DTE Hydrology Evolution

Final Report

Deliverable D7.1

This page has been intentionally left blank

Change Record

Date	Issue	Section	Page	Comment
31/05/2023	1.0			Final version

Control Document

Process	Name	Date
Written by:	Luca Brocca, Stefania Camici, Christian Massari, Angelica Tarpanelli, Luca Ciabatta, Jacopo Dari, Sara Mondanesi, Hamidreza Mosaffa, Paolo Filippucci (CNR-IRPI), Lorenzo Alfieri, Francesco Avanzi, Simone Gabellani (CIMA), Dominik Rains, Diego Miralles (UGent), Mariette Vreugdenhil, Raphael Quast, Bernhard Raml, Samuel Massart (TUWien), Simone Mantovani (MEEO), Alexander Jacob (EURAC)	
Checked by	Luca Brocca (CNR-IRPI)	
Approved by:		

Table of contents

Technical Note 1: DTE-Hydrology end-to-end description and demonstration plan

1.   Introduction

1.1.   The DTE Hydrology Evolution Project

The overarching objective of Digital Twin Earth (DTE) Hydrology Evolution project is to prototype a full advanced end-to-end demonstrator of DTE over the full Mediterranean region at high resolution in space and time (targeting 1 km and 1 hour) by extending the successful implementation of Digital Twin Earth Hydrology project that has been carried out over the Po River Valley in northern Italy. Specifically, DTE Hydrology has demonstrated the potential to advance towards an end-to-end reconstruction of the hydrological cycle at high resolution in space and time by an effective combination of state-of-the-art EO data, in situ observations, advanced hydrological and hydraulic modelling, AI and advanced digital platform capabilities. DTE Hydrology Evolution project we aim to further advance the DTE Hydrology concept (see **Figure 1.1.1**) through (1) the extension to the full Mediterranean region and hence the preparation to develop a Digital Twin of Hydrological processes and applications at continental and global scales, (2) the exploitation of the dedicated cloud-HPC infrastructure set up in ESRIN, (3) the engagement with a larger community of scientists and stakeholders, (4) the development of an advanced DTE Hydrology open science platform as a community tool offering an interactive access to the EO data, advanced products, in-situ data, and model results, and (5) the development of an attractive and efficient user interface, dashboards, and 3D immersive visualisation tools in an advanced digital environment for scientists, policy makers and citizens. It will be based on state-of-the-art digital platform capabilities and will represent a solid basis for a future ESA supported tool for Polar research and Earth system science.

Figure 1.1.1: DTE Hydrology Evolution concept.

1.2.   Scope of this Report

This report describes the end-to-end description of the satellite products that will be developed during the project and of the modelling systems that will be implemented to develop further modelled products and for addressing the projects case studies: (a) large scale water balance assessment, (b) flood\flooding prediction, (c) what-if scenario for flood risk assessment, and (d) what if scenario for water resources management.

1.3.   Applicable Documents

Not available.

1.4.   Reference Documents

RD-01   DTE Hydrology Evolution Technical Proposal - V1.0

2. DTE Hydrology Datacube

2.1. Precipitation

The satellite precipitation products developed within DTE Hydrology are based on the integration of multiple precipitation and rainfall datasets to generate a high spatial (1 km) and temporal (daily) resolution precipitation and rainfall products over the Mediterranean area. In the beta version of the dataset, the following precipitation and rainfall datasets are downscaled and merged together.

- **CPC**: This dataset is part of products suite from the CPC Unified Precipitation Project that are underway at NOAA Climate Prediction Center (CPC). CPC precipitation is obtained by combining all gauge information sources available at CPC and by taking advantage of the optimal interpolation (OI) objective analysis technique (Xie et al., 2007). See Chen et al. (2008), for further details. Precipitation data are available with a spatial resolution of .5 degree latitude x .5 degree longitude and they are accumulated for each day. This product is available in near real-time, with a latency of a few days.
- **GPM-LR**: The Integrated Multi-satellitE Retrievals for GPM (IMERG) algorithm use the data obtained from the GPM (Global Precipitation Measurement) mission to estimate precipitation over the majority of Earth's surface (Huffman et al., 2018). The resulting product spatial resolution is 0.1 degree, while the temporal resolution is 30 minutes. Here, the Late-run version of the dataset is adopted, characterized by 12–18 h latency. In this study, the rainfall data were accumulated to obtain daily precipitation measurements.
- **SM2RAIN-ASCAT**: This dataset is a new global scale rainfall product obtained from ASCAT satellite soil moisture data through the SM2RAIN algorithm (Brocca et al., 2014; 2019). The SM2RAIN-ASCAT rainfall dataset (in mm/day) is provided over a regular grid at 0.1-degree sampling on a global scale. The product represents the accumulated rainfall between the 00:00 and the 23:59 UTC of the indicated day. The SM2RAIN method was applied to the ASCAT soil moisture product (Wagner et al., 2013) for the period from January 2007 to December 2021 (15 years).

An additional precipitation reanalysis dataset is considered for obtaining the merging weights:

- **ERA5**: ERA5-Land (European Centre for Medium-Range Weather Forecasts, ECMWF, Reanalysis 5th Generation) provides hourly data of various land surface variables from 1950 onwards, combining models with observations. It was produced by replaying the land component of the ECMWF ERA5 climate reanalysis and it is characterized by an improved spatial resolution (0.1 degree), while the temporal resolution is 1 hour (Hersbach et al., 2020). The hourly rainfall was accumulated to daily scale in this study.

The output 1 km grid was obtained using the grid files provided by the project.

In the final version of the dataset, the rainfall data of SM2RAIN-ASCAT will be replaced by the ones from the application of SM2RAIN to high-resolution satellite soil moisture products, in order to increase the spatial accuracy of the final product.

The downscaling procedure is carried out by exploiting the information contained in CHELSA. CHELSA (Climatologies at high resolution for the earth's land surface areas; Karger et al., 2017) is a very high resolution

(30 arc sec, ~1km) global downscaled climate dataset currently hosted by the Swiss Federal Institute for Forest, Snow and Landscape Research WSL. It is based on a mechanistical statistical downscaling of global reanalysis data or global circulation model output and it includes climate layers for various time periods and variables. Monthly precipitation data from 2000 to 2019 were downloaded, extracted and averaged in order to obtain twelve values (one for each month in a year) of average daily precipitation for each pixel of the study area. This information is not useful for the value itself, but the relative precipitation patterns of CHELSA can be exploited to spatialize the coarse resolution information of precipitation products, being based on the modelling of orographic predictors of wind fields, valley exposition and boundary layer height .

The precipitation and rainfall datasets to be downscaled are first resampled to the project grid through a bilinear interpolation, in order to exploit the spatial information of the original product at its fullest. Hence, the downscaling procedure is executed for each product by considering each original coarse resolution pixel and the high resolution ones within it. First, relative CHELSA precipitation patterns are obtained for each month by dividing the mean CHELSA precipitation values of the selected high resolution pixels by their spatial average. Then, the mean monthly values of the selected product are calculated for each high and coarse resolution pixel. For each month, the data of the product to be downscaled are I) divided by their monthly average, II) multiplied for the relative CHELSA precipitation pattern and III) multiplied for the average value of the original coarse resolution pixel. In this way, three different results are achieved: 1) daily spatial patterns different every day, thanks to the use of the bilinear interpolation, 2) realistic monthly pattern of precipitation, thanks to the exploitation of the relative CHELSA precipitation pattern, 3) preservation of the original information thanks to the multiplication for the average value of the original coarse resolution pixel. Finally, a gaussian spatial filter is applied to the obtained downscaled product to reduce the boundary effects generated by the application of the procedure to each original coarse resolution pixel. The main steps of the downscaling process are summarized in **Figure 2.1.1**. The the bilinear interpolation step is here neglected to improve figure clarity.
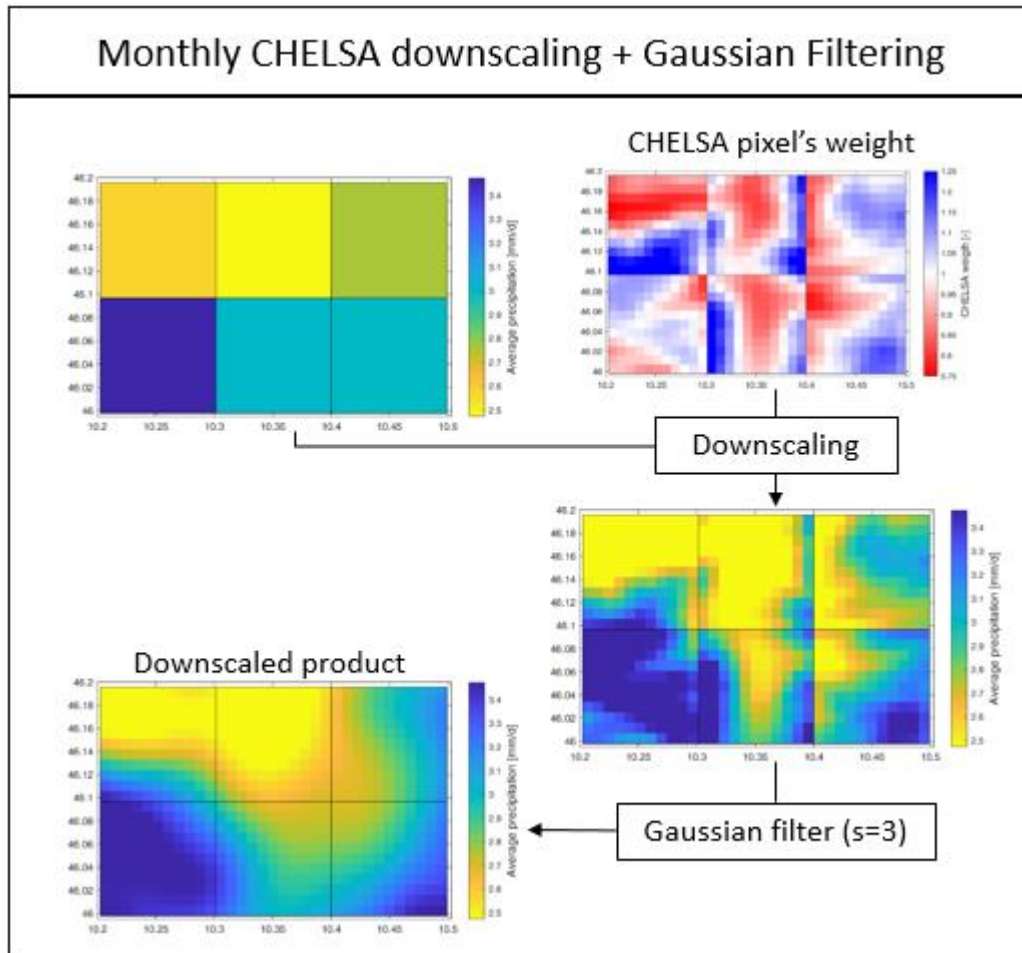
Figure 2.1.1: Downscaling procedure with CHELSA data and gaussian spatial filter. Bilinear interpolation step is neglected to improve figure clarity.

Two different products including the precipitation and rainfall (precipitation liquid fraction) are generated at 1° spatial and daily temporal resolution. 1) The precipitation product is provided by merging the CPC and GPM-LR based on an optimal interpolation technique, here we relied on a triple collocation (TC)-based merging using the signal-to-noise ratio (SNR) proposed in Gruber et al. (2017). In particular, to derive the merged dataset we seek the optimality in a least squares sense, so that the variance of residual random errors is minimized. This leads to a weighted average between CPC and GPM-LR as follows:

$$P_{Merged} = \omega 1 P_{CPC} + \omega 2 P_{GPM}$$

(1)

Where P_{Merged} is precipitation product, $\omega 1$ and $\omega 2$ are the CPC and GPM-LR weight. The weighted is calculated by using error variance estimates obtained from TC approach. The triple collocation approach requires three

independent datasets as input which are CPC, GPM-LR and ERA5 in this study. The mechanism of TC approach is established on a linear error model, which can be represented by the equation as:

$$X_i = \alpha_i + \beta_i t + \varepsilon_i \tag{2}$$

where Xi (i =1, 2, 3) are collocated measurement systems linearly related to the true underlying value t with additive random errors εi, respectively, while αi and βi are the ordinary least squares intercepts and slopes. Assuming that the errors from the independent sources have zero mean (E(εi)=0) and are uncorrelated with each other (Cov(εi,εj) =0, with i ≠ j) and with t(Cov(εi, t)=0) the variance of the error of each dataset can be expressed as:

$$\sigma_\varepsilon = \begin{cases} \sqrt{Q_{11} - \dfrac{Q_{12}Q_{13}}{Q_{23}}} \\ \sqrt{Q_{22} - \dfrac{Q_{12}Q_{23}}{Q_{13}}} \\ \sqrt{Q_{11} - \dfrac{Q_{12}Q_{23}}{Q_{12}}} \end{cases} \tag{3}$$

where Qij = Cov(Xi ,Xj) is the covariance within the variables Xi. Let θi = βiσ(t); then, we can also solve θi as:

$$\theta = \begin{cases} \sqrt{\dfrac{Q_{12}Q_{13}}{Q_{23}}} \\ \sqrt{\dfrac{Q_{12}Q_{23}}{Q_{13}}} \\ \sqrt{\dfrac{Q_{12}Q_{23}}{Q_{12}}} \end{cases} \tag{4}$$

Equations (3) and (4) are the finally derived formal formulas of TC approach for estimating the error of precipitation. The weights are solved as follows:

$$\omega 1 = \frac{\sigma^2_{CPC}}{\sigma^2_{CPC} + \sigma^2_{GPM}} \qquad \omega 2 = \frac{\sigma^2_{GPM}}{\sigma^2_{CPC} + \sigma^2_{GPM}} \tag{5}$$

Assumes that there are no systematic differences between the data sets. This assumption can be circumvented by deriving the weights from signal-to-noise ratio (SNR = (θ^2/σ^2)) properties instead from random error variances.

$$\omega 1 = \frac{SNR_{CPC}}{SNR_{CPC} + SNR_{GPM}} \qquad \omega 2 = \frac{SNR_{GPM}}{SNR_{CPC} + SNR_{GPM}} \tag{6}$$

2) the rainfall product is provided by merging the two precipitation products including CPC and GPM-LR and rainfall product SM2RAIN-ASCAT. Before using the precipitation products (CPC and GPM-LR), it is needed to extract the rainfall part by masking the snow using the Snow data obtained from the 4DMED-Hydrology project.

2.2. Evaporation

The Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011a; 2011b) uses satellite observations to estimate the actual, terrestrial evaporation as the sum of the following components: transpiration (Et), interception loss (Ei), bare-soil evaporation (Eb), snow sublimation (Es) and open-water evaporation (Ew). Intermediate outputs of the model include: potential evaporation (Ep), root-zone soil moisture ($SMroot$), surface soil moisture ($SMsurf$), and evaporative stress (S).

A Priestley and Taylor equation calculates Ep based on observations of surface net radiation and near-surface air temperature (Priestley and Taylor, 1972). Potential evaporation estimates are converted into actual evaporation based on the multiplicative, evaporative stress factor S. The derivation of S is based on microwave observations of the vegetation optical depth – used as a proxy for the vegetation water content – and simulations of root-zone soil moisture. The latter is calculated using a multi-layer running water balance that describes the infiltration of observed precipitation through the vertical soil profile. To correct for random forcing errors, microwave observations of surface soil moisture are assimilated into the soil profile. Interception loss is calculated separately based on the Gash analytical model of rainfall interception (Gash, 1979; Valente et al., 1997) driven by observations of precipitation and both vegetation and rainfall characteristics. Estimates of actual evaporation for regions covered by ice and/or snow are derived using a Priestley and Taylor equation adapted for ice and super-cooled waters. For a detailed description, the reader is directed to Miralles et al. (2011a; b) and Martens et al. (2017).

For this project, a beta version GLEAM 1km dataset was produced for the study period (2015–2021) on a daily time-scale for each sub-domain defined by the project. This beta version used forcing data from the GLEAM v3.6a that were bilinearly interpolated to 0.01°. Precipitation data was replaced with high-resolution data which was internally produced during the project. Also soil properties and land cover fractions were replaced using high resolution data. After having run GLEAM, the output was rescaled to 1 km using grid files as provided by the project. In the final GLEAM 1km version, additional forcing variables such as radiation and temperature will be replaced with observations available at higher resolution (Rains et al., 2022). In addition, soil moisture will be assimilated, which was not done for the beta-version. Table 2.2.1 provides more information on the forcing variables used to produce both datasets.

Table 2.2.1: Overview of forcing datasets.

Forcing Variable	GLEAM-1km (beta version)	GLEAM-1km (final version)
Radiation	MSWX (Beck et al., 2022)	LST-Planet & LSAF (Rains et al., 2022)
Air Temperature	MSWX (Beck et al., 2022)	LST-Planet & LSAF (Rains et al., 2022)
Precipitation	SM2Rain (Brocca et al, 2014)	SM2Rain (Brocca et al, 2014)
Snow Water Equivalent	GLOBSNOW L3v2 (Luojus et al., 2013) & NSIDC v01 (Armstrong et al, 2005)	GLOBSNOW L3v2 (Luojus et al., 2013) & NSIDC v01 (Armstrong et al, 2005)
Vegetation Optical Depth	VODCA (Moesinger et al., 2020)	VOD-Planet
Surface Soil Moisture*	n/a	SM-Planet, SM-TUW
Vegetation Fractions	MOD44B v6.0 (DiMiceli et al., 2015)	MOD44B v6.0 (DiMiceli et al., 2015)
Soil Properties	HiHydroSoil v2.0 (Simons et al., 2020)	HiHydroSoil v2.0 (Simons et al., 2020)

* Note that the surface soil moisture is assimilated into GLEAM thus is not a forcing variable as such.

Initial results based on the beta version of GLEAM 1km seem promising. The seasonal evaporation for summer and winter months is illustrated in Fig. 2.2.1. This dataset was validated against nine eddy-covariance stations available within the study region and period. Their median correlation was equal to 0.81 with a maximum correlation of 0.92 corresponding to the station visualised in Fig. 2.2.2.
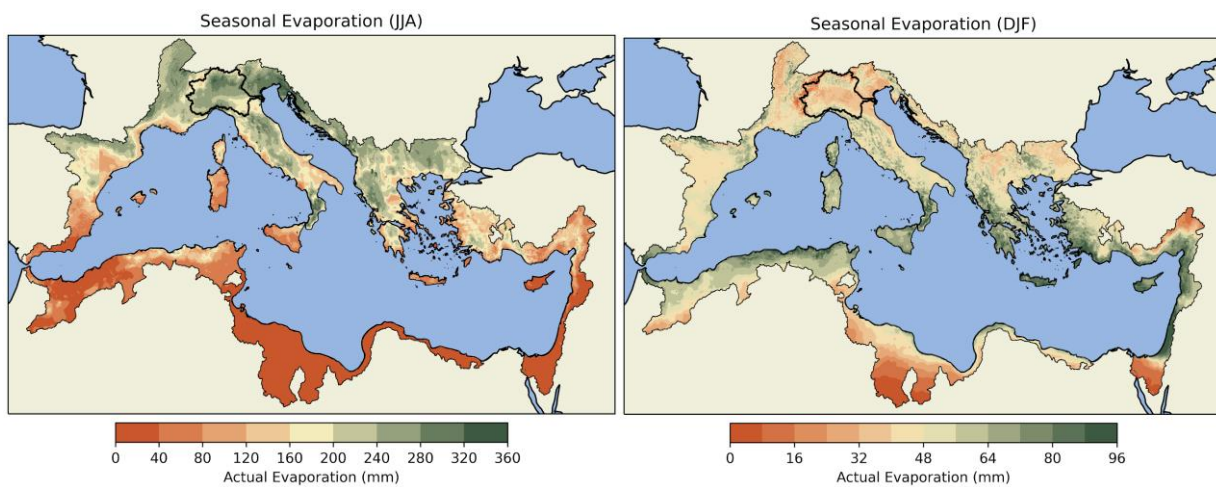
Figure 2.2.1: Seasonal evaporation for the summer (left) and winter (right) months.
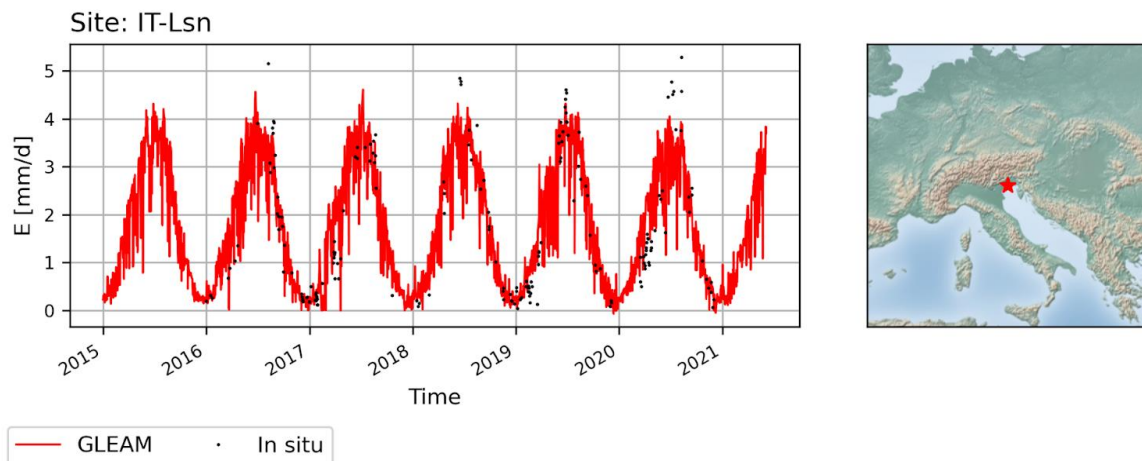
Figure 2.2.2: Daily evaporation according to GLEAM 1km (red) and field observations (black) at an eddy-covariance station close to the Po valley in Italy (45.74048°N, 12.7503°E).

2.3.   Soil moisture

The satellite soil moisture is retrieved from VV-polarized Sentinel-1 backscatter at 500m sampling (Equi7Grid) using the RT1 model (Quast and Wagner, 2019). The original 20m VV backscatter is resampled to 500m, only using pixels that are sensitive to soil moisture. At 20m, pixels that are located over cities, lakes and are subject to sub-surface scattering are masked out. In the dry regions, such as the Mediterranean, sub-surface scattering can affect the backscatter signal (Wagner et al., 2022). With dry soils, the assumption that backscatter is linearly related to soil moisture is violated. As the backscatter signal penetrates the soil deeper, multiple scattering can start to occur through interaction with rocks in the sub-surface. This can lead to higher backscatter, even though soil moisture is at its minimum. As no model currently accounts for sub-surface backscatter, pixels that show this behaviour are masked out at 20m.

To identify pixels which violate the linearity assumption, we compute the Pearson correlation coefficient between four years of backscatter signal and ERA5-Land soil moisture (Muñoz-Sabater et al., 2021), where the ERA5-Land data is resampled to match the 20m grid. This correlation map is then used to mask out pixels at 20m, which have a correlation coefficient lower than 0.2. An example of regions masked over the Iberian Peninsula is shown in **Figure 2.3.1**. These regions display close correspondence to areas with strong subsurface scatter effects identified by Wagner et al., 2022.
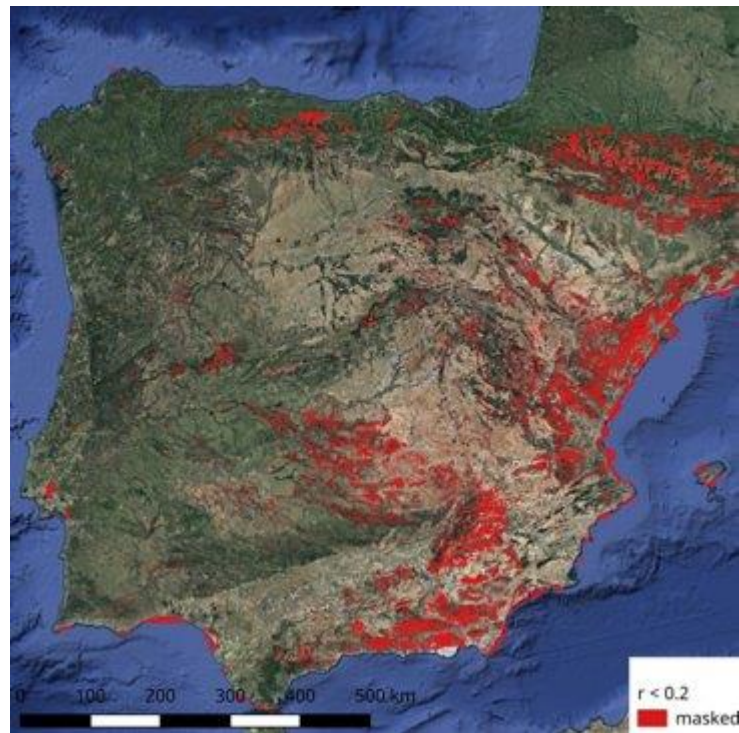
Figure 2.3.1: Masked pixels where the Pearson correlation coefficient between the backscatter signal and ERA5-Land soil moisture is below 0.2.

Soil moisture is then retrieved from the 500m VV backscatter data using the RT1 model. The RT1 algorithm is based on a first-order expansion of the radiative transfer equation, modelling backscatter as the result of scattering from a (rough) surface covered by vegetation. The retrieval is performed via a non-linear least squares optimization that minimizes the difference between the incidence-angle dependent measured and modelled σ_0 timeseries for each pixel individually. To correct for seasonal effects introduced by changing vegetation coverage, the vegetation optical depth is parametrized dynamically using Leaf Area Index from the Copernicus Global Land Service at 300 m, resampled to the EQUI7Grid at 500m.

The resulting soil moisture product has a spatial and temporal resolution of 1 km and nearly 3 days, respectively, and a temporal coverage from 2017 to June 2022. The high spatial detail allows obtaining more reliable soil moisture information over areas which are usually affected by topography in the coarse resolution soil moisture products. The soil moisture product shows good performance over Italy, compared to in situ and model data. Validation was done with ERA5-Land soil moisture at 9 km using backscatter with and without sub-surface scattering masking. Results show a clear increase in performance, especially in drier regions such as Sicily, Sardinia and Puglia (**Figure 2.3.2**). Lower correlations are still observed over the Italian sclerophyllous and deciduous forests ecoregion running from the Alps to the Apennines.
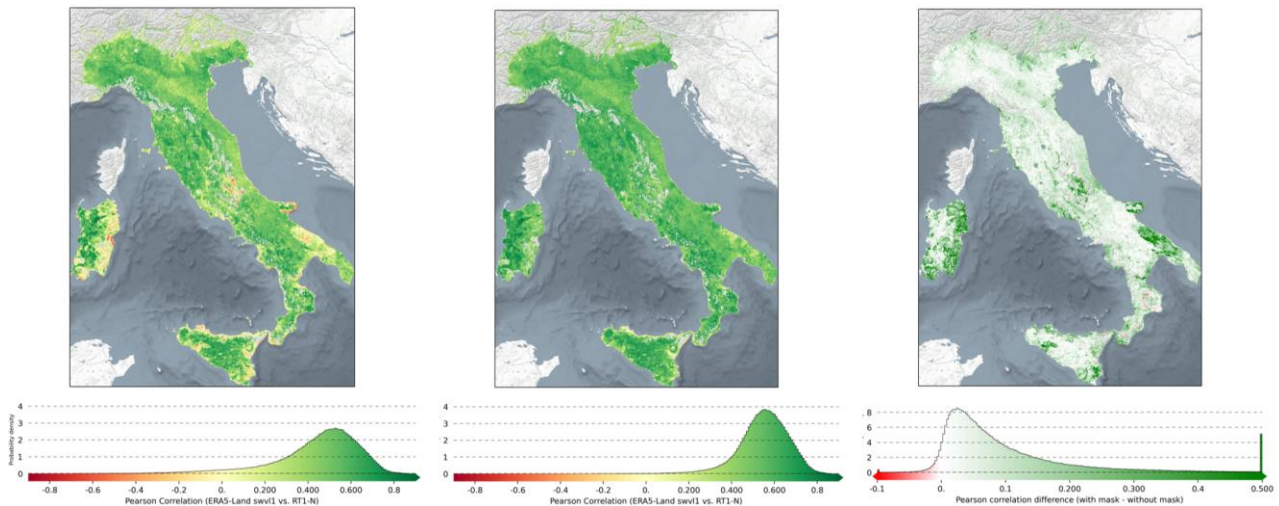
Figure 2.3.2: Pearson correlation coefficient between RT1 soil moisture, without (left) and with (middle) masking for subsurface scattering, and ERA5-Land soil moisture. The right image shows the difference in Pearson correlation between the two RT1 soil moisture products with ERA5-Land.

To have an independent validation of the soil moisture data, in situ validation was done over all International Soil Moisture Network stations in Italy which had data availability at a 0-10cm depth during the period 2017-2022. The RT1 soil moisture shows similar temporal dynamics to in situ soil moisture capturing main rainfall events and dry plateaus (**Figure 2.3.3**). Moderate correlations, varying between 0.453 and 0.572, are found with in situ soil moisture data (**Table 2.3.1**), demonstrating the ability of RT1 soil moisture to capture actual surface soil moisture dynamics. Lower correlations are found over in situ stations within pixels with forest as the dominant land cover type. An example is the Oltrepo station, where correlations are moderate. It is clear that RT1 soil moisture shows less dynamics than the in situ soil moisture. This is a result of the complex vegetation geometry of the forested areas which reduces RT1 sensitivity of surface soil moisture and induces noise in the retrievals.

Table 2.3.1: Performance of soil moisture products with in situ stations in Italy.

Network	Station	Pearson R
UMBRIA	Petrelle	0.572
	Torre-Olmo	0.505
	Cerbara	0.562
STEMS	Can_all	0.542
Oltrepo	Oltrepo	0.453
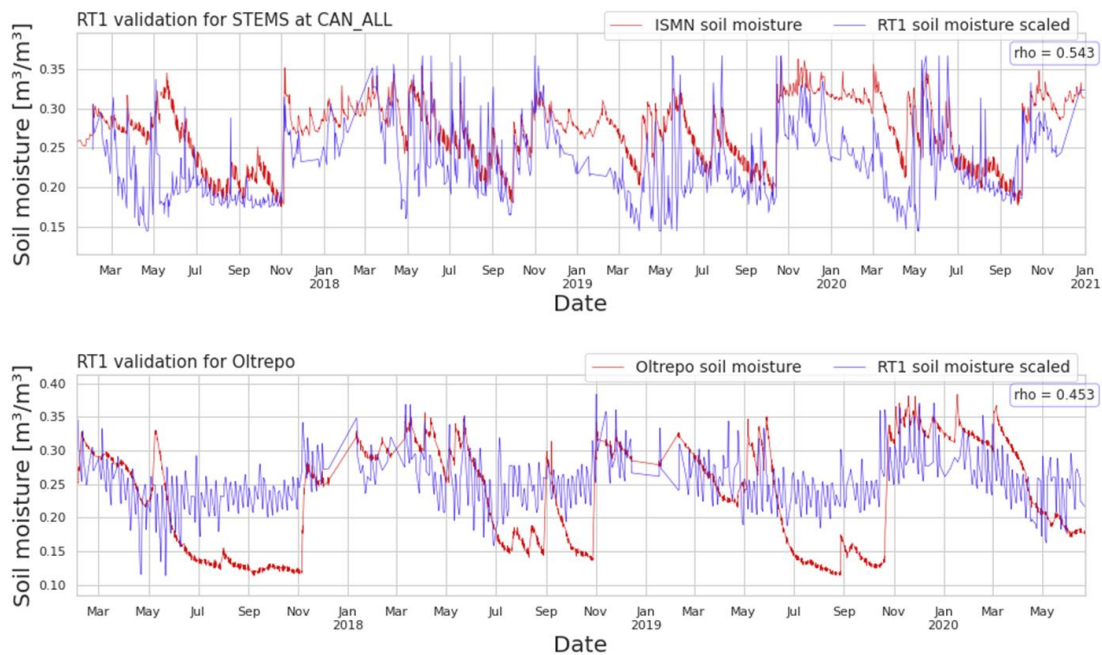
Figure 2.3.3: time series and Pearson correlation coefficient between RT1 soil moisture and in situ soil moisture for two stations.

2.4. River discharge

The river discharge product is derived completely by satellite through the combination of two different satellite data products (Tarpanelli et al., 2015; Tarpanelli et al., 2019): water level derived by altimeter and reflectance indices derived by multispectral images. It is well established that from elevation level measurements it is possible to estimate river discharge through the well-known rating curve (Belloni et al., 2021). The L-2 product of water level is downloaded from the portals available from different laboratories (DAHITI, https://dahiti.dgfi.tum.de/en/; THEIA, http://hydroweb.theia-land.fr/) and from the Copernicus Global Land Service (https://land.copernicus.eu/global/products/wl).

Similarly, for the multispectral images, the reflectance signal of an area near the river is sensitive to the increase in water due to a flood event in that area. The reflectance ratio (called reflectance index) between a constant signal coming from an urban area (or a bare soil), dry, and this wet signal, is correlated with the river discharge and provides good indicator of flow variation in the river (Tarpanelli et al., 2013). The L-2 reflectance index is extracted from the multi-temporal analysis of MSI/Sentinel-2 images (Filippucci et al., 2022). This high-resolution sensor is expected to provide useful information for the rivers in the Mediterranean areas characterised by width on average of about 50-100 m.

The algorithm to derive river discharge, Q [m³/s], is directly based on the combination of the L-2 products as in the following:

where A_b is the basin area, H is the water surface elevation derived by altimetry, H_0 is the minimum level observed by altimetry for the specific site, *CM* is the reflectance index extracted for the specific site. *a*, *b*, and *g* are three parameters that need to be calibrated by using the ground observation of river discharge (modelled discharge is also suitable if reliable enough) through the minimization of the error (typically, Nash Sutcliffe or Root Mean Square Error).

Because the rivers are narrow the estimation of river discharge by the combination of satellites is challenging and it is possible that one of the two L-2 products is not available for the selected sites or for brief periods (for example during winter the snow can affect the reflectance). In such cases, the combination of products is replaced by the estimation of discharge by single sensors. The procedure has been already tested within the ESA STREAMRIDE project.

2.5.   Irrigation

2.5.1 The SM-based inversion approach

The approach was born as an evolution of the SM2RAIN algorithm (Brocca et al., 2014), originally developed to estimate rainfall from soil moisture observations. Currently, the method can be implemented to retrieve irrigation amounts as well by inverting the soil water balance (Brocca et al., 2018; Dari et al., 2020; 2022), expressed as:

$$nZdS(t)/dt \; = \; i(t) + r(t) - g(t) - sr(t) - e(t) \qquad (2.5.1.1)$$

in which n [-] is the soil porosity, Z [mm] is the depth of the soil layer, S(t) [-] indicates the relative soil moisture, i(t) [mm/day] represents the irrigation rate, r(t) [mm/day] is the rainfall rate, g(t) [mm/day] is the drainage rate, sr(t) [mm/day] represents the surface runoff, and e(t) [mm/day] indicates the actual evapotranspiration rate. Eq. (2.5.1.1) is equivalent to the following:

$$Win(t) \; = \; nZdS(t)/dt + g(t) + sr(t) + e(t) \qquad (2.5.1.2)$$

where Win(t) indicates the algorithm output, i.e., the sum of irrigation and rainfall rates. By assuming the contribution of the surface runoff negligible (Brocca et al., 2015) and by expressing the drainage term as a soil moisture function, $g(t) = aS(t)^b$ with a [mm] and b [-] indicating drainage parameters, the solving equation can be expressed as:

$$Win(t) \; = \; Z^* \, dS(t)/dt + aS(t)^b + e(t) \qquad (2.5.1.3)$$

where $Z^*= nZ$ [mm] represents the water capacity of the soil layer. A soil-moisture-limited approach is used to compute the actual evapotranspiration term, which is expressed as the potential evapotranspiration rate, e0

(t), multiplied by the available water content; an adjustment factor, F, ranging between 0.6 and 1.4 is adopted. As a result, Eq. (2.5.1.3) can be written as follows:

$$Win(t) = Z^* \, dS(t)/dt + aS(t)^b + F * S(t) * e0 \, (t) \qquad (2.5.1.4)$$

with Z*, a, b, and F parameters to be calibrated. Once the parameters are calibrated, the total amount of water entering the soil can be estimated through the Eq. (2.5.1.4). Hence, irrigation is estimated by removing rainfall rates from the output, $i(t) = Win(t) - r(t)$. Negative irrigation rates (if any) are set equal to zero. In order to remove negligible irrigation rates attributable to random errors, the results are discarded if the ratio between weekly estimated irrigation and weekly rainfall is lower than 1.2. The SM-based inversion approach has been implemented in the same configuration as in the ESA Irrigation+ project, i.e., through 1 km surface soil moisture from the RT1 Sentinel-1 data set (Quast et al., 2019), rainfall from ERA5-Land at ~9 km spatial resolution (Muñoz-Sabater et al., 2021), and potential evapotranspiration data from the GLEAM (Global Land Evaporation Amsterdam Model) v3.5b product at 0.25° spatial resolution (Martens et al., 2017; Miralles et al., 2011). The algorithm parameters have been calibrated through the iterative procedure summarised in Fig. 2.5.1.
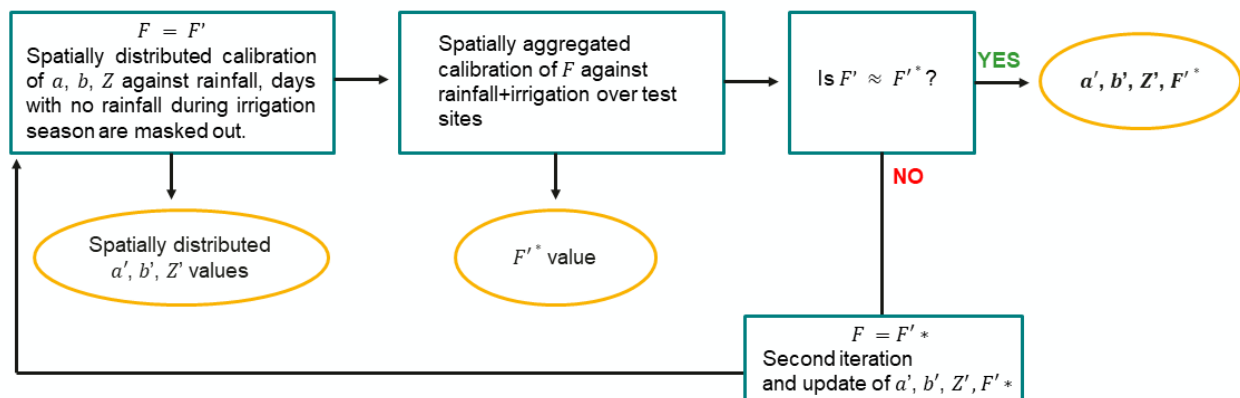
Figure 2.5.1. Iterative procedure adopted for calibrating the parameters of the SM-based inversion approach.

In a first step, the F value is set equal to 1 and the a, b, and Z* parameters are calibrated against rainfall; potential irrigation periods (i.e., days with no rainfall during the irrigation season) are masked out. The evapotranspiration adjustment factor, F, is then calibrated against the sum of rainfall plus irrigation over selected pilot sites where irrigation rates are known. At this point, the calibrated F value is compared with the first guess (F = 1) and, in case of disagreement, the values of a, b, and Z* are re-calibrated by repeating the first step and adopting the updated F value.

3.   Modelling system

3.1.   Hydrological Modelling

Two different hydrological models will be implemented in the project: (1) the fully distributed model CONTINUUM and (2) the modular model MISDc.

3.1.1. Fully Distributed Model: CONTINUUM

Continuum (Silvestro et al., 2013) is a distributed hydrological model relying on a morphological approach based on the identification of the drainage network components (Giannoni et al., 2000). Continuum is a trade-off between empirical models, which are easier to implement but sometimes lack realistic features, and complex physically based models, which try to reproduce all the hydrological processes in detail, heavily relying on parameterisation. The main drawback of the latter type is that such parameterisations are likely to introduce considerable uncertainties, especially when observations are sparse, which results in a non-robust estimate of the parameters.

The Continuum model was developed keeping the physical description of the hydrological processes as simple as possible. This resulted in good performance comparable to existing models, with an increased computational efficiency. In particular, the reduced complexity of the mathematical modelling and the relatively small number of parameters lead to a considerably lower calibration effort, increasing model robustness and portability to data-scarce environments. Thanks to the increased efficiency of the code, Continuum can be easily implemented in an ensemble configuration, enabling the modeller to directly estimate the prediction uncertainties.

Continuum is able to reproduce the spatial-temporal evolution of runoff, soil moisture, energy fluxes, surface soil temperature, snow accumulation and melting, and evapotranspiration. Moreover, it can account for the vegetation seasonal variability in terms of interception and evaporation. Deep flow and water table evolution are modelled with a simple scheme that reproduces the main physical characteristics of the processes, while a distributed interaction between water table and soil surface is represented with a simple parameterisation. The introduction of the so-called force-restore equation for the surface energy balance allows the calculation of the land surface temperature, which can be used for calibration and/or assimilation of remote sensing data.

In this project, Continuum was set up over the Mediterranean coast using the procedure described by Alfieri et al. (2022) and divided into 28 hydrologically consistent domains. Raster maps have constant grid spacing of 0.009° (~1 km along the latitudes and ~800m along the longitudes) and a time resolution of 1 h.
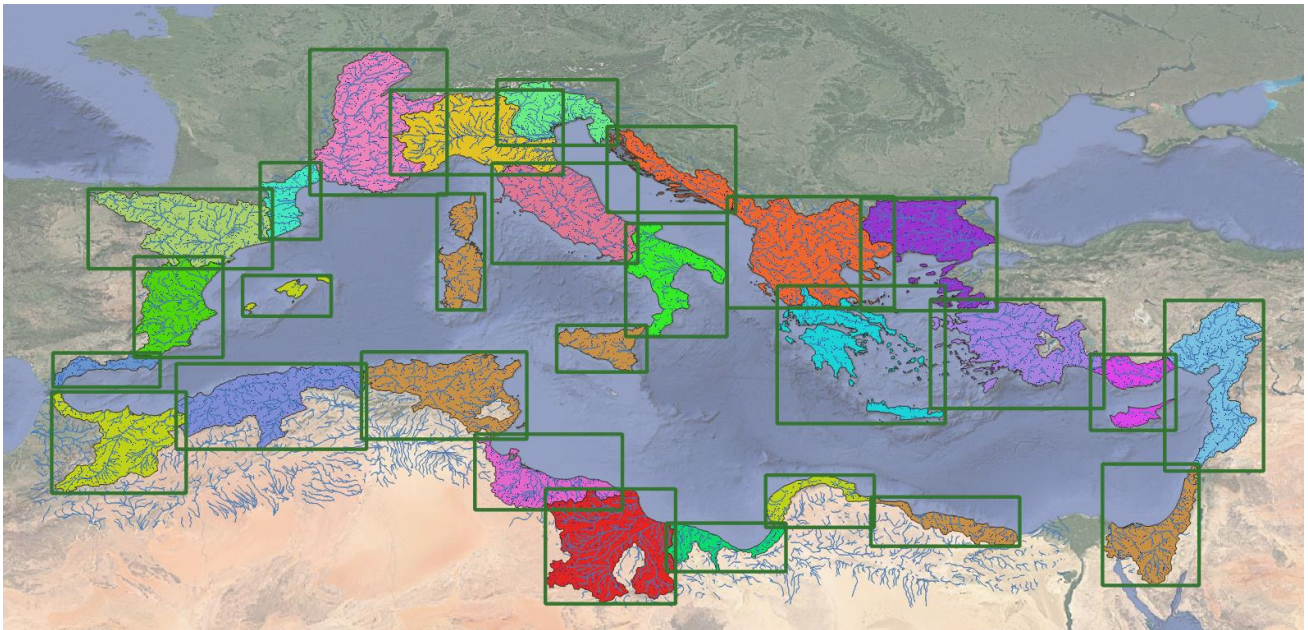
Figure 3.1.1: The 28 simulation domains of Continuum along the Mediterranean coast.

Static data

To produce the Curve Number map used in Continuum to model direct runoff or infiltration from rainfall excess, we combined information on the land use, land cover and land hydrological behaviours. In detail, we extracted land use and land cover information at the 300 m resolution grid from the global dataset ESA-CCI 2018 Land Cover map (ESA, 2017). The hydrologic soil type map was obtained considering the soil fraction in Sand and Clay with a 250 m spatial resolution from the ISRIC SoilGrids global dataset (Hengl et al., 2017). The same dataset has been used to estimate the first-attempt field capacity with the Saxton et al. (1986) pedotransfer function.

To produce the static layers related to the vegetation coverage we used the global land cover map ESA-CCI 2018, by relating the types of land use to different vegetation properties by means of look-up tables based on literature values.

Point data

A set of 186 reservoirs and 52 natural lakes have been included in the model setup (Figure 3.2). The selection was based on taking water bodies exceeding a maximum storage of 100Mm^3. Information on the dams and the corresponding reservoirs were taken from the GranD database (Lehner et al., 2011). Data ingested for each dam by the Continuum model include geographic coordinates, maximum stored volume, initial volume, maximum non-damaging discharge at the outflow gates, weir length, maximum storage level, outflow coefficient, and coordinates of the release point. For lakes, required metadata are the outlet coordinates, minimum volume inducing outflow discharge, initial volume, and emptying coefficient.
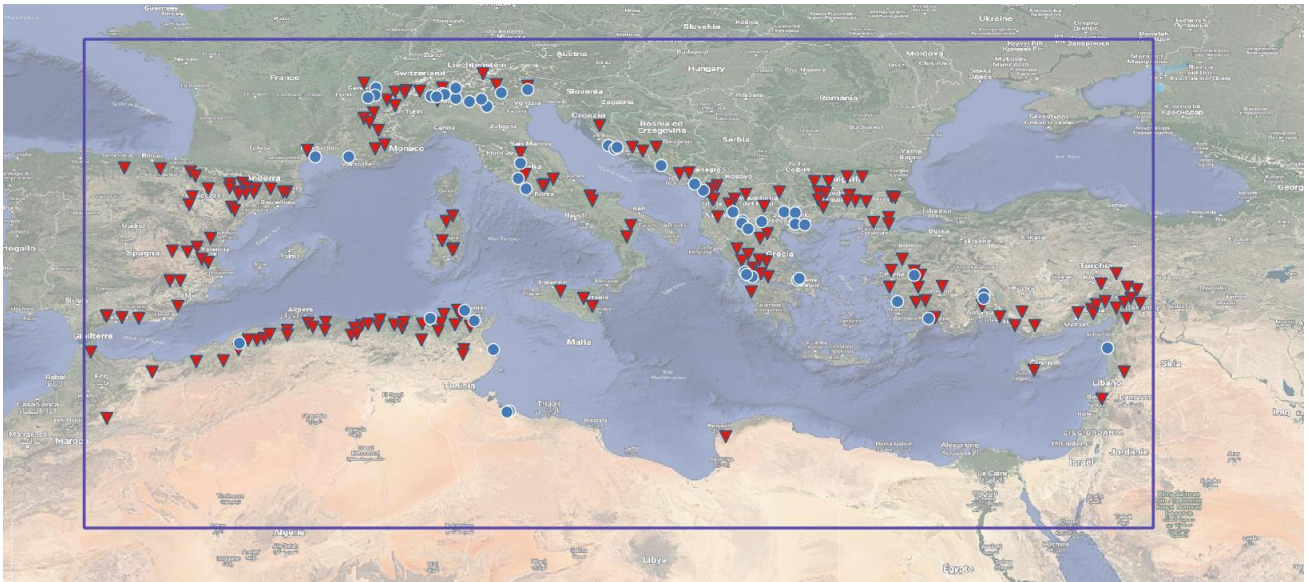
Figure 3.1.2: Reservoirs (red triangles) and natural lakes (blue circles) included in the Continuum hydrological model setup for the project. Bounding box of the simulation domain is shown with a blue rectangle.

Hydrological data

Hydrological data, in the form of river discharge, will be used for validation and if possible for calibration. Data is being collected in the framework of the sister project 4DMed and shared to the partners through a dedicated portal (https://edp-portal.eurac.edu/cdb_doc/4dmed/).

Meteorological data

Meteorological data is a key input to the hydrological model. For this project we will combine datasets coming from different sources. ERA5, the latest atmospheric reanalysis released by ECMWF will be used for the baseline simulations over the 28 domains, while local higher resolution data may be used for assessing the model sensitivity to different accuracy and spatial resolution. In addition satellite data produced in this project will be tested to evaluate their benefits in comparison to state of the art reanalysis data.

3.1.2. Modular Model: MISDc

MISDc, "Modello Idrologico SemiDistribuito in continuo" (Brocca et al., 2011; Camici et al., 2020) is a two-layer continuous hydrological model characterised by three components: a snow module estimating the temporal pattern of snow water equivalent, a soil module simulating the temporal pattern of soil moisture content and a rainfall–runoff transformation component modelling river discharge time series.

The MISDc model uses as input daily precipitation and air temperature data and models the temporal evolution of river discharge, actual evapotranspiration, soil moisture content and snow water equivalent for a surface and a root-zone soil layer. Solid and liquid precipitation are separated through a degree-day method modelling the melting process of snowpack and glaciers. Water is extracted from the first layer by evapotranspiration, which is calculated by a linear function between the potential evaporation and soil moisture. A non-linear relation is used for computing the percolation from the surface to the root-zone layer. The rainfall excess is calculated by a power law relationship as a function of the first layer soil moisture while base flow is a non-linear function of the soil moisture of the second layer. The MISDc has twelve parameters to be estimated by calibration against ground-based river discharge observations.

The model, tested in many different basins in Italy (see e.g. Brocca et al., 2011, 2013a, Massari et al., 2015a, Masseroni et al., 2017, and Cislaghi et al., 2020), in Europe (see e.g. Brocca et al., 2013b, Massari et al., 2018, and Camici et al., 2018) and for different applications (e.g. climate change impact studies; see Camici et al., 2014, soil moisture data assimilation; see Massari et al. 2015b, flood modelling Camici et al., 2020) provided very good results for flood prediction in near natural basins. As the model does not include any component to represent the human–water interactions (e.g., presence of water reservoirs or irrigation), it could fail in reproducing river discharge over non natural basins. However, being a modular model the MISDc can be easily adapted to modeling river discharge over anthropized basins. For the purpose of the DTE analyses, specific components will be added to MISDc model to predict the reservoir water storage and the releases (see eg., Burek et al., 2013 or Dang et al., 2020) as well as to model the irrigation timing and volume (see eg., ….).

In this project, MISDc will be set up over selected basins (Po, Tiber, ..) to build "what-if scenarios" for flood prediction and water resources management. For flood prediction a procedure similar to the one described in Brocca et al., (2013a) will be adopted. The MISDc model, forced by satellite precipitation data produced in this project, will estimate the soil moisture conditions and the corresponding streamflow for different precipitation events occurred in the study basin. The results of these simulations will be stored into a database that, including a large number of rainfall-runoff scenarios, will allow us to estimate the basin response to a rainfall event taking into account the uncertainties of rainfall forecasts and soil moisture conditions of the basin.

In a similar fashion, "what-if scenarios" will be developed for water resources management. In this part will be crucial to add the reservoir and irrigation modules to accurately reproduce the human impact on water resources.

3.2.   Hydraulic modelling

TELEMAC-2D (Galland et al., 1991, see http://www.opentelemac.org) is a 2D hydrodynamic model that uses the Saint Venant equation and models free-surface flows in two dimensions of horizontal space. At each point of the mesh, it calculates the depth of water and the two velocity components. TELEMAC-2D is used for a wide spectrum of applications including the design of hydraulic structures, dam breaches, and flood analysis. The main outputs of the model are water depth and water velocity, spatial and temporal evolution within the river bed and on the floodplain. In this work the TELEMAC-2D model was coupled with the Continuum hydrological model for specific case studies, primarily the case of the Medicane Apollo occurred in the Sicily region (Italy) in October 2021. TELEMAC-2D uses the stream-flow simulated by the Continuum as boundary conditions. A

very-high-resolution Digital Terrain Model, has been used for the implementation of TELEMAC-2D. The DTM from Lidar with a resolution of 2m is freely available from the Sicily Region (https://www.sitr.regione.sicilia.it/).

4. References

Alfieri, L., Avanzi, F., Delogu, F., Gabellani, S., Bruno, G., Campo, L., Libertino, A., Massari, C., Tarpanelli, A., Rains, D., Miralles, D. G., Quast, R., Vreugdenhil, M., Wu, H., and Brocca, L.: High-resolution satellite products improve hydrological modeling in northern Italy, Hydrol. Earth Syst. Sci., 26, 3921–3939, https://doi.org/10.5194/hess-26-3921-2022, 2022.

Armstrong, R., Brodzik, M., Knowles, K., and Savoie, M. (2005). Global Monthly EASE-Grid Snow Water Equivalent Climatology, Version 1, doi: http://dx.doi.org/10.5067/KJVERY3MIBPS.

Avanzi, F., Ercolani, G., Gabellani, S., Cremonese, E., Pogliotti, P., Filippa, G., Morra di Cella, U., Ratto, S., Stevenin, H., Cauduro, M., and Juglair, S. (2020). Learning about precipitation orographic enhancement from snow-course data improves water-balance modeling, Hydrol. Earth Syst. Sci. Discuss. [preprint], https://doi.org/10.5194/hess-2020-571, in review.

Beck, H. E., van Dijk, A. I. J. M., Larraondo, P. R., McVicar, T. R., Pan, M., Dutra, E., and Miralles, D. G. (2022). MSWX: Global 3-Hourly 0.1° Bias-Corrected Meteorological Data Including Near-Real-Time Updates and Forecast Ensembles, Bulletin of the American Meteorological Society, 103(3), E710-E732. Retrieved from http://www.gloh2o.org/mswx/

Belloni R., Camici S., Tarpanelli A. (2021) Towards the continuous monitoring of the extreme events through satellite radar altimetry observations. Journal of Hydrology, 603, Part A, 126870. https://doi.org/10.1016/j.jhydrol.2021.126870

Brocca L., Liersch S., Melone F., Moramarco T., Volk M. (2013a). Application of a model-based rainfall-runoff database as efficient tool for flood risk management. Hydrol. Earth Syst. Sci. 17, 3159.

Brocca L., Melone F., MoramarcoT. (2011). Distributed rainfall-runoff modeling for flood frequency estimation and flood forecasting. Hydrol. Process. 25 (18), 2801–2813.

Brocca L., Moramarco T., Dorigo W., Wagner W. (2013b). Assimilation of satellite soil moisture data into rainfall-runoff modelling for several catchments worldwide. In: Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International. IEEE, pp. 2281–2284.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., Levizzani, V., 2014. Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. J. Geophys. Res., 119, 5128–5141.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., Wagner, W. (2019). SM2RAIN-ASCAT (2007-2018): global daily satellite rainfall from ASCAT soil moisture. Earth System Science Data, 11, 1583–1601, doi:10.5194/essd-11-1583-2019.

Brocca, L., Massari, C., Ciabatta, L., Moramarco, T., Penna, D., Zuecco, G., Pianezzola, L., Borga, M., Matgen, P., Martínez-Fernández, J.: Rainfall estimation from in situ soil moisture observations at several sites in Europe: an evaluation of the SM2RAIN algorithm. J. Hydrol. Hydromechanics 63, 201–209. https://doi.org/10.1515/johh- 2015-0016, 2015.

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., Fernández-Prieto, D., 2018. How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. Int. J. Earth Obs. Geoinformation, 73, 752-766.

Burek, P., Van Der Knijff, J.M., Ad de Roo, 2013. LISFLOOD – Distributed Water Balance and flood Simulation Model – Revised User Manual. Doi: http://dx.doi.org/10.2788/24719.

Camici, S., Brocca, L., Melone, F., Moramarco, T. (2014). Impact of climate change on flood frequency using different climate models and downscaling approaches. Journal of Hydrologic Engineering, 19(8), 04014002, doi:10.1061/(ASCE)HE.1943-5584.0000959. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000959.

Camici, S., Ciabatta, L., Massari, C., Brocca, L. (2018). How reliable are satellite precipitation estimates for driving hydrological models: a verification study over the Mediterranean area. Journal of Hydrology, 563, 950-961, doi:10.1016/j.jhydrol.2018.06.067. https://doi.org/10.1016/j.jhydrol.2018.06.067.

Camici, S., 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.

Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Wayne Higgins, R., and Janowiak, J. E.: Assessing objective techniques for gauge-based analyses of global daily precipitation, J. Geophys. Res., 113, D04110, https://doi.org/10.1029/2007JD009132, 2008.

Cislaghi, A., Masseroni, D., Massari, C., Camici, S., Brocca, L. (2020). Combining rainfall-runoff model and regionalization approach for flood and water resource assessment in the western Po-Valley (Italy). Hydrological Science Journal, in press, doi:10.1080/02626667.2019.1690656. https://doi.org/10.1080/02626667.2019.1690656.

Dang, T. D., Chowdhury, A. F. M., and Galelli, S. (2020). On the representation of water reservoir storage and operations in large-scale hydrological models: implications on model parameterization and climate change impact assessments. *Hydrology and Earth System Sciences*, *24*(1), 397-416.

Dari, J., Brocca, L., Quintana-Seguí, P., Escorihuela, M.J., Stefan, V., Morbidelli, R., 2020. Exploiting high-resolution remote sensing soil moisture to estimate irrigation water amounts over a Mediterranean region. Remote Sens., 12, 2593.

Dari, J., Quintana-Seguí, P., Morbidelli, R., Saltalippi, C., Flammini, A., Giugliarelli, E., Escorihuela, M.J., Stefan, V., Brocca, L., 2022. Irrigation estimates from space: Implementation of different approaches to model the

evapotranspiration contribution within a soil-moisture-based inversion algorithm. Agric. Water Manag., 265,107537.

DiMiceli, C., Carroll, M., Sohlberg, R., Kim, D., Kelly, M., and Townshend, J. (2015). MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD44B.006

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T. and Lecomte, P. (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.

ESA (2017) Land Cover CCI Product User Guide Version 2. Tech. Rep., 2017.

Faroux, S., Kaptué Tchuenté, A. T., Roujean, J.-L., Masson, V., Martin, E. and Le Moigne, P. (2013) ECOCLIMAP-II/Europe: a twofold database of ecosystems and surface parameters at 1 km resolution based on satellite information for use in land surface, meteorological and climate models, Geoscientific Model Development, 6(2), 563–582, https://doi.org/10.5194/gmd-6-563-2013.

Filippucci, P., Brocca, L., Bonafoni, S., Saltalippi, C., Wagner, W., & Tarpanelli, A. (2022). Sentinel-2 high-resolution data for river discharge monitoring. *Remote Sensing of Environment*, *281*, 113255. https://doi.org/10.1016/j.rse.2022.113255

Gash, J.H.C. (1979). An analytical model of rainfall interception by forests, Quarterly Journal of the Royal Meteorological Society, 105, 43–55, doi: 10.1002/qj.49710544304.

Giannoni, F., Roth, G. and Rudari, R. (2000) A semi-distributed rainfall-runoff model based on a geomorphologic approach, Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, 25(7–8), 665–671, https://doi.org/10.1016/s1464-1909(00)00082-4.

Hengl, T., Jesus, J. M. de, 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. and Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning, PLOS ONE, 12(2), e0169748, https://doi.org/10.1371/journal.pone.0169748.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Huffman, G., Bolvin, D., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E., and Xie, P.: Algorithm Theoretical Basis

Document (ATBD) Version 4.5, NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) NASA, 2018.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, P., Kessler, M.: Climatologies at high resolution for the Earth land surface areas, Scientific Data, 4, 170122, https://doi.org/10.1038/sdata.2017.122, 2017.

Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., Thépaut, J. N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020

Luojus, K., Pulliainen, J., Takala, M., Lemmetyinen, J., Kangwa, M., Smolander, T., and Derksen, C.(2013). Global snow monitoring for climate research: Algorithm theoretical basis document (ATBD) - SWE-algorithm, Tech. Rep. Version/Revision 1.0/02.

Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., Verhoest, N.E.C., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev. 10, 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Massari C., Brocca L., Ciabatta L., Moramarco T., Gabellani S., Albergel C., De Rosnay P., Puca S., Wagner W. (2015a). The use of H-SAF soil moisture products for operational hydrology: flood modelling over Italy. Hydrology 2 (1), 2–22.

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

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, doi: 10.3390/rs10020292. https://doi.org/10.3390/rs10020292.

Masseroni, D., Cislaghi, A., Camici, S., Massari, C., Brocca, L. (2017). A reliable rainfall-runoff model for flood forecasting: review and application to a semiurbanized watershed at high flood risk in Italy. Hydrology Research, 48(3), 726-740, doi:10.2166/nh.2016.037. http://dx.doi.org/10.2166/nh.2016.037.

Miralles, D.G., de Jeu, R.A.M., Gash, J.H., Holmes, T.R.H. and Dolman, A.J. (2011b). Magnitude and variability of land evaporation and its components at the global scale, Hydrology and Earth System Sciences, 15, 967–981, doi: 10.5194/hess-15-967-2011.

Miralles, D.G., Holmes, T.R.H., De Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J., 2011a. Global land-surface evaporation estimated from satellite-based observations. Hydrol. Earth Syst. Sci. 15, 453–469. https://doi.org/10.5194/hess-15-453-2011

Moesinger, L., Dorigo, W., Jeu, R. D., Schalie, R. V. D., Scanlon, T., Teubner, I., and Forkel, M. (2020). The global long-term microwave vegetation optical depth climate archive (VODCA). Earth System Science Data, 12(1), 177-196. doi: 10.5194/essd-12-177-2020.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications (preprint). Data, Algorithms, and Models. https://doi.org/10.5194/essd-2021-82

Priestley, J. H. C. and Taylor, J. (1972). On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Monthly Weather Review, 100, 81–92.

Quast, R., Albergel, C., Calvet, J.-C., Wagner, W., 2019. A Generic First-Order Radiative Transfer Modelling Approach for the Inversion of Soil and Vegetation Parameters from Scatterometer Observations. Remote Sens. 11, 285. https://doi.org/10.3390/rs11030285

Rains, D., Trigo, I., Dutra, E., Ermida, S., Ghent, D., Hulsman, P., Gómez-Dans, J., and Miralles, D. G. (2022). High-resolution all-sky land surface temperature and net radiation over Europe. Earth System Science Data Discussions. doi: 10.5194/essd-2022-302.

Ross, C. W., Prihodko, L., Anchang, J., Kumar, S., Ji, W. And Hanan, N. P. (2018). Global Hydrologic Soil Groups (HYSOGs250m) for Curve Number-Based Runoff Modeling, ORNL Distributed Active Archive Center..

Saxton, K. E., Rawls, W. J., Romberger, J. S., and Papendick, R. I.: Estimating Generalized Soil-water Characteristics from Texture, Soil Science Society of America Journal, 50, 1031–1036, https://doi.org/10.2136/sssaj1986.03615995005000040039x, 1986.

Silvestro, F., Gabellani, S., Delogu, F., Rudari, R. and Boni, G. (2013). Exploiting remote sensing land surface temperature in distributed hydrological modelling: the example of the Continuum model, Hydrology and Earth System Sciences, 17(1), 39–62, https://doi.org/10.5194/hess-17-39-2013.

Simons, G.W.H., Koster, R., and Droogers, P. (2020). HiHydroSoil v2.0 - A high resolution soil map of global hydraulic properties. FutureWater Report 213.

Tarpanelli A., Brocca L., Barbetta S., Faruolo M., Lacava T., Moramarco T. (2015) Coupling MODIS and radar altimetry data for discharge estimation in poorly gauged river basin. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8(1), 141-148. http://dx.doi.org/10.1109/JSTARS.2014.2320582

Tarpanelli A., Brocca L., Melone F., Moramarco T., Lacava T., Faruolo M., Pergola N., Tramutoli V. (2013) Toward the estimation of river discharge variations using MODIS data in ungauged basins. Remote Sensing of Environment, 136, 47–55. http://dx.doi.org/10.1016/j.rse.2013.04.010

Tarpanelli A., Santi E., Tourian M.J., Filippucci P., Amarnath G., Brocca L. (2019). Daily river discharge estimates by merging satellite optical sensors and radar altimetry through artificial neural network. IEEE Transactions on Geoscience and Remote Sensing, 57(1), 329-341. https://doi.org/10.1109/TGRS.2018.2854625

Valente, F., David, J. S., and Gash, J. H. C. (1997). Modelling interception loss for two sparse eucalypt and pine forests in central Portugal using reformulated Rutter and Gash analytical models, Journal of Hydrology, 190, 141–162, doi: 10.1016/S0022-1694(96)03066-1.

Verdin, K. L. (2017). Hydrologic Derivatives for Modeling and Analysis—A new global high-resolution database, Hydrologic Derivatives for Modeling and Analysis—A new global high-resolution database, U.S. Geological Survey, Reston, VA, https://doi.org/10.3133/ds1053.

Wagner, W., Hahn, S., Kidd, R., Melzer, T., Bartalis, Z., Hasenauer, S., Figa, J., de Rosnay, P., Jann, A., Schneider, S., Komma, J., Kubu, G., Brugger, K., Aubrecht, C., Zuger, J., Gangkofner, U., Kienberger, S., Brocca, L., Wang, Y., Bloeschl, G., Eitzinger, J., Steinnocher, K., Zeil, P., Rubel, F.: The ASCAT Soil Moisture Product: A Review of its Specifications, Validation Results, and Emerging Applications, Meteorologische Zeitschrift, 22(1), 5-33, https://doi.org/10.1127/0941-2948/2013/0399, 2013.

Xie, P., A. Yatagai, M. Chen, T. Hayasaka, Y. Fukushima, C. Liu, and S. Yang: A gauge-based analysis of daily precipitation over East Asia, J. Hydrometeorol., 8, 607–626, https://doi.org/10.1175/JHM583.1, 2007

Technical Note 2: DTE-Hydrology Datasets User Manual v1.0

1. Introduction

1.1. The DTE Hydrology Evolution Project

The overarching objective of Digital Twin Earth (DTE) Hydrology Evolution project is to prototype a full advanced end-to-end demonstrator of DTE over the full Mediterranean region at high resolution in space and time (targeting 1 km and 1 hour) by extending the successful implementation of Digital Twin Earth Hydrology project that has been carried out over the Po River Valley in northern Italy. Specifically, DTE Hydrology has demonstrated the potential to advance towards an end-to-end reconstruction of the hydrological cycle at high resolution in space and time by an effective combination of state-of-the-art EO data, in situ observations, advanced hydrological and hydraulic modelling, AI and advanced digital platform capabilities. DTE Hydrology Evolution project we aim to further advance the DTE Hydrology concept (see **Figure 1**) through (1) the extension to the full Mediterranean region and hence the preparation to develop a Digital Twin of Hydrological processes and applications at continental and global scales, (2) the exploitation of the dedicated cloud-HPC infrastructure set up in ESRIN, (3) the engagement with a larger community of scientists and stakeholders, (4) the development of an advanced DTE Hydrology open science platform as a community tool offering an interactive access to the EO data, advanced products, in-situ data, and model results, and (5) the development of an attractive and efficient user interface, dashboards, and 3D immersive visualisation tools in an advanced digital environment for scientists, policy makers and citizens. It will be based on state-of-the-art digital platform capabilities and will represent a solid basis for a future ESA supported tool for Polar research and Earth system science.
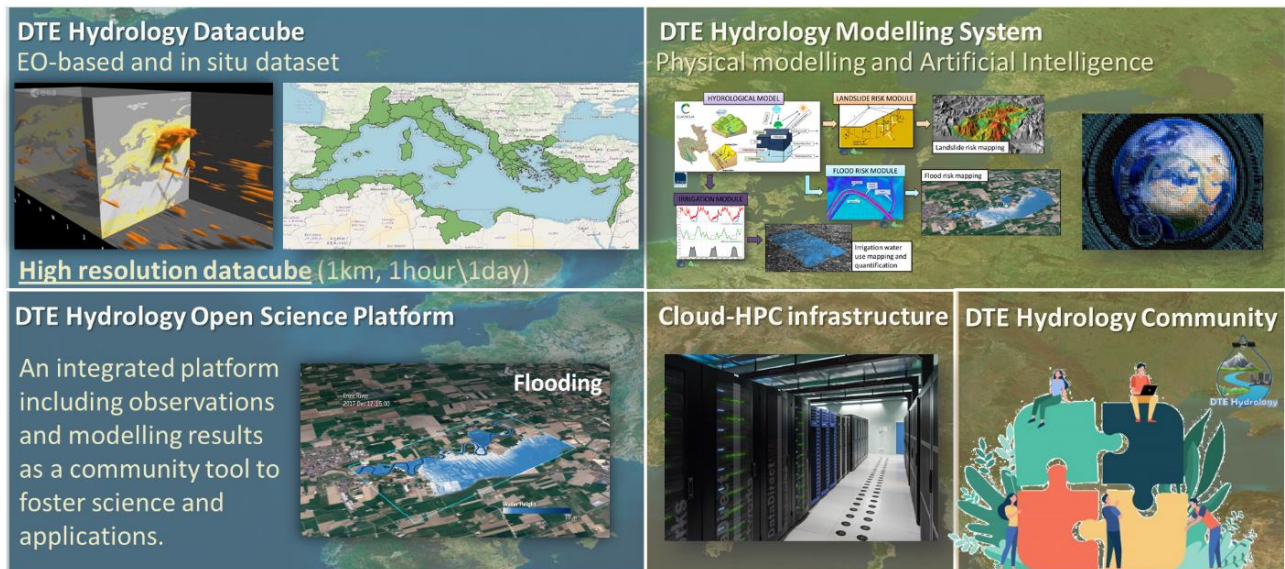
Figure 1: DTE Hydrology Evolution concept.

1.2.   Scope of this Report

This report describes the products contained in the DTE Hydrology Datacube with the indication of the specifications for each of them: time coverage, spatial-temporal resolution, variables, units, … **The DTE Hydrology Datacube will be evolving during the project with the addition of the modelled products, of EO-based discharge data and the further development of the products** to ensure a greater consistency between the different components of the water balance. Therefore, the DTE Hydrology Datacube released here represents the version 1.0 to be updated.

1.3.   Applicable Documents

Not available.

1.4.   Reference Documents

RD-01   DTE Hydrology Evolution Technical Proposal - V1.0

RD-02   Technical Note 1: DTE Hydrology end-to-end description and demonstration plan - V1.0

2. DTE Hydrology Datacube

2.1. Precipitation

The satellite precipitation products developed within DTE Hydrology are organised in netcdf files. One netcdf file was produced for each of the 28 sub-domains of the Mediterranean area. Precipitation data are stored as **3D array with dimensions nlat x nlon x n-days**, with *nlat* the number of cells in the latitudinal direction, *nlon* the number of cells in the longitudinal direction and *n-days* the total number of days within the study period (01/2015-02/2022). The number of pixels considered within each sub-domain is related to its spatial extent.

The file name includes the number of the analysed sub-domain, e.g "med_01_merged_precip.nc". In this first version of the precipitation dataset, only coarse resolution precipitation (CPC and GPM-LR) and rainfall (SM2RAIN-ASCAT) products were used to generate the high resolution precipitation, through the downscaling and merging procedures described in Deliverable 1.1. Each subset precipitation dataset is produced using the 1 km grid files provided by the project.

The following variables are provided in each netcdf file:

- *lon*: contains the longitude of each cell of the considered subdomain. Longitude is given in east degrees, from -180 to +180;
- *lat*: contains the latitude of each cell of the considered subdomain. Latitude is given in north degrees, from -90 to +90;
- *time*: contains the dates of the study period. Dates are given as the number of days from 2000-01-01.
- *precip*: contains the precipitation estimates obtained within DTE Hydrology from the merging of CPC and GPM-LR data. This product is exploited in WP 310 for snow modelling. Precipitation is given in mm/d.
- *rain*: contains the rainfall estimates (precipitation liquid fraction) obtained within DTE Hydrology from the merging of both precipitation (CPC and GPM-LR) and rainfall (SM2RAIN-ASCAT) data by masking the snow period (using the Snow data obtained within the project). Rainfall is given in mm/d.

Missing values in the files and sea pixels are masked with −9999.

In the final version of the dataset, the rainfall data of SM2RAIN-ASCAT will be replaced by the ones from the application of SM2RAIN to high-resolution satellite soil moisture products, in order to increase the spatial accuracy of the final product.

2.2. Evaporation

Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011a; 2011b) datasets are organised in netcdf files. There is one netcdf file per sub-domain and per year, and they are stored as a **3D array with dimensions *n-days* x nlat x nlon** with *n-days* the number of days in the corresponding year, *nlat* the number of cells in the latitudinal direction, and *nlon* the number of cells in the longitudinal direction. The number of cells in the latitudinal and longitudinal direction depends on the spatial extent of the sub-domain.

The file name includes the year and number of the sub-domain. For example, "GLEAM_2015_01.nc" contains daily data for the year 2015 for sub-domain number 1. The same spatial extent was used for each sub-domain as defined within the project. GLEAM was run at 0.01° resolution and then rescaled to 1 km using grid files provided by the project.

The beta version of GLEAM 1km covers the entire study period (2015–2021) for all sub-domains. It uses the bilinearly interpolated forcing data from GLEAM v3.6a, with the exception of precipitation, soil properties and land cover fractions. In the final version of GLEAM 1km, additional forcing variables such as temperature and radiation will be replaced using sources available at higher resolutions following the approach in Rains et al. (2022).

The following 2 variables are available with the beta version of GLEAM 1km:
1. *E* - Actual evaporation [mm/day]
2. *Ei* - Interception loss [mm/day]

The following variables will be included in the final version:
1. *Ep* - Potential evaporation [mm/day]
2. *Eb* - Bare-soil evaporation [mm/day]
3. *Es* - Snow sublimation [mm/day]
4. *Et* - Transpiration [mm/day]
5. *Ew* - Open-water evaporation [mm/day]
6. *S* - Evaporative stress factor [-]
7. *SMroot* - Root-zone soil moisture [m3/m3]
8. *SMsurf* - Surface soil moisture; 0-10 [m3/m3]

Note that by definition: *E = Et + Eb + Ew + Ei + Es* and *S = E/Ep*. Missing values in the files are masked with −999.

2.3. Soil moisture

The RT1 soil moisture product is obtained from the application of the RT1 algorithm (Quast et al., 2019) to 500m resampled Sentinel-1 data in Equi7Grid projection. Results have been reprojected and resampled to the ~1km grid provided by the project using bilinear resampling. The data is organised in zlib-compressed NetCDF files. A single NetCDF is produced for each Sentinel-1 timestamp for the whole of Italy for the time-period from 1.10.2017 to 29.6.2022.

The used file-naming convention is as follows:

<datetime>_<satellite>_<orbit-id>_<ascending/descending>.nc

- <datetime>: the observation date: YYYYMMDD_HHMMSS
- <satellite>: short-name for the Sentinel-1 satellite (e.g. S1A, S1B)
- <orbit-id>: the Sentinel-1 orbit number (e.g. 168, 88, etc.)

- <ascending/descending>: "A" for ascending and "D" for descending orbit

The data is provided both in the original grid (Equi7_grid) at 500m spatial sampling and in the 1km grid provided by the project (4DMED_grid).

The following variables are contained in each NetCDF file.

- **N**: RT1 soil moisture in % saturation scaled to the 0-1 range
  (integer-encoded with a precision of 0.0001)
- 4DMed_grid:
    - **x**: Longitude in east degrees, from -180 to +180;
    - **y**: Latitude in north degrees, from -90to +90;
- Equi7_grid:
    - **x**: x coordinate in the Equi7 EU grid
    - **y**: y coordinate in the Equi7 EU grid

The **N** parameter time series have been scaled for each pixel individually with respect to the 5 and 95 percentiles and clipped to the 0-1 range.
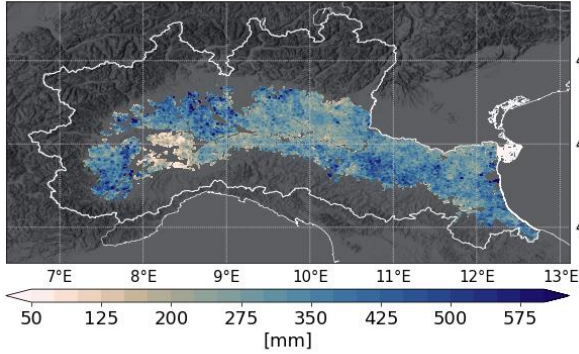
2.4.   Irrigation

The satellite-derived irrigation product contained in the DTE Hydrology Datacube has been developed through the SM-based inversion approach (**Brocca et al., 2018**; **Dari et al., 2020**; **2022**) within the ESA Irrigation+ project. The method relies on the backwards estimate of the total amount of water entering the soil by inverting the soil water balance. Over agricultural areas, irrigation rates can be obtained by removing precipitation from the algorithm's output.

Daily irrigation estimates at 1 km spatial resolution have been produced for the period January 2016 to July 2020 over the agricultural areas of the Po basin; the Corine Land Cover referring to 2018 (CLC2018) has been considered for selecting the agricultural domain of interest. The SM-based inversion approach has been implemented through 1 km surface soil moisture from the RT1 Sentinel-1 data set (**Quast et al., 2019**), rainfall from ERA5-Land at ~9 km spatial resolution (**Muñoz-Sabater et al., 2021**), and potential evapotranspiration data from the GLEAM (Global Land Evaporation Amsterdam Model) v3.5b product at 0.25° spatial resolution (**Martens et al., 2017**; **Miralles et al., 2011**). However, the GLEAM-derived 1 km product developed within this project will replace the currently adopted version in order to benefit from a finer modeling of the evapotranspiration contribution within the SM-based inversion algorithm.

The dataset is delivered on a regular 1 km grid as daily irrigation rates during the period 2016-July 2020. However, given the typical timing of irrigation practices and the associated uncertainties, an at least weekly or bi-weekly aggregation is recommended. As a demonstrative example, the cumulated irrigation amounts during the irrigation seasons (May-September 2016, 2017, 2018, and 2019) retrieved over the Po basin are provided in Figure 2.5.1, while Figure 2.5.2 shows the minimum and maximum values of the seasonal cumulated irrigation rates both distributed and aggregated by region.
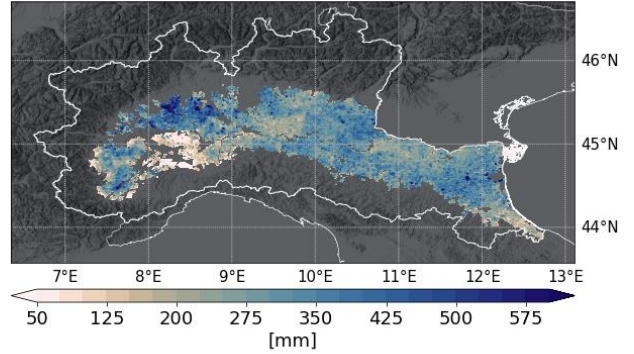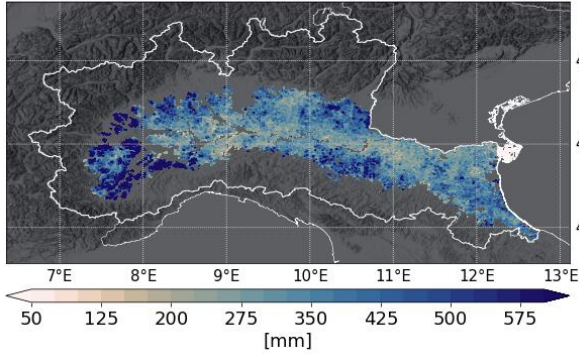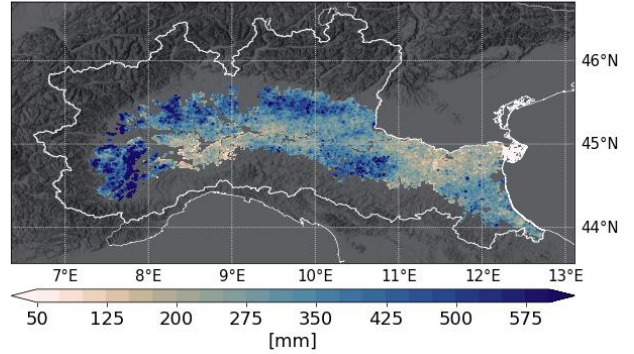
Figure 2.5.1. Cumulated irrigation amounts over the Po basin (Italy) in the periods May-September 2016, 2017, 2018, and 2019 estimated through the SM-based inversion approach.
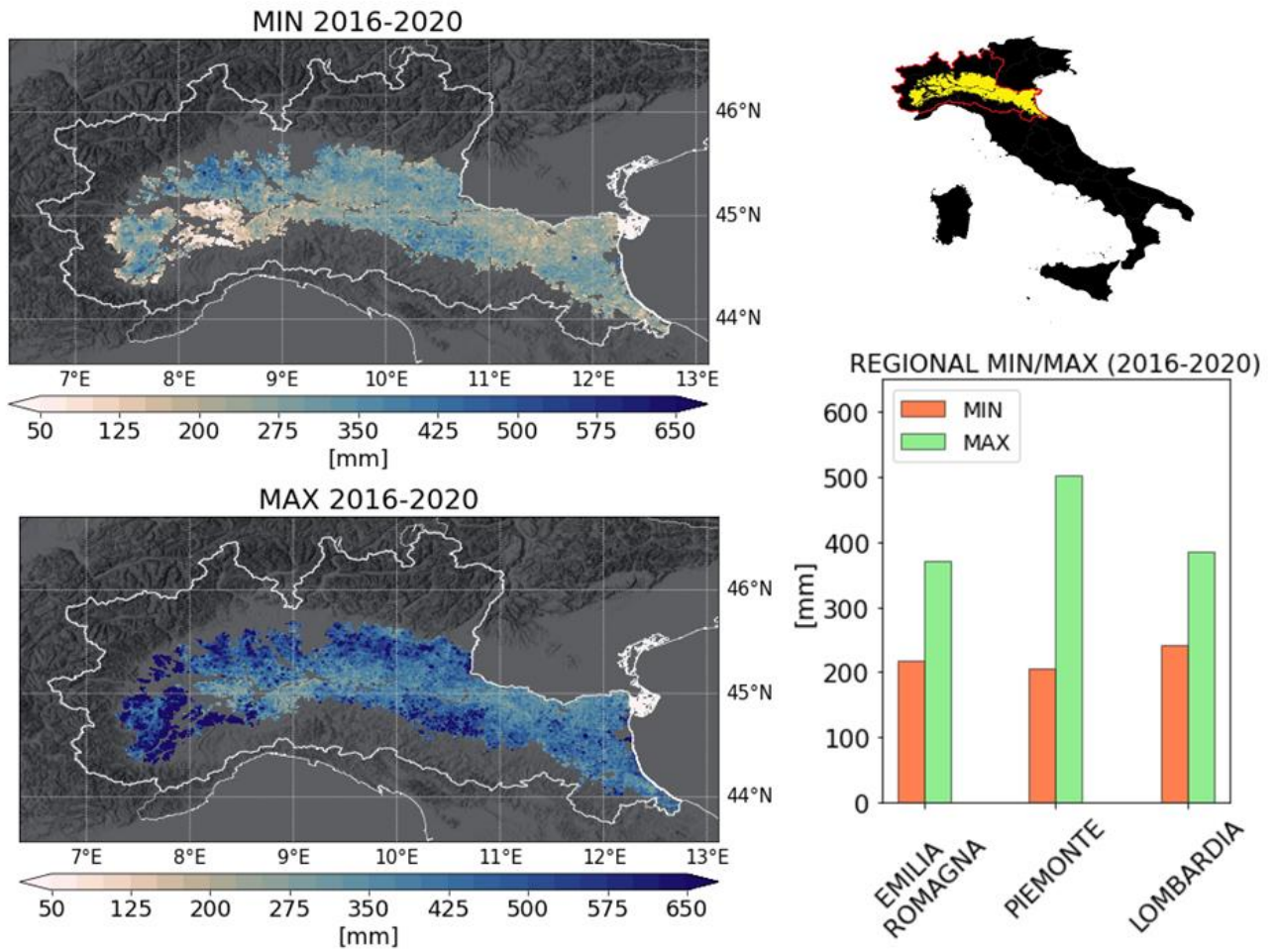
Figure 2.5.2. Minimum and maximum seasonal cumulated irrigation rates, both spatially distributed and aggregated by region information, are provided.

3.   References

Avanzi, F., Ercolani, G., Gabellani, S., Cremonese, E., Pogliotti, P., Filippa, G., Morra di Cella, U., Ratto, S., Stevenin, H., Cauduro, M., and Juglair, S. (2020). Learning about precipitation orographic enhancement from snow-course data improves water-balance modeling, Hydrol. Earth Syst. Sci. Discuss. [preprint], https://doi.org/10.5194/hess-2020-571, in review.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., Wagner, W. (2019). SM2RAIN-ASCAT (2007-2018): global daily satellite rainfall from ASCAT soil moisture. Earth System Science Data, 11, 1583–1601, doi:10.5194/essd-11-1583-2019.

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., Fernández-Prieto, D., 2018. How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. Int. J. Earth Obs. Geoinformation, 73, 752-766.

Dari, J., Brocca, L., Quintana-Seguí, P., Escorihuela, M.J., Stefan, V., Morbidelli, R., 2020. Exploiting high-resolution remote sensing soil moisture to estimate irrigation water amounts over a Mediterranean region. Remote Sens., 12, 2593.

Dari, J., Quintana-Seguí, P., Morbidelli, R., Saltalippi, C., Flammini, A., Giugliarelli, E., Escorihuela, M.J., Stefan, V., Brocca, L., 2022. Irrigation estimates from space: Implementation of different approaches to model the evapotranspiration contribution within a soil-moisture-based inversion algorithm. Agric. Water Manag., 265,107537.

Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., Verhoest, N.E.C., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev. 10, 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Miralles, D.G., Holmes, T.R.H., De Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J., 2011. Global land-surface evaporation estimated from satellite-based observations. Hydrol. Earth Syst. Sci. 15, 453–469. https://doi.org/10.5194/hess-15-453-2011

Miralles, D.G., de Jeu, R.A.M., Gash, J.H., Holmes, T.R.H. and Dolman, A.J. (2011b). Magnitude and variability of land evaporation and its components at the global scale, Hydrology and Earth System Sciences, 15, 967–981, doi: 10.5194/hess-15-967-2011.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications (preprint). Data, Algorithms, and Models. https://doi.org/10.5194/essd-2021-82

Quast, R., Albergel, C., Calvet, J.-C., Wagner, W., 2019. A Generic First-Order Radiative Transfer Modelling Approach for the Inversion of Soil and Vegetation Parameters from Scatterometer Observations. Remote Sens. 11, 285. https://doi.org/10.3390/rs11030285

Rains, D., Trigo, I., Dutra, E., Ermida, S., Ghent, D., Hulsman, P., Gómez-Dans, J., and Miralles, D. G. (2022). High-resolution all-sky land surface temperature and net radiation over Europe. Earth System Science Data Discussions. doi: 10.5194/essd-2022-302.

Technical Note 3: DTE Hydrology modelling experiments, validation and assessment v1.0

1. Introduction

1.1. The DTE Hydrology Evolution Project

The overarching objective of Digital Twin Earth (DTE) Hydrology Evolution project is to prototype a full advanced end-to-end demonstrator of DTE over the full Mediterranean region at high resolution in space and time (targeting 1 km and 1 hour) by extending the successful implementation of Digital Twin Earth Hydrology project that has been carried out over the Po River Valley in northern Italy. Specifically, DTE Hydrology has demonstrated the potential to advance towards an end-to-end reconstruction of the hydrological cycle at high resolution in space and time by an effective combination of state-of-the-art EO data, in situ observations, advanced hydrological and hydraulic modelling, AI and advanced digital platform capabilities. DTE Hydrology Evolution project we aim to further advance the DTE Hydrology concept (see **Figure 1**) through (1) the extension to the full Mediterranean region and hence the preparation to develop a Digital Twin of Hydrological processes and applications at continental and global scales, (2) the exploitation of the dedicated cloud-HPC infrastructure set up in ESRIN, (3) the engagement with a larger community of scientists and stakeholders, (4) the development of an advanced DTE Hydrology open science platform as a community tool offering an interactive access to the EO data, advanced products, in-situ data, and model results, and (5) the development of an attractive and efficient user interface, dashboards, and 3D immersive visualisation tools in an advanced digital environment for scientists, policy makers and citizens. It will be based on state-of-the-art digital platform capabilities and will represent a solid basis for a future ESA supported tool for Polar research and Earth system science.
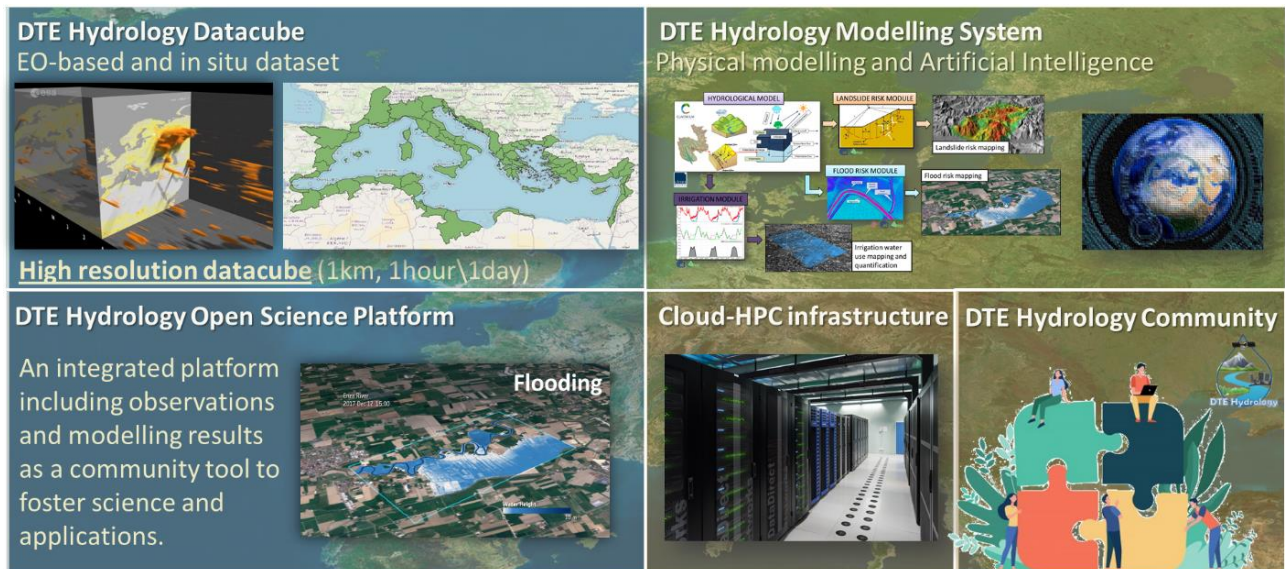
Figure 1: DTE Hydrology Evolution concept.

1.2.   Scope of this Report

This report describes the results of the two modelling simulations obtained with the fully distributed continuous rainfall-runoff model CONTINUUM over the whole Mediterranean Basin and with the modular continuous rainfall-runoff model MISDc for the development of the what-if scenario.

1.3.   Applicable Documents

Not available.

1.4.   Reference Documents

RD-01   DTE Hydrology Evolution Technical Proposal - V1.0

RD-02   Technical Note 1: DTE Hydrology end-to-end description and demonstration plan - V1.0

2. Fully continuous modelling experiments

Hydrological simulations with the distributed model Continuum (Silvestro et al., 2013) were performed over the entire Mediterranean basin, using different combinations of input data as described in the following sections.

2.1. Input data

In addition to the data described in Deliverable 1.1, the following data was collected and used in the distributed modelling experiments.

Hydrological data

Discharge data were extracted from the centralised database set up by EURAC, for use in the two projects DTE Hydrology Evolution and 4DMed Hydrology. Overall, 53 discharge time series were extracted: 19 in the Ebro, 3 in the Herault, 4 in the Medjerda and 27 in the Po basin (**Figure 2.1**).
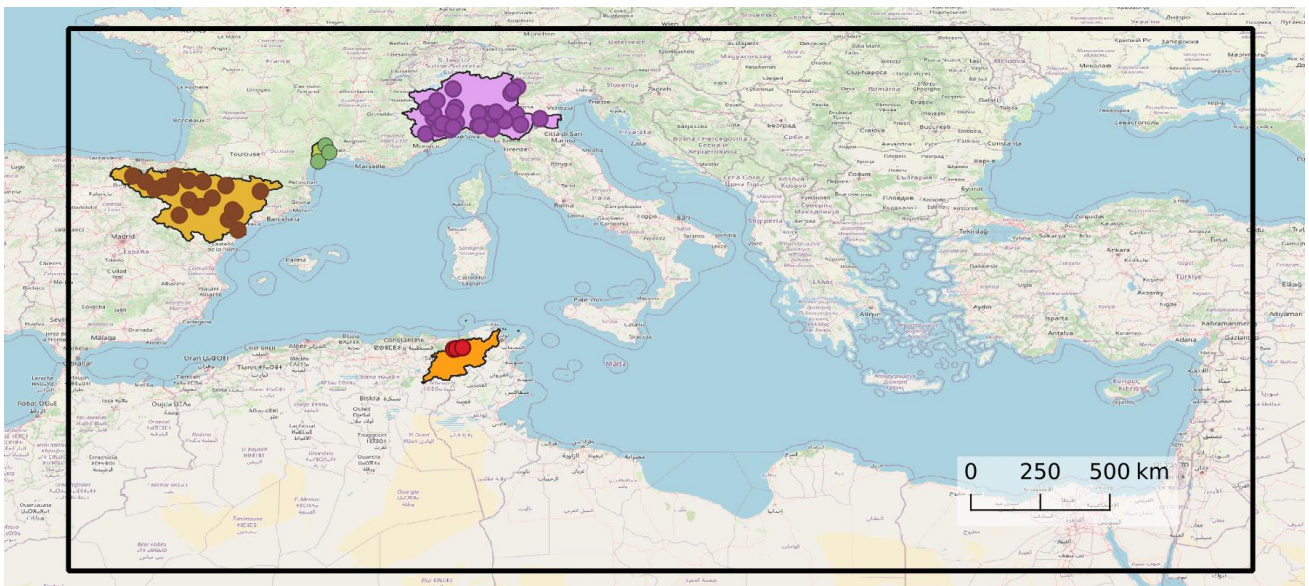

Figure 2.1: Location of the 53 stations with available discharge data used for model calibration.

Meteorological data

The hydrological model Continuum requires, as input, maps of precipitation, air temperature, humidity, wind speed and incoming solar radiation, typically at hourly resolution. The baseline hydrological simulation uses input meteorological data from ERA5 (Hersbach et al., 2020), ECMWF's latest atmospheric reanalysis. ERA5 data is available for the past few decades, it has 1 hour and ~30 km resolution, and it is based on ECMWF IFS Cycle 41r2 (2016). ERA5 is a global reanalysis, hence it covers entirely the simulation domain.

High resolution precipitation data derived from Earth Observations is used in a set of experiments for comparison with those derived by reanalysis data. Here, we used the rainfall data described in Deliverable D1.1, obtained by merging the two precipitation products CPC and GPM-LR with the SM2RAIN-ASCAT rainfall product. Hereinafter, it is referred to as SM2RAIN. Data originally generated at daily resolution for the entire Mediterranean basin was created for some test cases also at hourly resolution and keeping the 1 km spatial resolution, for use in hydrological simulations. Specifically, satellite based precipitation was generated for domain 3, 5, 6, 10 and 26, which include three of the four project case studies, i.e., the Ebro river basin, Herault and Medjerda. Similar satellite-based precipitation dataset was already available for the fourth case study, the Po river basin, from the previous project DTE Hydrology.

In addition, high-resolution (1 km, 1 hour) meteorological data was collected for the Sicilia Region, Italy, for model output comparison during the Medicane Apollo and the riverine flood induced by the Simeto river in the Catania plain, in October 2021. Point observations were initially taken from the Regional Centres of the Italian Department of Civil Protection (DPC) and then spatially interpolated to the entire regional territory. Precipitation fields were estimated with the Modified Conditional Merging (MCM) technique (Bruno et al., 2021), which incorporates precipitation gauges and radar estimates. The MCM is an improvement of the conditional merging (CM) technique proposed by Sinclair and Pegram (2005), which estimates the structure of covariance and the length of spatial correlation at every gauge, taking it from the cumulated radar precipitation fields.

2.2. Modelling experiments

Model calibration

To improve the representation of the hydrological states, Continuum was calibrated in the hydrological domains where observed discharge time series were available as benchmark, i.e., in the Ebro, Herault and Medjerda, adding up to the Po river basin which parameters were calibrated in the DTE Hydrology project. The approach implemented follows the procedure proposed by Alfieri et al. (2022), based on a multi-site calibration procedure that iteratively searches the model parameterization that best matches the available discharge observations over the calibration period at the considered calibration stations, through minimization of a cost function. Hydrological simulations run for the model calibration cover the 2 years starting on 2016-01-01, while the calibration period starts on 2016-07-01, leaving out the initial 6 months for model warm-up. Meteorological input used in the calibration is based on SM2RAIN precipitation and ERA5 for the remaining variables needed by the model. The calibration tool perturbs four scalar parameters related physical hydrological features: field capacity (ct), Curve Number (CN), flow motion coefficient in hillslopes (uh), and maximum water capacity of the aquifer (WTableHbr).

While the calibrated value of uh and WTableHbr are scalar numbers for the entire region of interest, for ct and CN the calibration consists in a rescaling of their default maps to the best value, thus preserving their spatial pattern, which depends on geographic spatial datasets of soil characteristics and land cover. The cost function, based on the Kling-Gupta Efficiency (Gupta et al., 2009), computes an error between simulated and observed discharge time series, weighted with the logarithm of the upstream area, to give higher weight to the downstream stations without neglecting the contribution of the most upstream ones.

The calibration procedure was performed through the implementation of a parallel search algorithm. The algorithm performs an iterative exploration of the 4-dimensional parameter space; the exploration starts with N=40 initial values sampled with a Gaussian Latin Hypercube approach. For each of these N parameter sets, a hydrological simulation is performed over the calibration period, and the cost function is computed to map the error hypersurface. The point that minimises the cost function is used as the centre of the following iteration, until the algorithm converges to an optimal solution.

For parameters which are not kept constant across the selected domain but are defined as the product of a default map times a scaling factor, we have specifically improved the perturbation strategy to improve the sampled parameter maps. Such an approach has the advantage of preserving the spatial distribution of selected parameters, linking it to a physically based quantity, and defining the relation and ranking among such quantities across the domain. However, such an approach is complicated by the need for constraining the calibrated parameter map to physically based constraints derived from the literature, which forces the procedure to a non-linear scaling between the calibration factor and the final (i.e., calibrated) parameter map. Among the chosen calibration parameters of Continuum, such considerations apply to two parameters: the Curve Number (CN) and the Field Capacity (Ct).

The trigonometric arc-tangent function is suitable for applying a scaling of the values in a map within a predefined range. However, an analysis of past applications showed that such function tends to select perturbed parameter maps at the edges of the range (hence with little physical meaning), leaving wide ranges of realistic values undersampled. Our work was focused on addressing such limitations and producing perturbed parameter maps following a more uniform distribution, hence enabling more efficient search of the best values. Results obtained were successful and the current algorithm produces quasi uniform sampling which narrows the sampling range at each iteration and speeds up the convergence to optimal values. Sample results are shown in **Figure 2.2**. In addition, we have adapted the sampling method so that the number of perturbations is set higher in the first iteration (default n=40 samples), which is then reduced by 20% at each subsequent iteration (i.e., 40, 32, 25, 20, 16). Such addition enables a thorough sampling at the initial iterations, yet an efficient use of the computing resources by reducing the number of runs in subsequent iterations. The calibration algorithm now performs a minimum of two iterations (40+32, i.e., a minimum of 72 model runs), where at the end of each iteration after the first it evaluates the improvement in the objective function. The calibration stops when the improvement in the objective function is smaller than a predefined threshold, which default value is set to 1%. In addition, a maximum of 5 iterations was imposed, leading to a maximum of 133 model runs for each calibrated domain.
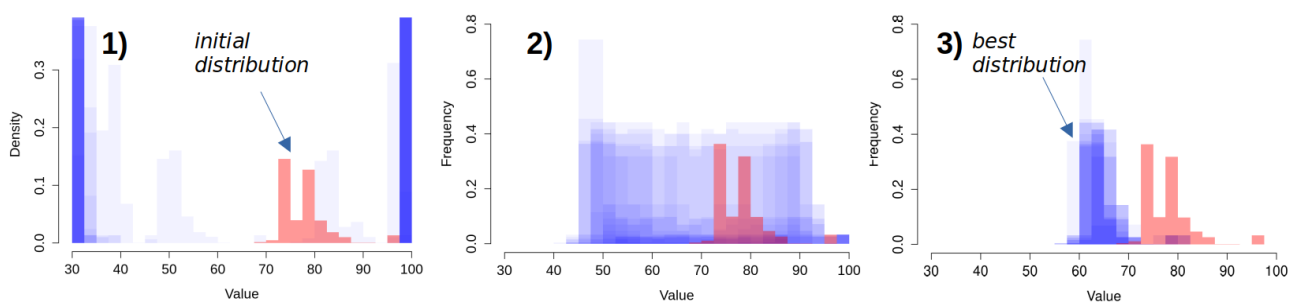
Figure 2.2: Curve Number for a test basin. 1) Unbalanced sampling leading to unrealistic values (in blue). 2) Improved balanced sampling around the initial distribution (iteration #1). 3) Narrower sampling towards the best parameter distribution (iteration #5).

Model runs

The hydrology in the Mediterranean basin was then simulated with the Continuum model, by running it in parallel in each of the 28 domains described in Deliverable D1.1. Model runs, all at 1 km and 1 hour resolution, include the following set of experiments:

- 28 model runs forced by ERA5 for 2015-2021, (one run per domain)
- 6 model runs forced by SM2RAIN (precipitation) and ERA5 (other variables), for the Ebro, Herault, Po, Rhone, Sicilia region and Medjerda.
- 1 model run for the Sicilia region (domain #7) forced by high-resolution ground-based data for the year 2021.

2.3. Results and discussion

Baseline simulations

The experiments for all the 28 domains are performed on the cloud and the machines managed by the Cloudferro company that provides innovative services in computing and in storing and processing big datasets.

The structure of the dte-evo cloud is based on:

- a management and control node (named dtm-control);
- 5 computing nodes (named dtm-compute[0-4]);
- a storage volume of 12 tbytes (mounted in the /home/ folder).

For each node, the available resources consist of 15 virtual cpus (INTEL Xeon Processor) and 16 gbytes of RAM; the total amount of the virtual cpus for the entire cloud structure is equal to 102 with 124 gbytes of RAM. The summary of the cloud resources is shown in the following figure.
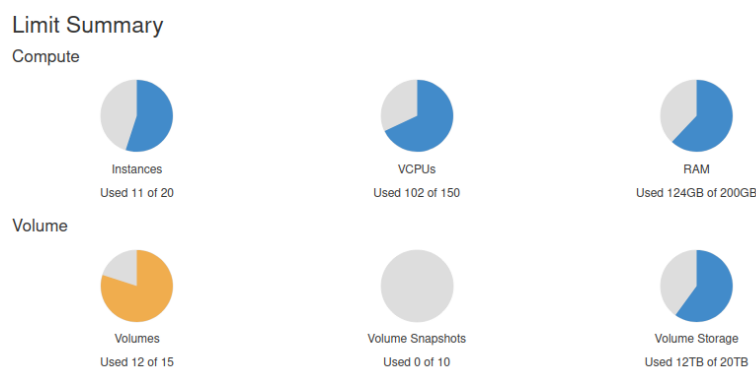

Figure 2.3: Summary of the cloud infrastructure related to the computing and the storage resources.

The operating system is based on the Rocky Linux (version 8.7) that is an open-source enterprise operating system designed to be 100% bug-for-bug compatible with Red Hat Enterprise Linux. The information about the operating system is reported in the next section.

Figure 2.4: Information about the operating system and cpus installed on the nodes.

The computing distribution is managed using the slurm workload manager; for this reason the processes must be submitted using the commands implemented by slurm to obtain better performances in terms of resource availability management. Although the computing resources are available for the single node, the volume storage is shared between the nodes and exploiting this idea it is possible to store and save all the datasets in the same mount point.

Figure 2.5: folder structure of the node.

The libraries, applications and datasets structure are set in the shared storage volume and consist of the following main parts:

- libraries;
- applications and core;
- utils;
- slurm;
- log;
- source volume storage;
- destination volume storage.

In the libraries part are installed all the dependencies needed by the applications, by the cores and by the utils; the application and core part is related to the hydrological model executables and configuration files. The utils location is used for the monitoring memory tools and the transferring datasets tools (to synchronise the datasets from remote machine to the cloud infrastructure).

In the slurm folder are saved the standard output and errors derived by the jobs submitted from the management node.

The log folder is used by the scripts and procedures to save the logging file to check errors or unexpected behaviours.

The source volume represents the geographical and forcing datasets used by the hydrological model; in the destination volume storage are saved all the results produced by the experiments performed in the cloud.

The root directory of the libraries, the datasets and the applications is located in /home/dte/ folder and it is common and shared between the nodes.

Each of the 28 domains (named generally med_dNN domain) is set to run in the related folder (for instance /home/dte/core/med_dNN), get the datasets from the source volume (/home/dte/volume_source/med_dNN) and save the output to the destination volume (/home/dte/volume_destination_core/med_dNN/). Following this settings of the applications and the datasets, each simulation can be run in parallel and independently from the others.

```
# ----------------------------------------------------------------
# script usage generic:
# sbatch -J {job_name} job_manager.slurm slurm_process_dir slurm_process_file slurm_job_dir
# script usage example:
# sbatch -J med_d02 job_manager.slurm /home/dte/core/med_d02/exec/ hmc_tool_running_core.sh /home/dte/slurm/
# ----------------------------------------------------------------

# ----------------------------------------------------------------
# sbatch directives
##SBATCH --job-name=hmc_core                    # Job name
#SBATCH --nodes=1                               # Run all processes on a single node
#SBATCH --ntasks=1                              # Run a single task
#SBATCH --partition=small                       # Partition name
#SBATCH --time=10-00:00:00                      # Time limit days-hrs:min:sec
##SBATCH --cpus-per-task=5                      # Number of CPU cores per task
##SBATCH --mem=5gb                              # Job memory request
#SBATCH --error=execution_job_%x_%j.err         # standard error file
#SBATCH --output=execution_job_%x_%j.out        # standard output file
##SBATCH --nodelist=dtm-compute-[0,2,3,4]       # node list included for executing jobs
##SBATCH --exclude=dtm-compute-[0,1,2,3]        # node list excluded for executing jobs
# ----------------------------------------------------------------
```

Figure 2.6: example of sbatch configuration for submitting the job to the cloud.

The script to submit the job to the cloud infrastructure is stored In the management node; using the folders configuration as previously shown, the command to configure the slurm manager, related to the "job_manager.slurm" script, can be used for starting the experiments using the following command-line:

>> sbatch -J {job_name} job_manager.slurm "slurm_process_dir" "slurm_process_file" "slurm_job_dir"

Or, if it is applied to the domain "med_d02":

>> sbatch -J "med_d02" job_manager.slurm "/home/dte/core/med_d02/exec/" "hmc_tool_running_core.sh" "/home/dte/slurm/"

Once the job is submitted, the job status can be checked by using the "squeue" commands:

```
[rocky@dtm-control ~]$ squeue
  JOBID PARTITION NAME                            USER ST      TIME   TIME_LEFT  CPUS MIN_MEM NODE NODELIST(REASON)
    469 small     med_d17_core                    rocky  R  7-22:14:06 22-01:45:54    1       0    1 dtm-compute-2
    468 small     med_d16_core                    rocky  R  7-22:14:15 22-01:45:45    1       0    1 dtm-compute-2
    467 small     med_d14_core                    rocky  R  7-22:14:27 22-01:45:33    1       0    1 dtm-compute-2
    466 small     med_d08_core                    rocky  R  7-22:14:38 22-01:45:22    1       0    1 dtm-compute-2
    465 small     med_d07_core                    rocky  R  7-22:15:17 22-01:44:43    1       0    1 dtm-compute-2
    464 small     med_d06_core                    rocky  R  7-22:15:22 22-01:44:38    1       0    1 dtm-compute-2
```

Figure 2.7: example of "squeue" command to checking the state of the job.

Moreover, it is possible to check the status of the jobs making the tailing of the standard output and errors stored in the home directory of the node where the job is running:

>> tail -f /var/lib/rocky/hmc_core_med_d02.out

>> tail -f /var/lib/rocky/hmc_core_med_d02.err

Generally the computation time for the experiments can take:

- 1 or 2 days for the smallest domains;
- 8 or 9 days for the biggest domains.

Each job to run the hydrological model uses 600 - 900 Gbytes of RAM. As a consequence, each computing node can be used to run 10 domains at the same time (about 12-13 virtual cpus) to avoid errors or unexpected behaviours in the simulations.

Simulations forced by EO

Continuum hydrological simulations forced by EO data were evaluated at the 53 river gauges where observed discharges were available. The evaluation period is set to the simulation extent except for the first year, which is removed to enable adequate model warm-up, hence it includes the range 2017-2019 for the Po basin and 2017-2021 for the Ebro, Herault and Medjerda basins. Best mean KGE are found in the Herault (KGE_{val}=0.51), Po (KGE_{val}=0.38), while they are poorer in the Ebro (KGE_{val}=0.06) and Medjerda (KGE_{val}=-3.04), the latter with all four stations below the skill threshold of -0.41 (see Knoben et al., 2019). See Table 2.1 for a comprehensive overview of the validation performance. As expected, validation results are generally less skillful than the corresponding ones obtained in calibration, which mean (median) values are KGE_{cal}=0.56(0.53) for the Herault, KGE_{cal}=0.48(0.54) for the Po, KGE_{cal}=0.11(0.21) for the Ebro and KGE_{cal}=0.12(0.33) for the Medjerda. The Medjerda shows the largest difference between calibration and validation results, where the drop in performance is largely attributed to a considerable discharge overestimation, about 400% in comparison with observed values. This is likely attributed to the challenges of precipitation estimates in very dry areas. In addition, visual inspection of the observed time series at Raghay Plaine suggests the presence of issues in the observations, resulting in data after 2015 being flagged as unreliable. This is likely to have negatively affected both the calibration and the validation phase. In the Ebro basin, the sub-optimal skills are mostly attributed to the significant reservoir impact on downstream flows and the challenges in modelling the unknown operating rules of flow release, an issue already reported by Arnal et al. (2019) in the calibration of the hydrological model underpinning the European Flood Awareness System.

On the other hand, stations in all four basins show skillful correlations, with less than 10% of stations (i.e., 5 out of 53) having correlation values smaller than 0.4. By pooling together all station results, median scores are KGE_{val}=0.33, r_{val}=0.62, Bias rate$_{val}$=0.77 and CVrate$_{val}$=0.77. This indicates generally good performance with very good model correlation and moderate underestimation (-23%) of observed values. Examples of simulated versus observed discharges in the four basins are shown in the following 4 figures (3 sample stations for each basin), for the entire validation period.

Table 2.1: Validation performance of simulated discharges in the four case studies, with Continuum forced by EO meteorological input over 2017-2021 (2017-2019 for the Po Basin). Benchmark data is taken from observed discharges in the available river gauges.

	UpsArea [km²]	KGE	r	Bias rate	CV rate
Po					
Gaiola	515	0.33	0.60	0.47	1.00

	UpsArea [km^2]	KGE	r	Bias rate	CV rate
Cassine	1,364	0.18	0.51	1.17	0.36
Carignano	3,649	0.50	0.71	0.60	1.08
Torino Murazzi	6,134	0.69	0.77	0.81	1.09
Lanzo	541	0.46	0.60	0.65	0.87
Ponte Alto	1,077	0.37	0.62	0.71	0.60
S Secondo	1,422	0.35	0.57	0.60	0.71
Ostia Parmense	387	0.32	0.52	0.53	0.89
Candoglia	1,480	0.44	0.80	0.53	0.76
Quinto Vercellese Cervo	178	-0.11	0.47	0.06	0.73
Palestro	2,168	0.44	0.69	0.66	0.67
Casale Monferrato	12,882	0.62	0.79	1.10	0.69
Ponte Verdi	485	0.12	0.47	0.32	0.82
Valsigiara	192	0.43	0.44	0.92	0.92
Tavagnasco	3,096	0.62	0.88	0.66	0.88
Spessa	35,976	0.63	0.85	0.67	1.05
Piacenza	39,195	0.60	0.86	0.64	1.08
Pontelagoscuro	67,487	0.61	0.86	0.72	1.22
Cremona	47,616	0.60	0.85	0.67	1.15
Borgoforte	59,169	0.60	0.85	0.68	1.20
Farigliano	1,379	0.43	0.60	0.64	1.20
Alba Q A	3,180	0.53	0.62	0.72	1.04
Salsominore	171	0.20	0.34	0.89	1.44
Ragoli	504	0.08	0.24	0.95	0.49
Ponte dei Tedeschi	361	0.18	0.52	0.41	0.70
Cimego	233	-0.04	0.11	1.20	0.50
Piana Crixia	229	0.01	0.45	0.50	0.35
Cartosio	180	0.42	0.57	0.70	0.74
Mean	**10,402**	**0.38**	**0.61**	**0.68**	**0.87**
SD	**19,603**	**0.22**	**0.20**	**0.24**	**0.27**
Ebro					
Río Ega en Andosilla	1,457	0.44	0.73	1.30	0.61
Río Ebro en Zaragoza	40,738	0.20	0.80	1.63	0.55
Río Cinca en Fraga	9,618	0.49	0.81	1.27	0.61
Río Segre en Serós	12,848	-0.53	0.66	2.49	1.01
Río Ebro en Tortosa	84,476	-0.08	0.77	1.96	0.57
Río Iregua en Islallana	561	0.50	0.56	0.89	1.20
Río Ara en Boltaña	652	0.48	0.66	0.68	0.78
Río Subordán en	345	0.50	0.66	0.71	0.78

|	UpsArea [km^2]	KGE	r	Bias rate	CV rate
Javierregay					
Río Irati en Liédena	1,584	0.31	0.62	1.21	0.47
Río Arga en Etxauri	1,635	0.31	0.66	1.05	0.40
Río Gallego en Zaragoza	3,963	-0.18	0.72	1.99	0.42
Río Segre en Organya	2,331	0.04	0.71	1.92	0.95
Río Ebro en Mendavia	11,936	-0.10	0.71	1.95	0.54
Río Ebro en Palazuelos	4,535	-0.10	0.62	1.91	0.51
Río Matarraña en Nonaspe	947	-0.19	0.15	0.26	0.61
Río Flumen en Barbués	565	-0.01	0.41	0.24	1.31
Río Arba en Tauste	2,161	-0.03	0.50	0.23	1.46
Río Jalón en Calatayud	7,330	-0.77	-0.14	0.47	2.25
Río Tirón en Haro	1,530	-0.20	0.70	2.13	0.75
Mean	**9,959**	**0.06**	**0.60**	**1.28**	**0.83**
SD	**20,308**	**0.36**	**0.23**	**0.72**	**0.46**
Herault					
Herault Agde	2,511	0.50	0.60	1.08	0.70
Herault Laroque	845	0.46	0.54	0.84	0.77
Lez	229	0.57	0.57	1.03	1.00
Mean	**1,195**	**0.51**	**0.57**	**0.98**	**0.82**
SD	**1,181**	**0.06**	**0.03**	**0.13**	**0.16**
Medjerda					
Ghardimaou	1,501	-0.81	0.40	2.61	0.44
Raghay Plaine	1,859	-9.72	0.50	11.70	1.34
Jendouba	2,442	-0.65	0.68	2.59	0.70
Raghay Sup	2,283	-0.99	0.40	2.88	0.70
Mean	**1,451**	**-3.04**	**0.50**	**4.94**	**0.79**
SD	**1,040**	**4.45**	**0.13**	**4.51**	**0.39**
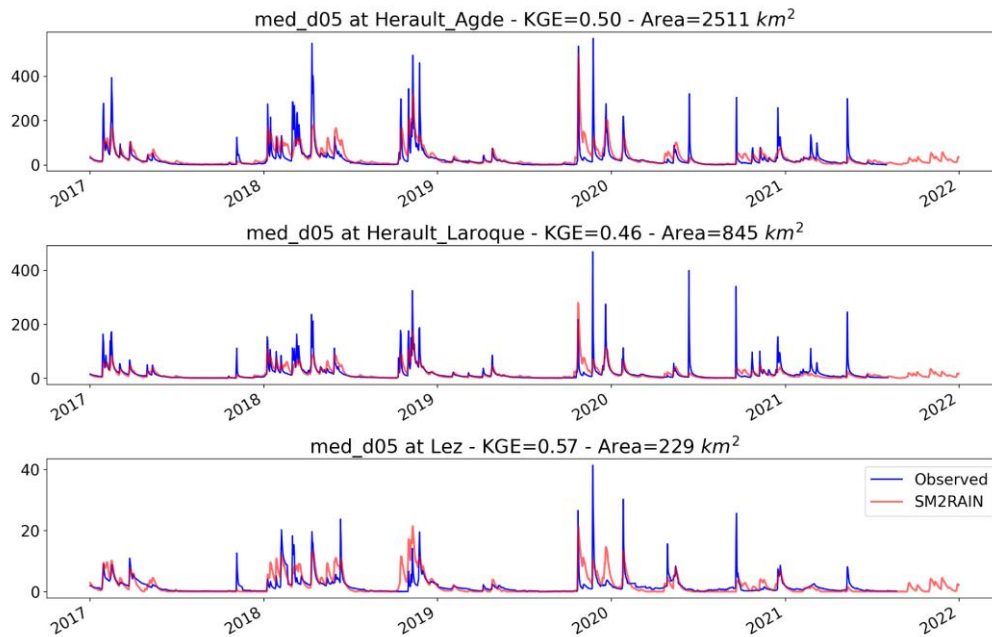
Figure 2.3: Modeled versus observed river discharges at 3 sample stations in the Herault river basin.
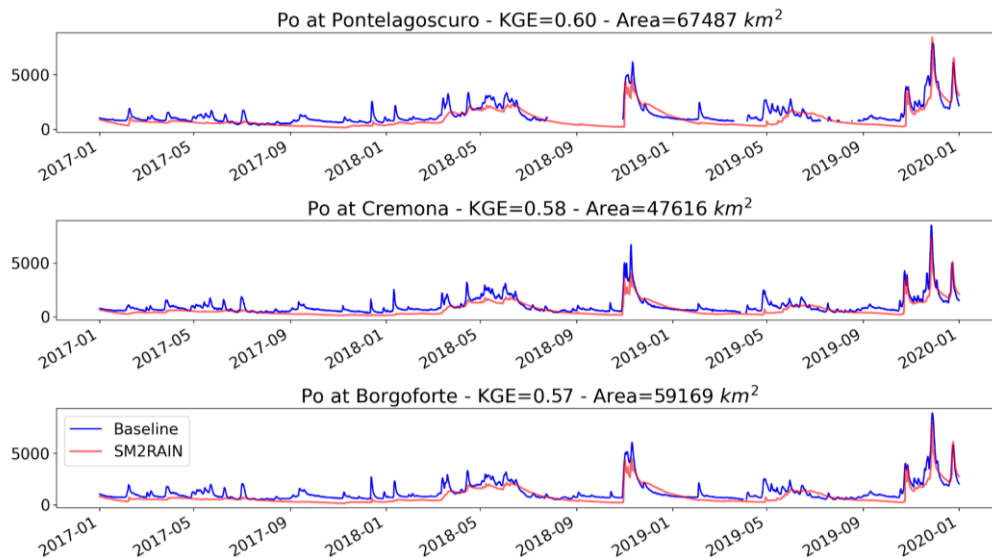

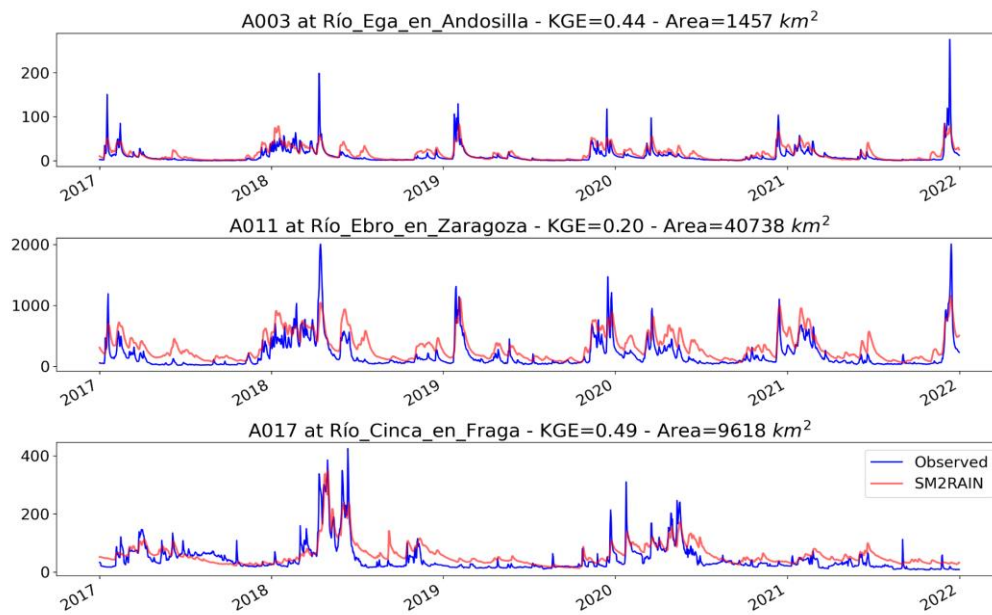***Figure 2.4***: Modeled versus observed river discharges at 3 sample stations in the Po river basin.

Figure 2.5: Modeled versus observed river discharges at 3 sample stations in the Ebro river basin.
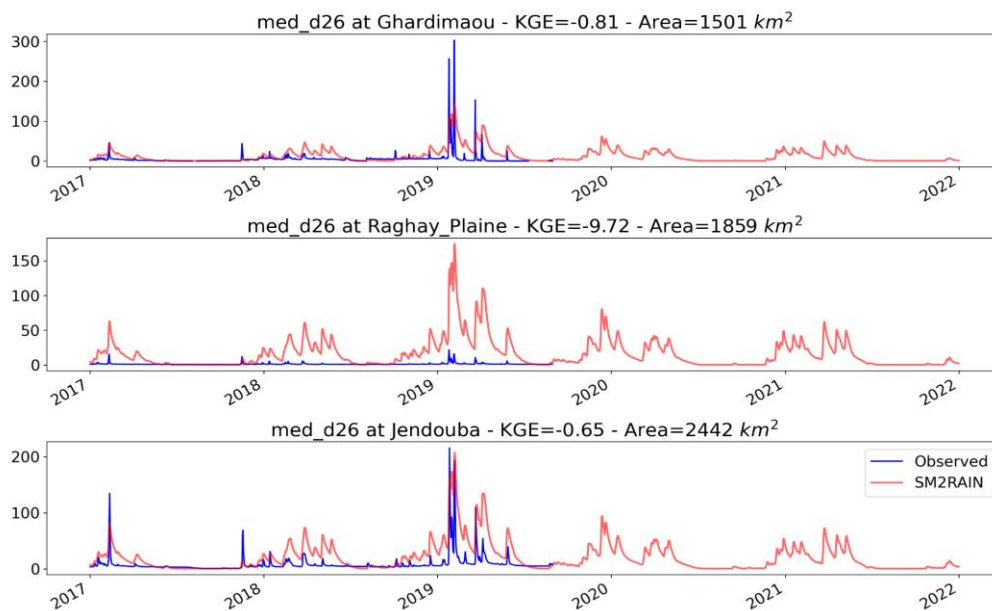

Figure 2.6: Modeled versus observed river discharges at 3 sample stations in the Medjerda river basin.

Comparison of ERA5 with EO hydrological simulations in Sicily

For the Sicilia region, Italy (domain #10), discharge observations are not available for evaluating the simulation performances. However, high resolution ground-based observations collected by the Italian Department of Civil Protection (see Sect. 2.1) were used to force the same hydrological model setup used for the ERA5 and SM2RAIN simulations. The advantage of having such high resolution benchmark simulation (hereafter DPC-MCM) is twofold. First, it enables a fair evaluation between ERA5 and SM2RAIN over the region, considering that the model parameters were not calibrated on either of the two precipitation products. Second, the evaluation can be done on the entire regional river network rather than just at points, being based on a spatially distributed model output. The goodness of the model parameterization used in Continuum was not evaluated on the output discharges, due to the lack of observations. However, the simulation forced by high-resolution meteorological data was then used as input to a bidimensional hydraulic model for the case of the Medicane Apollo 2021. Output inundation areas skillfully reproduced the observed inundation areas in the Catania Plain along the Simeto river, as also shown in the deliverables of WP400 of the project.

A summary of the performance of the hydrological model simulations for the 5-year period 2017-2021 is shown in Figure 2.7 (Figure 2.8) for the ERA5- (SM2RAIN-) driven simulations. The 166 points in the figures are representative of the outlet of each river reach and were chosen as a subset of the entire set of points along the modelled river network to reduce the auto-correlation effect along the same reaches. On average SM2RAIN simulations outperform those forced by ERA5 only, with mean KGE at the 166 stations of $KGE_{SM2RAIN}$=0.44 versus KGE_{ERA5}=0.28. Looking more in detail at the three decomposition terms of the KGE, one can note that the increase in performance of SM2RAIN is moderate, but persistent in all three terms (+0.05 in correlation, +0.03 in bias rate and +0.06 in CV rate). It is interesting to note that while ERA5 on average has a positive bias (bias_rate$_{ERA5}$=1.30), i.e., 30% overestimation, SM2RAIN simulations have similar but negative bias (bias_rate$_{SM2RAIN}$=0.73), i.e., 27% underestimation.

Sample comparison of the ERA5- (SM2RAIN-) driven simulations with the benchmark simulation MCM is shown in Figure 2.9 (Figure 2.10). See point locations in 2.11.
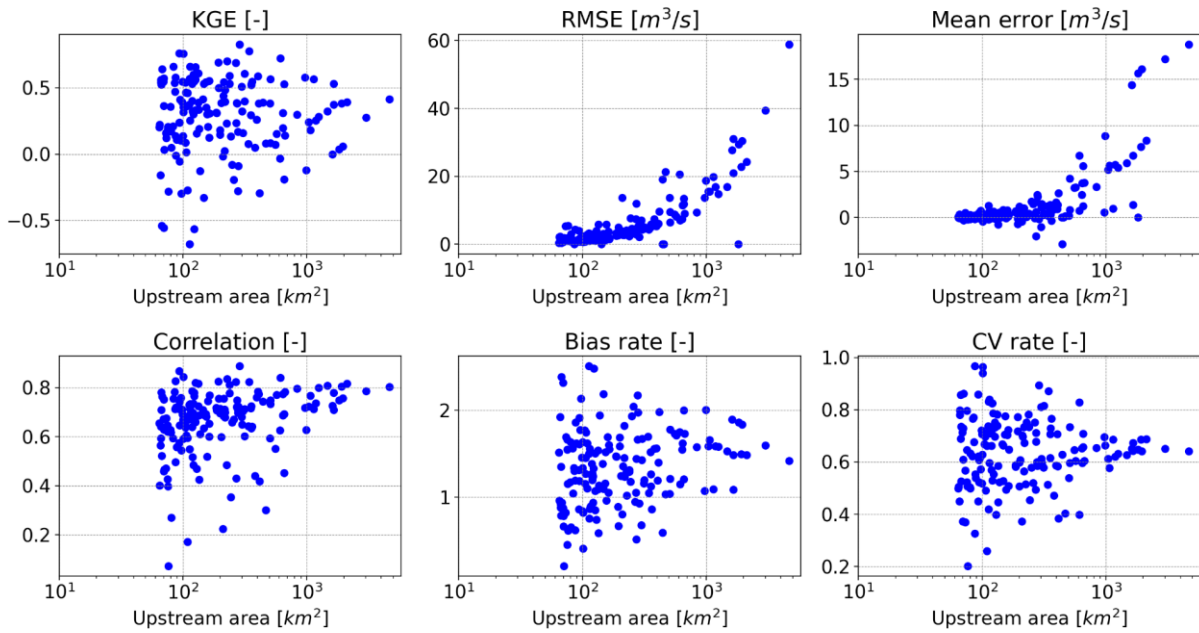
Figure 2.7: Skills of discharge time series simulated with Continuum forced with ERA5 precipitation versus the original MCM 1km, 1h resolution. Scores are calculated for all 166 river reaches of the modelled river network in the Sicily region, 5 years of hourly data starting on 1 January 2017.
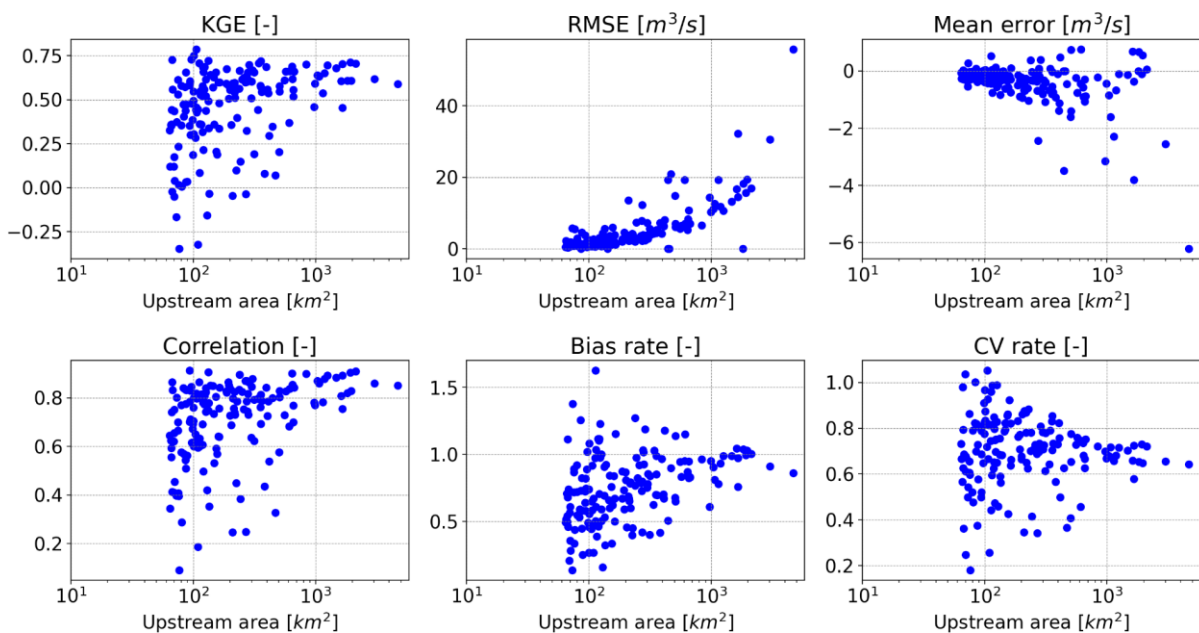
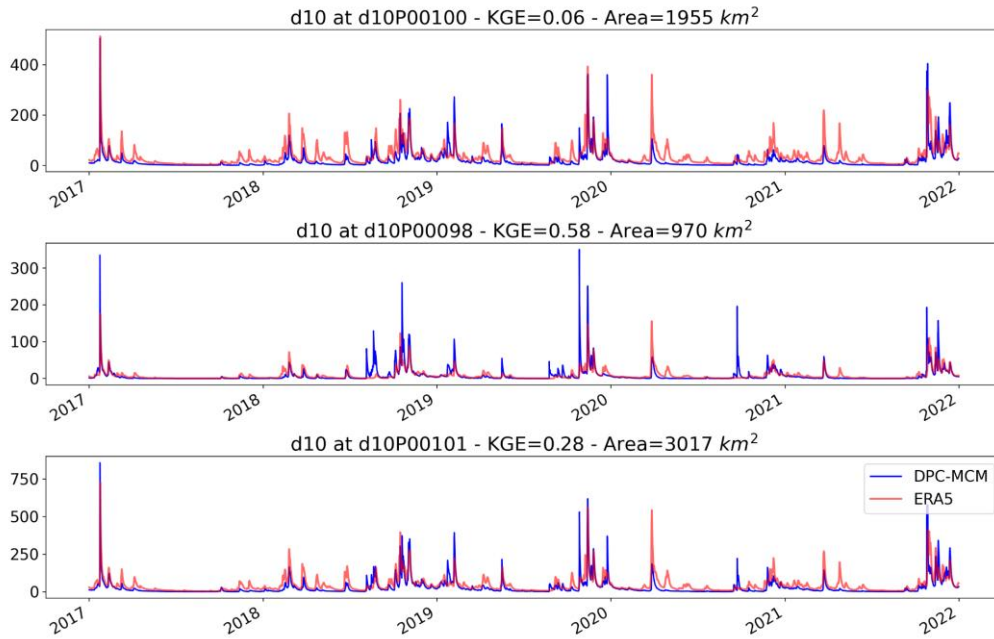**Figure 2.8**: As Figure 2.7 using Continuum forced with SM2RAIN precipitation.

Figure 2.9: Comparison of simulated discharge produced by Continuum at three sample stations using ERA5 precipitation input versus the original MCM 1km 1h resolution (blue). Station locations are shown in Figure 2.11.
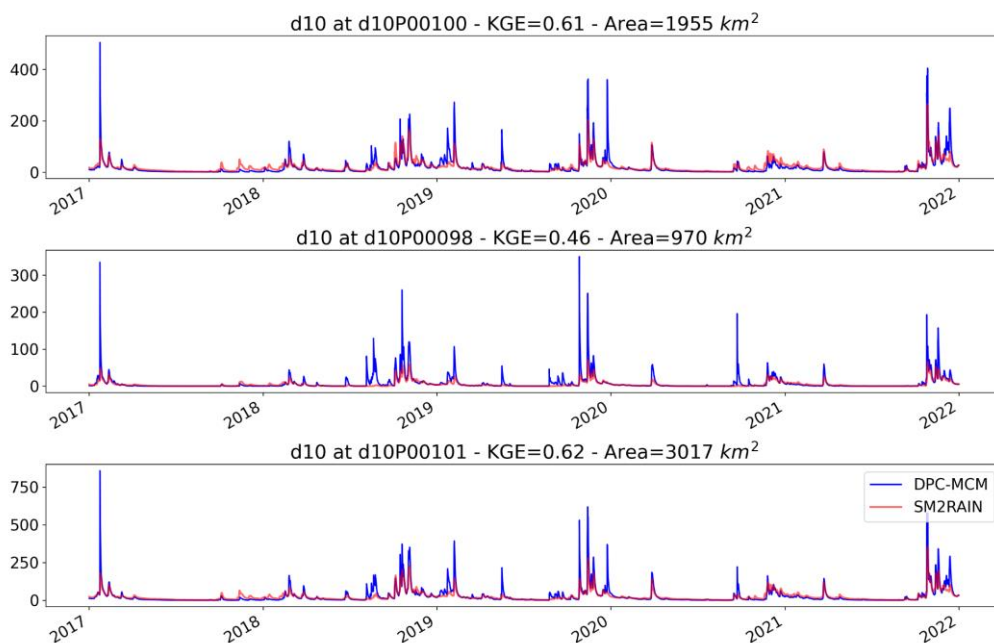
Figure 2.10: Comparison of simulated discharge at the same stations of Figure 2.9 using SM2RAIN precipitation versus the original MCM 1km 1h resolution (blue). Station locations are shown in Figure 2.11.
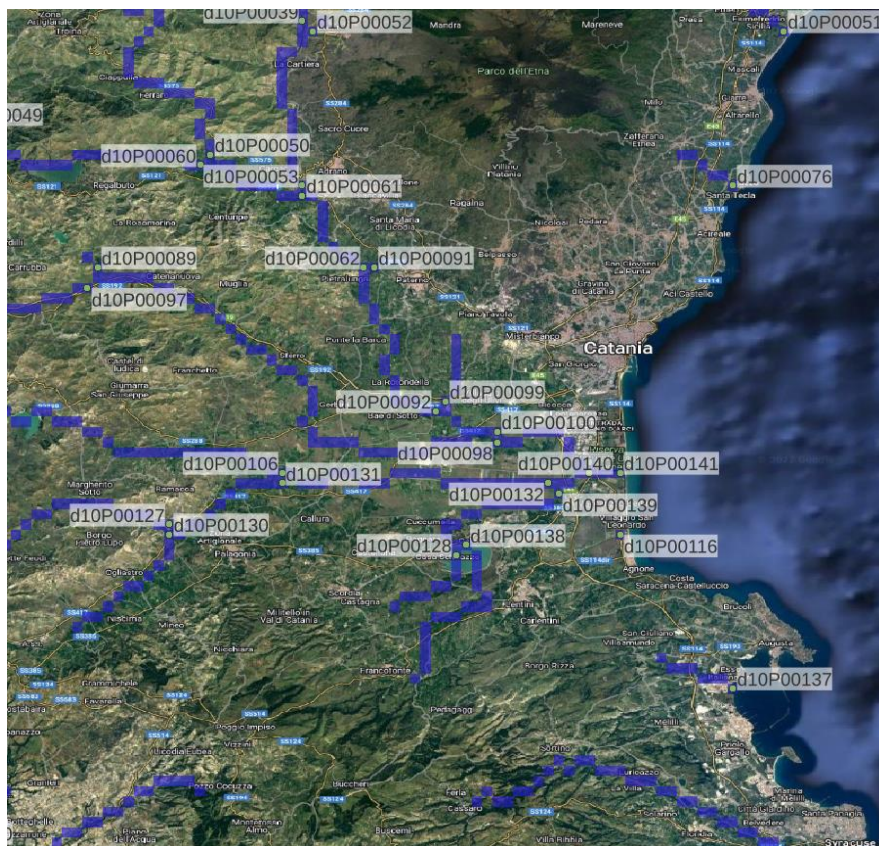

Figure 2.11: 1 km resolution modelled river network in Continuum over the Catania plain, Italy. Labels show the identifier of the outlet of each river reach.

3. Modular modelling experiments

Hydrological simulations with the modular modelling "Modello Idrologico SemiDistribuito in continuo, MISDc model (Brocca et al., 2011) were performed over the Po basin to build the what-if scenarios for flood risk assessment and water resources management. For the purpose, different input data, hydrological model configurations and validation data were used. Details are given in the following paragraphs.

3.1.   Input data

Here a brief description of the input data used to build both the what-if scenarios for flood risk assessment and water resources management is reported. Overall, the MISDc hydrological model requires both hydrological and meteorological data as input and they are sufficient if the target of the analysis is the flood risk assessment. Additional information on water uses have to be considered for the purpose of water resources management.

Hydrological data

River discharge data were extracted from the centralised database set up by EURAC, for use in the two projects DTE Hydrology Evolution and 4DMed Hydrology. Overall, 19 discharge time series were extracted in the Po basin as illustrated in Figure 3.1.
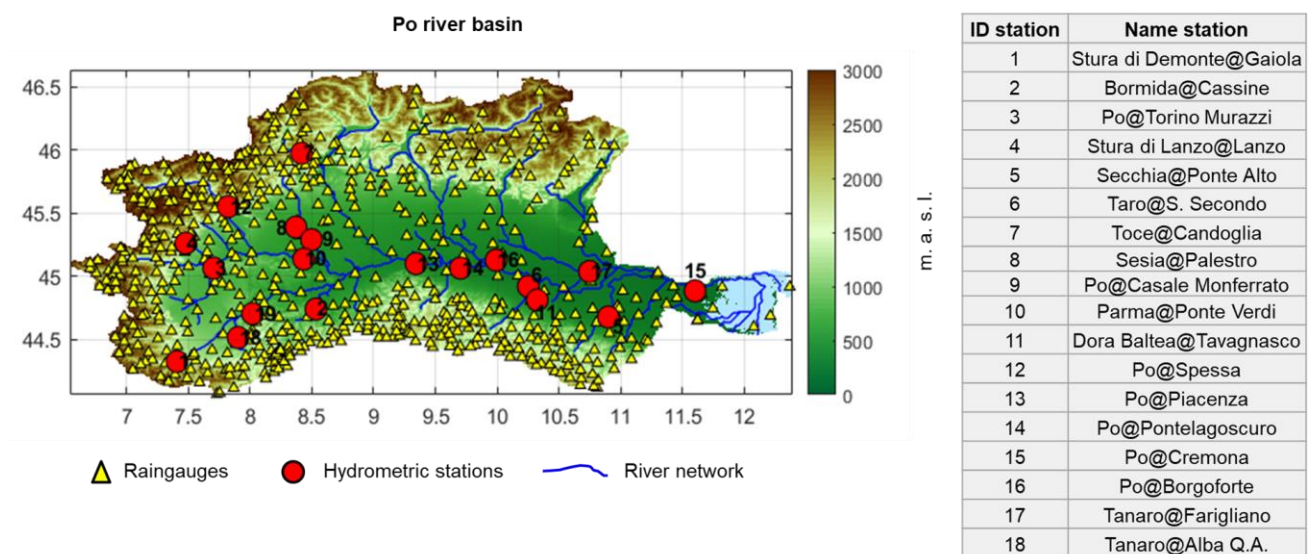

ID station	Name station
1	Stura di Demonte@Gaiola
2	Bormida@Cassine
3	Po@Torino Murazzi
4	Stura di Lanzo@Lanzo
5	Secchia@Ponte Alto
6	Taro@S. Secondo
7	Toce@Candoglia
8	Sesia@Palestro
9	Po@Casale Monferrato
10	Parma@Ponte Verdi
11	Dora Baltea@Tavagnasco
12	Po@Spessa
13	Po@Piacenza
14	Po@Pontelagoscuro
15	Po@Cremona
16	Po@Borgoforte
17	Tanaro@Farigliano
18	Tanaro@Alba Q.A.

Figure 3.1: Location of the 18 stations with available discharge data used for model calibration over the Po river basin.

Meteorological data

The modular hydrological model MISDc requires, as input, time series of precipitation and air temperature.

For precipitation, here we used high resolution precipitation data derived from Earth Observations described in Deliverable D1.1, obtained by merging two precipitation products CPC and GPM-LR with the SM2RAIN-ASCAT rainfall product. Hereinafter, it is referred to as SM2RAIN. Data were generated at 1 km /hourly spatial/temporal resolution for the period 2015-2021.

For air temperature we used the high resolution data developed within the ESA 4DMED project, i.e., daily /1 km data downscaled from ERA5-Land (European Centre for Medium-Range Weather Forecasts, ECMWF, Reanalysis 5th Generation). First, hourly temperature data were downloaded and averaged at daily frequency for the whole basin, then the dataset was downscaled to 1 km spatial resolution by using an high spatial resolution climatology derived by CHELSA (Climatologies at high resolution for the earth's land surface areas; Karger et al., 2017). For further information regarding the downscaling procedure, see "4DMED-Hydrology Deliverable 3.2: Algorithm Theoretical Basis Document (ATBD)".

Additional data

A set of additional data were used as input or to validate the MISDc model results. Specifically, the high-resolution data of Snow Water Equivalent, the RT1 Sentinel-1 soil moisture, the GLEAM surface and root zone soil moisture and the GLEAM Evapotranspiration described in Deliverable D1.1 were collected over the Po basin. Moreover, the high-resolution satellite based irrigation dataset (Dari et al., 2023) was collected.

Additionally, data related to water uses over the study area have been collected from the Water Balance Plan for the Po River District (PdgPo, 2010 and 2016) prepared by the authority the Po Authority Basin (Autorità di Bacino del Fiume Po, AdBPO) and used as input into the MISDc modular model to simulate the water uses. Table 3.1 summarises the civil, Industrial and agricultural water uses for the years 2010 and 2018 and the percentage values with respect to the total water uses. Looking at the table it can be noted that about $20*10^9$ m^3 of water is used for human activities, and among them, about 75-80% of the water is used for agricultural purposes.

Table 3.1: Water uses in the Po river basin as in the Water Balance Plan for the Po River District (PdgPo) for the year 2010 and 2016.

Water uses	Volume ($10^9\,m^3$/y)* *PdgPo2010	Volume ($10^9\,m^3$/y)* *PdgPo2018	% vs Tota *PdgPo2010l	% vs Tota *PdgPo2018l
Civil	2.5	2.8	12.20	14.58
Industrial	1.5	1.98	7.32	10.31
Agricultural	16.5	14.4	80.49	75.00
Total	**20.5**	**19.2**	100.00	100.00

The PdgPo2018 provides also the monthly average river discharge withdrawn from the Po basin upstream some sections along the Po river. By merging this information, reported in Table 3.2, with the one in Table 3.1 is possible to reconstruct the monthly water volume used for civil and industrial purposes over the study area.

Table 3.2: Monthly average river discharge (in m³/s) withdrawn from the Po basin upstream the section of Casale Monferrato, Cremona, Pontelagoscuro.

	Casale Monferrato (m³/s)	**Cremona (m³/s)**	**Pontelagoscuro (m³/s)**
Jan	24	104	111
Feb	23	103	110
Mar	41	126	131
Apr	134	297	359
May	165	362	476
Jun	322	807	1361
Jul	294	756	1306
Aug	160	363	544
Sept	119	289	402
Oct	43	135	143
Nov	30	117	123
Dec	26	110	117

3.2. Modelling experiments

The MISDc hydrological model is a semidistributed conceptual hydrological model characterised by a flexible modular structure. The modules implemented so far for MISDc are six and they are described in the following (see Figure 3.2):

- **snow module** able to divide the total precipitation into a liquid and solid component. By using as input the precipitation and the air temperature data it allows to estimate the snow water equivalent and the snowpack;
- **soil module** able to simulate the main processes acting into the soil (e.g., infiltration, evaporation, percolation) and consequently it provides as output, the evolution in time of the water content into the

soil and the runoff. It comprises two-layer soils to simulate the surface and the root zone soil moisture and runoff;

- **routing module** able to transfer the runoff through the basins and the river network. It provides as output river discharge time series at the closure section of each sub-basin as well as at the outlet and inner sections of the entire basin;
- **reservoir module** able to mimic the releases downstream the reservoirs according some operation rules defined as function of the storage volume (Burek et al., 2020);
- **irrigation module** able to simulate the irrigation practice. It is activated during irrigation periods when the soil moisture on the first layer is lower than a certain threshold and it is stopped when the field capacity is reached;
- **water uses module** that simulates the water extracted for civil and industrial purposes.
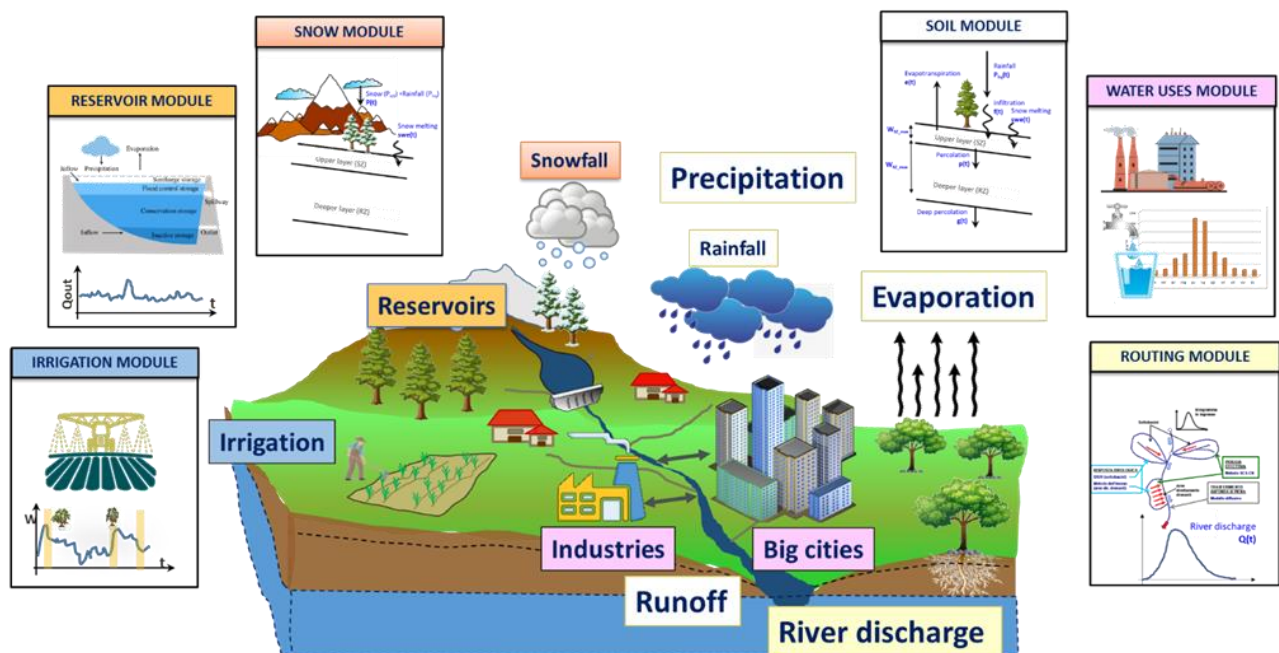
Figure 3.2: Structure of the modular approach MISDc applied to build the what-if scenario for flood and water resources management.

Depending on the target of the analysis, the six modules can be added or removed to adapt the structure of the MISDc model in reproducing flood or drought events. In particular, to simulate floods three modules are strictly needed: the snow, the soil and the routing modules (see Figure 3.2). Conversely, to correctly represent the water availability and the water demand during summer periods a major number of modules have to be considered. In addition to the three modules needed for floods, the water uses, the reservoir and the irrigation module have to be included in the structure of the MISDc model for the water resources management. Hereinafter we will refer to the two configurations of the modular MISDc as "Modular modelling for flood risk assessment" or "Modular modelling for water resources management".

Model calibration

The MISDc model includes 12 model parameters: the initial condition of soil water content, $W(t_0)$; the maximum water capacity of the soil layer, W_{max} (i.e. field capacity term); the initial abstraction coefficient, λ; the parameter of the relationship between the saturation degree and the soil retention, a; the correction factor for actual evapotranspiration b; the saturated hydraulic conductivity Ks; the exponent of drainage component, m; the fraction of drainage that transforms into subsurface runoff, θ; the degree-day coefficients, Cm-snowpack and Cm-glaciers for the melting process of snowpack and glaciers; the coefficient of lag–area relationship, η; the celerity C and the diffusivity, D. Each model parameter can vary within a range of admissible values derived by other applications of MISDc over many catchments worldwide (Camici et al., 2018, Massari et al., 2015, Masseroni et al., 2017). The calibration process consists in adopting a standard gradient based automatic optimization method implemented in the MATLAB software package ('fmincon' function; MATLAB R2016b, The MathWorks, Inc., Natick, Massachusetts, United States). This algorithm is particularly suitable and efficient for a limited number of model parameters and enables one to maximize an objective function. In this case, the objective function is the difference between the unity and the modified Kling-Gupta efficiency statistic dimensionless, KGE, proposed by Gupta et al. (2009).

Both the "MISDc for flood simulation" and "MISDc for water resources management" module structures were calibrated against observed river discharge data at several sections along the Po river for the period 2015-2021. For both the configurations, the MISDc model has been calibrated through a sequential calibration (De Lavenne et al., 2019) over 11 out 19 sections, whereas the remaining 8 stations have been used for validation.

3.3. Results and discussion

Hereinafter the results related to the modular modelling adopted for the flood risk assessment and for the water resources management are presented.

Modular modelling for flood risk assessment

Figure 3.3 illustrates the river discharge time series simulated by the MISDc for some stations along the Po river, whereas the results for all the sections are summarised in Table 3.3.
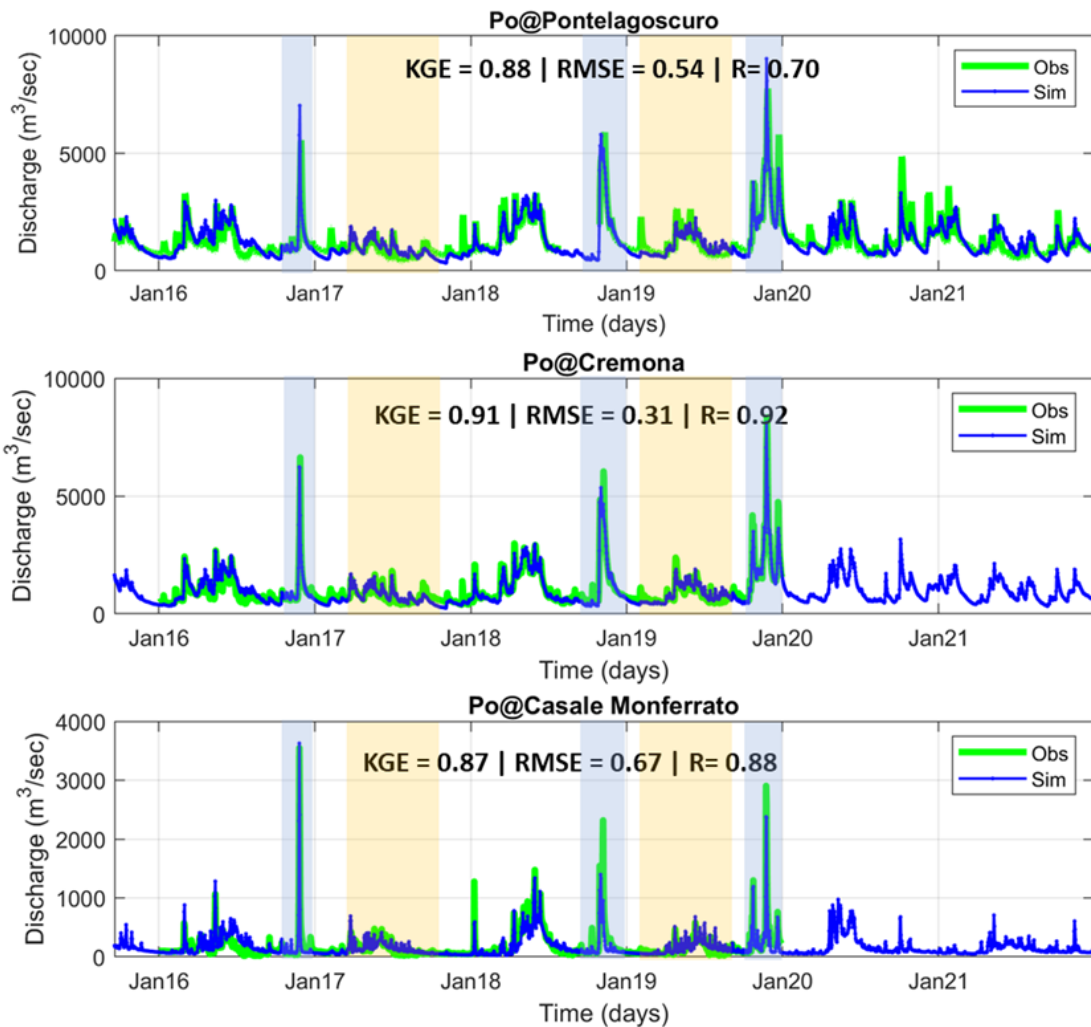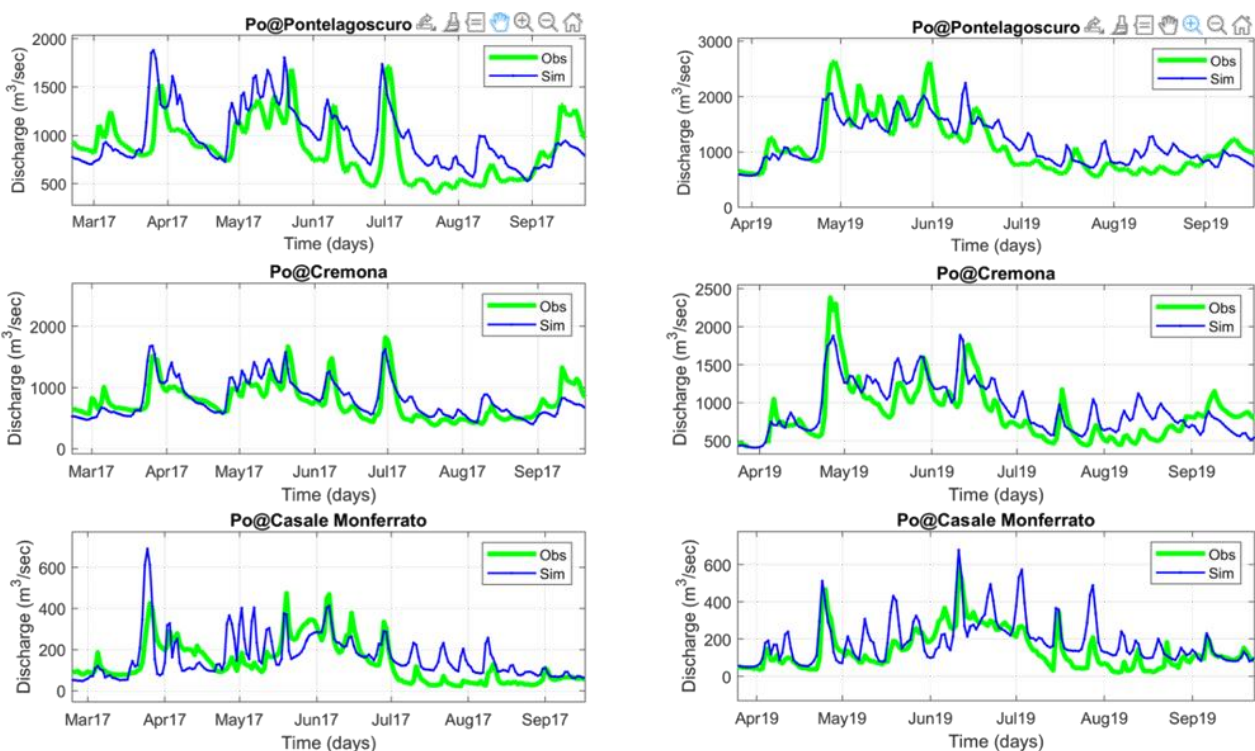
Figure 3.3: Comparison between observed and modelled river discharge time series at Pontelagoscuro, Cremona and Casale Monferrato along the Po river. Blue and orange rectangles identify floods and drought periods.

As it can be noted the model is able to accurately reproduce the river discharge time series not only over the calibration sections but also over the validation sections (bold stations in Table 3.3). In particular the model is able to accurately reproduce the flood events that affected the Po basin in November 2016, November 2018 and 2019 (see blue rectangles in Figure 3.3).

Table 3.3: Skill scores in terms of Kling-Gupta Efficiency (KGE) over the period 2015-2021 for the 18 gauging stations over the Po river basin. Results for the 8 validation stations are in bold.

ID station	Name station	KGE	ID station	Name station	KGE
1	Stura di Demonte@Gaiola	0.78	10	Parma@Ponte Verdi	0.72
2	Bormida@Cassine	0.77	11	Dora Baltea@Tavagnasco	0.86
3	**Po@Torino Murazzi**	**0.31**	12	Po@Spessa	0.89
4	Stura di Lanzo@Lanzo	0.80	**13**	**Po@Piacenza**	**0.88**
5	**Secchia@Ponte Alto**	**0.38**	14	Po@Pontelagoscuro	0.88
6	**Taro@S. Secondo**	**0.36**	**15**	**Po@Cremona**	**0.91**
7	Toce@Candoglia	0.85	**16**	**Po@Borgoforte**	**0.83**
8	Sesia@Palestro	0.71	17	Tanaro@Farigliano	0.85
9	Po@Casale Monferrato	0.87	**18**	**Tanaro@Alba Q.A.**	**0.77**

Modular modelling for water resources management

Figure 3.4 shows the river discharge time series simulated by MISDc for some stations along the Po river during the dry summer periods that occurred in 2017 and 2019. As it can be noted the model is able to reproduce the low flow periods for the stations of Pontelagoscuro, Cremona and Casale Monferrato.

Figure 3.4: Comparison between observed and modelled river discharge time series at Pontelagoscuro, Cremona and Casale Monferrato along the Po river, during the 2017 and 2019 drought periods highlighted in Figure 3.3 (orange boxes).

However, in order to assess the suitability of modular modelling for water resources management, it is not sufficient to validate the results in terms of river discharge, but additional output variables should be considered. To this end, modelled surface and root zone soil moisture, evapotranspiration and snow water

equivalent were compared with their respective time series extracted from external EO datasets. Figure 3.5 illustrates the comparison between modelled and GLEAM surface (SM first layer) and root-zone (SM second layer) soil moisture, modelled and GLEAM evapotranspiration (Evap) and modelled and CIMA snow water equivalent (SWE) at Pontelagoscuro section.
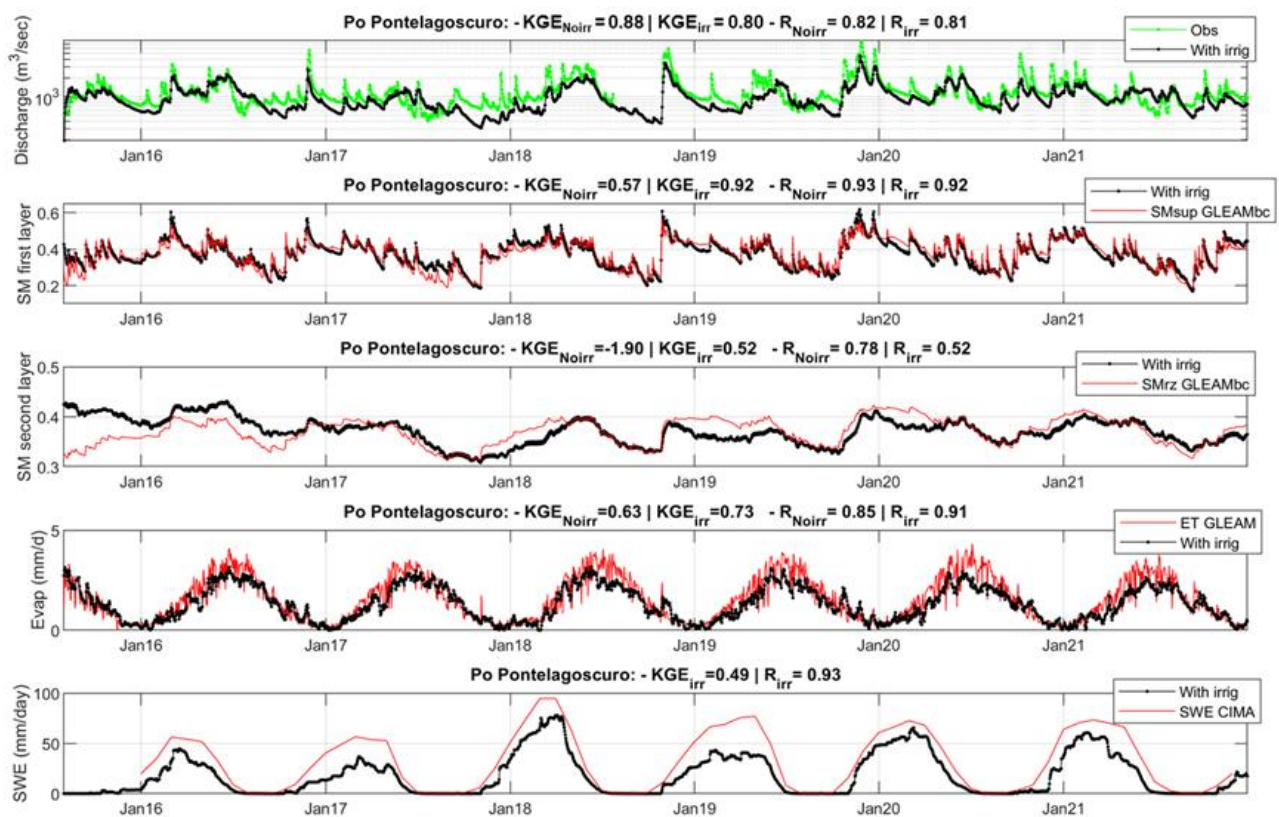
Figure 3.5: Observed and modelled river discharge, modelled and GLEAM surface (SM first layer) and root-zone (SM second layer) soil moisture, modelled and GLEAM evapotranspiration (Evap) and modelled and CIMA snow water equivalent (SWE) at Pontelagoscuro section.

An additional comparison has been carried out in terms of irrigation in Figure 3.6 between the modelled and the satellite-based estimates at monthly and yearly scale. Overall, the good agreement between the modelled and the EO datasets indicates the suitability of the modular modelling for the water resources management. In particular, the annual irrigation estimates averaged over the period 2015-2021 are equal to 510 mm, which corresponds to 16.8 km^3, in line with the AdBPO estimates.
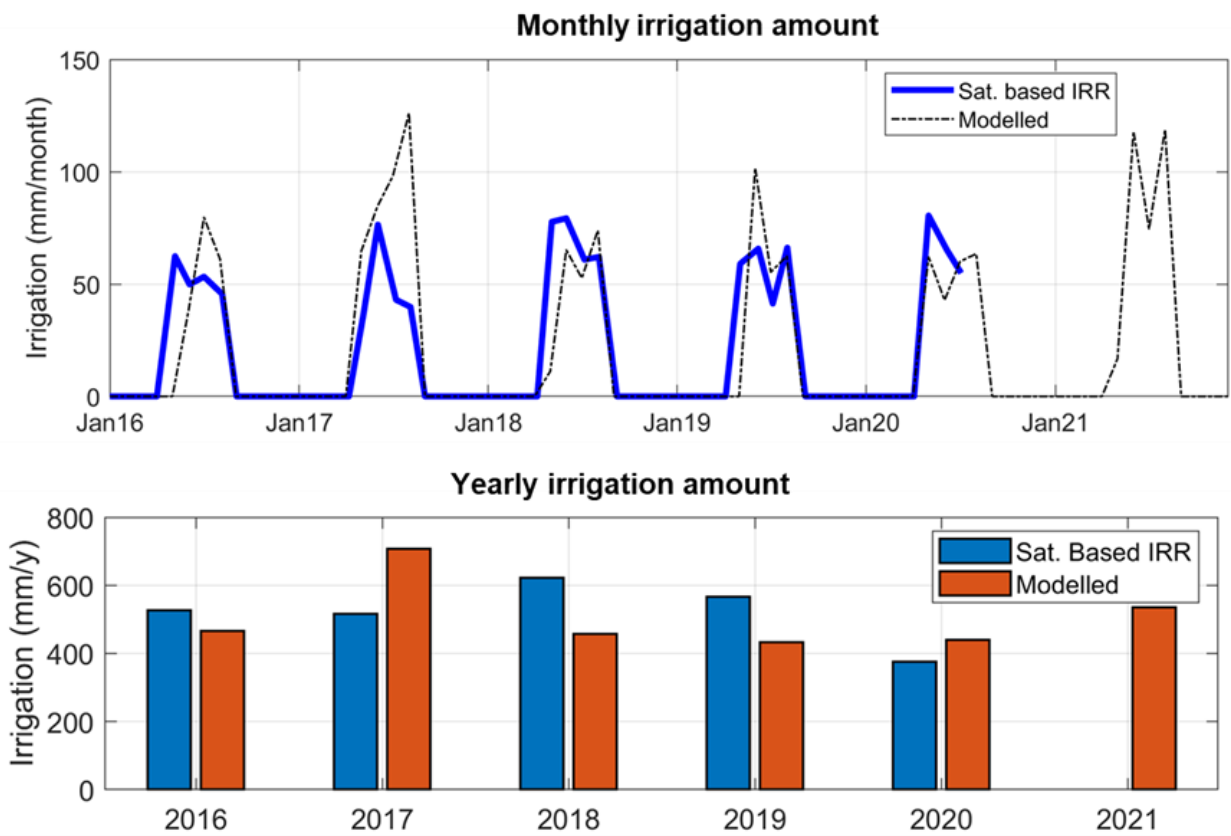
Figure 3.6: Satellite-based and modelled irrigation estimates at monthly and yearly scale for the Po river basin.

4.   References

Brocca, L., Melone, F. and Moramarco, T., 2011. Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting. Hydrological Processes, 25(18), 2801–2813.

Camici, S., Ciabatta, L., Massari, C. and Brocca, L., 2018. How reliable are satellite precipitation estimates for driving hydrological models: a verification study over the Mediterranean area. Journal of Hydrology, 563, 950–961. doi:10.1016/j.jhydrol.2018.06.067

De Lavenne, A., Andréassian, V., Thirel, G., Ramos, M. H., & Perrin, C. (2019). A regularization approach to improve the sequential calibration of a semidistributed hydrological model. Water Resources Research, 55(11), 8821-8839.

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

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

Masseroni, D., Cislaghi, A., Camici, S., Massari, C. and Brocca, L., 2017. A reliable rainfall–runoff model for flood forecasting: review and application to a semi-urbanized watershed at high flood risk in Italy. Hydrology Research, 48(3), 726–740. doi:10.2166/nh.2016.037.

Avanzi, F., Ercolani, G., Gabellani, S., Cremonese, E., Pogliotti, P., Filippa, G., Morra di Cella, U., Ratto, S., Stevenin, H., Cauduro, M., and Juglair, S. (2020). Learning about precipitation orographic enhancement from snow-course data improves water-balance modeling, Hydrol. Earth Syst. Sci. Discuss. [preprint], https://doi.org/10.5194/hess-2020-571, in review.

Brocca, L., Filippucci, P., Hahn, S., Ciabatta, L., Massari, C., Camici, S., Schüller, L., Bojkov, B., Wagner, W. (2019). SM2RAIN-ASCAT (2007-2018): global daily satellite rainfall from ASCAT soil moisture. Earth System Science Data, 11, 1583–1601, doi:10.5194/essd-11-1583-2019.

Brocca, L., Tarpanelli, A., Filippucci, P., Dorigo, W., Zaussinger, F., Gruber, A., Fernández-Prieto, D., 2018. How much water is used for irrigation? A new approach exploiting coarse resolution satellite soil moisture products. Int. J. Earth Obs. Geoinformation, 73, 752-766.

Burek, P., Satoh, Y., Kahil, T., Tang, T., Greve, P., Smilovic, M., ... & Wada, Y. (2020). Development of the Community Water Model (CWatM v1. 04) A high-resolution hydrological model for global and regional assessment of integrated water resources management, GMD, 13, 3267–3298, https://doi.org/10.5194/gmd-13-3267-2020.

Dari, J., Brocca, L., Quintana-Seguí, P., Escorihuela, M.J., Stefan, V., Morbidelli, R., 2020. Exploiting high-resolution remote sensing soil moisture to estimate irrigation water amounts over a Mediterranean region. Remote Sens., 12, 2593.

Dari, J., Quintana-Seguí, P., Morbidelli, R., Saltalippi, C., Flammini, A., Giugliarelli, E., Escorihuela, M.J., Stefan, V., Brocca, L., 2022. Irrigation estimates from space: Implementation of different approaches to model the evapotranspiration contribution within a soil-moisture-based inversion algorithm. Agric. Water Manag., 265,107537.

Dari, J., Brocca, L., Modanesi, S., Massari, C., Tarpanelli, A., Barbetta, S., Quast, R., Vreugdenhil, M., Freeman, V., Barella-Ortiz, A., Quintana-Seguí, P., Bretreger, D., Volden, E. 2023. Regional data sets of high-resolution (1 and 6 km) irrigation estimates from space. Earth System Science Data, 15(4), 1555-1575, doi:10.5194/essd-15-1555-2023.

Martens, B., Miralles, D.G., Lievens, H., van der Schalie, R., de Jeu, R.A.M., Fernández-Prieto, D., Beck, H.E., Dorigo, W.A., Verhoest, N.E.C., 2017. GLEAM v3: satellite-based land evaporation and root-zone soil moisture. Geosci. Model Dev. 10, 1903–1925. https://doi.org/10.5194/gmd-10-1903-2017

Miralles, D.G., Holmes, T.R.H., De Jeu, R.A.M., Gash, J.H., Meesters, A.G.C.A., Dolman, A.J., 2011. Global land-surface evaporation estimated from satellite-based observations. Hydrol. Earth Syst. Sci. 15, 453–469. https://doi.org/10.5194/hess-15-453-2011

Miralles, D.G., de Jeu, R.A.M., Gash, J.H., Holmes, T.R.H. and Dolman, A.J. (2011b). Magnitude and variability of land evaporation and its components at the global scale, Hydrology and Earth System Sciences, 15, 967–981, doi: 10.5194/hess-15-967-2011.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C., Thépaut, J.-N., 2021. ERA5-Land: A state-of-the-art global reanalysis dataset for land applications (preprint). Data, Algorithms, and Models. https://doi.org/10.5194/essd-2021-82

Quast, R., Albergel, C., Calvet, J.-C., Wagner, W., 2019. A Generic First-Order Radiative Transfer Modelling Approach for the Inversion of Soil and Vegetation Parameters from Scatterometer Observations. Remote Sens. 11, 285. https://doi.org/10.3390/rs11030285

Rains, D., Trigo, I., Dutra, E., Ermida, S., Ghent, D., Hulsman, P., Gómez-Dans, J., and Miralles, D. G. (2022). High-resolution all-sky land surface temperature and net radiation over Europe. Earth System Science Data Discussions. doi: 10.5194/essd-2022-302.

Alfieri, L., Avanzi, F., Delogu, F., Gabellani, S., Bruno, G., Campo, L., Libertino, A., Massari, C., Tarpanelli, A., Rains, D., Miralles, D. G., Quast, R., Vreugdenhil, M., Wu, H., and Brocca, L.: High-resolution satellite products improve hydrological modeling in northern Italy, Hydrology and Earth System Sciences, 26, 3921–3939, https://doi.org/10.5194/hess-26-3921-2022, 2022.

Bruno, G., Pignone, F., Silvestro, F., Gabellani, S., Schiavi, F., Rebora, N., Giordano, P., and Falzacappa, M.: Performing Hydrological Monitoring at a National Scale by Exploiting Rain-Gauge and Radar Networks: The Italian Case, Atmosphere, 12, 771, https://doi.org/10.3390/atmos12060771, 2021.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Joint Research Centre (European Commission), Ziese, M., Garcia, R., Dottori, F., Salamon, P., Schweim, C., Arnal, L., Disperati, J., Prudhomme, C., Gomes, G., Krzeminski, B., De Roo, A., Baugh, C., Sprokkereef, E., Lorini, V., Kalas, M., Mazzetti, C., Muraro, D., Wetterhall, F., Smith, P., Thiemig, V., Mikulickova, M., Asp, S.-S., Skoien, J. O., Rehfeldt, K., Garcia-Padilla, M., Gelati, E., Rauthe-Schöch, A., and Latini, M.: EFAS upgrade for the extended model domain: technical documentation, Publications Office of the European Union, LU, 2019.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash-Sutcliffe and Kling-Gupta efficiency scores, Hydrology and Earth System Sciences Discussions, 1–7, https://doi.org/10.5194/hess-2019-327, 2019.

Silvestro, F., Gabellani, S., Delogu, F., Rudari, R., and Boni, G.: Exploiting remote sensing land surface temperature in distributed hydrological modelling: the example of the Continuum model, Hydrology and Earth System Sciences, 17, 39–62, https://doi.org/10.5194/hess-17-39-2013, 2013.

Sinclair, S. and Pegram, G.: Combining radar and rain gauge rainfall estimates using conditional merging, Atmospheric Science Letters, 6, 19–22, 2005.

Technical Note 4: DTE Hydrology community platform end to end system integration and final functional description v1.0

1. Introduction

1.1. The DTE Hydrology Evolution Project

The overarching objective of Digital Twin Earth (DTE) Hydrology Evolution project is to prototype a full advanced end-to-end demonstrator of DTE over the full Mediterranean region at high resolution in space and time (targeting 1 km and 1 hour) by extending the successful implementation of Digital Twin Earth Hydrology project that has been carried out over the Po River Valley in northern Italy. Specifically, DTE Hydrology has demonstrated the potential to advance towards an end-to-end reconstruction of the hydrological cycle at high resolution in space and time by an effective combination of state-of-the-art EO data, in situ observations, advanced hydrological and hydraulic modelling, AI and advanced digital platform capabilities. DTE Hydrology Evolution project we aim to further advance the DTE Hydrology concept (see **Figure 1**) through (1) the extension to the full Mediterranean region and hence the preparation to develop a Digital Twin of Hydrological processes and applications at continental and global scales, (2) the exploitation of the dedicated cloud-HPC infrastructure set up in ESRIN, (3) the engagement with a larger community of scientists and stakeholders, (4) the development of an advanced DTE Hydrology open science platform as a community tool offering an interactive access to the EO data, advanced products, in-situ data, and model results, and (5) the development of an attractive and efficient user interface, dashboards, and 3D immersive visualisation tools in an advanced digital environment for scientists, policy makers and citizens. It will be based on state-of-the-art digital platform capabilities and will represent a solid basis for a future ESA supported tool for Polar research and Earth system science.

Figure 1: DTE Hydrology Evolution concept.

1.2.   Scope of this Report

This report describes the technical background of the Open Platform developed under the DTE Hydrology Evolution project for showcasing the 4 use cases: (1) large scale water balance assessment, (2) flooding simulations, (3) what-if scenario for flood risk assessment, and (4) what-if scenario for water resources management.

1.3.   Applicable Documents

Not available.

1.4.   Reference Documents

RD-01   DTE Hydrology Evolution Technical Proposal - V1.0

RD-02   Technical Note 1: DTE Hydrology end-to-end description and demonstration plan - V1.0

RD-03   Technical Note 2: DTE-Hydrology Datasets User Manual v1.0

RD-04   Technical Note 3: DTE Hydrology modelling experiments, validation and assessment v1.0

2. Technical background of DTE Hydrology platform

The platform is a conglomeration of technologies and languages that have made possible both the creation of the user interface and the background services that can be queried from it in order to provide data that can be analyzed and explored.

In general, the application can be divided into layers: one related to the GUI, the other related to data.

The former is responsible for displaying the data served upon request and also perform the requests themselves while the latter serves the only purpose of outputting the data, provided the requests are properly performed.

2.1. Graphical User Interface (GUI)

The user interface is crafted combining three languages (HTML5, CSS3 and ECMAScript 2015+) and multiple libraries and frameworks. The components are built with of Bootstrap, a popular, open-source front-end framework designed to make designing and building responsive, mobile-first websites, web applications, and UI components faster and easier. It includes pre-designed CSS styles, JavaScript components, and HTML templates for creating responsive layouts, typography, forms, buttons, and more. It also supports a responsive grid system for building flexible, adaptive layouts, and provides extensive documentation and community support. Such components are then brought to life with jQuery, a fast, small, and feature-rich JavaScript library that simplifies HTML document traversal and manipulation, event handling, and animation.

Combining these tools it is possible to interact with the components to perform any operation on the data, based on the ability to send/receive requests to/from the server, so that data can be retrieved, manipulated and sent back.
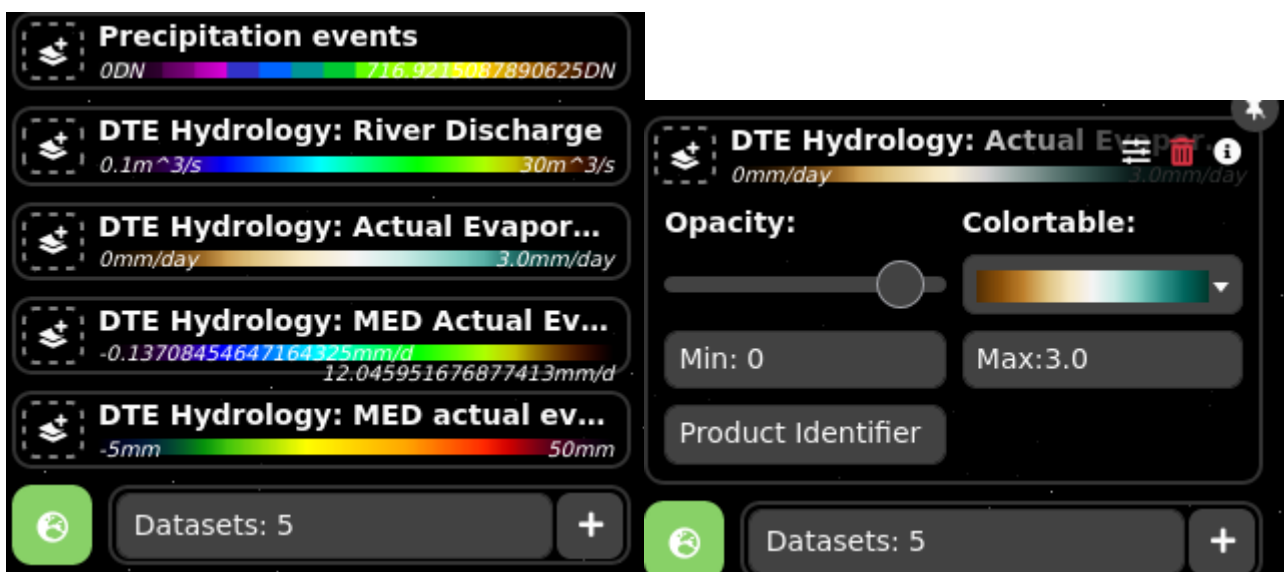
Figure 2.1.1: An example of components on the platform built to toggle one or more collection(s) on the map (left) and the equalization panel of a single dataset (right).

Another type of component that can be spotted quite often around the GUI is the plot component. Such components are rendered, updated and customized taking advantage of Plotly, a JavaScript graphing library that allows data visualization through various types of charts, including scatter plots, bar charts, and heat maps. It offers interactive charts with customizable features such as zooming, panning, and hovering over data points for additional information; the library can also be used with various programming languages such as Python, R, and MATLAB.

Figure 2.1.2: Plotly in action in four different scenarios (and counting!).

The platform offers an interactive map, initially presented as a 3D globe, thanks to NASA's WebWorld Wind - a state-of-the-art web-based virtual globe application that allows users to explore and interact with the world in three dimensions. It is an open-source and free tool that runs on most web browsers and personal computers.

WebWorld Wind is designed to provide users with a high-resolution, detailed, and interactive view of the Earth's surface and atmosphere. It uses advanced geospatial data processing and imaging technologies to render the globe with stunning accuracy and clarity, displaying features such as continents, countries, cities,

oceans, mountains, and even meteorological data in real-time. One of its major features is its ability to provide users with data layers that can be added to the application, displaying a mosaic with the preferred color table, but it's also possible to render points and polygons on the planet's surface.

WebWorld Wind can be used for a wide variety of purposes, including scientific research, education, tourism, and civil services planning. It provides an intuitive and interactive interface that can easily engage users, encouraging exploration, learning, and discovery. In summary, WWW is a powerful and innovative web-based virtual globe application that allows users to explore and interact with the world in three dimensions.

Figure 2.1.3: Tilted globe showing custom points, polygons and a rivers layer.

In addition, the map can be manipulated in terms terrain exaggeration, projection and base layer.
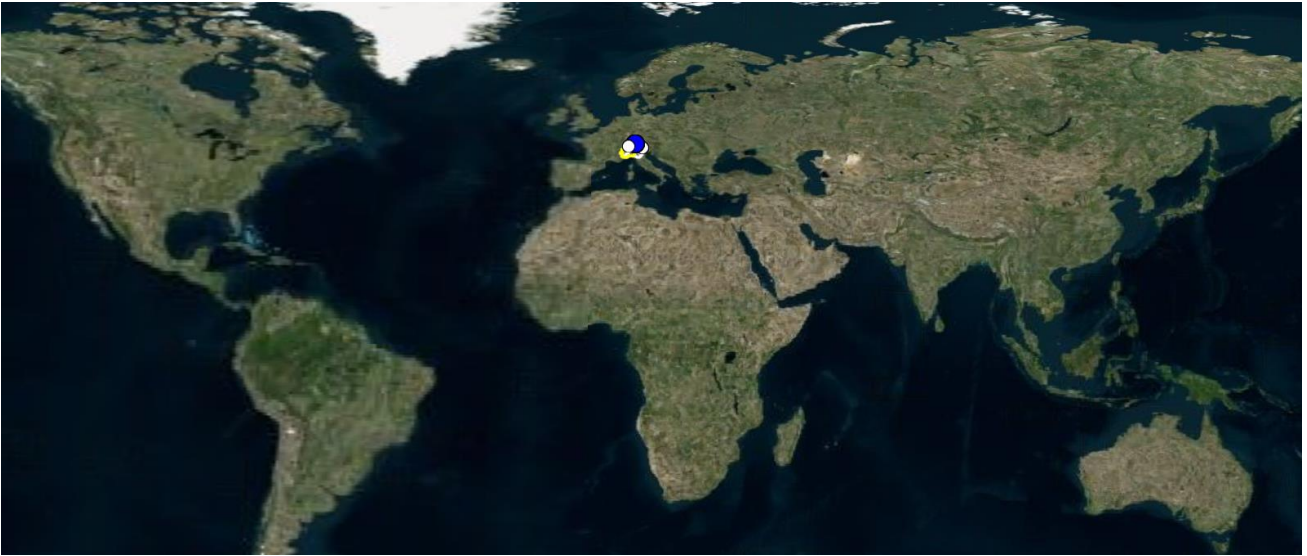
Figure 2.1.4: Map view switched from 3D to equirectangular.
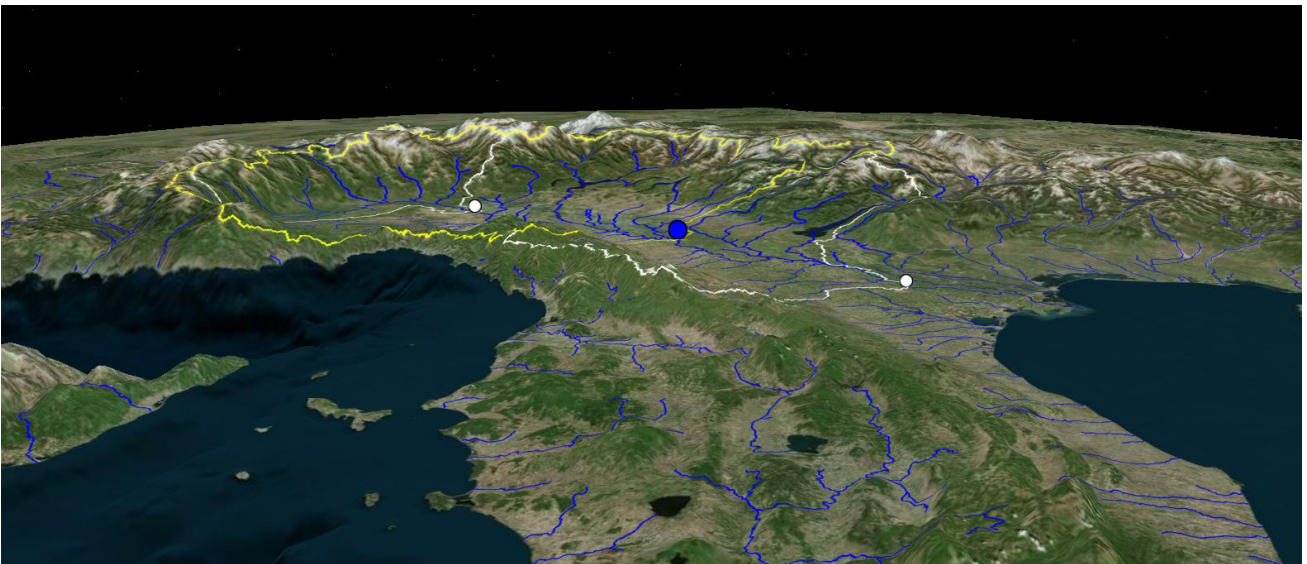
Figure 2.1.5: Tilted globe showing terrain exaggeration.

Figure 2.1.6: An example of alternative base layer.

2.2.   Data

The data is served through Django, a high-level Python web framework that makes it easy to build web applications quickly and efficiently. It provides a lot of built-in functionality, such as an ORM for database access, a templating engine for rendering HTML, and a routing system for handling URLs. Django is also known for its emphasis on security and scalability, making it a popular choice for building large-scale web applications. It is used to craft RESTful APIs to take incoming requests and generating responses, putting data under the shape of a JSON object.

Django is used to expose many geospatial data services such as OGC OpenSearch, allowing users to search for and discover geospatial data from a variety of sources, including maps, satellite imagery, and other geospatial data.

Figure 2.2.1: An example of the results of an OpenSearch query being rendered.

Besides OS, once the data has been discovered/searched, the Web Map Service comes into play as a standard protocol for serving geospatial data over the internet. It allows users to request maps of specific areas and layers of data, which are then generated and returned as images.

Each image is generated taking care of the filters, time parameters and color equalization, if present.
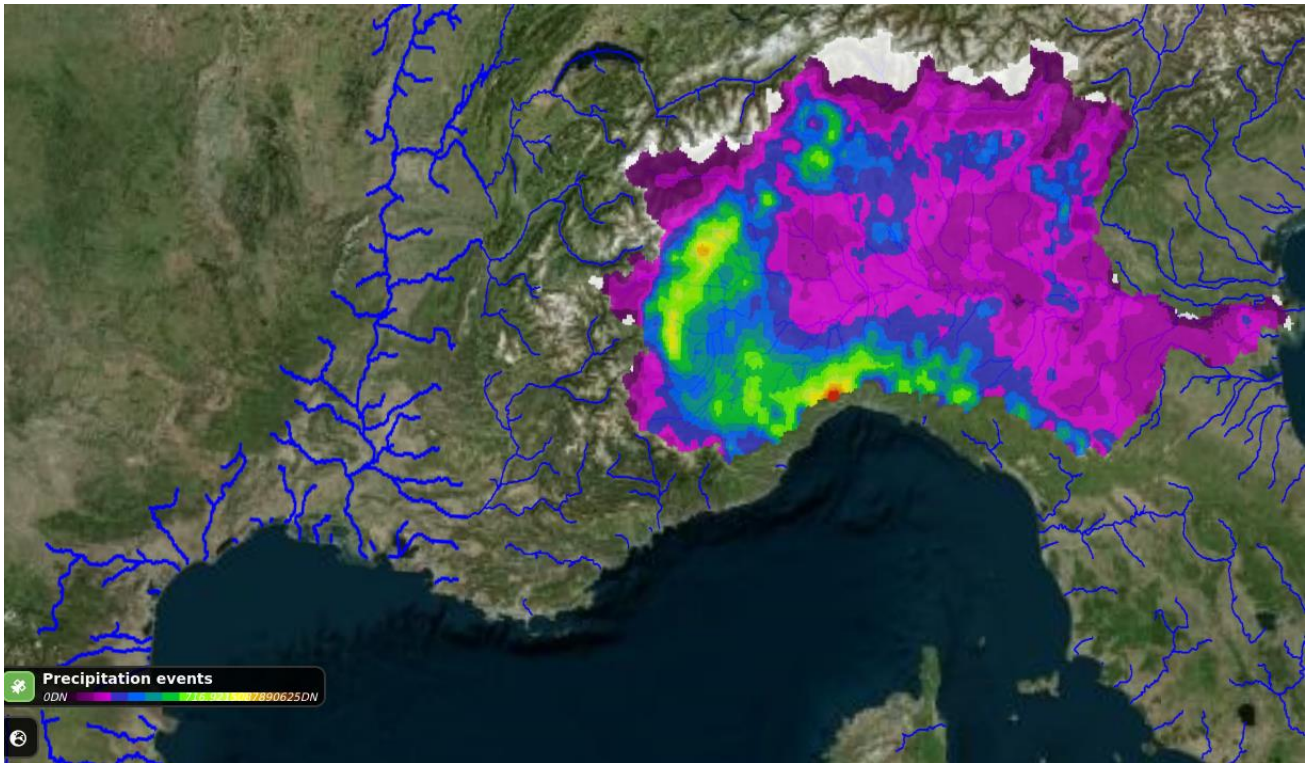
Figure 2.2.2: A WMS layer in action over the Po river's basin.

Technical Note 5: DTE Hydrology Live Demonstration v1.0

1. Introduction

1.1. The DTE Hydrology Evolution Project

The overarching objective of Digital Twin Earth (DTE) Hydrology Evolution project is to prototype a full advanced end-to-end demonstrator of DTE over the full Mediterranean region at high resolution in space and time (targeting 1 km and 1 hour) by extending the successful implementation of Digital Twin Earth Hydrology project that has been carried out over the Po River Valley in northern Italy. Specifically, DTE Hydrology has demonstrated the potential to advance towards an end-to-end reconstruction of the hydrological cycle at high resolution in space and time by an effective combination of state-of-the-art EO data, in situ observations, advanced hydrological and hydraulic modelling, AI and advanced digital platform capabilities. In the DTE Hydrology Evolution project we aim to further advance the DTE Hydrology concept (see **Figure 1**) through (1) the extension to the full Mediterranean region and hence the preparation to develop a Digital Twin of Hydrological processes and applications at continental and global scales, (2) the exploitation of the dedicated cloud-HPC infrastructure set up in ESRIN, (3) the engagement with a larger community of scientists and stakeholders, (4) the development of an advanced DTE Hydrology open science platform as a community tool offering an interactive access to the EO data, advanced products, in-situ data, and model results, and (5) the development of an attractive and efficient user interface, dashboards, and 3D immersive visualisation tools in an advanced digital environment for scientists, policy makers and citizens. It will be based on state-of-the-art digital platform capabilities and will represent a solid basis for a future ESA supported tool for Polar research and Earth system science.

Figure 1: DTE Hydrology Evolution concept.

1.2.   Scope of this Report

This report describes the DTE Hydrology Live Demonstration, meant as an "immersive and interactive" live demonstration of DTE Hydrology offering an intuitive and attractive 3D advanced visualisation of the DTE data and simulation (what if scenarios) results over the full Mediterranean region.

The demonstration was provided in October 2022 at the ESA Programme Board in Frascati, at the final meeting in Assisi (April 2023), at a number of internal meetings with stakeholders, as well as disseminated through social media. This demonstration emphasised the use of attractive state-of-the-art 3D interactive visualisation capabilities with interactive options for the users and an intuitive navigation and access to the data and results. In doing so, it demonstrated the potential of using HPC capabilities in the context of DTE Hydrology.

In the following, four case studies available on the online platform will be described:

- Large scale water balance assessment (Section 2);
- Flood prediction (Section 3);
- What-if scenario for flood risk assessment (Section 4);
- What-if scenario for water resources management (Section 5).

For each case study, we will both briefly describe the underlying data and provide a "user manual" on how to access and use the data on the online platform (https://explorer.dte-hydro.adamplatform.eu/).

1.3.   Applicable Documents

Not available.

1.4.   Reference Documents

RD-01   DTE Hydrology Evolution Technical Proposal - V1.0

RD-02   Technical Note 1: DTE Hydrology end-to-end description and demonstration plan - V1.0

RD-03   Technical Note 2: DTE-Hydrology Datasets User Manual v1.0

RD-04   Technical Note 3: DTE Hydrology modelling experiments, validation and assessment v1.0

RD-05   Technical Note 4: DTE Hydrology community platform end to end system integration and final functional description v1.0

2.   Large scale water balance assessment

In hydrology, a "water balance" is the application of the conservation of mass to the flow of water through a water basin. in mathematical terms:

$$Q = P - ET - dS$$

where P is precipitation, Q is discharge, ET is actual evapotranspiration, and dS is basin storage (whether in the form of groundwater, snow, glacier ice, and the like). The water balance can be computed at various scales, from the daily to the yearly (or multi-yearly) one. In this regard, monthly to yearly water balance terms can provide insights into, e.g., the propagation of precipitation deficit into streamflow deficit via groundwater (Bruno et al., 2022), available water resources in snow and implications of snow droughts (Avanzi et al., 2023), the role of irrigation in sustaining agriculture (Dari et al., 2023), or precipitation or ET estimate from satellite data (Brocca et al., 2014, Miralles et al., 2011). The water balance equation is at the foundation of modern hydrology, to the extent that it can be used even to diagnose issues with weather input data into hydrologic models (Avanzi et al., 2021) or shifts in catchment functioning (Massari et al., 2022).

For the DTE case study 1 on large scale water balance assessment, we chose to compute and display monthly terms of the Mediterranean water balance, as a trade-off between maximising temporal resolution of the DTE case study 1 information and reducing non-linearities in water-balance computations at very fine temporal scales.

For precipitation, we used an integrated product derived by downscaling and merging the CPC (Climate Prediction Center, Chen et al., 2008) and GPM-LR (Global Precipitation Measurement – Late Run, Zhou et al., 2019) products. CPC combines gauge sensor data to retrieve daily precipitation at 0.5 degree spatial resolution, while GPM-LR combines satellite data from the GPM mission to estimate precipitation with a spatial resolution of 0.1 degree and a temporal resolution of 30 minutes. Both the products are available in near real-time.

The spatial resolution of the two products is coarse with respect to the target (1 km), therefore they were first downscaled to 1 km spatial resolution by using an high spatial resolution climatology derived by CHELSA (Climatologies at high resolution for the earth's land surface areas; Karger et al., 2017). The two downscaled products were then integrated using a spatial weight matrix derived by a Triple collocation (TC) analysis applied to CPC, GPM-LR and ERA5 (Hersbach et al., 2020). For further information regarding the downscaling and the merging procedure, see "4DMED-Hydrology Deliverable 3.2: Algorithm Theoretical Basis Document (ATBD)". **Figure 2** shows the performance of the CPC, GPM and the integrated 4DMED product. Specifically, Pearson's correlation and Root Mean Square Error were calculated against high spatial resolution observed data from gauge and radar data in three study areas: the Ebro basin, the Herault basin and the Po River basin. The combined 4DMED product outperforms both CPC and GPM_LR in all study areas, both in terms of Pearson

Correlation and Root Mean Square Error. More information about the performance can be found in the document "4DMED-Hydrology Deliverable 3.2: Product Validation Report (PVR)".
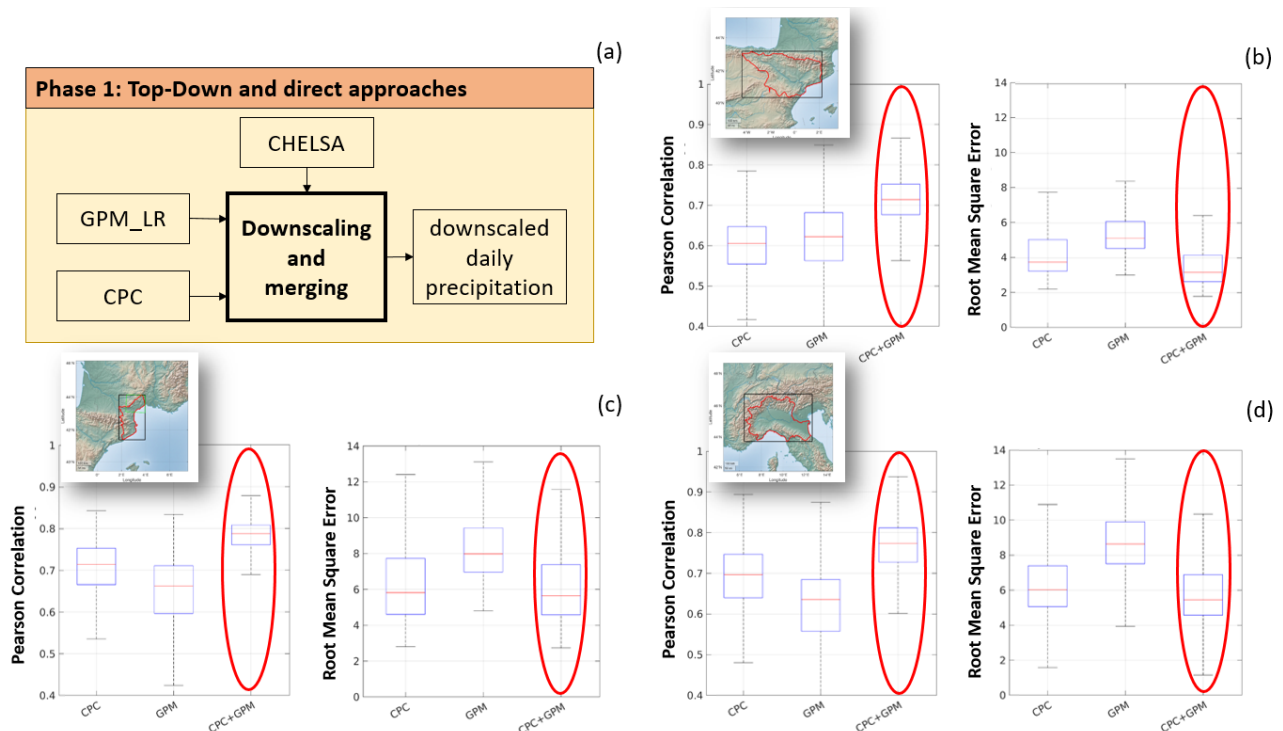

Figure 2: scheme of the rainfall product generation (panel a) and performance indices (Pearson correlation and Root mean square error) of the CPC, GPM and the integrated 4DMED products against high resolution observed data collected in the study areas of Ebro basin (panel b), Herault basin (panel c) and Po River basin (panel d).

For ET, we used the Global Land Evaporation Amsterdam Model (GLEAM, Miralles et al., 2011). GLEAM uses satellite observations to estimate the actual, terrestrial evaporation as the sum of the following components: transpiration (Et), interception loss (Ei), bare-soil evaporation (Eb), snow sublimation (Es) and open-water evaporation (Ew). A Priestley and Taylor equation calculates Ep based on observations of surface net radiation and near-surface air temperature (Priestley and Taylor, 1972). Potential evaporation estimates are converted into actual evaporation based on the multiplicative, evaporative stress factor S. The derivation of S is based on microwave observations of the vegetation optical depth – used as a proxy for the vegetation water content – and simulations of root-zone soil moisture. The latter is calculated using a multi-layer running water balance that describes the infiltration of observed precipitation through the vertical soil profile. To correct for random forcing errors, microwave observations of surface soil moisture are assimilated into the soil profile. Interception loss is calculated separately based on the Gash analytical model of rainfall interception (Gash, 1979) driven by observations of precipitation and both vegetation and rainfall characteristics. For a detailed description, the reader is directed to Miralles et al. (2011) and Martens et al. (2017). More information on the

project specific forcing data used to run GLEAM can be found in the document "4DMED-Hydrology Deliverable 3.1: Algorithm Theoretical Basis Document (ATBD)". GLEAM 1km evaporation estimates show high spatial detail compared to the default GLEAM v3.6b dataset which is at 25 km resolution (**Figure 3**). The evaporation dataset has been validated with respect to nine eddy-covariance stations in the study domain which have been obtained from the European Fluxes Database Cluster (EFDC). Results indicate the GLEAM-1km dataset represents the evaporation reasonably well with a median correlation of 0.70; see "4DMED-Hydrology Deliverable 3.2: Product Validation Report (PVR)" for more information on the performance.
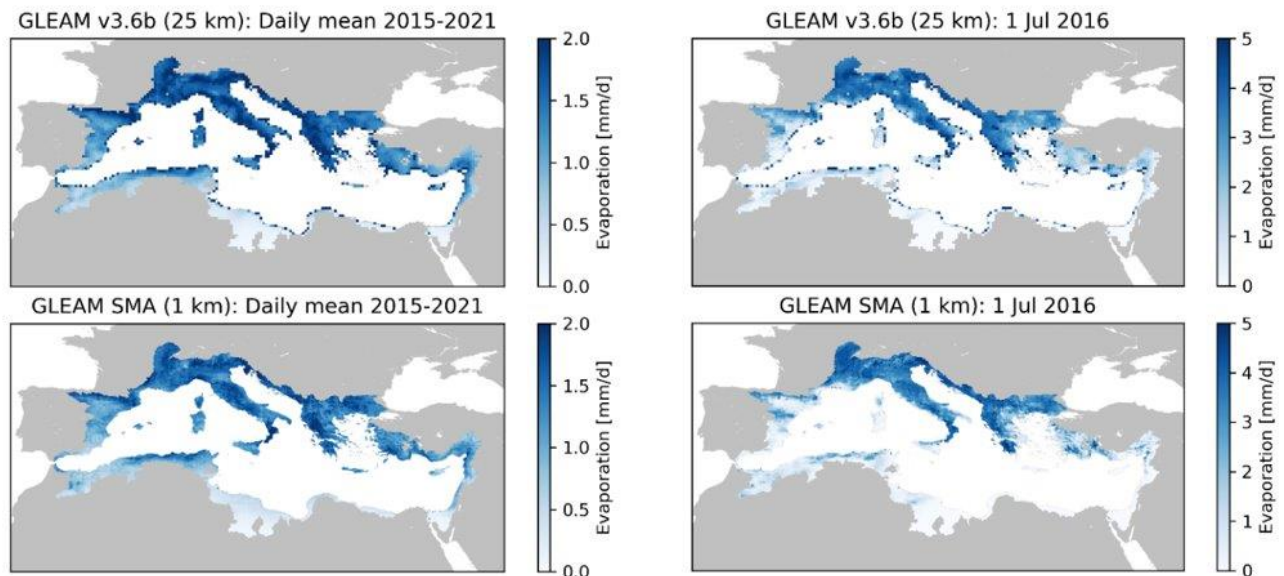
Figure 3: comparison of GLEAM v3.6b (25km, top) versus GLEAM SMA (1km, bottom). Daily mean over 2015-2021 (left) and sample data for 1 July 2016 (right).

For soil moisture, we used the RT1 surface soil moisture product retrieved from Sentinel-1 VV backscatter at 500m sampling (Quast and Wagner, 2019). RT1 is a radiative transfer model that uses non-linear least squares optimization between the incidence-angle dependent measured and modelled backscatter time series. Within the RT1 model, Copernicus Global Land Service data on Leaf Area Index is used to correct for vegetation dynamics.

The soil moisture retrieval with RT1 is done at 500m, but benefits from using 20m backscatter data from Sentinel-1 through strict masking for non-sensitive pixels. When resampling backscatter to 500m, only 20m pixels that are sensitive to soil moisture are used. Thus, pixels that are located over cities, lakes and are subject to sub-surface scattering (Wagner et al., 2022) are masked out. In the Mediterranean, sub-surface scattering may affect the backscatter signal, where under dry conditions multiple scattering can start to occur through interaction with rocks in the sub-surface (**Figure 4**). This can lead to higher backscatter, even though soil moisture is at its minimum. Sub-surface scattering pixels are identified based on the Pearson correlation coefficient between four years of backscatter signal and ERA5-Land soil moisture (Muñoz-Sabater et al., 2021). Validation with in situ stations and model data shows that soil moisture dynamics as reasonably well captured

by the RT1 soil moisture dataset, see "4DMED-Hydrology Deliverable 3.2: Product Validation Report (PVR)" for more information on the performance.
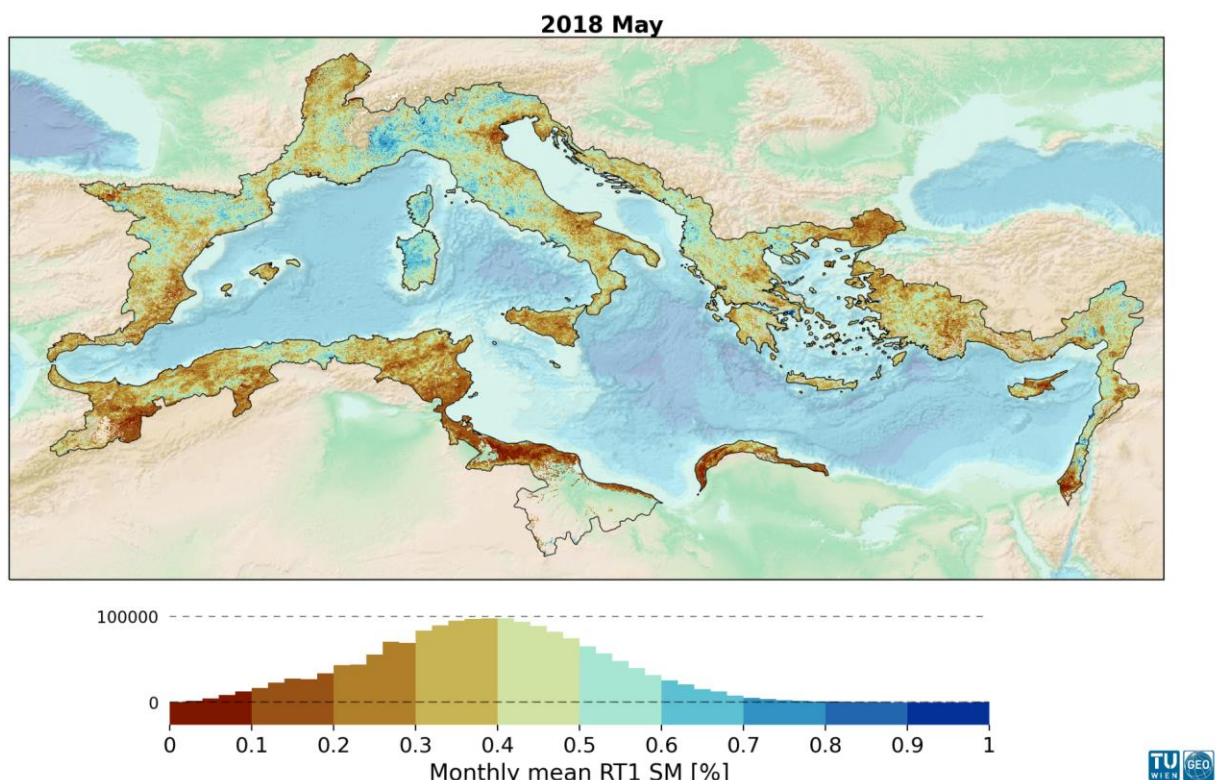
Figure 4: *RT1 soil moisture for the Mediterranean region for May 2018.*

For snow, we used a blended product obtained by assimilating the Sentinel-1-based product proposed by Lievens et al. (2019), C-SNOW, into the Snow Multidata Mapping and Modeling (S3M) 5.1 (Avanzi et al., 2022). S3M is a spatially distributed, physics-based snow and glacier model providing continuous-time estimates of snow water resources by solving the following processes: precipitation phase partitioning, snow mass balance, albedo aging, settling, snowmelt estimate based on air temperature and radiation, liquid water accumulation in snow and snow wetness, snow outflow generation based on an estimate of permeability, and snow refreezing. S3M includes a representation of glacier melt based on the same approach used for snow, but with different parameters. S3M is open source and available at https://github.com/c-hydro/s3m-dev . More information on this product is available in the ESA 4DMED deliverables 3.1 and 3.2. **Figure 5** shows an example of mean seasonal Snow Water Equivalent across the 4DMED domain, hydrological year 2018 (September 2017 through August 2018).

ESA 4DMED: S3M + C-SNOW, mean SWE for water year 2018
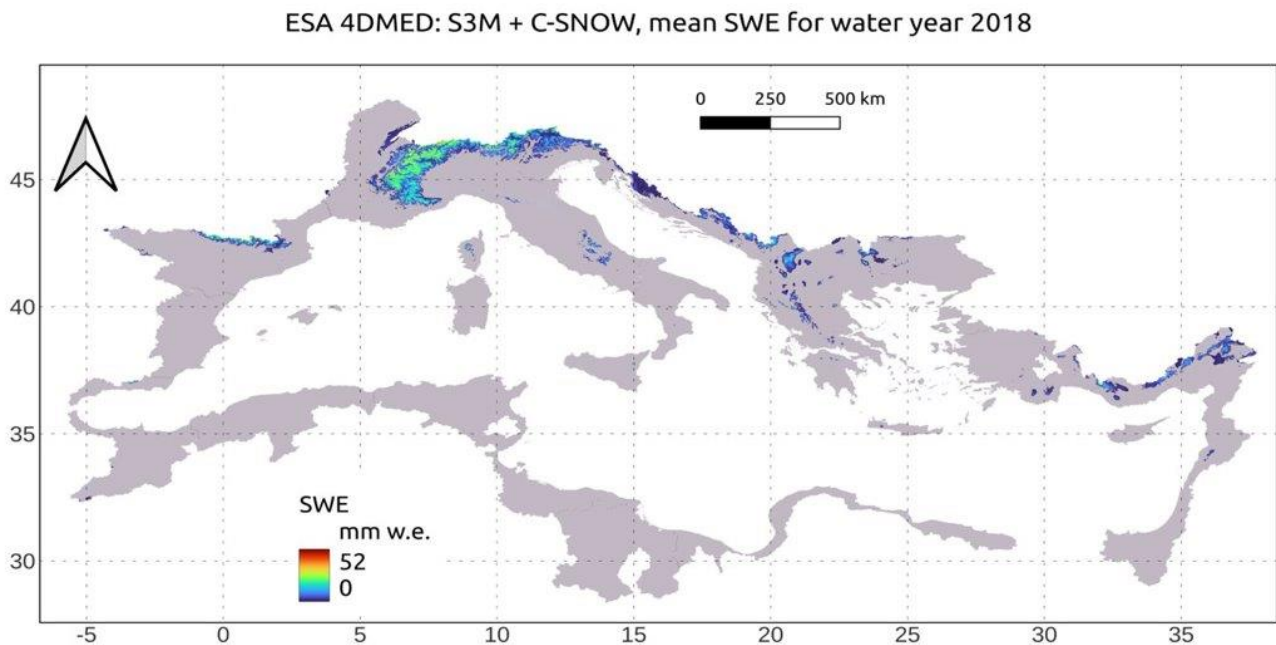
Figure 5: *Mean seasonal Snow Water Equivalent across the 4DMED domain, hydrological year 2018 (September 2017 through August 2018).*

Discharge was taken from the Continuum model and was obtained by running it in parallel in each of the 28 domains described in Deliverable D1.1. Model runs, all at 1 km and 1 hour resolution, include the following set of experiments: 28 model runs forced by ERA5 for 2015-2021, (one run per domain), 6 model runs forced by SM2RAIN (precipitation) and ERA5 (other variables), for the Ebro, Herault, Po, Rhone, Sicilia region and Medjerda, 1 model run for the Sicilia region (domain #7) forced by high-resolution ground-based data for the year 2021 (**Figure 6**).
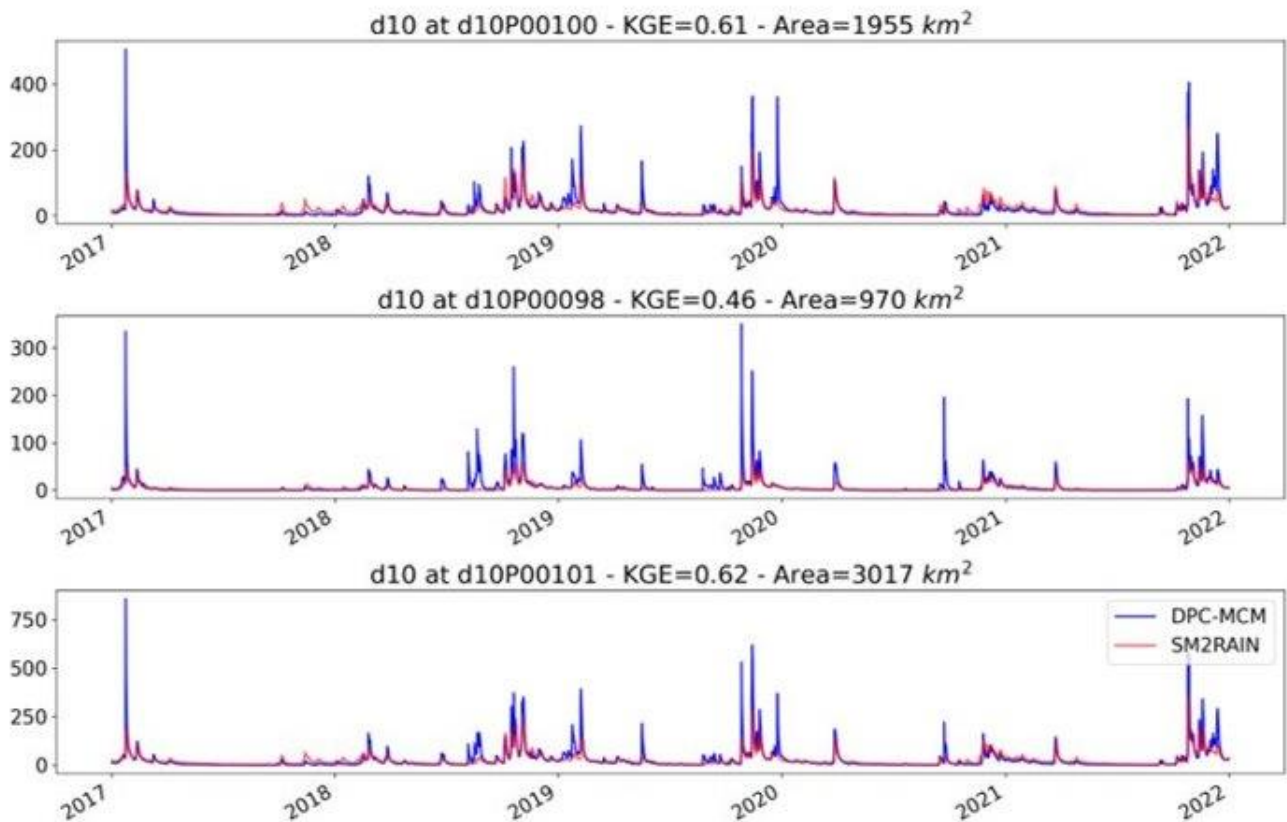
Figure 6: *Comparison of simulated discharge produced by Continuum at three sample stations in Sicily using DTE EO precipitation input ("SM2RAIN" in the plot) versus the original DPC-MCM 1km 1h resolution (blue).*

All the above datasets were shared by project partners and processed by MEEO srl to compute monthly anomalies and then displayed on an online platform (https://explorer.dte-hydro.adamplatform.eu/?use_case=1).

2.1. Web application

The DTE platform allows users to easily view the statistics of the different components of the water balance: precipitation, soil moisture, evaporation, snow water equivalent and runoff (**Figure 7**). This is achieved with an interactive map on the left part of the screen:
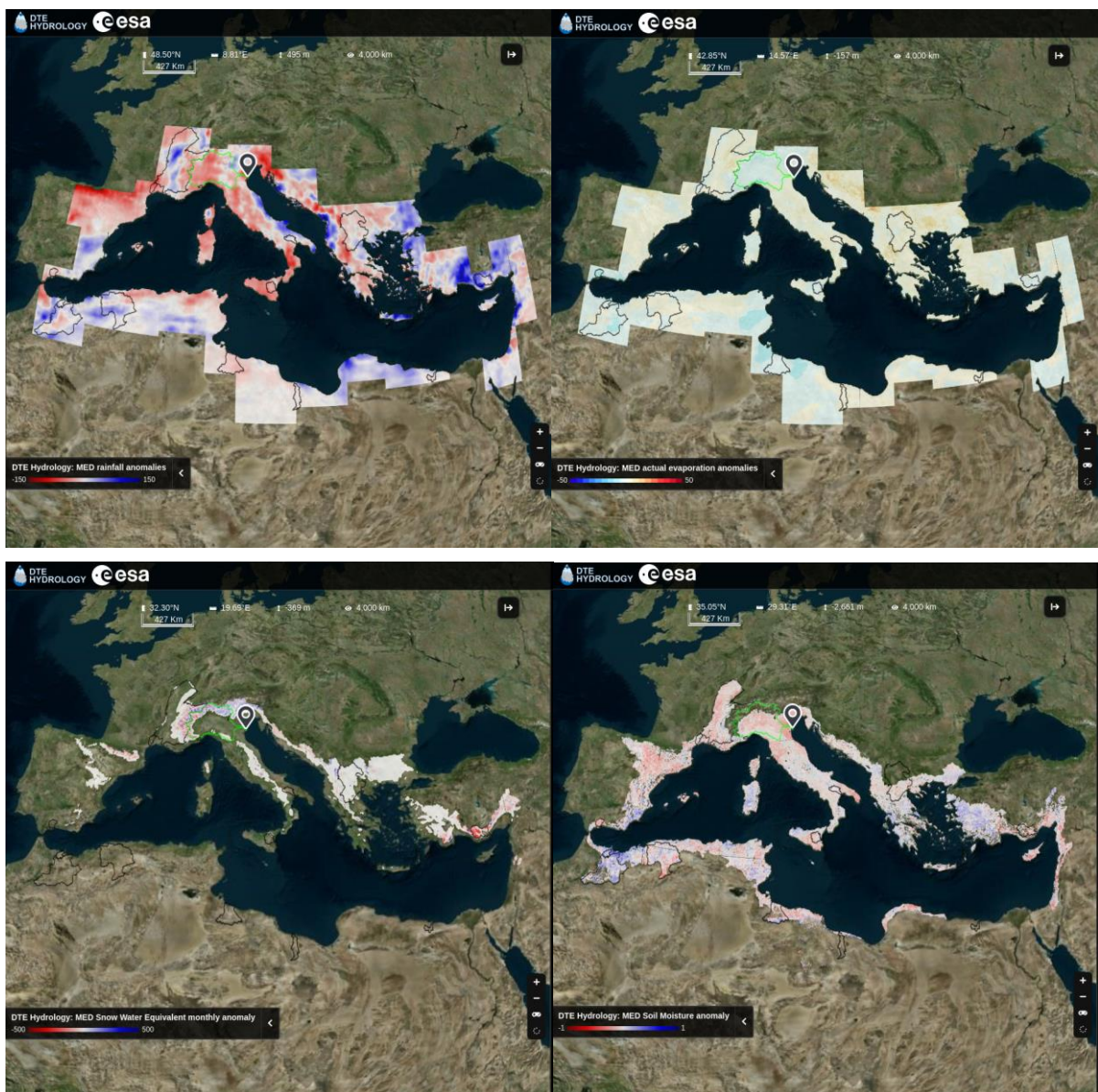
Figure 7: *Interactive map on different situations.*

From here, the user can select a basin to inspect, and the right part of the screen will react with the respective data (**Figure 8**).
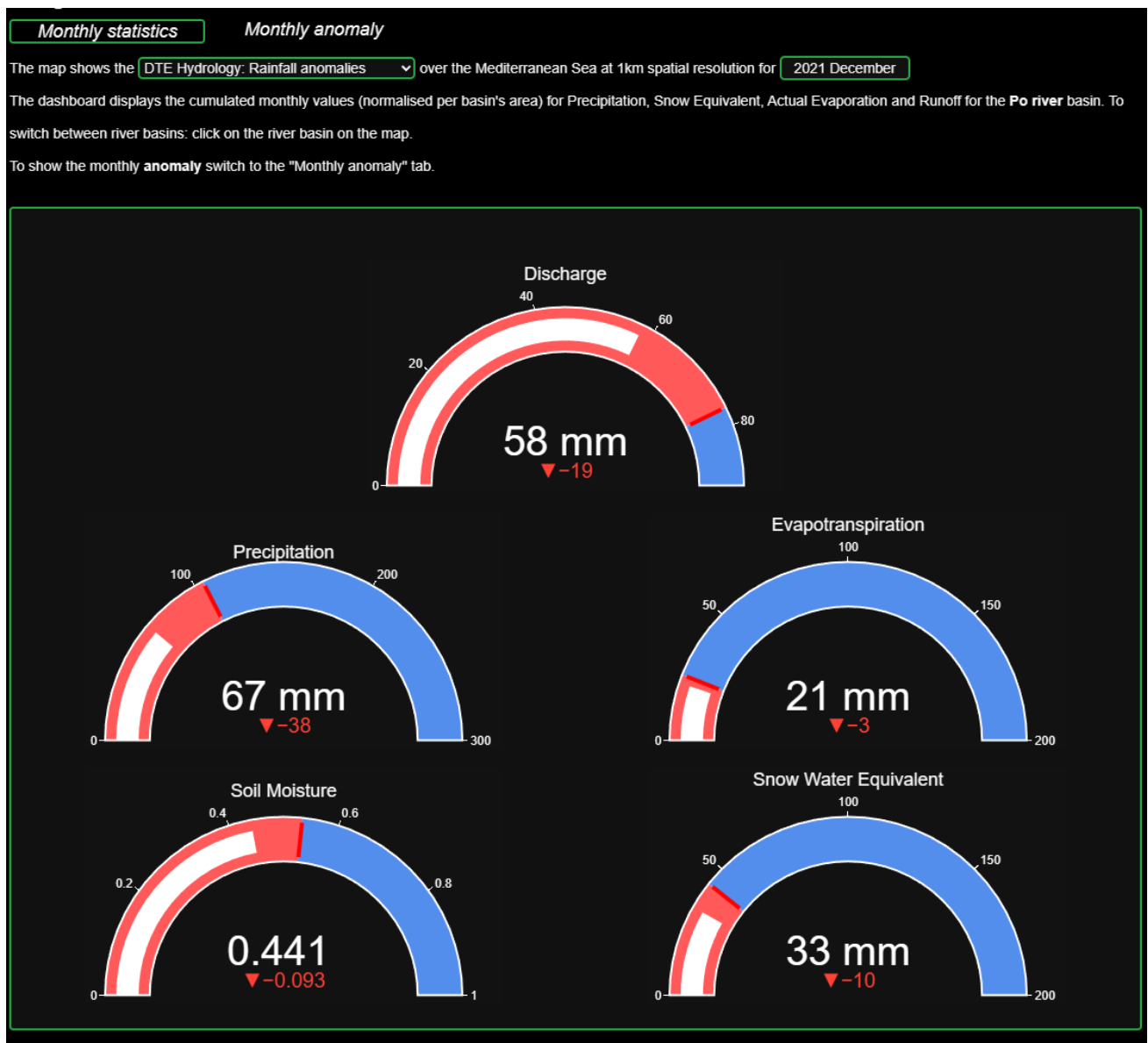
Figure 8: *An example of monthly statistics displayed for the rainfall anomalies dataset in Dec. 2021.*

The sidebar allows the user to change the current dataset and/or the time period. When any of these parameters change, the map updates as well with the respective data. It is also possible to switch from monthly statistics to monthly anomalies (**Figure 9**).
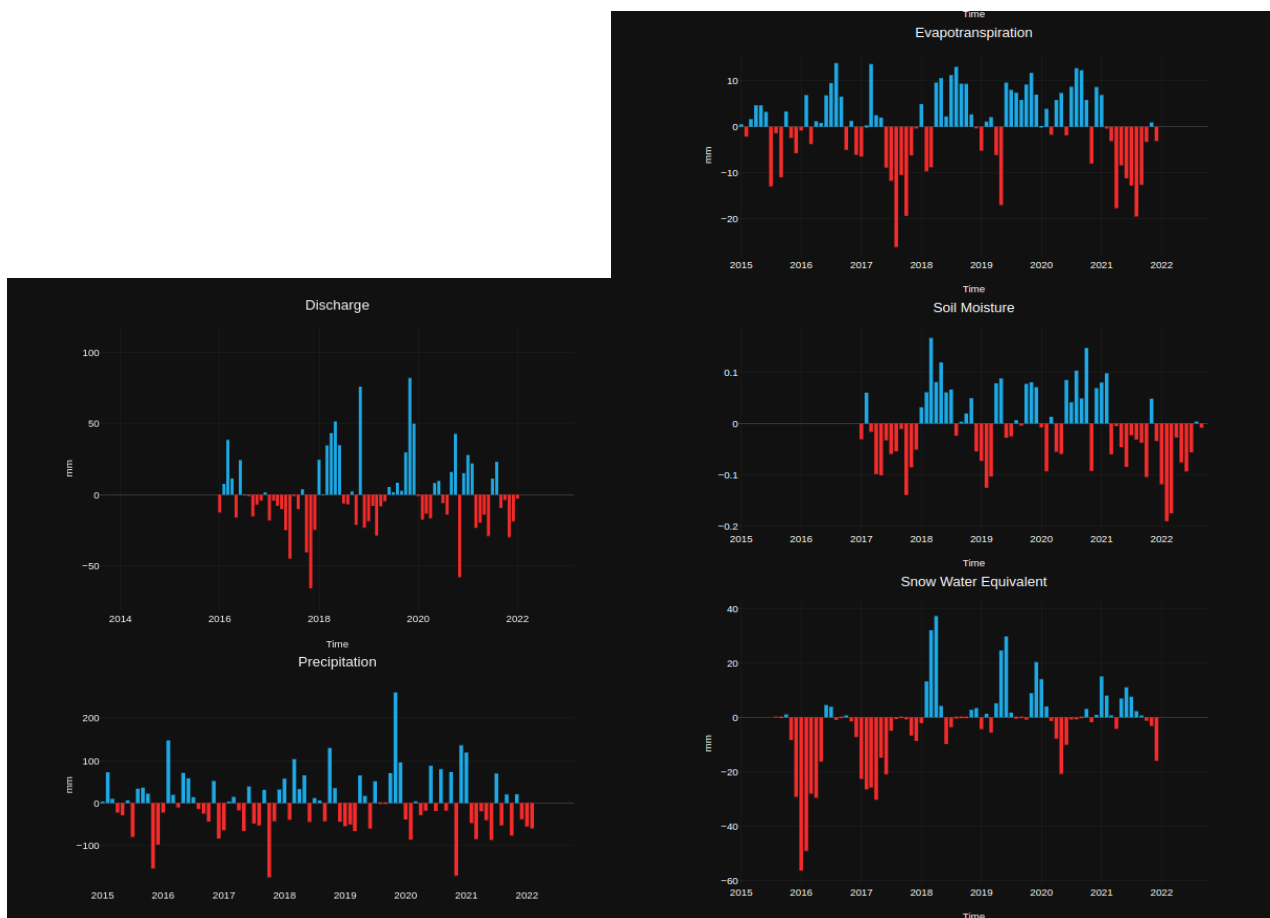
***Figure 9**: An example of plotted monthly anomalies for all the water balance components.*

3. Flooding prediction

3.1. The Apollo Medicane

In October 2021, a powerful Medicane, named Apollo, affected many countries on the Mediterranean coast, especially southern Italy, causing a death toll of seven people. A low-pressure area, centred over the Strait of Sicily, led to disturbed weather conditions on the southernmost regions of the Italian peninsula, especially on the Ionian sectors of Calabria and Sicily. The genesis of the Medicane was a low-pressure system which isolated

near the Balearic islands around 22 October 2021 and then moved towards the Algerian and Tunisian coastline, already producing damages on 23 October 2021. During its movement towards the southeast, it affected Sicily by producing self-regenerating thunderstorms around the Catania area in the late afternoon of 24 October 2021. It then moved towards Libya, but on 27 October, it changed direction, again approaching Sicily. Finally, on 29 October 2021, it moved towards the southeast, fading on 2 November off the Turkish coast. The cold air intrusion at 500 hPa caused a strong convective instability, and the wind convergence over eastern Sicily maintained the precipitation over the same area for hours. Moreover, the barotropic structure of the cyclone at this stage contributed to keep the phenomena over the same place. Between Sunday 24 and Tuesday 26 October 2021, the amount of rain recorded by the Regional Agrometeorology Service of Sicily (SIAS— http://www.sias.regione.sicilia.it/, see **Figure 10**) was 180 mm in Catania (CA), 280 mm in Lentini (LE), about 230 mm in Fabrizia (FA), and 520 mm in Linguaglossa-Monte Conca (LI). The latter is placed at an altitude of 1875 m on the northeastern side of Etna, and it is particularly prone to humid eastern winds because the mountain acts as an orographic barrier. The situation was correctly predicted and a red alert (the most Severe Weather Warning in the Italian system) was issued by the National Department of Civil Protection. Nevertheless, the runoff produced an extended flood over the Catanian plain with several casualties. The event hit, with harmful effects, the eastern side of Sicily and the south of the Calabria Region. The area is mainly characterized by several small coastal basins of a few hundred square kilometers with steep slopes and low concentration times. Within the area of interest it is located a medium-sized basin: the Simeto River, in

Sicily, with an area of about 4186 km² and an elevation that ranges from the sea up to a maximum of 3274 m asl.
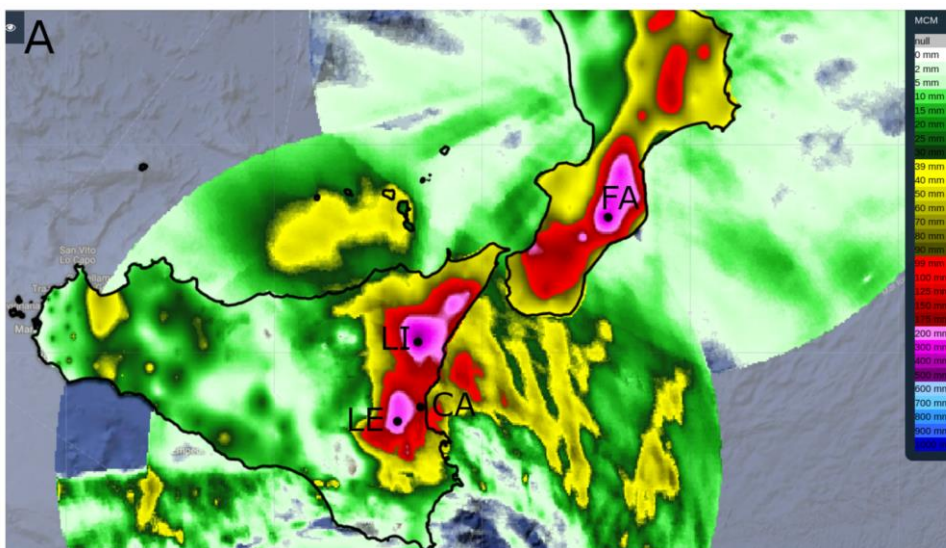
Figure 10: *Observed 48-h accumulated precipitation on 24–26 October 2021. Black dots with labels represent the rain gauges of Lentini (LE), Linguaglossa (LI), Catania (CA), and Fabrizia (FA). Adapted from* Lagasio et al. (2022).

3.2.   Hydrological simulations

Hydrological simulations for the Apollo Medicane event over the entire Sicily region are described in the project Deliverable D3.1, and include:

-    a 7-year simulation (2015-2021) forced by ERA5
-    a 6-year simulation (2016-2021) forced by the DTE EO precipitation data and ERA5 for other variables
-    a 1-year simulation (2021) forced by high-resolution ground-based data collected from the regional network

3.3.   Benchmark data

For this event, observed river discharge data was not available to assess the modeling performance. However, inundation extent was estimated from satellite imagery acquired by Sentinel 1 on 25 October 2021, through the AUTOmatic Water Areas DEtector (AUTOWADE, Pulvirenti et al., 2021, see **Figure 11**). The inundation extent in the Catania Plain was estimated at 69km² with the AUTOWADE algorithm. In addition, flood depths were estimated using the Google Earth Engine implementation of the Floodwater Depth Estimation Tool

(FwDET-GEE, Peter et al., 2022) and the Digital Elevation Model (DEM) provided by the Sicily Region at 2m resolution (**Figure 12**).
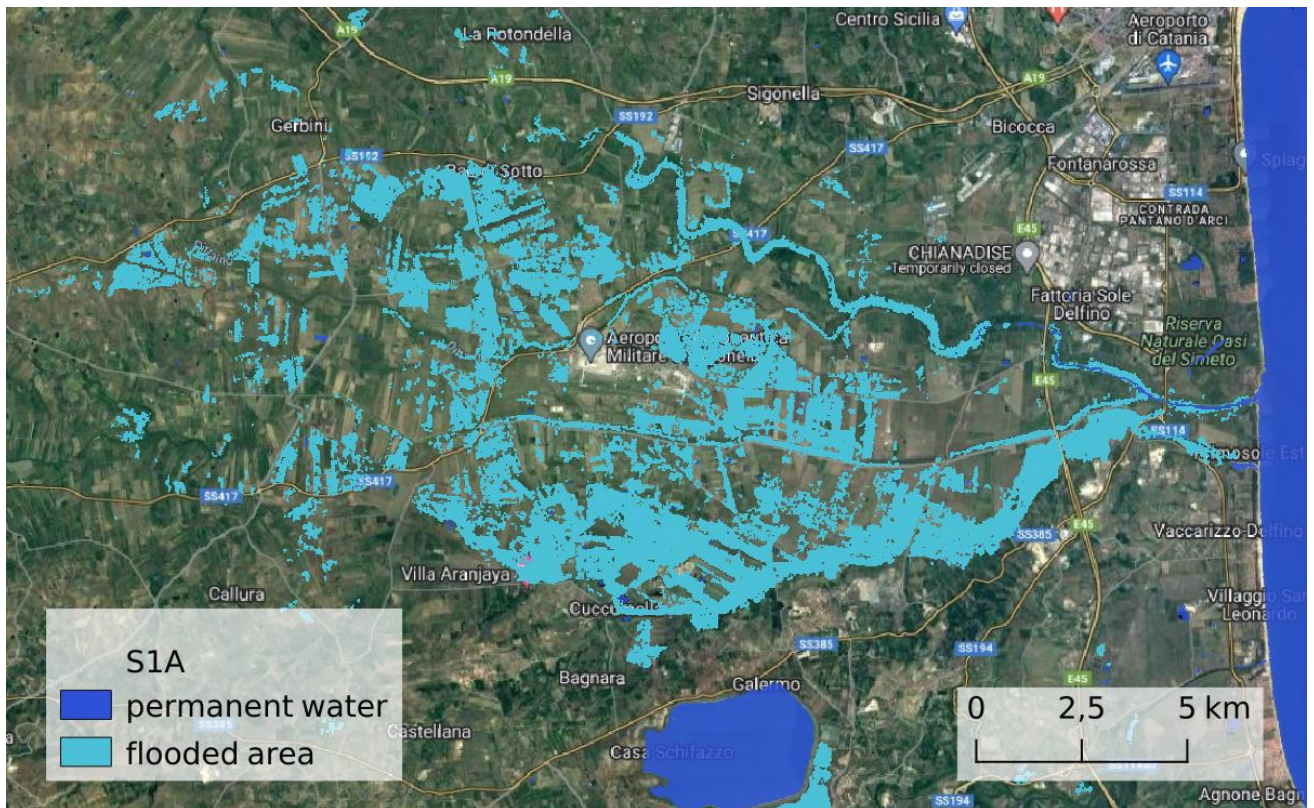
Figure 11: Inundation extent (in cyan) in the Catania Plain estimated with AUTOWADE using the Sentinel 1 acquisition on 25 October 2021.
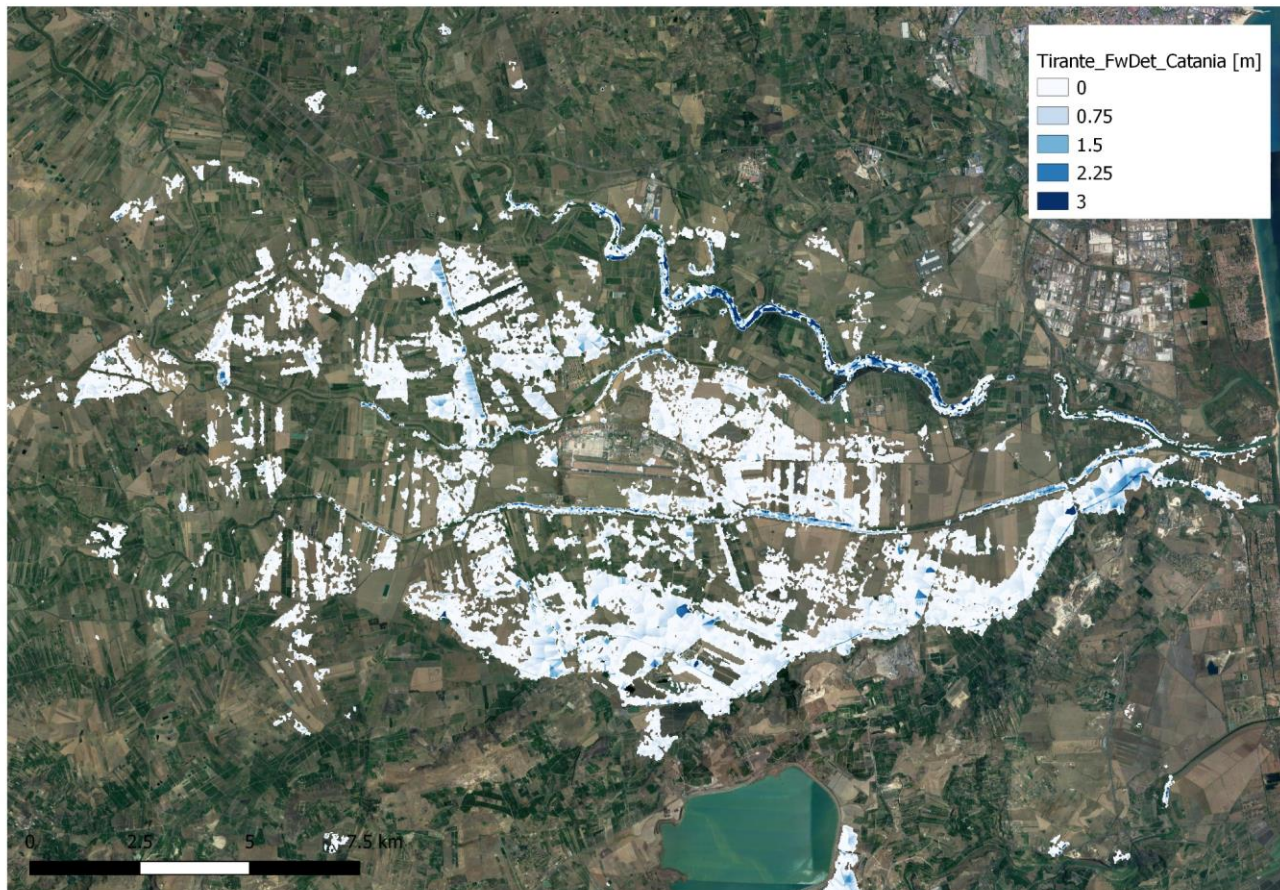
Figure 12: *Flood depths of the Sentinel 1 acquisition on 25 October 2021, estimated with FwDET-GEE.*

3.4.   Hydraulic simulations

Hydraulic simulations were performed with the hydrodynamic model TELEMAC-2D, described in the project Deliverable D1.1. We ran two configurations: one forcing TELEMAC-2D with the flood hydrographs derived from the high-resolution ground-based data (hereinafter MCM); the second forcing the model with the DTE EO precipitation data (hereinafter DTE). Each simulation covered the most intense 3 days of the event and took about 3 days to run in CIMA's servers, using a parallel implementation distributed on 90 CPUs.

The maximum flood extent simulated in the two configurations are shown in **Figure 13**, while contingency maps of the two simulations versus the Sentinel 1 observed inundation are shown in **Figure 14** and **15**.

Results denote a relatively good representation of the observed inundation extent in the proximity of the main river network, yet with a marked underestimation of inundation extent in the case forced by DTE input. This is due to a considerably lower discharge peak, which at the outlet of the Simeto was simulated at 540 m^3/s in the DTE-driven simulation, while it was 1500 m^3/s in the MCM-driven simulation. We use the Critical Success

Index (CSI) as a performance indicator of the modelled simulation extent. The CSI was chosen because it penalises both false alarms and missed events, yet without considering the True Negative values which would bias the resulting performance score. The resulting CSI is 0.35 for the MCM simulation, while it is 0.28 for the DTE simulation, confirming the visual impression of the figures. Both simulations do not capture the observed inundation in the northern part of the plain, enclosed between the Simeto and the Dittaino rivers. This was likely caused by surface water flooding in the area due to heavy rainfall rates.

MCM   **DTE**
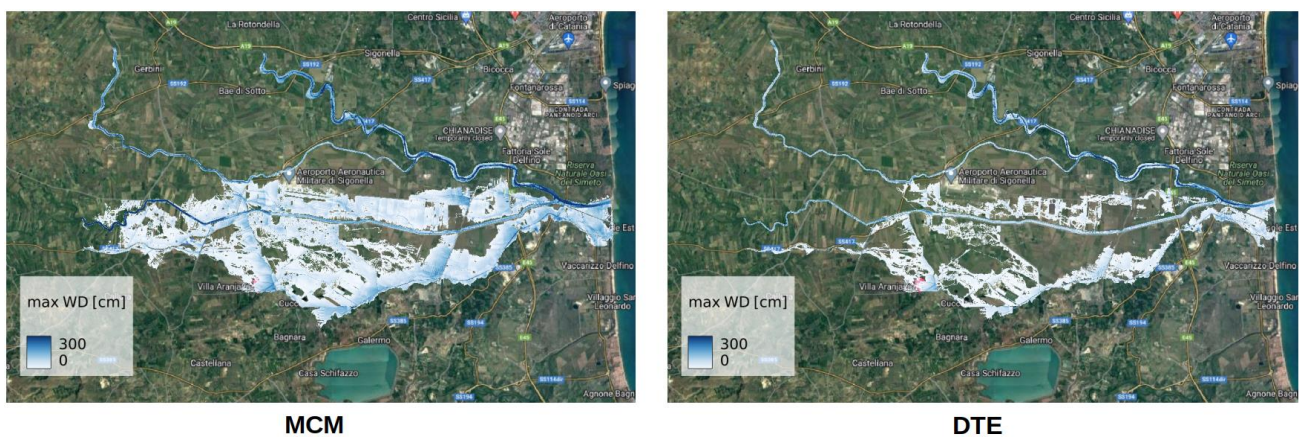
Figure 13: *Maximum flood extent simulated with TELEMAC-2D forced with MCM (left) and DTE (right) input.*
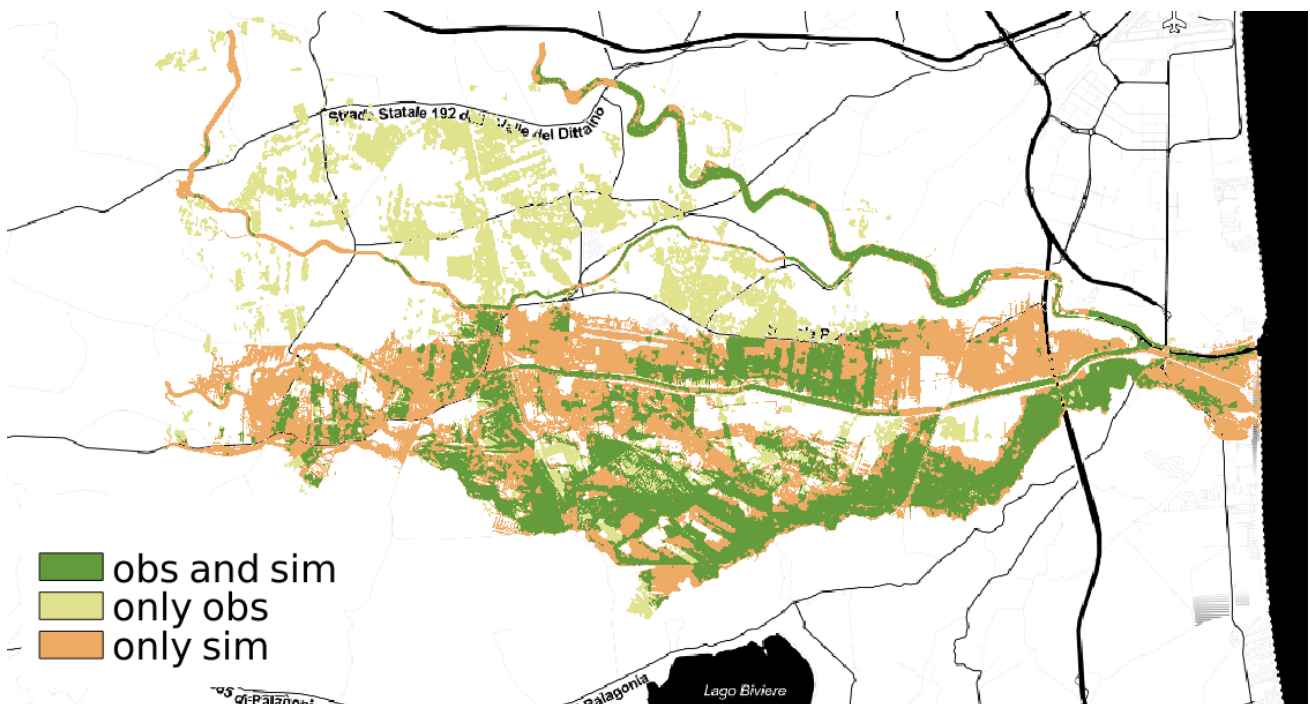
Figure 14: *Contingency map of MCM-driven hydraulic simulation versus Sentinel 1 observed inundation.*
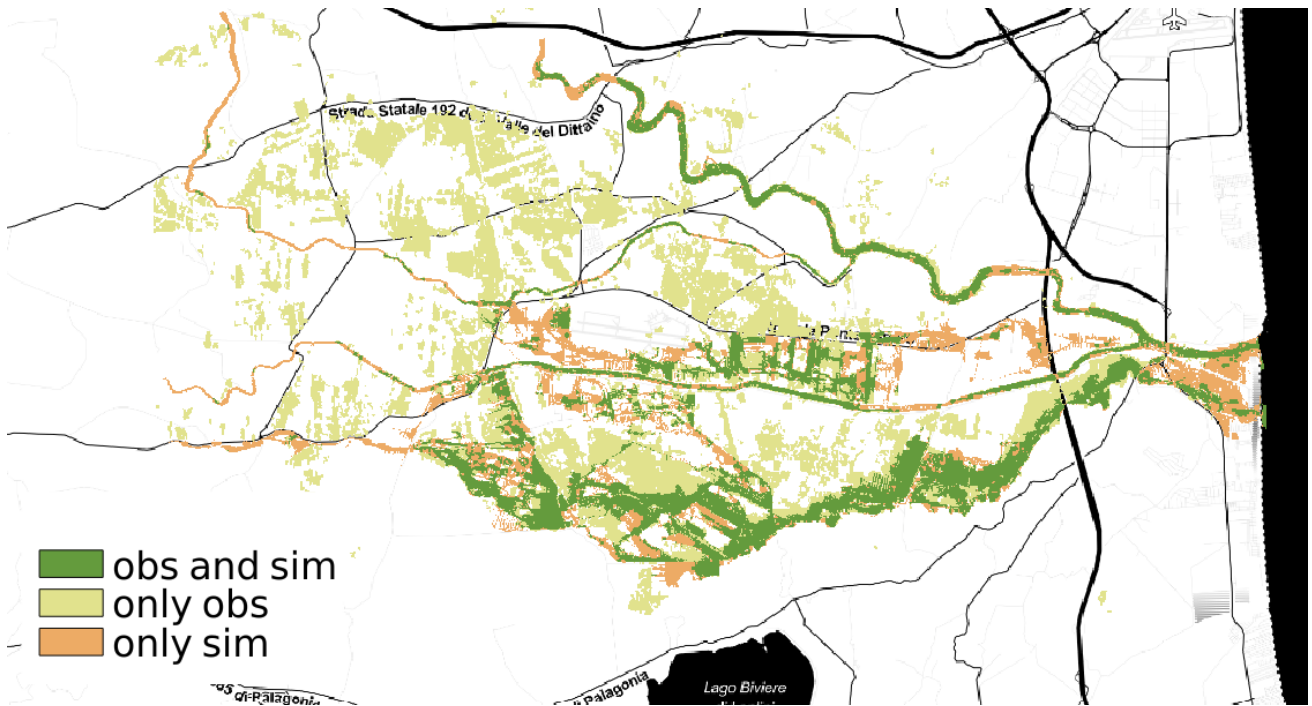
Figure 15: *Contingency map of DTE-driven hydraulic simulation versus Sentinel 1 observed inundation.*

3.5.   Web application

Through the DTE platform it is possible to inspect the flooding simulations for extreme events. Each event can be selected from the floating panel on the left (see **Figure 16**).
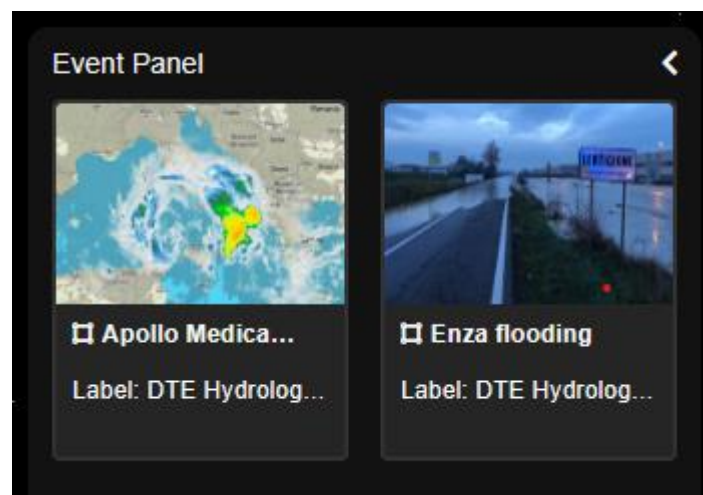
Once an event is clicked, the globe will start loading the scenario, animating it over time (**Figure 17**).

Figure 17: *The event's data loaded on the map.*

The "video" can be paused/resumed using the time bar on the bottom-right corner of the screen.

4. What-if scenario for flood risk assessment

The what if scenario for flood risk assessment and water resources management was developed for the Po river, which has been affected by flood and drought events on many occasions over time. In the last 10 years, the Po river basin has experienced at least three major flood events in November 2016, October 2018 and October 2019, and droughts in the summer of 2020, 2021 and 2022, with huge economic consequences and fatalities. A simple tool, such as a "what-if" scenario analysis could help stakeholders, local authorities and citizens be prepared for managing extreme hydrological events.

Before describing the procedure for flood risk assessment and its results, a few words will be spent describing what a what-if scenario analysis is. What-if scenario analysis is a data-intensive simulation that aims to explore the behaviour of a complex system under a set of given hypotheses (called scenarios). More pragmatically, what-if analysis measures how changes in a set of independent variables affect a set of dependent variables with respect to a given simulation model (Rizzi, 2009). Several steps are required to carry out the analysis:

1) <u>scenario definition</u>: identification of the key variables that control the phenomena and definition of their range of variability;

2) <u>model simulation</u>: building and calibrating a model on past events; running the model for different scenarios;

3) <u>prediction</u>: understanding how the system will evolve over time as a consequence of a hypothetical action.

In the following, the procedure adopted to build the what-if scenario for flood risk assessment will be detailed according to the three steps listed above.

4.1. Scenario definition

The variables controlling the flood formation process are many and are related to climate (air temperature, total precipitation, precipitation intensity/duration, snow cover and snowmelt, short and long-wave radiation), the catchment (infiltration capacity, runoff coefficient, water storage capacity, evapotranspiration) and the river characteristics (morphology, conveyance, roughness, water level, runoff, floodplain storage, river channel vegetation). The interaction of the three compartments influence the characteristics of river floods; the relative importance of each driver depends on the local situation and on the boundary conditions.

In the Po basin, river floods are mainly caused by extreme precipitation and excess soil moisture (Berghuijs et al., 2019) and in particular, the characteristics of these two variables in terms of magnitude and duration as well as spatial/temporal distribution define the characteristics of a flood event. Therefore, for the scenario definition, the following procedure was adopted (see **Figure 18**):

1. Step 1: analysis of observed river discharge data and extraction of flood peak events. A Peak Over Threshold (POT) analysis was carried out for the time series recorded at Pontelagoscuro station (outlet section of Po river) to extract the most significative flood events occurred over the Po river in the analysis period (2015-2021, see **Figure 18a**);

2. Step 2: extraction of precipitation maps for each flood event. For that, the high-resolution precipitation data derived from Earth Observations (described in Deliverable D1.1) were used;

3. Step 3: identification of initial soil moisture conditions before the flood event. For that, RT1 Sentinel-1 data were used. For each flood event, the spatial variation of soil moisture over the Po river basin was analysed and the initial soil moisture map corresponding to average (0.5) soil moisture condition was taken as a benchmark for the entire basin. This soil moisture map was multiplied in a range between 0.5 to 1.8 to represent all the possible soil moisture scenarios from dry (0.2) to wet (0.9) conditions.
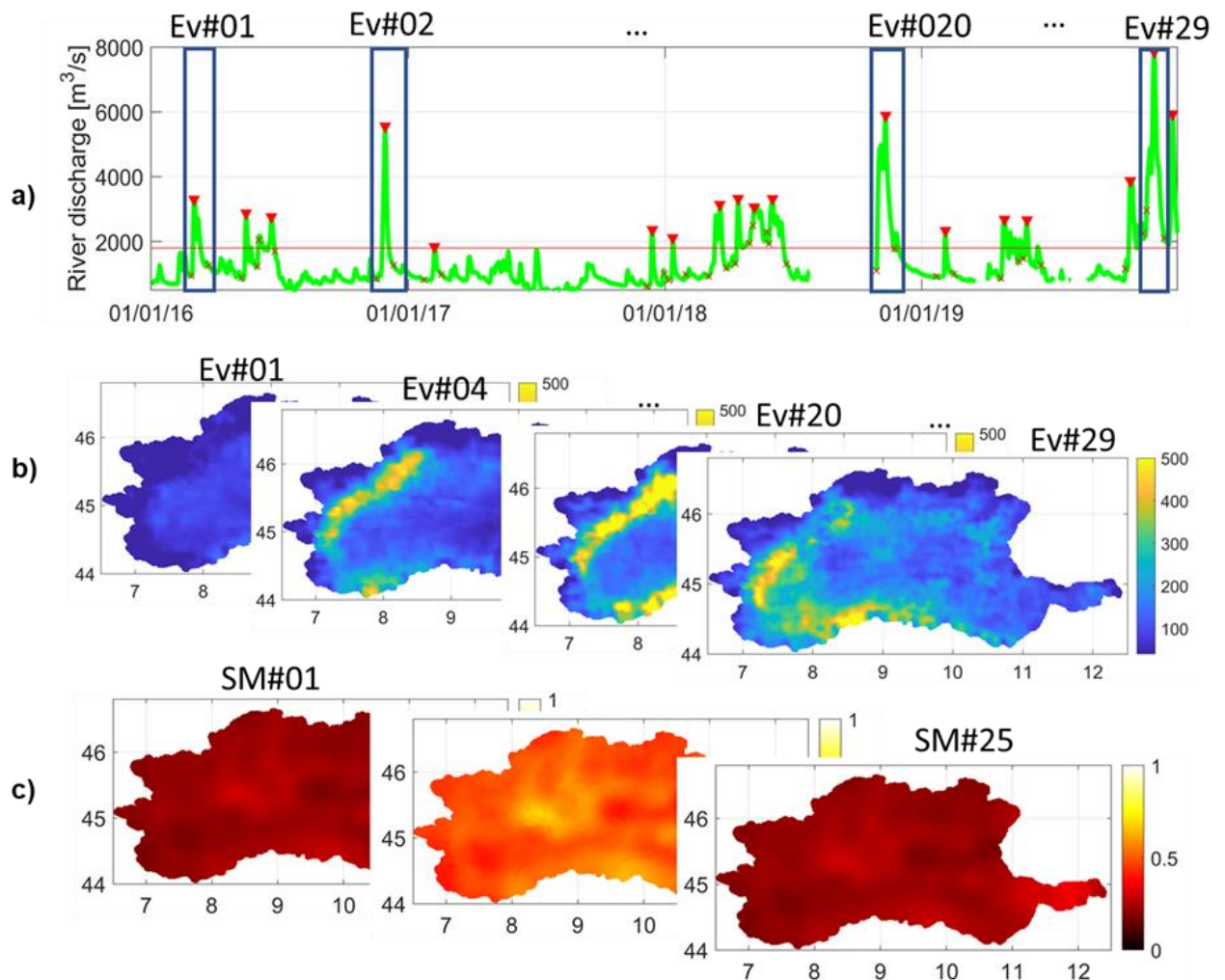
Figure 18: *Steps needed for the flood scenario definition: a) analysis of observed river discharge data and extraction of flood peak events, b) extraction of precipitation maps for each flood event; c) definition of initial soil moisture maps from satellite data.*

For that, the mean soil moisture conditions of the entire basin were varied in the range 0.2 to 0.89. Similarly, average precipitation data cumulated over three days ranged in the interval 36 mm to 189 mm. A total of 25 initial soil moisture conditions and 29 cumulated precipitation events were collected (see **Figure 18b** and **c**) and run through the hydrological model, to obtain 725 flood scenarios.

4.2.   Model simulation

The modular model MISDc, described in the Deliverable 3.1 "Technical Note 3: DTE Hydrology modelling experiments, validation and assessment", was used to run the 725 combinations of initial soil moisture

conditions and accumulated precipitation events. In particular, among the modules, the snow, soil and routing module (see **Figure 19**) were coupled to provide accurate flood estimates both at the outlet and the inner sections of the Po river. For more details about the calibration results of the MISDc model, the reader is referred to Deliverable 3.1.
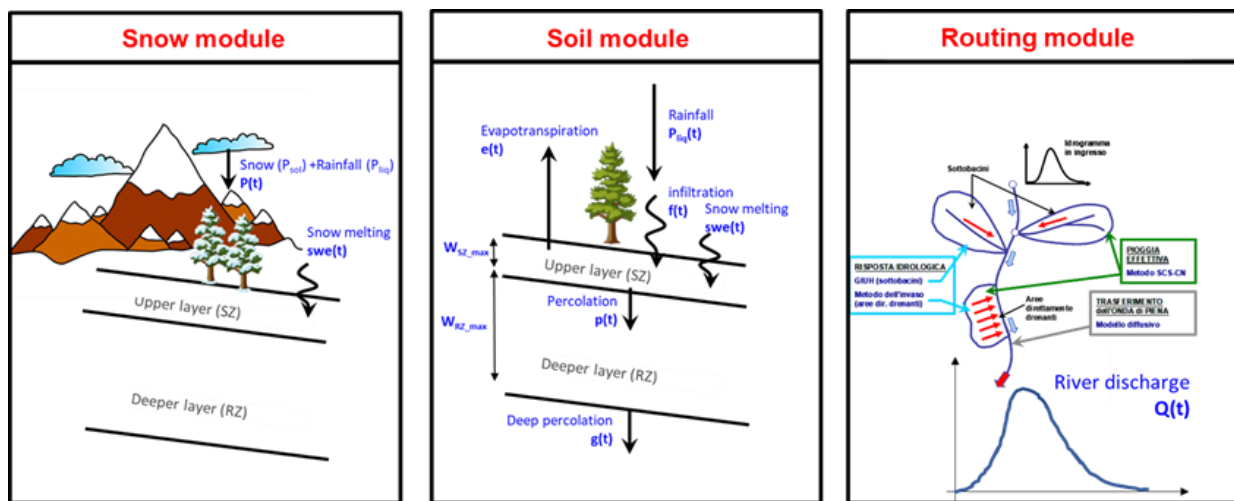
Figure 19: *Structure of the modular MISDc hydrological model adopted for the what-if scenario analysis for flood risk assessment.*

4.3. Predictions

By running the hydrological model, a database of 725 flood hydrographs can be obtained for each value of precipitation and initial soil moisture conditions, for different sections along the Po river. **Figure 20** illustrates the ensemble of flood hydrographs observed and simulated for the sections of Casale Monferrato, Piacenza, Cremona and Pontelagoscuro. The 5th, 50th and 95th percentiles are evaluated for both the ensembles. As it can be noted, for all the analysed sections the simulated and the observed flood hydrographs are in good agreement and observed statistics are well represented by the simulated data. The same analysis has been carried out for other two sections along the Po river (Torino Murazzi, Borgoforte) with similar results (not reported here for sake of brevity).
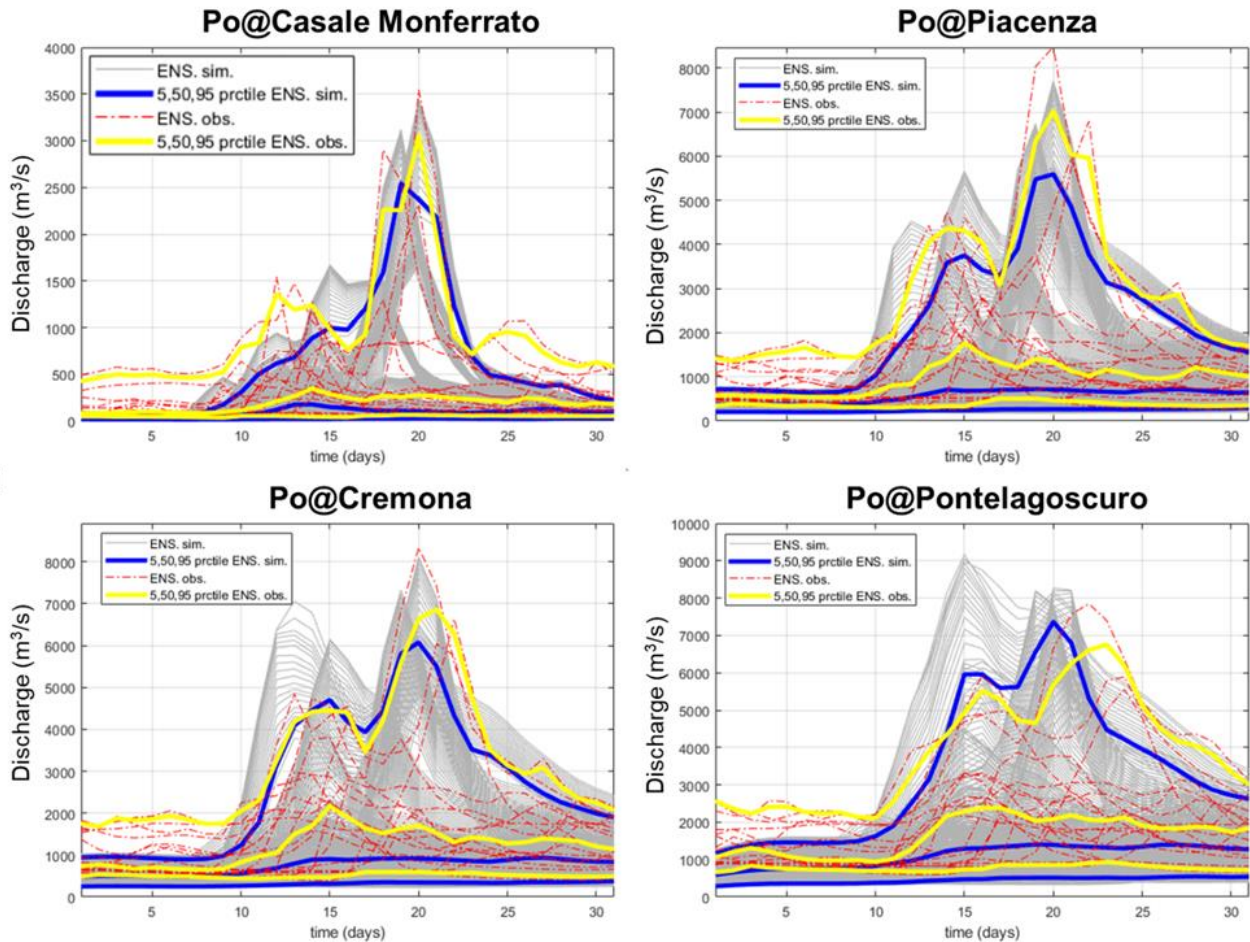
Figure 20*: Flood hydrographs observed (red lines) and simulated (gray lines) for the sections of Casale Monferrato, Piacenza, Cremona and Pontelagoscuro. Yellow and blue lines represent the statistics on the observed and simulated flood hydrographs.*

To additionally validate the flood database, past flood events occurred in the Po basin were investigated. Specifically, by knowing the value of occurred precipitation amount and the initial soil moisture conditions the database can be queried to extract the occurred flood hydrograph. The characteristics of some recent flood events observed in the sections of Casale Monferrato, Piacenza and Pontelagoscuro are listed in **Table 1**. In particular, for each event, the cumulated precipitations, the soil moisture conditions prior to the event and the observed peak discharge value is reported.

Table 1*: Characteristics of recent flood events observed in the sections of Casale Monferrato, Piacenza and Pontelagoscuro. For each event, the cumulated precipitations (P), the soil moisture conditions prior to the event (SM$_{ini}$) and the observed peak discharge value is reported.*

sections	Casale Monferrato			Piacenza			Pontelagoscuro		
event	P (mm)	SM_{ini} [-]	Q_{maxobs} (m³/s)	P (mm)	SM_{ini} [-]	Q_{maxobs} (m³/s)	P (mm)	SM_{ini} [-]	Q_{maxobs} (m³/s)
30 November 2016	-	-	-	-	-	-	130	0.5	2000
7 November 2018	-	-	-	-	-	-	164	0.25	7800
24 October 2019	-	-	-	-	-	-	171	0.25	5000
22 November 2019	177	0.53	2300	186	0.48	6000	178	0.55	8000

By querying the flood database with the precipitation and soil moisture values of **Table 1**, the flood hydrographs in **Figure 21** can be obtained. Dark and light blue areas identify the 50 and 90 confidence bounds, the red and the dark dashed lines represent the observed and the median flood hydrograph, respectively. It can be noted that, for all the analysed events, the observed hydrograph is within the 50% confidence bound and follows quite well the observed hydrographs.
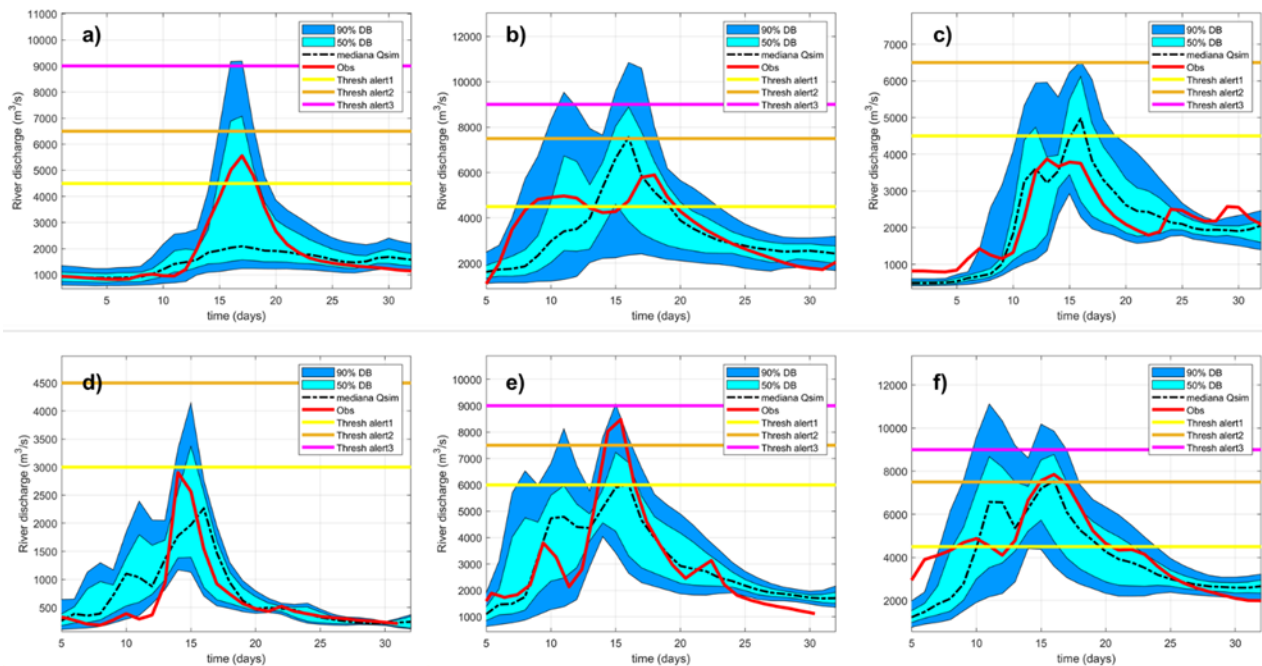
Figure 21: *Flood events occurred in the Po river on: a) 30 November 2016 at Pontelagoscuro, b) 7 November 2018 at Pontelagoscuro c) 24 October 2019 at Pontelagoscuro, d- e- f) 22 November 2019 at Casale Monferrato, Piacenza e Pontelagoscuro. Dark and light blue areas identify the 50 and 90 confidence bounds, the red and the dark dashed lines represent the observed and the median flood hydrograph, respectively.*

The tool is very efficient because, for example, by knowing the value of the expected rainfall (e.g. from a meteorological forecast) and the initial soil moisture conditions (e.g., from satellite observations), the database can be queried and the expected flood hydrograph can be extracted in a few seconds without running a hydrological model. The tool does not require any expertise to operate and can be easily accessed by stakeholders, students or citizens.

4.4.   Web application

The DTE Platform provides the flood hydrographs over the Po river's basin for 25 initial soil moisture conditions and 29 accumulated values of precipitation (see **Figure 22**). On the left, the map shows the selected initial conditions (soil moisture at the surface, precipitation at the top level) and respective alerts for 6 gauging stations along the river (Torino Murazzi, Casal Monferrato, Piacenza, Cremona, Borgoforte and Pontelagoscuro). The green markers represent low alert (0—4500), the yellow ones represent medium alert (4501—7000) and the red ones represent high alert (7001+).

On the right, the hydrograph displays the ensemble of river discharge on the picked station. It is possible to switch between stations by simply clicking over their marker on the map, while the initial soil moisture/precipitation initial condition's values can be edited in the respective fields in the right sidebar.
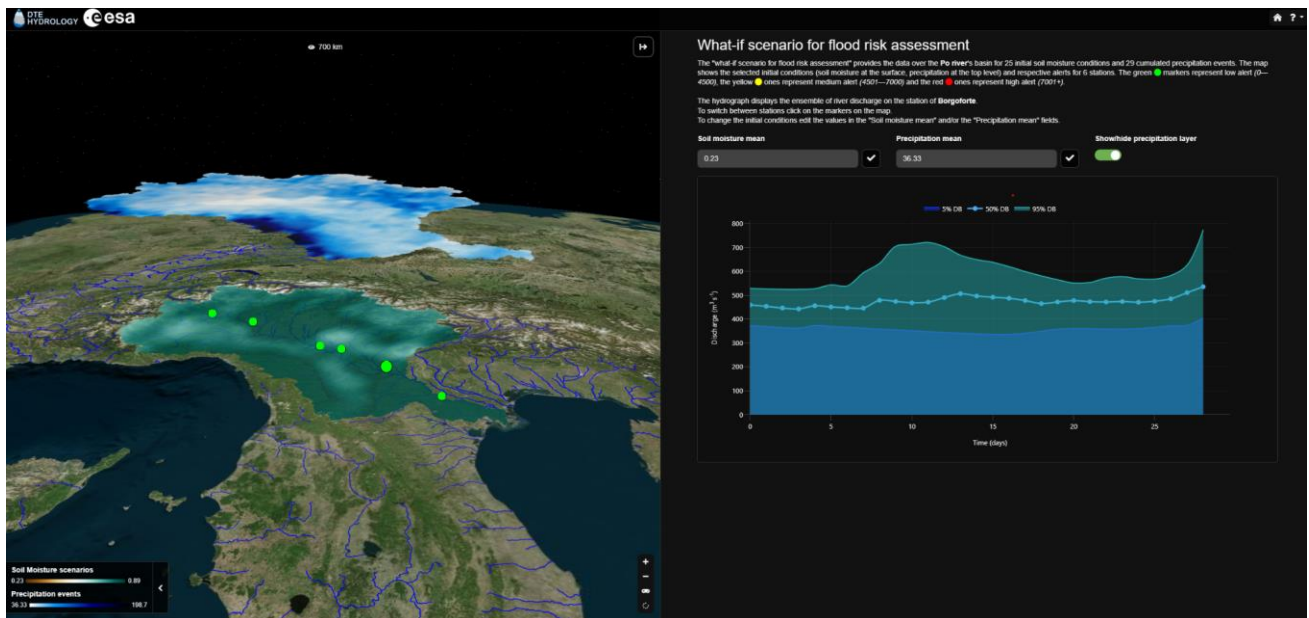
Figure 22: *What-if scenario for flood risk assessment over the Po river basin. On the left, the soil moisture and the precipitation maps are illustrated. On the right, the flood hydrographs for selected initial soil moisture and accumulated precipitation values are represented.*

When the data changes, the marker's color reacts accordingly.

5.   What-if scenario for water resources management

The Po river basin has been affected in the last years by huge drought events. Precipitation data analysis highlights that the average summer quantity of rain has considerably diminished in the Po River Basin in the last decades (Musolino et al., 2018). In the meantime, an increase in rainfall intensity (and, consequently, in the number of floods) was observed. The number of rainy days decreased, particularly in spring and summer, resulting in the decrease of river flow during the dry season (the worst reduction occurred in the years between 1980 and 2010). Moreover, since 1960, the average yearly temperature has increased by about 2°C, raising the water needs for agriculture and power generation (due mainly to air conditioning). Analysing the distribution of the water volume in the basin throughout the year, it appears that the whole storage capacity,

distributed in natural regulated alpine lakes and in artificial reservoirs, is enough to cover the water demand (agriculture, industry, power production, domestic uses) in absence of rainfall until the end of June. A dry period in spring, from March to April, is usually sufficient to trigger the activation of extraordinary water management measures for the next irrigation season, unless there is a successive occurrence of rainfalls. In the same way, a dry period during the start of irrigation season (April, May, and June) causes early release of all the spare accumulated water resources. The major drought events experienced in the last decades in the Po River Basin occurred in 2003, 2005/2007, 2012, and 2015, 2021 and 2022. In particular, the 2021-2022 winter has been one of the driest winters Italy experienced in the last 65 years; the situation has been more acute in the northern regions. As a result, the Po river recorded the lowest water levels in the last 16 years over several stations; water resources stored in lakes and snow in the Alps and Apennines are 43% lower than the 2006-2020 average. In this situation, stakeholders and policy makers are called to choose among different alternatives to cope with water shortages. A what-if scenario analysis could help stakeholders make the right choices to face the dramatic drought situation that the Po basin is continuously experiencing.

5.1. Scenario definition

Among the variables that control the availability of water resources in the Po basin, precipitation and air temperature are the key climatic variables, while soil moisture, snow water equivalent and reservoir levels are the key variables that control the initial state of the basin. For all these variables, the monthly values for the summer period (April to September) between 2015 and 2021 were evaluated on the basis of the available data. In **Figure 23**, the black lines and the black bars represent, for the different variables, the average values for the analysis period. Starting from these, and on the basis of their range of variability obtained from the observations, the variables were multiplied by different coefficient values to obtain 5 different scenarios. Specifically, the precipitation amount was varied between ±25/50%, air temperature between ± 30/50%, soil moisture between ± 10/30%, snow water equivalent between ± 30/60% while for the reservoirs only the average storage values were considered. In this way, a total of 625 scenarios can be explored, varying the state of each variable between very low, low, average, high and very high conditions.
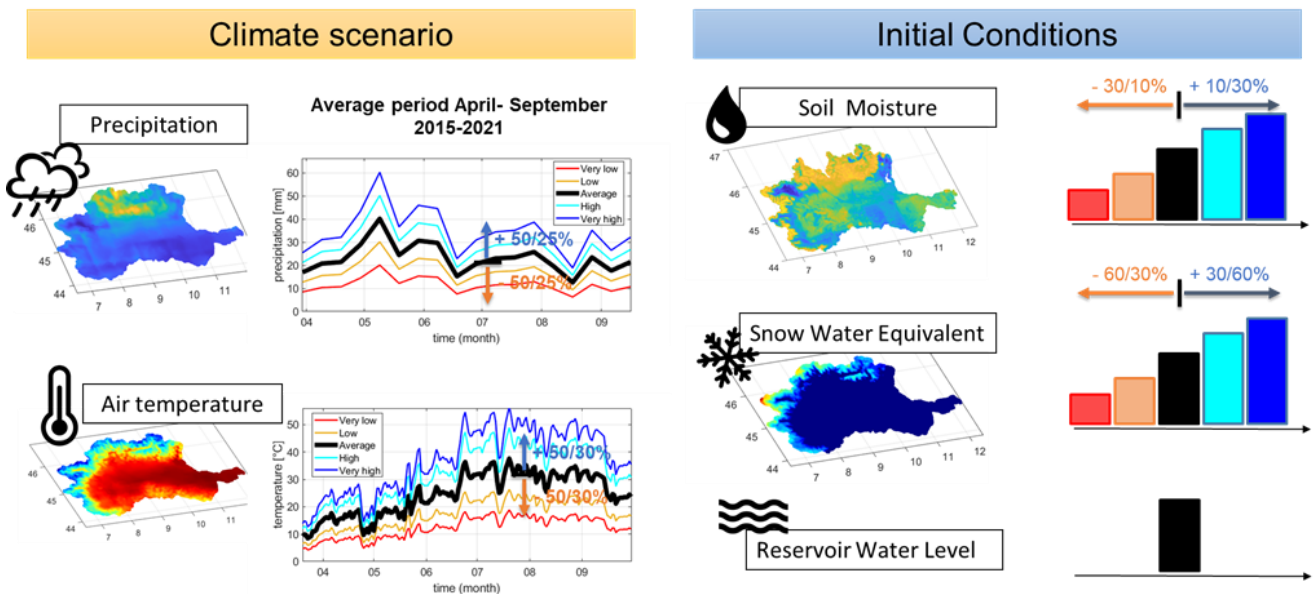
Figure 23: Controlling variables for the water resources management what-if scenario.

5.2. Model simulation

The modular model MISDc, described in the Deliverable 3.1 "Technical Note 3: DTE Hydrology modelling experiments, validation and assessment", was used to run the 625 combinations of precipitation, air temperature, initial soil moisture and snow water equivalent. In particular, among the modules, the snow, soil, irrigation, water uses and routing modules (see **Figure 24**) were coupled to provide accurate estimates of the fluxes (river discharge and evapotranspiration), the storage conditions (the soil water in the surface, in the root zone and the snow water equivalent) and the water can be used for agricultural, civil and industrial activities both at the outlet and the inner sections of the Po river. For major details about the calibration results of the MISDc model, the reader is referred to Deliverable 3.1.
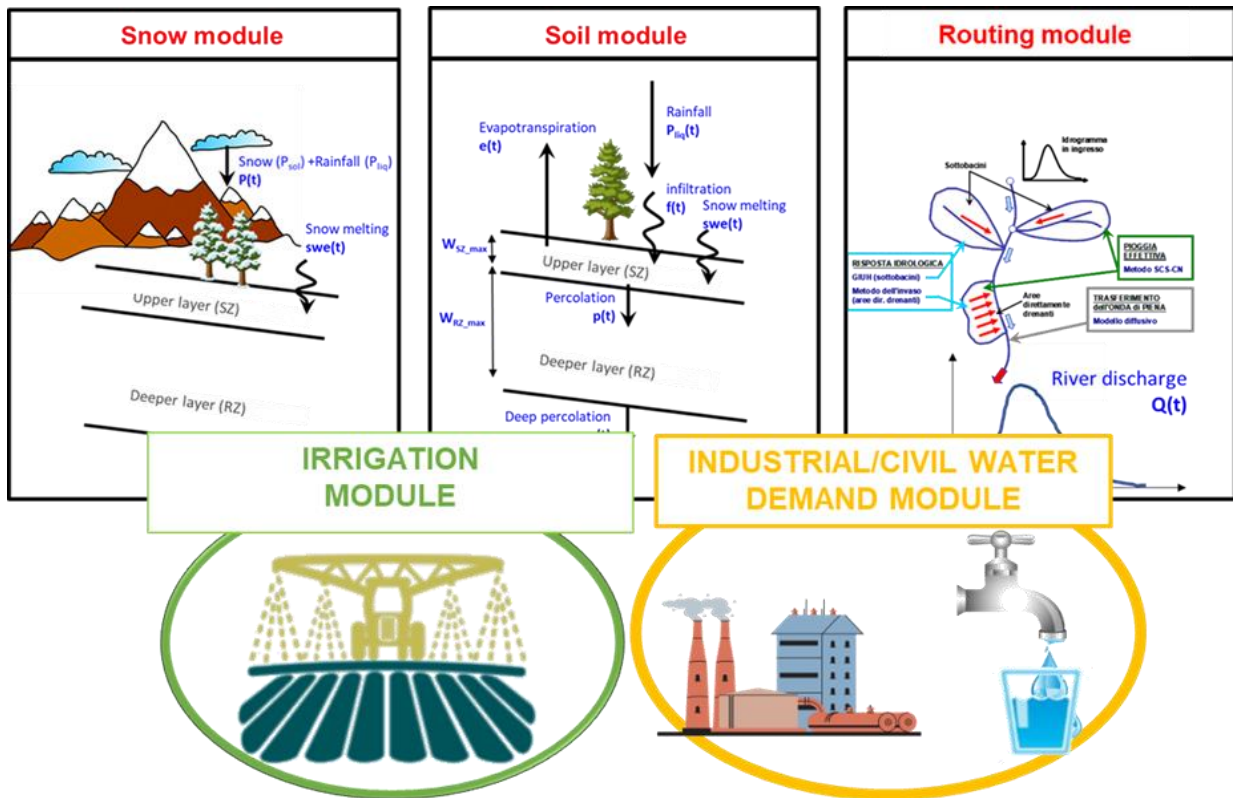
Figure 24: Structure of the modular MISDc hydrological model adopted for the water resources management what-if scenario analysis.

5.3. Predictions

By running the hydrological model, a database of 625 basin states can be obtained for the different scenarios of precipitation, air temperature, initial soil moisture and snow water equivalent conditions, for different sections along the Po river. As an example, **Figure 25** represents the what-if scenario for the next summer period obtained by querying the database on the base of the seasonal forecast of precipitation and air temperature and the initial conditions of the basin in terms of soil moisture and snow water equivalent. Specifically, the seasonal forecast of the European Centre for medium-range weather forecasts (ECMWF) predicts that the next summer will be sunny and dry, with cumulative precipitation below average and air temperature above average (but not very high). In terms of initial conditions, the soil is almost dry at the time of writing, with very low soil moisture and very little snow (see **Figure 25**, "Snow Water Equivalent" graph). As a result, it is expected that about 1.2 and 1.7 km^3 of water will be available for industrial and civil purposes while 20 km^3 can be used for agricultural purposes. In these conditions, a water deficit of about 1.3 km^3 will be recorded between May to September.
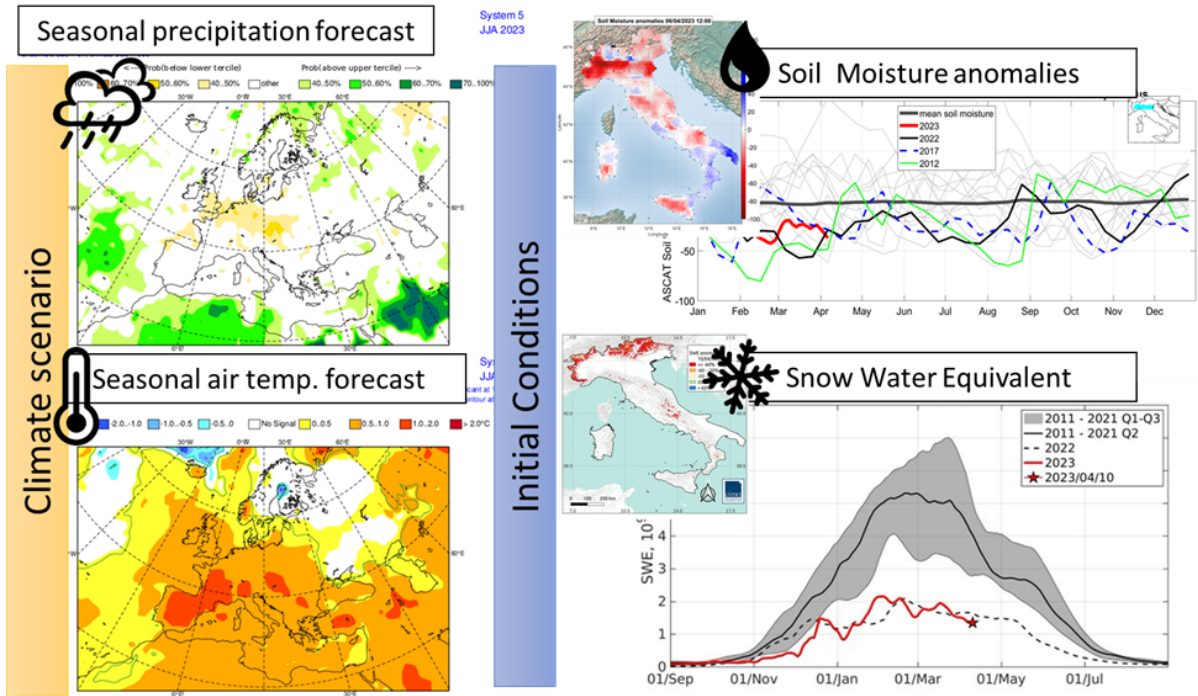
Figure 25*: What-if scenario for the next summer period based on the seasonal forecast of precipitation and air temperature and the initial state of soil moisture and snow water equivalent .*

5.4. Web application

The DTE platform provides the agricultural, civil and industrial water uses over the Po river basin from April to August as a function of different scenario (Very high, high, ..., very low) for precipitation, air temperature, initial soil moisture, initial snow water equivalent, and releases from reservoirs (currently only average).

To show the impact of the water resources management, agricultural, civil and/or industrial water uses can be toggled by clicking the respective green buttons on the top half of the sidebar (see **Figure 26**):
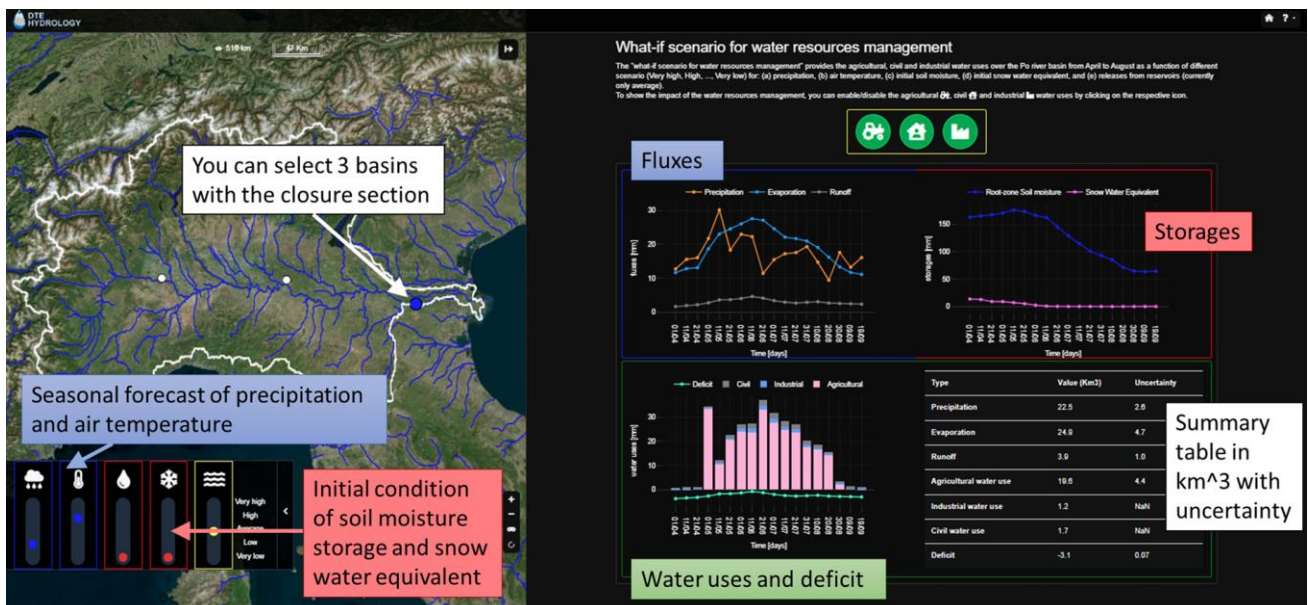
Figure 26: *What-if scenario for water resources management over the Po river basin. On the left the boundary of the basin at Casale Monferrato, Cremona and Pontelagoscuro is highlighted. On the right, the fluxes, the storages and the water uses are evaluated for different scenarios of precipitation, air temperature, initial soil moisture and snow water equivalent.*

The sliders on the bottom-left part of the screen will control the available scenarios level, and the plots and table on the right sidebar will react accordingly. It is also possible to switch over the inspection areas within the basin by clicking on the white dots (blue is selected).

6.   References

Avanzi, F., Ercolani, G., Gabellani, S., Cremonese, E., Pogliotti, P., Filippa, G., Morra di Cella, U., Ratto, S., Stevenin, H., Cauduro, M., and Juglair, S.: Learning about precipitation lapse rates from snow course data improves water balance modeling, Hydrology and Earth System Sciences, 25, 2109–2131, https://doi.org/10.5194/hess-25-2109-2021, 2021a.

Avanzi, F., Gabellani, S., Delogu, F., Silvestro, F., Cremonese, E., Morra di Cella, U., Ratto, S., and Stevenin, H.: S3M 5.1: a distributed cryospheric model with dry and wet snow, data assimilation, glacier mass balance, and debris-driven melt, Geoscientific Model Development Discussions, 1–50, https://doi.org/10.5194/gmd-2021-92, 2021b.

Avanzi, F., Gabellani, S., Delogu, F., Silvestro, F., Pignone, F., Bruno, G., Pulvirenti, L., Squicciarino, G., Fiori, E., Rossi, L., Puca, S., Toniazzo, A., Giordano, P., Falzacappa, M., Ratto, S., Stevenin, H., Cardillo, A., Fioletti, M., Cazzuli, O., Cremonese, E., Morra di Cella, U., and Ferraris, L.: IT-SNOW: a snow reanalysis for Italy blending modeling, in situ data, and satellite observations (2010–2021), Earth System Science Data, 15, 639–660, https://doi.org/10.5194/essd-15-639-2023, 2023.

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.

Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W., and Levizzani, V.: Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data, Journal of Geophysical Research: Atmospheres, 119, 5128–5141, https://doi.org/10.1002/2014JD021489, 2014.

Bruno, G., Avanzi, F., Gabellani, S., Ferraris, L., Cremonese, E., Galvagno, M., and Massari, C.: Disentangling the role of subsurface storage in the propagation of drought through the hydrological cycle, Advances in Water Resources, 169, 104305, https://doi.org/10.1016/j.advwatres.2022.104305, 2022.

Chen, M., Shi, W., Xie, P., Silva, V. B. S., Kousky, V. E., Wayne Higgins, R., and Janowiak, J. E.: Assessing objective techniques for gauge-based analyses of global daily precipitation, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD009132, 2008.

Dari, J., Brocca, L., Modanesi, S., Massari, C., Tarpanelli, A., Barbetta, S., Quast, R., Vreugdenhil, M., Freeman, V., Barella-Ortiz, A., Quintana-Seguí, P., Bretreger, D., and Volden, E.: Regional data sets of high-resolution (1 and 6 km) irrigation estimates from space, Earth System Science Data, 15, 1555–1575, https://doi.org/10.5194/essd-15-1555-2023, 2023.

Gash, J.: An analytical model of rainfall interception by forests, https://doi.org/10.1002/QJ.49710544304, 1979.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., and Kessler, M.: Climatologies at high resolution for the earth's land surface areas, Sci Data, 4, 170122, https://doi.org/10.1038/sdata.2017.122, 2017.

Lagasio, M., Fagugli, G., Ferraris, L., Fiori, E., Gabellani, S., Masi, R., Mazzarella, V., Milelli, M., Parodi, A., Pignone, F., Puca, S., Pulvirenti, L., Silvestro, F., Squicciarino, G., and Parodi, A.: A Complete Meteo/Hydro/Hydraulic Chain Application to Support Early Warning and Monitoring Systems: The Apollo Medicane Use Case, Remote Sensing, 14, 6348, https://doi.org/10.3390/rs14246348, 2022.

Lievens, H., Demuzere, M., Marshall, H.-P., Reichle, R. H., Brucker, L., Brangers, I., de Rosnay, P., Dumont, M., Girotto, M., Immerzeel, W. W., Jonas, T., Kim, E. J., Koch, I., Marty, C., Saloranta, T., Schöber, J., and De Lannoy, G. J. M.: Snow depth variability in the Northern Hemisphere mountains observed from space, Nat Commun, 10, 4629, https://doi.org/10.1038/s41467-019-12566-y, 2019.

Martens, B., Miralles, D. G., Lievens, H., van der Schalie, R., de Jeu, R. A. M., Fernández-Prieto, D., Beck, H. E., Dorigo, W. A., and Verhoest, N. E. C.: GLEAM v3: satellite-based land evaporation and root-zone soil moisture, Geoscientific Model Development, 10, 1903–1925, https://doi.org/10.5194/gmd-10-1903-2017, 2017.

Massari, C., Avanzi, F., Bruno, G., Gabellani, S., Penna, D., and Camici, S.: Evaporation enhancement drives the European water-budget deficit during multi-year droughts, Hydrology and Earth System Sciences, 26, 1527–1543, https://doi.org/10.5194/hess-26-1527-2022, 2022.

Miralles, D. G., Holmes, T. R. H., De Jeu, R. a. M., Gash, J. H., Meesters, A. G. C. A., and Dolman, A. J.: Global land-surface evaporation estimated from satellite-based observations, Hydrology and Earth System Sciences, 15, 453–469, https://doi.org/10.5194/hess-15-453-2011, 2011.

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth System Science Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Musolino, D., Vezzani, C., & Massarutto, A. (2018). Drought management in the Po river basin, Italy. Drought: science and policy, 201-215.

Peter, B. G., Cohen, S., Lucey, R., Munasinghe, D., Raney, A., and Brakenridge, G. R.: Google Earth Engine Implementation of the Floodwater Depth Estimation Tool (FwDET-GEE) for Rapid and Large Scale Flood

Analysis, IEEE Geoscience and Remote Sensing Letters, 19, 1–5, https://doi.org/10.1109/LGRS.2020.3031190, 2022.

Priestley, C. H. B. and Taylor, R. J.: On the Assessment of Surface Heat Flux and Evaporation Using Large-Scale Parameters, Monthly Weather Review, 100, 81–92, https://doi.org/10.1175/1520-0493(1972)100<0081:OTAOSH>2.3.CO;2, 1972.

Pulvirenti, L., Squicciarino, G., Fiori, E., Ferraris, L., and Puca, S.: A Tool for Pre-Operational Daily Mapping of Floods and Permanent Water Using Sentinel-1 Data, Remote Sensing, 13, 1342, https://doi.org/10.3390/rs13071342, 2021.

Quast, R., Albergel, C., Calvet, J.-C., and Wagner, W.: A Generic First-Order Radiative Transfer Modelling Approach for the Inversion of Soil and Vegetation Parameters from Scatterometer Observations, Remote Sensing, 11, 285, https://doi.org/10.3390/rs11030285, 2019.

Rizzi, S. (2009). What-If Analysis. In: LIU, L., ÖZSU, M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_466

Wagner, W., Lindorfer, R., Melzer, T., Hahn, S., Bauer-Marschallinger, B., Morrison, K., Calvet, J.-C., Hobbs, S., Quast, R., Greimeister-Pfeil, I., and Vreugdenhil, M.: Widespread occurrence of anomalous C-band backscatter signals in arid environments caused by subsurface scattering, Remote Sensing of Environment, 276, 113025, https://doi.org/10.1016/j.rse.2022.113025, 2022.

Zhou, Y., Nelson, K., Mohr, K. I., Huffman, G. J., Levy, R., and Grecu, M.: A Spatial-Temporal Extreme Precipitation Database from GPM IMERG, Journal of Geophysical Research: Atmospheres, 124, 10344–10363, https://doi.org/10.1029/2019JD030449, 2019.

Technical Note 6: DTE Hydrology Community Roadmap v1.0

1. Introduction

1.1. The DTE Hydrology Evolution Project

The overarching objective of Digital Twin Earth (DTE) Hydrology Evolution project has been to prototype a full advanced end-to-end demonstrator of DTE over the full Mediterranean region at high resolution in space and time (targeting 1 km and 1 hour) by extending the successful implementation of Digital Twin Earth Hydrology project that has been carried out over the Po River Valley in northern Italy. Specifically, DTE Hydrology has demonstrated the potential to advance towards an end-to-end reconstruction of the hydrological cycle at high resolution in space and time by an effective combination of state-of-the-art EO data, in situ observations, advanced hydrological and hydraulic modelling, AI and advanced digital platform capabilities.

1.2. Scope of this Report

This document summarises the main achievements obtained during the DTE Hydrology Evolution project and outlines the future steps that should be carried out in this activity.

1.3. Applicable Documents

Not available.

1.4. Reference Documents

RD-01 DTE Hydrology Evolution Technical Proposal - V1.0
RD-02 Technical Note 1: DTE Hydrology end-to-end description and demonstration plan - V1.0
RD-03 Technical Note 2: DTE-Hydrology Datasets User Manual v1.0
RD-03 Technical Note 3: DTE Hydrology modelling experiments, validation and assessment v1.0
RD-04 Technical Note 4. DTE Hydrology community platform end to end system integration and final functional description v1.0
RD-05 Technical Note 5: DTE Hydrology Live Demonstration v1.0
RD-06 DTE Hydrology Community Roadmap – Deliverable D5.1
RD-07 DTE Hydrology Final Report – Deliverable 7.1

2. DTE scientific and technological opportunities, current gaps/limitations and roadmap

This document builds on the activities already carried out in the DTE Hydrology project (RD-07) and specifically on the DTE Hydrology Community Roadmap (RD-06). Indeed, the overall concept developed under the DTE Hydrology Evolution project is the same as done in DTE Hydrology with several improvements (see **Figure 2.1**):

1) Extension of the spatial coverage to the whole Mediterranean Basin and the temporal coverage from 2016 to September 2022 and hence the preparation to develop a Digital Twin of Hydrological processes and applications at continental and global scales (CNR-IRPI, TUWien, UGENT, CIMA).
2) Extension of the modelling activities to two separate modelling concepts:
   a. Fully distributed modelling by CIMA
   b. Modular semi-distributed modelling by CNR-IRPI
3) Development of the DTE Hydrology Platform (https://explorer.dte-hydro.adamplatform.eu/) by MEEO not only for data visualization and exploration but mainly to showcase the uses cases developed in the project:
   a. Large scale water balance assessment
   b. Flooding simulation
   c. What-if scenario for flood risk assessment
   d. What-if scenario for water resources management
4) Connection with the High Performance Computing (HPC) facility set up in ESRIN for running the fully distributed hydrological modelling by CIMA and for developing the high-resolution evaporation products by EURAC.
5) Engagement with a wider community of scientists and stakeholders with the "federation" of multiple ESA (and non-ESA) projects and with dedicated meetings as exemplified at the final meeting.
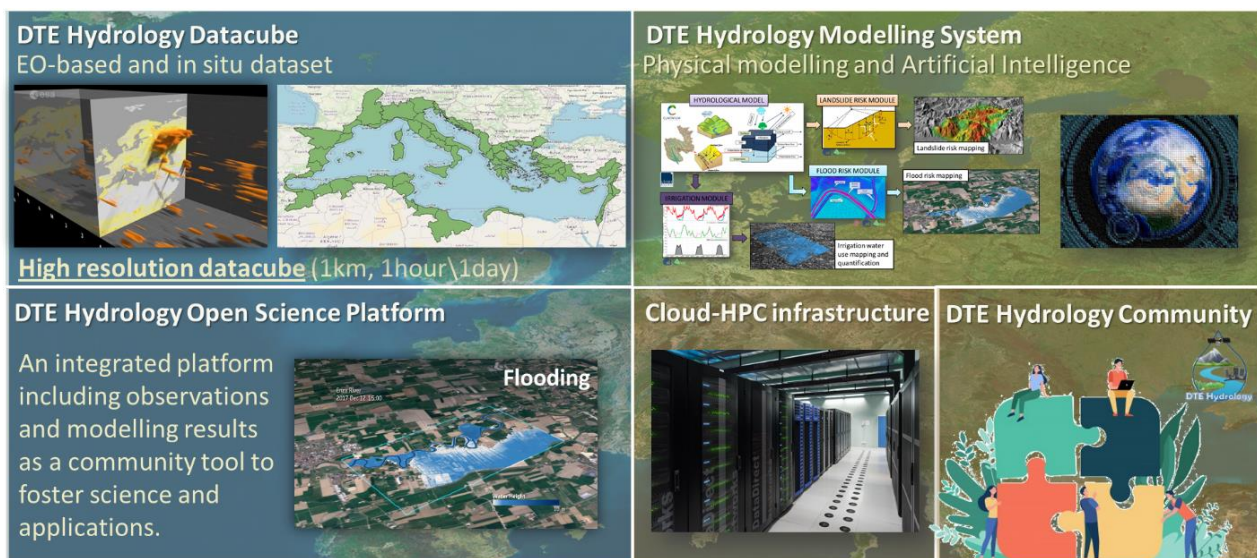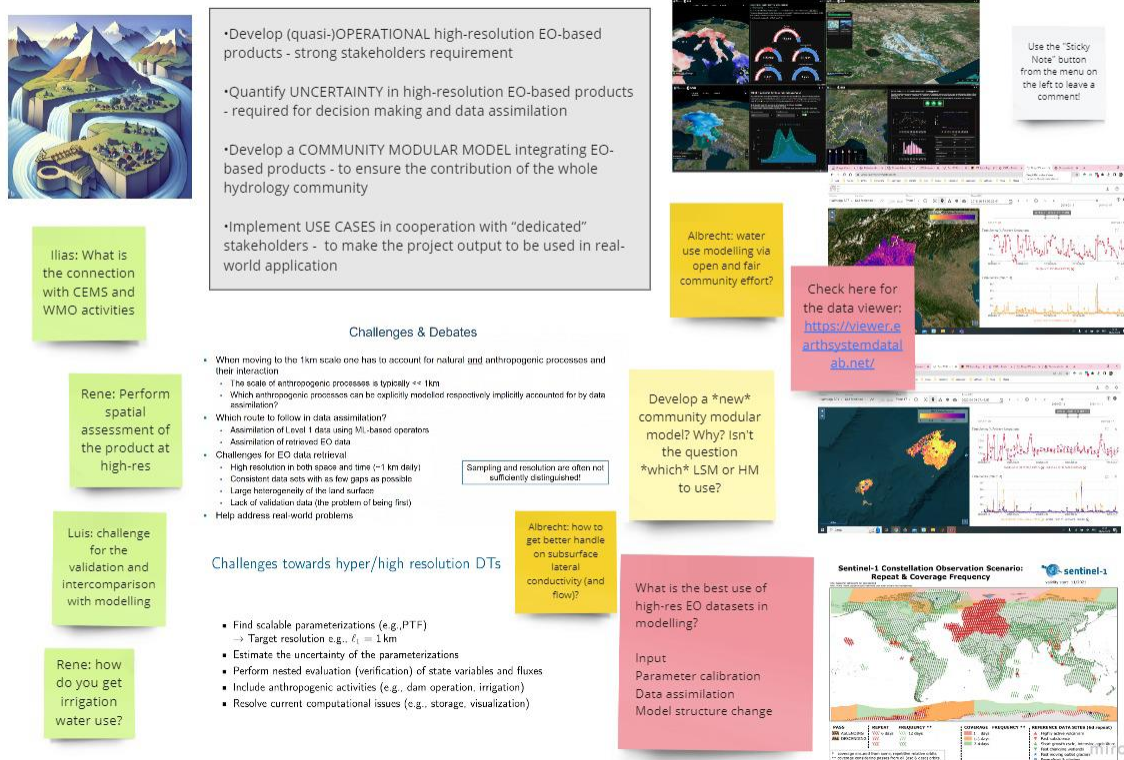
Figure 2.1: DTE Hydrology Evolution concept.

In this document we will focus only on scientific and technological gaps to be filled in future implementations of DT and we will mostly rely on the discussion carried out during the DTE Hydrology Evolution final meeting on 18th April, 2023 (see **Figure 2.2** and **Figure 2.3**). Moreover, we will also focus on the user needs based on the stakeholders tailored meeting carried out on 17th April, 2023 and on the separated meetings carried out during the project with each single stakeholders (see the list in **Figure 2.4**).

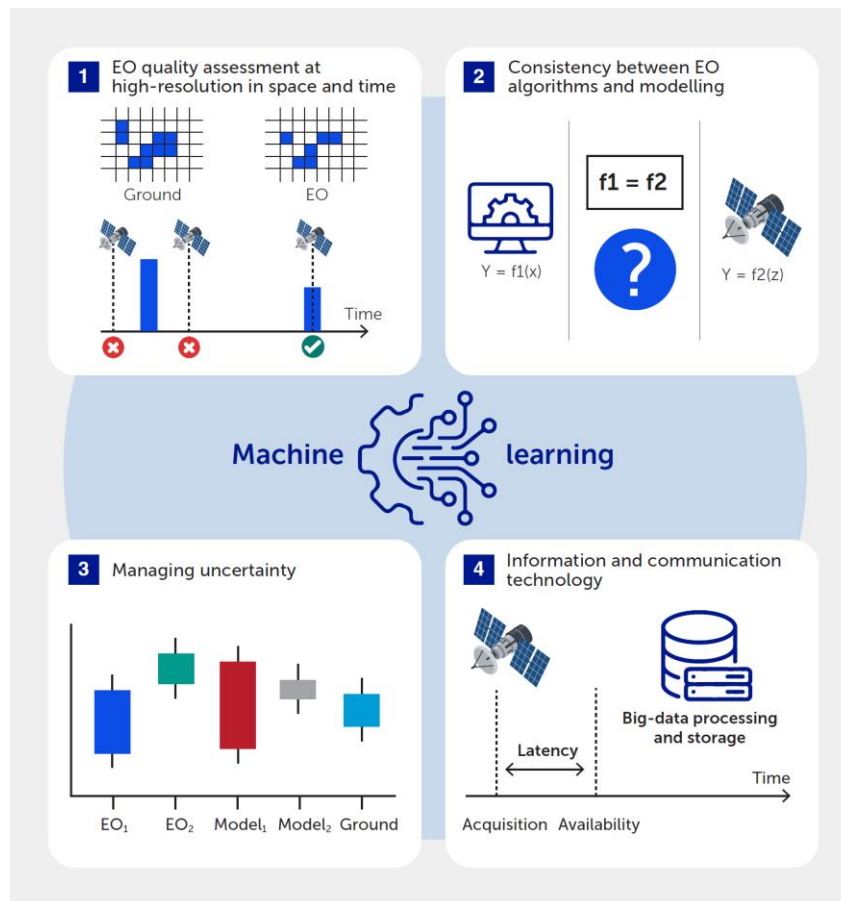*Figure 2.2*: MIRO Board built during the DTE Hydrology final meeting discussion.

Figure 2.3: Addressing challenges and limitations in earth observation (EO)-based datasets towards an improved Digital Twin Earth (DTE) Hydrology. The key challenges for DTE Hydrology are: 1) difficulties in assessing EO products' quality at high spatial and temporal resolution across different regions and climates; 2) ensuring consistency between satellite data retrieval algorithms and modeling, in both cases at high resolution 3) managing the inherent uncertainties in the accuracy of data from EO and model products, and; 4) availability of information and communication infrastructures that could support higher computational capacity while optimizing data latency and allowing long time-series. Artificial intelligence and machine learning techniques will play a vital role in addressing these challenges.
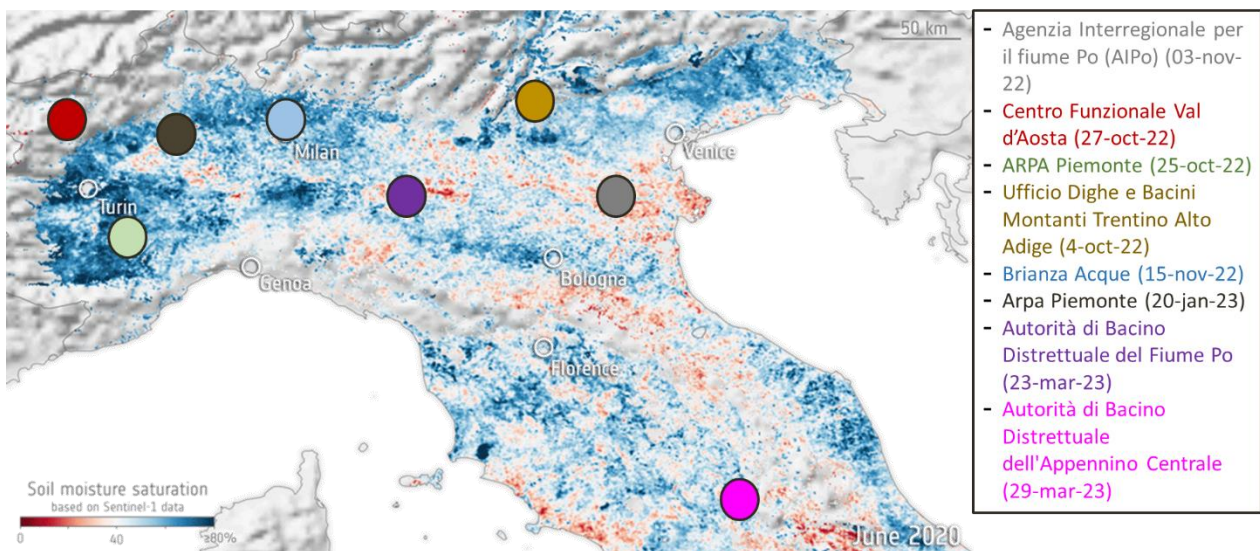
Figure 2.4: Location and name of the stakeholders involved during the DTE Hydrology Evolution project.

2.1. Development of high-resolution EO-based products in hydrology

For the development of high-resolution EO-based products in hydrology the following points are relevant:

- Assessment of uncertainty and error; the validation of high-resolution products is highly challenging and the development of in situ networks should be supported, particularly for evaporation and soil moisture, which are currently monitored at very few locations. The assessment of EO-based products performance in space should be performed to clearly understand their level of accuracy in reproducing high-resolution spatial patterns. All products need to be further tested in different regions and climates to comprehensively assess their quality, reliability, and usability for high-resolution applications.

- The consistency and independence of the different EO-based products describing the water cycle should be clearly assessed. Indeed, some of the products (e.g., evaporation, groundwater, runoff) have many dependencies on other products (**Figure 2.5**). The satellite datasets, and ancillary data, used for developing the products should be clearly defined to make the user aware of how best to use the products. This target is potentially achievable and it requires a strong collaboration between the modelling and the remote sensing communities.
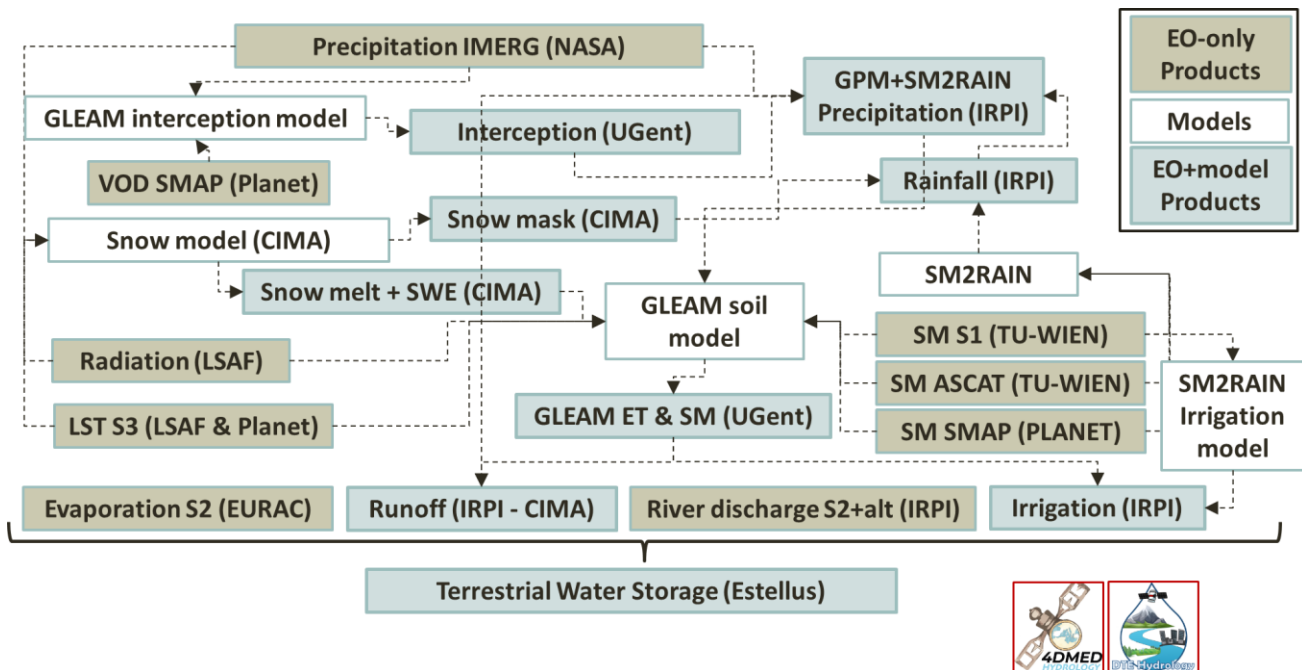
Figure 2.5: Diagram showing the dependencies of different EO-based products developed under DTE Hydrology and 4DMED Hydrology ESA projects.

- The availability of high-resolution EO-based products in a pre-operational context should be investigated. The possibility to develop operational chains for each of the products should be defined.

- Additional products that have not been fully exploited in DTE Hydrology projects and for which ESA is supporting the development are: river discharge, reservoir water level, land surface temperature and groundwater storage. The development of these EO-based products, the extension of current products over larger areas and temporal coverages, and the possibility to improve spatial and temporal resolution (**Figure 2.6**) needs to be investigated, too.
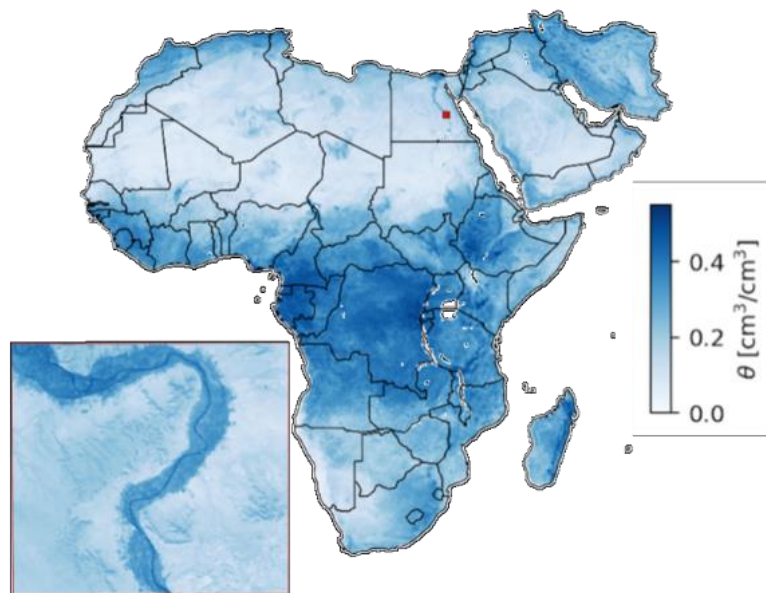
Figure 2.6: High-resolution (100 m) soil moisture product from WAPOR v3.0 for Africa.

2.2. Integration of high-resolution EO-based products and modelling

For the integration of high-resolution EO-based products in modelling the following points are relevant:

- There's the need to understand what is the best approach to exploit EO-based products in high-resolution modelling. Five different options are available for the use of EO-based products (some of these options will be investigated in 4DHYDROLOGY project but in-depth investigations are needed):

  - As input data, e.g., precipitation is found to be the most important variable if flood prediction is targeted
  - For the spatially distributed calibration of model parameter values including the calibration over the different components of the water cycle. The parameterization of high-resolution modelling is highly challenging and in this context the integration with EO-based products can be very effective.
  - In the framework of data assimilation.
  - For improving modelling structure, e.g., from the analysis of the model results in comparison with multiple EO-based products (**Figure 2.7**)
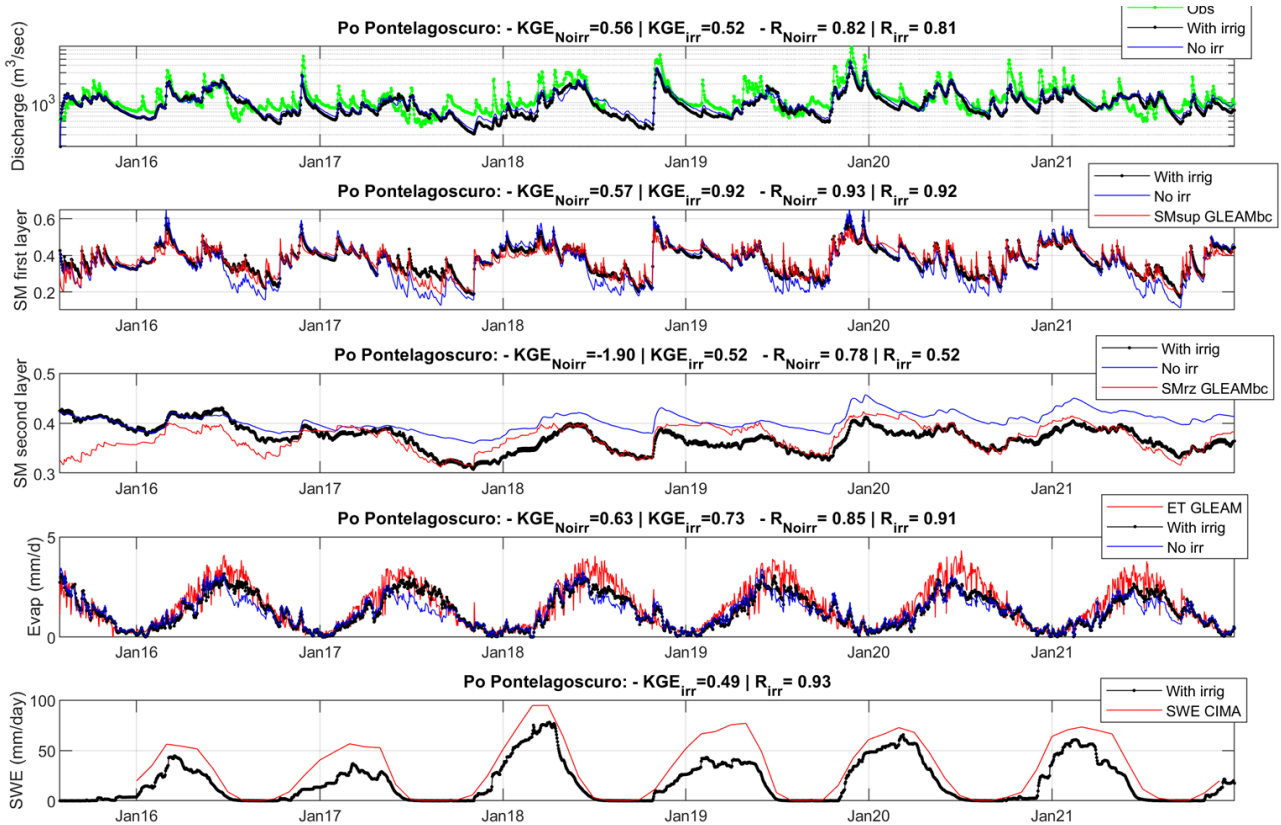 - For including unmodelled processes (e.g., irrigation).

Figure 2.7: Observed and modelled river discharge, modelled and GLEAM surface (SM first layer) and root-zone (SM second layer) soil moisture, modelled and GLEAM evapotranspiration (Evap) and modelled and CIMA snow water equivalent (SWE) at Pontelagoscuro section.

- The use of novel EO-based products (e.g., irrigation water use, reservoir water level) should be investigated to include anthropogenic activities in the modelling. Very first attempts have been made in ESA's DTE Hydrology and Irrigation+ project (**Figure 2.8** and **Figure 2.9**).
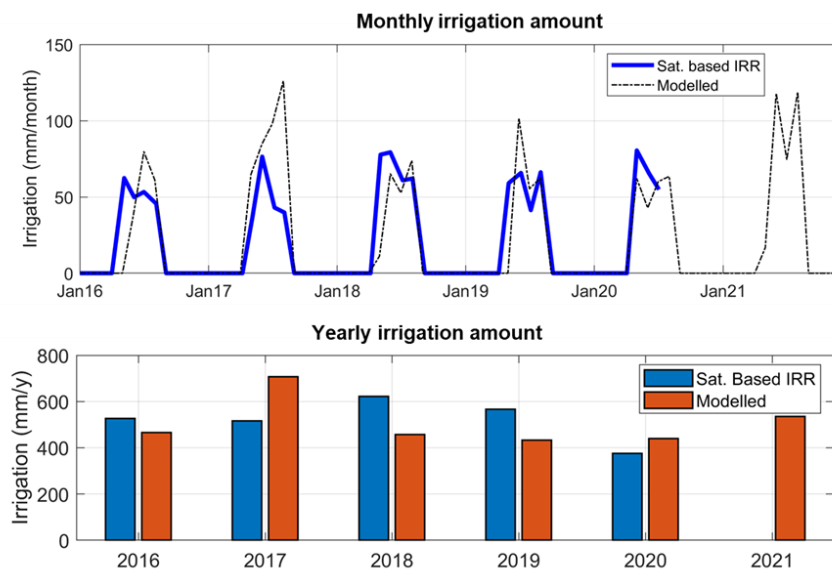
Figure 2.8: Satellite-based and modelled irrigation estimates at monthly and yearly scale for the Po river basin, comparison carried out under ESA DTE Hydrology Evolution project.
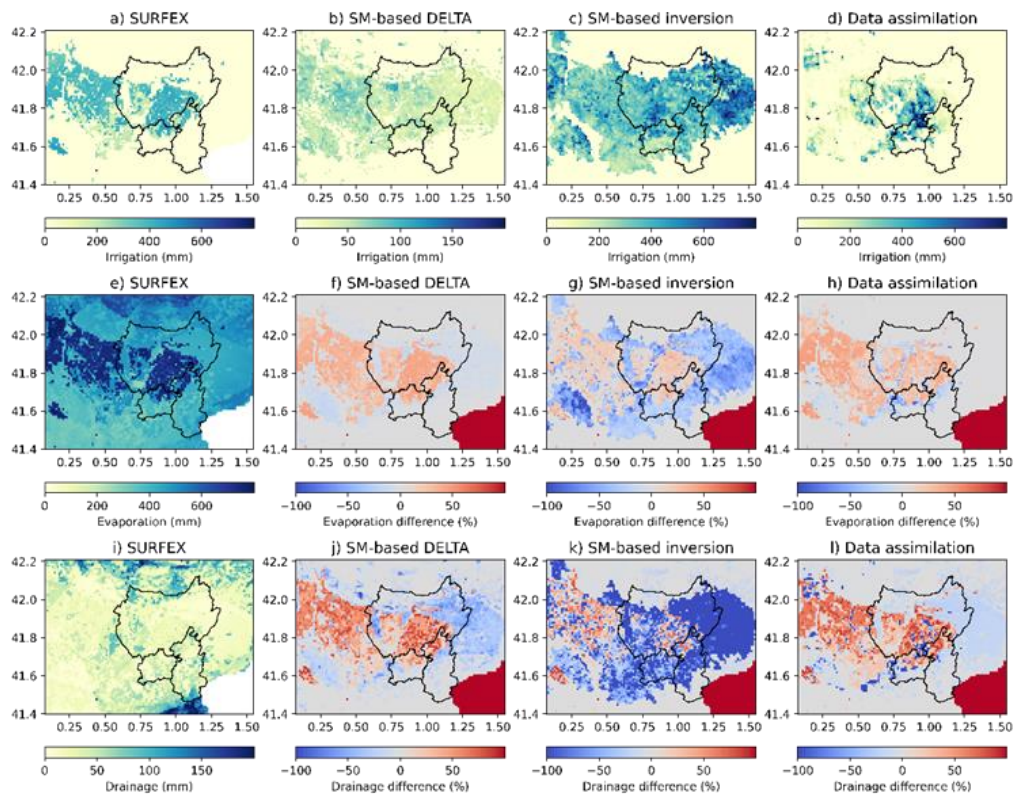
Figure 2.9: First row: SURFEX modelled mean annual irrigation (panel a) and mean annual irrigation provided by the IRRIGATION+ products: SM-based DELTA (panel b), SM- based inversion (panel c), and data assimilation (panel d). It should be noted that panel b's scale differs from the rest of the row's panels. Second and third rows: SURFEX modelled mean annual evaporation (panel e) and drainage (panel i). The rest of the panels from the second and third row show the difference as a percentage of modelled mean annual evaporation and drainage, respectively.

- High-resolution models should be calibrated and tested with spatially distributed data for fluxes (evaporation, river discharge) and states (soil moisture, snow water equivalent, groundwater storage), as shown in **Figure 2.7**. Sub-surface water dynamics (deep soil moisture, groundwater, infiltration, percolation) were a bit under-represented. The streamflow-only calibration should be avoided.

2.3. Models, data assimilation and machine learning needs and required developments

For the modelling of water cycle at high resolution the following issues are identified:
- The representation of physical processes over land (e.g., infiltration, surface and subsurface runoff, soil-plant-atmosphere interaction, groundwater dynamics) at high-resolution is significantly different from processes at coarse spatial resolution (20 km), as currently modelled by continental and global scale land surface and hydrological models. The modelling community must now collaborate to develop a modelling system that is reliable across multiple spatial scales.

- The spatial scale of anthropogenic processes is typically <<1 km, and modelling such processes is highly challenging (if feasible). The exploitation of data assimilation in this context is advisable but it's still not clear if data assimilation is able to fully account for unmodelled anthropogenic processes. EO-based products will be crucial in providing information on the human impact on the water, energy and carbon cycle, and hence such data need to be optimally and efficiently integrated into the modelling systems.

- The validity of 1 km/1 hour simulations should be investigated, especially as not all input data are available at this resolution, as the result is likely to be highly dependent on model assumptions.

- Uncertainties inherently affect both EO and model products' accuracy. Data assimilation is one of the most viable ways to optimally integrate EO and models and constrain uncertainties. The treatment of bias is an important aspect in this context, considering that most data assimilation methods assume that sources of information are unbiased.

- The assimilation of both level-1 EO-data (e.g., backscatter) and retrievals (e.g., soil moisture) should be investigated to have a better understanding on the impact of this choice on the results.

- The role of machine learning in the modelling development is to be investigated too. The development of hybrid modelling approaches, i.e., integrating physically based and machine learning modules, have been foreseen.

- Recent years have seen a continuous and rapid development of new artificial intelligence (AI) and machine learning techniques. These are expected to improve the outcomes of the DT for the water cycle, and need to be tested to:
    - Improve parameter estimation.
    - Elucidate poorly understood processes.
    - Emulate costly models.
    - Answer questions about the systems – questions not related to the 'what?', but to the 'why?' and the 'what if?' through explainable AI.

2.4. User needs

The interaction with stakeholders in the use of the novel and advanced EO-based products and modelling have been discussed at the final meeting and into the dedicated meetings. There's the need to co-develop use cases with interested stakeholders - "early adopters" - to build a system suitable for their purposes but at the same time exploiting the most recent advances in EO and modelling.

The development of the platforms to easy access to the data (e.g., https://viewer.earthsystemdatalab.net/, **Figure 2.10**, and https://explorer.dte-hydro.adamplatform.eu/, **Figure 2.11**), and easy data processing is highly encouraged and appreciated as the data volume has become unmanageable for most of the users.
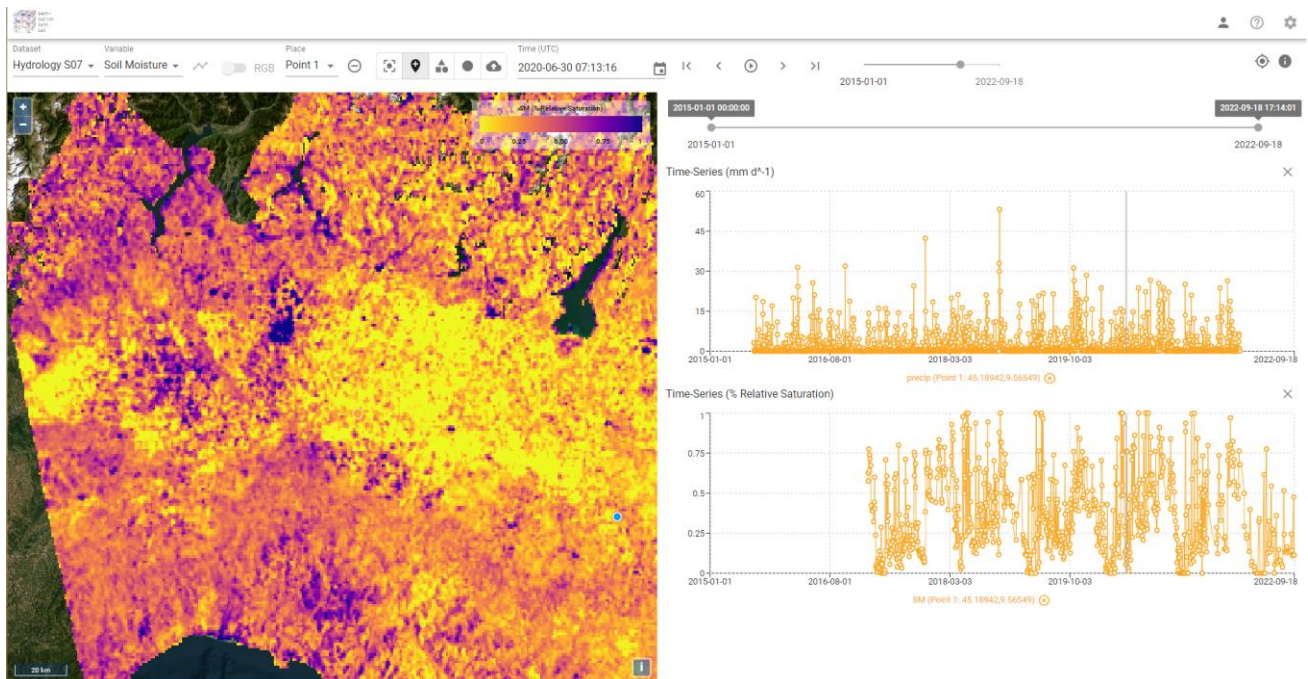

Figure 2.10: Example of visualization of spatial and temporal data for the Earth System Data Lab viewer.
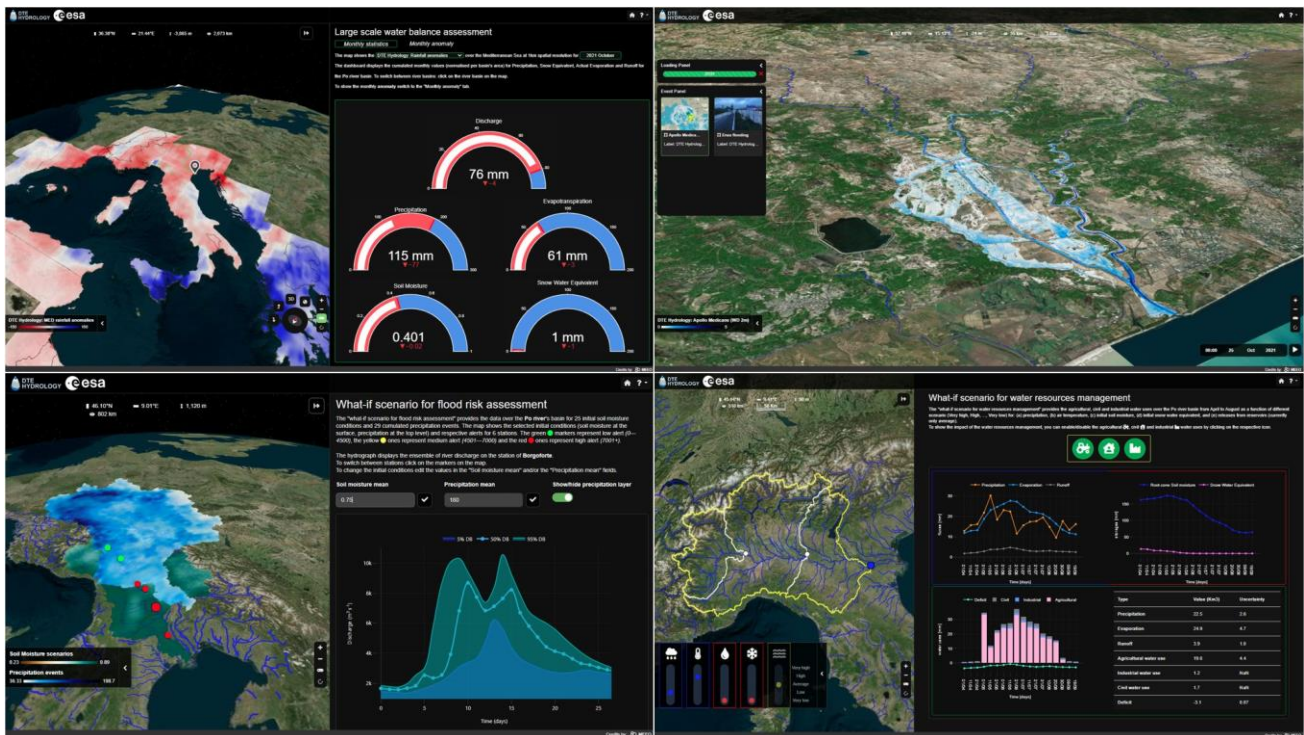
Figure 2.11: Example of visualisation of ADAM platform for drought monitoring, flooding monitoring, and what-if scenario for flood risk assessment and water resources management.

2.5. Technological needs and required developments

For the technological needs in the development of high resolution DTE for the water cycle, the following issues are identified:

- Data latency (i.e. the total time elapsed between data acquisition by a sensor and its availability) is an important limitation of EO-based datasets for some applications. An efficient, global, near-real time processing environment that could continuously provide higher-level data products on top of the acquired satellite data might require a specifically designed IT infrastructure optimized for the data acquisition, analysis ready data, and product processing and data distribution. The integration of the platform with High Performance Computing (HPC) infrastructures and centralized cloud systems is expected to address this issue.

- In recent years new satellite missions for the water cycle have been launched, such as the Meteosat Third Generation (MTG), the Surface Water and Ocean Topography satellite (SWOT) mission. Furthermore, there is a strong portfolio of guaranteed missions, such as the future Sentinels and the NASA-ISRO SAR (NISAR). Their joint exploitation for producing high-resolution products has to be investigated.

- There's the need to develop a platform\system facilitating the interaction of the modelling community in developing modular models, easily integrated with satellite observations (**Figure 2.12**). That system will facilitate the scientific community to test multiple modules integrated in the common framework. First attempts are available in the scientific literature and need to be investigated further (see EGU 2023 session: International community of practice on hydrological modelling - convenor Martyn Clark, https://meetingorganizer.copernicus.org/EGU23/EGU23-10527.html).
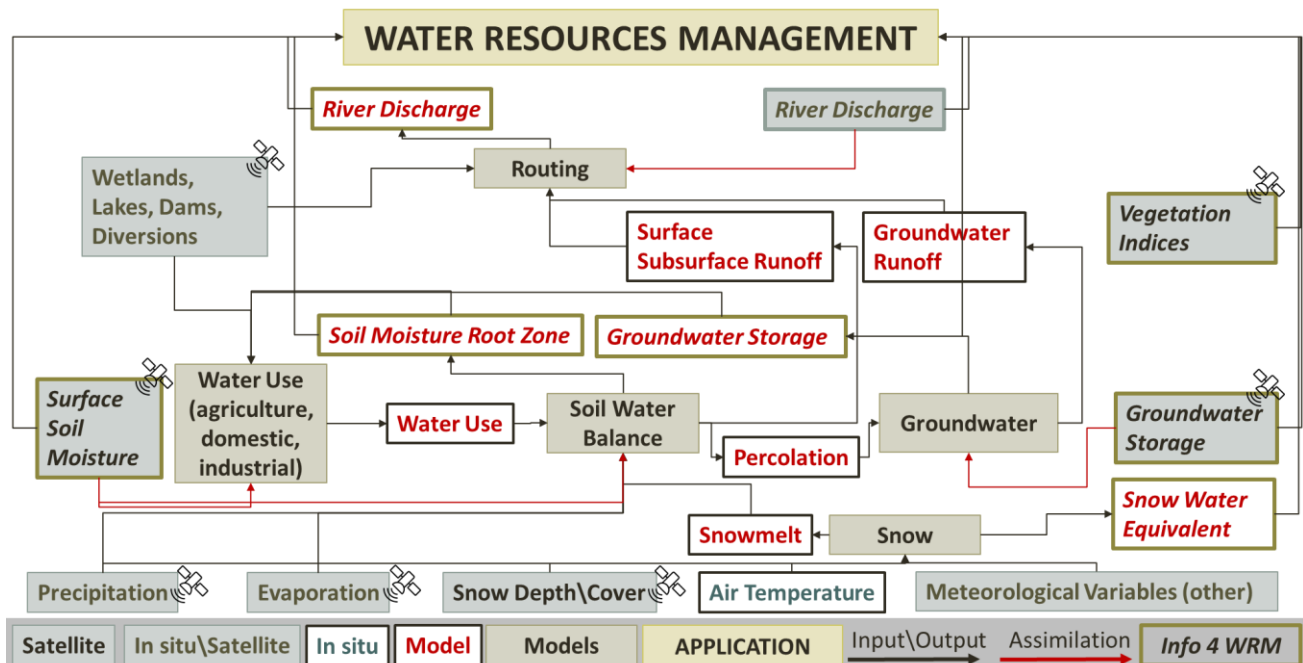

Figure 2.12: Diagram showing the different datasets and modules to be developed for addressing water resources management.

- There's the need to move forward in terms of data exchange between remote sensing and modelling communities with consistent data formatting meeting CF conventions. Data specifications should include time reference, updating frequency, latency, temporal validity, spatial and temporal resolution, and uncertainty/confidence. Data visualization, extraction, download should be easy and straightforward.