University of Tehran

Earth Observation and Geomatics Engineering

Online ISSN: 2588-4360

Homepage: https://eoge.ut.ac.ir/

Intelligent Methods for Soil Moisture Retrieval from Sentinel -1 SAR Data Based on a Designed Ground Sensor

Mohammad Maleki[1]✉ , Jalal Amini[2] , and Ali Younesi[3]

1. Corresponding author, Department of Civil Engineering, Razi University, Kermanshah, Iran. E-mail: mo.maleki@razi.ac.ir
2. University of Tehran, School of Surveying and Geospatial Engineering, College of Engineering, Tehran, Iran. E-mail: jamini@ut.ac.ir
3. University of Tehran, School of Surveying and Geospatial Engineering, College of Engineering, Tehran, Iran. E-mail: abcdef@ut.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 2025-12-07
Received in revised form 2026-02-23
Accepted 2026-04-03
Available online 2026-06-02

Keywords:
Soil moisture,
Deep learning,
Machine Learning,
Sentinel-1,
SAR Data,

ABSTRACT

Soil moisture is an important variable for many studies, such as crop yield estimation, drought monitoring, evapotranspiration, agricultural and water resources management. Today, remote sensing data are widely used to estimate soil moisture. The main purpose of this study is to estimate soil moisture using Sentinel 1 radar data. In the first step, 8 features were extracted from Sentinel 1 data (dual polarized radar backscatter). The ground probe was used for field measurement of soil moisture. In the next step, soil moisture retrieval was done using machine learning methods. These data were used to train and validate machine learning methods. In this research, support vector regression (SVR), decision tree, random forest (RF) and neural network (ANN) methods were performed to estimate soil moisture. Evaluation of the accuracy of different methods was done using field measurement and comparing it with the estimated moisture. The random forest method has the highest accuracy with, Root Mean Square (RMSE=0.03kg/kg) and correlation coefficient (R^2=98.6%). The results of this research showed that the use of different features extracted from Sentinel 1 data along with a suitable machine learning method can significantly increase the accuracy of soil moisture retrieval.

Cite this article: Maleki, M., Amini, J., & Younesi, A. (2025). Intelligent Methods for Soil Moisture Retrieval from Sentinel -1 SAR Data Based on a Designed Ground Sensor, *Earth Observation and Geomatics Engineering*, Volume 9, Issue 2, Pages 28-39. http//doi.org/ 10.22059/eoge.2026.407687.1193

Publisher: University of Tehran.

DOI: http//doi.org/ 10.22059/eoge.2026.407687.1193

1. Introduction

Soil moisture is a fundamental variable in global energy, carbon and water exchanges. Estimation of soil moisture is very important for many studies, such as estimation of the yield of agricultural products, drought monitoring, evaporation and transpiration, agricultural management, earthquake prediction and water resources management (Tronin , Hayakawa, & Molchanov , 2002; Tao, Ryu, Western, & Lee, 2022). Soil moisture has strong effects on vegetation and agricultural production (Maleki, Estimation of surface roughness using two physical models, IEM and fractal SPM, on X-band of SAR image, 2023). Today, hydrological researchers are very interested in simple indicators for monitoring drought events in short periods of time, which can be related to soil moisture (Tramblay & Quintana Seguí, 2022). In particular, soil moisture indices can be more relevant than climate indices for monitoring the potential effects of drought on agriculture and natural vegetation. Soil moisture that estimated from satellite data are valuable information for accurate monitoring of drought events; However, their use requires accurate ground data and computational resources (Tramblay & Quintana Seguí, 2022; Hong, Hernan, Laura , Zhi, & Yang , 2023).

Soil moisture is critical to climate and the hydrological cycle and is a spatial and temporal variable that plays a key role in understanding and predicting climatic trends, extreme runoff, weather, droughts, floods, and landslides (Nadeem, et al., 2022). The amount of water content in the soil can be divided into three depth sections: 1- surface soil moisture (with a depth of 0 to 10 cm), 2- soil moisture of the root zone (with a depth of 10 cm to 1 meter) and 3- deeper soil moisture (with greater depth from the root area to the level of groundwater or bedrock (Nadeem, et al., 2022).It has been very difficult to estimate soil moisture using traditional methods on the ground based on sampling. Estimating soil moisture using ground methods is time-consuming and expensive. Although modern hygrometers have facilitated the process of measuring soil moisture, this process is still time-consuming and expensive and requires calibration with laboratory methods. Installing permanent sensors is another way to facilitate the soil moisture measurement process, but this solution is not practical in large areas. In addition, one of the main problems related to ground gauges is that these measurements are only valid for a small area around the measurement location and it is difficult to prepare a soil moisture map using ground measurements. According to the cases mentioned in relation to ground methods of soil moisture measurement, researchers have shown great enthusiasm for using remote sensing in surface soil moisture estimation. The merits of using remote sensing in estimating soil moisture, such as speed, regular visits, wide coverage, and cost-effectiveness, have drawn much attention from researchers to the use of these data (Liu, Xu, Li, & Guo, 2021).

The use of electromagnetic waves in the visible region, or optical remote sensing, as the most basic earth observation technology, has always played an important role in estimating soil moisture (Ambrosone , et al., 2020). However, the characteristics of optical images have little sensitivity to this variable due to the lack of direct relationship with soil moisture and the high effectiveness of atmosphere and vegetation, and in general, the accuracy of soil moisture estimation using these images is not very suitable. On the contrary, by using the features of microwave remote sensing images, soil moisture can be estimated directly (Maleki, Amini, & Claudia , Fractal algorithm for surface roughness parametersretrieval using multi band/polarization AIRSAR data, 2019). In addition, one of the advantages of radar images is the very low effect of the atmosphere and vegetation (Zhang & Zhou , 2016; Sohrabinia , Rack , & Zawar-Reza , 2014). The use of radar images with synthetic aperture has been widely used in recent studies due to their ability to penetrate to vegetation. As a result, there are many models to estimate soil moisture using these images, which include: time series method (Moran, Hymer , & Qi J, 2000), experimental approaches (Mirsoleimani, Sahebi, Baghdadi, & El Hajj, 2019; Ezzahar, et al., 2020), machine learning and deep learning algorithms (Paloscia, et al., 2013; Baghdadi, El Hajj, El Hajj,, Ludwig, & La Jeunesse; Holtgrave, Förster, Greifeneder, Notarnicola, & Birgit , 2018; Xie , Xu , Zhao, Liu , & Wang , 2014), and changes detection algorithm (Balenzano , Mattia , Satalino , & Davidson , 2010).

In areas with barren soil, important features that affect radar images include soil moisture and roughness. In areas with vegetation, optical images can be used to eliminate the effect of vegetation (Tao, et al., 2019; Baghdadi, El Hajj , Zribi , & Fayad, 2015). In order to estimate soil moisture in these areas and combine optical and radar data, several algorithms have been developed. In general, these algorithms can be divided into two main categories:

1- Algorithms that remove the effect of vegetation based on the microwave scattering model (Liu & Shi , 2016; Kong , Yang , Zhen , Li , & Yang , 2018; Sekertekin , Marangoz , & Abdikan , 2020; Xing , et al., 2019).

2-Algorithms in which the effect of vegetation is shown using vegetation indices or polar decomposition characteristics, and then the inversion model is used to pair them with soil moisture (Attarzadeh , Amini , Notarnicola , & Greifeneder , 2018; Xie , et al., 2018).

Another effect of soil moisture is the destructive effect on the measurement of ground surface displacements using radar images. The change of soil surface moisture during the taking two images by satellite has an effect on the phase of radar interferometry. That result is an error in displacement estimation, which can be up to ten percent of the radar wavelength. Soil moisture can change significantly in a very small space-time scale. Spatial filtering methods cannot be used to reduce the effects of moisture changes on the interferometric phase. Another reason for the superiority of radar images over optic images is that the high contrast between water and land causes a large difference in reflection from the surface of the earth. The radar transmission coefficient in watery areas is almost equal to

80 due to the high electrical conductivity coefficient, but this value is less than 5 for dry soils (Schmugge , Kustas, Ritchie , Jackson, & Rango , 2002).

In this paper, field moisture data are measured using a ground hydrometer. After determining the study area and the time of the field operation according to the passage of the Sentinel 1 satellite over this area, the field operation began. Then by combining the features that extracted from the Sentinel 1 satellite images and ground truth data the moisture was estimated using machine learning and deep learning. finally, the soil moisture map was prepared.

2. Study Area and Dataset

2.1. Study Area

The study area is in the southern part of Tehran city, Nizam Abad village (near Islam Shahr city) of Iran in the longitude between 51^0 17'40" to 51^0 17'45" and latitude between 35^0 31' 8" to $35^0$31' 12". The study area is a barren land with an approximate area of one hectare near agricultural fields. After a request from one of the farmers, part of the study area was irrigated two nights before the passage of the satellite and another part was irrigated one night before the passage so that the soil reaches a stable moisture percentage. Figure 1 shows the location of the study area. Field operations include soil moisture measurement and topography mapping from the studied area.

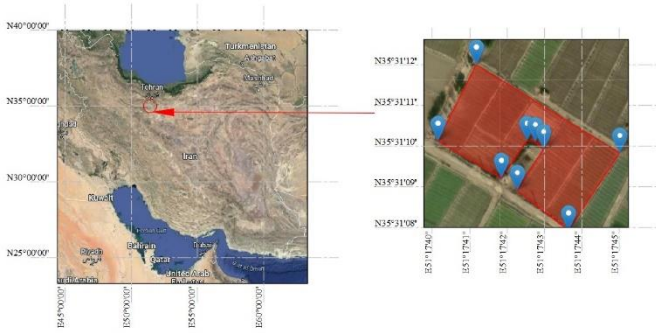

Figure 1. Location of the study area

2.2. Satellite Data

In this research, the Sentinel 1 radar satellite images in both VV and VH polarizations with a spatial resolution of 10 m, on August 21, 2023 were used. We downloaded the Sentinel data from website of the European Space Agency (https://scihub.copernicus.eu). The satellite transit time is 14:00 in the afternoon on August 21, 2023, and 2:00 in the morning on August 22, 2023 at night.

The mission of the Sentinel 1 satellite is an imaging system of two polar orbiting satellites and a day and night radar for land and ocean monitoring. The purpose of this mission is to provide continuity of data after the retirement of ERS-2 and Envisat. Compared to satellites such as Envisat, Sentinel 1 improves spatial, temporal and radiometric resolution, while operating at approximately the same frequency.

Another development is the use of cross-polarization to correct for seasonal vegetation effects in copolarized backscatter measurements. In general, HH polarization is preferred due to the better penetration waves through vertical vegetation (grasses, crops) (Pratola , Barrett , Gruber , & Dwyer , 2015).

In general, soil moisture retrieval from SAR data is challenging due to the confusing effect of surface roughness and vegetation on the signal. Although several approaches using different backscattering models have been developed, soil moisture retrieval from Sentinel 1 data remains an enigma, mainly due to significant roughness changes (e.g., due to agricultural practices) and the higher complexity of vegetation characteristics (Posner & Georgakakos , 2015). Due to the loss of the SMAP radar, Sentinel 1 is currently the only active sensor with high spatial resolution on a 6-day time horizon for surface soil moisture estimates (Posner & Georgakakos , 2015).

Table1: Information of Sentinel-1 data used in this study

Platform	Sentinel-1A	Sentinel-1B
Orbit direction	Descending	Ascending
Product type	GRD	GRD
Acquisition mode	IW	IW
Polarization mode	Dual (VH, VV)	Dual (VH, VV)
Incidence angle (°)	36–54	29–45
Revisit period	12 days	12 days
Spatial resolution	10 m	10 m
Acquisition period	Aug. 2015 to Dec. 2019	Sep. 2016 to Dec. 2019
Number of images	93	95

Table 1 shows the specification of satellite data used in this study.

2.3. Field Measurement

In order to collect field data, a soil moisture probe was designed and built. The process of designing and manufacturing the probe includes design, manufacturing, laboratory operation and calibration. The probe measures soil moisture (with an accuracy better than 0.9 percent of weighted moisture), displaying precise coordinate, and storing information for a range of 30 x 30 cm. (Figure 2). The full technical specifications and development procedure of the sensor have been previously published in (Amini & Younesi Sinaki, 2025)

Soil moisture in the field was measured simultaneously with the passage of the Sentinel 1 satellite over the study area with 2 meters interval using ground probe (figure3, b).

The topography mapping was done by using Raimand Ultra series GPS device (with receiving corrections from local system) at 2 meter intervals with 5 cm precision of elevation (figure3, a).

Figure 4 illustrated digital elevation model of study area.

Figure 2. ground probe for soil moisture measurement

Figure 3.a. The topography mapping using Raimand GPS

Figure 3.b. Measurement of Soil Moisture

3. Methodology

The flowchart of the proposed method to estimate soil moisture is illustrated in Figure 5.

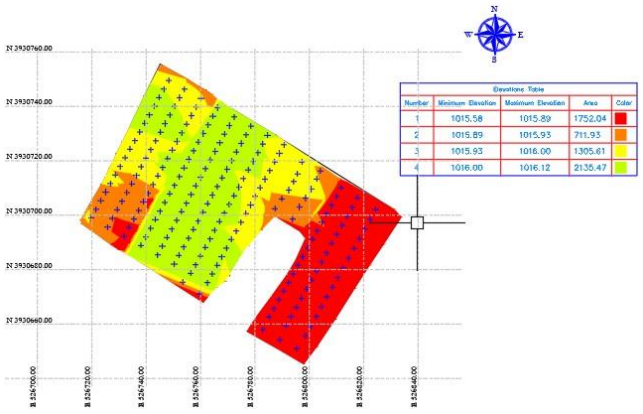

Figure 4. Digital Elevation Model of study area

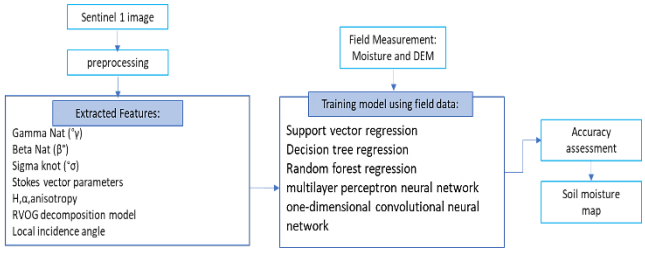

Figure5. The flowchart of propped method

3.1. Data Preprocessing

Sentinel-1 SLC data were preprocessed using the Sentinel Application Platform (SNAP). The preprocessing workflow followed the standard SLC processing chain and included: (1) application of the precise orbit file to improve orbital accuracy; (2) thermal noise removal; (3) TOPS deburst operation to merge individual bursts into a continuous image; (4) radiometric calibration (Amini & Younesi Sinaki, 2025) to convert the complex digital numbers into sigma nought (σ^0) backscatter coefficients; (5) speckle filtering to reduce SAR noise .This processing chain ensures radiometric consistency and geometric accuracy of the final backscatter products used in the analysis.

3.2. SAR Derived Features

In this study, machine learning methods were used to estimate soil moisture. In machine learning methods, better quality of input data plays an important role in increasing the accuracy of model prediction(SINGH nature reference). Converting raw data to other data forms can be effective in the better performance of machine learning methods. For this purpose, eight features were extracted from the Sentinel image. These are backscatter value(σ^0), gamma naught(Υ^0), betta naught(δ^0), stokes vector parameter, covariance matrix, (H, α, anisotropy), RVOG decomposition model and local incidence angle. Digital elevation model is produced from topography mapping. Some of these features are used for the first time to estimate soil moisture.

3.3. Machine Learning and Deep Learning Method for Estimating Soil Moisture

In this research, machine learning and deep learning

methods have been used to estimate soil surface moisture. Machine learning models include support vector regression, decision tree regression and random forest regression. The deep learning models used are multi-layer perceptron neural network and one-dimensional convolution neural network.

3.3.1. Support Vector Regression

The application of support vector machine in fitting regression function is based on the theory of structural risk minimization. Due to its strong generalizability and very good ability to work with large and small samples, support vector regression is widely used to extract surface parameters using satellite images, such as soil salinity mapping, soil moisture retrieval and forecasting the yield of agricultural products (Pekel , 2020).

Assuming that the soil moisture variable is SMC, it can be calculated from the equation (1):

$$SMC = f(x) + e \quad (1)$$

In the equation(1), f is the desired unknown mapping between the input data (x), which is equal to the features extracted from the satellite radar images and the elevation model, and the output data (the soil moisture), and e is a random variable that represents the impact noise.

The goal of this algorithm is to determine the mapping that brings the input data as close as possible to the output data, finally the function f can be approximated by considering the function $f^{'}$ using equation 2:

$$f^{'}(x) = \omega.\phi(x) + b \quad (2)$$

In the equation 2 ω is equal to the weight of the function, $\phi(x)$ is a mapping that maps samples from the original space to the feature space with higher dimensions, and b is the bias.

The optimal cost function can be estimated from equation 3:

$$\psi_{(\omega,\Sigma)} = c\sum_{i=1}^{N}(\Sigma_i + \Sigma_i^*) + \frac{1}{2}\left\|\omega\right\|^2 \quad (3)$$

In the equation 3, c is equal to the penalty parameter, which is a tuning parameter that allows adjusting the balance between function complexity and function tolerance $f^{'}$. Σ measures the distance of the training samples outside the target range.

When the $k(x_i, x)$ kernel function satisfies Mercer's theorem, the corresponding regression prediction function can be described as formula 4 according to the theory of universal functions:

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*).k(x_i, x) + b \quad (4)$$

where n is the number of training samples, α and α_i^* denotes the Lagrange coefficients, and b is the bias. The Gaussian radial basis function was chosen as the support vector regression kernel function in this study (Zhang & Zhou , 2016).

3.3.2. Decision Tree Regression

An effective approach to predict and classify variables in a model is to build decision trees. A decision tree consists of a root node, branches, decision nodes and leaf nodes. Tree-based modelling was introduced by Berryman (1984) and is a heuristic technique for discovering structure in data that is increasingly used to: (1) devise predictive rules that can be evaluated quickly and repeatedly, (2) assessing the adequacy of linear models, and (3) summarizing large multivariate data sets. Recursive partitioning starts from the root node where the target variable is, and attributes are tested at decision nodes, branching out different results as leaf nodes (also known as terminal nodes). or continue to generate more decision nodes (Breiman, , Friedman, Olshen, & Stone, 1984). The recursive partitioning method can be described as reducing "impurity", that is, the records of each leaf node have the same classification. Doing so results in the decision tree providing the information with the lowest misclassification (Nurdiawan , Kurnia , Solihudin , Hartati , & Suprapti , 2021).

3.3.3. Random Forest Regression

Random forest starts by drawing a certain number of bootstrap samples from the original data with replacement (where the number of bootstrap samples is equal to the number of trees growing in RF). The size of each bootstrap sample is the same as the training data set.

For each bootstrap sample, an unpruned unclassified tree is grown. The process of tree growth is similar to decision tree classification and regression. However, in order to produce a more diverse set of trees, the best split is selected from a random subset of all predictor variables rather than from all predictors in each split in each tree. Typically, for classification trees, the number of random subsets of predictors is equal to the square root of the total number of predictors. The rationale for random selection of partition variables is that lower correlation between individual trees leads to lower generalization error (Ahmed, et al., 2021). The overall classification is calculated by summing the classifications of each tree. Random forests have several advantages. It can overcome the instability problem of the confidence tree by averaging the predictions of several trees. The confidence tree creates a single classification tree, which may lead to unstable classification results. For example, for a data set with a large number of predictors, the classification tree will change if different partitioning variables are chosen. In addition, random forest provides a measure of the variable importance of all predictors, which identifies key factors in predicting the outcome variable. In addition, the random forest provides a graphical representation such as a confidence tree that can help us better understand the relationships between variables. Similarly, random forest can handle higher-order

interactions between variables in a dataset (Kornelsen & Coulibaly, 2014).

3.3.4. Multilayer Perceptron Neural Network

A neural network is a reasoning or logical model based on what happens in the human brain. The human brain consists of a number of interconnected nerve cells (or so-called main information processing units called neurons). The number of these neurons reaches 10 billion, and there are about 60 trillion connections between these neurons. By using only a few neurons, you can reach a speed that is better than the fastest computers and they have extremely high processing power. Each neuron has a relatively simple structure, but some of them can create high processing power. What the human brain can do using neurons and neural network is that the brain is a parallel information processing system, which means can process and perform several tasks at the same time and performs non-linear processing easily. The information processing system in the brain is complex; it is clear that the computer systems inspired by the human brain have not reached the ability of the human brain system. Learning is a fundamental part in the natural neural network, and this has motivated researchers to simulate on the computer and develop a natural neural network. The brain's nervous system can be trained and improved over time. An artificial neural network consists of a number of simple processing units called neurons that are connected to each other using weighted links, similar to neurons, and transmit signals. They transfer from one neuron to another (Ahmed, et al., 2021; Kornelsen & Coulibaly, 2014). In order to determine the best configuration of the network, the parameters that affect the prediction error value, including the number of pattern inputs, the delay value, the number of hidden layers and their number of neurons, activation functions and the learning algorithm through an iterative process to evaluate the minimum error Predictions are obtained when the training process is executed.

To start the prediction process, N is the number of observations, $y_1, y_2, \ldots, y_N$ are selected as the training set. $y_1, y_{N+1}, y_{N+2}, \ldots, y_{N+m}$ are considered as test sets. The number of input nodes corresponds to the number of lagged observations used to discover the underlying pattern in a time series. Different input nodes can affect the learning or predictive ability of the network. For example, a network with three nodes in the input layer, two nodes in the hidden layer and one node in the output layer has been proposed (Song , Zhang, Liu , Li D, Zhao , & Yang , 2016). In other words, all four observations in the training set form a pattern vector, three of which are input values and the last one is the output value. The training patterns in the proposed network are:

$$\begin{aligned} X_4 &= f(X_1, X_2, X_3) \\ X_5 &= f(X_2, X_3, X_4) \\ X_N &= f(X_{N-3}, X_{N-2}, X_{N-1}) \end{aligned} \quad (5)$$

The training process is performed to find the optimal connection weights in such a way that the prediction error is minimized. The PE can be written as equation 6:

$$PE = \sum (x_i - \hat{x}_i) \quad (6)$$

where x_i is the output of network. Figure 6 shows the basic architecture of a multi-layer forward perceptron for time series forecasting applications, which consists of a single output unit, k is hidden units, and n is input units. W_{ij} is the connection weight from input unit i to hidden unit j and T_j_j is the connection weight from jth hidden unit to output. The hidden layer consists of nodes that are connected to both the input layer and the output layer and are the most important part of a network (Wang , et al., 2019).

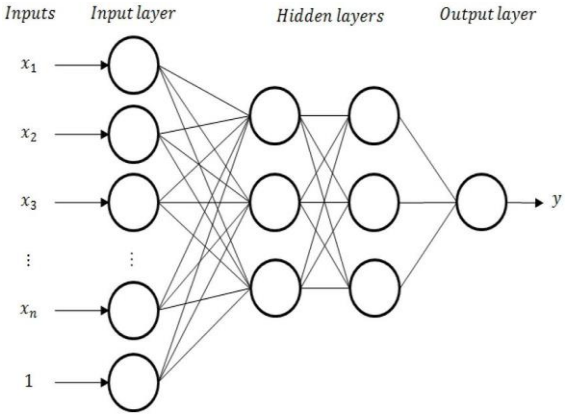

Figure 6. An example of a multilayer neural network

3.3.5. One-Dimensional Convolutional Neural Network

Convolutional neural network has the most outstanding performance in image recognition. The structure of a typical convolutional neural network consists of convolution, pooling, excitation function and fully connected layers. Mathematically, a filter operation is performed similarly as discrete convolution, hence the name convolutional neural network (Zhang & Zhou , 2016) .

The convolution layer is mainly used to extract features from the vector image. In practice, the number of filters or kernels, kernel size, and kernel stride are usually hyperparameters. and need artificial allocation. In a convolution layer, multiple kernels are used, and they are initialized differently, so the feature maps extracted by them are different from each other. The image or filtered sub vector of a kernel is named as its feature map (Carranza , Nolet , Pezij , & van der Ploeg , 2021).

In this research, one-dimensional convolutional neural network was used to estimate soil moisture. Figure 7 shows the structure of a one-dimensional convolutional neural network.

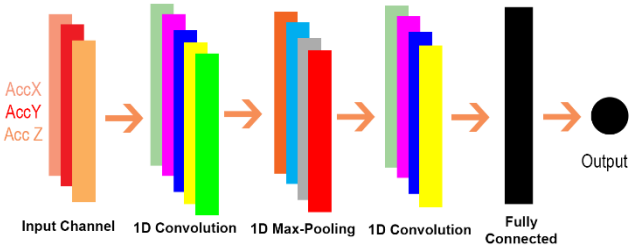

Figure7. An example of the convolutional neural network structure

3.4. Implementation and Evaluation of Results

Considering the non-linearity and complexity of the soil moisture estimation from satellite radar images and the influence of these images on environmental conditions, determining the parameters of artificial intelligence algorithms is a difficult task.

In this research, the parameters of different machine learning and deep learning models were determined using the network search method. The criterion for determining the accuracy of all models was the coefficient of determination (R2) and the root mean square error (RMSE), which are calculated from equations 7 and 8.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(Y_i - Y_i)^2}{n}} \quad (7)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (Yi - Y_i)^2}{\sigma} \quad (8)$$

Where Y_i is estimated value, Y_i is the measured value and n is the number of measured values.

3.4.1. Support Vector Regression

In this study, radial basis function kernel was used and the best penalty parameter (C) was determined as 1. The coefficient of determination (R2) is 79.3% and the root mean square is 0.495(kg/kg).

3.4.2. Decision Tree Regression

After the network search, the criterion for measuring the quality of the divisions in this research is the mean squared error, and the strategy used to select the division in each node is the best decision selection instead of random selection, and other parameters when in the default state were supposed to witness the best performance of the decision tree.

3.4.3. Random Forest Regression

In determining the number of estimators, the optimal number of trees in the forest, 200 trees were determined, and the standard for measuring the quality of division, is the mean square of errors, other parameters were set to default.

3.4.5. Multilayer Perceptron Neural Network

In the perceptron neural network, the number of neurons in the input layer is equal to the number of pixels, and in both hidden layers, the number of neurons is 64 and the activation function is RELU, the output layer, which is the fully connected layer, has one neuron. Finally, the number of unknown parameters is 5505, and all parameters can be trained. Figure 8 shows the absolute error in each period of the neural network.

Figure 9 shows the histogram of prediction error for all models used in this article. Figures 10 and 12 respectively show the prepared moisture map and the correlation coeffiecient between the field soil moisture data and the values estimated by different models.

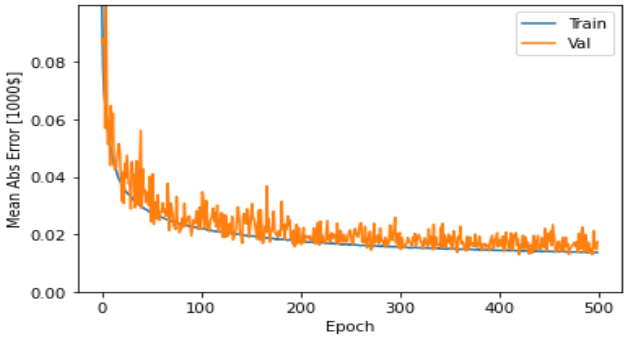

Figure 8. The absolute error in each period of the multilayer perceptron neural network.

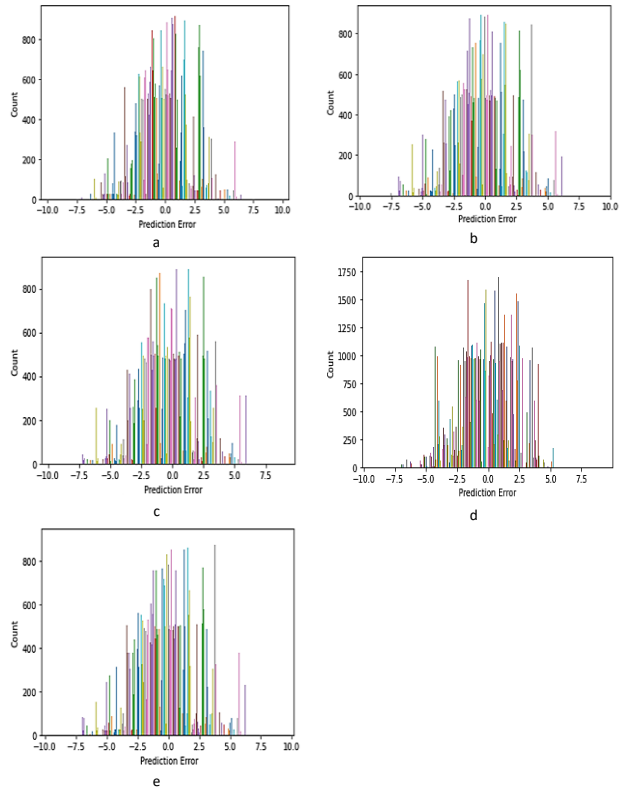

Figure 9. Histogram of model prediction error a) SVM b) Decision Tress c) Random Forest d) multilayer perceptron neural network e) one-dimensional convolutional neural network

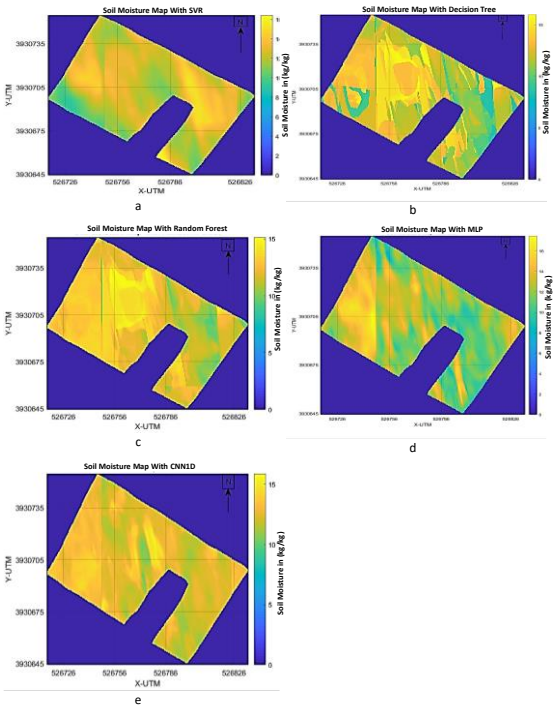

Figure 10. Soil Moisture Map a) SVM b) Decision Tress c) Random Forest d) multilayer perceptron neural network e) one-dimensional convolutional neural network

3.4.6. One-Dimensional Convolutional Neural Network

In the convolution neural network, the dimensions of the input layer are equal to the number of training data in the number of features, three layers of one-dimensional convolution with dimensions of 3, 5 and 7, a high integration layer and a random removal layer (with 0.5 removal) and a fully connected layer with 32 neurons are designed as hidden layers and the output layer is a fully connected layer with output dimensions. The activator function of all layers was considered RELU function. Figure 11 shows the absolute error in each period of the neural network.

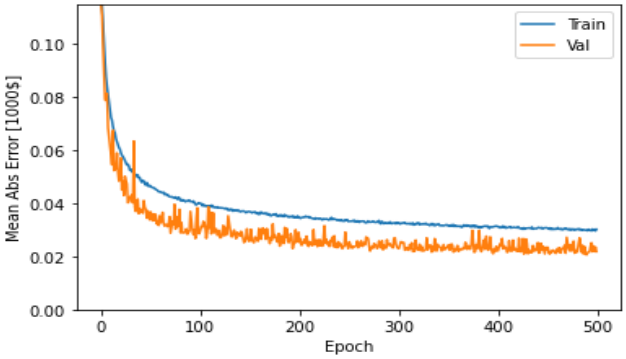

Figure11. Absolute error in each period of one-dimensional convolutional neural network

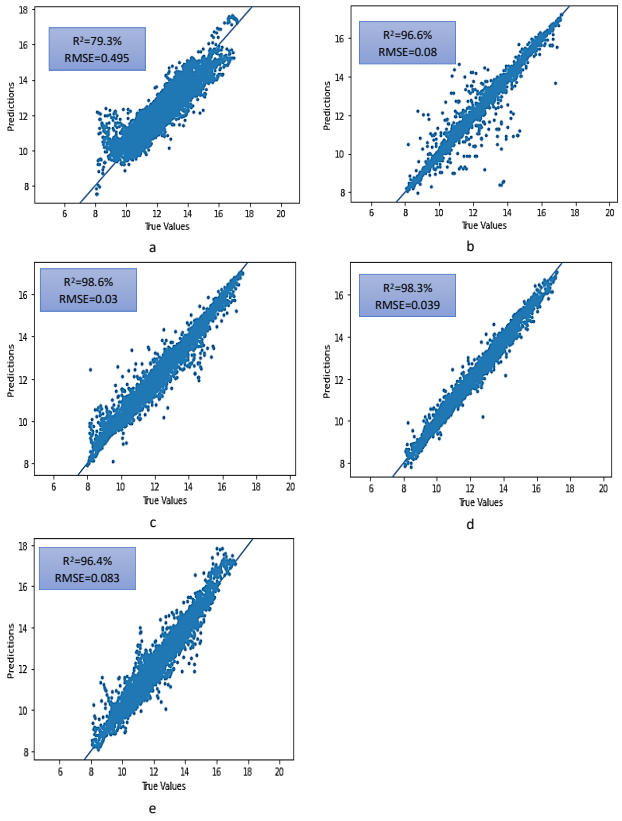

Figure 12. The correlation coefficient between ground truth data and predicted moisture a) SVM b) Decision Tress c) Random Forest d) multilayer perceptron neural network e) one-dimensional convolutional neural network.

Table 2 shows the RMSE and R2 for training and testing data for all model that used in this study.

Table 2. Comparison of machine learning and deep learning models

Models	Training		Testing	
	RMSE (kg/kg)	R^2 (%)	RMSE (kg/kg)	R^2(%)
Support vector regression	0.491	79.9	0.495	79.3
Decision tree regression	0.001	100	0.08	96.6
Random forest regression	0.004	99.8	**0.03**	**98.6**
multilayer perceptron neural network	0.033	98.6	0.039	98.3
one-dimensional convolutional neural network	0.08	96.7	0.083	69.4

4. Conclusion

Although ground-based soil moisture measurements provide high accuracy, they are time-consuming, labor-intensive, and costly for large-scale applications. The integration of radar satellite data with ground observations offers a practical alternative for spatial soil moisture estimation. In this study, machine learning and deep learning approaches were employed to estimate soil moisture using SAR-derived parameters from Sentinel-1 data. The results indicate that the Random Forest model achieved the highest performance (R^2 = 0.986), while Support Vector Regression showed comparatively lower performance (R^2 = 0.793). Despite the high predictive performance observed in this study, it should be noted that the results are influenced by the specific environmental conditions, sample size, and temporal coverage of the dataset. Therefore, the generalization of the models to other regions, soil types, or climatic conditions requires further validation. Additionally, the findings suggest that increasing the depth of convolutional neural networks does not necessarily improve model performance; rather, appropriate architectural design and parameter tuning play a critical role Overall, the proposed framework demonstrates promising potential for soil moisture estimation; however, further studies incorporating multi-seasonal data and independent validation datasets are recommended to fully assess its robustness and transferability.

References

Ahmed, A., Deo , R., Ghahramani , A., Raj , N., Feng , Q., Yin , Z., & Yang , L. (2021). LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4. 5 and RCP8. 5 global warming scenarios. *Stochastic Environmental Research and Risk Assessment*.

Ambrosone , M., Matese , A., Di Gennaro , S., Gioli , B., Tudoroiu , M., Genesio , L., & et al. (2020). Retrieving soil moisture in rainfed and irrigated fields using Sentinel-2 observations and a modified OPTRAM approach. *International Journal of Applied Earth Observation and Geoinformation.;., 89(8)*(102113). doi: https://doi.org/10.1016/j.jag.2020.102113

Amini, J., & Younesi Sinaki, A. (2025). Designing and construction a ground-based soil moisture sensor in order to provide ground truth data for remote sensing images. *JGST, 14*(3), 59-68.

Attarzadeh , R., Amini , J., Notarnicola , C., & Greifeneder , F. (2018). Synergetic use of Sentinel-1 and Sentinel-2 data for soil

moisture mapping at plot scale. *Remote Sensing., 10(18)*.

Baghdadi, N., El Hajj , Zribi , M., & Fayad, I. (2015). Coupling SAR C-band and optical data for soil moisture and leaf area index retrieval over irrigated grasslands. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing., 9(3)*, 1129-1243.

Baghdadi, N., El Hajj, El Hajj,, M., Ludwig, R., & La Jeunesse, I. (n.d.). Estimation of soil parameters over bare agriculture areas from C-band polarimetric SAR data using neural networks. *Hydrol. Earth Syst. Sci., 16*, 1607-1621. doi:https://doi.org/10.5194/hess-16-1607-2012, 2012.

Balenzano , A., Mattia , F., Satalino , G., & Davidson , M. (2010). Dense temporal series of C-and L-band SAR data for soil moisture retrieval over agricultural crops. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 4(2)*, 439-450.

Breiman, , l., Friedman, j., Olshen, R., & Stone, C. (1984). *Classification and Regression Trees.* Chapman and Hall/CRC. doi:https://doi.org/10.1201/9781315139470

Carranza , C., Nolet , C., Pezij , M., & van der Ploeg , M. (2021). Root zone soil moisture estimation with Random Forest. *Journal of hydrology., 593:125840*.

Ezzahar, J., Ouaadi, N, Zribi, M., Elfarkh, J., Aouade, G., Khabba, S., . . . Jarlan, , L. (2020). Evaluation of Backscattering Models and Support Vector Machine for the Retrieval of Bare Soil Moisture from Sentinel-1 Data. *Remote Sens, 12*(72). doi: https://doi.org/10.3390/rs12010072

Holtgrave, A., Förster, M., Greifeneder, F., Notarnicola, C., & Birgit , K. (2018). Estimation of Soil Moisture in Vegetation-Covered Floodplains with Sentinel-1 SAR Data Using Support Vector Regression. *86*, 85-101. doi: https://doi.org/10.1007/s41064-018-0045-4

Hong, Z., Hernan, A., Laura , V., Zhi, L., & Yang , H. (2023). Triple Collocation of Ground-, Satellite- and Land Surface Model-Based Surface Soil Moisture Products in Oklahoma Part II: New Multi-Sensor Soil Moisture (MSSM) Product. *Remote Sens, 15*(3450). doi: https://doi.org/10.3390/rs15133450

Kong , J., Yang , J., Zhen , P., Li , J., & Yang , L. (2018). A coupling model for soil moisture retrieval in sparse vegetation covered areas based on microwave and optical remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing, 56(12)*, 7162-7173.

Kornelsen , K., & Coulibaly, P. (2014). Root-zone soil moisture estimation using data-driven methods. *Water Resources Research, 50*(4), 2946-2962.

Liu , C., & Shi , J. (2016). Estimation of vegetation parameters of water cloud model for global soil moisture retrieval using time-series L-band Aquarius observations. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing., 912)*, 5621-5633.

Liu, J., Xu, Y., Li, H., & Guo, J. (2021). Soil Moisture Retrieval in Farmland Areas with Sentinel Multi-Source Data Based on Regression Convolutional Neural Networks. *Sensors, 21*(877). doi: https://doi.org/10.3390/s21030877

Maleki, M. (2023). Estimation of surface roughness using two physical models, IEM and fractal SPM, on X-band of SAR image. *Earth Observation and Geomatics Engineering,, 7(1)*, 26-35. doi: https://doi.org/10.22059/eoge.2023.365877.1143

Maleki, M., Amini, J., & Claudia , N. (2019). Fractal algorithm for surface roughness parametersretrieval using multi band/polarization AIRSAR data. *J. Appl. Remote Sens, 13*(1), 014525. doi: https://doi.org/10.1117/1.JRS.13.014525

Mirsoleimani, H., Sahebi, M., Baghdadi, N., & El Hajj, M. (2019). Bare Soil Surface Moisture

Retrieval from Sentinel-1 SAR Data Based on the Calibrated IEM and Dubois Models Using Neural Networks. *Sensors, 19*(3209). doi:https://doi.org/10.3390/s19143209

Moran, M., Hymer , D., & Qi J, S. (2000). Soil moisture evaluation using multi-temporal synthetic aperture radar (SAR) in semiarid rangeland. *Agricultural and Forest meteorology, 05(1-3):1*, 69-80. doi: 10.1016/S0168-1923(00)00189-1

Nadeem, A., Zha, Y., Shi, L., Ran, G., Ali, S., Jahangir, Z., . . . Awais, M. (2022). Multi-Scale Assessment of SMAP Level 3 and Level 4 Soil Moisture Products over the Soil Moisture Network within the ShanDian River (SMN-SDR) Basin, China. *Remote Sens, 14*(982). doi:https://doi.org/10.3390/rs14040982

Nurdiawan , O., Kurnia , D., Solihudin , D., Hartati , T., & Suprapti , T. (2021). Comparison of the K-Nearest Neighbor algorithm and the decision tree on moisture classification. *IOP Conference Series: Materials Science and Engineering.*

Paloscia, S., Pettinato, E., Santi, C., Notarnicola, L., Pasolli, A., & Reppucci, A. (2013). Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation. *Remote Sensing of Environment, 134*, 234-248. doi:https://doi.org/10.1016/j.rse.2013.02.027

Pekel , E. (2020). Estimation of soil moisture using decision tree regression. 2020;139(3-4):1111-9. *Theoretical and Applied Climatology, 139*(1), 1111-1119. doi:10.1007/s00704-019-03048-8

Posner , A., & Georgakakos , K. (2015). Soil moisture and precipitation thresholds for real-time landslide prediction in El Salvador. *Landslides, 12(6)*, 1179-1196.

Pratola , C., Barrett , B., Gruber , A., & Dwyer , E. (2015). Quality assessment of the CCI ECV soil moisture product using ENVISAT ASAR wide swath data over Spain, Ireland and Finland. *Remote Sensing, 7(11):15388-423*.

Schmugge , T., Kustas, W., Ritchie , J., Jackson, T., & Rango , A. (2002). Remote sensing in hydrology. *Advances in water resources., 25(8)*, 1367-1385.

Sekertekin , A., Marangoz , A., & Abdikan , S. (2020). ALOS-2 and Sentinel-1 SAR data sensitivity analysis to surface soil moisture over bare and vegetated agricultural fields. *Computers and Electronics in Agriculture, 171*(C).

Sohrabinia , M., Rack , W., & Zawar-Reza , P. (2014). Soil moisture derived using two apparent thermal inertia functions over Canterbury, New Zealand. *8*(1, 083624). doi: https://doi.org/10.1117/1.JRS.8.083624

Song , X., Zhang, G., Liu , F., Li D, Zhao , Y., & Yang , J. (2016). Modeling spatio-temporal distribution of soil moisture by deep learning-based cellular automata model. *Journal of Arid Land, 8*, 734-748.

Tao, L., Ryu, D., Western, A., & Lee, S.-G. (2022). Comparison of KOMPSAT-5 and Sentinel-1 Radar Data for Soil Moisture Estimations Using a New Semi-Empirical Model. *Remote Sens., 14* (4042). doi: https://doi.org/10.3390/rs14164042

Tao, L., Wang , G., Chen , W., Chen , X., Li , J., & Cai , Q. (2019). Soil moisture retrieval from SAR and optical data using a combined model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(2)*, 637-647.

Tramblay, Y., & Quintana Seguí, P. (2022). Estimating soil moisture conditions for drought monitoring with random forests and a simple soil moisture accounting scheme . *Nat. Hazards Earth Syst. Sci, 22*, 1325–1334. doi: https://doi.org/10.5194/nhess-22-1325-2022, 202

Tronin , A., Hayakawa, M., & Molchanov , O. (2002). Thermal IR satellite data application for earthquake research in Japan and China. *Journal of Geodynamics, 33*(4), 519-534.

Wang , X., Dou , X., Zhang , X., Liu , H., Li , H., & Meng , X. (2019). Development of soil spectral allocation models considering the effect of soil moisture. *Soil and Tillage Research, 195:104374*.

Xie , Q., Meng , Q., Zhang , L., Wang , C., Wang , Q., & Zhao , S. (2018). Combining of the H/A/alpha and freeman–durden polarization decomposition methods for soil moisture retrieval from full-polarization radarsat-2 data. *Advances in Meteorology, 2018*, 17 pages. doi: https://doi.org/10.1155/2018/9436438

Xie , X., Xu , J., Zhao, J., Liu , S., & Wang , P. (2014). Soil moisture inversion using AMSR-E remote sensing data: An artificial neural network approach. *Applied Mechanics and Materials*, 501-504.

Xing , M., He , B., Ni , X., Wang , J., An , G., Shang , J., & Huang , X. (2019). Retrieving Surface Soil Moisture over Wheat and Soybean Fields during Growing Season Using Modified Water Cloud Model from Radarsat-2 SAR Data. *Remote Sensing., 11(16)*. doi:https://doi.org/10.3390/rs11161956

Zhang , D., & Zhou , G. (2016). Estimation of soil moisture from optical and thermal remote sensing: A review. *Sensors, 16(8):*(1308).