Implementing Radar Technology on Drones for Subsurface Soil Moisture Detection: A Guide for Researchers, Agriculturists, and Environmentalists

Sponsored by SoilX

A Major Qualifying Project submitted to the faculty of
Worcester Polytechnic Institute in partial fulfillment of the requirement
for the degree of Bachelor of Science

Submitted by:

Nick Latsis, Ethan Reed, Josh Thurber

Advisors:

Prof. Seyed Zekavat

Prof. Doug Petkie

Worcester Polytechnic Institute

April 25th, 2024

Abstract

This project aims to develop intelligent moisture characterization by utilizing drone-based radar signal transmission and data analysis. With the increasing challenges posed by climate change, efficient water management in agriculture has become crucial. Farmers require precise information to optimize irrigation practices, minimize water waste, and maintain soil health. Our project addresses this need by leveraging radar technology mounted on drones to create root-zone moisture maps. By analyzing reflected radio frequency (RF) signals from Ground Penetrating Radar (GPR) systems, we aim to provide non-invasive, large-scale soil moisture measurements. This approach offers rapid, accurate assessments essential for agricultural decision-making, environmental monitoring, and hydrological research. Through the development of custom radar hardware and software, coupled with innovative data analysis techniques, we seek to establish a reliable and scalable solution for soil moisture detection. Our findings have the potential to revolutionize agricultural practices, leading to improved water resource management and enhanced sustainability in food production.

Table of Contents

List of Figures

Authorship Page

Nick Latsis was a co-author of this report. He focused on the implementation of the radar on the drone. This involved multiple meetings with the drone company to ensure all parameters for flight were followed, and the organization of test flights for the drone. Additionally, he conducted data collection at the Gateway Park SoilX Site for later analysis. He contributed to the writing and editing of all sections of the final report and final presentation.

Ethan Reed was a co-author of this report. He focused on the development of the sample data for analysis with gprMax. He also helped mount the radar to the drone.

Josh Thurber was a co-author of this report. He focused on the development of the Convolutional Neural Network (CNN) and mini-PC. This involved data cleaning of the Gateway Park SoilX Site datasets for use in the CNN. Additionally, he automated the parsing process with the mini-PC which simplified the measuring process. He contributed to the writing and editing of all sections of the final report and final presentation.

1: Introduction

1.1 Motivation

The goal of the project was to develop intelligent moisture characterization from a drone-based radar signal and data analysis. In recent years, the escalating impacts of climate change have profoundly affected agricultural landscapes worldwide. Agricultural communities are grappling with increasingly erratic weather patterns, prolonged droughts, and intensified floods, which pose significant challenges to crop production and food security. One of the critical concerns arising from these challenges is the efficient management of soil moisture. Farmers face the daunting task of optimizing irrigation practices to ensure adequate water supply for their crops while minimizing water, soil, and mineral waste. However, traditional methods of soil moisture monitoring often lack the precision and scalability required to meet the evolving needs of modern agriculture. Conventional techniques such as manual soil sampling or sensor-based monitoring systems are labor-intensive, time-consuming, and limited in spatial coverage. Moreover, the repercussions of inefficient irrigation practices extend beyond agricultural productivity to encompass economic, environmental, and social dimensions. Inefficient water usage not only inflates production costs and reduces profitability for farmers but also exacerbates environmental degradation and threatens the long-term sustainability of agricultural ecosystems.

Amidst these challenges, the development of innovative technologies presents a promising avenue for addressing the complexities of soil moisture management. Drone-based radar technology offers a unique solution by leveraging remote sensing capabilities to provide real-time, high-resolution data on soil moisture dynamics. By transmitting and receiving radar signals from aerial platforms, such as drones, it becomes possible to create detailed maps of root-zone moisture distribution across agricultural landscapes. The integration of radar technology with advanced data analysis techniques holds the potential to revolutionize irrigation management practices. Through the generation of comprehensive moisture characterization maps, farmers can make informed decisions regarding irrigation scheduling, water allocation, and crop management strategies. This not only enhances agricultural productivity and resource efficiency but also promotes environmental sustainability and resilience in the face of climate variability.In light of these considerations, our project seeks to bridge the gap between technological innovation and agricultural sustainability. By developing a drone-based radar system for soil moisture detection, we aim to empower farmers with the tools and insights needed to navigate the challenges of a rapidly changing climate. Ultimately, our efforts strive to foster a more resilient, equitable, and sustainable future for agricultural communities worldwide.

1.2 State of Art

This project harnesses the principles of radio wave science to investigate the intricate relationship between soil moisture levels and the amplitude of reflected radio frequency (RF) signals, employing Ground Penetrating Radar (GPR) technology as the primary tool. At its core, the project capitalizes on the fundamental understanding that water content in the soil exerts a pronounced influence on the reflection and propagation characteristics of RF signals. As radio waves penetrate the soil, they interact with moisture molecules, resulting in distinct changes in signal amplitude and travel time.Ground Penetrating Radar (GPR) technology serves as a sophisticated means of probing subsurface soil layers and capturing these subtle variations in RF signal properties. By analyzing the amplitude variations of signals reflected from the soil surface to a drone-mounted GPR system, the project endeavors to provide non-invasive, high-resolution measurements of soil moisture content across expansive agricultural landscapes. This innovative approach represents a paradigm shift in soil moisture monitoring, offering a rapid and accurate assessment of soil moisture dynamics over large spatial scales.The adoption of drone-mounted GPR technology, represents a significant advancement in the field of agricultural and environmental monitoring. Unlike conventional methods of soil moisture assessment, which are often constrained by limitations in spatial coverage, resolution, and scalability, our approach offers unparalleled capabilities for capturing detailed moisture characterization data across diverse terrains.

Furthermore, the integration of GPR technology with drone-based platforms enhances accessibility and flexibility, enabling efficient data collection over remote or inaccessible areas. This not only streamlines the monitoring process but also facilitates timely decision-making for farmers, land managers, and environmental researchers.The implications of this innovative methodology extend far beyond agricultural applications, encompassing a wide range of fields such as environmental monitoring, hydrological research, and land-use planning. The ability to accurately quantify soil moisture content non-invasively holds immense value for understanding ecosystem dynamics, predicting hydrological processes, and informing sustainable resource management practices.In summary, the utilization of Ground Penetrating Radar (GPR) technology in conjunction with drone-based platforms represents a pioneering approach to soil moisture monitoring. By offering rapid, accurate, and non-invasive measurements over large spatial scales, this methodology holds immense promise for revolutionizing agricultural practices, advancing environmental research, and promoting resilience in the face of climate change.

1.3 Radar Development

Our radar was crucial to determine the levels of moisture in the soil. We will go over the structure of the radar system we developed, detailing the various modifications it underwent during the project. We'll also delve into the capabilities of the radar, highlighting its operational features and the specific technological advances it incorporates. Additionally, we will discuss the software that was integral to the functioning of the radar, explaining how it facilitates data acquisition, processing, and analysis. This comprehensive overview will provide a clearer understanding of both the hardware and software components critical to our radar's performance.

1.4 Measurement Campaign

Our measurement campaign was designed to capture radar measurements of soil, simulating those from a drone flight. We begin by evaluating the radar's initial status, discussing our initial trials with an Akila radar and subsequent switch to a higher-resolution SFCW radar after facing challenges with data viability. Following the radar's assessment, we describe our methods for creating simulated sample data collected at various heights and depths at the Worcester Polytechnic Institute's Gateway Park, to mimic drone-based radar measurements. The chapter concludes with a discussion on the technical steps involved in securely attaching the reconfigured radar to the drone, ensuring balance and functionality for aerial data collection. This sequence of activities highlights our strategic approach to enhancing the accuracy and reliability of soil radar measurements from an airborne perspective.

1.5 Data Analysis and Machine Learning Techniques

The project utilizes a dataset comprising 500 spectrogram images and corresponding moisture data across various depths, employing advanced preprocessing techniques like interpolation and normalization to prepare the data for machine learning analysis. A Convolutional Neural Network (CNN) architecture is employed, featuring convolutional layers for feature extraction, pooling layers for dimensionality reduction, and dense layers for interpreting features for moisture prediction. Training prioritizes minimizing Mean Squared Error (MSE) to ensure accurate predictions across different soil layers. The study's significance lies in its potential to enhance environmental monitoring and agricultural practices by offering a reliable, non-invasive method for soil moisture assessment. Accurate moisture predictions from radar data aid in better water resource management and deepen understanding of soil properties, crucial for sustainable environmental stewardship and effective agricultural planning. This introduction lays the groundwork for a detailed exploration of the project's methodologies and findings, highlighting the innovative integration of machine learning with geophysical techniques to tackle key challenges in environmental and soil sciences.

2: Radar Development

In the following section, we will provide a detailed overview of our radar system, outlining its structure and the various configurations it underwent to enhance its functionality. We will also examine the diverse capabilities of the radar, highlighting its technological features and operational flexibility. Moreover, we will discuss the software integral to the radar's operation, detailing how it supports data management, signal processing, and overall system control.

2.1 Radar Structure

Our initial SFCW radar configuration utilized pulse signals to estimate delay and calculate distance. The setup included a PC connected to a Vector Signal Generator and a Spectrum Analyzer. The signal flow began with the Vector Signal Generator connected to a Low Pass Filter, which then linked to the first Low Noise Amplifier. This amplifier fed into a power splitter, which in turn connected to a power amplifier and then to the transmitting antenna. On the receiving end, the radar was connected to another low pass filter and a second low noise amplifier, followed by a mixer and the spectrum analyzer. The power splitter also connected directly to the mixer. Additionally, the power supply was hooked up to both low noise amplifiers and the power amplifier. However, this configuration was unable to change waveform and amplitude and could not be mounted to the drone. Hence, we switched to a different configuration. In the second radar configuration, which emitted a continuous signal tone, the only change was the removal of the power splitter and mixer, simplifying the setup. This setup allowed for a continuous signal to be transmitted and was mountable to the drone.

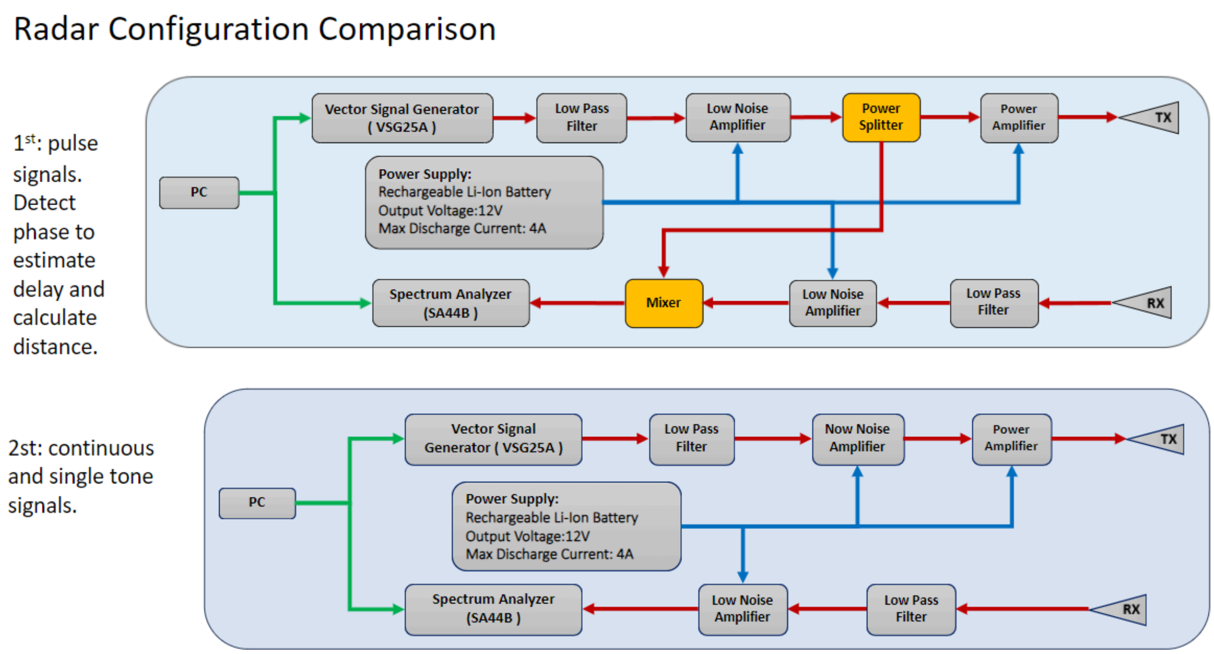

Figure 2.1.1 SFCW Radar Configurations

2.2 Capabilities of the Radar

The radar was equipped with the capability to emit a single continuous signal with a variable frequency. By adjusting the frequency, we could manipulate the wavelength of the emitted signal, which in turn influenced its interaction with the soil. Specifically, signals with longer wavelengths had the ability to penetrate deeper into the soil, providing insights into subsurface structures. Conversely, signals with shorter wavelengths were more precise and provided higher resolution data, allowing for the differentiation of various materials, such as distinguishing between sand and soil. This flexibility in frequency adjustment was crucial for optimizing the radar's performance according to the specific characteristics of the surveyed area and the nature of the materials being examined.

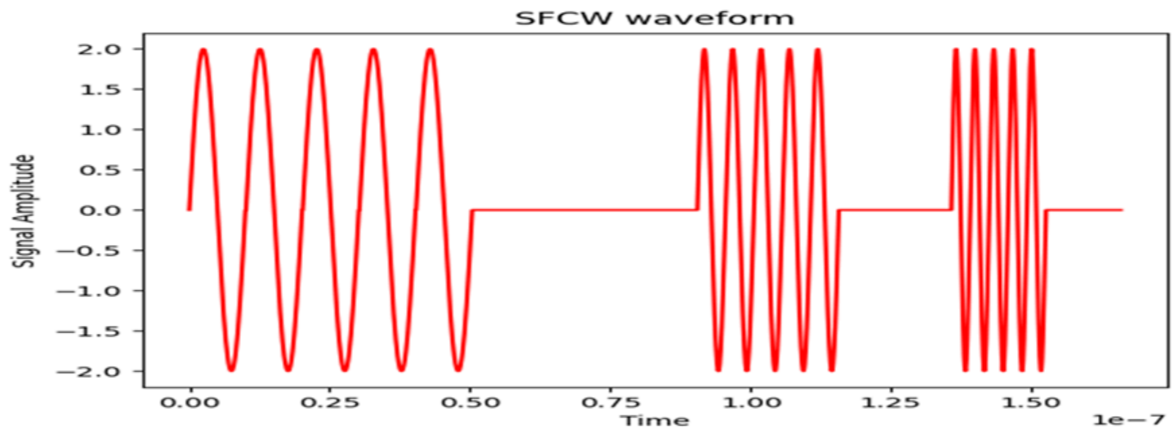

Figure 2.2.1 SFCW Waveform

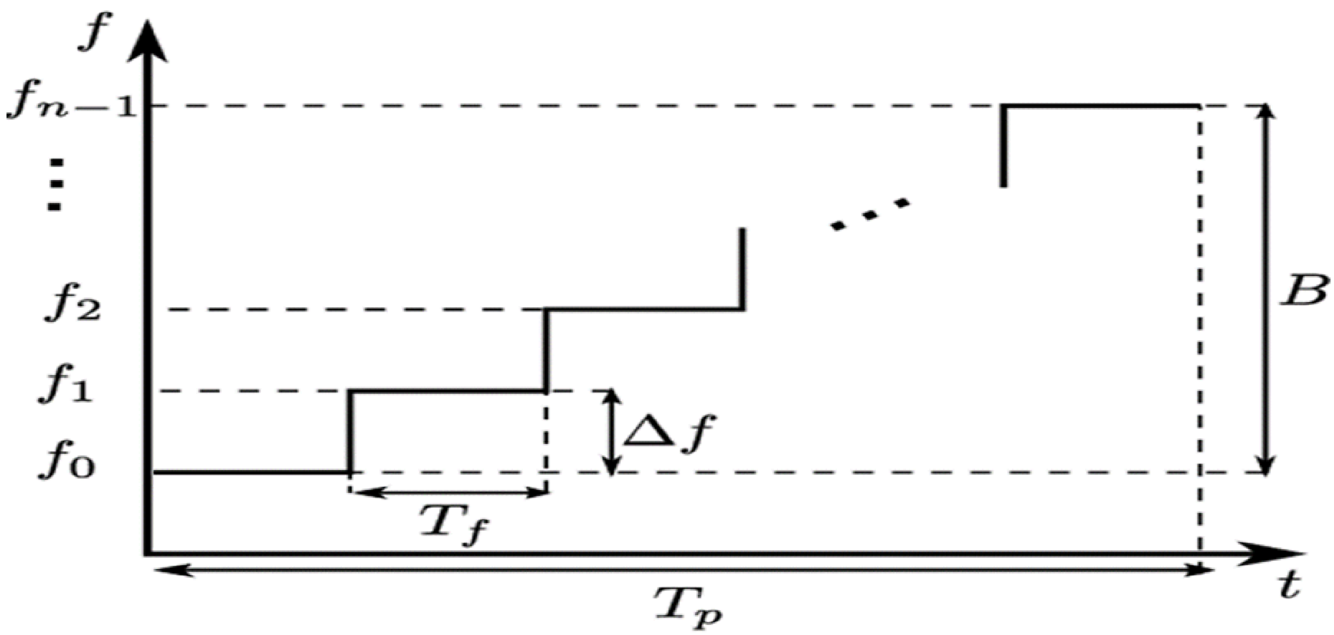

Figure 2.2.2 Frequency Steps

2.3 Radar Software

The developed radar system utilizes two separate software applications: Spike.exe and VSG25.exe, both produced by Signal Hound. Spike.exe operates the SA44B spectrum analyzer, allowing users to define frequency bandwidths to sweep from 1 Hz to 4.4 GHz in various modes. VSG25.exe controls the VSG25A signal vector generator, capable of generating diverse modulations from 100 MHz to 2.5 GHz, as well as emitting arbitrary waveforms created by third-party tools like Python and MATLAB. Additionally, a noteworthy feature of these devices is their API, which enables users to control both pieces of hardware with custom-developed Python or C++ code to perform any arbitrary operations.

3: Measurement Campaign

This chapter provides insight into our measurement campaign. The objective of our measurement campaign was aimed to get radar measurements of soil that would simulate those from a drone flight. First, we go over how we determined the status of the radar. Next, we cover how we created sample data that would simulate radar measurements from a drone. Lastly, we discuss how we attached the radar to the drone.

3.1 Creating Sample Data for Analysis

In order to better understand the characteristics of our actual data and refine our predictive models, we recognized the necessity of creating simulated data. This approach allowed us to systematically explore different scenarios and variables in a controlled environment, ensuring that our models were well-tuned before being applied to real-world data.

Simulated data served as a crucial step in our preparatory process, providing valuable insights into potential outcomes and behaviors that our actual data might exhibit. By leveraging simulated datasets, we were able to identify key patterns, test hypotheses, and adjust model parameters effectively, thereby enhancing the accuracy and reliability of our analytical outcomes once real data would be introduced. This method significantly mitigated the risks associated with direct real-world application, ensuring a more grounded and informed approach to model deployment.

3.1.1 GPRMax

We utilized the gprMax simulation tool to model electromagnetic wave interactions within a wide variety of subsurface environments. Starting with basic .in files, we generated detailed 3D models that incorporate an extensive range of materials and environmental conditions, closely mimicking real-world scenarios.

This process started by defining the domain ranges, mapping periods, waveform types, and the positions of transmitters and receivers. Materials were then meticulously created and placed within the simulation environment, ensuring accurate representation of physical properties. After the initial specifications were defined, soil layers were created at various depths using the Peplinski model based on values that were recorded with the moisture rod.

Subsequent adjustments to these models allowed for the exploration of different subsurface configurations, ranging from simple homogeneous layers to complex setups involving multiple object types and varying moisture levels within the soil. With each item added, both the wave interaction and pulse changed which was reflected in our model. Key to this approach is the generation of both output files and .vti files, the former of which can be used for AI quantitative analysis and visual numerical graphs. The .vti files are used for visual data display in ParaView, facilitating an intuitive understanding of wave propagation and interaction phenomena using the python code loop at the bottom of the image that creates an animation displaying the full motion of the pulse.

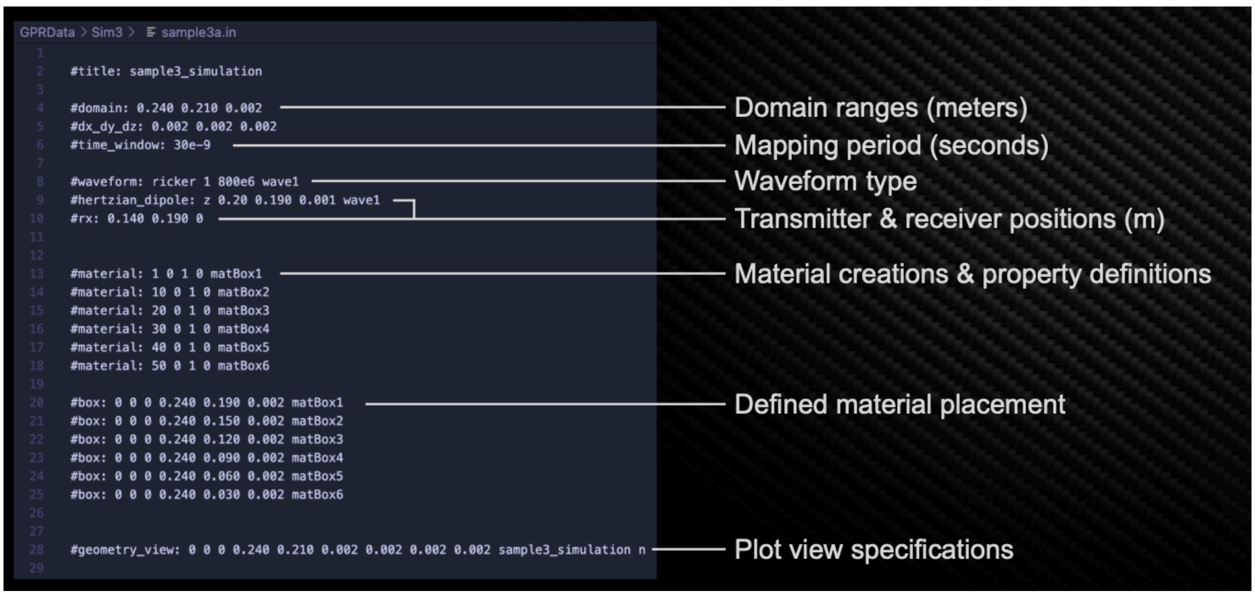

Figure 3.1.1.1 gprMax

```
#title: samp16e

#domain: 1.0 1.1 0.002
#dx_dy_dz: 0.002 0.002 0.002
#time_window: 30e-9

#waveform: ricker 1 400e6 wave1
#hertzian_dipole: z 0.45 1.05 0.001 wave1
#rx: 0.55 1.05 0

#soil_peplinski: 0.848 0.011 3.0 80 0.169 0.270 peplinskiLayer1
#soil_peplinski: 0.848 0.011 3.0 80 0.088 0.237 peplinskiLayer2
#soil_peplinski: 0.848 0.011 3.0 80 0.116 0.201 peplinskiLayer3
#soil_peplinski: 0.848 0.011 3.0 80 0.018 .257 peplinskiLayer4
#material: 12 0.03 1 0 largeRock

#box: 0.00 1.0 0 1 1.1 0.002 free_space
#fractal_box: 0.00 0.8 0 1 1.0 0.002 1.5 1 1 1 30 peplinskiLayer1 fractalBoxL1
#fractal_box: 0.00 0.7 0 1 0.8 0.002 1.5 1 1 1 30 peplinskiLayer2 fractalBoxL2
#fractal_box: 0.00 0.4 0 1 0.7 0.002 1.5 1 1 1 30 peplinskiLayer3 fractalBoxL3
#fractal_box: 0.00 0.0 0 1 0.4 0.002 1.5 1 1 1 30 peplinskiLayer4 fractalBoxL4
#box: 0.45 0.55 0 0.55 0.65 0.002 largeRock
#box: 0.283 0.35 0 0.383 0.45 0.002 largeRock
#box: 0.616 0.35 0 0.716 0.45 0.002 largeRock
#sphere: 0.167 0.683 0.0 .05 largeRock
#cylinder: 0.833 0.766 0 0.833 0.866 0.02 .05 largeRock

#python:
from gprMax.input_cmd_funcs import *
import os
import numpy as np
import random
N = 30
for i in range(1, N+1):
    snapshot(0, 0, 0, 1.0, 1.1, .002, .002, .002, .002, i*(30e-9/N), 'snapshot16e' + str(i))
#end_python:
#geometry_view: 0 0 0 1.0 1.1 0.002 0.002 0.002 0.002 samp16e n
```

Figure 3.1.1.2 Variation E of our 16th test model

3.1.2 Peplinski Model

To best suit accurate real-world conditions, the Peplinski model was employed to characterize the soil layers in our simulations due to its robust ability to accurately reflect the electromagnetic properties of soils as a function of their moisture content, soil type, and density. This model is particularly advantageous in GPR simulations for several reasons:

1. **Dielectric Property Accuracy**: The Peplinski model effectively calculates soil's complex dielectric constant, crucial for predicting how electromagnetic waves interact with varying soil layers.
2. **Moisture Sensitivity**: It adjusts for changes in soil moisture, vital for GPR since moisture substantially influences radar signal behavior.
3. **Soil Type Flexibility**: The model adapts to different soil compositions, essential for accurately simulating diverse geological conditions.

4. **Improved Predictive Performance**: Integrating this model enhances the learning efficiency of our machine learning algorithms, improving their ability to identify subsurface features.
5. **Reliable Calibration**: Employing a validated model ensures our simulations are realistic, boosting confidence in both the simulation outcomes and subsequent machine learning analyses.

3.1.2.1 Progression and Refinement of Simulation Scenarios

Initially, our simulations focused on establishing a baseline with a standard set of conditions. As the models progressed, we incorporated more complex variables, such as additional subsurface objects and adjusted soil moisture levels derived from actual field samples. This not only enhanced the realism of our simulations but also expanded the dataset used for subsequent machine learning analysis. Significant modifications were made to streamline the simulation process including: the reduction of layer count from six to four, simplification of the model to reduce the original line count by over two-thirds which enhanced computational efficiency, and strategic object count reduction and variation designed to systematically study the impact of different subsurface features on wave behavior.

These simulations provided a rich dataset, capturing a wide array of potential subsurface scenarios. This dataset serves as the foundation for our machine learning model development, aimed at classifying and interpreting complex waveforms.

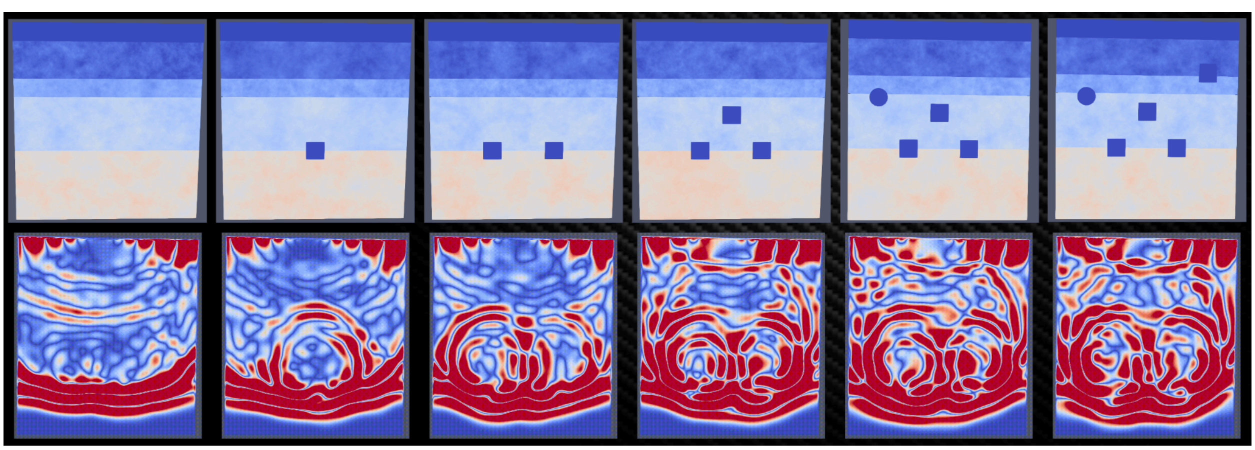

Figure 3.1.2.1.1 GPR Simulations

3.1.2.2 Integration of Machine Learning Models

Building on the synthetic dataset generated from gprMax simulations, we experimented with several machine learning models to classify subsurface features based on their effects on the GPR waveforms. Initial model experimentation employed included simple Multi-Layer Perceptrons (MLP), which later evolved into more complex architectures such as 1D Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). These models are particularly adept at handling the spatial and temporal data inherent in GPR waveforms.

- **MLP Models**: Served as a baseline for performance comparison.
- **CNN Models**: Leveraged for their ability to extract spatial patterns from waveforms.
- **RNN Models**: Utilized to capture temporal dependencies in the data, critical for understanding wave propagation over time.

Each model undergoes rigorous training and validation processes, with data split between training sets derived from simulated conditions and testing on unseen data to gauge generalization capabilities. The accuracy of these models were continuously benchmarked against new simulation data, driving iterative improvements in both the simulation parameters and the machine learning algorithms themselves.

The models were able to achieve this by analyzing the differences in the returned pulses, and examining the changes in return times. The green and red pulses in Figure 3.1.2.2.1 show an environment without objects, and the blue pulse shows the presence of an object. These measurements can then be inputted into our ML model to properly remove noise from the data for optimal analysis. The cleaned data can then finally be converted into the moisture values that can be interpreted through visual maps and other various outputs. Based on our experimentation with these models, we decided to proceed with a CNN model for our actual data.

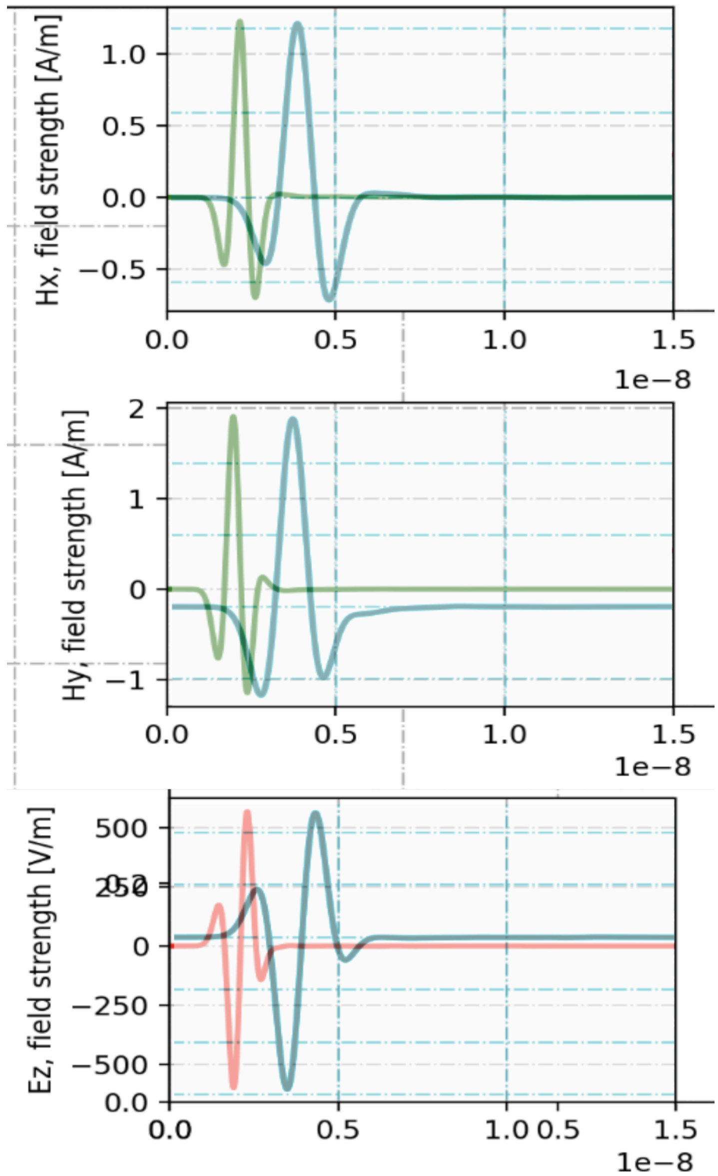

Figure 3.1.2.2.1 Simulated Environmental Scans

3.2 Determining the Status of the Radar

To assess the effectiveness of the radar in accurately measuring soil properties, we initially conducted tests using both the radar and moisture probe. The first radar model used was an Akila radar, but it failed to produce viable data. To address this issue, we switched to a SFCW radar, which offered higher resolution and a broader 10KHz band. After starting a new series of measurements with the SFCW radar and confirming the accuracy of its data, we proceeded with our sample data collection.

3.3 Creation of sample data

The sample data was gathered behind Gateway Park at Worcester Polytechnic Institute. SoilX had established a measurement site featuring six one-meter-deep holes for moisture probe assessments and two above-ground stands for radar setups, offering three different measurement heights: one, two, and three meters. Initially, we collected data from each in-ground hole using the moisture probe at varying depths: ten, twenty, thirty, sixty centimeters, and one meter. Subsequently, we conducted radar measurements at each height, taking three readings from the left, center, and right areas of the site, resulting in a total of nine measurements. These measurements were repeated with and without ground observers to filter out environmental noise. Finally, we correlated these radar measurements with the moisture probe data to evaluate how the radar readings corresponded to soil moisture levels.

Figure 3.3.1 Gateway Park SoilX Site Moisture Probe Measurements

Figure 3.3.2 Gateway Park SoilX Site Radar Measurements without observers

Figure 3.3.3 Gateway Park SoilX Site Radar Measurements with observers

3.4 Attaching the radar to the drone

The last phase of our measurement campaign involved mounting the radar onto the drone. SoilX had acquired an Aurelia X6 Max, capable of carrying up to six kilograms, which accommodated all the radar components. However, to ensure the drone's balance and prevent flight interference, we needed to reconstruct the radar into a more compact form. After reconfiguring, we successfully attached the radar to the drone.

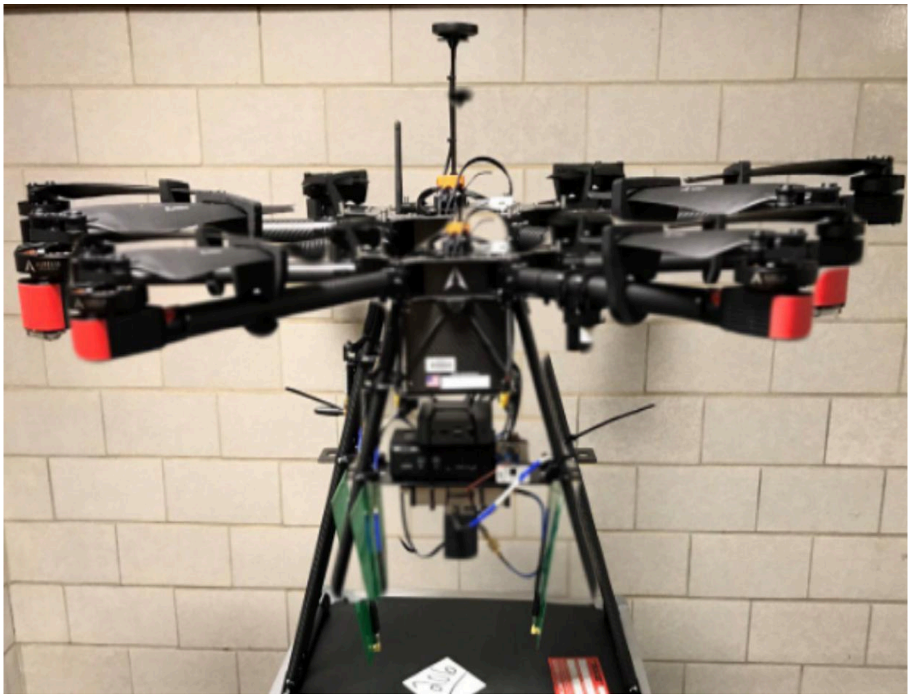

Figure 3.4.1 Aurelia X6 Max Drone with Radar mounted

3.4.1 Drone Capabilities

The Aurelia X6 MAX drone by Aurelia Aerospace is a versatile and robust UAV, designed to accommodate various operational needs and customization requirements. This model is particularly noted for its extended flight time and heavy payload capacity, making it suitable for our radar.

One of the standout features of the Aurelia X6 MAX is its flight time, which can reach up to 70 minutes with an advanced power system. This makes it one of the industry leaders in its category, allowing for prolonged operations without the need for frequent recharging or battery swaps. The drone also offers a substantial payload capability, able to carry up to 6 kg. This high capacity is facilitated by its robust frame and advanced motor setups, which are designed to ensure reliability and safety during flight, even in the event of a motor failure. The hexacopter configuration allows the drone to continue flying with only five motors operational, enhancing its resilience and operational security. This is optimal for our radar as it is within the payload capability and will not impact the flight capabilities of the drone when mounted.

In terms of operational range, the Aurelia X6 MAX can operate over distances of up to 5 km when equipped with advanced communication systems like Skydroid or HereLink. This extended range is beneficial for missions requiring broad area coverage or remote operations, such as surveying a multi-acre farm. Overall, the Aurelia X6 MAX is a high-performance drone suitable for a variety of demanding applications, from aerial surveying to payload delivery, bolstered by its significant customization capabilities and robust design.

4: Results and Analysis

In this chapter, we will discuss the outcomes of the radar measurements we conducted, focusing on the data collected and its implications. Additionally, we will delve into the machine learning model we employed to analyze this data, including a detailed exploration of a Convolutional Neural Network (CNN). We will cover how this model was implemented, its performance in interpreting the radar data, and the insights provided, which are crucial for understanding the effectiveness of our radar technology in real-world applications.

4.1 Results of the Radar and Moisture Probe

Our first type of data collected was radar data. A total of 500 high-resolution spectrogram images generated from SFCW radar sweeps were collected. Each spectrogram represents the reflected radar signals from subsurface structures, varying based on moisture content and soil composition.

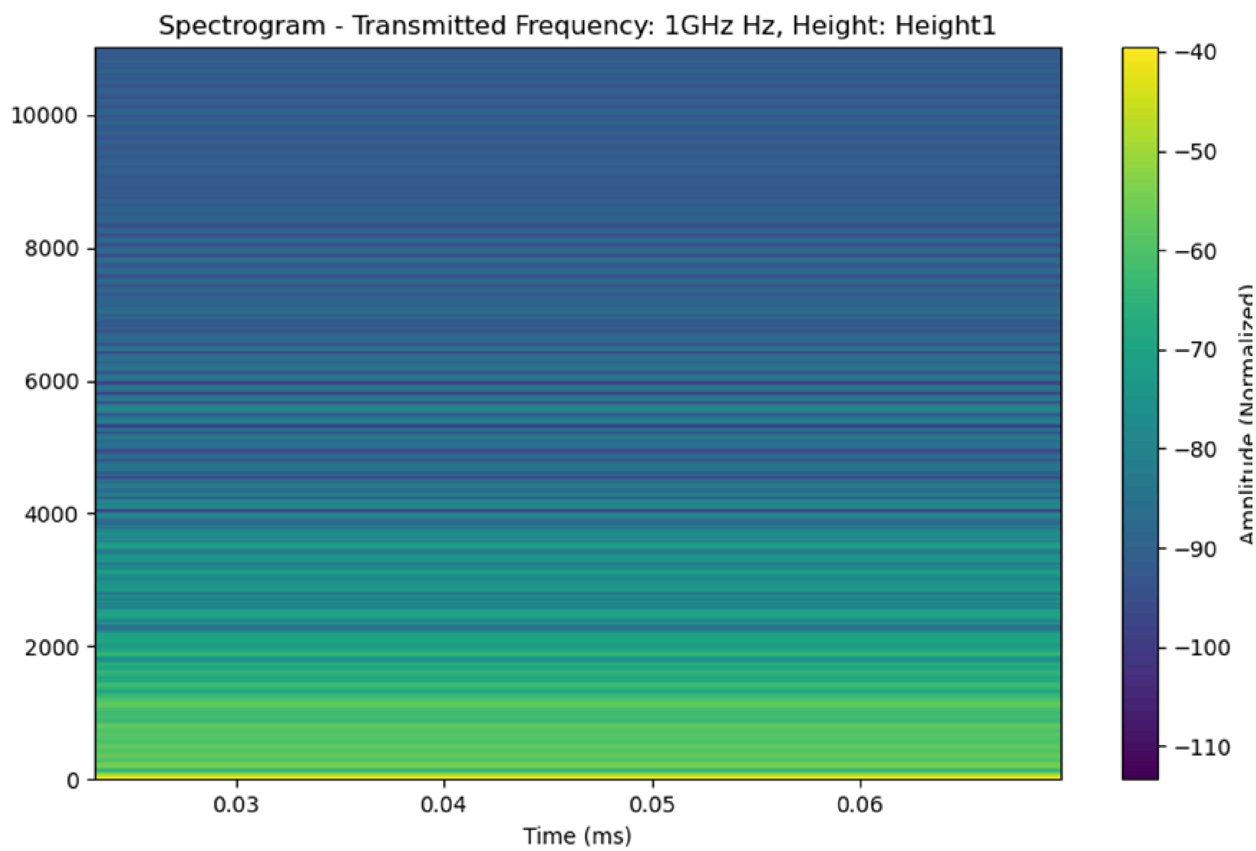

Figure 4.1.1 Spectrogram

The other type of data collected was moisture probe data. Moisture readings were meticulously recorded at six stratified depths ranging from 100 mm to 1000 mm beneath the surface, using advanced moisture probes. Each probe provided both percentage moisture content and electrical conductivity measurements in millivolts (mV), resulting in a comprehensive dataset comprising 12 distinct measurements per sampled location.

	A	B	C	D	E	F	G
1	Date & Time	10-30-2022	10-29-2022	10-23-2022	10-22-2022	10-21-2022	10-20-2022
2	A1_100_%	28.3	29.1	29.1	28.4	29.2	30.1
3	A1_200_%	32.3	34.7	32.8	31.9	33.8	33.7
4	A1_300_%	39	40.3	41.3	39.2	39.8	39.7
5	A1_400_%	50.7	52.2	52.5	50.8	54.2	52.4
6	A1_600_%	28.4	29.2	28.7	28.5	29.3	30.2
7	A1_1000_%	32.4	34.8	7.2	31.10	33.9	33.8

Figure 4.1.2 Moisture Data

4.2 Convolutional Neural Network

The primary objective of this phase was to develop an advanced predictive CNN model using Surface Frequency-Modulated Continuous-Wave (SFCW) radar data to accurately map subsurface moisture levels. This involved correlating complex patterns observed in radar spectrograms with moisture measurements taken at multiple depths to enable precise subsurface moisture monitoring.

A CNN is a specialized type of neural network that excels in processing data with a grid-like topology, such as images. CNNs are distinguished by their unique architecture, which includes one or more convolutional layers that automatically learn spatial hierarchies of features—from low-level edge features to high-level patterns specific to the task at hand—through a backpropagation algorithm. These layers use filters to perform convolution operations that capture the spatial relationships within the input data.

Following the convolutional layers, CNNs typically apply ReLU (Rectified Linear Unit) functions to introduce non-linear capabilities, allowing the model to learn more complex patterns. This is often complemented by pooling layers, which reduce the spatial size of the representation, decreasing the number of parameters and computation in the network, and thereby controlling overfitting.

CNNs have proven particularly effective in areas such as image and video recognition, image classification, and many tasks in computer vision, where they can identify patterns with extreme variability. The ability to automatically determine the best features for a given task, without needing manual feature extraction, makes CNNs highly versatile and scalable for image processing.

4.2.2 Data Preprocessing

The first step of data preprocessing was normalization and interpolation. Moisture data was first linearly interpolated to fill missing values and ensure completeness. Both datasets (percentage and mV) were then normalized using a MinMaxScaler to scale the features between 0 and 1, enhancing the neural network's convergence during training.

The next step in the data preprocessing was spectrogram processing. Each spectrogram image was processed to a uniform resolution of 128x128 pixels. Normalization was applied to adjust pixel intensity values to a [0, 1] scale, crucial for maintaining consistent input data format for CNN processing.

4.2.4 Feature Engineering and Integration

The feature engineering began with CNN feature extraction. The Convolutional Neural Network (CNN) was tasked with extracting and learning spatial patterns from the spectrograms, indicative of various moisture levels and soil compositions.

We then incorporated depth feature utilization. Instead of averaging depth readings, each depth-specific measurement was treated as a separate feature to retain detailed vertical moisture profiles. This approach allowed the model to learn and predict moisture content more accurately across different subsurface layers.

4.2.5 CNN Architecture for Spatial Feature Analysis

Layer (type)	Output Shape	Param #
conv2d_55 (Conv2D)	(None, 126, 126, 32)	320
max_pooling2d_55 (MaxPooling2D)	(None, 63, 63, 32)	0
conv2d_56 (Conv2D)	(None, 61, 61, 64)	18,496
max_pooling2d_56 (MaxPooling2D)	(None, 30, 30, 64)	0
flatten_29 (Flatten)	(None, 57600)	0
dense_58 (Dense)	(None, 128)	7,372,928
dense_59 (Dense)	(None, 73)	9,417

Figure 4.2.5.1 CNN Network Layers

The model included two convolutional layers. The first convolutional layer had 32 filters with a 3x3 kernel size, capturing low-level features such as edges and basic textures. The second layer increased to 64 filters, identifying more complex patterns as depth-specific moisture variations.

Each convolutional layer was followed by a MaxPooling layer with a 2x2 window, reducing feature dimensionality while retaining critical features. A flattening layer converted the 2D feature maps into a 1D feature vector for subsequent dense layers.

A dense layer with 128 neurons integrated and interpreted flattened features, followed by a final output layer predicting multiple moisture readings corresponding to various depths.

The model was trained using the Adam optimizer, chosen for its efficiency in handling sparse gradients and adaptive learning rate capabilities. The loss function was Mean Squared Error (MSE), penalizing the model for squared deviations between predicted and actual moisture readings. A validation split of 20% during training monitored and mitigated overfitting, which ensured the model generalized well to unseen data.

4.2.6 Performance Metrics and Evaluation:

Alongside MSE, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) provided different perspectives on prediction accuracy, aiding in understanding error magnitude and implications in practical scenarios.

The model underwent rigorous testing against a withheld set of data, with iterative adjustments made based on performance feedback, optimizing layer configurations, filter sizes, and learning

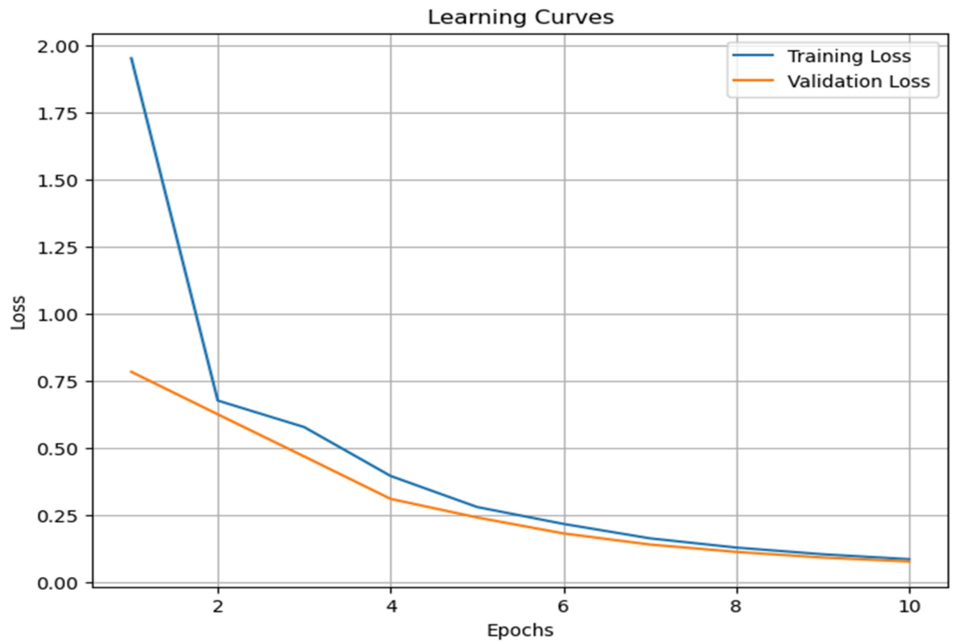

Figure 4.2.6.1 Learning Curves

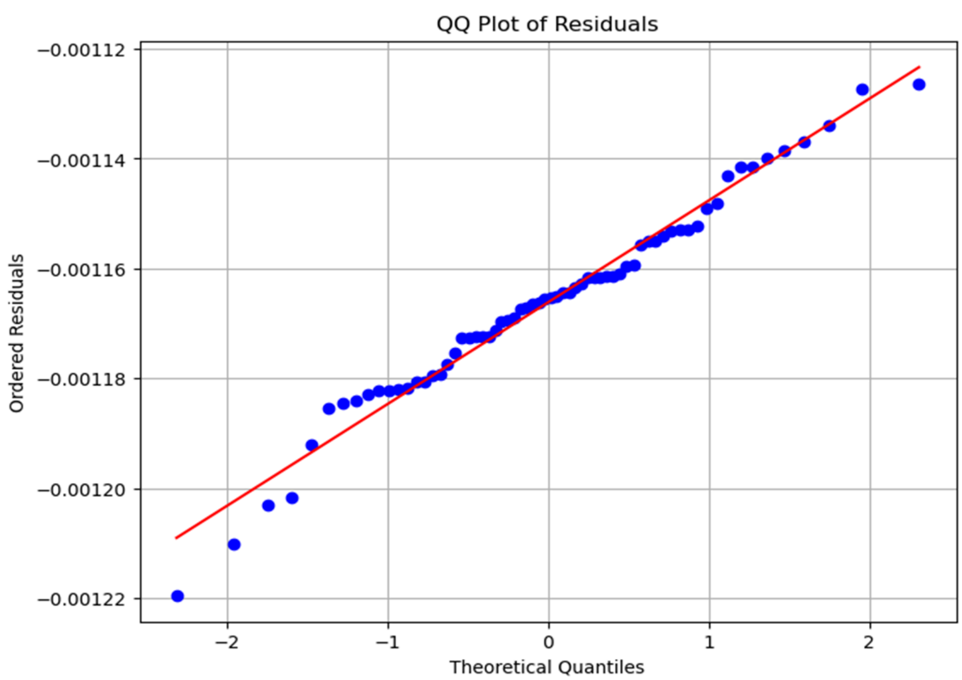

Figure 4.2.6.1 QQ Plot of Residuals

4.2.7 Significance and Implications

This project highlights the potential of machine learning in environmental and geotechnical applications, laying a foundation for future research into non-invasive soil analysis techniques. The developed model significantly aids water resource management, agricultural planning, and climate impact studies.

The successful development and validation of a CNN-based model to predict subsurface moisture levels using SFCW radar data represent a significant achievement in applying deep learning techniques to real-world environmental challenges. The model effectively interprets complex radar imagery and predicts moisture distribution with considerable accuracy, offering substantial benefits to scientific research and practical applications in soil and environmental sciences.

5: Conclusion

5.1 Summary of Current Work

Our project focuses on the development of a drone-based radar system for subsurface soil moisture detection. We have conducted extensive research and experimentation to design and implement this system, leveraging Ground Penetrating Radar (GPR) technology and advanced data analysis techniques. Our efforts have resulted in the successful creation of a dataset comprising 500 spectrogram images and corresponding moisture data across various depths. We have employed advanced preprocessing methods, including interpolation and normalization, to prepare the data for machine learning analysis.

Additionally, we have designed a Convolutional Neural Network (CNN) architecture to extract spatial features from the spectrograms and predict moisture levels accurately. Through rigorous training and validation, we have prioritized minimizing Mean Squared Error (MSE) to ensure precise predictions across different soil layers. The significance of our study lies in its potential to revolutionize environmental monitoring and agricultural practices by providing a reliable, non-invasive technique for assessing soil moisture. Our work underscores the innovative integration of radar technology and machine learning to address key challenges in environmental and soil sciences, paving the way for enhanced water resource management and sustainable agricultural practices.

5.2 Expanding Current Work

As the integration of Ground Penetrating Radar (GPR) simulations with machine learning continue to develop, future initiatives are aimed at significantly advancing subsurface exploration technologies. The focus will be on enhancing our simulation algorithms to more accurately model the complexities of subsurface environments, taking into account the diverse soil conditions and environmental factors that impact radar signals.

Moving forward, SoilX plans to refine our simulated models extensively and apply these improvements to real-world data. This approach is poised to revolutionize the methods used for detecting and analyzing subsurface features. The project will rely on continuous feedback from field tests to make necessary adjustments, ensuring the models perform effectively not only in simulations but also under varied and unpredictable real-world conditions.

Field testing will remain a crucial part of the development strategy, providing ongoing validation and refinement of the radar technology. By continuously integrating field data, we aim to ensure that the models are reliable across different environmental settings.

A significant aspect of the future work will involve merging advanced simulated models with more real-world data, thus bridging the gap between controlled laboratory conditions and the unpredictability of field environments. This synergy is expected to fine-tune the accuracy of the predictive models and extend their applicability across different geological settings.

More ongoing work includes the development of adaptive learning systems that can make real-time adjustments based on new data, enhancing the responsiveness of the models during live operations. Furthermore, fostering collaborative initiatives with both academic and industrial partners will be essential for broadening our data access and driving innovation, potentially leading to major breakthroughs in both GPR technology and machine learning applications.

Lastly, as we perfect and validate our models, preparing them for broader commercial use will be crucial. We envision facilitating the widespread adoption of the drone mounted GPR, particularly in the agricultural industry and other sectors dependent on precise subsurface analysis. Through these concerted efforts, we aim to significantly improve decision-making and resource management, transforming how industries engage with the subsurface world.

References

Lee, Kicheol, et al. “Prediction of Water Content in Subgrade Soil in Road Construction Using Hyperspectral Information Obtained Through UAV.” *Applied Sciences*, vol. 14, no. 3, Feb. 2024, p. 1248. https://doi.org/10.3390/app14031248.

Pramudita, Aloysius Adya, et al. “Soil Water Content Estimation With the Presence of Vegetation Using Ultra Wideband Radar-Drone.” *IEEE Access*, vol. 10, Jan. 2022, pp. 85213–27. https://doi.org/10.1109/access.2022.3197636.

Signal Hound, Inc. *VSG25 Programming Interface: Programmers Reference Manual*. 2019. Signal Hound, PDF.

Analog Devices, Inc. *PLL Software Installation Guide UG-476*. Rev. D, 2019, Analog Devices, PDF.

Analog Devices, Inc. *Evaluation Board User Guide UG-110*. Rev. 0, 2010, Analog Devices, PDF.

Analog Devices, Inc. *Evaluation Board User Guide UG-109*. Rev. 0, 2010, Analog Devices, PDF.

Analog Devices, Inc. *AD9915 Evaluation Board User Guide*. Rev. A, Analog Devices, PDF.

Analog Devices, Inc. *AD9914: 3.5 GSPS Direct Digital Synthesizer with 12-Bit DAC*. Rev. G, Analog Devices, PDF.

Signal Hound. *Spike User Manual*. Signal Hound, 2023, PDF.

“Aurelia X6 MAX.” *Aurelia Aerospace*, aurelia-aerospace.com/product/aurelia-x6-max.