An evaluation of the potential of Sentinel 1 for improving flash flood predictions via soil moisture–data assimilation

Luca Cenci\textsuperscript{1,2,3}, Luca Pulvirenti\textsuperscript{2}, Giorgio Boni\textsuperscript{2,4}, Marco Chini\textsuperscript{5}, Patrick Matgen\textsuperscript{5}, Simone Gabellani\textsuperscript{2}, Giuseppe Squicciarino\textsuperscript{2}, and Nazzareno Pierdicca\textsuperscript{3}

\textsuperscript{1}Scuola Universitaria Superiore IUSS Pavia, Pavia, 27100, Italy
\textsuperscript{2}CIMA Research Foundation, Savona, 17100, Italy
\textsuperscript{3}Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Rome, 00184, Italy
\textsuperscript{4}Department of Informatics, Bioengineering, Robotics and Systems Engineering, University of Genoa, Genoa, 16145, Italy
\textsuperscript{5}Luxembourg Institute of Science and Technology, Belvaux, 4422, Luxembourg

Correspondence to: Luca Cenci (luca.cenci@cimafoundation.org)

Received: 30 November 2016 – Revised: 25 July 2017 – Accepted: 11 September 2017 – Published: 1 November 2017

Abstract. The assimilation of satellite-derived soil moisture estimates (soil moisture–data assimilation, SM–DA) into hydrological models has the potential to reduce the uncertainty of streamflow simulations. The improved capacity to monitor the closeness to saturation of small catchments, such as those characterizing the Mediterranean region, can be exploited to enhance flash flood predictions. When compared to other microwave sensors that have been exploited for SM–DA in recent years (e.g. the Advanced SCATterometer – ASCAT), characterized by low spatial/high temporal resolution, the Sentinel 1 (S1) mission provides an excellent opportunity to monitor systematically soil moisture (SM) at high spatial resolution and moderate temporal resolution. The aim of this research was thus to evaluate the impact of S1-based SM–DA for enhancing flash flood predictions of a hydrological model (Continuum) that is currently exploited for civil protection applications in Italy. The analysis was carried out in a representative Mediterranean catchment prone to flash floods, located in north-western Italy, during the time period October 2014–February 2015. It provided some important findings: (i) revealing the potential provided by S1-based SM–DA for improving discharge predictions, especially for higher flows; (ii) suggesting a more appropriate pre-processing technique to be applied to S1 data before the assimilation; and (iii) highlighting that even though high spatial resolution does provide an important contribution in a SM–DA system, the temporal resolution has the most crucial role. S1-derived SM maps are still a relatively new product and, to our knowledge, this is the first work published in an international journal dealing with their assimilation within a hydrological model to improve continuous streamflow simulations and flash flood predictions. Even though the reported results were obtained by analysing a relatively short time period, and thus should be supported by further research activities, we believe this research is timely in order to enhance our understanding of the potential contribution of the S1 data within the SM–DA framework for flash flood risk mitigation.

1 Introduction

In small catchments, such as those characterizing the Mediterranean region, flash flood risk can be very high. In these basins, antecedent soil moisture (SM) conditions largely determine whether a given amount of rainfall produces conditions for a flash flood alert or not. Therefore, accurate and reliable SM estimates are of fundamental importance for reducing uncertainties of flash flood early warning systems (Borga et al., 2011; Massari et al., 2015a). Within this framework, microwave (MW) remote sensing offers an excellent opportunity for synoptical SM monitoring.

In recent decades, satellite-derived SM maps obtained from MW sensors characterized by low spatial/high temporal resolution (e.g. radiometers and scatterometers) have been exploited to improve discharge predictions of hydrological models via data assimilation (DA) techniques (SM–
DA) (e.g. Brocca et al., 2010; Matgen et al., 2012a; Laiolo et al., 2015a, b; Massari et al., 2015a, b; Cenci et al., 2016a). DA allows for combination of optimally different sources of information that may be characterized by different spatio-temporal resolution and accuracy. Via DA techniques, SM estimates obtained from satellite acquisitions and hydrological modelling can be merged to obtain more reliable state estimates and reduce the uncertainty of model forecasts (Walker and Houser, 2005; Liu et al., 2012). These techniques can thus be applied in an operational context to enhance, for example, flash flood predictions. However, SM–DA is a complex task and its performances are greatly influenced by several factors, such as the characteristics of the hydrological catchment, the properties of the satellite sensor (or of the satellite mission), the type of SM retrieval algorithm and the pre-processing techniques that are applied to allow for a correct assimilation of the satellite-derived SM maps (i.e. the observations). In fact, most of the time observed and modelled data have different spatial resolutions. Additionally, their SM estimates may refer to different soil layers (satellite: near-surface; model: root zone – RZ). Furthermore, systematic biases between the observing and the modelling system are usually present and an accurate characterization of their errors is generally difficult to achieve. These problems must be carefully accounted for and tackled before the assimilation is carried out (Drusch, 2005; Wagner et al., 2007; Massari et al., 2015a), and these tasks are usually achieved during the pre-processing step.

Notwithstanding the reported evidence on the utility of SM–DA for operational hydrology, further research is needed before an actual implementation in flash flood early warning systems can be envisaged (Wagner et al., 2007; Cenci et al., 2016a). In particular, it has been pointed out that operational agencies need to focus on the assimilation of new observations acquired by different sensors with different spatio-temporal resolutions (Liu et al., 2012). In this general context, it is worth understanding the added value of “new generation” high spatial resolution SM maps obtained using synthetic aperture radar (SAR) systems, e.g. Sentinel 1 (S1) (Wagner et al., 2007; Laiolo et al., 2015a; Cenci, 2016; Cenci et al., 2016a), characterized by a comparatively higher spatial and lower temporal resolution with respect to traditional MW sensors exploited for such purposes (e.g. the Advanced SCAtterometer, ASCAT; the Soil Moisture and Ocean Salinity, SMOS; and Soil Moisture Active Passive, SMAP). Recent analyses were carried out by assimilating S1 SAR and SMAP radiometer observations within the NASA Catchment Land Surface Model to improve the accuracy of SM estimates (Lievens et al., 2017) and by exploiting a S1-based SM map as input data for an event-based rainfall–runoff model for improving its hydrological simulation (Alexakis et al., 2017). However, to our knowledge, there is a lack of studies dealing with the assimilation of S1-derived SM products within hydrological modelling to improve continuous streamflow simulations and flash flood predictions (preliminary analyses – preparatory to this study – can be found in Cenci, 2016, Cenci et al., 2016b, c, and Cenci et al., 2017a, b).

In order to investigate this research topic, a SM–DA system was set up to assimilate different S1-derived SM products (S1 SM–DA) within a hydrological model that is currently implemented in Italy for flash flood forecasting, both at regional and national scale (Continuum; Silvestro et al., 2013). All the S1 SM maps (spatial resolution: ∼500 × 500 m²; temporal resolution 12 days) were produced by using the same SM retrieval algorithm but with different corrections of the vegetation effect on the radar backscattering (σ₀). The analysis was carried out in the Orba River catchment, an Italian basin prone to flash floods, during the time period October 2014–February 2015. The Nudging assimilation scheme (Walker and Houser, 2005) was chosen for its computational efficiency. The latter is of particular importance when dealing with operational applications. Since in situ SM monitoring stations were not available, the impact of the assimilation was evaluated by comparing the model discharge predictions (obtained with and without the assimilation) with observed discharge data, with a particular emphasis placed on the impact on high flow forecasts. Results were compared against those obtained by assimilating a more conventional high spatial (and temporal) resolution SM product derived from ASCAT acquisitions: H08 (SM-OBS-2: spatial resolution 1 × 1 km², temporal resolution 36 h; Wagner et al., 2013). The purpose of the comparison was to have a better understanding of the impact of the spatial and temporal resolutions of the observations in a SM–DA system. As the exploitation of S1-derived SM products for DA applications is a new research topic, the analysis was also aimed at defining the best pre-processing steps to be applied before their assimilation.

2 Study area, hydrological model and satellite data

2.1 Study area: the Orba River catchment

The Orba River catchment (Fig. 1) is located in the north-western (NW) Italian Apennines. Because of its small/medium dimension, its particular relief-driven organization in land cover distribution and its climatic and hydrological characteristics, it can be considered a representative catchment of the NW Mediterranean area. Its surface area is circa 800 km² and the river length is approximately 75 km. Its land cover is characterized by a forested, mountainous upper-slope area and a final alluvial part in which agricultural activities are present. Urban areas cover only a small percentage of the basin. As with most of the NW Mediterranean basin, the Orba catchment is characterized by a high annual variability of climatic characteristics, resulting in a marked seasonality of the hydrological regimes. The runoff is strongly influenced by the seasonal distribution of precipi-
Lower flows generally occur during the dry summers. The rainy season that generates higher flows and occasionally flash floods usually occurs in autumn and early winter, but late summer and spring rainfall/runoff can at times be relevant as well. Runoff generation strongly depends on rainfall intensity and land surface properties (e.g. geomorphology of slopes/topography, channel systems/hydrography and land cover) (Thornes et al., 2009; Cenci et al., 2016a; Merheb et al., 2016).

2.2 Hydrological model: Continuum

The hydrological model used in the experiment was Continuum, which is a time-continuous, spatially distributed, physically based model. It relies on a morphological approach based on the identification of drainage network components. The basin is represented using a regular square mesh based on a digital elevation model (DEM). Flow directions are identified on the basis of the directions of maximum slope derived by the DEM. Mass and energy balance equations are solved on each pixel of the domain. In this experiment, the spatial resolution of the model (i.e. of the DEM) was set to 100 x 100 m². Continuum needs five meteorological input datasets, namely precipitation, air temperature, air humidity, shortwave radiation and wind speed. In the experiment, the corresponding maps were obtained by interpolating to the model grid data recorded by the Italian Civil Protection Department network of in situ stations present in the catchment: 19 pluviometers, 19 thermometers, 10 hygrometers, 7 radiometers (shortwave radiation), and 8 anemometers. Interpolation algorithms were Kriging for precipitation, altitude relationship for air temperature and inverse distance weighting for the other parameters. The temporal resolution of the meteorological input was 1 h. The model requires six calibration parameters: two for the surface flow (which is distinguished between channel and hillslope flows), two for the subsurface flow and two for deep flow and the water table. Concerning the infiltration (process of major interest for the purpose of this work), the sub-surface flow component follows the drainage network direction, and it is modelled using a semi-empirical approach based on a modified Horton algorithm. Each model cell is represented as a tank with a maximum soil storage capacity ($V_{max}$ [L]), whose average value at catchment scale is 184 mm and the water contained within it is modelled as the actual water volume ($V(t)$ [L]). Consequently, the SM can be estimated, for each cell, at each time step, as the RZ saturation degree (SD):

$$SD_{MOD}(t) = \frac{V(t)}{V_{max}}.$$  

Instead, the deep-flow component follows the hydraulic head gradient, obtained by modelling the water table. The energy balance is solved explicitly at cell scale by means of the force–restore equation, which allows the land surface temperature (LST) to be a distributed state variable. The model was calibrated by minimizing the differences between model streamflow simulations and time series of observed discharge data. For any further information on the model structure and its calibration over the selected study, please refer to the detailed description provided by Silvestro et al. (2013).

2.3 Satellite data: Sentinel 1 and ASCAT-derived soil moisture products

The S1 constellation is composed of two satellites, S1 A and S1 B, each of them carrying a C-band SAR sensor. It has a spatial resolution of 5 x 20 m² (single look) in interferometric wide (IW) swath mode (250 km swath width) and a temporal resolution of 12 days if a single orbit of a single satellite is considered (6 days if the same orbit is taken into account by considering both satellites; Sentinel-ESA, 2017). The S1-derived surface SM maps that were utilized in this research were produced through the application of a multi-temporal SM retrieval algorithm based on a maximum likelihood approach developed by Pulvirenti et al. (2017). Level-1 ground range detected (GRD) images were exploited to achieve this task. Only S1-A acquisitions were used because S1-B had not been launched at the time of the analysis. S1 images were available starting from 10 October 2014, for a total of 12 images. Orbit 66 was selected because the Orba catchment is entirely contained within one single-image frame. Images were acquired in the early morning (at 05:35 UTC) to avoid daytime decoupling between surface and RZ soil layers (Jackson, 1980). Having considered a period where the soil can be assumed to be only slightly vegetated, a S1-
based SM product without the correction of the vegetation effect on $\sigma^0$ was produced (hereinafter defined as S1 No-CorVeg). Two additional products, in which the correction of the vegetation effect was accounted for, were generated as well. The correction was performed using both optical (Landsat 8) and X-band SAR (Cosmo SkyMed® – CSK®) images. These products will hereinafter respectively be defined as S1 CorVegOPTICAL and S1 CorVegCSK. In order to discriminate the contribution from soil and vegetation to $\sigma^0$, the semi-empirical water cloud model (WCM) was applied (Attema and Ulaby, 1978). WCM uses one parameter, i.e. the plant water content, to describe vegetation. This parameter was estimated through empirical relationships from both the Normalized Difference Vegetation Index (NDVI) obtained from Landsat 8 (Pierdicca et al., 2013) and from CSK® acquisitions (Paloscia et al., 2014). The WCM introduces two parameters ($A$ and $B$) that are generally crop dependent, so that a rigorous determination of $A$ and $B$ would imply taking into account both the crop type and its pheno logical cycle. However, in operational soil moisture applications, the availability of updated information about the vegetation stage of growth and the crop type is not guaranteed. Hence it was decided to use the values of $A$ and $B$ found by Bindlish and Baros (2001) and valid for different land uses. Note that pixels where the plant water content is larger than 5 kg m$^{-2}$ are masked since, for high values of plant water content (and of NDVI, which generally increases with it), the sensitivity of the radar measurement to soil moisture is very low.

Before the SM retrieval, S1 images were calibrated, miltlooked and geocoded using the SARscape software and the SRTM (3 arcsec) DEM in order to obtain $\sigma^0$ maps (VV polarization) with a pixel size of 100 × 100 m$^2$. Next, SM values were estimated by inverting a forward backscattering model (Oh et al., 2004), which relates the soil state, represented by its surface roughness and SM (in volumetric terms) to the SAR measurements. In practice, the forward model was used to generate a lookup table in which each record contains three values: volumetric SM, surface roughness and the corresponding $\sigma^0$ predicted by the forward model. After the correction of the vegetation effect, a search in the lookup table was performed by the algorithm to retrieve SM values from S1 $\sigma^0$ data (Kim et al., 2012). The volumetric SM range of validity of the aforementioned forward backscattering model was set between 0.05 and 0.4 m$^3$ m$^{-3}$. During the retrieval, areas with a surface slope higher than 15° (associated with complex topography) and areas with a CORINE Land Cover 2012 classification corresponding to forests, urban areas, and water bodies were excluded to avoid poor quality estimates. The percentage of pixels masked out during the retrieval was about 50%. The SM maps were then filtered with a low pass filter with a kernel size of 5 × 5 pixels in order to reduce the noise. Therefore, the actual spatial resolution of the S1-derived SM maps is of the order of 500 × 500 m$^2$, whereas their pixel size is 100 × 100 m$^2$ (the same as the RZ–SD maps produced by Continuum).

As previously explained, the performance of S1 SM–DA was compared with respect to the performance of the assimilation of another high spatial resolution SM product, namely H08. H08 is obtained by disaggregating to 1 × 1 km$^2$ SM estimates retrieved from ASCAT acquisitions by using the TU Wien change detection algorithm (spatial resolution: 25 × 25 km$^2$, temporal resolution: 36 h; Wagner et al., 2013). The disaggregation is achieved by adopting a linear approach in which time-invariant coefficients are used. The coefficients are derived from long Advanced Synthetic Aperture Radar (ASAR) backscatter time series acquired by the ENVISAT satellite (Wagner et al., 2013). The theoretical framework behind this disaggregation approach is that of “temporal stability”, i.e. SM temporal dynamics are similar across different spatial scales (Vachaud et al., 1985). Even though H08 spatial resolution is coarser than the one of S1, both SM products can be considered at high spatial resolution for hydrological applications.

3 Methods: data pre-processing and assimilation algorithm

3.1 Data pre-processing of Sentinel 1-derived soil moisture products

As stated in the introduction, satellite-derived SM products are usually pre-processed before the assimilation to be physically consistent with the SM state variable simulated by the hydrological model. Since S1 SM maps were produced with the same pixel size as Continuum RZ–SD maps, their pre-processing did not require any resampling. Following Albergel et al. (2012), S1 surface volumetric SM estimates were initially normalized between 0 and 1 using their own minimum and maximum values admissible from the retrieval model (i.e. 0.05 and 0.4 m$^3$ m$^{-3}$) to be expressed in terms of SD. Then, the exponential filter developed by Wagner et al. (1999) was applied to the SD data to obtain an estimation of the RZ (the generated product is often referred to as the Soil Water Index – SWI). For each pixel, the filter smooths the time series of surface SD values to obtain an estimate of the RZ–SD. The filter needs one soil parameter to be calibrated: the characteristic time length ($T$), which characterizes the SM temporal variation within the RZ profile. Following the reasoning reported in Laiolo et al. (2015b), where the $T$ parameter was estimated according to hydrological considerations related to the soil of the selected study area (i.e. the Orba basin), a value of $T = 10$ days was used. Since SWI estimates were still affected by systematic bias with respect to modelled data, a third-order polynomial pixel-based cumulative distribution function (CDF) matching approach was adopted to remove it (Drusch, 2005; Reichle and Koster, 2004). This pre-processing allowed three S1-derived RZ–SD

bias-free products to be obtained, defined as S1 SWI-CDF NoCorVeg, S1 SWI-CDF CorVegCSK, and S1 SWI-CDF CorVegOPTICAL, depending on the initial S1-derived SM product that was used.

Following Drusch (2005), three additional products were obtained by applying directly the pixel-based CDF matching approach to the initial S1 surface SD values. The underlying assumption behind this approach is that the differences between the soil moisture datasets, referring to different soil layers/depths, can be seen – in a statistical framework – as differences between the moments of their statistical distributions. Even though these differences can be mainly ascribed to the first two moments, the CDF matching approach – being an advanced non-linear technique – allows for reduction also of the differences between the higher moments of distributions (in addition to the first two, i.e. mean and variance; Reichle and Koster, 2004). Therefore, the exploitation of the CDF matching approach as an observation operator transformed the S1 surface SD values into RZ–SD estimates comparable to that produced by the Continuum model. In this exercise, a third-order, pixel-based CDF matching approach was used. The CDF matching was computed by taking into account S1 and Continuum maps acquired at the same instant of time. During its implementation, any pixel excluded during the SM retrieval from S1 data was also masked out in the Continuum RZ–SD maps. It was decided to evaluate this alternative pre-processing (in which the exponential filter was not used) because the SWI was initially designed for denser time series (with respect to S1 A), such as those provided by the ASCAT sensor. Therefore, it is also worth evaluating the performances of S1-derived RZ–SD maps obtained when removing the temporal component (i.e. SWI) from the pre-processing. These products were defined as S1 CDF NoCorVeg, S1 CDF CorVegCSK, and S1 CDF CorVegOPTICAL. A total number of six S1-derived RZ–SM products was thus analysed. A graphical representation of the S1-derived SM products’ data pre-processing is shown in Fig. 2.

### 3.2 Data pre-processing of ASCAT-derived soil moisture products

Since ASCAT is a C-band scatterometer, the resulting H08 estimates, expressed in terms of SD, refer to the surface soil layer. Hence, they must be pre-processed before being assimilated into Continuum. Following Laiolo et al. (2015a, b) and Cenci et al. (2016a), H08 data were firstly resampled to the model’s spatial resolution (100 x 100 m²) using the nearest neighbour approach. Subsequently H08 RZ–SD estimates were obtained by applying the SWI (as for S1 data, a T value equal to 10 days was used). Finally, H08 RZ–SD bias-free estimates were obtained by applying a minimum–maximum (MinMax) correction (Laiolo et al., 2015a, b; Cenci et al., 2016a). Again, only morning SM maps (ASCAT acquisition time: 09:30 UTC) were used in the experiment, for a total number of images equal to 63.

In order to allow for a better understanding of the role of the spatio-temporal resolution in a SM–DA framework, three H08 datasets were created for the comparison with the S1-derived SM products. In the first dataset, hereafter defined as H08 SWI-MinMax, the aforementioned pre-processing chain was applied to all 63 H08 (morning) images available for the period under investigation. A second dataset, called H08 SWI-MinMax subset [1], was created by selecting the 12 H08 images closest in time with respect to the 12 S1 acquisitions from the H08 SWI-MinMax dataset. The third dataset, named H08 SWI-MinMax subset [2], was instead composed by selecting the 12 H08 images closest in time with respect to the 12 S1 acquisitions from the raw H08 dataset (i.e. before the pre-processing) that were then pre-processed as a “standalone” dataset (thus computing the SWI and the MinMax correction only by using the selected images). The main difference between the H08 SWI-MinMax dataset and H08 SWI-MinMax subsets [1] and [2] is thus in the number of images comprising the datasets (respectively 63 vs. 12). This comparison will thus help to better understand the role of the temporal resolution (i.e. number of images) in a SM–DA system in which high spatial resolution data are assimilated. The main difference between H08 SWI-MinMax subsets [1] and [2] is in the role that the temporal resolution plays in the pre-processing (i.e. computation of the SWI and MinMax Correction). Note that only 11 images out of 12 have been assimilated because the first one was used to initialize the SWI filter for both S1 and the H08 SWI-MinMax subset [1] and [2] datasets. Consequently 62 images from the H08 SWI-MinMax dataset were finally assimilated. The list of these images is reported in Table 1, as well as those belonging to the S1 dataset. A graphical representation of the H08 products’ data pre-processing is shown in Fig. 2.

<table>
<thead>
<tr>
<th>Image ID</th>
<th>Acquisition day S1</th>
<th>Acquisition day H08</th>
</tr>
</thead>
<tbody>
<tr>
<td>0*</td>
<td>10/10/2014*</td>
<td>10/10/2014*</td>
</tr>
<tr>
<td>1</td>
<td>22/10/2014</td>
<td>22/10/2014</td>
</tr>
<tr>
<td>2</td>
<td>03/11/2014</td>
<td>01/11/2014</td>
</tr>
<tr>
<td>3</td>
<td>15/11/2014</td>
<td>15/11/2014</td>
</tr>
<tr>
<td>5</td>
<td>09/12/2014</td>
<td>10/12/2014</td>
</tr>
<tr>
<td>6</td>
<td>21/12/2014</td>
<td>21/12/2014</td>
</tr>
<tr>
<td>7</td>
<td>02/01/2015</td>
<td>03/01/2015</td>
</tr>
<tr>
<td>8</td>
<td>14/01/2015</td>
<td>14/01/2015</td>
</tr>
<tr>
<td>9</td>
<td>26/01/2015</td>
<td>27/01/2015</td>
</tr>
<tr>
<td>10</td>
<td>07/02/2015</td>
<td>07/02/2015</td>
</tr>
<tr>
<td>11</td>
<td>19/02/2015</td>
<td>17/02/2015</td>
</tr>
</tbody>
</table>
3.3 Data assimilation algorithm

A Nudging assimilation scheme was used to carry out the DA experiment. According to the Nudging approach, a weighted term (innovation) is added to the model forecast (model SM estimate) to allow it to be closer to the observation. The innovation is defined as the difference between the observation and the model forecast. The latter is updated every time step \(\hat{x}_{t|t-1} = \hat{x}_{t-1|t-1} + K_t (y_t - \hat{y}_{t|t-1})\); in its scalar form, it ranges from 0 to 1. The weighting term is called gain \(G\); in its scalar form, it ranges from 0 to 1. If \(G = 0\), the observation is not taken into account during the analysis step (open loop – OL – mode); if \(G < 1\), the observation replaces the model forecast (direct insertion). Different approaches can be exploited to quantify \(G\), which should take into account the uncertainty of both datasets. An error characterization is thus required. However, normally \(G\) is a constant value in both space and time. This allows Nudging to be an efficient (computationally inexpensive) assimilation algorithm, which is of primary importance for operation applications (Walker and Houser, 2005). For this exercise it was not possible to have a reliable error characterization and, thus, it was not possible to correctly define \(G\). This was for different reasons: (i) absence of SM in situ monitoring networks in the study area to which S1-based SM estimates could be compared; (ii) lack of validation studies for the S1-derived SM products in the literature; and (iii) impossibility of applying a triple collocation-based approach due to the lack of other simultaneously observed, independent datasets of sufficient temporal length. In fact, the S1 time series is (still) short and, in addition, H08 images are not always acquired on the same day as S1 or at the same time of day (refer to Table 1). All of this implies that it was not possible to have a reliable error characterization and, thus, it was not possible to define \(G\) with standard procedures. In order to circumvent these issues, it was decided to run the assimilation with different values of \(G\) (from 0 to 1, with steps of 0.1), thereby enabling a sensitivity analysis of its influence on the assimilation performances (Brocca et al., 2010). Finally, it must be pointed out that (i) the exact period of the assimilation ranged from 20 October 2014 to 20 February 2015, to exclude the initialization period that was necessary to compute the SWI; (ii) the assimilation algorithm implemented within the Continuum model prevents the assimilation of SM observation pixels whose modelled LST is lower than 0°C (i.e. frozen soil conditions).

4 Results and discussion

The impact of the assimilation was evaluated in terms of impact on discharge predictions, by comparing the modelled discharge (before and after the assimilation) with the observed one recorded by the hydrometer of the Casal Cermelli station, located at the outlet of the catchment (refer to Fig. 1). Different statistical scores were computed: the Nash–Sutcliffe (NS) efficiency coefficient (Nash and Sutcliffe, 1970), the efficiency (Eff) of assimilation (Massari et al., 2015a) and the normalized error reduction (NER, Chen et al., 2011). NS ranges from \(-\infty\) to 1 (perfect model). Since it is biased toward higher flows, if it is calculated from discharge values in a logarithmic scale (NS\(_{LS}\)), it can be used to specifically evaluate the model performances for lower flows (Massari et al., 2015a). NS and NS\(_{LS}\) were both used to evaluate results. Eff, ranging from \(-\infty\) to 100, represents the percentage of worsening (Eff<0) or improvements (Eff>0) after SM–DA. NER is a dimensionless value, ranging between \(-\infty\) and 100, representing the percentage of root mean squared error (RMSE) reduction (or increment) after SM–DA. The higher the NER value, the higher the RMSE reduction. Note that the RMSE is also biased toward higher flows.

NS, NS\(_{LS}\), NER and Eff were computed in the assimilation time period to better understand the overall impact of the assimilations (Analysis 1). Since the time period under investigation is short, a seasonal analysis was not feasible. However, it was possible to identify two distinct sub-periods characterized by contrasting discharge regimes: higher flows (20 October–22 December, Analysis 2) and lower flows (23 December–20 February, Analysis 3). A specific event was
evaluated, i.e. the highest peak recorded in the period under investigation (Analysis 4). The statistical scores in Analyses 1–4 were computed on hourly discharge values. Finally, the SM–DA impact on the daily maxima (hourly) discharge values recorded in the period under examination was investigated (Analysis 5). Analyses 1–5 were also carried out for the H08 SWI-MinMax dataset and SWI-MinMax subsets [1] and [2]. Results are shown in Fig. 3.

Concerning S1 SM–DA, almost all the analyses revealed that the assimilation enhanced the model streamflow simulations. In spite of the fact that negligible effects were recorded on lower flows (Analysis 3 in Fig. 3), SM–DA considerably improved Continuum discharge predictions (Analysis 1 in Fig. 3), especially of higher flows (Analysis 2, Analysis 4 and Analysis 5 in Fig. 3). These results support previous findings stating that (i) in regions not characterized by a proper...
Rainy season, such as central Italy, SM–DA is able to improve discharge predictions in periods characterized by high SM temporal variability; and (ii) improvements particularly affect higher flows (Massari et al., 2015a; Cenci et al., 2016a).

Considering the focus of this work (i.e. improving flash flood predictions via SM–DA), Analysis 4 is interesting because of the exceptional nature, in terms of both intensity and duration, of the observed flood event (refer to the hydrograph in Fig. 3). Furthermore, a S1 observation was available before the event started, thus providing the best conditions to evaluate its impact on the assimilation. The high improvements provided by the assimilation are a demonstration of the importance of SM–DA for reducing the uncertainty of (flash) flood predictions.

Also, findings provided by Analysis 5 are relevant for the research objectives of this study. In fact, it allows the assessment of the discharge predictions in a hypothetical scenario of flash flood early warning systems, in which the discharge is likely to be predicted in terms of daily maximum values. Results highlighted the positive impact of the assimilation on peak flow predictions.

The negligible effects of SM–DA on lower flows (Analysis 3) can be explained as follows. Since it is well known that SM estimates are not reliable when frozen soil conditions prevail (Wagner et al., 2013), the SM–DA system was set up in order not to assimilate the pixels of the observations whose LST predicted by the Continuum model was lower than 0°C. By observing Fig. 4, which shows the average catchment values of LST computed at the time of S1 acquisitions, it can be inferred that – during the lower flow period – many of the pixels of the S1-derived SM products were not assimilated because, most of the time, the average catchment LST was lower than 0°C. Consequently, the low number of assimilated pixels produced a negligible effect of S1 SM–DA on discharge predictions. Future versions of the SM retrieval algorithm will account for a quality check process able to correctly deal with the “frozen soil condition” issue during the production of the surface volumetric soil moisture maps (e.g. using the Continuum LST data to mask out, during the SM retrieval, S1 pixels with a high probability of having frozen soil).

The sensitivity analysis of $G$ values showed that all the aforementioned outcomes were more pronounced as the weight assigned to the observations increased. This was due to the fact that, even though the model performed well in its OL realization (refer to NS values in Fig. 3), it overestimated the streamflow simulations (Fig. 5). SM–DA allowed the predictions to be improved by reducing the entity of the overestimation. This effect was due to the fact that, during the higher flow period, S1-derived SM products estimated values of SM lower than those estimated by the Continuum model (Fig. 6). Therefore, the assimilation reduced the values of the modelled SM, which in turn produces lower discharge values. The higher the weight assigned to the innovation, the higher the “correction” of the overestimation of the

Continuum model (of both SM and discharge values) on the occasion of the flood events.

Concerning the S1-derived products, the analyses pointed out the enhanced performances provided by the CorVeg-OPTICAL maps pre-processed with the CDF matching approach. This emphasized (i) the importance of the correction of the vegetation effect during the SM retrieval algorithm, also for slightly vegetated soils (CorVegCSK data provided positive results as well, outperforming NoCorVeg maps), and (ii) the fact that the CDF matching approach is particularly suitable for S1 data because it does not account for the temporal component that characterizes the SWI.

The results obtained by assimilating the H08 SWI-MinMax products showed almost similar but more marked trends with respect to S1 SM–DA (refer to Analyses 1–5 in Fig. 3). This difference was due to the higher temporal resolution of the ASCAT observations that significantly improved the Continuum predictions on higher flows. The relevance of the temporal resolution in the SM–DA was evident from the comparison between results obtained by the assimilation of the H08 SWI-MinMax dataset and results obtained by assimilating SWI-MinMax subsets [1] and [2]. Moreover, it was also clear from the comparison between the SWI-MinMax subset [1] and [2] performances (Fig. 3). In spite of the same number of images that were used, the effect of the temporal resolution in SWI-MinMax subset [1], included in the data pre-processing through the application of the SWI, was able to provide enhanced results. Results obtained by Analysis 4 (Fig. 3) were also particularly interesting to support the aforementioned conclusion. In that event, only one S1 observation was assimilated (15 November 2014) and, coherently, only one H08 image was used in the assimilation of SWI-MinMax subsets [1] and [2] (15 November 2014). However, three H08 images were exploited during the assimilation of the H08 SWI-MinMax dataset (15, 16 and 17 November 2015) and the performances of their assimilation were clearly enhanced with respect to the others. These results allow the argument that, even though the spatial resolution of the observations is an important component of a SM–DA system, the importance of their temporal resolution seems to be even higher. In fact, it must be considered that
Figure 5. Mean discharge values computed for each model realization (i.e. a model with a different $G$ value) during the time period corresponding to Analyses 1–4. For the colour legend assigned to each line, please refer to the legend of Figs. 2 and 3.

The results of the sensitivity analysis carried out on the $G$ value for H08 are coherent with findings obtained from S1.

5 Conclusions

The aim of this research was to evaluate the potentialities of S1 for improving flash flood predictions via SM–DA. To this aim, different S1-derived SM products, characterized by high spatial resolution ($\sim 500 \times 500$ km$^2$) and moderate temporal resolution (12 days), were generated. The maps were obtained by applying the same SM retrieval algorithm, with the notable difference that the effect of vegetation on the radar backscattering was corrected in a different way. The maps were assimilated within a time-continuous, spatially distributed, physically based hydrological model to evaluate the impact of the assimilation on discharge predictions, with a particular emphasis on high flow predictions. In order to understand the impact of the spatio-temporal resolution, results were compared with those obtained by assimilating ASCAT-derived SM products (H08), characterized by high spatio-temporal resolution ($1 \times 1$ km$^2$ and 36 h, respectively). The experiment was carried out in an Italian catchment prone to flash floods during the time period October...
2014–February 2015. Results showed that the assimilation of S1-derived SM enhanced the model discharge predictions, especially of higher flows. Results also highlighted the better performances provided by the satellite-derived SM maps in which the vegetation effect was corrected using optical data and when the pixel-based CDF matching was used as an observation operator for both RZ estimation and bias reduction/removal. The comparison with the H08 product pointed out that, despite high spatial resolution being able to provide an important contribution in SM–DA systems, the temporal resolution has the most crucial role.

We believe these findings are important and represent an innovative contribution within the SM–DA framework because S1-derived SM maps are still a relatively new product. Moreover, to our knowledge this is the first work published in an international journal dealing with their assimilation within a hydrological model to improve continuous streamflow simulations and flash flood predictions. However, results were obtained by analysing a short time period and should be supported by further research activities. To this aim, future analyses will be based on longer time series of S1 data (including S1 B data and/or different orbits), other study areas, different pre-processing/assimilation techniques, and different hydrological models. Additional studies aimed at (i) obtaining an accurate error characterization of the aforementioned S1-derived SM products and (ii) comparing such products with respect to low spatial resolution/high temporal resolution SM data (e.g. H16, SMOS and SMAP) are also recommended.

**Data availability.** The Continuum hydrological model is distributed following the European Union Public License. Further information about the model and the source code can be found at http://continuum.cimafoundation.org. Meteorological data used to run the model belong to the Piedmont Region and were made available through the Italian Civil Protection Department. H08 data were obtained through the Italian Civil Protection Department/H-SAF project. Sentinel 1 data were downloaded at https://scihub.copernicus.eu/. Landsat data were downloaded at http://earthexplorer.usgs.gov/. Project carried out using CSK® products, © of the Italian Space Agency (ASI), delivered under a license to use by ASI.

**Competing interests.** The authors declare that they have no conflict of interest.

**Special issue statement.** This article is part of the special issue “Hydro-related hazardous events and their consequences in the Mediterranean area: knowledge management adaptation”. It is a result of the 15th Plinius Conference on Mediterranean Risks, Giardini Naxos, Italy, 8–11 June 2016.
Acknowledgements. This work was carried out within the MIDA project (Italian acronym for Assimilated Hydrological Soil Moisture Maps) funded by the Italian Space Agency (ASI). The contributions of Marco Chini and Patrick Matgen were supported by the National Research Fund of Luxembourg through the CAOS-2 (INTER/DFG/14/02/CAOS-2) and HYDRAS+ (FNR-INTER/Belspo) projects. The three anonymous reviewers are also acknowledged for their valuable suggestions.

Edited by: Giuseppe Tito Aronica
Reviewed by: David Mason and two anonymous referees

References