

IoT-based Low-Cost Time-Domain Reflectometry for Granular Soils: Sand-Specific Moisture Calibration and Machine-Learning Sand Identification

Golam Kibria

kibria314303@gmail.com

Rajshahi University of Engineering and Technology

Md. Mahmud Sazzad

Rajshahi University of Engineering and Technology

Ahsan Habib Pias

Rajshahi University of Engineering and Technology

Ritika Rani Bhakta

Rajshahi University of Engineering and Technology

Towhida Mostari

Rajshahi University of Engineering and Technology

Shoumik Ahmed

Rajshahi University of Engineering and Technology

Md. Rakib Hossen

Rajshahi University of Engineering and Technology

Tanvir Ahmed

Rajshahi University of Engineering and Technology

Article

Keywords: Time-domain reflectometry, Soil moisture, Sand classification, Machine learning, IoT soil sensor, Calibration, Random forest

Posted Date: September 15th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10866125/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

IoT-based Low-Cost Time-Domain Reflectometry for Granular Soils: Sand-Specific Moisture Calibration and Machine-Learning Sand Identification

Golam Kibria^{1*}, Md. Mahmud Sazzad², Ahsan Habib Pias³,
Ritika Rani Bhakta⁴, Towhida Mostari⁵, Shoumik Ahmed⁶,
Md. Rakib Hossen⁷, Tanvir Ahmed⁸

^{1*}Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

²Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

³Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

⁴Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

⁵Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

⁶Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

⁷Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

⁸Department of Civil Engineering, Rajshahi University of Engineering
& Technology, Rajshahi, 6205, Bangladesh.

*Corresponding author(s). E-mail(s): kibria314303@gmail.com;

Contributing authors: mmsruet@gmail.com;

habibahsan3002@gmail.com; ritikabhakta9@gmail.com;

towhidabookworm@gmail.com; shoumikahmed153@gmail.com;

rakibhossen549@gmail.com; tanvirahmedb37@gmail.com;

Abstract

Time-domain reflectometry (TDR) is a well-established technique for estimating soil water content, yet the universal calibration relationships developed for fine-textured soils perform poorly in sands, and the identification of soil type directly from dielectric measurements has received comparatively little attention. This study develops an integrated framework for sand-specific moisture calibration and automated sand identification using a low-cost, IoT-enabled TDR/7-in-1 soil sensor. Five natural sands (Sand 1-Sand 5) were characterized by grain-size analysis and index-property testing and were classified as poorly graded sands (SP, with Sand 4 as SP-SM; AASHTO group A-1-b). A laboratory dataset of 1,875 paired readings linking the raw sensor value to gravimetric water content was used to fit sand-specific ordinary-least-squares calibrations, which achieved coefficients of determination up to $R^2 = 0.943$ and root-mean-square errors of 1.6-2.7% moisture, substantially better than a single pooled model (RMSE = 2.84%). Independent field validation against oven-dry laboratory values gave a mean moisture accuracy of 96.5%. Three machine-learning classifiers-Random Forest, a support vector machine, and an artificial neural network were trained on the moisture and temperature channels to identify sand type, reaching an overall accuracy of 89.2%, with near-perfect recognition of the temperature-separated sands. The results confirm that sand-specific calibration markedly reduces moisture-estimation error and that dielectric sensor data carry sufficient information for automated material identification, providing a scalable methodology for TDR-based soil monitoring that can be extended to mixed and fine-grained soils.

Keywords: Time-domain reflectometry; Soil moisture; Sand classification; Machine learning; IoT soil sensor; Calibration; Random forest

1 Introduction

Time-Domain Reflectometry (TDR) has long been established as a powerful, non-destructive technique for determining soil dielectric properties and estimating volumetric water content with high temporal resolution (Jones et al. 2002). Owing to its reliability and repeatability under controlled conditions, TDR has been widely applied in agriculture, hydrology, precision irrigation, and geotechnical engineering. The method exploits the strong dependence of bulk dielectric permittivity on soil water content, enabling continuous monitoring without disturbing soil structure. However, despite its broad applicability, the reliable characterization of granular soils particularly sands remains challenging. Sands are characterized by high hydraulic conductivity, rapid water redistribution, and pronounced heterogeneity, which, together with complex air-water-solid interactions at the particle scale, produce highly variable dielectric responses and complicate accurate moisture estimation (Umenyiora et al. 2012; Yan et al. 2021).

A primary source of uncertainty in TDR-based measurements lies in the calibration models traditionally employed to relate dielectric permittivity to water content. Widely used empirical relationships, most notably the universal Topp equation, were

72 developed using fine-textured mineral soils and assume soil-independent behavior. As
73 a result, they often fail when applied to sands because they do not explicitly account
74 for critical textural and state parameters such as porosity, gradation, packing den-
75 sity, particle-size distribution, and degree of saturation (Umenyiora et al. 2012; Sakaki
76 et al. 2011). Although more physically based or semi-empirical approaches, such as
77 Hendrickx-type models, show improved agreement under specific conditions, they lack
78 explicit and repeatable formulations across varying saturation states and sand tex-
79 tures, limiting their generalizability (Umenyiora et al. 2012; Leão et al. 2020). These
80 limitations highlight the need for material-specific calibration strategies tailored to
81 granular soils.

82 In recent years, advances in signal processing, sensor technology, and machine
83 learning (ML) have been explored to address these challenges. Convolutional neural
84 networks have demonstrated potential for interpreting TDR waveforms and estimating
85 dielectric permittivity directly from reflected signals, although difficulties remain in
86 consistently identifying reflection points and standardizing probe geometry and mea-
87 surement procedures (Wang et al. 2023). Similarly, ensemble-learning methods such
88 as XGBoost and gradient-boosting regression trees have outperformed conventional
89 regression models in estimating soil water and salt content from TDR data, yet these
90 approaches are susceptible to overfitting and often lack physical interpretability when
91 soil properties are excluded from model inputs (Wan et al. 2023; Mikušová et al. 2024).
92 Parallel developments in TDR hardware have improved accuracy and enabled low-
93 cost, IoT-integrated systems for long-term monitoring; however, even with improved
94 sensors, calibration accuracy remains strongly dependent on soil texture, structure,
95 and experimental conditions (Frangi et al. 2009; Curioni et al. 2019; Comegna et al.
96 2025). Conversely, recently proposed soil-type-independent calibration relationships
97 may sacrifice accuracy by overlooking the decisive textural controls that dominate
98 dielectric behavior in granular media such as sands (Zhou et al. 2025).

99 A synthesis of the existing literature reveals two interconnected research gaps.
100 First, most TDR-based studies emphasize moisture-content estimation, while compar-
101 atively little attention has been given to the explicit identification or classification of
102 soil or sand type based on its dielectric fingerprint (Jones et al. 2002; Umenyiora et al.
103 2012). Second, even when machine-learning techniques are applied, detailed physical
104 characterization of sands such as fineness modulus, coefficient of uniformity, coeffi-
105 cient of curvature, and particle-size distribution is rarely integrated into calibration
106 or classification models, despite its well-documented influence on dielectric response
107 (Umenyiora et al. 2012; Sazzad et al. 2025). These limitations are further exacer-
108 bated by the frequent reliance on small or homogeneous datasets, which constrains the
109 robustness, transferability, and reproducibility of proposed models (Leão et al. 2020;
110 Comegna et al. 2025).

111 The present study directly addresses these gaps by developing an integrated frame-
112 work for sand-specific calibration and classification using Time-Domain Reflectometry.
113 The study employs a TDR sensor system capable of providing seven measurable
114 parameters, of which volumetric moisture content and temperature have been confi-
115 dently validated in preliminary investigations. Building on this foundation, the core
116 novelty of this work lies in systematically coupling detailed physical characterization of

sands with data-driven modelling, thereby extending TDR applications beyond generic moisture estimation toward intelligent material identification. Five distinct sand types are characterized using granulometric parameters including fineness modulus, gradation, particle-size distribution, coefficient of uniformity, and coefficient of curvature. Sand-specific calibration equations are developed from experimentally derived TDR responses, and multiple machine-learning algorithms - Random Forest, Support Vector Machines, and Artificial Neural Networks - are implemented to classify sand type based on TDR parameters, moisture content, and temperature.

The calibration and classification framework is rigorously validated using an extensive laboratory dataset obtained under controlled moisture and temperature conditions, with model performance evaluated using standard statistical metrics (R^2 , RMSE, and MAE). The anticipated outcomes of this work include refined, sand-specific calibration equations that reduce uncertainty relative to universal models, a robust and automated ML-based classification system capable of identifying unknown sand samples from their dielectric response, and a comprehensive experimental dataset that bridges soil physics with sensor analytics. By explicitly quantifying the relationship between granulometric properties and dielectric behavior, this study provides a solid foundation for improving TDR-based soil monitoring and establishes a scalable methodology that can be extended to mixed and fine-grained soils in future research.

2 Materials and Methods

2.1 System hardware architecture and data flow

The NPK 7-in-1 soil sensor integrates TDR-based moisture measurement with temperature, electrical conductivity (EC), pH, and NPK nutrient detection. When powered, the sensor activates its internal electrodes and generates a high-frequency electromagnetic pulse along its probes. The travel time and reflection characteristics of this pulse are used to calculate the soil's dielectric permittivity and derive volumetric water content, while the remaining parameters are obtained through ion-selective and conductivity-based sensing mechanisms.

All seven values are internally processed and stored in Modbus RTU registers for digital retrieval. The hardware configuration includes an RS-485 interface that connects the sensor to an ESP32 microcontroller, enabling stable long-distance communication. The ESP32 uses UART2 for Modbus queries, controlling transmit and receive functions via the DE/RE lines. A 20×4 I²C LCD is connected to the ESP32 for local display, receiving data through standard SDA and SCL lines. Power is supplied through a regulated 5-12 V source, with the ESP32 handling Wi-Fi connectivity for remote operation (Fig. 1).

During operation, the ESP32 periodically sends Modbus commands to read the sensor registers, converts the returned values into engineering units, and displays moisture, temperature, EC, pH, and NPK concentrations on the LCD in real time. Simultaneously, the processed data are transmitted over Wi-Fi to a predefined Google Apps Script endpoint. Error-handling and reconnection routines ensure reliable data acquisition and logging throughout continuous operation (Fig. 2).

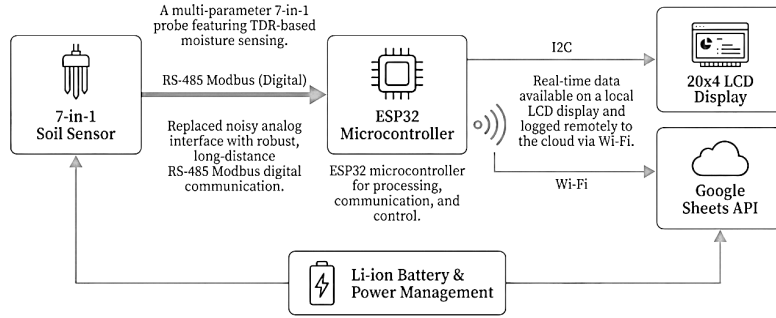


Fig. 1: System-level architecture of the IoT-based 7-in-1 soil-monitoring platform, showing the TDR-enabled NPK sensor, RS-485 Modbus communication, ESP32 processing unit, local LCD display, and cloud data logging via Wi-Fi.

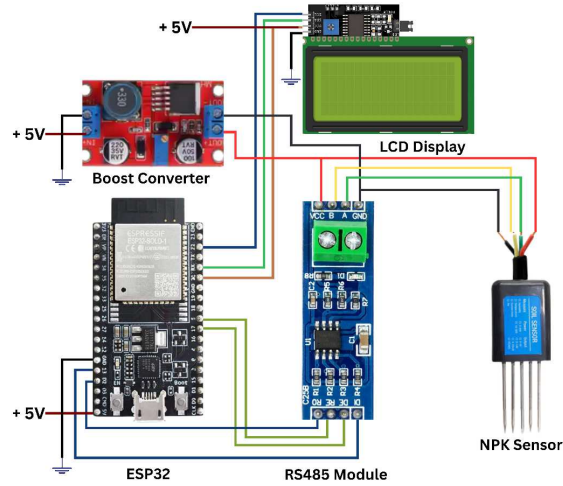


Fig. 2: Detailed hardware wiring and interfacing diagram of the ESP32-based soil-sensing system, illustrating RS-485 Modbus connectivity with the 7-in-1 NPK sensor, I²C LCD display integration, and regulated power management.

2.2 Working principle of the TDR sensor

TDR sensors determine soil moisture by analyzing the propagation characteristics of a high-frequency electromagnetic pulse transmitted along metallic waveguides inserted into the soil. When a fast step pulse travels through the probe, part of the signal is reflected whenever an impedance mismatch occurs between the transmission line and the surrounding medium. The magnitude of this reflection is quantified by the voltage reflection coefficient,

$$\rho = \frac{Z_L - Z_0}{Z_L + Z_0} \quad (1)$$

where Z_0 is the characteristic impedance of the probe and Z_L is the impedance of the surrounding medium. The sign and magnitude of ρ indicate whether the discontinuity corresponds to a higher- or lower-impedance boundary.

The key measurement in soil applications, however, is the travel time of the electromagnetic pulse. The TDR unit records the time t required for the pulse to propagate to the end of the probe and return. For a probe of length L , the apparent dielectric permittivity K_a of the soil is

$$K_a = \left(\frac{ct}{2L} \right)^2 \quad (2)$$

where c is the speed of light in vacuum. Because the propagation velocity is inversely proportional to $\sqrt{K_a}$, soils with higher water content having higher dielectric permittivity slow the pulse more strongly. Once K_a is determined, the volumetric water content (VWC, θ_v) is estimated using the widely accepted empirical relationship of [Topp et al. \(1980\)](#):

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

This polynomial reliably relates dielectric properties to soil water content for a broad range of mineral soils. Moreover, TDR's use of high-frequency pulses minimizes the influence of soil electrical conductivity, ensuring stable measurements even under varying salinity conditions (Fig. 3).

The original analogue sensing architecture was replaced with digital Modbus RTU communication via RS-485. This upgrade provides noise-immune differential signaling, long-cable support, elimination of ADC-induced nonlinearity, and the ability to poll multiple parameters from a single device.

2.3 Multi-parameter 7-in-1 sensor and communication

The 7-in-1 sensor simultaneously reports moisture (TDR-based), temperature (NTC thermistor), electrical conductivity (AC excitation), pH (glass electrode), and nitrogen, phosphorus and potassium concentrations (ion-selective membrane / colorimetric detection), as summarized in Fig. 4. The manufacturer-specified measurement ranges and accuracies are: temperature -40 to 80 °C (± 0.5 °C); moisture 0-100% ($\pm 3\%$); EC 0-10,000 $\mu\text{S cm}^{-1}$ ($\pm 10\%$); pH 0-10 (± 1); and nutrients 0-2,000 mg kg^{-1} ($\pm 2\%$ of full scale).

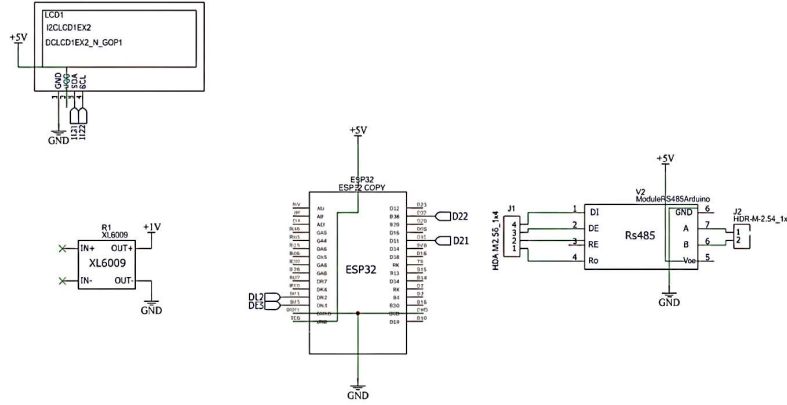


Fig. 3: Electrical schematic of the ESP32-based soil-monitoring system, showing RS-485 Modbus interfacing with the multi-parameter soil sensor and I²C-based LCD display connections.

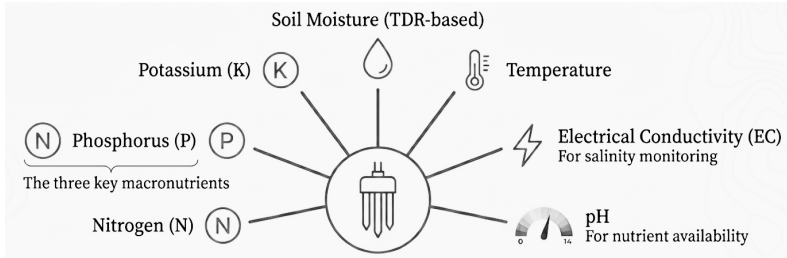


Fig. 4: Measurement capabilities of the TDR-based 7-in-1 sensor (moisture, temperature, EC, pH, N, P and K).

Each parameter is stored in a Modbus holding register. The ESP32 queries the sensor over RS-485 using function code 0x03 (read holding registers) at a baud rate of 9,600, with no parity and one stop bit.

2.4 IoT data management and local display

A 20×4 LCD provides in-field visibility, displaying moisture and temperature on the first row, EC and pH on the second, and N, P and K on the third. In parallel, the ESP32 transmits each reading via HTTP POST to a Google Apps Script endpoint, logging the timestamp together with moisture, temperature, EC, pH, and the three nutrient concentrations to a cloud spreadsheet for subsequent analysis, trend visualization, and alerting.

2.5 Test configuration and sample preparation

Laboratory-prepared specimens spanning a wide range of water contents were used to test the moisture sensor across the five sandy soils. Field samples were collected with

207 an auger or soil corer at predetermined depths and locations to ensure a representative
 208 characterization of the material under study. The assembled sensing device is shown
 209 in Fig. 5, and the end-to-end acquisition-and-logging process is summarized in Fig. 6.

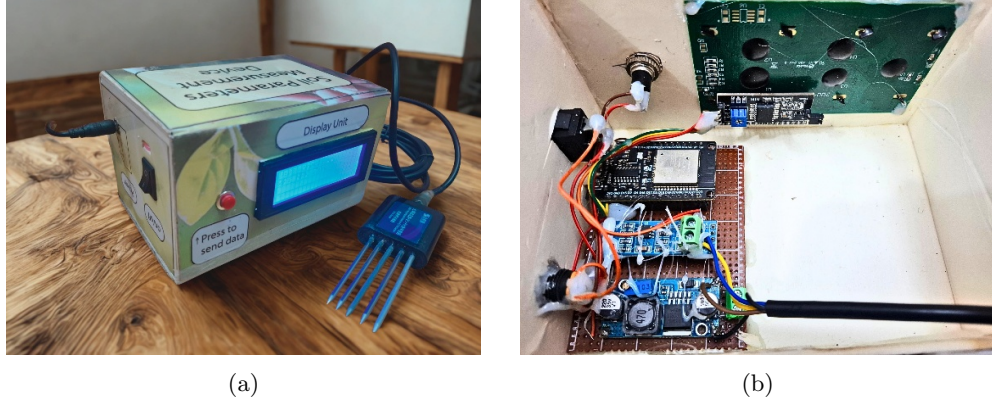


Fig. 5: The assembled soil-moisture sensing device: (a) external view of the enclosure incorporating the NPK sensor; (b) internal layout showing the ESP32 microcontroller, LCD display, and supporting circuitry for data acquisition and transmission.

210 For each specimen, the corresponding raw sensor reading (the electromagnetic-
 211 pulse response, hereafter the *Sensor Value*, S) was recorded. A calibration curve was
 212 established by plotting the raw sensor readings on the abscissa against the associated
 213 moisture contents, obtained from reference oven-dry measurements, on the ordinate.
 214 The best-fit calibration curve then yields an equation relating soil moisture content to
 215 the raw Sensor Value; this equation serves as the tool that converts raw sensor data
 216 into meaningful moisture values during subsequent measurements. The overall flow
 217 from physical measurement through local display to cloud storage is cyclic, with an
 218 operational cycle of under 15 s per iteration and a total setup time of approximately
 219 30-60 min.

220 2.6 Physical properties of the test sands

221 Five natural sand samples, designated Sand 1 to Sand 5, were used in this study. The
 222 physical and index properties of each sand were determined following standard proce-
 223 dures: the particle-size distribution was obtained by dry sieving (ASTM D6913), the
 224 specific gravity of solids was measured by the pycnometer method (ASTM D854), and
 225 each soil was classified using the Unified Soil Classification System (USCS; [ASTM
 226 International nd](#)) together with the AASHTO soil-classification system ([American
 227 Association of State Highway and Transportation Officials nd](#)). The measured prop-
 228 erties are summarized in Table 1, and the corresponding grain-size distribution curves
 229 are shown in Fig. 7.

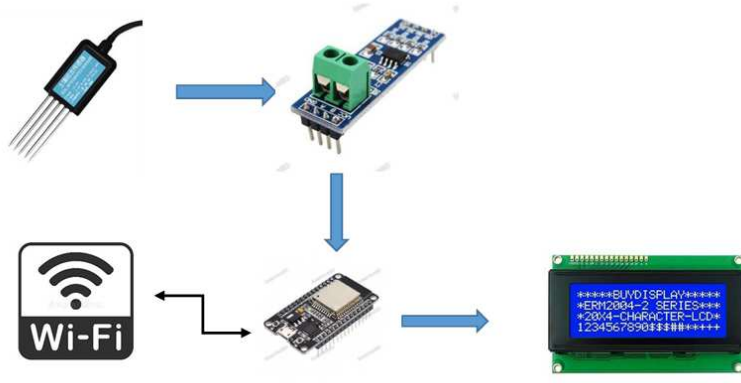


Fig. 6: End-to-end data-acquisition and logging process of the IoT soil-monitoring system, from in-soil measurement through ESP32 processing and local display to Wi-Fi cloud logging.

All five samples were coarse-grained and strongly sand-dominated: the sand fraction (4.75-0.075 mm) exceeded 92% in every case, while the gravel and fines contents remained below 1.1% and 6.4%, respectively. The effective size (D_{10}) ranged from 0.161 to 0.345 mm and the coarser characteristic size D_{60} from 0.518 to 1.296 mm, indicating fine-to-medium sands. The uniformity coefficient (C_u) varied within a narrow band of 2.52-3.76 and the coefficient of curvature (C_c) between 0.62 and 1.27. Since C_u was below 6 for all samples the threshold for a well-graded sand every soil was classified as poorly graded (uniform). The fineness modulus, computed from the gradation curves, ranged from 2.39 to 3.28; Sand 3 was the coarsest (FM = 3.28) whereas Sands 1 and 4 were the finest, giving a measurable spread in gradation across the sample set.

On the basis of these parameters, Sands 1, 2, 3 and 5 classified as poorly graded sand (SP), whereas Sand 4 whose fines content (6.37%) lies within the 5-12% range required the dual symbol SP-SM (poorly graded sand with silt). Under the AASHTO system all five sands fell in Group A-1-b (granular material; group index = 0), corresponding to an excellent-to-good subgrade rating. The specific gravity of solids ranged from 2.68 to 2.76, values typical of quartz-rich natural sands; the marginally higher figures for Sands 1 and 3 suggest a minor content of feldspar or heavy minerals. Collectively, the five sands are texturally and mineralogically similar all uniform, quartz-dominated sands of comparable specific gravity and near-identical engineering classification yet they differ systematically in their characteristic grain sizes and fineness modulus (Fig. 7). This combination of overall similarity with a controlled gradational spread makes the sample set well suited to assessing whether sensor-based dielectric measurements can resolve material differences that conventional index-property classification cannot.

Table 1: Physical properties and engineering classification of the five sand samples.

Property	Sand 1	Sand 2	Sand 3	Sand 4	Sand 5
Gravel, > 4.75 mm (%)	0.00	0.80	1.02	1.07	0.96
Sand, 4.75-0.075 mm (%)	98.97	98.85	94.88	92.55	98.17
Fines, < 0.075 mm (%)	1.03	0.35	4.10	6.37	0.87
Effective size, D_{10} (mm)	0.206	0.228	0.345	0.161	0.186
D_{30} (mm)	0.326	0.324	0.753	0.282	0.329
D_{60} (mm)	0.518	0.751	1.296	0.579	0.629
Uniformity coefficient, C_u	2.52	3.30	3.76	3.59	3.37
Curvature coefficient, C_c	0.99	0.62	1.27	0.85	0.92
Fineness modulus, FM	2.42	2.63	3.28	2.39	2.51
Specific gravity, G_s	2.70	2.69	2.70	2.68	2.72
USCS symbol	SP	SP	SP	SP-SM	SP
AASHTO group	A-1-b	A-1-b	A-1-b	A-1-b	A-1-b

D_{10} , D_{30} , D_{60} = particle sizes at 10%, 30% and 60% passing; $C_u = D_{60}/D_{10}$; $C_c = (D_{30})^2/(D_{10} \cdot D_{60})$; FM = fineness modulus (estimated from the gradation curves); G_s = specific gravity of solids; USCS = Unified Soil Classification System (ASTM D2487); SP = poorly graded sand; SP-SM = poorly graded sand with silt.

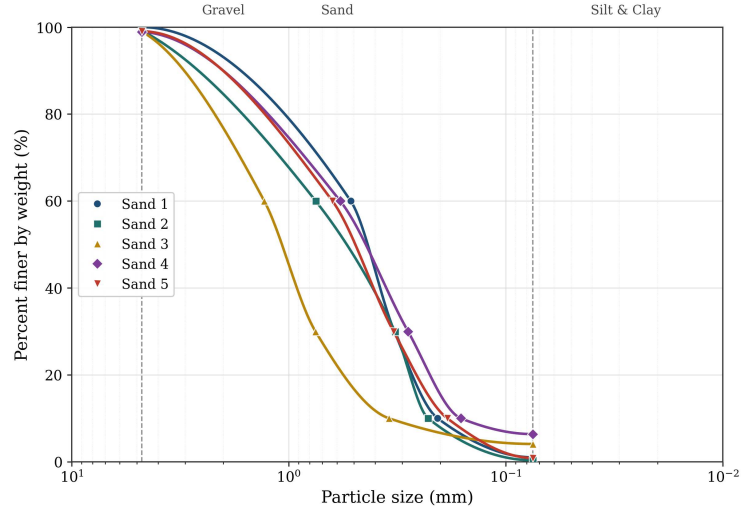


Fig. 7: Particle-size distribution curves of the five test sands (Sand 1-Sand 5). Dashed vertical lines mark the gravel-sand (4.75 mm) and sand-fines (0.075 mm) boundaries.

2.7 Dataset

The analysis uses 1,875 paired laboratory readings across the five sand samples. Each record pairs a raw Sensor Value (S) with the corresponding gravimetric water content (WC, %) measured in the laboratory, together with the ancillary channels logged by the device (temperature, EC, pH, and N-P-K). The per-sample sizes are $n = 318, 488,$

259 474, 297, and 298 for Sand 1-5, respectively (Table 2). All rows are measured data;
260 none were synthetically generated.

261 2.8 Analytical workflow

262 The complete analytical workflow is summarized in Fig. 8. The framework serves two
263 objectives moisture estimation and sand identification which share a common acqui-
264 sition front end but then follow deliberately different logic, reflecting the information
265 available in each case. Each stage is described below.

266 *Data acquisition:* Every measurement begins with the TDR/IoT probe, which returns
267 the raw Sensor Value (S) together with the moisture, temperature, EC, pH and N-P-K
268 channels. A labelled reference set of the five characterized sands the 1,875 calibration
269 readings and the tabulated physical properties (Tables 1 and 2) is stored and provides
270 the basis for both the regression calibrations and the trained classifier.

271 *Pre-processing:* The raw readings are cleaned and standardized so that the models
272 receive inputs on a consistent scale.

273 *Objective selection:* A decision node then routes the record according to the user's
274 objective, and the two branches proceed as follows.

275 *Branch 1 - Moisture estimation (soil type known):* In routine monitoring the soil
276 type is already known, because the material at a given site or borrow source has been
277 characterized in advance. The operator therefore specifies the soil type; the workflow
278 selects that sand's calibration equation (Eq. 4, $WC = a \cdot S + b$); and a single sensor
279 reading is converted directly into a moisture estimate. Because the correct, material-
280 specific calibration is applied, the reported moisture carries only that sand's calibration
281 uncertainty (its RMSE in Table 2) rather than the larger error of a generic model.

282 *Branch 2 - Sand identification (soil type unknown):* When the soil is unknown, its
283 type must be inferred from its dielectric behavior rather than assumed. The speci-
284 men is measured at several known moisture contents at least seven levels spanning
285 the working range and at each level the sensor reports its moisture and temperature,
286 which are collected into a reading matrix (a worked example is given in Section 3.3.2,
287 Table 6). This matrix of readings, rather than any single reading, is passed to the
288 trained machine-learning classifier (Random Forest, SVM, or ANN), which compares
289 it against the learned moisture-temperature responses of the five reference sands and
290 aggregates the per-reading predictions by soft voting into a single, robust decision.
291 Sampling across at least seven moisture levels is essential because the responses of
292 different sands overlap for any one reading, making a single measurement ambiguous;
293 the multi-reading matrix resolves this and yields a reliable identification.

294 *Output:* For moisture estimation the workflow returns the water content (%). For sand
295 identification it returns the predicted sand type, which is then linked to the physical-
296 properties database to report the specimen's probable index properties (C_u , C_c , G_s ,
297 fineness modulus, and USCS/AASHTO group).

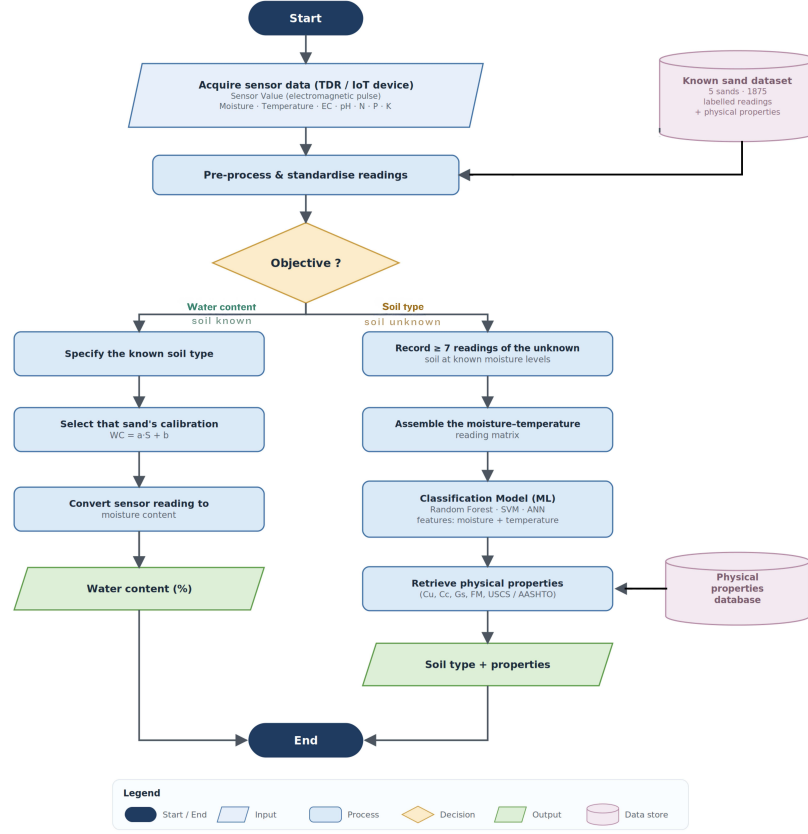


Fig. 8: Analytical workflow of the soil-sensor framework. Moisture estimation (left branch) applies the calibration of a soil whose type is already known, converting a single sensor reading into a moisture value; sand identification (right branch) takes an unknown soil, records at least seven readings across known moisture levels, and lets the machine-learning classifier identify the matching sand from this signature before returning the associated physical properties.

2.9 Regression model for moisture calibration

For each sample, gravimetric water content was regressed on the Sensor Value by ordinary least squares (Montgomery et al. 2012). A first-order (linear) relationship was fitted between the sensor reading S and the measured moisture content WC :

$$WC = a \cdot S + b \quad (4)$$

where a is the slope (the sensor sensitivity, in percent moisture per sensor unit) and b is the intercept. The moisture content was measured independently by the gravimetric (oven-dry) method to serve as the ground truth. A separate calibration was fitted for each sand because the sensor response depends on the soil's texture

306 and mineralogy; a single pooled model was also fitted across all 1,875 points as a
307 generalization check. The fitted equations are reported in Section 3.1.1 and the full
308 statistics in Table 2.

309 2.10 Machine-learning classification models

310 Sand identification is treated as a supervised multi-class problem on two sensor
311 features moisture content and temperature. Three learners were implemented and com-
312 pared: a Random Forest (RF) classifier (Breiman 2001), a Support Vector Machine
313 (SVM) with a radial-basis-function kernel (Cortes and Vapnik 1995), and an Artificial
314 Neural Network (ANN) implemented as a multilayer perceptron, with a k-nearest-
315 neighbours model as a baseline (Cover and Hart 1967). Random Forest constructs an
316 ensemble of decision trees on bootstrap samples and returns the majority vote; the
317 SVM separates classes by a maximum-margin hyperplane in a transformed feature
318 space; and the ANN learns a nonlinear decision boundary through successive weighted
319 layers. All models were built with scikit-learn (Pedregosa et al. 2011), evaluated by
320 five-fold stratified cross-validation, and combined through soft voting (probability
321 averaging) for the final decision. The predicted class is then mapped to tabulated phys-
322 ical properties (C_u , C_c , G_s , fineness modulus, and USCS/AASHTO group; ASTM
323 International nd; American Association of State Highway and Transportation Officials
324 nd; Das 2010).

325 2.11 Evaluation metrics

326 Model performance was assessed with standard statistical metrics. The coefficient of
327 determination (R^2) measures the proportion of variance in measured water content
328 explained by the fitted line. The Pearson correlation (r) measures the strength and
329 direction of the linear association between S and WC ; for simple linear regression, the
330 magnitude of r equals the square root of R^2 . The root-mean-square error (RMSE),
331 the square root of the mean squared residual (in % moisture), penalizes large errors
332 and serves as the headline calibration-accuracy figure, while the mean absolute error
333 (MAE) is less sensitive to outliers and is reported alongside it (Willmott and Matsuura
334 2005). For classification, accuracy is the fraction of readings assigned to the correct
335 sand; precision, recall, and the F1-score (their harmonic mean) are reported per class
336 and as macro- and support-weighted averages (Sokolova and Lapalme 2009).

337 3 Results and Discussion

338 The analysis comprised two linked stages: a spreadsheet-based regression analysis
339 of the measured calibration data, which establishes the direct sensor-moisture rela-
340 tionships and provides a transparent baseline (Section 3.1); and a machine-learning
341 analysis in Python, which quantifies how reliably the sand type can be recovered
342 from the sensor data (Section 3.2). The calibration statistics were computed with live
343 ordinary-least-squares spreadsheet formulas, and all diagnostic figures were produced
344 in Python.

3.1 Sensor-moisture calibration

3.1.1 Calibration equations

Ordinary least squares yields a distinct straight-line calibration for each sand, converting a raw Sensor Value (S) into an estimate of gravimetric water content (WC, %). The five sample-specific equations, together with the pooled (all-sample) equation, are:

- Sand 1: $WC = 0.6815 S - 5.0271$ ($R^2 = 0.943$, RMSE = 1.59%)
- Sand 2: $WC = 0.4612 S - 0.6513$ ($R^2 = 0.833$, RMSE = 2.67%)
- Sand 3: $WC = 0.3747 S - 0.1423$ ($R^2 = 0.805$, RMSE = 2.05%)
- Sand 4: $WC = 0.6740 S - 3.4556$ ($R^2 = 0.836$, RMSE = 1.90%)
- Sand 5: $WC = 0.7299 S - 5.8478$ ($R^2 = 0.808$, RMSE = 1.72%)
- Pooled: $WC = 0.4666 S - 0.5718$ ($R^2 = 0.792$, RMSE = 2.84%)

In each equation the slope a is the sensor sensitivity the rise in water content (%) per one-unit increase in the Sensor Value and the intercept b is the offset, i.e. the notional reading at which the estimated water content falls to zero. The slopes range from 0.375 (Sand 3, the least responsive) to 0.730 (Sand 5, the most responsive), and the intercepts range from near zero (Sands 2 and 3) to about -5 to -6 (Sands 1 and 5). Because these coefficients differ appreciably between sands, each sand must be read with its own equation; applying one sand's calibration to another introduces systematic error (Fig. 10). Operationally, once the sand has been identified (Section 3.2), the corresponding equation converts any new sensor reading into a moisture estimate, with that sand's RMSE (Table 2) as the expected uncertainty. The pooled equation is a single-curve fallback only; its larger error (RMSE = 2.84%) quantifies the penalty for ignoring sand type. Individual-sample R^2 ranges from 0.805 (Sand 3) to 0.943 (Sand 1); every relationship is strong and highly significant, given the large samples and high correlations ($r = 0.89$ - 0.97).

Table 2: Ordinary-least-squares calibration of gravimetric water content (WC, %) on Sensor Value (S) for each sand and for the pooled dataset. All statistics were computed with live spreadsheet formulas.

Sample	n	Slope (a)	Intercept (b)	R^2	Pearson r	RMSE (%)	MAE (%)
Sand 1	318	0.6815	-5.0271	0.942	0.9710	1.5933	1.2804
Sand 2	488	0.4612	-0.6513	0.833	0.9129	2.6735	2.1457
Sand 3	474	0.3747	-0.1423	0.805	0.8972	2.0518	1.5474
Sand 4	297	0.6740	-3.4556	0.836	0.9144	1.8981	1.4270
Sand 5	298	0.7299	-5.8478	0.807	0.8987	1.7150	1.3100
Pooled (all 5)	1875	0.4666	-0.5718	0.792	0.8900	2.8408	2.2366

3.1.2 Per-sample and pooled calibration

Figure 9 presents the six calibration panels (five sands plus the pooled set). In every panel the measured water content rises approximately linearly with the Sensor Value, and the OLS line (red) with its 95% confidence band (shaded) tracks the point cloud closely. Sand 1 is the tightest fit ($R^2 = 0.943$, $\text{RMSE} = 1.59\%$) and Sand 3 the loosest among the individual sands ($R^2 = 0.805$, $\text{RMSE} = 2.05\%$). The narrow confidence bands reflect the large sample sizes, while the pooled panel ($R^2 = 0.792$, $\text{RMSE} = 2.84\%$) shows visibly wider scatter, foreshadowing the case for sample-specific calibration.

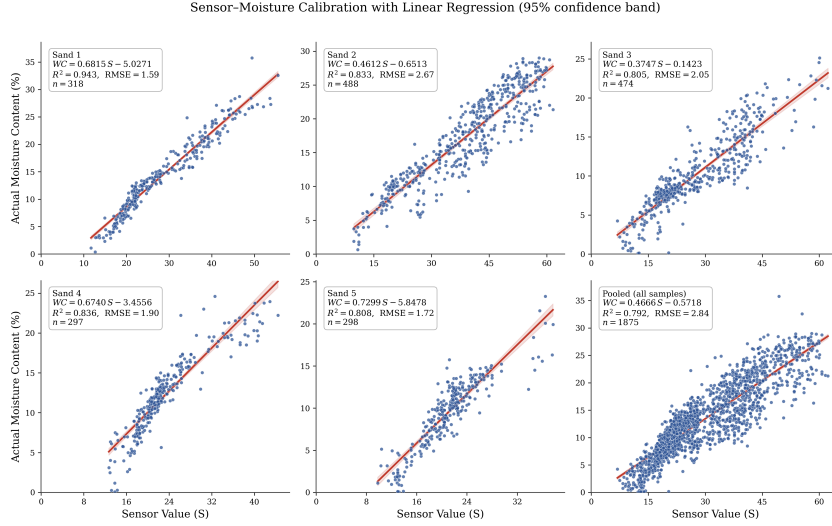


Fig. 9: Sensor-moisture calibration with linear regression and 95% confidence bands. Points are measured (gravimetric) water content against Sensor Value; each panel is annotated with its fitted equation, R^2 , RMSE and sample size n .

3.1.3 Comparison of calibration curves

Figure 10 overlays the five sand-specific regression lines on the combined (faint) scatter. The lines pass through a common region near $S \approx 15-20$ but fan out at higher sensor values: Sand 3 has the shallowest slope ($a = 0.375$) and predicts the lowest water content at a given reading, whereas Sand 5 and Sand 1 are the steepest ($a = 0.730$ and 0.682). This divergence is the visual basis for preferring one calibration per sand over a single global curve at $S = 50$ the sand-specific lines differ by several percentage points of water content.

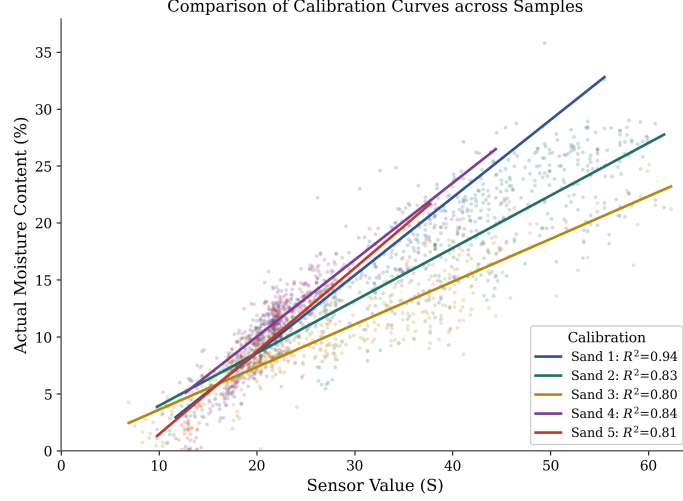


Fig. 10: Comparison of the five sand-specific calibration curves on shared axes; the legend reports each sample's R^2 . The spread of slopes shows that a single global calibration would introduce sample-dependent bias.

3.1.4 Residual analysis

Figure 11 plots residuals (measured – predicted) against predicted water content for each model. The residuals are centred on zero (solid red line), and the dashed lines mark the $\pm 1.96\sigma$ envelope ($\approx 95\%$ of residuals). Sand 1, Sand 4 and Sand 5 show fairly even scatter, consistent with the homoscedastic, unbiased-error assumption of OLS. Sand 2 and the pooled panel show mild structure a slight funnel and a hint of curvature at the extremes indicating that a strictly linear, constant-variance model is a small simplification for those cases, though not a severe one.

3.1.5 Predicted-versus-measured agreement

Figure 12 compares predicted and measured water content against the 1:1 line of perfect agreement. Points cluster tightly along the diagonal for every sample, with Sand 1 the closest and the pooled model the most dispersed. No systematic departure from the 1:1 line is evident there is no consistent over- or under-prediction across the range confirming that the linear calibrations are well-centred.

3.1.6 Performance summary

Figure 13 summarizes the two-headline metrics across the five sands. R^2 is highest for Sand 1 (0.943) and clusters between 0.805 and 0.836 for the remaining sands; RMSE is lowest for Sand 1 (1.59%) and highest for Sand 2 (2.67%). Sand 2 combines a moderate R^2 with the largest RMSE, reflecting its wider absolute spread of water contents. All five sand-specific models achieve RMSE below about 2.7%, which is the prediction error to quote as the calibration accuracy of the device.

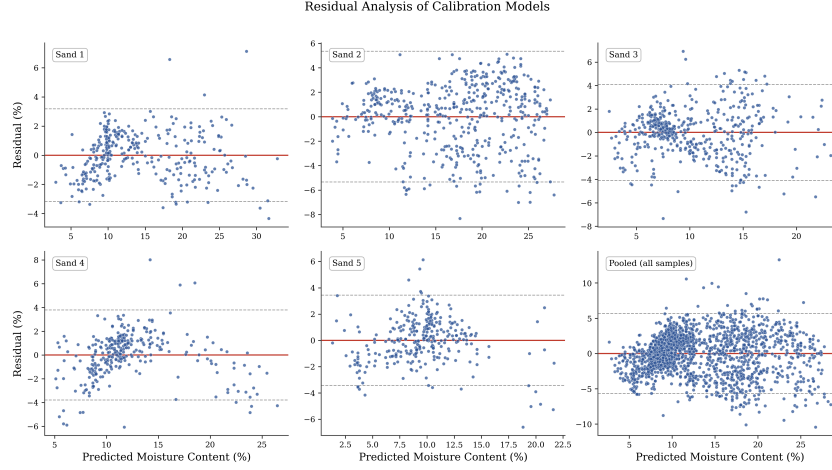


Fig. 11: Residual analysis of the calibration models. The solid line is zero error and the dashed lines are the $\pm 1.96\sigma$ limits; even scatter about zero supports the linear-regression assumptions.

3.1.7 Residual distributions

Figure 14 shows the residual distribution for each model with a fitted normal curve. All distributions are essentially zero-mean ($\mu \approx 0.00\%$) and approximately Gaussian, and the fitted standard deviations equal the corresponding RMSE values, as expected for unbiased residuals (Table 3). Sand 2 is the most spread ($\sigma = 2.67\%$) and shows mild left-tail asymmetry, echoing the structure in its residual plot; Sand 1 and Sand 5 are the narrowest ($\sigma = 1.59\%$ and 1.72%). The near-normal, zero-centred residuals support using the fitted lines for prediction with the RMSE-based error bounds.

Table 3: Residual-distribution parameters corresponding to Fig. 14. The residual standard deviation (σ) equals the RMSE for every model, confirming zero-mean (unbiased) residuals.

Sample	Residual mean μ (%)	Residual SD σ (%)	RMSE (%)
Sand 1	0.00	1.59	1.59
Sand 2	0.00	2.67	2.67
Sand 3	0.00	2.05	2.05
Sand 4	0.00	1.90	1.90
Sand 5	0.00	1.72	1.72
Pooled (all 5)	0.00	2.84	2.84

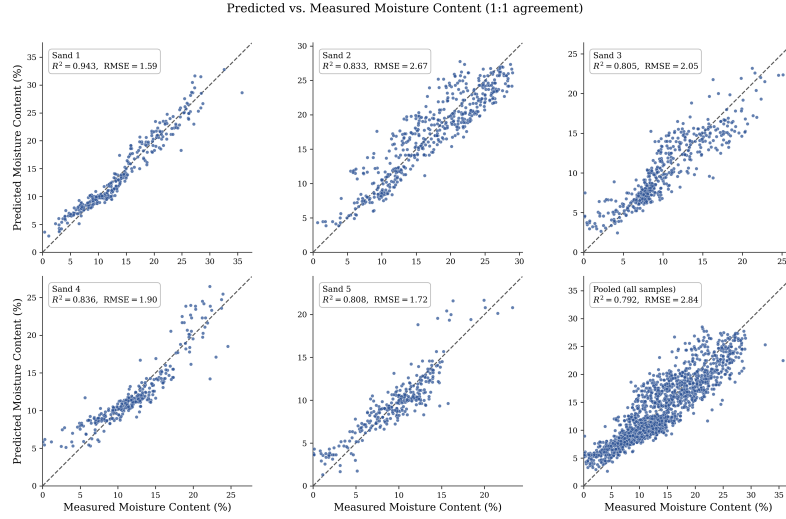


Fig. 12: Predicted versus measured moisture content with the 1:1 agreement line. Each panel repeats the model's R^2 and RMSE.

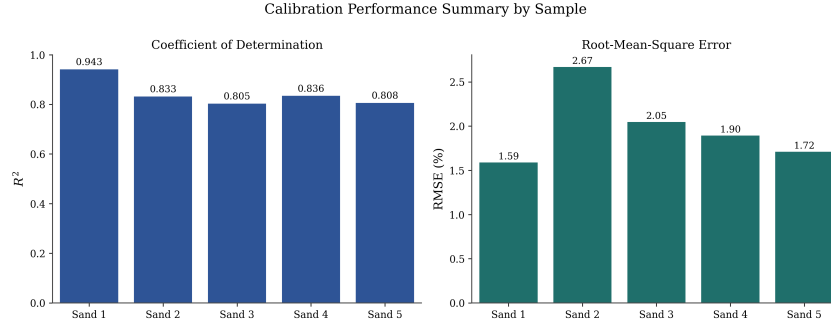


Fig. 13: Calibration performance by sample: coefficient of determination (left) and root-mean-square error (right). Values are the per-sand fits from Table 2.

3.2 Sand identification

3.2.1 Classifier performance

The three classifiers were trained on the moisture and temperature features and evaluated by five-fold stratified cross-validation. The Random Forest and ANN models each achieved an overall classification accuracy of approximately 89%, and the SVM approximately 86% well above the 20% expected from random assignment among five classes. Because a single reading is inherently ambiguous the calibration lines of different sands overlap accuracy improves substantially when several readings of the same sample are combined: averaging the classifier probabilities over 20-40 readings raises

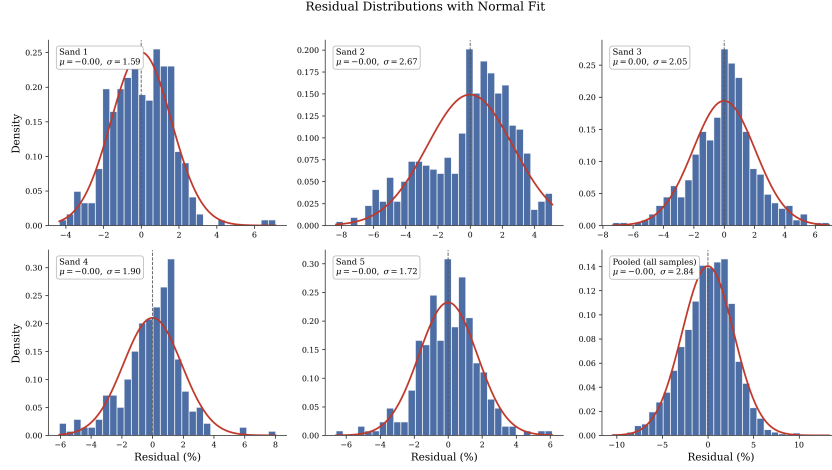


Fig. 14: Residual distributions with normal fit; each panel reports the residual mean (μ) and standard deviation (σ).

the identification accuracy to roughly 90%. This indicates that reliable material identification from the sensor is best achieved from a short series of readings rather than a single measurement.

3.2.2 Feature space

Figure 15 plots every reading in the moisture-temperature plane, colored by sand. Temperature is the dominant separating dimension. Sand 2 occupies a distinct low-temperature band (≈ 15 - 16.5°C) and Sand 3 a second band (≈ 18.5 - 20.5°C); both are well isolated from the other classes across the full moisture range. Sands 1, 4 and 5 share the warmer 24 - 31°C region and overlap substantially Sand 4 (≈ 24 - 27°C) lies between Sand 5 (≈ 26 - 29°C) and Sand 1 (≈ 28 - 31°C). Moisture alone does not separate the classes (all sands span roughly 0 - 30%), so the classifier relies mainly on temperature; this geometry anticipates where the errors occur.

3.2.3 Confusion matrix and per-class metrics

Figure 16 is the confusion matrix of the classifier over the 1,875 labelled readings; rows are the true sand and columns the predicted sand, so the leading diagonal counts correct predictions. The overall accuracy is 89.2% (1,672 of 1,875 correctly classified). Table 4 lists the per-class precision, recall and F1-score. The pattern of errors mirrors the feature space: the well-separated classes are recognized almost perfectly Sand 1 (F1 = 0.950), Sand 2 (F1 = 0.950) and Sand 3 (F1 = 0.918). The overlapping warm-region classes are weaker: Sand 4 is the hardest (recall 72.1%, F1 = 0.738), with 41 of its readings misassigned to Sand 5 and 25 to Sand 3, while 28 Sand-5 readings are in turn assigned to Sand 4. This mutual Sand 4 $\leftrightarrow$ Sand 5 confusion, together with the Sand 3 $\leftrightarrow$ Sand 4 leakage, accounts for most of the misclassifications exactly what the overlapping clusters in Fig. 15 predict. Adding a further discriminating feature

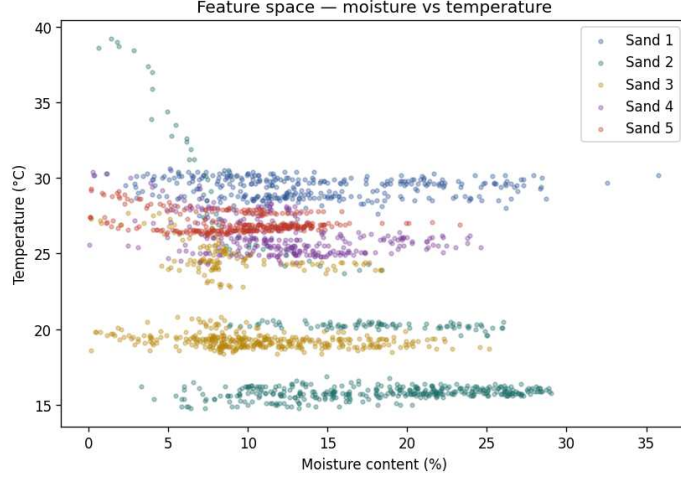


Fig. 15: Feature space used for classification: temperature versus moisture content, colored by sand. Sands 2 and 3 form isolated temperature bands, whereas Sands 1, 4 and 5 overlap in the warmer region.

449 for example the calibrated Sensor Value or the electrical-conductivity channel is the
 450 natural way to separate these three sands.

Table 4: Per-class classification performance derived from the confusion matrix (Fig. 16). Overall accuracy = 0.892 (1,672 / 1,875); macro-average F1 = 0.880; support-weighted F1 = 0.892.

Class	Support n	Precision	Recall	F1-score
Sand 1	318	0.947	0.953	0.950
Sand 2	488	0.972	0.928	0.950
Sand 3	474	0.914	0.922	0.918
Sand 4	297	0.756	0.721	0.738
Sand 5	298	0.808	0.889	0.847
Macro average	1875	0.879	0.883	0.880

451 3.3 Testing and validation

452 The framework was validated in two complementary ways: the moisture calibration
 453 was checked against independent laboratory measurements, and the sand-identification
 454 model was tested end-to-end on an unknown specimen under realistic field conditions.

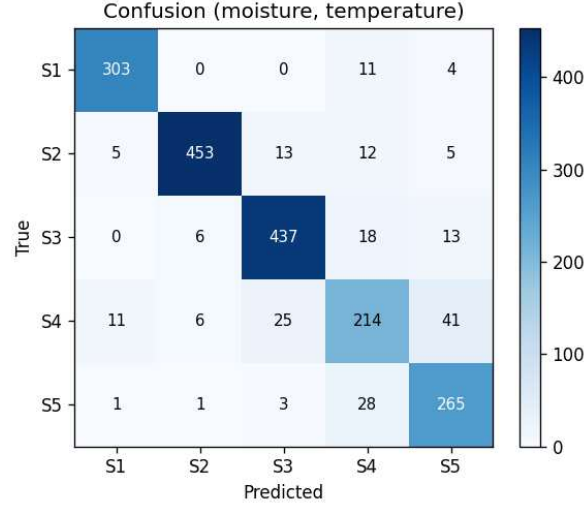


Fig. 16: Confusion matrix for the moisture-temperature classifier (rows = true sand, columns = predicted). Diagonal cells are correct classifications; the overall accuracy is 89.2%.

3.3.1 Moisture-calibration validation

The calibrated device was validated against independent laboratory reference values by comparing device-reported moisture with the corresponding oven-dry laboratory measurements on a set of field specimens. Accuracy was defined as $\left[1 - \frac{|\text{device} - \text{lab}|}{\text{lab}}\right] \times 100\%$. As summarized in Table 5, the device reproduced the laboratory moisture with a mean accuracy of 96.5%, the individual accuracies ranging from 93.5% to 99.7%. This close agreement across the operational moisture range confirms that the sand-specific calibrations transfer reliably from the laboratory to practical field measurement.

Table 5: Validation of device moisture readings against oven-dry laboratory values on field specimens.

Sample no.	Device reading (%)	Lab value (%)	Accuracy (%)	Mean 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	

3.3.2 Sand-identification validation

To assess the classification model under realistic conditions, an unknown sand specimen was identified from field-collected readings rather than from curated laboratory standards, so that the complete inference chain was exercised on genuinely unseen data. Following the procedure of Section 2.8, the specimen was measured at seven moisture levels, giving the moisture-temperature reading matrix in Table 6.

Table 6: Field reading matrix for the unknown specimen: seven paired moisture-temperature measurements used as the classifier input.

Reading	Moisture (%)	Temperature (°C)
1	6.3	26.0
2	12.0	24.4
3	8.5	19.1
4	16.0	18.9
5	18.2	19.3
6	22.4	19.2
7	8.2	25.9

The reading matrix was passed to the trained soft-voting ensemble. The model assigned the specimen to Sand 3 with a probability of 72.2%, far ahead of every other class (Table 7); the next-closest candidate, Sand 4, received only 14.8%. An independent k-nearest-neighbours classifier, applied as a cross-check, returned the same class. Because two classifiers that make errors in different ways agree here, the identification can be regarded as reliable.

Table 7: Class-probability distribution returned by the soft-voting ensemble for the specimen of Table 6; the decision (Sand 3) was confirmed by an independent k-nearest-neighbours cross-check.

Rank	Predicted class	Probability (%)
1	Sand 3	72.2
2	Sand 4	14.8
3	Sand 2	9.9
4	Sand 5	3.1
5	Sand 1	0.1

Once the class was assigned, it was linked to the physical-properties database to report the specimen’s probable engineering properties (Table 8). Two estimates are

provided: the most-likely properties, read directly from the identified class (Sand 3), and a probability-weighted expectation obtained by averaging each property across all candidate classes in proportion to their predicted probabilities. The two sets of values are close for instance, the coefficient of uniformity is 3.68 against 3.76, and the fineness modulus 3.06 against 3.28 because the leading candidates are all texturally similar, poorly graded sands. This closeness is practically important: even the 27.8% of probability mass not assigned to Sand 3 shifts the reported properties only slightly, so the engineering consequences of a borderline classification remain small.

Table 8: Predicted physical properties of the unknown specimen: most-likely values (from the identified class, Sand 3) and probability-weighted expected values across all candidate classes.

Property	Most-likely (Sand 3)	Probability-weighted expected
USCS	SP	SP
AASHTO	A-1-b	A-1-b
C_u	3.76	3.68
C_c	1.27	1.13
G_s	2.70	2.74
Fineness modulus	2.58	2.54
Fines (%)	3.28	3.18
D_{60} (mm)	1.296	1.12

Taken together with the cross-validated classification accuracy of 89.2% (Section 3.2.3) and the 96.5% moisture accuracy reported above, this worked example confirms that the system operates dependably from end to end, turning a short series of field readings on an unknown soil into a labelled sand type reported with its probable engineering properties.

3.4 Interpretation and data-quality notes

Several observations frame the interpretation of these results. Across all five sands, the sensor index is a strong linear proxy for gravimetric water content ($r = 0.89$ - 0.97), confirming that the low-cost probe carries genuine moisture information suitable for calibration. Because the slopes differ markedly between sands (0.375-0.730% per unit), a sample-specific calibration is more accurate than a single global fit; the pooled curve ($R^2 = 0.792$, RMSE = 2.84%) is therefore retained only as a robustness check, not as the recommended operational model. The RMSE of 1.6-2.7% for the individual sands is the value to cite as the calibration accuracy of the device.

For classification, the moisture-temperature model reaches 89.2% overall accuracy, with excellent performance for the temperature-separated sands ($F1 \geq 0.918$ for Sands 1-3) but weaker performance for the overlapping warm-region sands, Sand 4 being the least separable ($F1 = 0.738$). Introducing an additional feature most naturally the calibrated Sensor Value or the EC channel would most likely lift the accuracy of these three classes. Two consistency points are noted for completeness: the calibration

target throughout is gravimetric water content, and because dielectric probes are often calibrated against volumetric water content (Topp et al. 1980), the moisture basis should be stated consistently throughout the manuscript; and the raw acquisition for Sand 1 should be confirmed against the original logger export to ensure that its high R^2 (0.943) is not affected by channel mapping. All 1,875 records are measured laboratory data; no values were synthetically generated.

4 Conclusion

This study developed and validated an integrated framework that couples a low-cost, IoT-enabled TDR/7-in-1 soil sensor with data-driven modelling to achieve both accurate moisture estimation and automated sand identification. Five natural sands were characterized in detail and found to be texturally and mineralogically similar all uniform, poorly graded sands (USCS SP/SP-SM; AASHTO A-1-b) of comparable specific gravity yet distinguishable by their characteristic grain sizes and fineness modulus. The principal findings are as follows.

First, sand-specific linear calibrations relating the raw sensor value to gravimetric water content are strong and highly significant, with coefficients of determination up to $R^2 = 0.943$ and root-mean-square errors of 1.6-2.7% moisture. Because the calibration slopes differ appreciably between sands, these sample-specific models substantially outperform a single pooled calibration (RMSE = 2.84%), quantifying the error incurred by ignoring soil type. Independent field validation against oven-dry laboratory values confirmed a mean moisture accuracy of 96.5%.

Second, the same sensor stream supports automated material identification. Machine-learning classifiers trained on the moisture and temperature channels identified the sand type with an overall accuracy of 89.2%, recognizing the temperature-separated sands almost perfectly while confusing the three sands that overlap in the warm-temperature region. Combining several readings of a specimen, or introducing an additional discriminating feature such as the calibrated sensor value or electrical conductivity, is expected to raise this accuracy further.

Taken together, the results demonstrate that sand-specific calibration markedly reduces moisture-estimation uncertainty relative to universal models, and that dielectric sensor data carry sufficient information for reliable, automated sand classification. By explicitly linking granulometric properties to dielectric behavior, the framework bridges soil physics and sensor analytics and offers a scalable, reproducible methodology for TDR-based soil monitoring. Future work will extend the approach to mixed and fine-grained soils, incorporate additional sensor channels and larger, multi-site datasets, and evaluate cross-validated generalization to previously unseen specimens.

Declarations

Acknowledgements

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

Competing Interests

The authors declare that they have no competing interests as defined by Springer Nature, nor any other interests that could be perceived to influence the results and/or discussion reported in this paper.

Dual Publication

The authors confirm that the results, data, and figures presented in this manuscript have not been published elsewhere and are not under consideration by another publisher.

Authors' Contributions

Golam Kibria: Conceptualization, investigation, formal analysis, data analysis, methodology, validation, machine learning model development, dataset compilation, data visualization, writing – original draft, writing – review & editing (LaTeX formatting).

Md. Mahmud Sazzad: Supervision, writing – review & editing.

Ahsan Habib Pias, Shoumik Ahmed, Ritika Rani Bhakta, and Towhida Mostari: Data curation, visualization, literature review.

Md. Rakib Hossen: IoT device development.

Tanvir Ahmed: Visualization.

All authors reviewed and approved the final manuscript.

Third-Party Material

All material included in this manuscript is original to the authors. No permissions are required for reuse.

Data Availability

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Clinical trial number

Not applicable

Funding

This research did not receive any external funding. All expenses were covered by the authors.

Ethical Responsibilities of Authors

The authors declare that the submitted work is original, has not been published previously, is not under consideration elsewhere, and has been approved by all co-authors.

References

American Association of State Highway and Transportation Officials (n.d.). AASHTO M145: Classification of soils and soil–aggregate mixtures for highway construction purposes. AASHTO.

ASTM International (n.d.). ASTM D2487: Standard practice for classification of soils for engineering purposes (unified soil classification system). ASTM International.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

- 587 Comegna, A., Di Prima, S., Hassan, S. B. M., and Coppola, A. (2025). A novel
588 time domain reflectometry (TDR) system for water content estimation in soils:
589 Development and application. *Sensors*, 25(4):1099.
- 590 Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*,
591 20(3):273–297.
- 592 Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE*
593 *Transactions on Information Theory*, 13(1):21–27.
- 594 Curioni, G., Chapman, D. N., Royal, A. C. D., Metje, N., Dashwood, B., Gunn, D. A.,
595 Inauen, C. M., Chambers, J. E., Meldrum, P. I., Wilkinson, P. B., Swift, R. T., and
596 Reeves, H. J. (2019). Time domain reflectometry (TDR) potential for soil condition
597 monitoring of geotechnical assets. *Canadian Geotechnical Journal*, 56(7):942–955.
- 598 Das, B. M. (2010). *Principles of Geotechnical Engineering*. Cengage Learning, 7th
599 edition.
- 600 Frangi, J.-P., Richard, D.-C., Chavanne, X., Bexi, I., Sagnard, F., and Guilbert, V.
601 (2009). New in situ techniques for the estimation of the dielectric properties and
602 moisture content of soils. *Comptes Rendus Géoscience*, 341(10–11):831–845.
- 603 Jones, S. B., Wraith, J. M., and Or, D. (2002). Time domain reflectometry
604 measurement principles and applications. *Hydrological Processes*, 16(1):141–153.
- 605 Leão, T. P., da Costa, B. F. D., Bufon, V. B., and Aragón, F. F. H. (2020). Using time
606 domain reflectometry to estimate water content of three soil orders under savanna
607 in Brazil. *Geoderma Regional*, 21:e00280.
- 608 Mikušová, D., Suchorab, Z., Paśnikowska-Lukaszuk, M., Zaburko, J., and Trník, A.
609 (2024). Applying the machine learning method to improve calibration quality of
610 TDR measuring technique. *Advances in Science and Technology Research Journal*,
611 18(3):270–279.
- 612 Montgomery, D. C., Peck, E. A., and Vining, G. G. (2012). *Introduction to Linear*
613 *Regression Analysis*. Wiley, 5th edition.
- 614 Pedregosa, F., Varoquaux, G., Gramfort, A., et al. (2011). Scikit-learn: Machine
615 learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.
- 616 Sakaki, T., Limsuwat, A., and Illangasekare, T. H. (2011). A simple method for
617 calibrating dielectric soil moisture sensors: Laboratory validation in sands. *Vadose*
618 *Zone Journal*, 10(2):526–531.
- 619 Sazzad, M., Ahmed, T., Kibria, G., et al. (2025). IoT-based soil moisture measure-
620 ment and type prediction using advanced regression and machine learning models.
621 *Scientific Reports*, 15:35730.

- 622 Sokolova, M. and Lapalme, G. (2009). A systematic analysis of performance measures
623 for classification tasks. *Information Processing & Management*, 45(4):427–437.
- 624 Topp, G. C., Davis, J. L., and Annan, A. P. (1980). Electromagnetic determination
625 of soil water content: Measurements in coaxial transmission lines. *Water Resources*
626 *Research*, 16(3):574–582.
- 627 Umenyiora, C. A., Druce, R. L., Curry, R. D., Norgard, P., McKee, T., Bowders,
628 J. J., and Bryan, D. A. (2012). Dielectric constant of sand using TDR and FDR
629 measurements and prediction models. *IEEE Transactions on Plasma Science*,
630 40(10):2408–2415.
- 631 Wan, H., Qi, H., and Shang, S. (2023). Estimating soil water and salt contents
632 from field measurements with time domain reflectometry using machine learning
633 algorithms. *Agricultural Water Management*, 285:108364.
- 634 Wang, Z., Hua, S., Timlin, D., Kojima, Y., Lu, S., Sun, W., Fleisher, D., Horton,
635 R., Reddy, V. R., and Tully, K. (2023). Time domain reflectometry waveform
636 interpretation with convolutional neural networks. *Water Resources Research*,
637 59(2):e2022WR033895.
- 638 Willmott, C. J. and Matsuura, K. (2005). Advantages of the mean absolute error
639 (MAE) over the root mean square error (RMSE) in assessing average model
640 performance. *Climate Research*, 30(1):79–82.
- 641 Yan, G., Bore, T., Li, Z., Schlaeger, S., Scheuermann, A., and Li, L. (2021). Application
642 of spatial time domain reflectometry for investigating moisture content dynamics in
643 unsaturated loamy sand for gravitational drainage. *Applied Sciences*, 11(7):2994.
- 644 Zhou, Z.-Y., Li, L., Yu, W.-T., Zhang, R.-S., Tang, H., and Mu, Q.-Y. (2025). Devel-
645 opment of soil type-independent calibration relationships for water content and
646 dry density measurements using time-domain reflectometry. *Frontiers in Built*
647 *Environment*, 11:1653550.