A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

ARTICLES FOR FACULTY MEMBERS

Assessment of low-cost and higher-end soil moisture sensors across various moisture ranges and soil textures / Nandi, R., & Shrestha, D.

Sensors
Volume 24 Issue 18 (2024) 5886 Pages 1-13
https://doi.org/10.3390/s24185886
(Database: MDPI)

Enhancing smart agriculture using internet of things and transformer model / Sonata, I., & Arifin, Y.

International Journal of Innovative Computing, Information and Control
Volume 21 Issue 2 (2025) Pages 515–531
https://doi.org/10.24507/ijicic.21.02.515
(Database: ICIC International)

Enhancing soil moisture prediction with explainable AI: Integrating IoT and multi-sensor remote sensing data through soft computing / Mallik, S., Chakraborty, A., Podder, K., Talukdar, S., Rahman, A., & Mishra, U.

Applied Soft Computing
Volume 180 (2025) 113406 Pages 1-19
https://doi.org/10.1016/j.asoc.2025.113406
(Database: ScienceDirect)

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

ARTICLES FOR FACULTY MEMBERS

IoT based soil moisture measurement and type prediction using advanced regression and machine learning models / Sazzad, M. M., Ahmed, T., Kibria, G., & Khan, I.

Scientific Reports
Volume 15 (2025) 35730 Pages 1-14
https://doi.org/10.1038/s41598-025-19444-2
(Database: Nature Portfolio)

IoT-driven smart agricultural technology for real-time soil and crop optimization / Shahab, H., Naeem, M., Iqbal, M., Aqeel, M., & Ullah, S. S.

Smart Agricultural Technology
Volume 10 (2025) 100847 Pages 1-9
https://doi.org/10.1016/j.atech.2025.100847
(Database: ScienceDirect)

Printed RFID sensing system: The cost-effective way to IoT smart agriculture / Gómez-Gijón, S., Salmerón, J. F., Falco, A., Loghin, F. C., Lugli, P., Morales, D. P., Rodríguez, N., & Rivadeneyra, A.

Computers and Electronics in Agriculture
Volume 232 (2025) 110116 Pages 1-10
https://doi.org/10.1016/j.compag.2025.110116
(Database: ScienceDirect)

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

ARTICLES FOR FACULTY MEMBERS

Smart farming for a sustainable future: Implementing IoT-based systems in precision agriculture / Kumar, S. N., Suriyan, K., Jacob, A. T., Varghese, A., & Francis, E.

Bulletin of the National Research Centre
Volume 49 (2025) 71 Pages 1-13
https://doi.org/10.1186/s42269-025-01366-8
(Database: Springer Nature)

Smart system for automated irrigation using internet of things devices / Ferrarezi, R. S., Peng, T. W., Ferrarezi, R. S., & Peng, T. W.

HortTechnology
Volume 31 Issue 6 (2021) Pages 642–649
https://doi.org/10.21273/HORTTECH04860-21
(Database: American Society for Horticultural Science)

Using soil moisture information to better understand and predict wildfire danger: A review of recent developments and outstanding questions / Krueger, E. S., Levi, M. R., Achieng, K. O., Bolten, J. D., Carlson, J. D., Coops, N. C., Holden, Z. A., Magi, B. I., Rigden, A. J., & Ochsner, T. E.

International Journal of Wildland Fire
Volume 32 Issue 2 (2023) Pages 111–132
https://doi.org/10.1071/wf22056
(Database: CSIRO Publishing)

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

Assessment of low-cost and higher-end soil moisture sensors across various moisture ranges and soil textures / Nandi, R., & Shrestha, D.

Sensors
Volume 24 Issue 18 (2024) 5886 Pages 1-13
https://doi.org/10.3390/s24185886
(Database: MDPI)

sensors

MDPI

Article

Assessment of Low-Cost and Higher-End Soil Moisture Sensors across Various Moisture Ranges and Soil Textures

Rajesh Nandi and Dev Shrestha *

Department of Chemical & Biological Engineering, University of Idaho, Moscow, ID 83844, USA; rajesh@bau.edu.bd

* Correspondence: devs@uidaho.edu

Abstract: The accuracy and unit cost of sensors are important factors for a continuous soil moisture monitoring system. This study compares the accuracy of four soil moisture sensors differing in unit costs in coarse-, fine-, and medium-textured soils. The sensor outputs were recorded for the VWC, ranging from 0% to 50%. Low-cost capacitive and resistive sensors were evaluated with and without the external 16-bit analog-to-digital converter ADS1115 to improve their performances without adding much cost. Without ADS1115, using only Arduino's built-in analog-to-digital converter, the low-cost sensors had a maximum RMSE of 4.79% (v/v) for resistive sensors and 3.78% for capacitive sensors in medium-textured soil. The addition of ADS1115 showed improved performance of the low-cost sensors, with a maximum RMSE of 2.64% for resistive sensors and 1.87% for capacitive sensors. The higher-end sensors had an RMSE of up to 1.8% for VH400 and up to 0.95% for the 5TM sensor. The RMSE differences between higher-end and low-cost sensors with the use of ADS1115 were not statistically significant.

Keywords: low-cost sensor; moisture measurement; soil moisture sensor; water management

Citation: Nandi, R.; Shrestha, D. Assessment of Low-Cost and Higher-End Soil Moisture Sensors across Various Moisture Ranges and Soil Textures. *Sensors* **2024**, *24*, 5886. https://doi.org/10.3390/s24185886

Academic Editors: Asim Biswas and Carlos Campillo

Received: 16 July 2024
Revised: 31 August 2024
Accepted: 5 September 2024
Published: 11 September 2024

1. Introduction

Maintaining optimal soil moisture is crucial for plant growth and productivity. Soil moisture plays a vital role in various plant processes including nutrient uptake, photosynthesis, and transpiration. Insufficient moisture can hinder plant processes and reduce yields, while excessive moisture may lead to root diseases and increased greenhouse gas emissions, such as nitrous oxide, which contributes to climate change. The soil denitrification rate is positively related to soil water-filled pore space [1]; according to a model, denitrification becomes more important at soil moisture contents greater than 60% water-filled pore space due to a decreased O_2 supply [2]. N_2O is a 300 times more potent greenhouse gas than carbon dioxide (CO_2) over a 100-year time frame [3]. Increased soil N_2O emissions contribute to climate change. Effective irrigation management avoids under- or overwatering, but it is complex due to varying factors like soil type, climate conditions, soil characteristics, crop species, cultivation practices, and crop needs [4]. According to the 2024 UNESCO report, agriculture accounts for roughly 70% of freshwater withdrawals from freshwater resources [5] and contributes to the freshwater crisis in parts of the world [6]. Irrigation water use is increasing due to the increased demand for agricultural production from population growth [7]. Distributed soil water measurement is required for precision irrigation applications. Understanding the volumetric water content (VWC) of soil is not only important for agriculture production but also for catchment hydrology, flood forecasting, landslide prediction, and other ecosystem services [8]. When it comes to deploying a large number of soil moisture sensors to minimize water waste, the availability of accurate and low-cost soil moisture sensors can be a game changer.

Various methods and sensors are employed to measure the VWC. Traditional VWC measurement using the thermo-gravimetric method is labor-intensive and lacks real-time

data. In contrast, modern soil moisture sensors provide real-time measurements and use different techniques, such as neutron thermalization, time domain reflectometry (TDR), time domain transmissometry (TDT), electrical capacitance, and impedance [9]. There is also proximal technology like the cosmic ray neutron sensor (CRNS) which provides estimates of field-averaged soil moisture within a radius of up to 240 m from the sensor [10]. While high-end sensors like TDR offer greater accuracy, they are more expensive, whereas low-cost sensors, resistive and most capacitive types, are more affordable but may be less precise. According to a market research report [11], the global soil moisture sensor market is expected to reach an estimated USD 0.39 billion by 2028 with a compound annual growth rate of 12% from 2023 to 2028. The major drivers for this market are the increasing use of soil moisture sensors in agriculture to enhance productivity and sustain groundwater levels and the growing adoption of smart agricultural equipment globally. Studies indicate that sensor performance can vary with soil type and calibration, making it essential to select the appropriate sensor and calibrate it properly for accurate irrigation management.

Current research in in situ soil moisture measurement focuses on evaluating and improving the accuracy of various sensors used to monitor the VWC. Traditional methods like thermo-gravimetric analysis are time-consuming and laborious. Modern sensors such as neutron thermalization, TDR, and TDT are quick and accurate but costly. Research indicates that low-cost sensors, such as resistive and most capacitive types, may be less precise. Low-cost sensors' accuracy may vary depending on soil textures [12] and inherent variation in sensor impedances. Selecting a sensor for a specific soil type is important for proper irrigation scheduling. Singh et al. [13] evaluated the performance of six commercial sensors, and Kargas and Soulis [14] investigated the performance of the CS655 soil moisture sensor under ten different soil conditions and showed that sensor performance varied in different soil conditions. Dong et al. [15] used EC5 and CS616 sensors under three different soil conditions and found that the sensors did not work satisfactorily in all soil types. Ferrarezi et al. [16] evaluated the performance of various types of soil moisture sensors in sandy soil in Florida and found that some sensors performed well with factory-based calibration and some sensors required soil-specific calibration. They concluded that soil-specific calibration improves the measuring accuracy significantly.

We hypothesize that the precision and accuracy of low-cost sensors can be enhanced by utilizing a higher-bit-depth analog-to-digital converter in conjunction with a low-pass analog filter. The purpose of this study is to test the hypothesis by comparing the accuracy and precision of low-cost (lower-end) and high-cost (higher-end) soil moisture sensors in measuring the VWC with and without the external chip. Four different sensors were compared. A low-cost resistive soil moisture sensor (FC-28), a generic capacitive soil moisture sensor (v.1.2), an intermediate-cost capacitance-based soil moisture sensor (VH400), and a higher-end soil moisture sensor (5TM) (www.metergroup.com) also based on capacitive sensing were used.

2. Materials and Methods

2.1. Soil Moisture Sensors

Four sensors were selected for this research based on their cost: a low-cost resistive sensor, a low-cost capacitive sensor, an intermediate-cost VH400 (Vegetronix, Sandy, UT, USA) sensor, and a high-cost 5TM sensor (Decagon, Pullman, WA, USA). A brief description of each sensor is provided in the following subsections.

2.1.1. Low-Cost Resistive Sensor

Resistive sensors measure the soil moisture by correlating electrical resistance to water content. They have two exposed probes which are placed directly into the soil (Figure 1). For this study, an FC-28 resistive sensor (Shenzhen Jiexing Weiye Electronic Co., Ltd., Shenzhen, China) was used to measure the VWC. The product description claims the sensor is made of corrosion-resistant materials for a long service life, but the exact material composition is not specified.

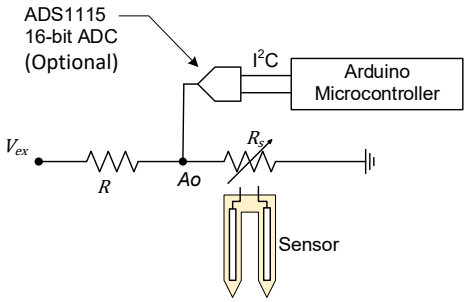


Figure 1. FC-28 resistive sensor circuit diagram (V_{ex} = 5 V and R = 10 kΩ. R_s represents the actual sensor. Analog output *Ao* can be either directly read from a microcontroller's analog-to-digital converter (ADC) or read digitally if an external ADC is used).

The sensor measures the resistance between the probes by sending an electrical current. The higher the VWC of the soil, the lower the resistance. The resistance change translates to the voltage drop across the probes, which is determined by measuring voltage A_o in Figure 1, and is given by the following:

$$A_o = V_{ex} \frac{R_s}{R_s + R} \quad (1)$$

where V_{ex} is the excitation voltage, and R_s and R are the resistances, as shown in Figure 1. The sensor's output is an analog voltage that varies nonlinearly to the soil's VWC. A microcontroller can be used to read the voltage. This type of sensor costs around USD 1.50 per unit.

2.1.2. Low-Cost Capacitive Sensor

Capacitive soil moisture sensors measure a change in the capacitance to determine the VWC. These sensors are commonly built with two parallel plates separated by a dielectric medium. The dielectric constant varies with the VWC. A low VWC in the soil corresponds to a low dielectric constant and so is the capacitance between the plates. If the VWC in the soil increases, the dielectric constant also increases, and consequently, the capacitance between the two plates also increases. A capacitive soil moisture sensor (version 1.2) (www.dfrobot.com) was used in this study.

The capacitor plate was excited with an astable 555 timer with 3 V square wave at 870 kHz. The generic sensor had a built-in voltage regulator that regulated the supply voltage between 3.3 and 5.5 V down to 3 V and a 555 timer in a stable configuration (Figure 2).

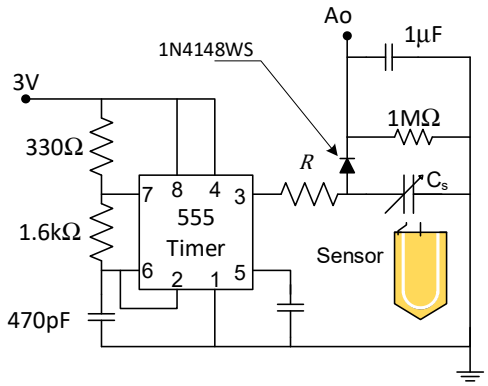


Figure 2. The capacitive soil probe circuit diagram. R = 10 kΩ. Adapted from [17]. Physical pin numbers are indicated on 555 Timer.

By coupling the sensor with a timer circuit that produces digital pulses with frequency f, we obtain an analog voltage (A_o) that an Arduino board can read, which is given by the following:

$$A_o \approx \left(\frac{3}{\sqrt{1 + (2\pi f R C_s)^2}} - 0.07 \right) \text{V} \quad (2)$$

R is 10 kΩ and C_s is the sensor capacitance (Figure 2). This voltage can then be converted to the VWC. This type of sensor costs about USD 2.00 per unit or less in bulk.

2.1.3. VH400 Moisture Sensor

The VH400 series (Vegetronix, Sandy, UT, USA) soil moisture sensor probes are based on capacitance, with two electrodes creating a capacitor that allows the electric field to penetrate the surrounding soil. The sensor uses an unbalanced design with one electrode generating the electric field and the second electrode internally fixed to ground potential. Stray capacitances can affect the measurement result, which can be reduced by shielding the cable. The circuit diagram of this sensor is proprietary and not published. Standard cable lengths for the probes are 2 m, 5 m, and 10 m long. The price of a VH400 with a 2 m cable is around USD 42.00.

2.1.4. 5TM Moisture Sensor

The Decagon 5TM sensor (Decagon, Pullman, WA, USA) has an accuracy of $\pm 0.03\ \text{m}^3/\text{m}^3$ (±3%) with a measurement range of 0–1 m^3/m^3. The electrical conductivity and temperature in soils can also be monitored using this sensor. Each of the three measurements is performed separately. Using capacitance/frequency domain technology, this sensor measures the media's dielectric constant to determine the VWC. The 5TM sensor measures the dielectric permittivity of the surrounding medium using an electromagnetic field. The sensor uses 70 MHz frequency, with a claim that this range reduces the impacts of salt and soil texture, enhancing the sensor's accuracy in the majority of soils. A thermistor is integrated into the probe for temperature measurement, and the use of three prongs allows for the measurement of resistivity and capacitance simultaneously. The cost of a 5TM sensor is around USD 350.00. The VWC is derived from the Topp equation [18].

$$\theta = 4.3 \times 10^{-6} \times \varepsilon^3 - 5.5 \times 10^{-4} \times \varepsilon^2 + 2.92 \times 10^{-2} \times \varepsilon - 5.3 \times 10^{-2} \quad (3)$$

where θ is the VWC and ε is the relative dielectric permittivity to air. The Topp equation results in measurements within 3% VWC of the actual soil VWC for mineral soil.

2.2. Signal Stability Improvement Using an ADC

To increase the analog-to-digital conversion resolution and to filter the high-frequency noise, we used an external 16-bit ADC (Texas Instruments' ADS1115). ADS1115 is a 16-bit, low-power, I2C-compatible analog-to-digital converter that includes an oscillator and a low-drift voltage reference. The ADS1115 breakout board costs about USD 5.00 apiece. In addition to ADC functionality, the chip consists of a low-pass filter, a programmable gain amplifier (PGA), and a programmable data rate. Furthermore, the ADS1115 model has a multiplexer that can measure either two differentials or four single-ended inputs. A block diagram of ADS1115 is shown in Figure 3 with the circuit diagram for the connection.

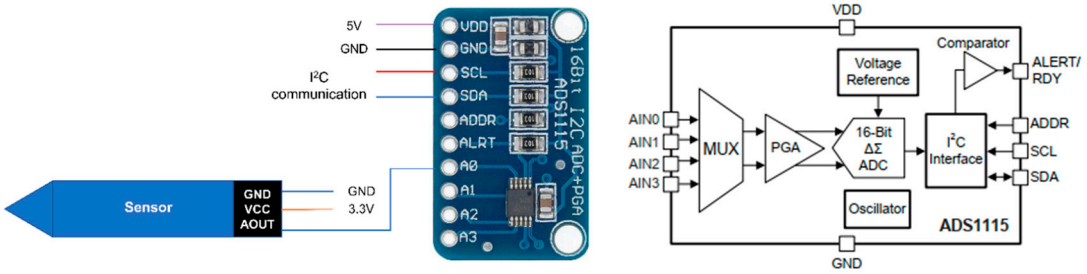


Figure 3. Sensor connected to ADS1115 breakout board (**left**) and ADS1115 block diagram (**right**), adapted from the user manual of ADS1115 [19]. Only one channel was used with unity gain. Reprinted with permission courtesy of Texas Instruments.

Arduino is a trademarked yet open architecture family of microcontrollers. Arduino Nano, one of its models, features an ATmega328P microprocessor (Microchip, Chandler, AZ, USA), equipped with eight multiplexed 10-bit analog-to-digital converters, labeled A0 through A7. The microcontroller has an ADC noise reduction mode that stops the CPU and all I/O modules except the asynchronous timer and ADC to minimize switching noise during ADC conversions. In contrast, ADS1115 utilizes a delta–sigma analog-to-digital converter. Delta–sigma ADCs offer low noise, are tolerant to a wide range of frequencies, and provide a more stable high-resolution output compared to the successive approximation ADC used in the Nano.

2.3. Sample Collection and Preparation

An aggregate soil sample was collected from the University of Idaho's Parker Farm (latitude 46.723099, longitude −116.962002). The soil is dark brown silty loam. This is a typical soil found in the Inland Northwest region of the United States, commonly referred to as the Palouse Series Soil. The soil taxonomic class is "Fine-silty, mixed, superactive, mesic Pachic Ultic Haploxerolls" [20]. The soil textural analysis showed 26% sand, 58% silt, and 16% clay.

The collected soil sample was ground and oven dried at 105 °C following the ANSI/ASABE S633 standard [21] until a constant weight was observed which took about 24 h. The dried soil was then separated into coarse and fine fractions using 1.981 mm and 0.147 mm sieves. The coarse soil represents soil particles of sizes between 1.981 mm and 0.147 mm and fine soil represents a soil size of less than 0.147 mm. The demarcation was selected around particle size for sand which is from 2 mm to 0.2 mm. Fine and coarse soils were mixed at a 1:1 ratio (by weight) and thoroughly mixed to prepare a medium-textured soil. Figure 4 shows the visual difference between the fine- and coarse-textured soils used in this study.

Figure 4. The fine (**left**) and coarse (**right**) soil used in this study.

2.4. Measurement of Soil Bulk Density

Soil bulk density was measured in the laboratory from a disturbed sample. The soil was oven-dried at 105 °C for 24 h and placed into a container of known volume, and the mass of the soil was recorded using a weighing scale. The container was filled while gently shaking, and physical compaction was not used. Variability in bulk density measurement is unavoidable; however, we expect that the variability in bulk density measurement is random error and reduced when averaged from replicates. The bulk density was determined by dividing the mass by the volume of the soil. The bulk density and standard deviation (SD) of course-, fine-, and medium-textured soil was 1.11 (SD = 0.02), 1.07 (SD = 0.02), and 1.26 (SD = 0.04) g/cm^3, respectively.

2.5. Experimental Setup

The circuit connection diagram using different sensors is shown in Figure 5. One of the capacitive, resistive, or VH400 sensors was connected to the analog input A0 pin of Arduino Nano. The data were transmitted to the host computer connected to the Arduino Nano via RS232 serial communication. All experiments were conducted at a room temperature of 20 °C.

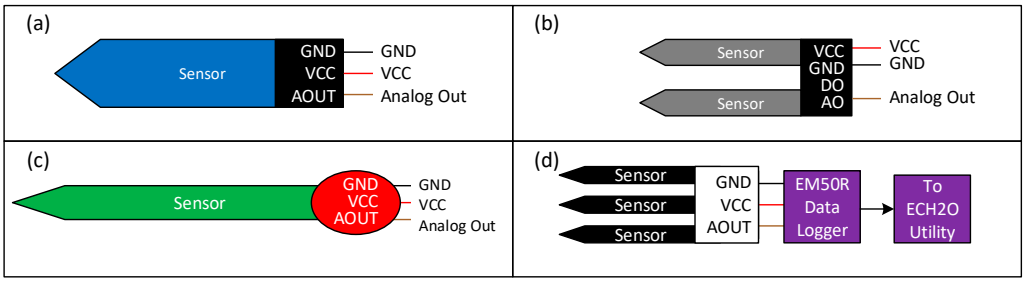


Figure 5. Circuit connection diagram of (**a**) low-cost capacitive sensor (v.1.2), (**b**) FC-28 low-cost resistive sensor, (**c**) VH400 sensor, and (**d**) 5TM sensor.

The 5TM sensor was directly read from an Em50R data logger [22]. The processed data from the signal processing unit were received in a computer using RS232 serial communication or the ECH20 utility (Decagon, Pullman, WA, USA) for the Em50R data logger. Figure 6 shows the experimental setup used in this study.

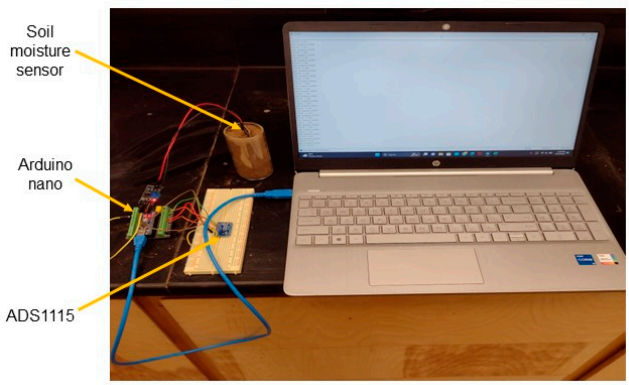


Figure 6. Experimental setup for this study with ADS1115.

2.6. Experimental Procedure

Three different soil textures were used to assess the accuracy of four soil moisture sensors. A calculated amount of water by volume was added to dry soil to prepare 0%, 10%, 20%, 30%, 40%, and 50% of soil VWC. The water volume was measured using a 50 mL burette (±0.1 mL), and the soil mass was measured using a digital balance with an accuracy of 0.01 g (Sartorius laboratory balance, Model L2200S, BioArrow Technology Ltd.,

Göttingen, Germany). To maintain homogeneity, the calculated amount of water was added incrementally to the dry soil, and the soil was thoroughly mixed to achieve a consistent distribution of moisture throughout the soil sample. The soil was manually compacted as tightly and uniformly as possible to ensure optimal sensor contact and to prevent the formation of large voids. This process was repeated for each targeted VWC level. The sensors were then inserted into the soil to measure the VWC. The analog sensor output value was recorded against the VWC. For measurement using the 5TM sensor, an EM50R data logger was used with the ECH2O utility, where the output was the percentage VWC. Each experiment with a combination of soil and moisture was replicated three times for the same input to evaluate the sensor's precision. More than 100 repeated measurements were taken within each replication to calculate the statistics.

2.7. Data Analysis

The sensor performances were evaluated for measurement accuracy and precision. The closer the measured values are, the higher the sensor precision. The difference between the actual and the measured values represents accuracy. A third-order polynomial was fitted to obtain calibration curves for the selected sensors in fine and coarse soils using Microsoft Excel's Data analysis tool pack (Office 360-2024 version, Microsoft, Redmond, WA, USA). A third-order polynomial was chosen to be consistent with the 5TM sensor which uses Equation (3). A good sensor should have a high accuracy and a high precision. An ANOVA on the Root Mean Squared Error (RMSE) was performed to find the statistical significance of differences in sensor performance among soil groups.

3. Results

3.1. Performance Comparison without ADS1115

Figure 7 shows the actual versus predicted moisture content for various soil textures. From the figure, it is evident that without additional signal processing, the sensor output exhibited greater deviations for the low-cost capacitive and resistive sensors, whereas the VH400 and 5TM sensors consistently demonstrated a more precise output. This observation underscores the notable differences in precision among the sensors which may be stemming from the digital noise of the internal ADC of the Arduino board. With no modifications, the VH400 and 5TM sensors exhibited more consistent readings compared to the low-cost capacitive and resistive sensors. Table 1 shows the results of regression analysis including different statistical parameters for sensor performance in different types of soil texture. For coarse soil, the resistive sensor had the lowest R^2 values for the linear regression line between the actual and predicted VWC (Table 1). The VH400 and 5TM sensors showed a higher R^2. The higher the value of R^2, the greater the degree of fit of the model to the data; however, it does not indicate whether the estimates and predictions are biased or not [23].

Table 1. Results of regression analysis (RMSE and SD are % VWC).

Soil Texture	Sensor Type	R^2	RMSE	SD
Coarse	Capacitive	0.99	1.88	1.43
	Capacitive with ADS1115	0.99	1.27	0.10
	Resistive	0.95	3.83	2.06
	Resistive with ADS1115	0.99	1.55	0.09
	VH400	1.00	0.85	0.51
	5TM	1.00	0.82	0.34

Table 1. *Cont.*

Soil Texture	Sensor Type	R^2	RMSE	SD
Fine	Capacitive	0.96	3.22	2.24
	Capacitive with ADS1115	0.99	1.81	0.08
	Resistive	0.93	4.39	2.59
	Resistive with ADS1115	1.00	0.64	0.04
	VH400	1.00	1.17	0.71
	5TM	1.00	0.95	0.29
Medium	Capacitive	0.95	3.78	1.48
	Capacitive with ADS1115	0.99	1.87	0.06
	Resistive	0.92	4.79	2.07
	Resistive with ADS1115	0.98	2.64	0.14
	VH400	0.99	1.80	0.69
	5TM	1.00	0.47	0.28

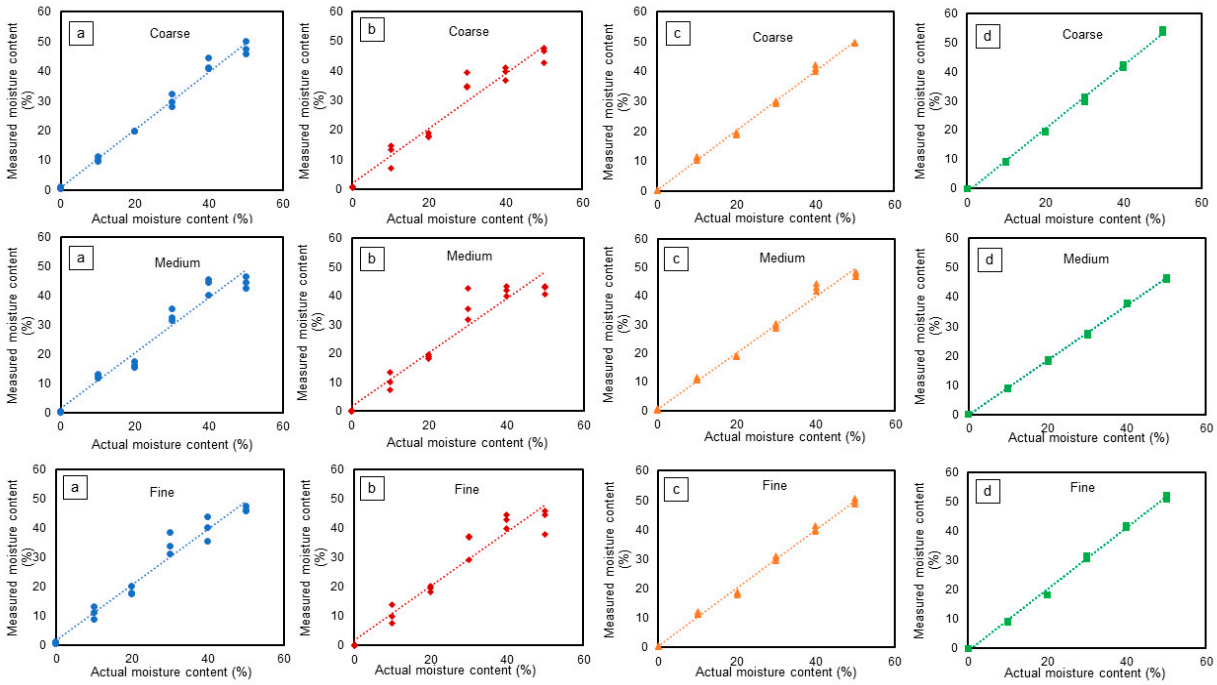


Figure 7. Calibration curve for sensors in coarse, medium, and fine soil: (**a**) capacitive, (**b**) resistive, (**c**) VH400, and (**d**) 5TM, each color representing a specific sensor. Each experiment was conducted with three replications.

The result of this study agrees with Okasha et al. [24] who obtained R^2 values for low-cost capacitive sensors from 0.96 to 0.99. Adla et al. [9] reported an R^2 value of 0.89 with a first-order linear line in sandy soil for low-cost resistive sensors similar to the one used in this research. Praštalo et al. [25] evaluated low-cost capacitive and resistive sensors in gravel and sandy soil with linear fit regression R^2 values of 0.9036 and 0.9023, respectively. The difference between the current study and the previous studies may be due to the difference in sensor models, the order of polynomials (first- vs. third-order polynomial) used to fit the data, and experimental conditions.

The RMSE values for the capacitive, resistive, VH400, and 5TM sensors were 1.88%, 3.83%, 0.85%, and 0.82% VWC, respectively (Table 1). The RMSE is a measure of accuracy;

therefore, the VH400 and 5TM sensors showed less error in estimating the VWC than the capacitive and resistive sensors. The SD for the capacitive, resistive, VH400, and 5TM sensors was found to be 1.43%, 2.06%, 0.51%, and 0.34% VWC, respectively. The SD is a measure of precision. The lower the value of the SD, the higher the precision. The VH400 and 5TM sensors showed higher precision than the capacitive and resistive sensors. ANOVA showed that the capacitive and resistive sensors' error was significantly higher than that of the VH400 and 5TM sensors ($p = 0.007$).

The resistive sensor had the lowest R^2 of 0.93 compared to the VH400 and 5TM sensors at 0.995 and 0.997. The generic capacitive sensor had an intermediate R^2 value of 0.96. The RMSE and SD values for the capacitive and resistive sensors were significantly higher compared to the VH400 and 5TM sensors ($p = 0.007$). Therefore, the precision of the VH400 and 5TM sensors was also higher in fine soil.

In medium-textured soil, a similar trend was observed as in coarse and fine soils. The 5TM sensor showed a better fit with an R^2 value of 1.0 followed by the VH400 ($R^2 = 0.99$), capacitive ($R^2 = 0.95$), and resistive ($R^2 = 0.92$) sensors. The results of the current study are in agreement with Praštalo et al. [25], who showed an R^2 value of 0.9856 using the VH400 sensor. The R^2 value for the 5TM sensor was lower (0.95) in a previous study by Li et al. [26] and improved with a linear calibration model (0.99).

Comparing the actual VWC and predicted value (Figure 7), a simple capacitive and resistive soil moisture sensor with no intermediate A/D converter frequently overestimated or underestimated the VWC. The RMSE for the resistive sensor was highest at 4.8%. For instance, at 50% moisture content, the predicted value of the capacitive moisture sensor was between 42.52% and 46.43% in medium-textured soil. The resistive sensor, on the other hand, predicted between 40.5% and 43.40%. Overall, the capacitive soil moisture sensor had ±0.09% to 4.68%, 0.08% to 7.48%, and 0.08% to 8.54% error in coarse-, medium-, and fine-textured soils, respectively. Conversely, the resistive soil moisture sensor overestimated or underestimated the VWC in coarse-, medium-, and fine-textured soils by up to 9.35%, 12.47%, and 11.97%, respectively. However, the VH400 and 5TM sensors showed close values to the actual ones. The maximum difference between the actual VWC and the VH400 predicted value was 4.31%. The 5TM sensor overestimated or underestimated the VWC by up to 4.23%. The specification for the 5TM sensor's accuracy was ±3% VWC in mineral soils with an electrical conductivity of less than 10 dS/m. However, the specifications are usually provided for a 95% confidence interval or RMSE, and some data points are expected to be outside of the 95% confidence interval. It should be noted that the error of 4.23% was the extreme we observed.

The discrepancy in the performance of the simple resistive and capacitive sensors in different soil textures could be attributed to insufficient soil contact. The accuracy of the VWC measurement depends on the sensor installation technique. The installation should be conducted carefully because the soil moisture sensor only detects a tiny amount of the soil surrounding the sensor. Many previous studies [27–30] highlighted that the sensor and soil must make good contact in order to prevent an air gap. However, the error of estimation in fine, coarse, and medium soil textures was not significantly different ($p = 0.21$). The results of this study suggest that the VH400 and 5TM sensors performed better than the simple resistive and capacitive sensors.

3.2. Performance Comparison with ADS1115

The use of ADS1115 provided more stable readings and improved the accuracy for low-cost sensors. The R^2 value using ADS1115 with the capacitive sensor for all soil types was around 0.99 (Table 1). The R^2 value using ADS1115 with the resistive sensor for coarse-, medium-, and fine-textured soils was 0.99, 0.98, and 1.0, respectively. The capacitive sensor showed an R^2 value of 0.99 for all types of soil textures. It is evident from Table 1 that the addition of ADS1115 improved the R^2 values of both the low-cost sensors. The RMSE values from the capacitive and resistive sensors also reduced with ADS1115. The RMSE

values with ADS1115 were between 1.27% and 1.87% moisture content for the capacitive sensor and between 0.64% and 2.64% VWC for the resistive sensor.

It is important to note that the results presented in Table 1 pertain to soil at 20 °C. Since both the resistivity and dielectric constant of water vary with temperature, soil moisture measurements taken across different temperatures require corrections to account for these changes, especially if such corrections have not already been applied, as is often the case with low-cost sensors.

The SD for the capacitive sensor was 0.10%, 0.06%, and 0.03% of the measured VWC for coarse-, medium-, and fine-textured soils. On the other hand, the resistive sensor showed an SD of 0.09%, 0.14%, and 0.04% for coarse-, medium-, and fine-textured soils, respectively. The addition of ADS1115 lowered the SD and improved the R^2. ANOVA showed that the predicted SD using ADS1115 was significantly lower for both the resistive and capacitive sensors ($p = 0.00$) compared to using the built-in Arduino ADC.

The measured vs. predicted VWC with the capacitive and resistive sensors with the addition of ADS1115 is shown in Figure 8. With ADS1115, both the capacitive and resistive sensors showed a consistent output. The capacitive sensor showed a deviation of up to 2.16%, 3.2%, and 2.47% of VWC in coarse-, medium-, and fine-textured soils, respectively. The deviation in the resistive sensor's predicted VWC value in coarse-, medium-, and fine-textured soils was up to 2.6%, 4.73%, and 1.06%, respectively, with ADS1115. Therefore, a significant precision improvement in the low-cost capacitive and resistive sensors was observed with ADS1115.

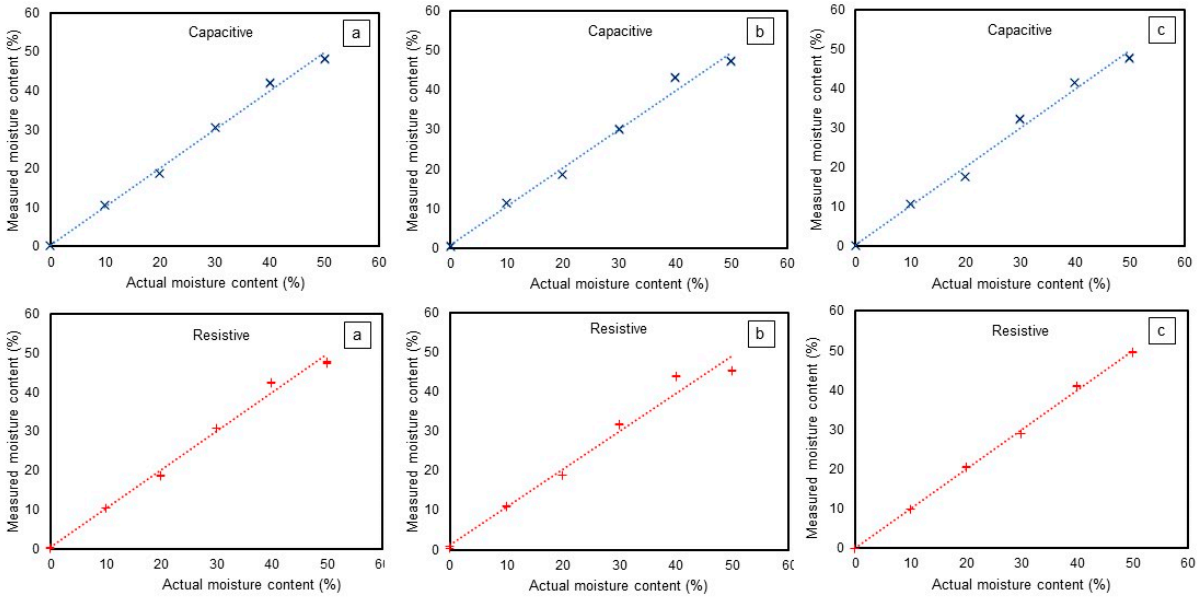


Figure 8. Calibration curve using ADS1115 with capacitive and resistive sensors, also differentiated by color in (**a**) coarse-, (**b**) fine-, and (**c**) medium-textured soil. Each experiment was conducted with three replications.

The average voltage fluctuation using the low-cost capacitive sensor in fine-textured soil with a constant VWC was around 27 mV using the Arduino A/D converter. The addition of ADS1115 reduced the signal fluctuation to only 0.95 mV. Similarly, the generic resistive sensor showed an average fluctuation of 59 mV, and the addition of ADS1115 reduced the fluctuations to 0.77 mV. The significant reduction in the voltage fluctuation observed after adding ADS1115 indicates that ADS1115 effectively improves the stability and accuracy of the analog measurements. The high fluctuations in voltage observed with the low-cost capacitive and resistive soil moisture sensors likely include noise from various sources such as electrical interference, sensor imperfections, and environmental factors.

ADS1115's higher resolution and potential filtering capabilities help reduce the impact of noise on the sensor readings, resulting in much lower voltage fluctuations.

One possible reason for why ADS1115 improved the predictive performance could be its built-in low-pass filter, which acts as an anti-alias filter to reduce high-frequency noise and interference in the analog signal. The filter limits the impact of electrical noise on measurements prior to converting the signal to a digital value. The filter had a cutoff frequency of about 5 Hz (read from Figure 9 at −3.01 dB gain). Thus, the analog input filter contributes to a more stable analog signal by filtering away high-frequency noise. Conversely, without ADS1115, the inconsistency in sensor readings may have originated from digital noise in the internal ADC of the Arduino board. Another factor that could have contributed to improved performance is the bit resolution. The Arduino Nano has a 10-bit ADC with an absolute accuracy of 2 Least Significant Bit (LSB). For 0–5 V conversion at 10 bit, a 2-bit error corresponds to an error of 9.8 mV. ADS1115 has a 16-bit ADC resolution, which provides 64 times higher resolution at 0.16 mV than a 10-bit ADC.

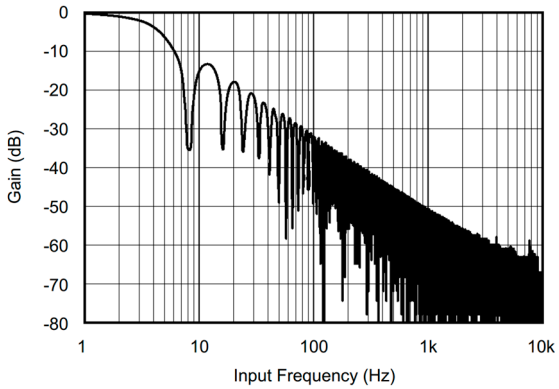


Figure 9. ADS1115 frequency response at 8 samples per second data rate [19]. Reprinted with permission courtesy of Texas Instruments.

One of the important considerations that determine the choice of sensors by commercial growers is the cost, particularly if the field with significant spatial variability [31]. The cost of the resistive and capacitive sensors is relatively lower than that of the VH 400 and 5TM sensors. Therefore, comparing the costs, the resistive and capacitive sensors with ADS1115 could be a good choice for a soil moisture monitoring system. Further study needs to be conducted to evaluate the long-term durability and reliability of these low-cost sensors in real-world field conditions.

4. Conclusions

The VH400 and 5TM sensors showed better precision and accuracy than the low-cost capacitive and resistive sensors with direct analog output measurement using Arduino's analog-to-digital converter. The low-cost sensors exhibited a higher signal fluctuation and offset when connected directly to the Arduino microcontroller. The addition of an external high-bit-resolution analog-to-digital converter (ADS1115) with a built-in low-pass filter significantly enhanced the prediction performance of these low-cost sensors. Specifically, the RMSE for the resistive sensor reduced from 4.79% to 2.64%, and the standard deviation of measurements reduced from 2.58% to 0.14%. Similarly, for the capacitive sensor, the RMSE was reduced from 3.87% to 1.87%, with the standard deviation decreasing from 2.24% to 0.10%. Notably, the difference in the RMSE for the low-cost and high-end sensors with ADS1115 was statistically insignificant.

Therefore, low-cost sensors, when paired with affordable analog-to-digital converters like ADS1115, offer a viable and cost-effective alternative for soil moisture measurement in the field, providing performance comparable to that of more expensive options. Furthermore, the sensors demonstrated consistent performance across different soil textures with no significant variation. However, it is important to note that the testing of low-cost sensors

in this research was limited to a short-term evaluation. Long-term sensor performance remains an area for future investigation.

Author Contributions: Conceptualization, R.N. and D.S.; methodology, R.N. and D.S.; software, R.N.; validation, R.N. and D.S.; formal analysis, R.N.; investigation, R.N.; resources, D.S.; data curation, R.N.; writing—original draft preparation, R.N.; writing—review and editing, R.N. and D.S.; visualization, R.N.; supervision, D.S.; project administration, D.S.; funding acquisition, D.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the USDA National Institute of Food and Agriculture, Hatch project 1009342, and Department of Chemical and Biological Engineering at the University of Idaho.

Data Availability Statement: Data can be made available upon request.

Acknowledgments: The authors acknowledge the Department of Soil and Water Systems at the University of Idaho for providing the 5TM sensors.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pan, B.; Xia, L.; Lam, S.K.; Wang, E.; Zhang, Y.; Mosier, A.; Chen, D. A Global Synthesis of Soil Denitrification: Driving Factors and Mitigation Strategies. *Agric. Ecosyst. Environ.* **2022**, *327*, 107850. [CrossRef]
2. Davidson, E.A. Fluxes of Nitrous Oxide and Nitric Oxide from Terrestrial Ecosystems. In *Microbial Production and Consumption of Greenhouse Gases: Methane, Nitrogen Oxide, and Halomethanes*; Rogers, J.E., Whitman, W.B., Eds.; ASM Press: Washington, DC, USA, 1991; pp. 219–235.
3. IPCC. *IPCC Second Assessment Report: Climate Change 1995 (SAR)—Working Group I: The Science of Climate Change*; IPCC: New York, NY, USA, 1996.
4. NRCS. Irrigation Water Management. In *National Engineering Handbook: Irrigation Guide*; NRCS, Ed.; United States Department of Agriculture: Washington, DC, USA, 1997; pp. 9.1–9.223.
5. UNESCO. Executive Summary. In *Water for Prosperity and Peace*; United Nations Educational, Scientific and Cultural Organization: Paris, France, 2024; pp. 1–8, ISBN 9789231006579.
6. Schulz, S.; Darehshouri, S.; Hassanzadeh, E.; Tajrishy, M.; Schüth, C. Climate Change or Irrigated Agriculture—What Drives the Water Level Decline of Lake Urmia. *Sci. Rep.* **2020**, *10*, 236. [CrossRef] [PubMed]
7. UNESCO. *Nature-Based Solutions for Water*; United Nations Educational, Scientific and Cultural Organization: Paris, France, 2018; ISBN 9789231002649.
8. Hardie, M. Review of Novel and Emerging Proximal Soil Moisture Sensors for Use in Agriculture. *Sensors* **2020**, *20*, 6934. [CrossRef] [PubMed]
9. Adla, S.; Rai, N.K.; Karumanchi, S.H.; Tripathi, S.; Disse, M.; Pande, S. Laboratory Calibration and Performance Evaluation of Low-Cost Capacitive and Very Low-Cost Resistive Soil Moisture Sensors. *Sensors* **2020**, *20*, 363. [CrossRef] [PubMed]
10. Flynn, K.D.; Wyatt, B.M.; McInnes, K.J. Novel cosmic ray neutron sensor accurately captures field-scale soil moisture trends under heterogeneous soil textures. *Water* **2021**, *13*, 3038. [CrossRef]
11. Lucintel. Soil Moisture Sensor Market: Market Size, Trends and Growth Analysis 2023–2028. 2023. Available online: https://www.lucintel.com/soil-moisture-sensor-market.aspx (accessed on 1 November 2023).
12. Serrano, D.; Ávila, E.; Barrios, M.; Darghan, A.; Lobo, D. Surface Soil Moisture Monitoring with Near-Ground Sensors: Performance Assessment of a Matric Potential-Based Method. *Meas. J. Int. Meas. Confed.* **2020**, *155*, 107542. [CrossRef]
13. Singh, J.; Lo, T.; Rudnick, D.R.; Dorr, T.J.; Burr, C.A.; Werle, R.; Shaver, T.M.; Muñoz-Arriola, F. Performance Assessment of Factory and Field Calibrations for Electromagnetic Sensors in a Loam Soil. *Agric. Water Manag.* **2018**, *196*, 87–98. [CrossRef]
14. Kargas, G.; Soulis, K.X. Performance Evaluation of a Recently Developed Soil Water Content, Dielectric Permittivity, and Bulk Electrical Conductivity Electromagnetic Sensor. *Agric. Water Manag.* **2019**, *213*, 568–579. [CrossRef]
15. Dong, Y.; Miller, S.; Kelley, L. Performance Evaluation of Soil Moisture Sensors in Coarse-and Fine-Textured Michigan Agricultural Soils. *Agriculture* **2020**, *10*, 598. [CrossRef]
16. Ferrarezi, R.S.; Nogueira, T.A.R.; Zepeda, S.G.C. Performance of Soil Moisture Sensors in Florida Sandy Soils. *Water* **2020**, *12*, 358. [CrossRef]
17. DFROBOT. *Capacitive Soil Moisture Sensor*; DFROBOT: Shanghai, China, 2015.
18. Topp, G.C.; Davis, J.L.; Annan, A.P. Electromagnetic Determination of Soil Water Content: Measurements in Coaxial Transmission Lines. *Water Resour. Res.* **1980**, *16*, 574–582. [CrossRef]
19. Texas Instruments. *ADS111x Ultra-Small, Low-Power, I2C-Compatible, 860-SPS, 16-Bit ADCs with Internal Reference, Oscillator, and Programmable Comparator*; Texas Instruments: Dallas, TX, USA, 2018.
20. USDA. *Official Series Description—PALOUSE Series*; USDA: Hutchinson, KS, USA, 2013.
21. ASABE. *Testing Protocol for Landscape Irrigation Soil Moisture-Based Control Technologies*; ASABE: St. Joseph, MI, USA, 2020.
22. METER Group. *Em50 Series Loggers*; METER Group: Pullman, WA, USA, 2019.

23. Matula, S.; Bát'ková, K.; Legese, W.L. Laboratory Performance of Five Selected Soil Moisture Sensors Applying Factory and Own Calibration Equations for Two Soil Media of Different Bulk Density and Salinity Levels. *Sensors* **2016**, *16*, 1912. [CrossRef] [PubMed]
24. Okasha, A.M.; Ibrahim, H.G.; Elmetwalli, A.H.; Khedher, K.M.; Yaseen, Z.M.; Elsayed, S. Designing Low-Cost Capacitive-Based 401 Soil Moisture Sensor and Smart Monitoring Unit Operated by Solar Cells for Greenhouse Irrigation Management. *Sensors* **2021**, *21*, 5387. [CrossRef] [PubMed]
25. Praštalo, P.; Brajović, L.; Prodanović, D. Using of Low Cost Moisture Sensors in Laboratory Experiments. In Proceedings of the International Conference on Contemporary Theory and Practice in Construction XV, Banja Luka, Bosnia and Herzegovina, 10 June 2022; pp. 530–539.
26. Li, B.; Wang, C.; Ma, M.; Li, L.; Feng, Z.; Ding, T.; Li, X.; Jiang, T.; Li, X.; Zheng, X. Accuracy Calibration and Evaluation of Capacitance-Based Soil Moisture Sensors for a Variety of Soil Properties. *Agric. Water Manag.* **2022**, *273*, 107913. [CrossRef]
27. Patrignani, A.; Ochsner, T.E.; Feng, L.; Dyer, D.; Rossini, P.R. Calibration and Validation of Soil Water Reflectometers. *Vadose Zone J.* **2022**, *21*, e20190. [CrossRef]
28. Saito, T.; Oishi, T.; Inoue, M.; Iida, S.; Mihota, N.; Yamada, A.; Shimizu, K.; Inumochi, S.; Inosako, K. Low-Error Soil Moisture Sensor Employing Spatial Frequency Domain Transmissometry. *Sensors* **2022**, *22*, 8658. [CrossRef] [PubMed]
29. Walker, J.P.; Willgoose, G.R.; Kalma, J.D. In Situ Measurement of Soil Moisture: A Comparison of Techniques. *J. Hydrol.* **2004**, *293*, 85–99. [CrossRef]
30. Yan, G.; Bore, T.; Bhuyan, H.; Schlaeger, S.; Scheuermann, A. The Technical Challenges for Applying Unsaturated Soil Sensors to Conduct Laboratory-Scale Seepage Experiments. *Sensors* **2022**, *22*, 3724. [CrossRef] [PubMed]
31. Kukal, M.S.; Irmak, S.; Sharma, K. Development and Application of a Performance and Operational Feasibility Guide to Facilitate Adoption of Soil Moisture Sensors. *Sustainability* **2020**, *12*, 321. [CrossRef]

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

Enhancing smart agriculture using internet of things and transformer model / Sonata, I., & Arifin, Y.

International Journal of Innovative Computing, Information and Control
Volume 21 Issue 2 (2025) Pages 515–531
https://doi.org/10.24507/ijicic.21.02.515
(Database: ICIC International)

International Journal of Innovative Computing, Information and Control
Volume **21**, Number **2**, April **2025**

 ISSN 1349-4198
pp. 515–531

ENHANCING SMART AGRICULTURE USING INTERNET OF THINGS AND TRANSFORMER MODEL

Ilvico Sonata and Yulyani Arifin

Computer Science Department, BINUS Graduate Program – Doctor of Computer Science
Bina Nusantara University
Jl. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
ilvico@binus.ac.id; yarifin@binus.edu

Received May 2024; revised September 2024

Abstract. *Agricultural production levels are highly dependent on weather and soil conditions. To ensure the sustainability of agricultural production can use smart agriculture with Internet of Things (IoT) technology that currently develops rapidly along with the availability of various sensors to detect weather and soil conditions. This paper will propose a smart agriculture through monitoring agricultural soil moisture, temperature, and air humidity, as well as predicting temperature and air humidity data using IoT and the Transformer model. Data taken from these sensors will be analyzed using Transformer time series forecasting to see temperature and humidity trends for determining suitable plant types and predicting agricultural yield. The experimental results show that the Transformer model has a better level of accuracy compared to the machine learning model with MAE, MSE and RMSE of 0.027, 0.0011 and 0.034 for temperature prediction results and 0.057, 0.0053 and 0.073 for humidity prediction results.*
Keywords: Deep learning, Internet of Things, Smart agriculture, Transformer

1. **Introduction.** Agriculture is a very important aspect for a country. Agricultural development can improve the welfare of society in a country. Some examples of agricultural products include fruit, vegetables, rice, corn, etc. For this reason, sustainable agriculture needs to be developed. One way to achieve sustainable agriculture is through the application of technology in monitoring agricultural land [1]. The use of technology in monitoring agricultural land allows action to be taken to prevent crop failure and damage to agricultural land due to weather or unfavorable soil conditions. One of the developments in sustainable agriculture is switching to smart agriculture [2]. Smart agriculture changes the traditional agricultural model where data collection for decision making and action is minimal and monitoring is done manually into an agricultural system with automatic data collection and processing to obtain information that is important for automatic decision making and monitoring [3]. Smart agriculture is also suitable for supporting the sustainable of the agricultural production like the Sustainable Development Goals (SDGs) from The United Nations Educational, Scientific and Cultural Organization (UNESCO) especially Goals No 12 that is related with responsible consumption and production [4]. Using digital technologies can leverage for sustainability of agricultural production based on the United Nations Development Programme (UNDP) Report [5].

2. **Literature Review.** Mandal et al. [3] and Qazi et al. [6] in their research explains the use of IoT in the application of smart agriculture. In this smart agricultural application, IoT is used to collect soil and weather condition data using various sensors for the monitoring process and storing data in a database. Data from various sensors is used to

DOI: 10.24507/ijicic.21.02.515

regulate soil conditions such as adding water and fertilizer remotely and automatically. In addition, data from various sensors is used to observe air temperature and humidity. The use of IoT in smart agriculture can increase the efficiency of agricultural processes and increase agricultural production. Several other studies on smart agriculture using IoT have been carried out, among others, by Ayaz et al. [7], Saraswathi et al. [8], and Prathibha et al. [9].

To obtain more in-depth information regarding soil conditions and trends in changes in weather conditions based on information obtained from various sensors used in IoT, the data collected from these sensors is then analyzed to predict certain conditions based on historical data. This data is very important to use to predict agricultural yields and determine the types of crops that are suitable for planting according to the soil and weather conditions needed to produce maximum harvests [10]. Several methods for predicting these conditions currently use artificial intelligence technology such as machine learning and deep learning. The use of machine learning to process data obtained from IoT in smart agricultural applications has been carried out, among others, by Sanjana et al. [11]. In their research, they used K-Nearest Neighbors (KNN), Support Vector Machine (SVM) and logistic regression machine learning models to predict crop yields. The results of the experiments carried out show that the SVM model has the highest accuracy. Machine learning has also been used to process data from IoT through research conducted by Reddy et al. [12]. In their research, they used decision trees to predict the level of water requirements in smart agriculture. Both studies used temperature, humidity, and soil moisture sensors to measure weather and soil conditions.

The use of deep learning models in processing data obtained from various sensors used by IoT for smart agricultural applications has been discussed in depth by Zhu et al. [13]. In their discussion, they explained the concepts, applications, and opportunities for using deep learning models such as Artificial Neural Network (ANN), Back Propagation (BP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) and Generative Adversarial Network (GAN). They concluded that the use of deep learning to process data obtained from various sensors is very useful and provides greater value in increasing agricultural productivity through plant disease detection, crop classification, weed identification, weather forecasting, and crop yield prediction. Altalak et al. [14] have also discussed the use of deep learning technology for smart agriculture. In their research, they discuss the use of various deep learning models such as CNN, RNN and Long Short-Term Memory (LSTM) as well as several combinations such as the combination of CNN and LSTM.

The combination of machine learning models and deep learning models in smart agriculture applications using IoT has been carried out by Durai and Shamili [15]. In their research, they used this artificial intelligence model for crop recommendations, weed identification, pesticide recommendations, and cost estimation. The development of the artificial intelligence model used was carried out through learning process using datasets obtained from various sensors used by IoT. The artificial intelligence models used include machine learning models such as SVM, Decision Tree Classifier, Random Forest Classifier, and XGBoost Classifier for the plant recommendation and pesticide recommendation processes. Deep learning models are used for weed identification. This deep learning model uses a pre-trained model from ResNet152V2. Meanwhile, for the cost estimation process, a regression model in machine learning, namely the XGBoost Regressor, is used.

The Transformer model was first introduced by Vaswani et al. [16] increasingly used as a sequence-to-sequence model that is very suitable for processing time series data. Research conducted by Sabililah and Adytia [17] shows that time series forecasting using a Transformer model has better accuracy results compared to other sequence to sequence

models such as RNN and LSTM in predicting sea level. The use of Transformer for forecasting time series data has also been carried out by Fasvazahra et al. [18]. In their research, they used a Transformer to predict electricity utility loads. Experimental results show that the Transformer model outperforms the LSTM model. Research conducted by Li et al. [19] in predicting electricity consumption also shows that the Transformer model provides higher accuracy values than machine learning models such as Multi-Layer Perceptron (MLP), SVM and XGBoost.

From the results of the literature review above, the use of IoT combined with artificial intelligence models can be applied to creating smart agricultural models. With a better level of accuracy than RNN, CNN, LSTM models and several machine learning models such as SVM and XGBoost, Transformer models are still very little used in IoT applications for smart agriculture. The Transformer model is very suitable for estimating data taken from sensors connected to IoT because data from sensors connected to IoT is time series data such as humidity, temperature, and soil moisture data. This paper will propose the use of Transformer models to predict data obtained from IoT-connected sensors to implement smart agriculture to increase crop yields.

The next chapter of this paper will discuss the proposed method for developing smart agriculture including the types of sensors used, data processing and transmission using IoT, and the Deep Learning Transformer model. Chapter 4 will discuss each of the devices and deep learning models that will be used in more detail, Chapter 5 will discuss the experimental results, and finally the implementation plan and conclusions will be formulated from the experimental results in the implementation and conclusion chapters.

3. Proposed Method.

We propose a smart agricultural model using IoT combined with a Transformer model to provide accurate information based on data received by IoT-connected sensors. The model we propose uses an ESP8266 Microcontroller Unit (MCU) as data acquisition whose task is to retrieve data from sensors connected to it. Sensors connected to the microcontroller will provide the data required for predictive analysis by the Transformer model. Figure 1 shows the proposed model.

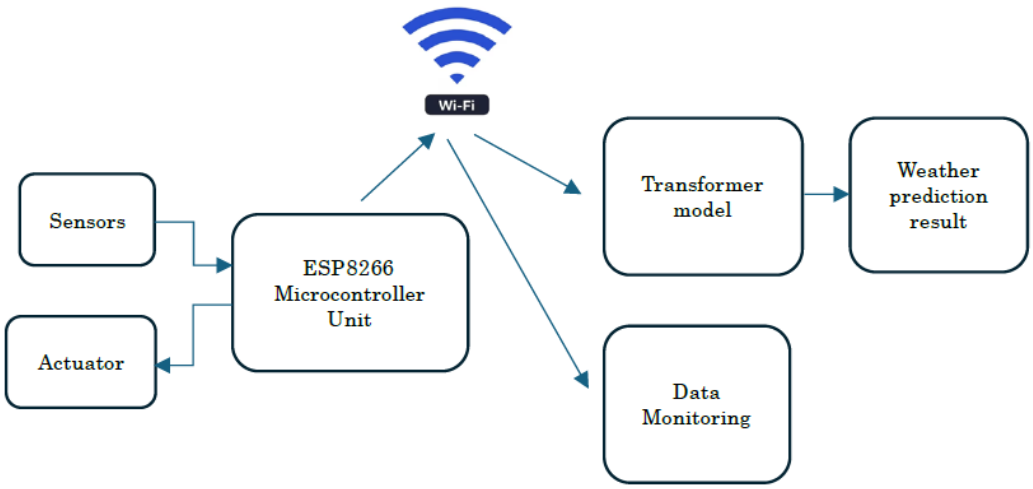


Figure 1. Proposed method

ESP8266 is a microcontroller that uses the nodeMCU platform which is an open source IoT platform and is very suitable for use as an IoT system controller because it can connect to Wi-Fi via 2.4 GHz Wi-Fi, using IEEE 802.11 BGN. The ESP8266 MCU will retrieve data from the sensors connected to it [20]. This data is processed by the ESP8266

as a basis for decision making in moving actuators such as motors or switches. All data is also sent via Wi-Fi to be displayed on the monitoring system and processed by the Transformer model for the prediction process.

The Transformer model will first be trained using a public dataset of temperature and humidity time series data. This Transformer model will continue to be trained using temperature and humidity data obtained directly through the IoT system. Through a continuous training process using real data, the accuracy level of the Transformer model in predicting temperature and humidity will increase.

4. **Technical Explanation.**

The next sub-chapters will discuss each part of the proposed method.

4.1. **Sensor and actuator.**

The sensors used in the proposed smart farming model prototype are temperature sensors, humidity sensors, and soil moisture sensors. The temperature and humidity sensor uses a DHT22 sensor which is an integrated sensor for measuring air temperature and humidity [21]. To measure soil humidity levels, the YL-69 soil moisture sensor is used [22].

DHT22 is a digital sensor that uses a voltage of 3.5V-5.5V. The resolution of this sensor is 16 bits for both temperature and humidity. The resulting measurement accuracy is $\pm 0.5^{\circ}$C for temperature and $\pm 1\%$ for humidity. The output from DHT22 is serial data with a temperature range of -40°C to 80°C and a humidity range of 0% to 100%. Figure 2 shows the DHT22 sensor. The pins on the DHT22 consist of VCC (1), Data (2) and ground (3) pins.

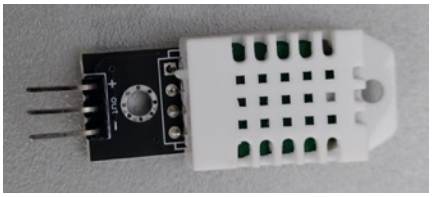


FIGURE 2. DHT22 temperature and humidity sensor

Soil moisture sensors are basically hygrometers. The sensor used to develop this smart agriculture model is the YL-69 sensor. This sensor has an operating voltage of 3.5V-5V. This sensor uses two electrodes embedded in the ground. These two electrodes will measure the resistance that occurs between the two electrodes, including due to the presence of water. In this way, this electrode will indirectly measure the water content in the soil or soil moisture. Apart from the electrodes, this sensor has an electronic part that functions to regulate the sensitivity of the measurements made. Sensitivity setting is done using a potentiometer. The output from this sensor is an analogue number 0-1023 which represents the measured soil moisture level. The YL-69 also has a digital output pin. This pin will be High and Low depending on the adjustment of the measured humidity level. This sensor has 4 pins, namely analogue output A0 (1), digital output D0 (2), ground (3) and VCC (4). Figure 3 shows the YL-69 sensor.

4.2. **Microcontroller and data acquisition.**

The ESP8266 MCU is a microcontroller device that works using the nodeMCU platform. The ESP8266 MCU module is integrated with power saving Wi-Fi IEEE 802.11 b/g/n compliance modul so it can be directly connected to a Wi-Fi network with WPA/WPA2 security protocols. With this condition, the ESP8266 MCU is very suitable for use as an IoT device [23]. The ESP8266 MCU works with a voltage of 3.3V but can produce a voltage of 5V at its digital output and operates at a frequency of 80MHz (default) to 160MHz. The ESP8266

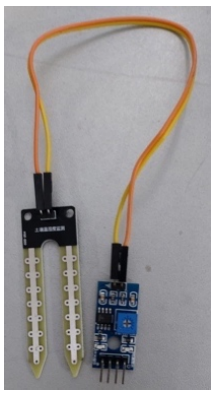


FIGURE 3. YL-69 soil moisture sensor

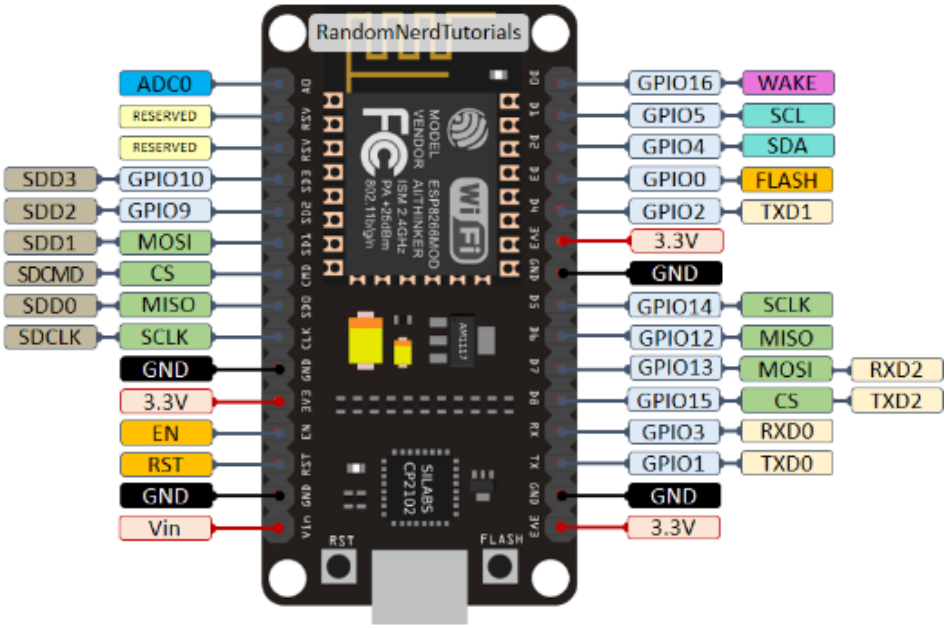


FIGURE 4. ESP8266 MCU

MCU consists of one analogue pin and 16 digital pins. For programming, the ESP8266 MCU can be programmed using the Arduino IDE including the syntax used and its libraries. Figure 4 shows the ESP8266 MCU along with its pin assignments taken from https://randomnerdtutorials.com/esp8266-pinout-reference-gpios/.

The ESP8266 MCU will retrieve data from temperature and humidity sensors, namely DHT22 and soil moisture sensors YL-69 connected to the analogue port. The ESP8266 MCU will send data via the Wi-Fi connected to it. The ESP8266 MCU will also check whether the soil moisture value is at the desired level or not. If soil moisture is below the desired level, the ESP8266 MCU will drive a servo motor connected to the digital output to open the water valve to agricultural land that needs it. If the soil moisture value matches the desired soil moisture level, the ESP8266 MCU will stop the motor so that the water flow stops as shown in Figure 5. Figure 6 shows the flowchart of the ESP8266 MCU program used.

4.3. **The monitoring system.** The monitoring system in the proposed smart farming system is used to visually display data trends obtained from the sensors used. The data that will be displayed includes data on soil moisture, air humidity and temperature, water

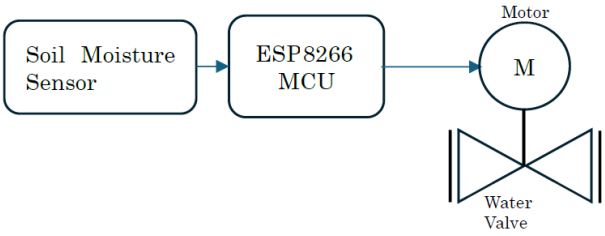


FIGURE 5. Water level monitoring system

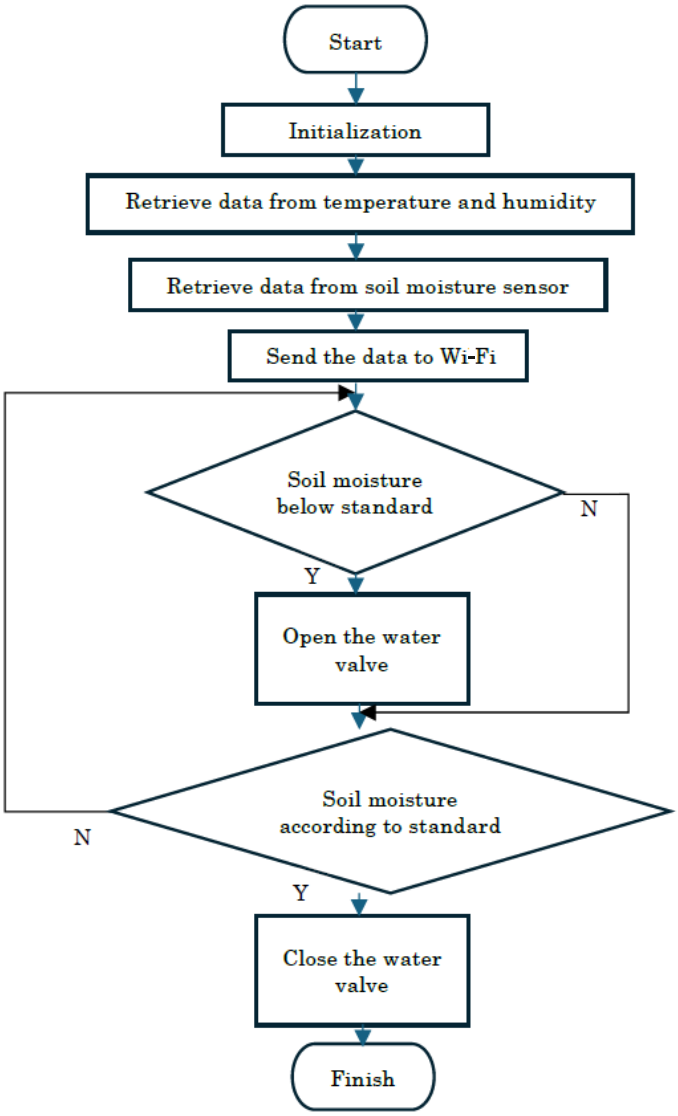


FIGURE 6. ESP8266 MCU flowchart

level and valve or sluice gate status. This display is in the form of a graph to see trends in air and ground conditions. This monitoring system uses LabVIEW software which is a virtual instrument developed by National Instrument [24].

This virtual instrument developed by LabVIEW has two interfaces for programming, namely block diagram and front panel. The block diagram consists of a function and connectivity of the instruments to be used. The front panel consists of an instrument panel and measurements that will be displayed in real time. Figure 7 shows an example

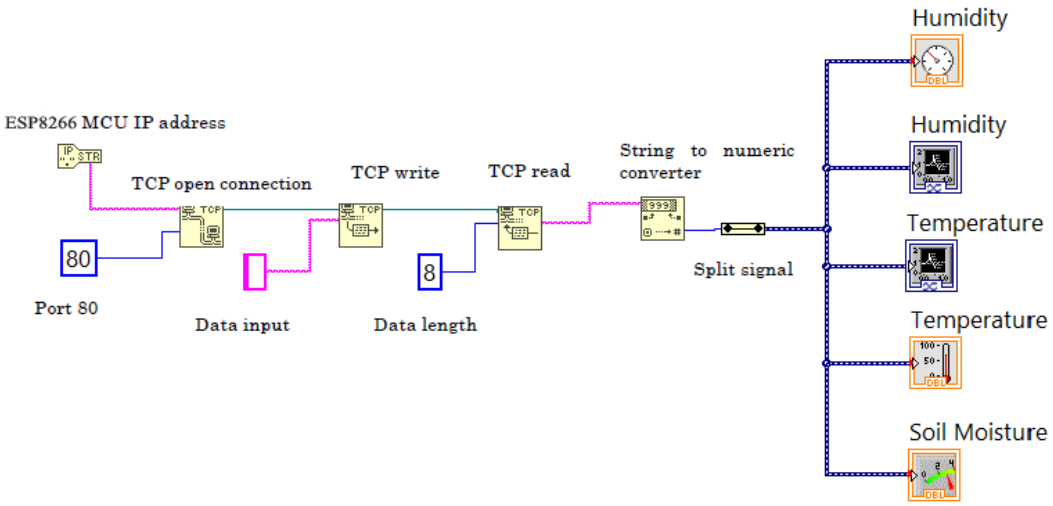


FIGURE 7. Example of a LabVIEW block diagram for smart agriculture

block diagram of the virtual instrument that will be used. The data from the sensor sent by the ESP8266 MCU via Wi-Fi is received by the LabVIEW virtual instrument using the TCP/IP port 80.

The communication process between LabVIEW and the ESP8266 MCU begins when LabVIEW opens a TCP open connection process according to the IP address and port used by the ESP8266 MCU. After the connection between LabVIEW and the ESP8266 MCU is established, the ESP8266 MCU will then send and write temperature, humidity and soil moisture data to LabVIEW via the TCP write process in LabVIEW. Then LabVIEW will read the data via TCP read with a maximum data length of 8 bytes. Next, the data in string form is converted into numeric and then broken down to be displayed on each appropriate instrument on the front panel.

The LabVIEW front panel used can be seen in Figure 8. The front panel used consists of monitoring graphs to monitor temperature and humidity history and display the current values of soil moisture, temperature, and humidity.

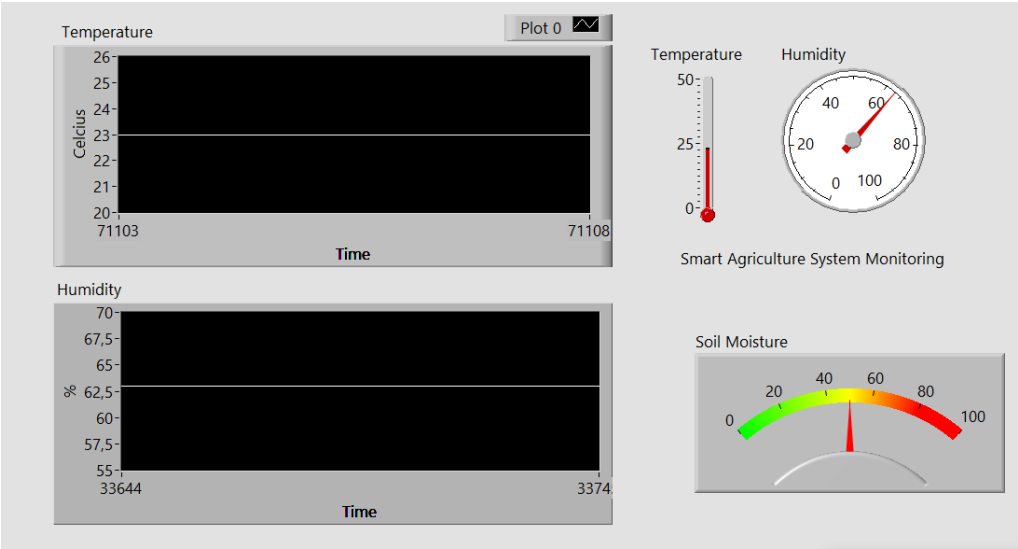


FIGURE 8. LabVIEW smart agriculture front panel

4.4. **Transformer model.** The Transformer model [16] in the proposed smart agriculture is used to predict air humidity and temperature through regression tasks. These prediction results will determine the crop yield predictions. The Transformer model consists of an encoder and decoder as shown in Figure 9. Air temperature and humidity data are received by the ESP8266 MCU and sent using the Internet to then be received and processed by the Transformer model. The Transformer model was trained using a dataset consisting of temperature and humidity data. This data is accompanied by a time label. The dataset S consisting of data on temperature T and humidity H equipped with a time

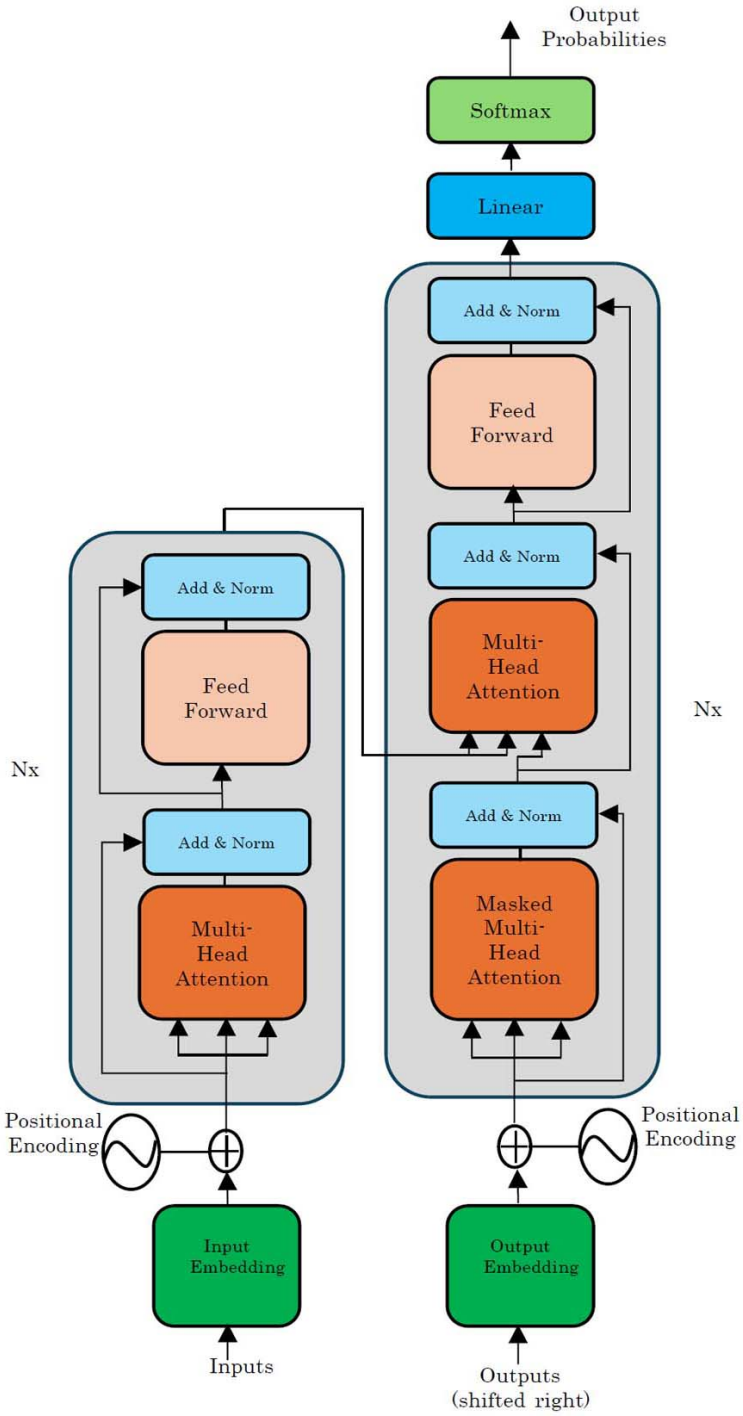


FIGURE 9. Transformer model architecture

label t can be written as $S = \{(T_1, H_1, t_1), (T_2, H_2, t_2), \ldots, (T_i, H_i, t_i)\}$. The data is then converted into a collection of vectors using the embedding process. These vectors represent the values of temperature and humidity. Each of these values can be called as a token that will be processed by the Transformer model. Because the data is time series data, each data has a close relationship with the nearest data. For this reason, the existence of Positional Encoding (PE) is very important, where positional encoding functions to provide relative positions between tokens [25]. In this Transformer model, the positional encoding used is the sine and cosine function to produce a unique vector for each token at each different position in the time series data as seen in Equations (1) and (2).

$$PE_{(pos,2i)} = \sin\left(\frac{pos}{1000^{\frac{2i}{d_{model}}}}\right) \tag{1}$$

$$PE_{(pos,2i+1)} = \cos\left(\frac{pos}{1000^{\frac{2i}{d_{model}}}}\right) \tag{2}$$

where d_{model} is the dimension of the embedding output. The next process is to give weight to each token. The weighting of each token is based on Query (Q), Key (K) and Value (V) which is calculated using a scaled dot product through an attention mechanism as shown in Equation (3). Figure 10 shows the attention mechanism process in the Transformer model.

$$Attention\,(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{3}$$

where Q, K, and V are Query, Key, and Value vectors, and d_k is the dimensions of vector K. As can be seen in Figure 10, the dot product process is first carried out for vectors Q and K via matrix multiplication (MatMul). The results of matrix multiplication are then continued with a scaling process through a division process with the dimensions of vector K. The results of the scaling process are then entered into the Softmax activation function, and a dot product process is carried out with vector V to get the attention value for each token.

Increasing the performance of this attention mechanism is carried out by parallel work of several attention mechanisms through the process of linear projection Q, K, and V

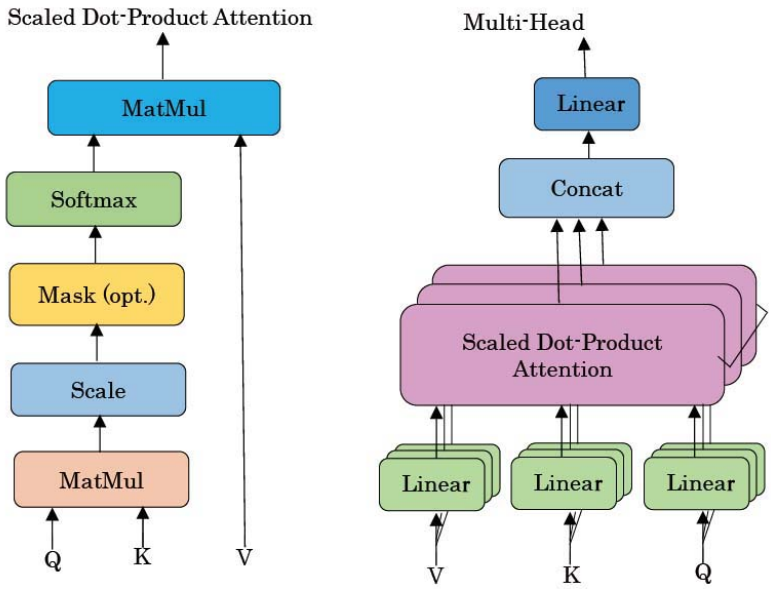


FIGURE 10. Attention mechanism

to the dimension of the vectors K (dk), K (dk), and V (dv) h times. Several attention mechanisms are concatenated to form Multi-Head Attention as shown in Figure 10 and can be calculated using Equations (4) and (5).

$$MultiHead\left(Q,K,V\right) = Concat\left(head_1, head_2, \ldots, head_h\right)W^O \tag{4}$$

$$head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right) \tag{5}$$

where $W_i^Q \in \mathbb{R}^{d_{model} \times d_k}$, $W_i^K \in \mathbb{R}^{d_{model} \times d_k}$, $W_i^V \in \mathbb{R}^{d_{model} \times d_v}$, $W^O \in \mathbb{R}^{hd_v \times d_{model}}$, W is the weight of vectors Q, K, and V obtained from the results of the Softmax function, and h refers to the number of heads used.

After going through the attention process, the next step is to carry out the normalization process on the Layer Normalization. This Layer Normalization is found in the encoder and decoder. The function of the Layer Normalization is to stabilize the training process by adjusting the gradient scale through standardizing the input value scale. In the Layer Normalization, there is also a process of adding up the Multi-Head Attention output with the original input from Multi-Head Attention which forms a residual connection. Through this process, values that are outside a certain scale can be adjusted to the desired scale so that the training process becomes more effective and efficient and produces training results with smoother gradients [26].

From the decoder side, the target label will enter the decoder which is shifted right and goes through the same process as the encoder. The final process of this Transformer is the process of predicting temperature and humidity values based on the training process using a dataset that has time labels. This prediction process uses a feed forward neural network with a Softmax activation function which can be calculated using Equation (6).

$$FeedForward\left(T,H,t\right) = softmax\left(0, (T,H,t)W_1 + b_1\right)W_2 + b_2 \tag{6}$$

where b_1 and b_2 are bias.

In general, the Transformer algorithm used in this model can be seen in Algorithm 1. When the program is first run, the Transformer model will load the dataset used. Next, pre-processing will be carried out on the dataset in the form of checking whether all the data is numeric and there are no empty fields. The dataset is then split into three parts, namely for the training, validation, and testing processes. The training process takes up 90% of the dataset and the validation process takes up 10% of the dataset. At this data pre-processing stage, the dataset scale is also changed using logarithmic return (log-return) to convert it into a relative value within a certain period. This change makes data distribution normally distributed and smoother. The goal of this change is to make the forecasting process more accurate.

After the split dataset process is complete, the next step is to set the hyper parameters of the Transformer model used. The hyper parameter settings that are carried out include setting the number of heads, the number of Transformer layers, and the number of neural network layers used. In addition, the batch size used for the training process, learning rate, number of epochs, activation function and loss function are also set.

After setting the hyper parameters of the Transformer model that will be used, the training and testing process is carried out for the entire dataset. In this training process, the weight metric parameters of the Transformer model are also updated. After the training process ends, the Transformer model has weight metric parameters that can generalize new data and can predict temperature and humidity according to time labels.

5. **Result and Discussion.** The proposed Transformer model was trained using 3,910 data records consisting of temperature, humidity, and time data. Dataset taken from https://www.kaggle.com/datasets/greegtitan/indonesia-climate. This dataset contains

Algorithm 1. Transformer model

load Temperature, humidity dataset
for each Temperature, humidity dataset
    delete empty records or field
    delete non-numeric data
    convert data using log-return
set train_size to 0.9
set val_size to 0.1
split dataset into training and validation according to train_size and val_size
set number of head to 3, number of layer to 2, number of dense to 256
set batch size to 250, learning rate to 0.0005, epoch to 50, activation function to Softmax, Loss function to Sparse Categorical Cross Entropy
for each record in dataset, **do** feature extraction using multi-head attention
  for epoch = 1: number of epochs
    for batch = 1: number of batches
      Generate another batch
      Train the model
      Validation the model
      Backpropagate the loss
update weight metric parameter
generate Temperature and humidity prediction
test the model

temperature, humidity, and rainfall data from 2010 to 2020 for several cities in Indonesia. In this experiment, temperature and humidity data were taken from the city of Bandung-West Java. Figure 11 shows the dataset of temperature graph and Figure 12 shows the dataset of humidity graph. The dataset was divided into 90% for training and 10% for validation. The training process used Google Colaboratory with VT100 GPU mode. In this experiment, the number of heads used was 3, 2 Transformer layers, 128 hidden units in the feed forward artificial neural network layer, 250 batch sizes and a learning rate of 0.0005. The time required for the training process using 50 epochs took 2 minutes.

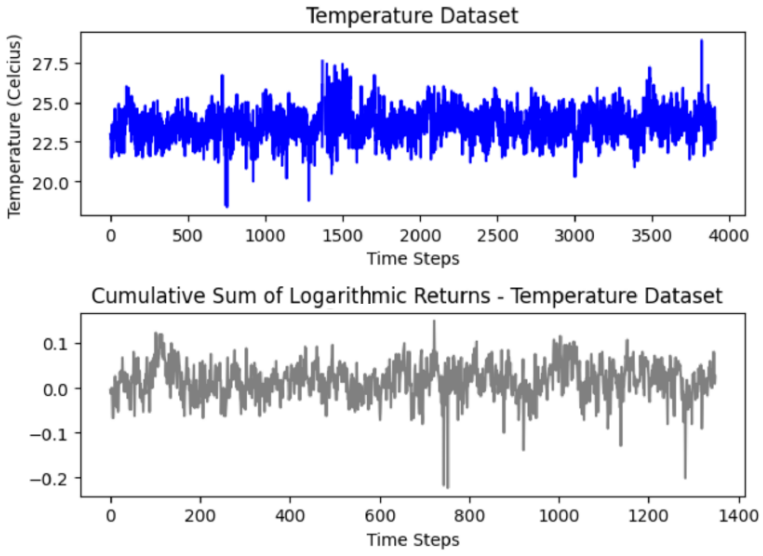


FIGURE 11. Temperature dataset

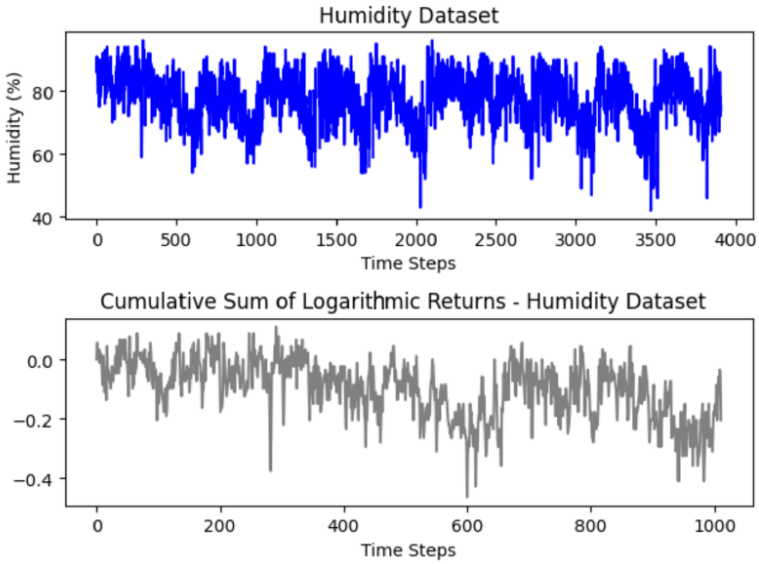


FIGURE 12. Humidity dataset

Figures 13 and 14 show the results of training the Transformer model to predict temperature and humidity. From Figure 13, the loss value resulting from the temperature prediction model validation process is 0.043, which is close to the loss value from the training process. Likewise in Figure 14, the loss value resulting from the validation process of the air humidity prediction model is 0.051, which is close to the loss value from the training process. This condition shows that the proposed Transformer model can be learned well using this dataset and there is no overfitting condition. An overfitting condition can occur if the loss value from the validation results is further away from the loss value from the training results [27]. In this way, the developed Transformer model can generalize and predict new data well [28].

Figure 15 shows the temperature prediction results using test data derived from 10% of the dataset. The prediction results in the form of a red graph look close to the black graph which is the actual temperature value. A similar thing can also be seen in Figure 16, where the red graph which is the result of the humidity prediction is close to the black graph which is the actual humidity value.

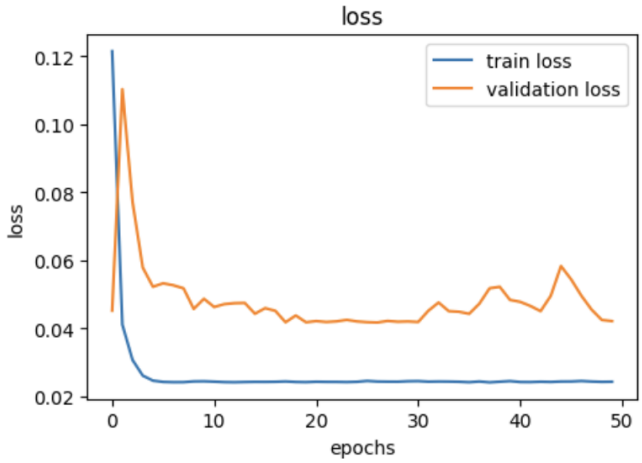


FIGURE 13. Temperature training result

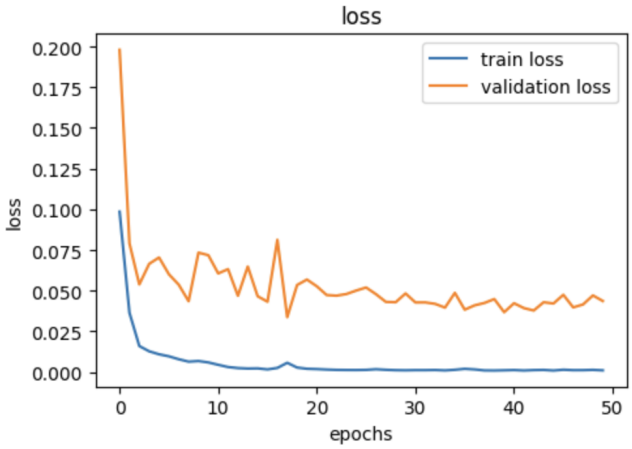


FIGURE 14. Humidity training result

Apart from using loss calculations from the training and validation process, the accuracy measurement process was also carried out using Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE) which are calculated using Equations (7)-(9).

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \tag{7}$$

$$MSE = \frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n} \tag{8}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}} \tag{9}$$

where y is the actual temperature and humidity value and x is the predicted temperature and humidity value. The calculation of MAE, MSE and RMSE values using 10% of the dataset for testing is used to create prediction graphs in Figures 15 and 16. To see the level of accuracy of the proposed Transformer model, a comparison of accuracy was carried out

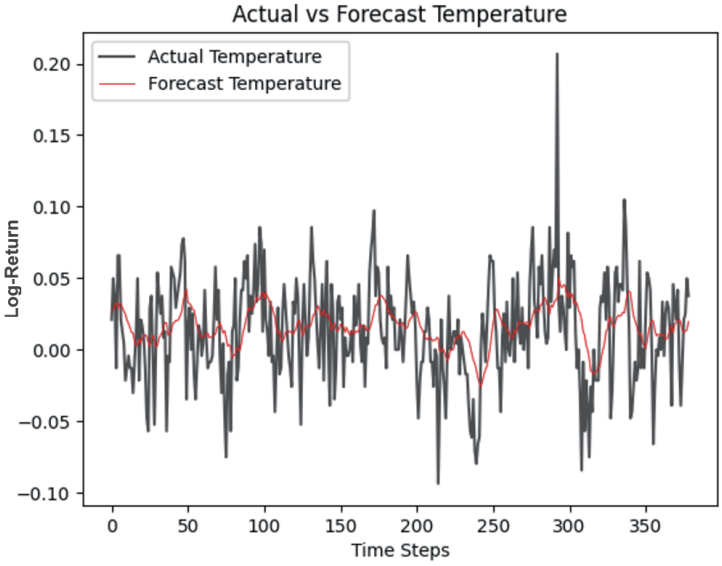


FIGURE 15. Temperature forecast

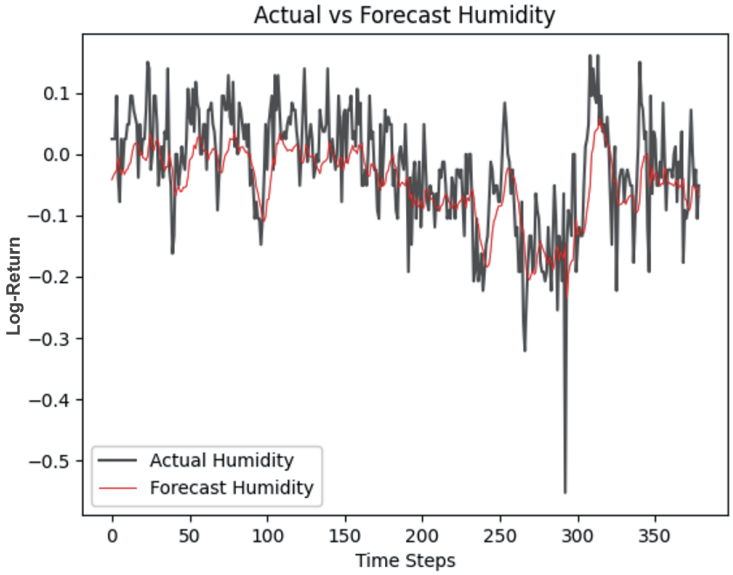


FIGURE 16. Humidity forecast

TABLE 1. Experimental result

Model		MAE	MSE	RMSE
Temperature	Transformer	**0.027**	**0.0011**	**0.034**
	SVR	0.032	0.0016	0.039
	XGBoost Regression	0.029	0.0014	0.037
Humidity	Transformer	**0.057**	**0.0053**	**0.073**
	SVR	0.063	0.0057	0.079
	XGBoost Regression	0.063	0.0055	0.077

using machine learning models, namely Support Vector Regression (SVR) and XGBoost Regression using the same dataset. The MAE, MSE and RMSE results for temperature and humidity using the Transformer, SVR and XGBoost models can be seen in Table 1.

From Table 1, the MAE, MSE and RMSE results from the Transformer model are smaller than those from the machine learning models, namely SVR and XGBoost Regression, both for temperature prediction and humidity prediction. This shows that the Transformer model has better accuracy compared to the SVR and XGBoost Regression models. The accuracy level of the Transformer model can outperform the SVR and XGBoost models because the Transformer model is very suitable for predicting time series data by paying attention to the closest data during the training process.

6. **Implementation.** Based on the results of the experiment above, this model can be implemented directly to farmers. Farmers who have land in an area can estimate the temperature and humidity for some time in the future. The results of the estimated temperature and humidity can be used to determine which plants are suitable for planting. For example, agricultural land in the Lembang – Bandung area, West Java – Indonesia, according to the forecasting results using the Transformer model for the following year will have a temperature and humidity ranging from 17°C-20°C and humidity between 80%-90%. Based on these data, farmers can determine which plants are suitable for planting. For example, strawberries are plants that are suitable for these conditions.

Daily operations of agricultural land require a long process [29]. By using this system, daily operations become simpler. The monitoring process is carried out using the

LabVIEW dashboard. This dashboard will display actual and historical temperature and humidity data. This monitoring is very important for predicting crop yields where changes in temperature and humidity greatly affect crop yields. Soil moisture control can also be done using an actuator in the form of a water valve drive motor. If the monitoring results show that soil moisture is below standard, the drive motor will open the water valve. If the soil moisture is in accordance with the needs, the drive motor will close the water valve. The use of this system can provide the right humidity value according to plant needs. With this system, it is hoped that the harvest results can be maximized.

7. **Conclusions.** Based on the experimental result, the proposed model has better accuracy and smaller prediction metric error compared with the comparison model. The monitoring of soil moisture, temperature and air humidity can be done effectively via IoT and LabVIEW virtual instrument. The irrigation process on agricultural land can be carried out more precision according to the needs and conditions of the soil. The Transformer model can be used with a good level of accuracy to predict temperature and humidity. The results of this prediction are very important for predicting agricultural yields and determining the type of plants to be planted so that they suit the temperature and humidity level requirements and can produce maximum agricultural production. The proposed model for smart agriculture using IoT and Transformer model can support the sustainability of the agricultural production.

The Transformer model training process can be carried out periodically using temperature and humidity data taken through the sensors used so that it will provide prediction results regarding temperature and humidity for the next period.

The main limitation in the implementation model on agricultural land is the lack of telecommunications infrastructure needed by IoT such as Internet and Wi-Fi networks. In addition, large agricultural land requires many sensors where the data obtained must be grouped and processed into several areas to be effective in the automation process. The wider the area of agricultural land, the more complicated the process of collecting data from sensors will be and requires special handling.

Data Availability Statement. The dataset used in this study uses a public dataset that can be accessed at the link https://www.kaggle.com/datasets/greegtitan/indonesia-climate.

Open Contributor. Ilvico Sonata conducted the design and experimentation of IoT and Transformer models. Yulyani Arifin conducted the research on SDG and the overall review of the paper.

REFERENCES

[1] M. Dhanaraju, P. Chenniappan, K. Ramalingam, S. Pazhanivelan and R. Kaliaperumal, Smart farming: Internet of Things (IoT)-based sustainable agriculture, *Agriculture (Switzerland)*, vol.12, no.10, DOI: 10.3390/agriculture12101745, 2022.

[2] Y. Abbassi and H. Benlahmer, The Internet of Things at the service of tomorrow's agriculture, *Procedia Computer Science*, pp.475-480, DOI: 10.1016/j.procs.2021.07.060, 2021.

[3] S. Mandal, A. Yadav, F. A. Panme, K. M. Devi and S. M. S. Kumar, Adaption of smart applications in agriculture to enhance production, *Smart Agricultural Technology*, vol.7, DOI: 10.1016/j.atech.2024.100431, 2024.

[4] K. Takács-György and I. Takács, Towards climate smart agriculture: How does innovation meet sustainability?, *Ecocycles*, vol.8, no.1, pp.61-72, DOI: 10.19040/ECOCYCLES.V8I1.220, 2022.

[5] D. Burra, J. Hildebrand, J. Giles, T. Nguyen, E. Hasiner, K. Schroeder, D. Treguer, A. Juergenliemk, A. Horst, A. Jarvis and W. Kropff, *Digital Agriculture Profile: Viet Nam*, Report, FAO, Rome, Italy, 2022.

[6] S. Qazi, B. A. Khawaja and Q. U. Farooq, IoT-equipped and AI-enabled next generation smart agriculture: A critical review, current challenges and future trends, *IEEE Access*, vol.10, pp.21219-21235, DOI: 10.1109/ACCESS.2022.3152544, 2022.

[7] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour and E. H. M. Aggoune, Internet-of-Things (IoT)-based smart agriculture: Toward making the fields talk, *IEEE Access*, vol.7, pp.129551-129583, DOI: 10.1109/ACCESS.2019.2932609, 2019.

[8] R. V. Saraswathi, J. Sridharani, P. S. Chowdary, K. Nikhil, M. S. Harshitha and K. M. Sai, Smart farming: The IoT based future agriculture, *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, pp.150-155, DOI: 10.1109/ICSSIT53264.2022.9716331, 2022.

[9] S. R. Prathibha, A. Hongal and M. P. Jyothi, IoT based monitoring system in smart agriculture, *2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT)*, Bangalore, India, pp.81-84, DOI: 10.1109/ICRAECT.2017.52, 2017.

[10] C. Sagana, P. Keerthika, M. Thangatamilan, R. Kamali, K. Nanthini and S. Maghathani, Identification of suitable crop based on weather condition, *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, DOI: 10.1109/ICCCI54379.2022.9740905, 2022.

[11] G. Sanjana, N. M. Davasam and N. M. Krishna, Smart farming using IoT and machine learning techniques, *2020 IEEE Bangalore Humanitarian Technology Conference (B-HTC)*, Vijiyapur, India, DOI: 10.1109/B-HTC50970.2020.9297916, 2020.

[12] K. S. P. Reddy, Y. M. Roopa, K. Rajeev L.N. and N. S. Nandan, IoT based smart agriculture using machine learning, *2020 2nd International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, pp.130-134, DOI: 10.1109/ICIRCA48905.2020.9183373, 2020.

[13] N. Zhu et al., Deep learning for smart agriculture: Concepts, tools, applications, and opportunities, *International Journal of Agricultural and Biological Engineering*, vol.11, no.4, pp.21-28, DOI: 10.25165/j.ijabe.20181104.4475, 2018.

[14] M. Altalak, M. A. Uddin, A. Alajmi and A. Rizg, Smart agriculture applications using deep learning technologies: A survey, *Applied Sciences (Switzerland)*, vol.12, no.12, DOI: 10.3390/app12125919, 2022.

[15] S. K. S. Durai and M. D. Shamili, Smart farming using machine learning and deep learning techniques, *Decision Analytics Journal*, vol.3, 100041, DOI: 10.1016/j.dajour.2022.100041, 2022.

[16] A. Vaswani et al., Attention is all you need, *arXiv Preprint*, arXiv: 1706.03762, 2023.

[17] R. N. Sabililah and D. Adytia, Time series forecasting of sea level by using Transformer approach, with a case study in Pangandaran, Indonesia, *2023 IEEE 8th International Conference for Convergence in Technology (I2CT 2023)*, DOI: 10.1109/I2CT57861.2023.10126216, 2023.

[18] I. A. Fasvazahra, D. Adytia and A. A. Simaremare, Electricity time series forecasting by using Transformer with case study in Jakarta Banten, *2022 5th International Conference on Computer and Informatics Engineering (IC2IE 2022)*, pp.167-172, DOI: 10.1109/IC2IE56416.2022.9970104, 2022.

[19] X. Li, Y. Zhong, W. Shang, X. Zhang, B. Shan and X. Wang, Total electricity consumption forecasting based on Transformer time series models, *Procedia Computer Science*, pp.312-320, DOI: 10.1016/j.procs.2022.11.180, 2022.

[20] M. Kashyap, V. Sharma and N. Gupta, Taking MQTT and NodeMcu to IOT: Communication in Internet of Things, *Procedia Computer Science*, pp.1611-1618, DOI: 10.1016/j.procs.2018.05.126, 2018.

[21] M. J. Alam, S. A. Rafi, A. A. Badhan, M. N. Islam, S. I. Shuvo and A. M. Saleque, Low cost IoT based weather station for real-time monitoring, *2020 IEEE 2nd International Conference on Circuits and Systems (ICCS2020)*, pp.127-133, DOI: 10.1109/ICCS51219.2020.9336596, 2020.

[22] W. Y. Tan, Y. L. Then, Y. L. Lew and F. S. Tay, Newly calibrated analytical models for soil moisture content and pH value by low-cost YL-69 hygrometer sensor, *Measurement (Lond)*, vol.134, pp.166-178, DOI: 10.1016/j.measurement.2018.10.071, 2019.

[23] J. Mesquita, D. Guimarães, C. Pereira, F. Santos, L. Almeida and L. Almeida, Assessing the ESP-8266 WiFi module for the Internet of Things, *2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA)*, Turin, Italy, pp.784-791, DOI: 10.1109/ETFA.2018.8502562, 2018.

[24] K. Kishkin, B. Ganev, D. Arnaudov and M. B. Marinov, LabVIEW based system for determining energy flows in energy storage system, *Proc. of 2021 30th International Scientific Conference Electronics*, DOI: 10.1109/ET52713.2021.9579768, 2021.
[25] Q. Zhang et al., An empirical study on the impact of positional encoding in Transformer-based monaural speech enhancement, *2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp.1001-1005, DOI: 10.1109/ICASSP48485.2024.10446337, 2024.
[26] R. Xiong et al., On layer normalization in the Transformer architecture, *arXiv Preprint*, arXiv: 20 02.04745, 2020.
[27] A. B. Shanmugavel, V. Ellappan, A. Mahendran, M. Subramanian, R. Lakshmanan and M. Mazzara, A novel ensemble based reduced overfitting model with convolutional neural network for traffic sign recognition system, *Electronics (Switzerland)*, vol.12, no.4, DOI: 10.3390/electronics12040926, 2023.
[28] C. M. Gevaert and M. Belgiu, Assessing the generalization capability of deep learning networks for aerial image classification using landscape metrics, *International Journal of Applied Earth Observation and Geoinformation*, vol.114, DOI: 10.1016/j.jag.2022.103054, 2022.
[29] D. Wang, H. Shi, L. Sun and Y. Du, A collaborative allocation approach for sorting resources in rural cold chain warehouses, *International Journal of Innovative Computing, Information and Control*, vol.19, no.4, pp.1221-1238, DOI: 10.24507/ijicic.19.04.1221, 2023.

Author Biography

Ilvico Sonata S.T., M.M., MT., MBA, has experience in the Satellite Telecommunications industry for more than 27 years. He has a passion for teaching and is currently a lecturer at Doctor of Computer Science Program, BINUS University. He has scientific expertise in the fields of artificial intelligence, electrical engineering, autonomous systems and computer vision. He also teaches in Bachelor's and Master's programs. Some of the courses taught include Advanced Softcomputing and Internet of Things. He has professional certification in the fields of Satellite Telecommunications and Data Science. He is also active in writing several conference papers and Scopus indexed journals.

Yulyani Arifin S.Kom, M.M., currently serves as Deputy Head of Doctor of Computer Science, BINUS University. She completed her Doctoral Education in 2021 at Doctoral of Computer Science, BINUS University and studied Master of Management specializing in Information Systems in 2005 and Bachelor of Information Management in 1998. In 2022, she completed his Professional Engineer degree. Before focusing on academics, she had 11 years of professional experience in the ERP field. She is also active in the Indonesian Chapter of the ACM-SIGCHI organization and is a member of the international organization ISACA. Since 1998, she has been actively teaching until now. The knowledge groups taught and research topics include human-computer interaction, natural language processing, text processing, multimedia and game virtual reality, augmented reality and mixed reality, IT risk management and software engineering.

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

Enhancing soil moisture prediction with explainable AI: Integrating IoT and multi-sensor remote sensing data through soft computing / Mallik, S., Chakraborty, A., Podder, K., Talukdar, S., Rahman, A., & Mishra, U.

Applied Soft Computing
Volume 180 (2025) 113406 Pages 1-19
https://doi.org/10.1016/j.asoc.2025.113406
(Database: ScienceDirect)

Applied Soft Computing Journal 180 (2025) 113406

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Enhancing soil moisture prediction with explainable AI: Integrating IoT and multi-sensor remote sensing data through soft computing

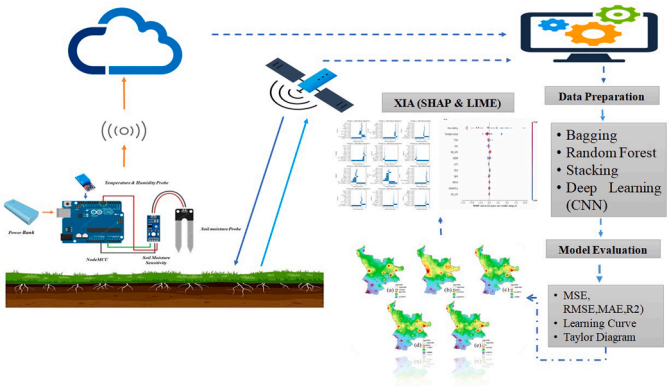
Santanu Mallik [a,e,*], Abhigyan Chakraborty [b,g], Krishanu Podder [c], Swapan Talukdar [d,f], Atiqur Rahman [d], Umesh Mishra [e]

[a] Department of Civil Engineering, Poornima College of Engineering,Sitapura, Jaipur. Rajasthan 302022, India
[b] Department of Space, North Eastern Space Application Center, Goverment of India, Umiam, Meghalaya 93103, India
[c] Department of Elementary Education, Goverment of Tripura, Agartala, India
[d] Department of Geography, Faculty of Sciences, Jamia Millia Islamia, New Delhi 110025, India
[e] Department of Civil Engineering, National Institution of Technology Agartala, Barjala, Jirania, Tripura 799046, India
[f] Department of Geography, Asutosh College, University of Calcutta, Kolkata-700026
[g] Department of Civil Engineering, Indian Institute of Technology Hyderabad, Sangareddy, Telangana, 502285 India

HIGHLIGHTS

- A framework integrating IoT and AI models was developed for accurate prediction of surface soil moisture.
- Stacking model achieved superior soil moisture prediction (RMSE: 0.574, R^2: 0.996)
- Spatially predicted soil moisture ranged from 6.45% to 68.27%
- XAI indicated humidity and atmospheric temperature as key soil moisture predictors

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Soil Moisture
Internet of Things
Remote sensing
Deep Learning
Explainable artificial intelligence

ABSTRACT

Soil moisture (SM) is a crucial variable for comprehending various ecosystem processes on Earth. However, the collection of SM and other field data is often burdensome and fails to accurately represent local spatial variability. In order to address this contemporary challenge, a framework has been developed to predict SM content at a spatial scale by amalgamating Internet of Things (IoT) and geospatial computing approaches. Additionally, this study aimed to investigate the behaviour and contribution of explanatory variables in predicting SM content using explainable AI (XAI). Field data acquired through sensors and IoT, along with proxy variables derived from Landsat 9, Sentinel-1A data, and terrain data were utilized to develop models at a spatial scale. The Boruta algorithm was employed as a feature selection method, followed by the use of three ensemble machine learning models (random forest, bagging, and stacking) and a convolutional neural network (CNN) model to predict soil moisture based on the selected parameters. Statistical performance indicators such as mean square error (MSE),

* Corresponding author at: Department of Civil Engineering, Poornima College of Engineering,Sitapura, Jaipur, Rajasthan 302022, India.
E-mail address: coolshan02@gmail.com (S. Mallik).

https://doi.org/10.1016/j.asoc.2025.113406
Received 6 July 2023; Received in revised form 4 January 2025; Accepted 27 May 2025
Available online 5 June 2025

root mean square error (RMSE), mean absolute error (MAE), and correlation coefficient (R^2) were used to validate all the models. Furthermore, XAI based SHapley Additive exPlanations (SHAP) and Local Interpretable Model Explanation (LIME) method is used for global and local site-specific sensitivity analysis. The validation results indicate that the stacking model outperforms all other models, yielding an MSE of 0.330, RMSE of 0.574, MAE of 0.233, and R^2 of 0.996. The spatial map of SM content reveals a range varying between 6.45 % and 68.27 %. The outputs of XAI models highlights that humidity and atmospheric temperature have the highest influence on predicting SM content both globally and locally. Overall, the proposed framework holds promise as an alternative to traditional approaches for estimating surface SM content. The obtained spatial map can prove useful in sustainable water resources management at the field scale.

1. Introduction

Soil moisture (SM) content is a crucial variable in understanding the global water, energy, and carbon cycles, and it plays a vital role in studying the Earth's climate and managing water resources [1,2]. SM content represents the amount of water held in the soil pores at a specific matrix potential, typically focusing on the top 5 cm of soil surface [3]. It serves as a "memory of precipitation" and is widely used as a proxy indicator for disaster prevention, meteorological and agricultural applications, and decision-support tools [2,4–7]. Consequently, soil moisture is considered an essential climate variable, and efforts have been made to improve soil moisture estimation.

Traditionally, laboratory-based direct or indirect methods have been used to estimate soil moisture [8]. However, these methods are prone to human error, time-consuming, labor-intensive, and may not always be practical [9]. Recent advancements in sensor technology and the Internet of Things (IoT) have led to more efficient, cost-effective, and resourceful methods of collecting time-series data compared to traditional laboratory-based methods. Previous studies have demonstrated the application of open-source technology-based systems for smart irrigation and precision farming, utilizing sensors to measure soil moisture, soil temperature, and environmental conditions along with weather forecast data [10–13]. Kumar and Pallikonda Rajasekaran, (2016) and Puengsungwan, (2020) have developed low-cost IoT-based sensors for collecting soil moisture data, validating their accuracy against conventional laboratory-based methods [13,14]. Sensors such as time domain reflectometry (TDR), time domain transmission (TDT), capacitance probe, and resistance probe are commonly used for soil moisture estimation [15,16]. However, the selection of appropriate sensors and the preparation of spatiotemporal data at various scales using IoT may not always be a cost-effective approach.

In light of technological advancements, remote sensing has emerged as a viable alternative for soil moisture estimation. Remote sensing allows for continuous data collection with higher spatial and temporal coverage, making it effective for mapping land surface soil moisture using optical, thermal, and microwave data from active and passive earth observation remote sensing [4,5,17,18–20]. However, the accuracy of remote sensing-based soil moisture retrieval can be influenced by factors like cloud cover, surface roughness, and vegetation [21,22]. To address these challenges, some researchers have combined active and passive Soil Moisture (SMAP) data to improve retrieval accuracy [23]. However, the limited spatial resolution of SMAP data restricts its usage in regional or local studies [24]. Additionally, handling the increasing amount of satellite observation data solely with theoretical and empirical models may not yield satisfactory results.

To overcome these limitations, data-driven models, particularly those based on artificial intelligence (AI), offer significant advantages over traditional approaches by enabling objective and purely data-driven retrieval models that integrate multiple data sources. Machine learning (ML) methods such as Multivariate Linear Regression, Artificial Neural Networks (ANN), Support Vector Machine (SVM), and Random Forest (RF) have been applied to estimate soil moisture [25–27]. Advanced ML models like ensemble, stacking, bagging, boosting, and averaging methods have gained popularity for predicting soil moisture and other soil parameters (Zounemat-Kermani et al., 2021; Zhang et al., 2022). Ensemble models have shown improved prediction performance in various studies [17,28,29].

However, the selection of an appropriate algorithm from numerous advanced ML models is a non-trivial task that requires experience due to the unique characteristics and challenges of each algorithm. Hence, considering the limitations of standalone and advanced ML models, deep learning (DL) models with self-learning and adaptability capabilities could offer a viable alternative. DL models, with their stacked non-linear layers, can capture complex relationships between variables [29,30]. Studies in various domains, such as water quality prediction, irrigation water suitability, groundwater levels, crop yield, and soil organic carbon, have demonstrated the efficacy of DL-based models over standalone and hybrid ML models [29–32]. Among the DL paradigms, Convolutional Neural Networks (CNNs) have shown great potential in emulating various parameters, including soil moisture content [33–36]. However, limited research has been conducted on the efficacy of CNN models for soil moisture prediction.

Although, the studies discussed above used various ML, DL or ensemble model are considered to be black box model as they lag model explainability. Explainable Artificial Intelligence (XAI) techniques, such as Shapley Additive Explanations (SHAP), have shown promise in remote sensing-based geoscience research. SHAP values provide valuable insights into the contribution of individual features, enhancing land use and land cover mapping [37]. Similarly, study reported by Jena et al., (2023) and Pradhan et al., (2023) used XAI based SHAP method to analyze the contribution and behavior of different features in earthquake spatial probability assessment, landslide susceptibility modeling aiding in proposing management strategies[38,39]. On the other hand XAI based Local Interpretable Model Explanation (LIME) method is used for explain the outcomes of the complex black box model for the local/site specific condition [40]. However, there is limited research available on the application of explainable Artificial Intelligence (XAI) techniques to soil science research, particularly in the context of soil moisture content and nutrient content. Further investigation is required to explore the potential of XAI techniques, such as SHAP values for global analysis and LIME for local and site-specific analysis, in soil science research. This research can play a crucial role in enhancing predictions and providing valuable insights into soil property dynamics, ultimately benefiting agriculture and environmental management. Given the existing gaps in the literature, this study aims to address the lack of comprehensive research on cutting-edge techniques for predicting surface soil moisture content.

The overall objective of this study is multifaceted. Firstly, it aims to develop a framework that integrates IoT data, multi-sensor remote sensing, and terrain information using ML and DL models. This framework will enable the mapping and prediction of soil moisture. Additionally, an optimized bagging ensemble model-based XAI approach will be developed to analyze the impact of variables on soil moisture prediction and assess its sensitivity at both the global and site-specific levels. Collectively, the proposed study has the potential to make significant contributions to the field of soil moisture prediction and sustainable development. By employing state-of-the-art techniques, including ensemble models, DL, XAI, and leveraging integrated IoT data

and remote sensing information, this study aims to advance our understanding of soil moisture dynamics and contribute to more effective and sustainable land management practices.

The proposed study is novel in several aspects. Firstly, the extensive use of integrated IoT data for collecting base data to predict surface soil moisture content has not been extensively explored in the literature. This approach offers advantages in terms of accuracy, speed, and cost-effectiveness over traditional data collection methods like manual measurements. By leveraging IoT data, the study can capture real-time and continuous environmental information, enabling more dynamic and accurate predictions of soil moisture content. Secondly, integration of multiple data sources, including integrated IoT data, remote sensing data, terrain, and meteorological parameters, using ensemble ML and DL models, is a unique approach to predicting soil moisture content. By combining various data types, the study can capture the complex interactions and dependencies that influence soil moisture dynamics. This integrated approach can lead to more accurate predictions, enabling better water resource management, sustainable agricultural practices, and informed environmental monitoring. Lastly, the application of XAI techniques, specifically SHAP and LIME to analyze the impact of variables on soil moisture prediction is a novel approach. Existing studies have primarily focused on the accuracy of ML and DL models in predicting soil moisture content, neglecting the examination of variable influence. This novel approach provides insights into the sensitivity of different variables and their contributions to soil moisture variations, aiding in the development of effective management strategies. The insights gained from this research can benefit stakeholders in agriculture, water resource management, and environmental conservation by providing actionable information for optimizing soil moisture levels and enhancing ecological sustainability.

The present article is organized as follows: Section 2 provides a comprehensive description of the material and method considering various important aspects. All the obtained results are presented and discussed in Section 3 and Section 4 respectively. Finally, Section 5 encapsulates the key conclusions derived from the study.

2. Material and method

2.1. Study area selection and rationale for research

This article explores the essential considerations involved in selecting a study area and the rationale that guides the research. The meticulous and thoughtful selection of a study area is crucial for ensuring the credibility, significance, and applicability of research findings. The study was conducted in Agartala city, which serves as the gateway to the northeastern state of India. Spanning a surface area of 76.5 km^2, the city is located within the longitude range of 91°15'- 91°20' E and the latitude range of 23°45'- 23°55'. Agartala experiences a subtropical climate characterized by high temperatures, humidity, and significant rainfall, as documented by the Indian Meteorological Department (IMD), Government of India. The soil composition in Agartala City predominantly falls under the classification of Sandy clay loam, with organic carbon content ranging from 0 % to 2.87 % [41].

The people of the region rely heavily on agriculture for their livelihood, underscoring the crucial importance of accurate soil moisture prediction and estimation. Such information provides valuable insights into the water content of agricultural fields, which is essential for ensuring optimal crop growth and yield. Moreover, this information holds significant implications for various state government agencies, including the water resources department, department of Agriculture, and farmer welfare, as it enables them to optimize water use efficiency, minimize water wastage, and maximize crop yields. Surprisingly, no previous studies have been reported to date from this study area focusing on soil moisture estimation and prediction utilizing geo-computational tools.

2.2. IoT based field data collection

The term IoT, short for the Internet of Things, refers to the integration of physical objects embedded with sensors, software, and processing capabilities that enable them to interact and share information with other systems and devices via the Internet or other communication networks [42]. Recent advancements in IoT technology have garnered significant attention from researchers worldwide. [14,43–45] have utilized IoT for collecting field parameters such as soil moisture, temperature, and humidity. These studies have demonstrated that sensor-based data collection for soil moisture is accurate, cost-effective, and offers ease of repeatability. Hence, in this study, the authors successfully developed an IoT-based platform to collect field data on soil moisture, temperature, and humidity from eighty-four (84) different locations using a random sampling technique within the Agartala municipality area, as illustrated in Fig. 1.

The architectural overview of the proposed IoT system is depicted in Fig. 2. A fork-shaped soil moisture probe is employed to measure soil moisture content, while an LM393 comparator connected to the soil moisture probe is utilized for comparing soil moisture levels. Additionally, a DHT11 sensor module is employed to measure temperature and humidity [46]. Within this system, the NodeMCU V1.0 board acts as a data collector, gathering information from the soil moisture, temperature, and humidity sensors and functioning as a client that publishes this sensor data to the Thinkspeak cloud platform [14,43]. Power to operate the entire system is provided by a portable power bank.

In the study, a total of seven sensors were developed and distributed to seven teams, with each team consisting of two individuals. Each team was assigned to collect data from 12 locations on the same date. Consequently, data from 84 locations were collected during February 2022 over one week, considering an average of seven days of data for each station. To validate the accuracy of the soil moisture sensor, an on-site analog soil moisture sensor was utilized, and laboratory-based gravimetric soil moisture content tests were conducted on soil samples obtained from 50 % of the sampling locations [47]. Fig. 1S depicts the field sampling process and the laboratory setup for gravimetric soil moisture analysis. Additionally, Fig. 2S presents the results of the correlation analysis between the IoT sensor and the gravimetric-based soil moisture content, utilizing a linear regression plot that demonstrates the high reliability of the IoT sensor in collecting soil moisture data.

The overall framework adopted in this study is shown in Fig. 3.The framework developed mainly divided into three part first field data collection (IoT sensors) and data preparation using multi spectral remote sensing, terrain data, followed by feature selection method. Secondly, development of ensemble and deep learning models for SM prediction. Lastly, use of explainable AI (XAI) based SHAP and LIME method for global interpretability and site specific/local analysis.

The moisture content of soil primarily determines its dielectric constant and it is directly proportional, which in turn greatly influences the microwave signal [48]. In this study, the dielectric constant map was quantified by leveraging field soil moisture information obtained from SMAP/Sentinel-1 (SMAPS1) Level-2 soil moisture data (available at https://search.earthdata.nasa.gov/), while soil texture information was obtained from a previous study conducted by Mallik et al., [41]. The estimation of soil dielectric constant at a specific microwave frequency can be achieved using the equation proposed by Dobson et al., (1985) and Hallikainen et al., (1985) as follows [49,50]:

$$DC = a_0 + a_1 s + a_2 c + b_0 + b_1 s + b_2 cMv + c_0 + c_1 s + c_2 cMv \quad (1)$$

In the equation, 'DC' is the soil dielectric constant, 's' represents the percentage of sand by weight, 'c' denotes the percentage of clay by weight, and 'Mv' represents the volumetric soil moisture content. The coefficients 'ai', 'bi', and 'ci' are frequency-dependent coefficients.

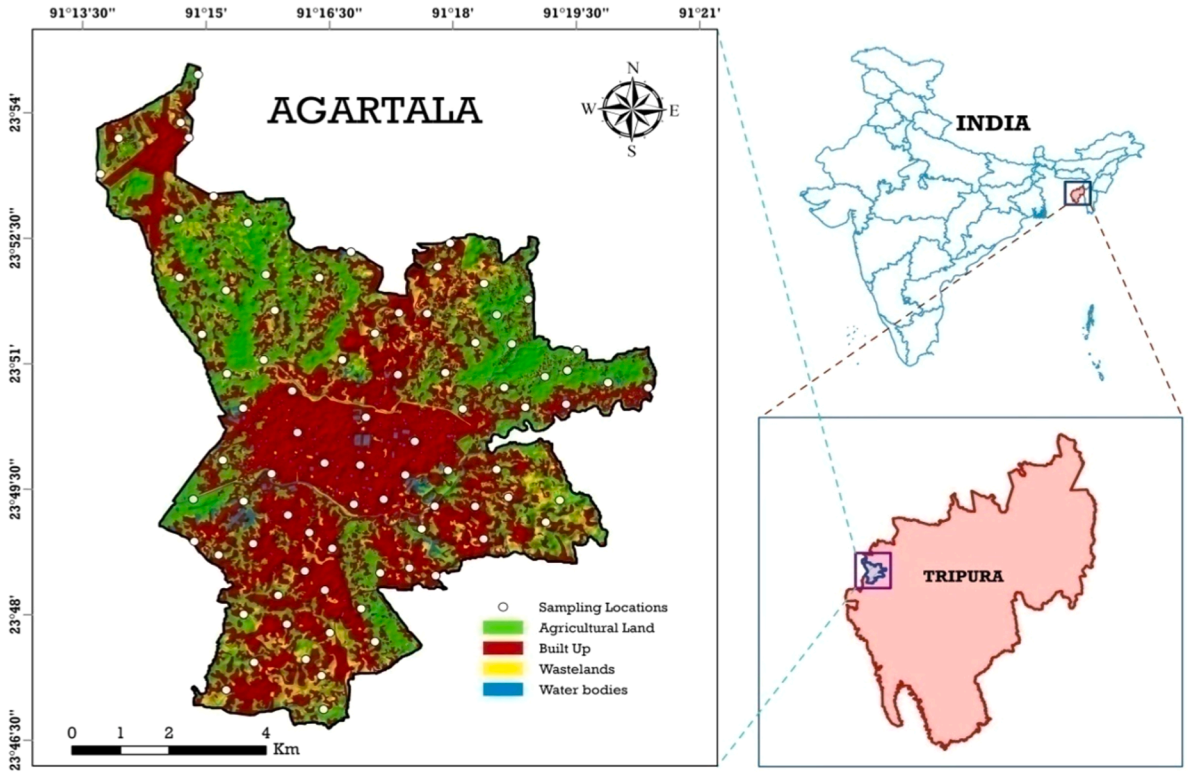


Fig. 1. Study area map with land use information and sampling location.

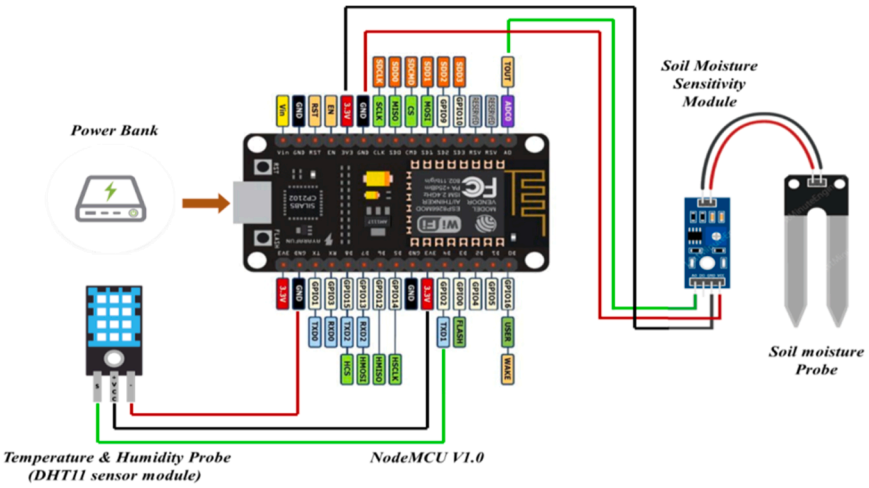


Fig. 2. Skeleton of in situ data collection using sensor.

2.3. Remote sensing data preparation

The core concept of remote sensing revolves around utilizing various platforms to capture and analyze the radiation emitted and reflected from a specific location. Particularly, for retrieving of soil moisture through utilization of different remote sensing data types, as extensively discussed by Nichols, (2011) and Wang and Qu, (2009)[6,51]. In this study, a combination of active and passive remote sensing data has been employed to estimate soil moisture.

2.3.1. Active remote sensing derivatives

Active remote sensing employs its energy source for illumination. One of the key advantages of active sensors is their ability to acquire measurements at any time, irrespective of the time of day or season. In this study, active remote sensing data from the Sentinel-1A satellite, specifically dated February 18, 2022, is assimilated from the Alaska Synthetic Aperture Radar (SAR) Facility, accessible at (https://search.asf.alaska.edu/#/).

However, SAR images are susceptible to noise, as well as geometric and radiometric distortions, as acknowledged by [18]. Therefore, the data undergoes preprocessing to address these distortions using SNAP software V7.0, following the established practices outlined by various researchers [4,17,52].Furthermore, the backscatter coefficient or sigma naught (σ0) data, which quantifies the reflective strength of radar targets in dual polarization (VV and VH), is extracted in the desired format using ArcGIS v 10.5.

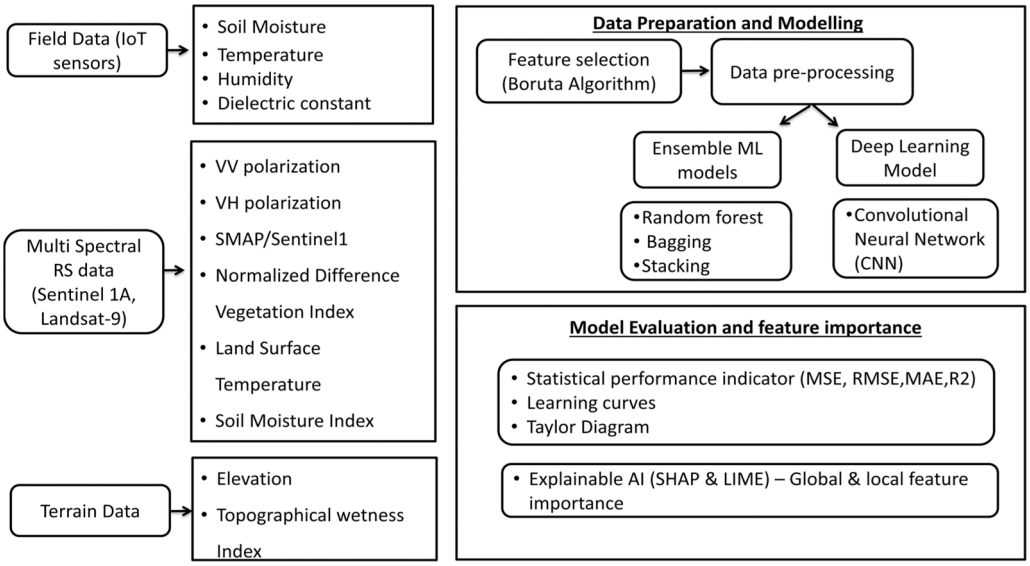


Fig. 3. Flowchart of the work.

2.3.2. Passive remote sensing derivatives

Passive sensors, which measure naturally occurring energy, can independently retrieve soil moisture information and provide insights into various land properties. The conjecture made by Sandholt et al., (2002) suggests the existence of a triangular relationship among vegetation cover, surface temperature, and soil moisture [53]. Hence, in this study, three passive remote sensing indices, namely the Normalized Difference Vegetation Index (NDVI), Land Surface Temperature (LST), and Soil Moisture Index (SMI), derived from optical and thermal bands of Landsat-9 data using ArcGIS software were utilized.

The Landsat-9 image, acquired on February 17, 2022, was obtained from(https://earthexplorer.usgs.gov/). NDVI serves as a simple graphical indicator that quantifies healthy green vegetation by measuring the difference between near-infrared (NIR) and red light. The range of NDVI values varies from $+1$ to -1, with -1 typically representing water bodies and $+1$ indicating dense green leafy vegetation, as described by [54]. LST represents the radiative skin temperature of the land surface, which is influenced by soil moisture and rainfall, as discussed by [28]. The methodology proposed by Avdan and Jovanovska (2016) was employed in this study to calculate LST using Landsat-9 data [55].

The Soil Moisture Index (SMI), is a metric that measures the ratio of the difference between the current soil moisture and the permanent wilting point to the difference between the field capacity and the residual soil moisture. SMI index values range from 0 to 1, where 0 signifies extremely arid conditions and 1 indicates exceptionally moist conditions. The process of retrieving the SMI involves direct utilization of LST, as described by [55]. The following equations were employed to calculate NDVI and SMI:

$$\mathbf{NDVI} = \mathbf{(NIR - Red)\ /\ (NIR + Red)} \tag{2}$$

$$\mathbf{SMI}=\mathbf{(LST_{max}-LST)/(LST_{max}-LST_{min})} \tag{3}$$

The calculation of the Normalized Difference Vegetation Index (NDVI) involves the utilization of the near-infrared band (Band 5) and the red band (Band 4). The Soil Moisture Index (SMI) represents the calculated metric for soil moisture. Additionally, the terms LST max, LST min, and LST refer to the maximum, minimum, and retrieve values of Land Surface Temperature (LST) respectively.

2.4. Topographic data preparation

The estimation of soil moisture content is significantly influenced by topography. However, there have been limited studies that incorporate topographic data in the development of predictive models for soil moisture estimation. Therefore, in this study, in addition to field and remote sensing data, topographic data including Elevation (ELV) and Topographic Wetness Index (TWI) were considered for the development of a soil moisture predictive model.

The Topographic Wetness Index (TWI) provides insights into how topography impacts the presence and distribution of saturated source areas that generate runoff, and it is utilized as an indirect indicator of soil moisture, as discussed by Naghibi et al., (2016) and Pal et al., (2020) [56,57]. TWI maps, along with Elevation (ELV) data, were generated using the Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM) with a resolution of 30 m. The following equation was employed for the estimation of TWI:

$$\mathbf{TWI}=\mathbf{ln}(\alpha/\tan\beta) \tag{4}$$

In this equation, the symbol α represents the cumulative upslope area that drains through a specific point, while tan β denotes the slope angle at that particular point.

2.5. Method for developing models for SM content prediction

In this research, ML and DL models were developed to predict Soil Moisture content by integrating multiple data sources to prepare eleven explanatory variables namely atmospheric temperature (TM), Humidity (HM), Dielectric constant (DC), SMAP/Sentinel 1 soil moisture (SMAPS1), Normalized difference vegetation index (NDVI), Land surface temperature (LST), Soil Moisture Index (SMI), VV polarization (σ_0VV), VH polarization (σ_0VH), Elevation (ELV), and Topographic Wetness Index (TWI). The specifications of these datasets, including their spatial and temporal resolutions, units, and sources, are detailed in Table 1.

These datasets were chosen for their high resolution and frequent temporal coverage, which are crucial for monitoring and predicting surface soil moisture content. The development of these models used for prediction Soil Moisture content involved a series of distinct steps, which are outlined in detail below:

2.5.1. Methods for feature selection

The Boruta algorithm, as described by Das et al., (2022) is a wrapper feature selection technique aimed at identifying the most crucial features for a given prediction task [17]. This algorithm operates by comparing the importance of each feature in the original dataset with

Table 1
Satellite derivative product specification used for prediction of soil moisture.

Dataset	Spatial Resolution	Temporal Resolution	Units	Frequency/Wavelength	Source
Sentinel−1 GRD (Ground Range Detected) Backscatter coefficient	10 m (VV, VH polarization)	6–12 days	Decibels (dB)	C-band (5.405 GHz / 5.6 cm)	European Space Agency (ESA)
SMAP/Sentinel−1 L2 Radiometer/Radar 30-Second Scene 3 km EASE-Grid Surface Soil Moisture	3 km	2–3 days	m^3/m^3	SMAP: L-band radiometer (1.4 GHz / 21 cm), Sentinel−1: C-band radar (5.405 GHz / 5.6 cm)	NASA/ESA
Landsat 9 derived NDVI using optical bands	30 m	16 days	Dimensionless (−1−1)	Red: 0.64–0.67 μm, NIR: 0.85–0.88 μm	USGS/NASA
Landsat 9 derived LST using Thermal and Optical bands	30 m	16 days	Kelvin (K)	Band 10: 10.6–11.19 μm, Band 11: 11.5–12.51 μm	USGS/NASA
Landsat 9 derived SMI using Thermal and Optical bands	30 m	16 days	Dimensionless (0−1)	Uses multiple bands	USGS/NASA
Elevation from SRTM 1 Arc-Second Global	30 m	Static	Meters (m)	C-band: 5.6 cm, X-band: 3.1 cm	NASA/USGS
Topographic Wetness Index (TWI)	30 m	Static	Dimensionless	N/A (derived from DEM)	Derived from SRTM

the importance of randomly generated shadow features. The shadow features are created by permuting the values of each feature independently, as outlined by Shahriar et al., (2020) [58]. The Boruta algorithm categorizes each feature into one of three labels: "confirmed," "tentative, " or "rejected." Features labeled as "confirmed" are deemed important by the algorithm. Features labeled as "tentative" require further analysis to determine their significance. On the other hand, features labeled as "rejected" are not considered further for the prediction of SM.

The Boruta algorithm can be applied using a range of machine learning or deep learning models, such as random forests, support vector machines, or neural networks. In the present, the random forest model has been used for Boruta algorithm. The selected features can then be used as inputs to the chosen model for predicting SM content.

2.5.2. *Methods for ML and DL models used for SM content prediction*

In recent years, the use of standalone machine learning techniques has become prevalent due to their ease of implementation. However, these standalone ML models may encounter challenges related to overfitting or underfitting during the testing phases, which can compromise their accuracy. To address these issues and improve generalization capacity and prediction robustness, ensemble models have gained significant popularity. By combining the outputs of multiple models, ensemble-based approaches offer valuable insights into the relationships between variables. Each model, referred to as a weak learner, focuses on different aspects of the data, while the combined model, known as a strong learner, provides an optimal predictive model [59]. Despite the advantages of ensemble models, selecting the most suitable algorithm from the vast array of available options can be challenging. To address this challenge, this study aims to compare the performance of various ensemble machine learning models, namely Bagging, Random Forest, and Stacking, with a deep learning-based CNN for predicting surface SM. By evaluating these different approaches, the study seeks to identify the most effective model for the task at hand.

2.5.2.1. Random forest (RF). Random forest is an ensemble learning method that combines multiple decision trees to make predictions [31, 48]. It utilizes feature selection techniques to reduce the interdependence between trees and employs bootstrapping to generate different subsets of data for each decision tree. The random forest algorithm is widely employed in regression and classification tasks due to its ability to handle large datasets and high-dimensional feature spaces. To optimize the performance of the random forest model, various parameters can be adjusted, including the number of trees, the maximum depth of each tree, and the minimum number of samples required to split a node. The precision of the random forest model is assessed using the out-of-bag (OOB) error metric [60].

2.5.2.2. Support vector machine (SVM). The Support Vector Machine (SVM) algorithm is widely recognized and employed in the field of machine learning for addressing classification problems [61]. Its fundamental objective is to construct a hyperplane in an n-dimensional space, leveraging kernel functions, to classify input points based on their attributes [62]. The SVM model encompasses four types of kernel functions, namely linear, polynomial, radial basis function, and sigmoid kernels that consider the input data as input and transforms it into the desired form [48]. In this particular study, a radial basis function was selected as the kernel function. To mitigate overfitting and regulate the level of nonlinearity, the associated hyperparameters, specifically the regularization parameter (C) and the kernel width (γ), were carefully adjusted.

2.5.2.3. Generalized linear model (GLM). The Generalized Linear Model (GLM) is a statistical methodology rooted in the principles of logistic regression [63]. GLM establishes a connection between dependent and explanatory variables, albeit with distinct assumptions compared to logistic regression [63,64]. The GLM model comprises three essential elements: a link function, a systematic component, and a random component. The systematic component delineates the linear predictor, whereas the random component characterizes the probability. The link function serves as the bridge connecting these two components together.

2.5.2.4. Bagging. Bagging is a technique that enhances the accuracy of machine learning methods utilized in regression and classification tasks. By employing the bootstrap repeated sampling methodology on the training dataset, bagging generates multiple classifiers, effectively reducing the variance of the dataset [65]. The individual output models obtained from the bootstrap samples are combined using the majority voting approach, which ensures that they are on a consistent scale. This process further enhances the precision of the machine learning approach.

2.5.2.5. Stacking. Stacking is an ensemble-based algorithm designed to enhance prediction accuracy by combining multiple base learners. The stacking algorithm operates in two distinct stages: level-0 and level-1 [66]. During the level-0 stage, diverse base learners are employed and trained on the input data. In the subsequent level-1 stage, the prediction outputs generated by these base learners are combined using meta-models [17,65]. This two-stage process enables stacking to leverage the strengths of different base learners and produce more accurate predictions by utilizing the meta-models for combining their outputs.

2.5.2.6. Convolutional neural networks (CNN). CNNs (Convolutional Neural Networks) are extensively utilized deep learning algorithms known for their exceptional feature extraction capabilities, particularly for high-dimensional data [67]. A typical CNN model architecture

consists of several components, including an input layer, convolutional layers, pooling layers, fully connected layers, and an output layer [68,69]. CNN operates in two stages: the first stage encompasses the feature extraction process, often referred to as the training stage, while the second stage involves the classification step [35]. One of the primary advantages of CNN is its ability to process entire raster data, such as image files, rather than being limited to vector or point-based data. This characteristic enables CNN to achieve faster processing times compared to other ML models [34].

2.5.3. Methods for evaluation of ML and DL models

To validate the performance of the models in predicting surface soil moisture (SM) content, several evaluation metrics can be used, including R^2, MAE, RMSE, and MSE.R^2 is a statistical measure that indicates the proportion of the variance in the dependent variable (SM content) that is explained by the independent variables (input features) in the model. A higher R^2 value indicates a better fit of the model to the data. MAE (mean absolute error) is a metric that measures the average absolute difference between the predicted values and the true values of the SM content. It provides a measure of the magnitude of the errors in the predictions. RMSE (root mean squared error) is a metric that measures the square root of the average squared differences between the predicted values and the true values of the SM content. It provides a measure of the magnitude of the errors in the predictions, with larger errors being penalized more heavily than smaller errors. MSE (mean squared error) is a metric that measures the average of the squared differences between the predicted values and the true values of the SM content.Equation in line with statistical error metrics are well explained by [31,45,70]. Furthermore, this study attempt to advocate the competency and robustness of Taylor diagram to compare the performance of the models. Taylor diagram graphically explains the relative skill of all the models that predict SM content for the study area [60].

2.6. Method for explainable AI (XAI) for SHAP and sensitivity analysis

2.6.1. Ensemble bagging model construction

The study begins by constructing an ensemble bagging model. Bagging (Bootstrap Aggregating) is an ensemble learning technique that aims to improve model performance and reduce overfitting by combining multiple base estimators. In this case, the base estimator used is a Random Forest Regressor, which is a decision tree-based model known for its robustness and accuracy in handling regression tasks.

2.6.2. Hyperparameter tuning through grid search

To optimize the ensemble bagging model, hyperparameter tuning is performed using GridSearchCV, a cross-validated grid search technique. Hyperparameters are parameters that cannot be learned during model training and need to be set beforehand. In this study, the hyperparameters to be tuned are the number of estimators (n_estimators), the maximum number of samples used for training each base estimator (max_samples), and the maximum number of features considered for each split in the decision trees (max_features).

2.6.3. Model evaluation and selection

After hyperparameter tuning, the bagging model is evaluated using a suitable metric, in this case, the R^2 score. The R^2 score quantifies the proportion of the variance in the target variable that can be explained by the model. The bagging model with the highest R2 score is selected as the best model.

2.6.4. XAI implementation

With the best bagging model identified, the study proceeds to implement Explainable Artificial Intelligence (XAI) techniques to gain insights into the model's behavior and improve interpretability.

a. **Global Interpretability using SHAP:** SHAP (SHapley Additive exPlanations) is a concept comes under game theory that is used for global interpretability [39]. SHAP values provide a unified measure of feature importance and quantify the contribution of each feature to the model's predictions across the entire dataset. This enables researchers to understand which features have the most significant impact on the model's overall performance.
b. **Site-Specific Analysis using LIME:** For site-specific analysis, LIME (Local Interpretable Model-agnostic Explanations) is employed. LIME is a post hoc model-agnostic technique that approximates the behavior of the complex model with a simpler, interpretable model around individual data points. It helps to explain the predictions made by the ensemble bagging model at specific sites, allowing researchers to understand the model's reasoning for each site's prediction. Sixteen sites are carefully chosen for the site-specific analysis. These sites are selected based on varying magnitudes of soil moisture to capture different scenarios and challenges in soil moisture management. The LIME analysis provides localized insights, helping to propose tailored and effective management plans for each specific location.

3. Results

3.1. Spatial mapping of IoT based collected parameters

To gain a comprehensive understanding of the distribution and variability of datasets across the study area, raster maps all the explanatory variables were generated. These maps, generated using the data obtained using both IoT and in-situ data, are presented in Fig. 4. The temperature map displays a consistent spatial distribution across the study area, except the southeastern region. Notably, a negative correlation was observed between temperature and humidity, whereby higher temperatures corresponded to drier weather conditions and subsequently lower relative humidity levels. Likewise, similar spatial variation patterns were observed between SM and the dielectric constant. This can be attributed to the fact that SM is influenced by the dielectric constant, with dry soil having a lower dielectric constant compared to wet soil or water. As soil moisture content increases, the dielectric constant of the soil also rises, indicating a positive relationship between the two variables.

3.2. Analysis and visualization of remote sensing based parameters

Figs. 5 and 6 present the spatial distribution of parameters derived from remote sensing data and terrain derivatives. Both NDVI and SMI exhibit similar spatial patterns, with lower to higher values observed as we move away from the central zone in the study area. In contrast, LST shows an inverse pattern, with higher values in the central region. This distribution can be explained by the fact that the central region is predominantly characterized by urban development, while the peripheral regions exhibit a greater presence of vegetation [41,71]. The NDVI values range from 0.481 to −0.059, where a value of −1 indicates non-vegetated land and a value of + 1 represents dense green foliage [72]. The SMI values indicate overall dry conditions throughout the study area, with values near 0 [73,74]. The soil moisture product derivate obtained from SMAP/sentinel1 indicate that the soil moisture the study area varies between 55(%) to 26(%). A huge change in the minimum value of SM is observed between SMAP/sentinel1 as compare to actual field SM value obtained using IoT.

The spatial distribution maps of backscatter coefficients, specifically VV and VH polarization, highlight higher values in the central and southern regions, while lower values are observed in the northeastern and northwestern regions. The higher backscatter values correspond to areas with significant urban development. It should be noted that SAR energy is influenced by the presence of vegetation [4]. Consequently, regions with dense vegetation having higher NDVI values in the raster,

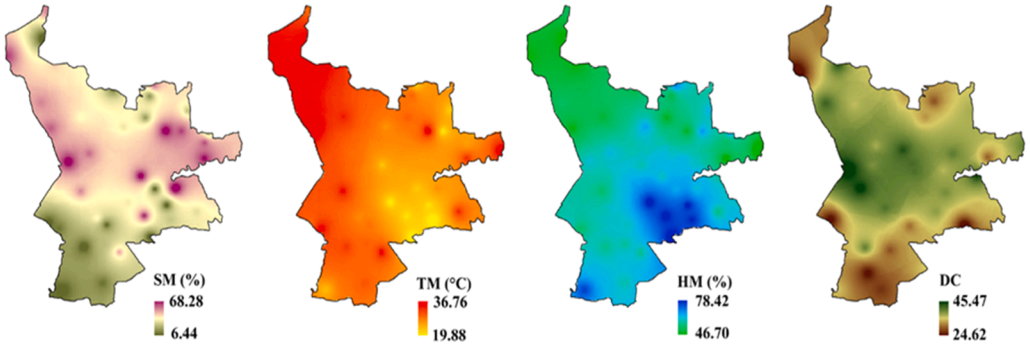


Fig. 4. In-situ field data (a) soil moisture (b) temperature (c) humidity (d) dielectric constant.

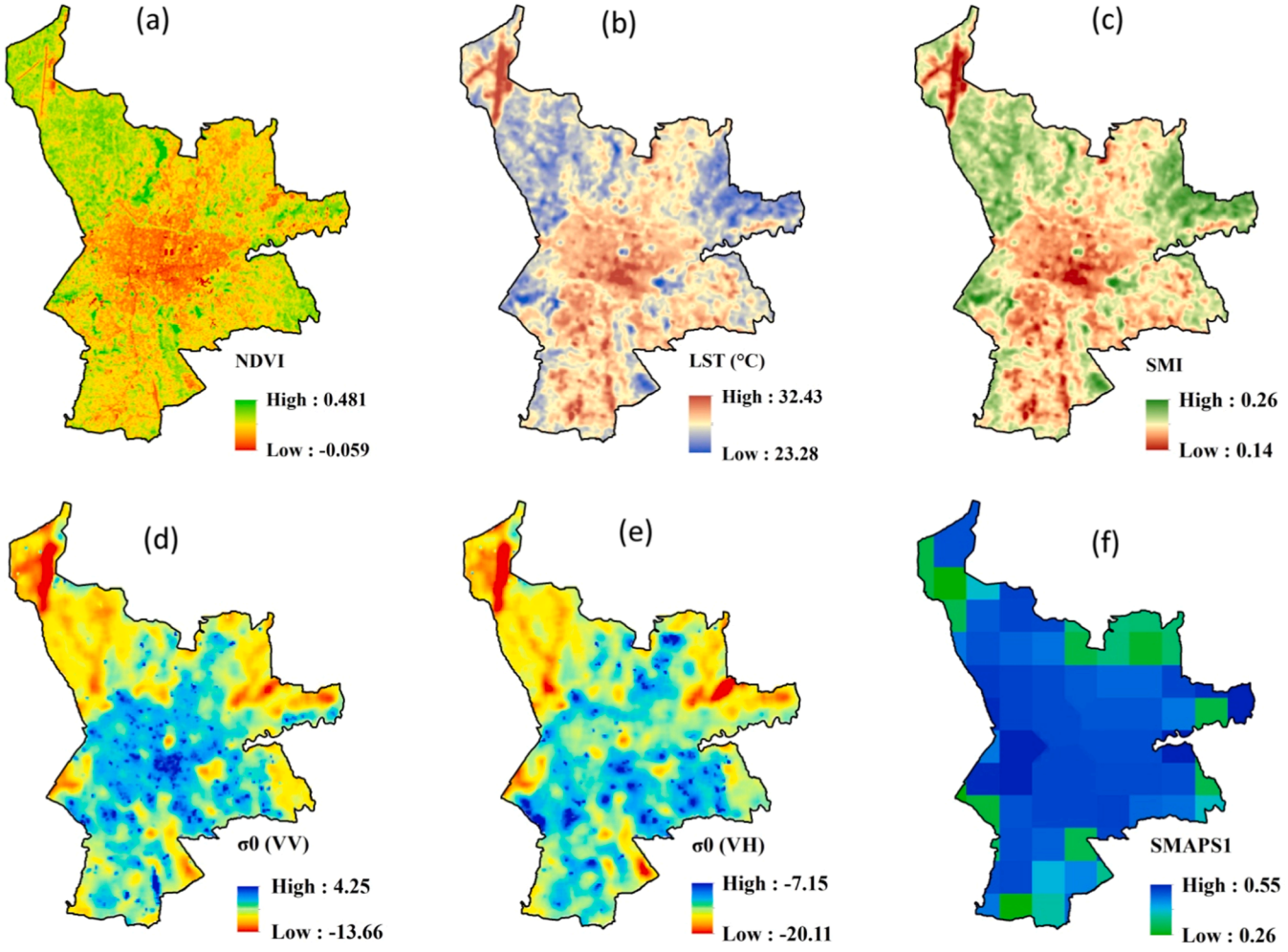


Fig. 5. Maps obtained from multi-sensor RS data (a) normalized difference vegetation index (b) land surface temperature (c) soil moisture index (d) VV polarization (σ_0VV) (e) VH polarization (σ_0VH) (f) SMAP/Sentinel 1 soil moisture.

exhibit lower values of the backscatter coefficient. The terrain derivatives, namely ELV and TWI derived from SRTM-DEM, indicate that the study area is predominantly characterized by low to moderate elevation. The TWI values range from 4.5 to 18.25, and it should be noted that the choice of the dry or wet season for field data collection of soil moisture introduces bias in the association between TWI and soil moisture estimation [75].

3.3. Development of SM content prediction models

3.3.1. Feature selection analysis

Data pre-processing, specifically feature selection, plays a vital role in improving performance and reducing the complexity of machine learning models. In this study, we employed the Boruta algorithm to select relevant features. The implementation of the Boruta algorithm and model preparation was carried out using Python libraries, namely BorutaPy, and scikit-learn. The Boruta algorithm identified the significant variables, which were then utilized as inputs for various models to

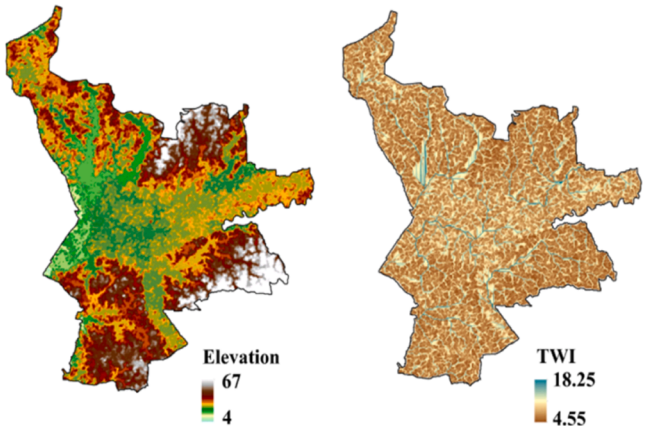


Fig. 6. Maps obtained from topographic data (a) elevation (b) topographical wetness index.

predict surface SM content. Fig. 3S illustrates the output generated by the Boruta algorithm, indicating that none of the variables were rejected and all variables were considered for predicting surface SM content.

3.3.2. Implementation of ML and DL models for predicting SM content

In the implementation of the study, an ensemble-based bagging model was first applied using the scikit-learn package in Python. The dataset had 12 columns, with the first 11 columns representing input parameters (DC, ELV, HM, LST, NDVI, σ0VH, σ0VV, SMAPS1, SMI, TM, and TWI), and the last column representing the output parameter IoT-based (SM). The dataset was then split into an 80:20 ratio for training and testing sets using the train_test_split method from scikit-learn. The bagging model consisted of three individual models, namely Random Forest (RF), Support Vector Machine (SVM), and Generalized Linear Model (GLM), combined using the Voting Regressor method. The bagging model was iterated several times to obtain the best possible parameters, including the number of jobs to run in parallel, the verbosity level, and the weights of each model. Similarly, each individual model within bagging was repeated multiple times to obtain the best possible model. A total of 110 iterations were performed to obtain the best possible model. The best combination of parameters for the bagging model is summarized in Table 1S.

In this study, a diverse set of machine learning models was employed to predict surface SM content. Initially, a random forest regression model was developed, and 30 iterations were conducted to determine the best-performing model, which happened to be the 13th iteration. This particular model consisted of 50 trees, a minimum depth of 6, a maximum depth of 12, a minimum number of leaves set at 62, and a maximum number of leaves set at 100.

Following that, an ensemble-based stacking model was implemented, incorporating three base learners: RF Regressor, SVM, and GLM. The optimal combination of parameters for the stacking model was determined by utilizing RF, SVM, and GLM, as specified in Table 2S. To obtain the most favorable parameter combination with minimal error, the RF and SVM models were subjected to 50 iterations. The SVM model employed a radial-based kernel function, with a C value of 1, indicating improved generalization to new data. The kernel width (γ) gamma value was determined by considering the number of features and their variance through the gamma scale. Consequently, GLM (Generalized Linear Model) was employed as a linear model to gain insights into the relationships between the input features and the target variable. A total of 25 iterations were performed to obtain the best coefficients for the GLM model. The resulting coefficients for the input parameters were as follows: [1.37, −0.08, −0.19, 43825.18, −5.88, −0.43, −0.02, −10.32, 3379474.37, 0.09, 0.01]. Additionally, the slope-intercept value was determined as (-1889427.64).

A new data frame was created using the predictions from each model, and these features were arranged from left to right in the column as RF, SVM, and GLM predictor. The data frame was then scaled using Standard Scaler, and a Linear Regression model was trained on the stacked features using the training data. The models were fitted on the training data, and predictions were made on the testing data using the stacked model. The equation of regression of the best-stacked model obtained after 27 iterations had coefficients for the input parameters as: [8.57, −0.03, −0.002] and the value of the slope-intercept value was 33.55.

Furthermore, In this study, a 1D_CNN (One-Dimensional Convolutional Neural Network) model was employed to predict a continuous target variable. The implementation was carried out using the Keras and Tensorflow packages in the Python programming language. The CNN architecture consisted of three Conv1D layers followed by a Flatten layer. The first Conv1D layer had 32 filters with a kernel size of 3 and employed a ReLU activation function. The output of this layer was then passed through a MaxPooling1D layer with a pool size of 1. Similarly, the second Conv1D layer comprised 64 filters with a kernel size of 3 and a ReLU activation function, followed by a MaxPooling1D layer with a pool size of 1. The third Conv1D layer contained 128 filters with a kernel size of 3 and a ReLU activation function. It was succeeded by a MaxPooling1D layer with a pool size of 1.

Following the third MaxPooling1D layer, the output was fed into a Flatten layer. This layer converted the three-dimensional tensor output from the third Conv1D layer into a one-dimensional tensor, enabling it to be utilized as input for the Dense layers. The Flatten layer achieved this by concatenating all the feature maps into a single long vector. The resulting vector was then passed through two Dense layers with ReLU activation functions. The first Dense layer had 64 units, while the second Dense layer had 1 unit and employed a linear activation function to predict the continuous target variable. The model was compiled with the Adam optimizer and utilized the mean squared error loss function. To prevent overfitting, early stopping was implemented with patience of 10. The model was trained on the training data for 100 epochs, with a batch size of 32 and a validation split of 0.2.

In summary, the Conv1D layers in this study utilized small kernel sizes and an increasing number of filters to extract higher-level features from the input data. The MaxPooling1D layers were employed to down sample the data and reduce the number of parameters in the model. Finally, the Flatten layer was used to transform the output of the Conv1D layers into a format suitable for the Dense layers.

3.3.3. Computing variable importance associated with all models

Variable importance analysis is a crucial technique used in machine learning models to determine the significance of input variables [76]. This technique plays a crucial role in the model building process, as it helps to identify which features contribute the most to the model's predictive power. By incorporating variable importance analysis, the robustness and interpretability of machine learning models can be enhanced. In this regard, Fig. 7 illustrates the variable importance of four models, including Bagging, Random Forest, Stacking, and CNN.

In all the models examined, the variable DC emerged as significantly more important for predicting soil moisture compared to other characteristics. Additionally, it was observed that the relative importance of features followed a consistent pattern across the models. Notably, in the case of the CNN model, SMI was found to be relatively more significant than LST. However, given that their importance values were similar; this slight variation in the ordering of importance was considered to be of minimal significance.

3.3.4. Evaluation metrics and model comparison

Before proceeding with the evaluation and implementation of models for raster generation using various approaches, learning curves were employed to depict the learning performance progression with experience. A learning curve is a visual representation that illustrates the

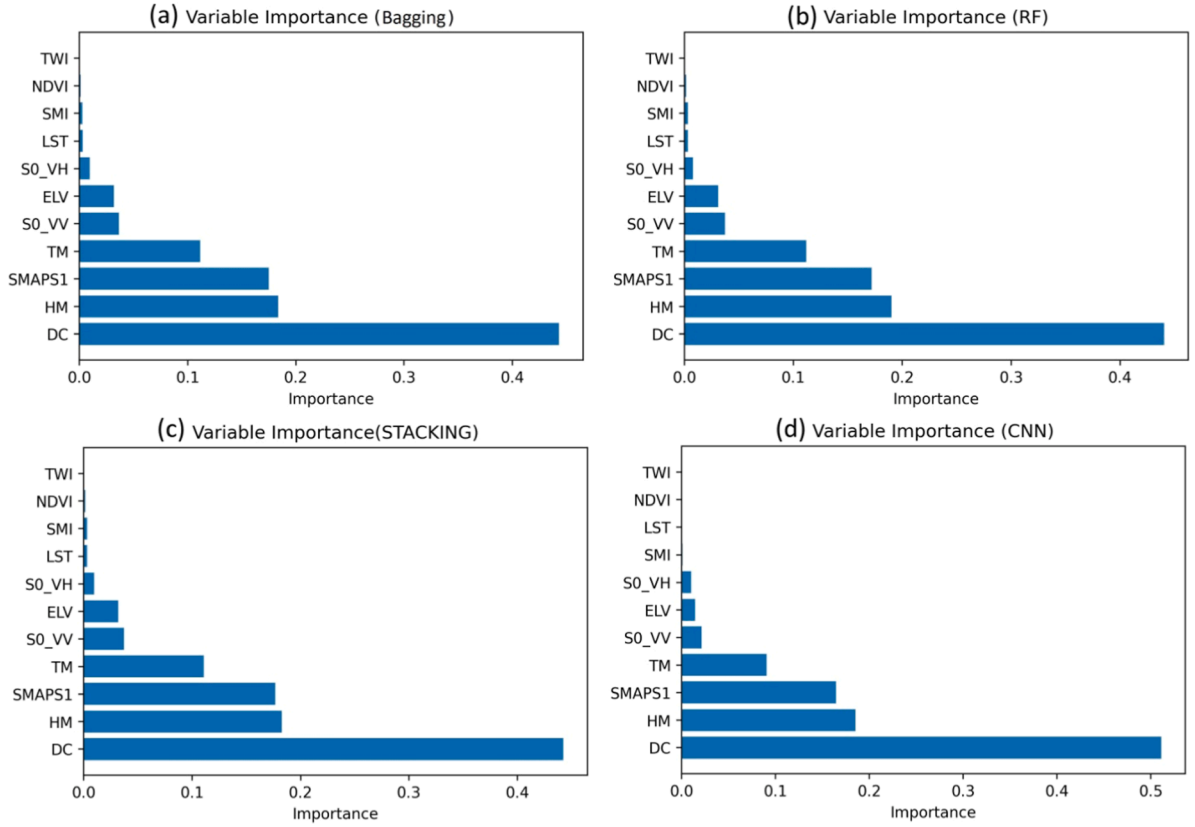


Fig. 7. Variable importance diagram for (a) bagging (b) random forest (c) stacking (d) CNN.

degree of learning or improvement attained by a machine learning model on the y-axis, concerning the size of the training dataset on the x-axis. In this study, the Mean Square Error (MSE) value is depicted on the y-axis of the learning curve. A high MSE error in the learning curve signifies inadequate fitting of the model to the training dataset. Models with high bias tend to oversimplify the learning process and do not adequately capture the nuances of the training data, resulting in elevated errors on the training data and diminished generalization performance on the test data, and vice versa.

Fig. 8 displays the learning curves for Bagging, RF, Stacking, and CNN models. It is evident from the figure that all the models do converge well, and MSE decreases as the number of datasets increases, indicating that the selected model obtained after numerous iterations were well-fitted models.

The credibility of the developed models was assessed by evaluating various statistical metrics such as R^2, RMSE, MSE, and MAE for the RF, Bagging, Stacking, and CNN models in both the training and testing phases. Figs. 9 and 10 depict the linear regression plots, which were used to assess the fitness of the data patterns and present the values of the statistical evaluation metrics for the training and testing phases. From

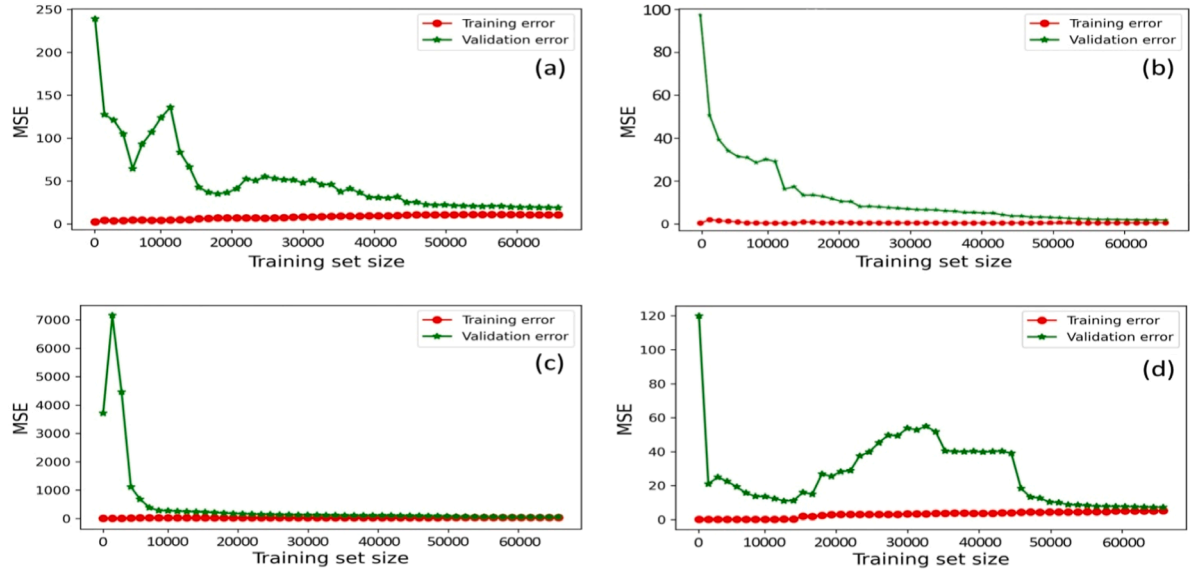


Fig. 8. Learning curves for a) bagging, b) random forest, c) stacking, d) CNN.

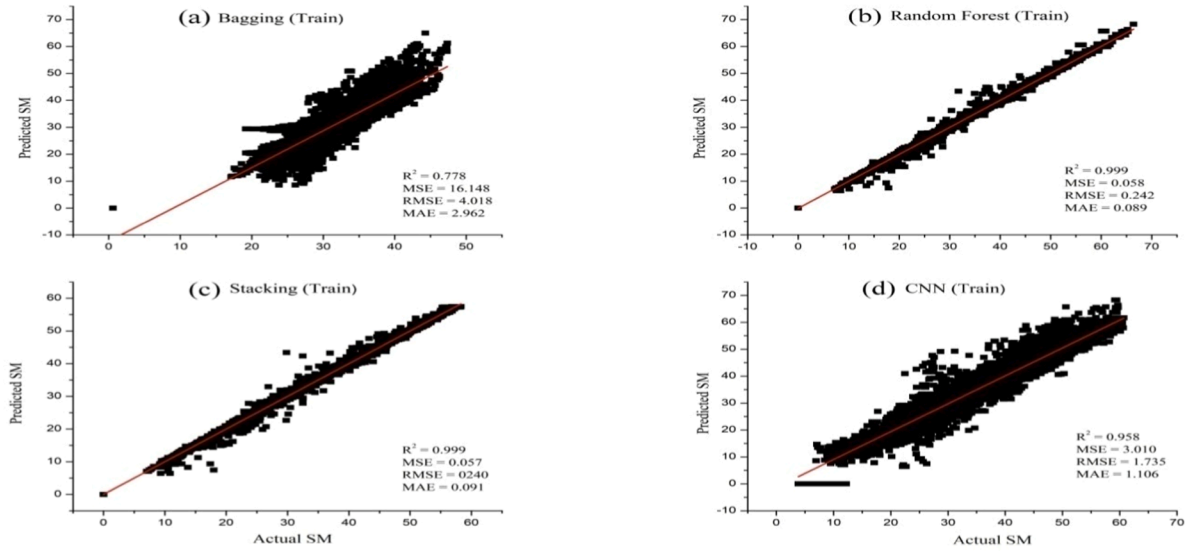


Fig. 9. The linear relationship curve between actual and predicted SM content for training data set (a) bagging (b) random forest (c) stacking (d) CNN.

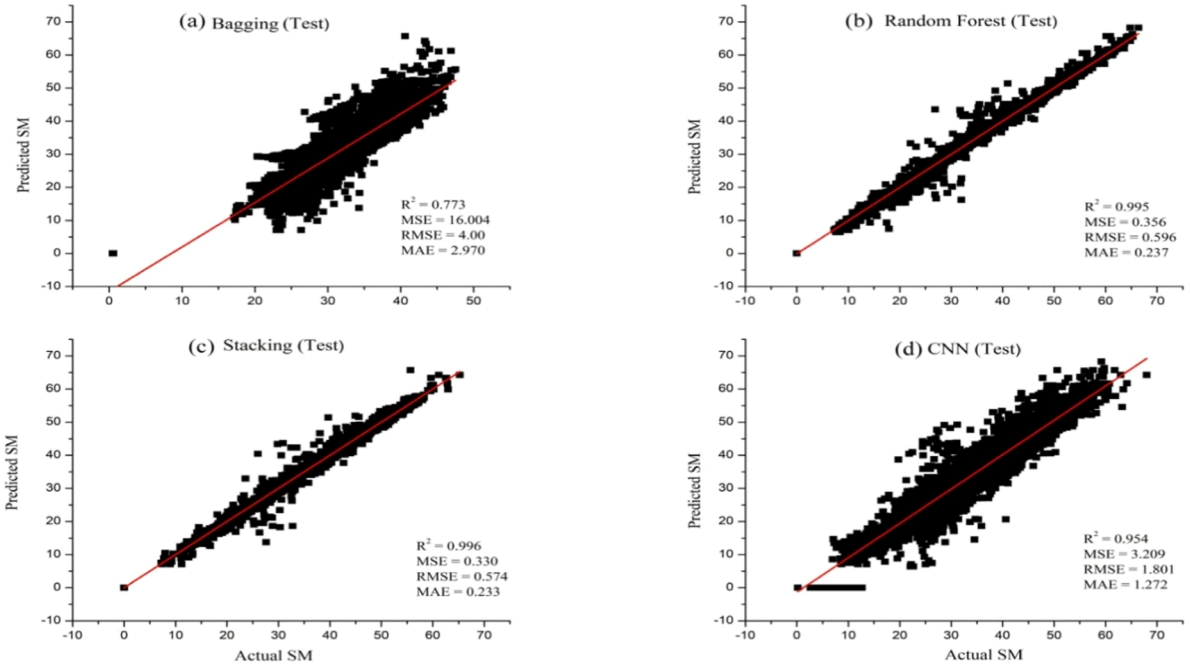


Fig. 10. The linear relationship curve between actual and predicted SM content for testing data set (a) bagging (b) random forest (c) stacking (d) CNN.

the figures, it is evident that the RF and Stacking models exhibit similar performance, while the Bagging and CNN models demonstrate comparatively lower and moderate performance. The R^2 values for the training phase are 0.999, 0.999, 0.778, and 0.958 for RF, Stacking, Bagging, and CNN models, respectively. The RMSE values are 0.242, 0.240, 4.018, and 1.735, the MSE values are 0.058, 0.057, 16.148, and 3.010, and the MAE values are 0.089, 0.091, 2.962, and 1.106 for the same models.

A similar trend is observed for the testing data, with minor differences in the values of the evaluation metrics. The Stacking model outperforms all other models, exhibiting the best fit and consistent co-linearity with the predicted values in both the training and testing phases. Based on the evaluation criteria, the models can be ranked in ascending order as follows: Stacking > RF > CNN > Bagging.

Furthermore, after conducting model diagnosis and evaluation, the robust Taylor diagram was employed to facilitate a comparative graphical assessment of the alignment between the model's output and the actual soil moisture (SM) data. This diagram utilizes correlation coefficients on the X-axis and standard deviations on the Y-axis to quantitatively measure the correlations between the predicted and actual SM content data. The distance from each point to the origin represents the root-mean-square error (RMSE). Fig. 11 showcases the Taylor diagram obtained for all the models using both the training and testing datasets.

From the figure, it is evident that the Stacking and RF models closely align with the highest correlation and RMSE values of the actual SM data. The models can be ranked as Bagging < CNN < Random Forest < Stacking based on their performance. Throughout the model evaluation processes and diagnostic tests, the Stacking model emerges as a robust AI tool for accurately and reliably emulating actual SM content in the study area. The enhanced performance of the stacked model may be attributed to the complementary nature of the base learners, resulting in

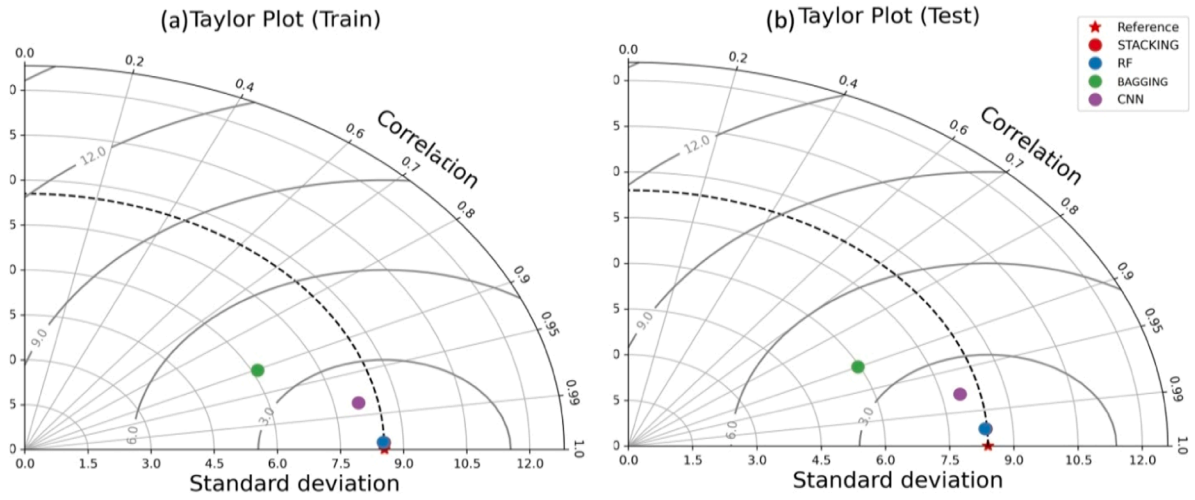


Fig. 11. Taylor diagram for the graphical presentation of the performance of models, namely, bagging, RF, staking, and CNN for (a) training and (b) testing phase.

reduced overfitting and bias compared to individual ML models [17]. The relatively lower performance of the Bagging model compared to others could be attributed to it being one of the earliest and simplest ensemble machine learning techniques, which is most suitable for problems with small training datasets [65]. The CNN model may not have outperformed the stacking ensemble or random forest models due to the nature of the data used in this study, which were tabular data rather than image data. Additionally, no specific relationships were considered while ordering the columns of the tabular data.

3.4. Spatial mapping of predicted SM content

Agartala is the capital of Tripura and has a humid subtropical climate with hot, humid summers, mild to cool winters, and moderate rainfall due to its location in northeastern India and proximity to the Bay of Bengal and Himalayan Mountains. The region generally has moderate to high soil moisture content levels. To study soil moisture distribution in Agartala during the pre-monsoon season, 84 IoT sensor readings were collected and an IDW interpolation-based soil moisture map was created as shown in Fig. 11(a). The map showed varying soil moisture levels

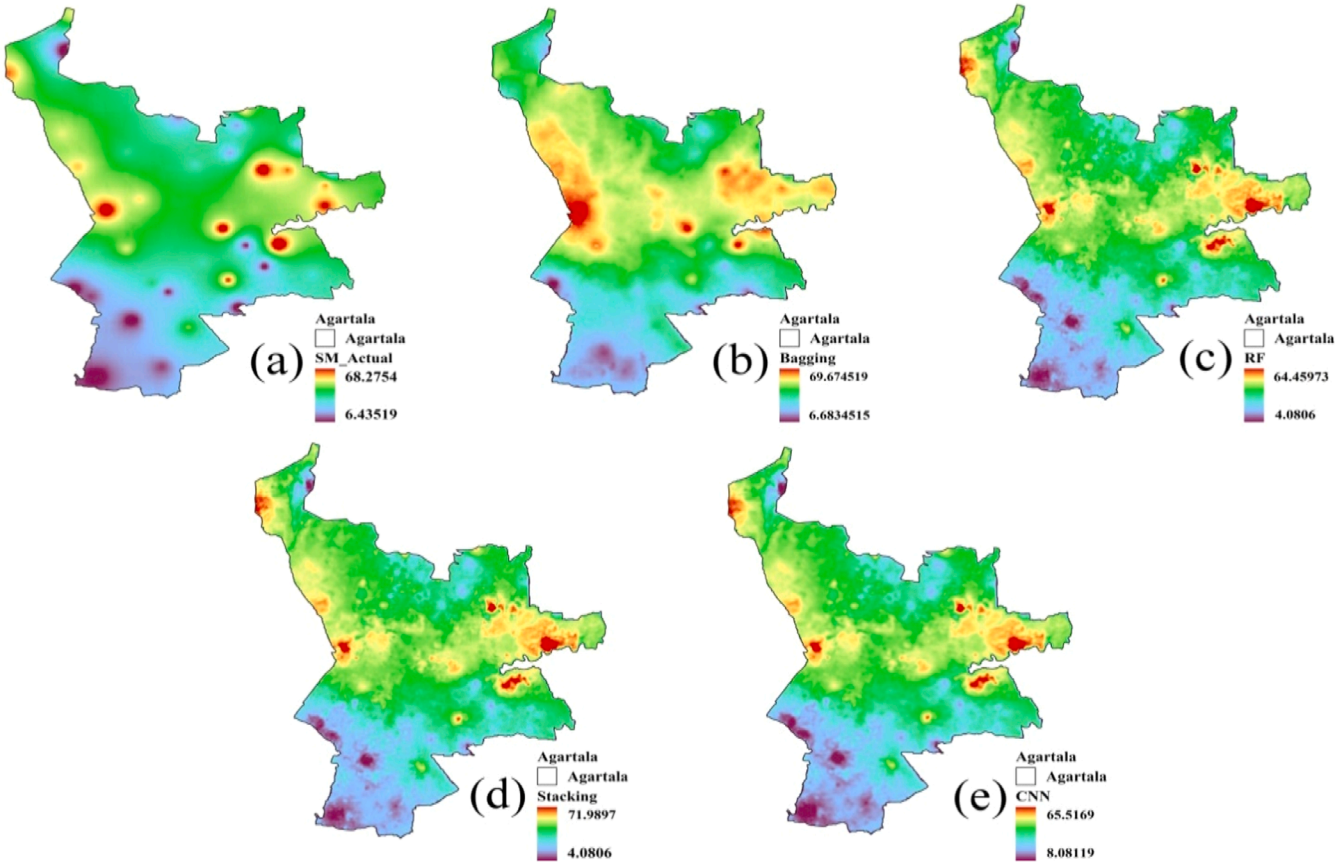


Fig. 12. Soil moisture using (a) actual IoT data, (b) bagging, (c) random forest, (d) stacking, (e) CNN.

across the region, with the highest soil moisture content in the northeastern region and the lowest in the southwestern region. Soil moisture distribution is influenced by factors such as vegetation, canopy cover, slope, soil texture, and built-up area composition, with finer soil texture composition regions displaying higher soil moisture content than coarser soil texture regions. The distribution of soil moisture in Agartala is dynamic and reflects the influence of humidity and temperature on soil moisture content.

To study the spatial dynamics of predicted soil moisture, the best models of the different machine learning algorithms were used to generate raster maps, as shown in Fig. 12(b,c,d,e). The raster maps generated from the machine learning algorithms exhibited spatial distributions closely resembling the actual soil moisture map. The most accurate representation was provided by the predictions of the Stacking model, which portrayed distinct pockets of very high soil moisture in the northeastern and northwestern regions, transitioning to moderate soil moisture in the central region, and culminating in very low soil moisture values in the southern region. In addition to accurately capturing the spatial distribution of soil moisture, the Stacking model demonstrated exceptional accuracy and precision in predicting the range of soil moisture values across the study region, estimating a range of 4.0806–71.9897 compared to the actual range of 6.43519–68.2754. The Random Forest model, while performing well in capturing the distribution of soil moisture, slightly underestimates the high and low values of the soil moisture range. It predicts a range of 4.0806–64.45973. On the other hand, the CNN model accurately represents the spatial distribution of soil moisture but overestimates the low value and underestimates the high value of the soil moisture range. Its predicted range is 8.08119–65.5169. Lastly, the Bagging model, although not entirely inaccurate in depicting soil moisture distribution, exhibits some discrepancies when compared to the other models.

3.5. Implementation of XAI for improving management strategy

3.5.1. Development of bagging ensemble model

The study focused on implementing explainable artificial intelligence (XAI) using an ensemble bagging model. The bagging model was optimized through grid search to ensure high accuracy. Once the best bagging model was identified, XAI techniques were employed to gain insights and improve decision-making.

Firstly, the ensemble bagging model was constructed using a RandomForest Regressor as the base estimator, and hyperparameter tuning was performed using GridSearchCV. The parameter grid for tuning included the number of estimators (n_estimators), the maximum number of samples to draw from the training dataset (max_samples), and the maximum number of features to consider for each split (max_features).

To gain a better understanding of the hyperparameter space and the model's performance, a 3D plot of the hyperparameters was created (Fig. 13). The x, y, and z axes represented the values of n_estimators, max_samples, and max_features, respectively, while the mean test score (R^2 score) was depicted as the surface plot. The best combination of hyperparameters was highlighted with a red marker, indicating the optimal configuration for the bagging model (max_features': 0.7, 'max_samples': 0.9, 'n_estimators': 200).

The best model was selected based on its performance, measured by the R^2 score. The best model and its corresponding hyperparameters were stored for further analysis. The best model's performance was evaluated using various metrics on the testing data. The MAE of 2.58 indicated that, on average; the model's predictions deviated by approximately 2.58 units from the actual values. The MSE of 11.43

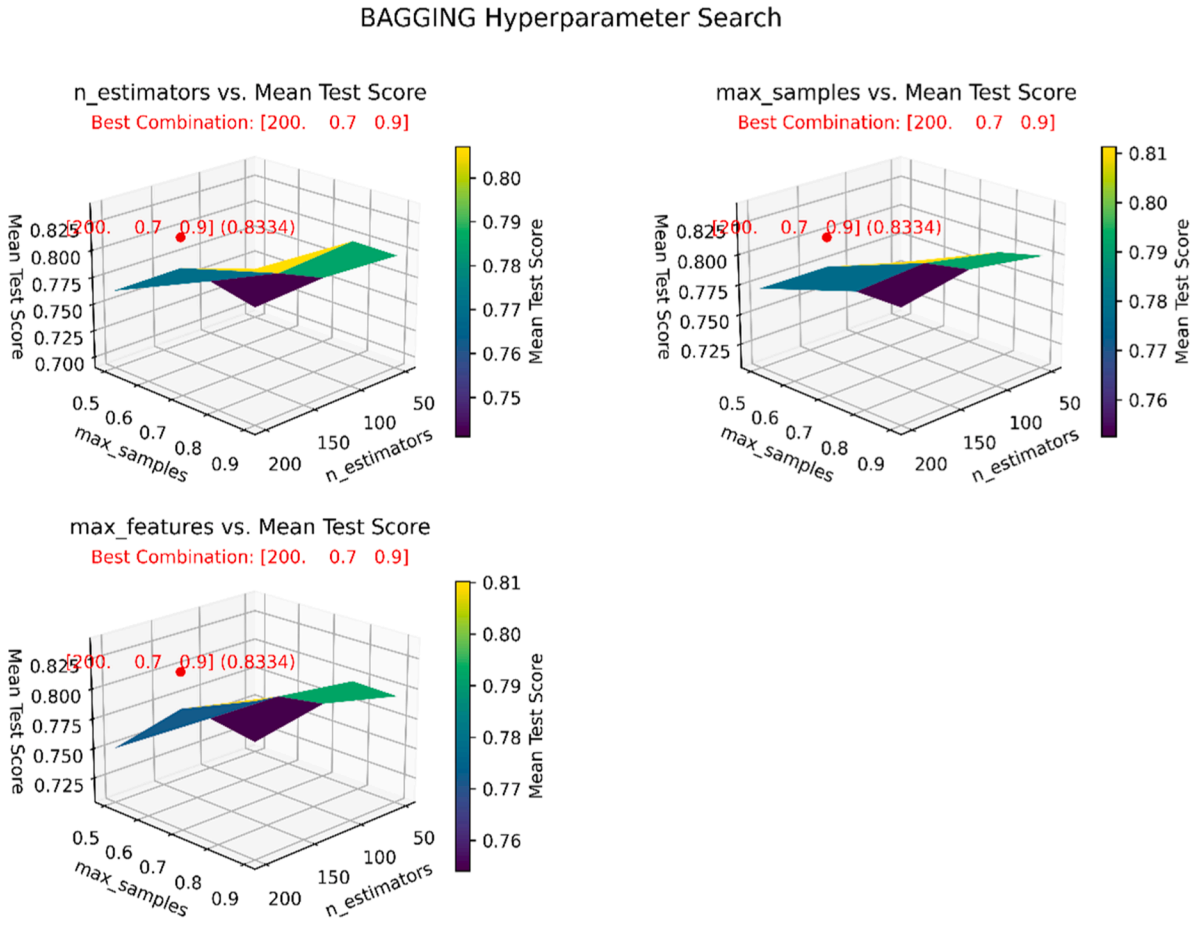


Fig. 13. Hyper parameter search for ensemble bagging model.

represented the average squared difference between the model's predictions and the true values, giving more weight to larger errors. The RMSE of 3.38 indicated the average magnitude of the prediction errors in the original unit, providing a meaningful measure of the model's accuracy. The R^2 Score of 0.86 demonstrated that approximately 86 % of the variance in the dependent variable was explained by the model's predictions, signifying a good fit to the data. Moreover, the Explained Variance Score of 0.86 revealed that the model accounted for about 86 % of the total variance in the target variable, indicating a reliable and robust performance. These metrics collectively demonstrated the best model's high accuracy and its ability to effectively capture the underlying patterns in the data, making it a valuable tool for making informed decisions and predictions.

3.5.2. Global level analysis using SHAP

After computing SHAP and variable importance from the best bagging model (Fig. 14), the results reveal valuable insights into the factors influencing soil moisture predictions. The SHAP model indicates that lower values of humidity, atmospheric temperature, and TWI have a higher influence on the prediction of soil moisture (Fig. 14a). Conversely, higher values of σ_0VH have a greater impact on the soil moisture prediction. Additionally, lower elevation and DC values exhibit a moderate influence on the soil moisture prediction. The variable importance analysis further supports these findings, indicating that humidity, atmospheric temperature, and TWI have a significant impact on soil moisture (Fig. 14b). This suggests that changes in these variables can lead to substantial fluctuations in soil moisture levels. The interpretation of these results is scientifically meaningful and can be used to make informed decisions for soil moisture management in agriculture. Furthermore, modifying the landscape, such as altering the drainage patterns or implementing soil conservation measures, and land use planning can potentially affect the TWI resulting in controlling the soil moisture content.

Therefore, the scientific analysis of SHAP and variable importance from the best bagging model provides valuable insights into the key factors influencing soil moisture predictions. By understanding the significant variables, stakeholders in agriculture and land management can take targeted steps to maintain optimal soil moisture levels. The findings will aid in the development of more effective and sustainable soil moisture management strategies, ultimately benefiting agricultural productivity and promoting sustainable land use practices.

3.5.3. Local level analysis using LIME

For the local level analysis, we employed LIME to estimate the responsible and important parameters for each site. LIME allows us to generate site-specific explanations that can aid in localized management decisions. We selected 16 sites, each with varying magnitudes of soil moisture (S1, S2, S5, S9, S10, S16, S20, S21, S23, S24, and S26) for analysis (Fig. 15).

At Site 10, we observed a very high importance of Humidity, with a value of 11.17. This indicates that Humidity has a significant impact on soil moisture at this site. On the other hand, Temperature exhibited a negative impact with an importance of −1.25, suggesting that higher temperatures are associated with lower soil moisture levels. The other features such as σ_0VH, LST, DC, TWI, ELV, DEM, σ_0VV, SMAPS1, NDVI, and SMI also contributed to soil moisture variations, though to a lesser extent. Site 16 displayed a similar pattern, with Humidity being the most influential factor with an importance of 11.35. Moreover, Temperature exhibited a negative influence (-1.50) on soil moisture, showing that higher temperatures were correlated with lower soil moisture content. Other factors like TWI, σ_0VH, ELV, DC, SMAPS1, LST, NDVI, σ_0VV, SMI, and DEM contributed to soil moisture variations at this site as well. At Site 26, Humidity was again highly significant, with an importance of −7.77. This indicates that high humidity levels are associated with low soil moisture. Temperature also had a notable impact, with an importance of 2.55, suggesting that lower temperatures are related to higher soil moisture content. Additionally, DC, σ_0VH, σ_0VV, LST, SMAPS1, NDVI, ELV, TWI, DEM, and SMI all played roles in determining soil moisture levels at this site. Site 5 exhibited a significant negative impact of Humidity with an importance of −4.84, indicating that higher humidity levels were associated with lower soil moisture content. On the other hand, Temperature displayed a positive impact (2.42), suggesting that lower temperatures were related to higher soil moisture levels. Other contributing factors at this site were LST, DC, TWI, σ_0VH, ELV, NDVI, σ_0VV, DEM, SMAPS1, and SMI.

Based on the site-specific analysis using LIME for the 16 sites, we have quantified the importance of different environmental parameters in relation to soil moisture at the local level. The analysis reveals the following key findings. Humidity has the most significant impact on soil moisture levels across all sites. An increase in humidity by 1 unit (on the specified scale) is associated with a decrease in soil moisture by approximately 4.71 units, on average. Conversely, a decrease in humidity is linked to an increase in soil moisture. Managing and monitoring humidity levels can thus be crucial in regulating soil moisture content. Temperatures above 0.08 on the scale have a negative influence on soil moisture, with an average impact of approximately −1.27 units. This suggests that higher temperatures can lead to decreased soil moisture. Addressing temperature fluctuations and heat stress in the soil could play a role in maintaining optimal moisture conditions. DC and NDVI has least impact on soil moisture, with an average importance of

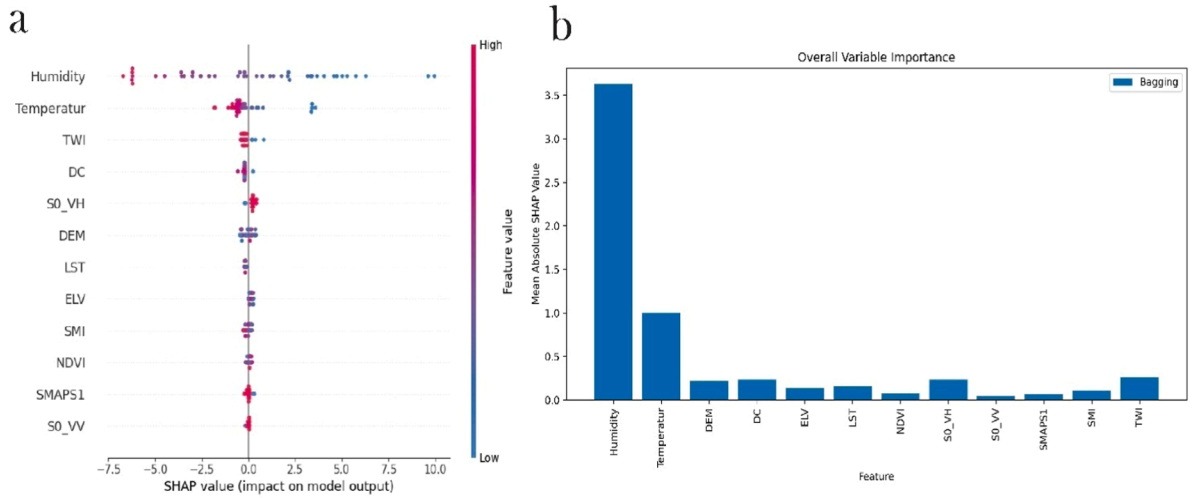


Fig. 14. Computation of SHAP (a) for variable importance (b) for global level analysis.

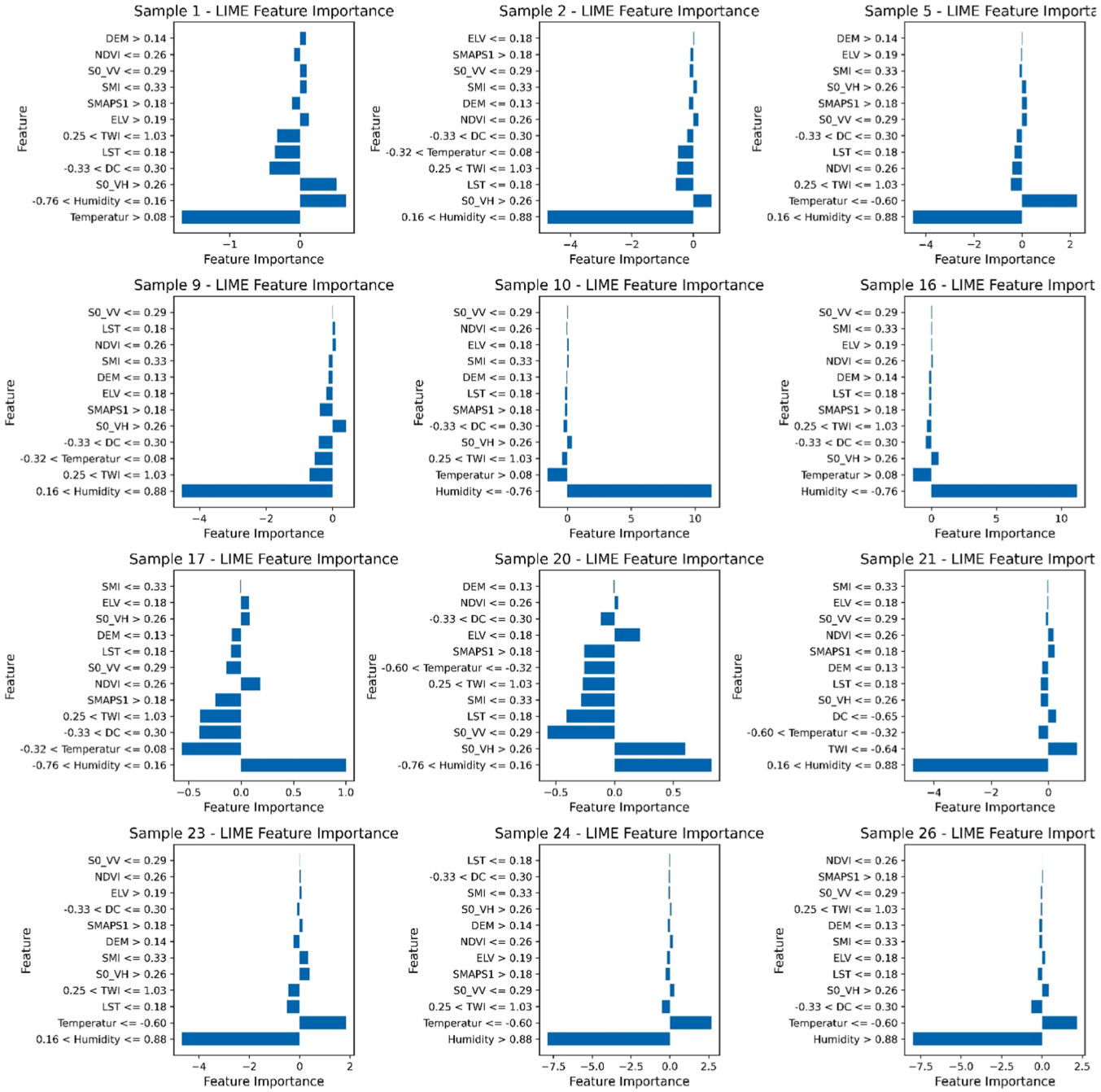


Fig. 15. Site specific analyses for better management using LIME.

approximately −0.3, 0.26 units. This implies that other factors influencing the dielectric constant of the area and ensuring adequate vegetation cover and healthier vegetation in these areas might contribute to better soil moisture retention. LST values less than or equal to 0.18 are negatively correlated with soil moisture, with an average importance of approximately −0.27 units. Managing urban heat islands and mitigating temperature rise in specific regions could positively impact soil moisture. TWI within the range of 0.25–1.03 has a negative impact on soil moisture, with an average importance of approximately −0.31 units. Considering the topography and hydrology of these areas can be essential in moisture management. ELV greater than 0.19 is positively associated with soil moisture, with an average importance of around 0.14 units. Higher elevation areas might experience better moisture conditions, and conservation measures in these regions could help preserve soil moisture. SMAPS1 greater than 0.18 has a negative influence on soil moisture, with an average impact of approximately −0.12 units. Monitoring and managing soil moisture using remote sensing data can aid in addressing moisture deficits.

These findings provide critical insights for local-level management of soil moisture specifically in agriculture area. By focusing on site-specific parameters and their impacts, appropriate strategies can be developed for each site to optimize soil moisture levels. Implementing such tailored management approaches could contribute to improved agricultural productivity, water conservation, and ecological sustainability in each region.

4. Discussion

This study endeavors to develop a machine learning-based model to predict surface soil moisture content in the Agartala region of India

using IoT sensor remote sensing data. Feature selection was performed using the Boruta algorithm, and the chosen features were employed to construct machine learning and deep learning models, namely Bagging, Random Forest, Stacking, and CNN. The findings from the spatial mapping of IoT-gathered parameters demonstrated a consistent spatial distribution of temperature across the study area, except the southeastern region. A comparable pattern was observed between soil moisture SM and the dielectric constant. Moreover, the results unveiled an inverse relationship between temperature and humidity. The spatial distribution of remote sensing-based parameters indicated that NDVI and SMI exhibited similar spatial distributions, while LST exhibited an opposite pattern. The backscatter coefficient maps displayed higher values in the central and southern regions and lower values in the northeastern and northwestern regions. The terrain derivatives ELV and TWI indicated that the study area was primarily characterized by low to moderate elevation.

The study presented a comparative analysis of different machine learning models for predicting SM content at a spatial scale in this study area. The learning curves were employed to assess the learning performance of the models, and it was observed that all models (Bagging, RF, Stacking, and CNN) exhibited favorable convergence and achieved a low MSE value as the dataset size increased, indicating their suitability as well-fitted models. Furthermore, this study employed the performance of the models based on statistical evaluation metrics such as (R^2, RMSE, MSE, MAE) for all the models in both the training and testing phases.

The evaluation metrics demonstrated that the Stacking model exhibited superior performance compared to all other models, displaying the best fit and consistent correlation with predicted values in both the training and testing phases. The RF model demonstrated a similar level of performance to the Stacking model, while the Bagging and CNN models exhibited lower performance. A comparative graphical assessment was conducted using a robust Taylor diagram to quantitatively analyze the correlations between actual and predicted SM content. It was observed that the Stacking and RF models closely aligned with the highest correlation and RMSE value of the actual SM data. The models were ranked as Bagging $<$ CNN $<$Random forest$<$ Stacking. These findings align with previous studies, as reported by Das et al., (2022), where optical, thermal, and microwave remote sensing data were employed to map soil surface moisture in a semi-arid region at a spatial resolution of 30 m[17]. Machine learning techniques, including bagging, boosting, and stacking, were utilized, and the authors concluded that the stacking of cubist, gradient boosting machine (GBM), and RF algorithms yielded superior results compared to individual algorithms. Similarly, Granata et al., (2022) reported that the stacked model outperformed all other model in predicting daily volumetric soil water content for both sets of input variables, achieving the highest performance with an R^2 of 0.962[77].

The inference drawn from the rigorous analysis of various model evaluation processes and diagnostic tests is that the Stacking model serves as a powerful artificial intelligence tool for accurately and reliably emulating actual SM content in a study area. The superior performance of the Stacking model can be attributed to the synergy among the base learners, which helps mitigate overfitting and bias that individual machine learning models may exhibit. The underperformance of the Bagging model can be attributed to its relative simplicity and its better suitability for problems with small training datasets [65]. As relatively large number dataset was generated in this study ensemble methods like Stacking and RF were better suited for the task.

While the CNN model's performance was comparatively lower than that of the Stacking and RF models, it does not imply that it performance is poor. The CNN model's performance was likely influenced by the dataset size, as CNN models typically require large datasets to achieve optimal performance. A study conducted by Ng et al., (2019) demonstrated that when the training sample size was small, the Partial Least Squares Regression (PLSR) model outperformed the CNN model[67]. However, as the sample size exceeded 2000, the CNN model began to surpass the PLSR and cubist models, exhibiting improved performance with further increases in sample size. This difference in performance can be attributed to the data-intensive nature of deep learning models, where a limited number of training samples may not provide sufficient information for effective model training. Consequently, predictions may be less accurate with smaller training sets [67,78]. Nonetheless, the CNN model yielded relatively good results in this study despite the small dataset (84 samples), indicating that tuning its hyperparameters, a critical step in building any machine learning model, may have contributed to its success. Additionally, the utilization of nonlinear activation functions in deep learning models enabled the extraction of nonlinear features and processing of complex relationships between soil moisture and mixed spectra [79]. Overall, the analysis highlights the Stacking model as the most suitable AI tool for accurately emulating actual SM content in the study area, considering the available dataset and other relevant factors.

This study provides valuable insights into the dynamic nature of soil moisture distribution in Agartala City and the factors influencing soil moisture content. The spatial mapping of predicted SM content, generated using the best model, offers a clear depiction of areas with high and low soil moisture content in the study area. The northwestern region of Agartala city exhibited higher soil moisture content, primarily due to factors such as the dense canopy and vegetation cover, as well as loamy and clayey soil texture. Conversely, the northern part of the region demonstrated lower soil moisture content, mainly influenced by higher elevation, sandy and loamy soil texture, and sparse vegetation.

The study successfully implemented XAI using an ensemble bagging model to improve soil moisture management strategies. The ensemble bagging model was optimized through grid search, and a 3D plot of hyperparameters was used to identify the best configuration for the model. The best model demonstrated high accuracy, with metrics such as MAE, MSE, RMSE, R^2 Score, and Explained Variance Score indicating its robust performance in capturing underlying patterns in the data.

Furthermore, global-level analysis using SHAP revealed that humidity, atmospheric temperature, TWI, σ_0VH, and σ_0VV were the most influential factors affecting soil moisture predictions. This scientific analysis provides valuable insights into the key factors influencing soil moisture, guiding the development of effective management strategies. Maintaining optimal humidity and temperature levels, considering TWI for water retention, are highlighted as crucial steps in soil moisture management. Additionally, the study employed LIME for local-level analysis, providing site-specific explanations and parameter importance for 16 sites. The findings showed that humidity had the most significant impact on soil moisture across all sites, and other factors like temperature, DC, NDVI, LST, and TWI also influenced soil moisture content at varying degrees. These results offer essential information for localized decision-making and will help stakeholders implement tailored strategies to optimize soil moisture levels.

Finally, the scientific application of XAI, ensemble bagging model, SHAP, and LIME provided valuable insights into the factors influencing soil moisture predictions at both global and local levels. The study's findings hold significant promise for the development of effective soil moisture management strategies, ultimately benefiting agricultural productivity and promoting sustainable land use practices. These findings affirm that NDVI and atmospheric temperature play crucial roles in comprehending the variations in soil moisture under diverse climatic conditions. This could be due to the significant impact of vegetation growth, which is highly influenced by the availability of soil moisture for plant development. Moreover, various human activities may also lead to an increase in atmospheric temperature, substantially influencing evaporation rates and causing changes in soil moisture content [80]. Additionally, the relationships between these variables can be complex and depend on several factors, such as regional climate patterns, soil characteristics, specific vegetation types present, and land management practices.

The significance of reading and undertaking this study lies in its

focus on a crucial issue related to sustainable agriculture. Notably, the study employs advanced machine learning techniques and XAI analysis, representing an innovative approach to understanding the complex relationships between input parameters and soil moisture content. This approach encourages stakeholders and researchers to adopt similar methodologies in other regions to enhance water resource management practices. Furthermore, this study contributes significant insights into the dynamic nature of soil moisture distribution in Agartala City and the factors influencing soil moisture content. The study underscores the importance of considering various factors such as vegetation cover, soil texture, and elevation when formulating strategies for sustainable agriculture practices. The findings and results presented in this study can guide policymakers in developing effective strategies for sustainable agriculture, preventing water scarcity, and improving water resource management.

5. Conclusion

Surface soil moisture is a vital parameter for the proper functioning of our ecosystem. In this study, we utilized a deep-learning approach with multi-sensor remote sensing and terrain data to estimate and predict surface soil moisture. By collecting in-situ data through IoT-based sensors and using ensemble-based machine learning techniques, such as bagging, random forest, and convolutional neural networks, we were able to identify 'stacking' as the best model for surface soil moisture prediction, as evidenced by various evaluation matrices, linear regression graphs, Taylor diagrams, and learning curves. To enhance transparency and accountability in soil moisture prediction for water management and decision-making processes, an explainable artificial intelligence (XAI) based SHapley Additive exPlanations (SHAP) and Local Interpretable Model Explanation (LIME) method were used to understand the influence on variables in predicting SM content both globally and locally. The result of SHAP suggest humidity and atmospheric temperature has the most sensitive parameters and have the highest influence for predicting surface soil moisture globally. While, LIME model also support the SHAP result while considering humidity as the most sensitive parameter locally. The framework developed holds significant potential for decision-makers and water resource management personnel, offering valuable support for promoting digital agriculture among farmers. Future studies should integrate soil hydraulic parameters into surface soil moisture prediction models to enhance accuracy. Incorporating various satellite-based soil moisture products such as ESA SMOS (Soil Moisture and Ocean Salinity), Sentinel-1 SAR (Synthetic Aperture Radar), EOS-04 (RISAT-1A), GNSS-R, AMSR-E (Advanced Microwave Scanning Radiometer for EOS), and model-based products like GLDAS (Global Land Data Assimilation System) and NLDAS (North American Land Data Assimilation System) will further improve reliability. Additionally, developing of a real-time monitoring and prediction system, possibly in the form of a mobile application, can provide farmers with timely information on soil moisture trends and patterns in soil moisture levels, enabling them to make informed decisions about irrigation and water management. As well as the framework developed can also be used to estimate and predict various field parameters related to irrigation, water resources and hydrology.

CRediT authorship contribution statement

Abhigyan Chakraborty: Writing – review & editing, Writing – original draft, Data curation. **Krishanu Podder:** Software. **Swapan Talukdar:** Writing – original draft, Methodology. **Atiqur Rahman:** Writing – review & editing, Writing – original draft, Supervision. **Umesh Mishra:** Supervision, Resources. **Santanu Mallik:** Writing – review & editing, Writing – original draft, Methodology.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Acknowledgements

The authors extend their gratitude to the College of Agriculture, Lembhuchera, Tripura for generously providing the necessary resources to conduct laboratory testing and analysis. Additionally, they would like to acknowledge the North Eastern Space Application Center (NESAC), Department of Space, and Government of India for sharing the Land Use/Land Cover (LULC) map of the designated study area.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.asoc.2025.113406.

Data Availability

Data will be made available on request.

References

[1] R.L. Bras, Complexity and organization in hydrology: a personal view, Water Resour. Res. 51 (2015) 6532–6548, https://doi.org/10.1002/2015WR016958.
[2] I. Ali, F. Greifeneder, J. Stamenkovic, M. Neumann, C. Notarnicola, Review of machine learning approaches for biomass and soil moisture retrievals from remote sensing data, Remote Sens 7 (2015) 16398–16421, https://doi.org/10.3390/rs71215841.
[3] E. Babaeian, S. Paheding, N. Siddique, V.K. Devabhaktuni, M. Tuller, Estimation of root zone soil moisture from ground and remotely sensed soil information with multisensor data fusion and automated machine learning, Remote Sens. Environ. 260 (2021), https://doi.org/10.1016/j.rse.2021.112434.
[4] P. Gururaj, P. Umesh, A. Shetty, Modeling of surface soil moisture using C-band Sar data over bare fields in the tropical semi-arid region of India, Appl. Geomat. 13 (2021) 555–564, https://doi.org/10.1007/s12518-021-00370-7.
[5] H. Adab, R. Morbidelli, C. Saltalippi, M. Moradian, G.A.F. Ghalhari, Machine learning to estimate surface soil moisture from remote sensing data, Water (Switz.) 12 (2020) 1–28, https://doi.org/10.3390/w12113223.
[6] S. Nichols, Review and evaluation of remote sensing methods for soil-moisture estimation, J. Photonics Energy 2 (2011) 28001, https://doi.org/10.1117/1.3534910.
[7] S.I. Seneviratne, T. Corti, E.L. Davin, M. Hirschi, E.B. Jaeger, I. Lehner, B. Orlowsky, A.J. Teuling, Investigating soil moisture–climate interactions in a changing climate: a review, EarthSci. Rev. 99 (2010) 125–161, https://doi.org/10.1016/J.EARSCIREV.2010.02.004.
[8] A.S. Abowarda, L. Bai, C. Zhang, D. Long, X. Li, Q. Huang, Z. Sun, Generating surface soil moisture at 30 m spatial resolution using both data fusion and machine learning toward better water resources management at the field scale, Remote Sens. Environ. 255 (2021), https://doi.org/10.1016/j.rse.2021.112301.
[9] J. Yue, J. Tian, Q. Tian, K. Xu, N. Xu, Development of soil moisture indices from differences in water absorption between shortwave-infrared bands, ISPRS J. Photogramm. Remote Sens 154 (2019) 216–230, https://doi.org/10.1016/j.isprsjprs.2019.06.012.
[10] A. Goap, D. Sharma, A.K. Shukla, C. Rama Krishna, An IoT based smart irrigation management system using machine learning and open source technologies, Comput. Electron. Agric. 155 (2018) 41–49, https://doi.org/10.1016/j.compag.2018.09.040.
[11] B. Prasath, M. Akila, IoT-based pest detection and classification using deep features with enhanced deep learning strategies, Eng. Appl. Artif. Intell. 121 (2023) 105985, https://doi.org/10.1016/j.engappai.2023.105985.
[12] J.Á. Martín-Baos, L. Rodriguez-Benitez, R. García-Ródenas, J. Liu, IoT based monitoring of air quality and traffic using regression analysis, Appl. Soft Comput. 115 (2022) 108282, https://doi.org/10.1016/j.asoc.2021.108282.
[13] R. Kumar, M.Pallikonda Rajasekaran, An IoT based patient monitoring system using raspberry pi, Int. Conf. Comput. Technol. Intell. Data Eng. ICCTIDE 2016. 2016 (2016), https://doi.org/10.1109/ICCTIDE.2016.7725378.
[14] S. Puengsungwan, IoT based soil moisture sensor for smart farming, proc. 2020, Int. Conf. Power Energy Innov. ICPEI 2020. (2020) 221–224, https://doi.org/10.1109/ICPEI49860.2020.9431455.
[15] H. Vereecken, J.A. Huisman, Y. Pachepsky, C. Montzka, J. van der Kruk, H. Bogena, L. Weihermüller, M. Herbst, G. Martinez, J. Vanderborght, On the spatio-temporal dynamics of soil moisture at the field scale, J. Hydrol. 516 (2014) 76–96, https://doi.org/10.1016/j.jhydrol.2013.11.061.

[16] D. Zhang, R. Tang, W. Zhao, B. Tang, H. Wu, K. Shao, Z.L. Li, Surface soil water content estimation from thermal remote sensing based on the temporal variation of land surface temperature, 2014, Vol. 6, Pages 3170-3187, Remote Sens 6 (2014) 3170–3187, https://doi.org/10.3390/RS6043170.

[17] B. Das, P. Rathore, D. Roy, D. Chakraborty, R.S. Jatav, D. Sethi, P. Kumar, Comparison of bagging, boosting and stacking algorithms for surface soil moisture mapping using optical-thermal-microwave remote sensing synergies, Catena 217 (2022) 106485, https://doi.org/10.1016/j.catena.2022.106485.

[18] F. Greifeneder, C. Notarnicola, W. Wagner, A machine learning-based approach for surface soil moisture estimations with google earth engine, Remote Sens 13 (2021), https://doi.org/10.3390/rs13112099.

[19] T.J. Schmugge, Remote sensing of soil moisture: recent advances, IEEE Trans. Geosci. Remote Sens. GE 21 (1983) 336–344, https://doi.org/10.1109/TGRS.1983.350563.

[20] D. Zhang, G. Zhou, Estimation of soil moisture from optical and thermal remote sensing: a review, Sens. (Basel) 16 (2016), https://doi.org/10.3390/S16081308.

[21] K.C. Kornelsen, P. Coulibaly, Advances in soil moisture retrieval from synthetic aperture radar and hydrological applications, J. Hydrol. 476 (2013) 460–489, https://doi.org/10.1016/J.JHYDROL.2012.10.044.

[22] Y. Zhang, S. Liang, Z. Zhu, H. Ma, T. He, Soil moisture content retrieval from landsat 8 data using ensemble learning, ISPRS J. Photogramm. Remote Sens 185 (2022) 32–47, https://doi.org/10.1016/j.isprsjprs.2022.01.005.

[23] S.K. Chan, R. Bindlish, P.E. O'Neill, E. Njoku, T. Jackson, A. Colliander, F. Chen, M. Burgin, S. Dunbar, J. Piepmeier, S. Yueh, D. Entekhabi, M.H. Cosh, T. Caldwell, J. Walker, X. Wu, A. Berg, T. Rowlandson, A. Pacheco, H. McNairn, M. Thibeault, J. Martinez-Fernandez, A. Gonzalez-Zamora, M. Seyfried, D. Bosch, P. Starks, D. Goodrich, J. Prueger, M. Palecki, E.E. Small, M. Zreda, J.C. Calvet, W.T. Crow, Y. Kerr, Assessment of the SMAP passive soil moisture product, IEEE Trans. Geosci. Remote Sens 54 (2016) 4994–5007, https://doi.org/10.1109/TGRS.2016.2561938.

[24] P. Abbaszadeh, H. Moradkhani, X. Zhan, Downscaling SMAP radiometer soil moisture over the CONUS using an ensemble learning method, Water Resour. Res. 55 (2019) 324–344, https://doi.org/10.1029/2018WR023354.

[25] S. Ahmad, A. Kalra, H. Stephen, Estimating soil moisture using remote sensing data: a machine learning approach, Adv. Water Resour. 33 (2010) 69–80, https://doi.org/10.1016/j.advwatres.2009.10.008.

[26] C. Carranza, C. Nolet, M. Pezij, M. van der Ploeg, Root zone soil moisture estimation with random forest, J. Hydrol. 593 (2021), https://doi.org/10.1016/j.jhydrol.2020.125840.

[27] K.C. Kornelsen, P. Coulibaly, Root-zone soil moisture estimation using data-driven methods, Water Resour. Res. 50 (2014) 2946–2962, https://doi.org/10.1002/2013WR014127.

[28] L. He, Y. Cheng, Y. Li, F. Li, K. Fan, Y. Li, An improved method for soil moisture monitoring with ensemble learning methods over the Tibetan plateau, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens 14 (2021) 2833–2844, https://doi.org/10.1109/JSTARS.2021.3058325.

[29] M. Emadi, R. Taghizadeh-Mehrjardi, A. Cherati, M. Danesh, A. Mosavi, T. Scholten, Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran, Remote Sens 12 (2020), https://doi.org/10.3390/rs12142234.

[30] J. Cao, Z. Zhang, F. Tao, L. Zhang, Y. Luo, J. Zhang, J. Han, J. Xie, Integrating Multi-Source data for rice yield prediction across China using machine learning and deep learning approaches, Agric. For. Meteor. 297 (2021) 108275, https://doi.org/10.1016/j.agrformet.2020.108275.

[31] S. Mallik, A. Chakraborty, U. Mishra, N. Paul, Prediction of irrigation water suitability using geospatial computing approach: a case study of agartala city, India, Environ. Sci. Pollut. Res. (2022), https://doi.org/10.1007/s11356-022-21232-8.

[32] J.B. Mohapatra, P. Jha, M.K. Jha, S. Biswal, Efficacy of machine learning techniques in predicting groundwater fluctuations in agro-ecological zones of India, Sci. Total Environ. 785 (2021) 147319, https://doi.org/10.1016/j.scitotenv.2021.147319.

[33] H. Afzaal, A.A. Farooque, F. Abbas, B. Acharya, T. Esau, Groundwater estimation from major physical hydrology components using artificial neural networks and deep learning, Water (Switz.) 12 (2020), https://doi.org/10.3390/w12010005.

[34] M. Panahi, N. Sadhasivam, H.R. Pourghasemi, F. Rezaie, S. Lee, Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR), J. Hydrol. 588 (2020) 125033, https://doi.org/10.1016/j.jhydrol.2020.125033.

[35] A. Passah, K. Amitab, D. Kandar, Sar image despeckling using deep CNN, IET Image Process 15 (2021) 1285–1297, https://doi.org/10.1049/ipr2.12104.

[36] T. van Klompenburg, A. Kassahun, C. Catal, Crop yield prediction using machine learning: a systematic literature review, Comput. Electron. Agric. 177 (2020) 105709, https://doi.org/10.1016/j.compag.2020.105709.

[37] A. Temenos, N. Temenos, M. Kaselimi, A. Doulamis, N. Doulamis, Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP, IEEE Geosci. Remote Sens. Lett. 20 (2023) 1–5.

[38] R. Jena, A. Shanableh, R. Al-Ruzouq, B. Pradhan, M.B.A. Gibril, M.A. Khalil, O. Ghorbanzadeh, G.P. Ganapathy, P. Ghamisi, Explainable artificial intelligence (XAI) model for earthquake spatial probability assessment in Arabian Peninsula, Remote Sens 15 (2023) 2248.

[39] B. Pradhan, A. Dikshit, S. Lee, H. Kim, An explainable AI (XAI) model for landslide susceptibility modeling, Appl. Soft Comput. 142 (2023) 110324, https://doi.org/10.1016/j.asoc.2023.110324.

[40] R. Saleem, B. Yuan, F. Kurugollu, A. Anjum, L. Liu, Explaining deep neural networks: a survey on the global interpretation methods, Neurocomputing 513 (2022) 165–180, https://doi.org/10.1016/j.neucom.2022.09.129.

[41] S. Mallik, T. Bhowmik, U. Mishra, N. Paul, Mapping and prediction of soil organic carbon by an advanced geostatistical technique using remote sensing and terrain data, Geocarto Int (2020), https://doi.org/10.1080/10106049.2020.1815864.

[42] S. Li, L.Da Xu, S. Zhao, The Internet of things: a survey, Inf. Syst. Front 17 (2015) 243–259, https://doi.org/10.1007/S10796-014-9492-7/FIGURES/7.

[43] R.K. Kodali, A. Sahu, An IoT based soil moisture monitoring on Losant platform, (2017). https://doi.org/10.1109/IC3I.2016.7918063.

[44] A. Bhardwaj, V. Dagar, M.O. Khan, A. Aggarwal, R. Alvarado, M. Kumar, M. Irfan, R. Proshad, Smart IoT and machine Learning-based framework for water quality assessment and device component monitoring, Environ. Sci. Pollut. Res. 29 (2022) 46018–46036, https://doi.org/10.1007/s11356-022-19014-3.

[45] H. Khurshid, R. Mumtaz, N. Alvi, A. Haque, S. Mumtaz, F. Shafait, S. Ahmed, M. I. Malik, A. Dengel, Bacterial prediction using Internet of things (IoT) and machine learning, Environ. Monit. Assess. 194 (2022), https://doi.org/10.1007/s10661-021-09698-4.

[46] S.V. Gaikwad, A.D. Vibhute, K.V. Kale, S.C. Mehrotra, An innovative IoT based system for precision farming, Comput. Electron. Agric. 187 (2021), https://doi.org/10.1016/j.compag.2021.106291.

[47] M. Tomar, T. Patidar, Development of a low cost soil moisture sensor, Proc. Int. Conf. Vis. Towar. Emerg. Trends Commun. Netw. ViTECoN 2019 (2019), https://doi.org/10.1109/VITECON.2019.8899399.

[48] Y. Jia, S. Jin, P. Savi, Q. Yan, W. Li, Modeling and theoretical analysis of gnss-r soil moisture retrieval based on the random forest and support vector machine learning approach, Remote Sens 12 (2020) 1–24, https://doi.org/10.3390/rs12223679.

[49] M.T. Hallikainen, F.T. Ulabz, M.C. Dobson, M.A. El-Rayes, L.K. Wu, Microwave dielectric behavior of wet Soil-Part I: empirical models and experimental observations, IEEE Trans. Geosci. Remote Sens. GE 23 (1985) 25–34, https://doi.org/10.1109/TGRS.1985.289497.

[50] M.C. Dobson, F.T. Ulaby, M.T. Hallikainen, M.A. El-Rayes, Microw. Dielectr. Behav. Wet. SoilPart II Dielectr. Mixing Models (1985).

[51] L. Wang, J.J. Qu, Satellite remote sensing applications for surface soil moisture monitoring: a review, Front. Earth Sci. China 2009 32. 3 (2009) 237–247, https://doi.org/10.1007/S11707-009-0023-7.

[52] K.S. Rawat, S.K. Singh, R.L. Ray, An integrated approach to estimate surface soil moisture in agricultural lands, Geocarto Int 36 (2021) 1646–1664, https://doi.org/10.1080/10106049.2019.1678674.

[53] I. Sandholt, K. Rasmussen, J. Andersen, A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status, Remote Sens. Environ. 79 (2002) 213–224, https://doi.org/10.1016/S0034-4257(01)00274-7.

[54] S. Mallik, T. Bhowmik, U. Mishra, N. Paul, Mapping and prediction of soil organic carbon by an advanced geostatistical technique using remote sensing and terrain data, Geocarto Int (2020), https://doi.org/10.1080/10106049.2020.1815864.

[55] U. Avdan, G. Jovanovska, Algorithm for automated mapping of land surface temperature using LANDSAT 8 satellite data, J. Sens. 2016 (2016), https://doi.org/10.1155/2016/1480307.

[56] S.A. Naghibi, H.R. Pourghasemi, B. Dixon, GIS-based groundwater potential mapping using boosted regression tree, classification and regression tree, and random forest machine learning models in Iran, Environ. Monit. Assess. 188 (2016) 1–27, https://doi.org/10.1007/s10661-015-5049-6.

[57] S. Pal, S. Kundu, S. Mahato, Groundwater potential zones for sustainable management plans in a river basin of India and Bangladesh, J. Clean. Prod. 257 (2020) 120311, https://doi.org/10.1016/j.jclepro.2020.120311.

[58] S.A. Shahriar, I. Kayes, K. Hasan, M.A. Salam, S. Chowdhury, Applicability of machine learning in modeling of atmospheric particle pollution in Bangladesh, Air Qual. Atmos. Heal (2020), https://doi.org/10.1007/s11869-020-00878-8.

[59] J.P. Monteiro, D. Ramos, D. Carneiro, F. Duarte, J.M. Fernandes, P. Novais, Meta-learning and the new challenges of machine learning, Int. J. Intell. Syst. 36 (2021) 6240–6272, https://doi.org/10.1002/int.22549.

[60] S. Mallik, S. Das, A. Chakraborty, U. Mishra, S. Talukdar, S. Bera, G.V. Ramana, Prediction of non-carcinogenic health risk using hybrid monte Carlo-machine learning approach, Hum. Ecol. Risk Assess. Int. J. 0 (2023) 1–24, https://doi.org/10.1080/10807039.2023.2188417.

[61] O. Rahmati, S.A. Naghibi, H. Shahabi, D.T. Bui, B. Pradhan, A. Azareh, E. Rafiei-Sardooi, A.N. Samani, A.M. Melesse, Groundwater spring potential modelling: comprising the capability and robustness of three different modeling approaches, J. Hydrol. 565 (2018) 248–261, https://doi.org/10.1016/j.jhydrol.2018.08.027.

[62] W. Chen, P. Tsangaratos, I. Ilia, Z. Duan, X. Chen, Groundwater spring potential mapping using population-based evolutionary algorithms and data mining methods, Sci. Total Environ. 684 (2019) 31–49, https://doi.org/10.1016/j.scitotenv.2019.05.312.

[63] H.R. Pourghasemi, M. Rossi, Landslide susceptibility modeling in a landslide prone area in mazandarn province, north of Iran: a comparison between GLM, GAM, Mars, and M-AHP methods, Theor. Appl. Clim. 130 (2017) 609–633, https://doi.org/10.1007/s00704-016-1919-2.

[64] M.D. de Menezes, F.H.A. Bispo, W.M. Faria, M.G.M. Gonçalves, N. Curi, L.R. G. Guilherme, Modeling arsenic content in Brazilian soils: what is relevant? Sci. Total Environ. 712 (2020) 136511 https://doi.org/10.1016/j.scitotenv.2020.136511.

[65] M. Zounemat-Kermani, O. Batelaan, M. Fadaee, R. Hinkelmann, Ensemble machine learning paradigms in hydrology: a review, J. Hydrol. 598 (2021) 126266, https://doi.org/10.1016/j.jhydrol.2021.126266.

[66] R. Taghizadeh-Mehrjardi, K. Schmidt, A. Amirian-Chakan, T. Rentschler, M. Zeraatpisheh, F. Sarmadian, R. Valavi, N. Davatgar, T. Behrens, T. Scholten, Improving the spatial prediction of soil organic carbon content in two contrasting

climatic regions by stacking machine learning models and rescanning covariate space, Remote Sens 12 (2020), https://doi.org/10.3390/rs12071095.
[67] W. Ng, B. Minasny, M. Montazerolghaem, J. Padarian, R. Ferguson, S. Bailey, A. B. McBratney, Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra, Geoderma 352 (2019) 251–267, https://doi.org/10.1016/j.geoderma.2019.06.016.
[68] Y. Chen, L. Song, Y. Liu, L. Yang, D. Li, A review of the artificial neural network models for water quality prediction, Appl. Sci. 10 (2020), https://doi.org/10.3390/app10175776.
[69] T. van Klompenburg, A. Kassahun, C. Catal, Crop yield prediction using machine learning: a systematic literature review, Comput. Electron. Agric. 177 (2020) 105709, https://doi.org/10.1016/j.compag.2020.105709.
[70] S. Saha, R. Sarkar, G. Thapa, J. Roy, Modeling gully erosion susceptibility in phuentsholing, Bhutan using deep learning and basic machine learning algorithms, Environ. Earth Sci. 80 (2021) 1–21, https://doi.org/10.1007/s12665-021-09599-2.
[71] S. Mallik, T. Bhowmik, U. Mishra, N. Paul, Local scale groundwater vulnerability assessment with an improved DRASTIC model, Nat. Resour. Res. (2021), https://doi.org/10.1007/s11053-021-09839-z.
[72] D. Doljak, G. Stanojević, Evaluation of natural conditions for site selection of ground-mounted photovoltaic power plants in Serbia, Energy 127 (2017) 291–300, https://doi.org/10.1016/j.energy.2017.03.140.
[73] C. Cammalleri, F. Micale, J. Vogt, A novel soil moisture-based drought severity index (DSI) combining water deficit magnitude and frequency, Hydrol. Process 30 (2016) 289–301, https://doi.org/10.1002/hyp.10578.
[74] A. Saha, M. Patil, V.C. Goyal, D.S. Rathore, Assessment and impact of soil moisture index in agricultural drought estimation using remote sensing and GIS techniques, in: MDPI AG, 2018, p. 2, https://doi.org/10.3390/ecws-3-05802.
[75] M.W. Raduła, T.H. Szymura, M. Szymura, Topographic wetness index explains soil moisture better than bioindication with Ellenberg's indicator values, Ecol. Indic. 85 (2018) 172–179, https://doi.org/10.1016/j.ecolind.2017.10.011.
[76] M. Dahim, S. Alqadhi, J. Mallick, Enhancing landslide management with hyper-tuned machine learning and deep learning models: Predicting susceptibility and analyzing sensitivity and uncertainty, (2023) 1–22. https://doi.org/10.3389/fevo.2023.1108924.
[77] F. Granata, F. Di Nunno, M. Najafzadeh, I. Demir, A stacked machine learning algorithm for Multi-Step ahead prediction of soil moisture, Hydrology 10 (2022) 1.
[78] J. Padarian, B. Minasny, A.B. McBratney, Using deep learning to predict soil properties from regional spectral data, Geoderma Reg. 16 (2019) e00198.
[79] W. Rawat, Z. Wang, Deep convolutional neural networks for image classification: a comprehensive review, Neural Comput. 29 (2017) 2352–2449.
[80] M. Sharma, P. Bangotra, A.S. Gautam, S. Gautam, Sensitivity of normalized difference vegetation index (NDVI) to land surface temperature, soil moisture and precipitation over district gautam buddh nagar, UP, India, Stoch. Environ. Res. Risk Assess. 36 (2022) 1779–1789, https://doi.org/10.1007/s00477-021-02066-1.

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

IoT based soil moisture measurement and type prediction using advanced regression and machine learning models / Sazzad, M. M., Ahmed, T., Kibria, G., & Khan, I.

Scientific Reports
Volume 15 (2025) 35730 Pages 1-14
https://doi.org/10.1038/s41598-025-19444-2
(Database: Nature Portfolio)

scientific reports

Check for updates

OPEN

IoT based soil moisture measurement and type prediction using advanced regression and machine learning models

Md. Mahmud Sazzad[1,3], Tanvir Ahmed[1,3]✉, Golam Kibria[1,3] & Ishmam Khan[2,3]

Measuring soil moisture has a big impact on resource efficiency and decision-making in geotechnical engineering, agriculture, and environmental sustainability. This study presents a revolutionary Internet of Things-based method for predicting soil type and moisture in real time. Water content was verified using the traditional oven-dry method for accuracy evaluation, and capacitance measurements were gathered using a sensor to generate a custom dataset in the lab. The method outperformed linear regression in water content prediction, achieving 96.49% accuracy using an Excel logarithmic regression equation. Furthermore, a machine learning model that employed polynomial regression was able to measure the water content of the soil and predict values for fine, medium-coarse, and coarse sand types. With an R2 score of 0.79, the model can account for almost 79% of the variation in water content and produce 1.71% of MAE(Mean Absolute of Error), indicating a strong relationship between capacitance and water content. On the expanded dataset, Random Forest classifier was chosen for classification, which correctly identified the intended soil types with an accuracy of roughly 97.77%. By combining sensor data with sophisticated algorithms, the suggested methodology makes it possible to analyze soil qualities effectively and non-destructively. This scalable method offers substantial potential for environmental management, soil monitoring, and precision agriculture and is flexible enough for the creation of mobile applications. Predictive modeling and real-time data processing combined improve resource management effectiveness while lowering the need for human intervention. In order to improve forecast accuracy and application and support more environmentally friendly farming methods and environmental monitoring, further research will surely give efforts to enrich the dataset, incorporate a variety of soil types, and take environmental elements like temperature and salinity into consideration.

Keywords Soil moisture content, Laboratory tests, Capacitive sensor, Oven dry method, Calibration, Artificial intelligence, Machine learning, Classification, Random forest classifier, Regression, Polynomial regression, Internet of things

The moisture content of soil is essential to many physical, chemical and biological processes in both natural ecosystems and man-made structures. It indicates the percentage of water in the soil, and a deep understanding of its dynamics is necessary to maintain the fragile soil-water balance. Accurate and real-time soil moisture monitoring is essential to improve agricultural productivity, ensure infrastructure durability, and promote environmental sustainability. In agriculture, it is essential to maintain soil health, increase crop yields, and protect water supplies. Monitoring soil moisture is crucial in civil engineering to prevent freeze-thaw cycles and chloride intrusion from damaging structures. Therefore, in order to protect the integrity of the environment and the health of the infrastructure, precise and efficient methods for measuring soil moisture must be developed.

Soil moisture was crucial to increase crop productivity, soil health, and agricultural water use efficiency. Although traditional gravimetric methods were believed to be accurate, their inability to provide real-time data made them inappropriate for dynamic scenarios. Capacitance-based sensors were developed as a continuous and economical means of monitoring the dielectric behavior of soils that were affected by the presence of moisture

[1]Department of Civil Engineering, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh. [2]Department of Computer Science and Engineering, Rajshahi University of Engineering & Technology, Rajshahi 6204, Bangladesh. [3]These authors contributed equally to this work. ✉email: tanvirahmedb37@gmail.com

levels[1]. Those sensors translate moisture levels into measurable signals (e.g., frequency or pulse width modulation) for real-time monitoring. While[2] proposes a new cylindrical electrode design for bulk materials, optimizing sensor geometry for granular media such as soil,[3] evaluates the practical deployment challenges of such sensors in IoT-based irrigation systems, including calibration drift and field variability. Together, these studies highlight the trade-offs between sensor innovation[2] and operational reliability[3] in agricultural applications. Their long-term dependability and suitability for clay and silty soils were still restricted. To develop universal, real-time sensing systems that might be used in a variety of agricultural contexts, more study was required. CWM-four and other calibration-free models demonstrated potential[4]. For the compaction management process in geo-technical engineering, precise soil moisture measurement was essential. For fine-grained soils, capacitive and resistive sensors provide a portable, affordable, and real-time substitute for oven-drying methods. However, environmental elements such as soil type, compaction, porosity, temperature, and salt may affect their accuracy. These sensors demonstrated the ability to enhance geo-technical processes, despite problems with coarse-grained soils and calibration difficulties for various soil types[5]. Despite the dearth of existing calibration methodologies, it was demonstrated that volumetric water content accuracy was improved by soil-specific calibration, ensuring great repeatability and low variability. Small, high-frequency capacitive sensors can be used to assess the moisture content of urban soil[6]. Among gravimetrically calibrated soil moisture sensors, low-cost irrigation sensors were shown to have the lowest error, especially for sandy and sandy loam soils[7].

The electromagnetic fields produced by electrodes, which were influenced by the dielectric characteristics of the soil, were tracked using capacitance-based sensors. Because of their adaptability, they were able to determine the moisture content of the soil at different depths[8]. Due to their accuracy, speed, and low energy consumption, these sensors were widely utilized in agricultural and grain storage[9]. However, their efficacy was occasionally hindered by salinity, soil texture, environmental diversity, and sensor-to-sensor variability[6,10]. It has been shown that calibration techniques like quadratic regression increase the accuracy of these sensors when compared to linear models, particularly when soil-specific changes are present[11,12]. In order to improve real-time monitoring and lower mistakes, techniques like logistic regression were used to lessen the impact of electrical conductivity (EC) on sensor accuracy, even in large agricultural settings. Despite being inexpensive, this method was ineffective in high-salinity and organic-rich soils and ignored elements such soil porosity and texture. Despite these limitations, the technique shown potential for accurate soil moisture monitoring[13].

Recent studies (e.g.,[3,12]) have used piecewise linear regression and temperature compensation algorithms (e.g., based on Arrhenius equation adjustments) to mitigate sensor drift caused by temperature fluctuations (±0.5% accuracy loss per ∘C in uncontrolled environments). These refinements improved volumetric water content estimates across sand to loam ratios, though challenges persist in clay-rich soils.

For civil engineering projects to last, it was essential to keep an eye on soil moisture throughout freeze-thaw cycles and chloride intrusion. Advanced methods such nuclear magnetic resonance (NMR), electrical impedance spectroscopy (EIS), and X-ray imaging were used to generate high-resolution data; nevertheless, their field applicability was constrained by their cost, intricacy, and limited usefulness[14]. Recent advancements have combined low-dimensional materials, like molybdenum disulfide (MoS_2), with hybrid sensing techniques to increase sensor sensitivity and reliability across a range of environmental conditions[15]. In order to solve problems specific to soil, data-driven calibration methods and hybrid models that integrate machine learning and sensor data showed promise[16]. Water-cut meters for near-linear transfer functions[17] , frequency-based grain probe moisture meters for agricultural use[18], and Internet of Things-enabled hydroponic drip fertigation systems[19]are examples of recent advancements in capacitance-based technologies. The need for additional study was highlighted by the persistence of issues such soil-type dependence and decreasing sensitivity at high volumetric water content (VWC) in spite of these developments[20]. In a range of environmental conditions, new materials such carbon-based composites could improve sensor performance and stability[21]. Combining IoT-enabled soil moisture monitoring with renewable energy sources increased the sustainability of agricultural practices[22].

While existing studies demonstrate the efficacy of capacitance-based sensors for soil moisture monitoring[1–7], several critical limitations remain unaddressed. First, most calibration models–such as linear or quadratic regression[11,12]–fail to capture the nonlinear relationship between capacitance and moisture across diverse soil types, particularly in high-plasticity or mixed-grain soils. Second, although hybrid machine learning and sensor approaches show promise[16,19], they often overlook practical real-time deployment due to high computational complexity. Third, the classification of soil types using low-cost sensors is still underexplored in laboratory settings, despite its relevance to geotechnical and environmental engineering. Additionally, most IoT-based systems focus solely on moisture content and either assume the soil type is known or do not offer a mechanism for concurrent identification, limiting their adaptability in real-world field scenarios[19,22].

This study is distinct in bridging these gaps through three key contributions: (i) it proposes a polynomial regression model tailored to nonlinear capacitance–moisture relationships, achieving improved accuracy (96.49%) over traditional linear/logarithmic models; (ii) it uniquely integrates real-time field measurements with laboratory-based soil classification using the same low-cost capacitance sensor, a feature not present in prior IoT-enabled systems; and (iii) it enables minimal recalibration needs through AI-driven adjustments, reducing the dependency on soil-specific calibration[5,11]. Unlike previous studies that treat moisture and soil type prediction as separate tasks, our approach offers simultaneous, dual-functionality–real-time water content estimation and soil type classification–within a unified, cost-effective framework. This makes it especially valuable for both agricultural applications and geotechnical evaluations.

The study presents a unique capacitance-based method for laboratory soil type detection and real-time, on-site soil moisture measurement. Using a custom dataset of three soil types and an AI-driven polynomial regression model, the process provides real-time water content readings on-site and predicts soil types in the lab with high accuracy. This approach, which has uses in environmental monitoring, agriculture, and geotechnical

engineering, blends complex calibration processes with machine learning algorithms to offer an economical and effective solution.

Methodology

In order to accurately detect and classify soil moisture, this study uses a methodical methodology to assess the performance of capacitive soil moisture sensors in combination with cutting-edge machine learning approaches. Key steps including sensor selection and calibration, soil sample preparation, controlled laboratory testing, and the use of both statistical and AI-driven models are all part of the technique, which is intended to guarantee reliable data collection and interpretation. Three different soil types are the specific focus of the study, and each phase is carefully planned to preserve scientific rigor and guarantee repeatability and reliability in measurements under a range of moisture circumstances.

Sensor identification and data recording process

Modern irrigation systems make considerable use of capacitive soil moisture sensors because of their remarkable dependability and longevity. These sensors' electrodes are covered in materials that resist corrosion, guaranteeing their durability and reliable operation. Compared to resistive sensors, capacitive sensors provide greater precision and accuracy while retaining dependable performance in a variety of soil types and field circumstances. They are perfect for long-term irrigation management applications because of their continuous sensitivity to changes in temperature and salinity as well as their increased longevity because of materials that are resistant to corrosion[23]. An electromagnetic method called capacitive sensing uses capacitive coupling to gauge the soil's dielectric permittivity. The apparent permittivity of soil is closely related to its water content because water's relative permittivity ($\varepsilon r \sim 80$) is much higher than that of other soil components, including organic matter ($\varepsilon r \sim 4$), mineral soil ($\varepsilon r \sim 2$-9), and air ($\varepsilon r \sim 1$)[5]. Bogena et al.'s (2007)[24] article goes into great detail into the electronic configurations and circuit diagrams of the majority of capacitive sensors.

The capacitance sensor's (Fig. 1) integrated voltage regulator chip, which guarantees steady performance across a wide input voltage range of 3.3 to 5.5 VDC, makes it compatible with 3.3V Arduino main control boards. It offers an output voltage range of 0 to 3.0 VDC and a PH2.54-3P interface for convenient connection. The sensor is adaptable to a wide range of applications due to its small size (98 mm × 23 mm). Using less than 5mA of current, the sensor produces about 1.5V for wet soil and 3.0V for dry soil when powered at 5V[25].

Capacitance sensors solely provide analog output signals, and the collection of these raw signals adheres to a specified protocol. The sensors transform the physical attribute of capacitance into an electrical signal represented as voltage, which is then converted into digital data via an analog-to-digital converter (ADC)

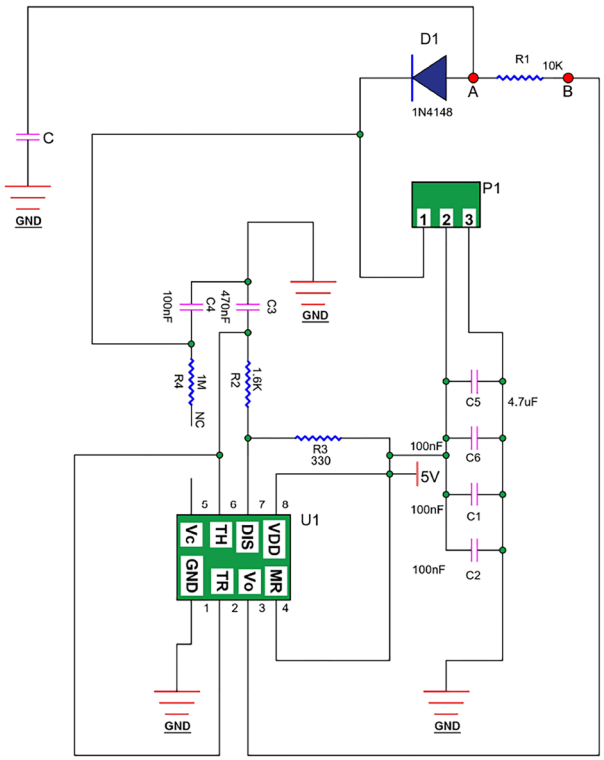


Fig. 1. The capacitive soil moisture sensor (version 1.2) circuit diagram, which includes a signal conditioning connection (P1), resistors, capacitors, and a diode (1N4148). The circuit produces voltage signals that are appropriate for microcontroller-based processing and correlate to the moisture content of the soil.

incorporated within an Arduino-compatible microcontroller board. This study employs the commonly deployed ESP32 microcontroller[26].

The ESP32 microcontroller is integrated into the bespoke solution for the capacitive sensor. These cost-effective, low-power 32-bit microcontrollers contain a 12-bit analog-to-digital converter (ADC), which simplifies the conversion of voltage levels from sensor output (ranging from 0 to 4.2 V) into 2^{12} (4096) discrete values, assuring precise digital representation.

The system employs an ESP32 microcontroller to interact with a Capacitive Soil Moisture Sensor v1.2 for real-time monitoring of soil moisture. The analog output of the sensor is read by the ESP32, which is powered by a 3.7V 18650 lithium-ion battery. This battery is managed using a Battery Management System (BMS) to ensure safe operation[27] . The ESP32 processes the sensor data before sending it over Wi-Fi to the Blynk server, where it is shown on the Blynk mobile app for remote control and monitoring (Fig. 2).

Soil properties

Four types of soil were selected for testing with water content sensors, consisting of four coarse-grained soils: Sand 1, Sand 2, Sand 3, and Sand 4. These soils are representative of materials commonly used in the construction of roadways, foundations, and embankment dams. Sand 1 and Sand 2 were sourced from a building construction site, while Sand 4 was previously utilized in a runway foundation in Rajshahi, Bangladesh. The grain size distribution of the soil samples was determined through sieve analysis, a standard method for characterizing coarse-grained soils. In this study, sieve analysis was conducted for Sand 1, Sand 2, Sand 3, and Sand 4. These soils are classified as uniform sands with minimal fine content. Sand 1 is coarse-grained, with a grain size (D50) of 0.439 mm, whereas Sand 2 is medium coarse-grained, with a grain size (D50) of 0.268 mm. Sand 3 and Sand 4 are classified as fine sands, with grain sizes (D50) of 0.190 mm and 0.138 mm, respectively. Among these, Sand 1 is well-graded, while the other sands are classified as poorly graded. After this for dataset purpose, Sand 3 and Sand 4 are merged. Fig. 3 presents the grain-size distribution curves of the selected soils.

Test configurations and methods

Laboratory-prepared specimens encompassing a wide range of water contents are used to test the moisture sensors for four sandy soil types. An auger or soil corer is used to gather soil samples from the field at predetermined depths and locations, guaranteeing an exact representation of the land being studied. Table 1 represents the physical properties of the selected soils used in this study, including their classification, fines content, and grain size distribution parameters.

Key terms:D50 = median grain size; Cu = uniformity coefficient (range of particle sizes); Cc= curvature coefficient (gradation smoothness).

Note: Soil samples are classified based on the Unified Soil Classification System (USCS). Sand 1 is well-graded, while Sand 2, Sand 3, and Sand 4 are poorly graded. Percent fines influence permeability and compaction. Bulk density and specific gravity indicate weight and mineral composition. $Cu(> 4 \text{ for sands})$ measures the range of particle sizes, and Cc (1-3 for well-graded soils) assesses gradation smoothness.

Soil samples are placed on aluminum or ceramic plates, ensuring the containers are clean and dry to prevent contamination. The initial wet weight of each soil sample is determined using a precision balance, accounting for both soil particles and water content. The containers holding the soil samples are placed in an oven maintained at a constant temperature, typically $105 \circ \mathrm{C}$ ($221 \circ \mathrm{F}$). The samples are dried for a specified period, usually 24 to 48 hours, or until their mass remains constant. Once the drying process is complete, the containers are removed from the oven and cooled in a desiccator to prevent moisture absorption from the air. After cooling,

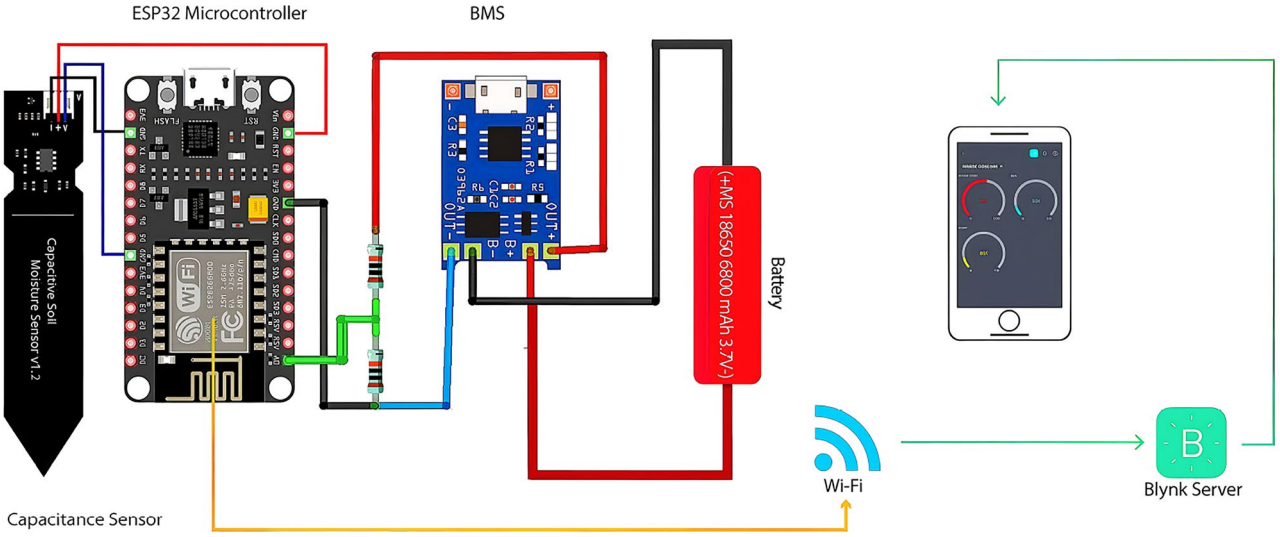


Fig. 2. Circuit diagram of the IoT-based soil moisture monitoring system using an ESP32 microcontroller, capacitive sensor, and Blynk server for wireless data transmission.

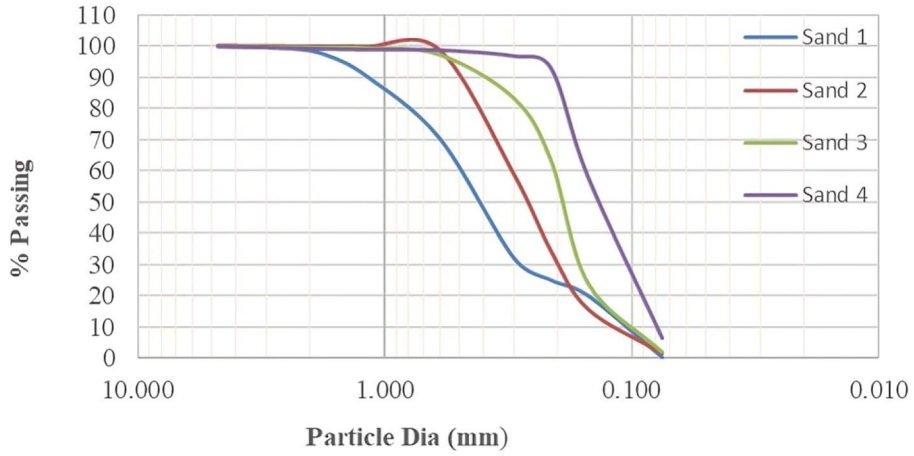


Fig. 3. Variation in particle size and gradation characteristics based on the percentage passing versus particle diameter is depicted by the grain size distribution curves for Sand 1 (Coarse Sand), Sand 2 (Medium-Coarse Sand), Sand 3 (Fine Sand), and Sand 4 (Fine Sand).

Soil type	USCS Class	% Fines	% Sand	Bulk density	Sp. gravity	D_{50}	Cu	Cc
				(g/cc)		(mm)		
Sand 1	SW-Well graded	0.42	99.58	1.88	2.67	0.439	4.65	1.29
Sand 2	SP-Poorly graded	1.51	98.50	1.75	2.64	0.268	2.58	1.05
Sand 3	SP-Poorly graded	2.06	97.95	1.74	2.63	0.190	2.03	1.20
Sand 4	SP-Poorly graded	6.30	93.71	1.70	2.63	0.138	1.91	0.97

Table 1. Physical properties of selected soil.

the dried soil samples are weighed to determine their dry weight. The soil moisture content can be found using the following formula:

$$\text{Moisture Content } (\%) = \frac{\text{Wet Weight} - \text{Dry Weight}}{\text{Dry Weight}} \times 100$$

The difference between the initial wet weight and the dry weight represents the mass of water in the soil sample. Dividing this difference by the dry weight and multiplying the result by 100 yields the soil's moisture content as a percentage. The oven-drying method is employed to calibrate the soil moisture measuring device. Achieving accurate results requires careful handling of the soil samples and precise control of the drying conditions, which can be time-intensive.

The sensor readings of each section of the soil sample is measured using the device(Fig. 4a and b). For each sample, the corresponding raw sensor readings (capacitance measurements) are recorded. A calibration curve is established by plotting the raw sensor readings on the x-axis and the associated moisture content values, derived from reference measurements, on the y-axis.

A calibration curve (logarithmic regression) correlates raw sensor readings (x-axis) with oven-dry moisture values (y-axis). This curve translates capacitance measurements into moisture content. The best-fit calibration curve is then used to derive an equation that relates soil moisture content to raw sensor readings (capacitance measurements). This equation serves as a tool to convert raw sensor data into meaningful soil moisture values during subsequent measurements.

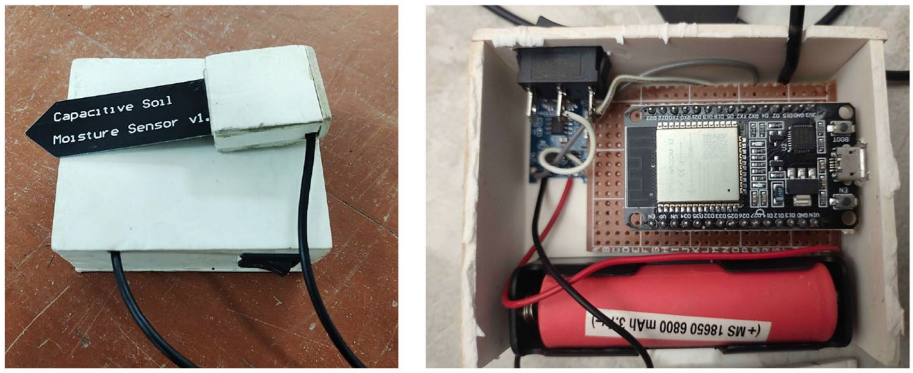
The soil moisture content is determined by measuring the sensor readings (Fig. 5) of the samples and applying the calibration equation. To ensure the accuracy and reliability of the moisture readings, they are compared with reference measurements.

AI model implementation:

Linear regression

Linear regression is a machine learning method used to predict a continuous target variable based on one or more input features. It is categorized as a supervised learning algorithm, where labeled training data, consisting of input-output pairs, is utilized to train the model. The general representation of the linear regression model is given by: y=mx+b

In this equation, b represents the y-intercept, while m denotes the slope or coefficient of the line.



(a) (b)

Fig. 4. (**a**) External view of the assembled soil moisture sensing device using a capacitive sensor. (**b**) Internal layout showing the ESP32 microcontroller, lithium-ion battery, and supporting circuit components for data acquisition and transmission.

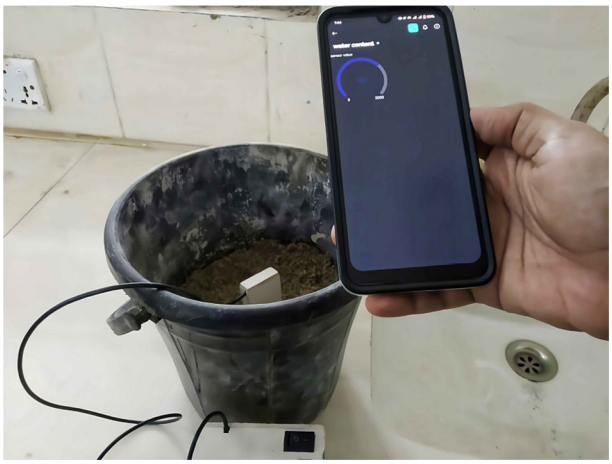


Fig. 5. Monitoring soil moisture using the Blynk app. The mobile app shows the raw sensor value indicating water content after receiving real-time moisture data from the sensor embedded in the soil.

For capacitance-based moisture devices, the purpose of linear regression is to establish a linear relationship between capacitance measurements (input features) and the corresponding soil moisture levels (target variable).

Polynomial regression

Polynomial regression is applied after generating the dataset to model the relationship between variables with a polynomial equation. The general form of the polynomial regression model is:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \ldots + \beta_n X^n + \varepsilon$$

Where:

- YY represents the soil moisture content.
- XX denotes the capacitance measurement.
- $\beta_0, \beta_1, \beta_2 \ldots \beta_n$ are the coefficients of the polynomial terms.
- ϵ is the error term.

The degree of the polynomial (n) is typically determined through testing or cross-validation, depending on the complexity of the relationship between capacitance and soil moisture.

In order to effectively estimate unknown data, machine learning models are used to forecast soil capacitance based on different percentages of water content. In order to examine the connection between soil capacitance and water content, a linear regression model is first implemented. A simplified polynomial regression model is then implemented to get predictive water content values.

Data points that correlate percentages of water content with the relevant values of soil capacitance are used in the training phase. The model forecasts soil capacitance for a variety of water content levels, including 2%,

4%, 6%, 8%, 10%, 12%, 14%, and 16%, when training is finished. If we want to perceive water content, we need to use a regression model.

The model uses the given training data to classify the kind of soil after predicting the soil capacitance. It designates particular classifications, such medium-coarse sand, coarse soil, etc. The model also produces a range of the soil and parameters that help to identify the soil type. This technology supports thorough soil analysis and related applications by enabling precise predictions of soil capacitance and efficient soil type classification.

Random forest classification

To predict soil types based on capacitance and water content parameters, a Random Forest classifier was used on the supplemented dataset. The mode of the classes (for classification) is the final prediction of the Random Forest ensemble learning technique, which builds several decision trees during training.

The following is a summary of a Random Forest classifier's general methodology:

- Multiple decision trees are trained on random subsets of the training data and features.
- Each tree creates an independent prediction.
- The final classification output is determined by majority voting among the trees.

A dataset of different soil types is used to train the algorithm. Initially, we have to select whether we need to know water content or capacitance. If we need capacitance, we need to give the machine a classification model and then it will determine the soil type. This technique enhances generalization and lessens overfitting, particularly for small to medium-sized datasets. In our instance, the Random Forest model correctly and reliably categorized the appropriate soil types, achieving an accuracy of roughly 97.77% on the test set.

Method execution

A flowchart that graphically describes the process of applying machine learning models to forecast soil type and water content is shown in the Fig. 6. It describes how inputs are processed to provide either a soil categorization using a trained classifier or an estimation of the water content using regression, with a focus on capacitance data from an Internet of Things device. This framework's two main functions are soil type recognition and water content prediction.

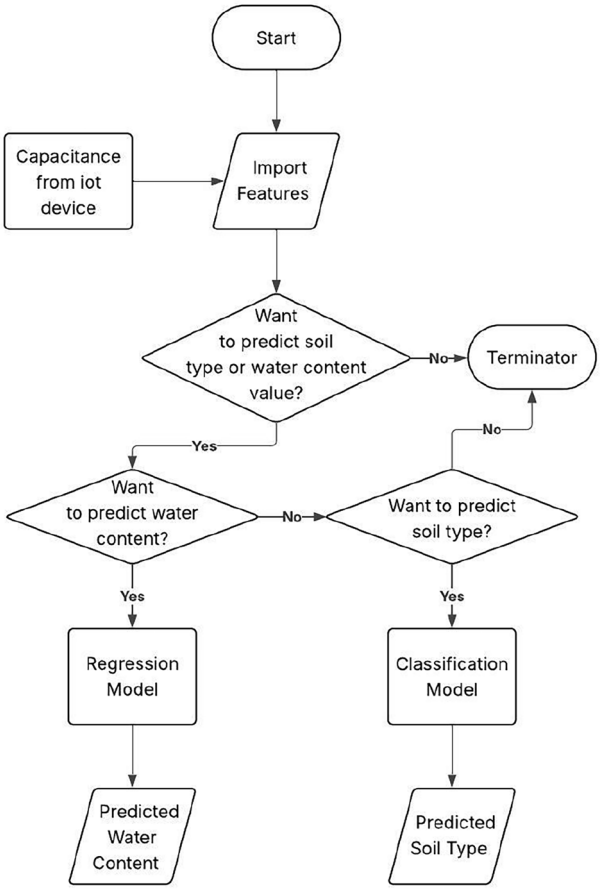


Fig. 6. Flowchart showing the decision process for predicting either soil type or water content based on sensor capacitance readings and input parameters.

Dataset preparation and augmentation for generalization
Three different soil types–fine sand, medium coarse sand, and coarse sand–were experimentally measured for moisture and capacitance in order to produce a specific dataset for this study. There were 120 samples of fine sand, 119 samples of medium coarse sand, and 218 samples of coarse sand in the original dataset. We used the Synthetic Minority Over-sampling Technique (SMOTE) to address the class imbalance problem and enhance the machine learning models' capacity for generalization. Using their feature space, this technique creates new samples for the underrepresented classes in a synthetic manner. Following augmentation, each of the three classes received 218 samples, for a final dataset of 654 samples. In addition to avoiding overfitting during classification[28], this augmentation step made sure the model was equally exposed to each type of soil, which improved the prediction robustness and dependability of the model in practical applications.

Selection of input:
The user starts the process by entering necessary data, such as capacitance readings from the Internet of Things device and other pertinent characteristics. The user's goal–whether it's to identify the type of soil or forecast the amount of water–determines the next stage. The flowchart's course is determined by this choice.

Prediction for water content:
The user must enter the capacitance value and the necessary features if they decide to predict the water content. These inputs are processed by the regression model, which was trained on the relationship between capacitance and moisture, to determine the soil's water content. To improve the prediction accuracy, both linear and polynomial regression models have been applied.

Soil type prediction:
When the goal is to estimate the kind of soil, the user inputs the capacitance values that correspond to known percentages of water content (e.g., 2%, 4%, 6%, up to 16%) together with other required information. The pre-recorded capacitance patterns of various soil types under identical water content conditions are then compared with these input values. Using a machine learning-based classification process, the trained Random Forest Classifier determines the most likely soil type based on the closest match between the input pattern and the stored data.

Training and prediction of the model
The model was trained using datasets that showed the percentage of water content correlated with the corresponding capacitance values for various types of soils. The model determines the most likely soil category for soil type prediction by comparing the input capacitance to the training data. Additionally, we used SMOTE (Synthetic Minority Over-sampling Technique) to improve generalization, reduce the danger of overfitting, and augment the dataset after gathering all the data. We used label encoding to translate categorical soil categories into numerical labels for the classification job of predicting soil type based on capacitance or water content. We divided the dataset into training and testing sets using an 80-20 stratified split to guarantee that the class distribution held steady across subsets. Furthermore, to lessen bias and improve the robustness of our findings, we used Stratified K-Fold Cross-Validation (K=5). Given the short and perhaps unbalanced size of our dataset, our approach ensures that every fold preserves the original class distribution.

Generated insights:
To improve the accuracy of soil categorization, the model additionally produces a variety of related metrics in addition to the main outputs. These metrics offer more information about the properties and behaviors of soil.

This process is visually represented by the flowchart, which guarantees logical progression and clarity. This methodical technique allows for precise soil categorization and enables reliable estimates of soil capacitance and water content, making it appropriate for a variety of applications in geo-technical engineering, agriculture, and environmental monitoring.

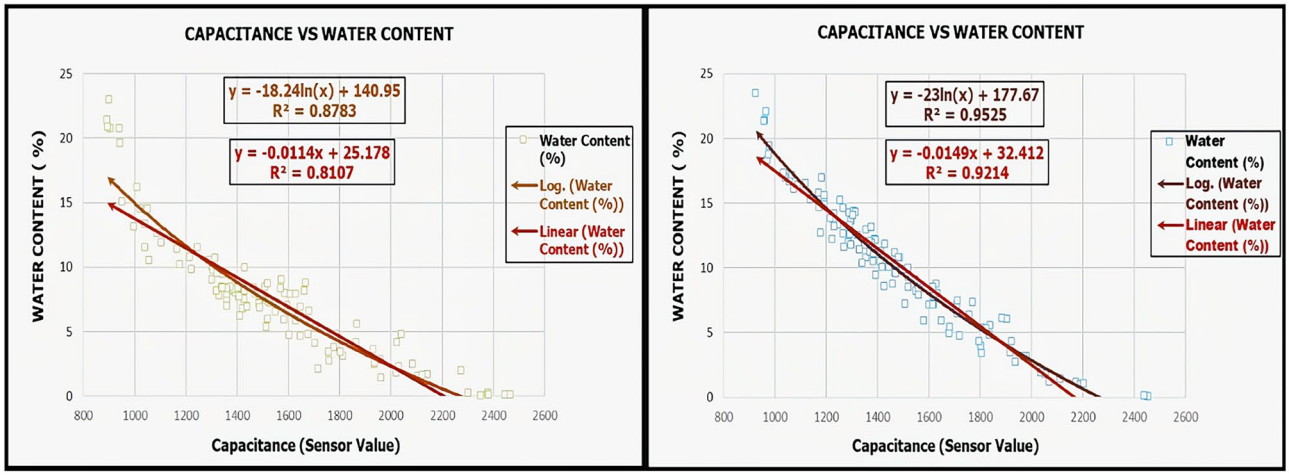
Test results and analysis

The forecasts of soil capacitance and water content are compared in this part using two different approaches: machine learning (ML)-based predictions and manual computation using Excel-based datasets.

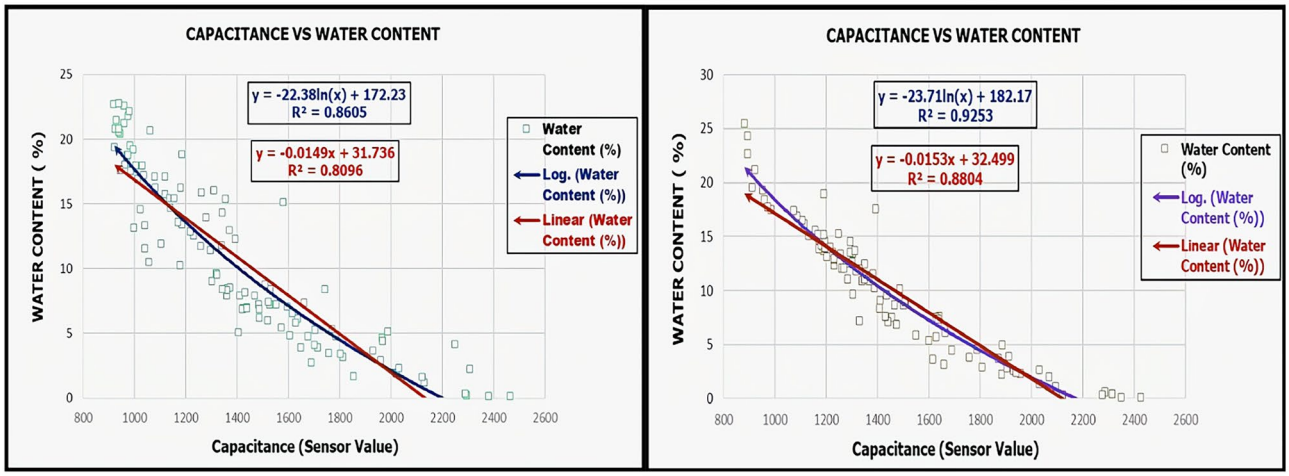
The first section provides a detailed analysis of the Excel results, emphasizing the direct correlations between capacitance values and percentages of water content in various soil types. By using established datasets, this method provides a baseline for evaluating prediction accuracy. The second section evaluates machine learning (ML) models trained on these datasets. Key metrics–including a confusion matrix and classification report–assess their performance in predicting soil water content and type. These metrics offer a thorough evaluation of the model's overall efficacy, recall, precision, and classification accuracy. To confirm the model's effectiveness and spot any inconsistencies, the ML-based predictions are contrasted with Excel-based outcomes, highlighting the benefits and drawbacks of both methods for soil analysis and classification.

A linear regression analysis was conducted using capacitance sensor values (analog signals converted into digital numbers) as the independent variable (x-axis) and the corresponding water content values, expressed as percentages, as the dependent variable (y-axis). Furthermore, a computational algorithm was developed to predict water content based on a given capacitance value. The regression analysis was carried out using Microsoft Excel.

Figure 7 illustrate the relationships between sensor values and moisture content for Sand 1, Sand 2, Sand 3, and Sand 4, respectively. For each sand type, two types of regression equations were developed: a linear equation and a logarithmic equation.



(a) Sand 1 (b) Sand 2



(c) Sand 3 (d) Sand 4

Fig. 7. Capacitance vs. Water Content in Sands. For (**a**) Sand 1 (Coarse Sand), (**b**) Sand 2 (Medium-Coarse Sand), (**c**) Sand 3 (Fine Sand), and (**d**) Sand 4 (Fine Sand), graphs illustrate the correlation between capacitance (sensor value) and water content (%). For every type of sand, linear and logarithmic trendlines are provided together with their equations and R^2 values.

For Sand 1 (Fig. 7a), the logarithmic equation $y = -18.24\ln(x) + 140.95$ ($R^2 = 0.8783$) outperforms the linear equation $y = -0.0114x + 25.178$ ($R^2 = 0.8107$).

Similarly, for Sand 2 (Fig. 7b), the logarithmic equation $y = -23\ln(x) + 177.67$ ($R^2 = 0.9525$) demonstrates a stronger correlation than the linear equation $y = -0.0149x + 32.412$ ($R^2 = 0.9214$).

In Sand 3 (Fig. 7c), the logarithmic equation $y = -22.38\ln(x) + 172.23$ ($R^2 = 0.8605$) again shows a better fit compared to the linear equation $y = -0.0149x + 31.736$ ($R^2 = 0.8096$).

Lastly, for Sand 4 (Fig. 7d), the logarithmic equation $y = -23.71\ln(x) + 182.17$ ($R^2 = 0.9253$) outperforms the linear equation $y = -0.0153x + 32.499$ ($R^2 = 0.8804$).

The negative coefficients in the equations indicate that a decrease in water content corresponds to an increase in sensor values. Meanwhile, the intercept term represents the estimated water content when the sensor capacitance is zero. Although this intercept may lack direct physical significance, it is included for mathematical completeness. The coefficient of determination (R^2) obtained from the regression analysis demonstrates a strong correlation between the sensor values and the corresponding water content for all sands. For all four sand types, the logarithmic equations consistently exhibit higher R^2 values than their linear counterparts. This suggests that logarithmic models are more effective in capturing the relationship between sensor values and moisture content, making them the preferred choice for analysis. Therefore, the logarithmic equation is used as the calibration model for this device.

Testing and validation

Using a spreadsheet program, trendline fitting of the experimental data was utilized to develop the linear and logarithmic equations used for soil moisture prediction. Because the logarithmic equation was better able to

Sample No	Device reading (%)	Lab testing value (%)	Accuracy(%)	Average accuracy(%)
1	2.85	3.24	88.00	94.47
2	5.22	5.20	99.61	
3	6.40	6.16	96.10	
4	7.95	7.41	92.71	
5	8.74	8.40	95.95	

Table 2. Device accuracy assessment via field and lab comparisons using a linear equation.

Sample NO	Device reading (%)	Lab testing value (%)	Accuracy(%)	Average accuracy(%)
1	3.03	3.24	93.52	96.49
2	4.96	5.20	95.38	
3	5.97	6.16	96.92	
4	7.39	7.41	99.73	
5	8.14	8.40	96.90	

Table 3. Device accuracy assessment via field and lab comparisons using a logarithmic equation.

capture nonlinear correlations between device readings and lab testing values, it attained a little greater accuracy (96.49%) than the linear equation (94.47%). The logarithmic model showed more consistency, with accuracies between 93.52% and 99.73%, while the linear model performed well, with accuracies ranging from 88.00% to 99.61%, as indicated in Tables 2 and 3. These findings validate the dependability of the suggested IoT system for practical uses by demonstrating the logarithmic model's stability and applicability for accurate soil moisture prediction.

Tables 2 and 3 show field data collected from the Rajshahi Airport runaway, where device readings were systematically compared with laboratory test results to evaluate accuracy and reliability.

Note: The accuracy of soil moisture prediction using a linear regression model is evaluated by comparing IoT device readings with lab-tested values. The linear model provides a baseline for assessing prediction performance but may not fully capture nonlinear soil moisture variations.

Note: The logarithmic regression model enhances prediction accuracy by capturing nonlinear relationships in soil moisture variations. This approach reduces error compared to linear regression, making it more suitable for real-world applications where soil moisture behavior is complex.

To create reliable predictive models for the classification of soil types and water content, the ML-based approach first trained data on soil datasets. Using test data, soil type and water content predictions were developed following the training phase. A confusion matrix (Fig. 8) and a classification report(Table 4) that emphasized important metrics including accuracy, precision, recall, and F1-score were used to assess the ML models' performance[29] . Based on the input capacitance values, these findings shed light on the model's capacity to accurately forecast soil type.

In order to show the potential benefits of machine learning for accurate and effective soil analysis, this section evaluates the consistency and dependability of the predictions by contrasting the manual and ML-based results.

A comparison of the expected and actual water content using linear (blue) and polynomial (red) regression models is shown in Fig 9. The ideal 1:1 relationship, where model predictions and actual values match exactly, is shown by the diagonal dashed black line. Although there are noticeable variations, particularly at lower and higher water content levels, both models generally follow the optimum trend. In comparison to the linear model, the polynomial model shows a comparatively tighter grouping of points along the 1:1 line, indicating better prediction ability and less bias. This distribution demonstrates how well the polynomial regression model can reflect the dataset's underlying nonlinear relationship.

The matrix indicates that the model showed high prediction skills for both Medium Coarse Sand and Fine Sand, despite the small number of misclassifications for Coarse Sand. In addition to the metrics shown in the classification report, this visualization offers a more thorough assessment of the model's performance and identifies possible areas for development.

Table 4 presents the classification report for soil sample prediction, detailing precision, recall, and F1-score for different sand types, along with overall accuracy and averaged performance metrics.

Note: The classification report evaluates model performance in predicting soil types. Precision measures the proportion of correctly identified positive cases among predicted positives, recall assesses how many actual positive cases were correctly identified, and the F1-score balances precision and recall. The support column indicates the number of instances per class, providing insight into dataset distribution and model reliability.

The model did remarkably well for Fine Sand, attaining flawless precision, recall, and F1-score, according to the classification report. With a high F1-score and flawless recall, Medium Coarse Sand also showed excellent performance. The recall was high for Coarse Sand, though, suggesting that the model only occasionally misclassified a few occurrences that belonged to this group. The model's overall accuracy was 97.73%, and its weighted and macro averages showed balanced performance for all soil types. These findings highlight how well

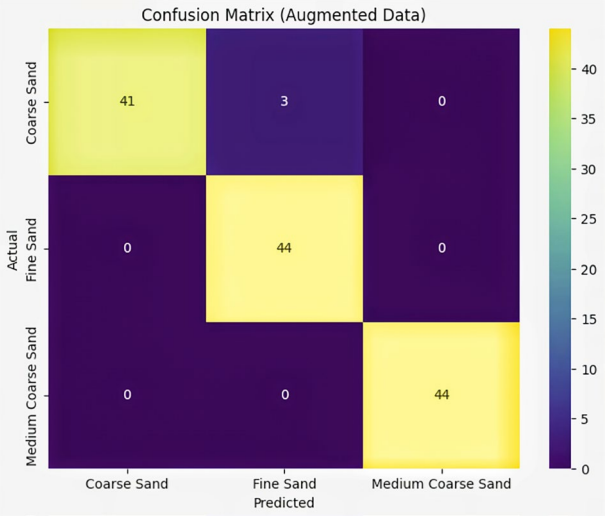


Fig. 8. Confusion matrix showing how well the machine learning model classified three different types of sand–coarse, fine, and medium-coarse–using supplemented data. Rows reflect actual labels, whereas columns represent anticipated labels. While incorrect classifications are displayed in the off-diagonal cells, correct classifications are displayed along the diagonal. With only a few misclassifications (three cases of coarse sand being forecasted as fine sand, for example), the model demonstrated great accuracy. The number of forecasts in each category is shown by the color intensity.

Class	Precision	Recall	F1-score	Support
Coarse sand	1.00	0.93	0.96	44
Fine sand	0.94	1.00	0.97	44
Medium coarse sand	1.00	1.00	1.00	44
Accuracy			0.98	132
Macro-avg	0.98	0.98	0.98	132
Weighted avg.	0.98	0.98	0.98	132

Table 4. Classification report of soil sample prediction using a polynomial regression model in machine learning.

the model classifies soil types; however, there is still an opportunity for improvement in some areas, such as coarse sand.

Conclusion

An AI-driven polynomial regression model and a capacitance-based soil moisture sensor are combined in this study to provide a reliable and affordable method for real-time soil moisture assessment and soil type categorization. Utilizing the sensor's non-destructive, in-situ measurement capabilities, the system showed great suitability for geotechnical engineering applications like foundation design and slope stability studies. The nonlinear relationship between capacitance values and moisture content was well captured by the polynomial regression model, which performed better than the linear model with an accuracy of 96.49% as opposed to 94.47%. To further confirm this nonlinear pattern, a logarithmic regression model was employed.

With an overall accuracy of 98%, the Random Forest classifier demonstrated remarkable accuracy in classifying soil types. All three tested classes–Coarse Sand (Precision: 1.00, Recall: 0.93, F1-score: 0.96), Fine Sand (Precision: 0.94, Recall: 1.00, F1-score: 0.97), and Medium Coarse Sand (Precision: 1.00, Recall: 1.00, F1-score: 1.00)–showed high consistency with the classifier. All evaluation indicators had weighted and macro averages of 0.98, which further supported the model's generalizability and dependability. These findings demonstrate the model's ability for precise and effective soil categorization and water content estimation under real-world field circumstances.

The findings highlight the importance of integrating capacitance-based sensors with machine learning algorithms to increase soil monitoring systems' accuracy. To improve prediction performance and increase real-world applicability in civil infrastructure, agriculture, and environmental monitoring, it is essential to choose regression models that take nonlinear behavior into consideration.

It is still to be confirmed how well the model performs on more complex soils like clay, silt, and high-organic-content soils, even though it showed excellent accuracy across the investigated sand types–fine sand, medium coarse sand, coarse sand, and sand-silt mixture. In order to enable wider generalization, future work

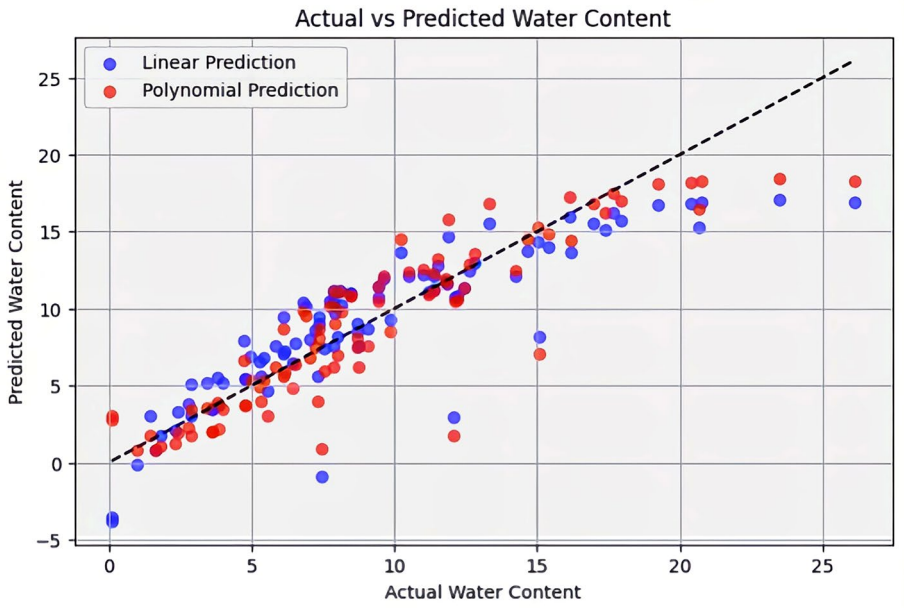


Fig. 9. Using linear (blue) and polynomial (red) regression models, the actual and predicted water content values are compared. The ideal 1:1 connection, where predicted and actual values are exactly equal, is represented by the dashed black line. The models' prediction performance is shown by the distribution of points around the line.

will concentrate on enlarging the dataset to cover various soil types. The effects of additional variables, including mineralogy, clay concentration, water chemistry, temperature, and salinity, will be investigated further. These improvements will further improve system accessibility and unlock more value for geotechnical, agricultural, and environmental applications when incorporated into clear, user-friendly applications.

Data availability

The dataset generated during this study is not publicly available but can be obtained from the corresponding author upon reasonable request.

Received: 13 February 2025; Accepted: 9 September 2025
Published online: 13 October 2025

References

1. Sheng, Z. et al. A portable pull-out soil profile moisture sensor based on high-frequency capacitance. *Sensors* **23**(8), 3806 (2023).
2. Nasirov, T. & Jabborov, K. Microprocessor moisture measuring device of bulk materials based on the capacitance transformer with cylindrical electrodes. In: E3S Web of Conferences, 386,03004 (2023). EDP Sciences
3. Pramanik, M. et al. Evaluation of capacitance-based soil moisture sensors in iot based automatic basin irrigation system (2023)
4. Deng, X. et al. A calibration-free capacitive moisture detection method for multiple soil environments. *Measurement* **173**, 108599 (2021).
5. Teixeira, J. & Santos, R. Exploring the applicability of low-cost capacitive and resistive water content sensors on compacted soils. *Geotech. Geol. Eng.* **39**(4), 2969–2983 (2021).
6. Kanso, T., Gromaire, M.-C., Ramier, D., Dubois, P. & Chebbo, G. An investigation of the accuracy of ec5 and 5te capacitance sensors for soil moisture monitoring in urban soils-laboratory and field calibration. *Sensors* **20**(22), 6510 (2020).
7. Al, J. & Smith, J. Low-cost soil moisture sensors' assessment for their accuracy after calibration through the gravimetric method. *Journal of Soil Research.* **56**(1), 353–369 (2024).
8. Wu, S., Zhang, B., Tian, Y., Zhou, S. & Ma, H. A grain moisture model based on capacitive sensor. In: Journal of Physics: Conference Series, 1074,012120 (2018). IOP Publishing
9. Nathk, D. & Ramanathan, P. A simplified and portable capacitance sensor to measure moisture content of paddy. *Gazi Univ. J. Sci.* **31**(3), 821–829 (2018).
10. Domínguez-Niño, J. M., Oliver-Manera, J., Arbat, G., Girona, J. & Casadesús, J. Analysis of the variability in soil moisture measurements by capacitance sensors in a drip-irrigated orchard. *Sensors* **20**(18), 5100 (2020).
11. Silva, C. R. et al. Calibração de uma sonda de capacitância em um argissolo. *Sci. Agricola* **64**(6), 636–640 (2007).
12. Mittelbach, H., Casini, F., Lehner, I., Teuling, A.J. & Seneviratne, S.I. (2011) Soil moisture monitoring for climate research: Evaluation of a low-cost sensor in the framework of the swiss soil moisture experiment (swisssmex) campaign. Journal of Geophysical Research: Atmospheres **116**(D5)
13. Deng, X. et al. A method of electrical conductivity compensation in a low-cost soil moisture sensing measurement based on capacitance. *Measurement* **150**, 107052 (2020).
14. Wang, W., Zhao, K., Zhang, P., Bao, J. & Xue, S. Application of three self-developed ect sensors for monitoring the moisture content in sand and mortar. *Construct. Build. Mater.* **267**, 121008 (2021).
15. Surya, S. G. et al. An in-field integrated capacitive sensor for rapid detection and quantification of soil moisture. *Sensors Actuators B: Chem.* **321**, 128542 (2020).

16. Archer, N., Rawlins, B., Machant, B., Mackay, J. & Meldrum, P. Approaches to calibrate in-situ capacitance soil moisture sensors and some of their implications. *SOIL Discussions* **2016**, 1–29 (2016).
17. Zabolotnyi, O., Zabolotnyi, V. & Koshevoy, N. Capacitive water-cut meter with robust near-linear transfer function. *Computation* **10**(7), 115 (2022).
18. Joaquin, A. C., Ramos, M. & Martinez, R. C. Development and performance evaluation of a grain probe moisture meter for paddy. *Int. J. Environ. Sci. Dev.* **10**(10), 350–355 (2019).
19. Suranata, I. W. A. et al. Arsitektur moisture meter dengan capacitive sensing dan serverless iot untuk hidroponik fertigasi. *J. RESTI (Rekayasa Sistem dan Teknologi Informasi)* **5**(2), 292–300 (2021).
20. Smith, W.B., Jung, H.-S. & Association, W.D.K., et al. Effect of high temperature drying on moisture content determination with electronic meters (1993)
21. Kostadinović, D. M., Dimitrijević-Jovanović, D., Vučićević, B. S., Jovanović, M. P. & Konjikušić, S. M. Smart capacitive moisture sensor calibration in mineral wool and green roof soil substrate. *Therm. Sci.* **25**, 1827–1836 (2021).
22. Patil, K.R., Eickhoff, S.B. & Langner, R. Predictive data calibration for linear correlation significance testing. arXiv preprint arXiv:2208.07081 (2022)
23. Adla, S. et al. Laboratory calibration and performance evaluation of low-cost capacitive and very low-cost resistive soil moisture sensors. *Sensors* **20**(2), 363 (2020).
24. Bogena, H. R., Huisman, J. A., Oberdörster, C. & Vereecken, H. Evaluation of a low-cost soil water content sensor for wireless network applications. *J. Hydrol.* **344**(1–2), 32–42 (2007).
25. Engineers, L.M. Interfacing Capacitive Soil Moisture Sensor with Arduino. https://tinyurl.com/5bb2bydd. Accessed: 2024-12-30 (2024)
26. Espressif. ESP32 Technical Reference Manual. https://www.espressif.com/sites/default/files/documentation/esp32_technical_reference_manual_en.pdf. Accessed: 2024-12-30 (2024)
27. Haldar, S., Gol, S., Mondal, A. & Banerjee, R. Iot-enabled advanced monitoring system for tubular batteries: Enhancing efficiency and reliability. *E-Prime-Adv. Elect. Eng., Elect. Energy* **9**, 100709 (2024).
28. Khatun, M.S. & Khan, I. Bangla counterfeit news identification: Using the power of bert. In: 2024 IEEE International Conference on Power, Electrical, Electronics and Industrial Applications (PEEIACON), pp. 518–522 (2024). https://doi.org/10.1109/PEEIACON63629.2024.10800650
29. Khan, I. et al. Leveraging the robust capability of modified multi-layer gated recurrent units for fake news detection. In: 2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), pp. 986–991 (2024). https://doi.org/10.1109/ICEEICT62016.2024.10534407

Acknowledgements
The authors acknowledge the support of the Department of Civil Engineering, Rajshahi University of Engineering & Technology, Bangladesh, for providing laboratory facilities.

Author contributions
Md. Mahmud Sazzad (1st author): Conceptualization, supervision, review. Tanvir Ahmed (2nd author): Data curation, IoT device development, methodology, original draft. Golam Kibria (3rd author): Visualization, validation. Ishmam Khan (4th author): Machine learning model development, editing. All authors reviewed the manuscript.

Funding
The authors did not receive any financial support for the research, authorship, or publication of this article.

Declarations

Competing interests
The authors declare no conflicts of interest relevant to this study.

Ethical approval
This study does not involve human participants or animals, and ethical approval is not applicable.

Consent to participate
Not applicable, as no human participants were involved.

Consent for publication
All authors consent to the publication of this manuscript.

Additional information
Correspondence and requests for materials should be addressed to T.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

IoT-driven smart agricultural technology for real-time soil and crop optimization / Shahab, H., Naeem, M., Iqbal, M., Aqeel, M., & Ullah, S. S.

Smart Agricultural Technology
Volume 10 (2025) 100847 Pages 1-9
https://doi.org/10.1016/j.atech.2025.100847
(Database: ScienceDirect)

Smart Agricultural Technology 10 (2025) 100847

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

IoT-driven smart agricultural technology for real-time soil and crop optimization

Hammad Shahab [a], Muhammad Naeem [a], Muhammad Iqbal [a,b,*], Muhammad Aqeel [a], Syed Sajid Ullah [c,*]

[a] Institute of Computer and Software Engineering, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan, 64200, Pakistan
[b] School of Interdisciplinary Engineering and Sciences (SINES), National University of Sciences and Technology (NUST), Islamabad 44000, Pakistan
[c] Department of Information and Communication Technology, University of Agder (UiA), N-4898 Grimstad, Norway

ARTICLE INFO

Keywords:
Soil monitoring
Fertilizer management
IoT
Smart farming
Precision agriculture
Sustainability
Crop management
Rice

ABSTRACT

To address world food and agricultural sustainability, the UN's Sustainable Development Goals (SDGs) place strong emphasis on a transformative agricultural sector. By 2050, there will be about 10 billion people on the planet, making it imperative to increase about 60 percent in food production efficiency and security in order to meet the growing population's demands. In this critical situation, IoT-based smart farming offers a promising solution that can achieve 12 SDGs out of 17. This study unveils a groundbreaking Internet of Things (IoT) system designed to revolutionize agricultural practices by providing real-time monitoring of soil parameters, including temperature, moisture, salinity, EC, pH, nitrogen, potassium, and phosphorus levels. Field testing conducted with rice crops in Rahim Yar Khan, Pakistan, demonstrated the system's remarkable ability to accurately measure all eight critical parameters. The soil analysis showed a temperature between 30.5 °C and 33.2 °C, a moisture content between 60.6 % and 94.1 % and a pH value between 7.13 and 8.33. Nutrient content varied with nitrogen (71–103 mg/kg), phosphorus (15–19 mg/kg), and potassium (101–141 mg/kg). Leveraging this data, an AI-driven mobile application was used to deliver best recommendations for optimizing crop management, particularly in fertilization, irrigation and disease diagnoses practices. By integrating advanced IoT technologies, cloud computing, predictive algorithms, and a smart soil sensor, this system revolutionizes agriculture by enabling real-time monitoring of critical factors influencing rice crops metabolism. It empowers farmers with data-driven insights to enhance productivity, optimize resource use, and improve sustainability. This scalable solution addresses modern agricultural challenges by reducing environmental impact and strengthening crop resilience. Focused on unlocking the potential of data-driven decision-making, this research paves the way for sustainable, smart farming practices to meet the growing demands of a rapidly increasing global population.

1. Introduction

With a growing global population comes the rising issue of food security. One of the solutions, which will definitely carry the day, is optimizing agricultural productivity by smarter resource management. However, four significant challenges hold back the efficient and sustainable development of agriculture. The first one is that labor-intensive and time-consuming manual labor severely curbs productivity. The second is poor decision-making; decision-making frequently quite relies on guesswork to impede the attainment of optimal crop management. Specific concerns are also complicated by issues such as climate change, whose impacts include irregular weather patterns and degraded soil quality. Finally, inefficient use of resources, including overuse of water and fertilizers, not only raises costs but also adversely affects the environment [1,2].

Agriculture-based industries act as backbones of economies in developing countries, such as Pakistan [3]. Healthy soils, crop growth analysis and optimization of crop metabolism determine sustainable crop production and food security in developing countries. Productivity increases due to improved farming practices for soil management and crop health monitoring, while conserving environmental resources for future generations [4]. Agriculture is a cornerstone of Pakistan's economy and a vital occupation for human survival. Its success hinges on healthy soil and farmers' knowledge of crop and land management.

* Corresponding authors.
E-mail addresses: muhammad.iqbal@sines.nust.edu.pk (M. Iqbal), syed.s.ullah@uia.no (S.S. Ullah).

https://doi.org/10.1016/j.atech.2025.100847
Received 5 December 2024; Received in revised form 16 February 2025; Accepted 19 February 2025
Available online 21 February 2025

Healthy soil also supports sustainable crop production by acting as a reservoir for nutrients and moisture, essential for plant growth and microbial activity. Informed farmers can make better decisions regarding planting, irrigation, and fertilization, ensuring optimal productivity [5]. These factors collectively sustain agricultural output and contribute to the country's economic stability [6,7]. It also makes a significant contribution to our daily lives [8]. For farmers, understanding soil nutrients is critical to success [9–11].

The literature highlights various IoT-driven farming systems for soil monitoring and crop management. In study of Zhang et al., 2017 IoT-based systems with sensors offer high accuracy for parameters like soil temperature and nutrients but lack mobile application integration [12]. Abishek et al. (2021) analyzed smart irrigation systems and IoT kits that mainly monitor soil moisture and weather conditions. Their research found moderate accuracy and cost-effectiveness, but also found that there is no cloud dashboard integration [13]. In the study by Mesas-Carrascosa et al. (2015), IoT-based irrigation systems were examined for their efficiency in monitoring soil and climate parameters. The study found that while mobile app support was available, cloud dashboard functionality was not included [14]. In the research of Ayaz et al. (2019) and Pitakaso et al. (2023), Decision Support Systems and AMIS optimize farming decisions but miss mobile support [15,16]. The extensive research and development and literature on IoT-based nutrient management also highlights innovative approaches namely precision farming—spanning automated devices [17], drones [18], satellite imagery [19], big data [20], and machine learning [21] to improve agricultural efficiency and sustainability. A research study by Hartono et al. (2024) presents a portable IoT system for real-time monitoring of soil nutrients, focusing on nitrogen, phosphorus, and potassium, which helps farmers make informed fertilizer decisions [22]. In the study of Islam et al. (2023), an advanced machine learning IoT-based device has been developed to monitor soil conditions and provide custom-made crop recommendations, optimizing resource use and increasing productivity [23]. Furthermore, the integration of IoT and machine learning into nutrient management models enables effective yield prediction and nutrient loss analysis, addressing the challenges of environmental factors [24]. While numerous studies highlight the transformative potential of IoT in nutrient management, the lack of comprehensive insights into crop health and key parameters of soil leaves a critical research gap. Soil key parameters must be monitored: temperature, moisture, salinity, pH, and nitrogen (N) to allow for the best environment for plant growth, nutrient uptake, and soil condition. All these make up the foundation of farming. Nevertheless, by also measuring electrical conductivity (EC), potassium (K), and phosphorus (P), farmers can improve their agricultural productivity management to a large extent. The inclusion of these extra parameters makes it possible to obtain a better understanding of the state of the soil, and therefore, healthier crops, higher yields, and, again, higher yields.

To address these challenges, it is essential to monitor not only soil nutrition but also crop growth and metabolism. Knowledge as to how plants metabolize nutrients as well as the impact of soil nutrition on their growth permits a farmer to make an informed judgment regarding fertilizer, irrigation practices, and overall crop management. Monitoring key soil parameters that include nitrogen, phosphorus, and potassium counterpart environmental parameters such as temperature, moisture, and pH levels increase the healthfulness of crop metabolism. Proper nutrient management skills ensure that crops attain all the elements necessary for photosynthesis, energy production, and protein synthesis, as crop health and yield are directly impacted. The application of latest technologies in the context of soil science in respect of sensors technology, IoT-based systems, machine learning, data analytics and cloud computing enable farmers to gain real-time insights into both soil conditions as well as plant metabolism through optimize resource use efficiency.

While IoT offers significant benefits in smart farming, challenges like real-time soil nutrient monitoring and accurate fertilizer and irrigation planning remain critical. The overall goal of this research is to show how the use of a smart system that measures eight key parameters of soil and the use of an AI-driven mobile app can change the world of eco-agriculture and increase food security. This study introduces a cutting-edge IoT system for real-time monitoring of soil parameters, including temperature, moisture, and nutrients, enhancing rice crop management. Field tests in Rahim Yar Khan, Pakistan, showcased its accuracy in measuring critical factors affecting crop metabolism. Supported by an AI-driven mobile app, it empowers farmers with data-driven insights to optimize productivity, sustainability, and resource efficiency, addressing modern agricultural challenges. The main objective of the study is to evaluate the effectiveness and accuracy of the proposed IoT-based real-time soil monitoring and data collection system for agricultural applications.

1.1. Key concepts in smart farming

Understanding the basic ideas behind soil, nutrients, and crop health is crucial to comprehending smart agriculture and its uses. An overview of these important soil and crop concepts is provided in this section.

1.1.1. Important parameters for rice crop

Rice cultivation requires specific environmental conditions to ensure optimal growth and yield [25]. Among these, soil temperature, moisture, and pH are critical factors influencing the various stages of rice growth, from germination to ripening. Maintaining appropriate levels of these parameters promotes nutrient availability, root development, and plant health, ensuring a sustainable and productive crop cycle. In Table 1 important parameters of rice crop soil are presented with the summary:

1.1.2. Soil nutrients of rice crop for sustainable production

Soil nutrients are vital for sustainable crop production, influencing plant growth, health, and yield [29]. Key nutrients like nitrogen, phosphorus, and potassium play essential roles in metabolism, structure, and stress resistance. Monitoring and managing these nutrients effectively can optimize crop productivity while ensuring environmental sustainability [30,31]. The following Fig. 1 illustrates various diseases due to nutrient deficiencies in plants. Nutrient deficiencies specially of nitrogen (N), phosphorus (P), and potassium (K) in rice crops can lead to various issues, impacting overall growth and yield. Nitrogen deficiency causes yellowing of older leaves (chlorosis), reduced tillering, and stunted growth, leading to poor grain production. Phosphorus deficiency results in dark green or purplish leaves, delayed maturity, and underdeveloped roots and grains. Potassium deficiency manifests as yellowing or browning of leaf edges, weak stems, and poor grain filling, leaving the crop more vulnerable to diseases and lodging. Addressing these deficiencies is essential to ensure healthy rice crop development and maximize productivity.

1.1.3. The importance of observing plant metabolism in rice cultivation

The metabolism of rice crops is shaped by factors like nutrient levels, water availability, and environmental conditions [32,33]. By using a sensor that tracks key elements such as temperature, moisture, salinity, pH, electrical conductivity, nitrogen, potassium, and phosphorus, farmers can gain real-time insights into the soil condition and crop health. This data helps ensure the rice plants are getting the right nutrients, moisture, and conditions needed for optimal growth. When integrated into a smart farming system with an online cloud and mobile application, the sensor allows farmers to make more informed decisions about when and how to irrigate, fertilize, and manage pests, ultimately boosting rice productivity and promoting more sustainable farming practices. The following Table 2 summarize the crucial constraints for growth of rice crop.

Table 1
Essential parameters of rice crop soil.

Crop Growth Phase	Growth Stage	Soil Temperature (°C)	Soil Moisture Level	Soil pH	Description	References
Establishment	Seed Germination	20–35	70–80 % field capacity	5.5–6.5	Optimal for rapid and uniform germination; excess moisture or improper pH reduces germination rates.	[25]
Vegetative Phase	Tillering	25–30	80–90 % field capacity	5.5–6.5	Ideal for root development and tiller formation; consistent moisture supports nutrient uptake.	[25,26]
Reproductive Phase	Panicle Initiation	22–30	70–80 % field capacity	5.5–6.5	Crucial for panicle development; temperature extremes may impair flowering and grain setting.	[27]
	Flowering	22–30	70–80 % field capacity	5.5–6.5	Stable moisture and temperature required for pollination and fertilization; sensitive to water stress.	[27]
Ripening Phase	Grain Filling	20–27	60–70 % field capacity	5.5–6.5	Cooler temperatures aid uniform grain filling; too much water may delay ripening.	[28]
	Maturation	20–27	50–60 % field capacity	5.5–6.5	Supports grain drying and harvest readiness; low soil moisture is preferred.	[28]

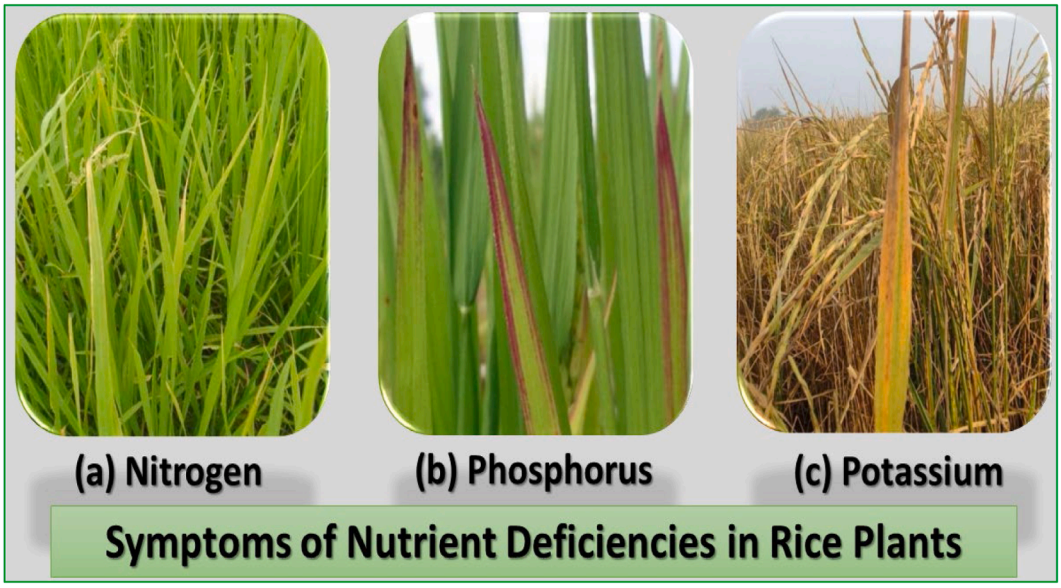


Fig. 1. Nutrient deficiencies in rice crop.

Table 2
Crucial constraints for growth of rice crop.

Parameter	The key feature in management healthy soil	Reference
Phosphorus	Enhances root development, energy transfer, and early growth in rice	[34]
pH	Optimal pH ensures nutrient availability and root health in rice fields	[35]
Moisture	Essential for maintaining flooded conditions and nutrient absorption	[36]
Temperature	The specific level is crucial for germination, vegetative growth, and grain filling in rice crops	[37–40]
Nitrogen	Promotes tillering, leaf growth, and chlorophyll production in rice	[30, 41–44]
Potassium	Improves disease resistance, water regulation, and grain quality in rice	[30]

1.1.4. Internet of Things (IoT) and fertilizer requirement for rice crop growth

Internet of Things (IoT), a modern smart technology, describes a network of smart devices that are linked to the Internet and have the ability to collect, share, and analyse data. These devices often have sensors, software, and other technologies that enable them to communicate with main control systems and other devices, facilitating the automation and remote monitoring of many processes via Internet clouds [45–47]. IoT devices used in agriculture can include soil moisture sensors, weather stations, and NPK sensors, which measure the soil's nutrient content [48,49]. They improve farm management by providing farmers with the ability to obtain real-time information about their farms. For example, as shown in Fig. 2 below, IoT devices can be used to schedule irrigation, manage livestock, and keep an eye on fertilizer and crop health. Farmers can increase productivity, reduce waste, and enhance sustainability by utilizing this cutting-edge IoT technology [50, 51].

The challenges of climate change, food insecurity, and sustainability call for innovative agricultural solutions. For rice crop growth, nutrients like nitrogen are vital for robust leafy development, while phosphorus and potassium support root health and grain quality. Overreliance on generalized fertilizer practices without real-time monitoring can harm soil health, reduce productivity, and cause environmental damage. IoT-based smart farming offers a transformative solution, enabling continuous monitoring of soil parameters through cloud technologies. The IoT approach empowers farmers to make data-driven decisions, optimize fertilizer use, and enhance crop health while minimizing environmental impact [52,53].

2. Methodology & materials

This section outlines the methodology and materials for the proposed IoT-based system aimed at enhancing rice crop management, particularly optimizing fertilizer use and irrigation practices.

Fig. 2. IoT applications in agriculture.

2.1. Proposed methodology and IoT-based system

The proposed system comprises a smart soil sensor, data acquisition system, wireless communication network, and cloud-based platform. The IoT-based system that is proposed in this article includes a smart sensor, data acquisition system, wireless communication network, and a cloud-based platform. This smart soil sensor accurately measures temperature, moisture, conductivity, pH, salinity, and NPK levels. It is built with stainless steel probes, sealed with epoxy resin, and runs on a 5–24 V power supply. The sensor has stable performance and high sensitivity, and can directly and stably reflect the actual moisture content of different soils and the nutritional status of the soil in time, providing a data basis for scientific planning about planting. Data is sent via an Internet connectivity module for efficient real-time monitoring to the cloud platform for analysis and visualization in agriculture. The sensor, equipped with advanced probe technology, simultaneously measures soil temperature, moisture, salinity, EC, pH, nitrogen, phosphorus, and potassium levels. The collected data is transmitted wirelessly. Fig. 3 shows the methodology of the overall system.

Table 3 describes the details of equipment and sensors used in the development of this IoT-based fertilizer and crop management system.

2.2. Experimental setup and soil sampling

A field study was conducted at different locations of rice fields in Rahim Yar Khan, Pakistan to assess soil health and identify potential

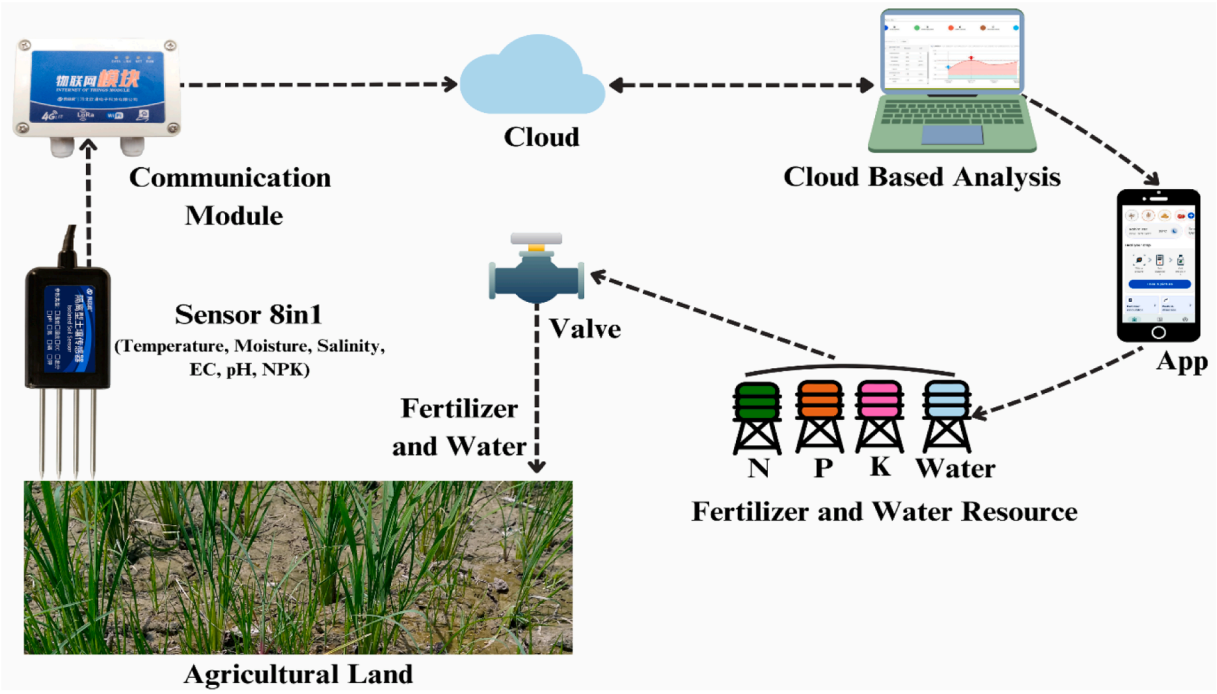


Fig. 3. Proposed framework.

Table 3
Smart soil nutrients monitoring sensor details.

Parameter	Range/Description
Communication Module	Hebei Ousu Electronic Technology Co., Ltd.
Cloud Integration	Jiousu IoT Cloud Server, real-time monitoring, analysis, visualization, CSV data export
IoT Module	ESP8266 kit-based WiFi Enabled Circuit
Sensor Type	8in1 Soil Nutrients Monitoring Sensor
Applications	Agriculture, precision farming, soil monitoring, smart farming, environmental monitoring
Supply Voltage	DC 5–24V
Power Consumption	15 mA @ 12 V DC
Communication Method	RS485
Baud Rate	9600bps
Temperature Range	−40 °C to 85 °C
Temperature Accuracy	±1.0 °C
Soil Moisture Range	0 % to 100 % (Relative Humidity)
Soil Moisture Accuracy	±5 %
Electrical Conductivity (EC) Range	0 to 10,000 µS/cm
EC Accuracy	±5 %
Nutrient Range (NPK)	Nitrogen (N): 0 to 200 mg/kg Potassium (K): 0 to 200 mg/kg Phosphorus (P): 0 to 200 mg/kg
NPK Accuracy	±10 %
pH Range	0 to 14
pH Precision	±10 %
Salt (Salinity) Range	0 to 10,000 µS/cm
Salt Accuracy	±5 %
Sealing Material	Black epoxy resin
Material	Four-pin design with three solid stainless-steel probes, one hollow stainless-steel probe, and zinc alloy. Total length: 138 ± 1 mm; Probe length: 65 ± 1 mm; Diameter: 3 ± 0.2mm

growth constraints using the proposed system. The proposed system was tested through external (laboratory/workshop) and on-site evaluations. Prior to deployment, a site visit was conducted to assess conditions on the farm. The rice crop were in the leaf development stage (end of July to beginning of August). The system was deployed in a 1-acre rice field (28.391257°N, 70.336044°E) under normal weather conditions. A mini solar panel of 20 W with a power bank ensured self-sufficiency, and the internet connection was via a mobile WiFi hotspot. 14 soil samples (0–6 inches deep) were field tested in real time at various locations and selected samples were cross-checked at a local soil testing laboratory for accuracy assessment. The primary objective of the study was to evaluate the effectiveness and accuracy of the proposed IoT based real-time environmental monitoring and data collection system for agricultural applications. Fig. 4 shows the laboratory and field experiment activities, including soil sampling and equipment setup.

3. Experimental results: decision-making with real-time monitoring and AI-driven mobile application

Data on temperature, pH, soil moisture, salinity (µS/cm), EC (µS/cm), nitrogen, phosphorus and potassium levels were recorded for each sample (Table 4). Temperatures ranged from 30.5 °C to 33 °C, reflecting typical seed growing season conditions. Soil moisture content varied between samples, indicating potential water management problems. Salinity was generally moderate, but effects on crop growth require further investigation. Nutritional analysis showed adequate nitrogen and potassium content, while phosphorus content showed some variation. Particularly, soil pH was consistently high in all samples, indicating alkaline conditions and suitable for rice crop growth and productivity.

The performance of the proposed system is further analyzed through a web-based cloud platform that stores data from rice fields, as shown in Fig. 5. Smart soil sensors collect data from soil samples and send them to the online cloud via an Internet connection, where farmers can analyze the data and make smart and effective decisions about irrigation and fertilization activities.

3.1. Algorithm for fertilizer and irrigation recommendation of rice crop

- Function IrrigationAndFertilizerRecommendation(soilMoistureData, nitrogenData, currentDate)
- // Initialize variables
- consecutiveDryDays=0
- consecutiveLowNitrogenDays=0
- irrigationNeeded = False
- fertilizerNeeded = False

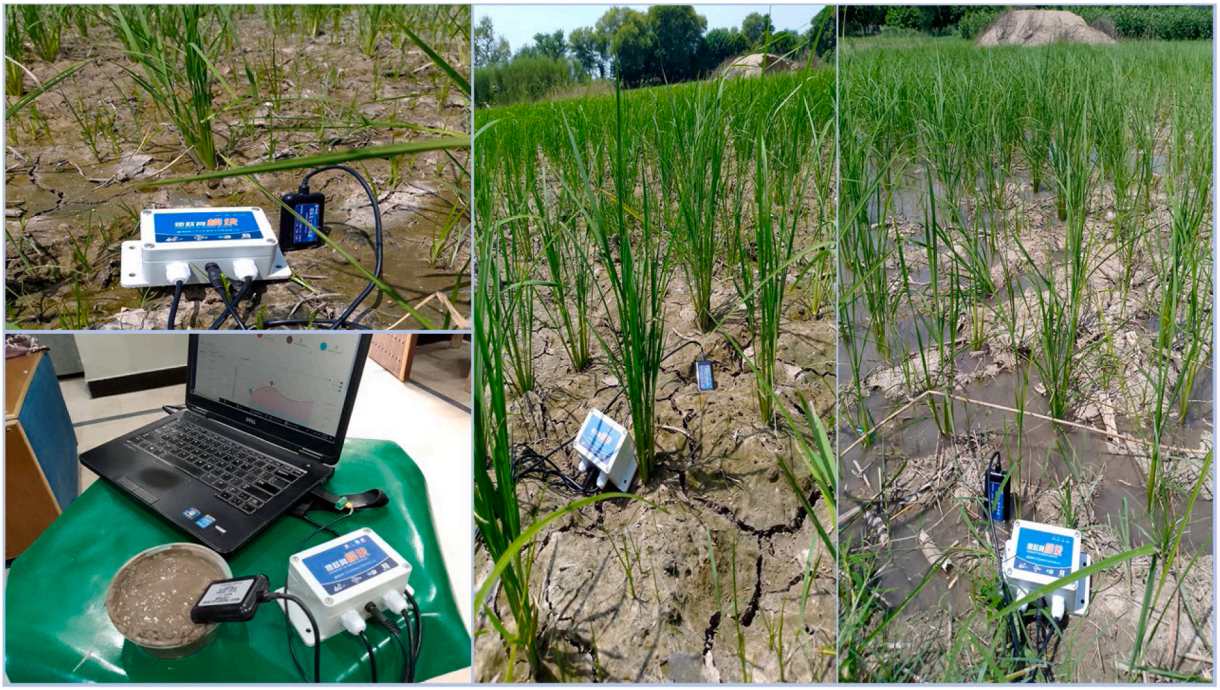


Fig. 4. Field testing in rice crop.

Table 4
Summary of experimental results.

Rice Soil	Temp (°C)	EC (µS/cm)	Soil Moisture (%)	Salinity (µS/cm)	Soil pH	Nitrogen (mg/kg)	Phosphorus (mg/kg)	Potassium (mg/kg)
Soil Sample 1	31.1	1884	60.6	1712	7.87	71	18	101
Soil Sample 2	33.2	1852	65.2	1701	7.13	83	16	119
Soil Sample 3	30.5	1864	62.3	1718	8.17	103	15	116
Soil Sample 4	30.8	1672	67.2	1217	7.18	94	16	109
Soil Sample 5	31.8	1271	81.2	1244	8.11	78	15	117
Soil Sample 6	32.1	1184	90.6	1212	7.81	98	18	141
Soil Sample 7	32.2	1191	94.1	1231	8.05	101	18	126
Soil Sample 8	32.9	1246	80.0	1388	7.19	77	19	124
Soil Sample 9	32.8	1252	87.2	1301	8.33	81	16	139
Soil Sample 10	31.5	1211	81.8	1362	8.16	84	15	121
Soil Sample 11	31.7	1316	80.0	1488	7.79	77	19	125
Soil Sample 12	32.9	1276	82.0	1388	7.52	75	18	123
Soil Sample 13	31.1	1724	83.2	1733	7.91	88	16	131
Soil Sample 14	32	1733	79.4	1714	8.21	83	19	136

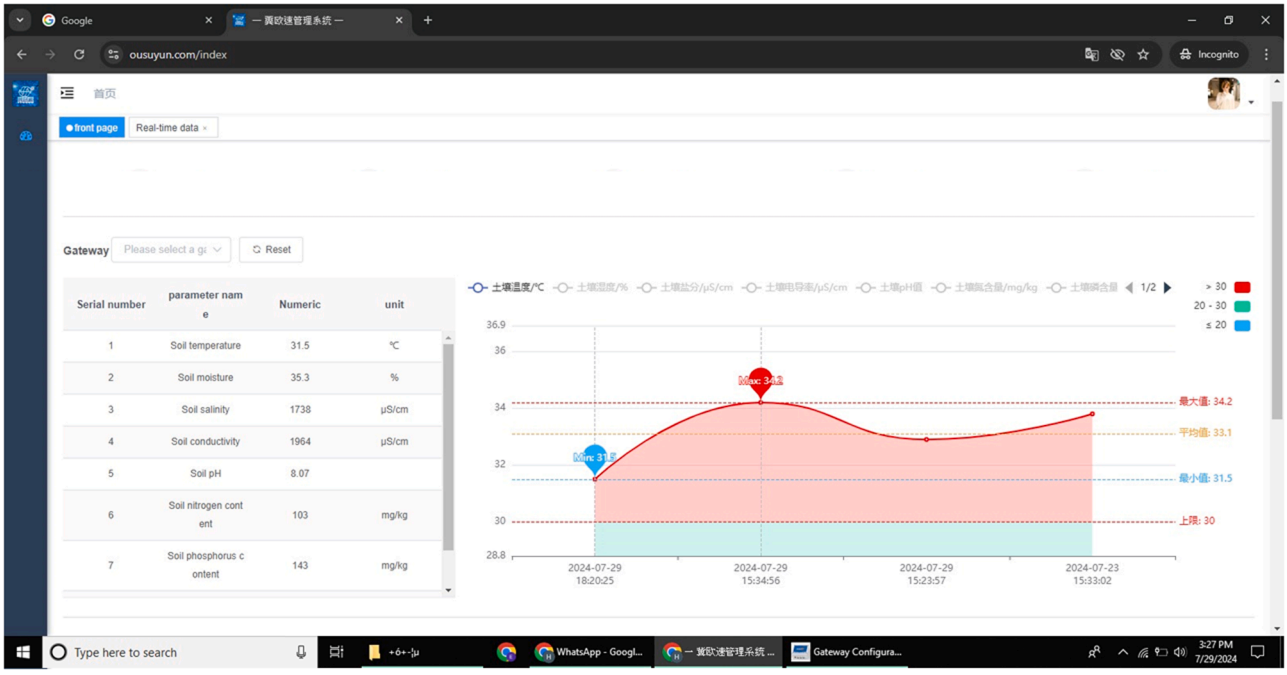


Fig. 5. Cloud platform for real-time monitoring.

- fieldCapacityThreshold = 70 // Optimal soil moisture for rice
- nitrogenDropThreshold = 100 // Significant drop in nitrogen levels

// Check soil moisture and nitrogen levels for the past 7 days

- for *i* = currentDate - 6 to currentDate
- if soilMoistureData[i] < fieldCapacityThreshold
- consecutiveDryDays=consecutiveDryDays /1
- else
- consecutiveDryDays=0
- if nitrogenData[i] < nitrogenData[i-1] - nitrogenDropThreshold
- consecutiveLowNitrogenDays=consecutiveLowNitrogenDays+ 1
- else
- consecutiveLowNitrogenDays=0

// Check if irrigation is needed

- if consecutiveDryDays≥7 // Trigger irrigation after 7 consecutive dry days
- irrigationNeeded = True
- CalculateIrrigationAmount(soilType, cropStage, weatherConditions)
- ApplyIrrigation()

// Check if fertilizer is needed

- if consecutiveLowNitrogenDays≥ 7 // Trigger fertilizer after 7 consecutive low nitrogen days
- fertilizerNeeded = True
- RecommendNitrogenFertilizer(cropStage, soilConditions)
- ApplyFertilizer()
- End Function

Explanation of Algorithm: The function observes soil moisture and nitrogen contents for 7 days. On consecutively dry days exceeding 7 or nitrogen contents dropping too low, it invokes irrigation or fertilizer applications. This IoT-based data-driven approach confirms rice crops obtain timely care for optimal growth and sustainable production. Mobile applications and software can be developed considering this idea and algorithm.

3.2. Plantix: the AI-Driven plant doctor app

Plantix is like having a crop expert in your pocket. This smart app uses AI to diagnose plant health issues, recommend fertilizers, and offer pest control tips [54,55]. Analyzing soil and environmental data, gives farmers easy-to-understand, tailored advice to help improve their crops and boost sustainability. To calculate the required amount of fertilizer and rice plant disease diagnosis using Plantix mobile app, data from the proposed system was entered into this mobile application Plantix. This application use soil test data to determine the precise fertilizer needs for rice crops. Figs. 6 and 7 illustrate snapshots of this application, showing recommended fertilizer rates and diagnosis based on the soil data analyzed.

4. Findings and discussion of essential requirements

The results of soil analysis of rice crops reveal several important insights regarding soil temperature, pH, moisture content, salinity, EC and nutrient levels. The data offers insights into the variability of these parameters across different soil samples, which are crucial for understanding soil health and optimizing agricultural practices. These factors jointly affect crop metabolism, growth, and productivity. The most important parameters for rice crops are discussed below.

Soil Temperature: The soil temperature across the samples ranged from 30.5 °C to 33.2 °C, with an average of approximately 31.8 °C. This range is typical for rice paddies, where higher temperatures can influence microbial activity and nutrient availability. Temperature is an important factor affecting the growth of rice, and the measured temperatures suggest a generally favorable condition for rice cultivation.

Soil Moisture: Soil moisture values varied from 60.6 % to 94.1 %, showing a wide range of water retention across samples. Sample 7 exhibited the highest moisture content at 94.1 %, which indicates well-watered conditions, potentially beneficial for rice cultivation. However, excess moisture can also lead to waterlogging, which may affect root respiration and nutrient uptake. Conversely, sample 1 had a lower moisture content (60.6 %), which may indicate dryer conditions, requiring irrigation management for optimal rice growth.

Soil pH: The soil pH ranged from 7.13 to 8.33, indicating a generally neutral to slightly alkaline pH across the samples. Rice plants typically thrive in slightly acidic to neutral soil (pH 5.5–7.0), but the observed pH values are within acceptable limits for rice cultivation. Samples with pH values above 8 (e.g. 3, 6, 7, and 13) indicate a risk of reduced absorption of nutrients, especially phosphorus, because high alkalinity may hinder its availability, potentially affecting crop metabolism and growth.

Nutrient Content:

- **Nitrogen:** The nitrogen content ranged from 71 mg/kg to 103 mg/kg. Nitrogen is a vital nutrient for rice growth, and the measured levels are adequate for promoting vegetative growth. Sample 3 had the highest nitrogen content at 103 mg/kg, which could support strong vegetative development.
- **Phosphorus:** Phosphorus values were relatively uniform, ranging from 15 mg/kg to 19 mg/kg, with sample 14 having the highest phosphorus content at 19 mg/kg. Phosphorus is essential for root development and flowering in rice, and the measured values are within a typical range for healthy rice soils.
- **Potassium:** Potassium levels ranged from 101 mg/kg to 141 mg/kg, with sample 6 having the highest value of 141 mg/kg. Potassium is crucial for regulating plant water balance and improving disease resistance, and the observed levels are sufficient for supporting optimal growth and yield.

The variation in these parameters across the different soil samples underscores the need for tailored management practices in rice farming. While some samples exhibit favorable conditions for rice growth, others may require adjustments in irrigation, fertilization, or pH management to optimize productivity. For instance, soil samples with higher EC and salinity may need strategies to manage soil salinity, while those with lower moisture content might require additional irrigation. The dataset highlights the complex interplay of environmental factors and nutrients in determining the sustainability and productivity of rice farming. The 8-in-1 soil sensor dataset provides valuable insights into the soil conditions and nutrient availability for sustainable rice farming. These parameters

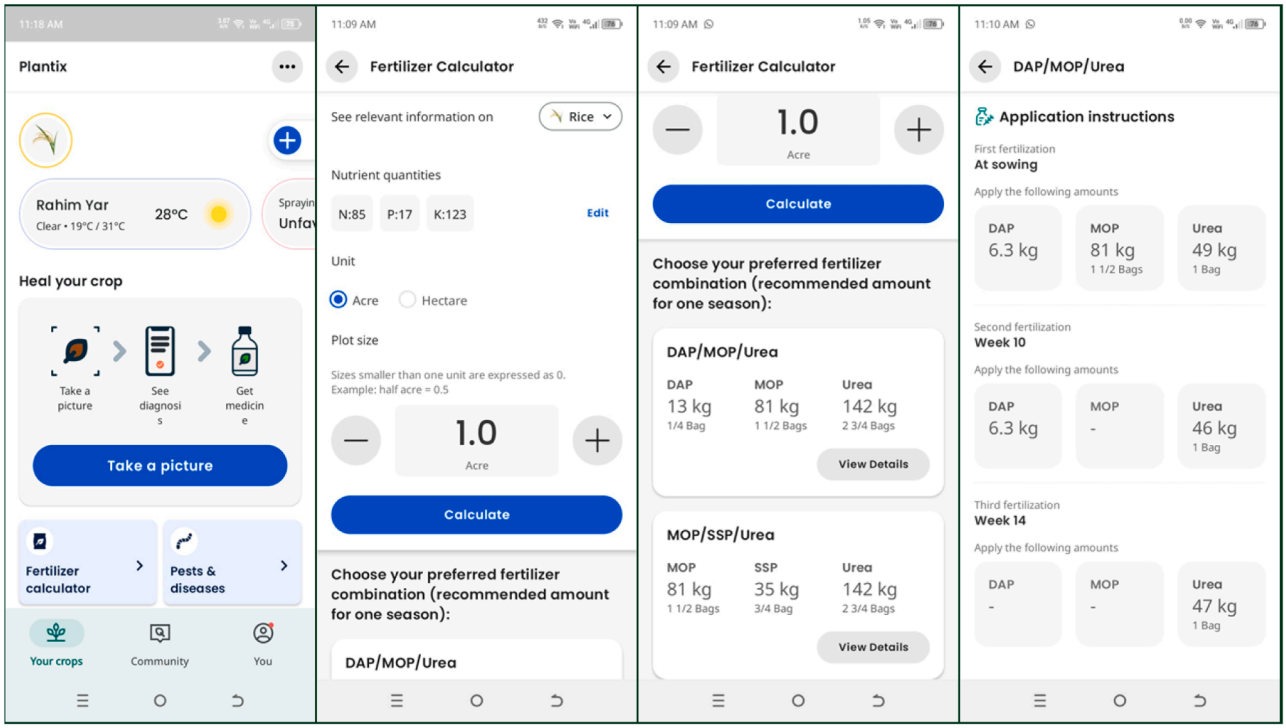


Fig. 6. Plantix mobile app.

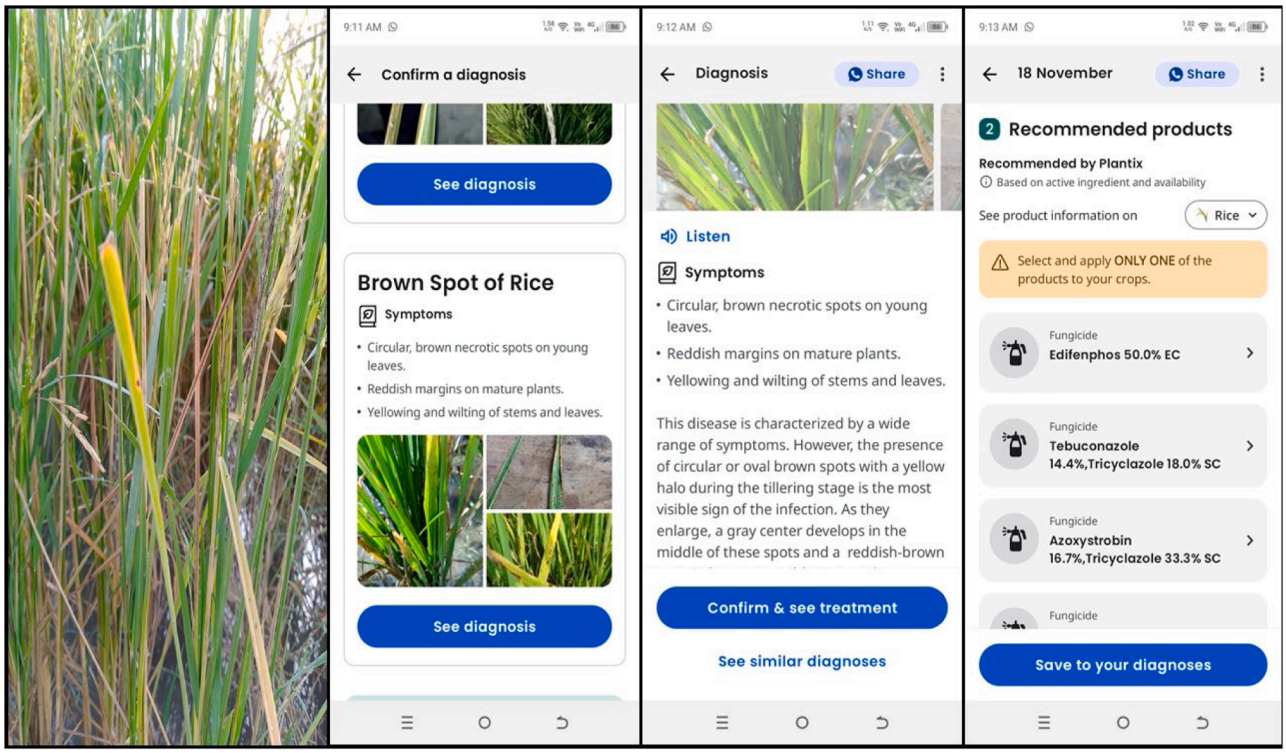


Fig. 7. Rice plants disease diagnosis using Plantix mobile app.

can guide informed decision-making for improving crop management practices, enhancing productivity, and ensuring environmental sustainability in rice cultivation.

5. Conclusions

IoT-based smart farming research presented in this article integrates metabolic monitoring of rice plants with key soil parameters: temperature, moisture, salinity, electrical conductivity (EC), pH, nitrogen, phosphorus, and potassium. By providing farmers with actionable insights, this system empowers them to make data-driven decisions, thereby increasing crop growth, development and grain yield. Field tests on rice crops in Rahim Yar Khan, Pakistan, validate the system's effectiveness, showing how real-time data monitoring enables precise adjustments in crop management to optimize growing conditions. The variation in soil parameters observed during testing, such as moisture levels and pH, shows the importance of continuous monitoring to maintain optimal conditions for rice crop metabolism and productivity. By integrating IoT with cloud computing, predictive algorithm and AI-driven mobile application, the system provides a scalable solution for modern agriculture, enabling efficient use of resources, reduced environmental impact, and improved crop resilience. This research addresses the need for data-driven decision-making in agriculture, paving the way for more sustainable and productive agricultural practices that can meet the demands of the world's growing population.

Ethical statement

The research described in this article was performed and conducted ethically, adhering to relevant guidelines and ensuring no harm to humans, animals, or the environment.

CRediT authorship contribution statement

Hammad Shahab: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Muhammad Naeem:** Software, Methodology, Formal analysis. **Muhammad Iqbal:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition. **Muhammad Aqeel:** Writing – original draft, Software, Methodology, Investigation, Data curation. **Syed Sajid Ullah:** Resources, Funding acquisition, Data curation.

Declaration of competing interest

The authors declare no conflicts of interest affecting the integrity of this work.

Acknowledgment

The authors acknowledge and thank the University of Agder, Norway, for their financial support, which was instrumental in enabling this research and achieving the results presented.

Data availability

No data was used for the research described in the article.

References

[1] S. Bhatnagar, et al., Exploring the dynamics of climate-smart agricultural practices for sustainable resilience in a changing climate, Environ. Sustain. Indic. 24 (2024) 100535, https://doi.org/10.1016/j.indic.2024.100535.
[2] N.A. Rust, et al., Have farmers had enough of experts? Environ. Manage. 69 (1) (2022) 31–44, https://doi.org/10.1007/s00267-021-01546-y.
[3] H. Shahab, M. Iqbal, A. Sohaib, A. Ur Rehman, A. Bermak, K. Munir, Design and implementation of an IoT-based monitoring system for early detection of lumpy skin disease in cattle, Smart Agric. Technol. 9 (2024) 100609 [Online]. Available: https://ssrn.com/abstract=4941578.
[4] B.M. Zerihun, T.O. Olwal, M.R. Hassen, Design and analysis of IoT-based modern agriculture monitoring system for real-time data collection. Computer Vision and Machine Learning in Agriculture, Volume 2, Springer, 2022, pp. 73–82, https://doi.org/10.1007/978-981-16-9991-7_5.
[5] V. Bhatnagar, R. Chandra, IoT-based soil Health monitoring and Recommendation system. Internet of Things and Analytics for Agriculture, Volume 2, Springer, 2020, pp. 1–21, https://doi.org/10.1007/978-981-15-0663-5_1.

[6] M. Cao, Y. Chen, W. Duan, Y. Li, J. Qin, Comprehensive evaluation of water–Energy–Food system security in the China–Pakistan Economic Corridor, Water 14 (12) (2022) 1900, https://doi.org/10.3390/w14121900.
[7] M.A. Khan, I. Uddin, N.S. Othman, Factors determining water stress: do environmental distress, energy consumption and industrialization matter? Environ. Dev. Sustain. (2024) 1–25, https://doi.org/10.1007/s10668-024-04653-y.
[8] F.Z. Jatoi, Agriculture in Pakistan and it's impact on economic growth, SSRN Electron. J. 103 (2021) 47–60, https://doi.org/10.2139/ssrn.3771914.
[9] A. Dobermann, et al., Responsible plant nutrition: a new paradigm to support food system transformation, Glob. Food Sec. 33 (2022) 100636, https://doi.org/10.1016/j.gfs.2022.100636.
[10] G. Lemaire, L. Tang, G. Bélanger, Y. Zhu, M.H. Jeuffroy, Forward new paradigms for crop mineral nutrition and fertilization towards sustainable agriculture, Eur. J. Agron. 125 (2021) 126248, https://doi.org/10.1016/j.eja.2021.126248.
[11] A.O. Klein, L. Carlisle, M.G. Lloyd, N.F. Sayre, T.M. Bowles, Understanding farmer knowledge of soil and soil management: a case study of 13 organic farms in an agricultural landscape of northern California, Agroecol. Sustain. Food Syst. 48 (1) (2024) 17–49, https://doi.org/10.1080/21683565.2023.2270451.
[12] X. Zhang, J. Zhang, L. Li, Y. Zhang, G. Yang, Monitoring citrus soil moisture and nutrients using an IoT based system, Sensors 17 (3) (2017) 447, https://doi.org/10.3390/s17030447.
[13] M.B. Abhishek, S. Tejashree, R. Manasa, T.G. Vibha, Smart agriculture management system using internet of things (IoT), Lect. Notes Netw. Syst 176 (3) (2021) 363–375, https://doi.org/10.1007/978-981-33-4355-9_28. LNNS.
[14] F.J. Mesas-Carrascosa, D. Verdú Santano, J.E. Meroño, M. Sánchez de la Orden, A. García-Ferrer, Open source hardware to monitor environmental parameters in precision agriculture, Biosyst. Eng. 137 (2015) 73–83, https://doi.org/10.1016/j.biosystemseng.2015.07.005.
[15] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, E.H.M. Aggoune, Internet-of-things (IoT)-based smart agriculture: toward making the fields talk, IEEE Access 7 (2019) 129551–129583, https://doi.org/10.1109/ACCESS.2019.2932609.
[16] R. Pitakaso, K. Sethanan, K.H. Tan, A. Kumar, A decision support system based on an artificial multiple intelligence system for vegetable crop land allocation problem, Ann. Oper. Res. (2023) 1–36, https://doi.org/10.1007/s10479-023-05398-z.
[17] M. Anjaneyulu, et al., Sustainable practices for plant nutrient control using IoT, E3S Web of Conf. 507 (2024) 1039, https://doi.org/10.1051/e3sconf/202450701039.
[18] R. Chin, C. Catal, A. Kassahun, Plant disease detection using drones in precision agriculture, Precis. Agric. 24 (5) (2023) 1663–1682, https://doi.org/10.1007/s11119-023-10014-y.
[19] Z.L. Li, et al., Satellite remote sensing of global land surface temperature: definition, methods, products, and applications, Rev. Geophys. 61 (1) (2023), https://doi.org/10.1029/2022RG000777.
[20] V.M. Ngo, T.V.T. Duong, T.B.T. Nguyen, C.N. Dang, O. Conlan, A big data smart agricultural system: recommending optimum fertilisers for crops, Int. J. Inf. Technol. 15 (1) (2023) 249–265, https://doi.org/10.1007/s41870-022-01150-1.
[21] T. Alahmad, M. Neményi, A. Nyéki, Applying IoT sensors and big data to improve precision crop production: a review, Agronomy 13 (10) (2023) 2603, https://doi.org/10.3390/agronomy13102603.
[22] R. Hartono, N. Maulana Yoeseph, F. Aji Purnomo, M. Asri Safi'ie, S. Alim Tri Bawono, Portable internet of things-based soil nutrients monitoring for precision and efficient smart farming, Bull. Electr. Eng. Informat. 13 (5) (Oct. 2024) 3326–3333, https://doi.org/10.11591/eei.v13i5.7928.
[23] M.R. Islam, K. Oliullah, M.M. Kabir, M. Alom, M.F. Mridha, Machine learning enabled IoT system for soil nutrients monitoring and crop recommendation, J. Agric. Food Res. 14 (2023) 100880, https://doi.org/10.1016/j.jafr.2023.100880.
[24] K.M. Sumana, K.S. Rekha, A study of efficient management of nutrients in fertilizers using IoT and ML, in: Proceedings - International Conference on Applied Artificial Intelligence and Computing, ICAAIC 2022, 2022, pp. 574–577, https://doi.org/10.1109/ICAAIC53929.2022.9793124.
[25] M. Ma, W. Cen, R. Li, S. Wang, J. Luo, The molecular regulatory pathways and metabolic adaptation in the seed germination and early seedling growth of rice in response to low o2 stress, Plants 9 (10) (2020) 1–14, https://doi.org/10.3390/plants9101363.
[26] Z. Ozbekova, A. Kulmyrzaev, Study of moisture content and water activity of rice using fluorescence spectroscopy and multivariate analysis, Spectrochim. Acta Part A Mol. Biomol. Spectrosc. 223 (2019) 117357, https://doi.org/10.1016/j.saa.2019.117357.
[27] M.S. Arshad, M. Farooq, F. Asch, J.S.V Krishna, P.V.V. Prasad, K.H.M. Siddique, Thermal stress impacts reproductive development and grain yield in rice, Plant Physiol. Biochem. 115 (2017) 57–72, https://doi.org/10.1016/j.plaphy.2017.03.011.
[28] C. Zhao, et al., Early flowering and rapid grain filling determine early maturity and escape from harvesting in weedy rice, Pest Manag. Sci. 74 (2) (Feb. 2018) 465–476, https://doi.org/10.1002/ps.4730.
[29] J. Sardans, J. Peñuelas, Potassium control of plant functions: ecological and agricultural implications, Plants 10 (2) (2021) 1–31, https://doi.org/10.3390/plants10020419.
[30] N. Ahmed, et al., Role of macronutrients in cotton production. Cotton Production and Uses: Agronomy, Crop Protection, and Postharvest Technologies, Springer Singapore, Singapore, 2020, pp. 81–104, https://doi.org/10.1007/978-981-15-1472-2_6.
[31] M.R. Thakur, V.M. Bhale, A.N. Paslawar, Mobility of N, P and K and root growth of bt and nonBt cotton in clayey soil under different NPK levels, Ecol. Environ. Conserv. 28 (2022) 419–424, https://doi.org/10.53550/eec.2022.v28i07s.069.
[32] S.I. Hassan, M.M. Alam, U. Illahi, M.A. Al Ghamdi, S.H. Almotiri, M.M. Su'ud, A systematic review on monitoring and advanced control strategies in smart agriculture, IEEE Access 9 (2021) 32517–32548, https://doi.org/10.1109/ACCESS.2021.3057865.
[33] R. Alfred, J.H. Obit, C.P.-Y. Chin, H. Haviluddin, Y. Lim, Towards paddy rice smart farming: a review on big data, machine learning, and rice production tasks, IEEE Access 9 (2021) 50358–50380, https://doi.org/10.1109/ACCESS.2021.3069449.
[34] N.A. Anjum, A. Masood, S. Umar, N.A. Khan, Introductory chapter: phosphorus in soils and plants. Phosphorus in Soils and Plants, IntechOpen, 2024, https://doi.org/10.5772/intechopen.113397.
[35] S. Dudala, S.K. Dubey, S. Goel, Microfluidic soil nutrient detection system: integrating nitrite, pH, and electrical conductivity detection, IEEE Sens. J. 20 (8) (2020) 4504–4511, https://doi.org/10.1109/JSEN.2020.2964174.
[36] M. Gavrilescu, Water, soil, and plants interactions in a threatened environment, Water 13 (19) (Oct. 2021) 2746, https://doi.org/10.3390/w13192746.
[37] J.L. Rotundo, et al., Development of a decision-making application for optimum soybean and maize fertilization strategies in Mato Grosso, Comput. Electron. Agric. 193 (2022) 106659, https://doi.org/10.1016/j.compag.2021.106659.
[38] A. Pathak, M.A. Uddin, M. Jainal Abedin, K. Andersson, R. Mustafa, M.S. Hossain, IoT based smart system to support agricultural parameters: a case study, Procedia Comput. Sci. 155 (2019) 648–653, https://doi.org/10.1016/j.procs.2019.08.092.
[39] P.K. Cheung, C.Y. Jim, N. Tapper, K.A. Nice, S.J. Livesley, Daytime irrigation leads to significantly cooler private backyards in summer, Urban Clim. 46 (2022) 101310, https://doi.org/10.1016/j.uclim.2022.101310.
[40] vol. 4, D. B. T.-E. of S. in the C.L. Ping, Cold-region soils, in: E. Hillel (Ed.), Encyclopedia of Soils in the Environment, Elsevier, Oxford, 2004, pp. 268–276, https://doi.org/10.1016/B0-12-348530-4/00024-2. vol. 4, D. B. T.-E. of S. in the.
[41] M. Hawkesford, et al., Functions of macronutrients. Marschner's Mineral Nutrition of Higher Plants: Third Edition, Elsevier, 2011, pp. 135–189, https://doi.org/10.1016/B978-0-12-384905-2.00006-6.
[42] A. B. T.-E. M A. Batool, et al., Effect of macronutrients management on nutrients uptake, partitioning, growth, and yield attributes in plants, in: P.-S.S. Husen (Ed.), Essential Minerals in Plant-Soil Systems, Elsevier, 2024, pp. 129–144, https://doi.org/10.1016/b978-0-443-16082-0.00015-1. A. B. T.-E. M.
[43] Y. Liu, et al., Genomic basis of geographical adaptation to soil nitrogen in rice, Nature 590 (7847) (Feb. 2021) 600–605, https://doi.org/10.1038/s41586-020-03091-w.
[44] A.S. Elrys, et al., Patterns and drivers of global gross nitrogen mineralization in soils, Glob. Chang. Biol. 27 (22) (Nov. 2021) 5950–5962, https://doi.org/10.1111/gcb.15851.
[45] H. Shahab, T. Abbas, M.U. Sardar, A. Basit, M.M. Waqas, H. Raza, Internet of things implications for the adequate development of the smart agricultural farming concepts, Big Data Agric 3 (1) (Nov. 2020) 12–17, https://doi.org/10.26480/bda.01.2021.12.17.
[46] F. Qaisar, H. Shahab, M. Iqbal, H.M. Sargana, M. Aqeel, M.A. Qayyum, Recent trends in Cloud computing and IoT platforms for IT management and development: a review, Pak. J. Eng. Technol. 6 (1) (Mar. 2023) 98–105, https://doi.org/10.51846/vol6iss1pp98-105.
[47] H. Shahab, et al., Real-time health monitoring smart system for cardiac patients using internet of things (IoT), Int. J. Electr. Eng. Emerg. Technol. 6 (1) (2023) 31–37.
[48] K. Obaideen, et al., An overview of smart irrigation systems using IoT, Energy Nexus 7 (2022) 100124, https://doi.org/10.1016/j.nexus.2022.100124.
[49] E. Vories, K. Sudduth, Determining sensor-based field capacity for irrigation scheduling, Agric. Water Manag. 250 (2021) 106860, https://doi.org/10.1016/j.agwat.2021.106860.
[50] D. Deshpande, S. Jadhav, R. Chounde, T. Kachare, K. Bhale, P. Waso, A survey on the role of IoT in agriculture for smart farming, Proc. 1st IEEE Int. Conf. Artif. Intell. Mach. Vision, AIMV 2021 7 (2021) 156237–156271, https://doi.org/10.1109/AIMV53313.2021.9670901.
[51] H. Shahab, M.M. Waqas, M. Muthmainnah, Revolutionizing manufacturing. Machine Vision and Industrial Robotics in Manufacturing, CRC Press, Boca Raton, 2024, pp. 385–404, https://doi.org/10.1201/9781003438137-21.
[52] L. Horrigan, R.S. Lawrence, P. Walker, How sustainable agriculture can address the environmental and human health harms of industrial agriculture, Environ. Health Perspect. 110 (5) (2002) 445–456, https://doi.org/10.1289/ehp.02110445.
[53] Y. Zhang, C. Xia, X. Zhang, Y. Sha, G. Feng, Q. Gao, Quantifying the relationships of soil properties and crop growth with yield in a NPK fertilizer application maize field, Comput. Electron. Agric. 198 (2022) 107011, https://doi.org/10.1016/j.compag.2022.107011.
[54] A. Balkrishna, R. Pathak, S. Kumar, V. Arya, S.K. Singh, A comprehensive analysis of the advances in Indian Digital agricultural architecture, Smart Agric. Technol. 5 (Oct. 2023) 100318, https://doi.org/10.1016/j.atech.2023.100318.
[55] F.B.N. Tonle, et al., A road map for developing novel decision support system (DSS) for disseminating integrated pest management (IPM) technologies, Comput. Electron. Agric. 217 (Feb. 2024) 108526, https://doi.org/10.1016/j.compag.2023.108526.

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

Printed RFID sensing system: The cost-effective way to IoT smart agriculture / Gómez-Gijón, S., Salmerón, J. F., Falco, A., Loghin, F. C., Lugli, P., Morales, D. P., Rodríguez, N., & Rivadeneyra, A.

Computers and Electronics in Agriculture
Volume 232 (2025) 110116 Pages 1-10
https://doi.org/10.1016/j.compag.2025.110116
(Database: ScienceDirect)

Computers and Electronics in Agriculture 232 (2025) 110116

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Printed RFID sensing system: The cost-effective way to IoT smart agriculture

Sonia Gómez-Gijón [a], José F. Salmerón [b], Aniello Falco [c], Florin C. Loghin [b], Paolo Lugli [c], Diego P. Morales [a], Noel Rodríguez [a], Almudena Rivadeneyra [a,*]

[a] Department of Electronics and Computer Technology, University of Granada, Granada, Spain
[b] Institute for Nanoelectronics, Technical University of Munich, Munich, Germany
[c] Faculty of Science and Technology, Free University of Bozen-Bolzano, Bozen-Bolzano, Italy

ARTICLE INFO

Keywords:
Humidity
Internet of Things (ioT)
Oxygen concentration
Robot
Screen printing
Sensor tag
Spray deposition
Temperature

ABSTRACT

The increase in global population, climate change and the scarcity of resources are some of the reasons why the improvement of the agricultural sector has become a crucial issue for the future of our society. Recently, the growth of Internet of Things (IoT) has meant a revolution in many industries, including the agricultural sector, where the employment of this technology is considered the ideal formula to boost crop efficiency. In this work we present a complete high performance IoT solution, consisting of an innovative printed smart radio frequency identification (RFID) tag and an autonomous measuring system based on a robotic platform, designed for soil monitoring purposes. The smart RFID tags are employed to obtain soil parameters, crucial for the correct development of the crops: humidity, oxygen concentration and temperature. The tag is novelly developed on a humidity dependent and flexible transparent foil substrate. The former feature enables the direct measurement of relative humidity (RH) by a capacitive structure defined on the substrate while the latter characteristic allows the sensing of oxygen content via a membrane whose luminescence intensity changes with the influence of this parameter. Effects on the antenna response are measured when the tag is placed in soil environment at root depth to validate the excellent sensing behavior of the tag.

Finally, to advance the scope of this work and to show the promising performance of the tags as part of a monitoring system in the field of smart agriculture, a robot equipped with an UHF reader and IoT features, capable of automating the finding and reading of the designed RFID tags, is designed. This robot sends the information via internet for its further processing, thus constituting a low-cost, large-scale and environmentally friendly solution, ideal for enhancing crops performance.

1. Introduction

Given the steady increase of human population – forecasted to reach 9 billion individuals by mid of the 21st century – optimizing the allocation of resources and minimizing the wastes suddenly become compelling priorities (Godfray et al., 2010). Water, electrical energy, food, and agricultural lands are, possibly, the most important resources, which our constantly progressing society needs to manage effectively.

Food, particularly, is strictly intertwined to the other resources, as the production chain of edibles – plant or animal based – relies on the exploitation of land and transforming water, energy and raw materials into a finite product. Since the available land reduces, the food producers must face the need to intensify their crops and create more efficient processes (Godfray et al., 2010).

A first step in this direction is creating smart sensing tags, fabricated with cost-effective and green processes and at the same time sustainable and affordable in high volume for the food producers. Such tags should be employed to detect vital parameters of plants crops – or plants, in general – to be able to irrigate or fertilize on demand, only when certain vital parameters exceed tolerable ranges.

Plants' roots are affected by different factors such as temperature (it has influence in roots permeability), aeration (it alters radical metabolism) and salinity (this property modifies water and nutrients absorption). When radical volume decreases, the temperature range experimented by the roots increases. This behavior might be caused by oxygen depletion in the roots and the microflora (Roberts et al., 2006). For this reason, some important parameters to be monitored in soils at typical root depth are the root temperature (Urrestarazu and Gavilán, 2004), aeration (Urrestarazu and Gavilán, 2004) and humidity (Tibbitts, 1979).

Previous works have already described the integration of different sensors in RFID tags to perform soil monitoring (da Fonseca et al., 2017; Korošak et al., 2019). However, all this studies are based on a

* Correspondence to: University of Granada, Dept. Electronics and Computer Technology Campus Fuentenueva s/n, 18071, Granada, Spain
E-mail address: arivadeneyra@ugr.es (A. Rivadeneyra).

https://doi.org/10.1016/j.compag.2025.110116
Received 15 April 2024; Received in revised form 30 December 2024; Accepted 7 February 2025
Available online 21 February 2025

conventional technological approach using rigid printed circuit boards. In Kim et al. (2014) Kim et al. already presented a inkjet-printed paper based RFID tag to analogically sense moisture content in agricultural applications. Paper, however, gives less flexibility in terms of possible measurement techniques, as it is an opaque substrate. For these reasons, we present a novel flexible and ecologically sustainable wireless printed soil tag based on cellulose acetate with the objective of monitoring different soil parameter at root depth. In our case, we found cellulose acetate to be the perfect candidate due to its transparency, ideal for the oxygen sensor, and humidity dependency of its dielectric constant, allowing the construction of the moisture sensor on it. By using this biodegradable substrate, we ensure that the tags can degrade harmlessly in the soil after their functional lifespan, avoiding ecological issues associated with traditional plastic materials. While the majority of previous studies have primarily concentrated on sensors for soil moisture, temperature, or a combination of both, our system advances this approach by incorporating these measurements alongside oxygen concentration within a single platform. Additionally, hence, we employ scalable deposition techniques, with potentially high throughput to develop a fully functional sensor tag, whose measurement capabilities can be easily extended to a plethora of printed sensors on semitransparent substrates (Qian and Long, 2018). The tag was developed employing printing techniques and the minimum amount of conventional integrated circuits, in order to minimize its costs and printed area. In particular, radiofrequency identification (RFID) technology is used as method of data capture, while the monitored parameters are temperature, relative humidity and oxygen content. On the other hand, IoT technology has become increasingly important in recent years, capable of making systems and devices more complex and intelligent. In the agricultural sector, not only monitoring parameters through sensors is enough but also the use of IoT is indispensable to store the acquired data, process that information and act as quickly as possible according to the crop needs. Applications of IoT to agriculture are abundant; irrigation management system, pest and disease control, greenhouse condition or water quality and soil monitoring. Previous works have shown the rising potential of IoT for this type of applications (Dholu and Ghodinde, 2018; Rao and Sridhar, 2018,?; Dagar et al., 2018; Yoon et al., 2018) but all of them are based on wired sensors with wireless connectivity. Our purpose here, is to make a complete wireless system, including from the sensors to the reading and visualization of the parameters. However, our work shares with all the other RFID works presented so far a main common drawback: the relatively low reading range of RFID and similar technologies. This renders the practical realization of a IoT multi-node smart agriculture platform significantly difficult. To tackle this issue, we also introduce a strategy that allows the monitoring of many buried nodes, controlling big farming areas without wiring. In fact, while long distance wireless reading is not possible in this context, we changed the perspective and realized a moving reader, which can scan the complete terrain. This has been achieved by means of a mobile robot, which is capable of receiving information from each and every RFID and transmit the input to a database via Wi-Fi. These data can then be remotely processed and enable the application of novel agriculture methodologies, such as greenhouse automation, crop management and predictive analytics for smart farming (Ratnaparkhi et al., 2020; Ayaz et al., 2019). In summary, our research presents several innovative aspects that distinguish our RFID tags from existing solutions. Firstly, the integration of a cellulose-based transparent substrate that simultaneously measures temperature, soil moisture, and oxygen concentration offers a rather complete monitoring system, which is rarely found in the literature. Secondly, the use of biodegradable materials aligns with current environmental directives, enhancing the sustainability of agricultural practices. Additionally, our proof of concept with a mobile robot not only demonstrates the practical application of these tags but also highlights their potential for scalable monitoring in diverse agricultural environments.

2. Materials and methods

2.1. Tag design

The block diagram of the proposed RFID tag we employed in this work is illustrated in Fig. 1(a). It is a microcontroller-based system designed for a wireless transmission of sensing parameters. This scheme is aimed at transmitting the measured data using the protocol EPC Gen 2 which operates in the ultra-high frequency (UHF) band (868 MHz) that is widely employed for RFID applications. This work develops the antenna and all the interconnection as well as the humidity sensor over a flexible substrate.

Fig. 1(b) presents the architecture of the printed tag, including the footprints of the required components. This tag is a semi-passive architecture based on SL900A RFID chip (AMS AG, Unterpremstaetten, Austria) compatible with EPC Gen 2 RFID standard (AG, 2014). This RFID chip has been selected because it integrates a sensor front-end (SFE) that offers different sensor conditioning stages and a 10-bits analog to digital converter (ADC). Thanks to this extra circuitry, the relative humidity sensor has been directly integrated in the tag, as presented in Salmerón et al. (2014d). Moreover, the chip includes a built-in temperature sensor. A typical dipole antenna resonating at 868 MHz (European Band for RFID UHF) composed the radiofrequency interface and a RF Surface Mount Device (SMD) inductor is used to match the input impedance of the chip to the antenna (Rao et al., 2005).

The system for soil monitoring described here can determine oxygen concentration, moisture content and temperature. The first parameter is obtained by the subsystem formed by the oxygen sensing membrane together with the excitation and detection electronics. As presented by Martínez-Olmos et al. (2013), we have used a color detector where the red (R) coordinate of the Red Green Blue (RGB) color space obtained directly as the output of a digital color detector quantifies the intensity (Figure S1a in supporting information).

In particular, the membrane is optically excited by a surface-mount light emitting diode (LED) (OCU-400UE390, OSA Opto Light, Berlin, Germany) with peak emission at 385 nm, corresponding to the UV band. The phosphorescence intensity is related to the oxygen concentration. This emission is captured by a color detector (S9706, Hamamatsu Photonics, Japan) that has been already employed to determine the color of virgin olive oil (Salmerón et al., 2012) as well as for oxygen monitoring (Martínez-Olmos et al., 2013). Although the three RGB components are simultaneous obtained, only the red and blue coordinates are taken into account. The R coordinate is associated exclusively with the luminescence generated by the sensitive layer, depending on the intensity of this emission and, in other words, on the oxygen concentration (Lopez-Ruiz et al., 2012; Park et al., 2010). The B coordinate is correlated with the visible emission of the LED and, thus, we use it as reference signal to compensate fluctuations in the source illumination (Wang et al., 2010). A PIC18LFK22 microcontroller (Microchip Technology Inc., Chandler, AZ, USA) with low power consumption (nanoWatt XLP Technology) is used to bias the LED in a pulsed mode and control the S9706 detector (Martínez-Olmos et al., 2013).

The humidity sensor is a printed planar interdigitated electrodes (IDE) capacitor. This kind of structures provides more direct interaction between the surrounding environment and the sensor compared to others (Mamishev et al., 2004; Rivadeneyra et al., 2016). The traditional approach to provide gas sensitivity is to deposit a sensing layer onto the electrodes with some dependent electrical property, in the case of capacitive sensors, frequently its electrical permittivity. Nevertheless, an easier strategy is to directly use a sensitive substrate (Virtanen et al., 2010; Briand et al., 2011; Rivadeneyra et al., 2014a,b, 2015), which guarantees faster processing and lower material consumption (Figure S1b).

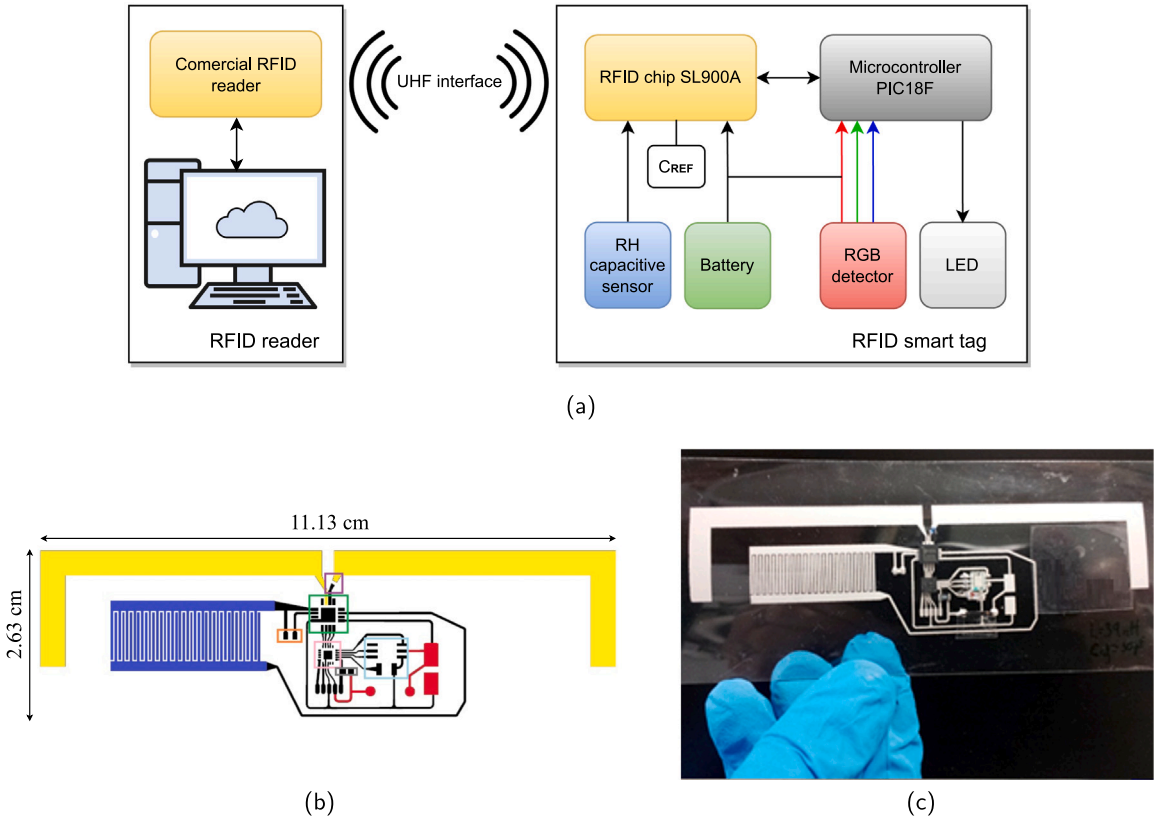


Fig. 1. (a) Block diagram of the developed RFID tag. Black connections represent data lines, red connections are principal power lines and blue connections are secondary power lines. Battery is optional. (b) Schematics of the fabricated RFID tag. (c) Photograph of the printed tag.

Cellulose acetate is a well-known chemical whose electrical behavior exhibits a high dependence to humidity (Sakai et al., 1996; Molina-Lopez et al., 2012). In particular, the relationship between the electrical permittivity of this substance and the relative humidity is given by:

$$\epsilon_r = \epsilon_{r0} + \alpha \cdot RH(\%) \quad (1)$$

where RH is the relative humidity in percentage and ϵ_{r0} and α are material dependent parameters.

This relative permittivity also varies, among other parameters, due to work frequency and environmental temperature. These factors should be taken into account in order to reduce or compensate the interferences with the humidity measurement.

The designed capacitive sensor occupies an area of 3.48 cm^2 (L = 1.10 cm x W = 3.16 cm) composed of 40 fingers (20 fingers for each electrode) with 200 μm width and inter-spacing (see Table S1). These dimensions could be decreased introducing a different screen-printing mesh density as stated in Salmerón et al. (2014a) but, in our case, these are the minimum dimensions that ensure reproducibility of the designed structure. The humidity sensor is directly connected to the SFE of the RFID chip, which has been configured in capacitance mode. In this mode, an external reference capacitor (C_{REF}) in series with the printed sensor (C_{SENS}) is required to interface the sensor, as depicted in supporting information (Figure S1c).

The measurement of the capacitive sensor is done by exciting it with a 100 kHz square wave signal whose amplitude is equal to V_{ref1} voltage, being the input voltage on the analog-to-digital converter (ADC) associated with the capacitance according to:

$$V_{ADC} = V_{REF1} \cdot \frac{C_{REF}}{C_{REF} + C_{SENS}} + V_{REF1} \quad (2)$$

Further explanation of this capacitive mode is found in Salmerón et al. (2014d).

Temperature value comes from a conversion in the on-chip A/D converter of the SL900A. Two internal voltage references, V_{ref1} and V_{ref2}, whose value can be selected through the internal registers of the chip in steps of 50 mV between 160 and 610 mV, set the lower and upper limits of the conversion. The lower limit is given by V_{ref2} while the upper limit is given by $2V_{ref2} - V_{ref1}$. The difference between them defines the input voltage range and the limits of operation. As the voltage references can be set by the user, a concrete resolution and range of temperature can be selected. The minimum resolution is 0.18 °C in a range of 189.9 °C while a resolution of 0.23 °C is obtained with the widest range of 237.2 °C.

In addition to this, the SL900A chip can be powered up by either an external battery or the RF field radiated by the reader, a RFID strategy know as semi-passive. This implies being possible to operate in active (battery assisted) or passive (no battery) modes. The tag could perfectly work without battery, but we have included it to add the capability of data logging: sensor data are periodically stored in the tag including a timestamp, but an RFID reader is needed to retrieve the information. Therefore, it is possible to monitor the soil state when no external RFID reader is present (red lines in Fig. 1(a)). It is also possible to store the information only when the sensed parameters are out of threshold values by setting the highest and lowest limits values of each parameter (Salmerón et al., 2014c).

For this purpose, a rechargeable polymer lithium-ion battery PGEB01 (General Electronics Co, Shenzhen, P.R. China) battery has been included in the tag with 3.7 V and a capacity of 40 mAh. Power consumption of this tag is virtually similar to the one presented in Martínez-Olmos et al. (2013), 50 μA in idle mode and 3.8 mA while performing the measurements. It is important to note that the battery is not present in Fig. 1(c) where the photographs of the designed tags are shown. To configure the tag, read data and properly present them, a Visual Basic application (Microsoft Corp., Redmond, WA, USA) has been used.

2.2. Fabrication process

As it can be seen in Fig. 1(c), the tag layout is fabricated on a transparent flexible substrate. Apart from the luminophore, all the components are printed and placed on the same side. The luminophore is deposited on the opposite side, just covering the area of both color detector and LED. In this way, both the exciting emission and the luminescence response travel from the LED to the membrane and from the membrane to the color detector, as illustrated in Figure S1a.

The selected substrate was 100-μm tick cellulose acetate (AC311100 from Goodfellow, Cambridge, UK). This film has been chosen for two main reasons: primarily, the film must be transparent enough to visible light to allow the LED to stimulate the luminophore and secondarily, this substrate presents a high sensitivity to humidity and thus, no other sensing layer is required to develop the moisture sensor. Apart from the lower complexity of the fabrication process, this approach yields another tangible advantage over competing strategies: as the sensible layer is on the opposite side with respect to the electronics and the IDEs, the former can be exposed to the soil, while the latter can be encapsulated and protected from the environmental agents. Furthermore, the biodegradability of the substrate make the tag environmentally friendly though it would affect its long term usability. However, as we selected a relatively thick substrate and the printed part is protected, this should not be an issue for many applications, such as greenhouses, where the growing season is typically shorter than six months and environmental conditions are carefully controlled. Additionally, studies suggest that cellulose acetate films with a high degree of substitution can maintain structural integrity for several months under similar conditions (Rumi et al., 2024; Yadav and Hakkarainen, 2022), and the protective ABS layer further enhances durability by shielding the printed components from environmental factors.

The printing technique to manufacture the tags has been screen printing due to higher performance of the antennas, as shown in our previous work (Salmerón et al., 2014a), where a minimum resistivity of ∼ 40 μΩ cm was obtained for a single layer of screenprinted ink on (PI) with 120 nylon threads/cm (Colella et al., 2017; Salmerón et al., 2014b). For this purpose, the fabrication of the prototypes has been performed with a screen-printing machine (FLAT-DX 100 from Siebdruck-Versand, Germany) with the optimal conditions previously reported to fabricate UHF smart tags: a 120 T/cm mesh was used with a silver conductive paste (Sigma Aldrich, USA) with a solid content higher than 75%. Finally, a three-step process to do the assembly of the chip to the foil was performed using the adhesive epoxy EPO-TEK H20E (Epoxy Technology, Inc., Billerica, USA), as detailed in Salmerón et al. (2014c). The preparation of the sensitive membrane is detailed in the supporting information.

The whole tag should be insulated from the surrounding soil in order to avoid the degradation of the different elements, only the backside of the interdigitated electrode should be open for environment moisture measurement. The enclosing of the tag should, however, allow the gas exchange. The insulating solution was obtained dissolving 2% wt. Acrylonitrile Butadiene Styrene (ABS, Sigma-Aldrich) in 1,2-Dichlorobenzene (Sigma-Aldrich), placed for 30 min in a bath sonicator and magnetically stirred for 2 h. The deposition of the insulating ABS layer on the IDE structure was performed by means of a semi-automated spray-coater. The setup is composed of a commercial spray-deposition head (Krautzberger GmbH, Germany) and an electromechanical pneumatic controller (self-made). The atomization pressure was kept at 2 bars, the substrate-to-nozzle distance was set to 10 cm, while the deposition temperature was set to 80 °C. The parameters were optimized in a previous work, in order to obtain a 5 μm thick smooth film onto arbitrary surfaces (Falco et al., 2016, 2015; Abdellah et al., 2010). All chemicals were weighed with a DV215CD balance (Ohaus Co., Pine Brook, NJ, USA) with a precision of ±0.01 mg.

2.3. Characterization

The antenna characterization has been carried out following the procedure described in Salmerón et al. (2014c). A UHF band commercial reader, DK-UHF RFID HP2 (IDS Microchip AG, Wollerau, Switzerland), fully compatible with EPC Gen 2 RFID standard has been employed to test the read range of the developed RFID tag. The electrical characterization of the capacitive humidity sensors has been performed by measuring their capacitances and dissipation factors using the four-wire measurement technique. The excitation voltage applied in all measurements has been VDC = 0 and VAC = 500 mV. The frequency sweep of analysis was from 100 kHz to 10 MHz. A complete compensation method has been implemented to eliminate the contribution of parasitic capacitances (Rivadeneyra et al., 2014a). The humidity and temperature responses of these humidity sensors have been measured in a climatic chamber VCL 4006 (Vötsch Industrietechnik GmbH, Reiskirchen, Germany) with a humidity range from 10% RH to 98% RH in a temperature range of + 10 °C to + 95 °C. The humidity deviation in time is ±1% to ±3% and the temperature deviation in time ±0.3 °C to ±0.5 °C. A commercial sensor (SHT15, Sensirion AG, Switzerland) has been employed to measure temperature and RH values in order to verify the data provided by the chamber. LabVIEW 2016 software (National Instruments Corporation, Texas, USA) has been used to control the full experimental set-up and record data.

The standard mixtures for the characterization of the oxygen sensor have been prepared with N2 as inert gas and controlling the flow rates of both O2 and N2 (>99%) gases (Air Liquid S.A., Madrid, Spain). The gases have been introduced in a mixing chamber using a computer-controlled mass flow controller (Air Liquid España S.A., Madrid, Spain) at a pressure of 760 Torr and a flow rate of 500 cm^3 min^{-1}. A thermostatic chamber has maintained a controlled temperature between −50 °C and +50 °C with ± 0.1 °C of accuracy for the thermal characterization of the system. For the electrical characterization of the system, a mixed signal oscilloscope (MSO4101, Tektronix, Beaverton, OR, USA), an $8^{1/2}$-bit Digital Multimeter 3158A (Keysight Technologies, Santa Clara, CA, USA), a 15 MHz waveform generator 33 120A (Keysight Technologies, Santa Clara, CA, USA), a precision Impedance Analyser 4294A and an impedance probe kit (4294A1) (Keysight Technologies, Santa Clara, CA, USA), a DC power supply E3630A (Keysight Technologies, Santa Clara, CA, USA) and a RFID reader DK-UHF RFID HP2 (IDS Microchip AG, Wollerau, Switzerland) have been used. A Visual Basic user interface was developed for the calibration and data analysis in a computer.

2.4. Autonomous navigation and measurement system design

Once the RFID sensor tag is designed, the final step is the automation of the reading process. The aim is to design a system capable of independently measure each and every node, transmit the readout data to a cloud storage and allow the final user to access it from anywhere. This would allow to reduce the costly repetitive human work, while at the same time increasing the reproducibility of the measurements.

For this purpose, we have used OSOYOO V2.1 Robot Car Kit (OSOYOO, 2020), a robot designed for beginners to learn programming and get hands-on experience on robot design and assembly. This robot is designed to be used with Arduino platform, but, when it comes to IoT, the ESP32 microcontroller is a more suitable option. Therefore, ESP32 module has been used as the heart of our system, performing communication with UHF reader and controlling the robot at the same time. In addition, due to its Wi-Fi and Bluetooth feature, the ESP32 will be able to send the sensor data read from the RFID tag to a web server where all the data will be saved and ready to be processed. Although in remote areas, the network may be unstable or there may be signal interference, the system is designed to store data locally on the mobile robot, allowing continuous data collection and storage even in areas with poor or no network coverage. When the robot returns to a location

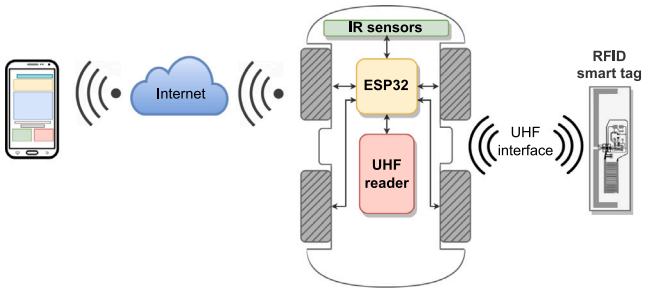


Fig. 2. System design diagram.

with stable connectivity, the stored data can be uploaded to the cloud for further processing and analysis. This flexibility ensures that the system remains functional even in remote or challenging environments. The proposed design is shown in Fig. 2, which consists of the robot with ESP32 microcontroller connected to the motors and IR sensors for trajectory control, and a common UHF reader to read the RFID smart tags. Fig. 3 shows two photos of the robot final assembly from different views. In addition, in Fig. 3(a) it can be seen the UHF reader module and the ESP32 board as the main components of our system. An overview of the IoT platform and connectivity design is shown in Figure S3.

3. Results and discussion

3.1. Antenna performance

The RFID antenna design has been replicated from our previous work (Salmerón et al., 2014c). It consists of a dipole antenna of 5.5 mm width and 79 mm length whose arms have been bended to optimize the occupied area. The antenna impedance must be the complex conjugate of the chip impedance, 31-j286 Ω, to maximize the power transfer between the antenna and the chip (Rao et al., 2005; Nikitin et al., 2009). An SMD inductor of 39 nH has been placed on the antenna feed point to obtain the necessary large imaginary part whilst the width of the dipole arms has been designed to match the real part. The antenna layout on cellulose acetate substrate have no differences with the design presented in Salmerón et al. (2014c), the substrates influence can be negligible in both cases due to the thin films employed.

Fig. 4 illustrates the measurement of the antenna response in two different scenarios: free air condition and when prototype is placed in soil at root depth (about 20 cm from the surface). In the latter scenario, two different prototypes have been buried in soil, as shown in Fig. 4. We have detected a shift of the antenna resonance frequency of 250–290 MHz below the expected value in air and a slightly decrease in the peak value. This behavior was expected due to the different relative permittivity of soil compared with free air conditions (Robinson et al., 1999). Regarding the maximum read range, the RFID tag could be detected at distances smaller than 70 cm in the worst case. This implies a reduction of the read range of 56% compared with the measured range presented in Salmerón et al. (2014c), attributable to the change of the resonance frequency and impedance of the antenna. These two factors increase the loss power between chip and antenna. Moreover, moisture content present in soil will affect the radiofrequency link between the reader and the tag because of the absorption of RF power by water. The correct operation of this prototype has been tested satisfactorily at this read distance for moisture content until 80%. Nevertheless, if logging is enabled, the prototype can store the sensor information in the RFID chip memory at any moisture level and read by the RFID when the humidity level would allow it.

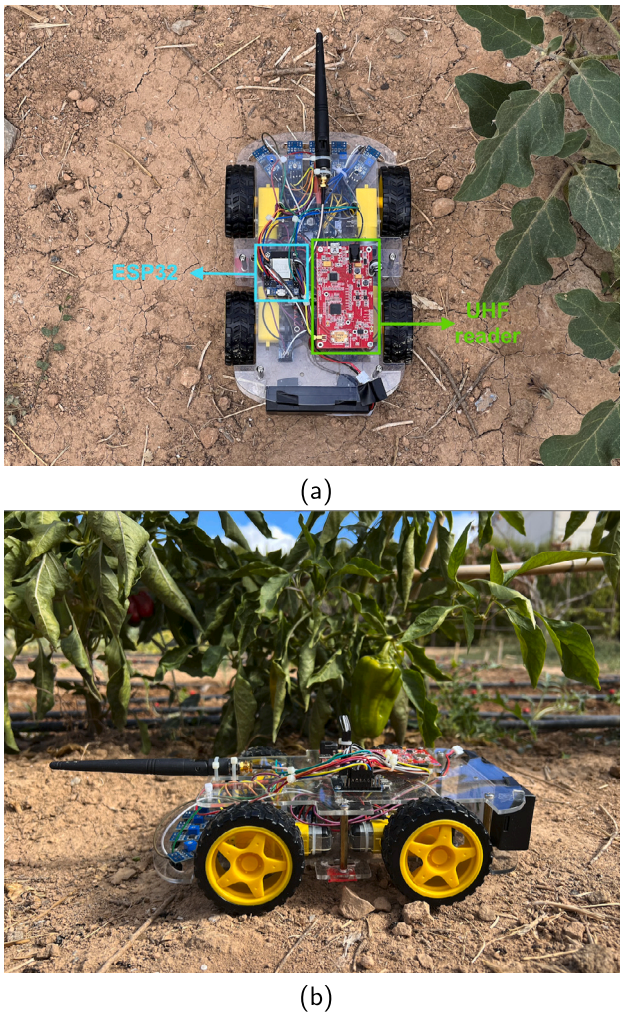


Fig. 3. Robot final assembly. (a) Top view. (b) Side view.

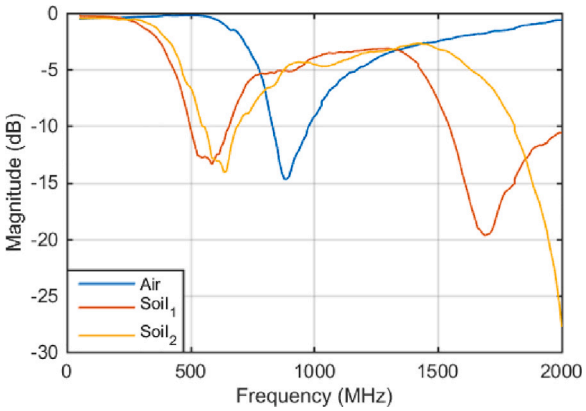


Fig. 4. Magnitude of the dipole antenna measured on free air and while the antenna is placed at root depth.

In particular cases, when the specific characteristics of soil (electrical permittivity and loss factor) and plant (root depth) are known, the antenna could be readily redesigned to optimize its response.

Table 1 summarizes key performance parameters, including resonant frequency and magnitude for the antenna under different conditions both in free air and when buried in soil.

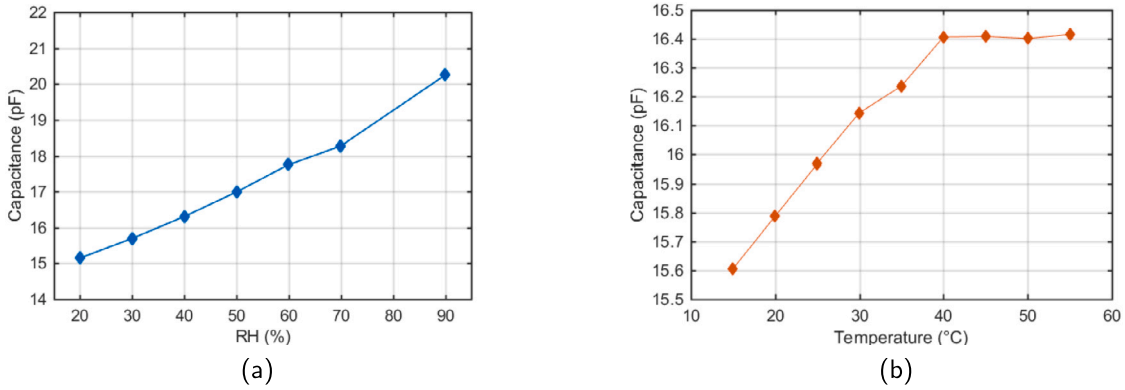


Fig. 5. (a) Capacitance vs. Relative humidity at 45 °C and 100 kHz. (b) Capacitance vs. Temperature at 45%RH and 100 kHz.

Table 1
Main characteristics of the antenna under different conditions.

	Magnitude (dB)	Resonant frequency (MHz)
Air	−14.7	889
Soil 1	−13.4	586
Soil 2	−14.1	635

3.2. Humidity and temperature responses

The responses of the sensor were measured with and without the deposition of a thick ABS layer (water barrier) on top of the IDEs. The protecting layer was deposited by spray deposition through a stencil, a method able to guarantee both a uniform deposition and the selection of the areas to cover. Albeit this test proved the feasibility of the selective deposition of an encapsulating layer to prevent the degradation of the electrodes over time, it was associated to a reduction in the sensitivity of the sensor of 60%. This is likely due to the reduced interaction between water molecules and the sensitive substrate, when the top side is insulated. In scenarios where the sensing tag is intended to be used for a prolonged period of time, selective protection of some areas is a solution to prevent rapid deterioration. However, as the current case did not require extended periods of testing, the results that follow are relative to samples with both sides exposed.

Although capacitive sensors have been measured in a wide range of frequencies, we present in this section the characterization at 100 kHz since the capacitive mode of the SFE of the RFID chip works at this frequency. Fig. 5(a) shows the variation of the capacitance with humidity at constant temperature. Eq. (3) presents the calibration curve after a linear regression, obtaining 73.4 fF/%RH of sensitivity with a correlation coefficient higher than 0.99. Errors were smaller than 2%, calculated as the standard deviation of 3 complete cycles with 2 different sensors. The tests were performed increasing and decreasing RH and temperature, respectively.

$$C(pF) = 0.0676 \cdot RH(\%) + 14.707 \qquad R^2 = 0.9932 \tag{3}$$

It is, however, necessary to take in account the possible occurrence of the thermal drift. For this reason, Fig. 5(b) presents the behavior of the capacitance with temperature at constant RH. The influence of temperature is quite significant up to 40 °C, following a roughly linear trend, finally subject to sudden saturation.

Eq. (4) shows the calibration curve after a linear regression, obtaining 31.6 fF/°C of thermal drift below 40 °C with a correlation coefficient higher than 0.99.

$$C(pF) = 0.0316 \cdot T(^{\circ}C) + 15.155 \tag{4}$$
$$R^2 = 0.9921, \quad T < 40\ ^{\circ}C$$

Due to this substantial interference of temperature in the capacitance, we have studied in detailed its response, sweeping the humidity from 20 °C to 45 °C in different temperature set points, as shown in Figure S2.

All these data have been processes in order to model the sensitivity to humidity considering the temperature. Eq. (6) shows the calibration curve obtained that corresponds to a second order function with a linear correlation coefficient higher than 0.98. The compensation in temperature is needed to obtain an accurate value of RH. In this regard, once the temperature is known, the humidity sensitivity can be calculated and applied to Eq. (4) to process the final value.

$$S_{RH}\left(\frac{fF}{\%RH}\right) = 4 \cdot 10^{-5} T(^{\circ}C)^2 \tag{5}$$
$$- 2.1 \cdot 10^{-3} T(^{\circ}C) + 0.0858$$
$$R^2 = 0.9805$$

$$C(pF) = S_{RH}\left(\frac{fF}{\%RH}\right) \cdot RH(\%) + 14.7127 \tag{6}$$
$$R^2 = 0.9805$$

The thermal compensation can be directly performed by means of the internal temperature sensor available in the RFID chip. The data processing can be carried out on the RFID reader-PC side, once the data stored in the tag are extracted, avoiding their converting in the microcontroller and, in this way, reducing the power consumption of the tag.

One of the main advantages of this strategy compared to the one presented by Kim et al. (2014) is the digital transmission of the data that avoids any uncertainty to the magnitude measured. In the case of an analog determination of the moisture content, there are many other factors that can modify the resonance frequency of the tag (Fernandez-Salmeron et al., 2015).

Due to the limitations of the climatic chamber, the calibration curves has not been estimated for temperatures below 10 °C but the RFID chip is prepared to work below −60 °C (AG, 2014). Therefore, the functioning of the tag is also guaranteed during the cold weather seasons.

3.3. Oxygen response

A full characterization of this sensor on flexible substrate is described in a previous work (Martínez-Olmos et al., 2013). The oxygen value is obtained with a 12-bit digital color detector that is able to measure the intensity of the emitted luminescence of the LED, quenched by the gaseous oxygen. Thanks to the high resolution of the color detector, oxygen concentration can be detected at very low levels ($5 \cdot 10^{-3}\%O_2$) with errors in the order of $10^{-5}\%O_2$. As for the humidity sensor, thermal drift can be easily compensated with the built-in temperature

Table 2
Main characteristics of the sensors.

Sensor	Input/Output range	Sensitivity
RH	20%–80% RH 14.707–19-09 pF	73.4 fF/%RH
Temp	20–40 °C 15.63–16.41 pF	31.6 fF/°C
Oxygen concentration	0.005–0.1% O_2	10^{-5} % O_2

sensor in the RFID chip. As demonstrated by Martínez-Olmos et al. (2013), this sensory system is really suitable for its application in the determination of oxygen content in modified atmosphere with very low oxygen concentration.

Table 2 provides an overview of the main characteristics of the sensors presented in our study, including their input and output linear ranges, as well as the sensitivity of the relative humidity (RH), temperature, and oxygen concentration sensors.

3.4. System verification

To assess the reproducibility of the developed RFID tags, we tested three different tags, each evaluated five times under soil conditions. We calculated the standard deviations of key parameters for the printed components. Specifically, the antenna showed a standard deviation of ±2.2 MHz for the resonant frequency and ±0.8 dB for the magnitude, while the capacitive sensor exhibited a standard deviation of ±0.6 pF. The consistent results across trials confirm the reliability of both the sensor integration and the overall system design.

The practical implementation of the tags will largely be shaped by the specific needs of end users. Variables such as the size of the farm, the type of crops, and the monitoring requirements will play a key role in determining how the system is configured and deployed. For example, in larger farms where crops are more dispersed, careful planning of the robot's path and tag placement will be necessary to ensure comprehensive coverage of the area. Different crops and irrigation methods can also influence soil conditions and the depth of root development. In densely planted areas, roots may be nearer to the surface. Similarly, in fields with specialized irrigation systems like drip irrigation, soil moisture and other parameters may vary significantly based on how and where water and nutrients are applied. As a result, the tag placement will need to be tailored to these specific factors, depending on the user's needs.

In relation to the proof of concept with the robot car, the first prototype of our mobile robot consists of a line-follower based on the use of five infrared sensors to follow the lines, which allows it to move along a certain path. These lines can be black on a white background or white on a black background. In addition, we have used four optical position encoders with opto-interrupters to keep track of the wheel spins. Although this prototype consists of a very simple design, it illustrates the capacities of the system and its promising results. Therefore, providing the robot with more resources and intelligence, it could adapt to autonomous navigation with SLAM techniques, as described in detail in Santos et al. (2020), Aguiar et al. (2020) and Melzer et al. (2016). The sensor data from commercial RFID tags, which follow the EPC Class 1 Gen 2 standard, have been read using the Cottonwood UHF reader from LinkSprite. After conducting tests, we confirmed that the reader, along with its omnidirectional antenna, was capable of reading tags at a distance of 25 cm. The UHF module has been connected to the ESP32 via UART (Universal Asynchronous Receiver Transmitter), allowing the collected data to be automatically sent to the ESP32 and stored in an array. Once all the data from different RFID tags have been collected, the ESP32 can transmit them via Wi-Fi to a web server, accessible from any authorized smart device. Due to the ESP32's dual-core feature, communication with the UHF reader module, wireless transmission, and robot control have been implemented on different CPUs, enhancing system performance. Fig. 6 illustrates an example of how the system works. As mentioned earlier, the robot follows the black line while reading the designed RFID tags.

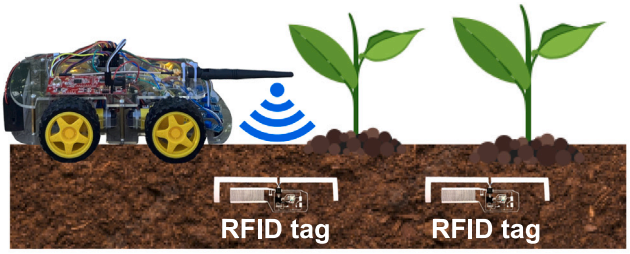


Fig. 6. Example of system operation.

In order to test the system connectivity, we implemented our IoT application using ThingSpeak, an IoT analytics platform offered by Mathworks, which allows us to visualize and analyze live data in the cloud. In addition, it is possible to execute Matlab code in ThingSpeak, which makes easier to perform online analysis and process data as it arrives. For this reason, we have considered ThingSpeak the best method to accomplish the proof of concept of our IoT design.

To make the system work, the first step is to install the ThingSpeak Library for ESP32, thus, the mobile robot will send the acquired data from RFID sensor tags to ThingSpeak Cloud via Internet through the router. At the same time, a channel on ThingSpeak must be created to visualize our data, comprising tree fields: temperature, relative humidity and oxygen concentration. The API Key provided by ThingSpeak is the password that allows the robot to send data to our channel and visualize it on any device through the ThingSpeak website or a specific application developed for this purpose.

Fig. 7 shows the results obtained through the ThingSpeak platform, displaying how the readings from the RFID tag sensors would appear. The first graph (Fig. 7(a)) represents the soil temperature data, which would be obtained from the SL900A chip. The second graph (Fig. 7(b)) depicts the soil relative humidity, as measured by the printed interdigitated electrode capacitor, and the third (Fig. 7(c)) illustrates the oxygen concentration, which would be detected by the O_2 sensing membrane.

In this way, we demonstrate the correct behavior of the entire system, from the design of the RFID tags to the design of the IoT platform based on the mobile robot, verifying that the data flow is adequate.

3.5. Cost-effectiveness analysis

Some of the highlights of this work is that it constitutes a high-resolution and long-life system, offering an IoT solution to monitor root plants state through oxygen concentration, humidity and temperature. However, the primary feature of the proposed system is its low-cost, which makes our system a technology available to all, both for large and small crops and especially helpful for developing countries or countries with drought problems.

The employment of scalable manufacturing techniques, seamlessly integrated in a single tag, opens the door to a variety of applications. The features we selected in this contribution are immediately related to soil monitoring; nevertheless, we are showing a proof of concept, which could be easily expanded to enclose, for instance, ionic and general chemical sensors — as we and other groups showed in other works employing similar manufacturing techniques (Qian and Long, 2018; Fernandez-Salmeron et al., 2015).

The cost-effectiveness of the proposed system is shown in Table 3 where the price of the different components and the complete system is detailed. It should be noted that certain aspects, such as regular maintenance, like battery replacement and cleaning sensor labels, have

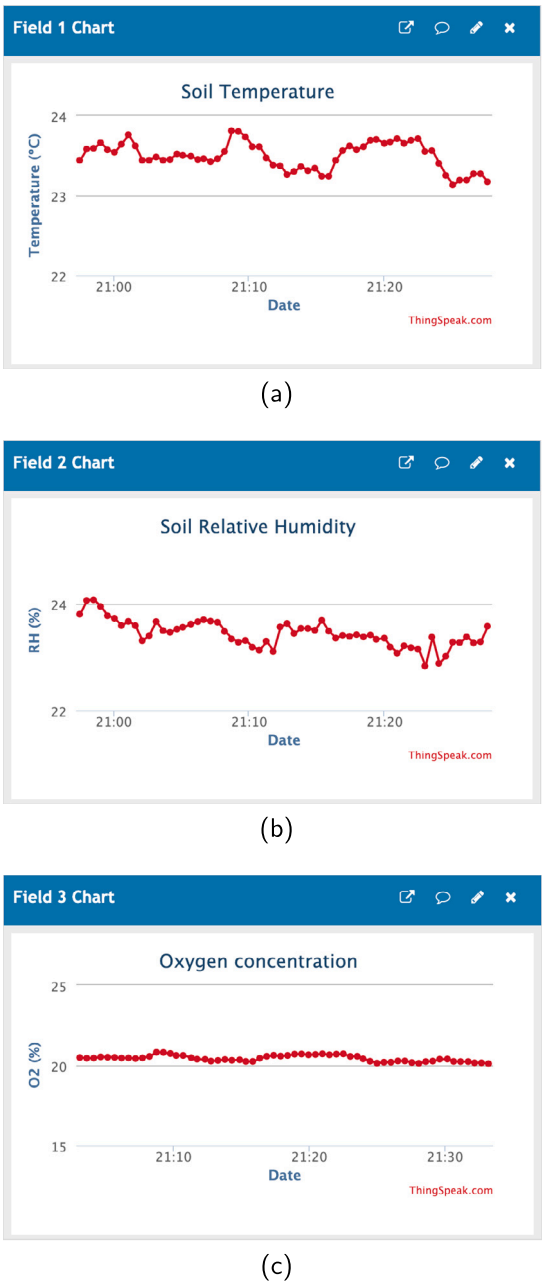


Fig. 7. Results visualization at ThingSpeak platform. (a) Temperature (b) Relative Humidity (c) Oxygen concentration.

not been included in the cost estimation because these tasks are expected to be infrequent and low-cost. The system is supposed to require minimal maintenance. Battery replacement will rarely be necessary, as the batteries are expected to last for extended periods, and the cost of replacing them is low. However, factors such as high temperatures and humidity can affect the performance and longevity of batteries. In environments with high heat or humidity, we anticipate that battery life could be shortened due to increased internal resistance and potential leakage, which may reduce the operational lifespan. Cleaning is also not a major concern, since the performance of the sensor labels is more impacted by their natural degradation than by surface cleanliness. Overall, the system is intended to be low-maintenance, keeping operational costs and complexity low, making it an affordable and practical solution for various agricultural applications.

Table 3
System cost estimation (retail price).

Components	Price (€)
OSOYOO Robot Kit	78.45
ESP32	9.99
UHF RFID reader module	70.74
UHF tag Si components	<40
UHF manufacturing	<10
Total:	<209.18

4. Conclusions

In this work, we developed a complete low-cost and large-scale system for smart farming based on low-impact fabrication techniques, totally prepared for its direct application in the agricultural sector. The system is comprised of the innovative smart printed RFID tags and a robot capable of autonomously reading and exporting to the cloud the information from the sensing tags as a way to confirm the potential of the whole system.

The RFID tags are designed to contain oxygen, humidity and temperature sensors, making the integration of the humidity sensor into the substrate itself along with the oxygen and temperature sensor a design never made before, turning our tags a high value product. The oxygen concentration is determined through the intensity of the luminescent emission generated in a sensitive membrane quenched by O_2, using a LED to excite the membrane and a color digital detector to extract the values of interest and requiring a transparent substrate. The relative humidity is obtained with a novel capacitive interdigitated electrode directly printed on the substrate. The electrical permittivity depends on the moisture content, so that no other sensitive layer is required, achieving a sensitivity of 0.0734 pF/%RH. The last parameter, temperature, is obtained by means of a built-in sensor in the RFID chip. The system is covered with ABS, but a small window is open behind the humidity sensor to allow the moisture sensing but avoiding the degradation of the oxygen membrane (but allowing oxygen to reach it) and enlarging the lifetime of the whole RFID tag. Both oxygen sensor and humidity sensor require temperature compensation that can be performed thanks to the temperature sensor.

The operation of the RFID tag in soil has been tested satisfactorily when buried at distance of 20 cm for a relative humidity below 80%. Thus, the sensor tag has been shown successful in the environment where a final user would employ it, demonstrating that, differently from other printed sensor tags in literature, the result of this work could be used in a real agricultural environment. If data logging is enabled by connecting a battery, the prototype can store the sensor information in the RFID chip memory at any moisture level. Compared to previous works (Kim et al., 2014; Aroca et al., 2018), our system demonstrates a buried reading range of 70 cm. In works where a higher reading range is achieved, the antennas are not buried (da Fonseca et al., 2017; Korošak et al., 2019). However, burying the tags completely underground has advantages such as greater accuracy in terms of the actual value of moisture, oxygen and temperature present in the roots, greater protection against external influences such as UV radiation, and not hindering the movement of machinery and agricultural work.

Finally, the work is perfectly completed with the design of the reading system based on a mobile robot with the ability to navigate through a simplified terrain with markers, find the RFID tags, read them accurately and interface with an IoT application, providing a complete, effective and easy-to-use solution to increase crop efficiency.

CRediT authorship contribution statement

Sonia Gómez-Gijón: Writing – review & editing, Writing – original draft, Software, Investigation, Conceptualization. **José F. Salmerón:** Writing – original draft, Methodology. **Aniello Falco:** Writing – review

& editing. **Florin C. Loghin:** Investigation. **Paolo Lugli:** Investigation. **Diego P. Morales:** Validation, Supervision, Conceptualization. **Noel Rodríguez:** Validation, Supervision. **Almudena Rivadeneyra:** Writing – review & editing, Writing – original draft, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was partially funded by the Andalusian regional projects, Spain: B-RNM-680-UGR20, P20_00265 and P21_00105, by the Spanish Ministry of Sciences and Innovation through the National Projects PID2020-117344RB-I00 and TED2021-129949A-I00 and the Ramón y Cajal fellow, Spain RYC2019-027457-I.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compag.2025.110116.

Data availability

Data will be made available on request.

References

Abdellah, A., Fabel, B., Lugli, P., Scarpa, G., 2010. Spray deposition of organic semiconducting thin-films: Towards the fabrication of arbitrary shaped organic electronic devices. Org. Electron. 11 (6), 1031–1038.

AG, A., 2014. Demo kit for the SL900a smart EPC sensor tag IC. http://www.ams.com/eng/Support/Design-Resources/Demoboards/UHF-RFID/UHF-Interface-and-Sensor-Tags/SL900A-DK-STQFN16.

Aguiar, A.S., dos Santos, F.N., Cunha, J.B., Sobreira, H., Sousa, A.J., 2020. Localization and mapping for robots in agriculture and forestry: A survey. Robotics 9, 97. http://dx.doi.org/10.3390/robotics9040097.

Aroca, R.V., Hernandes, A.C., Magalhães, D.V., Becker, M., Vaz, C.M.P., Calbo, A.G., 2018. Calibration of passive UHF RFID tags using neural networks to measure soil moisture. J. Sensors 2018 (1), 3436503.

Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A., Aggoune, E.-H.M., 2019. Internet-of-things (IoT)-based smart agriculture: Toward making the fields talk. In: IEEE Access. vol. 7, pp. 129551–129583. http://dx.doi.org/10.1109/ACCESS.2019.2932609.

Briand, D., Molina-Lopez, F., Quintero, A.V., Ataman, C., Courbat, J., de Rooij, N.F., 2011. Why going towards plastic and flexible sensors? Procedia Eng. 25, 8–15.

Colella, R., Rivadeneyra, A., Palma, A.J., Tarricone, L., Capitan-Vallvey, L.F., Catarinucci, L., Salmeron, J.F., 2017. Comparison of fabrication techniques for flexible UHF RFID tag antennas [wireless corner]. IEEE Antennas Propag. Mag. 59 (5), 159–168.

Dagar, R., Som, S., Khatri, S.K., 2018. Smart farming–IoT in agriculture. In: 2018 International Conference on Inventive Research in Computing Applications. ICIRCA, IEEE, pp. 1052–1056.

Dholu, M., Ghodinde, K.A., 2018. Internet of things (iot) for precision agriculture application. In: 2018 2nd International Conference on Trends in Electronics and Informatics. ICOEI, IEEE, pp. 339–342.

Falco, A., Petrelli, M., Bezzecchери, E., Abdelhalim, A., Lugli, P., 2016. Towards 3D-printed organic electronics: Planarization and spray-deposition of functional layers onto 3D-printed objects. Org. Electron. 39, 340–347.

Falco, A., Zaidi, A., Lugli, P., Abdellah, A., 2015. Spray deposition of polyethylenimine thin films for the fabrication of fully-sprayed organic photodiodes. Org. Electron. 23, 186–192.

Fernandez-Salmeron, J., Rivadeneyra, A., Carvajal Rodriguez, M.A., Capitan-Vallvey, L.F., Palma, A.J., 2015. HF RFID tag as humidity sensor: Two different approaches. Sensors J. 15 (10), 5726–5733, IEEE.

da Fonseca, N.S.S.M., Freire, R.C.S., Batista, A., Fontgalland, G., Tedjini, S., 2017. A passive capacitive soil moisture and environment temperature UHF RFID based sensor for low cost agricultural applications. In: 2017 SBMO/IEEE MTT-S International Microwave and Optoelectronics Conference. IMOC, pp. 1–4. http://dx.doi.org/10.1109/IMOC.2017.8121099.

Godfray, H.C.J., Beddington, J.R., Crute, I.R., Haddad, L., Lawrence, D., Muir, J.F., Pretty, J., Robinson, S., Thomas, S.M., Toulmin, C., 2010. Food security: the challenge of feeding 9 billion people. Sci. 327 (5967), 812–818.

Kim, S., Le, T., Tentzeris, M.M., Harrabi, A., Collado, A., Georgiadis, A., 2014. An RFID-enabled inkjet-printed soil moisture sensor on paper for smart agricultural applications. In: SENSORS, 2014 IEEE. IEEE, pp. 1507–1510.

Korošak, Ž., Suhadolnik, N., Pleteršek, A., 2019. The implementation of a low power environmental monitoring and soil moisture measurement system based on UHF RFID. Sensors 19, 5527. http://dx.doi.org/10.3390/s19245527.

Lopez-Ruiz, N., Martinez-Olmos, A., de Vargas-Sansalvador, I.M.P., Fernandez-Ramos, M.D., Carvajal, M.A., Capitan-Vallvey, L.F., Palma, A.J., 2012. Determination of O-2 using colour sensing from image processing with mobile devices. Sensors Actuators B- Chem. 171, 938–945.

Mamishev, A.V., Sundara-Rajan, K., Yang, F., Du, Y., Zahn, M., 2004. Interdigital sensors and transducers. Proc. IEEE 92 (5), 808–845.

Martínez-Olmos, A., Fernández-Salmerón, J., Lopez-Ruiz, N., Rivadeneyra Torres, A., Capitan-Vallvey, L.F., Palma, A., 2013. Screen printed flexible radiofrequency identification tag for oxygen monitoring. Anal. Chem. 85 (22), 11098–11105.

Melzer, K., Bhatt, V., Schuster, T., Jaworska, E., Maksymiuk, K., Michalska, A., Scarpa, G., Lugli, P., 2016. Multi ion-sensor arrays: Towards an electronic tongue. In: 2016 IEEE 16th International Conference on Nanotechnology. IEEE-NANO, IEEE, pp. 475–478.

Molina-Lopez, F., Briand, D., de Rooij, N., 2012. All additive inkjet printed humidity sensors on plastic substrate. Sensors Actuators B: Chem. 166, 212–222.

Nikitin, P.V., Rao, K.S., Martinez, R., Lam, S.F., 2009. Sensitivity and impedance measurements of UHF RFID chips, microwave theory and techniques. IEEE Trans. 57 (5), 1297–1302.

2020. OSOYOO V2.1 robot car kit for arduino: Introduction model#2019012400. Osoyoo.com.

Park, J., Hong, W., Kim, C.S., 2010. Color intensity method for hydrogel oxygen sensor array. Ieee Sensors J. 10 (12), 1855–1862.

Qian, R.C., Long, Y.T., 2018. Wearable chemosensors: A review of recent progress. ChemistryOpen 7 (2), 118–130.

Rao, K.V.S., Nikitin, P.V., Lam, S.F., 2005. Antenna design for UHF RFID tags: a review and a practical application. IEEE Trans. Antennas and Propagation 53 (12), 3870–3876.

Rao, R.N., Sridhar, B., 2018. IoT based smart crop-field monitoring and automation irrigation system. In: 2018 2nd International Conference on Inventive Systems and Control. ICISC, IEEE, pp. 478–483.

Ratnaparkhi, Sanika, Khan, Suvaid, Arya, Chandrakala, Khapre, Shailesh, Singh, Prabhishek, Diwakar, Manoj, Shankar, Achyut, 2020. Smart agriculture sensors in IOT: A review. In: Materials Today: Proceedings. (ISSN: 2214-7853) http://dx.doi.org/10.1016/j.matpr.2020.11.138.

Rivadeneyra, A., Fernández-Salmerón, J., Agudo, M., López-Villanueva, J., Capitan-Vallvey, L., Palma, A., 2014a. Design and characterization of a low thermal drift capacitive humidity sensor by inkjet-printing. Sensors Actuators B: Chem. 195, 123–131.

Rivadeneyra, A., Fernández-Salmerón, J., Agudo-Acemel, M., López-Villanueva, J.A., Capitan-Vallvey, L.F., Palma, A.J., 2016. Printed electrodes structures as capacitive humidity sensors: A comparison. Sensors Actuators A: Phys. 244, 56–65.

Rivadeneyra, A., Fernández-Salmerón, J., Agudo-Acemel, M., López-Villanueva, J.A., Palma, A.J., Capitan-Vallvey, L.F., 2015. A printed capacitive-resistive double sensor for toluene and moisture sensing. Sensors Actuators B: Chem. 210, 542–549.

Rivadeneyra, A., Fernández-Salmerón, J., Banqueri, J., Lopez-Villanueva, J.A., Capitan-Vallvey, L.F., Palma, A.J., 2014b. A novel electrode structure compared with interdigitated electrodes as capacitive sensor. Sensors Actuators B: Chem. 204, 552–560.

Roberts, J., Jackson, N., Smith, M., 2006. Tree Roots in the Built Environment. The Stationery Office.

Robinson, D., Gardner, C., Cooper, J., 1999. Measurement of relative permittivity in sandy soils using TDR, capacitance and theta probes: comparison, including the effects of bulk soil electrical conductivity. J. Hydrol. 223 (3), 198–211.

Rumi, S.S., Liyanage, S., Abidi, N., 2024. Soil burial-induced degradation of cellulose films in a moisture-controlled environment. Sci. Rep. 14 (1), 6921.

Sakai, Y., Sadaoka, Y., Matsuguchi, M., 1996. Humidity sensors based on polymer thin films. Sensors Actuators B: Chem. 35 (1), 85–90.

Salmerón, J.F., Gómez-Robledo, L., Carvajal, M.Á., Huertas, R., Moyano, M.J., Gordillo, B., Palma, A.J., Heredia, F.J., Melgosa, M., 2012. Measuring the colour of virgin olive oils in a new colour scale using a low-cost portable electronic device. J. Food Eng. 111 (2), 247–254.

Salmerón, J.F., Molina-Lopez, F., Briand, D., Ruan, J.J., Rivadeneyra, A., Carvajal, M.A., Capitán-Vallvey, L., de Rooij, N.F., Palma, A.J., 2014a. Properties and printability of inkjet and screen-printed silver patterns for RFID antennas. J. Electron. Mater. 43 (2), 604–617.

Salmerón, J.F., Molina-Lopez, F., Briand, D., Ruan, J.J., Rivadeneyra, A., Carvajal ..., M.A., Palma, A.J., 2014b. Properties and printability of inkjet and screen-printed silver patterns for RFID antennas. J. Electron. Mater. 43, 604–617.

Salmerón, J.F., Molina-Lopez, F., Rivadeneyra, A., Quintero, A.V., Capitán-Vallvey, L.F., Rooij, N.F.d., Ozáez, J.B., Briand, D., Palma, A.J., 2014c. Design and development of sensing RFID tags on flexible foil compatible with EPC gen 2. IEEE Sensors 14 (12), 4361–4371.

Salmerón, José F, Rivadeneyra, Almudena, Agudo, Manuel, Capitan-Vallvey, Luis Fermín, Banqueri, Jesús, Carvajal, Miguel A., Palma, A.J., 2014d. Printed single-chip UHF passive radiofrequency identification tags with sensing capability. Sensors Actuators A: Phys..

Santos, L.C., Aguiar, A.S., Santos, F.N., Valente, A., Petry, M., 2020. Occupancy grid and topological maps extraction from satellite images for path planning in agricultural robots. Robotics 9, 77. http://dx.doi.org/10.3390/robotics9040077.

Tibbitts, T.W., 1979. Humidity and plants. BioSci. 29 (6), 358–363.

Urrestarazu, M., Gavilán, M.U., 2004. Tratado de cultivo sin suelo. Mundi-Prensa Libros.

Virtanen, J., Ukkonen, L., Bjorninen, T., Sydanheimo, L., 2010. Printed humidity sensor for UHF RFID systems, sensors applications symposium (SAS). In: 2010 IEEE. IEEE, pp. 269–272, Sensors Applications Symposium.

Wang, X.D., Meier, R.J., Link, M., Wolfbeis, O.S., 2010. Photographing oxygen distribution. Angew. Chem. 49 (29), 4907–4909.

Yadav, N., Hakkarainen, M., 2022. Degradation of cellulose acetate in simulated aqueous environments: One-year study. Macromol. Mater. Eng. 307 (6), 2100951.

Yoon, C., Huh, M., Kang, S.G., Park, J., Lee, C., 2018. Implement smart farm with IoT technology. In: 2018 20th International Conference on Advanced Communication Technology. ICACT, IEEE, pp. 749–752.

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

Smart farming for a sustainable future: Implementing IoT-based systems in precision agriculture / Kumar, S. N., Suriyan, K., Jacob, A. T., Varghese, A., & Francis, E.

Bulletin of the National Research Centre
Volume 49 (2025) 71 Pages 1-13
https://doi.org/10.1186/s42269-025-01366-8
(Database: Springer Nature)

Kumar *et al.*
Bulletin of the National Research Centre (2025) 49:71
https://doi.org/10.1186/s42269-025-01366-8

Bulletin of the National Research Centre

RESEARCH

Open Access

Smart farming for a sustainable future: implementing IoT-based systems in precision agriculture

S. N. Kumar[1*], Kannadhasan Suriyan[2], Andrew Thomas Jacob[1], Amal Varghese[1] and Emma Francis[1]

Abstract

This research work comprises an IoT-based multi-sensor system that is capable of collecting a wide variety of parameters from soil as well as weather and displays these on a mobile application, which can be easily accessible to farmers for their daily cultivation needs. Real-time monitoring of various soil parameters with in-field implementation of the multi-sensor system helps in quick and accurate estimation and relaying of this collected data to the farmers for soil quality analysis, and corrections or adjustments can be made thereafter. It will be a huge benefit to the agricultural sector as it will bring about a new horizon to precision farming, as against the old agricultural practices where trained personnel and separate laboratory analysis are required. This research work presents a new IoT-based precision agriculture system with multi-sensor technology to quantify important soil parameters: pH, moisture, temperature, and NPK level. A case study on Kottayam rubber plantations proved the efficiency of the system in detecting site-specific nutrient deficiencies and enhancing crop suitability estimation. This paper fills the existing gap for integrating low-cost IoT-based monitoring with actionable feedback and mobile access. The system has future extensions in terms of integrating with weather forecasts, pest prediction modules, and machine learning-based yield prediction. A mobile app, 'Dhristi,' was developed for data-visualizing and presenting farmers with tailored cultivation recommendations. The system achieved high accuracy with a mean error margin of below 2% and offers a green solution for traditional soil testing, enhancing yield predictability and resource utilization.

Keywords IoT—Internet of things, NPK—Nitrogen, Phosphorus, Potassium, pH—Potential of hydrogen, ESP-01—Wi-Fi module, ADC—Analog-to-digital converter, RE/DE—Receiver/driver enable pins

Introduction

IoT (Internet of Things) technology has the potential to revolutionize agriculture and enable precision farming. Precision farming involves the use of technology to optimize agricultural practices, improve yields, and reduce waste. NPK (nitrogen, phosphorus, and potassium) sensors are a type of soil sensor that can be used for soil characterization in precision farming. These sensors are designed to measure the levels of nitrogen, phosphorus, and potassium in the soil, which are three essential nutrients that are required for plant growth and development. For monitoring the hydrometeorological data, a low-cost Internet of Things-based system was proposed in Celicourt and Piasecki 2014. With the use of the CC3200 LaunchPad, technologies such as the Internet of Things (IoT), cloud computing, and mobile computing, a cost-effective solution was conceived and constructed. Experimental data and commercially available data are compared in this study. There is a maximum 2% margin of error for the results. The findings for two different soil

*Correspondence:
S. N. Kumar
appu123kumar@gmail.com
[1] Department of EEE, Amal Jyothi College of Engineering, Koovappally, Kerala 686518, India
[2] Study World College of Engineering, Coimbatore, Tamil Nadu 641105, India

types are obtained, uploaded to the cloud, accessed via various devices in the remote area, and then presented. The main applications of cloud technology are cost savings and efficient resource use. With the help of this technology, the individual may view the data from any location and respond to it appropriately (Vani and Rao 2016).The ratio of heat absorbed to heat dissipated in the soil is known as the soil temperature. The chemical, physical, and biological characteristics of the plant are influenced by soil temperature (Onwuka and Mang 2018). Using an analog pH sensor, a prototype model for soil analysis provides real-time data for the current soil sample and indicates whether or not fertilization is currently required. The model makes use of an analog pH sensor to determine the soil's current pH values and whether fertilizers are necessary for the crop being cultivated right now (Oberoi et al. 2017). The soil's acidic character is indicated by a pH value between 4 and 7, which is correlated with the macronutrients potassium, phosphorus, and nitrogen. (K). A comprehensive and economical method called environmental monitoring system (EMS) is designed to detect and monitor environmental factors such as precipitation, atmospheric pressure, temperature, humidity, wind speed, wind direction, solar energy, and water level. Due to its distinguishing trait of being freely accessible in close to real time, the importance of these novel systems and sensors (cheap devices) will increase over time (Sudantha et al. 2018). A prototype model was created for testing and obtaining pH and soil moisture content findings in real time. A MATLAB program provides results of certain crops that can be produced in the current location and soil condition. Selected crop information is stored in a database (Balakrishna et al. 2018). The use of IoT for smart farming to communicate with farmers via a variety of platforms is suggested in this study. In order to improve agricultural output and conserve resources, the device is designed to provide real-time field data (temperature, humidity, moisture, UV index, and IR) (Doshi et al. 2019). For hydrologic monitoring, tracking, and transition, a novel class of inexpensive sensors has been created. An open-source, low-cost platform for measuring, recording, and wireless data transfer for hydrological tracking was proposed in Segovia-Cardozo et al. (2021). A two-legged lead (probe) is put into the area where the soil's water content needs to be assessed in order to gauge its moisture level. The moisture level is then determined by reading the resistance from the lead (Kukal et al. 2019). The development of an agronomic data logger device that uses open-source hardware platforms to capture information on soil moisture and water table levels is covered in this work (López et al. 2022). Finally, it demonstrates how moisture sensor predictions are enhanced over conventional approaches by nonlinear machine learning techniques. A better automatic irrigation system is able to keep an eye on and manage activity on the application field (Winston and Osikibo 2022). A detailed review on applications of IoT on precision agriculture is highlighted in Abu et al. (2022); Kagan et al. (2022). Leveraging Internet of Things (IoT) and machine learning (ML) technologies can revolutionize farming practices in smart cities. These technologies support crucial agricultural tasks like plant irrigation and monitoring of rainfall and drought conditions, assisting farmers in making informed decisions regarding their agricultural activities. The study delves into the multifaceted role and opportunities presented by smart cities, urban farming, communication technologies, as well as the integration of IoT and ML in agriculture (Singh et al. 2022). In order to determine applications, methods, and difficulties, the authors examined the literature on machine learning (ML), data analytics, and the Internet of Things (IoT) in agriculture (Rodrigues 2022). Likewise, they talked about the several kinds of sensors—soil, plant, and environmental—that are utilized in precision agriculture. Also they examined various data analytics methods, including feature selection, statistical analysis, machine learning algorithms, and data preprocessing. For many agricultural applications, including supervised learning, disease detection, pest identification, crop clustering, anomaly detection, and reinforcement learning, they described appropriate machine learning techniques (Rodrigues 2022).

Three steps make up the proposed Internet of Things (IoT) model for precision agriculture: data collection, data analysis, and crop prediction (Krishnakumar and Potty 1989a). Numerous characteristics, including temperature, humidity, soil moisture, pH, nitrogen, phosphorus, and potassium levels, and rainfall, are used to gather data. In order to anticipate crops, the data is preprocessed to eliminate noise and other characteristics. These models are trained using machine learning methods such as random forests, ANNs, and K-nearest neighbors (KNN). After that, the models are used to forecast the best crop based on soil qualities and environmental parameters, taking profitability and market demand into account. Farmers receive recommendations for the best yield via an interface (Krishnakumar and Potty 1989a). A thorough analysis of the body of research on IoT and UAV applications in smart farming has been performed (Krishna Kumar and Potty 1992). The analysis emphasized how IoT and UAVs have made major contributions to smart farming. IoT sensors measure a wide range of characteristics, including crop health, nutrient levels, temperature, humidity, and soil moisture. Along with data analytics methods like machine learning and their applications in jobs like disease detection and

agricultural yield prediction, data networking and communication are covered. A variety of data collection techniques are investigated, and UAV systems are examined. IoT integration is also covered in order to provide a more thorough understanding and decision-making (Krishna Kumar and Potty 1992). An Internet of Things (IoT)-based smart irrigation system that maximizes water use and raises agricultural productivity is presented in Rajendran 1992. A network of sensors, including ones for temperature, humidity, rainfall, and soil moisture, is part of the system. These sensors use wireless technology to gather data and send it to a cloud platform. The irrigation system is managed by the microcontroller unit through the use of machine learning algorithms or preprogrammed regulations. The cloud platform allows for remote monitoring and control by storing sensor data and analyzing it. For dynamic modifications, the system can interface with weather forecasting systems. According to experimental results, crop output can be increased by 10% to 15% while using 50% less water than with conventional approaches. Using a mobile app, farmers may remotely check the temperature, moisture content of the soil, and irrigation status. This methodology is flexible and adjustable to accommodate varying crop varieties and field sizes (Rajendran 1992). An Internet of Things (IoT)-based smart irrigation system that maximizes water use and raises agricultural productivity is presented in Akhter and Sofi (2022). A network of sensors, including ones for temperature, humidity, rainfall, and soil moisture, is part of the system. These sensors use wireless technology to gather data and send it to a cloud platform. The irrigation system is managed by the microcontroller unit through the use of machine learning algorithms or preprogrammed regulations. The cloud platform allows for remote monitoring and control by storing sensor data and analyzing it (Akhter and Sofi 2022). The application of Agriculture 4.0 technology to enhance intelligent farming methods is examined in Bakthavatchalam et al. (2022). It focuses on possible gains in productivity, sustainability, and efficiency. Increased agricultural output, a smaller environmental effect, better farm management, more labor efficiency, and better product quality and traceability are among the expected benefits. The advantages of implementing these technologies are emphasized in the study, including higher agricultural output, less of an impact on the environment, better farm management, and enhanced product quality and traceability. The potential for better labor efficiency and environmental effects is also highlighted in the research (Bakthavatchalam et al. 2022). In previous studies, it has been shown that the existence of a robust IoT-enabled smart irrigation system that utilizes embedded system real-time environment monitoring sensors and server-sent events (SSE) to enhance water management in agricultural practices. By using the EPS32 microcontroller along with a Web-based interface, the system allows for continuous monitoring and efficient control of irrigation based on real-time soil moisture and climatic data, further resulting in the conservation of water that is wasted during irrigation (Morchid et al. 2025, 2024a). Furthermore, the researchers point out the IoT-based fire management system in managing the environmental risk of agricultural fires. The system comprises Message Queuing Telemetry Transport (MQTT), enabling real-time communication between sensors and the monitoring interface (Morchid et al. 2024b, 2024c; Vijayaraghavan 2020). In (Krishna Kumar and Potty 1992), it gives the outlines of the progressive significance of IoT in converting traditional agricultural practice to independent data-driven systems. Moreover, it classifies the IoT architecture into perception, network, and application layers.

The core objective is to offer an economical, real-time monitoring system to evaluate the soil fertility parameters using IoT sensors and provide actionable information to farmers via a mobile app. It seeks to empower precision agriculture by facilitating data-driven decision-making to realize maximum crop yield and resource efficiency. Ongoing monitoring enhances sensitivity to soil deficits or surpluses, decreases fertilizer and water loss, facilitates timely interventions, and enables site-specific soil management, improving productivity and sustainability.

Key Contributions and Research Gaps: Despite increasing demand for smart agriculture, existing systems are often limited to combining multiple soil parameters with real-time actionable information. This paper bridges this gap by introducing a low-cost, multi-sensor IoT system with cloud connectivity and mobile interface, tested in real-world settings. The key contributions are the development of a multi-sensor platform for real-time soil monitoring (NPK, pH, temperature, moisture), designing a farmer-friendly mobile application ('Dhristi') for actionable insights, and a 10-variety field sites case study to confirm system reliability and comparative soil analysis for optimizing rubber plantation in Kerala.

Sect. "Materials and methods" focuses on the architecture of IoT-based system for agriculture, results and discussion are highlighted in Sect. "Results and discussion", and finally, conclusion is drawn in Sect. "Conclusion".

Materials and methods

The flow diagram of the proposed IoT-based precision farming system was depicted in Fig. 1. The sensor values are fed to the cloud storage for further analysis. The description of each sensor used in the research work is as follows.

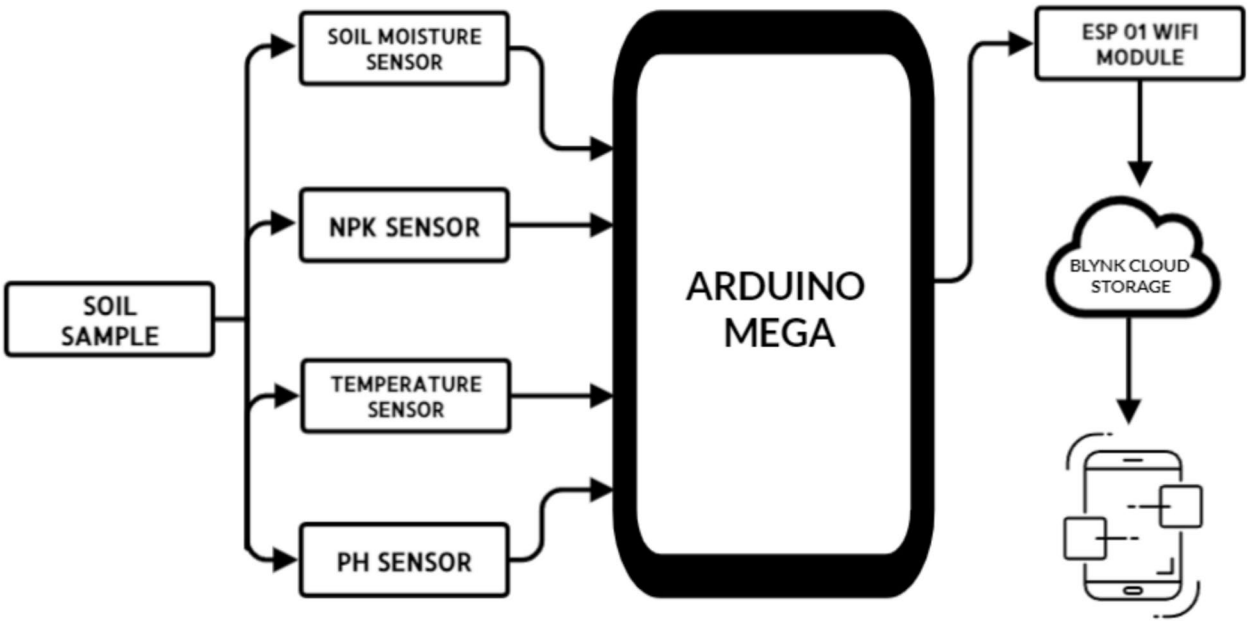


Fig. 1 Block diagram of the IoT-based precision agriculture

The main components that make up the system proposed include sensors such as NPK sensor, pH sensor, moisture sensor, and temperature sensor (DS18B20), microcontroller: Arduino Mega 2560, communication module: MAX485 (RS-485 communication) and ESP-01 (Wi-Fi module), cloud storage: for mobile access and uploading of data, and mobile app (Dhristi): end-user interface.

NPK Sensor: It works by using an electrical conductivity measurement technique to measure the levels of these nutrients in the soil. An A.C. voltage is introduced into the soil, causing a variable current in the soil. Changing current results in varying conductivity, and the concentration conductivity is used to determine the macronutrient concentration. The data obtained from the sensor is then sent to a computer, which interprets the data and gives an accurate measure of the NPK levels in the soil. The data can then be used to determine what nutrients need to be added to the soil in order to optimize plant growth. Using this information, the ratio of fertilizers needed to be added can be adjusted. By measuring the levels of nitrogen, phosphorus, and potassium in the soil, farmers can optimize their fertilizer use and avoid over- or under-fertilization, which can lead to reduced crop yields and environmental damage. It helps farmers identify areas of the field that are deficient in nitrogen, phosphorus, or potassium, allowing them to apply targeted fertilizers and improve crop yields. It can be used to monitor the uptake of nutrients by crops over time, allowing farmers to adjust their fertilizer application rates and timing.

pH Sensor: A pH sensor is a type of soil sensor that can be used for soil characterization in precision farming. pH is a measure of the acidity or alkalinity of soil and is an important factor in determining soil health and plant growth. pH sensors work by measuring the electrical potential difference between two electrodes placed in the soil. The pH of the soil can then be calculated based on this potential difference. pH sensors can be used to measure the pH of soil in real time, allowing farmers to quickly identify areas of the field that may have pH imbalances. pH sensors can be useful for a variety of applications in precision farming. pH sensors can help farmers adjust the pH of the soil to an optimal range for specific crops. For example, some crops may grow better in more acidic soil, while others may require more alkaline soil. It help farmers identify areas of the field that may be deficient in certain nutrients, as nutrient availability is often influenced by soil pH.

Soil moisture sensor: A moisture sensor is a type of soil sensor that can be used for soil characterization in precision farming. Soil moisture is an important factor in plant growth and development, as it determines the availability of water to plants. Moisture sensors work by measuring the water content of soil. There are several types of moisture sensors, including capacitance sensors, tensiometers, and gypsum blocks. Capacitance sensors measure the dielectric constant of the soil, which is related to its water content. Tensiometers measure the tension of water in the soil, while gypsum blocks measure the electrical conductivity of the soil, which is related to its water content. Moisture sensors can be useful for a variety of applications

in precision farming. Moisture sensors can help farmers optimize their irrigation practices by providing real-time information about soil moisture levels. This can help farmers avoid over- or underwatering, which can lead to reduced crop yields and water waste. Moisture sensors can help farmers improve their water management practices by providing information about soil moisture levels across the field. This can help farmers identify areas of the field that may require additional water or have excessive water content.

Temperature sensor: A temperature sensor is a type of soil sensor that can be used for soil characterization in precision farming. The temperature sensor uses analog-to-digital converter to change the analog voltage output into a digital value. This digital value will represent the temperature reading of the soil. This digital value can then be used to control a variety of systems such as irrigation systems or climate control systems. The accuracy of the sensor is improved by calibrating it with a reference temperature source. Soil temperature is an important factor in plant growth and development, as it influences seed germination, root growth, and nutrient uptake. Temperature sensors work by measuring the temperature of the soil. There are several types of temperature sensors, including thermistors, thermocouples, and resistance temperature detectors (RTDs). Thermistors and thermocouples are electronic sensors that measure changes in electrical resistance or voltage with changes in temperature. A comparative analysis of traditional soil analysis and IoT-based system is depicted in Table 1.

Results and discussion

The prototype model uses an Arduino Mega as a microcontroller board with ATmega2560 as the microcontroller. The model uses 4 sensors—JXBS 3001 NPK, JXBS 3001 PH, SEN0114 Moisture sensor, and DS18B20 temperature sensor. The NPK sensor (JXBS 3001 NPK) consists of 2 power cables and 2 communication cables—positive power supply (brown), negative power supply (black), RS485A (yellow/gray), and RS485B (blue). Sensor data is not directly read by the Arduino, it is communicated with using an MAX485 Serial Communicator. The MAX485 consists of 8 pins: GND, A, B, 5 V, receiver output pin (RO), driver output enable pin (DE), receiver output enable pin (RE), and driver input pin (DI). The positive power supply is connected to a 12 V DC supply, and the negative power supply is connected to the GND pin of the MAX485 which is commonly connected to the GND of the Arduino. RS485A is connected to pin A and RS485B is connected to pin B of the MAX485. The RE pin is connected to the digital pin 9 of the Arduino and DE pin is connected to digital pin 10. The RO pin is connected to RX3 of Arduino and DI is connected to TX3 of Arduino. The VCC pin of the MAX485 is connected to 5 V of Arduino. The soil pH sensor (JXBS 3001 PH) includes 2 power cables and 2 communication cables—positive power supply (brown), negative power supply (black), RS485A (yellow/gray), and RS485B (blue). It also uses MAX485 similar to NPK sensor. The positive power supply is connected to a 12 V DC supply, and the negative power supply is connected to the GND pin of the MAX485 which is commonly connected to the GND of the Arduino. RS485A is connected to pin A and RS485B is connected to pin B of the MAX485. The RE pin is connected to the digital pin 8 of the Arduino and DE pin is connected to digital pin 7. The RO pin is connected to RX2 of Arduino and DI is connected to TX2 of Arduino. The VCC pin of the MAX485 is connected to 5 V of Arduino. The SEN0114 moisture sensor contains two components: One is the sensing probe and the other is the sensor module having an LM393 comparator. It consists of 4 pins: VCC, GND, analog output pin, and digital output pin. Connect the analog pin of the sensor to the A0 analog pin of the Arduino. Connect the VCC pin to the 5 V supply and GND pin to GND pin of the Arduino which is shown in Fig. 2.

The final sensor used is the DS18B20 temperature sensor probe. It is a 1-wire communication sensor having 3 wires and uses a pull-up resistor. The 3 wires are power pin (red), data pin (yellow), and ground pin (black). The data pin is connected to digital pin 6, power pin is connected to 5 V pin and ground pin is connected to GND pin. A pull-up resistor of 220 Ω is connected between the data pin and supply pin. An ESP-01 pin is used for Wi-Fi connection for uploading the data into the cloud and it has 8 pins: VCC, GND, reset pin, data send pin (TXD), data receiver pin (RXD), chip enable pin, GPIO0 (internal

Table 1 Conventional soil analysis technique with the suggested IoT-based system

Aspect	Traditional methods	IoT-based system
Accuracy	High (laboratory-based)	Moderate to high (validated with < 2% error)
Time efficiency	Slow (sample collection and laboratory analysis)	Real-time results
Cost	High per test	Low per unit, scalable
Accessibility	Limited (laboratory-dependent)	Farmer-friendly app access
Customization	Low (generalized recommendations)	High (field-specific suggestions)
Limitation	Not dynamic; infrequent measurements	May face connectivity/power issues in rural areas

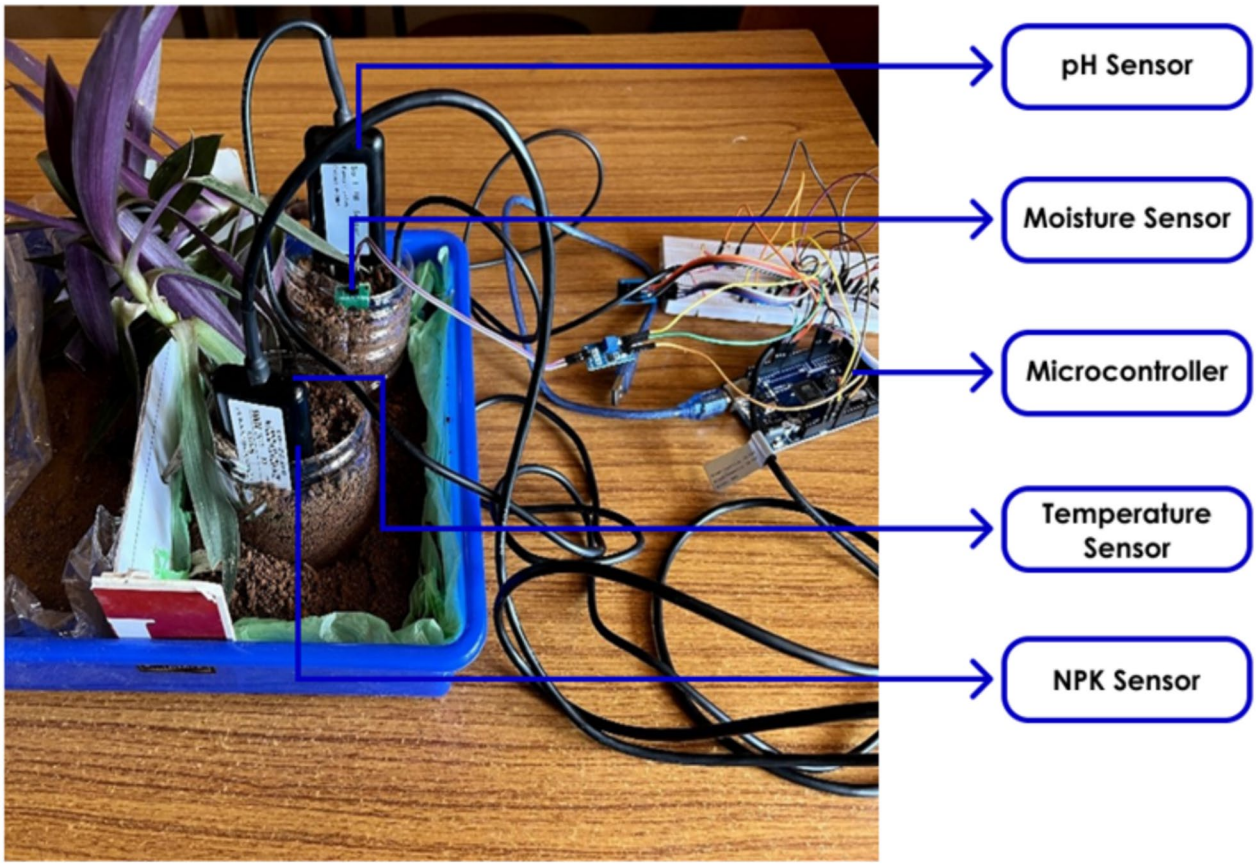


Fig. 2 Initial prototype model

pull-up pin), and GPIO2. The RXD and TXD of the module are connected to RX1 and TX1 of the Arduino, the GND pin is connected to the ground, and the enable and VCC pin is connected to the 3.3 V of the Arduino. The specification in the system is shown in Fig. 3. The program is then uploaded into the microcontroller and the data is displayed simultaneously.

Challenges for farm producers using IoT technology in rural areas: limited Internet connectivity, lack of technical literacy, sensor calibration issues, initial cost barriers, and power supply disruptions affecting data transmission.

Agriculturalists apply the information acquired through the sensor system to enhance their crop yields by fertilizer amount and type according to NPK readings, irrigation timing control with moisture levels, adjusting pH imbalances to promote maximum nutrient uptake, and choosing crops suitable for the prevailing soil conditions. Soil samples are collected from different locations in Kerala and are termed S1 to S10. The characteristics of the soil are estimated by the sensors and tabulated in Tables 2 and 3. In this research work, soil samples from 10 locations are considered. The final prototype is depicted in Fig. 4.

Rubber cultivation plays a vital role in the agricultural landscape of Kottayam, Kerala. However, understanding the relationship between soil fertility parameters and rubber crop performance is crucial for sustainable and optimized production. This case study aims to investigate the variations in soil fertility parameters across different areas and their influence on rubber crop growth and yield. In this extensive case study, ten different samples of soils were collected from different areas of a rubber plantation in Kottayam, Kerala, and these were the results obtained. The analysis of soil fertility parameters revealed notable variations across different locations within the area. pH levels exhibited variations from acidic to slightly alkaline, indicating diverse soil conditions. Soil pH from 4.5 to 5.5 is ideal, and it thrives well under an acidic environment in the soil. The optimum pH for rubber is reported to be in the range of 4 to 6.5, and it can tolerate up to a pH of 3.8 at the low (Rodrigues 2022) and 7.0 at the higher side (Krishnakumar and Potty 1989b). The soil NPK values are also in the ratio 4:2:1, which was in alignment with standard reference values of NPK in Kottayam (Krishna Kumar and Potty 1992; Rajendran 1992). Table 4 represents the soil parameters taken from different areas of a farm land with rubber crop in Kottayam district.

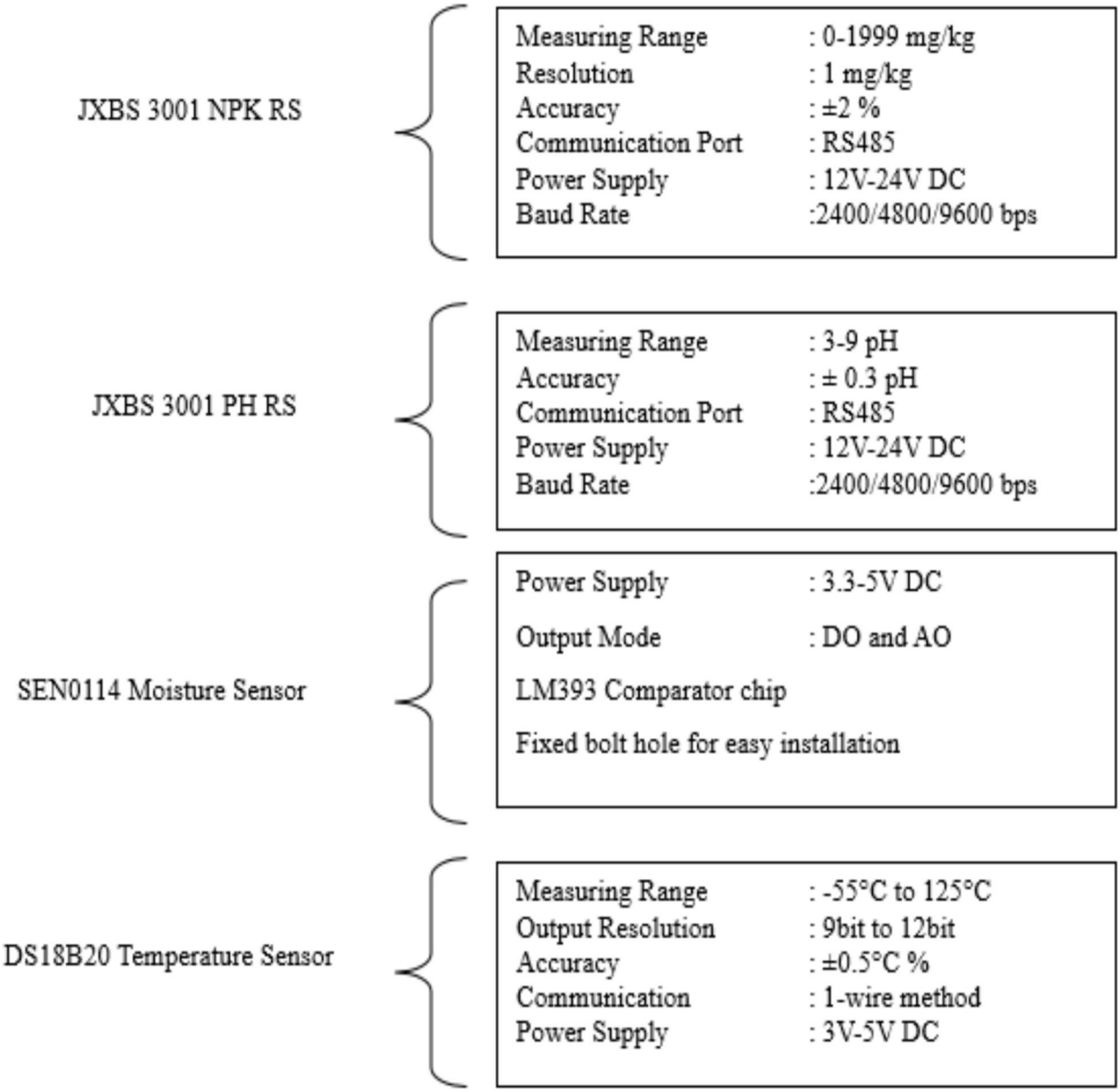


Fig. 3 Specifications of sensors used in the system

Organic matter content varied significantly, suggesting differences in soil health and nutrient availability. Nitrogen, phosphorus, and potassium levels showed distinct patterns, reflecting variations in fertilizer application and agricultural practices. The moisture levels also varied between places. The findings of this study indicate that soil fertility parameters vary significantly across different locations within the same area. These variations highlight the importance of understanding local soil conditions to implement tailored soil management practices for rubber cultivation. Adjusting fertilizer application rates, liming, and soil amendment strategies can help address nutrient deficiencies and optimize soil health. A thorough analysis allows us to provide a complete picture of the fertile parts of a specific area and to make the necessary preparations when sowing a crop, as shown in Table 3.

Some examples of IoT that can be used for precision farming are as follows:

Soil monitoring: IoT sensors can be used to monitor soil moisture, pH levels, temperature, and nutrient content. This data can be used to optimize irrigation and fertilization, leading to improved crop yields and reduced water and fertilizer waste.

Table 2 Soil characterization measurement (mean values of 20 observations from each sample)

Soil samples	Crop details	Temperature	pH	Moisture (%)
S1	Tapioca	32.00	6.40	66
S2	Banana	32.25	5.60	52
S3	Peas	32.19	4.25	27
S4	Tapioca	31.69	6.45	58
S5	Drumstick	30.81	5.65	52
S6	Coconut	30.75	5.15	56
S7	Banana	30.50	6.20	43
S8	Rubber	32.56	6.70	40
S9	Yam	32.38	5.50	22
S10	Rubber	32.69	6.70	37

Table 3 NPK values measurement by sensors (mean values of 20 observations from each sample)

Soil samples	N	P	K
S1	225	109	61
S2	170	82	46
S3	129	63	35
S4	228	111	62
S5	148	72	40
S6	170	82	46
S7	159	77	43
S8	136	66	40
S9	118	57	32
S10	125	61	34

Weather monitoring: IoT weather stations can be used to monitor local weather conditions and forecast changes in the weather. This data can be used to adjust irrigation and fertilization schedules, protect crops from frost, and reduce the risk of weather-related crop damage.

Crop monitoring: IoT sensors can be used to monitor crop growth, including plant height, leaf area, and fruit size. This data can be used to optimize irrigation, fertilization, and pest control, leading to improved crop yields and reduced pesticide use.

Equipment monitoring: IoT sensors can be used to monitor the condition of farming equipment, including tractors, combines, and other machinery. This data can be used to optimize maintenance schedules and reduce downtime.

Livestock monitoring: IoT sensors can be used to monitor the health and behavior of livestock, including cows, pigs, and chickens. This data can be used to optimize feeding schedules, detect illness early, and improve overall animal welfare.

Overall, IoT technology has the potential to transform agriculture and enable farmers to optimize their operations, reduce waste, and improve yields.

The pseudocode of the system is as follows:

```
BEGIN

// 1. Initialize Sensors and Modules
Initialize NPK_Sensor          // Reads Nitrogen (N), Phosphorus (P), Potassium (K)
Initialize Soil_pH_Sensor      // Reads soil acidity or alkalinity
Initialize Moisture_Sensor     // Reads soil water content
Initialize Temperature_Sensor  // Reads ambient or soil temperature
Initialize LCD_Display or IoT_Module (e.g., ESP8266, LoRa, Bluetooth)
Initialize Relay_Module (optional for irrigation)


// 2. Setup Communication
Begin Serial Communication (9600 bps)
Begin I2C Communication for LCD (if used)
Connect to WiFi/LoRa network (if using IoT dashboard)


// 3. Loop for Continuous Monitoring
LOOP forever

  // 3.1 Read NPK Sensor Values
  Read N_value from NPK_Sensor
  Read P_value from NPK_Sensor
  Read K_value from NPK_Sensor

  // 3.2 Read Environmental Sensor Values
  Read soil_pH from Soil_pH_Sensor
  Read moisture_level from Moisture_Sensor
  Read temperature_value from Temperature_Sensor

  // 3.3 Process or Calibrate Readings
  Calibrate pH using offset (if required)
  Convert moisture_level to % using sensor range
  Ensure NPK values are within sensor-specified limits

  // 3.4 Display Data Locally (Optional)
  Display "N: ", N_value on LCD
  Display "P: ", P_value
  Display "K: ", K_value
  Display "pH: ", soil_pH
  Display "Moisture: ", moisture_level%
  Display "Temp: ", temperature_value°C

  // 3.5 Transmit Data to IoT Platform (Optional)
  Prepare data packet (JSON or key-value)
  Send data to cloud (e.g., via MQTT or HTTP POST)

  // 3.6 Provide Decision Support Feedback
  IF N_value < N_threshold THEN
     Show "Add Nitrogen Fertilizer"
  ENDIF

  IF P_value < P_threshold THEN
     Show "Add Phosphorus Fertilizer"
  ENDIF

  IF K_value < K_threshold THEN
     Show "Add Potassium Fertilizer"
  ENDIF

  IF soil_pH < 6.0 THEN
     Show "Soil too acidic – Add lime"
  ELSE IF soil_pH > 7.5 THEN
     Show "Soil too alkaline – Add sulfur"
  ENDIF

  IF moisture_level < moisture_threshold THEN
     Show "Low moisture – Irrigate soil"
     Turn ON Relay for Irrigation
  ELSE
     Turn OFF Relay
  ENDIF

END LOOP

END
```

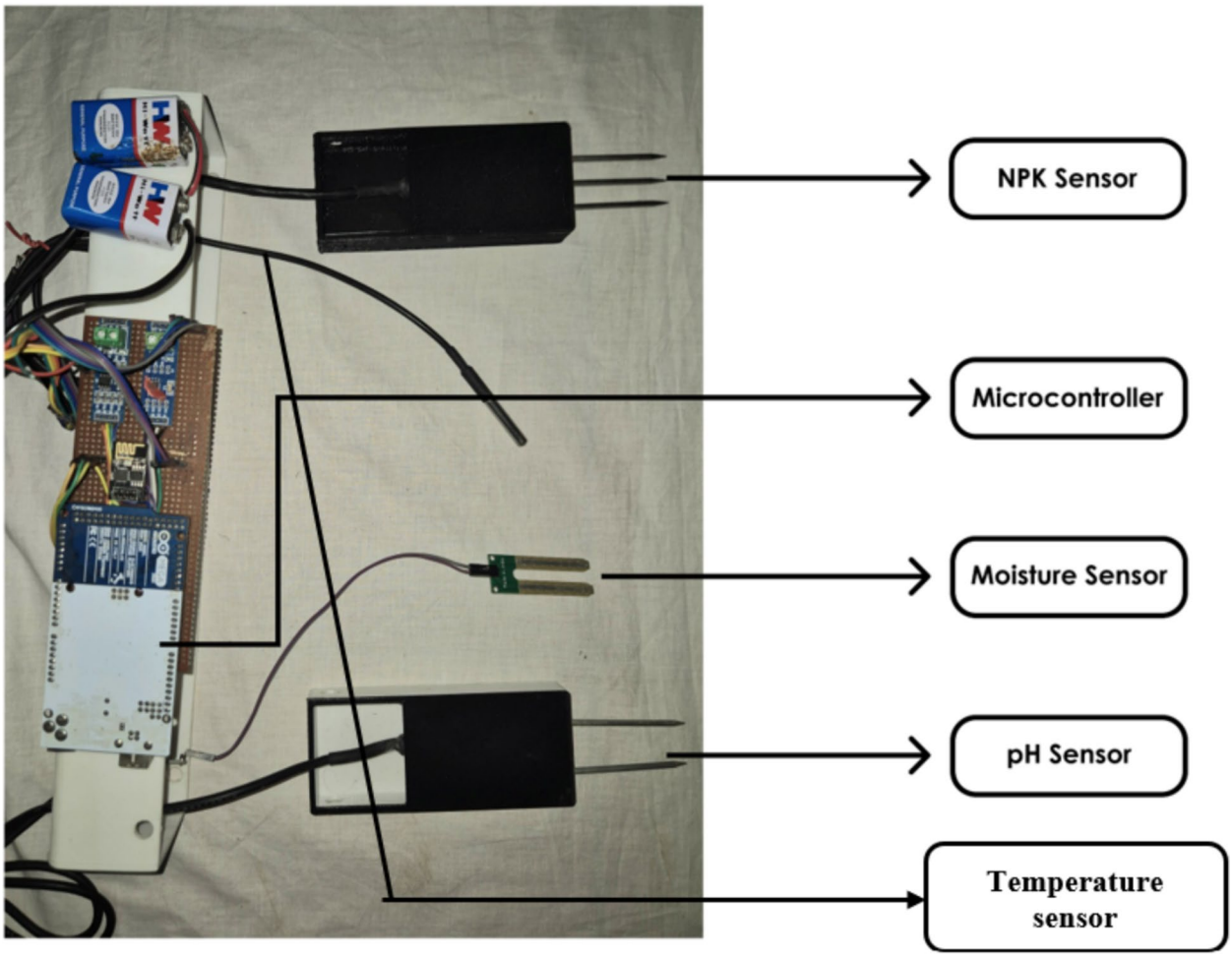


Fig. 4 Final prototype of the proposed system

Table 4 Soil parameters taken from different areas of a farm land with rubber crop in Kottayam district, Kerala

Soil samples	Temperature	N	P	K	pH	Moisture
S1	28.22	124.5	73.4	35	6.56	27%
S2	26.98	124.2	75.8	34	5.36	27%
S3	26.19	120	71	34	6.5	13%
S4	28.68	128.2	75.2	36	6.77	25%
S5	29.07	117.4	71.6	34	6.23	25%
S6	28.74	120.8	71.6	34	6.47	27%
S7	27.5	131	80	34	5.68	28%
S8	28.69	117.1	69.8	33	7.12	30%
S9	27.75	128	77.7	37	7.04	22%
S10	29.16	117.4	71.6	34	5.94	25%

Dhristi is a mobile application developed as an outcome of the research work designed to revolutionize the conventional soil fertility monitoring methods by providing real-time data analysis using integrated sensors. This innovative app offers farmers, agronomists, and land managers an efficient and convenient solution for monitoring and managing soil fertility parameters on their mobile devices. With Dhristi as a digital companion,

farmers can make informed decisions and achieve greater success in nurturing healthy soils and thriving crops. The Dhristi app utilizes advanced sensor technology to collect accurate and precise data on key soil fertility indicators. The mobile app can help a farmer determine when and how to water a field in several ways. It displays current moisture levels, provides recommendations for irrigation duration and volume based on the specific needs of the crops, and offers access to historical moisture patterns. Additionally, the app can alert farmers to the risks of overwatering or underwatering, helping them make more informed irrigation decisions. These sensors, integrated into the mobile device or connected wirelessly, measure parameters such as pH level, NPK values, moisture levels, and temperature.

The data is then transmitted to the app in real time, allowing users to access and analyze the information instantly. One of the primary features of the Dhristi app is its intuitive user interface, which displays the collected data in a user-friendly format. Users can easily navigate through different sections to view specific soil fertility parameters and track their variations over time. The app also provides graphical representations and customizable charts to visualize trends, enabling users to make informed decisions about soil management practices. Figure 5 demonstrates the recommendations page of the proposed application. Based on the collected data, Dhristi generates customized recommendations tailored to each farmer's specific crop and field conditions. By following these recommendations, farmers can maximize yields while minimizing the use of chemical inputs. The system is helpful in correcting soil quality issues in the field by determining specific nutrient deficiencies, prescribing the right fertilizers, recommending pH correction measures such as the application of lime, and informing the farmer when there is water stress.

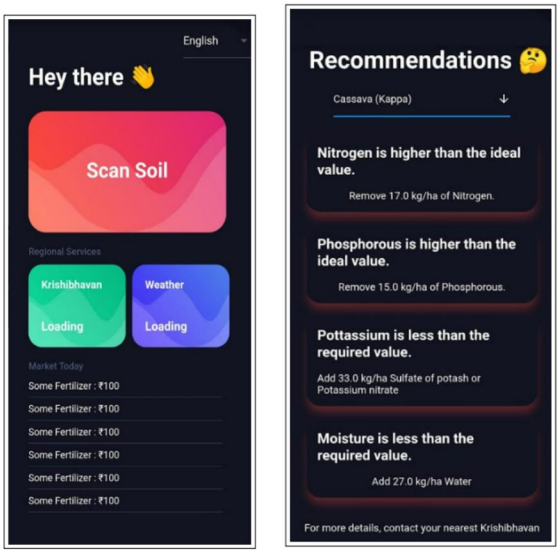
In addition to real-time data monitoring, the Dhristi app offers advanced features to enhance soil fertility management. It includes a comprehensive database of crop-specific soil fertility requirements, allowing users to compare their current soil conditions with optimal levels. The app provides personalized recommendations and suggestions based on the collected data, guiding users in nutrient management, irrigation scheduling, and other soil-related activities. Dhristi stores and analyzes historical data collected from sensors, allowing farmers to track trends and gain deeper insights into their soil's health over time. Moreover, the Dhristi app promotes collaboration and knowledge sharing among users. It allows farmers and experts to connect and share their experiences, best practices, and success stories through a community forum. This platform fosters a supportive environment where users can seek advice, learn from each other, and stay updated with the latest advancements in soil fertility management. Dhristi ensures that farmers have access to accurate and timely information, enabling them to achieve higher productivity and long-term soil health sustainability. Figure 6 depicts the analytics dashboard sensor data. The system reduces the need for specialist staff or laboratory tests with the framework: As it automates soil data acquisition and analysis along with real-time feedback without human sampling, it reduces dependence on outside laboratory services, and requires minimal training due to the easy-to-use interface.

Features added or modified for enhanced usability and overall performance include the integration of solar power for remote or off-grid applications, the incorporation of pest and disease sensors, and the use of predictive analytics through machine learning. Spoken guidance is provided in local languages, and the system coordinates with government agricultural advisory services while also supporting offline data synchronization for areas with poor connectivity. The comparative analysis of the proposed system with the typical existing systems is presented in Table 5.

Validation Strategy: The reading from each of the 10 field locations was averaged based on 20 readings and cross-checked against laboratory standards for soil testing. All the error margins of pH, moisture, and NPK readings were less than 2%, justifying the system's deployment in the field. The cell phone application, Dhristi, enables farmers to see live sensor readings and receive customized recommendations for fertilization and irrigation scheduling. Visualize trends using charts, access crop-specific nutrient guides, and monitor archival information. Collaborate with experts and other farmers via a community forum.

Restrictions: The system has been tested based on a limited number of Kerala soils, and performance might differ across different agro-climatic zones. Real-time detection of pests or diseases, which may be essential for integrative farm management, is not available in the current configuration. This system is particularly useful in regions with unpredictable weather. It supports adaptive irrigation management based on moisture and temperature levels, integration of weather forecast data (as a future upgrade), and the implementation of preventive measures to protect crops from weather-related issues.

Real-time monitoring makes sure that farmers can identify soil deficiencies, moisture stress, or pH imbalance in real time, allowing for on-time interventions in irrigation, fertilization, or liming. Laboratory analysis gives static information that can no longer reflect the field conditions by the time the results are obtained. IoT continuous monitoring supports dynamic, site-specific



(a) (b)

Fig. 5 **a**, **b** Home screen layout of the mobile application

soil management, lowering fertilizer loss, preventing water overuse, and crop stress. This results in increased productivity, lower costs, and less environmental impact.

The system supports sustainability by:

Reducing excessive use of fertilizers and water, thus reducing runoff pollution and saving resources.

Reducing reliance on chemical-intensive blanket approaches, rather favoring site-specific inputs.

Minimizing carbon footprint through fewer field visits, reduced laboratory tests, and voluntary solar-powered operations.

Enabling long-term soil health and farmer resilience.

Therefore, the system is a green option to conventional testing, which supports international Sustainable Development Goals (SDGs) on responsible production and climate action.

The possible future extensions of this work are as follows.

Weather forecast: Correlating soil status with climate forecasts for adaptive irrigation and fertilization.

Pest prediction: Predicting pest infestations using soil and weather information to inform preventive measures.

Yield prediction: Applying machine learning on past soil–crop information for predicting crop yields.

These characteristics will take the system from a monitoring tool to a predictive advisory platform that assists farmers in optimizing input use, reducing risks, and better planning for market demand.

Fig. 6 Analytics dashboard sensor data

Conclusion

Soil fertility is an essential factor that determines the success of farming. It is the soil's ability to provide essential nutrients to the crops, including nitrogen, phosphorus, potassium, and other micronutrients. Soil fertility can be affected by various factors, including soil type, weather conditions, and farming practices. As a result, farmers need to monitor their soil fertility levels regularly to ensure optimal crop growth and yield. Real-time soil fertility analysis provides farmers with an accurate and immediate understanding of their soil's nutrient content. This analysis is done using various techniques, such as

Table 5 Comparative analysis of the proposed system with the some of the typical existing systems

Study reference	Parameters monitored	Real-time data	Mobile/App interface	Hardware platform	Field testing
Khanal et al., (2024)	Moisture, EC, pH, UV, Temp, NPK	Yes	Web App	ESP32S3	Yes (multiple protocols)
Hartono et al., (2025)	NPK	Yes	Mobile App	Custom IoT System	Yes
Deshpande et al., (2022)	Moisture, Temp	Yes	Mobile App	Raspberry Pi	Yes (IITM, Pune)
Proposed Research Work	pH, NPK, Moisture, Temp	Yes	Mobile App (Dhristi)	Arduino Mega + ESP-01	Yes (10 Kerala farms)

soil sampling, laboratory analysis, and on-site testing. With real-time soil fertility analysis, farmers can quickly identify any deficiencies or imbalances in their soil and take corrective actions immediately, such as adjusting the fertilization regime. As an outcome of the research work, a mobile application was developed for the farmers' user interface. The proposed IoT-based precision agriculture system gravely improves the capability for real-time soil examination, thus facilitating timely decision-making on irrigation and fertilization by farmers. Through the employment of inexpensive sensors linked with cloud technology and an easy-to-use mobile application, the system offers a viable path toward sustainable agricultural practice. Future work will involve merging machine learning algorithms for yield forecasting and crop recommendation. Other features like pest/disease diagnosis, weather API integration, and solar-powered operations will be investigated to enhance system robustness and scalability.

Abbreviations

IoT	Internet of things
NPK	Nitrogen, phosphorus, and potassium
pH	Potential of hydrogen
ADC	Analog-to-digital converter
RE/DE	Receiver enable/driver enable
ESP-01	Wi-Fi module based on ESP8266
MAX485	RS-485 transceiver module
DS18B20	Digital temperature sensor (1-wire interface)
Dhristi	Mobile application for farm data visualization
RX/TX	Receive/transmit pins for serial communication
VCC	Voltage at the common collector (power supply)
GND	Ground
RO/DI	Receiver output/driver input (MAX485)
JXBS 3001 NPK	Soil NPK sensor model
JXBS 3001 PH	Soil pH sensor model
SEN0114	Capacitive soil moisture sensor
Arduino Mega 2560	The microcontroller board used in the system

Acknowledgements
The authors would like to the acknowledge the support from AICTE Idea Lab (IDEA LAB 202000080) of Amal Jyothi College of Engineering for supporting this research work.

Author contributions
S.N Kumar, S. Kannadhasan, Andrew Thomas Jacob, and Amal Varghese were responsible for design, conceptualization, implementation, writing, editing, and result. Emma Francis validated the results and carried out comparative analysis.

Funding
The authors declare that they have no funding with other sources.

Data availability
The authors declare that they have no data availability statement.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 21 December 2024 Accepted: 29 September 2025
Published: 8 October 2025

References

Abu NS, Bukhari WM, Ong CH, Kassim AM, Izzuddin TA, Sukhaimie MN, Norasikin MA, Rasid AF (2022) Internet of things applications in precision agriculture: A review. J Robotics Control (JRC) 3(3):338–347
Akhter R, Sofi SA (2022) Precision agriculture using IoT data analytics and machine learning. J King Saud Univ-Comput Inform Sci 34(8):5602–5618
Bakthavatchalam K et al (2022) IoT framework for measurement and precision agriculture: predicting the crop using machine learning algorithms. Technologies 10(1):13
Balakrishna K, Rao M, Anupama KP, Chaitra B, Pooja L (2018) Automatic testing of soil moisture, pH using arduino and selection of specific crop. Int J Comput Sci Eng 6(6):837–841
Celicourt P, Piasecki M (2014) Hydrometeorological data collection, publication and analysis using open-source hardware and software.
Deshpande G, Sharma P, Bhatt V. IoT-Based Low-Cost Soil Moisture and Soil Temperature Monitoring System. arXiv preprint. 2022; arXiv:2206.07488. [Available from: https://arxiv.org/abs/2206.07488]
Doshi J, Patel T, kumar Bharti S (2019) Smart farming using IoT, a solution for optimally monitoring farming conditions. Procedia Comput Sci 160(1):746–751
Hartono R, Pratama DR, Sari M, Nugroho W (2025) Portable internet of things-based soil nutrients monitoring for precision and efficient smart farming. Indones J Electr Eng Comput Sci 27(1):55–62
Kagan CR, Arnold DP, Cappelleri DJ, Keske CM, Turner KT (2022) Special report: The internet of things for precision agriculture (iot4ag). Comput Electron Agric 1(196):106742
Khanal K, Ojha G, Chataut S, Ghimire UK (2024) Iot-based real-time soil health monitoring system for precision agriculture. Int J Sci Eng Res 15(7):145–152
Krishna Kumar AK, Potty SN. Nutrition of hevea. In "Natural Rubber: biology, cultivation and technology." (Eds. Sethuraj MR, Mathew NM) (Crop Science 23. Elsevier); 1992
Krishnakumar AK, Potty SN (1989a) A new fertilizer recommendation for NE REgion. Rubber Ed Bull 24:5–8
Krishnakumar AK, Potty SN (1989b) A new fertilizer recommendation for NE region. Rubber Ed Bull 24:5–8

Kukal MS, Irmak S, Sharma K (2019) Development and application of a performance and operational feasibility guide to facilitate adoption of soil moisture sensors. Sustainability 12(1):321
López E, Vionnet C, Ferrer-Cid P, Barcelo-Ordinas JM, Garcia-Vidal J, Contini G, Prodolliet J, Maiztegui J (2022) A Low-Power IoT Device for Measuring Water Table Levels and Soil Moisture to Ease Increased Crop Yields. Sensors 22(18):6840
Morchid A, Jebabra R, Qjidaa H, El Alami R, Jamil MO (2024) Agri-tech innovations for sustainability: A fire detection system based on MQTT broker and IoT to improve environmental risk management. Res Eng 1(24):103683
Morchid A, Oughannou Z, El Alami R, Qjidaa H, Jamil MO, Khalid HM (2024c) Integrated internet of things (IoT) solutions for early fire detection in smart agriculture. Results Eng 24(1):103392
Morchid A, Et-taibi B, Oughannou Z, El Alami R, Qjidaa H, Jamil MO, Boufounas EM, Abid MR (2025) IoT-enabled smart agriculture for improving water management: A smart irrigation control using embedded systems and Server-Sent Events. Scientific African 1(27):e02527
Morchid A, Said Z, Abdelaziz AY, Siano P, Qjidaa H (2025) Fuzzy logic-based IoT system for optimizing irrigation with cloud computing: enhancing water sustainability in smart agriculture. Smart Agric Technol 29:100979
Morchid A, Jebabra R, Alami RE, Charqi M, Boukili B. Smart agriculture for sustainability: the implementation of smart irrigation using real-time embedded system technology. In2024 4th International conference on innovative research in applied science, engineering and technology (IRASET) 2024 May 16 (pp. 1–6). IEEE.
Oberoi A, Basavaraju S, Lekshmi S. Effective implementation of automated fertilization unit using analog pH sensor and Arduino. In 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC) 2017 Dec 14 (pp. 1–5). IEEE.
Onwuka B, Mang B (2018) Effects of soil temperature on some soil properties and plant growth. Adv Plants Agric Res 8(1):34–37
Rajendran PV. Comparative study of fertilizer recommendations based on soil and leaf analysis vis a vis blanket recommendations of rubber board (Doctoral dissertation, Department of Plantation Crops and Spices, College of Horticulture, Vellanikkara).
Rodrigues GC (2022) Precision agriculture: Strategies and technology adoption. Agriculture 12(9):1474
Segovia-Cardozo DA, Rodríguez-Sinobas L, Canales-Ide F, Zubelzu S (2021) Design and field implementation of a low-cost, open-hardware platform for hydrological monitoring. Water 13(21):3099
Singh DK, Sobti R, Jain A, Malik PK, Le DN (2022) Lora based intelligent soil and weather condition monitoring with internet of things for precision agriculture in smart cities. IET Commun 16(5):604–618
Sudantha BH, Warnakulasooriya KM, Jayasuriya YP, Ratnayaka GR, Mahanama PK, Warusavitharana EJ, Weerasinghe SN (2018) Open-source implementation of an integrated low-cost environmental monitoring system (EMS) for developing countries. Bhumi, Planning Res J 6(1):23–29
Vani PD, Rao KR (2016) Measurement and monitoring of soil moisture using cloud IoT and android system. Indian J Sci Technol 9(31):1–8
Vijayaraghavan V (2020) IoT and cloud hinged smart irrigation system for urban and rural farmers employing MQTT protocol. In2020 5th International Conference on Devices, Circuits and Systems (ICDCS) 2020 Mar 5 (pp. 71–75). IEEE.
Winston GI, Osikibo LT (2022) Design and Construction of an IMPROVED Automatic Irrigation System. Am J Eng Res (AJER) 11(05):144–154

Publisher's Note

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

Smart system for automated irrigation using internet of things devices / Ferrarezi, R. S., Peng, T. W., Ferrarezi, R. S., & Peng, T. W.

HortTechnology
Volume 31 Issue 6 (2021) Pages 642–649
https://doi.org/10.21273/HORTTECH04860-21
(Database: American Society for Horticultural Science)

Smart System for Automated Irrigation Using Internet of Things Devices

Rhuanito Soranz Ferrarezi[1,2] and Tzu Wei Peng[1]

Additional index words. embedded microcomputer, low-cost, management, open-source, water

Summary. Precision agriculture involves applying artificial intelligence, computers, sensors, and automation to improve crop field productivity while monitoring environmental conditions to conserve soil, water, and other natural resources focusing on agricultural sustainability. Despite many applications in agriculture, data monitoring and recording technologies have limited use due to the price. Low-cost open-source systems, like the ones available with the Internet of things (IoT) world, can potentially be developed as a universal-fit and cloud-connected technology for multiple applications. We designed and built a basic data collecting system using a commercial standalone embedded computer with Python programming language, serial data interface (SDI)-12/analog sensor adaptor, and digital sensors to monitor soil moisture and transmit the data remotely. SDI-12 is a standard communication protocol that transfers digital sensor measurements to a data recorder. We set up a pilot study that automatically collected and uploaded the data into the Internet to allow remote data transfer and access. The system performed reliably over 1 week with potting soil under field conditions without maintenance and successfully recorded data in real-time. The volumetric water content ranged from 0.03 to 0.23 $m^3 \cdot m^{-3}$, dielectric permittivity from 3.3 to 18.9 (unitless), EC from 0.0 to 0.3 $dS \cdot m^{-1}$, and soil temperature from 20.7 to 44.8 °C. All the data were successfully collected and uploaded to the cloud every 20 min, allowing users to remotely monitor the data using a free online application. However, heavy rainfall and high insolation could damage the system through excessive moisture or overheating, requiring a waterproof and heavy-duty protection case. The ThingSpeak channel allows customizing to suit a user's specific requirements or adding more features for further development, such as automated irrigation, which can improve irrigation and fertilization efficiency by applying water and fertilizers at the right time based on sensor readings.

The United Nations predicts that the world's population will reach 9.7 billion by 2050 (United Nations Department of Economic and Social Affairs, 2015), which will require global food production to rise rapidly to meet consumer demand. Understanding the future environmental challenges of global crop production and how to achieve greater yields with fewer inputs require quantitative assessments of environmental variables such as soil moisture and water availability (Tilman et al., 2011). The efficient use of irrigation water in agricultural production is vital to farming operations' economic and environmental sustainability. Precision irrigation technologies allow farmers to apply water when, where, and in the amount needed to maximize profits while protecting the environment (Payero et al., 2017). According to a study with large-scale commercial corn (*Zea mays*) in Nebraska, irrigation scheduling based on sensors reduced water application rates by 33%. It also reduced pumping cost by $28.5/acre per year, with no significant reduction in yield compared with the farmer-managed fields (Irmak et al., 2012). With the advance of Internet connectivity and the popularization of technology adoption in agriculture, new tools are becoming available to improve the decision process in farm operations.

Internet of things (IoT) describes the network of devices or "things" embedded with sensors, software, and other technologies to smartly connect and exchange data between equipment and systems over the Internet. In smart farming, automated systems monitor horticultural crops with sensors that measure environmental parameters such as solar radiation or light intensity, relative humidity, air temperature, soil moisture, electric conductivity (EC), wind speed, and direction, and control field operations (Ravindra, 2020). IoT-based smart farming allows users to automatically collect real-time information remotely as long as an Internet connection is available. Data are stored in cloud computing, which can assist users in making more intelligent decisions. IoT systems are expected to increase the operation efficiency compared with conventional agriculture due to the potential to lower dependence on frequent labor for decision processes. IoT devices are becoming more popular since consumers connect and integrate more equipment online, making this a dynamic and evolving market suitable for different operations.

With the introduction of industrial IoT in agriculture, more advanced sensors are being used. Sensors are either analog or digital. An analog sensor measures the external parameters and returns analog voltage as an output. The output voltage is usually in the range of 0 to 5 V direct current (DC). Unlike analog sensors, digital sensors produce discrete values (0 and 1); these values are often called digital (binary) signals in digital communication. Analog signals are highly affected by external factors such as noise and signal degradation, creating errors in the output signal and affecting data integrity. Noise is an unwanted electrical or magnetic phenomenon that corrupts a signal from the sensor. Digital signals are less susceptible to noisy environments, and hence digital sensors are usually preferred by users over analog ones. Digital sensors use a standard communication protocol known as serial/digital interface at 1200 baud [serial data interface (SDI)-12] that quickly allows user interfacing data recorders with microprocessor-based sensors. The

Units

To convert U.S. to SI, multiply by	U.S. unit	SI unit	To convert SI to U.S., multiply by
0.4047	acre(s)	ha	2.4711
0.0283	ft^3	m^3	35.3147
2.54	inch(es)	cm	0.3937
1	µmho/cm	$\mu S \cdot cm^{-1}$	1
1	mmho/cm	$dS \cdot m^{-1}$	1
(°F − 32) ÷ 1.8	°F	°C	(°C × 1.8) + 32

"intelligent" sensor typically takes a measurement, makes computations based on the raw sensor reading, and outputs the measured data in meaningful units that are easy to understand. These multiparameter digital sensors help battery-powered controllers with minimal current drain, allowing several sensors to connect to a single datalogger. As a result, they can be wired to the same control port, and less processing power is required from the datalogger as complex algorithms are executed by the sensor. In addition, SDI-12 sensors can measure several soil parameters such as volumetric water content (VWC), dielectric permittivity (ε), temperature, and EC. Although various data logging systems and soil moisture sensors are commercially available, the large-scale application of these technologies to make irrigation scheduling decisions among farmers is still limited.

Factors affecting the limited adoption of these technologies include the lack of knowledge about the system applications, little information in system design and operation, high cost, difficulty installing and maintaining equipment, and problems transmitting data in real-time from sensors in the field to the farmer (Payero et al., 2017). Most of these elements are centered on the microcontroller or datalogger used to process the data, as they are usually proprietary systems. The most common devices cost more than $1500 and are generally used in scientific research settings. The programming is relatively simple but requires proprietary software. One potential solution is using low-cost open-source platforms such as the Arduino (Arduino, Ivrea, Italy) and Arduino-like boards (Ferrarezi et al., 2015). These prototyping boards are prevalent for inexpensive projects as parts are readily available, and the cost is appealing. For example, Ferrarezi et al. (2015) described an automated system for monitoring soil moisture and controlling irrigation using low-cost open-source microcontrollers. Open-source commonly refers to software that uses an open development process and is licensed to include the source code, which means the source code is freely available to the public. Anyone can modify the code to customize the system based on user needs. The program is written using the Arduino integrated development environment (IDE) freeware that employs C/C++ programming language (Ferrarezi et al., 2015). However, the Arduino boards are made for small-scale prototyping projects. The addition of multiple accessory boards and algorithms to precisely control the timestamp, save data into a memory card, and transmit the data to the cloud reduces the system performance for various sensors to handle large, complex systems.

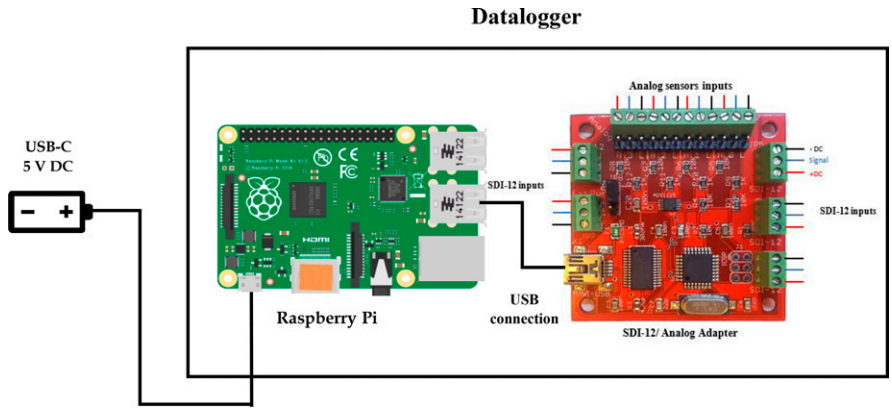


Fig. 1. Essential datalogger components to monitor sensors and control devices: embedded computer with a credit-card size and dual-display desktop [Raspberry Pi 4 Model B with 8 Gb of random-access memory (RAM); Raspberry Pi Foundation, Cambridge, United Kingdom] and serial data interface (SDI)-12/ analog adaptor with larger terminals [SDI-12 universal serial bus (USB) adaptor; Liudr's Blog, St. Cloud, MN] connected via the USB type-C port without the need of stacking or soldering the boards; DC = direct current.

Embedded computers with a credit-card size became popular due to the promotion of teaching introductory computer science. They are used in agriculture to monitor environmental parameters because of their low cost, modularity, and open-source design, and the system is more robust due to the higher processing capacity. There are different brands and models available in the market. To make smart farming more publicly accessible, a reliable, robust, low-cost, open-source IoT device can be developed as a universal-fit and cloud-connected data-collecting system integrated by computer programing.

Our objective is to describe a low-cost open-source IoT intelligent system using digital soil moisture sensors connected to an embedded computer to make moisture monitoring and irrigation control more efficient and affordable in specialty crop applications. This material was presented at the American Society of Horticultural Sciences 2020 Conference Workshop "Using Sensors to Inform and Control Irrigation Research and Management" as part of the Water Utilization and Management Professional Interest Group.

Pilot study: A low-cost open-source IoT system for irrigation control

Logic. The intelligent, automated irrigation platform described in

Received for publication 1 Apr. 2021. Accepted for publication 2 Aug. 2021.

Published online 27 October 2021.

[1]Horticultural Sciences Department, Indian River Research and Education Center, University of Florida, Fort Pierce, FL 34945

[2]Department of Horticulture, University of Georgia, Athens, GA 30602

This paper is based on information presented during the "Using Sensors to Inform and Control Irrigation Research and Management" workshop at the ASHS 2020 Virtual Conference as part of the Water Utilization and Management (WUM) Professional Interest Group (PIG).

Funding for this research was provided by the U.S. Department of Agriculture's (USDA) National Institute of Food and Agriculture (NIFA) Hatch/Multistate W4128 Project #1021596 and USDA-NIFA Emergency Citrus Disease Research and Extension Program (ECDRE) project #2018-70016-27387.

We thank the ASHS WUM PIG for the exciting presentations at the workshop, University of Florida's Institute of Food and Agricultural Sciences (UF/IFAS) Indian River Research and Education Center, and the HortTechnology editorial staff.

Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the authors.

R.S.F. is the corresponding author. E-mail: ferrarezi@uga.edu.

https://doi.org/10.21273/HORTTECH04860-21

Table 1. Information of the sensors used in the development of automated irrigation systems. The soil-specific calibrations for mineral soils ranging from 0 to <5 dS·m^{-1} were obtained from the sensors' user manuals and applied to the code for the datalogger's program code to output specific data accordingly.

Manufacturer and sensor model	Measuring technique[z]	Response variables[z]	Factory calibration[z] VWC	Factory calibration[z] ε
Meter Group, Pullman, WA				
10HS	FDR	Voltage	$2.97 \times 10^{-9} mV^3 - 7.37 \times 10^{-6} mV^2 + 6.69 \times 10^{-3} mV - 1.92$	$2.589 \times 10^{-10} mV^4 - 5.010 \times 10^{-7} mV^3 + 3.523 \times 10^{-4} mV^2 - 9.135 \times 10^{-2} mV + 7.457$
GS3	FDR	ε, EC_b, T	$5.89 \times 10^{-6} \varepsilon_a^3 - 7.62 \times 10^{-4} \varepsilon_a^2 + 3.67 \times 10^{-2} \varepsilon_a - 5.3 \times 10^{-2}$	
Acclima, Meridian, ID				
ACC-SEN-SDI	TDT	ε, EC_b, T	$4.3 \times 10^{-6} \varepsilon_a^3 - 5.5 \times 10^{-4} \varepsilon_a^2 \div 2.92 \times 10^{-2} \varepsilon_a - 5.3 \times 10^{-2}$	

[z]VWC = volumetric water content, ε = dielectric permittivity, FDR = frequency domain reflectometry, EC_b = bulk EC, T = temperature, TDT = time-domain transmission.

this paper integrates open-source software and hardware, Internet cloud components to allow remote data transfer and access, soil moisture sensors, and an irrigation manifold with a solenoid valve. A program was written and uploaded to a programmable single-board computer to perform data collection and transmission functions to an Internet cloud-based service.

Hardware. The low-cost open-source automated irrigation control system consisted of three major components: a datalogger, soil moisture sensors, and an irrigation manifold. Three soil moisture sensors were wired together and transmitted the data to the datalogger. The system stored the information locally in a microsecure digital (SD) card, uploaded it in real-time to the cloud wirelessly, and controlled the irrigation based on soil moisture data. This last portion was not tested in this pilot trial as it had been validated before (Ferrarezi et al., 2015; Nemali and van Iersel, 2006). In addition, data were also available remotely through a website that can be fully customized on the user backend.

Datalogger. We used an embedded computer with a credit-card size and dual-display desktop [Raspberry Pi 4 Model B with 8 GB of random-access memory (RAM); Raspberry Pi Foundation, Cambridge, UK] and an SDI-12/analog adaptor with larger terminals [SDI-12 universal serial bus (USB) adaptor; Liudr's Blog, St. Cloud, MN] to read both SDI-12 and analog sensors (Fig. 1). More info about the board and compatible soil moisture sensors are available at Liu (2021). The system was housed in a 20 × 12 × 5.6-cm waterproof and dustproof enclosure (LeMotech ABS, Shenzhen, China). The desktop microcomputer is an open-source single-board embedded computer used as both a microcontroller and microprocessor. The model used features Gigabit Ethernet, onboard wireless networking, Bluetooth, four USB ports, a micro-SD card slot for data storage, and 40 physical general-purpose input/output (GPIO) pins to interact with different components. The device consumes low power and requires a standard USB Type-C (USB-C) 5-V 3-A DC power supply and has an internal 5-V battery to power the SDI-12/analog adaptor via USB port connection without the need of stacking or soldering the boards.

Sensors. We used three sensors: an analog (10HS; Meter Group, Pullman, WA) and two SDI-12 digital sensors [GS3 (Meter Group) and ACC-SEN-SDI (Acclima, Meridian, ID)]. The 10HS analog sensor was assigned to address z, GS3 digital sensor to address 1, and ACC-SEN-SDI digital

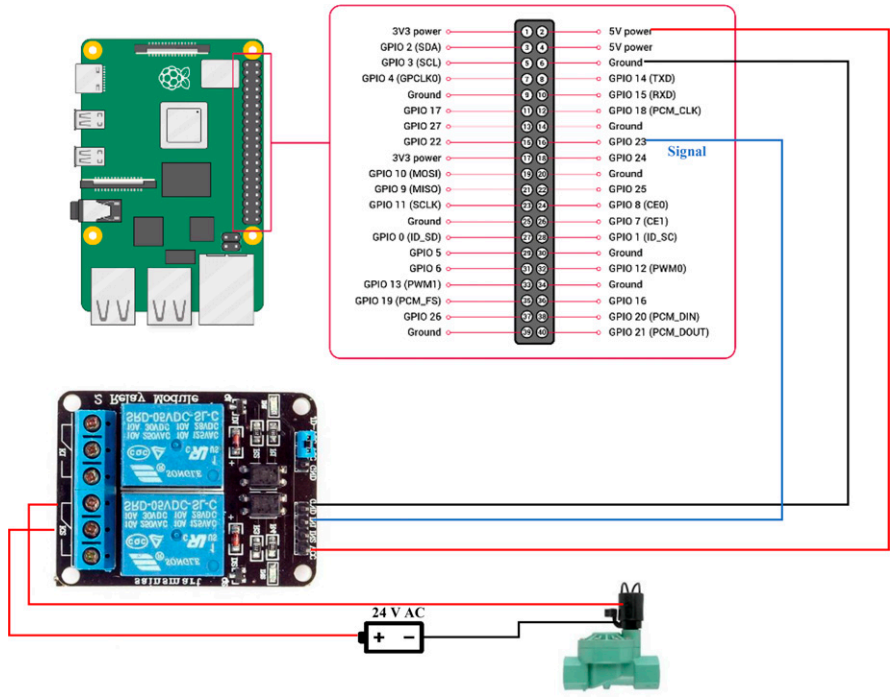


Fig. 2. Basic wiring diagram of automated irrigation controller with embedded computer with a credit-card size and dual-display desktop [Raspberry Pi 4 Model B with 8 Gb of random-access memory (RAM); Raspberry Pi Foundation, Cambridge, United Kingdom], relay module [one channel 5-V direct current (DC); SaintSmart, Las Vegas, NV], and 1-inch (2.5 cm) 24-V alternate current (AC) solenoid valve (Orbit, North Salt Lake, UT) (Raspberry Pi Foundation, 2021); GPIO = general purpose input/output.

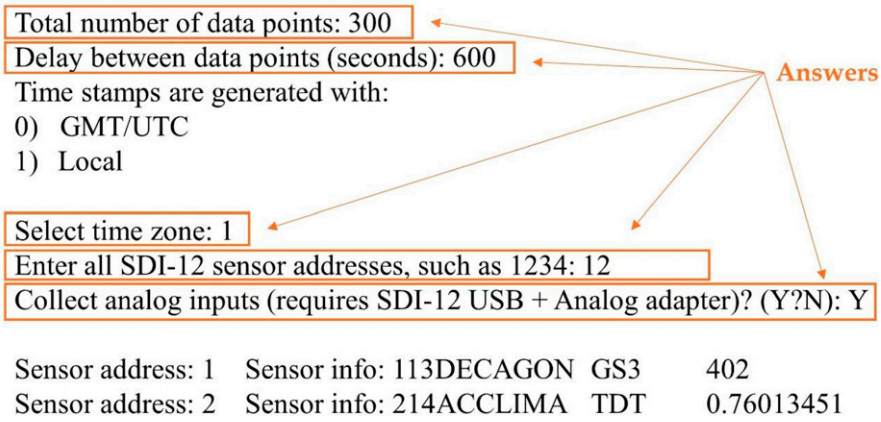


Fig. 3. Example of the computer code used to run the automated system. Items with an orange box are direct outputs of the computer instruction; GMT/UTC = Greenwich mean time and coordinated universal time, SDI = serial data interface, USB = universal serial bus.

sensor to address 2. The 10HS analog sensor uses frequency domain reflectometry (FDR) for the measurements. The principle is based on the electrical capacitance that uses the soil as a dielectric pole to measure the soil VWC. When connecting the capacitor with an oscillator to form an electrical circuit, changes in soil moisture can be detected by changes in the operating frequency (Ferrarezi et al., 2020). The analog sensor determines voltage, which is manually converted into VWC and ε by using the sensor manufacturer's calibration equations. The calibration was added to the datalogger to calculate the values automatically (Table 1).

The GS3 digital sensor also uses the same FDR technology by sending out a 70-MHz frequency pulse to the soil. The difference between the output wave and the return wave frequency is measured to determine the soil moisture. In addition, the GS3 sensor measures ε, bulk EC, and temperature.

The ACC-SEN-SDI digital sensor uses a different operation principle known as time-domain transmission (TDT), which measures the speed from the start to the end of the loop to propagate an electrical pulse. A pulse emitted by the ACC-SEN-SDI sensor is slower in wetter soils than drier soils. This sensor measures ε, EC, and temperature. Both sensors, the GS3 and ACC-SEN-SDI, do not require previous conversions since they output all the data as a string that quickly allows users to obtain the data microprocessed by the sensors. ε, EC, and temperature are generated directly using a generic equation built into the sensors.

Soil moisture sensors were wired to the SDI-12/analog adaptor. The 10HS sensor was powered by the internal 5-V DC power supply, and an external 12-V DC power supply was used to power the ACC-SEN-SDI and GS3 sensors as they require more than 5 V to function. Analog sensors were assigned to address z, and SDI-12 sensors to addresses starting from number 1 on the SDI-12/analog adaptor. The data collected from the sensors were stored in the desktop microcomputer as comma-separated values (CSV) file for local backup and also uploaded to the cloud for real-time monitoring.

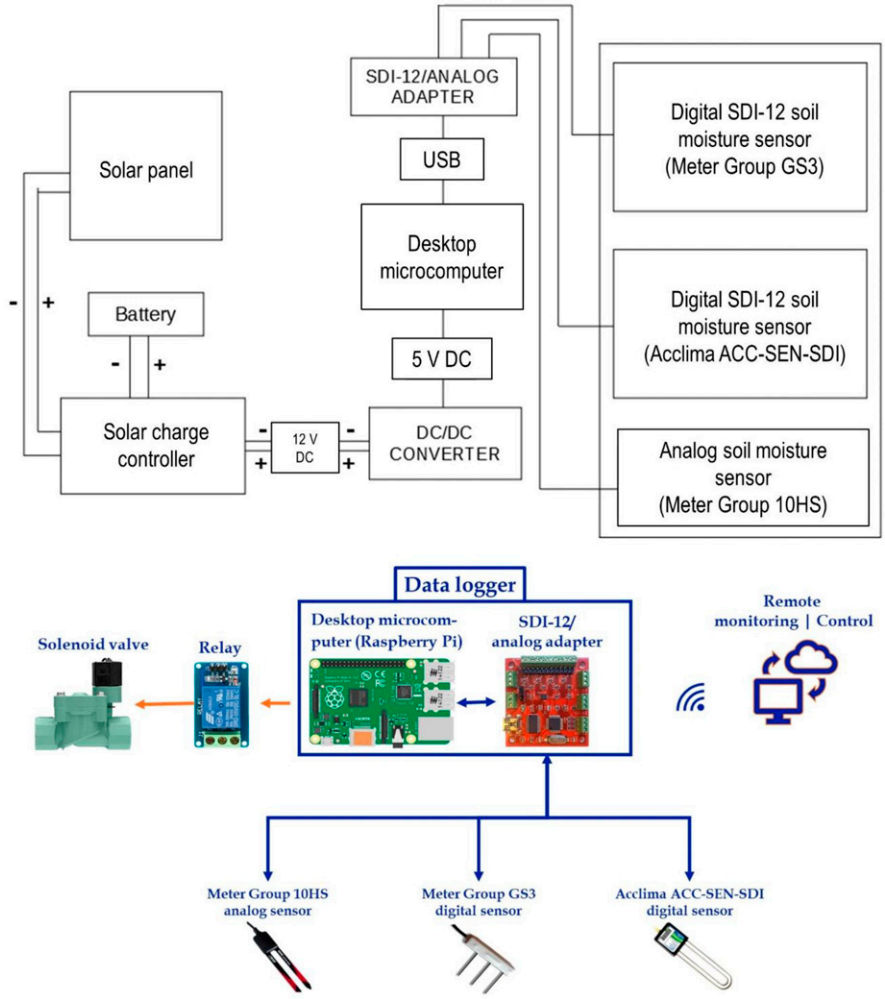


Fig. 4. Diagram of a smart system for automated irrigation using Internet of things devices: embedded computer with a credit-card size and dual-display desktop [Raspberry Pi 4 Model B with 8 Gb of random-access memory (RAM); Raspberry Pi Foundation, Cambridge, United Kingdom] and serial data interface (SDI)-12/ analog adaptor with larger terminals [SDI-12 universal serial bus (USB) adaptor; Liudr's Blog, St. Cloud, MN] connected to three soil moisture sensors {an analog (10HS; Meter Group, Pullman, WA) and two SDI-12 digital sensors [GS3 (Meter Group) and ACC-SEN-SDI (Acclima, Meridian, ID)]}; DC = direct current.

Irrigation manifold. The ability of microcontrollers to successfully turn solenoid valves on and off has been previously described (Ferrarezi et al., 2015; Nemali and van Iersel, 2006). For that reason, we decided not to trigger the irrigation automatically in this study. Instead, we built a manifold consisting of one 5-V DC-operated relay module (SainSmart, Lenexa, KS) and a 24-V alternate current (AC) solenoid valve (Orbit, North Salt Lake, UT). The 5-V DC relay module was wired to the desktop microcomputer's GPIO pins #2, #6, and #23; GPIO pin #2 acts as the internal 5-V power supply to operate the relay, pin #6 is the ground, and pin #23 is the signal output to control the relay. The 5-V DC relay module was equipped with a high-current AC relay up to 30 V and 10 A, allowing to connect the 24-V AC external power supply to trigger the solenoid valve. The solenoid valve was not connected to a water source (Fig. 2).

Software. The desktop microcomputer runs its unique operating system (Debian GNU/Linux version 10.8; Debian Project, San Francisco, CA) based on the Linux kernel, a free and open-source community-developed operating system. The program was built using the desktop microcomputer's built-in open-source software IDE (Thonny version 3.3.6; University of Tartu, Tartu, Estonia), which is available for Python programming language to write codes and upload the compiled data program into the desktop microcomputer. Our automated system was designed to run the script with Python version 3.4 or higher. We used a built-in virtual networking computing software (RealVNC Viewer version 6.20.817; RealVNC, Cambridge, UK) to access the desktop microcomputer for remote control and troubleshooting. Additional instructions for setting up the desktop microcomputer and its software are accessible on the manufacturer's website. We used a cloud server to store the data online and allow real-time monitoring (ThingSpeak, 2020), which is an open-source IoT analytics platform that enables users to aggregate, visualize, and analyze live data streams in the cloud. Users have a unique channel identification and key code through the application programming interface provided by the website after creating an account. The website also provides instructions to assist in setting up and configuring the channel's webpage. Users can decide to set the streaming channel in public or private mode, and the webpage could be viewed at any time from any place via a web browser or an app on a computer or a mobile device.

Code. The Python script running the system is exemplified in Fig. 3 and fully accessible as Supplemental Material. The code was modified from the open-source script available at Liu (2021).

Experimental setup. To test the power consumption of the system in a field where an outlet is not accessible, we used a 12-V 7-A DC rechargeable battery, solar panel, charge controller (CMP-12; Unique Bargains, San Francisco, CA), and a DC/DC 12-V to 5-V adaptor (Homree, Hong Kong, China). The solar panel was connected to the charge controller to regulate the incoming voltage, and the charge controller was wired to the 12-V DC rechargeable battery, which powered the entire system. Finally, the DC/DC adaptor was used to convert the 12-V output from the battery to 5 V to power the desktop microcomputer.

The desktop microcontroller, charge controller, and the DC/DC adaptor were installed in a waterproof enclosure, and the system was placed in an outdoor garden in central Florida. We used an all-natural potting mix for outdoor containers with pH ranging from 5.8 to 7.5 (All-Natural Potting Mix Premium for Outdoor Container; Kellogg Garden Organics, Carson, CA).

The system diagram is shown in Fig. 4, and the hardware components and the approximate cost are listed in Table 2. If users decide on using a regular 110-V AC outlet as a power

Table 2. List of materials and approximate costs for the automated irrigation system.

Main components	Part model[z]	Source	Cost
Desktop microcomputer or datalogger	Raspberry Pi 4 Model B	Raspberry Pi Foundation, Cambridge, UK	$35.00
	Serial data interface (SDI)-12/analog adaptor	Liudr's Blog, St. Cloud, MN	$79.90
Sensors (3)	10HS analog sensor	Meter Group, Pullman, WA	≈$200.00 each ($600.00)
	GS3 SDI-12 digital sensor	Meter Group	
	ACC-SEN-SDI SDI-12 digital sensor	Acclima, Meridian, ID	
Irrigation manifold	One channel 5-V direct current (DC) relay module	SaintSmart, Las Vegas, NV	$3.70
	1-inch 24-V alternate current (AC) female to female solenoid valve	Orbit, Bountiful, UT	$11.00
	Dual spray half-circle sprinkler	Rain Bird, Azusa, CA	$11.40
Power supply	12-V 7-A DC rechargeable battery	Generic Power System, Waukesha, WI	$30.00[y]
	20-W solar panel	Generic Power System	$25.00[y]
	CMP-12 5-A solar charge controller	SaintSmart	$18.00[y]
	DC/DC 12-V to 5-V adaptor	SSLHONG, Shenzhen, China	$14.00
Others	20 × 12 × 5.6-cm waterproof enclosure[z]	LeMotech ABS, Shenzhen, China	$30.00
	SanDisk 32 GigaByte (GB) microsecure digital (SD) card	Western Digital, San Jose, CA	$8.00
Total			$866.00

[z]1 inch = 2.54 cm, 1 cm = 0.3937 inch.
[y]Items not needed if the system is powered using electricity.

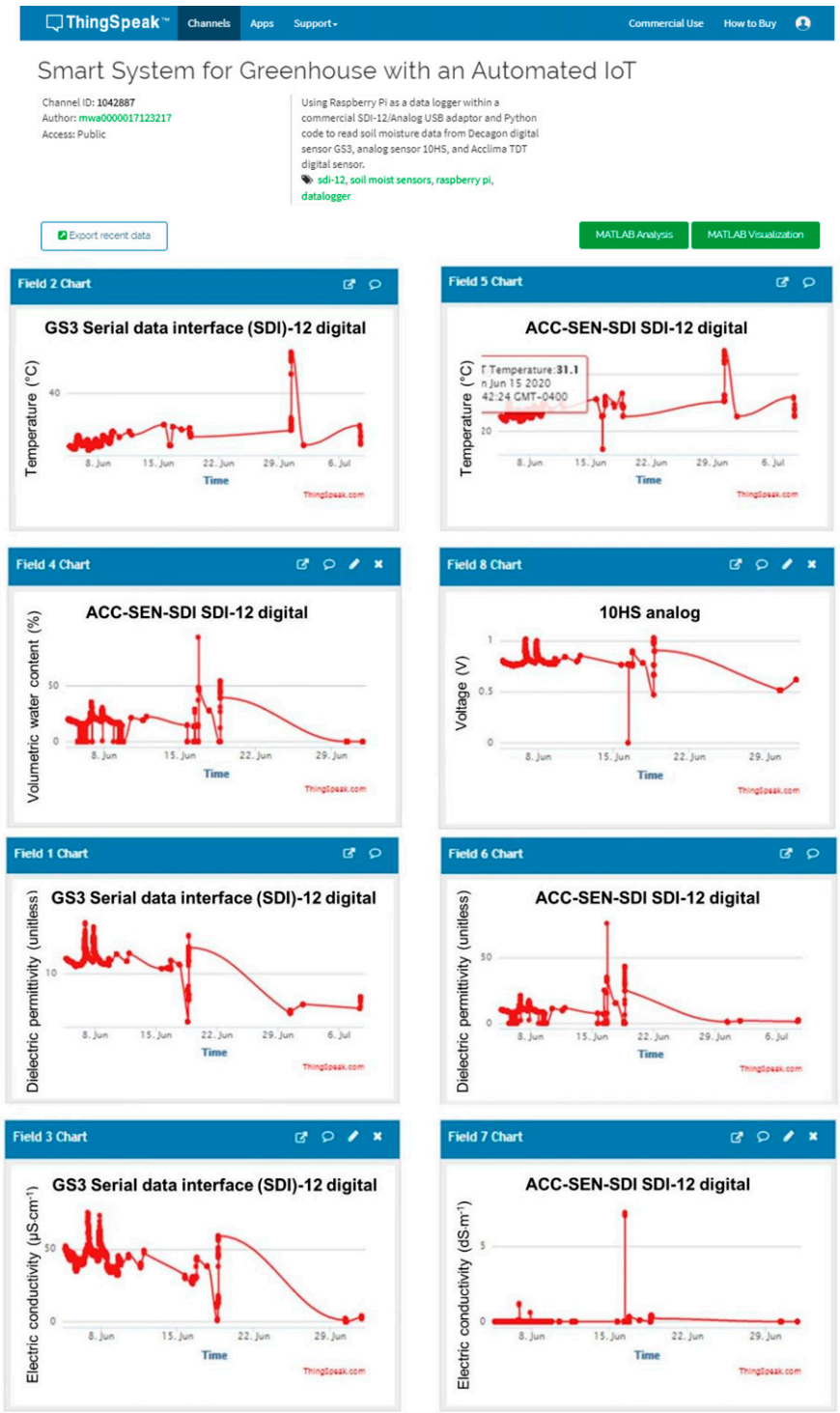


Fig. 5. Snapshot of cloud server storing the data and allowing real-time data monitoring for on demand irrigation in June 2020 (ThingSpeak, 2020); 1 $m^3 \cdot m^{-3}$ = 100%, (1.8 × °C) + 32 = °F, 1 $\mu S \cdot cm^{-1}$ = 1 µmho/cm, 1 $dS \cdot m^{-1}$ = 1 mmho/cm, 1 cm = 0.3937 inch.

supply, the price for the rechargeable battery set up can be reduced. The cost of soil moisture sensors and the irrigation manifold can vary depending on the users' choice. For this experiment, a one-channel 5-V DC-operated relay was used. We programed the relay to be controlled based on the sensor readings. If the VWC values dropped below 20%, the desktop microcomputer output 5 V DC to turn the relay on for 10 s. When the relay is on, it opened the solenoid valve to simulate letting the water flow through the system. At the same time, the system updated its irrigation status to the live streaming platform.

Results

Our goal was to search for a robust, low-cost open-source setup to read SDI-12 digital sensors and test if the system can run with the rechargeable solar power supply system for 1 week without maintenance.

Sensor readings. The ε expresses a substance's ability to hold an electrical charge and is represented as a substance's permittivity ratio by the vacuum permittivity. The ε equals 1 in air and ≈80 in water (Ferrarezi et al., 2020), easily discriminating moisture in different soil types and soilless growing media. We tested how the sensor readings responded in water and air and an open outdoor setting, and the embedded system successfully recorded the sensor readings. The ε readings in water for 10HS analog sensor = 81.4, GS3 digital sensor = 76.0, and ACC-SEN-SDI digital sensor = 73.3. In air, the ε reading for 10HS = 2.34, GS3 = 1.42, and ACC-SEN-SDI = 1 (Table 3). VWC is the ratio of water volume to soil volume. The 10HS sensor's reading is about 65% in water, GS3 equals 90%, and ACC-SEN-SDI 89%. Typically, the VWC ranges from 0% to 100% or 0 to 1 m^3 of water per cubic meter of soil. However, the range for the 10HS sensor is from 0% to 69% in soilless media (Meter Group, 2021). The VWC reading of sensors in the air all resulted in 0. Some sensors outputted negative values, but that results from the calibration equation terms since there is no negative water volume in soils.

Data collection and monitoring. The experiment was conducted from Mar. to June 2020. All data were successfully collected and uploaded to the cloud every 20 min, allowing users to monitor the data through the ThingSpeak channel remotely. Sample results of the sensors installed in potting soils in June 2020 on display at the ThingSpeak channel are shown in Fig. 5, highlighting the system responded as expected with the soil moisture conditions (dry-down periods followed by sharp increases after a rainfall event). ThingSpeak created a status window to show each sensor's

Table 3. Sensor voltage, dielectric permittivity (ε), and volumetric water content (VWC) readings in completely dry (air) and wet (water) conditions to test sensors' accuracy. The negative values of VWC obtained from sensors in the air represent the dehydrated condition, which could round up to zero.

			Reading value			
					VWC	
Sensor	Address	Location	Voltage (mV)	ε (unitless)	$(m^3 \cdot m^{-3})^z$	(%)
10HS analog	z	In air	430.27	2.34	−0.169	−16.9
		In water	1213.2	81.4	0.652	65.2
GS3 serial data interface (SDI)-12 digital	1	In air	–	1.42	−0.025	−2.50
		In water	–	76.0	0.899	89.9
ACC-SEN-SDI SDI-12 digital	2	In air	–	1.00	0.000	0.00
		In water	–	73.3	0.893	89.3

[z]1 $m^3 \cdot m^{-3}$ = 100%.

VWC reading and irrigation status every 20 min (Fig. 6). The status window is helpful for quickly checking if sensors are working correctly after adding water or moving the sensors for troubleshooting or other reasons. The data collected were analyzed from the CSV files stored in the desktop microcomputer. Fig. 7 shows the three sensors' graphic charts in June 2020 made from data stored locally in the desktop microcomputer. The peaks indicate the quick response of the sensors to an increase in soil moisture due to rainfall. The data shows the soil temperature readings from GS3 and ACC-SEN-SDI sensors were consistent, and the ACC-SEN-SDI sensor registered a higher EC value than the GS3 sensor. Sensor differences are expected since each one has a different measurement and processing technology. Several studies have been conducted to test other sensors in various soil types and conditions, indicating the need for soil-specific calibrations to increase accuracy (Ferrarezi et al., 2020). Furthermore, the wireless communication between the desktop microcomputer and the ThingSpeak website proved reliable and never failed to report data to the cloud.

System performance. We set up the system in a waterproof hard case and tested it outdoor for 3 months. The system experienced the weather variability of subtropical conditions in Florida without significant maintenance requirements. The only issue that occurred was the overheating of the desktop microcomputer due to direct exposure to sunlight, which caused the power consumption to go up. The increased power consumption drained the rechargeable battery faster than the solar panel charging speed, which resulted in the system running out of power in 3 d. After the first attempt, we moved the system to a shaded area to test the power supply performance. The system ran longer than the first test, but the solar panel's charging speed still could not keep up with the entire system's power consumption. We found the system depleted power quickly at night when the solar panel could not be charged. That was caused by the frequent data collection and high power consumption of the desktop microcomputer to collect data and send it in real-time to the cloud. The data collection and solenoid valve activation ran smoothly after switching the rechargeable solar power supply to a stable 110-V power outlet.

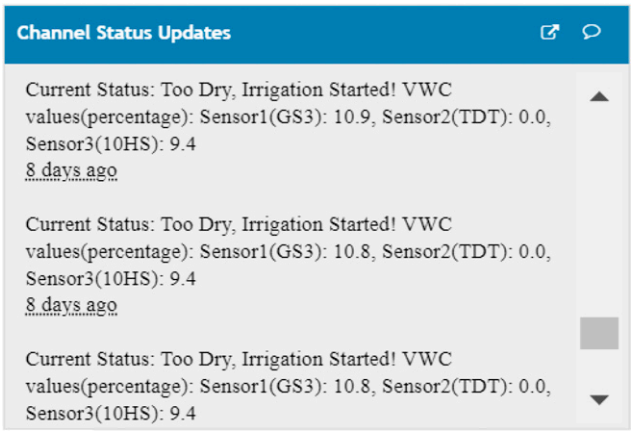


Fig. 6. Status window sample of cloud server platform to store the data and allow real-time volumetric water content (VWC) monitoring (ThingSpeak, 2020); TDT = time-domain transmission.

Conclusion

In this study, a system was developed to periodically collect data using three commercial sensors installed in an outdoor garden with potting soil and automatically send the data to the cloud via Wi-Fi. The system was developed using open-source devices (desktop microcomputer and SDI-12/analog adaptor) with Python script and an IoT platform to live streaming the data (ThingSpeak, 2021). Before installing the system to collect data, laboratory tests were conducted to check if the microcontroller could respond to the sensor readings in air and water based on the manufacturers' calibrations. The system described in this study can be built for $198 to $266 without considering the sensors, and every component of the system is commercially available (Table 2). Heavy-duty and waterproof protection cases are required if the system is running outdoor. Potential future improvements for the system could include more power-efficient instructions to save battery. In addition, the system can become more field-friendly where a stable power outlet is not available or difficult to reach. Making the system more power

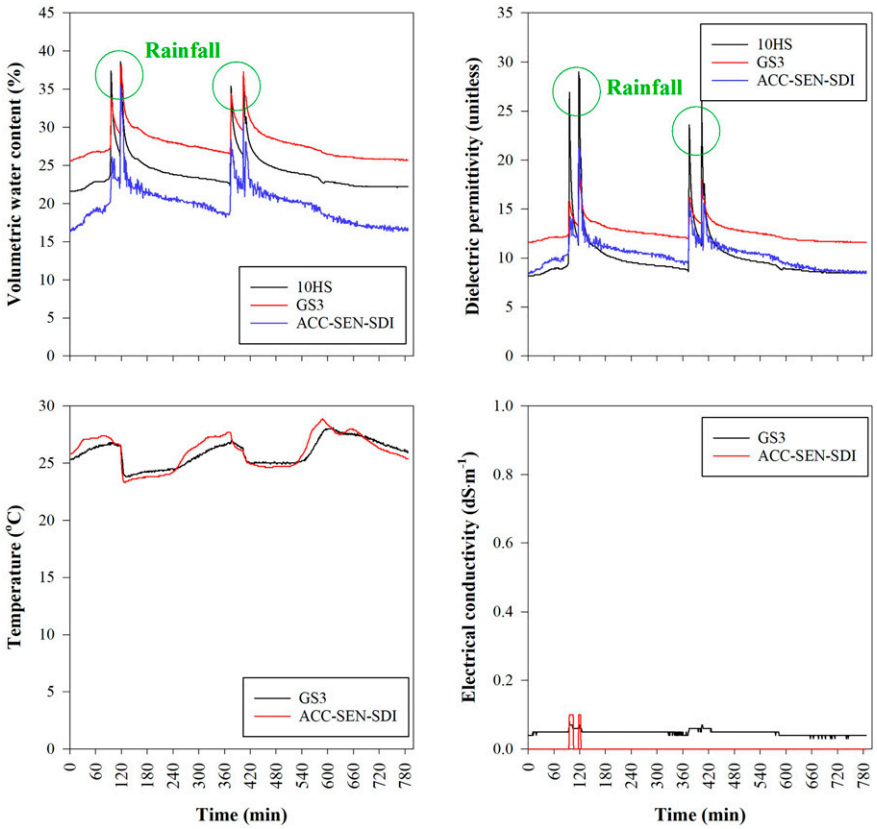


Fig. 7. Volumetric water content, dielectric permittivity, temperature, and electrical conductivity daily sample data extracted from the local comma-separated values (CSV) file; 1 $m^3 \cdot m^{-3}$ = 100%, (1.8 × °C) + 32 = °F, 1 $dS \cdot m^{-1}$ = 1 mmho/cm, 1 cm = 0.3937 inch.

efficient can be accomplished by programming the system only to turn on when data collection is required or using a microcontroller that uses less energy. The system's development and improvement can increase smart farming's popularity for scientists, home users, hobbyists, and small growers.

Literature cited

Ferrarezi, R.S., S.K. Dove, and M.W. van Iersel. 2015. An automated system for monitoring soil moisture and controlling irrigation using low-cost open-source microcontrollers. HortTechnology 25(1): 110–118, https://doi.org/10.21273/HORTTECH.25.1.110.

Ferrarezi, R.S., T.A.R. Nogueira, and S.G.C. Zepeda. 2020. Performance of soil moisture sensors in Florida sandy soils. Water 12(358):1–20, https://doi.org/10.3390/w12020358.

Irmak, S., M.J. Burgert, H.S. Yang, K.G. Cassman, D.T. Walters, W.R. Rathje, J.O. Payero, P. Grassini, M.S. Kuzila, K.J. Brunkhorst, D.E. Eisenhauer, W.L. Kranz, W.B. Van, J.M. Rees, G.L. Zoubek, C.A. Shapiro, and G.J. Teichmeier. 2012. Large-scale on-farm implementation of soil moisture-based irrigation management strategies for increasing maize water productivity. Trans. ASABE 55(3): 881–894, https://doi.org/10.13031/2013.41521.

Liu, J. 2021. SDI-12 USB adapter. 21 June 2021. <https://liudr.wordpress.com/gadget/sdi-12-usb-adapter>.

Meter Group. 2021. 10HS manual. 20 June 2021. <http://publications.metergroup.com/Manuals/20426_10HS_Manual_Web.pdf>.

Nemali, K.S. and M.W. van Iersel. 2006. An automated system for controlling drought stress and irrigation in potted plants. Scientia Hort. 110(3):292–297, https://doi.org/10.1016/j.scienta.2006.07.009.

Payero, J.O., A. Mirzakhani-Nafchi, A. Khalilian, X. Qiao, and R. Davis. 2017. Development of a low-cost Internet-of-Things (IoT) system for monitoring soil water potential using Watermark 200SS Sensors. Adv. Internet of Things 7(3):71–86, https://doi.org/10.4236/ait.2017.73005.

Raspberry Pi Foundation. 2021. GPIO. 29 July 2021. <https://www.raspberrypi.org/documentation/usage/gpio/>.

Ravindra, S. 2020. IoT applications in agriculture. 3 Mar. 2021. <https://www.iotforall.com/iot-applications-in-agriculture>.

ThingSpeak. 2020. ThingSpeak for IoT projects. 20 July 2020. <https://thingspeak.com>.

ThingSpeak. 2021. ThingSpeak for smart farming. 20 Mar. 2021. <http://54.88.253.164/pages/smart_farming>.

Tilman, D., C. Balzer, J. Hill, and B. Befort. 2011. Global food demand and the sustainable intensification of agriculture. Proc. Natl. Acad. Sci. USA 108(50): 20260–20264, https://doi.org/10.1073/pnas.1116437108.

United Nations Department of Economic and Social Affairs. 2015. World population projected to reach 9.7 billion by 2050. 20 Mar. 2021. <https://www.un.org/en/development/desa/news/population/2015-report.html>.

Supplemental material

Computer code used to automate the sensor measurement, irrigation control, and data uploading to an online Cloud computing service. Text in black represents the actual code, and text in blue/italics after # sign denotes explanation.

```
# !/usr/local/opt/python-3.5.1/bin/python3.5
# Serial data interface (SDI)-12 sensor data logger copyright Dr. John Liu 2017-11-06, modified by Rhuanito Ferrarezi & Tzu-Wei Peng

import os # For running command line commands
import time # For delaying in seconds
import datetime # For finding the system's real-time
import serial.tools.list_ports # For listing available serial ports
import serial # For serial communication
import re # For regular expression support
import platform # For detecting operating system
import urllib.parse # For encoding data to be uniform resource locator (URL) safe
import signal # For trapping Ctrl+c or interrupt signal (SIGINT)
import sys # For exiting the program with exit code
import RPi.GPIO as GPIO # For controlling relays
import math # For sensor calibration

GPIO.setmode(GPIO.BOARD) # General purpose Input/Output (GPIO) output for relays
GPIO.setup(23,GPIO.OUT)
GPIO.output(23,GPIO.LOW)
time.sleep(3)

def SIGINT_handler(signal, frame):
    ser.close()
    data_file.close()
    GPIO.output(23,GPIO.LOW)
    print('Quitting program!')
    sys.exit(0)
signal.signal(signal.SIGINT, SIGINT_handler)

channelID = "1042887" # User ThinkSpeak Channel Identification
api_key = "4VGNVVW32D5JMWBG" # User ThingSpeak channel write application programming interface (API) key
curl_command_format='curl "https://api.thingspeak.com/update/?api_key=%s%s%s"' # This is the client URL (cURL) upload command to thingspeak.com
RasPi_A=platform.node() # Use computer name as unit_id. For a Raspberry Pi, change its name from raspberrypi to something else to avoid confusion
curl_exists=False # The code will test whether cURL exists. If yes, it will be used to upload data
adapter_sdi_12_address='z'
no_data=False # This is the flag to break out of the inner loops and continue the next data point loop in case no data is received from a sensor such as the Global Positioning System (GPS)

if (os.system('curl -V')==0):
    curl_exists=True

print('+-'*40)
print('SDI-12 Sensor and Analog Sensor Python Data Logger with Telemetry V1.5.0')
print('Designed for Dr. Liu\'s family of SDI-12 USB adapters (standard,analog,GPS)\n\tDr. John Liu Saint Cloud MN USA 2017-11-06\n\tFree software GNU GPL V3.0\n\t\tModified by Rhuanito Ferrarezi and Tzu-Wei Peng')
print('\nCompatible with Windows, GNU/Linux, Mac OSX, and Raspberry PI')
print('\nThis program requires Python 3.4, Pyserial 3.0, and cURL (data upload)')
print('\nData is logged to YYYYMMDD.CVS in the Python code\'s folder')
print('\nVisit https://thingspeak.com/channels/%s to inspect or retrieve data' %(channelID))
print ('\nFor assistance with customization, telemetry etc., contact Dr. Liu.\n\thttps://liudr.wordpress.com/gadget/sdi-12-usb-adapter/')
print('+-'*40)
ports=[]

VID_FTDI=0x0403;

a=serial.tools.list_ports.comports()
for w in a:
    ports.append((w.vid,w.device))

ports.sort(key= lambda ports: ports[1])

print('\nDetected the following serial ports:')
i=0
for w in ports:
    print('%d)\t%s\t(USB VID=%04X)' %(i, w[1], w[0] if (type(w[0]) is int) else 0))
    i=i+1
total_ports=i # Now i= total ports

user_port_selection=input('\nSelect port from list (0,1,2...). SDI-12 adapter has USB VID=0403:')
if (int(user_port_selection)>=total_ports):
    exit(1) # Port selection out of range

ser=serial.Serial(port=(ports[int(user_port_selection)])[1],baudrate=9600,timeout=10)
time.sleep(2.5) # Delay for Arduino bootloader and the 1 second delay of the adapter
total_data_count=int(input('Total number of data points:'))
delay_between_pts=int(input('Delay between data points (second):'))

print('Time stamps are generated with:\n0) GMT/UTC\n1) Local\n')
time_zone_choice=int(input('Select time zone:'))
if time_zone_choice==0:
    now=datetime.datetime.utcnow() # Use coordinated universal time (UTC) time instead of local time
elif time_zone_choice==1:
    now=datetime.datetime.now() # Use local time, not recommended for multiple data loggers in different time zones

data_file_name="%04d%02d%02d.csv" %(now.year,now.month,now.day)
data_file = open(data_file_name, 'a') # Open yyyymmdd.csv for appending

sdi_12_address=''
user_sdi_12_address=input('Enter all SDI-12 sensor addresses, such as 1234:')
user_sdi_12_address=user_sdi_12_address.strip() # Remove any \r from an input file typed in windows
analog_inputs=input('Collect analog inputs (requires SDI12-USB + Analog adapter)? (Y/N)')
analog_inputs=(analog_inputs.strip()).capitalize() # Remove any \r from an input file typed in windows and capitalize answer

for an_address in user_sdi_12_address:
    if ((an_address>="0") and (an_address<="9")) or ((an_address>="A") and (an_address<="Z")) or ((an_address>="a") and (an_address<="z")):
        print("Using address:",an_address);
    sdi_12_address=sdi_12_address+an_address
    else:
        print('Invalid address:',an_address)

if analog_inputs=='Y':
    sdi_12_address=sdi_12_address+adapter_sdi_12_address
if len(sdi_12_address)==0:
    sdi_12_address=adapter_sdi_12_address # Use default address

for an_address in sdi_12_address:
    ser.write(an_address.encode()+b'I!') # Send identification
    sdi_12_line=ser.readline()
    print('Sensor address:',an_address,' Sensor info:',sdi_12_line.decode('utf-8').strip())

print('Saving to %s' %data_file_name)
print('='*40)

for j in range(total_data_count):
```

```
    i=0 # This counts to 6 to truncate all data to the 6 values set up in Sparkfun's Phant server upload
    value_str='' # This stores &value0=xxx&value1=xxx&value2=xxx-&value3=xxx&value4=xxx&value5=xxx and is only reset after all sensors are read
    if time_zone_choice==0:
        now=datetime.datetime.utcnow()
    elif time_zone_choice==1:
        now=datetime.datetime.now()
    output_str="%04d/%02d/%02d %02d:%02d:%02d%s" %(now.year,now.-month,now.day,now.hour,now.minute,-now.second,' GMT' if time_zone_choice==0 else '')
# Formatting date and time
for an_address in sdi_12_address:
    ser.write(an_address.encode()+b'M!');
# Start the SDI-12 sensor measurement
    # print(an_address.encode()+b'M!'); # Start the SDI-12 sensor measurement
    sdi_12_line=ser.readline()
    # print(sdi_12_line)
    sdi_12_line=sdi_12_line[:-2] # Remove \r and \n since [0-9]$ has trouble with \r
    m=re.search(b'[0-9]$',sdi_12_line) # Having trouble with the \r
    total_returned_values=int(m.group(0)) # Find how many values are returned
    sdi_12_line=ser.readline()
# Read the service request line
    ser.write(an_address.encode()+b'D0!') # Request data
    # print(an_address.encode()+b'D0!') # Request data
    sdi_12_line=ser.readline()
# Read the data line
    # print(sdi_12_line)
    sdi_12_line=sdi_12_line[1:-2]
# Remove address, \r and \n since [0-9]$ has trouble with \r

    values=[] # Clear cache before each sensor
    for iterator in range(total_returned_values): # Extract the returned values from SDI-12 sensor and append to values[]
        m=re.search(b'[+-][0-9.]+',sdi_12_-line) # Match a number string
        try: # If values found is less than values indicated by return from M, report no data found. This is a simple solution to GPS sensors before they acquire lock. For sensors that have lots of values to return, you need to find a better solution
            values.append(float(m.group(0))) # Convert into a number
            sdi_12_line=sdi_12_line[len(m.group(0)):]
        except AttributeError:
            print("No data received from sensor at address %c\n" %(an_address))
            time.sleep(delay_between_pts)
            no_data=True
            break
    if (no_data==True):
      break;

    output_str=output_str+'\n'+an_address

    for value_i in values:
      output_str=output_str+",%s" %(value_i) # Output returned values depends on your sensor choices and its calibration
        if (i==0): # GS3 dielectric value
          GS3_VWC=round(((0.118* ((value_i)**0.5))-0.117)*100,1) # GS3 sensor volumetric water content (VWC) mineral soil calibration
        if (i==3): # ACC-SEN-SDI sensor VWC value in percent
          TDT_VWC=value_i
        if (i==7 and i<8): # 10HS analog value
          mV=(value_i)*1000
          ANLG_VWC=round((((7.05*(10**-7))*((mV)**2))-((1.23*(10**-4))*(mV))-(3.03*(10**-2)))*100,1) # 10HS sensor VWC mineral soil calibration
        if (i<8):
          value_str=value_str+"&field%d=%s" %(i+1,value_i) # Format values for posting. Field starts with field1
          i=i+1

    if (no_data==True):
      no_data=False
      continue;
    while (i<8): # Pad with zeros in case we do not have six characters. This is only necessary for certain servers
        value_str=value_str+"&field%d=0" %(i+1) # Format values for posting. Field starts with field1, not field0
        i=i+1

    status_str='Current Status:'
    if (GS3_VWC<20 or TDT_VWC<20 or ANLG_VWC<20): # Put your own desire irrigation status
        irrigation_status="Too Dry, Irrigation Started! VWC values(percentage): Sensor1(GS3): %s, Sensor2(TDT): %s, Sensor3(10HS): %s" %(GS3_VWC, TDT_VWC, ANLG_VWC)
        irrigation_counts=1
        GPIO.output(23,GPIO.HIGH)
        time.sleep(10)
        GPIO.output(23, GPIO.LOW)
    else:
        irrigation_status='Do Not Need Irrigation. VWC values: Sensor1(GS3): %s, Sensor2(TDT): %s, Sensor3(10HS): %s' %(GS3_VWC, TDT_VWC, ANLG_VWC)
        irrigation_counts=0
        GPIO.output(23, GPIO.LOW)
    status_str_curl="&status=Current Status: %s" %(irrigation_status)
    status_str=status_str+"%s" %(irrigation_status)
    irrigation_system=output_str+"\n %s" %(irrigation_counts)

    print(output_str)
    print(status_str)
    print(irrigation_counts)
    output_str=output_str+'\n'
    irrigation_system=irrigation_system-+'\n'
    data_file.write(irrigation_system)
    if (curl_exists==True):
    curl_command=curl_command_format %(api_key,value_str,status_str_curl)
# Format cURL command

    print('=='*40)

    print(curl_command) # Debug information
    print(os.system(curl_command)) # Send data to ThinkSpeak.com using cURL

    values=[] # Clear values for the next iteration (ThinkSpeak 3.2.3 does not support clear and 3.4.3 and 3.5.1 does)
    data_file.flush() # Make sure data is written to the disk so stopping the script with ctrl+C will not cause data loss
    time.sleep(delay_between_pts)
    ser.close()
    data_file.close()
```

ARTICLES FOR FACULTY MEMBERS

A COST-EFFECTIVE IOT-BASED SOIL MOISTURE MONITORING SYSTEM FOR WILDFIRE-PRONE COASTAL SOILS

Using soil moisture information to better understand and predict wildfire danger: A review of recent developments and outstanding questions / Krueger, E. S., Levi, M. R., Achieng, K. O., Bolten, J. D., Carlson, J. D., Coops, N. C., Holden, Z. A., Magi, B. I., Rigden, A. J., & Ochsner, T. E.

International Journal of Wildland Fire
Volume 32 Issue 2 (2023) Pages 111–132
https://doi.org/10.1071/wf22056
(Database: CSIRO Publishing)

CSIRO PUBLISHING

REVIEW
https://doi.org/10.1071/WF22056

International Journal of WILDLAND FIRE
International Association of Wildland Fire

Using soil moisture information to better understand and predict wildfire danger: a review of recent developments and outstanding questions

Erik S. Krueger[A,*], Matthew R. Levi[B], Kevin O. Achieng[B,I], John D. Bolten[C], J. D. Carlson[D], Nicholas C. Coops[E], Zachary A. Holden[F], Brian I. Magi[G], Angela J. Rigden[H,§] and Tyson E. Ochsner[A]

For full list of author affiliations and declarations see end of paper

*Correspondence to:
Erik S. Krueger
Plant and Soil Sciences, Oklahoma State University, Stillwater, Oklahoma, USA
Email: erik.krueger@okstate.edu

§During a portion of this work, Rigden was with Earth and Planetary Sciences, Harvard University.

Received: 21 April 2022
Accepted: 21 October 2022
Published: 5 December 2022

Cite this:
Krueger ES *et al.* (2023)
International Journal of Wildland Fire
32(2), 111–132. doi:10.1071/WF22056

ABSTRACT

Soil moisture conditions are represented in fire danger rating systems mainly through simple drought indices based on meteorological variables, even though better sources of soil moisture information are increasingly available. This review summarises a growing body of evidence indicating that greater use of *in situ*, remotely sensed, and modelled soil moisture information in fire danger rating systems could lead to better estimates of dynamic live and dead herbaceous fuel loads, more accurate live and dead fuel moisture predictions, earlier warning of wildfire danger, and better forecasts of wildfire occurrence and size. Potential uses of soil moisture information in existing wildfire danger rating systems include (1) as a supplement or replacement for drought indices, (2) for live and (3) dead fuel moisture modelling, (4) for estimating herbaceous fuel curing, and (5) for estimating fuel loads. We identify key remaining research questions and note the logistical challenge of convincing wildfire professionals of the importance of soil moisture compared with more familiar wildfire danger metrics. While obstacles remain, the path forward is clear. Soil moisture information can and should be used to improve fire danger rating systems and contribute to more effective fire management for the protection of communities and ecosystems worldwide.

Keywords: fuel properties, *in situ*, modelling, remote sensing, review, soil moisture, wildfire, wildfire danger index, wildfire danger rating systems.

Introduction

At 6:33 am on the morning of 8 November 2018, a small fire was reported under electrical power lines near Camp Creek Road outside the town of Pulga in northern California, USA. Dry conditions and strong downslope winds with gusts $>25\ m\ s^{-1}$ (Brewer and Clements 2020) rapidly transformed that small fire into the deadliest and most costly wildfire in California's history. The Camp Fire burned >62 000 ha, destroyed >18 000 structures, and resulted in 85 fatalities (California Department of Forestry and Fire Protection 2019). This tragedy powerfully illustrates the importance of fire danger rating systems and the need to provide earlier and more accurate warnings for fire management agencies and the public. Toward that end, this review explores recent developments, knowledge gaps, and challenges in applying previously underutilised soil moisture information to better understand, assess, and predict wildfire danger. Up until now, the incorporation of soil moisture information into existing fire danger rating systems has been limited to simplistic models or drought indices that use standard weather variables to estimate soil moisture, even though better information is becoming increasingly available via *in situ* measurements, remote sensing, and more sophisticated modelling. One week prior to the tragic Camp Fire, for example, satellite observations showed strong negative soil moisture anomalies across northern California (Fig. 1), conditions that are known to substantially increase the probability of large wildfires

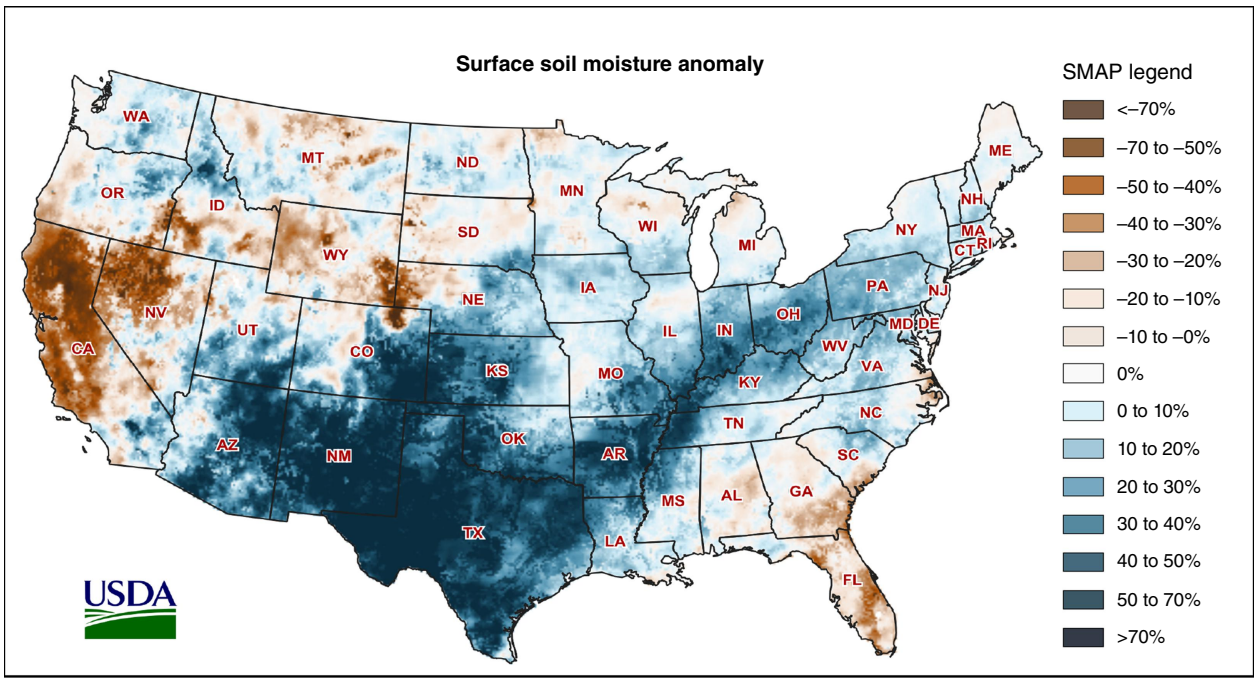


Fig. 1. Measured surface soil moisture anomaly in approximately the top 15 cm of the soil profile on 1 November 2018, 1 week prior to the Camp Fire in northern California, the deadliest and most destructive wildfire in the state's history. Soil moisture anomaly is calculated as deviation from average for that day, as reported by NASA's SMAP satellite mission. The map indicates exceptionally dry soil conditions conducive to high fire danger in northern California (image: USDA NASS Crop Condition and Soil Moisture Analytics system, https://nassgeo.csiss.gmu.edu/CropCASMA/).

(Krueger *et al.* 2015, 2016; Sazib *et al.* 2022), but the tools needed to effectively put this information into action are currently lacking.

It is not a recent revelation that soil moisture conditions are important for fire danger rating. Prominent fire danger rating systems in Canada, Australia, and the United States each use approximations of the moisture of mineral and/or organic soil horizons to quantify wildfire danger (Kumar and Dharssi 2015). For example, the Canadian Forest Fire Danger Rating System (CFFDRS) (Stocks *et al.* 1989; Wotton 2009) includes three moisture indices, termed moisture codes, to represent moisture stored in the organic layers of the forest floor. The Fine Fuel Moisture Code (FFMC) represents fuel moisture of fine surface litter with a depth of 10–20 mm, the Duff Moisture Code (DMC) represents fuel moisture of loosely compacted duff with a depth of about 50–100 mm, and the Drought Code (DC) represents fuel moisture of deep organic materials having a depth of about 100–200 mm (de Groot 1987) (Fig. 2). While intended to represent moisture of surface organic layers, DMC and DC are strongly correlated with soil moisture of mineral horizons near the surface (Pellizzaro *et al.* 2007; D'Orangeville *et al.* 2016), likely in part because of capillary and vapour flow between mineral and organic soil layers (Zhao *et al.* 2022). When considering soils with deep organic layers at the surface, i.e. deep O horizons in soil science terminology, the water stored in those layers may be viewed as either soil moisture or fuel moisture because the organic layer itself can become combustible at low water contents.

In the recently modified Australian Fire Danger Rating System (AFDRS, Matthews 2022), fire danger ratings for dry eucalypt forests are dependent in part on soil moisture deficit estimated using the Keetch–Byram Drought Index (KBDI, Keetch and Byram 1968). The KBDI uses temperature and precipitation data to estimate the soil moisture deficit in the upper soil layers (mineral and organic, if present) using a water balance approach. The KBDI was designed to represent the moisture deficit in approximately the top 760–890 mm for a fine-textured soil and greater depths for coarser-textured soils (Keetch and Byram 1968) (Fig. 2). Similarly, in the National Fire Danger Rating System (NFDRS) used in the United States (Deeming *et al.* 1972; Bradshaw *et al.* 1983; Burgan 1988; Jolly 2018), KBDI helps determine fire danger ratings through its influence on dead fuel load. The inclusion of these moisture indices in widely used fire danger rating systems makes it clear that their developers recognised the importance of soil moisture for understanding wildfire danger. However, at the time these systems were developed, large-scale soil moisture measurement systems and physically-based hydrologic models were not sufficiently developed. Instead, meteorological observations from existing weather stations were used to calculate soil moisture indices. Given advances in soil moisture measurement and modelling systems in recent decades, there is

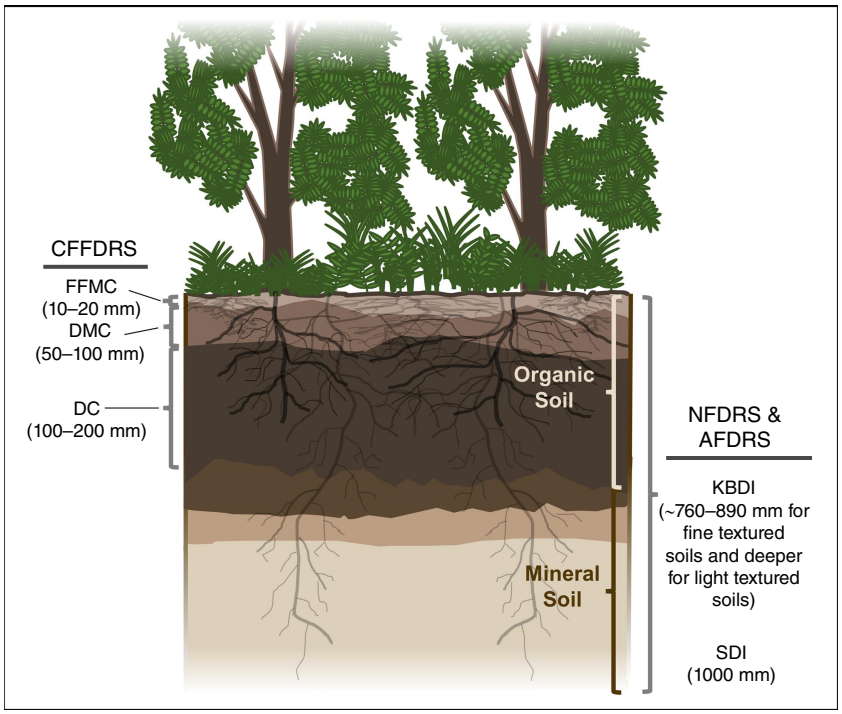


Fig. 2. Diagram of a hypothetical forest soil profile that approximates the relationships of mineral and organic soil layers to moisture indices used in fire danger rating systems in Canada (CFFDRS), the United States (NFDRS), and Australia (AFDRS). Moisture indices in the CFFDRS include the Fine Fuel Moisture Code (FFMC), the Duff Moisture Code (DMC), and the Drought Code (DC), which represent only the organic layers. The AFDRS uses the Keetch–Byram Drought Index (KBDI) and the Soil Dryness Index (SDI), each of which represent soil moisture in mineral and organic layers, and the NFDRS uses KBDI. None of these indices use measured soil moisture, account for physical properties of the soil, or directly account for impacts of overlying vegetation. Instead, moisture content is estimated using simplistic water balance approaches based on weather variables.

a need to reassess how to best represent the moisture conditions of organic and mineral soil layers in fire danger rating systems and to better understand the effects of those conditions on fire danger.

The effectiveness of fire danger rating systems can be determined through retrospective analyses of the relationship between fire danger ratings and important wildfire metrics including occurrence and size. For example, a recent analysis of the NFDRS showed generally positive correlations between fire danger ratings and fire sizes across the contiguous US, but there were important spatial inconsistencies. Notably, there was poorer performance in the eastern half of the country compared to the western half, possibly due to regional differences in soil–vegetation–climate interactions and in the timing and length of the fire season (Walding *et al.* 2018). Furthermore, large areas in the central US lacked the necessary data to generate fire danger ratings because those areas contained only a limited number of designated reporting stations for the Weather Information Management System, which provides weather data for the NFDRS (Walding *et al.* 2018). Thus, improvements to the NFDRS will likely need to consider both model structural improvements, as well as new and better sources and types of input data. Currently, the NFDRS and most other fire danger rating systems in use around the world rely on a relatively standard set of input variables that are routinely measured at weather stations, chief among them being air temperature, relative humidity, precipitation, and wind speed (de Groot *et al.* 2015), as well as the aforementioned simplistic moisture indices derived from these measurements.

A major hindrance to the integration of improved sources of soil moisture information into fire danger rating systems has been the inadequate duration and spatial extent of those sources. That situation is rapidly changing as a variety of new sources of soil moisture information are becoming available, each with unique strengths (Fig. 3). These new data sources include (1) soil moisture measured *in situ*, (2) soil moisture measured remotely by satellites, and (3) soil moisture data that are generated using physically-based models. This groundswell of information began with the advent of state and national automated soil moisture monitoring networks in the US in the late 1990s and the subsequent emergence of similar networks in other countries around the world (Dorigo *et al.* 2021). In parallel, satellite missions capable of monitoring soil moisture and closely-related variables have been developed and launched by NASA and other space agencies, with substantial increases in daily coverage of the Earth's surface since the late 1990s (Karthikeyan *et al.* 2017). These advances in soil moisture measurements have occurred alongside advances in numerical models, which can now provide soil moisture estimates for large domains with sub-kilometre resolution (Holden *et al.* 2019). Using these three types of soil moisture information, researchers began to generate first glimpses of the strong relationships between wildfire and *in situ* soil moisture (Krueger *et al.* 2015), remotely-sensed soil moisture (Bartsch *et al.* 2009), and modelled soil moisture (Slocum *et al.* 2010) (Table 1). Further studies have provided new insights into the relationships between soil moisture and fuel characteristics including fuel loads (e.g. Ellsworth *et al.* 2013; Sharma *et al.* 2018), curing (e.g. Wittich 2011;

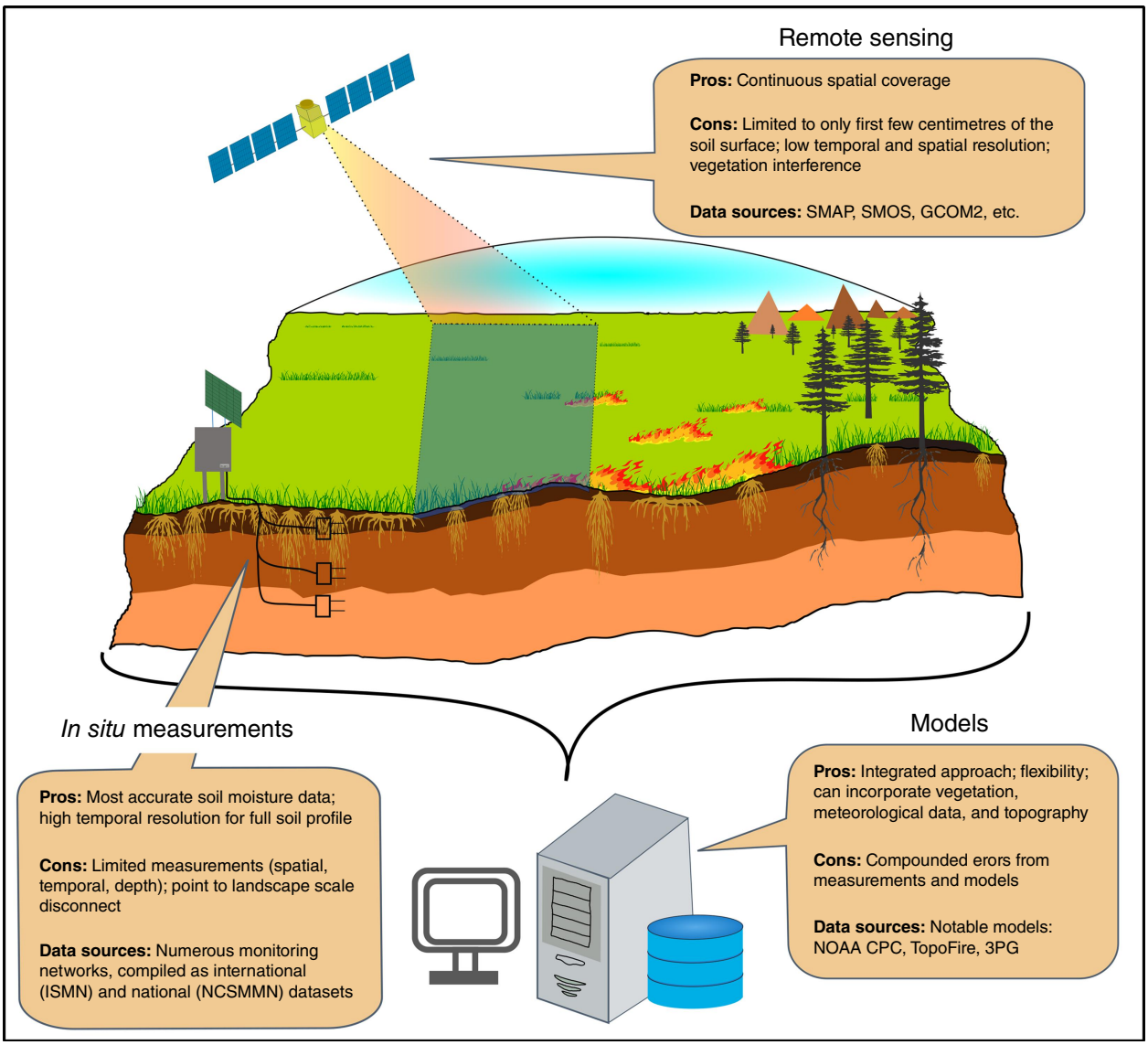


Fig. 3. A variety of *in situ*, remotely sensed, and modelled soil moisture data sources have been recently developed, with each having unique qualities making them well suited for wildfire danger modelling.

Sharma *et al.* 2021), and live (e.g. Bianchi and Defossé 2015; Fan *et al.* 2018) and dead fuel moisture (e.g. Masinda *et al.* 2021; Rakhmatulina *et al.* 2021). Other studies have directly related soil moisture to fire occurrence (Jensen *et al.* 2018; Ambadan *et al.* 2020) and fire size (e.g. Slocum *et al.* 2010; Forkel *et al.* 2012; Krueger *et al.* 2015), while still others have identified the impact of vegetation type on soil moisture–wildfire relationships (e.g. Schaefer and Magi 2019; Rigden *et al.* 2020). These and other important contributions (Table 1) to our understanding of soil moisture–wildfire relationships have emerged across a wide variety of scientific disciplines, which often are not well-connected, making the accelerating progress difficult to track and synthesise.

A further roadblock complicating the use of soil moisture information for fire danger ratings is that soil moisture conditions can be expressed in a variety of ways, making it more difficult to compare results across studies. For example, soil moisture can be expressed simply as soil volumetric water content (e.g. Schaefer and Magi 2019; Ambadan *et al.* 2020; Vinodkumar *et al.* 2021) or water content summed over some soil depth, i.e. soil water storage (e.g. Slocum *et al.* 2010; Krawchuk and Moritz 2011; Chikamoto *et al.* 2015). Alternatively, soil moisture can be formulated to represent the amount of soil moisture available to plants (Krueger *et al.* 2019), and it may also be normalised to allow for comparison across sites or across different soil moisture metrics. This normalisation procedure may be based on the

Table 1. Chronological list of some important contributions to improving our understanding of soil moisture–wildfire relationships.

Authors	Country	Wildfire or fuel metrics	Vegetation types	Soil moisture
Studies focused on fire				
Bartsch *et al.* (2009)	Russia	Wildfire extent	Forest (boreal)	Remotely sensed
Slocum *et al.* (2010)	USA	Wildfire size	Forest, grass, marsh	Modelled
Krawchuk and Moritz (2011)	Global	Wildfire occurrence	Multiple types	Remotely sensed
Forkel *et al.* (2012)	Russia	Wildfire extent	Forest (larch)	Remotely sensed
Chikamoto *et al.* (2015)	North America	Wildfire frequency	Multiple types	Modelled
Krueger *et al.* (2015)	USA	Wildfire size	Forest, shrub, grass	*In situ*
Chaparro *et al.* (2016)	Spain	Wildfire extent	Multiple types	Remotely sensed
Krueger *et al.* (2016)	USA	Wildfire probability	Forest, shrub, grass	*In situ*
Waring and Coops (2016)	Canada, USA	Wildfire occurrence	Forest	Modelled
Forkel *et al.* (2017)	Global	Wildfire extent	Multiple types	Remotely sensed
Krueger *et al.* (2017)	USA	Wildfire occurrence	Forest, shrub, grass	*In situ*
Jensen *et al.* (2018)	USA	Wildfire occurrence, extent	Multiple types	Remotely sensed
Holden *et al.* (2019)	USA	Wildfire occurrence	Multiple types	Modelled
Schaefer and Magi (2019)	Global	Wildfire occurrence	Multiple types	Remotely sensed
Vinodkumar and Dharssi (2019)	Australia	Wildfire radiative power	Forest, shrub, grass	Modelled
Ambadan *et al.* (2020)	Canada	Wildfire occurrence	Multiple types	Remotely sensed
O *et al.* (2020)	Global	Wildfire occurrence	Multiple types	Remotely sensed
Rigden *et al.* (2020)	USA	Wildfire occurrence	Forest, shrub, grass	Remotely sensed
Sazib *et al.* (2022)	Australia, USA	Wildfire occurrence	Multiple types	Remotely sensed
Studies focused on fuel				
Pook and Gill (1993)	Australia	Fuel moisture (dead)	Forest (pine)	*In situ*
Samran *et al.* (1995)	Canada	Fuel moisture (dead)	Forest (aspen)	*In situ*
Pellizzaro *et al.* (2007)	Italy	Fuel moisture (live)	Shrub	*In situ*
Wittich (2011)	Germany	Fuel moisture (live, dead)	Grass	*In situ*
Qi *et al.* (2012)	USA	Fuel moisture (live)	Shrub	*In situ*
Ellsworth *et al.* (2013)	USA	Fuel load, moisture (live, dead)	Grass	*In situ*
Bianchi and Defossé (2015)	Argentina	Fuel moisture (live)	Forest (ñire and cypress), other	*In situ*
Myers-Smith *et al.* (2015)	Multiple (arctic)	Fuel Load	Shrub	Remotely sensed
Burapapol and Nagasawa (2016)	Thailand	Fuel moisture (dead)	Forest (dipterocarp, deciduous)	Remotely sensed
D'Orangeville *et al.* (2016)	Canada	Drought code	Forest (maple, fir, spruce)	*In situ*
McGranahan *et al.* (2016)	South Africa	Fuel moisture, fire danger	Grass	*In situ*
Ackerman *et al.* (2017)	USA	Fuel load	Shrub	*In situ*
Elmes *et al.* (2018)	Canada	Fuel moisture (dead)	Forest, peatland	*In situ*
Fan *et al.* (2018)	France	Fuel moisture (live)	Shrub	Remotely sensed
Ruffault *et al.* (2018)	France	Fuel moisture (live)	Shrub	Modelled
Jia *et al.* (2019)	USA	Fuel moisture (live)	Shrub	Remotely sensed
Krueger *et al.* (2019)	USA	Fuel Load	Grass	*In situ*
Sharma *et al.* (2021)	USA	Fuel moisture (live, dead)	Grass	*In situ*

(*Continued on next page*)

Table 1. (*Continued*)

Authors	Country	Wildfire or fuel metrics	Vegetation types	Soil moisture
Krueger *et al.* (2021)	USA	Fuel Load	Grass	*In situ*
Lu and Wei (2021)	USA	Fuel moisture (live)	Multiple types	Remotely sensed
Lyons *et al.* (2021)	USA	Fuel moisture (live)	Forest, shrub, grass/forb	Modelled
Masinda *et al.* (2021)	China	Fuel moisture (dead)	Forest (pine)	*In situ*
Rakhmatulina *et al.* (2021)	USA	Fuel moisture (dead)	Forest (conifer)	*In situ*
Vinodkumar *et al.* (2021)	Australia	Fuel moisture (live)	Forest, shrub, grass	Modelled
Zhao *et al.* (2021)	Australia	Fuel moisture (dead)	Forest, shrub, woodland	*In situ*
Zhao *et al.* (2022)	Australia	Fuel moisture (dead)	Woodland	*In situ*

Studies are separated by those focused on soil moisture relationships with wildfire and those focused on relationships with fuels. The publication rate for studies on soil moisture–wildfire relationships has increased substantially in the past decade.

physical properties of the soil (e.g. Krueger *et al.* 2015; Waring and Coops 2016; Vinodkumar *et al.* 2017) or use statistical techniques (Lyons *et al.* 2021). To further complicate the situation, soil moisture may be expressed across different soil depths (Fan *et al.* 2018; Vinodkumar *et al.* 2021) and as absolute values or anomalies (O *et al.* 2020; Sazib *et al.* 2022). These varied formulations of soil moisture may have inadvertently prevented its widespread use and obscured the growing body of literature establishing the important relationships between soil moisture and wildfire.

Therefore, our objectives are to (1) summarise the rapidly growing body of research on soil moisture–wildfire relationships, (2) broaden the community of researchers aware of and engaged in this line of research, and (3) make a convincing case for more widespread use of soil moisture information in operational fire danger rating systems. This review is organised into four primary sections. The first three sections summarise what is known about the relationships of wildfire and fuel bed properties to (1) *in situ* soil moisture measurements, (2) remotely sensed soil moisture, and (3) modelled soil moisture. The fourth section explains potential links between soil moisture information and existing fire danger rating systems, using NFDRS as one specific example. We conclude by describing primary challenges and opportunities for using soil moisture information to better understand and predict wildfire danger, including the identification of key areas of needed future research.

In situ soil moisture measurements

In situ soil moisture measurements are the gold standard of soil moisture information (Levi *et al.* 2019) against which remote sensing and modelled soil moisture values are evaluated (Fig. 3). In some geographic areas, *in situ* soil moisture data are available at sufficient spatial and temporal resolutions to inform wildfire management. The International Soil Moisture Network (ISMN) houses publicly available data from nearly 2700 *in situ* soil moisture monitoring stations across 65 networks worldwide, a number that is steadily growing (Dorigo *et al.* 2021). The United States has an especially prolific collection of *in situ* soil moisture monitoring networks, and the ongoing National Coordinated Soil Moisture Monitoring Network (NCSMMN) initiative aims to produced harmonised data products from *in situ* soil moisture measurements from approximately 2000 sites across the nation (Cosh *et al.* 2021). One of the longest running and densest large-scale soil moisture monitoring networks in the US, and in the world, is the Oklahoma Mesonet (McPherson *et al.* 2007; Ochsner *et al.* 2013). Oklahoma is also consistently among the top 10 states in the US for wildfire risk (III 2021); accordingly, data from Oklahoma have proven valuable for understanding soil moisture–wildfire relationships.

A striking example of the connections between soil moisture, fuel bed properties, and wildfire comes from the Marena, Oklahoma, *In Situ* Sensor Testbed (MOISST) located in north-central Oklahoma. The MOISST site was established in 2010 to compare *in situ* soil moisture sensing technologies (Cosh *et al.* 2016) and measure vegetation dynamics in tallgrass prairie (PhenoCam 2021), with fuel bed properties repeatedly measured at and around the site (Sharma *et al.* 2018). PhenoCam images collected at the site showed markedly different vegetation conditions during August of 2012 and 2013 (Fig. 4). Drought conditions for May–July 2012 resulted in a fuel moisture content for mixed live and dead fuels of only 27% in early August when the photo in Fig. 4*a* was taken. The severity of the drought was reflected in the measured soil moisture, expressed as fraction of available water capacity (FAW). FAW is a calculated measure of plant-available water based on measured volumetric water content and the available water capacity of the soil (Krueger *et al.* 2015), and it can be determined for any landscape (e.g. grassland, forest, and cropland) for which these variables are known. It is defined as the ratio of measured plant available water to the maximum plant available water capacity of the soil, and it typically ranges from 0 (no plant available water) to 1 (maximum plant available water).

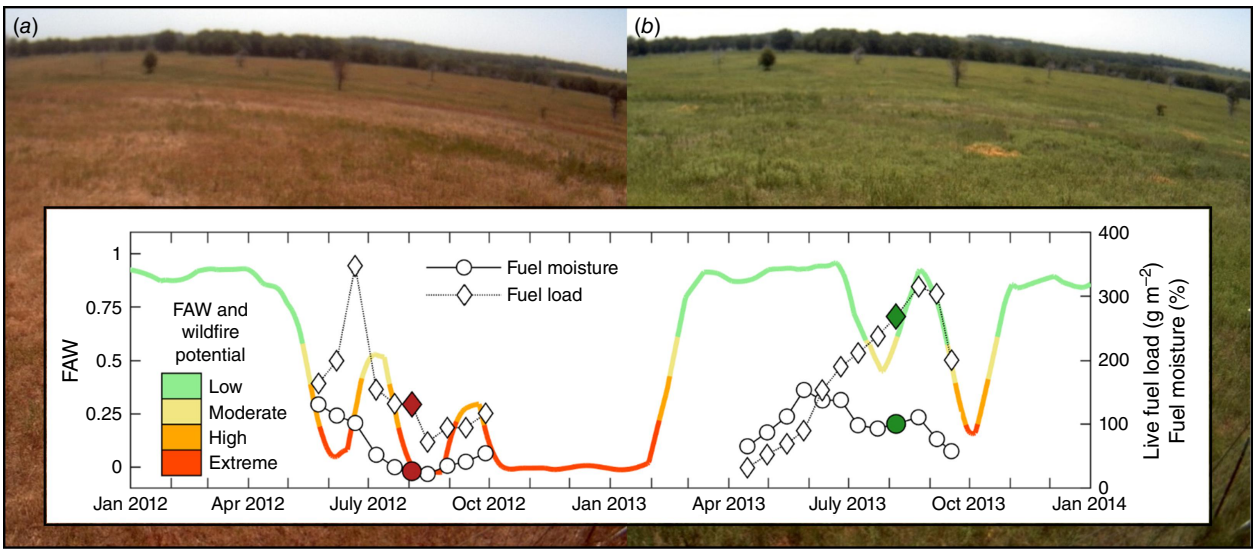


Fig. 4. PhenoCam images collected over grassland near Stillwater, Oklahoma on (*a*) 2 August 2012 and (*b*) 6 August 2013 show the influence of soil moisture on vegetation, and by extension, fire danger. The graph shows the measured fraction of available water capacity (FAW) at the image location, with colours indicating relative wildfire danger and solid circles and diamonds representing the mixed (live + dead) fuel moisture and fuel load, respectively, on days the images were collected. Photo credits: University of New Hampshire Phenocam Network (adapted from Levi *et al.* 2019).

During May–July 2012, FAW averaged only 0.23 (i.e. plant available water was at 23% of its possible maximum), levels indicative of severe drought (Sridhar *et al.* 2008). In contrast, FAW averaged 0.82 (i.e. 82% of possible maximum) over the same period in 2013, which corresponded with green vegetation in August 2013 (Fig. 4*b*) and a mixed fuel moisture content of 101%. The low fuel moisture contents in August 2012 contributed to extreme wildfire danger and the devastating Freedom Hill Fire, which ignited approximately 80 km east of the MOISST site the same day the photo was taken. This fire burned nearly 24 000 ha of mostly shrubland and forest over a 2-week period, destroyed more than 300 homes, and resulted in Federal Emergency Management Agency assistance claims totalling more than US$7 million.

The qualitative soil moisture–fuel bed relationships that are clear in Fig. 4, and may be intuitive to fire managers, have been described in detail by recent research based on *in situ* soil moisture measurements. The soil moisture–fuel moisture relationship was quantified for various shrub species in Italy by Pellizzaro *et al.* (2007), who found that soil moisture was a better predictor of live fuel moisture than weather variables or weather-derived drought indices. Their finding was corroborated by Qi *et al.* (2012), who found that soil moisture explained 66% of the variability in live fuel moisture for oak and sagebrush in northern Utah, and soil moisture was more strongly correlated with live fuel moisture than were remotely sensed vegetation indices. Similar linear relationships between soil moisture and fuel moisture have also been reported for grassland fuels in South Africa (McGranahan *et al.* 2016).

These findings have been corroborated by a series of studies in Oklahoma, the key results of which are summarised in Fig. 5. Sharma *et al.* (2021), using data from a grassland field study close to the MOISST site, reported that when soil moisture was plentiful (FAW values of at least 0.59), mixed fuel moisture was not related to soil moisture, but it declined as FAW decreased below this threshold. When FAW dropped below 0.40, the transpiration and growth rate of grassland live fuels declined, reflecting the intensification of drought stress (Krueger *et al.* 2021). When FAW declined below 0.36, the greenness of the vegetation, as indicated by the normalised difference vegetation index (NDVI), began to decrease (Sharma *et al.* 2021) (Fig. 5). At a still lower FAW threshold of 0.30, the transition of live fuel to dead (i.e. curing rate) increased rapidly, from near $0\,g\,m^{-2}\,day^{-1}$ for FAW $>$ 0.30 to more than $10\,g\,m^{-2}\,day^{-1}$ as FAW approached 0.20 (Sharma *et al.* 2021). This drought-induced curing is vividly depicted in Fig. 4, with extremely low soil moisture corresponding with vegetation that was almost completely cured by early August 2012, while little curing had occurred by the same time in 2013 when soil moisture was plentiful. A perhaps subtler distinction in fuel bed characteristics between these years is that the live fuel load in 2013 was more than double that in 2012, which portended potentially high wildfire activity if dry and windy conditions prevailed during the subsequent dormant season. These findings offer a physical explanation for the observed dependence of growing season wildfire size and probability on soil moisture conditions (Fig. 5).

In a different study that used *in situ* soil moisture data from the entire state of Oklahoma, Krueger *et al.* (2015)

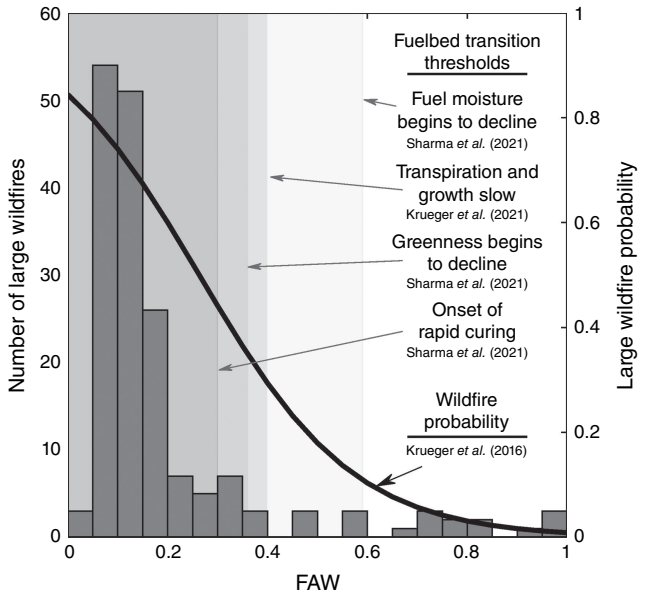


Fig. 5. Frequency distribution (histograms) and probabilistic relationship (solid black curve) between fraction of available water capacity (FAW) and large growing-season wildfires in Oklahoma from 2000 to 2012, adapted from Krueger *et al.* (2015) and Krueger *et al.* (2016). Subsequent research provided physical explanations and thresholds for empirical soil moisture–wildfire links (Sharma *et al.* 2021; Krueger *et al.* 2021). These thresholds describe how live grassland fuels transition to dead fuels as soil moisture declines, beginning with a drop in live fuel moisture (FAW = 0.59) followed by decreased transpiration and growth (FAW = 0.40). Next, vegetative greenness declines (FAW = 0.36), which culminates in rapid fuel curing as soil moisture conditions continue to deteriorate (FAW = 0.30).

showed that 90% of large growing season wildfires across all Oklahoma landscapes (forest, shrubland, and grassland) occurred for FAW < 0.40, which matches the threshold for transpiration reduction due to moisture stress in grassland vegetation (Krueger *et al.* 2021). These soil moisture–wildfire relationships were further described using probabilistic models in a subsequent study (Krueger *et al.* 2016). When plant available soil moisture was near its maximum, the probability of a large growing season wildfire across all Oklahoma landscapes was near zero even when temperature, wind speed, and relative humidity conditions were ripe for wildfires (Fig. 4 and Krueger *et al.* 2016). As FAW decreased to 0.59, the soil moisture threshold below which grassland fuel moisture decreased, wildfire probability increased to 0.10, and for a FAW value of 0.30, the threshold for rapid fuel curing, wildfire probability more than tripled to 0.44 (Fig. 5). These results suggest that soil moisture and weather conditions work in concert to support high growing season wildfire probability. Low soil moisture is associated with decreased fuel moisture and accelerated curing, while high temperatures, low relative humidity, and high wind speed facilitate fire ignition and spread.

When vegetation is dormant, however, current FAW levels were not a strong predictor of the probability of large wildfires in Oklahoma (Krueger *et al.* 2016), which may be related to the fact that dead fuel moisture content across most Oklahoma landscapes is not strongly dependent on soil moisture (Sharma *et al.* 2021). However, dormant season wildfire probability was increased by high soil moisture during the previous growing season. For example, when FAW during the growing season was at least 0.40, the probability of a large wildfire during the subsequent dormant season was approximately double compared with growing season FAW values near 0.20 (Krueger *et al.* 2016). Vegetation productivity, at least for Oklahoma grasslands, is maximised when FAW is >0.40 (Krueger *et al.* 2021), contributing to increased fine fuel loads in the subsequent dormant season.

Although there is a lack of evidence for soil moisture effects on dead fuel moisture in grasslands, *in situ* measurements from a diverse array of sites around the world reveal important links between soil moisture and dead fuel moisture for surface fuels in forests. In Australia, the influence of soil moisture on the fuel moisture content of fine dead fuels (i.e. leaf litter) was observed in plantations of Monterey pine approximately three decades ago (Pook and Gill 1993). The fuel moisture content for the pine needle litter on the surface was positively correlated with measured soil moisture in the 0–40 cm soil layer, and the correlation was stronger for unthinned and un-pruned stands ($r = 0.91$) than in thinned and pruned stands ($r = 0.45$). The fuel moisture content of the surface leaf litter was predicted more accurately when soil moisture data were included along with temperature and humidity data in a multiple regression model compared to a similar model without soil moisture data.

More recently in Australia, *in situ* soil moisture measurements have been linked to the fuel moisture content of the surface and subsurface litter layer under various eucalyptus species (Zhao *et al.* 2021). A follow-up experiment showed that dry soil had a limited influence on the fuel moisture content of the litter, primarily through vapour flow between the soil and the litter (Zhao *et al.* 2022). In contrast, wet soil had a stronger influence on litter moisture content, with evidence for both vapour and capillary flow between the soil and the litter. Similarly, *in situ* measurements from forested sites in the foothills of the Sierra Nevada in central California showed that soil moisture had a stronger influence than any other environmental or meteorological factor on the fuel moisture of 10-h fuels (6–25 mm diameter dead fuels) for wet soil conditions (Rakhmatulina *et al.* 2021). A dominant influence of soil moisture on the moisture content of dead fine fuels was also documented through *in situ* measurements in Korean pine and Scots pine stands in northeastern China (Masinda *et al.* 2021). These reports of the connection between moisture of mineral soils and that of overlying organic layers corroborate previous studies correlating soil moisture measurements with moisture codes from

the CFFDRS. For example, correlation coefficients of 0.6–0.8 were reported between measured soil moisture and soil moisture estimated from the DC index in Canadian forests (D'Orangeville *et al.* 2016).

These findings from diverse ecosystems and geographies highlight the dependencies of fuel loads, fuel moisture content, and wildfire probability on soil moisture. They have also laid the groundwork for a new generation of wildfire danger assessment tools that use *in situ* soil moisture information. However, even with expanding national- and regional-scale soil moisture monitoring networks, using *in situ* data for wildfire danger monitoring and management decisions is still constrained by the limited number of measurement sites in some locations, e.g. boreal forest, and most of South America, Africa, and Australia (Dorigo *et al.* 2021). Because soil moisture can vary greatly across even small distances (Famiglietti *et al.* 2008), point measurements of soil moisture are not necessarily representative of soil moisture at the landscape scale (Fig. 3). Finally, these datasets are often of limited duration, generally spanning less than 20 years (Cosh *et al.* 2021; Dorigo *et al.* 2021), which can make it challenging to use them for soil moisture–wildfire modelling. Therefore, there is a clear need for supplemental strategies for quantifying soil moisture, which include remotely sensed and modelled soil moisture information.

Remotely sensed soil moisture

Remote sensing technology has advanced rapidly since the first photograph of Earth was taken from space in 1946. Since that time, improvements in sensor fidelity, satellite and rocket launch technology, data storage, and aperture development have enabled many new capabilities, including near-real-time operations related to earth sciences and hydrology (McCabe *et al.* 2017). The ability to characterise the land surface using strategic regions of the electromagnetic spectrum has resulted in opportunities to remotely monitor and assess near-surface soil moisture and vegetation dynamics (Kumar *et al.* 2020; Mladenova *et al.* 2020), which are key to understanding the risks and impacts of wildfires. With the advent of refined satellite-based microwave sensors such as the European Space Agency's Soil Moisture Ocean Salinity (SMOS) mission, which launched in 2009 (Kerr *et al.* 2010), and NASA's Soil Moisture Active-Passive (SMAP) mission (Entekhabi *et al.* 2010), which launched in 2015, evidence is beginning to emerge that satellite-based soil moisture data can provide value for understanding and predicting wildfire danger in many ecosystems (O *et al.* 2020).

Remotely-sensed soil moisture data have proven useful for assessing fuel bed properties including biomass accumulation (i.e. fuel production) and fuel moisture content. For example, in southern France, live fuel moisture measurements for Mediterranean shrub species were significantly correlated with the preceding 15-day average remotely sensed soil moisture from the European Space Agency's Climate Change Initiative Soil Moisture dataset (ESA CCI SM, formerly known as ESV SM) (Fan *et al.* 2018). A subsequent study used soil moisture data from SMAP to estimate live fuel moisture of chamise at 12 chaparral sites in southern California (Jia *et al.* 2019). At those sites, a statistical model using weighted, accumulative soil moisture and growing degree days outperformed models using vegetation optical depth or other optical indices. There is also some evidence that remotely sensed soil moisture might be useful for estimating dead fuel moisture. Burapapol and Nagasawa (2016) reported that remotely sensed soil moisture based on Landsat and MODIS was closely linked with fuel moisture of dead leaves in dipterocarp and deciduous forests in Thailand. Soil moisture based on microwave remote sensing may be preferable to optical reflectance indices commonly used to characterise fuel moisture [see reviews by Gale *et al.* (2021), Yebra *et al.* (2013), and Arroyo *et al.* (2008)] because microwave sensors are less prone to disturbances from unfavourable weather (e.g. clouds) and because soil moisture is physiologically linked to plant processes (Nolan *et al.* 2020).

The results of the above regional studies (Fan *et al.* 2018; Jia *et al.* 2019) were supported by a nationwide analysis of the ESA CCI SM data and live fuel moisture at >1000 sites across the contiguous US (Lu and Wei 2021). That analysis spanned numerous vegetation types and climate zones and revealed that the correlations between soil moisture and live fuel moisture were typically strongest when soil moisture was measured 10–50 days in advance. Important vegetation types showing a relatively high sensitivity to soil moisture included pine, red cedar, sagebrush, oak, manzanita, chamise, mesquite, and juniper. The SMAP Level-4 surface and root zone soil moisture products, which result from assimilation of SMAP observations into a land surface model, and *in situ* soil moisture measurements at selected sites both showed somewhat stronger correlations with live fuel moisture than did the ESA CCI SM data.

The links between remotely sensed soil moisture data and fuel bed characteristics make those data useful for assessing wildfire danger. For example, positive soil moisture anomalies observed by Earth Resources Satellite 1 and 2 corresponded with a lower burned area of forest fires in the boreal forest of Siberia (Bartsch *et al.* 2009). Furthermore, extreme fire events in this region were more closely associated with remotely sensed soil moisture [AMSR-E (Njoku *et al.* 2003)] than precipitation anomalies or fire danger indices (Forkel *et al.* 2012). More recently, SMOS observations over boreal forest areas of Canada revealed that wildfires occurred more frequently in anomalously low soil moisture conditions (Ambadan *et al.* 2020). At more southerly latitudes, models using SMOS-derived soil moisture, in conjunction with temperature and site specific variables, such as land cover type, explained 68% of variability of

maximum fire area burned on the Iberian Peninsula (Chaparro *et al.* 2016). The inclusion of SMAP soil moisture observations increased skill in predicting wildfire occurrence in the western US relative to the use of vapour pressure deficit alone, particularly in grasslands (Rigden *et al.* 2020).

Because current soil moisture conditions can influence future fuel moisture and fuel load, soil moisture observations may be particularly helpful for forecasting wildfire danger. In Australia and California, for example, Sazib *et al.* (2022) found that soil moisture from SMAP was negatively correlated with wildfires at 1–2 month lead times in moist regions where fuels are typically plentiful, and it was positively correlated with wildfires in drier regions where fuel is scarce. These trends were attributed to a decrease in moisture of surface fuels in moist regions and increased biomass accumulation in dry regions. In an analysis that spanned the globe, O *et al.* (2020) found that soil moisture from the Essential Climate Variable Soil Moisture (ECV-SM) project was an important early predictor of wildfires. They reported that, in arid regions, positive soil moisture anomalies corresponded with increased biomass accumulation followed by wildfire outbreaks at lead times of 5 months. In humid regions, negative soil moisture anomalies were related to wildfires at lead times of 4 months, presumably because of decreased moisture of surface fuels. Likewise, soil moisture inferred from NASA's Gravity Recovery and Climate Experiment (GRACE) mission was often positively correlated with wildfire occurrence in herbaceous vegetation, shrublands, and forests at seasonal lead times, indicating that a wetter pre-fire-season can lead to increased plant (i.e. fuel) production in these landscapes (Jensen *et al.* 2018).

The large spatial extent of remote sensing datasets provides natural opportunities to explore how soil moisture–wildfire relationships vary across different land cover types. Schaefer and Magi (2019) used satellite-based fire count data from NASA (Giglio *et al.* 2018), land-use and land cover maps (Hurtt *et al.* 2020), and the ESA CCI SM product (Dorigo *et al.* 2017) with a biome map (Levavasseur *et al.* 2012) to study how fires behave relative to soil moisture variability within land cover types and across biomes. They found that the fire–productivity curve shape, which describes resource and climate limits surrounding a zone of optimal fire conditions (Krawchuk and Moritz 2011), was captured within the phase-space of fire and soil moisture. Fire counts were generally greatest when remotely-sensed average monthly soil moisture was relatively low, often around $0.1\ m^3\ m^{-3}$. At lower soil moisture levels, the average number of fires decreased with decreasing soil moisture, presumably because of resource limitations (i.e. decreasing fuel availability). At higher soil moisture levels, the average number of fires decreased with increasing soil moisture, likely due to increased fuel moisture contents. But the shape of the fire–soil moisture curve differed as a function of biomes and land cover types. For example, the occurrence of fires in boreal forests (Fig. 6*b*), which have a shallower rooting depth than forests in other biomes (Fan *et al.* 2017), relates to soil moisture availability in a way that is similar to grasslands (Fig. 6*a*), which also have shallow root depths. This apparent effect of root depth on the sensitivity of fire occurrence to soil moisture under different biomes reinforces the value of soil moisture as a predictor of fire danger. Consistent with these results, Forkel *et al.* (2017) showed that across the world, biophysical models of fire activity (e.g. Rabin *et al.* 2015) performed better when remotely sensed soil moisture (and moisture state in general) was considered.

These global scale analyses are possible because, unlike *in situ* soil moisture measurements, remotely sensed measurements provide data on soil moisture conditions across large spatial domains. However, remotely sensed measurements typically represent soil moisture conditions in only the top few centimetres of the soil (Abbaszadeh *et al.* 2021) and have lower temporal resolution compared to *in situ* networks. Furthermore, remotely sensed soil moisture measurements have historically shown a limited ability to accurately monitor soil moisture conditions where a dense vegetative canopy is present (Djamai *et al.* 2015; Dorigo *et al.* 2015), although recent advances provide unequivocal evidence that remote sensing measurements are sensitive to soil moisture under forest canopies (Colliander *et al.* 2020; Ayres *et al.* 2021). There is a clear need to focus future research on remotely sensed soil moisture–wildfire relationships at higher spatial resolution and for specific land cover types. Such studies may enhance the relevance of satellite-based soil moisture data to fire managers. These types of studies may also be particularly well-suited for linking with model-based approaches as described in the next section.

Modelled soil moisture

Given the historical lack of *in situ* and satellite measurements, proxies and estimates of soil moisture conditions have long been used in the context of wildfire danger. Approaches have ranged from drought indices based on simplistic soil water balance models (e.g. Palmer 1965; Keetch and Byram 1968; Mount 1972), to actual soil moisture values simulated using more complex process-based models (Carrega 1991; Holden *et al.* 2019), to hybrid approaches that incorporate measured soil moisture data into plant growth models (Krueger *et al.* 2021). These approaches have been applied across widely-varying time horizons, with some showing the possibility to facilitate predictions of soil moisture, and subsequently wildfire, for time frames potentially spanning decades (Chikamoto *et al.* 2015). The KBDI (Keetch and Byram 1968), in particular, has been used extensively to address the challenges of representing moisture deficits and their influence of wildfire

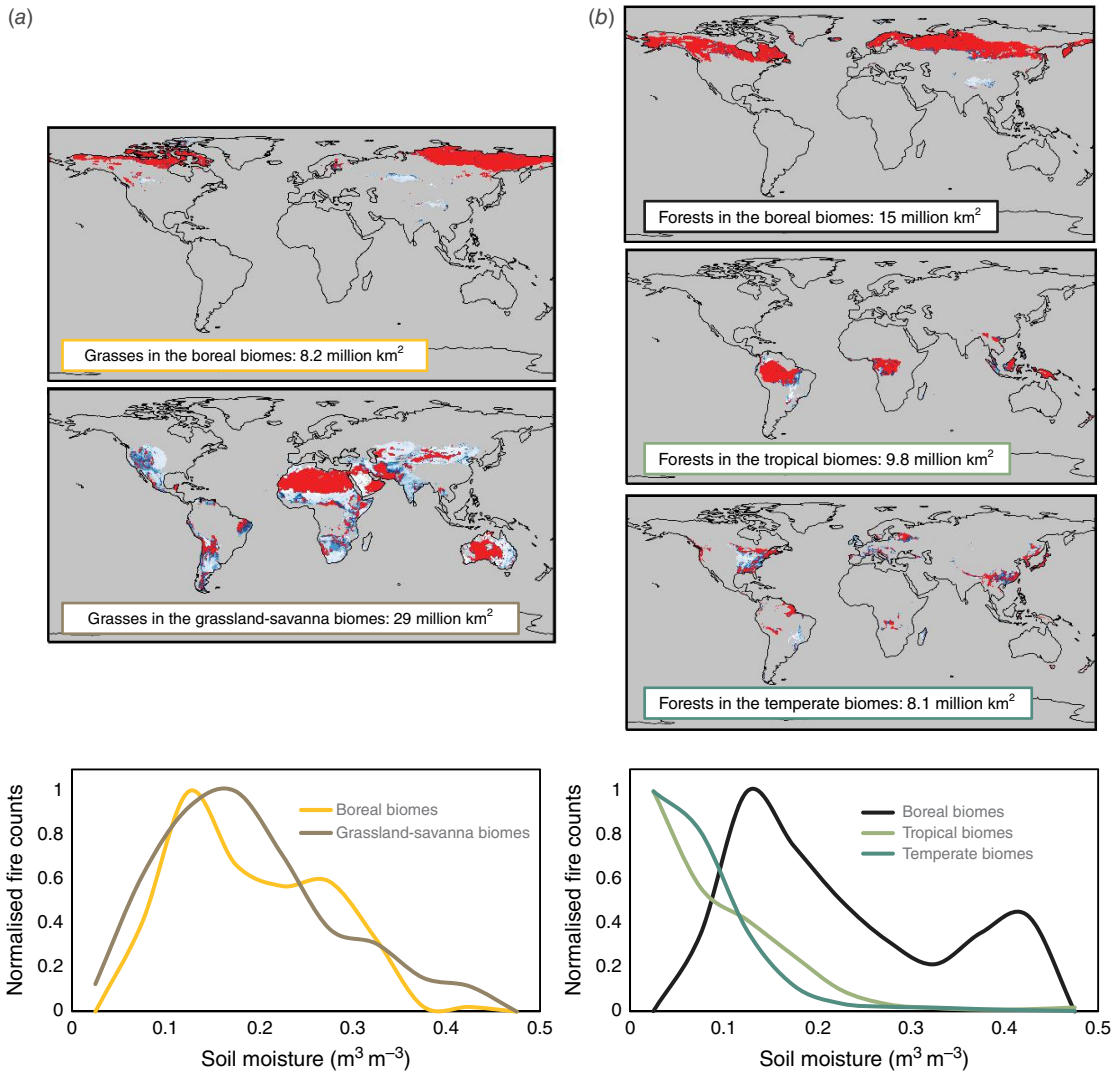


Fig. 6. Fire–soil moisture curves for different land cover types, (*a*) grasses and (*b*) forests, in different biomes (boreal, grassland-savanna, temperate, and tropical) showing the resource and climate limits at low and high soil moisture values. The curves are derived from monthly-averaged soil moisture for the 0–5 cm soil layer from the European Space Agency Climate Change Initiative (version 4.2) and monthly fire counts from MODIS Collection 6. The area analysed in each biome is shown on the maps in red, where this shading denotes where greater than 75% of the grid cell is a single land cover type. The area covered by grasses in temperate and tropical biomes (about 5.7 million km^2, or 13% of global grasses) and area covered by forests in grassland–savanna biomes (about 4.0 million km^2, or 11% of global forests) were excluded due to a weaker statistical signal. The shape of the fire–soil moisture curves varies across land cover types and across biomes, suggesting that soil moisture may be a viable predictor of biome-scale fire danger for different land cover types (adapted from Schaefer and Magi 2019).

danger. For example, KBDI has been used in the McArthur Mark 5 forest fire danger index (Holgate *et al.* 2017), the Fosberg fire weather index (Goodrick 2002), and the NFDRS (Burgan 1988).

Developed in the southern United States in the 1960s to predict moisture deficits in organic and mineral soil layers, KBDI is a unitless index ranging from 0 to 800. The KBDI calculation attempts to address important physical processes such as canopy interception of precipitation and the effects of biomass on rates of soil water loss. However, it has significant limitations. For example, it does not include humidity, wind, or radiation in its estimate of soil water loss. The model also uses climatological average precipitation as a surrogate for both leaf area and canopy interception, based on the assumption that wetter sites support more vegetation. Finally, KBDI does not consider variability in soil properties, instead assuming a water holding capacity of 8 inches (19.32 cm) for all soils. Given these limitations, it is not surprising that *in situ* and remotely sensed soil moisture are more strongly related to wildfires than KBDI

in grasslands in the western US (Rigden *et al.* 2020) and across grasslands, shrublands, and forests in Oklahoma (Krueger *et al.* 2017).

When compared to *in situ* soil moisture observations in Australia, KBDI showed a large wet bias relative to measurements in the 0–30 and 0–90 cm soil layers, had correlations with measured soil moisture that vary widely across climate zones and were sometimes negative, tended to dry down too slowly after wet periods, and performed more poorly than simulations from a physically-based land surface model (Holgate *et al.* 2017; Vinodkumar *et al.* 2017). Although KBDI can be calibrated to represent temporal variations in live fuel moisture at specific sites, it is unable to accurately represent spatial variations in live fuel moisture, and thus is not recommended for use in operational fire management (Ruffault *et al.* 2018). Replacing drought indices like KBDI with more robust soil moisture models has been noted as a priority for improving fire danger rating in the US (Jolly 2018) and is well underway in Australia (Vinodkumar and Dharssi 2019; Vinodkumar *et al.* 2021).

Process-based models link vegetation growth and functioning with soil properties and climate information and are sometimes referred to as land surface models or soil–vegetation–atmosphere-transfer models (Moran *et al.* 2004). These models can be particularly useful because they represent plant physiological processes, allowing vegetation to be modelled accurately over space and time, and thus capture many of the vegetation fuel attributes that are relevant for fire spread models (Landsberg *et al.* 2003). One example is the TOPOFIRE model, which was recently developed to provide high spatial resolution daily estimates of soil moisture, fuel moisture, and fire danger data and maps for the conterminous US (Holden *et al.* 2019). Another recent example is the modelling system developed by the Australian Bureau of Meteorology based on the Joint UK Land Environment Simulator (JULES) called the JULES-based Australian Soil Moisture Information (JASMIN) system (Vinodkumar and Dharssi 2019). The JASMIN system was specifically designed for application in operational fire prediction and risk management.

Such models hold promise, not only for wildfire decision support, but also for revealing a new foundational understanding of soil moisture–wildfire relationships. For example, modelled soil moisture values from the US National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA CPC) (e.g. Fig. 7) have been used with remotely sensed active fire data to understand global patterns in the constraints of fine fuel loads and fuel moisture on wildfire occurrence (Krawchuk and Moritz 2011). The NOAA CPC soil moisture (NOAA CPC 2022) is modelled using a water balance approach based on measured temperature and precipitation, as well as calibrated models that account for precipitation lost to streamflow and subsurface drainage (Huang *et al.* 1996). In another recent study, simulations from a physics-based model showed the close interaction of soil moisture with the fuel moisture of the litter layer in shrublands, woodlands, and forests (Zhao *et al.* 2021).

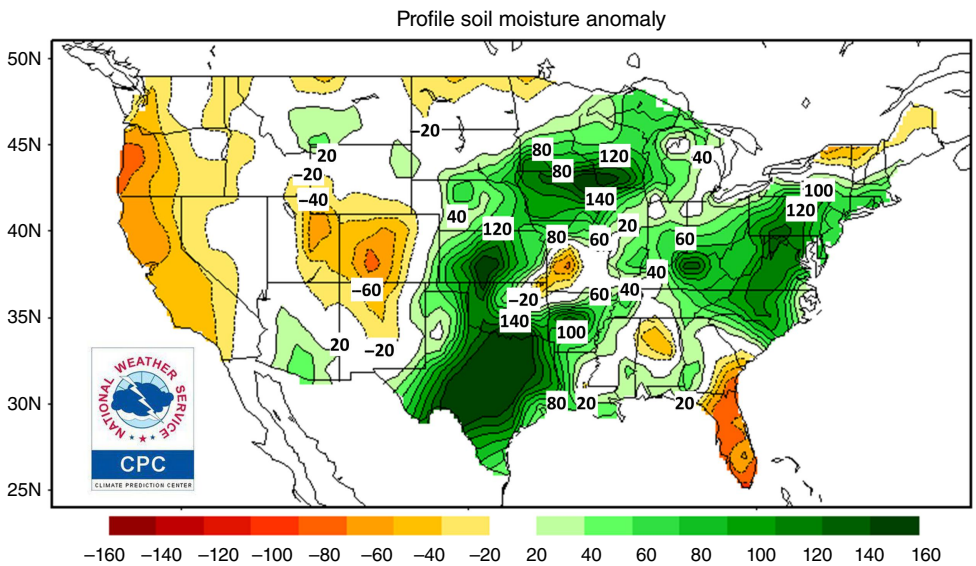


Fig. 7. Modelled soil moisture in top 160 cm of soil profile as reported by the National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA CPC) for 1 November 2018, 1 week prior to the Camp Fire in northern California. Daily and monthly soil moisture maps for the conterminous United States, as well as monthly soil moisture maps for the world, are produced by NOAA CPC. As with measured surface soil moisture in Fig. 1, this map indicates dry soil conditions throughout the soil profile in northern California prior to the Camp Fire (image: NOAA CPC).

Likewise, soil moisture modelled using TOPOFIRE has been shown to be a better predictor of canopy water content across the western US than is atmospheric vapour pressure deficit (Lyons *et al.* 2021). In fact, gridded 5-km resolution live fuel moisture estimates in grasslands, shrublands, and forests have been generated for Australia based on soil moisture values simulated with the JASMIN system (Vinodkumar *et al.* 2021). These live fuel moisture predictions used soil moisture as a leading indicator with a 14-day lag period. The 0–35 cm soil layer was determined to be the best layer for live fuel moisture prediction. This is similar to the 0–40 cm layer used for *in situ* soil moisture measurements in several prior wildfire-related studies (Pook and Gill 1993; Krueger *et al.* 2015, 2016, 2017; Sharma *et al.* 2021). There is a clear need for further development and refinement of process-based models specifically designed to capture soil moisture–fuel load–fuel moisture–fire danger relationships and for the application of those models for greater scientific understanding and improved fire danger ratings.

One limitation to process-based modelling approaches relative to simple drought indices is the increased complexity of model inputs and sometimes intensive calibration needs. Necessary inputs typically include gridded data sets for climate conditions, soil properties, and vegetation type and condition. Obtaining these input data at the necessary spatial and temporal scale and resolution can be challenging. For example, soil maps are often compiled at broad spatial scales, often do not cross political boundaries, and sometimes use inconsistent nomenclatures (Zheng *et al.* 1996; Mulder *et al.* 2011). Some critical soil attributes like soil depth and available water capacity can be hard to derive using traditional soil mapping techniques, although this may be overcome by using process-based models in an inverse fashion to estimate soil physical properties from plant growth data (Coops *et al.* 2012). Levi and Bestelmeyer (2018) summarise available spatial soil information datasets for fire modelling in the US and suggest that advances in soil modelling can lead to improved soil property maps and therefore more accurate fire predictions.

There appears to be great potential for hybrid approaches that incorporate *in situ* or satellite soil moisture measurements into process-based models. For example, predictions of grassland fuel loads can be improved by direct insertion of *in situ* soil moisture observations into a simulation model's soil water balance routine (Krueger *et al.* 2021), or a soil moisture model can be improved by assimilating satellite-based soil moisture observations, as demonstrated in Bolten *et al.* (2010). Hybrid approaches have also proven useful when predicting areas of vegetation stress, which may be more prone to wildfires. For example, areas of increased land surface temperature and decreased greenness are likely to be subject to lower vegetation growth and increased stress (Nemani *et al.* 1996). If prolonged, these stresses can result in increased litter fall, increased non-photosynthetic vegetation, and drier soil, which in turn correspond with increased fuel load. Based on this concept, Mildrexler *et al.* (2009) developed a global disturbance index using remotely sensed land surface temperature and greenness and demonstrated that this index could identify areas of broad scale vegetation stress. Waring *et al.* (2011) applied this index over western North America and demonstrated that increases in the area designated as stressed were positively correlated with the areas of increased simulated soil water stress and wildfire. Waring and Coops (2016) then compared simulated soil moisture with satellite derived area burned (Fig. 8*a*). Using a decision tree approach, they identified four seasonal combinations of current and antecedent soil moisture conditions that predicted where forest fires $>1\,km^2$ occurred with 69% accuracy (Fig. 8*b*).

These studies add to the growing body of evidence that an accurate accounting of soil moisture status, either by *in situ* measurements, remote sensing, or modelling, can improve our ability to anticipate when and where wildfires will occur. While soil moisture models can suffer from errors caused by inaccuracies in input data and the model structure, they are appealing because of their capability to incorporate diverse data sources including measured soil moisture and vegetation condition (Fig. 3). Yet as described in the following section, soil moisture information has thus far been largely absent from major fire danger rating systems.

Potential for inclusion of soil moisture information into fire danger rating systems

In this section we explore the potential for integration of soil moisture information into fire danger rating systems. We begin with a brief review of some of the leading fire danger rating systems and how they incorporate weather and other information to estimate fuel bed properties, estimates that could potentially be improved by incorporating soil moisture information.

National fire danger rating systems

Fire danger rating systems integrate inputs representing multiple fire danger factors, often via a model, into one or more qualitative or numerical indices of fire danger. Some systems also model physical characteristics of the fire, such as fire intensity, rate of spread, and flame length. Fire danger rating systems provide assessments of fire danger over broad geographical areas, encompassing up to millions of hectares, and are typically not designed to provide detailed fire danger information at the field scale. Spatially, when calculated over a grid, the fire danger for each grid cell represents the average fire danger at a given time over that cell, assuming homogeneous fuels, weather, and topography within the cell. Such systems are used to

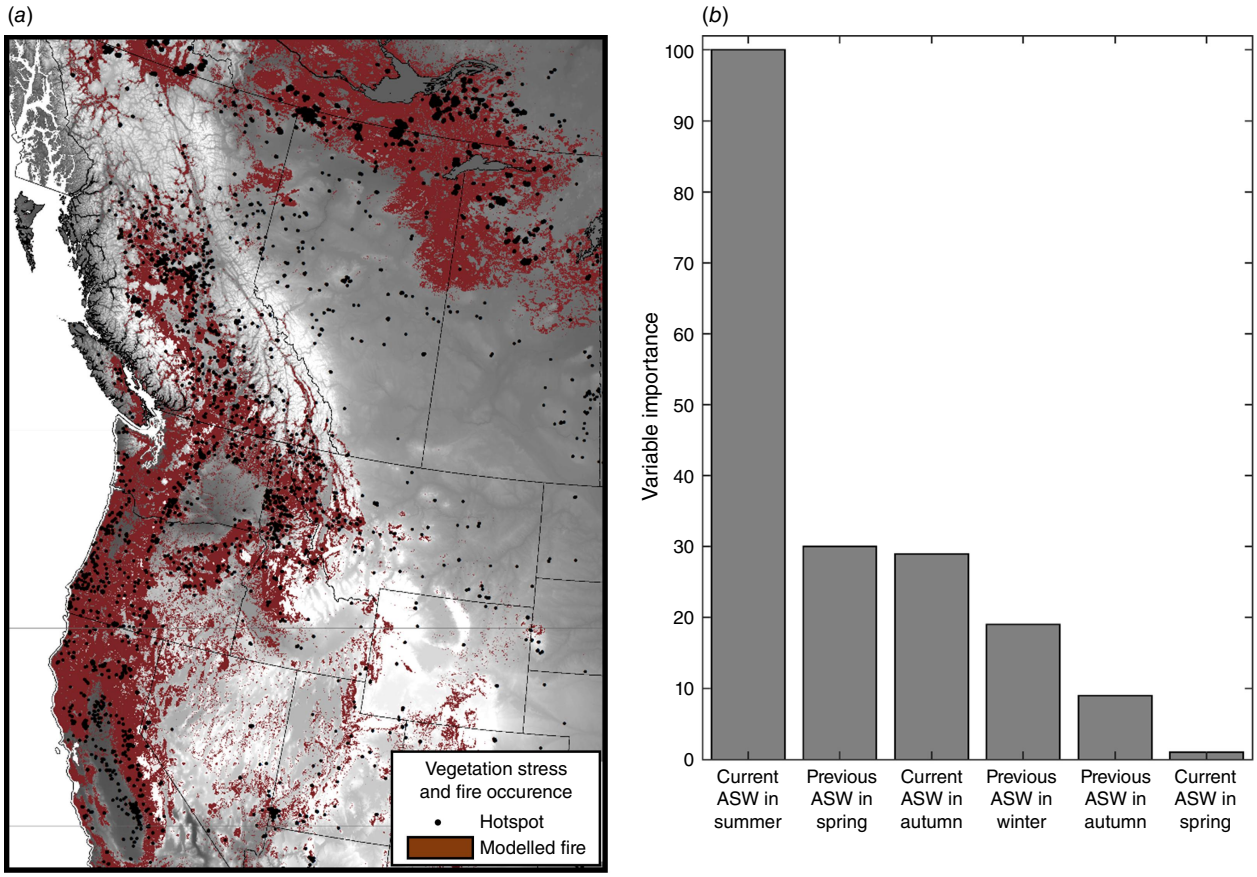


Fig. 8. (*a*) Example of model predictions of wildfires (red) based on available soil water (ASW) for forested portions of western North America in 2004, along with the locations of MODIS active fire hotspots (black dots) for the same period. (*b*) Relative importance of different seasonal functions of available soil water used to predict MODIS active hotspot occurrence of wildfires in 2001, 2004, and 2007 (adapted from Waring and Coops 2016).

provide public warnings, set preparedness levels, provide a good indication of the difficulty of fire suppression over a wide range of conditions, and to help wildfire managers make wise tactical and strategic management decisions (NWCG 2002).

While there are a number of fire danger ratings systems across the world, ranging from national to regional to local scales, it is instructive to look at three national systems that have been widely used for many decades, those of Australia, Canada, and the United States. While a new Australian system (AFDRS) is becoming operational in 2022 (AFAC 2022), the previous system consisted of two fire danger indices, each with six fire danger categories: the McArthur Forest Fire Danger Index (FFDI) and the McArthur Grassland Fire Danger Index (GFDI) (McArthur 1966, 1967; Noble *et al.* 1980). The FFDI and GFDI each required temperature, relative humidity, wind speed, and rainfall as weather inputs. The FFDI assumed a standard eucalypt forest in its calculations, while the GFDI assumed a standard grassland. For each index, the fuel type and total fuel load (live + dead) were constant. For the FFDI, fuel availability (drought factor) was calculated from soil moisture deficit, time since last rain, and rainfall amount (Matthews 2009). The soil moisture deficit used KBDI or the Soil Dryness Index (SDI, Mount 1972). For the GFDI, the degree of curing was also an input, which was typically estimated by ground-based visual observations, satellite imagery, or a combination of both.

The CFFDRS in Canada (Stocks *et al.* 1989) has been in its current form since 1992, with a series of improvements planned for release beginning in 2025 (CFSFDG 2021). It consists of two major subsystems, the Fire Weather Index (FWI) System (Van Wagner 1987) and the Fire Behaviour Prediction (FBP) System (FCFDG 1992). The FWI System uses a standard jack pine forest and consists of six components: three moisture indices (fine fuel moisture code, duff moisture code, and drought code) that represent three organic layers at or beneath the forest floor (Fig. 2), and three fire danger indices, including FWI itself. Weather inputs are temperature, relative humidity, wind speed, and rainfall. The FBP System consists of 16 available fuel models, including a grass model that requires degree of curing as an input (de Groot 1993). The system incorporates three outputs from the FWI System and uses topography in its

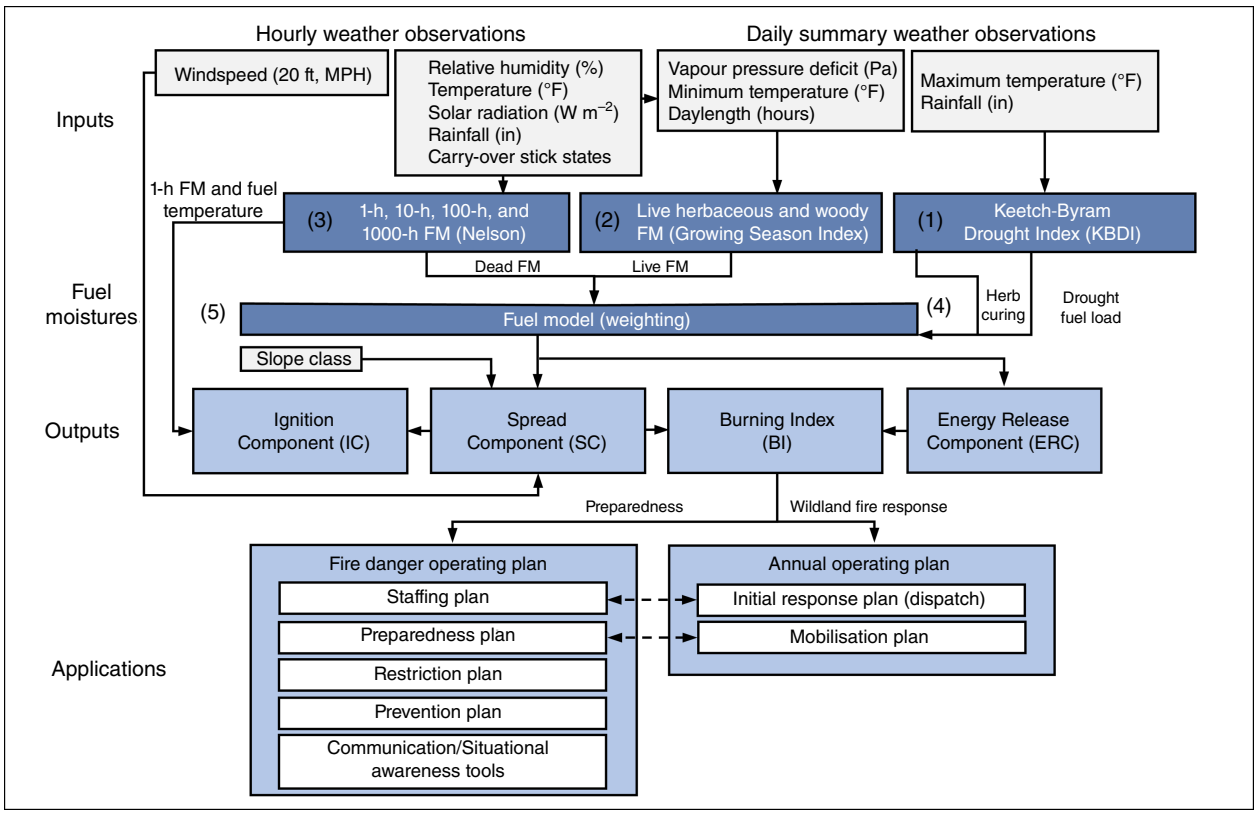


Fig. 9. Structure of the United States National Fire Danger Rating System NFDRS2016 adapted from Jolly (2018). Possible uses of soil moisture information in NFDRS2016 are numbered and in dark blue boxes, and potential downstream effects of the inclusion of soil moisture information are in light blue boxes. These potential uses include (1) supplementing or replacing KBDI, (2) live fuel moisture modelling, (3) dead fuel moisture modelling, (4) estimating herbaceous curing, and (5) fuel load modelling.

calculations. Foliar (live) fuel moisture is modelled using elevation, geographical location, and date; thus, the foliar moisture remains the same on a given date and location from year to year. Outputs of the FBP System include physical characteristics of the wildfire (e.g. rate of spread and fire intensity).

The NFDRS of the United States was originally released in 1972 (Deeming *et al.* 1972) with major updates in 1978 (Bradshaw *et al.* 1983) and 1988 (Burgan 1988). The latest version (NFDRS2016) includes five standard fuel models, reduced from 20 in the 1978 and 1988 versions (Jolly 2018). As with the two former versions, NFDRS2016 separately calculates live and dead fuel moisture as well as the dynamic fuel load transfer (i.e. curing or green-up) between 1-h dead (<6 mm diameter dead fuels) and live herbaceous fuels (Fig. 9). The live fuel moisture and dynamic fuel load transfer calculations in NFDRS2016 are a function of Growing Season Index (GSI), which is a function of temperature, relative humidity, and photoperiod (Jolly *et al.* 2005). As with the 1988 NFDRS, the new system uses KBDI as a drought surrogate to linearly increase the dead fuel loads when KBDI increases above a threshold of 100. The inputs to NFDRS2016 are temperature, relative humidity, rainfall, wind speed, solar radiation, and photoperiod (based on latitude and day of year). The outputs from NFDRS consist of four components describing the wildfire danger: Spread Component, Energy Release Component, Burning Index, and Ignition Component.

Potential pathways for inclusion of soil moisture information

Within fire danger rating systems like those described above, there are at least five potential uses for soil moisture information: (1) as a replacement or supplement for drought indices; (2) as an input for live fuel moisture modelling; (3) as an input for dead fuel moisture modelling; (4) as an input to estimate curing for herbaceous fuels; and as (5) as an input for estimation of fuel loads. We now briefly discuss each of these potential uses within the context of NFDRS2016 (Fig. 9), which provides a representative example for how soil moisture could potentially be used in fire danger rating systems worldwide.

First, soil moisture measurements or simulations from process-based models could be used to replace drought indices in fire danger rating systems. Moisture indices that can represent soil moisture have been used in fire danger rating systems across the world, including KBDI and SDI in

the Australian FFDI system, the drought code in the Canadian FWI System, and KBDI in the US NFDRS system. A growing body of evidence indicates that new sources of soil moisture information are useful for predicting wildfire danger across a variety of landscapes including grasslands, shrublands, and temperate and boreal forests, and soil moisture information can be more closely related to wildfire danger than traditional drought indices (e.g. Bartsch *et al.* 2009; Forkel *et al.* 2012; Krueger *et al.* 2015; Chaparro *et al.* 2016; Schaefer and Magi 2019; Ambadan *et al.* 2020; Rigden *et al.* 2020). For example, in Oklahoma, *in situ* soil moisture measurements provided an average of 10 days earlier warning than KBDI for the largest growing-season wildfires (Krueger *et al.* 2017). For one of the largest wildfires in that study, the Chester Fire, soil moisture reached dangerously low levels (FAW ≤ 0.2) more than 3 weeks before the actual fire, while KBDI never reached levels considered dangerous (≥ 600), thus providing no advance warning at all. When FAW is < 0.2, as it was when leading up to the Chester Fire, grassland curing rates of $\sim 13\,g\,m^{-2}\,d^{-1}$ can occur (Sharma *et al.* 2021), which could result in the accumulation of $> 270\,g\,m^{-2}$ of dead fuel in 3 weeks, or near 100% curing given typical grassland fuel loads for the region (Krueger *et al.* 2021). Analyses from Australia indicate that SDI is on average more strongly correlated with *in situ* soil moisture measurements than is KBDI, but like KBDI, SDI exhibits slower dry downs than *in situ* soil moisture and greater variation in performance across regions than more advanced process-based models (Holgate *et al.* 2016). These results suggest that soil moisture measurements or simulations from process-based models could effectively supplement or replace drought indices in fire danger rating systems.

A second potential use of soil moisture information is for live fuel moisture modelling in fire danger rating systems. Soil moisture has been shown to be a strong predictor of live fuel moisture in grasslands, (Sharma *et al.* 2021), shrublands (Pellizzaro *et al.* 2007; Qi *et al.* 2012), and forest understory (Bianchi and Defossé 2015). In fact, soil moisture observations have shown stronger correlations with live fuel moisture than drought indices in some Mediterranean shrub species (Pellizzaro *et al.* 2007) and stronger than remotely sensed vegetation indices in Gambel oak and sagebrush (Qi *et al.* 2012). In NFDRS2016, live fuel moisture is estimated using GSI, a simple empirical index for vegetation phenology based on photoperiod, vapour pressure deficit, and air temperature (Jolly *et al.* 2005 and Fig. 8). We are not aware of any peer-reviewed evaluations of the accuracy of live fuel moisture estimates based on GSI, although GSI has shown temporal trends similar to live fuel moisture content in sagebrush and chamise (Jolly 2018). Based on the evidence from the literature, we hypothesise that inclusion of soil moisture information as an additional input variable in the GSI calculation would lead to improved live fuel moisture estimates. Alternatively, live fuel moisture could be directly estimated from soil moisture information as has been successfully demonstrated in Australia (Vinodkumar *et al.* 2021).

Third, soil moisture could also be useful for dead fuel moisture estimation in fire danger rating systems. Soil moisture influences near-surface air temperature and humidity (McKinnon *et al.* 2021), and water movement between the soil and dead surface fuels has been observed in shrubland, eucalypt forests (Zhao *et al.* 2021, 2022), and aspen forests (Samran *et al.* 1995). The NFDRS2016 estimates dead fuel moisture using the Nelson model, which uses temperature, relative humidity, solar radiation, and precipitation as inputs (Nelson 2000). The Nelson model has shown reasonable accuracy in estimating dead fuel moisture with r^2 values of 0.51–0.79 (Carlson *et al.* 2007), but for some landscapes like conifer forests there is evidence that dead fuel moisture models incorporating soil moisture information provide better estimates than those that omit soil moisture information (Pook and Gill 1993; Masinda *et al.* 2021; Rakhmatulina *et al.* 2021). These studies highlight the potential to improve fire danger rating systems by using soil moisture information for estimating dead fuel moisture, particularly for dead surface fuels at forested sites.

Improving representation of the curing of herbaceous fuels is a fourth promising use of soil moisture information. Few studies have directly considered the relationship between soil moisture and curing, but the limited available data suggest a strong relationship between soil moisture conditions and the rate of curing in grasslands (Sharma *et al.* 2021). Likewise, positive correlations between the fuel moisture content and soil moisture content in grasslands have been observed during the curing period from the end of the growing season into mid-winter (McGranahan *et al.* 2016). The degree of curing in herbaceous fuels can be determined through direct measurements, visual estimates, remote sensing, or soil moisture deficit or plant phenology models (Duff *et al.* 2019). For example, in NFDRS2016 the dead herbaceous fuel load transfer is estimated as a function of the degree of curing, which is estimated from the GSI plant phenology model. Unpublished data show a negative relationship between GSI and grass curing ($r^2 = 0.41$) for one site in North Dakota, USA (Jolly 2018), but few, if any, published studies have compared GSI with curing levels measured *in situ*. Given that soil moisture deficits enhance curing (Wittich 2011), soil moisture information could perhaps be used as an additional input for the estimation of GSI and therefore curing, or the curing rate could be directly estimated from soil moisture observations.

A fifth potential use of soil moisture information in fire danger rating systems is for the estimation of fuel loads. Current fire danger rating systems assume a constant fuel load (live + dead) regardless of the differences in weather from one growing season to the next. But loads can vary substantially year-to-year, especially for herbaceous fuels. For example, incorporation of soil moisture observations into a simple, process-based plant growth model can provide

improved predictions of grassland productivity and fuel loads (Krueger *et al.* 2021). Likewise, soil moisture is a significant predictor of live fine fuel loads at guinea grass (*Megathyrsus maximus*) dominated sites in Hawaii (Ellsworth *et al.* 2013). There is also evidence for an important role of soil moisture conditions in regulating the growth rates of shrubland fuels in the Arctic (Myers-Smith *et al.* 2015; Ackerman *et al.* 2017; Martin *et al.* 2017). Accounting for soil moisture effects on fuel production rates could lead to better approaches to represent dynamic fuel loads and could potentially improve the performance of fire danger rating systems.

While research provides evidence for these proposed uses of soil moisture in wildfire danger rating, the supporting studies have been relatively few and often of limited geographic scope. Substantiating research across diverse geographic locations and biomes is essential to support implementation on a large scale. Furthermore, the usefulness of soil moisture information in fire danger rating systems is dependent on the way such information is generated. *In situ* soil moisture measurements can monitor conditions deep into the soil profile rather than just the top few centimetres. Therefore, these *in situ* measurements effectively represent root zone conditions, and they can be located in diverse vegetation types (e.g. grasslands, shrublands, and forests). A main limitation of *in situ* measurements is that each measurement typically represents only a small area and may not adequately reflect heterogeneous soil moisture conditions in the surrounding landscape.

Unlike *in situ* observations, which are lacking in many regions, satellite remote sensing is available globally and can provide useful large-scale estimates of soil moisture conditions. But remotely sensed soil moisture measurements have limited capacity to monitor conditions below the 5-cm depth, reduced accuracy beneath dense forest canopies, and lower temporal resolution as compared to *in situ* measurements. In contrast, simulated soil moisture information from process-based models can represent the entire root zone, can be extended to almost any land cover and land use type, and have flexible spatial and temporal resolution. Yet, the accuracy of these simulated values is limited by the availability and quality of the necessary soil, vegetation, and weather input data and by uncertainties in the model structure and parameters. Another limiting factor is the sometimes large computational requirements for running the simulations. The use of soil moisture information in fire danger rating systems may need to rely on a combination of all three sources to represent the best available information across a range of relevant scales.

Challenges and opportunities

We have described a steadily growing body of evidence indicating the need for and potential benefits of using soil moisture for wildfire danger assessments. While this research is promising, many questions remain. First, and perhaps most important, while the current body of research supports a litany of *potential* uses of soil moisture in fire danger rating systems, the practical benefits of these uses remain largely untested and logistical challenges likely remain. Pioneering efforts in the operational use of soil moisture information in fire danger rating systems include the use of *in situ* soil moisture measurements in OK-FIRE, a weather-based decision support system for wildland fire managers in Oklahoma that produces maps of growing-season wildfire danger, updated every 30 min, based on soil moisture (Oklahoma Mesonet 2021). These maps supplement similar maps based on KBDI for operational fire management decisions. Similarly, the operational use of remotely-sensed soil moisture data is being explored by the Barcelona Expert Centre (BEC), which downscales SMOS soil moisture data to create near-real-time fire risk maps (BEC Team 2018) that are currently used by Barcelona Provincial Council to provide wildfire early warning (Chaparro *et al.* 2016). In Australia, modelled soil moisture values are being used to generate dynamic nationwide live fuel moisture estimates designed for use in operational fire danger ratings (Vinodkumar and Dharssi 2019; Vinodkumar *et al.* 2021). Further research specifically aimed at techniques for incorporating soil moisture into wildfire danger systems is critically needed, as well as evaluation of fire danger ratings with and without soil moisture information.

Other important research needs and opportunities abound in this context. Some key research questions include: (1) What representations of soil moisture (e.g. absolute values, scaled values like FAW, anomalies, and percentiles) are best suited for wildfire danger assessment? (2) What are the soil depths for which moisture conditions are most strongly related to fuel production rates, fuel moisture, and wildfire occurrence and size? (3) How can the various sources of soil moisture information (*in situ*, remotely sensed, modelled, or a combination of these) best be leveraged for improving operational fire danger assessments? (4) How can soil moisture information be used to produce accurate dynamic estimates of live and dead fuel loads in fire danger rating systems? (5) How are soil moisture conditions related to and predictive of wildfire occurrence and severity in organic soil layers, where the soil itself is the fuel (Reardon *et al.* 2007; Rein *et al.* 2008; Prat-Guitart *et al.* 2016; Elmes *et al.* 2018)? (6) How do pre-fire soil moisture conditions influence burn severity, soil heating, and post-fire impacts of both wildfire and prescribed fire across different landscapes? These questions must all be answered in parallel with continued research aimed at refining and expanding *in situ*, remotely sensed, and modelled soil moisture products.

After clearing these scientific hurdles, there remains the further challenge of convincing wildfire professionals of the importance of soil moisture compared with more familiar

wildfire danger metrics. For example, the importance of KBDI has been engrained in generations of wildfire professionals, and it benefits from widespread familiarity and is inherently understood. It is critical that soil moisture be distinguished from this and other drought indicators, or it risks being overlooked as just another drought metric. The challenge for scientists is to formulate soil moisture information into a form that is easily understood and used by fire managers. Once that occurs, use of soil moisture information by fire managers should encourage greater acceptance, which in turn should encourage greater use. If available soil moisture information had been included in operational fire danger rating systems in the US, would it have resulted in earlier warning of extreme fire danger prior to the Camp Fire in 2018? Would it have helped save the lives of any of the 85 people who died from the fire? We do not know. But we know that we now have sufficient soil moisture information and adequate scientific evidence to begin using that information to improve fire danger rating systems around the world. So, let's begin.

References

Abbaszadeh P, Moradkhani H, Gavahi K, Kumar S, Hain C, Zhan X, Duan Q, Peters-Lidard C, Karimiziarani S (2021) High-Resolution SMAP Satellite Soil Moisture Product: Exploring the Opportunities. *Bulletin of the American Meteorological Society* **102**(4), 309–315. doi:10.1175/bams-d-21-0016.1

Ackerman D, Griffin D, Hobbie SE, Finlay JC (2017) Arctic shrub growth trajectories differ across soil moisture levels. *Global Change Biology* **23**(10), 4294–4302. doi:10.1111/gcb.13677

AFAC (2022) Australian Fire Danger Rating System Frequently Asked Questions (FAQs). Australian and New Zealand National Council for fire and emergency services. East Melbourne, Victoria, Australia. Available at https://www.afac.com.au/initiative/afdrs/afdrs-faqs [accessed 5 April 2022]

Ambadan JT, Oja M, Gedalof Z, Berg AA (2020) Satellite-Observed Soil Moisture as an Indicator of Wildfire Risk. *Remote Sensing* **12**(10), 1543. doi:10.3390/rs12101543

Arroyo LA, Pascual C, Manzanera JA (2008) Fire models and methods to map fuel types: The role of remote sensing. *Forest Ecology and Management* **256**(6), 1239–1252. doi:10.1016/j.foreco.2008.06.048

Ayres E, Colliander A, Cosh MH, Roberti JA, Simkin S, Genazzio MA (2021) Validation of SMAP soil moisture at terrestrial National Ecological Observatory Network (NEON) sites show potential for soil moisture retrieval in forested areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **14**, 10903–10918. doi:10.1109/JSTARS.2021.3121206

Bartsch A, Balzter H, George C (2009) The influence of regional surface soil moisture anomalies on forest fires in Siberia observed from satellites. *Environmental Research Letters* **4**(4), 045021. doi:10.1088/1748-9326/4/4/045021

BEC Team (2018) 'SMOS-BEC Land Products Description.' (Barcelona Expert Centre: Barcelona, Spain)

Bianchi LO, Defossé GE (2015) Live fuel moisture content and leaf ignition of forest species in Andean Patagonia, Argentina. *International Journal of Wildland Fire* **24**(3), 340–348. doi:10.1071/WF13099

Bolten JD, Crow WT, Zhan X, Jackson TJ, Reynolds CA (2010) Evaluating the Utility of Remotely Sensed Soil Moisture Retrievals for Operational Agricultural Drought Monitoring. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **3**(1), 57–66. doi:10.1109/JSTARS.2009.2037163

Bradshaw LS, Deeming JE, Burgan RE, Cohen JD (1983). The 1978 National Fire-Danger Rating System: technical documentation. General Technical Report INT-169. (United States Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station Ogden: UT, USA)

Brewer MJ, Clements CB (2020) The 2018 Camp Fire: Meteorological Analysis Using In Situ Observations and Numerical Simulations. *Atmosphere* **11**(1), 47. doi:10.3390/atmos11010047

Burapapol K, Nagasawa R (2016) Mapping soil moisture as an indicator of wildfire risk using Landsat 8 images in Sri Lanna National Park, northern Thailand. *Journal of Agricultural Science* **8**, 107. doi:10.5539/jas.v8n10p107

Burgan RE (1988) 1988 revisions to the 1978 National Fire-Danger Rating System. Research Paper SE-273. (United States Department of Agriculture, Forest Service, Southeastern Forest Experiment Station: Asheville, NC, USA)

California Department of Forestry and Fire Protection (2019) CAL FIRE investigators determine cause of the Camp Fire. CAL FIRE News Release, 15 May 2019. CAL Fire, Sacramento, California, USA. Available at https://www.fire.ca.gov/media/5121/campfire_cause.pdf

Carlson JD, Bradshaw LS, Nelson RM, Bensch RR, Jabrzemski R (2007) Application of the Nelson model to four timelag fuel classes using Oklahoma field observations: model evaluation and comparison with National Fire Danger Rating System algorithms. *International Journal of Wildland Fire* **16**(2), 204–216. doi:10.1071/WF06073

Carrega P (1991) A Meteorological Index of Forest Fire Hazard in Mediterranean France. *International Journal of Wildland Fire* **1**(2), 79–86. doi:10.1071/WF9910079

CFSFDG (2021) An overview of the next generation of the Canadian Forest Fire Danger Rating System. Information Report GLC-X-26. (Canadian Forest Service Fire Danger Group (CFSFDG), Natural Resources Canada, Canadian Forest Service, Great Lakes Forestry Centre: Sault Ste. Marie, Ontario, Canada)

Chaparro D, Vall-Llossera M, Piles M, Camps A, Rüdiger C, Riera-Tatché R (2016) Predicting the extent of wildfires using remotely sensed soil moisture and temperature trends. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **9**(6), 2818–2829. doi:10.1109/JSTARS.2016.2571838

Chikamoto Y, Timmermann A, Stevenson S, DiNezio P, Langford S (2015) Decadal predictability of soil water, vegetation, and wildfire frequency over North America. *Climate Dynamics* **45**(7), 2213–2235. doi:10.1007/s00382-015-2469-5

Colliander A, Cosh MH, Kelly VR, Kraatz S, Bourgeau-Chavez L, Siqueira P, Roy A, Konings AG, Holtzman N, Misra S, Entekhabi D, O'Neill P, Yueh SH (2020) SMAP detects soil moisture under temperate forest canopies. *Geophysical Research Letters* **47**(19), e2020GL089697. doi:10.1029/2020GL089697

Coops NC, Waring RH, Hilker T (2012) Prediction of soil properties using a process-based forest growth model to match satellite-derived estimates of leaf area index. *Remote Sensing of Environment* **126**, 160–173. doi:10.1016/j.rse.2012.08.024

Cosh MH, Ochsner TE, McKee L, Dong J, Basara JB, Evett SR, Hatch CE, Small EE, Steele-Dunne SC, Zreda M, Sayde C (2016) The Soil Moisture Active Passive Marena, Oklahoma, In Situ Sensor Testbed (SMAP-MOISST): Testbed Design and Evaluation of In Situ Sensors. *Vadose Zone Journal* **15**(4), 1–11. doi:10.2136/vzj2015.09.0122

Cosh MH, Caldwell TG, Baker CB, Bolten JD, Edwards N, Goble P, Hofman H, Ochsner TE, Quiring S, Schalk C, Skumanich M, Svoboda M, Woloszyn ME (2021) Developing a strategy for the national coordinated soil moisture monitoring network. *Vadose Zone Journal* **20**(4), e20139. doi:10.1002/vzj2.20139

de Groot WJ (1987) Interpreting the Canadian Forest Fire Weather Index (FWI) System. Paper presented at the Fourth Central Regional Fire Weather Committee Scientific and Technical Seminar, 2 April 1987, Winnipeg, Manitoba.

de Groot WJ (1993) 'Examples of fuel types in the Canadian Forest Fire Behavior Prediction (FBP) System.' (Forestry Canada Northwest Region Edmonton: Alberta, Canada)

de Groot WJ, Wotton BM, Flannigan MD (2015) Wildland fire danger rating and early warning systems. In 'Wildfire hazards, risks and disasters'. (Eds JF Shroder, D Paton) pp. 207–228. (Elsevier: Oxford)

Deeming JE, Lancaster JW, Fosberg MA, Furman WR, Schroeder MJ (1972) The National Fire Danger Rating System. Research Paper RM-84. (F. S. United States Department of Agriculture, Rocky Mountain Forest and Range Experiment Station: Fort Collins, CO)

Djamai N, Magagi R, Goïta K, Hosseini M, Cosh MH, Berg A, Toth B (2015) Evaluation of SMOS soil moisture products over the CanEx-SM10 area. *Journal of Hydrology* **520**, 254–267. doi:10.1016/j.jhydrol.2014.11.026

D'Orangeville L, Houle D, Duchesne L, Côté B (2016) Can the Canadian drought code predict low soil moisture anomalies in the mineral soil? An analysis of 15 years of soil moisture data from three forest ecosystems in Eastern Canada. *Ecohydrology* **9**(2), 238–247. doi:10.1002/eco.1627

Dorigo WA, Gruber A, De Jeu RAM, Wagner W, Stacke T, Loew A, Albergel C, Brocca L, Chung D, Parinussa RM, Kidd R (2015) Evaluation of the ESA CCI soil moisture product using ground-based observations. *Remote Sensing of Environment* **162**, 380–395. doi:10.1016/j.rse.2014.07.023

Dorigo W, Wagner W, Albergel C, Albrecht F, Balsamo G, Brocca L, Chung D, Ertl M, Forkel M, Gruber A, Haas E, Hamer PD, Hirschi M, Ikonen J, de Jeu R, Kidd R, Lahoz W, Liu YY, Miralles D, Mistelbauer T, Nicolai-Shaw N, Parinussa R, Pratola C, Reimer C, van der Schalie R, Seneviratne SI, Smolander T, Lecomte P (2017) ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment* **203**, 185–215. doi:10.1016/j.rse.2017.07.001

Dorigo W, Himmelbauer I, Aberer D, Schremmer L, Petrakovic I, Zappa L, Preimesberger W, Xaver A, Annor F, Ardö J, Baldocchi D, Blöschl G, Bogena H, Brocca L, Calvet JC, Camarero JJ, Capello G, Choi M, Cosh MC, Demarty J, van de Giesen N, Hajdu I, Jensen KH, Kanniah KD, de Kat I, Kirchengast G, Rai PK, Kyrouac J, Larson K, Liu S, Loew A, Moghaddam M, Martínez Fernández J, Mattar Bader C, Morbidelli R, Musial JP, Osenga E, Palecki MA, Pfeil I, Powers J, Ikonen J, Robock A, Rüdiger C, Rummel U, Strobel M, Su Z, Sullivan R, Tagesson T, Vreugdenhil M, Walker J, Wigneron JP, Woods M, Yang K, Zhang X, Zreda M, Dietrich S, Gruber A, van Oevelen P, Wagner W, Scipal K, Drusch M, Sabia R (2021) The International Soil Moisture Network: serving Earth system science for over a decade. *Hydrology and Earth System Sciences Discussions* **2021**, 5749–5804. doi:10.5194/hess-2021-2

Duff TJ, Bessell R, Cruz MG (2019) Grass Curing/Cured Fuels. In 'Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires'. (Ed. SL Manzello) pp. 1–7. (Springer International Publishing: Cham)

Ellsworth LM, Litton CM, Taylor AD, Kauffman JB (2013) Spatial and temporal variability of guinea grass (*Megathyrsus maximus*) fuel loads and moisture on Oahu, Hawaii. *International Journal of Wildland Fire* **22**(8), 1083–1092. doi:10.1071/WF12051

Elmes MC, Thompson DK, Sherwood JH, Price JS (2018) Hydrometeorological conditions preceding wildfire, and the subsequent burning of a fen watershed in Fort McMurray, Alberta, Canada. *Natural Hazards and Earth System Sciences* **18**(1), 157–170. doi:10.5194/nhess-18-157-2018

Entekhabi D, Njoku EG, O'Neill PE, Kellogg KH, Crow WT, Edelstein WN, Entin JK, Goodman SD, Jackson TJ, Johnson J, Kimball J, Piepmeier JR, Koster RD, Martin N, McDonald KC, Moghaddam M, Moran S, Reichle R, Shi JC, Spencer MW, Thurman SW, Tsang L, Van Zyl J (2010) The soil moisture active passive (SMAP) mission. *Proceedings of the IEEE* **98**(5), 704–716. doi:10.1109/JPROC.2010.2043918

Famiglietti JS, Ryu D, Berg AA, Rodell M, Jackson TJ (2008) Field observations of soil moisture variability across scales. *Water Resources Research* **44**(1), doi:10.1029/2006WR005804

Fan Y, Miguez-Macho G, Jobbágy EG, Jackson RB, Otero-Casal C (2017) Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences* **114**(40), 10572–10577. doi:10.1073/pnas.1712381114

Fan L, Wigneron JP, Xiao Q, Al-Yaari A, Wen J, Martin-StPaul N, Dupuy JL, Pimont F, Al Bitar A, Fernandez-Moran R, Kerr YH (2018) Evaluation of microwave remote sensing for monitoring live fuel moisture content in the Mediterranean region. *Remote Sensing of Environment* **205**, 210–223. doi:10.1016/j.rse.2017.11.020

FCFDG (1992) Development and structure of the Canadian Forest Fire Behavior Prediction System. Information Report ST-X-3. (Forestry Canada Fire Danger Group, Forestry Canada: Ottawa, Canada)

Forkel M, Thonicke K, Beer C, Cramer W, Bartalev S, Schmullius C (2012) Extreme fire events are related to previous-year surface moisture conditions in permafrost-underlain larch forests of Siberia. *Environmental Research Letters* **7**(4), 044021. doi:10.1088/1748-9326/7/4/044021

Forkel M, Dorigo W, Lasslop G, Teubner I, Chuvieco E, Thonicke K (2017) A data-driven approach to identify controls on global fire activity from satellite and climate observations (SOFIA V1). *Geoscientific Model Development* **10**(12), 4443–4476. doi:10.5194/gmd-10-4443-2017

Gale MG, Cary GJ, Van Dijk AIJM, Yebra M (2021) Forest fire fuel through the lens of remote sensing: Review of approaches, challenges and future directions in the remote sensing of biotic determinants of fire behaviour. *Remote Sensing of Environment* **255**, 112282. doi:10.1016/j.rse.2020.112282

Giglio L, Boschetti L, Roy DP, Humber ML, Justice CO (2018) The Collection 6 MODIS burned area mapping algorithm and product. *Remote Sensing of Environment* **217**, 72–85. doi:10.1016/j.rse.2018.08.005

Goodrick SL (2002) Modification of the Fosberg fire weather index to include drought. *International Journal of Wildland Fire* **11**(4), 205–211. doi:10.1071/WF02005

Holden ZA, Jolly WM, Swanson A, Warren DA, Jencso K, Maneta M, Burgard M, Gibson C, Hoylman Z, Landguth EL (2019) TOPOFIRE: A Topographically Resolved Wildfire Danger and Drought Monitoring System for the Conterminous United States. *Bulletin of the American Meteorological Society* **100**(9), 1607–1613. doi:10.1175/bams-d-18-0178.1

Holgate CM, De Jeu RAM, van Dijk AIJM, Liu YY, Renzullo LJ, Vinodkumar, Dharssi I, Parinussa RM, Van Der Schalie R, Gevaert A, Walker J, McJannet D, Cleverly J, Haverd V, Trudinger CM, Briggs PR (2016) Comparison of remotely sensed and modelled soil moisture data sets across Australia. *Remote Sensing of Environment* **186**, 479–500. doi:10.1016/j.rse.2016.09.015

Holgate CM, van Dijk AIJM, Cary GJ, Yebra M (2017) Using alternative soil moisture estimates in the McArthur Forest Fire Danger Index. *International Journal of Wildland Fire* **26**(9), 806–819. doi:10.1071/WF16217

Huang J, van den Dool HM, Georgarakos KP (1996) Analysis of Model-Calculated Soil Moisture over the United States (1931–1993) and Applications to Long-Range Temperature Forecasts. *Journal of Climate* **9**(6), 1350–1362. doi:10.1175/1520-0442(1996)009<1350:AOMCSM>2.0.CO;2

Hurtt GC, Chini L, Sahajpal R, Frolking S, Bodirsky BL, Calvin K, Doelman JC, Fisk J, Fujimori S, Klein Goldewijk K, Hasegawa T, Havlik P, Heinimann A, Humpenöder F, Jungclaus J, Kaplan JO, Kennedy J, Krisztin T, Lawrence D, Lawrence P, Ma L, Mertz O, Pongratz J, Popp A, Poulter B, Riahi K, Shevliakova E, Stehfest E, Thornton P, Tubiello FN, van Vuuren DP, Zhang X (2020) Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geoscientific Model Development* **13**(11), 5425–5464. doi:10.5194/gmd-13-5425-2020

III (2021) 'Facts + statistics: Wildfires.' (Insurance Information Institute: New York, NY) Available at https://www.iii.org/fact-statistic/facts-statistics-wildfires [accessed 13 August 2021]

Jensen D, Reager JT, Zajic B, Rousseau N, Rodell M, Hinkley E (2018) The sensitivity of US wildfire occurrence to pre-season soil moisture conditions across ecosystems. *Environmental Research Letters* **13**(1), 014021. doi:10.1088/1748-9326/aa9853

Jia S, Kim SH, Nghiem SV, Kafatos M (2019) Estimating Live Fuel Moisture Using SMAP L-Band Radiometer Soil Moisture for Southern California, USA. *Remote Sensing* **11**(13), 1575. doi:10.3390/rs11131575

Jolly WM (2018) Overview of NFDRS2016. Paper presented at the National NFDRS2016 Workshop, 28 April 2018, Tucson, AZ, USA.

Jolly WM, Nemani R, Running SW (2005) A generalized, bioclimatic index to predict foliar phenology in response to climate. *Global Change Biology* **11**(4), 619–632. doi:10.1111/j.1365-2486.2005.00930.x

Karthikeyan L, Pan M, Wanders N, Kumar DN, Wood EF (2017) Four decades of microwave satellite soil moisture observations: Part 2. Product validation and inter-satellite comparisons. *Advances in Water Resources* **109**, 236–252. doi:10.1016/j.advwatres.2017.09.010

Keetch JJ, Byram GM (1968) 'A drought index for forest fire control. Vol. 38.' (US Department of Agriculture, Forest Service, Southeastern Forest Experiment Station)

Kerr YH, Waldteufel P, Wigneron J-P, Delwart S, Cabot F, Boutin J, Escorihuela M-J, Font J, Reul N, Gruhier C, Juglea SE, Drinkwater MR, Hahne A, Martín-Neira M, Mecklenburg S (2010) The SMOS mission: New tool for monitoring key elements of the global water cycle. *Proceedings of the IEEE* **98**(5), 666–687. doi:10.1109/JPROC.2010.2043032

Krawchuk MA, Moritz MA (2011) Constraints on global fire activity vary across a resource gradient. *Ecology* **92**(1), 121–132. doi:10.1890/09-1843.1

Krueger ES, Ochsner TE, Engle DM, Carlson JD, Twidwell D, Fuhlendorf SD (2015) Soil moisture affects growing-season wildfire size in the Southern Great Plains. *Soil Science Society of America Journal* **79**(6), 1567–1576. doi:10.2136/sssaj2015.01.0041

Krueger ES, Ochsner TE, Carlson JD, Engle DM, Twidwell D, Fuhlendorf SD (2016) Concurrent and antecedent soil moisture relate positively or negatively to probability of large wildfires depending on season. *International Journal of Wildland Fire* **25**(6), 657–668. doi:10.1071/WF15104

Krueger ES, Ochsner TE, Quiring SM, Engle DM, Carlson JD, Twidwell D, Fuhlendorf SD (2017) Measured soil moisture is a better predictor of large growing-season wildfires than the Keetch–Byram Drought Index. *Soil Science Society of America Journal* **81**(3), 490–502. doi:10.2136/sssaj2017.01.0003

Krueger ES, Ochsner TE, Quiring SM (2019) Development and Evaluation of Soil Moisture-Based Indices for Agricultural Drought Monitoring. *Agronomy Journal* **111**(3), 1392–1406. doi:10.2134/agronj2018.09.0558

Krueger ES, Ochsner TE, Levi MR, Basara JB, Snitker GJ, Wyatt BM (2021) Grassland productivity estimates informed by soil moisture measurements: Statistical and mechanistic approaches. *Agronomy Journal* **113**, 3498–3517. doi:10.1002/agj2.20709

Kumar V, Dharssi I (2015) Sources of soil dryness measures and forecasts for fire danger rating. Bureau Research Report No. 009. (Australian Government, Bureau of Meteorology)

Kumar SV, Holmes TR, Bindlish R, de Jeu R, Peters-Lidard C (2020) Assimilation of vegetation optical depth retrievals from passive microwave radiometry. *Hydrology and Earth System Sciences* **24**(7), 3431–3450. doi:10.5194/hess-24-3431-2020

Landsberg JJ, Waring RH, Coops NC (2003) Performance of the forest productivity model 3-PG applied to a wide range of forest types. *Forest Ecology and Management* **172**(2), 199–214. doi:10.1016/S0378-1127(01)00804-0

Levavasseur G, Vrac M, Roche DM, Paillard D (2012) Statistical modelling of a new global potential vegetation distribution. *Environmental Research Letters* **7**(4), 044019. doi:10.1088/1748-9326/7/4/044019

Levi MR, Bestelmeyer BT (2018) Digital soil mapping for fire prediction and management in rangelands. *Fire Ecology* **14**(2), 11. doi:10.1186/s42408-018-0018-4

Levi MR, Krueger ES, Snitker GJ, Ochsner TE, Villarreal ML, Elias EH, Peck DE (2019) Rating fire danger from the ground up. *Eos* **100**, doi:10.1029/2019EO137858

Lu Y, Wei C (2021) Evaluation of microwave soil moisture data for monitoring live fuel moisture content (LFMC) over the coterminous United States. *Science of The Total Environment* **771**, 145410. doi:10.1016/j.scitotenv.2021.145410

Lyons DS, Dobrowski SZ, Holden ZA, Maneta MP, Sala A (2021) Soil moisture variation drives canopy water content dynamics across the western U.S. *Remote Sensing of Environment* **253**, 112233. doi:10.1016/j.rse.2020.112233

Martin AC, Jeffers ES, Petrokofsky G, Myers-Smith I, Macias-Fauria M (2017) Shrub growth and expansion in the Arctic tundra: an assessment of controlling factors using an evidence-based approach. *Environmental Research Letters* **12**(8), 085007. doi:10.1088/1748-9326/aa7989

Masinda MM, Li F, Liu Q, Sun L, Hu T (2021) Prediction model of moisture content of dead fine fuel in forest plantations on Maoer Mountain, Northeast China. *Journal of Forestry Research* **32**(5), 2023–2035. doi:10.1007/s11676-020-01280-x

Matthews S (2009) A comparison of fire danger rating systems for use in forests. *Australian Meteorological and Oceanographic Journal* **58**, 41–48. doi:10.22499/2.5801.005

Matthews S (2022) 'Fire Behaviour Index Technical Guide.' (Australian and New Zealand National Council for Fire and Emergency Services East Melbourne: Vic., Australia)

McArthur AG (1966) 'Weather and grassland fire behaviour (Leaflet no. 100).' (Forestry and Timber Bureau: Australia)

McArthur AG (1967) 'Fire Behaviour in Eucalypt Forests (Leaflet no. 107).' (Forestry and Timber Bureau: Australia)

McCabe MF, Aragon B, Houborg R, Mascaro J (2017) CubeSats in Hydrology: Ultrahigh-Resolution Insights Into Vegetation Dynamics and Terrestrial Evaporation. *Water Resources Research* **53**(12), 10017–10024. doi:10.1002/2017WR022240

McGranahan DA, Ramaano R, Tedder MJ, Kirkman KP (2016) Variation in grassland fuel curing in South Africa. *Fire Ecology* **12**(3), 40–52. doi:10.4996/fireecology.1203040

McKinnon KA, Poppick A, Simpson IR (2021) Hot extremes have become drier in the United States Southwest. *Nature Climate Change* **11**(7), 598–604. doi:10.1038/s41558-021-01076-9

McPherson RA, Fiebrich CA, Crawford KC, Kilby JR, Grimsley DL, Martinez JE, Basara JB, Illston BG, Morris DA, Kloesel KA, Melvin AD, Shrivastava H, Wolfinbarger JM, Bostic JP, Demko DB, Elliott RL, Stadler SJ, Carlson JD, Sutherland AJ (2007) Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet. *Journal of Atmospheric and Oceanic Technology* **24**(3), 301–321. doi:10.1175/JTECH1976.1

Mildrexler DJ, Zhao M, Running SW (2009) Testing a MODIS global disturbance index across North America. *Remote Sensing of Environment* **113**(10), 2103–2117. doi:10.1016/j.rse.2009.05.016

Mladenova IE, Bolten JD, Crow W, Sazib N, Reynolds C (2020) Agricultural Drought Monitoring via the Assimilation of SMAP Soil Moisture Retrievals Into a Global Soil Water Balance Model. *Frontiers in Big Data* **3**, 10. doi:10.3389/fdata.2020.00010

Moran MS, Peters-Lidard CD, Watts JM, McElroy S (2004) Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Canadian Journal of Remote Sensing* **30**(5), 805–826. doi:10.5589/m04-043

Mount AB (1972) 'KBDI and SDI are numerical representations of soil dryness estimated from temperature and precipitation (Bulletin No. 4).' (Tasmania Forestry Commission Tasmania: Australia)

Mulder VL, de Bruin S, Schaepman ME, Mayr TR (2011) The use of remote sensing in soil and terrain mapping — A review. *Geoderma* **162**(1–2), 1–19. doi:10.1016/j.geoderma.2010.12.018

Myers-Smith IH, Elmendorf SC, Beck PSA, Wilmking M, Hallinger M, Blok D, Tape KD, Rayback SA, Macias-Fauria M, Forbes BC, Speed JDM, Boulanger-Lapointe N, Rixen C, Lévesque E, Schmidt NM, Baittinger C, Trant AJ, Hermanutz L, Collier LS, Dawes MA, Lantz TC, Weijers S, Jørgensen RH, Buchwal A, Buras A, Naito AT, Ravolainen V, Schaepman-Strub G, Wheeler JA, Wipf S, Guay KC, Hik DS, Vellend M (2015) Climate sensitivity of shrub growth across the tundra biome. *Nature Climate Change* **5**(9), 887–891. doi:10.1038/nclimate2697

Nelson Jr RM (2000) Prediction of diurnal change in 10-h fuel stick moisture content. *Canadian Journal of Forest Research* **30**(7), 1071–1087. doi:10.1139/x00-032

Nemani RR, Running SW, Pielke RA, Chase TN (1996) Global vegetation cover changes from coarse resolution satellite data. *Journal of Geophysical Research: Atmospheres* **101**(D3), 7157–7162. doi:10.1029/95JD02138

Njoku EG, Jackson TJ, Lakshmi V, Chan TK, Nghiem SV (2003) Soil moisture retrieval from AMSR-E. *IEEE Transactions on Geoscience and Remote Sensing* **41**(2), 215–229. doi:10.1109/TGRS.2002.808243

NOAA CPC (2022) 'Soil Moisture (mm).' (National Oceanic and Atmospheric Administration Climate Prediction Center (NOAA CPC): College Park, Maryland, United States of America) Available at https://www.cpc.ncep.noaa.gov/products/Soilmst_Monitoring/US/Summary_anom.shtml [accessed 2 August 2022]

Noble IR, Gill AM, Bary GAV (1980) McArthur's fire-danger meters expressed as equations. *Australian Journal of Ecology* **5**(2), 201–203. doi:10.1111/j.1442-9993.1980.tb01243.x

Nolan RH, Blackman CJ, de Dios VR, Choat B, Medlyn BE, Li X, Bradstock RA, Boer MM (2020) Linking Forest Flammability and Plant Vulnerability to Drought. *Forests* **11**(7), 779. doi:10.3390/f11070779

NWCG (2002) 'Gaining an Understanding of the National Fire Danger Rating System.' (National Wildfire Coordinating Group)

O S, Hou X, Orth R (2020) Observational evidence of wildfire-promoting soil moisture anomalies. *Scientific Reports* **10**(1), 11008. doi:10.1038/s41598-020-67530-4

Ochsner TE, Cosh MH, Cuenca RH, Dorigo WA, Draper CS, Hagimoto Y, Kerr YH, Larson KM, Njoku EG, Small EE, Zreda M (2013) State of the Art in Large-Scale Soil Moisture Monitoring. *Soil Science Society of America Journal* **77**(6), 1888–1919. doi:10.2136/sssaj2013.03.0093

Oklahoma Mesonet (2021) OK-FIRE. 16-inch % plant available soil moisture. Available at https://www.mesonet.org/index.php/okfire/map/16_inch_plant_available_soil_moisture/current_maps [accessed 15 December 2021]

Palmer WC (1965) Meteorological drought, Research paper No. 45. (W. B. U.S. Department of Commerce: Washington, D.C.)

Pellizzaro G, Cesaraccio C, Duce P, Ventura A, Zara P (2007) Relationships between seasonal patterns of live fuel moisture and meteorological drought indices for Mediterranean shrubland species. *International Journal of Wildland Fire* **16**, 232–241. doi:10.1071/WF06081

PhenoCam (2021) Provisional data for site Marena, Oklahoma ROI 2012-2018. Available at http://phenocam.sr.unh.edu/ [accessed 11 August 2021]

Pook E, Gill A (1993) Variation of live and dead fine fuel moisture in *Pinus radiata* plantations of the Australian-Capital-Territory. *International Journal of Wildland Fire* **3**(3), 155–168. doi:10.1071/WF9930155

Prat-Guitart N, Rein G, Hadden RM, Belcher CM, Yearsley JM (2016) Effects of spatial heterogeneity in moisture content on the horizontal spread of peat fires. *Science of The Total Environment* **572**, 1422–1430. doi:10.1016/j.scitotenv.2016.02.145

Qi Y, Dennison PE, Spencer J, Riaño D (2012) Monitoring Live Fuel Moisture Using Soil Moisture and Remote Sensing Proxies. *Fire Ecology* **8**(3), 71–87. doi:10.4996/fireecology.0803071

Rabin SS, Magi BI, Shevliakova E, Pacala SW (2015) Quantifying regional, time-varying effects of cropland and pasture on vegetation fire. *Biogeosciences* **12**(22), 6591–6604. doi:10.5194/bg-12-6591-2015

Rakhmatulina E, Stephens S, Thompson S (2021) Soil moisture influences on Sierra Nevada dead fuel moisture content and fire risks. *Forest Ecology and Management* **496**, 119379. doi:10.1016/j.foreco.2021.119379

Reardon J, Hungerford R, Ryan K (2007) Factors affecting sustained smouldering in organic soils from pocosin and pond pine woodland wetlands. *International Journal of Wildland Fire* **16**(1), 107–118. doi:10.1071/WF06005

Rein G, Cleaver N, Ashton C, Pironi P, Torero JL (2008) The severity of smouldering peat fires and damage to the forest soil. *CATENA* **74**(3), 304–309. doi:10.1016/j.catena.2008.05.008

Rigden AJ, Powell RS, Trevino A, McColl KA, Huybers P (2020) Microwave Retrievals of Soil Moisture Improve Grassland Wildfire Predictions. *Geophysical Research Letters* **47**(23), e2020GL091410. doi:10.1029/2020GL091410

Ruffault J, Martin-StPaul N, Pimont F, Dupuy J-L (2018) How well do meteorological drought indices predict live fuel moisture content (LFMC)? An assessment for wildfire research and operations in Mediterranean ecosystems. *Agricultural and Forest Meteorology* **262**, 391–401. doi:10.1016/j.agrformet.2018.07.031

Samran S, Woodard PM, Rothwell RL (1995) The Effect of Soil Water on Ground Fuel Availability. *Forest Science* **41**(2), 255–267. doi:10.1093/forestscience/41.2.255

Sazib N, Bolten JD, Mladenova IE (2022) Leveraging NASA Soil Moisture Active Passive for Assessing Fire Susceptibility and Potential Impacts over Australia and California. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **15**, 779–787. doi:10.1109/JSTARS.2021.3136756

Schaefer AJ, Magi BI (2019) Land-Cover Dependent Relationships between Fire and Soil Moisture. *Fire* **2**(4), 55. doi:10.3390/fire2040055

Sharma S, Ochsner TE, Twidwell D, Carlson JD, Krueger ES, Engle DM, Fuhlendorf SD (2018) Nondestructive Estimation of Standing Crop and Fuel Moisture Content in Tallgrass Prairie. *Rangeland Ecology & Management* **71**(3), 356–362. doi:10.1016/j.rama.2018.01.001

Sharma S, Carlson JD, Krueger ES, Engle DM, Twidwell D, Fuhlendorf SD, Patrignani A, Feng L, Ochsner TE (2021) Soil moisture as an indicator of growing-season herbaceous fuel moisture and curing rate in grasslands. *International Journal of Wildland Fire* **30**, 57–69. doi:10.1071/WF19193

Slocum MG, Beckage B, Platt WJ, Orzell SL, Taylor W (2010) Effect of Climate on Wildfire Size: A Cross-Scale Analysis. *Ecosystems* **13**(6), 828–840. doi:10.1007/s10021-010-9357-y

Sridhar V, Hubbard KG, You J, Hunt ED (2008) Development of the soil moisture index to quantify agricultural drought and its "user friendliness" in severity-area-duration assessment. *Journal of Hydrometeorology* **9**(4), 660–676. doi:10.1175/2007JHM892.1

Stocks BJ, Lawson BD, Alexander ME, Van Wagner CE, McAlpine RS, Lynham TJ, Dubé DE (1989) The Canadian Forest Fire Danger Rating System: An Overview. *The Forestry Chronicle* **65**(6), 450–457. doi:10.5558/tfc65450-6

Van Wagner CE (1987) Development and structure of the Canadian Forest Fire Weather Index System. Technical Report 35. (Canadian Forest Service: Ottawa, Canada)

Vinodkumar, Dharssi I (2019) Evaluation and calibration of a high-resolution soil moisture product for wildfire prediction and management. *Agricultural and Forest Meteorology* **264**, 27–39. doi:10.1016/j.agrformet.2018.09.012

Vinodkumar, Dharssi I, Bally J, Steinle P, McJannet D, Walker J (2017) Comparison of soil wetness from multiple models over Australia with observations. *Water Resources Research* **53**(1), 633–646. doi:10.1002/2015WR017738

Vinodkumar V, Dharssi I, Yebra M, Fox-Hughes P (2021) Continental-scale prediction of live fuel moisture content using soil moisture information. *Agricultural and Forest Meteorology* **307**, 108503. doi:10.1016/j.agrformet.2021.108503

Walding NG, Williams HTP, McGarvie S, Belcher CM (2018) A comparison of the US National Fire Danger Rating System (NFDRS) with recorded fire occurrence and final fire size. *International Journal of Wildland Fire* **27**(2), 99–113. doi:10.1071/WF17030

Waring RH, Coops NC (2016) Predicting large wildfires across western North America by modeling seasonal variation in soil water balance. *Climatic Change* **135**(2), 325–339. doi:10.1007/s10584-015-1569-x

Waring RH, Coops NC, Running SW (2011) Predicting satellite-derived patterns of large-scale disturbances in forests of the Pacific Northwest Region in response to recent climatic variation. *Remote Sensing of Environment* **115**(12), 3554–3566. doi:10.1016/j.rse.2011.08.017

Wittich K-P (2011) Phenological observations of grass curing in Germany. *International Journal of Biometeorology* **55**(3), 313–318. doi:10.1007/s00484-010-0338-9

Wotton BM (2009) Interpreting and using outputs from the Canadian Forest Fire Danger Rating System in research applications. *Environmental and Ecological Statistics* **16**(2), 107–131. doi:10.1007/s10651-007-0084-2

Yebra M, Dennison PE, Chuvieco E, Riaño D, Zylstra P, Hunt ER, Danson FM, Qi Y, Jurdao S (2013) A global review of remote sensing of live fuel moisture content for fire danger assessment: Moving towards operational products. *Remote Sensing of Environment* **136**, 455–468. doi:10.1016/j.rse.2013.05.029

Zhao L, Yebra M, van Dijk AIJM, Cary GJ, Matthews S, Sheridan G (2021) The influence of soil moisture on surface and sub-surface litter fuel moisture simulation at five Australian sites. *Agricultural and Forest Meteorology* **298–299**, 108282. doi:10.1016/j.agrformet.2020.108282

Zhao L, Yebra M, van Dijk AIJM, Cary GJ, Hughes D (2022) Controlled field experiment clarifies the influence of soil moisture on litter moisture content. *Agricultural and Forest Meteorology* **314**, 108782. doi:10.1016/j.agrformet.2021.108782

Zheng D, Hunt ERJr, Running SW (1996) Comparison of available soil water capacity estimated from topography and soil series information. *Landscape Ecology* **11**(1), 3–14. doi:10.1007/BF02087109

Data availability. This is a review manuscript and contains no original data.

Conflicts of interest. The authors identify no conflicts of interest.

Declaration of funding. This work was supported in part by the US Department of Interior South Central Climate Adaptation Science Center (Grant no. G18AC00278) and the National Integrated Drought Information System. This research was also supported in part by the USDA National Institute of Food and Agriculture through Hatch Project OKL03123 and by the Division of Agricultural Sciences and Natural Resources at Oklahoma State University.

Acknowledgements. We thank the NOAA National Integrated Drought Information System for co-sponsoring the symposium (https://www.drought.gov/events/using-soil-moisture-information-better-understand-and-predict-wildfire-danger-symposium) that brought together many of the coauthors and launched the development of this review paper. We also thank Marina Skumanich (NOAA NIDIS) and Stephanie Connolly (USFS) for their contributions as co-organisers for that symposium. Finally, we thank NOAA-CPC for sharing archived soil moisture maps.

Author affiliations

[A]Plant and Soil Sciences, Oklahoma State University, Stillwater, Oklahoma, USA.

[B]Crop and Soil Sciences, University of Georgia, Athens, Georgia, USA.

[C]NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

[D]Biosystems and Agricultural Engineering, Oklahoma State University, Stillwater, Oklahoma, USA.

[E]Faculty of Forestry, University of British Columbia, Vancouver, British Columbia, Canada.

[F]United States Forest Service, Missoula, Montana, USA.

[G]Geography and Earth Sciences, University of North Carolina Charlotte, Charlotte, North Carolina, USA.

[H]Earth System Science, University of California, Irvine, USA, Irvine, California, USA.

[I]Present address: Department of Civil Engineering, DedanKimathi University of Technology, Private Bag 10143, Nyeri, Kenya.

ABOUT UMT FACULTY

SDI

Selective Dissemination of Information (SDI) service is a current-awareness service offered by the PSNZ for UMT Faculty Members. The contents selection criteria include current publications (last 5 years), highly cited and most viewed/downloaded documents. The contents with pdf full text from subscribed databases are organized and compiled according to a monthly theme which is determined based on the topics of specified interest.

For more information or further assistance, kindly contact us at 09-6684185/4298 or email to psnz@umt.edu.my/sh_akmal@umt.edu.my

Thank you.

Perpustakaan Sultanah Nur Zahirah
Universiti Malaysia Terengganu
21030 Kuala Nerus, Terengganu.

Tel. : 09-6684185 (Main Counter)
Fax : 09-6684179
Email : psnz@umt.edu.my