

Paper Received 24 March 2015, received in revised form 29 July 2015, accepted 6 August 2015, Available online 6 September 2015.
Published on International Journal of Applied Earth Observation and
Geoinformation 48 (2016) pp. 61–73.
<http://dx.doi.org/10.1016/j.jag.2015.08.002>

Application of Artificial Neural Networks for the Soil Moisture Retrieval from Active and Passive Microwave Spaceborne sensors

Emanuele Santi, Simonetta Paloscia, Simone Pettinato, Giacomo Fontanelli

Institute of Applied Physics –National Research Council (IFAC-CNR)

via Madonna del Piano, 10 - Firenze (Italy)

Contact Author email e.santi@ifac.cnr.it; Tel. +39 055 5226431; Fax +39 055 5226434.

Abstract: Among the algorithms used for the retrieval of SMC from microwave sensors (both active, such as Synthetic Aperture Radar-SAR, and passive, radiometers), the artificial neural networks (ANN) represent the best compromise between accuracy and computation speed. ANN based algorithms have been developed at IFAC, and adapted to several radar and radiometric satellite sensors, in order to generate SMC products at a resolution varying from hundreds of meters to tens of kilometers according to the spatial scale of each sensor. These algorithms, which are based on the ANN techniques for inverting theoretical and semi-empirical models, have been adapted to the C- to Ka- band acquisitions from spaceborne radiometers (AMSR-E/AMSR2), SAR (Envisat/ASAR, Cosmo-SkyMed) and real aperture radar (MetOP ASCAT).

Large datasets of co-located satellite acquisitions and direct SMC measurements on several test sites worldwide have been used along with simulations derived from forward electromagnetic models for setting up, training and validating these algorithms. An overall quality assessment of the obtained results in terms of accuracy and computational cost was carried out, and the main advantages and limitations for an operational use of these algorithms were evaluated.

This technique allowed the retrieval of SMC from both active and passive satellite systems, with accuracy values of about $0.05 \text{ m}^3/\text{m}^3$ of SMC or better, thus making these applications compliant with the usual accuracy requirements for SMC products from space.

Keywords: *Artificial Neural Networks, Soil Moisture Content, Synthetic Aperture Radar (SAR), Scatterometer, Microwave Radiometers*

1. Introduction

The amount of water stored in the soil is an essential variable controlling many biophysical processes that impact water, energy and carbon exchanges at the land-atmosphere interface. In-situ soil moisture (SMC) measurements are labor intensive and site-specific, moreover, frequent and spatially distributed soil moisture measurements, at different spatial scales, are advisable for the most part of the applications related to the environmental disciplines, such as climatology, meteorology, hydrology and agriculture. The possibility of observing soil moisture and its temporal evolution from space is, therefore, regarded as being extremely attractive.

Among the instruments operating from space for the observation of the Earth surface, the sensors operating in the microwave portion of the spectrum have received most attention because this frequency range has the unique ability to return information on media (atmosphere, vegetation, soil) that are opaque to the much shorter visible/near-infrared and thermal wavelength and, that is most important, because microwave scattering and emission are directly related to the water content of the observed target. In particular, remote sensing from active (Synthetic Aperture Radar – SAR and scatterometer) and passive sensors (radiometers) have demonstrated to be good and flexible tools for observing the moisture content of the first centimeter layer of soil and for detecting its spatial and temporal variations from radar (e.g. Barret et al. 2009; Notarnicola et al.

2006; Paloscia et al. 2004; Pathe et al. 2009; Pierdicca et al. 2010; Wagner et al. 1999; Wagner et al. 2007) and radiometric microwave sensors (e.g. Jackson, 1993; Jackson et al. 2010; Mladenova et al. 2014; Paloscia et al. 2006).

These two types of sensors (radiometers and SAR) retrieve information of the Earth surface at different scales. SAR, in fact, due to antenna synthesis, reaches a very high ground resolution, in the order of a few meters; whereas radiometers from space have a much coarser spatial resolution, in the order of tens of kilometers. These characteristics allowed the use of these sensors for different applications. Fine ground resolutions allowed in fact a more detailed investigation of the surface and can give information on small scale phenomena, useful for agricultural applications and water management at a farm or basin scale. The large scale of spaceborne radiometers is instead very useful for climatic applications and for detecting the trend of large scale phenomena, as global changes, e.g. deforestation and desertification processes.

The retrieval of soil parameters from active and/or passive microwave measurements is nonetheless not trivial, due to the non-linearity of the relationships between radar or radiometric acquisitions and ground parameters, and because, in general, more than one combination of soil parameters (soil moisture, roughness, vegetation cover, etc.) has the same electromagnetic response. Thus, in order to minimize the uncertainties and enhance the performance of soil parameter retrieval from remote sensing data, statistical approaches based on the Bayes theorem and learning machines are widely adopted for implementing the retrieval algorithms (e.g. Notarnicola et al. 2014; Pasolli et al. 2011; Pierdicca et al. 2014).

In this framework, the Artificial Neural Networks (ANNs) represent an interesting tool for implementing accurate and flexible retrieval algorithms, which able to operate with radar and radiometric satellite measurements and to easily combine information coming from different sources. ANN can be considered a statistical minimum variance approach for addressing the retrieval problem and they can be trained to represent arbitrary input-output relationships (Hornik, 1989; Linden and Kinderman, 1989). In the training phase, training patterns are sequentially presented to the network and the interconnecting weights of each neuron are adjusted according to a learning algorithm. The trained ANN can be considered as a type of non-linear least mean square interpolation formula for the discrete set of data points in the training set.

The effectiveness of ANNs in solving remote sensing problems has been well demonstrated, since they can easily merge data coming from different sources into a unique retrieval algorithm. ANNs have been successfully applied to many inverse problems in the remote sensing field, and in particular to retrieve soil moisture at local scale from SAR (e.g. Del Frate et al. 2003; Elshorbagy and Parasuraman 2008; Paloscia et al. 2008) or radiometric (e.g. Santi et al. 2012) observations. The comparison of retrieval algorithms carried out in Paloscia et al. 2008 demonstrated that ANNs, with respect to other widely adopted statistical approaches based on Bayes theorem and Nelder-Mead minimization, offer the best compromise between retrieval accuracy and computational cost.

ANNs in some cases have been used essentially as a black box, without further effort for understanding the underlying processes and the physics behind them. The strategy for minimizing these problems is mainly based on the use of both extensive datasets and model simulations for the training phase of ANN. The studies presented in Paloscia et al. 2010 and Santi et al. 2013, pointed out the potential of the ANN technique in easily and effectively ingesting information extracted from different sources for improving the retrieval process, such as NDVI derived from optical remote sensing imagery for taking into account the presence of vegetation on the ground. In Paloscia et al. 2013 nevertheless, the importance of a robust and extensive reference dataset for the training was confirmed, in order to obtain a retrieval algorithm able to work at large and global scale with a satisfactory accuracy.

An application of retrieval algorithms to MetOp-A ASCAT backscatter data can be found in Gruber et al. 2014, where the ANN techniques are compared with other retrieval algorithms, in their ability to retrieve soil

moisture on a global scale. Correlation and triple collocation analysis were performed using in situ and land surface model data as a reference, pointing out again the effectiveness of the ANNs compared with other inversion approaches.

In this paper, an overview of the results obtained with ANN algorithms applied to different microwave sensors for the retrieval of soil moisture at both local and global scales is presented. We have summarized and homogenized here the results presented in other papers published in both international journals and conference proceedings, by improving and tuning the algorithms used in past research, and validating them with new datasets. In some cases, more accurate results have been obtained, with improved accuracy of the estimated parameter. In this work, a unique procedure for all the algorithms was thus implemented, by choosing the same strategy for training, test, and validation.

In section 2 a detailed discussion of the methods and in particular of the training of ANN is carried out. In the following sections, the application of the ANN to microwave radiometers, scatterometers and SAR sensors is presented along with the main findings of the methods. Finally, some soil moisture maps obtained at different spatial scales by using the data from these sensors are shown for an overall validation of the implemented algorithms.

2. Implementing and training the Artificial Neural Networks

The algorithms presented in this work are based on feed-forward multilayer perceptron (MLP) ANNs, with a certain number of hidden layers of neurons between the input and the output. In MLPs, successive layers of neurons are fully interconnected, with trainable connection weights that control the strength of the connections. The ANN training was based on the back-propagation learning rule, which is an iterative gradient descent algorithm that is designed to minimize the mean square error between the desired target vectors and the actual output vectors. It should be noted that the gradient-descent method sometimes suffers from slow convergence, due to the presence of one or more local minima, which may also affect the final result of the training. This problem can be solved by repeating the training several times, with a resetting of the initial conditions and a verification that each training process led to the same convergence results in terms of R^2 and RMSE, by increasing it until negligible improvements were obtained.

In order to define the optimal ANN architecture in terms of number of neurons and hidden layers, the most suitable strategy is to start with a simple ANN architecture, generally with one hidden layer of few neurons. This ANN is trained by means of a subset of the available data, tested on the rest of the dataset and the training and testing errors are compared. The ANN configuration is then increased by adding neurons and hidden layers, training and testing are repeated and errors compared again, until a further increase of the ANN architecture is found to have a negligible decrease of the training error and an increase in the test error. This procedure allows defining the minimal ANN architecture capable of providing an adequate fit for the training data, so as to prevent overfitting problems. Overfitting is related to the oversizing of the ANN, and may cause considerable errors when testing ANN with input data that is not included in the training set (Moody et al. 1992; Tetko et al. 1995). In other words, the ANN is able to reproduce the training set with high accuracy but fails the test procedure.

Another key issue for defining the ANN best architecture is in the selection of the most appropriate transfer function: in general linear transfer functions give less accurate results in training and testing; however, they are less prone to overfitting and are more robust to outliers, i.e. input data out the range of the input parameters included in the training set. Logistic Sigmoid (logsig) and Tangential Sigmoid (tansig) transfer functions are instead characterized by higher accuracies in the training and test; however, they may lead to large errors when the trained ANN is applied to new datasets. Logsig generates outputs between 0 and 1 as the neuron's net input

131 goes from negative to positive infinity and describes the non-linearity, $g(a)$, as $1/(1+e^{-a})$. Alternatively,
132 multilayer networks can use the tansig function, $\tanh(a) = (e^a - e^{-a}) / (e^a + e^{-a})$.

133 Besides these problem, however, the main constraint for obtaining good accuracies with the ANN approach,
134 as it has been demonstrated in Paloscia et al. 2013, is represented by the statistical significance of the training
135 set, which shall be representative of a variety of surface conditions as wide as possible, in order to make the
136 algorithm able to operate at a global scale. Several efforts have been carried out in the last years to create large
137 databases of simultaneous ground SMC measurements and satellite acquisitions (see as example the Climate
138 Change Initiative - CCI - soil moisture activities promoted by the European Space Agency <http://www.esa-cci.org/>
139 or the SMEX experiments in the USA). However, the datasets derived from these experimental
140 activities are not large enough for generating the “robust” set required to train the ANNs for global monitoring
141 applications.

142 A suitable strategy for filling in any gap in the experimental datasets and to better characterize the microwave
143 signal dependence on SMC for a variety of surface conditions as wide as possible is to combine the
144 experimental measurements with simulated data obtained from electromagnetic models (e.m.). The
145 consistency between experimental data and model simulations can be obtained by deriving the range of input
146 parameters (namely soil moisture, surface roughness and vegetation biomass) from the experimental
147 measurements. This task can be achieved by applying the Nelder-Mead optimization algorithm (Nelder and
148 Mead, 1965), to find the values that minimize some appropriate cost functions between microwave
149 measurements and corresponding model simulations. After defining the minimum and maximum of each input
150 parameter, the input vectors for the e.m model are generated by using a pseudorandom function, rescaled in
151 order to cover the range of each parameter. Thousands of inputs vectors for running the model simulations can
152 be generated by iterating this procedure, thus obtaining datasets of surface parameters and corresponding
153 simulated microwave data for training and testing the ANN. The flowchart of Figure 1 represents the main
154 steps for generating the training from the experimental data. The same procedure allows generating the
155 independent dataset for validating the ANN after training. In general, the available data are divided in two
156 subsets with a random sampling, the first subset is divided again in 60% - 20% - 20% for training, test and
157 validation phases, respectively and the second subset is reserved for an independent test of the algorithm. The
158 random sampling of the dataset is reiterated 5-6 times and the training is repeated each time, in order to avoid
159 any dependence of the obtained results on the sampling process.



160

161 **Figure 1.** Definition of training and validation datasets from the experimental data using the forward e.m.
162 models.

163 Several ANN based algorithms have been developed at IFAC following this strategy, and adapted to
164 SAR/scatterometric and radiometric satellite sensors, in order to generate SMC products at both local and
165 global scales with a resolution varying from hundreds of meters to tens of kilometers.

3. ANN algorithm for spaceborne multifrequency radiometers

3.1 *HydroAlgo* algorithm for AMSR-E and AMSR2

The “HydroAlgo” algorithm (Santi et al. 2012) applies the ANNs for estimating the surface SMC from the acquisitions of the low resolution spaceborne radiometers, like the Advanced Microwave Scanning Radiometer for the Earth observing system (AMSR-E) (Lobl, 2001), which is no more operating, and its successor, AMSR2 (Imaoka et al. 2012). The main characteristic of the algorithm is the exclusive use of AMSR-E/2 data, for estimating and compensating the effects of vegetation conditions and surface roughness on the SMC retrieval, thanks to the inclusion of data acquired at the higher frequencies. Brightness temperature data collected over the areas of interest were extracted from the Hierarchical Data Format (HDF) files delivered by National Snow and Ice Data Center (NSIDC) and containing the calibrated and geocoded acquisitions of AMSR-E from AQUA satellite (Level 2 data) at C, X, Ku and Ka bands in both polarizations (H, V). A check of data for possible miscalibration (Paloscia et al. 2006) and for the presence of the Radio Frequency Interference (RFI) at C and X bands was carried out, by using a simple threshold method (Njoku et al. 2005) at both C and X bands, and all data over this threshold were eliminated from the dataset. HydroAlgo includes a disaggregation procedure, based on the SFIM filtering, (Santi, 2010; Parinussa et al. 2013), which is able to enhance the spatial resolution of the output SMC product up to the nominal sampling of AMSR-E ($\sim 10 \times 10$ km²), thus minimizing the limit of these sensors for the monitoring of heterogeneous landscapes, due to their coarse resolution. The core of the algorithm is composed by two feedforward multilayer perceptron (MLP) ANNs, trained independently for the ascending and descending orbits using the back-propagation learning rule. Inputs of the algorithm are the brightness temperatures at C band in V polarization, the polarization indices (PI) at 10.65 GHz and 18.7 GHz (X and Ku bands), defined as $PI = 2 \cdot (TbV - TbH) / (TbV + TbH)$, and the Tb at Ka band (36.5 GHz) in V polarization. The C band, i.e. the lowest AMSR-E frequency, was chosen for its sensitivity to the SMC, which is slightly influenced by sparse vegetation. PI at X and Ku bands were taken into account for compensating the effect of vegetation on soil emission (Paloscia and Pampaloni, 1988), and for flagging out the densely vegetated targets, where SMC cannot be retrieved at these frequencies. Tb at Ka band, in V polarization, was assumed as a proxy of the surface physical temperature, to account for the effect of diurnal and seasonal variations of the surface temperature on microwave emission (Njoku and Li, 1999).

HydroAlgo was developed and tested using a set of several thousand of data that was obtained by combining the experimental data collected in Mongolia (Yang et al. 2008) and the Murrumbidgee watershed, in Australia (Smith et al. 2012), with 10.000 values of Tb simulated by the simple implementation of the Radiative Transfer Theory known as “tau-omega” model (Mo et al. 1982). These experimental datasets were kindly provided by JAXA, within the framework of the JAXA ADEOS-II/AMSR-E and GCOM/AMSR2 research programs. The inputs of the model were randomly varied in a range derived from experimental data as follows: SMC (0.05-0.45 m³/m³), surface temperature (275-320 K), vegetation optical depth (τ) at C band ranging between 0.16 and 1.1, and vegetation single scattering albedo (ω) ranging between 0.03 and 0.08. At the higher frequencies, τ and ω values were derived from C band values using the relationship established in Santi et al. (2012). Model simulations were iterated for each input vector of surface parameters, obtaining a training set of 10,000 brightness temperatures at all the AMSR-E bands and polarizations. Figure 2 shows the behavior of both, experimental and simulated Tb at C and Ka bands in V. pol. as a function of volumetric SMC (m³/m³).

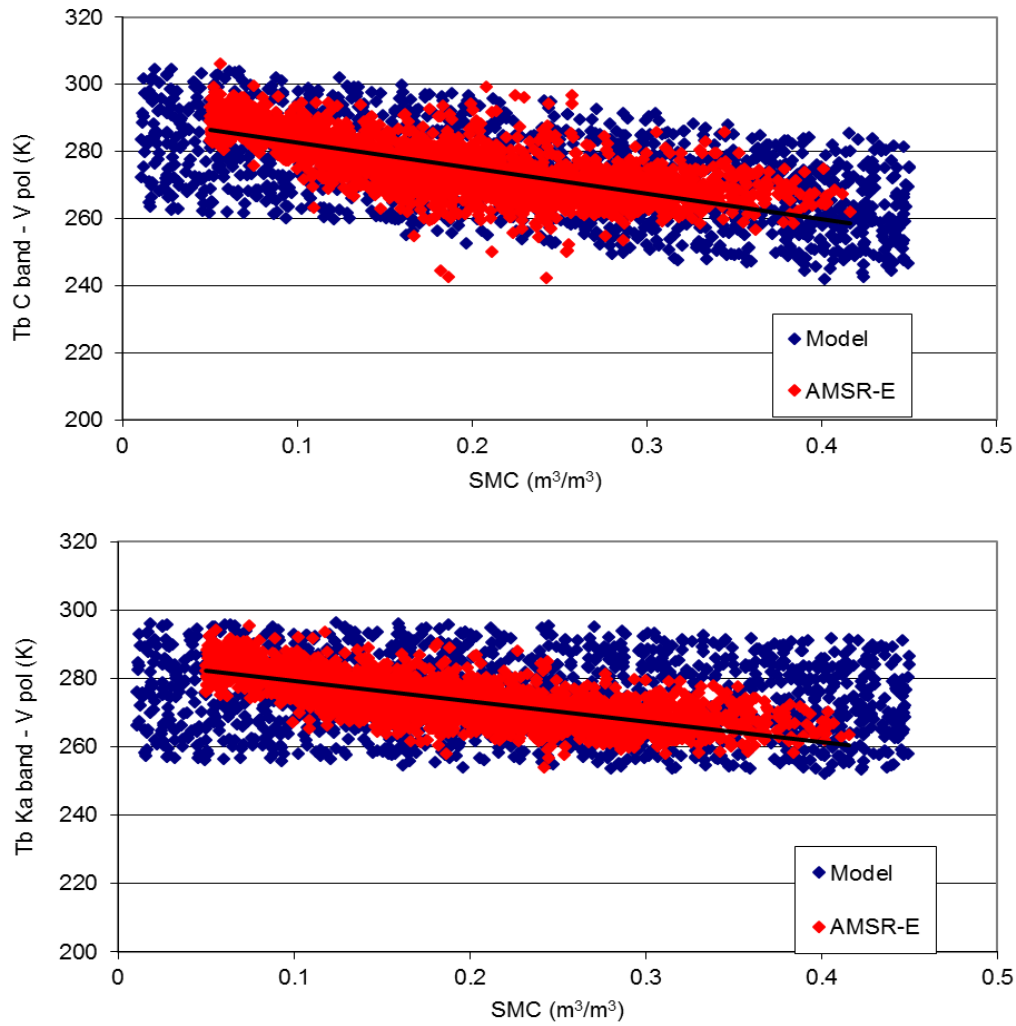


Figure 2. Experimental and simulated Tb values at C- and Ka bands, in V- polarization as a function of SMC derived from the dataset considered for training the ANN algorithm.

The flowchart of the whole processing chain is displayed in Figure 3 along with a detail on the ANN implemented for HydroAlgo.

The ANN independent test on the Australian and Mongolian data, which were not used for the training, produced the diagram of Figure 4, where the soil moisture estimated by the algorithm (SMC_{est}) is compared with the soil moisture measured on ground (SMC_{meas}). The corresponding statistics are: determination coefficient $R^2=0.8$, Root Mean Square Error $RMSE = 0.03 \text{ m}^3/\text{m}^3$, and $BIAS = 0.02 \text{ m}^3/\text{m}^3$. This results was obtained thanks to the contribution of the higher frequencies in accounting for the vegetation conditions, trough PI at X and Ku bands, and in estimating the surface temperature, trough Tb at Ka band, thus reducing the uncertainties.

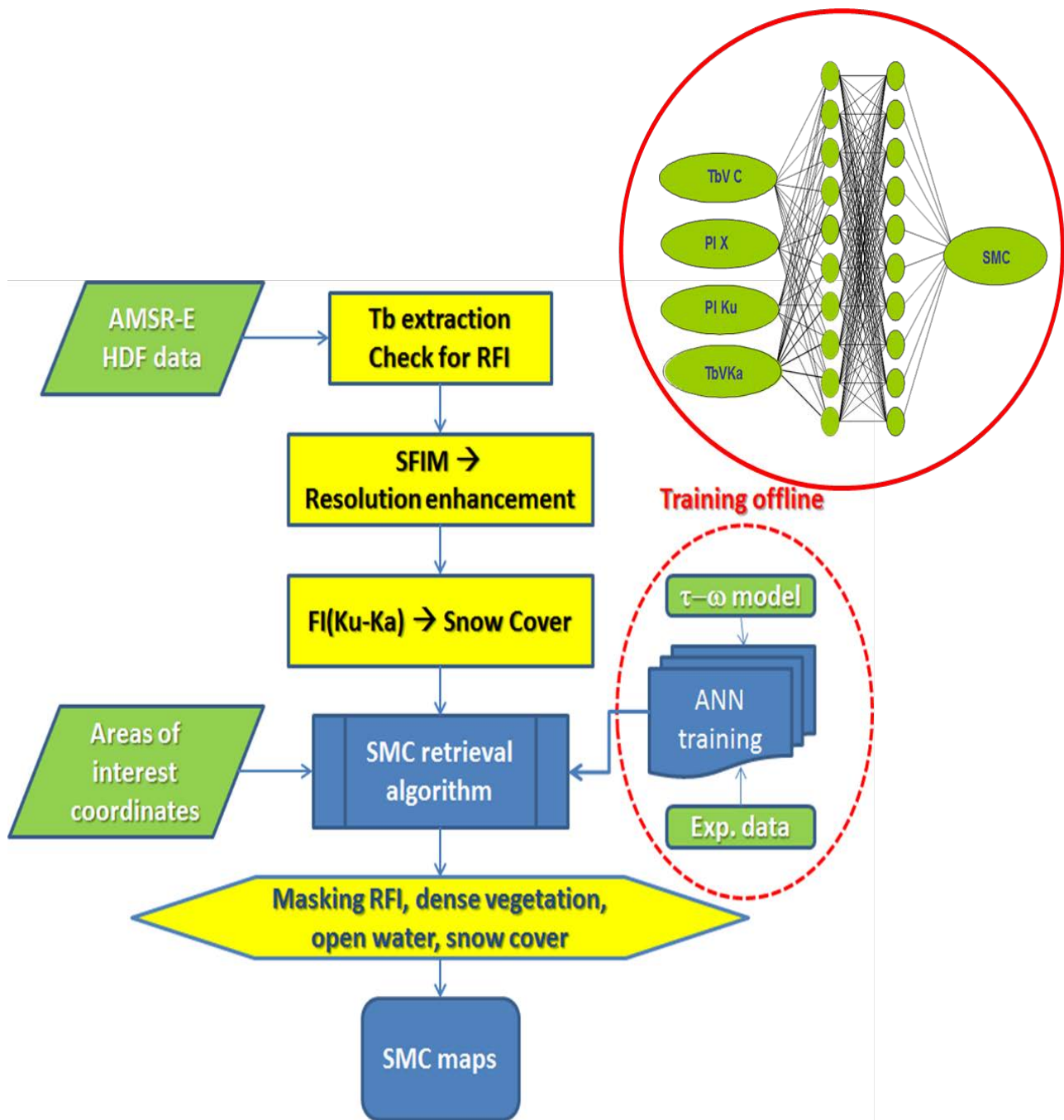


Figure 3. HydroAlgo flowchart with the detail of the ANN architecture

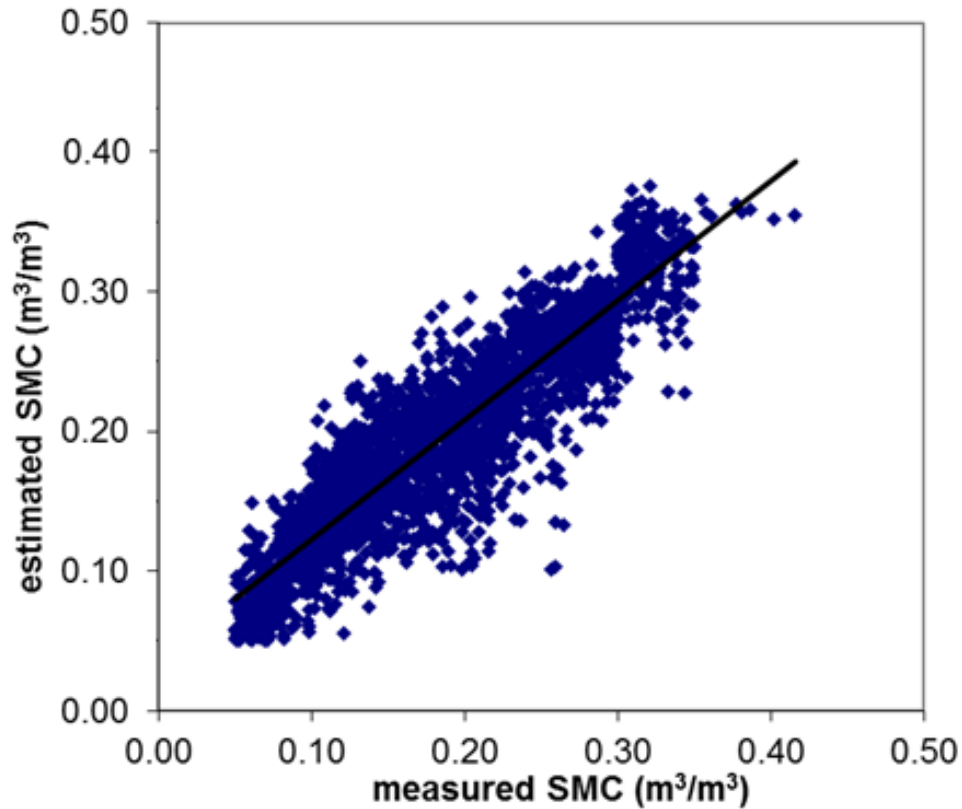


Figure 4. Test of the ANN algorithm with the Australian and Mongolian data not used for the training (Santi et al. 2012).

A further and new validation of the algorithm was carried out by using the datasets collected on the four Agricultural Research Service (ARS) watershed sites in the US from 2002 to 2009 (Jackson, 1993). These sites represent a wide range of ground conditions and precipitation regimes (from natural to agricultural surfaces and from desert to humid regions) and provide long-term in-situ data. The dimensions of each watershed are compatible with the AMSR-E footprint, close to 10x10km² at all frequencies with the application of the SFIM procedure. These sites are well-instrumented and characterized and have been the focus of several AMSR-E validation campaigns (http://nsidc.org/data/amsr_validation/). The vegetation characteristics in the selected watersheds represent typical conditions in specific climate regions, which cover a range of semi-arid to humid. All watersheds have multiple surface soil moisture and temperature sensors (5 cm depth). Characteristics of the individual watersheds are:

- Little Washita (OK): 20 soil moisture stations in a region dominated by the presence of rangeland and pasture plus some agricultural crops. The climate is sub-humid with about 750 mm annual precipitation. The size of the area is 610 km².
- Little River (GA): 29 soil moisture stations and heavily vegetated (predominantly forests, croplands, and pasture). The climate is characterized by hot humid summers and short mild winters (with about 1200 mm annual precipitation). The size of the area is 334 km².
- Walnut Gulch (AZ): 19 soil moisture stations, with brush and grass cover and characterized by a semiarid climate (with about 324 mm annual precipitation). The size of the area is 148 km².
- Reynolds Creek (ID): 19 soil moisture stations in a rangeland dominated area, with snow dominated precipitation. The size of the area is 238 km².

248 AMSR-E data collected on the ARS watersheds from 2002 to 2009 were used for this comparison, together
 249 with the corresponding ground truth data, which have been kindly provided by Dr. Tom Jackson. Due to the
 250 diurnal variations in surface soil moisture that occur during the day and the sensitivity of retrievals to
 251 assumptions about the near surface temperature and moisture profiles, data from the ascending and descending
 252 orbits were analyzed separately.

253 **Table 1.** Performances of HydroAlgo soil moisture algorithms terms of R^2 , RMSE, and bias for both ascending
 254 and descending orbits.

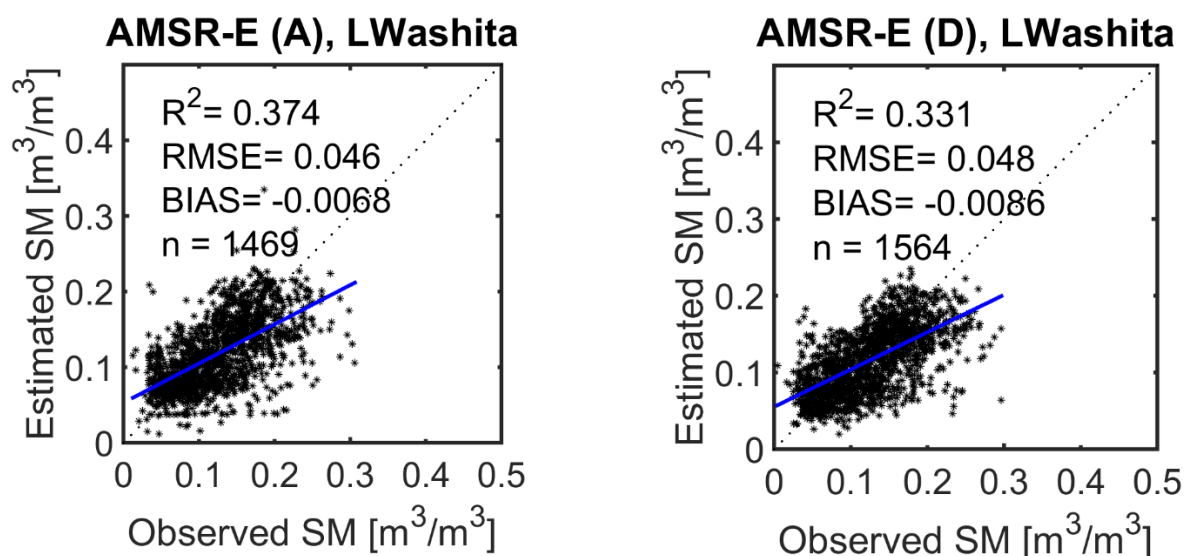
	Ascending			Descending		
	R^2	RMSE (m^3/m^3)	Bias (m^3/m^3)	R^2	RMSE (m^3/m^3)	Bias (m^3/m^3)
Little Washita	0.374	0.046	-0.007	0.331	0.048	-0.009
Walnut Gulch	0.297	0.019	-0.0003	0.255	0.02	0.0011
Little River	0.279	0.043	0.017	0.358	0.039	0.0021
Reynolds Creek	0.606	0.048	0.012	0.406	0.06	0.019

255

256 The overall performances of the algorithm are summarized in Table 1, in which the correlation coefficient
 257 (R^2), RMSE and Bias obtained for the different sites are listed, while the SMC estimated for each test area
 258 using HydroAlgo is plotted against the SMC measured on the ground in Figure 5.

259 The overall accuracy ($\text{RMSE} \leq 0.06 \text{ m}^3/\text{m}^3$ and $\text{Bias} < 0.02 \text{ m}^3/\text{m}^3$) is compliant with the AMSR-E mission
 260 SMC accuracy requirement. The spread of data can be attributed to the spatial variability within the test areas
 261 and the rather large time interval covered by the dataset (8 years).

262



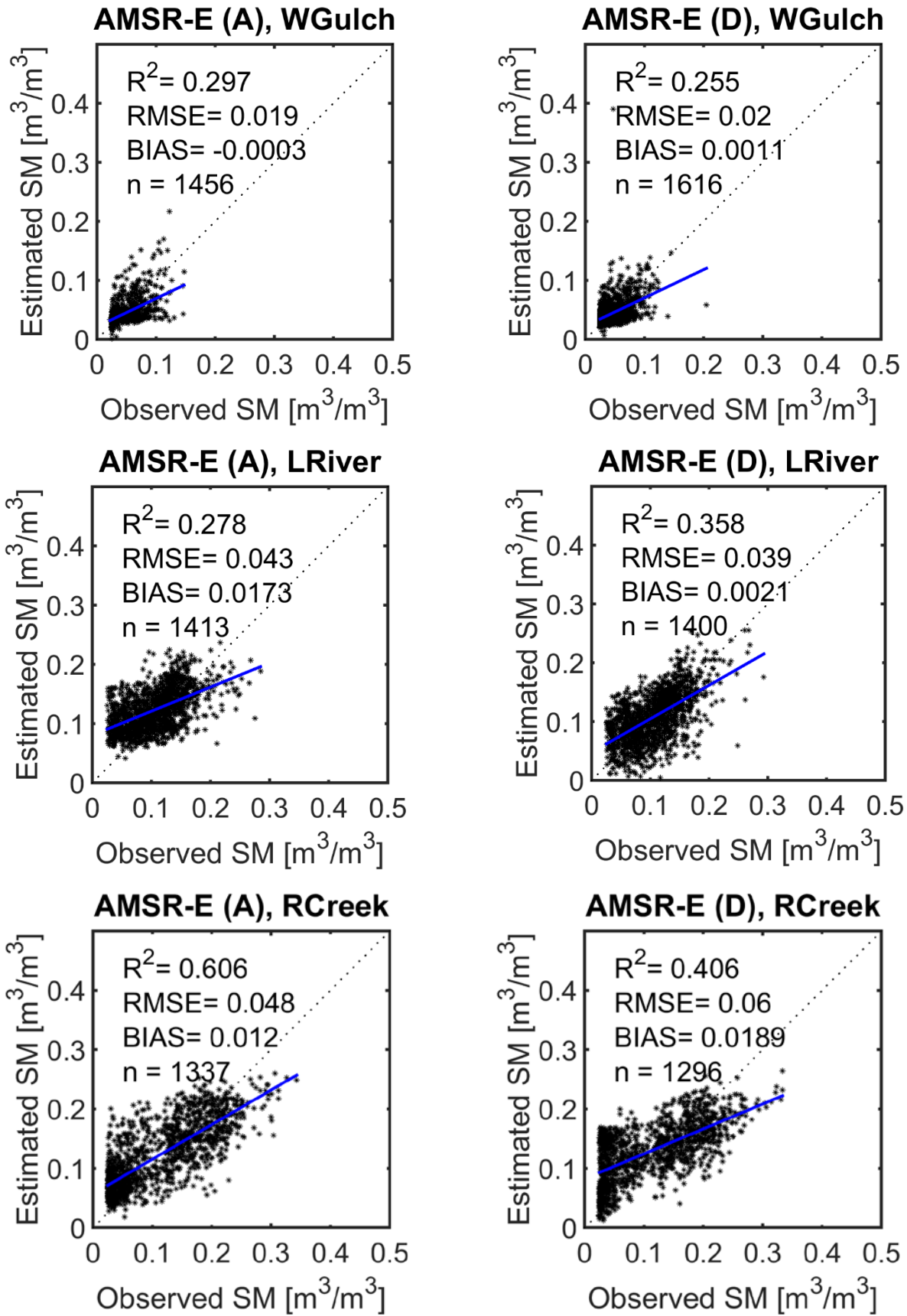


Figure 5: Comparison between the HydroAlgo-derived SMC and the ARS in situ data, where the upper [A] and the lower [B] half of the image summarize the results obtained using the Ascending (A) and Descending (D) overpasses. Continuous lines represent the regression equations.

4. ANN algorithm for real and synthetic aperture radars

4.1 ANN algorithm for ASCAT/SCA

An ANN algorithm for SMC retrieval from the MetOp ASCAT scatterometer (Figa-Saldaña et al. 2002) was developed in the framework of the Round Robin exercise, belonging to the Climate Change Initiative (CCI) soil moisture activities promoted by the European Space Agency (ESA). Data from 150 test sites of the International Soil Moisture Network (ISMN – Dorigo et al. 2011), containing co-located satellite acquisition and ground measurements of SMC were provided to the Round Robin participants for calibration and validation of the algorithms. The calibration dataset contained ASCAT daily acquisitions over 75 test sites for the five considered years, the measurements of top layer SMC from the ground stations and the Global Land Data Assimilation System (GLDAS) simulated 10 cm soil surface temperature, precipitation, and snow water equivalent, while the validation dataset contained only the ASCAT acquisitions on the remaining 75 test sites.

A preliminary analysis of the relationships between ASCAT measurements and SMC measured on ground confirmed the sensitivity of C- band radar measurements to the latter parameter, pointing out however that the geographic locations and therefore the different climatic and vegetation conditions of each test site significantly affected these relationships. This is evident from the example of Figure 6, representing the ASCAT backscattering coefficient (σ^0 , in dB) measured in forward beam as a function of the SMC (m^3/m^3) measured for two test sites. The first (ANMA network) is in Europe (Spain) and is characterized by seasonal vegetation cycles while the second (MAQU network) is in the semi-arid Mongolian plateau. The diagram well demonstrates that the presence of vegetation affects the sensitivity to SMC, as indicated by the two different slopes. A correlation of the measured signal to SMC is nonetheless confirmed in both cases, as it is pointed out by the determination coefficients, $R^2=0.24$ and $R^2=0.30$, respectively. This geographical dependence has been accounted for in the ANN inversion algorithm by adding the geographical position as ancillary input.

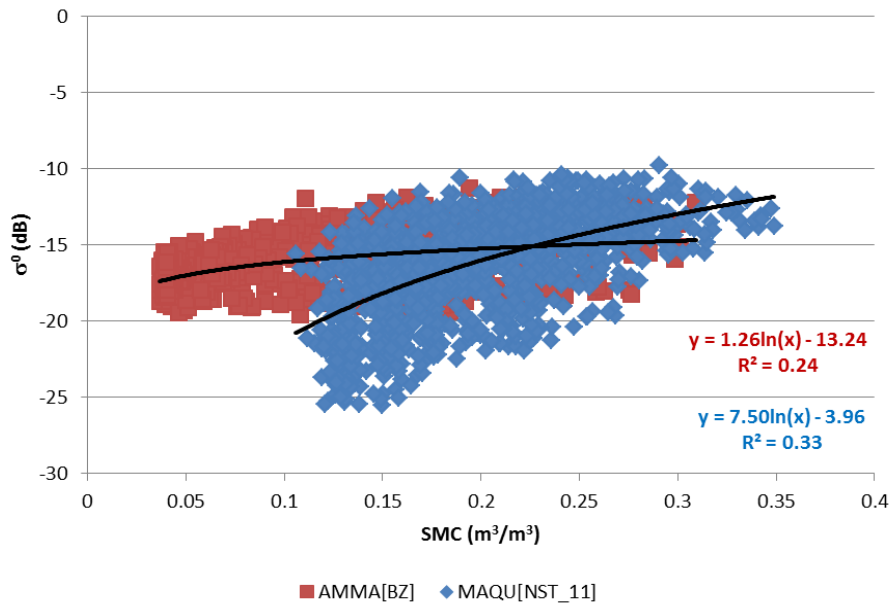
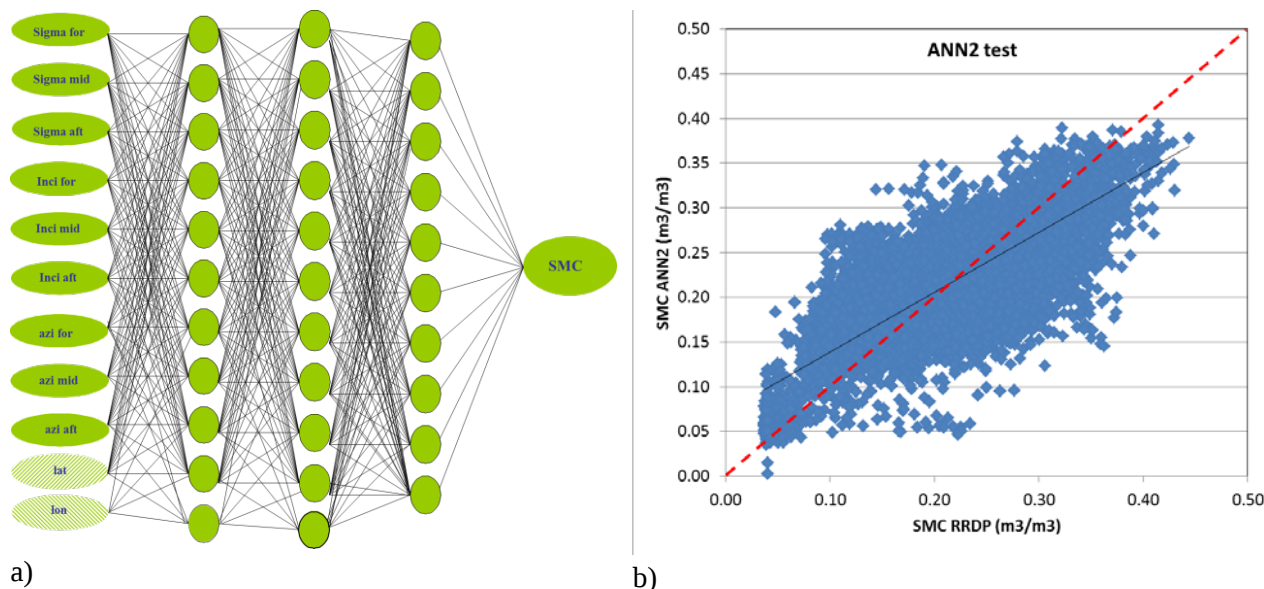


Figure 6. ASCAT measurements in forward beam as a function of the measured SMC for the AMMA (Spain) and MAQU (Mongolia) test sites.

The ANN optimization process for matching the ASCAT RRDP (Round Robin Data Package) dataset resulted in an architecture with three hidden layers of 11+11+10 neurons was selected (Figure 7a). A subset of the RRDP calibration dataset, composed by about the 25% of the data available was considered for generating the subset for training, testing, and validating the ANN (60% - 20% - 20%, randomly sampled), and the remaining 75% of data was considered for the independent validation of the algorithm. ANN inputs were the ASCAT

296 backscatter acquired at the three beams and the corresponding incidence and azimuth angles. Output of the
 297 ANN was the retrieved SMC (Figure 7b).
 298
 299



300 **Figure 7.** a) Architecture of the ANN after optimization. b) ANN independent validation: SMC estimated by
 301 ANN as a function of the target SMC

302 The overall performance of the algorithm is summarized in Table 2. It should be remarked that, in this case,
 303 the training was only based on the experimental data, since the CCI data considered for this application were
 304 representative of the different climatic areas of the world. Moreover ancillary information on soil roughness
 305 and vegetation cover was unavailable for most part of the test sites, making challenging the definition of the
 306 input vectors for running the model simulations. However, by adding the geographic position to the ANN
 307 inputs we noticed a substantial increase of the accuracy (R^2 from 0.49 to 0.67 and RMSE from 0.052 to 0.042
 308 m^3/m^3), with respect to considering backscattering and observation angles only.
 309

310 **Table 2.** Main statistical parameters of the retrieval obtained with the ANN algorithm.

RMSE	R^2	slope	intercept
0.042	0.67	0.67	0.072

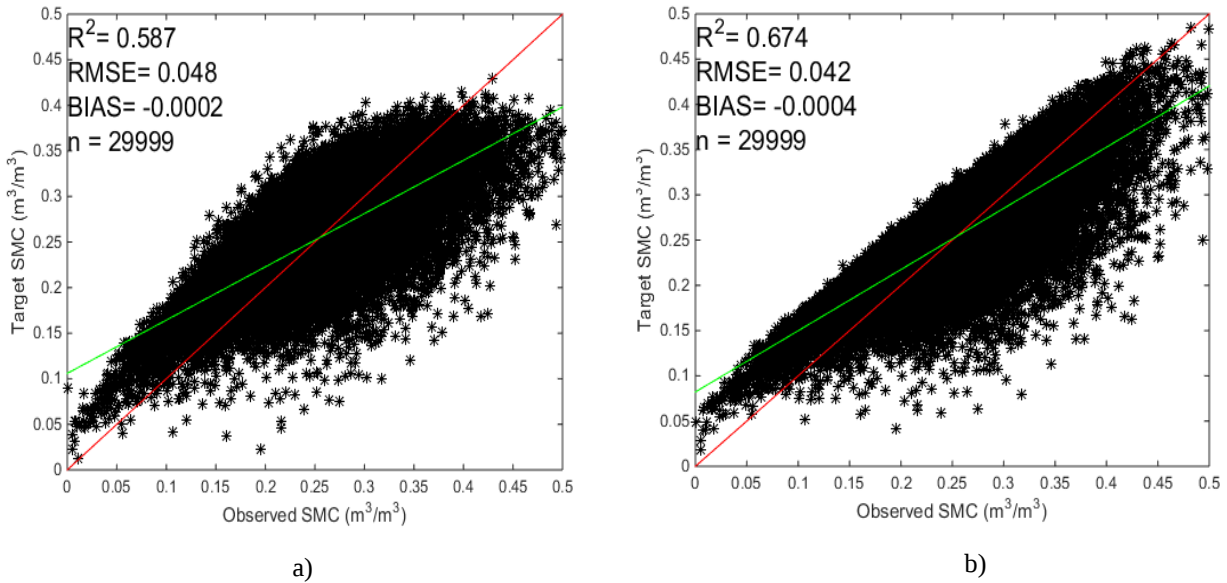
311

312 Model simulations have been instead considered for understanding the potential contribution of a VH polarized
 313 channel in improving the SMC retrieval accuracy from scatterometric acquisitions.

314 The EUMETSAT Polar System of second generation is indeed under study to replace the existing satellite
 315 system and provide continuity of observations in the 2020 timeframe. The inclusion of a cross- polarized (VH)
 316 channel on the mid-beam antenna in the EPS-SG SCA instrument, heir of ASCAT, is at present under study.
 317 The contribution of VH for correcting the effects of vegetation cover and surface roughness on soil moisture
 318 retrieval was then evaluated through the implementation of another ANN algorithm. Basing on the ANN
 319 architecture implemented for ASCAT, which included the backscatter acquired at the three beams and the
 320 corresponding incidence and azimuth angles at VV-polarized data only, a further input was added for the VH
 321 channel that will be acquired in mid-beam by SCA sensor. In this case, training and test of the algorithm were
 322 carried out using only data generated by a simple implementation of RTT, based on the coupling of Oh model
 323 (Oh et al. 1992) and the Vegetation Water Cloud (WCM) model (proposed by Attema and Ulaby, 1978),

324 accounting for the different observation geometries of the 3 SCA beams. The latter model considers the volume
 325 scattering of vegetation and the attenuation effect on the surface scattering of soil under vegetation. By
 326 coupling these models, the backscattering sensitivity to soil moisture for a wide range of vegetation and
 327 roughness conditions can accurately be evaluated.

328 A dataset of about 60 000 backscattering coefficients was generated with this model considering volumetric
 329 SMC values ranging between 0.05 m³/m³ and 0.45 m³/m³, Plant Water Content (PWC) between 0 and 5 Kg/m²
 330 and roughness (expressed as standard deviation of the surface heights, Hstd) between 0.5 and 2 cm. Training
 331 and test were carried out by dividing the available data in two subsets of 30 000 samples each. The first subset
 332 was again divided randomly in 60% - 20% - 20%, for training, test, and validation phase, respectively, while
 333 the second subset was retained for an independent validation. The results of this independent validation are
 334 represented in Figure 8a, for the ASCAT architecture (VV data only) and Figure 8b for the SCA architecture
 335 (VV+VH). This comparison demonstrated that the inclusion of the VH polarized channel is able to improve
 336 the retrieval accuracy, with an increase of the determination coefficient R² from 0.59 to 0.67 and a decrease of
 337 RMSE from 0.048 to 0.042 m³/m³, as displayed in the figures. This result is quite interesting since the
 338 simulations confirmed that the addition of a cross polarized acquisition to the future SCA sensor may help in
 339 improving the SMC retrieval.



340 **Figure 8.** Volumetric SMC (m³/m³) estimated by the ANN algorithm as a function of the target SMC for
 341 ASCAT configuration (VV polarization only) and b) SCA configuration (VV+VH).

342 4.2 ANN Algorithm for C- and X- band SAR

343 Following the approach proposed in Paloscia et al. 2013, an ANN algorithm able to estimate the SMC from
 344 SAR acquisitions at C- and X- bands in different acquisition geometries and polarizations has been
 345 implemented, tested and validated. The algorithm flowchart is represented in Figure 9. Each image was
 346 processed by applying a multilook process, in order to average the intensity in range and azimuth direction, by
 347 using the appropriated window size according the considered SAR sensor. The radiometric calibration was
 348 performed considering the local incidence angle based on the orbital parameters and the DEM, that combined
 349 with satellite orbital parameters allowed also the identification of layover and shadow effects. Layover and
 350 shadowing were almost negligible in flat areas. The geocoded images have a pixel size between 10×10 m² and
 351 30×30 m², and were co-registered in order to be comparable ‘pixel by pixel’ to each other and with other
 352 information, such as the local incidence angle (LIA).

353 The core of the algorithm is composed by 6+6 ANNs, trained for working with backscattering in VV or HH
354 polarization with and without the ancillary information on vegetation, represented by co-located NDVI from
355 optical sensor, and VV+VH or HH+HV combinations, at C- and X- band respectively. In order to generate an
356 algorithm able to work on a large/global scale, the dataset implemented for the ANN training was obtained by
357 combining experimental satellite measurements of backscattering coefficients (σ°), corresponding ground
358 parameters, and data simulated using e.m. forward models. The backscattering of the bare rough surfaces was
359 obtained, in this case, by using the Advanced Integral Equation model (AIEM, by Wu et al. 2004) for co-
360 polarized SAR signal, and the Oh model (Oh et al. 1992) for deriving the cross-polarized backscatter from the
361 co-polarized simulated by AIEM. The contribution of light vegetation was accounted for by using the WCM,
362 deriving the information on vegetation water content (VWC) from the Normalized Difference Vegetation
363 Index (NDVI), measured from the available optical sensors (e.g. Landsat and MODIS), through semi-empirical
364 relationships.

365 Minimum and maximum values of the soil parameters measured during the experimental campaigns (SMC,
366 Hstd and correlation length, Lc) were considered in order to define the range of variability of each soil
367 parameter. Using a pseudorandom function drawn from the standard uniform distribution on the open interval
368 (0,1), rescaled in order to cover the range of each soil parameter, we generated input vectors for the coupled
369 AIEM + WCM models, in order to simulate the backscattering at VV, HH and HV/VH polarizations. More in
370 detail, the input parameters were:

- 371 a) Incidence angle, θ = random between 20 and 50 degrees
- 372 b) SMC range between $0.05 \text{ m}^3/\text{m}^3$ and $0.45 \text{ m}^3/\text{m}^3$
- 373 c) Hstd = random between 1 and 3 cm for C band and 0.5 – 2 cm for X band.
- 374 d) Lc = random between 4 and 8 cm
- 375 e) Dielectric constant derived from random values of SMC between 5% and 45% using the Dobson
376 Model (Dobson et al. 1985).
- 377 f) NDVI = random between 0 and 0.8.

378

379 Since the relationship between Hstd and Lc is rather complicated and reliable measurements of the Lc
380 parameter are barely available, we decided to keep these two quantities independent, associating one random
381 variable with each of these. This procedure was then iterated 10 000 times, thus obtaining a set of
382 backscattering coefficients for each input vector of the soil parameters. The consistency between the
383 experimental data and the model simulation was verified before proceeding to the training phase. The ANN
384 training was carried out by considering the simulated σ° at the various polarizations and the incidence angle
385 as input of the ANN, and the soil parameters as outputs.

386 After training, the ANNs were tested on a different dataset that was obtained by re-iterating the model
387 simulations as described above. The use of a pseudorandom function prevented a correlation between these
388 two datasets: this fact was particularly important in order to evaluate the capabilities of ANN to generalize the
389 training phase and to prevent the overfitting problem. Incorrect sizing of the ANN or inadequate training could
390 cause the overfitting: the ANN returns outputs outside the training range (outliers) when tested with input data
391 that are not included in the training set.

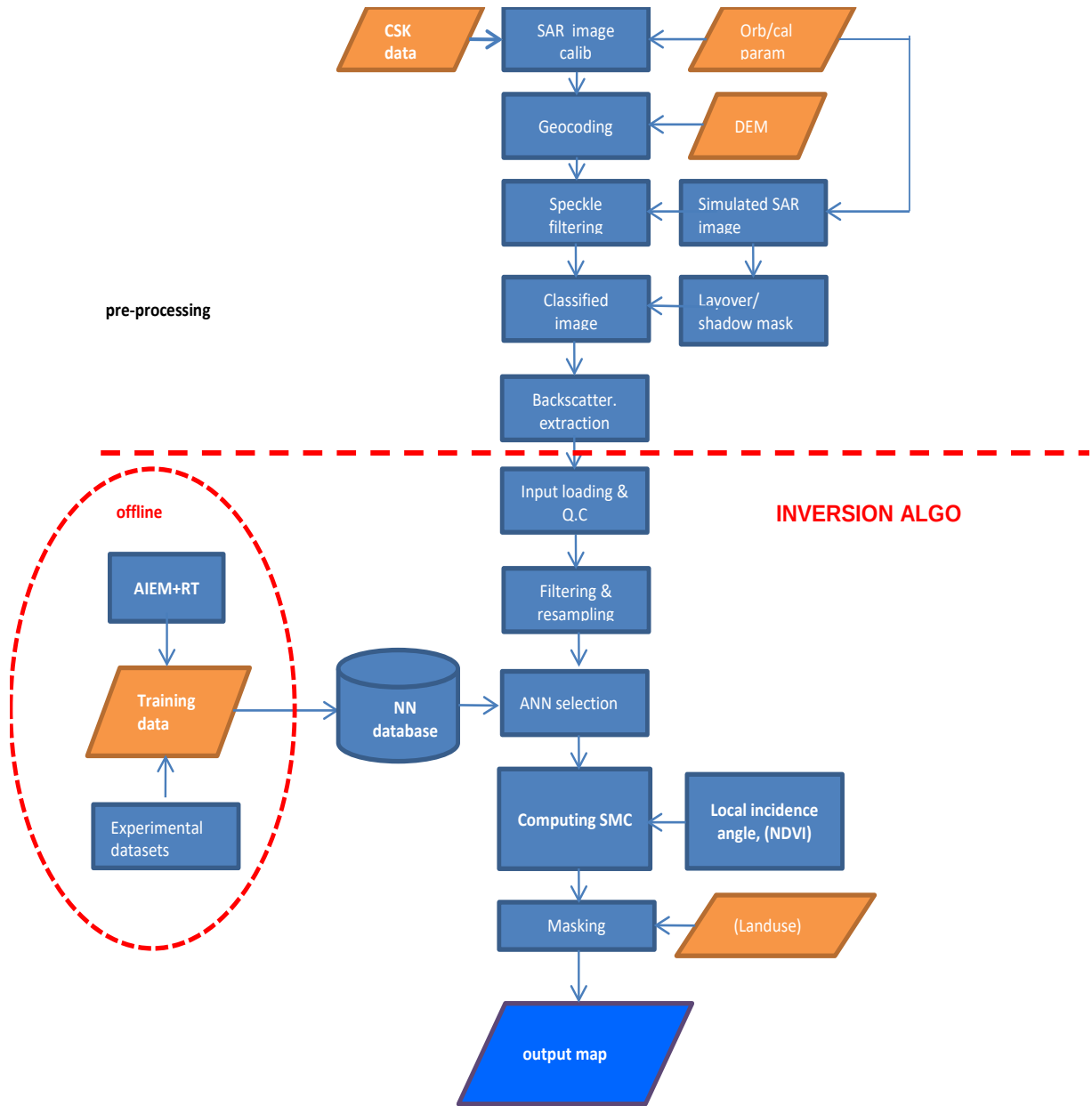


Figure 9. ANN SMC algorithm for SMC retrieval from C- band SAR. The blocks in the circle are performed offline during the training phase of the ANNs.

The algorithm was finally validated by considering the SAR images and corresponding ground truth in several test areas in Italy and Australia, for a total of about 700 field-averaged values of σ° at various polarizations and corresponding SMC measurements at C band and about 600 values at X band.

The results of the overall validation, obtained by comparing all the SMC values retrieved by the algorithm with the corresponding ground truth pointed out that the worst case corresponds to the VV polarization without ancillary information of vegetation (co-located NDVI). In this case the algorithm is able to provide only a rough estimate of SMC, since it cannot account for all the possible surface and vegetation conditions that affect the backscattering sensitivity to SMC, when only one polarization is available.

The HH polarization appears more related to the SMC than the VV, and the NDVI contributes at increasing the retrieval accuracy. The combination of co and cross polarizations, either VV+VH or HH+HV, offers the

best performance with a noticeable accuracy improvement with respect to the other combinations. The obtained results in terms of R^2 , RMSE and bias are summarized in Table 3 at both C and X bands. As expected the best results are achieved at C band, which is more sensitive to SMC and less influenced by the vegetation than X band. At the latter frequency instead, the vegetation effect is dominant, although some sensitivity to SMC is detectable at least for bare and scarcely vegetated surfaces.

Table 3. Statistical parameters (R^2 , RMSE and Bias) of the SMC retrieval by using the ANN algorithm at C and X bands.

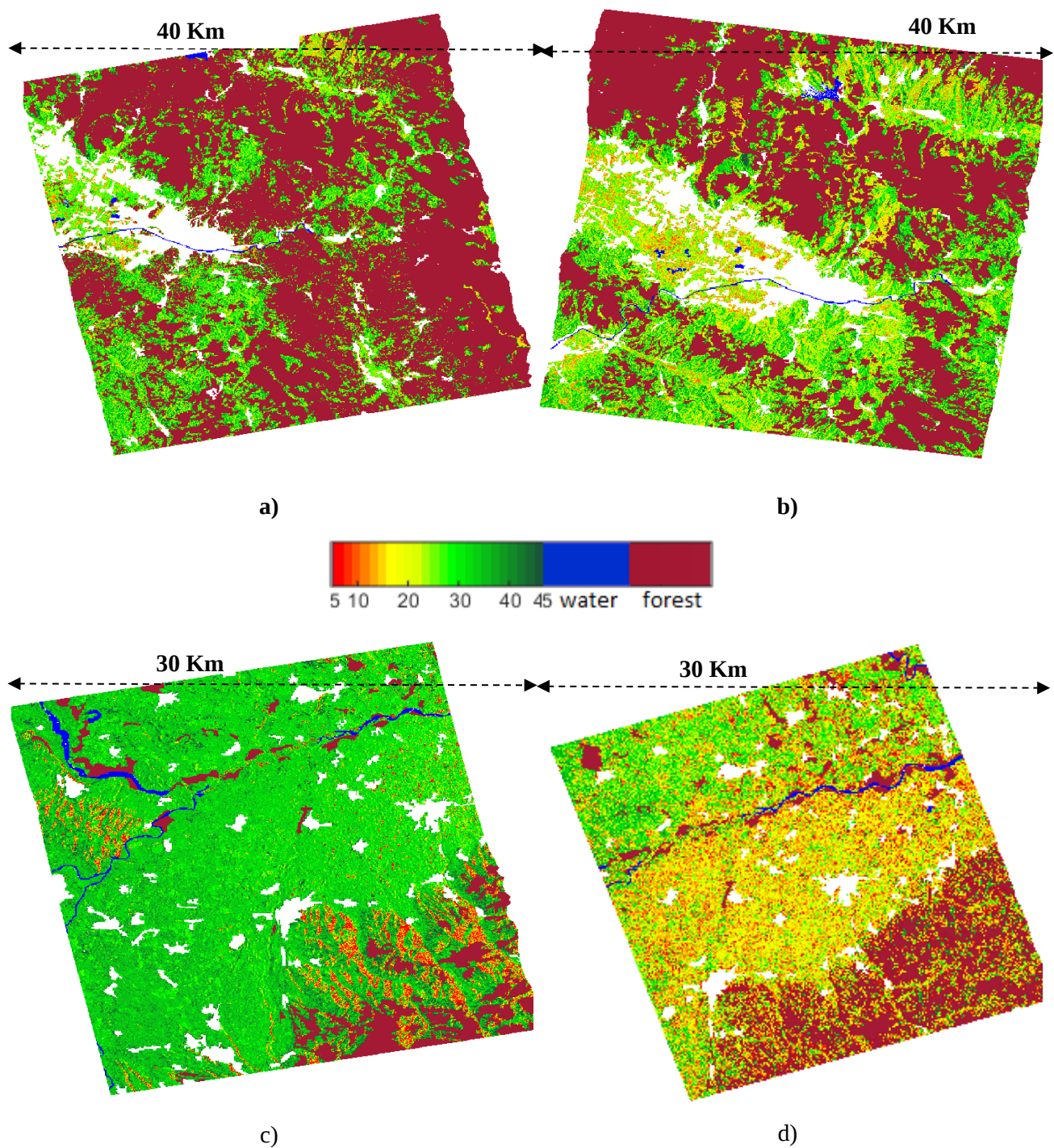
	C band			X band		
ANN	R^2	RMSE	Bias	R^2	RMSE	Bias
VV	0.49	0.067	-0.014	0.33	0.056	0.012
VV + NDVI	0.65	0.052	0.009	0.61	0.046	0.003
HH	0.85	0.026	-0.003	0.57	0.040	0.005
HH + NDVI	0.88	0.023	0.010	0.77	0.034	-0.003
VV+VH	0.59	0.016	-0.009	0.71	0.041	0.008
HH+HV	0.87	0.028	-0.009	0.79	0.032	0.008

413

414 5. Generation of SMC maps at local to global scale

Looking at the operational application of these algorithms, the generation of SMC maps in real or near real time at different resolutions, depending on the input sensor characteristics, takes advantage of the reduced computational cost of the ANN techniques with respect to other statistical methods. In Paloscia et al. 2013, the ANN retrieval algorithm was demonstrated to be able to processing 200 000 pixels/sec, which correspond to about 80 seconds for generating a SMC map at 25x25 m² resolution from an input SAR image of 100x100 km².

Although the SMC maps cannot be considered a real validation of the retrieval algorithms, since adequate ground truth for comparing the algorithm outputs at large and global scale is barely available, these maps represent an effective tool for verifying qualitatively the validity of the training process. Maps characterized by extreme SMC variations from a pixel to an adjacent one, by large percentages of outliers, and maps in which SMC spatial patterns are not detectable indicate that the training was not successful or that the training set was not well representative of the observed target, although the SMC estimated in the control points was close to the ground truth. In these cases, the ANN should be retrained and it should be verified if the training set was representative of the whole range of the input microwave data and output SMC. As an example of the operational capabilities of these algorithms, four SMC maps at basin scale are represented in Figure 10. These maps were derived from COSMO-SKYMED SAR images, which were collected on two test sites, in central (Sesto) and northern (Scrivia) Italy. Map dimensions are 40x40 Km², and white and blue colours represent masking for urban areas and water bodies respectively.



433

434 **Figure 10.** SMC maps of:

- 435 a) Sesto test area, 30 April 2011, obtained using the HH polarization only
- 436 b) Sesto test area, 19 September 2011, obtained using the HH polarization and NDVI
- 437 c) Scrivia test area, 11 April 2010, obtained using the HH-HV polarizations
- 438 d) Scrivia test area, 28 September 2010, obtained using the HH-HV polarizations
- 439

440 The diagram of Fig. 11 was inserted here as a partial validation of the previous maps. Estimated values of

441 SMC at X band in HH, VV, and HV polarizations, are represented as a function of SMC measured on ground

on selected fields. Since all the COSMO-SKYMED configurations have been taken into account, we can note that the resulting determination coefficient and slope are rather high ($R^2=0.8$, $s=0.78$), thus confirming a good correlation between measured and estimated soil moisture. The RMSE is $0.04 \text{ m}^3/\text{m}^3$, pointing also out that the role played at this frequency by surface roughness and vegetation is significant.

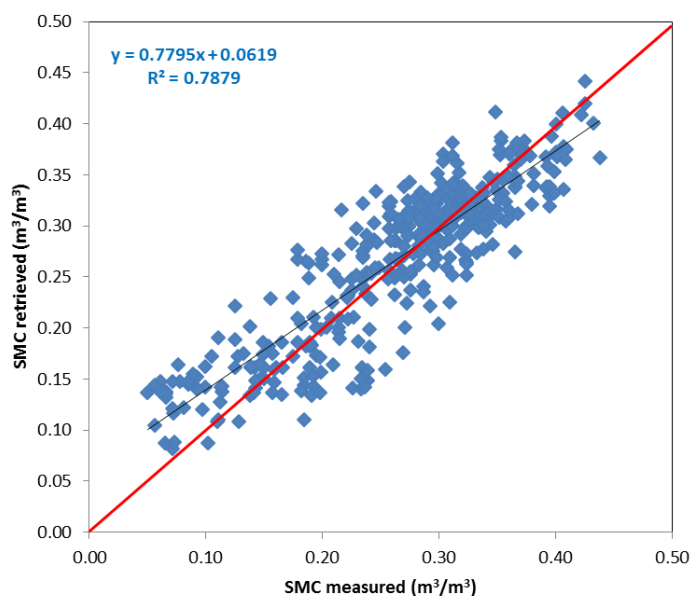


Figure 11. ANN estimated SMC is plotted vs. the corresponding ground measured SMC (%). The regression line is $\text{SMC retrieved} = 0.78 * \text{SMC measured} + 0.06$, with $R^2 = 0.79$.

Moving from basin to global scale, Figure 12 represents six SMC maps of a portion of the world (Europe and Africa) obtained as weekly average of AMSR-E acquisitions collected in different seasons between December 2009 and September 2010. Masking for snow cover areas (white), dense vegetation (dark green), desert (red), and water (blue) were carried out in the maps. A quantitative point-to-point validation of each image is obviously problematic; however, the algorithm appears to be able to reproduce the average moisture conditions of the different climatic areas of the Earth, following the seasonal variations of SMC. This is particularly evident in the areas covered by light vegetation, e.g. Sahel region and Spain, where in summer there is a general decrease of SMC (orange-red colors). An opposite behavior is instead visible in the areas in the austral hemisphere (e.g. South Africa). The reliability of these maps can be also evaluated looking at the presence of outliers, i.e. estimated SMC values outside the training range, which indicate an inappropriate training. For all the displayed maps indeed the percentage of outliers was really negligible ($< 0.1\%$).

6. Conclusions

An application of Artificial Neural Networks (ANN) techniques for retrieving the SMC from active and passive microwave satellite acquisitions was presented here. ANN based algorithms were developed and tuned for working with active and passive microwave acquisitions at several frequencies varying from C (5GHz) to Ka (37GHz) bands. Large datasets of co-located satellite acquisitions and direct SMC measurements on several test sites situated worldwide were used along with simulations from forward electromagnetic models for setting up, training and validating these algorithms. In detail, ANN based algorithms for the SMC retrieval developed and validated with AMSR-E, ASCAT, ENVISAT and Cosmo-SkyMed data were presented.

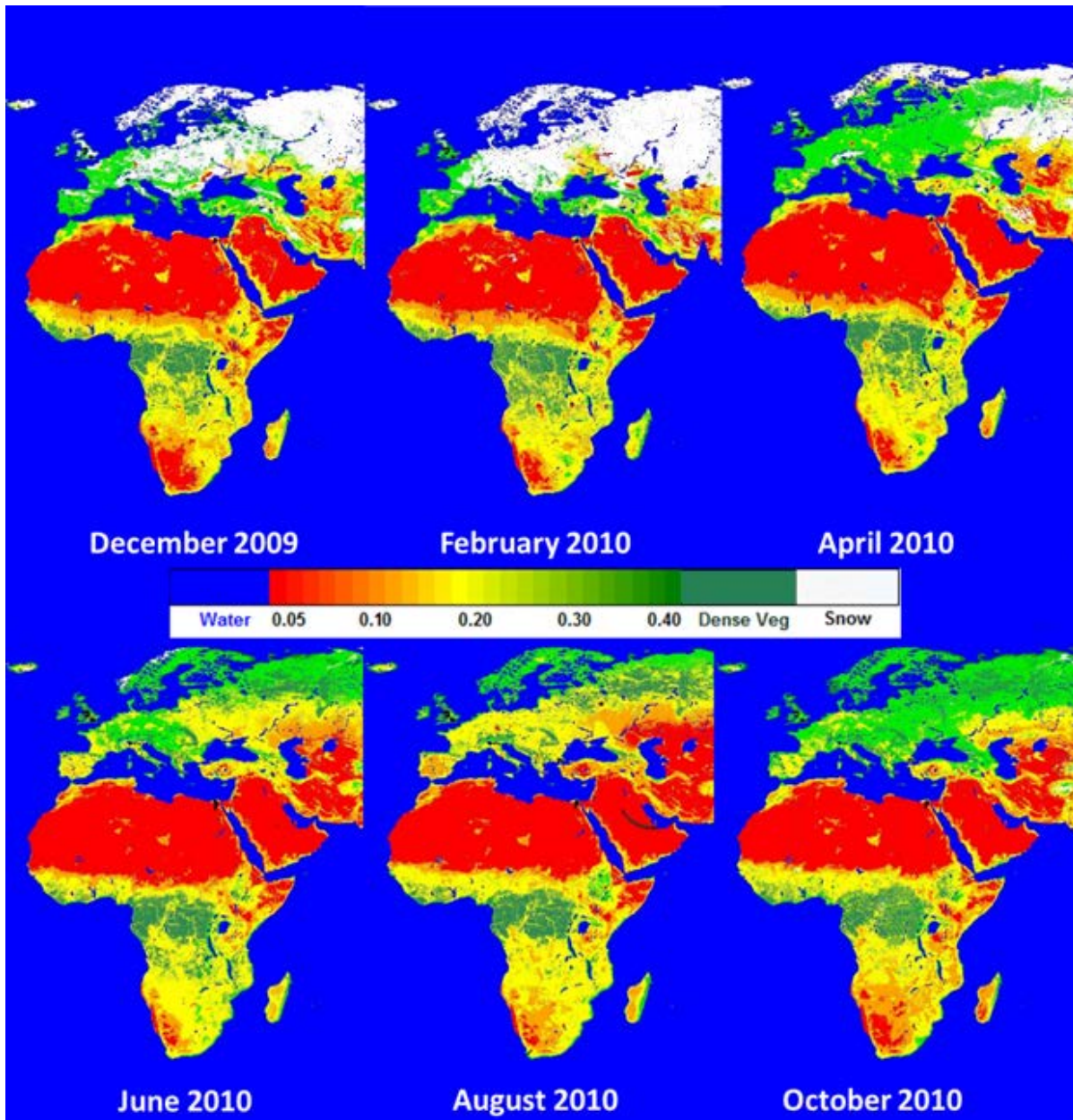


Figure 12. SMC maps of Europe and Africa obtained as weekly average of AMSR-E acquisitions collected in different seasons between December 2009 and September 2010.

The overview of the retrieval algorithms presented here, demonstrated that ANN are a very powerful tool for estimating soil moisture at both local and global scale, provided they have been trained with consistent datasets made up by both experimental and theoretical data. The performed test confirmed the flexibility of this method and the possibility of using it for both active and passive sensors with high accuracy and computational speed. Test of the algorithms returned accuracy values of about $0.05 \text{ m}^3/\text{m}^3$ of SMC or better, making these applications compliant with the usual accuracy requirements for SMC products from space. Moreover the possibility of repeating the training with new datasets enables easily the improvement of the retrieval accuracy, making this technique adaptable to new data and sensors and flexible. Another advantage of these algorithms is in the capability of merging data coming from different sources, as other sensors or ancillary information, into a unique retrieval approach. It was the case of the algorithm implemented for C- and X- band SAR that

485 takes advantage of the NDVI information from optical sensors (Landsat/Modis) when available for improving
486 the SMC retrieval accuracy.

487 The main constraint for accurate retrievals is due to the training process: the retrieval error may be large if the
488 ANN is tested with data not correctly represented in the training. Large datasets are therefore needed for
489 properly training the ANN, in order to cover the whole range of the microwave data and corresponding
490 moisture condition of the observed surface. It should be noted that there is not a unique way for defining the
491 training set. Some a priori knowledge and the support of model simulations help in setting the range of each
492 surface parameter, in order to make the training set as representative as possible of the observed surface.
493 Testing and validation on independent datasets (i.e. scarcely related to the data considered for training) may
494 indicate if the training has been achieved properly. In particular, the use of electromagnetic models for
495 generating large training dataset is one of the best methods for avoiding the danger of 'black box' algorithm
496 and to make sure that the results are based on physical assumptions. Since the training is performed off-line,
497 before starting the data processing, the computational speed of ANN is not hampered by this procedure.

498 Next steps should be the gathering of extended datasets for a more accurate validation of the results, especially
499 on a large scale where the possibility of validation is limited by the scarcity of ground truth data. Moreover,
500 this method should be applied to new sensors, whose more advanced characteristics will allow the retrieval of
501 more accurate results in terms of RMSE, Bias and other statistical parameters. An important further step, even
502 in view of the new SMAP mission, is the integration of active and passive data in a single algorithm, capable
503 of retrieving soil moisture at different scales with improved accuracy with respect to that one of separate
504 algorithms.

505 **7. Acknowledgments**

506 This research work was partially supported by the JAXA ADEOS-II/AMSR-E and GCOM/AMSR2 research
507 programs, by the ESA/ESTEC contract n°4000103855/11/NL/MP/fk on GMES Sentinel-1 Soil Moisture
508 algorithm development, by the EUMETSAT contract n. EUM/C0/14/4600001368/JF on the use of SCA cross-
509 pol for the Soil Moisture retrieval and by the ASI Hydrocosmo proposal n. 1720 on the retrieval and monitoring
510 of Land Hydrological parameters for Risk and Water Resources Management.

511 The authors wish to thank JAXA for providing the Mongolian and Australian datasets, Tom Jackson and Iva
512 Mladenova for providing the ground data from the ARS watersheds and Wolfgang Wagner for the involvement
513 in the ESA CCI Round Robin exercise.

514

References

- Attema, E. P. W. & Ulaby, F. T. 1978. Vegetation modeled as a water cloud. *Radio Science*, 13, 357-364.
- Barret, B.W.; Dwyer, E.; Whelan, P. 2009. Soil moisture retrieval from active spaceborne microwave observations: an evaluation of current techniques. *Remote Sensing*, 1, 210–242.
- Del Frate F., P. Ferrazzoli, G. Schiavon. 2003. Retrieving soil moisture and agricultural variables by microwave radiometry using neural networks. *Remote Sensing of Environment*, 84(2), 174–183. doi: 10.1016/S0034-4257(02)00105-0
- Dobson M. C., F. T. Ulaby, M. T. Hallikainen, and M. A. El-Rayes. 1985. Microwave dielectric behavior of wet soil—Part II: Dielectric mixing models. *IEEE Transactions on Geoscience and Remote Sensing*, 23, 35–46.
- Dorigo, W.A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M., Mecklenburg, S., van Oevelen, P., Robock, A., & Jackson, T. 2011. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, 15, 1675-1698.
- Elshorbagy and K. Parasuraman. 2008. On the relevance of using artificial neural networks for estimating soil moisture content. *Journal of Hydrology*, 362, 1–18. Available at: <http://www.sciencedirect.com/science/article/pii/S0022169408004204>.
- Figa-Saldaña J., J.J.W. Wilson, E. Attema, R. Gelsthorpe, M.R. Drinkwater and A. Stoffelen. 2002. The Advanced Scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers. *Canadian Journal of Remote Sensing*, 28(3), 404-412.
- Gruber A., S. Paloscia, E. Santi, C. Notarnicola, L. Pasolli, T. Smolander, J. Pulliainen, H. Mittelbach, W. Dorigo, W. Wagner. 2014. Performance Inter-Comparison of Soil Moisture Retrieval Models for the Metop-A ASCAT Instrument. *Proceedings of Geoscience and Remote Sensing Symposium IGARSS 2014* *Proceedings*, 2455-2458. ISBN: 978-1-4799-5774-3
- Hornik K. 1989. Multilayer feed forward network are universal approximators. *Neural Networks*, 2(5), 359-366.
- Imaoka, K.; Takashi, M.; Misako, K.; Marehito, K.; Norimasa, I.; Keizo, N. 2012. Status of AMSR2 instrument on GCOM-W1, earth observing missions and sensors: Development, implementation, and characterization II. *Proc. SPIE* 2012, 852815, doi:10.1117/12.977774.
- Jackson, T.J. 1993. Measuring surface soil moisture using passive microwave remote sensing. *Hydrological Processes*, Copyright © 1993 John Wiley & Sons, Ltd, 7(2), 139–152.
- Jackson T. J., Cosh M. H., Bindlish R., Starks P. J., Bosch D. D., Seyfried M. S., Goodrich D. C., and Moran M. S. 2010. Validation of Advanced Microwave Scanning Radiometer soil moisture products. *IEEE Transactions on Geoscience and Remote Sensing*, 48(12), 4256-4272.
- Linden A. and J. Kinderman. 1989. Inversion of multi-layer nets. *Proc. Int. Joint Conf. Neural Networks*, 2, 425-43.
- Lobl, E. 2001. Joint Advanced Microwave Scanning Radiometer (AMSR) Science Team meeting. *Earth Observer* 13(3), 3-9.
- Mladenova I.E., T. J. Jackson, E. Njoku, R. Bindlish, S. Chan, M.H. Cosh, T.R.H. Holmes, R.A.M. de Jeu, L. Jones, J. Kimball, S. Paloscia, E. Santi. 2014. Remote monitoring of soil moisture using passive microwave-

555 based techniques — Theoretical basis and overview of selected algorithms for AMSR-E. *Remote Sensing of*
556 *Environment*, 144, 197–213.

557 Mo T., B. J. Choudhury, T. J. Schmugge, J. R. Wang, and T. J. Jackson. 1982. A model for microwave emission
558 from vegetation covered fields. *Journal of Geophysical Research*, 87, 11229–11237.

559 Moody, J.E., Hanson, S.J., and Lippmann, R.P. 1992. The Effective Number of Parameters: An Analysis of
560 Generalization and Regularization in Nonlinear Learning Systems, *Advances in Neural Information Processing*
561 *Systems* 4, 847-854.

562 Nelder, J. A. & Mead, R. 1965. A Simplex Method for Function Minimization. *Computer Journal*, 7, 308-313.

563 Njoku E. G. and L. Li. 1999. Retrieval of Land Surface Parameters Using Passive Microwave Measurements
564 at 6–18 GHz. *Transactions on Geoscience and Remote Sensing*, 37(1), pp. 79-93.

565 Notarnicola, C.; Angiulli, M.; Posa, F. 2006. Use of radar and optical remotely sensed data for soil moisture
566 retrieval on vegetated areas. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 925–935.

567 Notarnicola C., M. Angiulli, and F. Posa. 2008. Soil Moisture Retrieval From Remotely Sensed Data: Neural
568 Network Approach Versus Bayesian Method. *IEEE Transaction on Geoscience and Remote Sensing*, 46(2),
569 547–557.

570 Notarnicola C. 2014. A Bayesian Change Detection Approach for Retrieval of Soil Moisture Variations Under
571 Different Roughness Conditions. *IEEE Geoscience and Remote Sensing Letters*, 11(2), 414-418.

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

574 Paloscia S. and P. Pampaloni. 1988. Microwave Polarization Index for Monitoring Vegetation Growth. *IEEE*
575 *Transactions on Geoscience and Remote Sensing*. 26(5), 617-621.

576 Paloscia, S.; Macelloni, G.; Pampaloni, P.; Santi, E. 2004. The contribution of multi-temporal SAR data in
577 assessing hydrological parameters. *IEEE Transactions on Geoscience and Remote Sensing Letters* 1, 201–205.

578 Paloscia S., G. Macelloni, E. Santi. 2006. Soil moisture estimates from AMSR-E brightness temperatures by
579 using a dual-frequency algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 3135–3144.

580 Paloscia S., P. Pampaloni, S. Pettinato, E. Santi. 2008. A Comparison of Algorithms for Retrieving Soil Moisture
581 From ENVISAT/ASAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 46(10), 3274-3284.

582 Paloscia S., P. Pampaloni, S. Pettinato, E. Santi. 2010. Generation of soil moisture maps from
583 ENVISAT/ASAR images in mountainous areas: a case study. *International Journal of Remote Sensing*, 31(9-
584 10), 2265 – 2276.

585 Paloscia S., S. Pettinato, E. Santi, C. Notarnicola, L. Pasolli, A. Reppucci. 2013. Soil moisture mapping using
586 Sentinel-1 images: Algorithm and preliminary validation. *Remote Sensing of Environment*, 134, 234-248.
587 doi.org/10.1016/j.rse.2013.02.027.

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

591 Pasolli, L., Notarnicola, C., Bruzzone, L., Bertoldi, G., Della Chiesa, S., Hell, V., Niedrist, G., Tappeiner, U.,
592 Zebisch, M., Del Frate, F., Vaglio Laurin, G. 2011. Estimation of Soil Moisture in an Alpine Catchment with
593 RADARSAT2 Images. Hindawi Publishing Corporation, *Applied and Environmental Soil Science*, Article ID
594 175473, 12 pages, doi:10.1155/2011/175473

595 Pathe, C., W. Wagner, D. Sabel, M. Doubkova, J.B. Basara. 2009. Using ENVISAT ASAR Global Mode data
596 for surface soil moisture retrieval over Oklahoma, USA. *IEEE Transactions on Geoscience and Remote*
597 *Sensing*, 47, 468-480.

598 Pierdicca N., L. Pulvirenti, G. Pace. 2014. A Prototype Software Package to Retrieve Soil Moisture From
599 Sentinel-1 Data by Using a Bayesian Multitemporal Algorithm. *IEEE Journal of Selected Topics in Applied*
600 *Earth Observations and Remote Sensing*, 7(1), 153-163.

601 Pierdicca N., L. Pulvirenti, C. Bignami, 2010. Soil moisture estimation over vegetated terrains using
602 multitemporal remote sensing data. *Remote Sensing of Environment*, 114, 440-448.

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

605 Santi E., S. Pettinato, S. Paloscia, P. Pampaloni, G. Macelloni, and M. Brogioni, 2012, “An algorithm for
606 generating soil moisture and snow depth maps from microwave spaceborne radiometers: HydroAlgo”,
607 *Hydrology and Earth System Sciences*, 16, 3659-3676. doi: 10.5194/hess-16-3659-2012.

608 Santi E., S. Paloscia, S. Pettinato, C. Notarnicola, L. Pasolli and A. Pistocchi. 2013. Comparison between SAR
609 Soil Moisture Estimates and Hydrological Model Simulations over the Scrivia Test Site. *Remote Sensing*, 5,
610 4961-4976. Doi: 10.3390/rs5104961.

611 Smith A. B., J. P. Walker, A. W. Western, R. I. Young, K. M. Ellett, R. C. Pipunic, R. B. Grayson, L.
612 Siriwardena, F. H. S. Chiew, and H. Richter. 2012. The Murrumbidgee soil moisture monitoring network data
613 set. *Water Resources Research*, 48, W07701. Doi: 10.1029/2012WR011976.

614 Tetko, I. V.; Livingstone, D. J.; Luik, A. I. 1995. Neural network studies. 1. Comparison of Overfitting and
615 Overtraining. *J. Chem. Inf. Comput. Sci.* 35 (5): 826–833. [Doi: 10.1021/ci00027a006](https://doi.org/10.1021/ci00027a006)

616 Wagner W., J. Noll, M. Borgeaud, H. Rott. 1999. Monitoring soil moisture over the Canadian prairies with the
617 ERS scatterometer. *IEEE Transactions on Geoscience and Remote Sensing*, 37(1), 206– 216.

618 Wagner W., G. Blöschl, P. Pampaloni, J.C. Calvet, B. Bizzarri, J.P. Wigneron, Y. Kerr. 2007. Operational
619 readiness of microwave remote sensing of soil moisture for hydrologic applications. *Nord. Hydrol.* 38, 1–20.

620 Wu, T.D. & Chen, K.S. 2004. A Reappraisal of the Validity of the IEM Model for Backscattering From Rough
621 Surfaces. *IEEE Transactions on Geoscience and Remote Sensing*, 42(4), 743-753.

622 Yang K., T. Koike, I. Kaihotsu, J. Qin. 2009. Validation of a Dual-Pass Microwave Land Data Assimilation
623 System for Estimating Surface Soil Moisture in Semiarid Regions. *Journal of Hydrometeorology*, 10, 780-793.

624