

# SoilCares: Towards Low-cost Soil Macronutrient and Moisture Monitoring Using RF-VNIR Sensing

Juexing Wang<sup>1,\*</sup>, Yuda Feng<sup>2,\*</sup>, Gouree Kumbhar<sup>3</sup>, Guangjing Wang<sup>1</sup>, Qiben Yan<sup>1</sup>, Qingxu Jin<sup>1</sup>, Robert C. Ferrier<sup>1</sup>, Jie Xiong<sup>2</sup>, Tianxing Li<sup>1</sup>

<sup>1</sup>Michigan State University, East Lansing, Michigan, USA

<sup>2</sup>University of Massachusetts Amherst, Amherst, Massachusetts, USA

<sup>3</sup>The Dow Chemical Company, Deer Park, Texas, USA

wangjuex@msu.edu, yudafeng@umass.edu,

{kumbharg, wanggu22, qyan}@msu.edu, {billjin, ferrier5}@egr.msu.edu, jxiong@cs.umass.edu, litanx2@msu.edu

## ABSTRACT

Accurate measurements of soil macronutrients (i.e., nitrogen, phosphorus, and potassium) and moisture play a key role in smart agriculture. However, existing commodity soil sensors are often expensive and the achieved accuracy is unsatisfactory. To address these issues, we present SoilCares, a low-cost soil sensing system enabling accurate and simultaneous monitoring of the concentration levels of soil moisture and macronutrients. SoilCares overcomes key challenges of accommodating diverse soil types and soil textures by introducing a novel membrane-based scheme. For moisture sensing, SoilCares leverages the multi-modal fusion of RF and NIR signals to significantly increase the sensing accuracy. Through delicate hardware design, we enable negligible-cost sensor data transmission using the existing sensing hardware, building up a complete end-to-end soil sensing system. SoilCares is cost-effective (\$63.5), portable (0.5 kg), and low-power (236  $\mu$ W), making it suitable for in-situ deployment. On-site experimental results show that SoilCares achieves high macronutrient sensing accuracy with a low RMSE of 0.138, and extremely low moisture estimation error of 1%, outperforming the state-of-the-art research and expensive commodity moisture sensors on the market.

## CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools.**

## KEYWORDS

Smart agriculture, Pervasive sensing, Multi-modality, Soil sensing  
Low-power and low-cost sensing

### ACM Reference Format:

Juexing Wang, Yuda Feng, Gouree Kumbhar, Guangjing Wang, Qiben Yan, Qingxu Jin, Robert C. Ferrier, Jie Xiong, Tianxing Li. 2024. SoilCares: Towards Low-cost Soil Macronutrient and Moisture Monitoring Using RF-VNIR Sensing. In *The 22nd Annual International Conference on Mobile Systems, Applications and Services (MOBISYS '24)*, June 3–7, 2024, Minato-ku, Tokyo, Japan. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3643832.3661868>

## 1 INTRODUCTION

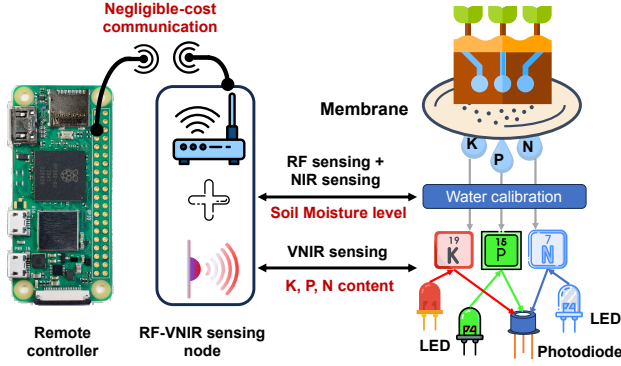
Precision agriculture, which refers to precise water, nitrogen, phosphorus, and potassium control in different farm regions, depends

on accurate soil moisture and macronutrient sensing. Proper fertilization enhances grain yields, while overusing fertilizers can lead to pollution of aquatic systems [74] and groundwater [15]. Moreover, proper soil moisture level facilitates plant absorption of essential nutrients, and precision irrigation has been proposed to save precious water resources and enable sustainable agriculture [24, 27, 33, 39]. Therefore, the ability to monitor the concentration levels of macronutrients and moisture has become a vital component in smart agriculture.

In recent years, several low-cost soil moisture sensing systems have been proposed, and they adopt various RF signals, including Wi-Fi, LoRa, LTE, and RFID techniques [13, 20, 22, 70]. However, these systems still lack the critical fertilizer sensing capability. The fundamental drawback that prohibits them from accurate fertilizer sensing is that RF waves with centimeter-level or even millimeter-level wavelengths are too coarse-grained to detect the variation of nutrients. In comparison, Vis-NIR (visible–near-infrared) waves have a nanometer wavelength, which is at the same scale as nutrient molecules and manifests unique physical properties of absorption/reflectance. As a result, Vis-NIR (VNIR) reflectance spectroscopy has become a prevalent method for analyzing soil properties [50, 55, 59]. However, this method requires a high-end spectrometer (~\$20,000 [8]) for generating a super-high-resolution reflectance spectrum and extracting nutrient-relevant information, which also introduces significant maintenance overhead. Moreover, it requires intricate pre-processing, such as grinding and drying, to mitigate the impacts of soil moisture, particle size, and other environmental factors. This has hindered the wide deployment of on-site fertilizer sensing.

To address the issues of reflectance spectroscopy while still leveraging the advantages of VNIR sensing, low-cost LEDs and photodiodes (PDs) are emerging as viable alternatives in agriculture sensing [5]. Combinations of LEDs can cover the discrete spectrum range from VIS to NIR, and photodiodes can serve as receivers of the reflected light. These devices are cheap, portable, and adaptable to various working environments with fewer restrictions. However, there are three major challenges associated with low-cost soil sensing systems. First, pre-processing steps like drying and grinding are critical for soil sensing due to significant influences of soil moisture and particle size on estimating other soil properties [49–51, 59, 72]. This is because soil moisture significantly affects the accuracy of macronutrient sensing. We need to totally remove it (i.e., drying) or estimate it accurately to remove its effect on macronutrient sensing. However, current low-cost RF-based solutions [13, 20, 22, 70] are

\*Both authors contributed equally to the paper.



**Figure 1: SoilCares system for macronutrients (N, P, K) and soil moisture measurements, with negligible-cost communication capability**

not able to achieve highly-accurate (<2%) soil moisture estimation. Without the grinding step, the large and random soil particle size also introduces unintended reflectance surfaces, affecting the accuracy of reflectance measurement for macronutrient sensing. Second, simultaneously sensing multiple soil macronutrients (i.e., nitrogen, phosphorus, and potassium) is challenging since they interfere with each other. Third, long-range and low-power communication is essential for practical deployment in vast farm fields. However, incorporating the communication capability into existing sensing hardware without incurring additional hardware and energy costs is challenging.

To address these challenges, we propose SoilCares, a generalized low-cost sensing solution (with the basic idea illustrated in Fig. 1) by leveraging the fusion of VNIR reflectance spectroscopy and RF technique to achieve in-situ soil moisture and macronutrient sensing. First, we design a novel denoising module to eliminate confounding factors (i.e., moisture and soil particle size) through a custom-designed membrane that removes the effect of particle size and a mapping function to eliminate the influence of moisture on soil macronutrient sensing. Second, we propose a multi-modal system that combines LED and RF sensing to further improve the accuracy of soil moisture sensing. Third, we develop an LED array-based system that can accurately sense the individual concentration levels of nitrogen (N), phosphorus (P), and potassium (K). Finally, we enable LoRa transmission by programming the clock generator of the computing unit in existing sensing hardware and outputting the generated clock signal through IO ports, mimicking the function of an ordinary transmitter. These approaches make SoilCares a complete end-to-end system ready for real-world deployment.

We implement the SoilCares prototype in a small form factor using cheap (\$63.5) and low-power (236  $\mu$ W) components. Extensive real-world experiments show that SoilCares achieves the root-mean-square error (RMSE) of 0.144, 0.147, and 0.125 for N, P, and K concentration measurements, respectively, and an extremely low error of 1% for moisture estimation. To the best of our knowledge, SoilCares is the first system demonstrating the capability of low-cost, accurate, in-situ soil moisture and macronutrient sensing. The main contributions of SoilCares are as follows:

- We propose to monitor soil moisture and soil macronutrients at the same time leveraging both RF and LED signals.
- We develop a denoising module to eliminate the effect of confounding factors. Specifically, we employ a novel membrane design to remove the effect of particle size and soil type. We also obtain the mapping function to remove the effect of soil moisture on macronutrient sensing.
- We propose a multi-modal model to combine VNIR and RF sensing for highly accurate soil moisture estimation and eventually remove the effect of moisture on macronutrient monitoring.
- We enable negligible-cost LoRa communication with delicate software design on existing sensing hardware, achieving a transmission range of 80 m and a coverage over 20,000  $m^2$ .
- We design and implement a compact, low-cost prototype. Extensive experiments in the lab and in the wild show that SoilCares achieves an average root-mean-square error of 0.138 for macronutrient sensing under varying conditions including different soil types/mixtures and various moisture levels. The proposed soil moisture sensing module achieves 1% mean absolute error, outperforming the state-of-the-art research [13, 20, 22] and expensive commodity moisture sensors on the market.

## 2 BACKGROUND

### 2.1 NIR-based Soil Macronutrient Sensing

The Beer-Lambert law [46] describes the attenuation of light intensity as it traverses through a substance, linking it to the material's constituent. This law is frequently utilized in chemical analysis to evaluate the concentration of chemical components capable of light absorption and scattering [48]. As the light of a specific spectrum passes through substances such as N, P, and K, it provokes the molecular bonds of each component to vibrate. Due to its unique molecular structure and bond, every chemical species generates a distinctive absorption spectrum, which can be used for element analysis [26, 64]. The absorbance of substance  $A$  can be derived in as below [65]:

$$A = \log\left(\frac{I_0}{I_r}\right) = \gamma \ell c, \quad (1)$$

where  $I_0$  is the emitted light intensity,  $I_r$  is the received light intensity after propagating through the optical length  $\ell$ ,  $\gamma$  is the molar attenuation coefficient, and  $c$  is the concentration of the attenuating species.

### 2.2 RF-based Soil Moisture Sensing

Volumetric water content (VWC)  $\theta_v$ , the volume of water per unit volume of soil, is a common metric for soil moisture measurement. Prior studies [34, 62, 67] have revealed the dependence of the *dielectric constant*  $\epsilon$  on VWC. The empirical formula to quantify the relationship between  $\epsilon$  and  $\theta_v$  [67] is depicted below:

$$\theta_v = 0.1138\sqrt{\epsilon} - 0.1758. \quad (2)$$

Based on Eq. 2, RF-based soil moisture sensing has been proposed [20]. By measuring the RF wave propagation speed  $c$  in the target soil, the dielectric constant can be obtained as  $\sqrt{\epsilon} = c_0/c$ , where  $c_0$  denotes the RF signal propagation speed in the air. So, soil

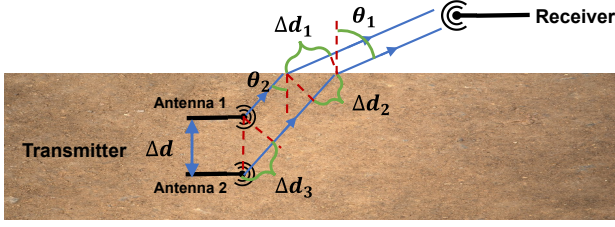


Figure 2: RF propagation path across soil and air

moisture can be estimated by measuring the RF signal propagation speed in the soil.

**TDoF-based soil moisture sensing.** Conventional RF solutions such as Time Domain Reflectometry (TDR) radar [38] estimate the RF wave speed based on accurate Time-of-Fight (ToF) measurements, which, however, require ultra-wide bandwidth, escalating the hardware cost. To enable affordable solutions, Time-Difference-of-Flight (TDoF) based approaches have been proposed [13, 20, 22].

As shown in Fig. 2, the same RF signal is transmitted by two antennas through soil and air and eventually received at the receiver. Since the distance between transmitter and receiver is much larger than twice the wavelength (e.g. 33 cm for LoRa at 915 MHz), the propagation paths can be considered as parallel [53]. The TDoF of the two paths can then be obtained as:

$$\Delta t = \frac{\Delta d_3 + \Delta d_2}{c} - \frac{\Delta d_1}{c_0} = \frac{\Delta d_3 + \Delta d_2 - \frac{\Delta d_1}{\sqrt{\epsilon}}}{c}. \quad (3)$$

According to Snell's law [73] and the geometric relationship, we have:

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{\lambda_1}{\lambda_2} = \frac{c}{c_0} = \sqrt{\epsilon}, \quad \frac{\sin \theta_1}{\Delta d_1} = \frac{\sin \theta_2}{\Delta d_2}, \quad (4)$$

where  $\lambda_1$  and  $\lambda_2$  are the RF wavelength in the air and in the soil, respectively.

With Eq. 4, the term  $\Delta d_2 - \Delta d_1/\sqrt{\epsilon}$  in Eq. 3 is canceled out, and the relationship between dielectric constant  $\epsilon$  and TDoF  $\Delta t$  can be obtained as:

$$\epsilon = \left( \frac{c_0 \Delta t}{\Delta d_3} \right)^2 = \left( \frac{c_0 \Delta t}{\Delta d \cos \theta_2} \right)^2. \quad (5)$$

In Eq. 5,  $\Delta d$  denotes the spacing between two antennas, which is predefined. The incident angle  $\theta_2$  does not need to be known and can be approximated with a constant value [13]. Therefore, the soil moisture  $\theta_v$  could be estimated from the TDoF  $\Delta t$  given Eq. 2 and 5. TDoF  $\Delta t$  can be accurately calculated by measuring the phase difference  $\Delta\phi$  of the received signals transmitted from the two antennas [20, 22]:

$$\Delta t = \frac{\Delta\phi}{2\pi f_c}, \quad (6)$$

where  $f_c$  denotes the carrier frequency of the RF signal.

### 3 SOILCARES DESIGN

In this section, we first describe the overview of SoilCares illustrated in Fig. 3, followed by the four design components. The first component involves techniques to eliminate the impact of soil moisture and particle size in macronutrient monitoring, allowing SoilCares

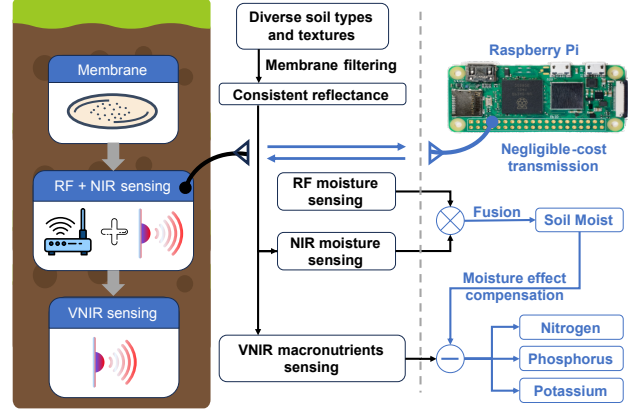


Figure 3: Overview of SoilCares: in-soil sensing node (left) and analyzer + negligible-cost controller (right)

to operate outside laboratory scenarios. Second, we present multi-modal soil moisture monitoring, combining RF and NIR sensing. With the multi-modal fusion, we enhance the accuracy and reliability of soil moisture sensing and further improve the sensing accuracy of macronutrients by eliminating the significant effect of moisture. Third, we address the monitoring of soil nitrogen, phosphorus, and potassium by designing a cost-effective LED-PD array. Last, negligible-cost LoRa transmission is realized on the existing sensing hardware, facilitating real-world deployments.

#### 3.1 Elimination of Confounding Factors for Soil Macronutrients Sensing

Recent studies have demonstrated the potential of macronutrient monitoring using VNIR optical systems, thanks to the rich soil spectral information contained within this spectrum [51, 55, 64]. Most of these studies, however, are predominantly confined to laboratory environments due to the requirement of pre-processing steps such as drying and grinding. These steps are crucial for mitigating the influence of soil moisture and particle size on the reflectance spectroscopy of soil samples [51, 55]. Soil moisture impacts the absorption spectra of soil elements, and different particle sizes lead to inconsistent reflectance, resulting in varying levels of accuracy [49–51, 59, 72].

We introduce a module to eliminate noise from soil moisture and particle size. To tackle the issue of varying soil particle sizes, we employ a membrane as depicted in Fig. 10(g). The membrane acts as an exchange platform, absorbing moisture and macronutrients until the equilibrium state is reached, i.e., the moisture and macronutrient levels are the same in the membrane and in the soil. Compared to directly measuring the moisture and macronutrient levels in the soil, the membrane presents a surface with more evenly distributed moisture and macronutrients, allowing for more accurate and stable measurements. Instead of selecting irregular soil particles as the surface for reflectance, this membrane absorbs macronutrients and water from the soil and provides a uniform surface for VNIR light reflectance measurements. Since the membrane is made of Poly-Vinylidene-Fluoride-co-Hexafluoropropylene (PVDF-co-HFP), it has minimal impact on Radio Frequency (RF) signals. We describe the membrane manufacturing details in Sec. 4.

To accommodate the noise caused by soil moisture, we break down the spectral absorbance  $A_i$  of the soil sample at the  $i^{th}$  wavelengths into three subcomponents as illustrated in Eq. (7):

$$A_i = A_{ti} + A_{mi} + A_{oi}, \quad (7)$$

where  $A_{ti}$  is the spectral absorbance of the target macronutrient,  $A_{mi}$  is the absorbance of soil moisture content, and  $A_{oi}$  is the absorbance of other elements. Given that the water component significantly influences the absorption spectrum at 1450 nm due to the O-H bond's vibration [7, 11, 41, 44], we establish a mapping function  $\phi(\cdot)$ . This function links the absorbance value at 1450 nm, represented as  $A_{m1450}$ , to each  $A_{mi}$  at various soil moisture levels.

By eliminating the mapped component  $\phi(A_{m1450})$  from the total absorbance  $A_i$  at the  $i^{th}$  wavelength, we obtain the remaining absorbance  $A_{ri}$  in Eq. 8:

$$A_{ri} = A_i - \phi_i(A_{m1450}) = A_{ti} + A_{oi}. \quad (8)$$

Subsequently, the concentration values  $y$  and the absorbance  $A_{ri}$  are projected into a latent space using Least Squares Support Vector Regression (LS-SVR) as:

$$Y = \sum_{j=1}^k \alpha_j f(A_{ri}) + b_2, \quad (9)$$

where  $Y$  represents the projected concentration values  $y$  in the latent space,  $f(\cdot)$  is the linear function that projects the value of  $A_{ri}$  into a new latent space,  $k$  is the total count of new latent variables based on the original number of wavelengths  $n$ ,  $\alpha_j$  represents the loading vector for the latent variable  $f(A_{ri})$ , and  $b_2$  is the bias of the new variable in the latent space. The initial absorbance of other elements, denoted as  $A_{oi}$ , exhibits inconsistent trends across different concentration levels, and its magnitude is substantially lower than the absorbance of the target macronutrient,  $A_{ti}$ . The impact of unrelated variables is further diminished through the projection process, resulting in a reduced amplitude. Consequently, we can disregard the value of  $A_{oi}$  in Eq. 7 and approximate the value of  $A_{ri}$  to  $A_{ti}$  in Eq. 9.

### 3.2 NIR-RF-based Soil Moisture Sensing

While both NIR and RF measurements can be used for soil moisture estimation, they are prone to errors caused by environmental factors. Specifically, NIR measurements can be affected by soil granularity and porosity [10, 64]. RF-based soil sensing is biased by uneven surfaces and obstacles such as stones [13, 22]. Interestingly, the factors affecting NIR and RF sensing are orthogonal, which facilitates fusing NIR and RF sensing for better performance.

The learning-based sensing techniques [32, 61, 68, 69, 71] have been widely studied. We explore the characteristics of both NIR and RF modalities and propose a multi-modal moisture sensing model. Specifically, we use the raw measurements, i.e., signal phase from RF receivers, and NIR reflectance from photodiode as input features to train the machine learning model. The obtained model is then used to predict soil moisture, which efficiently mitigates the bias from each individual modality and achieves higher accuracy than either of them.

For NIR-based moisture sensing, we use an LED transmitting 1450 nm light with a photodiode receiver since water molecules

**Table 1: Mean absolute error of moisture level estimation using different machine learning models**

Categories	ML Method	Soil 1	Soil 2	Soil 3	All
Linear	Ridge	1.37	1.33	2.83	2.11
	LS-SVR	1.21	1.33	2.49	1.80
Basic nonlinear	Polynomial	1.37	1.33	2.47	2.11
Proximity based	Nearest Neighbors	1.07	1.18	1.66	1.39
Probability based	Gaussian Process	3.66	2.48	4.10	3.76
Tree-structured	Decision Tree	0.66	0.95	1.36	1.26
	Random Forest	0.88	0.90	1.42	1.03
Neural Network	1 layer	2.89	1.80	4.09	2.35
	2 layers	1.90	1.77	3.78	2.36
Dimensionality Reduction	PCR	2.15	1.70	3.53	2.11

majorly absorb 1450 nm light waves [47]. NIR reflectance manifests a monotonic empirical relationship [41, 42] with the soil moisture level. For RF sensing, we use a LoRa transmitter with a single-input-out RF switch to enable two-antenna transmission. By switching channels within a chirp, we see a sudden phase jump due to the propagation path difference between two antennas, thus enabling TDoF sensing at the receiver side. The RF dielectric permittivity change due to different soil moisture levels is nonlinear based on Eq. 2. Consequently, the fusion task is modeled as a monotonic regression problem with two independent features.

To minimize the latency, we develop a fusion model based on small-size machine learning models instead of deep neural networks. Specifically, we compare ten machine learning models, including six linear and nonlinear models, three ensemble learning methods, and a dimensionality reduction method. The architecture of both 1-layer and 2-layer neural networks are determined by grid search on our collected dataset. The training dataset comprises around 600 soil samples, including three soil types across 5-50% soil moisture levels. We used 6-fold cross-validation in the training. The 2-layer neural network has two inputs, i.e., NIR reflectance and RF signal phase, ten neurons for each of the two hidden layers along with ReLU activation, and one output of the soil moisture estimation. The 1-layer neural network has the same inputs/output and 50 neurons in the hidden layer. Adam Solver is used for both neural networks, with a constant learning rate of 0.001.

Table 1 shows the mean absolute error of soil moisture level estimation across three soil types (see Sec. 5.2 for detailed comparison) and their combinations. The Decision Tree and Random Forest methods outperform others. Decision Tree achieves the highest accuracy on a single soil type but performs not as good as Random Forest considering all soil types. It indicates that Decision Tree is more prone to overfitting. Thus, we choose the Random Forest model for multi-modal fusion, and it achieves much higher accuracy (an error of 1.03%) than the high-end commodity soil sensor [3] (an error of 2.54%). Sec. 5.3.1 will present more evaluation on other practical considerations.

### 3.3 Spectral Channel Selection for Soil Macronutrient Sensing

Having addressed the interference of soil moisture and particle size on soil macronutrient sensing, the subsequent challenge we face

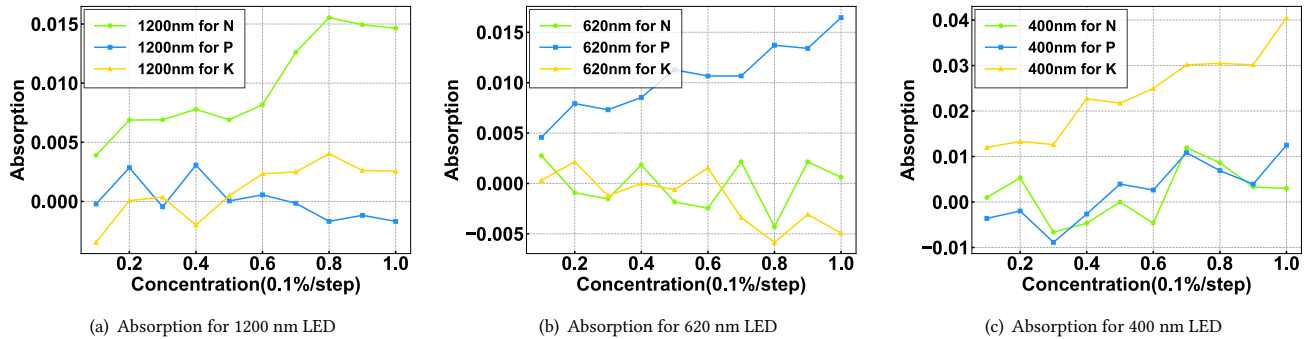


Figure 4: Absorption of Selected wavelength among different macronutrient

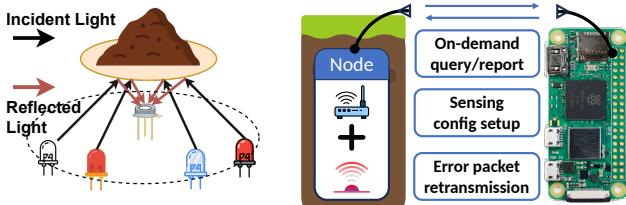


Figure 5: Placement of LEDs and photodiodes

is to simultaneously determine the concentration levels of multiple macronutrients [5, 50, 59]. In SoilCares, we design a low-cost LED array to address the challenge. Specifically, we employ three spectral regions to detect N, P, and K concentrations separately. We utilize seven LEDs at 850, 950, 1150, 1200, 1300, 1550, and 1650 nm for N monitoring. We employ four LEDs at 400, 460, 470, and 525 nm for K monitoring. We utilize four LEDs at 620, 640, 660, and 720 nm for P monitoring. The selection of these spectral regions is strategically made to optimize the light attenuation by the target nutrient while limiting the impact of light attenuation by other nutrients [5, 28, 30, 49]. Fig. 4 illustrates the absorption properties of the target macronutrients N, P, and K at the selected wavelengths. We observe that the absorption of the targeted element is at least three times the absorption of other macronutrients and significantly more distinctive at the chosen wavelength, indicating the effectiveness of the chosen spectrum on soil macronutrient sensing. Fig. 5 shows the placement of the LEDs, photodiode, membrane, and soil, where the LEDs are placed in a circle with the photodiode positioned at the center.

### 3.4 Negligible-cost LoRa Transmission

The hardware of SoilCares consists of two main subsystems: RF-VNIR sensing node in soil and remote controller module. We built the controller module with a low-cost RF receiver RTL-SDR (\$16) and a Raspberry Pi Zero (\$5). To trigger the operation of the sensing node in the soil, an extra transmitter is required, which would increase the hardware cost. Instead, we propose to exploit existing hardware to enable negligible-cost LoRa signal transmission

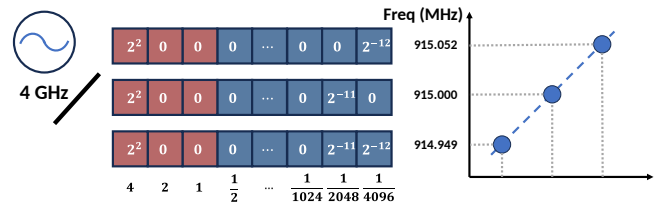


Figure 7: Illustration of frequency resolution problem

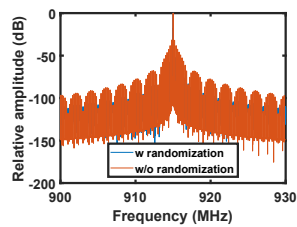
capability with delicate software programming. The key idea is to leverage existing RF-band signals in the current hardware for transmission. Fortunately, the Raspberry Pi has a programmable clock generator to drive its CPU up to 1 GHz, and the generated clock signal can be output to general-purpose input/output (GPIO) ports [2].

**Fine-grained frequency control.** Although sweeping the frequency to create chirp signals is straightforward, precise frequency control becomes the major challenge due to hardware limitations. The clock generator in Raspberry Pi is based on a 4GHz clock [54] and generates the target frequency through dividing the 4-GHz clock by a positive real number. The number is composed of a 12-bit integer part and a 12-bit fraction part [2]. Therefore, we can not obtain continuous target frequency due to the limited resolution. As shown in Fig. 7, for a 915 MHz channel with 125 kHz bandwidth, we can get only three discrete frequency points in the range by tuning the least significant bits, which is far from enough to generate a chirp signal.

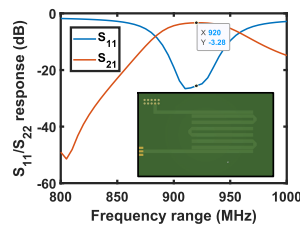
To enable fine-grained frequency sweeping, we leverage a basic concept of signal processing: the instantaneous frequency is the time-domain derivative of the instantaneous phase:

$$f(t) = \frac{1}{2\pi} \frac{d\phi(t)}{dt}. \quad (10)$$

We further derive that the instantaneous frequency can be tuned by adjusting the “speed” of phase rotation, *i.e.*, the accumulated phase rotation in a unit time. For example, we may quickly switch between the two discrete frequency points, 915 MHz and 915.052 MHz, half by half in a unit time. In this case, the accumulated phase rotation would equal that from a single-tone signal at the middle



**Figure 8: Fine-grained frequency tuning**



**Figure 9: PCB filter and frequency response**

frequency, 915.026 MHz. By adjusting the ratio of the two discrete frequencies, we obtain the target “fractional” frequency in the LoRa frequency band (902-928 MHz), as shown in Fig. 8. To eliminate the spurs due to periodic switching, we randomize the switching order while keeping the ratio of the two frequencies not changed. As a result, the spurs are eliminated, and we obtain a clean single-tone wave.

**Frequency harmonics reduction at GPIO.** Fine-grained frequency control is achieved with randomized frequency switching. However, the GPIO port of the Raspberry Pi is digital, i.e., it can only output “0” or “1”. In consequence, the output signal from the GPIO port is a square wave, which is the combination of the target 915 MHz wave and harmonics at 1830 MHz, 2745 MHz, etc. The unwanted harmonics at 1830 MHz and 2745 MHz fall in the 3G and 4G bands, while unlicensed transmission can cause legal issues. A 915 MHz bandpass filter is thus necessary after the GPIO output. Commodity 915 MHz bandpass filter is usually tens to hundreds of dollars [19], which is higher than the total cost of our system. To meet the requirement of a negligible-cost transmitter, we design a printed hairpin filter (\$0.5, see Fig. 9) [12] without any RF components to enable a low cost, which costs less than 1% of a commodity counterpart [19] and incurs only 1.3 dB more loss. It is a 2-layer PCB filter (the size of the PCB is 98 mm by 60 mm) with the base material of FR4. The thickness of the PCB board is 1.6 mm, and the outer copper weight is 1 oz.

## 4 IMPLEMENTATION

This section presents the implementation details for the proposed SoilCares prototype. The total cost of SoilCares is \$63.5, whereas the high-end commodity soil sensor costs more than \$500 [3], not to mention that SoilCares is capable of measuring the individual concentration of macronutrients while the commodity counterparts only provide the overall estimation.

**Membrane.** We manufacture the membrane using Poly-Vinylidene-Fluoride-co-Hexafluoropropylene (PVDF-co-HFP) with a number-average molecular weight ( $M_n$ ) of 130 kg/mol and a weight-average molecular weight ( $M_w$ ) of 400 kg/mol. Our membrane design has a thickness of about 100 microns, and the appearance of the membrane is shown in Fig. 10(g). The swelling ratio, which measures the change in the membrane’s size between wet and dry states, is an indicator of physical stability [25]. Our membrane design exhibits a 12.5% swelling ratio at 25 degrees Celsius, signifying that the membrane retains its shape well when exposed to moisture [25, 40].

Additionally, the impact of minor size variations can be effectively mitigated by considering the mapping relationship between the swelling ratio and soil moisture. The dimensions of the membrane can be adjusted based on needs. We use a circular membrane with a diameter of 41 mm for the VNIR-based soil sensing and a rectangular shape with 100 mm  $\times$  100 mm for soil moisture sensing. Both membranes are identical in composition, and the cost is negligible.

The membrane used in the experiment must have a smooth, flat surface and uniform thickness to guarantee that the VNIR signals, regardless of their optical paths, are uniformly reflected by the membrane. Variations in thickness can introduce additional noise into the system, negatively impacting the accuracy of predictions related to soil moisture and macronutrients. Furthermore, physical damage to the membrane, whether during the manufacturing or use process, will impair the overall performance of the system.

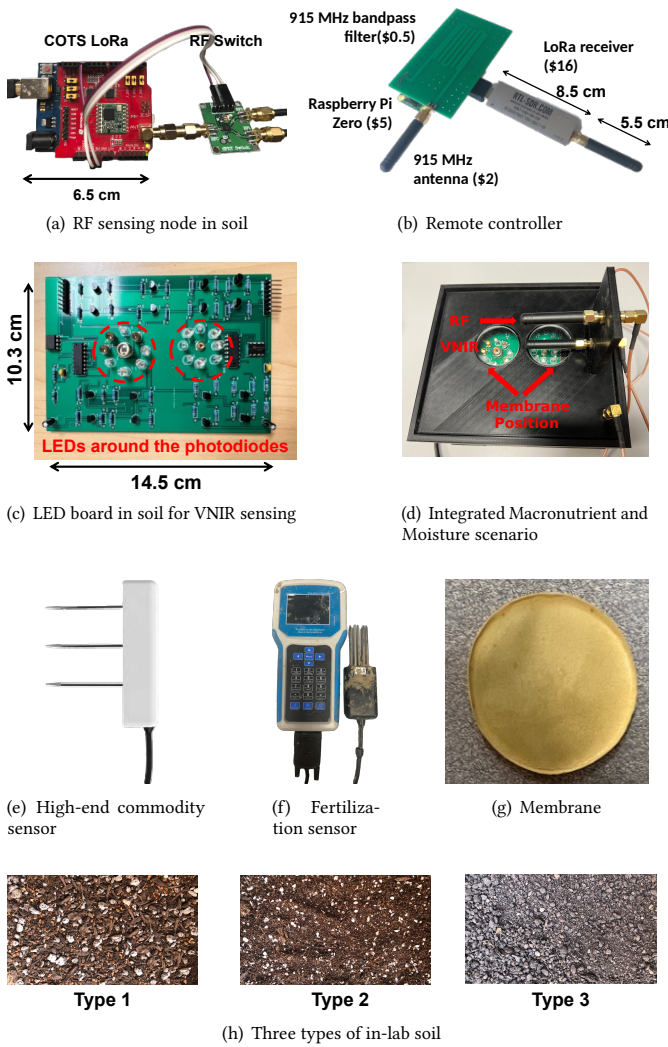
The durability of the membrane is influenced by its surrounding environment. In a farm environment, a membrane made from PVDF demonstrates remarkable resilience against destructive substances like acids [60]. While strong bases and extremely high temperatures may impair its function, such conditions are rare in agricultural soil. The membrane can remain effective for more than five years in the lab and 18 months in the wild. Note that during the growing season, the membrane can be taken out and replaced if it suffers physical damage in the soil.

**VNIR sensing module.** We utilize an Arduino UNO microcontroller to drive the VNIR transceiver (\$6) shown in Fig. 10(c). The PCB board consists of an amplification circuit and a multiplexer control circuit. The amplification circuit, employing an LM358P amplifier, amplifies the photodiode’s voltage changes by a factor of ten. The multiplexer, a CD74HC4051E type 8:1 device, can control the LED’s alternating on/off pattern. Fig. 10(c) illustrates the configuration of our PCB board design, where two LED arrays are placed in two circular patterns around the Vis- and NIR- photodiodes. Fig. 10(d) demonstrates the assembly of our VNIR-based sensing module. The light produced by the LED is reflected off the membrane and detected by the photodiode.

**RF sensing node.** Fig. 10(a) shows the RF sensing node, including a commodity LoRa transceiver board based on Arduino UNO board with Dragino LoRa shield [21] (\$24), and an RF switch HMC849 [17] (\$4). LoRa transmission is enabled using an open-source Arduino library [9]. We modify the timer codes of the library to enable the antenna switching function during transmission to support TDoF sensing with two transmitter antennas.

**Remote controller.** Fig. 10(b) shows the remote controller with a Raspberry Pi zero (\$5) as the backend. We use amateur radio, RTL-SDR (\$16), with an open-source library [56] to receive the raw LoRa signal for the sensing node. The collected raw data is processed with Python to obtain the moisture and macronutrient concentration levels. To control the sensing node remotely, we have the transmitter part consisting of another 915 MHz antenna, the customized filter, and our codes running in the Raspberry Pi. The customized filter (\$0.5) is connected to the Raspberry Pi to avoid illegal out-of-band signal leakage.

**Power consumption.** To support long-term deployment without replacing the battery, the in-soil sensing node is set to sleep mode



**Figure 10: Implementation of SoilCares prototype.**

whose power consumption is only  $35 \mu\text{W}$ . In the active mode, the Arduino UNO board consumes around  $26 \text{ mW}$  after removing unused components on the board. The VNIR board powers one LED and one photodiode at a time to measure the reflectance of the corresponding wavelength. Each measurement takes one second, and all measurements take 16 seconds with an average power consumption of  $11 \text{ mW}$ . After VNIR measurements, the LoRa shield is powered on for one second to transmit the VNIR reflectance values, whose power consumption is  $100 \text{ mW}$ . By applying the default setting of the commodity sensor (measured once per hour), the average power consumption becomes  $236 \mu\text{W}$ , i.e., a common  $5\text{V}$   $2250 \text{ mAh}$  Lithium Ion Battery can provide a five-year battery life for SoilCares.

## 5 EVALUATION

In this section, we initially introduce the three evaluation metrics that will be utilized. This is followed by a detailed presentation of

both in-lab and on-site experiments, which focus on monitoring soil moisture, nitrogen (N), phosphorus (P), and potassium (K) under a diverse range of experimental settings.

### 5.1 Evaluation Metric

For the soil moisture experiments in Sec. 5.3.1, we use the Mean Absolute Error (MAE) to facilitate the comparison with the state-of-the-arts [13, 22]. For the experiments on soil macronutrients, we leverage the coefficient of determination ( $R^2$ ) as the primary criterion to enable comparison with existing studies [49, 50, 59]. The formula of  $R^2$  is defined as Eq. 11:

$$R^2 = 1 - \frac{\sum_{k=1}^m (\hat{y}_k - f(x_k))^2}{\sum_{k=1}^m (\hat{y}_k - \bar{y})^2} \quad (11)$$

where  $\hat{y}_k$  is the ground truth of the  $k^{\text{th}}$  macronutrient concentration and  $m$  is the total number of experiment samples. The function  $f$  is the adapted model function that transforms the input raw data  $x_k$  to the predicted value of macronutrient concentration  $f(x_k)$ ,  $x_k$  is derived from the intensity of reflected light received at the photodiode, and  $\bar{y}$  is the mean value of the macronutrient concentration ground truth. The coefficient of determination  $R^2 = 1$  if the predicted value  $f(x_k)$  exactly matches the observed value  $\hat{y}_i$ . A negative value of  $R^2$  indicates that the adapted model is worse than simply calculating the observed values' mean.

We also leverage the mean square error of cross-validation (RMSECV) and root mean square error of prediction (RMSEP) as the evaluation metrics. These two criteria are considered as indicators of error in model predictions. The formulas of RMSECV and RMSEP are shown in Eq. 12:

$$\begin{aligned} \text{RMSECV} &= \sqrt{\frac{\sum_{l=1}^{L_c} (\hat{y}_l - y_{cl})^2}{L_c}} \\ \text{RMSEP} &= \sqrt{\frac{\sum_{l=1}^{L_p} (\hat{y}_l - y_{pl})^2}{L_p}} \end{aligned} \quad (12)$$

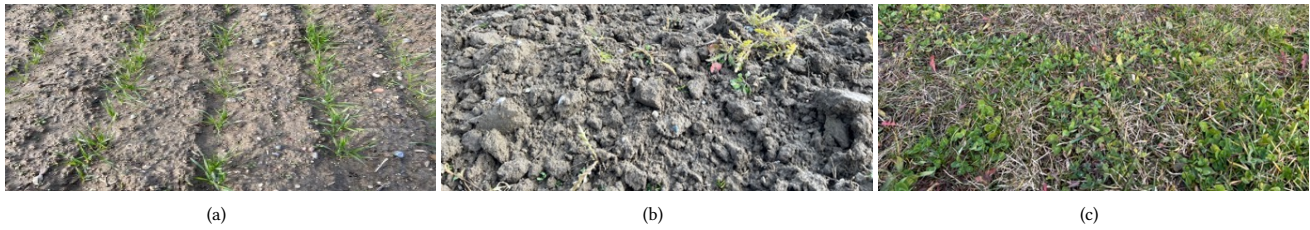
where  $l$  is the sample number,  $L_c$  and  $L_p$  are the total number of samples in the validation and prediction groups,  $y_{cl}$  and  $y_{pl}$  are the predicted concentration values of the target macronutrient from the validation group and prediction group.

### 5.2 Experiment Setting

**Soil samples for in-lab evaluation.** To validate the versatility of SoilCares across diverse soil types, we use three types of soil with distinct mixtures, shown in Fig. 10(h).

### 5.3 In-lab Evaluation

To evaluate the sensing performance across different moisture levels and soil types, we use transparent acrylic boxes ( $10 \text{ cm} \times 10 \text{ cm} \times 10 \text{ cm}$ ) to hold the above three types of soil samples with different moisture levels, spanning from the natural moisture level to saturation level (50%). For uniform soil moisture distribution, we cover the soil boxes with plastic wraps with uniform holes and let the specific amount of water drip through the holes uniformly. The high-end commodity soil sensor is used for comparison, shown in Fig. 10(e).



**Figure 11: On-site scenarios. (a): with new-planted wheat. (b): turned ground. (c): with grass covered**

For macronutrient evaluation, in alignment with the procedures outlined by [43, 50], we use raw samples with a weight of 10 g, mixed with the target macronutrient and distilled water (an additional weight of 4 g), which are then transferred to a circle transparent acrylic board (38.5 mm in diameter) and covered by a cylindrical acrylic box (20 mm in depth) for experimentation.

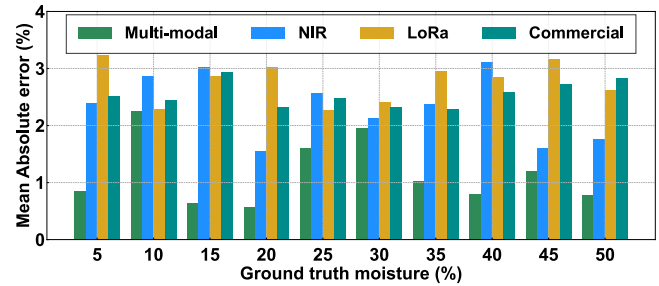
We make 471 in-lab soil samples, encompassing all soil types and various concentrations. For the in-lab Soil NPK monitoring, 360 soil samples are prepared with macronutrients ranging from 0.1% to 1%. We also gather 81 samples, which are mixed with three concentration levels for each macronutrient: 0.2% for Low, 0.5% for Medium, and 0.8% for High. We establish three distinct levels for soil moisture: 10% for low moisture level, 20% for medium moisture level, and 28% for high moisture level to perform the concurrent soil N/P/K and moisture monitoring.

**Soil and fertilization preparation for on-site study.** To test SoilCares in practical scenarios, we conduct 108 on-site experiments at local farm fields. We select three field types for our study, as depicted in Fig. 11. To achieve the goal of assessing the effects of excessive soil fertilization, we adjust the macronutrient levels in the test fields to approximately 1.5g/kg for nitrogen (N) [1, 52, 58], 1.5g/kg for phosphorus (P) [6, 14, 58], and 2g/kg for potassium (K) [52], while maintaining soil moisture levels between 20% and 45% throughout the duration of the experiment. The ground truth is measured after fertilization and full water infiltration [63] using the device in Fig. 10(e) and Fig. 10(f).

### 5.3.1 Performance of soil moisture and NPK monitoring.

We first evaluate SoilCares when only one soil substance (soil moisture, N, P, or K) is present in soil samples.

**Soil moisture estimation.** We conduct experiments to demonstrate the efficiency of our soil moisture monitoring module across a wide range of soil moisture levels (5-50% [13]) in agriculture. We repeat measurements 30 times for each soil sample and compute the mean output. We compare SoilCares with three baselines: the state-of-the-art LoRa-based method [13], a NIR-based method [42], and a high-end commodity device [3]. Fig. 12 shows the overall MAE across different moisture levels are 2.35%, 2.76%, 1.03%, and 2.54% for LoRa, NIR, SoilCares (LoRa + NIR), and the commodity device, respectively. SoilCares consistently achieves the highest accuracy for each moisture level and outperforms the high-end commodity sensor.



**Figure 12: Comparison of soil moisture error among different moisture estimation methods**

**Soil NPK estimation.** We then assess the ability of SoilCares to predict soil N, P, and K concentrations across three soil types. Table 2 summarizes the performance of SoilCares and three baselines [49, 50, 59]. Overall, SoilCares achieves a coefficient of determination of 0.811 for N, 0.803 for P, and 0.837 for K in soil sensing, comparable to existing methods that require soil pre-processing to enhance the prediction performance or expensive spectrometers. SoilCares only employs a COTS LoRa device, a few low-cost LEDs, and photodiodes to perform the function of sensing and data communication, which are more affordable and portable than the spectrometer-based method. Contrary to the methods utilized in [49, 50], our system avoids soil pre-processing (e.g., drying and grinding) and screening of collected soil samples. Finally, we observe that the existing LED-based method [49] achieves higher  $R^2$  than SoilCares. However, the soil macronutrient concentration in their study (10–50%) substantially exceeds the concentration range specified in soil fertilization guideline [57], and there was no evidence that the system was evaluated under the practical macronutrient concentration level.

**Impact of membrane across different soil types.** For soil moisture estimation, we observe an MAE of 1.03%, reflecting a notable error reduction of 51.8% compared to the baseline outcome of 2.14% without membrane. This improvement is mainly attributed to the uniform surface provided by the membrane, facilitating consistent NIR reflection measurements across different soil types. For macronutrient sensing, Table 3 compares soil NPK estimation accuracy with and without the membrane. SoilCares achieves an  $R^2 = 0.811$  for N,  $R^2 = 0.802$  for P, and  $R^2 = 0.837$  for K, significantly outperforming the performance without the membrane. To assess the performance of SoilCares across diverse soil types, we



**Table 2: Experimental settings: we evaluate SoilCares performance and comparing with three related works**

Work	Device	# of samples	Preprocessing	Concentration for N	Concentration for P	Concentration for K	Prediction Method	$R^2$ for N	$R^2$ for P	$R^2$ for K
[50]	Spectrometer	300	Required	0-3%	0-5%	0-0.3%	PCR / PLSR	0.87	0.99	0.90
[59]	Spectrometer	280	Required	0.0032% - 0.0208%	0.0025% - 0.0343%	0.0041% - 0.03454%	PLSR / LS-SVR	0.81	0.78	0.80
[49]	LEDS	150	Required	10% - 50%	10% - 50%	10% - 50%	Linear / Unique Curve	0.91	0.99	0.99
SoilCares	LEDS	360	Not Required	0-1%	0-1%	0-1%	PCR/ PLSR / LS-SVR	0.81	0.80	0.84

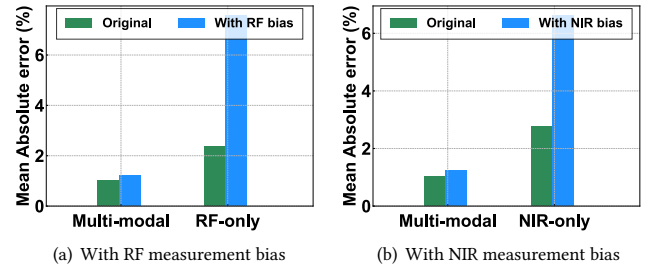
**Table 3: Impact of membrane and diverse soil types on soil macronutrient monitoring**

Components	Calibration Method	$R^2$	RMSECV	RMSEP
N w/ membrane	LS-SVR	0.811	0.132	0.135
N w/o membrane	PCR	0.439	0.210	0.266
P w/ membrane	PCR	0.802	0.128	0.130
P w/o membrane	PCR	0.426	0.201	0.239
K w/ membrane	LS-SVR	0.837	0.095	0.103
K w/o membrane	PCR	0.436	0.216	0.278

employ two evaluation metrics: RMSECV and RMSEP. RMSEP results are based on leave-one-soil-type-out cross-validation. We train our model on two types of soil samples and use the data collected from the third type of soil as the test set to compute RMSEP. This process is repeated three times to cover all the possible divisions of train and test sets. Meanwhile, RMSECV results are based on cross-validation, which mixes all soil samples of different types and randomly divides the train and test samples at a fixed ratio of 4:1 in our experiment. We observe that the difference between RMSEP and RMSECV is less than 0.008, showing the applicability of our method to various soil types. The results demonstrate the effectiveness of the membrane design in accommodating diverse soil types and textures.

**Combating single modality bias in soil estimation.** A single modality can be easily affected by environmental factors, such as soil granularity for NIR sensing and uneven surface for RF sensing. Our multi-modal method is proposed to address this issue. In Fig. 13(a), we artificially add a bias of 0.55 rad to each RF measurement. The moisture estimation error increases are 5.22% for the RF-only method and 0.20% for our multi-modal method. Similarly, we add a bias of 0.01 to each NIR measurement in Fig. 13(b). The moisture error increases are 3.96% and 0.21% for NIR-only and multi-modal methods, respectively.

**Impact of regression methodologies for macronutrient estimation.** Fig. 14 illustrates the Cumulative Distribution Function (CDF) of absolute error across three regression methodologies: Principal Component Regression (PCR), Partial Least Squares Regression (PLSR), and Least Squares Support Vector Regression (LS-SVR). The LS-SVR method demonstrates superior performance for soil N and K estimation with the membrane due to its regression process with kernel-based grid search. Conversely, in scenarios where the light absorption of the target element is not particularly prominent,

**Figure 13: Multi-modal resistance to RF/NIR bias**

PLSR and PCR produce better outcomes. A significant advantage of PCR and PLSR is their ability to discard dimensions with lower variance during the regression process. This characteristic helps mitigate the influence of noise [35], leading to more accurate predictions.

**5.3.2 Performance of concurrent soil NPK and moisture monitoring.** In this experiment, we assess the performance of our model, referred to as SoilCares, in scenarios where multiple confounding factors—including soil moisture, nitrogen (N), phosphorus (P), and potassium (K)—are concurrently present in soil samples. We partitioned the results into three groups based on their respective moisture levels: Low, Medium, and High, as illustrated in Sec. 5.2. Fig. 15 presents the SoilCares’s performance in relation to these four factors. Under medium soil moisture level, a sum absolute error of 2.36% for N, P, and K is achieved, which outperforms the sum error achieved under low moisture level (3.30%) and high moisture level (2.99%). This superior accuracy is mainly attributed to the lower error in moisture predictions around the medium level, which, after offsetting the effect of soil moisture, results in more accurate predictions for macronutrients.

The RMSE values under various moisture levels differ for each macronutrient element. We observed an RMSE of 0.139 for N, 0.150 for P, and 0.119 for K under the low moisture level. In the case of medium moisture level, the RMSE values are 0.101 for N, 0.109 for P, and 0.079 for K. Under the high moisture level, the RMSE values are 0.132 for N, 0.129 for P, and 0.137 for K. These results show that the RMSE of K varies significantly across the three settings. Apart from the factors related to moisture, we also observed that in the selected spectrum region for K monitoring, the chosen LEDs exhibit high sensitivity on N. This implies that the concentration level of N can influence the prediction of the concentration level of K to some extent [75]. In contrast, the spectrum regions selected

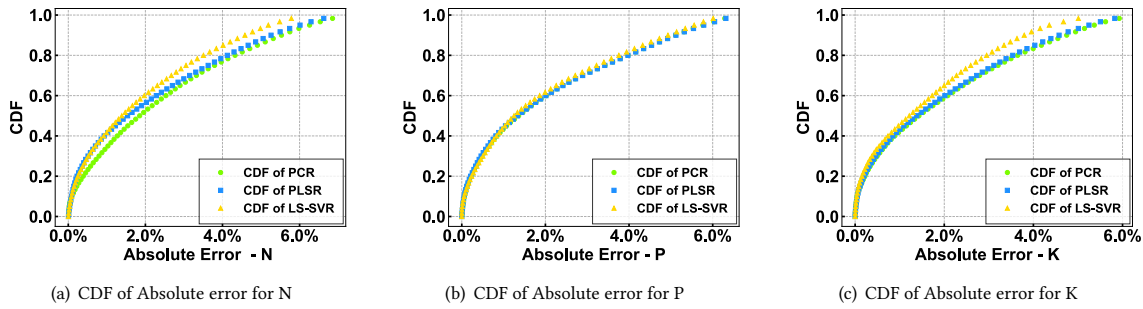


Figure 14: The CDFs of absolute error across three regression methodologies: PCR, PLSR, and LS-SVR.

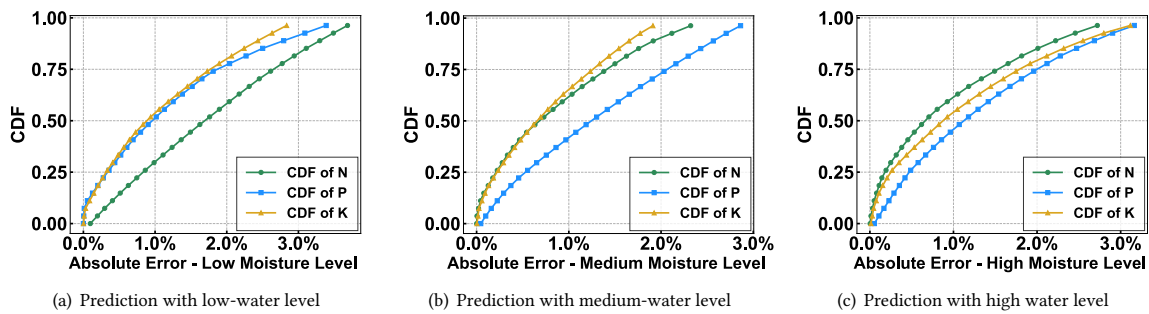


Figure 15: Prediction for three macronutrients simultaneously

for N and P do not exhibit sensitivity to other elements. Overall, our system achieves an average RMSE of 0.107 across three distinct soil moisture levels, each involving three macronutrients, indicating that SoilCares performs well in concurrently monitoring soil moisture and NPK levels.

## 5.4 On-site Study

**Impact of device depth in the soil.** We conducted on-site experiments at three depths: 0–10 cm, 10–20 cm, and 20–40 cm, as shown in Fig. 18(b). Most existing works [13, 20, 22] evaluated the performance of their systems at a depth up to 30 cm. The maximum testing depth (i.e., 40 cm) in this paper is determined based on the biological properties of plants. The roots of most plants are beneath the ground for about 20–30 cm, such as maize and celery [23, 45, 66]. We extend the sensing depth to 40 cm to cover more agricultural applications. Fig. 16(a) shows the RMSE of soil NPK at the three depths. We observe that the shallowest depth, 0–10 cm, yields the most accurate estimates, with an average RMSE of 0.112. As the depth increases, our system continues to deliver reliable results, with the RMSE always smaller than 0.182. Specifically, the predictions for N do not vary much across various depths, exhibiting an average RMSE of 0.145. For P and K, the average RMSE values are 0.147 and 0.129, respectively. These findings affirm that SoilCares

can work effectively at various depths to meet the objectives of diverse applications in real-world scenarios.

**Impact of terrains.** We conducted on-site experiments to assess the effectiveness of SoilCares on various terrains, as depicted in Fig. 11. Our test sites included farmland with newly planted wheat, newly turned soil land, and land with wild grass. Fig. 16(b) shows that SoilCares is comparably effective across all three farmlands. Across the three terrains and macronutrients, the overall performance stands at a mean RMSE of 0.139. This result demonstrates that SoilCares is adaptable and functional across various soil types and top coverings.

**Impact of weather conditions.** To verify the consistency of our system's performance under varying weather conditions, particularly under extreme weather (e.g., snow), we conduct experiments for a total of 60 hours. We start the experiment 12 hours after soil fertilization and water infiltration. After the inorganic fertilizers are applied, the macronutrient levels change rapidly in the next 72 hours and need to be closely monitored to avoid negative impacts on plants and the potential risk of contaminating groundwater. After 72 hours, the macronutrient levels change much more slowly. Fig. 16(c) illustrates the changes in absolute error and the temperature fluctuations. Notably, there are two significant increases in error during the 36–48th hours and 72nd hour intervals, influenced by abrupt drops in soil temperature. In the 36–48th hours, the soil temperature dropped below 0 °C. There was snowfall in the 72nd hour. These weather changes caused the moisture in the soil and the

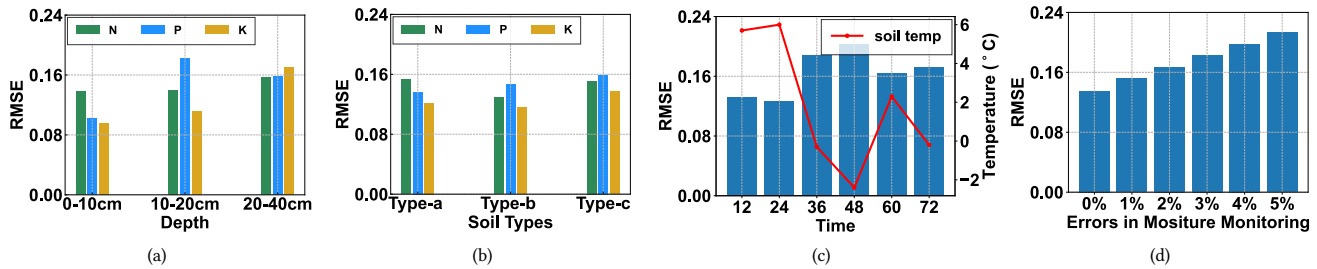


Figure 16: On-Site Soil Macronutrients level estimation across (a) different depths, (b) soil types, (c) soil temperature in time series (d) effect of moisture prediction error

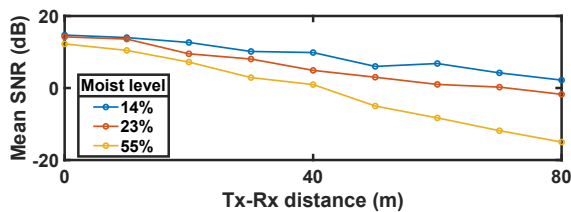


Figure 17: Reported SNR when correctly decoded

membrane to freeze, resulting in an RMSE exceeding 0.20. Despite these challenges, our system still produces reliable predictions with an RMSE lower than 0.165 for all other cases when the soil temperature is higher than 0 °C, the weather condition for SoilCares to provide reliable monitoring results.

**Impact of soil moisture estimation on soil NPK estimation.** Since the accuracy of soil moisture prediction is closely linked to the accuracy of macronutrient estimation, we investigate the impact of soil moisture estimation errors on soil macronutrient prediction. In this experiment, we focus on soil N prediction. We manually add noise to our soil moisture module to increase the errors in moisture sensing. Fig. 16(d) shows the RMSE of soil N prediction under various moisture estimation errors. The prediction error significantly escalates as moisture estimation error increases. Reducing moisture estimation error from 5% to 1% can significantly increase the prediction accuracy of nitrogen. This indicates that accurate soil moisture estimation plays a critical role in the performance of macronutrient prediction. The achieved soil moisture estimation accuracy in our system is adequate for reliably monitoring soil macronutrients [49, 50, 59]

**Performance of negligible-cost LoRa transmission.** We also evaluate the performance of the proposed negligible-cost LoRa transmission of the remote controller in the on-site study. The sensing node is buried in the soil, with the membrane and RF antennas positioned outside the box to ensure direct contact with the soil, while the circuits are placed inside the box. The sensing node is buried 60 cm below the soil surface at different locations under different soil moisture levels (14%, 23%, and 55% by adding water). Then, we move the controller away from the sensing node at a step size of 10 m from 0 m (right above the sensing node) to 80 m, as shown in Fig. 18(a). To evaluate the performance of employing

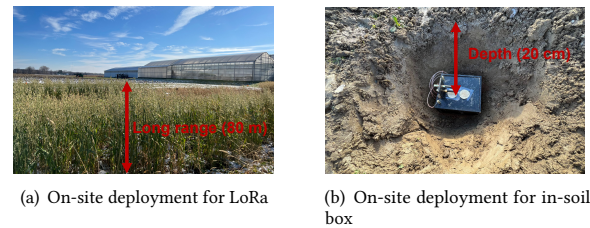


Figure 18: On-site deployment and communication

the controller for LoRa signal transmission, we adopt the default parameter setting for LoRa transmission, i.e., a central frequency of 915 MHz with a channel bandwidth of 125 kHz, a spreading factor of 12, and a 4/5 coding rate. The controller is programmed to send 100 LoRa packets. At the sensing node side, the commodity LoRa node acts as a receiver and decodes the received LoRa packets. Since the software [9] used in the sensing node only outputs when a packet is correctly decoded, we cannot obtain the bit error rate. Therefore, we plot the reported signal-to-noise-ratio (SNR) of those packets correctly decoded, as shown in Fig. 17. When the remote controller is right above the under-soil sensing node, the mean SNR across different moisture levels is all above 12 dB. Even in the most challenging case under a moisture level of 55%, we can still achieve successful packet reception at a distance of 80 m, where the mean SNR drops to -15 dB. Based on this result, the proposed negligible-cost LoRa transmission can achieve a transmission range larger than 80 m in real-world farmland, covering an area over 20,000 m<sup>2</sup>.

## 6 RELATED WORK

### 6.1 Light-based soil macronutrient sensing

Reflectance spectroscopy has found extensive usage in soil analysis for a large range of elements, such as nitrogen [5, 28], phosphorus [18, 49], potassium [16, 30, 43], and other critical elements like organic carbon [31]. The majority of accurate soil element measurements are conducted under laboratory settings using expensive and bulky spectrometers, which provide high-quality spectral resolution for analyses [29, 50, 59]. These lab-oriented studies often involve pre-processing steps such as drying and grinding the soil to mitigate the effects of confounding factors like soil moisture and particle size. This helps ensure higher measurement accuracy and

applicability to different soil types or textures [50, 59]. However, the high cost of spectrometers and the need of pre-processing severely limit the wide adoption of reflectance spectroscopy for everyday use. Recent studies have tried to circumvent these issues by adopting a combination of LEDs and photodiodes as an alternative to spectrometers [5, 49, 75]. Nevertheless, these works have not fully addressed the problem of generalizing across different soil types and still require tedious pre-processing. Moreover, they do not offer simultaneous prediction of multiple elements.

## 6.2 RF-based soil moisture sensing

RF-based soil moisture sensing is an emerging field. To replace the expensive ultra-wide-band ground penetrating radar (GPR) [4, 37], RF-based soil moisture sensing has been proposed to achieve low-cost sensing. A variety of RF signals have been exploited for soil moisture sensing including WiFi [20], RFID [70], LoRa [13] and LTE [22]. The WiFi-based solution proposes to use WiFi channel state information (CSI) to realize both moisture and salinity sensing. The RFID-based approach leverages RFID signal attenuation for moisture sensing. It is limited to container cases since the RFID tag needs to be attached to the outer surface of the container. In comparison, our proposed system can be applied to much broader application scenarios and achieves a higher accuracy (1% error vs. 3% error). The LoRa- and LTE-based solutions mainly focus on extending the range of soil moisture sensing and they are not capable of sensing macronutrients. A recent work [36] achieves a low moisture estimation error (1.1%) leveraging customized RF signals with a large GHz-bandwidth and expensive high-end software-defined radio platform (\$8000). In comparison, our system achieves a similar accuracy of 1% with low-cost hardware.

## 7 DISCUSSION

**On-site soil macronutrients and moisture sensing.** The primary distinction between laboratory and field experiments lies in the additional variables present in the wild that are not encountered in the lab, such as soil temperature and surface vegetation. It is essential to maintain the temperature above 0 degrees Celsius to ensure reliable sensing. Lower soil temperatures can alter the state of water, complicating the measurement of macronutrients and degrading the sensing performance. Moreover, stones and sharp objects could damage the membrane's surface, affecting the VNIR reflection. Furthermore, improper setup of the system, such as having large air gaps above the membrane, can also negatively impact system performance.

**Over 50% soil moisture.** Soil moisture levels in agriculture usually fall in the range 5% to 45% and the proposed system works well in this range. We also notice that in extreme cases (e.g., rice paddy), the soil moisture can be higher than 50%. In this case, the principles of the proposed system still hold. However, the high moisture can attenuate RF signals significantly, reducing the sensing range.

## 8 CONCLUSION

In this paper, we present SoilCares, a system capable of simultaneously measuring soil moisture and soil macronutrients using

RF-VNIR sensing. To accurately sense soil moisture and macronutrients across diverse soil types and textures, we design a novel membrane to provide a uniform reflectance surface for Vis-NIR sensing. Meanwhile, by leveraging the COTS LoRa hardware and low-cost LEDs and photodiodes, we propose a multi-modal model combining reflectance spectroscopy and RF sensing for soil moisture and macronutrient sensing. Extensive experiments show that SoilCares can achieve a root-mean-square error of 0.138 on soil macronutrient monitoring and 1% mean absolute error on soil moisture monitoring in complex real-world scenarios.

## REFERENCES

- [1] Francisco Albornoz. 2016. Crop responses to nitrogen overfertilization: A review. *Scientia horticulturae* 205 (2016), 79–83.
- [2] Arm. 2012. BCM2835 ARM Peripherals. <https://datasheets.raspberrypi.com/bcm2835/bcm2835-peripherals.pdf>
- [3] Aroya 2023. SUBSTRATE SENSOR. <https://aroya.io/en/aroya-sensors/products/substrate-sensor>
- [4] Ebrahim Babaian, Morteza Sadeghi, Scott B Jones, Carsten Montzka, Harry Vereecken, and Markus Tuller. 2019. Ground, proximal, and satellite remote sensing of soil moisture. *Reviews of Geophysics* 57, 2 (2019), 530–616.
- [5] Sneha J Bansod and Shubhada S Thakare. 2014. Near Infrared spectroscopy based a portable soil nitrogen detector design. *Int. J. Comput. Sci. Inf. Technol.* 5, 3 (2014), 3953–3956.
- [6] Elisabetta Barberis, Franco Ajmone Marsan, R Scalenghe, A Lammers, U Schwertmann, AC Edwards, R Maguire, MJ Wilson, A Delgado, and J Torrent. 1995. European soils overfertilized with phosphorus: Part 1. Basic properties. *Fertilizer research* 45 (1995), 199–207.
- [7] Marion F Baumgardner, LeRoy F Silva, Larry L Biehl, and Eric R Stoner. 1986. Reflectance properties of soils. *Advances in agronomy* 38 (1986), 1–44.
- [8] Bimedis.com [n. d.]. Spectrometer. <https://bimedis.com/a-item/spectroscopy-equipment-bfirl-wqf-510-1946874>
- [9] bngesp. [n. d.]. MQTTSN-over-LoRa. <https://github.com/bngesp/MQTTSN-over-LoRa>
- [10] PR Bullock, X Li, and L Leonardi. 2004. Near-infrared spectroscopy for soil water determination in small soil volumes. *Canadian journal of soil science* 84, 3 (2004), 333–338.
- [11] Hans Büning-Pfaue. 2003. Analysis of water in food by near infrared spectroscopy. *Food Chemistry* 82, 1 (2003), 107–115.
- [12] Richard J Cameron, Chandra M Kudsia, and Raafat R Mansour. 2018. *Microwave filters for communication systems: fundamentals, design, and applications*. John Wiley & Sons.
- [13] Zhaoxin Chang, Fusang Zhang, Jie Xiong, Junqi Ma, Beihong Jin, and Daqing Zhang. 2022. Sensor-free soil moisture sensing using lora signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–27.
- [14] Xiaohui Chen, Xiaojun Yan, Mingkuang Wang, Yuanyang Cai, Xuefan Weng, Da Su, Jiuxin Guo, Weiqi Wang, Yong Hou, Delian Ye, et al. 2022. Long-term excessive phosphorus fertilization alters soil phosphorus fractions in the acidic soil of pomelo orchards. *Soil and Tillage Research* 215 (2022), 105214.
- [15] ATMA Choudhury and IR Kennedy. 2005. Nitrogen fertilizer losses from rice soils and control of environmental pollution problems. *Communications in soil science and plant analysis* 36, 11-12 (2005), 1625–1639.
- [16] M Demiss, Sh Beyene, and S Kidanu. 2020. Comparison of soil extractants and spectral reflectance measurement for estimation of available soil potassium in some Ethiopian soils. *Eurasian Soil Science* 53 (2020), 1100–1109.
- [17] Analog Devices/Devices. [n. d.]. HMC849A Datasheet and Product Info | Analog Devices. <https://www.analog.com/en/products/hmc849a.html#product-overview>
- [18] Nandkishor M Dhawale, Viacheslav I Adamchuk, R Viscarra, S Prasher, Joann K Whalen, and A Ismail. 2013. Predicting extractable soil phosphorus using visible/near-infrared hyperspectral soil reflectance measurements. *Paper No. 13 47* (2013).
- [19] Digikey. 2023. 915MHz Bandpass Filter. <https://www.digikey.com/en/products/filter/rf-filters/844?s=N4IgjCBcoEwdIDGUBmBDANgZwKYBoQB7KAbRAGYwB2ATgFYAOEAXQLAcAXKEAZQ4CcAlgDsA5iAC%2BBMORpRQySOmz4ipCgAYN5GFRbukXgJHij5oA>.
- [20] 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.
- [21] Dragino [n. d.]. Arduino Shield featuring LoRa® technology. <https://www.dragino.com/products/lora/item/102-lora-shield.html>

- [22] Yuda Feng, Yaxiong Xie, Deepak Ganesan, and Jie Xiong. 2022. LTE-Based Low-Cost and Low-Power Soil Moisture Sensing. In *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*. 421–434.
- [23] Yuhong Gao, Yaping Xie, Hanyu Jiang, Bing Wu, and Junyi Niu. 2014. Soil water status and root distribution across the rooting zone in maize with plastic film mulching. *Field Crops Research* 156 (2014), 40–47.
- [24] Tiziano Gomiero, David Pimentel, and Maurizio G Paoletti. 2011. Is there a need for a more sustainable agriculture? *Critical reviews in plant sciences* 30, 1-2 (2011), 6–23.
- [25] Shuang Gu, Gaohong He, Xuemei Wu, Zhengwen Hu, Leilei Wang, Gongkui Xiao, and Lin Peng. 2010. Preparation and characterization of poly (vinylidene fluoride)/sulfonated poly (phthalazine ether sulfone ketone) blends for proton exchange membrane. *Journal of applied polymer science* 116, 2 (2010), 852–860.
- [26] Jon H Hardesty and Bassam Attili. 2010. Spectrophotometry and the Beer-Lambert Law: An important analytical technique in chemistry. *Collin College, Department of Chemistry* (2010).
- [27] Richard R Harwood. 2020. A history of sustainable agriculture. In *Sustainable agricultural systems*. CRC Press, 3–19.
- [28] Yong He, Shupai Xiao, Pengcheng Nie, Tao Dong, Fangfang Qu, and Lei Lin. 2017. Research on the optimum water content of detecting soil nitrogen using near infrared sensor. *Sensors* 17, 9 (2017), 2045.
- [29] Stefanie Heinze, Michael Vohland, Rainer Georg Joergensen, and Bernard Ludwig. 2013. Usefulness of near-infrared spectroscopy for the prediction of chemical and biological soil properties in different long-term experiments. *Journal of Plant Nutrition and Soil Science* 176, 4 (2013), 520–528.
- [30] T Hussain, MA Gondal, ZH Yamani, and MA Baig. 2007. Measurement of nutrients in green house soil with laser induced breakdown spectroscopy. *Environmental monitoring and assessment* 124 (2007), 131–139.
- [31] Christopher Hutengs, Michael Seidel, Felix Oertel, Bernard Ludwig, and Michael Vohland. 2019. In situ and laboratory soil spectroscopy with portable visible-to-near-infrared and mid-infrared instruments for the assessment of organic carbon in soils. *Geoderma* 355 (2019), 113900.
- [32] Mohd Javid, Abid Haleem, Ibrahim Haleem Khan, and Rajiv Suman. 2023. Understanding the potential applications of Artificial Intelligence in Agriculture Sector. *Advanced Agrochem* 2, 1 (2023), 15–30.
- [33] Qingxu Jin and Victor C Li. 2019. Development of lightweight engineered cementitious composite for durability enhancement of tall concrete wind towers. *Cement and Concrete Composites* 96 (2019), 87–94.
- [34] Scott B Jones, JM Blonquist, David A Robinson, V Philip Rasmussen, and Dani Or. 2005. Standardizing Characterization of Electromagnetic Water Content Sensors Part 1. Methodology. *Vadose Zone Journal* 4, 4 (2005), 1048–1058.
- [35] Richard B Keithley, R Mark Wightman, and Michael L Heien. 2009. Multivariate concentration determination using principal component regression with residual analysis. *TrAC Trends in Analytical Chemistry* 28, 9 (2009), 1127–1136.
- [36] 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* (Sydney, NSW, Australia) (*MobiCom '22*). Association for Computing Machinery, New York, NY, USA, 242–254. <https://doi.org/10.1145/3495243.3517025>
- [37] Anja Klotszsch, François Jonard, Majken Caroline Looms, Jan van der Kruk, and Johan A Huisman. 2018. Measuring soil water content with ground penetrating radar: A decade of progress. *Vadose Zone Journal* 17, 1 (2018), 1–9.
- [38] J Ledieu, P De Ridder, P De Clerck, and S Dautrebande. 1986. A method of measuring soil moisture by time-domain reflectometry. *Journal of Hydrology* 88, 3-4 (1986), 319–328.
- [39] Jiaqi Li, Qingxu Jin, Wenxin Zhang, Chen Li, and Paulo JM Monteiro. 2022. Microstructure and durability performance of sustainable cementitious composites containing high-volume regenerative biosilica. *Resources, Conservation and Recycling* 178 (2022), 106038.
- [40] Tuti Mariana Lim, M Ulaganathan, and Q Yan. 2015. Advances in membrane and stack design of redox flow batteries (RFBs) for medium-and large-scale energy storage. In *Advances in Batteries for Medium and Large-Scale Energy Storage*. Elsevier, 477–507.
- [41] Weidong Liu, Frédéric Baret, Xingfa Gu, Bing Zhang, Qingxi Tong, and Lanfen Zheng. 2003. Evaluation of methods for soil surface moisture estimation from reflectance data. *International Journal of Remote Sensing* 24, 10 (2003), 2069–2083.
- [42] David B Lobell and Gregory P Asner. 2002. Moisture effects on soil reflectance. *Soil Science Society of America Journal* 66, 3 (2002), 722–727.
- [43] Mila Ivanova Luleva, Harald Van der Werff, Victor Jetten, and Freek Van der Meer. 2011. Can infrared spectroscopy be used to measure change in potassium nitrate concentration as a proxy for soil particle movement? *Sensors* 11, 4 (2011), 4188–4206.
- [44] DF Malley, L Yesmin, and RG Eilers. 2002. Rapid analysis of hog manure and manure-amended soils using near-infrared spectroscopy. *Soil Science Society of America Journal* 66, 5 (2002), 1677–1686.
- [45] SA Materchera and HR Mloza-Banda. 1997. Soil penetration resistance, root growth and yield of maize as influenced by tillage system on ridges in Malawi. *Soil and Tillage Research* 41, 1-2 (1997), 13–24.
- [46] Thomas G Mayerhöfer, Susanne Pahlow, and Jürgen Popp. 2020. The bouguer-beer-Lambert law: Shining light on the obscure. *ChemPhysChem* 21, 18 (2020), 2029–2046.
- [47] Charles E Miller. 2001. Chemical principles of near-infrared technology. *Near-infrared technology in the agricultural and food industries* 2 (2001).
- [48] Abdul Mounem Mouazen, Josse De Baerdemaeker, and Herman Ramon. 2005. Towards development of on-line soil moisture content sensor using a fibre-type NIR spectrophotometer. *Soil and Tillage Research* 80, 1-2 (2005), 171–183.
- [49] Subra Mukherjee and Shakuntala Laskar. 2019. Vis-NIR-based optical sensor system for estimation of primary nutrients in soil. *Journal of Optics* 48, 1 (2019), 87–103.
- [50] Agus Arip Munawar, Yuswar Yunus, Purwana Satriyo, et al. 2020. Calibration models database of near infrared spectroscopy to predict agricultural soil fertility properties. *Data in Brief* 30 (2020), 105469.
- [51] Cargele Nduwamungu, Noura Ziadi, L-É Parent, Gaetan F Tremblay, and Laurent Thuriès. 2009. Opportunities for, and limitations of, near infrared reflectance spectroscopy applications in soil analysis: A review. *Canadian Journal of Soil Science* 89, 5 (2009), 531–541.
- [52] Shaojun Qiu, Jiagui Xie, Shicheng Zhao, Xinpeng Xu, Yunpeng Hou, Xiufang Wang, Wei Zhou, Ping He, Adrian M Johnston, Peter Christie, et al. 2014. Long-term effects of potassium fertilization on yield, efficiency, and soil fertility status in a rain-fed maize system in northeast China. *Field crops research* 163 (2014), 1–9.
- [53] T.S. Rappaport. 2002. *Wireless Communications: Principles and Practice*. Prentice Hall PTR. <https://books.google.com/books?id=TbgQAQAAMAAJ>
- [54] Source Code Raspberry Pi. 2021. Clock Driver. <https://github.com/raspberrypi/linux/tree/rpi-6.1.y/drivers/clk>
- [55] RA Viscarra Rossel, DJJ Walvoort, AB McBratney, Leslie J Janik, and JO Skjemstad. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* 131, 1-2 (2006), 59–75.
- [56] RTL-SDR.com [n. d.]. rtl-sdr. <https://github.com/osmocom/sdr/rtl-sdr>
- [57] Aliou Saidou, I Balogoun, EL Ahoton, AM Igué, S Youl, G Ezui, and A Mando. 2018. Fertilizer recommendations for maize production in the South Sudan and Sudano-Guinean zones of Benin. *Improving the Profitability, Sustainability and Efficiency of Nutrients Through Site Specific Fertilizer Recommendations in West Africa Agro-Ecosystems: Volume 2* (2018), 215–234.
- [58] M Schärer, C Stamm, T Vollmer, E Frossard, A Oberson, H Flüher, and S Sinaj. 2007. Reducing phosphorus losses from over-fertilized grassland soils proves difficult in the short term. *Soil Use and Management* 23 (2007), 154–164.
- [59] Yongni Shao and Yong He. 2011. Nitrogen, phosphorus, and potassium prediction in soils, using infrared spectroscopy. *Soil Research* 49, 2 (2011), 166–172.
- [60] Prem P Sharma, Vikrant Yadav, Swati Gahlot, Oksana V Lebedeva, Alexandra N Chesnokova, Divesh N Srivastava, Tatiana V Raskulova, and Vaibhav Kulshrestha. 2019. Acid resistant PVDF-co-HFP based copolymer proton exchange membrane for electro-chemical application. *Journal of membrane science* 573 (2019), 485–492.
- [61] Olya Skulovich and Pierre Gentine. 2023. A long-term consistent artificial intelligence and remote sensing-based soil moisture dataset. *Scientific data* 10, 1 (2023), 154.
- [62] GC Starr, B Lowery, and ET Cooley. 2000. Soil water content determination using a network analyzer and coaxial probe. *Soil Science Society of America Journal* 64, 3 (2000), 867–872.
- [63] M Stehlik, A Czako, M Mayerová, and M Madaras. 2019. Influence of organic and inorganic fertilization on soil properties and water infiltration. (2019).
- [64] Bo Stenberg, Raphael A Viscarra Rossel, Abdul Mounem Mouazen, and Johanna Wetterlind. 2010. Visible and near infrared spectroscopy in soil science. *Advances in agronomy* 107 (2010), 163–215.
- [65] S Svanberg. 2012. Atomic and Molecular Spectroscopy: Basic Aspects and Practical Applications, 6 Springer Science & Business Media. Berlin, Heidelberg (2012).
- [66] H Tomlinson, A Traore, and ZJFE Teklehaimanot. 1998. An investigation of the root distribution of *Parkia biglobosa* in Burkina Faso, West Africa, using a logarithmic spiral trench. *Forest ecology and management* 107, 1-3 (1998), 173–182.
- [67] 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.
- [68] Guangjing Wang, Qiben Yan, Shane Patrarungrong, Juexing Wang, and Huacheng Zeng. 2023. FacER: Contrastive Attention based Expression Recognition via Smartphone Earpiece Speaker. In *IEEE INFOCOM 2023-IEEE Conference on Computer Communications*. IEEE, 1–10.
- [69] Guangjing Wang, Lan Zhang, Zhi Yang, and Xiang-Yang Li. 2018. Socialite: Social activity mining and friend auto-labeling. In *2018 IEEE 37th International Performance Computing and Communications Conference (IPCCC)*. IEEE, 1–8.
- [70] 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.

- [71] Juexing Wang, Guangjing Wang, Xiao Zhang, Li Liu, Huacheng Zeng, Li Xiao, Zhichao Cao, Lin Gu, and Tianxing Li. 2023. PATCH: A Plug-in Framework of Non-blocking Inference for Distributed Multimodal System. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 7, 3 (2023), 1–24.
- [72] Juexing Wang, Xiao Zhang, Li Xiao, and Tianxing Li. 2023. Survey for Soil Sensing with IOT and Traditional Systems. *Network* 3, 4 (2023), 482–501.
- [73] Wikipedia. 2023. Snell's law — Wikipedia, The Free Encyclopedia. <http://en.wikipedia.org/w/index.php?title=Snell's%20law&oldid=1158093539>. [Online; accessed 13-June-2023].
- [74] Yong Yang, Ying Chen, Xiaolan Zhang, Edwin Ongley, and Lei Zhao. 2012. Methodology for agricultural and rural NPS pollution in a typical county of the North China Plain. *Environmental Pollution* 168 (2012), 170–176.
- [75] Khairunnisa Mohd. Yusof, Suhaila Isaak, Nor Hafizah Ngajikin, and Nurfatihah Che Abd Rashid. 2016. LED based soil spectroscopy.