Feasibility of Radio Frequency Based Wireless Sensing of Lead Contamination in Soil

Yixuan Gao
yixuan@cs.cornell.edu
Cornell Tech
New York, USA

Tanvir Ahmed
tanvir@infosci.cornell.edu
Cornell Tech
New York, USA

Mikhail Mohammed
mikhail.mohammed05@bcmail.cuny.edu
Brooklyn College of the City
University of New York
New York, USA

Zhongqi Cheng
ZCheng@brooklyn.cuny.edu
Brooklyn College of the City
University of New York
New York, USA

Rajalakshmi Nandakumar
rajalakshmi.nandakumar@cornell.edu
Cornell Tech
New York, USA

ABSTRACT

Widespread Pb (lead) contamination of urban soil significantly impacts food safety and public health and hinders city greening efforts. However, most existing technologies for measuring Pb are labor-intensive and costly. In this study, we propose SoilScanner, a radio frequency-based wireless system that can detect Pb in soils. This is based on our discovery that the propagation of different frequency band radio signals is affected differently by different salts such as NaCl and $Pb(NO_3)_2$ in the soil. In a controlled experiment, manually adding NaCl and $Pb(NO_3)_2$ in clean soil, we demonstrated that different salts reflected signals at different frequencies in distinct patterns. In addition, we confirmed the finding using uncontrolled field samples with a machine learning model. Our experiment results show that SoilScanner can classify soil samples into low-Pb and high-Pb categories (threshold at 200 ppm) with an accuracy of 72%, with no sample with >500 ppm of Pb being misclassified. The results of this study show that it is feasible to build portable and affordable Pb detection and screening devices based on wireless technology.

CCS CONCEPTS

• **Computer systems organization** → **Sensor networks**; • **Hardware** → **Sensor applications and deployments**; • **Computing methodologies** → **Machine learning approaches**.

KEYWORDS

Sensing Application, Urban Health, Soil Lead Contamination, Signal Processing, Machine Learning

1 INTRODUCTION

Urban soils are significant resources and provide essential ecological services such as growing produce, assimilation of organic waste, stormwater management, greening cities, improving air and water quality, and combating urban heat island effects [4]. However, numerous studies have shown that urban soils are often contaminated, primarily due to historical and current anthropogenic activities [28].

Lead (Pb), an invisible, odorless neurotoxin, is of particular concern, given its widespread presence in the environment and strong association with neurocognitive disorders and aggression in adolescents - especially for children [1, 24]. Pb is found to be present at elevated levels in urban soils worldwide. In New York City, it was found that over 50% of the garden soils tested contained more than 400 parts per million (ppm) of Pb – the previous general threshold set by the U.S. Environmental Protection Agency (EPA) and the New York State Department of Environmental Conservation [5]. In an official statement released on January 17, 2024, to strengthen the safeguards to protect families and children from Pb-contaminated soil, the U.S. EPA lowered the screening level for Pb in soil at residential properties from 400ppm to 200ppm [10], which is now the current screening standard. There is an urgent need to screen urban soils for traces of metal contaminants (such as Pb) as they can significantly impact public health and the safety of food grown in urban community gardens [22, 25].

Currently, composite soil samples are commonly sent to commercial or academic laboratories for analysis, utilizing chemical processing techniques and advanced instrumentation. Such analysis tends to require substantial labor and incur significant expenses [34], and thus poses challenges for many urban communities. These communities, often characterized by economic disadvantages, a prevalence of minority or marginalized populations, and a disproportionate burden of environmental contamination, face particular difficulties in accessing related resources. Furthermore, soil Pb is highly heterogeneous at a small scale, even within the same garden [12, 15]. Pb levels can vary by more than an order of magnitude at different locations within the same garden [3]. Thus, a composite soil sample is not able to reveal such variations and will miss hotspots that may pose the most health risks. While Portable X-Ray Fluorescence (pXRF[40]) has emerged as a handy tool for in-situ (or lab) screening of Pb and other metals in soils [11, 19, 27, 30, 33, 35, 42], the instrument typically costs $20,000-60,000, which is not affordable for most communities. Therefore, there is a need to develop low-cost alternatives to detect Pb in situ, so as to be able to accurately map sites for contamination. This study aims to address this need by examining the feasibility of developing an accessible and affordable Radio Frequency (RF) based wireless sensor that can monitor Pb.

As a rapidly emerging technology, radio frequency-based wireless signals have shown capacity in material-level sensing in recent years, such as liquid classification/testing [6, 7, 20, 29], food/fruit quality monitoring [14, 36]. However, in the field of soil sensing,

Table 1: Related work of RF-based soil sensing.

System	Method	Moisture	Salinity	Lead	Contribution Sumary
UHF RFID tags [2]	RFID - signal strength	Y	N	N	First COTS system testing soil moisture
Strobe [8]	Wifi - relative ToF	Y	Y	N	First work considering salinity of soil
GreenTag [39]	RFID - signal strength	Y	N	N	Greenhouse setting
CoMEt [16]	RF - ToF 2GHz-4GHz	Y	N	N	Moisture estimation without burying parts
SoilScanner[This Work]	Wifi + RFID	N/A	N/A	Y	First work considering individual salt in soil

material-level sensing is still challenging due to the complexity of the soil medium [43].

In this study, we aim to examine the feasibility of measuring contamination of Pb in soil through RF-based technologies. The key idea of this study is that when RF signals, such as Wi-Fi (2-5GHz) and RFID (Radio Frequency Identification) System [41] (∼ 900MHz) are transmitted by an RF transmitter, they are affected by the medium through which these signals propagate [31]. The propagation medium reflects, refracts, or absorbs the RF signal, and this property varies depending on the composition of the medium. By examining the signal at the RF receiver, we can infer the composition of the medium. We have observed that different frequencies of RF signals react differently to different salts. Thus, a sensing system that incorporates RF signals at various frequency bands is proposed here. Our experiments include the following: First, software-defined radio is used to demonstrate that Pb (lead chloride salt) has a distinct effect on different frequencies' RF signals than NaCl (sodium chloride, a common salt in soils) [37]. Second, a commercially accessible RFID device is used to see if the aforementioned observations remain valid. Next, a simple regression system is developed to demonstrate the feasibility of measuring the contents of two different salts, in this case $Pb(NO_3)_2$ and NaCl in the soil. We then built a machine learning model trained on 22 field samples and showed that we could classify Pb level in soil (at the current U.S. EPA screening level of 200ppm [10]) with an accuracy of 72% and recall of 80%. Finally, we discussed potential future work to make the sensor device robust and applicable to real-world soil conditions.

The main contributions of the work are as follows.

- We introduce SoilScanner, the first RF-based soil components analysis tool that can detect Pb contamination in soil. Existing RF-based wireless sensing systems have only been used for moisture and salinity sensing.
- We mathematically proved the feasibility of sensing different salts using RF signals. To our knowledge, this is the first work on modeling individual salt components in the soil.
- We open-source the first RF soil dataset that contains 23 lab-prepared "control" soil and 22 field soil samples. The dataset includes soil properties and RF signatures from our SoilScanner system.

2 RELATED WORK

Existing soil heavy metal (including Pb) detection: Widely adopted high-accuracy soil contamination detection techniques are all non-RF based, including inductively coupled plasma mass

spectrometry (ICP-MS) [23], Atomic Absorption, X-Ray Fluorescence [9], etc. The trade-off of these high-accuracy methods is the high labor costs and lab test requirements. While the portable XRF instrument [32] enables the in-situ testing capacity, its price is in the range of $20,000-60,000, which is not accessible to most economically disadvantaged communities or gardeners.

Terahertz spectroscopy is another emerging technology utilizing signals between RF range and laser optical region for heavy metal detection in soil [21]. By analyzing the reflectance of terahertz radiation, researchers can generalize the pattern of the signal. However, the terahertz sensor is also expensive ($1,500-3,000), and the portability is also a concern.

RF-based soil sensing: We summarized the existing RF-based soil sensing techniques in Table 1. Bechet et al. utilized a commercial off-the-shelf (COTS) RF system to test soil moisture [2]. Their work demonstrated the use of an RFID tag with a reader to estimate soil moisture. The moisture level is estimated from the signal power reflected from the tag. RFID has also been applied to greenhouse settings with improved accuracy [39]. However, salts in soil can affect the accuracy of single-band power-based moisture measurements because salt changes the electronic conductivity of the soil and affects the power of the signal received by the receiver. To overcome the salt interference of the power-based RF systems, researchers proposed methods that used the signal Time of Flight (ToF) instead of the power. Strobe [8] utilized a limited bandwidth WiFi frequency signal to measure the relative time of flight between the three antennae buried underground and then measured moisture by ToF and salinity by signal power. Note that the salinity measurements reveal the total concentration of all salts in the soil, instead of individual salt or contaminants. Additionally, Comet [16] proposed a reflection-based RF sensing approach for detecting moisture without disturbing the soil, such as excavation or embedding antennas underground.

To our knowledge, this study is the first attempt to differentiate individual salts in soil using RF signals, and this is an essential step for building an in-situ Pb measuring system in the future

3 BACKGROUND

3.1 RF signal propagation

RF signals are electromagnetic waves in the frequency range of 3KHz to 300GHz, including the bands that carry data through wireless communication systems such as Wi-Fi, FM radio, and Bluetooth. An RF system consists of a sender that transmits an RF signal, a receiver that receives both the direct and indirect reflections of the RF signal and a transmitting medium in between. Consider an RF signal: a single-tone cosine wave of amplitude A_s, frequency f, and

phase ϕ transmitted by the sender, the signal received is given by the following equation:

$$S_r(f, d) = A_s e^{-(\alpha+j\beta)d} \tag{1}$$

Where S_r represents the received signal, d represents the distance traversed by the signal, α represents the attenuation coefficient (signal loss per unit of distance traveled caused by the medium), and β represents the phase coefficient (phase shift per unit of distance traveled caused by the medium).

In any transmitting medium, the power of the signal decreases exponentially as distance (d) increases. If the transmission medium is solely air, the signal will spread out naturally and be attenuated along the free space path. For other transmission mediums, a portion of the signal is absorbed, and the rest either penetrates through or is reflected and scattered by the medium. The absorption/reflection ratio depends on the frequency of the transmitted signal and the characteristics of the medium. For example, a metal object is a good reflector of RF signal; RF signals of the lower frequency range can pass through an object while high-frequency RF signals are more reflected by materials.

Hence, the RF signal received by the receiver depends on three main factors: 1. frequency of the RF signal sent by the transmitter; 2. distance (d) between transmitter and receiver; and 3. characteristics of the transmission medium. In an RF sensing system transmitting signal and distance (d) between the sender and receiver are usually known, thus signal propagation will only be affected by the attenuation coefficient (α) and phase coefficient (β) of the transmitting medium. This can be used to infer and differentiate the composition of different mediums.

3.2 RF-Signal propagation in soil

In the context of soil medium, both moisture and salinity exert notable influences on the signal. Moisture affects the signal by predominantly changing the **dielectric constant ϵ_r'** of the medium. The dielectric constant is a fundamental property that governs the capacity of a material to store electrical energy. The dielectric constant of pure water is relatively high, at 80, in contrast to other prominent constituents found in soil, such as sand (ranging from 3 to 5) and air (1). The significant disparity in the dielectric constant of water and other materials allows for the possibility of determining moisture level through dielectric constant measurement. Directly measuring the dielectric constant is difficult; however, it can be estimated by utilizing the apparent dielectric permittivity of the soil, represented as ϵ_a, which can be determined through in-situ measurements. Apparent dielectric permittivity can be directly calculated through the velocity of the signal penetrating through the soil by equation $\epsilon_a = (c/v)^2$, where c represents the speed of light. In practice, determining velocity (v) involves measuring a signal's ToF as it traverses through the soil. Given that the distance, denoted as d, is known, the velocity (v) can be calculated by dividing the d by ToF. After apparent dielectric permittivity is measured, moisture content can be determined using the Topp equation [38].

Soils commonly consist of both moisture and various substances, including salts like NaCl and $Pb(NO_3)_2$. The presence of salt in soil has an impact on its electrical conductivity (σ), which refers to the soil's capacity to conduct an electrical current. Electrical conductivity impacts the signal power. In the controlled experiment,

we observed that different frequency signals exhibit different responses to different salts. Therefore, a multi-band RF-based system is designed to understand the potential for differentiating various soluble salts in the soil, including Pb which is a ubiquitous contaminant.

4 DESIGN

An overview of the SoilScanner system is shown in Fig. 1. The prepared soil (detailed in the implementation section) is placed in a 1L plastic container between the fixed transmission and receiving antenna. The transmitter emits short periods of single-tone RF signal at 700-1000MHz and 2.3-2.5GHz with a step of 0.5MHz. Then, the received power spectrum is passed to the machine learning (ML) model for further analysis.

4.1 RF signal affected by different soil salts

The insight of the multi-frequency-based power spectrum collection system can be summarized as follows.

(1) **The electrical conductivity (σ) [18] of each salt exhibits variations based on its density and exhibits distinct responses to alterations in the frequency of the surrounding electromagnetic field.**
The relationship between the density and electrical conductivity of different salt solutions exhibits variations when subjected to constant magnetic field frequency, temperature, and pressure. The varying current-carrying capacities and saturation levels of different ions offer an explanation.

(2) **The impact of the ambient magnetic field on various salt solutions varies.** For example, the application of a magnetic field leads to an increase in the conductivity of NaCl, $MgCl_{26}H_2O$, and KCl. The solubility of $MgSO_4$, $CaCl_{22}H_2O$, and $NaHCO_3$ exhibits a contrasting decrease. [13]

In a controlled environment with constant temperature and pressure, the conductivity of each salt i can be expressed as a function $\sigma_i(f)$ of the varying frequencies f of the applied electromagnetic wave. In addition, it is important to note that soil is composed of a diverse array of salts. The electrical conductivity (σ) of soil is a comprehensive measure that encompasses the collective impact of all these salts.

$$\sigma_{soil}(f) = \sigma_{S(\vec{d})}(f)$$

The density of each salt is denoted by the vector $\vec{d}$. And the function S is the combination function of all salts' conductivity.

When a soil sample is placed between a transmitting and a receiving antenna, the signal attenuation percentage can be measured and defined by the attenuation factor. [8]

$$\alpha = \frac{2\pi f}{c} \sqrt{\frac{\epsilon_r'}{2} \left[\sqrt{1 + \tan^2 \delta} - 1 \right]} \tag{2}$$

where ϵ_r' is the dielectric constant, c is the speed of light, and $\tan \delta$ is the loss tangent of the transmitting medium soil.

$$\tan \delta = \frac{\epsilon_r''(f) + \frac{\sigma_{soil}(f)}{2\pi f \epsilon_0'}}{\epsilon_r'} = l \tag{3}$$
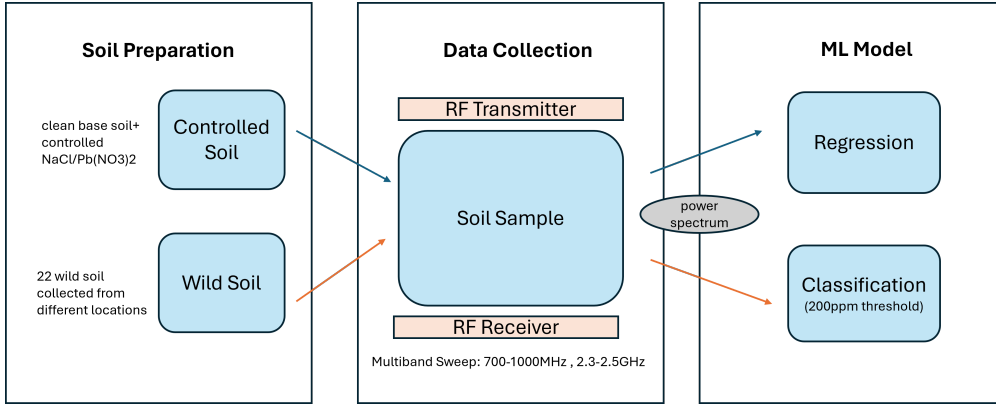
Figure 1: Flow chart of SoilScanner.

where $\epsilon_r''(f)$ is the dielectric loss, which is the energy required to heat a dielectric material in an alternating electric field. Loss tangent is noted using parameter l.

Therefore, at each frequency, when dielectric constant ϵ_r' is known, the power of the signal transmission is related to the conductivity of the materials and the dielectric loss of the materials.

$$Power_{soil}(f) \propto l_{soil}(f) = l_{S(\vec{d})}(f)$$

For each salt i, the loss tangent varies as the frequency changes due to the fact that the dielectric loss of each salt varies as the frequency changes [17] as well as insight 1 and 2 that electric conductivity of each salt varies differently when at different frequency magnetic field.

Since the objective is to determine each salt concentration in the soil, the question can be expressed as the following mathematical problem. Outputs consist of the densities $\{d_i\}_{all\ salts}$ of each distinct salt i. The input is the power transmitted and received at each frequency f, $Power_{soil}(f)$, which is proportional to the loss tangent of the combined salts when the moisture level and environmental factors (temperature, pressure) are known.

A learning system can then be developed. The system will accept $Power_{soil}(f)$ at distinct frequencies f as inputs and will output the density of each salt in the soil. The number of inputs will vastly exceed the number of outputs, resulting in a high capacity for estimation.

4.2 Learning algorithm for lab prepared controlled soil samples

We begin our analysis with controlled soil samples prepared in the lab. For these samples, we spiked "clean soil" with two different salts—NaCl and Pb(NO$_3$)$_2$, in varying quantities. Details of the soil preparation are provided in Section 5.3.

We first analyze the power spectrum of the data and then apply two key measures to build a regression model that estimates the concentration of Pb in the soil.

This section aims to develop a method for choosing parameters to distinguish between two different salts using data analysis. Since all other properties of the lab-prepared soil were the same except for the NaCl or Pb(NO$_3$)$_2$ content, a narrow frequency band is

sufficient to characterize the variation in Pb concentration. The selected frequency band showed the most significant correlation with Pb concentration.

Specifically, we introduce two key measures: Diff800 and Diff2300. Diff800 shows how much the power of the signal changes between 810.1MHz and 790.1MHz, while Diff2300 shows the same between 2408.6MHz and 2401.1MHz. By subtracting readings between two adjacent frequencies, the changes in the slope caused by Pb become more apparent and can be better evaluated. This process emphasizes the relative relationship between frequencies.

By concentrating on these measures, we hope to go beyond environmental influences and better understand how soil composition affects RF signal behavior. This structured approach provides a clear method for selecting parameters, making it easier to classify and differentiate between different salt types in soil samples. This contributes to the advancement of RF signal analysis in environmental research and monitoring.

4.3 Learning algorithm for uncontrolled field samples

Next, we analyze 22 field soil samples collected from various locations around a large metropolitan area. The details of the sample preparation are provided in Section 5.3. An ML-based model, shown in Fig. 12, was built for binary classifying soil by criteria of 200ppm of Pb, the new screening level for Pb in the residential area set by U.S. EPA in early 2024 [10].

4.3.1 Feature Engineering. To make the learning process easier and reduce near-frequency data redundancy, we design the following feature engineering techniques:

- **Frequency hopping**: instead of using all the frequencies', we select one frequency every certain interval.
- **Frequency aggregation**: similar to the hopping technique, we select the average value in that interval instead of selecting only one frequency every certain interval.
- **Weighted frequency aggregation**: as the name suggests, instead of giving every interval a uniform weight, we selectively weigh the intervals where we can see the most variance in the data.
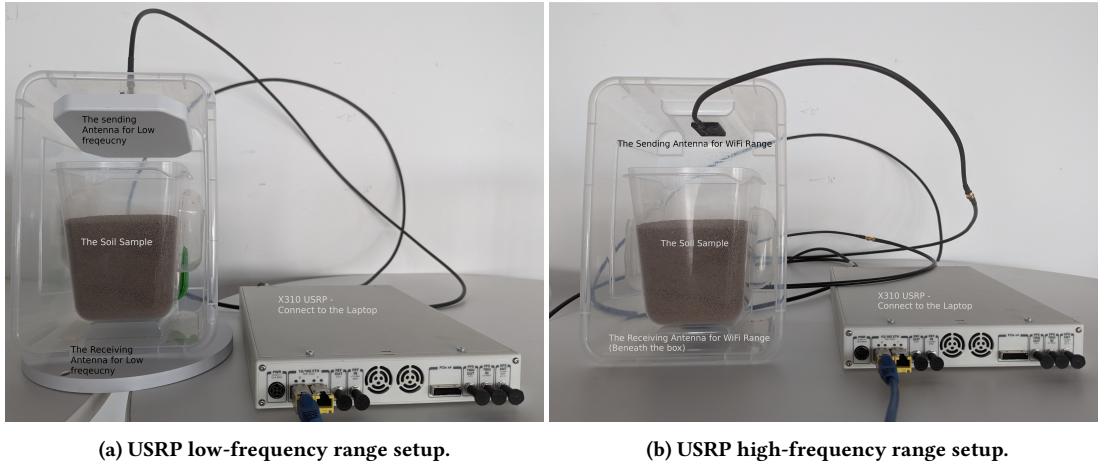
(a) USRP low-frequency range setup.

(b) USRP high-frequency range setup.

Figure 2: USRP setup.

4.3.2 Train-test-split.
Before splitting the dataset, we binary encoded each sample's Pb level as 0 and 1 by separating at the 200 ppm threshold. After encoding, 10 samples were classified as exceeding the amount of Pb (1), and 12 samples were classified as within the safe threshold (0).

We utilized Leave-one-out cross-validation (LOOCV) for the train-test-split for our dataset. LOOCV is a technique for evaluating a machine learning model's performance by training it on all but one data point and then testing it on the one held-out data point. This process is repeated for each data point, and the average performance across all iterations is used to estimate the model's generalization error. For smaller datasets, LOOCV has the benefit of providing an unbiased estimate of the model's performance, and it maximizes the information used for model assessment without sacrificing data. In our system, we repeat LOOCV k times and calculate the average performance of all k iterations for reduced bias of the noise added in the Data Augmentation step below.

4.3.3 Data Augmentation for Training Data.
For the model to be location-invariant and robust against noise, we need to increase the number of training samples. To do that, we designed a simple data augmentation technique that can up-sample the training dataset by a factor of r. First, all training samples were copied r times. Then, random noise, n, was added at each frequency power of each sample where, $n \sim \mathcal{N}(0, \sigma^2)$, and $\sigma \sim \mathcal{U}(0, 3)$. The reason for such a range for the standard deviation is that when we tested the same sample at different locations, we observed a standard deviation in the data that is similar to that range.

$$X \xrightarrow{\text{up-sample by r}} X_u \xrightarrow{\text{add Gaussian noise}} X_{augmented}$$

4.3.4 Model Training.
To achieve high robustness and performance, we utilized a soft voting classifier. A voting classifier is a type of ensemble learning method in which several base models are trained independently on the same dataset, and their predictions are combined through a voting mechanism to make the final prediction. In soft voting, the output probabilities of each base model are averaged for each class, and the class with the highest average probability is chosen as the final prediction. This method takes into account the confidence of each base model in its prediction rather than just the raw class labels. By combining multiple models with potentially different strengths and weaknesses, a voting classifier achieves better generalization and robustness compared to any individual base model. Our diversified base model includes Logistic Regression, SVM, Decision Tree, and Naive Bayes. Detailed model implementation is shown in Fig. 12.

5 IMPLEMENTATION

5.1 Hardware

A software-defined Universal Software Radio Peripheral (USRP), USRPX310 with two UBX-160 daughterboards operating at wideband frequencies ranging from 10MHz-6GHz was used. One daughterboard is the sender and is connected to a Laird S9025PL antenna for testing frequencies less than 1GHz. For frequencies above 1GHz, the sender daughterboard is connected to Taoglas Limited's ultrawideband FR4 antenna. The other daughterboard is the receiver and is connected to an EC742IOM06H-PIM-NF wideband antenna for frequencies less than 1GHz and Taoglas Limited's ultra-wideband FR4 antenna for testing frequencies above 1GHz. The two antennas are separated by 18cm and attached to the inner top of a plastic box with a dimension of 15cm*15cm*20cm. During an experiment, the sender antenna is attached to the top of a box, and the receivers are attached to the bottom. For all experiments, the soil is placed in a closed plastic box in the middle, 5-8cm from the sender and 2cm to the receiver, as shown in Fig. 2. The USRPX310 device is connected to a Lenovo Thinkpad laptop with a 1 GHz ethernet interface.

5.2 Software

The data acquisition system is programmed using GNU Radio version 3.10.5.1, an open-source software that controls the USRP with a graphic interface.

A sending signal comprises a base signal modulated on a carrier wave. The base wave is a cosine wave at a frequency of 100k. For each frequency range (modulated wave frequency 700-1000Mhz and

2300-2500Mhz), the initial carrier wave frequency is set to be the lower bound of that specific range, with a sample rate of 1 million samples per second and a gain value as the maximum value of the antenna capacity, 6dBm. The frequency shift is done on the carrier wave controlled by a global parameter $freq\_contorl$, updated at the receiver side.

On the receiver end, the raw signal is first received and demodulated at the USRP source block. Then, it is passed on to a Fast Fourier Transform (FFT) unit, which performs an FFT on every 1024 samples received. The FFT converts the time domain signal into the frequency domain. The power and standard deviation are then calculated by averaging 100 data points ($\approx$ 0.2sec) processed from the FFT processing unit. If the standard deviation exceeds a threshold of 0.02dBm, the program will recollect the signal for that specific frequency. The system will skip to the next frequency if it misses the standard threshold five consecutive times and report to the user. If successful, the data collection program will increase the $freq\_control$ variable by a step of 0.5MHz and collect data for the next frequency.

After all the data on the frequencies are collected, both R and Python are used for data analysis.

5.3 Soil Preparation

For evaluating our system, we prepared two sets of samples: 23 lab-prepared controlled soil samples, and 22 field-collected uncontrolled soil samples.

Lab prepared controlled soil samples: For the soil experiments, we chose loamy sand topsoil with low salt (<75 ppm), low Pb content (<5 ppm), and low organic content (<3%) as the base soil. The soil texture (loamy sand) is a common natural soil in the New York City (NYC) metropolitan area. The soil was air-dried for five days, passed through a 2mm sieve, and fully homogenized before being used. This soil was then divided into 1kg each and placed in a 1.7L IKEA Pruta plastic food container. The Pruta containers are built with thin plastic walls, reducing the signal loss when transmitting through the container. We then prepared two different salts. NaCl (table salt, with $97 - 99\%$ NaCl) and $Pb(NO_3)_2$ (a common soluble Pb compound). It should be noted that once $Pb(NO_3)_2$ is mixed with soil, Pb can react with various components in the soil and become less soluble.

Finally, we carefully spiked each sample using varying amounts of the above two salts, separately or as a combination of two salts. The different compositions are listed in the Table 2. The soil in each container was thoroughly mixed after spiking. All these samples were rested for at least 24 hours before testing.

Table 2: Soil variables for experiments in this study.

Sample ID	Base Soil mass(g)	NaCl (ppm)	$Pb(NO_3)_2$(ppm)
1-7	1600	[0,50,100,200,400,1000,2000]	0
8-14	1600	0	[200,400,1000,2000]
15-23	1600	[100,400,2000]	[100,400,2000]

Field collected uncontrolled soil samples: 22 distinct soil samples were collected from the field (gardens, farms, and other land uses in the NYC metropolitan area). The composition of these soil samples, including moisture, organic content, gravel, salt, pH, etc.,
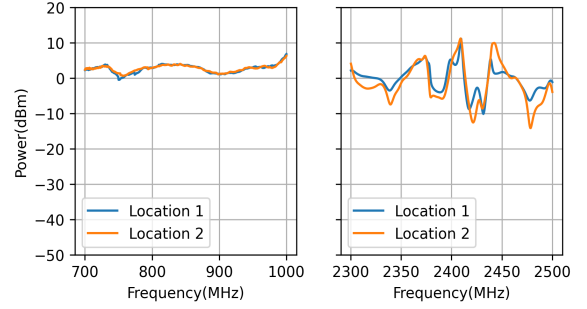

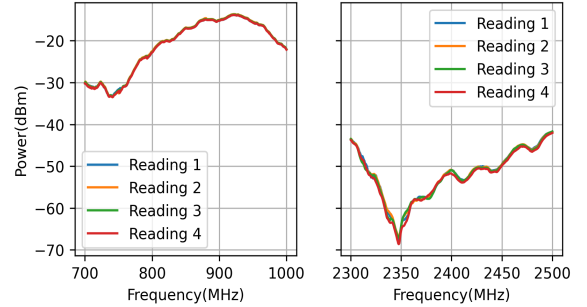**Figure 3: Location dependency test.**

Figure 4: Remounting test.

was carefully examined at the Urban Soils Lab at Brooklyn College. The uncontrolled samples offer substantial diversity beyond the lab-spiked samples. These analyses revealed significant variability among the samples, which will be discussed in detail in Section 6.4.1.

6 EVALUATION

We first demonstrate the robustness of the sensing system through its reprehensibility through two tests: the remounting test and the location test. Then, we analyze the effects of $Pb(NO_3)_2$, NaCl, and their combined effect on signals using controlled soil samples. Next, we verify whether these patterns are also observable in COTS RFID devices. Subsequently, we evaluate the performance of regression models built on controlled soil samples. Finally, we evaluate the machine learning model's performance for the uncontrolled field soil samples.

6.1 Location Robustness

One field sample's data was collected at two locations and then compared to understand the effect of location dependency in our RF sensing system. After removing the background signal, they are plotted in Fig. 3. Except for the small variation towards the higher frequencies, both signals are well-matched. The variation is expected because of interference with the WiFi signals present in the environment. However, this higher frequency interference can be further reduced by placing the SoilScanner system inside a metal box.
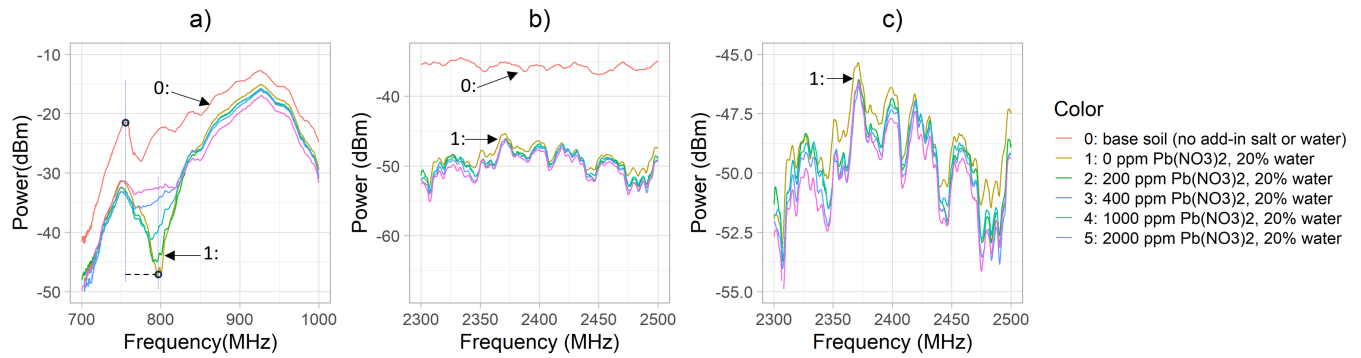
Figure 5: Received signal power (in dBm) variation with testing frequency a) 700-1000MHz, b)2300-2500MHz. Different colors represent different samples with varying Pb(No₃)₂ contents. c) is a zoomed-in view of b).
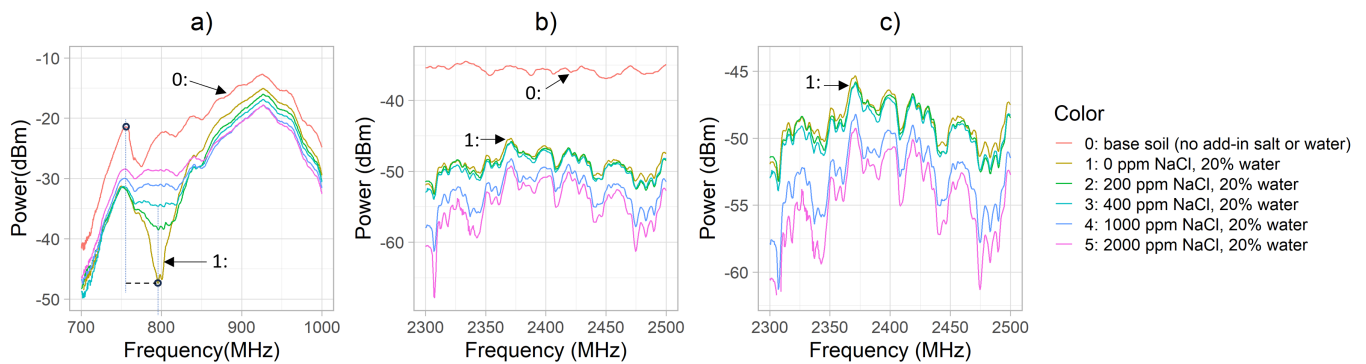

Figure 6: Received signal power (in dBm) variation with testing frequency a) 700-1000MHz, b)2300-2500MHz. Different color represent different samples with varying NaCl contents. c) is a zoomed-in view of b).

6.2 Remounting

A remounting test was performed to check the performance of our system. One field soil sample was tested four times at the same place, remounting the box every time between data collection. The signals are plotted in Fig. 4. As evident from the data, there is no visible difference in the received signal in the four different trials.

6.3 Analysis on lab prepared controlled soil samples

6.3.1 RFID and WiFi range signals for soils spiked with Pb(NO₃)₂.
Soil samples spiked with only $Pb(NO_3)_2$ (samples 15-18, see Table 2) were first tested with the RF setup at two distinct frequency ranges: 700-1000MHz with the Laird antenna and 2.3-2.5GHz with the WiFi high-frequency antenna. In addition to the spiked samples, two soil samples were tested to establish a baseline - 0: the base soil sample with no added water represented by the red line, and 1: the base soil sample with 20% of water represented by the lime green line. The average power of the received signal is plotted against frequency and shown in Fig. 5 and the average power of the received signal
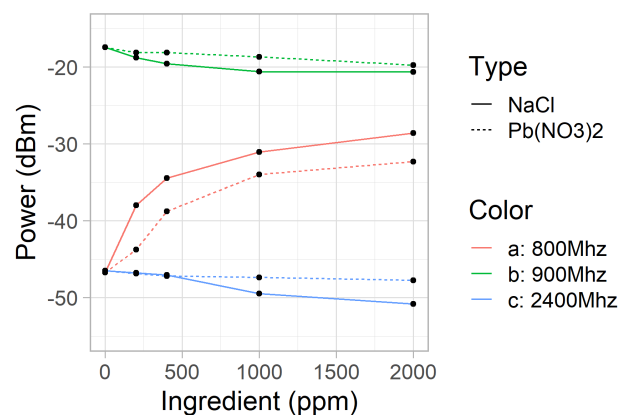

Figure 7: Received signal power (in dBm) variation with NaCl and Pb(NO₃)₂ concentrations. Different colors represent different frequencies of the tested signal.
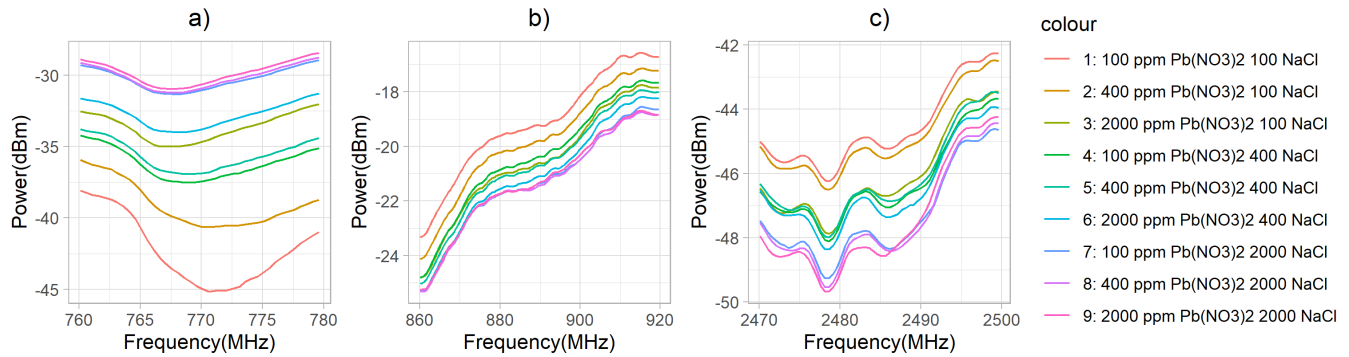
Figure 8: Different frequency range (x-axis) versus the RSSI (y-axis).

is plotted against $Pb(NO_3)_2$ concentration for selected frequency bands in Fig. 7. These results, while obtained in the same setup, were repeated over multiple days and were found to be consistent.

In Fig. 5a, When comparing the curve for the base soil (no water) and the curve for all other soils with water (moisture > 0), peaks for the latter slightly shifted towards 800MHz. This is because moisture can lead to higher permittivity of soil and a decrease in the speed of the signal that causes the peak shift. This is consistent with the findings in the literature that measured moisture content using RF technologies [8]. In Fig. 5a, if focused on 800MHz, The power of the received RF signal increases with $Pb(NO_3)_2$ concentration in soil (Fig. 7). The magnitude of increase between samples (e.g., ~ 5dBm for Pb concentration from 200ppm to 400ppm) is significant. This is because, with an increase of every 6dBm, the signal's amplitude is doubled. However, at frequencies higher than 825MHz, the received RF power appears to decrease with increased soil Pb content, shown as line b in Fig. 7. This shows that the effect of $Pb(NO_3)_2$ on received RF power varies at different frequencies, even within the same RFID band in the current test setup.

Fig. 5b plots the average power of the received RF signal in the WiFi frequency range (2.3–2.5GHz). It should be noted that the received RF power of base soil is very high (more than 10dBm higher) compared to all the other samples with moisture. Fig. 5c is a closer view of all the sample curves with 20% moisture. It can be seen that the power of the RF signal decreases with the increase in Pb content in the soil. This pattern is consistent within this frequency range. However, the magnitude of the decrease in signal power due to $Pb(NO_3)_2$ is slightly smaller than that in the RFID 900MHz band. Selected 2.4GHz band signal power response to soil Pb concentrations is plotted as line c in Fig. 7. Here the difference in reflected signal power between soil samples with 2000ppm Pb and that with no Pb is only about 3dBm.

This experiment demonstrates that variations in salt concentration impact the power of the received RF signal, and it varies at different frequencies.

6.3.2 RFID and WIFI range signals for soils spiked with NaCl . The above experiment was repeated with soil samples spiked with only NaCl salt, and the results are shown in Figs. 6 and 7. Overall, in the 700 to 1000MHz range, the observations are similar to $Pb(NO_3)_2$ spiked soils. For example, soil moisture shifted the peak towards 800MHz (Fig. 6a). The power of the RF signal increases with the NaCl content in soil at around 800MHz but decreases when above 825MHz (Fig. 6a and Fig. 7).

Within the WiFi frequency range (Figs. 6b and 6c), received power decreases with increasing content of NaCl, which is consistent with findings from the prior experiment with $Pb(NO_3)_2$. However, the magnitude of the decrease is higher for NaCl than $Pb(NO_3)_2$; this can be observed from the greater slope of the lines c in Fig. 7.

6.3.3 RFID and WiFi signals affected by a combination of $Pb(NO_3)_2$ and NaCl. The experiments with individual $Pb(NO_3)_2$ and NaCl showed that RF signals at different frequencies have different sensitivities to the two studied salts.

However, most soils from the field usually contain many different types of salt. So, an additional experiment was performed with soil samples spiked with varying concentrations of both NaCl and $Pb(NO_3)_2$ (Table 2, Samples 21-29).

Fig. 8 shows the power for the RF signal for all frequencies. RFID frequencies are plotted in two bands: 760-780MHz and 860-920MHz. The power increases with salt concentration in the 700-800MHz band, while the opposite trend can be observed in the 800-900MHz range. In the 2.4GHz spectrum, the signals are grouped into three clear ranges where the power is more attenuated when the NaCl content is at 2000 ppm. This is similar to the observation in the previous experiment (see Section 6.3.2) where NaCl affects the RF signal more than $Pb(NO_3)_2$ at 2.4GHz. This experiment shows that the effect of individual salts on reflected power signals is maintained with similar trends even when other salt(s) are present in the soil. While the effect by $Pb(NO_3)_2$ was maintained at the lower frequency band (700-900MHz), the effect by NaCl was maintained at the higher frequency band (2.4GHz).

6.3.4 Potential on commodity RFID. The above experiments were performed on expensive software-defined radios. The same experiment with a commercial RFID setup was repeated to prove the feasibility of using commercial off-the-shelf devices.

In this setup, a commercial Impinj R420 RFID tag Reader (Rev 3), which operates at 918-926MHz, was used. No additional hardware-level modification for the RFID reader was needed. The Laird S9025PL antenna was connected to the RFID reader using a Vulcan RFID

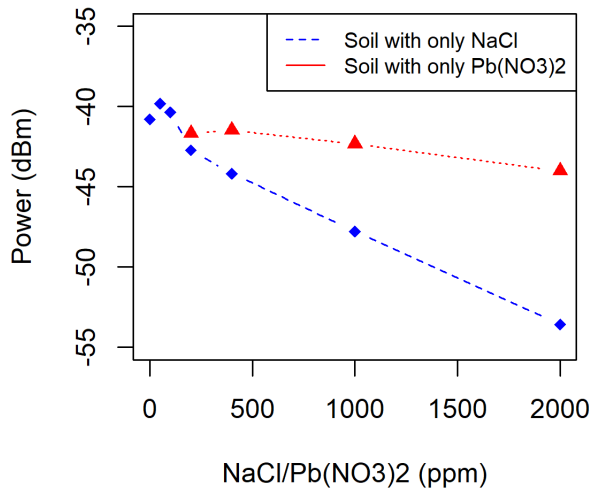**Figure 9: Experiment setup with commercial RFID.**

Figure 10: x-axis is the content of NaCl or $Pb(NO_3)_2$ in ppm, and y-axis is the power reading of RFID setup in dBm. The red line represents $Pb(NO_3)_2$ samples with 20% moisture, and the blue line represents NaCl samples with 20% moisture.

6 ft antenna cable and installed on the inner top of a plastic box with a dimension of 15cm*15cm*20cm (same as the USRP setting). The antenna has a 5.5dB gain and an elevation/azimuth beamwidth of 100 degrees. One EPC C1 GEN2 73X17mm UHF tag is attached to the inner bottom of the box, as shown in Fig. 9. The reader and the tag are separated by 18cm. We used Octane SDK with Java and ItemTest, an Impinj software, to control the reader activity. The reader's power was set at 30dB, and Rx sensitivity was set at -80dB. Each data collection session was 10 seconds, with tag-read activity

occurring at least 100 times. One average power is calculated using all the power readings of the specific tag (attached to the box) within the 10s period.

Fig. 10 shows the RF signal's power for different soil samples with commodity RFID setup. In a commercial RFID setup, only one average power value was recorded for each soil sample. For both NaCl and $Pb(NO_3)_2$, the power of the received signal is negatively proportional to the amount of salt in the sample. Furthermore, the slope representing how sensitive the RF power is changed due to the presence of salt is clearly different for the two studied salts. It is important to note that this observation is the same as that of the expensive software-defined radio setup at the 900MHz range.

6.3.5 Regression on controlled soil samples. The regression model shows the feasibility of separating Pb inside the soil. The R^2 of each pair of features and ingredient is shown in Table 3. Diff2300 can be used to estimate the level of NaCl because the linear relationship between NaCl and Diff2300 is robust with an R^2 value of 0.954; Diff800 can be used to estimate the level of $Pb(NO_3)_2$ because the linear relationship between $Pb(NO_3)_2$ and Diff800 is robust with an R^2 value of 0.920.

Table 3: Regression performance R^2 for Diff2300 and Diff800.

Ingredient	Diff800	Diff2300
$Pb(NO_3)_2$	0.920	0.017
NaCl	0.776	0.954

6.4 Analysis on uncontrolled field samples

6.4.1 Composition of the field samples. The studied field soil samples cover a wide range of moisture, pH, organic matter, gravel percentage, and heavy metals(including Pb). (Fig. 11). The average Pb content in those samples varied from 36ppm (minimum) to 1550ppm (maximum), with 12 samples below and 10 samples above the 200ppm threshold, moisture from 1% to 40%, and organic from 4% to 26%. The pH varied from 4.70 to 7.07, and the total salt amount from 58g to 710g. The composition of these samples is illustrated in Fig. 11.

6.4.2 Model performance for the field samples. The classification model's performance is evaluated by analyzing the confusion matrix. In this evaluation, we set model parameters $r = 200$ and $k = 100$.

Confusion Matrix:

Table 4 is the confusion matrix that visually summarizes the performance of a classification model. It compares the actual labels

Table 4: Confusion matrix for classifying Pb with the threshold of 200ppm.

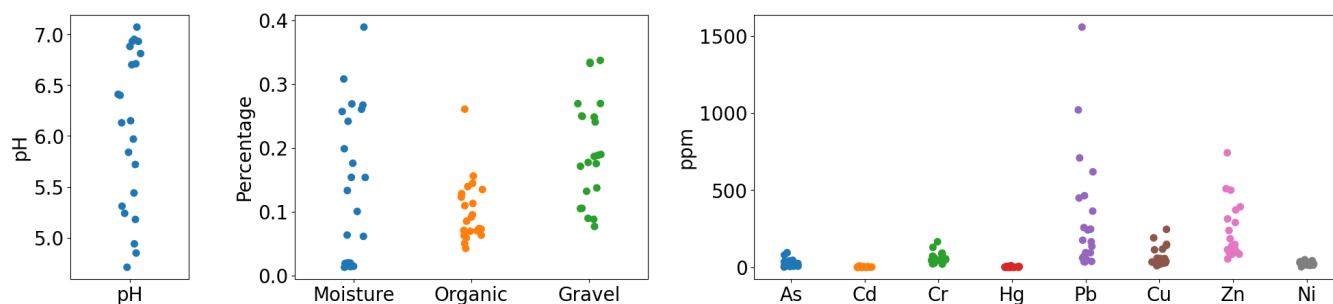|	Predicted Label	
True Label	Pb: <200 ppm	Pb >= 200 ppm
Pb: <200 ppm	0.36	0.18
Pb: >= 200 ppm	0.091	0.36

Figure 11: Composition of the field samples.

of the data (ground truth) with the labels predicted by the model. Rows represent the actual labels(Pb content from pXRF measurement) in the data. Columns represent the labels predicted by the model. Parameters in the table include:

True Negative (TN) - upper left: 36% where the model correctly predicted a negative class.

True Positive (TP) - lower right: 36% where the model correctly predicted a positive class.

False Positive (FP) - upper right: 18% where the model incorrectly predicted a positive class.

False Negative (FN) - lower left: 9% where the model incorrectly predicted a negative class.

we analyze the confusion matrix by following metrics:

Accuracy: 72%, the overall proportion of correct predictions $(TP + TN)/(Total)$.

Recall: 80%, defined as $(TP/(TP + FN)$, the ratio of true positives (TP) to the sum of true positives and false negatives (TP + FN), is a crucial metric for evaluating classification models. In the context of Pb detection, a high recall value is particularly significant. It indicates that the model effectively identifies a high proportion of samples with elevated Pb levels, reducing the likelihood of overlooking critical cases. This characteristic strengthens the system's potential as a screening mechanism, prompting further investigation for potential Pb contamination.

In addition, the false negative data points (which contain an excess amount of Pb but are classified as less than 200ppm) do not contain any of the four samples higher than 500ppm of Pb. We show that in our $22 * k$ iterations (where k is a parameter which indicates the number of times the ensemble model has been trained and tested), only the samples with 257ppm and 449ppm Pb were classified as false negatives. The low false negative rate of the model allows the system to be an excellent screening system for Pb detection. If our system's output is "above the threshold", the soil sample can be sent to a lab for further investigation.

7 LIMITATIONS AND FUTURE WORK

SoilScanner demonstrates the feasibility of utilizing wireless-based technologies for detecting Pb (and other salts or trace metals) in soil. Given the widespread presence of Pb as a metal contaminant in urban soils, the capability to detect it holds significant implications for public health, urban agriculture, and sustainability. However,

more work will be needed to achieve the goal of creating an affordable, accessible, portable, and non-invasive wireless sensor for in-situ Pb detection in soil.

7.1 Consideration of data collection and analysis approaches

The SoilScanner was constructed and tested using two sets of samples: 23 controlled, lab-prepared soil for regression, and 22 uncontrolled, natural field samples for classification. These two approaches represent distinct strategies for addressing the detection problem.

Controlled Sample Testing: Bottom-Up Approach

In the regression model, the goal was to establish a mapping comparison between the predicted values of the system and the exact Pb concentration with known methods (in this case, pXRF). This involved understanding the impact of various slats at different frequencies and determining the influence of soil composition. Calibration curves were developed for individual salts in this study NaCl and $Pb(NO_3)_2$ in a controlled setting. However, field soils have greater complexity and variability in properties. Additional experiments are necessary to examine the effects of other soil components.

Uncontrolled Sample Testing: Top-Down Approach

In contrast, for uncontrolled field samples, we used a binary classification approach based on a threshold rather than mapping exact Pb amounts. Acknowledging the variability in field soil samples, this approach considered the impact of Pb on signal propagation alongside other factors. The model can be generalized and further developed to support multi-class classification by incorporating various samples with varying properties. This requires lab experiments with soil samples collected from different locations to serve as validation for the developed model.

Integration of Approaches: Future Directions

While the two approaches necessitate different types of data collection and analysis, merging them offers a powerful strategy for future research. Leveraging the robust relationships established in controlled sample testing, this information can inform the binary classification model, reducing the complexity and sample size required for building an accurate and robust model. This integration of approaches represents a promising direction for advancing the capabilities of SoilScanner. We acknowledge the limited sample size in this study, and data augmentation, while helpful, is not a

perfect substitute for real samples. Despite this, the experiment yielded promising results, which we expect to improve with greater accuracy and generalizability in future research.

7.2 Complex environmental variables

The experiments in this study were conducted in a controlled lab setting. The application of the system to the real world requires consideration of other challenges, such as the presence of plants and other surface coverages, WiFi interference or objects that may affect reflection, as well as weather conditions that may affect RF propagation.

7.3 Portable chip manufacturing

This study has yielded promising results regarding the use of commercial RFID devices. Similar to RFID setup, the use of existing WiFi chipsets in smart devices that are smaller in size and work on batteries will be explored. If feasible, this hardware can then be put together to form a small sensor that is portable and can run on battery power. Research [26] has shown that it is possible to build a multiband RF system of size $15mm^2$ that can operate on three frequency bands namely RFID (800MHz), WiFi (2.4 and 5GHz), and run on button cell battery.

8 CONCLUSIONS

In this paper, we present SoilScanner, a multiband RF-based system for screening Pb contamination in soil. SoilScanner utilized the mathematically proved assumption that the signal power spectrum carries the information of individual salts in the soil. We show the feasibility of separating Pb in soil by building a regression model for controlled lab-prepared soil samples and a classification model for uncontrolled field soil samples based on the 200ppm threshold set by the U.S. EPA. Our results show that both models are robust, and the classification model has a zero error rate when Pb is > 500ppm. We are confident that SoilScanner is a significant step toward building a low-cost, easy-to-use, and high-accuracy Pb detection/screening system.

9 APPENDIX

The implementation details of the model designed for uncontrolled field samples are illustrated in Fig. 12.

ACKNOWLEDGMENTS

Figure 12: Pb classification system model.

REFERENCES

[1] Anjum Ara, Jawed Ahmad Usmani, et al. 2015. Lead toxicity: a review. *Interdisciplinary toxicology* 8, 2 (2015), 55–64.

[2] Rafael V Aroca, André C Hernandes, Daniel V Magalhães, Marcelo Becker, Carlos MP Vaz, and Adonai G Calbo. 2016. Application of standard EPC/GEN2 UHF RFID tags as soil moisture sensors. In *Proceedings*, Vol. 1. MDPI, 10.

[3] Béatrice Bechet, Sophie Joimel, Liliane Jean-Soro, Andrew Hursthouse, Alaba Agboola, Teresa E Leitão, Hugo Costa, Maria do Rosário Cameira, Cécile Le Guern, Christophe Schwartz, et al. 2018. Spatial variability of trace elements in allotment gardens of four European cities: assessments at city, garden, and plot scale. *Journal of soils and sediments* 18 (2018), 391–406.
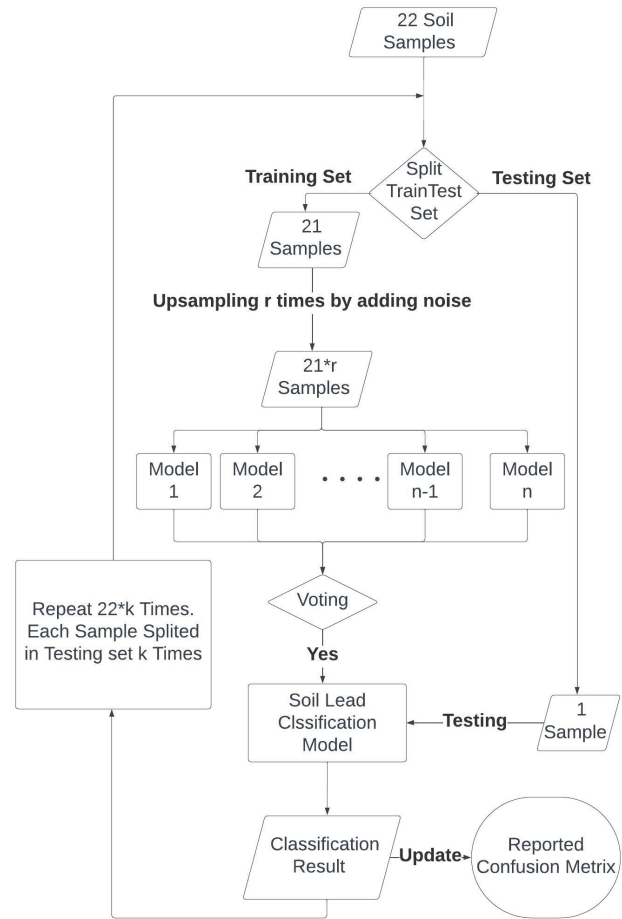
[4] Zhongqi Cheng, Ganga M Hettiarachchi, and Kye-Hoon Kim. 2021. Urban soils research: SUITMA 10. *Journal of Environmental Quality* 50, 1 (2021), 2–6.

[5] Zhongqi Cheng, Anna Paltseva, Ireyena Li, Tatiana Morin, Hermine Huot, Sara Egendorf, Zulema Su, Roxanne Yolanda, Kishan Singh, Leda Lee, et al. 2015. Trace metal contamination in New York City garden soils. *Soil Science* 180, 4/5 (2015), 167–174.

[6] Ashutosh Dhekne, Mahanth Gowda, Yixuan Zhao, Haitham Hassanieh, and Romit Roy Choudhury. 2018. Liquid: A wireless liquid identifier. In *Proceedings of the 16th annual international conference on mobile systems, applications, and services.* 442–454.

[7] Paul H Dietz, Darren Leigh, and William S Yerazunis. 2002. Wireless liquid level sensing for restaurant applications. In *SENSORS, 2002 IEEE*, Vol. 1. IEEE, 715–720.

[8] Jian Ding and Ranveer Chandra. 2019. Towards low cost soil sensing using Wi-Fi. In *The 25th Annual International Conference on Mobile Computing and Networking.* 1–16.

[9] MJ Dos Anjos, RT Lopes, EFO De Jesus, JT Assis, R Cesareo, and CAA Barradas. 2000. Quantitative analysis of metals in soil using X-ray fluorescence. *Spectrochimica Acta Part B: Atomic Spectroscopy* 55, 7 (2000), 1189–1194.

[10] EPA.gov. 2024. *EPA Strengthens Safeguards to Protect Families and Children from Lead in Contaminated Soil at Residential Sites in Region 7.* https://www.epa.gov/newsreleases/epa-strengthens-safeguards-protect-families-and-children-lead-contaminated-soil

[11] Asa Gholizadeh, Luboš Borvka, Radim Vašát, Mohammadmehdi Saberioon, Aleš Klement, Josef Kratina, Václav Tejnecký, and Ondřej Drábek. 2015. Estimation of potentially toxic elements contamination in anthropogenic soils on a brown coal mining dumpsite by reflectance spectroscopy: A case study. *PLoS One* 10, 2 (2015), e0117457.

[12] Christopher R Gonzales, Anna A Paltseva, Trevor Bell, Eric T Powell, and Howard W Mielke. 2021. Agreement of four analytical methods applied to Pb in soils from the small city of St. John's, Newfoundland, Canada. *International Journal of Environmental Research and Public Health* 18, 18 (2021), 9863.

[13] Shaoshuai Guo, Xueyun Han, Yufeng Peng, and Jiangting Li. 2016. Frequency Dependence of Conductivity Characteristics of Seawater Ionic Solution under Magnetic Field. In *MATEC Web of Conferences*, Vol. 67. EDP Sciences, 02021.

[14] Wen-Ding Huang, Sanchali Deb, Young-Sik Seo, Smitha Rao, Mu Chiao, and JC Chiao. 2011. A passive radio-frequency pH-sensing tag for wireless food-quality monitoring. *IEEE Sensors Journal* 12, 3 (2011), 487–495.

[15] Albert L Juhasz, Nicholas T Basta, and Euan Smith. 2013. What is required for the validation of in vitro assays for predicting contaminant relative bioavailability? Considerations and criteria. *Environmental pollution* 180 (2013), 372–375.

[16] Usman Mahmood Khan and Muhammad Shahzad. 2022. Estimating soil moisture using RF signals. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 242–254.

[17] Tsylya M Levitskaya and Ben K Sternberg. 2019. Parameters describing the material behavior in an electromagnetic field. *Electrical Spectroscopy of Earth Materials* (2019), 3–15.

[18] Edward Lewis. 1980. The practical salinity scale 1978 and its antecedents. *IEEE Journal of Oceanic Engineering* 5, 1 (1980), 3–8.

[19] Zhenhua Liu, Ying Lu, Yiping Peng, Li Zhao, Guangxing Wang, and Yueming Hu. 2019. Estimation of soil heavy metal content using hyperspectral data. *Remote Sensing* 11, 12 (2019), 1464.

[20] Humberto Lobato-Morales, Alonso Corona-Chávez, José Luis Olvera-Cervantes, Ricardo Arturo Chávez-Pérez, and José Luis Medina-Monroy. 2014. Wireless sensing of complex dielectric permittivity of liquids based on the RFID. *IEEE Transactions on Microwave Theory and Techniques* 62, 9 (2014), 2160–2167.

[21] Wei Lu, Hui Luo, Linxuan He, Wenxuan Duan, Yilin Tao, Xinyi Wang, and Shuaishuai Li. 2022. Detection of heavy metals in vegetable soil based on THz spectroscopy. *Computers and Electronics in Agriculture* 197 (2022), 106923.

[22] Marcella Malavolti, Susan J Fairweather-Tait, Carlotta Malagoli, Luciano Vescovi, Marco Vinceti, and Tommaso Filippini. 2020. Lead exposure in an Italian population: Food content, dietary intake and risk assessment. *Food research international* 137 (2020), 109370.

[23] Joseph M Matong, Luthando Nyaba, and Philiswa N Nomngongo. 2016. Fractionation of trace elements in agricultural soils using ultrasound assisted sequential extraction prior to inductively coupled plasma mass spectrometric determination. *Chemosphere* 154 (2016), 249–257.

[24] Prasenjit Mitra, Shailja Sharma, Purvi Purohit, and Praveen Sharma. 2017. Clinical and molecular aspects of lead toxicity: An update. *Critical reviews in clinical laboratory sciences* 54, 7-8 (2017), 506–528.

[25] Rajat Nag and Enda Cummins. 2022. Human health risk assessment of lead (Pb) through the environmental-food pathway. *Science of the Total Environment* 810 (2022), 151168.

[26] Rajalakshmi Nandakumar, Vikram Iyer, and Shyamnath Gollakota. 2018. 3d localization for sub-centimeter sized devices. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. 108–119.

[27] Anna A Paltseva, Maha Deeb, Erika Di Iorio, Luana Circelli, Zhongqi Cheng, and Claudio Colombo. 2022. Prediction of bioaccessible lead in urban and suburban soils with Vis-NIR diffuse reflectance spectroscopy. *Science of The Total Environment* 809 (2022), 151107.

[28] Yu Qian, Frank Gallagher, Yang Deng, Meiyin Wu, and Huan Feng. 2017. Risk assessment and interpretation of heavy metal contaminated soils on an urban brownfield site in New York metropolitan area. *Environmental Science and Pollution Research* 24 (2017), 23549–23558.

[29] Yili Ren, Sheng Tan, Linghan Zhang, Zi Wang, Zhi Wang, and Jie Yang. 2020. Liquid level sensing using commodity wifi in a smart home environment. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–30.

[30] Guy Schwartz, Gil Eshel, and Eyal Ben Dor. 2011. Reflectance spectroscopy as a tool for monitoring contaminated soils. *Soil Contam* 6790 (2011).

[31] John S Seybold. 2005. *Introduction to RF propagation*. John wiley & sons.

[32] Stephen Shefsky. 1997. Comparing Field Portable X-Ray Fluorescence (XRF) to laboratory analysis of heavy metals in soil. In *International symposium of field screening methods for hazardous wastes and toxic chemicals, Las Vegas, Nevada, USA*.

[33] Yinxian Song, Fengling Li, Zhongfang Yang, Godwin A Ayoko, Ray L Frost, and Junfeng Ji. 2012. Diffuse reflectance spectroscopy for monitoring potentially toxic elements in the agricultural soils of Changjiang River Delta, China. *Applied Clay Science* 64 (2012), 75–83.

[34] Rajneet Kour Soodan, Yogesh B Pakade, Avinash Nagpal, and Jatinder Kaur Katnoria. 2014. Analytical techniques for estimation of heavy metals in soil ecosystem: A tabulated review. *Talanta* 125 (2014), 405–410.

[35] Weichao Sun and Xia Zhang. 2017. Estimating soil zinc concentrations using reflectance spectroscopy. *International journal of applied earth observation and geoinformation* 58 (2017), 126–133.

[36] Sheng Tan, Linghan Zhang, and Jie Yang. 2018. Sensing fruit ripeness using wireless signals. In *2018 27th international conference on computer communication and networks (ICCCN)*. IEEE, 1–9.

[37] Ehsan Tavakkoli, Pichu Rengasamy, and Glenn K McDonald. 2010. High concentrations of Na+ and Cl−ions in soil solution have simultaneous detrimental effects on growth of faba bean under salinity stress. *Journal of experimental botany* 61, 15 (2010), 4449–4459.

[38] G Clarke Topp, JL Davis, and Aa P Annan. 1980. Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. *Water resources research* 16, 3 (1980), 574–582.

[39] Ju Wang, Liqiong Chang, Shourya Aggarwal, Omid Abari, and Srinivasan Keshav. 2020. Soil moisture sensing with commodity RFID systems. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services*. 273–285.

[40] David C. Weindorf, Noura Bakr, and Yuanda Zhu. 2014. Chapter One - Advances in Portable X-ray Fluorescence (PXRF) for Environmental, Pedological, and Agronomic Applications. Advances in Agronomy, Vol. 128. Academic Press, 1–45. https://doi.org/10.1016/B978-0-12-802139-2.00001-9

[41] Ron Weinstein. 2005. RFID: a technical overview and its application to the enterprise. *IT professional* 7, 3 (2005), 27–33.

[42] Yunzhao Wu, Jun Chen, Junfeng Ji, Peng Gong, Qilin Liao, Qingjiu Tian, and Hongrui Ma. 2007. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils. *Soil Science Society of America Journal* 71, 3 (2007), 918–926.

[43] Yihan Xu, Mikhail Smirnov, Michael C Kohler, Ziqian Dong, Reza K Amineh, Fang Li, and Roberto Rojas-Cessa. 2024. A Survey of Wireless Soil Sensing Technologies. *IEEE Access* (2024).