Student thesis series INES nr 607

Evaluation and merging of multiple gridded surface soil moisture products in Europe using ICOS measurements and triple collocation analysis

Hugo Bergman

2023

Department of
Physical Geography and Ecosystem Science
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Hugo Bergman (2023).

Evaluation and merging of multiple gridded surface soil moisture products in Europe using ICOS measurements and triple collocation analysis

Master degree thesis, 30 credits in *Geomatics*

Department of Physical Geography and Ecosystem Science, Lund University

Level: Master of Science (MSc)

Course duration: *January* 2023 until *June* 2023

Evaluation and merging of multiple gridded surface soil moisture products in Europe using ICOS measurements and triple collocation analysis

---

Hugo Bergman

Master thesis, 30 credits, in *Geomatics*

Supervisor:

Zheng Duan
Dep. of Physical Geography and Ecosystem Science, Lund University

Exam committee:

Wenxin Zhang
Dept. of Physical Geography and Ecosystem Science, Lund University

Xueying Li
Dept. of Physical Geography and Ecosystem Science, Lund University

Acknowledgements

To my supervisor, Zheng Duan, thank you for the excellent and invaluable guidance you offered throughout this project. A big thanks also goes to Jian Peng from Helmholtz Centre for Environmental Research in Germany for sharing code on TCA-weighted merging.

I would also like to show my appreciation for the soil moisture remote sensing community, including the Pytesmo contributors, for making software publicly available and user-friendly, thus opening the doors to the field. In addition, this project would not have been possible without the data providers, including ICOS, BEC, NASA, NSIDC, ESA with Copernicus and CCI, and EUMETSAT, which allowed the project to be conducted without any funding or additional licenses.

On a more general note, I would like to thank my professors, course mates, friends, and partner for the generous support I have received during the last couple of years in Lund.

ABSTRACT

Surface soil moisture (SM) is an essential climate variable that plays a key role in ecosystems and the energy, water, and carbon cycles. SM can be accurately measured using in situ measurements. However, these measurements are globally not densely located over large areas, which would be required for accurate large-scale SM estimation due to the high spatial variability of SM. Instead, global atmospheric models and satellite remote sensing in the microwave range are commonly utilised for large-scale SM monitoring. Both model and satellite approaches have resulted in multiple gridded SM products at regional or global scales at various spatial resolutions (typically between 1 and 40 km). The accuracy of the gridded products varies over different regions, climates, and land covers, necessitating their evaluation. Evaluation with in situ data is limited to areas where measurements are available. Over the past 15 years, triple collocation analysis (TCA) has been extensively applied to evaluate gridded SM products among different geophysical variables, as it can estimate the error structure of three independent datasets without the need for in situ measurements. TCA has also been used to successfully merge gridded products to generate more accurate SM estimates. This study evaluated and ranked eight gridded SM products, including SMOS L4, SMAP L3E, SMAP L4, Sentinel-1, ASCAT, ESA CCI SM, ERA5-Land, and GLDAS Noah, using in situ measurements of SM taken during 2020-2021 from the Integrated Carbon Observation System (ICOS) station network. SMAP L4 and ERA5-Land generally performed the best with similar statistical scores. When comparing the products against absolute SM on collocated dates, SMAP L4 had a median ubRMSD of ca 0.047 m^3/m^3 and a median correlation coefficient of 0.73. ESA CCI SM and SMAP L3 gave slightly worse scores, while GLDAS Noah showed a relatively poor correlation against short-term SM anomalies. Sentinel-1 generally had the worst performance, together with ASCAT and SMOS L4. In addition, TCA was performed using a triplet consisting of SMAP L3E, ASCAT, and GLDAS Noah. The TCA reinforced the results of the in situ-based evaluation, with the lowest ubRMSE and highest SNR found for SMAP L3E among the three products. The TCA of this triplet was also used to produce a weighted merged dataset of SM estimates. When compared to ICOS measurements, the merged product performed better than GLDAS Noah and ASCAT but similar to or worse than SMAP L3E. Additionally, the TCA-weighted product performed similarly to a simple arithmetic mean, indicating the merging process was not worthwhile in the study area.

Keywords*: evaluation, soil moisture, remote sensing, ICOS, triple collocation analysis, drought monitoring*

Table of Contents

Abbreviations

ASCAT	Advanced Scatterometer
BEC	Barcelona Expert Center
CCI	Climate Change Initiative
ECMWF	European Centre for Medium-Range Weather Forecasts
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GLDAS	Global Land Data Assimilation System
ICOS	Integrated Carbon Observation System
NSIDC	National Snow and Ice Data Center
SM	Soil Moisture
SMAP	Soil Moisture Active Passive
SMOS	Soil Moisture Ocean Salinity
SNR	Signal to Noise Ratio
TCA	Triple Collocation Analysis
ubRMSD	Un-biased Root-Mean-Squared-Deviation
ubRMSE	Un-biased Root-Mean-Squared-Error

1. Introduction

Surface soil moisture (SM) plays a vital role in the water (Jung et al., 2010), carbon (Trugman et al., 2018), and energy cycles (Seneviratne et al., 2010) and is classified as an Essential Climate Variable (GCOS, n.d.). Europe is a densely populated region of the world, with strong dependence on soil water, and droughts are predicted to increase across the continent due to climate change (Grillakis, 2019). Thus, monitoring of surface soil moisture is considered vital, particularly within the context of soil water stress (Bogena et al., 2022).

Currently, three main approaches, namely in situ measurements, satellite remote sensing, and atmospheric reanalysis models are widely used for estimating SM. In situ measurements are usually considered as the most accurate (Babaeian et al., 2019; Gruber & Peng, 2022). Some stations and station networks have been harmonized into databases, e.g., the global scale International Soil Moisture Network (ISMN) (Dorigo et al., 2011), but installing stations is expensive and many parts of the world lack a comprehensive coverage (Babaeian et al., 2019). Therefore, using only in situ measurements cannot well monitor the spatial and temporal variability of SM at a regional or global scale.

Satellite remote sensing can estimate SM in large spatial and long temporal scales. This approach is based on measuring the surface's emission and backscattering of electromagnetic (EM) signals, particularly in the microwave domain (Dorigo et al., 2017). The retrieval by the microwave signals can be conducted via passive (radiometers) or active (real-aperture radar, synthetic-aperture radar) systems (Gruber & Peng, 2022). Sensors on satellite missions such as Sentinel-1, Soil Moisture and Ocean Salinity (SMOS), and Soil Moisture Active Passive (SMAP) allow for the generation of several regularly released gridded SM products (Babaeian et al., 2019). In addition, various data, including environmental and meteorological variables, can be merged for generating gridded reanalysis products, e.g., ERA5-Land (Albergel et al., 2018).

As gridded SM products have been found to vary in quality depending on location (Al-Yaari et al., 2019; Cammalleri et al., 2017), the products need to be evaluated according to the user's desired utilization. Evaluation also benefits the products themselves by providing guidance toward their improvement (Dorigo et al., 2021). Despite SM's high spatial variability, in situ measurements are commonly used for evaluating the coarse gridded SM products (Louvet et al., 2015; Zheng et al., 2022) through sparse networks (Chen et al., 2018). As denser observation networks lead to more reliable evaluation, efforts have been made to upscale the sparse stations (Crow et al., 2012).

Evaluation of SM products using in situ measurements have been conducted over various areas and scales, from regional to global (Louvet et al., 2015; Ma et al., 2023; Stillman & Zeng, 2018; Zeng et al., 2015; Zheng et al., 2022). The in situ measurements utilized in this type of study have often been from the ISMN (Al-Yaari et al., 2019), or from designated core validation sites for the specific SM product to be evaluated (Colliander et al., 2017). The Integrated Carbon Observation System (ICOS) manages a sparse network of ecosystem stations across Europe that, among other environmental variables, measure SM. The network has the advantage that all measurements are relatively standardised and the ICOS infrastructure

employs rigorous measurement quality controls (ICOS, n.d.), while being spread over a large part of Europe. Given that much ICOS data were made available in just recent years, ICOS data have not previously been used for bias correction in the generation of gridded SM products, meaning that the SM products and ICOS measurements are independent from existing products and allow for performing a fair evaluation. To the author's knowledge, there are no previously published studies that have evaluated SM products primarily using ICOS data.

Evaluation of gridded products with situ observations cannot be performed for regions where no measurements exist. An alternative method is using Triple Collocation Analysis (TCA) which evaluates the error structure of three data sets simultaneously without requiring additional validation data (Scipal et al., 2008). Previous studies employing TCA have resulted in performance evaluation and ranking of SM products, e.g., over China (Wu et al., 2020). There have also been efforts to merge multiple pre-existing gridded products based on their TC estimated errors (Gruber et al., 2017).

Studies have previously been made on comparing several SM products over Europe, using either evaluation with ISMN stations (Al-Yaari et al., 2019), just TCA (Cammalleri et al., 2017; Fascetti, Pierdicca, Pulvirenti, & Crapolicchio, 2014; Pierdicca et al., 2015), or both in situ-based evaluation and TCA (Brocca et al., 2011; Xu et al., 2021). These studies have either been limited to Europe (Brocca et al., 2011; Fascetti, Pierdicca, Pulvirenti, Crapolicchio, et al., 2014; Pierdicca et al., 2015), or included the area as part of a multi-region or global assessment (Al-Yaari et al., 2019; Cammalleri et al., 2017; Xu et al., 2021). However, in the global studies, results are generally not reported with high detail for Europe. Further, developments are regularly made to products, and to the author's knowledge, no published studies have evaluated these products with a particular focus on Europe in recent years. In addition, previous studies have generally employed in situ data from the ISMN, which lacks sites in certain areas of Europe, that however do have coverage by ICOS sites, e.g., in Sweden. Moreover, to the author's knowledge, TCA based merging of soil moisture has not previously been attempted over large European areas.

The aim of this study is to evaluate and compare the performance of multiple current and widely used gridded soil moisture products, that employ different technologies and data, to identify the best-performing products in Europe. This study further aims to investigate if merging of multiple products can generate a new improved product. To achieve the aims of this study, the evaluation is conducted using ICOS measurements, and Triple Collocation Analysis (TCA) This study will answer the following three research questions:

1) Which gridded SM products perform the best based on evaluation with new in situ measurements from ICOS?
2) Among a selection of independent products, which SM products exhibit the best performance based on TCA evaluation?
3) Can existing SM products be merged to produce an improved SM product using TCA?

2. Background

2.1. General introduction to soil moisture

Soil moisture is commonly defined as the water content of the unsaturated (vadose) zone of the soil (Seneviratne et al., 2010). The SM unit varies is usually either gravimetric, volumetric, i.e., m^3_{water}/m^3_{soil}, or relative, based on the saturation of the soil in %. Conversion between saturation and volumetric data requires information on the porosity of the soil. This study focuses on surface SM, normally roughly referring to the top 5 cm, as opposed to root zone SM (ca 0-100 cm).

At the surface, the soil connects the terrestrial and atmospheric components of the hydrological cycle (Daly & Porporato, 2005). Water mainly enters the soil though precipitation and leaves through evapotranspiration and runoff, with SM affecting the processes that occur the processes in between (Ochsner et al., 2013). Water is essential to most plants and SM has been shown to be a controller of local environmental niches (Silvertown, 2004). Additionally, SM has been classified as a controller of net primary productivity which can be limited by drought induced soil water stress (Moyano et al., 2013; Trugman et al., 2018). Consequently, SM also strongly influences agriculture and food production, as well as soil's role as a global carbon pool/sink. Anthropogenic climate change is predicted to increase the temporal variability in soil moisture (Taylor et al., 2013), for example via droughts. This will have a strong impact on the soil processes (Grillakis, 2019). Although, it has been indicated that there is large uncertainty in model prediction on the effect of soil moisture changes on soil carbon storage (Falloon et al., 2011). Besides its importance for the natural environment and agriculture, SM also affects infrastructure development (Nguyen et al., 2023).

2.2. Remote sensing of soil moisture

2.2.1. General principles of remote sensing of soil moisture

Remote sensing, i.e., retrieval of information on objects from a distance, is within the field of Earth Observation mainly based on the different responses of different surface features to electromagnetic energy as it backscatters or is emitted from the Earth surface objects (Gruber & Peng, 2022). Commonly, the information is collected by sensors on sun-synchronous polar orbiting satellites (Entekhabi et al., 2010; Kerr et al., 2010). SM was first estimated using satellite data in the mid-1980's, with the first dedicated missions being launched in the late 00's (Gruber & Peng, 2022). The advancements have been numerous since then, in issues ranging from accuracy to resolution and spatial coverage (Liu & Yang, 2022). Besides moisture, other soil properties such as organic carbon, iron, and carbonate can also be retrieved using remote sensing techniques (Mulder et al., 2011).

Remote sensing of soil moisture is predominantly based on the soil's response to electromagnetic energy in the microwave range due to the dielectric properties of the soil materials in this range (Gruber & Peng, 2022; Louvet et al., 2015). Microwave frequencies range from 0.3 GHz to 300 GHz, and SM specific sensors normally employ the L- (1-2 GHz) or C-(4-8 GHz) bands (Gruber & Peng, 2022; Ochsner et al., 2013). In addition to efficiently

capturing the soil's dielectric properties, microwave frequencies are mostly unaffected by atmospheric conditions, and the signals can reach depths of 5 cm into the soil. However, microwave signals have the drawback of low spatial resolution, it generally being in the range of tens of kilometres (Gruber & Peng, 2022; Liu & Yang, 2022). Higher frequency results in higher spatial resolution but is also more affected by disturbance, such as vegetation, while lower frequencies can penetrate deeper into the soil, but result in coarser spatial resolution (Gruber & Peng, 2022; Ochsner et al., 2013). There have been attempts to employ significantly higher frequencies in order to increase the spatial resolution, such as visible-near infrared and shortwave-infrared waves (Lobell & Asner, 2002), but these are currently not widespread in use. There are two main methods of signal retrieval: passive and active. The are further detailed in the subsequent sections.

2.2.2. Passive microwave remote sensing of soil moisture

Passive remote sensing of soil moisture, which is based on the soil's naturally emitted signals, generally utilizes radiometers active in the L-band, commonly at ca 1.4 GHz (Wigneron et al., 1998). The large-scale possibility to use these microwave wavelengths for capturing the dielectric soil properties was presented by Jackson (1993). The passive models are generally based on retrieving the dielectric permittivity of the soil through brightness temperature values of the radiometer, but the specifics vary by model, which generally include several parameters, including considerations for roughness (Entekhabi et al., 2010; Gruber & Peng, 2022; Mironov et al., 2013; Wigneron et al., 1998). Compared to radars, radiometers have the advantage of being relatively cheap, having high temporal resolution, and utilizing the high SM-sensitivity at the low frequencies (Edokossi et al., 2020)

The main passive soil moisture retrieval missions of today include Soil Moisture Active Passive (SMAP) (Entekhabi et al., 2010) and Soil Moisture Ocean Salinity (SMOS) (Kerr et al., 2010). These missions are described further in section 3.

2.2.3. Active microwave remote sensing of soil moisture

Active remote sensing is based on artificially emitted EM pulses, with real-aperture- (Gruber & Peng, 2022) or synthetic-aperture radars (SAR) (Ulaby et al., 1996). As opposed to the L-bands of the radiometers, radar sensors usually utilize the C-band, for example the Advanced Scatterometer (ASCAT), which operates at 5.255 GHz. Instead of brightness temperature retrieval, radars retrieve the backscatter coefficient, i.e., change in magnitude of the emitted signal at return to sensor (Gruber & Peng, 2022).

Compared to passive sensors, active SM retrieval has the advantage of having higher spatial resolution (~100 m), but it is less sensitive to SM and require relatively complex data processing (Edokossi et al., 2020). SAR data is known to have limitations, still needing improvements regarding incidence angle, surface roughness, and vegetation, especially for further downsampling applications (Gruber & Peng, 2022). Active missions included in this study are Sentinel-1 and ASCAT, which are further detailed in section 3.

2.3. Reanalysis products

Reanalysis products provide extensive temporal and spatial coverage of numerous environmental variables and are mainly based on atmospheric model components coupled with ground-based measurements, and can further assimilate satellite information, including microwave data (Sabater et al., 2008; Sawada et al., 2015). That the models use various data sources allows for high temporal resolution, e.g., hourly as in ERA5-Land, with practically gap-filled spatial coverage. Reanalysis products have potential of being as reliable as in situ observations (Tarek et al., 2020). This study includes ERA5-Land (Hersbach et al., 2020) and GLDAS Noah (Rodell et al., 2004). Another common reanalysis product not included in this study is MERRA from NASA (Gelaro et al., 2017).

Note that in this study, the reanalysis products are treated the same as the satellite based gridded products and are evaluated against the ICOS measurements. However, due to their commonly high accuracy and similar spatial support to the satellite-based products, they can be used for comparisons with gridded SM estimates, as long as they are independently generated from the SM product that is being evaluated (Al-Yaari et al., 2014).

2.4. Evaluation of SM products

2.4.1. Background to in situ measurements and related evaluation

Traditionally, SM product validation is done by comparing the product estimates to various types of in situ measurements that are assumed to represent the true soil moisture. There are a few different methods of directly measuring SM. Besides gravimetric measurements, which involve collection of samples, drying, and weighting, which is very labour intensive and mainly done for calibration purposes (Dorigo et al., 2011), it is common to use dielectric probes placed a few centimetres into the soil (Mortl et al., 2011; Seyfried et al., 2005). The instruments generally provide accurate measurements, even if they have the risk of degrading over time or can be affected by soil disturbances (Dorigo et al., 2011).

The networks that the measurement stations belong to are either dense or sparse. Dense networks have several stations per gridded product cell, which allows for spatial interpolation of SM values. Sparse networks generally only have one point measurement in a gridded SM product cell (Gruber et al., 2020). Many of the networks included in the ISMN, as well as ICOS, are sparse. The main problem associated with using sparse networks in evaluation studies is the unknown representativeness of the point-based measurement for the coarser area (in this study up to ~25 km by ~25 km) represented by the gridded product cell, especially considering the high spatial variability of SM (Crow et al., 2012). While the representativeness might particularly affect the differences in magnitude of SM, it has also been shown that temporal dynamics at point-based stations remain relatively similar to that at the larger surrounding area, even at coarse scales (Brocca et al., 2011). This advises which evaluation metrics were employed (see section 4). The relatively homogeneity in temporal dynamics means that at the least the Pearson correlation coefficient R can be justified (Brocca et al., 2009; Ma et al., 2019; Vachaud et al., 1985), and sparse networks are commonly used for evaluation (e.g., Dorigo et al., 2015; Högström et al., 2018; Hu et al., 2022).

Nevertheless, the representativeness issue causes a large degree of uncertainty to the in situ-measurement-based evaluation results. Further attention is given to it throughout the study, in particular via discussions on land cover heterogeneity which is known to be related to SM variability and product error (Panciera, 2009). Due to the unknown uncertainty of large-scale representativeness of ICOS, the results from the in situ observation-based evaluation are particularly interesting to compare with the results from TCA, which is independent of the SM measurement issues (see section 2.4.2.).

It should be noted that there are, in addition to the in-soil instrument point measurements, cosmic-ray neutron measurements which are non-invasive, i.e., not placed in the soil itself (Zreda et al., 2008). They provide measurements on a scale of hundreds of metres, which generally means that they avoid some of the representativeness issues of point-based measurements (Zreda et al., 2012). One network of such measurements is COSMOS-Europe (Bogena et al., 2022), but it was not included in this study.

The methods employed for the in situ-measurement-based evaluation are further explained in section 4.

2.4.2. Triple collocation-based evaluation

Triple collocation analysis (TCA) was first used for wind modelling, introduced by Stoffelen (1998). It was later brought into soil moisture research through Scipal et al. (2008). Since then, it has become a staple method for gridded SM product evaluation (F. Chen et al., 2018; Gruber, Su, Zwieback, et al., 2016). TCA has also been used for other environmental variables, including evapotranspiration (Li et al., 2023), and precipitation (Alemohammad et al., 2015; Dong et al., 2020).

TCA utilizes three temporally and spatially collocated independent sets of data to perform error estimation. Within SM product TCA error estimation, it is common practice that the product triplet consists of one passive, one active, and one independent product, such as a land surface model, or in situ measurements (Gruber, Su, Zwieback, et al., 2016).

The common notation of the TCA error model is in accordance with Eq. 1 below (Gruber, Su, Zwieback, et al., 2016).

$$i = \alpha_i + \beta_i \Theta + \varepsilon_i \qquad (1)$$

where i represents the dataset value, Θ is the true SM, α_i is the additive bias, β_i is the multiplicative bias, and ε_i represents the random error, which has a mean of 0. From three products assumed to have this error structure, the aim of triple collocation is to estimate the individual parts, in particular the random error (Stoffelen, 1998).

For triple collocation to be valid, the following requirements/assumptions need to be true (F. Chen et al., 2018; Gruber, Su, Zwieback, et al., 2016; Peng et al., 2021):

i. All datasets have a linear relationship to the true SM
ii. Errors and true SM are not correlated
iii. Errors of all three datasets are independent of each other

iv. There is stationarity in the error statistics and signal

For full error retrieval and TCA, see formulas in section 4.2.

The data sets in the triplet can be merged using a weighted mean, with the weights determined by the TCA estimated errors. It has been shown that TCA merged gridded products can outperform the parent data (Peng et al., 2021). The method has been employed by the European Space Agency (ESA), for the ESA CCI SM product (Dorigo et al., 2017) which can be considered a benchmark TCA based merging product.

Further, there are additional techniques which are based on similar assumptions as triple collocation, which have also become common for cross-error estimation of products. These approaches include Extended Collocation (Gruber, Su, Crow, et al., 2016), Quadruple Collocation (Min et al., 2022), and Extended Double Instrumental Variable (Kim et al., 2021).

3. Data

3.1. Gridded soil moisture products

3.1.1. Sentinel-1

The Sentinel-1 SM product is part of ESA's Copernicus program. The product is based on C-band SAR data at 5.405 Hz from the two Sentinel-1 satellites Sentinel-1A and Sentinel-1B, that were launched in April 2014 and April 2016, respectively (Bauer-Marschallinger et al., 2019). Thus, data for April 2014 and onwards is available. The SM retrieval from the raw SAR data is based on the change detection algorithm developed by TU Wien (Wagner et al., 1999). After calibration, it is downsampled to 1 km resolution for all of Europe. The revisit frequency is approximately 1.5 – 4 days, with SM values representing days, and the product represents SM in the top 5 cm of the soil (Bauer-Marschallinger et al., 2019).

The provided SM unit is degree of saturation (%). Thus, the data was converted to volumetric soil water content (m^3/m^3) using the global 0.25° GLDAS Land Surface Model porosity dataset, derived from Reynolds et al. (2000) (Rodell et al., 2004). The porosity dataset was retrieved at https://ldas.gsfc.nasa.gov/gldas/soils (LDAS, n.d.). Sentinel-1's recommended quality masking was applied. This included removing for flags of extreme slopes, dry extremes, wet extremes, water surfaces, and low sensitivity. The Sentinel-1 SM product was retrieved from the ESA Copernicus Global Land Service portal (https://land.copernicus.eu/global/products/ssm).

3.1.2. SMAP L3E and SMAP L4

The Soil Moisture Active Passive (SMAP) satellite was launched by NASA in January 2015. The satellite carries both an L-band radar sensor, at 1.26 GHz, and an L-band radiometer, at 1.41 GHz (Entekhabi et al., 2010). However, only the radiometer was functioning during the study period, meaning that SMAP was classified as a passive product in this study. From the radiometer brightness temperatures, several products are generated (Colliander et al., 2017), two of which were selected for this study.

The SMAP Level 3 Enhanced product (SMAP L3E) (version 5) is a daily composite of the SMAP Level-2 SM (O'Neill et al., 2021). From the approximate 36 km radiometer resolution, the product employs the Backus-Gilbert optimal interpolation algorithm to resample it to a 9 km grid. Temporally, SMAP L3E has SM estimates at AM and PM. These were averaged to get the daily SM value. General recommendations were followed for the quality masking. Snow, permanent ice, frozen ground, and static water pixels were all filtered out. However, several ICOS stations were located on pixels that were flagged for other reasons, in particular vegetation density, and to keep the validation data, pixels with these flags were not masked out.

SMAP L4 (version 7), assimilates the SMAP sensor data in a geophysical land surface model (Reichle et al., 2022a; Reichle et al., 2022b; Reichle et al., 2022c). The assimilation model is based on the GEOS-5 land data assimilation system (LDAS) and includes precipitation observations and GEOS-5 surface meteorology in addition to the SMAP brightness temperatures (Reichle et al., 2017). The quality flagging and filtering is done within the assimilation, and pixels with unfavourable flags for snow cover, frozen ground, vegetation density etc. are filtered out. The model gives output 3-hourly at the 9 km grid. The 8 daily values were averaged for one daily mean value that was used in the evaluation.

Both SMAP L3E and SMAP L4 represents the SM in the top 5 cm of the soil and have coverage from March 2015 and onwards. The data were retrieved from the National Snow and Ice Data Center (NSIDC) (O'Neill et al., 2021) (https://nsidc.org/data/spl3smp_e/versions/5, https://nsidc.org/data/spl4smgp/versions/7).

3.1.3. SMOS L4

ESA launched the Soil Moisture Ocean Salinity (SMOS) mission in 2009 to monitor SM over land and salinity at the ocean surface (Kerr et al., 2010; Pablos et al., 2022). The satellite employs a radiometer in the L-band at 1.4 GHz. The level 4 product, used in this study, was generated via algorithms from the Barcelona Expert Center that downscaled the raw data, at a 10s of km scale, to a fine resolution of 1 km. Besides L-band brightness temperature, the downscaled SM is obtained from land surface temperature (LST) data from the European Centre For Medium-Range Weather Forecasts (ECMWF) model and Normalized Difference Vegetation Index (NDVI) data from the Terra Moderate Resolution Imaging Spectroradiometer (MODIS) to perform the downsampling (Pablos et al., 2022; Piles et al., 2011).

SMOS L4 is released with daily values for ascending and descending overpass, these were averaged to get daily SM values. Only excellent quality data was used. Barcelona Expert Center (BEC) provided the product via their portal at https://www.catds.fr/Products/Available-products-from-CEC-SM/L4-Land-research-products.

3.1.4. ASCAT

ASCAT is an active product and is produced within the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) H SAF soil moisture project. Various versions of the ASCAT product exists, this study used version ASCAT SSM CDR v5 EXT 12.5 km, H116.

The mission is based on the Metop-satellites (A, B, and C) which contain the Advanced Scatterometer (ASCAT). ASCAT uses the C-band, at 5.255 GHz to retrieve the backscatter information. The TU-Wien developed algorithm from Wagner et al. (1999) is then used in order to generate the surface soil moisture from backscatter data. The H116 version is resampled to 12.5 km from original resolution of 25-34 km. The product contains descending and ascending values that are averaged for daily SM. The data is released as degree of saturation (%) in the top 5 cm. It was converted to VSWC (m^3/m^3) using the same data as for Sentinel-1 (see section 3.1.1). Quality flag wise, the confidence flag "0" was used, i.e., indicating no data issues. The ASCAT data was retrieved from the EUMETSAT portal (https://navigator.eumetsat.int/product/EO:EUM:DAT:0308).

3.1.5. ESA CCI SM

The ESA Climate Change Initiative (CCI) SM product (Dorigo et al., 2017; Gruber et al., 2019) project began in 2010, and exists within the broader aim to collect of long-term climate data records. A new version is released yearly. This study used the daily images combined version, v.07.1, which has released SM estimates covering 1978 to 2021. The product includes both passive and active SM retrievals, using data from, among other sensors; ASCAT, SMOS, and SMAP. However, inclusion of sensors vary over time depending on their availability. 12 passive and 5 active sensors are used at different points in time. Data from the sensors are merged, partly using triple collocation analysis (EODC, 2022). Further, ESA CCI SM also uses the reanalysis GLDAS-Noah, both as a scaling reference and for uncertainty estimation (Gruber et al., 2019). The product is resampled to a 0.25° grid and has daily values. The unit is VSMC (m^3/m^3). Only recommended quality flags were kept, meaning that estimates flagged for snow, frozen soil, or dense vegetation were filtered out. While different sensors have different soil sensing depths, the product has been considered to represent the top 5 cm of the soil in previous research (Wu et al., 2020).

The data was retrieved from the ESA SM CCI website (https://www.esa-soilmoisture-cci.org/data).

3.1.6. ERA5-Land

The European Centre for Medium-Range Weather Forecasts (ECMWF) produces the global atmospheric product ERA5 (Hersbach et al., 2020; Muñoz-Sabater et al., 2021) within the Copernicus Climate Change Service (C3S). ERA5 merges meteorological models with various measurements. In this study, ERA5-Land (Muñoz Sabater, 2019) was used. ERA5-Land is produced from a replay of the land component of ERA5 and is resampled to an enhanced horizontal resolution of 0.1°. The temporal coverage is from 1950 to present. The hourly data version was used in this study, with values averaged to daily values. The shallowest available depth was used, at 0-7 cm.

The data was accessed at the Climate Data Store and Copernicus Climate Change Service (C3S, 2022)(https://cds.climate.copernicus.eu/cdsapp#!/dataset/10.24381/cds.e2161bac?tab=overview).

3.1.7. GLDAS Noah

The second reanalysis dataset included in this study was GLDAS-Noah v.2.1 (Beaudoing & Rodell, 2020). It is part of the Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004), and uses Noah Model 3.6 to produce a global estimate of environmental variables at 0.25° spatial resolution with a temporal resolution of 3 hours. The soil moisture between 0-10 cm was selected and the 3-hourly data were averaged to create daily values. It contains data from 1978 to present. The product was retrieved at the NASA Earth data portal (https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH025_3H_2.1/summary).

Table 1. The Gridded Soil Moisture Products

Gridded product	Grid resolution	Temporal coverage	Coverage	Sampling depth (cm)	Frequency (GHz)
Sentinel-1	1 km	April 2014 – present	Europe	0-5	5.405 (C)
SMAP L3E	9 km	March 2015 – present	Global	0-5	1.41 (L)
SMAP L4	9 km	March 2015 – present	Global	0-5	1.41 (L)
SMOS L4	1 km	June-2010 – present	Europe	0-5	1.4 (L)
ASCAT	12.5 km	January 2019 – present	Global	0-5	5.255 (C)
ESA CCI SM	0.25°	January 1978 - December 2021	Global	0-5	Various
ERA5-Land	0.1°	January 1950 – present	Global	0-7	-
GLDAS Noah	0.25°	January 2000 - present	Global	0-10	-

3.2. In situ measurements from ICOS

The Integrated Carbon Observation System (ICOS) is a European collaboration and infrastructure to monitor the carbon cycle. ICOS hosts atmosphere stations, ocean stations, and ecosystem stations. This study uses data from the ecosystem stations, of which there are currently 87. However, many of these stations do not have coverage during 2020 and 2021. Instead, only 38 ICOS stations were found to provide SM data for most of the study period (see Table A1 in appendices for detailed station information). Geographically, the stations are located from northern Finland and Sweden, through central and western Europe, to southern France and northern Italy (Figure 1). ICOS coverage is lacking in the Iberian Peninsula, Iceland, Great Britain, and Eastern Europe.

ICOS utilizes dielectric sensors to measure the SM and the top SM measurement, which is what was used in this study, is instructed to be fixed at 5 cm (De Beeck et al., 2018). The data is collected half-hourly by ICOS and were averaged to provide a daily value which could be compared with the daily gridded SM estimates. All ICOS data was retrieved from the ICOS Data Portal (https://data.icos-cp.eu/portal/).

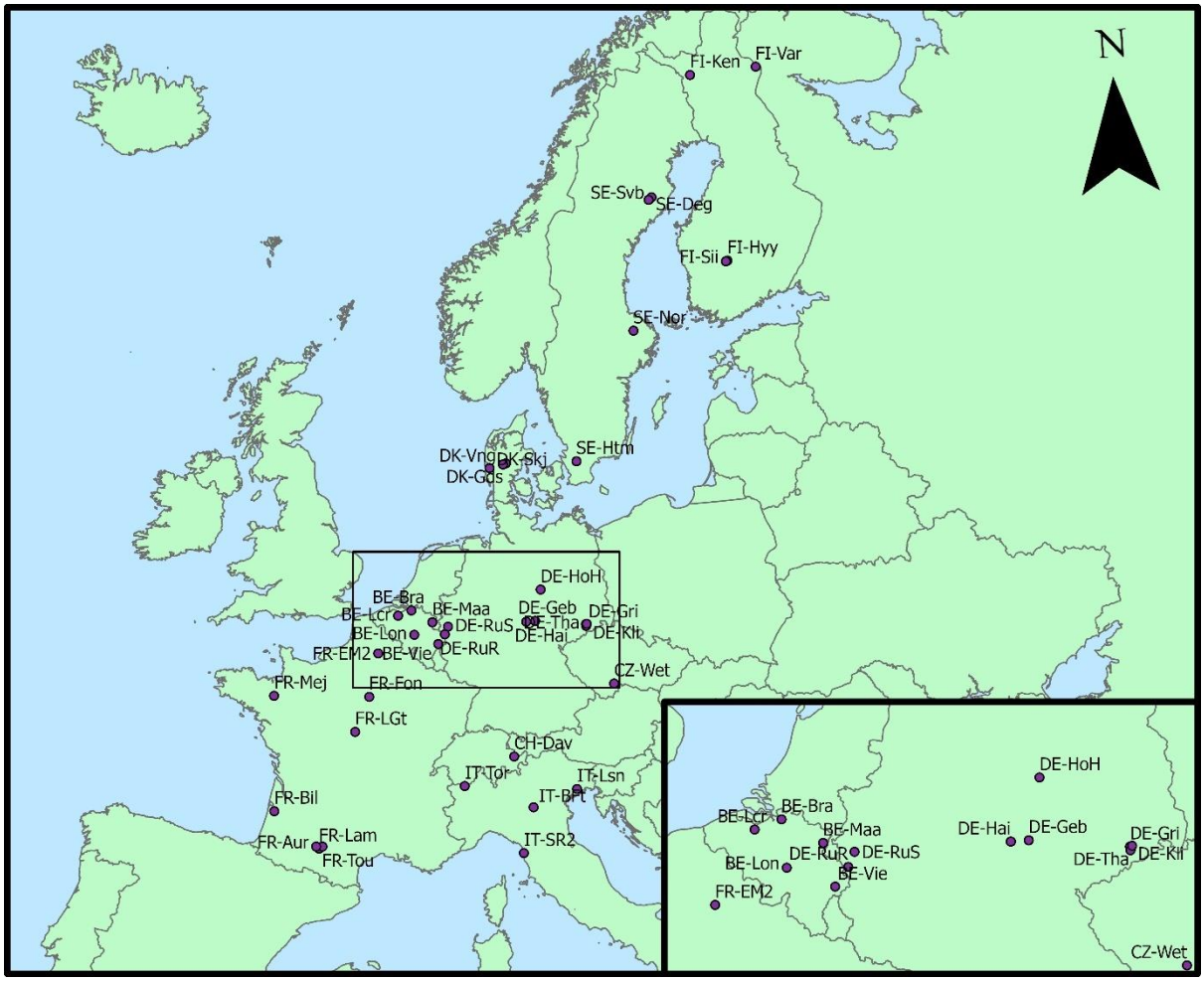

Figure 1. Map of the 38 ICOS stations. Inset for the stations in Belgium and central Germany. Made with Natural Earth.

The representativeness of the land cover that the ICOS stations were located on within the surrounding area was investigated using 1 and 9 km horizontal squares, the pixel sizes of the SMOS and SMAP products respectively, centred on each station. The 2018 100 m CORINE Land Cover (CLC) dataset v.2020_20u1, was used (Copernicus, 2020). The CLC data indicated that SE-Svb was surrounded by 100% coniferous forest in the 1 km square, and 85% coniferous forest in the 9 km square, while SE-Deg was surrounded by 73% peat bog in the 1 km square and only 5% in the 9 km square, see Figure 2 below and Table A1 in the appendices.

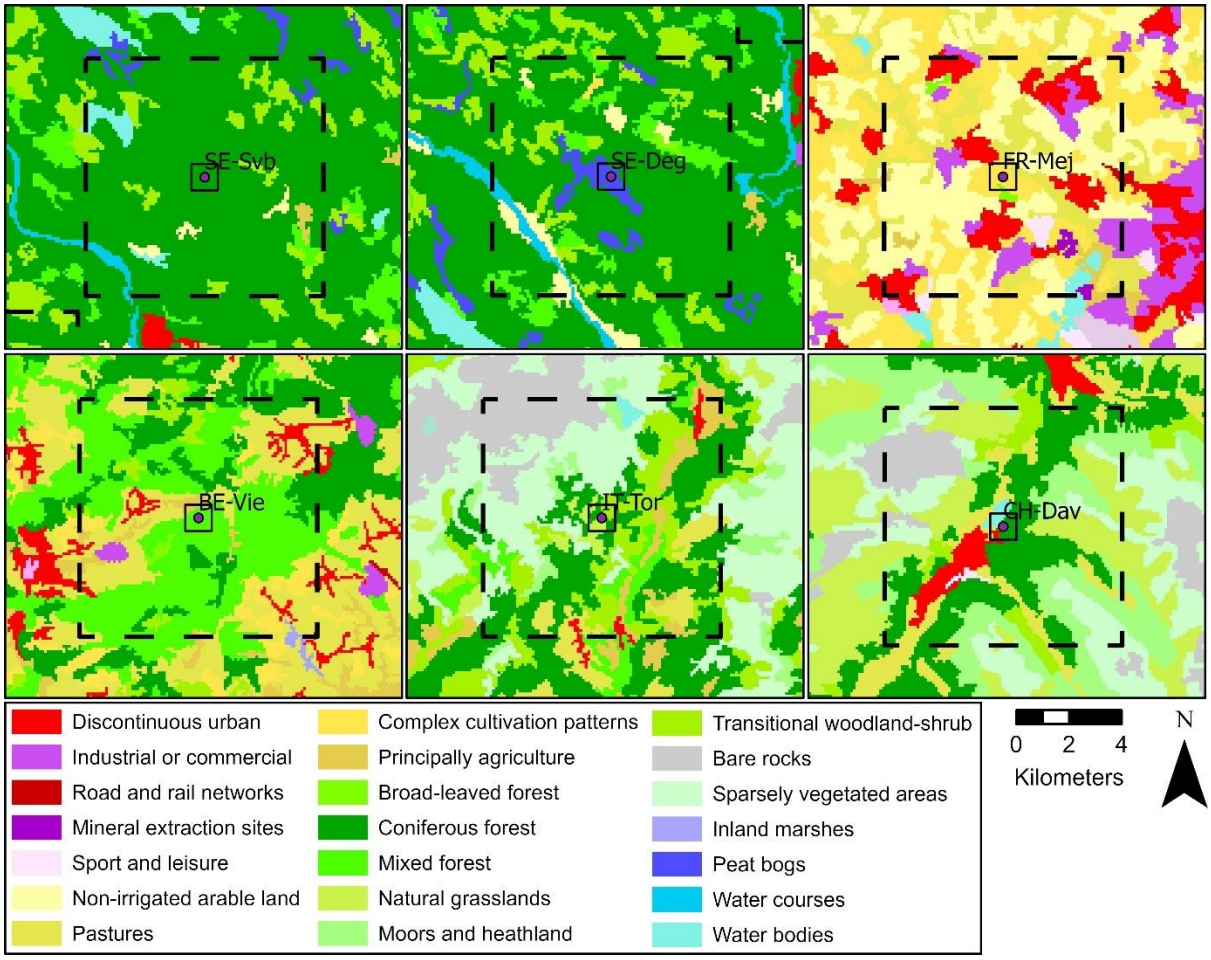

Figure 2. Corine Land Cover around 6 selected ICOS stations with 1 km horizontal square (solid outline) and 9 km horizontal square (dashed outline).

4. Methodology

4.1. In situ measurement-based evaluation

4.1.1. General workflow and pre-processing

The in situ measurement-based evaluation mainly follows the guidelines and framework set by (Gruber et al., 2020), including data acquisition, spatial and temporal collocation of measurements and gridded data, masking of flagged retrievals, decomposition, i.e., retrieval of anomalies, rescaling, and metric calculation.

The general workflow, from data collection to metric calculation is displayed in Figure 3.

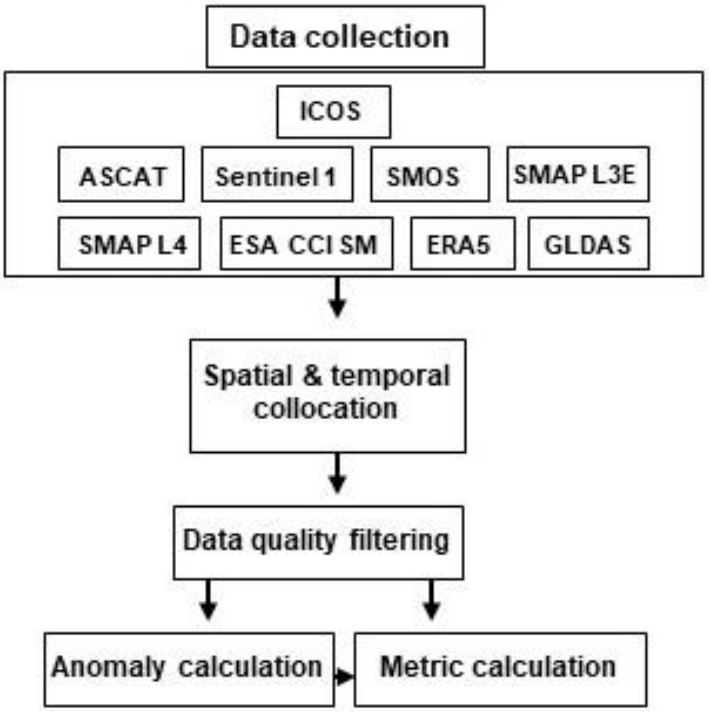

Figure 3. General workflow of the in situ-measurement based evaluation.

Data management were mainly handled in Python using standard packages, while the open-access program the Python Toolbox for the Evaluation of Soil Moisture Observations, Pytesmo (https://doi.org/10.5281/zenodo.596422), was employed for much of the analysis.

Gridded product data and ICOS data were spatially matched via a nearest neighbour operation, i.e., by using the SM value of the gridded product-cell that the ICOS-station was located inside. The selected grid cell SM could then be compared to the in situ SM, following standard practice (Dorigo et al., 2015; Gruber et al., 2020). As the stations were sparsely located, there was no need to interpolate SM from different in situ measurements. In Germany and Finland, some pixels contained several (2 or 3) ICOS stations at the coarser grids. One method to deal with this issue is to only use the station that is assumed to be the most representative over the grid cell, this being the station with the highest mean correlation at the pixel for several products (Dorigo et al., 2017; Xu et al., 2021). However, as the products, besides SMAP L4 and SMAP L3E, did not use the same grids or grid sizes, it was decided to select and keep the station which had the highest correlation coefficient with the gridded data, for each product individually. The only exception was between SMAP L3E and SMAP L4, where the station with the highest mean correlation between both products was kept.

Each instance of a removed station is shown in Table 2, but no more than two stations were removed for a single product, meaning that the removal method likely did not have a major impact on the results.

Table 2. The removed stations for each product in the in situ measurement-based evaluation.

Gridded product	Removed stations
ASCAT	DE-Tha, FI-Sii
ERA5-Land	None
ESA CCI SM	DE-Tha, DE-Kli
GLDAS Noah	DE-Tha, DE-Kli
Sentinel 1	None
SMAP L3E	DE-Kli
SMAP L4	DE-Kli
SMOS	None

To give a most fair representation of each products' individual performance as released, they were not resampled to the same grid. Thus, the impact of resampling artifacts on the result was avoided.

Temporal collocation of measurements was done to facilitate comparisons only done for dates that all products had data for at a station. Data was converted to being daily for all products where necessary, and then matched between ICOS and products.

The products' respective quality flags were used to filter out unreliable data (see section 3 for quality filtering employed for each product). However, as all products had different quality flags it is not certain that the same thresholds were used. This is a well-known issue when comparing different product and needs to be considered (Gruber et al., 2020). However, as the bulk of inter-product comparisons was done for collocated dates, i.e., dates with values for all products after quality flag filtering at each of them, it can be assumed that most of the data with significant uncertainty was consistently filtered out.

To further ensure avoidance of frozen soil, for which satellite retrieval is not possible (Babaeian et al., 2019), the dates were further filtered by ICOS soil temperature measurements, removing dates with below 0 °C soil temperature.

4.1.2. Anomalies and evaluation metrics

Short-term SM-anomalies represent the high frequency dry and wet events, i.e., brief fluctuations in the SM, and were calculated to enable evaluation of gridded SM product's skill at capturing these variations (Gruber et al., 2020). Anomalies were calculated using the same method as Peng et al. (2021), using a 35-day moving mean. The anomaly A on day i for product P's measurement over the 35-day mean $\bar{P}$ is presented in Eq. 2 (Peng et al., 2021).

$$A_i = P_i - \bar{P} \quad (2)$$

Gruber et al. (2020) suggested a threshold requiring at least 25% valid data within the window for the anomalies to be calculated (i.e., 8.75 days in a 35-day window) but as there was a lot of missing gridded product data for many periods, it was decided to use a 7-day threshold instead. A mean consisting of only 7 values is not necessarily representative of the conditions during the 35-day period but was a necessary compromise. Also note that when calculating anomalies,

only data from collocated dates between ICOS and the products were included to ensure that gaps in temporal coverage for the gridded product SM did not affect the scores.

The main metrics that are used in in situ measurement-based evaluation, and that are suggested by Gruber et al. (2020), are the linear Pearson correlation R (eq. 5), Root-Mean-Squared-Deviation (RMSD), and bias. It is also recommended to use the scaled versions of the metrics when comparing absolute soil moisture data, i.e., with removal of inherent product bias. Thus, for RMSD, the unbiased RMSD (ubRMSD) was used instead (eq. 6). Further, following recommendations, bias was not presented in this study as the footprints of compared measurements were different (Gruber et al., 2020).

Statistical metrics were only calculated when there were more than 30 dates available, resulting in exclusion of some stations for specific products. The equations for the statistical metrics as well as variance and covariance follow below. Overbars represent the temporal mean, n represents the total number of measurements, i represents a day or a single measurement:

Variance σ_X^2, which represents dispersion of data from product X, was calculated according to Eq. 3.

$$\sigma_X^2 = \frac{\sum(X_i - \bar{X})^2}{n-1} \quad (3)$$

Covariance σ_{XY}, which represents direction and size of relationship between X and Y, was calculated from Eq. 4.

$$\sigma_{XY} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{n-1} \quad (4)$$

Pearson's correlation coefficient R between data sets X and Y, was calculated from Eq. 5 (Gruber et al., 2020).

$$R_{XY} = \frac{\sigma_{XY}}{\sigma_X \, \sigma_Y} \quad (5)$$

The Un-biased Root-Mean-Squared-Deviation $ubRMSD$, was derived from Eq. 6 (Gruber et al. 2020).

$$ubRMSD = \sqrt{\sigma_X^2 + \sigma_Y^2 - 2\sigma_{XY}} \quad (6)$$

The R simply shows the correlation between the two datasets being compared. A perfect product would have a linear correlation with a dataset assumed to have a perfect match with the ground reality, which would result in an R of 1. ubRMSD shows how much the difference generally is between the product and ground reality, i.e., lower ubRMSD is closer to the reference data set and can be assumed to have lower error and lower uncertainty.

4.2. Triple collocation-based evaluation

4.2.1. Workflow and triplet selection

The triple collocation analysis workflow started with pre-processing, including data quality filtering, resampling the gridded products to the same grid, removing data when estimated soil temperatures were below 0 °C, and temporal collocation. Once this was done, the three data sets were acquired, and the TC metrics could be calculated (Figure 4).

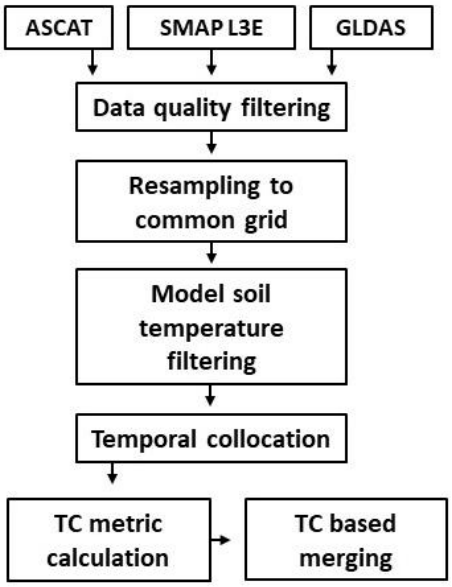

Figure 4. General workflow of the triple collocation (TC)-based evaluation.

The independent products selected to be used in the TC triplet were GLDAS Noah, ASCAT and SMAP L3E . ASCAT has been used in triplets together with SMAP previously, indicating that they should be independent (Al-Yaari et al., 2014; F. Chen et al., 2018; Peng et al., 2021).

It is also standard practice to select one active and one passive product together with a reanalysis product for the TC-based evaluation (Gruber et al., 2017). As mentioned previously, SMAP L3E only employed its radiometer during the study period, thus effectively being a passive product. These three products are as independent as gridded SM products can practically be.

4.2.2. Pre-processing

All quality flag filtering that was performed for each of the products in the in situ-measurement based evaluation was also conducted for the triple collocation. In addition, modelled soil temperature was used to filter out days with below 0 °C soil temperatures, i.e. This data came from ERA5-Land or GLDAS Noah respectively depending on resampling type, further detailed below. It is also common to filter out data with potential snow cover in the area (Gruber et al., 2019; Gruber, Su, Zwieback, et al., 2016). However, due to increased computer processing times that this brought, it had to be avoided. Note that many of the potential snow dates were

already excluded in the removal of below 0 °C soil temperatures, in addition to already being removed from the quality flag filters of the actual products.

For the TCA, it was necessary to resample the products to the same grid, as the products were evaluated on a pixel-by-pixel basis. There are several approaches for how to do this, the preferred method depending on factors such as the difference in grid cell size between different products, and the original satellite retrieval footprint size. The methods include area weighted averaging (Wu et al., 2020), which is an ideal method for upscaling products (Gruber et al. 2020), nearest neighbour interpolation (Al-Yaari et al., 2014) which is appropriate for similar sized grids (F. Chen et al., 2018), and bilinear interpolation (Hu et al., 2022). The selected approach was to interpolate the GLDAS Noah, SMAP L3E, and ASCAT products to the 0.1° ERA5-Land grid using the nearest neighbour interpolation to facilitate comparisons against ERA5-Land. 0.1° is on a similar scale to SMAP and ASCAT, thus, nearest neighbour retains a lot of information, largely avoiding artifacts (F. Chen et al., 2018). However, the grid resolution is significantly higher than the GLDAS Noah grid (0.25°), meaning that GLDAS Noah could be more strongly affected. Further, both ASCAT and SMAP's original satellite retrieval footprints were on a larger scale, at 25 km and 36 km scale respectively, meaning that the algorithmic downscaling methods of these products also affect the final product values.

To investigate if the difference in resolution between GLDAS Noah and the others, as well as the resampling method, make a significant difference in the TCA evaluation results, another approach of resampling was performed in addition to the nearest neighbour method. The second selected approach was to resample SMAP L3E and ASCAT to the 0.25° GLDAS Noah grid using inverse-distance-weighting (IDW). IDW has previously been used by F. Chen et al., (2018) that found it to be a good approach for ASCAT. With this method, the GLDAS Noah product remained unchanged, but there may have been resampling artifacts for SMAP and ASCAT respectively.

4.2.3. Evaluation metrics for the TCA

The products were compared on a pixel-by-pixel basis, for each of the three products. Using the assumed error structure of the data, as presented in section 2.4.2., the triple collocation could be workflow and formulas follow below.

From the series of SM values from each product, the signal to noise ratio (SNR) with decibel (dB) units was calculated. The SNR is directly related to the correlation to the assumed true SM, with higher SNR being preferable. Within TCA evaluation, EUMETSAT H SAF has specified three different levels of SNR based on their desirability: 0: threshold, 3: target, 6: optimal (Gruber et al., 2020).

For three products, indicated by M: model (GLDAS Noah), A: active (ASCAT), and P: passive (ASCAT), the estimated unbiased root mean squared error, the rescaled general difference from the true SM, also known as error standard deviation, ubRMSE $\sigma_{\varepsilon X}$ of each product were calculated according to Equations 7 to 9 (Peng et al., 2021).

$$\sigma_{\varepsilon M} = \sqrt{\sigma_M^2 - \frac{\sigma_{MA}\sigma_{MP}}{\sigma_{AP}}} \quad (7)$$

$$\sigma_{\varepsilon A} = \sqrt{\sigma_A^2 - \frac{\sigma_{AM}\sigma_{AP}}{\sigma_{MP}}} \quad (8)$$

$$\sigma_{\varepsilon P} = \sqrt{\sigma_P^2 - \frac{\sigma_{PM}\sigma_{PA}}{\sigma_{MA}}} \quad (9)$$

Further, the SNR_X for each product was calculated according to Eqs. 10 to 12 (Peng et al., 2021).

$$SNR_M = -10\log(\frac{\sigma_M^2\sigma_{AP}}{\sigma_{MA}\sigma_{MP}} - 1) \quad (10)$$

$$SNR_A = -10\log(\frac{\sigma_A^2\sigma_{MP}}{\sigma_{AM}\sigma_{AP}} - 1) \quad (11)$$

$$SNR_P = -10\log(\frac{\sigma_P^2\sigma_{MA}}{\sigma_{PM}\sigma_{PA}} - 1) \quad (12)$$

Note, that when presenting values for each of the three products in the results, only the pixels which also had ICOS stations were included instead of all processed pixels. This was done to avoid spatial bias when comparing the inter-product ranking in TCA to the inter-product ranking of the in situ observation-based evaluation.

4.2.4. Merging of the triplet products

Before the merging could be conducted, the bias of GLDAS Noah and ASCAT against SMAP L3E as baseline was calculated to facilitate rescaling of GLDAS Noah and ASCAT to SMAP L3E, with rescaling coefficients β_X (Eq. 13 to 15) (Gruber, Su, Zwieback, et al., 2016). SMAP L3E was selected as reference product as it was the better performing microwave-based product.

$$\beta_M = \frac{\sigma_{PA}}{\sigma_{MA}} \quad (13)$$

$$\beta_P = 1 \quad (14)$$

$$\beta_A = \frac{\sigma_{PM}}{\sigma_{AM}} \quad (15)$$

With the rescaling coefficients, GLDAS Noah and ASCAT could be adjusted to the SMAP L3E baseline as $M_{rescaled}$, and $A_{rescaled}$ respectively (Eqs. 16 and 17) (Peng et al., 2021).

$$M_{rescaled} = \beta_M * (M - \overline{M}) + \overline{P} \quad (16)$$

$$A_{rescaled} = \beta_A * (A - \bar{A}) + \bar{P} \quad (17)$$

The weight for each product, W_X, in the TCA merging, was then calculated according to Eqs. 18 to 20 (Peng et al., 2021).

$$W_M = \frac{\sigma_P^2 \sigma_A^2}{\sigma_M^2 \sigma_P^2 + \sigma_M^2 \sigma_A^2 + \sigma_P^2 \sigma_A^2} \quad (18)$$

$$W_P = \frac{\sigma_M^2 \sigma_A^2}{\sigma_P^2 \sigma_M^2 + \sigma_P^2 \sigma_A^2 + \sigma_M^2 \sigma_A^2} \quad (19)$$

$$W_A = \frac{\sigma_M^2 \sigma_P^2}{\sigma_A^2 \sigma_M^2 + \sigma_A^2 \sigma_P^2 + \sigma_M^2 \sigma_P^2} \quad (20)$$

The weighted merged SM estimate was then produced from Eq. 21.

$$SM_i = M_i * W_M + P_i * W_P + A_i * W_A \quad (21)$$

The simple arithmetic mean between the products was calculated according to Eq. 22.

$$Mean\ SM_i = \frac{M_i + P_i + A_i}{3} \quad (22)$$

Further, via a one-tailed student's t-test, the p-values of each duo of products were calculated, i.e., the duos being SMAP with GLDAS, SMAP with ASCAT, and GLDAS with ASCAT. The merging method chosen for the pixel was based on these p-values, according to the scheme proposed by Gruber et al. (2017), in particular the version used by (Peng et al., 2021), and it was done in order to increase the temporal coverage (Table 3). What merging method was used for each grid cell, as well as how many collocated dates there were at each grid cell, can be found in Figure A1, in the appendices.

Table 3. The TCA based merging scheme based on t-test p-values. Adapted from Peng et al. (2021) and Gruber et al. (2017). M represents the reanalysis GLDAS Noah, P represents the passive product SMAP L3E, and A represents the active product ASCAT. The number in parenthesis represents the merging method in Figure A1.

M - P $p < 0.05$	M - A $p < 0.05$	P - A $p < 0.05$	Merging method
Yes	Yes	Yes	TCA-weighted (0)
Yes	Yes	No	M (1)
No	Yes	Yes	P (2)
Yes	No	Yes	A (3)
Yes	No	No	Mean (M, P) (4)
No	Yes	No	Mean (M, A) (5)
No	No	Yes	Mean (P, A) (6)
No	No	No	No merging (7)

The TCA was only conducted where there were at least 100 common dates as recommended by previous studies (Dorigo et al., 2010; Scipal et al., 2008). In addition, any pixels that had data which statistically violated the TC assumptions were discarded.

5. Results

5.1. Evaluation with in situ measurements

5.1.1. Results without considering temporal collocation

Figure 5 displays Pearson's R and ubRMSD for the absolute SM of each product against ICOS measurements. This gives an indication of individual performance for each product but should not be used as a primary source to rank inter-product performance as the data is not collocated by dates or stations.

All products generally showed positive correlations with ICOS measurements, although for most products a few stations had R values of 0 or even negative. The lowest R value was approximately -0.3. The median R for the products ranged from 0.25 at Sentinel-1 to 0.71 at SMAP L4, with all products except Sentinel 1 and ASCAT having a highest correlation above 0.8. Most products had a median unbiased root-mean-squared-deviation (ubRMSD) at the stations of around 0.05 m^3/m^3, with some outliers reaching 0.20 m^3/m^3 and higher. Note that the target error, or required accuracy for many products is 0.04 m^3/m^3 (Gruber et al., 2020; Kerr et al., 2010) which the products generally fell slightly short of, even if it was reached by some products at individual stations.

ASCAT had medium to large correlation to ICOS data at most stations, with R values being generally between 0.3 and 0.6, and even higher correlation coefficients at some stations, for example at BE-Lcr and FR-Aur (Figure 6). The median R was 0.47. ASCAT's ubRMSD was under 0.14 m^3/m^3 at all stations with a median of 0.074 m^3/m^3. Sentinel-1 had a median of 0.25, with several stations significantly below 0. Its median ubRMSD was 0.068 m^3/m^3 and it had an outlier above 0.20 m^3/m^3. SMOS had most of its R values in the range between 0.4 and 0.8, with over 0.8 at FR-Aur and FR-Mej. The ubRMSD for SMOS L4 was in a relatively narrow range between 0.05 and 0.10 m^3/m^3, with a few higher outliers.

SMAP L3E had several large correlations with ICOS stations, with R values over 0.8, for example at DE-RuR, BE-Maa, and FR-Tou. Its median R was approximately 0.65. However, it showed high variability, as it also produced some negative Rs, such as at CH-Dav, and SE-Deg. As for ubRMSD, SMAP L3E mostly produced errors of around 0.05 m^3/m^3, but just as for most products, there were some higher outliers. SMAP L4 contained a large median R, at over 0.71, and had a consistently high performance with R values at most stations being between 0.6 and 0.8. The median ubRMSD for SMAP L4 was 0.044 m^3/m^3. ESA CCI SM had a median R of ca 0.66 but a wide range of values, including one negative outlier of nearly -0.2 at SE-Deg. Generally, the R was in between 0.5 and 0.8. The median ubRMSD was 0.045 m^3/m^3, with relatively small outliers.

As for the pure reanalysis products, ERA5-Land mostly had medium to large correlations, with R values between 0.55 and 0.8, and a median of ca 0.68. The ERA5-Land median ubRMSD

was approximately 0.049 m^3/m^3. GLDAS Noah had a median R of 0.63, with a relatively large range of values. Likewise, the ubRMSD had a range from 0.02 m^3/m^3 to 0.12 m^3/m^3 when excluding the outliers, with a median of ca 0.057 m^3/m^3.

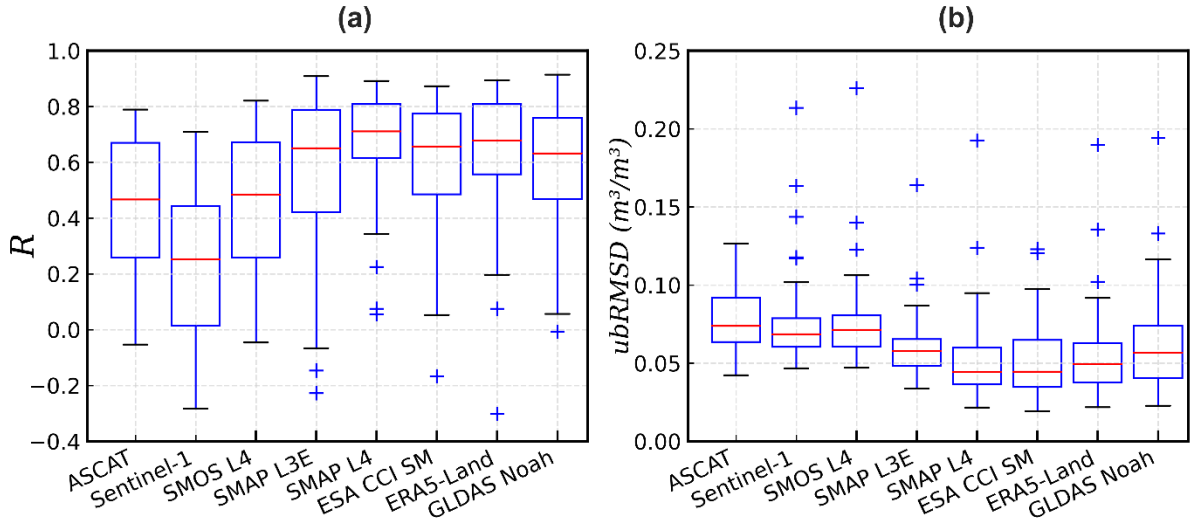

Figure 5. The evaluation metrics R (a) and ubRMSD (b) for each SM product at the ICOS stations. The values were calculated based on absolute SM and non-collocated dates between products.

Figure 6 displays the R for each product at each of the 38 ICOS stations based on the absolute SM. There was a large variation between the different products and stations. For some sites, such as FR-Mej, all products had a relatively large correlation (>0.7), while there were other stations such as FR-LGt and DE-Tha where all product's correlations were close to 0, or even negative. The stations where all correlations were low might be indicative of measurement stations that are less representative for the SM on the product grid cell scale.

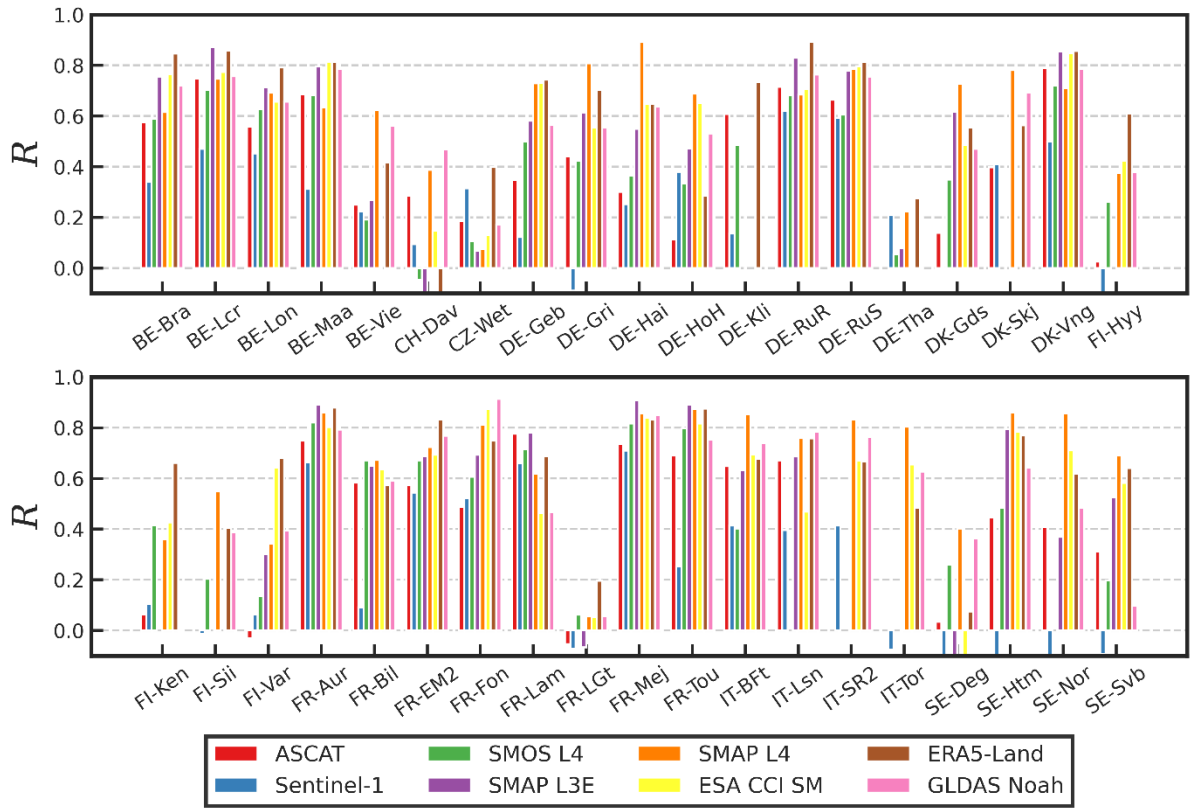

Figure 6. The evaluation metric R for each SM product at each of the ICOS stations. The values were calculated based on absolute SM and all available dates for each product.

Product correlation with short-term anomalies is displayed in Figure 7. Across all products, the R values were toward 0.1 lower than for the absolute SM. Some products, such as GLDAS Noah, showed even larger differences in correlation compared to the absolute SM. Most products and stations seemed to have a medium correlation for short-term anomalies. As for the ubRMSD, five out of eight products' median was below 0.04 m^3/m^3.

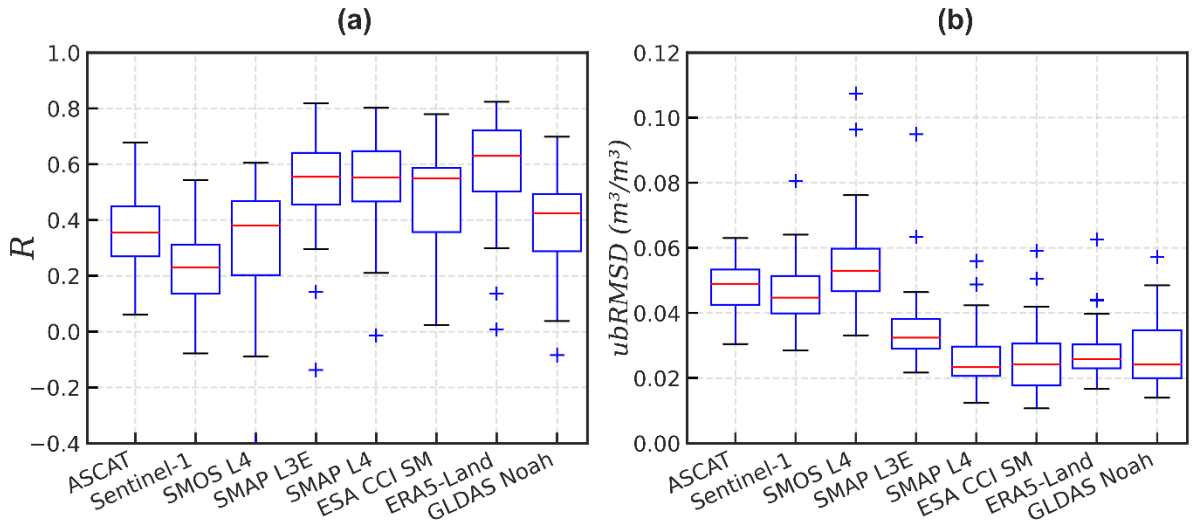

Figure 7. The evaluation metrics R (a) and ubRMSD (b) for each SM product at the ICOS stations. The values were calculated based on anomalies and all available dates for each product.

5.1.2. Results considering temporal collocation

Figure 8 shows the performance of the products as calculated for common dates for each ICOS station, and only stations that all products have data for, thus allowing for a fair inter-product comparison. Out of the eight products, ASCAT, Sentinel 1, and SMOS L4 had the lowest three R values while the remaining five products consistently seemed to produce higher correlations.

Sentinel-1 seemed to have the lowest scores of the three poorer products as it generally produced lower R values than ASCAT and SMOS L4, even if it had similar ubRMSD to ASCAT. SMOS had a higher median ubRMSD (0.072 m^3/m^3) than ASCAT (0.077 m^3/m^3) and Sentinel-1 which otherwise had a larger, and nearly identical range of values.

The five better performing products had similar values both for R and ubRMSD, excluding outliers, and it is difficult to say which ones performed better with certainty. The highest median R for the collocated dates was found in ESA CCI SM (and GLDAS Noah), at over 0.9. However, ESA CCI SM and GLDAS had lower medians for R than SMAP L4 and ERA5-Land. In addition, ERA5-Land was the only product with no R-values below 0. As for the ubRMSD, SMAP L4 had the lowest median, ca 0.044 m^3/m^3, while it was highest for SMAP L3E, at ca 0.052 m^3/m^3.

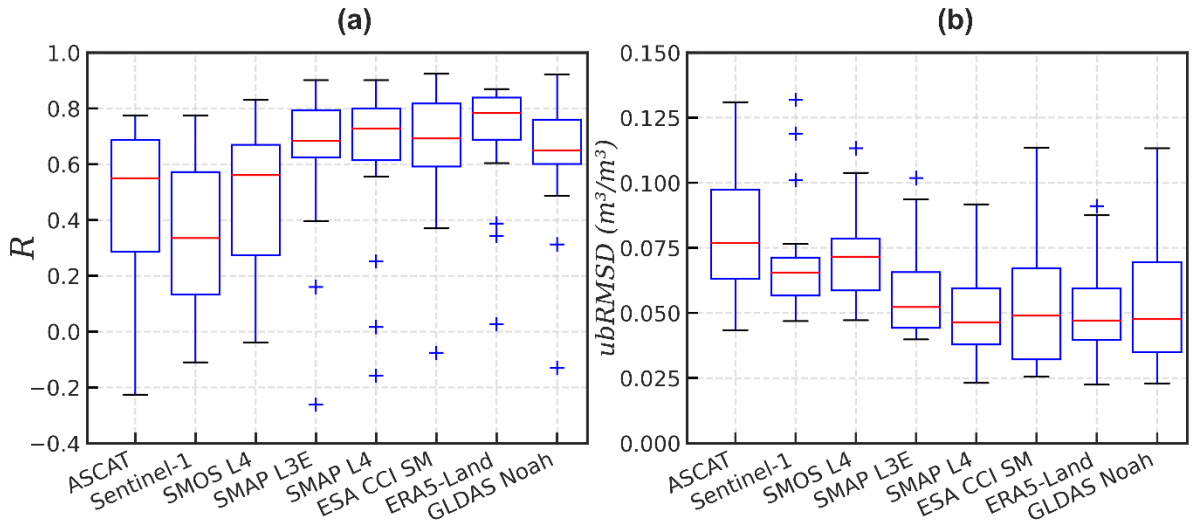

Figure 8. The evaluation metrics R (a) and ubRMSD (b) for each SM product at the ICOS stations. The values were calculated based on absolute SM and collocated dates between all eight products, and at the same stations where all products had valid data (23 common stations).

The five better performing products were temporally collocated to increase the underlying data without considering the three lower products, resulting in Figure 9. With the increased temporal coverage for the comparison, all five products had similar median R values to the previous comparison, perhaps with the main exception of GLDAS Noah which saw a slight increase. The range of values seemingly increased in the lower end for SMAP L4, ESA CCI SM, and ERA5-Land. SMAP L4 had the highest median R, and some of the higher lowest values of R. Further, SMAP L4 had the lowest median ubRMSD out of the five products, but all products had similar values, except for SMAP L3E which had a slightly higher median but a smaller range of values than the other four products.

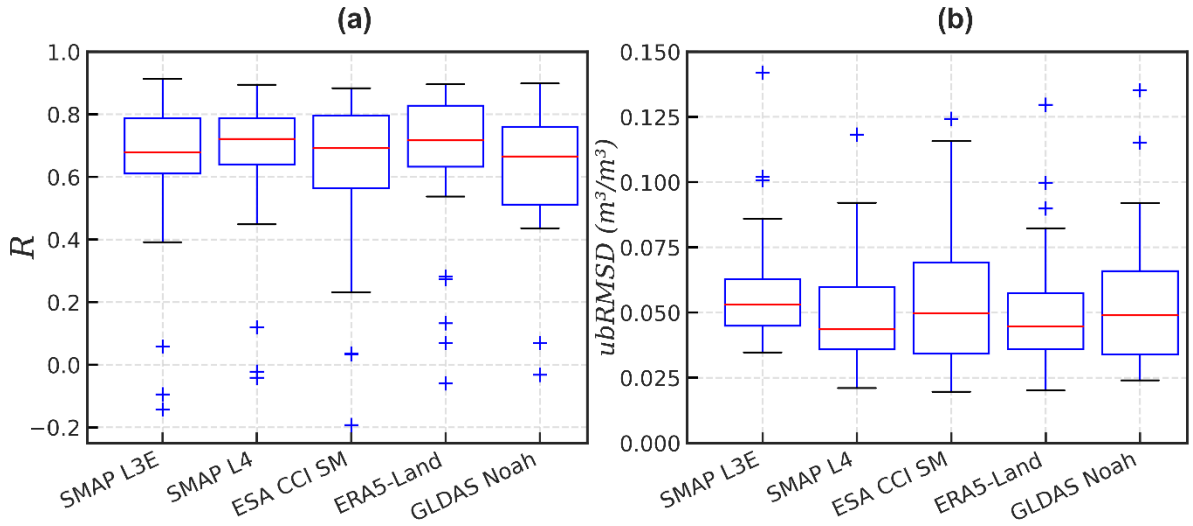

Figure 9. The evaluation metrics R (a) and ubRMSD (b) for each SM product at the ICOS stations. The values were calculated based on absolute SM and collocated dates between the five products in the figure (23 common stations).

As for short-term SM anomalies, GLDAS Noah generally captured these poorly compared to the others, with a low R, the median just above 0.4 (Figure 10). ERA5-Land had the highest median R by some margin and seemed to have the highest correlation overall, even if the lower end was quite low, while SMAP L3E and SMAP L4 performed relatively similarly in terms of R, L3E having higher median. ESA CCI SM had lower R values than the SMAP products. For ubRMSD, SMAP L3E had the highest values, while ESA CCI SM and GLDAS Noah had the lowest medians. Although, GLDAS Noah had a large range of values, while for example ERA had a much more consistent performance in terms of the errors.

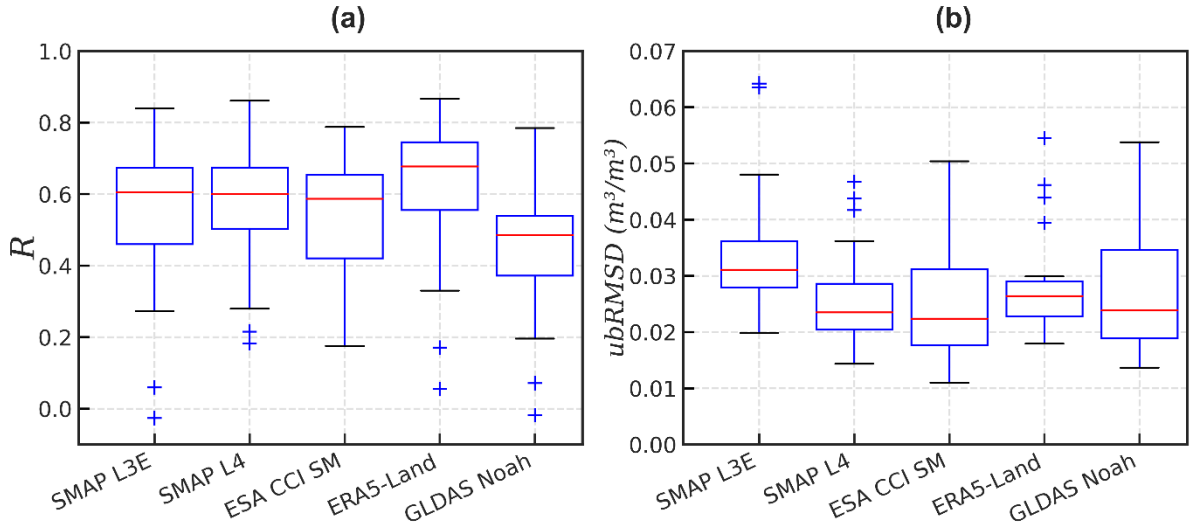

Figure 10. The evaluation metrics R (a) and ubRMSD (b) for each product at the ICOS stations. The values were calculated based on SM anomalies and collocated dates between the five products in the figure (at 29 common stations).

In general, GLDAS Noah performed similarly to the other four products for absolute SM but was clearly lacking terms of capturing SM variability with the anomalies compared to the other five products. While SMAP L3E showed comparable performance to the other products in terms of the correlation, it had the highest median ubRMSD both for absolute SM and anomalies. For absolute SM this was because SMAP L3E did not have any stations with ubRMSD below 0.03 m^3/m^3, unlike the other products. ESA CCI SM had much lower correlation with anomalies than ERA5-Land, while SMAP L4 was the only product to consistently perform in the top 2-3 for every metric considered. SMAP L4 and ERA5-Land performed similarly. The ranking of the products can vary depending on subjective prioritization factors such as consistency, statistical scores, temporal and spatial coverage, among others. However, if we consider a ranking based solely on their general statistical scores, it could be presented as follows (1 means the best):

1. SMAP L4, 2. ERA5-Land, 3. ESA CCI SM, 4. SMAP L3E, 5. GLDAS Noah, 6. SMOS L4, 7. ASCAT L4, 8. Sentinel-1.

5.2. Evaluation with triple collocation analysis (TCA) and results of merging

5.2.1. Results of TCA

The results presented in this section are valid for the locations of the ICOS stations, which was done in order to minimize spatial bias when comparing the results to those of the in situ measurement-based evaluation. Out of the three gridded products, SMAP L3E showed the highest SNR and lowest ubRMSE for both grid sizes / resampling methods at the ICOS sites. SMAP L3E was estimated to have a median SNR of 5.4 for the 0.25 ° resampling and 6.9 for the 0.1° resampling, as well as a ubRMSE of ca 0.10 m^3/m^3. That SMAP L3E performed better than ASCAT and GLDAS Noah at the ICOS site locations is in line with the results from the traditional evaluation. ASCAT and GLDAS Noah seemed to perform similarly according to the TCA (Figure 11). The threshold of 0 SNR was breached at least one station for all products, but the median was above the target of 3 for five of them, only ASCAT at the 0.25° size falling slightly short.

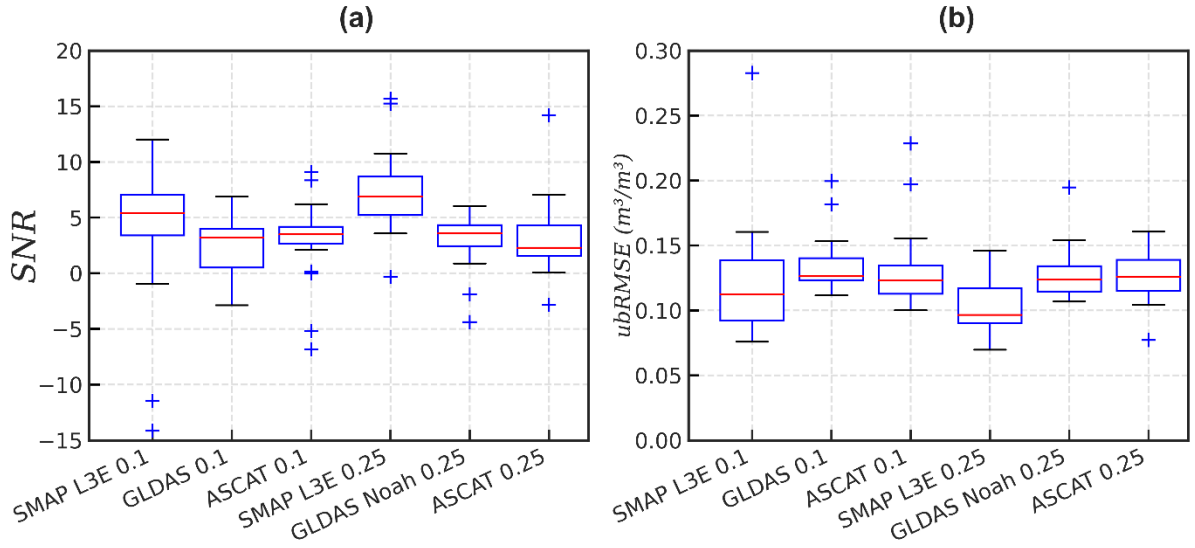

Figure 11. SNR (a) and ubRMSE (b) at each of the ICOS stations for the three products, and both using the ERA 0.1° grid and the GLDAS 0.25° grid. Calculations done on absolute SM.

5.2.2. Merging multiple products based on TCA

Due to limited time and data processing capabilities, it was not possible to perform the TCA based merging over all of Europe. Instead, it was done for southern Scandinavia, including Denmark, southern Sweden and Gotland, and southwestern Norway. This is an area with several land use types: large agricultural areas, especially in Denmark, large forests in much of Sweden and eastern Norway, and strong topography and high altitudes in the western Norwegian parts. Figure 12 shows the SM distribution for each of the parent products, the TCA-weighted merge, the arithmetic mean of the parent products, and ERA5-Land as a reference product, averaged over July 2021. Firstly, note the blank regions for the TCA-weighted merge, particularly in western Norway. Here, the TCA based merging was not conducted due invalid TCA conditions in the three parent products. Considering the large

topographical gradients and cold conditions at high altitudes, this was not unexpected. Besides southwestern Norway, the TCA-weighted merging was skipped at several locations in Sweden, especially in Värmland County, north of Lake Vänern.

Pattern wise, the TCA-weighted merge was most similar to SMAP L3E. Both SMAP L3E and the TCA-weighted merged model indicated relatively dry conditions, in this case SM of between 0.15 m^3/m^3 and 0.25 m^3/m^3 in predominantly agricultural regions such as Denmark, Skåne County, Gotland, as well as in between the lakes Vänern and Vättern. The rest of the area was mostly estimated to have wet conditions, with SM estimates generally above 0.40 m^3/m^3.

GLDAS Noah estimated dry conditions over much of the area with very minor spatial variation, and SM values consistently between 0.15 and 0.25 m^3/m^3. ASCAT seemed to produce SM values somewhere between SMAP L3E and GLDAS Noah, from 0.25 m^3/m^3 to 0.35 m^3/m^3. In fact, the arithmetic mean intuitively appeared to have similar estimates to the ASCAT.

Most of Denmark appeared to have relatively high soil moisture according to ERA5-Land, which was the opposite of what the other products indicated. In fact, ERA5-Land's pattern generally seemed to be inverted compared to SMAP L3E and the TCA-weighted merge, with relatively higher SM in the areas where the TCA-weighted SM was low, and vice versa.

Note that the descriptions for the spatial SM pattern in Figure 12 are only valid for July 2021, and might not be representative of the whole study period. See appendix Figure A1 for a further figure on which merging method was selected for each pixel.

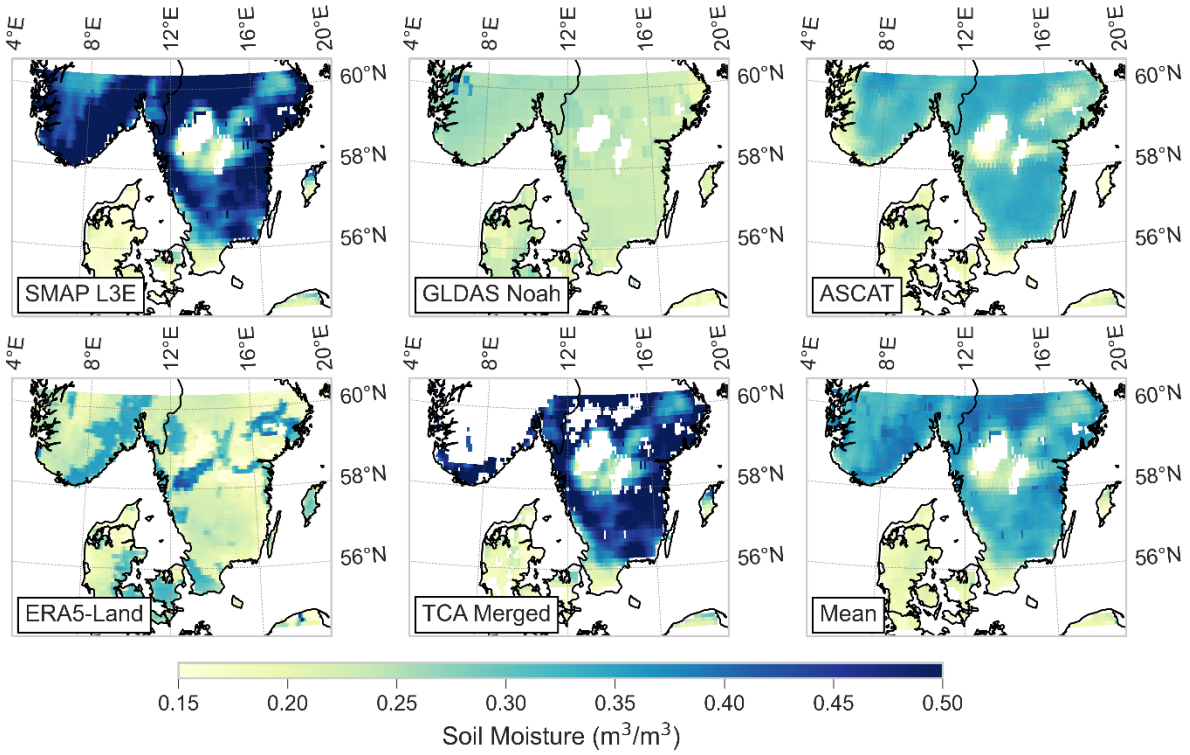

Figure 12. Comparison of SM estimates from three parent products (SMAP L3E, GLDAS Noah, ASCAT), TCA-weighted merged, the arithmetic mean of the three parent datasets, and ERA5-Land over southern Scandinavia. The values were calculated over the 0.1° grid and averaged over July 2021.

The merging model was also run over all the ICOS sites, allowing for it to be compared with the other gridded products. Figure 13 shows the R and ubRMSD for each of the merged/mean products absolute SM against ICOS for all available dates at all available stations. It should be noted that the dates used are essentially the same for the four merged products with maximum a few differences in dates per station, as a result of the resampling. Note that due to TCA violation over several ICOS stations, only ca 20 stations per product could be included.

The TC merge at 0.25° had the highest R values, with a median of circa 0.718, while R was generally lowest for the 0.1° mean product. 0.1° TC weighted merging and 0.25° means gave similar R values. For both 0.1° and 0.25° however, the TC weighted merging gave slightly higher correlations than the arithmetic means, but the difference was small. 0.1° mean had the lowest median ubRMSD, at ca 0.045 m^3/m^3. However, the ubRMSD was very similar across all products, between 0.04 and 0.05 m^3/m^3.

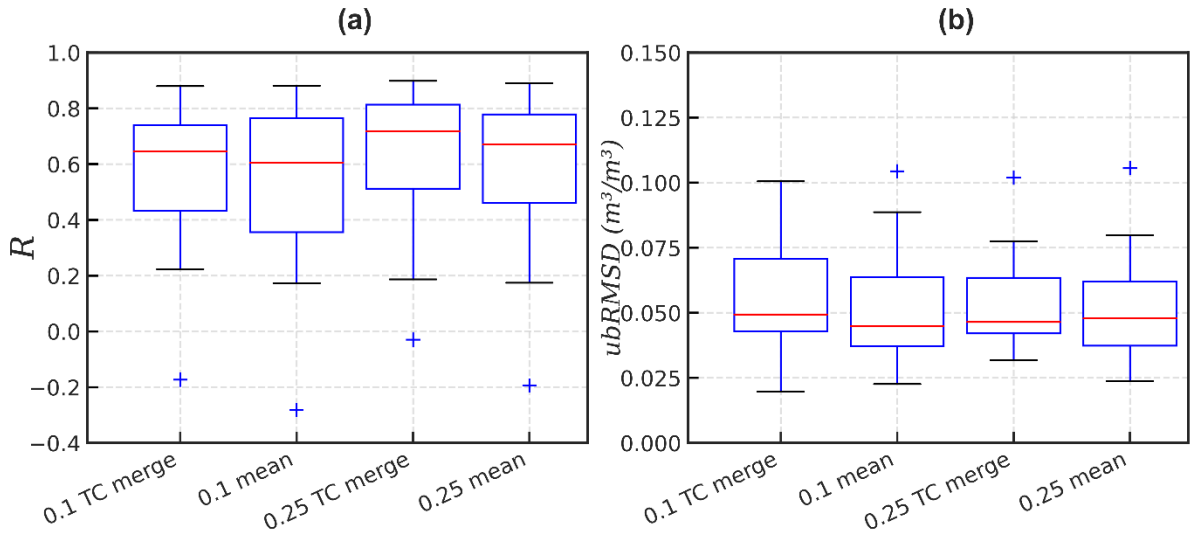

Figure 13. The evaluation metrics R (a) and ubRMSD (b) for the merged SM estimates using TC-based merging (denoted as TC merge) and using simple mean method (denoted mean) at the ICOS stations. The values were calculated based on the absolute SM.

As for the anomalies, the order of performance for R was nearly the same as for the absolute SM, with the highest correlation being found for the 0.25° TC weighted merging, at ca 0.6 (Figure 14). ubRMSD was slightly lower for the arithmetic mean produced products, but the difference was very small.

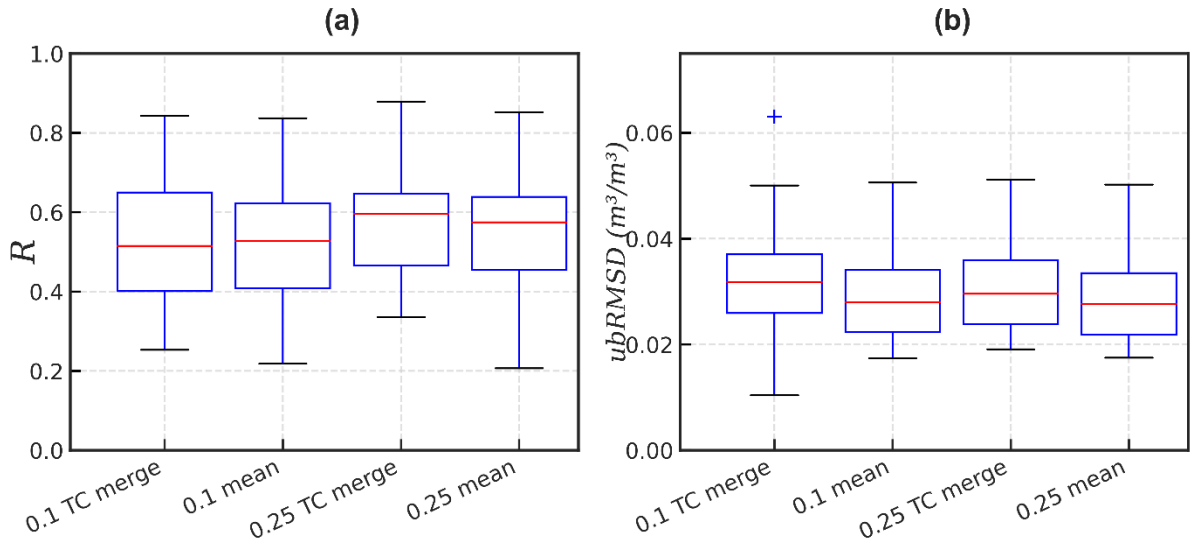

Figure 14. The evaluation metrics R (a) and ubRMSD (b) for the merged SM estimates using TC-based merging (denoted as TC merge) and using simple mean method (denoted mean) at the ICOS stations. The values were calculated based on short-term anomalies.

5.2.3. Comparison of merged product with other products

When comparing the TCA-weighted and mean products with the pre-existing products on collocated dates (Figure 15), only 13 stations could be used, hence why the inter-product

rankings are slightly different from previous comparisons. While the median R for the 0.1° weighted product was similar to GLDAS Noah, the rest of the distribution suggested that both 0.1° TC weighted, 0.1° mean, 0.25° TC weighted, and 0.25° mean, had larger correlations with absolute SM than both GLDAS Noah and in particular ASCAT. For the four created products, the correlation was similar to that of SMAP L3E, as well as the other three products included (SMAP L4, ESA CCI SM, ERA5). As for the ubRMSD for absolute SM, it was on the same scale for the new products as for the pre-existing ones, besides a clear improvement on ASCAT.

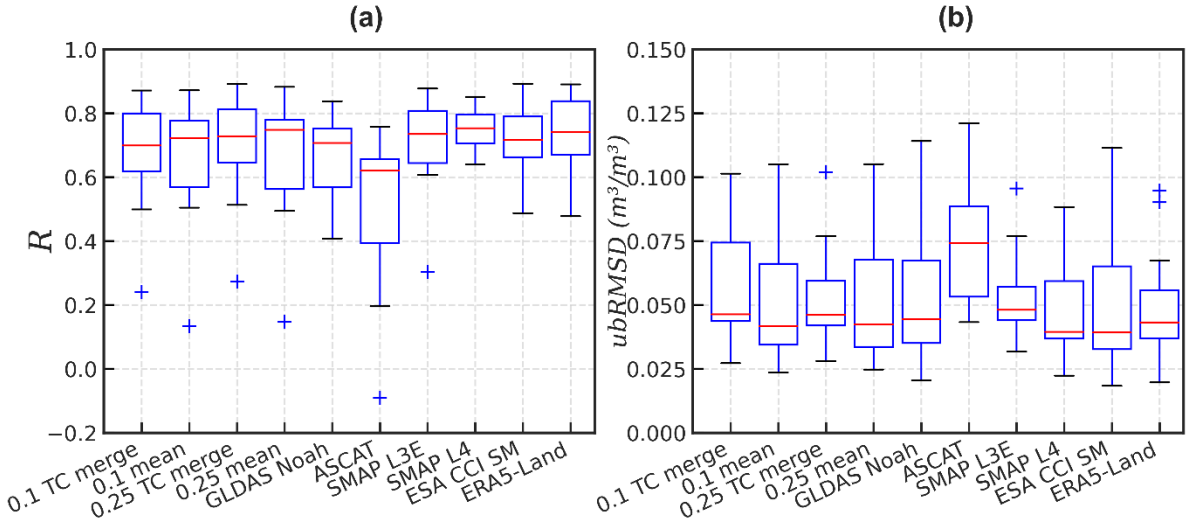

Figure 15. The evaluation metrics R (a) and ubRMSD (b) for the merged and mean products, as well as pre-existing products, against ICOS measurements. The values were calculated based on absolute SM and collocated dates.

For SM anomalies (Figure 16), the median correlation of all four created products was clearly higher than for GLDAS Noah and ASCAT. The 0.25° TC weighted merge also seemingly had slightly lower correlation than SMAP L3E and ERA5, but higher than SMAP L4. Overall, (Figures 15 and 16), there did not appear to be a clear difference in performance between the TCA-weighted merge, and the arithmetic mean of the three products.

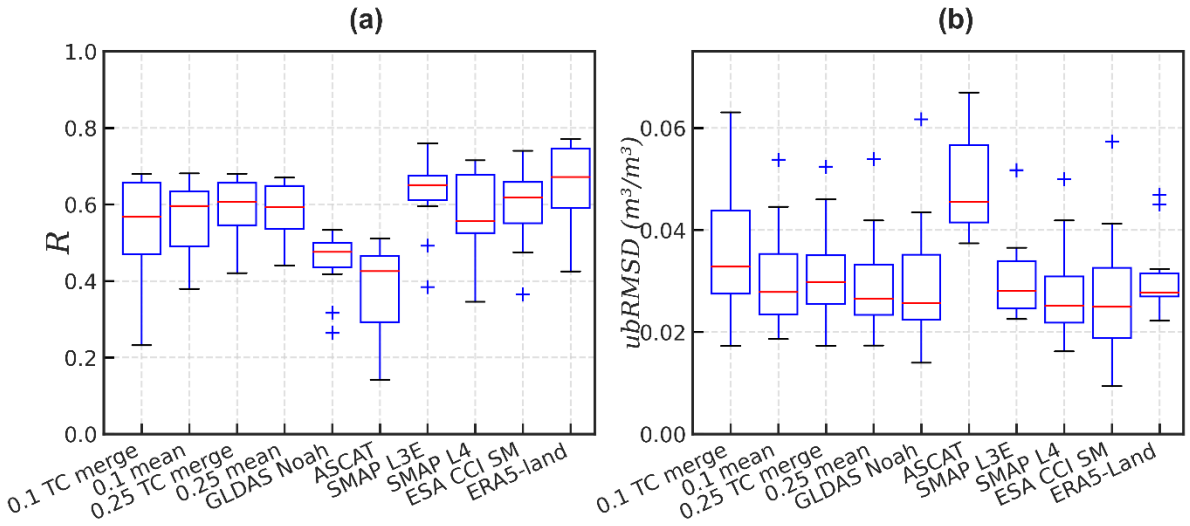

Figure 16. The evaluation metrics R (a) and ubRMSD (b) for the merged and mean products, as well as pre-existing products, against ICOS measurements. The values were calculated based on SM anomalies and collocated dates.

6. Discussion

6.1. Evaluation of products with in situ measurements

Firstly, it should be noted that when comparing the results of this study to results of other studies, location and time of the data used is different, which might account for any differences and always needs to be considered. On the results, it was directly relatively clear that the products could be grouped into a better performing group, which included the SMAP-products, ESA CCI SM, ERA5-Land, and GLDAS Noah, as well as group that had significantly lower statistical scores, consisting of Sentinel-1, SMOS L4, and ASCAT.

The two gridded products with 1 km resolution, SMOS L4 and Sentinel-1, seemingly performed worse than the coarser products in the in situ measurement-based evaluation. The smaller difference in scale between grid cell and in situ measurement could have been assumed to give ICOS measurements more representativeness over the product cell footprint, considering that the proportions of the land cover are generally higher over the 1 km square than the larger square, but this advantage was seemingly outweighed by other factors.

In previous validation studies of SMOS L4, its performance has varied, with Spatafora et al. (2020) finding satisfactory correlations at the REMEDHUS network in Spain that were larger than those found in this study. However, it can be noted that the better scores in the study by Spatafora were particularly found for SM only derived from ascending passes, with lower correlations and higher errors for SM derived from descending passes, which might account for some of the difference to this study in which equal weights were given to the two overpass times.

Portal et al. (2020) found relatively consistent and good performance between SMOS and SMAP's downsampled versions, including strong performance for SMOS L4 at the

REMEDHUS network. However, when compared to SMAP products over Great Britain, as done by Peng et al. (2021), SMOS L4 were shown to perform significantly worse than the sensor-wise comparable, but less significantly downsampled SMAP products. This difference in performance was attributed to the comparably poor radio frequency interference filtering (RFI) of SMOS (Peng et al., 2021), as well as the downscaling method and additional input data, including MODIS NDVI, as the input data had issues stemming from varying sensing depths (Lamptey et al., 2021). Ideally, SMOS L4 would also have been compared to the lower level SMOS products in this study for further estimation of the magnitude of differences caused by downsampling.

Peng et al. (2021) also found relatively low correlation and high ubRMSD for Sentinel-1 in Great Britain when compared to other products, i.e., their results for both Sentinel-1 and SMOS L4 were similar to those of this study, albeit with slightly better scores. Although, it should be noted that that was against cosmic-ray COSMOS measurements, which are assumed to be more consistent with reality over the 1 km cell than point measurements (Bogena et al., 2022) such as ICOS. Sentinel-1 is also currently missing a dynamic vegetation correction, an addition of which would potentially enhance its performance (Bauer-Marschallinger et al., 2019).

SAR C-bands, employed by Sentinel-1 and ASCAT, are less sensitive to SM than the radiometer L-bands of SMOS and SMAP (Edokossi et al., 2020). It is assumed that that is why ASCAT also showed relatively poor performance. Previously, ASCAT has been shown to perform better against in situ data than Sentinel-1 in Italy (Bauer-Marschallinger et al., 2019) and Great Britain (Peng et al., 2021). However, El Hajj et al. (2018) found Sentinel-1 to have significantly higher correlations against the SMOSMANIA network in southwestern France. In fact, El Hajj found the correlations to be higher than ASCAT. This indicates that performance of Sentinel-1 might have strong regional variability, perhaps caused by the absence of dynamic vegetation correction. The average R found for ASCAT by Hajj et al. (2018), was 0.44, which was close well to the median R of 0.47 it had to ICOS in this study. In fact, if using the average of station R values instead of the median, ASCAT's R was 0.44 against ICOS as well. Both Sentinel-1 and ASCAT needed to be converted from degree of saturation (%) to volumetric soil water content, using a coarse porosity dataset, which further may have increased the magnitude of errors.

The SMAP-products produced strong scores for essentially all investigated metrics, further confirming that the L-band successfully can be used to consistently produce relatively reliable SM estimates, which has also indicated by previous studies (Beck et al., 2021; Ma et al., 2019). SMAP L4, which assimilates the radiometer data with meteorological data in a geophysical model, seemed to perform particularly well, and it produced arguably the best scores out of any gridded products included in this study, together with ERA5-Land.

Al-Yaari et al. (2019) found the non-enhanced SMAP L3 to have higher correlations against ground-based measurements than ESA CCI SM in Europe, followed by ASCAT. This is consistent with the findings in this study, as SMAP L3E and ESA CCI SM perform better than ASCAT. However, SMAP L3E produced slightly inferior statistical scores than ESA CCI SM against the ICOS measurements. As for SMAP L3E, El Hajj et al. (2018) found its mean station correlation to be 0.65 which is also consistent with the results in this study.

ESA CCI SM performed considerably better than SMOS L4, Sentinel-1, and ASCAT, slightly better than SMAP L3E and slightly worse than SMAP L4 and ERA5-Land. Overall, the scores indicated that ESA CCI SM is relatively successful. That it performs better than ASCAT, and low level SMAP and SMOS products was also shown by Zheng et al. (2022) for the Shandian River Basin. Zheng et al. reported R values for ESA CCI SM at ca 0.6 which was similar to what was found against ICOS stations in this study, whereas ESA CCI SM produced a median of 0.66. However, Zheng et al. found a quite low ubRMSE, at below 0.03 m^3/m^3, compared to the 0.045 m^3/m^3 found against ICOS. ESA CCI SM has previously also been shown to perform well globally in relation to the comparable, blended SMOPS product (Wang et al., 2021), which, joint with the results presented here, further indicates that it can be considered to represent the current state of the art among merged SM products.

GLDAS Noah showed clear issues in capturing short-term SM variability despite producing good scores for most other metrics. ERA5-Land performed clearly better for the anomalies than GLDAS Noah, indicating that it is more sensitive to SM variation. In fact, ERA5-Land scored relatively highly for most metrics. Beck et al. (2021) as well as Zheng et al. (2022) also found ERA5-Land to outperform GLDAS Noah, with Zheng et al. also highlighting GLDAS Noah's issues with the short-time SM changes. This issue may be partly caused by GLDAS Noah estimating SM in the top 10 cm instead of the top 5 cm, as deeper soil is less susceptible to strong, sudden SM variations.

The ranking in descending order of performance: SMAP L4, ERA5-Land, ESA CCI SM, and GLDAS Noah, was in similar agreement with what Fan et al. (2022) found in Jiangsu Province, China, for correlation with in situ measurements, that however had ESA CCI behind both GLDAS Noah and ASCAT. Fan et al. (2022) also found that ASCAT performed better than SMAP L3, which goes against the findings of this study.

Overall, the findings were relatively consistent with other studies in the field, with similar R-values, and slightly higher ubRMSD-values. It can be noted that the three best performing products had different approaches, specifically: the assimilation of satellite data in a geophysical model of SMAP L4, the fully model-based ERA5-Land, and the various satellite data blending of ESA CCI SM. The conclusion from this is that there are several different ways and no definitive best method for generating accurate SM estimates in Europe.

Further, the relative similarity to scores found in literature indicates that the ICOS measurements are promising for use in future validation studies. That ubRMSD was generally above the target of 0.04 m^3/m^3 for the products is assumed to be an effect of representativeness issues for the in situ measurements compared to the coarse grid of the products (Al-Yaari et al., 2019).

The suitability of specific ICOS stations for the study is assumed to vary, as the inter-station variability in product performance was significant, with some stations seemingly having high correlations and low errors for all stations, and all product performance being poor at other stations. Examples of stations for which the differences against the gridded products were generally around or below 0.05 m^3/m^3 included BE-Maa, DE-Tha, and FR-Aur while stations with differences generally above 0.10 m^3/m^3 included FI-Sii, SE-Deg, and DE-Gri (Figure 17).

The pattern is similar to the stations that showed high and low correlations, respectively, indicating that station wise ubRMSD and R are related (see Figure 6 in section 5.1.1.).

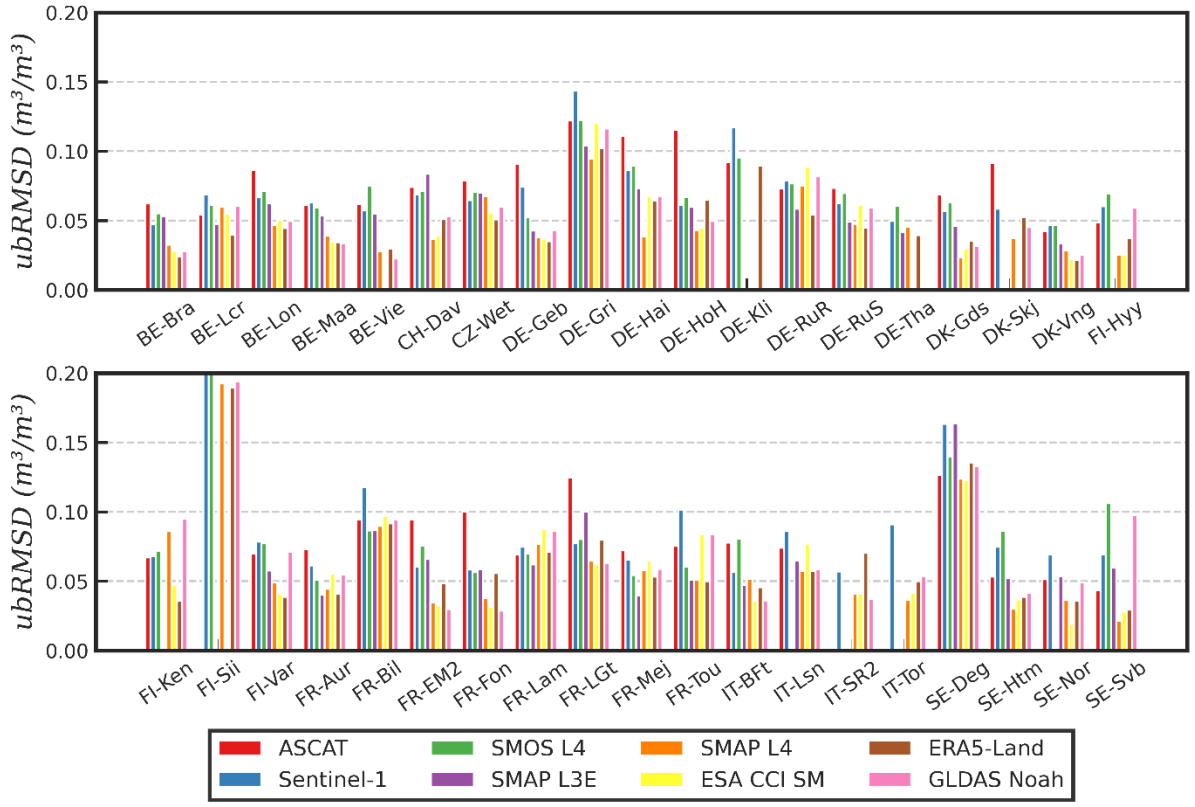

Figure 17. The evaluation metric ubRMSD for each SM at each of the ICOS stations. The values were calculated based on absolute SM and all available dates for each product.

The direct correlation between SM and land surface properties has been thoroughly investigated within the field, with studies finding various effects of land cover and spatial heterogeneity on soil moisture and SM product errors (Al-Yaari et al., 2014; Venkatesh et al., 2011; Wang et al., 2021). The consensus is that there is a relatively clear direct relationship (Chen et al., 2009; Hu et al., 2022; Panciera, 2009). Lakhankar et al. (2009) showed that sub-pixel heterogeneity in vegetation cover reduced microwave derived SM data accuracy. In particular vegetation density has an effect on the performance of the product (Ma et al., 2019).

To investigate this issue further, the correlation for SMAP L3E was grouped over different ICOS land covers. There was a large spread in correlation at evergreen deciduous forests, while it performed well over deciduous forests, croplands, and grasslands and poorly over wetlands (Figure 18). Wetlands have consistently different soil moisture to surrounding land cover types, not following the overarching temporal dynamics, which explains the poor performance here. High wetland proportion within the pixels is known to make retrieval more uncertain (Leroux et al., 2013). Thus, the suitability of wetland station in evaluation studies is dubious. However, it should be mentioned that wetland sites have been included in ISMN and validation studies previously, e.g., for product calibration (Dorigo et al., 2011; Marczewski et al., 2010).

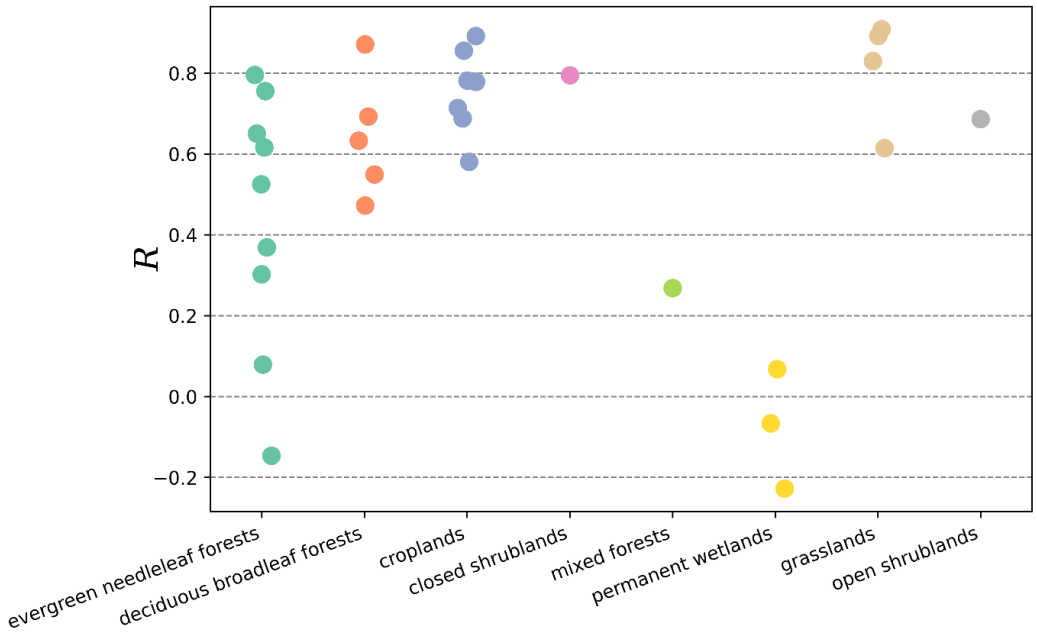

Figure 18. The evaluation metric R at each station for SMAP L3E, grouped by ICOS site.

A brief look at the performance of SMAP L3E at ICOS stations with different levels of land cover proportion within a 9 km horizontal square seemed to indicate that a higher proportion of the station land cover class in the surrounding area generally led to higher correlations (Figure 19), at least up to a certain point between 0.5 and 0.7, although there is a low sample size. Note that the land cover and representativeness in themselves are linked and related, and that no adjustment has been done to account for this due to the low sample size. To further illustrate this point, wetlands generally made up a small portion of the proportion while evergreen forest generally made up for a large part of the proportion in the square (see Table A1 in appendices for the land cover and its representativeness at each station).

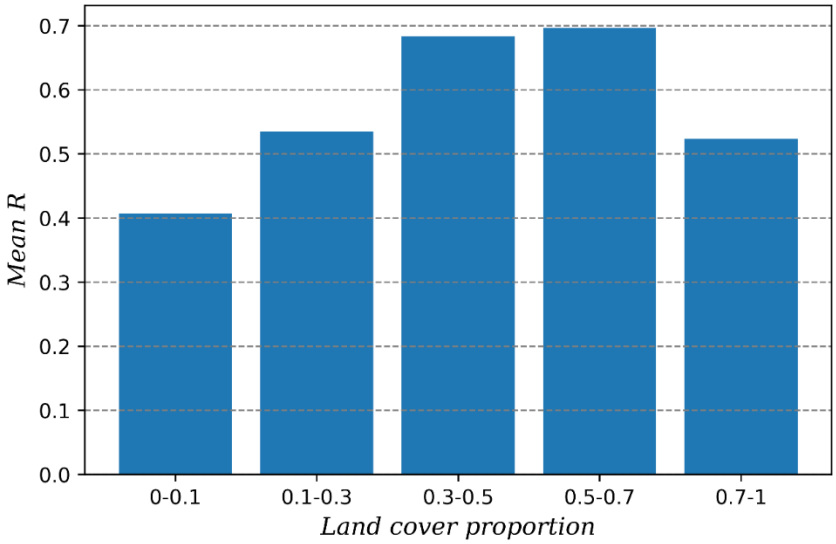

Figure 19. The average of the metric R for SMAP L3E at ICOS stations for different proportions of the station land cover in a 9 km horizontal square centered on the station.

6.2. Evaluation with the triple collocation analysis and merging

For ubRMSE, TCA reinforced the results from the in situ-observation based evaluation, that SMAP L3E performed better than ASCAT and GLDAS Noah. However, the ubRMSE were several orders of magnitude above the target of 0.04 m^3/m^3 for all three products, indicating that the targets errors were not reached for of them. While GLDAS Noah had markedly lower ubRMSD in the in situ measurement-based evaluation, the difference between ASCAT and GLDAS Noah was negligible in regard to TC estimated ubRMSE. However, the results from the presented TCA analysis were only done over the ICOS stations. When looking at error variance, i.e., ubRMSE squared, over southern Scandinavia instead (Figure 20), it seemed like the errors were more pronounced for ASCAT than GLDAS Noah over many regions, in particular south-eastern Sweden, but GLDAS Noah had larger relative error variance in southern Denmark, and Värmland County, Sweden. This further illustrates the spatial variability of performance of different products.

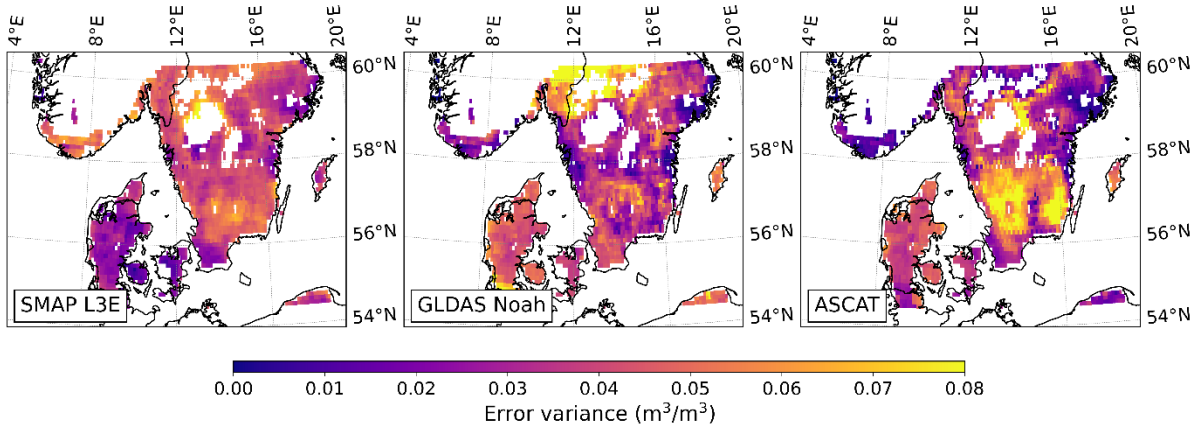

Figure 20. The error variance at each of the pixels for each product in Southern Scandinavia according to TCA of the triplet at the 0.1° grid.

The relative SNR was in line with the ubRMSE values for the products, with relatively high signals from SMAP L3E and minor differences between ASCAT and GLDAS Noah. For SMAP L3E, several locations showed SNR at above the optimal target threshold of 6, and seemingly, SNR was higher than in comparable study by Peng et al. (2021) which used a similar triplet. However, the retrieved ubRMSE was also significantly higher in this study. As the errors are still relatively high, the validity of the high SNR of SMAP L3E is assumed to be less robust. For that reason, it would be encouraged to further produce confidence interval estimation, based on bootstrapping methods (F. Chen et al., 2018; Gruber et al., 2020). However, due to time limitations, it was not done in this study. SNR is directly linked to the R between the product and the assumed true SM. The presentation of SNR allows for a comprehensive assessment of actual data quality, which provides an additional perspective to that of the TC-based error variances.

The reasons for the ranking in performance for SNR can likely be derived to much the same reasons as for the in situ observation-based evaluation, i.e., that the L-band radiometer, as well as the algorithms which SMAP L3E employ has advantages to SM, that in this region outweigh

the resolution and model advantages, compared to the C-band in ASCAT and the model aspects of GLDAS Noah, respectively. As for GLDAS Noah, the relatively poor performance is likely explained by it struggling with the short-term temporal variations, which further confirms that its SM sensitivity is relatively low.

Looking at the distribution of weight assignment for the merging (Figure 21), which is directly related to the inter-product relations of error variance, SMAP L3E did not seem to have significantly higher weight assignments in southern Scandinavia than the other products, besides in the agricultural dominated region in Denmark. Note that this not necessarily true for the rest ICOS stations, for which the product statistics were shown for previously, e.g., many of them were on agriculture land covers.

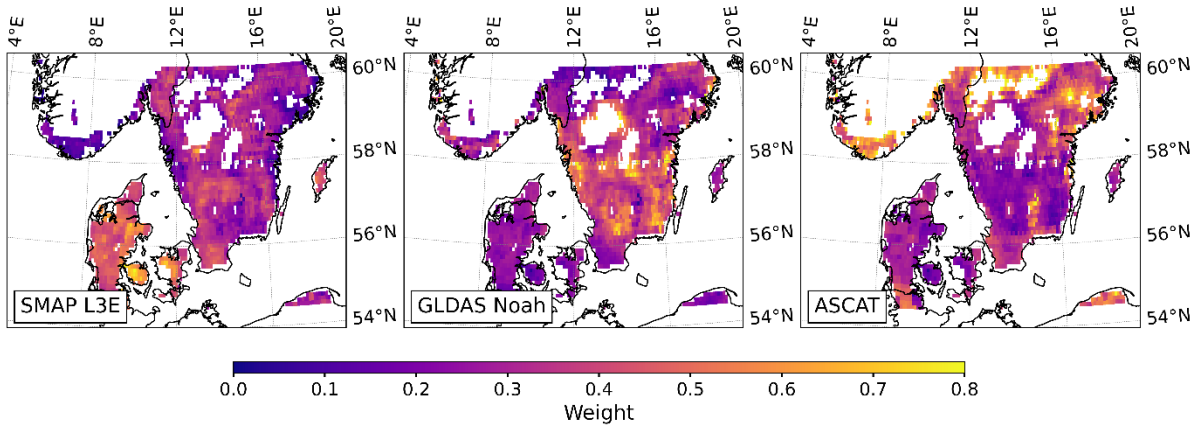

Figure 21. The weight assigned to each parent product at each of the pixels.

However, note that for a relatively large proportion of the area in Denmark, the selected merging method was not based on the TCA estimated weights but instead of the arithmetic mean of SMAP L3E and ASCAT or of SMAP L3E and GLDAS Noah (see Figure A1 in appendices), according to the merging scheme. As explained by Peng et al. (2021), this might have caused a decrease in accuracy in the product compared to a purely TCA-weighted merge, but it will have increased the coverage of the product, which is an important considerations for the product application, especially if one of the parent products has poor temporal coverage. It would have been interesting to also conduct this analysis without the merging scheme to further draw conclusions on the relationship between accuracy, selection of merging method, and temporal coverage.

The aim was that the merged product should perform better than all the parent products. Instead, it could not be said with certainty that the merged products even performed better than the arithmetic mean of the parent datasets. This has also occurred in previous studies (Peng et al., 2021; Yilmaz et al., 2012). It could be argued that it is worthwhile to pursue the merging regardless, as the merged product still performs well in relation to other products, while reducing the uncertainty due to the involvement of more data sources (Yilmaz et al., 2012). However, for most applications it is probably more convenient to simply use pre-existing products, such as ESA CCI SM. The failure to produce an improved product relative to the

parents, in particular against the best performing parent, SMAP L3E, is likely linked to the relatively poor performances of GLDAS Noah and ASCAT compared to SMAP L3E.

Zheng et al. (2022) raised the possibility of using a triplet consisting of a low level SMAP product, ASCAT, and ERA5-Land, instead of GLDAS Noah. This is suggested for future studies. Considering that ERA5-Land scored significantly better than GLDAS Noah for the in situ measurement-based evaluation, it is expected that a triplet including ERA-Land would perform better than the triplet employed in this study.

It should also be noted that the performance of the TCA merged product was better when using the IDW resampling of ASCAT and SMAP L3E to the coarse GLDAS Noah 0.25° grid rather than nearest neighbour interpolation to the finer 0.1°ERA5-Land grid. This indicates that a higher proportion of the relevant information was kept from the original products using the IDW interpolation, which agrees with Fan Chen et al. (2018) that suggested IDW to be appropriate for ASCAT.

6.3. Limitations of this study and recommendations for future studies

Because of the short time span of ICOS SM measurements, it was not possible to assess product performance for long-term anomalies, which is another important product-property (Gruber et al., 2020). Further, if expanding on the use of ICOS sites as common for SM validation, more research should be conducted on investigating each site's SM representativeness on the grid cell scale. This could for example be done by taking regular field measurements in the areas surrounding each station. Additionally, for more a more comprehensive estimate of SM product performance over Europe, larger coverage over the region is needed, e.g., in Iberia and Eastern Europe. Ideally, statistics on confidence should also be added to increase the reliability of the results (Gruber et al., 2020).

In general, a more comprehensive error-cross correlation analysis should be conducted to guide the selection of suitable triplets in the triplet collocation analysis. Another TCA-related limitation is that the datasets have varying level of coverage, which led to the usage of the t-test based merging scheme and occasional selection of alternate merging method to the TCA-weighted mean, which is assumed to be ideal. Future studies should explore alternative methods to address data coverage variations and determine the most appropriate merging scheme. The TCA requirements of independence also limit the potential of the TC approach and application as many SM products are related.

Practical limitations included the time limitations associated with the brief project period, which resulted in compromises being made regarding areas selected for TCA processing and filtering. For example, the study did not account for potential snow cover or include product-external surface roughness masking, as done in other studies (Kim et al., 2015). Future studies should consider these practical limitations and incorporate relevant factors to improve the overall analysis.

7. Conclusions

In this thesis, eight gridded SM products: SMAP L4, SMAP L3E, SMOS L4, Sentinel-1, ASCAT, ESA CCI SM, ERA5-Land, and GLDAS Noah were evaluated against in situ measurements from the relatively new European measurement network ICOS. The SMAP L4 product was found to have strong correlations with the ICOS measurements and relatively low errors and was potentially the product with best scores overall, together with the reanalysis ERA5-Land. ESA CCI SM, and SMAP L3E also produced high statistical scores, while GLDAS Noah performed poorly in capturing short-term SM anomalies. Overall, SMOS L4, ASCAT, and Sentinel-1 exhibited the poorest performance among the evaluated soil moisture products, receiving the lowest scores in descending order.

The ICOS stations seemed promising for future validation studies as the results were generally in line with previous research. ICOS is expected to be even more suitable in a few years when the temporal coverage is longer than it was for this study. It is recommended to conduct further research on the spatial variability of SM in the areas surrounding each ICOS station to guide the selection of appropriate stations for future studies.

Further, triple collocation analysis was performed with a triplet that included SMAP L3E, GLDAS Noah, and ASCAT. It mainly corroborated the results of the ICOS measurement-based evaluation, with SMAP L3E exhibiting a better signal- to- noise ratio and unbiased root mean square error compared to GLDAS Noah and ASCAT.

Triple collocation analysis was also used to estimate weights for merging of SMAP L3E, ASCAT, and GLDAS Noah. The merged product scored better statistically than ASCAT and GLDAS Noah, but similarly to SMAP L3E. Thus, pursuing the merging of this specific triplet using triple collocation was deemed unnecessary, as it in addition did not outperform a simple arithmetic mean between the three parent datasets. However, there is potential for exploring other combinations of products, and further efforts to achieve an improved TCA-weighted merging approach are encouraged.

References

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 Sensing of Environment, 224*, 289-303. https://doi.org/10.1016/j.rse.2019.02.008

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

Albergel, C., Dutra, E., Munier, S., Calvet, J. C., Munoz-Sabater, J., De Rosnay, P., & Balsamo, G. (2018). ERA-5 and ERA-Interim driven ISBA land surface model simulations: Which one performs better? *Hydrology and Earth System Sciences, 22*(6), 3515-3532. https://doi.org/10.5194/hess-22-3515-2018

Alemohammad, S. H., McColl, K. A., Konings, A. G., Entekhabi, D., & Stoffelen, A. (2015). Characterization of precipitation product errors across the United States using multiplicative triple collocation. *Hydrology and Earth System Sciences, 19*(8), 3489-3503. https://doi.org/10.5194/hess-19-3489-2015

Babaeian, E., Sadeghi, M., Jones, S. B., Montzka, C., Vereecken, H., & Tuller, M. (2019). Ground, Proximal, and Satellite Remote Sensing of Soil Moisture. *Reviews of Geophysics, 57*(2), 530-616. https://doi.org/https://doi.org/10.1029/2018RG000618

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., & Wagner, W. (2019). Toward Global Soil Moisture Monitoring with Sentinel-1: Harnessing Assets and Overcoming Obstacles . *IEEE Transactions on Geoscience and Remote Sensing, 57*(1), 520-539. https://doi.org/10.1109/TGRS.2018.2858004

Beaudoing, H., & Rodell, M. (2020). *GLDAS Noah Land Surface Model L4 3 hourly 0.25 x 0.25 degree V2.1*.

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 . *Hydrology and Earth System Sciences, 25*(1), 17-40. https://doi.org/10.5194/hess-25-17-2021

Bogena, H. R., Schrön, M., Jakobi, J., Ney, P., Zacharias, S., Andreasen, M., Baatz, R., Boorman, D., Duygu, M. B., Eguibar-Galán, M. A., Fersch, B., Franke, T., Geris, J., González Sanchis, M., Kerr, Y., Korf, T., Mengistu, Z., Mialon, A., Nasta, P., … Vereecken, H. (2022). COSMOS-Europe: a European network of cosmic-ray neutron soil moisture sensors. *Earth System Science Data, 14*(3), 1125-1151. https://doi.org/10.5194/essd-14-1125-2022

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 Sensing of Environment, 115*(12), 3390-3408. https://doi.org/10.1016/j.rse.2011.08.003

Brocca, L., Melone, F., Moramarco, T., & Morbidelli, R. (2009). Soil moisture temporal stability over experimental areas in Central Italy. *Geoderma, 148*(3), 364-374. https://doi.org/10.1016/j.geoderma.2008.11.004

Cammalleri, C., Vogt, J. V., Bisselink, B., & De Roo, A. (2017). Comparing soil moisture anomalies from multiple independent sources over different regions across the globe. *Hydrology and Earth System Sciences, 21*(12), 6329-6343. https://doi.org/10.5194/hess-21-6329-2017

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 Sensing of Environment, 214*, 1-13. https://doi.org/10.1016/j.rse.2018.05.008

Chen, X., Zhang, Z., Chen, X., & Shi, P. (2009). The impact of land use and land cover changes on soil moisture and hydraulic conductivity along the karst hillslopes of southwest China. *Environmental Earth Sciences, 59*(4), 811-820. https://doi.org/10.1007/s12665-009-0077-6

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., … Yueh, S. (2017). Validation of SMAP surface soil moisture products with core validation sites. *Remote Sensing of Environment, 191*, 215-231. https://doi.org/10.1016/j.rse.2017.01.021

Copernicus. (2020). *Corine Land Cover (CLC) 2018, Version 2020_20u1*. https://land.copernicus.eu/pan-european/corine-land-cover/clc2018

Copernicus Climate Change Service (C3S). *ERA5-Land hourly data from 1950 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac

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 . *Reviews of Geophysics, 50*(2). https://doi.org/10.1029/2011RG000372

Daly, E., & Porporato, A. (2005). A review of soil moisture dynamics: From rainfall infiltration to ecosystem response [Review]. *Environmental Engineering Science, 22*(1), 9-24. https://doi.org/10.1089/ees.2005.22.9

De Beeck, M. O., Gielen, B., Merbold, L., Ayres, E., Serrano-Ortiz, P., Acosta, M., Pavelka, M., Montagnani, L., Nilsson, M., Klemedtsson, L., Vincke, C., De Ligne, A., Moureaux, C., Marańon-Jimenez, S., Saunders, M., Mereu, S., & Hörtnagl, L. (2018). Soil-meteorological measurements at ICOS monitoring stations in terrestrial ecosystems. *International Agrophysics, 32*(4), 619-631. https://doi.org/10.1515/intag-2017-0041

Dong, J., Lei, F., & Wei, L. (2020). Triple Collocation Based Multi-Source Precipitation Merging. *Frontiers in Water, 2*, Article 1. https://doi.org/10.3389/frwa.2020.00001

Dorigo, W., Himmelbauer, I., Aberer, D., Schremmer, L., Petrakovic, I., Zappa, L., Preimesberger, W., Xaver, A., Annor, F., Ardö, J., Baldocchi, D., Bitelli, M., Blöschl, G., Bogena, H., Brocca, L., Calvet, J. C., Camarero, J. J., Capello, G., Choi, M., … Sabia, R. (2021). The International Soil Moisture Network: Serving Earth system science for over a decade [Review]. *Hydrology and Earth System Sciences, 25*(11), 5749-5804. https://doi.org/10.5194/hess-25-5749-2021

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., … Lecomte, P. (2017). ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment, 203*, 185-215. https://doi.org/10.1016/j.rse.2017.07.001

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 Sensing of Environment, 162*, 380-395. https://doi.org/10.1016/j.rse.2014.07.023

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., De Jeu, R. A. M., & Naeimi, V. (2010). Error characterisation of global active and passive microwave soil moisture datasets . *Hydrology and Earth System Sciences, 14*(12), 2605-2616. https://doi.org/10.5194/hess-14-2605-2010

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Van Oevelen, P., Robock, A., & Jackson, T. (2011). The International Soil Moisture Network: A data hosting facility for global in situ soil moisture measurements . *Hydrology and Earth System Sciences, 15*(5), 1675-1698. https://doi.org/10.5194/hess-15-1675-2011

Edokossi, K., Calabia, A., Jin, S., & Molina, I. (2020). GNSS-Reflectometry and Remote Sensing of Soil Moisture: A Review of Measurement Techniques, Methods, and Applications. *Remote Sensing, 12*(4), 614. https://www.mdpi.com/2072-4292/12/4/614

El Hajj, M., Baghdadi, N., Zribi, M., Rodríguez-Fernández, N., Wigneron, J. P., Al-Yaari, A., Al Bitar, A., Albergel, C., & Calvet, J. C. (2018). Evaluation of SMOS, SMAP, ASCAT and Sentinel-1 soil moisture products at sites in Southwestern France . *Remote Sensing, 10*(4), Article 569. https://doi.org/10.3390/rs10040569

Entekhabi, D., Johnson, J., Kimball, J., Piepmeier, J. R., Koster, R. D., Martin, N., McDonald, K. C., Moghaddam, M., Moran, S., Reichle, R., Shi, J. C., Njoku, E. G., Spencer, M. W., Thurman, S. W., Tsang, L., Van Zyl, J., O'Neill, P. E., Kellogg, K. H., Crow, W. T., … Jackson, T. J. (2010). The Soil Moisture Active Passive (SMAP) Mission [Periodical]. *Proceedings of the IEEE, Proc. IEEE, 98*(5), 704-716. https://doi.org/10.1109/JPROC.2010.2043918

Earth Observation Data Centre for Water Resources Monitoring (EODC) GmbH. (2022). *ESA Climate Change Initiative Plus – Soil Moisture. Product Validation and Intercomparison Report (PVIR) Supporting Product version v07.1.*

Falloon, P., Jones, C. D., Ades, M., & Paul, K. (2011). Direct soil moisture controls of future global soil carbon changes: An important source of uncertainty. *Global Biogeochemical Cycles, 25*(3). https://doi.org/10.1029/2010GB003938

Fan, L., Xing, Z., Lannoy, G. D., Frappart, F., Peng, J., Zeng, J., Li, X., Yang, K., Zhao, T., Shi, J., Ma, H., Wang, M., Liu, X., Yi, C., Ma, M., Tang, X., Wen, J., Chen, X., Wang, C., … Wigneron, J. P. (2022). Evaluation of satellite and reanalysis estimates of surface and root-zone soil moisture in croplands of Jiangsu Province, China. *Remote Sensing of Environment, 282*. https://doi.org/10.1016/j.rse.2022.113283

Fascetti, F., Pierdicca, N., Pulvirenti, L., & Crapolicchio, R. (2014). ASCAT and SMOS soil moisture retrievals: A comparison over Europe and Northern Africa. *13th Specialist Meeting on Microwave Radiometry and Remote Sensing of the Environment, MicroRad 2014 - Proceedings*, Pasadena, California, USA.

Fascetti, F., Pierdicca, N., Pulvirenti, L., Crapolicchio, R., & Sabater, J. M. (2014). Soil moisture comparison through triple and quadruple collocation between: Metop, ERA, SMOS and in-situ data. *International Geoscience and Remote Sensing Symposium (IGARSS),* Quebec City, QC, Canada.

GCOS. (n.d.). *Soil Moisture*. GCOS. Retrieved 2023-01-30 from https://gcos.wmo.int/en/essential-climate-variables/soil-moisture

Gelaro, R., McCarty, W., Suárez, 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, … Zhao, B. (2017). The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate, 30*(14), 5419-5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Grillakis, M. G. (2019). Increase in severe and extreme soil moisture droughts for Europe under climate change. *Science of the Total Environment, 660*, 1245-1255. https://doi.org/10.1016/j.scitotenv.2019.01.001

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., … Wagner, W. (2020). Validation practices for satellite soil moisture retrievals: What are (the) errors? [Review]. *Remote Sensing of Environment, 244*, Article 111806. https://doi.org/10.1016/j.rse.2020.111806

Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W. (2017). Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. *IEEE Transactions on Geoscience and Remote Sensing, 55*(12), 6780-6792. https://doi.org/10.1109/TGRS.2017.2734070

Gruber, A., & Peng, J. (2022). Remote sensing of soil moisture. In *Reference Module in Earth Systems and Environmental Sciences*. Elsevier. https://doi.org/10.1016/B978-0-12-822974-3.00019-7

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 System Science Data, 11*(2), 717-739. https://doi.org/10.5194/essd-11-717-2019

Gruber, A., Su, C. H., Crow, W. T., Zwieback, S., Dorigo, W. A., & Wagner, W. (2016). Estimating error cross-correlations in soil moisture data sets using extended collocation analysis . *Journal of Geophysical Research, 121*(3), 1208-1219. https://doi.org/10.1002/2015JD024027

Gruber, A., Su, C. H., Zwieback, S., Crow, W., Dorigo, W., & Wagner, W. (2016). Recent advances in (soil moisture) triple collocation analysis. *International Journal of Applied Earth Observation and Geoinformation, 45*, 200-211. https://doi.org/10.1016/j.jag.2015.09.002

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., & Chiara, G. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society, 146*(730), 1999-2049. https://doi.org/10.1002/qj.3803

Högström, E., Heim, B., Bartsch, A., Bergstedt, H., & Pointner, G. (2018). Evaluation of a MetOp ASCAT-Derived Surface Soil Moisture Product in Tundra Environments. *Journal of Geophysical Research: Earth Surface, 123*(12), 3190-3205. https://doi.org/10.1029/2018JF004658

Hu, F., Wei, Z., Yang, X., Xie, W., Li, Y., Cui, C., Yang, B., Tao, C., Zhang, W., & Meng, L. (2022). Assessment of SMAP and SMOS soil moisture products using triple collocation method over Inner Mongolia. *Journal of Hydrology: Regional Studies, 40*, 101027. https://doi.org/10.1016/j.ejrh.2022.101027

ICOS. (n.d.). *Ecosystem stations*. ICOS. Retrieved 2023-01-26 from https://www.icos-cp.eu/observations/ecosystem/stations

Jackson, T. J. (1993). III. Measuring surface soil moisture using passive microwave remote sensing . *Hydrological Processes, 7*(2), 139-152. https://doi.org/10.1002/hyp.3360070205

Jung, M., Reichstein, M., Ciais, P., Seneviratne, S. I., Sheffield, J., Goulden, M. L., Bonan, G., Cescatti, A., Chen, J., De Jeu, R., Dolman, A. J., Eugster, W., Gerten, D., Gianelle, D., Gobron, N., Heinke, J., Kimball, J., Law, B. E., Montagnani, L., … Zhang, K. (2010). Recent decline in the global land evapotranspiration trend due to limited moisture supply . *Nature, 467*(7318), 951-954. https://doi.org/10.1038/nature09396

Kerr, Y. H., Gruhier, C., Juglea, S. E., Drinkwater, M. R., Hahne, A., Martin-Neira, M., Mecklenburg, S., Waldteufel, P., Wigneron, J. P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M. J., Font, J., & Reul, N. (2010). The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle [Periodical]. *Proceedings of the IEEE, Proc. IEEE, 98*(5), 666-687. https://doi.org/10.1109/JPROC.2010.2043032

Kim, S., Liu, Y., Johnson, F. M., Parinussa, R. M., & Sharma, A. (2015). A global comparison of alternate AMSR2 soil moisture products: Why do they differ? . *Remote Sensing of Environment, 161*, 43-62. https://doi.org/10.1016/j.rse.2015.02.002

Kim, S., Pham, H. T., Liu, Y. Y., Marshall, L., & Sharma, A. (2021). Improving the Combination of Satellite Soil Moisture Data Sets by Considering Error Cross Correlation: A Comparison Between Triple Collocation (TC) and Extended Double Instrumental Variable (EIVD) Alternatives [Periodical]. *IEEE Transactions on Geoscience and Remote Sensing, Geoscience and Remote Sensing, IEEE Transactions*

on, IEEE Trans. Geosci. Remote Sensing, PP(99), 1-11. https://doi.org/10.1109/TGRS.2020.3032418

Lakhankar, T., Ghedira, H., Temimi, M., Azar, A. E., & Khanbilvardi, R. (2009). Effect of land cover heterogeneity on soil moisture retrieval using active microwave remote sensing data . *Remote Sensing, 1*(2), 80-91. https://doi.org/10.3390/rs1020080

Lamptey, P. N. L., Petropoulos, G. P., & Srivastava, P. K. (2021). Smos L4 downscaled soil moisture product evaluation over a two year - period in a mediterranean setting. In *Advances in Remote Sensing for Natural Resource Monitoring* (pp. 111-131). https://doi.org/10.1002/9781119616016.ch8

Leroux, D. J., Kerr, Y. H., Richaume, P., & Fieuzal, R. (2013). Spatial distribution and possible sources of SMOS errors at the global scale. *Remote Sensing of Environment, 133*, 240-250. https://doi.org/10.1016/j.rse.2013.02.017

LDAS. (n.d.). *GLDAS* Soil Land Surface. Retrieved 2023-04-18 from https://ldas.gsfc.nasa.gov/gldas/soils

Li, X., Zhang, W., Vermeulen, A., Dong, J., & Duan, Z. (2023). Triple collocation-based merging of multi-source gridded evapotranspiration data in the Nordic Region. *Agricultural and Forest Meteorology, 335*, 109451. https://doi.org/10.1016/j.agrformet.2023.109451

Liu, Y., & Yang, Y. (2022). Advances in the Quality of Global Soil Moisture Products: A Review [Review]. *Remote Sensing, 14*(15), Article 3741. https://doi.org/10.3390/rs14153741

Lobell, D. B., & Asner, G. P. (2002). Moisture effects on soil reflectance . *Soil Science Society of America Journal, 66*(3), 722-727. https://doi.org/10.2136/sssaj2002.7220

Louvet, S., Pellarin, T., Al Bitar, A., Cappelaere, B., Galle, S., Grippa, M., Gruhier, C., Kerr, Y., Lebel, T., Mialon, A., Mougin, E., Quantin, G., Richaume, P., & de Rosnay, P. (2015). SMOS soil moisture product evaluation over West-Africa from local to regional scale. *Remote Sensing of Environment, 156*, 383-394. https://doi.org/10.1016/j.rse.2014.10.005

Ma, H., Li, X., Zeng, J., Zhang, X., Dong, J., Chen, N., Fan, L., Sadeghi, M., Frappart, F., Liu, X., Wang, M., Wang, H., Fu, Z., Xing, Z., Ciais, P., & Wigneron, J. P. (2023). An assessment of L-band surface soil moisture products from SMOS and SMAP in the tropical areas. *Remote Sensing of Environment, 284*, Article 113344. https://doi.org/10.1016/j.rse.2022.113344

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 Sensing of Environment, 231*, 111215. https://doi.org/10.1016/j.rse.2019.111215

Marczewski, W., Slominski, J., Slominska, E., Usowicz, B., Usowicz, J., Romanov, S., Maryskevych, O., Nastula, J., & Zawadzki, J. (2010). Strategies for validating and directions for employing SMOS data, in the Cal-Val project SWEX (3275) for wetlands. *Hydrol. Earth Syst. Sci. Discuss., 2010*, 7007-7057. https://doi.org/10.5194/hessd-7-7007-2010

Min, X., Shangguan, Y., Li, D., & Shi, Z. (2022). Improving the fusion of global soil moisture datasets from SMAP, SMOS, ASCAT, and MERRA2 by considering the

non-zero error covariance. *International Journal of Applied Earth Observation and Geoinformation, 113*, 103016. https://doi.org/10.1016/j.jag.2022.103016

Mironov, V., Kerr, Y., Wigneron, J. P., Kosolapova, L., & Demontoux, F. (2013). Temperature-and texture-dependent dielectric model for moist soils at 1.4 GHz. *IEEE Geoscience and Remote Sensing Letters, 10*(3), 419-423, Article 6268319. https://doi.org/10.1109/LGRS.2012.2207878

Mortl, A., Muñoz-Carpena, R., Kaplan, D., & Li, Y. (2011). Calibration of a combined dielectric probe for soil moisture and porewater salinity measurement in organic and mineral coastal wetland soils. *Geoderma, 161*(1), 50-62. https://doi.org/10.1016/j.geoderma.2010.12.007

Moyano, F. E., Manzoni, S., & Chenu, C. (2013). Responses of soil heterotrophic respiration to moisture availability: An exploration of processes and models. *Soil Biology and Biochemistry, 59*, 72-85. https://doi.org/10.1016/j.soilbio.2013.01.002

Mulder, V. L., de Bruin, S., Schaepman, M. E., & Mayr, T. R. (2011). The use of remote sensing in soil and terrain mapping - A review [Review]. *Geoderma, 162*(1-2), 1-19. https://doi.org/10.1016/j.geoderma.2010.12.018

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J. N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications [Data paper]. *Earth System Science Data, 13*(9), 4349-4383. https://doi.org/10.5194/essd-13-4349-2021

Muñoz Sabater, J. (2019). *ERA5-Land hourly data from 1950 to present.* Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.e2161bac

Nguyen, T. M., Walker, J. P., Ye, N., & Kodikara, J. (2023). Use of an L-band radiometer for proximal moisture measurement in road construction. *Transportation Geotechnics, 38*, Article 100876. https://doi.org/10.1016/j.trgeo.2022.100876

O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., Bindlish, R., Chaubell, J., & Collinader, A. (2021). *SMAP Enhanced L3 Radiometer Global and Polar Grid Daily 9 km EASE-Grid Soil Moisture, Version 5.* NASA National Snow and Ice Data Center.

Ochsner, T. E., Cosh, M. H., Cuenca, R. H., Dorigo, W. A., Draper, C. S., Hagimoto, Y., Kerr, Y. H., Larson, K. M., Njoku, E. G., Small, E. E., & Zreda, M. (2013). State of the art in large-scale soil moisture monitoring [Review]. *Soil Science Society of America Journal, 77*(6), 1888-1919. https://doi.org/10.2136/sssaj2013.03.0093

Pablos, M., González-Haro, C., Piles, M., & Portal., G. (2022). *BEC SMOS Soil Moisture Products Description*. https://bec.icm.csic.es/doc/BEC_SMOS_PD_SM_L3v4_L4v6.pdf

Panciera, R. (2009). *Effect of Land Surface Heterogeneity on Satellite Near-Surface Soil Moisture Observations* [PhD, The University of Melbourne].

Peng, J., Tanguy, M., Robinson, E. L., Pinnington, E., Evans, J., Ellis, R., Cooper, E., Hannaford, J., Blyth, E., & Dadson, S. (2021). Estimation and evaluation of high-resolution soil moisture from merged model and Earth observation data in the Great

Britain. *Remote Sensing of Environment, 264*, Article 112610. https://doi.org/10.1016/j.rse.2021.112610

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

Piles, M., Camps, A., Vall-Llossera, M., Corbella, I., Panciera, R., Rudiger, C., Kerr, Y. H., & Walker, J. (2011). Downscaling SMOS-derived soil moisture using MODIS visible/infrared data . *IEEE Transactions on Geoscience and Remote Sensing, 49*(9), 3156-3166, Article 5756664. https://doi.org/10.1109/TGRS.2011.2120615

Portal, G., Jagdhuber, T., Vall-llossera, M., Camps, A., Pablos, M., Entekhabi, D., & Piles, M. (2020). Assessment of multi-scale SMOS and SMAP soil moisture products across the Iberian Peninsula. *Remote Sensing, 12*(3), Article 570. https://doi.org/10.3390/rs12030570

Reichle, R. H., De Lannoy, G., Koster, R. D., Crow, W., Kimball, J. S., Liu, J., & Bechtold, M. (2022a). *SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Geophysical Data, Version 7.* NASA National Snow and Ice Data Center Distributed Active Archive Center.

Reichle, R. H., De Lannoy, G., Koster, R. D., Crow, W., Kimball, J. S., Liu, J., & Bechtold, M. (2022b). *SMAP L4 Global 9 km EASE-Grid Surface and Root Zone Soil Moisture Land Model Constants, Version 7.* National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/KN96XNPZM4EG

Reichle, R. H., De Lannoy, G., Koster, R. D., Crow, W. T., Kimball, J. S., Liu, Q., & Bechtold, M. (2022c). *SMAP L4 Global 3-hourly 9 km EASE-Grid Surface and Root Zone Soil Moisture Analysis Update, Version 7.* NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/LWJ6TF5SZRG3

Reichle, R. H., De Lannoy, G. J. M., Liu, Q., Ardizzone, J. V., Colliander, A., Conaty, A., Crow, W., Jackson, T. J., Jones, L. A., Kimball, J. S., Koster, R. D., Mahanama, S. P., Smith, E. B., Berg, A., Bircher, S., Bosch, D., Caldwell, T. G., Cosh, M., González-Zamora, Á., … Zeng, Y. (2017). Assessment of the SMAP Level-4 Surface and Root-Zone Soil Moisture Product Using In Situ Measurements. *Journal of Hydrometeorology, 18*(10), 2621-2645. https://doi.org/10.1175/JHM-D-17-0063.1

Reynolds, C. A., Jackson, T. J., & Rawls, W. J. (2000). Estimating soil water-holding capacities by linking the Food and Agriculture Organization soil map of the world with global pedon databases and continuous pedotransfer functions. *Water Resources Research, 36*(12), 3653-3662. https://doi.org/10.1029/2000WR900130

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C. J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D., & Toll, D. (2004). The Global Land Data Assimilation System. *Bulletin of the American Meteorological Society, 85*(3), 381-394. https://doi.org/10.1175/BAMS-85-3-381

Sabater, J. M., Rüdiger, C., Calvet, J. C., Fritz, N., Jarlan, L., & Kerr, Y. (2008). Joint assimilation of surface soil moisture and LAI observations into a land surface model.

Agricultural and Forest Meteorology, 148(8-9), 1362-1373. https://doi.org/10.1016/j.agrformet.2008.04.003

Sawada, Y., Koike, T., & Walker, J. P. (2015). A land data assimilation system for simultaneous simulation of soil moisture and vegetation dynamics. *Journal of Geophysical Research, 120*(12), 5910-5930. https://doi.org/10.1002/2014JD022895

Scipal, K., Holmes, T., De Jeu, R., Naeimi, V., & Wagner, W. (2008). A possible solution for the problem of estimating the error structure of global soil moisture data sets. *Geophysical Research Letters, 35*(24), Article L24403. https://doi.org/10.1029/2008GL035599

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture-climate interactions in a changing climate: A review [Review]. *Earth-Science Reviews, 99*(3-4), 125-161. https://doi.org/10.1016/j.earscirev.2010.02.004

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

Silvertown, J. (2004). Plant coexistence and the niche. *Trends in Ecology & Evolution, 19*(11), 605-611. https://doi.org/10.1016/j.tree.2004.09.003

Spatafora, L. R., Vall-Llossera, M., Camps, A., Chaparro, D., Alvalá, R. C. S., & Barbosa, H. (2020). Validation of smos l3 and l4 soil moisture products in the remedhus (Spain) and cemaden (brazil) networks. *Revista Brasileira de Geografia Fisica, 13*(2), 691-712. https://doi.org/10.26848/rbgf.v13.2.p691-712

Stillman, S., & Zeng, X. (2018). Evaluation of SMAP soil moisture relative to five other satellite products using the climate reference network measurements over USA. *IEEE Transactions on Geoscience and Remote Sensing, 56*(11), 6296-6305, Article 8370803. https://doi.org/10.1109/TGRS.2018.2835316

Stoffelen, A. (1998). Error Modeling and Calibration; Towards the true surface wind speed. *J. Geophys. Res., 103*, 7,755-757,766. https://doi.org/10.1029/97JC03180

Tarek, M., Brissette, F. P., & Arsenault, R. (2020). Evaluation of the ERA5 reanalysis as a potential reference dataset for hydrological modelling over North America. *Hydrology and Earth System Sciences, 24*(5), 2527-2544. https://doi.org/10.5194/hess-24-2527-2020

Taylor, R. G., Scanlon, B., Döll, P., Rodell, M., van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J. S., Edmunds, M., Konikow, L., Green, T. R., Chen, J., Taniguchi, M., Bierkens, M. F. P., MacDonald, A., Fan, Y., Maxwell, R. M., Yechieli, Y., … Treidel, H. (2013). Ground water and climate change. *Nature Climate Change, 3*(4), 322-329. https://doi.org/10.1038/nclimate1744

Trugman, A. T., Medvigy, D., Mankin, J. S., & Anderegg, W. R. L. (2018). Soil Moisture Stress as a Major Driver of Carbon Cycle Uncertainty. *Geophysical Research Letters, 45*(13), 6495-6503. https://doi.org/10.1029/2018GL078131

Ulaby, F. T., Dubois, P. C., & van Zyl, J. (1996). Radar mapping of surface soil moisture. *Journal of Hydrology, 184*(1), 57-84. https://doi.org/10.1016/0022-1694(95)02968-0

Vachaud, G., Passerat De Silans, A., Balabanis, P., & Vauclin, M. (1985). Temporal stability of spatially measured soil water probability density function . *Soil Science Society of*

America Journal, 49(4), 822-828. https://doi.org/10.2136/sssaj1985.03615995004900040006x

Venkatesh, B., Lakshman, N., Purandara, B. K., & Reddy, V. B. (2011). Analysis of observed soil moisture patterns under different land covers in Western Ghats, India . *Journal of Hydrology, 397*(3-4), 281-294. https://doi.org/10.1016/j.jhydrol.2010.12.006

Wagner, W., Lemoine, G., & Rott, H. (1999). A method for estimating soil moisture from ERS Scatterometer and soil data . *Remote Sensing of Environment, 70*(2), 191-207. https://doi.org/10.1016/S0034-4257(99)00036-X

Wang, Y., Leng, P., Peng, J., Marzahn, P., & Ludwig, R. (2021). Global assessments of two blended microwave soil moisture products CCI and SMOPS with in-situ measurements and reanalysis data. *International Journal of Applied Earth Observation and Geoinformation, 94*, Article 102234. https://doi.org/10.1016/j.jag.2020.102234

Wigneron, J.-P., Schmugge, T., Chanzy, A., Calvet, J.-C., & Kerr, Y. (1998). Use of passive microwave remote sensing to monitor soil moisture. *Agronomie, 18*(1), 27-43. https://doi.org/10.1051/agro:19980102

Wu, X., Lu, G., Wu, Z., He, H., Scanlon, T., & Dorigo, W. (2020). Triple collocation-based assessment of satellite soil moisture products with in situ measurements in China: Understanding the error sources. *Remote Sensing, 12*(14), Article 2275. https://doi.org/10.3390/rs12142275

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 Sensing of Environment, 254*, Article 112248. https://doi.org/10.1016/j.rse.2020.112248

Yilmaz, M. T., Crow, W. T., Anderson, M. C., & Hain, C. (2012). An objective methodology for merging satellite- and model-based soil moisture products. *Water Resources Research, 48*(11), Article W11502. https://doi.org/10.1029/2011WR011682

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 Sensing of Environment, 163*, 91-110. https://doi.org/10.1016/j.rse.2015.03.008

Zheng, J., Zhao, T., Lü, H., Shi, J., Cosh, M. H., Ji, D., Jiang, L., Cui, Q., Lu, H., Yang, K., Wigneron, J. P., Li, X., Zhu, Y., Hu, L., Peng, Z., Zeng, Y., Wang, X., & Kang, C. S. (2022). Assessment of 24 soil moisture datasets using a new in situ network in the Shandian River Basin of China. *Remote Sensing of Environment, 271*, Article 112891. https://doi.org/10.1016/j.rse.2022.112891

Zreda, M., Desilets, D., Ferré, T. P. A., & Scott, R. L. (2008). Measuring soil moisture content non-invasively at intermediate spatial scale using cosmic-ray neutrons. *Geophysical Research Letters, 35*(21), Article L21402. https://doi.org/10.1029/2008GL035655

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., & Rosolem, R. (2012). COSMOS: The cosmic-ray soil moisture observing system. *Hydrology and Earth System Sciences, 16*(11), 4079-4099. https://doi.org/10.5194/hess-16-4079-2012

Appendices

Table A1. The ICOS stations used for evaluating the gridded products with location, SM retrieval start date, ICOS site type, CORINE land cover type, and proportion of the CORINE land cover type for boxes at 1 km and 9 km size around the site.

Site-ID	Lon.	Lat.	SM start date	ICOS site type	CORINE land cover	Prop. 9 km	Prop. 1 km	Dataset
BE-Bra	4.5198	51.3076	2020-10-26	evergreen needleleaf forests	mixed forests	0.21	0.71	Janssens et al. (2022)
BE-Lcr	3.8504	51.1122	2019-01-01	deciduous broadleaf forests	complex cultivation patterns	0.38	0.66	De Meulder et al. (2022)
BE-Lon	4.7462	50.5516	2018-08-03	croplands	non-irrigated arable land	0.55	0.89	Dumont et al. (2022)
BE-Maa	5.6319	50.9799	2020-05-04	closed shrublands	moors and heathland	0.07	0.97	Roland et al. (2022)
BE-Vie	5.9981	50.3050	2020-10-26	mixed forests	mixed forest	0.49	0.84	Vincke et al. (2022)
CH-Dav	9.8559	46.8153	2019-11-18	evergreen needleleaf forests	coniferous forest	0.27	0.61	Gharun et al. (2022)
CZ-Wet	14.7704	49.0247	2020-01-01	permanent wetlands	inland marshes	0.03	0.39	Dusek et al. (2022)
DE-Geb	10.9146	51.0997	2020-10-26	croplands	non-irrigated arable land	0.88	1	Brümmer et al. (2022)
DE-Gri	13.5126	50.9500	2017-01-01	grasslands	pastures	0.09	0.73	Bernhofer et al. (2022a)
DE-Hai	10.4521	51.0794	2019-01-01	deciduous broadleaf forests	broad-leaved forest	0.66	1	Knohl et al. (2022)
DE-HoH	11.2224	52.0866	2019-01-17	deciduous broadleaf forests	broad-leaved forest	0.19	1	Rebmann et al. (2023)
DE-Kli	13.5224	50.8931	2018-05-07	croplands	non-irrigated arable land	0.57	0.76	Bernhofer et al. (2022b)
DE-RuR	6.3041	50.6219	2011-05-13	grasslands	pastures	0.46	0.97	Schmidt et al. (2022a)

DE-RuS	6.4471	50.8659	2019-04-29	croplands	non-irrigated arable land	0.54	1	Schmidt et al. (2022b)
DE-Tha	13.5652	50.9626	2020-10-26	evergreen needleleaf forests	coniferous forest	0.07	0.99	Bernhofer et al. (2023)
DK-Gds	9.3341	56.0737	2021-03-12	evergreen needleleaf forests	coniferous forest	0.50	1	Friborg et al. (2022a)
DK-Skj	8.4048	55.9127	2021-05-11	permanent wetlands	inland marshes	0.09	0.99	Friborg et al. (2022b)
DK-Vng	9.1607	56.0375	2020-10-26	croplands	non-irrigated arable land	0.45	0.82	Friborg et al. (2022c)
FI-Hyy	24.2948	61.8474	2018-11-06	evergreen needleleaf forests	coniferous forest	0.73	0.94	Mammarella et al. (2022)
FI-Ken	24.2430	67.9872	2020-01-01	evergreen needleleaf forests	coniferous forest	0.64	1	Laurila et al. (2022)
FI-Sii	24.1929	61.8327	2019-02-27	permanent wetlands	peat bogs	0.14	0.59	Tuittila et al. (2022)
FI-Var	29.6100	67.7549	2017-01-01	evergreen needleleaf forests	coniferous forest	0.86	1	Kolari et al. (2022)
FR-Aur	1.1061	43.5497	2019-01-01	croplands	non-irrigated arable land	0.70	0.84	Tallec et al. (2022)
FR-Bil	-0.9561	44.4937	2019-10-28	evergreen needleleaf forests	coniferous forest	0.38	0.53	Loustau et al. (2023)
FR-EM2	3.0207	49.8721	2018-12-31	croplands	non-irrigated arable land	0.87	0.69	Leonard et al. (2022)
FR-Fon	2.7801	48.4764	2019-10-28	deciduous broadleaf forests	broad-leaved forest	0.37	1	Berveiller et al. (2022)
FR-Lam	1.2379	43.4964	2020-10-26	croplands	non-irrigated arable land	0.61	0.84	Brut et al. (2022)
FR-LGt	2.2841	47.3229	2017-01-01	permanent wetlands	inland marshes	0.00	0.23	Jacotot et al. (2022)

FR-Mej	-1.7964	48.1184	2019-02-05	grasslands	complex cultivation patterns	0.23	0.81	Flechard et al. (2022)
FR-Tou	1.3747	43.5729	2018-01-01	grasslands	complex cultivation patterns	0.04	0.59	Calvet et al. (2022)
IT-BFt	10.7420	45.1978	2019-01-01	deciduous broadleaf forests	broad-leaved forest	0.04	0.95	Gerosa et al. (2022)
IT-Lsn	12.7503	45.7405	2016-01-01	open shrublands	vineyards	0.27	0.78	Pitacco et al. (2022)
IT-SR2	10.2909	43.7320	2019-10-28	evergreen needleleaf forests	coniferous forest	0.10	0.79	Arriga et al. (2023)
IT-Tor	7.5781	45.8444	2016-01-01	grasslands	natural grasslands	0.02	0.34	Cremonese et al. (2022)
SE-Deg	19.5565	64.1820	2019-11-04	permanent wetlands	peat bogs	0.05	0.74	Nilsson et al. (2022)
SE-Htm	13.4190	56.0976	2018-04-19	evergreen needleleaf forests	coniferous forest	0.55	0.98	Heliasz et al. (2022)
SE-Nor	17.4795	60.0865	2018-11-06	evergreen needleleaf forests	coniferous forest	0.65	0.93	Mölder et al. (2022)
SE-Svb	19.7745	64.2561	2019-04-30	evergreen needleleaf forests	coniferous forest	0.85	1	Peichl et al. (2022)

Table A1 references

Arriga, N., Goded, I., Dell'Acqua, A., & Matteucci, M. (2023). *ETC L2 Meteo, San Rossore 2, 2018-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/uaXpUgcEYSvIR7sTFakvJ-k0

Bernhofer, C., Eichelmann, U., Grünwald, T., Hehn, M., Moderow, U., & Prasse, H. (2022a). *ETC L2 ARCHIVE, Grillenburg, 2016-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/QQ0wFBD2m8WhWohhMzqehgql

Bernhofer, C., Eichelmann, U., Grünwald, T., Hehn, M., Moderow, U., & Prasse, H. (2022b). *ETC L2 ARCHIVE, Klingenberg, 2017-12-31–2021-12-31*. Ecosystem Thematic Centre. https://hdl.handle.net/11676/DVG0EGHKcHwTqHEPjp8kk27y

Bernhofer, C., Eichelmann, U., Grünwald, T., Hehn, M., Mauder, M., Moderow, U., & Prasse, H. (2023). *ETC L2 Meteo, Tharandt, 2019-12-31–2022-10-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/kkx9RhfbPkfC8-Hd3bRE_F6C

Berveiller, D., Dufrêne, E., Delpierre, N., Morfin, A., Clotilde, P.-G., Vincent, G., Bazot, S., & Soudani, K. (2022). *ETC L2 Meteo, Fontainebleau-Barbeau, 2018-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/3I3BIuFYCqw7agXNwY_O5cmS

Brut, A., Tallec, T., Granouillac, F., Zawilski, B., Claverie, N., Lemaire, B., & Ceschia, E. (2022). *ETC L2 Meteo, Lamasquere, 2019-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/Pqwg35HZOlNo78rCiWcn5v6L

Brümmer, C., Delorme, J.-P., & Schrader, F. (2022). *ETC L2 Meteo, Gebesee, 2019-12-31–2022-08-31*. Ecosystem Thematic Centre. https://hdl.handle.net/11676/Rke1fU8V1g1FEyisv3f9qrhO

Calvet, J.-C., Maurel, W., & Paci, A. (2022). *ETC L2 Meteo, Toulouse, 2017-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/Lv1CzmVV1c7l1Te2YhHDGKVL

Cremonese, E., Galvagno, M., & Morra di Cella, U. (2022). *ETC L2 Meteo, Torgnon, 2015-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/n5cwUxYPI0yrwrCp34nbA1F3

De Meulder, T., Lefevre, L., Roland, M., Segers, J., & Van Look, J. (2022). *ETC L2 Meteo, Lochristi, 2018-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/tq9zPYUzM5sJSiQIJPPZU5rG

Dumont, B., Heinesch, B., Bodson, B., Bogaerts, G., Chopin, H., De Ligne, A., Demoulin, L., Douxfils, B., Engelmann, T., Faurès, A., Longdoz, B., Manise, T., Orgun, A., Piret, A., & Thyrion, T. (2022). *ETC L2 ARCHIVE, Lonzee, 2016-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/aXMAXzSREkyzYrYSWVr4dfzZ

Dusek, J., Kivalov, S., Pavelka, M., Šigut, L., Jocher, G., Czerný, R., Staník, K., Stellner, S., & Trusina, J. (2022). *ETC L2 Meteo, Trebon, 2019-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/L4q-TS_QftExbNq9B6Xovt42

Flechard, C., Fauvel, Y., & Hamon, Y. (2022). *ETC L2 Meteo, Mejusseaume, 2018-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/9O7kkIlNTkTG8kYjO0dGk8Wq

Friborg, T., Jensen, R., & Larmanou, E. (2022a). *ETC L2 ARCHIVE, Gludsted Plantage, 2019-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/aXez8c3uOE2OLuD-o-Y8APo1ed

Friborg, T., Jensen, R., & Larmanou, E. (2022b*). ETC L2 ARCHIVE, Skjern, 2019-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/yfzNxLPEIiDtn9ZiJz4uLb1Q

Friborg, T., Jensen, R., Larmanou, E., & Rasmussen, L. (2022). *ETC L2 Meteo, Voulundgaard, 2019-12-31–2022-08-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/mDyqhNpWiqAOOpg55QaMd4_c

Gerosa, G., Bignotti, L., Finco, A., & Marzuoli, R. (2022). *ETC L2 Meteo, Bosco Fontana, 2018-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/JC0q-2uTJMeHeTJNUSuSqtYz

Gharun, M., Baur, T., Buchmann, N., Etzold, S., Eugster, W., Feigenwinter, I., Gessler, A., Hug, C., Häni, M., Hörtnagl, L., Kumar, S., Liechti, K., Marty, M., Meier, P., Schmitt

Oehler, M., Staudinger, M., Stutz, T., Sutter, F., Thimonier Rickenmann, A., … Zweifel, R. (2022). *ETC L2 ARCHIVE, Davos, 2018-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/EzUhZI2nT3Hnop-m5xduRzSw

Heliasz, M., Kljun, N., Biermann, T., Holst, J., Holst, T., Linderson, M.-L., Mölder, M., & Rinne, J. (2022). *ETC L2 Meteo, Hyltemossa, 2017-12-31–2022-08-31*. Ecosystem Thematic Centre. https://hdl.handle.net/11676/bswp_irhm0hhFn09Q3_Uj3mT

Jacotot, A., Paroissien, J.-B., & Perdereau, L. (2022). *ETC L2 Meteo, La Guette, 2016-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/Kz343rOR8d-9UP18mlCaxjOQ

Janssens, I., De Meulder, T., Lefevre, L., Roland, M., Segers, J., & Van Look, J. (2022). *ETC L2 Meteo, Brasschaat, 2019-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/NJ5ZmcK3eo3-bKhLz_GpeQ7r

Knohl, A., Klosterhalfen, A., Markwitz, C., Siebicke, L., & Tiedemann, F. (2022). *ETC L2 Meteo, Hainich, 2018-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/_lH2WA5i_4sAdm-0g4Bue0jB

Kolari, P., Aalto, P., Keronen, P., Kulmala, L., & Sahoo, G. (2022). *ETC L2 Meteo, Varrio, 2016-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/ep9peFD4fyWn0nKJ0E6rIl58

Laurila, T., Aurela, M., Hatakka, J., & Rainne, J. (2022). *ETC L2 ARCHIVE, Kenttarova, 2019-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/a3tAkRKa4b7jNx-CE9fGxjuG

Leonard, J., Bornet, F., & Grehan, E. (2022). ETC L2 Meteo, Estrees-Mons A28, *2016-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/qxYIVA2pE1F6wuWpIHrBiqQg

Loustau, D., Aluome, C., Chipeaux, C., Denou, J.-L., Depuydt, J., Garrigou, C., Kruszewski, A., & Lafont, S. (2023). *ETC L2 Meteo, Bilos, 2018-12-31–2022-10-31*. Ecosystem Thematic Centre. https://hdl.handle.net/11676/mTL6zPfGu4VDKXWJW-8bta25

Mammarella, I., Aalto, J., Back, J., Kolari, P., Laakso, H., Levula, J., Matilainen, T., Pihlatie, M., Pumpanen, J., Taipale, R., & Vesala, T. (2022). *ETC L2 Meteo, Hyytiala, 2017-12-31–2022-10-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/3AOes0w1K-0sOeUU1He1QVMx

Mölder, M., Kljun, N., Lehner, I., Båth, A., Holst, J., & Linderson, M.-L. (2022). *ETC L2 Meteo, Norunda, 2017-12-31–2022-08-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/CjcspM-TQbJRxOs8aMKpd2wi

Nilsson, M., Peichl, M., Marklund, P., De Simon, G., Smith, P., Löfvenius, P., Dignam, R., Holst, J., Mölder, M., Andersson, T., Larmanou, E., Linderson, M.-L., & Lindgren, K. (2022). *ETC L2 Meteo, Degero, 2018-12-31–2022-08-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/fnfdFvLChAlSbitlFNv3vHLB

Peichl, M., Nilsson, M., Smith, P., Marklund, P., De Simon, G., Löfvenius, P., Dignam, R., Holst, J., Mölder, M., Andersson, T., Larmanou, E., Linderson, M.-L., Lindgren, K., Ottosson-Löfvenius, M., Tülp, H., & Öquist, M. (2022). *ETC L2 Meteo, Svartberget, 2018-12-31–2022-08-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/j1N-Y0zZM_MClImO0Q8pLxoQ

Pitacco, A., Tezza, L., Meggio, F., Peressotti, A., & Vendrame, N. (2022). *ETC L2 Meteo, Lison, 2015-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/jDJwh-ib9WqawcnbYG-jtgOs

Rebmann, C., Dienstbach, L., Schmidt, P., Wiesen, R., Meis, J., García Quirós, I., Gimper, S., & Paasch, S. (2023). *ETC L2 ARCHIVE, Hohes Holz, 2018-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/cuvG4mwumg9zZzhQ6GIFr49H

Roland, M., De Meulder, T., Lefevre, L., Segers, J., & Van Look, J. (2022). *ETC L2 ARCHIVE, Maasmechelen, 2019-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/u7oh-s9uhOHWzjIAyYGigNWa

Schmidt, M., Becker, N., Dolfus, D., Graf, A., Kettler, M., & Mattes, J. (2022a). *ETC L2 ARCHIVE, Rollesbroich, 2010-12-31–2021-12-31*. Ecosystem Thematic Centre. https://hdl.handle.net/11676/zGXCWDo6GQQjxiYlbA47Ddnp

Schmidt, M., Becker, N., Dolfus, D., Esser, O., Graf, A., Kettler, M., & Mattes, J. (2022b). *ETC L2 ARCHIVE, Selhausen Juelich, 2018-12-31–2022-09-30*. Ecosystem Thematic Centre. https://hdl.handle.net/11676/GWgWkZJCp6nvA6m00ZE8wVYg

Tallec, T., Ceschia, E., Granouillac, F., Claverie, N., Zawilski, B., Brut, A., Lemaire, B., & Gibrin, H. (2022). *ETC L2 Meteo, Aurade, 2018-12-31–2021-12-31.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/czwXRHCz3VqEXX4Eb6mVlhpb

Tuittila, E.-S., Kolari, P., Korrensalo, A., Laakso, H., Levula, J., Mammarella, I., Matilainen, T., Männistö, E., Taipale, R., & Vesala, T. (2022). *ETC L2 Meteo, Siikaneva, 2016-12-31–2022-10-31*. Ecosystem Thematic Centre. https://hdl.handle.net/11676/MyW0S6ayzrpl-tLa1B70-JKP

Vincke, C., Bogaerts, G., Chebbi, W., Chopin, H., Demoulin, L., Douxfils, B., Engelmann, T., Faurès, A., Heinesch, B., Manise, T., Orgun, A., Piret, A., & Thyrion, T. (2022). *ETC L2 ARCHIVE, Vielsalm, 2019-12-31–2022-09-30.* Ecosystem Thematic Centre. https://hdl.handle.net/11676/kzDSeZDZdSQ7TaeJWw1Affll

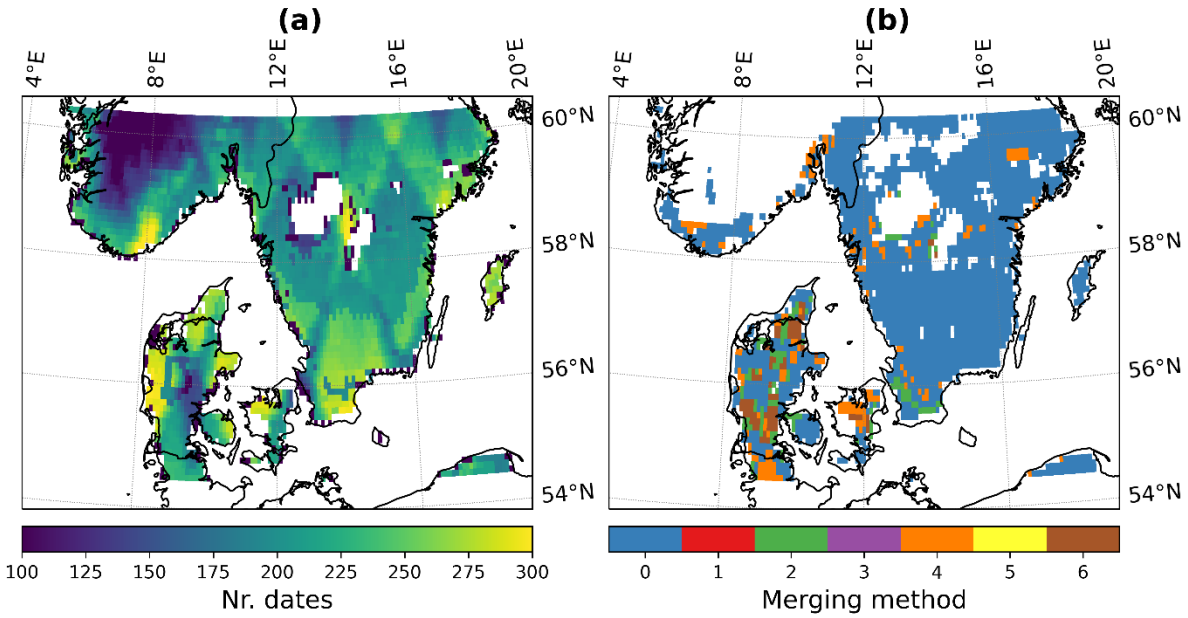

Figure A1. The number of collocated dates between parent products included in the TCA analysis for each pixel (a), and (b) the merging method for each pixel based on the merging scheme: 0: TCA-weighted mean, 1:Only GLDAS Noah, 2: Only SMAP L3E, 3: Only ASCAT, 4: Mean of GLDAS Noah and SMAP L3E, 5: Mean of GLDAS Noah and ASCAT, 6: Mean of SMAP L3E and ASCAT.