The global SMOS Level 3 daily soil moisture and brightness temperature maps

Ahmad Al Bitar1,2, Arnaud Mialon1,2, Yann H. Kerr1,3, François Cabot1,3, Philippe Richaume1, Elsa Jacquette3, Arnaud Quesney4, Ali Mahmoodi1, Stéphane Tarot5, Marie Parrens1, Amen Al-Yaar6, Thierry Pellarin7, Nemesio Rodriguez-Fernandez1, and Jean-Pierre Wigneron6

1Centre d’Etudes Spatiales de la Biosphère, Université de Toulouse, CNES/CNRS/IRD/UPS, Toulouse, France
2Centre National de Recherche Scientifique, Paris, France
3Centre National d’Etudes Spatiales, Paris, France
4CapGemini Sud, 109 Avenue du Général Eisenhower, 31000 Toulouse, France
5Ifremer, BP 70, 29280 Plouzane, France
6INRA, UMR1391 ISPA, Villenave d’Ornon, France
7IGE, University Grenoble Alpes, CNRS/G-INP/IRD/UGA, Grenoble, France

Correspondence to: Ahmad Al Bitar (ahmad.albitar@cesbio.cnes.fr)

Received: 4 January 2017 – Discussion started: 10 February 2017
Revised: 18 April 2017 – Accepted: 19 April 2017 – Published: 6 June 2017

Abstract. The objective of this paper is to present the multi-orbit (MO) surface soil moisture (SM) and angle-binned brightness temperature (TB) products for the SMOS (Soil Moisture and Ocean Salinity) mission based on a new multi-orbit algorithm. The Level 3 algorithm at CATDS (Centre Aval de Traitement des Données SMOS) makes use of MO retrieval to enhance the robustness and quality of SM retrievals. The motivation of the approach is to make use of the longer temporal autocorrelation length of the vegetation optical depth (VOD) compared to the corresponding SM autocorrelation in order to enhance the retrievals when an acquisition occurs at the border of the swath. The retrieval algorithm is implemented in a unique operational processor delivering multiple parameters (e.g. SM and VOD) using multi-angular dual-polarisation TB from MO. A subsidiary angle-binned TB product is provided. In this study the Level 3 TB V310 product is showcased and compared to SMAP (Soil Moisture Active Passive) TB. The Level 3 SM V300 product is compared to the single-orbit (SO) retrievals from the Level 2 SM processor from ESA with aligned configuration. The advantages and drawbacks of the Level 3 SM product (L3SM) are discussed. The comparison is done on a global scale between the two datasets and on the local scale with respect to in situ data from AMMA-CATCH and USDA ARS Watershed networks. The results obtained from the global analysis show that the MO implementation enhances the number of retrievals: up to 9% over certain areas. The comparison with the in situ data shows that the increase in the number of retrievals does not come with a decrease in quality, but rather at the expense of an increased time lag in product availability from 6 h to 3.5 days, which can be a limiting factor for applications like flood forecast but reasonable for drought monitoring and climate change studies. The SMOS L3 soil moisture and L3 brightness temperature products are delivered using an open licence and free of charge using a web application (https://www.catds.fr/sipad/). The RE04 products, versions 300 and 310, used in this paper are also available at ftp://ext-catds-cpdc:catds2010@ftp.ifremer.fr/Land_products/GRIDDED/L3SM/RE04/.

Published by Copernicus Publications.
1 Introduction

Surface soil moisture (SM) is a control physical parameter for many hydrological processes like infiltration, runoff, precipitation and evaporation (Koster et al., 2004). Estimates of SM are needed for many applications concerned with monitoring droughts (Keyantash and Dracup, 2002), floods (Brocca et al., 2010; Lievens et al., 2015), weather forecast (Drusch, 2007; de Rosnay et al., 2013), climate (Jung et al., 2010) and agriculture (Guérif and Duke, 2000). It is identified among the 50 Essential Climate Variables (ECVs) for the Global Climate Observing System (GCOS). It has also been selected for the creation of decadal (10 years) time series from remote sensing in the ESA Climate Change Initiative (CCI) project (Hollmann et al., 2013).

SM can be obtained from several Earth observation (EO) techniques ranging from visible to microwave wavelengths using active (Ulaby et al., 1996) and passive (Kerr and Njoku, 1990) instruments. Retrieval of SM from passive microwave sensors is a challenging task because features like surface heterogeneity (water surfaces and land use), vegetation cover (vegetation density and distribution), climatic conditions (freezing and snow), acquisition configurations (angle, frequency and polarisation) and topography (multiple scattering) need to be carefully considered while upscaling to the sensor coarse resolution. Several approaches like regression models (Njoku et al., 2003; Wigneron et al., 2004; Saleh et al., 2006), statistical and contextual methods (Verhoest et al., 1998), neural networks (Liu et al., 2002; Rodríguez-Fernández et al., 2015), and radiative-transfer-based approaches (Kerr and Njoku, 1990; Wigneron et al., 2007; Owe et al., 2008; O’Neill et al., 2015) have been developed to retrieve SM based on the sensor frequency, acquisition modes and richness of information (multi-angular, full polarisation and active). The Soil Moisture and Ocean Salinity (SMOS) mission of ESA (Kerr et al., 2001, 2010) with contributions from Centre National d’ Études Spatiales (CNES) in France and Centro para el Desarrollo Tecnológico Industrial (CDTI) in Spain is the first Earth observation mission dedicated to SM mapping. The SMOS Level 2 (L2) SM retrieval algorithm (Kerr et al., 2012) minimizes the squared differences between L-MEB (Wigneron et al., 2007) forward simulations of multi-angular dual-polarisation TB and corresponding SMOS measurements using the Levenberg–Marquardt optimisation algorithm to retrieve physical parameters, mainly SM and VOD.

The L-MEB radiative transfer model is based on the optical depth single-scattering albedo (τ→0) emission model (Mo et al., 1982) combined with specific parameterisations to take into account the impact of vegetation and soil roughness on polarisation mixing and angular signature. The Soil Moisture Active Passive (SMAP) mission, launched by NASA in January 2015, delivers TB observations at a fixed (40°) incidence angle (Entekhabi et al., 2010). The SMAP soil moisture processor currently relies on a single-channel algorithm (SCA) (O’Neill et al., 2015) for its main product. This algorithm uses a forced VOD in a single-orbit configuration. Mienecki et al. (2014) and Wigneron et al. (2017) presented a review and a comparison of the different retrieval approaches for L-band microwave from EO missions (SMOS, SMAP and AQUARIUS).

Passive microwave sensors have a high revisit frequency: 1 day for Advanced Microwave Scanning Radiometer - Earth Observing System (AMSR-E) (Njoku and Entekhabi, 1996) and 2–3 days for SMOS and SMAP. In this study the multi-orbit (MO), multi-angular and dual-channel horizontal and vertical (H/V) operational retrieval algorithm implemented at the CATDS (Centre Aval de Traitement des Données SMOS) by CNES is presented. Retrieval using temporal series is becoming increasingly common in operational EO retrieval algorithms for optical and to some extent microwave technologies. Some examples in the optical domain are the correction of aerosol impact for visible images (Hagolle et al., 2008, 2015), cloud detection (Hagolle et al., 2010) and the use of MO for land cover classification (Inglada and Mercier, 2007). The previous methodologies are being implemented for high-end level 2-A and level 3 products for the Copernicus Sentinel-2 mission. The use of MO in the radar community is a standard approach. The SM retrievals from ERS (European Remote Sensing), Advanced Scatterometer (ASCAT), RADARSAT-2 and Sentinel-1 are based on a change detection algorithm (Wagner et al., 1999, 2013; Naeimi et al., 2009). Similarly, Mattia et al. (2006) introduced a priori surface parameters and multi-temporal synthetic aperture radar (SAR) data to reduce the impact of vegetation and soil roughness in SM retrieval from SAR. Recently, a generalisation of change detection to multiple regression using cumulative distribution function (CDF) transformations was applied to RADARSAT-2 time series data and validated over the Bebrambi watershed, South India (Tomer et al., 2015). In microwave radiometry, Konings et al. (2016) presented a time series retrieval of vegetation optical depth based on AQUARIUS L-band acquisitions.

Here a detailed presentation of the products and retrieval algorithm of an inter-comparison between the SMOS SO (single-orbit) and the SMOS MO (multi-orbit) operational products is given. More specifically, the objective of this paper is to present the daily L3 SM and TB V310 products and associated algorithms and to compare the SMOS MO level 3 retrievals to the level 2 single-orbit operational retrievals obtained using V600 L1 ESA-SMOS products. Since the SMOS mission launch in November 2009, this is the first reprocessing to have an aligned version of the processors from Level 1 up to Level 3, enabling a direct comparison of the products. In the next sections the MO retrieval SM algorithm and the L3 TB are presented. The datasets used for the assessment, the results of the comparison and conclusions are presented.
2 The CATDS Level 3 soil moisture processor

2.1 Algorithm overview

The Level 3 SM (L3SM) processor consists of a set of several algorithms. The forward model in L3SM uses the same physically based forward models as the ESA SMOS Level 2 SM processor, but in a MO retrieval context. A short summary of the main features of this processor is provided here and a detailed description can be found in Kerr et al. (2012). The SMOS L2 retrieval can be divided into two main components:

1. The first component is a physical model that computes TB at the antenna reference frame forced by ancillary data (land classification and soil properties) and physical parameters (skin or near-surface temperature and soil temperature). The selected physical model for the SMOS mission is L-MEB from Wigneron et al. (2007). The main features of the L-MEB physical model implementation in the SMOS operational processor are as follows:
   - Effective scattering albedo is considered.
   - SM and VOD are jointly retrieved over nominal (bare soil and low vegetation) surfaces using angular signature and polarisation information.
   - Dual polarisation is used. Full polarisation data are only used to take into account the Faraday rotation and geometric rotation to transform modelled TB from the top of atmosphere (TOA) to the antenna reference frame.
   - The mean antenna pattern (Kerr et al., 2012) is used in the iterative retrieval algorithm. The mean weighting function expresses the average contributions for all angular acquisitions. The $-3 \text{ dB}$ footprints are about 20 km in radius. This corresponds to the nominal resolution of the synthetic aperture. This also corresponds to 86% of the signal if a homogeneous surface is considered (Al Bitar et al., 2012).
   - Surface heterogeneity is considered through aggregated TB contributions from $4 \times 4 \text{ km}^2$ surface units. The contributions are then convoluted with the mean antenna pattern. A total area of $125 \times 125 \text{ km}^2$ is considered at each retrieval node to compute the total emissions.
   - Dynamic changes in surface state (freezing, rainfall, etc.) are considered through the use of ancillary weather data from ECMWF reanalysis products. Since the mission launch, many improvements have been implemented in the operational processing model. Some examples include, for instance, the improved parametrization of the forest albedo in Rahmoune et al. (2014) or the choice of dielectric mixing models in Mialon et al. (2015).

2. The second component of the retrieval algorithm is an iterative optimisation scheme that minimises a Bayesian cost function constructed from the observed and the modelled TBs in order to retrieve the physical parameter values. Preprocessing and post-processing steps are implemented to filter the input and output data for undesired effects like the decrease in quality due to spatial sampling or radio frequency interferences (RFIs) (Oliva et al., 2012; Richaume et al., 2014).

The physical approach at Level 3 MO is the same as that of Level 2 SO. In fact the core processing uses the same implementation of the L-MEB radiative transfer model. The main difference in Level 3 is the use of several orbits, rather than one, to retrieve SM and VOD. This has an impact first on the post-processing steps for selecting the orbits and second on the optimisation scheme to retrieve the parameters. Since the Level 2 retrieval is a multi-parameter retrieval, the Level 3 is thus a multi-orbit multi-parameter retrieval. The reasons that motivated the use of the MO approach are the following:
   - The angular sampling and radiometric accuracy at the border of the swath are reduced. Figure 1 shows the cumulative number of records for several descending orbits. The asterisk in each panel represents the same location in the La Plata region in South America. The orange regions inside the orbits observed on 18, 20 and 23 May 2010 depict the mild decrease in the number of TB measurements (15–35) during the instrument calibration phases. However, most important is the low number of TB measurements (35) observed on 21 May when the
point of interest is at the border of the swath. A low number of TB measurements spanning a narrow range of incidence angles generates failures in the iterative retrieval of SM and VOD. The use of MO can help improve the number of successful retrievals at the border of the swath.

- The VOD is expected to vary slowly in time and thus to be highly correlated between two consecutive ascending or descending orbits or over a short period of time (a few days). In fact, at L band the VOD is mainly correlated to vegetation water content (Jackson and Schmugge, 1991), which is expected to vary slowly in time compared with temporal variability in SM.

Other general motivations for Level 3 products are to provide a global gridded product, in contrast to swath-based products and to provide fixed-angle-binned TB products. The 25 km Equal-Area Scalable Earth Grid version 2.0 (EASE-Grid 2.0) (Brodzik and Knowles, 2002), which was selected for the Level 3 MO product also has a spatial sampling closer to the sensor’s nominal resolution. The main input TB for the processing is generated from the snapshot-based L1B products, which are TBs in the Fourier domain. This consists of an inverse fast Fourier transform (IFFT) to make the transition from the Fourier domain to the spatial domain using the L3 EASE-Grid 2.0. In a subsequent step, TB measurements corresponding to the same grid point are selected from the different snapshots (for a given grid point, the incidence angle of the observation is different for each snapshot) to construct a grid-point-based product similar to the ESA L1C TB product but in EASE-Grid 2.0. The alternative is to interpolate the ESA L1C TB dataset from the 15 km Icosahedral Snyder Equal Area (ISEA) grid to the 25 km EASE-Grid 2.0 grid. This option was excluded because it could have generated interpolation artefacts on the TB products that would have propagated through the processing chain.

2.2 Orbit selection

The selection of orbits is needed to select TBs at high latitudes where a sub-daily revisit is available and to generate the time series dataset on the EASE-Grid 2.0 as input to the MO retrieval. The following criteria are applied for the selection of revisits:

- Ascending and descending orbits are processed separately since the impact of RFI (Oliva et al., 2012) and sun corrections (Khazâal et al., 2016) between ascending and descending orbits are very different.

- TB products are filtered at high latitudes where more than one revisit per day occurs (latitudes above 60° N and 60° S). A maximum of one revisit per day is considered. The selection criterion is the minimum distance from the centre of the swath because the radiometric accuracy and resolution is best at the centre. This criterion is applied for each grid node individually.

At this level the acquisitions for a given day for ascending and descending orbits are separately stored in a three-dimensional matrix accounting for snapshots, longitude and latitude. A snapshot is an image associated to the acquisition of SMOS during a given integration time (epoch). Snapshots have different epochs and polarisation following a preprogrammed acquisition sequence. From this product a fixed-angle-binned TB product is generated as presented in Sect. 3. The product is also used in the next processing steps of L3SM MO.

- For each retrieval and over each node a 7-day period is considered in which three revisits are selected from the complete list of revisits (Fig. 2). The first coincides with the central date (date of main product). The two others correspond to selected dates either before (previous 3.5 days) or after (3.5 days posterior) the considered date. Like in the previous processing step, the selection is done based on minimum distance from the swath centre for each node.

2.3 Cost function and retrieval

Observed TBs at the antenna reference frame from the precedent, actual and succeeding dates are assembled for each node. The forward algorithm is run to generate the modelled TB for each of the TB dataset records. The ancillary data and parameters are independently considered for each record. A Bayesian cost function that includes the aforementioned MO observed TB and modelled TB is then constructed. This is achieved by incorporating in the retrieval approach a temporal autocorrelation function for the VOD. The cost function
is as follows:

$$\text{Cost} = (TB_M - TB_F)^T \cdot \text{COV}_{TB}^{-1} \cdot (TB_M - TB_F) + \sum_P (P - P_0)^T \cdot \text{COV}_P^{-1} \cdot (P - P_0),$$

(1)

where \(\text{COV}_{TB} = \sigma^2_{TB}\) is the error covariance matrix of TB data when assuming no auto-temporal correlation, \(TB_M\) is the measured TB from SMOS, \(TB_F\) is the forward modelled TB using L-MEB, \(P\) is the vector of retrieved parameters (SM and VOD) at the three times of acquisition, \(\text{COV}_P\) is the error covariance matrix for parameter \(P\) and \(P_0\) is the a priori value of parameter \(P\).

It is important to note that three SM values are retrieved simultaneously at each node: SM\(_P\) for the preceding date, SM\(_A\) for the actual date and SM\(_F\) for the succeeding date. The same applies to VOD. In the case of SM, the a priori values are retrieved from ECMWF reanalysis data.

Where \(P = [SM_P, SM_A, SM_F]\), the error covariance matrix considering no cross- or autocorrelation is given by

$$\text{COV}_{SM} = \sigma^2_{SM0} \cdot I,$$

(2)

where \(\sigma^2_{SM0}\) is the standard-deviation error associated with SM. It is set to a high value: 0.7 m\(^3\) m\(^{-3}\). \(I\) is the \((3 \times 3)\) identity matrix.

When \(P\) is equal to VOD, the error covariance matrix, considering temporal autocorrelation and no cross correlation between the different parameters, is given by

$$\text{COV}_{VOD} = \sigma^2_{VOD0} \begin{bmatrix} 1 & ... & ... \\ \rho(t_P, t_A) & 1 & ... \\ ... & ... & ... \\ \rho(t_P, t_F) & \rho(t_A, t_F) & 1 \end{bmatrix}.$$

(3)

where \(\sigma^2_{VOD0}\) is the standard-deviation error associated with VOD, and \(\rho\) is the correlation function modelled assuming a Gaussian autocorrelation distribution:

$$\rho_{VOD}(t_1, t_2) = \rho_{max}(t_1, t_2) \cdot \exp\left(-\frac{(t_1 - t_2)^2}{T_c^2}\right),$$

(4)

where \(t_1\) and \(t_2\) are the times (expressed in days) corresponding to the VOD retrieval dates (P, A or F), \(\rho_{max}(t_1, t_2)\) is the maximum amplitude of the correlation function between \(t_1\) and \(t_2\) and \(T_c\) is the characteristic correlation time for VOD (\(T_c = 30\) days for forests and \(T_c = 10\) days for low vegetation).

Figure 3 shows the shape of the correlation function for the two correlation lengths used in the processing. The green curve corresponds to the forested surfaces and the blue one to the nominal surfaces (bare soil and low vegetation).

The parameter values namely (SM\(_P\), SM\(_A\), SM\(_F\), VOD\(_P\), VOD\(_A\) and VOD\(_F\)) are retrieved by minimising the cost function in an iterative procedure using the Levenberg–Marquardt optimisation algorithm. Thus, at the end of each daily retrieval, three SM values are available. The retrieval associated with the best goodness of fit (\(X^2\)) value is then selected and delivered in the 1-day product. This product is only available when the filtering is finished, and thus with 7 days of lag time. Using the daily maps, time synthesis products (3 days, 10 days and monthly) are then provided. A detailed description of the algorithm is presented in the CATDS L3 Algorithm Theoretical Basis Document (Kerr et al., 2013).

### 3 The CATDS Level 3 angle-binned TB processor

The objective of this algorithm is to generate a product containing fixed-angle full-polarisation brightness temperatures at top of atmosphere (TOA) but with the polarisations expressed in the ground reference frame (horizontal and vertical components) over the EASE-Grid 2.0. The main input for this algorithm is the snapshot dataset mentioned in the previous section. The algorithm consists of four steps: (a) filtering, (b) interpolation, (c) reference frame transformation and (d) angle binning. However, note that before being projected to a ground reference frame, the data are processed in the instrument reference frame. Thus, TBs are labelled TB\(_H\) and TB\(_V\) to express that the polarisations are at satellite level, while once processed they will be provided in the ground reference frame and will be labelled TB\(_H\) and TB\(_V\).

#### 3.1 TB filtering

The filtering eliminates brightness temperatures that are impacted by anthropogenic effects (such as RFIs), or spurious effects (such as sun impact). The filtering criteria, shown in Table 1, are similar to those for L3 MO SM and L2 SO retrievals. A detailed description of the filtering criterion is provided in the SMOS L2 ATBD (Algorithm Theoretical Basis Document). The reader can refer to Khazaal et al. (2016) for a more detailed evaluation of the impact of sun corrections and Richaume et al., 2014 and Soldo et al., 2014 for the im-
Table 1. List of applied filtering criterion used on brightness temperature products prior to interpolation.

<table>
<thead>
<tr>
<th>Filtering criteria</th>
<th>Applied test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thresholds</td>
<td>50 K &lt; TBX and TBY &lt; 340 K</td>
</tr>
<tr>
<td></td>
<td>−50 K &lt; TB_{xy} &lt; +50 K</td>
</tr>
<tr>
<td>Amplitude</td>
<td>50 K &lt; \sqrt{TB_X^2 + TB_Y^2} &lt; 500 K</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>TB - 2 · ATB &lt; TB &lt; TB + 2 · ATB</td>
</tr>
<tr>
<td>First Stokes</td>
<td>ST1 = (ST1 over each dwell line (angular elements):</td>
</tr>
<tr>
<td></td>
<td>ear interpolation is used to compute the missing</td>
</tr>
<tr>
<td></td>
<td>acquisitions when there is no complete dataset for</td>
</tr>
<tr>
<td></td>
<td>any epoch. A weighted linear cross interpolation</td>
</tr>
<tr>
<td></td>
<td>is performed (each epoch an acquisition can be</td>
</tr>
<tr>
<td></td>
<td>co-polarised (X, Y) moved.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>X discarded, the associated cross polarisation is</td>
</tr>
<tr>
<td></td>
<td>discarded. In case a cross polarisation is</td>
</tr>
<tr>
<td></td>
<td>active the flag means that the pixel is located in</td>
</tr>
<tr>
<td></td>
<td>a zone where a sun alias was reconstructed (after</td>
</tr>
<tr>
<td></td>
<td>sun removal, measurement may be degraded). The</td>
</tr>
<tr>
<td></td>
<td>spatial resolution eliminates records that are</td>
</tr>
<tr>
<td></td>
<td>impacted by aliasing (only alias-free field of view</td>
</tr>
<tr>
<td></td>
<td>considered). The sun tail is considered when the</td>
</tr>
<tr>
<td></td>
<td>pixel is located in the hexagonal alias directions</td>
</tr>
<tr>
<td></td>
<td>centred on a sun alias.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spatial resolution</td>
</tr>
<tr>
<td></td>
<td>S$\text{SMEF}$ &lt; (55 × 55) km$^2$</td>
</tr>
<tr>
<td></td>
<td>$L_{\text{maj}}/L_{\text{mi}} &lt; 1.5$</td>
</tr>
<tr>
<td></td>
<td>B\text{ORDER FOV} (flag is off)</td>
</tr>
<tr>
<td>RFI</td>
<td>L1A STRONG RFI (flag is off)</td>
</tr>
<tr>
<td></td>
<td>L1B STRONG RFI (flag is off)</td>
</tr>
<tr>
<td></td>
<td>POINT SOURCE RFI (flag is off)</td>
</tr>
<tr>
<td></td>
<td>TAILS RFI (flag is off)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sun correction</td>
</tr>
<tr>
<td></td>
<td>SUN_POINT (flag is off)</td>
</tr>
<tr>
<td></td>
<td>SUN_TAILS (flag is off)</td>
</tr>
</tbody>
</table>

ATB is the radiometric accuracy of SMOS TB, ST1 is the first Stokes parameter, ST is the average of ST1 over each dwell line (angular signature), ST4 is the fourth Stokes parameter, SMEF is the area of the half-maximum contour of the mean synthetic antenna pattern, $L_{\text{maj}}$ is the length of the major axis of the synthetic antenna pattern and $L_{\text{mi}}$ is the length of the minor axis of the synthetic antenna pattern.

3.2 TB interpolation

The acquisition sequence of SMOS is shown in Table 2. At each epoch an acquisition can be co-polarised (X, Y) or combined cross (X, Y, X) and co-polarised. The table shows that there is no complete dataset for any epoch. A weighted linear interpolation is used to compute the missing acquisitions based on adjacent ones.

The weighting function accounts for the two following elements:

- The TB acquisitions have different accuracy levels since the integration time is longer when only co-polarisation is acquired (pure acquisition) when compared to the case where combined cross and co-polarisation are acquired.

- The time span between two acquisitions in the same mode is not constant. Acquisitions closer in time are considered more reliable than farther ones, taking into consideration that the synthetic antenna weighting function rotates and that the incidence angle changes.

The time interpolation function of TB at time $i$ ($TB_i$) is as follows:

$$TB_i = \frac{W_{i-1} \cdot TB_{i-1} + W_{i+1} \cdot TB_{i+1}}{W_{i-1} + W_{i+1}},$$

with

$$W_{i-1} = \frac{\sigma_{i-1} \cdot \text{nb}_e\text{p}_o_{i-1}}{\sigma_{i+1} \cdot \text{nb}_e\text{p}_{i+1}},$$

where $\text{nb}_e\text{p}_o$ is the number of epochs between acquisitions at time $i$, $\sigma$ is the corresponding radiometric accuracy and $W_i$ is the weighting coefficient at time $i$. The standard deviation of the interpolated field is computed based on the square root of the weighted variances of the adjacent acquisition. We assume that the acquisitions are not correlated; therefore, no cross correlation term is considered in the equation. The following formulation is used:

$$\sigma_i = \sqrt{\left(Q_{i-1} \cdot \sigma_{i-1}\right)^2 + \left(Q_{i+1} \cdot \sigma_{i+1}\right)^2} \cdot \frac{1}{\text{nb}_e\text{p}_o},$$

$$Q_i = \frac{\sigma_i}{\text{nb}_e\text{p}_o}.$$

The same approach as Eq. (5), while applying a constant weight, is used to compute the interpolated values of auxiliary information such as major and minor semi-axis length, incidence angle, Faraday angle and geometric angle.

3.3 Transformation from antenna to ground reference frame

In this step, the TBs are transformed from the antenna reference frame ($X$, $Y$) to the ground reference frame ($H$, $V$). This is done without accounting for atmospheric and galactic contributions. They are considered as TOA TBs. The TB components at antenna reference frame exhibit polarisation mixing due to the geometry of the acquisition (Fig. 4). Faraday rotation will also slightly alter the polarisations.

The inverse of the rotation matrix is used to transform the TB data from antenna to ground reference frame:

$$\begin{pmatrix} TB_H \\ TB_V \end{pmatrix} = \text{IRM} \begin{pmatrix} TB_X \\ TB_Y \end{pmatrix} = \begin{pmatrix} \cos^2 a & \sin^2 a & \cos a \cdot \sin a \\ \sin^2 a & \cos^2 a & -\cos a \cdot \sin a \\ -\sin 2a & \sin 2a & \cos 2a \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

where $TB_3$ and $TB_4$ are the Stokes 3 and Stokes 4 components. The inverse of rotation matrix (IRM) is given by

$$\text{IRM} = \begin{pmatrix} \cos^2 a & \sin^2 a & \cos a \cdot \sin a \\ \sin^2 a & \cos^2 a & -\cos a \cdot \sin a \\ -\sin 2a & \sin 2a & \cos 2a \end{pmatrix}.$$
Table 2. Acquisition sequences of SMOS in full polarisation mode (capital letters are used for pure acquisition).

<table>
<thead>
<tr>
<th>Snapshot number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB (real/imaginary)</td>
<td>X/Y</td>
<td>Y/X</td>
<td>X/Y</td>
<td>X/Y</td>
<td>Y/X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TB (co-polarisation)</td>
<td>X</td>
<td>X</td>
<td>Y</td>
<td>y</td>
<td>X</td>
<td>x</td>
<td>Y</td>
<td>X</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. Transformation from antenna (S) to ground reference frame (G). $\omega_f$ is the Faraday rotation angle and $\Theta_g$ is the geometric rotation angle (adapted from SMOS L2 ATBD).

where

\[ a = \Theta_g + \omega_f, \]  

with $\Theta_g$ being the geometric angle and $\omega_f$ being the Faraday rotation angle as shown in Fig. 4.

The accuracies of the TB data are then computed by propagating the accuracies using the matrix below:

\[
\begin{align*}
\sigma TB_H &= (IRM_{2,1}^2 \cdot \sigma TB_{X}^2 + IRM_{3,2}^2 \cdot \sigma TB_{Y}^2)
+ 4 \cdot (IRM_{1,3}^2 + IRM_{1,4}^2) \cdot \sigma TB_{XY}^0.5 \\
\sigma TB_V &= (IRM_{2,1}^2 \cdot \sigma TB_{X}^2 + IRM_{1,4}^2 \cdot \sigma TB_{Y}^2)
+ 4 \cdot (IRM_{1,3}^2 + IRM_{2,4}^2) \cdot \sigma TB_{XY}^0.5 \\
\sigma TB_3 &= (IRM_{2,1}^2 \cdot \sigma TB_{X}^2 + IRM_{1,4}^2 \cdot \sigma TB_{Y}^2)
+ 4 \cdot (IRM_{1,3}^2 + IRM_{2,4}^2) \cdot \sigma TB_{XY}^0.5 \\
\sigma TB_4 &= (IRM_{2,1}^2 \cdot \sigma TB_{X}^2 + IRM_{1,4}^2 \cdot \sigma TB_{Y}^2)
+ 4 \cdot (IRM_{1,3}^2 + IRM_{2,4}^2) \cdot \sigma TB_{XY}^0.5
\end{align*}
\]  

where $IRM_{i,j}$ are the $i$th column and $j$th row components of the IRM matrix.

3.4 Angle binning

This step consists in averaging the TOA TBs at fixed-angle intervals using an arithmetic mean. The selected incidence angle bins, shown in Table 3, are designed to also cover the SMAP acquisition angle ($40^\circ$).

All TB values outside the interval defined by mean (TB) ± 2 SD (TB) are considered as outliers and removed from the binning. The SD (TB) corresponds to the standard deviation of TB values inside each angle bin, not to be confused with the radiometric accuracy. The filtered outlier values are mainly associated with low RFI effects. If one component of TB ($TB_H$, $TB_V$ and $TB_{HV}$) is filtered out, all the other components are disregarded.

4 Datasets

4.1 Remote sensing datasets

4.1.1 SMOS CATDS Level 3 soil moisture products

The CATDS Level 3 user data products (CLF3UA/D) are MO soil moisture retrieval products. They contain 1-day global maps of geophysical parameters (SM, VOD, imaginary and real part of the dielectric constant, etc.) retrieved as described above, processing parameters (percentage of forest cover, choice of physical model, etc.) and quality indicators (probability of RFI, goodness of fit between modelled TB from L-MEB and observed TB $X^2$, etc.) over continental surfaces for ascending and descending orbits separately. They are in the netCDF format over the EASE-Grid 2.0 25 km and generated at the Institut Français de Recherche pour l’Exploitation de la Mer (IFREMER) for CNES and distributed via the CATDS web portal (http://www.catzds.fr) and ftp server. The operational production of L3SM started in 2010 and it is currently ongoing. The time span used in this study covers 2010–2015 for the global maps and 2010–2016 for the time series analysis. The user has access to the latest versions of the products from reprocessing and operational processing. The current study uses the latest data corresponding to reprocessing RE04, which uses CATDS V300 corresponding to ESA V620 Levels 1 and 2. It is the first simultaneous Level 2 and Level 3 reprocessing campaign since the start of the mission. Previous versions of the L3SM products where compared to soil moisture products from AMSR-E (Al-Yaari et al., 2014a) and ASCAT (Al-Yaari et al., 2014b) missions, but this is the first comparison enabling an aligned configuration of the L2SM SO and L3SM MO. It has homogenised inputs (L1B/C) and physical parameterization. It uses the Mironov model to relate soil liquid water content with the effective permittivity of the ground (Mialon et al., 2015), enhanced forest parameterization for albedo (Rahmoune et al., 2014), enhanced global soil texture map consistent with the one used for the SMAP mission and the latest RFI detection tech-
techniques (Richaume et al., 2014). It also uses the latest (V620) brightness temperature products at Level 1B. The SM maps, RFI probabilities and mean forest cover are extracted in the present study from the L3 product.

The mean forest cover provides the percentage of forest cover, taking into account the mean antenna pattern. It is obtained by convoluting the ECOCLIMAP (Masson et al., 2003) forest cover using the SMOS antenna weighting function at a resolution of 4 km over an area of 125 × 125 km². The RFI map was obtained by averaging the RFI probability field in the L3SM product. This information includes strong RFI and moderate RFI depicted from the SMOS full-polarisation brightness temperatures (Richaume et al., 2014). After extraction, RFI filtering is applied with probability of RFI < 10% and goodness of fit with a probability of $X^2 > 0.95$.

### 4.1.2 SMOS DPGS Level 2 soil moisture product

The ESA L2 Soil Moisture User Data Product (SMUDP; Kerr et al., 2012), which is a SO retrieval product, is used in this study for comparison purposes. This product is a half-orbit swath-based dataset of physical variables (SM, VOD, dielectric constant imaginary and real parts, etc.), processing parameters (percentage of forest cover, type of surface model, etc.) and quality indicators (probability of RFI, $X^2$, etc.) over continental surfaces. Ascending and descending orbits are processed separately in the current configuration. The SMUDP product is delivered in the BinX format over the ISEA discrete global grid (Carr et al., 1997), with a hexagonal partitioning of aperture 4 at a resolution of 9 km known as ISEA4H9. The grid point centres have a fixed separation distance of around 15 km. Products are generated at the ESA SMOS Data Processing Ground Segment (DPGS) and disseminated by ESA via Earth Online. The DPGS and CATDS share the same reprocessing dissemination strategy: the most recent version of the processor is implemented in the operational processing before the end of the reprocessing campaign. Version 620 of SMUDP is used in this study. The selected time span is 2010–2015 for the global analysis and 2010–2016 in the time series analysis.

The main characteristics and differences between the L2SM SO retrieval and L3SM MO retrieval products are summarised in Table 4.

### 4.1.3 SMOS CATDS Level 3 brightness temperature products

The SMOS CATDS full-polarisation angle-binned daily brightness temperature products (CDF3TA/D) version 310, were downloaded from the same database as the L3 MO SM. These products consist of global 1-day maps of full-polarisation TB over fixed-angle bins with their associated accuracies. Detailed computation was described above in Sect. 3. The product also contains auxiliary data like the geometric angles, Faraday angles, length of major semi-axis and length of minor semi-axis. Quality flags are also provided in the product. The $T_{BH}$ and $T_{BV}$ records are extracted for the $40^\circ$ bin. No additional filtering is done over these products.

### 4.1.4 SMAP NSIDC (National Snow and Ice Data Center) L1C brightness temperature

The SMAP mission from NASA was launched in January 2015. It operates like SMOS in L-band using a radiometer and a radar (that was operational for about 80 days). It has a local overpass time at 18:00 UTC and 06:00 UTC for ascending and descending orbits, respectively, but the acquisitions are not necessarily synchronous with SMOS. In this study we use the SMAP TB derived from the radiometer acquisitions. The SMAP L3B_SM_P product is downloaded from the National Snow and Ice Data Center (NSIDC) website (O’Neil et al., 2016). The SMAP L3 TB is used as input for the SM retrievals and it is corrected for water contribution.

### Table 3. Selected incident angle bins.

<table>
<thead>
<tr>
<th>Bin ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bin centre</td>
<td>2.5°</td>
<td>7.5°</td>
<td>12.5°</td>
<td>17.5°</td>
<td>22.5°</td>
<td>27.5°</td>
<td>32.5°</td>
<td>37.5°</td>
<td>40°</td>
<td>42.5°</td>
<td>47.5°</td>
<td>52.5°</td>
<td>57.5°</td>
<td>62.5°</td>
</tr>
<tr>
<td>Bin width</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
<td>5°</td>
</tr>
</tbody>
</table>

### Table 4. Main characteristics of the SMOS Level 3 and Level 2 SM products.

<table>
<thead>
<tr>
<th>Product</th>
<th>L3SM</th>
<th>L2SM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of product</td>
<td>MIR_CLF3A/D</td>
<td>MIR_SMUDP</td>
</tr>
<tr>
<td>Gridding system</td>
<td>EASEv2</td>
<td>ISEA 4H9</td>
</tr>
<tr>
<td>Product sampling</td>
<td>25 km</td>
<td>15 km fixed</td>
</tr>
<tr>
<td>Resolution</td>
<td>SMOS nominal resolution of 40 km</td>
<td></td>
</tr>
<tr>
<td>Multi-parameter retrieval</td>
<td>SM, VOD</td>
<td>SM, VOD</td>
</tr>
<tr>
<td>Angular signature</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Polarisation impact</td>
<td>H/V</td>
<td>H/V</td>
</tr>
<tr>
<td>Multi-orbit</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Forward model</td>
<td>L-MEB (tau omega)</td>
<td></td>
</tr>
<tr>
<td>Availability</td>
<td>3.5–7 days</td>
<td>6 h</td>
</tr>
<tr>
<td>Processing centre</td>
<td>CATDS (CNES)</td>
<td>DPGS (ESA)</td>
</tr>
<tr>
<td>Format</td>
<td>NetCDF</td>
<td>BinX</td>
</tr>
<tr>
<td>Version</td>
<td>V300</td>
<td>V620</td>
</tr>
<tr>
<td>Coverage</td>
<td>Global grid</td>
<td>Swath based</td>
</tr>
</tbody>
</table>
Table 5. Properties of the in situ sites used for the evaluation.

<table>
<thead>
<tr>
<th>Network</th>
<th>Location</th>
<th>Vegetation/climate</th>
<th>Soil texture</th>
<th>Topography</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walnut Gulch watershed</td>
<td>Southeastern Arizona, USA</td>
<td>Brush and grass covered, desert shrubs, rangeland, cattle grazing, semi-arid</td>
<td>Range/sandy loam</td>
<td>Rolling</td>
</tr>
<tr>
<td>Little Washita watershed</td>
<td>Southwest Oklahoma, USA</td>
<td>Rangeland and pasture (63%), winter wheat, subhumid</td>
<td>Range, wheat/silt or sand</td>
<td>Rolling</td>
</tr>
<tr>
<td>AMMA-CATCH network Niger</td>
<td>Niger</td>
<td>South Sahelian climate with semi-arid vegetation and crops (millet, fallows and tiger bush)</td>
<td>sandy loam, 91% sand and 9% clay</td>
<td>–</td>
</tr>
<tr>
<td>AMMA-CATCH network Ouémé</td>
<td>Benin</td>
<td>Sudanian climate with different types of rain systems and Guinean savanna vegetation</td>
<td>77% sand and 19% clay</td>
<td>–</td>
</tr>
</tbody>
</table>

and atmospheric effects. It is provided on the EASE-Grid 2.0 with a 36 km resolution in HDF5 format. The TB_H and TB_V records are extracted for the year 2015. At level 3 processing, only ascending orbits, coinciding with afternoon overpasses at 18:00 UTC and thus SMOS descending overpasses, are available from the SMAP mission.

4.2 In situ datasets

In this study, the SMOS SM products are evaluated against in situ SM from two networks with spatially distributed SM data at the footprint scale (USDA Watershed and AMMA-CATCH). The in situ soil moisture data from probes installed near the surface are used. These sites provide a soil moisture reading, representative of the first 5 cm of the top soil layer, as they are vertically installed. This may lead to a mismatch between the sensor sampling depth and the expected representative depth 0–2 or 0–3 cm of the L-band microwave radiometers (Escorihuela et al., 2010). This mismatch induces errors that are to be considered in conjunction with the sampling errors due to the spatial heterogeneity. The choice of the sites is made to cover contrasting environments over two different continents to provide an overview of the SM processor performances. The statistics over the sites are computed for data available within 1 h of space-borne acquisitions (SMOS and SMAP).

4.2.1 AMMA dataset

The AMMA long-term observing system (AMMA-CATCH, 1996 and 2005) includes three mesoscale sites located in Niger, Benin and Mali that are representative of the West African ecoclimatic gradient (Cappelaere et al., 2009; Mougin et al., 2009). The AMMA-CATCH soil moisture network is a well-established network in terms of satellite product assessment (de Rosnay et al., 2009; Pellarin et al., 2009; Louvet et al., 2015). The Niger and Benin sites are selected for this study. The Niger site, centred at 13.645°N–2.632°E, is mainly composed of tiger bush on the plateau and fallow savannah and pearl millet crop fields on the sandy slopes (Cappelaere et al., 2009). The Benin site, located at 1.5–2.8°E, 9–10.2°N, is mainly composed of woody savannah and tropical forest. Most ground-based instruments are located in the north-western part of the Ouémé catchment (9.745°N–1.653°E). The observed annual rainfall amount was 1578 mm in 2010, 1093 mm in 2011 and 1512 mm in 2012.

4.2.2 USDA – watersheds

The United States Department of Agriculture (USDA) Agricultural Research Service monitors a network of watersheds across the US using a high number of instruments. Surface soil moisture (5 cm) is monitored across the watersheds and has been recorded on an hourly basis since 2002. The USDA provides estimates of the average soil moisture over an area that is approximately the size of a SMOS footprint. Two of the watersheds have been selected for this study: Walnut Gulch (WG), Arizona, USA (Keever et al., 2008), and Little Washita (LW), Oklahoma, USA (Elliott et al., 1993). Soils in WG can be classified as sandy loam. The original datasets are available from https://www.tucson.ars.ag.gov/dap/ for WG and from http://ars.mesonet.org/webrequest/ for LW. Over LW the soil properties are more heterogeneous with a loam, clay and sand textures. Previous studies on calibration and scaling have quantified the uncertainty of the in situ measurements over the sites to be lower than 0.01 m³ m⁻³ when compared to gravimetric measurements. The basin-scale weighted average is based on the Thiessen polygon method and has a standard deviation between 0.05 and 0.10 m³ m⁻³. A detailed description of the site characteristics is provided in Jackson et al. (2010), and details on the averaging procedure are provided in Jackson et al. (2012). This network has been used for validation of remote sensing soil moisture datasets (including SMOS) in many studies.
5 Methodology of evaluation

5.1 Global comparison of SMOS and SMAP TB

In order to compare the SMOS TB product to SMAP TB, the SMOS daily product was averaged following the same interpolation procedure as the one suggested in the SMAP mission. The method consists of using an inverse distance weighting for all the SMOS EASE 2.0 25 km grid points at the limits of the EASE 2.0 36 km grid of the SMAP product. The TB_H and TB_V from SMAP products are extracted and used without modification. The comparison is done over the pixels with a water fraction of less than 0.01 (i.e. 1%) since the SMAP TBs are provided with subtracted open surface water. The contribution of the water surface is computed considering surface fraction from MODIS MOD44W and the emission of water using the Klein–Swift (1977) dielectric constant model forced by the surface soil layer temperature from GSFC (Goddard Space Flight Center) (O’Neil et al., 2015).

5.2 Global soil moisture maps comparison

Global comparison is done over the EASE-Grid 2.0 25 km used for the L3 MO SM product. The L3 MO SM field is extracted directly from the product. The L2 SO SM field is interpolated to the EASE-Grid 2.0 25 km using a three-stage interpolation strategy where the availability of the products inside the limits of the grid node is considered

- bilinear, if more than two soil moisture retrievals are available;
- linear, if two soil moisture retrievals are available;
- nearest point, if one soil moisture retrieval is available.

The L2 SO SM is also filtered at high latitude where several soil moisture retrievals are available. The selection criterion is minimum distance from the swath centre, the same as for the L3 MO SM algorithm.

5.3 Local evaluations

No interpolation is used after the extraction of the SM time series. The comparison is based on the following statistical indicators:

- mean bias: (in situ – retrieved soil moisture) (m^3 m^-3)
- standard error of the estimate (SEE) (m^3 m^-3)
- Pearson correlation coefficient (R)
- RMSE (m^3 m^-3)

6 Results and discussions

6.1 SMOS and SMAP brightness temperatures

Figures 5a, b and 6a, b show the comparison between the SMOS L3 TB and SMAP L3 TB at a 40° incidence angle. Figure 5a shows the average of SMOS and SMAP TB_H and TB_V for the winter (January, February and March) and summer (July, August and September) seasons for 2016. The gaps (in dark blue) in the SMOS images are due to RFI with a differentiated impact for ascending and descending orbits. The difference in TBs between H/V acquisitions is smaller than between ascending and descending configurations. The main explanations for these differences are that, first, the L1 algorithm in SMOS and SMAP does not use the same configuration for the computation of the Faraday rotation. The Faraday rotation is impacted by the TEC (total electronic content) in the ionosphere. The SMAP algorithm uses the STOKES 3 parameters to account for the Faraday rotation. The SMOS algorithm uses auxiliary TEC files to compute the Faraday rotation. The ionosphere TEC is very different between ascending and descending orbits as the heating during the day increases the TEC. The second explanation is that the RFI probabilities are very different between ascending and descending orbits due to directional aspects and they are closer between H/V polarisations. The SMAP products show a higher coverage because SMAP has on-board RFI filtering and mitigation, which enables a better coverage but at the cost of a lower radiometric accuracy. The spatial patterns of TB are highly consistent for the two missions. Figure 6a and b show the distribution of difference of TB_H and TB_V from SMOS and SMAP for the winter (January, February and March) and summer (July, August and September) seasons during 2016. As described in Sect. 5.1, only nodes with a water fraction of less than 0.01 (i.e. 1%) are considered. The mean difference is about −3.67 to −4.16 K, with SMAP being colder independent of polarisation or season. The standard deviation of all comparisons is about 3.65 K. This value is due to differences in calibration of the sensors and to the impact of differences in the acquisition time.

6.2 Soil moisture retrievals on a global scale

Based on the aforementioned evaluation methodology, the L3SM MO retrievals are compared to those of L2SM SO on the global scale over the 2010–2015 period. The auxiliary maps of mean forest cover percentage (Fig. 7a) and average RFI probabilities (Fig. 7b) for 2011 are provided as complementary information. These maps are obtained from the L3SM product.

Figure 8a and b show the mean number of successful retrievals per year (2010–2015) obtained from L3SM and the empirical cumulative distribution function (Cox and Oakes, 1984).
Figure 5. The 3-month average maps of SMOS L3 TB at 40° (left panels) and SMAP L3 TB (right panels) for H polarisation and V polarisation considering January, February, and March (a) and July, August, and September (b).

L2SM, respectively. White (blank) pixels in panel (a) show the areas where no successful soil moisture retrieval is available. These pixels are mostly located in areas of dense vegetation (Congo), areas that are seasonally inundated (Amazon Basin) and/or areas with high RFI (South-East Asia and the Middle East). From Fig. 10a it is clear that the coverage area of the L3SM product is higher in these areas.

Figure 9a and b show the difference (MO–SO) in the number of successful soil moisture retrievals between the L3SM and L2SM products. The general behaviour shows a systematic increase in the number of retrievals of the MO with respect to the SO retrievals. The number of retrievals moderately increases in desert and plain areas (10–20 retrievals per year per orbit). The increase is much higher for forested areas. The L2SM showed a higher number of successful retrievals in the area between 62–70° longitude and 35–55° latitude. This is due to an anomaly in the processing of TB products. The ancillary data containing the TEC are not properly used over this region. This has been corrected and all operational products are now properly processed. The archive products will be corrected for this error in the next processing campaign. Also, from Figs. 7 and 8 it is clear that no enhancement in number of retrievals has been observed in areas with very high RFI probabilities in descending orbits (not shown here) like the north Asian region.

The mean soil moisture from L3SM and L2SM for ascending orbits is provided in Fig. 10a and b. These figures show that the soil moisture spatial patterns are very similar between the SO and MO SM retrievals. The coverage of the multi-orbit product is higher, as already shown in the previous figures. Nevertheless, some discrepancies can be observed from the difference map (Fig. 10c). The L3SM MO soil moisture values are generally higher than those of L2SM SO. This is most visible in forested areas (Fig. 7a), and this is
Figure 6. Distribution of bias between SMAP and SMOS L3 TB for pixels with less than 0.01 (1%) water fraction for January–February–March (a) and July–August–September (b), H polarisation (right panels) and V polarisation (left panels).

consistent with climatic conditions over these areas. They are also higher in areas with high RFI pollution (Fig. 7b). This generally leads to a decrease in the value of the retrieved soil moisture values. Thus, the higher L3SM can be due to the positive impact of using multiple dates during RFI prone periods.

6.3 In situ comparison

The statistics for the comparison of L2SM SO and L3SM MO with in situ networks is shown in Tables 3 and 4 for ascending and descending orbits, respectively. The number of retrievals is systematically higher for the L3SM than the L2SM as expected from the global analysis. Note that, contrary to the global analysis, the in situ analysis is done without any grid interpolation by considering the closest node. Tables 6 and 7 show the statistics for the on-site comparison for ascending and descending orbits, respectively. The skills are of similar magnitudes for the LW and Niger sites and the lowest skill is obtained for the Benin site in descending overpasses. No site showed a lower number of successful retrievals for L3SM. The bias values are not much improved by the L3SM. They seem to increase at the majority of the sites. The comparison shows a slight negative bias for the two datasets. The absolute value of bias is less than 0.04 m$^3$m$^{-3}$, except for the L3 retrieval over the Benin site, which is 0.058 m$^3$m$^{-3}$. Thus, in general the bias is within the uncertainty of the in situ data if we consider the combined errors from sensor errors, spatial heterogeneity and sensing depth mismatch. The correlation values range from 0.65 to 0.88 for the different sites. Increased correlation was found for the L3SM products over the Niger site and slightly over WG in descending overpasses. The majority of the correlation values remain high with L3SM retrieval with no significant difference between L2SM and L3SM.

More in-depth analysis can be obtained by inspecting the time series of soil moisture. Figures 11 and 12 show the time series for the selected sites for the period 2010–2016 and for ascending and descending overpasses. The Niger and Benin sites present a very pronounced seasonal signal typical of the Sahelian sites. Over these sites the L3SM shows consistently lower soil moisture than L2SM for high soil moisture values. The L3SM is closer in this case to the site data. The time
Figure 7. Global map of the mean forest-cover percentage used in the SMOS L2 SO and L3 MO soil moisture retrievals (a) and map of the radio frequency interference (RFI) probabilities (b) for ascending orbit from the L3MO soil moisture processor.

Table 6. Statistics of the in situ vs. SMOS L3SM and L2SM for ascending orbits.

<table>
<thead>
<tr>
<th>Site</th>
<th>R</th>
<th>Bias (m$^3$ m$^{-3}$)</th>
<th>SEE (m$^3$ m$^{-3}$)</th>
<th>RMSE (m$^3$ m$^{-3}$)</th>
<th>Number of retrievals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L2</td>
<td>L3</td>
<td>L2</td>
<td>L3</td>
<td>L2</td>
</tr>
<tr>
<td>AMMA-CATCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benin</td>
<td>0.84</td>
<td>0.74</td>
<td>−0.039</td>
<td>−0.058</td>
<td>0.056</td>
</tr>
<tr>
<td>Niger</td>
<td>0.82</td>
<td>0.81</td>
<td>−0.006</td>
<td>−0.003</td>
<td>0.052</td>
</tr>
<tr>
<td>Watersheds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little Washita</td>
<td>0.83</td>
<td>0.82</td>
<td>−0.021</td>
<td>−0.03</td>
<td>0.041</td>
</tr>
<tr>
<td>Walnut Gulch</td>
<td>0.81</td>
<td>0.73</td>
<td>0.005</td>
<td>−0.007</td>
<td>0.038</td>
</tr>
</tbody>
</table>

Table 7. Statistics of the in situ vs. SMOS L3SM and L2SM for descending orbits.

<table>
<thead>
<tr>
<th>Site</th>
<th>R</th>
<th>Bias (m$^3$ m$^{-3}$)</th>
<th>SEE (m$^3$ m$^{-3}$)</th>
<th>RMSE (m$^3$ m$^{-3}$)</th>
<th>Number of retrievals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L2</td>
<td>L3</td>
<td>L2</td>
<td>L3</td>
<td>L2</td>
</tr>
<tr>
<td>AMMA-CATCH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benin</td>
<td>0.74</td>
<td>0.61</td>
<td>−0.029</td>
<td>−0.037</td>
<td>0.069</td>
</tr>
<tr>
<td>Niger</td>
<td>0.63</td>
<td>0.65</td>
<td>−0.011</td>
<td>−0.008</td>
<td>0.049</td>
</tr>
<tr>
<td>Watersheds</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Little Washita</td>
<td>0.81</td>
<td>0.80</td>
<td>−0.001</td>
<td>−0.012</td>
<td>0.042</td>
</tr>
<tr>
<td>Walnut Gulch</td>
<td>0.69</td>
<td>0.72</td>
<td>−0.019</td>
<td>−0.029</td>
<td>0.047</td>
</tr>
</tbody>
</table>
Figure 8. Mean number of successful SM retrievals per year (2010–2015) for ascending orbits from L3SM MO (a) and L2SM SO (b).

Figure 9. Global map of the difference in the mean number of SM successful retrievals per year over the 2011–2015 period (L3SM_{MO}–L2SM_{SO}) for ascending orbits (a) and descending orbits (b).
series for LW show that the SMOS data closely follow the behaviour of the soil moisture dynamics over this site. One of the reasons for this is that the rainfall events are well separated, enabling the remote sensing data to capture the dynamics of physical processes (e.g. infiltration and evaporation) on a coarse scale. Thus, the exponential behaviour typical of a drying soil is well depicted.

Figures 13 and 14 show the CDF of the in situ, L2SM and L3SM data for ascending and descending orbits. From these figures it can be concluded that the SMOS soil moisture is drier than the 5 cm in situ data across the different values of soil moisture. This can be explained by the SMOS penetration depth with respect to the depth of the installation of the in situ sensors. Nevertheless, the shape of the distribution function, describing the extreme and seasonal cycles, is well captured in most cases. The Niger site’s Sahelian climate is well captured, with a high probability of low soil moisture values and a small number of extreme values. The differences between the L2SM and L3SM data are mainly observed for the Benin and LW sites. When comparing Figs. 13 and 14, small differences can be noted between ascending and descending orbits.

Figure 10. Mean soil moisture map over 2011–2015 for ascending orbits from CATDS L3SM MO (a), DPGS L2SM SO (b) and the difference (MO–SO) map between L3SM MO and L2SM SO (c).
Figure 11. Time series for the validation sites for the ascending (06:00 UTC) overpasses.
Figure 12. Time series for the validation sites for the descending (18:00 UTC) overpasses.
Figure 13. Cumulative distribution function (CDF) for the validation sites for ascending overpasses.

Figure 14. Cumulative distribution function (CDF) for the validation sites for descending overpasses.
7 Data availability

The main datasets can be accessed as follows:

- MIR_CLF31A/D: SMOS-CATDS Level 3 1-day soil moisture maps for ascending (06:00 UTC) and descending (18:00 UTC) orbits version 300, link: ftp://ext-catds-cpdc@ftp.ifremer.fr/Land_products/GRIDDED/L3SM/RE04/MIR_CLF31A/;

- MIR_CDF3TA/D: SMOS-CATDS Level 3 1-day fixed-angle bin full-polarisation brightness temperatures maps for ascending (06:00 UTC) and descending (18H00) orbits version 310, link: ftp://ext-catds-cpdc@ftp.ifremer.fr/Land_products/GRIDDED/L3SM/RE04/MIR_CLF31A/.

8 Conclusions

The level 3 daily maps of soil moisture and brightness temperatures are presented in this paper. A multi-orbit soil moisture retrieval algorithm for SMOS data is used to obtain the soil moisture product. The main feature of the algorithm is the use of MO and of temporal autocorrelation of optical vegetation depth in the cost function. The algorithm is implemented operationally at CATDS. The processing chain delivers gridded products over the EASE 2.0 grid at 25 km in netCDF format. The L3 angle-binned TB product is compared to SMAP brightness temperature maps at 40°. The results show small differences in mean TB between the products for H/V polarisation and ascending and descending orbits. The SMAP product presents a wider coverage due to the on-board RFI filtering. The L3SM MO product is compared to the L2SM SO product. The best improvements in algorithm performances are in terms of the number of successful retrievals observed over forested and RFI-prone areas. Also, the L3SM MO product shows, on average, wetter soil moisture retrievals than the L2SM SO. The comparison with local sites showed that the quality of the retrievals is comparable between L2SM SO and L3SM MO. This shows that the increase in the number of successful retrievals does not degrade quality, but rather comes at the expense of an increased time lag in product availability (6 h for L2SM SO versus 3.5 to 7 days for L3SM MO). The SO and MO products show a slight dry bias except for the AMMA Benin site, which is smaller than the in situ data uncertainty (\(< 0.04 \text{m}^3\text{m}^{-3}\)).

More accurate auxiliary files like soil maps from SoilGrids (https://www.soilgrids.org/) may improve the retrieval quality, but more densely instrumented sites will be needed to access the improvements. Future works will concentrate on the associated optical depth product not presented in this paper. An application of the algorithm to the SMAP data has been envisioned.
Appendix A: List of abbreviations

ARS  Agricultural Research Service
AMMA  Analyse Multidisciplinaire de la Mousson
AMSR-E  Advanced Microwave Scanning Radiometer – Earth Observing System
ASCAT  Advanced Scatterometer
CATDS  Centre Aval de Traitement des Données SMOS
CNES  Centre National d’Etudes Spatiales
CCI  Climate Change Initiative
CDTI  Centro para el Desarrollo Tecnológico Industrial
DPGS  Data Processing Ground Segment
EASE-Grid  Equal-Area Scalable Earth Grid
ECMWF  European Centre for Medium-Range Weather Forecasts
ECV  Essential Climate Variables
EO  Earth observation
ESA  European Space Agency
IFREMER  Institut Français de Recherche pour l’Exploitation de la Mer
ISEA  Icosahedral Snyder Equal Area
L-MEB  L-band Microwave Emission of the Biosphere
MO  Multi Orbit
MODIS  Moderate-Resolution Imaging Spectroradiometer
NASA  National Aeronautics and Space Administration (USA)
SM  Soil Moisture
SMAP  Soil Moisture Active Passive
SMOS  Soil Moisture and Ocean Salinity
SMUDP  Soil Moisture User Data Product
SO  Single Orbit
TOA  Top of Atmosphere
USDA  United States Department of Agriculture
VOD  Vegetation Optical Depth
ERS  European Remote Sensing
ATBD  Algorithm Theoretical Basis Document
NSIDC  National Snow and Ice Data Center (USA)
GSFC  Goddard Space Flight Center
UTC  Coordinated Universal Time
Acknowledgements. The SMOS L3SM products were obtained from the Centre Aval de Traitement des Données SMOS (CATDS), operated for the “Centre National d’Études Spatiales” (CNES, France) by IFREMER (Brest, France). This study was supported by the CNES “Terre, Océan, Surfaces Continentales, Atmosphère” program. The authors would like to thank the USDA ARS Hydrology and Remote Sensing Laboratory, AMMA-CATCH project for the in situ datasets.

Edited by: D. Carlson
Reviewed by: M. Schwank and one anonymous referee

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

References


Mougin, E., Hiernaux, P., Kergoat, L., Grippo, M., de Rosnay, P., Timouk, F., Le Dantec, V., Demarez, V., Lavenu, F., Ar-


