A global-scale intercomparison of Triple Collocation Analysis- and ground-based soil moisture time-variant errors derived from different rescaling techniques

Kai Wu [a, *], Dongryeol Ryu [b], Wolfgang Wagner [c], and Zhongmin Hu [a]

[a] College of Ecology and Environment, Hainan University, Haikou 570000, China

[b] Department of Infrastructure Engineering, The University of Melbourne, Parkville, Victoria 3010, Australia

[c] Department of Geodesy and Geoinformation, TU Wien, 1040 Vienna, Austria

Corresponding author: Kai Wu, kaiwu@hainanu.edu.cn

Abstract: Accurate specification of spatiotemporal errors of remotely sensed soil moisture (SM) data is essential for a correct assessment of their utility and optimally integrating multiple SM products or assimilating them into hydrological models. Although Triple Collocation Analysis (TCA) has been widely used to provide SM errors, the impact of rescaling technique on the TCA error estimates has not received major attention, which can lead to biased and inaccurate error estimates. Moreover, current knowledge about time-variant SM errors derived from TCA is still very limited, which hampers the advance of applying time-variant errors in data merging and data assimilation studies efficiently. Based on these considerations, this work aims to advance the use of the TCA for characterizing errors with a focus on the rescaling techniques, and validating TCA-based time-variant errors using global ground measurements in 759 grid cells. To this end, the Advanced Scatterometer (ASCAT) and four passive-based SM products, including Soil Moisture and Ocean Salinity Level-3 (SMOSL3), SMOS INRA-CESBIO (SMOSIC), Soil Moisture Active Passive Level-3 (SMAPL3), and SMAP INRAE BORDEAUX (SMAPIB) SM products were considered. The time-variant error term here denotes an aggregate error magnitude over a 101-day moving-time-window. It is found that different selection of the rescaling technique considered in TCA led to TCA error estimates with significantly different accuracy when ground-based errors are regarded as the benchmark. The optimal combination strategy to implement TCA is applying TCA to SM anomalies and rescaling the errors by coefficients derived from the TCA model. Pearson's correlation with ground-based time-variant errors is 0.62, 0.72, 0.83, 0.89, and 0.93 for SMAPIB, SMAPL3, SMOSIC, SMOSL3, and ASCAT SM, respectively. Considering time-variant errors in applications is necessary since time-variant errors deviate from time-invariant errors by 50% when errors are rescaled by TCA model parameters. Time-invariant errors are greater than time-variant errors when SM products are rescaled against a reference dataset while the opposite conclusion can be drawn when errors are rescaled by the TCA coefficients. TCA- and ground-based methods provide consistent evaluations in 74.7% (77.3%), 75.8% (79.8%), 79.6% (81.1%), and 78.6% (79.7%) of the analysis period on average (median) for the TCA implementations with SMAPL3, SMAPIB, SMOSL3, and SMOSIC SM, respectively. The error analysis reveals that TCA typically underestimated ASCAT errors while overestimated passive SM errors when considering ground-based evaluation as the benchmark. Moreover, TCA was found to have relatively less power to efficiently characterize SM errors in croplands when compared with other land cover types. This study validated TCA time-variant errors using ground measurements and compared TCA- and ground-based evaluation performances on a global scale. Our work arouses particular attention to the rescaling technique selection considered in TCA, which is crucial for accurately characterizing SM errors and efficiently using them in various hydrometeorological applications.

Keywords: soil moisture; rescaling technique; Triple Collocation Analysis; time-variant error

1 Introduction

Soil moisture (SM) plays an important role in modeling hydrological processes such as runoff and evapotranspiration, and links water, energy, and carbon cycles (Jackson, 1993; Houser et al., 1998; Western et al., 2002; Daly and Porporato, 2005). SM data can be applied in many disciplines such as drought monitoring, flood prediction, crop productivity forecasts, irrigation planning, and weather forecasting (Sims et al., 2002; Narasimhan and Srinivasan, 2005; Bolten et al., 2010; Wanders et al., 2014). SM observations can be obtained from ground measurements, hydrological modeling, and satellite observations, and each of them has its distinctive error characteristic. Learning about the error characteristics of various SM products is important as it significantly influences the uncertainties of the hydrological models driven by the SM content and therefore has a great impact on the interpretation of the model simulation. Furthermore, stochastic data assimilation relies on accurate error specifications for the observations and model predictions (Reichle, 2008). Data fusion studies also require accurate error specifications (Crow et al., 2015; Gruber et al., 2017) or signal-to-noise ratio information (Kim et al., 2022) to optimally integrate multiple SM products. An accurate specification of SM error variability in space (such as Gruber et al., 2015 and Wu et al., 2018) and in time (such as Loew and Schlenz, 2011; Zwieback et al., 2012; Su et al., 2014a; Wu et al., 2021) may lead to further improvements for data assimilation and data fusion studies.

Triple Collocation Analysis (TCA) is a popular evaluation method to provide relative errors for SM products derived from different platforms without requiring an absolute ‘truth’ (Dorigo et al., 2010). It was initially proposed by Stoffelen (1998) to solve the issue of error estimation for sea wind and wave height and was introduced to provide SM error estimates by Scipal et al. in 2008a. Besides, SM uncertainty can be obtained by a three-corned hat method, which was recently applied by Liu et al. (2021) to evaluate eleven SM products in the Qinghai-Tibet Plateau. Many studies have applied TCA to evaluate remotely sensed SM retrievals (such as Leroux et al., 2013; Su et al., 2014a; Su et al., 2014b; Chakravorty et al., 2016; Kim et al., 2018), model simulations (such as Dorigo et al., 2010; Al-Yaari et al., 2014), reanalysis SM (such as Scipal et al., 2008a; Scipal et al., 2010; Yilmaz and Crow 2014; Miyaoka et al., 2017), or ground-based SM measurements (Miralles et al., 2010; Chen et al., 2018; Wu et al., 2021). Currently, TCA-based SM error estimates have already been widely used in SM validation (Dorigo et al., 2010; Miyaoka et al., 2017; Kim et al., 2021), data assimilation (Gruber et al., 2019), and data fusion (Gruber et al., 2017; Peng et al., 2021) studies.

TCA has been typically applied to provide time-invariant errors in the whole investigation period. However, SM retrieval errors are known to vary with time due to vegetation phenology, changes in surface roughness, and variable environmental conditions (Ulaby et al., 1983; Famiglietti et al., 2008; Zwieback et al., 2018). To account for the time-variant feature of SM products, Loew and Schlenz (2011), Zwieback et al. (2012), Su et al. (2014a), and Wu et al. (2021) proposed various time-window-based TCA schemes to estimate time-variant errors at different timescales by relaxing the stationary assumption underlying the TCA. Several studies tried to consider and include time-variant error characteristics in applications. For example, Khan et al. (2014) proposed a multi-model data merging approach that considers monthly covariance matrix of the forecast errors, which was successfully applied to the sea surface temperature forecasts and achieved better performance than time-invariant forecasts. In addition, Kim et al. (2016) combined multiple satellite-based SM datasets by taking time-variant weights into account and reported this method outperforms the time-invariant approach. Therefore, an extension of TCA to temporal domain may greatly benefit many applications. Nevertheless, accuracy of the time-variant errors derived from window-based TCA scheme is not well known by the SM community, which should be investigated before applying time-variant SM errors in actual applications.

As increasing number of applications use TCA to characterize errors for large spatial and temporal datasets, the reliability of TCA has drawn attention from several fields, such as soil freeze/thaw (Li et al., 2022), root zone SM (Xu et al., 2021), and satellite-based surface albedo (Wu et al., 2019). However, more work is still needed to verify TCA-based SM spatiotemporal errors using ground measurements as such validation on a global scale has not been fully investigated in previous studies. We argue that this validation is important as it helps researchers learn about the accuracy and reliability of the TCA, and better interprets the TCA error estimates. Previous work on this topic had mostly a regional focus or did not consider time-variant errors. For example, Brocca et al. (2011) and Chen et al. (2016) reported that TCA- and ground-based estimates, i.e., error and satellite-versus-truth correlation coefficient, are strongly consistent in Europe and USA, respectively. By comparing the TCA- and ground-based errors, Loew and Schlenz (2011) investigated the point-to-area sampling error in Southern Germany. Notably, Kim et al. (2020) compared TCA-based fMSE (fractional Mean-Square-Error) with the conventional fMSE derived from ground measurements and showed these two kinds of methods yielded consistent evaluation results. Similarly, Zhang et al. (2021) investigated the TCA- and ground-based satellite-versus-truth correlation coefficients and reported that they yielded consistent spatial distribution. However, studies by Kim et al. (2020) and Zhang et al. (2021) only focus on the spatial domain, i.e., error metrics are assumed to be constant over time. The performance of TCA in the temporal domain remains unclear currently. Moreover, Yilmaz and Crow (2014) presented numerical and analytical comparisons for ground- and TCA-based SM errors and found that ground-based errors are often higher than TCA-based errors. A similar result was also confirmed by Dorigo et al. (2015) who pointed out that TCA errors were consistently lower than ground-based errors. Based on current studies, a global-scale intercomparison of TCA- and ground-based SM errors is highly needed to advance the TCA applications, especially for time-variant errors as they have great potential to improve the performance of data merging and data assimilation systems.

To accurately estimate time-variant SM errors using the TCA method, two key issues are typically ignored in current TCA studies, which may lead to inaccurate and biased SM error estimates. The first issue is the rescaling technique selection considered in TCA. The rescaling process is an essential step in TCA as it removes relative differences between the considered datasets and makes the resulting error estimates comparable for the statistical comparison purpose. The rescaling technique applied in TCA not only can minimize the impact of representativeness errors derived from different spatial resolution, spatiotemporal mis-alignment, and different vertical measurement support (Gruber et al., 2013; Chen et al., 2017; Molero et al., 2018) but also can implicitly compensate for different units used in the considered datasets (Gruber et al., 2020). The rescaling techniques applied in TCA can be divided into two categories: methods that rescale SM inputs against a selected reference dataset prior to TCA implementation (Scipal et al., 2008a), and methods that rescale errors by parameters derived from the TCA model after the TCA implementation (Gruber et al., 2016a). There are several statistical techniques to rescale SM observations prior to TCA, such as variance and mean matching (Dorigo et al., 2010; Miralles et al., 2010; Kim et al., 2020), normalization (Fascetti et al., 2016; Pierdicca et al., 2015; Wu et al., 2021), CDF (Cumulative Distribution Function) matching (Doubkova et al., 2012; Gruber et al., 2014; An et al., 2016; Zhuang et al., 2020), and linear regression method (Scipal et al., 2008a). Or normalizing all datasets into a common observation space prior to TCA by calculating their z-scores. Generally, both these two kinds of rescaling techniques can address the first-order (additive) biases for the considered datasets. However, Gruber et al. (2020) reported that the statistical matching cannot address the second-order (multiplicative) biases while the method that rescales errors using the TCA model parameters can. Most current studies ignored the impact of second-order (multiplicative) biases on validation and error characterization. Such impact of second-order (multiplicative) biases resulted from the rescaling technique selection on the TCA error characterization remain unclear and needs further investigation (Gruber et al., 2020). Nevertheless, selection of the rescaling technique has not received major attention in current TCA studies. Recently, both Kim et al. (2020) and Gruber et al. (2020) pointed out the necessity to investigate the impact of rescaling technique selection on the final TCA error estimates.

The other important issue that influences TCA is the selection of SM inputs. There are two options for the SM inputs. One uses the original SM measurements directly (such as Dorigo et al., 2015; Gruber et al., 2017; Wu et al., 2018; Wu et al., 2021) while the other one uses SM anomalies (such as Dorigo et al., 2010; Draper et al., 2013; Su et al., 2014a; Miralles et al., 2010; Chakravorty et al., 2016; Kim et al., 2020). In practice, a popular technique to obtain short-term anomalies relies on moving-window-based averages over the investigation period, carried out by Miralles et al. (2010), Crow et al. (2012), and Draper et al. (2013). The anomaly-based approach removes the seasonal effects underlying SM time-series that can artificially enhance the correlations between two SM time-series (Scipal et al., 2005; Scipal et al., 2008b) and therefore, reveals the ability of the SM products to capture the short-term events of drying and wetting (Dorigo et al., 2010; Al-Yaari et al., 2014). Working on the original time series also makes sense as it captures other properties of the data sets, even though the errors cross-correlations may be higher (Miralles et al., 2010; Draper et al., 2013). These two methods are complementary as both are needed to describe the quality of the SM data. However, no study to date makes a quantitative comparison for the final TCA error estimates derived from these two input selections. If TCA errors have small differences for the two methods, absolute-based approach is more favorable given its simplicity in practice.

Based on the considerations above, we first explored the impact of temporal interpolation on the accuracy of TCA error estimates in Section 3.1. Second, the impacts of rescaling technique and SM inputs selection on the accuracy of TCA time-invariant and time-variant errors were investigated in Section 3.2. Correlations and RMSE (Root Mean Squared Error) between TCA- and ground-based time-variant errors were also compared for the ASCAT and four passive-based SM data, i.e., SMOSL3, SMOSIC, SMAPL3, and SMAPIB SM products, in six land cover types. Third, the relative difference between time-invariant and time-variant errors was quantified in Section 3.3. Then, the evaluation consistency between TCA- and ground-based methods is explored in Section 3.4. Finally, time-variant errors were compared for ASCAT and the four passive SM products based on the TCA- and ground-based methods in Section 3.5, along with TCA- vs. ground-based evaluations over six land cover types.

2 Data and Methodology

This section introduces SM products and methods used. Sections 2.1 and 2.2 provide information about Soil Moisture and Ocean Salinity (SMOS), Soil Moisture Active Passive (SMAP), ASCAT SWI (Soil Water Index), Global Land Data Assimilation System Version 2.1 (GLDAS2), and ERA-Interim SM products, along with International Soil Moisture Network (ISMN) ground measurements. After briefly describing TCA and Quadruple Collocation Analysis (QCA) methods in Section 2.3, Section 2.4 details four rescaling techniques used in TCA implementation. Conventional ground-based method to estimate SM errors is presented in Section 2.5 and the approach to test the robustness of temporal interpolation applied in the TCA time-variant scheme is described in Section 2.6. Finally, Section 2.7 defines the evaluation metrics used in our work: Overall Relative Difference (ORD), Relative Difference (RD), RMSE, and Root Mean Squared Difference (RMSD).

2.1 SMOS, SMAP, ASCAT, GLDAS2, and ERA-Interim SM products

The Soil Moisture and Ocean Salinity (SMOS) mission is designed to observe SM content in the global landmass surface (0-5 cm) and salinity over the oceans based on multi-angular brightness temperature data observed by a radiometer operating at L-band (Kerr et al., 2012). It was launched by European Space Agency (ESA) in November 2009 and contributes to improving our understanding of the global water cycle and weather/seasonal climate forecasting. SMOS collects brightness temperature data at a local overpass time of 06:00 PM for descending pass and 06:00 AM for ascending pass, respectively. Here, the SMOS Level-3 (SMOSL3) SM product with version 3.3 was used (Al Bitar et al., 2017). This product was developed by CATDS-PDC (Centre Aval de Traitement des Données SMOS - Production & Dissemination Center) and provides SM observations projected on Equal-Area Scalable Earth Grid (Version 2) with a spatial resolution of 40 km.

Data screening for SMOSL3 SM was based on the DQX (data quality index) affiliated with the CATDS-PDC SMOSL3 product and SM observations were filtered out in our work when its associated DQX value is greater than 0.007 (Chen et al., 2018; Wang et al., 2021). The DQX value quantifies the error in SM observations in volumetric SM units (Al-Yaari et al., 2014). The RFI (Radio Frequency Interference) issue was not taken into consideration for SMOSL3 data screening as we expect more SM observations participated in the time-variant error estimation, especially in regions such as Eurasian.

Besides SMOSL3, SMOS INRA-CESBIO (SMOSIC) SM product of version 2 was also considered in our work. The SMOSIC SM product was constructed by INRA (Institute National de la Recherche Agronomique) and provides SM values derived from the SMOS L3 brightness temperature data with a spatial resolution of 25 km. One of the main characteristics of the SMOSIC SM is the maximal independence of auxiliary data during its retrieval process, which is different from the SMOSL3 SM that strongly relies on auxiliary data (Wigneron et al., 2021). The SMOSIC SM product was used here to increase the robustness of our TCA conclusions since the SMOSL3 SM product is not independent with model-based or reanalysis SM data and therefore may hinder the independence assumption required by the TCA and may have potential influence on the final intercomparison and validation. To guarantee the data quality of SMOSIC SM data, only SM values associated with scene flags $\leqslant 1$ were retained in the following analyses. The scene flags were affiliated with the SMOSIC product and can be used as an indicator to detect events that influence the SMOS L3 brightness temperature observations used in the retrieval process, such as strong topography, frozen condition, and water body contamination. Similar to the data filtering of the SMOSL3, the RFI problem was not considered for the SMOSIC SM in our analysis. Readers are recommended to refer to Wigneron et al. (2021) to find out more information about the SMOSIC SM product.

The SMOSL3 and SMOSIC SM datasets were reconstructed from local time-based into UTC time-based to match the ASCAT, GLDAS2, and ISMN ground measurements. This reconstruction was completed by considering the navigational time zones based on longitude values and local statutory deviation was not considered in such transformation (Kim et al., 2018). After the transformation, SM values at ascending and descending orbits were averaged when they are found on the same day. Consequently, SMOSL3 and SMOSIC SM daily datasets with UTC stamps were constructed and used here. Readers can find the conceptual map (Fig. S1) of UTC zones for converting the local time, i.e., ascending for 06:00 AM and descending for 06:00 PM, of SMOS dataset to match other SM data.

The Soil Moisture Active Passive (SMAP) satellite is designed by NASA (National Aeronautics and Space Administration) to provide global surface (0-5 cm) SM observations with a spatial resolution of 40 km (Entekhabi et al., 2010). Since a malfunction was found in the SMAP radar system, only the L-band radiometer onboard SMAP satellite collects observations at a local overpass time of 06:00 AM for descending pass and 06:00 PM for ascending pass, respectively (Wu et al., 2020). The SMAP Level-3 SM product in version 8 (SMAPL3) was used here (O'Neill et al., 2021). Begin with this version, the Dual Channel Algorithm (DCA) is applied as a new baseline algorithm in the retrieval process, which departs from prior versions that used the Single Channel Algorithm-Vertical Polarization (SCA-V) as the baseline algorithm. Data screening for the SMAPL3 SM product was implemented using the retrieval quality flag affiliated with the SMAPL3 product. Only SM observations with recommended quality were considered in our analysis. The quality flag indicates whether unfavorable environmental conditions, such as frozen soil, snow cover, flood, steeply sloped topography, urban area, and dense vegetation, occurred during the SM retrievals.

Besides SMAPL3, a recently developed SM product by INRAE Bordeaux (Li et al., 2022), i.e., SMAP INRAE BORDEAUX (SMAPIB) SM data, was also considered in our analysis. The SMAPIB SM product applied a new 2-Parameter retrieval algorithm, i.e., SM and vegetation optical depth, to the dual-polarized brightness temperature observations collected from SMAP satellite during its SM retrieval process. Currently, only SM observations acquired from the descending (06:00, local overpass time) orbits are available and can be freely accessed from https://ib.remote-sensing.inrae.fr/. Similar to the SMOSIC data, this SM product was used here as it does not use any modeled or reanalysis SM data as input in its retrieval algorithm, which makes it a favorable SM dataset to be applied to the TCA method that requires a strong independence assumption. As suggested by Li et al. (2022), only SM observations with scene flag value $\leqslant 1$ were kept in the following analyses as corresponding SM values are less affected by the frozen soil, strong topography, and water body contamination. To match with other SM data, the SMAPL3 and SMAPIB SM datasets were reconstructed from local time-based to UTC time-based using the same method applied in the SMOSL3 and SMOSIC data preprocessing.

The Advanced Scatterometer (ASCAT) is a C-band (5.2 GHz) active sensor onboard the MetOp satellite (Bartalis et al., 2007). The ASCAT SWI (Soil Water Index) product (Wagner et al., 1999) with T=1 was used here, which describes SM content in the top soil layer (Albergel et al., 2008; Paulik et al., 2014). This product was obtained from CGLS (Copernicus Global Land Service, https://land.copernicus.eu/global/products/swi), which provides SM observations projected on a regular latitude/longitude grid with a spatial resolution of $0.1° \times 0.1°$ in units of percentage. The ASCAT SWI is a daily product with a time stamp at 12:00 UTC regardless of the actual observation time (Paulik et al., 2014). Previous studies have validated this SM product using remotely sensed SM or ground measurements, such as Albergel et al. (2009), Brocca et al. (2010), Brocca et al. (2011), and Paulik et al. (2014).

The ASCAT SWI values were masked out in our work when its surface state flag indicates frozen soil conditions. The surface state flag contains three different soil conditions, namely frozen, unfrozen, and melting states of the soil surface based on the temperature observations. Moreover, the SWI product provides a quality flag (QFLAG) that describes the number of available ASCAT SM observations considered during the SWI calculation. Here, SWI values were filtered out when the corresponding QFLAG value is smaller than 10% as these SWI values become unreliable due to limited ASCAT SM inputs (Wu et al., 2021).

SM product simulated from the GLDAS2 was used as a third independent estimate of SM applied in our TCA. GLDAS2 SM product was created by the Noah land surface model that uses a combination of model and observation data (excluding SM products) as its forcing data (Rodell et al., 2004). Only SM simulations in the top soil layer (0-10 cm) were used here since ASCAT and SMOS/SMAP SM observations only represent SM content of the topsoil layer in several centimeters. The GLDAS2 provides 3-hourly SM simulations projected on a regular latitude/longitude grid of $0.25^\circ \times 0.25^\circ$. Here, a daily GLDAS2 SM product was constructed by averaging all available SM simulations within the same day. To exclude SM values simulated under frozen soil conditions, GLDAS2 SM values were screened if the corresponding GLDAS2 soil temperature value is lower than 0℃.

The ERA-Interim SM product is a global reanalysis dataset derived from an advanced data assimilation system (Balsamo et al., 2015). It was used as an independent SM here to test the spatial representativeness of the ground measurements following Dorigo et al. (2015). This dataset is projected on a regular latitude/longitude grid with a spatial resolution of $0.25^\circ \times 0.25^\circ$ and provides SM analysis values with UTC stamps in the four soil layers, i.e., 0-7 cm, 7-28 cm, 28-100 cm, and 100-289 cm. Here, only SM data in the soil layer of 0-7 cm were considered. A daily ERA-Interim SM dataset was reconstructed by averaging all available SM values within the same day. Given that the ERA-Interim product is only available until 31 August, 2019, ERA-Interim SM data from July 3, 2012 to 31 August, 2019 were used in our analysis. The ERA-Interim SM values were screened when the associated GLDAS2 soil temperature value is lower than 0℃ to exclude SM values simulated under frozen soil conditions.

In our work, the projection grid of the GLDAS2 SM product was selected as the reference projection. ASCAT, SMOS, SMAP, and ERA-Interim SM observations were resampled on the GLDAS2 grid using the nearest neighbor interpolation. Since multiple SM products were considered here, their different overpass time stamps were accounted for by binning all SM products to a regularized time step (Gruber et al., 2020), i.e., daily UTC time here. A summary of the SM datasets used in our work is provided in Table 1. The timeframe of our analysis starts from July 3, 2012 to August 31, 2020 with a total of 2982 days. Although the discrepancy in measurement depth is considerable for ASCAT, SMOS, SMAP, GLDAS2, and ERA-Interim, Albergel et al. (2008) and Brocca et al. (2011) pointed out that SM in the upper 10 cm is strongly correlated with surface SM (e.g., 0-5 cm).

Table 1 A summary of the SM products used in this work

SM products	Spatial resolution	Temporal resolution	Nominal observation depth	Temporal coverage	Version	Projection grid
GLDAS2	0.25°×0.25°	3-hourly, UTC	0-10 cm	January 2000 to present	2.1	regular lat/lon grid
ASCAT SWI	0.1°×0.1°	daily at 12:00, UTC	0-5 cm	January 2007 to present	3.0	regular lat/lon grid
SMOSL3	40 km	6 AM and 6 PM local time	0-5 cm	January 2010 to present	3.3	EASE-Grid 2.0
SMOSIC	25 km	6 AM and 6 PM local time	0-5 cm	January 2010 to present	2.0	EASE-Grid 2.0
ERA-Interim	0.25°×0.25°	3-hourly, UTC	0-7 cm	January 1979 to August 2019	2.0	regular lat/lon grid
SMAPL3	36 km	6 AM and 6 PM local time	0-5 cm	March 2015 to present	8	EASE-Grid 2.0
SMAPIB	36 km	6 AM local time	0-5 cm	March 2015 to present	1.0	EASE-Grid 2.0

Note: Equal-Area Scalable Earth Grid, Version 2.0 (EASE-Grid 2.0)

In addition, SMOS and SMAP SM data were temporally interpolated here by filling the gaps existed in their time series with the average of neighboring SM values included in a 3-day time window. This will significantly increase the statistical power of the TCA method since more SM triplet samples can participate in the TCA calculation (Leroux et al., 2013; Wu et al., 2021). Limited samples, e.g., 30 samples, can lead to many error estimates that are not convergent, which reduces the robustness of the moving-window-based TCA method. However, the temporal interpolation influences the final TCA errors inevitably due to the additional interpolation errors and such impact should be investigated before applying the TCA time-variant scheme to the temporally interpolated SM datasets (detailed in Sec. 2.6 and Sec. 3.1).

2.2 ISMN ground measurements

Ground-based SM measurements are often considered to be the most accurate representation of the true SM content even though their spatial support is very small. The ISMN ground measurements from 759 stations belonging to 25 sparse networks (Dorigo et al., 2011) were used here to validate the TCA evaluation performance. To be consistent with the UTC time of other SM datasets, we took an average of all available SM values on the same day.

Stations were selected using several criteria. First, only stations of which the measurement interval was restricted within the upper 10 cm of the top soil layer were considered. Station with the shallowest depth was selected when stations with multiple depths are available in the same location (e.g., one taken at 5 cm and one taken at 10 cm). Second, only stations where the number of available SM values is greater than 500 were selected to guarantee sufficient data samples applied in the

TCA. Third, an areal representative station (Dorigo et al., 2015) was selected using a 3rd independent dataset, i.e., the ERA-Interim SM product, if more than one station was found in the experimental grid cell. Correlation coefficients between ground measurements and ERA-Interim, GLDAS2, SMOS/SMAP, and ASCAT SM data were calculated and the station with the highest correlation average was selected in the following analysis. Finally, ISMN ground measurements were screened by the ISMN quality flags (Dorigo et al., 2013) to guarantee their data quality. After filtering ISMN stations in the top soil, 759 stations remained available in our work. Fig. 1 shows the 759 experimental pixels that include these selected ISMN stations and Table 2 summarizes the ISMN sparse networks, number of stations used, and associated references.

Referring to the classification scheme of the MODIS IGBP dataset (Friedl et al., 2002), land cover types associated with the experimental grid cells mainly include grasslands (47.4%), croplands (15.0%), woody savannas (12.1%), savannas (8.8%), forests (7.5%), and open shrublands (4.9%). These six land cover types are used to categorize our results in the following analyses.

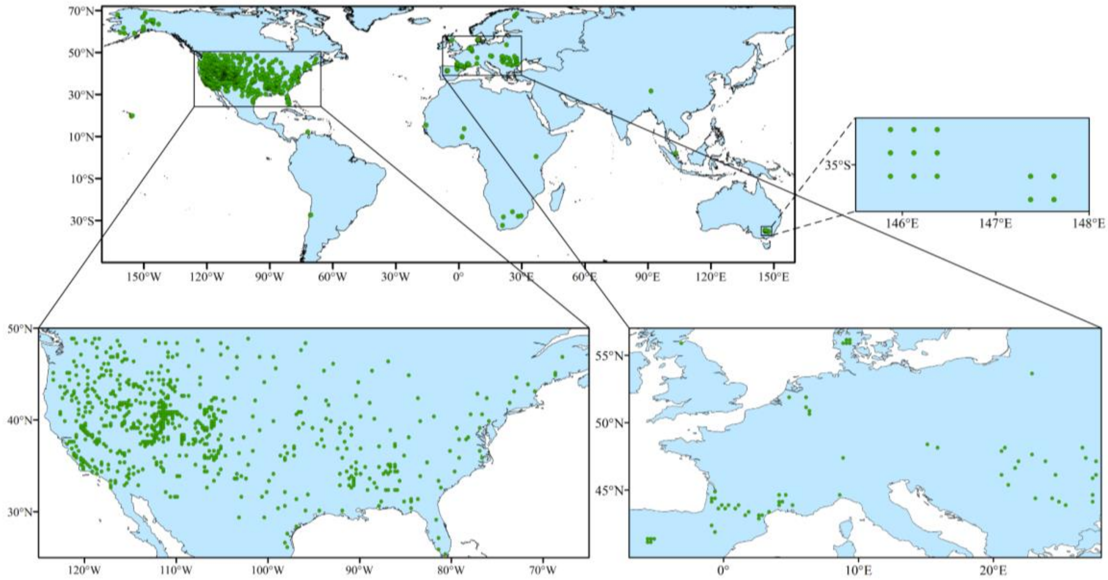


Fig. 1 Spatial distribution of 759 grid cells that include selected ISMN stations.

Table 2 ISMN ground stations from sparse networks

Network	Number of stations	Reference
AMMA-CATCH	3	Pellarin et al. (2009), Mougin et al. (2009), Cappelaere et al. (2009), de Rosnay et al. (2009)
ARM	4	http://www.arm.gov/
BIEBRZA_S-1	1	http://www.igik.edu.pl/en
COSMOS	5	Zreda et al. (2012)
CTP_SMTMN	1	Yang et al. (2013)
DAHRA	1	Tagesson et al. (2015)
FMI	3	http://fmiarc.fmi.fi/
FR_Aqui	3	Institute of Agricultural Research
GROW	7	https://growobservatory.org/index.html
HOBE	5	Bircher et al. (2012)
IMA_CAN1	1	Biddoccu et al. (2016)
IPE	2	Instituto Pirenaico de Ecologia (IPE-CSIC)
iRON	3	Osenga et al. (2019)
LAB-net	2	Mattar et al. (2016)
MySMNet	1	University Technology Malaysia
OZNET	13	Smith et al. (2012)
PBO_H2O	123	Larson et al. (2008)
REMEDHUS	5	http://campus.usal.es/~hidrus/
RSMN	19	http://assimo.meteoromania.ro/
SCAN	174	http://www.wcc.nrcs.usda.gov/
SMOSMANIA	19	Calvet et al. (2007), Albergel et al. (2008)
SNOTEL	256	http://www.wcc.nrcs.usda.gov/
SOILSCAPE	6	Moghaddam et al. (2010)

TERENO	3	Zacharias et al. (2011)
USCRN	99	Bell et al. (2013)

2.3 Triple/Quadruple Collocation Analysis

This section briefly introduces two Collocation-based Analyses, i.e., TCA and QCA, and the approach to estimate time-variant errors that have the same meaning of the TCA errors using ground measurements. TCA provides SM error estimates for three SM datasets without requiring an absolute truth. It assumes a linear relationship between SM observation θ_i and hypothetical unknown SM truth θ, which can be written as follows:

$$\theta_i = \alpha_i + \beta_i\theta + \varepsilon_i \quad i \in \{\text{GLDAS2}, \text{ASCAT}, \text{passive SM}\} \tag{1}$$

where the ε_i represents zero-mean random noise in θ_i, the α_i and β_i are additive and multiplicative coefficients that represent systematic errors in θ_i. The passive SM includes SMOSL3, SMOSIC, SMAPL3, and SMAPIB SM products. Based on this error model, SM errors can be obtained by calculating corresponding (co-)variances of the three SM datasets and simplifying with the other two assumptions, i.e., error orthogonality and zero Error Cross Correlation (ECC). Detailed formula derivation can be found in Gruber et al. (2016a).

Rescaling coefficients can be obtained through TCA along with error estimates. They can be used to make errors comparable after TCA or rescale SM data into a preselected reference dataset (Yilmaz and Crow, 2014; Gruber et al., 2017). Here, GLDAS2 was regarded as the reference dataset that was already perfectly calibrated. Therefore, the rescaling coefficients can be written as

$$\begin{cases} \beta_{\text{ASCAT}}^{\text{GLDAS2}} = \dfrac{\text{Cov}\left(SM_{\text{GLDAS2}}, SM_{\text{passive}}\right)}{\text{Cov}\left(SM_{\text{ASCAT}}, SM_{\text{passive}}\right)} \\ \beta_{\text{passive}}^{\text{GLDAS2}} = \dfrac{\text{Cov}\left(SM_{\text{GLDAS2}}, SM_{\text{ASCAT}}\right)}{\text{Cov}\left(SM_{\text{passive}}, SM_{\text{ASCAT}}\right)} \end{cases} \tag{2}$$

where $\beta_{\text{ASCAT}}^{\text{GLDAS2}}$ and $\beta_{\text{passive}}^{\text{GLDAS2}}$ are used to linearly rescale the ASCAT and passive-based SM errors against the GLDAS2 errors; $\text{Cov}(X, Y)$ represents covariance between X and Y time series.

The QCA extends TCA to four SM products and can account for the existence of non-zero ECC within a certain pair of collocated SM datasets. The QCA was used here to provide ECC values between GLDAS2, ASCAT, and SMOS/SMAP SM products and serve as an alternative error estimate method. The formulation reported in Gruber et al. (2016b) was conducted at the 759 sparse ground observation sites where ground measurements can serve as the fourth SM product. The least square solution for the QCA problem is given by

$$\boldsymbol{y} = \begin{bmatrix} \sigma_a^2 \\ \sigma_b^2 \\ \sigma_c^2 \\ \sigma_d^2 \\ \sigma_{ab} \\ \frac{\sigma_{ac}\sigma_{ad}}{\sigma_{cd}} \\ \frac{\sigma_{bc}\sigma_{bd}}{\sigma_{cd}} \\ \frac{\sigma_{ac}\sigma_{cd}}{\sigma_{ad}} \\ \frac{\sigma_{bc}\sigma_{cd}}{\sigma_{bd}} \\ \frac{\sigma_{ad}\sigma_{cd}}{\sigma_{ac}} \\ \frac{\sigma_{bd}\sigma_{cd}}{\sigma_{bc}} \\ \frac{\sigma_{ac}\sigma_{bd}}{\sigma_{cd}} \\ \frac{\sigma_{ad}\sigma_{bc}}{\sigma_{cd}} \end{bmatrix} \quad \boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \boldsymbol{x} = \begin{bmatrix} \beta_a^2\sigma_\theta^2 \\ \beta_b^2\sigma_\theta^2 \\ \beta_c^2\sigma_\theta^2 \\ \beta_d^2\sigma_\theta^2 \\ \beta_a\beta_b\sigma_\theta^2 \\ \sigma_{\varepsilon_a}^2 \\ \sigma_{\varepsilon_b}^2 \\ \sigma_{\varepsilon_c}^2 \\ \sigma_{\varepsilon_d}^2 \\ \sigma_{\varepsilon_a\varepsilon_b} \end{bmatrix} \tag{3}$$

where a, b, c, and d denote ASCAT, passive-based, GLDAS2, and ground SM, respectively. To ensure the robustness of error estimates, TCA and QCA values were used only when Pearson's correlation coefficients between the triplet in TCA or quadruplet in QCA are greater than 0.2 and passed a t-test ($\alpha < 0.05$) (Scipal et al., 2008a, Su et al., 2014a). The least squares solution for the parameter $\boldsymbol{x}$ is given as

$$\widehat{\boldsymbol{x}} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \tag{4}$$

Ideally, ECC values between the SM triplet should be zero, which is a strong assumption in TCA method. In practice, researchers try to solve this problem by choosing three independent SM datasets, e.g., one active microwave, one passive microwave, and one model-based SM (Scipal et al., 2008a; Gruber et al., 2015; Gruber et al., 2016a; Wu et al., 2018). However, a certain level of ECC is inevitable in practice, especially for active and passive SM retrievals as reported in Gruber et al. (2016b) and Pierdicca et al. (2017). Therefore, it is necessary to inspect ECC values between these two kinds of SM datasets prior to TCA to guarantee reliable results.

As reported in Gruber et al. (2016b) and Chen et al. (2018), the ECC values can be obtained by the QCA. In particular, a selected data pair (e.g., ASCAT and passive-based SM here) is allowed to be correlated during the QCA while ECC between other data pairs are still required to be zero. Consequently, ECC value between the selected data pair can be obtained via their error covariance (i.e., $\sigma_{\varepsilon_a \varepsilon_b}$) and their error variance (i.e., $\sigma^2_{\varepsilon_a}$ and $\sigma^2_{\varepsilon_b}$) values derived from the Eq. (3). Alternatively, ECC values can be calculated based on the ground measurements directly. The boxplots of ECC values between ASCAT and passive SM datasets derived from QCA and ground measurements were shown in Fig. S2. For the QCA implementations consider

SMOSL3, SMOSIC, SMAPL3, and SMAPIB as the passive SM in the quadruplet, the interquartile range is [-0.08, 0.34] and [-0.16, 0.35] for ECC values derived from QCA and ground measurements, respectively, and the associated median values are 0.16 and 0.20 for these two cases. The small ECC values suggest that the strong assumption in TCA is generally fulfilled in our analysis. It is notable that ECC values can also be obtained by an extended double instrumental variable algorithm using only two independent products, which was recently proposed by Dong et al. (2020).

TCA can provide time-variant errors by relaxing the stationary assumption. For example, the time-variant errors can be obtained by applying TCA to SM time series for a 30-day moving window advancing by 15-day steps over the experimental periods (Loew and Schlenz, 2011). However, we argue that a moving-time-window advancing by daily step can better capture the temporal variability of SM errors and provide more inherent information about time-variant errors. Moreover, Su et al. (2014) proposed an approach of estimating multi-annual window-based errors for each day of the year. But this may overlook subtle temporal variability in SM errors as the interannual variation of environmental conditions, such as vegetation and rainfall, may change a lot in different years. Based on these considerations, a moving-window-based TCA scheme was considered here.

Following the scheme proposed by Wu et al. (2021), we obtained ASCAT and SMOS/SMAP time-variant daily errors by applying TCA to SM time series with a sliding 101-day moving window advancing by daily step over the whole experimental period. The 101-day time window was used here to estimate time-variant errors with sufficient statistical power while keeping the kernel size of the window short enough to capture the seasonal variability of SM errors. Since temporal interpolation was considered for the SMOS and SMAP SM datasets here, the number of available triplet samples used in the 101-day window is typically large enough to obtain reliable time-variant errors. To avoid unreliable error estimates derived from limited samples (e.g., 5 triplets), the minimum sample requirement was defined as 100 (Scipal et al., 2008a) and 90 (Wu et al., 2021) for the time-invariant and time-variant error estimates, respectively.

2.4 Rescaling techniques and input selection in TCA

Rescaling is an essential technique included in TCA as it eliminates bias between different SM datasets prior to the TCA or adjusts the TCA error estimates into a preselected reference space for a comparison purpose after the TCA implementation. Here, we investigated four popular rescaling techniques generally applied in TCA: three rescaling techniques applied prior to TCA, including variance and mean matching (VAR), normalization (NORM), and CDF matching (CDF), and one rescaling technique considers coefficients derived from TCA (TCA_Self). For convenience, these abbreviations will be used to refer to the four rescaling techniques described above. Since Section 2.3 has described the TCA_Self in equation (2), we briefly introduced the other three rescaling techniques here.

VAR is a linear rescaling technique that forces SM time series to have the same mean and standard deviation as the reference SM time series (Draper et al., 2009; Brocca et al., 2010; Dorigo et al., 2010). Equation (5) describes this rescaling approach.

$$SM_{\text{rescaled}} = \frac{SM - \mu(SM)}{\sigma(SM)} \sigma(SM_{\text{GLDAS2}}) + \mu(SM_{\text{GLDAS2}}) \tag{5}$$

where SM and SM_{rescaled} denote SM time series before and after rescaling, respectively, $\mu(\cdot)$ and $\sigma(\cdot)$ represent mean and standard deviation.

NORM is a standardization method that forces SM time series to have the same maximum and minimum as the reference SM time series (Rüdiger et al., 2009; Albergel et al., 2010; Albergel et al., 2012; Su et al., 2014a; Wu et al., 2021). This transformation can be written as the following equation (6).

$$SM_{\text{rescaled}} = \frac{SM - \min(SM)}{\max(SM) - \min(SM)} [\max(SM_{\text{GLDAS2}}) - \min(SM_{\text{GLDAS2}})] + \min(SM_{\text{GLDAS2}}) \tag{6}$$

where SM and SM_{rescaled} have the same meaning as those in equation (6), $\min(\cdot)$ and $\max(\cdot)$ represent minimum and maximum values in corresponding time series, respectively.

CDF can be considered as an enhanced nonlinear rescaling technique, which is rescaled in such a way that the cumulative distribution function of SM time series is matched with that of reference SM data (Reichle and Koster, 2004; Drusch et al. 2005; Brocca et al. 2011; Brocca et al., 2013; Su et al., 2013). Equation (7) describes this transformation.

$$CDF(SM_{\text{rescaled}}) = CDF(SM) \tag{7}$$

where $CDF(\cdot)$ represents cumulative distribution function, SM and SM_{rescaled} have the same meaning as those in equation (5). Following Brocca et al. (2013), a fifthorder polynomial function was considered in our CDF rescaling to match the two cumulative distribution functions.

Besides the aforementioned four rescaling techniques, QCA errors were also considered to make a comparison with the TCA errors. The ASCAT and SMOS/SMAP errors estimated by QCA were linearly rescaled against GLDAS2 errors using the following equation

$$\begin{cases} \sigma_{ASCAT_{\mathrm{after}}} = \dfrac{\beta_{GLDAS2}}{\beta_{ASCAT}} \times \sigma_{ASCAT_{\mathrm{before}}} \\ \sigma_{passive_{\mathrm{after}}} = \dfrac{\beta_{GLDAS2}}{\beta_{passive}} \times \sigma_{passive_{\mathrm{before}}} \quad passive \in \{SMOS, SMAP\} \end{cases} \tag{8}$$

where $\sigma_{ASCAT_{\mathrm{after}}}$ and $\sigma_{passive_{\mathrm{after}}}$ denote rescaled errors for ASCAT and passive SM data, i.e., SMOSL3, SMOSIC, SMAPL3, and SMAPIB SM products, $\sigma_{ASCAT_{\mathrm{before}}}$ and $\sigma_{passive_{\mathrm{before}}}$ represent ASCAT and passive SM errors before rescaling. The rescaling coefficients, $\frac{\beta_{GLDAS2}}{\beta_{ASCAT}}$ and $\frac{\beta_{GLDAS2}}{\beta_{passive}}$, can be obtained from equation (3).

TCA can be applied to original SM values or anomalies. Here, following Dorigo et al. (2010) and Albergel et al. (2012), anomalies were obtained by subtracting averages of SM values in a sliding time window from original SM values. The sliding window has a kernel size of 35-day and advances by daily step over the whole investigation period.

2.5 Conventional ground-based approach to estimate SM errors

To assess the evaluation power of TCA, TCA errors derived from the four rescaling techniques were validated by conventional SM errors obtained from ground measurements that are considered as the benchmark during the validation. As reported in Yilmaz and Crow (2014), ground-based error variance (regarded as the error term in our work) can be obtained by equation (9), which has the same meaning as the TCA errors.

$$\sigma_{\varepsilon_X}^2 = \frac{1}{N}\sum_{t=1}^{N}(X_t - SM_{\mathrm{Ground2}_t})^2 \quad X \in \{SM_{\mathrm{SMOS}}, SM_{\mathrm{SMAP}}, SM_{\mathrm{ASCAT}}\} \tag{9}$$

where $\sigma_{\varepsilon_X}^2$ denotes ground-based error variance for the given dataset X, X_t and SM_{Ground_t} represent SM values of dataset *X* and ground measurement at time step *t*, respectively, *N* is the number of whole investigation days.

In Eq. (9), the pre-processing of the ground measurements is different for each rescaling technique. For VAR, NORM, and CDF applied prior to TCA, ground measurements and other SM products were jointly rescaled by the same rescaling technique. By contrast, ground measurements were directly used without further process in TCA_Self and QCA_Self, and the resulting ground-based errors were rescaled by the same parameter and the same way as in the rescaling of TCA- and QCA-based errors.

2.6 Robustness of temporal interpolation applied in TCA time-variant scheme

Here, the TCA time-variant scheme was applied to passive-based SM datasets, i.e., SMOSL3, SMOSIC, SMAPL3, and SMAPIB, with temporal interpolation. The temporal interpolation introduces additional errors into SM time series and therefore may have potential impacts on the final TCA error estimates. Before applying TCA to temporally interpolated SM data, it is necessary to investigate such impact on the TCA time-variant errors.

To this end, first, we found a few pixels that have enough SM samples without interpolation to calculate the TCA time-variant errors using a 101-day moving-time-window. Since multiple passive SM datasets were considered in our work, the pixels were separately selected for each TCA implementation that uses SMOSL3, SMOSIC, SMAPL3, and SMAPIB SM as one of the triplet inputs. Second, the passive SM time series were resampled without replacement using a different percentage of the original data amount, i.e., 95%, 90%, 80%, 70%, 60%, and 50%. The passive SM time series were then temporally interpolated and time-variant TCA errors were calculated based on the resampled and interpolated SM datasets. The resulting time-variant errors were compared with the time-variant errors derived from the original SM data (without interpolation) using the Pearson's correlation and RMSD metrics. To guarantee reliable conclusions, the resampling was repeated 1000 times and we took an average of the resulting correlation coefficients and RMSD values to represent the overall performance of interpolation on TCA error estimates. Finally, correlations and RMSD values between time-variant errors calculated from original SM data and SM time series resampled by different percentage values can be obtained for each selected pixel. Based on the correlation and RMSD results, the impact of temporal interpolation on TCA errors can be quantified.

2.7 Evaluation metrics

The Relative Difference (RD) and Overall Relative Difference (ORD) metrics were used here to evaluate the relative difference between time-invariant and time-variant errors. In a given pixel, the ORD is defined as:

$$\mathrm{ORD} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\frac{TVE_t - TIE}{TIE}\right)^2} \times 100\% \tag{10}$$

where TIE denotes time-invariant error and TVE_t represents time-variant error at time step t, N denotes the actual number of available common pairs of TIE and TVE in the given pixel. ORD quantifies the overall relative difference between time-invariant and time-variant errors. But it fails to provide intuitive information about the positive or negative sign of the relative difference between these two kinds of errors. Therefore, the RD metric was constructed to address this issue. For a given pixel, the RD is defined as follows:

$$\mathrm{RD} = \frac{1}{N}\sum_{t=1}^{N}\frac{TVE_t - TIE}{TIE} \times 100\% \tag{11}$$

where symbols in equation (11) are the same as those in equation (10).

Pearson's correlation and RMSE metrics were used to evaluate the accuracy of TCA-based error estimates and ground-based errors were regarded as the benchmark. The RMSE is defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\sigma_{\mathrm{TCA}_t} - \sigma_{\mathrm{ground}_t})^2} \tag{12}$$

where σ_{ground} is ground-based errors and σ_{TCA} represents TCA-based errors, N denotes the number of associated common pairs of σ_{ground} and σ_{TCA} during the investigation period.

The calculation of RMSD metric is similar to RMSE and is defined as:

$$\mathrm{RMSD} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(\sigma_{\mathrm{resampled}_t} - \sigma_{\mathrm{original}_t}\right)^2} \tag{13}$$

where $\sigma_{\mathrm{resampled}}$ denotes TCA time-variant errors derived from SM time series resampled by the different percentage values whereas $\sigma_{\mathrm{original}}$ represents TCA time-variant errors estimated from original SM data without interpolation.

3 Results

3.1 Impact of temporal interpolation on TCA error estimates

As noted in Sec. 2.6, a potential error source of the TCA estimates in our work is the temporal interpolation applied to the passive SM datasets, which may lead to a potential impact on the final TCA error estimates. Fig. 2 explores such impact using the method described in Sec. 2.6. SMAPIB results are not included here as limited available samples (< 30) were found. Three ASCAT results are expected from the TCA implementations with SMOSL3, SMAPL3, and SMOSIC, and all of them are used to construct the boxplots in Fig. 2 (d).

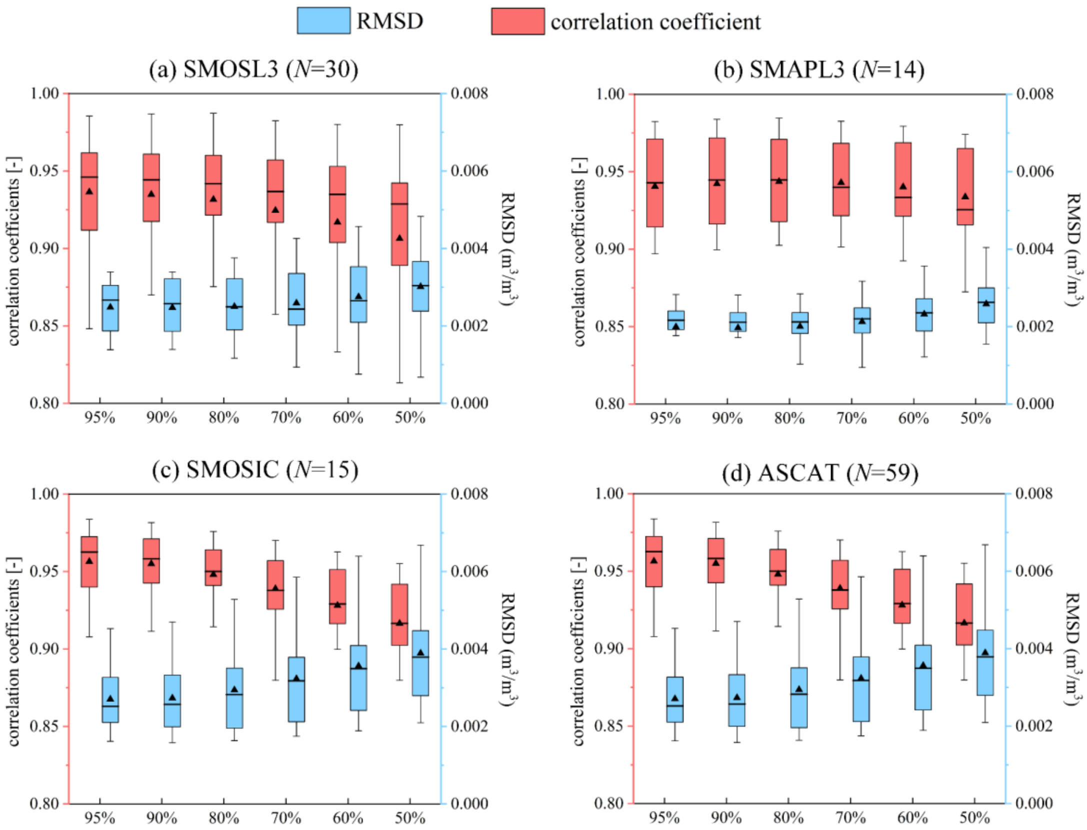


Fig. 2 Boxplots of correlation coefficients (red) and RMSD values (blue) between time-variant errors derived from the original SM data without interpolation and the SM time series resampled by different percentage values of the original data. (a-d) exhibits the results of applying interpolation to SMOSL3, SMAPL3, SMOSIC, and ASCAT SM data, respectively. The x-axis denotes the percentage values of the original data considered in the resampling. The y-axes on the left (red color) and right (blue color) describe correlation coefficients and RMSD, respectively. The '*N*' above each subfigure denotes the number of available samples included in the corresponding boxplot.

Boxplots in Fig. 2 demonstrate that the temporal interpolation has a small impact on the final TCA error estimates. In line with expectations, for all cases shown in Fig. 2 (a-d), correlation displays a decreasing trend while RMSD yields an increasing trend as the sampling percentage of the original dataset decreases, which indicates the temporal interpolation indeed introduces additional errors into SM time series and consequently degenerates the accuracy of the final TCA error estimates. However, it is worthwhile mentioning that this degeneration is relatively small as correlation coefficients were typically high (> 0.9) and RMSD values were generally small (< 0.004 m^3/m^3) for all the cases. Moreover, as the sampling percentage decreases from 95% to 50%, the correlation decreased by a value smaller than 0.04 and the RMSD typically increased by a value smaller than 0.001 m^3/m^3 regarding their average and median values. These results demonstrate that the temporal interpolation applied in our work has a small impact on TCA error estimates and it is an efficient way to address the limited sample issue lies in the TCA time-variant scheme.

3.2 Accuracy of TCA errors derived from different rescaling techniques

Pearson's correlation coefficients and RMSE values between ground- and TCA-based time-invariant errors for ASCAT and multiple passive SM products are summarized in Table 3 and Table 4, respectively. The impacts of the four rescaling techniques (i.e., VAR, NORM, CDF, and TCA_Self) and the two input selections (i.e., original values or anomalies) on the final TCA time-invariant errors are compared in these two tables. Moreover, QCA error estimates (marked as QCA_Self) were also considered to make a comparison with the TCA errors. The correlation and RMSE values of ASCAT are averaged from multiple ASCAT results that are obtained from the TCA implementations with multiple passive SM data used here.

Table 3 Pearson's correlation coefficients between ground- and TCA-based time-invariant errors for SMOSL3, SMOSIC, SMAPL3, SMAPIB, and ASCAT SM products. The time-invariant errors are derived from multiple rescaling techniques. The first and second recommendation strategies (excluding QCA_Self) to implement TCA are highlighted with green and yellow colors, respectively. All the ρ values passed a t-test ($\alpha < 0.05$)

		NORM	VAR	CDF	TCA_Self	QCA_Self
original values	SMOSL3	0.25	0.39	0.42	0.65	0.92
	SMOSIC	0.15	0.31	0.40	0.54	0.97
	SMAPL3	0.19	0.20	0.34	0.38	0.96
	SMAPIB	0.21	0.19	0.16	0.50	0.74
	ASCAT	0.49	0.70	0.73	0.86	0.89
anomalies	SMOSL3	0.69	0.75	0.74	0.95	0.96
	SMOSIC	0.64	0.66	0.63	0.89	0.90
	SMAPL3	0.56	0.58	0.60	0.70	0.79
	SMAPIB	0.47	0.58	0.54	0.70	0.67
	ASCAT	0.74	0.79	0.79	0.93	0.90

Table 4. The same as Table 3 but for the RMSE values

		NORM	VAR	CDF	TCA_Self	QCA_Self
original values	SMOSL3	0.0256	0.0191	0.0176	0.0178	0.0220
	SMOSIC	0.0295	0.0188	0.0178	0.0276	0.2208
	SMAPL3	0.0265	0.0227	0.0214	0.0263	0.0575
	SMAPIB	0.0265	0.0207	0.0203	0.0241	0.0266
	ASCAT	0.0231	0.0156	0.0147	0.0194	0.0217
anomalies	SMOSL3	0.0086	0.0072	0.0074	0.0047	0.0046
	SMOSIC	0.0099	0.0075	0.0077	0.0062	0.0062
	SMAPL3	0.0102	0.0085	0.0087	0.0083	0.0086
	SMAPIB	0.0163	0.0184	0.0214	0.0158	0.0157
	ASCAT	0.0079	0.0075	0.0077	0.0060	0.0070

Tables 3 and 4 jointly show that the highest correlation and the smallest RMSE values between ground- and TCA-based time-invariant errors are obtained by applying TCA to SM anomalies and rescaling the resulting errors using TCA_Self.

Table 3 shows the correlations between these two kinds of errors were 0.95, 0.89, 0.70, 0.70, and 0.93 for the SMOSL3, SMOSIC, SMAPL3, SMAPIB, and ASCAT SM products, respectively. This combination strategy of applying anomalies and TCA_Self was also recommended based on the RMSE results shown in Table 4, as the smallest RMSE value was typically observed when considering this combination strategy. Furthermore, Tables 3 and 4 indicate that there is a small difference between errors derived from TCA_Self and QCA_Self when considering anomalies as the SM inputs. By contrast, errors estimated from TCA_Self and QCA_Self exhibited evident discrepancies for the correlation and RMSE metrics when considering absolute values as inputs. In particular, as shown in Table 3, errors derived from QCA_Self appeared higher correlation coefficient than those derived from TCA_Self when using absolutes as SM inputs.

The validation of TCA time-variant errors derived from different rescaling techniques using ground-based errors are revealed in Fig. 3. Specifically, TCA- and ground-based time-variant errors are compared using Pearson's correlation and RMSE metrics for each experimental grid cell. The correlation coefficients and RMSE values were gleaned and construct the boxplots shown in Fig. 3. Since correlation and RMSE values of ASCAT SM product can be obtained for each TCA implementation that applied multiple passive SM datasets, all the ASCAT results are considered in Fig. 3 (e).

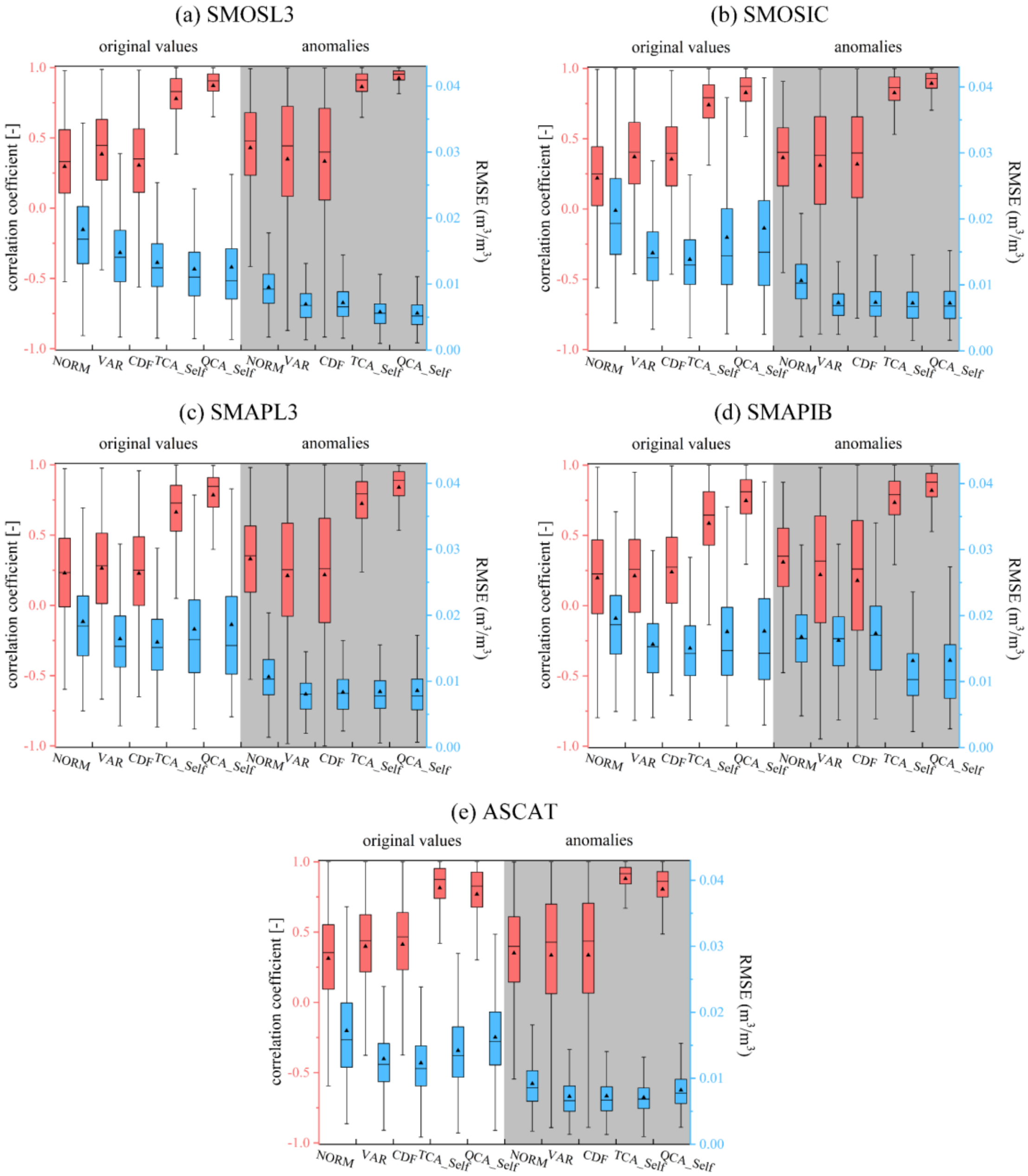


Fig. 3 Boxplots of the Pearson's correlation coefficients (red) and RMSE values (blue) between ground- and TCA-based time-variant errors for the TCA implementations that consider (a) SMOSL3, (b) SMOSIC, (c) SMAPL3, (d) SMAPIB, and (e) ASCAT as the triplet inputs. In each subfigure, results derived from absolutes and anomalies are shown in the white and grey areas, respectively. The x-axis is the multiple rescaling techniques considered in the TCA. The y-axes on the left (red color) and

right (blue color) describe correlation coefficients and RMSE, respectively.

Similar to the conclusion of time-invariant errors drawn from Tables 3 and 4, boxplots in Fig. 3 also demonstrate that the highest correlation and the smallest RMSE values between ground- and TCA-based time-variant errors are typically obtained by the combination strategy that considers SM anomalies as TCA inputs and rescaling the resulting errors with TCA_Self. Based on this optimal combination strategy, TCA errors were strongly correlated with ground-based errors as associated correlation coefficients were 0.88 (0.92), 0.83 (0.87), 0.86 (0.91), 0.73 (0.79), and 0.72 (0.79) for the ASCAT, SMOSIC, SMOSL3, SMAPIB, and SMAPL3 cases regarding the average (median) values. Corresponding scatterplots are illustrated in Fig. 4 and all available samples were included in such comparison.

In contrast with the correlation metric, RMSE values derived from TCA_Self were significantly smaller than those obtained from the other three rescaling techniques (excluding QCA_Self) only for the SMOSL3 and SMAPIB cases. However, it has to be stressed that TCA_Self RMSE values were generally small for all cases, i.e., typically smaller than 0.015 m^3/m^3 for SMAPIB and smaller than 0.01 m^3/m^3 for SMOSL3, SMOSIC, and SMAPL3. Furthermore, Fig. 3 implies that ground-based errors had a stronger correlation with QCA_Self errors than TCA_Self errors for all cases except for the ASCAT case. However, one can see that the discrepancies were relatively small for the TCA_Self and QCA_Self RMSE values.

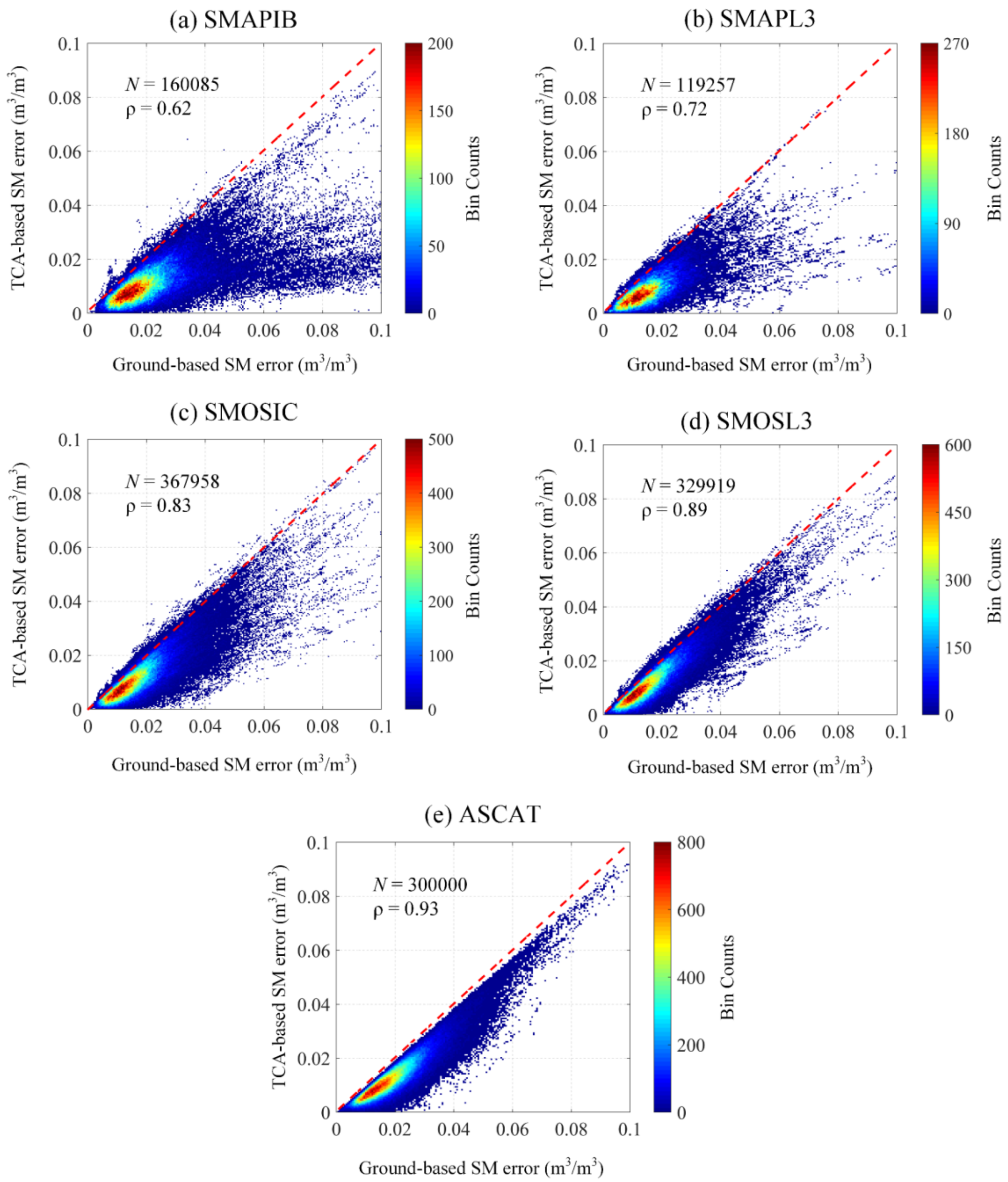


Fig. 4 Scatterplots of ground- and TCA-based time-variant errors for (a) SMAPIB, (b) SMAPL3, (c) SMOSIC, (d) SMOSL3, and (e) ASCAT SM data. The original scatter points are binned in the increase of x-axis for ground-based error and y-axis for TCA-based error, respectively, and colored based on the number of points included in the bin. All the ρ values passed a t-test ($\alpha < 0.05$). N describes the number of experimental grid cells considered in each scatterplot.

The scatterplots in Fig. 4 exhibit that TCA-based errors are strongly correlated with ground-based errors as their ρ values were in the range from 0.62 to 0.93. However, it is notable that TCA- and ground-based errors had a relatively weak correlation ($\rho = 0.62$) for the SMAPIB case, and considerable points with an underestimation of the TCA-based errors were observed in Fig. 4 (a). Moreover, Fig. 4 demonstrates that ground-based errors were generally greater than TCA-based errors. Given the above results, only TCA errors obtained from the optimal combination strategy, i.e., applying TCA to SM anomalies with TCA_Self, were considered in the following Sections 3.4 and 3.5.

At the end of this section, the correlation and RMSE values between TCA- and ground-based time-variant errors are further investigated in Fig. 5 by a classification with six land cover types, which explores the TCA performance in different land covers. Since multiple passive SM data were applied in TCA, all the resulting correlation and RMSE values of ASCAT SM data were considered in Fig. 5.

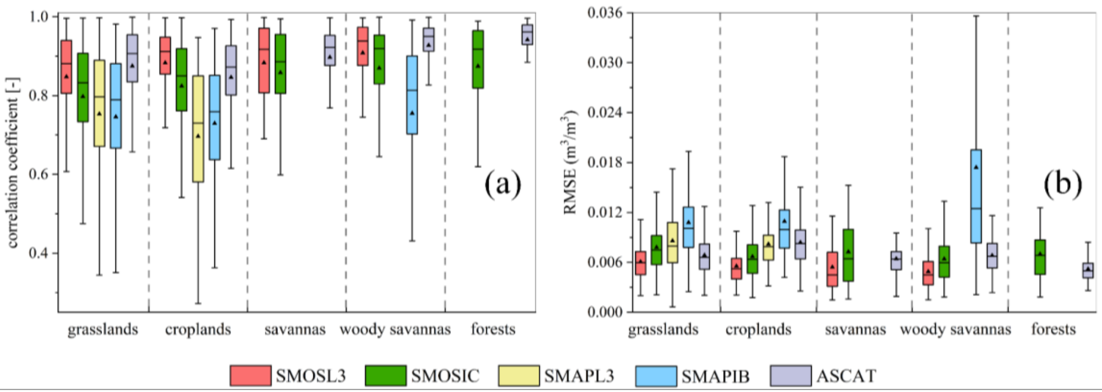


Fig. 5 Boxplots of (a) correlation coefficients and (b) RMSE between TCA- and ground-based errors classified by six land cover types for ASCAT and multiple passive SM products. Boxplots with limited available samples (< 30) are not

considered in such comparisons.

In general, Fig. 5 reveals that TCA can accurately estimate ASCAT and SMOS/SMAP time-variant errors in all the six land cover types as associated correlations are typically greater than 0.6 and the RMSE values are mostly smaller than 0.01 m^3/m^3. Nevertheless, Fig. 5 (a) shows relatively weak correlations between TCA- and ground-based errors in croplands for all SM datasets except for SMOSL3 SM, which implies TCA method has relatively less power to efficiently characterize errors of satellite-based SM products in croplands.

In comparison, TCA provides more accurate error estimates for SMOS SM than SMAP SM when considering ground-based errors as the benchmark. Even though several boxplots were not included in Fig. 5 due to their limited sample issue, it appears that the correlations presented a decreasing trend and the RMSE gave an increasing trend for SM products in the order of SMOSL3, SMOSIC, SMAPL3, and SMAPIB. Notably, the discrepancy between TCA- and ground-based errors is evident for SMAPIB SM, especially in woody savannas.

3.3 Relative difference between time-variant and time-invariant errors

This section identifies the relative difference between time-variant and time-invariant SM errors derived from different rescaling techniques to find out (i) the necessity to consider and include time-variant SM errors in applications and (ii) which one is larger, time-variant or time-invariant error? To these ends, the ORD and RD metrics were used to address these two issues. For a given combination of passive SM dataset and rescaling technique, the ORD value can be calculated for each experimental grid cell. Mean and median of the ORD values collected from all experimental pixels were used to describe the overall deviation between time-invariant and time-variant errors. Note that only TCA errors obtained from SM anomalies were considered in this section.

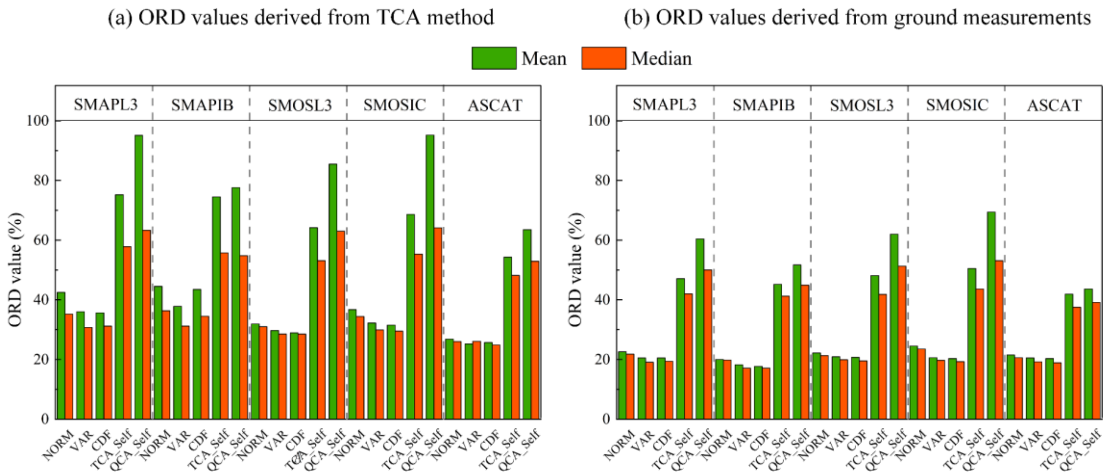


Fig. 6 Mean and median of the ORD values derived from the NORM, VAR, CDF, TCA_Self, and QCA_Self rescaling techniques for SMAPL3, SMAPIB, SMOSL3, SMOSIC, and ASCAT SM datasets. (a) and (b) represent ORD values obtained from the TCA- and ground-based evaluations, respectively.

The bar charts in Fig. 6 demonstrate that it is necessary to consider and include time-variant errors in actual applications. The ORD average and median values derived from different rescaling techniques were typically greater than 25% and 20% for the TCA-based and ground-based errors, respectively. In particular, ORD average and median values derived from the TCA_Self and QCA_Self methods were mostly greater than 50% and 40% for errors obtained from TCA and ground measurements, respectively. Comparing Fig. 6 (a) and (b), one can see that TCA-based ORD values were typically smaller than the ground-based ORD values, which implies TCA-based time-variant errors have a larger variance than ground-based time-variant errors.

The first question proposed at the beginning of this section is explored using the ORD metric in the above analysis. However, the ORD metric only explores overall magnitude of the relative differences between time-variant and time-invariant errors. The RD metric is adopted here to answer the second question put forward in this section.

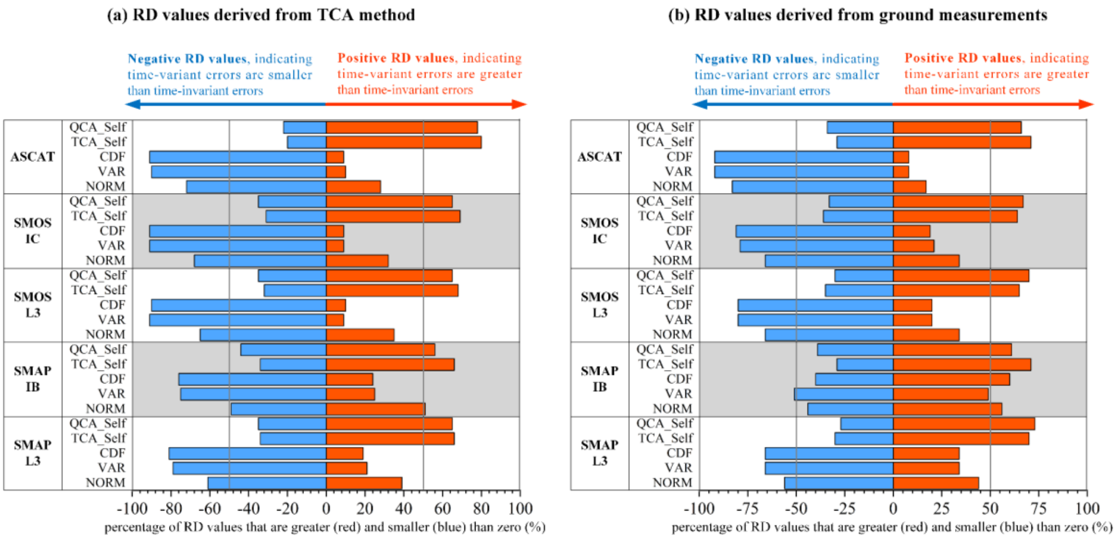


Fig. 7 Fractions of the RD values that are greater (red) or smaller (blue) than 0 for the combination of five SM datasets, i.e., ASCAT, SMOSIC, SMOSL3, SMAPIB, and SMAPL3, and five rescaling techniques, i.e., QCA_Self, TCA_Self, CDF, VAR, and NORM. (a) and (b) present associated results based on the RD values derived from the TCA- and ground-based evaluations, respectively.

The RD results in Fig. 7 reveal that the relative magnitude relationship between time-invariant and time-variant errors is varied with the selection of rescaling technique used in TCA. In general, negative RD values were observed for the CDF, VAR, and NORM rescaling techniques, implying time-variant errors were smaller than time-invariant errors. By contrast, positive RD values were found for the TCA_Self and QCA_Self rescaling techniques, which indicates time-variant errors were greater than time-invariant errors. It is worth mentioning that the RD results of SMAPIB SM are different from other cases, especially for the ground-based RD values shown in Fig. 7 (b).

The noticeable difference between time-variant and time-invariant errors demonstrated in Fig. 6 suggests that considering time-variant errors rather than time-invariant errors in applications is necessary. Time-variant error better characterizes the temporal variability of SM errors and consequently, a more accurate output is expected from applications that strongly rely on an accurate error specification, such as data merging and data assimilation studies. Contrary to the expectation, time-invariant errors do not provide an average reference for time-variant errors derived from TCA and their relative magnitude depends on the rescaling technique used. Simple rescaling techniques, such as VAR, CDF, and NORM, tend to underestimate SM time-variant errors when window-based TCA was applied.

3.4 Evaluation consistency between TCA- and ground-based methods

This section assesses the evaluation consistency between TCA- and ground-based methods regarding the time-variant errors of ASCAT and multiple passive SM data used in our work. For each experimental grid cell, the evaluation consistency was quantified as a percentage value, which represents the fraction of days that TCA- and ground-based methods provide consistent evaluation results. Four assessment results can be obtained for the four TCA implementations that consider SMOSL3, SMOSIC, SMAPL3, and SMAPIB as the passive SM in the triplet. In each TCA implementation, the consistent evaluation includes two cases: both TCA and ground-based methods suggest (i) ASCAT or (ii) the given passive SM dataset has the smallest time-variant error.

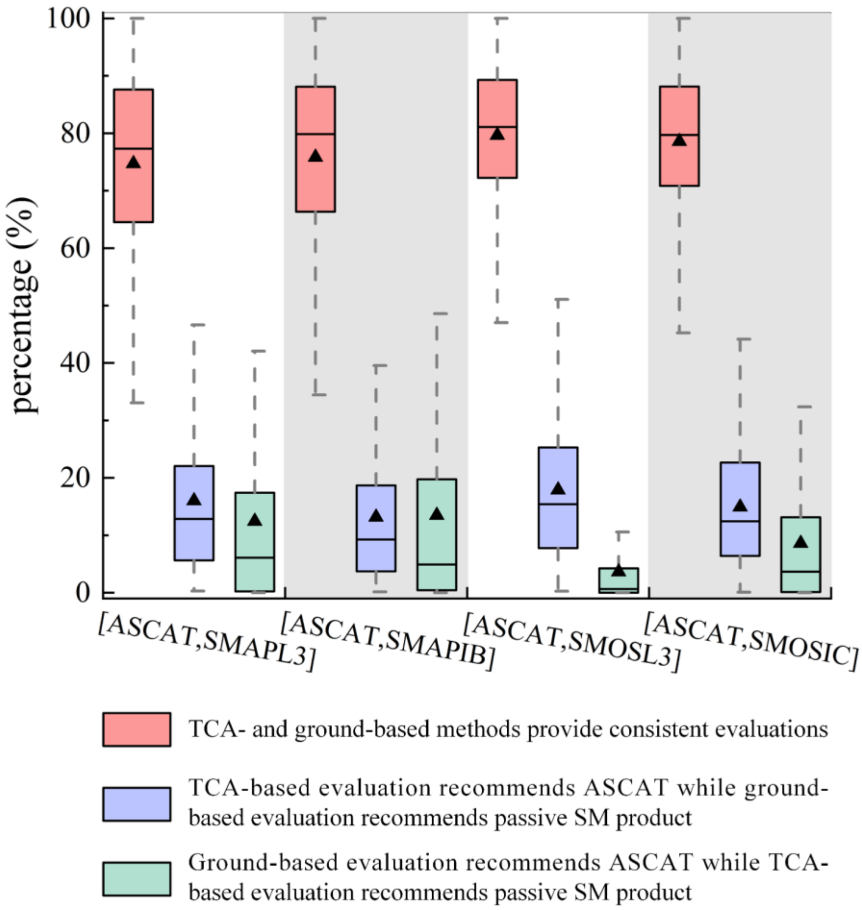


Fig. 8 Boxplots of a binary assessment of the evaluation consistency between TCA- and ground-based methods for ASCAT and four passive-based SM time-variant errors. The four columns illustrated in this figure denote the four TCA implementations that consider [GLDAS2, ASCAT, SMAPL3], [GLDAS2, ASCAT, SMAPIB], [GLDAS2, ASCAT, SMOSL3], and [GLDAS2, ASCAT, SMOSIC] as the triplet.

Boxplots in Fig. 8 show that there is a high consistency between TCA- and ground-based evaluations. These two kinds of methods provided consistent evaluations in 74.7% (77.3%), 75.8% (79.8%), 79.6% (81.1%), and 78.6% (79.7%) of the investigation days on the global average (median) for the TCA implementations with SMAPL3, SMAPIB, SMOSL3, and SMOSIC SM, respectively. However, the boxplots with dark blue and cyan in Fig. 8 also exhibit that the TCA- and ground-based evaluations provided inconsistent evaluation results in certain periods. Notably, TCA more or less underestimated ASCAT errors and overestimated passive-based SM errors and this phenomenon was more pronounced for the SMOS data than the SMAP data.

3.5 Comparisons of time-variant errors by land cover classification for ASCAT and four passive SM products

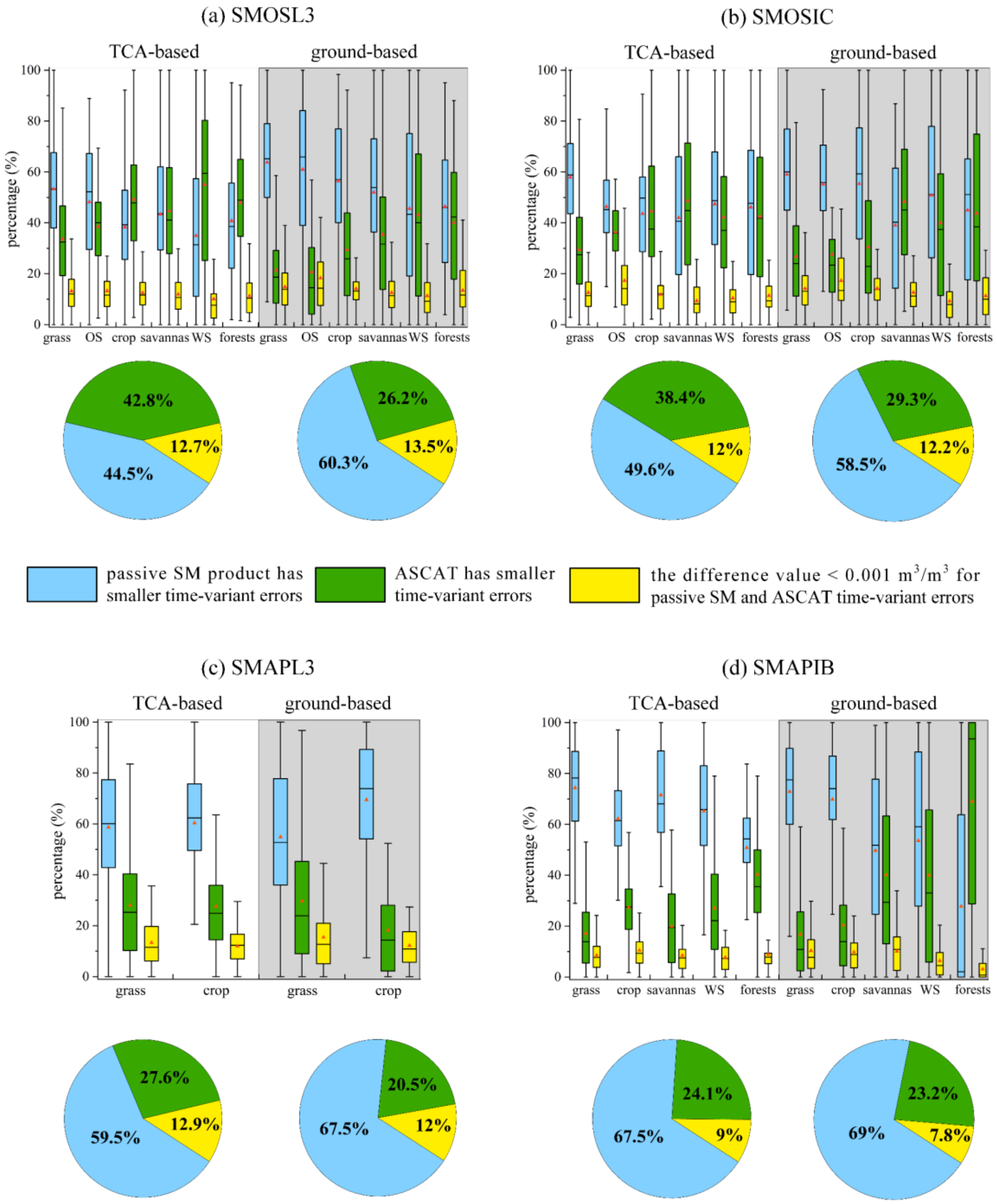
This section is an extensive analysis of the evaluation consistency shown in Sec. 3.4 by pairwise comparing TCA- and ground-based time-variant errors for ASCAT and the four passive SM products, i.e., SMOSL3, SMOSIC, SMAPL3, and SMAPIB. However, this section focuses on the error comparison while Sec. 3.4 aims at quantifying the evaluation consistency between TCA- and ground-based methods. Also, this section provides insights into the relative strengths of ASCAT and the four passive SM products in different land cover types regarding their time-variant errors. Four comparison results are expected as TCA was applied to the four above passive SM products. In such comparisons, three cases were considered: (i) ASCAT or (ii) passive-based SM has the smallest time-variant error, and (iii) ASCAT has similar performance with passive SM as the difference between their errors is smaller than 0.001 m^3/m^3. For a given TCA implementation that uses SMOS or SMAP SM as the

826 triplet, three percentage values corresponding to the above three cases can be

827 calculated for each grid cell. The percentage values were collected from all pixels and

828 were further categorized by six land cover types including grasslands (grass), open

829 shrublands (OS), croplands (crop), savannas, woody savannas (WS), and forests.



830

831 Fig. 9 Boxplots of the percentage values classified by six land cover types for the

832 three cases: ASCAT (green boxes) or passive SM (blue boxes) has the smallest time-

variant error and the difference between their errors is smaller than 0.001 m^3/m^3 (yellow boxes). (a-d) represent the TCA- (white areas) and ground-based (grey areas) assessments that consider SMOSL3, SMOSIC, SMAPL3, and SMAPIB SM data, respectively. To guarantee reliable results, boxplots associated with limited available samples (< 30) are excluded. The two pies below each subfigure exhibit the overall performance of the three cases that include all available samples for the TCA- (on the left) and ground-based (on the right) assessments.

Overall, both TCA- and ground-based comparisons demonstrate that ASCAT provides more and more reliable SM observations as vegetation cover increases while passive-based SM provides more and more reliable observations as vegetation cover decreases. The boxplots in Fig. 9 exhibit that the percentage value typically becomes smaller for passive SM while getting larger for ASCAT SM as vegetation cover increases. For all the four cases illustrated in Fig. 9 (a-d), percentage values of the case that ASCAT provided similar performance with passive SM (the yellow boxes) had relatively small differences in the six land cover types, and associated percentage values were mostly smaller than 20%.

The pie charts in Fig. 9 confirm that TCA underestimated ASCAT errors and overestimated SMOS errors when considering ground-based evaluation as the benchmark. Comparing the pie charts on the left and right sides, the fraction value that ASCAT had better performance decreased by 16.6%, 9.1%, 7.1%, and 0.9% for the SMOSL3, SMOSIC, SMAPL3, and SMAPIB cases, respectively. Conversely, fraction value that passive-based SM provided smaller time-variant error increased by

15.8%, 8.9%, 8%, and 1.5% for the above four cases, respectively. However, this phenomenon is not evident for SMAPIB when compared with the SMOSL3, SMOSIC, and SMAPL3.

Based on the land cover analysis in Fig. 9, we can further investigate the land cover types where TCA was found to underestimate ASCAT errors whereas overestimate passive SM errors. As shown in Fig. 9 (a), this phenomenon was observed in all the six land cover types for the SMOSL3 SM. However, this mismatch phenomenon was only evident in open shrublands and croplands for the SMOSIC SM illustrated in Fig. 9 (b). Although the available samples are relatively few for SMAPL3 in Fig. 9 (c), one can still see that the mismatch phenomenon was obvious in croplands. Compared with the above three SM products, the TCA performance is more complex for the SMAPIB SM. Fig. 9 (d) shows this mismatch phenomenon was prominent in croplands for the SMAPIB case. However, TCA was found to overestimate ASCAT errors and underestimate SMOS errors for the SMAPIB in savannas, woody savannas, and forests when considering ground-based errors as the benchmark.

4 Discussion

It is necessary to consider the rescaling technique used in TCA as the selection of rescaling method has a great impact on the accuracy of the final TCA error estimates, which is generally ignored in current TCA studies. Ground- and TCA-based errors achieve the highest correlation and the smallest RMSE when TCA is applied to SM anomalies with the TCA_Self rescaling technique. Rescaling the inputs against a reference prior to TCA can implant the distribution information of the reference dataset to all TCA inputs. Consequently, the input SM products to TCA may show spurious large cross-correlation and make the resulting TCA errors smaller as well. By contrast, TCA_Self does not affect the original SM observations during the TCA calculation. It provides optimal rescaling parameters as it considers the individual random error properties and matches the variability of the jointly observed signal (Gruber et al., 2017). Although TCA_Self and other suboptimal rescaling techniques can address the first-order (additive) biases for considered datasets. However, the TCA_Self can address the second-order (multiplicative) biases while other suboptimal rescaling techniques cannot (Gruber et al., 2020). This can explain the better performance of TCA_Self than other rescaling techniques.

TCA_Self was also recommended to use in data assimilation to remove the systematic bias between model simulation and observations when compared with VAR and CDF rescaling techniques (Yilmaz and Crow, 2013). However, TCA_Self was applied prior to TCA to rescale SM inputs in Yilmaz and Crow (2013) and error estimates were derived from the rescaled SM time-series. By contrast, the optimal strategy proposed in our work applied TCA_Self after the TCA implementation to rescale the error estimates. These two kinds of approaches are supposed to lead to the same error estimates (Gruber et al., 2016a). Our result is complementary to the study by Yilmaz and Crow (2013) and a synergy result can be drawn that TCA_Self is an optimal rescaling technique not only for removing systematic bias between SM datasets but also for estimating errors using the TCA method.

The ORD results suggest considering and including time-variant SM errors in applications as time-variant errors deviate from time-invariant errors by over 50% when TCA_Self is used. Distinct temporal variability of SM errors at short time scales such as seasonally, monthly, or daily time scales was confirmed in Loew and Schlenz (2011), Zwieback et al. (2012), Su et al. (2014a), and Wu et al. (2021). However, we found the relative magnitude relationship between time-invariant and time-variant errors is not the same for different rescaling techniques used in TCA, which has not been fully addressed by previous studies. Loew and Schlenz (2011) and Wu et al. (2021) found time-invariant errors are greater than time-variant errors. Our results confirmed this conclusion when SM products are rescaled against a reference dataset prior to TCA. However, time-invariant errors are smaller than time-variant errors when error estimates are rescaled by TCA_Self. This result indicates that simple matching technique, including NORM, VAR, and CDF, tends to underestimate time-variant errors when less and less data pairs are considered in the moving-window-based TCA. Moreover, SM inputs are rescaled against a reference dataset in the simple matching techniques, which can significantly influence the magnitude of the final TCA errors (Draper et al., 2013; Dorigo et al., 2015).

Even though we found a strong consistency between TCA- and ground-based evaluations, these two evaluations have distinct differences in two aspects. First, ground-based errors are larger than TCA-based errors for both time-invariant and time-variant cases, which is in line with Yilmaz and Crow (2014) and Dorigo et al. (2015). Second, TCA typically overestimates passive-based SM errors and underestimates ASCAT errors when considering ground-based errors as the benchmark. The better TCA results for ASCAT SM could be due to its relatively better match with GLDAS2 SM as they have more consistent spatial representativeness in grid cells. The spatial representativeness of ASCAT and passive SM datasets is kept since we used the nearest neighbor resampling. This resampling method can preserve original SM values in an unaltered way but may introduce the difference of spatial representativeness between ASCAT and passive SM data in the resampling grid cells.

Moreover, errors derived from SM anomalies instead of raw values were considered in our analysis. The anomalies are more sensitive to capture single events of drying and wetting resulting from rainfall (Dorigo et al., 2010). The coarser spatial resolution of passive-based SM data can act as a lowpass filtering and may make the passive SM datasets less sensitive to the drying and wetting events. By contrast, the finer spatial resolution of ASCAT allows to better capture drying and wetting, and this leads to a better consistency to GLDAS2. This is a plus point for the ASCAT, and consequently TCA deems GLDAS2 and ASCAT SM to provide more consistent observations, which may make the above phenomenon more obvious. It is worth noting that ASCAT is now moving to a 6.25 km grid, and then this comparative advantage of ASCAT may become even more prominent. Considering careful upscaling strategy, such as studies in Albergel et al. (2010), Crow et al. (2012), and Colliander et al. (2017), for ASCAT observations overlapping the GLDAS2 grids may account for this problem. However, the impact of spatial resampling approaches on the TCA error estimates are not reported in current literature and this point of view needs further investigation in future studies. Recently, Gruber et al. (2020) also proposed this concern. This result also reminds researchers to carefully process SM datasets prior to TCA and prudently interpret the resulting errors in their studies.

It is worth mentioning that SMAPIB is a newly developed SM product and it was applied in TCA only in a few studies, such as Zheng et al. (2022). Our analysis suggests that TCA errors derived from SMAPIB have relatively different performances compared with SMOSL3, SMOSIC, and SMAPL3 SM products, such as the relatively low correlations with ground-based errors in the scatterplot of Fig. 4 (a), the correlation and RMSE comparisons in Fig. 5, the RD analysis in Fig. 7, and the binary assessments in Fig. 9. The discrepancy between SMAPIB and other three L-band SM products results from the fact that currently SMAPIB only provides SM observations acquired from the descending (06:00 AM) orbits whereas other SM products are reprocessed and binned into daily observations in our analysis. TCA requires the inputs to describe the same physical quantity (Scipal et al. 2008a) and therefore TCA deems SMAPIB SM to provide larger and distinctive error characterizations when compared with other SM products.

The land cover analysis in Fig. 9 suggests that TCA has relatively less power to efficiently characterize SM errors in croplands. The anthropogenic activities, such as agricultural burning and harvesting, make it challenging to efficiently retrieve SM content in croplands due to the noticeable temporal variability in canopy structure, surface roughness, and diversity of crop types (Lawrence et al., 2014; Momen et al., 2017; Colliander et al., 2017). This leads to uncorrected time-variant errors for SMOS (Patton and Hornbuckle, 2013) and SMAP (Colliander et al., 2017) SM products, and increases the difficulty of decoupling SM content from other environmental information in the backscatter measurements for the ASCAT data. The imperfect description of vegetation cover variability and anthropogenic activities is also a concern for GLDAS2 SM data in croplands. Consequently, TCA- and ground-based errors exhibit less consistency in croplands. Considering more accurate temporal information of ancillary data in SM retrieval algorithms, such as irrigation and harvesting activities, may reduce time-variant SM errors and improve the TCA performance in the croplands.

Both TCA- and ground-based evaluations suggest ASCAT provides more and more reliable observations as vegetation cover increases while passive-based SM has more reliable observations in land cover types featured with a relatively thin vegetation cover. The relatively poor performance of ASCAT data in sparsely vegetated areas may be explained by the fact that ASCAT observations are sensitive to subsurface scatterers during dry soil conditions and this phenomenon can extend into semi-arid environments with sparse to low vegetation cover, which is recently revealed by Wagner et al. (2022). The better performance of SMOS/SMAP than ASCAT data regarding the time-variant errors in our analysis may be explained by the fact that the L-band sensor makes SMOS/SMAP less sensitive to atmospheric effects such as rainfall events (Reul et al., 2012), which is a key variable that influences time-variant errors in SM products (Wu et al., 2021). However, it must be mentioned that

47% of the investigation pixels belong to grasslands in our study and the associated ranking conclusions may not hold true when global landmass pixels are investigated. Different choices in the geographical area, processing, and data screening can lead to different conclusions and rankings of SM products. Besides, it is notable that SMOS and SMAP SM products are up to the latest standard in this work. By contrast, this is not the case for ASCAT SWI data as the Copernicus service did not invest in updating the algorithm recently.

The robustness of applying TCA to temporally interpolated SM inputs was validated and we found this interpolation has a small impact on the final TCA error estimates. This is encouraging as one of the problems in the TCA time-variant scheme is the limited available samples that can be used in a moving-time-window. A small sample size used in TCA would underestimate the true value of random uncertainties (Tsamalis, 2022) and therefore makes the resulting errors unreliable. Our results filled this gap and the temporal interpolation is recommended to be applied in the moving-window-based TCA scheme to address the limited sample issue. Moreover, Chen et al. (2018) applied the bootstrap method to guarantee reliable TCA estimates, and this may be another way to solve the sample issue in TCA.

There are two limitations in this study. First, the value of 0.001 m^3/m^3 was selected as the threshold to distinguish the case that ASCAT and passive-based SM provide comparable errors from other possible cases. This threshold is significantly smaller than the value of 0.005 m^3/m^3 used in Al-Yaari et al. (2014). The evaluation results will change as the threshold value varies. Second, the depth discrepancy is considerable as satellite-based SM products generally present SM content in the topsoil layer of several centimeters whereas the GLDAS2 SM data represent SM simulations in a soil layer depth of 0-10 cm. During the ground-based evaluation, passive SM observations retrieved from L-band brightness temperature data are more consistent with the deeper sub-surface ground measurements (typically deeper than 5 cm) regarding the soil layer. This may make the ground-based evaluation put more trust in the passive SM data compared with the ASCAT data.

5 Conclusions

In this study, we aimed at optimizing and validating the TCA technique with a focus on the rescaling approaches and applied it to compare time-variant errors for ASCAT and four passive-based SM products, i.e., SMOSL3, SMOSIC, SMAPL3, and SMAPIB. Based on the obtained results the following conclusions can be drawn:

(1) Temporal interpolation introduces additional errors in the TCA error estimates. Nevertheless, it has a relatively small impact on the accuracy of the final TCA error estimates.

(2) Rescaling technique has a great impact on the final TCA error estimates. The optimal combination strategy to implement TCA is applying TCA to SM anomalies with the TCA_Self rescaling technique. Errors derived from this optimal combination strategy achieve the highest correlation and the smallest RMSE when considering conventional ground-based errors as the benchmark.

(3) Based on the optimal combination strategy, TCA-based errors are strongly correlated with ground-based errors and their Pearson’s correlation coefficients are

0.62, 0.72, 0.83, 0.89, and 0.93 for SMAPIB, SMAPL3, SMOSIC, SMOSL3, and ASCAT SM data, respectively. The RMSE values are typically smaller than 0.01 m^3/m^3 for the above five SM products.

(4) Considering and including time-variant errors in applications is necessary as time-variant errors typically deviate from time-invariant errors by a value greater than 50%. The relative magnitude between time-invariant and time-variant errors relies on the rescaling technique used in TCA. Time-invariant errors are greater than time-variant errors when SM products are rescaled against a reference dataset prior to TCA. By contrast, time-invariant errors are smaller than time-variant errors when error estimates are rescaled by the TCA_Self after the TCA implementation.

(5) The evaluation performances are strongly consistent for the TCA- and ground-based methods. They provide consistent evaluation results in 74.7% (77.3%), 75.8% (79.8%), 79.6% (81.1%), and 78.6% (79.7%) of the investigation period on global average (median) for the TCA implementations with SMAPL3, SMAPIB, SMOSL3, and SMOSIC SM, respectively. However, they have two evident differences. First, TCA-based errors are mostly smaller than ground-based errors. Second, TCA generally underestimates ASCAT errors and overestimates passive-based SM errors when considering ground-based errors as the benchmark.

(6) Both TCA- and ground-based methods suggest that ASCAT provides more and more reliable SM observations with smaller time-variant errors as vegetation cover increases while passive-based SM, i.e., SMAP and SMOS data here, provides more and more reliable observations with smaller time-variant errors in land cover types

featured with a relatively thin vegetation cover.

(7) Compared with other land cover types, TCA has relatively less power to efficiently characterize SM errors in croplands.

Error characterization is crucial for correctly interpreting and efficiently using SM observations obtained from satellites, hydrological modeling, and ground measurements. Our results reveal that the selection of rescaling techniques has a great impact on the TCA error estimates, which did not receive much attention in current TCA studies. TCA can accurately evaluate SM products only when TCA is properly implemented. Our study is a strong reaffirmation of previous work in related studies that uses TCA and our study is crucial for communities who want to use TCA (and extends) for error estimation, data merging and data assimilation of satellite-based SM data.

Acknowledgments

We thank the editors and three anonymous reviewers for their enthusiastic and constructive suggestions, which leads to numerous clarifications in the revision. This research is funded by the High-Level Talents Introduced to the Hainan University Program. Key techniques to estimate time-variant errors in microwave-based soil moisture data using Triple Collocation Analysis. Grant number: KYQD(ZR)-22084. The SMOSL3 data were obtained from the "Centre Aval de Traitement des Données SMOS" (CATDS), operated for the "Centre National d'Etudes Spatiales" (CNES, France) by IFREMER (Brest, France).

References

Albergel, C., Rüdiger, C., Carrer, D., Calvet, J.C., Fritz, N., Naeimi, V., Bartalis, Z., Hasenauer, S., 2009. An evaluation of ASCAT surface soil moisture products with in-situ observations in Southwestern France. Hydrol. Earth Syst. Sci. 13, 115–124. https://doi.org/10.5194/hess-13-115-2009

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J.C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., Piguet, B., Martin, E., 2008. From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on in-situ observations and model simulations. Hydrol. Earth Syst. Sci. 12, 1323–1337. https://doi.org/10.5194/hess-12-1323-2008

Albergel, C., Calvet, J., Rosnay, P. De, Balsamo, G., Wagner, W., Hasenauer, S., Naeimi, V., Martin, E., 2010. Cross-evaluation of modelled and remotely sensed surface soil moisture with in situ data in southwestern France. Hydrol. Earth Syst. Sci. 2177–2191. https://doi.org/10.5194/hess-14-2177-2010

Albergel, C., Rosnay, P. De, Gruhier, C., Muñoz-sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., Wagner, W., 2012. Evaluation of remotely sensed and modelled soil moisture products using global ground-based in situ observations. Remote Sens. Environ. 118, 215–226. https://doi.org/10.1016/j.rse.2011.11.017

Al Bitar, A., Mialon, A., Kerr, Y.H., Cabot, F., Richaume, P., Jacquette, E., Quesney, A., Mahmoodi, A., Tarot, S., Parrens, M., Al-Yaari, A., Pellarin, T., Rodriguez-Fernandez, N., Wigneron, J.P., 2017. The global SMOS Level 3 daily soil moisture and brightness temperature maps. Earth Syst. Sci. Data 9, 293–315. https://doi.org/10.5194/essd-9-293-2017

Al-Yaari, A., Wigneron, J.P., Ducharne, A., Kerr, Y.H., Wagner, W., De Lannoy, G., Reichle, R., Al Bitar, A., Dorigo, W., Richaume, P., Mialon, A., 2014. Global-scale comparison of passive (SMOS) and active (ASCAT) satellite based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land). Remote Sens. Environ. 152, 614–626. https://doi.org/10.1016/j.rse.2014.07.013

An, R., Zhang, L., Wang, Z., Quaye-ballard, J.A., You, J., Shen, X., Gao, W., Huang, L., Zhao, Y., Ke, Z., 2015. Validation of the ESA CCI soil moisture product in China. Int. J. Appl. Earth Obs. Geoinf. 48, 28-36. https://doi.org/10.1016/j.jag.2015.09.009

Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., and Anderson, C., 2007. Initial soil moisture retrievals from the METOP-A

Advanced Scatterometer (ASCAT). Geophys. Res. Lett. 34, L20401, doi:10.1029/2007GL031088.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F.: ERA-Interim/Land: a global land surface reanalysis data set, Hydrol. Earth Syst. Sci., 19, 389–407, doi:10.5194/hess-19-389-2015, 2015.

Bell, J.E., Palecki, M.A., Baker, C.B., Collins, W.G., Lawrimore, J.H., Leeper, R., Hall, M.E., Kochendorfer, J., Meyers, T.P., Wilson, T., et al., 2013, U.S. Climate Reference Network Soil Moisture and Temperature Observations. J. Hydrometeorol. 14, 977–988. https://doi.org/10.1175/JHM-D-12-0146.1

Biddoccu, M., Ferraris, S., Opsi, F., Cavallo, E., 2016. Long-term monitoring of soil management effects on runoff and soil erosion in sloping vineyards in Alto Monferrato (North–West Italy). Soil Tillage Res. 155, 176–189. https://doi.org/10.1016/j.still.2015.07.005

Bircher, S., Skou, N., Jensen, K.H., Walker, J.P., Rasmussen, L., 2012. A soil moisture and temperature network for SMOS validation in Western Denmark. Hydrol. Earth Syst. Sci. 16, 1445–1463. https://doi.org/10.5194/hess-16-1445-2012

Bolten, J.D., Crow, W.T., Jackson, T.J., Zhan, X., Reynolds, C.A., 2010. Evaluating the Utility of Remotely Sensed Soil Moisture Retrievals for Operational Agricultural Drought Monitoring. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 57–66. https://doi.org/10.1109/JSTARS.2009.2037163

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Hasenauer, S., 2010. ASCAT soil wetness index validation through in situ and modeled soil moisture data in central Italy. Remote Sens. Environ. 114, 2745–2755. https://doi.org/10.1016/j.rse.2010.06.009

Brocca, L., Hasenauer, S., Lacava, T., Melone, F., Moramarco, T., Wagner, W., Dorigo, W., Matgen, P., Martínez-fernández, J., Llorens, P., Latron, J., Martin, C., Bittelli, M., 2011. Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. Remote Sens. Environ. 115, 3390–3408. https://doi.org/10.1016/j.rse.2011.08.003

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Albergel, C., 2013. Scaling and filtering approaches for the use of satellite soil moisture observations. Remote Sensing of Energy Fluxes and Soil Moisture Content. CRC Press. 411–426, http://dx.doi.org/10.1201/b15610-21.

Calvet, J. -C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., & Piguet, B., 2007. In situ soil moisture observations for the CAL/VAL of SMOS: the SMOSMANIA network. IEEE Int. Geosci. Remote Sens. Symp., IGARSS, Barcelona, Spain, 1196–1199, doi:10.1109/IGARSS.2007.4423019.

Cappelaere, B., Descroix, L., Lebel, T., Boulain, N., Ramier, D., Laurent, J., Favreau, G., Boubkraoui, S., Nazoumou, Y., Oi, M., Ottlé, C., Quantin, G., 2009. The AMMA-CATCH experiment in the cultivated Sahelian area of south-west Niger – Investigating water cycle response to a fluctuating climate and changing environment. J. Hydrol. 375, 34–51. https://doi.org/10.1016/j.jhydrol.2009.06.021

Chakravorty, A., Ram, B., Prakash, O., Dhanya, C.T., 2016. A regional scale performance evaluation of SMOS and ESA-CCI soil moisture products over India with simulated soil moisture from MERRA-Land. Remote Sens. Environ. 186, 514–527. https://doi.org/10.1016/j.rse.2016.09.011

Chen, F., Crow, W.T., Bindlish, R., Colliander, A., Burgin, M.S., Asanuma, J., Aida, K., 2018. Global-scale evaluation of SMAP, SMOS and ASCAT soil moisture products using triple collocation. Remote Sens. Environ. 214, 1–13. https://doi.org/10.1016/j.rse.2018.05.008

Chen, F., Crow, W.T., Colliander, A., Cosh, M.H., Jackson, T., Bindlish, R., Reichle, R., Chan, S.K., Bosch, D., Starks, P., Goodrich, D., Seyfried, M., 2016. Application of triple collocation in ground-based validation of soil moisture active/passive (SMAP) level 2 data products. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (2), 489–502. https://doi.org/10.1109/JSTARS.2016.2569998

Chen, F., Crow, W.T., Colliander, A., Cosh, M.H., Jackson, T.J., Bindlish, R., Reichle, R.H., Chan, S.K., Bosch, D.D., Starks, P.J., Goodrich, D.C., Seyfried, M.S., 2017. Application of Triple Collocation in Ground-Based Validation of Soil Moisture Active/Passive (SMAP) Level 2 Data Products. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 489–502. https://doi.org/10.1109/JSTARS.2016.2569998

Colliander, A., Jackson, T.J., Bindlish, R., Chan, S., Das, N., Kim, S.B., Cosh, M.H., Dunbar, R.S., Dang, L., Pashaian, L., Asanuma, J., Aida, K., Berg, A., Rowlandson, T., Bosch, D., Caldwell, T., Caylor, K., Goodrich, D., al Jassar, H., Lopez-Baeza, E., Martínez-Fernández, J., González-Zamora, A., Livingston, S., McNairn, H., Pacheco, A., Moghaddam, M., Montzka, C., Notarnicola, C., Niedrist, G., Pellarin, T., Prueger, J., Pulliainen, J., Rautiainen, K., Ramos, J., Seyfried, M., Starks, P., Su, Z., Zeng, Y., van der Velde, R., Thibeault, M., Dorigo, W., Vreugdenhil, M., Walker, J.P., Wu, X., Monerris, A., O'Neill, P.E., Entekhabi, D., Njoku, E.G., Yueh, S., 2017. Validation of SMAP surface soil moisture products with core validation sites. Remote Sens. Environ. 191, 215–231. https://doi.org/10.1016/j.rse.2017.01.021

Crow, W. T., A. A. Berg, M. H. Cosh, A. Loew, B. P. Mohanty, R. Panciera, P. de Rosnay, D. Ryu, and J. P. Walker, 2012. Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products. Rev. Geophys. 50, RG2002, doi:10.1029/2011RG000372.

Crow, W. T., Su, C.-H., Ryu, D., and Yilmaz, M. T., 2015. Optimal averaging of soil moisture predictions from ensemble land surface model simulations. Water Resour. Res. 51, 9273–9289. https://doi.org/10.1002/ 2015WR016944

Daly, E., Porporato, A., 2005. A review of soil moisture dynamics: From rainfall infiltration to ecosystem response. Environ. Eng. Sci. 22, 9–24. https://doi.org/10.1089/ees.2005.22.9

Dorigo, W.A., Gruber, A., Jeu, R.A.M. De, Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R.M., Kidd, R., 2015. Evaluation of the ESA CCI soil moisture product using ground-based observations. Remote Sens. Environ. 162, 380-395. https://doi.org/10.1016/j.rse.2014.07.023

Dorigo, W.A., Scipal, K., Parinussa, R.M., Liu, Y.Y., Wagner, W., De Jeu, R.A.M., Naeimi, V., 2010. Error characterisation of global active and passive microwave soil moisture datasets. Hydrol. Earth Syst. Sci. 14, 2605–2616. https://doi.org/10.5194/hess-14-2605-2010

Dorigo, W.A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A., Jackson, T., 2011. The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements. Hydrol. Earth Syst. Sci. 15, 1675–1698. https://doi.org/10.5194/hess-15-1675-2011

Dorigo, W., Xaver, A., Vreugdenhil, M., Gruber, A., A, H., Sanchis-Dufau, A., Zamojski, D., Cordes, C., Wagner, W., Drusch, M., 2013. Global automated quality control of in situ soil moisture data from the international soil moisture network. Vadose Zone J. 12 (3). https://doi.org/10.2136/vzj2012.0097

Doubkova, M., Dorigo, W., Hegyiova, A., and Wagner, W., 2012. Evaluation of the ASAR GM soil moisture product, IEEE Int. Geosci. Remote Sens. Symp., 2012, 1227-1230. https://doi.org/10.1109/IGARSS.2012.6351323

Dong, J., Wei, L., Chen, X., Duan, Z., Lu, Y., 2020. An instrument variable based algorithm for estimating cross-correlated hydrological remote sensing errors. J. Hydrol. 581, 124413. https://doi.org/10.1016/j.jhydrol.2019.124413

Draper, C.S., Walker, J.P., Steinle, P.J., de Jeu, R.A.M., Holmes, T.R.H., 2009. An evaluation of AMSR-E derived soil moisture over Australia. Remote Sens. Environ. 113, 703–710. https://doi.org/10.1016/j.rse.2008.11.011

Draper, C., Reichle, R., Jeu, R. De, Naeimi, V., Parinussa, R., Wagner, W., 2013. Estimating root mean square errors in remotely sensed soil moisture over continental scale domains. Remote Sens. Environ. 137, 288–298. https://doi.org/10.1016/j.rse.2013.06.013

Drusch, M., Wood, E.F., Gao, H., 2005. Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. Geophys. Res. Lett. 32, 32–35. https://doi.org/10.1029/2005GL023623

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., Van Zyl, J., 2010. The soil moisture active passive (SMAP) mission. Proc. IEEE 98, 704–716. https://doi.org/10.1109/JPROC.2010.2043918

Famiglietti, J.S., Ryu, D., Berg, A.A., Rodell, M., Jackson, T.J., 2008. Field observations of soil moisture variability across scales. Water Resour. Res., 44, W12602, doi:10.1029/2008WR007323.

Fascetti, F., Pierdicca, N., Pulvirenti, L., Crapolicchio, R., Muñoz-Sabater, J., 2016. A comparison of ASCAT and SMOS soil moisture retrievals over Europe and Northern Africa from 2010 to 2013. Int. J. Appl. Earth Obs. Geoinf. 45, 135–142. https://doi.org/10.1016/j.jag.2015.09.008

Friedl, M.A., McIver, D.K., Hodges, J.C.F., Zhang, X.Y., Muchoney, D., Strahler, A.H., Woodcock, C.E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., Schaaf, C., 2002. Global land cover mapping from MODIS: Algorithms and early results. Remote Sens. Environ. 83, 287–302. https://doi.org/10.1016/S0034-4257(02)00078-0

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., Wagner, W., 2016a. Recent advances in (soil moisture) triple collocation analysis. Int. J. Appl. Earth Obs. Geoinf. 45, 200–211. https://doi.org/10.1016/j.jag.2015.09.002

Gruber, A., Su, C.-H., Crow, W.T., Zwieback, S., Dorigo, W. a, 2016b. Estimating error cross-correlations in soil moisture data sets using extended collocation analysis. J. Geophys. Res. Atmos. 1208–1219. https://doi.org/10.1002/2015JD024027

Gruber, A., Crow, W., Dorigo, W., Wagner, W., 2015. The potential of 2D Kalman filtering for soil moisture data assimilation. Remote Sens. Environ. 171, 137–148. https://doi.org/10.1016/j.rse.2015.10.019

Gruber, A., Dorigo, W.A., Crow, W., Wagner, W., Member, S., 2017. Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. IEEE Trans Geosci Remote Sens. 55(12), 6780–6792. https://doi.org/10.1109/TGRS.2017.2734070.

Gruber, A., Lannoy, G. De, Crow, W., 2019. A Monte Carlo based adaptive Kalman filtering framework for soil moisture data assimilation. Remote Sens. Environ. 228, 105–114. https://doi.org/10.1016/j.rse.2019.04.003

Gruber, A., Paloscia, S., Santi, E., Notarnicola, C., Pasolli, L., Smolander, T., Pulliainen, J., Mittelbach, H., Dorigo, W., Wagner, W., 2014. Performance inter-comparison of soil moisture retrieval models for the MetOp-A ASCAT instrument. Geoscience and Remote Sensing Symposium (IGARSS). 2014, 2455–2458. https://doi.org/10.1109/IGARSS.2014.6946969

Gruber, A., De Lannoy, G., Albergel, C., Al-Yaari, A., Brocca, L., Calvet, J.C., Colliander, A., Cosh, M., Crow, W., Dorigo, W., Draper, C., Hirschi, M., Kerr, Y., Konings, A., Lahoz, W., McColl, K., Montzka, C., Muñoz-Sabater, J., Peng, J., Reichle, R., Richaume, P., Rüdiger, C., Scanlon, T., van der Schalie, R., Wigneron, J.P., Wagner, W., 2020. Validation practices for satellite soil moisture retrievals: What are (the) errors? Remote Sens. Environ. 244, 111806. https://doi.org/10.1016/j.rse.2020.111806

Gruber, A., Dorigo, W.A., Zwieback, S., Xaver, A., Wagner, W., 2013. Characterizing Coarse-Scale Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture Network. Vadose Zo. J. 12, vzj2012.0170. https://doi.org/10.2136/vzj2012.0170

Houser, P.R., Shuttleworth, W.J., Famiglietti, J.S., Gupta, H. V., Syed, K.H., Goodrich, D.C., 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. Water Resour. Res. 34, 3405–3420. https://doi.org/10.1029/1998WR900001

Jackson, T.J., 1993. III. Measuring surface soil moisture using passive microwave remote sensing. Hydrol. Process. 7, 139–152. https://doi.org/10.1002/hyp.3360070205

Kerr, Y.H., Waldteufel, P., Richaume, P., Wigneron, J.P., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S.E., Leroux, D., Mialon, A., Delwart, S., 2012. The SMOS soil moisture retrieval algorithm. IEEE Trans. Geosci. Remote Sens. 50, 1384–1403. https://doi.org/10.1109/TGRS.2012.2184548

Khan, M.Z.K., Mehrotra, R., Sharma, A., Sankarasubramanian, A., 2014. Global sea surface temperature forecasts using an improved multimodel approach. J. Clim. 27, 3505–3515. https://doi.org/10.1175/JCLI-D-13-00486.1

Kim, S., Parinussa, R.M., Liu, Y.Y., Johnson, F.M., Sharma, A., 2016. Merging alternate remotely-sensed soil moisture retrievals using a non-static model combination approach. Remote Sens. 8, 1–16. https://doi.org/10.3390/rs8060518

Kim, H., Parinussa, R., Konings, A.G., Wagner, W., Cosh, M.H., Lakshmi, V., Zohaib, M., Choi, M., 2018. Global-scale assessment and combination of SMAP with ASCAT (active) and AMSR2 (passive) soil moisture products. Remote Sens. Environ. 204, 260–275. https://doi.org/10.1016/j.rse.2017.10.026

Kim, H., Wigneron, J., Kumar, S., Dong, J., Wagner, W., Cosh, M.H., Bosch, D.D., Holifield, C., Starks, P.J., Seyfried, M., Sol, I., Atmosphère, P., Recherche, U.M. De, National, I., Recherche, D., Inrae, A., 2020. Global scale error assessments of soil moisture estimates from microwave-based active and passive satellites and land surface models over forest and mixed irrigated/dryland agriculture regions. Remote Sens. Environ. 251, 112052. https://doi.org/10.1016/j.rse.2020.112052

Kim, S., Sharma, A., Liu, Y.Y., Young, S.I., 2022. Rethinking Satellite Data Merging: From Averaging to SNR Optimization. IEEE Trans. Geosci. Remote Sens. 60. https://doi.org/10.1109/TGRS.2021.3107028

Kim, S., Dong, J., Sharma, A., 2021. A Triple Collocation-Based Comparison of Three L-Band Soil Moisture Datasets, SMAP, SMOS-IC, and SMOS, Over Varied Climates and Land Covers. Front. Water 3, 1–12. https://doi.org/10.3389/frwa.2021.693172

Larson, K. M., Small, E. E., Gutmann, E. D., Bilich, A. L., Braun, J. J., & Zavorotny, V. U., 2008. Use of GPS receivers as a soil moisture network for water cycle studies. Geophysical Research Letters, 35, 24. https://doi.org/10.1029/2008GL036013

Lawrence, H., Wigneron, J.P., Richaume, P., Novello, N., Grant, J., Mialon, A., Al Bitar, A., Merlin, O., Guyon, D., Leroux, D., Bircher, S., Kerr, Y., 2014. Comparison between SMOS Vegetation Optical Depth products and MODIS vegetation indices over crop zones of the USA. Remote Sens. Environ. 140, 396–406. https://doi.org/10.1016/j.rse.2013.07.021

Leroux, D.J., Kerr, Y.H., Richaume, P., Fieuzal, R., 2013. Spatial distribution and possible sources of SMOS errors at the global scale. Remote Sens. Environ. 133, 240–250. https://doi.org/10.1016/j.rse.2013.02.017

Li, X., Wigneron, J.P., Fan, L., Frappart, F., Yueh, S.H., Colliander, A., Ebtehaj, A., Gao, L., Fernandez-Moran, R., Liu, X., Wang, M., Ma, H., Moisy, C., Ciais, P., 2022. A new SMAP soil moisture and vegetation optical depth product (SMAP-IB): Algorithm, assessment and inter-comparison. Remote Sens. Environ. 271, 112921. https://doi.org/10.1016/j.rse.2022.112921

Li, H., Chai, L., Crow, W., Dong, J., Liu, S., Zhao, S., 2022. The reliability of categorical triple collocation for evaluating soil freeze/thaw datasets. Remote Sens. Environ. 281, 113240. https://doi.org/10.1016/j.rse.2022.113240

Liu, J., Chai, L., Dong, J., Zheng, D., Wigneron, J.P., Liu, S., Zhou, J., Xu, T., Yang, S., Song, Y., Qu, Y., Lu, Z., 2021. Uncertainty analysis of eleven multisource soil moisture products in the third pole environment based on the three-corned hat method. Remote Sens. Environ. 255, 112225. https://doi.org/10.1016/j.rse.2020.112225

Loew, A., Schlenz, F., 2011. A dynamic approach for evaluating coarse scale satellite soil moisture products. Hydrol. Earth Syst. Sci. 15, 75–90. https://doi.org/10.5194/hess-15-75-2011

Mattar, C., Santamaría-Artigues, A., Durán-Alarcón, C., Olivera-Guerra, L., Fuster, R., Borvarán, D., 2016. The LAB-net soil moisture network: application to thermal remote sensing and surface energy balance. Data 1, 6. https://doi.org/10.3390/data1010006

Miralles, D.G., Crow, W.T., Cosh, M.H., 2010. Estimating spatial sampling errors in coarse-scale soil moisture estimates derived from point-scale observations. J. Hydrometeorol. 11, 1423–1429. https://doi.org/10.1175/2010JHM1285.1

Miyaoka, K., Gruber, A., Ticconi, F., Hahn, S., Wagner, W., Figa-Saldana, J., Anderson, C., 2017. Triple Collocation Analysis of Soil Moisture from Metop-A ASCAT and SMOS Against JRA-55 and ERA-Interim. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 2274–2284. https://doi.org/10.1109/JSTARS.2016.2632306

Moghaddam, M., Entekhabi, D., Member, S., Goykhman, Y., Li, K., Liu, M., Member, S., Mahajan, A., Nayyar, A., Shuman, D., 2010. A Wireless Soil Moisture Smart Sensor Web Using Physics-Based Optimal Control: Concept and Initial Demonstrations. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 522–535. https://doi.org/10.1109/JSTARS.2010.2052918

Mougin, E., Hiernaux, P., Kergoat, L., Grippa, M., Rosnay, P. De, Timouk, F., Dantec, V. Le, Demarez, V., Lavenu, F., Arjounin, M., Lebel, T., Soumaguel, N., Ceschia, E., Mougenot, B., Baup, F., Frappart, F., Frison, P.L., Gardelle, J., Gruhier, C., Jarlan, L., Mangiarotti, S., Sanou, B., Tracol, Y., Guichard, F., Serça, D., Galy-lacaux, C., Seghieri, J., Becerra, S., Dia, H., Gangneron, F., Mazzega, P., 2009. The AMMA-CATCH Gourma observatory site in Mali: Relating climatic variations to changes in vegetation, surface hydrology, fluxes and natural resources. J. Hydrol. 375, 14–33. https://doi.org/10.1016/j.jhydrol.2009.06.045

Momen, M., Wood, J.D., Novick, K.A., Pangle, R., Pockman, W.T., McDowell, N.G., Konings, A.G., 2017. Interacting Effects of Leaf Water Potential and Biomass on

Vegetation Optical Depth. J. Geophys. Res. Biogeosciences 122, 3031–3046. https://doi.org/10.1002/2017JG004145

Molero, B., Leroux, D.J., Richaume, P., Kerr, Y.H., Merlin, O., Cosh, M.H., Bindlish, R., 2018. Multi-Timescale Analysis of the Spatial Representativeness of In Situ Soil Moisture Data within Satellite Footprints. J. Geophys. Res. Atmos. 123, 3–21. https://doi.org/10.1002/2017JD027478

Narasimhan, B., Srinivasan, R., 2005. Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. Agric. For. Meteorol. 133, 69–88. https://doi.org/10.1016/j.agrformet.2005.07.012

Osenga, E.C., Arnott, J.C., Endsley, K.A., Katzenberger, J.W., 2019. Bioclimatic and soil moisture monitoring across elevation in a mountain watershed: opportunities for research and resource management. Water Resour. Res. 55, 2493– 2503. https://doi.org/10.1029/2018WR023653

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, R. Bindlish, and J. Chaubell. (2021). SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 8 [Data Set]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/OMHVSRGFX38O. Date Accessed 07-20-2022.

Paulik, C., Dorigo, W., Wagner, W., Kidd, R., 2014. Validation of the ASCAT soil water index using in situ data from the International Soil moisture network. Int. J. Appl. Earth Obs. Geoinf. 30, 1–8. https://doi.org/10.1016/j.jag.2014.01.007

Patton, J., Hornbuckle, B., 2013. Initial validation of SMOS vegetation optical thickness in Iowa. IEEE Geosci. Remote Sens. Lett. 10, 3791–3794. https://doi.org/10.1109/IGARSS.2012.6350491

Pellarin, T., Laurent, J.P., Cappelaere, B., Decharme, B., Descroix, L., Ramier, D., 2009. Hydrological modelling and associated microwave emission of a semi-arid region in South-western Niger. J. Hydrol. 375, 262–272. https://doi.org/10.1016/j.jhydrol.2008.12.003

Peng, J., Tanguy, M., Robinson, E.L., Pinnington, E., Evans, J., Ellis, R., Cooper, E., Hannaford, J., Blyth, E., Dadson, S., 2021. Estimation and evaluation of high-resolution soil moisture from merged model and Earth observation data in the Great Britain. Remote Sens. Environ. 264, 112610. https://doi.org/10.1016/j.rse.2021.112610

Pierdicca, N., Fascetti, F., Pulvirenti, L., Crapolicchio, R., Muñoz-sabater, J., 2015. Analysis of ASCAT, SMOS, in-situ and land model soil moisture as a regionalized variable over Europe and North Africa. Remote Sens. Environ. 170, 280–289. https://doi.org/10.1016/j.rse.2015.09.005

Pierdicca, N., Fascetti, F., Pulvirenti, L., Crapolicchio, R., 2017. Error Characterization of Soil Moisture Satellite Products: Retrieving Error Cross-Correlation Through Extended Quadruple Collocation. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 4522–4530. https://doi.org/10.1109/JSTARS.2017.2714025

Reichle, R.H., Koster, R.D., 2004. Bias reduction in short records of satellite soil moisture. Geophys. Res. Lett. 31. L19501. https://doi.org/10.1029/2004GL020938

Reichle, R.H., 2008. Data assimilation methods in the Earth sciences. Adv. Water Resour. 31, 1411–1418. https://doi.org/10.1016/j.advwatres.2008.01.001

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., Toll, D., 2004. The Global Land Data Assimilation System. Bull. Am. Meteorol. Soc. 85, 381–394. https://doi.org/10.1175/BAMS-85-3-381

Rosnay, P. De, Gruhier, C., Timouk, F., Baup, F., Mougin, E., Hiernaux, P., Kergoat, L., Ledantec, V., 2009. Multi-scale soil moisture measurements at the Gourma meso-scale site in Mali. J. Hydrol. 375, 241–252. https://doi.org/10.1016/j.jhydrol.2009.01.015

Rüdiger, C., Calvet, J., Gruhier, C., Holmes, T. R. H., de Jeu, R. A. M., & Wagner, W., 2009. An Intercomparison of ERS-Scat and AMSR-E Soil Moisture Observations with Model Simulations over France. J. Hydrometeorol. 10(2), 431-447.

Reul, N., Tenerelli, J., Chapron, B., Vandemark, D., Quilfen, Y., and Kerr, Y., 2012. SMOS satellite L-band radiometer: A new capability for ocean surface remote sensing in hurricanes. J. Geophys. Res. 117, C02006, doi:10.1029/2011JC007474.

Scipal, K., Holmes, T., De Jeu, R., Naeimi, V., Wagner, W., 2008a. A possible solution for the problem of estimating the error structure of global soil moisture data sets. Geophys. Res. Lett. 35, 2–5. https://doi.org/10.1029/2008GL035599

Scipal, K., Drusch, M., Wagner, W., 2008b. Assimilation of a ERS scatterometer derived soil moisture index in the ECMWF numerical weather prediction system. Adv. Water Resour. 31, 1101–1112. https://doi.org/10.1016/j.advwatres.2008.04.013.

Scipal, K., Scheffler, C., Wagner, W., 2005. Soil moisture-runoff relation at the catchment scale as observed with coarse resolution microwave remote sensing. Hydrol. Earth Syst. Sci. 9, 173–183.

Scipal, K., Dorigo, W., De Jeu, R., 2010. Triple collocation - A new tool to determine the error structure of global soil moisture products. Int. Geosci. Remote Sens. Symp. 3, 4426–4429. https://doi.org/10.1109/IGARSS.2010.5652128.

Sims, A.P., Niyogi, D.D.S., Raman, S., 2002. Adopting drought indices for estimating soil moisture: A North Carolina case study. Geophys. Res. Lett. 29, 24-1-24–4. https://doi.org/10.1029/2001GL013343

Smith, A.B., Walker, J.P., Western, A.W., Young, R.I., Ellett, K.M., Pipunic, R.C., Grayson, R.B., Siriwardena, L., Chiew, F.H.S., Richter, H., 2012. The Murrumbidgee Soil Moisture Monitoring Network data set. Water Resour. Res. 48, 1–6. https://doi.org/10.1029/2012WR011976.

Stoffelen, A., 1998. Toward the true near-surface wind speed: Error modeling and calibration using triple collocation. J. Geophys. Res. C Ocean. 103, 7755–7766. https://doi.org/10.1029/97jc03180

Su, C., Ryu, D., Crow, W.T., Western, A.W., 2014a. Stand-alone error characterisation of microwave satellite soil moisture using a Fourier method. Remote Sens. Environ. 154, 115–126. https://doi.org/10.1016/j.rse.2014.08.014

Su, C.H., Ryu, D., Crow, W.T., Western, A.W., 2014b. Beyond triple collocation: Applications to soil moisture monitoring. J. Geophys. Res. 119, 6419–6439. https://doi.org/10.1002/2013JD021043

Su, C., Ryu, D., Young, R.I., Western, A.W., Wagner, W., 2013. Inter-comparison of microwave satellite soil moisture retrievals over the Murrumbidgee Basin, southeast Australia. Remote Sens. Environ. 134, 1–11. https://doi.org/10.1016/j.rse.2013.02.016

Tagesson, T., Fensholt, R., Guiro, I., Rasmussen, M.O., Huber, S., Mbow, C., Garcia, M., Horion, S., Sandholt, I., Holm-Rasmussen, B., et al., 2015. Ecosystem properties of semiarid savanna grassland in West Africa and its relationship with environmental variability. Glob. Chang. Biol. 21, 250–264. https://doi.org/10.1111/gcb.12734

Tsamalis, C., 2022. Clarifications on the equations and the sample number in triple collocation analysis using SST observations. Remote Sens. Environ. 272, 112936. https://doi.org/10.1016/j.rse.2022.112936

Tugrul Yilmaz, M., Crow, W.T., 2013. The optimality of potential rescaling approaches in land data assimilation. J. Hydrometeorol. 14, 650–660. https://doi.org/10.1175/JHM-D-12-052.1

Tugrul Yilmaz, M., Crow, W.T., 2014. Evaluation of assumptions in soil moisture triple collocation analysis. J. Hydrometeorol. 15, 1293–1302. https://doi.org/10.1175/JHM-D-13-0158.1

Ulaby, F.T., Dobson, M.C., Brunfeldt, D.R., 1983. Improvement of Moisture Estimation Accuracy of Vegetation-Covered Soil by Combined Active/Passive Microwave Remote Sensing. IEEE Trans. Geosci. Remote Sens. GE-21, 300–307. https://doi.org/10.1109/TGRS.1983.350557

Wagner, W., Lemoine, G., Rott, H., 1999. A method for estimating soil moisture from ERS Scatterometer and soil data. Remote Sens. Environ. 70, 191–207. https://doi.org/10.1016/S0034-4257(99)00036-X

Wagner, W., Lindorfer, R., Melzer, T., Hahn, S., Bauer-marschallinger, B., Morrison, K., Calvet, J., Hobbs, S., Quast, R., Greimeister-pfeil, I., Vreugdenhil, M., 2022. Widespread occurrence of anomalous C-band backscatter signals in arid environments caused by subsurface scattering. Remote Sens. Environ. 276, 113025. https://doi.org/10.1016/j.rse.2022.113025

Wang, X., Lü, H., Crow, W.T., Zhu, Y., Wang, Q., Su, J., Zheng, J., Gou, Q., 2021. Assessment of SMOS and SMAP soil moisture products against new estimates combining physical model , a statistical model , and in-situ observations : A case study over the Huai River Basin , China. J. Hydrol. 598, 126468. https://doi.org/10.1016/j.jhydrol.2021.126468

Wanders, N., Karssenberg, D., De Roo, A., De Jong, S.M., Bierkens, M.F.P., 2014. The suitability of remotely sensed soil moisture for improving operational flood forecasting. Hydrol. Earth Syst. Sci. 18, 2343–2357. https://doi.org/10.5194/hess-18-2343-2014

Western, A.W., Grayson, R.B., Blöschl, G., 2002. Scaling of Soil Moisture: A Hydrologic Perspective. Annu. Rev. Earth Planet. Sci. 30, 149–180. https://doi.org/10.1146/annurev.earth.30.091201.140434

Wigneron, J.P., Li, X., Frappart, F., Fan, L., Al-Yaari, A., De Lannoy, G., Liu, X., Wang, M., Le Masson, E., Moisy, C., 2021. SMOS-IC data record of soil moisture and L-VOD: Historical development, applications and perspectives. Remote Sens. Environ. 254. https://doi.org/10.1016/j.rse.2020.112238

Wu, K., Ryu, D., Nie, L., Shu, H., 2021. Time-variant error characterization of SMAP and ASCAT soil moisture using Triple Collocation Analysis. Remote Sens. Environ. 256, 112324. https://doi.org/10.1016/j.rse.2021.112324

Wu, K., Shu, H., Nie, L., Jiao, Z., 2018. Triple collocation-based estimation of spatially correlated observation error covariance in remote sensing soil moisture data assimilation. J. Appl. Remote Sens. 12, 1–19. https://doi.org/10.1117/1.JRS.12

Wu, K., Nie, L., Shu, H., 2020. A comparison of SMAP and SMOS L-band brightness temperature observations over the global landmass. Int. J. Remote Sens. 41, 399–419. https://doi.org/10.1080/01431161.2019.1641759

Wu, X., Xiao, Q., Wen, J., You, D., 2019. Direct Comparison and Triple Collocation: Which Is More Reliable in the Validation of Coarse-Scale Satellite Surface Albedo Products. J. Geophys. Res. Atmos. 124, 5198–5213. https://doi.org/10.1029/2018JD029937

Xu, L., Chen, N., Zhang, X., Moradkhani, H., Zhang, C., Hu, C., 2021. In-situ and triple-collocation based evaluations of eight global root zone soil moisture products. Remote Sens. Environ. 254, 112248. https://doi.org/10.1016/j.rse.2020.112248

Yang, K., Qin, J., Zhao, L., Chen, Y., Tang, W., Han, M., Chen, Z., Lv, N., Ding, B., Wu, H., & Lin, C., 2013. A Multiscale Soil Moisture and Freeze–Thaw Monitoring Network on the Third Pole. Bull. Amer. Meteor. Soc. 94(12), 1907-1916. https://doi.org/10.1175/BAMS-D-12-00203.1

Zacharias, S., Bogena, H., Samaniego, L., Mauder, M., Fuß, R., Pütz, T., Frenzel, M., Schwank, M., Baessler, C., Butterbach-Bahl, K., Bens, O., Borg, E., Brauer, A., Dietrich, P., Hajnsek, I., Helle, G., Kiese, R., Kunstmann, H., Klotz, S., Munch, J.C., Papen, H., Priesack, E., Schmid, H.P., Steinbrecher, R., Rosenbaum, U., Teutsch, G., Vereecken, H., 2011. A Network of Terrestrial Environmental Observatories in Germany. Vadose Zo. J. 10, 955–973. https://doi.org/10.2136/vzj2010.0139.

Zhang, R., Kim, S., Sharma, A., Lakshmi, V., 2021. Identifying relative strengths of SMAP, SMOS-IC, and ASCAT to capture temporal variability. Remote Sens. Environ. 252, 112126. https://doi.org/10.1016/j.rse.2020.112126

Zheng, J., Zhao, T., Lü, H., Shi, J., Cosh, M.H., Ji, D., Jiang, L., Cui, Q., Lu, H., Yang, K., Wigneron, J.P., Li, X., Zhu, Y., Hu, L., Peng, Z., Zeng, Y., Wang, X., Kang, C.S., 2022. Assessment of 24 soil moisture datasets using a new in situ network in the Shandian River Basin of China. Remote Sens. Environ. 271. https://doi.org/10.1016/j.rse.2022.112891

Zhuang, R., Zeng, Y., Manfreda, S., Su, Z., 2020. Quantifying long-term land surface and root zone soil moisture over Tibetan plateau. Remote Sens. 12, 509. https://doi.org/10.3390/rs12030509

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R., COSMOS: the COsmic-ray Soil Moisture Observing System. Hydrol. Earth Syst. Sci., 16, 4079–4099, https://doi.org/10.5194/hess-16-4079-2012, 2012.

Zwieback, S., Colliander, A., Cosh, M.H., Martínez-Fernández, J., McNairn, H., Starks, P.J., Thibeault, M., Berg, A., 2018. Estimating time-dependent vegetation biases in the SMAP soil moisture product. Hydrol. Earth Syst. Sci. 22, 4473–4489. https://doi.org/10.5194/hess-22-4473-2018

Zwieback, S., Dorigo, W., Wagner, W., 2012. Temporal error variability of coarse scale soil moisture products - Case study in central Spain. Int. Geosci. Remote Sens. Symp. 722–725. https://doi.org/10.1109/IGARSS.2012.6351463

Supplementary

Matlab codes with an example to estimate time-variant errors based on the TCA_Self rescaling technique is provided along with this manuscript. Please cite this manuscript if you intend to use the provided Matlab codes.

List of Figure Captions

Fig. 1 Spatial distribution of 759 grid cells that include selected ISMN stations.

Fig. 2 Boxplots of correlation coefficients (red) and RMSD values (blue) between time-variant errors derived from the original SM data without interpolation and the SM time series resampled by different percentage values of the original data. (a-d) exhibits the results of applying interpolation to SMOSL3, SMAPL3, SMOSIC, and ASCAT SM data, respectively. The x-axis denotes the percentage values of the original data considered in the resampling. The y-axes on the left (red color) and right (blue color) describe correlation coefficients and RMSD, respectively. The 'N' above each subfigure denotes the number of available samples included in the corresponding boxplot.

Fig. 3 Boxplots of the Pearson's correlation coefficients (red) and RMSE values (blue) between ground- and TCA-based time-variant errors for the TCA implementations that consider (a) SMOSL3, (b) SMOSIC, (c) SMAPL3, (d) SMAPIB, and (e) ASCAT as the triplet inputs. In each subfigure, results derived from absolutes and anomalies are shown in the white and grey areas, respectively. The x-axis is the multiple rescaling techniques considered in the TCA. The y-axes on the left (red color) and right (blue color) describe correlation coefficients and RMSE, respectively.

Fig. 4 Scatterplots of ground- and TCA-based time-variant errors for (a) SMAPIB, (b) SMAPL3, (c) SMOSIC, (d) SMOSL3, and (e) ASCAT SM data. The original scatter points are binned in the increase of x-axis for ground-based error and y-axis for TCA-based error, respectively, and colored based on the number of points included in the bin. All the ρ values passed a t-test ($\alpha < 0.05$). N describes the number of experimental grid cells considered in each scatterplot.

Fig. 5 Boxplots of (a) correlation coefficients and (b) RMSE between TCA- and ground-based errors classified by six land cover types for ASCAT and multiple passive SM products. Boxplots with limited available samples (< 30) are not considered in such comparisons.

Fig. 6 Mean and median of the ORD values derived from the NORM, VAR, CDF, TCA_Self, and QCA_Self rescaling techniques for SMAPL3, SMAPIB, SMOSL3, SMOSIC, and ASCAT SM datasets. (a) and (b) represent ORD values obtained from the TCA- and ground-based evaluations, respectively.

Fig. 7 Fractions of the RD values that are greater (red) or smaller (blue) than 0 for the combination of five SM datasets, i.e., ASCAT, SMOSIC, SMOSL3, SMAPIB, and SMAPL3, and five rescaling techniques, i.e., QCA_Self, TCA_Self, CDF, VAR, and NORM. (a) and (b) present associated results based on the RD values derived from the TCA- and ground-based evaluations, respectively.

Fig. 8 Boxplots of a binary assessment of the evaluation consistency between TCA- and ground-based methods for ASCAT and four passive-based SM time-variant errors. The four columns illustrated in this figure denote the four TCA implementations that consider [GLDAS2, ASCAT, SMAPL3], [GLDAS2, ASCAT, SMAPIB], [GLDAS2, ASCAT, SMOSL3], and [GLDAS2, ASCAT, SMOSIC] as the triplet.

Fig. 9 Boxplots of the percentage values classified by six land cover types for the three cases: ASCAT (green boxes) or passive SM (blue boxes) has the smallest time-variant error and the difference between their errors is smaller than 0.001 m^3/m^3 (yellow boxes). (a-d) represent the TCA- (white areas) and ground-based (grey areas) assessments that consider SMOSL3, SMOSIC, SMAPL3, and SMAPIB SM data, respectively. To guarantee reliable results, boxplots associated with limited available samples (< 30) are excluded. The two pies below each subfigure exhibit the overall performance of the three cases that include all available samples for the TCA- (on the left) and ground-based (on the right) assessments.

List of Table Captions

Table 1 A summary of the SM products used in this work

Table 2 ISMN ground stations from sparse networks

Table 3 Pearson's correlation coefficients between ground- and TCA-based time-invariant errors for SMOSL3, SMOSIC, SMAPL3, SMAPIB, and ASCAT SM products. The time-invariant errors are derived from multiple rescaling techniques. The first and second recommendation strategies (excluding QCA_Self) to implement TCA are highlighted with green and yellow colors, respectively. All the ρ values passed a t-test ($\alpha < 0.05$)

Table 4 The same as Table 3 but for the RMSE values