Assessment of 24 soil moisture datasets using a new in situ network in the Shandian River Basin of China

Jingyao Zheng, Tianjie Zhao, Haishen Lü, Jiancheng Shi, Michael H Cosh, Dabin Ji, Lingmei Jiang, Qian Cui, Hui Lu, Kun Yang, et al.

Remote Sensing of Environment 271 (2022) 112891

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Assessment of 24 soil moisture datasets using a new *in situ* network in the Shandian River Basin of China

Jingyao Zheng [a], Tianjie Zhao [b,*], Haishen Lü [a], Jiancheng Shi [c], Michael H. Cosh [d], Dabin Ji [b], Lingmei Jiang [e], Qian Cui [f], Hui Lu [g], Kun Yang [g], Jean-Pierre Wigneron [h], Xiaojun Li [h], Yonghua Zhu [a], Lu Hu [i], Zhiqing Peng [b], Yelong Zeng [b], Xiaoyi Wang [a], Chuen Siang Kang [j]

[a] *State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, National Cooperative Innovation Center for Water Safety and Hydro-science, College of Hydrology and Water Resources, Hohai University, China*
[b] *State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, China*
[c] *National Space Science Center, Chinese Academy of Sciences, China*
[d] *USDA Agricultural Research Service, Beltsville Agricultural Research Center, Hydrology and Remote Sensing Laboratory, 10300 Baltimore Avenue, Beltsville, MD 20705, United States*
[e] *Faculty of Geographical Science, Beijing Normal University, China*
[f] *Information Center of Ministry of Water Resources of China, China*
[g] *Department of Earth System Science, Tsinghua University, China*
[h] *INRAE, UMR1391 ISPA, F-33140 Villenave d'Ornon, France*
[i] *International Institute for Earth System Science, Nanjing University, China*
[j] *Department of Geoinformation, Faculty of Built Environment & Surveying, Universiti Teknologi Malaysia, Malaysia*

ARTICLE INFO

Editor: Jing M. Chen

Keywords:
Validation
In situ network
Satellite-based soil moisture
Model-based soil moisture
Triple collocation
Climate change initiative
Copernicus climate change service

ABSTRACT

A new soil moisture and soil temperature wireless sensor network (the SMN-SDR) consisting of 34 sites was established within the Shandian River Basin in 2018, located in a semi-arid area of northern China. In this study, *in situ* measurements of the SMN-SDR were used to evaluate 24 different soil moisture datasets grouped according to three categories: (1) single-sensor satellite-based products, (2) multi-sensor merged products, and (3) model-based products. Triple collocation analysis (TCA) was applied to all possible triplets to verify the reliability and robustness of the results. Impacts of different factors on the accuracy of soil moisture products were also investigated, including local acquisition time, physical surface temperature, and vegetation optical depth (VOD). The results reveal that the latest Climate Change Initiative (CCI) -combined product (v06.1, merging extra low-frequency passive microwave data) had the best agreement with *in situ* measurements from the SMN-SDR, with the lowest ubRMSE (< 0.04 m^3/m^3) and highest R (> 0.6). Among all single-sensor retrieved soil moisture products, the Soil Moisture Active Passive (SMAP) products performed best in terms of R (> 0.6) and ubRMSE (close to 0.04 m^3/m^3), with the SMAP-MDCA (Modified Dual Channel Algorithm) being slightly better than the baseline SCA-V (Single Channel Algorithm-Vertical polarization). Importantly, the newly developed SMAP-IB product, which does not use auxiliary data, delivered the best bias statistics and higher VOD values compared with the drier SMAP retrievals, suggesting that the low VOD values (underestimated vegetation effects) may be the major factor causing the dry bias of SMAP products in this study area. It was also found that TCA may systematically overestimate the correlation and underestimate the ubRMSE of soil moisture products as compared with ground-based metrics. TCA-based metrics may vary considerably when using different triplets, due to the TCA assumptions being violated even with the most conservative triplets (in this case an active product, a passive product, and a model-based product). Redundant TCA-based metrics from multiple independent triplets could be averaged to increase the accuracy of final TCA estimates. This study is the first to use *in situ* measurements from the SMN-SDR to conduct a comprehensive evaluation of commonly used, multi-source soil moisture products. These results are expected to further promote the improvement of satellite- and model-based soil moisture products.

* Corresponding author at: State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China.
E-mail address: zhaotj@aircas.ac.cn (T. Zhao).

https://doi.org/10.1016/j.rse.2022.112891
Received 16 July 2021; Received in revised form 31 December 2021; Accepted 3 January 2022
Available online 13 January 2022

1. Introduction

Soil moisture is a key element of the terrestrial water cycle, controlling various hydrological, ecological, and biogeochemical processes (Babaeian et al., 2019). The acquisition of accurate surface soil moisture data is critical for many applications, including drought monitoring, crop yield estimation, flood prediction, and heat-wave forecasting (Jackson et al., 2010).

Considering the strong spatial heterogeneity of soil moisture, point-scale *in situ* measurements can only obtain precise information over a very small area. Accordingly, *in situ* measurements are usually too sparse to specifically describe the spatial-temporal variation of soil moisture over large scales (Crow et al., 2012). To meet research and application needs, various global-scale soil moisture products have been developed, which can be roughly divided into the following three major groups:

(1) Remotely sensed soil moisture directly retrieved from active/passive satellite microwave observations. Microwave remote sensing, especially passive microwave, has become an effective way to estimate soil moisture at large scales due to its strong sensitivity to soil moisture, as well as the advantages of short-term access and all-weather monitoring. Several satellite platforms carrying passive/active microwave sensors have been successfully operated for the past decades, including (but not limited to) low-frequency passive microwave sensors such as SMOS (Soil Moisture and Ocean Salinity) (Kerr et al., 2010) and SMAP (Soil Moisture Active Passive) (Entekhabi et al., 2010), high-frequency passive microwave sensors such as AMSR2 (Advanced Microwave Scanning Radiometer-2) (Imaoka et al., 2010) and Chinese Fengyun series satellites (Zhu et al., 2019), and active microwave sensors such as ASCAT (the Advanced Scatterometer) (Wagner et al., 2013) and Sentinel-1 (Paloscia et al., 2013). Based on the observed passive (brightness temperature, TB) and/or active (backscatter coefficient) microwave signals, these platforms have developed corresponding retrieval algorithms and released soil moisture products to the public.

(2) Merged soil moisture products estimated by blending multiple separately released microwave remotely sensed products. The most typical case is the CCI (Climate Change Initiative) project (Dorigo et al., 2015).

(3) Model-derived soil moisture, including (but not limited to) NASA's Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004), NASA's Modern-Era Retrospective analysis for Research and Applications (MERRA), and the European Center for Medium-Range Weather Forecasts (ECMWF) Land Surface Reanalysis v5-Land (ERA5-Land) soil moisture products.

There are different retrieval procedures, design accuracy, input datasets, and spatial-temporal resolution between the different soil moisture products, with their accuracy affected by different factors. For instance, the quality of satellite-based soil moisture products is hindered by retrieval algorithms, microwave instrument characteristics, topography, climate, and land cover conditions, while the model structure, input parameters, and atmospheric forcing errors are the main factors for the uncertainty of model-based products (Chen et al., 2013). Hence, performance evaluation of soil moisture products before use is incredibly important, since it can both guide practical application and improve the corresponding retrieval process.

The validation of soil moisture products is commonly based on comparison with reference data, which is assumed to be a good representation of ground truth, aiming to calculate their systematic and random errors and obtain quantitative statistics about the quality of soil moisture products (Gruber et al., 2020). The reference data commonly used for satellite soil moisture validation can be broadly classified into the following categories: (1) *in situ* measurements, (2) model simulations, and (3) satellite products. Each type of reference data has its limitations. *In situ* measurements can be divided into (1) field and airborne campaigns during discrete periods, and (2) long-term *in situ* networks from ground monitoring stations. The inevitable differences in spatial-temporal scale, and the construction of a dense network which is labor-intensive (*e.g.* repeated measuring) and costly, are the main challenges for ground-based evaluation (Babaeian et al., 2019). Numerous studies have assessed soil moisture products by comparing them with remotely sensed retrievals or model simulations of similar spatial-temporal resolution (Al-Yaari et al., 2014; Chen et al., 2018). Although this approach is capable of solving scale issues and achieving a global-scale assessment, it can only provide relative conclusions between different soil moisture products, since these non-fiducial reference data may inherently contain unknown biases and uncertainties (Ma et al., 2019).

The random error variance of three collocated datasets can be estimated through triple collocation analysis (TCA), without the need for a single high-quality reference data, which is why TCA has been widely used in soil moisture validation studies (Chen et al., 2018; Dorigo et al., 2015; Kim et al., 2020). The use of TCA requires some basic assumptions (see Section 3.3), which are often violated. Many studies have been devoted to diagnosing the robustness of TCA, including: (1) implementing multiple independent triplets on the same soil moisture product and checking their consistency, TCA estimates should be identical if the basic assumptions are not violated (Gruber et al., 2020); (2) applying the least-squares quadruple collocation analysis (QCA) or more data sets to calculate the error cross-correlations (Chen et al., 2018; Pierdicca et al., 2017); and (3) comparing the consistency of performance metrics through TCA and dense *in situ* measurements (Dorigo et al., 2015; Kim et al., 2020).

In recent years, scientists have carried out considerable soil moisture validation work through different methods, including field campaigns, *in situ* networks, mutual-comparison of other remotely sensed retrievals or model-based products, and calculation of TCA- and QCA-based metrics (Chen et al., 2018; Wu et al., 2014; Ye et al., 2019, 2021). However, only a few validation studies have had a large quantity (> 5) of soil moisture products (Al-Yaari et al., 2019; Beck et al., 2021; Sabaghy et al., 2020). Besides, with the rapid development of soil moisture retrieval algorithms, there are new soil moisture products (SMAP-IB) and the latest versions (newly-released CCI and SMOS-IC products) that have not been assessed.

Inter-comparison between soil moisture products and dense *in situ* measurements (namely fiducial reference data or ground-truth) is currently the primary approach used to assess soil moisture products. Several scientific teams have established dedicated soil moisture networks, including the SMAP core validation sites (Chan et al., 2016; Jackson et al., 2014) and monitoring networks in the Tibetan Plateau (Yang et al., 2013). To provide long-term ground reference data for future soil moisture missions, a new soil moisture network was constructed during the SMELR (Soil Moisture Experiment in the Luan River) (Zhao et al., 2020). Although there have been previous studies using ground data from this new network to verify soil moisture simulations (Kang et al., 2021; Wang et al., 2021), this is the first study to apply *in situ* measurements from the new soil moisture network to assess multi-source and commonly used soil moisture products. Moreover, the robustness of TCA was examined by comparing the similarity of TCA- and ground-based metrics, as well as the spread between redundant TCA estimates of multiple triples. The effects of factors such as local acquisition time, temperature input, and vegetation on soil moisture products were also investigated.

2. Datasets

2.1. Shandian River network and in situ measurements

The Shandian River flows through Hebei Province and Inner Mongolia in North China, with a total length of 877-km, and the watershed is classified as having a temperate continental climate and a seasonally frozen ground. Rainfall is mainly concentrated in summer and the annual average precipitation in most areas is 300–500 mm. The soil moisture and temperature wireless sensor network within the

Shandian River basin (referred to as the SMN-SDR hereafter) was established during the SMELR (Zhao et al., 2020), from July 18, 2018 to September 28, 2018. The topography of the SMN-SDR is relatively flat, with various land cover types including grasslands, croplands, shrublands, and forests.

The SMN-SDR currently consists of 34 stations within an area of ~10,000 km^2 (115.5–116.5°E, 41.5–42.5°N). To match the different scales of various SM products, the spatial distribution of these 34 stations (see Fig. 1) was designed as a three-sampling-scale nested structure, including 100-km (large-scale), 50-km (medium-scale), and 10-km (small-scale). At each station, five Decagon 5TM sensors were installed in a horizontal orientation at different measuring depths (3, 5, 10, 20, and 50-cm) to measure soil moisture and soil temperature. The nominal accuracy (and resolution) of the 5TM sensor are ±3% m^3/m^3 (0.0008 m^3/m^3) for soil moisture and ± 1 K (0.1 K) for soil temperature (Zhao et al., 2020). Field calibration of the 5TM sensor is also a key step to ensure the measurement accuracy of the probes. Thus undisturbed soil samples at each measuring layer were collected for laboratory analysis and calibration. The calibration function of the 5TM senor within the SMN-SDR was derived from Zhao et al. (2020). Of the 34 stations, 20 stations were equipped with the HOBO rain gauge, mainly concentrated in the small- and medium-scale. The solar panels and IoT links were used to provide remote wireless access to all ground data. The sampling interval of the data recording time is 10- (before June 2019) or 15-min (after June 2019).

Although the SMN-SDR is ideal for matching different scales (different depths) of soil moisture products, the main objective of this study was to evaluate near-surface coarse-resolution soil moisture products. Considering the shallower sampling depths of microwave remote sensing, the top-layer (3-cm) *in situ* measurements were used here for soil moisture validation. The research period of this study is from September 1, 2018 to December 31, 2020.

2.2. Soil moisture datasets

A total of 24 near-surface, coarse-resolution soil moisture datasets were used in this study, including 15 based on single-sensor satellite-based products (seven for L-band low-frequency passive microwave observations, seven for C/X-band higher-frequency passive microwave observations, and one for C-band active microwave observation), six based on multi-sensor merged products, and three based on land surface models. Each category of the soil moisture datasets was evaluated to explore the performance across and within product categories. Table 1 summarizes the main characteristics of all 24 soil moisture datasets assessed in this study. More details are provided in Appendix A.

3. Methods

3.1. Data pre-processing

3.1.1. Quality control

In practice, only the retrievals considered as “good” in soil moisture products are usually used (Gruber et al., 2020). Hence, before validation, quality control procedures are used to mask out the unreliable

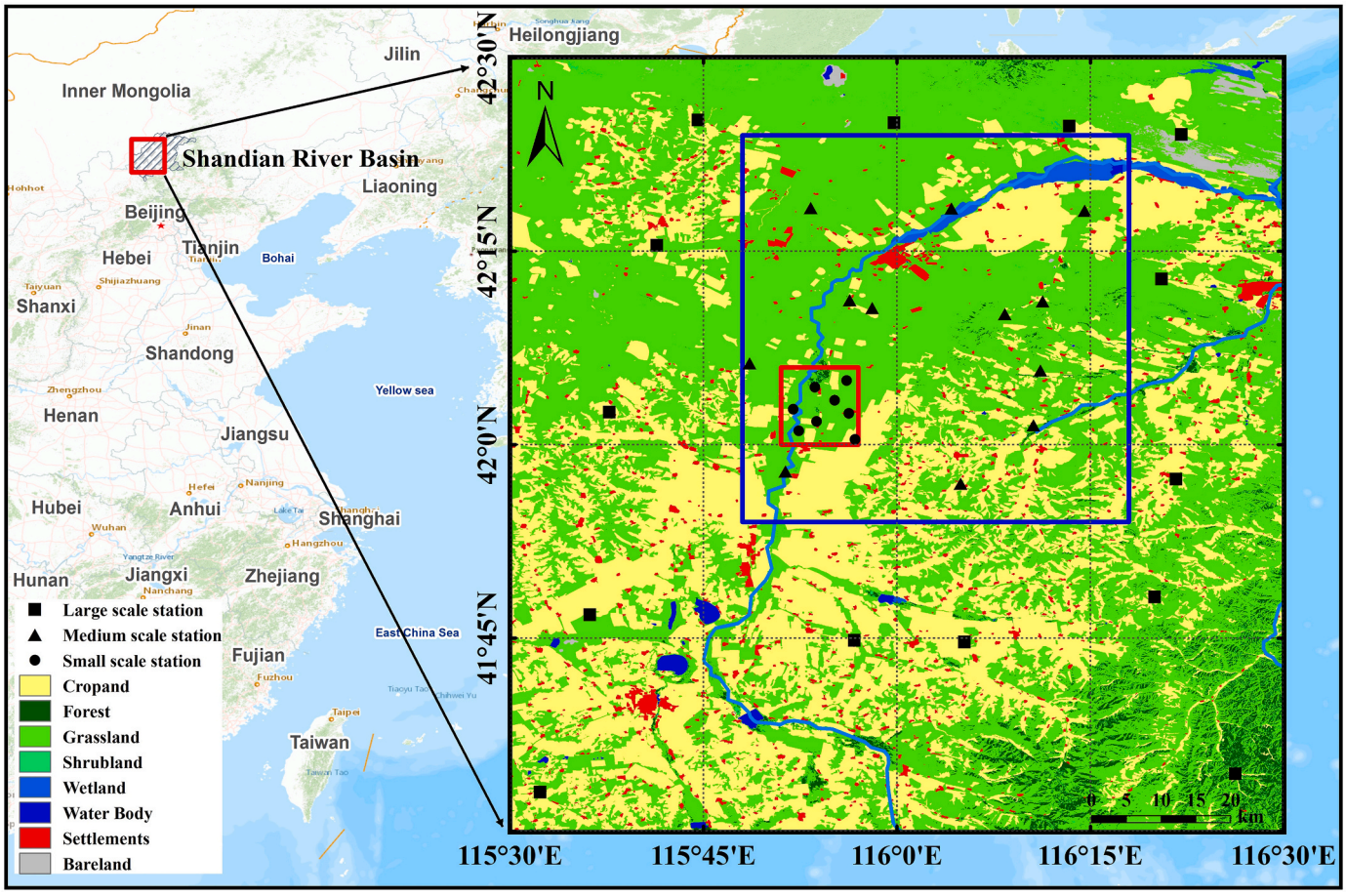

Fig. 1. Location of the SMN-SDR and spatial distribution of 34 stations in the SMN-SDR; land cover (from GlobeLand30) is also shown. The different symbols represent stations at different sampling scales including 100-km (large-scale, represented by the bold black square), 50-km (medium-scale, represented by the bold blue square), and 10-km (small-scale, represented by the bold red square). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Presentation of the soil moisture products evaluated in this study. The observation time (local time, divided into daytime and nighttime) refers to the timestamps of *in situ* data used to assess soil moisture products. For remotely sensed soil moisture data, the observation time also corresponds to the satellite overpass time. "Asc" and "Des" represent ascending and descending, respectively.

Property	Product (Version)	Grid resolution	Represent depth	Observation time		Reference(s)
				daytime	nighttime	
Low-frequency passive microwave product (L-band)	SMAP-L2 - R17000 (SCA-V\SCA-H\MDCA)	36 km	0–5 cm	(Des) 06:00	(Asc) 18:00	Entekhabi et al. (2010) O'Neill et al. (2020)
	SMAP-IB - V1	36 km	0–5 cm	(Des) 06:00	–	Li et al. (n.d.)
	SMOS-L2 - V650	15 km				Kerr et al. (2012)
	SMOS-L3 - V300	25 km	0–5 cm	(Asc) 06:00	(Des) 18:00	Al Bitar et al. (2017)
	SMOS-IC - Version 2.0	25 km				Wigneron et al. (2021)
High-frequency passive microwave product (C\X-band)	AMSR2-LPRM L3 - V001 (C1\C2\X)	0.25°	0–2 cm	(Asc) 13:30	(Des) 01:30	Owe et al. (2001)
	AMSR2-JAXA - Version 3					Fujii et al. (2009)
	FY-3B L2 - V1.0			(Des) 13:30	(Asc) 01:30	
	FY-3C L2 - V1.0	25 km	0–2 cm	(Des) 10:00	(Asc) 22:00	Shi et al. (2006)
	FY-3D L2 - V1.0			(Asc) 14:00	(Des) 02:00	Kang et al. (2021)
Active microwave product	ASCAT (H115\H116)	0.1°	0–2 cm	(Des) 09:30	(Asc) 21:30	Wagner et al. (2013) Kern (2021)
Merged product	CCI - v06.1 (active\passive\combined)	0.25°	0–3 cm	08:00	–	Gruber et al. (2019)
	C3S - v201912.0.0 (active\passive\combined)					C3S (2020)
Model-based product	GLDAS-Noah −2.1	0.25°	0–10 cm	08:00	20:00	Beaudoing and Rodell (2020)
	ERA5-Land	0.1°	0–7 cm	08:00	20:00	Muñoz Sabater (2019)
	MERRA-2 -M2T1NXLND 5.12.4	0.5° x 0.625°	0–5 cm	08:30	20:30	Gelaro et al. (2017)

estimates. However, care should be taken when selecting the quality flag thresholds, as they are a trade-off between the evaluation quality and available data quantity. The inherent quality flags of soil moisture products and corresponding thresholds used in this study are as follows: 1) SMAP-IB retrievals were masked out if the scene flag exceeded 1. 2) SMAP-L2 data were masked out if the retrieval quality flag had a value other than 0. 3) SMOS-L2 retrievals were masked out if the Radio Frequency Interferences (RFI) probability exceeded 0.2 or if the Chi-2 probability dropped below 0.05. 4) SMOS-L3 retrievals were masked out if the data quality index exceeded 0.06 or if RFI exceeded 0.3. 5) SMOS-IC retrievals were masked out if the scene flag exceeded 1 or if RMSE exceeded 8. 6) ASCAT retrievals were masked out if the soil moisture noise exceeded 10%, or if wetland fraction exceeded 15%, or if topographic complexity exceeded 20%. Note that the quality control procedures applied to different soil moisture products were not the same, and some datasets do not even have a quality flag (for instance AMSR2-JAXA and FY-3). The final validation result is thus representative of the product accuracy obtained by users that apply available quality control procedures.

The incorrect values of *in situ* measurements were also filtered by examining their time-series graphs (not shown). Additionally, all soil moisture products were assessed under unfrozen soil conditions. For this purpose, *in situ* soil temperature measurements (3-cm) were used as auxiliary information (*i.e.*, soil moisture estimates were retained only when the areal average *in situ* soil temperatures were over 273.15 K), to avoid uncertainties associated with model-based temperature data.

3.1.2. Spatial-temporal resampling

Considering the scale difference between point-based *in situ* measurements and grid-based soil moisture products, the spatial averaging method was used to obtain areal average soil moisture estimates within the SMN-SDR. Specifically, soil moisture products were resampled to a 0.25° equal-spaced lat-lon grid using nearest-neighbor search. Then all 16 resampled pixels nested within the 1° × 1° SMN-SDR (see Section 2.1) were averaged to represent the areal average retrievals of ~100 km domain. The areal mean values of soil moisture products were discarded if less than 50% of the pixels contained valid data.

The main reasons for averaging 4 × 4 pixels used for validation include: 1) Many of the satellite-based soil moisture products have a spatial resolution of around 30–50 km and can be even larger when considering the side lobe. The common solution to validating such a coarse-scale soil moisture product is to approximate the "true" values from both satellite observations and ground measurements for a particular region, which is larger than the native resolution of satellite observations and is 100 km in this study (Jackson et al., 2014). In addition, a single grid/footprint may only contain a small number of ground sites, making it difficult to eliminate scale differences between *in situ* points and satellite pixels. The single-point/footprint observations are also likely to have higher variability and noise (Jackson et al., 2012). 2) As the native/gridded spatial resolution varied for a number of the 24 datasets used in this study, it was necessary is to resample different products to a consistent grid resolution to avoid the mismatch in the targeted/study area. This solution has been often used in previous studies, especially in TCA validation (Gruber et al., 2020; Xu et al., 2021). Therefore, the resampling pre-processor used the 0.25° equal-spaced lat-lon grid as the reference grid in this study, as most products were stored at a 0.25° grid.

The *in situ* measurements from the 34 stations of the SMN-SDR were also averaged to represent the areal average "truth" of the ~100 km research domain, as there are sufficient sample sites within the SMN-SDR to approximate the full network average (Coopersmith et al., 2021). Since the SMN-SDR is an area where mixed land use is present (dominated by grassland and cropland), a weighted average was used rather than a simple arithmetic average (Yee et al., 2016), whereby the measurements from stations were first averaged across a regular 0.0625° pixel (a quarter of the 0.25° resampled grid) and the 0.0625°pixels then averaged to represent the areal average ground-truth of the SMN-SDR.

The areal average retrievals of soil moisture products were thus compared with the areal average ground "truth" within the SMN-SDR, and all evaluation (both ground-based and TCA-based) metrics calculated at the regional scale (1° × 1° in this study). Similar validation of soil moisture products at the regional scale has been applied in many previous studies (Chen et al., 2013, 2017; Jackson et al., 2010, 2012; Zeng et al., 2015).

Given the high temporal frequency of *in situ* measurements within the SMN-SDR (10 or 15 min), soil moisture products were evaluated using ground data consistent with the product's inherent timestamp (*i.e.*, satellite overpass time or model time-step). Moreover, the daytime

and nighttime performance of soil moisture products were evaluated separately (see Table 1 for the detailed division).

3.1.3. Short-term anomalies

Since the performance of soil moisture products at different time scales may vary greatly (Gruber et al., 2020), validation metrics should be separately calculated for different frequency components (in this case the raw soil moisture estimates and short-term anomalies). The raw soil moisture estimates denote the non-decomposed data, whereby the short-term anomalies refer to the residuals from a seasonality, which is not averaged across years into a climatology as the study period is short (2018–2020). The assessment of short-term anomalies aims to obtain the capability of soil moisture products to capture the day-to-day fluctuation (or individual drying or wetting events), which can remove seasonal effects on skill metrics and is more useful than raw soil moisture estimates for many applications (such as data assimilation) (Al-Yaari et al., 2019). In this study, short-term anomalies were computed as the deviation from the mean for a 35-day moving window, which has been widely used in previous studies (Albergel et al., 2012; Gruber et al., 2020). The central moving mean was estimated only when there are at least 25% valid data within the 35-day window (Gruber et al., 2020). The function of short-term anomalies is given by

$$SM_{anom}(t) = SM(t) - \overline{SM(t-17:t+17)} \tag{1}$$

where $SM(t)$ and $SM_{\mathrm{anom}}(t)$ represent the raw soil moisture estimates and short-term anomalies at time t in days, respectively. The over-bar is the temporal mean operator.

3.2. Validation metrics

The error characteristics of soil moisture products are unable to be fully described by a single validation metric. Accordingly, four widely used statistical metrics for soil moisture validation were computed to obtain a more comprehensive skills description of soil moisture products.

First, the Pearson correlation coefficient (R) was used to evaluate the ability of soil moisture products to capture temporal variations of *in situ* measurements. Then the dryness or wetness of soil moisture products was measured relative to ground data by the temporal mean bias (Bias). Moreover, the root mean squared error (RMSE) and unbiased RMSE (ubRMSE) were interpreted as the standard deviation of random error (Montzka et al., 2020). These skill metrics were calculated as

$$\mathrm{R} = \sum_{i=1}^{n}\left(SM_{estimated(i)} - \overline{SM_{estimated}}\right)\left(SM_{in\ situ(i)} - \overline{SM_{in\ situ}}\right) \Big/ \sqrt{\sum_{i=1}^{n}\left(SM_{estimated(i)} - \overline{SM_{estimated}}\right)^2 \sum_{i=1}^{n}\left(SM_{in\ situ(i)} - \overline{SM_{in\ situ}}\right)^2} \tag{2}$$

$$\mathrm{Bias} = \overline{\left(SM_{estimated} - SM_{in\ situ}\right)} \tag{3}$$

$$\mathrm{RMSE} = \sqrt{\overline{\left(SM_{estimated} - SM_{in\ situ}\right)^2}} \tag{4}$$

$$\mathrm{ubRMSE} = \sqrt{\overline{\left[\left(SM_{estimated} - \overline{SM_{estimated}}\right) - \left(SM_{in\ situ} - \overline{SM_{in\ situ}}\right)\right]^2}} \tag{5}$$

where the $SM_{in\ situ}$ and $SM_{estimated}$ denote the *in situ* measurements and soil moisture products, respectively. The over-bar represents the temporal mean operator, and n is the number of comparison soil moisture data pairs.

3.3. Triple collocation metrics

In this study, TCA was used to verify soil moisture products, and TCA robustness was tested by comparing the consistency between the TCA- and ground-based metrics. There are some basic assumptions in TCA (Gruber et al., 2016): 1) The three collocated datasets required by TCA should describe the same geophysical variable (in this case the areal-average near-surface soil moisture of the SMN-SDR). 2) All three datasets should be linearly related to the true state. 3) The errors of triplets are independent of each other. 4) The errors of triplets are independent of the true state.

In practice, different triplets of TCA may violate the above basic assumptions for various reasons (Kim et al., 2020). Consequently, the commonly used triplet combination of an active-based product, a passive-based product, and a model-based product is applied here to calculate TCA-based metrics from the range of alternatives. The focus of this study was on the performance of the TCA-based R (the linear correlation against the unknown truth) (McColl et al., 2014) and ubRMSE (the temporal standard deviation of errors in datasets) (Montzka et al., 2020). Their calculation formulas are given by

$$R_x = \sqrt{\frac{\sigma_{xy}\sigma_{xz}}{\sigma_{xx}\sigma_{yz}}} \tag{6}$$

$$ubRMSE_x = \sqrt{\sigma_{xx} - \frac{\sigma_{xy}\sigma_{xz}}{\sigma_{yz}}} \tag{7}$$

where the subscripts x, y, and z refer to the triplets in TCA (in this case an active product, a passive product, and a model-based product). The covariance between datasets is defined as σ. In the present study, the TCA-based metrics are considered to be robust only when: 1) the triplets had more than 100 comparison samples, and 2) the correlation of any two datasets was greater than 0.3 (Kim et al., 2020).

3.4. Confidence intervals from moving block bootstrap

Calculation of the above metrics is based on finite-sized samples, with the sample size directly affecting the statistical performance. To quantify the statistical uncertainties caused by sampling errors, this study relied on the moving block bootstrap (MBB) re-sampling technique (Ólafsdóttir and Mudelsee, 2014) to construct confidence intervals (CI) for validation metrics. The relative differences between different cases (in this study performance metrics of different soil moisture products) were considered to be greater than sampling uncertainty (*i.e.*, they are statistically significant) only when the CI of various cases do not overlap.

Compared with resampling single data points, the advantage of MBB (resampling data blocks composed of a certain amount of consecutive observations) is that it can prevent the autocorrelation in the soil moisture time-series from generating an enlarged CI (Gruber et al., 2020). The specific steps to apply MBB to estimate CI of validation metrics are summarized as follows: 1) Resample blocks of collocation input datasets 1000 times, with replacement and preserve the original sample size. For detailed information about the calculation of the optimal block length, readers are kindly referred to Gruber et al. (2020). 2) Repeatedly calculate validation metrics in each resampling procedure. 3) Construct the empirical probability distribution of these iteration

metrics and obtain the corresponding 90% CI to visually express sampling errors (*i.e.*, the range between the 5th and 95th percentile of the block-bootstrapped sampling distribution).

4. Results

4.1. Soil moisture products vs. in situ measurements

In this study, all available soil moisture estimates for each product (after quality control) during the study period were evaluated to obtain the accuracy when the end-user uses these products separately (Al-Yaari et al., 2019). Note that all evaluation metrics were calculated during the non-frozen period (see Section 3.1.1). In addition to being presented directly in the scatterplots, the ground-based metrics of raw soil moisture estimates are also listed in Table 2.

4.1.1. Evaluation of satellite-based soil moisture products

Fig. 2 presents the time series and scatterplots of low-frequency passive microwave products (SMOS and SMAP) and areal *in situ* measurements within the SMN-SDR. According to the time series, areal *in situ* measurements showed a moderate variation from 0.1 to 0.3 m^3/m^3. The soil surface in SMN-SDR was frozen approximately in November and gradually thawed in April with little variations of soil liquid water content during this period. The time series also showed that SMAP and SMOS tended to underestimate soil moisture from the SMN-SDR. This underestimation has also been reinforced by their negative bias statistics. The ubRMSE values of SMAP (0.036–0.044 m^3/m^3) were generally lower than those of SMOS (0.041–0.064 m^3/m^3), which basically meets its expected accuracy goal (*i.e.*, ubRMSE <0.04 m^3/m^3). The higher R values (0.590–0.630) in SMAP also reflected that it can better capture the temporal variation of *in situ* measurements in the SMN-SDR.

During the comparison of different SMAP retrieval algorithms, time series (a1 and c1 in Fig. 2) indicated that SCA-H produced lower retrievals, while MDCA and SCA-V retrievals were very similar. Though SMAP-IB underestimated ground measurements, it is encouraging that SMAP-IB was wetter than SMAP-SCA and -MDCA. The bias statistics confirmed this phenomenon, showing an increasing trend of negative bias from SMAP-IB (−0.057 m^3/m^3) to SCA-V (−0.064 and − 0.070 m^3/m^3) to MDCA (−0.064 and − 0.071 m^3/m^3) and to SCA-H (−0.092 and − 0.097 m^3/m^3). The R values of the four SMAP algorithms were very similar, being close to 0.6. SCA-H had the largest error in the SMN-SDR, with the biggest RMSE, ubRMSE, and bias among the different SMAP algorithms. The ubRMSE (0.044 m^3/m^3) of SMAP-IB was similar to SMAP-SCA. The performance of MDCA was slightly better than SCA-V because of its smaller ubRMSE (0.036 and 0.041 m^3/m^3).

The soil moisture retrievals in the three SMOS products also showed an underestimated change from SMOS-L2 (−0.063 and − 0.067 m^3/m^3) to SMOS-L3 (−0.084 and − 0.090 m^3/m^3) and to SMOS-IC (−0.103 and − 0.110 m^3/m^3). Although SMOS-IC was the driest estimate overall for the SMN-SDR among the three SMOS products, it had the lowest ubRMSE (0.041 and 0.043 m^3/m^3) and acceptable R values (0.451 and 0.561). Using a multi-orbit approach, SMOS-L3 produced more available data and had a smaller ubRMSE (0.047 and 0.050 m^3/m^3) than SMOS-L2 (0.057 and 0.064 m^3/m^3).

The comparison between higher-frequency passive microwave products (AMSR2 and FY-3) and ground measurements is displayed in Fig. 3. From time series (a1 and c1 in Fig. 3), it can be seen that JAXA retrievals underestimated ground measurements and showed very low variations, maintaining an almost constant value of around 0.1 m^3/m^3 except for soil moisture fluctuations in summer (June to September). This phenomenon of JAXA was manifested in the quantitative metrics as the largest dry bias (−0.115 and − 0.117 m^3/m^3) and RMSE (0.121 and 0.123 m^3/m^3) among all evaluated products, leading to the unexpectedly lowest ubRMSE (0.036 and 0.037 m^3/m^3) among all satellite-based products. The time series plots also suggested that the soil moisture variation range of AMSR2-LPRM was much larger than JAXA. Except for the slight underestimation of the LPRM-C2 nighttime product, other LPRM products overestimated ground measurements. Mutual comparison of different datasets in the LPRM products generally showed that the lower the frequency, the better the metrics (*i.e.*, lower ubRMSE and higher R). For example, among the daytime products, LPRM-C1 (6.9-GHz) had the lowest ubRMSE of 0.051 m^3/m^3, LPRM-C2 (7.3-GHz) had an ubRMSE of 0.059 m^3/m^3, and LPRM-X (10.7-GHz) had the highest ubRMSE of 0.077 m^3/m^3. The low R values (< 0.5, especially less than 0.2 during daytime) of AMSR2-LPRM present that LPRM products did not reflect the temporal variation of ground measurements.

Table 2
Summary statistics (R, bias, RMSE, and ubRMSE) of comparison between raw soil moisture estimates of the 24 soil moisture datasets and *in situ* measurements (3-cm) within the SMN-SDR for the period 01/09/2018 to 31/12/2020. The daytime and nighttime performance were evaluated separately. "Num" is the number of comparison data pairs. The best metrics in each category of soil moisture product are indicated in bold.

Property	Product	daytime					nighttime				
		R	bias	RMSE	ubRMSE	Num	R	bias	RMSE	ubRMSE	Num
Low-frequency passive microwave product (L-band)	SMAP-IB	0.614	**−0.057**	**0.072**	0.044	221					
	SMAP-L2 (SCA-H)	**0.630**	−0.092	0.102	0.044	215	0.590	−0.097	0.105	0.039	224
	SMAP-L2 (SCA-V)	0.625	−0.064	0.077	0.043	219	**0.606**	−0.070	**0.079**	0.037	232
	SMAP-L2 (MDCA)	**0.630**	−0.064	0.076	**0.041**	200	0.591	−0.071	0.080	**0.036**	205
	SMOS-L2	0.577	−0.063	0.090	0.064	210	0.362	**−0.067**	0.088	0.057	165
	SMOS-L3	0.570	−0.084	0.097	0.050	**272**	0.440	−0.090	0.101	0.047	**273**
	SMOS-IC	0.561	−0.110	0.118	0.043	123	0.451	−0.103	0.111	0.041	116
High-frequency passive microwave product (C\X-band)	AMSR2-LPRM (C1)	0.182	0.074	0.090	0.051	**447**	**0.490**	0.066	0.088	0.059	418
	AMSR2-LPRM (C2)	0.155	0.021	**0.063**	0.059	**447**	0.416	−0.002	0.070	0.070	418
	AMSR2-LPRM (X)	0.130	0.027	0.082	0.077	**447**	0.445	0.080	0.103	0.065	418
	AMSR2-JAXA	**0.449**	−0.115	0.121	**0.037**	**447**	0.430	−0.117	0.123	**0.036**	**425**
	FY-3B	0.055	**0.007**	0.077	0.077	77	0.187	−0.031	0.076	0.069	126
	FY-3C	0.338	−0.013	0.065	0.064	170	0.257	−0.011	0.070	0.069	135
	FY-3D	0.361	−0.441	0.068	0.054	363	0.420	**−0.001**	**0.065**	0.065	235
Active microwave product	ASCAT	0.391	0.093	0.105	0.049	377	0.382	0.086	0.100	0.050	371
Merged product	CCI-active	0.460	0.033	0.073	0.065	385					
	CCI-passive	**0.616**	0.047	0.063	0.042	395					
	CCI-combined	0.609	0.006	**0.028**	**0.027**	416					
	C3S-active	0.477	0.071	0.094	0.063	456					
	C3S-passive	0.543	0.033	0.063	0.054	388					
	C3S-combined	0.570	**−0.001**	0.031	0.030	**494**					
Model-based product	ERA5-Land	**0.666**	0.072	0.080	0.037	507	**0.605**	0.066	0.076	0.037	**550**
	GLDAS-Noah	0.356	0.007	0.038	0.038	507	0.322	**−0.003**	0.038	0.038	**550**
	MERRA-2	0.464	**−0.003**	**0.028**	**0.027**	**509**	0.356	−0.009	**0.030**	**0.029**	547

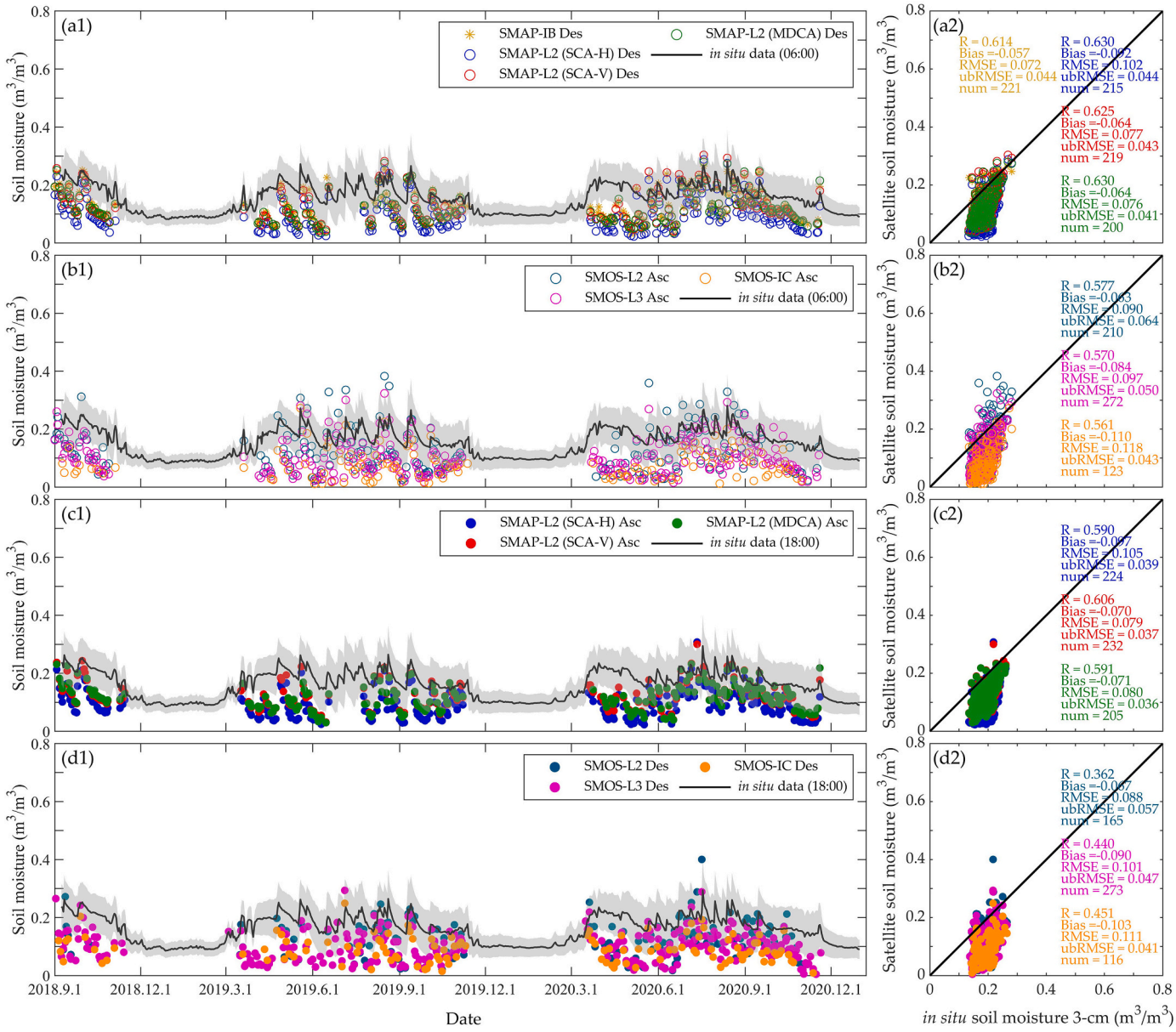

Fig. 2. The comparison between raw soil moisture estimates of L-band low-frequency passive microwave products (SMAP and SMOS) and *in situ* measurements (3-cm) within the SMN-SDR for the period 01/09/2018 to 31/12/2020, including time series (left, with the suffix −1) and scatterplots (right, with the suffix −2). The daytime (a and b, shown as hollow) and nighttime (c and d, shown as solid) performance were evaluated separately. "Asc" and "Des" represent ascending and descending orbits, respectively. The gray-shaded areas in the time-series refer to the standard deviation of stations measurements within the SMN-SDR.

FY-3 products presented low accuracy in the SMN-SDR, with large ubRMSE (higher than 0.06 m^3/m^3 in most cases) and small R values (lower than 0.4 in most cases). The performance of FY-3D was better than FY-3B and -3C. Due to the different overall lengths of FY datasets (mentioned in Section A.1.4 of the appendices), FY-3B and -3C are only available in small quantities (see Table 2). The ubRMSE values of ASCAT (see Fig. 4) were similar to those of SMOS L3 (close to 0.05 cm^3/cm^3), being lower than those of all high-frequency passive products (except for the underestimated JAXA), and higher than those of low-frequency passive SMAP and SMOS-IC. However, the performance of ASCAT in terms of correlation was low ($R < 0.4$), with ASCAT being greatly wetter than ground measurements (positive bias of 0.086 and 0.093 m^3/m^3).

In summary, SMAP-MDCA outperformed other single-sensor satellite-based soil moisture products within the SMN-SDR, indicated by the highest R of 0.630 and the lowest ubRMSE of 0.036 m^3/m^3. It's worth noting that among SMAP products, the latest SMAP-IB performed similarly to SMAP-MDCA as their R and ubRMSE metrics were close, but the dry bias of SMAP-IB was improved. Despite the consistent underestimation bias of SMOS and SMAP retrievals, the overall behavior of those L-band passive products was better than that of higher-frequency passive (AMSR2 and FY-3) and active microwave (ASCAT) products. For SMOS products, the better performance, particularly in terms of ubRMSE estimates, of SMOS-IC over SMOS-L2 and -L3 products. Additionally, the active-sensor ASCAT performed similar to AMSR2 C-band products and was superior to higher-frequency X-band products.

4.1.2. *Evaluation of merged soil moisture products*

The multi-sensor merged soil moisture products chosen in this study are the CCI and C3S datasets. Their broader bandwidth sensitivity and combination of microwave principles are expected to improve the performance statistics (González-Zamora et al., 2019). The analysis of CCI/C3S results has two main parts. In the first part, the impact of merging

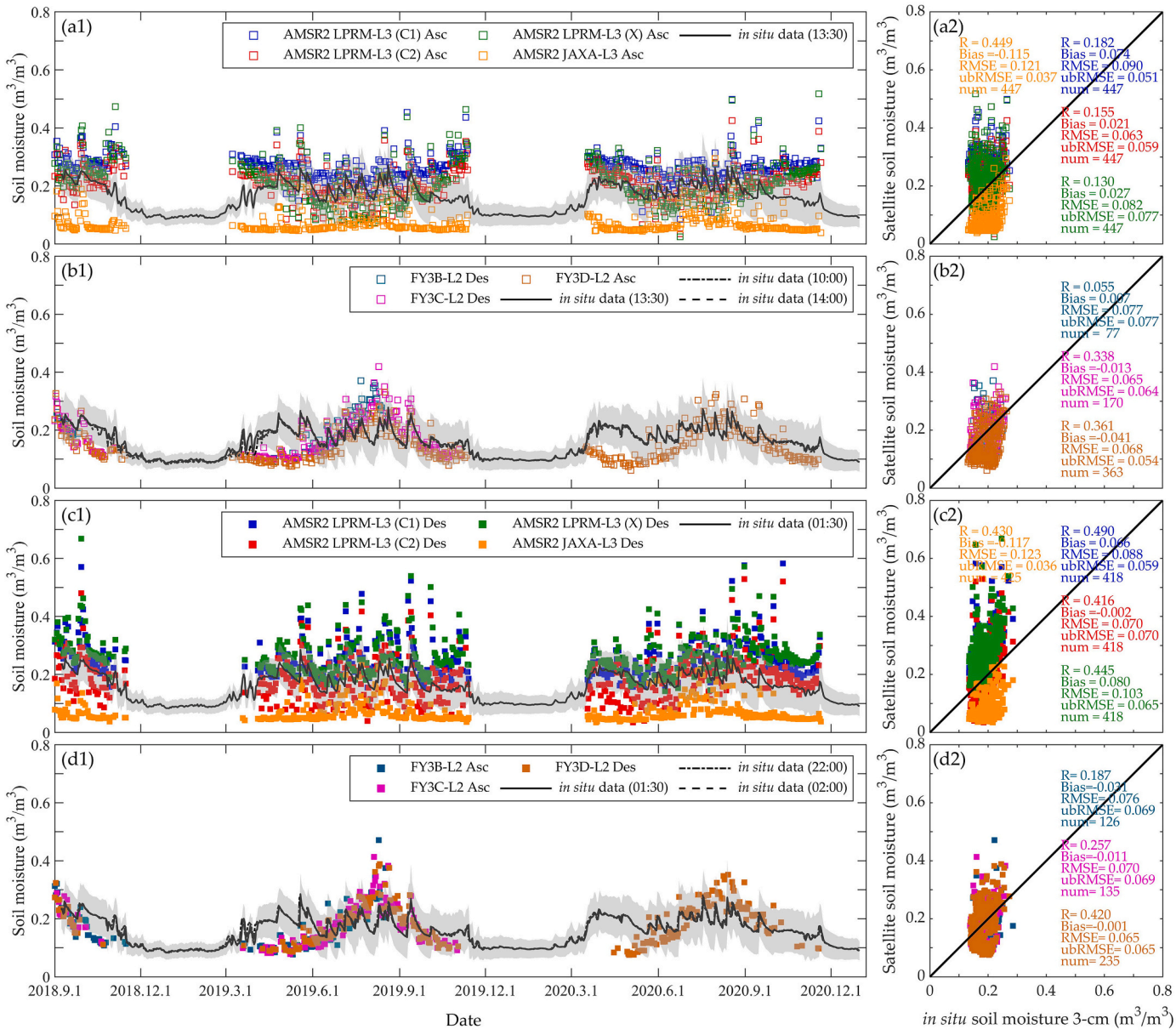

Fig. 3. Same as Fig. 2, but for the C/X-band higher-frequency passive microwave products (AMSR2 and FY-3).

SMAP retrievals on the performance of the latest version CCI was examined. Fig. 5 clearly shows that CCI-passive and CCI-combined obtained by merging extra SMAP data had higher R (0.616 and 0.609) and lower ubRMSE (0.042 and 0.027 m^3/m^3) values than C3S (without fusion of SMAP retrievals). The performance of CCI-active and C3S-active was similar since the merged active sensors are the same (ASCAT-A/B). Compared with ASCAT time-series products (5-day composite, Fig. 4), CCI/C3S-active products had a comparable R (0.460 and 0.477), but worse ubRMSE (> 0.05 m^3/m^3). In the second part, the performance differences of three merged data within CCI and C3S were compared. The ranking of the ubRMSE values (from best to worst) of the six merged CCI/C3S soil moisture data was (1) CCI-combined, (2) C3S-combined, (3) CCI-passive, (4) C3S-passive, (5) C3S-active, and (6) CCI-active.

Overall, CCI-combined performed better than the best-performing single-sensor satellite-based product (*i.e.*, SMAP-MDCA) in terms of ubRMSE values. It was noticeable that CCI-combined retrievals were equally distributed around the 1:1 line, with the comparable R (> 0.6) and the better bias (0.027 m^3/m^3). Compared to the single-senor products with dry or wet biases, a better bias value (close to zero) and the greater number of observations (see Table 2) for CCI-combined highlight the advantages of multi-sensor merging tools.

4.1.3. *Evaluation of model-based soil moisture products*

Fig. 6 shows the assessment results of three model-based products in the SMN-SDR. As mentioned in Table 1, *in situ* measurements at 8:30 and 20:30 were used to evaluate MERRA-2 daytime and nighttime performance, respectively. The *in situ* measurements at 8:00 and 20:00 corresponds to the other two soil moisture models (GLDAS-Noah and ERA-5 Land). Compared to GLDAS-Noah and MERRA-2 with close-to-zero bias, ERA5-Land (blue triangle) overestimated ground measurements, with the larger positive bias (0.066 and 0.072 m^3/m^3). However, ERA5-land could reflect well the dynamic change of ground measurements and display the highest R values (0.605 and 0.666). It is worth noting that among all the products evaluated, those with R values greater than 0.6 were only SMAP, CCI-passive, and CCI-combined products (including merging with SMAP retrievals), and ERA5-Land. The ubRMSE values of all three model products satisfied the SMAP scientific goal of ubRMSE

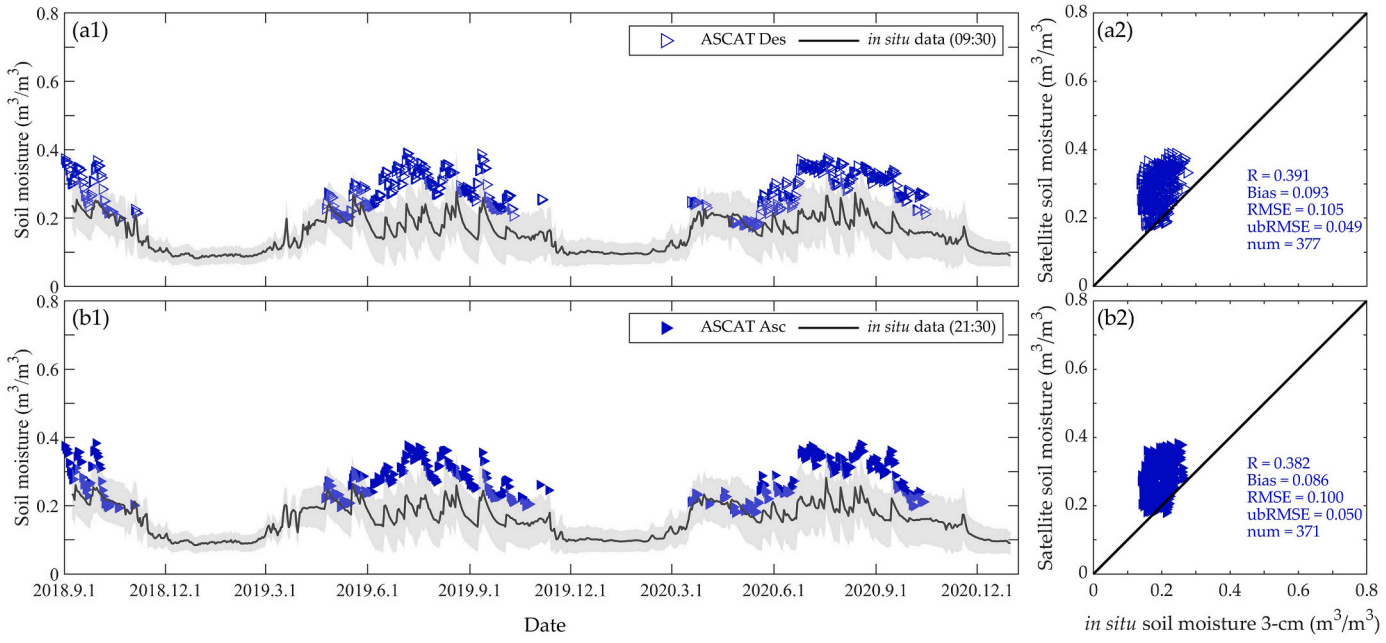

Fig. 4. Same as Fig. 2, but for the active microwave products (ASCAT).

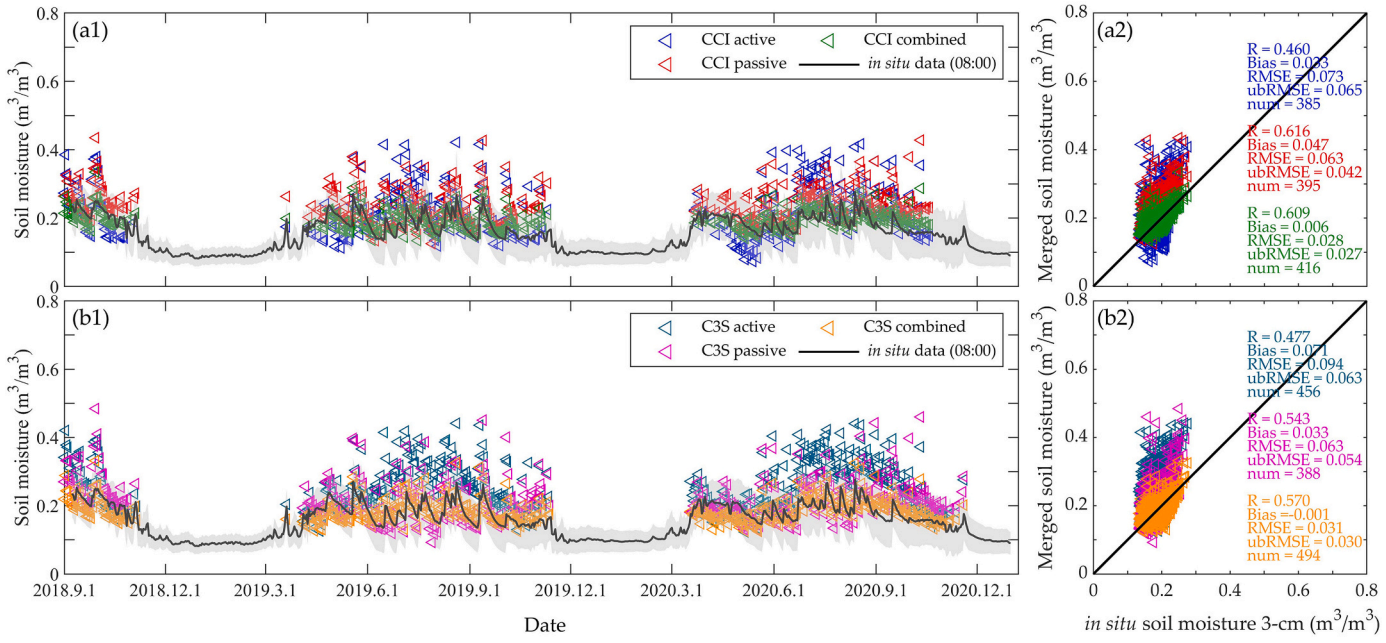

Fig. 5. Same as Fig. 2, but for the multi-sensor merged products (CCI and C3S).

<0.04 m^3/m^3. Although MERRA-2 had the lowest ubRMSE (0.027 and 0.029 m^3/m^3) among all the evaluated products, it was mainly linked to the relatively low temporal variability in MERRA-2 rather than its high accuracy. The relatively lower R (< 0.5) and narrower distribution (green triangle in Fig. 6) also confirm the interpretation for MERRA-2, that the 'flat' soil moisture estimates may result in underestimating the ubRMSE statistics.

To sum up, only ERA5-Land reflected similar performance to SMAP in terms of R (> 0.6) and ubRMSE (< 0.04 m^3/m^3), but ERA5-Land had a large wet bias. The GLDAS-Noah and MERRRA2 products should be improved in capturing time dynamics of *in situ* measurements within the SMN-SDR.

4.2. Validation of triple collocation analysis

Triple collocation analysis (TCA) was employed to evaluate soil moisture products by constructing the data collection as triplets: a passive microwave product, an active microwave product, and a model-based product. More than one independent triplet could be built to estimate TCA-based metrics of a particular soil moisture product since numerous soil moisture datasets were evaluated. It is worth noting that if a product is blended or assimilated into another product, the two datasets should not be considered in the same triplet to avoid cross-correlation errors (Kim et al., 2020). Hence, CCI/C3S-combined products were not used for TCA in this study. CCI/C3S-passive and -active products which use GLDAS-Noah data were also not used for TCA with

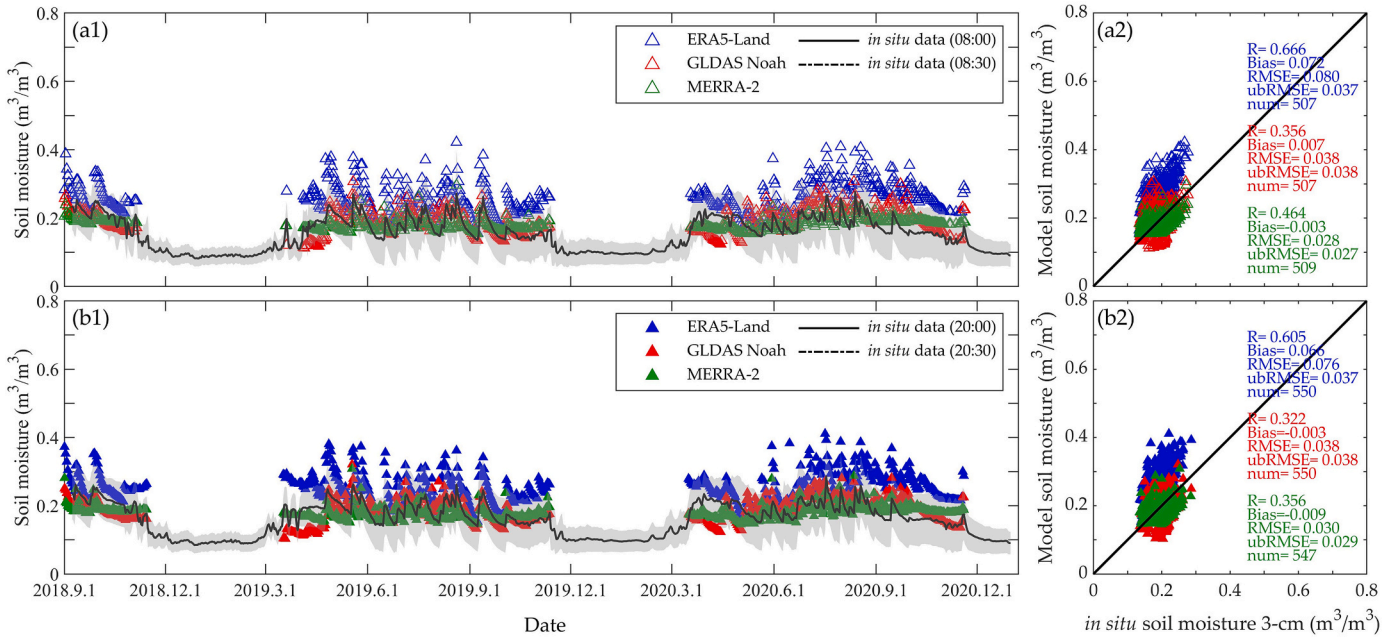

Fig. 6. Same as Fig. 2, but for the model-based products (ERA5-Land, GLDAS-Noah, and MERRA-2).

Table 3
Possible triplets (a passive microwave product, an active microwave product, and a model-based product) for daytime products.

Property	Passive product	Different combinations of possible triplets	
		Active product	Model product
Low-frequency passive microwave product (L-band)	SMAP-IB SMAP-L2 (SCA-V\SCA-H \MDCA) SMOS-IC	ASCAT	GLDAS-Noah ERA5-Land MERRA-2
		CCI-active C3S-active	ERA5-Land MERRA-2
	SMOS-L2 SMOS-L3	ASCAT	GLDAS-Noah MERRA-2
		CCI-active C3S-active	MERRA-2
High-frequency passive microwave product (C\X-band)	AMSR2-LPRM L3 (C1 \C2\X) AMSR2-JAXA FY-3B L2 FY-3C L2 FY-3D L2	ASCAT	GLDAS-Noah ERA5-Land MERRA-2
		CCI-active C3S-active	ERA5-Land MERRA-2
Merged passive product	CCI-passive C3S-passive	ASCAT CCI-active C3S-active	ERA5-Land MERRA-2

GLDAS-Noah. As, SMOS-L2 and -L3 products used ECMWF soil moisture data in the algorithm they were not considered as independent for use in TCA with ERA5-Land. The possible triplets of each daytime product are shown in Table 3. Note that only the active- or model-product TCA metrics from the triplets fusing SMOS or SMAP were adopted because the frequencies of ASCAT and low-frequency passive products are more dissimilar than high-frequency passive products (Kim et al., 2018).

To avoid violating the implicit assumption that triplets should describe the same geophysical variable, active and passive microwave products were collocated by taking the closest temporal observations (see Table 1). Moreover, the model-based products were taken from that at the central time between the two overpass times (satellite-active and -passive products) to construct the TCA triplet (Lei et al., 2015). Fig. 7 displays the comparison between the TCA- and ground-based metrics for the daytime products in 21 soil moisture datasets, including the raw soil moisture estimates and short-term anomalies. Since the CCI/C3S products were classified into daytime products (see Table 1), the number of triplets in the nighttime products was different from the daytime products (see Table A1).

The feasibility and robustness of TCA can be investigated not only by comparing the consistency of results with the ground-based metrics (blue circle in Fig. 7 and A1) but also by checking the differences between the redundant TCA estimates corresponding to multiple triplets of the same product (the height of the green box or error-bar in Fig. 7 and A1). It can be seen from Fig. 7 and A1 that the TCA-based R values were systematically higher than ground-based R, and the TCA-based ubRMSE values were consistently lower than ground-based ubRMSE, for both raw soil moisture estimates and short-term anomalies. TCA allows the calculation of unbiased error metrics of soil moisture products relative to the unknown truth, while the ground-based metrics inevitably contain errors of *in situ* measurements. Therefore, the above-mentioned difference between TCA- and ground-based metrics may be affected by the representativeness of the *in situ* network at the satellite footprint, such as the accuracy of ground measurements and scale differences (Dorigo et al., 2015). Besides, the cross-correlation errors between soil moisture products in TCA triplets may also lead to systematic deviations in TCA estimates (Chen et al., 2018).

It is also clear that TCA-based metrics were generally very similar to ground-based estimates by comparing the relative trends of product performance in Fig. 7. However, in a few cases, the TCA-based metrics showed an opposite trend to the ground-based calculation. That is, soil moisture products with lower (or higher) ground-based metrics had higher (or lower) TCA-based values. This opposite trend can cause a partial closed loop on the curve, such as the CCI/C3S raw soil moisture estimates in the upper two rows of Fig. 7. It is interesting to note that the above-closed-loop basically disappeared in the CCI/C3S short-term anomalies (see the bottom two rows of Fig. 7), indicating that the cross-correlation errors caused by the climatology can be eliminated to a certain extent by calculating the anomaly time-series (Chen et al., 2018). Furthermore, the TCA-based metrics from different triplets were quite different in most cases (especially raw soil moisture estimates, the higher box height in the upper two rows of Fig. 7 and A1). While the above difference may be caused by the sampling errors or the

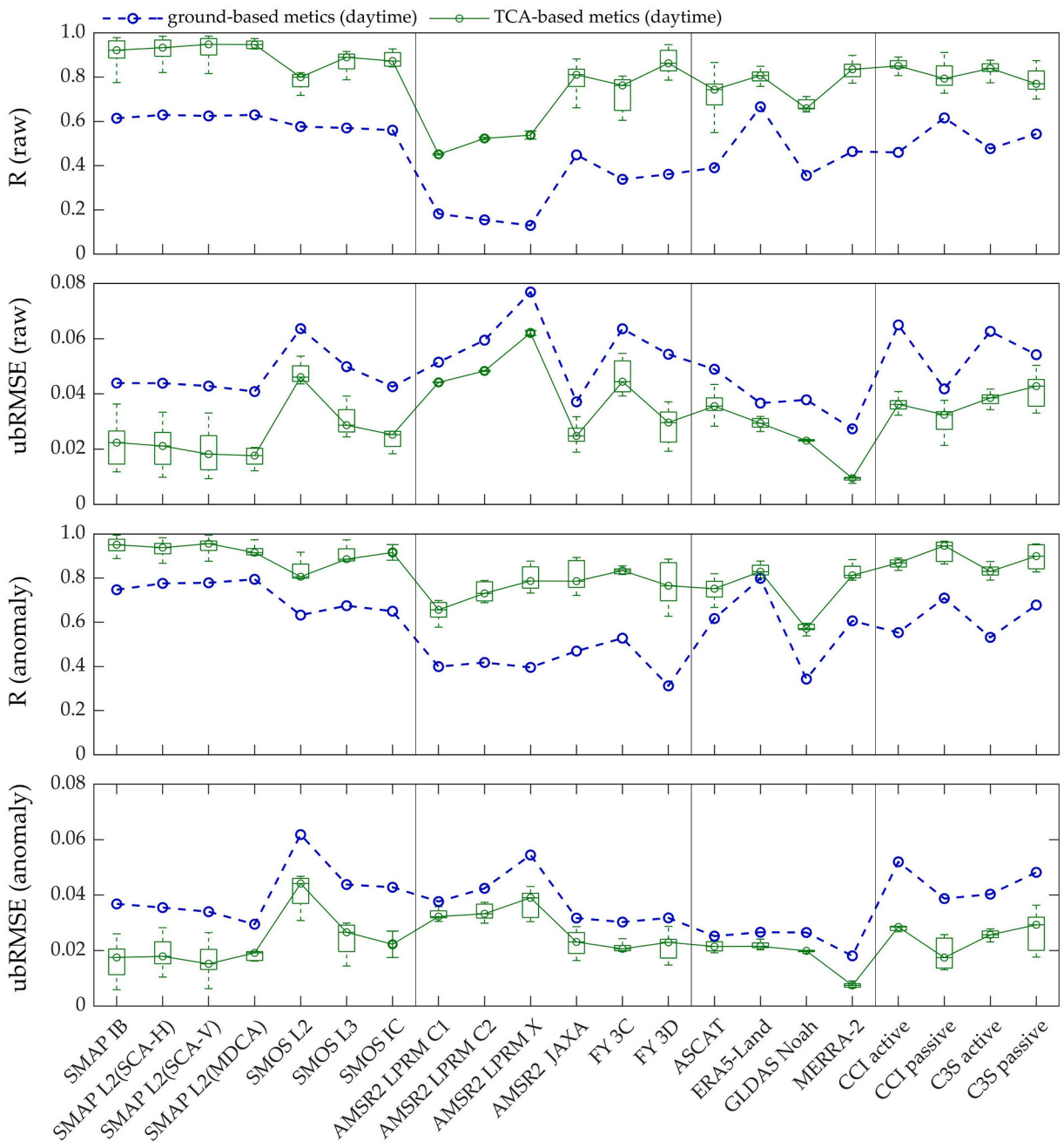

Fig. 7. Summary statistics of TCA-based (green boxplots for more than two robust TCA estimates, green error-bars for only two robust TCA estimates) and ground-based (blue circles) R and ubRMSE values, for raw soil moisture estimates (the first and second rows) and short-term anomalies (the third and fourth rows) of daytime products. The green line in the box (the green point in the error bar) is the TCA-based sample median, the bottom and top of each box represent the inter-quartile range, and whiskers represent the minimum and maximum values. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

destruction of the TCA assumption (*e.g.*, the sampling depth mismatch), recent studies have shown that the error correlation between active and passive soil moisture products is non-negligible (Chen et al., 2018; Gruber et al., 2016). However, the use of short-term anomalies seemed to reduce this difference, especially the ubRMSE values during nighttime (see the bottom row of Fig. A1).

The above difference in the redundant TCA-based estimates can also be shown in the error bars in Figs. 8 and A2 (the first row). The top and bottom of the error bar represent the maximum and minimum TCA values, and the midpoint represents the average of multiple TCA estimates. Compared to the raw soil moisture estimates (a and c in Fig. 8 and A2), the range of error bars was reduced to a certain extent in the short-term anomalies (b and d in Fig. 8 and A2). The skill estimate closest to the ground-based metric among multiple TCA triplets, calculated based on the condition that the quadratic sum of ΔR and ΔubRMSE (difference between the normalized TCA- and ground-based metrics) is the minimum, is also displayed in Fig. 8 and A2 (see e, f, g, and h). It is encouraging that the average of the multiple TCA estimates (midpoint of the error bar in the first row) shows similar performance to the scatters in the second row, or in other words, shows similar performance to ground-based metrics overall. Recent studies used four or more datasets to calculate the cross-correlation errors of soil moisture products within the TCA triplet, but this TCA extension also contains some basic assumptions (Chen et al., 2018). To sum up, under the premise of not

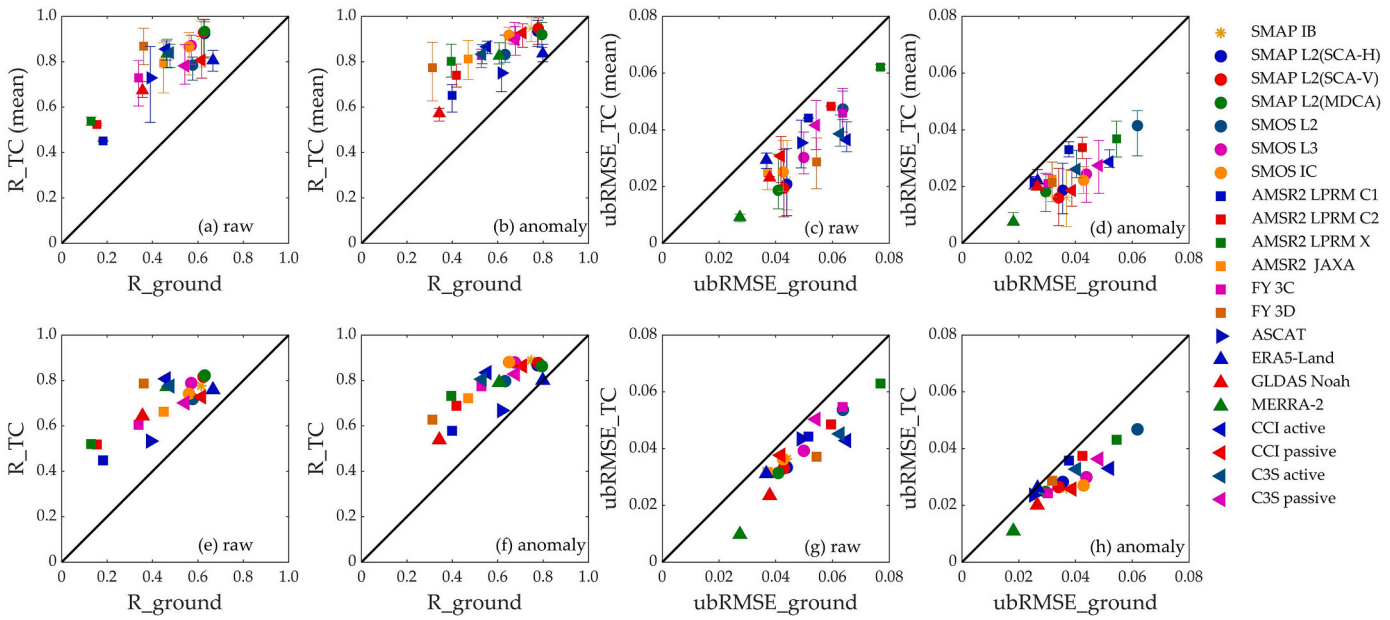

Fig. 8. TCA-based metrics *versus* ground-based statistics (R and ubRMSE), for raw soil moisture estimates and short-term anomalies of daytime products. The midpoint of the error bar (a, b, c, d) is the average of multiple TCA metrics. The bottom and top of the error bar capture the minimum and maximum TCA estimates. The points of scatterplots (e, f, g, h) represent the skill estimate closest to the ground-based metric among redundant TCA metrics, which was calculated based on the condition that the quadratic sum of the ΔR and ΔubRMSE (difference between the normalized TCA- and ground-based metrics) is the minimum.

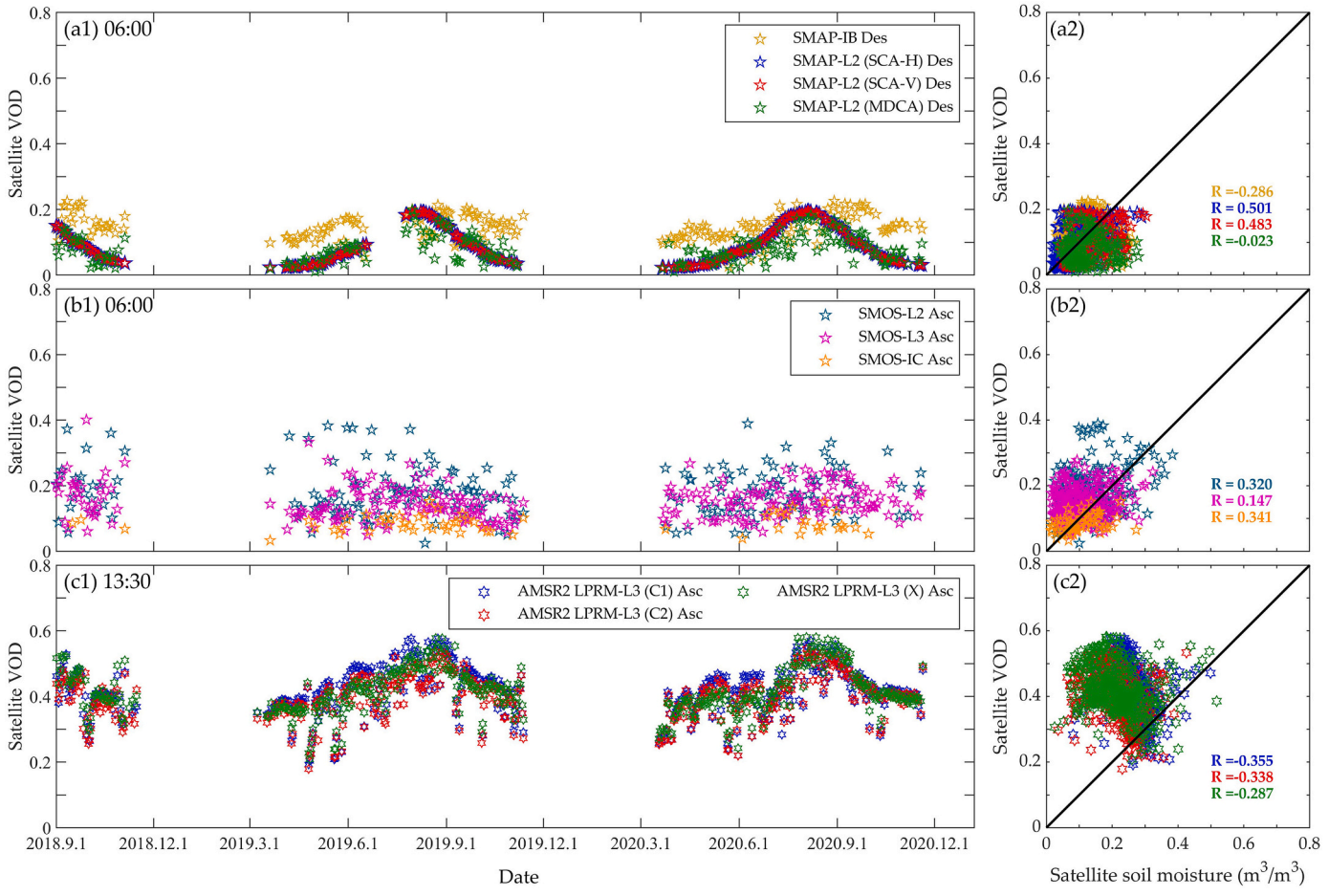

Fig. 9. Time-series of satellite VOD daytime products, and scatterplots of passive microwave soil moisture products against corresponding VOD daytime products within the SMN-SDR for the period 01/09/2018 to 31/12/2020. "Asc" and "Des" represent ascending and descending orbits, respectively.

obviously destroying the TCA assumption, averaging multiple valid TCA estimates is a simpler way to reduce the uncertainty of a single TCA value and increase the precision of the estimates.

5. Discussion

5.1. Impacts of vegetation optical depth

The uncertainty in vegetation parameters is considered to be the main cause of soil moisture retrieval error, due to the coupling effect between soil and vegetation contributions (Zhao et al., 2021). The time series of ten different satellite VOD datasets were presented to analyze the effect of vegetation parameters on soil moisture products. Both daytime (Fig. 9) and nighttime (Fig. A3) performance were evaluated in this study. The SMAP-SCA algorithm uses NDVI (Normalized Difference Water Index) to estimate VOD, while SMOS, SMAP-IB, SMAP-MDCA, and AMSR2-LPRM use the observed TB information to retrieve the VOD simultaneously with the soil moisture calculation.

Through comparisons in Fig. 9, LPRM achieved the highest VOD values (0.2 to 0.7) in the SMN-SDR, followed by SMOS-L2 and -L3 products (0 to 0.5), while SMAP and SMOS-IC had a very narrow range of VOD values (less than 0.2). The higher range of VOD retrievals for LPRM has also been observed on another network in China (Cui et al., 2017). Higher VOD estimates mean lower transmittance of the vegetation layer and thus a higher emissivity of the vegetation, resulting in a decrease in the soil emissivity as well as an increase in the soil dielectric constant, and finally an increase in the soil moisture values. The inter-comparison between VOD and soil moisture products in LPRM confirmed the above hypothesis (*i.e.*, the higher the VOD values, the wetter the soil moisture retrievals). For instance, the VOD values of LPRM-X, -C1, and -C2 showed a decreasing trend during nighttime (01:30 local time, see c1 in Fig. A3), and a similar decreasing trend in the bias values (c2 in Fig. 3) of LPRM products in different bands from LPRM-X (0.080 m^3/m^3), -C1 (0.066 m^3/m^3), and -C2 (-0.002 m^3/m^3). The above phenomenon illustrated that apart from the systematic overestimation of LPRM temperature input (see Fig. A4, the comparison between physical surface temperature products and *in situ* measurements within the SMN-SDR), the high VOD values could also be a key factor of the overestimation of ground measurements by the LPRM products.

The comparatively lower VOD values may be one of the reasons for the dryness of the low-frequency passive microwave products in the SMN-SDR. For example, the underestimation of ECMWF temperature (see Fig. A4) and the relatively smaller range of VOD values resulted in the largest negative bias of SMOS-IC (mentioned in Section 4.1.1). Although GEOS-5 temperature systematically overestimated *in situ* soil temperature (Fig. A4), the dryness of SMAP products (Fig. 2) indicated that vegetation correction (low VOD values) may be the main factor for the underestimation of ground measurements by the SMAP retrieval algorithms in the SMN-SDR. The above assumption can be confirmed in the inter-comparison between the VOD values estimated by SMAP-IB and other SMAP algorithms (SCA and MDCA). That is, the larger the SMAP-IB VOD estimates (see a1 in Fig. 9), the wetter the SMAP-IB soil moisture retrievals (see a1 and a2 in Fig. 2).

It can also be observed that the VOD time-series of SMAP-SCA (a1 and c1 in Fig. 9) were smoother than for the other VOD products since the SMAP-SCA algorithm estimates the VOD based on an empirical relationship with the NDVI climatology (Bindlish et al., 2011). Furthermore, the VOD products of SCA-H and SCA-V had a good ability to identify vegetation seasonal variability, reflected by the gradually increasing VOD values in spring (March), and gradually decreasing VOD after reaching the highest vegetation density in summer (August). This was similar to the LPRM VOD results, but the time to reach the LPRM VOD peak (September) was a little later than the SMAP-SCA. The VOD values of SMOS were a bit unstable and noisy, with weak seasonal variations. SMOS retrievals were as noisy as their VOD, which may be due to the influence of RFI or multiple minima issues (Wigneron et al., 2021).

5.2. Impacts of local acquisition time

Diurnal variations in ground conditions may affect the performance of soil moisture products at different local acquisition times. Hence, the full identification of the impact of local acquisition time on soil moisture product accuracy is of great significance to the design of future satellite missions and the development of retrieval algorithms. Fig. 10 presents the ground-based metrics (R, bias, and ubRMSE) of soil moisture datasets during the daytime and nighttime, including the raw soil moisture estimates (a, b, and c in Fig. 10) and short-term anomalies (d and e). Bias estimates were only calculated for raw soil moisture time-series since it is trivial for anomalies (Gruber et al., 2020). The statistical uncertainty of ground-based metrics was estimated by a 1000-member MBB distribution. Error bars in Fig. 10 represent the 5th and 95th percentile range of the above bootstrapping distribution. By analyzing the error metrics in Fig. 10, this study aims to shed light on the following two key points: (1) the impact of local acquisition time on soil moisture product accuracy within the SMN-SDR, (2) the performance summary of soil moisture products and possible causes of retrieval uncertainties (Section 5.3).

For low-frequency passive microwave soil moisture products, nighttime products (18:00 local time) had smaller ubRMSE (blue circle in c and e of Fig. 10) than daytime, while the daytime products (06:00 local time) had higher R values (red symbol in a and d of Fig. 10) than nighttime, especially for SMOS. Similarly, there is not a unified conclusion about the influence of local acquisition time on the high-frequency passive microwave products within the SMN-SDR. Although the nighttime products (01:30 and 02:00 local time) of LPRM-C1, -C2, and FY-3D had a relatively larger R, the ubRMSE values of their daytime (13:30 and 14:00 local time) products were smaller. The LPRM-X, JAXA, and FY-3B performed better during the nighttime (01:30 local time), with lower ubRMSE and higher R values. On the contrary, the daytime performance (10:00 local time) was more reliable for FY-3C, with smaller ubRMSE and higher R values than the nighttime products (20:00 local time). The daytime and nighttime products performed similarly for ASCAT and three model-based products in the SMN-SDR, except for the slightly higher R values in the short-term anomalies during the daytime (red symbol in d of Fig. 10).

It is generally believed that passive microwave soil moisture products at nighttime (or early morning) are more accurate, because the soil-vegetation temperature difference during the daytime may cause errors in the soil moisture retrievals (Jackson et al., 2010). However, the comparison results showed that isothermal conditions during nighttime (or early morning) do not have a decisive impact on soil moisture retrievals accuracy within the SMN-SDR, in agreement with previous studies (Lei et al., 2015; Zeng et al., 2015). For instance, the magnitude of RFI contamination at different overpass times may also affect the quality of soil moisture retrievals (Zhao et al., 2015b).

5.3. Other factors and performance summary

The evaluation uncertainties may suffer from the mismatch between *in situ* measurements and grid-based soil moisture products (Chen et al., 2017). Therefore, some measures have been taken to mitigate the mismatch in this study, including the spatial-temporal resampling method (Section 3.1.2), the sensor calibration, and the measuring depth selection (Section 2.1). Hence, uncertainties in retrieval algorithms (inevitable constraints and assumptions) and auxiliary parameters (such as physical surface temperature, vegetation parameters, surface roughness, and soil texture mainly in satellite soil moisture retrievals) are likely the remaining reasons for the different performances of soil moisture products. In addition, instrument characteristics (different bands) and external interference (RFI) may have affected the performance of satellite soil moisture products (Zhao et al., 2015a). According

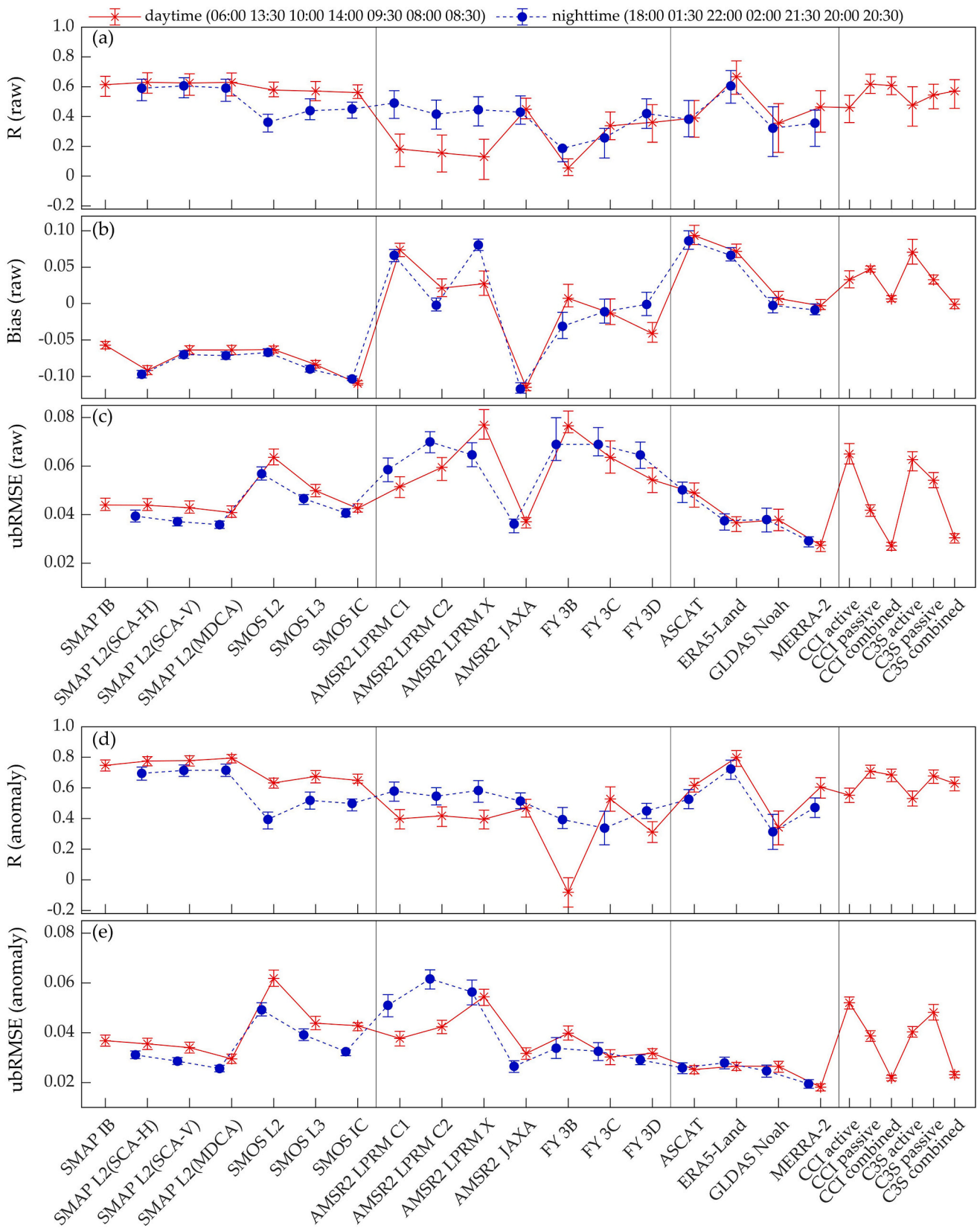

Fig. 10. Summary statistics of ground-based R, bias, and ubRMSE values, for raw soil moisture estimates (a, b, and c) and short-term anomalies (d and e) in 24 soil moisture datasets. The red and blue symbols represent the daytime and nighttime performance, respectively. The timestamps of soil moisture products are listed in the legend. Error bars represent the 5th and 95th percentile spread of a 1000-member moving block bootstrap sampling distribution. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

to the ground-based metrics from the SMN-SDR (Fig. 10), the inferences of the soil moisture products evaluated in this study are summarized as follows.

Among single-sensor satellite-based soil moisture products within the SMN-SDR, low-frequency passive microwave products (SMAP and SMOS) at L-band were generally better than C/X-band higher-frequency passive microwave retrievals in terms of R and ubRMSE. This is not surprising as L-band signals have a deeper penetration depth and the higher sensitivity to soil moisture. Although SMAP retrievals underestimated the ground measurements in the SMN-SDR, SMAP products performed best among all single-sensor satellite-based products, and its ubRMSE values basically met the mission requirement ($< 0.04\ m^3/m^3$) along with the higher R values (> 0.6). The good performance of SMAP has been validated by extensive research (Beck et al., 2021; Ma et al., 2019; Ye et al., 2021), and the dryness of SMAP is also in agreement with previous research (Al-Yaari et al., 2019; Chen et al., 2017). The deeper penetration depth of the L-band, the smaller radiometric error (about 1.3 K) of the sensor (Entekhabi et al., 2010), more accurate input parameters (such as GEOS-5 temperature data, see Fig. A4), and the utilization of RFI mitigation hardware and software (Chan et al., 2016) may be some of the reasons for the better performance of SMAP. The physical surface temperature is a key input in the passive microwave SM retrieval algorithms (Zhao et al., 2015a). The behavior of GEOS-5 and ECMWF temperature products were consistent with ground measurements in the SMN-SDR (see Fig. A4), while the temperature estimates from LPRM performed the worst with the highest ubRMSE (2.297 and 4.750 K) and the lowest R (0.825 and 0.939). The comparison result of the above three temperature products was consistent with previous studies (Beck et al., 2021; Cui et al., 2017; Zeng et al., 2015).

Remarkably, the ubRMSE values (around $0.04\ m^3/m^3$) of SMAP in the SMN-SDR were similar to that of validation in the dense networks (Chen et al., 2017; Colliander et al., 2017), which is better than that of the sparse network ($> 0.05\ m^3/m^3$) (Xing et al., 2021). These dense networks are considered to have reasonable representativeness for the satellite footprint scale (Chen et al., 2019; Gruber et al., 2020). This showed that the regional-scale validation in this study effectively reduced the inherent uncertainties in the scale difference (*i.e.*, spatial-scale mismatch and/or depth mismatch between *in situ* points and satellite pixels, see Section 3.1.2). The validation results at a 0.25° grid further confirmed the above inference. The selected validation grid is a 0.25° grid with the highest site density over the SMN-SDR (the bold purple square in Fig. A5), which installed a total of eight small-scale stations and three medium-scale stations. Fig. A6 and Table A3 (A4) present the performance metrics of soil moisture products over the selected 0.25° validation grid. Although different validation scales did not alter performance differences between soil moisture products (1° and 0.25° in this study), the error metrics calculated at the 0.25° grid were slightly worse, with smaller R and larger ubRMSE values compared to that at 1°.

In contrast to the satisfactory SMAP retrievals, the performance of SMOS products, which are also generated from L-band observations, was slightly inferior in the SMN-SDR. Larger radiometric error (about 4 K) and more RFI contamination lead to unstable SMOS retrievals, particularly in China (Wigneron et al., 2021). Among the three SMOS products, SMOS-IC had the best performance in terms of ubRMSE and R values because of some improvements in the algorithm and the application of an RFI filtering tool (Wigneron et al., 2021). The largest negative bias of SMOS-IC among all SMOS products was consistent with previous studies (Fernandez-Moran et al., 2017). Table A5 shows the comparison between FAO (the Food and Agriculture Organization) soil texture data (Nachtergaele et al., 2012) and ground-based soil texture measurements during the SMN-SDR maintenance in 2019. The comparison result within SMN-SDR indicated that the FAO data tended to underestimate sand content with a negative median bias of −26.72%, ranging from −58.21% to 29.39%. The clay and silt contents derived from the FAO data were found to be overestimated with a positive median bias of 17.68% and 6.07%, respectively. This performance of FAO soil texture information was consistent with a previous study on the Tibetan Plateau (Xing et al., 2021). The underestimated sand content would lead to an overestimation of soil moisture. Therefore, the underestimated soil moisture by SMOS is not due to the uncertainty in soil texture, but to the underestimation of VOD and soil temperature within the SMN-SDR.

The high-frequency passive microwave soil moisture products (AMSR2 and FY-3) did not obtain favorable results within the SMN-SDR. The C/X band shallow penetration depth, the possible RFI contamination, errors in auxiliary information (such as LPRM temperature data, see Fig. A4), and the uncalibrated model parameters (such as the single scattering albedo and surface roughness) are likely the reasons for the low accuracy of the high-frequency passive microwave products. The penetration depth dependence on frequency is a possible reason for the different performance of three LPRM products, while different satellite overpassing times (Liu et al., 2021), the amount of available data (see Table 2), the localization accuracy, and radiometer stability are likely to cause the quality difference between three FY-3 products.

The wetness of ASCAT within the SMN-SDR is in line with previous work (Morrison and Wagner, 2020; Zeng et al., 2015). The unit conversion (*i.e.*, the accuracy of soil texture information) and active-sensor characteristics (*i.e.*, the higher sensitivity to sub-surface heterogeneities and surface roughness) may cause the deviation of ASCAT (Brocca et al., 2011; Morrison and Wagner, 2020; Wagner et al., 2013). It is worth noting that the relative trends in multi-source product performance reflected by the raw soil moisture estimates and short-term anomalies were very similar (see Fig. 10). However, skill metrics of ASCAT short-term anomalies were improved compared with the raw soil moisture estimates, which were similar to SMOS and better than high-frequency passive products. This is likely due to the removal of seasonal vegetation signals when soil moisture short-term anomalies are applied (Gruber et al., 2020).

The performance of model-based products could suffer partly from the uncertainties of model structure, forcing inputs, and model parameters (Chen et al., 2013). The capabilities of GLDAS-Noah and MERRA-2 in predicting soil moisture dynamics within the SMN-SDR could be further refined by improving both the quality of meteorological forcing data and the soil property database (Beck et al., 2021). The mismatch between the depth of the model soil layer and the ground observation depth also contributes to the low scores of modeled soil moisture datasets. The deeper layer soil moisture simulation (0–10 cm), compared to 3-cm measurements, may be attributed to the worst skill metrics of the GLDAS-Noah product (Beck et al., 2021). Since in each grid, the model uses prescribed soil properties, vegetation data (such as leaf area index and the fraction of photosynthetically active radiation climatology), topography, and so on, the capability of representation of horizontal heterogeneity in each model also influence the soil moisture simulation skill. MERRA-2 has the coarsest resolution and produces 'flat' soil moisture with the smallest dynamics in the SMN-SDR. The resolution of ERA5-Land is around 10-km and its soil moisture varies considerably. Moreover, models generally use a prescribed vegetation data, being the vegetation climatology derived from AVHRR (the Advanced Very High Resolution Radiometer) or MODIS (the Moderate Resolution Imaging Spectroradiometer), which may contribute to the errors of soil moisture simulation.

Some results that emphasize the advantages of multi-sensor merging technology were shown in Fig. 10, with CCI-combined having better performance, including the highest R and smallest ubRMSE values among all the evaluated products. However, the R values of CC-combined were slightly lower than SMAP, likely due to the uncertainty of the GLDAS-Noah model during the unit scaling and the TCA hypothetical destruction during the merging procedure (Kim et al., 2020).

6. Conclusion

Dense measuring networks are considered to provide the best possible reference data for grid-based soil moisture products. To provide long-term ground measurements for satellite- and model-based soil moisture products, a new soil moisture and temperature wireless sensor network (the SMN-SDR) was established within the Shandian River basin, consisting of 34 stations within an area of ~10,000 km^2 and presenting a three-sampling-scale nested structure. This study is the first to evaluate and compare 24 different soil moisture datasets based on the *in situ* measurements from the SMM-SDR. The validation metrics were separately calculated for both raw soil moisture estimates and short-term anomalies. In addition, the moving block bootstrap re-sampling method was used to estimate the 90% confidence interval of ground-based metrics to eliminate sampling uncertainty. The main findings in this study were as follows:

(1) Inter-comparison of soil moisture products within the SMN-SDR: The newly-released CCI-combined product (v06.1, merging extra SMAP retrievals) had the best performance among all the evaluated products, with the lowest ubRMSE (< 0.03 m^3/m^3), highest R (> 0.6), and close-to-zero bias. Although SMAP underestimates the ground measurements, it had the lowest ubRMSE (close to 0.04 m^3/m^3) along with the highest R values (> 0.6) among all the single-sensor products. Among the three retrieval algorithms of SMAP-L2, MDCA was slightly better than SCA-V (baseline algorithm for the beta release), while SCA-H had the largest error. The latest SMOS-IC had the lowest ubRMSE and bias among the three SMOS products. Conversely, the high-frequency passive microwave products (AMSR2 and FY-3), the active-sensor ASCAT, and three model-based products (ERA5-land, GLDAS-Noah, and MERRA2) did not get favorable results.

(2) The performance of SMAP-IB within the SMN-SDR: This study is the first attempt to compare the recently developed SMAP-IB with other soil moisture products. The R (around 0.6) and ubRMSE (close to 0.04 m^3/m^3) values of SMAP-IB were basically at the same level as SMAP. However, the improved bias metrics (and higher VOD values) of SMAP-IB compared with other drier SMAP products were very encouraging as they show the great potential of the SMAP-IB product development. Moreover, the feature of SMAP-IB not requiring any auxiliary data is key to developing products independent of additional information sources such as vegetation optical index and carrying out applications of SMAP L-VOD in monitoring vegetation dynamics.

(3) The impact of local acquisition time: The ubRMSE values of SMOS and SMAP nighttime products (18:00) were small, while the R values of daytime products (06:00) were larger than nighttime. For high-frequency passive microwave retrievals, the nighttime products had higher R values in most cases. Obvious differences were not observed between the daytime and nighttime performance of ASCAT and three model-based products.

(4) The impacts of VOD: The fact that SMAP and SMOS-IC underestimated *in situ* measurements indicates that the vegetation correction (the smaller VOD values) might be a major factor causing the dryness of SMAP in the SMN-SDR.

(5) Validation of TCA: The TCA-based metrics were generally very similar to the ground-based metrics (see Fig. 7 and A1), showing the feasibility of TCA for evaluating different soil moisture products without *in situ* information. Although the most conservative triplet collection (an active product, a passive product, and a model-based product) was applied, the TCA assumptions may still be undermined, caused by the depth mismatch between different soil moisture products or inevitable error cross-correlations. Averaging the TCA-based metrics of multiple freedom degrees is an effective method to reduce the uncertainty of a single TCA estimate.

To eliminate the spatial-scale mismatch, both *in situ* measurements and grid-based soil moisture products were converted into areal average estimates within the SMN-SDR through the spatial match-up (see Section 3.1.2). Therefore, all the above evaluation results were based on regional scale comparisons, including the TCA estimates. The above analysis of TCA can be extended to pixel-scale evaluation to obtain more accurate error characteristics of soil moisture products on the satellite footprint scale. Investigating the error source of soil moisture retrievals is set on a basis that is more qualitative than quantitative in this study. Different ways such as sensitivity analysis of retrieval algorithms and/or applying the QCA method to calculate the error cross-correlations between datasets may be able to provide quantitative evidence to promote the development of future soil moisture products and satellite missions.

Data availability

The recent two-year (01/09/2018 to 31/12/2020) *in situ* data are freely accessible online *via* the International Soil Moisture Network (https://ismn.geo.tuwien.ac.at/en/networks/?id=SMN-SDR) and the National Tibetan Plateau/Third Pole Environment Data Center (doi:10.11888/Soil.tpdc.271425; doi:10.11888/Soil.tpdc.271434). The user guide document and metadata (including geographic location, elevation, land use type, soil texture, site photographs, and calibration function of sensors, *etc.*) for interpretation of the SMN-SDR measurements are also available for download. Data collection of all stations within the SMN-SDR is ongoing with 15-min time steps. The subsequent data publishing will be completed in due course, with roughly a one-year period for data protection.

CRediT authorship contribution statement

Jingyao Zheng: Software, Writing – original draft, Data curation, Investigation, Writing – review & editing, Validation. **Tianjie Zhao:** Conceptualization, Software, Data curation, Investigation, Writing – review & editing. **Haishen Lü:** Supervision, Funding acquisition, Resources. **Jiancheng Shi:** Supervision, Funding acquisition, Resources. **Michael H. Cosh:** Formal analysis, Writing – review & editing. **Dabin Ji:** Data curation, Resources. **Lingmei Jiang:** Formal analysis, Writing – review & editing. **Qian Cui:** Data curation, Resources. **Hui Lu:** Formal analysis, Writing – review & editing. **Kun Yang:** Formal analysis, Writing – review & editing. **Jean-Pierre Wigneron:** Formal analysis, Writing – review & editing. **Xiaojun Li:** Formal analysis, Writing – review & editing. **Yonghua Zhu:** Investigation. **Lu Hu:** Investigation. **Zhiqing Peng:** Investigation. **Yelong Zeng:** Formal analysis. **Xiaoyi Wang:** Formal analysis. **Chuen Siang Kang:** Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (NSFC) (Grant nos. 42090014, 41830752, 42071033, and 41971042), National Key Research and Development Program (Grant Nos. 2018YFA0605400), and the U.S. Department of Agriculture, Agricultural Research Service. USDA is an equal opportunity provider and employer. The author sincerely thanks Dr. Stefan Kern of ICDC, Center for Earth System Research and Sustainability, and the University of Hamburg for providing ASCAT data. All H-SAF products are owned by EUMETSAT, and the EUMETSAT SAF Data Policy applies. They are available for all users free of charge.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2022.112891.

References

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.

Albergel, C., de Rosnay, P., 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-Yaari, A., Wigneron, J.-P., Ducharne, A., Kerr, Y., de Rosnay, P., de Jeu, R., Govind, A., Al Bitar, A., Albergel, C., Muñoz-Sabater, J., Richaume, P., Mialon, A., 2014. Global-scale evaluation of two satellite-based passive microwave soil moisture datasets (SMOS and AMSR-E) with respect to Land Data Assimilation System estimates. Remote Sens. Environ. 149, 181–195. https://doi.org/10.1016/j.rse.2014.04.006.

Al-Yaari, A., Wigneron, J.-P., Dorigo, W., Colliander, A., Pellarin, T., Hahn, S., Mialon, A., Richaume, P., Fernandez-Moran, R., Fan, L., Kerr, Y.H., De Lannoy, G., 2019. Assessment and inter-comparison of recently developed/reprocessed microwave satellite soil moisture products using ISMN ground-based measurements. Remote Sens. Environ. 224, 289–303. https://doi.org/10.1016/j.rse.2019.02.008.

Babaeian, E., Sadeghi, M., Jones, S.B., Montzka, C., Vereecken, H., Tuller, M., 2019. Ground, proximal, and satellite remote sensing of soil moisture. Rev. Geophys. 57, 530–616. https://doi.org/10.1029/2018RG000618.

Beaudoing, H., Rodell, M., 2020. GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1. https://doi.org/10.5067/E7TYRXPJKWOQ.

Beck, H.E., Pan, M., Miralles, D.G., Reichle, R.H., Dorigo, W.A., Hahn, S., Sheffield, J., Karthikeyan, L., Balsamo, G., Parinussa, R.M., van Dijk, A.I.J.M., Du, J., Kimball, J. S., Vergopolan, N., Wood, E.F., 2021. Evaluation of 18 satellite- and model-based soil moisture products using in situ measurements from 826 sensors. Hydrol. Earth Syst. Sci. 25, 17–40. https://doi.org/10.5194/hess-25-17-2021.

Bindlish, R., Jackson, T., Zhao, T., 2011. A MODIS-based vegetation index climatology. In: Gao, W., Jackson, T.J., Wang, J., Chang, N.-B. (Eds.), Remote Sensing and Modeling of Ecosystems for Sustainability VIII. SPIE, pp. 41–48.

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.

C3S, 2020. Soil moisture gridded data from 1978 to present available at: https://cds.climate.copernicus.eu.

Chan, S.K., Bindlish, R., O'Neill, P.E., Njoku, E., Jackson, T., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Yueh, S., Entekhabi, D., Cosh, M.H., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., González-Zamora, Á., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Palecki, M., Small, E.E., Zreda, M., Calvet, J.-C., Crow, W.T., Kerr, Y., 2016. Assessment of the SMAP passive soil moisture product. IEEE Trans. Geosci. Remote Sens. 54, 4994–5007. https://doi.org/10.1109/TGRS.2016.2561938.

Chen, Y., Yang, K., Qin, J., Zhao, L., Tang, W., Han, M., 2013. Evaluation of AMSR-E retrievals and GLDAS simulations against observations of a soil moisture network on the central Tibetan Plateau. J. Geophys. Res.-Atmos. 118, 4466–4475. https://doi.org/10.1002/jgrd.50301.

Chen, Y., Yang, K., Qin, J., Cui, Q., Lu, H., La, Z., Han, M., Tang, W., 2017. Evaluation of SMAP, SMOS, and AMSR2 soil moisture retrievals against observations from two networks on the Tibetan Plateau. J. Geophys. Res. Atmos. Res 175, 238. https://doi.org/10.1038/175238c0.

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., Cosh, M.H., Colliander, A., Asanuma, J., Berg, A., Bosch, D.D., Caldwell, T.G., Collins, C.H., Jensen, K.H., Martínez-Fernández, J., McNairn, H., Starks, P.J., Su, Z., Walker, J.P., 2019. Uncertainty of reference pixel soil moisture averages sampled at SMAP core validation sites. J. Hydrometeorol. 20, 1553–1569. https://doi.org/10.1175/JHM-D-19-0049.1.

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 https://doi.org/10.1016/j.rse.2017.01.021.

Coopersmith, E.J., Cosh, M.H., Starks, P.J., Bosch, D.D., Holifield Collins, C., Seyfried, M., Livingston, S., Prueger, J., 2021. Understanding temporal stability: a long-term analysis of USDA ARS watersheds. Int. J. Digital Earth 14, 1243–1254. https://doi.org/10.1080/17538947.2021.1943550.

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

Cui, H., Jiang, L., Du, J., Zhao, S., Wang, G., Lu, Z., Wang, J., 2017. Evaluation and analysis of AMSR-2, SMOS, and SMAP soil moisture products in the Genhe area of China. J. Geophys. Res.-Atmos. 122, 8650–8666. https://doi.org/10.1002/2017JD026800.

Dorigo, W.A., Gruber, A., De Jeu, R.A.M., 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.

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.

Fernandez-Moran, R., Al-Yaari, A., Mialon, A., Mahmoodi, A., Al Bitar, A., De Lannoy, G., Rodriguez-Fernandez, N., Lopez-Baeza, E., Kerr, Y., Wigneron, J.-P., 2017. SMOS-IC: an alternative SMOS soil moisture and vegetation optical depth product. Remote Sens. https://doi.org/10.3390/rs9050457.

Fujii, H., Koike, T., Imaoka, K., 2009. Improvement of the AMSR-E algorithm for soil moisture estimation by introducing a fractional vegetation coverage dataset derived from MODIS data. J. Remote Sens. Soc. Jpn. 29, 282–292.

Gelaro, R., McCarty, W., Suarez, M.J., Todling, R., Molod, A., Takacs, L., Randles, C.A., Darmenov, A., Bosilovich, M.G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A.M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J.E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S.D., Sienkiewicz, M., Zhao, B., 2017. The Modern-era retrospective analysis for research and applications, Version 2 (MERRA-2). J. Clim. 30, 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1.

González-Zamora, Á., Sánchez, N., Pablos, M., Martínez-Fernández, J., 2019. CCI soil moisture assessment with SMOS soil moisture and in situ data under different environmental conditions and spatial scales in Spain. Remote Sens. Environ. 225, 469–482. https://doi.org/10.1016/j.rse.2018.02.010.

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., Wagner, W., 2016. 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., Scanlon, T., van der Schalie, R., Wagner, W., Dorigo, W., 2019. Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology. Earth Syst. Sci. Data 11, 717–739. https://doi.org/10.5194/essd-11-717-2019.

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.

Imaoka, K., Kachi, M., Fujii, H., Murakami, H., Hori, M., Ono, A., Igarashi, T., Nakagawa, K., Oki, T., Honda, Y., Shimoda, H., 2010. Global change observation mission (GCOM) for monitoring carbon, water cycles, and climate change. Proc. IEEE 98, 717–734. https://doi.org/10.1109/JPROC.2009.2036869.

Jackson, T.J., Cosh, M.H., Bindlish, R., Starks, P.J., Bosch, D.D., Seyfried, M., Goodrich, D.C., Moran, M.S., Du, J., 2010. Validation of advanced microwave scanning radiometer soil moisture products. IEEE Trans. Geosci. Remote Sens. 48, 4256–4272. https://doi.org/10.1109/TGRS.2010.2051035.

Jackson, T.J., Bindlish, R., Cosh, M.H., Zhao, T., Starks, P.J., Bosch, D.D., Seyfried, M., Moran, M.S., Goodrich, D.C., Kerr, Y.H., Leroux, D., 2012. Validation of soil moisture and ocean salinity (SMOS) soil moisture over watershed networks in the U.S. IEEE Trans. Geosci. Remote Sens. 50, 1530–1543. https://doi.org/10.1109/TGRS.2011.2168533.

Jackson, T.J., Cosh, M., Crow, W., 2014. In: Lakshmi, V. (Ed.), Some Issues in Validating Satellite-Based Soil Moisture Retrievals from SMAP with In Situ Observations in Remote Sensing of the Terrestrial Water Cycle. American Geophysical Union, Wiley Books.

Kang, C.S., Zhao, T., Shi, J., Cosh, M.H., Chen, Y., Starks, P.J., Collins, C.H., Wu, S., Sun, R., Zheng, J., 2021. Global soil moisture retrievals from the Chinese FY-3D microwave radiation imager. IEEE Trans. Geosci. Remote Sens. 1–15 https://doi.org/10.1109/tgrs.2020.3019408.

Kern, S., 2021. ASCAT Global Maps of daily running 5-daily mean surface soil moisture. https://doi.org/10.25592/UHHFDM.8681.

Kerr, Y.H., Waldteufel, P., Wigneron, J., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M., Font, J., Reul, N., Gruhier, C., Juglea, S.E., Drinkwater, M.R., Hahne, A., Martín-Neira, M., Mecklenburg, S., 2010. The SMOS mission: new tool for monitoring key elements ofthe global water cycle. Proc. IEEE 98, 666–687. https://doi.org/10.1109/JPROC.2010.2043032.

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.

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.P., Kumar, S., Dong, J., Wagner, W., Cosh, M.H., Bosch, D.D., Collins, C.H., Starks, P.J., Seyfried, M., Lakshmi, V., 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.
Lei, F., Crow, W.T., Shen, H., Parinussa, R.M., Holmes, T.R.H., 2015. The impact of local acquisition time on the accuracy of microwave surface soil moisture retrievals over the contiguous United States. Remote Sens. https://doi.org/10.3390/rs71013448.
Li, X., Wigneron, J.P., Fan, L., Frappart, F.,Simon H.Y., Colliander, A., Ebtehaj A., Gao L., Fernandez-Moran R., Liu X.Z., Wang M.j., Ma H.l., Moisy C., Ciais P. A new SMAP soil moisture and vegetation optical depth product (SMAP-IB): Algorithm, assessment and inter-comparison. Remote Sens. Environ. In Review.
Liu, Y., Zhou, Y., Lu, N., Tang, R., Liu, N., Li, Y., Yang, J., Jing, W., Zhou, C., 2021. Comprehensive assessment of Fengyun-3 satellites derived soil moisture with in-situ measurements across the globe. J. Hydrol. 594, 125949 https://doi.org/10.1016/j.jhydrol.2020.125949.
Ma, H., Zeng, J., Chen, N., Zhang, X., Cosh, M.H., Wang, W., 2019. Satellite surface soil moisture from SMAP, SMOS, AMSR2 and ESA CCI: a comprehensive assessment using global ground-based observations. Remote Sens. Environ. 231, 111215 https://doi.org/10.1016/j.rse.2019.111215.
McColl, K.A., Vogelzang, J., Konings, A.G., Entekhabi, D., Piles, M., Stoffelen, A., 2014. Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. Geophys. Res. Lett. 41, 6229–6236. https://doi.org/10.1002/2014GL061322.
Montzka, C., Cosh, M., Bayat, B., Al Bitar, A., Berg, A., Bindlish, R., Bogena, H.R., Bolten, J.D., Cabot, F., Caldwell, T., 2020. Soil moisture product validation good practices protocol version 1.0. Good Pract. Satellite Derived Land Product Validat. 123.
Morrison, K., Wagner, W., 2020. Explaining anomalies in SAR and scatterometer soil moisture retrievals from dry soils with subsurface scattering. IEEE Trans. Geosci. Remote Sens. 58, 2190–2197. https://doi.org/10.1109/TGRS.2019.2954771.
Muñoz Sabater, J., 2019. ERA5-Land hourly data from 1950 to present. https://doi.org/10.24381/cds.e2161bac.
Nachtergaele, F., van Velthuizen, H., Verelst, L., Batjes, N., Dijkshoorn, K., van Engelen, V., Fischer, G., Jones, A., Montanarella, L., Petri, M., Prieler, S., Shi, X., Teixeira, E., Wiberg, D., 2012. Harmonized World Soil Database Version 1.2. Food and Agriculture Organization of the United Nations (FAO). 19th World Congress of Soil Science, Soil Solutions for a Changing World.
Ólafsdóttir, K.B., Mudelsee, M., 2014. More accurate, calibrated bootstrap confidence intervals for estimating the correlation between two time series. Math. Geosci. 46, 411–427. https://doi.org/10.1007/s11004-014-9523-4.
O'Neill, P.E., Chan, S., Njoku, E.G., Jackson, T., Bindlish, R., Chaubell, J., 2020. SMAP L2 Radiometer Half-Orbit 36 km EASE-Grid Soil Moisture. Version 7. https://doi.org/10.5067/F1TZ0CBN1F5N.
Owe, M., de Jeu, R., Walker, J., 2001. A methodology for surface soil moisture and vegetation optical depth retrieval using the microwave polarization difference index. IEEE Trans. Geosci. Remote Sens. 39, 1643–1654. https://doi.org/10.1109/36.942542.
Paloscia, S., Pettinato, S., Santi, E., Notarnicola, C., Pasolli, L., Reppucci, A., 2013. Soil moisture mapping using Sentinel-1 images: algorithm and preliminary validation. Remote Sens. Environ. 134, 234–248. https://doi.org/10.1016/j.rse.2013.02.027.
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. Select. Top. Appl. Earth Observ. Remote Sens. 10, 4522–4530. https://doi.org/10.1109/JSTARS.2017.2714025.
Rodell, M., Houser, P.R., Jambor, U.E.A., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., 2004. The global land data assimilation system. Bull. Am. Meteorol. Soc. 85, 381–394.
Sabaghy, S., Walker, J.P., Renzullo, L.J., Akbar, R., Chan, S., Chaubell, J., Das, N., Dunbar, R.S., Entekhabi, D., Gevaert, A., Jackson, T.J., Loew, A., Merlin, O., Moghaddam, M., Peng, Jian, Peng, Jinzheng, Piepmeier, J., Rüdiger, C., Stefan, V., Wu, X., Ye, N., Yueh, S., 2020. Comprehensive analysis of alternative downscaled soil moisture products. Remote Sens. Environ. 239, 111586 https://doi.org/10.1016/j.rse.2019.111586.
Shi, J., Jiang, L., Zhang, L., Chen, K.S., Wigneron, J.P., Chanzy, A., Jackson, T.J., 2006. Physically based estimation of bare-surface soil moisture with the passive radiometers. IEEE Trans. Geosci. Remote Sens. 44, 3145–3153. https://doi.org/10.1109/TGRS.2006.876706.
Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Züger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Blöschl, G., Eitzinger, J., Steinnocher, K., 2013. The ASCAT soil moisture product: a review of its specifications, validation results, and emerging applications. Meteorol. Z. 22, 5–33. https://doi.org/10.1127/0941-2948/2013/0399.
Wang, Z., Zhao, T., Qiu, J., Zhao, X., Li, R., Wang, S., 2021. Microwave-based vegetation descriptors in the parameterization of water cloud model at L-band for soil moisture retrieval over croplands. GIScience Remote Sens. 58, 48–67. https://doi.org/10.1080/15481603.2020.1857123.
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, 112238 https://doi.org/10.1016/j.rse.2020.112238.
Wu, X., Walker, J.P., Das, N.N., Panciera, R., Rüdiger, C., 2014. Evaluation of the SMAP brightness temperature downscaling algorithm using active–passive microwave observations. Remote Sens. Environ. 155, 210–221. https://doi.org/10.1016/j.rse.2014.08.021.
Xing, Z., Fan, L., Zhao, L., De Lannoy, G., 2021. Remote Sensing of Environment A first assessment of satellite and reanalysis estimates of surface and root-zone soil moisture over the permafrost region of Qinghai-Tibet Plateau, p. 265. https://doi.org/10.1016/j.rse.2021.112666.
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., Zhu, L., 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. Am. Meteorol. Soc. 94, 1907–1916. https://doi.org/10.1175/BAMs-d-12-00203.1.
Ye, N., Walker, J.P., Bindlish, R., Chaubell, J., Das, N.N., Gevaert, A.I., Jackson, T.J., Rüdiger, C., 2019. Evaluation of SMAP downscaled brightness temperature using SMAPEx-4/5 airborne observations. Remote Sens. Environ. 221, 363–372. https://doi.org/10.1016/j.rse.2018.11.033.
Ye, N., Walker, J.P., Wu, X., de Jeu, R., Gao, Y., Jackson, T.J., Jonard, F., Kim, E., Merlin, O., Pauwels, V.R.N., Renzullo, L.J., Rüdiger, C., Sabaghy, S., von Hebel, C., Yueh, S.H., Zhu, L., 2021. the soil moisture active passive experiments: validation of the SMAP products in Australia. IEEE Trans. Geosci. Remote Sens. 59, 2922–2939. https://doi.org/10.1109/TGRS.2020.3007371.
Yee, M.S., Walker, J.P., Monerris, A., Rüdiger, C., Jackson, T.J., 2016. On the identification of representative in situ soil moisture monitoring stations for the validation of SMAP soil moisture products in Australia. J. Hydrol. 537, 367–381. https://doi.org/10.1016/j.jhydrol.2016.03.060.
Zeng, J., Li, Z., Chen, Q., Bi, H., Qiu, J., Zou, P., 2015. Evaluation of remotely sensed and reanalysis soil moisture products over the Tibetan Plateau using in-situ observations. Remote Sens. Environ. 163, 91–110. https://doi.org/10.1016/j.rse.2015.03.008.
Zhao, T., Shi, J., Bindlish, R., Jackson, T., Cosh, M., Jiang, L., Zhang, Z., Lan, H., 2015a. Parametric exponentially correlated surface emission model for L-band passive microwave soil moisture retrieval. Physics Chem. Earth, A/B/C 83–84, 65–74. https://doi.org/10.1016/j.pce.2015.04.001.
Zhao, T., Shi, J., Bindlish, R., Jackson, T.J., Kerr, Y.H., Cosh, M.H., Cui, Q., Li, Y., Xiong, C., Che, T., 2015b. Refinement of SMOS multiangular brightness temperature toward soil moisture retrieval and its analysis over reference targets. IEEE J. Select. Top. Appl. Earth Observ. Remote Sens. 8, 589–603. https://doi.org/10.1109/JSTARS.2014.2336664.
Zhao, T., Shi, J., Lv, L., Xu, H., Chen, D., Cui, Q., Jackson, T.J., Yan, G., Jia, L., Chen, L., Zhao, K., Zheng, X., Zhao, L., Zheng, C., Ji, D., Xiong, C., Wang, T., Li, R., Pan, J., Wen, J., Yu, C., Zheng, Y., Jiang, L., Chai, L., Lu, H., Yao, P., Ma, J., Lv, H., Wu, J., Zhao, W., Yang, N., Guo, P., Li, Y., Hu, L., Geng, D., Zhang, Z., 2020. Soil moisture experiment in the Luan River supporting new satellite mission opportunities. Remote Sens. Environ. 240 https://doi.org/10.1016/j.rse.2020.111680.
Zhao, T., Shi, J., Entekhabi, D., Jackson, T.J., Hu, L., Peng, Z., Yao, P., Li, S., Kang, C.S., 2021. Retrievals of soil moisture and vegetation optical depth using a multi-channel collaborative algorithm. Remote Sens. Environ. 257 https://doi.org/10.1016/j.rse.2021.112321.
Zhu, Y., Li, X., Pearson, S., Wu, D., Sun, R., Johnson, S., Wheeler, J., Fang, S., 2019. Evaluation of Fengyun-3C soil moisture products using in-Situ data from the Chinese Automatic Soil Moisture Observation Stations: A case study in Henan Province, China. Water 11, 248.