Assimilation of Satellite Soil Moisture Products for River Flow Prediction: An Extensive Experiment in Over 700 Catchments Throughout Europe

D. De Santis1, D. Biondi1, W. T. Crow3, S. Camici2, S. Modanesi2, L. Brocca2, and C. Massari2

1Department of Computer, Modeling, Electronics and Systems Engineering, University of Calabria, Arcavacata di Rende, Italy, 2Research Institute for Geo-Hydrological Protection, National Research Council, Perugia, Italy, 3Hydrology and Remote Sensing Laboratory, USDA ARS, Beltsville, MD, USA

Abstract In this study, we perform a data assimilation (DA) experiment on a very large number (~700) of small- and medium-scale (150–10,000 km²) European catchments to assess the impact of satellite soil moisture (SM) DA on streamflow simulations for different climatic and hydrologic conditions. In the experiment, Climate Change Initiative SM active, passive and combined products are assimilated over a time period 2003–2016 via an Ensemble Kalman Filter (EnKF). The results show that, on average, the assimilation of the three products provides relatively small improvements as compared to analogous open loop (OL) results (i.e., with an increase on median Kling-Gupta Efficiency equal to 0.0048, 0.0033, and 0.0022 [−] for the active, the passive, and the combined products, respectively). OL performance itself is found to be a significant driver of the assimilation results: greater improvements are observed in catchments with poor OL streamflow predictions and inaccurate precipitation estimates. The remotely sensed product accuracy also emerges as relevant for assimilation efficiency, while factors affecting SM retrievals such as vegetation density, topographic complexity and basin area are found to have only a limited impact on the spatial pattern of performance. Small and detrimental effects of SM assimilation are observed over fully humid catchments and at high latitudes where pre-storm soil moisture has reduced control on runoff generation as well as in basins where the hydrological model contains structural limitations.

1. Introduction

Soil moisture (SM) pre-storm conditions are a key factor in runoff production and can explain much of the observed hydrological response of a basin (e.g., Berthet et al., 2009; Penna et al., 2011). For this reason, the integration of SM observations into hydrological models is considered a valuable practice to improve streamflow predictions (e.g., Coustau et al., 2012; Tramblay et al., 2010).

In recent years, SM estimates from remote sensing measurements have experienced growth in their availability and accuracy (Karthikeyan et al., 2017; McCabe et al., 2017). Multiple operational and research satellite SM data sets from active to passive microwave sensors are currently available (e.g., Mecklenburg et al., 2016; Wagner et al., 2013) with coarse-scale surface retrievals and approximately daily coverage.

Data assimilation (DA) is a theoretically powerful way to integrate remotely sensed SM observations into hydrological models (Y. Liu et al., 2012); however, its actual benefit is still controversial regarding streamflow prediction, with several studies showing neither systematic nor significant improvements on streamflow prediction, due to a number of technical, mathematical and hydrological factors (e.g., Alvarez-Garreton et al., 2015; Chen et al., 2011; Draper et al., 2011; Mao et al., 2019; Massari et al., 2015; Matgen et al., 2012; Ridler et al., 2014). In other words, although it appears to be a logical step to update modeled SM via DA—as a more accurate knowledge of the state of the system should lead to a better streamflow forecast—there are many underlying factors responsible for the relative success or the failure of the assimilation task, such as: The actual connection between SM and runoff and its model representation, the accuracy of assimilated remote sensing retrievals, and the procedural choices within the DA system configuration—including the fulfillment of DA method theoretical assumptions.
The sensitivity of the catchment response to SM can be varied and influenced by the dominant runoff mechanism operating in a particular catchment. For instance, SM DA usually has a negligible effect in catchments where the hydrological response is snowmelt-dominated (e.g., Koster et al., 2018; Reichle et al., 2019).

Moreover, models may not adequately describe the link between SM and runoff generation - thus limiting SM DA efficiency. This applies not only to conceptual hydrological models, but also to Land Surface Models (LSM) that do not always reflect the best hydrologic process understanding (Crow et al., 2018). Depending on model structure, simulated SM can be considered as an “index” of the wetness state that does not necessarily provide a measure corresponding to absolute soil water content (Koster & Milly, 1997). In this respect, hydrological models can often produce good runoff results in cases where their SM estimates contain systematic error, and a SM reanalysis may not necessarily benefit streamflow predictions. However, despite the model-specific nature of simulated SM, models tend to produce very similar temporal SM variability (Koster et al., 2009), confirming a high potential for SM DA provided that systematic differences between modeled and observed SM are treated appropriately. Furthermore, if pre-storm SM conditions in the subsurface layers are the most critical for runoff generation, then DA needs to rely on adequate vertical coupling between surface and root zone model layers for the satellite surface SM assimilation to be effective (e.g., Brocca et al., 2012; Chen et al., 2011).

Another key factor in SM DA performances is the accuracy of satellite SM for the purpose of hydrological simulations. In addition to errors in sensor measures and retrieval model, both pre-retrieval and post-retrieval quantities are impacted by the presence of urban areas, dense vegetation, open water, complex topography, snow coverage or frozen soils. These factors result in different types of retrieval errors (gross, systematic, and random) which must be properly corrected and/or characterized and may constitute a limit on DA results. In addition, the procedures used to deal with observation errors (e.g., approaches for bias correction and random error characterization) have been shown to have a significant impact on assimilation results (e.g., Massari et al., 2015).

Finally, different DA approaches rely on their own assumptions (e.g., observation errors uncorrelated in time, error normality, absence of bias, etc.), that, if violated, can preclude realizing any material advantage from SM retrieval skill.

Most studies addressing the impact of DA on streamflow prediction have been carried out on a small number of basins within a limited geographical area. Very few studies have reported results from more extensive analyses considering a large number of basins representing different (climatic, physiographic, etc.) conditions and, therefore, truly capable of providing truly general (i.e., not case-specific) results. In Kumar et al. (2014), the effects of assimilating satellite SM (retrieved from a range of radiometers including AMSR-E) were evaluated over the continental United States. In their experiment, a LSM coupled with a distributed routing module was employed. Minor streamflow improvements were obtained across most basins—perhaps due to the high skill of model SM estimates prior to SM assimilation. Fairbairn et al. (2017) evaluated the impact of assimilating ASCAT-derived SM into a LSM coupled with a distributed hydrogeological model. The investigation was conducted on over 500 stream-gauge stations in France, with the aim of correcting random errors in model initial conditions; however, the experiment resulted in spurious increases in runoff, which degraded the model performances, partly attributed to limitations in DA implementation. In Lievens et al. (2016), SMOS SM was instead assimilated into a LSM coupled with a lag-routing module. The experiment was conducted on a large basin in Australia with over 100 streamflow stations, generally obtaining minor improvements. Small assimilation impacts were attributed to: Limited streamflow sensitivity to antecedent SM conditions; vegetation, climatic, and orographic effects on satellite retrievals; and ambiguous procedural choices for bias correction and random errors characterization.

The extensive experiments in the above-mentioned studies have analyzed the use of data from only one type of microwave sensor (i.e., from only scatterometers or radiometers). For example, Fairbairn et al. (2017) assimilated ASCAT satellite soil moisture data; Kumar et al. (2014) used only a passive microwave sensor and Lievens et al. (2016) assimilated passive SMOS soil moisture retrievals. As a result, these studies have neglected active versus passive variations in SM retrieval quality (Chen et al., 2018). Likewise, past studies have generally assimilated satellite SM within complex, data-demanding physically based land-surface/hydrological models. While there are relevant examples of applying complex land models in operational
contexts (e.g., Spyrou et al., 2020; Viterbo et al., 2020), they are generally much more difficult to apply in near-real-time than simpler conceptual models (Wijayarathne & Coulibaly, 2020). For instance, the Sacramento model (SAC-SMA, Burnash et al., 1973) remains a key model used by the United States National Weather Service River Forecast System to issue river forecasts across the country. Likewise, the HBV model (Bergström, 1995) is widely used by the Swedish Meteorological and Hydrological Institute for operational hydrological forecasting and spillway design.

Other SM DA experiments involved a large number of basins, but as opposed to the above-mentioned studies, were performed on a decadal or monthly scale or without a routing module (e.g., Albergel et al., 2017; Koster et al., 2018; Reichle et al., 2019). These studies also tend to report a relatively small impact on streamflow simulations.

The primary task of this study is to contribute to the evaluation of satellite SM assimilation in a conceptual rainfall-runoff model and, by considering a large sample of catchments (>700) located throughout Europe, to allow for an exhaustive investigation of factors impacting SM assimilation performance. In this way, we also expand upon the geographic coverage of past studies that were mainly confined within national boundaries (e.g., Draper et al., 2011; Fairbairn et al., 2017; Ridler et al., 2014); moreover, the set of basins, we explore is representative of a wide range of climates, as well as vegetation and orography conditions for which the generalization of research findings and conclusions in different contexts is not obvious and cannot be automatically extended.

Here, indications about the added value of SM retrievals from both active and passive sensors under different catchment conditions are thoroughly investigated, focusing on the way in which the accuracy evaluated for the satellite SM data reflects on improvement or deterioration of simulated discharges. This systematic investigation is conducted using SM retrievals derived from active and passive microwave sensors with sufficiently extensive temporal data coverage. For this purpose, SM products from European Space Agency Climate Change Initiative (ESA CCI) are employed. These products are derived by merging observations from different sensors to generate an active, a passive, and a combined data product - each covering a multi-decadal period.

Besides basic catchment properties, data assimilation results also depend strongly on the baseline assimilation model used. Here, remotely sensed observations over the period 2003–2016 are assimilated via an Ensemble Kalman Filter (Evensen, 1994) into a conceptual rainfall-runoff model MISDc-2L (Massari et al., 2018). The central research question is whether remotely sensed observations (soil moisture in this case) can compensate for the process-level simplicity of a conceptual model—relative to the more comprehensive representation provided by LSMs. To further explore the effect of satellite SM DA on discharge predictions, the enhancement in efficiency indexes with reference to different streamflow conditions (high, medium, and low flows) are also evaluated.

Furthermore, we carry out a systematic investigation of different conditions possibly influencing the assimilation impact on runoff modeling. This analysis focuses on multiple factors including: (a) the role of baseline open loop (OL) model performance, (b) sensitivity to the use of three different ESA CCI SM data sets, and (c) the accuracy of remotely sensed data. Moreover, the influence of catchment characteristics (e.g., vegetation, topography, and climate conditions) as well as of runoff generation dynamics (e.g., snow-dominated catchments) are also considered.

The study is structured as follows. The materials employed are described in Section 2, and our methods are illustrated in Section 3. The results are presented and discussed in Section 4, and concluding remarks are provided in the last section.

2. Materials

2.1. Study Catchments and Hydrological Data

The DA experiment is conducted on a set of 775 European catchments, with areas ranging from 150 to 10,000 km² (Table 1 and Figure S1). Daily discharge measurements for these catchments are obtained from five different hydrological archives: The Global Runoff Data Base, the European Water Archive, and
national databases from Italy (ISPRA HIS), Spain (CEDEX CEH), and Portugal (SNIRH). Catchment boundaries, if not made available by data product developers, are obtained following the procedure described in Camici et al. (2020). The study period starts in 2003 (in order to match the daily availability of satellite SM retrievals) and ends on a basin-specific date between 2011 and 2016 (depending on the particular streamflow data record available within each catchment). Potential anthropogenic impacts on discharge time series are checked through an accurate visual inspection identifying clear patterns of flow regulations like constant flow regime to variable precipitation inputs and sudden changes in the flood hydrographs. Moreover, since the model simulates only natural flows the performance over highly regulated basins are very low and the chance of considering these basins is significantly reduced by verifying that the model's streamflow predictions reach a minimum performance benchmark. The list of basins employed in this experiment is reported in Table S1.

### 2.2. Climatic Data

Basin average precipitation and mean temperature at the daily scale are estimated from the gridded 0.25° E-OBS v19.0e data set (Haylock et al., 2008) provided by the European Climate Assessment & Data set (ECA&D) project for the period 2002–2016. E-OBS gridded data are obtained from the ground station observations collected as part of the ECA&D initiative, with station density varying significantly over the European domain. The current spatial interpolation method is based on the generation of a 100-member ensemble of realizations of each daily field. Unfortunately, the uncertainty directly predicted by this ensemble spread, which is closely related to the station density, turned out to be an underestimation in data-sparse regions (Cornes et al., 2018). For this reason, the provided ensemble spread values are only qualitatively considered in this study to capture spatial patterns in expected precipitation errors. Figure S2 plots the temporal average of this ensemble spread over all study catchments for the study period.

### 2.3. Remotely Sensed Soil Moisture Data

Within the ESA CCI project, multiple single-sensor active and passive microwave SM data sets are blended into three separate harmonized global surface soil moisture products: A merged active, a merged passive, and a combined one (Dorigo et al., 2017; Gruber et al., 2019; Y. Y. Liu et al., 2012). All ESA CCI SM v04.4 products employed in this study are provided with a daily temporal sampling (reference time at 0:00 UTC) on a regular 0.25° grid.

Ancillary data are provided by the ESA CCI project containing information on the spatial distribution of factors affecting the accuracy of SM retrievals including vegetation density, expressed as temporally averaged vegetation optical depth (VOD), and topographic complexity. The spatial distributions of basin-averaged VOD and topographic complexity within the study domain are plotted in Figure 1.

SM observations for the period 2003–2016, consisting of retrievals not flagged for data inconsistency (e.g., temporary snow coverage or dense vegetation), are considered for the DA experiment. Until December 2006, active microwave retrievals are obtained from the ERS wind scatterometers (revisit time of several days) and, subsequently, from the ASCAT sensor on board MetOp satellites (with near-daily time coverage). Passive microwave data are based on observations from the AMSR-E sensor on the Aqua satellite, the WindSat sensor on the Coriolis satellite, the SMOS radiometer, and the AMSR2 sensor on the GCOM-W1 satellite. Combined, these sources guarantee nearly daily data cover over our European spatial domain.

A preliminary evaluation of the accuracy of ESA CCI combined SM is performed (Figure S3) in terms of Pearson correlation with modeled SM estimates acquired versus surface SM estimates acquired from the Global Land Data Assimilation System (GLDAS) (Rodell et al., 2004). These results show good agreement in the Mediterranean (particularly the Iberian Peninsula). However, mountain ranges (e.g., the Alps, Pyrenees) and Nordic areas (e.g., Scandinavia) are characterized by relatively poor temporal correlation between GLDAS and ESA CCI.

| Catchment area (km²) | Mean annual rainfall (mm) | Vegetation optical depth (–) | Topographic complexity index (%) |
|---------------------|---------------------------|-----------------------------|---------------------------------|
| Minimum value       | 150                       | 316                         | 0.18                            |
| Maximum value       | 9,948                     | 3,388                       | 0.64                            |

| Table 1 | Main Characteristics of the Study Catchments | Mean annual rainfall (mm) | Vegetation optical depth (–) | Topographic complexity index (%) |
|---------|---------------------------------------------|---------------------------|-----------------------------|---------------------------------|
| Minimum value | 150                        | 316                       | 0.18                        | 0                               |
| Maximum value | 9,948                      | 3,388                     | 0.64                        | 50                              |
2.4. Hydrological Model

The rainfall-runoff model employed in this study is MISDc-2L (Brocca et al., 2012; Massari et al., 2018), which allows for the reproduction of different runoff generation mechanisms and simulates several hydrological processes including evapotranspiration, infiltration, percolation, snow melting, surface, and subsurface runoff.

The model consists of two components; the first one is a soil water balance model that simulates soil moisture temporal patterns and sets the initial conditions for the second component which is a rainfall-runoff model for flood hydrograph simulation.

The soil water balance component subdivides the soil into two zones: At each time step, the model simulates the temporal evolution of the soil water content in the upper (layer 1) and lower (layer 2) soil layers—expressed in terms of water depth ranging from 0 to \( W_{\text{max},1} \) and \( W_{\text{max},2} \), respectively. To model the dynamics of fluxes and water storage processes, the two layers are assumed as interconnected: Specifically, a percolation rate between the two layers, calculated via a non-linear relation proposed by Famiglietti and Wood (1994), links the one-dimensional representation of the soil water balance in the upper soil zone to that of the lower soil zone. This two-layer approach allows for the simulation of SM control on the partitioning of rainfall into infiltration and surface runoff (infiltration excess) in the upper layer, as well as a saturation excess mechanism when water content exceeds the maximum capacity. Surface runoff generated from infiltration and saturation excess is convoluted through a Geomorphological Instantaneous Unit Hydrograph, while subsurface runoff is transferred to the outlet by a linear reservoir approach. For both routing schemes, a common lag time is applied using the relationship with catchment area proposed by Melone et al. (2002).

The model is applied in lumped mode, with a daily time step. \( W_{\text{max},1} \) fixed at 45 mm (which roughly corresponds to 10 cm soil depth at a porosity of 0.45 m\(^3\)m\(^{-3}\)) for each catchment; this assumption is a tradeoff between the soil depths investigated by satellite sensors and the necessity to avoid numerical instabilities at a daily scale due to an excessively reduced soil thickness (Manfreda et al., 2014). Reducing the gap between modeled and observed surface layers allows for the direct assimilation of remotely sensed SM without a preprocessing step such as the exponential filter (Wagner et al., 1999) that will introduce time-correlated errors into observations. The soil water capacity of the lower layer, \( W_{\text{max},2} \), is calibrated for each catchment within the range 250–3000 mm, together with eight other parameters (Table S2). All model SM variables are expressed in relative saturation terms by normalizing water content with respect to layer capacity, and referred to as \( W_{1\%} \) and \( W_{2\%} \), for the upper and lower soil layers, respectively.
3. Methods

DA methods aim to merge forecasted model state (i.e., the background) and observations by considering their respective uncertainties to minimize the error variance in the updated model states (i.e., the analysis). For its computational efficiency and robustness, the EnKF, a sequential DA method well-suited for dealing with moderate nonlinearities typically found in hydrologic models, is employed for assimilating satellite data. The EnKF requires an appropriate ensemble of model states to represent the background error covariance at any given time.

The EnKF assumes the absence of bias in model and observational data sets and is designed to handle random, zero-mean, normally distributed errors. Following typical practice, systematic differences between modeled and remotely sensed SM are corrected prior to DA. A common approach is to rescale the observations to match some statistical properties of the model estimates (e.g., Reichle & Koster, 2004), with the implicit outcome that all bias is attributed to the observations. Here, a preprocessing phase is implemented with remotely sensed SM data being mapped into the space of the $W_{1\%}$ state variable through the CDF-matching approach. The “preprocessed” surface soil moisture observations (hereinafter referred to as SSM*)—properly characterized in terms of their random error variance $\sigma^2_{z,\text{obs}}$—are then assimilated to update $W_{1\%}$ and $W_{2\%}$ model states.

Calibration and DA evaluation are both performed over the entire period 2003–2016. The reason for considering a single time period is twofold: On the one hand there is a limited overlap between actual data availability of ESA CCI SM products and observed streamflow time series (that in many cases do not go beyond 2012); on the other hand, there is a need to perform the various calibration steps (including observations rescaling parameters, here evaluated on a monthly scale) over a sufficiently long time-window to obtain robust estimates.

Figure 2 provides a schematic overview of the steps applied to perform DA, which are described in detail in the following sections.

3.1. Ensemble Kalman Filter Implementation

The model state vector for the $i$-th ensemble member consisting of the soil water contents in the two soil layers is:

$$x^i = \begin{bmatrix} W_{1\%}^i \\ W_{2\%}^i \end{bmatrix}$$

(1)

When an observation $y$ (i.e., SSM*) is available at time $t_k$, the model state vector is updated as follows:

$$x_k^{i+} = x_k^{i-} + G_k \left( y_k + \eta_k^i - Hx_k^{i-} \right)$$

(2)

where:

1. $x_k^{i-}$ and $x_k^{i+}$ refer to the background and analysis states, respectively
2. $H = \begin{bmatrix} 1 & 0 \end{bmatrix}$ is the linear observation operator that maps between model states and SM observations
3. $\eta_k^i$ is the observation perturbation for the $i$-th ensemble member, with values drawn from a normal distribution with zero-mean and known variance (i.e., $\sigma^2_{z,\text{obs}}$)
4. $G_k$ is the $2 \times 1$ Kalman gain matrix.

The Kalman gain matrix is defined as follows:

$$G_k = P_k^- H^T \left( HP_k^- H^T + R \right)^{-1}$$

(3)
where $H^T$ is the transpose of the observation operator, and $P_k^r$ and $R$ are the covariance matrix of background and observation errors respectively. For the assimilation of only one type of observation, the $R$ matrix reduces to the scalar $\sigma_{e,obs}^2$, while the time-variant $2 \times 2$ matrix is defined as:

$$P_k^r = \frac{1}{N-1} D_k D_k^T$$

(4)

with:

$$D_k = \begin{bmatrix} x_k^1 - \bar{x}_k \\ \vdots \\ x_k^N - \bar{x}_k \end{bmatrix}$$

(5)

$$\bar{x}_k = \frac{1}{N} \sum_{i=1}^{N} x_k^i$$

(6)

and $N$ is the ensemble size.

3.2. Model Calibration and Error Representation

The model calibration is performed by optimizing the Kling-Gupta Efficiency index (KGE; Gupta et al., 2009) based on the Euclidian distance from the ideal point of three components, that is, correlation and ratios in mean and standard deviation:

$$\text{KGE} = 1 - \sqrt{(r - 1)^2 + (\sigma_r / \sigma_o - 1)^2 + (\mu_r / \mu_o - 1)^2}$$

(7)
where \( r \) is the Pearson correlation coefficient, \( \sigma_s \) and \( \mu_s \) are the standard deviation and the mean of simulated streamflow, \( \sigma_o \) and \( \mu_o \) are the standard deviation and the mean of observed streamflow. KGE ranges between \(-\infty\) and 1 (optimal value). An initial observation period of 1 year (2002) is used for model spin-up, and the period 2003–2016 is considered for parameter calibration.

To represent the background error covariance in the DA framework, a proper ensemble of SM states is generated considering an adequate model error characterization. Errors in model representation of SM state variables are a consequence of both errors in forcing data and errors in estimated parameters and model structure.

In this study, model errors are simulated through direct perturbation of model climate forcing (i.e., precipitation and temperature inputs) and SM states (\( W_{15} \) and \( W_{25} \)) according to appropriate error distributions. The approach implicitly represents both parameter and model-structural errors by directly perturbing model states (e.g., Alvarez-Garreton et al., 2014; Mao et al., 2019; Massari et al., 2018). All generated errors are assumed temporally independent, and the ensemble size is set to 100 members. An open loop (OL) run without data assimilation, with ensembles generated as described above, is also considered as a reference for SM DA assessment (considering the ensemble mean).

A multiplicative error is assumed for precipitation (Tian et al., 2013) based on a lognormal distribution with mean one and a standard deviation of \( \sigma_{eP} \). Inaccuracies in precipitation data are typically considered as the main random error source in hydrologic model forecasts (e.g., Han et al., 2014; Massari et al., 2018); therefore, given the high impact of rainfall on SM and runoff variability, the dimensionless \( \sigma_{eP} \) value is calibrated for every catchment to satisfy an ensemble test on simulated discharges. A test, proposed by De Lannoy et al. (2006) and adapted to river discharge, checks the consistency between the simulated streamflow uncertainty, quantified by the temporal average of the ensemble spread, \( \langle sp \rangle \), and the observed forecasting error, expressed by the temporal average of the ensemble skill, \( \langle sk \rangle \), with:

\[
\langle sp \rangle = \frac{1}{T_p} \sum_{k=1}^{T_p} \left[ \frac{1}{N} \sum_{i=1}^{N} (Q_{i,k} - \langle Q \rangle_k)^2 \right]
\]

\[
\langle sk \rangle = \frac{1}{T_p} \sum_{k=1}^{T_p} \left[ (\langle Q \rangle_k - Q_{o,k})^2 \right]
\]

where \( Q_{i,k} \) is the simulated discharge of the \( i \)-th ensemble member at time \( t_k \), \( \langle Q \rangle_k \) is the mean over the ensemble, \( Q_o \) the observed discharge, \( N \) is the ensemble size and \( T_p \) the number of time intervals. The precipitation error parameter \( \sigma_{eP} \) is calibrated to ensure that \( \langle sp \rangle/\langle sk \rangle \) approaches unity.

For the other forcing and state errors, an additive model is used based on an assumed zero-mean Gaussian distribution. Regarding temperature uncertainties, the error standard deviation \( \sigma_{e,T} \) is fixed equal to 1°C. The state vector consisting of \( W_{15} \) and \( W_{25} \) model variables is instead perturbed according to the error covariance matrix \( \Sigma \):

\[
\Sigma = \begin{bmatrix} \sigma_{e,W_{15}}^2 & 0 \\ 0 & \gamma^2 \sigma_{e,W_{25}}^2 \end{bmatrix}
\]

where error variance \( \sigma_{e,W_{15}}^2 \) corresponds to an error standard deviation of 0.01, and \( \gamma \) is the ratio of standard deviation in \( W_{25} \) to standard deviation in \( W_{15} \) time series, produced during the deterministic run. Defining \( \gamma \) in this way provides noise magnitudes in the two layers that is coherent in terms of their ratio with unperturbed SM variability.

### 3.3. Observations Rescaling and Error Characterization

For each ESA CCI SM product, daily SM data are averaged over the catchments. Time series of the \( W_{15} \) ensemble mean obtained during the OL simulations are then considered as reference data for the remotely sensed SM rescaling.
A widely used rescaling technique is CDF-matching (Reichle & Koster, 2004) that seeks to match the complete cumulative density function (CDF) of satellite and model estimates by applying a nonlinear operator. The CDF-matching is applied here following Drusch et al. (2005) and using a sixth order polynomial function to correct CDF differences.

CDF-matching is generally used in a “lumped” mode, by considering, as a whole, the model and observational multi-annual data sets. To account for possible changes in climatological differences at seasonal time scale (Kumar et al., 2015), we calculated the rescaling parameters separately for each month.

During the period considered, the actual temporal coverage of the remotely sensed SM data can be highly variable due to several factors: The temporal frequency of ESA CCI contributing missions (i.e., poor availability of active SM retrievals prior to 2007), seasonal factors (e.g., frozen or snow covered soil at high altitudes and latitudes during winter period) or data gaps attributable to the ESA CCI processing and quality assessment chain (sometimes found in the passive and combined data sets). Therefore, a minimum sample size of 150 remotely sensed SM retrievals is imposed to obtain reliable polynomial coefficients at a monthly scale.

Prior to DA, errors in rescaled SSM* must be characterized in statistical terms. Observation random errors are modeled as additive, temporally uncorrelated and Gaussian-distributed with zero-mean and stationary error variance $\sigma^2_{e,obs}$ which needs to be estimated.

Several error characterization methods are available (e.g., Crow & Reichle, 2008; Dorigo et al., 2010; Gruber et al., 2016). Here, the error characterization is performed with the aim of evaluating the optimal relative weight of satellite-based observations with respect to the background modeled SM as to improve streamflow simulations. In other words, the observation error variance is selected by individually maximizing assimilation efficiency (expressed through KGE index) in each study catchment (e.g., Loizu et al., 2018; Massari et al., 2015; Matgen et al., 2012; Patil & Ramsankaran, 2017).

### 3.4. Performance Metrics

The aforementioned KGE index is used as main indicator for DA assessment. Since it is sensitive to large errors (Garcia et al., 2017; Pool et al., 2018; Santos et al., 2018) commonly observed at flow peaks, an KGE index value computed on the untransformed discharges is considered here as a criterion for evaluating the high-flow conditions. To evaluate the performance of the hydrological simulations under different flow conditions, two transformed variants of the KGE skill metric are also adopted to put more emphasis on specific flow ranges. Specifically, the KGE index computed on root-squared-transformed discharges, referred to as $\Delta KGE_{sqr}$ provides more balanced information (Garcia et al., 2017) and is considered here to be more suitable for the evaluation of average-flow conditions. Likewise, the KGE index is also applied to the inverse of the discharge, referred to as $\Delta KGE_{inv}$, to put more weight on low-flow periods (Santos et al., 2018). To deal with zero discharge values, a small constant equal to 1/100 of the mean observed streamflow is added to the discharge values before transformation. Mutual differences in performance metrics between the DA and OL cases are also evaluated and indicated as $\Delta KGE$, $\Delta KGE_{sqr}$, and $\Delta KGE_{inv}$, respectively; positive values imply an improvement over the OL performance.

### 4. Results and Discussion

#### 4.1. Model Performances

Following individual calibration in each catchment, relatively good KGE index values (median value equal to 0.773) are obtained. Calibration runs also exhibit satisfactory $KGE_{sqr}$ (median equal to 0.763). However, low-flow conditions reveal poor $KGE_{inv}$ results (median equal to −0.888). The distributions of performance metrics across all catchments are presented in Figure 3.

The model OL performance metrics are then estimated considering the ensemble mean of simulated discharges in the OL run, with model ensembles obtained as described in Section 3.2. OL KGE values are always above a threshold value of 0.5, indicating model robustness in the selected catchments, and generally
show very limited differences compared to KGE values for the calibration run. Although still unsatisfactory, larger improvements (vs. the calibrated run) are noted for KGE_{inv} results (Figure 3). The spatial distribution of the OL KGE index, shown in Figure 4, is largely explained by that of the calibrated precipitation error $\sigma_{\varepsilon,P}$—which, in turn, show spatial similarities with the independent rainfall uncertainty estimates provided with the E-OBS rainfall product (Figure S2). OL performance within each study catchment is summarized in Table S3.

### 4.2. Remotely Sensed Soil Moisture Rescaling

The rescaling between satellite observations and the ensemble mean of $W_{1\%}$ for the OL run is performed individually within each catchment during the 2003–2016 period with the described monthly CDF-matching approach.

![Figure 3](image1.png)

**Figure 3.** Boxplots of: (a) KGE, (b) KGE_{sqr}, and (c) KGE_{inv} values in the study basins for the calibration and OL runs. For each box, the central mark refers to the median, while the bottom and top edges indicate the 25 and 75 percentiles, respectively. The whiskers extend from minimum to maximum values. KGE, Kling-Gupta Efficiency.

![Figure 4](image2.png)

**Figure 4.** OL KGE index values obtained from the ensemble mean of simulated discharges over the 2003–2016 period. KGE, Kling-Gupta Efficiency; OL, open loop.
The rescaling of satellite data is not performed on catchments that do not meet the data availability threshold imposed above to obtain reliable polynomial coefficients. As a result, the number of catchments considered in DA experiments is 775, 716, and 770 for the assimilation of active, passive, and combined products, respectively. The final availability of rescaled observations for each ESA CCI SM product is shown in Figure S4.

Correlation coefficients \( r_{SM} \) between the modeled and satellite-derived surface SM for the ESA CCI: (a) Active, (b) Passive, and (c) Combined SM products within each study catchment. ESA CCI, European Space Agency Climate Change Initiative; SM, soil moisture.

The distribution of correlation coefficients within the catchment set is illustrated in Figure 5. Median \( r_{SM} \) values are equal to 0.63, 0.44, and 0.56 for the active, passive, and combined products, respectively. The active-derived SSM* product captures the most temporal soil moisture variability in the reference modeled SM, with \( r_{SM} \) values that are generally above 0.5. Correlation coefficients exhibit distinguishable spatial patterns, with the best agreement obtained in Spain and United Kingdom. The lowest \( r_{SM} \) values are found at the highest latitudes (i.e., Iceland, Scandinavia) and in mountainous regions (i.e., the Alps, Pyrenees).
where climatic and topographic factors negatively influence the quality of satellite-based SM retrievals. The use of the combined product essentially confirms the same results—although with slightly lower $r_{SM}$ values overall. The assimilation of passive-only SM retrievals leads to significant deviations in SM temporal patterns, resulting in correlation coefficients that are generally below 0.5—with widespread higher exceptions in Spain.

### 4.3. Observation Error Variance Calibration Results

The observation error variance employed during assimilation, $\sigma_{\varepsilon,\text{obs}}^2$, is calibrated for each catchment to maximize the KGE index of the DA streamflow analysis over the period 2003–2016. To explore the DA effects, KGE is optimized considering a $\sigma_{\varepsilon,\text{obs}}^2$ range between zero and one with SM expressed in terms of saturation degree. The calibration of $\sigma_{\varepsilon,\text{obs}}^2$, as well as the extension of its search range, are supported by the goal of assessing the actual magnitude of improvements obtainable in streamflow prediction, as well as the need to search for $\sigma_{\varepsilon,\text{obs}}$ values that can improve the—already relatively good—performance achieved by the baseline OL case.

Figure 6. Calibrated observation standard error values $\sigma_{\varepsilon,\text{obs}}$ for ESA CCI: (a) Active, (b) Passive, and (c) Combined SM products within each study catchment. ESA CCI, European Space Agency Climate Change Initiative.
Figure 6 describes the spatial distributions of calibrated $\sigma_{\text{cal}}$ values over the study domain for each ESA CCI SM product. The median $\sigma_{\text{cal}}$ values are 0.33, 0.45, and 0.41 for the active, passive, and combined products, respectively. The high variability in calibrated $\sigma_{\text{cal}}$ underscores the importance of accurately specifying $\sigma_{\text{cal}}$ to represent spatially varying observation errors for optimized DA performance.

4.4. Data Assimilation Impacts on Simulated Streamflow and Relation With SM Accuracy

The analysis of DA results is carried out by comparing the skill obtained for each of the three ESA CCI SM products within the DA experiment with respect to the OL both in space (i.e., spatial patterns) and time (i.e., seasonal variations).

Boxplots of $\Delta KGE$, $\Delta KGE_{\text{sqr}}$, and $\Delta KGE_{\text{inv}}$ metrics for each ESA CCI SM product are shown in Figure 7, with whiskers being extended between the 10% and 90% quantiles. Very slight improvements in the KGE index are generally found following DA, and median values of $\Delta KGE$ are equal to 0.0048, 0.0022, and 0.0033 for active, passive, and combined products, respectively. KGE differences larger than 0.01 are found in between only 15% and 35% of the study catchments—depending on the ESA CCI SM product assimilated. Conversely, deterioration in streamflow predictions are observed in about 25% of the catchments, regardless of the product used. These results confirm previous outcomes observed in the assimilation of satellite SM literature (e.g., Kumar et al., 2014; Lievens et al., 2016), in terms of the limited improvement in streamflow simulations associated with soil moisture DA. Specific DA performances within each study catchment are reported in Table S3. In this experiment, the SM assimilation mainly improves the KGE index through an enhancement in correlation between simulated and observed discharge; however, the latter is not generally due to an improvement in streamflow covariance, but rather to a decrease in the simulated discharge variance.

The assimilation of the active-derived retrievals generally leads to better DA results than the combined retrievals. While the assimilation of the passive retrievals provides the worst DA performance (as expected from its relatively poor quality over many catchments located in Central Europe and at high latitudes, see Figures 5 and S5). In this respect, Figure 8 highlights the agreement between distributions of $r_{\text{sm}}$ and $\Delta KGE$ values (i.e., higher correlations with modeled SM correspond to greater improvements in streamflow simulations). In other words, the assimilation of better products tends to produce better streamflow outcomes.

Figure 9 shows the spatial distribution of differences in KGE between the DA and OL cases; blue shading indicates skill improvements and red shading skill degradations. The spatial pattern of improvements in KGE are quite similar across the three ESA CCI SM data sets. However, active and combined retrieval assimilation results in a net positive impact in central Europe and in the Mediterranean area (i.e., Spain, Italy) providing greater benefits to the discharge model predictions compared to passive product. These regional findings are consistent with those reported in other DA studies investigating the influence of different satellite soil moisture products on hydrological predictions (e.g., Azimi et al., 2020; Cenci et al., 2016;
Laiolo et al., 2016), which also highlighted the positive effect of active SM retrievals for improving runoff prediction over Mediterranean climates. In particular, Figure 9 illustrates that passive retrieval assimilation is somewhat less effective in central Europe. However, across all of Europe, the assimilation of the passive product essentially improves KGE within the same portion of catchments as the active product.

In general, at northern latitudes (e.g., the UK, Scandinavia) DA improvements are minimal and degradation is common. This could be due to snowmelt-dominated runoff in these regions (Berghuijs et al., 2019) that intuitively results in negligible SM data assimilation efficiency (e.g., Koster et al., 2018; Reichle et al., 2019). Some performance deteriorations after DA, not fully explained by the above-mentioned factor, can also be interpreted as a consequence of local climate conditions (as the deteriorations are also observed at lower latitudes during the cold season, see Figure 10).

Poor results can also be observed over western and northern hilly areas of UK. These areas are characterized by a relatively high rainfall accumulation and high soil moisture (for instance, in these areas, average values for the ensemble $W_1$ in OL range between 0.5 and 0.8, while in the rest of the study domain they usually vary between 0.05 and 0.35). A potential explanation for the less satisfactory performance after DA relies upon the sensitivity of runoff generation to SM: When SM is nearly saturated, runoff is largely controlled by precipitation inputs and less by pre-storm SM. This is confirmed by the study of Massari and Camici (2020) who show—in the same study area and with similar data—that, within the UK, pre-storm river discharge is a better proxy of basin initial conditions prior flood events than SM.

In Figure S5, the distribution of $\Delta$KGE values is represented with reference to the main climate types present within the catchment data set (according to the Koppen climate classification provided in Chen & Chen, 2013). Mediterranean climates with dry summer are characterized by the greatest DA benefits, while the added value of satellite SM is less significant in fully humid climates, such as those occurring in the higher-middle latitudes (a condition characterizing more than half of the basins) and even more in high-latitude areas (i.e., Scandinavian basins).

Similar conclusions can be drawn for the KGE$_{avg}$ index (focusing on average-flow conditions) despite values of $\Delta$KGE$_{avg}$ that are generally smaller than $\Delta$KGE. However, a negative assimilation impact is found for low-flow conditions, with a worsening of the KGE$_{inv}$ index in more than half of the study catchments. This
degradation is most common for the active product and less common for the assimilation of passive retrievals. The enhanced impact of SM DA for simulating floods, as opposed to low flow conditions, is consistent with Baugh et al. (2020).

The impact of DA during different seasons of the year is assessed by splitting the analysis of streamflow time series into warm- (April-September) and cold- (October-March) season halves. A seasonal analysis of OL performances (not shown) exhibits higher cold-season KGE values in some areas (e.g., the UK, France, Central Europe), while elsewhere no significant seasonal differences are found. In Figure 10, the DA results in terms of ΔKGE are reported for the combined data set; the other ESA CCI SM products confirm similar outcomes. Positive assimilations impacts are generally found during the warm season—reflecting the same spatial pattern found when considering the whole DA period. However, degradation is much more common during the cold season when most catchments register a negative effect after DA—with the exception of catchments in Spain and other Mediterranean areas. Variations in OL performances (see Section 4.5) can also help to explain seasonal variations in DA efficiency: Positive DA impact are more common in the warm season when the OL streamflow KGE values are generally lower. Naz et al. (2019), who evaluated the effects of ESA CCI combined SM assimilation in a LSM over a gridded European domain, similarly obtained larger DA improvements in the simulation of runoff generation during the summer vs. the winter.

Figure 9. ΔKGE values for the ESA CCI (a) Active, (b) Passive, and (c) Combined SM products within each study catchment. ESA CCI, European Space Agency Climate Change Initiative; KGE, Kling-Gupta Efficiency; SM, soil moisture.
4.5. Model Skills in Open Loop and Remotely Sensed Product Accuracy

Spatial and seasonal patterns within the study area suggest a strong relationship between OL model performance and DA effects. This is a well-known behavior when looking at SM assimilation impacts on LSM SM state estimates (e.g., Reichle et al., 2008). However, this effect has not been described in detail for the impact of SM assimilation on LSM fluxes.

Figure 11 summarizes the results for assimilation of the active SM product, with other ESA CCI SM products showing similar outcomes. By comparing ΔKGE and ΔKGE_sqr values with corresponding OL model scores, we find the largest DA improvement in medium-high flows within catchments characterized by poor OL performance. However, the opposite is true for ΔKGE_inv results—which show the largest degradation in catchments where the OL performance is already poor during low-flow conditions. This suggests that during low-flow conditions the model structure can exert a non-negligible role on the DA efficiency. This is a noteworthy example that shows that the DA cannot correct for model structural deficiencies and can, in certain cases, even degrade already poor results.

4.6. Vegetation and Topographic Characteristics and Area of the Catchments

The impact of vegetation density and topographic complexity conditions known to affect the quality of satellite SM retrieval results is also analyzed.

The accuracy of satellite data is generally expected to be inversely proportional to vegetation density and topographic complexity. In Figure 12, the frequency of ΔKGE values corresponding to the ancillary data provided within ESA CCI SM products is reported. The color patterns for both active and passive products show only limited evidence of the expected trends for differences in runoff simulation performance. For both indices, larger improvements should concentrate in the upper left corner, while smaller ΔKGE or performance decrements should be located in the lower right corner.

In some cases, other factors appear to have a larger impact than vegetation density or topographic characteristics. For example, catchments with VOD and topographic complexity index higher than 0.5% and 12%, respectively, mostly located in the Alps, Pyrenees, Cantabrian Mountains, northern Apennines, are characterized by DA enhancements attributable to poor OL predictions. On the other hand, catchments...
with the lowest VOD (localized in Iceland, Scandinavia, Denmark, UK) and topographic complexity values (concentrated in Finland, southern England, northern Germany) exhibit negligible DA improvements, due to either climatic effects or good underlying OL performance.

To understand potential impacts of the area of the basins on assimilation results, and a possible relationship with SM spatial resolution limitations, we plotted ΔKGE as a function of basin area in Figure 13. However, results do not suggest a clear relationship exists between catchment area and ΔKGE.

5. Conclusions

In this study, the suitability of remotely sensed SM DA to improve hydrological model performances is evaluated through an experiment involving more than 700 European catchments with different climate, topographic and landcover characteristics. The use of the different ESA CCI SM products from active and passive microwave sensors is also investigated. Such an extensive analysis allows for the examination of exhaustive and non-case-specific indications concerning key conditions that determine the assimilation impact on runoff modeling in the study domain, as well as to the extent the streamflow simulations can generally be improved via SM DA.

The added value of remotely sensed SM assimilation compared to OL simulation appears to be weak for the majority of investigated catchments. The assimilation provided noticeable KGE improvements (i.e., larger than 0.01) in only 15%–35% of the catchments, depending on the ESA CCI SM product, whereas ~25% of the catchments showed small, almost negligible, deteriorations in streamflow predictions. The DA improvement in the forecast skill is found to be higher in the warm season from April to September. These results are substantially confirmed when also considering the KGEinv index, computed on root-squared-transformed discharges to avoid applying excessive weight to high-discharge values. DA generally worsens the KGEinv index focusing on low flow values; nevertheless, it should be noted that this result

Figure 11. Comparison between OL scores and the distribution of differences after DA of the ESA CCI active SM product, for the: (a) KGE, (b) KGEinv, and (c) KGEsqr metrics. The whiskers are limited to 1.5 times the distance between upper and lower quartiles. DA, Data assimilation; ESA CCI, European Space Agency Climate Change Initiative; KGE, Kling-Gupta Efficiency; OL, open loop; SM, soil moisture.
may be related to model structure and calibration practices. In this perception, SM assimilation can have a different impact if different model structures are considered as also found by Loizu et al. (2018). Thus, care should be taken in the result interpretation when using a fixed conceptual modeling of processes (under the “one model fits all” paradigms, i.e., a single general model applicable to every catchment, Fenicia et al., 2011) and a small sample of catchments.

Special attention is paid to the evaluation of various factors potentially affecting DA impacts on streamflow simulation. In this regard, our main findings can be summarized as follows.

• The OL model skill is found to play an important role in the assessment of DA effects and contributes significantly to differences in assimilation performances observed within the study area at a seasonal scale. In this sense, the integration of satellite SM information contributes more added value to poorer OL streamflow simulations. However, even in this case, only limited improvements are seen (e.g., for KGE values in OL lower than 0.55, the median ΔKGE is 0.0295).

• Remotely sensed product accuracy is also found to play a role in DA effects on runoff. Among the investigated ESA CCI SM products, the active data set is generally the most accurate in reproducing true SM values and, when assimilated, provides the highest runoff improvement within the study catchments (median ΔKGE = 0.0048), followed by the combined (median ΔKGE = 0.0033) and passive (median ΔKGE = 0.0022) ESA CCI SM products.

Figure 12. Frequency of ΔKGE values with respect to both VOD (top row) and topographic complexity index (bottom row) for the assimilation of ESA CCI active (left column) and passive (right column) retrievals. ESA CCI, European Space Agency Climate Change Initiative; KGE, Kling-Gupta Efficiency; VOD, vegetation optical depth.

Figure 13. ΔKGE values after DA of the ESA CCI active product for different catchment area classes. The bar heights represent the mean ΔKGE value for each class, while the markers correspond to the 5 and 95 percentiles. DA, Data assimilation; ESA CCI, European Space Agency Climate Change Initiative; KGE, Kling-Gupta Efficiency.
• Climatic factors (e.g., snow-dominated or fully humid conditions) are responsible for some local inefficiency of satellite observation integration related to negative effects on retrievals; moreover, the modification of runoff generation dynamics or prolonged periods of high saturation can also reduce the sensitivity of runoff generation on pre-storm SM.

• Vegetation and orographic characteristics of the catchments provide only a partial explanation of differences in runoff estimation skill between the DA and OL.

Despite the large number of catchments considered in our analysis, we cannot exclude the possibility that additional factors (e.g., rescaling technique, choice of hydrological model and choice of assimilation filter) also significantly affected our results. Therefore, future investigations should consider the inclusion of different experiment configurations addressing specific issues to provide a clearer picture of the benefit of SM assimilation into hydrological models and the appropriate parameterization of a DA scheme based on catchment characteristics and data availability.

Data Availability Statement

Some results reported in the supporting information were generated using the QA4SM service (https://qa4sm.eodc.eu/). The service was developed by TU Wien GEO and AWST GmbH. The results were generated on November 2019. The supporting data sets are available in Zenodo repository (http://doi.org/10.5281/zenodo.3908392).

Acknowledgments

The authors acknowledge the data providers and specifically: (a) for satellite soil moisture data, the ESA Climate Change Initiative for Soil Moisture (https://www.esa-soilmoisture-cci.org); (b) for precipitation and temperature data, the E-OBS data set from the EU-FP6 project UERRA (http://www.uerra.eu) and the Copernicus Climate Change Service, and the data providers in the EC4D project (https://www.ec4d.eu); (c) for river discharge data and geospatial data products, the Global Runoff Data Centre (https://www.bafg.de/GRDC), Federal Institute of Hydrology (BfG), Koblenz, Germany, that hosts the Global Runoff Data Base (GRDB) and the European Water Archive (EWA); the Italian Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRRA) for Hydrological Information System (HIS) data (http://www.hiscentral.isprambiente.gov.it); the Spanish Centro de Estudios hidrográficos (CEH) of Centro de Estudios y Experimentación de Obras Públicas (CEDEX) (http://ceh-flumen64.cedex.es); and the Portuguese Sistema Nacional de Informação de Recursos Hídricos (SNIRH) (https://snirh.pt). Figure S3 was generated using the QA4SM service (https://qa4sm.eodc.eu/). The service was developed by TU Wien GEO and AWST GmbH. The figure was generated on November 2019.

References

Albergel, C., Munier, S., Leroux, D. J., Dewaele, H., Fairbairn, D., Barbuc, A. L., et al. (2017). Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX v.6.0. LDAS-Monde assessment over the Euro-Mediterranean area. Geoscientific Model Development, 10(10), 3883–3912. https://doi.org/10.5194/gmd-10-3889-2017

Alvarez-Garretón, C., Ryu, D., Western, A. W., Crow, W. T., & Robertson, D. E. (2014). The impacts of assimilating satellite soil moisture into a rainfall-runoff model in a semi-arid catchment. Journal of Hydrology, 519, 2763–2774. https://doi.org/10.1016/j.jhydrol.2014.07.041

Alvarez-Garretón, C., Ryu, D., Western, A. W., Su, C.-H., Crow, W. T., Robertson, D. E., & Leahy, C. (2015). Improving operational flood ensemble prediction by the assimilation of satellite soil moisture: Comparison between lumped and distributed systems. Hydrology and Earth System Sciences, 19(4), 1659–1676. https://doi.org/10.5194/hess-19-1659-2015

Azimi, S., Dariane, A. B., Modanesi, S., Bauer-Marschallinger, B., Bindlish, R., Wagner, W., & Massari, C. (2020). Assimilation of Sentinel 1 and SMAP—based satellite soil moisture retrievals into SWAT hydrological model: The impact of satellite revisit time and product spatial resolution on flood simulations in small basins. Journal of Hydrology, 581, 124367. https://doi.org/10.1016/j.jhydrol.2019.124367

Baugh, C., de Rosnay, P., Lawrence, H., Jurina, T., Drusch, M., Zoeter, E., & Frundhomme, C. (2020). The impact of SMOS soil moisture data assimilation within the operational global flood awareness system (GloFAS). Remote Sensing, 12(9), 1490. https://doi.org/10.3390/rs12091490

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The relative importance of different flood-generating mechanisms across Europe. Water Resources Research, 55(6), 4582–4593. https://doi.org/10.1029/2019wr024841

Bergström, S. (1995). The HBV model. In V. P. Singh (Ed.), Computer models of watershed hydrology (pp. 443–476). Highlands Ranch, CO: Water Resources Publications.

Berthet, L., Andréassian, V., Perrin, C., & Javelle, P. (2009). How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. Hydrology and Earth System Sciences, 13(6), 819–831. https://doi.org/10.5194/hess-13-819-2009

Brocca, L., Moramarco, T., Melone, F., Wagner, W., Hasenauer, S., & Hahn, S. (2012). Assimilation of surface- and root-zone ASCAT soil moisture products into rainfall-runoff modeling. IEEE Transactions on Geoscience and Remote Sensing, 50(7), 2542–2555. https://doi.org/10.1109/tgrs.2011.2177468

Burnash, R. J. C., Ferral, R. L., & McGuire, R. A. (1973). Computer models of watershed hydrology (pp. 443–476). Highlands Ranch, CO: Water Resources Publications.

Cenci, L., Massari, C., Ciabatta, L., Marchesini, I., & Brocca, L. (2020). Which rainfall score is more informative about the performance in river discharge simulation? A comprehensive assessment on 1318 basins over Europe. Hydrology and Earth System Sciences, 24(10), 4869–4885. https://doi.org/10.5194/hess-24-4869-2020

Chen, D., & Chen, H. W. (2013). Using the Köppen classification to quantify climate variation and change: An example for 1901–2010. Environmental Development, 6, 69–79. https://doi.org/10.1016/j.envdev.2013.03.007

Chen, D., Crow, W. T., Bindlish, R., Collander, A., Burgin, M. S., Asanuma, J., & Aida, K. (2018). Global-scale evaluation of SMAP, SMOS, and ASCAT soil moisture products using triple collocation. Remote Sensing of Environment, 214, 1–13. https://doi.org/10.1016/j.rse.2018.05.008

Chen, F., Crow, W. T., Starks, P. J., & Moriasi, D. N. (2011). Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. Advances in Water Resources, 34(4), 526–536. https://doi.org/10.1016/j.adwres.2011.01.011
Crow, W. T., Chen, F., Reichle, R. H., Xia, Y., & Liu, Q. (2018). Exploiting soil moisture, precipitation, and streamflow observations to evaluate soil moisture/precipitation coupling in land surface models. *Geophysical Research Letters*, 45(10), 4869–4878. https://doi.org/10.1029/2018gl077193

Crow, W. T., & Reichle, R. H. (2008). Comparison of adaptive filtering techniques for land surface data assimilation. *Water Resources Research*, 44(8). https://doi.org/10.1029/2008wr006883

De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., & Veroehst, N. E. C. (2006). Assessment of model uncertainty for soil moisture through ensemble verification. *Journal of Geophysical Research*, 111. https://doi.org/10.1029/2005jd006367

Dorigo, W., Wagner, W., Albergel, C., Albrecht, P., Balsamo, G., Brocca, L., et al. (2017). ESA CCI soil moisture for improved Earth system understanding: State-of-the art and future directions. *Remote Sensing of Environment*, 203, 185–215. https://doi.org/10.1016/j.rse.2017.07.001

Dorigo, W. A., Scipal, K., Parinussa, R. M., Liu, Y. Y., Wagner, W., de Jeu, R. A. M., & Naeimi, V. (2012). Trend-preserving blending of passive microwave soil moisture and brightness temperature products into a land surface model. *Remote Sensing of Environment*, 123(10), 2605–2616. https://doi.org/10.1016/j.rse.2012.09.016

Drusch, M., Wood, E. F., & Gao, H. (2005). Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture. *Geophysical Research Letters*, 32, L15403. https://doi.org/10.1029/2005gl023623

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(CS5), 10143–10162. https://doi.org/10.1029/94js0572

Fairbairn, D., Barbu, A. L., Napoly, A., Albergel, C., Mahfouf, J.-F., & Calvet, J.-C. (2017). The effect of satellite-derived surface soil moisture and leaf area index land data assimilation on streamflow simulations over France. *Hydrology and Earth System Sciences*, 21(4), 2015–2033. https://doi.org/10.5194/hess-21-2015-2017

Famiglietti, J. S., & Wood, E. F. (1994). Multiscale modeling of spatially variable water and energy balance processes. *Water Resources Research*, 30(11), 3061–3078. https://doi.org/10.1029/94wr01498

Fenicia, F., Kavetski, D., & Savenije, H. H. G. (2011). Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, 47(11), W11510. https://doi.org/10.1029/2010wr010174

Garcia, F., Folton, N., & Oudin, L. (2017). Which objective function to calibrate rainfall-runoff models for low-flow index simulations? *Hydrological Sciences Journal*, 62(7), 1149–1166. https://doi.org/10.1080/02626667.2017.1308511

Gruber, A., Scanlon, T., van der Schalie, R., Wagner, W., & Dorigo, W. (2019). Evolution of the ESA CCI soil moisture climate data records and their underlying merging methodology. *Earth System Science Data*, 11(2), 717–739. https://doi.org/10.5194/essd-11-717-2019

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., & Wagner, W. (2016). Recent advances in (soil moisture) triple collocation analysis. *International Journal of Applied Earth Observation and Geoinformation*, 45, 200–211. https://doi.org/10.1016/j.jag.2015.09.002

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modeling. *Journal of Hydrology*, 377(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Han, E., Crow, W. T., Holmes, T., & Bolten, J. (2014). Benchmarking a soil moisture data assimilation system for agricultural drought monitoring. *Journal of Hydrometeorology*, 15(3), 1117–1134. https://doi.org/10.1175/jhm-d-13-0125.1

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, 113, D20119. https://doi.org/10.1029/2008jd010201

Karthikeyan, L., Pan, M., Wanders, N., Kumar, D. N., & Wood, E. F. (2017). Four decades of microwave satellite soil moisture observations: Part 1. A review of retrieval algorithms. *Advances in Water Resources*, 109, 106–120. https://doi.org/10.1016/j.advwatres.2017.09.006

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K., & Puma, M. J. (2009). On the nature of soil moisture in land surface models. *Journal of Climate*, 22(16), 4322–4335. https://doi.org/10.1175/2009jcli2832.1

Koster, R. D., Liu, Q., Mahanama, S. P. P., & Reiche, R. H. (2018). Improved hydrological simulation using SMAP data: Relative impacts of model calibration and data assimilation. *Journal of Hydrometeorology*, 19(4), 727–741. https://doi.org/10.1175/jhm-d-17-0238.1

Koster, R. D., & Milly, P. C. D. (1997). The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *Journal of Climate*, 10(7), 1578–1591. https://doi.org/10.1175/1520-0442(1997)10(1578:tibtao)2.0.co;2

Kumar, S. V., Peters-Lidard, C. D., Mocke, D., Reichle, R., Liu, Y.,Arsenault, K. R., et al. (2014). Assimilation of remotely sensed soil moisture and snow depth elevations for drought estimation. *Journal of Hydrometeorology*, 15(6), 2446–2469. https://doi.org/10.1175/jhm-d-13-0132.1

Kumar, S. V., Peters-Lidard, C. D., Santanello, J. A., Reiche, R. H., Draper, C. S., Koster, R. D., et al. (2015). Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes. *Hydrolgy and Earth System Sciences*, 19(11), 4463–4478. https://doi.org/10.5194/hess-19-4463-2015

Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Rudari, R., et al. (2016). Impact of different satellite soil moisture products on the predictions of a continuous distributed hydrological model. *International Journal of Applied Earth Observation and Geoinformation*, 58, 131–145. https://doi.org/10.1016/j.jag.2015.06.002

Lievens, H., De Lannoy, G. J. M., Al Bitar, A., Drusch, M., Dumeadah, G., Hendricks Franssen, H.-J., et al. (2016). Assimilation of SMOS soil moisture and brightness temperature products into a land surface model. *Remote Sensing of Environment*, 180, 292–304. https://doi.org/10.1016/j.rse.2015.10.033

Liu, Y., Weerts, A. H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., et al. (2012). Advancing data assimilation in operational hydrologic forecasting: Progresses, challenges, and emerging opportunities. *Hydrolgy and Earth System Sciences*, 16(10), 3863–3887. https://doi.org/10.5194/hess-16-3863-2012

Liu, Y. Y., Dorigo, W. A., Parinussa, R. M., de Jeu, R. A. M., Wagner, W., McCabe, M. F., et al. (2012). Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment*, 123, 280–297. https://doi.org/10.1016/j.rse.2012.03.014

López, I., Massari, C., Álvarez-Mozos, J., Tarpanelli, A., Brocca, L., & Casali, J. (2018). On the assimilation set-uped ASCAT soil moisture data for improving streamflow catchment simulation. *Advances in Water Resources*, 111, 86–104. https://doi.org/10.1016/j.advwatres.2017.10.034
Manfreda, S., Brocca, L., Moramarco, T., Melone, F., & Sheffield, J. (2014). A physically based approach for the estimation of root-zone soil moisture from surface measurements. *Hydrology and Earth System Sciences, 18*(3), 1199–1212. https://doi.org/10.5194/hess-18-1199-2014

Mao, Y., Crow, W. T., & Nijssen, B. (2019). A framework for diagnosing factors degrading the streamflow performance of a soil moisture data assimilation system. *Journal of Hydrometeorology, 20*(1), 79–97. https://doi.org/10.1175/jhm-d-18-0115.1

Massari, C., Brocca, L., Tarpanelli, A., & Moramarco, T. (2015). Data assimilation of satellite soil moisture into rainfall-runoff modelling: A complex recipe? *Remote Sensing, 7*(9), 11403–11433. https://doi.org/10.3390/rs70911403

Massari, C., & Camici, S. (2020). Antecedent wetness conditions of European floods: A comprehensive study. Paper presented at 2020 IEEE international geoscience and remote sensing symposium, IGARSS 2020, Virtual Symposium. IEEE. https://doi.org/10.1109/igars39084.2020.9324459

Massari, C., Camici, S., Ciabatta, L., & Brocca, L. (2018). Exploiting satellite-based surface soil moisture for flood forecasting in the Mediterranean area: State update versus rainfall correction. *Remote Sensing, 10*(2), 292. https://doi.org/10.3390/rs10020292

Matgen, P., Fenicia, F., Heitz, S., Plaza, D., de Keyser, R., Pauwels, V. R. N., et al. (2012). 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. *Advances in Water Resources, 44*, 49–65. https://doi.org/10.1016/j.adwatres.2012.03.022

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., et al. (2017). The future of Earth observation in hydrology. *Hydrology and Earth System Sciences, 21*(7), 3879–3914. https://doi.org/10.5194/hess-21-3879-2017

Mechlenburg, S., Drusch, M., Kaleschke, L., Rodríguez-Fernandez, N., Reul, N., Kerr, Y., et al. (2016). ESA’s soil moisture and ocean salinity mission: From science to operational applications. *Remote Sensing of Environment, 180*, 3–18. https://doi.org/10.1016/j.rse.2015.12.025

Melone, F., Corradini, C., & Singh, V. P. (2002). Lag prediction in ungauged basins: An investigation through actual data of the upper Tiber River valley. *Hydrological Processes, 16*(5), 1085–1094. https://doi.org/10.1002/hp.313

Nac, B. S., Curtis, W., Montzka, C., Sharples, W., Goergen, K., Keune, J., et al. (2019). Improving soil moisture and runoff simulations at 3 km over Europe using land surface data assimilation. *Hydrology and Earth System Sciences, 23*(1), 277–301. https://doi.org/10.5194/hess-23-277-2019

Patil, A., & Ramsankaran, R. (2017). Improving streamflow simulations and forecasting performance of SWAT model by assimilating remotely sensed soil moisture observations. *Journal of Hydrology, 555*, 683–696. https://doi.org/10.1016/j.jhydrol.2017.10.058

Penza, D., Tromp-van Meerveld, H. I., Gobbi, A., Borga, M., & Dalla Fontana, G. (2011). The influence of soil moisture on threshold runoff generation processes in an alpine headwater catchment. *Hydrology and Earth System Sciences, 15*(3), 689–702. https://doi.org/10.5194/hess-15-689-2011

Pool, S., Vis, M., & Seibert, J. (2018). Evaluating model performance: Towards a non-parametric variant of the Kling-Gupta efficiency. *Hydrological Sciences Journal, 63*(1–2), 1365–2041. https://doi.org/10.1080/02626667.2018.1552002

Reichle, R. H., Crow, W. T., Koster, R. D., Sharif, H. O., & Mahanama, S. P. P. (2008). Contribution of soil moisture retrievals to land data assimilation products. *Geophysical Research Letters, 35*(1), L01404. https://doi.org/10.1029/2007gl031966

Reichle, R. H., & Koster, R. D. (2004). Bias reduction in short records of soil satellite soil moisture. *Geophysical Research Letters, 31*, L19501. https://doi.org/10.1029/2004gl020938

Reichle, R. H., Liu, Q., Koster, R. D., Crow, W. T., De Lannoy, G. J. M., Kimball, J. S., et al. (2019). Version 4 of the SMAP level-4 soil moisture algorithm and data product. *Journal of Advances in Modeling Earth Systems, 11*(10), 3106–3130. https://doi.org/10.1029/2019ms001729

Rüdiger, M.-E., Madsen, H., Stisen, S., Bircher, S., & Fensholt, R. (2014). Assimilation of SMOS-derived soil moisture in a fully integrated hydrological and soil-vegetation-atmosphere transfer model in Western Denmark. *Water Resources Research, 50*(11), 8962–8981. https://doi.org/10.1002/2014wr015392

Rodell, M., Houser, P. R., Jambor, U., Gottschalch, J., Mitchell, K., Meng, C.-J., et al. (2004). The global land data assimilation system. *Bulletin of the American Meteorological Society, 85*(3), 381–394. https://doi.org/10.1175/bams-85-3-381

Santos, L., Thirié, G., & Perrin, C. (2018). Technical note: Pitfalls in using log-transformed flows within the KGE criterion. *Hydrology and Earth System Sciences, 22*(8), 4583–4591. https://doi.org/10.5194/hess-22-4583-2018

Spyrou, C., Vlasák, G., Pappa, A., Mentzaíou, A., Katsafados, F., Papadopoulos, A., et al. (2020). Implementation of a nowcasting hydro-meteorological system for studying flash flood events: The case of Mindria, Greece. *Remote Sensing, 12*(17), 2784. https://doi.org/10.3390/rs12172784

Tian, Y., Huffman, G. J., Adler, R. F., Tang, L., Sapiano, M., Maggioni, V., & Wu, H. (2013). Modeling errors in daily precipitation measurements: Additive or multiplicative? *Geophysical Research Letters, 40*(10), 2060–2065. https://doi.org/10.1002/2013gl05320

Tramblay, Y., Bovier, C., Martin, C., Didon-Lescot, J.-F., Todorovik, D., & Domergue, J.-M. (2010). Assessment of initial soil moisture conditions for event-based rainfall-runoff modelling. *Journal of Hydrology, 387*(3–4), 176–187. https://doi.org/10.1016/j.jhydrol.2010.04.006

Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., et al. (2020). A multiscale, hydrometeorological forecast evaluation of national water model forecasts of the May 2018 Ellicott City, Maryland, Flood. *Journal of Hydrometeorology, 21*(3), 475–499. https://doi.org/10.1175/jhm-d-19-0125.1

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., et al. (2013). The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications. *Meteorologische Zeitschrift, 22*(1), 53–55. https://doi.org/10.1127/0941-2948/2013/0399

Wagner, W., Lemoine, G., & Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sensing of Environment, 70*(2), 191–207. https://doi.org/10.1016/s0034-4257(99)00316-x

Wijayarathne, D. B., & Coulibaly, P. (2020). Identification of hydrological models for operational flood forecasting in St. John's, Newfoundland, Canada. *Journal of Hydrology: Regional Studies, 21*, 100646. https://doi.org/10.1016/j.jhrse.2019.100646