Trakya University Journal of Engineering Sciences
http://dergipark.org.tr/tujes
ISSN 2147–0308

Research Article / Araştırma Makalesi
DOI : 10.59314/tujes.1464575

EFFICIENT SOIL MOISTURE MONITORING WITHOUT IN-SITU PROBES: LSTM-BASED BLUETOOTH SIGNAL STRENGTHS ANALYSIS

Selçuk YAZAR[1]*, Deniz TAŞKIN[2], Erdem BAHAR[3]

[1]*Kirklareli Üniversitesi, Mühendislik Fakültesi Yazilim Mühendsiliği Bölümü*
[2]*Trakya Üniversitesi, Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü*
[3]*T.C. Tarim ve Orman Bakanliği, Gelibolu İlçe Müdürlüğü*

Makale Künye Bilgisi: Yazar, S., Taşkın, D., Bahar, E. (2022). Efficient Soil Moisture Monitoring Without In-Situ Probes: Lstm-Based Bluetooth Signal Strengths Analysis, *Trakya Üniversitesi Mühendislik Bilimleri Dergisi*, 25(1), 21-38.

Highlights

- Monitoring soil moisture levels by tracking changes in Bluetooth signal strength over time, eliminating the need for in-situ probes or specialised sensors.
- Utilising artificial intelligence techniques from the LSTM architecture, this research achieves a remarkable level of accuracy in predicting soil moisture content changes.
- This study presents a method that enables the use of Bluetooth signals, traditionally used in wireless communication, as a means to non-invasively measure soil moisture content.

Article Info

Article History:
Received:
April 11, 2023
Accepted:
July 18, 2023

Keywords:
Bluetooth Low Energy;
Active microwaves;
Long Short-Term Memory;
Deep learning;
Soil moisture.

Abstract

Measuring soil moisture without damaging the soil structure is important in agriculture. Electrical conductivity and microwaves are widely used for this purpose. Recently, there has been increasing interest in using artificial neural networks and time series forecasts to determine soil moisture content. This study investigates the possibility of determining soil moisture content, especially for soil samples with different pH values, using neural network architecture and Bluetooth signal power values with a transmission power of 0.001 Watt. The aim is to assess the soil moisture change state directly using Bluetooth signal levels without an in-situ probe. In an experimental study, a machine learning model based on Bluetooth signal strengths from alkaline soil samples was used to estimate the soil moisture content change with a root-mean-square error of 15%. This method eliminates the need for a dedicated sensor, as soil moisture can be measured reliably by monitoring signal level changes over time.

Yerinde Problar Olmadan Etkili Toprak Nemi İzleme: LSTM Tabanlı Bluetooth Sinyal Gücü Analizi

Makale Bilgileri

Makale Tarihçesi:
Geliş:
11 Nisan 2023
Kabul:
18 Temmuz 2023

Anahtar Kelimeler:
Bluetooth Düşük Enerji;
Derin öğrenme;
Uzun-Kısa Süreli Bellek;
Aktif mikrodalgalar;
Toprak nemi

Öz

Toprak yapısına zarar vermeden toprak nemi ölçümü tarımda önemlidir. Elektriksel iletkenlik ve mikrodalgalar bu amaçla yaygın olarak kullanılmaktadır. Son zamanlarda, toprak nem içeriğini belirlemek için yapay sinir ağları ve zaman serisi tahminlerinin kullanılmasına olan ilgi artmaktadır. Bu çalışmada, bir yapay sinir ağı mimarisi ve 0,001 Watt iletim gücüne sahip Bluetooth sinyal güç değerleri kullanılarak, özellikle farklı pH değerlerine sahip toprak örnekleri için toprak nem içeriğinin belirlenme olasılığı araştırılmaktadır. Amaç, toprak nem değişim durumunu yerinde bir prob olmadan doğrudan Bluetooth sinyal seviyelerini kullanarak değerlendirmektir. Yapılan deneysel bir çalışmada, alkali toprak örneklerinden elde edilen Bluetooth sinyal güçlerine dayalı bir yapay öğrenme modeli kullanılarak toprak nem içeriği değişimi %15'lik bir kök-ortalama-kare hata (RMSE) değeri ile tahmin edilmiştir. Bu yöntem, toprak nemi zaman içindeki sinyal seviyesi değişiklikleri izlenerek güvenilir bir şekilde ölçülebildiğinden, özel bir sensör ihtiyacını ortadan kaldırmaktadır.

*Sorumlu Yazar: selcukyazar@klu.edu.tr

1. Introduction

Soil moisture retrieval methods which are frequently used today to measure the amount of water in the soil is a critical element that should be known mainly in agricultural production. In addition to agricultural production, the amount of water in the soil is important information for hydrological applications, disaster prediction, and environmental monitoring. There are many remote and proximal detection methods for calculating this value. All these measurement techniques used today have been developed primarily considering the chemical and physical properties of the soil. Different variables, including vegetation, need to be taken into account to determine the moisture content of the soil. Characterization and monitoring of soil characteristics are necessary to perform site-specific farming practices that are important for matching human activities with local environmental requirements.

Measuring the water content has a great interest in many disciplines, especially in porous environments such as soil. Although gravimetric sampling is the best way to measure the water content in the soil, samples must be removed from a soil mass. The methods used in the '50s and '60s to measure soil water content were mostly radioactive (Gardner and Kirkham 1952; Reginato and Bavel 1964). These methods were quite accurate and harmful to the soil; however, they required special care to avoid calibration and potential health hazards for each soil sample. A novel and innovative technique for quantifying the moisture content in soil without causing any damage or alteration was devised by employing Time Domain Reflectometry (TDR) as the primary means of measurement (Davis and Chudobiak 1975). TDR determines the dielectric constant of an object using simple electrodes placed on the sample in which moisture content is to be measured (Topp et al. 1980) proposed an empirical relationship between the dielectric constant of the soils with various tissues and the volumetric water content. One advantage of TDR is that the water content and collective electrical conductivity of the soil are measured simultaneously with a single probe. The methods which are used to determine the hydrological properties of the soil are in two main titles. Reference sampling methods or TDR at small scales such as (0.1 m) passive microwave radiometry or active radar method used in large scales such as (> 10 m and 100 m) (Lambot et al. 2010).

Small-scale techniques are often invasive, sometimes require drilling holes, and they may not represent soil characteristics on a management scale. Electrical drilling of the soil, which is an invasive method, can be done with conventional geoelectric or electromagnetic induction techniques (Carriere et al. 2021). However, the electrical conductivity of the soil is highly variable depending on the water content, water salinity, texture, and structure at the same time (Noborio 2001). For large-scale techniques, the characterization in the measured area is limited with the first few centimeters of soil and the temporal resolution is relatively poor. Non-invasive field scale techniques are required for applications involving agricultural water management, soil and water conservation, and to close the existing scale gap between ground accuracy measurements and remote sensing.

One of the most often measured soil parameters is soil reaction, or pH, which has an impact on a variety of chemical, biological, and physicochemical soil processes. Nutrient availability, root development, microbial activity, mineral solubility, and adsorption phenomena are a few examples of these activities. Despite how simple it is to measure a pH meter's output; it can be challenging to ascertain the true pH of the soil or soil solution. Because it influences many chemical and microbiological activities as well as plant growth, soil pH is a crucial factor. Valdez et al. show that it is crucial to understand that pH is a dynamic variable that is influenced by a variety of

circumstances. In irrigated soils and soils that experience wetting and drying in natural ecosystems, significant fluctuations in pH are expected to happen throughout wetting/drying cycles (Zarate-Valdez, 2006). Cycles of wet and dry weather are likely to cause changes in the pH of the soil. Some other articles (Gascho, Parker, Gaines,1996; Scheberl et al., 2019) shows a relationship with pH value and soil moisture. In those papers for all soil textures and moisture contents, it was discovered that the glass electrode sensors could measure soil pH very precisely and substantially. The relationship between some specific water solutions and soil pH levels was statistically determined.

Outside of the traditional methods used for measurement, a variety of machine learning techniques, such as the K-neighbors Regressor (KNR), Random Forest Regressor (RFR), Gradient Boosting (GB), Multi-layer Perceptron Regressor (MLPR), and Stochastic Gradient Descent Regressor (SGDR), have shown to be very beneficial for assessing soil moisture content. In a paper (Manfreda et al, 2023), the authors examined accurate global surface soil moisture (SSM) data, which is vital for hydrological and climatological needs. Machine learning (ML) techniques using various data sources were used to estimate daily SSM. Eight ML algorithms and ten ensemble models were tested. Gradient Boosting (GB), Multi-layer Perceptron Regressor (MLPR), Stochastic Gradient Descent Regressor (SGDR) and RFR showed promise in multiple climates. In particular, ensemble models combining KNR, RFR, and XB have demonstrated their potential in water management and crop yield forecasting by improving forecast accuracy.

Some other research (Pekel, 2020) discusses the development of a hybrid method that combines particle swarm optimization (PSO) and artificial neural network (ANN) for estimating soil moisture (SM) in different parameters such as air temperature, time, relative humidity, and soil temperature. The PSO algorithm is used to change the weights of the ANN in order to optimize the estimation process. The proposed hybrid PSO-ANN method shows promise in accurately estimating soil moisture, which can have implications for agriculture and climate research.

Furthermore, several Artificial Neural Networks (ANN) (Zonghan 2023; Singh and Gaurav 2023; Luo et al. 2023; Mu et al. 2022; Batchu et al. 2023) are commonly used inversion techniques for soil moisture retrieval. In the literature, several studies have been done using machine-learning-based inversion models. For instance (Ma et al. 2023) proposed a soil moisture prediction neural network guided by the water transport driving mechanism, which reduced the need for large datasets and training capability while achieving high accuracy. (Singh and Gaurav 2023) developed a fully connected feed-forward artificial neural network model to estimate surface soil moisture using satellite images, outperforming other machine learning algorithms. (Luo et al. 2023) introduced a back propagation neural network model to determine the relationship between characteristic bands/indices and soil moisture measurements, achieving high accuracy and applicability. (Mu et al. 2022) developed a nonlinear Erf-BP neural network method using multiple-resource remote-sensing data, which improved the accuracy of soil moisture estimation compared to linear models. (Batchu et al., 2023) developed a deep learning convolutional-regression model that estimated soil moisture using various predictors, achieving high correlation and accuracy.

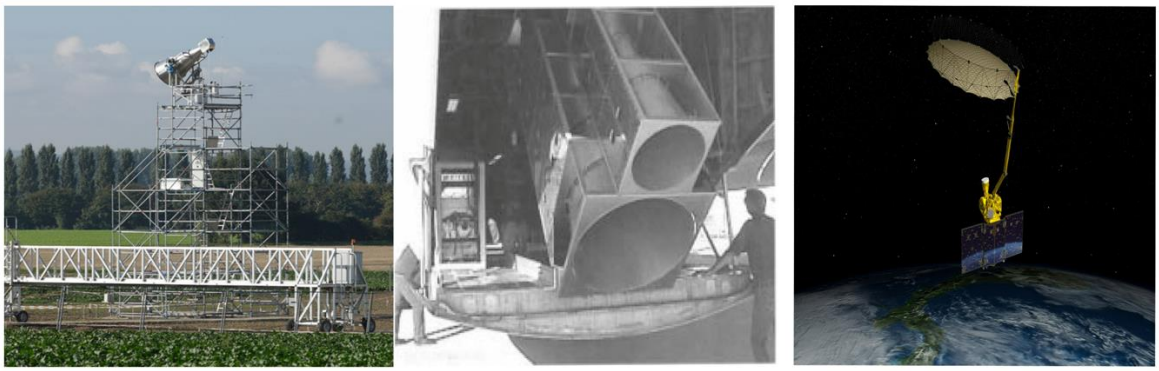
The aim of this study is the possibility of determine soil moisture content using Long Short-Term Memory (LSTM) neural network architecture and Bluetooth Low Energy (BLE) signal strengths with a maximum transmission power of 1 mW. Here, the pH values of soil samples have been used as distinctive features. Our study will contribute to soil moisture detection by

presenting a novel technique based on neural networks with satisfactory efficiency. With our method, we have shown that the moisture of the soil can be measured directly with BLE signals without using a special sensor.

2. Material and Method

2.1 Microwave Sensing

Microwaves are now frequently employed to study the structure of objects in a variety of domains, including astronomy and food (Martin et al. 2022). Energy is reflected, transmitted, or absorbed by the substance when microwaves are directed at it. Microwaves also scatter in soil and rocks due to various factors such as roughness, water content, polarization, and angle of incidence. These three types of energy ratios are referred to as material properties (Nguyen and S. Songsermpong 2022). The fundamental variables that describe how materials interact with electromagnetic fields like microwaves are dielectric constant and permeability. Diverse frequency regions of a material's dielectric profile are studied. The dielectric constant of materials can be measured at microwave frequencies using a variety of non-resonance and resonant techniques, such as transmission lines, free space, coaxial probes, and cavities. The permeability of microwaves in soil measurements has been discovered to be influenced by several factors. These factors include not only frequency, but also density, water content, sampling depth, mineral composition, grain size distribution, porosity, boundary conditions, vegetation canopies, and geographical conditions. Some of these parameters, especially the last few have very typical features. Remote measurements with microwaves can be grouped as active and passive microwave remote sensing. Images of some devices used in active and passive microwave soil water moisture measurements are shown in Figure1.



(a) (b) (c)

Figure 1. (a, b). Passive microwave sensors at performing at sites, (c) SMAP satellite launch in January 2015, for both active and passive microwave earth sensing.2.2 Dielectric measurements

Dielectric permeability is a measure of the change in the electric charge distribution caused by the electric field applied to any material. It is often expressed in relation to free space and for these reasons, it is called as complex relative permeability and is expressed as ε_r. ε_r describes the behavior of the material in the electric field as shown in Equation 1. The value shown as ε_r' in the same equation is called the dielectric constant. The second part, ε_r'', is called the loss factor. The value of ε_r varies between 2.5 for dry soil and 25 for very moist soil (Newman 1964).

$$\varepsilon_r = \varepsilon_r' - j\varepsilon_r'' \quad (1)$$

In addition, these properties differ accordingly soil density and texture. The ε_r value in Eq.(1) affects variables such as the grain size of the soil and water content in the spaces between them. The sand and clay content of the soil should also add more empirical expressions to dielectric constant.

The loss factor represents the loss of electric field energy in the material, whereas the dielectric constant represents a material's capacity to hold electrical energy. The loss tangent, as described in equation (2), is another widely used parameter. It is the ratio of the loss factor to the dielectric constant.

$$\tan\delta = {\varepsilon_r''}\Big/{\varepsilon_r'} \qquad (2)$$

Soil as a general definition can be defined as the main material in which land plants grow in the world, providing a natural environment. In the process of soil formation, it is subjected to climatic and genetic changes, and as a result, an altered material emerges with the effects of micro and macro-organisms. The content of this material is effective to a particular extent in determining its properties.

Microwaves, which are used to determine the moisture properties of the soil, can penetrate the vegetation and soil deeply. For dry vegetation and dry soil, microwave penetration

depth is larger. With an increase in vegetation and soil moisture content, penetration declines. The power of the waves will weaken as they travel through the soil due to this characteristic of microwaves. This is brought on by the dry material's variable dielectric constant in the presence of water. At ambient temperature, pure water has a dielectric constant of between 80 and 1 GHz (Calla 2002). Therefore, by computing the soil's dielectric constant using these microwave properties, the moisture content in the soil can be determined. The electromagnetic (EM) properties of the soil are impacted not only by frequency dependence but also by the dependency on density and water content. Surface soil moisture, which is an important component especially in drought studies and agricultural activities, is the water content of the top 10 cm of the soil (Dong et al. 2016). Knowing the moisture value at this depth provides the necessary information in the decision-making stages for the studies to be carried out.

2.3 Long-Short Term Memory

For many learning issues involving sequential input, recurrent neural networks with long short-term memory (LSTM) have proven to be an efficient and scalable solution. LSTM was originally designed by Hochreiter and Schmidhuber in 1997 (Hochreiter and Schmidhuber 1997). Deep learning techniques have led to the recent rediscovery of this kind of neural network. Many conventional methods for deep learning can be replaced with LSTM-based neural networks, which excel at predicting and categorizing temporal sequences. A memory cell with the ability to maintain its status over time and nonlinear door units that control the flow of information into and out of the cell make up the core of the LSTM architecture. The three different types of gates found in each LSTM block the input gate, output gate, and forget gate perform writing, reading, and resetting to the cell memory, respectively. These gates are not binary, but they are analogical (usually mapped in [0, 1], representing 0 total inhibition and 1 total activation) are managed by a sigmoidal activation function.

These gates make it possible for LSTM cells to retain information indefinitely. The cell keeps its prior state if the input gate protects it on the activation threshold, and if the current state is activated, it is coupled with the entry value. The output gate determines whether or not the value in the cell will be carried out, while the forget gate, as the name suggests, resets the present state of the cell when its value is cleared to 0.

All repetitive neural networks have the shape of a chain of the network's repetitive modules. This repeating module in conventional RNNs has a fairly straightforward structure similar to a single tanh layer. Although the repeated module of LSTMs also has a chain-like structure, it is structured differently. Four layers interact in a highly unique way instead of just one, like in a neural network. Figure 2 depicts a typical LSTM cell's general structure.

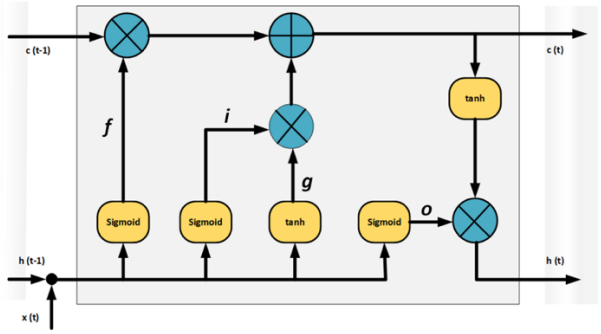


Figure 2: Sample diagram of standard LSTM cell and contents

The line shown above Figure 2 is c, which indicates the status of an LSTM cell and represents the internal memory of the unit. The technique that works around the gradient problem of LSTM is the hidden state, and it uses the i, f, o, and g gates. The parameters for these gates are learned by the LSTM during the recurrent neural network's training. i, f, and o are the input, forget, and output gates in Figure 2. The same equations are used to calculate them, but with different parameter matrices. The output vector is utilized to specify how much of the second vector can travel through the first since the sigmoid function modulates the output of these gates between zero and one. During this process, another vector can multiply the output element. Briefly, we can explain how LSTM works as follows.

The forget-gate equation (3) defines how much of the previous case h (t-1) you want to pass.

$$f = \sigma\left(W_f h_{t-1} + U_f x_t\right) \quad (3)$$

The input gate, an essential component within the architecture of a recurrent neural network, determines the degree to which the recently computed state for the present entry x(t) shall be considered, while the exit output, another crucial element, denotes the extent of the internal state that one desires to exhibit to the subsequent layer in the network.

$$i = \sigma(W_i h_{t-1} + U_i x_t) \quad (4)$$

$$o = \sigma(W_o h_{t-1} + U_o x_t) \quad (5)$$

The internal hidden state g shown in Figure 2 is calculated based on the current entry x (t) and the previously hidden state h (t-1).

$$g = tanh\left(W_g h_{t-1} + U_g x_t\right) \quad (6)$$

Equation (6) is the same as the SimpleRNN cell used in recurrent neural network methods, but in this case, its output is modulated by the output of the input gate i. Considering i, f, o and g, the cell state at any time t are expressed in c (t). Here, c(t) value can be calculated as the sum of c(t-1) multiplied by forget-gate and g value multiplied by input-gate during the training of the neural network in equation (7).

$$c_t = (c_{t-1} \otimes f) \oplus (g \otimes i) \quad (7)$$

This method is essentially a technique for merging the significance of the preceding memory and the fresh input when it comes to the learning process of the network. By assigning a value of 0 to the forget-gate, the previous memory is disregarded, while setting the input-gate to 0 results in the disregard of the newly computed state. Ultimately, the hidden-state h(t) at any given time t is obtained through the multiplication of the memory value c (t) by the output-gate.

$$h_t = \tanh(c_t) \otimes o \quad (8)$$

On the other hand, LSTMs are effective in catching long-term temporal dependencies in general. Optimization problems encountered in Simple Recurrent Networks (SRN) (Gao et al. 2020; Allen-Zhu, Li, and Song 2019) do not appear here, and thanks to these features, they have provided new approaches to solve many difficult problems. Some of these problems can be listed as handwriting recognition and production (Carbune et al. 2019; Paul et al. 2019; Ren and Ganapathy 2019), language modeling and translation (Adate and Tripathy 2019), speech acoustic modeling (Zia and Zahid 2019), speech synthesis (Hanzlíček, Vít, and Tihelka 2019), protein secondary structure prediction (Hu et al. 2019), the analysis of audio (Ertam 2019) and video data (Hussain et al. 2019), indoor based location (Zhang, Qu, and Wang 2020).

2.4 Bi-directional LSTM and Gated Recurrent Unit

In general behavior, RNN's are unidirectional and move along the directions of the original time series data. However, in some applications, capturing ordered information in reversed series greatly improves prediction. RNN structures like this that use both forward and backward pass-through improves the ability of the network to capture memory over long ranges and are called bi-directional (Schuster and Paliwal 1997). Bi-directional LSTM also acts on the time series with the same logic.

The so-called gated recurrent unit (GRU) method is quite similar to LSTM, and compared to LSTM, which has three gates, the GRU has only two processing gates. It does not have the output gate and internal memory found in LSTM. The update gate in the GRU determines how the previous memory is combined with the available memory and combines the functionality achieved by the LSTM's input and forget gates. Combining the effect of the previous memory with the effect of the current input, the reset gate is applied directly to the previous memory. Despite a few differences in how memory is transmitted across the data series within the time series, the gate mechanisms in both LSTM and GRU aim to learn long-range dependencies in the data. On the other hand, GRU has the advantage of less trainable weights compared to LSTM. Tuning model hyperparameters, such as the dimensionality of hidden units, improves the predictions of both. GRU has more advantages in situations with less training data as it requires fewer trainable weights (Cho et al. 2014).

2.4 Bluetooth Low Energy

Bluetooth Low Energy (BLE), commercially known as Bluetooth Smart which, is a low power wireless technology evolves for short-range communication. BLE operates on the 2.4 GHz Industrial Scientific Medicine (ISM) band. Since the main feature of BLE technology is lower power consumption, it is very suitable for systems running on small batteries (Darroudi, Caldera-Sànchez, and Gomez 2019). That is the reason why devices operating on BLE have the ability to utilize a coin cell as a source of power and function continuously for many years. The advantageous attributes of BLE, including its low power consumption and economical cost, render it an optimal selection for sensor devices that require power sensitivity. Furthermore, BLE can be seamlessly integrated into a diverse range of applications, such as efficient monitoring of environmental and health conditions (Ghori, Wan, and Sodhy 2019; Al Mamun and Yuce 2019; Wu, Wu, and Yuce 2019), smart home automation (Ali and Ali 2019), indoor localization (Nagarajan et al. 2020), as well as labeling and services based on proximity. It is also used in combination with other close-range communication protocols such as Near Field Communication (NFC) or RFID (Wang et al. 2019). In BLE applications, multiple BLE slave devices known as advertisers can be connected to the master BLE devices known as scanners.

There are three class devices in BLE with radio transmission power specified as 100mW (20dBm), 2.5mW (4dBm) and 1mW (0dBm). For Class 3 devices with 0dBm radio transmission power, the communication range exceeds 10 meters, and for 20dBm, the communication range exceeds 100 meters. Bluetooth power classes shown in Table 1.

Table 1: Bluetooth power classes.

Power Class	Maximum Output Power (mW)	Range (m)
1	100	100
2	2.5	15
3	1	10

In this study, the signals produced by the simple advertiser BLE beacon device were used. The devices that used are in the 3rd power class and their transmitter power is maximum 1mW.

2.4 Soil Samples Preparation

The soil samples used in this study were prepared in the laboratories of Kırklareli Agricultural Research Institute. Soil samples were prepared for analysis by drying, beating and sieving and their reactions were examined using purified water. The pH determination method here is based on potentiometric measurement of the amount of hydrogen ions in the medium created by mixing the soil with deionized purified water using a pH meter. No measurements were made for moisture during soil sample preparation but moisture measurements made during experiment. Soil samples were prepared at the institute, brought to the laboratory of Trakya University for the experiments and the measurement system was installed there.

3. Experiment

During this study, we used soil samples with pH values of 4.47, 5.31, 6.64, and 7.52 to determine the soil water content. We did not make a special selection study for properties such as texture, hardness except pH value that can be found in soil samples. We placed beacons with Texas Instruments CC2541 BLE chip in plastic containers where we put soil samples. We placed the transmitting devices at a depth of 15 cm and in the horizontal plane. Beacon devices that we use BLE chips have antennas in the "inverted F" format and produce an approximately spherical electromagnetic field. The BLE receiver is the BLE receiver on the RaspberryPi card. We recorded the Received Signal Strength Indicator (RSSI) values of beacons at 15 seconds intervals, to a database over the network with a scanner application running on Raspberry Pi device. Soil samples and plastic containers are shown in Figure 3.

In our experiment, for 8 weeks, we added 200 ml of water to each soil sample with the different pH properties on Mondays. We assumed that all soil samples were found in the same atmospheric conditions since the containers with the samples were open-mouthed. During this process, we constantly recorded the values of Beacon devices.

BLE receiver software is developed by Python with pyBluez, the Python port of the Bluez protocol stack, and the basic Python libraries, which transform each BLE node into a container based virtual nodes. The Python port of the Bluez protocol stack is preferred due to its flexibility and ease of use in implementing Bluetooth-based systems. Additionally, the Bluez protocol stack, offers direct access to the Host Controller Interface (HCI) layer, reducing overhead from higher layers and enabling efficient data transfer operations. The Python application scans and receives RSSI information from beacon devices.

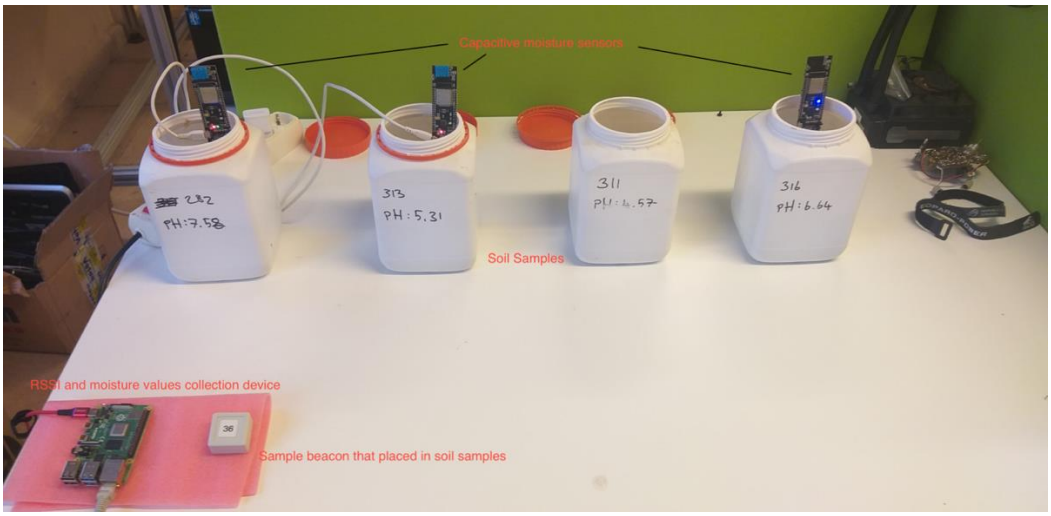


Figure 3: Setup of sample soils for BLE for RSSI and moisture values collecting.

4. Results and Discussion

First, we reduced the signal strength data we collected in our experiment to 30 minutes intervals. We reduced the number of data we collected at intervals of 15 seconds to a reasonable level. Before creating the artificial intelligence model for the collected RSSI data, we examined the mean values of this information according to the days of the week. The mean values we obtained are shown in Table 2.

Table 2: Table of calculated mean values of daily measurements based on the pH value of the soil samples.

Weekdays / Means of RSSI values	pH 4.47	pH 5.31	pH 6.64	pH 7.52
1	-80.997520	-79.096771	-73.359802	-80.034737
2	-80.539352	-80.437500	-76.238426	-82.134262
3	-79.851852	-81.171295	-75.680557	-82.004631
4	-78.689812	-80.858795	-72.793983	-80.821762
5	-78.392250	-78.968521	-72.450363	-79.595642
6	-77.174477	-79.648438	-71.059898	-80.125000
7	-79.463539	-79.473961	-70.554688	-78.958336

When we examine the data in Table 2, we found that the rate of change of RSSI values based on days of the week is closer to each other in alkali soil samples. To see this result better, we graphically analyzed the mean values we obtained. The results are shown in Figure 4. According to the results shown in Figure 4, we have seen that the means of RSSI values are close to each other as the pH value of the soil samples increases. When we pay particular attention to the graphics (c) and (d), we observed a similar change in RSSI levels from the first day of the week when we added water to soil samples with pH 6.64 and pH 7.52. This shows that in our experiment, we can use BLE signals to estimate soil moisture value/level.

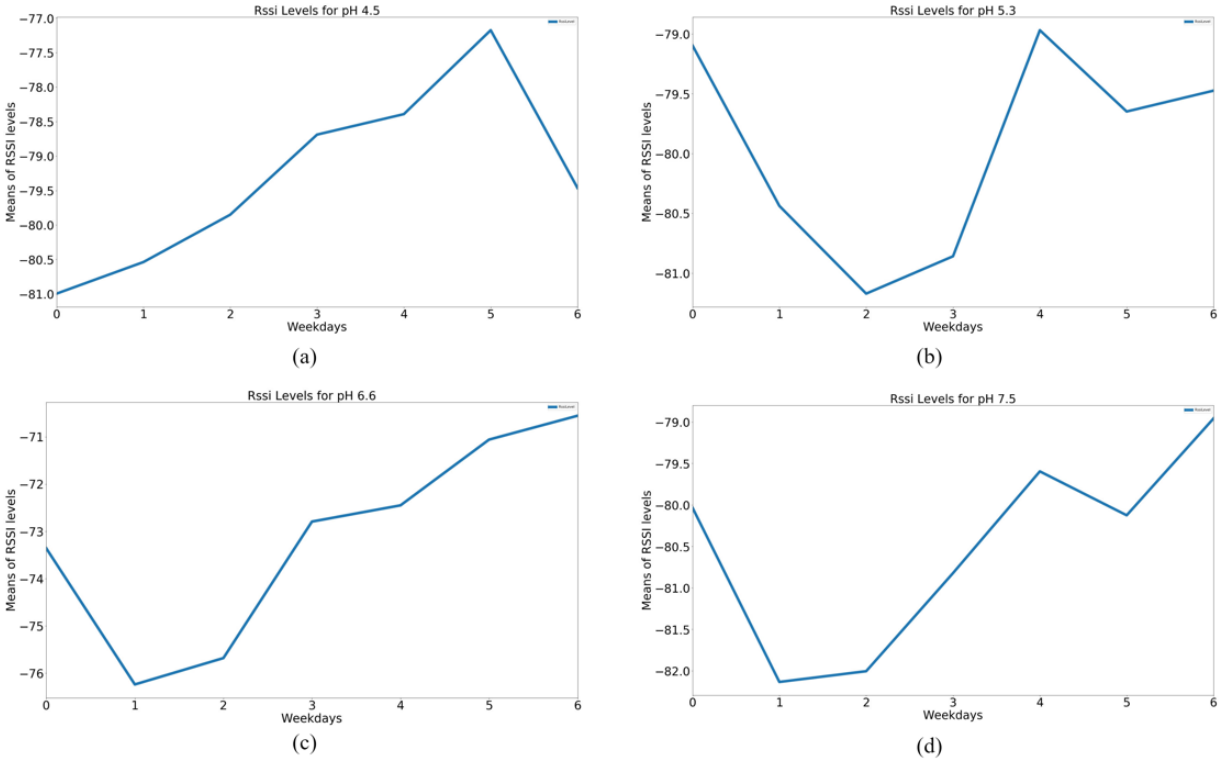


Figure 4: (a) RSSI levels for pH 4.5 daily manner, (b) RSSI levels for pH 5.3 daily manner, (c) RSSI levels for pH 6.6 daily manner, (d) RSSI levels for pH 7.5 daily manner, Graphs showing the change of RSSI levels of all soil samples on days of the week (0:Monday, 6:Sunday).

We aimed to develop an artificial neural network model where we can estimate the RSSI values here based on time, especially since we observed the gradual change in the soil at pH 7.52. While creating the artificial neural network, we chose to send the time-bound BLE signals to the 1D convolution layer before sending them to the LSTM cells for higher performance. One dimensional convolution artificial neural network layer is used especially in time series analysis. In our studies, we have achieved the most successful result of the prediction models we have created to find the lowest RMSE value for four different pH soil samples, thanks to the model shown in Figure 5.

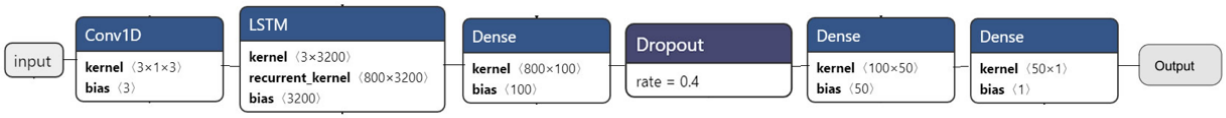


Figure 5: Artificial Neural Network model for predicting for RSSI values.

In the model we have developed, there are 800 time steps in the LSTM layer we feed after the first convolution layer. In this study, the total data set size that we use for the training of the neural network is 2880 BLE signal data. The filter that we set here as 1x3x1, is moved by a stride of the one-time unit over the on data set. Thus, 1 x 3 feature maps are created. The rest of the time steps are calculated similarly. We used an approach equivalent to the third degree autoregressive model training using a 1 x 3 convolution filter. We also added a 40% drop out layer to our model to prevent over-fitting of training data. This style of approach produces features on short-term subsets of time series. Progress of filter shown in Figure 6.

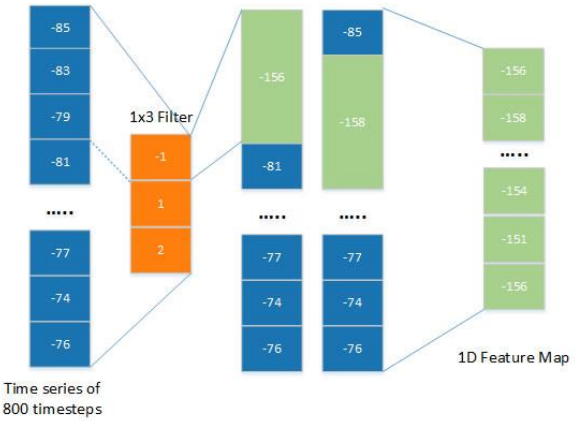


Figure 6: Illustration of 1D convolution.

In the LSTM layer, we preferred hyperbolic tangent (tanh) as the activation function. We used the 3200-dimensional vector we obtained from the LSTM layer to feed 100, 50 and 1 dimensional fully connected layers, respectively. In this way, we tried to predict BLE signal strength values we obtained with 30 minutes intervals. By running this model for 100 epochs, we compared the results for all different pH values. In the model training, no significant change was observed for more than 100 epochs, and to avoid overfitting, an optimum number of 100 was found to be appropriate. A lower value was found to be insufficient considering the frequency of signal recording.

During the training of the artificial neural network, we used the Mean squared Error function for the loss function and the adaptive momentum (ADAM) approach to optimize the Loss function. We also set the value of batch size to 256. The training and prediction results of the soil sample at pH 7.52 with the lowest RMSE value we obtained are shown in Figure 7. As seen here, even though the train and predict values of our artificial intelligence model do not fully capture the signal levels, we have seen that the RSSI values are close to the peak values generated over time with a certain offset difference.

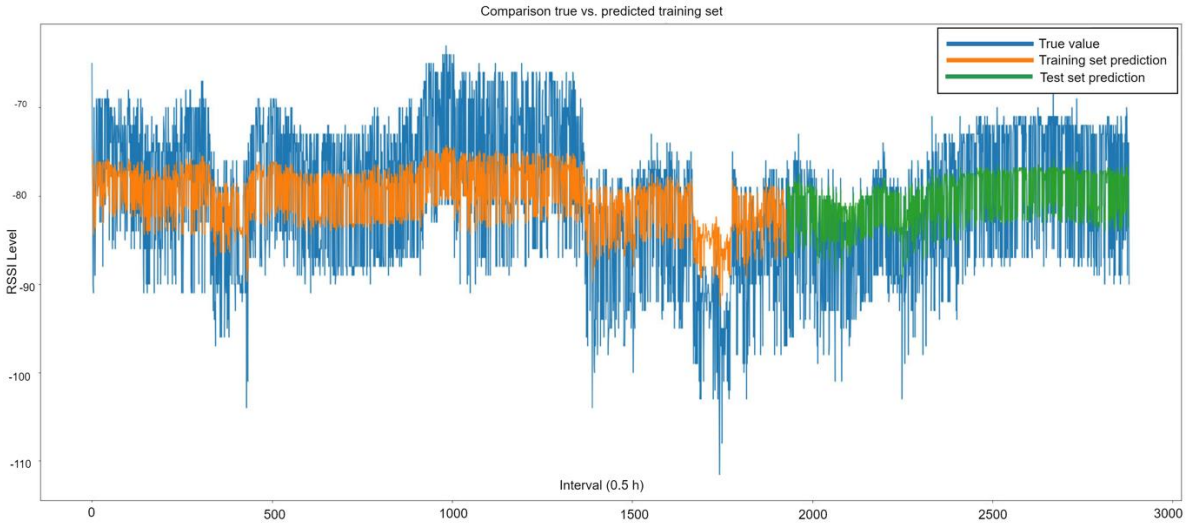


Figure 7: Best train and prediction results for pH 7.52 soil sample

Also, as seen in Figure 7, there is a technical artifact that we think our dataset originates in measurement devices between 1500 –2000 values. However, this distortion did not make any difference to the mean values we calculated at the signal levels. The model results that we obtained for all different pH values are shown in Table 3. Considering the results in Table 3, it is seen that the GRU and the Bi-directional LSTM methods give very good results in soils at the acidity limit. But the other values in the table also show that the method used works well.

Table 3: RMSE results for all soil samples with other LSTM types

Soil pH value	RMSE value with LSTM	RMSE value with GRU	RMSE value with Bi-directional LSTM
4.47	8.42	8.61	8.59
5.31	8.12	7.75	7.79
6.64	8.75	8.82	8.73
7.52	8.01	0.15	0.20

Another metric used in regression estimation such as LSTM is the R2 value. R2, also known as the coefficient of determination, is a statistical measure of the proportion of variance in the dependent variable that can be estimated from the independent variables in a regression model. R2 is a scale-free score, meaning that its value is always less than one, regardless of how big or tiny the values are. The calculation of R2 is shown in equation (9).

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})}{\sum(y_i - \bar{y})} \qquad (9)$$

Table 4 shows the R2 values of the measured RSSI values. Again, the highest and consistent values were obtained in the alkaline soil sample, confirming the RMSE results. In terms of R2 and RMSE relationship, the prediction made with GRU unit for the sample with a pH value of 7.52 fits the data better than other models.

Table 4: R2 results for all soil samples with other LSTM types

Soil pH value	R^2 value with LSTM	R^2 value with GRU	R^2 value with Bi-directional LSTM
4.47	0.06432	0.08620	0.086182
5.31	0.04050	0.04041	0.04048
6.64	-1.6210	0.08264	0.09006
7.52	0.08626	0.08625	0.08624

During the experiments, RSSI values were measured periodically but not frequently with a capacitive moisture meter. This device measures the moisture in the soil in % value. This value, which is found by proportioning two discrete values defined as dry and wet, is used to determine the moisture content in the soil in simple applications. On device sensor there is a

capacitor and when the capacitor is charged, it begins to discharge. The duration of this process is measured by the sensor. The variance of the moisture values we obtained during the study period is shown in Table 5.

Table 5: Variance of moisture sensor values

Soil pH value	Variance of moisture sensor values
4.47	1.06122449
5.31	1.346938776
6.64	1.530612245
7.52	1.204081633

As can be seen in Table 5, the variance of the change between the measured values increases gradually in alkaline samples. Although the variation of soil moisture is not related to pH in terms of certain minerals, the result seen here is that moisture values change visibly. The distinctive properties of the soil samples were determined only on the basis of pH and other components were not taken into account. Therefore, although it is not clear whether this change is due to other characteristics of the samples, there is a significant change. Different pH values were used here as a distinguishing feature of the four different soil samples. The effects of pH values in relation to water or various salts have been studied in many studies. However, this study did not set up a setup to directly examine these relationships.

In our study, we observed that the results of the artificial intelligence model and it also gave good results in clay soils, even if we did not take into account the clay rate in soil samples.

Considering Fig. 4.b, although the change in RSSI values here occurs later than acidic soil samples, the RMSE values obtained as a result of the artificial intelligence model are closer to the alkali soil sample (pH 7.52). Under normal conditions, water diffusion in clay soils is slow due to the very small gaps in the clay (Haria et al. 1994). Our model results support this fact. Another result is that the clay rate in the soil increases and behaves similarly to acidic properties. The prediction of the RSSI values of the clay soil sample at pH 5.31 is shown in Figure 8.

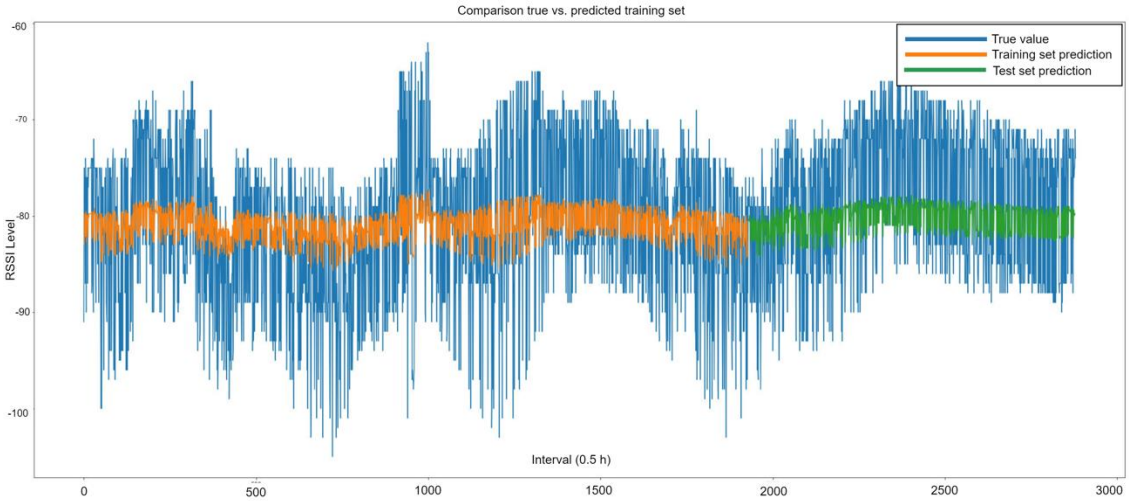


Figure 8: Train and prediction results for pH 5.3 clay soil sample

Here, it was seen that the offset difference between the measured signals and the predicted values found during the training of the model was higher. This difference is due to the fact that the 1D feature map values are lower than the measured RSSI values. However, the results of the artificial intelligence model still follow the pattern of the signal levels.

Numerous theoretical, empirical and semi-empirical methods have been developed to detect the moisture level of soil with active and passive microwave remote

sensing technics. In particular, assimilation of frequently referenced Radiometry and Synthetic Aperture Radar (SAR) data is a new technique used in soil moisture measurements, and there is a lot of research on this subject. The equipment in this type of works is quite large and energy-consuming devices. Today, researchers are constantly developing studies with active and passive microwaves. It can be said that these have some advantages and disadvantages compared to each other. A summary of the advantages and disadvantages of the different methods is available in Table 6.

Table 6: Passive and active microwave sensing methods comparation

Studies	Method	Power	Advantages	Disadvantages
(De Jeu et al. 2014)	Passive Microwave	300 kW	Can penetrate through different objects. Not limited by clouds, dust and daytime conditions, etc.	The energy level being emitted is quite low.
(Sun et al. 2019)		400W		
(Panciera et al. 2014)	Active Microwave	500 W	Detailed spatial resolution, not limited by clouds, dust and daytime conditions, etc.	Complicated analysis, cost-intensive.
(Abdel-Wahab et al. 2019)		10W		
Our Proposal	Active Microwave (BLE)	1mW	Low Power, Low Cost Devices. In-situ detection. Not limited by clouds, dust and daytime conditions, etc.	Low spatial resolution. Best results for acidic soils.

5 Conclusion

In this paper, we have presented the soil moisture content prediction method using LSTM recurrent neural networks. It has shown promising results on acidic soil samples and clay soil sample. We believe that using BLE signal strengths predicting soil moisture content proposed method has contributed to environmental and agricultural improvement. We foresee that the artificial intelligence solution applied in this study will make the method of obtaining information about the soil more widespread with the help of easily accessible and cheap devices.

Today, passive microwave measurements that used to determine soil moisture content are made at wavelengths called L-Band. These measurements are usually carried out via satellites. However, many factors should be taken into account during the remote sensing of soil moisture content. These factors can be listed as soil content, surface hardness, vegetation, backscattering angle of microwaves. There are theoretically and empirically constantly developed approaches to explain the effects of all these components. Even though there are effects to be neglected, even the effects of dipole moment, which will occur by the interaction of water molecules and microwaves in the soil, should be examined to determine the exact value in the measured sites. In this study, we tried to find the water content by the machine learning method by examining the BLE signal behavior in acidic soils without ignoring the importance of all these physical effects. Our results show that the artificial neural

network method we created is feasible.

On the other hand, BLE signals are similar to microwaves defined as L and S bands in satellite measurement systems, which are used frequently today. Although the transmission power of BLE signals is lower than the L and S bands, we think that the applications to be carried out using these signals will vary depending on the water content of the soil in the long term by the method of temporal analysis. With BLE signals, an estimate of water moisture content can also bring economic gain. The market prices of these BLE devices that are incomparable degree with similar microwave radar and measuring devices and methods are quite low. The properties of BLE and other L, S-band microwaves are shown in Table 7 (Sengupta and Liepa 2005).

Table 7: Frequency and wavelength values for L, S microwave bands and simple BLE signal values

Microwave Bands	Frequency Range (GHz)	Wavelength (cm)
L	1 – 2	30 – 15
S	2 – 4	15 – 7.5
BLE	2.4	~12

References

Abdel-Wahab, W., Al-Saedi, H., Ehsandar, A., Palizban, A., Raeis-Zadeh, M., & Safavi-Naeini, S. (2019). Efficient integration of scalable active-phased array antenna based on modular approach for MM-wave applications. Microwave and Optical Technology Letters, 61(5), 1333–1336. https://doi.org/10.1002/mop.31744

Adate, A., & Tripathy, B. K. (2019). S-LSTM-GAN: Shared Recurrent Neural Networks with Adversarial Training. In A. J. Kulkarni, S. C. Satapathy, T. Kang, & A. H. Kashan (Eds.), Proceedings of the 2nd International Conference on Data Engineering and Communication Technology (Vol. 828, pp. 107–115). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-1610-4_11

Allen-Zhu, Z., Li, Y., & Song, Z. (2019). On the convergence rate of training recurrent neural networks. In Proceedings of the 33rd International Conference on Neural Information Processing Systems (pp. 6676–6688). Red Hook, NY, USA: Curran Associates Inc.

Batchu, V., Nearing, G., & Gulshan, V. (2023). A Deep Learning Data Fusion Model Using Sentinel-1/2, SoilGrids, SMAP, and GLDAS for Soil Moisture Retrieval. Journal of Hydrometeorology, 24(10), 1789–1823. https://doi.org/10.1175/JHM-D-22-0118.1

Calla, O. P. N. (2002). Application of Microwave Remote Sensing In Ocean Studies. 2, 623–632. Kochi, India: Allied Publishers.

Carbune, V., Gonnet, P., Deselaers, T., Rowley, H. A., Daryin, A., Calvo, M., … Gervais, P. (2020). Fast multi-language LSTM-based online handwriting recognition. International Journal on Document Analysis and Recognition (IJDAR), 23(2), 89–102. https://doi.org/10.1007/s10032-020-00350-4

Carrière, S. D., Martin-StPaul, N. K., Doussan, C., Courbet, F., Davi, H., & Simioni, G. (2021). Electromagnetic Induction Is a Fast and Non-Destructive Approach to Estimate the Influence of Subsurface Heterogeneity on Forest Canopy Structure. Water, 13(22), 3218. https://doi.org/10.3390/w13223218

Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014, September 2). Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. arXiv. Retrieved from http://arxiv.org/abs/1406.1078

Darroudi, S., Caldera-Sànchez, R., & Gomez, C. (2019). Bluetooth Mesh Energy Consumption: A Model. Sensors, 19(5), 1238. https://doi.org/10.3390/s19051238

Davis, J. L., & Chudobiak, W. J. (1975). In Situ Meter for Measuring Relative Permittivity of Soils. 75-1A. https://doi.org/10.4095/104349

De Jeu, R. A. M., Holmes, T. R. H., Parinussa, R. M., & Owe, M. (2014). A spatially coherent global soil moisture product with improved temporal resolution. Journal of Hydrology, 516, 284–296. https://doi.org/10.1016/j.jhydrol.2014.02.015

Dong, J., Steele-Dunne, S. C., Ochsner, T. E., & Van De Giesen, N. (2016). Determining soil moisture and soil properties in vegetated areas by assimilating soil temperatures. Water Resources Research, 52(6), 4280–4300. https://doi.org/10.1002/2015WR018425

Ertam, F. (2019). An effective gender recognition approach using voice data via deeper LSTM networks. Applied Acoustics, 156, 351–358. https://doi.org/10.1016/j.apacoust.2019.07.033

Gao, T., Gong, X., Zhang, K., Lin, F., Wang, J., Huang, T., & Zurada, J. M. (2020). A recalling-enhanced recurrent neural network: Conjugate gradient learning algorithm and its convergence analysis. Information Sciences, 519, 273–288. https://doi.org/10.1016/j.ins.2020.01.045

Gardner, W., & Kirkham, D. (1952). DETERMINATION OF SOIL MOISTURE BY NEUTRON SCATTERING: Soil Science, 73(5), 391–402. https://doi.org/10.1097/00010694-195205000-00007

Gascho, G. J., Parker, M. B., & Gaines, T. P. (1996). Reevaluation of suspension solutions for soil pH. Communications in Soil Science and Plant Analysis, 27(3–4), 773–782. https://doi.org/10.1080/00103629609369594

Ghori, M. R., Wan, T.-C., & Sodhy, G. C. (2020). Bluetooth Low Energy 5 Mesh Based Hospital Communication Network (B5MBHCN). In M. Anbar, N. Abdullah, & S. Manickam (Eds.), Advances in Cyber Security (Vol. 1132, pp. 247–261). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-15-2693-0_18

H. Ali, M., & K. Ali, N. (2019). IoT based security system and intelligent home automation multi monitoring and control systems. IAES International Journal of Robotics and Automation (IJRA), 8(3), 205. https://doi.org/10.11591/ijra.v8i3.pp205-210

Han, Q., Zeng, Y., Zhang, L., Cira, C.-I., Prikaziuk, E., Duan, T., … Su, B. (2023). Ensemble of optimised machine learning algorithms for predicting surface soil moisture content at global scale [Preprint]. Earth and space science informatics. https://doi.org/10.5194/gmd-2023-83

Hanzlíček, Z., Vít, J., & Tihelka, D. (2019). LSTM-Based Speech Segmentation for TTS Synthesis. In K. Ekštein (Ed.), Text, Speech, and Dialogue (Vol. 11697, pp. 361–372). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-27947-9_31

Haria, A. H., Johnson, A. C., Bell, J. P., & Batchelor, C. H. (1994). Water movement and isoproturon behaviour in a drained heavy clay soil: 1. Preferential flow processes. Journal of Hydrology, 163(3–4), 203–216. https://doi.org/10.1016/0022-1694(94)90140-6

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hu, H., Li, Z., Elofsson, A., & Xie, S. (2019). A Bi-LSTM Based Ensemble Algorithm for Prediction of Protein Secondary Structure. Applied Sciences, 9(17), 3538. https://doi.org/10.3390/app9173538

Hussain, T., Muhammad, K., Ullah, A., Cao, Z., Baik, S. W., & De Albuquerque, V. H. C. (2020). Cloud-Assisted Multiview Video Summarization Using CNN and Bidirectional LSTM. IEEE Transactions on Industrial Informatics, 16(1), 77–86. https://doi.org/10.1109/TII.2019.2929228

Lambot, S., Slob, E., Minet, J., Jadoon, K. Z., Vanclooster, M., & Vereecken, H. (2010). Full-Waveform Modelling and Inversion of Ground-Penetrating Radar Data for Non-invasive Characterisation of Soil Hydrogeophysical Properties. In R. A. Viscarra Rossel, A. B. McBratney, & B. Minasny (Eds.), Proximal Soil Sensing (pp. 299–311). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-90-481-8859-8_25

Luo, D., Wen, X., & He, P. (2023). Surface Soil Moisture Estimation Using a Neural Network Model in Bare Land and Vegetated Areas. Journal of Spectroscopy, 2023, 1–10. https://doi.org/10.1155/2023/5887177

Ma, Z., Wu, B., Chang, S., Yan, N., & Zhu, W. (2023). Developing a physics-guided neural network to predict soil moisture with remote sensing evapotranspiration and weather forecasting [Other]. pico. https://doi.org/10.5194/egusphere-egu23-10597

Mamun, M. A. A., & Yuce, M. R. (2019). Sensors and Systems for Wearable Environmental Monitoring Toward IoT-Enabled Applications: A Review. IEEE Sensors Journal, 19(18), 7771–7788. https://doi.org/10.1109/JSEN.2019.2919352

Martín, F., Vélez, P., Muñoz-Enano, J., & Su, L. (2023). Planar microwave sensors. Hoboken, New Jersey: Wiley-IEEE Press.

Mu, T., Liu, G., Yang, X., & Yu, Y. (2022). Soil-Moisture Estimation Based on Multiple-Source Remote-Sensing Images. Remote Sensing, 15(1), 139. https://doi.org/10.3390/rs15010139

Nagarajan, B., Shanmugam, V., Ananthanarayanan, V., & Bagavathi Sivakumar, P. (2020). Localization and Indoor Navigation for Visually Impaired Using Bluetooth Low Energy. In A. K. Somani, R. S. Shekhawat, A. Mundra, S. Srivastava, & V. K. Verma (Eds.), Smart Systems and IoT: Innovations in Computing (Vol. 141, pp. 249–259). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-8406-6_25

Newman, A. L. (1964). Soil Survey (Vol. 17). US Department of Agriculture, Soil Conservation Service.

Nguyen, T. P., & Songsermpong, S. (2022). Microwave processing technology for food safety and quality: A review. Agriculture and Natural Resources, 56(1), 57–72. Retrieved from https://li01.tci-thaijo.org/index.php/anres/article/view/253973

Noborio, K. (2001). Measurement of soil water content and electrical conductivity by time domain reflectometry: a review. Computers and Electronics in Agriculture, 31(3), 213–237. https://doi.org/10.1016/S0168-1699(00)00184-8

Panciera, R., Walker, J. P., Jackson, T. J., Gray, D. A., Tanase, M. A., Ryu, D., … Hacker, J. M. (2014). The Soil Moisture Active Passive Experiments (SMAPEx): Toward Soil Moisture Retrieval From the SMAP Mission. IEEE Transactions on Geoscience and Remote Sensing, 52(1), 490–507. https://doi.org/10.1109/TGRS.2013.2241774

Paul, I. J. L., Sasirekha, S., Vishnu, D. R., & Surya, K. (2019). Recognition of handwritten text using long short term memory (LSTM) recurrent neural network (RNN). 030011. Kurdistan, Iraq. https://doi.org/10.1063/1.5097522

Pekel, E. (2020). EVALUATION OF ESTIMATION PERFORMANCE FOR SOIL MOISTURE USING PARTICLE SWARM OPTIMIZATION AND ARTIFICIAL NEURAL NETWORK. Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi. https://doi.org/10.28948/ngumuh.529418

Reginato, R. J., & Van Bavel, C. H. M. (1964). Soil Water Measurement with Gamma Attenuation. Soil Science Society of America Journal, 28(6), 721–724. https://doi.org/10.2136/sssaj1964.03615995002800060014x

Ren, G., & Ganapathy, V. (2019). Recognition of Online Handwriting with Variability on Smart Devices. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 7605–7609. Brighton, United Kingdom: IEEE. https://doi.org/10.1109/ICASSP.2019.8682706

Scheberl, L., Scharenbroch, B. C., Werner, L. P., Prater, J. R., & Fite, K. L. (2019). Evaluation of soil pH and soil moisture with different field sensors: Case study urban soil. Urban Forestry & Urban Greening, 38, 267–279. https://doi.org/10.1016/j.ufug.2019.01.001

Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. IEEE Transactions on Signal Processing, 45(11), 2673–2681. https://doi.org/10.1109/78.650093

Sengupta, D. L., & Liepa, V. V. (2005). Applied Electromagnetics and Electromagnetic Compatibility (1st ed.). Wiley. https://doi.org/10.1002/0471746231

Singh, A., & Gaurav, K. (2023). Deep learning and data fusion to estimate surface soil moisture from multi-sensor satellite images. Scientific Reports, 13(1), 2251. https://doi.org/10.1038/s41598-023-28939-9

Sun, H., Cai, C., Liu, H., & Yang, B. (2019). Microwave and Meteorological Fusion: A method of Spatial Downscaling of Remotely Sensed Soil Moisture. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(4), 1107–1119. https://doi.org/10.1109/JSTARS.2019.2901921

Topp, G. C., Davis, J. L., & Annan, A. P. (1980). Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. Water Resources Research, 16(3), 574–582. https://doi.org/10.1029/WR016i003p00574

Wang, T., Zhou, J., Wang, W., Zhang, G., Huang, M., & Lai, Y. (2019). A personal local area information interaction system based on NFC and Bluetooth technology. International Journal of High Performance Computing and Networking, 13(4), 455. https://doi.org/10.1504/IJHPCN.2019.099268

Wu, F., Wu, T., & Yuce, M. (2018). An Internet-of-Things (IoT) Network System for Connected Safety and Health Monitoring Applications. Sensors, 19(1), 21. https://doi.org/10.3390/s19010021

Zárate-Valdez, J. L., Zasoski, R., & Läuchli, A. (2006). SHORT-TERM EFFECTS OF MOISTURE CONTENT ON SOIL SOLUTION pH AND SOIL EH. Soil Science. Retrieved from https://www.semanticscholar.org/paper/SHORT-TERM-EFFECTS-OF-MOISTURE-CONTENT-ON-SOIL-pH-Z%C3%A1rate-Valdez-

Zasoski/ba3aba909b76ba66b9be0cc8bfec3c897ae25f32

Zhang, Y., Qu, C., & Wang, Y. (2020). An Indoor Positioning Method Based on CSI by Using Features Optimization Mechanism With LSTM. IEEE Sensors Journal, 20(9), 4868–4878. https://doi.org/10.1109/JSEN.2020.2965590

Zia, T., & Zahid, U. (2019). Long short-term memory recurrent neural network architectures for Urdu acoustic modeling. International Journal of Speech Technology, 22(1), 21–30. https://doi.org/10.1007/s10772-018-09573-7