Soil moisture retrieval at 1-km resolution making a synergistic use of Sentinel-1/2/3 data

Remi Madelon[1], Nemesio J. Rodríguez-Fernández[1], Hassan Bazzi[2], Nicolas Baghdadi[2],
Clement Albergel[3], Wouter Dorigo[4], and Mehrez Zribi[1]

[1]CESBIO (Université de Toulouse, CNES, CNRS, IRD, INRAE), 18 av. Edouard Belin, bpi 2801, 31401 Toulouse, France
[2]TETIS, INRAE, Université de Montpellier, 34090 Montpellier, France
[3]European Space Agency Climate Office, ECSAT, Harwell Campus, Oxfordshire, Didcot OX11 0FD, United Kingdom
[4]TU Wien, Vienna, Austria

Correspondence: Nemesio J. Rodríguez-Fernández (nemesio.rodriguez@cesbio.cnes.fr)

Abstract.

High-resolution (HR) surface soil moisture (SM) observations are important for applications in hydrology and agriculture, among other purposes. For instance, the S^2MP (Sentinel-1/Sentinel-2 derived Soil Moisture Product) algorithm was designed to retrieve surface SM at agricultural plot scale using simultaneously Sentinel-1 (S1) backscatter coefficients and Sentinel-2 (S2) NDVI (Normalized Difference Vegetation Index) as inputs to a neural network trained with Water Cloud Model simulations. However, for many applications, including future climate impact assessment at regional level, a resolution of 1 km is already a significant improvement with respect to most of the publicly available SM data sets, which have resolutions of about 25 km. Therefore, in this study, the S^2MP algorithm was adapted to work at a resolution of 1 km and extended from croplands (cereals and grasslands) to herbaceous vegetation types. A target resolution of 1 km also allows to explore the use of NDVI derived from Sentinel-3 (S3) instead of S2. The algorithm improvements are evaluated both over Europe and other regions of the globe, for which S1 coverage is poorer.

Two sets of SM maps at 1-km resolution were produced with S^2MP over six regions of $\sim 10^4$ km^2 in the southwest and southeast of France, Spain, Tunisia, North America, as well as Australia from 2017 to 2019. The first set of maps was derived from the combination of S1 and S2 data (S1+S2 maps), while the second one was derived from the combination of S1 and S3 (S1+S3 maps). S1+S2 and S1+S3 SM maps were compared to each other and to those of the 1-km resolution Copernicus Global Land Service (CGLS) SM and Soil Water Index (SWI) data sets as well as to the SMAP+S1 product. The S^2MP S1+S2 and S1+S3 SM maps are in very good agreement in terms of correlation ($R \geq 0.9$), bias (≤ 0.04 m^3 m^{-3}) and standard deviation of the difference ($STDD \leq 0.03$ m^3 m^{-3}) over the 6 domains investigated in this study. The S^2MP maps are well correlated to those from the CGLS SM product ($R \sim 0.7$-0.8), but the correlations with respect to the other HR maps (CGLS SWI and SMAP+S1) drop significantly over many areas of the 6 domains investigated in this study. In addition, higher correlations between the HR maps were found over croplands and when the 1-km pixels have a very homogeneous land cover. The bias in between the different maps was found to be significant over some areas of the six domains, reaching values of ± 0.1 m^3 m^{-3}. The S1+S2 maps show a lower STDD with respect to CGLS maps (≤ 0.06 m^3 m^{-3}) than with respect to the SMAP+S1 maps (≤ 0.1 m^3 m^{-3}) for all the 6 domains.

25   Finally, all the HR data sets were also compared to in situ measurements from 5 networks across 5 countries along with
coarse resolution (CR) SM products from SMAP, SMOS and the ESA Climate Change Initiative (CCI). While all the CR and
HR products show different bias and STDD, the HR products show lower correlations than the CR ones with respect to in situ
measurements.

1   Introduction

30   Surface soil moisture (SM) plays a key role in the Earth water cycle as it affects many hydrological processes such as infiltration,
runoff, evaporation and precipitation (Koster et al., 2004). SM is used to constrain numerical weather prediction (NWP) models
via data assimilation (De Rosnay et al., 2013; de Rosnay et al., 2014; Rodríguez-Fernández et al., 2019). It is also used for crop
yields forecasting, food security and agriculture management (Guerif and Duke, 2000). SM was identified as one of the 50
"essential climate variables" (ECVs) by the Global Climate Observing System in the context of the United Nations Framework

35   Convention on Climate Change (Plummer et al., 2017; GCOS, 2021). Building long time series of SM is crucial for climate
applications, and this is the goal of projects such as the European Space Agency's Climate Change Initiative (ESA CCI) for
SM (Gruber et al., 2019). Both active and passive microwave sensors can be used to estimate SM at coarse resolutions ($\sim$
25-40 km) including the active Advanced Scatterometer (ASCAT, Vreugdenhil et al., 2016), the passive Advanced Microwave
Scanning Radiometer 2 (AMSR2, Kim et al., 2015; Imaoka et al., 2000) and the two sensors that have specifically been

40   designed to measure SM at L-band: the Soil Moisture and Ocean Salinity (SMOS, Kerr et al., 2012) and Soil Moisture Active
Passive (SMAP, Entekhabi et al., 2010). However, despite the actual availability of these SM products they do not match
the requirements of a number of applications. Peng et al. (2020) have discussed a road-map and requirements for future SM
products. An optimal spatial resolution for data assimilation into NWP models and reanalysis would be 5-10 km (global models
are already running with resolutions better than 10 km, see for instance Muñoz-Sabater et al., 2021). The evaluation of climate

45   models and the assessment of climate change impacts at regional level would also benefit from a higher resolution than that of
the current generation of coarse resolution sensors. In addition, other applications in hydrology, agriculture and risk assessment
require even higher resolutions of $\sim$ 1 km (Massari et al., 2021).

Downscaling the coarse scale resolution data by merging them with higher resolution data is a possibility to achieve high
resolutions SM data sets. For example, it can be done using visible/infrared data (Merlin et al., 2012) or using Synthetic

50   Aperture Radar (SAR, Tomer et al., 2016; Das et al., 2019). SAR observations alone have also been tested to estimate SM using
different frequencies and instruments such as RadarSAT, ALOS-L or TerraSAR-X. Radar signal is sensitive to the dielectric
constant linked to SM but also to surface geometry (including roughness) and vegetation water content and structure (Ulaby
et al., 1986). Different inversion algorithms have been proposed considering principally three techniques: change detection
algorithms (Wagner et al., 1999; Balenzano et al., 2010; Bauer-Marschallinger et al., 2018), direct inversion of physical or

55   empirical models (Moran et al., 2000; Srivastava et al., 2009; Pierdicca et al., 2010; Hajj et al., 2014; Bousbih et al., 2017;
Şekertekin et al., 2018) and finally machine learning methods (El Hajj et al., 2017).

With the successive launches of the C-band SARs onboard Sentinel-1 A (S1-A, 2014) and Sentinel-1 B (S1-B, 2016), SM can be estimated at high spatial resolution and with a revisit time better than 6 days over Europe. Three operational high resolution (HR) SM data sets using S1 exist such as the S1 1-km SM and Soil Water Index (SWI) products from the Copernicus Global

60 Land Service (CGLS, Bauer-Marschallinger et al., 2018, 2019) and the SMAP+S1 downscaled product (Das et al., 2020) with resolutions of 1 or 3 km. SM estimates at a very high resolution (plot scale) at some locations in Europe (as well as in Lebannon and Morocco) are also distributed by the French continental surfaces data center (THEIA, https://www.theia-land.fr). In contrast to the CGLS data set, the THEIA SM data set is obtained making a synergistic use of S1 and Sentinel-2 (S2) using the S^2MP algorithm (El Hajj et al., 2017). This data set has been evaluated against in situ data in comparison to SMAP,

65 SMOS and ASCAT coarse resolution data sets (El Hajj et al., 2018) and with respect to the CGLS S1 SM data set (Bazzi et al., 2019), both in the south of France. The S^2MP SM showed the lowest unbiased root mean squared errors with respect to in situ measurements but also a moderate correlation, lower than that obtained for SMAP and ASCAT data sets (El Hajj et al., 2018). In this region, the S^2MP SM showed better performances with respect to in situ measurements than the CGLS SM for the classical metrics (Bazzi et al., 2019).

70 Taking into account the importance of having accurate HR SM data sets, a more comprehensive evaluation of existing products must be done over different types of surfaces and regions. It is worth noting that even if the current temporal coverage of the S1 data is relatively short, it is also essential to evaluate their potential future use for climatic applications. For this and other applications mentioned above a resolution of 1-km is a breakthrough with respect to existing data sets. Moreover, this target resolution also makes possible to use Sentinel-3 (S3) NDVI as input to the S^2MP algorithm in addition to S1 data.

75 The S1 spatial coverage and revisit time are optimized over Europe and the CGLS products only provided over Europe and Mediterranean region. However, it is interesting to evaluate the potential of Sentinel based products over other regions of the globe. Hence, six domains of $\sim 10^4$ km^2 were studied in the southwest and southeast of France, Tunisia, North America, Spain and Australia. The S^2MP S1+S2 and S1+S3 maps were compared to those of the SMAP+S1 product as well as those of the CGLS SM and SWI data sets. The comparison was carried out over the 6 study regions on a per pixel basis and the results were

80 analysed according to pixel homogeneity for areas covered by croplands and herbaceous vegetation. In addition, the HR time series were evaluated against in situ data along with those of coarser resolution SM data sets from SMAP, SMOS, and ESA CCI.

The paper is structured as follows. Section 2 presents the different remotely sensed and ground based data that are used in this study. Section 3 describes the methodology used. Section 4 shows the results of the comparisons of the HR maps as well

85 the assessment against in situ measurements. Section 5 discusses the interest of the S^2MP algorithm. Section 6 draws the conclusions of the study.

2 Data

2.1 High resolution soil moisture products

2.1.1 S^2MP

90 The S^2MP (Sentinel-1/Sentinel-2 derived Soil Moisture Product) algorithm (El Hajj et al., 2017) estimates surface SM at the scale to agricultural plots using synergistically radar signal and optical images from S1 and S2 satellites as inputs to a neural network. The neural network is trained using a synthetic database gathering *(i)* SAR C-band backscatter coefficients in the VV polarization *(ii)* incidence angles (from 20 to 45 degrees), and *(iii)* NDVI (Normalized Difference Vegetation Index) as input and SM examples as target. This synthetic database is built using a Water Cloud Model (Baghdadi et al., 2017) combined with

95 an Integral Equation Model (Baghdadi et al., 2006, 2011) that are specially modified and optimized for this application.

In the framework of this study, the trained neural networks were applied to a real database gathering SAR and NDVI measurements computed from S1 and S2 over croplands and herbaceous vegetation areas at 100-m resolution using the CGLS land cover map. It does not consider retrievals for other types of land cover. Afterwards, the data were averaged at 1-km resolution to provide HR SM maps (hereafter S^2MP_{S1S2}) over the six $\sim 10^4$ km^2 regions of the study. The S^2MP algorithm

100 was also adapted to work directly at 1-km resolution using NDVI from S3 instead of S2 to produce maps for the same six regions (hereafter S^2MP_{S1S3}, see Sect. 3.1).

2.1.2 SMAP/Sentinel-1 L2

Soil Moisture Active Passive (SMAP) was launched on 31 January 2015 by NASA (National Aeronautics and Space Administration). It carries on board a passive radiometer operating at 1.4 GHz (L-band), and a synthetic-aperture radar (active

105 instrument) operating at 1.2 GHz. The respective spatial resolutions of the two instruments are 40 km and 1 to 3 km, however the radar stopped working a few months after launch. SMAP provides passive measurements of the land surface SM in vertical and horizontal polarizations at a fixed incidence angle of 40 degrees (Entekhabi et al., 2014). SMAP ascending and descending orbits cross the equator at 6:00 pm and 6:00 am respectively, and the maximum revisit period is 3 days.

The SMAP+S1 L2 V001 product (hereafter $SMAPS1$) provides SM at 1-km and 3-km resolution that are estimated using

110 the SMAP Enhanced L3 V004 Half-Orbit at 9-km resolution and Copernicus S1A and S1B C-band SAR data (Das et al., 2020). Brightness temperatures from SMAP are disaggregated on the 1-km and 3-km EASE-Grid by using the S1 radar backscatter data and HR SM retrievals are obtained using the SMAP Active-Passive algorithm. Data are split in two different sets. The first set only uses the closest data in time from SMAP descending orbits to spatially match up with the S1 scene. While the second set uses the closest data in time between descending (6:00 am) and ascending (6:00 pm) orbits. The second data set at 1-km

115 resolution was considered in this study.

2.1.3 Copernicus Global Land Service products

The Copernicus Global Land Service S1 Surface Soil Moisture product (hereafter $CoperSSM$) is retrieved from the S1 radar backscatter images estimations over the European continent at 1-km resolution (Bauer-Marschallinger et al., 2019). The images are acquired at C-band SAR in VV polarization, and the retrieval approach is based on a change detection algorithm (Bauer-Marschallinger et al., 2018). Changes observed in the C-band SAR backscatter coefficient are interpreted as changes in the SM values, whereas other surface properties such as the geometry, surface roughness and vegetation cover are interpreted as static in time for each pixel. The algorithm provides local relative SM values in percentages ranging between 0% and 100% except in the case of extremely dry conditions, frozen soil, snow-covered soil and flooding. The data are sampled at 1-km resolution from 11°W to 50°E and from 35°N to 72°N.

The Copernicus Global Land Service S1 Soil Water Index product (hereafter $CoperSWI$) is derived from a fusion of surface SM observations from S1 C-band SAR and Metop ASCAT sensors (Bauer-Marschallinger et al., 2018). It uses a two-layer water balance model that is adapted to use a recursive formulation and does not account for soil texture. A Surface State Flag (SSF) that indicates frozen/unfrozen/melting state of the surface, depending on the temperature, is used to identify SM values under non-frozen conditions to be used for the SWI calculation. SWI and quality flag values are calculated based on a phenomenological formulation that depends on a parameter T. An increased T value is either due to an increase in reservoir depth or a decreased pseudo-diffusivity coefficient. This means that, for a fixed pseudo-diffusivity constant, an increased T value represents a deeper soil layer (Paulik et al., 2014). SWI estimations for eight different T values are provided within the product. Previous evaluations of SWI data by Paulik et al. (2014) and Albergel et al. (2008) showed that the best agreement with in situ measured surface SM is usually obtained with T values in the range of 5-10, therefore SWI data with $T = 5$ were used in this study.

2.2 Coarse resolution soil moisture products

2.2.1 SMAP products

In addition to the $SMAPS1$ product (see Sect 2.1), several coarse resolution data sets from SMAP were used for comparison purposes:

(*i*) the SMAP L3 V6 SM product (hereafter $SMAPL3$). It is a daily gridded composite of the SMAP L2 V5 SM files (O'Neill et al., 2018, 2019). Only SM retrievals derived from L1C brightness temperatures (Chan et al., 2018) using the Single Channel Algorithms V-polarization (Entekhabi et al., 2010) were considered. SMAP L3 data are sampled at 36-km resolution.

(*ii*) the SMAP Enhanced L3 V1 SM product (hereafter $SMAPL3E$), which is obtained by oversampling the L1C brightness temperatures from 36-km to 9-km resolution using the Backus-Gilbert interpolation algorithm. Only SM retrievals derived using the Single Channel Algorithms V-polarization were considered.

2.2.2 SMOS products

The Soil Moisture and Ocean Salinity (SMOS) mission is part of the Earth Explorer program from ESA (European Space Agency), with contributions from CNES (Centre National d'Etudes Spatiales, France) and CDTI (Centro Para el Desarrollo Tecnológico Industrial, Spain). The satellite was launched on 2 November 2009 to measure globally and frequently the surface

150 SM over land and sea surface salinity over the oceans (Kerr et al., 2001), with a maximum revisit time of 3 days and Equator crossings at 6:00 am and 6:00 pm for ascending and descending overpasses, respectively. The SMOS mission relies on a two dimensional interferometric instrument operating at L-band (21 cm, 1.4 GHz) able to measure brightness temperatures in both vertical and horizontal polarizations, at incidence angles from 0 to 60 degrees and with a spatial resolution of 25-50 depending on the position on the field of view (43 km on average). The following SMOS SM data sets were used in this study:

155 (*i*) The CATDS SMOS L3 V3 SM product (hereafter $SMOSL3$), which is a multi-orbit SM product provided by the Centre Aval de Traitement des Données (CATDS) with a grid resolution of 25 km (Al Bitar et al., 2017). The SM retrieval process is based on the algorithm used for the SMOS L2 product (Kerr et al., 2012), and focuses on the iterative minimization of the difference between a forward model and the brightness temperatures. The model uses the τ - ω (optical depth-single scattering albedo) approach to take into account the effect of vegetation. In contrast to the L2 algorithim, the L3 one uses simultaneously

160 three orbits within a one week period to better constrain the SM and VOD (vegetation optical depth) estimations. The data are sampled at 25-km resolution.

(*ii*) The ESA SMOS Near Neal Time (NRT) Neural Network SM product (hereafter $SMOSNRT$), provided on the icosahedral equal area grid (ISEA4H9) with 15-km resolution (Rodríguez-Fernández et al., 2017). It is designed to provide SM in less than 3.5 h after sensing. The algorithm uses a neural network (NN) trained using SMOS L2 SM data (Kerr et al., 2012).

165 The input data for the NN are SMOS brightness temperatures with incidence angles from 30 to 45 degrees for horizontal and vertical polarizations and soil temperature in the 0–7 cm layer from the European Centre for Medium-Range Weather Forecasts (ECMWF) models.

2.2.3 ESA CCI COMBINED product

In the COMBINED product of ESA SM CCI version 5.2 (hereafter $CCISM$) (Dorigo et al., 2017; Gruber et al., 2019) L2

170 data sets from different active and passive sensors are directly scaled by matching their Cumulative Density Functions (CDF) to that of the GLDAS (Global Land Data Assimilation System, Rodell et al., 2004) Noah land surface model in order to remove relative biases and harmonize their dynamical ranges. In the period of this study, ESA CCI Combined uses the H SAF active sensor products from ASCAT-A and -B (Advanced Scatterometer, Wagner et al., 2013) and the passive sensors AMSR2 (Advanced Microwave Scanning Radiometer 2, Kim et al., 2015; Imaoka et al., 2000), SMAP, and SMOS. SM data from the

175 passive sensors are estimated using the Land Parameter Retrieval Model (LPRM) version 6 (Van der Schalie et al., 2016, 2017). The data are sampled at 25-km resolution.

2.3 Land cover map

The Dynamic Land Cover Map product version 3 provided by CGLS was used to evaluate the different HR and CR data sets. As described above, it is the same land cover map used by the S^2MP algorithm for this study. This product delivers a global land cover map at 100-m resolution covering the period between 2015 to 2019 (Buchhorn et al., 2020) For each year, a land cover map is provided with three different levels of classes: 11 classes at level 1 (all types of forests are considered as an unique land cover class), 13 classes at level 2 (forests are splitted in two land cover classes: open and closed forests) and up to 22 classes at level 3 (all types of open and closed forests are considered). In this study, only the 2019 land cover map at level 1 was considered. Figure 1 shows the 7 land cover classes that are represented in the six regions of study.

2.4 in situ data

The evaluation against in situ measurements was performed using data from the REMEDHUS (Gonzalez-Zamora et al., 2018), SMOSMANIA (Calvet et al., 2007), OZNET (Smith et al., 2012; Young et al., 2008), USCRN (Bell et al., 2013) and the Merguellil networks (Gorrab et al., 2015) that are located within the six regions of this study. All the data, except those from the Merguellil network, were retrieved from the International Soil Moisture Network (ISMN, Dorigo et al., 2011, 2021). Only sensors between 0 and 5 cm depth were considered. In total, 43 ISMN and 6 Merguellil sites were used for the evaluation (Table 1).

3 Methods

3.1 Building S^2MP maps using NDVI derived from Sentinel-2 and Sentinel-3

The S^2MP algorithm for soil moisture estimation requires as input values the SAR backscatter coefficient in VV polarization, the incidence angle of the SAR acquisition and the NDVI derived from optical images. The NDVI could be either derived from S2 or S3 images. In this study, the use of the S2 and S3 NDVI was tested for soil moisture estimation using the S^2MP. In the first version of the S^2MP algorithm, NDVI values required as input for the neural network are derived from S2 optical images. The NDVI was first calculated at 10-m spatial resolution (native resolution of the S2) using atmospherically and ortho-rectified S2 images. To overcome the cloud cover issue present in optical images, a gap filling was performed using the linear interpolation to obtain two cloud-free NDVI images per month (1st and 15th of each month). Over each study site and at each S1 acquisition, the 10-m resolution S1 backscattering signal, incidence angle and S2 NDVI (nearest NDVI date to the S1 date) are averaged for each 100-m masked land cover pixels (cropland and herbaceous vegetation). The SM estimation using the S^2MP algorithm is thus obtained at 100-m spatial resolution. The 100-m SM estimations were then aggregated at 1-km grid scale, by averaging the 100-m SM estimation within each 1-km grid cell, to obtain the 1-km SM product using the same 1-km grid extent of the $CoperSSM$ product. In the second version of the S^2MP algorithm, the S3 NDVI was used instead of the S2 NDVI. In this case, the S1 VV backscattering and the incidence angle at each S1 acquisition were averaged directly at 1-km scale only for croplands and herbaceous vegetation areas inside each 1-km grid cell and the NDVI values were derived directly

from the S3 images at 1-km scale. The SM estimation using the neural network inversion was thus directly applied at 1-km grid scale to obtain 1-km SM estimations.

210 ## 3.2 Data sets comparisons and evaluations

Comparisons between data sets and evaluations against in situ measurements were done from January 2017 to December 2019. Only ascending orbits from SMOS and descending orbits from SMAP within this time period were considered.

Relative indices from $CoperSSM$ and $CoperSWI$ were scaled against in situ data, for each region independently. This process is needed to transform the indices into SM values with volumetric units ($m^3\,m^{-3}$). The following scaling formula was 215 applied:

$$SM_n^* = SM_n \times [max(SM_n^{IS}) - min(SM_n^{IS})] + min(SM_n^{IS}) \tag{1}$$

where SM_n and SM_n^* are respectively the original and scaled SM values from $CoperSSM$. SM_n^{IS} include all the SM observations from all the in situ time series available for the current region n. The 5% lowest and highest values are discarded before applying the minimum and maximum functions to remove the effect of possible outliers. The same process was also undertaken 220 to scale the SWI values from $CoperSWI$. The Copernicus indices were also scaled using $SMOSL3$ or $SMAPL3$ to obtain the maximum and minimum references instead of in situ data. The final results were quite comparable regardless the reference used and thus only the scaling against in situ was used for the rest of the study.

Since all the HR data sets are sampled on the same 1-km regular grid, the different time series were compared on a per pixel basis for each region in terms of Pearson correlation (R), bias and standard deviation of the difference ($STDD$, also referred 225 to as unbiased Root Mean Square of the Difference by some authors) creating maps of each metric. The results were analyzed as a function of the land cover class. More precisely, S^2MP_{S1S2} were compared to $CoperSSM$, $CoperSWI$ and $SMAPS1$. Of course, the S^2MP_{S1S3} maps were also compared to those of S^2MP_{S1S2}.

Finally, all the different CR and HR data sets were evaluated against the in situ data available in the 6 regions of study. For each ground station, the closest time series of each remotely sensed data set was compared to the in situ measurements 230 by computing the correlation, bias and $STDD$. Only samples for which the difference in acquisition times with the in situ measurements does not exceed 1 hour were taken into account to compute those statistical metrics. Then, only in situ locations where all the data sets provide a time series were used (Tab. 1). Time series of anomalies in a 35-days time window were also derived as follows.

$$SM_t^a = (SM_t - \mu_t)/\sigma_t$$

$$\mu_t = 1/N \times \sum_{n=t_1}^{t_2} SM_n \tag{2}$$

$$\sigma_t = \sqrt{1/(N-1) \times \sum_{n=t_1}^{t_2} (SM_n - \mu)^2}$$

235 where SM_t^a and SM_t are the SM and anomalies values at time t, respectively. N is the number of observations from t minus 17 days (t_1) to t plus 17 days (t_2).

Finally, the HR data were also evaluated against in situ measurements after aggregation at a 25-km resolution in order to compare their performances to the CR data at a comparable resolution.

4 Results

240 ### 4.1 Sentinel-3 versus Sentinel-2 NDVI

The S2 and S3 NDVI were compared in each site at 1 km grid scale to help better understand the reliability of using the S3 NDVI instead of the S2 NDVI when estimating the SM at 1 km grid scale using the S^2MP algorithm. The S2 NDVI at 1 km grid scale was calculated by averaging the S2 NDVI pixels at 10 m scale corresponding to the cropland and herbaceous vegetation within each 1 km grid cell (same grid of the S3 NDVI) and compared to the S3 NDVI obtained at 1 km. Figure 2

245 shows scatter plots between the S2-NDVI and the S3-NDVI over the year 2019 for all the study sites. Figure 2 shows that the correlation between the S2-NDVI and the S3-NDVI varies across the studied sites with high correlations observed in Australia (0.86) and Tunisia (0.79), moderate correlations for both USA (0.68) and Spain (0.64) and no correlation in the west and east of Occitanie. The results indicate that in dry regions such as that in Australia, Tunisia and USA high correlation exists between S2 and S3 NDVI whereas low correlation between S2 and S3 NDVI is present in humid areas such as the Occitanie region.

250 The results also show that the S3-NDVI for all sites saturates at approximately between 0.6 and 0.7 whereas the S2-NDVI reaches higher values between 0.8 and 0.9. The lower NDVI values provided by S3 NDVI compared to the S2 NDVI could be mainly due to the mixture of the surface reflectance for different land cover classes in the 1 km S3-NDVI. Figure 3 shows the distribution of the correlation coefficient (R2) values for each site as a function of the months. The results show that in the dry periods of the year (summer season of each site), the correlation between the S2 and S3 NDVI is higher than that in the winter

255 and spring season. For example, the Australian site shows that the correlation between the S2-NDVI and the S3-NDVI is only high (0.72) in the summer and autumn season between January and June. In USA, months from March to July had a correlation between 0.25 and 0.53 whereas other months had no correlation between the S2 and S3 NDVI values. In Occitanie east and west, most of the months show no correlation between the S2 and S3 NDVI values except for some summer months such as June, July and August. In fact, winter and spring seasons encounters the highest NDVI values due to the development of the

260 vegetation cycles. On the other hand, summer seasons usually have low NDVI values corresponding to bare soil conditions except in the presence of irrigated summer crops. Thus, for low NDVI values (usually in summer) the S2 and S3 are highly correlated whereas for high NDVI values with the peak of the vegetation development (in spring) the correlation between S2 and S3 decreases.

4.2 Comparison between S^2MP_{S1S2} and S^2MP_{S1S3}

265 Figure 4 shows the Pearson correlation (R), bias and standard deviation of the difference ($STDD$) between the S^2MP_{S1S2} and S^2MP_{S1S3} maps. A very good agreement between the two data sets was found in the six regions with $R \geq 0.9$, bias $\leq$ 0.04 m^3 m^{-3} (S^2MP_{S1S3} minus S^2MP_{S1S2}) and $STDD \leq 0.03$ m^3 m^{-3} for most of the areas. However, some differences can be seen between the two data sets in the northwest of the Spanish region, in the areas with significant forests cover in the France southwest region and in narrow areas of the Tunisian region. The somewhat higher differences in the Spanish and

270 France southwest regions are seen over pixels covered by a mix of forests, croplands and herbaceous vegetation (dominated by forests), while the somewhat larger differences in some narrow areas of Tunisia is due to heterogeneous land cover around several river basins with rolling topography, sparse forests as well as grasslands.

As discussed above, these small differences are expected due to the use of S3 NDVI estimated at 1-km resolution instead of S2 NDVI at 100-m resolution, which are then aggregated only for croplands and herbaceous vegetation areas within the 1-km^2

275 pixels. However, taking into account the overall very good agreement of S^2MP_{S1S2} and S^2MP_{S1S3} maps, for the sake of simplicity and clarity, in the following sections only S^2MP_{S1S2} is compared to the other HR data sets.

4.3 Comparison of S^2MP_{S1S2} against the HR SM data sets

Figures 5, 6 and 7 present the comparison of S^2MP_{S1S2} against $CoperSSM$, $CoperSWI$ as well as $SMAPS1$ over the six study regions in terms of bias, standard deviation of the difference ($STDD$), and Pearson correlation (R), respectively. Some

280 diagonal structures can be seen in the maps comparing S^2MP_{S1S2} to $CoperSSM$ in Spain and in the southwest of France. These artifacts, most pronounced in the Pearson correlation maps but also present in the bias and $STDD$ maps, come from the $CoperSSM$ data as previously discussed by Bazzi et al. (2019). Indeed, the artifacts are seen on the sub-swaths of the S1 product showing a big difference between the SM estimations in the $CoperSSM$ at the same SM estimation date. Bazzi et al. (2019) showed that the difference of the SM estimation at both sides of the sub-swath at a given date of the $CoperSSM$ map

285 can reach 0.11 m^{3}m^{-3}.

4.3.1 Absolute values

Figure 5 shows that S^2MP_{S1S2} show a bias in the range from -0.1 to 0.1 m^3 m^{-3} with respect to the other HR products over most of the pixels within the six regions of study. However, there are areas in the Spanish, Tunisian and France southeast domains, where S^2MP_{S1S2} shows a dry bias of absolute value larger than 0.1. This is also particularly the case in the

290 southwest of France, with respect to $CoperSSM$ and $CoperSWI$ as well as with respect to $SMAPS1$ in Australia and North America. For these regions and HR data sets, the bias is negative over the whole area. For all the other combinations of regions and HR products, the bias values are both positive and negative. For instance, the bias distribution of S^2MP_{S1S2} with respect to $SMAPS1$ in the France southwest region is splitted in two (Fig. 5i). The bias is wet in the west part of the region and dry in the east part. There is no clear relationship between this behavior and the dominant land cover class. However, the dry

295   bias observed in the east part of the Tunisian region corresponds to an area of salted lakes, named Sebkha, whose water and moisture contents can vary significantly according to climate.

Figure 6 shows that the $STDD$ values of S^2MP_{S1S2} with respect to $CoperSSM$ and $CoperSWI$ are lower than 0.05 $\mathrm{m^3\,m^{-3}}$ over almost all the pixels of the 4 regions where the Copernicus data sets are available. However, the $STDD$ obtained with respect to $SMAPS1$ (Fig. 6c,i,l) are often much higher and can reach 0.12 $\mathrm{m^3\,m^{-3}}$. These significant differences are

300 specifically observed in the west part of the Spanish region, the east part of the France southwest region as well as the northwest part of the France southeast region. The $STDD$ obtained in Tunisia with respect to $SMAPS1$ (Fig. 6f) are comparable to those obtained with respect to the Copernicus data sets (Fig. 6d,e) in the north part of the same region, and do not exceed 0.06 $\mathrm{m^3\,m^{-3}}$. In Australia and North America (Fig. 6m,n), the $STDD$ with respect to $SMAPS1$ are quite similar to those found in Spain and France (Fig. 6c,i,l), where values can reach $\sim 0.08 - 0.1 \mathrm{m^3\,m^{-3}}$ in the southeast and northwest parts of

305 the Australia and North America regions, respectively. There is no clear and unique relationship with the dominant land cover class. For instance, the $STDD$ with respect to $SMAPS1$ in the southwest of France is higher over the forests than over the croplands dominated areas, while in the North America region the $STDD$ was found to be lower over the forest.

4.3.2   Temporal dynamics

Overall, S^2MP_{S1S2} and $CoperSSM$ show a high correlation (above 0.7-0.8) over almost all the pixels for all the regions of

310 study (Figs. 7a,d,g,j). In contrast, lower values are found for the correlation between S^2MP_{S1S2} and $CoperSWI$ (maximum values of 0.6-0.7 and values lower than 0.4 in many large areas, Figs. 7b,e,h,k). Finally, the lowest correlation values are obtained with respect to the $SMAPS1$ data set (Figs. 7c,f,i,l,m,n).

In the Spanish region, the highest correlation values are obtained in the areas dominated by croplands. The lowest values are found in the northwest and in a few spots in the south and the center of the region. In all cases, those are areas with

315 heterogeneous land cover dominated by forests. Similar spatial features are observed in the three maps comparing S^2MP_{S1S2} to $CoperSSM$, $CoperSWI$ and $SMAPS1$ (Figs. 7a-c) but with lower values of correlation with respect to $CoperSWI$. The areas of high correlation with respect to $SMAPS1$ show higher values than with respect to $CoperSWI$ but the areas with low correlation also show lower values than with respect to $CoperSWI$.

In Tunisia (Fig. 7d-f), the correlations obtained in the north are quite good with values of 0.8-0.9 with respect to $CoperSSM$

320 and 0.6-0.7 with respect to $SMAPS1$. The values drop in the south to values lower than 0.3, especially near the coasts, where mixed land cover pixels include urban areas. The correlation with respect to $CoperSWI$ (Fig. 7e) is only higher than 0.5 for the regions where the 1-km^2 pixels are dominated by croplands (Fig. 1).

In the France southwest region (Fig. 7g-i), there is a clear frontier from the areas where the pixels are dominated by croplands with respect to the areas dominated by forests (Fig. 1). In the case of the correlation with respect to $CoperSSM$, values remain

325 higher than 0.7 for most of the areas dominated by forests. However in the case of the correlation with respect to $CoperSWI$ and $SMAPS1$, values in those areas decrease to lower than 0.3. In the France southeast region (Fig. 7j-l), correlations are only significant with respect to $CoperSSM$ and in regions dominated by croplands.

The correlations between S^2MP_{S1S2} and $SMAPS1$ reach 0.6-0.7 in Australia (Fig. 7m). Similar moderate to low correlation values are obtained in North America with no clear relationship with the land cover type (Fig. 7n).

4.3.3 Comparison as a function of land cover

As discussed above, the correlation maps show some features related to the dominant land cover class, in particular, higher correlations are found for areas dominated by croplands and herbaceous vegetation. To get further insight into the effect of the dominant land cover and the homogeneity of the land cover within the 1-km^2 pixels, Table 2 gives the median Pearson correlation (R) of S^2MP_{S1S2} with respect to $CoperSSM$, $CoperSWI$ and $SMAPS1$ for pixels dominated by croplands and herbaceous vegetation to different extend: *(i)* "mixed" pixels in which croplands or herbaceous vegetation are the dominant land cover class in any proportion and *(ii)* "homogeneous" pixels in which croplands or herbaceous vegetation represent at least 75% of the 1-km^2 area.

In the Spanish region, S^2MP_{S1S2} is better correlated to $CoperSSM$ ($R \sim$ 0.7-0.8) than $CoperSWI$ ($\sim$ 0.55-0.65) and $SMAPS1$ ($R \sim$ 0.45-0.65) when pixels become more homogeneous. The same pattern can be observed in the southwest and southeast of France when S^2MP_{S1S2} is compared to $CoperSSM$ and $CoperSWI$ over croplands. Correlations between S^2MP_{S1S2} and $SMAPS1$ significantly drop in those regions ($R \sim$ 0.2-0.5) over croplands as well as herbaceous vegetation areas. In addition, S^2MP_{S1S2} shows very low correlations with respect to the other SM HR data sets in Tunisia ($R \sim$ 0.2-0.55) for the both land cover types. However, the correlations obtained between S^2MP_{S1S2} and $SMAPS1$ in Australia and North America ($R \sim$ 0.5-0.55) are higher than those found in France and Tunisia, both over croplands and herbaceous vegetation.

Compared to the other HR data sets, S^2MP_{S1S2} shows generally higher correlations over croplands than over herbaceous vegetation. In addition, there is a clear improvement in correlation of S^2MP_{S1S2} with respect to the other data sets when the land cover within the pixel is homogeneous, in particular over croplands.

4.4 Evaluation against in situ data

Table 3 presents the evaluation of the different CR and HR SM products with respect to in situ data in terms of bias, $STDD$ and Pearson correlation of the original time series (R) as well as Pearson correlation of the anomalies time series (R^a). The highest bias is obtained for $CCISM$ with 0.07 m^3 m^{-3}. The SMAP products, S^2MP_{S1S2} as well as $SMOSNRT$ show the lowest bias ($\leq 0.005 m^3$ m^{-3} in absolute value) with respect to in situ data.

The highest (0.08 m^3 m^{-3}) and lowest $STDD$ (0.04 m^3 m^{-3}) are obtained for $CoperSSM$ and $CCISM$, respectively. S^2MP_{S1S2}, as most of the data sets, shows an in between value with 0.05 m^3 m^{-3}.

The highest correlation values are obtained for the coarse resolution data (0.65-0.68) and the lowest for the S^2MP_{S1S2} and $CoperSSM$ Sentinel-only HR data sets (0.44-0.48). The HR data obtained from merging approaches ($CoperSWI$ and $SMAPS1$) show intermediate values (0.60-0.64).

Regarding the correlation of the anomalies time series, $CoperSWI$ gets the highest R^a with respect to in situ data with 0.49, closely followed by $SMAPL3E$ with 0.46. $CoperSSM$ and $SMAPS1$ show the lowest performances with R^a equal

360 to 0.25 and 0.26. S^2MP_{S1S2} gets an higher value with 0.36, which is also better than two CR data sets: $SMOSL3$ ($R^a{\sim}0.30$) and $SMOSNRT$ ($R^a{\sim}0.33$).

The performances of the two Sentinel-only HR data sets averaged at 25-km resolution with respect to in situ data are comparable to the performances obtained for the original 1-km data sets. The performance of the merged ASCAT+S1 ($CoperSWI^*$) averaged at 25-km resolution is also similar to that of the 1-km resolution data set ($CoperSWI$). In contrast, for the $SMAPS1^*$

365 data set, which is a downscaled product, the correlation increases from 0.60 at 1-km resolution ($SMAPS1$) to 0.69 at 25-km resolution. R^a also increases from 0.26 at 1-km resolution to 0.32 at 25-km resolution, but it does not reach the values of R^a obtained for the SMAP-only ($SMAPL3$ and $SMAPL3E$) products (0.43-0.46).

5 Discussion

The four HR data sets evaluated in this study have been produced with different approaches. Two products were obtained by

370 merging S1 data with ASCAT ($CoperSWI$) and SMAP ($SMAPS1$), respectively. S^2MP_{S1S2} and $CoperSSM$ are based on Sentinel-only. The last one is computed from local temporal variations of the S1 backscatter coefficients time series following the method of Wagner (1998). In contrast, S^2MP_{S1S2} uses electromagnetic modelling to compute a database predicting backscatter coefficients from surface characteristics such as SM and vegetation status (approximated by NDVI). This database is in turn used to train a neural network to predict SM from S1 backscatter coefficients and HR NDVI estimations.

375 Initially, the S^2MP algorithm by El Hajj et al. (2017) was limited to croplands and the use of NDVI derived from S2. In the framework of this study, the algorithm has been extended to herbaceous vegetation areas. In addition, two methodologies are used to mimic vegetation dynamics. In the first approach, S1 backscatter coefficients and S2 NDVI are estimated over croplands and herbaceous vegetation areas at a 100-m resolution and then they are aggregated within 1-km^2 pixels (S^2MP_{S1S2}). In the second approach, instead of using S2 NDVI, S3 NDVI are obtained directly at a resolution of 1 km (S^2MP_{S1S3}).

380 The results of this study show that when the target resolution is 1 km, it is possible to replace S2 NDVI by S3 NDVI in the S^2MP approach without losing skill. Regarding the comparison to the other Sentinel-only product ($CoperSSM$), high correlations have been found over croplands and herbaceous areas (Tab. 2). Different SM dynamics are actually noticeable, e.g., in the forested areas of southwestern France (Fig. 1 and 7).

In addition, slightly better results of S^2MP_{S1S2} with respect to in situ measurements compared to those of $CoperSSM$

385 were found (Table 3). This is also true for the bias even if the $CoperSSM$ retrievals have previously been scaled against the in situ data. Most of the ground sites are representative of croplands and herbaceous regions. Therefore, it is difficult to evaluate the relative performance against in situ measurements over other land cover types for which more differences would be expected, e.g., forest cover. On the one hand, by construction, the S^2MP algorithm starts being out of its application domain when considering pixels dominated by forests cover. On the other hand, the $CoperSSM$ indices computation does currently

390 not account for vegetation dynamics, which can lead to biases with respect to in situ data over areas covered by seasonal and dense vegetation. In addition, for most applications the $CoperSSM$ indices should also be transformed into SM time series

and this will be problematic without reference SM values under forest to scale them. Therefore, an extension of the S^2MP approach to forest areas would definitely be interesting to provide HR SM mapping over large regions.

SM retrievals using the S^2MP approach were already evaluated against in situ measurements along with other HR and CR
395 data sets by El Hajj et al. (2018) and compared to CGLS SM byBazzi et al. (2019). In their evaluation with respect to in situ data, El Hajj et al. (2018) found that the S^2MP_{S1S2} data set shows lower correlations than $SMAPL3$ and $SMAPL3E$ but higher correlations than $SMOSNRT$, $SMOSL3$ and $SMAPS1$. In contrast, in the current study, the S^2MP_{S1S2} shows a lower correlation against in situ data than the other products. These differences can be due to the fact that the analysis was only carried out in the south of France during a different time period (from January 2016 to June 2017). Moreover, in El Hajj
400 et al. (2018), SM estimates from the S^2MP algorithm were only derived over croplands while in our study, the S^2MP SM retrievals were performed both for croplands and herbaceous vegetation. However, the results of the current study are in good agreement with those of El Hajj et al. (2018) in terms of bias. In both studies, S^2MP_{S1S2} shows low bias compared to in situ measurements, similar to or lower than the CR data sets. Regarding the S^2MP_{S1S2} and $CoperSSM$ comparison by Bazzi et al. (2019), it is interesting to note that S^2MP_{S1S2} is better correlated to in situ data than $CoperSSM$, which is in agreement
405 with the results of our study (Table 3).

Taking into account the results discussed in the previous section, HR merged products ($SMAPS1$, $CoperSWI$) provide better estimations or temporal agreement with in situ data than the HR Sentinel-only data (S^2MP_{S1S2}, $CoperSSM$) (Tab. 3) but still show lower performances than the CR data. This was also found by Bauer-Marschallinger et al. (2019), who demonstrated that 25-km resolution SM estimates from ASCAT were better correlated to in situ measurements within the
410 Italian Umbria region that those of the 1-km resolution $CoperSSM$. The results of our study are also in perfect agreement with the findings by Ojha et al. (2021), who showed that two merged products, $SMAPS1$ and SMAP + DISPATCH (Merlin et al., 2012), were better correlated to in situ data than $CoperSSM$ over several regions in France and Spain.

When aggregated to CR, the performances of the $SMAPS1$ 1-km data increase to values similar to those of the $SMAPL3$ data. This implies that the gain in resolution brought by merging data of different resolutions comes at the expense of intro-
415 ducing uncertainties in the resulting HR data set.

Obviously, this study was limited to comparisons over six regions of $10^4\ km^2$ within a 3-years time period, so the results can not be straightforwardly extended to a global scale. However, it is noticeable that the HR products provide SM estimates using C-band measurements while the CR data sets used in this study are computed using L-band measurements. Indeed, SMOS and SMAP were specifically designed to measure surface SM, which was not the case for the Sentinel satellites.

420 **6 Conclusions**

The goal of this study was to adapt the S^2MP approach, originally designed to retrieve SM at the plot scale over agricultural fields, to a 1-km resolution, which allows to replace S2 by S3 NDVI and to significantly improve the temporal sampling. In addition, the approach was extended to herbaceous land cover areas in addition to croplands and tested in six regions over four continents to assess its performances beyond previous evaluations in Southern France.

Very good agreement was found between the S1+S2 and the S1+S3 S^2MP maps in terms of correlations ($R \geq 0.9$), bias ($\leq 0.04 \ m^3 m^{-3}$) and standard deviation of the difference ($STDD \leq 0.03 \ m^3 m^{-3}$) for the 6 regions.

The S^2MP maps were compared to those of the 1-km surface SM product provided by CGLS, which is also a Sentinel-only based data set. In contrast to S^2MP, the CGLS surface SM is a local index of SM variations and for many applications it has to be scaled against a reference to transform the variation index to actual SM in volumetric units ($m^3 m^{-3}$) before being used. S^2MP was also compared to two HR merged data sets, *i* the SWI from CGLS combining S1 and ASCAT measurements as well as *ii* the SMAP+S1. As for the surface SM data set, the SWI data had to be scaled into absolute SM values. CGLS products only provide retrievals over the European continent and the Mediterranean basin.

The results showed that the S^2MP data set is well correlated to the 1-km surface SM product provided by CGLS in large areas of the 6 regions of study: $R \sim 0.7\text{-}0.8$. However, the correlations significantly drop in some areas of the 6 domains when the S^2MP maps are compared to those of the SWI product and those of SMAP+S1. Over almost all the pixels within the 6 regions, the standard deviation of the difference ($STDD$) between S^2MP and the CGLS products are lower than $0.06 \ m^3 m^{-3}$ while the values obtained with respect to SMAP+S1 can reach $0.1 \ m^3 m^{-3}$. Then, the bias between S^2MP and the other HR data sets differ significantly inside a same region and can be strongly dry or wet ($\pm \ 0.1 \ m^3 m^{-3}$). The correlations between S^2MP and the other HR data sets improve over croplands when the 1-km pixels are homogeneous but a similar behaviour was not found for pixels where the dominant land cover class is herbaceous vegetation.

The HR products were also evaluated with respect to in situ measurements along with coarser resolution data sets from SMOS, SMAP and ESA CCI. The coarse resolution (CR) products show higher correlations ($0.65 \leq R \leq 0.68$) than the HR data sets ($0.44 \leq R \leq 0.64$), and the HR merged data sets showed higher correlations than the HR Sentinel-only ones. However, as several SMAP and SMOS data sets, S^2MP also showed a very low bias with respect to in situ measurements. Finally, the $STDD$ differ according to the data set as well as the spatial resolution and range from 0.04 to $0.08 \ m^3 m^{-3}$ for the ESA CCI product to the SWI product.

The S^2MP algorithm gave promising results over the 6 regions of study. Of course, for some applications the approach should be extended to consider other land cover types than croplands and herbaceous. For instance, a 1 km resolution SM data sets over all land cover types could be used to asses climate impact at regional level in the future when 1 km SM times series are longer. Finally, a remaining challenge is to provide HR SM data with comparable spatio-temporal coverage and retrieval quality across different land cover types than those of the state-of-the-art coarse resolution products, such as the SMOS, SMAP and ESA CCI products.

Author contributions. RM, NJR-F, NB and MZ designed the study. RM and NJR-F undertook the different evaluations and wrote the first version of the manuscript. HB and NB produced the S^2MP_{S1S2} and S^2MP_{S1S3} maps. CA, NB, WD and MZ participated in the analysis of the results. All the authors contributed to the final version of the manuscript.

Competing interests. The authors declare no conflict of interest.

Acknowledgements. This research made use of data from the Centre Aval de Traitement des Données SMOS (CATDS) operated for the Centre National d'Etudes Spatiales (CNES) by The Institut Français de Recherche pour l'Exploitation de la Mer (IFREMER) in France, as well as from the Copernicus Global Land Service (CGLS), the National Snow and Ice Data Center (NSIDC) and the ESA's Climate Change Initiative for Soil Moisture project. The authors acknowledge partial funding from the ESA's Climate Change Initiative for Soil Moisture project (Contract No. 4000104814/11/I-NB and 4000112226/14/I-NB). RM and NJR-F acknowledge partial funding by the Centre National d'Études Spatiales (CNES) APR TOSCA project SMOS-TE.
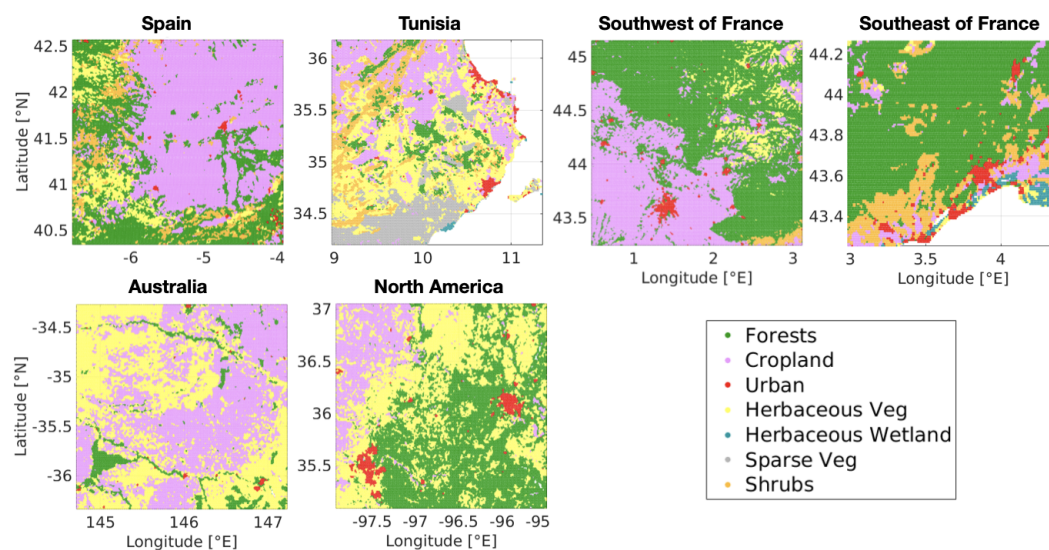
Figure 1. Copernicus land cover maps of the 6 regions of study at 1-km spatial resolution. Only the dominant land cover type within a 1-km^2 pixel is shown. For instance, a pixel characterised as forests can contain 27% of forests, 26% of croplands, 24% of herbaceous vegetation and 23% of shrublands, or 90% of forests and 10% of herbaceous vegetation.

Table 1. In-situ measurements that were used in this study. The depths are quoted as two numbers: the first one is the upper depth, and the second one is the lower depth of the sensor. Both numbers are equal when the sensor is placed horizontally. The fourth column gives the number of sensors available for each region.

Location	Measurements	Depth (m)	Sensors	Reference
Spain	REMEDHUS	0–0.05	20	Gonzalez-Zamora et al. (2018)
Southwest of France	SMOSMANIA	0.05–0.05	4	Calvet et al. (2007)
Southeast of France	SMOSMANIA	0.05–0.05	6	Calvet et al. (2007)
Australia	OZNET	0–0.05	11	Smith et al. (2012); Young et al. (2008)
North America	USCRN	0.05–0.05	2	Bell et al. (2013)
Tunisia	MERGUELLIL	0–0.05	6	Gorrab et al. (2015)
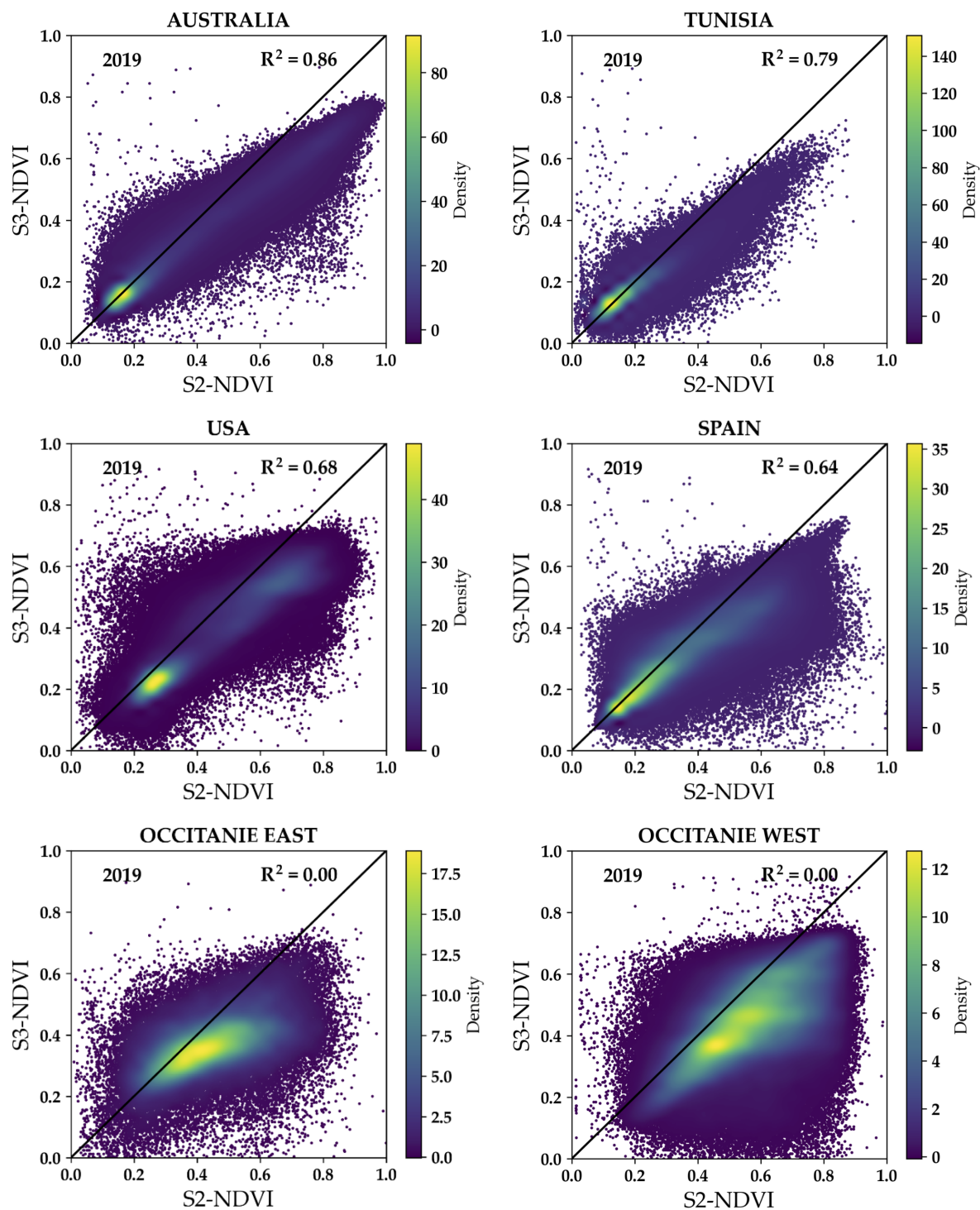
Figure 2. Correlation between S2 and S3 NDVI at 1-km grid scale over one year for the 6 study sites.
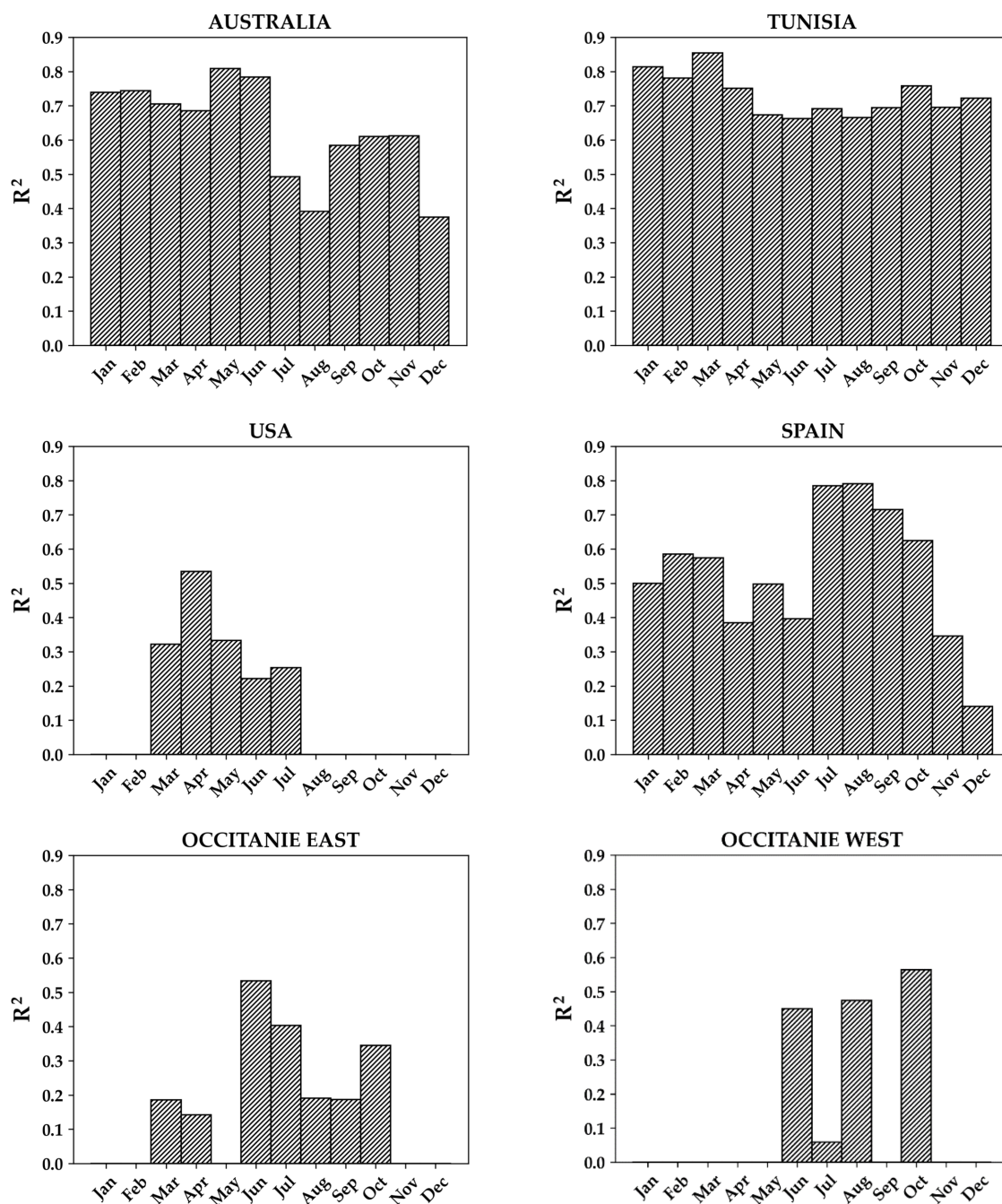
Figure 3. Correlation between S2 and S3 NDVI at 1-km grid scale each month for the 6 study sites. Months having no bars means that there is no correlation between S2 and S3 NDVI.
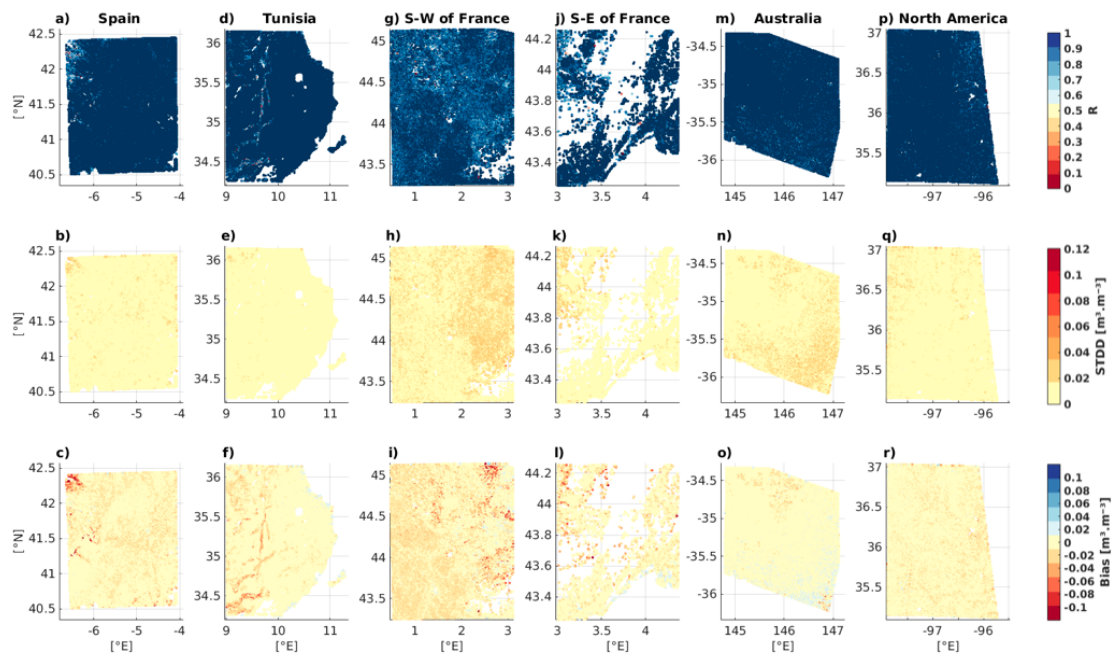
Figure 4. Comparison of S^2MP_{S1S3} with respect to S^2MP_{S1S2} over the regions of study in terms of Pearson correlation (R) as well as bias (S^2MP_{S1S2} minus S^2MP_{S1S3}) and standard deviation of the difference ($STDD$) in m^3 m^{-3}.
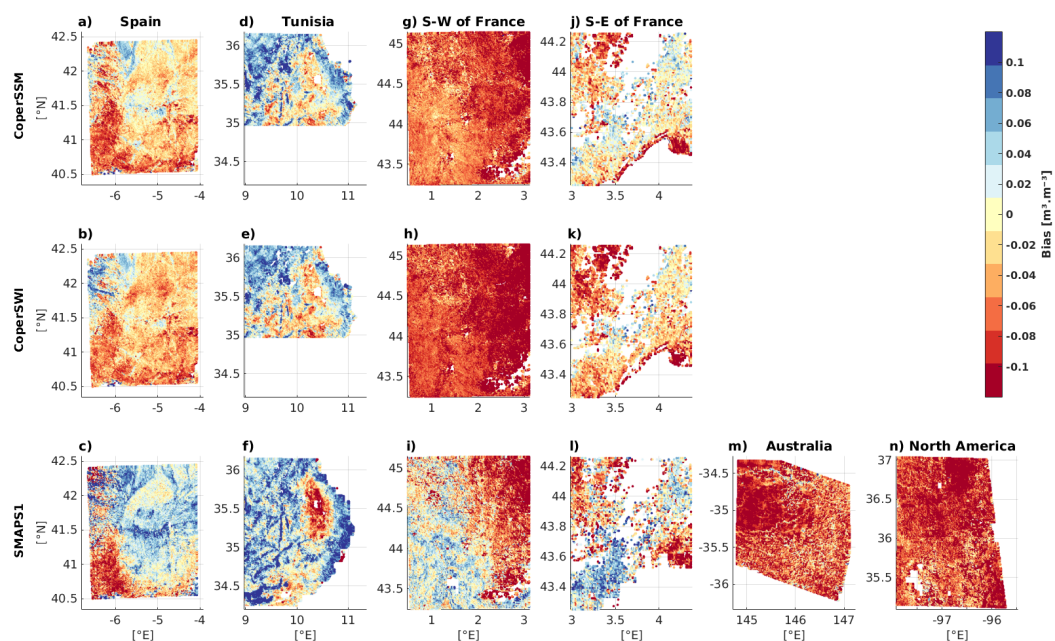
Figure 5. Comparison of S^2MP_{S1S2} with respect to $CoperSSM$ (S^2MP_{S1S2} minus $CoperSSM$), $CoperSWI$ (S^2MP_{S1S2} minus $CoperSWI$) and $SMAPS1$ (S^2MP_{S1S2} minus $SMAPS1$) over the regions of study in terms of bias in m^3 m^{-3}.
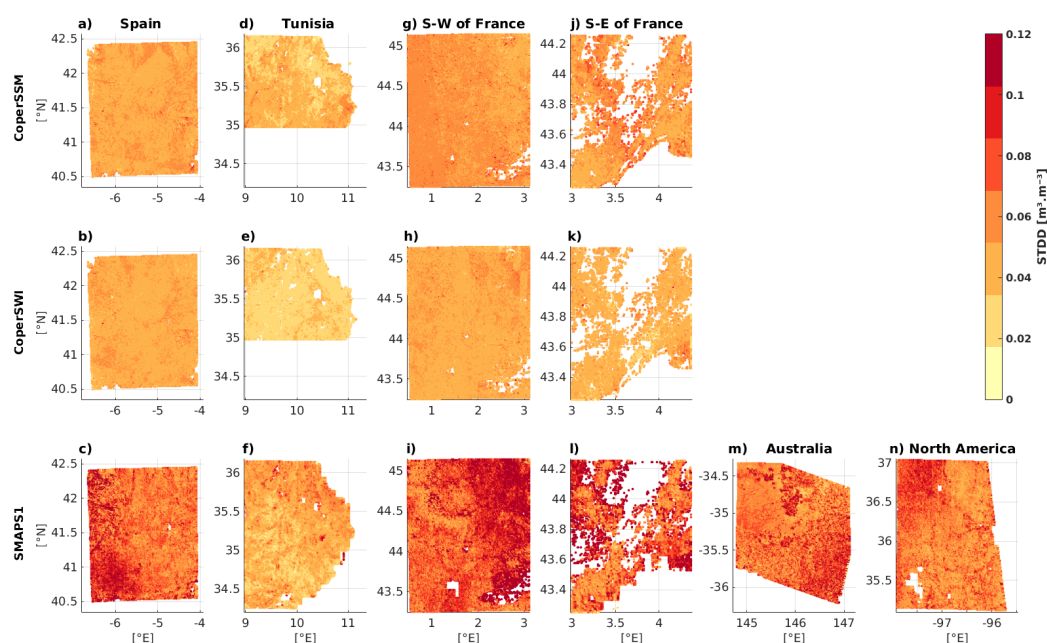
Figure 6. Comparison of S^2MP_{S1S2} with respect to $CoperSSM$, $CoperSWI$ and $SMAPS1$ over the regions of study in terms of standard deviation of the difference ($STDD$) in $\mathrm{m^3\,m^{-3}}$.
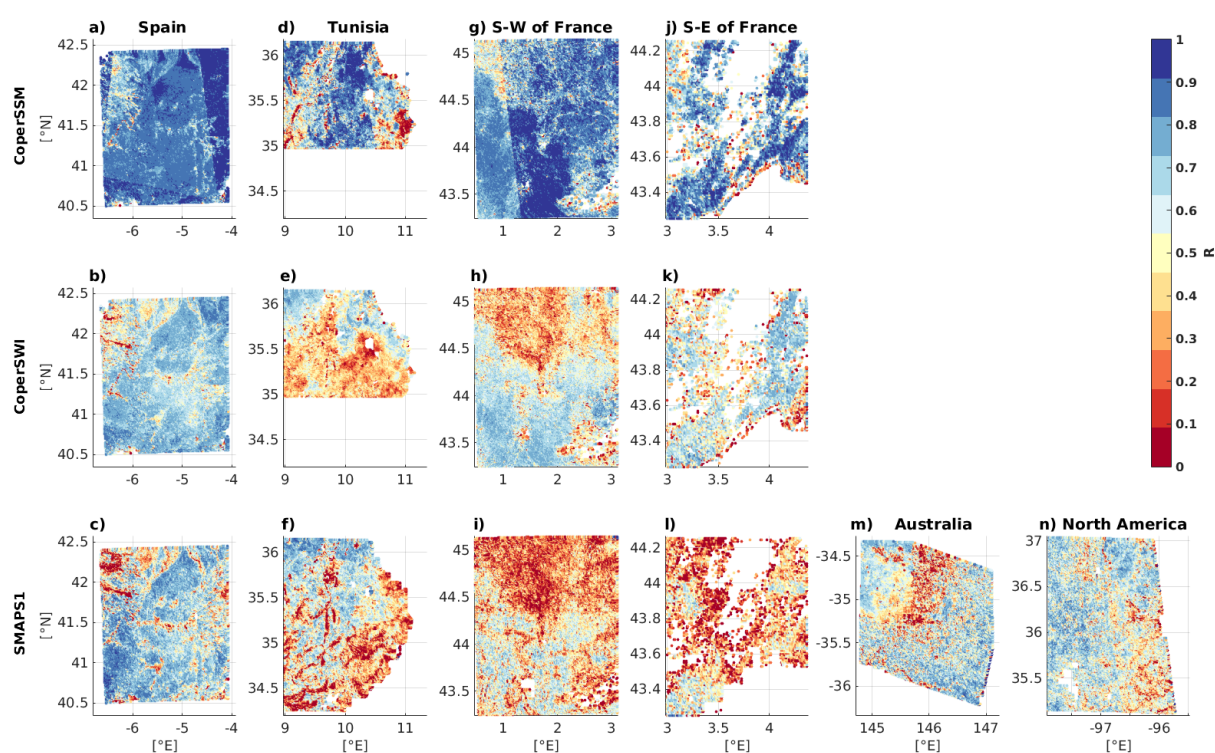
Figure 7. Comparison of S^2MP_{S1S2} with respect to $CoperSSM$, $CoperSWI$ and $SMAPS1$ over the regions of study in terms of Pearson correlation (R).

Table 2. Comparison of S^2MP_{S1S2} against the HR SM products, in terms of Pearson correlation (R), over 1-km^2 pixels where croplands and herbaceous vegetation are the dominant land cover classes. M stands for "mixed" pixels in which croplands or herbaceous vegetation are the dominant land cover classes in any proportion while H stands for "homogeneous" pixels, in which at least 75% of the 1-km^2 area is covered by croplands or herbaceous vegetation. Only median values are shown.

Regions	Products	Croplands		Herbaceous veg	
		M	H	M	H
Spain	$CoperSSM$	0.75	0.77	0.67	0.66
	$CoperSWI$	0.64	0.65	0.56	0.59
	$SMAPS1$	0.60	0.63	0.46	0.48
Tunisia	$CoperSSM$	0.41	0.54	0.26	0.24
	$CoperSWI$	0.34	0.44	0.23	0.24
	$SMAPS1$	0.40	0.48	0.30	0.31
Southwest of France	$CoperSSM$	0.69	0.77	0.64	0.52
	$CoperSWI$	0.58	0.64	0.52	0.56
	$SMAPS1$	0.36	0.48	0.27	0.18
Southeast of France	$CoperSSM$	0.68	0.78	0.39	-
	$CoperSWI$	0.53	0.56	0.27	-
	$SMAPS1$	0.32	0.29	0.22	-
Australia	$CoperSSM$	-	-	-	-
	$CoperSWI$	-	-	-	-
	$SMAPS1$	0.49	0.52	0.49	0.52
North America	$CoperSSM$	-	-	-	-
	$CoperSWI$	-	-	-	-
	$SMAPS1$	0.54	0.54	0.53	0.54

Table 3. Evaluation of the HR and CR SM time series against in situ data in terms of Pearson correlation (R,R^a), bias (remotely sensed minus ground based SM in [m3 m^{-3}]) and standard deviation of the difference ($STDD$ in [m^3 m^{-3}]). The metrics were computed by taking into account the 6 regions of study together and only the median values are shown here. The symbol * indicates the HR data sets averaged at 25-km resolution.

Products	R	R^a	$Bias$	$STDD$
Sentinel-only high resolution data				
S^2MP_{S1S2}	0.48	0.36	0.003	0.05
$CoperSSM$	0.44	0.25	0.05	0.08
Merged high resolution data				
$CoperSWI$	0.64	0.49	0.06	0.05
$SMAPS1$	0.60	0.26	-0.001	0.07
Coarse resolution data				
$SMAPL3$	0.68	0.43	0.005	0.05
$SMAPL3E$	0.68	0.46	0.002	0.05
$SMOSL3$	0.65	0.30	0.01	0.07
$SMOSNRT$	0.67	0.33	-0.004	0.06
$CCISM$	0.68	0.40	0.07	0.04
High resolution data aggregated to coarse resolution				
$S^2MP^*_{S1S2}$	0.49	0.33	0.01	0.05
$CoperSSM^*$	0.43	0.22	0.05	0.07
$CoperSWI^*$	0.63	0.48	0.06	0.05
$SMAPS1^*$	0.69	0.32	0.01	0.05

References

Al Bitar, A., Mialon, A., Kerr, Y., Cabot, F., Richaume, P., Jacquette, E., Quesney, A., Mahmoodi, A., Tarot, S., Parrens, M., Al-yaari, A., Pellarin, T., Rodriguez-Fernandez, N., and Wigneron, J.-P.: The Global SMOS Level 3 daily soil moisture and brightness temperature maps, Earth System Science Data, 9, 293–315, https://doi.org/10.5194/essd-2017-1, 2017.

Albergel, C., Rüdiger, C., Pellarin, T., Calvet, J. C., Fritz, N., Froissard, F., and Martin, E.: From near-surface to root-zone soil moisture using an exponential filter: an assessment of the method based on in-situ observations and model simulations., Hydrology and Earth System Sciences, 12(6), 1323–1337, 2008.

Baghdadi, N., Holah, N., and Zribi, M.: Calibration of the integral equation model for SAR data in C-band and HH and VV polarizations., Remote Sensing, 27, 805–816, 2006.

Baghdadi, N., Chaaya, J., and Zribi, M.: Semiempirical calibration of the integral equation model for SAR data in C-band and cross polarization using radar images and field measurements., IEEE Geosci. Remote Sens. Lett., 8, 14–18, 2011.

Baghdadi, N., El Hajj, M., and Zribi, M.and Bousbih, S.: Calibration of the Water Cloud Model at C-Band for Winter Crop Fields and Grasslands., Remote Sensing, 9, 969, 2017.

Balenzano, A., Mattia, F., Satalino, G., and Davidson, M.: Dense Temporal Series of C- and L-band SAR Data for Soil Moisture Retrieval Over Agricultural Crops., IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 4, 439–450, 2010.

Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L., and Wagner, W.: Soil Moisture from Fusion of Scatterometer and SAR: Closing the Scale Gap with Temporal Filtering., Remote Sensing, pp. 1 – 20, https://doi.org/DOI 10.3390/rs10071030, 2018.

Bauer-Marschallinger, B., Freeman, V., Cao, S., Paulik, C., Schaufler, S., Stachl, T., Modanesi, S., Massari, C., Ciabatta, L., Brocca, L., and Wagner, W.: Toward Global Soil Moisture Monitoring With Sentinel-1: Harnessing Assets and Overcoming Obstacles., IEEE Transactions on Geoscience and Remote Sensing, pp. 1 – 20, https://doi.org/DOI 10.1109/TGRS.2018.2858004, 2019.

Bazzi, H., Baghdadi, N., El Hajj, M., Zribi, M., and Belhouchette, H.: A Comparison of Two Soil Moisture Products S2MP and Copernicus-SSM over Southern France., Journal of Selected Topics in Applied Earth Observations and Remote Sensing, https://doi.org/doi: 10.1109/JSTARS.2019.2927430, 2019.

Bell, J., Palecki, M., Baker, C., Collins, W., Lawrimore, J., Leeper, R., Hall, M., Kochendorfer, J., Meyers, T., Wilson, T., and Diamond, H.: U.S. Climate Reference Network soil moisture and temperature observations., J. Hydrometeorol., 14, 977–988, 2013.

Bousbih, S., Zribi, M., Lili-Chabaane, Z., Baghdadi, N., El Hajj, M., Gao, Q., and Mougenot, B.: Potential of Sentinel-1 Radar Data for the Assessment of Soil and Cereal Cover Parameters., Sensors, 17, 2617, 2017.

Buchhorn, M., Bertels, L., Smets, B., De Roo, B., Lesiv, M., Tsendbazar, N. E., Masiliunas, D., and Li, L.: Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Algorithm Theoretical Basis Document, https://doi.org/DOI 10.5281/zenodo.3606361, 2020.

Calvet, J.-C., Fritz, N., Froissard, F., Suquia, D., Petitpa, A., and Piguet, B.: In situ soil moisture observations for the CAL/VAL of SMOS: The SMOSMANIA network, in: Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International, pp. 1196–1199, IEEE, 2007.

Chan, S., Njoku, E. G., and Colliander, A.: SMAP L1C Radiometer Half-Orbit 36 km EASE-Grid Brightness Temperatures, Version 4., Boulder, Colorado USA. NASA National Snow and Ice Data Center, https://doi.org/https://doi.org/10.5067/ZVILG0PS6CTI., 2018.

Das, N., Entekhabi, D., Dunbar, S., Chaubell, J., Colliander, A., Yueh, S., Jagdhuber, T., Chen, F., Crow, W. T., O'Neill, P. E., Walker, J., Berg,
500   A., Bosch, D., Caldwell, T., Cosh, M., Collins, C. H., Lopez-Baeza, E., and Thibeault., M.: The SMAP and Copernicus Sentinel 1A/B
    microwave active-passive high resolution surface soil moisture product, Remote Sensing of Environment. 233. 111380., https://doi.org/doi:
    https://doi.org/10.1016/j.rse., 2019.

Das, N., Entekhabi, D., Dunbar, R. S., Kim, S., Yueh, S., Colliander, A., O'Neill, P. E., Jackson, T., Jagdhuber, T., Chen, F., Crow, W. T.,
    Walker, J., Berg, A., Bosch, D., Caldwell, T., , and Cosh, M.: SMAP/Sentinel-1 L2 Radiometer/Radar 30-Second Scene 3 km EASE-
505   Grid Soil Moisture, Version 3, Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center.,
    https://doi.org/doi: https://doi.org/10.5067/ASB0EQO2LYJV., 2020.

De Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., and Isaksen, L.: A simplified Extended Kalman Filter for the global
    operational soil moisture analysis at ECMWF, Quarterly Journal of the Royal Meteorological Society, 139, 1199–1213, 2013.

de Rosnay, P., Balsamo, G., Albergel, C., Muñoz-Sabater, J., and Isaksen, L.: Initialisation of land surface variables for numerical weather
510   prediction, Surveys in Geophysics, 35, 607–621, 2014.

Dorigo, W., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., Oevelen, P. v., et al.: The
    International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements, Hydrology and Earth System
    Sciences, 15, 1675–1698, 2011.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., et al.: ESA CCI Soil
515   Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215,
    2017.

Dorigo, W., Himmelbauer, I., and Aberer, D.: The International Soil Moisture Network: serving Earth system science for over a decade,
    Hydrology and Earth System Sciences, https://doi.org/https://doi.org/10.5194/hess-2021-2, 2021.

El Hajj, M., Baghdadi, N., Zribi, M., and Bazzi, H.: Synergic use of Sentinel-1 and Sentinel-2 images for operational soil moisture mapping
520   at high spatial resolution over agricultural areas., Remote Sensing, 9, 1292, https://doi.org/doi:10.3390/rs9121292, 2017.

El Hajj, M., Baghdadi, N., Zribi, M., Rodríguez-Fernández, N., Wigneron, J. P., Al-Yaari, A., Al Bitar, A., Albergel, C., and Calvet, J. C.:
    Evaluation of SMOS, SMAP, ASCAT and Sentinel-1 Soil Moisture Products at Sites in Southwestern France., Remote Sensing, 10(4),
    569, 2018.

Entekhabi, D., Njoku, E. G., O'Neill, P. E., Kellogg, K. H., Crow, W. T., Edelstein, W. N., Entin, J. K., Goodman, S. D., Jackson, T. J.,
525   Johnson, J., et al.: The soil moisture active passive (SMAP) mission, Proceedings of the IEEE, 98, 704–716, 2010.

Entekhabi, D., Yueh, S., O'Neill, P. E., and Kellogg, K. H.: SMAP Handbook, Tech. rep., Jet Propulsion Laboratory, NASA, 2014.

GCOS: The Status of the Global Climate Observing System 2021, Tech. rep., Global Climate Observing System, World Meteorological
    Organization, Report 240, 2021.

Gonzalez-Zamora, A., Sanchez, N., Pablos, M., and Martinez-Fernandez, J.: CCI soil moisture assessment with SMOS soil moisture and in
530   situ data under different environmental conditions and spatial scales in Spain., Remote Sensing of Environment, 255, 2018.

Gorrab, A., Zribi, M., Baghdadi, N., Mougenot, B., and Lili-Chaabane, Z.: Retrieval of both soil moisture and texture using TerraSAR-X
    images., Remote Sensing, 7, 10 098–10 116, https://doi.org/doi:10.3390/rs70810098, 2015.

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and
    their underlying merging methodology, Earth System Science Data, pp. 1–37, 2019.

535 Guerif and Duke: Adjustment procedures of a crop model to the site specific characteristics of soil and crop using remote sensing data
    assimilation, Agriculture Ecosystems Environment, pp. 57–69, 2000.

Hajj, M., Baghdadi, N., Belaud, G., Zribi, M., Cheviron, B., Courault, D., Hagolle, O., and Charron, F.: Irrigated Grassland Monitoring Using a Time Series of TerraSAR-X and COSMO-SkyMed X-Band SAR Data., Remote Sens., 6, 10 002–10 032, 2014.

Imaoka, K., Kachi, M., Kasahara, M., Ito, N., Nakagawa, K., and Oki, T.: Instrument performance and calibration of AMSR-E and AMSR2, Remote Sensing and Spatial Information Sciences, 38., 2000.

Kerr, Y., Waldteufel, P., Richaume, P., Wigneron, J., Ferrazzoli, P., Mahmoodi, A., Al Bitar, A., Cabot, F., Gruhier, C., Juglea, S., Leroux, D., Mialon, A., and Delwart, S.: The SMOS Soil Moisture Retrieval Algorithm, IEEE Transactions on Geoscience and Remote Sensing, 50, 1384–1403, https://doi.org/10.1109/TGRS.2012.2184548, 2012.

Kerr, Y. H., Waldteufel, P., Wigneron, J. P., Martinuzzi, J., Font, J., and Berger, M.: Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission, IEEE Transactions on Geoscience and Remote Sensing, 39, 1729–1735, https://doi.org/10.1109/36.942551, 2001.

Kim, S., Liu, Y. Y., Johnson, F. M., Parinussa, R. M., and Sharma, A.: A global comparison of alternate AMSR2 soil moisture products: Why do they differ, Remote Sensing of Environment, 161, 43–62, 2015.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C. T., Kanae, S., Kowalczyk, E., Lawrence, D., Liu, P., Lu, C.-H., Malyshev, S., McAvaney, B., Mitchell, K., Mocko, D., Oki, T., Oleson, K., Pitman, A., Sud, Y. C., Taylor, C. M., Verseghy, D., Vasic, R., Xue, Y., and Yamada, T.: Regions of strong coupling between soil moisture and precipitation, Science, 305, 1138–1140, 2004.

Massari, C., Modanesi, S., Dari, J., Gruber, A., De Lannoy, G., Girotto, M., Quintana-Seguí, P., Le Page, M., Jarlan, L., Zribi, M., and Ouaadi, N.: A review of irrigation information retrievals from space and their utility for users, Remote Sensing, 13(20), 4112, 2021.

Merlin, O., Jacob, F., Wigneron, J. P., Walker, J., and Chehbouni, G.: Multidimensional Disaggregation of Land Surface Temperature Using High-Resolution Red, Near-Infrared, Shortwave-Infrared, and Microwave-L Bands, Geoscience and Remote Sensing, IEEE Transactions on, 50, 1864–1880, https://doi.org/10.1109/TGRS.2011.2169802, 2012.

Moran, M., Hymer, D., Qi, J., and E Sano, E.: Soil moisture evaluation using multi-temporal synthetic aperture radar (SAR) in semiarid rangeland., Agric. For. Meteorol., 105, 69–80, 2000.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., et al.: ERA5-Land: A state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349–4383, 2021.

Ojha, N., Merlin, O., Suere, C., and Escorihuela, M. J.: Extending the spatio-temporal applicability of DISPATCH soil moisture downscaling algorithm: A study case using SMAP, MODIS and Sentinel-3 data, Frontiers in Environmental Science, 9, 2021.

O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., and Bindlish, R.: SMAP L2 Radiometer Half-Orbit 36 km EASE-Grid Soil Moisture, Version 5., NASA National Snow and Ice Data Center Distributed Active Archive Center, 2018.

O'Neill, P. E., Chan, S., Njoku, E. G., Jackson, T., Bindlish, R., and Chaubell, J.: SMAP L3 Radiometer Global Daily 36 km EASE-Grid Soil Moisture, Version 6., NASA National Snow and Ice Data Center Distributed Active Archive Center, 2019.

Paulik, C., Dorigo, W., Wagner, W., and Kidd, R.: Validation of the ASCAT Soil Water Index using in situ data from the International Soil Moisture Network., International journal of applied earth observation and geoinformation, 30, 1–8, 2014.

Peng, J., Albergel, C., Balenzano, A., Brocca, L., Cartus, O., Cosh, M. H., and Loew, A.: A roadmap for high-resolution satellite soil moisture applications–confronting product characteristics with user requirements., Remote Sensing of Environment, 252, 112 162, 2020.

Pierdicca, N., Pulvirenti, L., and Bignami, C.: Soil moisture estimation over vegetated terrains using multitemporal remote sensing data., Remote Sens. Environ., 114, 440–448, 2010.

Plummer, S., Lecomte, P., and Doherty, M.: The ESA Climate Change Initiative (CCI): A European contribution to the generation of the
    Global Climate Observing System, Remote Sensing of Environment, 203, 2–8, 2017.

Rodell, M., Houser, P. R., U, J., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin,
    J. K., Walker, J. P., Lohmann, D., and Tol, D.: The global land data assimilation system, American Meteorological Society, pp. 381–394,
    2004.

Rodríguez-Fernández, N., de Rosnay, P., Albergel, C., Richaume, P., Aires, F., Prigent, C., and Kerr, Y.: SMOS neural network soil moisture
    data assimilation in a land surface model and atmospheric impact., Remote Sensing, 11(11), 1334, 2019.

Rodríguez-Fernández, N. J., Muñoz-Sabater, J., Richaume, P., Albergel, C., de Rosnay, P., Kerr, Y. H., Drusch, M., and Mecklenburg, S.:
    SMOS near real time soil moisture product: Processor overview and first validation results, Hydrology and Earth System Sciences, 21,
    5201–5216,, 2017.

Smith, A. B., Walker, J. P., Western, A. W., Young, R. I., Ellett, K. M., Pipunic, R. C., Grayson, R. B., Siriwardena, L., Chiew, F. H. S., and
    H, R.: The Murrumbidgee soil moisture monitoring network data set, Water Resour. Res., 48, https://doi.org/10.1029/2012WR011976,
    2012.

Srivastava, H., Patel, P., Sharma, Y., and Navalgund, R.: Large-Area Soil Moisture Estimation Using Multi-Incidence-Angle RADARSAT-1
    SAR Data., IEEE Trans. Geosci. Remote Sens., 47, 2528–2535, 2009.

Tomer, S. K., Al Bitar, A., Sekhar, M., Zribi, M., Bandyopadhyay, S., and Kerr, Y.: MAPSM: A spatio-temporal algorithm for merging soil
    moisture from active and passive microwave remote sensing, Remote Sensing, 8, 990, 2016.

Ulaby, F. T., Moore, R. K., and Fung, A. K.: Microwave Remote Sensing Active and Passive-Volume III: From Theory to Applications,
    Artech House, Inc, 1986.

Van der Schalie, R., Kerr, Y., Wigneron, J., Rodríguez-Fernández, N., Al-Yaari, A., and de Jeu, R.: Global SMOS soil moisture retrievals
    from the land parameter retrieval model, International Journal of Applied Earth Observation and Geoinformation, 45, 125–134, 2016.

Van der Schalie, R., de Jeu, R., Kerr, Y., Wigneron, J., Rodríguez-Fernández, N., Al-Yaari, A., Parinussa, R., Mecklenburg, S., and Drusch,
    M.: The merging of radiative transfer based surface soil moisture data from SMOS and AMSR-E, Remote Sensing of Environment, 189,
    180–193, 2017.

Vreugdenhil, M., Dorigo, W. A., Wagner, W., De Jeu, R. A., Hahn, S., and Van Marle, M. J.: Analyzing the vegetation parameterization in
    the TU-Wien ASCAT soil moisture retrieval, IEEE Transactions on Geoscience and Remote Sensing, 54, 3513–3531, 2016.

Wagner, W.: Soil Moisture Retrieval From ERS Scatterometer Data, Ph.D. thesis, Vienna University of Technology, 1998.

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

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa-Saldaña, J., de Rosnay, P., Jann, A., Schneider, S., et al.: The
    ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications, Meteorologische Zeitschrift,
    22, 5–33, 2013.

Young, R., Walker, J., Yeoh, N., Smith, A., Ellett, K., Merlin, O., and Western, A.: Soil moisture and meteorological observations from the
    murrumbidgee catchment., Department of Civil and Environmental Engineering, The University of Melbourne., 2008.

Şekertekin, A., Marangoz, A., and Abdikan, S.: Soil Moisture Mapping Using Sentinel-1A Synthetic Aperture Radar Data., Int. J. Environ.
    Geoinform., 5, 178–188, 2018.