JGR | Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2025JH000730

Leveraging Neural Operator and Sliding Window Technique for Enhanced Subsurface Soil Moisture Imputation Under Diverse Precipitation Scenarios

Abhilash Singh[1,2], **Vidhi Singh**[2], **and Kumar Gaurav**[2]

[1]School of Mathematics, Faculty of Engineering and Physical Sciences, University of Leeds, Leeds, UK, [2]Fluvial Geomorphology and Remote Sensing Laboratory, Department of Earth and Environmental Sciences, Indian Institute of Science Education and Research Bhopal, Bhopal, India

Key Points:

- Developed a novel Fourier Neural Operator (FNO) to enhance subsurface soil moisture imputation by employing a sliding window concept that seamlessly integrates rainfall, soil temperature, and normalised time data
- The FNO model achieves an average correlation coefficient of approximately 0.99 and an RMSE of around 1.2 which outperforms the traditional imputation methods
- Introducing temporal lags reduces imputation error by up to 15%, improving responsiveness to rapid rainfall-induced changes

Correspondence to:

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

Citation:

Singh, A., Singh, V., & Gaurav, K. (2025). 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*, e2025JH000730. https://doi.org/10.1029/2025JH000730

Received 10 APR 2025
Accepted 21 AUG 2025

Author Contributions:

Conceptualization: Abhilash Singh, Kumar Gaurav
Data curation: Vidhi Singh
Formal analysis: Vidhi Singh
Funding acquisition: Kumar Gaurav
Investigation: Abhilash Singh, Vidhi Singh, Kumar Gaurav
Methodology: Abhilash Singh, Vidhi Singh, Kumar Gaurav
Software: Abhilash Singh, Vidhi Singh
Supervision: Kumar Gaurav
Validation: Abhilash Singh

Abstract Environmental monitoring and decision-making are sometimes hampered by missing sensor data, particularly in soil moisture (SM) records that underpin hydrological modeling, agricultural management, and climate studies. We present a novel imputation framework based on Fourier neural operators (FNO) that robustly reconstructs missing SM values by learning global spatiotemporal dependencies directly in the frequency domain. Our approach segments high-frequency hydro-meteorological time series into overlapping windows, incorporating rainfall, soil temperature, and normalized time coordinates and leverages spectral convolution layers with a sliding window strategy to capture both short- and long-term dynamics. Through extensive experiments using multidepth SM measurements (10, 20, 30, and 40 cm), we systematically assess the impact of varying missing data ratios (from 5% to 50%) and temporal lag configurations. The FNO model demonstrates statistically better performance as compared to the traditional statistical and machine learning imputation methods. The FNO shows high correlation coefficients and low root mean square errors even under challenging rain and no-rain conditions. Although the imputation accuracy of all models decreases during rain events, we observed that incorporating a temporal delay marginally improves performance by reducing the imputation error by up to 15%. These results establish the FNO framework as a paradigm shift in environmental data imputation, demonstrating unparalleled accuracy by harnessing global spatiotemporal dependencies to effectively overcome data sparsity even under the most challenging conditions.

Plain Language Summary This study introduces a new, innovative method to fill in missing data from soil moisture (SM) sensors, which are vital for understanding water movement in soils, managing agriculture, and studying climate patterns. Traditional methods sometimes fall short because they struggle to capture the complex patterns in natural systems. In contrast, the new method uses Fourier neural operator, which looks at data in a different way by analyzing the data's overall patterns rather than just local details. The method works by splitting the sensor data into smaller, overlapping pieces and then combining additional information like temperature and rainfall. Tests show that this approach not only works better than older techniques but also adapts well to different conditions, even when a large portion of the data is missing. By slightly adjusting for the natural delay in how SM reacts to rain, the method further improves its accuracy.

1. Introduction

Time series records to soil moisture (SM) measurement are essential for applications such as hydrological modeling, agricultural management, and climate studies (Li et al., 2022; Singh & Gaurav, 2024). In particular, subsurface SM records are especially prone to frequent and prolonged gaps due to the challenges of underground sensor maintenance and exposure to harsh conditions, which can severely undermine the reliability of subsequent analyses if not properly addressed (Chavoshi et al., 2024). Traditional imputation methods, ranging from statistical interpolation to machine learning techniques, struggle to capture the complex spatial-temporal dependencies present in these data sets (Mitra et al., 2023).

Neural operators have emerged as a transformative framework for learning mappings between infinite-dimensional function spaces (Azizzadenesheli et al., 2024). In particular, the Fourier neural operator (FNO) has demonstrated remarkable efficacy in approximating solutions to complex partial differential equations. The FNO leverages the Fourier transform to learn global representations of data, effectively capturing long-range dependencies and dynamic patterns. Unlike conventional neural network architectures that rely on local

Visualization: Abhilash Singh, Kumar Gaurav
Writing – original draft: Abhilash Singh, Vidhi Singh, Kumar Gaurav
Writing – review & editing: Abhilash Singh, Kumar Gaurav

convolutions or recurrent mechanisms, FNO operates in the frequency domain, enabling it to directly model the underlying physical processes and multiscale phenomena inherent in the environmental data (Lu et al., 2021). This ability to learn global operators from data, coupled with its inherent resolution invariance and efficient parameterization, positions FNO as a uniquely powerful tool for imputation tasks where traditional methods often fall short (Cao et al., 2024).

This study focuses on SM measurements recorded at various depths (10, 20, 30, and 40 cm) to assess the efficacy of the FNO in filling data gaps. This multidepth analysis allows to capture the variability in SM dynamics at different layers. SM at shallow depths typically respond more rapidly to rainfall events, while deeper layers exhibit slower and more integrated responses. We analyze a wide range of experimental scenarios, including different percentages of missing data (ranging from 5% to 50%), multiple temporal lag settings (from no lag to a lag of 5 time steps) and a range of precipitation contexts specifically, periods of rainfall, dry periods, and overall aggregated conditions. Varying the missing data percentages tests the resilience of FNO under different levels of data sparsity, while exploring temporal lags allows us to understand how incorporating short-term historical context influences imputation accuracy. These settings are critical because they simulate the realistic data quality issues and temporal dynamics that practitioners encounter in environmental monitoring. In addition, we evaluate the performance of the FNO under different rain magnitude events to determine its sensitivity to abrupt changes in SM driven by precipitation. Event-based scenarios are particularly important because discrete rainfall events can cause rapid and significant changes in SM, especially in the shallow layers. By analyzing how the FNO performs during high-intensity rain events compared to dry periods, we can assess its ability to capture the underlying physics of moisture redistribution and its robustness under dynamic, nonstationary conditions.

This comprehensive approach not only provides insights into the overall imputation capability of the FNO but also sheds light on its performance under varying environmental conditions. By addressing these diverse experimental scenarios, this study addresses several key research questions.

1. How effective is the FNO in accurately imputing missing SM data across various sensor depths?
2. What impact do different configurations of temporal delays have on the accuracy of imputation under varying percentages of missing data?
3. How robust is the FNO in distinguishing between rain and no-rain scenarios, and how does it perform during events of different rainfall intensities?

By exploring these questions through extensive experimental analysis, we aim to provide a comprehensive evaluation of FNO as a novel tool for the imputation of environmental data, ultimately contributing to improved data reliability in environmental monitoring and decision-making processes.

2. Related Work

The field of missing data imputation in environmental sciences has evolved considerably, driven in part by the increasing availability of high-frequency temporal data and the growing need for accurate subsurface data imputation (Y. Zhang & Thorburn, 2022). Early efforts predominantly relied on traditional statistical techniques such as mean, median, and mode imputation, as well as basic regression methods, which often fell short in capturing the intricate spatial and temporal dynamics inherent in environmental processes (Pastorini et al., 2024; Richman et al., 2009). More recent approaches have leveraged advanced machine learning models, including Bayesian methods, K-nearest neighbors (KNN), and ensemble techniques such as random forest (RF) to better approximate these complex dependencies in high temporal time series data (Boomgard-Zagrodnik & Brown, 2022). However, the inherent limitations of these methods, particularly in scenarios with high data sparsity or abrupt changes induced by rainfall events, have highlighted the need for more sophisticated techniques.

Several studies have been conducted on data imputation for rainfall, groundwater, and other critical environmental variables (Boujoudar et al., 2024; L. Chen et al., 2019; Chivers et al., 2020; Pinthong et al., 2024). Recently, Chivers et al. (2020) introduced a two-step machine learning approach for imputing subhourly precipitation data. Their method employs ensemble decision trees (XGBoost and RF) for classification, followed by neural networks for regression, and achieves correlation coefficients between 0.56 and 0.95, with root mean square error (RMSE) values ranging from 0.064 to 0.389 mm. This combined approach demonstrates

considerable promise, but it appears that some rain events, ranging from light to intense, may not be fully captured. Furthermore, there is an anticipation that performance might face additional challenges when transitioning to higher temporal resolutions, such as 10-min intervals, primarily due to increased variability and rapid fluctuations in precipitation that are more difficult to model accurately at finer time scales (Marcy et al., 2022).

Similarly, groundwater data imputation has been widely studied to enhance the modeling and forecasting accuracy (Evans et al., 2020; Ramirez et al., 2022; Z. Zhang et al., 2017). A deep learning framework combining the Bayesian linear regression with long short-term memory and convolutional neural network models achieved high accuracy in regional groundwater level prediction (E. Chen et al., 2024). However, challenges remain, such as the need for long-term data sets and a larger set of predictor variables for reliable modeling. The study also highlighted the potential of extending this framework to incorporate advanced methods for model uncertainty quantification. Similarly, an artificial gap introduction approach used in assessing automated gap imputation in regional groundwater data sets compared missForest, imputePCA, and linear interpolation techniques (Bikše et al., 2023). Although imputation methods performed well for varying gap patterns, none of them accurately filled the missing data during extreme hydrological conditions. In such a condition, the imputation errors increases significantly. These studies highlight the need for a robust model that efficiently processes high-temporal resolution data (less than 30 min), employs imputation techniques adaptable to varying degrees of data sparsity and environmental conditions, and delivers improved performance, especially under extreme circumstances.

To the best of our knowledge, studies on SM imputation at high temporal resolutions that incorporates the data from surface and subsurface measurements while explicitly linking SM responses to rain events do not exist. This study addresses the key challenges of previous studies by developing an FNO framework. It transforms data into the frequency domain to capture the complex spatiotemporal dependencies and also integrates temporal lag adjustments to better align with the delayed response of subsurface moisture. The approach is specifically designed to reconstruct SM dynamics across multiple sensor depths under both rain and no-rain conditions. This provides a more accurate and reliable imputation solution for environmental monitoring and hydrological forecasting.

3. Data Sets

This study incorporates high-frequency hydro-meteorological measurements collected from an advanced monitoring station installed at the Indian Institute of Science Education and Research Bhopal campus (23.287°N, 77.278°E) in central India. The study site is characterized by black clayey soils with high moisture retention properties. The observatory integrates an automatic weather station (AWS) and SM sensors to capture key hydrological and atmospheric variables (Figure 1). We deploy four calibrated frequency domain reflectometry (FDR) sensors at depths of 10, 20, 30, and 40 cm in a soil pit near the weather station. The sensors are calibrated before deployment using the gravimetric method, achieving a root mean square error (RMSE) of approximately 0.07 m^3/m^3. Additionally, they are cross-calibrated to ensure consistency and minimize potential sensor bias. They continuously record volumetric water content (%) and soil temperature (ST) (°C) at each depth. They are designed to measure SM within a range of 0%–100%, with a ±2% precision for 0%–50% moisture content, and ±3% for 50%–100% SM values. Additionally, each sensor is equipped with temperature probes that measure ST across a broad range from −40°C to +60°C, ensuring accurate monitoring of both moisture and thermal conditions within the soil. Rainfall is concurrently measured by the AWS using a tipping bucket rain gauge with a maximum collection rate of 100 $mmhour^{-1}$ and a resolution of 0.20 mm. It ensures accurate temporal alignment between the rainfall data and soil variables. All instruments in AWS and FDR sensors are programmed to log data at 10-min intervals. A year-long time series data are used in this study, spanning March 2024 to February 2025 (Figure 1). The use of high-frequency measurements plays an important role in accurately capturing the rapid infiltration and redistribution dynamics within the 10–40 cm soil depth range, especially following rainfall events. Unlike coarser temporal resolutions (e.g., hourly), which tend to smooth out short-term fluctuations, the chosen frequency enables the model to learn and account for fast-changing subsurface processes with greater fidelity. This makes the data set particularly well-suited for evaluating imputation performance under realistic, transient conditions. Furthermore, the 1-year continuous time series provides a comprehensive representation of seasonal variability and diverse hydrological conditions, including dry spells and intense rainfall events. Together, the high temporal resolution and extended observation period offer a robust foundation for validating the proposed algorithm.

10.1029/2025JH000730

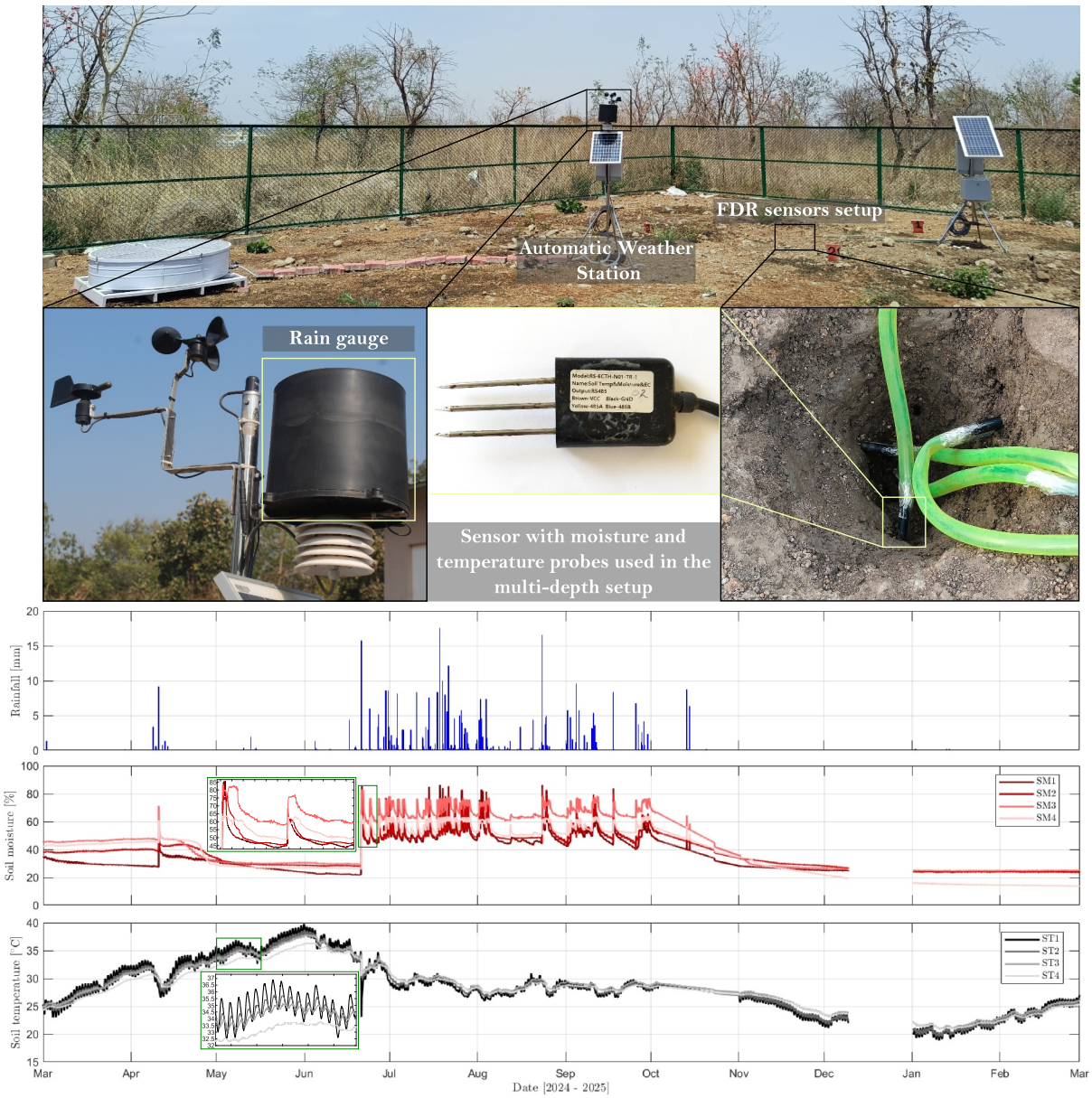

Figure 1. The observatory setup at Indian Institute of Science Education and Research Bhopal features an automatic weather station (AWS) and a frequency domain reflectometry (FDR) sensor system for continuous monitoring of meteorological conditions and soil properties. The top panel showcases the observatory along with key instrumentation, including a tipping bucket rain gauge within the AWS for measuring rainfall, an FDR sensor that records soil moisture (SM) and temperature. Field photograph shows the sensor configuration in a soil pit, where probes are installed at multiple depths (10, 20, 30, and 40 cm). The bottom panel presents time series data from March 2024 to February 2025, rainfall (blue bars), SM (shades of red), and soil temperature (shades of gray) at different depths (SM1, ST1: 10 cm; SM2, ST2: 20 cm; SM3, ST3: 30 cm; SM4, ST4: 40 cm). A highlighted period (green box) marks distinct variations in SM and temperature trends, effectively illustrating depth-dependent responses to environmental conditions.

Although all SM sensors were calibrated prior to deployment using the gravimetric method and cross-calibrated to minimize sensor bias, long-term deployments can still be susceptible to gradual drift or occasional anomalies due to environmental or hardware-related factors. However, no evidence of sensor drift was observed during the 1-year monitoring period. To ensure data quality, we manually inspected the time series across all depths and masked any clearly inconsistent or erroneous values prior to analysis. Although drift correction was not necessary in this study, we acknowledge that in longer-term or large-scale deployments, sensor drift can affect data integrity and model performance. In such cases, incorporating periodic recalibration, drift-aware preprocessing, or sensor fusion strategies would be essential. This represents an important avenue for future enhancements of the framework.

This study considers three distinct scenarios: overall instances, instances with rainfall, and instances without rainfall. The data set comprises a total of 47,808 observations, forming the basis of the first scenario. Among these, 1,266 instances (2.65% of the total data) recorded rainfall, constituting the second scenario. The remaining 46,542 observations (97.35%) represent instances where no rainfall occurred, framing the third scenario.

4. Methodology

Algorithm 1. FNO-Based Imputation Algorithm.

Input: Time series data $X \in \mathbb{R}^{B\times T\times 3}$ and ground truth soil moisture $Y \in \mathbb{R}^{B\times T}$.
Parameters: Fourier modes m, network width w, number of spectral layers L, sliding window size T_w, stride s, missing mask M, learning rate η, and epochs E.
Preprocessing:
  Compute normalized time grid grid $\in [0,1]^{B\times T\times 1}$.
  Concatenate: $\tilde{X} \leftarrow \text{concat}(X, \text{grid})$.
  Lift features: $X_0 \leftarrow \text{fc0}(\tilde{X})$, where $X_0 \in \mathbb{R}^{B\times T\times w}$.
for $l = 1$ **to** L **do**
  Permute: $x \leftarrow \text{permute}(X_{l-1})$ ▷ Now $x \in \mathbb{R}^{B\times w\times T}$
  Fourier transform: $\hat{x} \leftarrow \mathcal{F}(x)$
  Retain first m modes: $\hat{x}_m \leftarrow \hat{x}[:, :, 1:m]$
  Spectral convolution: $\hat{y} \leftarrow \hat{x}_m \odot W^{(l)}$ ▷ Element-wise multiplication with learned complex weights
  Inverse Fourier transform: $y \leftarrow \mathcal{F}^{-1}(\hat{y})$
  Residual connection: $z \leftarrow \text{conv1x1}(x)$
  Nonlinear activation: $x_l \leftarrow \text{ReLU}(y + z)$
  Permute back: $X_l \leftarrow \text{permute}^{-1}(x_l)$
end for
Output Projection:
$H \leftarrow \text{fc1}(X_L)$
$Y_{\text{pred}} \leftarrow \text{fc2}(\text{ReLU}(H))$
Loss Computation:
$\mathcal{L} \leftarrow \frac{\sum_{i,j}(Y_{\text{pred}}(i,j) - Y(i,j))^2 M(i,j)}{\sum_{i,j} M(i,j) + \epsilon}$
Optimization: Update model parameters via backpropagation using Adam with learning rate η for E epochs.
Imputation:
  Extract sliding windows of length T_w with stride s from the time series.
  Apply the trained FNO to each window.
  Aggregate overlapping window predictions (e.g., via averaging) to reconstruct the full-time series.
  Impute missing entries in Y using the reconstructed series.
Output: Imputed soil moisture time series.

We have measured the SM and ST at depths (below ground) at 10, 20, 30, and 40 cm, and rainfall. The Date_Time column is standardized to ensure consistent temporal alignment. To simulate realistic scenarios where sensor malfunctions or communication issues lead to missing data, we introduce missing values using a random mask $M \in \{0,1\}^T$ such that $M(i) = 0$ indicates a missing observation at time step i. Missing data percentages range from 5% to 50%, increasing uniformly by 5%, to evaluate the robustness of our imputation framework.

Table 1
Summary of Variables and Parameters Used in the FNO-Based Soil Moisture Imputation Algorithm

Variable	Description	Value/Default
B	Batch size	Variable
T	Number of time steps (window length)	128
m	Number of Fourier modes retained	16
w	Network width	64
L	Number of spectral layers	4
T_w	Sliding window size	128
s	Sliding window stride	64
M	Missing mask	[0,1] for [5%, 10%, …45%, 50%]
η	Learning rate	0.001
E	Number of epochs	5,000
ϵ	Stability constant	1×10^{-8}

4.1. Model Development and Architecture

To efficiently train the model on time-series data, we adopt a sliding window approach. The continuous time series is divided into overlapping windows of fixed length $T_w = 128$ time steps with a stride $s = 64$ time steps (details in Table 1). For each window, the input matrix is formed by stacking the rainfall, ST, and a normalized time coordinate (a linear grid ranging from 0 to 1). Thus, the input for each window is represented as follows:

$$X_{\text{window}} \in \mathbb{R}^{T_w \times 3}, \tag{1}$$

and the corresponding target SM values are denoted by

$$Y_{\text{window}} \in \mathbb{R}^{T_w}. \tag{2}$$

This segmentation captures both short-term and long-term dependencies while increasing the number of training samples.

At the core of our methodology is the FNO, which leverages the Fourier transform to learn a global mapping between the input and output. Our goal is to learn an operator $\mathcal{G}$ such that

$$Y_{\text{pred}} = \mathcal{G}(X), \tag{3}$$

where X represents the input features, and Y_{pred} denotes the imputed SM.

Our model first "lifts" the input window into a higher-dimensional space to better capture its complex features. Given an augmented input, where the normalized time grid $g \in \mathbb{R}^{T_w \times 1}$ is concatenated with X_{window}, we compute the lifted representation as

$$X_0 = \text{fc0}(\text{concat}(X_{\text{window}}, g)) \in \mathbb{R}^{T_w \times w}, \tag{4}$$

with $w = 64$ representing the network width.

The lifted features are then processed through a series of L spectral convolution layers. In each spectral layer, the following steps are performed. First, the input is permuted so that the feature dimension precedes the temporal dimension:

$$x \in \mathbb{R}^{w \times T_w}. \tag{5}$$

Next, a discrete Fourier transform converts the time-domain signal into the frequency domain:

$$\hat{x} = \mathcal{F}(x). \tag{6}$$

To focus on the most informative global features, only the first $m = 16$ Fourier modes are retained:

$$\hat{x}_m = \hat{x}[:, :, 1:m]. \tag{7}$$

The spectral convolution is performed by multiplying these truncated Fourier coefficients element-wise with learned complex weights $W^{(l)}$:

$$\hat{y} = \hat{x}_m \odot W^{(l)}, \tag{8}$$

where $\odot$ denotes element-wise multiplication. An inverse Fourier transform then converts the modified coefficients back into the time domain:

$$y = \mathcal{F}^{-1}(\hat{y}). \tag{9}$$

Finally, a residual connection is established by applying a 1×1 convolution (denoted as conv1x1) to the original permuted input x and adding the result to y. This sum is passed through a ReLU activation function to produce the output of the spectral layer:

$$x_l = \mathrm{ReLU}(y + \mathrm{conv1x1}(x)). \tag{10}$$

This sequence, lifting the input, transforming it into the frequency domain, performing a truncated spectral convolution, applying an inverse transform, and adding a residual connection with a nonlinear activation, enables the FNO to capture both local and global temporal dynamics.

After processing through all spectral layers, the final representation is permuted back to its original order and then passed through two additional fully connected layers to project the features to the target space, yielding the final prediction:

$$Y_{\mathrm{pred}} \in \mathbb{R}^{T_w \times 1}. \tag{11}$$

The complete FNO-based imputation process is presented in Algorithm 1 which methodically presents the steps from feature lifting and frequency domain transformation to the final imputation strategy.

4.2. Model Training

For training, we utilize a masked mean squared error loss function that computes the error only on the observed (nonmissing) data points (J. Lin et al., 2023). The loss function is defined as follows:

$$\mathcal{L} = \frac{\sum_{i=1}^{T_w} M(i)(Y_{\mathrm{pred}}(i) - Y(i))^2}{\sum_{i=1}^{T_w} M(i) + \epsilon}, \tag{12}$$

where $M(i)$ indicates whether a measurement is observed (1) or missing (0), and $\epsilon = 1 \times 10^{-8}$ is a small constant to ensure numerical stability.

We train the model using the Adam optimizer with a learning rate $\eta = 0.001$ for $E = 5000$ epochs. The Adam optimizer updates the parameters θ at each iteration t based on the gradient g_t of the loss function (Dubey et al., 2019). It maintains an exponential moving average of the gradients and their squares, given by

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \tag{13}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \tag{14}$$

where β_1 and β_2 are decay rates (typically $\beta_1 = 0.9$ and $\beta_2 = 0.999$) (Xie et al., 2024). To correct the bias introduced in the moving averages during the initial iterations, the estimates are bias-corrected (Kingma & Ba, 2014):

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1-\beta_2^t}. \tag{15}$$

The model parameters are then updated as follows:

$$\theta_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}. \tag{16}$$

This adaptive update mechanism allows for efficient and stable convergence, even when the gradients exhibit high variance.

Additionally, we explore various temporal lag settings ranging from no lag to a lag of five time steps (i.e., up to 50 min for 10-min resolution data) to incorporate different degrees of recent historical context (Reichstein et al., 2019). This is particularly important for capturing rapid infiltration and redistribution processes in SM following rainfall events. The inclusion of up to five lags was based on empirical testing, which indicated that additional lags beyond this point provide diminishing returns in accuracy while increasing computational complexity. Thus, this configuration offers a practical balance between predictive performance and model efficiency.

After the training phase, the FNO model is applied to the full-time series using the sliding window approach. Overlapping window predictions are aggregated (by averaging) to reconstruct the complete time series and impute the missing values. The imputation performance is evaluated using several metrics, including the Pearson correlation coefficient R, RMSE, and bias. These metrics are computed for the overall data set and separately for periods with rainfall and during dry periods, providing a detailed assessment of the model's sensitivity to environmental conditions.

In summary, our methodology integrates rigorous data preprocessing, a novel FNO model that leverages Fourier-domain operations to capture global spatiotemporal dependencies and a robust sliding window strategy for training and imputation. By varying the missing data percentage, temporal lag settings, and evaluating performance under different precipitation conditions, we provide a comprehensive framework that addresses the challenges of reconstructing missing SM data and offers insights into optimal imputation strategies for dynamic, event-driven environmental data.

5. Results

5.1. Model Performance

Figure 2, scatter plots comparing observed versus imputed SM at 10 cm depth were generated for overall data, rain conditions, and no-rain conditions across missing data percentages from 5% to 50% in 5% increments. At lower missing data levels (e.g., 5%), no-lag scenarios typically showed excellent agreement with observed values, with an average correlation coefficient (R) of approximately 0.99 and an RMSE near 1.21, as evidenced by data points clustering tightly along the y = x line. Notably, under rain conditions at 5% missing data, certain lag scenarios (lag_4) outperformed the no-lag case, achieving an R of 0.99 and an RMSE of 1.88, suggesting that incorporating a lag can, under favorable conditions, slightly enhance the model's responsiveness to recent temporal changes. However, as the percentage of missing data increases, a critical observation emerges: both lag and no-lag scenarios exhibit increased scatter, but the divergence becomes more pronounced in the lag cases. For instance, at 50% missing data, no-lag scenarios show a moderate decline in performance, with R values dropping to roughly 0.97 and RMSE rising to about 2.94, while lag scenarios (lag_4) further degrade to an R of approximately 0.96 and an RMSE near 3.07. This indicates that, although lags can capture temporal dependencies, they also tend to amplify the noise in the imputed estimates when data become sparse. The inset scales that plot RMSE and R offer a clear depiction of these performance differences: they reveal that the advantage of lagged imputation observed under lower missing data conditions diminishes at higher missing data levels, with lag one consistently serve as a robust benchmark. Overall, while specific lag configurations can yield marginal improvements under

10.1029/2025JH000730

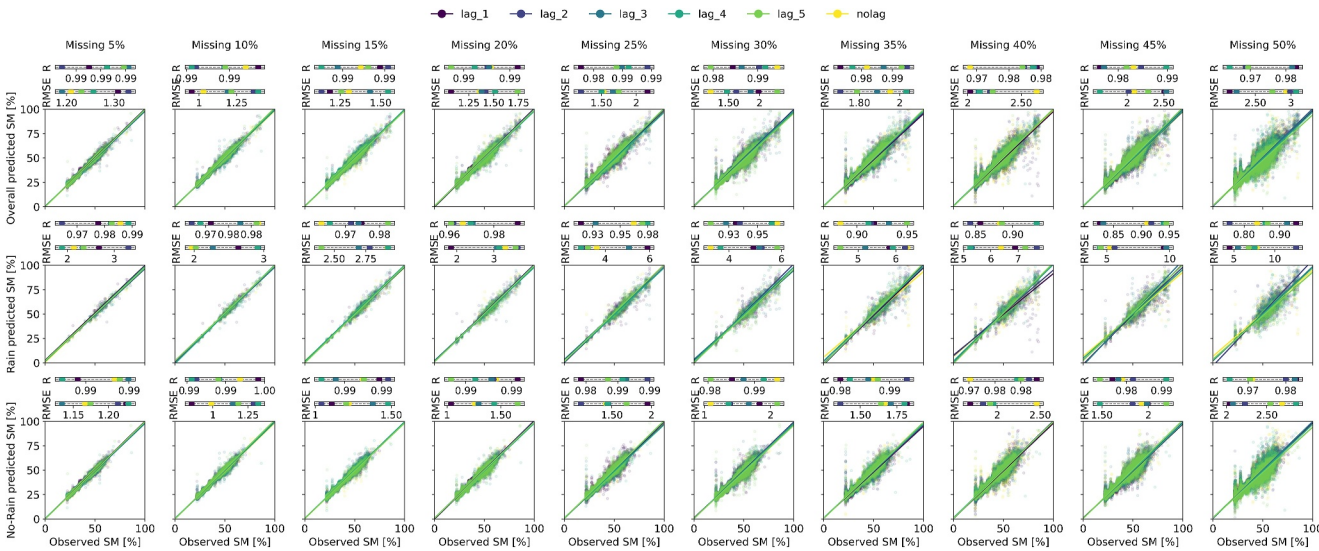

Figure 2. Scatter plots with corresponding linear regression fits of observed versus imputed soil moisture at a 10-cm depth under varying missing data percentages and rainfall conditions. The figure is arranged in a grid with three rows and multiple columns. The rows represent different categories: overall predictions (top row), predictions during rainfall (middle row), and predictions during no-rain conditions (bottom row). Each column corresponds to a specific missing data percentage, ranging from 5% to 50% in 5% increments. Within each subplot, individual scenarios are represented by distinct colors. A linear regression line is fitted to the imputed data points for each scenario, and the performance metrics (R and RMSE) are dynamically displayed in inset panels above the main plot.

certain conditions (notably during rainfall), the general trend is that the introduction of a lag increases variability and scattering, thus adversely affecting imputation accuracy as the missing data percentage grows.

We also performed an error histogram analysis, where we first computed the error for each imputed instance as the difference between the predicted (imputed) and observed SM (Figure 3). For each group (overall, rain, and no-rain), we generated histograms using 20 bins and normalized these histograms so they represent probability densities. A vertical dashed line is drawn at the zero-error mark, which serves as a visual benchmark for identifying bias. Overlaid on these histograms is a Gaussian fit, from which we extracted the mean and the standard

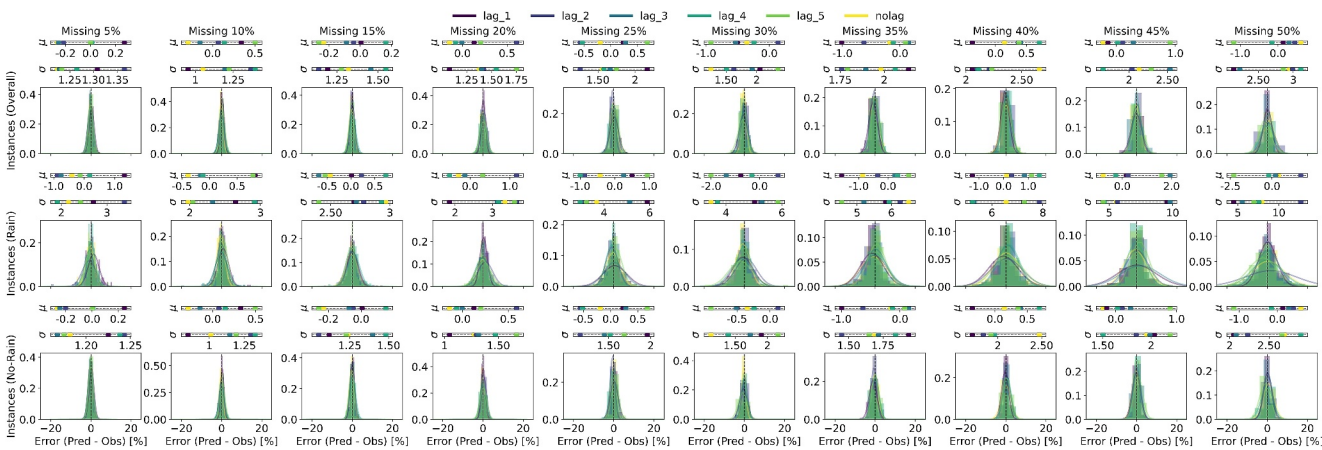

Figure 3. Error histogram analysis of the imputation performance for soil moisture (SM) at a 10-cm depth under varying missing data ratios. Each column corresponds to a specific missing data percentage, while the three rows depict error distributions for (top) overall instances, (middle) instances during rainfall, and (bottom) instances during no-rain conditions. For each subplot, the histogram of the error (defined as the difference between predicted and observed SM) is plotted with overlaid Gaussian curves fitted using the mean (μ) and standard deviation (σ) estimated via maximum likelihood. Different imputation scenarios (distinguished by a "no-lag" or "lag" label) are assigned unique colors to ensure consistency across plots. Inset panels above each subplot dynamically display the Gaussian parameters (μ and σ) for each scenario, facilitating a quantitative assessment of the imputation bias and dispersion. For better visualization, the error range is uniformly set from -25 to 25. A vertical dashed line at zero error is provided for reference.

deviations (σ) as a measure of error spread (or RMSE). These parameters allow us to quantify both the bias (through the closeness of mean to zero) and the dispersion (through the magnitude of σ).

For the overall error category, the results indicate that both lag and no-lag approaches yield mean errors very close to zero, typically within ±1%. For instance, at a 10-cm sensor depth and 10% missing data, the no-lag configuration might show a mean of approximately −0.29% with a σ around 1.05%, while the lag approach (lag_1) shows a mean of about +0.34% and a σ of roughly 0.95%. This represents a reduction in the error spread of about 10% when employing the lag method, demonstrating that even when the predictions are nearly unbiased, the lag approach helps in tightening the overall error distribution.

During rain events, the SM dynamics become more volatile, which naturally leads to higher absolute σ values. For example, at 10 cm depth and with moderate missing data (e.g., 20%), the no-lag scenario might record a σ of about 3.33% compared to roughly 1.85% with the lag approach (lag_1), a reduction of approximately 44.4%. Although the mean errors in both cases remain very close to zero (often within ±0.5%), the relative improvement in σ confirms that the lag adjustment provides a significant benefit by reducing the spread in predictions during periods of rainfall.

In the no-rain condition, although the inherent variability is somewhat lower, we observe a similar trend. At higher missing data levels, such as 40%, the no-lag case exhibits a σ of around 2.47%, while the lag method (lag_1) reduces this spread to about 1.71%, translating to an improvement in the error dispersion of roughly 30.8%. The mean error in these scenarios continues to be near zero (typically within ±0.5%), indicating that both approaches are unbiased; however, the lag technique consistently achieves a tighter error distribution.

Overall, the analysis reveals that although rain events naturally lead to higher RMSE values compared to no-rain periods, the incorporation of immediate lag corrections consistently reduces the error spread, although this benefit deteriorates with higher lags and missing ratios.

5.2. Comparison With Benchmark

The benchmarking centers on evaluating a broad spectrum of imputation techniques applied to SM data under various missingness scenarios. At the core of the study is the FNO, a model that transforms data into the frequency domain to capture complex spatial and temporal dependencies. This advanced framework is compared against more conventional methods. Traditional approaches such as mean, median, and mode imputation replace missing values using simple statistical summaries, providing computational efficiency but lacking in the ability to reflect inherent variability (Harel & Zhou, 2007). More refined techniques, namely Bayesian, KNN, and RF imputation leverage regression-based models and ensemble learning to better capture underlying patterns by iteratively refining estimates using observed data (W.-C. Lin & Tsai, 2020). Additionally, deep learning imputation is implemented via a simple multilayer perceptron that learns representations from the indices of available data, while methods like hotdeck and colddeck imputation employ either random selection of donor values or predetermined constants, respectively, to fill gaps.

The performance of various imputation methods was assessed by constructing Dolan-Moré performance profiles based on scaled RMSE metrics (Figure 4). First, benchmark summary metrics for each method were collected from experimental runs at a 10-cm sensor depth. For each unique missing ratio (considered as an independent problem instance), the raw RMSE values (overall, during rain events, and during no-rain periods) were normalized by the maximum RMSE observed within that group. This scaling produces a relative performance measure where a perfect imputation (0 RMSE) is assigned a value of 0, and the worst case is normalized to 1. This normalization step is essential to ensure that comparisons across different scenarios and missing ratios are both fair and robust. The normalized metrics is then used to construct Dolan-Moré performance profiles, which illustrate the fraction of problem instances for which each method's performance is within a given threshold (τ) of the best observed performance. Specifically, for a continuum of thresholds (τ ranging from 0 to 1), the proportion of cases where a method's scaled RMSE falls below each τ was computed. The resulting cumulative fractions characterize the relative efficiency of each imputation method.

Across all scenarios, the FNO model consistently emerges as the best performer. Under overall conditions (i.e., aggregating all cases), the FNO model shows a strong performance advantage. For instance, at a τ of 0.2, meaning that the scaled RMSE is within 20% of the best performance, FNO achieves a cumulative fraction of 100% of the problem instances. In comparison, traditional methods such as mean imputation and Bayesian imputation record

10.1029/2025JH000730

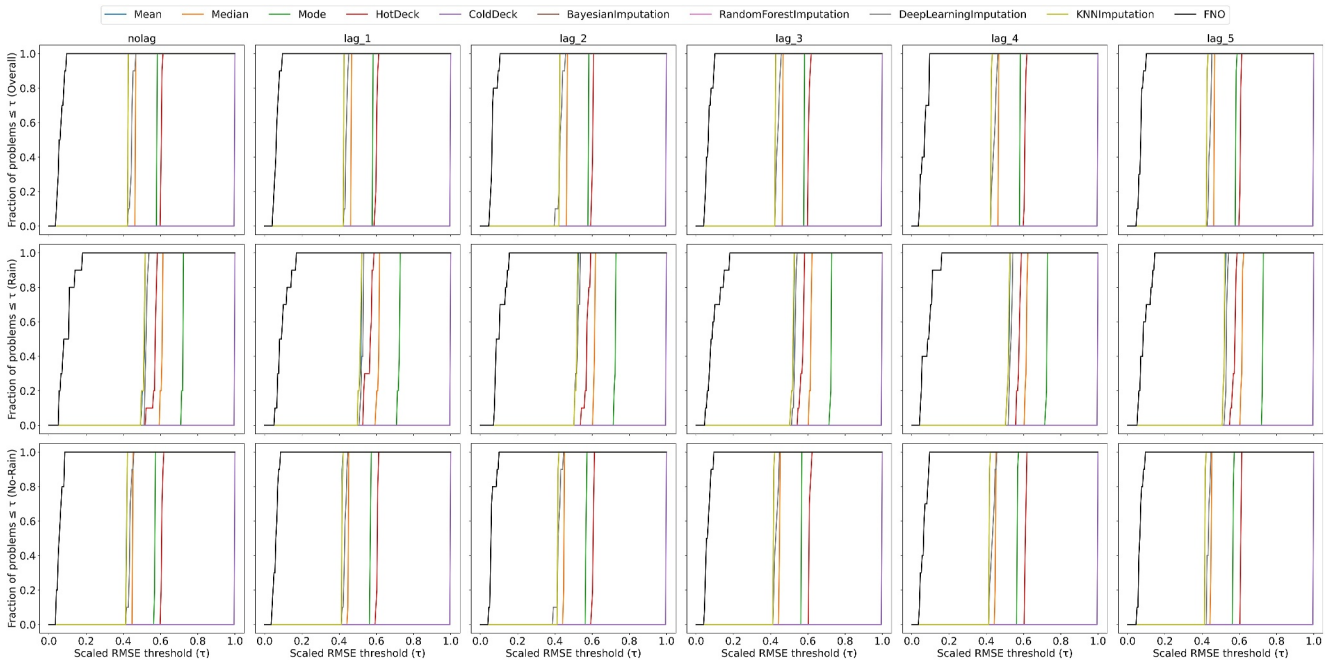

Figure 4. Dolan-Moré performance profiles based on scaled root mean square error (RMSE) for 10 cm soil moisture reconstruction across all scenarios (no-lag and lag conditions). The three rows represent, respectively, the overall, rainy, and no-rain conditions. Each curve shows, for a given imputation method, the fraction of missing data problems (i.e., unique missing ratios) for which the scaled RMSE is below a threshold τ (ranging from 0 to 1). Lower values indicate performance closer to the ideal zero RMSE. Methods such as mean imputation, along with various iterative and tree-based approaches, exhibit comparable performance, suggesting a similar convergence (overlapping) in their imputation outcomes.

cumulative fractions of 0% at this threshold. As τ increases to 0.5, FNO maintains its dominance with 100% coverage, whereas most other methods, including mean, KNN, RF, and deep learning-based imputations, reach 100% only at this relaxed threshold. This reinforces FNO's consistent and superior performance, even under stricter evaluation criteria. When the analysis is grouped by environmental conditions, the impact of rainfall becomes evident. During rain events, all models experience a drop in performance due to the abrupt fluctuations and increased variability in SM, likely due to rapid water infiltration and heterogeneous distribution across sensor depths, thereby increasing the difficulty of accurate imputation. At $\tau = 0.2$, the FNO model achieves a perfect 100% coverage under rain conditions, while all traditional methods perform poorly, each covering only 0% of the problem instances. Even at $\tau = 0.5$, FNO maintains its 100% performance, whereas conventional methods (Bayesian imputation, KNN imputation, mean, and random forest imputation) lag behind with coverage of just 20%. Conversely, under no-rain conditions, the performance gap remains wide, and FNO again covers 100% of the cases at both $\tau = 0.2$ and $\tau = 0.5$, while most other methods (Bayesian, Mean, KNN, etc.) reach full coverage only at the more relaxed $\tau = 0.5$ threshold, indicating that the deep learning-based approach is exceptionally robust when the SM dynamics are less disrupted by external factors. Furthermore, when evaluating across different scenarios (the "no-lag" scenario and various lag scenarios), these trends persist.

For a robust comparison of all the algorithms, we perform a statistical analysis using pairwise Welch's t-tests on the error distributions. This test does not assume equal variances or equal sample sizes between methods, making it particularly suitable for our heterogeneous data conditions. The analysis computes p-values for each pairwise comparison across different scenarios and conditions (overall, rain, no-rain), where a p-value below the significance threshold ($\alpha = 0.05$) indicates that the difference in performance between the two methods is statistically significant. This approach provides a rigorous quantitative assessment of whether the observed differences in error metrics are due to random chance or reflect true performance differences among the methods. From Figure 5, it is evident that the FNO method consistently outperforms the other methods, showing statistically significant improvements across most scenarios and conditions. Methods such as mean, median, and mode, as well as some imputation-based approaches like Bayesian, KNN, and RF, exhibit similar performance with no significant differences in many comparisons. This similarity could be attributed to these methods relying on comparable statistical assumptions or imputation frameworks, leading to analogous error distributions under

10.1029/2025JH000730

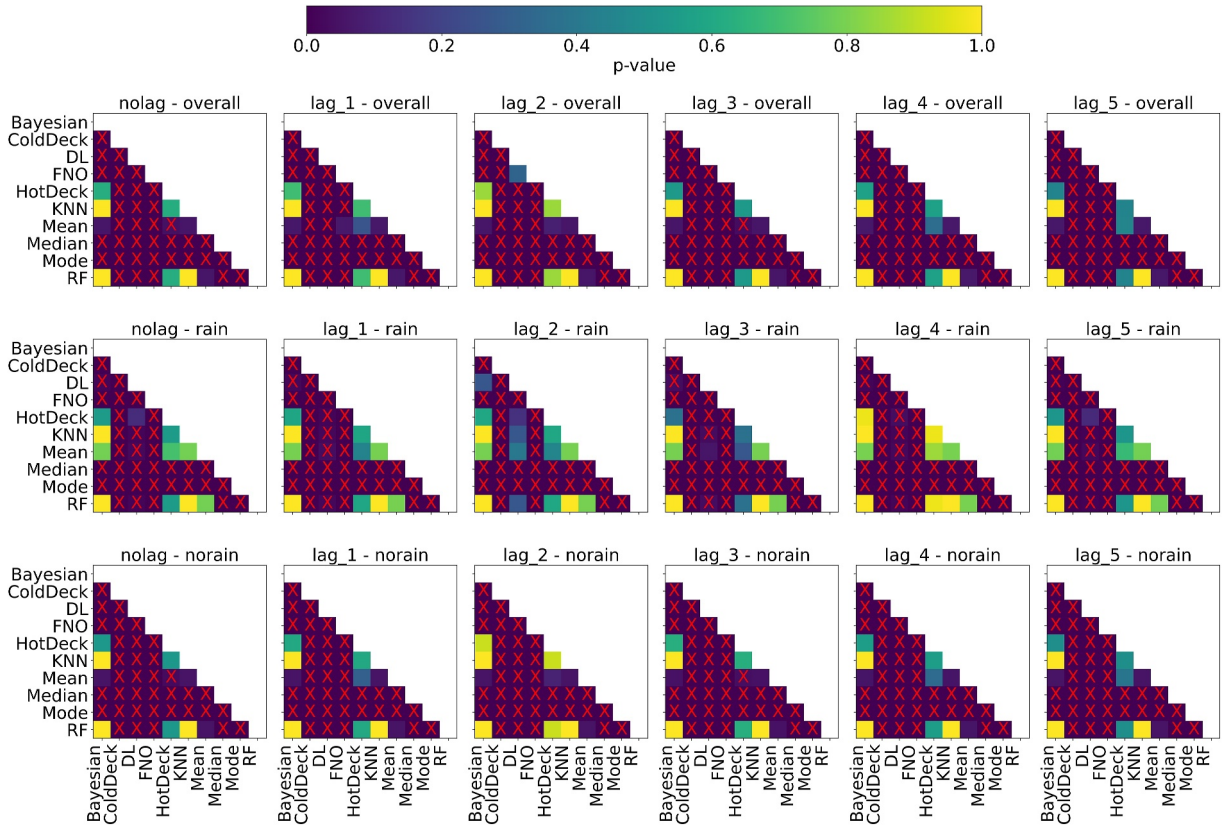

Figure 5. Heatmaps summarizing the results of pairwise Welch's t-tests comparing the reconstruction error defined as the difference between predicted (or imputed) and observed soil moisture values across various imputation methods under different scenarios at a10-cm depth. The panel is arranged as follows: the top row corresponds to the overall data set (all imputed observations), the middle row to rain events (imputed rows with rainfall > 0), and the bottom row to no-rain events (imputed rows with rainfall = 0). Each column represents a different lag scenario. The following imputation methods are compared: Fourier neural operator, mean, Bayesian (abbreviated from Bayesian imputation), K-nearest neighbors (abbreviated from KNN imputation), RF (abbreviated from random forest imputation), DL (abbreviated from deep learning imputation), median, mode, hotdeck, and colddeck. Cells with a p-value below 0.05 (indicating a statistically significant difference between the error distributions of the corresponding methods) are marked with a red "X." The color intensity in each cell indicates the magnitude of the p-value, as shown by the accompanying color bar.

certain conditions. Overall, the analysis robustly indicates that FNO is significantly better than all the other methods, reinforcing its potential as the preferred choice for SM prediction in our study.

5.3. Temporal Distribution Analysis

The temporal distribution analysis reveals that our results are robust and not the product of random chance (Figure 6). In this analysis, we varied the imputation time steps by altering the random seed from 1 to 30 used for model initialization, resulting in 30 separate experiments. For each seed, we plotted a split violin plot to compare the RMSE values under rain and no-rain scenarios. Each plot incorporates data from 60 different points (derived from 6 lag/no-lag configurations combined with 10 different missing ratios), thus encapsulating a wide range of imputation conditions that minimize the influence of outliers or singular instances.

We found that the differences in the median, 25th, and 75th percentiles across all seeds within each case (rain or no-rain) are minimal. For instance, the rain scenario generally has a median of around 3.78–4.56 (within ≈0.8%), while the no-rain scenario consistently falls around 1.47–1.69 (within ≈0.2%). Moreover, the interquartile ranges for both scenarios display a considerable overlap across the seeds, indicating that the variability within each condition is limited and the performance is stable over different imputation time steps. This stability implies that the observed differences in model performance between the rain and no-rain scenarios are systematic rather than a result of random variation.

10.1029/2025JH000730

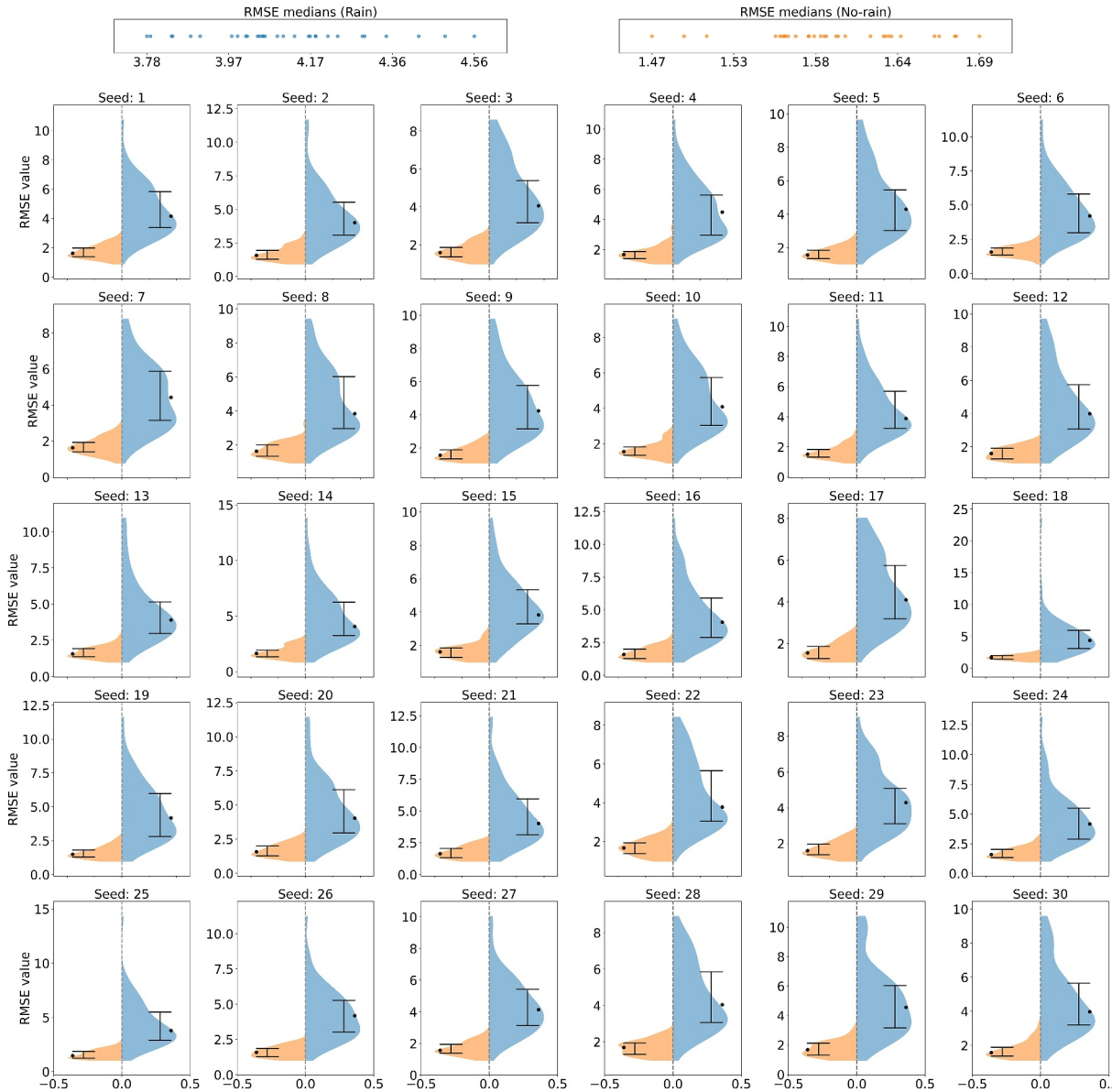

Figure 6. Temporal distribution analysis of root mean square error (RMSE) values for rain and no-rain conditions across 30 seeds, measured at a sensor depth of 10 cm. Each subplot (labeled with the seed number) displays a split violin plot where the right half represents the distribution of RMSE rain values, and the left half represents the RMSE no-rain values. Black markers denote the median for each distribution, with error bars indicating the interquartile range (25^{th} to 75^{th} percentiles). Inset panels above the main figure dynamically display the median values for both conditions.

6. Discussion

6.1. Multi-Depth Analysis

The analysis of the aggregated results reveals several key trends in the correlation (R) metrics between sensor depths, weather conditions, and missing data ratios. In general, increasing missing data ratios lead to a decrease in R values, indicating that higher levels of missing data degrade the precision and reliability of measured SM responses (Figure 7). In detail, the R metrics under rain conditions tend to exhibit sharper declines and greater variability compared to those under no-rain conditions, reflecting the complex dynamics introduced by rainfall events. Moreover, the comparison between no-lag and lag scenarios highlights the importance of temporal adjustments: the no-lag configuration, which compares measurements without any temporal offset, may underestimate the true relationship by neglecting the inherent delay in the SM response. Incorporating appropriate lags

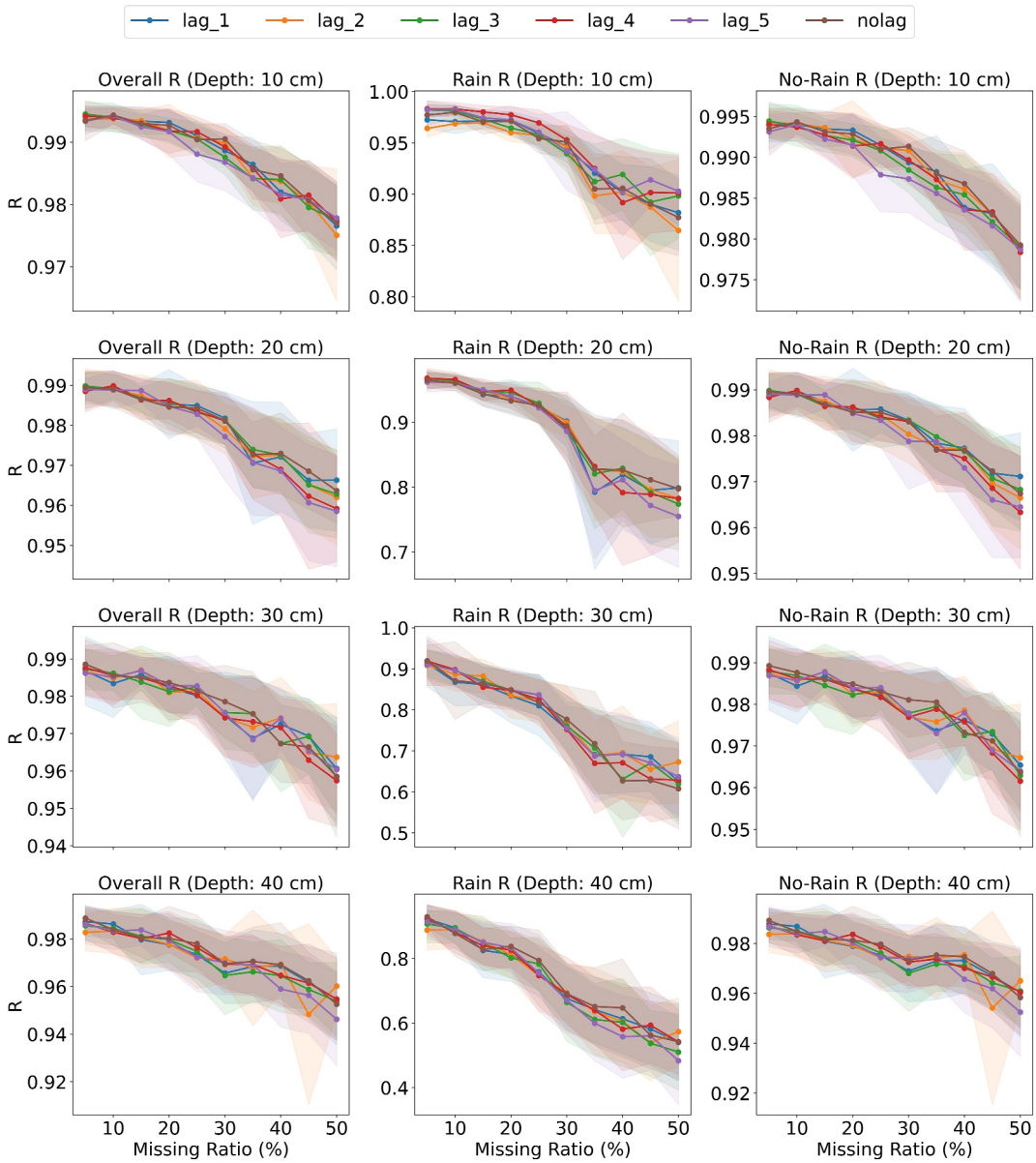

Figure 7. R metrics for overall, rain, and no-rain conditions across sensor depths under varying missing data ratios. This multipanel figure illustrates the effect of increasing missing data ratios on the R of model predictions. Each row corresponds to a specific sensor depth (in cm), while the three columns represent overall R (left), rain R (middle), and no-rain R (right). For each subplot, the solid lines depict the mean R calculated across multiple simulation seeds for each imputation scenario, with the shaded regions indicating $\pm 1\,\sigma$ to highlight variability across seeds. The x-axis shows the missing ratio expressed as a percentage, and the y-axis displays the corresponding R value.

aligns the response timing, particularly enhancing the correlation during rain events and for deeper soil layers. As sensor depth increases from 10 to 40 cm, the overall strength of the correlations diminishes, with the shallower layers showing more immediate and robust responses to rainfall, while deeper layers display attenuated and delayed moisture signals. For instance, the results indicate that at a 10-cm sensor depth in the no-lag scenario, the overall R value decreases from approximately 0.993 with 5% missing data to around 0.990 when the missing ratio reaches 30%, which corresponds to a 0.30% decline. Under rain conditions, a similar trend is observed, though the decline is steeper, dropping by nearly 2.66% (from 0.976 at 5% missing data to 0.950 at 30% missing data), while the no-rain conditions show a more moderate reduction. At deeper levels, such as 40 cm, the overall R starts at roughly 0.988 at 5% missing data and falls to about 0.970 at a 30% missing ratio, marking a reduction of nearly

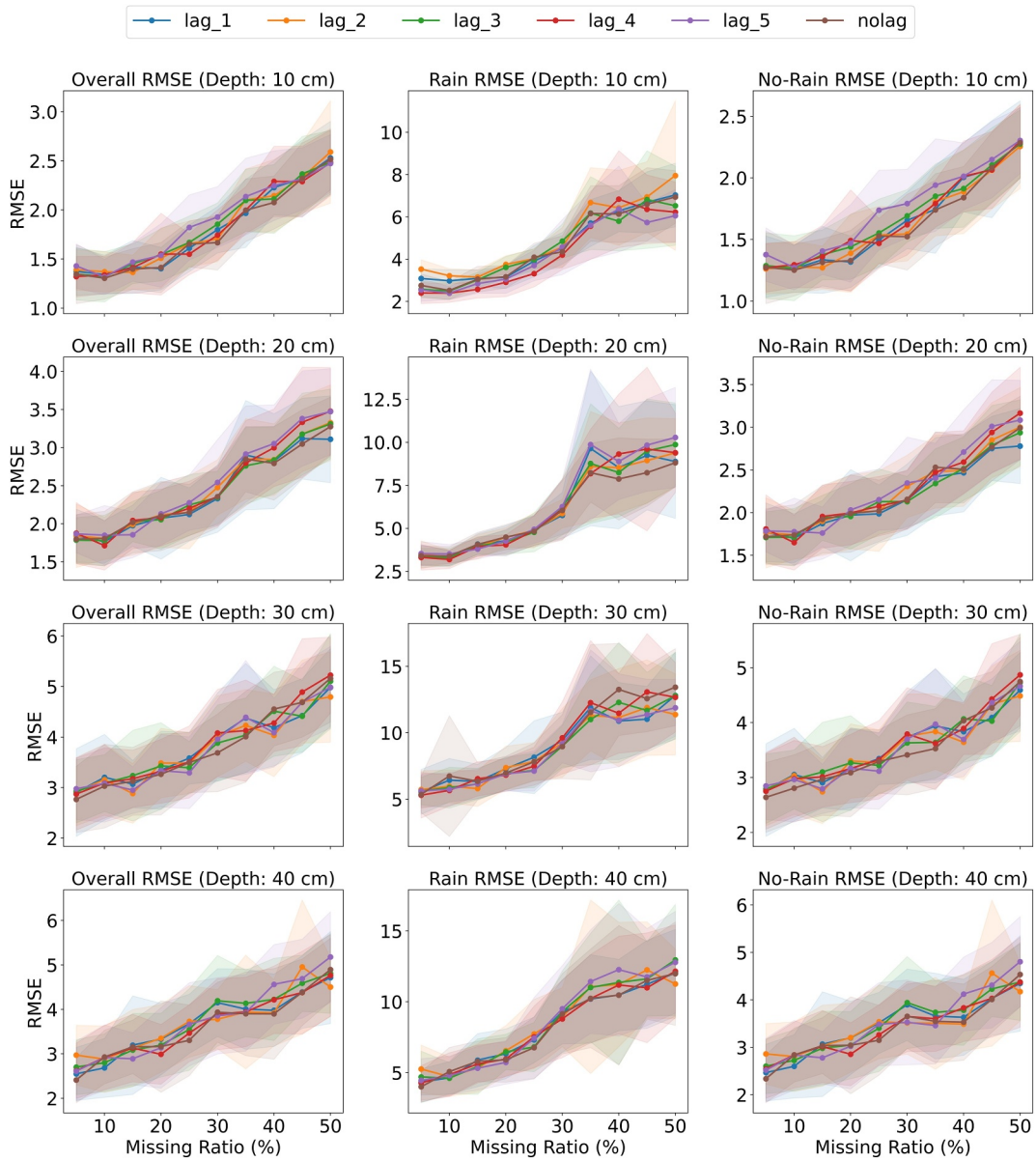

Figure 8. Root mean square error (RMSE) metrics for overall, rain, and no-rain conditions across sensor depths under varying missing data ratios. This multipanel figure illustrates the effect of increasing missing data ratios on the RMSE of model predictions. Each row corresponds to a specific sensor depth (in cm), while the three columns represent overall RMSE (left), rain RMSE (middle), and no-rain RMSE (right). For each subplot, the solid lines depict the mean RMSE calculated across multiple simulation seeds for each imputation scenario, with the shaded regions indicating $\pm 1\ \sigma$ to highlight variability across seeds. The x-axis shows the missing ratio expressed as a percentage, and the y-axis displays the corresponding RMSE value.

1.82%. Notably, when appropriate temporal lags are introduced, improvements of up to 0.025 in R values (approximately 2.5% enhancement) are observed, particularly under rain events, which suggests that accounting for the delayed SM response is crucial. This trend highlights how both increasing data loss and soil depth exacerbate the degradation of correlation metrics and they highlight the benefits of incorporating lags to better align with the intrinsic response times of SM dynamics.

The analysis of the RMSE metrics across sensor depths, weather conditions, and missing data ratios reveals clear trends in the prediction error (Figure 8). As the missing ratio increases, RMSE values also increase, indicating that higher levels of missing data lead to greater discrepancies between observed and predicted (imputed) SM. For

instance, at a 10-cm sensor depth, the overall RMSE in the no-lag scenario might increase from approximately 1.348 at 5% missing data to around 1.679 at a 30% missing ratio, a nearly 24.56% increase in errors. Under rain conditions, this trend is even more pronounced, while no-rain conditions display a comparatively moderate rise in RMSE. At deeper layers, such as 40 cm, the RMSE can start at around 2.413 and escalate to roughly 3.909 with increased missingness, marking an increase of approximately 62%. Moreover, when comparing no-lag with lag scenarios, the introduction of an appropriate temporal lag accounting for the delayed SM response often results in a reduction in RMSE by up to 15%, depending on the precipitation scenarios. For instance, at a 10-cm sensor depth, the no-lag RMSE for rain is approximately decreased from 2.777 to around 2.375, corresponding to a 14.47% reduction. These findings are consistent with the one observed for R, which found that both increased missing data and greater sensor depth significantly degrade prediction accuracy, while incorporating lags can mitigate some of this degradation by better aligning model predictions with the inherent delayed response of subsurface moisture dynamics.

The aggregated bias values quantify the systematic deviation between the observed and predicted SM values across different sensor depths, weather conditions, and missing data levels (Figure 9). Overall, as the missing data ratio increases, the bias tends to become more pronounced, suggesting that higher levels of missing information lead to greater systematic errors in the predictions. For instance, in the no-lag scenario at a shallow depth, the bias might start relatively low, but as the missing ratio increases, the bias can grow noticeably. In deeper layers, the baseline bias is higher, reflecting the increased complexity and delay in the SM response at greater depths. Furthermore, when comparing no-lag to lag scenarios, the introduction of a temporal lag generally results in a reduction of the bias in some scenarios. This improvement is particularly evident under rain conditions, where the delayed response of the soil to rainfall events is better captured, thereby reducing the systematic error. Quantitatively, the application of lags might reduce the bias by approximately 15%–20% on average. These findings emphasize that aligning the timing of SM responses through appropriate lag adjustments is crucial for minimizing systematic prediction errors, especially when dealing with missing data and when monitoring deeper soil layers.

6.2. Performance Across Varying Rain Magnitudes

To study and understand the key contributing factors behind the relatively low accuracy of FNO and other benchmark algorithms during rain events, we analyzed the response of FNO to different rainfall intensities (Figure 10). To do so, rain events were first categorized into three distinct groups (high, moderate, and low) based on the magnitude of rainfall. High-intensity events are defined as those with rainfall exceeding 12 mm, moderate-intensity events as those with rainfall between 6 and 12 mm, and low-intensity events as those with rainfall between >0 and 6 mm. We then automatically selected a 1-week window that satisfied these criteria. Rainfall recorded during the third week of July (monsoon period) is considered high, that observed during the second week of April (premonsoon) is considered moderate, and that captured during the fourth week of July is considered low (monsoon period) (Figure 10).

In our study, we observed that during high-intensity rainfall events (total 64 events recorded), the FNO exhibits higher RMSE values across multiple soil depths. This trend is likely due to the rapid and nonlinear changes in SM during peak monsoon conditions. The heavy rain characteristics of the Indian summer monsoon (JJAS) lead to rapid surface saturation and heterogeneous moisture distribution, making it challenging for imputation algorithms to accurately predict moisture dynamics. Additionally, the limited number of high-intensity events may reduce the statistical robustness of the training data, contributing further to the elevated prediction errors during these extremes.

Conversely, during moderate-intensity rainfall events (with 22 events recorded in the premonsoon period), the SM dynamics are less volatile. In April, soils are typically drier, and the gradual changes induced by moderate rains allow the FNO to capture temporal trends more accurately. The reduced intensity of these events generally results in lower RMSE values compared to high-intensity events, suggesting that the model can more reliably handle gradual transitions in SM. However, the relatively small sample size in this category might limit the overall representativeness of these performance metrics.

For low-intensity events (with 209 rain events), the FNO demonstrates comparatively lower RMSE values. The large number of these events provides a robust data set, enabling the model to effectively learn and generalize the underlying SM dynamics even in the presence of small rainfall-induced changes. In this category, because the changes in SM are not very dramatic, the predictions tend to be more consistent and accurate across all soil depths.

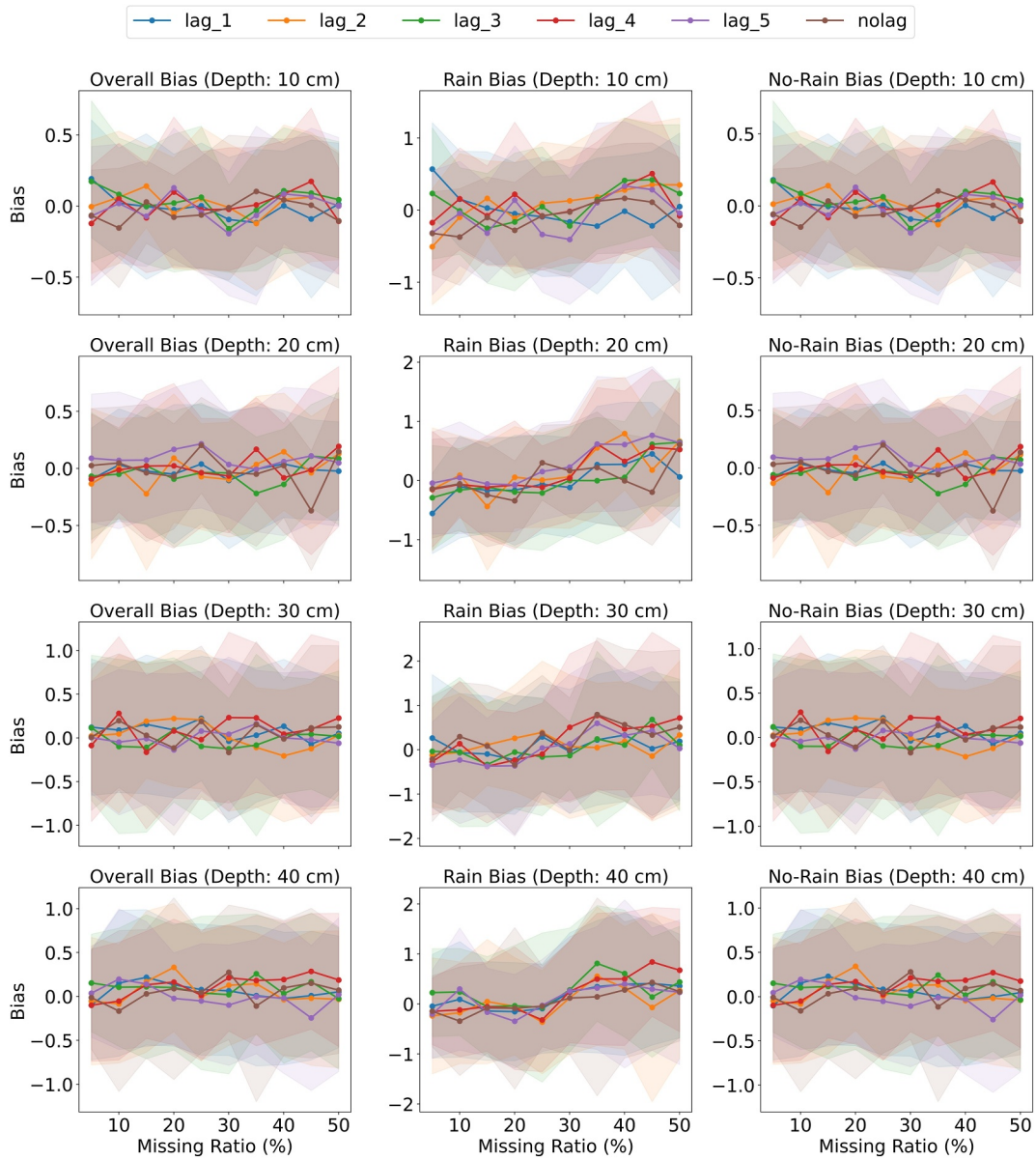

Figure 9. Bias metrics for overall, rain, and no-rain conditions across sensor depths under varying missing data ratios. This multipanel figure illustrates the effect of increasing missing data ratios on bias in model predictions, with each row corresponding to a specific sensor depth (in cm) and the three columns representing overall bias (left), rain bias (middle), and no-rain bias (right). For each subplot, the solid lines depict the mean bias calculated across multiple simulation seeds for each imputation scenario, while the shaded regions represent $\pm 1\ \sigma$, highlighting the variability across seeds. The x-axis shows the missing ratio expressed as a percentage, and the y-axis displays the corresponding bias value.

6.3. Cross-Site Validation and Generalization

To evaluate the generalizability of our model, we have tested its performance on the data of Kapululira in Zambia, which has a humid subtropical climate influenced by monsoons. The data are available from 30 November 2021, at 15:40 to 1 July 2025, at 15:30. It covers a total duration of 1,308 days (approximately 3.6 years). The observatory is located at latitude −15.96336 and longitude 28.88702 and is managed by the Zambia Meteorological Department (ZMD). It records data at 10-min intervals across several depths. In this study, we used data from depths of 5, 10, 20, and 30 cm.

10.1029/2025JH000730

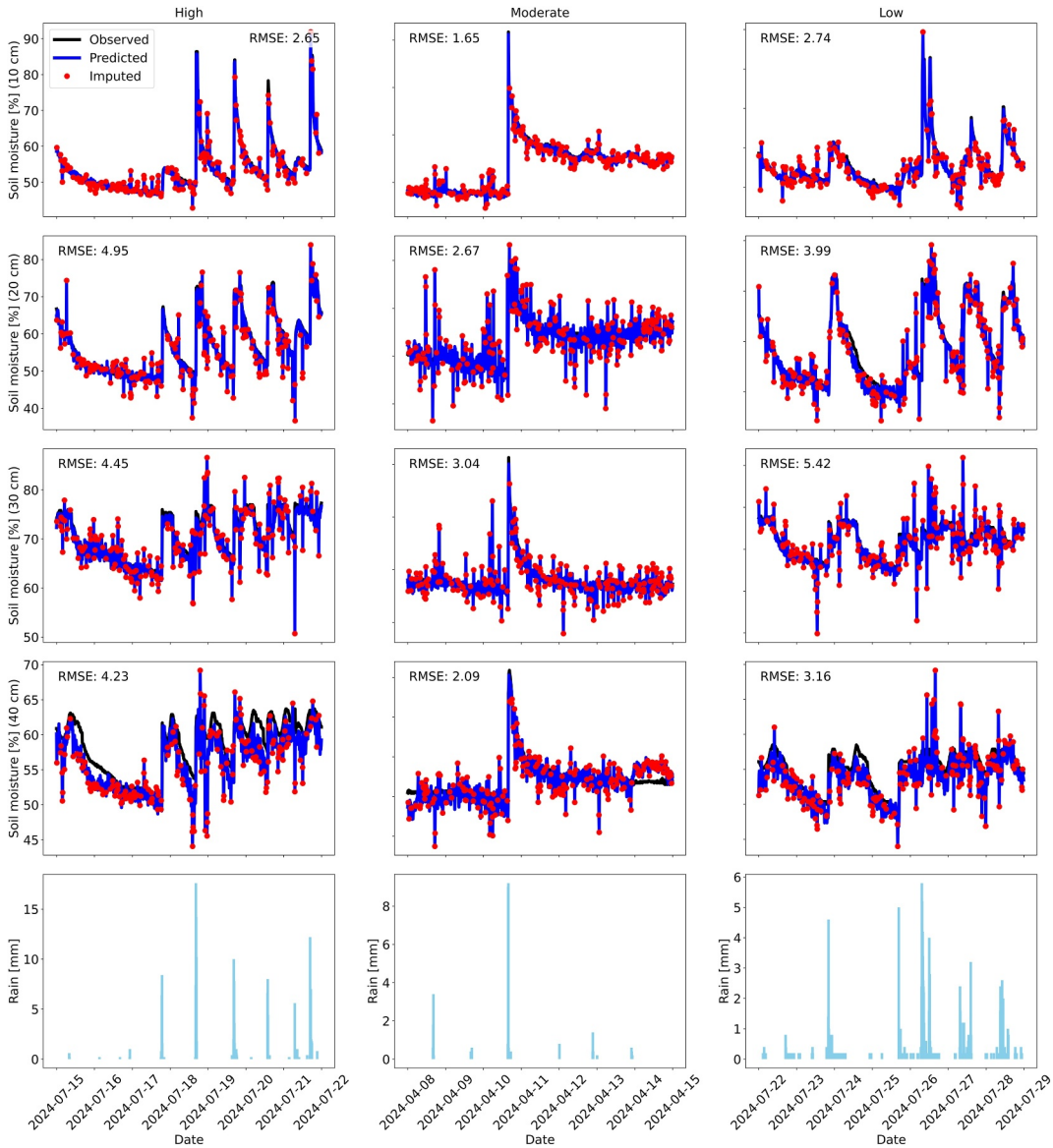

Figure 10. The performance of the Fourier neural operator model in reconstructing soil moisture (SM) across varying rain magnitudes. Each column represents a different precipitation regime: the first column displays data from the third week of July (15–21 July) under high rain intensity, the second column shows data from the second week of April (8–14 April) under moderate rain conditions, and the third column presents data from the fourth week of July (22–28 July) with low rain intensity. For each column, the top four rows provide time series plots of observed versus FNO-predicted SM at depths of 10, 20, 30, and 40 cm, with imputed values highlighted by red markers. The bottom row of each column depicts the corresponding rainfall data as vertical lines, linking rain events to the SM dynamics.

The result of our proposed imputation model at the Kapululira site is consistent with what we previously observed at Bhopal, indicating a strong generalizability (Figure 11). Across all depths, the model demonstrates a consistent trend of increasing RMSE with increasing missing data ratios, reflecting expected sensitivity to the extent of data unavailability.

At the shallowest depth of 5 cm, the model generally performs better, likely due to the higher correlation of surface layer SM with observable atmospheric inputs. As depth increases, the model encounters slightly greater difficulty in imputation, which is in line with the increasing complexity of moisture redistribution processes in deeper layers. However, this depth-wise variation remains consistent and controlled, indicating stable performance across the profile.

10.1029/2025JH000730

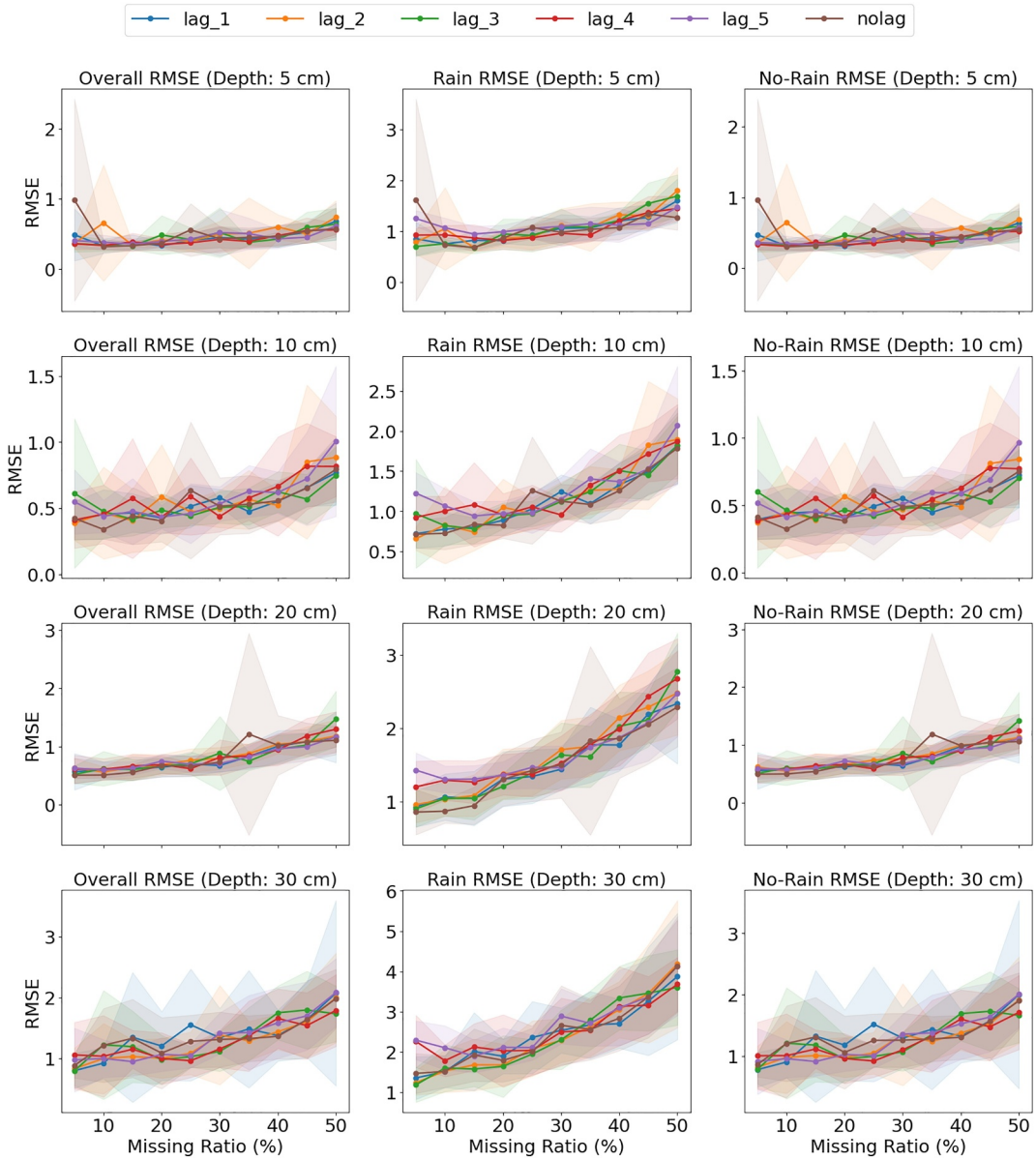

Figure 11. Performance at the Kapululira site (Zambia). Root mean square error (RMSE) metrics for overall, rain, and no-rain conditions across sensor depths under varying missing data ratios. This multipanel figure illustrates the effect of increasing missing data ratios on the RMSE of model predictions. Each row corresponds to a specific sensor depth (in cm), while the three columns represent overall RMSE (left), rain RMSE (middle), and no-rain RMSE (right). For each subplot, the solid lines depict the mean RMSE calculated across multiple simulation seeds for each imputation scenario, with the shaded regions indicating $\pm 1\ \sigma$ to highlight variability across seeds. The x-axis shows the missing ratio expressed as a percentage, and the y-axis displays the corresponding RMSE value.

Under rain and no-rain conditions, the model follows the same pattern observed in Bhopal: RMSE values are higher during rain periods compared to dry spells. This can be attributed to the increased temporal variability and infiltration dynamics associated with rainfall events, which introduce nonlinear fluctuations in SM, especially near the surface. In contrast, during no-rain conditions, SM changes tend to be more gradual, allowing the model to achieve better accuracy.

The evaluation under a short-lag scenario demonstrates the model's ability to effectively leverage recent temporal information for reconstructing missing values. The consistency in the trend across depths, rainfall conditions, and

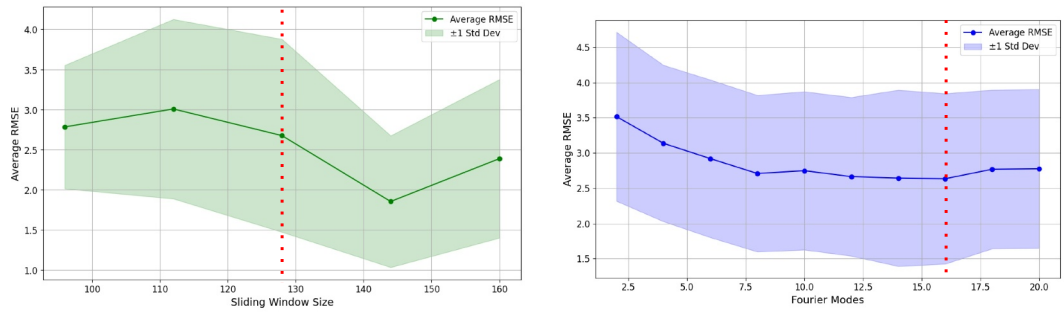

Figure 12. Sensitivity analysis of model performance with varying (a) sliding window sizes and (b) number of Fourier modes. The red dotted lines indicate the values used in this study (window size = 128, Fourier modes = 16).

missing data ratios further strengthens the evidence for the model's robustness and its potential applicability in diverse climatic regions with varying soil and moisture characteristics.

6.4. Hyperparameter Sensitivity and Operational Considerations

To evaluate the influence of key architectural choices on model performance and guide practical deployment, we conducted a systematic sensitivity analysis of two critical hyperparameters: the sliding window size and the number of Fourier modes. This evaluation was performed using the 10-min resolution data, with average RMSE computed across all configurations, including different rainfall conditions (rain/no-rain), temporal lags, and sensor depths.

As shown in Figure 12a, the left panel illustrates model performance across various sliding window sizes: (96, 112, 128, 144, and 160). While window sizes of 144 and 160 resulted in slightly lower average RMSE, we selected a window size of 128 for this study. This value offers a balanced trade-off between prediction accuracy, temporal context, and computational efficiency. Larger windows, though marginally better in accuracy, increase memory demands and can be less practical for real-time or resource-constrained applications. The selection of 128 ensures consistent performance without incurring excessive computational overhead, making it a suitable default configuration for high-frequency (10-min) data. Moreover, this window size can be adapted for other temporal resolutions with minimal tuning, for example, shorter windows for hourly data, where fewer steps are needed to capture equivalent durations.

The right panel of Figure 12b shows the effect of varying the number of Fourier modes from 2 to 20. We observed that the model accuracy improves significantly up to 8 modes, after which it saturates. Although we selected 16 Fourier modes in this study to ensure robustness across all depths and rainfall conditions, lower values (e.g., 8 or 10) can still deliver competitive performance with reduced computational cost. This is particularly beneficial when deploying the model in real-time systems or on limited hardware.

While the FNO model is relatively expensive to train, each run in our setup (using a single NVIDIA A100 GPU) took approximately 0.5 hr over 10,000 epochs, and its inference phase is highly efficient, making it practical for operational applications once trained. The sensitivity analysis presented here offers practical guidance on tuning model parameters to achieve an optimal balance between accuracy and computational efficiency. For instance, reducing the sliding window size from 128 to 112 or using 8 Fourier modes can significantly lower training time and memory usage, with minimal impact on performance. These insights provide a framework for adapting the model to diverse operational contexts, ensuring scalability across different temporal resolutions, computational environments, and data availability scenarios.

6.5. Limitations and Future Work

The limitations of this study include a noticeable performance degradation when the missing data ratio increases, especially under high levels of data sparsity and in deeper soil layers where moisture dynamics are more complex. The incorporation of temporal lags, while beneficial for capturing delayed responses during rain events, sometimes amplifies noise in the imputed estimates, leading to reduced accuracy in certain scenarios. Additionally, the

fixed sliding window approach may not be optimal for all environmental conditions, limiting the model's adaptability.

Future work should focus on enhancing the robustness and generalizability of the FNO framework for environmental data imputation. Potential directions include the development of adaptive windowing techniques and the incorporation of additional environmental variables such as soil type, vegetation cover, or other meteorological parameters to better capture complex dynamics. Furthermore, integrating uncertainty quantification methods and exploring multiscale modeling strategies could help address the noise amplification issues observed with higher missing data ratios.

7. Conclusion

This study demonstrates that the FNO framework offers significant improvements in the imputation of SM data over traditional statistical and machine learning methods. By transforming data into the frequency domain, the FNO effectively captures complex spatial and temporal dependencies, resulting in robust performance across a wide range of scenarios and environmental conditions. Our extensive evaluations across multiple sensor depths, varying missing data ratios, and distinct precipitation scenarios confirm that the FNO not only achieves higher correlation coefficients and lower RMSE values but also maintains its accuracy even under challenging conditions.

While all models exhibit sensitivity to rain events highlighting the need to incorporate additional contextual or auxiliary data when modeling SM under transient conditions, the application of temporal lags within our framework has shown promise in aligning model predictions with the inherent delays in SM responses during rainfall. Overall, the findings provide strong evidence that adopting modern deep learning-based imputation methods, such as the FNO, can lead to more accurate and reliable missing-value reconstructions. These advances are critical for improving environmental monitoring, hydrological forecasting, and ultimately, supporting informed decision-making in environmental management.

Data Availability Statement

In situ data can be downloaded from Zenodo (Singh et al., 2025b). The code developed for this study can be downloaded from Zenodo (Singh et al., 2025a). The same is also available for download at https://abhilashsingh.net/codes.html.

Acknowledgments
We gratefully acknowledge the Geohydrology Observatory Facility at IISER Bhopal, where the experiments were conducted, which played a key role in conceptualizing this study. A.S. would like to thank the Zambia Meteorological Department (ZMD) for providing the gauge station data sets used in this study (accessed through University of Leeds). K.G. would like to acknowledge the funding from the Indo-French Center for the Promotion of Advanced Research: IFCPAR/CEFIPRA through a research Grant: 6707-1, and an institutional Grant received from IISER Bhopal.

References

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), 320–328. https://doi.org/10.1038/s42254-024-00712-5

Bikše, J., Retike, I., Haaf, E., & Kalvāns, A. (2023). Assessing automated gap imputation of regional scale groundwater level data sets with typical gap patterns. *Journal of Hydrology*, *620*, 129424. https://doi.org/10.1016/j.jhydrol.2023.129424

Boomgard-Zagrodnik, J. P., & Brown, D. J. (2022). Machine learning imputation of missing mesonet temperature observations. *Computers and Electronics in Agriculture*, *192*, 106580. https://doi.org/10.1016/j.compag.2021.106580

Boujoudar, M., El Ydrissi, M., Abraim, M., Bouarfa, I., El Alani, O., Ghennioui, H., & Bennouna, E. G. (2024). Comparing machine learning algorithms for imputation of missing time series in meteorological data. *Neural Computing & Applications*, 1–15. https://doi.org/10.1007/s00521-024-10601-8

Cao, Q., Goswami, S., & Karniadakis, G. E. (2024). Laplace neural operator for solving differential equations. *Nature Machine Intelligence*, *6*(6), 631–640. https://doi.org/10.1038/s42256-024-00844-4

Chavoshi, A., Dashtian, H., Bakhshian, S., Young, M. H., & Niyogi, D. (2024). PINN-SM: A physics-informed neural networks model for Vadose zone soil moisture profile prediction. *Authorea Preprints*.

Chen, E., Andersen, M. S., & Chandra, R. (2024). Deep learning framework with Bayesian data imputation for modelling and forecasting groundwater levels. *Environmental Modelling & Software*, *178*, 106072. https://doi.org/10.1016/j.envsoft.2024.106072

Chen, L., Xu, J., Wang, G., & Shen, Z. (2019). Comparison of the multiple imputation approaches for imputing rainfall data series and their applications to watershed models. *Journal of Hydrology*, *572*, 449–460. https://doi.org/10.1016/j.jhydrol.2019.03.025

Chivers, B. D., Wallbank, J., Cole, S. J., Sebek, O., Stanley, S., Fry, M., & Leontidis, G. (2020). Imputation of missing sub-hourly precipitation data in a large sensor network: A machine learning approach. *Journal of Hydrology*, *588*, 125126. https://doi.org/10.1016/j.jhydrol.2020.125126

Dubey, S. R., Chakraborty, S., Roy, S. K., Mukherjee, S., Singh, S. K., & Chaudhuri, B. B. (2019). diffgrad: An optimization method for convolutional neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, *31*(11), 4500–4511. https://doi.org/10.1109/tnnls.2019.2955777

Evans, S., Williams, G. P., Jones, N. L., Ames, D. P., & Nelson, E. J. (2020). Exploiting earth observation data to impute groundwater level measurements with an extreme learning machine. *Remote Sensing*, *12*(12), 2044. https://doi.org/10.3390/rs12122044

Harel, O., & Zhou, X.-H. (2007). Multiple imputation: Review of theory, implementation and software. *Statistics in Medicine*, *26*(16), 3057–3077. https://doi.org/10.1002/sim.2787

Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
Li, W., Migliavacca, M., Forkel, M., Denissen, J. M., Reichstein, M., Yang, H., et al. (2022). Widespread increasing vegetation sensitivity to soil moisture. *Nature Communications*, *13*(1), 3959. https://doi.org/10.1038/s41467-022-31667-9
Lin, J., Gao, F., Shi, X., Dong, J., & Du, Q. (2023). SS-MAE: Spatial–spectral masked auto encoder for multisource remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, *61*, 1–14. https://doi.org/10.1109/tgrs.2023.3331717
Lin, W.-C., & Tsai, C.-F. (2020). Missing value imputation: A review and analysis of the literature (2006–2017). *Artificial Intelligence Review*, *53*(2), 1487–1509. https://doi.org/10.1007/s10462-019-09709-4
Lu, L., Jin, P., Pang, G., Zhang, Z., & Karniadakis, G. E. (2021). Learning nonlinear operators via deeponet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, *3*(3), 218–229. https://doi.org/10.1038/s42256-021-00302-5
Marcy, C., Goforth, T., Nock, D., & Brown, M. (2022). Comparison of temporal resolution selection approaches in energy systems models. *Energy*, *251*, 123969. https://doi.org/10.1016/j.energy.2022.123969
Mitra, McGough, S. F., Chakraborti, T., Holmes, C., Copping, R., Hagenbuch, N., et al. (2023). Learning from data with structured missingness. *Nature Machine Intelligence*, *5*(1), 13–23. https://doi.org/10.1038/s42256-022-00596-z
Pastorini, M., Rodríguez, R., Etcheverry, L., Castro, A., & Gorgoglione, A. (2024). Enhancing environmental data imputation: A physically-constrained machine learning framework. *Science of The Total Environment*, *926*, 171773. https://doi.org/10.1016/j.scitotenv.2024.171773
Pinthong, S., Ditthakit, P., Salaeh, N., Hasan, M. A., Son, C. T., Linh, N. T. T., et al. (2024). Imputation of missing monthly rainfall data using machine learning and spatial interpolation approaches in Thale sap Songkhla river basin, Thailand. *Environmental Science and Pollution Research*, *31*(41), 54044–54060. https://doi.org/10.1007/s11356-022-23022-8
Ramirez, S. G., Williams, G. P., & Jones, N. L. (2022). Groundwater level data imputation using machine learning and remote earth observations using inductive bias. *Remote Sensing*, *14*(21), 5509. https://doi.org/10.3390/rs14215509
Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat, F. (2019). Deep learning and process understanding for data-driven earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1
Richman, M. B., Trafalis, T. B., & Adrianto, I. (2009). Missing data imputation through machine learning algorithms. In *Artificial intelligence methods in the environmental sciences* (pp. 153–169). Springer.
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., Singh, V., & Gaurav, K. (2025a). Fno imputation algorithm [code]. *Zenodo*. https://doi.org/10.5281/zenodo.15178763
Singh, A., Singh, V., & Gaurav, K. (2025b). Fno imputation datasets [dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.15849521
Xie, X., Zhou, P., Li, H., Lin, Z., & Yan, S. (2024). Adan: Adaptive Nesterov momentum algorithm for faster optimizing deep models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *46*(12), 9508–9520. https://doi.org/10.1109/tpami.2024.3423382
Zhang, Y., & Thorburn, P. J. (2022). Handling missing data in near real-time environmental monitoring: A system and a review of selected methods. *Future Generation Computer Systems*, *128*, 63–72. https://doi.org/10.1016/j.future.2021.09.033
Zhang, Z., Yang, X., Li, H., Li, W., Yan, H., & Shi, F. (2017). Application of a novel hybrid method for spatiotemporal data imputation: A case study of the minqin county groundwater level. *Journal of Hydrology*, *553*, 384–397. https://doi.org/10.1016/j.jhydrol.2017.07.053