Preprint version

Paper Received 10 July 2017; Received in revised form 30 October 2017; Accepted 30 October 2017, Published on International Journal of Applied Earth Observation and Geoinformation 65 (2018) 114–123.
http://dx.doi.org/10.1016/j.jag.2017.10.010

ON THE SYNERGY OF SMAP, AMSR2 AND SENTINEL-1 FOR RETRIEVING SOIL MOISTURE.

E. Santi[1], S. Paloscia[1], S. Pettinato[1], L. Brocca[2], L. Ciabatta[2], D. Entekhabi[3]

[1]Institute of Applied Physics - National Research Council, Florence, Italy
[2]Research Institute for Geo-Hydrological Protection - National Research Council, Perugia, Italy
[3]Massachusetts Institute of Technology, Parsons Laboratory, Cambridge, US
e.santi@ifac.cnr.it

Abstract. An algorithm for retrieving soil moisture content (SMC) from synergic use of both active and passive microwave acquisitions is presented. The algorithm takes advantage of the integration of microwave data from SMAP, Sentinel-1 and AMSR2 for overcoming the SMAP radar failure and obtaining a SMC product at enhanced resolution (0.1° x 0.1°) and improved accuracy with respect to the original SMAP radiometric SMC product. A disaggregation technique based on the Smoothing filter based intensity modulation (SFIM) allows combining the radiometric and SAR data. Disaggregated microwave data are used as inputs of an Artificial Neural Networks (ANN) based algorithm, which is able to exploit the synergy between active and passive acquisitions. The algorithm is defined, trained and tested using the SMEX02 experimental dataset and data simulated by forward electromagnetic models based on the Radiative Transfer Theory. Then the algorithm is adapted to satellite data and tested using one year of SMAP, AMSR2 and Sentinel-1 co-located data on a flat agricultural area located in the Po Valley, in northern Italy. Spatially distributed SMC values at 0.1° x 0.1° resolution generated by the Soil Water Balance Model (SWBM) are considered as reference for this purpose. The synergy of SMAP, Sentinel-1 and AMSR2 allowed increasing the correlation between estimated and reference SMC from R ≅ 0.68 of the SMAP based retrieval up to R ≅ 0.86 of the combination SMAP+Sentinel-1+AMSR2. The corresponding Root Mean Square Error (RMSE) decreased from RMSE ≅ 0.04 m^3/m^3 to RMSE ≅ 0.024 m^3/m^3.

Keywords — soil moisture, SMAP, Sentinel-1, electromagnetic models, neural networks.

1 INTRODUCTION

Satellite remote sensing has become an essential tool for observing the main land surface parameters at several temporal and spatial scales. The ever-increasing number of satellite missions and the consequent augmented availability of long-term global data sets of land surface parameters have improved noticeably our knowledge on the carbon, energy and hydrological cycles. In this respect, the moisture content of soil (SMC), which is one of the main driving factors of the hydrological cycle, is receiving an increasing attention, and dedicated satellite sensors have been recently launched. Microwaves are the most suitable frequencies for such a purpose, due to their independence of cloud cover and solar illumination, and, especially in the low part of the microwave spectrum, for the direct sensitivity to the water content of the observed surfaces. Besides the Soil Moisture and Ocean Salinity (SMOS) (Silvestrin et al. 2001, Mecklenburg et al. 2012), operating at L band, one of most popular satellite sensors dedicated to the global SMC monitoring is SMAP (Entekhabi et al. 2010, Entekhabi et al. 2004), which is characterized by the unique feature of carrying on board both a radiometer and a radar operating in the same band (L band), although with different spatial resolutions.

The aim of SMAP was in improving the accuracy and the spatial resolution of SMC available products by fusing information from radar backscattering (σ^0) at higher resolution (1-3 Km) and radiometric brightness temperature (Tb) at lower resolution ($\cong$ 40 Km). SMAP mission was planned for estimating SMC at three different resolutions: 36 Km (radiometer only), 9 Km (radar + radiometer), and 3 Km (radar only). Das et al. (2011) presented an algorithm that estimates soil moisture by combining SMAP Radiometer and Radar data. The algorithm is based on the integration of the spatial heterogeneity at fine resolution detected by SAR observations with the coarse resolution radiometer measurements for generating a disaggregated Tb at intermediate-resolution. The algorithm was validated using airborne data from the passive and active L band system and simulated data. Results indicated that the algorithm allows retrieving soil moisture at 9-km resolution, within the SMAP accuracy requirements over slightly vegetated areas (Das et al. 2014).

Unfortunately, the radar onboard SMAP failed a few months after the launch, preventing the generation of the two SMC products at higher resolution. The research prosecuted therefore on the SMAP radiometric product only. This study uses as a reference the SMAP radiometric-only soil moisture product based on the application of the Backus Gilbert optimal interpolation to the SMAP radiometric data. The gridded radiometer and soil moisture products are posted at 9 km spacing but maintain the SMAP radiometer field-of-view resolution (half-power and -3 dB) of about 30 km to 40 km. These data are hosted at the National Snow and Ice Data Center (NSIDC) as enhanced Level-1C (L1C) 9 Km Brightness Temperature product. L2 and L3 enhanced SMC products at 9 Km posting derived from L1C SMAP Tb data are also distributed by NSIDC (O'Neill et al. 2016).

The possibility of exploiting the SMAP synergy with other microwave sensors, for replacing the SMAP radar data and restoring the combined radar and radiometer SMC product has been also investigated. Das et al. (2016) identified the Sentinel-1A SAR data as the most suitable for replacing the SMAP radar acquisitions, despite the different frequency, due to the similar orbit configuration, that allows overlapping with small time difference and to the availability of co-pol and cross-pol from Sentinel 1A observations.

The synergy of SMAP with Sentinel-1 and AMSR2 is evaluated in this study, which is focused on the development of an algorithm able to merge microwave data from SMAP and AMSR2 radiometers and Sentinel-1 (S-1) SAR for obtaining an improved SMC product at 0.1° x 0.1° resolution, which is similar to the discontinued SMAP radar/radiometer product.

This study is implemented in two steps: at first, the SMEX02 experimental dataset is considered for defining a retrieval algorithm based on Artificial Neural Networks (ANN). ANN are indeed able to take advantage of synergistic L band acquisitions from the JPL Passive Active L and S band Sensor (PALS) and to account for ancillary information on vegetation biomass (Plant water content – PWC) and surface temperature (LST). The algorithm was later on adapted to satellite data and validated using one year of combined SMAP, S-1 and AMSR2 acquisitions on a test area located in the Po valley, in

northern Italy. For this scope, all the SMAP, AMSR2 and S-1 acquisitions collected at the same date and closest time over the selected test area were considered. SMAP and AMSR2 data were disaggregated at the target resolution of 0.1° x 0.1° using the well-established Smoothing filter based intensity modulation disaggregation (SFIM - Santi 2010). Reference SMC data for comparison are derived from the Soil Water Balance Model (SWBM – Brocca et al. 2014 and 2008) that has been largely validated over Italy and Europe. The model SMC takes advantage of high fidelity precipitation information and it is spatially mapped for comparison with airborne and satellite retrievals. It represents a complimentary approach to the use of sparse in situ measurements, which suffer from spatial representativeness errors, for evaluating the retrieval algorithms. Such errors are difficult to distinguish from retrieval errors and they are limited in spatial coverage.

After a description of the two test areas and the experimental data used in this work, provided in section 2, the sensitivity of microwave data to the target SMC is presented in section 3. The ANN algorithm implementation and test on both test areas is described in section 4, that also contains some comparison with the official SMAP SMC product and discussion on the results.

2 TEST AREAS AND DATA

Two different test areas have been considered in this study. The popular Walnut Creek watershed in central Iowa, USA, on which the SMEX02 experiment was carried out, was considered for defining the retrieval concept and for implementing the algorithm, due to the availability of detailed measurements of a large set of surface parameters.

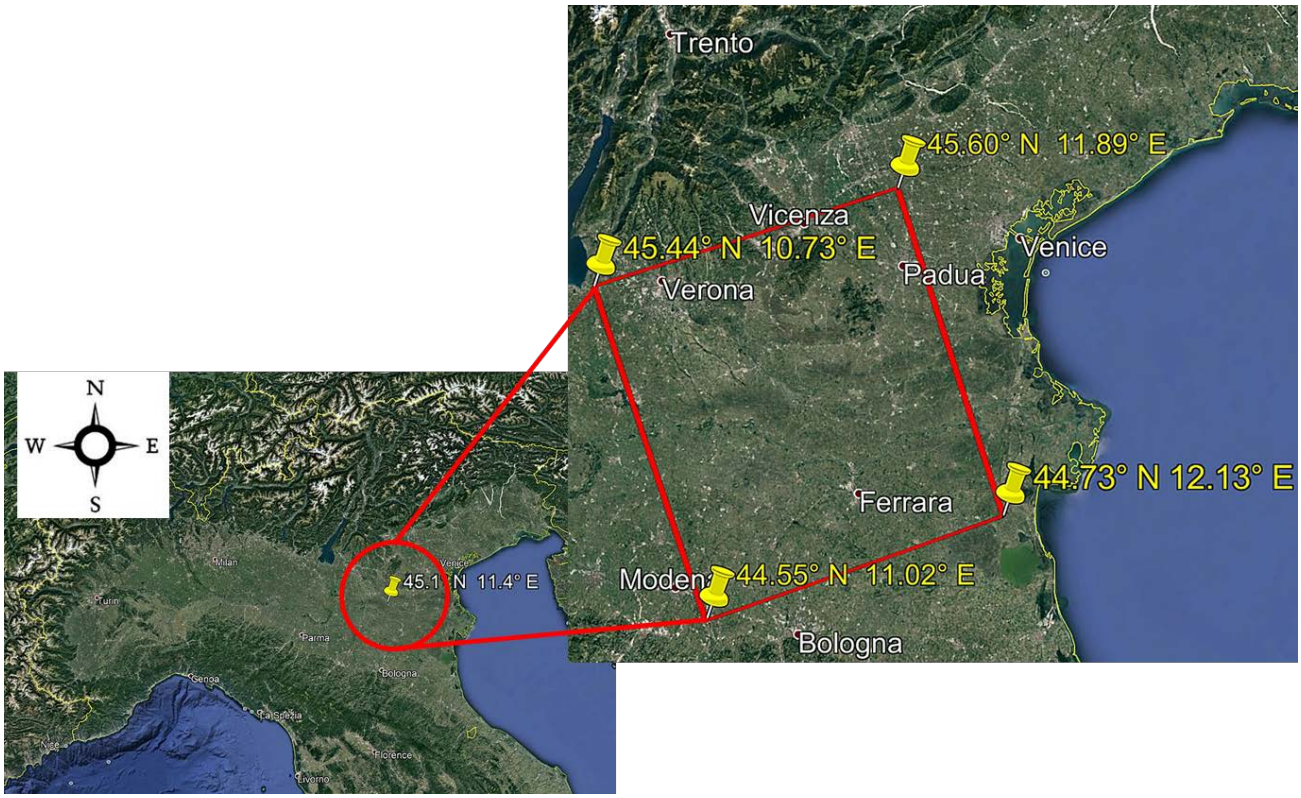


Figure 1. The test area location in Italy from Google Earth.

The possibility of adapting the algorithm to satellite data was then evaluated on a flat agricultural area in the Po Valley, northern Italy, for which time series SMAP, AMSR2 and Sentinel-1 acquisitions were available. The area has an extension of ~120Km x 120Km, and its center coordinates are 11.4° E and 45.1° N (Figure 1).

The area is mainly covered by large fields of wheat and corn, with a small percentage of sugarcane; the Po River crosses it from west to east. The town of Verona and its surroundings, in the upper left corner, is the only area enough urbanized to affect the SMAP radiometric acquisitions.

2.1 SMEX02 dataset

The SMEX02 experimental dataset (Njoku 2003, Bindlish and Jackson, 2002, Jackson and Cosh. 2003) was derived from a large experimental activity carried out in the Walnut Creek watershed in central Iowa, USA, between June and July 2002. It included radar and radiometric acquisitions at L band from the PALS (Passive Active L band System) airborne sensor, C and X band radiometric acquisitions from the Polarimetric Scanning Radiometer (PSR) and ground truth measurements of Soil

Moisture Content (SMC) and Plant Water Content (PWC) collected during the experimental campaign that was carried out in summer 2002. Five flights have been selected (June 25 and 27, July 01, 07 and 08) from the available archive, and all the data (airborne acquisitions and ground truth) have been resampled on a fixed grid of nodes equally spaced at 100 m x 100 m, using linear interpolation. This resampling resulted in a database of about 70.000 microwave measurements at the three frequencies and corresponding ground measurements of SMC and PWC, in a SMC range between 5% and 35% and a PWC range between 0.5 and 5 Kg/m^2. These data have been considered for defining the retrieval concept and training the algorithm.

2.2 Po Valley dataset

One year of overlapped SMAP, AMSR2 and S-1 acquisitions available on the area between April 2015 and March 2016 was considered for this analysis, which was carried out point by point over a grid of 144 nodes spaced at 0.1° x 0.1°. Fourty S-1 images were available in the selected period: for each S-1 image, the corresponding SMAP and AMSR2 acquisitions, collected at the closest time from each other and separated by ascending and descending passes, were considered. The total dataset resulted in about 5000 co-located acquisitions of the three satellites, with the radiometric data almost equally divided between ascending and descending orbits.

S-1 images were downloaded by The Copernicus Open Access Hub, which is the ESA official archive providing open access to Sentinel-1, Sentinel-2 and Sentinel-3 data products. The images were ground range detected (GRD) images in interferometric wide swath mode (IW), with an original spatial resolution of 20 m x 22 m and 5 x 1 looks in range and azimuth, respectively. Images were first mosaicked to cover entirely the test area and then radiometrically calibrated, multilooked and geocoded for obtaining georeferenced and calibrated σ^0 at a spatial resolution of 100 m x 100 m. Calibration was performed by taking into account the local incidence angle (LIA) obtained by terrain correction using a DEM of the area at 90 m x 90 m resolution, derived from SRTM. Finally, S-1 σ^0 data

were firstly degraded at 0.1° x 0.1° resolution using a low pass filtering and then resampled onto the fixed grid using nearest neighbor interpolation.

SMAP L1B Radiometer Half-Orbit Time-Ordered Brightness Temperatures were obtained from the NSIDC data portal, while AMSR2 L1B V2 multifrequency data were downloaded from the JAXA GCOM data portal. In this study, we considered only AMSR2 overpasses that overlapped completely those of SMAP on Italy, at the same date and in the same ascending or descending orbit, in order to avoid any problem in disaggregating both SMAP and AMSR2 data at the target resolution of 0.1° x 0.1°.

SMAP and AMSR2 C and X bands have been disaggregated at the target resolution using the SFIM filtering. This disaggregation technique, which was initially developed for fusing multispectral images and Panchromatic data (Liu, 2000), was adapted to microwave radiometers in (Santi 2010), with the aim of increasing the resolution of C and X bands of AMSR-E up to values close to the sampling rate (i.e. 10 km x 10 km). Other publications aimed at soil moisture and Snow Depth retrieval from AMSR-E and AMSR2 (Santi et al. 2012, Santi et al. 2014, Parinussa et al. 2013), reported extensive validations of SFIM, that is currently applied to the LPRM/AMSR2/GCOM-W1 L2 Downscaled Surface Soil Moisture product, hosted by NASA (https://data.nasa.gov/Earth-Science/LPRM-AMSR2-GCOM-W1-L2-Downscaled-Surface-Soil-Mois/kndx-bvi2). In the current study, SFIM has been modified and adapted to SMAP using the following strategy: for each couple of overlapped SMAP and AMSR2 data, SMAP data have been resampled on the lat/lon grid of the corresponding AMSR2 overpass and then managed as an “additional” channel of AMSR2, applying the same processing proposed in Santi (2010). A key characteristic of SFIM relies upon its capability of preserving the information content of the original data after disaggregation. This means that SFIM has effect in separating the different surface contributions in mixed pixels, while it keeps unchanged the original values on uniform areas.

The problem of obtaining SMC reference value representative of the 0.1° x 0.1° (approximately 100 km^2) pixel was addressed by using the well assessed Soil Water Balance hydrological Model

(SWBM), which was developed in Brocca et al. (2014 and 2008). SWBM accounts for precipitation and temperature data for estimating the water fluxes within the soil layer, and it requires the soil texture information for estimating the soil hydraulic properties (Santi et al., 2016). The model was largely described in Brocca et al. (2014), to which we refer for details, and it was extensively validated in Italy and Europe (Brocca et al. 2011, Lacava et al. 2012, Gumuzzio et al., 2016). Specifically, in Lacava et al. (2012), we compare model simulations with in situ soil moisture data for one station located in the Po Valley with good performance (RMSE=0.022 m^3/m^3). In Santi et al. (2016), the model was used to assess the accuracy of satellite soil moisture data at 10 km scale in central Italy. The good agreement with in situ and satellite soil moisture justified the use of the SWBM for obtaining spatially distributed SMC data from the punctual measurements of meteorological stations available in the area. SWBM is freely available as ®Matlab code (http://dx.doi.org/10.13140/2.1.1460.8323).

3 DATA ANALYSIS

The first step in defining the retrieval algorithm structure was to evaluate the sensitivity of the available microwave data to SMC for both datasets.

3.1 SMEX02

A first attempt to relate the radiometric and radar acquisitions at L band to the measured SMC is shown in Figure 2, where the L band Tb (a) and σ° (b) from PALS are represented as a function of SMC. The poor sensitivity demonstrated by both active and passive PALS acquisitions to SMC (R^2 between 0.17 and 0.19) suggested exploiting the contribution of ancillary parameters as PWC and LST for improving the SMC retrieval.

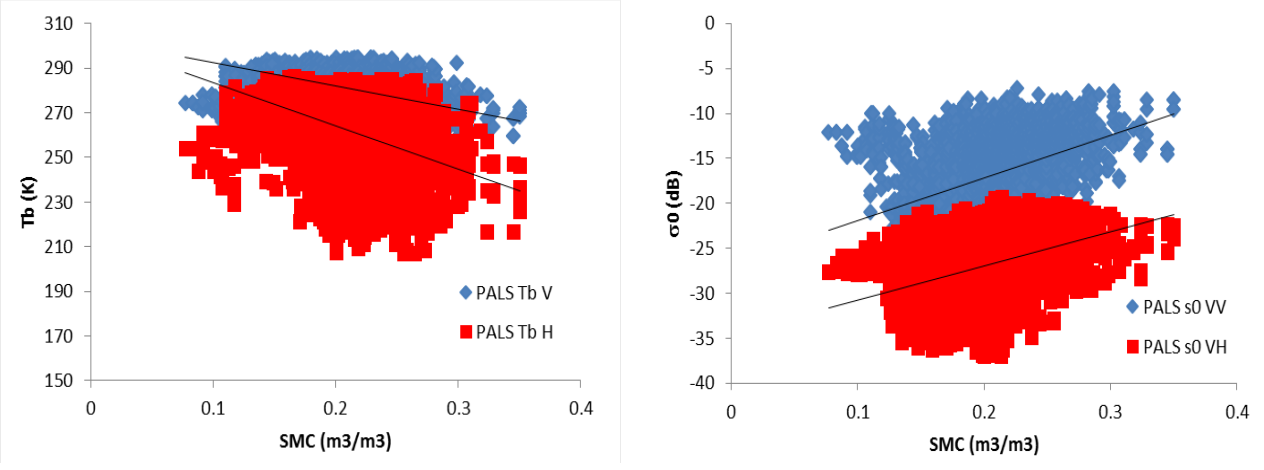


Figure 2 a) PALS Tb at L- band in both H and V polarizations as a function of SMC from SMEX02 database. 1b) PALS σ^0 at both VV and VH polarizations as a function of SMC from SMEX02 database.

SMEX02 dataset includes direct measurements of PWC; however, looking at an operational application of the algorithm for a global monitoring, we focused on the possibility to derive this information from the Polarization Index at X band, that can be obtained for instance from co-located SMAP and AMSR2 overpasses. This index, defined as

$$PI_X = 2\left(\frac{(TbV_X - TbH_X)}{(TbV_X + TbH_X)}\right)$$

was already demonstrated well correlated to the vegetation biomass (Paloscia and Pampaloni, 1988), and was successfully employed for estimating PWC on a global scale (Santi et al. 2012). PWC was related to PI_X using the following relationship, which was obtained comparing PI_X from AMSR-E and PWC from optical data on a wide portion of Africa, from the Sahara Desert to Equatorial forest, thus including a very wide range of vegetation types and landscapes.

$$PWC = -1.77ln(PI_X) - 2.39 \text{ (Kg/m}^2\text{)}$$

By applying this relationship to the PSR acquisitions at X band, the distribution of Figure 3 was obtained, in which the histogram of PWC derived from in-situ measurements and PWC derived from PI_X, using the above relationship, are compared.

Figure 3 demonstrates that PI_X is not able to reproduce the highest and lowest values of the ground measurements, due to the coarse resolution of PSR radiometer, which caused a significant oversampling of the acquired data and a subsequent smoothing of Tb dynamics. However, the mean

values of measured and PI_X-derived PWC are very close - respectively 2.2 Kg/m^2 and 2.5 Kg/m^2, thus demonstrating the PI capability to catch the average PWC value of the area. This relationship confirmed the sensitivity of PI_X to PWC, thus supporting the possibility of using PI_X as additional input of the ANN algorithm.

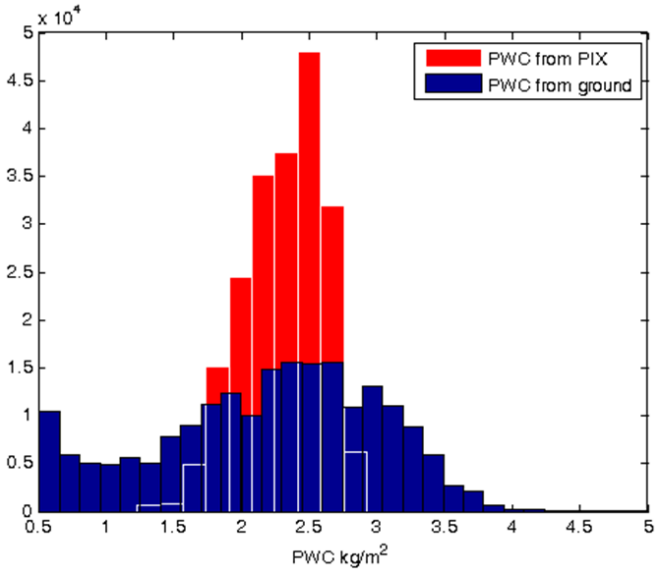


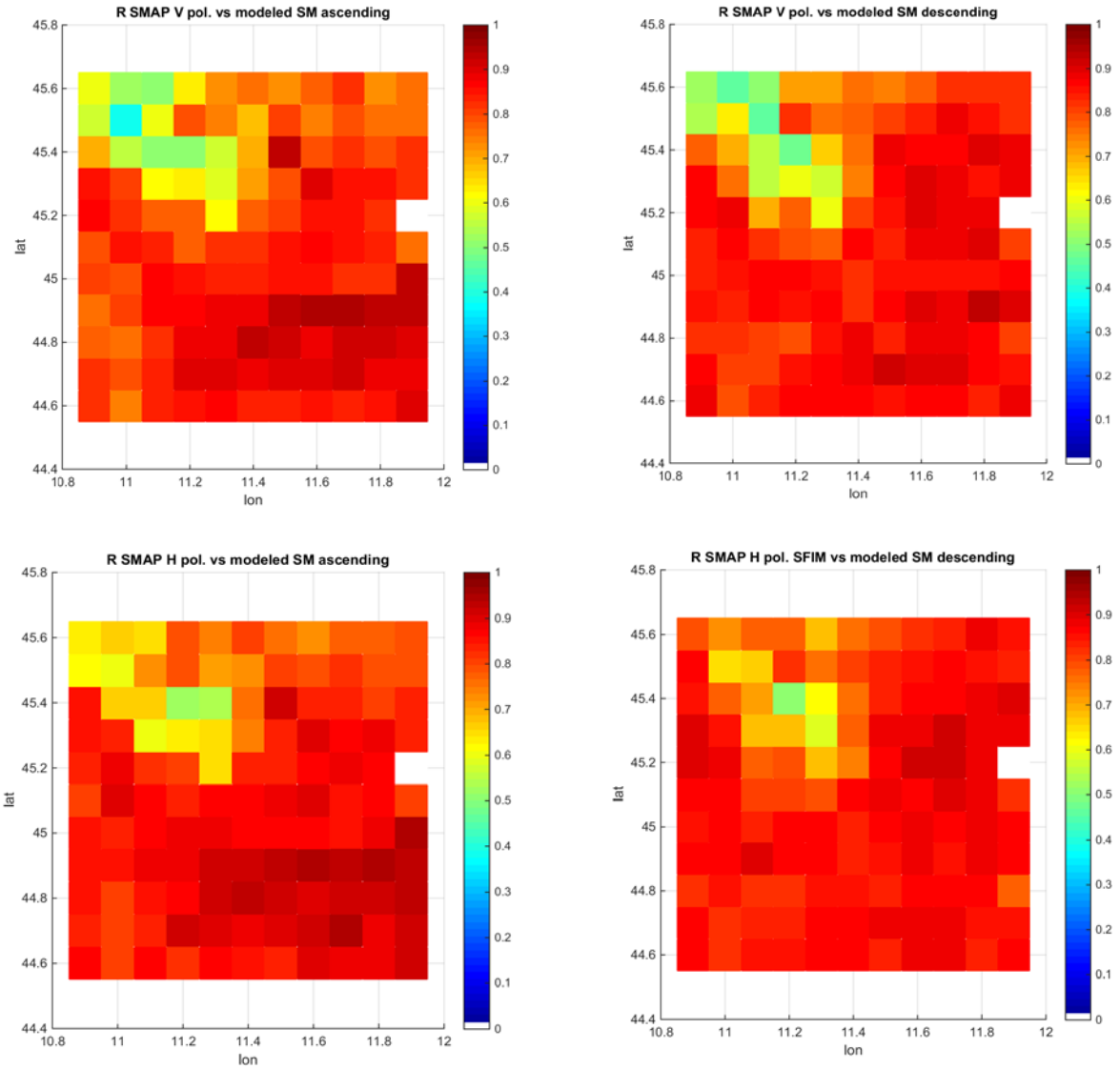
Figure 3. Histograms of PWC derived from PSR PI_X and from in-situ data

Another ancillary parameter strongly affecting the radiometric measurements is the LST. For developing the algorithm, we derived this information from MODIS LST data collected at the same dates of the SMEX02 campaigns. MODIS overflew the area 2-3 hours after PALS; therefore, the LST product was rescaled basing on ancillary information in order to account for the surface temperature variations between the MODIS and PALS acquisitions. In detail, the data collected by the SMEX02 Tower-Based Surface Temperature system, which was composed by 12 flux towers, were considered to compute the temperature difference between Modis LST and in-situ data at the time of flights. The Modis LST product was then scaled by this factor in order to have a distributed information on the surface temperature.

3.2 PO VALLEY

In Figure 4, the maps of the correlation coefficients (R) between the Tb from SMAP and SMC estimated from the SWBM, and σ^0 from S1 and SMC SWBM are shown. R values were computed at

each pixel of the grid considering the whole time series of data, while the overall R values are summarized in Table 1. Figure 4 points out that SMAP acquisitions in both H and V polarizations are very well correlated with SMC simulations, whose values ranged between very dry (0.05 m^3/m^3) and wet (0.3 m^3/m^3) conditions during the considered time period. The contribution of SFIM is limited, since the area is almost flat and homogeneous; however, the disaggregated SMAP Tb are slightly better correlated to SMC with respect to the original data, characterized by coarser resolution (Table 1). The lower R values in the upper left part of the images depend on the presence of urban areas (Verona and its suburbs) that are wide enough for influencing the SMAP acquisitions, while the highest R values in the lower-left part of the images do correspond to agricultural areas with negligible urbanization.



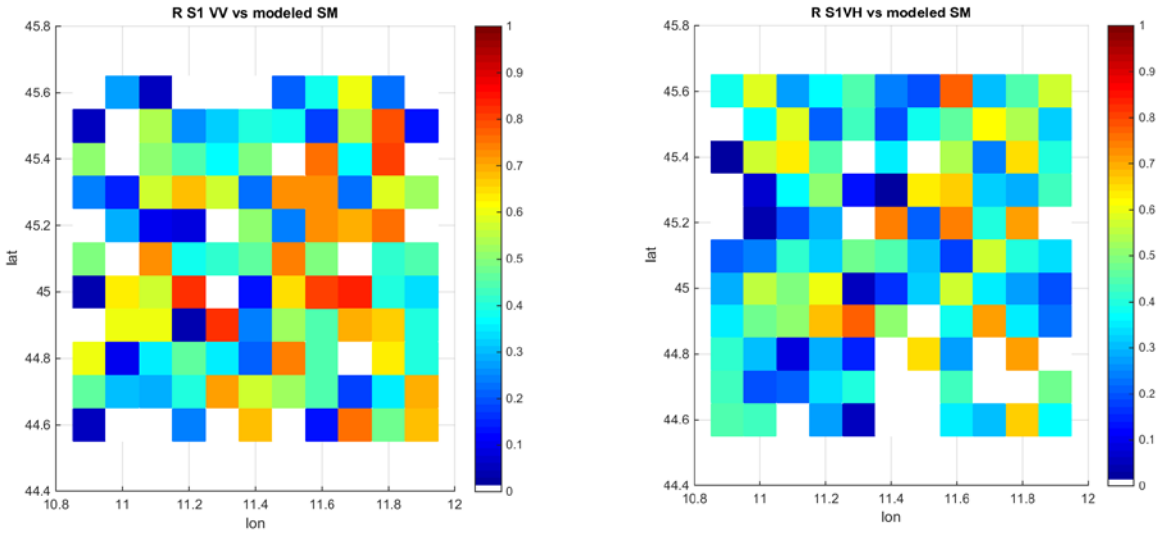


Figure 4. R maps: SMAP V and H in ascending and descending orbits and S-1 VV and VH vs. SWBM SMC

The R values reported in both Figure 4 and Table 1 do confirm the scarce sensitivity of S-1 C band data to SMC on the study area, unless no corrections for vegetation effects are performed.

Table 1. Correlation between SMAP and S-1 acquisitions and reference SMC

	SMAP TbV (original resolution)	*SMAP TbH (original resolution)*	*SMAP TbV (SFIM disaggregated at 0.1° x 0.1°)*	*SMAP TbH (SFIM disaggregated at 0.1° x 0.1°)*	*S-1 σ^0VV*	*S-1 σ^0VH*
R	***0.76***	***0.79***	***0.78***	***0.80***	***0.35***	***0.38***

4 ANN ALGORITHM

The SMC retrieval algorithm was based on Artificial Neural Networks (ANN). We considered the feed-forward multi-layer perceptron (MLP) ANN available in the Matlab ® toolbox, that use the back propagation (BP) learning rule for training. The proposed algorithm considers Tb and σ° together with ancillary PWC and LST data as inputs and SMC as output.

Two configurations, namely ANN1 and ANN2, have been implemented, depending on the availability of ancillary information on PWC and LST. The final architecture ANN, in terms of number of neurons and hidden layers, has been defined through an iterative process that repeated the ANN training and test by increasing the number of neurons and hidden layers at each iteration and compared

training and test errors (Santi et al. 2016). This process is needed to avoid underfitting (too simple architectures are not able to reproduce the problem with enough accuracy) and overfitting (too complex architectures return low training error but fail the test on other datasets). The outcome of this process was an "optimal" configuration with two hidden layers of 10 neurons each for both ANN. The architecture of the two ANNs is displayed in Figure 5 a) and b).

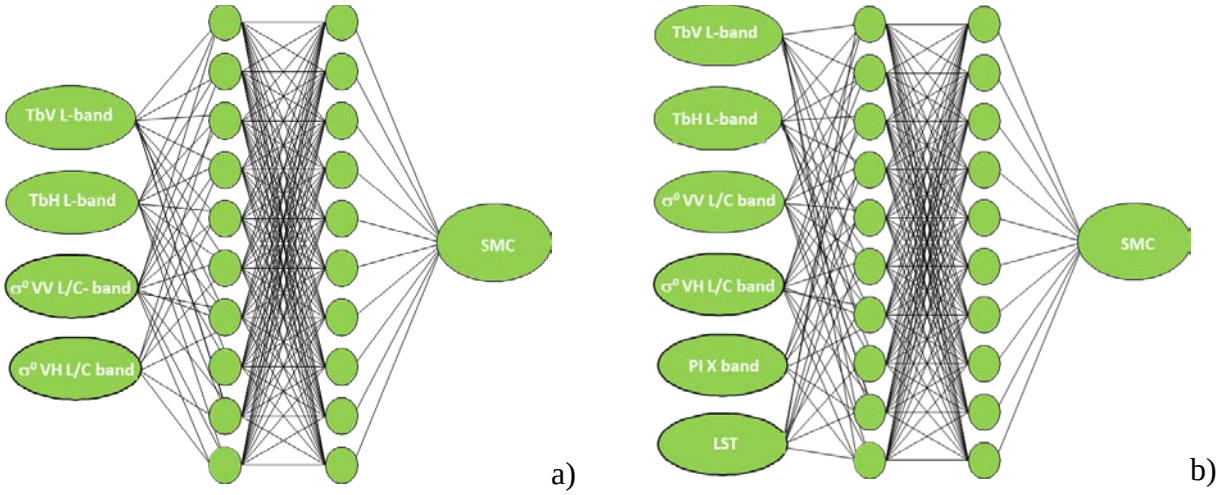


Figure 5. ANN architecture in the configuration without, ANN1 (a), and with, ANN2 (b), the ancillary inputs of LST and PWC.

4.1 ANN training

In order to extend the validity of the training and to make the ANN able to reproduce a wider range of observed surface conditions, data simulated by forward electromagnetic models based on the radiative transfer theory have been added to the SMEX02 experimental dataset. In detail, Tb at L and X bands have been simulated by the so called "tau-omega" model (Mo et al. 1982), while the Vegetation Water Cloud (VWC) model (Attema and Ulaby 1978) was considered for simulating the σ°values at L and C bands. The contribution of soil under vegetation was accounted for by coupling VWC with the OH semi-empirical model (Oh et al. 1992).

Model simulations have been iterated 30.000 times, with inputs, namely SMC, LST, PWC, the optical depth (tau), the single scattering albedo (omega), the height standard deviation (HSTD), and the correlation length (Lc), randomly varying in the range of experimental surface parameters, derived from the SMEX02 database. The data for training and independent test have been defined as follows:

- TRAINING: simulated + ½ experimental (approx. 65,000 samples)
- Independent TEST: remaining ½ experimental (approx. 35,000 samples)

The range of model inputs not available from direct measurements, namely tau, omega, and HSTD, was derived from the experimental data using a Nelder Mead optimization algorithm, according to the strategy proposed in Santi et al. (2016). In particular, Nelder Mead minimization was applied to find tau and omega values that minimize the following cost function between measured Tb from PALS and PSR and tau-omega model simulations:

$$\sqrt{\left(TbV_{PALS/PSR} - TbV_{model}\right)^2 + \left(TbH_{PALS/PSR} - TbH_{model}\right)^2}$$

The mean and variance of τ and ω resulting from this minimization are listed in table:

Table 2. Mean value and variance of tau and omega from the minimization process.

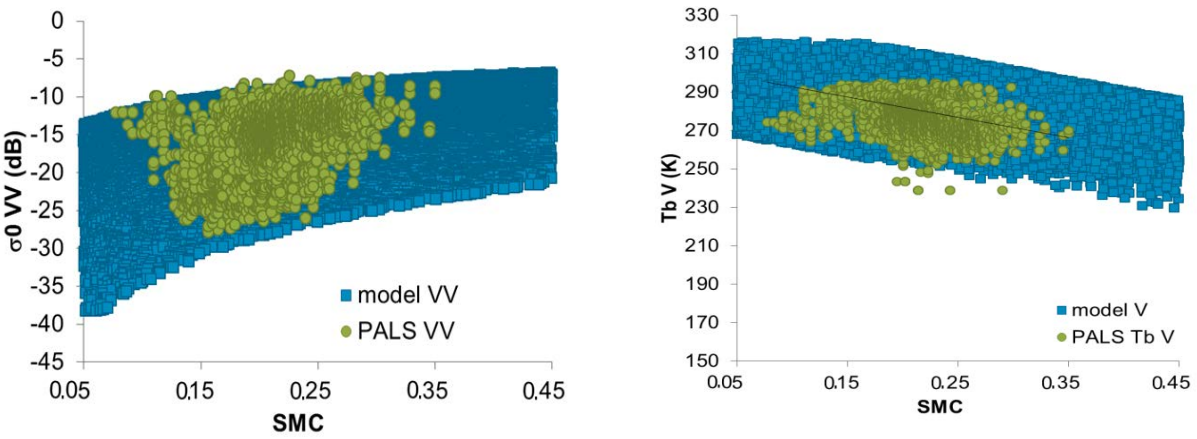
τ	L	C	X
mean	0.24	0.76	0.99
variance	0.003	0.012	0.009
ω			
mean	0.03	0.02	0.02
variance	0.002	0.000	0.000

A similar procedure allowed estimating the roughness parameter, HSTD. In this case, the Nelder Mead algorithm was applied to find HSTD values that minimize the cost function between measured σ^0 and VWC model simulations:

$$\sqrt{(\sigma^0 VV_{PALS} - \sigma^0 VV_{model})^2 + (\sigma^0 HH_{PALS} - \sigma^0 HH_{model})^2}$$

The function was minimized as a function of HSTD and of the empirical A and B parameters of WCM, which depend on the crop type. The minimization resulted in a mean HSTD of 1.55 cm, with a variance=0.36 cm.

The comparison between simulated and measured data is represented in Figure 6, where the simulated and measured σ^0 (a) and Tb (b) are represented as a function of the SMC measured on ground.



a) b)

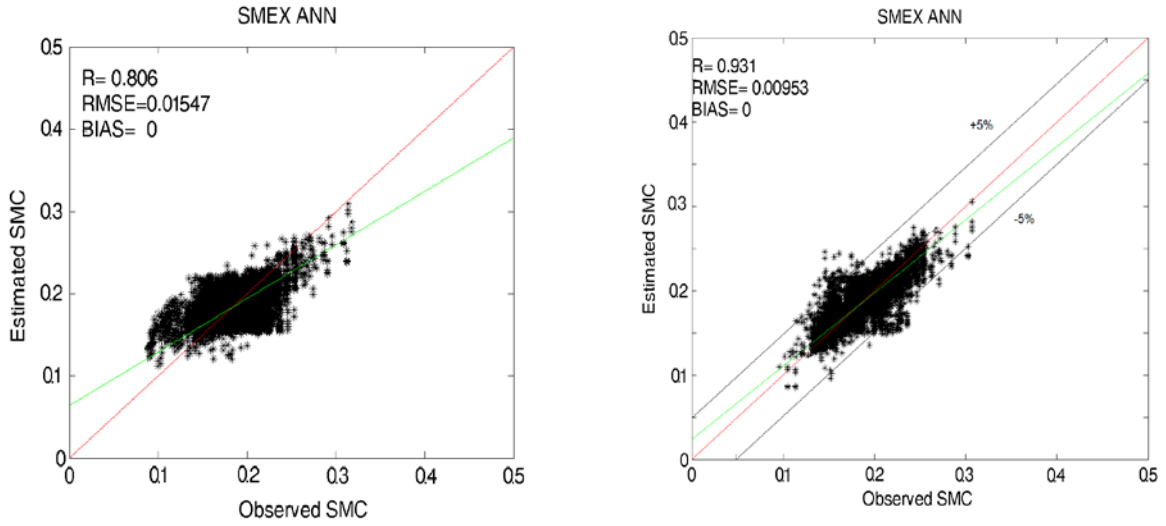
Figure 6. a) Simulated and measured σ^0 at L band, VV polarization, as a function of the measured SMC; b) simulated and measured Tb at L band, V polarization, as a function of the measured SMC.

The ANN training was carried out by splitting the training set in 60%, 20% and 20% for training itself, a posteriori test and validation. This has been done for applying the so-called "early stopping" rule, which stops the training as soon as the errors on the three subsets are diverging. This operation was needed to prevent overfitting. Training was repeated 100 times, each time by resetting the initial ANN configuration and only the ANN that provided best result was saved and considered for testing the algorithms. It has to be remarked that the tests of the algorithm presented in section 4.1 and section 4.2 does not refer to the ANN test and validation cited here, that is part of the training itself, but to the application of the already trained and saved ANN to the remaining 50% of experimental data, not considered for the training.

4.2 ANN test with airborne data (SMEX02)

After training, the algorithm was tested on SMEX02 data by applying the already trained ANN1 and ANN2 to the 35,000 experimental samples not involved in the training process, in order to keep this result as independent as possible of the training itself.

In Figure 7, the SMC estimated by the algorithm is represented as a function of the SMC measured on ground for the ANN1, only PALS data (a) and ANN2, PALS data + ancillary (b).



a) b)

Figure 7. a) SMC estimated by ANN1 vs. SMC measured; b) SMC estimated by ANN2 vs. SMC measured.

The results can be considered encouraging in both cases; however, from the statistics is evident the contribution of the ancillary information in enhancing the retrieval accuracy. The correlation coefficient increased indeed from R=0.81 of ANN1 to R=0.93 of ANN2, and the corresponding RMSE decreased from 0.015 (m^3/m^3) to 0.009 (m^3/m^3).

4.3 ANN application to satellite data (Po Valley)

The ANN algorithm was then adapted to satellite data and tested on the Po Valley area. With respect to the original implementation for SMEX02, σ^0 from S-1 replaced the PALS radar acquisitions in VV and VH polarization, and SMAP Tb at V and H polarization replaced and the PALS radiometric acquisitions.

PWC was derived from the Polarization Index computed from the X band channel of AMSR2 (Santi et al. 2012), and LST was obtained from the AMSR2 Tb at Ka band, in V pol (Owe et al. 2001). The replacement of the SMAP L band radar with the S-1 SAR at C band was evidently the most challenging issue, being C band more affected by the disturbing effect of vegetation than L band, which is instead more related to SMC. An L band SAR sensor as the ALOS-2 PALSAR would be more suitable for

this scope; however, the availability of ALOS-2 images over Italy is scarce, thus introducing a further limitation for validating the algorithm. On the other hand, Das et al. (2016) have pointed out the possibility of using Sl-1 for replacing the missing SMAP radar acquisitions.

These changes imposed a retraining of the two ANN (ANN1 and ANN2), in order to account for the differences between satellite and airborne acquisitions, keeping nevertheless unchanged the algorithm architecture with respect to the SMEX02 implementation. Considering the limited amount of data available, the definition of the training set was relatively challenging, since we had to comply with the opposite needs of having data enough to train the ANN and as many as possible to have a statistically significant test of the algorithm. As for SMEX02, data simulated by tau-omega and WCM models contributed to increase the training set: 10% of the available dataset (about 500 points) was combined with an equal number of simulated data for training, while the test was carried out on the remaining 90% of experimental data, not involved in the training process.

The obtained results are represented in the scatterplots of Figure 8, separated by ascending and descending orbits. The plots represent the estimated SMC vs. the SWBM distributed SMC. Along with ANN1 (3^{rd} column) and ANN2 (4^{th} column), the 2^{nd} column shows the results obtained by another ANN implementation that considers SMAP data only and the 1^{st} column shows official L2 enhanced SMAP SMC 9 Km product, included as a term of comparison for evaluating the effectiveness of the proposed technique. Although the limited amount of data available and the small extension of the test area do not allow drawing too general conclusions, the obtained results appear promising. The comparison between scatterplots in the 1^{st} and 2^{nd} columns shows that the ANN trained with SMAP data only provides slightly better results than the official product in the area. When including in the comparison the results shown in 3^{rd} and 4^{th} columns, is also evident that the inclusion of S-1 data (ANN1), although scarcely related to SMC on their own, allowed a certain improvement of the retrieval accuracy with respect to the use of SMAP data alone. The obtained R and RMSE are summarized in Table 3.

Table 3. R and RMSE obtained in ascending and descending orbits for the considered satellite combinations

		SMAP Official	SMAP ANN	SMAP+S-1	SMAP+AMSR2+S-1
Asc.	R	0.69	0.79	0.85	0.87
	RMSE (m^3/m^3)	0.039	0.029	0.025	0.023
Desc.	R	0.68	0.78	0.82	0.84
	RMSE (m^3/m^3)	0.04	0.029	0.026	0.025

ANN1 probably represents the most interesting implementation, since it has the same inputs of the original SMAP radar and radiometer algorithm. Finally, the inclusion of ancillary information on PWC and LST from AMSR2 (ANN2 - right column), allowed a further accuracy improvement, although smaller than the one obtained for SMEX02. This is mainly due to the homogeneity of the area, which makes the effect of ancillary information from AMSR2 less relevant. For inhomogeneous landscapes, characterized by the presence of different vegetation covers, the ancillary information on PWC and LST may instead have a significant effect on the algorithm performances. This is confirmed when considering the upper left part of the area, which is characterized by mixed agricultural and urban areas. The maps of R and RMSE, computed at each pixel of the grid and displayed in Figure 9 and Figure 10, show indeed that the synergy of SMAP with S-1 and AMSR2 does improve appreciably the retrieval accuracy in this sub-area, with respect of both (ANN and official) SMC products based on SMAP data only.

The main limitation for extending the applicability of the proposed technique to global scale relies on the availability of S-1 data and the satellite revisiting time, which is actually between 6 and 12 days and depends on the planning updates. The synergy between SMAP and AMSR2 is instead less problematic, being both satellites operating in sun synchronous orbits with a frequent coverage of the entire earth surface.

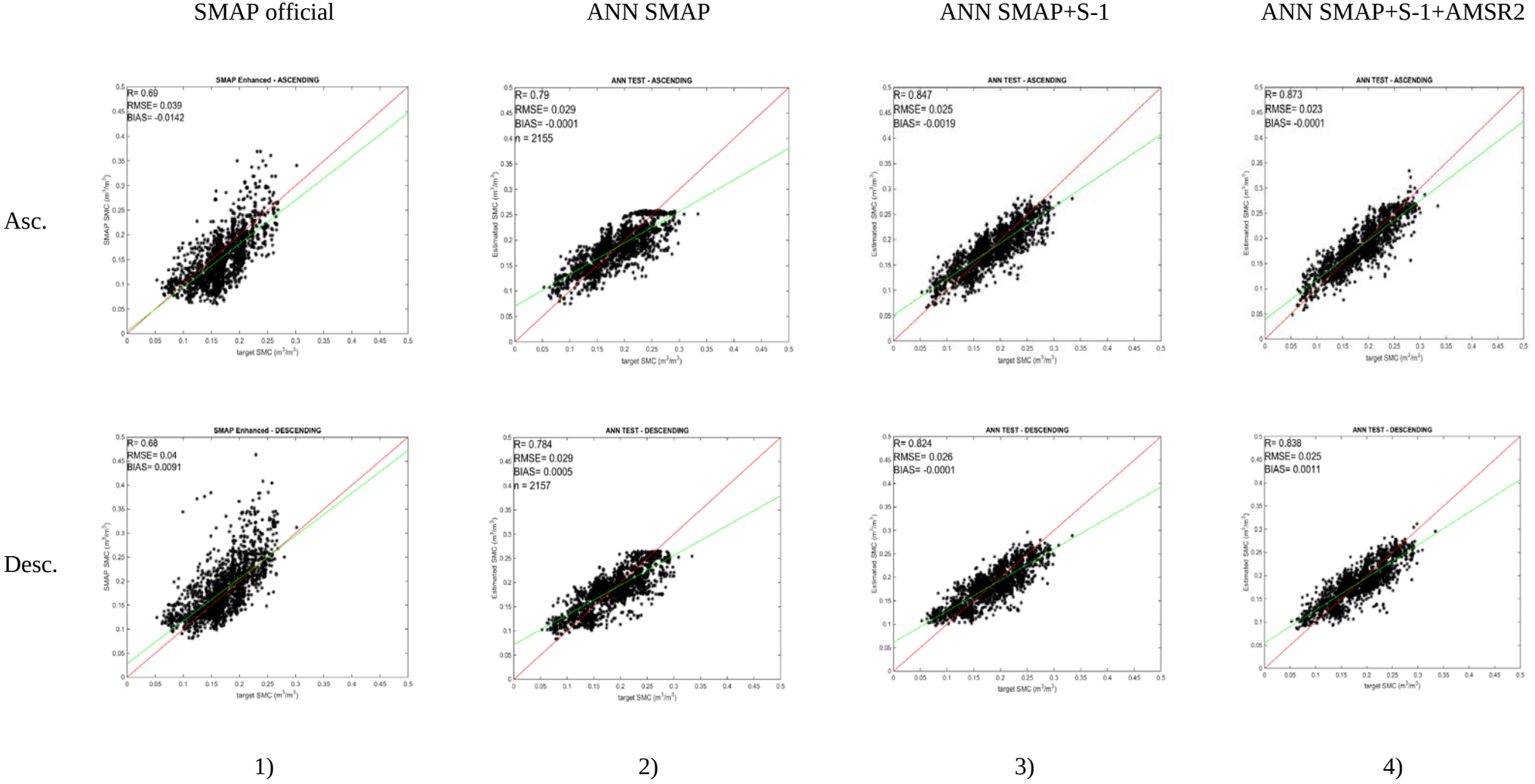
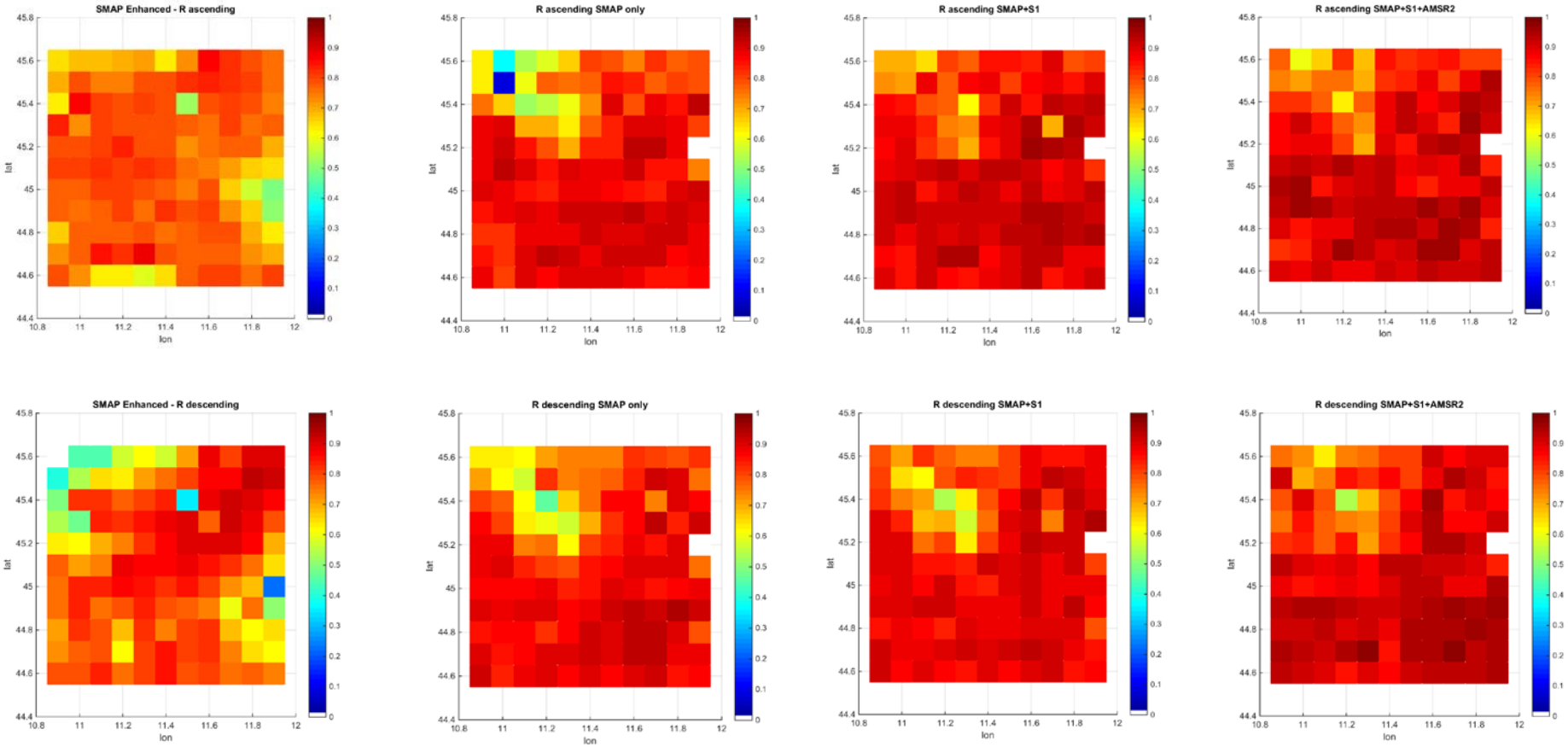


Figure 8. ANN algorithm test using satellite data in ascending (upper line) and descending (lower line) orbits, using disaggregated SMAP data alone 2), disaggregated SMAP + S1 (3 – ANN1 configuration), and adding also PI_X and TbV_{Ka} from AMSR2 (4 – ANN2 configuration). As term of comparison, the 1st column shows the results obtained using the official L2 enhanced SMAP 9Km SMC product.

SMAP official | ANN SMAP | ANN SMAP+S-1 | ANN SMAP+S-1+AMSR2

Asc.

Desc.



1) 2) 3) 4)

Figure 9. Maps of R in ascending (upper line) and descending (lower line) orbits, using disaggregated SMAP data alone 2), disaggregated SMAP + S1 (3 – ANN1 configuration), and SMAP + S-1 + AMSR2 (4 – ANN2 configuration). As term of comparison, the 1st column shows the results obtained using the official L2 enhanced SMAP 9Km SMC product.

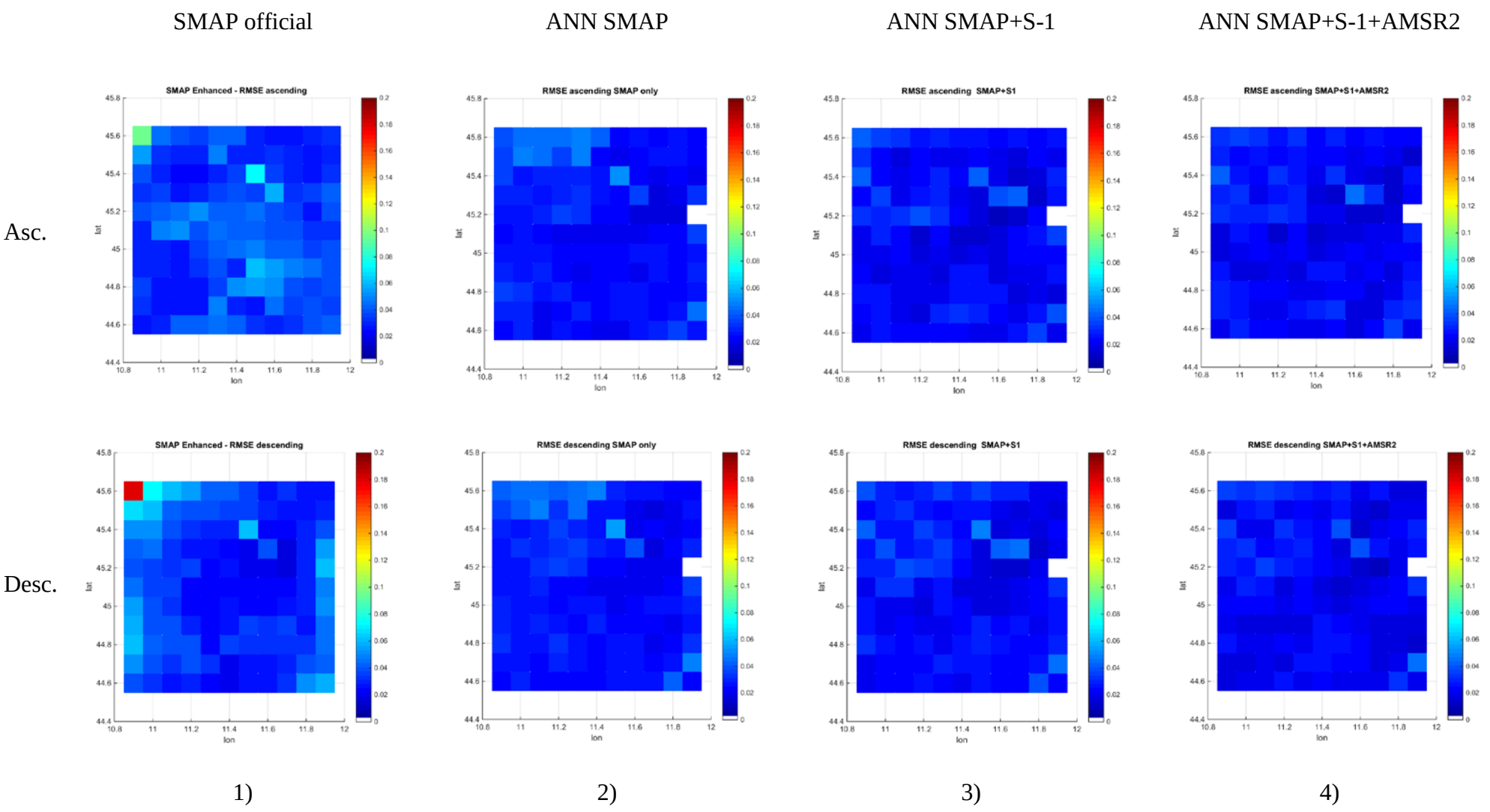


Figure 10. Maps of RMSE in ascending (upper line) and descending (lower line) orbits, using disaggregated SMAP data alone 2), disaggregated SMAP + S1 (3 – ANN1 configuration), and SMAP + S-1 + AMSR2 (4 – ANN2 configuration). As term of comparison, the 1st column shows the results obtained using the official L2 enhanced SMAP 9Km SMC product.

5 CONCLUSIONS

An algorithm based on Artificial Neural Networks (ANN) for retrieving SMC from synergic active and passive microwave acquisitions has been discussed. The algorithm was aimed at generating a SMC product at a resolution of 0.1°x0.1° (approx. 10Km x 10Km). PALS and PSR acquisitions from SMEX02 and forward EM models simulations were considered for defining and training the algorithm, which was implemented in two configurations, ANN1 and ANN2, the latter accounting for ancillary information on LST and PWC. The algorithm was then adapted to satellite data and tested considering one year of overlapped SMAP, S-1 and AMSR2 acquisitions collected on an agricultural area in northern Italy. The missing SMAP radar acquisitions have been replaced by S-1 data, while the ancillary information on PWC and LST has been derived from AMSR2. SMAP and AMSR2 data were disaggregated at a resolution of 0.1°x0.1° by using the SFIM disaggregation and S-1 acquisitions were downsampled to the same resolution. The "official" L2 SMAP enhanced SMC 9km product was considered in the same area and for the same dates as a term of comparison for evaluating the obtained results.

The obtained results were encouraging, and the synergy with S-1 and AMSR2 allowed increasing R from about 0.68 of the official SMAP enhanced product, to about 0.84 for ANN1 and 0.86 for ANN2 (SMAP+S-1+AMSR2), with a corresponding decrease of RMSE from about 0.040 m^3/m^3 to about 0.025 m^3/m^3 (ANN1) and 0.024 m^3/m^3 (ANN2). Although ANN1 reached slightly less accurate results than ANN2, it is probably the most interesting implementation, since it is the closest in terms of inputs to the original SMAP radar and radiometer algorithm. The synergy effect was particularly evident in a sub area in which the agricultural fields are mixed with urban areas, thus limiting the performance of the retrieval based on SMAP data only.

It should be remarked that the test area is rather small and the available dataset is not sufficient for drawing global considerations; however, the proposed technique appeared promising for obtaining a combined radar/radiometer product at improved spatial resolution and enhanced accuracy with respect

to the SMAP SMC product. The main limitation for generalizing this technique to a global application is represented by the limited revisiting of S-1.

6 REFERENCES

Attema, E. P. W. & Ulaby, F. T. 1978. Vegetation modeled as a water cloud. Radio Science, 13, 357-364.

Bindlish, R. and Thomas Jackson. 2002. SMEX02 Aircraft Polarimetric Scanning Radiometer (PSR) Data. Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.

Brocca L., F. Melone, T. Moramarco, (2008), "On the estimation of antecedent wetness conditions in rainfall-runoff modelling", Hydrological Processes, 22 (5), 629-642, doi:10.1002/hyp.6629.

Brocca L., S. Camici, F. Melone, T. Moramarco, J. Martinez-Fernandez, J.-F. Didon-Lescot, R. Morbidelli (2014), "Improving the representation of soil moisture by using a semi-analytical infiltration model", Hydrological Processes, 28(4), pp. 2103-2115, doi:10.1002/hyp.9766.

Brocca L., S. Hasenauer, T. Lacava, F. Melone, T. Moramarco, W. Wagner, W. Dorigo, P. Matgen, J. Martínez-Fernández, P. Llorens, J. Latron, C. Martin, M. Bittelli (2011) "Soil moisture estimation through ASCAT and AMSR-E sensors: an intercomparison and validation study across Europe," Remote Sensing of Environment, 115, 3390-3408, doi:10.1016/j.rse.2011.08.003.

Das N. N., D. Entekhabi, E.G. Njoku, J. J. C. Shi, J. T. Johnson, A. Colliander, 2014. Tests of the SMAP Combined Radar and Radiometer Algorithm Using Airborne Field Campaign Observations and Simulated Data, IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, 52, 4, pp.2018-2028.

Das N. N., D. Entekhabi, S. Kim, S. Yueh, and P. O'Neill, 2016. Combining SMAP and Sentinel Data for High-Resolution Soil Moisture Product. Proc. Of International Geoscience and remote sensing symposium, IGARSS 2016, July 2016, pp. 129-131.

Entekhabi D., E. G. Njoku, P. E. O'Neill, K. H. Kellogg, W. T. Crow, W. N. Edelstein, J. K. Entin, S. D. Goodman, T. J. Jackson, J. Johnson, J. Kimball, J. R. Piepmeier, R. D. Koster, N. Martin, K. C. McDonald, M. Moghaddam, S. Moran, R. Reichle, J. C. Shi, M. W. Spencer, S. W. Thurman, L. Tsang, J. Van Zyl. "The Soil Moisture Active Passive (SMAP) Mission," in Proceedings of the IEEE, vol. 98, no. 5, pp. 704-716, May 2010. doi: 10.1109/JPROC.2010.2043918

Entekhabi D., Njoku, E., Houser, P., Spencer, M., Doiron, T., Belair, S., Crow, W., Jackson, T. J., Kerr, Y., Kimball, J., Koster, R., McDonald, K., O'Neill, P., Pultz, T., Running, S., Shi, J. C., Wood, E. and van Zyl, J., 2004. The Hydrosphere State (HYDROS) Satellite Mission: an Earth System Pathfinder for Global Mapping of Soil Moisture and Land Freeze/Thaw. IEEE Trans. Geosci. Rem. Sens., 42(10): 2184-2195.N. N. Das, D. Entekhabi and E. G. Njoku, "An Algorithm for Merging SMAP Radiometer and Radar Data for High-Resolution Soil-Moisture Retrieval. IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 5, pp. 1504-1512, May 2011. doi: 10.1109/TGRS.2010.2089526

Gumuzzio, Á., Brocca, L., Sanchez, N., González-Zamora, Á., Martínez-Fernández, J., 2016. Comparison of SMOS, modelled and in situ long-term soil moisture series in the NW of Spain. Hydrological Science Journal, vol. 61, no. 14, pp. 2610-2625, July 2016. doi: 10.1080/02626667.2016.1151981

Jackson, T. and M. Cosh. 2003. SMEX02 Watershed Soil Moisture Data, Walnut Creek, Iowa. Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.

Lacava T., P. Matgen, L. Brocca, M. Bittelli, T. Moramarco (2012) “A first assessment of the SMOS soil moisture product with in-situ and modelled data in Italy and Luxembourg,” IEEE Transaction on Geoscience and Remote Sensing, 50(5), 1612-1622, doi:10.1109/TGRS.2012.2186819.

Liu, J. G., 2000. Smoothing Filter Based Intensity Modulation: a Spectral Preserve Image Fusion Technique for Improving Spatial Details. International Journal of Remote Sensing, 21, No. 18, 3461-3472.

Mecklenburg, S., Drusch, M., Kerr, Y. H., Font, J., Martin-Neira, M., Delwart, S., ... & Crapolicchio, R. (2012). ESA's soil moisture and ocean salinity mission: Mission performance and operations. IEEE Transactions on Geoscience and Remote Sensing, 50(5), 1354-1366.

Mo T., B. J. Choudhury, T. J. Schmugge, J. R. Wang, and T. J. Jackson (1982), “A model for microwave emission from vegetation covered fields,” J. Geophys. Res., vol. 87, pp. 11229–11 237.

Njoku, E. G. 2003. SMEX02 Passive and Active L and S band System (PALS) Data. Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.

Oh, Y., Sarabandi, K., Ulaby, F. T. 1992. An empirical model and an inversion technique for radar scattering from bare surfaces, IEEE Transactions on Geoscience and Remote Sensing, 30, 370-381.

O'Neill, P. E., S. Chan, E. G. Njoku, T. Jackson, and R. Bindlish. 2016. SMAP Enhanced L3 Radiometer Global Daily 9 km EASE-Grid Soil Moisture, Version 1. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: http://dx.doi.org/10.5067/ZRO7EXJ8O3XI.

Owe M. and A. A. Van De Griend, “On the relationship between thermodynamic surface temperature and high-frequency (37 GHz) vertically polarized brightness temperature under semi-arid conditions,” Int. J. Remote Sens., vol. 22, no. 17, pp. 3521–3532, Nov. 20, 2001.

Paloscia S., P. Pampaloni, 1988, "Microwave Polarization Index for Monitoring Vegetation Growth", IEEE Trans. on Geoscience and Remote Sensing, GE-26, n. 5, sept. 1988, pp. 617-621

Parinussa R. M., M. T. Yilmaz, M. C. Anderson, C. R. Hain and R. A. M. de Jeu. 2013. An intercomparison of remotely sensed soil moisture products at various spatial scales over the Iberian Peninsula. Hydrological Processes, 28(18), 4865–4876. DOI: 10.1002/hyp.9975.

Santi E, S Pettinato, S Paloscia, P Pampaloni, G Fontanelli, A Crepaz, M Valt, 2014. Monitoring of Alpine snow using satellite radiometers and artificial neural networks. Remote Sensing of Environment, 144, 179-186. http://dx.doi.org/10.1016/j.rse.2014.01.012

Santi E. 2010. An application of SFIM technique to enhance the spatial resolution of microwave radiometers. International Journal of Remote Sensing, 31(9-10), 2419-2428.

Santi E., S. Pettinato, S. Paloscia, P. Pampaloni, G. Macelloni, and M. Brogioni, 2012, “An algorithm for generating soil moisture and snow depth maps from microwave spaceborne radiometers: HydroAlgo”, Hydrol. Earth Syst. Sci., 16, pp. 3659-3676, doi:10.5194/hess-16-3659-2012.

Santi, E., S. Paloscia, S. Pettinato, G. Fontanelli. Application of artificial neural networks for the soil moisture retrieval from active and passive microwave spaceborne sensors. Int. J. Appl. Earth Observ. Geoinf. 2016, Volume 48, Pages 61–73. http://dx.doi.org/10.1016/j.jag.2015.08.002.

Silvestrin, P., Berger, M., Kerr, Y., & Font, J. (2001). ESA’s second earth explorer opportunity mission: The soil moisture and ocean salinity mission-SMOS. IEEE Geoscience and Remote Sensing Newsletter, 118, 11-14.