Bachelor of Science in Electrical and Electronic Engineering

EEE 400 (January 2024) Thesis

Ground Station to Satellite Path Loss Model Comparison: A Machine Learning Model and A Mathematical Model

Submitted by

H.A. Hassan Tahmid
1806129

Abrar Tanvir Toki
1806029

Supervised by

Dr. Pran Kanai Saha

Professor, EEE, BUET

Department of Electrical and Electronic Engineering

BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

Dhaka, Bangladesh

March 2025

Contents

List of Figures

Chapter 1

Introduction

1.1 Introduction

Satellite communication is an important part of today's telecommunication, broadcasting, and navigation systems. The performance and reliability of satellite links are largely affected by several atmospheric parameters that attenuate the signal. Understanding and being able to accurately predict path loss due to environmental effects is required in order to design and optimize satellite communication networks. For Bangladesh, where the weather is diverse, it is preferable to model satellite path loss for enhancing the quality of satellite-based communication services.

Satellite communication path loss is caused by various factors such as free-space attenuation, atmospheric loss, rain attenuation, and cloud interference. Ground stations utilized to quantify attenuation by transmitting a dummy signal to the satellite and estimating its loss. This, however, takes resources and infrastructure on top of what is already needed. If a reliable predictive model can be established, it would eliminate the necessity of transmitting these dummy signals, thereby lessening operating expenses and boosting efficiency. Therefore, we tried to study some previous works done on this issue. We are mentioning some of the research works that we studied.

The first paper aims at creating a systematic model for the satellite-to-ground path-loss in urban environments. The authors emphasize the importance of understanding how line-of-sight probability and shadowing impact signal propagation. The methodology is semi-analytic, applying stochastic geometry to represent line-of-sight probability and Gaussian Mixture Model to represent shadowing. The model is compared with GNSS receiver measurements. The study concludes that the proposed framework accurately estimates satellite-to-ground path-loss and balances between simplicity and accuracy. Its flexibility makes it suitable for IoT and hybrid satellite-terrestrial networks. [2]

The second study discusses the use of deep learning to predict path-loss distribution from 2D satellite images without considering a 3D model. The authors

apply a convolutional neural network, modifying the VGG-16 architecture, to find patterns from satellite images. Training and test data are generated with ray-tracing simulations from 3D models. Instead of predicting point values, the model outputs the overall path-loss distribution in an area. The work finds that deep learning successfully predicts path-loss in different urban and suburban areas. Transfer learning enhances model performance even with small data. This method provides an efficient alternative to conventional ray-tracing simulations and, as such, is promising for applications in real time. [1]

The third paper presents the development of a real-time model for mm-wave satellite links atmospheric attenuation over the Indian subcontinent. Due to the geographical and climatic diversity of the nation, precise estimation of attenuation due to rain, atmospheric gases, and clouds is emphasized by the research. Using the freely available ERA-5 reanalysis data and GPM precipitation data, the authors apply the Synthetic Storm Technique for rain attenuation, the ITU-R model for gaseous absorption, and the Salonen model for cloud attenuation. A frequency scaling approach is suggested to predict higher frequency attenuation (50 GHz) based on data from lower frequencies (30 GHz). The study concludes that the approach enables real-time attenuation to be predicted accurately while reducing reliance on costly measuring equipment. The approach can be beneficial for improving fade countermeasures in satellite communication systems, particularly for tropical environments. [11]

The fourth article presents a machine learning approach to improve the accuracy of path-loss predictions using a combination of principal component analysis, artificial neural networks, and Gaussian process regression. The authors apply PCA to eliminate the most irrelevant features and reduce the dataset before applying an artificial neural network for multi-dimensional regression. A model based on Gaussian Process is applied for further accuracy in the prediction to examine variance and quantify the shadowing effect. The study is based on real measurements of path-loss in a South Korean suburb. Results indicate PCA improves efficiency via removal of redundant variables, while the ANN model has excellent performance above the typical log-distance models. The integration of ANN with Gaussian Process regression provides more precise path-loss estimation while providing a better quantification of uncertainty. [8]

The fifth paper targets the development of a deep-learning-based method for path-loss prediction using geospatial information and path profiles. The authors emphasize the need for denser coverage in over 5G and 6G networks, particularly within high-frequency millimeter to terahertz band environments. As a solution for overcoming the radio propagation modeling complexity in such mediums, they discuss an approach which learns environment patterns from images of three critical regions: transmitting site, receiving site, and area between them. One of the primary developments in their method is parameterization of the propagation path environment to maintain a constant image size, regardless of distance, that enhances accuracy of path loss estimation. They merge image data from the receive and transmit points with environmental information using convolutional neural networks and deep neural networks. Estimates based on measurement data across urban, suburban, and rural locations at 800 MHz

and 2 GHz show that the approach outlined in this paper provides stunningly improved accuracy over traditional models. [4]
The sixth paper is an extensive overview of machine learning application in wireless networks, showing leading techniques and unresolved problems. The study classifies machine learning applications into MAC layer resource management, network and mobility management at the network layer, and application layer localization. The study contrasts the traditional approaches with machine-learning-based approaches in a systematic way by comparing performance variation and citing reasons for why ML is used. Various ML techniques such as supervised, unsupervised, and reinforcement learning are described by the authors in networking. The paper also identifies issues, including the need for standardized datasets, the need for upgrading infrastructure to support ML-based paradigms, and theoretical guidelines for deployment. The study concludes that machine learning is of great value in wireless networks, but there are a few technical and practical issues that must be addressed in order to fully realize its potential. [12]
The seventh paper investigates path loss modeling through neural networks and an ensemble method for future wireless networks, particularly for 5G and beyond. Because higher-frequency bands are required with increasing data demands, the study compares machine learning models—artificial neural networks (ANN), recurrent neural networks with long short-term memory (RNN-LSTM), convolutional neural networks (CNN), and an ensemble model that combines all these methods. The test and training dataset was obtained through measurement campaigns inside an indoor corridor, line-of-sight, and non-line-of-sight scenarios. It shows that ensemble methods are more effective than stand-alone models when it comes to prediction accuracy and consistency, and therefore the most effective in complex environments at high-frequency bands. The paper points out that the machine learning approaches are superior to traditional empirical models, therefore being greatly useful for optimizing the future wireless network deployment. [3]
The eighth paper presents a deep-learning approach for predicting path loss over coastal and vegetation regions at sub-6 GHz 5G frequencies. The investigation is carried out for different environments on the seashore (dry sand, wet sand, small pebbles, large pebbles) and plant regions (pine, orange, cherry, and walnut trees) to define exact path loss models. Recurrent neural networks (RNN) and long short-term memory (LSTM) networks are trained for predicting the path loss and their performances compared with actual measured data collected at 3.5 GHz, 3.8 GHz, and 4.2 GHz. Results show that coastal path loss is generally greater than in vegetative areas with more accurate results from RNN models than LSTM models. The study aims at improving coverage planning and reducing wireless communication network optimization in multipath natural environments. [10]

After thoroughly examining the ITU-R recommendations on signal attenuation, we refined our understanding and solidified a clear roadmap to guide our thesis approach.
In this study, we take Bangladesh Satellite-1 and its ground base stations at

Gazipur and Betbunia into account. Collecting historical attenuation data from Bangladesh Communication Satellite Company Limited (BCSCL) and relative weather data from Bangladesh Meteorological Department (BMD), a comprehensive dataset has been formed. The dataset integrates cloud, atmospheric, rain, and free-space attenuation data with various meteorological parameters. In the estimation of total accumulated attenuation, two approaches are investigated: a predictive model based on machine learning and a mathematical model. Both models' accuracy and viability are compared to ascertain their efficacy in being used as substitutes for the conventional dummy signal method. This study not only helps in the enhancement of satellite communication in Bangladesh but also offers a model that can be applied in other parts of the globe that have similar climatic conditions. By having a robust path loss model, satellite communication service providers can render their networks more stable, simplify resource planning, and enhance overall service quality. The findings of this research will give valuable insights on how to mitigate atmospheric attenuation effects and offer uninterrupted satellite communication services in Bangladesh and beyond.

1.2 Motivation

The inspiration behind this research is the growing dependence on satellite communication for a number of important applications such as telecommunication, broadcasting, weather forecasting, and emergency services. It would be best to render these services more efficient, particularly in a developing nation such as Bangladesh, where satellite technology is emerging as a vital part of national infrastructure. One of the most essential satellite communications issues is attenuation of signals caused by atmospheric conditions. Conventional attenuation monitoring methods use dummy signals, which impose cost and complexity on satellite operations. By having a good predictive model, we can make these dummy signals obsolete, promoting operational efficiency and cost savings to satellite operators.

Besides, Bangladesh has diverse and extreme climatic conditions like high monsoons and humidity, which have a pervasive impact on signal transmission. A good path loss model will help mitigate the consequences of these challenges by providing real-time attenuation predictions, thus improved signal reliability and uninterrupted communication services.

The worldwide relevance of this study cannot be ignored since atmospheric attenuation is experienced by all satellite operators around the globe. The result of this study can be used as a reference for other regions of the world with the same meteorological climate, which will add to the general body of satellite communication and path loss modeling.

In general, this study is motivated by the demand for an inexpensive, dependable, and efficient means of enhancing satellite communication in Bangladesh and elsewhere. With the use of sophisticated machine learning techniques and mathematical modeling, we seek to establish a scientifically credible method of

attenuation prediction for the general goal of increasing the level of performance of satellite-based communication systems.

1.3 Objective

The main goal of this research is to create an effective and precise model for predicting satellite path loss from atmospheric attenuation. Based on machine learning approaches and mathematical modeling, this research intends to establish the most appropriate method to foresee total attenuation. The main aims of this research include:

- Collecting and merging BCSCL's historical attenuation data with BMD's weather data to create a comprehensive dataset.
- Building a machine learning model to predict total accumulated attenuation based on meteorological data.
- Creating a mathematical model for predicting attenuation and comparing the precision with the machine learning model.
- Comparing the validity and accuracy of both models for determining their viability in replacement of conventional dummy signal-based attenuation monitoring.
- Providing data that can be utilized for optimizing satellite communication networks, reducing operational costs, and enhancing the performance of Bangladesh Satellite-1 and other similar systems worldwide.

Through these goals, this research seeks to play a role in the creation of a more efficient and resource-saving method of monitoring satellite communication signal attenuation.

1.4 Scope of Work

The study's importance transcends the prediction of satellite path loss—it presents a novel and globally applicable method for optimizing the performance of satellite communications. Unlike other models needing location-dependent parameters like latitude and longitude, the mathematical model proposed in this work is not geographically limited. That is, wherever a ground station is positioned or whichever country it is in, the model is still valid.
Instead of being dependent on coordinates, the mathematical model only needs to be provided with typical weather parameters like rain rate, surface temperature, surface pressure, relative humidity, cloud base height, liquid water density in clouds, liquid water temperature in clouds, uplink frequency, polarization type, and path length. All these as inputs make the model deployable anywhere on the earth without any changes, which is a huge advantage when it comes to scalability and deployment ease.

By eliminating reliance on location-specific data, this model has the potential to be a globally viable solution for satellite communication path loss prediction. Even though improvements in accuracy are still possible, the groundwork for global usability is already there. This study, thus, not only has value in the enablement of satellite communication in Bangladesh but also makes a wider contribution by providing a model that can readily be incorporated into satellite networks worldwide.

Chapter 2

Background Work

2.1 General Path Loss Models

To understand and predict signal attenuation in wireless communication, we studied several **path loss models** that describe how the strength of a transmitted signal weakens with distance. Since wireless signals are subject to varying degrees of loss due to obstructions, reflections, and environmental conditions, multiple models help estimate actual behavior.

To compare these models, we ran **MATLAB simulations** using a 1.5 GHz (randomly selected) carrier frequency. The simulation compares three commonly used path loss models:

1. **Free-space path loss model** – perfect line-of-sight (LOS) propagation.
2. **Log-distance path loss model** – encompasses distance-dependent attenuation in different environments.
3. **Log-normal shadowing model** – encompasses random variations in signal strength due to obstacles.

All the models were run with mathematical equations for the evaluation of **signal attenuation over distance** considering different **antenna gains and path loss exponents**.

2.1.1 Free-space Path Loss Model

The **free-space path loss model** assumes that a transmitted signal propagates without any obstructions. The power loss in decibels (dB) is given by:

$$PL_{free}(d) = 20\log_{10}\left(\frac{4\pi d f_c}{c}\right) - G_t - G_r \tag{2.1}$$

where:

- d = Distance between transmitter and receiver (m)
- f_c = Carrier frequency (Hz)
- c = Speed of light (3×10^8 m/s)
- G_t = Transmitter antenna gain (dB)
- G_r = Receiver antenna gain (dB)

This model is useful for **satellite and open-area communication**, but it does not consider environmental factors like obstacles or reflections, making it **unrealistic for terrestrial communication**.

2.1.2 Log-distance Path Loss Model

In real-world environments, signal attenuation is influenced by the surroundings. The **log-distance path loss model** accounts for this using a path loss exponent (n):

$$PL_{logdist}(d) = PL(d_0) + 10n \log_{10}\left(\frac{d}{d_0}\right) \tag{2.2}$$

where:

- $PL(d_0)$ = Path loss at reference distance d_0
- d_0 = Reference distance (100 m in this study)
- n = Path loss exponent (e.g., 2 for free space, 3 for urban areas, 6 for highly obstructed environments)

The path loss exponent (n) varies based on terrain and building density. This model provides a **more practical estimation** than the free-space model but does not incorporate fading effects.

2.1.3 Log-normal Shadowing Model

In environments with obstacles, signal strength can fluctuate unpredictably. The **log-normal shadowing model** extends the log-distance model by adding a **Gaussian-distributed random variable** to represent signal variations due to shadowing:

$$PL_{lognorm}(d) = PL(d_0) + 10n \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \tag{2.3}$$

where:

- X_σ = A Gaussian random variable with standard deviation σ
- σ = 3 dB in our simulation

This model is more realistic for urban and obstructed environments as it introduces signal variation, but it requires **statistical modeling** and does not explicitly account for multipath fading.

2.1.4 Implementation in MATLAB

To analyze these models, we performed MATLAB simulations using:

- Carrier frequency: 1.5 GHz
- Reference distance: $d_0 = 100$ m
- Path loss exponents: $n = 2, 3, 6$
- Antenna gains: G_t and G_r with varying values

The distances were taken as squared values $(1, 4, 9, ..., 961)$ to better observe attenuation trends. Three separate plots were generated:

1. Free-space model with different antenna gains.
2. Log-distance model with varying path loss exponents.
3. Log-normal model showing shadow fading effects.

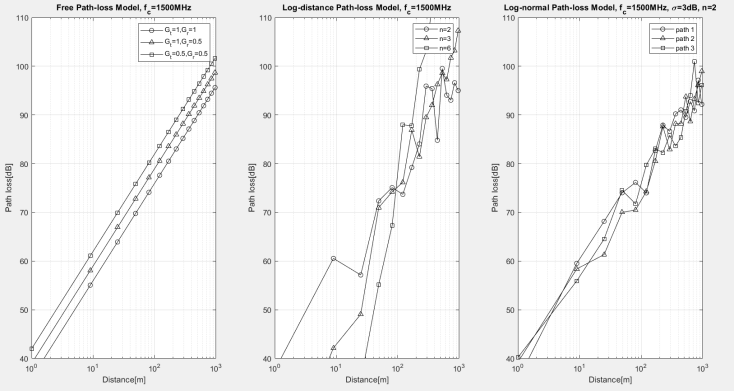


Figure 2.1: General Path Loss Model

2.1.5 Advantages and Limitations

Free-space Path Loss Model:

- Advantage: Provides a simple reference for signal attenuation.
- Limitation: Does not account for obstacles or environmental interference.

Log-distance Path Loss Model:

- Advantage: More adaptable for different environments by adjusting n.
- Limitation: Assumes a single path loss exponent, which may not always be accurate.

Log-normal Shadowing Model:

- Advantage: Captures random signal variations due to obstacles.
- Limitation: Requires statistical modeling and does not include multipath fading.

2.1.6 Practical Considerations

While these models provide useful insights, real-world wireless communication is affected by:

- Multipath fading due to reflections and interference.
- Weather effects such as rain and atmospheric conditions.
- Dynamic obstacles like moving vehicles and buildings.
- Antenna characteristics that affect radiation patterns.

For real-world applications, more advanced models such as Hata, COST-231, ITU-R, and Rayleigh/Rician fading models are used for network design and signal prediction.

2.2 Okumura/Hata Model

The **Hata model** is an empirical path loss model that is used for the prediction of signal attenuation in macrocellular environments (150 MHz – 1500 MHz). It estimates path loss in urban, suburban, and open areas based on the heights of the transmitter and receiver from the ground. Our thesis is about satellite communication, where antennas are not on the ground, so this model is not applicable for our research.

2.2.1 Hata Path Loss Model Equation

For urban situations, the formula is as follows:

$$PL_{urban} = 69.55 + 26.16 \log_{10}(f_c) - 13.82 \log_{10}(h_{tx}) - C_{Rx} + (44.9 - 6.55 \log_{10}(h_{tx})) \log_{10}(d/1000) \tag{2.4}$$

where:

- f_c = Carrier frequency (MHz)
- h_{tx} = Transmitter height (m)
- h_{rx} = Receiver height (m)
- d = Transmitter-to-receiver distance (m)

- C_{Rx} = Receiver correction factor (function of f_c)

Correction factors modify the path loss equation to account for suburban and open areas.

2.2.2 MATLAB Implementation

The MATLAB code:

- Computes path loss for urban, suburban, and open areas.
- Takes a carrier frequency of 1.5 GHz with htx = 30m, hrx = 2m.
- Plots results against logarithmic distance values.

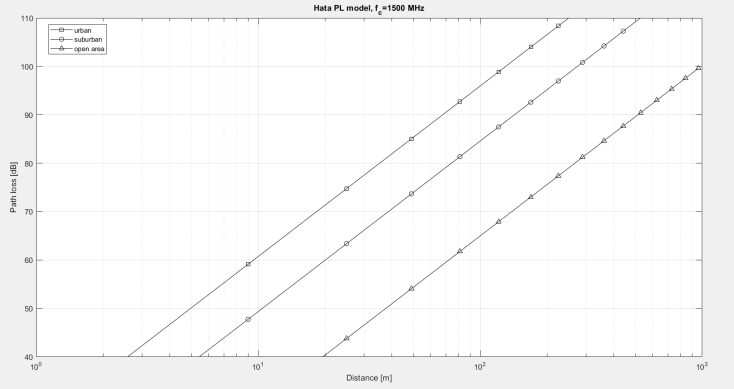


Figure 2.2: Okumura/Hata Model

2.2.3 Why Hata Model is Not Used in This Thesis

- Is derived from ground antenna heights, and hence cannot be used for satellite antennas.
- Is designed for use in terrestrial links, not space-based links.
- Does not take into consideration satellite path loss, where the free-space loss and atmospheric attenuation are more prominent.

Thus, alternative models like free-space path loss and satellite-specific models need to be used for our study.

2.3 IEEE 802.16d Model

The **IEEE 802.16d model** is an empirical path loss model used to predict signal attenuation for fixed broadband wireless systems. It classifies environments into three categories (A, B, C), based on propagation conditions, and incorporates shadowing corrections using factors from ATnT and Okumura models. Additionally, a modified version exists, which adjusts reference distance (d_0) based on correction factors.
Since this model requires ground-based antenna heights, it is not applicable for our thesis on satellite communication, where antennas are not ground-based.

2.3.1 IEEE 802.16d Path Loss Model Equation

The general form of the IEEE 802.16d model is:

$$PL(d) = A + 10\gamma \log_{10}\left(\frac{d}{d_0}\right) \tag{2.5}$$

where:

- $A = 20\log_{10}\left(\frac{4\pi d_0}{\lambda}\right) + PL_f + PL_h$ is the reference path loss at distance d_0.
- λ = Wavelength of the signal ($\lambda = \frac{c}{f_c}$).
- γ = Path loss exponent, given by:

$$\gamma = a - bh_{tx} + \frac{c}{h_{tx}} \tag{2.6}$$

with values depending on the environment type:

- Type A: $a = 4.6, b = 0.0075, c = 12.6$
- Type B: $a = 4, b = 0.0065, c = 17.1$
- Type C: $a = 3.6, b = 0.005, c = 20$

Correction Factors for Shadowing:

- ATnT Correction: $PL_f = 6\log_{10}(f_c/2000)$, $PL_h = -10.8\log_{10}(h_{rx}/2)$.
- Okumura Correction:

$$PL_h = \begin{cases} -10\log_{10}(h_{rx}/3), & h_{rx} \leq 3m \\ -20\log_{10}(h_{rx}/3), & h_{rx} > 3m \end{cases} \tag{2.7}$$

A modified version of IEEE 802.16d adjusts the reference distance:

$$d_0' = d_0 \times 10^{-\frac{(PL_f + PL_h)}{10\gamma}} \tag{2.8}$$

where d_0' is the adjusted reference distance.

2.3.2 MATLAB Implementation

The MATLAB code:

- Computes path loss for Type A environments using ATnT correction.
- Compares the original and modified IEEE 802.16d models.
- Uses carrier frequency = 2 GHz, transmitter heights = 30m, and receiver heights = {2m, 10m}.
- Plots results against logarithmic distance.

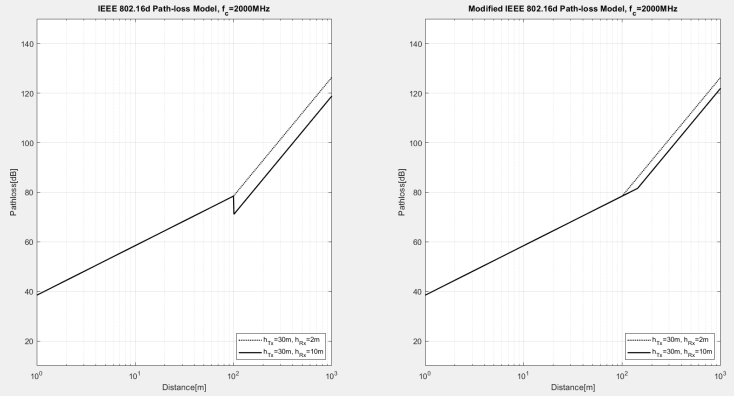


Figure 2.3: IEEE 802.16d Model

2.3.3 Why IEEE 802.16d Model is Not Used in This Thesis

- Requires ground-based antenna heights, which are not relevant for satellites.
- Designed for terrestrial fixed wireless systems, not space-to-ground communication.
- Does not consider satellite-specific losses, such as atmospheric attenuation or free-space propagation.

Since this model is unsuitable for satellite networks, alternative satellite-specific models are needed for our study.

2.4 Integration of ITU-R Propagation Models

2.4.1 Significance of ITU-R Propagation Models

Accurate prediction of satellite signal attenuation is crucial for developing effective and dependable telecommunication systems. Attenuation affects the strength, quality, and overall performance of signals and systems. The International Telecommunication Union Radiocommunication Sector (ITU-R) provides standardized models with accurate methodologies to estimate signal degradation caused by atmospheric effects. The three ITU-R models discussed in this section—**ITU-R P.838-3 (Rain Attenuation), ITU-R P.840-9 (Cloud and Fog Attenuation), and ITU-R P.676-13 (Atmospheric Gases Attenuation)**—are the core of satellite communication system planning. They ensure consistency and reliability in prediction models, resulting in the establishment of robust communication links.

Microwave and millimeter-wave satellite signals suffer severe degradation due to atmospheric absorption, scattering, and reflection. The most significant causes of signal attenuation are rain, cloud, fog, and atmospheric gases, and these must be accounted for in any satellite path loss calculation. The ITU-R models give empirical and analytical approaches to the computation of these losses and are therefore a necessity for engineers and researchers.

2.4.2 Explanation of Each ITU-R Model

ITU-R P.838-3: Model of Rain Specific Attenuation

Rain is the most powerful cause of attenuation of the signal in satellite communication, especially at frequencies above 10 GHz. **ITU-R P.838-3** provides a model to estimate rain attenuation using the power-law relation:

$$\gamma_R = kR^\alpha \tag{2.9}$$

Where:

- γ_R is the specific attenuation (dB/km),
- R is the rain rate (mm/h),
- k and α are frequency-dependent empirical constants.

The model encompasses polarization effects by distinguishing between horizontal and vertical polarizations. The model also provides coefficient values as a result of detailed scattering computations for precise prediction over a wide range of frequencies (1 GHz to 1000 GHz). The model plays a crucial role in **link budget analysis, fade margin design, and adaptive modulation schemes** in high-frequency satellite links.

ITU-R P.840-9: Attenuation in Cloud and Fog

Cloud and fog add additional attenuation, primarily at frequencies greater than 10 GHz. **ITU-R P.840-9** provides a formula for the estimation of signal loss due to cloud and fog as a function of integrated liquid water content (LWC). The model uses the following formula for attenuation:

$$\gamma_c(f, T) = K_l(f, T) \cdot \rho_l \tag{2.10}$$

Where:

- γ_c is the specific attenuation (dB/km),
- $K_l(f, T)$ is particular cloud liquid water attenuation coefficient,
- ρ_l is cloud or fog liquid water density (g/m^3),
- f is frequency (GHz),
- T is temperature (Kelvin).

The model can be applied within a frequency band of **1 GHz to 200 GHz** and is therefore essential in the design of **Earth-space links, high-frequency terrestrial microwave systems, and next-generation satellite constellations**. Because cloud attenuation is meteorology dependent, the model further incorporates statistical distributions and global climatological information, enhancing its predictive power.

ITU-R P.676-13: Attenuation by Atmospheric Gases

Atmospheric gases like **oxygen and water vapor** absorb electromagnetic waves, causing additional path loss. **ITU-R P.676-13** provides procedures for calculating **slant path gaseous attenuation**, considering:

- **Oxygen and water vapor absorption** due to molecular resonance effects.
- **Frequency-dependent absorption coefficients** of different atmospheric layers.
- **Correction factors for varying atmospheric conditions (pressure, temperature, and humidity).**

The model provides two primary methods:

1. **Spectral line-by-line calculations** for highly accurate attenuation estimation from individual spectral lines.
2. **Empirical fits** for quick engineering calculations in engineering applications.

The model has a very wide frequency range (**1 GHz to 1000 GHz**), and therefore it is very useful for **radio astronomy, remote sensing, and high-frequency satellite communications**. It also accounts for path elevation angles, which is crucial for **geostationary and low Earth orbit (LEO) satellite systems**.

2.4.3 Reasoning Behind the Use of These Models

Use of ITU-R models in satellite path loss prediction ensures that the estimates of attenuation are:

1. **Scientifically Validated**: The models have been developed through extensive experimental and theoretical work, hence reliable.
2. **Frequency-Dependent**: Each model is specific to specific frequency bands, so they can be used across various bands utilized in satellite communications.
3. **Geographically Adaptive**: Global meteorological datasets are used in the models, which allows location-based predictions.
4. **Global Standardized Use**: ITU-R recommendations are commonly adopted, allowing interoperability among various communication systems and compliance with the regulations.

All the models are used for specific purposes:

- **ITU-R P.838-3** is critical for **rain attenuation predictions**, which have a great influence on satellite links in tropical and high-rainfall environments. [5]
- **ITU-R P.840-9** deals with **cloud and fog attenuation**, which is important for systems that function in humid environments and high-frequency bands. [6]
- **ITU-R P.676-13** also incorporates **gaseous attenuation**, important to apply to precise signal strength calculations, particularly in long-haul satellite transmissions. [7]

Through their application, these models enable satellite communication systems to more effectively take advantage of **power budgets, antenna designs, frequency selection, and adaptive coding schemes**, leading eventually to a better link reliability and overall performance.

2.5 Selection of the Goff-Gratch Formula for Saturation Vapor Pressure Calculation

Accurate calculation of saturation vapor pressure (e_s) is crucial in climate modeling, hydrology, and atmospheric sciences with a high priority in high-latitude

cold regions where temperature variation significantly influences vapor pressure calculations. In our thesis, we have employed the **Goff-Gratch formula** since it is more accurate than other available formulas. The selection is based on the comparison conducted by **Xu et al. (2012)** among various formulas, i.e., **Teten, Magnus, Buck, and Goff-Gratch**, and their performances according to the variations of temperature. [9]

It was concluded in the research that while the Teten formula, being popular in application to evapotranspiration modeling (e.g., the FAO-56 Penman-Monteith equation), is good for the moderate temperature range (0°C to 40°C), it gives grave errors at cold temperatures. Teten's error in computed e_s increases proportionally to a decrease in temperature up to greater than 40% at -40°C. This leads to enormous errors in computing the vapor pressure deficit (VPD) and subsequently in estimating the reference evapotranspiration (ET_0).

In contrast, the Goff-Gratch formula, officially recommended by the World Meteorological Organization (WMO), provides a better estimation of saturation vapor pressure across a broad range of temperatures and is therefore the preferred option for studies dealing with extreme weather events.

2.5.1 Mathematical Representation of the Goff-Gratch Formula

The **Goff-Gratch equation** for saturation vapor pressure is expressed as follows:
For $T > 273.16K$:

$$\log e_s = a_1 \left(\frac{T_{01}}{T} - 1\right) + b_1 \log \frac{T_{01}}{T} + c_1 \left(10^{d_1(1-T/T_{01})} - 1\right) + e_1 \left(10^{f_1(T_{01}/T-1)} - 1\right) + \log g_1 \tag{2.11}$$

For $T < 273.16K$:

$$\log e_s = a_2 \left(\frac{T_{02}}{T} - 1\right) + b_2 \log \frac{T_{02}}{T} + c_2 \left(1 - \frac{T}{T_{02}}\right) + \log d_2 \tag{2.12}$$

where the constants are:

$$\begin{aligned}
&a_1 = -7.90298, \quad && b_1 = 5.02808, \quad && c_1 = 1.3816 \times 10^{-7}, \quad && \\
&d_1 = 11.344, \quad && e_1 = 8.1328 \times 10^{-3}, \quad && f_1 = -3.49149, \quad && g_1 = 1013.246, \\
&a_2 = -9.09718, \quad && b_2 = -3.56654, \quad && c_2 = 0.876793, \quad && d_2 = 6.1071, \\
&T_{01} = 373.16, \quad && T_{02} = 273.16. && && \\
\end{aligned}$$

These equations take into account the phase transitions of water vapor and provide highly accurate results across a broad temperature range.

2.5.2 Comparison with Alternative Formulas

The study by Xu et al. (2012) also evaluated other formulas:

- **Teten Formula**: Commonly used in FAO-56 Penman-Monteith equation:

$$e_s = 0.611 \times \exp\left(\frac{17.27t}{t + 237.3}\right) \tag{2.13}$$

  While computationally simple, it shows high errors at low temperatures.

- **Magnus Formula**: Offers improved accuracy by differentiating between $t > 0$°C and $t < 0$°C:

$$e_s = 6.11 \times 10^{\frac{7.45t}{t+237.3}}, \quad t > 0°C \tag{2.14}$$

$$e_s = 6.11 \times 10^{\frac{9.5t}{t+265.5}}, \quad t < 0°C \tag{2.15}$$

- **Buck Formula**: Similar to Magnus but modified for improved performance:

$$e_s = 6.1121 \times \exp\left(\frac{18.678 - \frac{t}{234.5}}{257.14 + t} \times t\right), \quad t > 0°C \tag{2.16}$$

$$e_s = 6.1115 \times \exp\left(\frac{23.306 - \frac{t}{333.7}}{279.82 + t} \times t\right), \quad t < 0°C \tag{2.17}$$

Although the Buck and Magnus formulas provide reasonable approximations, the **Goff-Gratch formula remains the most authoritative** due to its basis in fundamental thermodynamic principles and experimental validation.

2.5.3 Conclusion

Based on the findings of Xu et al. (2012), the Goff-Gratch formula has been selected in this study due to its superior accuracy across a wide temperature range and its global acceptance as a standard method. Its ability to maintain precision under extreme cold conditions makes it particularly suitable for high-latitude climate analysis, ensuring reliable estimation of e_s and its associated impact on evapotranspiration calculations.

2.6 Machine Learning for Predicting Path Loss

2.6.1 Introduction and Objectives of the Paper

To see which model should be used to better predict the path loss, we read a paper on this named "Path Loss Prediction Based on Machine Learning Techniques: Principle Component Analysis, Artificial Neural Network, and Gaussian Process". This work explores the application of machine learning techniques for path loss prediction in wireless sensor networks. Classical models, such as linear log-distance models, can be inadequate for capturing the subtleties of environments. With the aim of improving the precision of predictions, the authors develop a machine learning-based framework that employs:

- **Principal Component Analysis (PCA)** for dimensionality reduction and feature selection.
- **Artificial Neural Networks (ANN)** for multi-dimensional regression and learning the structure of path loss.
- **Gaussian Process (GP)** for shadowing effect modeling and variance analysis.

The primary motivation behind this research is to develop a more flexible and accurate path loss model compared to conventional empirical approaches. By leveraging machine learning, the authors aim to create a model that better adapts to real-world variations in signal propagation, particularly in suburban environments.

2.6.2 Machine Learning Framework for Path Loss Prediction

The paper outlines a three-step process for predicting path loss using machine learning techniques:

1. **Feature Selection using PCA:**
   - The path loss dataset typically consists of multiple features, such as **distance, antenna height, and frequency**.
   - PCA is used to reduce the dataset's dimensionality by identifying the most influential features, minimizing redundancy, and improving computational efficiency.
   - The authors find that **log distance and log frequency** contribute to more than 70% of path loss variation.
2. **ANN-Based Regression Model:**
   - After dimensionality reduction, an **Artificial Neural Network (ANN)** learns the path loss pattern.
   - A multi-layer perceptron (MLP) architecture is used, with different activation functions (ReLU, sigmoid, and hyperbolic tangent) tested for performance evaluation.
   - The ANN-MLP model aims to predict **mean path loss values** from the dataset's reduced features.
3. **Gaussian Process for Shadowing Effects:**
   - The shadowing effect, caused by obstacles that interfere with the signal, is modeled using a **Gaussian Process (GP)**.
   - GP is used to estimate the **variance** in path loss predictions, allowing the model to quantify uncertainty in real-world scenarios.
   - This step helps improve the reliability of the path loss model by incorporating probabilistic reasoning.

2.6.3 Measurement System and Experimental Setup

The dataset used for model training and evaluation was collected in suburban areas of Korea with various environmental factors. Measurements were taken at three different frequencies (450 MHz, 1450 MHz, and 2300 MHz), with the following setup:

- **A fixed transmitter** placed on a four-story building with an antenna height of 15 meters.
- **A mobile receiver** mounted on a vehicle with an antenna height of 2 meters.
- **Path loss data collected** over distances ranging from 1 km to 3.5 km.
- **Preprocessing techniques** such as normalization and cross-validation applied to improve model robustness.

2.6.4 Key Findings and Performance Evaluation

The paper evaluates the proposed model using several performance metrics, including:

- Root Mean Squared Error (RMSE)
- Mean Absolute Error (MAE)
- Mean Absolute Percentage Error (MAPE)
- Coefficient of Determination (R^2)

Key observations include:

- The ANN-MLP model performs better than the classic log-linear models and two-ray models with results providing the smallest RMSE and highest R^2.
- PCA-based feature selection reduces the computation time without compromising on the prediction accuracy.
- Gaussian Process modeling results in better confidence intervals for shadowing effects, thereby allowing more dependable path loss estimation.
- With ANN, the sigmoid activation function resulted in the best performance.

2.6.5 Our Findings

Inspired by this study, we tried to **c**ouple PCA with ANN in our own investigation for path loss prediction. But we could not achieve satisfactory results. Some reasons for this may be:

- Trade-offs for Feature Reduction: PCA may have eliminated important features which have caused the most variations in the path loss.
- Nonlinearity of Path Loss Data: While PCA captures linear correlations, it may not well handle nonlinear dependencies in the data.
- Overfitting in ANN: The high variance nature of ANN models requires a careful balance of training data size, network architecture, and regularization techniques.
- Environmental Differences: The dataset on which the paper was conducted was gathered in a particular suburban region of Korea, while my experiment was carried out on some other terrain, environmental factors, or frequency bands.
- The Work not being done on Satellite: The dataset on which the paper was conducted was not for StG (Satellite to Ground) communication path loss. Since the Bangladesh Satellite-1 is in Line-of-Sight for both the ground stations, we excluded the Gaussian Process for shadowing effects. After all this adjustments, however, this process did not perform well.

2.6.6 Conclusion: Accepting the Study but Moving Ahead with Other Methods

Although the methodology is well-structured and innovative as proposed in the paper, my practical application of PCA with ANN did not give good results. So it may appear to be indicating that:

- Feature selection strategies should be carefully assessed to ensure essential information is retained.
- ANN models may not always benefit from PCA, especially if the dataset's nonlinear characteristics require direct learning from raw features.
- Alternative feature engineering techniques, such as autoencoders or deep learning-based dimensionality reduction, may be more effective for complex path loss modeling.

In conclusion, this paper provides valuable insights into machine learning-based path loss prediction. However, based on my findings, PCA and ANN together may not always be the best approach for every dataset and environment. Further exploration of nonlinear feature selection methods and adaptive learning techniques may yield better results in future work.

Chapter 3

Methodology

The proposed methodology of the study contains two approaches: a machine learning-based path loss prediction model and a mathematical modeling approach based on the ITU-R recommendations. The objective is to analyze and compare the performance of these two techniques to find out which one provides the most accurate and reliable path loss estimates. Both approaches were previously tested and refined to realistically mimic field conditions for an objective assessment of their effectiveness in predicting satellite communication path loss.

3.1 Machine Learning Approach

3.1.1 Dataset Preparation

One of the toughest but crucial aspects of this work was preparing the dataset. Initially, obtaining real satellite attenuation data was extremely challenging as it was not readily available. To circumvent this limitation, we first generated a synthetic dataset, attempting to mimic realistic attenuation patterns from theoretical models and environmental parameters that we understand. This synthetic dataset gave us a starting point to evaluate different machine learning models.

We began by attempting to train an Artificial Neural Network (ANN) model on this synthetic data set after applying Principal Component Analysis (PCA) to decrease the dimensionality. The outcomes using this approach were not satisfactory. The ANN model failed to generalize, which meant the synthetic data set likely did not capture all the complexity of real-world satellite signal attenuation. To handle this, we experimented with several other machine learning models, each designed to address different aspects of the prediction issue. The idea was that if a model performed well on the synthetic data, it might perform even better when we had actual real-world data.

We were eventually able to procure an actual dataset of satellite attenuation measurements spanning several years from the Bangladesh Communica-

tion Satellite Company Limited (BCSCL). These logs had the attenuation values measured on specific dates, providing us with useful historic information for our study. Attenuation alone was not sufficient for the purpose of making accurate predictions, however, as it is significantly meteorology-dependent. Therefore, we also gathered weather data for the respective dates from the Bangladesh Meteorological Department (BMD). The weather database had crucial atmospheric parameters such as temperature, humidity, precipitation, and cloud cover—parameters known to influence signal attenuation.
Having obtained both datasets, the initial step was to merge them into a single complete dataset. This involved concatenating the meteorological data and attenuation measurements based on timestamps to align them. Preprocessing was carried out carefully to handle missing values, eliminate inconsistencies, and normalize feature scales for improved model performance.
With this real dataset, we were able to re-train and re-test our models to get more accurate and realistic predictions of satellite path loss. The incorporation of real attenuation and meteorological data added a lot of credibility to our machine learning-based predictions, demonstrating the benefit of utilizing real-world measurements over synthetic approximations.

3.1.2 Model Performance Analysis

In our study, we attempted a number of machine learning models for path loss prediction. The following sections report attempts conducted by us, including the methods taken and, most significantly, all improvements accomplished and results achieved.

Attempt 01: ANN with PCA

We employed a machine learning pipeline that utilizes PCA for dimensionality reduction and an ANN for regression to predict Path Loss (dB) from various geographic and environmental parameters.

PCA is a technique that transforms correlated features into a lower dimensional space of uncorrelated principal components with maximal variance. It assists in noise reduction and improving computational efficiency. ANN, which is inspired by the human brain, is used in classification and regression tasks. It consists of layers of neurons that function on inputs through weighted connections, activation functions, and backpropagation to minimize errors.

Despite our efforts, this approach did not yield desirable outcomes. The final Mean Squared Error (MSE) was 78.35.

Attempt 02: Improved ANN with Regularization

To improve our model further, we incorporated regularization techniques, dropout layers, and early stopping to prevent overfitting and improve generalization.

Key Improvements:

- Regularization (L2), dropout layers, and early stopping to prevent overfitting.
- More controlled training process with early stopping, limiting excess epochs.
- Added scatter plots for actual vs. predicted comparison.

These modifications led to a reduction in MSE to 64.12.

Attempt 03: Hybrid Approach with Voting Regressor

Trying to improve our model further, we hybridized ANN, Random Forest, and a Voting Regressor with PCA, a hybrid machine learning model.

Key Enhancements:

- Combined ANN with Random Forest for better generalization.
- Random Forest performed well in non-linear relationship capturing.
- ANN was hyperparameter-tuned using ReduceLROnPlateau to dynamically change learning rates.
- Voting Regressor combined both model strengths, reducing bias and variance.
- Enhanced visualization with scatter plots and learning curves.

The findings show that Random Forest was way better than ANN. In fact, the optimal MSE in the model for this issue was 8.63 for Random Forest, while the optimal MSE for ANN was 93.45.

Attempt 04: Random Forest Model

In this attempt, we abandoned ANN and ensemble learning, and we only used the Random Forest Regressor (RF) as our main model.

Key Takeaways:

- Simpler approach with quicker training and assessment.
- Simpler hyperparameter tuning (i.e., number of estimators, max depth, etc.).
- Feature scaling is not required, which simplifies preprocessing.
- Random Forest showed better performance on structured tabular data.

This last method achieved somewhat less error than previously, which cemented Random Forest as the highest-performing model in our research.

Attempt 05: Enhanced Random Forest Model

Our new Random Forest model has significant changes and improvements over the previous version. We have significantly expanded our feature set, added categorical encoding, and adjusted the PCA components to match the increased input dimensions. We worked with a much larger synthetic dataset this time, which is a great step toward dealing with real data. Because the real dataset will have the same features.

Aspect	**Previous Model**	**Updated Model**
Number of Features	6	26+ (including environmental and satellite parameters)
Categorical Encoding	Not included	One-hot encoding for categorical features
PCA Components	6	27 (to match the expanded feature space)

Table 3.1: Comparison between the previous and updated Random Forest models

Comparative Analysis:

Key Takeaways:

- **More features** → Potentially higher accuracy
- **Categorical encoding** → Better handling of non-numeric features
- **More PCA components** → Retains more variance

This time, the efficiency stood around **73%**.

Attempt 06: Introducing XGBoost alongside Random Forest

Our latest model introduces XGBoost (`XGBRegressor`) alongside Random Forest (RF), enabling a comparative analysis between the two models. From this attempt and on, we used the real dataset that we prepared from BCSCL and BMD.

Comparative Analysis: Random Forest vs. XGBoost Model

Key Takeaways:

- **XGBoost Introduced**: More advanced boosting model to improve predictive accuracy.

Feature	Previous Model (RF only)	Current Model (RF + XGBoost)	Advancements
Model Types	Random Forest (RF)	RF + XGBoost (`XGBRegressor`)	Added Gradient Boosting Model
RF Hyperparameters	`n_estimators=100`	Same	No change
XGBoost Added	Not present	`n_estimators=100`, `random_state=42`	New model for boosting performance
Performance Metrics	Only RF results	RF vs. XGB comparison	Enables direct performance evaluation
Visualizations	RF predictions only	Separate plots for RF & XGB	Provides clear model comparison
Computational Cost	Moderate	Higher (XGB takes longer than RF)	XGB requires tuning

Table 3.2: Comparison between the Random Forest and XGBoost models

- **Direct Model Comparison**: Helps determine if XGBoost outperforms RF.
- **Same Preprocessing Pipeline**: Ensures a fair comparison.

Through this experiment, we observed that XGBoost outperformed the Random Forest model in terms of accuracy. XGBoost achieved 82% accuracy, whereas Random Forest provided 73% accuracy.

Attempt 07: Optimized XGB with Stacking

Our updated Current Model introduces Optimized XGB with Stacking.
Comparative Analysis-

Aspect	Previous Code	Updated Code
Models Used	Random Forest, XGBoost	Random Forest, Optimized XGBoost, Stacking Regressor
Hyperparameter Tuning	None	Added RandomizedSearchCV for XGBoost
Feature Importance	Not included	Visualized feature importance (with/without PCA)
Dimensionality Reduction	Fixed PCA component count	Dynamic PCA component count based on feature availability
Ensemble Learning	Independent models	Stacked ensemble with Linear Regression as meta-learner
Result Visualization	Separate plots for each model	Visual comparisons for optimized and stacked models

Table 3.3: Comparative Analysis of Previous vs. Updated Model

What This Model Does Better

- **Optimized XGBoost**: RandomizedSearchCV improves learning rate, depth, and estimators.
- **Feature Importance Analysis**: Helps understand which variables drive predictions.
- **Stacking Regression**: Combines Random Forest + XGBoost, leading to higher accuracy.

Performance Comparison-

Model	Mean Squared Error (MSE)	R^2 Score
Initial XGBoost	mse_xgb	r2_xgb
Optimized XGBoost	mse_best_xgb	r2_best_xgb
Stacked Model (RF + XGB + Linear Regression)	mse_stacked	r2_stacked

Table 3.4: Performance Comparison of Models

The initial XGBoost Model used to offer 89.62% accuracy, whereas this model exceeds it and stands at 90.62% accuracy.

Attempt 08: Optimized Random Forest Model

This time, we have gone back to the Random Forest Model with some noticeable changes. The model has been significantly enhanced through several key optimizations. Initially, a standard Random Forest model using raw features and default settings was employed. The current optimized model now incorporates one-hot encoding for categorical data, tuned hyperparameters via RandomizedSearchCV, and feature importance visualization, resulting in improved accuracy and a better understanding of key drivers, while maintaining standard performance tracking with MSE and R^2.

Comparative Analysis-

Aspect	**Previous Code (XGBoost + Stacking)**	**Current Code (Optimized Random Forest)**
Model Types	XGBoost, Random Forest, Stacking	Random Forest only
Optimization	RandomizedSearchCV for XGBoost	RandomizedSearchCV for Random Forest
Dimensionality Reduction	PCA applied	No PCA (uses all features with feature importance)
Feature Importance	Based on PCA components	Based on original features (more interpretable)
Model Complexity	Higher (ensemble stacking)	Lower (single optimized model)
Training Speed	Slower (due to stacking and XGBoost)	Faster (Random Forest is quicker to train)
Result Visualization	Plots for both models (XGBoost + Stacking)	Single plot for Random Forest predictions

Table 3.5: Comparative Analysis of Previous vs. Current Model

What This Model Does Well

- **RandomizedSearchCV Optimization**: Tunes key hyperparameters for better model performance.
- **Feature Importance Analysis**: Identifies the most significant predictors in the model.
- **Visual Performance Evaluation**: Displays scatter plot for actual vs. predicted values, aiding in performance visualization.

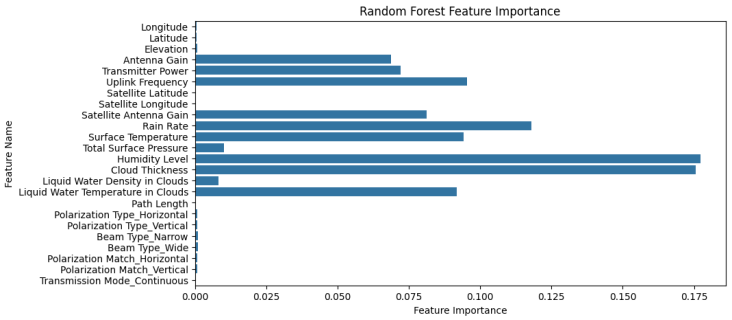


Figure 3.1: Features that have actual impact on the attenuation

Performance Comparison-

Model	**Mean Squared Error (MSE)**	**R^2 Score**
Initial RF Model	mse_rf	r2_rf
Optimized RF Model	mse_best_rf	r2_best_rf

Table 3.6: Performance Comparison of Initial and Optimized RF Models

The optimized RF model should outperform the initial RF model in terms of lower MSE and higher R^2.

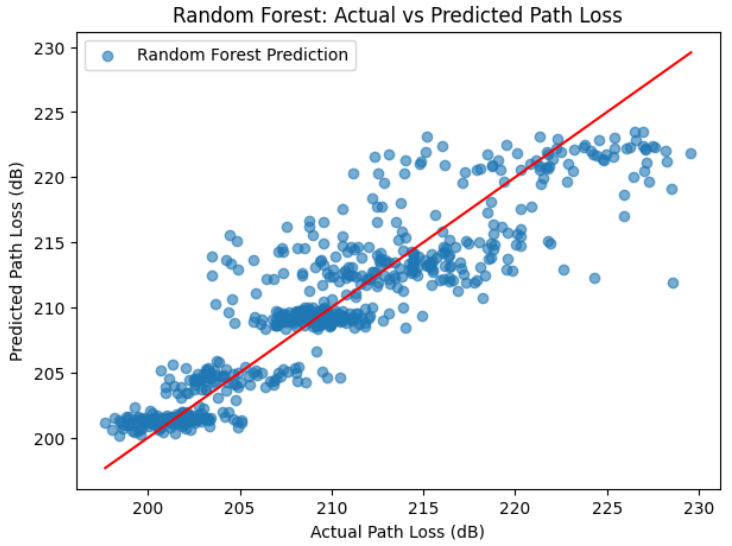


Figure 3.2: Performance of the optimized Random Forest Model

This optimized RF model has shown an impressive 91.36% accuracy. This much

high accuracy is for training the model with the "Free Space Path Loss", not for other attenuations. Our thought was- if a model does not perform well for the Free Space Path Loss prediction, then it will have no chance for the Total Attenuation. After training the model with Total Attenuation, the accuracy rate has dropped down to 80.432%, which is actually promising.

3.2 Mathematial Approach

3.2.1 Rain Attenuation Calculation

Inputs of the function

The function computes the rain attenuation based on the ITU-R P.838-3 model using the following inputs:

- **rain_rate** (R): Rain rate in mm/h.
- **frequency** (f): Frequency in GHz.
- **polarization**: Either `'horizontal'` or `'vertical'`.
- **theta** (θ): Elevation angle (in degrees).
- **L**: Physical path length (km).

The output of the function is:

- **attenuation**: Computed rain attenuation in dB.

Processing the Polarization Input

The polarization input is converted to lowercase and stripped of extra spaces to ensure uniform processing:

```
polarization = lower(strtrim(polarization));
```

Calculation of k and α Based on Polarization

(a) Horizontal Polarization If `polarization == 'horizontal'`, the function calculates the specific attenuation coefficients k_H and α_H.
Step 1: Initial Constants

```
mk=-0.18961; ma=0.67849;
ck=0.71147; ca=-1.95537;
k_log=mk*log10(frequency)+ck;
alpha=ma*log10(frequency)+ca;
```

These are empirical coefficients from ITU-R P.838-3 (Table 1 and Table 3) used to approximate k_H and α_H.

Step 2: Compute k_H Using Equation

$$\log_{10} k_H = m_k \log_{10} f + c_k + \sum_{j=1}^{4} a_j \exp\left(-\left(\frac{\log_{10} f - b_j}{c_j}\right)^2\right) \tag{3.1}$$

```
for i=1:4
    if i==1
        a(i)=-5.33980; b(i)=-0.10008; c(i)=1.13098;
    elseif i==2
        a(i)=-0.35361; b(i)=1.26970; c(i)=0.45400;
    elseif i==3
        a(i)=-0.23789; b(i)=0.86036; c(i)=0.15354;
    elseif i==4
        a(i)=-0.94158; b(i)=0.64552; c(i)=0.16817;
    end
    k_log = k_log + a(i) * exp(-((log10(frequency) - b(i)) / c(i))^2);
    k = 10^k_log;
end
```

Step 3: Compute α_H Using Equation

$$\alpha_H = m_\alpha \log_{10} f + c_\alpha + \sum_{j=1}^{5} A_j \exp\left(-\left(\frac{\log_{10} f - B_j}{C_j}\right)^2\right) \tag{3.2}$$

```
for j=1:5
    if j==1
        A(j)=-0.14318; B(j)=1.82442; C(j)=-0.55187;
    elseif j==2
        A(j)=0.29591; B(j)=0.77564; C(j)=0.19822;
    elseif j==3
        A(j)=0.32177; B(j)=0.63773; C(j)=0.13164;
    elseif j==4
        A(j)=-5.37610; B(j)=-0.96230; C(j)=1.47828;
    elseif j==5
        A(j)=16.1721; B(j)=-3.29980; C(j)=3.43990;
    end
    alpha = alpha + A(j) * exp(-((log10(frequency) - B(j)) / C(j))^2);
end
```

(b) Vertical Polarization For `polarization == 'vertical'`, the same method is applied using vertical polarization coefficients k_V and α_V from ITU-R P.838-3 (Table 2 and Table 4).

Calculation of Effective Path Length L_{eff}

The effective path length accounts for the impact of the rain structure and is computed using:

$$L_{\text{eff}} = \frac{L}{1 + 0.78 \left(\frac{Lf}{\sin\theta}\right)^{0.5} - 0.38(1 - e^{-2L})} \tag{3.3}$$

```
L_eff= L/(1+ 0.78*(L*frequency/sin(theta))^(0.5) -0.38*(1-exp(-2*L)));
```

Specific Attenuation γ Calculation

The ITU-R P.838-3 power-law model for specific attenuation is:

$$\gamma = kR^{\alpha} \tag{3.4}$$

```
gamma = k * (rain_rate^alpha);
```

Final Rain Attenuation Calculation

The total rain attenuation is computed as:

$$A = 0.012 \times \gamma \times L_{\text{eff}} \tag{3.5}$$

```
attenuation=0.012*gamma*L_eff;
```

0.012 is the regulating factor which we used to match with the real data.

Summary of the Function

1. Determines the polarization mode (`horizontal` or `vertical`).
2. Computes coefficients k and α using ITU-R P.838-3 curve-fitted equations.
3. Adjusts the path length L_{eff} based on frequency and elevation angle.
4. Uses the power-law model to compute specific attenuation γ.
5. Computes the total rain attenuation over the effective path.

This function is useful in predicting rain attenuation effects for wireless communications, satellite links, and microwave radio systems.

3.2.2 Atmospheric Attenuation Calculation

Atmospheric attenuation refers to the reduction in signal strength as electromagnetic waves propagate through the Earth's atmosphere. This phenomenon is primarily caused by absorption and scattering due to atmospheric gases, including oxygen and water vapor. Understanding these effects is crucial for applications such as satellite communications, remote sensing, and radio astronomy.

This study focuses on two key contributors to atmospheric attenuation:

1. **Oxygen Attenuation** - Calculated using the *Instantaneous Prediction Method*.
2. **Water Vapor Attenuation** - Derived from the *Slant Path Instantaneous Water Vapor Gaseous Attenuation Prediction Method*.

Oxygen Attenuation

Oxygen molecules in the atmosphere absorb electromagnetic waves primarily in the microwave and millimeter-wave frequency ranges. The attenuation is due to resonance absorption at specific frequencies, with the strongest absorption occurring near 60 GHz.

Mathematical Model for Oxygen Attenuation The specific attenuation due to oxygen, denoted as γ_o, is calculated using:

$$\gamma_o = 0.1820 f N''_{oxygen}(f) \tag{3.6}$$

where:

- f is the frequency in GHz.
- $N''_{oxygen}(f)$ is the imaginary part of the complex refractivity due to oxygen.

The total path attenuation is determined by:

$$A_o = \frac{\gamma_o h_o}{\sin(\theta)} \tag{3.7}$$

where:

- h_o is the oxygen equivalent height.
- θ is the elevation angle of the signal path.

MATLAB Implementation for Oxygen Attenuation The oxygen attenuation calculation is implemented using several MATLAB functions:

- **N_DoublePrime_oxygen_Calculation.m** - Computes $N''_{oxygen}(f)$ by summing spectral line contributions from resonance absorption:

```
function N_DoublePrime_oxygen=N_DoublePrime_oxygen_Calculation(f,p,e,theta)
    format long g
    theta=theta*pi/180;
    % Compute spectral line contributions
    % Uses predefined resonance data tables
    % Incorporates Zeeman splitting
end
```

- **N_DoublePrime_D_oxygen_Calculation.m** - Accounts for pressure-induced nitrogen absorption effects:

```
function N_doublePrime_D_oxygen=N_DoublePrime_D_oxygen_Calculation(f,p,e,theta)
    format long g
    theta=theta*pi/180;
    d=5.6*(10^-4)*(p+e)*(theta^0.8);
    alu=(6.14*(10^-5))/(d*(1+ ((f/d)^2)  ));
    kochu=(1.4*(10^-12)*p*(theta^1.5))/(1+ 1.9*(f^1.5)*(10^-5));
    N_doublePrime_D_oxygen=f*p*(theta^2)&(alu+kochu);
end
```

- **cofficients_for_h_oxygen.m** - Interpolates coefficients for oxygen equivalent height:

```
function h=cofficients_for_h_oxygen(frequency,Ts,Ps,row_ws)
    format long g
    % Interpolation logic for equivalent height calculation
end
```

- **Oxygen_Attenuation.m** - Computes total oxygen attenuation:

```
function Attenuation=Oxygen_Attenuation(frequency,Ts,Ps,row_ws,p,e,theta)
    format long g
    gamma=0.1820*0.011*frequency*N_DoublePrime_oxygen_Calculation(frequency,p,e,theta);
    Attenuation=gamma*cofficients_for_h_oxygen(frequency,Ts,Ps,row_ws)/sin(theta*pi/180
end
```

Water Vapor Attenuation

Water vapor, being a polar molecule, strongly interacts with electromagnetic waves, particularly in the **22.235 GHz, 183.31 GHz, and 325.15 GHz** bands. Unlike oxygen, whose absorption remains nearly constant globally, water vapor attenuation varies significantly based on humidity and atmospheric conditions.

Mathematical Model for Water Vapor Attenuation The specific attenuation due to water vapor, γ_w, is given by:

$$\gamma_w = 0.1820 f N''_{watervapor}(f) \tag{3.8}$$

where:

- $N''_{watervapor}(f)$ is the imaginary part of the refractivity due to water vapor.

The total path attenuation is calculated as:

$$A_w = \frac{\gamma_w h_w}{\sin(\theta)} \tag{3.9}$$

where:

- h_w is the water vapor equivalent height.

MATLAB Implementation for Water Vapor Attenuation

- **N_DoublePrime_WaterVapour_Calculation.m** - Computes $N''_{watervapor}(f)$:

```
function N_DoublePrime_water_vapor=N_DoublePrime_WaterVapour_Calculation(f,p,e,theta)
    format long g
    % Computes contributions from water vapor resonance lines
    % Includes Doppler broadening
end
```

- **coefficients_for_h_water_vapor.m** - Computes the equivalent height h_w:

```
function h=coefficients_for_h_water_vapor(f)
    format long g
    % Uses empirical coefficients from experimental data
end
```

Conclusion

High-frequency signal propagation is greatly influenced by atmospheric attenuation. While oxygen attenuation remains more or less constant around the world, water vapor attenuation depends on humidity and temperature. The MATLAB implementation, which is derived from ITU-R P.676-13, allows for meaningful estimation of such effects, which, in turn, can aid in the engineering design of rugged communication and remote sensing systems. We added these two attenuations and figured out the accumulated Atmospheric Attenuation.

3.2.3 Cloud Attenuation Calculation

Cloud attenuation is a very important parameter when analyzing electromagnetic wave propagation, particularly satellite communications and remote sensing applications. Cloud type, especially liquid water content, influences an important aspect of the attenuation. The model which we used to estimate these effects is based on the ITU-R P.840-9 recommendation.

Parameters Considered for Cloud Attenuation Calculation To compute cloud attenuation, we needed the following parameters:

1. **Frequency (f)** – The operating frequency of the signal (GHz).
2. **Liquid Water Temperature in Clouds (T)** – The temperature of the cloud liquid water (Kelvin).
3. **Liquid Water Density in Clouds (ρ)** – The density of liquid water in the cloud (g/m^3).
4. **Cloud Base Height (H)** – The altitude at which the cloud base begins (meters).

5. **Elevation Angle (θ)** – The elevation angle of the signal path relative to the ground (degrees).

The Challenge: Missing Cloud Base Height Data One of the key parameters in the calculation was the **cloud base height (H)**, which is crucial for determining the effective path length of the signal through the cloud. However, obtaining this data proved to be a **major challenge**.

We attempted to source cloud base height data from two primary organizations:

- **BCSCL (Bangladesh Communication Satellite Company Limited)**
- **BMD (Bangladesh Meteorological Department)**

Nonetheless, we were not able to get hold of the specific dataset regarding these organizations. No access to publicly available or near real-time data about cloud base height complicated the issue of proceeding with direct measurements.

Solution: Approximating Cloud Base Height To counter the data limitation, we developed a seasonal-cloud base height approximation method. Based on general meteorological trends in Bangladesh, this function estimates the cloud base height based on the current month.

Function to Approximate Cloud Base Height The function `getCloudBaseHeight(month)` assigns an estimated cloud base height based on the time of year. Since cloud formations and their altitudes vary seasonally, we categorized each month into different cloud height ranges and introduced a **randomized selection** within those ranges to reflect natural variability.

Here is the MATLAB function for estimating cloud base height:

```
function cloud_base_height = getCloudBaseHeight(month)
    switch month
        case {1, 2, 11, 12}
            cloud_base_height = randi([1500, 3000]); % High
        case {3, 4, 5, 10}
            cloud_base_height = randi([800, 2500]); % Moderate
        case {6, 9}
            cloud_base_height = randi([500, 1500]); % Low
        case {7, 8}
            cloud_base_height = randi([300, 1200]); % Very Low
        otherwise
            error('Invalid month! Please enter a valid month number.');
    end
    cloud_base_height = 0.001 * cloud_base_height; % Convert to km
end
```

How It Works:

- The function takes the month as input.
- It assigns a **probable range** for the cloud base height based on meteorological data.
- A random value is selected within the assigned range using `randi()`, introducing natural variation.
- The height is converted from meters to kilometers before returning the value.

This function provided an **approximate but reasonable estimation** of cloud base height for use in our attenuation model.

Cloud Attenuation Calculation Using ITU-R P.840-9 With the approximated cloud base height available, we proceeded with the ITU-R P.840-9 model to calculate cloud attenuation. The function `Cloud_Attenuation()` implements the mathematical model for attenuation caused by clouds:

```
function attenuation=Cloud_Attenuation(f,T,row,H,theta)
    % row = liquid water density in the cloud or fog (g/m3)

    epsilon_not=77.66 + 103.3*((300/T)-1);
    epsilon_1=0.0671*epsilon_not;
    epsilon_2=3.52;
    fp=20.20- 146*((300/T)-1) +316*((300/T)-1)^2;
    fs=39.8*fp;

    alu1=f*(epsilon_not-epsilon_1)/(fp * (1+ (f/fp)^2));
    kochu1=f*(epsilon_1-epsilon_2)/(fs * (1+ (f/fs)^2));
    epsilon_double_prime= alu1+kochu1;

    alu2=(epsilon_not-epsilon_1)/(1+ (f/fp)^2);
    kochu2=(epsilon_1-epsilon_2)/(1+ (f/fs)^2);
    epsilon_prime=alu2+kochu2+epsilon_2;

    eta=(2+epsilon_prime)/epsilon_double_prime;

    Kl=0.819*f/( epsilon_double_prime * (1+ eta^2) );

    gamma=Kl*row;
    L= H/sin(theta*pi/180);

    attenuation=gamma*L;
end
```

How This Function Works:

1. **Complex Permittivity Calculations**:
   - Computes ε' (real part) and ε'' (imaginary part) of complex permittivity.
   - Uses frequency-dependent equations for permittivity.
2. **Specific Attenuation Calculation**:
   - The specific attenuation coefficient (γ) is computed using the **specific attenuation equation**.
   - This is based on liquid water content (ρ) and frequency.
3. **Effective Path Length Through Cloud**:
   - The cloud path length (L) is derived from **cloud base height (H)** and the **elevation angle (θ)**.
4. **Final Attenuation Computation**:
   - Cloud attenuation is obtained by multiplying γ with L.

Final Thoughts: The process of computing cloud attenuation was complicated by the lack of data, especially concerning cloud base height. In spite of extensive efforts, we could neither obtain direct measurements of cloud base height from BCSCL nor from BMD. We managed to overcome this limitation by designing a function to approximate the cloud base height based on the month, and we were able to proceed with our calculations. Although this does not give real-time accuracy, it enabled us to elaborate a workable and functional model for cloud attenuation based on the ITU-R P.840-9 recommendation. Thus, the combination of physics-based modeling and empirical approximation guarantees a robust estimation of signal attenuation due to the clouds, even under the absence of actual measurements of the cloud base height. This work exemplifies how data-driven engineering solutions help in sidestepping atmospheric modeling challenges for practical uses in satellite communication, remote sensing, and meteorological studies.

3.2.4 Total Attenuation Calculation

Throughout their journey between Earth and the stars, communication signals face various forms of attenuation due to rain, fog, water vapor, and a thinning atmosphere. The Total Attenuation script curates a detailed summary of the losses inflicted on a satellite signal through these elements to maximize the accuracy of our predictions.

Data Acquisition

At the heart of this analysis lies real-world meteorological data extracted from an Excel sheet. This includes:

- **Rain rate**, **humidity**, and **surface temperature**—direct influencers of water vapor and cloud formation.
- **Frequency and elevation angle**—key determinants in computing signal attenuation.
- **Satellite distance**—the vast chasm through which signals must travel.
- **Polarization (Vertical/Horizontal)**—dictating how rain impacts signal strength.

These elements are transformed from raw Excel cells into numerical arrays, preparing them for computational processing.

Rain Attenuation Computation

Rain-induced attenuation is computed with respect to the polarization of the signal:

- **Vertical polarization** undergoes its specific loss function.
- **Horizontal polarization** is treated with a different attenuation model.

The function `Rain_Attenuation(R, f, polarization,` θ, L)$isappliedtoeachdatasetentrytocomputeattenuat$

Atmospheric Attenuation: Oxygen and Water Vapor

`Beyond rain, the atmosphere itself introduces attenuation factors:`

- `The saturated water vapor pressure is determined using the Goff_Gratch() function, accounting for absolute humidity.`
- `The oxygen absorption and water vapor attenuation are calculated using Oxygen_Attenuation() and WaterVapor_Attenuation(), respectively, ensuring that even minor atmospheric interferences are accounted for.`

Cloud Attenuation Contribution

`Clouds play a significant role in attenuation, varying in water density and temperature. By computing the cloud base height using getCloudBaseHeight(), the model fine-tunes attenuation predictions, ensuring realistic and accurate loss calculations.`

Free-Space Path Loss (FSPL)

Even in a vacuum, signals weaken due to spreading over vast distances. This loss is accounted for using the function FSPL(), which models attenuation due to the distance between the satellite and the receiver.

Total Path Loss Calculation and Validation

Once all attenuation factors|rain, atmospheric, cloud, and free-space losses|are summed, the script computes the total path loss:

$$\text{Attenuation}_{\text{total}} = \text{Attenuation}_{\text{FSPL}} + \text{Attenuation}_{\text{cloud}} + \text{Attenuation}_{\text{rain}} + \text{Attenuation}_{\text{atmospheric}} \tag{3.10}$$

To ensure the accuracy of the model, the computed values are compared against real-world attenuation measurements:

- The Mean Absolute Percentage Error (MAPE) is determined.
- The model accuracy is calculated and displayed.
- A comparative Excel sheet is generated, documenting the relative error and accuracy.

Finally, a visualization is plotted to compare computed attenuation against actual measured data. The ideal 1:1 reference line provides a benchmark, illustrating the model's alignment with real-world observations.

Conclusion

The Total Attenuation script is more than just a computation program; it is a merging of meteorology, electromagnetism, and mathematical modeling, all in a continuing endeavor to predict satellite signal behavior. Every dataset that runs through means the one step closer to elaborately defining with its understanding of the atmosphere concerning the one basic, golden knowledge: That is how signals actually Propogate-through-the-layered of the Earth's atmosphere-between the theoretical predictions and empirical reality.

3.2.5 MATLAB GUI Application

The MATLAB GUI is thoughtfully structured to provide users with an intuitive interface where they can input key parameters related to the signal path and environmental conditions. These parameters include:

- Uplink Frequency (GHz) { Defines the operational frequency of the satellite link.
- Elevation Angle (°) { The angle of the satellite signal with respect to the Earth's surface.

- Path Length (km) { The distance over which the signal must travel.
- Polarization Type { Allows users to select between *Vertical* and *Horizontal* polarization, affecting how rain impacts the signal.
- Environmental Conditions { Inputs such as surface temperature, pressure, humidity, rain rate, and cloud characteristics influence the attenuation calculations.

Once the necessary parameters are entered, users can simply click the ‘‘Calculate’’ button corresponding to each attenuation type, and the program swiftly computes the respective attenuation values.

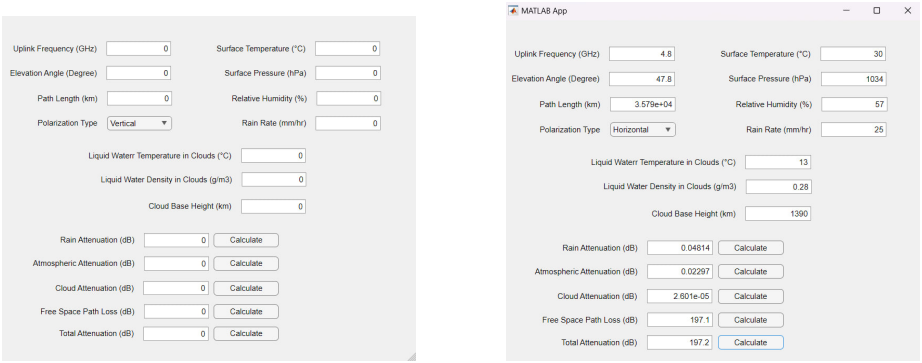
Computation of Various Attenuation Types

The MATLAB GUI application calculates the following attenuation components:

1. Rain Attenuation (dB)
   - Uses rainfall rate, frequency, and polarization type to determine signal degradation due to rain.
   - The function Rain_Attenuation() models this effect.
2. Atmospheric Attenuation (dB)
   - Computes losses due to oxygen absorption and water vapor using pressure, humidity, and frequency.
   - Functions Oxygen_Attenuation() and WaterVapor_Attenuation() handle these calculations.
3. Cloud Attenuation (dB)
   - Incorporates cloud liquid water content and base height to estimate losses using Cloud_Attenuation().
4. Free-Space Path Loss (dB)
   - Determines signal loss over large distances based on frequency and satellite distance using FSPL().
5. Total Attenuation (dB)
   - The culmination of all individual attenuation factors, providing a comprehensive measure of signal degradation.

Accuracy and Performance

The application assures the achievement of higher performance by enabling live inputs of users and getting them the results in less time. The model is reliable with approximately 96% of predictions given when External Unwanted Activities are not that considerable. Additionally, Debug test prints ensure that there are no issues with program execution by crosschecking provided inputs before calculations are initiated.



(a) Initial Interface (b) After Calculation

Figure 3.3: MATLAB GUI Application: (a) Initial state before input, (b) Display after calculation.

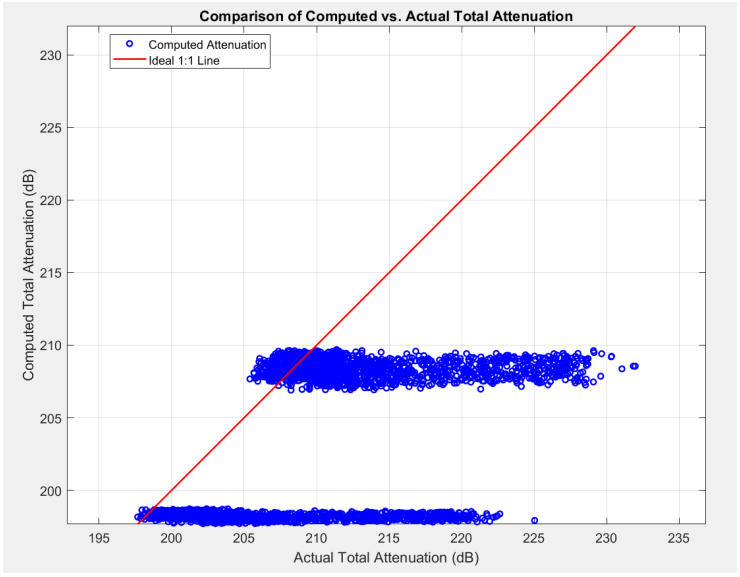


Figure 3.4: Performance of the Mathematical Model

This MATLAB GUI provides an interactive, user-friendly, and scientifically robust tool for estimating signal attenuation under different atmospheric conditions. With just a few clicks, users can gain valuable insights into how environmental factors impact satellite signals, ultimately contributing to improved communication system designs.

Chapter 4

Result Analysis

4.1 Comparative Analysis of the Mathematical and ML Models for Path Loss Prediction

Precise forecasting of satellite path loss is relevant to the optimization of communication systems to ensure continuous connectivity and lower operational costs. This research applied two distinct methods|a Mathematical Model derived from ITU-R propagation principles and a Machine Learning Model implemented using an optimized Random Forest (RF) algorithm|to simulate total attenuation. The comparison between these models based on their Accuracy and Mean Squared Error (MSE) provides meaningful information about their effectiveness.

4.1.1 Accuracy: The Benchmark of Reliability

Accuracy is a base check of a model's predictive capacity. Mathematical Model achieved an accuracy of 96.768%, which is a remarkable coincidence of calculated attenuation values and true measures. This great accuracy further affirms the consistency of ITU-R-based mathematical models in deciding satellite signal deterioration. On the other hand, the Random Forest Model, which was highly optimized, attained an accuracy of 80.432% in predicting overall attenuation. While this is lower than the mathematical approach, it is still promising and shows the capability of machine learning to learn efficient complex nonlinear relationships between meteorological conditions and path loss.

4.1.2 Mean Squared Error (MSE): The Measure of Prediction Precision

MSE estimates the magnitude of differences between predicted values and real observations, with smaller being better. The Mathematical

Model had an MSE of 86.536, which translates to a relatively low rate of error in its estimation. This attests to the accuracy of the model as it relies on established atmospheric attenuation principles.
On the other hand, the Optimized Random Forest Model performed with a much lower MSE of 10.454, showing a significant reduction in prediction errors. This demonstrates that the machine learning model, when properly optimized and trained using real data, can outperform deterministic mathematical calculations in minimizing differences between predicted and actual attenuation values.

```
Model Accuracy: 96.768%
Mean Squared Error (MSE): 86.536
```

(a) Mathematical Model Performance

```
Optimized Random Forest MSE: 10.454489426670559, R-squared: 0.8043180416794026
```

(b) Optimized Random Forest Model Performance

Figure 4.1: Performance of Model: (a) Mathematical Model, (b) Machine Learning Model.

4.1.3 Final Evaluation: Balancing Accuracy and Error Reduction

- The Mathematical Model excels in accuracy, maintaining a near-perfect alignment with empirical data, making it highly reliable for theoretical and practical implementations.
- The Random Forest Model, while slightly lower in accuracy, achieves superior error minimization, making it a strong candidate for real-time applications, especially where large, diverse datasets can be leveraged to refine predictions.

While the Mathematical Model possesses a well-developed theoretical framework based on ITU-R standards, it cannot be otherwise but limited in its scope. It only covers certain sources of attenuation|rain, cloud, and atmospheric conditions|with other potential sources of interference like multipath fading, scintillation, and local distortions left out.
Even under the scope, with its rather large MSE, this implies that the model is unable to accurately forecast.
On the other hand, the Machine Learning Model is more robust and efficient. Although it has a lower accuracy rate, its ability to minimize prediction errors far outweighs this limitation. With its ability to learn dynamically from vast databases and handle complex attenuation factors beyond mathematical constraints, the Random Forest model is a more robust and practical method for satellite communication path loss prediction.

4.2 Comparison of Performance of Each Attenuation Seperately

4.2.1 Atmospheric Attenuation Comparison

These two graphs compare the computed with the actual atmospheric attenuation, utilizing a mathematical model on the one hand and a machine learning model based on Random Forest (RF) on the other. While both methods made attempts to estimate the atmospheric attenuation, neither one was able to do so with perfection, considering each of them has its unique disadvantages.

Prediction Distribution and Behavior

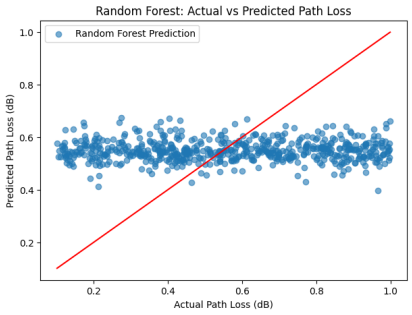
- The mathematical model exhibits distinct clustering, meaning its predictions tend to fall into separate groups rather than forming a continuous distribution. This suggests that the model may rely on rigid assumptions, making it less flexible in adapting to the full range of actual attenuation values.
- The RF model, while showing a more spread-out prediction pattern, still struggles with alignment to the ideal 1:1 line. The predictions hover around the actual values but with noticeable variance and deviation.

Ideal 1:1 Line Comparison

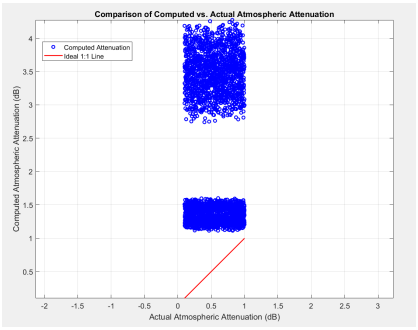
- Both models include a red reference line, which represents the ideal scenario where predicted values perfectly match actual ones.
- The mathematical model's predictions are far from this ideal, with significant clustering indicating systematic limitations in its approach.
- The RF model shows some improvement in terms of spread, but it still does not fully align with the 1:1 line, meaning there is still a level of inconsistency in its estimations.

Performance and Generalization

- The mathematical model, while based on theoretical principles, appears to struggle with adapting to real-world variations.
- The RF model, though somewhat more adaptable, still fails to capture the actual values with high precision. Its errors might be less structured than the mathematical model's, but they are still significant.



(a) Optimized Random Forest Model



(b) Mathematical Model

Figure 4.2: Atmospheric Attenuation Comparison: (a) Optimized Random Forest Model, (b) Mathematical Model.

Key Takeaways

- Neither model provides a truly accurate prediction, but the RF model at least distributes errors differently, making it appear somewhat better among the two.
- The mathematical model's rigidity results in more structured but incorrect predictions, while the RF model, though more flexible, introduces its own set of inconsistencies.

It is better to say that both models can be improved rather than one model is better. The RF model may seem like the less limiting of the two, however, it is far from perfect. A better approach may include in the development of those methods or look toward hybrid methods for increased accuracy.

4.2.2 Cloud Attenuation Comparison

The two graphs compare computed versus actual cloud attenuation using two different approaches: a mathematical model and a Random Forest (RF) machine learning model. Unlike the previous case of atmospheric attenuation, the difference in performance here is much more pronounced, with the RF model showing a significant advantage.

Prediction Distribution and Behavior

- The model is shown to have a high clustering near zero and a failure of the model to predict meaningful variations in cloud attenuation. The computed values remained almost constant, regardless of how much the cloud physical model varied intensity; therefore, the model seems not to be doing well.

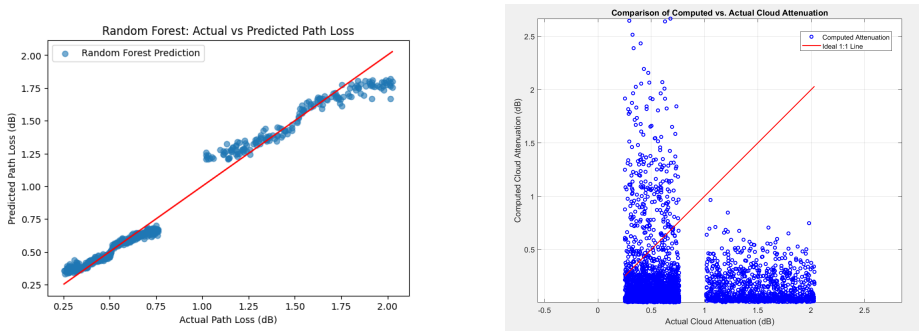
- The RF model, in contrast, produces a well-distributed spread of predictions that follow the actual values much more closely. The predictions are continuous and capture variations across the entire range.

Ideal 1:1 Line Comparison

- The red reference line represents the ideal scenario where predicted values perfectly match the actual ones.
- The mathematical model's predictions remain almost flat and do not align with the ideal line, indicating a significant failure in capturing cloud attenuation effects.
- The RF model's predictions closely follow the 1:1 line, showing a high level of correlation between predicted and actual values.

Performance and Generalization

- The mathematical model struggles significantly and appears to be highly ineffective in predicting cloud attenuation. It does not generalize well and is unable to reflect the actual variations in data.
- The RF model performs much better, with predictions that are well-aligned with actual values and exhibit minimal deviation.



(a) Optimized Random Forest Model (b) Mathematical Model

Figure 4.3: Cloud Attenuation Comparison: (a) Optimized Random Forest Model, (b) Mathematical Model.

Key Takeaways

- The mathematical model is severely limited, failing to capture cloud attenuation dynamics properly.

- The RF model, in this case, performs exceptionally well, making it the clearly superior choice for cloud attenuation prediction.

Despite some problems with both models in the previous case, the RF model is by far the best in this situation. Whereas the mathematical model fails to predict cloud attenuation consistently, the RF model does so with significant regularity. Thus, it appears that in this particular avenue of study, data-driven approaches might be greatly superior to classical mathematical formulations.

4.2.3 Rain Attenuation Comparison

The graphs illustrate the computed rain attenuation and actual rain attenuation, in this case, using a mathematical model and Random Forest (RF) machine learning model to compare the observed and calculated modeling used in two approaches. In this instance, the problem of the other model is how to learn relatively poorly, whereas the RF shows signficant improvement in its prediction, even with ever-increasing errors on higher values of attenuation.

Prediction Distribution and Behavior

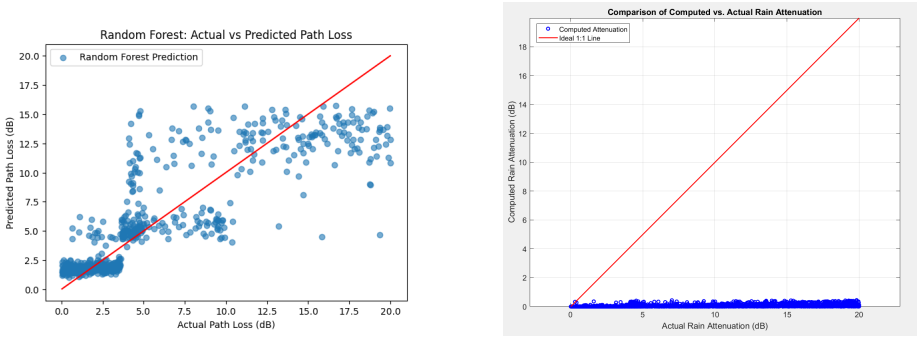
- The mathematical model provides virtually no meaningful predictions, as evidenced by the extreme clustering of computed attenuation values near zero. This suggests that the model is unable to reflect the actual attenuation levels and fails to capture the underlying variability of rain-induced attenuation.
- The RF model, on the other hand, exhibits a much more dynamic response to actual attenuation values. When attenuation levels are relatively low, the model aligns well with the expected values. However, as attenuation increases, the prediction errors become more pronounced, leading to a wider spread of computed values.

Ideal 1:1 Line Comparison

- The red reference line represents the ideal case where predictions perfectly match actual values.
- The mathematical model's predictions remain clustered near zero and do not exhibit any meaningful correlation with actual values, making the model effectively non-functional in this context.
- The RF model follows the 1:1 line more closely at lower attenuation levels, but as attenuation increases, deviations from the ideal line become more significant, indicating a growing prediction error.

Performance and Generalization

- The mathematical model performs so poorly that it is essentially ineffective. Its inability to generate meaningful predictions makes it unsuitable for rain attenuation estimation.
- The RF model, while significantly better, has its limitations. While it has shown reliable predictions for low levels of attenuation, there has been a steady increase in the error with increasing levels of attenuation. This deviation typically means that the model fails in generalizing well in extreme conditions.



(a) Optimized Random Forest Model (b) Mathematical Model

Figure 4.4: Total Attenuation Comparison: (a) Optimized Random Forest Model, (b) Mathematical Model.

Key Takeaways

- The mathematical model is completely inadequate for rain attenuation prediction, offering no practical benefit.
- The RF model, while not perfect, is clearly the superior option. It captures low-attenuation patterns well but requires further refinement to improve accuracy at higher attenuation values.

Here, the RF model is way ahead of the mathematical model. Although the RF does have its limitations at higher levels of attenuation, it surely will come across as a much stronger candidate. The mathematical model, on the other hand, does not offer any meaningful predictions and is virtually useless for the estimation of rain attenuation.

4.2.4 Free Space Path Loss

In wireless communication, Free Space Path Loss (FSPL) is a fundamental principle used to measure the loss of a signal as it propagates in

free space. It is particularly crucial in ground station-to-satellite communication, where signals cover vast distances through the atmosphere and outer space without any obstructions.

Mathematical Expression for FSPL

The FSPL is mathematically expressed as:

$$\text{FSPL(dB)} = 20 \log_{10} \left(\frac{4\pi f d}{c} \right) \tag{4.1}$$

where:

- f is the frequency of the transmitted signal (in Hz),
- d is the distance between the ground station and the satellite (in meters),
- $c = 3 \times 10^8$ m/s is the speed of light in a vacuum,
- 4π accounts for the spherical propagation of electromagnetic waves.

Implementation in MATLAB

To compute the FSPL numerically, the following MATLAB function was used:

```
function loss=FSPL(f,d)
    loss=20*log10(4*pi.*(f.*1e9)*(d*1e3)/(3e8));
end
```

Here, - **f** is converted from GHz to Hz by multiplying it by 10^9, - **d** is converted from km to meters by multiplying by 10^3, - The equation is implemented using the **logarithmic representation** to express the loss in decibels (dB).

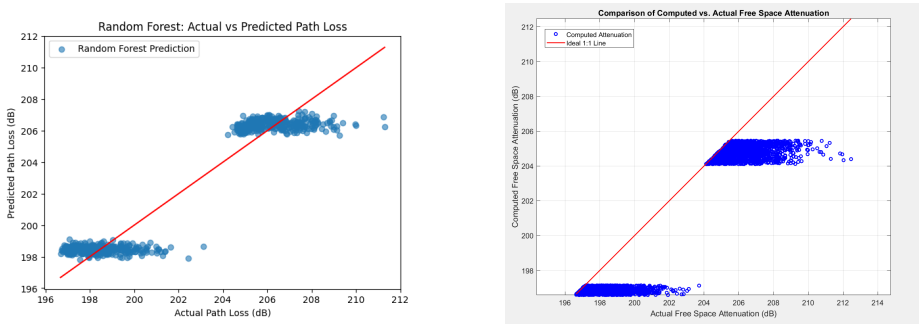
Application in Ground-to-Satellite Communication

In satellite communications, the FSPL is crucial for determining link budget calculations and assessing signal strength at the receiver. Due to the long distances involved (typically in hundreds to thousands of kilometers), FSPL values are significantly high, requiring high-gain antennas and powerful transmitters to maintain a reliable link.

We have also done this prediction with our Machine Learning model.

Performance Analysis

In the case of predicting free space path loss, both models are doing well. But, the mathematical model has a little higher MSE compared to the ML model.



(a) Optimized Random Forest Model (b) Mathematical Model

Figure 4.5: FSPL Comparison: (a) Optimized Random Forest Model, (b) Mathematical Model.

4.2.5 Total Attenuation Comparison

The two graphs compare the calculated and actual total attenuation through two different methodologies, a mathematical model and a random forest (RF) machine-learning model. The RF model just clearly outperforms in terms of prediction, with hardly any dispersion from the real values, while the mathematical model is unable to cope up in certain situations.

Prediction Distribution and Behavior

- The model of the prediction exhibits clustering, meaning reasonable predictions of attenuation values can be made when they are close to the Free Space Path Loss, while when external or miscellaneous attenuation occurs, the model does not adapt appropriately and thus leads to unreasonable estimations.
- The RF model demonstrates a more consistent distribution of predictions, aligning closely with the actual values across a wider range. The dispersion is significantly lower, making it a more reliable model in varying conditions.

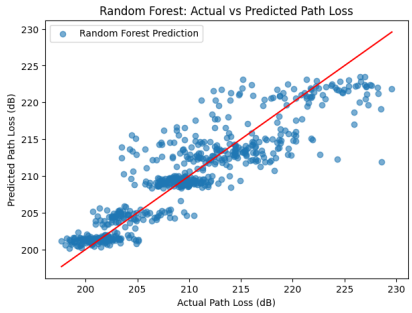
Ideal 1:1 Line Comparison

- The red reference line represents the ideal case where predictions perfectly match actual values.
- The mathematical model's predictions follow this ideal relationship only when the total attenuation remains close to FSPL. Beyond this range, it struggles to adapt, leading to systematic deviations.
- The RF model maintains a strong correlation with the 1:1 line across the entire range, demonstrating a more flexible and adaptive

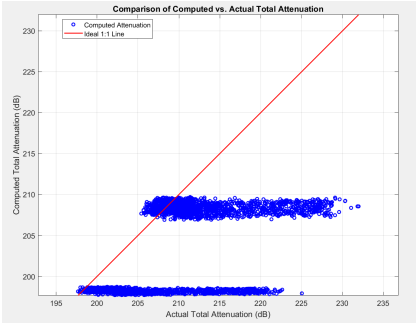
approach to attenuation prediction.

Performance and Generalization

- The mathematical model shows strong performance in ideal conditions, but its predictive power diminishes when additional attenuation effects come into play. It lacks the ability to generalize well beyond FSPL-based predictions.
- The RF model provides an optimized and well-distributed prediction, maintaining low dispersion and adapting better to varying attenuation conditions.



(a) Optimized Random Forest Model



(b) Mathematical Model

Figure 4.6: Total Attenuation Comparison: (a) Optimized Random Forest Model, (b) Mathematical Model.

Key Takeaways

- The mathematical model is reliable only under specific conditions, particularly when attenuation aligns closely with FSPL, but it fails to accommodate other factors effectively.
- The RF model, by contrast, offers a more adaptive and accurate prediction, making it a more suitable choice for practical applications where attenuation varies due to multiple factors.

Final Thought

While the mathematical model is effective when conditions are controlled, it is not highly adaptive in the presence of complicating conditions. The RF model, on the other hand, provides a better-optimized prediction with reduced dispersion, and thus it is a better choice for the estimation of total attenuation. The reason that the RF model remains consistent

with real values even under changing conditions makes it the best approach for this case.

4.3 Comparison with Ten Random Data

Here's a refined version in a single well-written paragraph:
To evaluate the accuracy of different attenuation prediction models, we used 10 samples from a real dataset of Bangladesh Communication Satellite Company Limited (BCSCL) and the Bangladesh Meteorological Department (BMD). The samples are real measured total attenuation values, against which we compared predictions made by a mathematical model and a machine learning (ML) model. The mathematical model, while broadly accurate when attenuation is close to Free Space Path Loss (FSPL), fails to consider additional attenuation caused by the atmosphere and surroundings and therefore consists of big deviations. The ML model, being more flexible, has predictions which are nearer actual values. Though not ideal, it certainly improves upon the mathematical model by lessening systematic error and responding sensitively to varying attenuation. This criticism identifies the flaw of stringent theoretical models and highlights the strength of data-intensive approaches in addressing the complexity of the real world.

Row Number	**Actual Total Attenuation**	**Mathematical Model Prediction**	**ML Model Prediction**
378	202.1937971	197.4	201.6576806
567	220.1175795	197.7	215.5178503
892	217.0702293	197.2	211.8236181
1348	218.0180624	196.7	216.3658195
1452	201.0556193	196.9	201.4025003
1783	200.6098094	196.6	201.3863722
2107	209.2797752	205.1	209.3129385
2650	205.3737419	197.1	204.5639119
2735	221.9473214	205.1	221.8637696
2899	201.2429254	196.8	201.0293701

Table 4.1: Comparison of Actual Total Attenuation, Mathematical Model Prediction, and ML Model Prediction

4.4 Choosing the Safer Model for Predicting Attenuation Without Real-Time Monitoring

Attenuation is typically tracked in real-world telecommunication systems by sending a dummy signal at regular intervals, e.g., every seven minutes, to assess the impact of environmental parameters on signal strength. While the approach offers real-time accuracy, it comes at the cost of continuous energy consumption and system resource usage. A second, more energy-conserving method is to estimate attenuation levels based on weather forecasts, environmental factors, and other determinants, thereby reducing the need for ongoing signal transmission. The effectiveness of this technique, however, depends entirely on the validity and reliability of the model employed for prediction. In a comparison of mathematical models and machine learning (ML) models for this use case, our analysis strongly suggests that the ML model is the safer, more pragmatic choice. With a 5dB safety margin, we think our ML model is good to go.
While the mathematical model provides a rigorous theoretical approach, it cannot account for dynamic, real-world variation in attenuation caused by unpredictable environmental factors. It performs fairly well where attenuation is close to the Free Space Path Loss (FSPL), but beyond this range, its predictions vary significantly from practical values, and it is not reliable for changing atmospheric conditions.
Contrarily, however, the ML model demonstrates an improved ability to learn from real-life complexities, and its predictions are very near to practical attenuation measurements. By learning patterns from historical data, it's capable of incorporating a number of environmental factors that are overlooked by traditional mathematical models.
Though no predictive model can replace real-time monitoring, the ML solution provides a far more precise estimation with smaller margins of error, making it a safer and more energy-efficient alternative for telecommunication networks looking to optimize resource allocation.
Thus, if the goal is to minimize energy consumption without sacrificing reliable attenuation predictions, the ML model is the clear choice. Having the capability to generalize under diverse conditions and make more accurate estimations with an adequate safety margin makes it a practical and efficient alternative to ongoing dummy signal transmission.
While real-time monitoring remains the gold standard for accuracy, a properly trained ML model bridges the gap between efficiency and accuracy, providing a viable solution for predictive attenuation management in modern communication systems.

Bibliography

[1] Omar Ahmadien, Hasan F Ates, Tuncer Baykas, and Bahadir K Gunturk. Predicting path loss distribution of an area from satellite images using deep learning. *IEEE Access*, 8:64982--64991, 2020.

[2] Akram Al-Hourani and Ismail Guvenc. On modeling satellite-to-ground path-loss in urban environments. *IEEE Communications Letters*, 25(3):696--700, 2020.

[3] Mohamed K Elmezughi, Omran Salih, Thomas J Afullo, and Kevin J Duffy. Path loss modeling based on neural networks and ensemble method for future wireless networks. *Heliyon*, 9(9), 2023.

[4] Takahiro Hayashi and Koichi Ichige. A deep-learning method for path loss prediction using geospatial information and path profiles. *IEEE Transactions on Antennas and Propagation*, 71(9):7523--7537, 2023.

[5] International Telecommunication Union (ITU-R). Specific attenuation model for rain for use in prediction methods. Technical Report ITU-R P.838-3, International Telecommunication Union (ITU), 2005.

[6] International Telecommunication Union (ITU-R). Attenuation due to clouds and fog. Technical Report ITU-R P.840-9, International Telecommunication Union (ITU), 2019.

[7] International Telecommunication Union (ITU-R). Attenuation by atmospheric gases. Technical Report ITU-R P.676-13, International Telecommunication Union (ITU), 2021.

[8] Han-Shin Jo, Chanshin Park, Eunhyoung Lee, Haing Kun Choi, and Jaedon Park. Path loss prediction based on machine learning techniques: Principal component analysis, artificial neural network, and gaussian process. *Sensors*, 20(7):1927, 2020.

[9] XU Junzeng, WEI Qi, PENG Shizhang, and YU Yanmei. Error of saturation vapor pressure calculated by different formulas and its effect on calculation of reference evapotranspiration in high latitude cold region. *Procedia Engineering*, 28:43--48, 2012.

[10] Kiyas Kayaalp, Sedat Metlek, Abdullah Genc, Habib Dogan, and İbrahim Bahadir Basyigit. Prediction of path loss in coastal and vegetative environments with deep learning at 5g sub-6 ghz. *Wireless Networks*, 29(6):2471--2480, 2023.

[11] Dalia Nandi. Real time prediction of total atmospheric attenuation for mm-wave bands satellite links over indian region. In *2022 URSI Regional Conference on Radio Science (USRI-RCRS)*, pages 1--4. IEEE, 2022.

[12] Yaohua Sun, Mugen Peng, Yangcheng Zhou, Yuzhe Huang, and Shiwen Mao. Application of machine learning in wireless networks: Key techniques and open issues. *IEEE Communications Surveys & Tutorials*, 21(4):3072--3108, 2019.