SMOS near real time soil moisture product: processor overview and first validation results

Nemesio Rodríguez-Fernández[1,2], Joaquin Muñoz Sabater[1], Philippe Richaume[2], Patricia de Rosnay[1], Yann Kerr[2], Clement Albergel[1,3], Matthias Drusch[4], and Susanne Mecklenburg[5]

[1]European Centre for Medium-Range Weather Forecasts, Shinfield Road, Reading RG2 9AX, UK
[2]CESBIO (Université de Toulouse, CNES, CNRS, IRD), 18 av. Edouard Belin, bpi 2801, 31401 Toulouse, France
[3]CNRM UMR 3589, Météo-France/CNRS, Toulouse, France
[4]European Space Agency, ESTEC, Noordwijk, The Netherlands
[5]European Space Agency, ESRIN, Frascati, Italy

Correspondence to: nemesio.rodriguez@cesbio.cnes.fr

Abstract. Measurements of the surface soil moisture (SM) content are important for a wide range of applications. Among them, operational hydrology and numerical weather prediction, for instance, need soil moisture information in near-real-time (NRT), typically not later than 3 hours after sensing. The European Space Agency (ESA) Soil Moisture and Ocean Salinity (SMOS) satellite is the first mission specifically designed to measure soil moisture from space. The ESA level 2 SM retrieval algorithm is based on a detailed geophysical modelling and cannot provide SM in NRT. This paper presents the new ESA SMOS NRT SM product. It uses a neural network (NN) to provide SM in NRT. The NN inputs are SMOS brightness temperatures for horizontal and vertical polarizations and incidence angles from 30° to 45°. In addition, the NN uses surface soil temperature from the European Centre for Medium Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS). The NN was trained on SMOS Level 2 SM. The swath of the NRT SM retrieval is somewhat narrower (~ 915 km) than that of the L2 SM dataset (~ 1150 km), which implies a slightly lower revisit time. The new SMOS NRT SM product was compared to the SMOS Level 2 SM product. The NRT SM data shows a standard deviation of the difference with respect to the L2 data of < 0.05 $\mathrm{m^3 m^{-3}}$ in most of the Globe and a Pearson correlation coefficient higher than 0.7 in large regions of the Globe. The NRT SM dataset does not show a global bias with respect the L2 dataset but can show local biases of up to 0.05 $\mathrm{m^3 m^{-3}}$ in absolute value. The two SMOS SM products were evaluated against in situ measurements of SM from more than 120 sites of the SCAN (Soil Climate Analysis Network) and the USCRN (United States Climate Reference Network) networks in North America. The NRT dataset obtains similar but slightly better results than the L2 data. In summary, the neural network SMOS NRT SM product exhibits performances similar to those of the Level 2 SM product but it has the advantage of being available in less than 3.5 hours after sensing, complying with NRT requirements. The new product is processed at ECMWF and it is distributed by ESA and via the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) multicast service (EUMETCast).

1 Introduction

Surface soil moisture (SM) represents less than 0.001 % of the global freshwater budget by volume but it plays an important role in the water, carbon and energy cycles (Lahoz and De Lannoy, 2013). SM is the water reservoir for plants and agriculture and it affects the propagation of diseases such as malaria (Montosi et al., 2012; Peters et al., 2014). The amount of moisture in the soil is an important variable to understand the coupling of the continental surface and the atmosphere (Koster et al., 2004; Seneviratne et al., 2006; Tuttle and Salvucci, 2016). Surface SM softens the effect of precipitations, affects the partitioning of the water cycle (infiltration and run-off, and therefore the groundwater storage, see McColl et al. (2017)) and it can can also be used to improve rainfall estimations (Pellarin et al., 2008; Crow et al., 2009; Brocca et al., 2016). SM measurements have been used to perform data assimilation into land surface models (Xu et al., 2015; Blankenship et al., 2016; Lievens et al., 2016), SVAT (Soil Vegetation Atmosphere Transfer) models (Martens et al., 2016; Ridler et al., 2014; Muñoz Sabater et al., 2007) and in carbon-cycle models (Scholze et al., 2016). SM data assimilation can improve river discharge predictions, and remote sensing measurements are useful in otherwise data-scarce catchments (Alvarez-Garreton et al., 2016; Chen et al., 2011; Pauwels et al., 2002). SM measurements are useful to monitor landslide risks (Hawke and McConchie, 2011) and remotely sensed SM has been used to compute landslide susceptibility maps (Ray et al., 2010).

Regarding flood forecasting, in the framework of the European Flood Awareness System (EFAS) the forecast accuracy improves significantly (5-10%) when remotely sensed SM is assimilated in addition to discharge data (Wanders et al., 2014). SM initial conditions are among the most important hydrological properties affecting flash flood triggering (Norbiato et al., 2008; Ponziani et al., 2012). The assimilation of soil moisture products from the Advanced Scatterometer (ASCAT) has been successfully used in the context of flash flood early warning systems in Mediterranean catchments (Cenci et al., 2016).

In addition to operational hydrology applications, operational numerical weather prediction also benefits from remotely sensed SM data assimilation. Meteorological agencies such as the European Centre for Medium Range Weather Forecasts (ECMWF) and the United Kingdom Met Office assimilate ASCAT surface SM into their operational numerical weather predictions models (de Rosnay et al., 2013; Dharssi et al., 2011). The approach has also been tested in off-line mode at Meteo France (Barbu et al., 2014). To be useful for operational applications, remotely sensed data should be available in near-real-time (typically less than 3-4 hours after sensing, hereafter NRT).

The Soil Moisture and Ocean Salinity (SMOS) European Space Agency (ESA) satellite (Kerr et al., 2010) is the first instrument that has been specifically designed to measure soil moisture from space. It carries an L-Band (1.4 GHz) radiometer to perform full polarization and multi-angular (0° - 60°) measurements of the Earth thermal emission. ECMWF uses SMOS NRT brightness temperatures (T_{b}) in their operational Integrated Forecasting System (Muñoz Sabater et al., 2012). The ESA SMOS operational Level 2 SM retrieval algorithm is based on a point-per-point iterative minimization of the difference of a physical model and the satellite measurements (Kerr et al., 2012). The free parameters are the soil moisture content and the 1.4 GHz opacity, which is mainly due to the water content of the vegetation in between the soil surface and the sensor (which some authors refer to as VOD, vegetation optical depth).

Many studies have evaluated the SMOS L2 SM dataset in comparison to other remote sensing datasets, models and in situ measurements (Al Bitar et al., 2012; Wanders et al., 2012; Albergel et al., 2012; Bircher et al., 2013; Al-Yaari et al., 2014a, b; Leroux et al., 2014; Louvet et al., 2015; Kerr et al., 2016). SMOS shows very good global performance although other remote sensing and model products can show better performances at some sites. In any case, datasets from the only two instruments specifically conceived to measure SM, SMOS and NASA's Soil Moisture Active Passive (SMAP), compare very well with each other (Jackson et al., 2016; Burgin et al., 2016).

As already mentioned, most operational users over land, in particular in numerical weather prediction and operational hydrology, require SM information to be available in NRT, typically referring to less than 3-4 h after sensing. For instance, ASCAT soil moisture data is distributed by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) within 135 minutes after data acquisition (Wagner et al., 2013), which allows to assimilate the data by operational numerical weather prediction centers such as ECMWF (de Rosnay et al., 2013). In the case of the current SMOS ground segment, the production of Level 1C T_b data from raw data takes typically 1 hour of processing time and the Level 2 SM inversion up to 80 minutes for a half-orbit. However, due to data handling operations, the synchronization of some operations and the dissemination orchestration the total latency time from data acquisition to SM dissemination is of the order of 6 hours. Therefore, this processing chain does not fulfil the NRT requirements. However, as already mentioned, SMOS T_b measurements are provided in NRT to ECMWF. In addition, with 6 years of SMOS measurements available, statistical algorithms can be exploited to provide faster retrievals and neural networks have been shown to be a promising technique to generate a SM dataset from SMOS T_b measurements (Rodríguez-Fernández et al., 2015). Based on the latter, a neural network processing chain to provide SMOS SM in NRT has been implemented by ESA. The requirements are that the NRT dataset should display at least the same accuracy as the geophysical level 2 soil moisture data product, the data should be retrieved over a large swath, and the retrieval should rely on a minimum of auxiliary data files. As the ASCAT NRT SM processing chain cited above, the SMOS NRT SM chain handles model parameters derived off-line using a database with a large number of past observations. The advantages are that the processing is robust and very fast. In case that significant changes in the Level 1 data used as input are available, then the model parameters should be updated correspondingly.

The new SMOS NRT SM product is available from 2016 onwards and it is distributed through the World Meteorological Organization's Global Telecommunication System (GTS) and the EUMETCast service from EUMETSAT in NetCDF format. EUMETCast is a dissemination system that uses commercial telecommunication geostationary satellites and research networks to multi-cast data files to a wide user community.

This paper describes the SMOS NRT SM processing chain and discuss the first evaluation results. It is organized as follows. Section 2 describes the data used for the implementation and the validation of the SMOS NRT SM product. Section 3 discusses the NRT SM processing chain (more details are given in the Appendix). Section 4 contains a description of the methods used to evaluate the NRT SM product. Section 5 presents the evaluation results. Finally, a summary is presented in Section 6.

2 Data

2.1 SMOS satellite

SMOS (Mecklenburg et al., 2012; Kerr et al., 2010) measures the thermal emission from the Earth at a frequency of 1.4 GHz in full-polarization and for incidence angles from $0°$ to $\sim 60°$. The full incidence angle range is accessible in the center of the swath. On the contrary, only angles in the $40°$-$45°$ range are accessible all across the swath. SMOS has 69 antennas to perform interferometry and synthesize an aperture of ~ 7.5 m (Anterrieu and Khazaal, 2008). The spatial resolution on the ground, defined as the projection of the full width at half maximum of the synthesized beam, is 43 km on average over the field of view (Kerr et al., 2010). The satellite follows a sun-synchronous polar orbit with 6:00 am/pm equator overpass time for ascending/descending half-orbits.

2.1.1 SMOS Level 2 soil moisture

The SMOS level 2 processor performs a detailed modelling of the Earth emission at 1.4 GHz at two polarizations and a large number of incidence angles using the $\tau - \omega$ (optical depth - single scattering albedo) approach to account for interaction of L-band radiation with the vegetation (Wigneron et al., 2007). The ground effective temperature is computed from the soil temperature at a deep layer (~ 1 m) and the surface layer (a few cm's) using the formulation of Choudhury et al. (1982) with the parametrization by Wigneron et al. (2001). The soil temperature for those two layers is taken from ECMWF IFS model simulations.

For each grid node, the surface is modelled with $4 \times 4\,\text{km}^2$ cells taking into account different land covers. Then the processor computes the contributions of those cells within $123 \times 123\,\text{km}^2$ accounting for the projection of the SMOS synthesized antenna power pattern on the Earth surface to model SMOS-like T_b values. The vegetation optical depth (τ) and the SM content are free parameters that are allowed to vary to minimize the difference of the simulated T_b values and SMOS Level 1C T_b measurements. In the case of forest, two contributions to the opacity are taken into account: one from the trees, which is estimated from the maximum Leaf Area Index (LAI, Ferrazzoli et al., 2002), and another from the understory vegetation. Soil temperature is obtained from ECMWF Integrated Forecast System (IFS) data. For footprints with mixed land cover, the SM content of the minor land cover is estimated from ECMWF IFS and its contribution to the T_b is fixed. For such cases, the SMOS SM retrieval is only performed for the dominant land cover class within the footprint (Kerr et al., 2012). ESA Level 2 SM data are provided in an Icosahedral Equal Area (ISEA) 4H9 grid (Sahr et al., 2003) with a sampling space of 15 km.

The version of the SMOS L2 SM dataset used in this study is v620, which became operational in May 2015. In order to have enough data for a robust training of the NN, an additional dataset from 01/June/2010 to 30/June/2012 was reprocessed with the same version v620 of the L2 SM algorithm. The evaluation of the NRT SM product has been done from May 2015 to the time of the NRT SM implementation (end of 2015).

2.1.2 SMOS near Real-Time soil moisture

The SMOS Near Real-Time SM product was obtained by training a neural network using SMOS T_b measurements and soil temperature from ECMWF models as input. SMOS T_b measurements are provided by ESA to ECMWF in NRT (less than 3 hours after sensing) for operational monitoring within the Integrated Forecast System (Muñoz-Sabater, 2015).

The training dataset used for the supervised learning phase of the neural network was the SMOS Level 2 SM product. The training was done using data from June 2010 to June 2012 and it is described in Sect. 3. The NRT SM processing chain was evaluated using data from May 2015 to November 2015. Taking into account the satisfactory results of the evaluation (presented in Sect. 5), the SMOS NRT SM product became operational in January 2016. The SMOS NRT SM product is computed at ECMWF and delivered to ESA, where the data are sent to EUMETSAT for dissemination via EUMETCast. SMOS NRT SM data can also be accessed via the SMOS online dissemination service from the ESA Earth Online portal. The SMOS NRT SM data are provided in NetCDF files in the same ISEA 4H9 grid of other ESA SMOS products. The version of the SMOS NRT SM data used in this study is version 100. More details on the NRT SM processor are presented in Sect. 3 and in the Appendix.

2.2 In situ soil moisture measurements

The SMOS NRT SM product was evaluated against in situ measurements of SM over a large number of sites. The same evaluation was done with the Level 2 SM product. The in situ data used for those evaluations are described below.

The Soil Climate Analysis Network (SCAN) of the United States Department of Agriculture (Schaefer et al., 2007) has been widely used to evaluate modelled and remote sensing soil moisture datasets and it contains over 100 sensors/sites. The sensors are located in agricultural regions with a relatively homogeneous landscape in many cases. The sensors used in this study are placed horizontally at 5 cm depth.

The United States Climate Reference Network (USCRN, Bell et al., 2013) is a network of climate monitoring stations with sites across the U.S.A., managed and maintained by the National Oceanic and Atmospheric Administration (NOAA). This network was designed with climate science in mind. The stations are placed in pristine environments expected to be free of development for many decades. There are around 140 stations with sensors at different depths. The sensors used in this study are horizontally installed at 5 cm.

The in situ data have been obtained directly from the teams operating both networks but these datasets are also available from the International Soil Moisture Network (Dorigo et al., 2011).

3 The SMOS Near-Real-Time soil moisture processor

The SMOS NRT SMOS processor is based on a neural network approach to retrieve soil moisture from SMOS observations. Rodríguez-Fernández et al. (2015) showed that SMOS T_b values binned in 5°-width incidence angle bins (the L3TB product, Al Bitar et al., 2017) can be used to retrieve SM on a daily basis. They used ECMWF SM model fields as reference dataset to train the NN. Rodríguez-Fernández et al. (2013) had previously shown that SMOS Level 3 SM (Al Bitar et al., 2017) can also
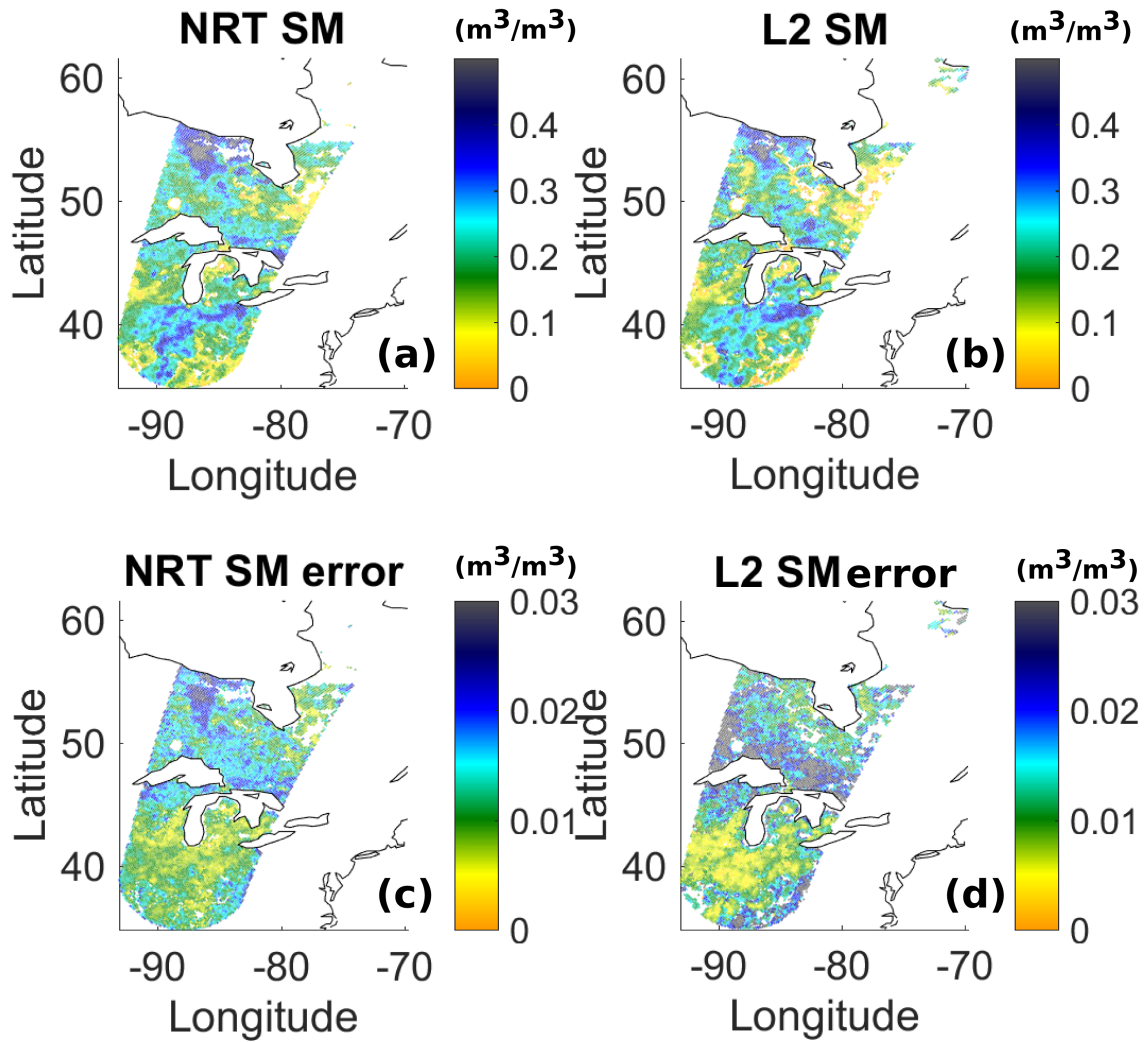
Figure 1. Comparison of the NRT-NN SM product (a) and the Level 2 SM (b) for one orbit of day 27/May/2012. The corresponding NRT-NN uncertainty is shown in panel (c), while the L2 SM uncertainty is shown in panel (d).
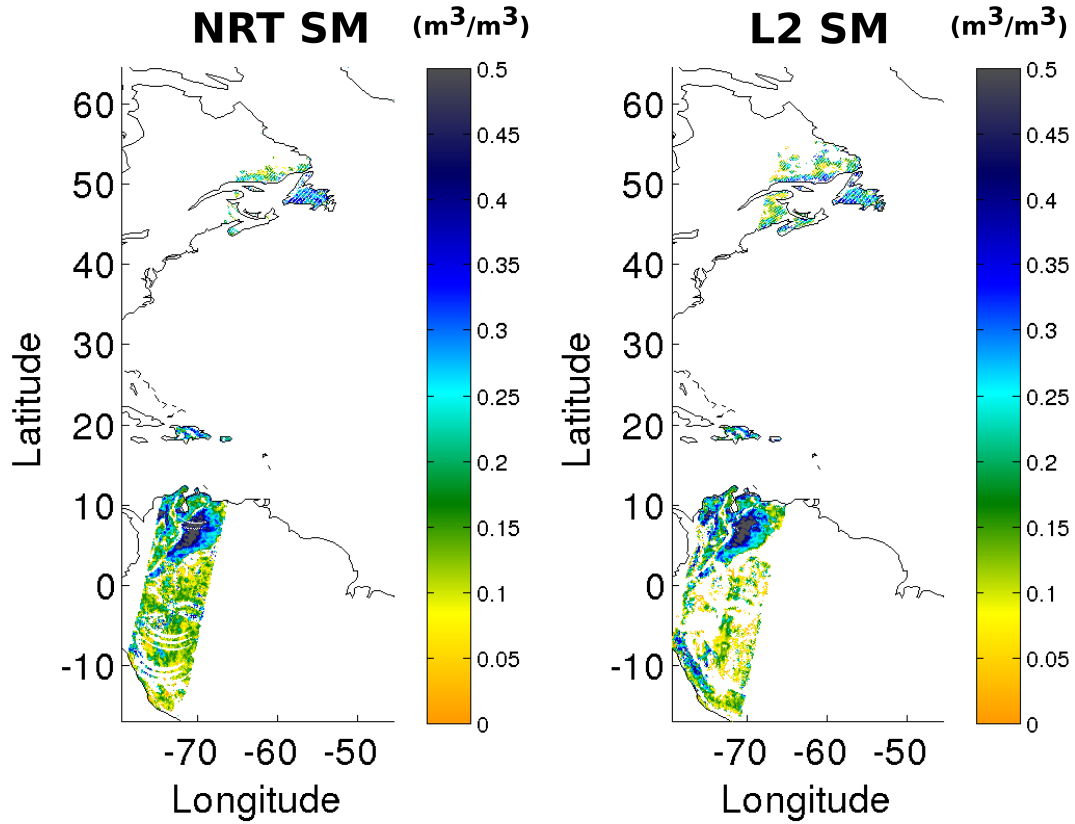
Figure 2. Comparison of the NRT-NN SM product (a) and the Level 2 SM (b) for one orbit of day 27/May/2012.

be used as reference data to train the NN instead of ECMWF modelled SM fields. They also showed that the additional input dataset with the most significant impact on the retrieval is soil temperature. In the context of the ESA operational NRT SM processor the goal was to obtain a SM dataset as similar as possible to the the ESA Level 2 SM dataset but in NRT. Therefore, the ESA SMOS Level 2 SM dataset (Kerr et al., 2012) was used as reference dataset for the training phase. Finally, taking into account operational constraints, the only complementary data used for the retrieval are soil temperature estimations for the first layer (0-7 cm) of ECMWF models, which is the complementary dataset with the most significant impact ($\sim 3\%$) on the retrieval (Rodríguez-Fernández et al., 2013). This dataset was chosen because it is the same model data used by the Level 2 SM algorithm (Sect. 2).

3.1 Input data

The input to the SMOS NRT SM processor are SMOS near-real-time T_{b} values, which are distributed by ESA to ECMWF in BUFR (Binary Universal Form for the Representation of meteorological data) format (Gutierrez and Canales Molina, 2010; de Rosnay et al., 2012). The T_{b} values are provided with the polarization referred to the antenna reference frame XY. Several
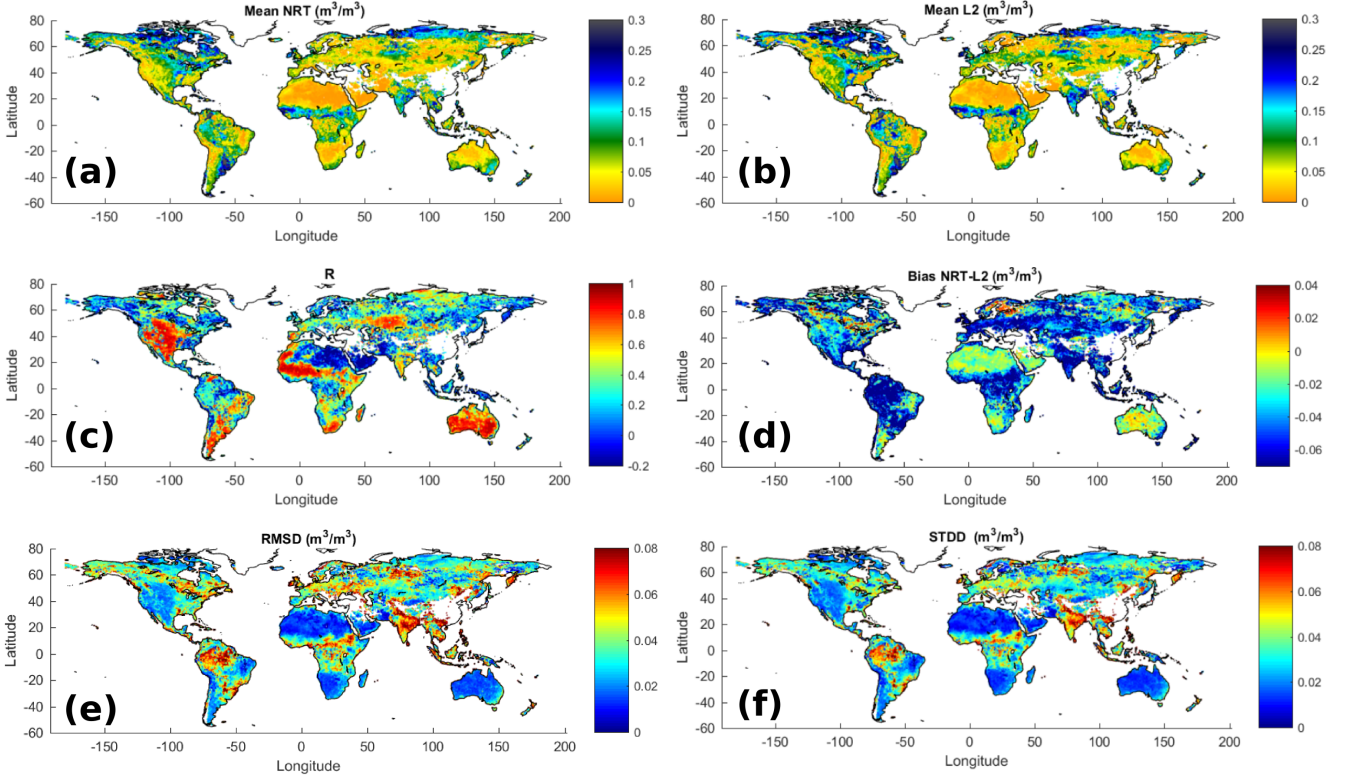
Figure 3. Mean SM for the NRT (a) and the L2 (b) SMOS products. Pearson correlation (c), bias (d), root mean square of the difference (e) and standard deviation (f) of the difference of the NRT SM and L2 SM.

quality checks are performed to filter the T_{b} values: $T_{\mathrm{b}X}$ and $T_{\mathrm{b}Y}$ should be in the expected physical range [80 K − 340 K] and the real and imaginary components of the cross-polarised measurements ($T_{\mathrm{b}XY}$) should be in the range [−50 K, 50 K], otherwise the T_{b} values are considered to be corrupted or affected by RFI (Radio Frequency Interference from human-built equipment). In order to keep information on the possible residual RFI contamination, a RFI probability was computed for each

5 observation as the number of BUFR T_{b} measurements filtered out due to the RFI flags with respect to the total number of T_{b} measurements. The observed T_{b} values are also filtered out if a specific BUFR flag indicates that the observation is located in a zone affected by the aliased image of the Sun. The selected NRT T_{b} values are transformed from the antenna-based XY reference frame to the ground-based horizontal and vertical (HV) reference frame as described by Al Bitar et al. (2017).

In a second step the HV T_{b} values are averaged in 5°-width incidence angle bins. Three angle bins are actually used for

10 training and applying the neural network: 30°– 35°, 35°– 40°, and 40°– 45°. As discussed by Rodríguez-Fernández et al. (2016), using these three angle bins is the best trade off of performances (which improve with a large angular signature) and swath-width of the retrieval (which decreases with an increasing number of angle bins used). With this configuration SM is retrieved in the central 914 km of the swath (the maximum possible swath is ∼ 1150 km). A SM retrieval can only be

done if there is a well-defined value of the T_{b} for all the three angle bins and the two polarizations H and V. The current implementation of the NRT SM processor does not perform any interpolation of the T_{b} versus incidence angle profiles.

Using the SMOS T_{b} values measured at a time t for a given latitude (λ) and longitude (ϕ) grid point and for each polarization and incidence angle bin, $T_{b_{\lambda\phi}}(t)$, a local normalized index can be computed as:

$$5 \quad I_{\lambda\phi}(t) = SM_{\lambda\phi}^{T_b^{min}} + [SM_{\lambda\phi}^{T_b^{max}} - SM_{\lambda\phi}^{T_b^{min}}]\frac{T_{b_{\lambda\phi}}(t) - T_{b_{\lambda\phi}}^{min}}{T_{b_{\lambda\phi}}^{max} - T_{b_{\lambda\phi}}^{min}} \tag{1}$$

Where T_b^{max} and T_b^{min} are the maximum and minimum values of the T_{b} in the local time (λ, ϕ) series, $SM^{T_b^{min}}$ and $SM^{T_b^{max}}$ are the associated SM in the SMOS Level 2 SM reference dataset. The index I is computed for each incidence angle bin and polarization at the time t of the SMOS acquisition and it contains a local information on the dynamic ranges of both the measured T_{b} and the reference SM. In the current version of the processor (v100), $T_b^{max,min}$ and $SM^{T_b^{max,min}}$ have been

10 computed using data from 01/June/2010 to 30/June/2012 (the same period used to train the neural network, see Sect. 3.2). This linear expectation index is inspired by the approach used to retrieve SM with the scatterometers such as ASCAT (Wagner et al., 1999; Bartalis et al., 2007) and helps to improve the performances of the NN retrieval (Rodríguez-Fernández et al., 2015).

The only auxiliary data used by the SMOS SM NRT processor are snow depth and soil temperature from the latest high-resolution forecast produced by the ECMWF IFS, with a typical latency of less than 1 h. The ECMWF IFS soil temperature in

15 the 0–7 cm layer is used as input to the NN, as it increases the performances of the retrieval (Rodríguez-Fernández et al., 2016). SMOS data from a given grid point are not used if snow is found in that point based on the latest ECMWF snow depth forecast field or if the soil temperature forecast of the top soil 7 cm is below 274 K. The SMOS NRT SM dataset is a land-only product and a SM retrieval is not provided if more than 50% of the SMOS footprint is covered by water. This filter avoids spurious soil moisture values near the coastlines, for instance. Of course, even if less than 50 % of the footprint surface is covered by water,

20 the SM retrieval can still be affected by the free water. Users interested in regions close to the coast or to water bodies are advised to use the land-sea mask available on the ESA SMOS Data Products portal. This mask was computed from the 1-km USGS (United States Geological Survey) land-sea mask aggregated into the ISEA grid common to Level 1 and Level 2 SMOS products.

3.2   The neural network processor

25 The HV angle-binned T_{b} values have been collocated with ECMWF IFS forecasts for the soil temperature and snow cover and finally they have been collocated with version 620 SMOS L2 SM data (Kerr et al., 2012) in the 01/June/2010 to 30/June/2012 period. As discussed above, a local normalized index I has been computed from extreme T_{b} values and the associated L2 SM. In addition to the filters discussed above (hard RFI, Sun tails, frozen or snow covered soil, water fraction), to compute the extreme values tables and for the training of the NN, the following filters have also been applied:

30   – The latitude is limited to the [-60°,75°] range.

  – A SMOS L2 SM value associated to the maximum or minimum T_{b} is required (otherwise I cannot be defined).

– The SM uncertainty provided by `Dqx` (Data Quality Index) parameter in L2 SM data files was required to be lower than 0.06 m^{3}m^{-3} to use the most reliable data for the training.

The input vectors contain T_{b} values and I indexes for H and V polarizations and the three angle bins from $30°$ to $45°$ and the soil temperature from 0 to 7 cm from ECMWF IFS forecast. Therefore, input vectors have a total of 13 elements. All 13 elements must be well-defined to train the NN and there must be a well-defined associated SM value.

One fifth of the vectors in the training data base were selected randomly to have $\sim 3 \times 10^5$ vectors. A subset of 60% of those vectors is used for the actual training, 20% is used for evaluation of the NN performances during the training and to avoid over-training, the final 20% is used to test the performances of the trained NN *a posteriori*. Gradient back-propagation and minimization with the Levenberg-Marquardt algorithm has been used. One single hidden layer with 5 neurons has been used, as it has been shown by Rodríguez-Fernández et al. (2016) that it is enough to capture the relationship between the input data and the reference SM and while keeping the NN as simple as possible. No signs of overtraining were found and the training was stopped after 50 iterations when the mean squared difference was asymptotically approaching to a minimum. When the trained NN was applied to the test subset and the NN output was compared to the SMOS L2 SM, the Pearson correlation R was 0.86, the standard deviation of the difference (STDD) was 0.068 m^{3}m^{-3} and the Root Mean Square Error or Difference (RMSE) was also 0.068 m^{3}m^{-3}, which implies that there was not a significant bias in between both SM datasets. These results show that the NN ability to capture the dynamics of the current L2 SM dataset is very good. The evaluation results discussed in Sect. 5 below confirm that the quality of the SM-NRT-NN product fulfil the specifications of the operational product.

NN NRT SM uncertainties were computed by error propagation through the neural network taking into account the error of the T_{b} measurements used as input as explained in the Appendix.

3.3 SMOS NRT SM processor output

The SMOS NRT SM product is a land-only product, collocated and delivered in the ISEA 4H9 grid (Sahr et al., 2003) common to other ESA SMOS products. The main characteristics of the product and the description of the fields are presented in Muñoz-Sabater et al. (2016). The processor output fields are shown in Table 1. Figure 1 shows the NRT-NN SM product and its associated uncertainty for a portion of an orbit of day 27/May/2012.

4 Methods

4.1 Global evaluation

Several metrics have been used to evaluate the NRT SM dataset from 15/May/2015 to 25/November/2015 against the SMOS L2 SM dataset. For all grid points $\lambda\phi$, the temporal means of both SM dataset, $\overline{SM}_{\lambda\phi}^{L2}$ and $\overline{SM}_{\lambda\phi}^{NRT}$, have been computed as:

$$\overline{SM}_{\lambda\phi}^{NRT} = \frac{1}{N_t} \sum_{i=1}^{N_t} SM_{\lambda\phi}^{NRT}(t_i) \tag{2}$$

Table 1. SMOS Near Real-Time soil moisture processor output fields

Fields in the NRT product
ISEA grid point number
Latitude
Longitude
Number of days since 1-Jan-2000
Seconds from midnight (all times are UT)
Soil moisture
Soil moisture uncertainty
RFI probability

Table 2. Comparison to in situ measurements over the USCRN and SCAN networks. The columns are: the SM product, the mean number of points in the time series, the mean and median Pearson correlation with respect to in situ measurements, the mean bias (mean in situ SM minus mean SMOS SM), the RMS and STD of the difference time series averaged over all sites, and the Pearson correlation of the anomalies time series (R_a). The statistics have been computed independently for the SM-NRT-NN and the SM-L2 product. The number of SM retrievals is, on average, larger for the SM-L2. The two lower rows show the results using only times for which both the SM-NRT-NN and the SM-L2 products are simultaneously available.

SM	Mean N_{pts}	Mean R	Median R	mean Bias	mean RMSD	Mean STDD	mean R_a
L2	186	0.63	0.64	0.035	0.100	0.065	0.56
NRT	94	0.70	0.71	0.036	0.095	0.058	0.48
L2	88.	0.67	0.69	0.026	0.092	0.062	0.59
NRT	88.	0.71	0.72	0.031	0.091	0.056	0.56

and

$$\overline{SM}_{\lambda\phi}^{L2} = \frac{1}{N_t} \sum_{i=1}^{N_t} SM_{\lambda\phi}^{L2}(t_i), \tag{3}$$

using only times (t_i) for which a well-defined value is present simultaneously in both datasets. This number is in principle different for each $\lambda\phi$ grid point, but it will be noted as N_t in the following instead of $N_{\lambda\phi_t}$ to simplify the notation.

5 A bias map has been computed from the local (λ and ϕ) mean of each dataset as follows:

$$Bias_{\lambda\phi} = \overline{SM}_{\lambda\phi}^{NRT} - \overline{SM}_{\lambda\phi}^{L2}. \tag{4}$$

In order to compare the temporal dynamics of the two datasets, the Pearson correlation R has also been computed as follows:

$$R_{\lambda\phi} = \frac{\sum\limits_{i=1}^{N_t}(SM_{\lambda\phi}^{NRT}(t_i) - \overline{SM}_{\lambda\phi}^{NRT})(SM_{\lambda\phi}^{L2}(t_i) - \overline{SM}_{\lambda\phi}^{L2})}{\sqrt{\sum\limits_{i=1}^{N_t}(SM_{\lambda\phi}^{NRT}(t_i) - \overline{SM}_{\lambda\phi}^{NRT})^2}\sqrt{\sum\limits_{i=1}^{N_t}(SM_{\lambda\phi}^{L2}(t_i) - \overline{SM}_{\lambda\phi}^{L2})^2}}, \tag{5}$$

where the sum runs for all the points available at a given position $\lambda\phi$: N_t.

The absolute values of the two datasets have been evaluated using the standard deviation of the difference as a metric. The local time series difference D of the two datasets was defined as:

$$D_{\lambda\phi}(t) = SM_{\lambda\phi}^{NRT}(t) - SM_{\lambda\phi}^{L2}(t) \tag{6}$$

The standard deviation of the difference time series (STDD) has been computed as:

$$STDD_{\lambda\phi} = \sqrt{\overline{D_{\lambda\phi}^2} - \overline{D_{\lambda\phi}}^2} = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}D_{\lambda\phi}^2(t_i) - \left(\frac{1}{N_t}\sum_{i=1}^{N_t}D_{\lambda\phi}(t_i)\right)^2} \tag{7}$$

In some studies, the STDD is calculated indirectly from the bias and the root mean squared difference (RMSD) and called unbiased-RMSD (ubRMSD).

4.2 Local evaluation against in situ measurements

Evaluating remote sensing measurements against in situ measurements is a difficult exercise. The spatial resolution of coarse scale remote sensing observations (~ 40 km) is very different to point-like measurements by in situ sensors. The large scale spatial representativeness of the in situ measurements is not guaranteed (see for instance Gruber et al., 2012). In addition, the depth of the microwave emitting layer can be different with respect to the sensing depth of the in situ measurements. The goal of the current study is not to deal with this open issues but to compare two different retrieval approaches using the same instrument, therefore spatial representativeness or sensing depth differences will not affect the comparison. More detailed evaluations of SMOS SM retrievals can be found in the following references: Al Bitar et al. (2017); Kerr et al. (2016); Leroux (2012); Van der Schalie et al. (2016); Jackson et al. (2012).

The SMOS NRT SM and the SMOS L2 SM datasets have been evaluated against the in situ measurements discussed in Sect. 2 in a consistent manner. First, for each station available, a quality check of the data was performed. Sites with suspicious data (e.g. measurements discontinuity, spurious jumps) were eliminated. The locations of the 127 retained stations is showed in Fig. 4.

The same metrics discussed in the previous section have been computed for the NRT SM dataset with respect to the in situ measurements and for the L2 SM dataset with respect to the in situ measurements. The Pearson correlation was used to compare temporal dynamics of two SM datasets. The long term (seasonal) dynamics were compared by computing the Pearson
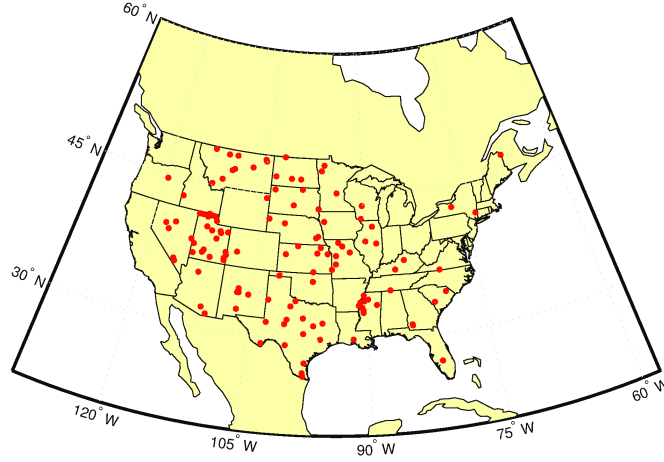
Figure 4. Location of the in situ measurements sites used in this study.

correlation coefficient R of SM^{L2} and SM^{NRT} with respect to SM^{inSitu}, site per site. In addition, the short-scale (1-30 days) dynamics were evaluated by computing site per site the Pearson correlation of the anomalies times series. Following Albergel et al. (2009), the SM anomaly at given time t, $SM_a(t)$, was computed using a 31 day window centred at t as follows:

$$SM_a(t) = \frac{SM(t) - \text{Mean}(SM(t-15, t+15))}{\text{STD}(SM(t-15, t+15))} \qquad (8)$$

5   where $SM(t-15, t+15)$ represents the ensemble of measurements in the 31-day window. The Pearson correlation coefficient R computed with the anomalies time series will be referred to as R_a in the following.

The metrics were computed independently for the NRT and the L2 datasets in a first step. In a second step, the metrics were recomputed only using times for which both the NRT and the L2 were simultaneously available, and thus, using the same number of points for the two time series.

10 **5   Results: SMOS NRT soil moisture evaluation**

5.1   Swath-level comparison to SMOS L2 SM

Figures 1a,c show the NRT-NN SM product and its associated uncertainty for a portion of an orbit of day 27/May/2012. The corresponding L2 SM and its associated uncertainty as given by the DQX (Data Quality Index) parameter (Kerr et al., 2012) are also shown (Figs. 1b,d). As discussed in Sect. 3, the swath width of the NRT-SM retrieval is somewhat narrower than the L2

15 SM one but both maps show similar spatial structures and numerical values. The uncertainties have similar numerical values as well, but the spatial patterns are not the same. This is expected as the two retrieval algorithms are different. Finally, it should be noted that the spatial coverage can be different for both products as shown in Fig. 2:
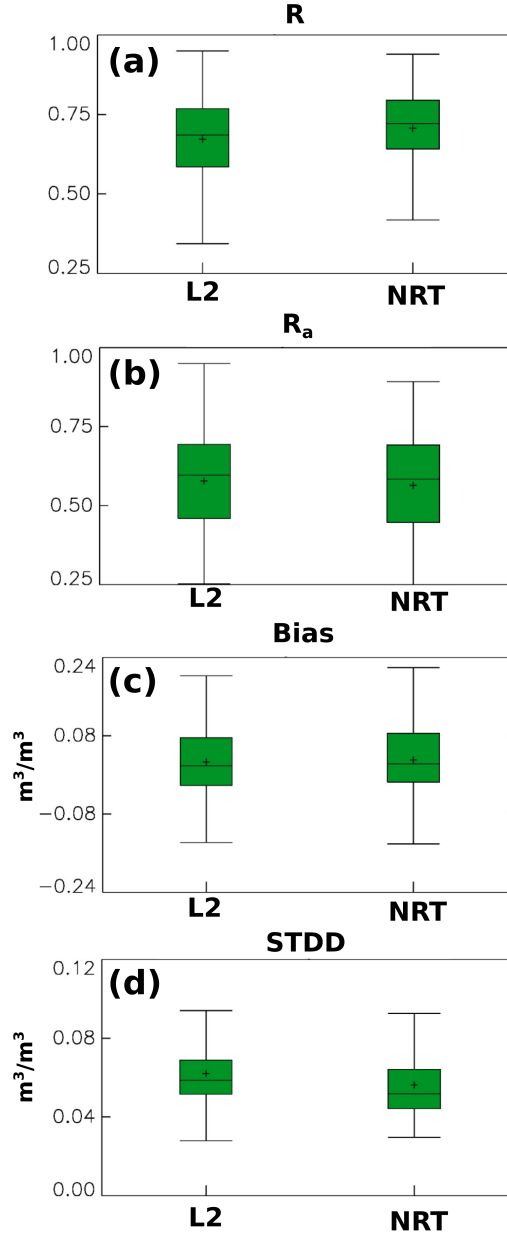
Figure 5. Boxplots for (a) the Pearson correlation coefficient (R) of the NRT and L2 time series with respect to the in situ measurements (b) Pearson correlation coefficient of the anomalies time series (R_a), (c) Bias (mean *in situ* minus mean SMOS SM), and (d) STDD of the two SMOS products in comparison to in situ measurements. The box contains the middle 50% of the data, the central bar represents the median value of the distribution. The upper edge (hinge) of the box indicates the 75th percentile of the data set (q_3), and the lower hinge indicates the 25th percentile (q_1). The mean values are also shown as black crosses. The upper and lower bars represent the minimum and maximum values of the distribution excluding outliers. Points are considered as outliers if they are larger than $q_3 + 1.5(q_3 - q_1)$ or smaller than $q_1 - 1.5(q_3 - q_1)$.

- the NRT SM product can show circle-arc gaps when not all of the angle bins have a well-defined T_b value, while in contrast the L2 algorithm can perform an inversion even if some T_b values have been filtered out.

- The NRT SM global retrieval algorithm can provide a SM estimate even when the local minimization of the L2 algorithm does not converge. This can happen mainly in dense forest areas.

5.2 Global evaluation with respect to SMOS L2 SM

The SMOS NRT SM product has also been compared to the SMOS L2 SM product globally and over the period mentioned in Sect. 4. Figs. 3a,b show the mean of the NRT and L2 SM products over the period of the study. Both maps show an overall excellent agreement, although it is possible to appreciate a significant negative bias (-0.05 m^{3}m^{-3}) in the NRT SM product in the regions with the highest L2 SM (tropical and boreal forest) . The typical number of points with both NRT-SM-NN and SM-L2 in the evaluation period is ~ 100. The correlation of both products is high (> 0.7) over a large part of North-America, the southernmost part of South-America, the Iberian peninsula, the Sahel and South-Africa, Australia and parts of central Eurasia. The correlation is significantly lower over forest (both tropical and boreal) and in deserts such as the Sahara. In the Sahara the low correlation is probably not significant because the SM values are very low and the variance is driven by the noise. Actually, Fig. 3f shows that the STDD is also very low in this region. Therefore, L2 SM and NRT SM have actually similar values. In contrast, dense forest regions show a high STDD in addition to a low R. Therefore, both products show some differences in these regions. Unfortunately, in situ measurements are not available to perform an independent evaluation of both dataset for dense forest sites. In conclusion, both products show similar dynamics over large parts of the Globe. The Bias map (Fig. 3d) shows that the NRT SM products show a tendency to underestimate the L2 SM dataset, which is a expected behaviour as it has been obtained using a regression technique and extreme values are under-represented in the reference dataset. The most significant effect of the bias is to increase the RMSD (Fig. 3e) with respect to the STDD in parts of Europe and Canada. However, one should note that both the RMSD and the STDD are lower than 0.04 m^{3}m^{-3} over most of the Globe (all except the reddish regions in Fig. 3e,f).

5.3 Evaluation with respect to in situ measurements

The SMOS NRT SM product was evaluated against in situ measurements from the SCAN Schaefer et al. (2007) and USCRN Bell et al. (2013) networks (Sect. 2). These networks of in situ measurements have been extensively used for the validation of remote sensing data (Albergel et al., 2009; Rodríguez-Fernández et al., 2015; Al Bitar et al., 2012; Albergel et al., 2012; Kerr et al., 2016).

The quality metrics discussed in Sect. 4 have been computed site per site independently for the SMOS NRT. The same evaluation was done for L2 products. The mean number of points in the time series from May 2015 to November 2015 is 186 for the L2 product while is only half of that value for the NRT product. The reason is a longer revisit time of the SM-NRT-NN product due to the narrower swaths of the retrievals and the lack of retrievals if not all the 6 T_b values are well defined for both polarizations and the three angle bins from 30° to 45°.
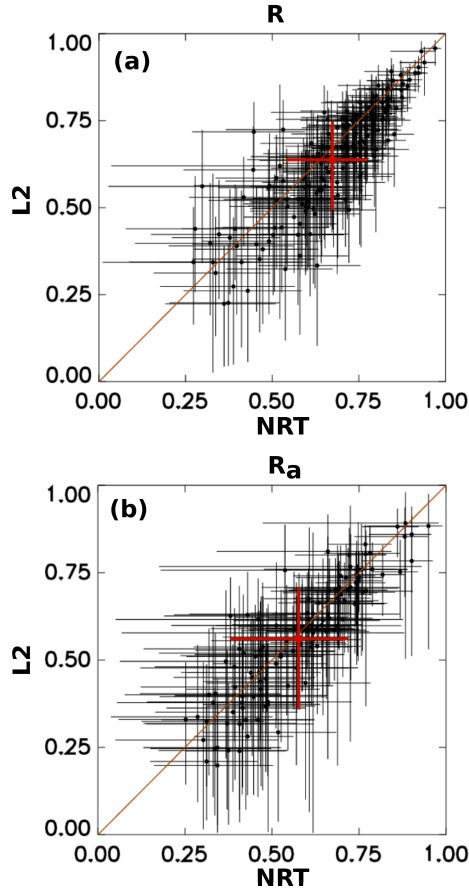
Figure 6. (a) Scatter plots showing the Pearson correlation coefficient of the NRT and L2 SM time series with respect to the in situ measurements (R). The errorbars account for the 95% confidence intervals. The red symbols represents an averaged value. (b) Same as (a) but for the anomalies time series (R_a).

Table 2 summarizes the results in the form of averages over all the sites (for the Pearson correlation also the median value is given). Both SMOS products show a similar mean bias with respect to the in situ measurements, while the mean STDD and RMSD are slightly lower for the NRT SM product. In order to get further insight into the intrinsic quality differences of both datasets, the same statistics have been computed but only using times for which both SMOS products are retrieved. The results are also shown in Table 2. The differences in the evaluation of both products decreases, but the NRT product still shows a larger correlation and lower STDD with respect to the in situ measurements than the L2 product.

Since the mean or median values alone do not show the full picture of the evaluation for more than 100 sites, Figs. 5a,b show boxplots for the Pearson correlation coefficient of the time series (R) and the anomalies times series (R_a), respectively. Figures 5c,d show boxplots for the bias and the STDD. As expected, there is a large variation from one site to another. The bias and STDD distributions are similar for both products. The correlation is as high as almost 1 for some sites both for the NRT SM

and L2 SM (the maximum is slightly higher for the later). Interestingly, the lower values of the distribution of the correlation are higher for the NRT product. In summary, the bias and R_a distributions are very similar for both products while the NRT product shows a lower STDD and a higher R for the central two quartiles of the distribution (green boxes in Fig. 5). This behaviour was already found by Rodríguez-Fernández et al. (2015), who analysed different NN models to retrieve SM from

5 SMOS observations after training the NN on ECMWF simulated SM fields. When comparing to in situ measurements, the best NNs models showed a higher Pearson R and a lower STDD than those obtained for the ECMWF SM model simulations. These results can be understood because, provided that the training is done with a large number of statistically representative samples, the NN will not be significantly affected by outliers or inconsistent values during the training phase. The NN output is the most likely (in the sense of the Bayes theorem) SM value taking into account a given set of input data. Thus, a good NN

10 model can show slightly better quality metrics when compared to in situ measurements than the dataset used as reference to train the NN.

Finally, Fig. 6 shows scatter plots of the correlation for the time series and for the anomalies time series taking into account the respective confidence intervals. For most of the sites, both products show the same statistics with respect to the in situ measurements and globally, the scatter plot points lie close to the 1:1 line.

15 **6 Conclusions**

This paper describes the ESA SMOS NRT SM processor and the first evaluation of this new operational dataset. This processor is based on a neural network algorithm that uses SMOS NRT brightness temperatures and ECMWF IFS soil temperature in the 0–7 cm layer as input. It has been trained with SMOS Level 2 SM data as reference. The SMOS NRT brightness temperatures have been transformed from the antenna reference frame to the ground reference frame to express the polarization as horizontal

20 and vertical components. In addition they have been binned in 5°-width incidence angle bins. Soil temperature and snow cover forecasts from ECMWF IFS are used to filter out frozen soil or soil covered by snow. The uncertainties of the NRT SM data were estimated from the input brightness temperature uncertainties.

The SMOS NRT SM product was evaluated with respect to the original SMOS Level 2 SM product using several months of data. The NRT SM product compares well with the L2 product. The most significant difference is that the NRT SM dataset

25 shows local negative bias at the positions were the highest SM values were found (basically under tropical forest).

The SMOS NRT SM product was also evaluated with respect to in situ measurements of SM over the SCAN and USCRN networks. The NRT product shows similar performances to those of the L2 product. Actually, the mean and median correlation are slightly higher than those obtained for the L2 product. In addition, the standard deviation of the difference with respect to the in situ measurements is lower for the NRT product than for the L2 product.

30 In summary, the SMOS NRT SM product shows similar performance to the Level 2 product but it has the advantage to be available in NRT. NRT brightness temperatures are received by ECMWF from ESA in less than three hours after sensing. The NRT SM production takes on average 15 minutes (the arrival of new NRT TBs is checked every 30 minutes and the actual NRT SM production takes a few minutes). The SMOS NRT SM product is delivered to ESA and EUMETSAT for dissemination via

EUMETCast. Therefore, the SMOS NRT SM data are available for a large range of operational applications such as numerical weather prediction, hydrological forecasting and crop modelling.

Data availability. The datasets used in this study (Sect. 2) are publicly available. The SMOS L2 SM and NRT SM data can be downloaded from ESA. The SMOS NRT SM data is also available via EUMETCast in NRT. The in situ measurements can be downloaded from the International Soil Moisture Network (Dorigo et al., 2011).

Appendix A: NRT SM algorithm

The SMOS NRT has been described qualitatively in Sect. 3. The current section describes the algorithm and the output uncertainties calculation in detail. Complementary information can be found in Rodríguez-Fernández et al. (2016) and Muñoz-Sabater et al. (2016).

A1 Neural network specification

The NN discussed in Sect. 3 has two layers. The first layer contains $j = 1, ..., n_{L1}$ nodes or neurons with an hyperbolic tangent as activation function. The second layer contains a single neuron with a linear function as activation function. The number of input elements n_{in} is 13: 6 T_b values (H and V for incidence angle bins from 30 to 45°), 6 index I (H and V for incidence angle bins from 30 to 45°), and ECMWF soil Temperature. The inputs range should be re-normalized to have values in the $[-1, 1]$ range. If for each input vector element, the minimum and maximum values found during the training phase are given by the vectors v_i^{min} and v_i^{max} $(i = 1, ..., n_{in})$, the normalization can be computed as follows:

$$v_i^{norm} = -1 + 2\frac{v_i - v_i^{min}}{v_i^{max} - v_i^{min}}, \forall i = 1...n_{in} \tag{A1}$$

The normalized input, together with the first layer weights (W_{L1}) and bias B_{L1} are used to compute the first layer outputs v^{L1} as follows:

$$v_j^{L1} = \tanh(\sum_{i=1}^{n_{in}} W_{L1}^{ij} v_i^{norm} + B_{L1}^j), \forall j = 1...n_{L1} \tag{A2}$$

The output of the second layer is computed from the first layer outputs, and the second layer weights (W_{L2}) and bias B_{L2} as follows:

$$v^{L2} = \sum_{j=1}^{n_{L1}} W_{L2}^j v_j^{L1} + B_{L2} \tag{A3}$$

The values of the weights W_{L1} and W_{L2} and the bias B_{L1} and B_{L2} are determined after the training phase. The exact values for the operational NRT SM processor can be found in Muñoz-Sabater et al. (2016). Finally, to obtain the NN output (v^{out}), the output of the second layer has to be re-normalized as follows:

$$v^{out} = v^{L2}_{newMin} + \frac{v^{L2}_{newMax} - v^{L2}_{newMin}}{v^{L2}_{oldMax} - v^{L2}_{oldMin}}(v^{L2} - v^{L2}_{oldMin}); \tag{A4}$$

A2   Neural network output uncertainties

From the definition of $I_{\lambda\phi}(t)$ (Eq. 1) it is possible to compute the uncertainties from the T_b and SM uncertainties. First, Eq. 1 can be rewritten as:

$$I_{\lambda\phi}(t) = SM^{T_b^{min}}_{\lambda\phi} + [SM^{T_b^{max}}_{\lambda\phi} - SM^{T_b^{min}}_{\lambda\phi}] \times I_{1_{\lambda\phi}}(t) \tag{A5}$$

where $I_{1_{\lambda\phi}}(t)$ is given by:

$$I_{1_{\lambda\phi}}(t) = \frac{T_{b_{\lambda\phi}}(t) - T^{min}_{b_{\lambda\phi}}}{T^{max}_{b_{\lambda\phi}} - T^{min}_{b_{\lambda\phi}}} \tag{A6}$$

The uncertainties $\Delta I_{\lambda\phi}(t)$ and $\Delta I_{1_{\lambda\phi}}(t)$ can be computed from uncertainties in T_b measurements, in the maximum and minimum T_b and the associated SM values as follows:

$$\Delta I^2_{\lambda\phi}(t) = [SM^{T_b^{max}}_{\lambda\phi} - SM^{T_b^{min}}_{\lambda\phi}]^2 \big(\Delta I_{1_{\lambda\phi}}(t)\big)^2 + \tag{A7}$$

$$+ [1 - I_{1_{\lambda\phi}}(t)]^2 \big(\Delta SM^{T_b^{min}}_{\lambda\phi}\big)^2 + \tag{A8}$$

$$+ [I_{1_{\lambda\phi}}(t)]^2 \big(\Delta SM^{T_b^{max}}_{\lambda\phi}\big)^2 \tag{A9}$$

Where $\Delta I_{1_{\lambda\phi}}(t)$ is given by:

$$\Delta I^2_{1_{\lambda\phi}}(t) = \frac{\Delta T_{b_{\lambda\phi}}(t)^2}{T^2_{D_{\lambda\phi}}} + \tag{A10}$$

$$+ \frac{(\Delta T^{max}_{b_{\lambda\phi}})^2}{T^2_{D_{\lambda\phi}}}\left(\frac{T_{m_{\lambda\phi}}(t)}{T_{D_{\lambda\phi}}}\right)^2 + \tag{A11}$$

$$+ \frac{(\Delta T^{min}_{b_{\lambda\phi}})^2}{T^2_{D_{\lambda\phi}}}\left(-1 + \frac{T_{m_{\lambda\phi}}(t)}{T_{D_{\lambda\phi}}}\right)^2 \tag{A12}$$

as a function of the uncertainty of the local instantaneous measurement $\Delta T_{b_{\lambda\phi}}(t)$ and the uncertainties of the local extreme T_b values ($\Delta T^{max}_{b_{\lambda\phi}}$ and $\Delta T^{min}_{b_{\lambda\phi}}$).

The uncertainties of the NN output given by Eqs. A1-A4 can be estimated from the uncertainties in the input vector elements (Δv_i) as follows. First the uncertainties of the normalized input vector can be computed as:

$$\Delta v^{norm}_i = 2\frac{\Delta v_i}{v^{max}_i - v^{min}_i}, \ \forall i = 1...n_{in} \tag{A13}$$

Using those quantities, the uncertainty of the two layers neural network given by Eqs. A2 and A3 can be expressed as:

$$(\Delta v^{L2})^2 = \sum_{i=1}^{n_{in}} \left\{ (\Delta v_i^{norm})^2 \left(\sum_{j=1}^{n_{L1}} W_{L2}^j W_{L1}^{ij} \sigma^j \right)^2 \right\} \tag{A14}$$

where σ^j is given by:

$$\sigma^j = 1 - \tanh^2 (\sum_{i=1}^{n_{in}} W_{L1}^{ij} v_i^{norm} + B_{L1}^j), \ \forall j = 1...n_{L1} \tag{A15}$$

5   It is worth noting that in the current implementation the neural network weights are assumed to be constant after training. There exist some methods to estimate the additional output uncertainty that originates from the neural network weight uncertainty that comes from the uncertainties in the reference data used for the training (see for instance Aires et al., 2004) but they are too complex to be implemented in the SMOS NRT SM operational processor. In contrast, some uncertainties in the reference data used for the training have already been taken into account through $\Delta SM_{\lambda\phi}^{T_b^{min}}$ and $\Delta SM_{\lambda\phi}^{T_b^{max}}$ in Eq. A9.

10   Finally, the uncertainty after the normalization of the output can be written as:

$$\Delta v^{out} = \frac{v_{newMax}^{L2} - v_{newMin}^{L2}}{v_{oldMax}^{L2} - v_{oldMin}^{L2}} \Delta v^{L2}; \tag{A16}$$

Expressing the output uncertainty as Eq. A14 implies that the vector elements v_i are independent. However, when using index I as input as well as the actual T_b measurements, some elements are not independent. Since the uncertainties in Eq. A14 are expressed in quadratic form, Eq. A14 gives an upper limit to the output uncertainty.

15   *Author contributions.* NJRF and JMS are the principal authors of this manuscript. The neural network approach and the uncertainties calculation were designed by NJRF and PR. JMS, PR and NJRF implemented the operational version of the SMOS NRT SM algorithm. The global evaluation and the comparison to in situ measurements have been done by NJRF and CA, respectively. YK, PdR, MD and SM reviewed the system design and the results.

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

20   *Acknowledgements.* The authors acknowledge useful discussions with Filipe Aires and Catherine Prigent on neural network algorithms to retrieve soil moisture from microwave observations. The authors acknowledge interesting and constructive comments from two anonymous referees and Prof. W. Wagner that improved the manuscript, in particular regarding the comparison to the NRT ASCAT processing chain.

References

Aires, F., Prigent, C., and Rossow, W.: Neural network uncertainty assessment using Bayesian statistics with application to remote sensing: 3. Network Jacobians, Journal of Geophysical Research: Atmospheres (1984–2012), 109, 2004.

5 Al Bitar, A., Leroux, D., Kerr, Y. H., Merlin, O., Richaume, P., Sahoo, A., and Wood, E.: Evaluation of SMOS Soil Moisture Products Over Continental US Using the SCAN/SNOTEL Network, Geoscience and Remote Sensing, IEEE Transactions on, 50, 1572–1586, doi:10.1109/TGRS.2012.2186581, 2012.

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

10 maps, Earth System Science Data, 9, 293–315, doi:10.5194/essd-2017-1, http://www.earth-syst-sci-data-discuss.net/essd-2017-1/, 2017.

Al-Yaari, A., Wigneron, J.-P., Ducharne, A., Kerr, Y., De Rosnay, P., De Jeu, R., Govind, A., Al Bitar, A., Albergel, C., Munoz-Sabater, J., et al.: Global-scale evaluation of two satellite-based passive microwave soil moisture datasets (SMOS and AMSR-E) with respect to Land Data Assimilation System estimates, Remote Sensing of Environment, 149, 181–195, 2014a.

Al-Yaari, A., Wigneron, J.-P., Ducharne, A., Kerr, Y., Wagner, W., De Lannoy, G., Reichle, R., Al Bitar, A., Dorigo, W., Richaume, P., et al.:

15 Global-scale comparison of passive (SMOS) and active (ASCAT) satellite based microwave soil moisture retrievals with soil moisture simulations (MERRA-Land), Remote Sensing of Environment, 152, 614–626, 2014b.

Albergel, C., Rüdiger, C., Carrer, D., Calvet, J., Fritz, N., Naeimi, V., Bartalis, Z., and Hasenauer, S.: An evaluation of ASCAT surface soil moisture products with in-situ observations in Southwestern France, Hydrology and Earth System Sciences, 13, 115–124, 2009.

Albergel, C., de Rosnay, P., Gruhier, C., Muñoz-Sabater, J., Hasenauer, S., Isaksen, L., Kerr, Y., and Wagner, W.: Evaluation of remotely

20 sensed and modelled soil moisture products using global ground-based in situ observations, Remote Sensing of Environment, 118, 215–226, 2012.

Alvarez-Garreton, C., Ryu, D., Western, A. W., Crow, W. T., Su, C.-H., and Robertson, D. R.: Dual assimilation of satellite soil moisture to improve streamflow prediction in data-scarce catchments, Water Resources Research, 52, 5357–5375, 2016.

Anterrieu, E. and Khazaal, A.: Brightness Temperature Map Reconstruction from Dual-Polarimetric Visibilities in Synthetic Aperture

25 Imaging Radiometry, IEEE Transactions on Geoscience and Remote Sensing, 46, 606–612, doi:10.1109/TGRS.2007.914799, http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4456807, 2008.

Barbu, A., Calvet, J.-C., Mahfouf, J.-F., and Lafont, S.: Integrating ASCAT surface soil moisture and GEOV1 leaf area index into the SURFEX modelling platform: a land data assimilation application over France, Hydrology and Earth System Sciences, 18, 173–192, 2014.

30 Bartalis, Z., Wagner, W., Naeimi, V., Hasenauer, S., Scipal, K., Bonekamp, H., Figa, J., and Anderson, C.: Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT), Geophysical Research Letters, 34, 20 401, doi:10.1029/2007GL031088, 2007.

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.

Bircher, S., Skou, N., and Kerr, Y. H.: Validation of SMOS L1C and L2 Products and Important Parameters of the Retrieval Algo-

35 rithm in the Skjern River Catchment, Western Denmark, IEEE Transactions on Geoscience and Remote Sensing, 51(5), 2969–2985, doi:10.1109/TGRS.2012.2215041, http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6355980, 2013.

Blankenship, C. B., Case, J. L., Zavodsky, B. T., and Crosson, W. L.: Assimilation of SMOS Retrievals in the Land Information System, IEEE Transactions on Geoscience and Remote Sensing, 54, 2016.

Brocca, L., Pellarin, T., Crow, W. T., Ciabatta, L., Massari, C., Ryu, D., Su, C.-H., Rüdiger, C., and Kerr, Y.: Rainfall estimation by inverting SMOS soil moisture estimates: A comparison of different methods over Australia, Journal of Geophysical Research: Atmospheres, 121, 2016.

5   Burgin, M., Colliander, A., Njoku, E. G., S, C., Cabot, F., Kerr, Y., Bindlish, R., Jackson, T., Entekhabi, D., and Yueh, S.: A Comparative Study of the SMAP Passive Soil Moisture Product with existing Satellite-Based Soil Moisture Products, IEEE Transactions on Geoscience and Remote Sensing, submitted, 2016.

Cenci, L., Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Boni, G., and Rudari, R.: Assimilation of H-SAF Soil Moisture Products for Flash Flood Early Warning Systems. Case Study: Mediterranean Catchments, IEEE Journal of Selected Topics in Applied

10   Earth Observations and Remote Sensing, 9, 5634–5646, 2016.

Chen, F., Crow, W. T., Starks, P. J., and Moriasi, D. N.: Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture, Advances in Water Resources, 34, 526–536, 2011.

Choudhury, B. J., Schmugge, T. J., and Mo, T.: A parameterization of effective soil-temperature for microwave emission, J. Geophys. Res. Oceans Atmos., 87, c2, 1301–1304, 1982.

15   Crow, W. T., Huffman, G. J., Bindlish, R., and Jackson, T. J.: Improving satellite-based rainfall accumulation estimates using spaceborne surface soil moisture retrievals, Journal of Hydrometeorology, 10, 199–212, 2009.

de Rosnay, P., Dragosavac, M., Drusch, M., Gutiérrez, A., Rodríguez López, M., Wright, N., Muñoz Sabater, J., and R., C.: SMOS NRT BUFR specification, SMOS-NRT-BUFR-ECMWF - v2.0, ECMWF, 2012.

de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C., and Isaksen, L.: A simplified Extended Kalman Filter for the global

20   operational soil moisture analysis at ECMWF, Quarterly Journal of the Royal Meteorological Society, 139, 1199–1213, 2013.

Dharssi, I., Bovis, K., Macpherson, B., and Jones, C.: Operational assimilation of ASCAT surface soil wetness at the Met Office, Hydrology and Earth System Sciences, 15, 2729–2746, 2011.

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

25   Sciences, 15, 1675–1698, 2011.

Ferrazzoli, P., Guerriero, L., and Wigneron, J.-P.: Simulating L-band emission of forests in view of future satellite applications, Geoscience and Remote Sensing, IEEE Transactions on, 40, 2700–2708, 2002.

Gruber, A., Dorigo, W. A., Zwieback, S., and Wagner, W.: Characterizing Coarse-Scale Representativeness of in situ Soil Moisture Measurements from the International Soil Moisture Network, doi:10.2136/vzj2012.0170, 2012.

30   Gutierrez, A. and Canales Molina, A.: SMOS NRT Product Format Specification, Tech. Rep. SO-ID-DMS-GS-0002, DEIMOS Space, 2010.

Hawke, R. and McConchie, J.: In situ measurement of soil moisture and pore-water pressures in an 'incipient'landslide: Lake Tutira, New Zealand, Journal of environmental management, 92, 266–274, 2011.

Jackson, T., O'Neill, P., Njoku, E., S, C., Bindlish, R., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez-Fernández, J., González-Zamora,

35   A., Seyfried, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Su, Z., van der Velde, R., Asanuma, J., Palecki, M., Small, E., Zreda, M., Calvet, J., Crow, W., Kerr, Y., Yueh, S., and Entekhabi, D.: Calibration and Validation for the L2/3-SM-P Version 3 Data Products, Tech. rep., Jet Propulsion Laboratory, Pasadena, CA. SMAP Project, JPL D-93720,, 2016.

Jackson, T. J., Bindlish, R., Cosh, M. H., Zhao, T., Starks, P. J., Bosch, D. D., Seyfried, M., Moran, M. S., Goodrich, D. C., Kerr, Y. H., et al.: Validation of Soil Moisture and Ocean Salinity (SMOS) soil moisture over watershed networks in the US, Geoscience and Remote Sensing, IEEE Transactions on, 50, 1530–1543, 2012.

5   Kerr, Y., Waldteufel, P., Wigneron, J.-P., Delwart, S., Cabot, F., Boutin, J., Escorihuela, M.-J., Font, J., Reul, N., Gruhier, C., Juglea, S., Drinkwater, M., Hahne, A., Martin-Neira, M., and Mecklenburg, S.: The SMOS Mission: New Tool for Monitoring Key Elements ofthe Global Water Cycle, Proceedings of the IEEE, 98, 666–687, doi:10.1109/JPROC.2010.2043032, 2010.

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,

10   1384–1403, doi:10.1109/TGRS.2012.2184548, 2012.

Kerr, Y. H., Al-Yaari, A., Rodriguez-Fernandez, N., Parrens, M., Molero, B., Leroux, D., Bircher, S., Mahmoodi, A., Mialon, A., Richaume, P., et al.: Overview of SMOS performance in terms of global soil moisture monitoring after six years in operation, Remote Sensing of Environment, 180, 40–63, 2016.

Koster, R. D., Dirmeyer, P. A., Guo, Z., Bonan, G., Chan, E., Cox, P., Gordon, C., Kanae, S., Kowalczyk, E., Lawrence, D., et al.: Regions of

15   strong coupling between soil moisture and precipitation, Science, 305, 1138–1140, 2004.

Lahoz, W. A. and De Lannoy, G. J.: Closing the gaps in our knowledge of the hydrological cycle over land: conceptual problems, Surveys in Geophysics, pp. 1–38, 2013.

Leroux, D.: Analyses statistiques et modèles d'inversion pour la validation des données d'humidité des sols de la mission SMOS, http://tel.archives-ouvertes.fr/tel-00776536/, 2012.

20   Leroux, D. J., Kerr, Y. H., Al Bitar, A., Bindlish, R., Jackson, T. J., Berthelot, B., and Portet, G.: Comparison between SMOS, VUA, ASCAT, and ECMWF soil moisture products over four watersheds in US, Geoscience and Remote Sensing, IEEE Transactions on, 52, 1562–1571, 2014.

Lievens, H., De Lannoy, G., Al Bitar, A., Drusch, M., Dumedah, G., Franssen, H.-J. H., Kerr, Y., Tomer, S. K., Martens, B., Merlin, O., et al.: Assimilation of SMOS soil moisture and brightness temperature products into a land surface model, Remote Sensing of Environment,

25   180, 292–304, 2016.

Louvet, S., Pellarin, T., al Bitar, A., Cappelaere, B., Galle, S., Grippa, M., Gruhier, C., Kerr, Y., Lebel, T., Mialon, A., et al.: SMOS soil moisture product evaluation over West-Africa from local to regional scale, Remote Sensing of Environment, 156, 383–394, 2015.

Martens, B., Miralles, D., Lievens, H., Fernández-Prieto, D., and Verhoest, N.: Improving terrestrial evaporation estimates over continental Australia through assimilation of SMOS soil moisture, International Journal of Applied Earth Observation and Geoinformation, 48, 146–

30   162, 2016.

McColl, K. A., Alemohammad, S. H., Akbar, R., Konings, A. G., Yueh, S., and Entekhabi, D.: The global distribution and dynamics of surface soil moisture, Nature Geoscience, 10, 100–104, 2017.

Mecklenburg, S., Drusch, M., Kerr, Y. H., Font, J., Martin-Neira, M., Delwart, S., Buenadicha, G., Reul, N., Daganzo-Eusebio, E., Oliva, R., and Crapolicchio, R.: ESA's Soil Moisture and Ocean Salinity Mission: Mission Performance and Operations, IEEE Transactions

35   on Geoscience and Remote Sensing, 50, 1354–1366, doi:10.1109/TGRS.2012.2187666, http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6175118, 2012.

Montosi, E., Manzoni, S., Porporato, A., and Montanari, A.: An ecohydrological model of malaria outbreaks, Hydrology and Earth System Sciences, 16, 2759, 2012.

Muñoz-Sabater, J.: Incorporation of passive microwave Brightness Temperatures in the ECMWF soil moisture analysis, Remote Sensing, 7, 5758–5784, 2015.

Muñoz Sabater, J., Jarlan, L., Calvet, J.-C., Bouyssel, F., and De Rosnay, P.: From near-surface to root-zone soil moisture using different assimilation techniques, Journal of Hydrometeorology, 8, 194–206, 2007.

Muñoz Sabater, J., Fouilloux, A., and de Rosnay, P.: Technical implementation of SMOS data in the ECMWF Integrated Forecasting System, Geoscience and Remote Sensing Letters, IEEE, 9, 252–256, 2012.

Muñoz-Sabater, J., Rodríguez-Fernández, N. J., Richaume, P., de Rosnay, P., and Kerr, Y. H.: SMOS Near-Real-Time Soil Moisture processor., Tech. Rep. ECMWF ESA report, ECMWF, Reading, UK, 2016.

Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E., and Anquetin, S.: Flash flood warning based on rainfall thresholds and soil moisture conditions: An assessment for gauged and ungauged basins, Journal of Hydrology, 362, 274–290, 2008.

Pauwels, V. R., Hoeben, R., Verhoest, N. E., De Troch, F. P., and Troch, P. A.: Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values, Hydrological Processes, 16, 995–1013, 2002.

Pellarin, T., Ali, A., Chopin, F., Jobard, I., and Bergès, J.-C.: Using spaceborne surface soil moisture to constrain satellite precipitation estimates over West Africa, Geophysical Research Letters, 35, 2008.

Peters, J., Conte, A., Verhoest, N. E., De Clercq, E., Goffredo, M., De Baets, B., Hendrickx, G., Ducheyne, E., et al.: On the relation between soil moisture dynamics and the geographical distribution of Culicoides imicola, Ecohydrology, 7, 622–632, 2014.

Ponziani, F., Pandolfo, C., Stelluti, M., Berni, N., Brocca, L., and Moramarco, T.: Assessment of rainfall thresholds and soil moisture modeling for operational hydrogeological risk prevention in the Umbria region (central Italy), Landslides, 9, 229–237, 2012.

Ray, R. L., Jacobs, J. M., and Cosh, M. H.: Landslide susceptibility mapping using downscaled AMSR-E soil moisture: A case study from Cleveland Corral, California, US, Remote sensing of environment, 114, 2624–2636, 2010.

Ridler, M.-E., Madsen, H., Stisen, S., Bircher, S., and Fensholt, R.: Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil-vegetation-atmosphere transfer model in Western Denmark, Water Resources Research, 50, 8962–8981, 2014.

Rodríguez-Fernández, N. J., Richaume, P., Aires, F., Prigent, C., Kerr, Y. H. Kolassa, J., Jiménez, C., Cabot, F., and Mahmoodi, A.: Soil moisture retrieval from SMOS observations using neural networks, Tech. Rep. SMOS Ground Segment SO-TN-CB-GS-038, CESBIO, Toulouse, France, 2013.

Rodríguez-Fernández, N. J., Aires, F., Richaume, P., Kerr, Y. H., Prigent, C., Kolassa, J., Cabot, F., Jiménez, C., Mahmoodi, A., and Drusch, M.: Soil moisture retrieval using neural networks: application to SMOS, IEEE Transactions on Geoscience and Remote Sensing, Volume:53 , Issue: 11, 5991–6007, doi:10.1109/TGRS.2015.2430845, 2015.

Rodríguez-Fernández, N. J., Richaume, P., Muñoz-Sabater, J., de Rosnay, P., and Kerr, Y. H.: SMOS Near-Real-Time Soil Moisture processor. Recommended neural network configuration and algorithm description, Tech. Rep. SMOS Ground Segment SO-TN-CB-GS-049, CESBIO, Toulouse, France, 2016.

Sahr, K., White, D., and J., K. A.: Geodesic discrete global grid systems cartography., Cartography and Geographic Information Science, 30(2), 30(2), 121–134, 2003.

Schaefer, G., Cosh, M., and Jackson, T.: The USDA natural resources conservation service soil climate analysis network (SCAN), Journal of Atmospheric and Oceanic Technology, 24, 2073–2077, 2007.

Scholze, M., Kaminski, T., Knorr, W., Blessing, S., Vossbeck, M., Grant, J., and Scipal, K.: Simultaneous assimilation of SMOS soil moisture and atmospheric CO_2 in-situ observations to constrain the global terrestrial carbon cycle, Remote Sensing of Environment, 180, 334–345, 2016.

Seneviratne, S. I., Lüthi, D., Litschi, M., and Schär, C.: Land–atmosphere coupling and climate change in Europe, Nature, 443, 205–209, 2006.

Tuttle, S. and Salvucci, G.: Empirical evidence of contrasting soil moisture–precipitation feedbacks across the United States, Science, 352, 825–828, 2016.

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.

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, doi:10.1016/S0034-4257(99)00036-X, http://www.sciencedirect.com/science/article/pii/S003442579900036X, 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.

Wanders, N., Karssenberg, D., Bierkens, M., Parinussa, R., de Jeu, R., van Dam, J., and de Jong, S.: Observation uncertainty of satellite soil moisture products determined with physically-based modeling, Remote Sensing of Environment, 127, 341–356, 2012.

Wanders, N., Karssenberg, D., Roo, A. d., De Jong, S., and Bierkens, M.: The suitability of remotely sensed soil moisture for improving operational flood forecasting, Hydrology and Earth System Sciences, 18, 2343–2357, 2014.

Wigneron, J. P., Laguerre, L., and Kerr, Y. H.: A simple parameterization of the L-band microwave emission from rough agricultural soils, Geoscience and Remote Sensing, IEEE Transactions on, 39, 1697–1707, 2001.

Wigneron, J.-P., Kerr, Y., Waldteufel, P., Saleh, K., Escorihuela, M.-J., Richaume, P., Ferrazzoli, P., De Rosnay, P., Gurney, R., Calvet, J.-C., et al.: L-band Microwave Emission of the Biosphere (L-MEB) Model: Description and calibration against experimental data sets over crop fields, Remote Sensing of Environment, 107, 639–655, 2007.

Xu, X., Tolson, B. A., Li, J., Staebler, R. M., Seglenieks, F., Haghnegahdar, A., and Davison, B.: Assimilation of SMOS soil moisture over the Great Lakes basin, Remote Sensing of Environment, 169, 163–175, 2015.