Remote Sensing of Environment 134 (2013) 234–248

Contents lists available at SciVerse ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

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

Article history:
Received 21 September 2012
Received in revised form 11 January 2013
Accepted 28 February 2013
Available online 9 April 2013

Keywords:
Sentinel-1
Backscattering coefficient
Soil moisture maps
Inversion algorithms
Artificial Neural Network

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.

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

* Corresponding author. Tel.: +39 0555226494.
E-mail address: S.Paloscia@ifac.cnr.it (S. Paloscia).

0034-4257/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.rse.2013.02.027

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, $\sigma^0{}_{dry}$ 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 $\sigma^0{}_{dry}$ 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 ($\leq$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: