Soil moisture mapping using Sentinel-1 images: Algorithm and preliminary validation

S. Paloscia [a,*], S. Pettinato [a], E. Santi [a], C. Notarnicola [b], L. Pasolli [b], A. Reppucci [c]

[a] *Institute of Applied Physics, National Research Council (IFAC-CNR), via Madonna del Piano, 10, 50019 Florence, Italy*
[b] *EURAC Research, Bolzano, Italy*
[c] *Starlab, Barcelona, Spain*

ARTICLE INFO

ABSTRACT

The main objective of this research is to develop, test and validate a soil moisture content (SMC) algorithm for GMES Sentinel-1 characteristics. The SMC product, which is to be generated from Sentinel-1 data, requires an algorithm capable of processing operationally in near-real-time and delivering the product to the GMES services within 3 h from observation. An approach based on an Artificial Neural Network (ANN) has been proposed that represents a good compromise between retrieval accuracy and processing time, thus enabling compliance with the timeliness requirements. The algorithm has been tested and subsequently validated in several test areas in Italy, Australia, and Spain.
In all cases the validation results were very much in line with GMES requirements (with RMSE generally <4%SMC – between 1.67%SMC and 6.68%SMC – and very low bias), except for the case of the test area in Spain, where the validation results were penalized by the availability of only VV polarized SAR images and MODIS low-resolution NDVI. Nonetheless, the obtained RMSE was slightly higher than 4%SMC.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Soil moisture content (SMC) is a key hydrological and climatic variable in various application domains (e.g. Entekhabi et al., 1994; Jackson, 1993). Unfortunately, the retrieval from local direct measurements of distributed, quantitative and accurate information relative to the moisture level of soils on a global scale is almost impracticable, due to the high spatial variability of the target variable. Such methods are moreover time consuming and expensive.

In recent decades, many spatially-distributed hydrological models have been developed and successfully applied at scales ranging from small catchments to the globe (e.g. Entekhabi & Eagleson, 1989; Famiglietti & Wood, 1994). However, accurate spatial prediction of soil moisture requires an appropriate accounting for the variability of soil characteristics and climate forcing. Moreover, an accurate assessment cannot be made without adequate spatially-distributed soil moisture measurements at the scale of interest.

The possibility of measuring SMC on a large scale from satellite sensors, with complete, repeated and frequent coverage of the Earth's surface is, therefore, extremely enticing (Beaudoin et al., 1990; Benallegue et al., 1995). Research activities carried out worldwide in the past have demonstrated that sensors operating in the low-frequency portion of the microwave spectrum (P- to L-band) are sensitive to variations in the moisture level of a soil layer, the depth of which depends on the soil characteristics, the moisture profile and the signal wavelength (Macelloni et al., 1999; Shi et al., 1992, 1997). However, at present, most SAR systems onboard remote sensing satellites (e.g., RADARSAT2, COSMO SkyMed and TerraSAR-X) operate at C- and X-bands, which, in terms of sensitivity to soil moisture variations over vegetated terrains, are not the best suited ones for retrieving SMC. Although some preliminary studies indicate the feasibility of retrieving soil moisture also by using the new generation X-band SAR sensors (Baghdadi et al., 2012), working at such high frequencies implies the challenge of coping with the interfering effects introduced by surface roughness and, above all, by vegetation coverage on the backscattering coefficient. Under these operational conditions, an estimate of spatial variations of moisture with the accuracy requirements of the end-user is still problematic and challenging. Even when a priori knowledge of the meteorological conditions, soil properties, and surface coverage are exploited together with correcting procedures for the effects of soil roughness and vegetation, the retrieval of soil moisture remains a challenge.

From an analytical point of view, the retrieval of soil parameters from radar measurements falls within the category of ill-posed problems, because, in general, more than one combination of soil characteristics (in terms of SMC, roughness, vegetation coverage, etc.) leads to the same electromagnetic response at the sensor. Multi-sensor techniques have a certain potential in distinguishing between different contributions of the soil features to the global system response. The rationale is that soil characteristics affect the microwave signal differently and to a different extent, depending on the sensor configuration. By using sensors at different frequencies, polarizations, and incidence

angles, it is thus possible to improve the extraction of information and the retrieval accuracy. However, the availability of multi-frequency and multi-angle sensors on satellite platforms is often difficult to find, and is sometimes even unfeasible, due to acquisition planning limitations and constraints, thus reducing the applicability of this approach in operational scenarios. The bistatic radar technique could provide a further advantage with respect to monostatic radar, but no spacecraft missions implementing that configuration are in operation (Brogioni et al., 2010; Pierdicca et al., 2008).

An effective approach for mitigating the ambiguity introduced by vegetation and roughness conditions on the ground involves focusing attention on temporal variations among subsequent remote sensing acquisitions. The rationale in this case lies in the assumption that the average characteristics of roughness and vegetation cover remain almost unaltered, whereas mainly soil moisture content variations affect the backscattering signal (Balenzano et al., 2011; Doubková et al., 2012; Macelloni et al., 1999; Paloscia et al., 2004; Pierdicca et al., 2010; van der Velde et al., 2012).

Following a multi-temporal change-detection approach, Wagner et al. developed a SMC retrieval algorithm for the ERS scatterometer (Wagner et al., 1999a,b,c). ERS backscattering is described in terms of empirical backscatter parameters and the relative surface SMC according to $\sigma^0(\theta, t) = \sigma^0_{dry}(\theta, t) + S(t)ms(t)$, where θ is the local incidence angle, t is the time, σ^0_{dry} is the backscattering coefficient observed under completely dry soil conditions in decibels, and S is the sensitivity in decibels of the σ^0 to changes in soil moisture. A change-detection approach was also applied by Zribi et al. (2011) to a semi-arid region using high-resolution ENVISAT/ASAR data for soil evaporation evaluation.

The relative SMC ranges from zero in dry soil to unity (or 100%) in a completely saturated soil. This algorithm has been adapted by Pathe et al. (2009) to the ENVISAT/ASAR Global Monitoring (GM) time series. By using the estimated model parameters σ^0_{dry} and S, a relative surface soil moisture index was retrieved from the extrapolated ASAR GM measurements. Very recently, the same change detection algorithm was tested using ENVISAT/ASAR wide-swath (WS) SAR time series within the framework of an evaluation of potential SMC operational algorithms for the ESA Sentinel 1 mission (Hornacek et al., 2012).

The aforementioned research suggests that multi-temporal approaches for retrieving SMC at regional scales from C-band SAR time series can successfully account for surface roughness effects and, to some extent, for low vegetation cover (≤ 1 kg/m^2 of the biomass). The main limitation is still the necessity for having sufficiently frequent SAR acquisitions in order to ensure that the assumption of the stability of roughness and vegetation conditions among the different acquisitions remains still valid.

As an alternative to change-detection-based approaches, a feasible strategy for developing single acquisition SMC retrieval algorithms is based on the inversion of physical-based forward electromagnetic models (Dubois et al., 1995; Fung, 1994; Oh et al., 1992, 2002; Shi et al., 1997; Wu & Chen, 2004). Such models provide a physical-based description of the interactions between microwave electromagnetic radiation and real objects (e.g., bare or vegetated soils), thus enabling a simulation of various experimental scenarios in terms of sensor configurations and soil characteristics. This is a crucial property for ensuring generality and for avoiding the dependence on local site and sensor conditions, which are often common when dealing with empirical-based algorithms. Once forward models have been validated, inversion algorithms that make use of single- or multi-frequency or multi-polarization radar measurements can be developed. However, the mathematical formulation of these models is complicated and makes a direct inversion difficult.

The actual potential of forward model inversion in providing accurate SMC maps has been investigated in various studies. In a recent paper, Paloscia et al. (2008) compared the performance of three inversion algorithms in providing SMC estimates from a series of co- and cross-polarized single acquisition SAR images acquired by the ENVISAT/ASAR sensor on an agricultural test site in Italy. The algorithms taken into consideration were a feed-forward multilayer perceptron (MLP) neural network, a statistical approach based on the Bayes theorem (1958) and an iterative optimization algorithm based on the Nelder and Mead (1965). The evaluation of the algorithms carried out on the basis of estimation accuracy (error percentage), computational complexity (the actual number of pixels processed per second) and potential critical aspects (e.g., the dependence on auxiliary information or the existence of boundary conditions) pointed out that:

- In general, the iterative optimization algorithm provides the highest estimation accuracy. However, it greatly depends on the initialization of the inversion procedure and is extremely slow. This makes it difficult to apply the procedure in vast regions, e.g., in the case of global operational SMC products.
- The statistical approach based on the Bayes theorem provides comparable accuracy in the retrieval, showing at the same time greater stability and fewer requirements for a priori information as compared to the iterative algorithm. Again, the main limitation is represented by the rather high computational burden.
- The MLP neural network (ANN) technique shows slightly poorer, but in general comparable, results compared to the iterative optimization algorithm. Its great advantage lies in the significantly reduced computational time required during the prediction phase. This makes the ANN technique particularly suitable for near real time and operational SMC products, provided that a sufficiently robust and representative set of samples is used during the training phase.

The effectiveness of the ANN inversion algorithm has been further investigated in a subsequent study, in which it is applied to the retrieval of soil moisture content from ENVISAT/ASAR imagery acquired in a mountain area in the Italian Alps (Paloscia et al., 2010). Again, the performances achieved were extremely promising even under these challenging operating conditions. The study also pointed out the potential of the ANN technique in easily and effectively ingesting information extracted from different sources in order to improve the retrieval process, such as the NDVI index derived from optical remote sensing imagery for taking into account the presence of vegetation on the ground.

The potential of machine learning methods for the inversion of forward analytical models and the retrieval of soil moisture was specifically investigated also in the work carried out by Pasolli et al. (2011). In this case, the ANN algorithm was compared with another state-of-the-art method, namely Support Vector Regression (SVR), for the retrieval of soil moisture in bare agricultural areas from C-band scatterometer data. The analysis points out once more the good and similar retrieval performances achieved by the two methods, despite the fact that the SVR showed greater robustness in the presence of outliers and a higher stability in the presence of a reduced number of reference training data. This suggests, again, the importance of a robust and extensive reference dataset for the training of the ANN technique.

The above-mentioned research clearly points out the potential of the theoretical forward model inversion for dealing with the retrieval of soil moisture content from SAR remote sensing data. The availability in the near future of a regular, global and frequent coverage of the Earth's surface with the upcoming satellite SAR systems, such as the Sentinel-1 family, offers on the one hand the potential for fine-scale and near-real-time SMC products and, on the other hand, calls for a better investigation and assessment of theoretical forward model based retrieval algorithms in a wide variety of land surface conditions.

Within the framework of the ESA-funded GMES Sentinel-1 Soil Moisture Algorithm Development (S1-SMAD) project, an algorithm based on the inversion of an analytical electromagnetic model through an ANN has been proposed and intensively validated in order to assess its feasibility for the derivation of near-operational soil moisture content estimates for the Sentinel-1 mission. ESA requirements for GMES SMC product to be matched were the following:

- SMC accuracy: ≤5% in volume
- Spatial resolution: 1 km or less
- Timeliness: 3 h from observation in all cases

The aim of this paper is to present the main outcome of this research activity, by detailing the main characteristics of the proposed retrieval algorithm and discussing the results of the validation activity that has been carried out.

The rest of the paper is organized as follows: Section 2 provides an overview of the developed algorithm, based on Artificial Neural Networks (ANN); in Section 3, the experimental activities have been described together with the test areas; and Section 4 describes the algorithm test and validation. Section 5 contains some conclusions.

2. The ANN approach

The approach that we proposed is based on the work presented in Paloscia et al. (2008), in which some popular inversion strategies, based on statistical and iterative methods (namely, the Bayesian approach, Nelder–Mead minimization, and Artificial Neural Network – ANN – algorithms) were compared. The result of this comparison was that ANN was found to represent the best compromise between retrieval accuracy and computational cost. The algorithm used and described in this paper is based on this first version, which has been subsequently developed and improved. The present algorithm introduced several innovations, which are not only represented by the extended validation, but also by the correction for the effect of vegetation and the presence of a set of ANN able to couple with several sensor configurations (e.g. only one polarization or the combination of some of them). The structure of the algorithm is shown in Fig. 1, which presents the flow chart of ANN.

This study pointed out that the main constraint for obtaining a good accuracy with the ANN approach is the "robustness" of the training process, which needs to be representative of all the surface conditions presented to the ANN during the operational use. The used ANN is a feedforward multilayer perceptron (MLP), with two hidden layers of ten neurons each (Hornik, 1989; Linden & Kindermann, 1989). In order to generate a training set that would meet these requirements, we followed these steps:

- Starting from the archive of past experiments on bare soils available at IFAC, a dataset composed of backscattering coefficients (σ°) and corresponding direct measurements of ground parameters was set up. This dataset did not include the data collected on the four areas (Scrivia, Cordevole, Matera and Alto Adige) that were considered for the algorithm test and validation. This dataset was described and used in several papers (e.g. Baronti et al., 1995; Macelloni et al., 1999, 2002; Paloscia, 2002; Paloscia et al., 2004). The ground parameters included soil moisture and soil roughness, expressed as the standard deviation of the heights (Hstd) and correlation length (Lc). In this phase, we considered only bare surfaces, and disregarded vegetation and vegetation effects. Since this dataset was not sufficient for training the ANN and completely setting the neurons and weights, the AIEM and Oh models (Fung, 1994; Oh et al., 2002; Wu & Chen, 2004) were used for simulating co-polar and cross-polar data, respectively, which were included in the training set. Minimum and maximum values of the soil parameters measured during the experimental campaigns (SMC, Hstd and Lc) were considered in order to define the range of variability of each soil parameter.
- Using a pseudorandom function drawn from the standard uniform distribution on the open interval (0,1), rescaled in order to cover the range of each soil parameter, we generated input vectors for the AIEM and Oh models, in order to simulate the backscattering at VV, HH and HV/VH polarizations. More in detail the input parameters were:
  a) Incidence angle, θ = random between 20 and 50°
  b) Hstd = random between 1 and 3 cm

c) Lc = random between 4 and 8 cm
d) Dielectric constant derived from random values of SMC between 5% and 45% using the Dobson Model (Dobson et al., 1985).

Since the relationship between Hstd and Lc is rather complicated and it is difficult to obtain reliable measurements of the Lc parameter, we decided to keep these two quantities independent, associating one random variable with each of these. This procedure was then iterated 10,000 times, thus obtaining a set of backscattering coefficients for each input vector of the soil parameters.

The ANN training was carried out by considering the simulated backscattering at the various polarizations and the incidence angle as input of the ANN, and the soil parameters as outputs. Two different ANNs were trained specifically for HH and VV polarizations. The consistency between the experimental data and the model simulation was verified before proceeding to the training phase.

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

The ANN performances were tested by considering several configurations of input/outputs, in accordance with the acquisition modes of Sentinel-1, i.e. in VV, HH or HV/VH polarizations.

The least favorable situation corresponded with the availability of images in single polarization only (VV or HH pol.), with no other information on vegetation and without ancillary data. In this case, ANN worked with the backscattering coefficient, σ°, and the incidence angle as inputs, and SMC as output. The results of the test in both polarizations were very close, and showed the following statistical parameters: the determination coefficient (R^2) is equal to 0.80, and RMSE < 4 (% of SMC).

2.1. Vegetation corrections

2.1.1. Single polarization + NDVI

The approach proposed in the previous section cannot be considered suitable when vegetation covered surfaces are observed. The different soil and vegetation conditions strongly affect the relationship between σ° and the moisture content of the soil. This is evident from the diagram of Fig. 2, in which some experimental relationships between σ° and SMC obtained on various test sites are presented. It can be noted that σ° exhibits the same sensitivity to soil moisture for all the analyzed datasets, but with different biases, depending on the different surface types and vegetation features of each test area. The problem of taking into account the effects of vegetation within the algorithm was therefore dealt with by adding a term related to vegetation cover to the formulation of the backscattering responses from bare surfaces. By using the Radiative Transfer Theory as a basis, the contribution of vegetation was modeled by considering both the volume scattering produced by vegetation and the attenuation effect of soil under vegetation on the backscattering. The total backscattering of the vegetated surface was derived from the backscattering of bare surfaces, which was simulated by AIEM, using the so-called "Water Cloud Model" (Attema & Ulaby, 1978; Bindlish & Barros, 2001; Joseph et al., 2010). This model, which is characterized by a rather simple implementation, simulates the backscattering of vegetated surfaces as a function of the soil backscattering and the Vegetation Water Content (VWC) by using two parameters (A and B) that have been set up at two polarizations (VV and HH) on the basis of the available experimental data. Model simulations have been obtained setting A = 0.0855 and B = 0.0126 at VV polarization, and A = 0.0467 and
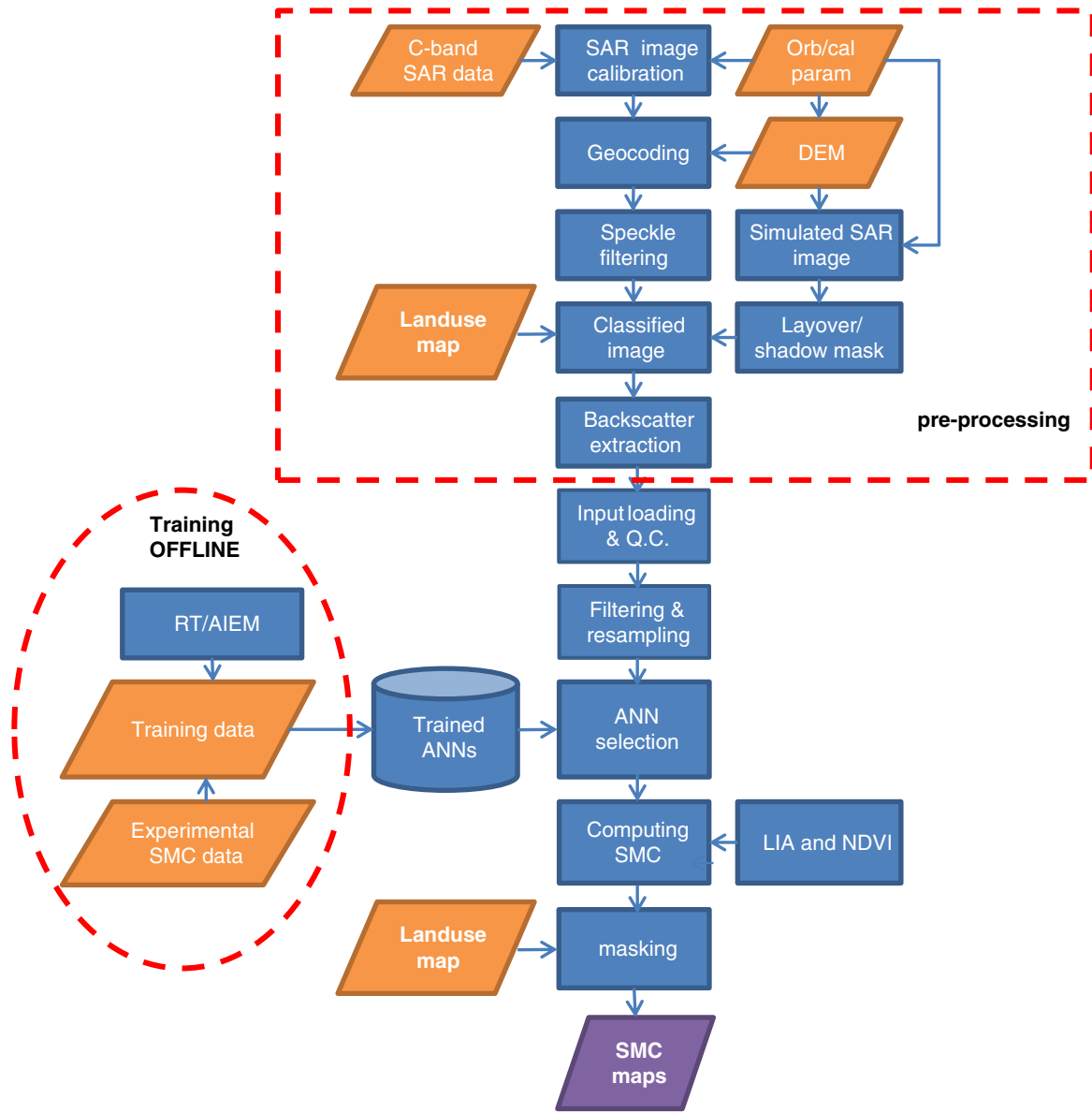
Fig. 1. The flow chart of the ANN algorithm.

B = 0.0155 at HH polarization. VWC was computed from the NDVI using the semi-empirical relationship below, which was obtained from our data archives:

$$VWC = 12.86*NDVI-2.25. \tag{1}$$

We can observe that this relationship, which was indeed validated mainly for agricultural crops (not only cereals) and grass, is close to others existing in literature (e.g. Jackson et al., 2004).

Fig. 3 represents the simulated values of $\sigma°$ in VV polarization from vegetated surfaces as a function of SMC, compared with the available archive data. The experimental data are those of all the available datasets, and were collected under different vegetation conditions, ranging from agricultural crops to grassland and to alpine pastures. The regressions obtained for model simulations and experimental data were very close:

$$\sigma° = 1.98*\ln(SMC)-13.19, R^2 = 0.38 \tag{2}$$

for model simulations and

$$\sigma° = 2.07*\ln(SMC)-13.668, R^2 = 0.25 \tag{3}$$

for experimental data

Analogous results were obtained in HH polarization, as shown in Fig. 4. In this case the regressions were:

$$\sigma° = 3.07*\ln(SMC)-15.42 \left(R^2 = 0.52\right) \tag{4}$$

for model simulations and

$$\sigma° = 2.91*\ln(SMC)-14.58 \left(R^2 = 0.25\right). \tag{5}$$

for experimental data

As for the case of bare soils, two datasets of 10,000 simulated values of $\sigma°$ at the various polarizations were obtained by iterating the model simulations with random inputs. The additional input
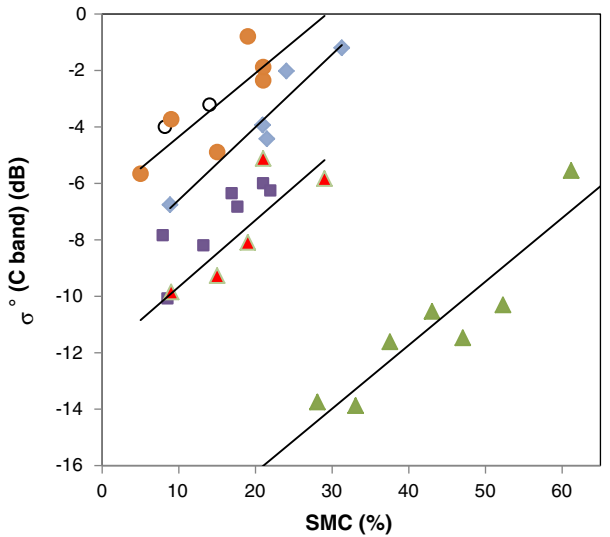
Fig. 2. Backscattering coefficient vs. SMC measured on ground for the archive data collected on different test sites on both bare and vegetated soils. The total number of investigated fields is 31. The sensitivity of the different regression equations ranges from 0.24 dB/%SMC (Vegetated soil − VV pol.) to 0.26 dB/%SMC (bare and vegetated soil − HH pol).

parameter VWC was derived from NDVI values between 0.2 and 0.8, according to Eq. (1). The first dataset was combined with half of the experimental data in order to generate the training set, while the second dataset was used to test the ANN after training. Two ANNs were trained and tested for the two configurations: VV + NDVI and HH + NDVI. The test resulted in the following figures: in the case of HH polarization, the regression slope was 0.998, the determination coefficient, $R^2 = 0.995$ and the RMSE was 0.93 (% of SMC). Poorer results were obtained in the case of VV polarization, where the RMSE increased up to 2.87 (% of SMC). The loss of accuracy in the case of
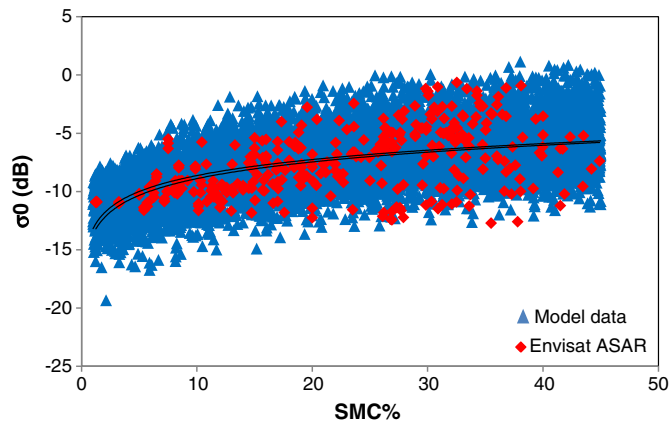

Fig. 3. Comparison between simulated (blue triangles) and measured (red diamonds) σ° in VV pol. as a function of SMC%. The continuous lines represent the regression equations: $\sigma°_{VV} = 1.98\ln(SMC) − 13.19$ ($R^2 = 0.38$) for model data, and $\sigma°_{VV} = 2.69\ln(SMC) − 13.66$ ($R^2 = 0.25$) for experimental data.
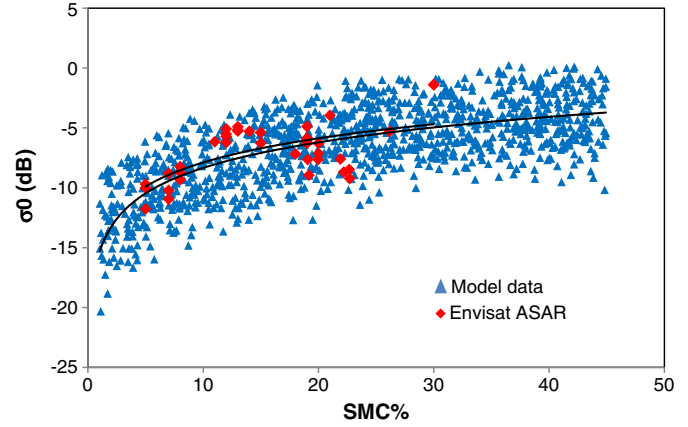

Fig. 4. Comparison between simulated (blue triangles) and measured (red diamonds) σ° in HH pol. as a function of SMC%. The continuous lines represent the regression equations: $\sigma°_{HH} = 2.91\ln(SMC)-14.58$ ($R^2 = 0.4$) for model data, and $\sigma°_{HH} = 3.08\ln(SMC)-15.42$ ($R^2 = 0.52$) for experimental data.

VV polarization was confirmed when the algorithm was tested with the experimental data.

2.1.2. Dual polarization without NDVI

When dual polarized (VV/VH or HH/HV) images are available, without NDVI ancillary information, the sensitivity of each polarization to the different surface features can be considered for distinguishing certain surface types and for evaluating the effect of the vegetation and soil conditions, in order to improve the accuracy of SMC retrieval.

From an analysis of archive data available at IFAC, we found that the Polarization Ratio, PR, i.e. the ratio between co- and cross-polarized backscattering coefficients (PR = $\sigma°_{HH}$ / $\sigma°_{VH}$), exhibits a rather good sensitivity to vegetation cover, expressed as NDVI (slope = −4.69), while the sensitivity to the SMC is basically negligible (slope = −0.03), as is evident from the diagrams of Fig. 5a and b.

This result was confirmed by model simulations. Fig. 6 represents a comparison between simulated and measured (from the archives) PR (HH/HV). The regressions obtained between PR and NDVI for both the experimental data and the model simulations were very close:

$$PR = −4.69*NDVI + 10.27 \tag{6}$$

for experimental data and

$$PR = −4.17*NDVI + 9.90 \tag{7}$$

for model simulations

Similar results were obtained for PR = VV/VH. By following the procedure described in the previous sections, two other ANNs were trained and tested for the two combinations VV + VH and HH + HV, while the PR was considered as a second input instead of NDVI. The results of the test gave a $R^2 > 0.98$ and a RMSE of about 1.3 (% of SMC) for both the ANNs.

It should be noted that the ANN training is carried out "offline", before applying the algorithm to the operational SMC retrieval. If new datasets are made available, the training of the ANN can be updated in order to improve the algorithm performances.
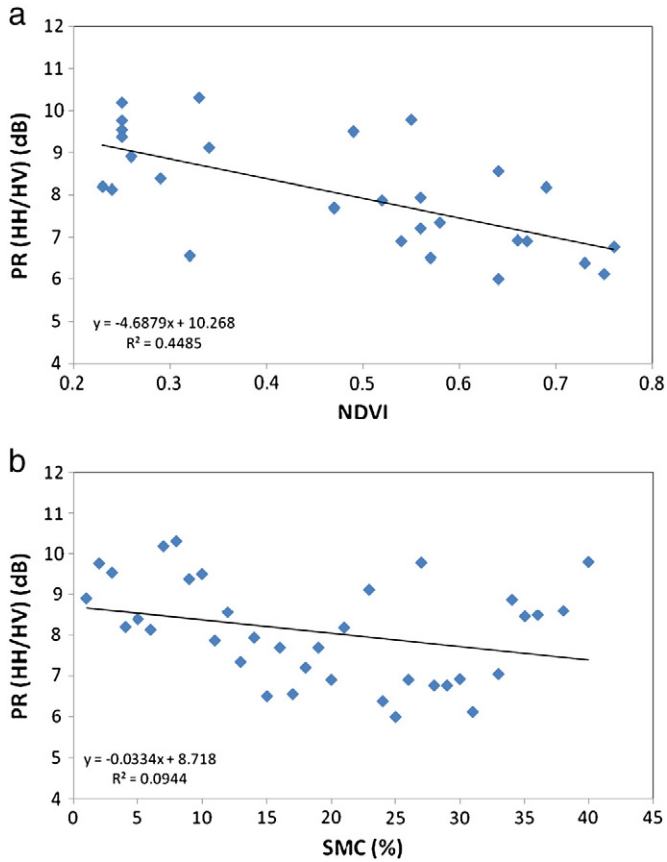
Fig. 5. a) and b): Polarizarion Ratio (PR) between σ°_{HH} and σ°_{HV} as a function of a) NDVI and b) SMC.

3. Experimental activities

The ANN algorithm performances were tested and subsequently validated in the following main test sites (four in Italy, one in Australia, and one in Spain), where SAR images and simultaneous ground truth data have been collected for several years:
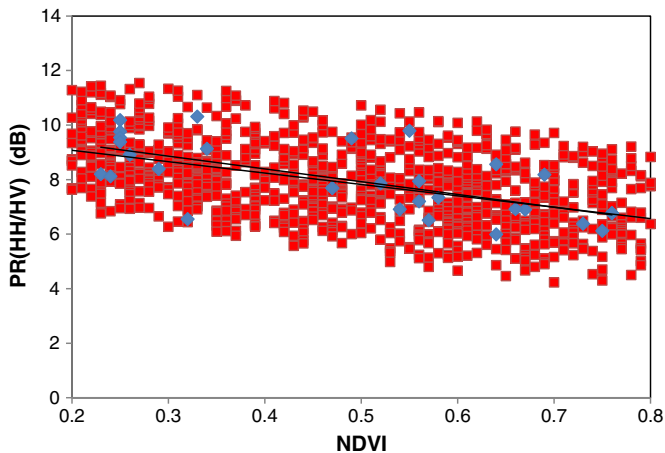
- *Scrivia* test site
- *Matera* test site

Fig. 6. Simulated (red squares) and measured (blue diamonds) PR (HH/HV) as a function of NDVI. The regression lines are the following: PR = −4.69 ∗ NDVI + 10.27, for experimental data, and PR = −4.17 ∗ NDVI + 9.90, for model simulations, respectively.

- *Cordevole* test site
- *Alto* Adige test site
- *Australian* test site
- *Spanish* test site

In Table 1 the main characteristics of the test areas have been summarized.

3.1. Scrivia test site

The Scrivia watershed, which is located in North-West Italy, is a flat alluvial plain measuring about 300 km², situated close to the confluence of the Scrivia and the Po Rivers. The site is characterized by large homogeneous agricultural fields of wheat, corn, sugarbeet, and potatoes (Paloscia et al., 2008).

From 2003 to 2009, the area was monitored by the ENVISAT/ASAR images. These images were mainly collected in HH/HV polarizations and at an incidence angle of 23°. Simultaneously with satellite acquisitions, ground campaigns were carried out in two sub-areas along the Scrivia river close to the town of "Castelnuovo Scrivia", which is at the confluence between Scrivia and Po rivers. The total extent of the two areas is approximately 10 km². The ground measurements involved all the significant vegetation and soil parameters, such as plant density, leaf and stalk dimensions, the number of leaves per plant, plant water content (PWC), soil moisture content (SMC), and surface roughness. The vegetation parameters were measured using conventional methods, while the roughness measurements were obtained by using a needle profilometer 1.2 m in length, with 0.5 cm interval between pins. The statistical parameters for characterizing the roughness, namely the standard deviation of the heights (Hstd) and correlation length (Lc), were computed from the digitalization of the measured profiles. To measure the average volumetric soil moisture of the first soil layer (10–15 cm) Time Domain Reflectometry (TDR) probes were used. At least 5–6 samples of SMC and vegetation were collected for each field considered, while roughness was measured along and across the rows. Table 2 contains a list of the SAR acquisitions used in this work.

3.2. Matera test site

The Matera area, which is located in the Basilicata Region of Southern Italy, has been intensively investigated in numerous remote sensing campaigns, such as that of 1994 during the SIRC/XSAR mission. Between 2000 and 2003, series of campaigns were carried out in order to monitor wheat growth. In this case, ground scatterometer and satellite (ENVISAT-ERS) data were acquired over the test site, along with extensive ground measurements (Mattia et al., 2003). This is a predominantly agricultural area mainly devoted to the cultivation of wheat, which is usually sown at the end of December, reaches its maximum growth at mid May, and is harvested approximately at mid June, depending on weather conditions.

During the 2008–09 campaigns, ground measurements of SMC, Hstd, Lc and PWC were carried out in 5 fields. The ground measurements were carried out following conventional methods similar to the ones indicated for the Scrivia test site. In this case the roughness profilometer has a length of 3 m and a distance between pins of 0.5 cm. The area investigated measured approximately 10 km by 10 km. The soil moisture levels changed from 7% in July to 35% in April, due to the heavy and frequent rain events that occurred in that period. The SAR images acquired are summarized in Table 3.

3.3. Cordevole test site

This watershed, which is located on the foothill of Mount Sella in Northern Italy (Veneto region) was selected because of its relatively smooth topography and the availability of its historical and topographic

Table 1
Synopsis of the main characteristics of the test areas.

Test site	Lat, Lon	Main characteristics	Climatic conditions	Field measurements available (ranges)
Scrivia	45.0 N, 8.8 E	Agricultural area (corn, wheat, sugarbeet, potatoes)	Continental climate with moderately cold winters and hot summers (annual rainfall: 600-700 mm)	Soil moisture (10–30%) Soil roughness (1–3 cm) Vegetation parameters
Matera	40.72 N, 16.61 E	Agricultural area (cereals, mainly wheat)	Mediterranean climate with dry/hot summer and rainy/mild winter (annual rainfall: <500 mm)	Soil moisture (7–35%) Soil roughness (1–3 cm) Vegetation parameters Meteorological data
Cordevole	46.51 N, 11.87 E	Mountain area (grassland and pastures)	Alpine climate (annual rainfall: 1220 mm,49% falls as snow)	Soil moisture (25–55%) Vegetation parameters Meteorological data
Alto Adige (Mazia Valley)	46.70 N, 10.62 E	Mountain area (grassland and pastures)	Dry cold continental climate with strong precipitation gradients. (annual rainfall: 600 mm)	Soil moisture (15–35%) Vegetation parameters Meteorological data
Australia (Yanco River)	−34.97 S, 146.02 E	Pastures and cropland	Semi-arid area (annual rainfall: 600 mm)	Soil moisture (10–45%) Meteorological data
Spain (Emporda)	42.2 N, 2.9 W	Agricultural area (cereals, sunflower, corn, alfalfa)	Mediterranean climate with dry/hot summer and rainy/mild winter (annual rainfall: 600–700 mm, falling mainly in winter and spring).	Soil moisture (5–25%) Vegetation parameters Meteorological data

data. The area covers 95 km^2 at a mean altitude of 1948 m asl, with a mean slope inclination of 51%. In the middle of the basin, at the top of Mount Cherz (2000 m asl), there is a fairly large plateau free of forests.

The site is equipped with a corner reflector and with a meteorological station. During the ENVISAT passes, three ground campaigns were carried out on 14 June, 19 July, and 27 September 2004. The measurements included a photographic survey and measurements of soil moisture (with a TDR probe) as well as fresh biomass of vegetation (gravimetric method) in different parts of the site. At least 90 punctual measurements were carried out during the campaign. In June and July, a significant portion of the area was covered by rather dense grass, which in September was almost dry. The mean value of the soil moisture was close to 40% in June over 45% in July, and less than 35% in September.

The ASAR images obtained over the area are summarized in Table 4.

3.4. Alto Adige/South Tyrol test site

The study area of the Mazia Valley, a side valley of the Venosta Valley, is located in the South Tyrol region (Northern Italy). The catchment area extends for about 100 km^2, with an altitudinal range of between 920 m asl (Sluderno) and 3738 m asl (Palla Bianca). The site is characterized by a mean annual precipitation (Mazia, 1580 m asl) of 525 mm. However, different moisture patterns can be observed, which are mainly due to the irrigation practices in highly intensive grasslands (in the valley floor) and the presence of small rivers that descend from the mountain tops. Grasslands and pastures with heterogeneous characteristics in terms of vegetation species can easily be found by moving from the lower altitudes to the higher ones. The area is equipped with permanent stations for the measurements of air and soil parameters. It is one of the test sites of the ESA AO project "SOFIA" (Pasolli et al., 2011), mainly devoted to estimating soil moisture by

using fully polarimetric RADARSAT2 images. Radarsat2 acquisitions over this area are listed in Table 5.

Two field measurement campaigns were carried out in the Mazia valley at the same time as the satellite acquisitions.

Samples of grass and soil were collected in order to obtain accurate measurements of biomass, vegetation water content, soil gravimetric moisture and bulk density. Moreover, also non-destructive measurements of SMC, by using a TDR sensor, were carried out.

3.5. Australian test site

The Australian test site was chosen in order to test the algorithm in meteorological and climatic conditions far from the Italian test sites. The soil moisture ground measurements are available through the website: http://www.ipf.tuwien.ac.at/insitu/index.php/in-situ-networks. html, which collects and shares out ground truth information in test sites distributed world-wide, which are connected through the *International Soil Moisture Network* (Dorigo et al., 2011).

The test site is part of the OzNet network (http://www.oznet.org.au/), which is located in the south-eastern part of the Australian continent. A 60 × 60 km^2 area to the south and west of the Yanco Research Station was considered for the algorithm test. The Yanco river basin is located in a partially semi-arid area of the Australian interior, west of Canberra, where the annual rainfall amounts to between 600 and 800 mm. The area is cropped with pastures and cereals. There are 37 soil moisture sites distributed across the Yanco study area with soil moisture data recorded since 2003 or, in some cases since 2009. The ground data available are:

- Soil moisture measurements (with an interval of 30 min at an integrated depth of 30 cm)
- Soil temperature (with an interval of 20 min)
- Rainfall gauge with a tipping bucket.

As auxiliary information, the land cover-land use from Globcover was used to select the main classes of land-use for which SMC retrieval is not possible (Globcover Product Manual, 2008) (Bicheron et al.,

Table 2
ENVISAT/ASAR acquisitions over the Scrivia test site.

Dates	Product	Polarization	Swath/θ
November 7, 2003	APP	HH/HV	2/23°
April 30, 2004	APP	HH/HV	2/23°
June 4, 2004	APP	HH/HV	2/23°
October 22, 2004	IMS	VV	2/23°
November 26, 2004	APP	HH/HV	2/23°
May 30, 2005	APP	HH/HV	2/23°
September 26, 2008	IMS	VV	2/23°
April 24, 2009	IMS	VV	2/23°
May 29, 2009	IMS	VV	2/23°

Table 3
ENVISAT/ASAR acquisitions over the Matera test site.

Dates	Product	Polarization	Swath/θ
July 13, 2008	APS	HH/HV	2/23°
October 10, 2008	APS	HH/HV	2/23°
May 7, 2008	IMG	VV	2/23°
April 11, 2009	APS	HH/VV	3/26°

Table 4
ENVISAT/ASAR acquisitions over the Cordevole test site.

Dates	Product	Polarization	Swath/θ
June 14, 2004	IMS	VV	2/23°
July 19, 2004	IMS	VV	2/23°
August 23, 2004	IMS	VV	2/23°
September 27, 2004	IMS	VV	2/23°
November 1, 2004	IMS	VV	2/23°

Table 6
ENVISAT/ASAR acquisitions over the Australian test site.

Dates	Product	Polarization (track)	Landsat (path/row)
12 July, 2004	IMS	VV-I2 (302)	10 June 2004 (92/84)
5 September, 2005	IMS	VV-I2 (302)	17 Sept 2005 (92/84)
19 December, 2005	IMS	VV-I2 (302)	20 Nov 2005 (92/84)
16 February, 2009	APS	VV–VH-I1 (302)	31 Jan 2009 (92/84)

2008). The Globcover is a product derived from an automatic and regionally-tuned classification of a time series of MERIS FR mosaics. Its 22 land-cover global classes are defined with the UN Land Cover Classification System (LCCS). The product covered the period from December 2004 to June 2006.

For EO data from the ESA–EOLI catalog, four ENVISAT/ASAR images are available for this area and are listed in Table 6. Images acquired in different seasons were selected, in order to have a wide range of soil moisture and vegetation conditions. Unfortunately, all images are in VV polarization, which was found not to be an optimal polarization for soil moisture retrieval. VV polarization is, in fact, strongly attenuated by vegetation characterized by thin vertical stems and narrow long leaves, such as cereals and grass. Landsat images acquired in periods close to the ENVISAT passes were used as auxiliary information in order to take into account the effect of the vegetation.

3.6. Spanish test site

Another area was selected in Spain (Emporda) for additional tests and for a validation of the algorithm. Six in-situ measurement campaigns of gravimetric soil moisture, soil texture and plant water content were organized in 24 fields.

The 6 campaigns took place on the following dates: 10 May, 14 June, 23 August, and 27 September 2010, 25 February and 22 June 2011 (Table 7). The fields were chosen to account for different types of cultivations and soil types. The evolution of soil moisture was followed during the period from May 2010 to June 2011 and was in line with the seasonal trend, typical of the Mediterranean environment, with a decrease in soil moisture between May and August and an increase during the autumn and winter months. The weather during the period was monitored by several meteorological stations. The 2010 spring and summer were relatively dry, with very few rain events. This was reflected in the soil moisture measurements, which were all quite low. The month of February 2011 was also relatively dry, with few precipitations and higher temperatures than the mean ones for that period.

4. Algorithm validation

The ANN algorithm was validated on the six test areas described in the previous section.

4.1. Scrivia

The first test and validation of the proposed algorithm was carried out on the Scrivia test site, by considering the available ENVISAT/ASAR images (Table 2). Unfortunately, due to several user conflicts for

Northern Italy, few dual polarized images could be obtained, and several single polarized images were in VV pol., instead than in HH pol. The latter was demonstrated to be more suitable than the VV for estimating soil moisture, due to the less pronounced masking effects of vegetation.

In order to compare the C-band backscatter with data collected on ground, the ASAR images were calibrated, and then geocoded using a regional map of the site (scale 1:10,000) so that the correct areas, i.e. where the ground measurements were carried out, could be identified with the precision of one pixel. In this case, ground truth and backscattering measurements were averaged field by field over at least 200 pixels each, making any further filtering meaningless.

Fig. 7 represents the relationships obtained between the SMC measured and estimated by the ANN algorithm in the two configurations – (a) single polarization and (b) dual polarization – on the 23 "reference" fields, by considering the data collected from November 2003 to June 2004. The comparison was obtained by averaging both ground and backscattering data at field resolution.

The regression obtained for the "single-pol." algorithm was $R^2 = 0.78$, and RMSE = 2.75% SMC, while the "dual-pol." algorithm made possible a retrieval with $R^2 = 0.85$ and RMSE = 2.15% SMC.

It should be noted that the results presented for the Scrivia test site were obtained by using only the SAR backscattering as input for the algorithm, without any ancillary information from optical/land cover to correct for the vegetation and roughness effects. In fact, the different response of the SAR backscattering in different polarizations to the surface features can be used for identifying certain surface types and evaluating the effect of the vegetation and soil conditions on the SMC retrieval. The contribution of the cross-polarized backscattering to enhancing the retrieval accuracy is evident from Fig. 7a, which represents the results for the "single polarization" algorithm. We can see that the retrieval error is higher for the intermediate SMC values, which were collected in extremely dissimilar conditions of roughness and vegetation, while the low and high SMC, which were recorded in similar conditions of vegetation and roughness, had a better retrieval. The "dual-pol." algorithm, is able to overcome this problem, thus increasing the retrieval accuracy of intermediate SMC values as well (Fig. 7b).

The SMC maps obtained from the single-pol. images for the 2009 spring are presented in Fig. 8 and refer to the area surrounding Castelnuovo Scrivia. Because of the limited dimensions of the area represented and the target resolution for the Sentinel application (1 km), the validation of these maps was limited in this case to the SMC estimates and corresponding ground truth averaged over the test area, which are highlighted by red circles. The average values obtained are those listed in the table beside the map.

Table 5
SAR (RADARSAT2) acquisitions over the Alto Adige/South Tyrol test site.

Dates	Polarization	Swath/θ
May 1, 2010 June 3, 2010	HH-VV-HV-VH	42°–45°

Table 7
ENVISAT/ASAR acquisitions over the Spanish test site.

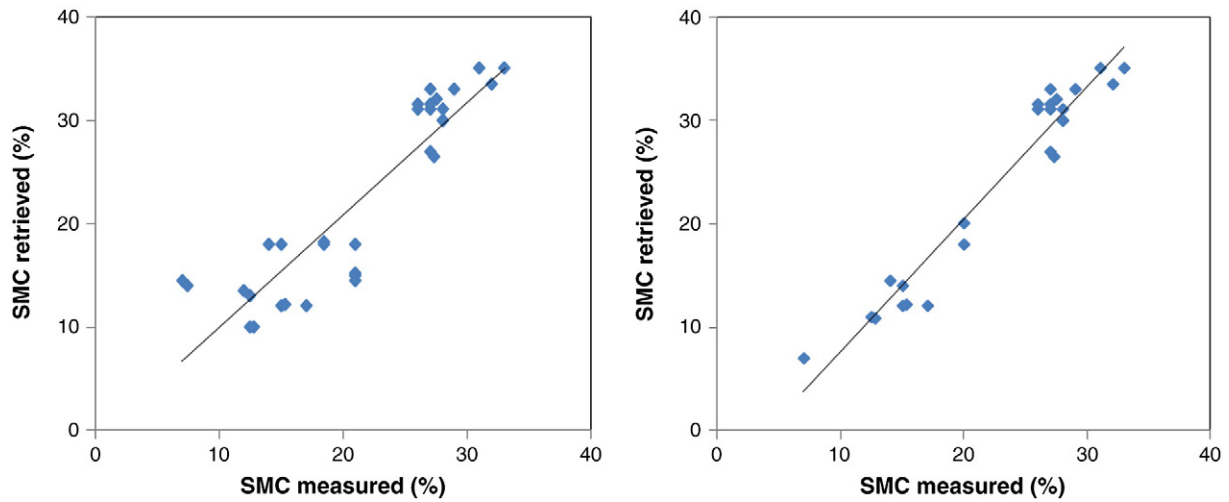| Dates	Product	Polarization	Track
10 May, 2010	WSM	VV	201
14 June, 2010	WSM	VV	201
23 August, 2010	WSM	VV	201
27 September, 2010	WSM	VV	201
25 February, 2011	WSM	VV	15
22 June, 2011	WSM	VV	403

Fig. 7. SMC retrieved vs. SMC measured on ground (left) "single-pol." algorithm and (right) "dual-pol." algorithm for the Scrivia test site. The computed regression equations are: SMCest = 1.013SMCmeas − 2.01 (R^2 = 0.78 and RMSE = 2.75% SMC) for the "single-pol." algorithm, and SMCest = 1.029SMCmeas − 1.69 (R^2 = 0.85 and RMSE = 2.15% of SMC) for the "dual-pol." algorithm.

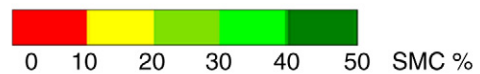| Date	SMC meas.	SMC ANN
24/04/09	~30%	~28%
29/05/09	~23%	~21%
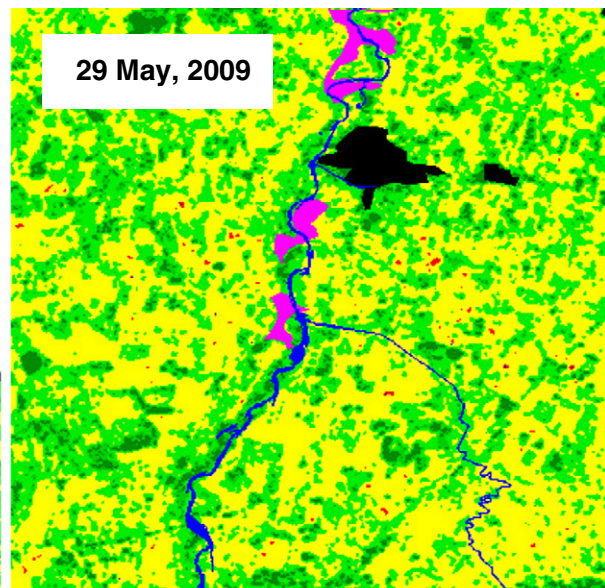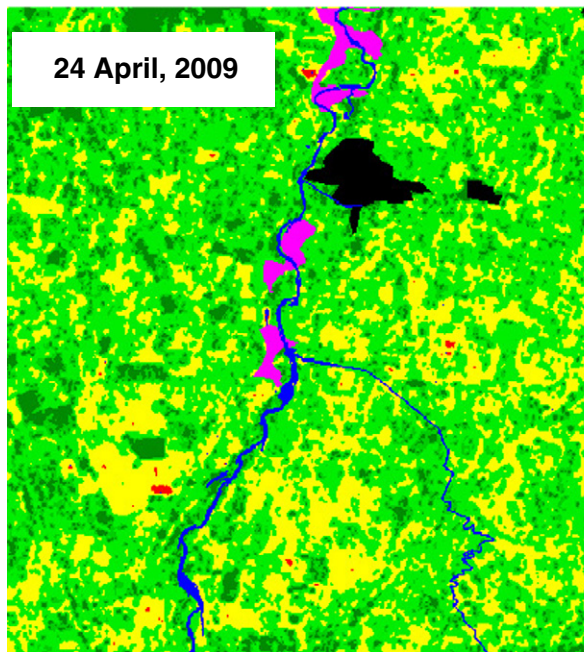
Fig. 8. SMC maps obtained from the single-pol. ENVISAT/ASAR images collected on the area of Castenuovo Scrivia in April and May 2009. Urban, water bodies, and dense vegetation areas have been masked in black, blue, and magenta, respectively. In the table on the left-top side of the image, the average values of SMC (measured on ground and estimated by the ANN algorithm) for the entire image were reported.

Table 8
SMC values estimated from the ANN with and without NDVI data, as compared with the average ground measurements for the Matera test site.

Date	Mean NDVI	SMC ANN w NDVI	SMC ANN w/o NDVI	Ground measurements
May 2008	0.55	15.0%	20.5%	13%
July 2008	0.27	7.9%	14.5%	7%
October 2008	0.29	27%	40.8%	25%
April 2009	0.50	22.5%	22.6%	25%

4.2. Matera

The algorithm test and validation were subsequently extended to the other test sites, in order to confirm the validity of the algorithm in areas characterized by different climatic and vegetation conditions. The methodology was first tested on the agricultural test site of Matera, using 4 available ENVISAT/ASAR images and the corresponding averaged ground measurements. The results, which are indicated in Table 8, were then further compared with the values of SMC obtained by considering the ANN for bare surfaces, without accounting for vegetation. The results indicate a notable improvement in the estimate when the NDVI was available. The values indicated in the table were averaged over the area pointed out by red circle in Fig. 9. A statistical analysis of the results gave a RMSE = 1.8%SMC and Bias = 0.75%SMC, for ANN with NDVI, and RMSE = 9%SMC and Bias = 7.25%SMC for ANN without NDVI.

Extensive SMC maps of the agricultural area surrounding Matera were generated by the algorithm and were represented in Fig. 9. The
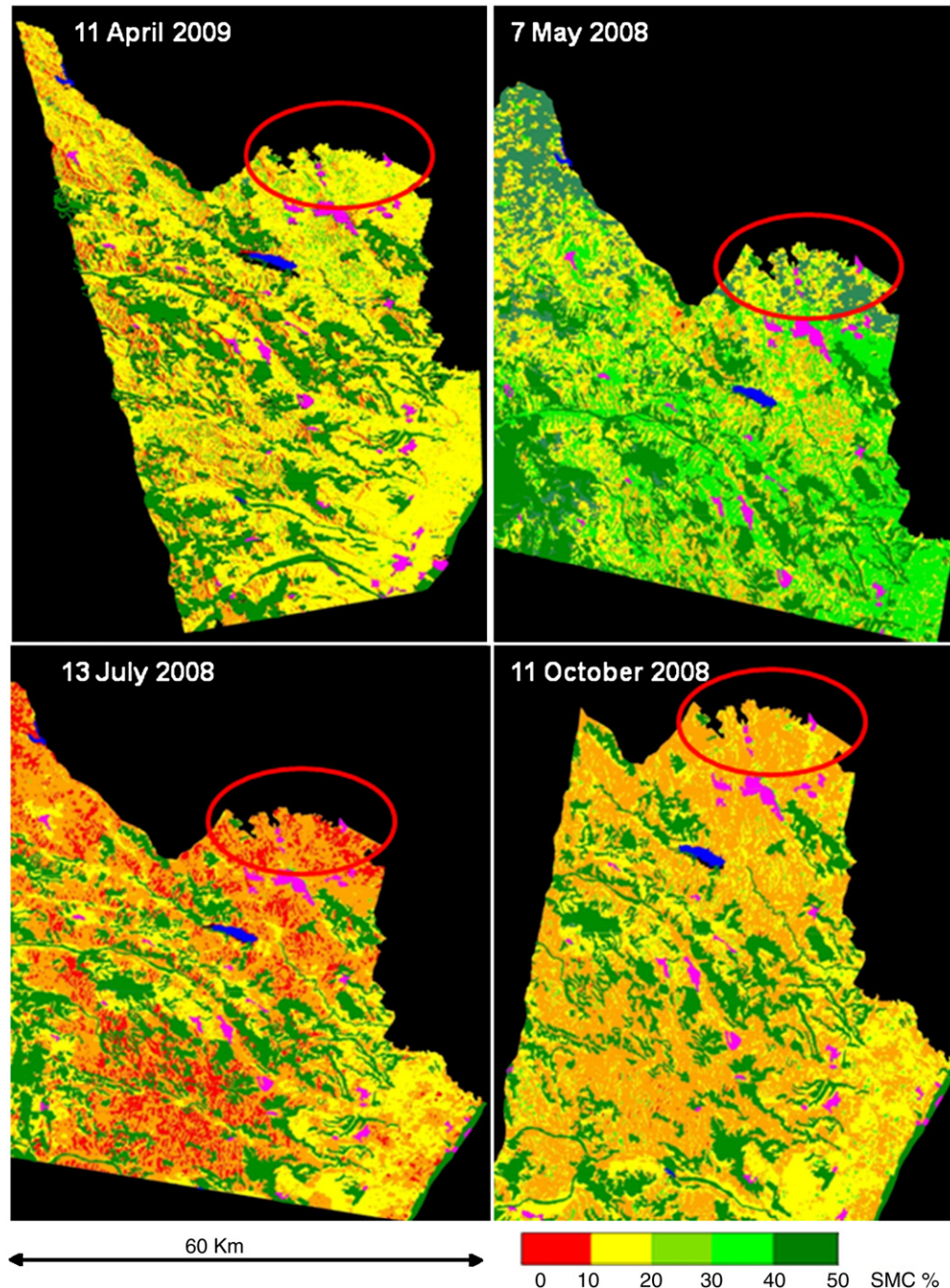

Fig. 9. SMC maps derived from the ENVISAT/ASAR images available for the Matera test area. Red circles identify the study area. Urban areas, water bodies and forests have been masked in magenta, blue, and dark green, respectively.
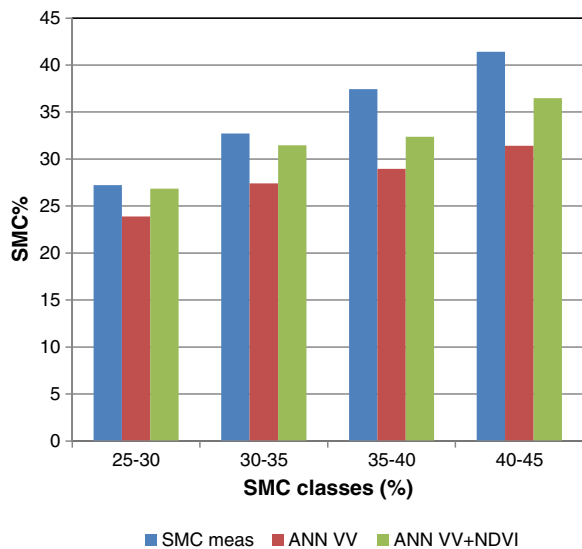
Fig. 10. Histogram of the comparison between the ANN-estimated and the ground-measured SMC over the Cordevole area. SMC was subdivided into 4 classes between 25% and 45%. Two ANNs were considered: one used only VV polarized data, and the other one also used the NDVI information.

values of estimated SMC are in line with the meteorological conditions of the area and with the typical Mediterranean climate, which is usually characterized by very rainy springs and dry summers.

4.3. Cordevole

A further test of the algorithm was attempted on the mountainous test area of the "Cordevole" basin, an area that is characterized by a completely different surface cover, in order to better evaluate its capabilities in adapting to the different vegetation conditions. The algorithm was applied to the three available ENVISAT/ASAR images of the Cordevole watershed (about 7 km × 15 km), which were collected on 14 June, 19 July and 27 September 2004. Once again, backscattering values were extracted and compared with ground measurements of SMC. The effect of topography was accounted for using a LIA (Local Incidence Angle) map as an input of the ANN. Data corresponding to incidence angles outside the training range (i.e. 20–50°) were disregarded.
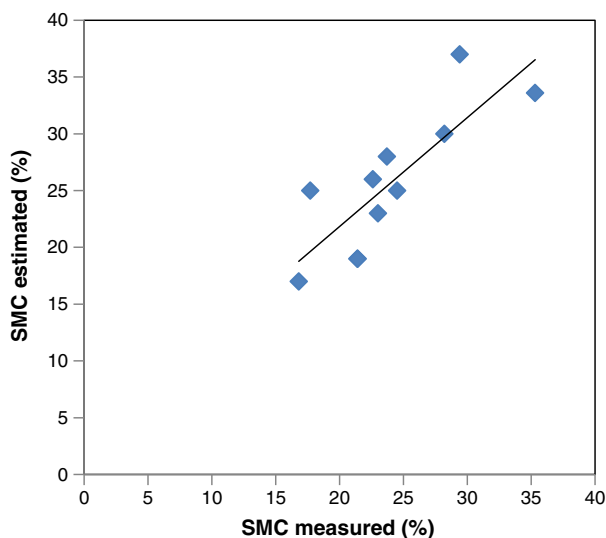
Moreover, areas showing layover and shadowing have been masked out before the processing. Unfortunately, the fairly constant SMC values recorded during the three field campaigns did not enable us to obtain a wide temporal variation, because the mean value of the measurements carried out for each campaign ranged between 25% and 30% for September and the 45% for July. Therefore, the algorithm output was represented after dividing the range of SMC measurements into 4 classes of between 25% and 45% in steps of 5%. A comparison between the ANN-estimated and the ground-measured SMC is shown in the histogram of Fig. 10. Two ANNs were considered: one used only VV polarized data, while the second one also used the vegetation information derived from NDVI. The ANN using also vegetation information was more capable of reproducing the level of the data, even if with a few underestimations.

4.4. Alto Adige/South Tyrol

The algorithm was applied also to the available images collected from RADARSAT2 and acquired on 3 June 2010 over a mountain test site in the South Tyrol region (Northern Italy). In this case, the average of the ground measurements was 27.7% and the corresponding average of the algorithm outputs was 26.3%. Also in this case, the LIA map was considered for compensating the topography effect and areas showing layover and shadowing were masked out before the processing. Fig. 11 represents a scatterplot of the comparison between the averaged values of the ground measurements and the corresponding algorithm estimates. In this case, $R^2 = 0.67$, and RMSE = 4.91%SMC, and Bias = 1.55%SMC.

4.5. Australia

The fourth validation was carried out on the Yanco River test areas in Australia. If we examine the rainfall and soil moisture data measured at the meteorological stations in the area during the winters of 2004 and 2005, and the summer of 2005, the highest values of soil moisture were found in September. The entire area shows very high NDVI values (>0.4) in almost all seasons, except July (dry season). The availability of VV polarization only and the high NDVI values limited the performances of the algorithm. Moreover, the presence of wheat or other cereals produced lower values of σ°, as compared with the expected ones, when the plants were well-developed.

Fig. 11. Algorithm outputs vs ground measurements for the Alto Adige test area. The regression equation is: SMCest = 0.88SMCmeas + 4.72 (R^2 = 0.67, RMSE = 4.91%SMC, and Bias = 1.55%SMC).
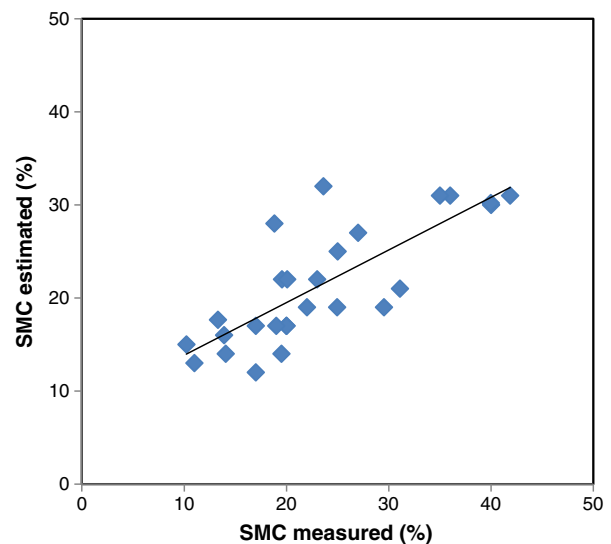

Fig. 12. Soil moisture measured on ground as a function of soil moisture estimated by the algorithm for the Australian test area. The regression equation is: SMCest = 0.57SMCmeas + 8.17 (R^2 = 0.62, RMSE = 5.58%SMC, and Bias = −1.85%SMC).

Fig. 12 contains a scatterplot of the soil moisture measured on ground at various meteorological stations, in different seasons, vs. the soil moisture estimated by the algorithm. From the diagram it is evident that, in spite of the non optimal observation conditions (VV pol. and high NDVI), the algorithm was able to retrieve the soil moisture values with a fairly good accuracy ($R^2 > 0.6$ and RMSE = 5.58%). The algorithm underestimated the soil moisture values but this was because

the SMC was measured on ground in an integrated layer of 30 cm, which is usually moister than the surface layer.

The corresponding soil moisture maps of the area obtained by using this algorithm are shown in Fig. 13: a) July 12, 2004; b) September 5, 2005; c) December 19, 2005; and d) February 16, 2009. A significant moisture variation depending on the season, the rainfall and the ground measurements, is evident from the maps. As expected, the February

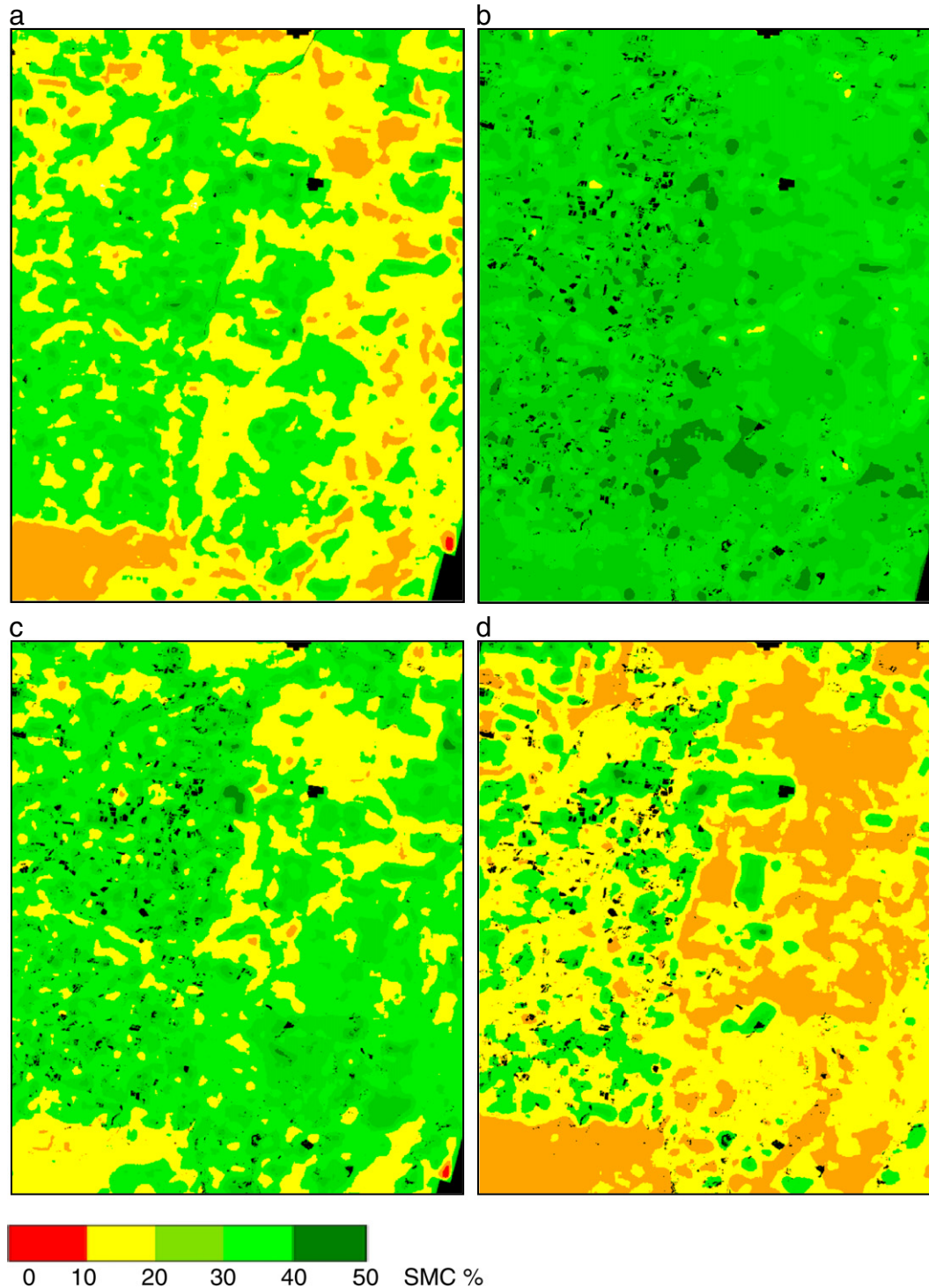**Fig. 13.** Soil moisture maps of the area obtained by using this algorithm (a: July 12, 2004; b: September 5, 2005; c: December 19, 2005; d: February 16, 2009). Black spots represent masked areas: i.e. dense vegetation, water bodies, urban areas.

Table 9
Validation results, obtained over the test site in Spain, by considering (a) or disregarding (b) NDVI information. In gray, the couples of values SMC estimated/SMC measured, for which the NDVI information worsened the results, have been singled out.

a

Acquisition date	SMC ANN (%)	SMC ground (%)	RMSE (%)	St. Dev. (%)
10 May, 2010	16.11	24.05	5.34	3.45
14 June, 2010	12.66	14.35	3.52	2.66
23 August, 2010	10.18	7.86	3.41	2.19
27 September, 2010	17.19	14.99	6.38	4.76
25 February, 2011	13.74	21.41	4.58	0.69
22 June, 2011	11.58	10.85	4.35	3.44
Mean			4.60	2.87

b

Acquisition date	SMC ANN (%)	SMC ground (%)	RMSE (%)	St. Dev. (%)
10 May, 2010	20.06	24.05	2.72	0.97
14 June, 2010	16.37	14.34	4.01	2.69
23 August, 2010	16.65	7.85	8.50	2.90
27 September, 2010	17.92	14.99	5.12	3.23
25 February, 2011	20.31	21.40	2.31	0.58
22 June, 2011	16.39	10.85	5.21	1.71
Mean			4.64	2.01

map, which corresponds to the end of the dry season, shows the lowest values of soil moisture, whereas September shows the highest ones. Pasture areas, which are spread along the banks of rivers, show the lowest values of soil moisture as compared to those of the agricultural areas.

4.6. Spain

A final validation was subsequently carried out between 2010 and 2011 in the test area of Emporda, in Spain. Six in-situ campaigns were organized in 24 fields, where ground data of SMC, soil texture and plant water content were collected. Six corresponding ENVISAT/ASAR images, in VV pol., were collected over the area, while the corresponding NDVI maps were derived from MODIS.

Table 9a summarizes the results obtained for this test area. It should be noted that, although the RMSE was in line with the results of the other test areas (i.e. <5%SMC), the availability of VV polarization only and the low resolution of NDVI maps (500 × 500 m^2) slightly hampered the retrieval accuracy. Indeed, the fields investigated are often smaller than 1 pixel in the NDVI maps and are surrounded by forested areas. The NDVI associated with each field also accounts for the surrounding forest, and therefore cannot be considered representative of the true field conditions. These inaccurate NDVI values influenced the SMC retrieval, especially during the spring and autumn, when the vegetation cover is extremely variable in the area. Moreover, the average NDVI values in February were extremely high (0.6–0.9), and were inconsistent with the bare soil conditions of most of the fields.

The tests were therefore repeated, with the NDVI being disregarded: the results are summarized in Table 9b. A comparison of the results between Tables 9a and b points out that, without NDVI, the accuracy improved in February and also in the autumn and spring, when the low resolution NDVI maps could not be representative of the high spatial variability of vegetation conditions. On the other hand, in the summer, when the vegetation was well-developed, and NDVI was consequently

almost uniform, the low spatial resolution was less critical and the use of NDVI made possible an improvement in the retrieval accuracy.

4.7. Summary of validation results

By considering all the ENVISAT/ASAR images available on the 6 test areas and the ground truth measurements, it was possible to set up a dataset of about 600 field-averaged values of backscattering and the corresponding SMC measured on ground. Depending on the available input data, the algorithm performed a selection of the most appropriate ANN among the ones developed and described in Section 2. Six different configurations were defined as follows:

1. VV polarization only, without NDVI
2. HH polarization only, without NDVI
3. VV polarization and NDVI
4. HH polarization and NDVI
5. VV and VH polarizations
6. HH and HV polarizations

The results obtained are presented in the following diagrams of Fig. 14(a–d). As expected, the VV polarization gave the poorest results, while ancillary information on NDVI, when available, and made possible an increase in the retrieval accuracy. Unfortunately, most available images were obtained in this configuration, due to ESA's ordering policy. The HH polarization was more correlated to the SMC, and also in this case, the use of NDVI increased the retrieval accuracy. The most favorable results were obtained when co- and cross-polarizations were available.

More in detail, Fig. 14 represents the SMC estimated by the algorithm as a function of the SMC measured on ground, at field scale, for the data available in VV polarization, with and without the NDVI. The latter input made possible an increase in the determination coefficient R^2 from 0.495 to 0.65 (see Table 10, where the statistical parameters obtained from this comparison – R^2, RMSE, and Bias – are listed for each ANN).

Fig. 14b represents the results obtained using the HH polarization and the HH polarization + NDVI. The HH polarization appeared to be more indicated for SMC monitoring purposes, and the NDVI contribution increases the retrieval accuracy also in terms of the regression slope from 0.78 to 0.88, whereas R^2 slightly improves from 0.85 to 0.86.

The case of the VV and VH polarizations is shown in Fig. 14c. The SMC estimates obtained by using the dual-pol. ANN are compared with those generated by the ANN for single-pol. and single-pol. + NDVI. Unfortunately, only very few data were available for this configuration and for a limited range of SMC measured on ground. However, the accuracy improvement produced by the simultaneous use of two polarizations is evident: R^2 changes from very negligible values up to as much as 0.59 (VV + VH polarizations).

Finally, the case of HH combined with d HV polarization is presented in Fig. 14d. R^2 changed from 0.80 for only HH pol. to 0.87 for HH + VH polarizations. As expected, this was the combination that provided the best performances, but satisfactory results were also obtained by using the two ANNs for HH polarization only and HH polarization + NDVI, respectively.

5. Conclusions

The SMC product, to be generated from Sentinel-1 data, requires an algorithm capable of processing operationally in near-real-time and of delivering the product to the GMES services within 3 h from the moment of observations. The proposed approach is based on an Artificial Neural Network (ANN) that was chosen as it represents a good compromise between retrieval accuracy and processing time, thus allowing compliance with the timeliness requirements. The algorithm was developed so as to ingest different inputs based on the possible configurations of the Sentinel-1 system.
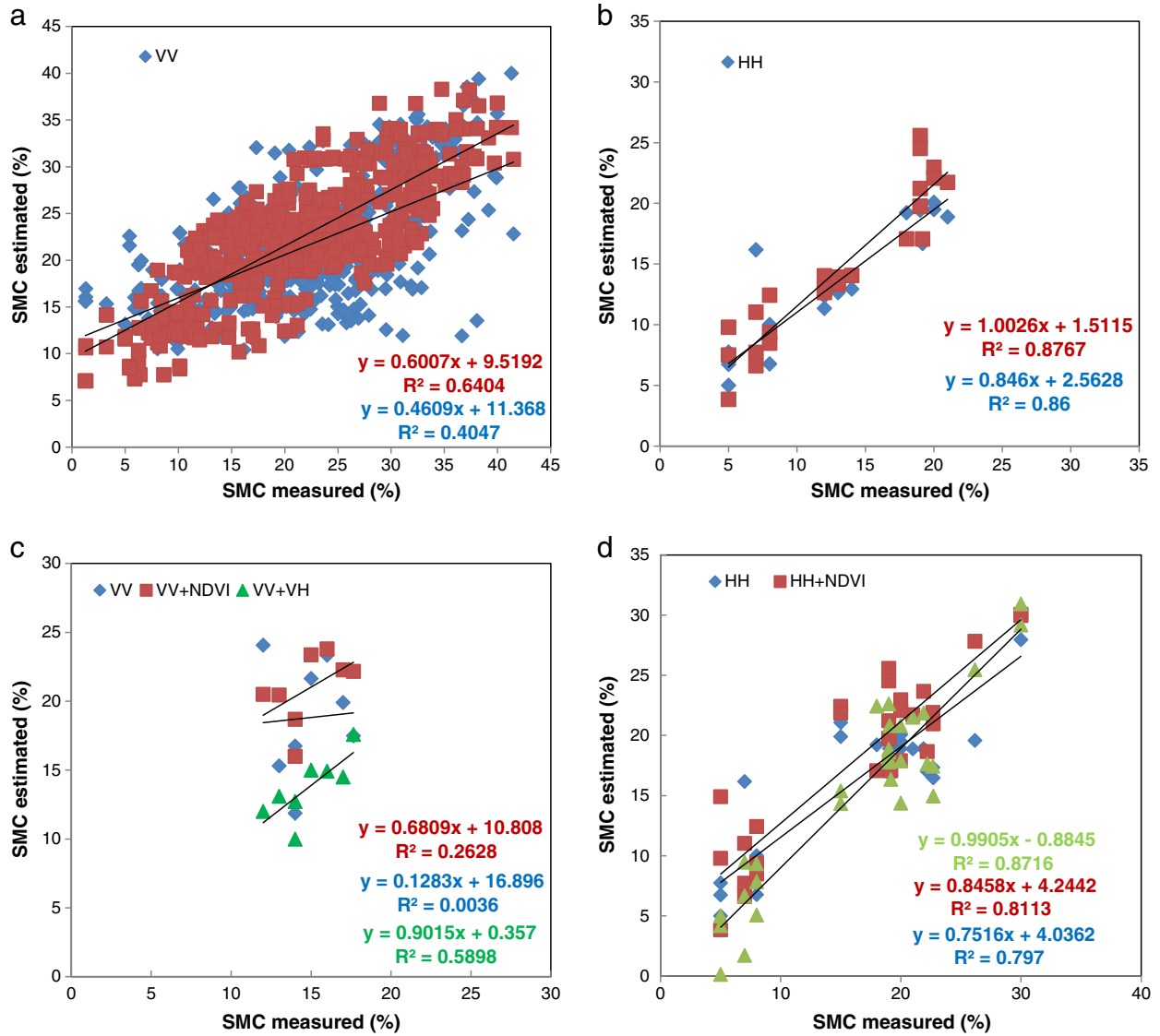
Fig. 14. a: SMC estimated by the algorithm as a function of the ground truth, in cases of VV polarization only, and VV pol. + NDVI. b: SMC estimated by the algorithm as a function of the ground truth in cases of HH polarization only and HH pol. + NDVI. c: SMC estimated by the VV + VH algorithm as a function of the ground truth. d: SMC estimated by the algorithm for the case of HH + HV, as a function of the ground truth.

Depending on the available input data, the algorithm performs a selection of the most appropriate ANN, among 6 different configurations: 1) VV polarization with and without NDVI, 2) HH polarization with and without NDVI, 3) VV polarization with NDVI, 4) HH polarization with NDVI, 5) VV and VH polarizations, and 6) HH and HV polarizations.

The ANN algorithm was validated in six test areas: Alessandria, Cordevole, Matera and Alto Adige in Italy, Australia, and Spain.

The retrieval accuracy for volumetric SMC, according to GMES S1 requirements, was ≤5%, and this was fulfilled by most of the SMC estimated values. However, the validation results obtained in Spain were penalized by the availability of only VV polarization SAR images and MODIS low-resolution NDVI, although the RMSE was in any case slightly >4%. The accuracy (RMSE) of the algorithm ranges indeed from around 2% of SMC, when even HV polarization is available, to 6% of SMC in the worst case, when only VV polarization is present. As far the processing time is concerned, the ANN algorithm makes a rapid inversion possible, in the order of the timeliness required (i.e. 3 h from observation). The inversion is carried out for each pixel of input image, and the algorithm is able to process about 200,000 pixels/s. For an input SAR image of 100×100 km^2 at 25 m resolution, the whole process takes about 80 s to generate the corresponding SMC map at the same spatial resolution. If a lower spatial resolution is required, resampling of the input backscattering can be performed in order to further reduce the processing time.

Table 10
Statistical parameters (R^2, RMSE and Bias) of the SMC retrieval by using different ANN, depending on to the available polarizations and ancillary information.

ANN	R^2	RMSE	Bias
Single polarization (HH or VV)			
VV	0.495	6.68	−0.14
VV + NDVI	0.650	5.27	0.99
HH	0.854	2.60	−0.33
HH + NDVI	0.881	2.32	1.06
Dual polarization (VV + VH or HH + HV)			
VV	0.004	6.08	5.40
VV + NDVI	0.011	5.47	3.53
VV + VH	0.590	1.67	−0.98
HH	0.797	3.15	1.57
HH + NDVI	0.811	3.44	−0.10
HH + HV	0.872	2.83	−0.98

Considering that the algorithm needs to act upon global conditions, the best performances are obtained when information on vegetation is available, in terms of either cross-polarized channel or NDVI information. The cross-polarized channel (e.g. HV) helps to disentangle the effect of vegetation. If optical images are available, NDVI information is able to improve the retrieval performances. It is worthwhile mentioning that NDVI can be considered as a key factor in determining a reliable SMC estimate, when only VV polarization is available. In this context, the forthcoming combination of Sentinel 1 and 2 satellites represents an ideal case to have reliable estimates of SMC at high temporal and spatial resolution.

Acknowledgments

References

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

Baghdadi, N., Cresson, R., El Hajj, M., Ludwig, R., & La Jeunesse, I. (2012). Estimation of soil parameters over bare agriculture areas from C-band polarimetric SAR data using neural networks. *Hydrology and Earth System Sciences, 16*, 1607–1621.

Balenzano, A., Satalino, G., Pauwels, V., & Mattia, F. (2011). Soil moisture retrieval from dense temporal series of C-band SAR data over agricultural sites. *Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, 3136–3139, Vancouver, Canada, 24–29 July, 2011.*

Baronti, S., Del Frate, F., Ferrazzoli, P., Paloscia, S., Pampaloni, P., & Schiavon, G. (1995). SAR polarimetric features of agricultural areas. *International Journal of Remote Sensing, 16*(14), 2639–2656.

Bayes, T. (1958). An essay towards solving a problem in the doctrine of chances. *Biometrika, 45*, 296–315.

Beaudoin, A., Le Toan, T., & Gwyn, Q. H. (1990). SAR observations and modeling of the C-band backscatter variability due to multiscale geometry and moisture. *IEEE Transactions on Geoscience and Remote Sensing, 28*(5), 886–895.

Benallegue, M., Taconet, O., Vidal-Madjar, D., & Normand, M. (1995). The use of radar backscattering signals for measuring soil moisture and surface roughness. *Remote Sensing of Environment, 53*, 61–68.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., et al. (2008). GLOBCOVER: Products Description and Validation Report, https://globcover.s3.amazonaws.com/LandCover_V2.2/GLOBCOVER_Products_Description_Validation_Report_I2.1.pdf (©MEDIAS-France 2008)

Bindlish, R., & Barros, A. P. (2001). Parameterization of vegetation backscatter in radar-based, soil moisture estimation. *Remote Sensing of Environment, 76*, 130–137.

Brogioni, M., Pettinato, S., Macelloni, G., Paloscia, S., Pampaloni, P., Pierdicca, N., et al. (2010). Sensitivity of bistatic scattering to soil moisture and surface roughness of bare soils. *International Journal of Remote Sensing, 31*(15), 4227–4255.

Dobson, M. C., Ulaby, F. T., Hallikainen, M. T., & El-Rayes, M. A. (1985). Microwave dielectric behavior of wet soil — Part II — Dielectric mixing models. *IEEE Transactions on Geoscience and Remote Sensing, 23*(1), 35–46.

Dorigo, W. A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., et al. (2011). The international soil moisture network: A data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences, 15*(5), 1675–1698.

Doubková, M., Albert, I. J. M., Dijk, Van, Sabel, Daniel, Wagner, Wolfgang, & Blöschl, Günter (2012). Evaluation of the predicted error of the soil moisture retrieval from C-band SAR by comparison against modelled soil moisture estimates over Australia. *Remote Sensing of Environment, 120*, 188–196.

Dubois, P., Van Zyl, J., & Engman, T. (1995). Measuring soil moisture with imaging radars. *IEEE Transactions on Geoscience and Remote Sensing, 33*(4), 915–926.

Entekhabi, D., & Eagleson, P. S. (1989). Land surface hydrology parameterization for atmospheric general circulation models including subgrid scale spatial variability. *Journal of Climate, 2*, 816–831.

Entekhabi, D., Nakamura, H., & Njoku, E. (1994). Solving the inverse problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations. *IEEE Transactions on Geoscience and Remote Sensing, 32*(2), 438–448.

Famiglietti, J. S., & Wood, E. F. (1994). Multiscale modeling of spatially variable water and energy balance processes. *Water Resources Research, 30*(11), 3061–3078.

Fung, A. K. (1994). *Microwave scattering and emission models and their applications.* Norwood, MA: Artech House.

Hornacek, M., Wagner, W., Sabel, D., Truong, H. -L., Snoeij, P., Hahmann, T., Diedrich, E., & Doubkova, M. (2012). Potential for high resolution systematic global surface soil moisture retrieval via change detection using Sentinel-1. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 5*(4), 1303–1311.

Hornik, K. (1989). Multilayer feed forward network are universal approximators. *Neural Networks, 2*(5), 359–366.

Jackson, T. J. (1993). Measuring surface soil moisture using passive microwave remote sensing. *Hydrological Processes, 7 (2).* (pp. 139–152): John Wiley & Sons, Ltd (Copyright © 1993).

Jackson, T. J., Chen, D., Cosh, M., Li, F., Anderson, M., Walthall, C., et al. (2004). Vegetation water content mapping using Landsat data derived normalized difference water index for corn and soybeans. *Remote Sensing of Environment, 92*, 475–482.

Joseph, A. T., van der Velde, R., O'Neill, P. E., Lang, R., & Gish, T. (2010). Effects of corn on C- and L-band radar backscatter: A correction method for soil moisture retrieval. *Remote Sensing of Environment, 114*(11), 2417–2430.

Linden, A., & Kindermann, J. (1989). *Proc. Int. Joint Conf. Neural Networks. Inversion of multi-layer nets, Vol. II.* (pp. 425–443).

Macelloni, G., Paloscia, S., Pampaloni, P., Ruisi, R., Dechambre, M., Valentin, R., et al. (2002). Modelling radar backscatter from crops during the growth cycle. *Agronomie, 22*, 575–579.

Macelloni, G., Paloscia, S., Pampaloni, P., Sigismondi, S., de Matthæis, P., Ferrazzoli, P., et al. (1999). The SIR-C/X-SAR experiment on Montespertoli: sensitivity to hydrological parameters. *International Journal of Remote Sensing, 20*(13), 2597–2612.

Mattia, F., Davidson, M. W. J., Le Toan, T., D'Haese, C. M. F., Verhoest, N. E. C., Gatti, A. M., et al. (2003). A comparison between soil roughness statistics used in surface scattering models derived from mechanical and laser profilers. *IEEE Transactions on Geoscience and Remote Sensing, 41*(7), 1659–1671.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal, 7*, 308–313.

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.

Oh, Y., Sarabandi, K., & Ulaby, F. T. (2002). Semi-empirical model of the ensemble-averaged differential Mueller matrix for microwave backscattering from bare soil surfaces. *IEEE Transactions on Geoscience and Remote Sensing, 40*(6), 1348–1355.

Paloscia, S. (2002). A summary of experimental results to assess the contribution of SAR for mapping vegetation biomass and soil moisture. *Canadian Journal of Remote Sensing, 28*(2), 246–261.

Paloscia, S., Macelloni, G., Pampaloni, P., & Santi, E. (2004). The contribution of multi-temporal SAR data in assessing hydrological parameters. *IEEE Geoscience and Remote Sensing Letters, 1*(3), 201–205.

Paloscia, S., Pampaloni, P., Pettinato, S., & Santi, E. (2008). A comparison of algorithms for retrieving soil moisture from ENVISAT/ASAR Images. *IEEE Transactions on Geoscience and Remote Sensing, 46*(10), 3274–3284.

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

Pasolli, L., Notarnicola, C., Bruzzone, L., Bertoldi, G., Della Chiesa, S., Hell, V., et al. (2011). Estimation of soil moisture in an Alpine catchment with RADARSAT2 images. *Applied and Environmental Soil Science*: Hindawi Publishing Corporation. http://dx.doi.org/10.1155/2011/175473 (2011, Article ID 175473, 12 pages).

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

Pierdicca, N., Castracane, P., & Pulvirenti, L. (2008). Inversion of electromagnetic models for bare soil parameter estimation from multifrequency polarimetric SAR data. *Sensors, 8*, 8181–8200.

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

Shi, J. C., van Zyl, J., Soares, J. V., & Engman, E. T. (1992). Development of soil moisture retrieval algorithm for L-band SAR measurements. *Proceedings of IEEE Geoscience and Remote Sensing Symposium IGARSS'92* (pp. 495–497).

Shi, J. C., Wang, J., Hsu, A. Y., O'Neill, P. E., & Engman, E. T. (1997). Estimation of bare surface soil moisture and surface roughness parameter using L-band SAR image data. *IEEE Transactions on Geoscience and Remote Sensing, 35*, 1254–1266 (September 1997).

van der Velde, R., Su, Z., van Oevelen, P., Wen, J., Ma, Y., & Salama, Mhd. S. (2012). Soil moisture mapping over the central part of the Tibetan Plateau using a series of ASAR WS images Original Research Article. *Remote Sensing of Environment, 120*, 175–187.

Wagner, W., Lemoine, G., Borgeaud, M., & Rott, H. (1999b). A study of vegetation cover effects on ERS scatterometer data. *IEEE Transactions on Geoscience and Remote Sensing, 37*(2), 938–948.

Wagner, W., Lemoine, G., & Rott, H. (1999c). A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sensing of Environment, 70*, 191–207.

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

Wu, T. D., & Chen, K. S. (2004). A reappraisal of the validity of the IEM model for backscattering from rough surfaces. *IEEE Transactions on Geoscience and Remote Sensing, 42*(4), 743–753.

Zribi, M., Chahbi, A., Shabou, M., Lili-Chabaane, Z., Duchemin, B., Baghdadi, N., et al. (2011). Soil surface moisture estimation over a semi-arid region using ENVISAT ASAR radar data for soil evaporation evaluation. *Hydrology and Earth System Sciences, 15*, 345–358.