Environmental Modelling & Software 102 (2018) 13–28

Contents lists available at ScienceDirect

Environmental Modelling & Software

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

A surface soil moisture mapping service at national (Italian) scale based on Sentinel-1 data

Luca Pulvirenti [a, *], Giuseppe Squicciarino [a], Luca Cenci [b, a], Giorgio Boni [a, c], Nazzareno Pierdicca [b], Marco Chini [d], Cosimo Versace [e, a], Paolo Campanella [f, a]

[a] CIMA Research Foundation, 17100 Savona, Italy
[b] Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, 00184 Rome, Italy
[c] Dipartimento di Informatica, Bioingegneria, Robotica e Ingegneria dei Sistemi, University of Genoa, 16145 Genoa, Italy
[d] Department of Environmental Research and Innovation, Luxembourg Institute of Science and Technology, L-4422 Belvaux, Luxembourg
[e] Acrotec Foundation, 17100 Savona, Italy
[f] Fadeout Software srl, 16124 Genoa, Italy

A R T I C L E I N F O

Article history:
Received 16 May 2017
Received in revised form
20 October 2017
Accepted 26 December 2017

Keywords:
Soil moisture
Sentinel-1
Multi-temporal algorithm
Plant water content
Landsat-8
Sentinel-2
MODIS

A B S T R A C T

This paper presents MULESME, a software designed for the systematic mapping of surface soil moisture using Sentinel-1 SAR data. MULESME implements a multi-temporal algorithm that uses time series of Sentinel-1 data and ancillary data, such as a plant water content map, as inputs. A secondary software module generates the plant water content map from optical data provided by Landsat-8, or Sentinel-2, or MODIS. Each output of MULESME includes another map showing the level of uncertainty of the soil moisture estimates. MULESME was tested by using both synthetic and actual data. The results of the tests showed that root mean square error is in the range between 0.03 m^3/m^3 (synthetic data) and 0.06 m^3/m^3 (actual data) for bare soil. The accuracy decreases in the presence of vegetation (root mean square in the range 0.08–0.12 m^3/m^3), as expected.

Software availability

Name of software: MULESME (MUltitemporal LEast Square Moisture Estimator)
Developer: Luca Pulvirenti, CIMA Research Foundation, via A. Magliotto 2, 17100 Savona, Italy (luca.pulvirenti@cimafoundation.org)
First year available: 2016
Programming language: IDL
Required hardware: 16 GB RAM minimum
Supported systems: Windows, Linux
Required software: IDL/ENVI. MULESME was designed and tested using the ENVI 5.4.1 version. The MODIS Conversion Toolkit must be used to process MODIS data, if used. The freely available ESA Sentinel Application Platform (SNAP) must be used for the pre-processing of Sentinel-1 data
Availability: mail to luca.pulvirenti@cimafoundation.org to request the IDL/ENVI source code of MULESME. A set of test data can be provided too
License: ENVI + IDL commercial license (Harris Geospatial)

1. Introduction

The role of soil moisture (*SM*) as a key variable for the characterization of the global climate is widely recognized within the international scientific community. Surface *SM* controls the partitioning of available energy at the ground surface into latent and sensible heat exchange through evaporation and transpiration processes (Anagnostopoulos et al., 2017; Petropoulos and McCalmont, 2017). Furthermore, the *SM* content of the root zone regulates the redistribution of precipitation into infiltration, runoff, storage in the root zone and percolation into deeper ground water storage (Sheikh et al., 2009).

* Corresponding author. Via Armando Magliotto 2, 17100 Savona, Italy.
E-mail address: luca.pulvirenti@cimafoundation.org (L. Pulvirenti).

https://doi.org/10.1016/j.envsoft.2017.12.022
1364-8152/

Coarse resolution (25–50 km) *SM* estimates, provided by satellite microwave radiometers or scatterometers, are useful in support of numerical weather prediction, climate monitoring and flood forecasting (Brocca et al., 2017; Hornacek et al., 2012). High-resolution (0.1–1 km) products, obtained from Synthetic Aperture Radar (SAR) data, can be useful even for applications such as monitoring of agricultural yield at field level, or irrigation management (Hornacek et al., 2012). Although C-band is not the ideal frequency for soil moisture retrieval applications, being also sensitive to soil roughness and the presence of vegetation (Fascetti et al., 2016), the availability of six-day repeat Sentinel-1 (S1) C-band SAR data currently represents the only opportunity to systematically produce surface (depth of ~5 cm) *SM* maps at high spatial resolution.

Recent literature studies demonstrated that S1 data can be suitable for *SM* mapping (Balenzano et al., 2012; Hornacek et al., 2012; Paloscia et al., 2013; Pierdicca et al., 2014). Nonetheless, the systematic use of S1 data for providing end users with an operational *SM* mapping service poses two problems. The first one regards the accuracy of *SM* estimates, because *SM* retrieval from SAR data is an ill-posed problem. In fact, even considering the most favorable condition, i.e., a bare terrain, SAR measurements are sensitive not only to *SM*, but also to soil roughness (Marzahn et al., 2012) and are also very noisy because of the speckle noise characteristic of any SAR image. This problem was dealt with in the literature by developing multi-temporal retrieval algorithms (Balenzano et al., 2011; Hornacek et al., 2012; Kim et al., 2012; Pierdicca et al., 2010), which assume that the temporal scale of variation of soil roughness is considerably slower than that of *SM*. Hence, if a dense time-series of SAR data is available, as expected using S1, short term changes in the backscattering coefficient σ^0 (that represents the SAR measurement) are basically related to *SM* variations (Balenzano et al., 2011). The situation is further complicated by the dependence of σ^0 on biomass parameters, as well as plant structure and geometry, if vegetation is present. To correct for the vegetation influence on σ^0, simple semi-empirical models, such as the Water Cloud model (Attema and Ulaby, 1978), using few "bulk" parameters such as the plant water content *W*, are commonly used in the literature. These models can be easily inverted to discriminate the soil contribution to the SAR measurement from that related to vegetation, but require reliable data about the bulk parameters. Various studies demonstrated the potential of retrieving *W* from optical images, in particular using semi-empirical relationships between *W* and the normalized difference vegetation index (*NDVI*) (e.g. Jackson et al., 1999; Liu and Shi, 2016; Pierdicca et al., 2010). However, it must be underlined that tackling the effects of vegetation is still a challenge for any estimation approach, because semi-empirical models may lack of generality. In the literature, accuracies in the order of 0.04–0.13 m^3/m^3 (Root Mean Square Error: RMSE) are reported (e.g. Hajnsek et al., 2009) for *SM* retrieval from SAR, but these scores often refer to specific test sites and/or case studies in which multi-frequency or fully polarimetric data (see section 4.2) were available. Using S1 data at large (e.g. national) scale, even higher RMSE can be expected, especially if SAR observations are performed under dense vegetation conditions, so that the need to improve the quality of *SM* estimates, for instance by assimilating them into a hydrological model (e.g. Brocca et al., 2012; Cenci et al., 2016a), clearly emerges. It should also be pointed out that in areas where *W* is very high, as well as in forested and urban areas and in areas with complex topography, *SM* retrieval from SAR is unfeasible, so that the corresponding maps have gaps (i.e., masked areas).

The second problem connected to the systematic use of S1 data for designing a *SM* mapping service, is related to the more general need of developing software tools that allow the community to really take advantage of the progresses achieved in Earth Observation (EO) technology. Examples of these tools are those developed by Petropoulos et al. (2013), for the pre-processing of the Spinning Enhanced Visible and Infrared Imager (SEVIRI) data, Keramitsoglou et al. (2006), for SAR-based oil spill detection, Boni et al. (2016) and Martinis et al. (2015), to produce SAR-based maps of flooded areas. For what concerns *SM* retrieval, the need of EO-based software tools was recently highlighted by Srivastava (2017) and Petropoulos et al. (2015). Tischler et al. (2007) designed a GIS tool to integrate *SM* predictions from a land surface model with EO measurements.

This paper presents MULESME (MUltitemporal LEast Square Moisture Estimator), a software implementing an automated processing chain designed for an operational SAR-based service whose aim is the production of daily high-resolution (~500 × 500 m^2) *SM* maps at national (Italian) scale. The paper is focused on the design of the software and does not propose a new *SM* retrieval algorithm. Note that an operational SAR-based *SM* mapping service does not exist to date and this prevents potential users from fully exploiting the advances achieved in retrieving *SM* (or at least its variations) from short revisit S1 data. Hence, a paper presenting a software able to implement this kind of service by systematically producing an updated high resolution *SM* map as soon as new S1 images are available, represents a novel contribution to the literature. Besides hydrologists and meteorologists, potential users interested in this software may be authorities or government agencies at national scale to monitor either antecedent soil wetness conditions in case of flood alert issues (Teng et al., 2017), or water resources consumption in areas affected by droughts. Even agricultural managers could be interested in a *SM* mapping service in order to get timely information about water requirements of the soil (Flores-Carrillo et al., 2017).

To our knowledge, a high-resolution *SM* mapping service was never proposed in the literature, as previously pointed out. A near-real-time *SM* distribution service is implemented by the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) using low resolution Advanced Scatterometer (ASCAT) data (Wagner et al., 2013). The European Space Agency (ESA) recently released the Soil Moisture Ocean Salinity (SMOS) Level 2 Soil Moisture Near Real Time Neural Network data product (Rodriguez-Fernandez et al., 2015); even in this case the spatial resolution is in the order of tens of km.

MULESME uses, as input data, time series of S1 Interferometric Wide Swath products (see section 2.1), as well as ancillary data, namely a land cover map, topographic slope information, local incidence angle maps and a map representing the state of the vegetation. The latter is generated by a secondary processor that uses optical data (Landsat-8, Sentinel-2, MODIS) as inputs.

The software was implemented using the IDL language and the ENVI routines that can be launched by means of specific IDL instructions. It was developed within the framework of the MIDA (Italian acronym of maps of soil moisture for hydrologic data assimilation) project, funded by the Italian Space Agency (ASI) and the WASDI (Web based ASI Spatial Data Infrastructure) project, funded by the European Space Agency (ESA) on behalf of ASI. In particular, the software was firstly designed in the framework of the MIDA project, whose aim is the generation of *SM* maps through the assimilation of S1 derived estimates into a hydrological model. In the near future, MULESME will be installed in the WASDI platform, connected to the Italian Sentinel Collaborative Ground Segment (Coll-It), in order to exploit the Coll-It storage capability and its computing resources, without the need of moving big amounts of data towards the processors.

Section 2 gives an overview on Sentinel-1 data and describes the algorithms that are implemented in MULESME to retrieve *SM*,