

A comprehensive approach to building a continuous hydrologic model with Soil Moisture Accounting using Earth Observation data

Vasiliki Thomopoulou ^{a,*}, Theano Iliopoulou ^a, Panagiotis Kossieris ^a, George Bariamis ^a, Ioannis Tsoukalas ^b, Andreas Efstratiadis ^a and Christos Makropoulos ^a

^a Department of Water Resources, Faculty of Civil Engineering, National Technical University of Athens, Heroon Polytechniou 5, GR-157 80 Zografou, Greece

^b Department of Civil Engineering, Democritus University of Thrace, Kimmeria Campus, GR-67100 Xanthi, Greece

*Corresponding author. E-mail: vasilikithomopoulou@mail.ntua.gr

 VT, 0009-0000-8361-6879; TI, 0000-0001-7053-4348; PK, 0000-0002-5201-3079; GB, 0000-0002-9548-3054; IT, 0000-0001-5272-7605; AE, 0000-0002-2907-1633; CM, 0000-0003-0308-4265

ABSTRACT

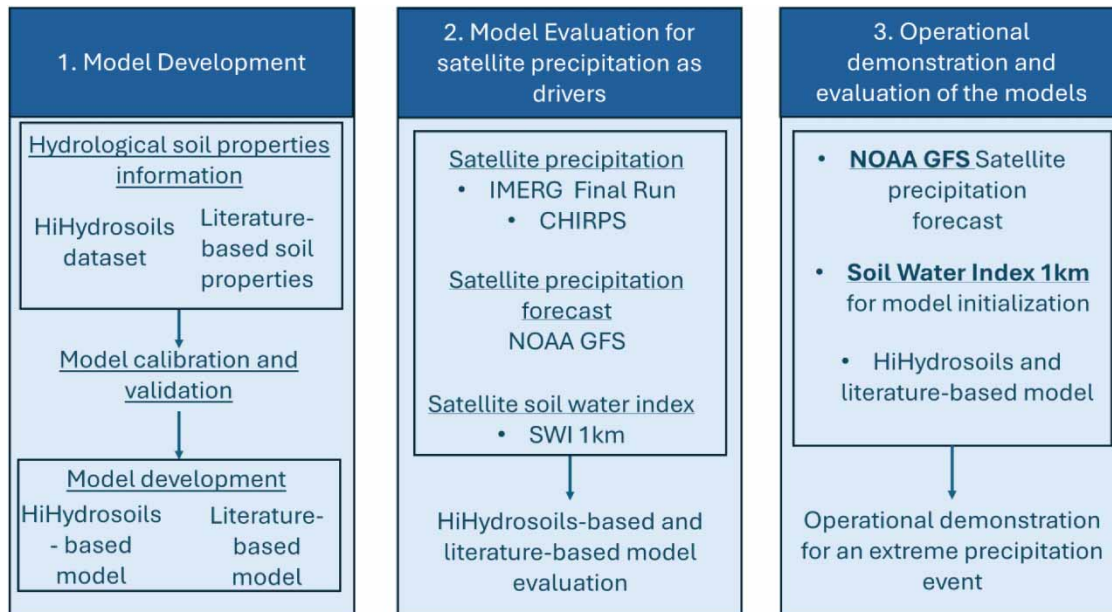
In recent years, both availability and interest in Earth Observations (EO) have increased due to their ability to provide information with extensive spatial coverage, which is valuable for data-scarce regions. This study provides a roadmap to exploit EO/satellite data (EO-based model) for continuous hydrological simulation using the Hydrologic Engineering Center–Hydrologic Modelling System model with a soil moisture accounting (SMA) component. As a case study, we consider the Boeotikos Kephisos River basin, in Greece. The SMA component is calibrated using the HiHydroSoils dataset, to map its parameters to the regionally varying soil hydraulic properties and a comparison is made to an alternative parameterization, following the standard literature approach (literature-based model). The effectiveness of satellite data in enhancing model performance is further assessed, by comparing three different satellite precipitation datasets, as model drivers, and by using satellite-based soil moisture for model initialization. The discussion extends to the potential for integration of EO/satellite data at the operational level, by simulating a significant precipitation event. Yet, the most promising result pertains to the opportunity to exploit satellite-derived estimates of soil hydraulic properties to base the calibration of the data-intensive SMA scheme, with the EO-based model significantly outperforming the literature-based model parameterization.

Key words: continuous hydrological simulation, EO data, satellite-based data, soil moisture, soil moisture accounting, soil properties

HIGHLIGHTS

- An open dataset related to soil properties is used to perform continuous hydrological simulation in a data-scarce region.
- Satellite data are incorporated in various steps of the process (precipitation, evapotranspiration, and soil moisture).
- Satellite precipitation and soil moisture data drive simulation.
- The model's efficiency is assessed in a real-world scenario (operational demonstration).

GRAPHICAL ABSTRACT



1. INTRODUCTION

Hydrological modelling, which entails the simulation of river basin processes in response to meteorological forcing, holds a prominent position in the planning, design, and management of water resource systems. The literature provides a plethora of hydrological models, which are differentiated with respect to various aspects (e.g., [Beven 2005](#); [Brocca et al. 2011](#); [Nalbantis et al. 2011](#); [Elga et al. 2015](#)), such as the level of detail of spatial structure (i.e., lumped, semi-distributed, or distributed), the temporal scale (i.e., event-based or continuous), the structure of the underlying process representation and parameterization (i.e., physical, empirical, conceptual, or data-driven), and the handling of uncertainty (deterministic or stochastic).

Over the last decades, advances in computational power have enabled an ever-increasing development and deployment of continuous hydrological models. In contrast to event-based rainfall-runoff models, continuous hydrological models consider the evolution of state variables, such as soil moisture storage and groundwater levels, to provide a long-term representation of the spatiotemporal dynamics of the involved processes. Essentially, continuous hydrological modelling is not limited to streamflow estimation, rather it allows the simulation of all key hydrometeorological processes, such as actual evapotranspiration (ET), soil moisture, and terrestrial water storage, thereby providing an estimation for the entire water balance at the basin scale, and its temporal dynamics ([Jajarmizadeh 2014](#)).

In this context, a variety of continuous hydrological models of different complexities and structures have been developed. A key representative of lumped conceptual approaches is the Sacramento model ([Burnash 1973](#)) which generates streamflow, using as input the precipitation and potential evapotranspiration (PET) data and incorporating a soil moisture accounting (SMA) scheme to estimate the water balance within the catchment. Another widely known model is the Stanford Watershed Model ([Crawford & Linsley 1966](#)) which uses a more complex conceptual structure by dividing the soil layer into four aquifers. [Efstratiadis et al. \(2008\)](#) developed the semi-distributed model HYDROGEIOS, to support decision-making in human-modified catchments. By integrating various levels of geographical information and following the hydrological response unit concept within model parameterization, HYDROGEIOS efficiently captures the key processes of the surface and groundwater hydrological cycle, also accounting for water abstractions and regulations. Another notable model is the SMA loss method that has been used to simulate water movement in various basins, including challenging semi-karst or karst regions ([Ries et al. 2015](#); [Katsanou & Lambrakis 2017](#); [Berthelin et al. 2023](#)). Simulating the water movement and groundwater recharge for an area dominated by karst formations is highly challenging as it necessitates additional data, in terms of soil moisture and properties such as thickness and texture, while discharge from karst springs may also need to be considered.

From the category of distributed-parameter models, which are more or less physically based, key representatives are the precipitation-runoff modelling system (Dawdy *et al.* 1983), the Soil Water Assessment Tool (Arnold *et al.* 1995), and the MIKE Système Hydrologique Européen (Refsgaard & Storm 1995). These models employ a detailed spatial discretization that allows the evaluation of the effect of different combinations of meteorological forcing, geomorphology, soil properties, and land use on discharge.

In the past few years, data-driven approaches have also been used for continuous streamflow prediction. These models employ machine learning algorithms using meteorological data such as forcing variables and static catchment characteristics as inputs to the model to estimate the discharge at the basin's outlet (Kratzert *et al.* 2019; Arsenault *et al.* 2023). Even more, in such cases, the absence of real-time data from meteorological stations as well as the widespread absence of accessible updated data related to soil moisture conditions render runoff forecasting at basin outlets difficult.

Although continuous simulation has been approached in many ways, it remains significantly more challenging than the standard event-based modelling. The data-intensive requirements (e.g., soil properties information, hydrometeorological conditions), make continuous hydrological simulation challenging, especially in basins with limited data availability.

A remedy to the above challenges could be provided by Earth Observation (EO) datasets (including both remote sensing and reanalysis data), which are currently becoming increasingly more available and openly accessible (Wagemann *et al.* 2021). The term 'Earth Observation data' incorporates data from satellites and *in situ* measurements, e.g., data from installed sensors, airborne platforms, and weather stations. Open EO data can facilitate the monitoring of the natural resources and provide information about many hydrological processes such as precipitation and soil water content, and related geophysical properties, such as vegetation cover and soil properties. Particularly, their wide spatial coverage, covering almost the entire globe, their high temporal resolution, and low latency highlight the significant potential of EO datasets for the development and deployment of continuous hydrological models, covering a wide range of hydrometeorological processes involved (Loumagne *et al.* 2001; Xu *et al.* 2014; Ali *et al.* 2023). In this context, EO datasets can provide information both for the meteorological forces (e.g., precipitation, temperature, and ET) and the basin's conditions (e.g., soil water content, land/vegetation cover and soil properties, and topographic relief).

Moreover, gridded EO datasets have the potential to provide detailed information on the spatial heterogeneity of variables of interest. These datasets have gained interest and are being used in various water-related fields beyond the hydrological simulation of streamflow. Araghi *et al.* (2021) compared different gridded precipitation products for simulating rainfed crops. Additionally, Saeedi *et al.* (2021) analysed different soil moisture products across various land covers, climatic conditions, and soil textures, comparing them with *in situ* measurements, while Escorihuela & Quintana-Seguí (2016) compared gridded satellite and simulated soil moisture products over the Mediterranean region. Additionally, Panahi *et al.* (2021) conducted a spatiotemporal assessment of various ET products and Tadesse *et al.* (2015) compared actual ET products and utilized them to identify growing zones in Ethiopia during the rainy season.

Recent applications of EO datasets in hydrological modelling concern both the regional and the global scale. EO data are used for the development of global hydrological models due to the spatial and temporal resolution that they offer. In these models, satellite precipitation is the most important input variable while other satellite-derived information such as water storage has also been used to force and validate global hydrological models (Güntner 2008; Xiang *et al.* 2021; Pimentel *et al.* 2023). On the regional scale, the developed models mostly concern event hydrological simulation, where EO data are used for flood mapping and flood impact assessment (Psomiadis *et al.* 2020; Schumann *et al.* 2022). In these studies, satellite data are used to delineate the flood extent, perform land use land cover classification, and evaluate potential flood impacts. Other studies have performed assimilation of satellite-derived soil moisture data in the continuous hydrological model to assess performance in terms of streamflow simulation (Baguis *et al.* 2017; Li *et al.* 2019) while satellite-based precipitation products are often used as model drivers in continuous simulations (Alazzy *et al.* 2017).

Yet, there is still ample potential in exploring the applicability of EO datasets to improve continuous hydrological simulation, particularly with respect to the spatial representation and heterogeneity of the soil hydraulic properties. The latter exerts a major control on simulated discharge particularly within the continuous simulation approach (Yang *et al.* 2018). Meanwhile, such information is usually not easily accessible, while it is costly and laborious to collect at the regional scale. EO datasets with information on soil properties present therefore an underexplored opportunity to enhance the physically based parameterization of continuous hydrological models. In this regard, this study uses spatial information on soil properties from the open soil hydraulic properties dataset, HiHydroSoils, to calibrate related model parameters within the Hydrologic Engineering Center-Hydrologic Modelling System (HEC-HMS) and compare the performance to the one

obtained using a standard literature approach. Furthermore, to further investigate the multi-faceted potential of open EO datasets in developing a continuous hydrological simulation model, satellite-based ET and precipitation series as well as satellite-derived soil moisture information, are also used in the calibration, validation, and initialization phases of the model, respectively. Finally, the combination of the above-mentioned applications is tested in a hypothetical operational scenario, to assess the model performance in near real-time operational simulation.

The case study is the Boeotikos Kephisos River basin located upstream of Lake Yliki, Greece, which is part of the water supply system of Attica, Greece, managed by the Athens Water Supply and Sewerage Company (EYDAP). The basin is characterized by karst aquifers and intensive agricultural activity, characteristics that underline the complexities of hydrological simulation in the basin and necessitate a detailed representation of soil properties at the regional scale (Efstratiadis 2008; Nalbantis *et al.* 2011).

The remainder of this work is organized as follows. Section 2 provides a brief description of the HEC-HMS model, the datasets that are employed in this study, as well as the methods that are used for the incorporation and evaluation of the EO datasets in the hydrological simulation. The results are presented in Section 3. Discussion of the methods and results is provided in Section 4, while conclusions are presented in Section 5.

2. MATERIAL AND METHODS

2.1. HEC-HMS model description and setup

In this study, we integrate EO data within the widely used HEC-HMS software (HEC, 2000) to drive tests on continuous hydrological simulation. The software includes a variety of sub-models for infiltration, rainfall-runoff transformation, baseflow estimation, hydrological routing, and modules that enable continuous simulation, such as the SMA loss method (Bennett & Peters 2004). The key elements and models involved in this work are presented in Figure 1, also indicating the components in which EO data are involved.

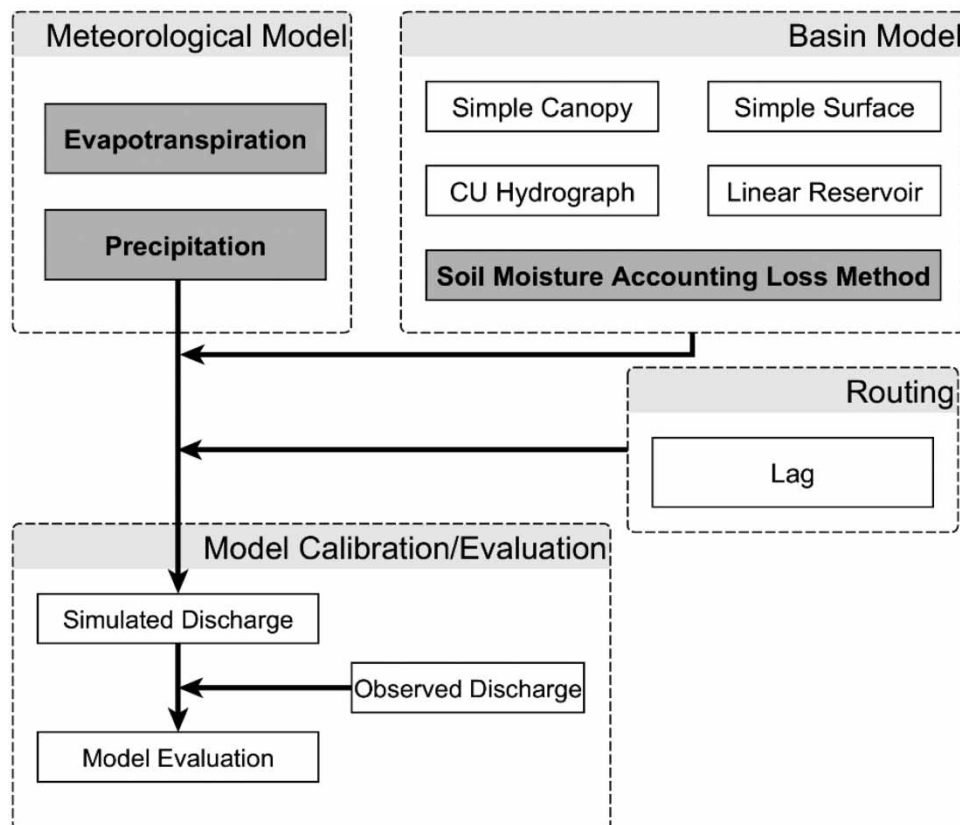


Figure 1 | General flowchart of continuous hydrologic simulation. Dark grey boxes indicate the components in which EO data are involved.

The modelling setup is conducted via two key sub-models, i.e., basin and meteorological. The latter, which is arguably simpler to parameterize, requires the definition of associated drivers, particularly, ET and precipitation. In our setup, the ‘monthly average’ method was selected to estimate losses due to ET. As detailed in the following, the average monthly PET was retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS) satellite-based dataset, while for the precipitation process, we used both ground-based measurements and satellite products (see Section 2.2).

With regards to the basin model, it includes the following sub-models: canopy storage, surface storage, a model to transform excess precipitation to direct runoff, and a model that simulates water losses in the soil and baseflow. In detail:

- Canopy storage is anticipated to vary based on the type of vegetation species and the density of vegetation cover. However, the connection between canopy storage and these factors is not well understood yet (Véliz-Chávez *et al.* 2014), and due to this, the estimation of maximum canopy storage was obtained in the calibration procedure. The crop coefficient was set to 1 and ET to ‘Wet and Dry periods’. For water uptake, we used the ‘Simple Canopy’ method, assuming that water is drawn from the soil at the PET rate.
- The estimation of surface depression storage values relies on the initial estimates provided by Bennet & Peters (2004) in combination with the slope percentage value from the Soil Survey Manual (U.S. Department of Agriculture 2017).
- To convert excess rainfall into a direct runoff, we employed the Clark unit hydrograph method (Clark 1945). This method requires the calculation of both the time of concentration and a storage coefficient, thus accounting for attenuation and diffusion processes. This approach is realistic since the lower part of the river course is characterized by almost negligible slopes.
- To simulate water loss in the soil, special focus was given to the proper identification of parameters of HEC-HMS associated with the simulation of water movement within the soil. To enable continuous simulation, we employed the SMA loss method, which divides the soil into five layers: the canopy interception layer, surface depression layer, soil profile storage layer, and two groundwater layers, termed GW1 and GW2, respectively. The method has seven parameters. Starting with maximum infiltration, it sets the upper bound on infiltration from surface storage into the soil. Next, soil percolation defines the upper bound of percolation from the soil storage layer into the upper groundwater layer. Finally, percolation rates in GW1 and GW2 set the upper bound on the percolation from the upper groundwater layer to the lower groundwater layer, and the upper bound on deep percolation, respectively. In addition to the above-mentioned parameters, the SMA method requires the specification of initial wetness conditions for both the soil and the groundwater layers. EO data were employed to support the calibration of all SMA parameters and the proper identification of initial conditions.
- Finally, the linear reservoir approach was chosen to model the baseflow recession after precipitation events. This is directly related to the loss method (the infiltration calculated from the loss method is the inflow to the linear reservoir). The GW1 and GW2 coefficients for the linear reservoir correspond to the GW1 and GW2 coefficients of the SMA loss method. The GW1 and GW2 fractions in the linear reservoir determine how the water from the loss method is divided into the groundwater layers. The attenuation during the routing in each one of the groundwater layers is accounted for by the number of reservoirs/layers, increasing with the increasing number of reservoirs.

2.2. EO datasets

EO datasets are employed to support the calibration of HEC-HMS, and particularly the SMA method, as well as to drive the model with gridded meteorological data input. Particularly, HiHydroSoils and the soil water index (SWI 1 km) datasets are used in the parameter mapping of the SMA method, by providing data on soil properties and initial soil moisture conditions, while Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS), Integrated Multi-satellite Retrievals for Global Precipitation Measurement (GPM IMERG), and the National Oceanic and Atmospheric Administration (NOAA) global forecast system (GFS) were used to drive the calibrated model with gridded precipitation estimates. Finally, ET estimates were obtained from the MODIS ET dataset. Each dataset employed is fully discussed in the following paragraphs, and a summary of the datasets used is provided in Table 1.

- HiHydroSoils v2.0 (Simons *et al.* 2020) builds upon the ISRIC SoilGrids 250 m (De Sousa *et al.* 2021), offering an enhanced, high-resolution global dataset for soil hydraulic properties. SoilGrids 250 m produces information combining soil observations from more than 200,000 locations (*in situ* data) and over 400 covariates related to vegetation, climate, and geology with machine learning models. It provides information about soil properties such as bulk density, soil organic carbon, and soil pH, in six different (standard) depths. The variables included in this dataset are used as input for deriving

Table 1 | Summary of the EO datasets used in the study

Dataset	Variable	Temporal resolution	Spatial resolution	Reference
HiHydroSoils v2.0	Soil hydraulic properties	–	250 m	Simons <i>et al.</i> (2020)
Soil water index 1 km (SWI 1 km)	Soil moisture	Daily	1 km	Bauer-Marschallinger <i>et al.</i> (2018)
MODIS ET/PET	Potential evapotranspiration	8 days	500 m	Running & Mu (2015)
CHIRPS	Precipitation	Daily	5,566 m (0.05°)	Funk <i>et al.</i> (2015)
GPM IMERG Final Run	Precipitation	Daily	11,132 m (0.1°)	Huffman <i>et al.</i> (2020)
NOAA GFS	Precipitation forecast	Daily	2,7830 m (0.25s°)	Clough <i>et al.</i> (2005)

soil hydraulic properties for the HiHydroSoils v2.0 dataset. For the conversion of the soil properties into soil hydraulic functions, pedotransfer functions were utilized, while for the calculation of the Hydrologic Soil Group, the absolute depth to bedrock and the simulated groundwater depth was used as input. The hydraulic properties in the HiHydroSoils dataset are presented in the six different depths for which ISRIC SoilGrids250 m provides the soil properties. Layer 1 extends from the surface to a depth of 5 cm, while layer 2 represents the next 10 cm (from 5 to 15 cm). Layer 3 includes the region from 15 to 30 cm and layer 4 includes the next 30 cm. Layer 5 spans from 60 to 100 cm deep, while the final (sixth) layer starts from the lower 100 cm and expands downwards to the lower 200 cm. The hydraulic soil properties included in the HiHydroSoils dataset are provided in Table A1 of the Appendix.

- The SWI at 1 km spatial resolution (SWI 1 km) product is based on the existing Copernicus Global Land Services SWI, including high-resolution Synthetic Aperture Radar (SAR) surface soil moisture (SSM) data from the 2014-launched Sentinel-1 mission (Potin *et al.* 2022), as well as SSM data from the Advanced Scatterometer (ASCAT) mission (Figa-Saldaña *et al.* 2002). The SWI algorithm was developed at Vienna University of Technology (Wagner *et al.* 2013; Paulik *et al.* 2014), based on soil moisture (surface and profile) estimates. It considers two soil layers for which the water balance equation is solved. The first layer is the one that is accessible from the C-band microwave sensors, while the second layer is lower, extending downwards from the first layer. The SWI algorithm simulates the temporal dependency of the soil moisture of the reservoir layer with the moisture conditions on the first layer.
- ET estimations were obtained from the National Aeronautics and Space Administration (NASA) MODIS ET dataset, which is an 8-day composite dataset, with a spatial resolution of 500 m. The estimations are based on the Penman–Monteith method (Pereira 1998) and reanalysis of meteorological and vegetation-related remote sensing data to calculate ET, PET, latent heat flux (LHF), and potential latent heat flux (PLHF). The algorithm used for the extraction of ET/PET considers both the surface energy and vegetation-related indices such as the leaf area index and the normalized vegetation index for the calculation of vegetation cover. The pixel values for both ET layers (ET and PET) are derived by aggregating the values from all eight days within the composite period (Running *et al.* 2017).
- Regarding precipitation, we employed three widely known satellite-based datasets, two precipitation reanalysis products (CHIRPS and GPM IMERG), used within simulations, and one forecast product (NOAA GFS), used in the context of streamflow prediction, described as follows:
 - CHIRPS is a quasi-global rainfall dataset, with a spatial resolution of 0.05°. It provides daily, pentadal, and monthly precipitation estimates from 1981 to the present. The dataset integrates data from rainfall gauges and infrared cold cloud duration observations to generate an initial product available with a 2-day latency and a final precipitation dataset accessible with an average latency of 3 weeks following the observation period (Funk *et al.* 2014).
 - GPM IMERG collects rainfall data from orbiting GPM satellites with microwave sensors (Hou *et al.* 2014). It also uses geostationary satellites with infrared sensors to fill in any missing data. There are three versions of the GPM IMERG data: Early, Late, and Final Run (Bolvin *et al.* 2020). The Final Run version enhances its rainfall estimates by combining data from ground stations in the Global Precipitation Climatology Centre network (Schneider *et al.* 2014). It has a spatial resolution of 0.1° and is available in half-hourly and daily time intervals. It becomes available three months after the actual observation date.

- The GFS (Wu *et al.* 2011) is a weather forecasting model developed by the National Centers for Environmental Prediction. The GFS integrates global models for the nexus of atmosphere, ocean, land/soil, sea, and ice, providing forecasts for various weather-related processes, such as wind, temperature, ozone concentration, and precipitation. It updates every 6 h, with new data available four times a day, and produces 384-h weather forecasts on a 28-km grid, with intervals of 1 and 3 h.

2.3. Coupling HEC–HMS model with EO soil properties data

In this work, we take advantage of the information provided by the HiHydroSoils dataset to support the calibration of the HEC–HMS model, and specifically the parameters of the SMA loss method associated with the simulation of the downward water movement in the soil and towards the outlet. This dataset provides information with respect to most of the SMA method variables, while its gridded data availability allows accounting for the spatial heterogeneity of the basin's soil hydraulic properties. To evaluate the added value derived from such an approach, we contrast it with another model setup, which uses *default* parameter estimation procedures proposed in the literature. The two approaches are detailed in the following sections.

2.3.1. SMA calibration using the HiHydroSoils dataset (EO-based calibration)

The HiHydroSoils dataset provides soil-related parameters in six standard depths. However, these depths do not necessarily represent either the total soil depth of an area (HiHydroSoils provides information up to the first 200 cm, while soils can be deeper or shallower) or the variation of the soil properties as we move to deeper layers. For example, the average soil conductivity profile from 15 to 30 cm, the third layer of the dataset, is probably not the same from 15 to 20 cm. Thus, reasonable assumptions should be made with respect to the selection of soil layers and parameters, also in accordance with the structure of the SMA model. For example, the Tension Zone layer will be considered as a part of the soil storage layer (as described in the SMA model) and the tension storage will be estimated for the first 15 cm of the soil.

To map the parameters provided by HiHydroSoils into the domain of the SMA method, we developed the methodology illustrated in Figure 2. Starting with the maximum infiltration parameter, the parameter values of each subbasin were derived from the saturated hydraulic conductivity layer within the upper 15 cm of the soil (Kirkham 2014). Soil percolation represents the speed at which water migrates to the initial groundwater layer and it was considered equal to the saturated hydraulic conductivity in the 15–100 cm range. When addressing percolation in two groundwater layers (termed GW1 and GW2), we refer to the saturated hydraulic conductivity for the deepest layer accessible in the HiHydroSoils dataset (100–200 cm). Soil storage values were estimated by multiplying the soil depth by the weighted soil porosity, and tension storage was calculated by

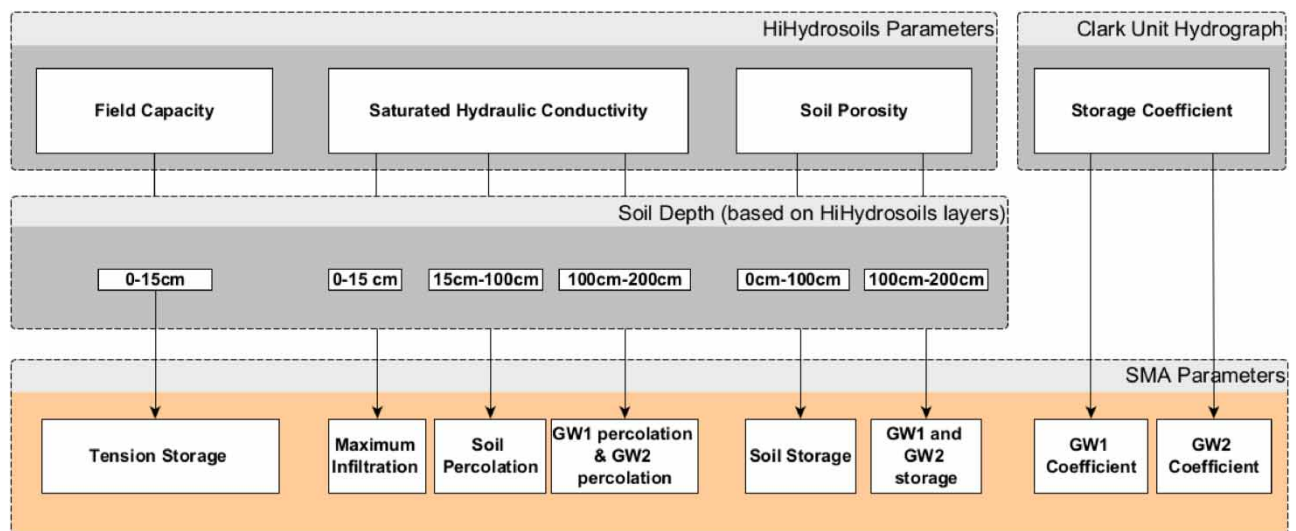


Figure 2 | Conceptualization of the use of the HiHydroSoils dataset for the SMA model calibration.

computing the weighted field capacity times the average soil depth in each subbasin. The GW1 and GW2 parameters were estimated based on the storage coefficient derived by the Clark unit hydrograph method. To introduce diversity in parameter values among subbasins in accordance with their spatial heterogeneity, the initial calculations for GW1 and GW2 storages were made in proportion to their soil storage. For the reasons mentioned above, these parameters underwent further refinement during the calibration procedure.

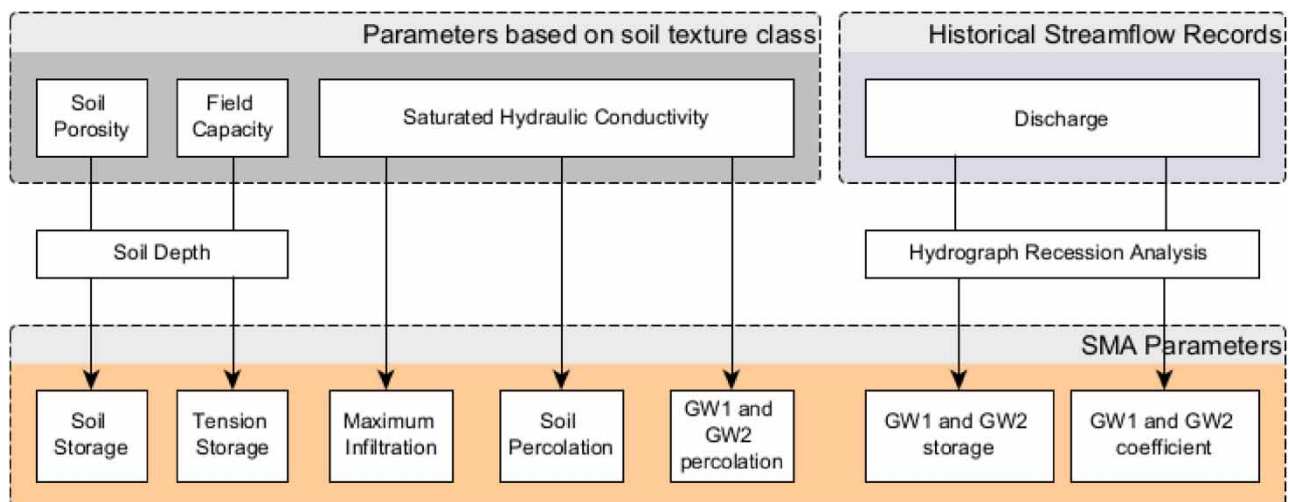
2.3.2. SMA calibration using the standard approach (literature-based calibration)

Typically, initial estimations for SMA parameters are obtained based on past studies and default values suggested in the literature. Here, we adopt such an approach to calibrate a model that serves as a benchmark scenario against the EO-based calibration. The associated methodology is visually displayed in Figure 3. Specifically, the selection of initial values for four of the parameters, namely, maximum storage depth for GW1 and GW2, and the respective storage coefficients, was guided by hydrograph recession analysis of the historical streamflow record as suggested by Fleming & Neary (2004). Six hydrograph events were selected across different seasons and years, and the recession curves for baseflow (represented by GW2) and interflow (represented by GW1) were identified under the linear reservoir assumption, which allowed estimating the corresponding recession constants. Subsequently, the storage coefficient and the maximum storage depth were derived as recommended by Fleming & Neary (2004) and Mahmood & Jia (2022) and the average values were used as representatives for the entire basin. Yet, these values were also subject to calibration. Soil porosity, field capacity, and saturated hydraulic conductivity were derived based on soil texture classes from the SoilGrids250 (Hengl *et al.* 2017), as suggested by Rawls *et al.* (1982).

Soil and tension storage values were estimated in the same way as in the first model (by multiplying the porosity and the field capacity by the soil depth). GW1 and GW2 percolation rates were obtained during the calibration process. Soil percolation and maximum infiltration values were estimated based on the average saturated hydraulic conductivity. All values were further refined during the calibration procedure. The representation of spatial heterogeneity in the second approach has largely to do with the local soil texture classes. This means that in areas that are dominated by a single soil texture class, even if the soil storage and tension storage values vary due to different depths in the area, the related values of the SMA variables representing water movement do not.

2.4. Embedding EO soil moisture information in model initialization

Initialization in hydrological modelling is of great importance as this influences the performance of the subsequent simulation steps, while inaccurate initialization can lead to unreliable results, particularly when the time horizon of the simulation is not long enough (Lepinas *et al.* 2018; Moumni *et al.* 2019). For example, in the case of a continuous hydrological model, the initial soil saturation conditions influence how soon the soil storage layers will be saturated. Often, it is preferred that



continuous simulations start at the beginning of the hydrological year, thus allowing the assigning of dry initial conditions. Under this premise, the water storage layers are set practically empty. Alternatively, a model warm-up period is considered for the model properties (e.g., soil storage) to stabilize. However, given that a warm-up period may range from one to several years, it can lead to data underuse. To avoid the assumption of dry conditions before the initialization and thus reduce the model's warm-up period, satellite soil moisture data can be incorporated to estimate the initial saturation of the soil (Laiolo *et al.* 2016).

To assess the importance of SSM in hydrological simulations, the SWI data were employed to establish the initial conditions for the SMA loss method. The SWI is a dimensionless parameter, representing the wilting point and the field capacity at 0 and 1, respectively. In our approach (using the SMA loss method to simulate the downward movement of water), the SWI effectively represents a portion of the Tension Zone storage, which is estimated using the soil's field capacity. Moreover, within the SMA method, tension storage accounts for only a fraction of the overall soil storage. As a result, the initial conditions of soil water content can be estimated as a percentage of the total soil storage. Moreover, due to the fine resolution of the SWI product ($1 \times 1 \text{ km}^2$), the spatial heterogeneity between the subbasins can also be preserved.

2.5. Study area and *in situ* data

The study area is the Boeotikos Kephisos River basin, which is located in the East Central Greece Region (Figure 4). The main attributes of the basin are given in Table A2 (Appendix Supplementary Tables and Figures for the data used in the study). In the past, this was a closed basin that discharged at Copais Lake, before the draining of the latter in 1886. Since then, an extended drainage network conveys the basin's runoff to a central channel, ending at a diversion tunnel that eventually drains into Lake Yliki, which is one of the key reservoirs providing drinking water to the metropolitan area of Athens.

The basin has a unique geological and hydrogeological structure, due to the dominance of limestone formations and, as a result, it is characterized by rich groundwater potential. The underground layers of the basin can be described as a two-layer structure, with an upper aquifer with Neogene deposits, and a larger karstic aquifer from limestone of various types. In addition, there are extensive fronts of underground losses, particularly at the southeast parts of the basin. These

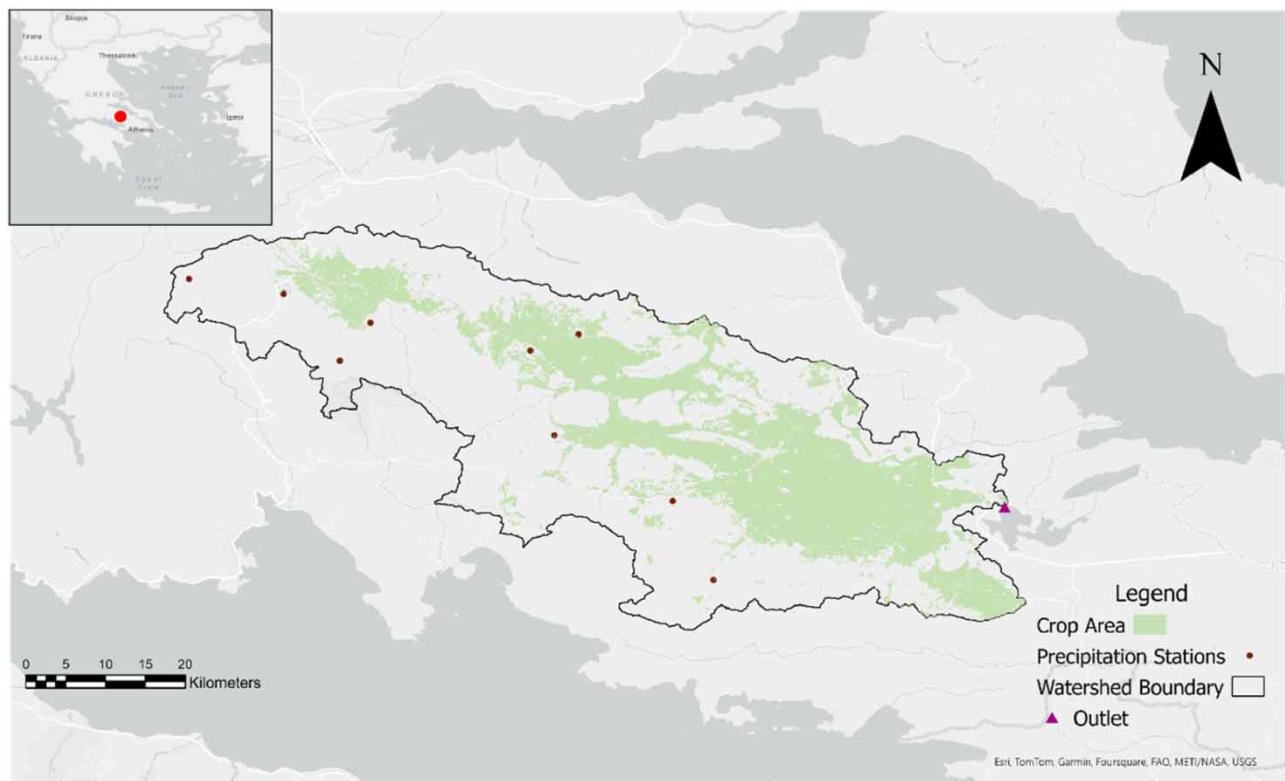


Figure 4 | Map of the basin along with the location of precipitation stations.

characteristics, along with the extensive groundwater outflows, induce discrepancies in the water balance (Efstratiadis 2008). Crops cover 33.6% of the total area, and the main crop types are winter wheat, olives, corn, and cotton. Due to intensive irrigation demands, the hydrological regime of the basin is heavily modified due to both groundwater pumping and surface abstractions. In fact, due to intensive abstractions and regulations along the channel network, the water diverted to Yliki Lake is significantly reduced or even eliminated during the summer months (from June to mid-September). This poses additional complexities and difficulties in the hydrological simulation of the basin, since irrigation withdrawals, from both surface and groundwater resources, are extended, irregular, and unmonitored, disturbing substantially the natural dynamics of the basin, especially during the summer months of the growing season (Efstratiadis 2008; Nalbantis *et al.* 2011).

The river basin is divided into eleven subbasins that are next represented through the HEC-HMS v.4.10 environment. The digital elevation model of the area was pre-processed to extract the watershed boundary, the stream network, and the subbasins. Subbasin delineation was carefully considered to account for recent changes in the hydraulic regime of the basin, particularly in the northeast parts of the basin (former lake Copais), where extended irrigation and drainage works have been constructed to serve agricultural development.

In situ measurements of daily precipitation and discharge were used both for model calibration and validation. The sole hydrometric station is located at the basin's outlet. The weather stations used in the simulation are shown in Figure 4. Daily precipitation data were available from 1 January 2005 to 31 December 2017 and from 1 January 2019 to 31 December 2021. The selection of the stations was performed considering the wider time for which consecutive precipitation measurements could be found. As a result, the spatial coverage of the river basin is somewhat limited since stations at the southeastern part of the basin had insufficient data for the period of interest.

The *in situ* precipitation records that were used for the calibration and validation of the EO-based and the literature-based models were spatially integrated, according to the Thiessen method, to provide areal estimations at each subbasin. HEC-HMS requires both spatial and temporal weights. Subbasin spatial weights were determined based on the Thiessen polygons, and temporal weights for each subbasin were determined based on the nearest gauge to this subbasin. Thus, each subbasin could have more than one spatial weight (all adding up to unity), but only one temporal weight, derived from the closest gauge.

3. RESULTS AND DISCUSSION

Following the typical split-sample approach, the simulation period is divided into calibration (1 October 2008–30 September 2017) and validation sub-sets (1 October 2019–30 September 2021). The evaluation of different models is conducted by comparing the observed and simulated discharges at the outlet of the basin, through the Nash–Sutcliffe efficiency (NSE) index (Nash & Sutcliffe 1970). The range of the NSE lies between $-\infty$ and 1, where the unit value indicates a perfect fit. An NSE lower than zero indicates that the mean of the observed time series is a better predictor than the model (Krause *et al.* 2005). Moreover, as evaluation metrics, we also use the correlation coefficient (CC), the root mean square error, and the percent bias.

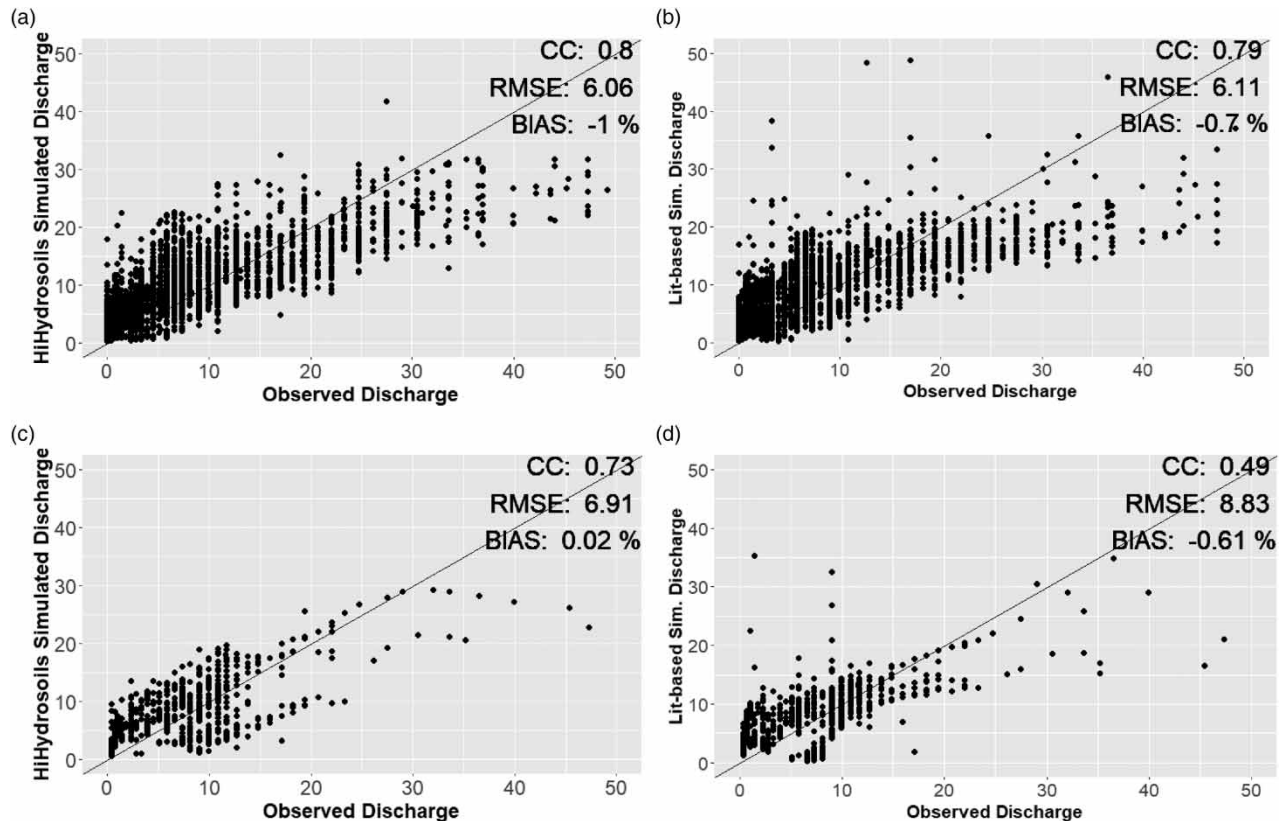
Section 3.1 presents the comparison of the literature-based and HiHydroSoils-based models' calibration using precipitation estimates from *in situ* measurements. In Section 3.2, the two models are further evaluated and compared, studying two different scenarios with respect to initial soil moisture conditions. One scenario assumes dry conditions at the beginning of the simulation period, while the second exploits EO-based soil moisture estimates to initialize the model. To further demonstrate the operational character of the latter approach, three EO products are used to drive the simulation in a hypothetical operational scenario.

3.1. Literature-based vs. HiHydroSoils-based model calibration

The results from the two models, using the HiHydroSoils dataset and the literature-based approach, respectively, for the calibration and validation period with the *in situ* precipitation, are summarized in Table 2 and Figure 5. In the calibration period, the HiHydroSoils-based model achieves a slightly better performance compared to the literature-based model (the corresponding NSE values are 0.621 and 0.615). However, the added value of exploiting information from the HiHydroSoils dataset is revealed in the validation period, where the EO-based model ensures an NSE equal to 0.569, which outperforms substantially the literature-based model. In general, as shown in Figure 5, the HiHydroSoils-based model exhibits a much better performance than the literature-based model with respect to all examined metrics.

Table 2 | Calibration and validation performance of the models

	NSE calibration	NSE validation
HiHydroSoils-based model	0.621	0.569
Literature-based model	0.615	0.217

**Figure 5** | Scatter plots of observed versus simulated discharge (m^3/s) for the calibration period with (a) HiHydroSoils dataset and (b) literature-based approach, and the validation period with (c) HiHydroSoils dataset and (d) literature-based approach.

To examine in more detail the notable discrepancy between the calibration and the validation performance for the literature-based model, a sensitivity analysis is also performed for both model versions. For the sensitivity analysis, four parameters of the SMA loss method are varied within $\pm 10\%$ of the calibrated values, including ± 1 , ± 3 , ± 5 , ± 8 , and $\pm 10\%$: the soil and tension storages, the maximum infiltration, and soil percolation, as they are found to be critical for the simulation of the soil saturation conditions (Bhuiyan *et al.* 2017). The analysis was performed for both models (HiHydroSoils-based and literature-based models) and the results are statistically summarized and presented in Table 3.

The sensitivity analysis suggests that the HiHydroSoils-based model consistently outperforms the literature-based model both in calibration and validation, with the latter performing poorly during validation, and it does not generalize well beyond the calibration data. More specifically, for the validation phase, the performance of the literature-based model is significantly poorer with the median and mean NSE values around 0.216–0.220 indicating low variability but consistently poor performance.

A direct comparison of the observed and simulated discharge time series from the two models for the entire period is given in Figure 6. The figure reveals a satisfactory match of the observed to simulated discharge for the HiHydroSoils-based model for the long-term runoff dynamics. Yet, none of the calibration approaches achieves a good performance with respect to accurately replicating the major discharge events. This could be in part attributed to a potential underestimation of areal precipitation, as derived from a relatively small number of stations that do not equally cover the entire area, but it is also

Table 3 | Summary statistics (median, mean, and standard deviation) of the models' NSE from the sensitivity analysis varying each parameter by $\pm 10\%$

	HiHydroSoils-based model				Literature-based model			
	Median NSE	Mean NSE	Min NSE	Max NSE	Median NSE	Mean NSE	Min NSE	Max NSE
Calibration								
Soil Storage	0.621	0.62	0.615	0.623	0.611	0.61	0.603	0.615
Tension Storage	0.621	0.62	0.614	0.623	0.603	0.604	0.603	0.615
Maximum Infiltration	0.619	0.619	0.616	0.621	0.615	0.612	0.605	0.618
Soil Percolation	0.621	0.621	0.617	0.623	0.615	0.614	0.604	0.617
Validation								
Soil Storage	0.517	0.516	0.498	0.528	0.219	0.22	0.211	0.220
Tension Storage	0.518	0.516	0.509	0.520	0.217	0.216	0.209	0.217
Maximum Infiltration	0.520	0.520	0.520	0.520	0.216	0.216	0.201	0.221
Soil Percolation	0.517	0.516	0.495	0.532	0.217	0.218	0.186	0.237

expected to be related to the complex response of the basin during significant storm events. The discharge recession, conversely, is adequately represented.

3.2. Model initialization with satellite soil moisture information

To assess the effect of the models' initial conditions, we re-run simulations from 1 October 2019 to 30 September 2021 (previously used as the validation period) with all four of the precipitation datasets as model drivers (both *in situ* and satellite-based), and now compare between a dry-state initialization (assuming zero soil wetness at the start of this simulation period) and an initialization using soil saturation information from the SWI dataset, available for the same starting period. The initial values for the SWI initialization are available in Table A3. Regarding the model results for various precipitation datasets, those for NOAA GFS are presented separately from the results of other datasets as they refer to precipitation forecasts rather than observed precipitation.

A comparison of the model's performance between a dry-state initialization and an initialization with the soil saturation information (the SWI initialization) for both (observed) precipitation datasets (CHIRPS and IMERG Final Run) and the *in situ* precipitation is given in Table 4 in terms of the NSE, while related discharge simulations are compared in Figure 7. This comparison was carried out for both model versions (HiHydroSoils and literature-based).

It is observed that the NSE for the SWI initialization is raised regardless of the model's precipitation driver (i.e., precipitation dataset). The best improvement is observed for IMERG-driven simulation (42% increase). In the case of the literature-based model, the NSE for the IMERG-driven simulation for the SWI initialization is less than the NSE for the simulation assuming dry initial conditions. Yet it should be noted that none of the combinations of the SWI initialization with precipitation data results in an acceptable NSE performance and therefore, there is limited potential for replacing the models' 'warm-up' period with a SWI initialization.

NOAA 24-precipitation accumulation performed better for both models when soil saturation information was incorporated. In the case of the HiHydroSoils model, the NSE reached almost 0.5, improving the model's performance compared to dry initialization by 20%. For the literature-based model, NOAA GFS performed well too, reaching 0.57 for the dry-state initialization and 0.6 for the SWI initialization. The results for dry-state and the SWI initialization for the NOAA GFS are provided in Table 5 and Figure 8. However, this increase in performance seems to be more attributable to the better performance of the NOAA GFS dataset than to the improvement due to the soil moisture initialization.

Yet, these initialization experiments, compared to the model performance in the validation period (Section 3.1) when run in continuous time, point to the fact that a 'warm-up' period is essential when applying the HEC-HMS SMA scheme for a baseflow-dominated basin and cannot be entirely replaced by the SWI initialization.

3.3. Operational demonstration

Finally, to assess the capacity of the EO data to be used in real time, an operational demonstration was conducted. More specifically, we estimated the flood runoff produced by an intense storm event named 'Ballos' that hit the study area on

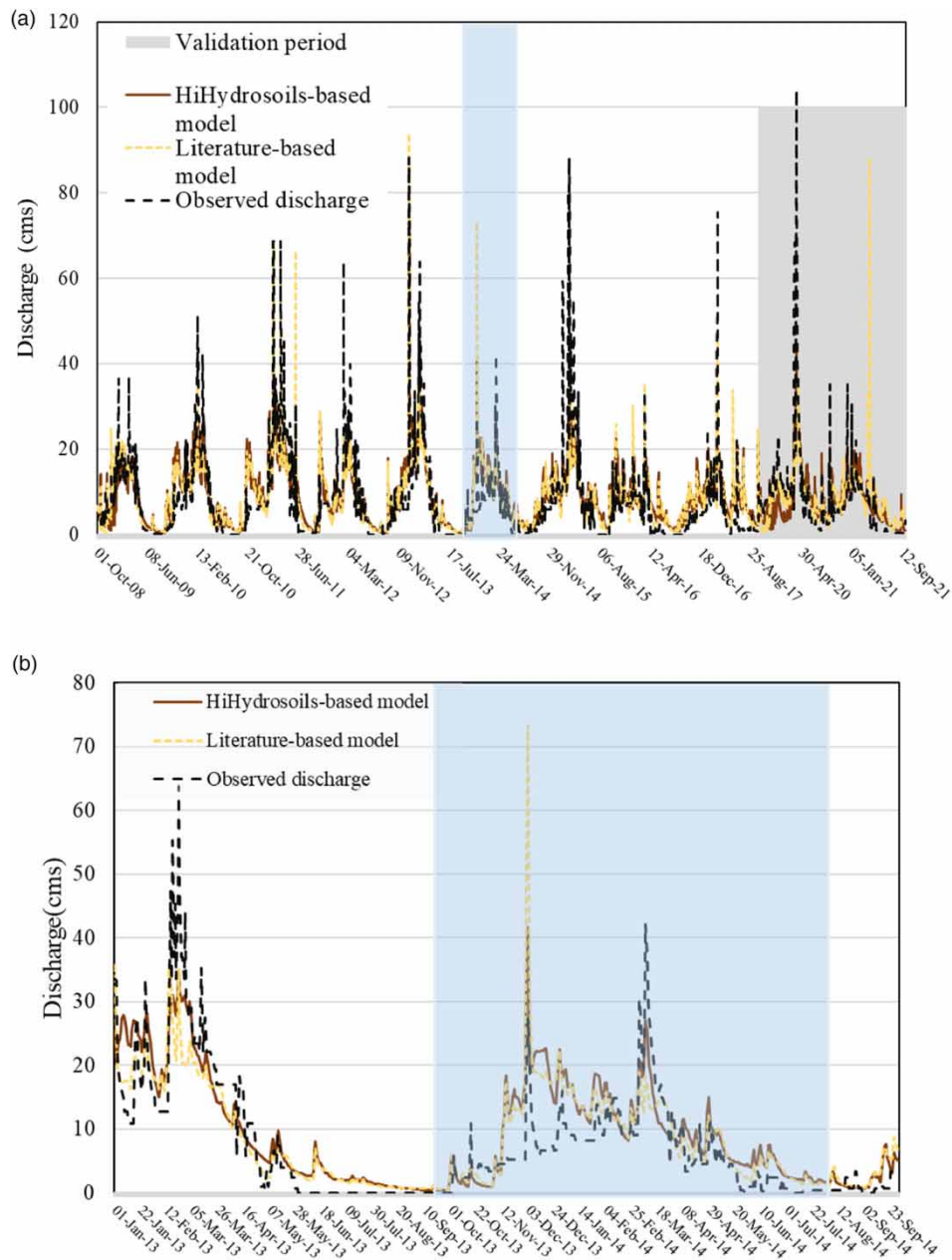


Figure 6 | (a) Discharge time series for the calibration and validation periods (grey shaded). Observed discharge (black), simulated with the EO-based model (blue) and the literature-based model (orange). (b) A closer look at the simulated (EO- and literature-based) and the observed discharge from January 2013 to September 2014.

14 October 2021. For the practical demonstration, which involves utilizing SSM and precipitation forecast products, a hydrological simulation was conducted, mimicking real-time conditions. We set the starting date of the simulation to 12 October 2021, and then extended the simulation for an additional day following the occurrence of the Ballos event (15 October 2021). This simulation accounted for the HiHydroSoils model, soil saturation initial conditions (using the SWI), and the 24-h precipitation accumulation forecast (NOAA), which is appropriate for operational use. The initial soil saturation values for each of the subbasins are given in Table A4 (see Appendix). The discharge results are compared with the observed and the discharge results of the literature-based model (for the same days and with the same drivers).

Table 4 | NSE results with the SWI initialization for the validation period for IMERG Final Run, CHIRPS, and *in situ* precipitation vs. the dry-state initialization

	HiHydroSoils model			Literature-based model		
	NSE Dry-state initialization	NSE SWI initialization	Percent change	NSE dry-state initialization	NSE SWI initialization	Percent change
IMERG precipitation	0.148	0.255	41.96%	0.256	0.226	11.72%
CHIRPS precipitation	0.194	0.28	30.71%	0.218	0.266	22.02%
<i>in situ</i> precipitation	0.569	0.584	10.96%	0.217	0.244	12.44%

As shown in Figure 9, at the operational simulation, both models were able to estimate an increased runoff volume at the basin's outlet. Since the time of concentration of the basin is approximately one day, we used the data from 15 October 2021, one day after the storm Ballos hit, to assess the model's performance.

Starting with the HiHydroSoils model, the simulated discharge from 11th October to 14th October closely approaches the observed discharge. On 14 October though, the simulated discharge is larger than the observed and continues to rise the following day, although with a milder slope, while at the same time the observed discharge declines. This can be attributed to precipitation volume errors in the NOAA forecast or to flaws in the calibrated model in the first place, while the possibility that there may be an error in the observed discharge itself cannot be excluded. In this respect, further information is sought from Yliki Lake, for which daily storage observations are publicly available on the operators' website (EYDAP). The storage difference between the 14th and 15th of October 2021 is 636,000 m³. Based on the observed and the simulated flows, the average discharge value for 14th October was 5.205 and 6.995 m³/s, respectively. This is equal to 449.712 m³ for the observed flow and 604.368 m³ for the simulated. The difference in the observed storage volume, however, does not precisely reflect the total volume of water that entered the lake, as the percentage of the incoming volume is not calculated, primarily due to sub-surface losses. Consequently, the difference in water volume recorded by the operator results in an underestimation of the total water volume that entered the lake. Nevertheless, the simulated discharge reaches the actual difference in the lake's water balance more accurately than the observed discharge. For the literature-based model, however, the volume error is significantly higher, with the model overestimating the observed discharge.

4. DISCUSSION

The goal of this work was to test the performance of EO and satellite datasets at various stages of hydrological modelling, when utilizing the SMA scheme of the HEC-HMS model. Satellite and EO data have the potential to support hydrological simulations, especially in areas with limited data availability, yet many limitations still exist. The model key drivers, particularly satellite precipitation, are not yet accurate enough to be used confidently in complex hydrological simulations as the NSE results show, but they can serve as a valuable data source not only when precipitation data are not available but also when forecasting precipitation.

Also, in terms of the methodology, a specific modelling approach may not be universally applicable to every case study, regardless of its sophistication. Regions with different properties may require different approaches, related to data availability and the characteristics of the study area. Equifinality, describing the fact that multiple model and parameter configurations can be considered acceptable for simulating real-world problems, also arises in the SMA model calibration. Here, to ensure a fair evaluation of the value of the HiHydroSoils dataset for model calibration, the authors limited the parameter adjustments to only those involved in the SMA, keeping the rest of the parameters the same between the two models. Calibrating two completely different models would not provide an objective evaluation of the HiHydroSoil dataset's performance, which in our case proved superior to the utilization of a literature-based parameterization.

The model initialization with soil saturation information from the SWI dataset improves the model's performance regardless of the precipitation input. This is due to the model's warm-up period being reduced as the soil storage layer is no longer considered to be completely empty at the beginning of the simulation, thus the appropriate soil saturation levels are reached sooner. However, the improvement was still not sufficient for the SWI initialization to be considered a viable alternative to

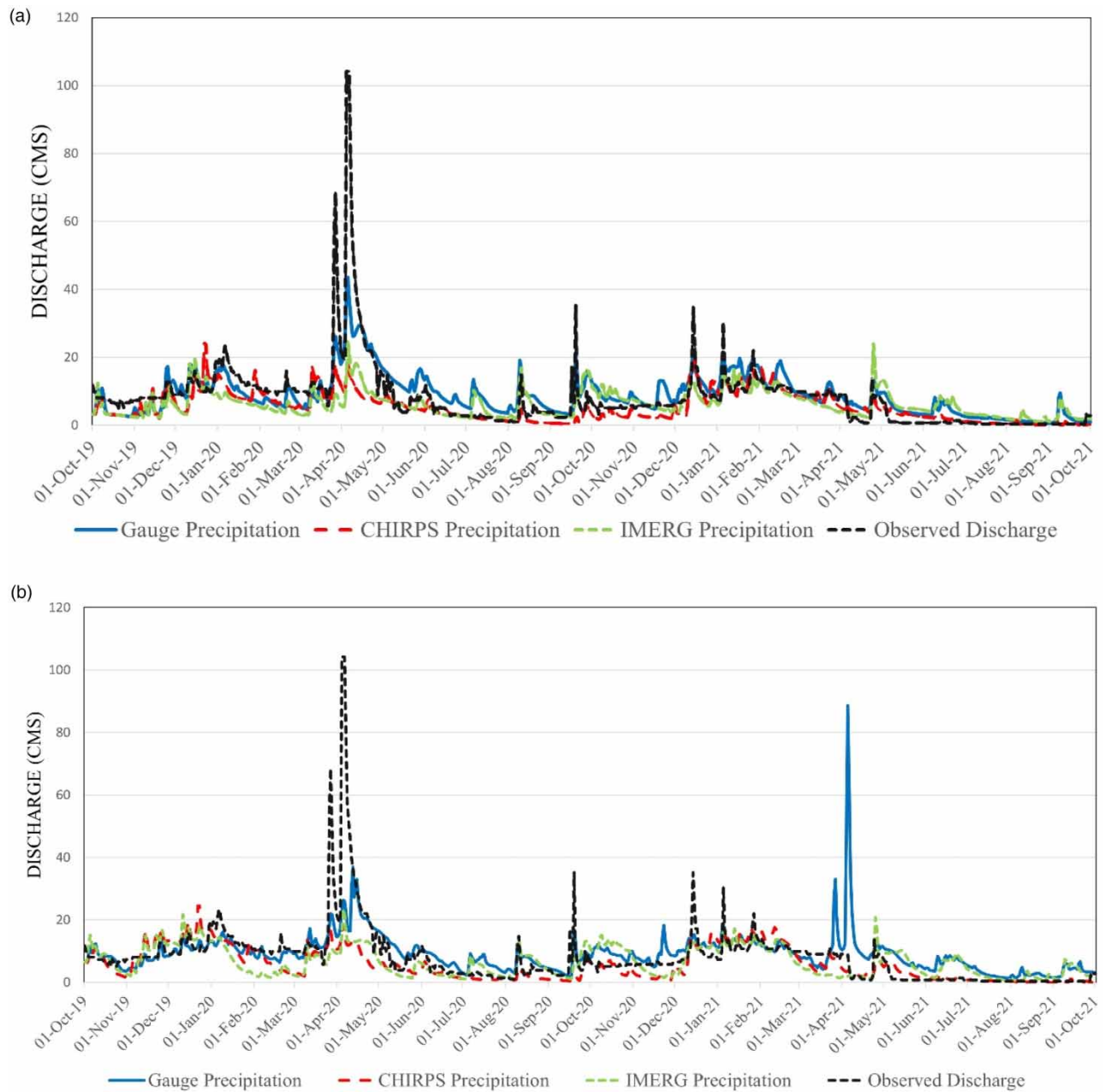


Figure 7 | Discharge results with satellite precipitation for the (a) HiHydroSoils and (b) the literature-based model with the SWI initialization for gauge (blue), CHIRPS (red), and IMERG Final Run (green) datasets.

Table 5 | NSE results with the SWI initialization for the validation period for NOAA 24-h accumulation forecast precipitation vs. the dry-state initialization

	HiHydroSoils model			Literature-based model		
	NSE dry-state initialization	NSE SWI initialization	Percent change	NSE dry-state initialization	NSE SWI initialization	Percent change
NOAA GFS	0.389	0.491	20.77%	0.567	0.599	5.64%

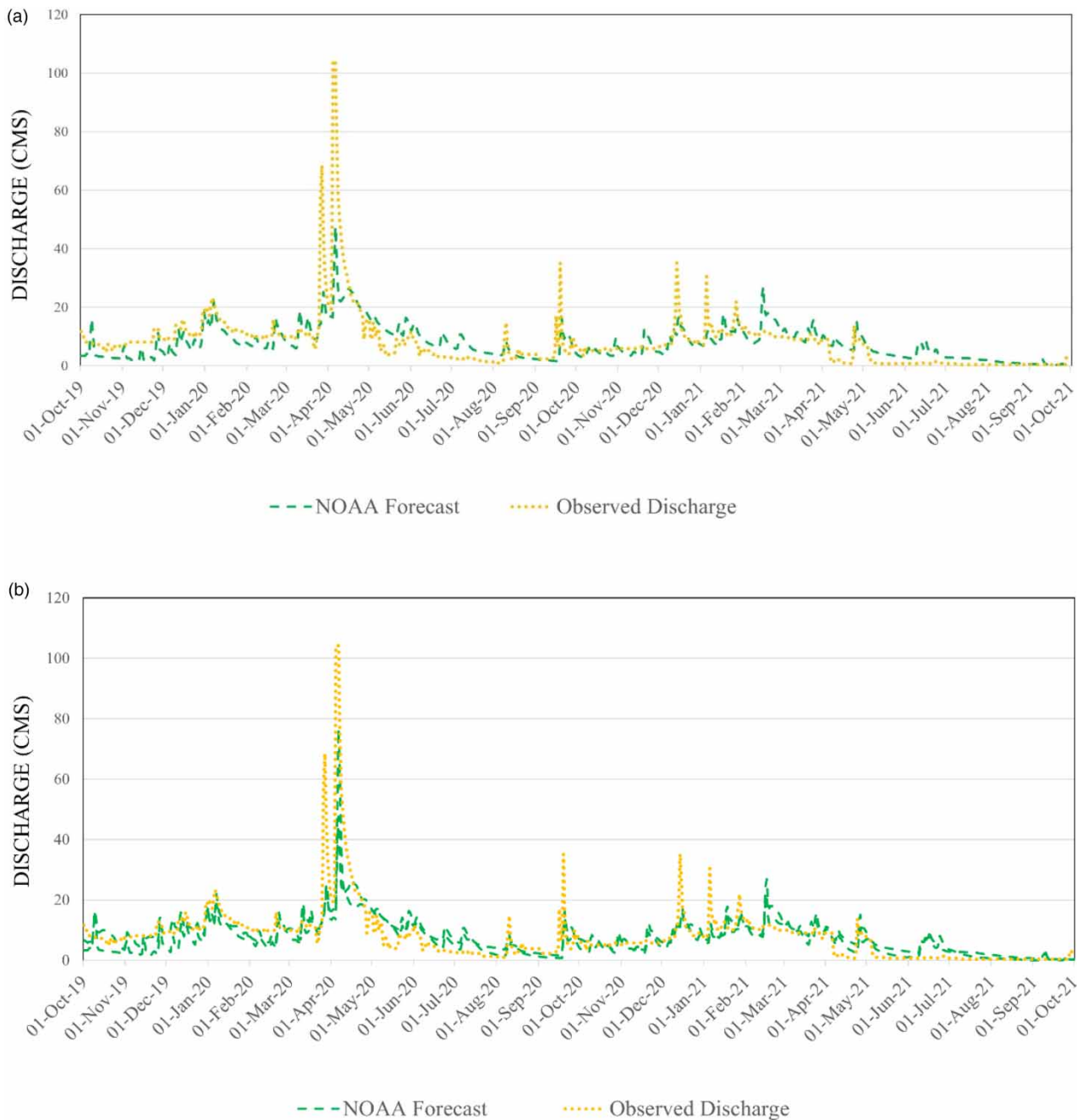


Figure 8 | Discharge results with satellite precipitation for the (a) HiHydroSoils and (b) the literature-based model with the SWI initialization for NOAA GFS dataset.

using a warm-up period for model setup. This might be due to the complex groundwater response of the studied basin, which is represented by a sophisticated parameterization and related internal state variables of the SMA module, which cannot be efficiently captured by the SWI information.

Although the HiHydroSoils-based model performance is satisfactory when driven with observed precipitation data, it is limited in accurately simulating the extreme streamflow events. This can be attributed to limitations in the models that may be linked to the challenges of modelling high peak flows in karstic areas such as the study area. The interaction between rapid interflow, subsurface storage, and the unpredictable response of karst systems to rainfall makes it difficult to capture the

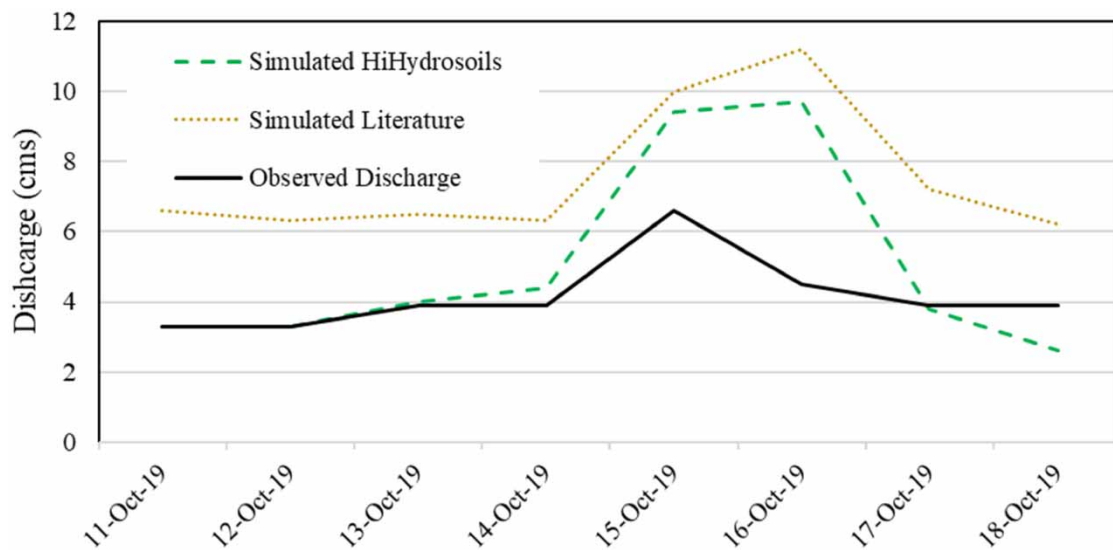


Figure 9 | Observed and simulated discharge results for the HiHydroSoils and the literature-based model.

timing and magnitude of major peak events. However, the underground formations of the study area may not be the only reason for the models' difficulty in predicting high flows. Instead, limitations in the models themselves, the data sources, or the calibration process could be considered as potential reasons for this shortcoming.

5. CONCLUSIONS

In this study, we aimed to further explore the value of integrating EO and satellite datasets in different stages of a continuous hydrological model simulation using an SMA scheme. Starting with model calibration and validation, we compared two distinct model setup approaches: one using the HiHydroSoils dataset (EO data set with soil properties information) and the second using parameters derived from standard literature approaches and prior studies. The first approach outperformed the second one in a challenging basin with an extensive karst aquifer system. The utilization of gridded datasets is particularly advantageous for simulating the spatial heterogeneity that characterizes river basins of complex topography and large extent. This stands out as the key benefit of EO datasets compared to the literature-based approach. Specifically, HiHydroSoils can be used in challenging hydrological simulations to provide a baseline estimation for the model parameters, provided that the basin dynamics and characteristics are understood.

Satellite-based precipitation data are not found to be accurate enough to replace *in situ* precipitation observations in hydrological simulation models, although the performance may be improved via adequate data assimilation or downscaling methods. Including satellite-derived soil moisture information prior to initiating the hydrological simulation can also shorten the warm-up period for model simulation, yet it cannot entirely replace the necessity of the latter, particularly for a groundwater-dominated basin. This opportunity, however, may be especially helpful in situations where extended records of observed discharge data are not available. Finally, regarding the operational potential for exploiting EO/satellite datasets, it is important to note that our preliminary results are promising. Exploring additional integration possibilities, including alternative model approaches and different datasets, may offer more insights.

The most promising direction, however, identified in this work is the potential for developing a sophisticated hydrological model for continuous hydrologic simulation, using the HEC-HMS data-intensive SMA scheme built solely based on EO data. This allows overcoming limitations regarding the usual lack of detailed *in situ* data for the estimation of the soil hydraulic properties, the acquisition of which is costly and time consuming. In data-scarce regions or areas characterized by rich groundwater potential and diverse soil formations, incorporating soil properties from EO data, such as the HiHydroSoils dataset, may provide an effective baseline for the calibration of parameters required for continuous hydrologic models.

ACKNOWLEDGEMENTS

This work was supported by the European Union under ToDrinQ Project (Project code 101082035).

DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Alazzy, A. A., Lü, H., Chen, R., Ali, A. B., Zhu, Y. & Su, J. (2017) Evaluation of satellite precipitation products and their potential influence on hydrological modeling over the Ganzi river basin of the Tibetan Plateau, *Advances in Meteorology*, **2017**, 1–23. <https://doi.org/10.1155/2017/3695285>.
- Ali, M. H., Popescu, I., Jonoski, A. & Solomatine, D. P. (2023) Remote sensed and/or global datasets for distributed hydrological modelling: A review, *Remote Sensing*, **15**, 1642. <https://doi.org/10.3390/RS15061642>.
- Araghi, A., Jaghargh, M. R., Maghrebi, M., Martinez, C. J., Fraisse, C. W., Olesen, J. E. & Hoogenboom, G. (2021) Investigation of satellite-related precipitation products for modeling of rainfed wheat production systems, *Agricultural Water Management*, **258**, 107222. <https://doi.org/10.1016/J.AGWAT.2021.107222>.
- Arnold, J. G., Williams, J. R. & Maidment, D. R. (1995) Continuous-time water and sediment-routing model for large basins, *Journal of Hydraulic Engineering*, **121**, 171–183. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1995\)121:2\(171\)](https://doi.org/10.1061/(ASCE)0733-9429(1995)121:2(171)).
- Arsenault, R., Martel, J. L., Brunet, F., Brissette, F. & Mai, J. (2023) Continuous streamflow prediction in ungauged basins: Long short-term memory neural networks clearly outperform traditional hydrological models, *Hydrology and Earth System Sciences*, **27**, 139–157. <https://doi.org/10.5194/HESS-27-139-2023>.
- Baguis, P., Roulin, E., Wagner, W., Notarnicola, C., Qu, J. J., Ryu, D., Baghdadi, N. & Thenkabail, P. S. (2017) Soil moisture data assimilation in a hydrological model: A case study in Belgium using large-scale satellite data, *Remote Sensing*, **9**, 820. <https://doi.org/10.3390/RS9080820>.
- Bauer-Marschallinger, B., Paulik, C., Hochstöger, S., Mistelbauer, T., Modanesi, S., Ciabatta, L., Massari, C., Brocca, L. & Wagner, W. (2018) Soil moisture from fusion of scatterometer and SAR: Closing the scale gap with temporal filtering, *Remote Sensing*, **10**, 1030. <https://doi.org/10.3390/RS10071030>.
- Bennett, T. H. & Peters, J. C. (2004) Continuous soil moisture accounting in the Hydrologic Engineering Center Hydrologic Modeling System (HEC–HMS). Joint Conference on Water Resource Engineering and Water Resources Planning and Management 2000. In: Hotchkiss, R. H. & Glade, M. (eds.) *Building Partnerships*, Minneapolis, MN: American Society of Civil Engineers, p. 104. [https://doi.org/10.1061/40517\(2000\)149](https://doi.org/10.1061/40517(2000)149).
- Berthelin, R., Olarinoye, T., Rinderer, M., Mudarra, M., Demand, D., Scheller, M. & Hartmann, A. (2023) Estimating karst groundwater recharge from soil moisture observations – A new method tested at the Swabian Alb, southwest Germany, *Hydrology and Earth System Sciences*, **27**, 385–400. <https://doi.org/10.5194/hess-27-385-2023>.
- Beven, K. J. (2005) Rainfall-runoff modeling: Introduction, *Encyclopedia of Hydrological Sciences*, **11**, 1857–1868. <https://doi.org/10.1002/0470848944.HSA130>.
- Bhuiyan, H. A. K. M., McNairn, H., Powers, J. & Merzouki, A. (2017) Application of HEC–HMS in a cold region watershed and use of RADARSAT-2 soil moisture in initializing the model, *Hydrology*, **4**, 9. <https://doi.org/10.3390/HYDROLOGY4010009>.
- Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., Sorooshian, S., Tan, J. & Xie, P. (2020) NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). In: Levizziani, V., Kidd, C., Kirschbaum, D. B., Kummerow, C. D. & Turk, J. (eds.), *Satellite Precipitation Measurement: Volume 1*, pp. 343–353. Cham: Springer. 10.1007/978-3-030-24568-9_19/FIGURES/3.
- Brocca, L., Melone, F. & Moramarco, T. (2011) Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting, *Hydrological Process*, **25**, 2801–2813. <https://doi.org/10.1002/HYP.8042>.
- Burnash, R. J. C. (1973) *A Generalized Streamflow Simulation System: Conceptual Modeling for Digital Computers*. Sacramento, California: US Dept. of Commerce, National Weather Service.
- Clark, C. O. (1945) Storage and the unit hydrograph, *Transactions of the American Society of Civil Engineers*, **110**, 1419–1446. <https://doi.org/10.1061/TACEAT.0005800>.
- Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., Boukabara, S. & Brown, P. D. (2005) Atmospheric radiative transfer modeling: A summary of the AER codes, *Journal of Quantitative Spectroscopy and Radiative Transfer*, **91**, 233–244. <https://doi.org/10.1016/J.JQSRT.2004.05.058>.
- Crawford, N. H. & Linsley, R. K. (1966) *Digital simulation in hydrology: Stanford watershed model IV* (Technical Report No. 39). Stanford, CA: Department of Civil Engineering, Stanford University.
- Dawdy, D. R., Lichty, R. W., Bergmann, J. M. & U.S. Geological Survey (1983) *Precipitation-runoff modeling system: User's manual*. U.S. Geological Survey Water-Resources Investigations Report 83-4238. Reston, VA: USGS. <https://doi.org/10.3133/wri834238>.
- De Sousa, L. M., Poggio, L., Batjes, N. H., Heuvelink, G. B. M., Kempen, B., Ribeiro, E. & Rossiter, D. (2021) SoilGrids 2.0: producing quality-assessed soil information for the globe, *Soil*, **7** (1), 217–240. <https://doi.org/10.5194/soil-7-217-2021>.

- Efstratiadis, A. (2008) Non-linear methods in multiobjective water resource optimization problems, with emphasis on the calibration of hydrological models. PhD Thesis. Athens: National Technical University of Athens.
- Efstratiadis, A., Nalbantis, I., Koukouvinos, A., Rozos, E. & Koutsoyiannis, D. (2008) **HYDROGEIOS: A semi-distributed GIS-based hydrological model for modified river basins**, *Hydrology and Earth System Sciences*, **12**, 989–1006. <https://doi.org/10.5194/HESS-12-989-2008>.
- Elga, S., Jan, B. & Okke, B. (2015) **Hydrological modelling of urbanized catchments: A review and future directions**, *Journal of Hydrology*, **529** (P1), 62–81. <https://doi.org/10.1016/j.jhydrol.2015.06.028>.
- Escorihuela, M. J. & Quintana-Seguí, P. (2016) **Comparison of remote sensing and simulated soil moisture datasets in Mediterranean landscapes**, *Remote Sensing of Environment*, **180**, 99–114. <https://doi.org/10.1016/J.RSE.2016.02.046>.
- EYDAP (n.d.) Raw Water Reserves [WWW Document]. Available at: <https://www.eydap.gr/en/TheCompany/Water/Savings>.
- Figa-Saldaña, J., Wilson, J. J. W., Attema, E., Gelsthorpe, R., Drinkwater, M. R. & Stoffelen, A. (2002) **The advanced scatterometer (ASCAT) on the meteorological operational (MetOp) platform: A follow on for European wind scatterometers**. *Canadian Journal of Remote Sensing*, **28** (3), 404–412. <https://doi.org/10.5589/M02-035>.
- Fleming, M. & Neary, V. (2004) **Continuous hydrologic modeling study with the hydrologic modeling system**, *Journal of Hydrologic Engineering*, **9**, 175–183. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2004\)9:3\(175\)](https://doi.org/10.1061/(ASCE)1084-0699(2004)9:3(175)).
- Funk, C. C., Peterson, P. J., Landsfeld, M. F., Pedreros, D. H., Verdin, J. P., Rowland, J. D., Romero, B. E., Husak, G. J., Michaelsen, J. C. & Verdin, A. P. (2014) **A quasi-global precipitation time series for drought monitoring: U.S. Geological Survey Data Series 832**. Reston, VA: USGS. <https://dx.doi.org/10.3133/ds832>.
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A. & Michaelsen, J. (2015) **The climate hazards infrared precipitation with stations – A new environmental record for monitoring extremes**, *Scientific Data*, **2**, 1–21. <https://doi.org/10.1038/sdata.2015.66>.
- Güntner, A. (2008) **Improvement of global hydrological models using GRACE data**, *Surveys in Geophysics*, **29**, 375–397. <https://doi.org/10.1007/S10712-008-9038-Y/TABLES/1>.
- Hengl, T., De Jesus, J. M., Heuvelink, G. B. M., Gonzalez, M. R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., Guevara, M. A., Vargas, R., MacMillan, R. A., Batjes, N. H., Leenaars, J. G. B., Ribeiro, E., Wheeler, I., Mantel, S. & Kempen, B. (2017) **Soilgrids250m: Global gridded soil information based on machine learning**, *PLoS One*, **12**. <https://doi.org/10.1371/JOURNAL.PONE.0169748>.
- Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K. & Iguchi, T. (2014) **The global precipitation measurement mission**, *Bulletin of the American Meteorological Society*, **95**, 701–722. <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K. L., Joyce, R. J., Kidd, C., Nelkin, E. J., Sorooshian, S., Stocker, E. F., Tan, J., Wolff, D. B. & Xie, P. (2020) **Integrated multi-satellite retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG)**, *Advances in Global Change Research*, **67**, 343–353. https://doi.org/10.1007/978-3-030-24568-9_19/FIGURES/3.
- Hydrologic Modeling System (2000) *HEC-HMS Hydrologic Modeling System*, Technical Reference Manual, CPD-74B. Davis, CA: Hydrologic Engineering Center.
- Jajarmizadeh, M. (2014) **A review on theoretical consideration and types of models in hydrology**. *Journal of Environmental Science and Technology*, **5** (5), 249–261. <https://doi.org/10.3923/jest.2012.249.261>.
- Katsanou, K. & Lambrakis, N. (2017) **Modelleren van de helleense karststroomgebieden met het sacramento soil moisture accounting model**, *Hydrogeology Journal*, **25**, 757–769. <https://doi.org/10.1007/s10040-016-1520-x>.
- Kirkham, M. B. (2014) Infiltration. In: Kirkham, M. B. (ed.), *Principles of soil and plant water relations*, pp. 201–227. London: Academic Press. <https://doi.org/10.1016/C2013-0-12871-1>.
- Kratzert, F., Klotz, D., Herrnegger, M., Sampson, A. K., Hochreiter, S. & Nearing, G. S. (2019) **Toward improved predictions in ungauged basins: Exploiting the power of machine learning**, *Water Resources Research*, **55**, 11344–11354. <https://doi.org/10.1029/2019WR026065>.
- Krause, P., Boyle, D. P. & Bäse, F. (2005) **Comparison of different efficiency criteria for hydrological model assessment**, *Advances in Geosciences*, **5**, 89–97.
- Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Rudari, R., Pulvirenti, L., Boni, G., Fascetti, F., Pierdicca, N., Crapolicchio, R., Hasenauer, S. & Puca, S. (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*, **48**, 131–145. <https://doi.org/10.1016/J.JAG.2015.06.002>.
- Lepinas, F., Dastoor, A. & Fortin, V. (2018) **Performance of the dynamically dimensioned search algorithm: Influence of parameter initialization strategy when calibrating a physically based hydrological model**, *Hydrology Research*, **49**, 971–988. <https://doi.org/10.2166/NH.2017.139>.
- Li, D., Liang, Z., Li, B., Lei, X. & Zhou, Y. (2019) **Multi-objective calibration of MIKE SHE with SMAP soil moisture datasets**, *Hydrology Research*, **50**, 644–654. <https://doi.org/10.2166/NH.2018.110>.
- Loumagne, C., Normand, M., Riffard, M., Weisse, A., Quesney, A., Hégarat-Masclé, S. L. & Alem, F. (2001) **Integration of remote sensing data into hydrological models for reservoir management**, *Hydrological Sciences Journal*, **46**, 89–102. <https://doi.org/10.1080/02626660109492802>.

- Mahmood, R. & Jia, S. (2022) A comprehensive approach to develop a hydrological model for the simulation of all the important hydrological components: The case of the three-River headwater region, China. *Water (Switzerland)*, **14**, 2778. <https://doi.org/10.3390/W14182778/S1>.
- Moumni, H., Laila, S., Sebari, K., Stour, L. & Ahbari, A. (2019) Estimation des valeurs initiales des paramètres du modèle HEC–HMS. application au bassin versant de Óouergha (Sebou, Maroc) Abdellatif Ahbari Agence du Bassin Hydraulique de Draa Oued Noun Estimation of the HEC–HMS model parameters in data-scarce regions. Application to the Ouergha watershed (Sebou, Morocco). *Journal of Materials and Environmental Science*, **9** (1), 305–317. <https://doi.org/10.26872/JMES.2018.9.1.34>.
- Nalbantis, I., Efstratiadis, A., Rozos, E., Kopsiafti, M. & Koutsoyiannis, D. (2011) Holistic versus monomeric strategies for hydrological modelling of human-modified hydrosystems, *Hydrology and Earth System Sciences*, **15**, 743–758. <https://doi.org/10.5194/hess-15-743-2011>.
- Nash, J. E. & Sutcliffe, J. V. (1970) River flow forecasting through conceptual models part I – A discussion of principles, *Journal of Hydrology (Amst)*, **10**, 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6).
- Panahi, D. M., Tabas, S. S., Kalantari, Z., Ferreira, C. S. S. & Zahabiyoun, B. (2021) Spatio-temporal assessment of global gridded evapotranspiration datasets across Iran, *Remote Sensing(Basel)*, **13**, 1816. <https://doi.org/10.3390/RS13091816/S1>.
- Paulik, C., Dorigo, W., Wagner, W. & Kidd, R. (2014) Validation of the ASCAT soil water index using in situ data from the international soil moisture network, *International Journal of Applied Earth Observation and Geoinformation*, **30**, 1–8. <https://doi.org/10.1016/J.JAG.2014.01.007>.
- Pereira, L. S. (1998) *Crop evapotranspiration: Guidelines for computing crop water requirements*. FAO Irrigation and Drainage Paper No. 56. Rome: FAO.
- Pimentel, R., Crochemore, L., Andersson, J. C. M. & Arheimer, B. (2023) Assessing robustness in global hydrological predictions by comparing modelling and earth observations, *Hydrological Sciences Journal*, **68**, 2357–2372. <https://doi.org/10.1080/02626667.2023.2267544>.
- Potin, P., Colin, O., Pinheiro, M., Rosich, B., O'Connell, A., Ormston, T., Gratadour, J. B. & Torres, R. (2022) 'Status and evolution of the Sentinel-1 Mission', *International Geoscience and Remote Sensing Symposium (IGARSS)*. July 2022, pp. 4707–4710. <https://doi.org/10.1109/IGARSS46834.2022.9884753>.
- Psomiadis, E., Diakakis, M. & Soulis, K. X. (2020) Combining SAR and optical earth observation with hydraulic simulation for flood mapping and impact assessment, *Remote Sensing*, **12**, 3980. <https://doi.org/10.3390/RS12233980>.
- Rawls, W. J., Brakensiek, C. L. & Saxton, K. E. (1982) Estimation of soil water properties, *Transactions of the ASAE*, **25**, 1316–1320. <https://doi.org/10.13031/2013.33720>.
- Refsgaard, J. C. & Storm, B. (1995) MIKE SHE. In: Singh, V. P. (Ed.), *Computer Models of Watershed Hydrology*, pp. 806–846. Littleton, CO: Water Resource Publications.
- Ries, F., Lange, J., Schmidt, S., Puhlmann, H. & Sauter, M. (2015) Recharge estimation and soil moisture dynamics in a Mediterranean, semi-arid karst region, *Hydrology and Earth System Sciences*, **19**, 1439–1456. <https://doi.org/10.5194/hess-19-1439-2015>.
- Running, S. & Mu, Q. (2015) MOD16A2 MODIS/Terra Evapotranspiration 8-day L4 Global 500m SIN Grid. Reston, VA: USGS. <https://doi.org/10.5067/MODIS/MOD16A2.006>.
- Running, S. W., Mu, Q., Zhao, M. & Moreno, A. (2017) 'MODIS Global Terrestrial Evapotranspiration (ET) Product (NASA MOD16A2/A3) Collection 5. NASA Headquarters'. Missoula, MT: Numerical Terradynamic Simulation Group Publications. Available from: https://scholarworks.umt.edu/ntsg_pubs/268.
- Saeedi, M., Sharafati, A. & Tavakol, A. (2021) Evaluation of gridded soil moisture products over varied land covers, climates, and soil textures using in situ measurements: A case study of Lake Urmia Basin, *Theoretical and Applied Climatology*, **145**, 1053–1074. <https://doi.org/10.1007/S00704-021-03678-X/FIGURES/14>.
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M. & Rudolf, B. (2014) GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theoretical and Applied Climatology*, **115**, 15–40. <https://doi.org/10.1007/S00704-013-0860-X/TABLES/3>.
- Schumann, G., Giustarini, L., Tarpanelli, A., Jarihani, B. & Martinis, S. (2022) Flood modeling and prediction using earth observation data, *Surveys in Geophysics*, **44**, 1553–1578. <https://doi.org/10.1007/S10712-022-09751-Y>.
- Simons, G. W. H., Koster, R. & Droogers, P. (2020) Hihydrosoil v2.0 – A high resolution soil map of global hydraulic properties. Wageningen, The Netherlands: FutureWater. <http://www.futurewater.eu/hihydrosoil>.
- Tadesse, T., Senay, G. B., Berhan, G., Regassa, T. & Beyene, S. (2015) Evaluating a satellite-based seasonal evapotranspiration product and identifying its relationship with other satellite-derived products and crop yield: A case study for Ethiopia, *International Journal of Applied Earth Observation and Geoinformation*, **40**, 39–54. <https://doi.org/10.1016/J.JAG.2015.03.006>.
- U.S. Department of Agriculture (2017) Chapter 4: Soil survey methods, In *Soil Survey Manual 2017* (Handbook No. 18). CreateSpace Independent Publishing Platform. <https://books.google.gr/books?id=5nAxswEACAAJ>.
- Véliz-Chávez, C., Mastachi-Loza, C. A., González-Sosa, E., Becerril-Piña, R. & Ramos-Salinas, N. M. (2014) Canopy storage implications on interception loss modeling, *American Journal of Plant Sciences*, **05**, 3032–3048. <https://doi.org/10.4236/AJPS.2014.520320>.
- Wagemann, J., Siemen, S., Seeger, B. & Bendix, J. (2021) Users of open big Earth data – An analysis of the current state, *Computers & Geosciences*, **157**, 104916. <https://doi.org/10.1016/j.cageo.2021.104916>.
- Wagner, W., Wien, T. U., Hahn, S., Melzer, T., Profile, S., Bartalis, Z., Kidd, R., Hasenauer, S., Figa, J., De Rosnay, P., Jann, A., Schneider, S., Rgen Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Zü Ger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Nter Blö Schl, G.,

- Eitzinger, J., Steinnocher, K., Zeil, P. & Rubel, F. (2013) The ASCAT soil moisture product: A review of its specifications, validation results, and emerging applications, *Meteorologische Zeitschrift*, **22**, 1–29. <https://doi.org/10.1127/0941-2948/2013/0399>.
- Wu, L., Seo, D. J., Demargne, J., Brown, J. D., Cong, S. & Schaake, J. (2011) Generation of ensemble precipitation forecast from single-valued quantitative precipitation forecast for hydrologic ensemble prediction, *Journal of Hydrology (Amst)*, **399**, 281–298. <https://doi.org/10.1016/J.JHYDROL.2011.01.013>.
- Xiang, Y., Chen, J., Li, L., Peng, T. & Yin, Z. (2021) Evaluation of eight global precipitation datasets in hydrological modeling, *Remote Sensing*, **13**, 2831. <https://doi.org/10.3390/RS13142831>.
- Xu, X., Li, J. & Tolson, B. A. (2014) Progress in integrating remote sensing data and hydrologic modeling, *Progress in Physical Geography*, **38**, 464–498. https://doi.org/10.1177/0309133314536583/ASSET/IMAGES/LARGE/10.1177_0309133314536583-FIG1.JPEG.
- Yang, Y., Yuan, H. & Yu, W. (2018) Uncertainties of 3D soil hydraulic parameters in streamflow simulations using a distributed hydrological model system, *Journal of Hydrology (Amst)*, **567**, 12–24. <https://doi.org/10.1016/J.JHYDROL.2018.09.042>.

First received 20 December 2023; accepted in revised form 7 October 2024. Available online 19 November 2024