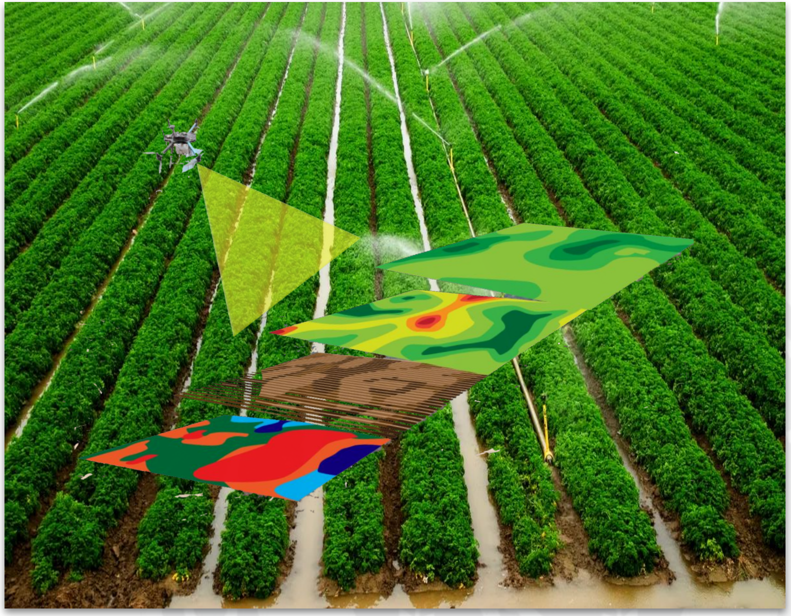
Comprehensive GPR Signal Analysis via Descriptive Statistics and Machine Learning

Himan Namdari*, Majid Moradikia*, Douglas Todd Petkie* , Radwin Askari†, Seyed Zekavat*
* Worcester Polytechnic Institute (WPI), Worcester, MA, USA
Email: {hnamdari, mmoradikia,dtpetkie, rezaz}@wpi.edu
† Michigan Tech Research Institute (MTRI), Ann Arbor, Michigan, USA
Email: {raskari}@mtu.edu

Unlocking Earth's Secrets with Ground Penetrating Radar

Subtimal farm Irrigation

- Stress on Plants.
- Soil Degradation.
- Nutrient Leaching.
- Water Wastage.



Solution

- Optimal irrigation by Estimating soil moisture.
- Air-coupled Ground Penetration Radar has a high potential to estimate the soil moisture in root zone.

Sources: https://www.afacademy.org/

GPR: Beyond the Agriculture Fields

Accurate soil moisture estimation plays a pivotal rule in many industries.

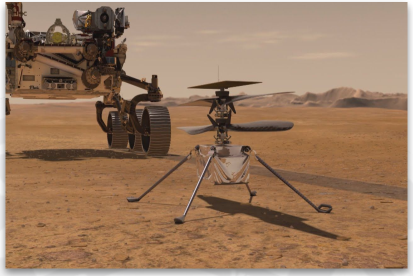
Forestry [2].

Civil Engineering & Construction [3].

Wildfire Assessment [4].

Planetary Exploration [5].

GPR Data collection.



Sources: https://www.afacademy.org/
1: Hubbard, S., Chen, J., Williams, K., Peterson, J., & Rubin, Y. (2005). Environmental and agricultural applications of GPR.
2: Lorenzo Cimadevila, H., Pérez Gracia, M. D. L. V., Novo, A., & Armesto, J. (2010). Forestry applications of ground-penetrating radar. Forest Systems (formerly: Investigación agraria: sistemas y recursos forestales), 19(1), 5-17.
3: Lai, Wallace Wai-Lok, Xavier Derobert, and Peter Annan. "A review of Ground Penetrating Radar application in civil engineering: A 30-year journey from Locating and Testing to Imaging and Diagnosis." Ndt & E International 96 (2018): 58-78.
4: Khanmohammadi, Sadegh, et al. "Prediction of wildfire rate of spread in grasslands using machine learning methods." Environmental Modelling & Software 156 (2022): 105507.
5: Bristow, Charlie S., et al. "Investigation of the age and migration of reversing dunes in Antarctica using GPR and OSL, with implications for GPR on Mars." Earth and Planetary Science Letters 289.1-2 (2010): 30-42.

GPR: A Glimpse Beneath the Surface

How does GPR work?

Transmitter emits EM waves into the ground. (Top figure)

Receiver records the reflected waves [6]. (Bottom figure)

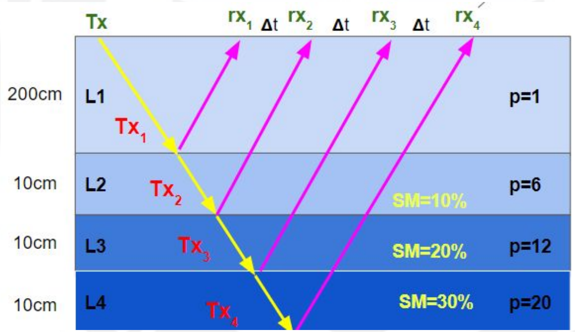
Why Soil Moisture?

Soil Moisture impacts and changes the received signal.

Signal power, speed, shape, and form.

GPR Radar

GPR Reflection Model

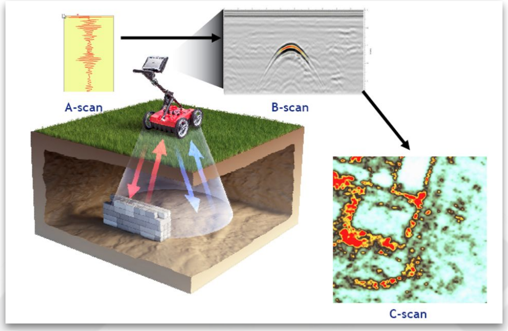


Sources: https://www.afacademy.org/
6: Jol, Harry M., ed. *Ground penetrating radar theory and applications*. elsevier, 2008.
7: Bradford, John H. "Measuring water content heterogeneity using multifold GPR with reflection tomography." Vadose Zone Journal 7.1 (2008): 184-193.

GPR Data Collection: Air vs. Surface Coupling

- Surface-Coupled Method
  - Soil and crop Destructive
  - Small survey area
  - Slow
- Air-Coupled Method
  - Non-Destructive
    - Altitude from the ground
  - Mega farms
  - Fast

Surface- Based



Air-Coupled

Sources: https://www.afacademy.org/
8: Wu, Kaijun, et al. "A new drone-borne GPR for soil moisture mapping." Remote Sensing of Environment 235 (2019): 111456.

GPR Principles and Soil Moisture

Uses electromagnetic waves to detect changes in sub-surface properties.

Soil moisture content by analyzing reflected signals.

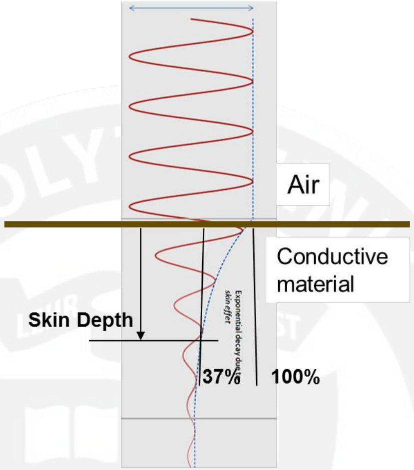
- Delta (δ) is depth
- Mu (μ) is permeability
- Sigma (σ) is conductivity
- Frequency (f)

Indirect relation between μ, σ, and f with δ

Higher frequency lower depth

Higher attenuation in deeper soil.

$$D = \sqrt{\frac{1}{\pi f \mu \varepsilon}}$$



http://blog.nutaq.com/blog/shielding

Sources: https://www.usnews.com/
https://impulseradargpr.com/
9: https://impulseradargpr.com/
10: https://www.sensoft.ca/range-of-gpr/
11: https://gpg.geosci.xyz/

Velocity, Permittivity, and Depth of Penetration

Analysing the impact of permittivity on the signal.

Permittivity of the layer was varied between 2-78.

- Slower signal propagation.
- Delayed reflection arrival.
- Attenuated faster at higher permittivity levels
- Direct relation between δ and v
- Indirect relation between v and ε

$$D = \frac{v}{2f}$$

$$v = \frac{c}{\sqrt{\varepsilon_r}}$$

v is the speed of the electromagnetic wave in the material.

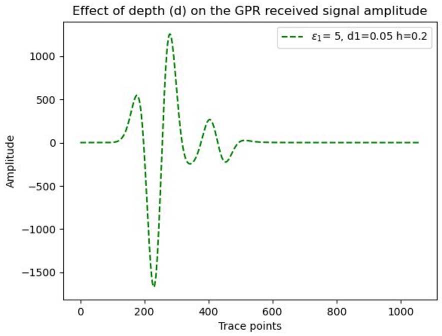
c is the speed of light in a vacuum, approximately 3×10^8 m/s.

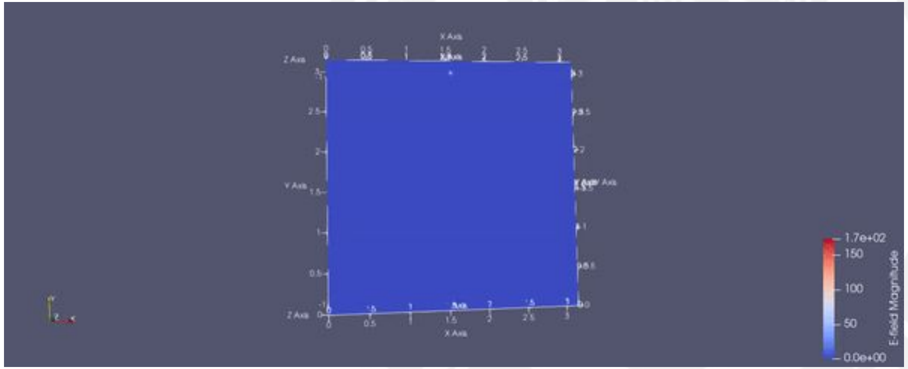
ε_r is the relative permittivity (or dielectric constant) of the material.

Sources: https://www.usnews.com/

gprMax Simulator

- gprMax
  - Open-source GPR simulator
- Application
  - Large-scale soil data generation.
- Method
  - Uses FDTD for accuracy.
- Goal
  - Generated labeled data
  - Train ML for soil moisture prediction.





Sources: https://www.sensoft.ca/
12: https://gi.copernicus.org/

Problem Definition

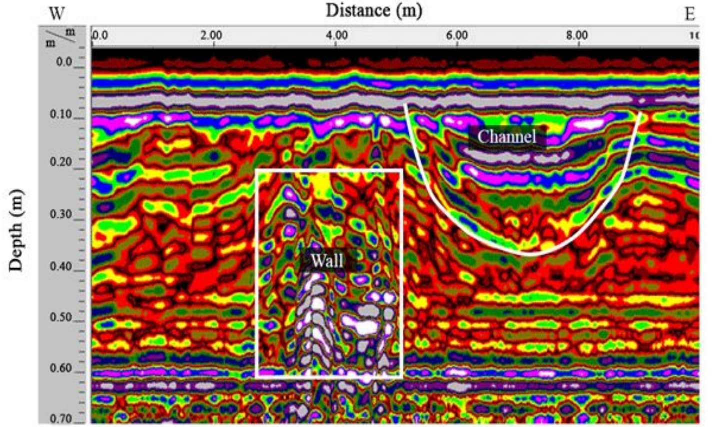
GPR Data Analysis Limitations

- Complex Data analysis process to extract features [12]...
- GPR received signal Labeling.

Our Contribution

- gprMax for data Augmentation and labeling.
- Feature extraction using descriptive statistic.
  - Input data size reduction and improving efficiency.
- Feature selection using random forest model.
  - higher data size reduction and more efficiency.

Example of Archaeological site scanned by GPR image



Sources: https://www.sensoft.ca/
12: https://gi.copernicus.org/

Our Contribution

Large-Scale Data Generation

- 10,000 GPR simulations
- 1-3 Layer Model with variation in permittivity.

Feature Extraction

Descriptive statistic (15 feature)

Feature Selection Random Forest (5 Top features)

Dimensionality Reduction

- 99% Training input size reduction

Machine Learning Modeling

Efficiency Enhancement

Sources: https://www.sensoft.ca/
12: https://gi.copernicus.org/

GPR Feature Types

Changes in soil properties can change the behavior of the received signal

Amplitude Patterns:
- Variations in signal strength.

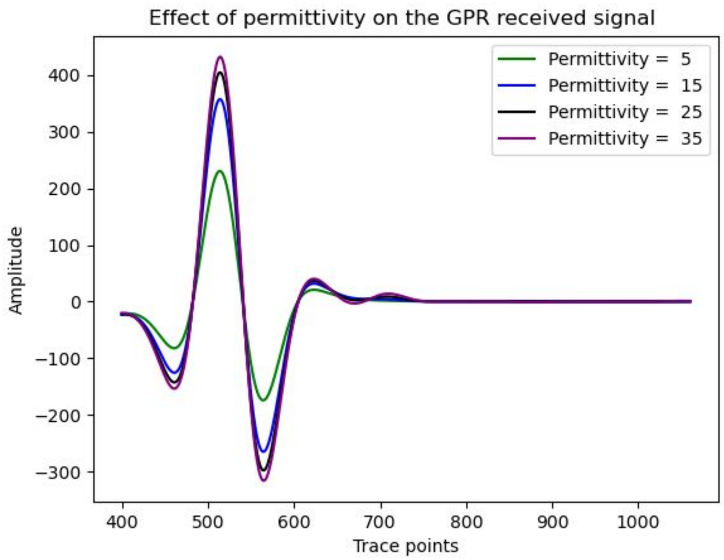
Waveform Shapes:
- Unique forms of the GPR signals.

Anomalies:
- Unexpected or unusual data points.

Time Delays:
- Time taken for signals to reflect back.

Frequency Components:
- Different frequency.

Texture Variations:
- Subsurface variations.



GPR Data Feature Extraction Methods

Feature extraction isolates the most important information from data for analysis.

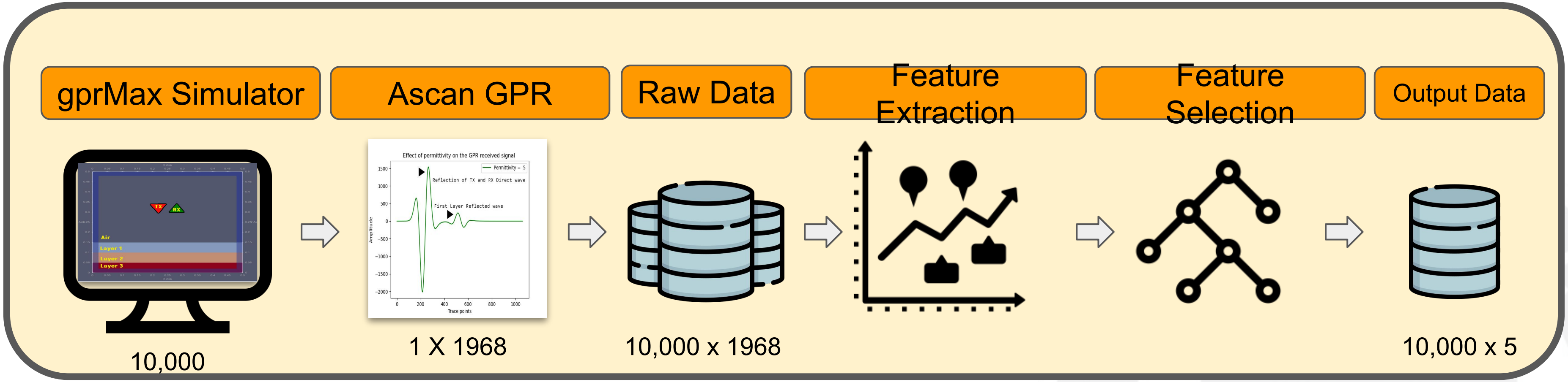
- Descriptive Statistic
  - Signal Amplitude
    - Mean, std
  - Signal Shape
    - Kurtosis
  - Signal Anomalies
    - Q1-Q4

Low Moderate High

Method	Complexity	Efficiency	Resolution	Adaptability
Wavelet Decomposition				
PCA				
Time-Frequency Analysis				
Statistical Measures				
Texture Analysis				
Descriptive Analysis				

Sources: https://openai.com/dall-e-2

Model Overview



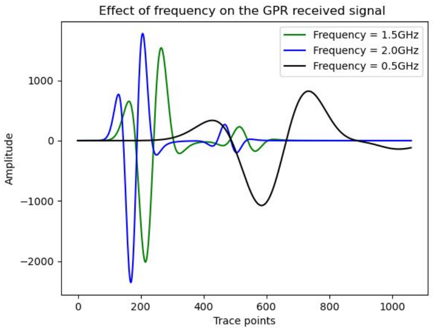
- Dataset Shape: a x b
- a = Number of simulation
- b = Number of features and labels

Sources: https://www.usnews.com/, https://www.flaticon.com/

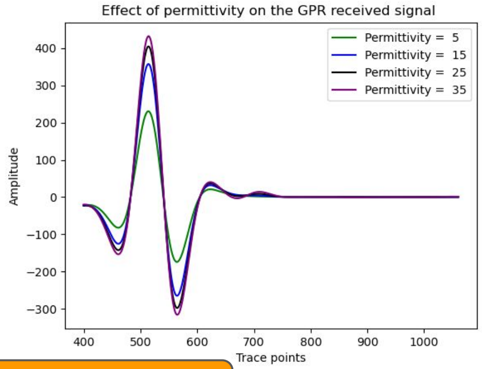
Methodology

1. Generate different soil scenarios
   a. gprMax
2. Analyse the impacts of
   a. Diverse permittivity
   b. Number of layers
   c. Depth Impact
3. Utilizing the DSFX method
   a. Extracting main features
4. Utilizing the RFFI method
   a. Extracting important features.
5. Evaluate the method

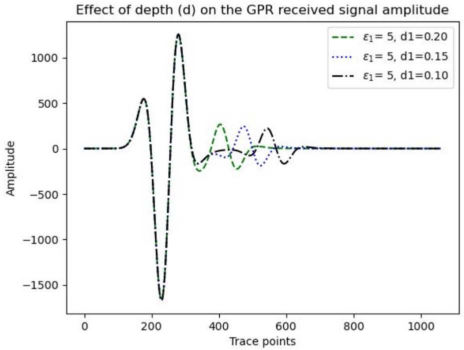
Frequency Impact



Permittivity Impact



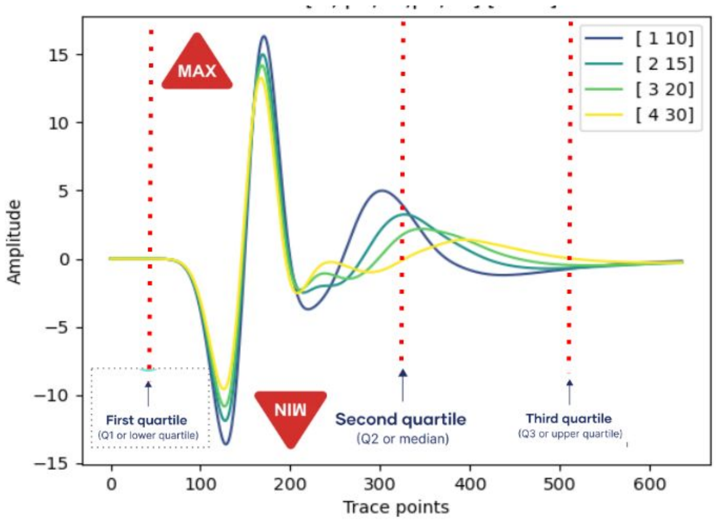
Depth Impact



Sources: https://www.usnews.com/

Descriptive Statistics Features Extraction (DSFX)

- Proposed Descriptive Statistical Feature extraction
  - In Time domain analysis
    - Extrating
      - Min, Max, Mean, Std, Q1-Q4
      - Skewness (SK), Kurtosis, entropy
  - In Frequency Domain
    - Min, Max, Mean, Std
  - Overall 15 Features
- Dimension Reduction
  - Raw features = 1967
  - Extracted features = 15
  - Percentage Reduction 99.24%

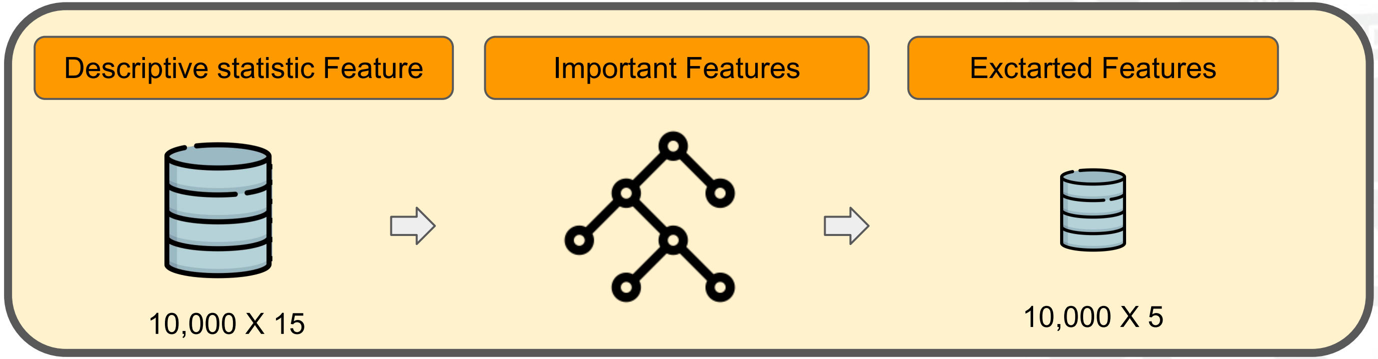


$$sk = \frac{3\,(\mu - Median)}{std}$$

Sources: https://www.usnews.com/

Random Forest Features Importance (RFFI)

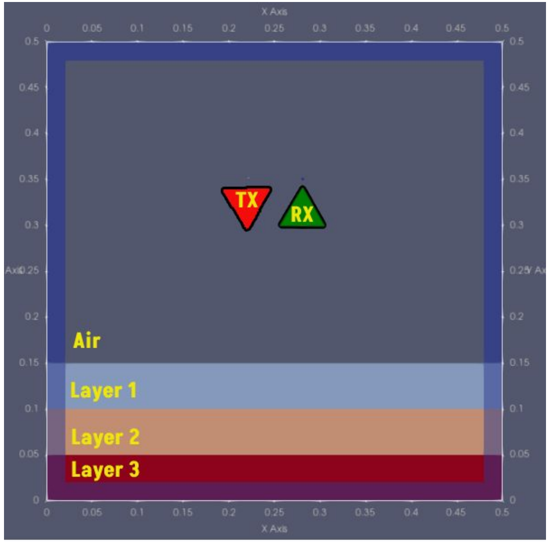
Setup
    gprMax, single layer, variable permittivity.
Input data
    15 GPR features.
    1000 simulation with labels
RF Model
    Trained and ranked features.
Outcome
    Top 5 features identified.



Sources: https://www.usnews.com/

Experimental Setup and gprMax Configuration

Simulated Environment



Simulation Model Parameters	Details
Model Type	1-3 Layer Model
Dimension	1.5 × 1.5 × 0.002
Propagation Time	5-8 ns
Frequency	1.5 GHz
Waveform	Ricker Pulse
Permittivity Range	2-78
Depth	0.01 - 0.1
Number of Simulations	1000+

Sources: https://www.usnews.com/

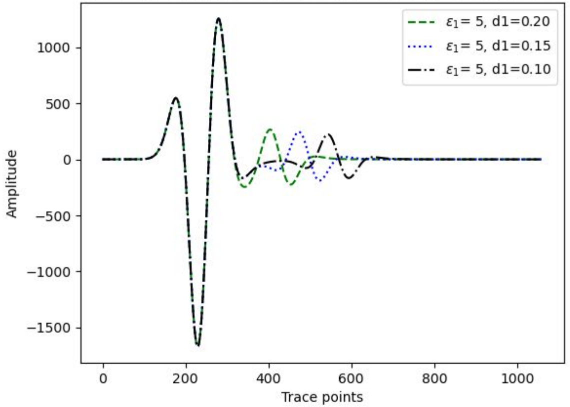
Experimental Results (Soil Depth)

Analysing the impact of depth on signal.

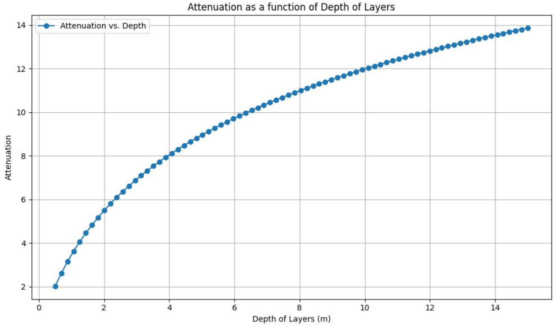
Depth of the layer was varied between 0.01-0.1 (m).

- Higher Attenuation
- Slower signal propagation.
- Delayed reflection arrival.
- Attenuated faster at higher permittivity levels.

Impact of Depth



Depth vs Amplitude



Sources: https://www.usnews.com/

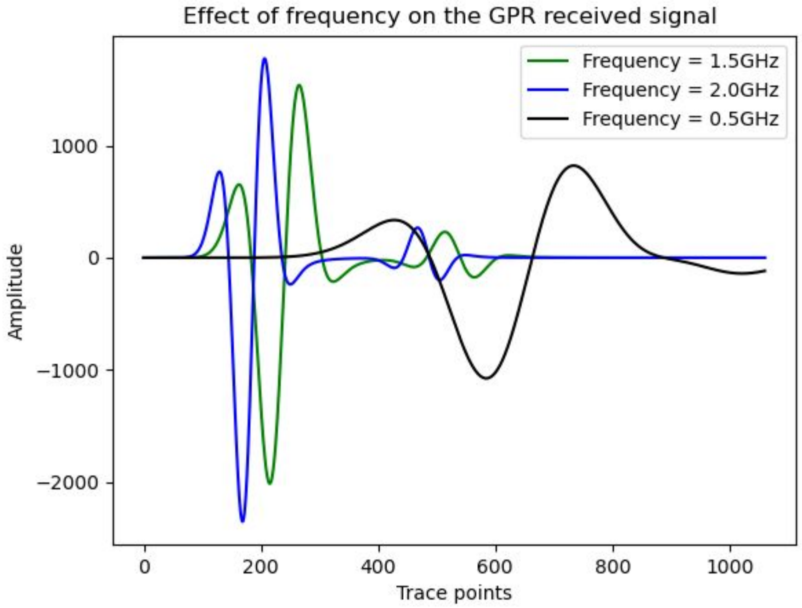
Signal Frequency Experimental Results

Analysing the impact of frequency on the signal.

Frequency was adjusted between 0.5 - 2.0 GHz.

- With increasing frequency:
  - Lower Signal propagation
  - Faster Signal attenuation
  - Lower depth of penetration.
  - Shape of the signal
    - 0.5 GHz harder to interpret

Impact of Frequency



Sources: https://www.usnews.com/

Descriptive Statistic Features Analysis

Features that Increase with Permittivity:

- Skewness
- Standard Deviation (Std)
- Third & Fourth Quartile

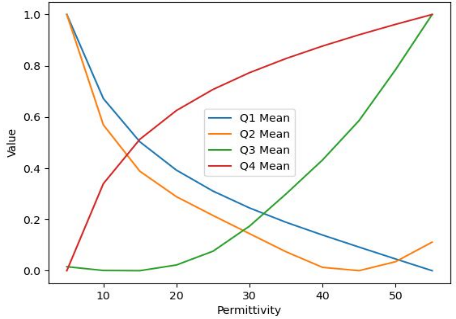
Features that Decrease with Permittivity:

- Second Quartile Mean (Q2 Mean)
- Third Quartile Mean (Q3 Mean)

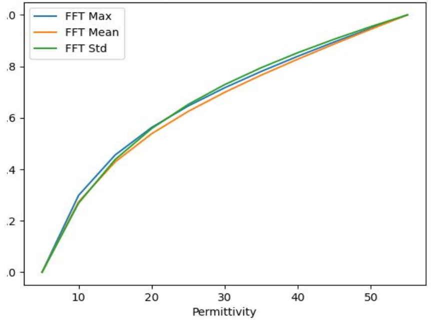
Features with Unpredictable Behavior:

- Minimum Value (MIN)
- Maximum Value (MAX)

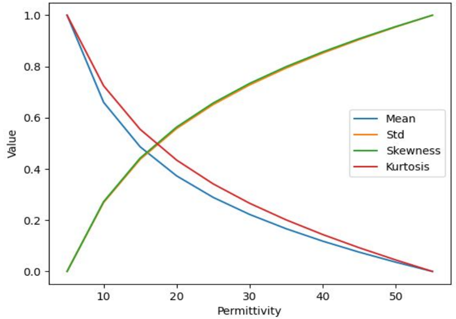
Q1-Q4



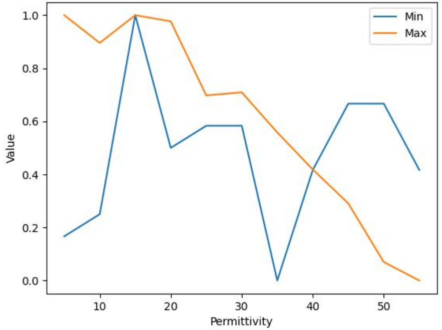
FFT



Mean -STD-Skewness



Min- Max



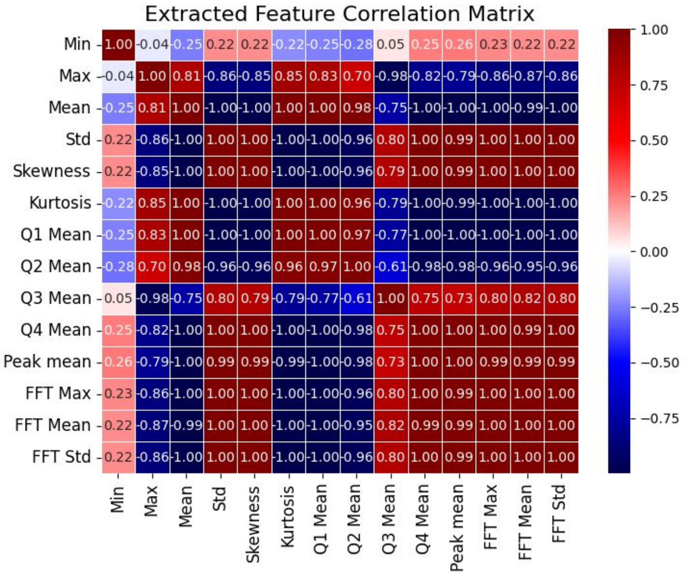
Sources: https://www.usnews.com/

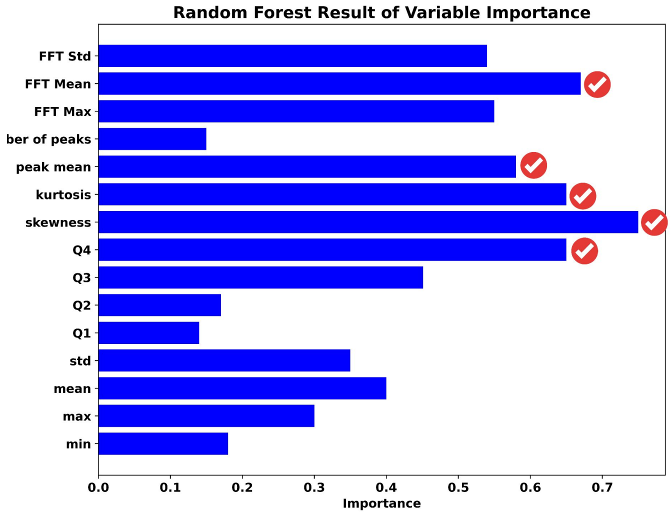
Random Forest Experimental Results

Top five Features:

FFT, Peak, and Q4 Mean

Kurtosis, Skewness





Sources: https://www.usnews.com/

Conclusion

Data Generation

Labeling

15 Features via Stats

Top 5 Features (RF)

NN: Minor RMSE Rise

Time Reduction

99% Data Size Cut

98% Faster Execution tTime

Feature Set	RMSE (Training)	RMSE (Test)	Execution Time (s)
Raw Feature Set	**0.003**	**0.004**	1.75
DSFX Feature Set	0.025	0.022	**0.025**

Sources: https://www.usnews.com/

Q & A

Sources: https://siwi.org/