Data in Brief 38 (2021) 107345

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

Dataset of Sentinel-1 surface soil moisture time series at 1 km resolution over Southern Italy

Anna Balenzano [a,*], Francesco Mattia [a], Giuseppe Satalino [a], Francesco P. Lovergine [a], Davide Palmisano [a], Malcolm W.J. Davidson [b]

[a] National Research Council of Italy (CNR), Institute for Electromagnetic Sensing of the Environment (IREA), UOS Bari, Italy
[b] European Space Agency, Mission Science Division, Noordwijk, the Netherlands

A R T I C L E I N F O

Article history:
Received 18 June 2021
Revised 1 September 2021
Accepted 2 September 2021
Available online 4 September 2021

Keywords:
Surface soil moisture
High resolution
Sentinel-1
Synthetic aperture radar (SAR)
Time series
Short term change detection

A B S T R A C T

This paper describes the specifications of the surface soil volumetric water content (Θ) [m^3/m^3] product derived from Sentinel-1 (S-1) data and assessed in the study “Sentinel-1 soil moisture at 1 km resolution: a validation study” [1]. The S-1 Θ product consists of Θ mean and standard deviation values at 1 km spatial resolution and is expected to support applications in agriculture and hydrology as well as the Numerical Weather Prediction at regional scale [2]. The retrieval algorithm is a time series based short term change detection that is implemented in the “Soil MOisture retrieval from multi-temporal SAR data” (SMOSAR) code (v2.0). The provided dataset represents an example of the developed S-1 Θ product and consists of a time series of 183 S-1 Θ images over Southern Italy from January 2015 to December 2018. The maps were produced for each ascending S-1 acquisition date on the Relative Orbit Number (RON) 146 and the temporal gap between consecutive maps is 6 days (when both S-1A and S-1B data are available) or 12 days.

DOI of original article: 10.1016/j.rse.2021.112554

* Corresponding author.
E-mail address: anna.balenzano@cnr.it (A. Balenzano).
Social media: (A. Balenzano)

https://doi.org/10.1016/j.dib.2021.107345

Specifications Table

Subject	Earth-Surface Process
Specific subject area	Remote sensing of high resolution surface soil volumetric water content from Synthetic Aperture Radar data at C-band.
Type of data	Raster images
How data were acquired	Θ maps were derived from S-1 Interferometric Wide Swath (IW) observations. The retrieval code is called SMOSAR ("Soil MOisture retrieval from multi-temporal SAR data") and was developed at the Institute for Electromagnetic Sensing of the Environment (IREA), National Research Council of Italy (CNR).
Data format	Processed
Parameters for data collection	The time series of the Θ maps was derived from ascending (RON 146) S-1 A and B IW data, acquired from January 2015 to December 2018 at a temporal resolution of either 6 or 12 days.
Description of data collection	The S-1 data were first pre-processed and, then, used as input into the SMOSAR code. SMOSAR transforms dense or quasi-dense time series (i.e., 6 or 12 days revisit) of N dual-polarized S-1 IW images at 40 m pixel size (∼100 m resolution) into N maps at 520 m pixel size (∼1 km resolution).
Data source location	Southern Italy
Data accessibility	Repository name: Zenodo Data repository: https://doi.org/10.5281/zenodo.5006307
Related research article	A. Balenzano, F. Mattia, G. Satalino, F.P. Lovergine, D. Palmisano, J. Peng, P. Marzahn, U. Wegmuller, O. Cartus, K. Dabrowska-Zielinska, J. Musial, M. W. J. Davidson, V. R.N. Pauwels, M. H. Cosh, H. McNairn, J. T. Johnson, J. P. Walker, S. H. Yueh, D. Entekhabi, Y. H. Kerr and T. J. Jackson, Sentinel-1 soil moisture at 1 km resolution: a validation study, Remote Sens. Environ. (2021), https://doi.org/10.1016/j.rse.2021.112554 [1].

Value of the Data

- The S-1 Θ dataset provides an example of surface soil volumetric water content [m^3/m^3] retrieval from S-1 constellation at high spatial resolution, as well as its co-registered uncertainty.
- The S-1 Θ information can bridge the gap between in situ field-scale observations and operational low-resolution Θ products delivered by missions, such as ESA Soil Moisture and Ocean Salinity (SMOS) and NASA Soil Moisture Active Passive (SMAP). In particular, land applications, such as hydrology and high resolution Numerical Weather Prediction, are expected to have beneficial impact from such a high resolution Θ product.
- The S-1 Θ product can be used to quantify the variability of Θ across scales and link it to the dynamics of the underlying water, energy and biological processes.

1. Data Description

The dataset consists of a time series of the Sentinel-1 (S-1) Θ product at 1 km spatial resolution assessed in [1] over Southern Italy. The Θ time series was obtained in correspondence of the ascending (RON 146) S-1 Interferometric Wide swath (IW) acquisition dates from January 2015 to December 2018. The temporal gap between consecutive Θ maps is 6 days (when both S-1A and S-1B data are available) or 12 days. On each date (183 in total), two co-registered layers are provided: mean Θ and its standard deviation, referred to as mean and stddev, respectively.

Table 1
Data format and range of the S-1 Θ maps.

Format	Envi image
Map Projection	Geographic lat/lon
Datum	WGS84
Pixel size	0.0052 Deg
Data Type	Integer*2
layers and units	Θ mean [m^3/m^3]: Θ at ~520 m pixel size, corresponding to a resolution of ~1 kmΘ stddev [m^3/m^3]: Θ standard deviation
Range	Θ mean in [0.015-0.60] m^3/m^3 Θ stddev > 0 m^3/m^3
Threshold	33% In SMOSAR, the masking and retrieval is carried out at 40 m pixel scale. The masked pixels are set to <no data>. Then, a low pass filter and a resampling leading to a (final) pixel size of ~520 m are applied. When the percentage of <no data> 40 m pixels is higher than a threshold (i.e., 33%) the final pixel is masked (i.e., set to <no data>).
Scale factor	10000 The physical values of Θ mean, Θ stddev are obtained dividing the Digital Number by 10000

Each layer is a geo-referenced map in lat/lon projection, with a square pixel of length of 0.0052 deg (~520 m). All maps are stored in a zip archive. The name convention adopted for the Θ maps is:

SSM_P_STCD_yyyymmdd_S1X_IW_GRDH_1SDV_yyyymmddThhmmss_MNxx_RLddmm.vv.
GTC_MAP_v2_0.img_type

where

- SSM_P_STCD = Precision Surface Soil Moisture product obtained by the Short Term Change Detection algorithm;
- yyyymmdd = Map date: year, month, day;
- S1X_IW_GRDH_1SDV_yyyymmddThhmmss = S-1 product type (string derived from the S-1 naming convention, i.e. X=A,B; IW=Interferometric Wide swath; GRDH=Ground Range Detected High resolution; 1SDV=Level 1 Standard product, dual VV+VH polarisation, yyyymmddThhmmss=sensing date and time of acquisition;
- MNxx_RLddmm = nr of frames into the S-1 mosaic and Reference Latitude crossing the northern frame in decimal (mm) degree (dd);
- vv.GTC = S-1 product (vv polarization, Geocoded Terrain Corrected);
- MAP_type =MAP_v2_0.img_{mean|stddev}.

The characteristics of the Θ maps are summarized in Table 1. The metadata are reported in a header file (*.hdr), associated to each Θ map. The header file (*.hdr) reports information on the data format and geographic projection of the map (Table 2). Fields in Table 2 have the same structure of the ENVI header files [3].

As an example, Fig. 1 shows the Θ mean and Θ stddev maps produced over Southern Italy on January, 9 2018. In correspondence of mountainous areas, the map is significantly masked due to the important presence of woods. Along the coast, the Θ field shows the pattern of precipitation events (dark blue) alternated to dryer areas. It can also be noted that the higher the Θ value, the higher the associated stddev, generally.

Table 2
Description of the ENVI header file.

Field	Content
description	{string}
samples	Integer
lines	Integer
bands	Integer
header offset	Integer
file type	ENVI Standard
data type	2 (type integer*2);
interleave	bsq (band sequential)
byte order	0 (0=first byte is the least significant)
map info	{“Geographic Lat/Lon”, reference pixel x location, reference pixel y location, pixel latitude, pixel longitude, x pixel size, y pixel size, “WGS84”, “units=Degrees”}
coordinate system string	coordinate system string = {GEOGCS[“GCS_WGS_1984”,DATUM[“D_WGS_1984”,SPHEROID[“WGS_1984”,6378137, 298.257223563]],PRIMEM[“Greenwich”,0],UNIT[“Degree”,0.017453292519943295]]}
Waveleng. units	Unknown
band names	{string}

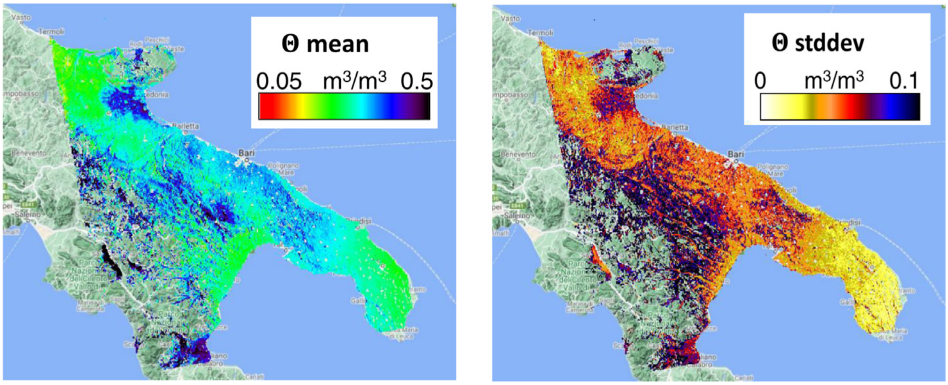

Fig. 1. Example of S-1 surface soil volumetric water content (Θ) product over Southern Italy on January, 9 2018. Left: Θ mean. Right: Θ stddev.

2. Experimental Design, Materials and Methods

The Θ maps are derived from time series of S-1 Synthetic Aperture Radar (SAR) data. The Θ retrieval implements a change detection algorithm (STCD) requiring a short revisit between SAR observations [4]. The rationale is that changes of surface parameters influencing the radar backscatter, apart from Θ, such as soil roughness, canopy structure, vegetation water content, usually take place at longer temporal scales than Θ changes (excluding periods of cultivation practices). Therefore, ratios of backscatter between subsequent SAR acquisitions, sufficiently close in time, are expected to track changes in Θ only. This approximation (also referred to as “alpha approximation”) makes robust the retrieval and expedites the processing.

The algorithm can be applied to bare and vegetated surfaces dominated by soil attenuated scattering that enables good radar sensitivity to Θ throughout the growing season. This restriction implies that, before the retrieval, those areas showing a poor radar sensitivity to Θ (i.e., areas dominated by volume scattering) are masked. The procedure is implemented in two-steps. The first one consists of using a quasi-static land cover map (e.g., CCI LandCover [5]) to obscure classes such as forests, urban areas, water bodies. The second step is the dynamic masking of environments characterized by volume scattering. The method exploits the Kittler-Illingworth [6] algorithm, applied to S-1 cross-polarized observations [7]. The result is that the Θ retrieval is applied only over areas dominated by soil attenuated scattering that show an adequate radar sensitivity to Θ (the masked pixels are labelled as <no data>).

The code implementing the STCD algorithm is referred to as "Soil MOisture retrieval from multi-temporal SAR data" (SMOSAR) [8]. SMOSAR works pixel-wise and transforms dense or quasi-dense time series (i.e., 6 or 12 days-revisit) of N S-1 IW images at 40 m pixel size (~100 m resolution) into N Θ maps. The final product is delivered at a lower spatial resolution in order to reduce the Θ variability and mitigate the impact of possible rapid and sparse changes of vegetation and/or roughness that can be wrongly interpreted as Θ changes. For this scope, SMOSAR finally performs a low pass filter and subsequent resampling of pixel spacing to ~520 m (corresponding to a spatial resolution of ~1 km). The resolution cells in which the percentage of <no data> pixels is equal or larger than 33% are masked. A co-registered layer, consisting of the Θ standard deviation (> 0 m^3/m^3) is also provided.

Ethics Statements

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

CRediT Author Statement

Anna Balenzano: Conceptualization, Methodology, Writing – original draft; **Francesco Mattia:** Methodology, Writing – original draft; **Giuseppe Satalino:** Software, Visualization; **Francesco P. Lovergine:** Software, Data curation; **Davide Palmisano:** Visualization, Writing – original draft; **Malcolm W.J. Davidson:** Supervision.

Acknowledgments

This research was supported by the Scientific Exploitation of Operational Missions (SEOM) program of the European Space Agency, through the project "Exploitation of S-1 for Surface Soil Moisture Retrieval at High Resolution (Exploit-S-1)" (contract 4000118762/16/I-NB). The authors would like to thank Maria Mottola for her technical assistance.

References

[1] A. Balenzano, F. Mattia, G. Satalino, F.P. Lovergine, D. Palmisano, J. Peng, P. Marzahn, U. Wegmuller, O. Cartus, K. Dabrowska-Zielinska, J. Musial, M.W.J. Davidson, V.R.N. Pauwels, M.H. Cosh, H. McNairn, J.T. Johnson, J.P. Walker, S.H. Yueh, D. Entekhabi, Y.H. Kerr, T.J. Jackson, Sentinel-1 soil moisture at 1 km resolution: a validation study, Remote Sens. Environ. 263 (2021) 112554, doi:10.1016/j.rse.2021.112554.

[2] J. Peng, C. Albergel, A. Balenzano, L. Brocca, O. Cartus, M.H. Cosh, … P. de Rosnay, A roadmap for high-resolution satellite soil moisture applications – confronting product characteristics with user requirements, Remote Sens. Environ. 252 (2021) 112162.

[3] ENVI - Environment for Visualizing Images, © 2017 Exelis Visual Information Solutions, a subsidiary of Harris Corporation, https://www.l3harrisgeospatial.com/docs/using_envi_home.html.

[4] A. Balenzano, F. Mattia, G. Satalino, M.W.J. Davidson, Dense temporal series of C- and L-band SAR data for soil moisture retrieval over agricultural crops, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 4 (2011) 439–450.

[5] P. Defourny, C. Lamarche, S. Bontemps, T. De Maet, E. Van Bogaert, I. Moreau, C. Brockmann, M. Boettcher, G. Kirches, J. Wevers, M. Santoro, ESA Land Cover CCI, Product User Guide, v2.0., 2017. http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf.

[6] J. Kittler, J. Illingworth, Minimum Error Thresholding, Pattern Recognit. 19 (1986) 41–47.

[7] G. Satalino, A. Balenzano, F. Mattia, M.W.J. Davidson, C-band SAR data for mapping crops dominated by surface or volume scattering, Geosci. Remote Sens. Lett. 11 (2014) 384–388.
[8] A. Balenzano, G. Satalino, F. Lovergine, M. Rinaldi, V. Iacobellis, N. Mastronardi, F. Mattia, On the use of temporal series of L- and X- band SAR data for soil moisture retrieval. Capitanata plain case study, Eur. J. Remote Sens. 46 (2013) 721–737.