Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Integration of SAR Data and Machine Learning for Soil Moisture Estimation using SMAP Validation

Dheeraj Bhima Raut[1*], Rajeshwari Pangarkar[1], Sidheshwar Raut[2], Shafiyoddin Sayyad[3]
[1] *Department of Physics, Mrs. K. S. K. Science College, Beed (MH), India-421122*
[2] *Department of Physics, Shri. Chhatrapati Shivaji College, Omerga (MH), India-413606*
[3] *Microwave & Imaging Spectroscopy Research Lab, Milliya College, Beed (MH), India-431122*
Corresponding Author: [1]Dheerajraut11@gmail.com

ARTICLE INFO

Received: 12 Oct 2024

Accepted: 10 Dec 2024

ABSTRACT

Soil moisture estimation is crucial for various applications, including precision agriculture, hydrology, and climate studies. Synthetic Aperture Radar (SAR) data has proven to be highly effective for soil moisture prediction due to its capability to provide high-resolution, all-weather observations. This study explores the potential of two machine learning techniques, first Support Vector Regression (SVR) and second one Random Forest Regression (RF), for soil moisture prediction using SAR-derived features. This research utilizes the Sentinel-1 SAR imagery, validated soil moisture data products given by NASA Soil Moisture Active Passive (SMAP) Mission which is a NASA satellite observatory focused on measuring and mapping soil moisture, and used the auxiliary data derived from Google earth engine. Feature selection techniques are applied to enhance model performance, and hyperparameter tuning is conducted to optimize predictive accuracy. The performance of SVR and RF models is evaluated based on correlation (r), Root Mean Square Error (RMSE), and R^2 scores. Experimental results indicate that both SVR and RF models demonstrate strong predictive capabilities, but the RF model exhibits slightly superior performance in terms of generalization and resistance to overfitting.

Keywords: Soil Moisture Prediction, Machine Learning, Synthetic Aperture Radar (SAR), Random Forest, Support Vector Regression, Sentinel-1.

1. INTRODUCTION

Soil moisture is important for environmental and agricultural applications such as precision agriculture, hydrological modelling, drought monitoring, and climate change assessment [1]. Traditional ground-based soil moisture measurements are accurate but are often spatially sparse and temporally inconsistent [2]. While remote sensing technologies like Synthetic Aperture Radar (SAR) can be used for large-scale soil moisture mapping due to their high spatial resolution and all-weather, day-night imaging capabilities. SAR systems, operating in the microwave spectrum, are sensitive to soil moisture content because of the contrast in dielectric properties between dry soil and water [3]. This sensitivity enables the detection of variations in soil moisture through backscatter coefficients [4]. Studies have demonstrated that increases in soil moisture result in significant increases in radar backscatter, making SAR a potent tool for soil moisture monitoring [5].

This study investigates the application of advanced machine learning algorithms, specifically Support Vector Regression (SVR) and Random Forest Regression (RF), to estimate surface soil moisture using backscattering coefficient (σ^0) which is derived from Sentinel-1 SAR data. Sentinel-1, operated by the European Space Agency (ESA), provides dual-polarized (VV and VH) C-band SAR imagery that is sensitive to surface roughness and dielectric properties, both affected by soil moisture [6]. The study selected is Beed district located in Maharashtra, India.

To build predictive models, the SAR-derived inputs are combined with validated reference soil moisture data obtained from the NASA Soil Moisture Active Passive (SMAP) mission. SMAP provides global soil moisture estimates at a coarse resolution, which are resampled and spatially matched with the SAR data [7]. Additional auxiliary variables, such as MODIS Temperature, Soil Moisture (from SMAP), DEM, RVI and Soil Dielectric Constant using simplified Dubois model, are used to improve model accuracy**.** Furthermore, the feature selection method like

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

correlation analysis, is used to identify the most relevant input variables [8]. Then hyperparameter tuning is conducted using grid search and cross-validation techniques to optimize each model's performance [9]. The models are trained and tested across multiple geographical locations to evaluate their robustness. Finally, Model evaluation is carried out using statistical metrics including the Pearson correlation coefficient (r) which measures the linear relation between two continuous variables, root mean square error (RMSE) which gives the average magnitude of the error between predicted and observed values also RMSE is very sensitive to large errors, and the coefficient of determination (R^2) which shows how well model's predictions match the actual soil moisture values [10].

By integrating advanced modeling techniques like SVR and RF with SAR observations, this study seeks to improve the accuracy of soil moisture predictions and develop a valuable tool for farmers and agricultural planners. The results will contribute to more efficient irrigation scheduling, better water resource management, and ultimately enhanced agricultural productivity [11]. Furthermore, this research emphasizes the potential of remote sensing in combined with ML to address challenges in precision agriculture which offers a scalable and cost-effective solution for soil moisture monitoring on a regional scale. This approach advances our understanding of soil moisture dynamics and leverages cutting-edge technology which helps to develop a sustainable agricultural practice.

The study uses two machine learning models to predict soil moisture from the sentinel-1A SAR backscattering data. The models are Support Vector Regression (SVR) and Random Forest Regression (RFR) where both SVR and RFR offer unique strengths in modeling the complex relationship between SAR backscattering coefficients (σ^0) and soil moisture [12, 13].

2. STUDY AREA AND DATASET

2.1 Study Area:

In this study, we focus on the Beed district, located in Maharashtra, India, which encompasses approximately 10,693 square kilometers. The region experiences a semi-arid climate with annual rainfall ranging from 600 to 800 millimeters, primarily during the southwest monsoon season. The predominant soil types include black soil, red soil, and alluvial soil, supporting the cultivation of crops such as cotton, pulses, sugarcane, and soybeans.

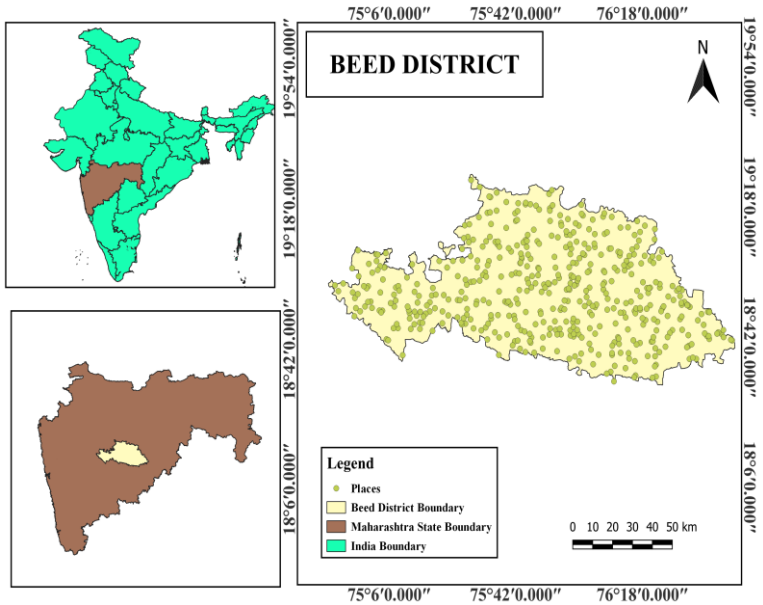

Figure 1. Study Area

2.2 Dataset:

The following Sentinel SAR data is used for the study, dating from 1st January 2023 to 30th January 2023.

S1A_IW_GRDH_1SDV_20230101T005459_20230101T005524_046583_059524_C538

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

S1A_IW_GRDH_1SDV_20230101T005524_20230101T005549_046583_059524_B5A4

S1A_IW_GRDH_1SDV_20230108T004653_20230108T004718_046685_059889_2E9C

S1A_IW_GRDH_1SDV_20230125T005458_20230125T005523_046933_05A0F3_A491

S1A_IW_GRDH_1SDV_20230125T005523_20230125T005548_046933_05A0F3_52DF

After mosaicking all these data sets, we have finally got the subset of our study area.

For auxiliary data, we have used the Google Earth Engine (GEE). Using GEE, we have calculated MODIS Temperature, Soil Moisture (from SMAP), DEM, RVI and calculated the Soil Dielectric Constant using simplified Dubois model for which soil moisture from SMAP is used.

3. METHODOLOGY

The present study focuses on using Sentinel-1A observed data for soil moisture prediction. Sentinel 1A is a European Space Agency (ESA) satellite equipped with SAR sensor which uses C band, so it is used to acquire images [14]. Using these images backscattering coefficients (σ^0) were calculated for σ^0VV (vertical transmit and receive polarization), σ^0VH (vertical transmit and horizontal receive polarization), σ^0VV + σ^0VH, and the Radar Vegetation Index (RVI) is also calculated [14]. RVI (Equation 1) serves as an alternative to NDVI and is important for monitoring vegetation health [15].

Backscattering values are used to calculate RVI because they are very sensitive towards soil moisture [16].

$$RVI = \frac{4 \times \sigma^0 VH}{\sigma^0 VV + \sigma^0 VH,} \quad (1)$$

3.1 SAR Data acquisition and processing:

The Sentinel 1A SAR data was acquired from the Copernicus Data Platform for the study area over the area of interest (AOI). The pre-processing involved several essential steps, including radiometric correction, thermal noise removal, calibration, the orbit file, speckle filtering, geometric correction, and terrain correction [17]. Following these steps, the backscattering coefficients (σ^0VV, σ^0VH, and σ^0VV + σ^0VH) were extracted from the preprocessed Sentinel-1 images of the study area. All pre-processing tasks were performed using the SNAP software [18]. The Radar Vegetation Index (RVI) is calculated using σ^0VV and σ^0VH, as previous research has demonstrated a strong correlation between RVI and soil moisture (SM). Sentinel-1 data, with its short temporal resolution, is good for land studies due to its short revisit time, this short revisit cycle makes Sentinel 1A highly suitable for monitoring dynamic surface processes like soil moisture [19].

These datasets are then used as inputs for machine learning models, namely support vector regression (SVR) and random forest (RF), to predict soil moisture. Then the data is split into training (80%) and testing (20%) sets [20]. Model performance is evaluated by comparing predicted soil moisture values against reference data (SMAP) using the correlation coefficient (r), root mean square error (RMSE), and R-squared (R^2) metrics [21, 22, 23].

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

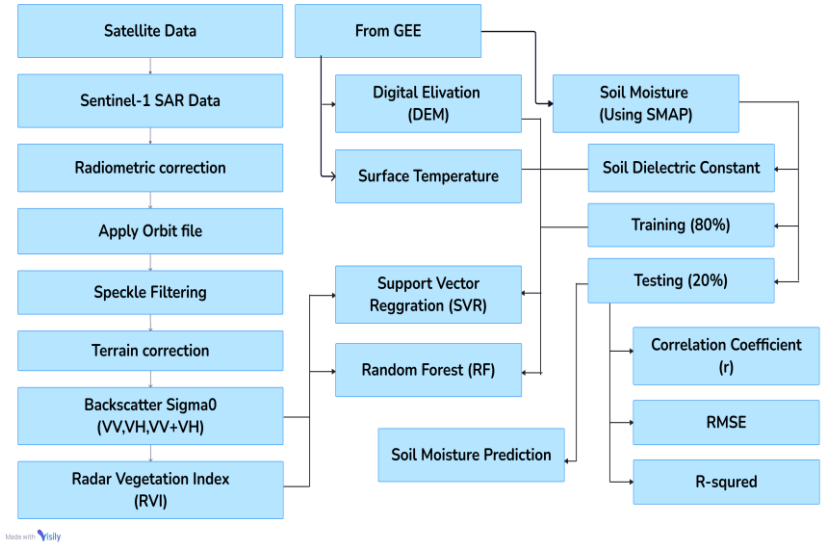

Fig. 2. Flowchart depicting the ML-based soil moisture predictive modeling used in this study.

3.2 Machine learning algorithms:

The main objective of the machine learning models is to estimate soil moisture using the backscattering coefficient (σ^0) which is obtained from Sentinel 1A SAR data, gives imformation about surface characteristics and soil properties like radar vegetation index (RVI) which separate the vegetation signal from the soil signal [24]. Also additional model input like MODIS-derived land surface temperature which shows how thermal environment can impact soil moisture [25], soil moisture information from the SMAP satellite which is used as a reference data for validation [26, 27], a digital elevation model (DEM) that gives topographic imformation [28] and the soil dielectric constant which is estimated using the simplified Dubois model are also used to train the models [29].

Supervised machine learning was chosen because it learns the relationship between input features and reference soil moisture values using training data. This is useful for remote sensing applications because in a remote sensing the development of machine learning model is based on pairing ground based soil moisture measurements (here SMAP data is used) with corresponding satellite and environmental observations. Once trained these machine learning models can predict soil moisture where direct observation is not possible which makes these model highly valuable for large scale soil moisture monitoring [30]. In the present study, two types of ML models support vector regression (SVR) and random forest regression (RFR) are used to estimate and predict the Soil Moisture over the study area.

3.2.1 Support vector Regression (SVR):

As given by vapnik V. (1998), the goal of SVR is to find a function $f(x)$ that approximates the relationship between input features (e.g., SAR backscatter coefficients, RVI, LST, DEM, Soil Dielectric) and the target variable (soil moisture).

For dataset with inputs $x_i \in \mathbb{R}^n$ (e.g. SAR features) and output y_i (Soil moisture) the SVR tries to learn. So, the linear regression function is defined as [31],

$$f(x) = \langle \mathrm{w}, \mathrm{x} \rangle + b \tag{2}$$

Where, x represents the feature vector (e.g., σ^0 VV, σ^0 VH, RVI, DEM, LST, Soil Dielectric), w is the weight vector that defines the direction of the regression function, b is the bias term and $\langle w, x \rangle$ is the inner product.

The model is trained such that predictions remain within an ε-insensitive margin, meaning deviations smaller than ε from the actual soil moisture are not penalized.

The optimization problem for SVR is:

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

$$\min_{w,b,\xi,\xi_i^*} \frac{1}{2} \|w\|^2 c \sum_{i=1}^{n} (\xi_i + \xi_i^*) \quad (3)$$

Subject to,

$$y_i - \langle w, x_i \rangle - b) \leq \varepsilon + \xi_i \quad (4)$$

$$\langle w, x_i \rangle + b - y_i \leq \varepsilon + \xi_i^* \quad (5)$$

$$\xi_i, \xi_i^* \geq 0 \text{ for all i} \quad (6)$$

Where, C is the regularization parameter that balances model complexity and prediction error, ϵ defines the error tolerance margin and ξ_i, ξ_i^* are slack variables that allow certain data points to fall outside the margin.

Since soil moisture retrieval is nonlinear (due to mixed effects of vegetation and soil dielectric properties), SVR uses a kernel function $K(x_i, x_j)$ to project the data into a higher-dimensional feature space [32].

The regression function is given as,

$$f(x) = \sum_{i=1}^{n} (\alpha_i - \alpha_i^*) K(x_i, x) + b \quad (7)$$

Where $f(x)$ is predicted soil moisture, α_i, α_i^* shows weights (decide importance of each training point), x_i istraining data point (features like σ^0VV, σ^0VH, RVI, etc.), x is new input data for prediction, $K(x_i, x)$ is kernel function (measures similarity between training data and input) and b is bias term (adjusts the prediction up or down).

Now The Radial Basis Function Kernel is given as,

$$K(x_i, x) = \exp(-y\|x_i - x\|^2) \quad (8)$$

Where $K(x_i, x)$ shows similarity score between training point x_i and new input x, $\|x_i - x\|^2$is a squared distance between the two points, y is the parameter that controls how far the influence of training point spreads. If two points are very close, $K(x_i, x) \approx 1$ and if they are far apart, $K(x_i, x) \approx 0$.

This kernel captures complex nonlinear dependencies between SAR backscatter and soil moisture, making SVR well-suited for remote sensing applications [31, 32].

3.2.2 Random Forest Regression (RFR):

Random Forest Regression (RFR), first introduced by Breiman (2001), was applied in this study to estimate soil moisture using a diverse set of remote sensing and environmental inputs. Unlike a single decision tree, which often suffers from overfitting, RFR constructs an ensemble of trees and averages their predictions, which substantially improves generalization performance. Each tree is trained on a bootstrap sample of the dataset, represented as,

$$D_i = \{(x_i, y_i), (x_2, y_2), \ldots, (x_n, y_n)\} \quad (9)$$

where x represents the predictor variables (Sentinel-1 SAR backscatter coefficients, MODIS-derived land surface temperature, radar vegetation index, digital elevation model, and soil dielectric constant from the Dubois model), while y denotes the reference soil moisture observations from SMAP. By using sampling with replacement, every tree receives a slightly different subset of the training data, which increases diversity and robustness in the forest.

Within each tree, a subset of features is randomly selected at every split. The best split is chosen by minimizing the prediction error in terms of the residual sum of squares, expressed as

$$j^*, t^* = arg\ min_{j,t} \left[\sum_{x_i \in R_1(j,t)} (y_i - \bar{y}R_1)^2 + \sum_{x_i \in R_2(j,t)} (y_i - \bar{y}R_2)^2 \right] \quad (10)$$

Here, R_1 and R_2 are the partitions created by a candidate split, and $\bar{y}R$ is the average soil moisture in region R. The leaf node prediction for each tree is then computed as the mean value of all training samples that fall into that node,

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

$$\hat{y}R = \frac{1}{|R|}\sum_{x_i \in R} y_i \tag{11}$$

Once all trees are built, the model combines their results by averaging the outputs, which reduces variance and stabilizes predictions. The aggregated prediction for an input x is given by

$$\hat{y}(x) = \frac{1}{B} t_b(x) \tag{12}$$

where B is the total number of trees, and $t_b(x)$ is the prediction of the b^{th} tree.

Breiman (2001) further demonstrated that the generalization error of a random forest depends on two factors, the strength of individual trees (s) and the correlation between them (ρ). This relationship is expressed as

$$E = \rho \frac{1-s^2}{s^2} \tag{13}$$

The equation highlights that RFR achieves better accuracy when the trees are strong predictors while remaining weakly correlated with one another. Beyond prediction, Random Forest provides valuable insights into variable importance [33].

The Random Forest Regression (RFR) regression model is developed to predict soil moisture by leveraging multisource remote sensing and environmental variables, including Sentinel-1 backscatter coefficients, MODIS derived temperature, digital elevation data, radar vegetation index (RVI) and dielectric constant based on the Dubois model. Finaly to optimize model performance, hyperparameters were tuned using a five-fold cross-validated grid search. The number of trees, maximum depth, and number of features considered at each split were carefully adjusted to strike a balance between bias and variance, thereby ensuring reliable and robust soil moisture predictions.

4. RESULT AND DISCUSSION

4.1 Analysis of the Auxiliary datasets and backscattering:

The relationships between soil moisture (SM) and radar-derived parameters, including the Modis temperature, RVI, backscattering VV, backscattering VH, dielectric constant, and DEM are shown in the scatterplot matrix (Figure 3). In scatterplot the relationships between the variables and their corresponding correlation coefficients are displayed in each cell.

This pair plot visualizes the relationships among seven geospatial and environmental variables: Backscatter VH, Backscatter VV, DEM (Digital Elevation Model), Dielectric Constant (Dubois Model), RVI (Radar Vegetation Index), Temperature, and Soil Moisture. On the diagonal, each variable's histogram shows its individual distribution. Below the diagonal, scatter plots with regression lines shows the pairwise relationships, while above the diagonal, the Pearson correlation coefficients (r) quantify the strength and direction of linear relationships. The strongest positive correlation is observed between dielectric constant and RVI (r = 0.89), indicating a high degree of association likely due to their mutual sensitivity to vegetation and moisture. Backscatter VV is strongly correlated with both dielectric constant (r = 0.76) and RVI (r = 0.75), demonstrating its effectiveness in measuring vegetation and moisture content. A moderate positive correlation exists between Backscatter VH and

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

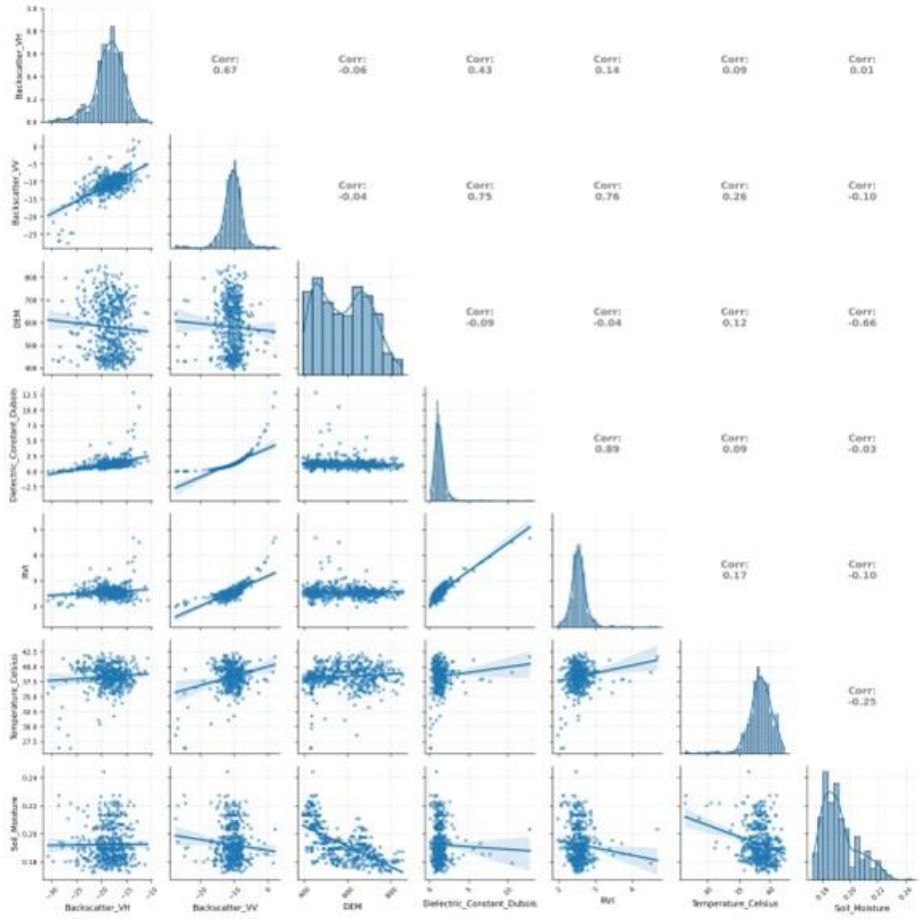

Fig. 3. Relationship between SMAP Soil Moisture, SAR Backscatter, and Auxiliary Features

Backscatter VV (r = 0.67), suggesting that both polarization channels capture related surface features. On the other hand, DEM exhibits a strong negative correlation with soil moisture (r = -0.66), indicating that higher elevations are generally associated with lower soil moisture levels. Similarly, temperature and soil moisture show a weak negative correlation (r = -0.25), reflecting drier conditions with increasing temperature. Other weaker correlations include Backscatter VH with Dielectric Constant (r = 0.43), RVI with Temperature (r = 0.17), and DEM with Temperature (r = 0.12). Several pairs show negligible correlation, such as Backscatter VH vs. DEM (r = -0.06), Backscatter VH vs. RVI (r = 0.14), Backscatter VV vs. DEM (r = -0.04), and Backscatter VH vs. Soil Moisture (r = 0.01).

Overall, the plot reveals valuable insights for environmental modeling. Dielectric Constant, RVI, and Backscatter VV are closely interrelated and may serve as effective predictors of vegetation or moisture, while DEM and temperature influence soil moisture inversely. These relationships are essential for hydrological, ecological, and remote sensing analyses.

4.2 Assessment of Machine Learning Model Performance:

A Taylor diagram (Figure 4) visually summarizes how well a model's predictions match observed data by combining three key statistics: the standard deviation (indicating variability), the correlation coefficient (measuring pattern similarity), and the centered root mean square difference (CRMSD) (representing the error in pattern shape). In this polar plot, models closer to the reference point with matching standard deviation and a small angle (high correlation) are considered more accurate. Taylor diagrams provide a concise and insightful evaluation of both accuracy and variability in a single view.

Taylor diagrams illustrate the performance of Support Vector Regression (SVR) and Random Forest Regression (RFR) models for soil moisture prediction on both training and testing datasets. Each point represents a model's standard deviation (SD), centered root mean square difference (CRMSD) and correlation coefficient relative to

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

observed soil moisture values. For the training data, RF achieved a correlation of 0.94, an SD of 0.0105, and a CRMSD of 0.0038, whereas SVR showed a correlation of 0.87, an SD of 0.0106, and a CRMSD of 0.0048. On the test data, RF maintained superior generalization with a correlation of 0.90.

Standard Deviation of 0.0104 and CRMSD of 0.0055, compared to SVR's correlation of 0.82, SD of 0.0095, and CRMSD of 0.0067. The reference point (black dot) denotes the observed data with SD values of 0.0123 (train) and 0.0142 (test). These results indicate that RF gives more accurate and stable predictions across both datasets.

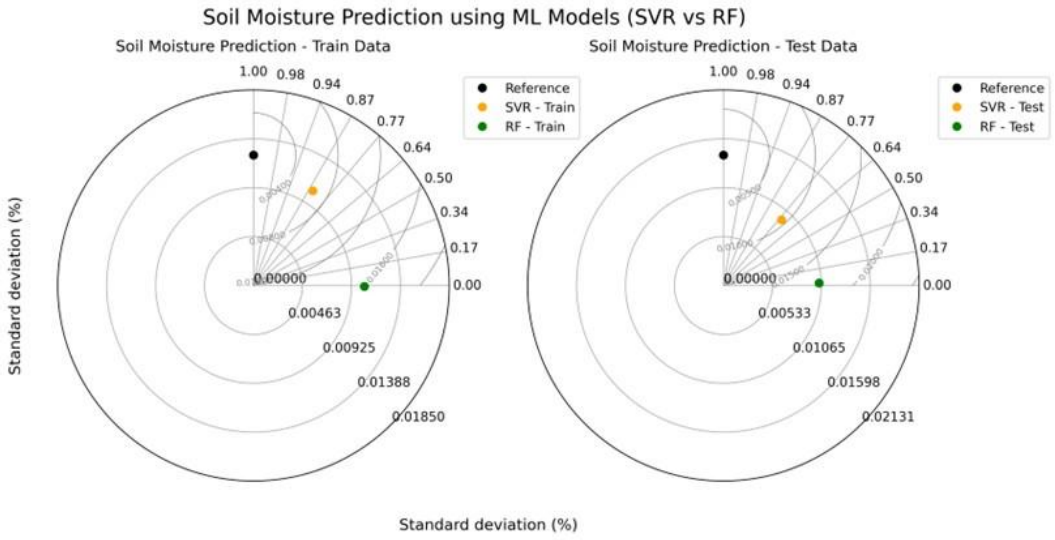

Fig. 4. Visualization of ML Model Performance via Taylor Diagrams

4.3 Evaluation of Machine Learning Models Using Statistical Metrics:

Table 1. Statistical Evaluation of the ML Models Implemented in This Study.

ML Models used	Training Results				Validation Results			
	r	RMES (%)	R-squared	Bias (%)	r	RMES (%)	R-squared	Bias (%)
SVR	0.85	3.41	0.72	0.29	0.75	4.93	0.55	0.55
RF	0.95	2.17	0.89	-0.001	0.72	4.97	0.52	-0.09

As shown in Table 1, Support Vector Regression (SVR) and Random Forest Regression (RFR) were evaluated to predict soil moisture based on backscatter and ancillary environmental features. For the training dataset, the SVR model achieved a correlation coefficient (r) of 0.85, RMSE of 3.41%, R^2 of 0.72, and a bias of 0.29%. During validation, its performance slightly decreased, with r = 0.75, RMSE = 4.93%, R^2 = 0.55, and bias = 0.55%. In contrast, the RF model demonstrated superior performance during training, with r = 0.95, RMSE = 2.17%, R^2 = 0.89, and a negligible bias of -0.001%. For validation, the RF model maintained competitive accuracy, recording r = 0.72, RMSE = 4.97%, R^2 = 0.52, and a bias of -0.09%. Overall, although both models showed strong performance, the Random Forest model demonstrated slightly superior training accuracy and generalization capability compared to SVR.

4.4 Assessment of the ML models:

Figure 5 illustrates the comparison between predicted and observed soil moisture (SM%) values for the two machine learning models, Support Vector Regression (SVR) and Random Forest Regression (RFR), across both training and testing datasets. The solid black one-to-one line represents ideal prediction accuracy, where predicted values perfectly match observed values. In the training data plot, both models demonstrate reasonably good performance, with RF outperforming SVR (RMSE = 0.004, R^2 = 0.887 vs. RMSE = 0.007, R^2 = 0.716), as evidenced by a tighter

Journal of Information Systems Engineering and Management
2024, 9(4)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

clustering of RF predictions around the one-to-one line. This suggests that RF effectively captured underlying patterns in the training set. In the testing dataset, the plots show a wider dispersion of points for both models, reflecting a decrease in prediction accuracy on unseen data. The performance drop is notable for both models (RF: $R^2 = 0.523$, SVR: $R^2 = 0.552$), highlighting potential generalization challenges. These discrepancies underscore the need for model tuning or additional validation approaches to ensure reliability in practical applications, especially where accurate soil moisture prediction is critical for environmental monitoring and agricultural planning.

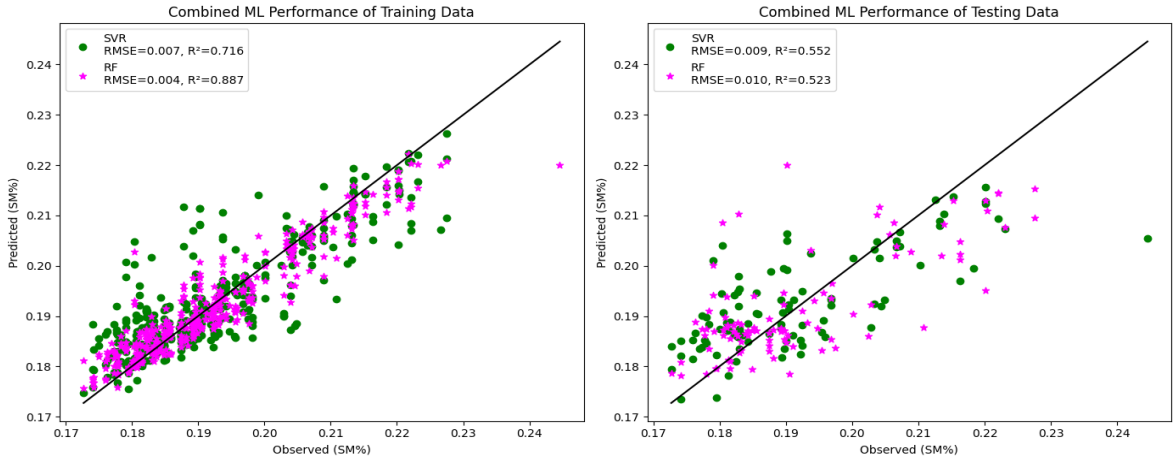

Fig. 5. Combined ML performance of training and testing

5. CONCLUSION

This study explored the predictive capability of radar backscatter and auxiliary environmental datasets for soil moisture estimation using machine learning techniques. The analysis of input variables revealed meaningful relationships, particularly the strong positive correlations among dielectric constant, radar vegetation index (RVI), and VV backscatter, underscoring their importance in moisture and vegetation modeling. Conversely, the Digital Elevation Model (DEM) and temperature exhibited inverse relationships with soil moisture, aligning with expected hydrological patterns in higher and warmer regions.

Evaluation through Taylor diagrams and statistical measures showed that RFR achieved superior performance compared to SVR in both training and testing phases. With improved correlation, reduced error, and lower bias, RFR proved more accurate and reliable in handling nonlinear relationships compared to SVR. However, both models exhibited performance declines during validation, emphasizing the need for further tuning and possibly incorporating additional features or ensemble strategies to enhance generalization.

Visual assessments of predicted versus observed values show RF's superior pattern capture in training data, though testing accuracy highlighted some generalization limitations. This study shows that the fusion of radar derived data and environmental variables enhances the accuracy of machine learning based soil moisture estimation. The findings support the broader application of RF-based approaches in environmental monitoring and precision agriculture while also suggesting the need for repeated refinement to ensure robust model deployment under varied conditions.

REFERENCES

1. Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., Ruiperez Gonzalez, M., & Kilibarda, M. (2017). SoilGrids250m: Global gridded soil information based on machine learning. PLoS ONE, 12(2), e0169748. https://doi.org/10.1371/journal.pone.0169748
2. Zhao-Liang Li, Pei Leng, Chenghu Zhou, Kun-Shan Chen, Fang-Cheng Zhou, Guo-Fei Shang, Soil moisture retrieval from remote sensing measurements: Current knowledge and directions for the future, *Earth-Science Reviews*, Volume 218, 2021, 103673, ISSN 0012-8252, https://doi.org/10.1016/j.earscirev.2021.103673.
3. Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., & Crow, W. T. (2010). The Soil Moisture Active Passive (SMAP) mission. *Proceedings of the IEEE*, *98*(5), 704–716. https://doi.org/10.1109/JPROC.2010.2043918

4. Wagner, W., Lemoine, G., & Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and SAR data. *Remote Sensing of Environment*, *70*(2), 191-207. https://doi.org/10.1016/S0034-4257(99)00036-X
5. Stanyer, C.; Seco-Rizo, I.; Atzberger, C.; Marti-Cardona, B. Soil Texture, Soil Moisture, and Sentinel-1 Backscattering: Towards the Retrieval of Field-Scale Soil Hydrological Properties. *Remote Sens.* 2025, 17, 542. https://doi.org/10.3390/rs17030542
6. Rajeev, H., Gururaj, P. & Pathak, A.A. Dynamic monitoring of surface soil moisture fluctuations using synthetic aperture radar and data-driven algorithms. *Appl Geomat* 17, 1–15 (2025). https://doi.org/10.1007/s12518-024-00606-2
7. Narendra N. Das, Dara Entekhabi, R. Scott Dunbar, Mario J. Chaubell, Andreas Colliander, Simon Yueh, Thomas Jagdhuber, Fan Chen, Wade Crow, Peggy E. O'Neill, Jeffrey P. Walker, Aaron Berg, David D. Bosch, Todd Caldwell, Michael H. Cosh, Chandra H. Collins, Ernesto Lopez-Baeza, Marc Thibeault, The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product, *Remote Sensing of Environment*, Volume 233, 2019, 111380, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2019.111380.
8. Kim, S., Kim, H., Kwon, Y., & Nguyen, H. H. (2025). A stand-alone framework for predicting spatiotemporal errors in satellite-based soil moisture using tree-based models and deep neural networks. *GIScience & Remote Sensing*, *62*(1). https://doi.org/10.1080/15481603.2025.2475572
9. Naeini, S. & Komaki, Chooghi Bairam & Verrelst, Jochem & Kakooei, Mohammad & Mahmoodi, Mohammad. (2023). Satellite-Based Estimation of Soil Moisture Content in Croplands: A Case Study in Golestan Province, North of Iran. *Remote Sensing*. 15. 2155. 10.3390/rs15082155.
10. Beck, H. E., Pan, M., Miralles, D. G., Reichle, R. H., Dorigo, W. A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R. M., van Dijk, A. I. J. M., Du, J., Kimball, J. S., Vergopolan, N., and Wood, E. F.: Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors, *Hydrol. Earth Syst. Sci.*, 25, 17–40, https://doi.org/10.5194/hess-25-17-2021, 2021.
11. Mohammad Amin Shahriari, Hossein Aghighi, Mohsen Azadbakht, Davoud Ashourloo, Ali Akbar Matkan, Foad Brakhasi, Jeffrey P. Walker, Soil moisture estimation using combined SAR and optical imagery: Application of seasonal machine learning algorithms, *Advances in Space Research*, Volume 75, Issue 8, 2025, Pages 6207-6221, ISSN 0273-1177, https://doi.org/10.1016/j.asr.2025.01.064.
12. Hoa, P.V.; Giang, N.V.; Binh, N.A.; Hai, L.V.H.; Pham, T.-D.; Hasanlou, M.; Tien Bui, D. Soil Salinity Mapping Using SAR Sentinel-1 Data and Advanced Machine Learning Algorithms: A Case Study at Ben Tre Province of the Mekong River Delta (Vietnam). *Remote Sens.* 2019, *11*, 128. https://doi.org/10.3390/rs11020128
13. Jianjun Wang, Fei Wu, Jiali Shang, Qi Zhou, Irshad Ahmad, Guisheng Zhou, Saline soil moisture mapping using Sentinel-1A synthetic aperture radar data and machine learning algorithms in humid region of China's east coast, *CATENA*, Volume 213, 2022, 106189, ISSN 0341-8162, https://doi.org/10.1016/j.catena.2022.106189.
14. https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1
15. Xueqian Hu, Li Li, Jianxi Huang, Yelu Zeng, Shuo Zhang, Yiran Su, Yujiao Hong, Zixiang Hong, Radar vegetation indices for monitoring surface vegetation: Developments, challenges, and trends, *Science of The Total Environment*, Volume 945, 2024, 173974, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2024.173974.
16. Li, J.; Wang, S. Using SAR-Derived Vegetation Descriptors in a Water Cloud Model to Improve Soil Moisture Retrieval. *Remote Sens.* 2018, *10*, 1370. https://doi.org/10.3390/rs10091370
17. https://step.esa.int/main/toolboxes/sentinel-1-toolbox/
18. Filipponi, F. Sentinel-1 GRD Preprocessing Workflow. *Proceedings* 2019, *18*, 11. https://doi.org/10.3390/ECRS-3-06201
19. Nasirzadehdizaji, R.; Balik Sanli, F.; Abdikan, S.; Cakir, Z.; Sekertekin, A.; Ustuner, M. Sensitivity Analysis of Multi-Temporal Sentinel-1 SAR Parameters to Crop Height and Canopy Coverage. *Appl. Sci.* 2019, *9*, 655. https://doi.org/10.3390/app9040655
20. Jia, Y.; Jin, S.; Savi, P.; Yan, Q.; Li, W. Modeling and Theoretical Analysis of GNSS-R Soil Moisture Retrieval Based on the Random Forest and Support Vector Machine Learning Approach. *Remote Sens.* 2020, *12*, 3679. https://doi.org/10.3390/rs12223679
21. Mohseni, F.; Mirmazloumi, S.M.; Mokhtarzade, M.; Jamali, S.; Homayouni, S. Global Evaluation of SMAP/Sentinel-1 Soil Moisture Products. *Remote Sens.* 2022, *14*, 4624. https://doi.org/10.3390/rs14184624

22. Tola, D.; Bustillos, L.; Arragan, F.; Chipana, R.; Hostache, R.; Resongles, E.; Espinoza-Villar, R.; Zolá, R.P.; Uscamayta, E.; Perez-Flores, M.; et al. High Spatial Resolution Soil Moisture Mapping over Agricultural Field Integrating SMAP, IMERG, and Sentinel-1 Data in Machine Learning Models. *Remote Sens.* 2025, *17*, 2129. https://doi.org/10.3390/rs17132129
23. Karamvand, A., Hosseini, S.A. & Sharafati, A. SMAP products for prediction of surface soil moisture by ELM network model and agricultural drought index. *Acta Geophys.* 71, 1845–1856 (2023). https://doi.org/10.1007/s11600-022-00973-7
24. Jyoti Sharma, Rajendra Prasad, Prashant K. Srivastava, Shubham K. Singh, Suraj A. Yadav, Dharmendra K. Pandey, Improved radar vegetation water content integration for SMAP soil moisture retrieval, *Remote Sensing Applications: Society and Environment*, Volume 37, 2025, 101443, ISSN 2352-9385, https://doi.org/10.1016/j.rsase.2024.101443.
25. Kang, J.; Tan, J.; Jin, R.; Li, X.; Zhang, Y. Reconstruction of MODIS Land Surface Temperature Products Based on Multi-Temporal Information. *Remote Sens.* 2018, *10*, 1112. https://doi.org/10.3390/rs10071112
26. Chuanxiang Yi, Xiaojun Li, Jiangyuan Zeng, Lei Fan, Zhiqing Xie, Lun Gao, Zanpin Xing, Hongliang Ma, Antoine Boudah, Hongwei Zhou, Wenjun Zhou, Ye Sheng, Tianxiang Dong, Jean-Pierre Wigneron, Assessment of five SMAP soil moisture products using ISMN ground-based measurements over varied environmental conditions, *Journal of Hydrology*, Volume 619, 2023, 129325, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2023.129325.
27. Hongliang Ma, Jiangyuan Zeng, Xiang Zhang, Jian Peng, Xiaojun Li, Peng Fu, Michael H. Cosh, Husi Letu, Shaohua Wang, Nengcheng Chen, Jean-Pierre Wigneron, Surface soil moisture from combined active and passive microwave observations: Integrating ASCAT and SMAP observations based on machine learning approaches, *Remote Sensing of Environment*, Volume 308, 2024, 114197, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2024.114197.
28. Chukwuma J. Okolie, Julian L. Smit, A systematic review and meta-analysis of Digital elevation model (DEM) fusion: pre-processing, methods and applications, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 188, 2022, Pages 1-29, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2022.03.016.
29. Tania Neusch, Manfred Sties, Application of the Dubois-model using experimental synthetic aperture radar data for the determination of soil moisture and surface roughness, *ISPRS Journal of Photogrammetry and Remote Sensing*, Volume 54, Issue 4, 1999, Pages 273-278, ISSN 0924-2716, https://doi.org/10.1016/S0924-2716(99)00019-2.
30. Lamichhane, M.; Mehan, S.; Mankin, K.R. Soil Moisture Prediction Using Remote Sensing and Machine Learning Algorithms: A Review on Progress, Challenges, and Opportunities. *Remote Sens.* 2025, *17*, 2397. https://doi.org/10.3390/rs17142397.
31. Vapnik V. Statistical learning theory. New York: Wiley; 1998.
32. Smola AJ, Schölkopf B, Muller K-R. The connection between regularization operators and support vector kernels. Neural Networks 1998;11: 637–49. doi:10.1016/S0893-6080(98)00032-X.
33. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324.