstructions.

3.4.2 Cross-validation

Cross-validation (CV) is a technique to validate a model using the same set of sample data for training and validation of the model. This is useful when only a small set of data is obtainable. A popular use of CV is k-fold cross-validation. In k-fold cross-validation the dataset is divided into k disjoint, equally sized subsets. One set is used for validation of the model and the remaining k-1 sets are jointly used for training of the model. The process is repeated k times, using a different set for validation each iteration. This way a large fraction of the data can be used for training and all the cases appear in the validation set. The drawback is that the model must be trained k-times instead of once. Typical values for k are between 3 and 10.

A special case of k-fold CV, where the size of a subset is equal to one, is leave one-out-cross-validation (LOO-CV). The advantage is that a bigger subset of the data is used for training, which could improve the accuracy of the model. The drawback is that it is computationally more expensive, as the number of iterations required to use all the subsets once for validation, is as many as the samples in the set.

Figure 3.15 illustrates an iteration of 4-fold cross-validation. The 12 samples in the sample set are randomly split into 4 disjoint, equally sized subsets. This is only done once and the subsets are fixed. For the first iteration one subset is used for validation while the combination of the other 3 subsets are used for training of the model. In the second, third and fourth iteration a different subset will be used for validation while the combination of the remaining subsets will be used for training.

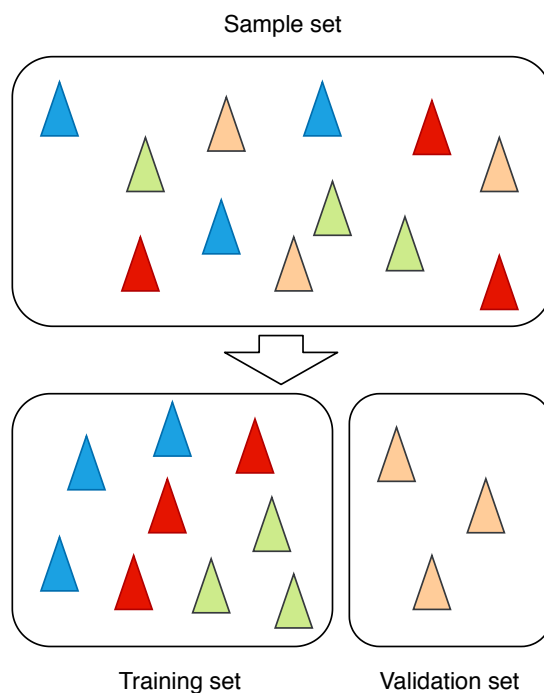


Figure 3.15: 4-fold CV example

3.4.3 Gradient descent method

For this implementation the function to be minimised is the RMSE. The RSME is calculated as shown in equation 3.3.

$$RMSE = \sqrt{\frac{\sum_{i=1}^k (M_i - P_i)^2}{k}} \quad (3.3)$$

, where M is the measured received power in dBm and P is the predicted received power in dBm for either the training or validation subsets, depending on which RMSE is calculated. k denotes the number of samples in the subset.

The prediction is a function of the DEM height, and also the land cover height added to it, and therefore the RMSE is also a function of the land cover height. The partial derivative is estimated as given in equation 3.4. Let the RMSE be denoted as $f(\mathbf{h})$.

$$\frac{\delta f(\mathbf{h})}{\delta h_i} \approx \frac{f(\mathbf{h})|_{h_i=H_{ni}+\Delta H} - f(\mathbf{h})|_{h_i=H_{ni}}}{\Delta H} \quad (3.4)$$

, where ΔH is the increase in height and H_{ni} is the current height of the land cover class, i . All heights are in metres.

The gradient vector contains the partial derivatives of the function $f(\mathbf{h})$ and is denoted by $\nabla f(\mathbf{h}_n)$.

In order to decrease the error one must move against the calculated gradient. Thus the new heights are estimated as given in equation 3.5.

$$\mathbf{h}_{n+1} = \mathbf{h}_n - \frac{\nabla f(\mathbf{h}_n)}{\|\nabla f(\mathbf{h}_n)\|} \Delta H \quad (3.5)$$

The new heights are rounded to the nearest integer since SPLAT! only uses integers for the DEM.

It can be shown that the function is not convex and the algorithm is not guaranteed to find a global minimum. The function is tested not to be convex by increasing a variable (one of the heights) monotonically and observing the output or RMSE. As the height is increased, the RMSE decreases, then increases, and then decreases again. If it was convex the output would not decrease for the second time.

3.4.4 Logical flow of the method

Figure 3.16 shows the flow diagram of the proposed and implemented gradient descent method to decrease the error. Each step is then explained.

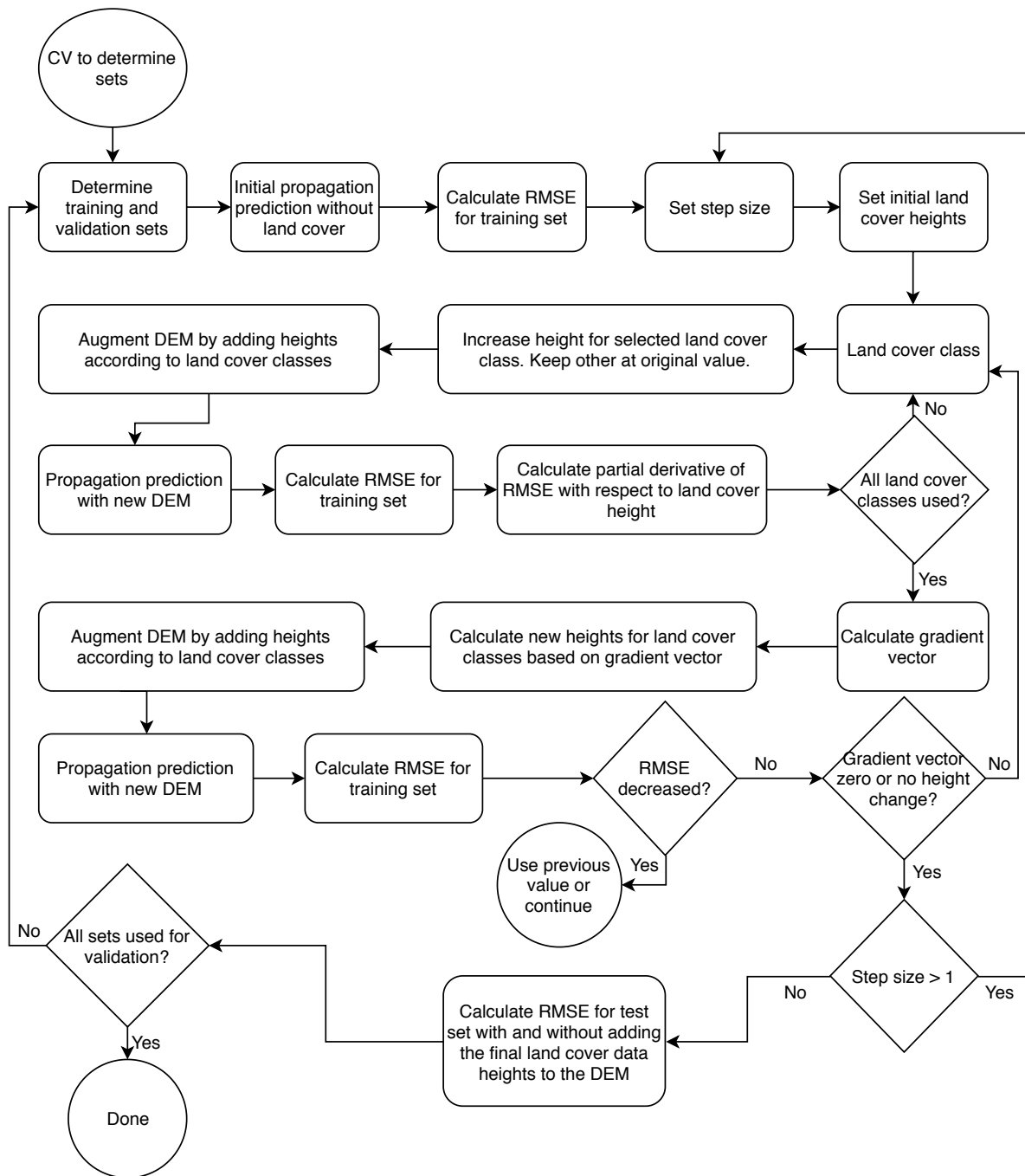


Figure 3.16: Flow diagram of the gradient descent method for reducing error

Cross-validation to determine sets

For the case of 30 transmitters, 5-fold cross-validation is done. The sample set of 30 transmitters is randomly ordered and split into 5 subsets of 6 transmitters each.

Determine training and validation sets

For iteration one, use the first subset as the validation set and combine the rest of the subsets for training. For iteration 2, use the second subset as validation and so forth until all 5 subsets were used for validation once.

Initial propagation prediction without land cover

Do the initial propagation prediction, using SPLAT!, with the original DEM from the SRTM data and store the predicted received power. This is done without adding land cover data and the initial heights can therefore be seen as zero.

Calculate RMSE

Calculate the baseline RMSE for both the validation and the training sets using equation 3.3. The baseline RMSE for the training set will be used in the calculation of the partial derivative for the first iteration.

Set step size

The step size refers to the height to add to the initial height, associated with the land cover class, of the iteration at hand. It is denoted as ΔH .

Set initial land cover heights

The initial heights are the heights associated with each land cover class. These heights are the starting point for the search of a local minimum. Changing the initial values may result in finding a better local minimum. The initial heights can be reset after the step size is reduced. This is typically the case when the step size is too large and the algorithm climbs out of the minimum, resulting in an increase in the error. The last point before the error increased is then used as the initial height after the step size is reduced.

Land cover class

One land cover class is selected and evaluated at a time in order to determine the effect

of that land cover class on the predicted path loss.

Increase height for selected land cover class. Keep other at the initial value

Add the height, as determined by the step size, to that (initial height) of the land cover class. Keep the other heights associated with the land cover classes not selected, as in the previous step, at the initial heights. Consider the initial height vector $\mathbf{h}_I = [h_1 \ h_2 \ \dots \ h_n]$, with h_1 associated with the first land cover class etc. For every iteration using the different land cover classes the height vector used will be $\mathbf{h}_i = [h_1 \ h_2 \ \dots \ h_i + \Delta H \ \dots \ h_n]$ with ΔH being the step size. Only one height is changed at a time while the others remain at the initial values. Importantly, it is the initial values that are kept. So, if for example the first height is changed and the process moves to the next land cover class, the first height is changed back to the initial value since the effect of the first height has already been observed. By not changing the height back to the initial value one will observe the effect of the combination of land cover classes.

Augment DEM by adding heights according to land cover classes

Heights are added to the DEM at the locations at which the land cover class is present. The original DEM is read into MATLAB and the appropriate grid is created. Another two grids are created for the land cover classes: one containing the class number, and another containing the associated heights. The height grid is added to the DEM grid and a new SDF file is generated and stored.

Propagation prediction with new DEM

The propagation prediction is performed using SPLAT! and the newly generated SDF file as input. All the other parameters remain the same.

Calculate the partial derivative of RMSE with respect to land cover heights

Calculate the partial derivative of the RMSE with respect to the land cover height. Since the propagation prediction is a function of the height of the DEM, which in turn is a function of the height of the land cover classes, the partial derivative with regards

to the height of a land cover class can be computed.

All land cover classes used

Determine whether all the land cover classes has been used and the partial derivatives calculated. If not, use the next land cover class and repeat the process.

Calculate gradient vector

The gradient vector contains the partial derivatives of the RMSE with respect to each land cover class height. The gradient vector is given as $\nabla f(\mathbf{h}_n)$.

Calculate new heights for land cover classes based on gradient vector

Calculate the new heights with equation 3.4.

Calculate RMSE for the training set

The new RMSE is calculated only for the training set. This will be used in the determination of the partial derivatives in the next iteration. This step considers all the height changes when determining the RMSE and not only that of one land cover class.

RMSE decreased

Check if the error decreased or increased. If the error decreased continue with the process. If the error increased it means that the algorithm is climbing out of a minimum. The user can either choose to use the previous point (heights) as the initial heights and reduce the step size to try and find a smaller minimum or the user can choose to let the algorithm continue. It could be that the algorithm climbs out of one minimum and into another in the next iteration, in which case the error could even be lower. It could be particularly useful with the initial search of local minimums. The user's discretion can be used. If the step size is already at its lowest; the last RMSE before an increase, together with the heights to achieve it, should be taken as the minimum.

It could happen that the algorithm climbs in and out of the same minimum and that the heights are repeated. In the cases of repetitive height the algorithm should be stopped.

Gradient vector zero or no height change?

If the gradient vector is zero (or very small) or there is no height change it means that a minimum is found and further iterations will yield the exact same result. The algorithm with the current parameters should then be stopped.

step size (ΔH) > 1

Decreasing the step size could improve the reduction in error since smaller, finer changes are made to the model. If the step size (ΔH) is greater than one, the land cover heights that resulted in the smallest error should be used as the initial heights and the step size reduced. The amount of reduction is up to the user's discretion. If the step size is equal to one, no finer steps can be taken and the process should continue to the next step.

Calculate RMSE for test set with all land cover classes heights used

When the model is trained and no further improvement can be made the RMSE for the test set should be calculated using equation 3.3. The DEM used should be augmented with the heights that resulted in the smallest error during the training of the model.

All sets used for validation

The sample set was divided into five subsets of six samples each. One set was to be used as the validation set while the other sets are used for training the model. A different set would be used for every cycle of the process. If all five subsets have been used once as the validation set, the process of training the model with the gradient descent method is complete. The average of the RMSE and other statistical indicators of the five sets are computed. The heights are also averaged and the result is used to augment the DEM for the evaluation set. If all the sets have not been used the process is started with the next subset as a validation set.

3.4.5 Initial parameters and assumptions

The gradient descent method requires initial parameters or a starting point to search for a local minimum. Without knowing anything of the function to minimise, any random starting point is good as the other. Such is the case for this study. Assumptions are made in order to reduce the complexity of the function and to reduce the computation time. Assumptions regarding the initial heights of the land cover classes are also made in order to guess a starting point which would probably yield a good result. The entire function space to search is too large and the computational time required too long. A global minimum is therefore not guaranteed. In addition to the assumptions that the land cover is homogeneous in a 90mx90m square and that each land cover has the same effect on each of the radiated EM waves, the following assumptions are made.

Due to the hypothesis that the land cover acts as an additional obstruction, the height associated with the land cover class is limited to positive heights for this study. Water can have an effect on radio wave propagation but is not considered or modelled by adding a height to the digital elevation model based on the premise that the modelled loss is largely due to the land cover being an obstruction.

Due to the process of obtaining SRTM data, impervious surfaces like dense tree canopies could be included in the DEM [69] as explained in chapter 2. The assumption is made that the DEM represents the bare earth ground level.

In order to keep the dimensionality of the function low and thereby reducing the computation time and complexity of the problem, only certain land cover classes in the paths are considered. Based on the proportion of the land cover present in all the paths as seen in Figure 3.17, only the developed and forest classes are considered.

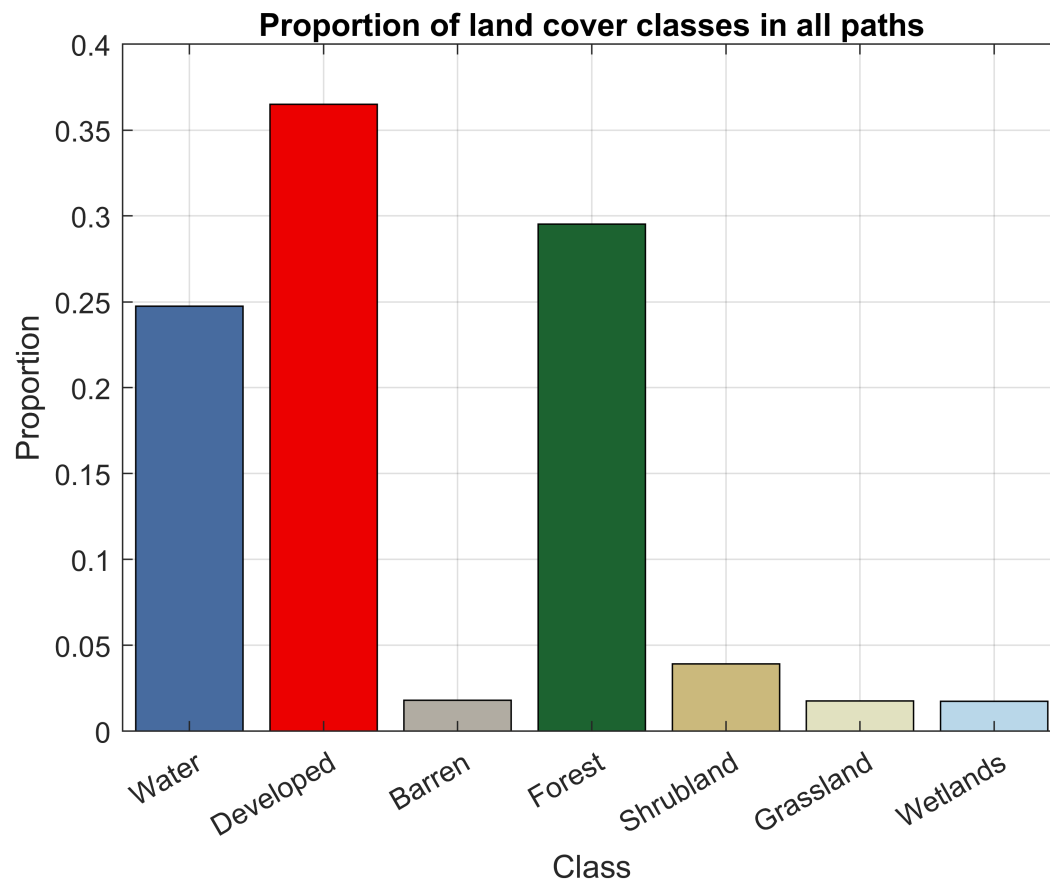


Figure 3.17: Proportion of land cover over all paths

Since the measurements were taken during the summer months in Seattle it is assumed that all the trees had leaves and the canopy of the forests were covered with green foliage. All the forest classes, i.e. classes 41 to 43, are grouped together to reduce the number of variables and computational time. The same does not apply to the first three developed classes (21, 22 & 23) since the percentage of impervious cover differ, even though the description of mixed vegetation and constructed material is similar. The dimensionality is thus reduced to 5 distinct classes. The five classes are thus the four different types of developed classes and the three types of forest classes grouped together as one. Although there is a lot of water present in the paths, it was not considered as obstructions and therefore not considered for this study. As will be discussed in chapter 6, the effect of water can also be modelled using this method and may be considered as future work.

To find initial starting points for the algorithm, all the transmitters can be used and no cross-validation is needed. Several approaches are taken to estimate good initial heights for the land cover.

1. The first approach is to start at heights extrapolated from similar research. Ballot [63] also used a gradient descent method to determine the effective heights of land cover classes. Although a different land cover database is used and the geographical location is different, the heights might be a good starting point. Using table IV in [63] as a guideline, the initial heights for the respective classes are chosen as 1m for classes 21 & 22 (Low intensity and open space developed), 4m for class 22 (medium intensity developed), 5m for class 24 (high intensity developed) and 2m for the combined forest class. The authors of [71] also tuned a propagation model for a DVB-T2 network using measurement. Their measurements were taken in an eastern part of Croatia and forest were identified as the main contributor to additional path loss due to land cover. The best results of their study were obtained when, amongst other considerations, adding an effective clutter of 25m to the DEM for the forest land cover class. The height of 25m is used as the initial height for the forest in this study since the trees in Washington are much larger and dense compared to the woodland of Ballot's [63] study.

2. Another approach is to start with all the heights set to zero and observe the systematic increase in heights as the error is reduced. A possible pitfall is that such small changes in the model may not have a significant impact on the error. Alternatively initial heights can be large within reason and a decrease in error would probably lead to a decrease in heights.

3. A third approach is to try several random initial heights and observe the changes to get a feeling for the solution. Intuitive thinking is required to guess good initial heights based on previous iterations' experience.

3.4.6 Implementation

The gradient descent method is implemented in MATLAB while the propagation prediction is implemented using SPLAT!'s Longley-Rice ITM and ITWOM. A Python script is used to scan the generated reports from SPLAT! and extract, calculate and store the predicted average received power. Figure 3.18 illustrates the interaction between MATLAB, SPLAT!, Python and the files.

MATLAB generates modified SDF files, SPLAT! uses it for the propagation prediction and produces reports. The Python script reads the reports and saves a file containing the predicted powers. MATLAB then reads that file.

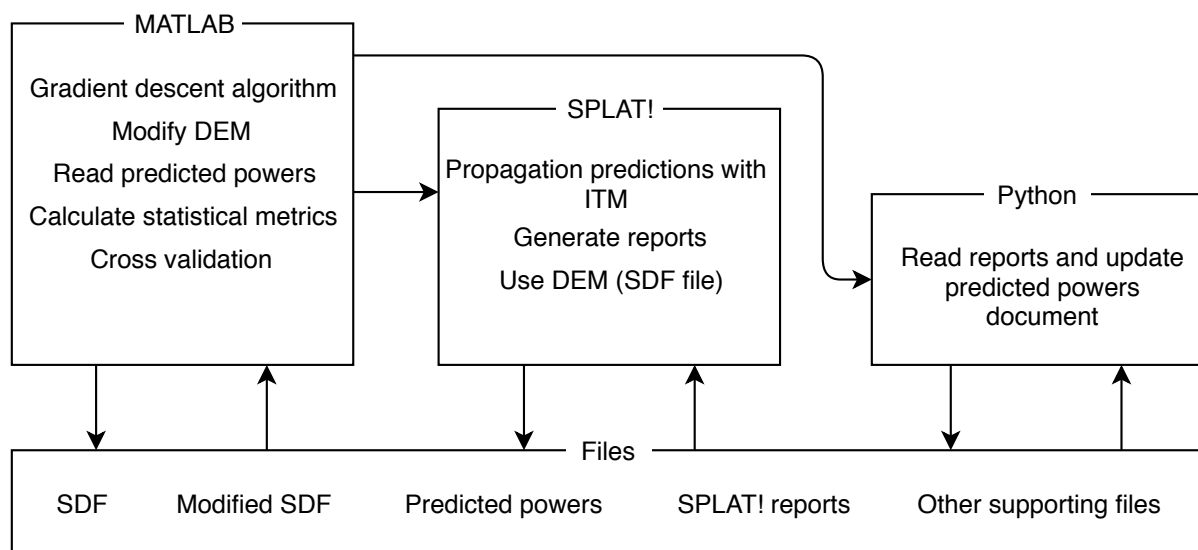


Figure 3.18: Interaction of software in the implementation of the method

SPLAT!

SPLAT! version 1.4.0 on Linux is used for this study. The computations and algorithms used are the same as in newer versions.

Input file generation

Both the site location (QTH) and ITM parameter (LRP) files are generated for every

transmitter beforehand. The QTH file is also generated for the receiver. The QTH and LRP files are generated from the search results of the FCC database as mentioned in the transmitter selection section.

The QTH files are generated by extracting the transmitter site's geographic location and the antenna's height above average ground level. The names of the QTH files are the same as the call sign for the transmitters. An example of a QTH file follows.

```
KBCS           ;Name
47.543472      ;Latitude
122.108694     ;Longitude
79m           ;Antenna height (AGL)
```

The latitude and longitude can be given in decimal degrees format or degree, minutes, second (DMS) format [11]. Latitudes north of the equator are positive and longitudes west of the Greenwich meridian are positive. The antenna height above ground level (AGL) is given in feet unless an "m" is appended to the height to indicate that it is given in meters.

The LRP files are generated with the default values. The frequency and ERP at which each transmitter broadcasts is changed respectively. The LRP file has the same name as the corresponding QTH file so that SPLAT! can find it when the propagation prediction is executed. An example of an LRP file follows. As mentioned earlier, the model is insensitive to the radio climate for paths shorter than 100km and the default values can be used.

```
15.000         ;Earth Dielectric Constant (Relative permittivity)
0.005          ;Earth Conductivity (Siemens per metre)
301.000        ;Atmospheric Bending Constant (N-units)
91.3           ;Frequency in MHz (20 MHz to 20 GHz)
5             ;Radio Climate (5 = Continental Temperate)
0             ;Polarization (0 = Horizontal, 1 = Vertical)
0.50          ;Fraction of situations (50% of locations)
```

0.50 ;Fraction of time (50% of the time)
1800 ;Effective Radiated Power (ERP) in Watts (optional)

Antenna pattern files can also be generated and will be taken into consideration when the propagation prediction is executed. Instead of adding antenna pattern files for the directional antennas, the bearing between each transmitter and the receiver is determined, and the gain of the directional antenna in that direction is determined and stored in a separate file. The antenna pattern is taken into account when the predicted received power is calculated by first multiplying the ERP with the gain and then subtracting the predicted path loss.

SPLAT! is a command-line driven program. A shell script is created that loops through all the transmitters and calls the SPLAT! command with the appropriate input parameters. An example of such a command follows.

```
splat -t Tx_name.qth -r Rx_name.qth -metric -d ../path_to_SDF_files/
```

The ITM is used in point-to-point mode. SPLAT! generates an output file containing all the necessary information regarding the link between the transmitter and receiver. The information includes the path loss calculated with the ITM. It also includes other information that is useful for analysis of the link like the path profile, distance and bearing from the receiver to the transmitter and vice versa, and also obstructions that are encountered in the path. A Python script extracts the path loss from this file, extracts the other information like the antenna pattern, ERP and other metadata from another file and calculates the received power for every transmitter. This information is then saved in a separate file that is accessed by MATLAB to read the predicted received power.

MATLAB

The SDF and land cover class grids are created once at the beginning of the method. The geographic locations of the transmitters and receiver are established. It is important to have the transmitters' and receiver's locations when the heights are added to

the SDF file. The pixels where transmitters and the receiver are located should remain unchanged when adding heights due to land cover. If any height is added at the above mentioned locations, it will ultimately change the transmitter's or receiver's height AMSL, which is incorrect.

The sample set is randomly ordered and split into five disjoint subsets of six transmitters each. The transmitters are identified by their operating frequency. Five binary masks are generated, one for each of the five iterations. Each mask contains zeros for one subset that will be used as the validation set and ones for the rest of the samples that will be used for training.

The gradient descent method is then implemented. After each height change a new modified SDF file needs to be generated and the propagation prediction needs to be done. A function is called that adds the generated land cover height grid to the SDF grid and writes a new SDF file in the format that SPLAT! requires. The shell script that calls the SPLAT! commands and also calls the Python script that calculates the received power, is called from MATLAB when the propagation prediction needs to be done. The resulting file containing the predicted received power is then read with MATLAB and the predictions are stored in a vector.

3.5 Conclusion

The research methodology was discussed in three main parts, namely the measurement study, processing of land cover and the method of reducing the error.

For the measurement study, the receiver setup was discussed, and the transmitters selected. The measurement approach and processing of the measurements were explained up to the point where the average received power value was calculated from the PSD.

The processing of the land cover data was discussed. It was explained how the land

cover data was manipulated and processed in order to be able to add it to the DEM, using MATLAB.

The method of reducing the error was also discussed. The gradient descent method, initial parameters and assumptions were explained. The logical flow of the method was explained. Finally, the implementation of the method using, a combination of MATLAB, SPLAT! and Python scripts were discussed.

The exploratory data analysis on the field strength measurements and land cover data will be discussed in the next chapter. The results of reducing the error between field strength measurements and propagation predictions will also be discussed.

Chapter 4

Exploratory data analysis and results

This chapter starts with an exploratory data analysis of the measurement study. The resulting average received power to be used as the ground truth is then given. The land cover data is then analysed and the paths consisting out of land cover classes per unit distance is then extracted for every transmitter-receiver pair. Finally the results of reducing the error between the field strength measurements and predictions are discussed.

4.1 Measurement study

The results of the measurement study, as explained in section 3.2, are discussed in this section. The measurement approach is explained in section 3.2.3, and the processing of the measurement data is explained in sections 3.2.4 and 3.2.5. Measurements for a period of one day, eight days and 93 days are discussed and analysed separately. The results are summarised, and the average received power per transmitter that is used as the baseline, is tabulated.

Figures 4.1 and 4.2 show the daily average received power for two separate transmit-

ters on the same date, 1st of July 2018. The blue line indicates the hourly measured power. The red line indicates the average received power for the day and the green line the average received power for eight consecutive days. From these two figures, the hourly variance in received power can be observed. For the transmitter, call sign KLSY, there is not much variance (note the scale). The standard deviation is 0.0967 dBm. For the other transmitter, call sign KBCS, the received power varies more during the day, with a standard deviation of 1.597 dBm. The average received power for the day is also different than that of an eight-day period. The difference can be small (just over 0.1 dBm) or it can be large (almost 2 dBm). Determining the average measured path loss to compare it with the predicted path loss is a challenge.

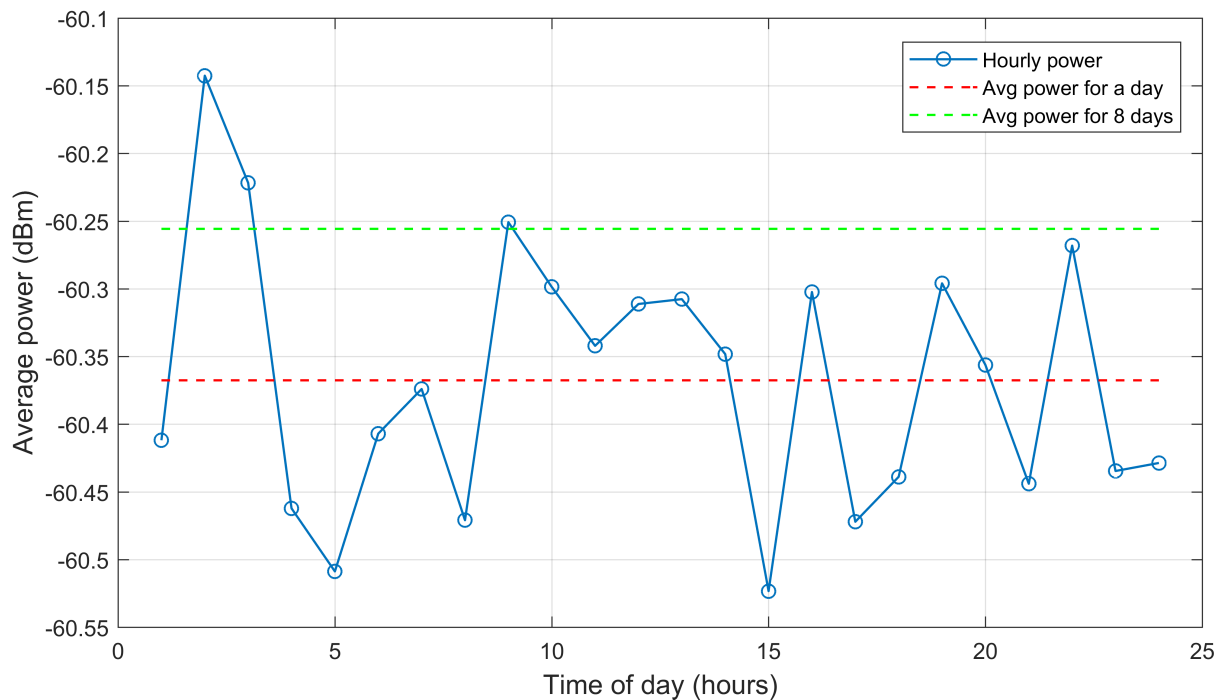


Figure 4.1: Average received power for transmitter KLSY for 1 day

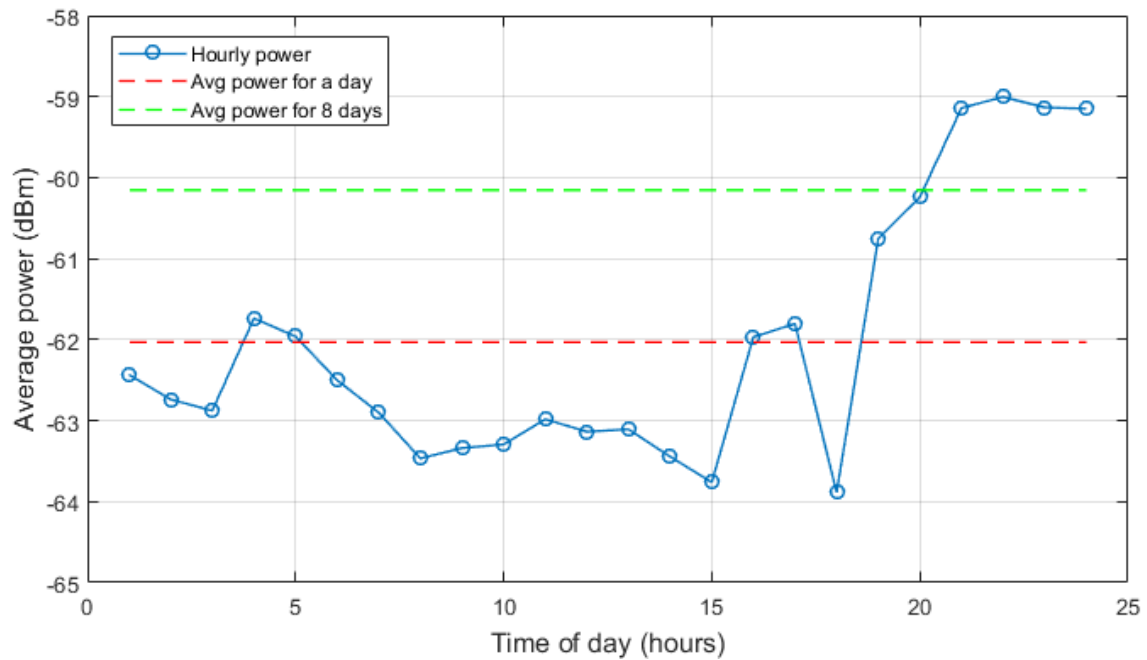


Figure 4.2: Average received power for transmitter KBCS for 1 day

Figure 4.3 shows the geographic locations of the transmitters (blue dots) and the UW station receiver (green dot). There are transmitters that are located on the same tower meaning they all have the exact same geographic location. The difference between the transmitters are the height of the antenna above ground level, the ERP, antenna patterns and broadcasting frequency. There are also transmitters which are located very close to each other and appears as one transmitter on the big map. The blue circle shows a zoomed area of the map that contains multiple transmitters even though it appears as only one on the bigger map. The transmitters are situated as such that they can be clustered together so that the distances between the receiver and transmitters in a cluster are similar. Five clusters are formed. The distances for each cluster from the receiver are approximately 4km, 20km, 30km, 40km and 100km respectively. The “cluster” at 40km is composed of only a single transmitter where the rest of the clusters contain multiple transmitters.

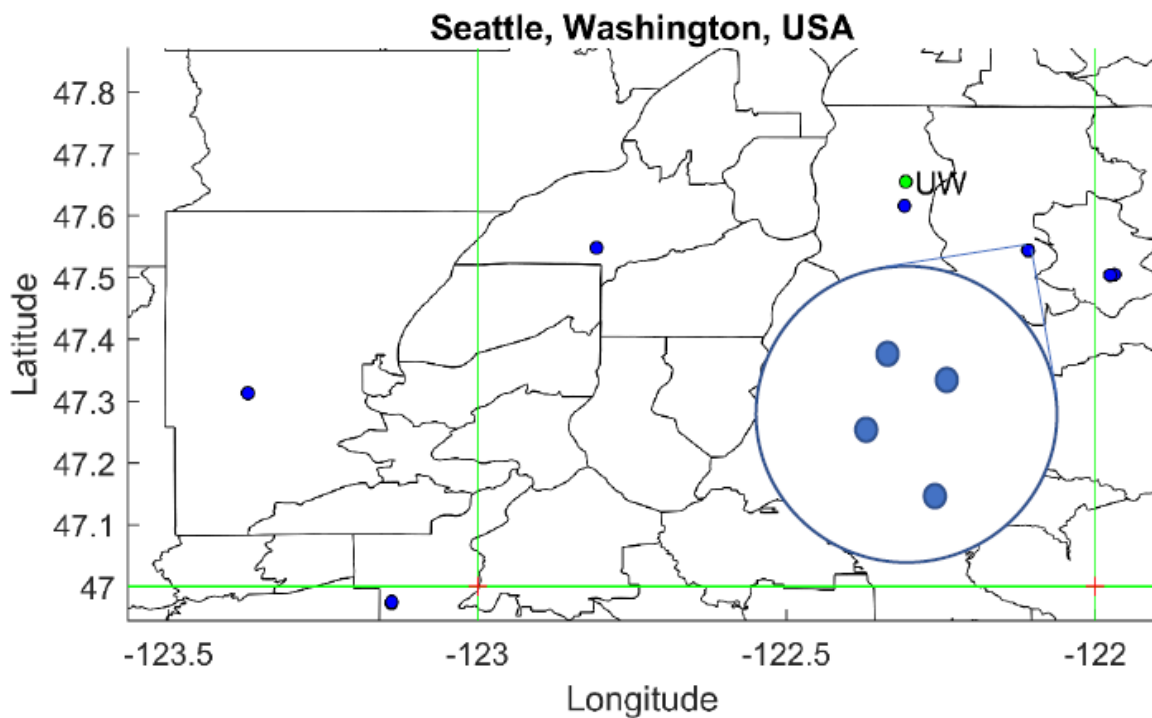


Figure 4.3: Geographic locations of transmitters and receiver

Measurements over a period of eight days, from 1 to 8 July 2018, are shown for three of the clusters that contain the majority of the transmitters. The daily average received power is plotted on the left while the standard deviation for the corresponding day is plotted on the right. The plot on the right therefore shows how much the signal strength varied during the day.

Figures 4.4, 4.5 and 4.6 show the 8-day period measurements for the clusters approximately 20km, 30km and 100km from the receiver respectively. The call signs shown in the legend corresponds to the call sign in Table 3.5. By comparing the plots, it can be seen that a greater path length does not correlate with greater variance in the signal strength and that a daily standard deviation of about 0.2-2dBm can be expected for normal operation.

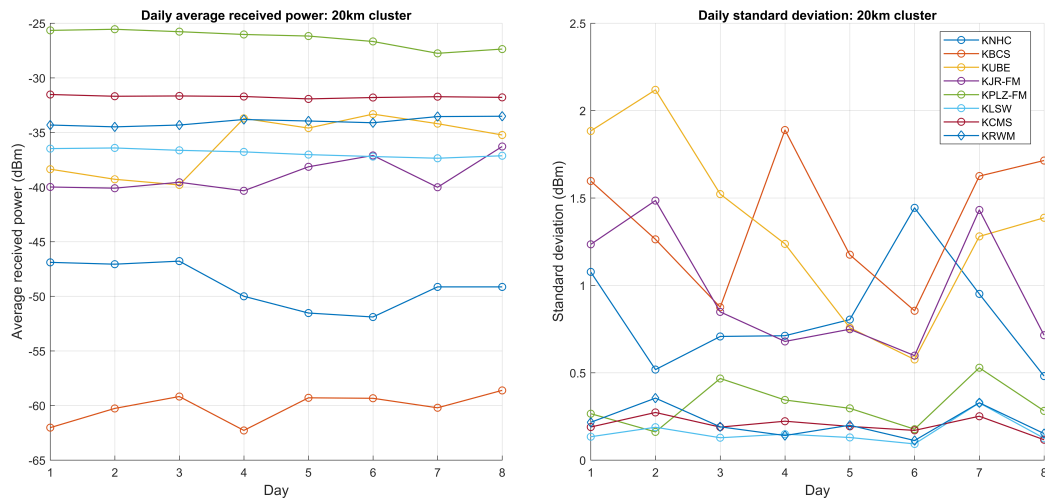


Figure 4.4: Average received power: 20km cluster for 8 days

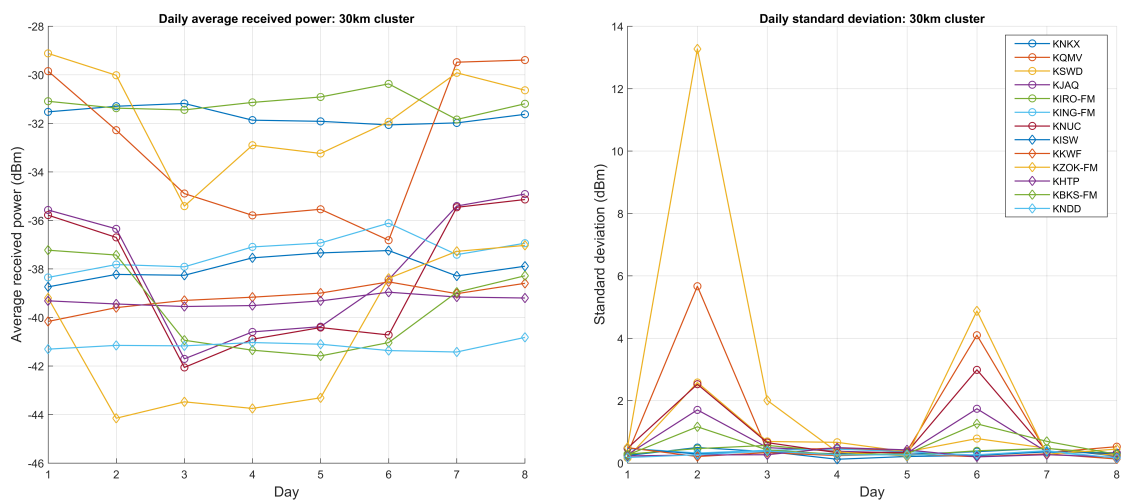


Figure 4.5: Average received power: 30km cluster for 8 days

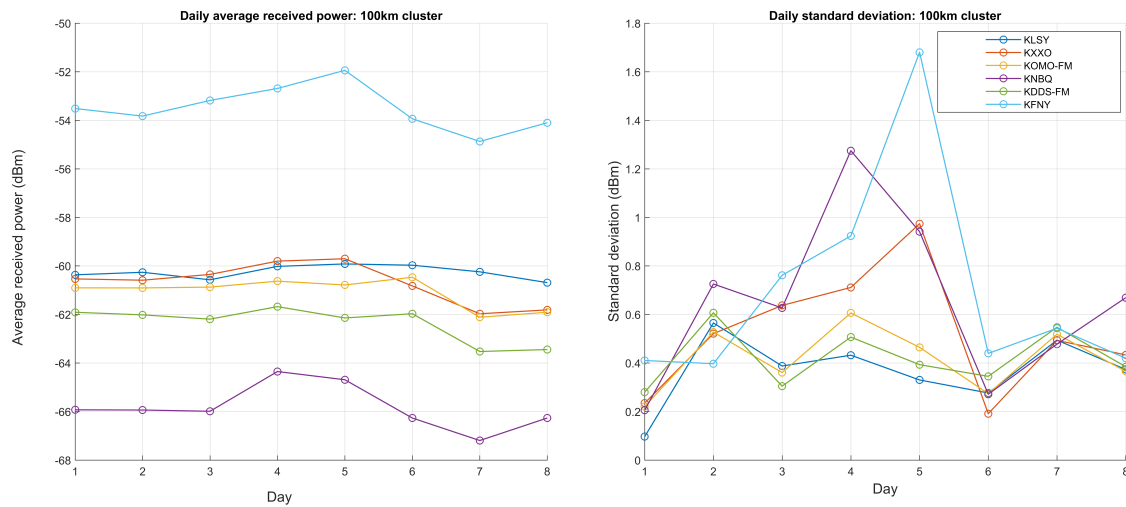


Figure 4.6: Average received power: 100km cluster for 8 days

Figure 4.5 is further analysed to attempt to determine the cause of the large standard deviations from some transmitters on the second and sixth day. The transmitters with the large standard deviation in received power are all located on the same physical tower. It is possible that maintenance on the tower is the cause for the varying signal strengths. Figure 4.7 shows the received power on day 2 from two of the transmitters, call-signs KZOK-FM and KQMV, with the greatest standard deviation. The received power shown on the left is unusually low between 07:00 and 16:00 and on the right between 21:00 and 23:00. Maintenance on the transmitters seems like a plausible explanation for the large variation in the signal strength. There are clearly days where the received power levels are out of the norm and these outlying measurements should be excluded when computing the long term average received powers.

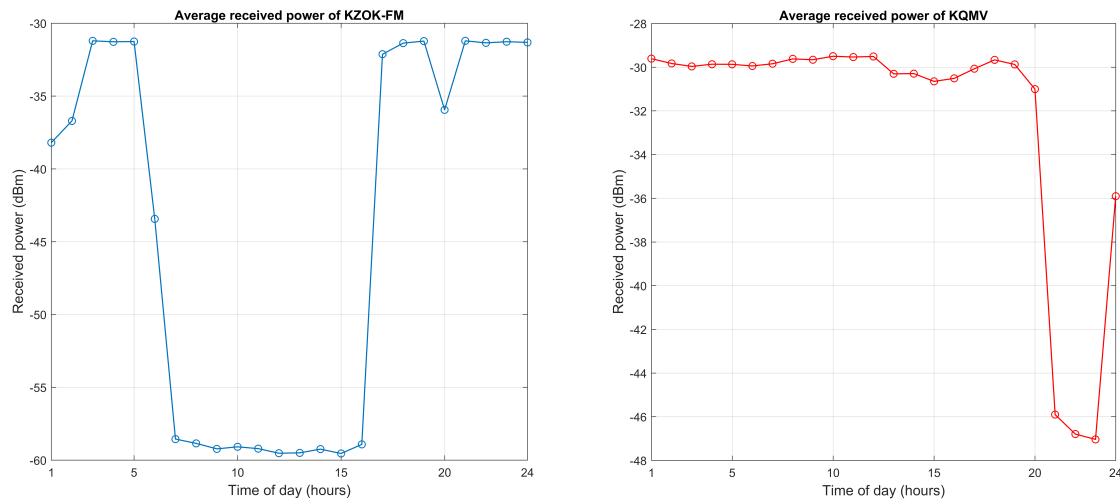


Figure 4.7: Received power (on day 2) from two of the transmitters with the greatest variance in signal strength.

Figures 4.8, 4.9, 4.10 & 4.11, 4.12 and 4.13 show the entire 93-day period measurements for the clusters approximately 4km, 20km, 30km, 40km and 100km from the receiver respectively. The 30km cluster is separated into two sets of clusters and plotted accordingly. The one set contains the transmitters from the tower that had a high standard deviation as discussed before when comparing the measurements of 8 days. The other set contains the remaining transmitters. The plots are less cluttered as a result. Outliers in the data are not yet removed. Figure 4.14 is a box plot of the averaged received power for all the transmitters. The red plus marks indicate the outliers. Apart from the obvious outliers in the data, especially between day 10 and 18 in the 30km cluster, the long term average power seems to be rather consistent.

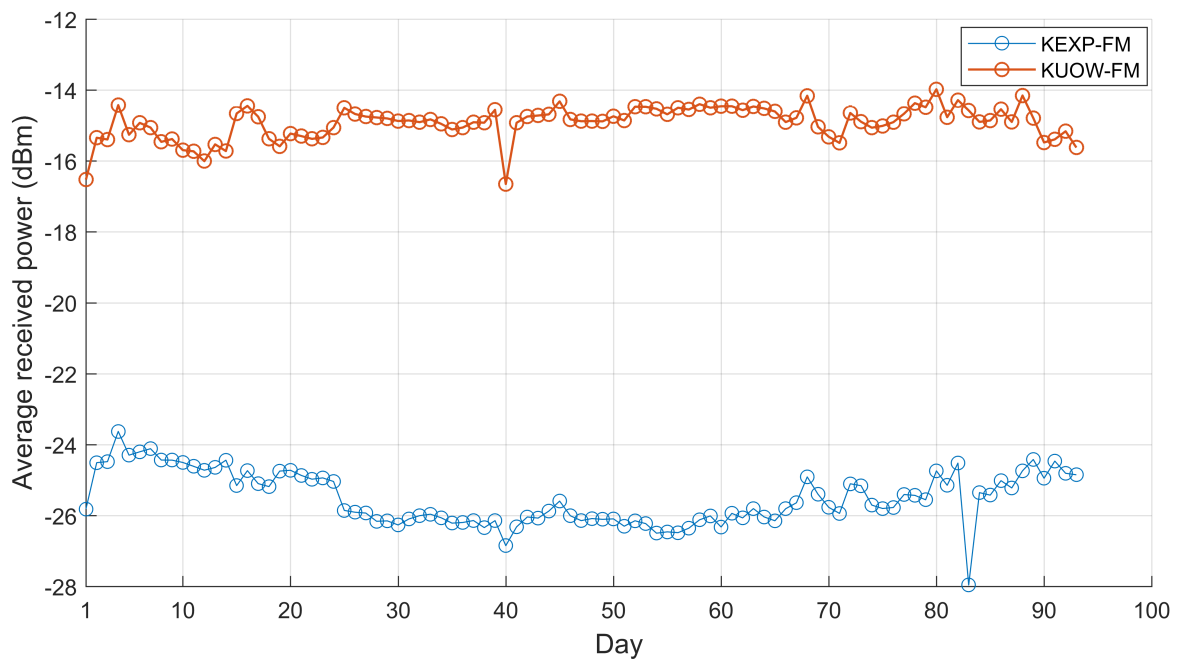


Figure 4.8: Average received power: 4km cluster for 93 days

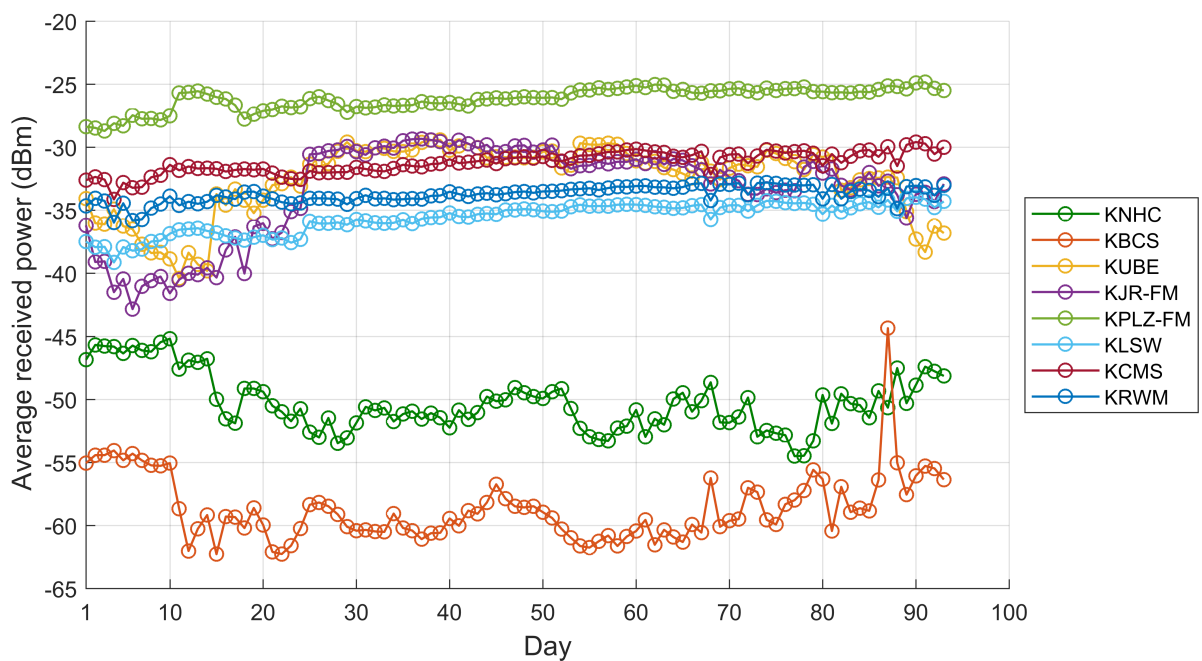


Figure 4.9: Average received power: 20km cluster for 93 days

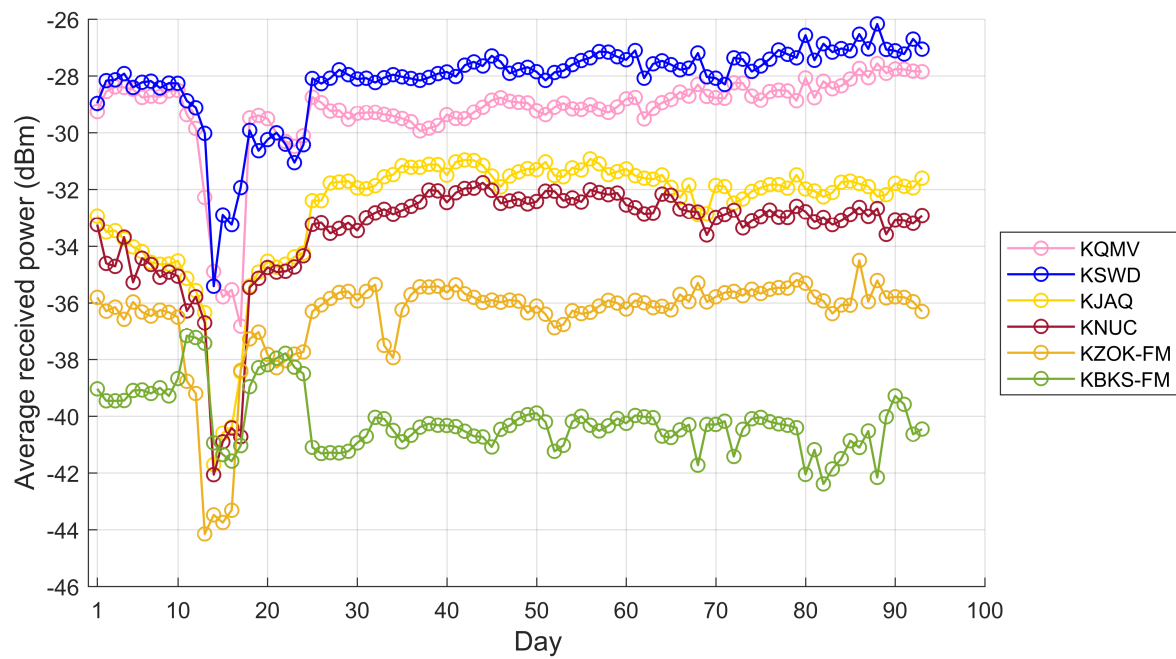


Figure 4.10: Average received power: 30km cluster-1 for 93 days

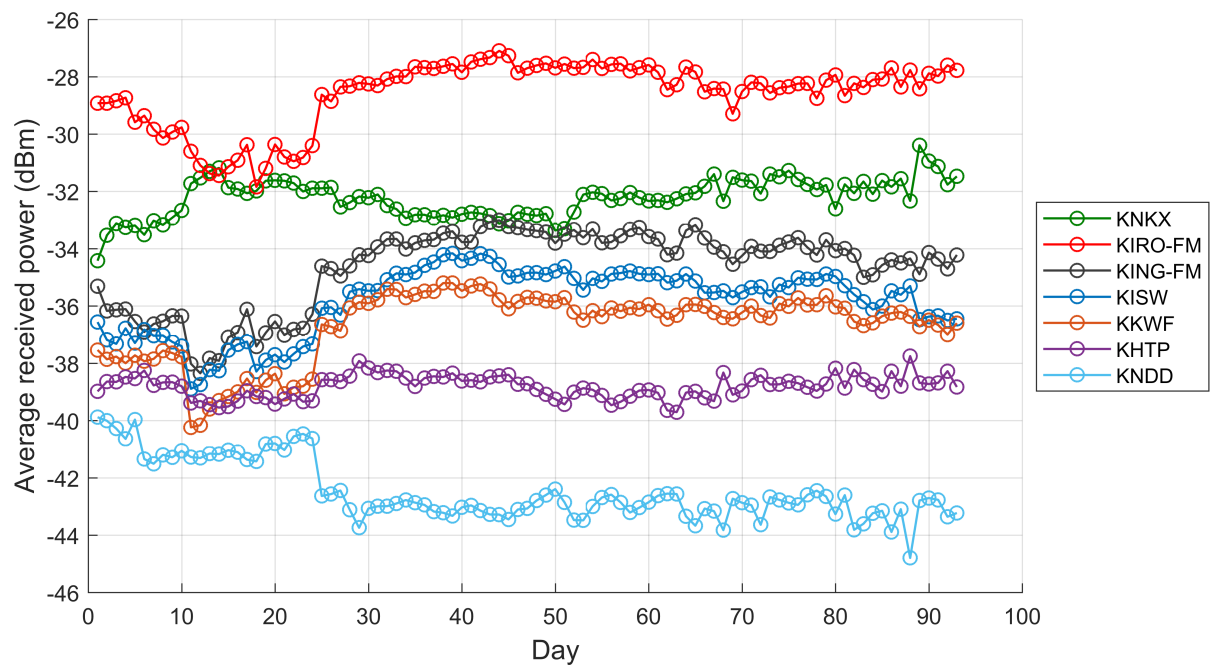


Figure 4.11: Average received power: 30km cluster-2 for 93 days

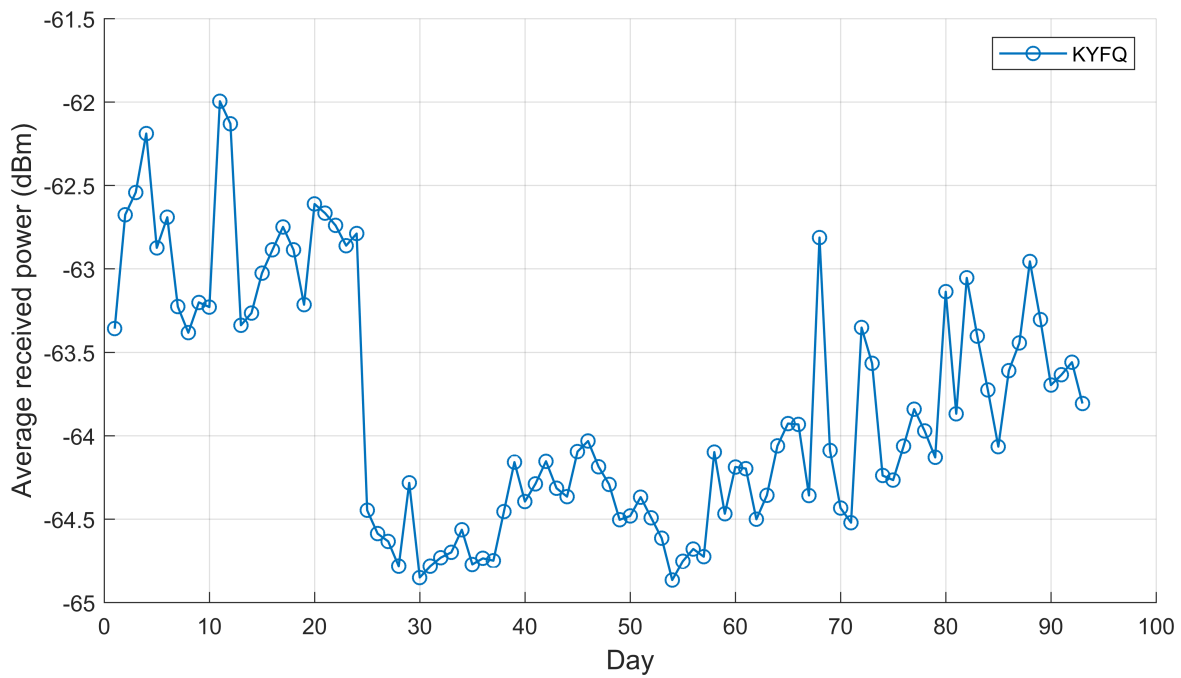


Figure 4.12: Average received power: 40km cluster for 93 days

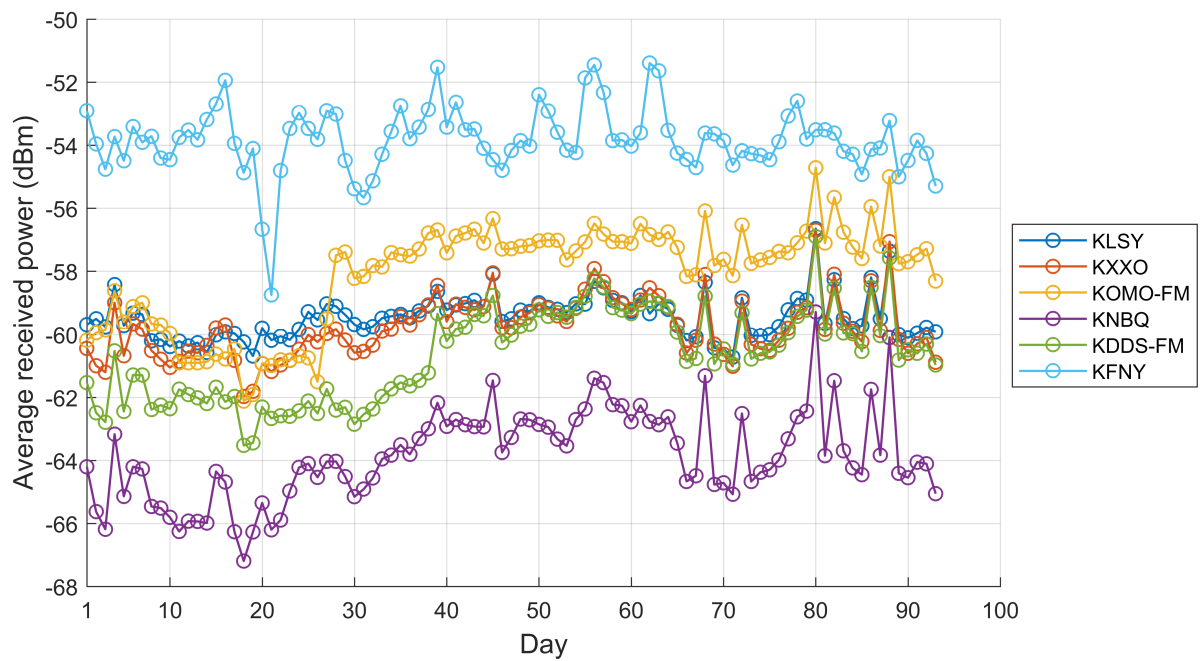


Figure 4.13: Average received power: 100km cluster for 93 days

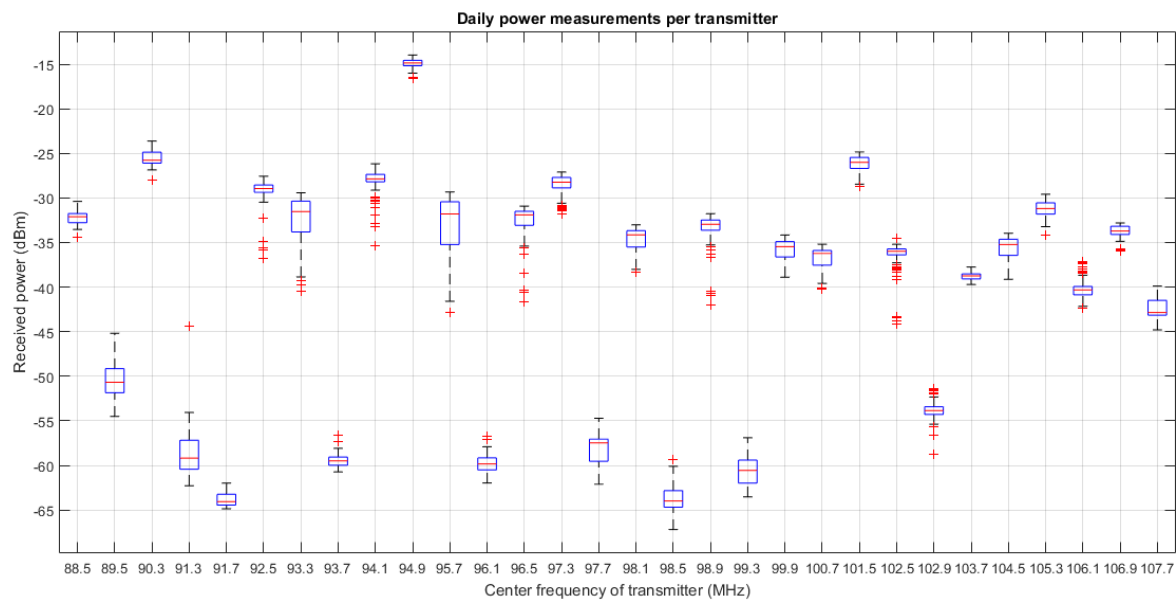


Figure 4.14: Box plot of averaged received power for all transmitters

The long term average received power per transmitter is calculated after the outliers in the measurement data are removed. Table 4.1 summarises the processed measurement results and shows the average received power and standard deviation per transmitter. This is used as the baseline signal strength value at the receiver location.

Table 4.1: Measurement results from UW station

Call sign	Frequency (MHz)	Average power (dBm)	Standard devi- ation (dBm)	Distance to receiver (km)	Distance cluster
KNKX	88.5	-32.2104	0.6321	29.98	30km
KNHC	89.5	-50.2789	2.2363	19.41	20km
KEXP-FM	90.3	-25.5154	0.7537	4.34	4km
KBCS	91.3	-58.8096	2.1878	19.33	20km
KYFQ	91.7	-63.8284	0.7422	39.39	40km
KQMV	92.5	-28.9265	0.5891	30.35	30km
KUBE	93.3	-31.7978	1.7989	19.31	20km
KLSY	93.7	-59.5127	0.5876	88.64	100km
KSWD	94.1	-27.696	0.5065	30.35	30km
KUOW-FM	94.9	-14.8823	0.3995	4.34	4km
KJR-FM	95.7	-32.1808	2.2246	19.31	20km
KXXO	96.1	-59.8151	0.9042	98.44	100km
KJAQ	96.5	-31.8224	0.5676	30.35	30km
KIRO-FM	97.3	-28.2799	0.7511	29.98	30km
KOMO-FM	97.7	-57.5343	1.0192	88.64	100km
KING-FM	98.1	-34.2557	0.9006	29.98	30km
KNBQ	98.5	-63.8371	1.4495	98.23	100km
KNUC	98.9	-33.0424	0.826	30.35	30km
KDDS-FM	99.3	-60.6616	1.4488	88.64	100km
KISW	99.9	-35.6733	0.9881	29.98	30km
KKWF	100.7	-36.271	0.6427	29.98	30km
KPLZ-FM	101.5	-26.1315	0.8604	19.26	20km
KZOK-FM	102.5	-35.964	0.4368	30.35	30km
KFNY	102.9	-53.8656	0.6919	98.23	100km
KHTP	103.7	-38.81	0.4011	29.98	30km
KLSW	104.5	-35.5719	1.1329	19.41	20km
KCMS	105.3	-31.1995	0.8112	19.31	20km
KBKS-FM	106.1	-40.3915	0.7229	30.35	30km
KRWM	4106.9	-33.681	0.532	19.33	20km
KNDD	107.7	-42.8651	0.536	29.98	30km

4.2 Land cover

The results of the processed land cover data are analysed and discussed in this section. The generated grids containing land cover class information are explained. The proportion of land cover in the paths is analysed and discussed.

The processed land cover data as discussed in section 3.3, are stored in four MATLAB matrices, one for each tile within the established bounds. Table 4.2 shows the bounds of the four tiles. Each of the four matrices has 1200 rows and 1200 columns. Every entry is an integer that represents a land cover class. The resolution of the land cover is 90m, meaning every pixel represents an area of 90mx90m.

Table 4.2: Land cover tile bounds

Tile	Latitude		Longitude	
	North	South	East	West
A	48	47	-123	-124
B	48	47	-122	-123
C	48	47	-121	-122
D	47	46	-123	-124

Figure 4.15 is a representation of a land cover tile situated around Seattle. The colour map corresponds with the type of land cover at the geographical location. Figure 4.16 shows the proportion of land cover that is present in all the paths combined. A path is defined as the path that the EM waves travel from a transmitter to a receiver. The water, developed and forest land cover classes dominate the proportion of land cover present in all the paths combined.

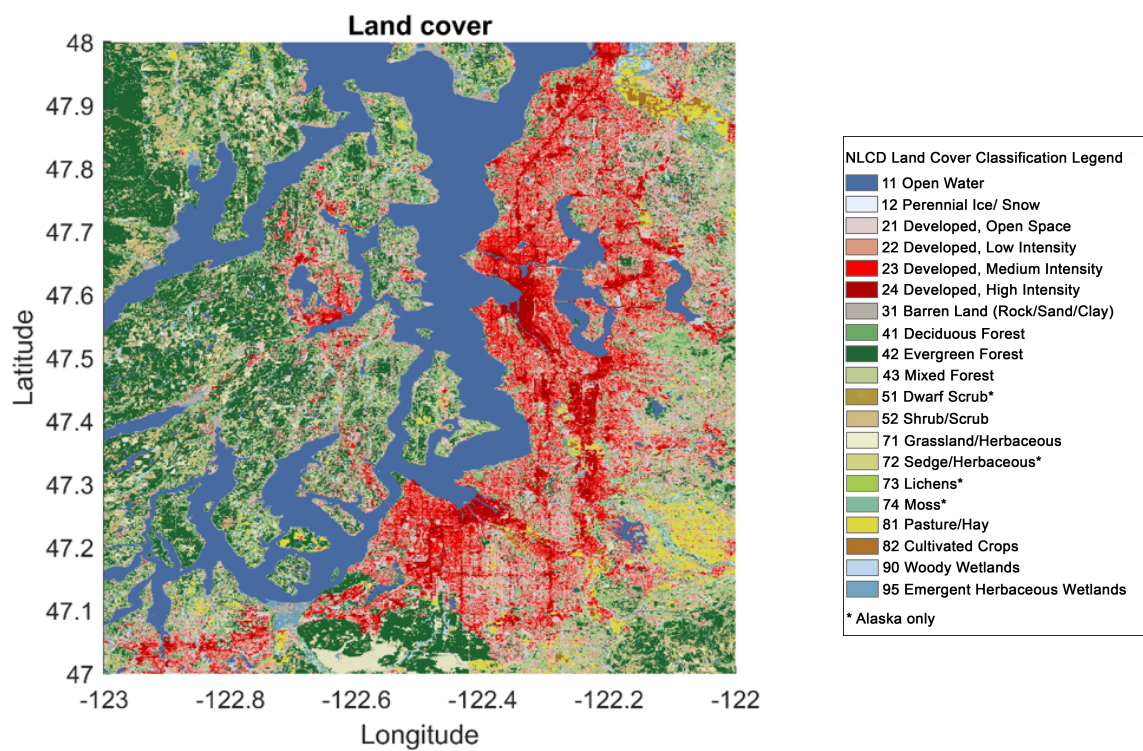


Figure 4.15: Example of a land cover tile over a region of interest

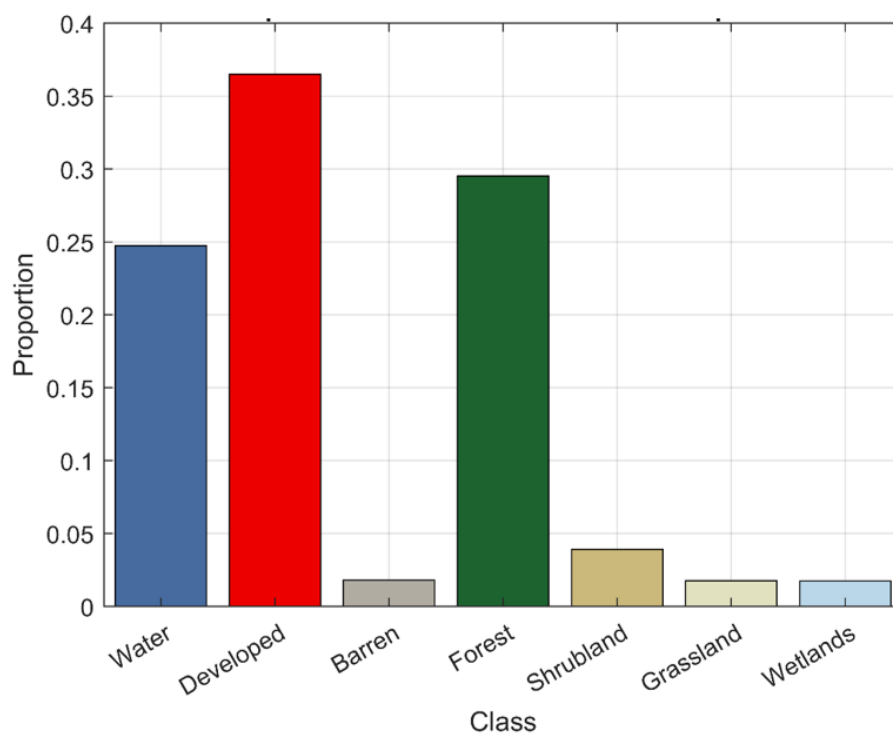


Figure 4.16: Proportion of land cover classes in all paths for UW station

Figure 4.17 shows the proportion of land cover subclasses that are present in all the paths combined. The land cover subclasses are more detailed than the land cover classes from Figure 4.16. For example the forest land cover class is comprised of classes 41, 42 and 43, which are different types of forest.

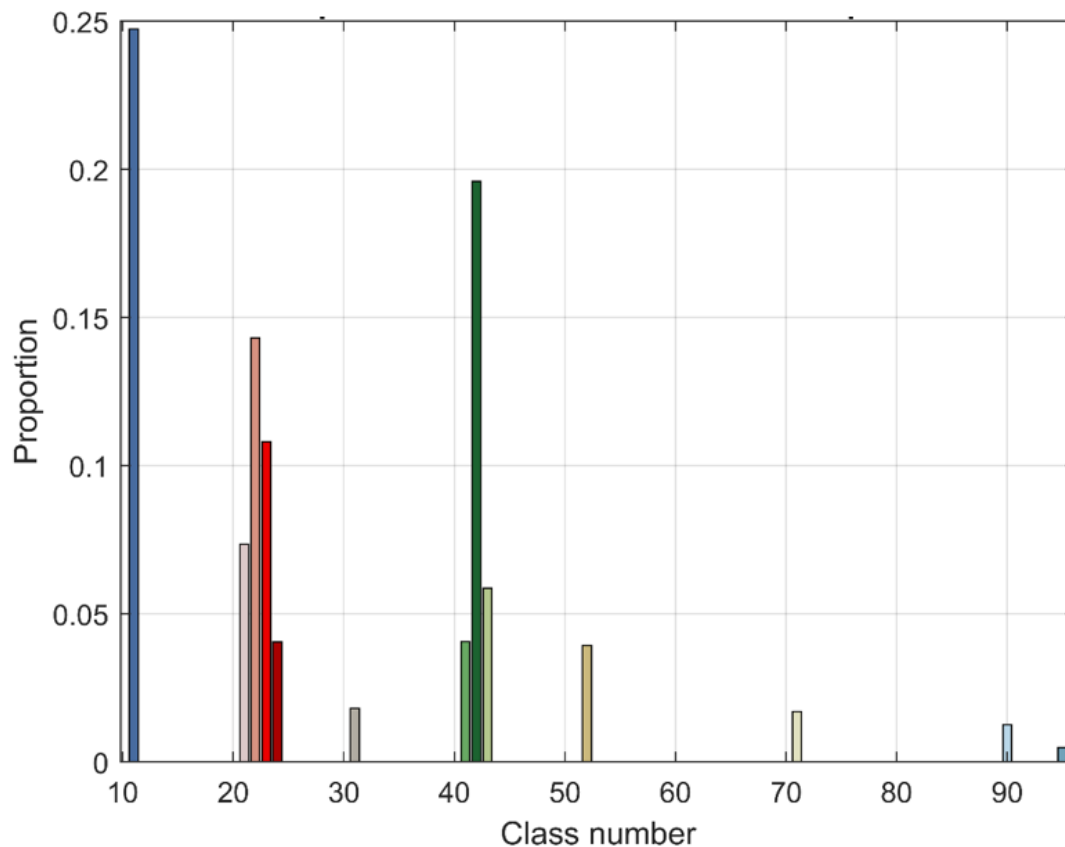


Figure 4.17: Proportion of land cover subclasses for unique paths to the UW station

There are eleven unique paths in terms of the combination of land cover classes on the path. That is there are eleven unique geographically located towers, some having more than one transmitter. What makes the radio paths unique for transmitters on the same tower are the different heights of the antennas. Figure 4.18 shows the proportion of land cover for the unique paths. The shorter paths are dominated by developed land cover and the longer paths by forest and water.

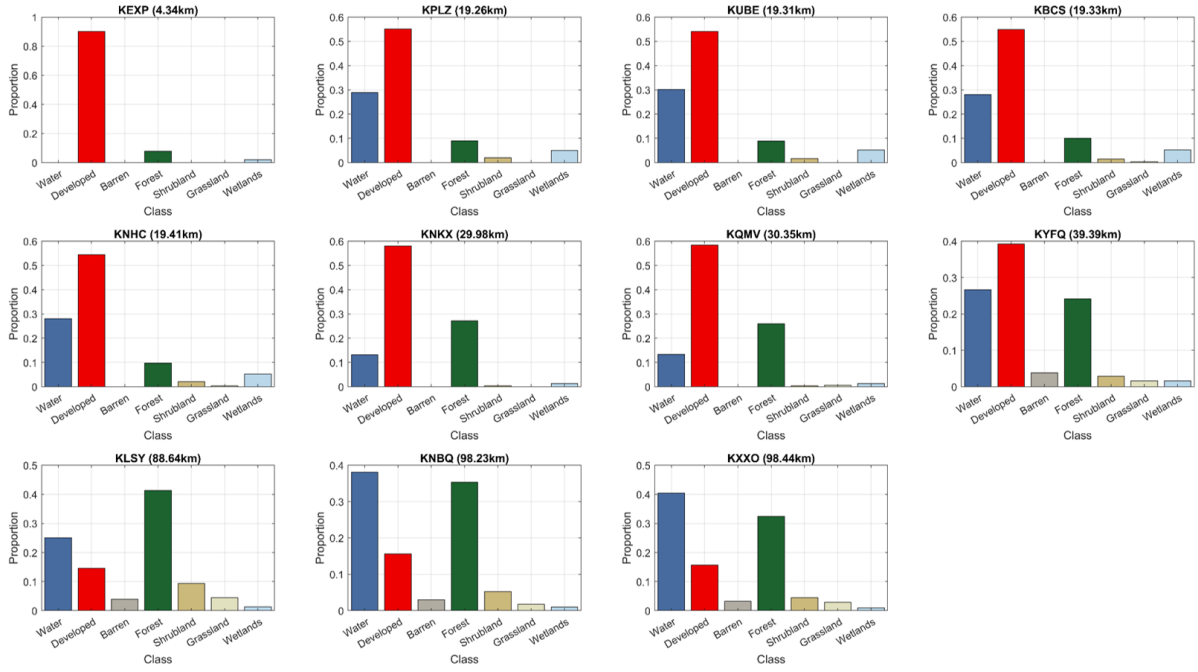


Figure 4.18: Proportion of land cover classes in all paths for UW station

For the distances considered, a path can be approximated by drawing a straight line from a transmitter to the receiver. The path is computed in MATLAB using the function, *track2*. The WGS84 reference ellipsoid is used. In MATLAB it is specified as *wgs84Ellipsoid*. Using the WGS84 reference ellipsoid means that the path between two points follows an arc on the surface of the earth. For small distances it can be approximated as a straight line. Given a land cover grid and a reference vector, a land cover value is extracted approximately every 50m starting at the transmitter and moving in a straight line (on the arc) to the receiver. If two values are extracted from the same pixel, the second one is discarded. Therefore every pixel is only counted once. Basically, if the path goes through the pixel it is counted.

4.3 Reducing the error

In this section the results from the method of reducing the error are analysed and discussed. The results from the three approaches of using different initial heights or start-

ing points are given, followed by the final results that produced a better decrease in error.

As discussed in section 3.4 and specifically section 3.4.4, the set of 30 transmitters are randomly split into five disjoint subsets of six transmitters each so that 5-fold cross-validation could be performed. The fixed subsets are used for all the training and validation that follows.

4.3.1 ITM

The RMSE between the measured and predicted power is the metric chosen to be minimised because larger errors are penalised more. The mean absolute error (MAE), mean, standard deviation and correlation coefficient are also calculated for both the training and validation sets for every iteration.

Table 4.3 displays all the initial predicted values for each would-be iteration before the method of reducing the error is implemented. This is the result of using the unmodified SDF files as input for SPLAT!. The iteration determines which subset is used for validation. Subset one is used for validation in iteration one, and so forth up to iteration five. The standard deviation is the deviation of the error.

Table 4.3: Initial parameters before the method is implemented for ITM

Iteration	RMSE (dBm)		MAE (dBm)		Mean Error (dBm)		STD (σ) (dBm)		Correlation	
	Train	Validate	Train	Validate	Train	Validate	Train	Validate	Train	Validate
1	11.8552	14.0897	10.3842	12.8677	-10.3842	-12.8677	5.8425	6.2873	0.9165	0.8064
2	12.5780	11.3084	11.0472	10.2156	-11.0472	-10.2156	6.1429	5.3127	0.8994	0.8377
3	13.0520	8.9041	11.5219	8.3170	-11.5219	-8.3176	6.2638	3.4831	0.8855	0.9823
4	12.4503	11.8599	11.1348	9.8655	-11.1348	-9.8655	5.6900	7.2108	0.9073	0.8625
5	11.6869	14.6410	10.3165	13.1387	-10.3165	-13.1387	5.6095	7.0767	0.9001	0.8910

Initial height parameter

The discussion that follows will make heavy use of class numbers when explaining the results. Refer to Table 2.5 for information regarding the type of land cover associated with a class number. This will assist in understanding the results. In short, classes 21 to 24 correspond with the various degrees of developed land cover. Classes 41 to 43 are grouped together and correspond with the forest land cover. Values given in a matrix, e.g. $[a \ b \ c \ d \ e]$, should be interpreted as follows. The height given by a is the associated height of the land cover class 21. The height given by b is the associated height of the land cover class 22 and so forth. The height given by e is the associated height of the land cover classes 41 to 43 which are grouped together.

Three approaches for finding initial heights for the land cover classes were discussed in section 3.4.5. The first approach was to guess initial heights based on the results of similar research. The initial heights for the land cover classes were chosen as follows. 1m for classes 21 and 22, 2m for class 23, 5m for class 24 and 25m for classes 41 to 43. The result was that the heights of classes stayed constant except for class 23 which was reduced to 0m and classes 41 to 43 of which the height increased to 42m. Only one class seemed to have a significant effect on the path loss which is unlikely. The reduction in the error was around 1.5dBm. The reasons that these heights do not give satisfactory results can be because a different propagation model is used than in the literature as well as a different land cover database. Because the geographic locations of the studies in the literature and this study differ, the land cover most likely also differ.

From the first approach it was seen that only one iteration of the 5-fold CV could be done to determine suitable initial heights since the reduction in the error is similar for different iterations. For the second approach the initial heights were set to zero. A step size of 4m was initially used and the process continued until a step size of 1m was finally used. Only the height of class 22 increased with this approach. It is possibly because the developed land cover is located in closer proximity to the receiver than the other classes. Increasing the height of those classes most likely increases the obstacles in the first Fresnel zone resulting in greater path loss which in turn reduces the error

between the measurements and predictions. The reduction in error happened very slowly per iteration. The final height of class 22 was 18m while the other heights were zero except for class 24, which height was 3m. The reduction in RMSE for the training set was 0.2dBm. Keeping the hypothesis in mind the model was incorrectly fitted.

The opposite was done and big initial heights were chosen. 10m for classes 21 and 22, 20m for class 23, 30m for class 24 and 50m for classes 41 to 43. Initially the error was much larger than the baseline error. The important result from this approach was not the reduction in error achieved but rather that relative upper limits for the initial heights were obtained. The RMSE for the training set was reduced to 9.3514dBm. By supervising the method and observing the height changes as the error was reduced the following was seen. The height for class 21 tended to reduce in the first iteration and didn't have an effect after that. It was also the case for class 23. Classes 41 to 43 also reduced to a height of around 40m and had limited effect after that. Classes 22 and 24 seemed to have the most influence on the method. Class 22 varied between 5 and 7m while class 24 varied between 26 and 30m.

The knowledge from the second approach is used in the third approach. The third approach is to start at random initial heights. The random heights are restricted to be between specified values for each class. It could be between 0m and 10m for classes 21 to 23, between 10m and 35m for class 24 and between 20m and 40m for classes 41 to 43. Tables A.1, A.2 and A.3 in Appendix A show the results for tuning the ITM when various random starting points or initial heights are used. These random initial heights did not provide the best improvement of all the approaches.

Using the results from the third approach and choosing initial heights accordingly, the following results were obtained. As was the case with the previous attempts, some classes had a limited effect on path loss. After a local minimum was reached those classes were manually tuned to attempt to find an even greater improvement. Table 4.4 displays all the final values for each iteration after the method is implemented. The metrics are the same as in Table 4.3 with addition to the heights obtained. The heights, given as a vector and separated by a space, are given per land cover class. The

classes corresponding to the heights from left to right are classes 21, 22, 23, 24, and 41 to 43 grouped together. Classes 21 to 24 are the various developed classes, from least developed to most developed. Classes 41 to 43, which are grouped together, are the forest classes.

Table 4.4: Final parameters after the method is implemented for ITM

Iteration	Heights (m)	RMSE (dBm)		MAE (dBm)		Mean Error (dBm)		STD (σ) (dBm)		Correlation	
		Train	Validate	Train	Validate	Train	Validate	Train	Validate	Train	Validate
1	[1 2 2 23 40]	8.2304	5.9002	6.1419	9.6864	-4.6967	-8.1877	6.9041	9.3089	0.8945	0.3705
2	[1 7 5 28 40]	9.4709	3.2362	7.1108	5.6490	-5.1793	-3.9023	8.0997	5.6565	0.8583	0.8773
3	[1 6 2 27 10]	9.0762	4.3551	6.8035	7.9623	-5.1915	-2.9120	7.6050	8.9925	0.8533	0.9587
4	[1 7 5 28 35]	8.8830	4.6275	7.3013	4.8946	-5.2381	-3.9088	7.3286	9.1898	0.8871	0.7499
5	[2 7 5 28 30]	9.1604	4.0612	6.8522	6.6960	-4.9760	-5.1287	7.8565	6.8994	0.8377	0.9159

The results are summarised in Table 4.5. The average of the five iterations for each metric is taken. The average is taken for the baseline results as well as the results from the method of reducing the error. The difference is also calculated. The RMSE is reduced by 7.7299 dBm, the MAE by 3.9032 dBm and the mean error by 6.073 dBm. The MAE is greater than the mean error meaning that with the augmented propagation model the path loss is sometimes over predicted. For the baseline case, the path loss was consistently under-predicted. The consequence of over predicting the path loss is possible interference for users outside of the the predicted coverage area. The standard deviation increased by 2.1353 dBm and the correlation decreased by 0.1006. There is still a moderate linear correlation between the measured and predicted signal powers. The correlation is only very weak for the fifth iteration but strong for the rest.

Table 4.5: Evaluation set results for ITM

Metric	Augmented		Difference (Baseline - Augmented)
	Baseline	with land cover	
RMSE (dBm)	12.166	4.4361	7.7299
MAE (dBm)	10.8809	6.9777	3.9032
Mean Error (dBm)	-10.8809	-4.8079	-6.0730
Standard deviation (dBm)	5.8741	8.0094	-2.1353
Correlation	0.8760	0.7754	0.1006

The heights are also averaged and the result is shown in Table 4.6. The DEM should be augmented with the corresponding land cover class information and accompanying heights.

Table 4.6: Final heights per land cover class for ITM

Class type	Class number	Height (m)
Developed - Open Space	21	1
Developed - Low intensity	22	6
Developed - Medium intensity	23	4
Developed - High intensity	24	27
Forest	41 to 43	31

4.3.2 ITWOM

Table 4.7 displays all the initial values for each would-be iteration before the method of reducing the error is implemented. This is the result of using the unmodified SDF files as input for SPLAT!.

Table 4.7: Initial parameters before the method is implemented for ITWOM

Iteration	RMSE (dBm)		MAE (dBm)		Mean Error (dBm)		STD (σ) (dBm)		Correlation	
	Train	Validate	Train	Validate	Train	Validate	Train	Validate	Train	Validate
1	9.8511	16.3385	8.1665	14.9304	-8.1665	-14.9304	5.6279	7.2689	0.8976	0.8966
2	11.7382	10.1972	9.7827	8.4655	-9.7827	-8.4655	6.6268	6.2247	0.8690	0.8931
3	12.4099	6.2534	10.5973	5.2070	-10.5973	-5.2070	6.5967	3.7933	0.8725	0.9805
4	11.8539	9.6469	9.8747	8.0973	-9.8747	-8.0973	6.6989	5.7443	0.8729	0.8019
5	11.2149	12.3300	9.1750	10.8960	-9.1750	-10.8960	6.5880	6.3219	0.8875	0.8232

The method of reducing the error is performed using the ITWOM as well. Learning from the results of reducing the error of the ITM, initial heights are estimated close to the final heights or output of the method for reducing the error with the ITM. The ITWOM behaves much in the same way as the ITM in the sense that there are two land cover types that have a significant impact on reducing the error. The approach to find a minimum in the error is therefore augmented with additional information to choose good initial heights. Three of the five land cover classes that did not have a significant effect on the error were held constant, while the two land cover classes that did have a significant effect, were varied within reasonable bounds. The result is plotted to visualise the search space and find good initial heights. Figure 4.19 shows the RMSE between the measured and predicted signal power when two land cover classes' heights are varied. It can be seen from the figure, that there is a valley where the RMSE is low. With this information, good initial heights for the land cover classes can be chosen. All 30 transmitters were used to make the plot for this initial step. The gradient descent method, as described in section 3.4.3, is used to reduce the error when using the ITWOM for predictions. The above mentioned initial heights for land cover classes are also used. All of the land cover classes used for reducing the error for the ITM is also used for the ITWOM. The two highlighted classes were for illustration purposes and to find good initial height for classes that made a major difference.

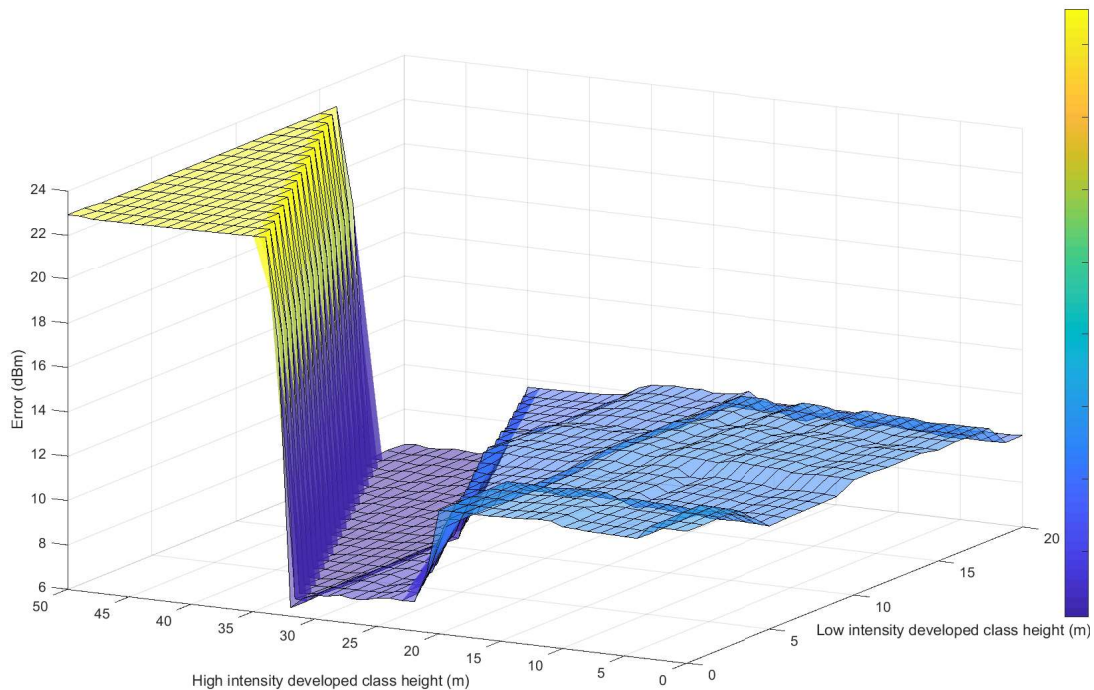


Figure 4.19: Comparison of RMSE between measured and predicted signal power when varying two land cover classes' heights.

Table 4.8 shows the values after the error is reduced. The averaged results are summarised in Table 4.9. The average of the five iterations for each metric is taken. The average is taken for the baseline results as well as the results from the method of reducing the error. The difference is also calculated. For the case of augmenting the ITWOM with land over data, the RMSE was reduced by 7.8347 dBm, from 10.9532 dBm to 3.1185 dBm. The MAE was reduced from 9.5192 dBm to 5.4327 dBm. The mean was significantly reduced and is almost zero. The mean is 0.2664 dBm. The big difference between the mean error and the MAE shows that the ITWOM over-predicts the path loss more often. The standard deviation is slightly worse after the method to reduce the error is applied. The standard deviation is 0.234 dBm more, and is 6.1046 dBm. The correlation coefficient between the measured and predicted signal powers increased from 0.8791 to 0.901. This means that the predicted signal powers are a better fit to the measured signal powers after the method is applied.

Table 4.8: Final parameters after the method is implemented for ITWOM

Iteration	Heights (m)	RMSE (dBm)		MAE (dBm)		Mean Error (dBm)		STD (σ) (dBm)		Correlation	
		Train	Validate	Train	Validate	Train	Validate	Train	Validate	Train	Validate
1	[2 4 0 35 16]	6.4397	2.9844	5.5300	4.5654	-0.2790	2.2363	6.5720	5.3468	0.9027	0.9863
2	[2 6 0 38 18]	6.1806	3.2447	5.1583	6.0289	0.6677	-1.3505	6.2766	6.9531	0.9384	0.9049
3	[3 9 0 40 16]	6.4498	2.6758	5.4994	4.994	-0.3302	2.3046	6.5799	5.2908	0.9150	0.9922
4	[5 4 0 35 17]	6.0829	3.4167	5.1575	6.0554	-0.3731	2.6127	6.2020	6.9168	0.9468	0.8520
5	[2 4 0 35 17]	6.0115	3.5407	5.2915	5.5197	1.3979	-4.4710	5.9725	6.0156	0.9472	0.7696

Table 4.9: Evaluation set results for ITWOM

Metric	Augmented		Difference (Baseline - Augmented)
	Baseline	with land cover	
RMSE (dBm)	10.9532	3.1185	7.8347
MAE (dBm)	9.5192	5.4327	4.0865
Mean Error (dBm)	-9.5192	0.2664	-9.7856
Standard deviation (dBm)	5.8706	6.1046	-0.234
Correlation	0.8791	0.9010	-0.0219

The heights are also averaged and the result is shown in Table 4.10. The DEM should be augmented with the corresponding land cover class information and accompanying heights in order to improve the prediction accuracy.

Table 4.10: Final heights per land cover class for ITWOM

Class type	Class number	Height (m)
Developed - Open space	21	3
Developed - Low intensity	22	6
Developed - Medium intensity	23	0
Developed - High intensity	24	37
Forest	41 to 43	17

The ITWOM behaved similarly to the ITM when the model was fitted. There were two dominant land cover classes that had the most effect on the predicted path loss. The baseline predictions of the ITWOM is more accurate than that of the ITM. The RMSE between field strength measurements and propagation predictions with the ITWOM is 10.9532 dBm while that of the ITM is 12.166 dBm. The standard deviations are similar although for this study the ITWOM just outperforms the ITM. The same applies to the correlation.

After the models have been augmented with land cover data, the ITWOM showed better improvements than the ITM. The RMSE for both models were reduced by around 7.7 dBm with the ITWOM outperforming the ITM again. The RMSE parameter for the ITWOM was 3.1185 dBm while it was 4.4361 dBm for the ITM. The standard deviation for the ITWOM was also better than that of the ITM, being 6.1046 dBm as opposed to 8.0094 dBm. More importantly, the standard deviation did not increase as much for the ITWOM as it did for the ITM. The ITWOM predictions also fit the measured data better and showed an improvement in correlation. The ITM showed a deterioration in the correlation coefficient. The ITWOM predictions still had a strong linear correlation with the measurements while the ITM predictions had a moderate linear correlation with the measurements. The results seem to indicate that the ITWOM is better suited to be augmented with land cover data than the ITM.

4.4 Conclusion

The result of the measurement study is an average received power measurement per transmitter over a period of 93 days, with the outliers removed. The standard deviations of the measurements are small, with the majority being under 1 dBm.

The results of the processed land cover data are grids containing land cover classes to augment the digital elevation model with. From analysing the proportion of land cover in the paths the dimensionality of the problem was reduced.

The measurements and land cover data were necessary to reduce the error between measured and predicted path loss. A gradient descent method was used to reduce the RMSE between measured and predicted power levels. 5-fold cross-validation was done, and the result of the average of the five validation sets was tabulated. Multiple initial heights were used in search for local minimums that provide a good decrease in error. Due the dimensionality of the problem the entire space was too big to search. The search for minimums was supervised, and multiple local minimums were found. By supervising the process, the interaction between the variables could be observed as the error was reduced. This led to better chosen initial heights that resulted in a better reduction in error.

For the case of using the ITM, the RMSE was reduced from 12.166 dBm to 4.4361 dBm. That is a decrease of 7.7299 dBm. The standard deviation did, however, increase by 2.1353 dBm. There is still a moderate linear correlation between the measured and predicted received power even though the correlation coefficient decreased. The ITWOM performed better, and the RMSE was reduced from 10.9532 dBm to 3.1185 dBm. That is a decrease of 7.8347 dBm. The standard deviation only increased by 0.234 dBm from 5.8706 dBm to 6.1046 dBm. The correlation coefficient improved to 0.901 resulting in a strong linear correlation between measurements and predictions using the ITWOM.

To ensure that the field strength measurements were processed correctly, the land cover data were extracted and manipulated correctly, and the gradient descent method was implemented correctly, all three aspects of the process are verified and discussed in the next chapter. In chapter 5, the trained model is validated on an independent evaluation set of paths between the transmitters and a different receiver that weren't used in the training of the model.

Chapter 5

Verification and validation

This chapter starts with the verification of field measurements. The process of extracting received signal power from the data obtained from the spectrum observatories is verified. Then the process of manipulating the land cover data is verified to ensure that the manipulated land cover data's geographical locations and class values match those of the original source. Verification is also done to ensure that the DEM is augmented correctly. The implementation of the gradient descent method is then verified by means of an example. Finally, validation is done on the augmented models to determine if the models can be used on new propagation paths, other than those for which it was trained.

5.1 Verification

5.1.1 Introduction

The research project can be divided into three main parts that have to be implemented correctly. For the measurement study, verification is done by confirming that the obtained data is interpreted correctly. For the land cover data, verification is done to

ensure that the land cover heights are added at the correct geographic locations. For the method of reducing the error, the gradient descent method is verified by means of an example. The output of the method is validated with an additional evaluation set of measurements and predictions. The validation technique that is used is known as historical data validation [89]. The additional set of measurements can be considered as part of the data that was used to build the model but kept aside specifically for testing the model. Cross-validation on the set of measurements used to train the model not only increases the number of training samples but also acts as validation for the model.

5.1.2 Measurements

To ensure that the extracted measurement data is interpreted correctly, the process is visually verified by plotting the extracted data in MATLAB and comparing it to the plot generated by the CityScape website. The PSD plots are plotted for the same date and time to be comparable. Figure 5.1 shows a screen capture of the PSD plot from the CityScape website. Figure 5.2 shows the PSD plot generated in MATLAB.

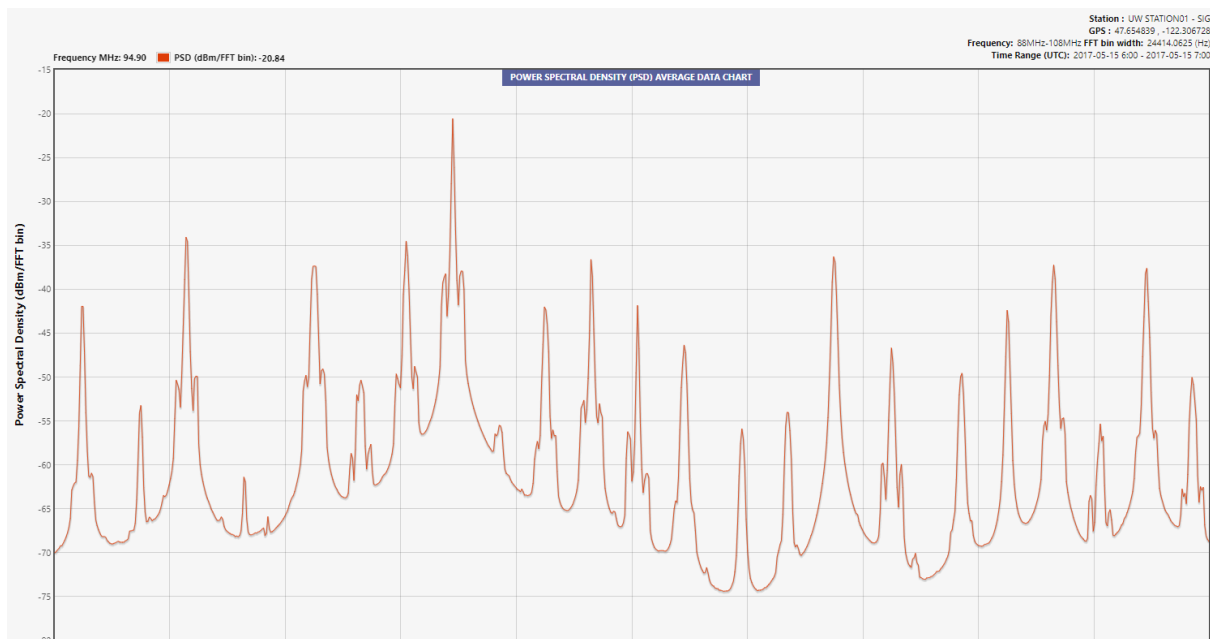


Figure 5.1: PSD plot from CityScape website

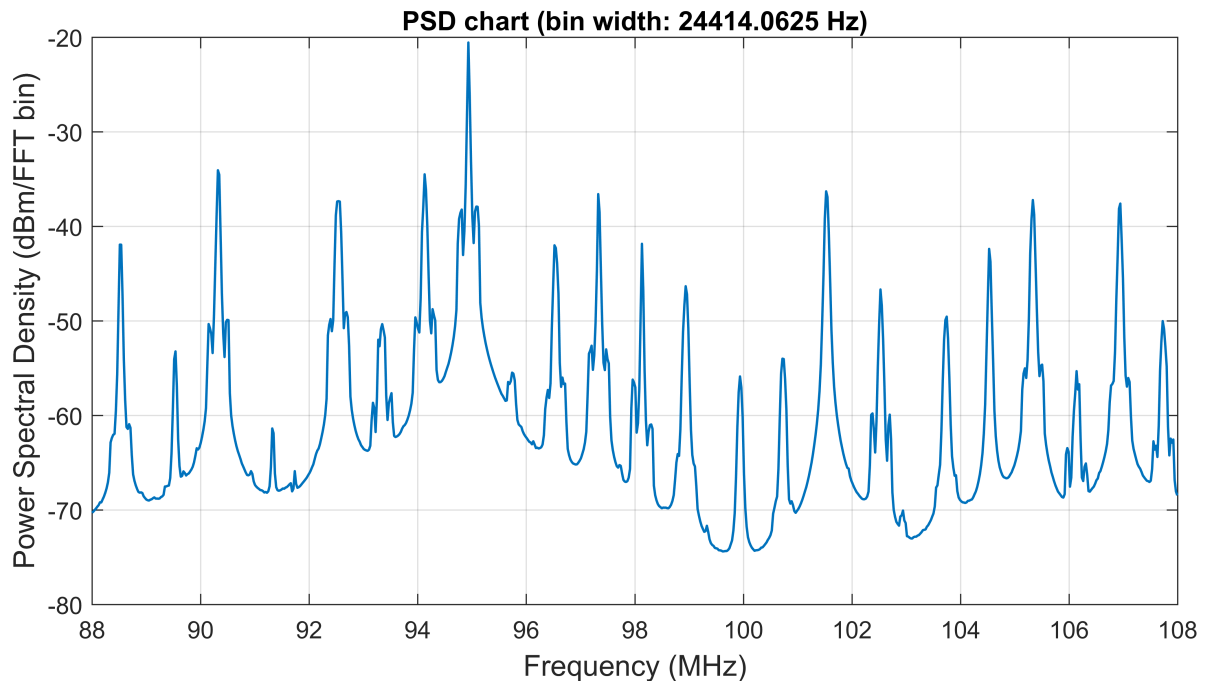


Figure 5.2: PSD plot generated in MATLAB

The PSD plots look exactly the same, and after closer inspection, the power levels are also exactly the same. This verifies that the measured data is interpreted and processed correctly up to the point where the PSD could be plotted. The bin size is verified which is an important parameter in extracting the average power from the PSD.

5.1.3 Land cover

Three aspects of processing the land cover data are verified. The first aspect that is verified is that the land cover class values are extracted correctly from the NLCD and that it is represented correctly in MATLAB. The second aspect to verify is that the DEM data (SDF file) is read and represented correctly in MATLAB. The third aspects to verify is that the function of writing new SDF files is correct.

Land cover class values

To ensure that the process of extracting land cover classes from the NLCD was implemented correctly, samples from the land cover grid in MATLAB is tested against the land cover data in QGIS. Ten thousand random samples from one land cover tile generated in MATLAB, are extracted. Each sample contains a latitude, longitude and class value. The samples are saved as a CSV file. The CSV file is used to create a new point layer in QGIS and the point sampling tool is used to extract the class values at the given coordinates. The class values for the two sets are then compared. The resolution of the land cover data in QGIS is still 30m.

The land cover data in MATLAB is downsampled from a 30m-resolution to a 90m-resolution using nearest-neighbour interpolation. Therefore the random samples are limited to coordinates that are multiples of 3-arc seconds and within the bounds of the one land cover tile. If any random coordinate is used, chances are that the land cover value will be different for the set used in QGIS when compared to that used in MATLAB, due to the resolution difference between the two sets.

Figure 5.3 shows the random samples overlaid on the land cover dataset in QGIS. All of the samples from MATLAB exactly matches all the samples from QGIS. The land cover data is therefore extracted correctly and the reference vector of the land cover grid in MATLAB is also hereby verified, meaning that the location of the land cover classes is correct.

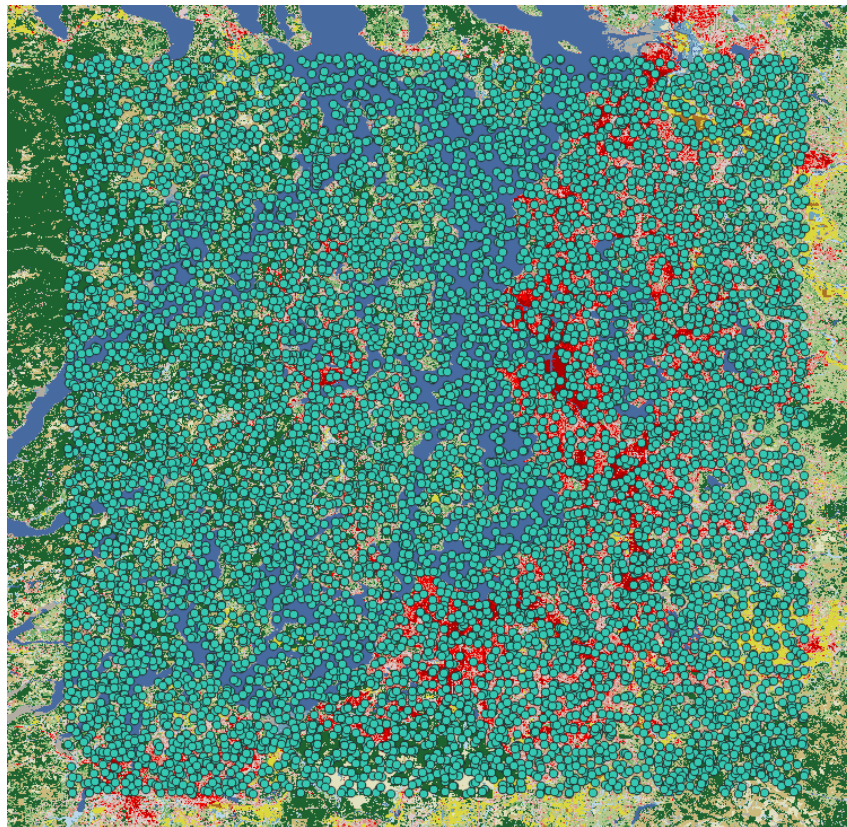


Figure 5.3: 10 000 samples (blue dots) used for verification of land cover extraction process overlaid on the land cover dataset in QGIS

SDF file read and write

In order to add the land cover data to the DEM, or the SDF file used by SPLAT!, the SDF data and land cover data formats should be identical. The addition of heights to the digital elevation model according to land cover is done by adding the data in two geographically identical grids in MATLAB. The SDF file is read into MATLAB. It is then augmented with the land cover data. Then a modified SDF file is written and stored. The process is illustrated in Figure 5.4.

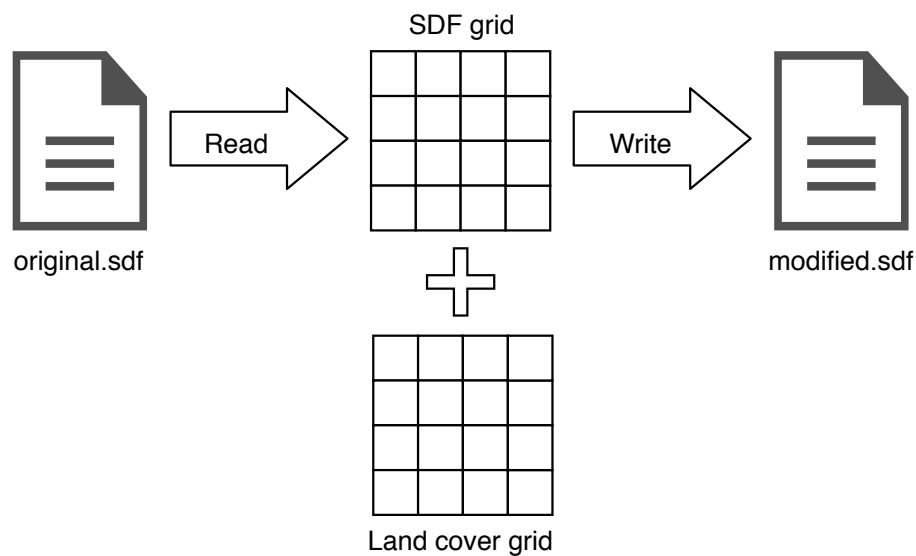


Figure 5.4: SDF file modification process

File read

The SDF file is a text file containing integers representing the heights of the DEM. It is not in a grid format. Instead the first four lines give the minimum and maximum degrees latitude and degrees longitude values followed by all the data, each in their own line. The data in the SDF file is extracted into a grid. The data representing the height in the South-West corner of the grid is given in row one and column one. The SDF grid is plotted using the reference vector as explained in section 3.3.2. The orientation of the data set is visually verified against a screen capture of Google Maps. Figure 5.5 is the screen capture from Google Maps and Figure 5.6 is the plot of the DEM in MATLAB. The rivers give a good indication that the orientation of the data is correct. The land cover data is already verified to be correct. In order to verify that the same reference vector and grid format is used for both the land cover data and the SDF data, the land cover data is plotted and visually compared to the plotted SDF data shown in Figure 5.6. The plot of the land cover grid can be seen in section 4.2, Figure 4.15. It can be seen from the layout of the rivers in the two plots, that the same reference vector and grid formats are used for both land cover and SDF data. After this process, it is verified that the SDF file is read and stored correctly into a grid in MATLAB. The land

cover data is extracted and stored correctly in an identical grid. The reference vector and grids are also verified to be correct. The heights from the land cover data that will be added to the DEM will therefore be on the correct geographic locations.



Figure 5.5: Google Maps screen capture

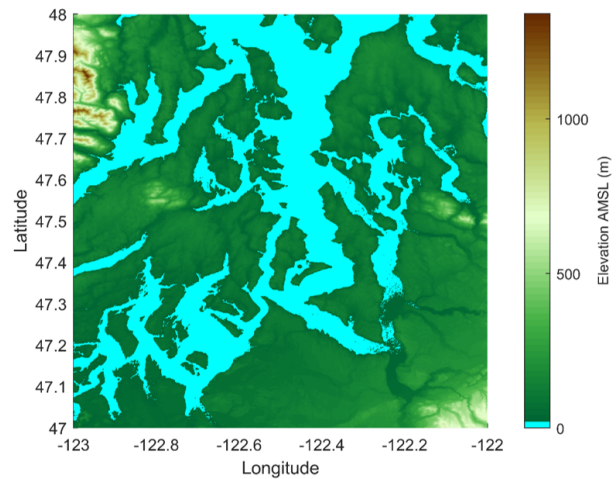


Figure 5.6: SDF grid output

Write file

The modified SDF file should be written in the correct format for SPLAT! to interpret. The geographic location information is written to the first four lines followed by the grid data. To verify that the file is written correctly, an SDF file is read using the function explained before. It is then immediately written to a text file using a self defined function in MATLAB. The newly written file is then read with the same function, and plotted. The writing function is visually verified and also numerically. The newly written and original SDF files are identical, verifying that the write function is implemented correctly.

5.1.4 Gradient descent method

The gradient descent method is verified by means of an example. Table 5.1 shows four iterations of the gradient descent method implemented to reduce the error between measurements and predictions. The gradient vector, heights per land cover class and

RMSE are shown. The five land cover classes used are classes 21, 22, 23, 24 and 41 to 43 grouped. The indices of the gradient vector correspond to the classes and so does the heights which are represented as a vector, separated by a space. For example in iteration 1, land cover class 21 corresponds to a height of 0m. Land cover class 22 corresponds to a height of 10m and so forth.

Table 5.1: Step-wise gradient descent example

Iteration	Gradient vector	Heights (m)	RMSE (dBm)
1	N/A	[0 10 8 28 40]	11.4879
2	[0; 0.0002; 0.007; -0.5036; 0.1168]	[0 10 8 32 39]	9.4688
3	[0.0003; 0.5056; 0.0084; 0.685; 0.7596]	[0 8 8 30 36]	9.3708
4	[0.0005; 0.5126; 0.0368; 0.6608; 0.0042]	[0 6 8 27 36]	9.2384

Consider iteration 2. The heights shown in iteration 1 are the initial heights. The RMSE in iteration 1 is therefore the initial error. The step size for this example is 4m. Each land cover class height is, on its turn, modified by adding 4m to the current height. The SDF file is modified and the propagation prediction is done using the ITM in SPLAT!. The new error is calculated and then the partial derivative. The gradient vector is updated to contain that partial derivative. The values in iteration 2 show the gradient vector after the iteration is completed. The heights in iteration 2 are the new heights after it is changed according to the gradient vector and the RMSE is the reduced error after all the heights are changed and the propagation prediction is done with the modified SDF file. The SDF file is modified with all the classes' heights in this step.

For example 4m is added to the height of class 24 and the new heights are 0m, 10m, 8m, 32m and 40m respectively. The partial derivative is calculated using the RMSE in iteration 1 and the newly calculated RMSE after the propagation prediction is done with the modified SDF file as input. The partial derivative is calculated to be -0.5036. Moving against the gradient the height of class 24 is expected to be higher in iteration 2 than in iteration 1. This is in fact the case. The height changes from 28m to 32m

in iteration 2. The height decreases for classes 41 to 43, from 40m to 39m, because the partial derivative of the RSME is positive. Note that the amount of height change differs for the two classes. It is because the partial derivative differs in size, meaning the change in one class has a greater effect on the error than the change of height in another class. The height of the first two classes doesn't change because the partial derivative of the RMSE is too small, meaning that the positive change in height did not have a significant impact on the error between measurements and predictions.

The heights are changing as expected by looking at the gradient vector. The RMSE is being reduced with every iteration. The gradient descent method is working as expected.

5.2 Validation

5.2.1 Introduction

Validation is done by testing the trained model on a set of data that is not part of the training set. This is done when implementing the 5-fold cross-validation. An additional measurement set from the nearby Microsoft station that isn't used in training of the model at all, is used for further validation. The DEM is modified by the average of the final heights from the result of the five iterations from implementing cross-validation. The modified SDF file is used as input for the propagation prediction on the evaluation set.

The purpose of validating the model on an independent set of transmitters that are not used in the training set, is to evaluate the performance of the trained model on new input data. The model may perform well on the data it is specifically trained for, but it is of little use when it can't perform well on new input data. If it can only be used on the training set, one would be better off just using the measurements, which are the ground truth. Instead of only needing a few measurements, the entire area needs to be

measured and then propagation prediction is not needed. This will however be a time consuming and expensive procedure. Another expected result of validating the model is to be able to evaluate if there is bias in the training of the model and give insight into improving the model.

5.2.2 Evaluation set

The Microsoft station is used as the receiver for the evaluation set. The station's parameters can be found in Table 3.2 in section 3.2.1. The same transmitters are used and the transmitters' information can be found in Table 3.5. The paths are different because the receiver is at a different geographic location. The proportion of land cover present on paths are similar to those for the case where the UW station is the receiver. The proportion of land cover types for the Microsoft station as the receiver is given in Figure 5.7. The type of area, for which the model is trained, is similar. The measurements from the Microsoft station was extracted and manipulated exactly the same as those from the UW station. The process and parameters are discussed in sections 3.2.4 and 3.2.5.

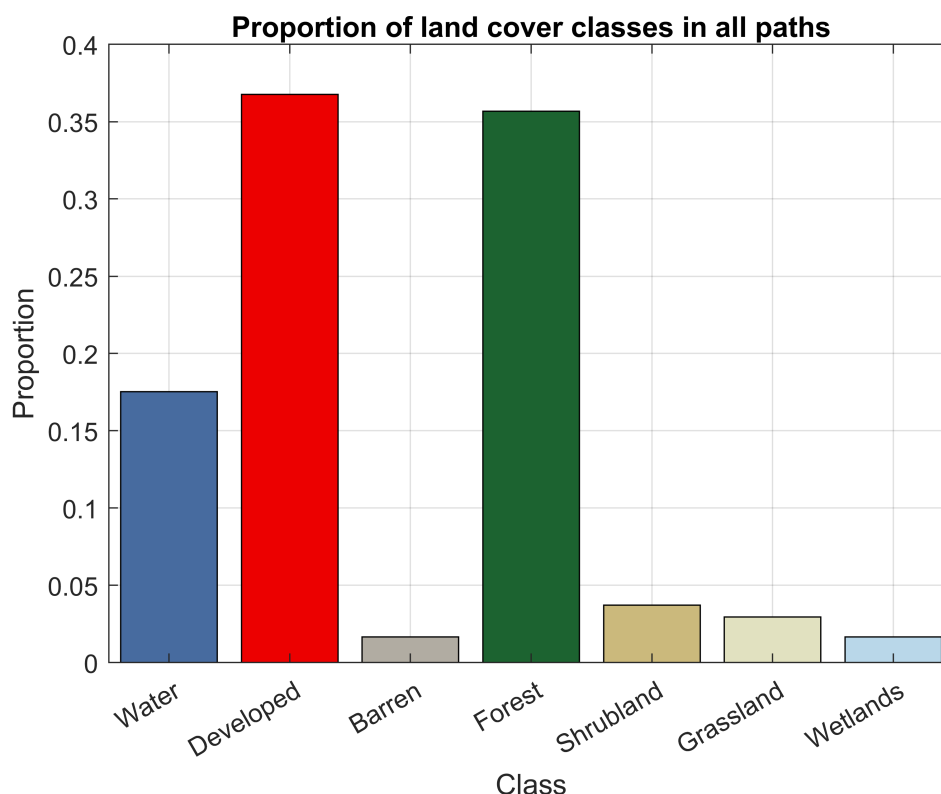


Figure 5.7: Proportion of Land Cover

ITM

SPLAT! is used to perform the ITM propagation prediction on the evaluation set. That is, for every transmitter-receiver pair. The predictions are first done using the unmodified SDF file to establish the baseline error. The predictions are repeated with the modified SDF files. The SDF files are augmented with the land cover data. The average of the heights from the results of the five cross-validation iterations, were added to the SDF file. The results are shown in Table 5.2.

Table 5.2 shows the results for the evaluation set, with the Microsoft station as the receiver. The heights per land cover class that were used, is given by the following vector, $[1\ 6\ 4\ 27\ 31]$, where again each entry in the vector corresponds to a land cover class. Therefore the first four entries are the different degrees of the developed class and the last entry is the forest class.

Table 5.2: Evaluation set results for ITM

Metric	Augmented with		Difference (Baseline - Augmented)
	Baseline	land cover	
RMSE (dBm)	51.2226	47.177	4.0456
MAE (dBm)	50.8416	44.5667	6.2749
Mean Error (dBm)	-50.8416	-44.5667	-6.2749
Standard deviation (dBm)	6.2354	15.4773	-9.2419
Correlation	0.9113	0.8924	0.0189

The measurements from the Microsoft station are drastically lower than expected. Power levels closer to that of the UW station are expected since the receivers are only 12.5km apart. The difference in error is what should be evaluated and the scale of the received power is therefore insignificant as long as it has a uniform offset. The differences in RMSE, mean and MAE are comparable to that of the UW station, validating the method. The standard deviation is significantly higher for the evaluation set but the correlation is still strong.

It could be that the settings of the spectrum observatory at the time of measurements are not reflected correctly. Physical access to the station was not available, only the measurement data. Assuming then that the PSD data from the Microsoft station is offset by a fixed amount, an attempt to explain the difference is made. a) The PSD data could be measured in dB but plotted in dBm. An additional 30dB should then be added. b) The station information indicates that a Mini Circuits Splitter ZX10R-14-S+ is used as a connector. That could account for an additional 6dB loss that should then be added. c) The cable loss could account for another 1dB to be added. d) The station information shows an Rx gain of 10dB. This value should be subtracted. The total to add to the PSD data is then 27dB. If all of the above is correct the values are more inline with what is expected as shown in Table 5.3. It is however, still high.

Table 5.3: Evaluation set results after the PSD is modified for ITM

Metric	Augmented with		Difference (Baseline - Augmented)
	Baseline	land cover	
RMSE (dBm)	23.8417	21.859	1.9827
MAE (dBm)	24.6436	23.4123	1.2313
Mean Error (dBm)	-23.8417	-17.5667	-6.275
Standard deviation (dBm)	6.2354	15.4773	-9.2419
Correlation	0.9113	0.8924	0.0189

Comparing Table 5.2 and Table 5.3 one can see that the reduction in the errors (RMSE and MAE) are lower when the power level is adjusted. From this result and further analysis into the individual path loss per transmitter-receiver pair, it is seen that the model more often over-predicts the path loss. The mean error is reduced because the negative and positive differences cancel out more often. For the evaluation set, the model makes large adjustments to a few transmitter-receiver pairs' path losses, while for the training set, the model makes smaller adjustments to the path losses but for more of the transmitter-receiver pairs. There seems to be training bias in the model. It could be because the transmitters are clustered together. It was also observed that a single meter in height change could have a significant effect on path loss. The path loss is therefore very sensitive to the height changes and the average heights used could lead to the result of poorer performance by the model.

Since the method of reducing the error changes the predicted path loss the scale of the actual measurement values aren't as important as the difference between the error before and after the method is implemented. The difference in error therefore validates the method. Adjusting the scale of the Microsoft station measurements is just an attempt to show that the evaluation set used is valid.

ITWOM

SPLAT! is used again to perform the ITWOM propagation prediction on the evaluation set. That is, for every transmitter-receiver pair. The predictions are first done using the unmodified SDF file to establish the baseline error. The predictions are repeated with the modified SDF files. The SDF files are augmented with the land cover data. The average of the heights from the five cross-validation iterations were added to the SDF file. The results are shown in Table 5.4.

Table 5.4 shows the results for the evaluation set, with the Microsoft station as the receiver. The heights per land cover class that were used, is given by the following vector, [3 6 0 37 17], where again each entry in the vector corresponds to a land cover class. Again, the first four entries are the different degrees of the developed class and the last entry is the forest class.

Table 5.4: Evaluation set results for ITWOM

Metric	Augmented with		Difference (Baseline - Augmented)
	Baseline	land cover	
RMSE (dBm)	48.7686	40.0024	8.7662
MAE (dBm)	48.0363	38.3753	9.661
Mean Error (dBm)	-48.0363	-38.3753	-9.661
Standard deviation (dBm)	8.4193	11.2928	-2.8735
Correlation	0.7285	0.7209	0.0076

The results of using the augmented model with the validation set are similar to the results of the cross-validation tests when fitting the model to the training set. The RMSE is reduced by 8.7662 dBm. The MAE and mean error are significantly lower. The standard deviation did increase but not as much as when the ITM was used. The correlation coefficient for the augmented model is almost equal to that of the baseline model. From this result it can be concluded that the augmented model reduces

the error between measured and predicted field strength values for new transmitter-receiver pairs that are in an area with similar land cover. As seen with the validation of the augmented ITM, the measurements from the Microsoft station seems to have a constant offset. The offset of 27 dB is added to the measurements and the validation performed again. The results are shown in Table 5.5.

Table 5.5: Evaluation set results after the PSD is modified for ITWOM

Metric	Augmented with		Difference (Baseline - Augmented)
	Baseline	land cover	
RMSE (dBm)	22.6586	16.0289	6.6297
MAE (dBm)	21.0363	11.6964	9.3399
Mean Error (dBm)	-21.0363	-11.3753	-9.661
Standard deviation (dBm)	8.4193	11.2928	-2.8735
Correlation	0.7285	0.7209	0.0076

The expected result was that the augmented ITWOM would perform much like the augmented ITM on the adjusted evaluation (Microsoft station) set. This is however not the case. Unlike the augmented ITM, the augmented ITWOM still reduced the RMSE significantly when compared to the baseline RMSE. The resulting RMSE, MAE, mean error, standard deviation and correlation coefficients are almost the same as the values using the original evaluation set (Microsoft station set without adjusting the power levels) with only the RMSE being noticeably lower.

5.2.3 Conclusion

The process of manipulating the field strength measurements and land cover was verified. The implementation of the gradient descent method was also verified. Therefore one can have confidence in the ground truth of the received power. One can also be confident that the land cover data was added to the DEM accurately. Furthermore one

can be sure that the process to reduce the error between field strength measurements and predictions worked as expected and that it reduced the error as expected.

The augmented models were validated by testing them on an evaluation set that wasn't used for training the models. The same transmitters but with a different receiver were used. That ensured that new transmitter-receiver pair paths were used. The field strength measurements of the evaluation set were much lower than expected. There seemed to be a constant bias in the measurements. The models were evaluated against the set. An attempt was made to explain the possible constant bias, and the models were evaluated against the evaluation set with the constant bias added. The effects of the augmented models were to be evaluated, meaning the difference between the baseline and the augmented model. The exact field strength measurement was thus of less importance. It did however still matter as can be seen from the results in Table 5.3. The higher received signal strengths meant that, for the ITM, sometimes the model over-predicted the path loss were it wouldn't have if the measured signal strength was lower.

On the initial evaluation set, the ITM performed as expected when compared with the training results in section 4.3.1. On the adjusted evaluation set, the model did not perform well. There seemed to be a training bias in the model. The ITWOM, which from the training seemed to be better suitable to be augmented with land cover data, did perform as expected on the initial evaluation set. Unlike the ITM, it still reduced the RMSE on the adjusted evaluation set, and it still performed well. There did not seem to be a training bias in the model. The ITWOM was validated, and the results showed that the augmented model could successfully reduce the error between measured and predicted field strengths, even for propagation paths that weren't in the training set.

The next chapter will give an overview of the study and the findings will be discussed. Recommended future work in this area of research will also be made.

Chapter 6

Conclusion

This chapter serves as a conclusion for the dissertation. An overview of the research is given. The research findings are then discussed. Following the findings of the research, recommendations for future work are made. The dissertation is then concluded with some closing remarks.

6.1 Dissertation overview

The problem identified, in short, is as follows. The need for available spectrum is increasing, and there is currently an artificial spectrum shortage due to current spectrum management policies. One solution is to utilise unused licensed bands. The unused spectrum is determined by means of a GLDB and propagation models are used to predict the coverage areas of primary transmitters in order to identify white spaces where the frequency bands can be reused. The problem is that large scale propagation models are inaccurate due to the many parameters with variability to account for. It is not always feasible to take all the parameters into account. There is thus a need to improve the accuracy of the large scale propagation models. Chapter 1 gave background information on the problem.

The aim of the research, as stated in section 1.5, was to develop a method to augment large scale propagation models in order to reduce the mean error between field strength measurements and predictions made by the models. Objectives were also set out in order to achieve the aim of the research.

A literature study was done in chapter 2. Important aspects of RF propagation were discussed. This included multipath effects on RF waves, different models to evaluate the loss of power due to diffraction, and the concepts of effective radiated power and gain of a transmitter. Different propagation models were discussed with the emphasis on the Longley-Rice ITM and the ITWOM. Background information was given on land cover data and considerations for taking field strength measurements as well as considerations for tuning large scale propagation models. Finally, related work on augmenting large scale propagation models with field measurements were discussed. The literature study provided insights into developing a method to augment large scale propagation models with field measurements. In the literature, the gradient descent method was identified as a candidate method to reduce the error between measurements and predictions. It also provided the necessary background information on aspects to consider when implementing the method.

The literature study was followed by a discussion of the research model and implementation in chapter 3. The measurements study was discussed in detail, followed by a discussion on processing the land cover data so that it can be used to augment the propagation models. Finally, the method to reduce the error was discussed. A process flow of the method was given in section 3.4.4, and the toolchain to implement it was explained in section 3.4.6. In short, land cover data were used to augment the DEMs that are inputs to the ITM and ITWOM. This was done by adding a height to the DEM according to the land cover class. The height represented the effect of the land cover on the path loss due to the presence of that land cover on the path. A gradient descent method was used to reduce the error between measurements and predictions, by tuning the heights associated with the land cover classes. The ITM and ITWOM were the chosen models to be augmented.

An exploratory data analysis followed in chapter 4. The results of the measurement study were analysed, and the processing thereof were discussed. The processed data were used in the implementation of the method. The land cover data was processed and analysed and then also discussed. Finally, the results of implementing the method to augment large scale propagation models were discussed. First the results for augmenting the ITM was discussed and then the results for augmenting the ITWOM.

Verification and validation followed in chapter 5. The three important aspects of this study were verified. First the processing of the field strength measurements were verified, followed by the verification of the processing of the land cover data. The implementation of the gradient descent method was also verified. After verifying that all the components of the research model were implemented correctly, the augmented propagations models were validated. The augmented models were validated on an independent set of transmitter-receiver pairs that weren't used in the training of the models. This was done by using a different receiver site. Finally, conclusions were made about the generalisation of the models as well as about the research model. The findings will be discussed next.

6.2 Findings

The research aimed to answer the following questions:

- Can large scale propagation models be augmented with field measurements in order to reduce the mean error between measurements and predictions, and if so, by how much can the error be reduced?
- What field measurements are needed to augment large scale propagation models?

The research ultimately showed that it is possible to reduce the mean error between field strength measurements and large scale radio propagation predictions by aug-

menting the propagation models with field measurements. In section 2.6, where the related work was discussed, it was found early that land cover data is a good field measurement to augment propagation models with. Numerous studies successfully reduced the error between measurements and predictions of different propagation models by adding land cover data that models the effect of the obstructions of the objects making up the land cover. In this study, the mean error between measurements and propagation predictions were also reduced by means of using land cover data. The amounts of the reduction in error for this study are discussed next.

The results of augmenting the ITM with land cover data showed that the RMSE between the field strength measurements and propagation predictions was reduced by around 7.7 dB, from 12.1 dBm to 4.3 dBm. This result was obtained by taking the average of the results of testing the model for five distinct subsets of the sample set, due to the 5-fold cross-validation used. The resulting standard deviation increased by around 2 dB, which means that the errors differed more from the mean than before. There was however still a moderately strong correlation between the predictions and measurements, indicated by a Pearson correlation coefficient of around 0.77. Therefore, by considering these results, it was initially found that the mean error for the ITM could be reduced significantly by augmenting the model with land cover data and training the model using field strength measurements.

The augmented model was validated using a set of independent transmitter-receivers pairs so that the paths were not in the training set. There was uncertainty regarding the accuracy of measurements from the second receiver site, and an attempt was made to explain the difference between the expected and actual field strength measurements. Nonetheless, both cases, expected and actual measurements, were used to validate the augmented ITM. Using the measurements as is, the RMSE for the augmented ITM was still lower than the baseline RMSE. It was however not as much of an improvement than the results from training the model and using cross-validation but this is to be expected. The standard deviation was much higher. It was found that there were only a few differences between the baseline errors and those from the augmented model, but when there were differences, the differences were large. This caused the RMSE to be

reduced but the standard deviation to increase dramatically. It was much worse when the adjusted measurements were used for validation. The RMSE was barely reduced, and the standard deviation was still much higher. It was found that now, the errors were still large but there were more positive errors than with the previous results. From this result it was found that the augmenting of the ITM produced a reduction in error on the trained data, but it did not generalise well. There seems to be a training bias. The location of the receiver seems to have a significant effect on the training of the model as the land cover class surrounding the receiver had the most impact on the error. Another explanation for the model performing well on the training data but not the new set is the fact that although the paths differ because of the unique locations and heights of the transmitters, the small set of unique land cover paths from transmitters to the receiver, is still too similar. The transmitters are located close to each other, and some are even on the same physical tower with only the heights differing. The evaluation set used in cross-validation can contain paths that are very similar to those in training, and that may produce good results. From this case study, it was found that more diversity in transmitter-receiver paths are needed to train the model in order for it to be generalised better so that it can be used in similar environments and still perform well.

Similarly, the results of augmenting the ITWOM with land cover data showed that the RMSE between the field strength measurements and propagation predictions was reduced by around 7.8 dB, from 10.9 dBm to 3.1 dBm. This result was obtained by taking the average of the results of testing the model for five distinct subsets of the sample set, due to the 5-fold cross-validation used. The resulting standard deviation increased by around 0.2 dB, which means that the errors differed more from the mean than before but very slightly. There was however still a very strong correlation between the predictions and measurements, indicated by a Pearson correlation coefficient of around 0.9, which is also a slight improvement on the baseline value. Therefore, by considering this results, it was initially found that the mean error for the ITWOM could be reduced significantly by augmenting the model with land cover data and training the model using field strength measurements. The improvement on the ITWOM is

more significant than the improvement on the ITM, partly due to the small increase in standard deviation.

As was done with the ITM, the augmented method was validated using a set of independent transmitter-receivers pairs so that the paths were not in the training set. The augmented ITWOM performed almost as well on the independent set of transmitter-receivers pairs as on the training data. A significant difference is that the standard deviation increased to around 2.8 dBm. Unlike the ITM, the ITWOM performed well on both the expected and actual measurements from the receiver used for validation.

Field strength measurements were taken over a period of three months. Variation in measurements was expected and also observed. There was a notable occurrence where there were significant variations in the measurements for a period of a few days. It was also seen that the field strength measurements were mostly consistent over the measurement period. It was therefore found that a lengthy measurement campaign is not a requirement for augmenting large scale propagation models like the ITWOM, with field measurements.

6.3 Recommendations for future work

This study showed that it is possible to augment large scale propagation models using field measurements in order to reduce the mean error between predicted and measured signal strength. For future work, the following recommendations are made.

In general, it would be advantageous to have many more transmitter-receiver pairs in order to have a lot more measurement samples to use for training and testing respectively. By having more transmitter-receiver pairs, the chances are that there will be more diversity in the path profiles, including the path profile of the terrain type. Specifically, more receiver locations are recommended since it seems like the terrain directly surrounding the receiver has a significant impact.

The broadcasting frequencies of transmitters used in this study were all in the FM radio band. Transmitters in the TV band were considered but since the receiver was stationary, the lack of a number of measurements from transmitters in this frequency range, made it infeasible. From the related work in section 2.6, it was seen that higher frequencies are affected more by the land cover on the paths. It is recommended that other frequency ranges are also explored when augmenting large scale propagation models with field measurements. The TV band would be a suitable frequency range, especially for TVWS applications.

The land cover data that was used in this study was upsampled from a resolution of 30x30m to 90x90m. The reason for upsampling the land cover data was for it to be similar to the resolution of the SRTM data. Using a resolution of 90x90m increased the speed of augmenting the SRTM data and also increased the speed of propagation predictions. Both were extensively used in augmenting the model and had a significant impact on the speed of execution. SRTM data is available in a resolution of 30x30m, and it is recommended that this, together with the original land cover data (30x30m), is used in future work.

Most of the transmitter-receiver paths from this study included a large amount of water in the path. This was due to the location of the measurement campaign which was conducted in Seattle, Washington. The land cover class, water, was not considered as one of the classes to train the model with. This is because of the rationale that the land cover is an additional obstruction to the EM wave, as discussed in section 3.4.1. Water can have a significant impact on the measured field strength since EM waves can easily be reflected off smooth surfaces. For future work, it is recommended that water is also included in the paths when augmenting large scale propagations with field measurements since the effect of the land cover class on the path is modelled, regardless if it is an obstruction or not.

From the literature study it was found that the elevation measured for SRTM over some land cover types like dense forests may be too high. This is due to the radar waves not being able to penetrate the top of a dense forest canopy. This could justify

negative heights for some land cover classes and should be explored in future work.

The gradient descent method was used to find heights associated with land cover classes in order to reduce the mean error between predicted and measured path loss. It was found that the search space displayed many sub-minima and the gradient descent method could not find a global minima. The gradient descent method could also not escape from local minima and supervision and intervention were required to escape from local minima. Future work should therefore include the use of more sophisticated optimisation algorithms that can climb out of local sub-minima and have a better chance of finding the global minima.

The augmented ITWOM performed better than the augmented ITM in terms of reducing the mean error. This suggests that some models are better suited for this type of augmenting, as discussed in this study. Other commonly used large scale propagation models could be augmented with the same land cover data and using the same method. In future work, the results can then be compared to the results of this study.

6.4 Closure

This study aimed to develop a method to augment large scale propagation models with field measurements in order to reduce the mean error between the propagation predictions and field strength measurements. It also attempted to quantify the achievable reduction in error. The results of the study showed that it is possible to augment large scale propagation models with field measurements to achieve a considerable reduction in error. It also made use of a stationary spectrum observatory, which with the equipment in place, means that continuous measurements can be collected with relative ease.

The augmented large scale propagation models can be incorporated into network planning tools to provide more accurate predictions and benefit the network operator in terms of cost, and also the clients in terms of quality of service. It can also assist in the

decision making regarding secondary unlicensed use of the spectrum, specifically in the TV band where TVWS with the use of a GLDB seems popular.

The augmented large scale propagation models can be used in spectrum monitoring and management applications. Better informed decisions regarding the use of spectrum can be made, to ultimately use the available spectrum more efficiently.

Bibliography

- [1] "Regionally harmonized bands," 2019. [Online]. Available: <https://www.itu.int/en/ITU-R/information/Pages/emergency-bands.aspx>
- [2] T. Chiwewe and G. Hancke, Jr, "A look at spectrum management policies for the efficient use of radio spectrum," Mar. 2015. [Online]. Available: <http://www.ee.co.za/article/look-spectrum-management-policies-efficient-use-radio-spectrum.html>
- [3] J. P. Linnartz, "Diffraction Loss." [Online]. Available: <http://www.wirelesscommunication.nl/reference/chaptr03/diffrac.htm>
- [4] K. Bullington, "Radio Propagation for Vehicular Communications," *IEEE Transactions on Vehicular Technology*, vol. 26, no. 4, pp. 295–308, Nov. 1977.
- [5] M. Willis, "Propagation Tutorial - Diffraction," May 2007. [Online]. Available: <http://www.mike-willis.com/Tutorial/PF7.htm>
- [6] C. Haslett, *Essentials of Radio Wave Propagation*. Cambridge University Press, 2008.
- [7] P. McKenna and C. Hammerschmidt, "Propagation measurement workshop Radio propagation models," Jul. 2016. [Online]. Available: <https://www.its.bldrdoc.gov/media/66385/propmeasworkshop3july282016.pdf>
- [8] G. Hufford, A. Longley, and W. Kissick, "A Guide to the Use of the ITS Irregular Terrain Model in the Area Prediction Mode," U.S. Department of Commerce,

-
- National Telecommunication and Information Administration, Tech. Rep., Apr. 1982. [Online]. Available: http://www.ntia.doc.gov/files/ntia/publications/ntia_82-100_20121129145031_555510.pdf
- [9] S. Roy, K. Shin, A. Ashok, M. McHenry, G. Vigil, S. Kannam, and D. Aragon, "CityScape: A Metro-Area Spectrum Observatory," in *2017 26th International Conference on Computer Communication and Networks (ICCCN)*, Jul. 2017, pp. 1–9.
- [10] P. L. Rice and A. G. Longley, "Prediction of Tropospheric Radio Transmission Loss Over Irregular Terrain. A Computer Method-1968," Institute for Telecommunication Sciences, Technical, 1968.
- [11] J. A. Magliacane, D. Macdonald, and R. Bentley, "Ubuntu Manpage: splat An RF Signal Propagation, Loss, And Terrain analysis tool." [Online]. Available: <http://manpages.ubuntu.com/manpages/trusty/man1/splat.1.html>
- [12] "National Land Cover Database 2011 (NLCD2011) Legend," Mar. 2018. [Online]. Available: <https://www.mrlc.gov/data/legends/national-land-cover-database-2011-nlcd2011-legend>
- [13] J. Devore and N. Farnum, *Applied Statistics for Engineers and Scientists*, 2nd ed. Curt Hinrichs, 2005.
- [14] M. Cave, C. Doyle, and W. Webb, *Essentials of Modern Spectrum Management*. Cambridge University Press, Aug. 2007.
- [15] "Collection of the basic texts of the International Telecommunication Union adopted by the Plenipotentiary Conference," 2015. [Online]. Available: <http://search.itu.int/history/HistoryDigitalCollectionDocLibrary/5.21.61.en.100.pdf>
- [16] "What is Spectrum Harmonization? - Definition from Techopedia." [Online]. Available: <https://www.techopedia.com/definition/30519/spectrum-harmonization>
- [17] A. M. Wyglinski, M. Nekovee, and T. Hou, *Cognitive Radio Communications and Networks: Principles and Practice*. Academic Press, Nov. 2009.

-
- [18] L. E. Doyle, *Essentials of Cognitive Radio*. Cambridge University Press, 2009.
- [19] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Computer Networks*, vol. 50, no. 13, pp. 2127–2159, Sep. 2006.
- [20] "TV White Space Databases," Sep. 2016. [Online]. Available: <https://www.ofcom.org.uk/spectrum/spectrum-management/TV-white-space-databases>
- [21] "White Space Database Administration," May 2011. [Online]. Available: <https://www.fcc.gov/general/white-space-database-administration>
- [22] "DBS-01 — White Space Database Specifications," Feb. 2015. [Online]. Available: <http://www.ic.gc.ca/eic/site/smt-gst.nsf/eng/sf10928.html>
- [23] "Regulations on the use of Television White Spaces," in *Governt Gazette, Independent Communication Authority of South Africa*, Mar. 2018.
- [24] "3.5 GHz Band Overview," Dec. 2015. [Online]. Available: <https://www.fcc.gov/rulemaking/12-354>
- [25] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the Error of Path Loss Models," in *2011 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, May 2011, pp. 71–82.
- [26] J. S. Seybold, *Introduction to RF Propagation*. John Wiley & Sons, Inc, Oct. 2005.
- [27] "IEEE Standard for Definitions of Terms for Antennas," *IEEE Std 145-2013 (Revision of IEEE Std 145-1993)*, pp. 1–50, Mar. 2014.
- [28] A. F. Molisch, *Wireless Communications*, 2nd ed. Chichester, West Sussex, U.K: Wiley, 2011.
- [29] J. R. Hallas, "Antenna Gain Specs—What Do They Really Mean?" *QST*, vol. 94, no. 5, pp. 40–40, Mar. 2010.
- [30] G. A. Jones, D. H. Layer, and T. G. Osenkowsky, *National Association of Broadcasters Engineering Handbook*, 10th ed. Taylor & Francis, Apr. 2013.
-

-
- [31] C. Huygens, *Traité de la Lumière*. (Leyden 1690) [English translation by S.P. Thompson, *Treatise on Light*]. London: Macmillan, 1912.
- [32] D. Miller, "Huygens's wave propagation principle corrected," *Optics Letters*, vol. 16, no. 18, pp. 1370–1372, 1991.
- [33] A. Jimoh J, J. O. Okoye, and E. Akinloye Bolanle, "Software for Fresnel-Kirchoff Single Knife-Edge Diffraction Loss Model," *Mathematical and Software Engineering*, Vol 2, Iss 2, Pp 76-84 (2016), no. 2, p. 76, 2016.
- [34] K. Bullington, "Radio Propagation at Frequencies above 30 Megacycles," *Proceedings of the IRE*, vol. 35, pp. 1122 – 1136, 1947.
- [35] J. Epstein and D. W. Peterson, "An Experimental Study of Wave Propagation at 850 MC," *Proceedings of the IRE*, vol. 41, no. 5, pp. 595–611, May 1953.
- [36] J. Deygout, "Multiple Knife-Edge Diffraction of Microwaves," *Antennas and Propagation, IEEE Transactions on*, vol. 14, pp. 480 – 489, 1966.
- [37] I. Poole, "Radio Wave Reflection." [Online]. Available: http://www.radio-electronics.com/info/propagation/em_waves/electromagnetic-reflection-refraction-diffraction.php
- [38] "Radio Mobile - RF propagation simulation software : Fresnel zones." [Online]. Available: http://radiomobile.pe1mew.nl/?Calculations:Propagation_calculation:Fresnel_zones
- [39] V. S. Abhayawardhana, I. J. Wassell, D. Crosby, M. P. Sellars, and M. G. Brown, "Comparison of Empirical Propagation Path Loss models for Fixed Wireless Access Systems," in *2005 IEEE 61st Vehicular Technology Conference*, vol. 1, May 2005, pp. 73–77 Vol. 1.
- [40] M. Hata, "Empirical Formula for Propagation Loss in Land Mobile Radio Services," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, Aug. 1980.

-
- [41] 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 *2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, Jun. 2014, pp. 1–4.
- [42] ITU-R, "Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3 000 MHz P Series Radiowave propagation," Technical, 2013.
- [43] S. E. Shumate, "Longley-Rice and ITU-P.1546 Combined: A New International Terrain-Specific Propagation Model," in *Vehicular Technology Conference Fall (VTC 2010-Fall)*, 2010 IEEE 72nd, Sep. 2010, pp. 1–5.
- [44] "Radio Mobile - RF propagation simulation software: ITM model propagation settings." [Online]. Available: http://radiomobile.pe1mew.nl/?Calculations:ITM_model_propagation_settings
- [45] "Longley-Rice Variability Values," 2003. [Online]. Available: https://www.softwright.com/faq/support/longley_rice_variability.html
- [46] "Notes on Longley-Rice Propagation," 2000. [Online]. Available: https://www.softwright.com/faq/engineering/prop_longley_rice.html
- [47] "SPLAT! for Windows." [Online]. Available: <https://www.dxzone.com/dx16760/splat-for-windows.html>
- [48] S. E. Shumate, "NTIA releases updated Version 7.0 of the Longley-Rice Irregular Terrain Model," *IEEE Broadcasting Technology Society Newsletter*, pp. 20–22, 2007, vol. 15, no. 3.
- [49] —, "More on Longley-Rice: Error Codes Decoded, Out-of-Date Instructions, and An Updated FreeWrap-Around," *IEEE Broadcasting Technology Society Newsletter*, pp. 18–21, 2007, vol. 15, no. 4.
- [50] —, "On Replacing a 30-Year-old Temporary Patch in Longley-Rice," *IEEE Broadcasting Technology Society Newsletter*, pp. 16–18, 2008, vol. 16, no. 1.
-

-
- [51] —, “Longley-Rice and the Irregular Terrain Model,” *IEEE Broadcasting Technology Society Newsletter*, pp. 22–23, 2008, vol. 16, no. 2.
- [52] —, “The Irregular Terrain Model (ITM) Averaging System,” *IEEE Broadcasting Technology Society Newsletter*, pp. 19–21, 2008, vol. 16, no. 3.
- [53] —, “Longley-Rice’s Terrain Irregularity Parameter, Delta-H,” *IEEE Broadcasting Technology Society Newsletter*, pp. 23–25, 2008, vol. 16, no. 4.
- [54] —, “Longley-Rice’s Faulty Subroutines, Part 1: Z1sq1,” *IEEE Broadcasting Technology Society Newsletter*, pp. 16–17, 2009, vol. 17, no. 1.
- [55] —, “Longley-Rice’s Faulty Subroutines, Part 3: Log vs. Ln Confusion,” *IEEE Broadcasting Technology Society Newsletter*, p. 26, 2009, vol. 17, no. 4.
- [56] —, “Longley-Rice’s Faulty Subroutines: Background on the c++ alos,” *IEEE Broadcasting Technology Society Newsletter*, pp. 28–31, 2010, vol. 18, no. 1.
- [57] —, “Longley Rice’s Most Complex Line of Code,” *IEEE Broadcasting Technology Society Newsletter*, pp. 29–33, 2010, vol. 18, no. 2.
- [58] —, “Testing the Fix for Longley-Rice’s Most Complex Line of Code,” *IEEE Broadcasting Technology Society Newsletter*, pp. 22–25, 2010, vol. 18, no. 3.
- [59] —, “An Error in Diffraction Calculation in the ITM?” *IEEE Broadcasting Technology Society Newsletter*, pp. 20–21, 2011, vol. 19, no. 1.
- [60] S. Kasampalis, P. I. Lazaridis, Z. D. Zaharis, A. Bizopoulos, S. Zetlas, and J. Cosmas, “Comparison of Longley-Rice, ITM and ITWOM propagation models for DTV and FM broadcasting,” in *2013 16th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Jun. 2013, pp. 1–6.
- [61] P. Lazaridis, A. Bizopoulos, S. Kasampalis, J. Cosmas, and Z. Zaharis, “Evaluation of prediction accuracy for the Longley-Rice model in the FM and TV bands,” Sep. 2013.
- [62] “Q-Rap: Radio Planning Tool.” [Online]. Available: <http://www.qrap.org.za/>
-

-
- [63] M. Ballot, "Radio-Wave Propagation Prediction Model Tuning of Land Cover Effects," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 8, pp. 3490–3498, Oct. 2014.
- [64] "Radio Mobile - RF propagation simulation software: Propagation calculation." [Online]. Available: http://radiomobile.pe1mew.nl/?Calculations:Propagation_calculation
- [65] "SPLAT! A Terrestrial RF Path Analysis Application For Linux/Unix." [Online]. Available: <https://www.qsl.net/kd2bd/splat.html>
- [66] "SRTM topography." [Online]. Available: https://dds.cr.usgs.gov/srtm/version2_1/Documentation/SRTM_Topo.pdf
- [67] "Index of SRTM data." [Online]. Available: <https://dds.cr.usgs.gov/srtm/>
- [68] "Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global," Jan. 2015. [Online]. Available: <https://lta.cr.usgs.gov/SRTM1Arc>
- [69] "Shuttle Radar Topography Mission Frequently Asked Questions," Nov. 2018. [Online]. Available: <https://www2.jpl.nasa.gov/srtm/faq.html>
- [70] "National Land Cover Data for SA." [Online]. Available: [/national_land_cover_data_sa](#)
- [71] K. Ruščić and A. Skenderović, "Measurements and Propagation Model Tuning in DVB-T2 network," in *Proceedings ELMAR-2014*, Sep. 2014, pp. 1–4.
- [72] "2013-14 National Land-Cover – 72 classes," 2015. [Online]. Available: <http://www.sasdi.net/metaview.aspx?uuid=0e1bda9b920f00f2a82adea0c3a905fd>
- [73] "MRLC NLCD 2011." [Online]. Available: <https://www.mrlc.gov/nlcd2011.php>
- [74] "NLCD 2011 Land Cover (CONUS) - Metadata." [Online]. Available: https://www.mrlc.gov/sites/default/files/metadata/NLCD_2011_Land_Cover_L48.xml

-
- [75] "Coordinate systems, map projections, and geographic (datum) transformations." [Online]. Available: <http://resources.esri.com/help/9.3/arcgisengine/dotnet/89b720a5-7339-44b0-8b58-0f5bf2843393.htm>
- [76] "Datums—Help | ArcGIS for Desktop." [Online]. Available: <http://desktop.arcgis.com/en/arcmap/10.3/guide-books/map-projections/datums.htm>
- [77] C. A. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. John Wiley & Sons, Apr. 2005, vol. 1, google-Books-ID: agWwQgAACAAJ.
- [78] M. Anderson and D. Hall, "Understanding RF Instrument Specifications Part 1." [Online]. Available: https://www.eetimes.com/document.asp?doc_id=1276291
- [79] J. A. Snyman, *Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms*, ser. Applied Optimization. Springer US, 2005.
- [80] P. Avez, P. V. Wesemael, A. Bourdoux, A. Chiumento, S. Pollin, and V. Moeyaert, "Tuning the Longley-Rice propagation model for improved TV white space detection," in *2012 19th IEEE Symposium on Communications and Vehicular Technology in the Benelux (SCVT)*, Nov. 2012, pp. 1–6.
- [81] S. Roy, "White Space Networking: Integrated Sensing and Database Architecture," Hyderabad, India, 2014.
- [82] A. Achtzehn, J. Riihijärvi, and P. Mähönen, "Improving Accuracy for TVWS Geolocation Databases: Results from Measurement-Driven Estimation Approaches," in *2014 IEEE International Symposium on Dynamic Spectrum Access Networks (DYS-PAN)*, Apr. 2014, pp. 392–403.
- [83] D. D. E. Michel and T. Emmanuel, "Optimization of Okumura Hata Model in 800MHz based on Newton Second Order algorithm. Case of Yaoundé, Cameroon," p. 10, Mar. 2015.

-
- [84] city_scape, "city-scape/CityScape_raw_data_decoder," Sep. 2017, original-date: 2016-11-29T00:16:29Z. [Online]. Available: https://github.com/city-scape/CityScape_Raw_Data_Decoder
- [85] "LMR-400 Datasheet." [Online]. Available: <https://www.timesmicrowave.com/documents/resources/LMR-400.pdf>
- [86] "FM Query Broadcast Station Search," Dec. 2015. [Online]. Available: <https://www.fcc.gov/media/radio/fm-query>
- [87] "CityScape Spectrum Observatory, CityScape Station Installation and Operation Manual," Aug. 2016.
- [88] P. J. Rousseeuw and C. Croux, "Alternatives to the Median Absolute Deviation," *Journal of the American Statistical Association*, vol. 88, no. 424, pp. 1273–1283, Dec. 1993.
- [89] R. G. Sargent, "Verification and validation of simulation models," in *Proceedings of the 2011 Winter Simulation Conference*, 2011, pp. 183–198.

Appendix A

Results of augmenting the ITM using random initial heights

Table A.1: Results for first random initial height set using ITM

Iteration	Heights (m)	RMSE (dBm)		MAE (dBm)		Mean Error (dBm)		STD (σ) (dBm)		Correlation	
		Train	Validate	Train	Validate	Train	Validate	Train	Validate	Train	Validate
1	[4 9 0 21 40]	10.5409	7.0448	8.9758	12.8677	-8.4867	-12.8677	6.3864	6.2873	0.9317	0.8064
2	[4 9 0 21 40]	11.3477	5.6542	9.8638	10.2156	-9.1497	-10.256	6.8566	5.3127	0.9135	0.8377
3	[4 9 0 21 40]	11.9937	4.1101	10.5011	6.7587	-10.0140	-6.7587	6.7426	5.1256	0.8980	0.9843
4	[4 9 0 21 40]	11.3229	5.7037	9.9341	9.0266	-9.4556	-8.9920	6.3629	7.6894	0.9212	0.8621
5	[4 9 9 20 42]	10.0964	6.1280	8.2190	9.4048	-6.6777	-7.4854	7.7355	10.6308	0.8343	0.8082

Table A.2: Results for second random initial height set using ITM

Iteration	Heights (m)	RMSE (dBm)		MAE (dBm)		Mean Error (dBm)		STD (σ) (dBm)		Correlation	
		Train	Validate	Train	Validate	Train	Validate	Train	Validate	Train	Validate
1	[2 3 3 24 31]	8.2471	5.9056	6.0921	9.7045	-4.4421	-8.2694	7.0980	9.2384	0.8946	0.3756
2	[2 3 3 24 31]	9.6076	3.2536	7.0922	5.7040	-5.5068	-4.0100	8.0421	5.6135	0.8547	0.8777
3	[2 3 3 24 31]	9.3100	4.0267	6.9472	6.2839	-5.3115	-4.7920	7.8107	7.0904	0.8405	0.9485
4	[2 3 3 24 31]	9.0060	4.6673	7.1942	5.2960	-5.4698	-4.1588	7.3085	9.1546	0.8833	0.7473
5	[2 4 3 25 31]	9.1322	4.2502	6.7746	6.8910	-5.0856	-4.5504	7.7482	7.8652	0.8374	0.9022

Table A.3: Results for third random initial height set using ITM

Iteration	Heights (m)	RMSE (dBm)		MAE (dBm)		Mean Error (dBm)		STD (σ) (dBm)		Correlation	
		Train	Validate	Train	Validate	Train	Validate	Train	Validate	Train	Validate
1	[5 6 3 27 26]	8.0982	5.9064	6.2250	9.7129	-3.9809	-8.2744	7.2039	9.2354	0.9005	0.3758
2	[5 6 3 27 26]	9.4789	3.2585	7.2231	5.7206	-5.0393	-4.0406	8.2010	5.6012	0.8610	0.8778
3	[5 9 3 30 26]	8.9939	4.6673	6.7455	8.7839	-4.8123	-2.0904	7.7615	9.9658	0.8548	0.9610
4	[5 6 3 27 26]	8.8889	4.6318	7.3954	5.0313	-5.1223	-3.7088	7.4209	9.2991	0.8894	0.7499
5	[5 4 3 25 26]	9.1339	4.2541	6.7734	6.9060	-5.1194	-4.5721	7.7271	7.8602	0.8379	0.9024