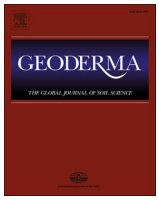
Geoderma 441 (2024) 116765

Contents lists available at ScienceDirect

Geoderma

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



Regional soil water content monitoring based on time-frequency spectrogram of low-frequency swept acoustic signal

Kangle Song [a], Jing Nie [a,b,*], Yang Li [a,b,*], Jingbin Li [a,b], Pengxiang Song [c], Sezai Ercisli [d]

[a] College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China
[b] Key Laboratory of Modern Agricultural Machinery of Xinjiang Production and Construction Corps, Shihezi 832003, China
[c] Engineering Training Center, Xinjiang University, Urumqi 830017, China
[d] Faculty of Agriculture, Ataturk University, Erzurum 25240, Turkey

ARTICLE INFO

Handling Editor: B. Minasny

Keywords:
Soil water content
Sweep frequency
Time–frequency spectrogram
Swin-Transformer
Image regression

ABSTRACT

Acoustic waves offer a non-destructive, safe, and cost-effective means of monitoring the environment, with a potential application in soil water content monitoring. However, extracting soil water information from acoustic signals is still challenging. To tackle this issue, we have developed a low-frequency swept acoustic signal detection device and system. We conducted soil penetration testing using low-frequency swept acoustic signals. The swept-frequency acoustic signals passing through the soil were transformed into time–frequency spectrogram. Using the Swin-Transformer model, we established a regression model between the time–frequency spectrogram of the swept frequencies and the soil water content. Predictions were made both on a laboratory test dataset and through field trials using the calibrated model. The results indicate that the RMSE, MAE, and R^2 values between the observed and the model's outputs of water content (%) for the test laboratory dataset are 0.191, 0.081, and 0.999, respectively, using the Swin-Transformer model. In the case of the field trials, the RMSE, MAE, and R^2 values between the predicted and observed values are 6.715 %, 1.829 %, and 0.711, respectively. These studies demonstrate that this method is highly effective in predicting soil water content, with the best results achieved at a resolution of 20 PPI (Pixels Per Inch) and within the frequency range of 260–360 Hz. It provides an efficient approach for acoustic soil water content detection, effectively resolves the difficulty in building models caused by the single-parameter limitation in traditional acoustic model.

1. Introduction

Against the backdrop of the escalating global water crisis, accurate and cost-effective soil water content detection devices have become key tools for providing services in precision and sustainable agriculture, aiming to optimize water use (Nie et al., 2022a). Soil water content measurement methods encompass drying-weighing, resistance, nuclear magnetic resonance (NMR), neutron, and electromagnetic wave-based approaches. The gravimetric or drying-weighing method is precise but operationally cumbersome. The resistance method is inexpensive and rapid but has a lower accuracy. NMR equipment is costly. The neutron method (Altdorff et al., 2023, Howells et al., 2023) is suitable for long-term field soil water content monitoring but exhibits significant errors in surface soil detection. Additionally, it is expensive and poses potential radiation hazards. Electromagnetic wave detection methods, categorized by wavelength, include the radio wave and microwave-based FDR method (Li et al., 2020), the TDR method (Qin et al., 2023), and the GPR method (Cheng et al., 2023). FDR method is characterized by high accuracy, affordability, and widespread application. The TDR method measures surface soil water content but is costly. However, neither can directly assess regional soil water content. The GPR method can measure a wide range of soil water content but comes with expensive equipment. Remote sensing can monitor large-scale soil water content, utilizing bands such as microwave (Urbina-Salazar et al., 2023), thermal infrared (TIR), shortwave infrared (SWIR), and near-infrared (NIR) for soil water content detection. Among these, microwaves exhibit minimal atmospheric interference and extensive applicability. However, remote sensing can only detect surface soil water content status, influenced by soil texture, bulk density, and vegetation, without the ability to assess deep soil water content. The shorter-wavelength gamma-ray detection method (Taylor et al., 2023) is suitable for regional soil water content detection but is susceptible to soil bulk density. Among the methods for

* Corresponding authors at: College of Mechanical and Electrical Engineering, Shihezi University, Shihezi 832003, China.
E-mail addresses: niejing19@shzu.edu.cn (J. Nie), liyang328@shzu.edu.cn (Y. Li).

https://doi.org/10.1016/j.geoderma.2023.116765
Received 19 October 2023; Received in revised form 27 December 2023; Accepted 27 December 2023
Available online 4 January 2024

regional soil water content assessment, neutron and gamma-ray methods involve radiation risks, GPR is expensive, and remote sensing has limited accuracy, making it challenging for widespread use as a low-cost and continuous measurement method. Although each method has its advantages and limitations, the demand for a low-cost, and environmentally stable regional soil water content detection method is increasingly urgent in modern agriculture. Therefore, there is a need for a more feasible approach that meets the requirements of modern agriculture.

Sound waves, as signals capable of propagating in solid, liquid, and gas phases, have widespread applications in fields such as ultrasonic detection, non-destructive testing, and sonar detection. For soil water content detection, the use of sound waves can be free from the influence of soil chemical properties, holding significant potential for development. Past research has conducted in-depth theoretical analyses of the propagation characteristics of sound waves in soil (Biot, 1956, Brutsaert, 1964), laying the foundation for subsequent model optimization (Xu et al., 2020, Gorthi et al., 2020). Changjie found a high correlation between the attenuation rate of sound waves in soil and soil water content (Chang-jie et al., 2015). However, models constructed using a single speed of sound and attenuation traditionally have shown significant variations across different soil types. (Xu et al., 2021). The use of frequency-sweeping sound signals for soil water detection using attenuation information at numerous frequencies, facilitates in obtaining more precise and robust models. Nevertheless, fitting the extensive data on the relationship between frequency-sweeping signals and water content remains a challenging issue.

In recent years, machine learning and deep neural networks have demonstrated unparalleled advantages in fitting relationships involving multidimensional and large parameters(Li and Chao, 2021), offering new possibilities to address this challenge (Chao and Li, 2022, Li et al., 2022c). Similarly, they have provided new approaches and methods for sustainable agriculture (Nie et al., 2022b, Yang et al., 2022), medical data analysis (Li and Ercisli, 2023), and more. In the field of soil water content prediction, researchers have successfully used the LSTM algorithm for water content forecasting (Datta and Faroughi, 2023, Li et al., 2022b), achieving good results through subsequent improvements (Li et al., 2022a, Kara et al., 2023, Yinglan et al., 2022). Studies also designed a soil layer recognition model based on deep neural networks (Jiang et al., 2021) and predicted soil organic carbon using DNN based on remote sensing images (Odebiri et al., 2022). Other methods include improving the prediction accuracy of soil properties using fractional-order derivative technology combined with one-dimensional convolutional neural networks on near-infrared spectra (Liu et al., 2023). Machine learning has been utilized to identify the spatiotemporal patterns of preferential flow (Kang et al., 2023) and construct a comprehensive drought monitoring model through deep learning (Shen et al., 2019). Deep neural networks have also demonstrated remarkable capabilities in the extraction and classification of acoustic features, exceptional performance in anomaly detection in time series data (Choi et al., 2021), and direct feature recognition using sound data (Romero et al., 2019). Hasan proposed deep neural network-based spectral imaging approach achieved a highly accurate classifier for fault diagnosis in sound emission signals (Hasan et al., 2019). Deep neural networks exhibit outstanding capabilities in the prediction of soil properties and the construction of complex models for acoustic signal analysis.

Combining the multiple-parameter features of frequency-sweeping signals with the powerful modeling capabilities of deep neural networks, this study adopts frequency-sweeping sound signals for soil transmission detection. It reflects soil water content through attenuation information at multiple frequencies and constructs a regression model for the relationship between water content and time–frequency spectrograms using a deep learning network. This achieves regression predictions for time–frequency spectrogram of different soil water content levels. This study has theoretical and practical value for addressing the limitations of existing soil water content detection methods. The main research contributions are as follows: (1) We introduce a method for soil water content detection based on time–frequency spectrogram of time-swept acoustic signals. (2) We design a set of equipment and a system for time-swept acoustic signal detection. (3) We establish a regression model for the relationship between time–frequency spectrogram and soil water content using the Swin-Transformer algorithm and compare the impact of different PPI and frequency ranges on prediction results.

2. Principle and methods

2.1. Principle of soil water content detection by sweeping acoustic signals

Different frequency sound waves exhibit varied attenuation characteristics under distinct soil water content conditions. Changes in soil water content induce variations in inter-particle forces, leading to differences in the attenuation of sound waves. Establishing a connection between sound signal attenuation and soil water content is achievable. However, the attenuation of traditional single-frequency sound signals struggles to adequately capture soil water content characteristics. In contrast, a sweep of sound waves over a certain duration can encompass multiple frequencies, allowing for the conveyance of more information in soil water content detection. Therefore, this paper adopts a frequency-sweeping transmission signal for soil water content measurement. The time–frequency spectrogram of sound signals penetrating the soil reflects the energy attenuation levels of different frequency sound waves, establishing a relationship with water content. The detection principle is illustrated in Fig. 1.

2.2. Detection method and model construction process

The process of using transmissive sweep-frequency acoustic signals for soil water content detection is divided into five parts: sweep-frequency signal construction, transmissive detection in soils with varying water content levels, data processing, model training, and result prediction. The specific process is illustrated in Fig. 2.

2.2.1. Construction of swept acoustic signal

In this study, sweep-frequency acoustic signals are generated by the Analog Discovery 2 portable oscilloscope. The waveform generator of the oscilloscope can produce diverse and high-quality waveform signals. The computer allows customization of waveform types, sweep time, and signal amplitude. Once the waveform is defined, it can control the oscilloscope to output the designed sweep-frequency signal. The custom sweep-frequency signal is then amplified by an amplifier, and the amplification factor remains constant throughout the measurement process. After amplification, the sweep-frequency acoustic signal drives the low-frequency speaker to produce acoustic, thus forming a high-definition sweep-frequency acoustic source.

The effective determination of the detection frequency range is a crucial step in using frequency-sweeping acoustic signals for soil water content detection. Experiments were conducted with acoustic sweeps in the 20–1500 Hz range for transmission testing. The attenuation of the frequency-sweep acoustic signal at a detection distance of 850 mm is shown in Fig. 4, where the color depth represents the distribution of residual energy in various frequency bands of the spectrogram. It can be observed that the signal energy is relatively high in the 50–550 Hz range. Excessive signal attenuation can significantly reduce the effectiveness of information carried by the signal. Therefore, we determined the efficient frequency-sweep detection range to be 50–550 Hz. Additionally, during the analysis of frequency-sweep energy, we discovered minimal attenuation in the signal around 310 Hz. As a result, we constructed continuous frequency-sweep signals in the range of 50–550 Hz for 10 s and concise narrow-range frequency-sweep signals in the 260–360 Hz range for 2 s with 310 Hz as the center frequency. This approach fully utilizes numerous parameters to reflect soil moisture content while exploring the signal performance of a small frequency-

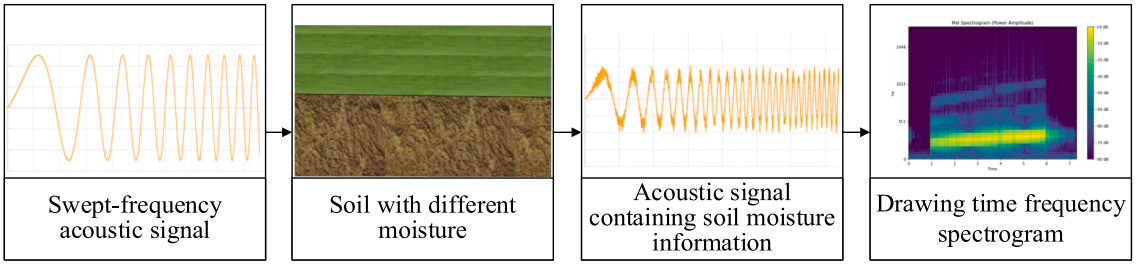


Fig. 1. Principle of detecting soil water content by sweeping frequency acoustic signal.

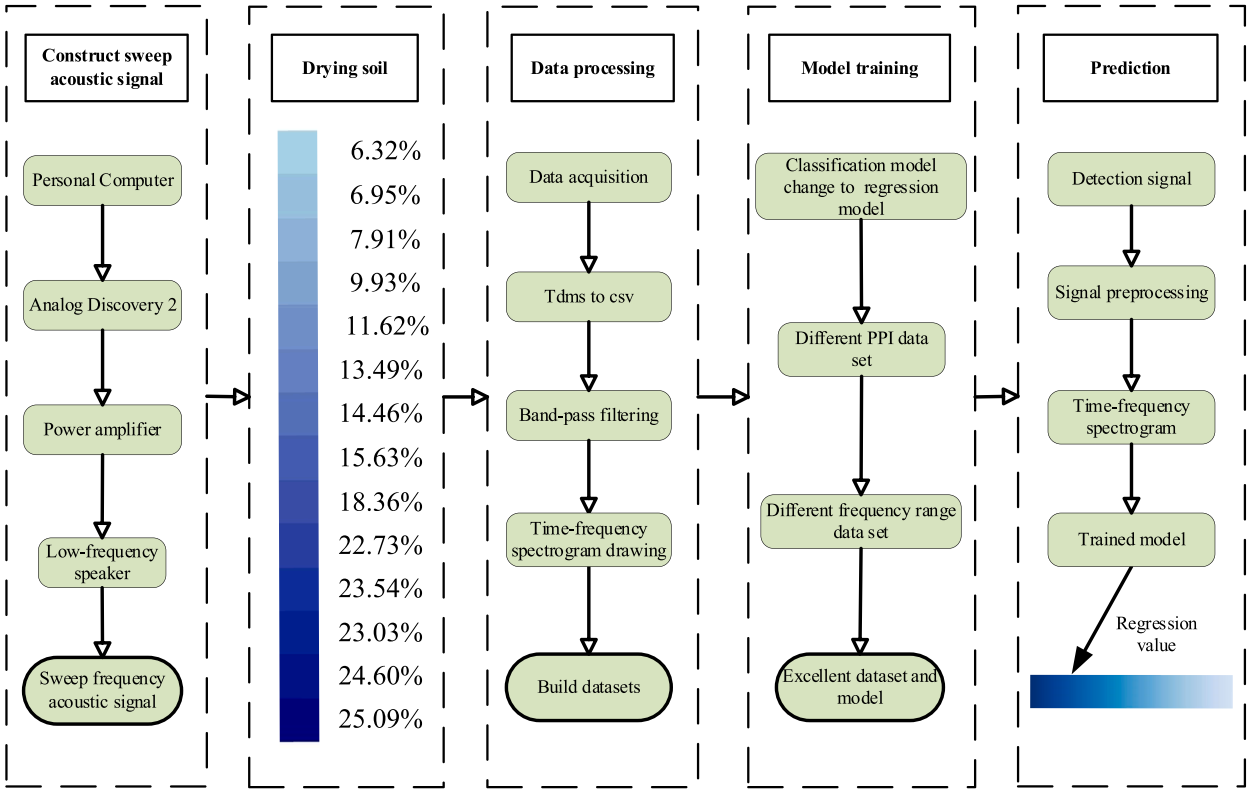


Fig. 2. Method and process diagram of detecting soil water content by sweeping frequency acoustic signal.

sweep interval. Ensuring consistent frequency density, we constructed wide and narrow frequency ranges. During data collection, we acquired data separately using these two frequency ranges to determine the impact of frequency range on the overall detection results.

2.2.2. Soil drying condition and detection process

The detection process has taken place in a constant temperature and quiet acoustic environment. The main soil texture under investigation consists of 13.20 % clay, 34.44 % silt, and 53.36 % sand, placed in a soundproofed soil tanker measuring 1000 × 300 × 300 mm. The research examines the variation in soil water content starting from (25.09 % mass water content) and continuing until the soil water content becomes unavailable for plant uptake (6.32 % mass water content). To capture the gradual decrease in water content, the study conducts measurements at 14 different water content states during continuous drying. Water content calibration at each stage is performed through multi-point sampling and drying. Immediately after the measurements, soil samples are collected from five evenly distributed points within the measurement range, with a depth of 200 mm. The soil water content of these samples was determined using the MA100Q moisture analyzer, utilizing the drying method. The average of the five-point measurements represents the soil water content for that specific measurement area. Considering the inevitable signal fluctuations in low-frequency speakers and the entire acoustic production system, multiple measurements were taken at each water content level to create a large dataset, allowing the model to distinguish between various states.

2.2.3. Data processing and time–frequency spectrograms generation

The sound signals transmitted through the soil are captured by the acoustic sensor, and the data acquisition card converts the received analog signals into tdms data. A Python algorithm is employed to systematically extract the sound wave data from the first channel of the tdms data, converting it into a CSV format. The sound wave data undergoes bandpass filtering to eliminate noise from both low and high-frequency ranges. Subsequently, using the Python audio processing library librosa, operations such as windowing, Fourier transformation, and power spectrogram computation are applied to the time-domain audio, ultimately producing the time–frequency spectrogram.

The time–frequency spectrogram of the frequency-sweeping sound signals illustrates the frequency components and their corresponding energy levels at different moments. During the spectrogram creation, the results of the Fourier transformation are subjected to modulus

operations, followed by logarithmic transformation into decibels (dB). The final presentation is in the form of an amplitude graph. Considering the frequency sweep, this study adopts the Mel spectrogram, offering higher resolution in the low-frequency region and gradually decreasing resolution as the frequency increases. The Mel spectrogram creation involves the following steps: first, the signal is framed into multiple small-time windows; next, a Hamming window function is applied to mitigate boundary effects; then, fast Fourier transformation is performed to obtain the spectral information for each window. The amplitude spectrum is converted to power spectrum to enhance spectral features, and Mel filters are utilized for filtering, particularly to enhance the resolution in the low-frequency range. The specific steps are outlined below:

1.Signal Framing: The audio signal is divided into multiple small time windows, or frames;

2.Windowing: A window function is applied to each time window to reduce boundary effects;

3.Fast Fourier Transform (FFT): The FFT is performed on each window to obtain the spectral information for each window;

4.Power Spectrogram: The amplitude spectrum is converted into a power spectrum to enhance spectral features;

5.Mel Filtering: The power spectrum is filtered using a set of Mel filters to enhance resolution in the lower-frequency region.

The specific formulas are as follows:

The Hamming window function is shown in Eq. 1

$$\omega(n) = 0.54 - 0.46 \cdot \cos\left(\frac{2 \cdot \pi \cdot n}{M-1}\right) \tag{1}$$

In this context, $\omega(n)$ represents the value of the window function at index n, where n denotes the index of the window's sampling point, and M represents the length of the window.

The Discrete Fourier transformation formula is shown in Eq. 2

$$X(\mathrm{m}) = \sum_{\mathrm{n}=0}^{N-1} \mathrm{x}(\mathrm{n}) \mathrm{e}^{-j \cdot 2\pi \cdot n \cdot m / N} \tag{2}$$

Spectrogram power spectrum is calculated by taking the square of the mode for each window power spectrum to get the power spectrum. As in Eq. 3

$$SP_X(m) = \left| \sum_{\mathrm{n}=0}^{N-1} \mathrm{x}(\mathrm{n}) \mathrm{e}^{-j \cdot 2\pi \cdot n \cdot m / N} \right|^2 \tag{3}$$

The Mel-filter bank is used for filtering. The more Mel filters in the filter bank, the higher the frequency resolution, but it also increases computational cost and processing time. Considering various factors, in this experiment, the parameter for the number of Mel filters is set to 128. Mel filter banks are used to obtain the Mel spectrogram. The Mel filter bank used in this study is based on the Slaney-style filter bank. Its purpose is to map the linear frequency range to the Mel scale range and evenly distribute these filters on the Mel scale according to the desired number of MEL filters. The Mel scale transformation formula for this type is as follows, as shown in Eq. 4.

$$Mel(l) = 2595 \cdot \log_{10}\left(1 + \frac{f}{700}\right) \tag{4}$$

In this paper, all time–frequency plots are generated using Python and libraries such as librosa. The process involves extracting audio signal data in CSV format and then creating time–frequency plots. An example of such a time–frequency plot is illustrated in Fig. 3.

Data processing was accomplished through the implementation of Python algorithms to perform batch calculations and plotting. The temporal data volume of the 10-second swept-frequency acoustic signal reached 64 MB. Even after signal processing and data transformation, the lossless data volume remained substantial. Consequently, when generating time–frequency spectrograms, custom resolutions were employed for saving to mitigate data size. Subsequently, datasets essential for further analysis were constructed, and a clear demarcation between training and validation sets was established.

2.2.4. Model training and regression prediction

In consideration of practical applications involving lightweight and computational cost, this study has undertaken modifications to the lightweight Swin-Transformer deep neural network for image regression

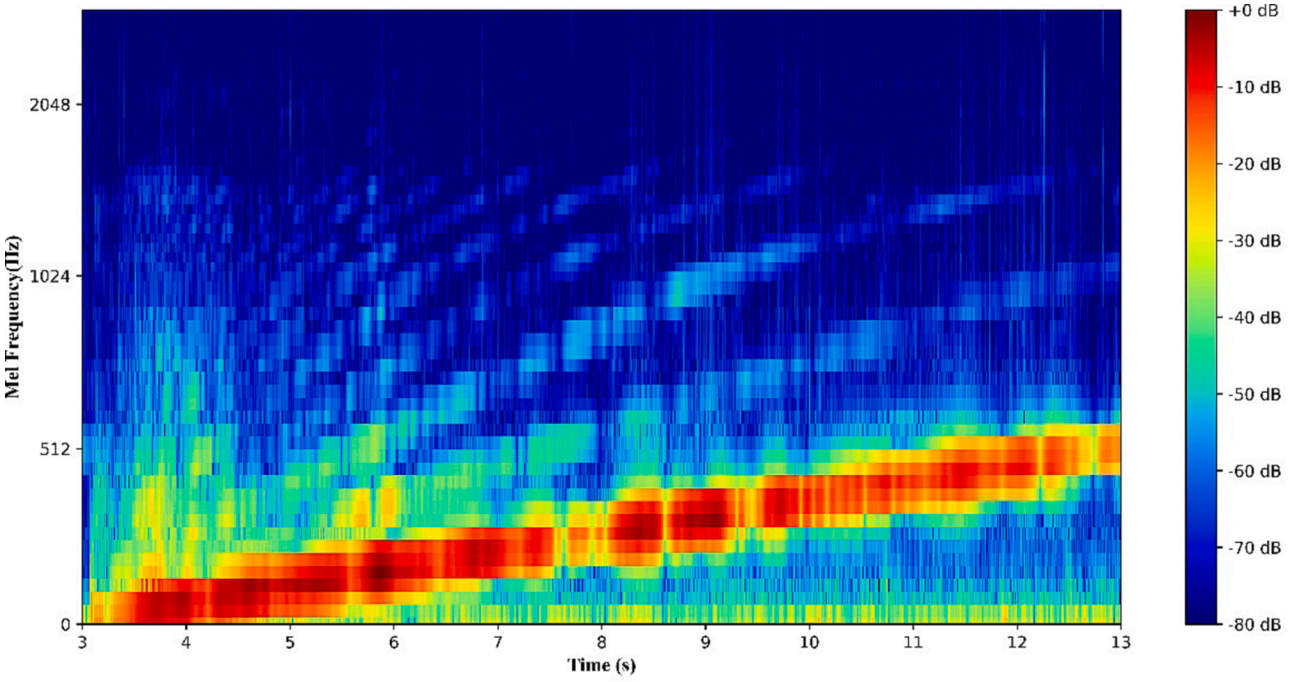


Fig. 3. Mel time-spectrogram plotted in Python. This figure is a Mel-time–frequency spectrogram drawn by python, where the horizontal axis is time, the vertical axis is the frequency, and the color depth in the figure represents the energy of a frequency component in a certain short time interval. The light area is the main frequency part of the time–frequency spectrogram, and the other upper multiple color bands are multifold spectral regions.

tasks. The selected models underwent subsequent training based on pre-trained model parameters. Training was conducted on the training set of time–frequency spectrogram datasets with varying soil water content, resulting in the development of a high-performing model and the identification of key parameters for dataset construction.

During prediction, the detection signals underwent signal processing with optimal parameters. Simultaneously, the signals were processed using dataset construction methods that yielded superior prediction results. The spectrograms were then input into the trained model to obtain soil water content levels.

3. Experimental setup

3.1. Soil material characteristics

The soil used in this study was obtained from a crop experimental field located at N44.365° latitude and E86.066° longitude in northwestern China. The soil texture for this study was determined using the hydrometer method. Multiple samples were collected and analyzed, yielding average results of 13.20 % clay, 34.44 % silt, and 53.36 % sand. The average soil bulk density was determined to be 1.48 g cm^{-3} using the cutting ring method. Soil water content measurements were conducted using the MA100Q moisture analyzer. This instrument employs the principle of the drying-weighing method. During the detection process, samples were dried at 105℃, and the water content was automatically calculated. The instrument is equipped with built-in weights and a repro TEST function, enabling self-calibration before each measurement. It boasts high precision, with an accuracy of 0.1 mg/0.001 %.

Soil samples were collected using a dedicated 8 mm diameter sampling tube at a depth of 150 mm. Real-time calibration of the data was performed using the detector. Throughout the data collection process, no secondary irrigation or overall disruption and reconstruction of the soil structure were conducted, aiming to maintain consistency with the natural drying process of agricultural field soil.

3.2. Detection devices and systems

The composition, specific connections, and functions of the detection system are illustrated in Fig. 5. Prior to testing, the low-frequency loudspeaker and the sound signal receiver were positioned at opposite ends of the soil under test. A computer controlled a digital waveform generator to produce the required swept-frequency acoustic signal. The digital oscilloscope produced a custom signal to the power amplifier, which linearly amplified the signal. Subsequently, the power amplifier drove the low-frequency loudspeaker to emit the designed swept-frequency acoustic wave signal. The sound signal receiver on the other side of the soil under test captured the transmitted swept-frequency acoustic signal through the soil. A data acquisition card collected the waveform of the sound waves after passing through the soil. The computer performed data conversion, filtering, and generated time–frequency spectrograms from the collected signals. Finally, the time–frequency spectrograms were inputted into a pre-trained model to obtain prediction of water content.

The specific models and main parameters of the experimental equipment are as follows: the preamplifier is the AWA14604 by Aihua, with a detection frequency range of 10–100 k Hz, and the sound signal receiver is model AWA14412. The swept-frequency acoustic signal is generated using the digital oscilloscope function of the Analog Discovery 2. The power amplifier is the BORIZSONIC PA-01, with a maximum power of 100 W. The loudspeaker is a subwoofer by Rei-fan Acoustics, with a rated power of 80 W, a frequency response range of 43–2.5 k Hz, and a sensitivity of 88 dB ± 3 dB. The data acquisition card is the VK702-Pro USB high-speed data acquisition card, featuring 8-channel preamplifier inputs, 14-bit resolution, and a maximum sampling rate of

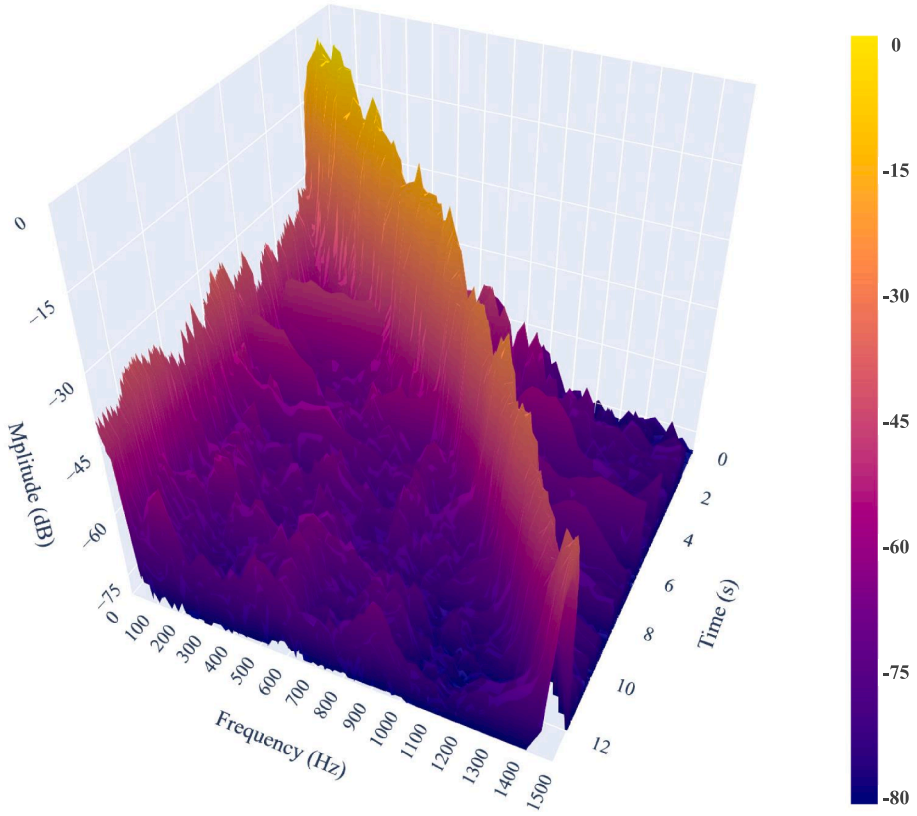


Fig. 4. Energy map of soil 20–1500 Hz transmission swept frequency acoustic signal(The figure represents the spectrum value and its energy intensity at a certain small time period).

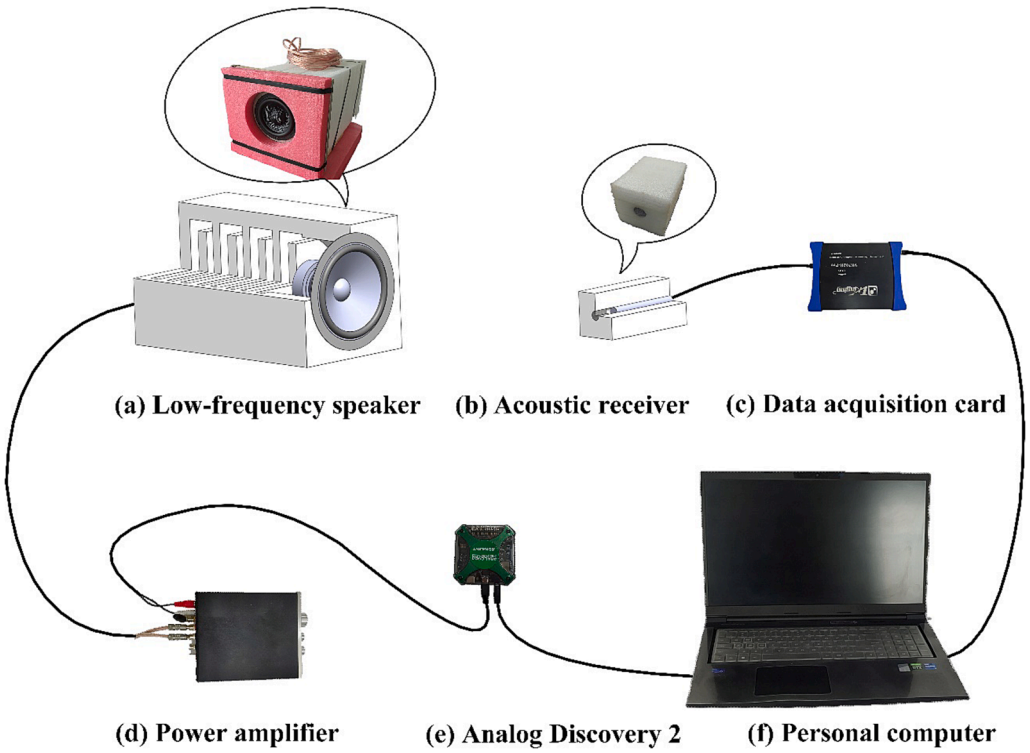


Fig. 5. Diagram of the data acquisition system. (a)Low-frequency speaker. The low-frequency speaker is driven by a power amplifier to play the swept frequency signal output by analog discovery2. (b)Acoustic receiver. Acoustic receiver receives the swept frequency acoustic signals with different humidity in the soil sample. (c)Data acquisition card. The data acquisition card converts the analog signal of the acoustic receiver into a digital signal for transmission to a personal computer. (d) Power amplifier. The amplifier receives the signal from the Analog discovery 2 to drive the speaker to emit acoustic waves. (e)Analog Discovery 2. The digital oscilloscope function of analog discovery 2 was used to construct the desired swept frequency signal. (f)Personal computer. Define the waveform emitted by analog discovery 2, control the acquisition card to receive and save data.

100ksps.

3.3. Structure of the soil tanker

A soil testing chamber was constructed to build a dataset for detecting soil water content using frequency-swept acoustic signals and optimizing related parameters. The experimental soil chamber used is depicted in Fig. 6 and consists of stainless-steel grids, acoustic-permeable fabric, waterproof film, soundproof sponge, and structural fasteners (note that some simplifications were made in the diagram for the grids, acoustic-permeable fabric, and waterproof film). The testing platform was divided into four layers from the inside out. The outermost layer is wrapped with PE foam to isolate environmental noise interference. There was only one circular opening on the speaker side to allow the acoustic waves generated by the speaker to pass through, while the other five sides were fully enclosed to isolate acoustic.

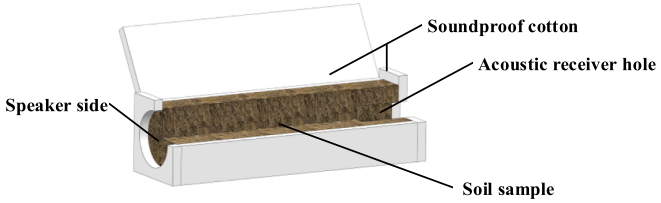


Fig. 6. The main components of the experiment soil platform. (The above picture omits the stainless-steel mesh structure, acoustic permeable cloth and waterproof membrane) Sound insulation cotton insulates the leakage of acoustic signal interference, and the side of the speaker is hollow to facilitate the passage of acoustic signals. The acoustic signal receiving hole is used to collect acoustic signals under the premise of noise isolation of the corresponding acoustic signal receiving components.

The inside of the soundproofing cotton is embedded with a double-layer plastic film to prevent water from seeping out of the soil chamber. Multiple stainless-steel grids were fastened together with zip ties to form the framework of the soil chamber. Compared to a solid plate-like structure, the architecture made of stainless steel grids had a smaller surface area that can weaken the propagation of acoustic waves, allowing for better air circulation and water permeability, making experiments easier to conduct. The medium with a significant difference in hardness and softness between the air and the grids acts as a barrier to the conduction of acoustic signals. Connecting multiple stainless-steel grids with zip ties, rather than welding them, reduces the transmission of acoustic signals. The dimensions of the stainless-steel chamber are 1000 × 300 × 300 mm. An acoustic-permeable fabric was laid around the inner perimeter of the stainless-steel grids' structure. The chamber was filled with the soil to be tested. The acoustic-permeable fabric allows acoustic waves to pass through but confines the soil within the chamber. The combination of the acoustic-permeable fabric and the stainless-steel grids created a complete chamber structure. On one side of the chamber, there was the speaker's acoustic emission side, facing the soundproof cotton with a circular opening reserved for it. The inner grids and soundproof fabric form the side of the chamber without soundproofing treatment. On the other side, within the plastic film and the stainless-steel grid, there was a central 100*100 mm square hole. The square hole was excavated inward by 150 mm, surrounded by a certain-sized acoustic-permeable fabric for support, and separated from the soil by acoustic-permeable fabric. The actual testing distance at the square hole is 850 mm. A slightly larger square-shaped PE foam was embedded in the square hole, with a smaller hole in the center, slightly smaller than the acoustic receiver. During measurements, the acoustic receiver was inserted into the small hole to receive the frequency-swept acoustic signals passing through the soil, even in the presence of

significant noise isolation.

3.4. Data collection and dataset construction

Waterproof measures were taken on the constructed experimental platform, and the soil was irrigated until it reached saturation. Subsequently, with proper measures to prevent evaporation, the excess free water was drained. Depending on the soil type and while maintaining a certain temperature and anti-evaporation measures, the soil was allowed to stand for 1.5 days to reach the maximum water-holding capacity state. At this point, water no longer freely seeped out. The detection system's acoustic speaker and acoustic signal receiver were positioned on opposite sides of the experimental soil chamber. The designed frequency-swept acoustic signal was played through the low-frequency speaker, while simultaneously, the computer controlled the data acquisition card to collect the acoustic signals passing through the soil. This process constituted one data collection cycle. To avoid signal fluctuations caused by dynamic changes during each cycle of playing and collecting the frequency-swept acoustic signal, multiple collections were performed for the same water content. All measurements were conducted at an average ambient temperature of 20.5 ℃. The detection process was maintained in a quiet acoustic environment with an average sound level of 23.7 dB, avoiding interference from various types of noise.

After the data collection was completed, soil samples were immediately taken from five evenly distributed points within the soil chamber. These samples were dried and tested using the portable MA100Q moisture meter. With this instrument, we can easily measure the mass water content of the sampled soil. The average value of three samples was used to calibrate the soil water content at that moment, thereby completing the acquisition of raw data for one water content dataset.

After allowing the soil to naturally dry for a period, the same procedure was repeated. In the end, a total of 3500 frequency-swept acoustic signal data points were obtained across 14 different water content levels. After generating time–frequency plots and removing data points with significant issues, 2776 valid data points remained. These data points were used to construct a dataset containing acoustic signal data for different PPI and frequency-swept segments. The dataset was divided into a testing set and a training set in a 2:8 ratio. The data collection frequencies and corresponding water content values are detailed in Table 1.

3.5. Model and training platform

In order to minimize computation time and save computational resources, a lightweight classification network called Swin-Transformer was chosen for model construction. Transformer architecture has shown excellent performance in various fields such as speech recognition, machine translation, and image processing (Vaswani et al., 2017, Han et al., 2022). In the field of speech recognition, its fundamental principle is to convert sequential audio signals into spectral information for feature extraction. The spectrograms used in this study share information characteristics similar to speech data, and the algorithm's related structures are suitable for learning and extracting audio spectral information (Nie et al., 2023). Additionally, this architecture easily incorporates global information into the analysis (He et al., 2022), which is crucial as the attenuation characteristics of multiple frequencies in time–frequency plots are distributed across the entire image. Therefore, this structure is suitable for feature extraction from time-frequency plots.

Furthermore, the Swin-Transformer employs the Shifted Windows Multi-Head Self-Attention (SW-MSA) structure, which facilitates information transfer between adjacent windows, benefiting the connectivity of information in time–frequency plots. This combination of Shifted Windows and Self-Attention has demonstrated excellent performance in deep feature extraction in image processing (Huang et al., 2022, Gao et al., 2022, Ma et al., 2022, Üzen et al., 2022). Hence, this model was chosen for feature extraction from time-frequency plots to perform regression on soil water content values. Of course, modifications were made to the model's input, loss function, and output to enable it to perform image regression on time–frequency plots related to soil water content.

The entire model training platform was based on the Windows 11 operating system, with a GPU of NVIDIA GeForce RTX 3060, a processor of Intel Core i7-12700H 2.3 GHz, and 16 GB of RAM. All model construction and training, as well as validation, were implemented using the Python language. The models were built on the Pytorch deep learning framework, and the development tool used was PyCharm. The research employs Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2) as evaluation metrics. The model was assessed based on these four metrics to evaluate its performance.

3.6. Field experiments

To validate the real-world effectiveness of using time–frequency plots for on-field soil water detection, multiple field experiments were conducted in the experimental fields where soil samples were obtained for this study. Soil water content was tested at five different time points. During the tests, precautions were taken to protect the speaker and acoustic receiver from dust. They were buried at depths of approximately 250 mm in different positions within the field, ensuring that the center of the speaker and the acoustic receiver were located about 200 mm below the ground surface. To reduce the impact of excessive attenuation during the propagation of acoustic waves on the test results, the reference distance for these validation experiments was set at 850 mm, following the measurements conducted on the experimental platform. Subsequent applications will require data collection at various distances to meet different detection requirements.

The experimental field was tested multiple times, and parameters such as actual testing depth and testing distance were recorded. Transmittance tests of soil were conducted using frequency-sweeping acoustic signals, and multiple sets of acoustic signal data were collected. After the tests, soil samples were immediately taken from the vicinity of the testing area, at a depth of approximately 200 mm, at five evenly distributed points. Soil mass water content was determined using an MA100Q moisture analyzer through the drying method. The average of the water content values at these five points was considered as the soil water content for the testing area.

By using the model to analyze the time–frequency plots of soil at this water content level, detection and identification results were obtained. Through a comparison between the model's identification results and the observed water content values, the practical effectiveness of using frequency-sweeping acoustic signals' time–frequency spectrogram for soil water content detection was verified.

4. Results

4.1. Prediction results of Swin-Transformer

The Swin-Transformer model takes images with an input size of

Table 1
The continuous variation values of 14 types of soil water content.

Number of experiments	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Soil water content (%)	25.09	24.60	23.03	23.54	22.73	18.36	15.63	14.46	13.49	11.62	9.93	7.91	6.95	6.32

224x224. When different images are input, they are resized before being fed into the model. Images with a size close to the input image size, specifically 200x200, were used with various models. The detection results on the test set are shown in Fig. 7.

From Fig. 7, the performance of the Swin-Transformer model on the test set reveals that the predicted values are relatively concentrated. The predicted values for the 14 water content levels exhibit minimal fluctuations compared to the observed values, with most predictions closely aligning with the y = x line. The statistical summary of the prediction results is presented in Table 2. Multiple evaluation metrics demonstrate excellent performance.

To validate the effectiveness of the method and model across various soil types, we conducted a similar training process on a small dataset of sandy soil and loamy clay using the model parameters trained on the original dataset. The soil compositions were as follows: Sandy soil (clay content 5.42 %, silt content 7.46 %, sand content 87.12 %), loamy clay (clay content 35.45 %, silt content 19.41 %, sand content 45.14 %). The statistical results of the model's predictions on the experimental soil and the predictions on the other two soil types are shown in the first two rows of Table 2. Additionally, we compiled the results of a study conducted by Cao and Xiao on the observed performance of eight types of FDR soil moisture detectors (Cao and Xiao, 2024). Cao and Xiao tested the performance of EC-5, GS-1, MAS-1, 5TM, 5TE, TEROS-11, TEROS-12, and SMT-100 soil moisture detectors. These eight detectors were developed by two major manufacturers, METER (USA) and Truebners (Germany). They conducted tests on sandy soil (with a physical composition of clay 0.25 %, silt 7.87 %, sand 91.88 %) and calibrated the measured soil's true water content using the drying-weighing method. The detection capabilities of the eight devices were evaluated. We compiled the performance of the eight detectors under various moisture levels, presenting the ranges of their RMSE, and R^2 in Table 2 (Cao and Xiao expressed it in volumetric water content, and for fair comparison, we converted the results).

From the specific predictions and statistical results on the test set, it is evident that the Swin Transformer algorithm performed well overall in predicting results for various soil test sets. Most predicted values closely align with the y = x line. The RMSE value of the model's predictions on the experimental soil test set is 0.191 %, indicating that the model achieves an accuracy of 0.191 % mass water content in the test set. The RMSE value for predictions on the other two soils is also 0.285 %. Compared to Cao You-song's statistical results for eight common FDR soil moisture detection devices (ranging from 0.453 % to 2.717 %), this

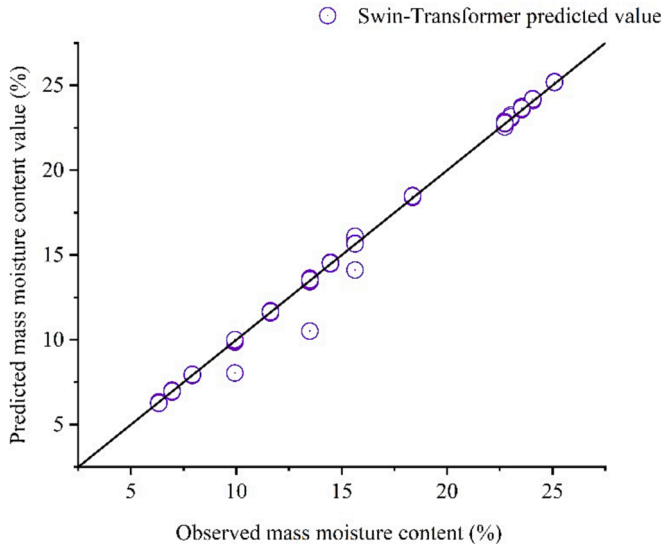


Fig. 7. Comparison of predicted and observed results of Swin-Transformer model on the test set.

Table 2
Predictive statistics of the model on three soil types and FDR method (MAE considers absolute values, making it more stringent compared to MBE.)

Detection Method.	MAE/MBE(%)	RMSE(%)	R^2
Results of experimental soil dataset	0.081(MAE)	0.191	0.999
Results of other soil types dataset	0.109(MAE)	0.285	0.995
Range eight FDR results	0.261 ~ 2.603 (MBE)	0.453 ~ 2.717	0.417 ~ 0.924

Please check Figs. 3 and 4 orders in MS (Interchanged ie, 1st Fig. 4 and then Fig. 3).

method demonstrates higher accuracy on the experimental soil and the other two soil test sets than the commonly used FDR methods. This suggests that this method can achieve high accuracy and has great detection potential.

The results indicate that the soil moisture detection method, combining deep neural networks with frequency-sweeping acoustic signals, can be applied to various soil types, and the model performs well on the experimental soil dataset. Compared to FDR detection methods, this method demonstrates certain advantages with the support of large datasets and deep neural networks. The Shifted Windows combined with Self-Attention structure in the Swin Transformer model exhibit strong feature extraction capabilities for time–frequency spectrograms. The characteristics of global information and information transfer between adjacent parts in this type of structure enhance its feature extraction and model fitting capabilities.

4.2. *The results of different swept-frequency ranges*

In this study, swept-frequency acoustic signals were generated using a digital oscilloscope. Two sets of signals, one ranging from 50 to 550 Hz lasting for 10 s and another from 260 to 360 Hz lasting for 2 s, were employed for experimental soil detection. While maintaining the same frequency density, the relationship between wide and narrow frequency bands and algorithm recognition was observed. Separate training was conducted for each frequency band, and the trained models were then applied to the test set.

The regression results of the time–frequency spectrograms and corresponding label values were analyzed for both frequency bands. Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and R-squared (R^2) values for the two frequency ranges were calculated. The results of the four-evaluation metrics are depicted in Fig. 8.

From Fig. 8, we can observe that the model demonstrates favorable results in the recognition of both frequency ranges. The dataset composed of swept-frequency acoustic signals in the 50–550 Hz frequency range exhibits relatively better performance in three out of four evaluation metrics, with only the MAE parameter slightly inferior to the 260–360 Hz frequency range. This suggests that the recognition of time–frequency spectrograms in the broader 50–550 Hz frequency range is slightly superior to that of the 260–360 Hz range. On the test set, the RMSE for the two frequency ranges reached 0.173 % and 0.286 %, respectively.

The results indicate that both frequency ranges yield satisfactory predictive outcomes, with slight differences that are not markedly significant. In terms of the entire detection process, the 2 s data in the 260–360 Hz range, compared to the 10 s data in the 50–550 Hz range, can reduce computational costs by approximately 80 % in the initial data processing stage. Despite the marginal disadvantage in predictive results, the advantage in computational cost reduction is particularly noteworthy. This is significant for practical deployment and application, contributing to a substantial reduction in computation and cost. The RMSE values for both frequency ranges suggest that the predictive capability of this method can achieve an accuracy of 0.17 % to 0.28 % in estimating water content.

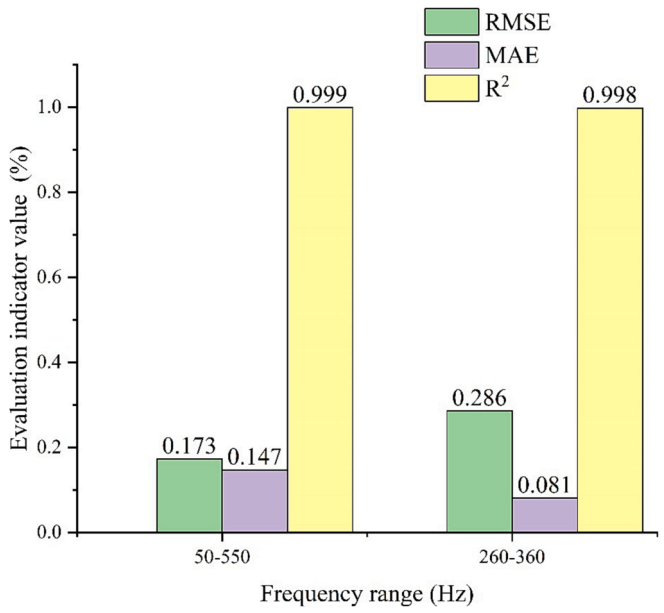


Fig. 8. Two frequency bands predicted data outcomes on Swin-Transformer models.

4.3. The results of different PPI

The value of the PPI parameter is crucial for the model input, as it not only determines the scale of the entire dataset but also influences the computational speed of the model. Therefore, datasets with five different resolutions, namely 10 (100x100), 20 (200x200), 30 (300x300), 40 (400x400), and 50 (500x500), were constructed. The Swin-Transformer model was trained on datasets corresponding to the PPI values, resulting in the corresponding model parameters. Predictions were made on the corresponding test sets. The prediction results and statistical summaries are presented in Fig. 9, with R^2 values consistently above 0.99, hence not explicitly shown in the figure.

From the prediction results, we observe that the predicted values for the five different PPIs mostly aggregate near the y = x line. There is a noticeable improvement in the prediction results for the 20 PPI (200 × 200) resolution. The models trained on the 10 PPI and 50 PPI datasets show significantly inferior predictive results compared to the 20 PPI, 30 PPI, and 40 PPI datasets. Excessively large or small PPI values both result in relatively weaker recognition performance. The resolution of the 10 PPI dataset drops sharply, while the changes in the 30 PPI, 40 PPI, and 50 PPI resolutions are more gradual.

This indicates that the transformation of original time-series data into time–frequency spectrograms with smaller PPI values is highly effective in preserving features. The algorithm is capable of training robust models based on this transformation. On the one hand, the results from the 10 PPI dataset suggest that datasets with excessively low resolutions significantly impact model construction and accuracy. In comparison, the datasets with 30, 40, and 50 PPI have a milder impact on the model. This implies that the model resizing process is less friendly to low-resolution data compared to high-resolution data. The application of this method requires careful consideration of the crucial value of PPI according to the algorithm's requirements. On the other hand, the process of converting the initial large-volume swept-frequency acoustic signal time-series data into time–frequency spectrograms, followed by feature extraction through windowing and Fourier transformation, and then saving as images with custom PPI, effectively and significantly reduces the data volume while ensuring sufficient features. The saved data values of the images are in a quadratic relationship with the PPI values. The file sizes for single images at 50, 40, 30, 20, and 10 PPI are only 59.5 KB, 43.1 KB, 26.8 KB, 13.9 KB, and 5.1 KB, respectively. Taking the optimal 20 PPI as an example, each image's data volume is only 13.9 KB, reducing the original 63 MB time-series data storage by a factor of 4641.

This method, which combines large-volume data detection with small-volume data recognition, significantly reduces data volume while ensuring effective features. In practical applications, the preprocessing of large-volume data can be rapidly completed by front-end hardware. By reducing data volume through this method, small-volume images can be transmitted to the cloud via the Internet of Things (IoT), where more powerful and larger models on the cloud can yield more accurate results. Existing IoT communication services are fully capable of handling the data volume of 13.9 KB images.

4.4. Field experiments results

In order to validate the practical application effectiveness of this method, field validation experiments were conducted on sampled soils in the field. Experiments were conducted at five different water content levels. Detection was carried out at a distance of 840 mm in a manner illustrated in the abstract. The scanning frequency range was set to 260–360 Hz, and time–frequency spectrograms were generated using a 20 PPI for the swept-frequency results. The results are depicted in Fig. 10.

The model's performance in the field experiment is slightly inferior to that on the test set, but overall, the predictions still cluster around the y = x line. During the field experiment, the model's predictions for high water content soil were slightly less accurate than those for low water content soil. The overall prediction results yield RMSE = 2.592 %, MAE = 1.829 %, and R2 = 0.711. The model trained on the test bench achieved an accuracy of 2.592 % in the actual field experiment. The

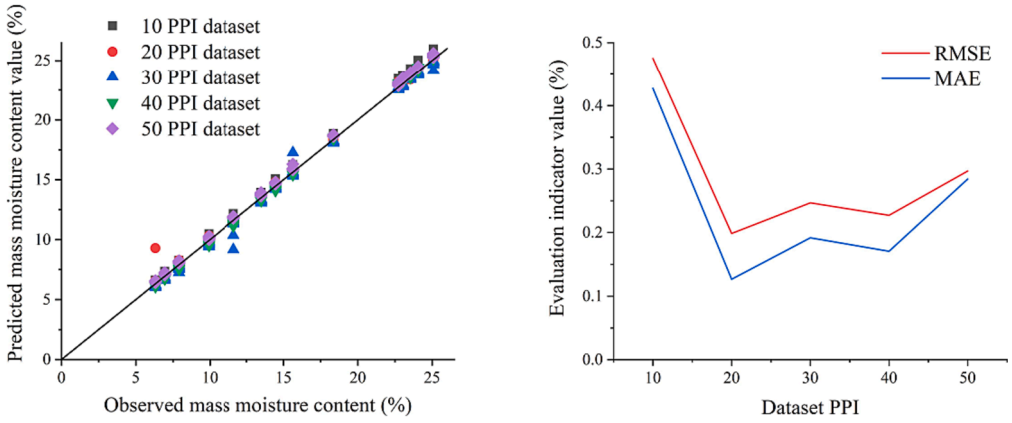


Fig. 9. Comparison of prediction results and statistical summaries for different PPI datasets.

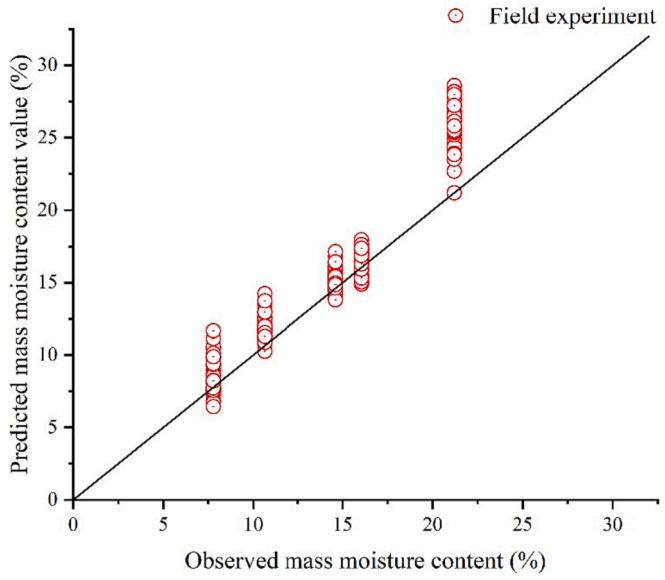


Fig. 10. Predicted and observed soil moisture content from field experiments.

precision was comparable to that of the FDR method, slightly inferior to its performance on the test set.

The model's superior performance on the test set compared to the field experiment is expected. There are several reasons for this. Firstly, there are differences between the actual soil conditions in the test bench experiment and the field experiment, including regional variations in physical structure, differences in soil compaction caused by cultivation and planting processes, and variations in organic matter content. Additionally, there are differences in acoustic environments between the soundproofing measures taken in the test bench during dataset construction and the actual acoustic environment in the field. Moreover, there is a certain amount of statistical error in the water content calibration based on the sampling method used during dataset construction and water content sampling in the field experiment.

5. Discussion

This study employs swept-frequency acoustic signals for soil water content detection, compressing the data through time–frequency spectrograms, and subsequently using the Swin-Transformer model for soil water content regression. The model's results on the test set and in the field, experiments demonstrate the feasibility of this method, achieving a prediction accuracy of 0.191 % on the dataset, showcasing its potential application. A comparison of the detection effects in two frequency bands, 260–360 Hz and 50–550 Hz, reveals that, under similar prediction accuracy, the short-wave band exhibits a distinct time advantage. Additionally, a comparison of multiple PPI values indicates that 20PPI, which is close to the model's input resolution, outperforms others in terms of detection effectiveness, while larger PPI time–frequency spectrograms do not show corresponding advantages. Field experiments further confirmed the practical applicability of this method. Cao Yousong conducted a study on the actual detection results of 8 common FDR-type instruments (Cao and Xiao, 2024). We compared our results with his. In contrast, our method and model exhibit certain advantages on the test set, but the results from field experiments are close to his findings without a significant edge. This indicates that our method and model possess robust detection capabilities, and there is ample room for improvement in the generalization ability of practical detection. By supplementing various datasets, detection accuracy can be enhanced without the need to replace the detection equipment.

The method is capable of soil water content detection, yet there are aspects that require further investigation. On one hand, although the dataset collection process underwent soundproofing, the acoustic environment during data collection differs from that in actual field detection. Interference such as low-frequency noise caused by wind passing through water pipes and acoustic signals disrupted by the vibrations of agricultural machinery engines during real field detection may impact the model's detection performance. Some animal movements and vocalizations near the detection area can interfere with the signal, thereby affecting the detection results. At the same time, the effectiveness of waterproof measures during the entire irrigation process, the impact of related accessories on the acoustic signal, and the effects of irrigation and heavy rain on the positional displacement of equipment in the soil also need further investigation. To enhance the model's resistance to interference, methods such as anti-interference training need to be explored for various common disturbances. On the other hand, while this method utilizes the signal attenuation to achieve soil water content detection and has constructed a substantial dataset, soils with different textures, organic content, bulk density,and stone content inevitably exhibit more diverse attenuation characteristics. Changes in these parameters to a certain extent affect the attenuation of the acoustic signal, ultimately influencing the effective detection bandwidth. This reduction in the effective frequency range makes model construction more challenging and affects accuracy. Improving the model's performance and generalization requires the collection of more diverse field datasets. Furthermore, our research is primarily aimed at monitoring crop growth. For the sake of convenience in comparison and considering detection speed, we calibrated using mass water content. However, for soil hydrology and hydraulic studies, evaluating volumetric water content might be more relevant. The actual performance needs further investigation. Therefore, future research can focus on constructing rapidly diverse datasets, optimizing detection devices and systems using technologies like the Internet of Things, and developing models with stronger generalization, fewer samples, and better detection capabilities under complex acoustic environments, as well as exploring the distance characteristics of this method.

Based on the Swin Transformer neural network, the method of constructing models for soil water content using soil-transmitted swept-frequency acoustic signals' time–frequency spectrograms demonstrates certain advantages compared to various traditional soil water content detection methods. Firstly, in terms of methodology, it explores a new approach for cost-effective regional soil water content detection. This method incorporates more soil water content information into the detection process. Relative to the prevalent multi-sensor point-based soil water content detection methods, it provides a more reasonable assessment of soil moisture content in the penetrated region. The detection range of methods such as FDR and TDR are very limited. For practical field measurements, a large number of sensors are required. Moreover, the sensors need to be arranged reasonably to accurately assess the soil moisture content in the area. Despite this, the volume of measured data is also limited, and the statistical results are significantly affected. Our approach is similar to transmission methods such as GPR and neutron. The detection results are influenced by various parts of the detection range, encompassing information from each part. In comparison, such methods require fewer sensors for practical field assessments of moisture content, making them more reasonable. More importantly, the detection system cost of this method is significantly advantageous compared to existing regional soil water content detection methods such as GPR, neutron method, gamma-ray method, etc. Furthermore, compared to traditional model-building methods, Swin Transformer neural network exhibit outstanding relationship fitting capabilities in this method. The feasibility of deep learning combined with time–frequency spectrograms in soil water content detection is evident. It is foreseeable that by enriching the variety of dataset types and developing rapid real-time data collection devices, the practical prediction accuracy and applicability of this method to various environments can be greatly enhanced. Finally, the method's proposal to transform time-series data into time–frequency spectrograms reduces the volume of 64 MB time-

series data to 13.9 KB, significantly lowering the data load. With this data volume, the use of Internet of Things technology can place the prediction part on the cloud, enabling rapid detection of soil water content by a massive detection device.

6. Conclusions

This study proposes a cost-effective method for soil water content detection using swept-frequency acoustic signals. It utilizes Swin-Transformer neural network to recognize the time–frequency spectrograms of low-frequency swept-frequency acoustic signals from soils with different water content.The method demonstrated a detection accuracy of 0.191 % on the test set and 2.592 % in field experiments, indicating its capability for soil water content detection. This method provides a novel approach for the detection of characteristics in porous and multiphase media such as soil.

CRediT authorship contribution statement

Kangle Song: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jing Nie:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Funding acquisition, Data curation. **Yang Li:** Writing – review & editing, Supervision, Project administration, Methodology, Formal analysis, Conceptualization. **Jingbin Li:** Supervision, Project administration, Funding acquisition. **Pengxiang Song:** Validation, Resources, Methodology, Investigation, Formal analysis. **Sezai Ercisli:** Supervision.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 31860333).

References

Altdorff, D., Oswald, S.E., Zacharias, S., Zengerle, C., Dietrich, P., Mollenhauer, H., Attinger, S., Schrön, M., 2023. Toward large-scale soil moisture monitoring using rail-based cosmic ray neutron sensing. Water Resour. Res. 59 e2022WR033514.

Biot, M.A., 1956. Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range. J. Acoust. Soc. Am. 28, 179–191.

Brutsaert, W., 1964. The propagation of elastic waves in unconsolidated unsaturated granular mediums. J. Geophys. Res. 69, 243–257.

Cao, Y. & Xiao, B. 2024. Performance and calibration of nine dielectric moisture sensors in measuring water content of aeolian sandy soil. *J. Northwest A&F Univ. (Nat. Sci. Ed.)*, 1-14.

Chang-Jie, X., Liang-Ying, X., Yuan-Ye, Y., 2015. Effect of parameters of three-phase unsaturated soils on wave propagation. Rock Soil Mech. 36, 340–344.

Chao, X., Li, Y., 2022. Semisupervised few-shot remote sensing image classification based on KNN distance entropy. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 15, 8798–8805.

Cheng, Q., Su, Q., Binley, A., Liu, J., Zhang, Z., Chen, X., 2023. Estimation of surface soil moisture by a multi-elevation UAV-based ground penetrating radar. Water Resour. Res. 59 e2022WR032621.

Choi, K., Yi, J., Park, C., Yoon, S., 2021. Deep learning for anomaly detection in time-series data: review, analysis, and guidelines. IEEE Access 9, 120043–120065.

Datta, P., Faroughi, S.A., 2023. A multihead LSTM technique for prognostic prediction of soil moisture. Geoderma 433, 116452.

Gao, L., Zhang, J., Yang, C., Zhou, Y., 2022. Cas-VSwin transformer: A variant swin transformer for surface-defect detection. Comput. Ind. 140, 103689.

Gorthi, S., Chakraborty, S., Li, B., Weindorf, D.C., 2020. A field-portable acoustic sensing device to measure soil moisture. Comput. Electron. Agric. 174, 105517.

Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., 2022. A survey on vision transformer. IEEE Trans. Pattern Anal. Mach. Intell. 45, 87–110.

Hasan, M.J., Islam, M.M., Kim, J.-M., 2019. Acoustic spectral imaging and transfer learning for reliable bearing fault diagnosis under variable speed conditions. Measurement 138, 620–631.

He, X., Zhou, Y., Zhao, J., Zhang, D., Yao, R., Xue, Y., 2022. Swin transformer embedding UNet for remote sensing image semantic segmentation. IEEE Trans. Geosci. Remote Sens. 60, 1–15.

Howells, O.D., Petropoulos, G.P., Triantakonstantis, D., Ioannou, Z., Srivastava, P.K., Detsikas, S.E., Stavroulakis, G., 2023. Examining the variation of soil moisture from cosmic-ray neutron probes footprint: experimental results from a COSMOS-UK site. Environ. Earth Sci. 82, 41.

Huang, X., Dong, M., Li, J., Guo, X., 2022. A 3-d-swin transformer-based hierarchical contrastive learning method for hyperspectral image classification. IEEE Trans. Geosci. Remote Sens. 60, 1–15.

Jiang, Z.-D., Owens, P.R., Zhang, C.-L., Brye, K.R., Weindorf, D.C., Adhikari, K., Sun, Z.-X., Sun, F.-J., Wang, Q.-B., 2021. Towards a dynamic soil survey: Identifying and delineating soil horizons in-situ using deep learning. Geoderma 401, 115341.

Kang, W., Tian, J., Bogena, H.R., Lai, Y., Xue, D., He, C., 2023. Soil moisture observations and machine learning reveal preferential flow mechanisms in the Qilian Mountains. Geoderma 438, 116626.

Kara, A., Pekel, E., Ozcetin, E., YıLDıZ, G.B., 2023. Genetic algorithm optimized a deep learning method with attention mechanism for soil moisture prediction. Neural Comput. Appl. 1–12.

Li, Y., Chao, X., 2021. Toward sustainability: trade-off between data quality and quantity in crop pest recognition. Front. Plant Sci. 12, 811241.

Li, Y., Ercisli, S., 2023. Explainable human-in-the-loop healthcare image information quality assessment and selection. CAAI Trans. Intelligence Technol.

Li, Y., Chao, X., Ercisli, S., 2022c. Disturbed-entropy: A simple data quality assessment approach. ICT Express 8, 309–312.

Li, Q., Li, Z., Shangguan, W., Wang, X., Li, L., Yu, F., 2022a. Improving soil moisture prediction using a novel encoder-decoder model with residual learning. Comput. Electron. Agric. 195, 106816.

Li, P., Zha, Y., Tso, C.-H.-M., Shi, L., Yu, D., Zhang, Y., Zeng, W., 2020. Data assimilation of uncalibrated soil moisture measurements from frequency-domain reflectometry. Geoderma 374, 114432.

Li, Q., Zhu, Y., Shangguan, W., Wang, X., Li, L., Yu, F., 2022b. An attention-aware LSTM model for soil moisture and soil temperature prediction. Geoderma 409, 115651.

Liu, Y., Lu, Y., Chen, D., Zheng, W., Ma, Y., Pan, X., 2023. Simultaneous estimation of multiple soil properties under moist conditions using fractional-order derivative of vis-NIR spectra and deep learning. Geoderma 438, 116653.

Ma, J., Tang, L., Fan, F., Huang, J., Mei, X., Ma, Y., 2022. SwinFusion: Cross-domain long-range learning for general image fusion via swin transformer. IEEE/CAA J. Autom. Sin. 9, 1200–1217.

Nie, J., Wang, Y., Li, Y., Chao, X., 2022a. Artificial intelligence and digital twins in sustainable agriculture and forestry: a survey. Turk. J. Agric. For. 46, 642–661.

Nie, J., Wang, Y., Li, Y., Chao, X., 2022b. Sustainable computing in smart agriculture: survey and challenges. Turk. J. Agric. For. 46, 550–566.

Nie, J., Jiang, J., Li, Y., Wang, H., Ercisli, S., Lv, L., 2023. Data and domain knowledge dual-driven artificial intelligence: Survey, applications, and challenges. Expert. Syst. e13425.

Odebiri, O., Mutanga, O., Odindi, J., 2022. Deep learning-based national scale soil organic carbon mapping with Sentinel-3 data. Geoderma 411, 115695.

Qin, P., Deng, Y., Cui, Y., Ye, W., 2023. Development and application of TDR mini-probes for monitoring moisture in small-scale laboratory tests. Int. J. Civil Eng. 1–10.

Romero, H. E., Ma, N., Brown, G. J., Beeston, A. V. & Hasan, M. Deep learning features for robust detection of acoustic events in sleep-disordered breathing. ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019. IEEE, 810-814.

Shen, R., Huang, A., Li, B., Guo, J., 2019. Construction of a drought monitoring model using deep learning based on multi-source remote sensing data. Int. J. Appl. Earth Obs. Geoinf. 79, 48–57.

Taylor, A., Kalnins, A., Koot, M., Jackson, R., Toloza, A., Ahmed, H.S., Goddard, R., Blake, W.H., 2023. Portable gamma spectrometry for rapid assessment of soil texture, organic carbon and total nitrogen in agricultural soils. J. Soil. Sediment. 23, 2556–2563.

Urbina-Salazar, D., Vaudour, E., Richer-De-forges, A.C., Chen, S., Martelet, G., Baghdadi, N., Arrouays, D., 2023. Sentinel-2 and Sentinel-1 bare soil temporal mosaics of 6-year periods for soil organic carbon content mapping in central France. Remote Sens. (Basel) 15, 2410.

Üzen, H., Türkoğlu, M., Yanikoglu, B., Hanbay, D., 2022. Swin-MFINet: Swin transformer based multi-feature integration network for detection of pixel-level surface defects. Expert Syst. Appl. 209, 118269.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I., 2017. Attention is all you need. Adv. Neural Inf. Proces. Syst. 30.

Xu, Y., Li, J., Duan, J., Song, S., Jiang, R., Yang, Z., 2020. Soil water content detection based on acoustic method and improved Brutsaert's model. Geoderma 359, 114003.

Xu, Y., Duan, J., Jiang, R., Li, J., Yang, Z., 2021. Study on the detection of soil water content based on the pulsed acoustic wave (PAW) method. IEEE Access 9, 15731–15743.

Yang, Y., Li, Y., Yang, J., Wen, J., 2022. Dissimilarity-based active learning for embedded weed identification. Turk. J. Agric. For. 46, 390–401.

Yinglan, A., Wang, G., Hu, P., Lai, X., Xue, B., Fang, Q., 2022. Root-zone soil moisture estimation based on remote sensing data and deep learning. Environ. Res. 212, 113278.