

Article

Data Assimilation of Satellite Soil Moisture into Rainfall-Runoff Modelling: A Complex Recipe?

Christian Massari *, Luca Brocca, Angelica Tarpanelli and Tommaso Moramarco

Research Institute for Geo-Hydrological Protection, National Research Council CNR,
Via di Madonna Alta 126, Perugia 06128, Italy; E-Mails: luca.brocca@irpi.cnr.it (L.B.);
angelica.tarpanelli@irpi.cnr.it (A.T.); tommaso.moramarco@irpi.cnr.it (T.M.)

* Author to whom correspondence should be addressed; E-Mail: christian.massari@irpi.cnr.it;
Tel.: +39-075-501-4417; Fax: +39-075-501-4420.

Academic Editors: Guy J-P. Schumann, Richard Gloaguen and Prasad S. Thenkabail

Received: 30 June 2015 / Accepted: 1 September 2015 / Published: 8 September 2015

Abstract: Data assimilation (DA) of satellite soil moisture (SM) observations represents a great opportunity for improving the ability of rainfall-runoff models in predicting river discharges. Many studies have been carried out so far demonstrating the possibility to reduce model prediction uncertainty by incorporating satellite SM observations. However, large discrepancies can be perceived between these studies with the result that successful DA is not only related to the quality of the satellite observations but can be significantly controlled by many methodological and morphoclimatic factors. In this article, through an experimental study carried out on the Tiber River basin in Central Italy, we explore how the catchment area, soil type, climatology, rescaling technique, observation and model error selection may affect the results of the assimilation and can be the causes of the apparent discrepancies obtained in the literature. The results show that: (i) DA of SM generally improves discharge predictions (with a mean efficiency of about 30%); (ii) unlike catchment area, the soil type and the catchment specific characteristics might have a remarkable influence on the results; (iii) simple rescaling techniques may perform equally well to more complex ones; (iv) an accurate quantification of the model error is paramount for a correct choice of the observation error and, (v) SM temporal variability has a stronger influence than the season itself. On this basis, we advise that DA of SM may be not a simple task and one should carefully test the optimality of the assimilation experiment prior to drawing any general conclusions.

Keywords: data assimilation; satellite soil moisture; rescaling technique; observation error

1. Introduction

It is well recognized that soil moisture (SM) is central to the repartition of the mass and energy fluxes between the land surface and the atmosphere, and, as a consequence, it plays a key role in the hydrological cycle. For instance, in rainfall-runoff modelling, it has been demonstrated that the wetness condition of the catchment before a flood event strongly controls the severity of the response of the catchment in terms of runoff [1–4]. In other words, different initial soil moisture conditions can discriminate between catastrophic and minor effects for the same amount of rainfall. Therefore, an accurate estimate of the initial SM conditions is paramount for reducing the uncertainties in early warning flood forecasting models and may help to mitigate the flood hazard.

Hydrological models are valid solutions for obtaining accurate estimate of the SM conditions and have been used worldwide for diminishing the impact of floods on the population by issuing warnings with appropriate lead-time before floods occur. However, hydrological simulations are “imperfect” [5] in the sense that they contain uncertainties which are mainly related to (i) the quality and quantity of the hydrological data used to drive the models [6–8] as well as representativeness errors due to scale incompatibility [9]; (ii) the simple representation of the real physical processes leading to inevitable assumptions and simplifications and thus unavoidably imperfect approximations to the complex reality [5] and (iii) errors in parameter estimates which can result in huge errors in the model outputs [10]. As a result, there is an urgent need for techniques that effectively and efficiently assimilate new sources of information into the models to produce less uncertain hydrological predictions. Due to the increased spatial-temporal resolution and good accuracy of new satellite SM products [11–14], several studies have recently performed SM data assimilation (DA) into hydrological models [15–22].

Beside the choice of the most appropriate DA technique [23] and the type of sensor, DA involves a series of methodological assumptions and choices—not directly related to these above two issues—that might significantly affect the final result. Indeed, in the scientific literature, different studies characterized by a number of conflicting conclusions can be found, thus making a general interpretation of the utility of DA of SM in rainfall-runoff modelling difficult. In particular, different factors were found to affect the DA like the particular model structure, the bias correction technique, the model and observation error quantification, the spatial mismatching between catchment scale processes and observations, the catchment topography and climatology, just to cite a few. The result is that DA of SM in rainfall-runoff modelling still presents a number of controversial issues that seem to be not yet fully understood [17,22].

As an example, the model structure (*i.e.*, the part of the model that links the surface and the root-zone SM) may play an important role in the success of the DA experiment. Chen *et al.* [20], in assimilating *in situ* SM observations into the Soil and Water Assessment Tool (SWAT) in the Cobb Creek Watershed (341 km²) in USA through Ensemble Kalman Filter (EnKF) [24] (using both synthetic and real case experiments), concluded that: “Assimilation of actual surface soil moisture data had limited success in the upper layers only and was generally unsuccessful in improving stream flow prediction mainly due to the SWAT decoupling between surface and root-zone layer that limits the ability of the EnKF to

update the soil moisture states of deeper layers”. Brocca *et al.* [17], in a small catchment in central Italy, assimilated the Advanced SCATterometer (ASCAT, [11]) SM observations into two layer continuous hydrological model obtaining a great improvement in discharge prediction particularly for the floods occurring during dry to wet transition periods. The authors explored the sensitivity of the assimilation of both surface and root zone satellite SM observations to the model structure via two different assimilation experiments: (i) assimilation of the surface SM observations in the upper layer of the hydrological model, and, (ii) assimilation of the root zone SM observations (obtained by the application of the Exponential filter, [25]) into the lower layer of the model. They showed that the second solution provides much better results because of the larger impact of the root zone SM on runoff simulation for the investigated basin and of the weak and highly non-linear linkage between the SM states of the two layers which prevent the EnKF from updating the state of the deeper layer when the surface layer is assimilated.

The accurate estimate of model and observation error along with its structure is another crucial issue in DA of SM [26–28]. The importance of model error assumption was initially highlighted by Crow and Van Loon [26] using statistics calculated using filter innovations (*i.e.*, differences between model forecast mapped on the observation space and observations themselves). The authors highlighted that inappropriate model error assumptions can lead to circumstances in which the assimilation of surface SM can degrade the performance of open loop simulations. De Lannoy *et al.* [29] proposed using two ensemble indexes statistics previously employed in weather forecast analysis for calibrating the model error parameters and for giving an estimate of the model error, while in a recent paper Alvarez-Garreton *et al.* [22] used a methodology that maximizes the likelihood of observing historical streamflow data given the calibrated model parameters and the model parameter errors. Despite the effort made by these studies, a clear guideline for a proper model error assessment appears to be a long way off from reaching a final solution.

The estimation of an appropriate observation error for DA techniques is also another open question [30]. Different studies have tried to characterize the satellite SM error either by comparing satellite SM data with ground observations [31], or by providing an estimate of the error independently [30,32]. A quantification of the error associated with the retrieval is also supplied as ancillary data by data providers (e.g., the European Organisation for the Exploitation of Meteorological Satellites, EUMETSAT, within the project “Satellite Application Facility on Support to Operational Hydrology and Water Management”, H-SAF, supplies ASCAT SM data along with flags of different types and the associated uncertainties). The difficulty stems from the fact that satellite SM observations have in general higher uncertainty than ground measurements with error characterized by non-stationary characteristics [32], a much coarser spatial resolution, and are related to the top few centimetres of the soil. Additionally, they are affected by the error associated with retrieval algorithms. All these factors represent a challenge not only in the mere quantification of the error but also in the observation error spatial correlation [33], error structure [27] and compatibility between error statistics with the underlying hypothesis of the DA technique [27,28]. Indeed, in many cases, SM satellite observation error is treated as an independent Gaussian noise with fixed time variance, and this leads to a number of known weaknesses which can affect the optimality of the DA experiment [28]. Despite these issues, a final choice is needed in DA and this choice has been made by different methodologies like a simple preliminary calibration of the error [17], by the Triple Collocation (TC) method [21] as well as by using ancillary data provided with the SM retrieval [34]. It has to be noted that these choices may lead to a different error quantification and it is still not clear how they are propagated, in terms of optimality, into the DA experiments.

Moreover, model predictions and satellite observations are characterized by systematic differences between hydrological variables and this is particularly true for SM [31,35]. This bias is not desired in the framework of sequential DA techniques and should be removed prior to integrating observations into the models [36]. For bias removal, several potential strategies have been adopted in the past [16,37,40,42,43] occasionally leading to contrasting results in terms of their optimality in improving discharge predictions. Indeed, it has been shown that the use of suboptimal rescaling schemes [38] can degrade the performance of the stochastic DA and that an optimal rescaling technique (RT) can be identified for DA (*i.e.*, the TC-based RT). However, even if an optimal RT can be theoretically found, its application to reality might not be straightforward. For instance, the successful application of TC requires sufficiently large numbers of coincident data points from three independent time series and, within the analysis period, homogeneity of their linear relationships and error structures [39]—two requirements not always fulfilled.

Beside purely methodological issues, topography, spatio-temporal resolution and climatology are other important factors that may affect the results of the DA experiments. For instance, Brocca *et al.* [40] assimilated different soil moisture products (from remote sensing and global modelling) into rainfall-runoff modelling for five catchments located in Italy, France, Luxembourg and US using a simple nudging scheme. The authors found that the reliability differs according to the climatic region and the accuracy of satellite retrievals. In particular, no improvements were seen in mountainous regions and in regions characterized by the presence of snow. Moreover, in very dry climates, the very low SM temporal variability precluded an improvement of the rainfall-runoff simulations when SM was assimilated. Matgen *et al.* [19], in a tiny catchment in Luxembourg, used the Particle Filter [41] for assimilating both *in-situ* and satellite SM observations into the Bibmodel [42]. They found that, while the assimilation of *in-situ* data was successful, the use of ASCAT data did not lead to significant advantages and that the assimilation of SM is more efficient for flood forecasting during the transition periods (*i.e.*, spring and fall seasons). However, the authors did not clearly identify the reasons for the low performance of satellite data (*e.g.*, spatial mismatch between catchment and observations, observation error estimation) with respect to the better performance of *in-situ* data. Alvarez-Garreton *et al.* [37] assimilated SM data derived from the Advanced Microwave Scanning Radiometer (AMSR-E) into simple hydrological model in a semi-arid catchment through the EnKF. They found that the assimilation improved the prediction of both moderate and major floods; however, the improvement was higher and more consistent for moderate events. Recently, Corato *et al.* [43] used the Particle Filter [44] for assimilating different satellite SM observations (*i.e.*, two active and one passive products) with different spatial and temporal resolutions in the Severn catchment in UK trying to analyse under what conditions (*i.e.*, sensors characteristics and climatology) the assimilation of SM provided a successful result. The authors found that the performance of the DA depends on the period of the year, and active and passive sensors show different “optimal” assimilation periods probably due to the effect of vegetation on the sensor accuracy.

In summary, large discrepancies have been seen in the studies published so far likely due to the differences in the implementation of the DA techniques. The result is that there exist no well-accepted guidelines about how to optimally choose appropriate settings within a DA experiment based on data and basin characteristics, and the “optimal” DA configuration still appears to be a complex task.

In this paper, making use of an experimental study, we discuss and highlight how some of the choices regarding the observations error, the RT and the climatic and topographic conditions can affect the results of the DA experiments in run-off simulations. It is worth noting that this paper does not attempt

to provide general guidelines for “optimal” choices of the settings involved in DA studies; instead, we aim to present to the readers a simple but comprehensive and robust case study which highlights the main difficulties encountered in the setup of a DA experiment. The paper has the general objective to emphasize the reasons why further research in this direction is necessary in order to draw a more general approach for solving the open issues in hydrologic DA. Secondary objectives are: (i) To provide new evidence of the improvement in flood prediction from assimilating real satellite SM in an operational context; (ii) to test different RTs and their impacts in improving discharge prediction; (iii) to explore the effect of an erroneous choice of the observation error through a sensitivity analysis; (iv) to study the effects of different catchments characteristics (e.g., size, land use and morphology) on DA and (v) to analyze the effect of the climatology and SM conditions on the results of the DA.

On this basis, in order to analyze in deep the above targets five sub-catchments of the Upper Tiber River Basin with area ranging from 137–2000 km² are considered for which high-quality hydro-meteorological hourly observations are available in the period 2000–2013.

The paper is organized as follow: In Section 2, we give an overview of the study areas and the *in-situ* and satellite data. Section 3 details the procedure adopted in the analysis, describing the rainfall-runoff model, the assimilation techniques, the procedures for bias correction, the model and the observation error and the measure of performance. In Section 4 we present the results and some considerations are drawn in the discussions. In the last section, perspectives and conclusions are given.

2. Materials

2.1. Study Area

The study area is located in Central Italy and includes five sub-catchments of the Upper Tiber River basin (Figure 1). The catchments were selected by considering those characterized by a (i) good quality of rainfall, temperature and discharge data and (ii) size of the water storage not larger than 500 mm (about 150 cm of soil depth considering the average porosity in the area) in order to exclude catchments characterized by too much variance in depth between the satellite observations (few centimeters) and the depth involved in the rainfall-runoff transformation. The latter was identified by a preliminary analysis on the catchments for which data were available in the area and by using information of past studies. Finally, we selected: Chiani at Morrano (CHI-MOR), Marroggia at Azzano (MA-AZ), Niccone at Miglianella (NI-MI), Tevere at Ponte Felcino (TV-PF), Nestore at Marsciano (NE-MA). The main features of these catchments are shown in Table 1. Specifically, the drainage area ranges from 137 km² to 2040 km² and the mean catchment elevation ranges between 350 and 604 m a.s.l. The catchments located in the western part, Nestore and Chiani, are characterized by low slopes, and very similar soil and land use, whereas Marroggia lies in the Apennine reliefs, with higher elevations and slopes (Table 1). The climate is Mediterranean with mean annual precipitation of about 800 mm. Higher monthly precipitation values are generally observed during the autumn–winter period when floods caused by widespread rainfall normally occur. Mean annual temperature ranges from around 12–13 °C. Snowfall represents a low percentage of precipitation and is unusual and ephemeral at altitudes below 500 m a.s.l.

Infiltration rates were calculated based on the soil type. Based on this information, catchments located in the western part (NE-MA, CHI-MO, NI-MI) are characterized by a higher percentage of soil with

lower infiltration rate while MA-AZ has a higher percentage of soil with high infiltration rate due a greater presence of debris and conoids typical of mountainous areas.

Table 1. Physical parameters of the selected catchments (A = area, P = annual precipitation amount, El = mean elevation, Sl = mean slope) along with the most important calibrated model parameters in the period 2000–2009 (W_{max} = maximum water storage, K_s = parameter of drainage, K_c = parameter of evapotranspiration, η = lag time parameter of Equation (1)) and the time averaged model error standard deviation $\bar{\sigma}_m$. # SM-OBS4 pixels represents the number of the satellite pixels falling inside the catchment area, SM-OBS4 noise is the spatial average of noise SM information provided as ancillary data with SM-OBS4 product while T refer to value of the characteristic time length of the Exponential filter. α_1 , α_2 , α_3 represent the coefficients used to multiply the value of the parameters for obtaining their model error standard deviation, while σ_p is the rainfall error standard deviation.

Catchment	TV-PF	NE-MA	CHI-MO	MA-AZ	NI-MI
A (km ²)	2040	725	457	258	137
P (mm)	818	739	724	835	830
El (m)	521	350	404	604	452
Sl. (%)	26	17	19	29	25
Land Use (%)					
Woods	53.6	37.7	49.5	55.7	65.1
Croplands	9.4	3.8	4.7	4.6	7.9
Grasslands	35.2	53.5	43.8	35.8	26.6
Urban areas	1.7	5.0	2.0	3.9	0.4
Hydrologic Soil Group—Soil Conservation Service (%)					
High infiltration rate	5.0	1.7	0.6	9.4	3.2
Moderate infiltration rate	19.5	17.4	19.9	11.1	12.4
Low infiltration rate	74.4	81.0	79.4	75.6	84.4
Very low infiltration rate	1.0	0.0	0.1	3.9	0.0
Parameters of the Rainfall—Runoff Model					
W_{max} (mm)	497	380	388	327	372
K_s (mm/h)	20.8	28.8	29	40	21.8
K_c (mm/h)	0.87	0.84	0.66	1.03	0.89
η (-)	0.59	0.94	1.18	0.37	1.51
$\bar{\sigma}_m$ (%)	4	4.4	2	3.3	5
α_1 (%)	40	40	80	50	80
α_2, α_3 (%)	30	10	30	20	30
σ_p (mm)	0.75	0.5	0.25	0.75	0.25
Satellite Data					
# SM-OBS4 pixels	12	5	1	1	1
SM-OBS4 noise (%)	11	9.5	8	13.5	9.3
T (days)	48	87	67	62	68

2.2. In Situ Data

In the study area, a dense hydro-meteorological monitoring network has been operating for more than 25 years consisting of 84 rain gauges, 36 thermometers and 43 hydrometric gauges. The data are

recorded with a time interval of 30 minutes, here aggregated at hourly time steps. Rainfall and temperature recorded from January 2000 to December 2013 were considered in this study. Rainfall data are aggregated at catchment scale by the Thiessen polygon method while temperature data are simply averaged for the sensors located inside the catchment. As regards streamflow data, reliable rating curves provide discharge values to be used for calibrating and validating the procedure.

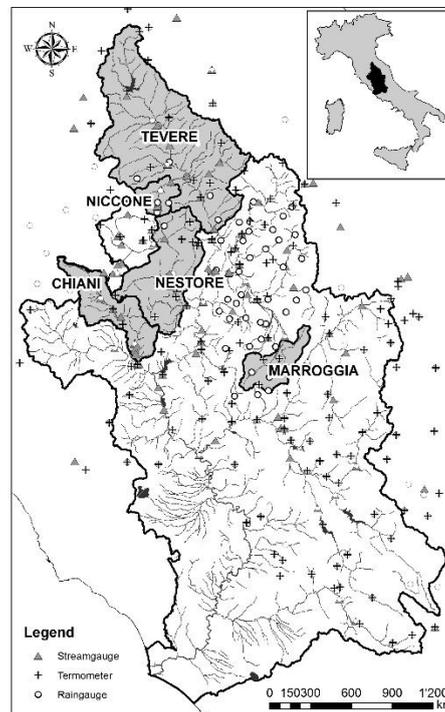


Figure 1. Study area, catchment locations and hydro-meteorological network.

2.3. Satellite Data

Satellite SM data from the scatterometer ASCAT were used in this study due their relatively good performance in Central Italy [17]. This allows to alleviate possible issues related to the general overall quality of satellite observations. In particular, the demonstrational product H25 SM-OBS-4 (Metop ASCAT Soil Moisture Time Series) distributed through H-SAF [45] project was used. The product cover all the globe and it is available from 2007. It is retrieved from the radar backscattering coefficients measured by ASCAT onboard the Metop A satellite using a change detection method, developed at the Research Group Remote Sensing, Department for Geodesy and Geoinformation, Vienna University of Technology. The relative SM data ranging between 0% and 100% are derived by scaling the normalized backscattering coefficients between the lowest/highest values corresponding to the driest/wettest soil conditions. The derived product represents the SM content in the first 5 cm of the soil in relative units between totally dry conditions and total water capacity. The unit is degree of saturation (%), but can be turned into volumetric units with the use of soil porosity information. In this study, the soil porosity information was not used since the product was rescaled using the model climatology (refer to Section 3.3 for further details). The spatial resolution of the product is 25×25 km, then resampled at 12.5 km and based on the WARP 5 grid (using a weighted mean with weighting coefficients computed according to the Hamming window function). Global coverage over Europe is achieved in ~ 1.5 days, while in Italy,

measurements are available about once a day. Along with the data, a number of ancillary information is provided, such as porosity, orbit direction, processing and surface state flag, soil moisture, noise, and so on.

For the analysis, only default values of the processing flags were used while frozen points were removed from the time series. Information on soil moisture noise varying between 0 and 100% was used as a reference for evaluating the accuracy of the product in the different catchments.

Given the different catchment extensions and spatial resolutions of the product, only data falling inside the catchment boundaries were selected and spatially averaged by using the inverse distance weighted (IDW) interpolation method with respect to the catchment centroid. Thanks to the property of temporal stability of SM [46] and as shown by other studies [47], the ASCAT product satisfactorily represents the ground SM state in the study area, also for small river basins, and, as a result, they can be considered spatially representative of the real SM variations. Associated soil moisture noise information was averaged as well in the same way and is shown in Table 1. As it can be seen, the values are very similar except MA-AZ for which soil moisture noise is slightly larger. Note that these values are used in the analysis only as a reference to estimate the uncertainty of the satellite observations and will not be directly used in the data assimilation experiments. Moreover, it is expected that when the satellite SM time series are averaged, their noise is somehow reduced, hence, there could be a non-straightforward relationship between the mean value of the SM noise in the catchment and the related real noise of the mean satellite SM time series (actually, this is true only for TV-PF and NE-MA since they contain more than one pixel in the catchment as it can be seen in Table 1).

3. Methods

3.1. Hydrological Model

The continuous and distributed rainfall–runoff model (MISDc) is a hydrological model developed for flood forecasting in the Upper Tiber River region at the hourly (or less) time scale. The lumped version of MISDc used in this study couples a Soil Water Balance model (SWB, [1]) to simulate the SM temporal pattern and a routing module [48] for transferring the rainfall excess to the outlet section of the catchment. The two models are linked through an experimentally-derived linear relationship between the soil potential maximum retention S of the Soil Conservation Service-Curve Number (SCS-CN, [49]) and the relative SM at the beginning of a rainfall event. It has to be pointed out that SWB simulates the soil moisture evolution considering a single soil layer (in a previous paper [17], the authors used a two layers model). Here, a single layer structure was used in order to reduce the complexity of the assimilation.

The SWB represents the main processes needed for SM simulation: infiltration, percolation and evapotranspiration. The processes are represented for infiltration through the SCS-CN approach, for drainage by a gravity-driven non-linear relationship and for actual evapotranspiration, by a linear relationship with the potential evapotranspiration calculated through a modified Blaney and Criddle method [50]. The SM simulated by the SWB is used to calculate the parameter S by means of an experimentally-derived relationship between S and SM [51].

Three different components contribute to generating discharge: the surface runoff, the saturation excess, and the subsurface runoff component. The first two are summed and routed to the catchment

outlet by the Geomorphological Instantaneous Unit Hydrograph (GIUH) while the subsurface runoff is transferred to the outlet section by a linear reservoir approach.

For the three components (*i.e.*, surface runoff, saturation excess and subsurface runoff), the lag time is evaluated through the relationship proposed by Melone *et al.* [52]:

$$L = \eta A^{0.33} \quad (1)$$

with L being the lag time in hours, A the area of the catchment (km^2), and η a parameter to be calibrated.

MISDC requires as input data rainfall and air temperature. The model outputs are both the total volume of runoff and the catchment mean relative SM.

In summary, MISDC, contains eight parameters to be estimated, *i.e.*, the maximum water capacity of the soil layer, W_{max} , the saturated hydraulic conductivity, K_s , the parameter controlling the fraction of drainage that transforms in subsurface runoff, ν , the pore size distribution index, m , the correction coefficient for the potential evapotranspiration, K_c , the lag-area relationship parameter η , the initial abstraction coefficient, λ , and the parameter a of the relationship between modelled SM and the S of the SCS-CN method.

3.2. Ensemble Kalman Filter

The EnKF was introduced by Evensen [24] as an alternative to the traditional Extended Kalman filter which has been shown to be problematic because of the strongly nonlinear dynamics. The EnKF is a sequential DA technique that uses the traditional update equation of the classical Kalman filter methods with the difference that, in the gain calculation, the error covariance is evaluated by using an ensemble of model states.

In particular, being $\mathbf{Y}(t)$ the vector of system states at time step t_k obtained via a generic model and \mathbf{Z}_k the observation vector at time t_k , then, the optimal updating of \mathbf{Y}_k , can be expressed as

$$\mathbf{Y}_k^{i+} = \mathbf{Y}_k^{i-} + \mathbf{K}_k \left[\mathbf{Z}_k + \mathbf{v}_k^i - \mathbf{H}_k \mathbf{Y}_k^{i-} \right] \quad (2)$$

where \mathbf{Y}_k^{i-} and \mathbf{Y}_k^{i+} refer to the forecast and analysis states for the i_{th} ensemble member, respectively, \mathbf{H}_k is the observation operator that maps the model states to the observations, \mathbf{v}_k is a synthetically generated error added to the observation \mathbf{Z}_k and represents the uncertainties of the observation process that is assumed to be a mean-zero Gaussian random variable with variance \mathbf{R}_k [53], and \mathbf{K}_k is the Kalman gain:

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} \quad (3)$$

In Equation (3) \mathbf{P}_k^- is the covariance matrix of the forecast error calculated by propagating an ensemble of model states (obtained by Monte Carlo or other ensemble generation methods) using the updated states from the previous time step.

$$\mathbf{P}_k^- = \frac{\left[\mathbf{Y}_k^- - \langle \mathbf{Y}_k^- \rangle \right] \left[\mathbf{Y}_k^- - \langle \mathbf{Y}_k^- \rangle \right]^T}{N-1} \quad (4)$$

A single deterministic EnKF prediction (*i.e.*, the “analysis”) is calculated by averaging model state predictions, \mathbf{Y}_k^- , at each time step across the N members of the ensemble (*i.e.*, the ensemble size). In Equation (4) $\langle \mathbf{Y}_k^- \rangle$ denotes the mean of \mathbf{Y}_k^- . For further details on the EnKF and its application to rainfall-runoff modelling, refer to Reichle *et al.* [16].

3.3. Filtering and Rescaling Techniques

While satellite SM observations represent only the top few centimeters of the soil column, the active soil layer involved in the rainfall-runoff transformation may be much deeper. Indeed, the hydrological model used in this study simulates the variation of the storage in a single soil layer that can assume depths spanning from the surface to the root zone depending on the catchment hydrology. To use surface observations in the DA experiments, a solution is to apply the exponential filter proposed by Wagner *et al.* [25] to the satellite time series in order to estimate the soil water index (*SWI*) of the root-zone. *SWI* has been widely used to represent SM at deeper layers based on surface observations through the following equation [54].

$$SWI_k = SWI_{k-1} + G_k [ms_k - SWI_{k-1}] \quad (5)$$

where ms_k is the surface SM observed by the satellite sensor at time t_k and SWI_k is the Soil Wetness Index representing the profile averaged saturation degree at time t_k . The gain G_k at time t_k is given by (in a recursive form):

$$G_k = \frac{G_{k-1}}{G_{k-1} + e^{-\frac{(t_k - t_{k-1})}{T}}} \quad (6)$$

where T is the characteristic time length and represents the time scale of SM variations to obtain the *SWI*. In this study, T was calibrated by maximizing the correlation between *SWI* obtained through ASCAT observations and modelled SM during the calibration period (*i.e.*, the common period between available satellite observations and model calibration results, 2007–2009).

Satellite SM observations and model predictions are generally characterized by considerable bias, hence, prior to the assimilation, this bias must be corrected in order to optimally merge the two time series [55]. The bias removal is in general carried out by rescaling the observations against model predictions [56]. In this study, three different rescaling schemes were implemented: linear regression (LR, [38]) mean and variance matching (VM, [17]) and cumulative distribution function matching (CDF, [57]). In the LR approach, the observations are regressed to the model predictions using a least squares regression technique, while for the VM the mean and the variance of the observations are matched to the ones of the model by using a simple formulation as described in [17]. VM involves the matching of only the first two moments of the probability distribution associated with the time series, hence as additional RT we used the CDF approach in which the matching is carried out for all the moments of the distribution. It is worth noting that other rescaling techniques could be used such as the CDF applied to the anomalies; however, the ones chosen for this work appear to be the most used in the literature. We have also to remark that we did not consider any effect of a bias seasonality as underlined by [58] and we applied the RTs to all data regardless of the season. The latter issue represents an additional complexity that may affect the results of the assimilation and will be the subject of future studies.

In principle, when the errors associated to observations and modelled time series are small with respect to the signals, these three RTs provide solutions close to optimal ones, however, this is hardly true in hydrological DA; hence, the three rescaling methods may present some problems [38]. As a result, an evaluation of their performances in a real DA experiment is beneficial for a better understanding of their behavior [34].

The setup of the rescaling schemes was carried out in order to guarantee the assimilation in an operational mode. That is, the parameters related to rescaling were “calibrated” prior to the assimilation in the calibration period. In this way, rescaling is always based on the past available data, while no “future” information is used. It has to be noted that a more optimal methodology can be used such as carrying out the calibration of the rescaling parameters at each assimilation step by using all the available past data. Although this is easily implementable, here we preferred to set a fixed-long (>2 years) time period in order to reduce additional influences from the calibration of the rescaling parameters on the analysis.

3.4. Model Error Representation and Observation Error

The determination of the model error is a crucial step in the DA experiment since model predictions and observations are merged through Equation (3)—which expresses the relative weight between the model and observation errors through the Kalman gain. If these errors are not properly represented, the determination of the \mathbf{K} can significantly affect the results of the assimilation. As remarked in the Introduction, the main sources of uncertainty in hydrologic models are the errors in the forcing data, the model structure and the incorrect specification of model parameters. A common way to represent these errors is by adding unbiased synthetic noise to forcing datasets (precipitation and evapotranspiration), model state variables and/or model parameters.

In this study, we adopted a multiplicative model error [59] for the rainfall data (assuming no autocorrelation in the rainfall error) and an additive perturbation for temperature [17]. In particular, rainfall was perturbed by a log-normally distributed, unit mean, spatially homogeneous and temporally uncorrelated multiplicative random noise σ_p , while for temperature a zero-mean normally distributed additive error was chosen with temperature standard deviation equal to σ_T . For the perturbation of model parameters, we first carried out a sensitivity analysis (not shown) in order to select for modelled SM the most sensitive parameters. Then, these parameters were perturbed with a normally distributed additive error with zero mean. In particular, only three parameters were selected: W_{max} , K_s and K_c , for which the relative standard deviations were expressed as percentage α of the absolute value of the parameter, *i.e.*, $\sigma_{W_{max}} = \alpha_1 W_{max}$, $\sigma_{K_s} = \alpha_2 K_s$, and $\sigma_{K_c} = \alpha_3 K_c$. The model error estimation was performed by varying the values of the percentages α_i ($i = 1,2,3$) along with σ_p (σ_T was assumed constant) in a way that two ensemble verification measures (Test #1 and Test #2 onward) commonly used in meteorology [29] were satisfied. That is, if the ensemble spread sp is large enough, the temporal mean of the ensemble skill $\langle sk \rangle$ should be similar to the temporal average of the ensemble spread $\langle sp \rangle$

$$\text{Test \#1 } \frac{\langle sp \rangle}{\langle sk \rangle} \cong 1 \quad (7)$$

and, if the observation is indistinguishable from a member of the ensemble,

$$\text{Test \#2 } \frac{\langle sk \rangle}{\langle mse \rangle} \cong \sqrt{\frac{(N+1)}{2N}} \quad (8)$$

where

$$\begin{aligned}
 \langle sp \rangle &= \frac{1}{T_p} \sum_{k=1}^T \left\{ \frac{1}{N} \sum_{i=1}^N (Q_k^i - \langle Q \rangle_k)^2 \right\} \\
 \langle sk \rangle &= \frac{1}{T_p} \sum_{k=1}^T \left\{ (\langle Q \rangle_k - Q_{obs,k})^2 \right\} \\
 \langle mse \rangle &= \frac{1}{T_p} \sum_{k=1}^T \left\{ \frac{1}{N} \sum_{i=1}^N (Q_k^i - Q_{obs,k})^2 \right\}
 \end{aligned} \tag{9}$$

In Equation (9), *mse* denotes the ensemble mean squared error, Q_k^i is the simulated discharge of the *i*th member at assimilation time *t_k*, $Q_{obs,k}$ is the observed discharge at time *t_k*, T_p is the length of the time period, while $\langle \rangle$ is the mean operator. Assuming σ_T constant, from different values of α_i ($i = 1,2,3$), σ_p and σ_T we obtained different ensembles and different values, $f_1(\alpha_1, \alpha_2, \alpha_3, \sigma_p)$ and $f_2(\alpha_1, \alpha_2, \alpha_3, \sigma_p)$ of Test #1 and Test #2, respectively. Then, the optimal selection of α_i ($i = 1,2,3$), σ_p was carried out by picking up the minimum of the following function:

$$F = \sqrt{\left(f_1(\alpha_1, \alpha_2, \alpha_3, \sigma_p) - 1 \right)^2 + \left(f_2(\alpha_1, \alpha_2, \alpha_3, \sigma_p) - \sqrt{\frac{(N+1)}{2N}} \right)^2} \tag{10}$$

The calibration of the model error standard deviations and σ_p was performed during 2000–2009, then the same perturbations were used during the validation period for generating the ensemble.

Test #1 and Test #2 serve to guarantee appropriate ensemble statistics. Indeed, if the spread of the ensemble is too large the filter tends to overfit the observations. By contrast, if the spread of the ensemble is too little, the filter tends to underuse the observations. Hence, Test#1 ensures that the ensemble is large enough while the Test #2 checks whether the truth is statistically indistinguishable from a member of the ensemble [17].

Another problem with perturbation methods is that perturbing restricted value variables (e.g., soil moisture, rainfall and parameters) may produce unrealistic values of such variables and bias in the ensemble that may affect the optimality of the assimilation [60]. To avoid this, we adopted the same techniques used by Turner *et al.* [60] in which restricted variables are perturbed by taking into account their domain of existence. This procedure guarantees realistic values and absence of a biased ensemble.

As remarked in the Introduction, a correct assumption of observation error is fundamental in sequential DA. In this study, we explore the sensitivity of the assimilation results as a function of the observation error for different RTs and different basins with different areas. Since the assimilation of the observation was carried out by considering the *SWI*, the observation error is referred to this variable. To this end, different observation error standard deviations σ_{obs} , were chosen ranging from 0.5%–20%. Moreover, as a reference we calculated the variance of rescaled *SWI* for VM, LR and CDF schemes and take the ratio between σ_{obs}^2 and σ_{SWI}^2 . In particular, as in [27], we considered an upper bound of the rescaled observation error for each catchment, and tested a selection of error variances within the bound. The rationality behind this is that if we express the rescaled observation (SWI_{ASCAT}^R as the sum of the “true” rescaled observation $\{SWI_{ASCAT}^R\}$ and the rescaled observation error.), and we assume error orthogonality, we can express the variance of the rescaled observation as $\sigma_{SWI_{ASCAT}^R}^2 = \sigma_{\{SWI_{ASCAT}^R\}}^2 + \sigma_{\epsilon_{ASCAT}^R}^2$

where $\sigma_{\varepsilon_{ASCAT}^R}^2$ represents the term R (from here onward not in bold since in the paper configuration it is a scalar) in Equation (3). Under this assumption $\sigma_{SWI_{ASCAT}^R}^2$ can be considered an upper bound for R . Specific settings for the model parameter errors and observation errors are discussed ahead in Section 4.2.

3.5. Performance Indexes

The performance of the DA experiments were evaluated using different scores. Each score was calculated by comparing the observed discharge time series (Q_{obs}) against (i) the discharge time series obtained during the calibration and the validation period, Q_c and Q_v ; (ii) the time series obtained by the mean of the different discharge ensemble members of the open loop simulations, (Q_{ol}) and (iii) against the mean of the different updated discharge ensemble members obtained in assimilation mode (Q_a).

In particular, the Nash-Sutcliffe efficiency index, (NS , [61]) was used in order to evaluate the agreement between Q_c , Q_v , Q_{ol} and Q_a (Q_{sim} in Equation (11)) and Q_{obs} :

$$NS = 1 - \frac{\sum_{t=1}^{Pa} (Q_{obs} - Q_{sim})^2}{\sum_{t=1}^{Pa} (Q_{obs} - \bar{Q}_{obs})^2} \quad (11)$$

where \bar{Q}_{obs} is the mean value of the observed discharge during the period of analysis Pa . NS was calculated for both the entire time series and for a number of selected events. The event selection was carried out by extracting the events with a continuous rainfall characterized by a total rainfall larger than 10 mm, and no rainfall in the preceding day. Only for discharges related to the selected events, we also calculated the error in volume by the following equation:

$$Err_v = \frac{\sum_{t=1}^{Pa} (Q_{sim} - Q_{obs})}{\sum_{t=1}^{Pa} Q_{obs}} 100 \quad (12)$$

To evaluate the benefit in DA, we calculated the efficiency, Eff , expressed as

$$Eff = 1 - \frac{\sum_{k=1}^{Pa} (Q_a - Q_{obs})^2}{\sum_{k=1}^{Pa} (Q_{ol} - Q_{obs})^2} \quad (13)$$

and represents the relative improvement in discharge estimation after the assimilation of SM. $Eff > 0$ denotes an improvement of the discharge estimation while $Eff \leq 0$ means no improvement or deterioration. Eff in Equation (13) was calculated by comparing Q_{obs} with the open loop ($Q_a = Q_{ol}$) and with respect to the validation ($Q_a = Q_v$).

To evaluate the specific improvement of the assimilation in low and high flows, we used two indexes as in [62]. The first one is a modified version of the Nash–Sutcliffe criterion and it is basically a calculation of the NS in logarithmic discharges in order to give more weight to low flows.

$$NS_{(\log Q)} = 1 - \frac{\sum_{t=1}^{P_a} [\log(Q_{sim} + \varepsilon) - \log(Q_{obs} + \varepsilon)]^2}{\sum_{t=1}^{P_a} [\log(Q_{obs} + \varepsilon) - \log(\bar{Q}_{obs} + \varepsilon)]^2} \quad (14)$$

while the second one is

$$ANSE = 1 - \frac{\sum_{t=1}^{P_a} (Q_{obs} + \bar{Q}_{obs})(Q_{sim} + Q_{obs})^2}{\sum_{t=1}^{P_a} (Q_{obs} + \bar{Q}_{obs})(\bar{Q}_{obs} + Q_{obs})^2} \quad (15)$$

In particular, we calculated the differences $\Delta_{NS(\log Q)}$, Δ_{ANSE} between the results obtained after and prior the assimilation. In Equation (14), ε is arbitrarily chosen as a small fraction of the inter-annual mean discharge (e.g., $\bar{Q}_{obs}/40$) and is introduced to avoid problems with nil observed or simulated discharges.

As additional score, we calculated the Normalized Root Mean Squared Error (NRMSE, [22]):

$$NRMSE = \frac{\frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{P_a} \sum_{k=1}^{P_a} (Q_a^i(t) - Q_{obs}(t))^2}}{\frac{1}{N} \sum_{i=1}^N \sqrt{\frac{1}{P_a} \sum_{k=1}^{P_a} (Q_{ol}^i(t) - Q_{obs}(t))^2}} \quad (16)$$

NRMSE greater than one denotes a deterioration of the results while the opposite means improvement. As highlighted by Alvarez-Garreton *et al.* [22], the *NRMSE* provides information about both the spread of the ensemble and the performance of the ensemble mean, which can be considered as the best estimate of the ensemble prediction.

By changing the time window where *NRMSE* is calculated, it is possible to evaluate the benefit of the assimilation in time. To this end, *NRMSE* was calculated over a time window of 90 days in order to analyse a possible seasonal effect on the assimilation of SM.

4. Results

4.1. Model Calibration and Validation

Two distinct periods were considered in the analysis: A calibration period of 10 years ranging from 1 January 2000 to 31 December 2009 and a validation period of four years where we run the model in assimilation mode. The calibration of the hydrological model was performed by a standard gradient-based automatic optimisation method (“*fmincon*” function in MATLAB®). The calibrated parameters (only W_{max} , K_c , K_s and η) are shown in Table 1 while the results in terms of *NS* of the calibration are reported in Table 2. A general good behaviour (in calibration) can be seen in the model for all the catchments.

The values of the parameters denote a homogenous behaviour for all catchments for W_{max} and K_c while different values of K_s and η were obtained for MA-AZ. The results obtained in validation—we are referring to a single deterministic run using the calibrated parameters of the previous step—provide relatively good performance with *NS* always above 0.6 (an example of the discharge time series in validation for TV-PF during 2013 is shown in Figure 2b). The analysis of the results was carried out also

by calculating the performance of the model on a number of selected flood events (Table 3) extracted from the time series as specified in Section 3.5. The results for validation are generally satisfactory ($NS > 0.7$ and error in volume lower than 30%) except for TV-PF and NI-MI. The poor performance of the model in these two cases was mainly determined by a significant mismatch existing between the model simulation and the observed discharge for the most important flood event which occurred in November 2012 (see the results for TV-PF in Figure 2c). Indeed, in this case, a levee overtopping occurred prior to the measurement section determining a flood peak much lower than expected (*i.e.*, there was a significant reduction of the river discharge upstream to the measurement section) and, as a consequence, a strange shape of the flood hydrograph formed.

Overall, the model performances can be assumed satisfactory considering the hourly temporal resolution, and thus the experiment can be considered a good test for the analysis of satellite SM DA.

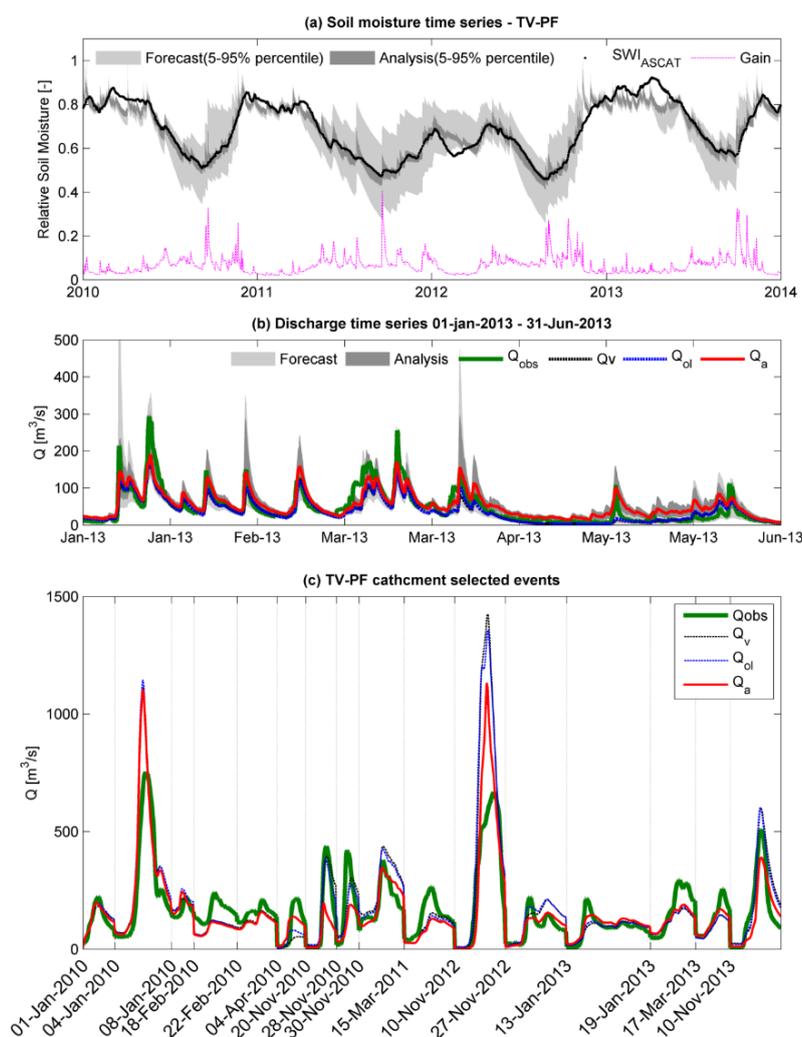


Figure 2. Results of the data assimilation experiment for Tiber River at Ponte Felcino using LS as RT and an observation error equal to 7%. **(a)** Plots the satellite SM time series along with the ensemble forecast and analysis (5%–95% percentiles). In purple, it is also shown the Kalman gain. **(b)** Plots a small part of the discharge time series: the plot of the entire time series is not shown because of problems of visualization. In the panel, Q_{obs} is the observed discharge, Q_v is the discharge obtained in validation, Q_{ol} is the ensemble mean and Q_a is the updated discharge obtained in assimilation. **(c)** Plots the discharge values for the selected flood events.

Table 2. Results of the calibration-validation procedures and of the data assimilation experiment for the five selected catchments. NS_c is the Nash Sutcliffe efficiency index obtained in calibration from 2000–2009, NS_v refers to the same index in validation during 2010–2013, while NS_{ol} is the Nash Sutcliffe calculated in the same validation period on the ensemble mean. On the right part of the table, σ_{obs} refers to the observation error for which the data assimilation provides the best performance in terms of Nash-Sutcliffe (NS_a). Eff_v and Eff_{ol} refer to the results in terms of efficiency in the validation and in the ensemble mean, respectively. Gray cells in the table denote a deterioration with respect to the open loop simulations.

		NS_c	NS_v	NS_{ol}	NS_a	Eff_v (%)	Eff_{ol} (%)	σ_{obs}
TV-PF	VM				0.79	39.64	35.86	0.05
	LR	0.80	0.65	0.67	0.80	42.33	38.72	0.05
	CDF				0.75	29.53	25.11	0.05
NE-MA	VM				0.91	29.30	31.56	0.20
	LR	0.88	0.87	0.87	0.91	28.72	31.00	0.20
	CDF				0.90	23.49	25.93	0.20
CHI-MOR	VM				0.87	24.64	31.66	0.09
	LR	0.76	0.83	0.81	0.88	26.03	32.93	0.07
	CDF				0.87	21.09	28.44	0.05
MA-AZ	VM				0.70	2.06	-13.30	0.20
	LR	0.77	0.69	0.74	0.71	3.67	-11.44	0.20
	CDF				0.70	3.35	-11.81	0.20
NI-MI	VM				0.79	39.58	41.67	0.03
	LR	0.85	0.65	0.64	0.79	39.27	41.37	0.05
	CDF				0.77	34.23	36.50	0.03

Table 3. Results of the data assimilation experiment for the five selected catchments on the selected flood events (number of events N_{ev}). NS_{val} is the Nash Sutcliffe efficiency index calculated on the selected flood events from 2000–2009, NS_{ol} refers to the same index calculated on the selected flood events on the ensemble mean, σ_{obs} refers to the observation error (the same of Table 2) while NS_a is the Nash-Sutcliffe efficiency index calculated after data assimilation on the selected flood events. Eff_v and Eff_{ol} refer to the results in terms of efficiency with respect to the validation and to the ensemble mean, respectively, calculated for the selected flood event. Gray cells in the table denote a deterioration with respect to the open loop simulations.

		NS_v	Err_{vol} (%)	NS_{ol}	NS_a	$Err_{vol(a)}$ (%)	Eff_v (%)	Eff_{ol} (%)	σ_{obs}	N_{ev}
TV-PF	VM				0.64	-2.56	52.55	50.18	0.05	16
	LS	0.25	26.09	0.29	0.67	-5.49	56.08	53.89	0.05	16
	CDF				0.56	2.84	41.53	38.61	0.05	16
NE-MA	VM				0.91	14.23	37.09	38.71	0.20	16
	LS	0.85	19.45	0.85	0.91	15.14	36.69	38.32	0.20	16
	CDF				0.90	16.04	32.83	34.56	0.20	16
CHI-MOR	VM				0.87	-1.18	27.11	37.54	0.09	24
	LS	0.83	1.69	0.8	0.88	-1.64	28.02	38.32	0.07	24
	CDF				0.87	-1.27	22.05	33.21	0.05	24

Table 3. Cont.

		NS_v	$Err_{vol}(\%)$	NS_{ol}	NS_a	$Err_{vol(a)}(\%)$	$Eff_v(\%)$	$Eff_{ol}(\%)$	σ_{obs}	N_{ev}
	VM				0.71	-7.31	7.69	-34.14	0.2	17
MA-AZ	LS	0.68	-0.1	0.78	0.72	-6.83	10.83	-29.57	0.2	17
	CDF				0.71	-5.25	8.51	-32.95	0.2	17
	VM				0.75	-9.65	55.56	57.98	0.03	30
NI-MI	LS	0.44	-3.81	0.41	0.75	-12.83	54.5	56.98	0.05	30
	CDF				0.7	-9.85	46.16	49.09	0.03	30

4.2. Ensemble Generation and Data Assimilation Experimental Setup

Model parameter errors were calibrated (in the period 2000–2009) by running the open-loop (*i.e.*, the run without assimilation of SM) with 50 ensemble members by varying the coefficient α_1 between 30 and 80% along with α_2 and α_3 between 0% and 30%. At the same time, σ_p varied between 0.25 and 1, while σ_T was assumed constant and equal to 3 °C. Figure 3 shows the value of the function F of Equation (10) in the $\alpha_1 - \sigma_p$ domain for different values of α_2 , α_3 for NE-MA catchment (the other catchments lead to the same conclusions and are not shown for brevity). It can be seen that there is a problem of equifinality in which different combinations of error parameters lead to a similar value of F . The main responsible of this behavior is α_1 which controls the perturbation of the storage parameter W_{max} . In order to make a selection, we ran the same assimilation experiments by using different optimal and close-optimal solutions (e.g., 2/3 solutions for each catchment) obtaining very similar results with some small differences in some cases. It is likely that the model-non-linearity plays a role in this case and allows obtaining similar model error covariance and in turn similar results. Finally, since a choice has to be made (e.g., think to the application of the DA at operational level), we selected the best combination of the error parameters as described in Section 3.4. Such values for all the catchments are shown in Table 1. In addition, it has to be stressed that the selection of the quadruplet was made during the calibration period, and hence, this does not guarantee that the quadruplet is optimal also during the validation phase (2010–2013). This eventuality was tested for all catchments and it was found that the tests are close to the required values except in some cases (e.g., NE-MA yields 0.68 *vs.* 1 for test #1 and 0.63 *vs.* 0.71 for test #2). Another issue that might potentially affect the methodology of the selection of the best error parameters is the reliability of the stream flow observations. Indeed, there is the assumption that their random error is much lower than the modelling error. Although the quality of the observations is very high—since they were successfully tested and cross-verified by many other papers in the past (e.g., [17,40,63])—we cannot guarantee this is always true. The above results suggest the complexity of the issue, which would require a separate and detailed analysis specifically addressed to the problem of the estimation of the model error covariance.

Table 1 also shows the mean model error standard deviation $\bar{\sigma}_m$ which ranges between 2% and 5% and does not seem to have any relation with the performance of the model in the validation period. The ensemble of model states generally contains a look similar to the one of TV-PF presented in Figure 2. It has to be noted also larger ensemble sizes were used (e.g., 100, 500) leading to the same overall conclusions.

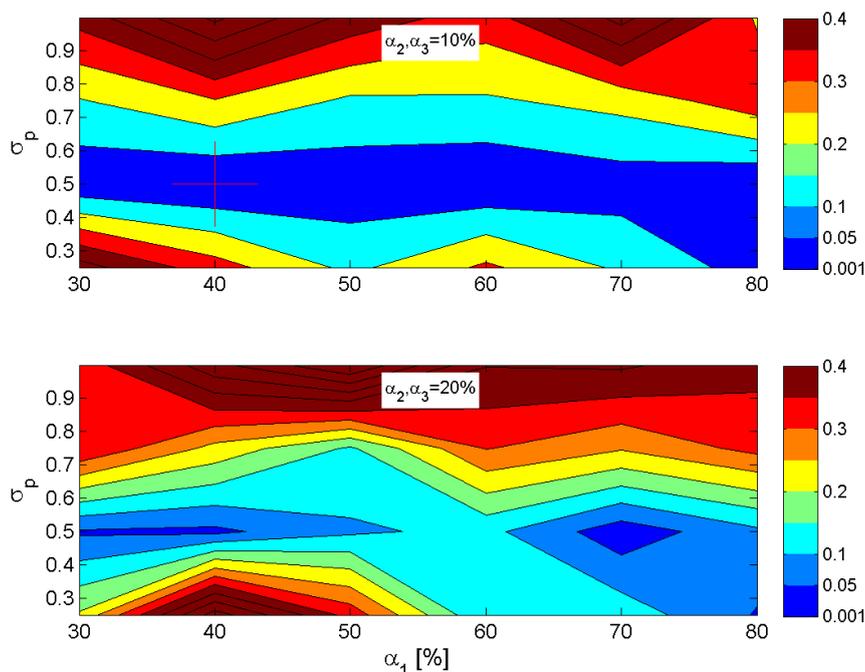


Figure 3. Contour plot of the function F of Equation (10) as a function of the standard deviation of the precipitation σ_p , the percentage α_1 of the perturbation of the maximum water storage W_{max} , and the percentages α_2 , α_3 of the parameters of drainage, K_s and evapotranspiration K_c , respectively (period 2000–2009, catchment NE-MA). The optimal value of these standard deviations is identified by the red cross and corresponds to the minimum value of the function F representing the error made in the identification of the optimal value of the tests. It can be clearly seen a problem of equifinality in the figure with darker blue areas (smaller F) covering a large part of the domain.

In Equation (2), v_k (the error representing the uncertainties of the observation process with variance R and hereafter as σ_{obs}) was made variable between 0.5% and 20% (*i.e.*, 0.005, 0.01, 0.03, 0.05, 0.07, 0.09, 0.12, 0.15, 0.18, 0.2) and the corresponding R was then compared against the variance of the rescaled SWI for different RTs as explained in Section 3.4. Assuming that the range of volumetric soil moisture variability is equal to 0.33 [1], these values correspond in volumetric terms to a range variable between 0.002 and 0.06 m^3/m^3 and are in the range of those found in the area by [31].

The determination of the characteristic time length T and the calibration of the rescaling parameters were performed in the period 1 January 2007 to 31 December 2009, then these values were used for filtering and rescaling as the observations from the satellite were made available. The objective was to reproduce a “real time operational” application of DA by using only information from the past for flood prediction. The results of the calibration of the parameters T are shown in Table 1 while the parameters related to the RTs are not shown. T values are consistent with those found by [64] in the area with larger values obtained for NE-MA. The filter setup with the above configuration was used for all the five selected catchments.

4.3. Data Assimilation Experiment

Figure 4 summarizes the results of the DA experiments for the selected catchments ordered from the largest (top) to the smallest (bottom) for the different RTs and different observation errors. It can be seen

that the assimilation generally provides positive efficiency indexes for a large range of observation errors except for MA-AZ. The improvements are significant for TV-PF, NI-MI and CHI-MO with maximum performances (Eff_a between 30% and 40%) obtained for σ_{obs} ranging between 0.03 and 0.07. NE-MA shows similar improvements but only after $\sigma_{obs} = 0.09$. The same efficiency index as a function of the ratio between the variance of the error and the rescaled observation $\sigma_{obs}^2 / \sigma_{SWI^R}^2$ (top axis of the panels) show similar results with TV-PF, CHI-MOR and NI-MI having higher performance with $\sigma_{obs}^2 / \sigma_{SWI^R}^2$ between 0.1 or less, and 0.5. Consistently, NE-MA shows improvements for $\sigma_{obs} / \sigma_{SWI^R}$ larger than 0.5 and best Eff for values of $\sigma_{obs}^2 / \sigma_{SWI^R}^2$ greater than one. It has to be pointed out that the latter results violate the assumption made in Section 3.4, which assumes the variance of the rescaled observation as an upper bound value of the variance of the observation error; hence, making this assumption may not lead to an optimal solution. For MA-AZ, Eff_a is always lower than zero although it increases by increasing the observation error.

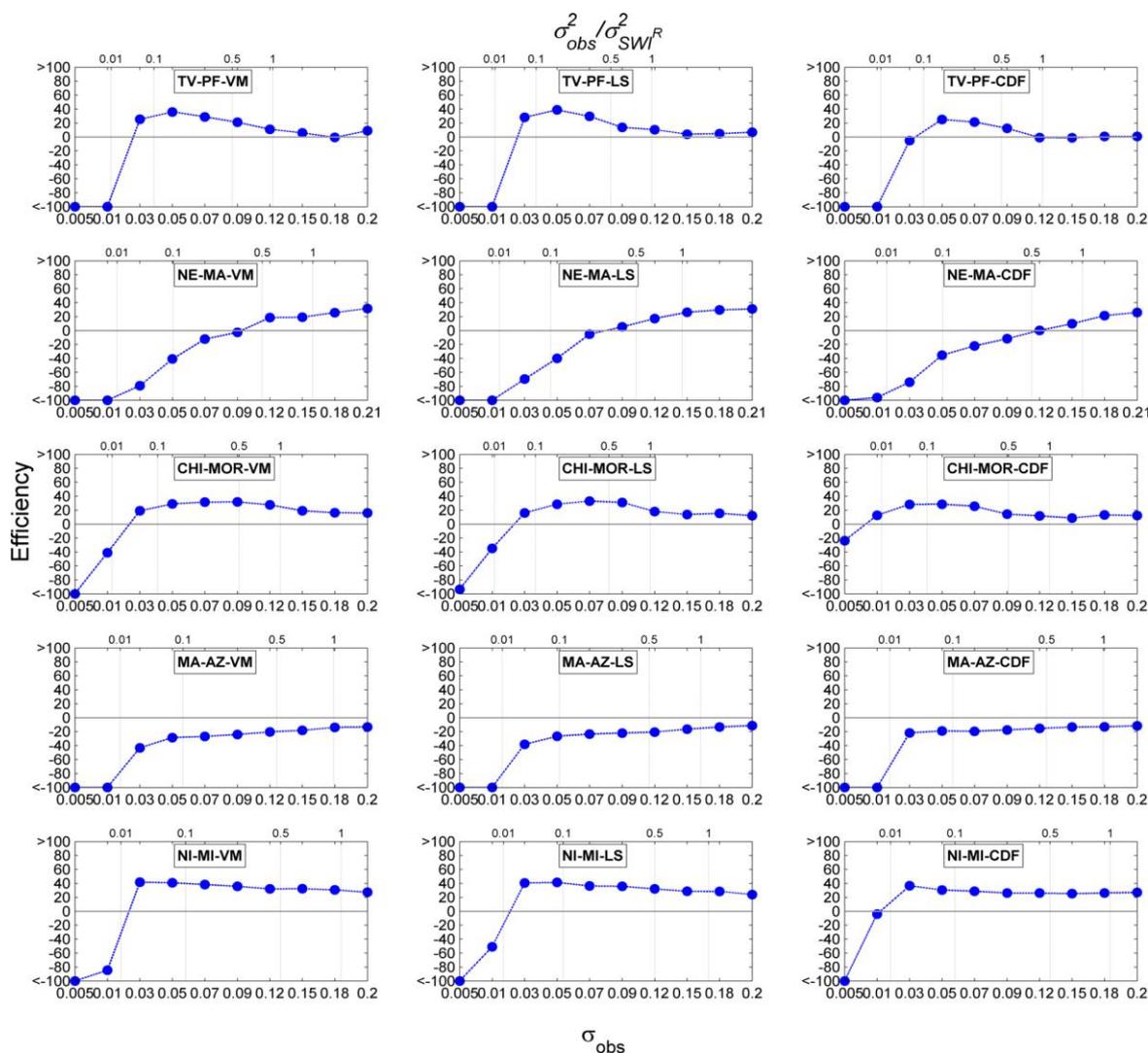


Figure 4. Efficiency index calculated with respect to the open loop simulation for the five catchments for the three RTs as a function of the observation error. σ_{obs} is the observation error standard deviation, σ_{obs}^2 is the variance observation error, $\sigma_{SWI^R}^2$ is the variance of the rescaled observation. Values below zero (black line) denote a deterioration of the results.

For a better understanding the best configurations obtained in DA for the five catchments, we summarize in Table 2 the experiments characterized by an optimal choice of the observation error (assuming it was possible its identification). Among other performance scores, the table shows the NS obtained during the assimilation run, NS_a , and Eff calculated by using Q_{ol} and Q_v . It can be seen that except for CHI-MO the optimal observation error does not change too much from one RT to another and lower NS and Eff are generally obtained with CDF.

The same analysis of Table 2 was carried out on the selected flood events and it is presented in Table 3. It can be seen a significant improvement of the flood prediction for TV-PF, CHI-MOR, NE-MA and NI-MI while for MA-AZ the assimilation provides a deterioration of the results with respect to the open loop. At the same time, the error in volume reduces for TV-PF, NE-MA and CHI-MO while it increases for the MA-AZ and NI-MI for all the RTs.

To assess the effectiveness of the assimilation specifically for low and high flows we calculated the gain ($\Delta_{NS(logQ)}$, Δ_{ANSE}) obtained in the $NS(logQ)$ and $ANSE$ indexes with respect to the open loop simulation for the three RTs considering the optimal choice of observation error (the same presented in Table 2). The results are plotted in Figure 5 and show significant improvement for both indexes only for NI-MI. For TV-PF, DA succeeds only for high flows and show contrasting results between the different RTs. MA-AZ always provides poor results while CHI-MO and NE-MA have a general consistent behavior. Overall, the improvement for high flows is more consistent.

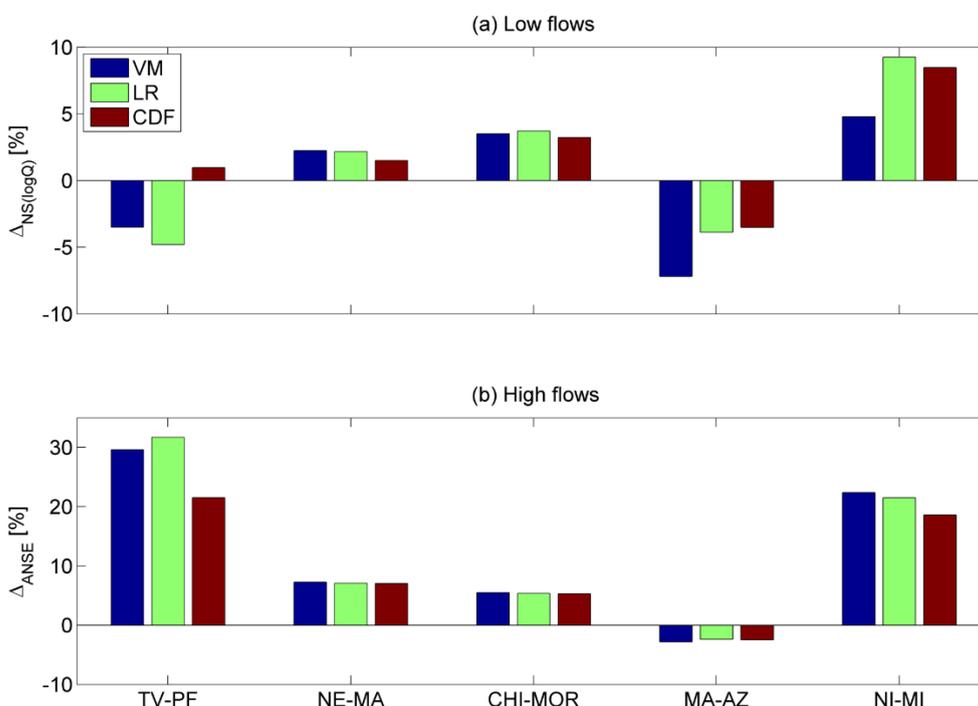


Figure 5. Differences ($\Delta_{NS(logQ)}$, and Δ_{ANSE}) between performance scores $NS(logQ)$ and $ANSE$ obtained prior and after the assimilation of SM for the five selected catchments. While (a) $NS(logQ)$ is well suited for highlighting the performance in the reproduction of low flows, (b) $ANSE$ is used for characterizing the agreement of high flows. The related observation errors are the same as selected in Table 2.

An example of the results of the DA for TV-PF (LR) is shown in Figure 2. Figure 2a shows the reduction of the uncertainty with respect to the open loop simulation of SM (*i.e.*, a reduction of the ensemble spread) determined by the assimilation of the SWI and the value of gain parameter K in time. As it can be seen, K shows a general increase in the last parts of the year meaning that the filter is disposed to using more observations during dry-to-wet periods. Figure 2b contains the open loop simulation, the validation run and the updated discharge during 2013 (we show only one year for a clearer visualization), while Figure 2c focuses on the selected flood events for the entire validation time period (*i.e.*, 2010–2013). It can be seen that while certain flood events (4 April and 30 November 2010, 27 November 2012) are better predicted with SM assimilation, others, even if close to the previous ones (20–28 November 2010), are inaccurate.

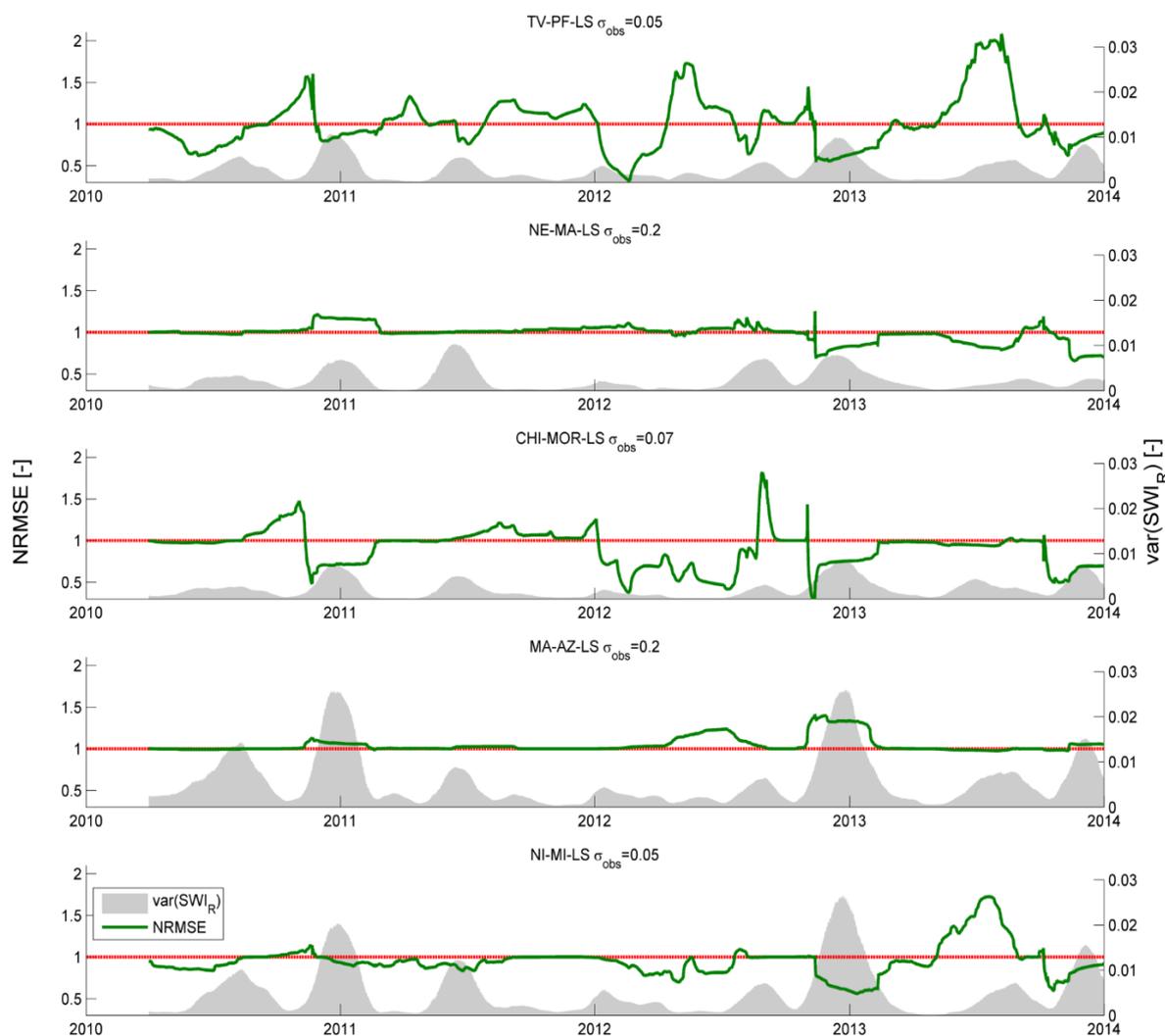


Figure 6. Pattern of the Normalized Root Mean Squared error (*NRMSE*) and variance of the rescaled Soil Water Index (rescaled with LR technique), $\sigma_{SWI^R}^2$ for the five selected catchments calculated with a time window of 90 days. Values above the red horizontal line (*NRMSE* = 1) denote deteriorations. Periods of increased $\sigma_{SWI^R}^2$ generally correspond with improvements except for MA-AZ.

Finally, to analyze the effect of the seasonality, we calculated the *NRMSE* within a time window of 90 days as a function of time (Figure 6) and we relate it with the SM temporal variability (by calculating

the variance of the rescaled *SWI* in the same time window). Figure 6 shows the results obtained for the LS RT and the optimal choice of the observation error (as above). Although *NRMSE* shows some random characteristics due to the occurrence of flood events of different magnitude during the year—which is much more pronounced for relatively small basins—a similar pattern can be observed for all catchments (except MA-AZ) which highlights positive performances ($NRMSE < 1$) of the assimilation during periods of higher SM temporal variability and negative during periods with smaller temporal variability ($NRMSE > 1$).

4.5. Discussion

The performed analysis highlights a good general behavior of the hydrological model for all the selected catchments both in calibration and in validation (Table 2). From the hydrological modelling point of view, MA-AZ is characterized by a larger value of the coefficient of drainage, K_s , and a much quicker response (smaller η) with respect to the catchments located in the western part of the study area (NE-MA, CHI-MO, NI-MI) while TV-PF shows an intermediate behavior. In particular, the larger K_s for MA-AZ is somehow consistent with the higher infiltration rate value (*i.e.*, MA-AZ is located in the Apennines relief with high permeable rocks, see Table 1) while the smaller η is likely due to the difficulty of the model to reproduce the larger baseflow component for this catchment (also noted for TV-PF).

The overall results of the assimilation are positive (similar results were obtained by [17,63] in the same area). Improvements of the performance after the assimilation are obtained for all catchments for proper configurations of observation error and RT, while for MA-AZ no pairs (σ_{obs} —RT) have been found to lead to an enhancement of the performances. The reason for the poor assimilation results of MA-AZ might be due to three main causes. First, the expected poor quality of the satellite observations (the catchment is located in a mountainous area where ASCAT may suffer from topography issues)—this problem is confirmed by the higher SM-OBS4 soil moisture noise with respect to the other catchments as shown in Table 1. Second, an overestimation of the model error covariance, which prevents the filter from obtaining an improvement in the range of the chosen observation errors. Finally, and most likely, it might be due to the runoff generation mechanism, which is different from the other catchments. Indeed, the higher permeability of MA-AZ catchment determines that water stored in the soil is less retained by surface tension (*i.e.*, gravity is dominant) and water is released in large quantities to the subsurface leading to a larger contribution of base flow and subsurface runoff. This contribution is not negligible with respect to the total runoff in this catchment and is much less controlled by SM variation at the surface (*i.e.*, where we have information of SM from the satellites). On the other hand, catchments located in the western part of the Umbria Region are characterized by less permeable soils with higher retention properties and show larger sensitivity to surface SM. This second hypothesis was tested through a synthetic experiment (not shown) which showed that the efficiencies obtainable for the other catchments were much larger than those obtainable for MA-AZ. Although DA provides positive results for most of the catchments, the improvements also contain a number of contradictions, which are highlighted below.

In the calibration of the model error, it was found that different parameter perturbations lead to the same values of the selected ensemble statistical tests (Figure 3) and the model error does not seem to have a relationship with the model performance (e.g., compare the model error of NI-MI and MA-AZ in Table 1 and their performance in Table 2). This highlights some difficulties in the identification of the optimal model error to perform the DA. This issue may also be exacerbated for a distributed model

characterized by a large number of model parameters. It must be said, however, that the entire procedure for the quantification of the model error covariance and the satisfaction of the ensemble statistical tests might be affected by the presence of random errors in the observed discharge time series. As a result, for a better understanding of the problems associated with this methodology, we recommend a more focused study (which makes use of a synthetic experiments) specifically addressed to cope with these problems.

The effect of the catchment area along with other basin characteristics was explored by comparing the results of the assimilation for the different catchments (Table 2, Figure 4). No particular pattern was found in the results of assimilation by changing the size of the catchment. Indeed, improvements were obtained for the biggest and the smallest basin. This raises two main issues related to the spatial mismatch between the catchment size and the spatial resolution of the satellite SM observations. That is, (i) coarse satellite SM observations (e.g., 25 km spatial resolution) can add a significant benefit also for small catchments (e.g., NI-MI), and (ii) the spatial mean of more satellite SM pixels (*i.e.*, taken among the pixels fallen inside the catchment area) can be successfully assimilated in a lumped model (e.g., TV-PF). Thus, the area of the catchment seems to play an insignificant role in the performances, at least for the explored range.

Assuming a correct evaluation of the model error, the success of the DA seems slightly affected by the choice of RT. The pattern of the improvements looks very similar for VM, LR and CDF, and the optimal error is most of the time similar for the different RTs. Some slight differences between the performances in terms of optimality in the choice of the observation error can be observed if we express the quantification of the observation error as a fraction of the variance of the rescaled observations (Figure 4.). In this case, VM and CDF provide almost coincident results and LR yields slightly lower values with similar efficiency. Moreover, it is clear in this case that considering the observation error as a fraction of the rescaled observation may be not always valid.

Once we selected the optimal observation error (the one related to the maximum NS_a , Table 2) for the three RTs, we obtained an improvement for all catchments except MA-AZ and a general lower performance for the CDF technique. The reason for the lower performance of CDF might be due to the small sample size, (*i.e.*, the length of the time series) where the CDF technique was calibrated. Similar results were found by [34] in the Murray Darling Basin in Australia who analyzed the impact of different RTs on the DA and showed sub-optimal performance of the CDF due to the relatively short two-year data record. As mentioned earlier, we have to stress the existence of other rescaling techniques (see [57,65] for further details); however, we found that the ones selected in this study are the most used in DA of SM in rainfall runoff modelling. Therefore, we prefer to focus our analysis by using them, and pursue the matter in a future study specifically addressed to analyzing the effect of the RT on the DA.

The analysis of the results in terms of flood events (Table 3) or by considering the indexes better suited for high and low flow conditions (Figure 5) offers an interpretation of the outcomes from another perspective and highlights some conflicting results. Indeed, by considering only the flood events, the assimilation deteriorates the performance in terms of error in volume for NI-MI while the analysis in terms of high ($ANSE$) and low flows ($NS_{(logQ)}$) highlights significant improvements for TV-PF for high flows but smaller and even negative scores for low flows. NE-MA, CHI-MO and NI-MI have consistent performances for low and high flows as well as MA-AZ for which the scores are always negative. Overall, a general lower performance is obtained with the CDF technique for high flow conditions, while for low flows CDF demonstrates an unexpected good behavior for TV-PF and provides similar results

to the other techniques for the other catchments. As before, this might be due to the larger number of SM values associated to low flow conditions with respect to those in high flow conditions that are used for calibrating the CDF parameters. The above considerations suggest that simpler rescaling techniques (e.g., VM) can be considered equally valid (or better), and, in most of the cases, preferable due to the simplicity of their implementation.

The choice of the optimal observation error for a user interested in performing DA in this area also appears to be a complex task. For instance, the optimal error is very different among adjacent catchments characterized by similar climatology, slopes, rainfall and altitude and similar calibration and validation scores (e.g., NE-MA and CHI-MO). In this case, the improvements are similar but the observation error is very different (20% for NE-MA, 5% for CHI-MO, or twice the rescaled observation for NE-MA and about one third of the rescaled observation for CHI-MO). One could argue that this may depend on the different quality of the satellite observations, but if we consider as reliable the SM-OBS4 soil moisture noise information (see Section 2.2), we find that it is very similar between the two (9% for CHI-MO, 9.5% for NE-MA). A justification of this result might be found in the accuracy of the model error estimation. Indeed, while the model performs very similarly for the two catchments, their model errors are quite different (Table 1); hence, the procedure used in Section 4.3 for the estimation of model error may be not completely reliable. Moreover, in validation (2010–2013), the standard deviations calibrated during 2000–2009 could not lead to the same optimal solutions. The result is that a proper choice of the observation error may strongly depend on a correct model error estimation. This is somehow consistent with what has been already found in the literature (e.g., [28]).

The analysis of the seasonality highlights some other interesting points. In general, it has been observed that the assimilation provides good results when SM temporal variability is higher (except for MA-AZ). Indeed, the periods of positive performances match with the periods of larger SM temporal variability for NI-MI, TV-PF, NE-MA and CHI-MO. The higher variability coincides in Central Italy with the wet season where, in fact, in November–March, the natural increase of SM (due to the seasonality) is accompanied by a high interchange between short-lived wet and dry periods, and the non-existence of a proper rainy season determines soil conditions which hardly remain constantly saturated. In these conditions, SM variability increases and the wetness conditions of the catchment prior to the occurrence of a flood event play a crucial role in determining the magnitude of the flood response. By contrast, in summer, rainfall is scarce and the relatively high temperature of the Mediterranean climate lowers the SM toward wilting point conditions for longer periods with a reduction of the SM variability. However, when SM temporal variability is low in the wet season (e.g., end of 2011), it does not guarantee that positive results are obtained by assimilation of SM. This suggests that rather than considering the season, one should consider the SM temporal variability for a good interpretation of the results. In order to find a mathematical proof of this tendency, we calculated the correlation coefficient between the SM temporal variability and the NRMSE, however, due to the randomness of the flood events especially in relatively small river basins (as explained in the Results Section), we found low absolute correlation values (around -0.3 for all the catchments except MA-AZ which yields $+0.45$). Although these values are low, the presence of an inverse correlation value (*i.e.*, higher SM variability *vs.* positive performance) somehow suggests the existence of the tendency of a reduction of the NRMSE with higher SM temporal variability. We expect that using a longer time period of analysis and focusing the study on much larger river basins can provide more convincing results. It should also be pointed out that satellite ASCAT SM observations are affected

by larger errors during the dry season in the Mediterranean areas due to the effect of the vegetation and other factors [66,67], and this might have an effect on the result interpretation. The contrasting behavior of MA-AZ catchment (*i.e.*, the lower performance of the DA when the SM temporal variability is higher, considering also the positive correlation between the SM variability and *NRMSE*), highlights the importance of possible local effects or a much larger observation error, which prevents an improvement in flood predictions.

Table 4 summarizes the main points addressed in the discussion and may facilitate the reader to go through the conclusion section.

Table 4. Analyzed factors and findings in the DA of soil moisture in Central Italy.

Factors	What Has Been Done	Findings
Model error	We performed a sensitivity analysis and calibrated the perturbation of the parameters and of the rainfall for calculating the optimal value of the ensemble statistical indexes	The results are good but we observed a problem of equifinality in which different perturbations lead to same values of the ensemble statistical indexes.
Catchment area, topography and soil type	We compared the performance of the assimilation for different catchments characterized by different areas (140–2000 km ²) and soil type	We did not observe any effect of the catchment area but catchment specific characteristics may prevent positive results of the DA
Rescaling technique	We compared the performance of the assimilation for three different RTs (LR, VM, CDF)	We obtained a uniform pattern of performance of the DA for LR, VM and CDF. We concluded that simple techniques may be equally valid as easier to implement
Observation error	We chose different observation errors and tested their optimality in the DA experiments	We obtained that adjacent and similar catchments have very different optimal observation errors, therefore its appropriate choice may strongly depend on a correct model error estimation
Flood magnitude	We explored the performance of the DA for low and high flows	We found larger improvements for high flow conditions.
Seasonality	We performed a seasonal analysis and compared it against the SM temporal variability	We obtained that DA performance during the year that has a relation with the SM temporal variability. Thus, we advise that this parameter must be taken into consideration when analyzing possible effects of seasonality in the DA of SM.

5. Conclusions

In this paper, making use of an experimental study, we discuss and highlight how the implementation of a DA experiment involves a complex setup phase characterized by a number of choices that can have non-negligible effects on the success of the DA. Assuming the validity of the assumptions for the application of the EnKF—which are not the goal of this study—we analyzed the effects of some important factors which can introduce complexities, uncertainties and non-optimality conditions for the application of DA and which can be propagated across the whole DA process. In particular, we found a number of contradictions and optimality issues that make the DA of soil moisture into rainfall-runoff modelling a complex task. In general, we found a general low effect for size but a strong effect for

catchment specific characteristics. Apparent insensitivity of the RT to the DA performance might be the result of a lack of analysis specifically focused on high, low flow or singular flood events and simple rescaling techniques may be equally valid to complex ones, with the advantage of an easier implementation. The selection of the observation error also provides some evidence that its optimal choice has a strong connection with an accurate representation of the model error and that the latter might play a central role in the success of the DA experiment. The analysis of the seasonality has provided evidence that SM temporal variability rather than the season is an important parameter to consider for a correct interpretation of the results of the DA of SM in terms of seasonal analysis.

Overall, while we provided new evidence that assimilation of SM can significantly improve discharge predictions, at the same time we showed that a number of factors can lead to non-optimal solutions with the result that DA of SM in rainfall-runoff modelling appears to be still a complex task and seems very much affected by local conditions and methodology issues. To this end, we believe that the setup phase must be tailored considering specific catchment characteristics and it needs a preliminary filter calibration operation. Hence, we encourage that further research in this field not only provides general guidelines or rules for a proper DA everywhere, but also supports catchment-data and application specific solutions, which involve new filter calibration methodologies as part of iterative DA setup improvement. This could be effectively reached through the development of open source tools for an easy and simple setup of the DA experiments, which would allow testing sensors, techniques and DA results worldwide. Only then we will have a clearer picture of whether a sensor or a technique is better than another.

Acknowledgments

We would like to acknowledge Umbria region, the Italian Civil Protection Department (<http://www.cfumbria.it/>, <http://www.protezionecivile.gov.it/>) and the EUMETSAT Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF, <http://hsaf.meteoam.it/>) for providing the data. Cristiano Corradini and Stefania Camici are also acknowledged for preprocessing the data.

Author Contributions

All authors contributed extensively to the work presented in this paper. Specific contributions included (i) the development of the data assimilation code (Christian Massari) and of the hydrological model (Luca Brocca and Christian Massari), (ii) the implementation of the filtering, rescaling and processing of the observations along with the run of the simulations and the manuscript preparation (Christian Massari, Angelica Tarpanelli and Luca Brocca) and (iii) the analysis of the results (Christian Massari, Luca Brocca, Angelica Tarpanelli and Tommaso Moramarco). All co-authors contributed to the editing of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Brocca, L.; Melone, F.; Moramarco, T. On the estimation of antecedent wetness conditions in rainfall–runoff modelling. *Hydrol. Process.* **2008**, *642*, 629–642.
2. Tramblay, Y.; Bouaicha, R.; Brocca, L.; Dorigo, W.; Bouvier, C.; Camici, S.; Servat, E. Estimation of antecedent wetness conditions for flood modelling in northern Morocco. *Hydrol. Earth Syst. Sci.* **2012**, *16*, 4375–4386.
3. Coustau, M.; Bouvier, C.; Borrell-Estupina, V.; Jourde, H. Flood modelling with a distributed event-based parsimonious rainfall-runoff model: Case of the karstic Lez river catchment. *Nat. Hazard. Earth Syst. Sci.* **2012**, *12*, 1119–1133.
4. Van Steenbergen, N.; Willems, P. Increasing river flood preparedness by real-time warning based on wetness state conditions. *J. Hydrol.* **2013**, *489*, 227–237.
5. Liu, Y.; Gupta, H.V. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resour. Res.* **2007**, *43*, 1–18.
6. Tramblay, Y.; Bouvier, C.; Ayral, P.; Marchandise, A. Impact of rainfall spatial distribution on rainfall-runoff modelling efficiency and initial soil moisture conditions estimation. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 157–170.
7. Bárdossy, A.; Das, T. Influence of rainfall observation network on model calibration and application. *Hydrol. Earth Syst. Sci. Discuss.* **2006**, *3*, 3691–3726.
8. Zehe, E.; Becker, R.; Bárdossy, A.; Plate, E. Uncertainty of simulated catchment runoff response in the presence of threshold processes: Role of initial soil moisture and precipitation. *J. Hydrol.* **2005**, *315*, 183–202.
9. Grayson, R.; Blöschl, G. *Spatial Patterns in Catchment Hydrology: Observations and Modelling*, 1st ed.; Cambridge University Press: New York, NY, USA, 2001.
10. Gupta, H.V.; Sorooshian, S.; Yapo, P.O. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res.* **1998**, *34*, 751–763.
11. Bartalis, Z.; Wagner, W.; Naemi, V.; Hasenauer, S.; Scipal, K.; Bonekamp, H.; Figa, J.; Anderson, C. Initial soil moisture retrievals from the METOP-A Advanced Scatterometer (ASCAT). *Geophys. Res. Lett.* **2007**, *34*, L20401.
12. Owe, M.; de Jeu, R.; Holmes, T. Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res.* **2008**, *113*, F01002.
13. Kerr, Y.H.; Waldteufel, P.; Wigneron, J.-P.; Delwart, S.; Cabot, F.; Boutin, J.; Escorihuela, M.-J.; Font, J.; Reul, N.; Gruhier, C.; *et al.* The SMOS mission: New tool for monitoring key elements of the global water cycle. *IEEE Proc.* **2010**, *98*, 666–687.
14. Entekhabi, D.; Njoku, E.G.; O’Neill, P.E.; Kellogg, K.H.; Crow, W.T.; Edelstein, W.N.; Entin, J.K.; Goodman, S.D.; Jackson, T.J.; Johnson, J.; *et al.* The soil moisture active passive (SMAP) mission. *IEEE Proc.* **2010**, *98*, 704–716.
15. Pauwels, V.R.N.; Hoeben, R.; Verhoest, N.E.C.; De Troch, F.P.; Troch, P.A. Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values. *Hydrol. Process.* **2002**, *16*, 995–1013.
16. Reichle, R.; McLaughlin, D.B.; Entekhabi, D. Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Weather Rev.* **2002**, *130*, 103–114.

17. Brocca, L.; Moramarco, T. Assimilation of surface-and root-zone ASCAT soil moisture products into rainfall–runoff modeling. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 2542–2555.
18. Draper, C.; Mahfouf, J.F.; Calvet, J.C.; Martin, E.; Wagner, W. Assimilation of ASCAT near-surface soil moisture into the SIM hydrological model over France. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 3829–3841.
19. Matgen, P.; Fenicia, F.; Heitz, S.; Plaza, D.; de Keyser, R.; Pauwels, V.R.N.; Wagner, W.; Savenije, H. Can ASCAT-derived soil wetness indices reduce predictive uncertainty in well-gauged areas? A comparison with *in situ* observed soil moisture in an assimilation application. *Adv. Water Resour.* **2012**, *44*, 49–65.
20. Chen, F.; Crow, W.T.; Starks, P.J.; Moriasi, D.N. Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. *Adv. Water Resour.* **2011**, *34*, 526–536.
21. Chen, F.; Crow, W.T.; Ryu, D. Dual forcing and state correction via soil moisture assimilation for improved rainfall-runoff modeling. *J. Hydrometeorol.* **2014**, *15*, 1832–1848.
22. Alvarez-Garreton, C.; Ryu, D.; Western, A.W.; Su, C.-H.; Crow, W.T.; Robertson, D.E.; Leahy, C. Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: Comparison between lumped and semi-distributed schemes. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 1659–1676.
23. Zhou, Y.; McLaughlin, D.; Entekhabi, D. Assessing the performance of the ensemble kalman filter for land surface data assimilation. *Mon. Weather Rev.* **2006**, *134*, 2128–2142.
24. Evensen, G. The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Syst.* **2009**, 83–104.
25. Wagner, W.; Lemoine, G.; Rott, H. A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sens. Environ.* **1999**, *70*, 191–207.
26. Crow, W.T.; Van Loon, E. Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture. *J. Hydrometeorol.* **2006**, *7*, 421–432.
27. Alvarez-Garreton, C.; Ryu, D.; Western, A.W.; Crow, W.; Robertson, D.; States, U. Impact of observation error structure on satellite soil moisture assimilation into a rainfall-runoff model. In Proceedings of the 20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December 2013; pp. 1–6.
28. Crow, W.T.; Van Den Berg, M.J. An improved approach for estimating observation and model error parameters in soil moisture data assimilation. *Water Resour. Res.* **2010**, *46*, 1–12.
29. De Lannoy, G.J.M.; Houser, P.R.; Pauwels, V.R.N.; Verhoest, N.E.C. Assessment of model uncertainty for soil moisture through ensemble verification. *J. Geophys. Res.* **2006**, *111*, D10101.
30. Dorigo, W.A.; Scipal, K.; Parinussa, R.M.; Liu, Y.Y.; Wagner, W.; de Jeu, R.A.M.; Naeimi, V. Error characterisation of global active and passive microwave soil moisture datasets. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 2605–2616.
31. Brocca, L.; Hasenauer, S.; Lacava, T.; Melone, F.; Moramarco, T.; Wagner, W.; Dorigo, W.; Matgen, P.; Martínez-Fernández, J.; Llorens, P.; *et al.* Soil moisture estimation through ASCAT and AMSR-E sensors: An intercomparison and validation study across Europe. *Remote Sens. Environ.* **2011**, *115*, 3390–3408.
32. Su, C.-H.; Ryu, D.; Western, A.W.; Wagner, W. De-noising of passive and active microwave satellite soil moisture time series. *Geophys. Res. Lett.* **2013**, *40*, 3624–3630.

33. Reichle, R.H.; Koster, R.D. Assessing the impact of horizontal error correlations in background fields on soil moisture estimation. *J. Hydrometeorol.* **2003**, *4*, 1229–1242.
34. Lievens, H.; Tomer, S.K.; Al Bitar, A.; De Lannoy, G.J.M.; Drusch, M.; Dumedah, G.; Hendricks Franssen, H.-J.; Kerr, Y.H.; Martens, B.; Pan, M.; *et al.* SMOS soil moisture assimilation for improved hydrologic simulation in the Murray Darling Basin, Australia. *Remote Sens. Environ.* **2015**, *168*, 146–162.
35. Koster, R.D.; Guo, Z.; Yang, R.; Dirmeyer, P.A.; Mitchell, K.; Puma, M.J. On the nature of soil moisture in land surface models. *J. Clim.* **2009**, *22*, 4322–4335.
36. Reichle, R.H.; Koster, R.D. Bias reduction in short records of satellite soil moisture. *Geophys. Res. Lett.* **2004**, *31*, L19501.
37. Alvarez-Garreton, C.; Ryu, D.; Western, A.W.; Crow, W.T.; Robertson, D.E. The impacts of assimilating satellite soil moisture into a rainfall–runoff model in a semi-arid catchment. *J. Hydrol.* **2014**, *519*, 2763–2774.
38. Yilmaz, M.T.; Crow, W.T. The optimality of potential rescaling approaches in land data assimilation. *J. Hydrometeorol.* **2013**, *14*, 650–660.
39. Su, C.; Ryu, D.; Crow, W.T.; Western, A.W. Beyond triple collocation: Applications to soil moisture monitoring. *J. Geophys. Res. Atmos.* **2014**, *119*, 6419–6439.
40. Brocca, L.; Moramarco, T.; Dorigo, W.; Wagner, W. Assimilation of satellite soil moisture data into rainfall-runoff modelling for several catchments worldwide. In Proceedings of 2013 IEEE Geoscience and Remote Sensing Symposium, Melbourne, Australia, 21–26 July 2013; pp. 2281–2284.
41. Djuric, P.; Kotecha, J.; Zhang, J.; Huang, Y.; Ghirmai, T.; Bugallo, M.F.; Miguez, J. Particle filtering. *Signal Process. Mag.* **2003**, *20*, 19–38.
42. Matgen, P.; Heitz, S.; Hasenauer, S.; Hissler, C.; Brocca, L.; Hoffmann, L.; Wagner, W.; Savenije, H.H.G. On the potential of MetOp ASCAT-derived soil wetness indices as a new aperture for hydrological monitoring and prediction: A field evaluation over Luxembourg. *Hydrol. Process.* **2012**, *26*, 2346–2359.
43. Corato, G.; Matgen, P.; Fenicia, F.; Schläffer, S.; Chini, M. Assimilating satellite-derived soil moisture products into a distributed hydrological model. In Proceedings of 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec, QC, Canada, 14 July 2014; pp. 3315–3318.
44. Moradkhani, H.; Hsu, K.-L.; Gupta, H.; Sorooshian, S. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resour. Res.* **2005**, *41*, W05012.
45. H-SAF. Available online: <http://hsaf.meteoam.it> (accessed on 7 September 2015).
46. Vachaud, G.; Passerat De Silans, A.; Balabanis, P.; Vauclin, M. Temporal stability of spatially measured soil water probability density function. *Soil Sci. Soc. American J.* **1985**, *49*, 822–828.
47. Brocca, L.; Melone, F.; Moramarco, T.; Wagner, W.; Hasenauer, S. ASCAT soil wetness index validation through *in situ* and modeled soil moisture data in central Italy. *Remote Sens. Environ.* **2010**, *114*, 2745–2755.
48. Melone, F.; Neri, N.; Morbidelli, R.; Saltalippi, C. A conceptual model for flood prediction in basins of moderate size. In *Applied Simulation and Modeling*; Hamza, M.H., Ed.; IASTED Acta Press: Anaheim, CA, USA, 2001; pp. 461–466.

49. USDA, SCS. *National Engineering Handbook: Hydrology, Section 4*; US Department of Agriculture, Soil Conservation Service; Washington, DC, United States, 1985; Chapters 4–10.
50. Pruitt, W.O.; Doorenbos, J. Background and development of methods to predict reference crop evapotranspiration (ET_o). *Guidel. Predict. Crop Water Requir.* **1977**, *24*, 108–119.
51. Brocca, L.; Melone, F.; Moramarco, T.; Morbidelli, R. Antecedent wetness conditions based on ERS scatterometer data. *J. Hydrol.* **2009**, *364*, 73–87.
52. Melone, F.; Corradini, C.; Singh, V.P. Lag prediction in ungauged basins: An investigation through actual data of the upper Tiber River valley. *Hydrol. Process.* **2002**, *16*, 1085–1094.
53. Burgers, G.; Jan van Leeuwen, P.; Evensen, G. Analysis scheme in the ensemble kalman filter. *Mon. Weather Rev.* **1998**, *126*, 1719–1724.
54. Albergel, C.; Rüdiger, C.; Pellarin, T.; Calvet, J.-C.; Fritz, N.; Froissard, F.; Suquia, D.; Petitpa, A.; Piguet, B.; Martin, E. From near-surface to root-zone soil moisture using an exponential filter: An assessment of the method based on *in-situ* observations and model simulations. *Hydrol. Earth Syst. Sci.* **2008**, *5*, 1603–1640.
55. Dee, D.P.; Da Silva, A.M. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* **1998**, *124*, 269–295.
56. Reichle, R.H.; Crow, W.T.; Keppenne, C.L. An adaptive ensemble Kalman filter for soil moisture data assimilation. *Water Resour. Res.* **2008**, *44*, W03423.
57. Drusch, M. Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. *Geophys. Res. Lett.* **2005**, *32*, L15403.
58. Su, C.-H.; Ryu, D. Multi-scale analysis of bias correction of soil moisture. *Hydrol. Earth Syst. Sci.* **2015**, *19*, 17–31.
59. Tian, Y.; Huffman, G.J.; Adler, R.F.; Tang, L.; Sapiano, M.; Maggioni, V.; Wu, H. Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophys. Res. Lett.* **2013**, *40*, 2060–2065.
60. Turner, M.; Walker, J.; Oke, P. Ensemble member generation for sequential data assimilation. *Remote Sens. Environ.* **2008**, *112*, 1421–1433.
61. Nash, J.; Sutcliffe, J. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290.
62. Hoffmann, L.; El Idrissi, A.; Pfister, L.; Hingray, B.; Guex, F.; Musy, A.; Humbert, J.; Drogue, G.; Leviandier, T. Development of regionalized hydrological models in an area with short hydrological observation series. *River Res. Appl.* **2004**, *20*, 243–254.
63. Moramarco, T.; Singh, V.P.; Alta, V.M.; Rouge, B. Simple method for relating local stage and remote discharge. *J. Hydrol. Eng.* **2001**, *6*, 78–81.
64. Brocca, L.; Melone, F.; Moramarco, T.; Wagner, W.; Naeimi, V.; Bartalis, Z.; Hasenauer, S. Improving runoff prediction through the assimilation of the ASCAT soil moisture product. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 1881–1893.
65. Albergel, C.; de Rosnay, P.; Gruhier, C.; Muñoz-Sabater, J.; Hasenauer, S.; Isaksen, L.; Kerr, Y.; Wagner, W. Evaluation of remotely sensed and modelled soil moisture products using global ground-based *in situ* observations. *Remote Sens. Environ.* **2012**, *118*, 215–226.

66. Massari, C.; Brocca, L.; Barbetta, S.; Papathanasiou, C.; Mimikou, M.; Moramarco, T. Using globally available soil moisture indicators for flood modelling in Mediterranean catchments. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 839–853.
67. Zwieback, S.; Dorigo, W.; Wagner, W. Temporal error variability of coarse scale soil moisture products-case study in central Spain. In Proceedings of the 2012 IEEE International: Geoscience and Remote Sensing Symposium , Munich, Bavaria, Germany, 22–27 July 2012; pp. 722–725.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).