https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

Evaluation of ASCAT soil moisture retrievals and their potential to detect intraday variability

Lan Anh Dinh[1], Filipe Aires[2,1], and Victor Pellet[3,2]

[1]Estellus, Paris, France
[2]LIRA, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, Paris, France
[3]LMD, École Polytechnique, Palaiseau, France

Correspondence: Lan Anh Dinh (lananh.dinh@hotmail.com)

Abstract. Accurate sub-daily soil moisture (SM) retrievals from satellite observations remain a major challenge due to sparse temporal sampling and retrieval uncertainties. This study introduces a localized convolutional neural network (CNN-l) framework designed to enhance SM estimates from Advanced SCATterometer (ASCAT) observations by exploiting spatial features and adapting to local conditions. The proposed approach achieves strong agreement with ERA5 reference SM, with total correlation coefficients exceeding 0.9, even at a sub-daily scale. Validation against *in situ* measurements from 568 monitoring sites across the contiguous United States (CONUS) shows a median temporal correlation of 0.65, compared to 0.59 for the operational ASCAT H120 product. Our CNN-based retrievals also reveal meaningful intraday variability when SM signals exceed retrieval uncertainty, particularly during heavy precipitation events (> 10 mm day^{-1}), offering new insight into short-term hydrological responses. Future efforts should prioritize the integration of complementary satellite observations from multiple instruments to enhance retrieval accuracy, robustness, and temporal resolution. Additionally, strategies to improve retrieval of extremes (such as localization strategies or variable augmentation) should be further developed.

1 Introduction

Surface soil moisture (hereafter referred to as soil moisture, SM) represents a relatively small component of the hydrological cycle (Trenberth et al., 2007), but plays a crucial role in land–atmosphere interactions and is vital for a variety of applications, including flood forecasting, agriculture, and water management (Bateni and Entekhabi, 2012; Ochsner et al., 2013; McColl et al., 2017). Reflecting its importance, significant efforts have been made to provide consistent SM datasets such as *in situ* SM measurements (e.g., through the International Soil Moisture Network, ISMN (Dorigo et al., 2013, 2021)), or SM estimates from satellite remote sensing (Aires et al., 2001; Kolassa et al., 2013; Araya et al., 2021; Singh and Gaurav, 2023; Han et al., 2023). Long-term operational products have been delivered by missions such as the European Remote Sensing satellite (ERS) scatterometer (Wagner et al., 1999a); Advanced SCATterometer (ASCAT) onboard the Metop satellites (Bartalis et al., 2007); the Soil Moisture and Ocean Salinity (SMOS) mission (Kerr et al., 2010); and the Soil Moisture Active Passive (SMAP) mission (Entekhabi et al., 2010). Apart from these SM products, directly retrieved from single satellite platforms, the European Space Agency Climate Change Initiative for Soil Moisture (ESA CCI SM) is performing the data fusion of all these satellite observations to obtain a consistent satellite-based long-term Climate Data Records (CDRs) of SM (Dorigo et al., 2017).

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

If, historically, the retrieval of SM from active instruments (e.g., ERS and ASCAT) is partly based on a statistical approach (Wagner et al., 1999a), in recent decades, there has been a growing interest in employing purely statistical machine learning approaches for SM retrieval. Among these, neural networks (NNs) have emerged as powerful tools due to their ability to capture complex, nonlinear relationships between satellite observations and surface SM. Numerous studies have demonstrated the effectiveness of NNs in improving retrieval accuracy by exploiting large datasets and learning hierarchical representations of input features (Aires et al., 2001; Kolassa et al., 2017; Rodríguez-Fernández et al., 2017; Yao et al., 2021; Pellet et al., 2025), in particular for ASCAT observations (Aires et al., 2021a). Many existing retrieval products, however, operate at the pixel level (e.g., ASCAT soil moisture product and NN retrievals (Aires et al., 2021a)), neglecting the spatial dependencies present in satellite observations. Convolutional neural networks (CNNs) offer a promising alternative by exploiting spatial patterns, potentially improving retrieval performance beyond pixel-based approaches.

Despite the wealth of research on SM retrieval, most existing models operate at a daily temporal resolution, mainly due to too high uncertainties in retrieval (i.e., too low signal-to-noise ratio). Although daily estimates can be sufficient for many applications, the availability of sub-daily (e.g., hourly) SM retrievals could unlock additional insight into dynamic processes such as soil–plant–atmosphere interactions, diurnal variation in evapotranspiration, and short-term hydrological forecasting (Vilà-Guerau de Arellano et al., 2020; Vermunt et al., 2022). However, relatively few studies have focused on this finer temporal scale (Kim et al., 2021; Wang et al., 2024). Notably, the Metop ASCAT CDR (H SAF, 2021a), which we use later in this study for comparison, already provides SM estimates at sub-daily resolution (i.e., at native overpass times). Here, we investigate whether a CNN-based approach can further enhance retrieval performance and better capture intraday variability.

This study aims to explore the potential of deep learning to improve sub-daily SM from ASCAT observations over the contiguous United States (CONUS). Specifically, we develop a localized convolutional neural network (CNN-l) framework designed to exploit spatial patterns and adapt to regional conditions, thereby reducing systematic biases.

The remainder of this study is organized as follows. Section 2 introduces the datasets employed in this study. Section 3 details the model architecture and evaluation metrics for sub-daily retrievals. Section 4 presents the model assessment, and section 5 discusses its ability to capture intraday SM dynamics. Finally, section 6 summarizes the findings and discusses implications for future applications.

2 Datasets

The core database used in this study consists of spatiotemporal coincidences of ASCAT level-2 observations and ERA5 reanalysis data (i.e., the fifth generation ECMWF (European Centre for Medium-term Weather Forecast) reanalysis (Hersbach et al., 2023)). Both datasets are projected onto a common 0.25° resolution grid covering the study domain, the CONUS, which spans from 125°W to 70°W in longitude and from 25°N to 50°N in latitude, encompassing a climatically diverse region (Bernhardt et al., 2018). The study period includes four years, from 1 January 2016 to 31 December 2019.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

2.1 ASCAT information

ASCAT is a C-band (5.255 GHz) vertically polarized scatterometer known for its high radiometric accuracy (Figa-Saldaña et al., 2002). The ASCAT instruments are onboard three Metop-series polar-orbiting satellites, Metop-A, Metop-B, and Metop-C, launched on October 19^{th}, 2006; September 17^{th}, 2012; and November 7^{th}, 2018, respectively. These satellites operate in a low Earth, sun-synchronous polar orbit at an altitude of approximately 817 km, enabling near-global coverage every 12 hours. In our study area, ASCAT measurements are typically acquired during two daily time windows: from 00:00 to 05:00 UTC (ascending orbits) and from 14:00 to 19:00 UTC (descending orbits). The ASCAT backscattering coefficient has been used to retrieve SM, as it is strongly influenced by the soil's dielectric properties, which vary with moisture content (Srivastava et al., 2009; El Hajj et al., 2016; Wagner et al., 2013, 1999b).

We used the Metop ASCAT SSM CDR (H120 version 7), which provides a consistent data record of SM-related products at 12.5 km resolution (H SAF, 2021a), covering the entire duration of Metop satellite missions (2007-present). This dataset is derived using the latest version of the SM retrieval algorithm developed by TU Wien under the EUMETSAT H SAF program (H SAF, 2021b). Two key variables are used here:

Backscatter (σ_{40}) - The ASCAT instrument measures backscatter at various incidence angles. To facilitate intercomparison and reduce vegetation-related effects, the Metop ASCAT CDR provides backscatter normalized to a reference incidence angle of 40° (in dB). This normalized value, σ_{40}, is more consistent and less sensitive to surface heterogeneity compared to raw backscatter observations.

ASCAT surface soil moisture (SSM) - The SSM or SM product represents the relative water content of the top few centimeters of the soil (approximately the top 5 cm). The retrieval algorithm scales the normalized backscatter (σ_{40}) between predefined dry and wet reference values to estimate the relative SM (H SAF, 2021b). These relative SM values are expressed as degrees of saturation, ranging from 0% (completely dry) to 100% (fully saturated).

The Metop ASCAT SSM CDR provides SM and σ_{40} at the native temporal resolution of the satellite overpasses, without any temporal aggregation. The native spatial resolution of the ASCAT Level-1b product varies from 25 km (near swath) to 34 km (far swath). However, the CDR's Level-1c and Level-2 products are resampled on a discrete global grid with a uniform spatial sampling of 12.5 km at both latitude and longitude. To collocate ASCAT with ERA5 data, the CDR was regridded onto the ERA5 regular latitude-longitude grid at a 0.25° spatial resolution and an hourly temporal frequency.

To convert the relative SM values to volumetric SM ($m^3\ m^{-3}$), we multiplied the ASCAT SM values by soil porosity, following the approach of Saxton and Rawls (2006). Porosity estimates were obtained from the Global Land Data Assimilation System (GLDAS) dataset (Rodell et al., 2004), accessible at https://ldas.gsfc.nasa.gov/gldas/soils.

2.2 ERA5 database

In this study, we utilized the ERA5 reanalysis dataset to benefit from the comprehensive representation of the land-atmosphere system. ERA5 is based on the ECMWF operational numerical weather prediction system and incorporates a coupled land-atmosphere assimilation scheme, providing reliable datasets for our SM study.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

Soil moisture (SM) - ERA5 provides volumetric SM (m^3 m^{-3}) at 0-7 cm, 7-28 cm, 28-100 cm, and 100-289 cm depths. To ensure consistency with ASCAT satellite measurements, we focus on the topmost layer of 0-7 cm, commonly referred to as surface SM. This variable is used here as our primary target because it serves as a widely adopted benchmark in machine-learning–based SM retrieval development, owing to its spatial coherence and overall accuracy, consistent with established practice in previous studies (Aires et al., 2005; Rodríguez-Fernández et al., 2019; Aires et al., 2021b; Pellet et al., 2025; Dinh, 2025).

Soil temperature (ST) - ERA5 also offers ST at the same four depth intervals. As with ERA5 SM, we used the 0-7 cm layer to align with the surface sensitivity of the satellite observations. ST is considered an important auxiliary variable in SM retrieval, given its indirect relationship with soil water content through thermal properties (Campbell, 1985; Zhang et al., 2004).

Leaf area index (LAI) - Vegetation structure significantly influences ASCAT backscatter-incidence angle (Petchiappan et al., 2022). Additionally, many previous studies have shown the usefulness of vegetation indices in SM retrieval (Aires et al., 2021a; Han et al., 2023; Pellet et al., 2025). Among various vegetation indices—such as the normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI)—we selected the LAI. To maintain consistency with the objective of exploiting ASCAT-based SM information within the ECMWF assimilation framework, we used the ERA5 LAI product rather than externally sourced satellite datasets (Myneni et al., 2015; Yan et al., 2024).

Original ERA5 data are available at an hourly rate. Thus, they will be used directly for sub-daily SM retrieval.

2.3 *In situ* data from the International Soil Moisture Network

We used *in situ* SM data from the ISMN (Dorigo et al., 2013, 2021), which can be downloaded from https://ismn.earth/en/data/, to evaluate retrieval performance in 2019. The ISMN is a collaborative initiative supported by several international organizations and sustained through funding from the European Space Agency (ESA) and voluntary contributions from scientists and monitoring networks worldwide. It provides a consistent, globally accessible database of *in situ* SM measurements and serves as a key reference for validating both model-based and satellite-derived SM products (Kolassa et al., 2013; Batchu et al., 2023).

Within the study area, SM data from 568 *in situ* sites were extracted from three major networks: the Soil Climate Analysis Network (SCAN, Schaefer et al. (2007)); the SNOwpack TELemetry (SNOTEL, Leavesley et al. (2010)); and the United States Climate Reference Network (USCRN, Bell et al. (2013)). Although many stations report SM profiles at multiple depths (up to 2 m), only the topmost measurement—taken at 0 to 7 cm depth—was used in this study for validation to align with the depth-sensitivity of ASCAT and ERA5 data. Sites with more than 100 days of missing observations in 2019 were excluded from the analysis. The *in situ* measurements are reported at hourly intervals and expressed in volumetric units (m^3 m^{-3}).

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

3 Retrieval methods

3.1 Convolutional neural networks

CNNs, a class of deep learning models, were originally designed for tasks such as image and speech recognition (Lecun et al., 1998; Hinton et al., 2012). Due to their ability to efficiently process data with spatial or grid-like structures (i.e., images), CNNs have become increasingly popular in remote sensing applications, for either classification, segmentation, or retrieval (Maggiori et al., 2017; Rezaee et al., 2018; Aires et al., 2021a), see Ghanbari et al. (2021) for a meta-analysis.

CNNs operate on input data structured as multidimensional tensors. Here, input images are represented as a tensor $\mathbf{X}$ of dimensions (height) $\times$ (width) $\times$ (depth), where the height and width correspond here to 100 $\times$ 200 pixels, representing the spatial grid of latitude and longitude cells over the CONUS. The depth corresponds to the number of input channels, such as the backscattering coefficient σ_{40} and other auxiliary variables. A CNN transforms these inputs through a series of layers, typically composed of convolution operations and nonlinear activation functions.

The core component of a CNN is the convolutional layer, which applies learnable kernels or filters—small matrices of weights, typically of size of 3 $\times$ 3 $\times$ depth or 5 $\times$ 5 $\times$ depth—across the input tensor. Each kernel performs an element-wise multiplication followed by a summation, effectively extracting localized spatial features. The result of this operation is a feature map. Multiple kernels are used per layer to learn diverse features.

Activation functions are applied after each convolution to introduce nonlinearity, which is crucial for learning complex patterns. The Rectified Linear Unit (ReLu), defined as $f(x) = max(0, x)$, is often used due to its computational efficiency and ability to mitigate vanishing gradient issues (Nair and Hinton, 2010; Goodfellow et al., 2016).

In addition, several architectural parameters influence the operations and outputs of convolutional layers:

- *Stride* determines the number of pixels the kernel moves at each step. Larger strides reduce the output resolution.
- *Padding* adds extra pixels around the border of the input image. “Same” padding preserves input dimensions, while “valid” padding does not, resulting in a smaller output.

Local connection: To achieve finer spatial specialization, locally connected layers have been proposed (Chen et al., 2015) and adopted in remote sensing applications (Boucher et al., 2023; Pellet et al., 2025). Unlike conventional convolutional layers, they do not share weights across space; instead, each spatial region is assigned its own set of filters, allowing the network to learn location-specific patterns.

3.2 Proposed model architecture

We consider a localized CNN variant, incorporating a locally connected convolutional layer in which filters learn distinct weights for each spatial location. To preserve spatial dimensions, zero-padding is applied prior to a single 5 $\times$ 5 locally connected layer without weight sharing. This localized architecture has been shown to perform particularly well in extreme cases (Dinh, 2025), motivating its use here for the sub-daily estimates.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

The input to the model comprises ASCAT-derived σ_{40}, along with two auxiliary variables—ST and LAI—identified in section 2 as highly relevant to the target SM variable. While ST and LAI were selected for this study, other auxiliary variables may also contribute to SM retrieval, such as the antecedent precipitation evaporation index (Han et al., 2023) and the amplitude of the diurnal cycle of surface temperature (Prigent et al., 2005). However, the selection and ranking of such variables fall outside the scope of this work.

Our model is trained using the Adam optimizer (Kingma and Ba, 2017) with an initial learning rate of 0.001, and an epsilon value of 1 e^{-7} for numerical stability. Training runs for a maximum of 200 epochs with early stopping applied after 5 epochs if no improvement in validation loss, restoring the best weights. Weight initialization is controlled using the Glorot Uniform initialization with a fixed random seed to ensure reproducibility. A batch size of 16 is used in the training. Since CONUS-scale training made computationally intensive optimization impractical, these hyperparameters were selected using a structured manual tuning approach.

The dataset spanning 2016 to 2018 is partitioned into 80% for training and 20% for validation. Data from 2019 is held out for independent testing and performance evaluation. In addition, separate models are trained for ascending and descending satellite passes.

3.3 Evaluation metrics

To evaluate the performance of SM retrievals, we used four statistical evaluation metrics: Pearson's correlation coefficient (r, unitless), root mean square error (RMSE, $m^3\ m^{-3}$), and bias (Bias, $m^3\ m^{-3}$). These metrics are computed as follows:

$$r = \frac{\sum_{i=1}^{N}(y_{pred,i} - \overline{y_{pred}})(y_{ref,i} - \overline{y_{ref}})}{\sqrt{\sum_{i=1}^{N}(y_{pred,i} - \overline{y_{pred}})^2}\sqrt{\sum_{i=1}^{N}(y_{ref,i} - \overline{y_{ref}})^2}} \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_{pred,i} - y_{ref,i})^2}{N}} \tag{2}$$

$$Bias = \frac{\sum_{i=1}^{N}(y_{pred,i} - y_{ref,i})}{N}, \tag{3}$$

where $y_{pred,i}$ is the predicted SM, $y_{pred,i}$ is the reference data (i.e., ERA5 or *in situ* SMs), N is the number of samples of SM data, and $\overline{y_{pred}}$ is the mean value of the predicted SM data. The standard deviation (Std, $m^3\ m^{-3}$) of the difference between two datasets is also reported:

$$Std = \sqrt{\frac{\sum_{i=1}^{N}(d_i - \overline{d})^2}{N-1}} \tag{4}$$

where $d_i = y_{pred,i} - y_{ref,i}$ is the difference between the predicted and reference SMs, and $\overline{d} = \frac{1}{N}\sum_{i=1}^{N} d_i$.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

4 Retrieval model assessment

As detailed in section 3, we employed a localized CNN architecture that integrates three physically relevant inputs: σ_{40}, ST, and LAI. This configuration is denoted here as CNN-lo, where “l" indicates its localized setup and “o" refers to the orbit-based temporal alignment. The subsequent analysis assesses the performance of CNN-lo relative to ERA5 reanalysis and *in situ* measurements, and examines its improvements over the ASCAT H120 product.

All ASCAT overpasses (ascending and descending) were temporally collocated with the nearest hourly ERA5 reanalysis SM, hereafter referred to as ERA5-o. The corresponding ASCAT-based SM H120 product is denoted as H120-o. Likewise, *in situ* measurements (ISMN-o) were aligned to the closest ERA5-o timestamps to ensure consistent temporal reference across the datasets.

4.1 Evaluation against ERA5 SM

Figure 1 represents the ERA5-o and CNN-lo SM estimates and the difference between ASCAT-based retrievals (CNN-o and the H120 product) with respect to ERA5-o for multiple time slots on January 20^{th} 2019, over ascending passes. Although separate models were trained for ascending and descending orbits, both are analyzed to maximize sub-daily coverage. In this example, five ascending and five descending (not shown) passes covered the study area, with each grid cell typically receiving up to two observations per orbit type—still limited, but sufficient to explore intraday variability under certain conditions.

The CNN-derived SM estimates show strong spatial consistency with the ERA-o reference, effectively capturing similar patterns and value ranges across all observed orbits (Figure 1, first and second columns). This agreement is further reflected in the relatively small differences observed in the third column, where CNN-lo deviates minimally from ERA5-o. On the other hand, H120-o, which was not trained using ERA5 data, shows substantially larger discrepancies relative to ERA5-o, as indicated by the prevalence of dark red and blue regions in the fourth column. When assessed over the full set of available orbits for the year 2019, CNN-lo achieves high correlation with ERA5-o (r = 0.92 for ascending and r = 0.91 for descending passes), confirming its ability to capture SM dynamics at sub-daily scales. In contrast, the H120 product achieves significantly lower correlations (r = 0.58 and 0.59, respectively). It should be noted that H120 was not trained on ERA5 and is therefore more independent from ERA5 than CNN-lo. However, ERA5 provides the most reliable estimates of SM (Aires et al., 2001), achieving comparable results using only ASCAT data in our retrieval is a promising outcome. For a more independent assessment, we next compare all products against *in situ* data.

4.2 Evaluation against *in situ* measurements

We evaluated three SM estimates (ERA5-o, CNN-lo, and H120-o) against measurements from 568 ISMN stations across CONUS in 2019 (see section 2). Each ISMN station was collocated to the nearest 0.25° grid cell to match the resolution of the gridded SM estimates. Figure 2(a) presents probability density functions (PDFs) of temporal correlation (r), bias (Bias, $m^3\ m^{-3}$), and standard deviation (Std, $m^3\ m^{-3}$). ERA5-o achieves the highest overall agreement with *in situ* data (median r = 0.75, Bias = 0.056, and Std = 0.061), which is expected given that ERA5 assimilates a wide range of satellite observations

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

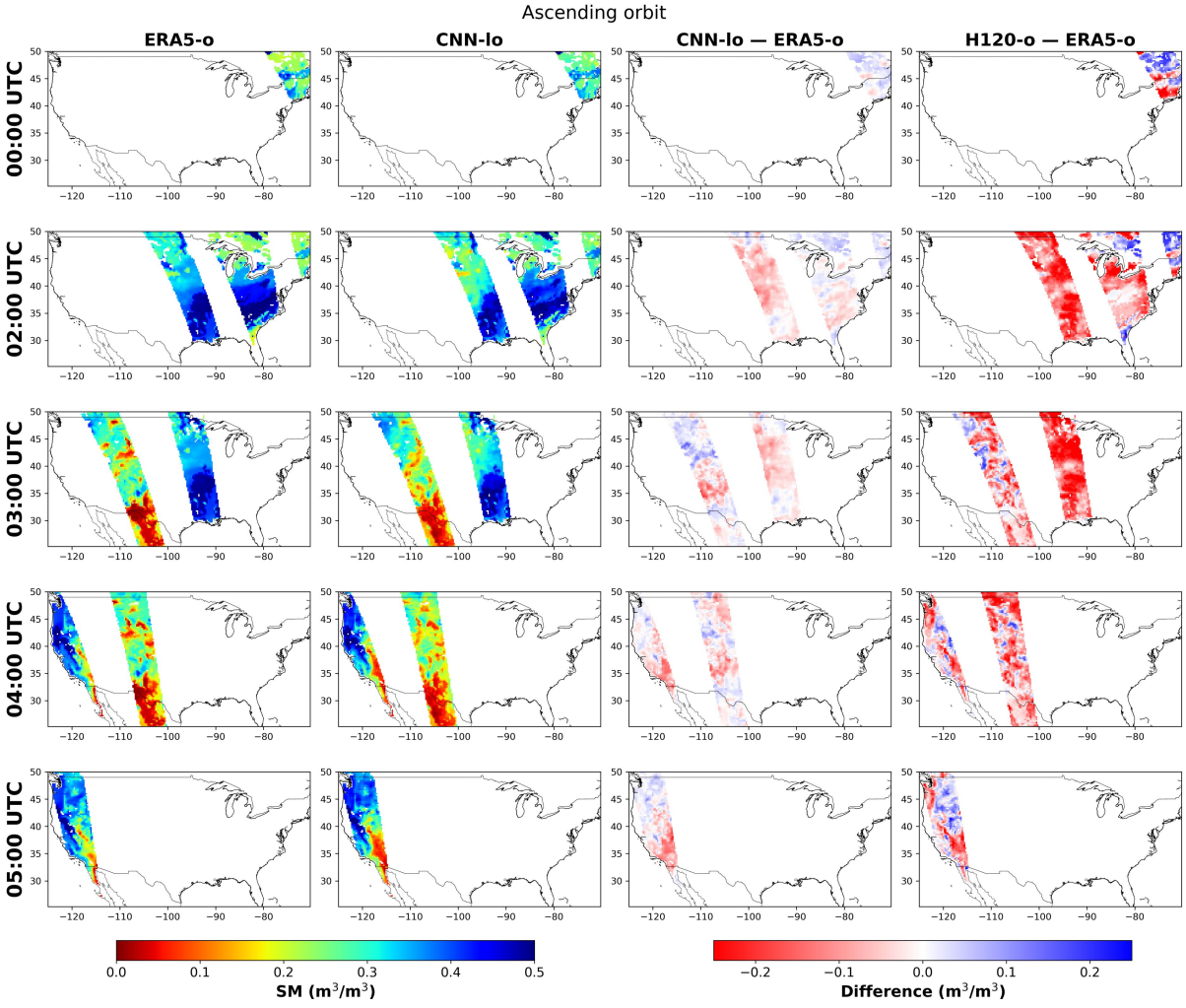

Figure 1. Hourly soil moisture estimates corresponding to five ascending ASCAT overpasses (00:00-05:00 UTC) on January 20^{th}, 2019. The first and second columns display SM values from ERA5-o and CNN-lo, respectively. The third column shows the SM differences between CNN-lo and ERA5-o, while the fourth column presents differences between H120-o and ERA5-o.

and *in situ* measurements. ERA reanalysis has long been shown to outperform SM accuracy compared to any individual satellite retrieval (see, for instance, (Aires et al., 2001)). The H120 product shows the lowest median correlation (r = 0.59) and exhibits a long negative tail in the bias distribution, despite having the lowest median bias error (0.024 m^3 m^{-3}). Its higher error variability (Std) indicates reduced skill in capturing SM dynamics. CNN-lo significantly outperforms H120, achieving a median correlation of 0.65 and error levels comparable to ERA5, as shown in the distributions in Figure 2(a).

The spatial distribution of correlation coefficients (Figure 2(b)) indicates that CNN-lo generally performs well, with particularly strong agreement in the western and southeastern regions. The difference map of Δ_r (middle panel) demonstrates that

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

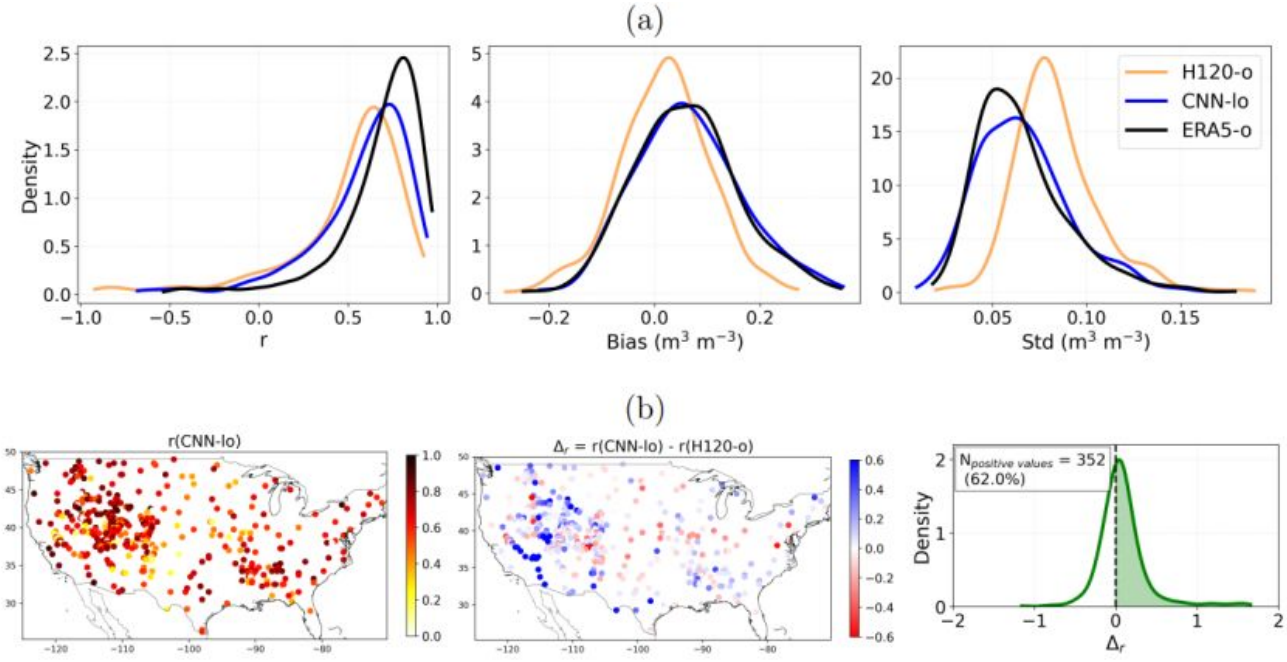

Figure 2. (a) Probability density functions (PDF) of temporal correlation (r), error bias (Bias, m^3 m^{-3}), and standard deviation of error (Std, m^3 m^{-3}) for ERA5-o, CNN-lo, and H120-o against *in situ* measurements at 568 sites across CONUS in 2019. (b) Maps showing (left) correlation of CNN-lo with *in situ* data and (middle) improvement in correlation with *in situ* data for CNN-lo relative to H120-o (Δ_r = r(CNN-lo) - r(H120-o)). Blue (red) indicates improved (degraded) correlation. The PDF of this Δ_r is also presented in the right panel.)

CNN-lo improves upon H120-o in a large part of the domain, as indicated by the predominance of blue points. Quantitatively, CNN-lo outperforms H120-o at 352 of 568 sites, about 62% (green area in the last panel of Figure 2(b)). The PDF of Δ_r further confirms this improvement, showing a pronounced positive tail approaching 2. This pattern reflects substantial performance gains at many sites where H120-o exhibits a low or even negative correlation. Although about 38% of the sites show a reduced correlation, the magnitude of these degradations is generally much smaller than the observed improvements.

To better illustrate the comparison at the site level, Figure 3 shows hourly SM time series for six randomly selected ISMN sites. All products capture the temporal dynamics and amplitude of SM reasonably well. ERA5-o consistently achieves the highest correlations (> 0.8), reflecting the advantage of assimilating multiple data sources. CNN-lo closely tracks the ERA5 SM ranges and temporal patterns. The H120 product exhibits higher temporal variability than the other retrievals.

5 Analysis of intraday variability

The previous section demonstrated that CNN-lo outperforms the ASCAT H120 product, exhibiting strong agreement with both ERA5 and *in situ* measurements. Here, we investigate the capability of this NN-based approach to capture diurnal SM varia-

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

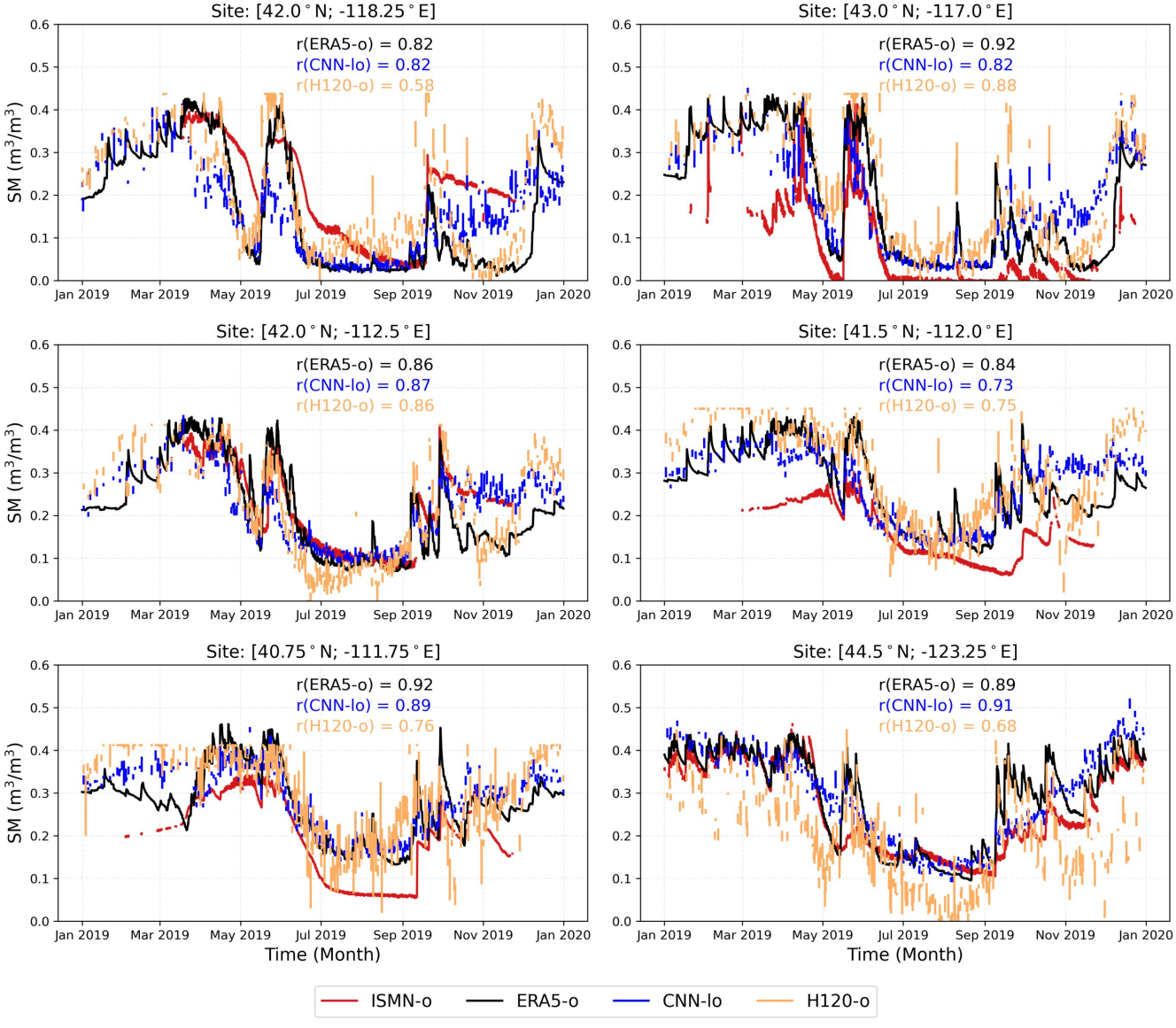

Figure 3. Hourly soil moisture (SM) time series in 2019 at six sites, comparing *in situ* data (ISMN-o) with ERA5-o, CNN-lo, and ASCAT H120 product (H120-o). Each panel includes the temporal correlation coefficient (r) between each SM estimate and *in situ* measurements.

tions, despite the inherent limitation in the satellite sampling frequency. Since only a few satellite observations are available per pixel per day (e.g., Figure 1), extracting a reliable diurnal signal is a true challenge!

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

5.1 Analysis of one case study

To illustrate the capacity of CNN-lo in capturing short-term SM dynamics, we analyzed a day with a significant precipitation event: November 20^{th}, 2019. Figure 4(a) displays spatial maps of daily SM amplitude (Δ_{SM}) derived from CNN-lo and ERA5-o. For reference, the full 24-hour SM amplitude of ERA5 (ERA5-24h) and the corresponding daily accumulated precipitation are also represented. High Δ_{SM} values appear in the southwest region, coinciding with a precipitation event captured in the ERA5 precipitation data.

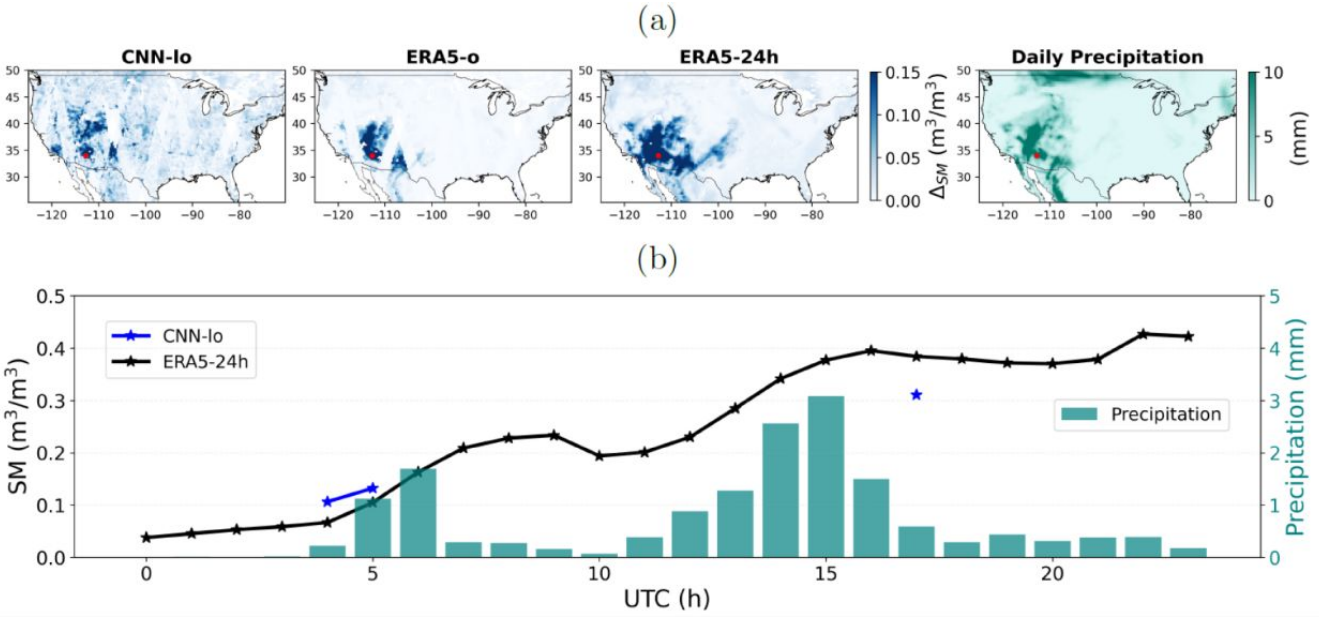

Figure 4. (a) Maps of daily SM amplitude (Δ_{SM}) on November 20^{th}, 2019, derived from CNN-lo and ERA5-o. ERA5-24h (full 24-hour amplitude) and daily accumulated precipitation are also shown for reference. (b) Hourly SM time series at [34°N, 112.75°W] (marked with a red dot in the maps), comparing CNN-lo retrievals to ERA5-24h. CNN-lo estimates correspond to three ASCAT overpasses: two ascending (04:00–05:00 UTC) and one descending (17:00 UTC). The precipitation is also presented.

We then examined the hourly SM time series in a representative grid cell located at [34°N, 112.75°W], shown in red in Figure 4(b). According to ERA5-24h, the rain started around 04:00 UTC and continued throughout the day. Prior to rainfall (00:00–04:00 UTC), ERA5 SM values remained low, then gradually increased, reaching approximately 0.42 m^3 m^{-3} by the end of the day.

On this day, ASCAT provided three overpasses: two during the ascending orbit (04:00–05:00 UTC) and one during the descending orbit (17:00 UTC). CNN-lo retrieved SM estimates for these overpasses, showing good agreement with ERA5. The CNN-lo daily SM amplitude ($\Delta_{SM} = 0.2$ m^3 m^{-3}) is significant but lower than ERA5-24h (0.31 m^3 m^{-3}). The errors are due to 1) retrieval uncertainties and 2) limitations in temporal sampling.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

This particular case study demonstrates the model's ability to capture intraday SM dynamics during heavy precipitation events, despite sparse observations. Such information is critical for hydrological applications, as rapid soil saturation can provoke flood risks. However, the spatial depiction of intraday variability remains strongly constrained by ASCAT's sampling frequency. Integrating observations from multiple instruments (e.g., SMOS, SMAP) is expected to enhance both retrieval accuracy and temporal resolution.

5.2 Δ_{SM} sensitivity to precipitation level

The key question is: What level of intraday SM variation can be reliably detected from NN-based retrieval? To address this, we compared the diurnal SM amplitude (Δ_{SM}) from CNN-lo with ERA5-o for all pixels and days in 2019 having at least two ASCAT observations. Results are shown in Figure 5.

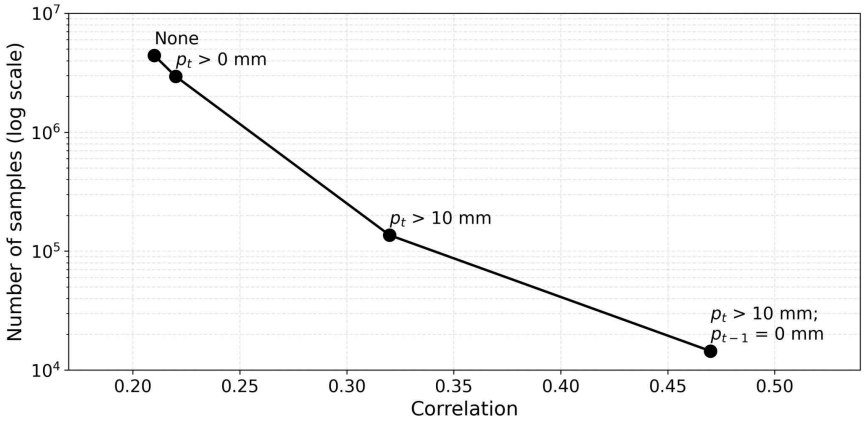

Figure 5. Correlation between diurnal SM amplitude (Δ_{SM}) estimated by CNN-lo and the ERA5-o reference as a function of daily precipitation thresholds (p_t in mm, where t denotes the day index). "None" indicates no threshold applied. The x-axis presents the correlation values, and the y-axis shows the corresponding number of samples on a logarithmic scale for each threshold condition.

When considering all 4,419,791 samples (i.e., threshold is none), the correlation between CNN-lo and ERA5-o Δ_{SM} is modest (r = 0.21), indicating that most observed differences reflect retrieval noise rather than true diurnal variability. This is expected given ASCAT's sparse temporal sampling and the inherent difficulty of resolving small sub-daily changes. However, as shown in section 55.1, to isolate such cases, we progressively applied precipitation thresholds. For days with any measurable rainfall ($p_t > 0$ mm), correlation improves slightly to 0.22. For heavier events ($p_t > 10$ mm), correlation rises to 0.32, and further increases to 0.47 when excluding cases with rain on the previous day ($p_{t-1} = 0$ mm). This subset (14,445 samples) primarily reflects isolated heavy rainfall events, suggesting that CNN-lo can capture SM responses when the signal-to-noise ratio is favorable.

Figure 6 shows the scatter plots for this subset ($p_t > 10$ mm, $p_{t-1} = 0$ mm), comparing CNN-lo and ERA5-o for: (1) maximum SM value (SM_{max}), (2) minimum SM value (SM_{min}), and (3) diurnal SM amplitude (Δ_{SM}). Each point represents

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

one single pixel-day. CNN-lo tends to underestimate SM_{max} and slightly overestimate SM_{min}—a typical behavior of statistical models that dampen extremes (Hastie et al., 2009). Despite this, strong agreement with ERA5-o is achieved for SM_{max} and SM_{min} (r = 0.82 and r = 0.88, respectively). As expected, Δ_{SM} exhibits lower correlation (r = 0.47) because amplitude errors propagate from both extremes. The third panel of Figure 6 indicates that the regression fit (dashed red line) is skewed by high-amplitude cases. However, the denser cluster of points in the 0–0.05 m^3 m^{-3} range (i.e., yellow points) suggests that the model captures stable SM conditions well. These low-amplitude cases likely correspond to days with minimal diurnal variation, where the model performs more reliably.

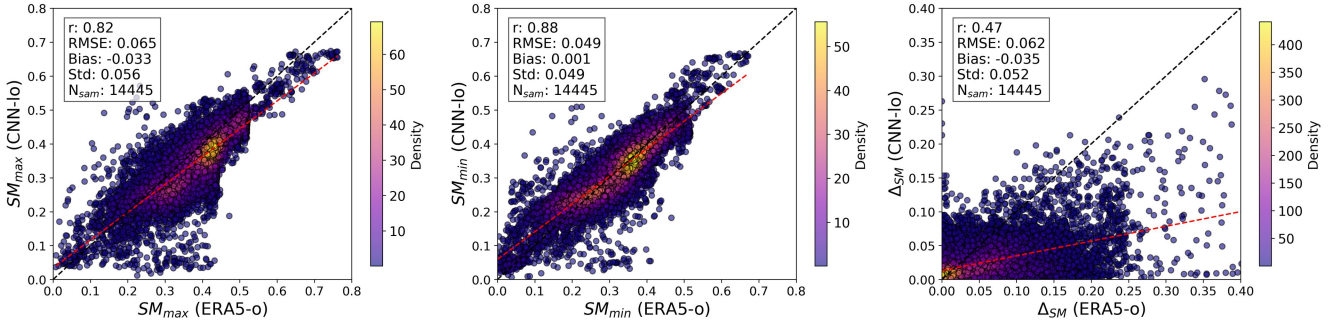

Figure 6. Scatter plots comparing CNN-lo and ERA5-o SM for days with isolated precipitation events ($p_t > 10$ mm and $p_{t-1} = 0$ mm) in 2019. Panels show: (1) maximum SM value (SM_{max}); (2) minimum SM value (SM_{min}); and (3) the resulting diurnal SM amplitude (Δ_{SM}=SM_{max}-SM_{min}), all in m^3 m^{-3}. Each point represents one pixel-day. Statistical metrics include total correlation (r), root mean square error (RMSE, m^3 m^{-3}), bias (Bias, m^3 m^{-3}), standard deviation (Std, m^3 m^{-3}), and number of samples ($N_{samples}$). The dashed black line represents the 1:1 line; the red dashed line is the linear regression fit. The colorbar indicates the local density of dots in the scatterplot.

5.3 Evaluation of intraday SM variations

We now examine the spatial and temporal structure of these intraday variations over the CONUS. Figure 7 summarizes key intraday SM signals for the 14,445 cases identified above: timing of SM_{max} and SM_{min}, their respective values, and the diurnal amplitude Δ_{SM} for CNN-lo, ERA5-o, and ERA5-24h. Note that higher precipitation is concentrated in the eastern CONUS, resulting in limited sampling in western regions (white pixels).

CNN-lo and ERA5-o exhibit similar spatiotemporal patterns. Note, however, that the estimation of the SM maximum and minimum times is degraded when limited by the ASCAT time sampling. For CNN-lo and ERA5-o, in the eastern part, SM typically reaches its maximum in the late afternoon (around 17:00 UTC) and its minimum in the early morning (around 03:00 UTC), aligning with ASCAT overpass times. In contrast, ERA5-24h—benefiting from full hourly coverage—shows SM_{max} occurring later in the evening and SM_{min} earlier, indicating a more complete description of the diurnal cycle.

Maps of SM_{max} and SM_{min} are largely consistent across CNN-lo and ERA-o, supported by the high correlation reported in Figure 6. However, CNN-lo underestimates Δ_{SM} compared to ERA5-o, consistent with the tendency of statistical models to dampen variability (Hastie et al., 2009; Skafte et al., 2019). Interestingly, ERA5-o also exhibits lower amplitudes relative to

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

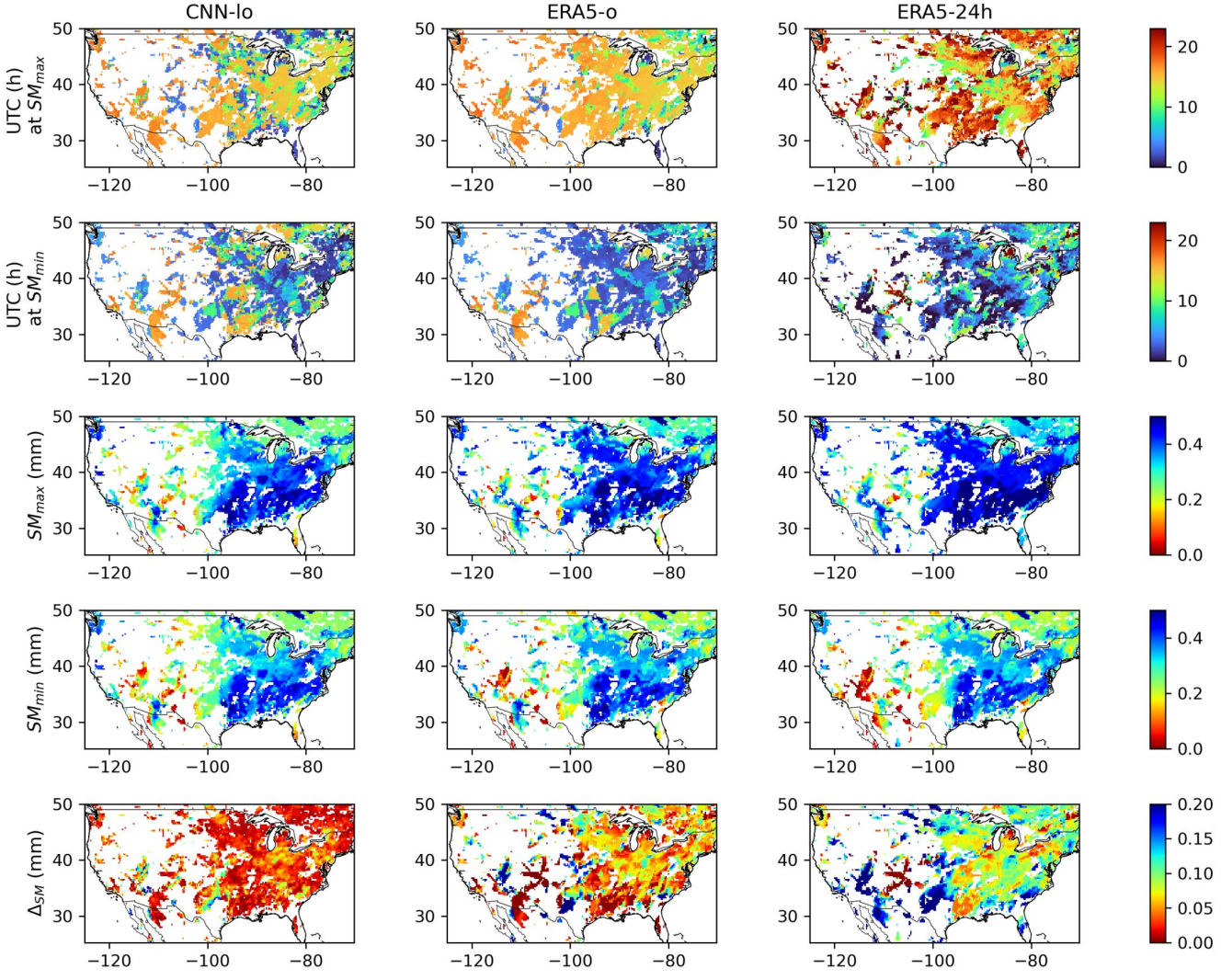

Figure 7. From top to bottom: (1) Time of maximum SM in UTC (hours); (2) time of minimum SM in UTC (hours); (3) maximum SM value (SM_{max}, m^3 m^{-3}); (4) minimum SM value (SM_{min}, m^3 m^{-3}); and (5) diurnal SM amplitude (Δ_{SM}, m^3 m^{-3}). Columns represent CNN-lo (left), ERA5-o (middle), and ERA5-24h (right). Only days in 2019 with total daily precipitation exceeding 10 mm and no rainfall on the preceding day are included (i.e., 14,445 selected cases). Each pixel reflects the average signal computed over all available days.

ERA5-24h, which has a broader dynamic range due to full diurnal coverage. Therefore, the underestimation by CNN-lo likely reflects both retrieval smoothing and ASCAT's temporal sampling limitations—a constraint that could be mitigated through multi-sensor data fusion, as discussed in section 6.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

6 Conclusion and perspectives

This study demonstrates that a localized CNN can improve SM retrieval from ASCAT observations. The CNN approach achieves correlations exceeding 0.9 with ERA5 and a median temporal correlation of 0.65 against *in situ* measurements—outperforming the ASCAT H120 product (0.58 and 0.59, respectively). Importantly, we show that intraday SM signals are detectable under specific conditions, particularly when the diurnal amplitude (Δ_{SM}) exceeds retrieval uncertainty. Such favorable signal-to-noise conditions typically occur during intense precipitation events over unsaturated soils. These findings reveal a pathway toward sub-daily SM monitoring, with direct value for hydrological forecasting, agricultural management, and early-warning systems (Brocca et al., 2010; Vilà-Guerau de Arellano et al., 2020; Vermunt et al., 2022).

Despite this promise, significant challenges remain. Firstly, like most statistical models, CNN retrievals tend to attenuate extreme values, leading to underestimation of Δ_{SM} (Hastie et al., 2009). Future work should explore architectures better suited for nonlinear dynamics, potentially through enhanced localization strategies or input feature augmentation (Boucher and Aires, 2023). For instance, including soil texture, precipitation-related data, or other auxiliary variables may further improve retrieval skill, given their strong relationship with SM (Abbaszadeh et al., 2019). Additionally, integrating dynamic components (O and Orth, 2020) into the retrieval framework could enable better representation of temporal dependencies, thereby improving overall model accuracy at a sub-daily scale.

Secondly, the temporal sampling of satellite observations remains a major limitation. In our example of ASCAT observations, with typically two overpasses per day per orbit type, key intraday dynamics are easily missed. Multi-sensor integration offers one possible solution, but introduces added complexity, including inter-sensor calibration, heterogeneous retrieval uncertainties, and the need for temporal regularization (Samadzadegan et al., 2025). Conditioning retrievals on external information, such as precipitation or other ancillary information, could also improve sensitivity to rapid SM changes. Advances in machine learning, such as foundation models, may facilitate these developments by enabling integration of heterogeneous data sources, dynamic downscaling, and spatiotemporal consistency (Hong et al., 2024; Yu et al., 2025).

Beyond SM, the proposed CNN framework has broader applicability for retrieving geophysical variables characterized by spatial structure and incomplete radiative transfer constraints, such as land surface temperature, surface emissivity, vegetation cover, and surface water extent. Moreover, these methods could extend to forward modeling, such as predicting satellite brightness temperatures directly from geophysical states. By leveraging spatial patterns and localized adaptation, deep learning offers a transformative opportunity to increase the temporal and spatial fidelity of remote sensing products. This temporal and spatial downscaling opens new opportunities for real-time environmental monitoring at scales previously unattainable.

Data availability. The ASCAT data supporting this study can be obtained from the Metop ASCAT SSM CDR H SAF (2021a), https://hsaf.meteoam.it/ (last access: 22 Jan 2025). Porosity data Rodell et al. (2004) are from https://ldas.gsfc.nasa.gov/gldas/soils (last access: 31 May 2025). The ERA5 reanalysis dataset can be downloaded from https://cds.climate.copernicus.eu Hersbach et al. (2023) (last access: 15 Mar 2025). Finally, the International Soil Moisture Network data are available at https://ismn.earth/en/data/ Dorigo et al. (2021) (last access: 3 Mar 2025).

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

Author contributions. All authors conceptualized this research. LAD and VP collected the data and prepared the programming and computational environment. LAD and FA conducted the formal analysis and contributed to methodology development. LAD implemented the models and carried out the validation and visualization. FA supervised the project and obtained the funding. LAD and FA wrote the paper. All authors revised the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The CERISE project (grant agreement N° 101082139) is funded by the European Union. Views and opinions expressed are, however, those of the author(s) only and do not necessarily reflect those of the European Union or the Commission. Neither the European Union nor the granting authority can be held responsible for them. We thank Patricia de Rosnay and Peter Weston (ECMWF) for valuable discussions.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

References

Abbaszadeh, P., Moradkhani, H., and Zhan, X.: Downscaling SMAP Radiometer Soil Moisture Over the CONUS Using an Ensemble Learning Method, Water Resources Research, 55, 324–344, https://doi.org/https://doi.org/10.1029/2018WR023354, 2019.

Aires, F., Prigent, C., Rossow, W. B., and Rothstein, M.: A new neural network approach including first guess for retrieval of atmospheric water vapor, cloud liquid water path, surface temperature, and emissivities over land from satellite microwave observations, Journal of Geophysical Research: Atmospheres, 106, 14 887–14 907, https://doi.org/https://doi.org/10.1029/2001JD900085, 2001.

Aires, F., Prigent, C., and Rossow, W. B.: Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: 2. Global statistical relationships, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/https://doi.org/10.1029/2004JD005094, 2005.

Aires, F., Boucher, E., and Pellet, V.: Convolutional neural networks for satellite remote sensing at coarse resolution. Application for the SST retrieval using IASI, Remote Sensing of Environment, 263, 112 553, https://doi.org/https://doi.org/10.1016/j.rse.2021.112553, 2021a.

Aires, F., Weston, P., de Rosnay, P., and Fairbairn, D.: Statistical approaches to assimilate ASCAT soil moisture information—I. Methodologies and first assessment, Quarterly Journal of the Royal Meteorological Society, 147, 1823–1852, https://doi.org/https://doi.org/10.1002/qj.3997, 2021b.

Araya, S. N., Fryjoff-Hung, A., Anderson, A., Viers, J. H., and Ghezzehei, T. A.: Advances in soil moisture retrieval from multispectral remote sensing using unoccupied aircraft systems and machine learning techniques, Hydrology and Earth System Sciences, 25, 2739–2758, https://doi.org/10.5194/hess-25-2739-2021, 2021.

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., and Anderson, C.: Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), Geophysical Research Letters, 34, https://doi.org/https://doi.org/10.1029/2007GL031088, 2007.

Batchu, V., Nearing, G., and Gulshan, V.: A Deep Learning Data Fusion Model Using Sentinel-1/2, SoilGrids, SMAP, and GLDAS for Soil Moisture Retrieval, Journal of Hydrometeorology, 24, 1789–1823, https://journals.ametsoc.org/view/journals/hydr/24/10/JHM-D-22-0118.1.xml, 2023.

Bateni, S. M. and Entekhabi, D.: Relative efficiency of land surface energy balance components, Water Resources Research, 48, https://doi.org/https://doi.org/10.1029/2011WR011357, 2012.

Bell, J. E., Palecki, M. A., Baker, C. B., Collins, W. G., Lawrimore, J. H., Leeper, R. D., Hall, M. E., Kochendorfer, J., Meyers, T. P., Wilson, T., and Diamond, H. J.: U.S. Climate Reference Network Soil Moisture and Temperature Observations, Journal of Hydrometeorology, 14, 977–988, https://journals.ametsoc.org/view/journals/hydr/14/3/jhm-d-12-0146_1.xml, 2013.

Bernhardt, J., Carleton, A. M., and LaMagna, C.: A Comparison of Daily Temperature-Averaging Methods: Spatial Variability and Recent Change for the CONUS, Journal of Climate, 31, 979 – 996, https://doi.org/10.1175/JCLI-D-17-0089.1, 2018.

Boucher, E. and Aires, F.: Improving remote sensing of extreme events with machine learning: land surface temperature retrievals from IASI observations, Environmental Research Letters, 18, 024 025, https://doi.org/10.1088/1748-9326/acb3e3, 2023.

Boucher, E., Aires, F., and Pellet, V.: Towards a new generation of artificial-intelligence-based infrared atmospheric sounding interferometer retrievals of surface temperature: Part I – Methodology, Quarterly Journal of the Royal Meteorological Society, 149, 1180–1196, https://doi.org/https://doi.org/10.1002/qj.4447, 2023.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrology and Earth System Sciences, 14, 1881–1893, https://doi.org/10.5194/hess-14-1881-2010, 2010.

Campbell, G. S.: Soil Physics with Basic: Transport Models for Soil-Plant Systems, Elsevier, Amsterdam, ISBN 978-0-444-42557-7, 1985.

Chen, Y., López-Moreno, I., Sainath, T. N., Visontai, M., Álvarez, R., and Parada, C.: Locally-connected and convolutional neural networks for small footprint speaker recognition, in: Interspeech, https://api.semanticscholar.org/CorpusID:6623788, 2015.

Dinh, L. A.: ASCAT soil moisture retrieval using deep learning: A focus on localization strategy, Front. Remote Sens., https://www.frontiersin.org/journals/remote-sensing/articles/10.3389/frsen.2025.1718353/, accepted, 2025.

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., Dostálová, A., Sanchis-Dufau, A. D., Zamojski, D., Cordes, C., Wagner, W., and Drusch, M.: Global Automated Quality Control of In Situ Soil Moisture Data from the International Soil Moisture Network, Vadose Zone Journal, 12, vzj2012.0097, https://doi.org/10.2136/vzj2012.0097, 2013.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215, https://doi.org/https://doi.org/10.1016/j.rse.2017.07.001, earth Observation of Essential Climate Variables, 2017.

Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W., Xaver, A., Annor, F., Ardö, J., et al.: The International Soil Moisture Network: serving Earth system science for over a decade, Hydrology and earth system sciences, 25, 5749–5804, Accessed 3 Mar 2025, 2021.

El Hajj, M., Baghdadi, N., Zribi, M., Belaud, G., Cheviron, B., Courault, D., and Charron, F.: Soil moisture retrieval over irrigated grassland using X-band SAR data, Remote Sensing of Environment, 176, 202–218, https://doi.org/https://doi.org/10.1016/j.rse.2016.01.027, 2016.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Spencer, M. W., Thurman, S. W., Tsang, L., and Van Zyl, J.: The Soil Moisture Active Passive (SMAP) Mission, Proceedings of the IEEE, 98, 704–716, https://doi.org/10.1109/JPROC.2010.2043918, 2010.

Figa-Saldaña, J., Wilson, J. J., Attema, E., Gelsthorpe, R., Drinkwater, M. R., and and, A. S.: The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers, Canadian Journal of Remote Sensing, 28, 404–412, https://doi.org/10.5589/m02-035, 2002.

Ghanbari, H., Mahdianpari, M., Homayouni, S., and Mohammadimanesh, F.: A Meta-Analysis of Convolutional Neural Networks for Remote Sensing Applications, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 3602–3613, https://doi.org/10.1109/JSTARS.2021.3065569, 2021.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, http://www.deeplearningbook.org, 2016.

H SAF: ASCAT Surface Soil Moisture Climate Data Record v7 12.5 km sampling - Metop, https://doi.org/10.15770/EUM_SAF_H_0009, Accessed 22 Jan 2025, 2021a.

H SAF: Algorithm Theoretical Baseline Document (ATBD) Metop ASCAT Surface Soil Moisture Climate Data Record v7 12.5 km sampling (H119) and Extension (H120), 2021b.

Han, Q., Zeng, Y., Zhang, L., Wang, C., Prikaziuk, E., Niu, Z., and Su, B.: Global long term daily 1 km surface soil moisture dataset with physics informed machine learning, Scientific Data, 10, 101, https://doi.org/10.1038/s41597-023-02011-7, 2023.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

Hastie, T., Tibshirani, R., and Friedman, J.: Neural Networks, pp. 389–416, Springer New York, New York, NY, ISBN 978-0-387-84858-7, https://doi.org/10.1007/978-0-387-84858-7_11, 2009.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, https://doi.org/10.24381/cds.adbb2d47, Accessed on 15 Mar 2025, 2023.

Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., and Kingsbury, B.: Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups, IEEE Signal Processing Magazine, 29, 82–97, https://doi.org/10.1109/MSP.2012.2205597, 2012.

Hong, D., Zhang, B., Li, X., Li, Y., Li, C., Yao, J., Yokoya, N., Li, H., Ghamisi, P., Jia, X., Plaza, A., Gamba, P., Benediktsson, J. A., and Chanussot, J.: SpectralGPT: Spectral Remote Sensing Foundation Model, IEEE Transactions on Pattern Analysis and Machine Intelligence, 46, 5227–5244, https://doi.org/10.1109/TPAMI.2024.3362475, 2024.

Kerr, Y. H., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., Juglea, S. E., Drinkwater, M. R., Hahne, A., Martín-Neira, M., and Mecklenburg, S.: The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle, Proceedings of the IEEE, 98, 666–687, https://doi.org/10.1109/JPROC.2010.2043032, 2010.

Kim, H., Lakshmi, V., Kwon, Y., and Kumar, S. V.: First attempt of global-scale assimilation of subdaily scale soil moisture estimates from CYGNSS and SMAP into a land surface model, Environmental Research Letters, 16, 074 041, https://doi.org/10.1088/1748-9326/ac0ddf, 2021.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, https://arxiv.org/abs/1412.6980, 2017.

Kolassa, J., Aires, F., Polcher, J., Prigent, C., Jimenez, C., and Pereira, J. M.: Soil moisture retrieval from multi-instrument observations: Information content analysis and retrieval methodology, Journal of Geophysical Research: Atmospheres, 118, 4847–4859, https://doi.org/https://doi.org/10.1029/2012JD018150, 2013.

Kolassa, J., Gentine, P., Prigent, C., Aires, F., and Alemohammad, S.: Soil moisture retrieval from AMSR-E and ASCAT microwave observation synergy. Part 2: Product evaluation, Remote Sensing of Environment, 195, 202–217, https://doi.org/https://doi.org/10.1016/j.rse.2017.04.020, 2017.

Leavesley, G. H., David, O., Garen, D. C., Nrcs-Usda, N., Goodbody, A. G., Lea, J. K., Marron, J. K., and Strobel, M.: A MODELING FRAMEWORK FOR IMPROVED AGRICULTURAL WATER-SUPPLY FORECASTING, https://api.semanticscholar.org/CorpusID:129416417, 2010.

Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P.: Gradient-based learning applied to document recognition, Proceedings of the IEEE, 86, 2278–2324, https://doi.org/10.1109/5.726791, 1998.

Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P.: Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification, IEEE Transactions on Geoscience and Remote Sensing, 55, 645–657, https://doi.org/10.1109/TGRS.2016.2612821, 2017.

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

Myneni, R., Knyazikhin, Y., and Park, T.: MCD15A3H MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006, https://doi.org/10.5067/MODIS/MCD15A3H.006, 2015.

Nair, V. and Hinton, G. E.: Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML'10, p. 807–814, Omnipress, Madison, WI, USA, ISBN 9781605589077, 2010.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

O, S. and Orth, R.: Global soil moisture from in-situ measurements using machine learning – SoMo.ml, https://arxiv.org/abs/2010.02374, 2020.

Ochsner, T. E., Cosh, M. H., Cuenca, R. H., Dorigo, W. A., Draper, C. S., Hagimoto, Y., Kerr, Y. H., Larson, K. M., Njoku, E. G., Small, E. E., and Zreda, M.: State of the Art in Large-Scale Soil Moisture Monitoring, Soil Science Society of America Journal, 77, 1888–1919, https://doi.org/https://doi.org/10.2136/sssaj2013.03.0093, 2013.

Pellet, V., Aires, F., Boucher, E., and Volden, E.: Enhancing Soil Moisture Statistical Retrieval from SMOS using Partial Convolutions and Localization Strategies, Journal of Applied Meteorology and Climatology, https://doi.org/10.1175/JAMC-D-25-0041.1, 2025.

Petchiappan, A., Steele-Dunne, S. C., Vreugdenhil, M., Hahn, S., Wagner, W., and Oliveira, R.: The influence of vegetation water dynamics on the ASCAT backscatter–incidence angle relationship in the Amazon, Hydrology and Earth System Sciences, 26, 2997–3019, https://doi.org/10.5194/hess-26-2997-2022, 2022.

Prigent, C., Aires, F., Rossow, W. B., and Robock, A.: Sensitivity of satellite microwave and infrared observations to soil moisture at a global scale: Relationship of satellite observations to in situ soil moisture measurements, Journal of Geophysical Research: Atmospheres, 110, https://doi.org/https://doi.org/10.1029/2004JD005087, 2005.

Rezaee, M., Mahdianpari, M., Zhang, Y., and Salehi, B.: Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11, 3030–3039, https://doi.org/10.1109/JSTARS.2018.2846178, 2018.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., and Toll, D.: The Global Land Data Assimilation System, Bulletin of the American Meteorological Society, 85, 381–394, https://journals.ametsoc.org/view/journals/bams/85/3/bams-85-3-381.xml, Accessed 31 May 2025, 2004.

Rodríguez-Fernández, N., de Rosnay, P., Albergel, C., Richaume, P., Aires, F., Prigent, C., and Kerr, Y.: SMOS Neural Network Soil Moisture Data Assimilation in a Land Surface Model and Atmospheric Impact, Remote Sensing, 11, https://doi.org/10.3390/rs11111334, 2019.

Rodríguez-Fernández, N. J., de Souza, V., Kerr, Y. H., Richaume, P., and Al Bitar, A.: Soil moisture retrieval using SMOS brightness temperatures and a neural network trained on in situ measurements, in: 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), pp. 1574–1577, https://doi.org/10.1109/IGARSS.2017.8127271, 2017.

Samadzadegan, F., Toosi, A., and Javan, F. D.: A critical review on multi-sensor and multi-platform remote sensing data fusion approaches: current status and prospects, International Journal of Remote Sensing, 46, 1327–1402, https://doi.org/10.1080/01431161.2024.2429784, 2025.

Saxton, K. E. and Rawls, W. J.: Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions, Soil Science Society of America Journal, 70, 1569–1578, https://doi.org/https://doi.org/10.2136/sssaj2005.0117, 2006.

Schaefer, G. L., Cosh, M. H., and Jackson, T. J.: The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN), Journal of Atmospheric and Oceanic Technology, 24, 2073–2077, https://journals.ametsoc.org/view/journals/atot/24/12/2007jtecha930_1.xml, 2007.

Singh, A. and Gaurav, K.: Deep learning and data fusion to estimate surface soil moisture from multi-sensor satellite images, Scientific Reports, 13, 2251, https://doi.org/10.1038/s41598-023-28939-9, 2023.

Skafte, N., Jø rgensen, M., and Hauberg, S. r.: Reliable training and estimation of variance networks, in: Advances in Neural Information Processing Systems, edited by Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R., vol. 32, Curran Associates, Inc., https://proceedings.neurips.cc/paper_files/paper/2019/file/07211688a0869d995947a8fb11b215d6-Paper.pdf, 2019.

https://doi.org/10.5194/egusphere-2026-2360
Preprint. Discussion started: 28 April 2026

Srivastava, H. S., Patel, P., Sharma, Y., and Navalgund, R. R.: Large-Area Soil Moisture Estimation Using Multi-Incidence-Angle RADARSAT-1 SAR Data, IEEE Transactions on Geoscience and Remote Sensing, 47, 2528–2535, https://doi.org/10.1109/TGRS.2009.2018448, 2009.

Trenberth, K. E., Smith, L., Qian, T., Dai, A., and Fasullo, J.: Estimates of the Global Water Budget and Its Annual Cycle Using Observational and Model Data, Journal of Hydrometeorology, 8, 758–769, https://doi.org/10.1175/JHM600.1, 2007.

Vermunt, P. C., Steele-Dunne, S. C., Khabbazan, S., Judge, J., and van de Giesen, N. C.: Extrapolating continuous vegetation water content to understand sub-daily backscatter variations, Hydrology and Earth System Sciences, 26, 1223–1241, https://doi.org/10.5194/hess-26-1223-2022, 2022.

Vilà-Guerau de Arellano, J., Ney, P., Hartogensis, O., de Boer, H., van Diepen, K., Emin, D., de Groot, G., Klosterhalfen, A., Langensiepen, M., Matveeva, M., Miranda-García, G., Moene, A. F., Rascher, U., Röckmann, T., Adnew, G., Brüggemann, N., Rothfuss, Y., and Graf, A.: CloudRoots: integration of advanced instrumental techniques and process modelling of sub-hourly and sub-kilometre land–atmosphere interactions, Biogeosciences, 17, 4375–4404, https://doi.org/10.5194/bg-17-4375-2020, 2020.

Wagner, W., Lemoine, G., and Rott, H.: A Method for Estimating Soil Moisture from ERS Scatterometer and Soil Data, Remote Sensing of Environment, 70, 191–207, https://doi.org/https://doi.org/10.1016/S0034-4257(99)00036-X, 1999a.

Wagner, W., Noll, J., Borgeaud, M., and Rott, H.: Monitoring soil moisture over the Canadian Prairies with the ERS scatterometer, IEEE Transactions on Geoscience and Remote Sensing, 37, 206–216, https://doi.org/10.1109/36.739155, 1999b.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., and Steinnocher, K.: The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications, Meteorologische Zeitschrift, 22, 5–33, https://doi.org/10.1127/0941-2948/2013/0399, 2013.

Wang, J., Zhang, Y., Song, P., and Tian, J.: Estimating sub-daily resolution soil moisture using Fengyun satellite data and machine learning, Journal of Hydrology, 632, 130 814, https://doi.org/https://doi.org/10.1016/j.jhydrol.2024.130814, 2024.

Yan, K., Wang, J., Peng, R., Yang, K., Chen, X., Yin, G., Dong, J., Weiss, M., Pu, J., and Myneni, R. B.: HiQ-LAI: a high-quality reprocessed MODIS leaf area index dataset with better spatiotemporal consistency from 2000 to 2022, Earth System Science Data, 16, 1601–1622, https://doi.org/10.5194/essd-16-1601-2024, 2024.

Yao, P., Lu, H., Shi, J., Zhao, T., Yang, K., Cosh, M. H., Gianotti, D. J. S., and Entekhabi, D.: A long term global daily soil moisture dataset derived from AMSR-E and AMSR2 (2002–2019), Scientific Data, 8, 143, https://doi.org/10.1038/s41597-021-00925-8, 2021.

Yu, Z., Idris, M. Y. I., Wang, H., Wang, P., Chen, J., and Wang, K.: From Physics to Foundation Models: A Review of AI-Driven Quantitative Remote Sensing Inversion, https://arxiv.org/abs/2507.09081, 2025.

Zhang, S.-W., Qiu, C.-J., and Xu, Q.: Estimating Soil Water Contents from Soil Temperature Measurements by Using an Adaptive Kalman Filter, Journal of Applied Meteorology, 43, 379–389, https://journals.ametsoc.org/view/journals/apme/43/2/1520-0450_2004_043_0379_eswcfs_2.0.co_2.xml, 2004.