AGU ADVANCING EARTH AND SPACE SCIENCES

JGR Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2026JH001322

Transferable Neural Operators to Estimate Subsurface Soil Moisture

Abhilash Singh[1], M. Ravikumar[2], Vidhi Singh[3], Shivprasad More[4], Virendra M. Tiwari[4], and Kumar Gaurav[3]

[1]School of Mathematics, Faculty of Engineering and Physical Sciences, University of Leeds, Leeds, UK, [2]CSIR-National Geophysical Research Institute, Hyderabad, India, [3]Fluvial Geomorphology and Remote Sensing Laboratory, Department of Earth and Environmental Sciences, Indian Institute of Science Education and Research Bhopal, Bhopal, India, [4]CSIR-North East Institute of Science and Technology, Jorhat, India

Key Points:

- Developed a transferable neural operator to estimate subsurface soil moisture from surface soil moisture
- Fine-tuning with just 5% of data yields near-optimal model adaptation, minimizing the need for extensive training data at new locations
- Progressive fine-tuning reduces bias uncertainty, enhancing prediction stability and reliability across environmental settings

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. Singh and K. Gaurav,
a.singh4@leeds.ac.uk;
kgaurav@iiserb.ac.in

Citation:

Singh, A., Ravikumar, M., Singh, V., More, S., Tiwari, V. M., & Gaurav, K. (2026). Transferable neural operators to estimate subsurface soil moisture. *Journal of Geophysical Research: Machine Learning and Computation*, *3*, e2026JH001322. https://doi.org/10.1029/2026JH001322

Received 4 MAR 2026
Accepted 17 JUL 2026

Author Contributions:

Conceptualization: Abhilash Singh, Vidhi Singh, Virendra M. Tiwari, Kumar Gaurav
Data curation: M. Ravikumar, Shivprasad More, Virendra M. Tiwari
Funding acquisition: Abhilash Singh, Virendra M. Tiwari, Kumar Gaurav
Investigation: Abhilash Singh, Vidhi Singh, Kumar Gaurav
Methodology: Abhilash Singh, Vidhi Singh, Kumar Gaurav
Software: Abhilash Singh, Vidhi Singh

Abstract

Access to accurate and timely soil moisture information is critical for effective water resource management, ensuring food security, and advancing climate modeling; however, such information is often limited. We propose a novel transferable Fourier Neural Operator (FNO) framework designed to estimate subsurface soil moisture utilizing surface-level data (5 cm). This framework demonstrates robust generalization capabilities for soil moisture estimates for deeper soil layers, wherein prediction tasks are inherently more complex. The model has been trained on high-resolution temporal data from six distinct sites encompassing both arid and humid subtropical regions, as well as various soil types. It reliably predicts soil moisture at depths of 15 and 50 cm. By leveraging transfer learning, pre-training, and fine-tuning with only 5% of the target domain data, we achieved improved accuracy and stability. This innovative framework provides a scalable and data-efficient solution for delivering subsurface soil moisture data to users with constrained data sets.

Plain Language Summary

This research introduces a new artificial intelligence method that estimates subsurface soil moisture at various depths by using surface measurements. In simple terms, instead of digging deep into the soil to measure moisture at every level, the model collects data from just the top layer (about 5 cm deep) and uses this information to predict moisture deeper down (at 15 and 50 cm). The approach has been tested in diverse scenarios, covering two different climates, and it showed that even with a limited amount of local data, the model could quickly adjust and deliver strong performance, proving its transferability.

1. Introduction

Subsurface soil moisture information is important in understanding water availability, root zone processes, and the response of the ecosystem to climatic variability (Blume et al., 2009). However, it is often challenging and expensive to measure subsurface soil moisture due to the use of specialized sensors and invasive measurement techniques (Hummel et al., 2001). On the other hand, surface soil moisture is accessed relatively easily using remote sensing and in situ observations (Kumar et al., 2009; McColl et al., 2017; Singh & Gaurav, 2024). This asymmetry motivates the development of new computational models that can make good estimates of subsurface soil moisture profiles from easily measurable surface soil moisture information (Singh, Singh, & Gaurav, 2025b, 2026).

Deep learning models have shown their capability to learn complex patterns in high-resolution spatial-temporal data (S. Wang et al., 2020). Among them, the Fourier Neural Operator (FNO) has shown strong potential because it learns mappings between function spaces using spectral representations, allowing it to capture both local variability and long-range dependencies (Alipour et al., 2026; Azizzadenesheli et al., 2024; Li et al., 2026; Liu et al., 2023; Singh, Singh, & Gaurav, 2025a; Y. Wang et al., 2024). This property is particularly relevant for soil moisture estimation, where surface and subsurface moisture are connected through nonlinear, time-dependent processes influenced by rainfall, evapotranspiration, soil texture, and climatic variability (Xi et al., 2025). However, most deep learning models, including neural-operator-based models, are commonly trained and evaluated within the same site or domain (Nunez et al., 2025). Their ability to generalize across locations with different soil moisture distributions and climatic conditions remains less explored, especially for subsurface soil moisture estimation (Singh, Niranjannaik, & Gaurav, 2025).

Supervision: Kumar Gaurav
Validation: Abhilash Singh, Vidhi Singh
Visualization: Abhilash Singh, Vidhi Singh, Virendra M. Tiwari, Kumar Gaurav
Writing – original draft: Abhilash Singh, Vidhi Singh, Kumar Gaurav
Writing – review & editing: Abhilash Singh, Vidhi Singh, Virendra M. Tiwari, Kumar Gaurav

This limitation motivates the use of transfer learning. Transfer learning is a valuable solution to mitigate the problem of a lack of direct subsurface soil moisture information in data-scarce regions (Hutchinson et al., 2017). By reusing models trained on one domain or site (known as the source domain) and fine-tuning them on data from another (known as the target domain), transfer learning leverages rich, pre-learned representations effectively (Theodoris et al., 2023; Yang et al., 2024). For soil moisture estimation, this method addresses the lack of subsurface data by relying on the easily available surface measurements from many different areas. Adapting a model trained on data from a well-instrumented site to new, less densely monitored sites enables the extension of accurate subsurface soil moisture estimation across broader spatial scales. This can efficiently solve the problem of data scarcity.

This study aims to develop a transferable FNO model designed to estimate subsurface soil moisture at 15 and 50 cm depths using surface soil moisture measurements recorded at 5 cm. The FNO leverages spectral methods to capture both local interactions and long-range dependencies in the high-resolution temporal data, which is collected at 15-min intervals over a period of 3 years. Data from six diverse sites are used to rigorously evaluate the model's transferability. By fine-tuning the pre-trained FNO model on a varying subset of data from target sites, we comprehensively assess its ability to generalize across different environmental conditions. This approach directly addresses the need for a model that can learn complex temporal soil moisture dynamics while remaining adaptable to new locations where only limited target-domain data are available. Our work addresses the following research questions:

1. How accurately does the proposed FNO model estimate subsurface soil moisture using surface soil moisture measurements across different study sites?
2. What is the baseline transferability performance of the model when it is trained on one site and directly applied to other sites without any fine-tuning?
3. What is the optimal fine-tuning percentage of available data that achieves the best trade-off between computational cost and estimation accuracy, particularly for deeper soil layers?
4. Does increasing the proportion of fine-tuning data lead to a reduction in the model's prediction bias?

To the best of our knowledge, no comprehensive study has yet evaluated the transferability of the FNO for soil moisture estimation, particularly when it comes to subsurface soil moisture. This study demonstrates the practical benefits of transfer learning in hydrological applications.

2. Material and Methods

2.1. Data Sets

We used high-resolution soil moisture data from six different locations in India. The soil moisture monitoring stations are operational at Jawahar Navodaya Vidyalayas (JNV) in Sirsa (Haryana), Gorakhpur (Uttar Pradesh), Aurangabad and Sitamarhi (Bihar), at Bundelkhand University in Jhansi (Uttar Pradesh), and the CSIR-Central Building Research Institute (CBRI) in Roorkee (Uttarakhand). These sites span latitudes from 24.871°N to 29.863°N and longitudes from 74.903°E to 85.529°E. Based on the Köppen-Geiger climate classification (Beck et al., 2018), the Sirsa station in Haryana falls within an arid region (BWh), while the remaining stations are located in areas characterized by a humid subtropical climate (Cwa) (Figure 1a).

Details regarding the HydraProbe sensor specifications, calibration procedures, and installation methodology are provided in Section S1 in Supporting Information S1. A time-series assessment of soil moisture across all stations reveals strong seasonal variability driven by the Indian Summer Monsoon, with distinct differences between arid and humid subtropical zones (Figure S1 in Supporting Information S1). These temporal dynamics, along with depth-wise variation and site-specific anomalies, are discussed in detail in Section S2 in Supporting Information S1 (Seyfried et al., 2005; Tian et al., 2023).

2.2. Model Architecture, Optimization and Development

In this study, our primary objective is to develop a transferable FNO model that accurately estimates subsurface soil moisture at two distinct depths (15 and 50 cm) by using readily available surface soil moisture measurements recorded at 5 cm. The goal is to construct a mapping function from the surface measurement to deeper soil moisture profiles so that the model can be effectively adapted (or transferred) to estimate subsurface conditions

10.1029/2026JH001322

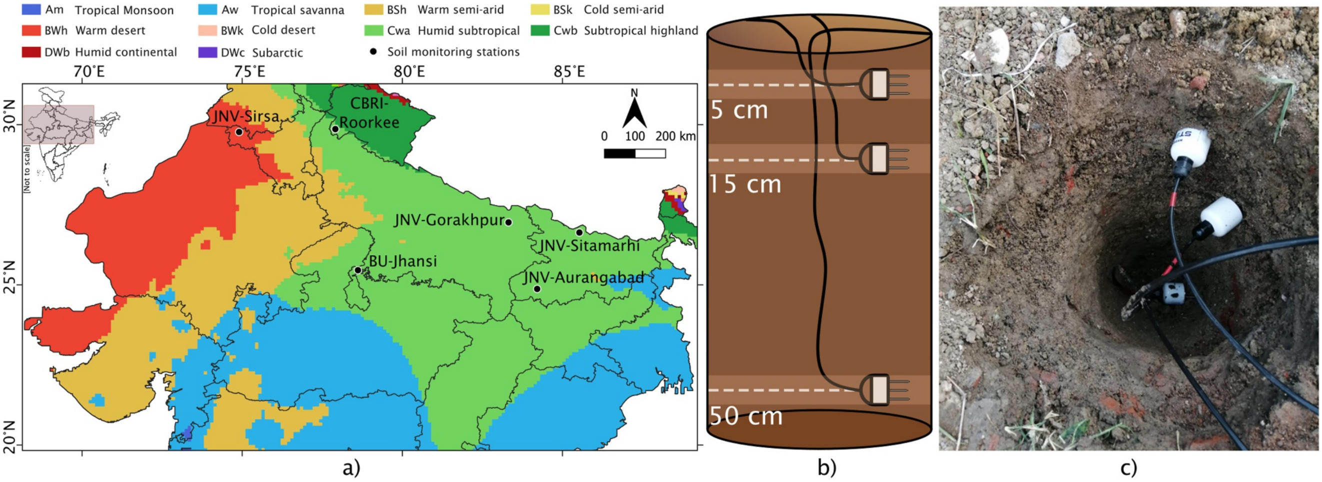

Figure 1. (a) Map showing the locations of six soil monitoring stations across different Indian states overlaid on the Köppen-Geiger climate classification, representing diverse climatic zones. (b) Schematic diagram illustrating soil moisture sensor installation at 5, 15, and 50 cm depths to capture surface-to-subsurface moisture dynamics. (c) Field photograph showing the top-view of sensors deployed in a vertical profile for continuous in situ monitoring.

across different geographical locations. This is particularly valuable in scenarios where direct subsurface measurements are sparse or expensive to obtain.

The algorithm is applied to high temporal resolution time series data, with observations recorded every 15 min over a continuous span of 3 years. Data have been collected from all six sites. Each site provides diverse environmental conditions, ensuring that the developed FNO model has a robust and transferable representation of the characteristics.

The input to the FNO consists of two components. The primary feature is the surface soil moisture measurement, denoted as x_{SM5cm}. The secondary, auxiliary feature is a normalized grid value x_{grid}, derived from a linear spacing between 0 and 1, which serves as a proxy for temporal context. This feature does not represent calendar time, season, or time of day; instead, it provides a positional reference that helps the FNO identify the relative location of each observation within the temporal sequence. These two features are concatenated to form the complete input vector for each measurement point:

$$\mathbf{x} = \begin{bmatrix} x_{\mathrm{SM5cm}} \\ x_{\mathrm{grid}} \end{bmatrix}. \tag{1}$$

The objective of the model is to learn a functional mapping that predicts subsurface soil moisture measurements at 15 and 50 cm. Mathematically, this mapping is defined as:

$$\mathcal{F} : \begin{bmatrix} x_{\mathrm{SM5cm}} \\ x_{\mathrm{grid}} \end{bmatrix} \rightarrow [y_{SMd}], d \in \{15, 50\}\mathrm{cm} \tag{2}$$

where y_{SM15cm} and y_{SM50cm} represent the subsurface soil moisture at 15 cm and 50 cm, respectively.

This formulation enables the FNO to capture both the local interactions from direct surface measurements and the global dependencies along the temporal dimension (e.g., x_{grid}). The auxiliary grid feature helps in capturing systematic variations and supports the generalization process across different sites. Moreover, owing to the high temporal resolution (15-min timesteps), the model is capable of capturing rapid dynamics in soil moisture, thereby providing a detailed and robust estimation that is transferable across the diverse environmental conditions observed in the six different study sites.

2.2.1. Data Processing

The raw data undergo several essential preprocessing steps to ensure its quality and suitability for model training. Initially, a data cleansing procedure is performed where non-numeric entries (primarily indicating missing values) are converted to NaN and subsequently removed from the data set. In addition, outlier values are filtered to enhance overall data quality. Following this, the data is sorted according to time to preserve the inherent sequential structure. Subsequently, a normalized grid is constructed to serve as a temporal coordinate. This grid, when concatenated with the primary feature, forms the complete two-channel input required by the model. Finally, to develop a trained model at each site using its data, we divided the data into a 70:30 split between train and test. To test the fine-tuned model's performance, the other sites' data set is randomly shuffled and divided into two equal parts. One-half constitutes the training pool (50%), which is later used to fine-tune during transfer learning experiments, while the other half is set aside as the test set (50%) to ensure a fair final evaluation.

2.2.2. Model Architecture and Forward Pass

The model leverages spectral methods to capture both global and local dependencies. In the first stage, the low-dimensional input $\mathbf{x}$ is "lifted" into a higher-dimensional latent space via a fully connected layer, yielding

$$\mathbf{h}^{(0)} = \mathrm{fc0}(\mathbf{x}), \tag{3}$$

where $\mathbf{h}^{(0)} \in \mathbb{R}^{N\times w}$ and w is the chosen channel width. This transformation enriches the representation before subsequent spectral processing. The core of the architecture involves a spectral convolution layer. Here, the latent representation is transformed into the frequency domain using the one-dimensional real-valued fast Fourier transform,

$$\hat{\mathbf{h}}(k) = \sum_{n=0}^{N-1} \mathbf{h}(n)\, e^{-i\frac{2\pi}{N}kn}. \tag{4}$$

Here, k ranges from zero to the integer part of $N/2$, corresponding to the nonredundant Fourier coefficients returned by the RFFT. For computational efficiency, only the first M Fourier modes are retained, such that

$$\hat{\mathbf{h}}_{\mathrm{trunc}}(k) = \begin{cases} \hat{\mathbf{h}}(k), & k < M, \\ 0, & k \geq M. \end{cases} \tag{5}$$

In the frequency domain, a set of learnable complex weights $\mathbf{W}(k)$is then applied via element-wise multiplication:

$$\hat{\mathbf{y}}(b,o,k) = \sum_{i=1}^{w} \hat{\mathbf{h}}(b,i,k)\, \mathbf{W}(i,o,k), \tag{6}$$

where b denotes the batch index, i the input channels, o the output channels, and k the Fourier modes. The modified frequency-domain representation is transformed back to the temporal domain by the inverse real-valued fast Fourier transform:

$$\mathbf{y}(n) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{\mathbf{y}}(k)\, e^{i\frac{2\pi}{N}kn}. \tag{7}$$

Parallel to this spectral operation, a pointwise convolution (a one-dimensional convolution with a kernel size of 1) is applied to the latent representation. The outputs of the spectral convolution and the pointwise convolution are summed together, and the resulting feature map is passed through a ReLU activation function to obtain the next layer's output:

$$\mathbf{h}^{(l+1)} = \mathrm{ReLU}\left(\mathrm{SpectralConv1d}\left(\mathbf{h}^{(l)}\right) + \mathrm{Conv1d}_{\mathrm{pointwise}}\left(\mathbf{h}^{(l)}\right)\right). \tag{8}$$

Multiple such convolutional blocks (i.e., four) are stacked to build a rich hierarchical representation that effectively captures both long-range dependencies and local details. Finally, the latent representation is projected back to the target output dimension using additional fully connected layers. Specifically, if $\mathbf{h}^{(L)}$ is the output of the final convolutional block, the projected output is computed as

$$\mathbf{z} = \mathrm{fc2}\left(\mathrm{ReLU}\left(\mathrm{fc1}\left(\mathbf{h}^{(L)}\right)\right)\right), \tag{9}$$

where $\mathbf{z}$ represents the final predicted value (e.g., soil moisture at the target depth).

2.2.3. Backward Pass and Optimization

To quantify the error between the model predictions $\hat{y}$ and the ground truth y, we employ the commonly used Mean Squared Error (MSE) loss defined as

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left(y_i - \hat{y}_i\right)^2. \tag{10}$$

This differentiable loss function is essential for optimizing the model parameters. In the backward pass, the gradient $\nabla_\theta \mathcal{L}$ with respect to the parameters θ is computed via the chain rule. For optimization, we used the Adam optimizer, which updates the model parameters using adaptive estimates of the first and second moments of the gradients (Reyad et al., 2023; Zhang, 2018). Specifically, at iteration t, the parameter update is governed by:

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + \left(1-\beta_1\right)\nabla_\theta \mathcal{L}, \\ v_t &= \beta_2 v_{t-1} + \left(1-\beta_2\right)\left(\nabla_\theta \mathcal{L}\right)^2, \\ \hat{m}_t &= \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t}, \\ \theta_{t+1} &= \theta_t - \zeta\frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}, \end{aligned} \tag{11}$$

where ζ is the learning rate and β_1, β_2 are hyperparameters controlling the momentum and variance estimates, respectively.

2.2.4. Transfer Learning and Fine-Tuning

In many real-world scenarios, a model trained on one domain (source) must be adapted to perform well on a related but distinct domain (target). Transfer learning leverages the knowledge learned from the source by fine-tuning the model on the target data set, which is particularly valuable when the target data are limited. In our approach, each pre-trained model, developed on data from one of the six sites, is evaluated for its transferability by fine-tuning with varying percentages of the available training pool from the target site and subsequently tested on a fixed test set. The target data set is first randomly divided into two equal parts: a training pool (50%) for fine-tuning (from 5% to 50%) and a fixed test set (50%) for evaluation. The random target-site split was used to evaluate transfer-learning adaptation under limited randomly available calibration samples, not to simulate an operational forward-in-time forecasting experiment. Multiple fine-tuning scenarios are explored by varying the percentage $p\%$ of the full target site data set used; specifically, the number of fine-tuning samples is determined as

$$N_{\text{fine}} = \min\left(\left\lfloor \frac{p}{100} \times N \right\rfloor, N_{\text{train pool}}\right), \tag{12}$$

where N represents the total number of observations and $N_{\text{train pool}}$ is the size of the training pool. To maintain fairness in comparison, the pre-trained source model is reinitialized to its original state before each fine-tuning experiment. This ensures that each fine-tuning run starts from the same baseline. The selected subset of the target data is then used to adapt the model parameters by minimizing the MSE loss over a fixed number of epochs.

Importantly, we explore all possible transfer learning scenarios by testing the transferability of the model from each source site to every other target site. This enables us to assess the model adaptability under diverse environmental conditions. Model performance is evaluated using metrics such as Root Mean Squared Error (RMSE), bias, and the Pearson correlation coefficient (R) on the target test set.

Algorithm 1 in Supporting Information S1 details the pseudo-code for our transferable algorithm, outlining its core procedural steps and innovative mechanisms. Table S1 in Supporting Information S1 comprehensively summarizes the model architecture and training parameters used in this study.

3. Results and Discussion

3.1. Subsurface Soil Moisture Estimation

The FNO model effectively captures the relationship between surface and subsurface soil moisture across all data sets, as evidenced by extremely high correlation coefficients ($R > 0.99$) in every case (Figure 2a). Such high R values indicate that, regardless of climatic conditions, the model robustly detects the linear relationship between the measurements at 5 cm (surface) and 15 cm (subsurface). In addition, the RMSE values remain low, suggesting that the predicted values closely match the observed data across the board.

For the humid subtropical environments, represented by BU-Jhansi, CBRI-Roorkee, JNV-Aurangabad, JNV-Gorakhpur, and JNV-Sitamarhi, the metrics are consistently strong. BU-Jhansi records an R value of approximately 0.99, with a modest negative bias of −0.08, an RMSE of 0.39, and a σ around 0.38. Similarly, JNV-Aurangabad shows a slight positive bias (0.05) with an RMSE near 0.40 and a comparable error spread. CBRI-Roorkee and JNV-Gorakhpur yield R values just below 0.995, with RMSE values in the 0.39–0.40 range and biases that indicate a slight tendency to overestimate (positive biases of 0.04 and 0.07, respectively). JNV-Sitamarhi stands out with the highest correlation ($R \approx 0.998$) and the lowest RMSE (0.30) among the humid data sets, along with a bias of 0.05 and a σ of 0.30, indicating particularly precise performance in that specific region. Across all the data sets, the mean error (μ) is close to zero, indicating that the model's errors are evenly balanced around zero and that there is no significant systematic bias in the predictions (Figure 2b).

In contrast, the data set from JNV-Sirsa, which represents an arid region, demonstrates equally remarkable results with an R value of about 0.99. Its RMSE is 0.24, and the bias is almost negligible (−0.01), with an error spread (σ) of 0.24. This suggests that even in an arid setting, where a priori one might expect more challenging conditions due to reduced and highly variable moisture content, the model is capable of delivering precise predictions. The lower RMSE and σ in JNV-Sirsa might partly reflect a smaller absolute range of moisture values typical of arid environments, and the near-zero bias signifies an absence of systematic over- or underestimation.

Overall, the results from all six data sets demonstrate that the FNO model produces accurate and precise estimations of subsurface soil moisture using surface measurements. Although each humid subtropical data set shows consistent performance with only minor variations in bias and error spread, the arid data set (JNV-Sirsa) not only matches but, in some metrics, outperforms its humid counterparts. This indicates that despite expectations that arid conditions might complicate the estimation process, the approach used handles both climatic extremes effectively.

3.2. Comparison With the Benchmark Algorithms

We compare the FNO model with a diverse set of benchmark algorithms to evaluate its performance and generalizability. This ensures that our proposed methodology stands up to a wide array of alternative algorithms, thereby increasing confidence in its application for estimating subsurface soil moisture from surface measurements.

To benchmark the efficacy of our proposed FNO model, we systematically compared its performance against 15 widely used regression algorithms spanning linear models (Linear Regression, Ridge, Lasso, ElasticNet, Bayesian Ridge), tree-based and ensemble methods (Decision Tree, Random Forest, ExtraTrees, Gradient Boosting, AdaBoost), kernel-based methods (SVR, Kernel Ridge), instance-based learning (KNN), neural network-based approaches (MLP), and probabilistic methods (Gaussian Process). This collection was chosen

10.1029/2026JH001322

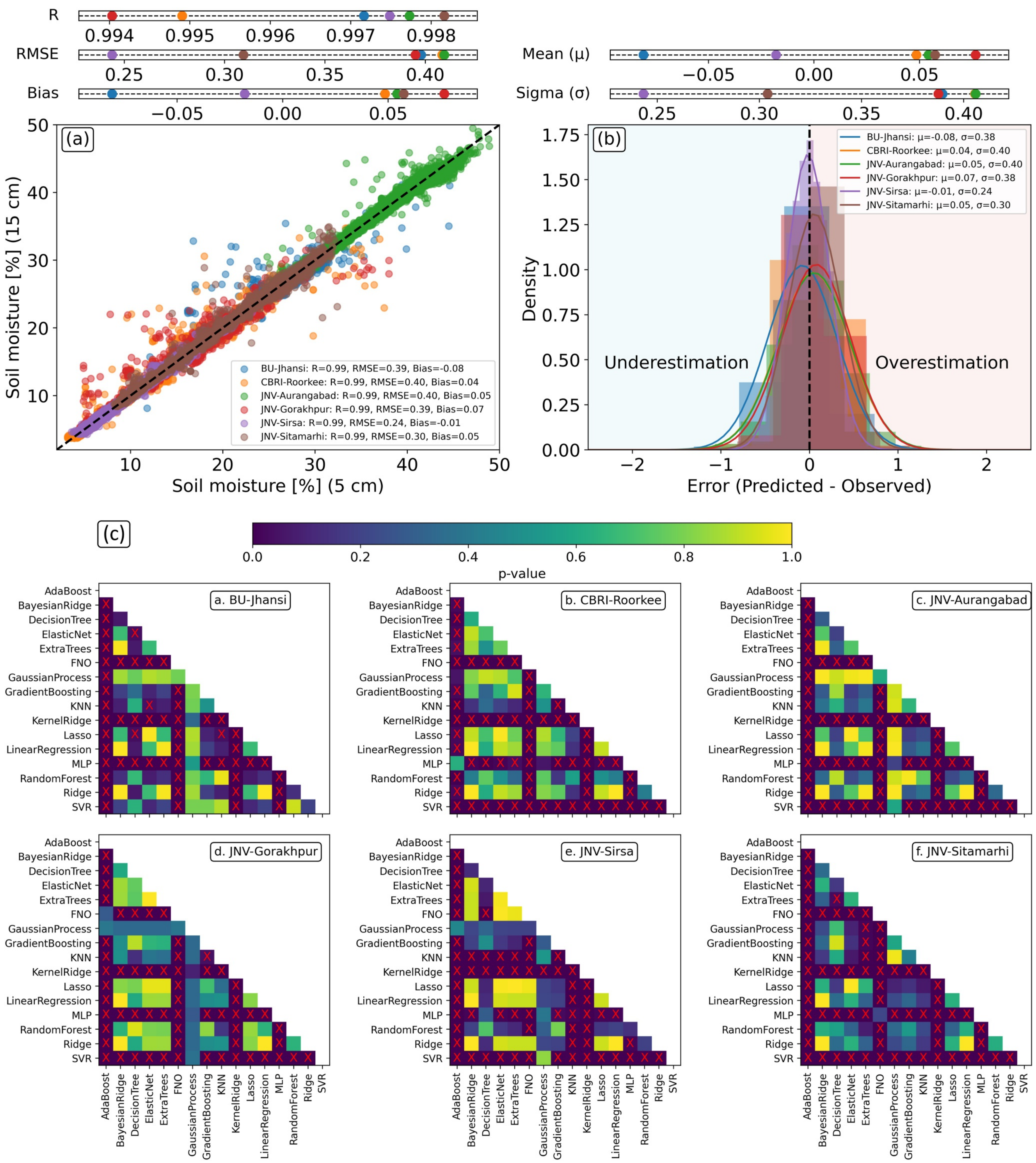

Figure 2. Evaluation of soil moisture predictions across six study sites. (a) Scatter plots of observed versus FNO-predicted soil moisture at 15 cm depth, with each site shown in a different color. Performance metrics (R, RMSE, bias) are included in the legend. The black dashed line denotes perfect agreement. (b) Histograms of prediction errors with Gaussian fits (mean μ, standard deviation σ). Shaded areas indicate underestimation (sky blue) and overestimation (light coral), with the zero-error line marked in black. Insets summarize dynamic ranges of the metrics. (c) Pairwise Welch's t-tests comparing prediction errors between algorithms across data sets. Heatmaps display p-values in the lower triangle; red "X" marks denote statistically significant differences ($p < 0.05$).

because each represents a well-established paradigm in regression analysis, and together, they provide a comprehensive picture of current best practices in environmental modeling and development.

We implemented all benchmark algorithms in the scikit-learn framework and evaluated them using the same input features as the FNO model, namely surface soil moisture at 5 cm depth and the normalized grid coordinate. To ensure a fair comparison, all models are trained and tested using the same 70:30 train-test split for each data set, with a fixed random seed for reproducibility. The benchmark models were evaluated using a consistent baseline configuration rather than extensive dataset-specific hyperparameter tuning, so that the comparison reflected their standard predictive capability under the same data availability conditions. For most models we use the default scikit-learn settings, while applying simple regularization or complexity controls where appropriate. The FNO model is evaluated on the same test indices used for the benchmark models to avoid differences arising from different data splits.

Analyzing the results for individual data sets, the FNO model consistently outperforms the benchmarks in terms of R and RMSE, although several benchmark models exhibit smaller absolute bias values (Table S2 in Supporting Information S1). For example, in the BU-Jhansi data set, the FNO achieves an RMSE of 0.39, a bias of -0.08, and an R value of 0.99, significantly better than the linear models (e.g., Linear Regression and Ridge with RMSE ≈ 1.29 and $R \approx 0.96$) and even many non-linear models. Similarly, for the CBRI-Roorkee data set, FNO attains an RMSE of 0.40, bias of 0.04, and R of 0.99, while models like Gradient Boosting and KNN perform competitively but still do not reach the level of precision exhibited by FNO. In the data sets from JNV-Aurangabad and JNV-Gorakhpur, FNO maintains its advantage with RMSE values of 0.40 and 0.39, respectively, and high correlation coefficients that confirm its capability to capture the underlying soil moisture relationships more accurately than traditional techniques.

Particularly noteworthy is the performance on the JNV-Sirsa data set where FNO records an RMSE of 0.24, near-zero bias (-0.01), and an R of 0.99. In contrast, the conventional linear models and many kernel-based approaches show much higher errors and considerably lower correlation values, highlighting the challenges these methods face in arid conditions. The JNV-Sitamarhi data set further ensures the FNO's superiority, with an RMSE of 0.30 and R of 0.99, outperforming even the best among the benchmark models in a similar manner.

Overall, our comparative analysis demonstrates that the FNO model not only provides consistent and highly accurate estimates of subsurface soil moisture across diverse climatic conditions but also outperforms a wide range of established benchmark algorithms with a rank of 1/16. This robust performance across both humid subtropical and arid regions indicates that FNO is particularly effective at capturing the complex, often nonlinear dynamics between surface and subsurface moisture levels, making it a promising tool for environmental monitoring and water resource management.

To rigorously validate our findings, we applied Welch's t-test to compare the error distributions of the FNO model with those of each benchmark algorithm (Figure 2c). Welch's t-test is particularly appropriate in this context because it does not assume equal variances, a common situation when comparing different predictive models. This statistical test allows us to determine whether the observed differences in mean errors are due to inherent differences in model performance rather than random chance.

For each pairwise comparison between FNO and the other models, the null hypothesis posited that there is no significant difference in the mean error of the models. The alternative hypothesis, on the other hand, suggested that the observed differences are significant. The computed p-values represent the probability of obtaining the observed differences under the null hypothesis. In this study, a p-value lower than 0.05 (significance threshold) is taken as strong evidence against the null hypothesis, indicating that the differences in model errors are statistically meaningful. The significant pairs are marked with a red "X."

We observe that the majority of the comparisons yield p-values well below 0.05. This finding indicates that the differences in error metrics between the FNO model and the benchmark algorithms are statistically significant. Although the FNO model outperforms all 15 benchmark algorithms at the JNV-Sirsa site, the relatively low number of statistically significant paired differences indicates that its performance advantage in this arid region is less pronounced compared to that observed at other sites.

In conclusion, Welch's t-test provides strong evidence of significant differences in mean prediction errors between the FNO and most benchmark algorithms, supporting the observed differences in predictive performance.

3.3. Zero-Shot Transfer Learning

Our transfer learning experiments were designed to test how well the baseline (or zero-shot) model, without any fine-tuning, could generalize when directly applied to new target sites, and how its performance evolves as the volume of target domain data increases. This analysis is crucial because it allows us to understand not only the overall generalization ability of the model but also its performance stability when faced with a progressively larger and more representative target data set.

For example, when the baseline model trained at a source location such as BU-Jhansi is directly applied to a target site, the initial evaluation using a small fraction of the target data (e.g., 10%–25%) typically shows a moderate R_{mean} value accompanied by a higher R_{std}. This variability indicates that with limited target data the predictions are less stable and more influenced by data heterogeneity. As the amount of target data used for evaluation increases say, moving from 25% to 75% or even to 100%, there is an apparent trend of improvement, with the estimated performance metrics becoming more stable, as reflected by their reduced standard deviations (Figure 3a). The R_{mean} values steadily increase while the R_{std} values decrease. This pattern suggests that the increased volume of target data mitigates randomness and enhances the model's ability to capture the target domain's intrinsic moisture dynamics.

A closer look at the results reveals that this trend holds for most sources. In the BU-Jhansi case, the R_{mean} might increase from values in the low 0.924 to nearly 0.949, with R_{std} diminishing from approximately 0.008 to below 0.0009 as the target data volume expands. Similar improvements are observed for other data sets such as CBRI-Roorkee and JNV-Aurangabad, where the reliability of the predictions becomes markedly enhanced with more extensive target data. However, it is notable that for sites like JNV-Sirsa, which represent arid conditions, the initially lower performance gradually improves with an increasing amount of target data. Nonetheless, a marked difference is observed when a model trained in an arid region is applied to a humid subtropical region. While JNV-Sirsa consistently exhibits relatively lower R values, the trend of increasing R with additional target data is consistent across all locations.

Our analysis of the RMSE as a function of increasing target data reveals a slightly mixed trend (Figure 3b). As more target data is incorporated, the RMSE values in most of the scenarios generally decrease, indicating improved predictive accuracy. However, a notable exception is observed for the baseline model trained on JNV-Aurangabad. When applied to other locations, this model consistently yields higher RMSE values compared to models from other sources. This poorer performance appears to be linked to the unique characteristics of the soil moisture distribution at JNV-Aurangabad. In particular, the range of the 5 cm soil moisture at this site, especially for moisture levels below 12.5%, differs from that of the other locations, leading to a mismatch that the model has not learned to handle effectively.

To further investigate this phenomenon, we calculated the pair-wise overlapping area between the 5 cm soil moisture distributions across all sites (Figures 4a and 4b). The overlapping area is a measure of similarity between the distributions, with values closer to 100 indicating high similarity. Our results demonstrate that JNV-Aurangabad has relatively low overlapping areas with other locations, for example, overlaps of approximately 22.4% with JNV-Gorakhpur and 16% with JNV-Sirsa which results in significant RMSE value as compared to other sites. This reduced overlap clearly indicates that the soil moisture regime at JNV-Aurangabad is distinct from those at the other sites, leading to higher RMSE values when the model trained at JNV-Aurangabad is directly transferred without fine-tuning.

We observed that bias behaves similarly to RMSE when plotted against increasing target data (Figure 4c). These findings strongly suggest that while the baseline transfer learning approach is generally effective, improving with increased target data, the performance of the model trained on JNV-Aurangabad is suboptimal in new domains due to the inherent differences in soil moisture distributions. Consequently, the observed higher bias (or RMSE) highlights the need for fine-tuning the transfer learning model to better adapt to the specific characteristics of each target domain. This targeted adaptation is likely essential for achieving more robust and accurate subsurface soil moisture predictions across diverse environmental conditions.

3.4. Transfer Learning With Fine-Tuning

Our extended analysis reveals that even a modest incorporation of target data, specifically around 5%, yields a substantial improvement in the correlation coefficient (Figure 5a). When the baseline model is fine-tuned with

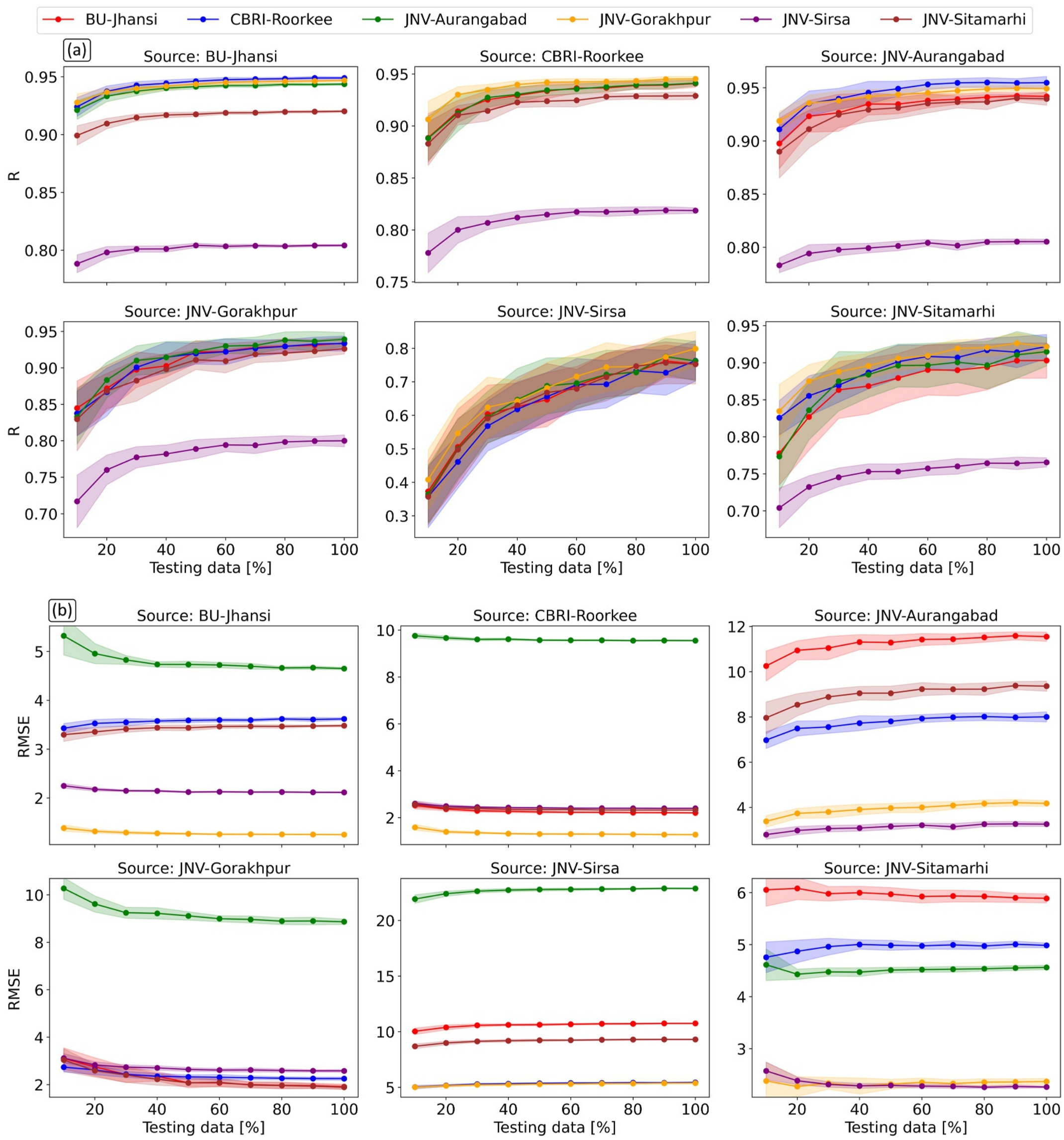

Figure 3. Zero-shot transfer learning performance across sites illustrating (a) *R* and (b) RMSE as a function of the percentage of data used for testing (ranging from 10% to 100%) for various target sites. Each subplot represents a different source model, and the colored curves correspond to target sites with shaded bands indicating ±1 standard deviation. These results assess the transfer learning accuracy achieved without any fine-tuning, highlighting how error metrics vary with increasing test data proportions.

just 5% of the available target domain data, *R* increases markedly compared to its performance without fine-tuning. This early stage improvement indicates that the model is highly responsive to domain-specific cues, allowing it to recalibrate its internal representations to better capture the relationships inherent in the target environment and effectively bridging the gap between the source and target distributions. Beyond this initial boost, further increases in the fine-tuning percentage tend to produce only incremental enhancements in *R*, reflecting a diminishing returns effect. The performance improvements saturate as the model has already used up all the useful information it can get from the target domain. This trend confirms that the critical domain-specific characteristics that drive the increase in *R* are largely captured with just 5% fine-tuning, and additional data contributes marginally to further performance gains.

We repeated the same analysis with RMSE as a function of increasing fine-tuning percentage and observed a trend similar to that observed with *R* (Figure 5b). In the baseline scenario (0% fine-tuning), the model trained on JNV-

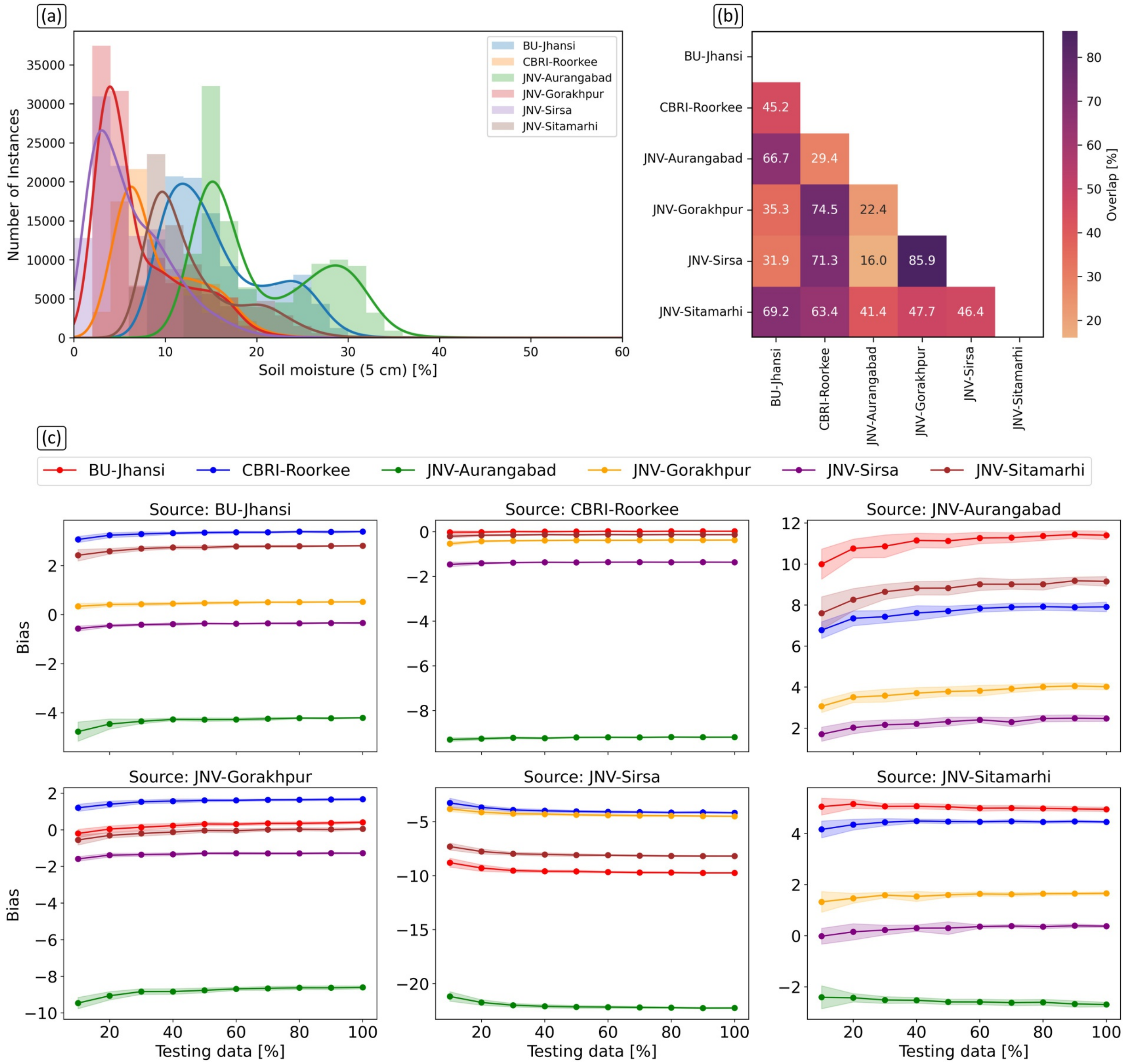

Figure 4. (a) Frequency distribution of surface soil moisture (5 cm depth) across six monitoring stations using histograms and KDE (Kernel Density Estimate) curves. (b) The heatmap shows the percentage overlap in soil moisture distribution between pairs of sites, indicating the spatial similarity in moisture regimes. (c) This figure illustrates the bias as a function of the percentage of data used for testing (ranging from 10% to 100%) for various target sites. Each subplot represents a different source model, and the colored curves correspond to target sites with shaded bands indicating ±1 standard deviation. These results assess the transfer learning accuracy achieved without any fine-tuning, highlighting how error metrics vary with increasing test data proportions.

Aurangabad exhibited significantly higher RMSE values compared to models trained on other sites. This discrepancy clearly highlights the challenges posed by the distinct soil moisture distribution at JNV-Aurangabad, which results in poor generalization when the baseline model is directly applied to new target domains. Remarkably, when we introduced as little as 5% of the target data for fine-tuning, the RMSE for the JNV-Aurangabad model dropped dramatically. The substantial reduction in RMSE brought its performance in line with that of other models, highlighting the model's ability to quickly adapt to domain-specific characteristics with a minimal amount of additional data. This finding confirms that the significant prediction errors observed in the baseline case, due to the mismatch in soil moisture distribution are effectively mitigated by even a small degree of fine-tuning. Using a quantitative saturation criterion, 5% fine-tuning was considered near-optimal because it achieved 53.35% average RMSE reduction relative to the 0% baseline, corresponding to approximately 93.15% of the total improvement obtained at 50% fine-tuning. The additional gain from 5% to 10% fine-tuning was only

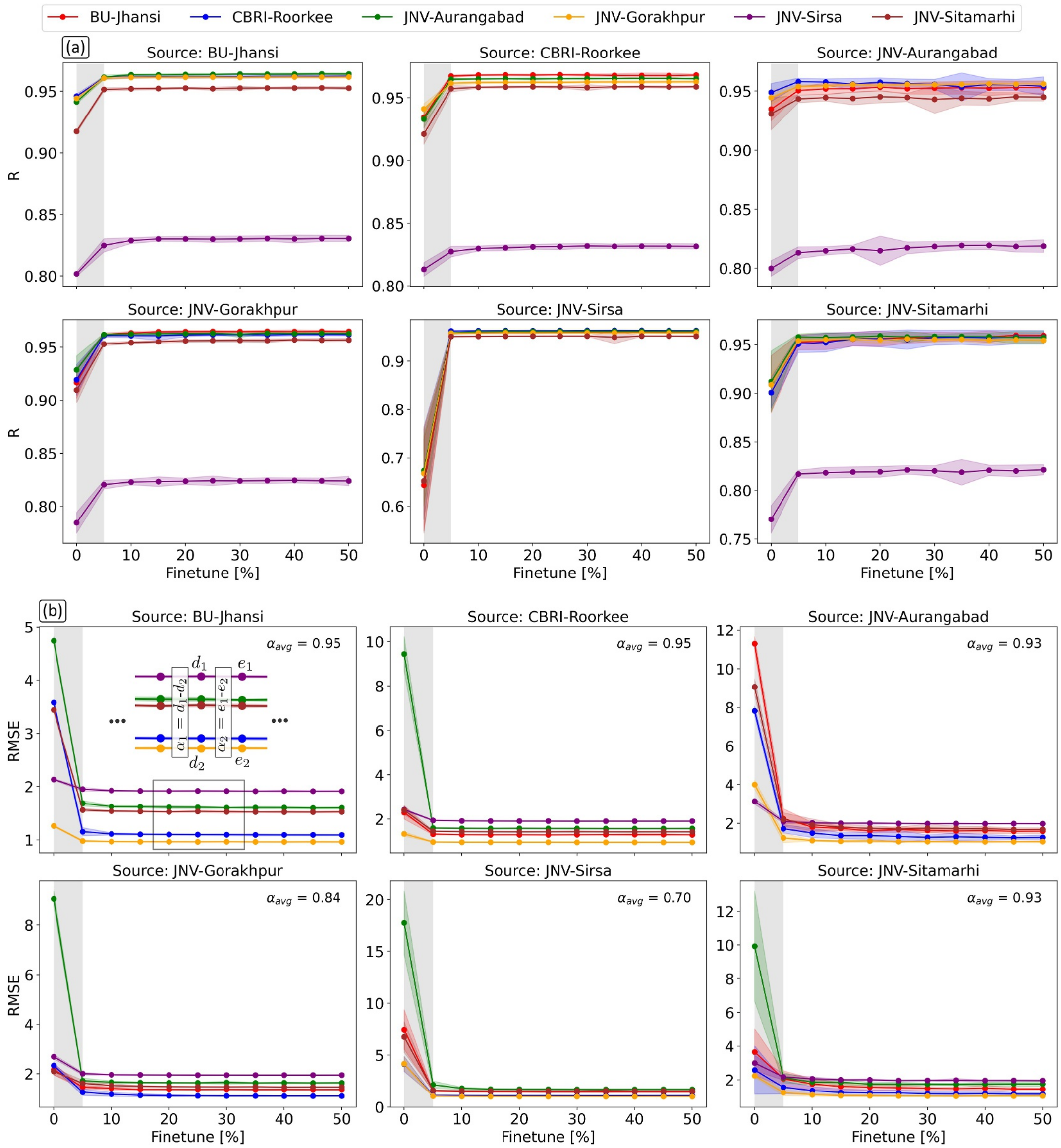

Figure 5. Transfer learning fine-tuning performance across sites. This multi-panel figure evaluates the transfer learning accuracy quantified by (a) R and (b) RMSE as a function of the percentage of data used for fine-tuning. Each subplot corresponds to a different source model, with colored curves representing various target data sets. The curves show the mean R and RMSE values, while the shaded bands denote ±1 standard deviation, illustrating how incremental fine-tuning affects the transfer learning performance without complete retraining.

2.29% points, equivalent to 4.0% of the total 0%–50% improvement, indicating that the incremental benefit beyond 5% fine-tuning was already below the 5% threshold.

To further analyze how close the fine-tuned models perform across different locations once stability is achieved (e.g., after 10% of fine-tuning), we introduce a parameter α_{avg} defined by

$$\alpha_{\text{avg}} = \frac{1}{n_s}\sum_{i=1}^{n_s}(||PM_{\text{worst}}(i)| - |PM_{\text{best}}(i)||) \tag{13}$$

Here, n_s represents the total number of stable fine-tuning levels considered in the post-saturation region. In this study, the stable region is defined as fine-tuning percentages greater than 10%, giving ($n_s = 8$) for the fine-tuning levels 15%, 20%, 25%, 30%, 35%, 40%, 45%, and 50%. $PM_{\text{best}}(i)$ represents the performance metric (such as

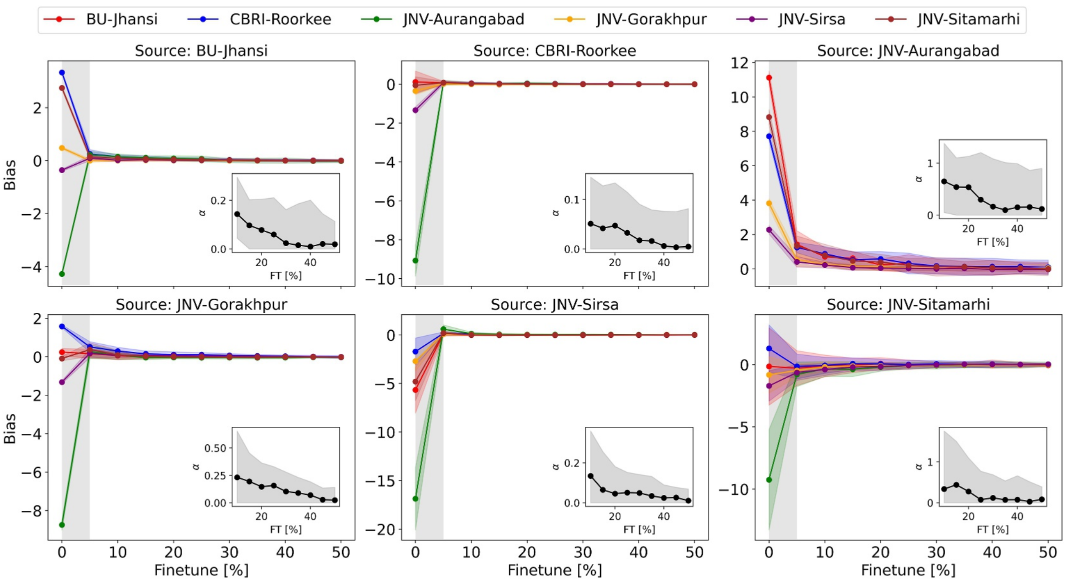

Figure 6. Transfer learning fine-tuning performance across sites. This multi-panel figure evaluates the transfer learning accuracy quantified by bias as a function of the percentage of data used for fine-tuning. Each subplot corresponds to a different source model, with colored curves representing various target data sets. The curves show the mean bias values, while the shaded bands denote ±1 standard deviation, illustrating how incremental fine-tuning affects the transfer learning performance without complete retraining.

RMSE or bias) for the best case, and $PM_{\text{worst}}(i)$ represents the corresponding metric for the worst case at the i^{th} instance. In essence, α_{avg} quantifies the average spread in model performance once it has reached a stable state after the introduction of target data. A lower α_{avg} implies that the discrepancy between the best-case and worst-case performances is minimal, indicating that the model behaves uniformly across different regions.

Our analysis yielded α_{avg} values of 0.95 for BU-Jhansi and CBRI-Roorkee, 0.93 for both JNV-Aurangabad and JNV-Sitamarhi, 0.84 for JNV-Gorakhpur, and 0.70 for JNV-Sirsa. These results demonstrate that, once the models become stable through fine-tuning, the performance spread is relatively narrow, particularly for regions with lower α_{avg} values. Interestingly, the lowest α_{avg} observed at JNV-Sirsa suggests that, in this domain, the fine-tuned model's performance is highly consistent with minimal variability across different metrics. Overall, the low values of α_{avg} across the various sites confirm that the transfer learning approach achieves similar performance characteristics once stabilized.

In our analysis of bias for different percentages of fine-tuning, the results reveal a trend that mirrors those observed for RMSE and R (Figure 6). At the baseline level (0% fine-tuning), the model's bias is noticeably elevated, reflecting a systematic error when the source model is directly applied to the target domain. However, with the inclusion of just 5% of target data for fine-tuning, we observe a dramatic reduction in bias. This rapid improvement indicates that the fine-tuning process is highly effective in correcting the systematic errors of the model by allowing it to quickly assimilate the unique soil moisture characteristics.

Further, our findings reveal that although the model's bias is substantially reduced following an initial 5% fine-tuning, subsequent fine-tuning primarily serves to diminish the uncertainty in bias. The inset plots in each subplot, which illustrate the parameter α (quantifying the average spread in model performance), show a gradual decrease with increasing percentages of fine-tuning. This reduction in α indicates that, while the primary systematic error is corrected early in the fine-tuning process, additional fine-tuning enhances the consistency of the bias correction, thereby reducing bias uncertainty. This dual benefit of the fine-tuning process is particularly significant from a methodological standpoint. The initial fine-tuning rapidly aligns the model with the target domain, substantially lowering the mean bias. However, further fine-tuning is instrumental in homogenizing the bias across various evaluations, leading to more stable and reliable predictions. Such refinement is critical in practical applications where both accuracy and the predictability of error bounds are important.

10.1029/2026JH001322

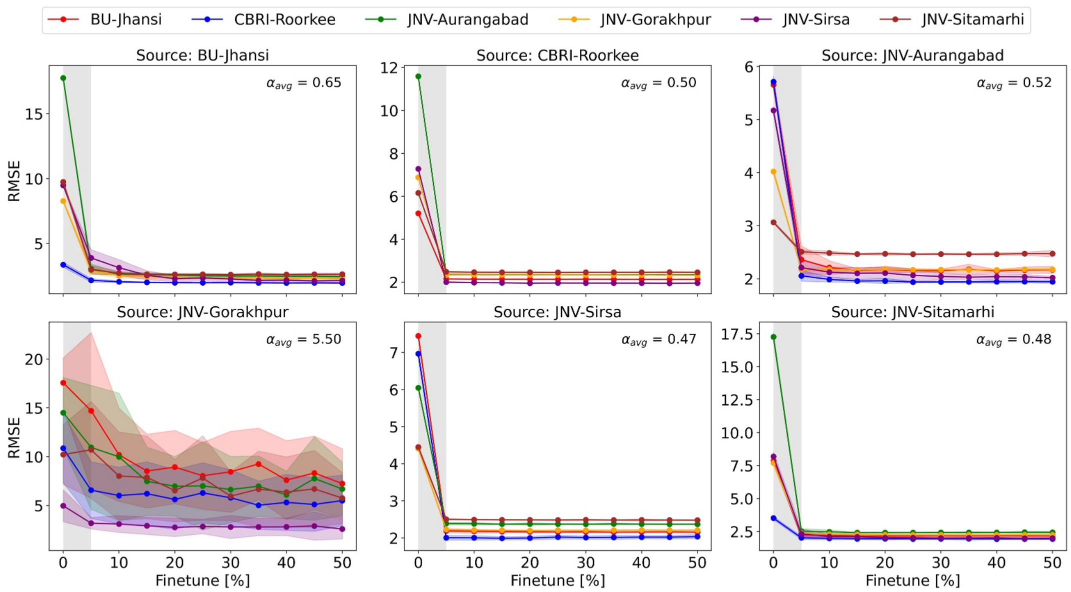

Figure 7. Transfer learning fine-tuning performance across sites at 50 cm. This multi-panel figure evaluates the transfer learning accuracy quantified by RMSE as a function of the percentage of data used for fine-tuning. Each subplot corresponds to a different source model, with colored curves representing various target data sets. The curves show the mean RMSE values, while the shaded bands denote ±1 standard deviation, illustrating how incremental fine-tuning affects the transfer learning performance without complete retraining.

We also performed stochastic robustness analysis at 15 cm depth by computing the median RMSE over 30 different random seeds for each target location. It confirms the model's stability across 30 random seeds, with consistently narrow RMSE ranges across all sites (Figure S2 in Supporting Information S1), indicating reproducible performance and minimal sensitivity to random initialization. Site-wise RMSE distributions and variability metrics are detailed in Section S3 in Supporting Information S1.

3.5. Performance at Deeper Layer

When extending our analysis to a deeper layer, estimating soil moisture at 50 cm from measurements at 5 cm, we observed a clear trend of decreasing mean RMSE as the fine-tuning percentage increases (Figure 7). For instance, when using BU-Jhansi as the source and CBRI-Roorkee as the target, the baseline (0% fine-tuning) yields a mean RMSE of approximately 3.37 with a standard deviation of 0.19. With just 5% of target data incorporated for fine-tuning, the mean RMSE drops substantially to around 2.17 (std ≈ 0.14), and further increments in fine-tuning (up to 50%) gradually improve the performance to about 1.96 (std ≈ 0.03). This pattern suggests that even minimal fine-tuning is highly effective in aligning the model with the target domain's deeper soil moisture characteristics, and additional fine-tuning further consolidates these gains with increasingly consistent predictions similar to what we have observed in the 15 cm soil depth.

A more dramatic example is observed when BU-Jhansi is used as the source and JNV-Aurangabad as the target. Here, the baseline RMSE is extremely high, around 17.76, reflecting the significant mismatch in soil moisture distributions between the source at 5 cm and the deep profile at 50 cm in JNV-Aurangabad. However, once 5% of the target data is introduced, the mean RMSE significantly drops to approximately 3.05, followed by a steady decline to about 2.45 by the time 50% fine-tuning is applied. This significant reduction indicates that even a small amount of domain-specific data can correct for the large discrepancies inherent in the baseline model due to different moisture ranges, especially in sites like JNV-Aurangabad where the moisture profile deviates considerably from other locations.

Similar trends are noted for other source-target pairs. For example, with BU-Jhansi as the source and JNV-Gorakhpur as the target, the baseline RMSE is around 8.28, which falls to roughly 2.92 with 5% fine-tuning, and further declines to about 2.32 at 50% fine-tuning. For BU-Jhansi to JNV-Sirsa and JNV-Sitamarhi, the

10.1029/2026JH001322

baseline RMSE values of approximately 9.48 and 9.75, respectively, also exhibit marked improvements with fine-tuning, stabilizing around 2.15 and 2.64 at higher fine-tuning percentages. Across these various pairs, the accompanying standard deviations decrease as fine-tuning increases, reflecting not only improved accuracy but also enhanced consistency in model predictions.

When using JNV-Gorakhpur as the source, the performance transferability to various target domains reveals significant baseline discrepancies that are effectively mitigated through fine-tuning. For example, when predicting soil moisture at 50 cm for BU-Jhansi, the baseline RMSE is exceptionally high at approximately 17.58, reflecting the marked distributional differences between JNV-Gorakhpur and BU-Jhansi. However, incorporating just 5% of target data for fine-tuning reduces the RMSE substantially to around 14.70, and further increments-up to 50% progressively lower the RMSE to about 7.24. This substantial reduction is accompanied by a decrease in the standard deviation, indicating not only an improvement in accuracy but also a stabilization of model performance. A similar trend is observed with other target data sets. These patterns indicate that the inherent differences between the moisture distributions in JNV-Gorakhpur and the target sites cause large initial errors, which can be rapidly corrected with even minimal fine-tuning.

To compare the consistency of the predictions to estimate soil moisture at 15 and 50 cm from measurements at 5 cm, we used α_{avg}, which we define to quantify the average spread in the performance of the model once stability is achieved after fine-tuning. For the 15 cm estimates, α_{avg} values ranged from 0.70 to 0.95 across the six locations, indicating a relatively narrow spread in performance once the model became stable. Conversely, for the 50 cm estimates, most locations exhibited even lower α_{avg} values (ranging between 0.47 and 0.65), suggesting that fine-tuning leads to a more uniform performance across evaluation runs when estimating deeper soil moisture layers. However, an exception is noted for JNV-Gorakhpur, where the α_{avg} for 50 cm estimation was anomalously high at 5.50. This elevated value implies a considerably higher variability in the model's performance for this specific site when predicting soil moisture at a deeper layer. The contrast between JNV-Gorakhpur and the other locations highlights potential challenges related to site-specific factors that may cause deeper soil moisture dynamics to be inherently more complex or sensitive to the fine-tuning process. Overall, while the consistent, low α_{avg} values at most sites validate the reliability of the fine-tuning approach for both 15 and 50 cm estimations, the outlier observed at JNV-Gorakhpur points to the need for further investigation into the underlying causes of variability in deeper soil moisture predictions at this location.

4. Conclusion

In this study, we developed a transferable FNO framework for estimating subsurface soil moisture from surface measurements, leveraging high-resolution data from six climatically diverse sites. The model demonstrated robust performance across heterogeneous conditions. FNO outperforms the benchmark algorithms at both 15 and 50 cm depths. Notably, fine-tuning with as little as 5% of target data effectively adapted the model, correcting systematic errors and reducing uncertainty. Stochastic robustness analysis confirmed the robustness and reproducibility of predictions, while comparative assessments using the α_{avg} parameter highlighted the improved consistency of fine-tuned estimates, even at deeper depths.

Our approach addresses the critical challenge of data scarcity, offering a scalable and efficient solution for subsurface moisture monitoring. It holds strong potential for climate-resilient agriculture and water resource forecasting in data-sparse regions. Future work will explore expanded domain adaptation strategies and broader spatiotemporal generalization.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

In situ data can be downloaded from Zenodo (Singh et al., 2025). The code developed for this study can be downloaded from Zenodo (Singh et al., 2026).

Acknowledgments

The authors gratefully acknowledge the University of Leeds and IISER Bhopal for providing institutional support. We gratefully acknowledge the GRACE Network Project and National Geospatial Policy (NGP), Department of Science and Technology (DST) for funding the deployment of these stations. V. M. Tiwari acknowledges support from the ANRF J. C. Bose Fellowship. We also thank the Directors of CSIR-NGRI and CSIR-NEIST for their support. Krishan Vishwakarma and Prashant Kumar Gupta for their invaluable assistance in installing soil moisture sensors and dedicated fieldwork. This work is permitted to be published with reference number (Ref. No. NGRI/Lib/2026/Pub-92).

References

Alipour, A., Daneshvar, F., Mani, S., Seroka, G., Shi, L., Yang, Z., et al. (2026). Neural operators for continuous bias correction of water level forecast guidance. *Journal of Geophysical Research: Machine Learning and Computation*, *3*(1), e2025JH000853. https://doi.org/10.1029/2025jh000853

Azizzadenesheli, K., Kovachki, N., Li, Z., Liu-Schiaffini, M., Kossaifi, J., & Anandkumar, A. (2024). Neural operators for accelerating scientific simulations and design. *Nature Reviews Physics*, *6*(5), 1–9. https://doi.org/10.1038/s42254-024-00712-5

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 1–12. https://doi.org/10.1038/sdata.2018.214

Blume, T., Zehe, E., & Bronstert, A. (2009). Use of soil moisture dynamics and patterns at different spatio-temporal scales for the investigation of subsurface flow processes. *Hydrology and Earth System Sciences*, *13*(7), 1215–1233. https://doi.org/10.5194/hess-13-1215-2009

Hummel, J. W., Sudduth, K. A., & Hollinger, S. E. (2001). Soil moisture and organic matter prediction of surface and subsurface soils using an NIR soil sensor. *Computers and Electronics in Agriculture*, *32*(2), 149–165. https://doi.org/10.1016/s0168-1699(01)00163-6

Hutchinson, M. L., Antono, E., Gibbons, B. M., Paradiso, S., Ling, J., & Meredig, B. (2017). Overcoming data scarcity with transfer learning. *arXiv preprint arXiv:1711.05099*.

Kumar, S. V., Reichle, R. H., Koster, R. D., Crow, W. T., & Peters-Lidard, C. D. (2009). Role of subsurface physics in the assimilation of surface soil moisture observations. *Journal of Hydrometeorology*, *10*(6), 1534–1547. https://doi.org/10.1175/2009jhm1134.1

Li, X., Jiang, Y., Zhuang, Y., Hu, X., & Zhao, B. (2026). A physics-informed multi-scale Fourier neural operator framework for snow avalanche dynamics simulation. *Journal of Hydrology*, *675*, 135573. https://doi.org/10.1016/j.jhydrol.2026.135573

Liu, N., Jafarzadeh, S., & Yu, Y. (2023). Domain agnostic Fourier neural operators. *Advances in Neural Information Processing Systems*, *36*.

McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., & Entekhabi, D. (2017). The global distribution and dynamics of surface soil moisture. *Nature Geoscience*, *10*(2), 100–104. https://doi.org/10.1038/ngeo2868

Nunez, A., Misra, S., & Falola, Y. (2025). Transfer learning for geological carbon storage forecasting using neural operator. *Advances in Water Resources*, *199*, 104948. https://doi.org/10.1016/j.advwatres.2025.104948

Reyad, M., Sarhan, A. M., & Arafa, M. (2023). A modified Adam algorithm for deep neural network optimization. *Neural Computing & Applications*, *35*(23), 17095–17112. https://doi.org/10.1007/s00521-023-08568-z

Seyfried, M., Grant, L., Du, E., & Humes, K. (2005). Dielectric loss and calibration of the hydra probe soil water sensor. *Vadose Zone Journal*, *4*(4), 1070–1079. https://doi.org/10.2136/vzj2004.0148

Singh, A., & Gaurav, K. (2024). PIML-SM: Physics-informed machine learning to estimate surface soil moisture from multi-sensor satellite images by leveraging swarm intelligence. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1–13. https://doi.org/10.1109/TGRS.2024.3502618

Singh, A., Niranjannaik, M., & Gaurav, K. (2025). Overcoming data scarcity: A transfer learning framework with fine-tuned neural networks and multi-sensor satellite image fusion for soil moisture estimation. *Engineering Applications of Artificial Intelligence*, *159*, 111636. https://doi.org/10.1016/j.engappai.2025.111636

Singh, A., Ravikumar, M., Singh, V., Shivprasad, M., Tiwari, V. M., & Gaurav, K. (2025). Transferable neural operators to estimate sub-surface soil moisture datasets. *Zenodo*. https://doi.org/10.5281/zenodo.16106763

Singh, A., Ravikumar, M., Singh, V., Shivprasad, M., Tiwari, V. M., & Gaurav, K. (2026). Transferable neural operators to estimate sub-surface soil moisture code. *Zenodo*. https://doi.org/10.5281/zenodo.18867777

Singh, A., Singh, V., & Gaurav, K. (2025a). Leveraging neural operator and sliding window technique for enhanced subsurface soil moisture imputation under diverse precipitation scenarios. *Journal of Geophysical Research: Machine Learning and Computation*, *2*(3), e2025JH000730. https://doi.org/10.1029/2025jh000730

Singh, A., Singh, V., & Gaurav, K. (2025b). Physics-aware probabilistic modeling of subsurface soil moisture using diffusion processes across different climate settings. *Geophysical Research Letters*, *52*(20), e2025GL118607. https://doi.org/10.1029/2025gl118607

Singh, A., Singh, V., & Gaurav, K. (2026). Weak physics-guided multi-agent learning for surface to subsurface moisture estimation across diverse climate and soil conditions. *Journal of Geophysical Research: Machine Learning and Computation*, *3*(1), e2025JH001039. https://doi.org/10.1029/2025jh001039

Theodoris, C. V., Xiao, L., Chopra, A., Chaffin, M. D., Al Sayed, Z. R., Hill, M. C., et al. (2023). Transfer learning enables predictions in network biology. *Nature*, *618*(7965), 616–624. https://doi.org/10.1038/s41586-023-06139-9

Tian, J., Zhang, B., Wang, Y., Pei, L., & He, C. (2023). Vertical partition patterns of infiltration within soil profile and its control factors at large-scale arid mountainous areas. *Journal of Hydrology*, *618*, 129190. https://doi.org/10.1016/j.jhydrol.2023.129190

Wang, S., Cao, J., & Philip, S. Y. (2020). Deep learning for spatio-temporal data mining: A survey. *IEEE Transactions on Knowledge and Data Engineering*, *34*(8), 3681–3700.

Wang, Y., Zhang, H., Lai, C., & Hu, X. (2024). Transfer learning Fourier neural operator for solving parametric frequency-domain wave equations. *IEEE Transactions on Geoscience and Remote Sensing*, *62*, 1–11. https://doi.org/10.1109/TGRS.2024.3440199

Xi, X., Zhuang, Q., & Liu, X. (2025). A hybrid physics-guided deep learning modeling framework for predicting surface soil moisture. *Journal of Geophysical Research: Machine Learning and Computation*, *2*(3), e2025JH000682. https://doi.org/10.1029/2025jh000682

Yang, J., Yang, Q., Hu, F., Shao, J., & Wang, G. (2024). A climate-adaptive transfer learning framework for improving soil moisture estimation in the Qinghai-Tibet plateau. *Journal of Hydrology*, *630*, 130717. https://doi.org/10.1016/j.jhydrol.2024.130717

Zhang, Z. (2018). Improved Adam optimizer for deep neural networks. In *2018 IEEE/ACM 26th international symposium on quality of service (IWQOS)* (pp. 1–2).