SMOSAR ALGORITHM FOR SOIL MOISTURE RETRIEVAL USING SENTINEL-1 DATA

Anna Balenzano (1), Francesco Mattia (1), Giuseppe Satalino (1), Valentijn Pauwels (2) and Paul Snoeij (3)

(1) Consiglio Nazionale delle Ricerche (CNR) – Istituto di Studi sui Sistemi Intelligenti per l’Automazione (ISSIA), Bari, Italy
(2) Ghent University – Laboratory of Hydrology and Water Management (LHWM), Ghent, Belgium
(3) European Space Agency (ESA) – European Space Research and Technology Centre (ESTEC), Noordwijk, The Netherlands

ABSTRACT
This paper describes and assesses the quality of the algorithm, “Soil MOisture retrieval from multi-temporal SAR data” (SMOSAR), developed in view of the forthcoming European Space Agency (ESA) Sentinel-1 (S-1) mission. SMOSAR retrieves soil moisture \( m \) products at high spatial resolution (i.e. less than 1km) from dense time series of either single (i.e. HH or VV) or dual polarized (i.e. HH+HV or VV+VH) S-1 data. The assessment of the algorithm performance is based on time series of ENVISAT/ASAR data collected over the DEMMIN site (Germany) in 2006 and over the Matera site (Italy) in 2003 and 2005 and RADARSAT-2 data acquired over the Flevoland site (The Netherlands) in 2009. Results indicate that \( m \) can be retrieved with an accuracy of 5% at HH polarization, whereas at VV polarization more investigations are required to provide reliable figure for the retrievable accuracy.

Index Terms— Sentinel-1, soil moisture, retrieval algorithm, Synthetic Aperture Radar (SAR), hydrology.

1. INTRODUCTION

The forthcoming European Space Agency (ESA) Sentinel-1 (S-1) C-band SAR constellation will provide continuous all-weather day/night global coverage, with six days exact repetition time (near daily coverage over Europe and Canada) and with radar data delivery within 3 hours [1]. These features open new possibilities for the monitoring of the soil moisture content \( m \) at high spatial (100-1000m) and temporal resolution, with important consequences on operational land services.

Recently, the \( m \) retrieval algorithm, “Soil MOisture retrieval from multi-temporal SAR data” (SMOSAR) intended to use the S-1 Interferometric Wide Swath image data, has been developed in the framework of the ESA project “GMES Sentinel-1 soil moisture algorithm development” [2], which was aimed at developing global near-real-time and off-line \( m \) products at high spatial resolution. The strength of the SMOSAR algorithm lies on its capability of monitoring \( m \) of agricultural soils throughout the growing season, at least underneath those crops not dominated by volume scattering (e.g. cereal crops). This is accomplished by inverting temporal changes of radar backscatter rather than absolute backscatter values [3]. Indeed, under the hypothesis that the time series of N SAR images is dense (revisiting time 6-12 days) the backscatter change between two subsequent acquisitions is mainly related to the temporal change of \( m \), which is characterized by a temporal scale of few days, and poorly related to the temporal changes of vegetation biomass or crop canopy or surface roughness, which also affect the radar response but are usually characterized by a temporal scale of a few weeks. In SMOSAR, the approximation that backscatter changes over a short time interval depend only on \( m \) changes, is made. This approach simplifies the \( m \) retrieval and allows strongly reducing the computation time, as required by the timeliness constraints of a near-real time product.

In this context, the objective of this paper is to present the implemented SMOSAR prototype and its assessment, which has focused on investigating the extension to which SMOSAR fulfills the performance goals in terms of accuracy and spatial resolution based on available satellite C-band SAR data.

2. SAR AND GROUND DATA

The algorithm has been assessed using the rich and well-documented data sets collected during the AgriSAR 2006 campaign over the Görmin farm at the DEMMIN site (Germany) [4], the AgriSAR 2009 campaign over Flevoland (The Netherlands) site [5], and the Italian ENVISAT 2003 and 2005 campaigns over Matera site (Italy) [6]. The three
selected sites are characterized by significantly different agronomical and topographical features and locations. The SAR data acquired over the Görmin site consists of 12 descending ASAR images in alternating polarization (12 images at HH and 3 images at VV) acquired at mean incidence angle between 20° and 34° and at a mean temporal resolution of 7 days from May to July 2006. The AgriSAR 2009 data set over Flevoland includes HH/HV and VV/VH Sentinel-1 (S-1) Interferometric Wide Mode products simulated by 16 ascending RADARSAT-2 data at FQ5, FQ9, FQ13 modes (from 24° to 33° mean incidence angle) and at a mean temporal resolution equal to 8 days from June to September. The SAR data set over Matera encompasses 20 descending ASAR images in alternating polarization (12 images at HH and 20 images at VV) acquired at incidence angle between 23° and 38° and at a mean temporal resolution of 7 days from September 2003 to December 2005. It is worth noting that the ASAR time series acquired over Görmin and Matera contain several temporal gaps, do not always include both HH and VV polarizations (e.g. an important lack of VV time series is faced over the Görmin site). As a consequence, in case of Matera and Görmin, groups of N=4 SAR images have been processed separately.

For the validation of the $m_c$ products, spatially and temporally distributed $m_c$ data are needed. The $m_c$ observations used for the SMOSAR validation have been provided by a hydrological model [7], calibrated using the in situ measurements. This is due to the fact that the in situ measurements are available only at specific dates not in coincidence with the SAR acquisitions and/or at a restricted number of locations for the selected test sites. The advantage of this approach is the possibility to validate the $m_c$ product over quite large areas and at various spatial resolutions, extending the comparison between retrieved versus the “observed” $m_c$ values well beyond the limited number of point-scale in situ measurements.

3. SMOSAR ALGORITHM

The SMOSAR algorithm transforms dense time series of N dual (i.e. HH & HV or VV & VH) polarized C-band SAR images into time series of N $m_c$ maps over agricultural sites. Fig. 1 shows the SMOSAR prototype architecture. Ancillary data such as land cover and soil texture information are also required. SMOSAR consists of two main processors: the on-line and off-line processors.

The on-line processor provides $m_c$ maps, referred to as Fast Delivered (FD) $m_c$ products, in near-real time, whereas the off-line processor requires additional SAR acquisitions to perform averages over time and relative error estimates on a specific date, and to produce Precision (P) $m_c$ and error products.

The main components of the SMOSAR on-line processor are:

• the masking block, which masks the stack of N co-polarized SAR images in correspondence of areas where the $m_c$ retrieval algorithm cannot be applied (i.e. areas dominated by volume scattering);

• the retrieving block, which uses the N masked co-polarized images to produce N FD $m_c$ maps.

The masking block works at pixel basis and is based on an adaptive threshold scheme, using the cross-polarized SAR data. The retrieving block works over agricultural or short vegetated areas and inverts temporal changes of radar backscatter, under the assumption that backscatter changes over a short time intervals depend on $m_c$ changes only. Finally, the on-line processor performs a spatial averaging over the retrieved $m_c$ maps in order to reduce the root mean
square error between the observed and the retrieved \(m_v\) maps.

Under the assumption that the algorithm works for a long period covering a temporal series of \(M\) images, the temporal series of \(M\) images are processed in \(M-N+1\) groups of \(N\) (e.g. \(N=4\)) images with overlapping. Each group is obtained from the previous one by shifting forward of one image. As a result, a generic image on a specific date is processed \(N\) times and then \(N\) correlated but not identical FD \(m_v\) maps for the same date are obtained as output.

The off-line processor converts the \(N\) FD correlated \(m_v\) maps referring to the same date into a unique \(P\) \(m_v\) map and the corresponding error map. This is accomplished by computing the average and the standard deviation at pixel scale. The single \(m_v\) map produced is expected to be characterized by a higher accuracy. Furthermore, the standard deviation map can be a worthwhile additional information layer useful to provide an internal assessment to the quality of the \(m_v\) maps. More precisely, a low pixel value (e.g. between 0 and 3%) implies that the conditions underlying the algorithm assumptions were substantially met and then the \(m_v\) estimate is considered reliable. Conversely, a high pixel value likely means that one or more algorithm assumptions were not fulfilled for that pixel. Intermediate values can be used as an indication of \(m_v\) reliability.

A first description of the classification and \(m_v\) retrieval approaches adopted in SMOSAR has been illustrated in [8], where also the algorithm assessment over the DEMMIN site was presented. In the following the assessment is extended to three sites with various climatic, hydrologic and agronomic characteristics.

4. SMOSAR VALIDATION

The SMOSAR validation has been separated into two main parts, with investigations on:

- the overall accuracy (OA) that can be achieved by the SMOSAR masking block over the various sites and the robustness of the rule based classification approach adopted in the masking block. The results have been presented in [9];

- the accuracy of the \(m_v\) product of the SMOSAR retrieving block as a function of spatial resolution, polarization, and site location.

The assessment of the SMOSAR retrieving block performances has been focused on the stability properties of the algorithm versus changes of site location (i.e. global scale applicability), polarization and pixel size scale. Temporal series of FD \(m_v\) maps, at 100m, 400m and at 700m pixels size, have been produced for Görmin, Matera and Flevoland sites. In order to evaluate the \(m_v\) product accuracy, the rmse between SAR retrieved and modelled \(m_v\) maps has been computed at HH and VV polarization and at the three aforementioned pixel scales. In addition, a further estimate of the rmse has been carried out over the entire area over each test site (i.e. approximately 30km\(^2\) over Görmin, 16km\(^2\) over Matera and 5km\(^2\) over Flevoland). In general, for all the sites and polarizations, the decrease of the spatial resolution generally produces an improvement (though moderate) of SMOSAR performances (i.e. the rmse decreases, the correlation increases and also the parameters of the linear fit improve).

Based on the available data set, the rmse errors across the various spatial scales and sites range between 4% and 6% at HH polarization, whereas they range between 1% and 10% at VV polarization. In this respect, while at HH polarization the observed rmse figures appear fairly stable, at VV polarization the variability is too large indicating either an intrinsic difficulty to generalize the retrieval approach over different sites or an insufficient number of images and sites analyzed. Further investigations are then necessary to ascertain this aspect. As an example, Fig. 2 shows the scatter plots between the total series of HH (upper panel) and VV (lower panel) SAR retrieved and modeled \(m_v\) maps at 400m pixels size. The FD \(m_v\) products are compared with the hydrologic model maps. The parameters of the linear fit (A and B), the rmse, correlation coefficient (R), the mean modeled and retrieved \(m_v\) values are reported on each plot. The various colored symbols refer to various sites as reported in the legend.

As far as the \(P\) \(m_v\) maps validation is concerned, in order to produce a sufficiently long time series of \(P\) products a longer series of SAR images is required. For this reason, only the Flevoland site over which the longest series of SAR data both at HH and HV and at VV polarizations was acquired, has been investigated. At HH polarization, a systematic improvement in the retrieved accuracy of \(P\) \(m_v\) products (with respect to the correspondent FD ones) is observed, whereas at VV polarization there is no improvement and results are substantially stable. This behavior is related to the existence of important biases in the FD products at VV polarization, whereas at HH polarization results were unbiased. A more systematic analysis of FD versus \(P\) products of SMOSAR will be addressed in a future paper.

5. CONCLUSIONS

The SMOSAR \(m_v\) retrieval algorithm, based on the use of backscatter temporal changes observed within a short revisiting time, has been presented. The performances of the retrieval algorithm have been assessed using temporal series of ASAR data (collected during the AgriSAR 2006 campaign and the Italian ENVISAT 2003 and 2005 campaigns) and of simulated S-1 data produced in the framework of the AgriSAR 2009 campaign. The assessment of the algorithm has been performed by comparing the SAR-derived \(m_v\) maps over the selected sites with hydrologically modelled \(m_v\) maps. A critical aspect faced in this analysis is that the available data sets were not homogeneous in terms of ground data documentation and of SAR images acquired.
Nevertheless, the analyzed data have allowed the assessment of the SMOSAR performance, focusing on its stability properties versus changes of site location and on the dependence of $m_v$ product accuracy on SAR polarization and spatial resolution. Based on the available data set, the accuracy of FD products has been estimated as 5% at a pixel size of 400m (approximately corresponding to a spatial resolution of 800m) and HH polarization. Conversely, at VV polarization the analyzed SAR data are not sufficient to provide a reliable figure for the retrievable accuracy. Physical arguments and the analysis of backscatter data at coarse resolution versus $m_v$ indicate that at VV polarization the expected accuracy is lower than at HH. However, this crucial point requires more investigations and more experimental data to be realistically quantified. The analysis of SMOSAR P product has been carried out at HH polarization and on the Flevoland site only. Results confirm an improvement in the accuracy of these products with respect to the FD ones. However, the information content of P maps and associated error maps require further investigations as the available data set was too limited to extrapolate the obtained results.

To conclude, the results of this paper support the conclusion that the time series of S-1 SAR data acquired by the forthcoming S-1 constellation will make the development of $m_v$ products at high spatial resolution with improved robustness possible.

6. ACKNOWLEDGEMENT

The research in this paper was supported by the European Space Agency (ESA) contract “GMES Sentinel-1 Soil Moisture Algorithm Development” funded by the European Union. The team is grateful to ESA for providing the ground and SAR data collected during the AgriSAR 2006 and 2009 campaigns.

7. REFERENCES