The GIRI global flood hazard model

Lorenzo Alfieri, Lorenzo Campo, Simone Gabellani, Tatiana Ghizzoni, Christian Herold, Andrea Libertino, Eva Trasforini, and Roberto Rudari

[1]CIMA Research Foundation, Savona, Italy
[2]UNEP/GRID, Geneva, Switzerland

Correspondence to: Lorenzo Alfieri (lorenzo.alfieri@cimafoundation.org)

Abstract. In the recent years, global flood models have emerged as practical tools to transform our understanding of global flood risk. However, the large computational efforts needed to produce them limit the existing applications to a few scenarios, partial coverage, or coarse resolution products. In this article, we present a methodological approach for producing 90m resolution global flood hazard maps for different flood magnitudes under present and future scenarios. The approach relies on a cascade of calibrated meteorological-hydrological-hydraulic models and integrates global datasets of atmospheric variables for the present climate and from bias corrected projections of future climate from the ISIMIP3b initiative, enabling the creation of comprehensive and detailed flood hazard maps for different return periods. The significance of such mapping lies in its ability to address the challenges posed by local and global-scale flood events, as well as the impact of climate change on flood risk management. Results contribute to the Global Infrastructure Risk Model and Resilience Index with an advanced hazard product with key implications for improved financial loss assessment, aid in disaster risk reduction efforts, and for global impact assessments.

1 Introduction

Flooding has traditionally been tackled at the local scale. Yet, there is a growing understanding that several flood events are connected to, or driven by, short and long-term global weather patterns (Hagos et al., 2016; Fan et al., 2015). Globalization has made flood events in one part of the world capable of causing significant economic and social impacts in all parts of the globe (Trigg et al., 2013). Addressing the effects of climate change, which are felt globally, presents an additional challenge to flood risk management efforts. Hence, there is an increasing need for assessments of flood risk at the global scale. Such a global need has become evident on several fronts: scientific studies to simulate the effects of general circulation modeling, insurance catastrophe modeling to understand risk and exposure (Bates et al., 2018) and intergovernmental efforts in disaster risk reduction (e.g., Desai et al., 2015; UNISDR, 2015).

Over the past decade, the development of Global Flood Model (GFM) initiatives has rapidly evolved from research experiments into practical tools that are transforming our comprehension of global flood risk (Ward et al., 2015). These initiatives have benefited from advancements in computational power, global data availability, and remotely sensed products, leading to accurate and high resolution globally available datasets. GFMs are often based on a cascade of meteorological-

hydrological-hydraulic models. They are particularly suitable for estimating potential inundation under different flood probabilities, hence, they can project potential future flood hazard under different climatic conditions. Relevant examples of GFMs include CaMa-UT from the University of Tokyo (Yamazaki et al., 2011), CIMA-UNEP developed for the UNISDR Global Assessment Report (GAR) 2015 (Rudari et al., 2015), the ECMWF model (Pappenberger et al., 2012), the FATHOM
35 Global flood hazard model (Sampson et al., 2015), GLOFRIS by Deltares (Winsemius et al., 2013), and the European Commission - Joint Research Centre (JRC) model (Dottori et al., 2016).

The objective with flood hazard mapping in the context of financial loss assessment is to have comprehensive and detailed flood hazard maps for different return periods, typically a relatively small set of maps with constant probability of occurrence, often expressed in terms of return periods (usually between 10 and 1,000 years). The challenges when mapping flood hazard
40 are the resolution required and the spatial coverage. Local topography conditions can significantly influence the damages sustained in the properties. For instance, Wojtkiewicz et al. (2013) estimated that around 30% of the National Flood Insurance Program (NFIP) claims in the United States are located outside of the 100-year flood zones.

In this article, we present the methodological approach devised to produce global flood hazard maps under present and future scenarios, based on hydrological and hydraulic modeling and global datasets of atmospheric valiables as forcing input. The
45 procedure presented here builds upon and improve the work by Rudari et al. (2015) for the Global Assessment Report on Disaster Risk Reduction and contribute to the Global Infrastructure Risk Model and Resilience Index (GIRI), an ongoing initiative of the Coalition for Disaster Resilient Infrastructure (CDRI).

2 Material and methods

2.1 The flood modelling framework

50 Several literature approaches use a set of stochastic rainfall scenarios and a simplified basin response model to generate flood intensity scenarios. This approach has some advantages, including the ability to produce a large sample of possible rainfall histories. However, the simplification brought by considering only simple rainfall-runoff transformation is often unacceptable when working over a large domain that includes different climatological and hydrological regimes. Therefore, it is more robust to use a more complete hydrologic model that allows the simulation of all components of the water cycle in a continuous and
55 physically based manner. However, such models cannot be driven by rainfall fields alone, hence different meteo-hydrological variables fields need to be dynamically consistent with one another. This is often tackled by using the output of General Circulation Models (GCM), which offer all the above-mentioned features at the appropriate spatial and temporal scales for large-scale studies. Such an approach has several advantages:

• modelling hydrological events in a continuous manner, therefore taking into account the initial hydrological conditions
60 before heavy rainfall events;

• modeling complex hydrological interactions among the different flood formation processes (e.g., groundwater contributions, snow melt contributions, exfiltration in the peri river system).

• Including the effects of climate change by using state of the art climate projections available from GCM model output forced by projected greenhouse gas concentrations in the atmosphere.

65 The workflow chosen for this work is to exploit the output of numerical climate models and post-processes the output of the hydrologic and hydraulic simulations to expand the number of scenarios to be used through a probabilistic approach.

2.2 Meteorological data

The meteorological dataset used in the present study for simulating basins' response during the historical period is composed of the W5E5 climate data (Karger et al., 2022), which include precipitation, air temperature, air humidity, wind velocity, and
70 solar radiation. This dataset is a merged product that incorporates WFDE5 data (Cucchi et al., 2020) over land and ERA5 data (Hersbach et al., 2020) over the ocean. The WFDE5 dataset was generated by implementing the WATCH Forcing Data (WFD) methodology to surface meteorological variables obtained from the ERA5 reanalysis. Bias-adjusted monthly precipitation totals of WFDE5 results in more plausible global hydrological water balance components as analyzed in an uncalibrated hydrological model (WaterGAP), compared to using raw ERA5 data for model forcing (see Cucchi et al., 2020). In addition,
75 the W5E5 dataset was compiled to support the bias adjustment of climate input data for the impact assessments carried out in phase 3b of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP3b, Lange and Büchner, 2021), which guarantee consistency between the present and the future climate forcing used in this study.

Future climate projections to feed the hydrological model were taken from the ISIMIP3b dataset. The latter is composed by 30 model runs, resulting from the combination of 10 models with 3 Shared Socioeconomic Pathways (SSP) scenarios: SSP126,
80 SSP370, SSP585. The denomination of SSP scenarios comprises the name of the basic pathway (i.e., 1 to 5) defined by O'Neill et al. (2014), followed by two numerals indicating the additional radiative forcing achieved by the year 2100 in units of tenths of watts per square meter.

To reduce the computational load, yet providing a comprehensive evaluation of the possible future scenarios, hydrological simulations were performed with two representative climatic scenarios which include the majority of the possible future
85 combinations. The selection was made by taking the model runs closest to the 20th and the 80th percentiles of projected global surface air temperature, selected between the 15 scenarios (see Figure 1) derived from the 5 ISIMIP3b models considered as "primary"[1] following analysis on their process representation, structural independence, climate sensitivity, performance in the historical period as well as the special input data needs of the fisheries and marine ecosystems sector (FishMIP). The two resulting scenarios chosen to model the future climate are the SSP126/IPSL-CM6A-LR and SSP585/IPSL-CM6A-LR, both
90 resulting from the CM6A-LR model of the Institut Pierre-Simon Laplace (IPSL).

The Clausius-Clapeyron equation indicates that air temperature is directly linked to atmospheric water vapour content, which in turn determines the total precipitable water. Hence the choice of climate projections based on average air temperature change is relevant for hazards like flooding, tropical cyclones, rainfall-induced landslides, as well as for droughts, which are related

---

[1]See https://www.isimip.org/documents/413/ISIMIP3b_bias_adjustment_fact_sheet_Gnsz7CO.pdf

to both precipitation and temperature. The selected scenarios of ISIMIP3b include a complete bias adjustment and statistical
95 downscaling for all the atmospheric variables based on ISIMIP3BASD v2.5 (Lange, 2021), which in turn relies on the W5E5
dataset. This approach enables achieving both physical coherence between the variables and the correction of their bias based
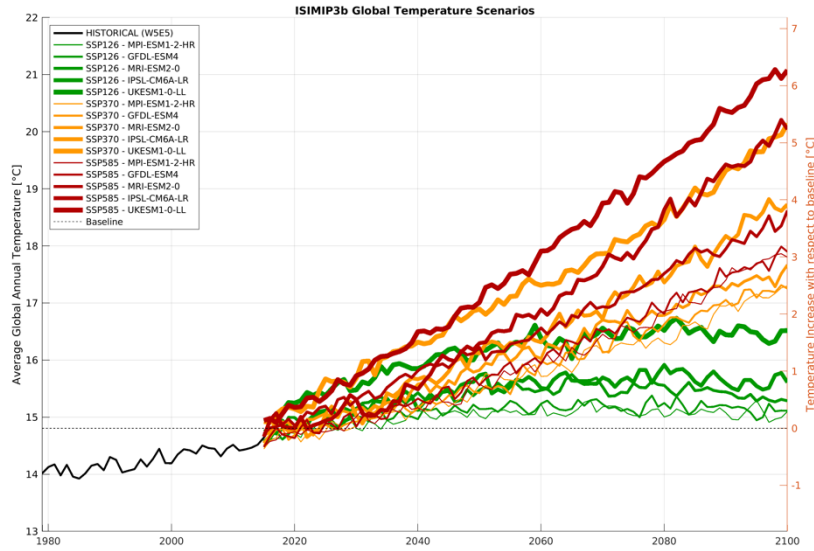on global observational datasets.

Figure 1: Trends of global air temperature in the 15 model-SSP combinations of ISIMIP3b primary models.

100

2.3 Hydrological Modelling

Hydrological processes in the study region are simulated with the Continuum model (Silvestro et al., 2013, 2015). Continuum is a semi-physically based rainfall-runoff-routing distributed hydrological model, which completely solves the mass and energy balance at the land surface. It relies on a morphological approach placing the Digital Elevation Model (DEM) as the key

105   element, from which the drainage network and other hydrological derivatives are computed (Giannoni et al., 2000). Continuum reproduces the spatio-temporal evolution of runoff, soil moisture, energy fluxes, surface soil temperature, snow accumulation and melting, by reproducing all the main processes of the hydrological cycle. The model can implement the presence of dams, lakes, diversions and other hydraulic structures. The infographic below summarizes inputs and outputs of the model.
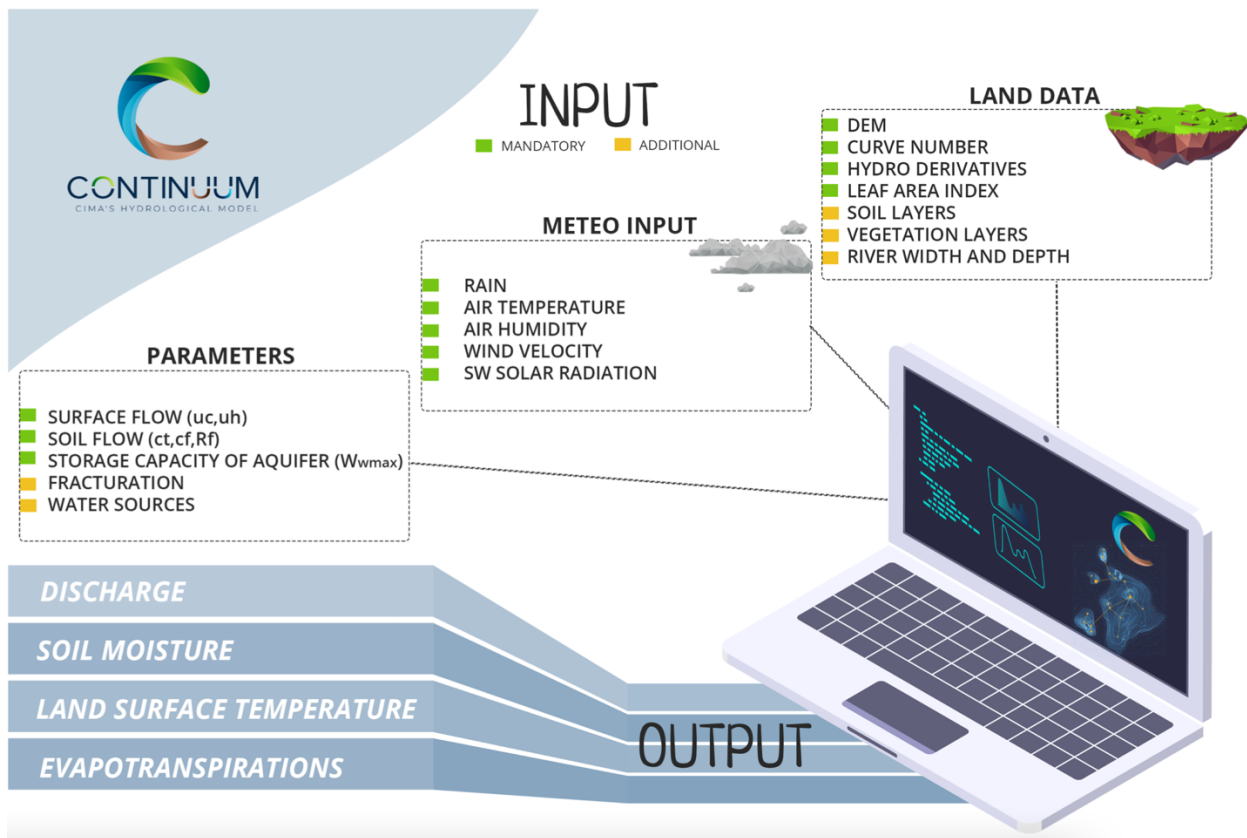

Figure 2: Overview of the input and output data of Continuum.

110

Continuum was implemented in operational flood forecasting systems and in several research studies (e.g., Laiolo et al., 2013; Cenci et al., 2016; Corral et al., 2019; Alfieri et al., 2023) as well as at the global scale for the GAR2015 global flood model (Rudari et al., 2015). The model is open source and its code is available on the GitHub page https://github.com/c-hydro/hmc-

dev together with several pre- and post-processing tools. An overview documentation of the Continuum model and its ancillary tools can be found at https://hmc.readthedocs.io/en/latest/.

In this work, Continuum was set up into 40 computation domains, covering all continents except Antarctica. Model domains were implemented at two spatial resolutions (Figure 3): a) 2 arc-minutes (~3600 m at the equator), including Africa, South America, Central America (Mexico, Cuba), and Oceania (Fiji-Vanatu-New Caledonia-Solomon, New Zeland); and b) 4 arc-minutes (~7200 m at the equator), including North America, Europe, Asia, and Oceania (Australia).

The Digital Elevation Model (DEM) is taken from the MERIT Hydro dataset (Yamazaki et al., 2019), with spatial resolution of 3 arc-second (~90 m at the equator). MERIT Hydro comes with a pre-computed and corrected set of hydrological derivatives, including channel network and basin partitioning. Ancillary data including flow accumulation and drainage direction were computed from the DEM with GRASS GIS (https://grass.osgeo.org/). The DEM was upscaled at the chosen domain resolution and carved using the 90 m stream network. Land use and land cover information at 300 m resolution are taken from the ESA-CCI Land Cover map v2 (ESA, 2017), which was used to estimate the soil characteristics and the vegetation cover. Further, we applied the USDA method for soil texture identification and hydrologic soil type classification (Shirazi and Boersma, 1984) by combining the ISRIC SoilGrids (Hengl et al., 2017) maps of soil fraction in sand and clay at 250 m spatial resolution. River widths and depths are estimated by identifying power-law functions depending on the drainage area, consistently with the data derived by Andreadis et al. (2013).

Point features implemented include the largest 2074 reservoirs and 4125 lakes (Figure 4), extracted from the Global Dam Watch (Mulligan et al., 2021), the Global Reservoir and Dam Database (GRanD v1.3) (Lehner et al., 2011), the FAO-AQUASTAT-Dams (https://www.fao.org/aquastat/en/databases/dams) and the HydroLAKES (Messager et al., 2016) datasets. Both sets of lakes and reservoirs were selected among those having total storage larger than 300 Mm^3, with additional screening to filter out lakes at the outlet of endorheic basins, coastal lakes, as well as those with insufficient data quality.

Figure 3: Continuum model domains at 2 (green) and 4 (red) arc-minute resolution.

Figure 4: Dams and lakes included in the Continuum model implementation.

2.4 Model calibration

To improve the representation of the hydrological states, Continuum was calibrated in the model domains using discharge, soil moisture and evapotranspiration data as benchmark. The aim is to accurately representing the statistical characteristics of discharges across the entire river network, with particular focus on high and low extremes. Daily discharge data was taken from 3068 gauging stations chosen among the more than 10,000 stations of the Global Runoff Data Centre[2] (GRDC) following

---

[2]https://www.bafg.de/GRDC/EN/Home/homepage_node.html

7

quality control criteria. Daily soil moisture and evapotranspiration maps were taken from the GLEAM (Martens et al., 2017) dataset. We deployed a multi-site calibration procedure that iteratively searches the model parameterization that best matches the available observations over the calibration period, through minimization of a cost function. Hydrological simulations run for the model calibration range between 2 and 3 years, chosen according to criteria of data quality, number of available discharge stations and preference to more recent periods. Each calibration is preceded by a run of about 5 years to enable model warm-up and improvement of the initial model states. The calibration tool aims to optimize four physical hydrological features: infiltration velocity at saturation (cf), field capacity (ct), Curve Number (CN), and water sources (ws). This is done by perturbing four dimensionless scalar parameters: k_{ct}, k_{cf} and k_{CN}, used to rescale the default maps of ct, cf, and CN respectively, and a multiplier k_{ws} of the ws map. In particular, each map of ct, cf and CN results from the rescaling of an initial map using an arc-tangent function, which maintains the original pattern and ensures the physical range of the resulting values. The cost function J is composed of three terms, related to the modeling error on river discharges, soil moisture and evapotranspiration, respectively. The first term is based on the maximization of the Kling-Gupta Efficiency (Gupta et al., 2009), by computing an error between the flow duration curves at each percentile between 1 and 100, weighting with the logarithm of the station upstream area, to give higher weight to the downstream stations without neglecting the contribution of the most upstream ones. The error terms on soil moisture and evapotranspiration are based on a minimization of their root mean square error (RMSE) on a seasonal basis. The cost function J to be minimized is defined as:

$$J = \frac{\sum_{i=1}^{N_{flow}} \log(A_i) f\left(1 - KGE\left(Q_{mod,i}, Q_{obs,i}\right)\right)}{\sum_{i=1}^{N_{flow}} \log(A_i)}$$

\+ (1a)

$$+ \; w_{SM} \sum_{k}^{Nseason} RMSE\left(SM_{mod,k}, SM_{sim,k}\right) + w_{ET} \sum_{k}^{Nseason} RMSE\left(ET_{mod,k}, ET_{sim,k}\right)$$

(1b)

where:

• $Q_{mod,i}$, $Q_{obs,i}$ are modelled and observed quantiles of the duration curves of discharge flows in the hydrometric station i

• A_i is the upstream area of the hydrometric station i

• KGE is the Kling-Gupta Efficiency

• $SM_{mod,k}$, $Sm_{sim,k}$ are modelled and simulated soil moisture in the season k {DJF=winter, MAM=spring, JJA=summer, SON=autumn}

• $ET_{mod,k}$, $ET_{sim,k}$ are modelled and simulated evapotranspiration in the season k

170   • w_{SM}, w_{ET} are weights to assign to the calibration component for soil moisture and evapotranspiration (tipically we used $w_{SM}=w_{ET}=0.5$).

The calibration algorithm performs an iterative exploration of the 6-dimensional parameters space, with N sets per iteration sampled through a gaussian Latin Hypercube methodology (N~30-35). For each of these sets of parameters, hydrological

175   simulations are performed, and the cost function J is computed to map the error hypersurface. The point that minimizes J is used as the new centre for the next exploration (Figure 5), whose range is progressively reduced until a convergence is reached, i.e., the reduction of J between two consecutive iterations is under a predefined range.
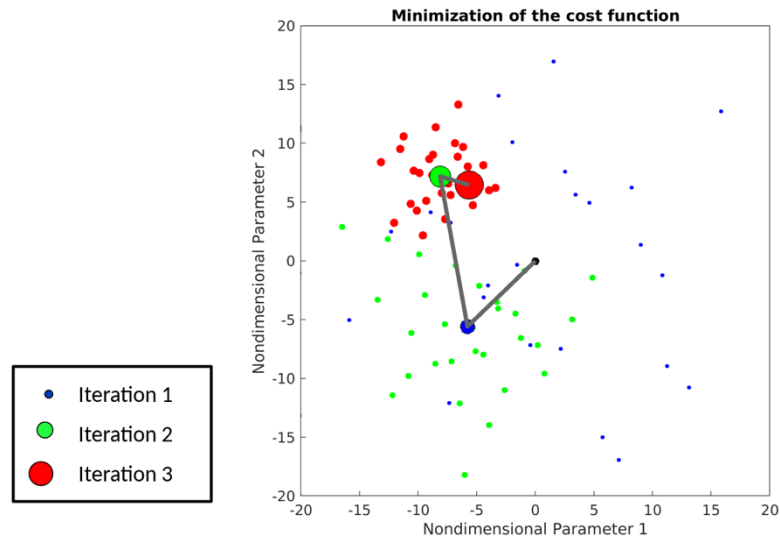
Figure 5: Representation of the search algorithm in a 2-dimensional space

180

2.5 Discharge Quantiles

Calibrated parameters were then used in a set of long-term hydrological simulations for all domains and over the entire period of availability of the meteorological forcing in the historical period 1979-2016, as well as for a 50-year time slice in the future climate chosen as 2051-2100. Long term simulations are analyzed statistically to extract discharge annual maxima and estimate
185 extreme value distributions for about 140,000 representative river sections along the simulated world river network for both the historical and future climate. We tested various analytical probability distribution functions on each set of discharge peaks, including the Generalized Extreme Value (GEV), Log-normal, Gamma, Weibull, Gumbel, Normal, Exponential, Generalized Pareto, and Log-Pearson (see example in Figure 6). For each fitted probability distribution we calculated 40 quantiles, with finer refinement around both tails of the distribution, and compared them versus the empirical ones. The probability distribution
190 with the minimum root mean square error between the empirical and the fitted quantiles was then selected to estimate the discharge quantiles to use in the subsequent hydraulic modeling.
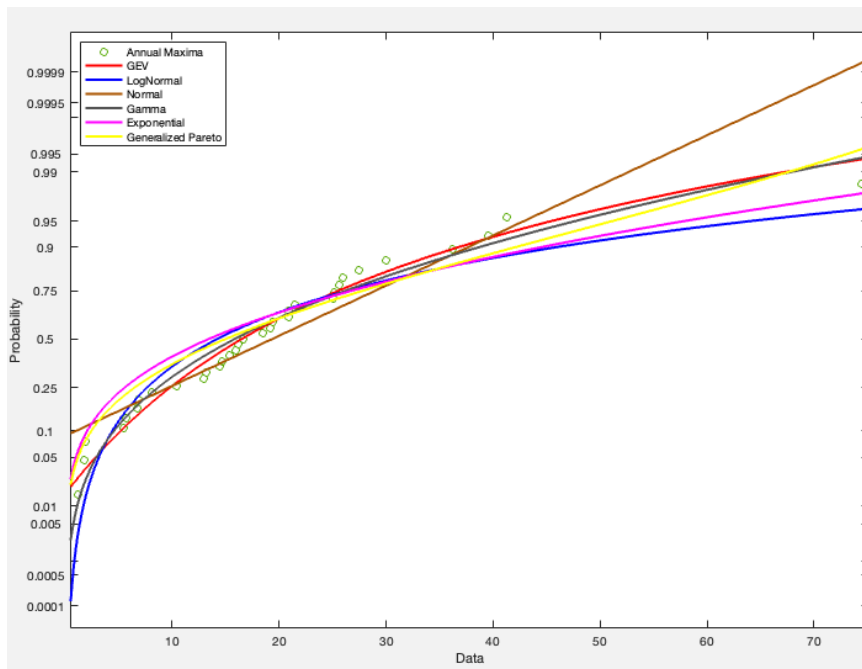

Figure 6: annual maxima and analytical distributions for a sample section.

1.4 Hydraulic modelling

Central component of a GFM is the hydraulic model that simulates, to varying degrees of complexity, the physics of fluid flows. The hydraulic model applied here to compute flood hazard is based on a simplified approach which operates in one-dimension along the global river network, by solving the Manning equation at regularly spaced points along the centerline of each river channel (Rudari et al., 2015). The approach is based on the computation of the channel uniform flow depth, by assuming that slope, geometry and friction of each river reach can be considered as constant values. This assumption is reasonable for relatively short reaches, meaning having densely distributed river sections along the river network. In such case, the flow depth and magnitude can also be considered as constant in time. The Manning equation can be expressed as:

$$Q = \frac{1}{n} A \cdot R^{\frac{2}{3}} \cdot S^{\frac{1}{2}} \tag{2}$$

where:

- n is the Manning roughness coefficient and reflects the friction factor of the riverbed. Manning n values are empirical and are related to the dominant land cover in the simulation area.

- A is the surface of the flow through the river cross section.

- R is the hydraulic radius. It is computed as the flow area divided by the wetted perimeter.

- S is the channel section slope.

River cross sections are generated at approximately every kilometer from the original MERIT Hydro DEM at 90m resolution, centered on the riverbed line. An iterative process is used to solve the equation (2) for the considered modelled discharges at each river section and for a set of nine return periods including the 1 in 2, 5, 10, 25, 50, 100, 200, 500, 1000 years. Each cross section is incrementally flooded using the chosen discharge peak, starting from the center point, until water can no longer extend in any direction. Water levels are then used to compute the maximum water depths of flooding in relation to each return period in present and future climate conditions. In a subsequent phase, backwater effects are considered by ensuring that absolute water levels deacrease along the river network and by raising non-compliant values with the corresponding maximum levels found downstream each section. To improve the quality of the output, resulting maps are then used to estimate water levels at a tenfold number of sections equally spaced along the original ones. Flood extent is then recomputed at each new section to refine the resulting map and the process to account for backwater effects is repeated.
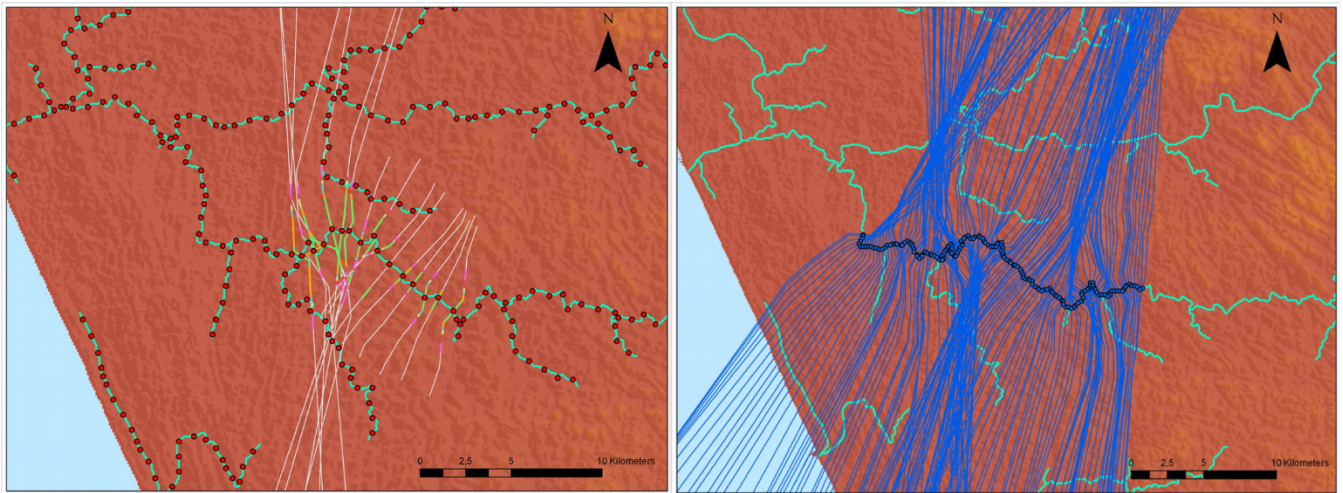
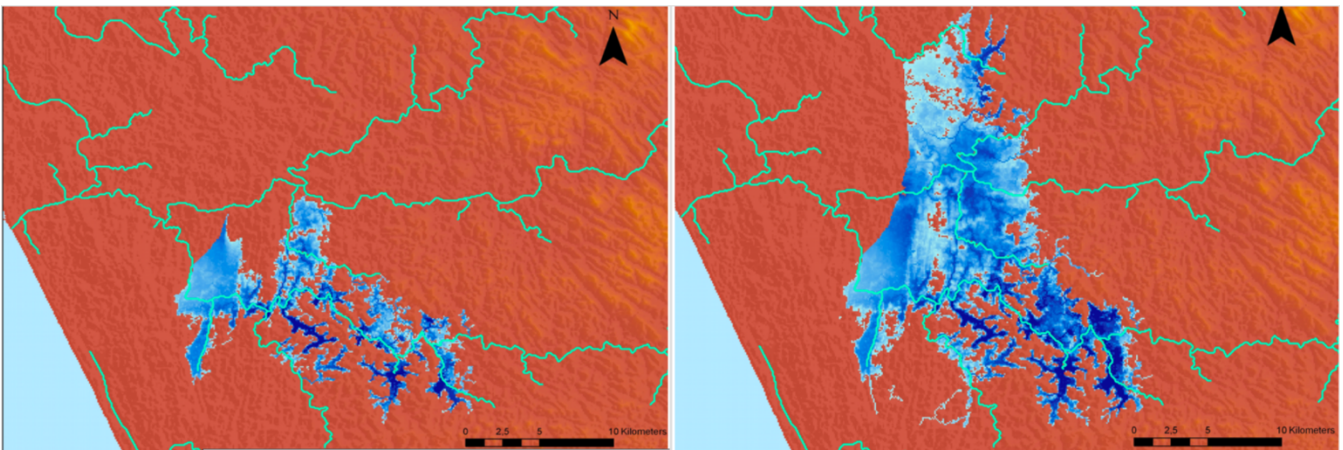**Figure 7: Initial (left) and densified (right) cross sections for a sample stream section in the Kerala region, India.**

220

Figure 8: 1 in 25 year (left) and 500 year (right) return period maps of flood hazard for a sample river section in the Kerala region, India.

3 Results

3.1 Hydrological model calibration

The calibration process yielded good agreement between the modeled discharge and the observed data, resulting in accurate reconstruction of the flow duration curves across the considered watersheds and sections. This improvement during calibration is visually demonstrated in Figure 9 for 3 sample sections, each displaying a scatterplot of observed discharge (Q observed) versus modeled discharge (Q modeled) for a specific hydrometric station. The scatterplots include blue dots representing previous simulations conducted during the iterative exploration of the calibration process, while red dots indicate the optimal results. The bisector indicates perfect match. The represented values in the scatterplots correspond to predetermined quantiles of the observed and modeled discharge time series, which form the respective duration curves.
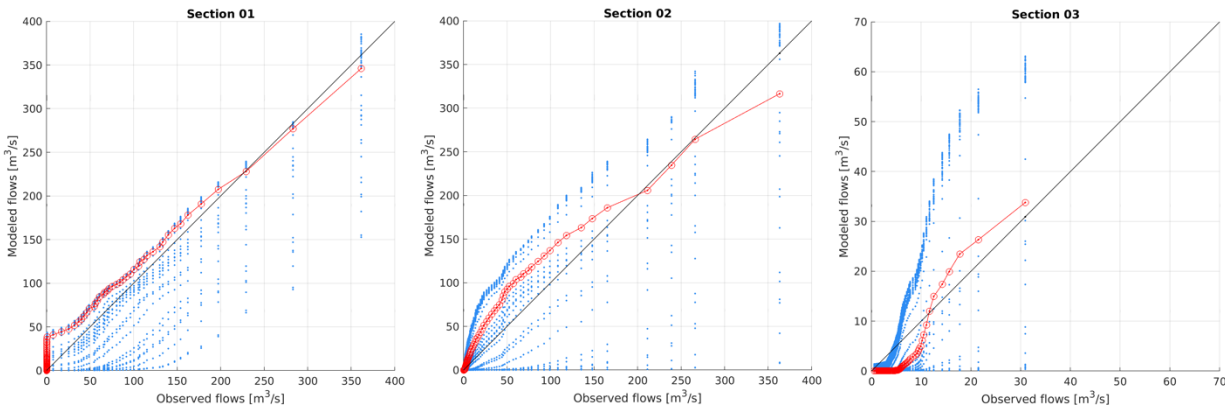

Figure 9: Scatterplot of observed versus modelled quantiles of the flow duration curves at 3 sample river sections. Blue dots represent simulation results obtained during the exploration phase, while the red curve is the calibration result.

Given the scale of application and the relatively coarse resolution of the meteorological input data, performance in larger river basins are on average better. In addition, we found higher quality of observed discharge time series in the downstream sections of most rivers. Examples of observed versus simulated time series in the Congo River, Nile River and Amazon River are shown in Figure 10, together with the corresponding Q-Q plot.
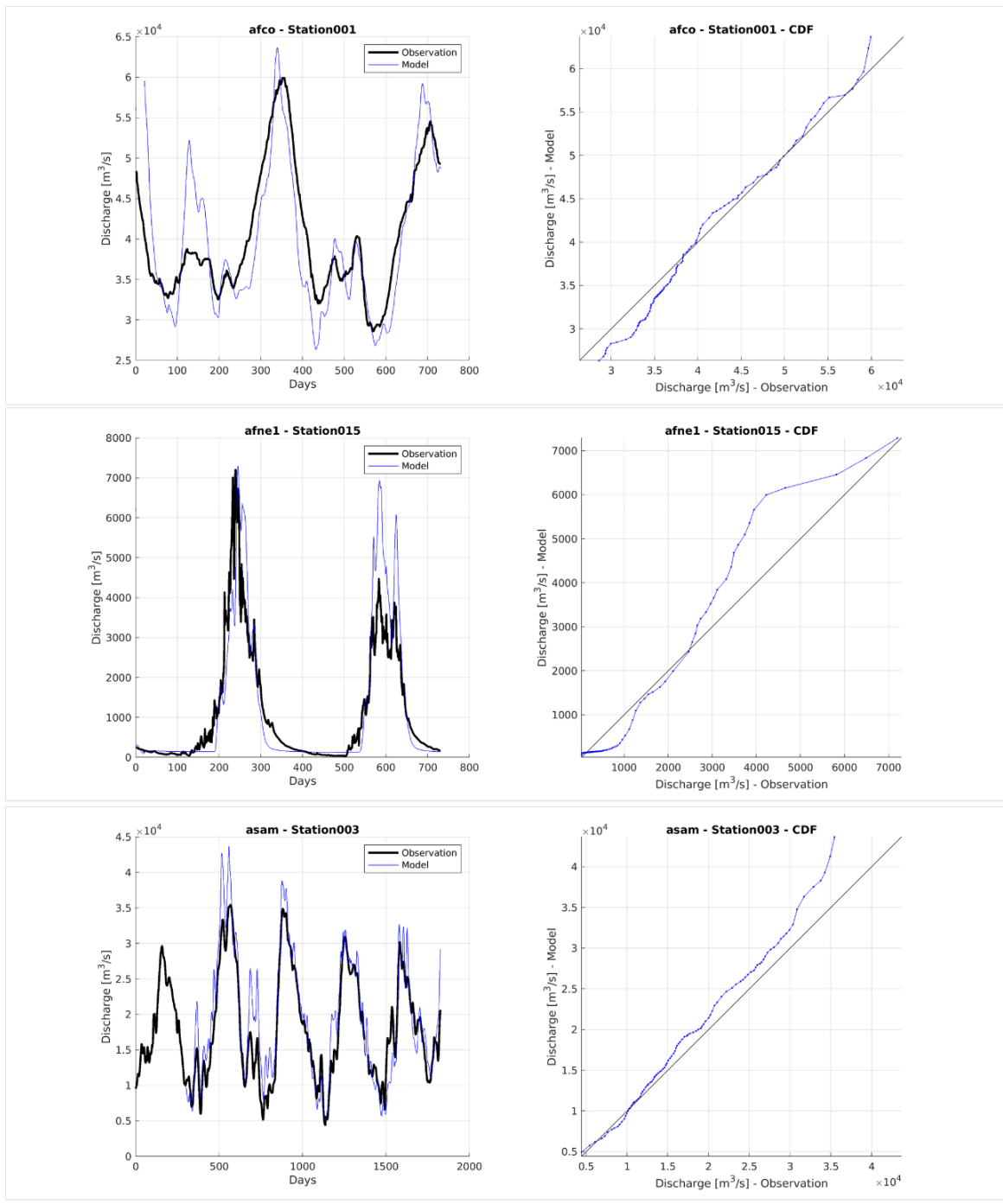
Figure 10: Observed versus simulated discharge time series and corresponding Q-Q plot for 3 sample stations in the Congo River (afco – Station001), Nile River (afne – Station005) and Amazon River (asam – Station003).

Overall across the 3066 calibration stations, the median KGE of discharges was -0.11, while 69% of stations achieved skillful

245 performance, above the no-skill threshold value of -0.41, as defined by Knoben et al., (2019). A median correlation of 0.59 indicates particularly good correlations between observed and modeled discharges (
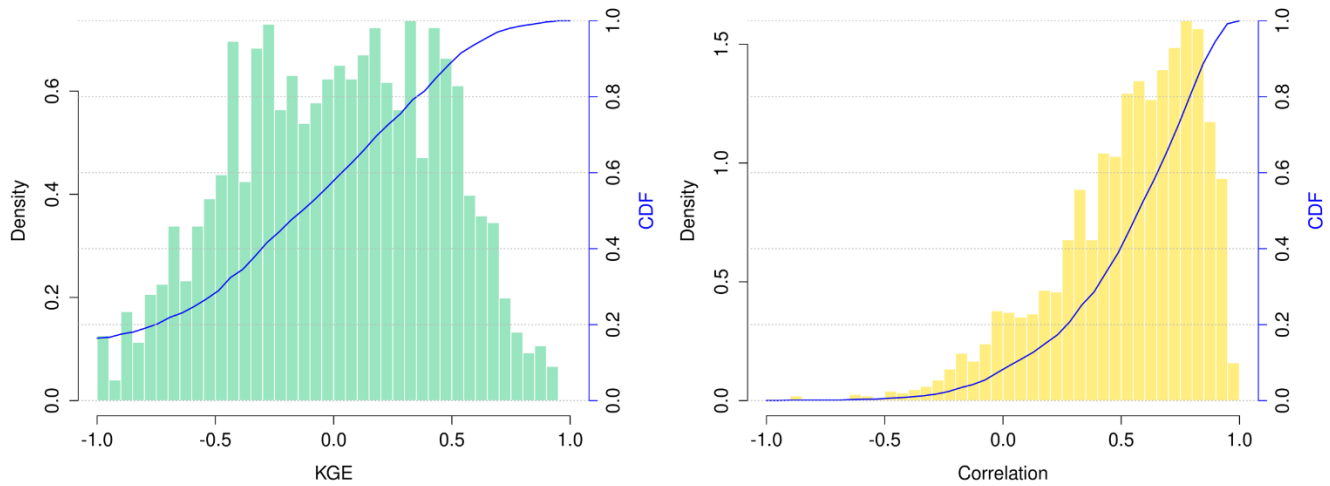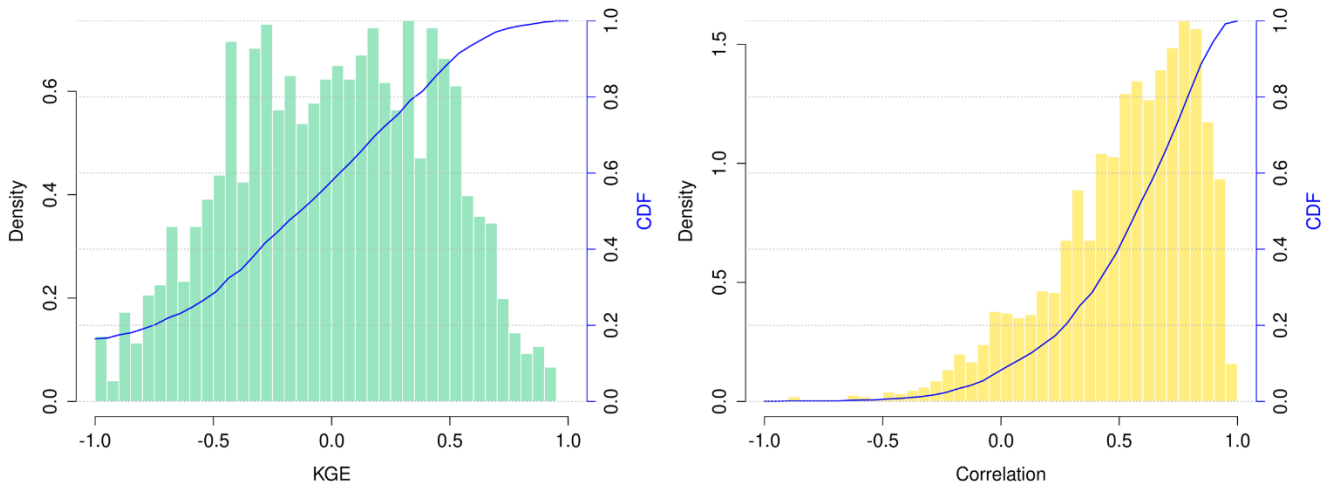

Figure 11), indicating that the sub-optimal KGE scores are related with some residual simulation bias. The spatial distribution

250 of the simulation performance is relatively evenly distributed across the globe (Figure 12 and Figure 13). A small but persistent deterioration of simulation skills is generally visible in the driest areas in Southern Africa, Argentina, Rocky Mountains, Mexico, Middle East and central Asia, where also the correlation is reduced. In addition, some stations with poorer performance are those where the observed discharges are the farthest in the past, when the atmospheric data have on average lower performance, particularly at the local level and for small river basins, due to the limited coverage of satellite data. This

255 is a recurrent situation, given the persistent decline in the global-scale availability of observed discharge data, witnessed from the 1980s onwards (see e.g., Do et al., 2018). Overall, calibration performance compare within the range of published state of the art global hydrological models (Alfieri et al., 2020; Beck et al., 2016; Harrigan et al., 2020), though in this work we have included a significantly higher number of hydrometric stations (3066 versus ~1000-2000 in similar published works). On average, multi-site calibrations as those performed in this work exhibit lower performance during the calibration period

260 compared to cascading calibrations. However, they produce enhanced and consistent basin-wide performance, resulting in significant skill improvement in uncalibrated rivers and overall outside of the calibration period, as it is most needed in operational systems (Wi et al., 2015).

Figure 11: histogram (Density) and cumulative distribution (CDF) of KGE (left) and correlation (right) of modeled versus observed discharges at the 3066 calibration sections.
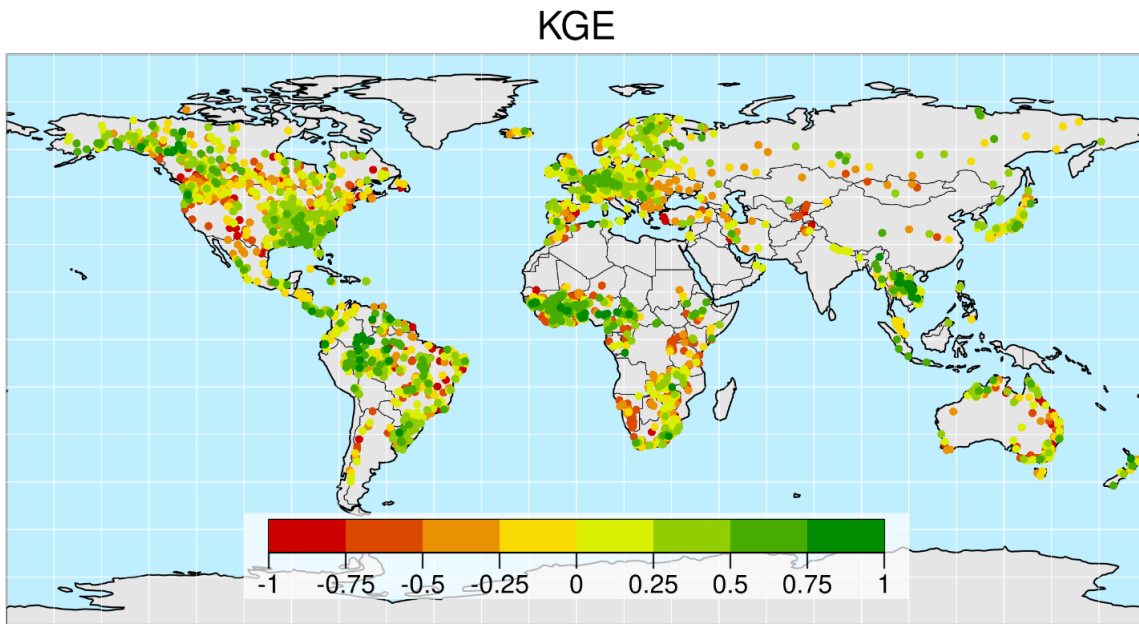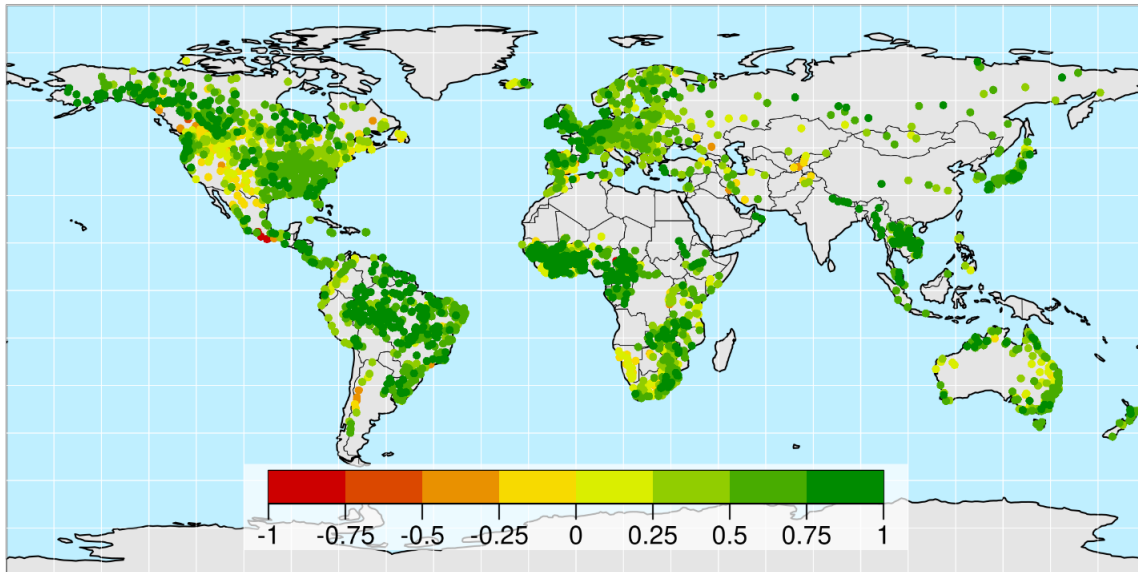

Figure 12: Map of the KGE of modeled versus observed discharges at the 3066 calibration sections.

Correlation

Figure 13: Map of the correlation of modeled versus observed discharges at the 3066 calibration sections.

An additional example, related to soil moisture comparison after model calibration is given in Figure 14 for Madagascar, while Figure 15 shows a comparison between observed and simulated evapotranspiration in the Amazon river basin after calibration. In the latter case the model shows a tendency to underestimate average seasonal evapotranspiration with respect to the GLEAM dataset used as benchmark.
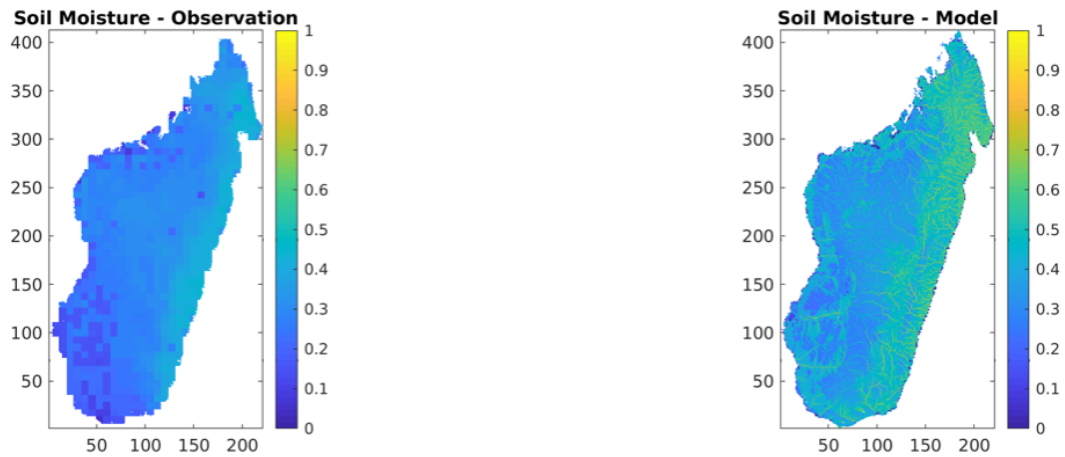
Figure 14: Observed and modeled soil moisture for Madagascar after calibration.
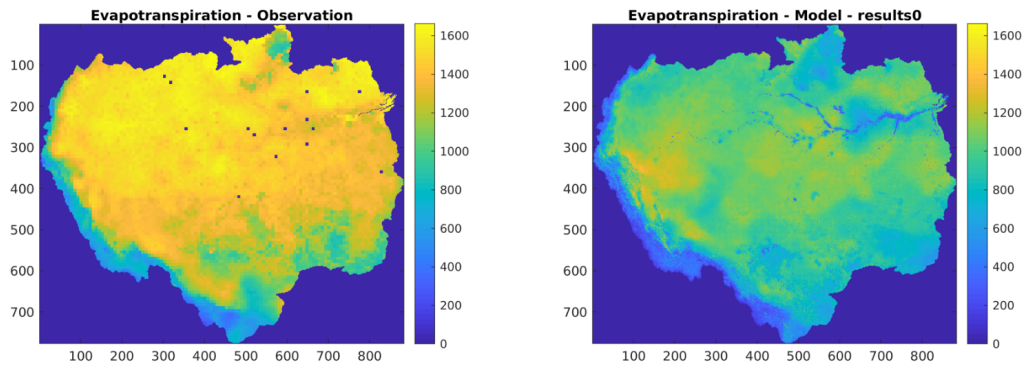

Figure 15: Observed and modeled evapotranspiration in the Amazon river basin after calibration.

After calibrating the model parameters, the hydrological model Continuum was run for a period of 120 years in total (1979-2100) to produce a seamless hydrological simulation, for the historical and both selected future scenarios. The historical period (1979-2016) was run with the meteorological input provided by W5E5 datasets, while the future period (2017-2100) was run with both the SSP126/IPSL-CM6A-LR and SSP585/IPSL-CM6A-LR inputs. From these simulations the discharge values in each point of the simulated river network are extracted so that extreme value statistics are used to extract quantiles of discharge peaks to feed the hydraulic model simulations.

4 Conclusions

The procedure presented in this paper provides a valuable tool for producing global flood inundation maps under present and future climate conditions. A sample of the output at six locations is shown in Figure 16 for the historical climate and the highest of the two chosen warming scenarios. The hydro-meteorological and hydraulic modeling chain used has the advantage of being able to simulate all components of the water cycle in a continuous and physically-based manner. This allows for a more accurate assessment of flood risk in different regions of the world, which is crucial for disaster risk reduction and flood early warning systems. The ability to produce impact assessments under different climate change scenarios further emphasizes the importance of this product for decision-making and planning purposes. Results of this study will contribute to a better understanding of global flood risk and support efforts to reduce the impact of flooding on vulnerable populations and infrastructure, both within the GIRI initiative as well in research and operational products.
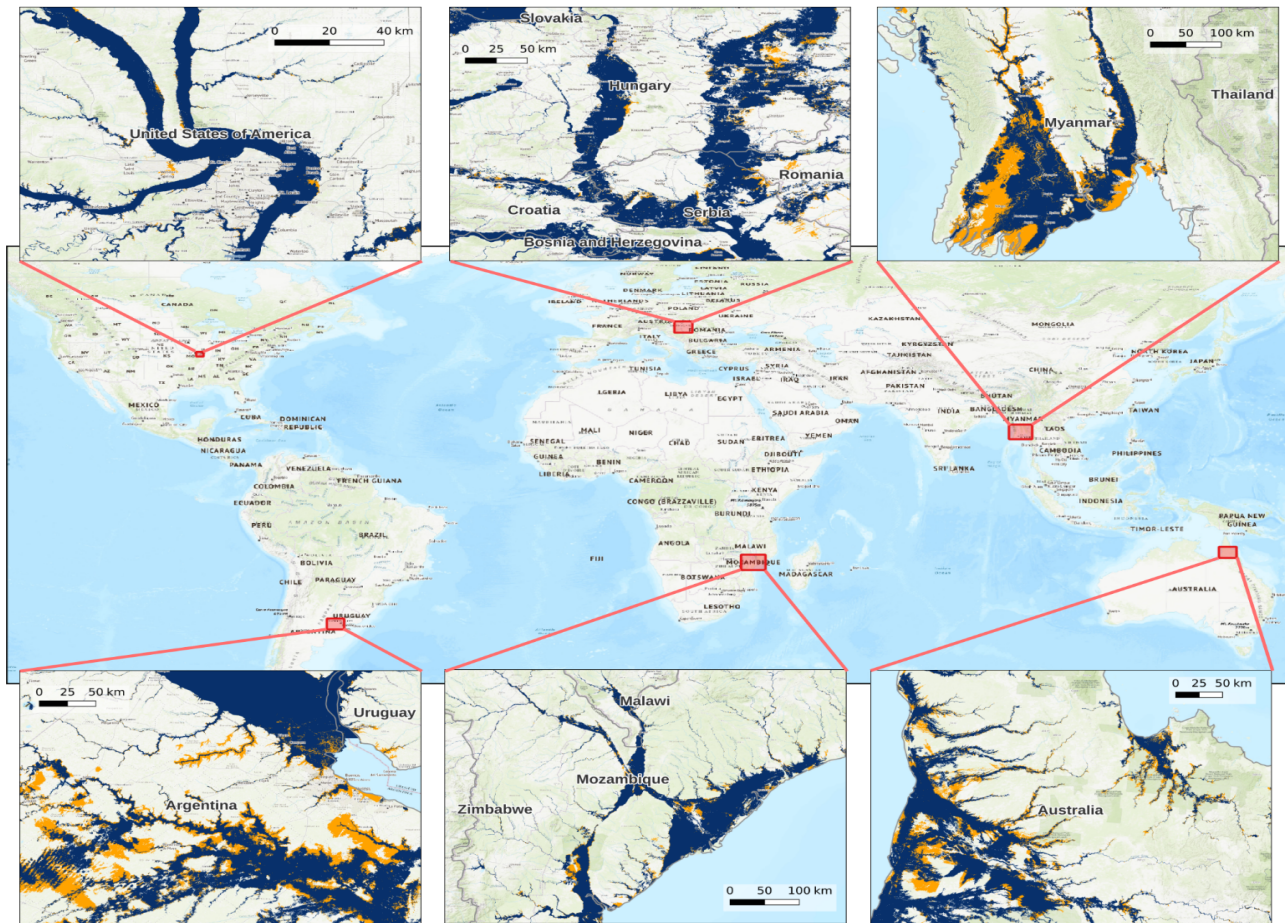
Figure 16: modeled 1 in 100 year maximum flood extent in the historical (1979-2016, in blue) and future scenario (1951-2100) under SSP585 (in orange).

Flood scenarios generation

300   The Hazard maps created in combination with the hydrological simulation performed for their computation can be combined for the generation of all possible flood events that can affect the areas of interest: the hazard maps provide water levels in flood prone areas for different return periods, but they do not represent flood events. A flood event or flood scenario usually affects only a portion of the country. The distinction between flood map and flood scenario is fundamental: flood risk estimates only based on flood maps are reliable if the area of interest is relatively small but, if the area is wide (e.g., country or regional level),

305   it is necessary to generate all possible flood scenarios that can affect the area of interest with their probability of occurrence.

To apply the scenarios' simulation engine, the domain was divided in Pertinence Areas (PA): PA are polygons defined at the resolution of the inundation model which define a univocal link with each stream of the river network in the hydrological model. PA retain hydrologically coherence by assigning each portion of the flooded area to the corresponding stream of the causative inflow hydrograph which produced the highest flood depth among all hydraulic simulations.

310 For each of the PA, an analysis of the modelled discharge time series was done to select independent flood events.

The methodology used for the events generation relies on a multivariate statistical approach that takes in input the selected events and, by preserving their spatial correlation, it is able to simulate events not yet observed both in terms of intensities as well as geographical distribution. The approach used for the events generation covers all the possible range of intensities and spatial dependencies and assures that:

315
- the spatial correlation of small- and large-scale events is preserved in the simulated event set;
- the statistical properties of the observed events at each location is preserved in the simulated event set.

The scenario generation process consists of two components: the first one is the event definition and selection and the second one is the probabilistic events generation.

In Figure 17 an example of selected and simulated events is shown for a couple of Pertinence Areas: the selected events are
320 shown in red, the simulated ones in blue. The behavior of the two marginals is provided for the probability domain (on the left), the space of the physical variables (in the middle) and in terms of return period (on the right). The algorithm allows us to capture and describe accurately the correlation between the different PA. The strength of this approach is linked with the capability of the scenarios modelling of preserving the statistical properties (i.e., the simulated event assures that the marginal distributions are preserved during the simulation process) of what has been observed, on one side, but going beyond that, on
325 the other side.

21
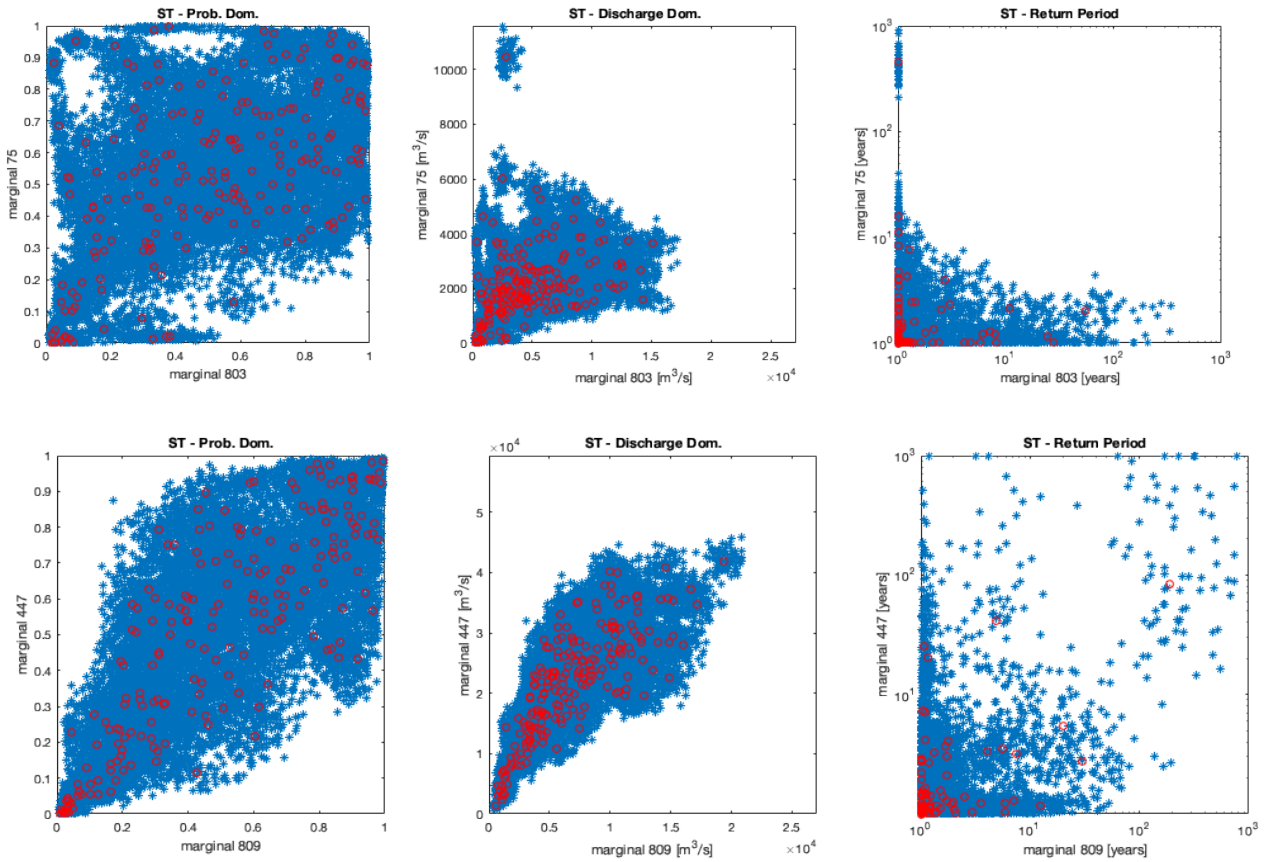
Figure 17 Example of flood events for two Pertinence Areas in the probability domain (column 1), in the space of the physical variables – discharges (column 2) and in terms of return periods of the flood events (column 3).

This flood hazard analysis is aimed at the computation of the loss of all the synthetic scenarios that are generated according to the described procedure and, ultimately, to compute the loss statistics (AAL, etc.). However, two issues arise: 1) the loss statistics are of interest at country or province scale, that are territorial subdivisions not coincident with the hydrological watershed, 2) the computation of the loss of a large number of synthetic scenarios, that is required in order to obtain reliable statistic, is a computationally challenging task. For these reasons, the CIMA procedure is optimized by performing a pre-computation of the minimum possible number of losses that, combined, can describe all possible loss scenarios and, at the same time, are shaped in order to consider the territorial characterization of interest.

This is achieved by furtherly subdividing the domain of interest in the Minimal Units of Loss (MUL): the MUL are a territorial subdivision obtained by crossing the map of Pertinence Areas (that retain all the hydrological information) with the map of the administrative boundaries of interest for the analysis (e.g., Admin1 boundaries of the countries that overlap the hydrological domain of interest). The result is a finer mapping of the territory in which different MUL can belong to the same administrative

22

territorial unit because they can be interested by floods from different streams of the hydrological network, or, at the same time, different MUL can belong to the same Pertinence Area but to different administrative units, that is desired because the aggregation level at which loss data are of interest requires this level of territorial detail. Figure 18 shows an example of MUL map obtained by the combination of the maps of Pertinence Areas and administrative boundaries.

Pertinence areas   **Administrative boundaries**   **MUL (Minimal Units of Loss)**
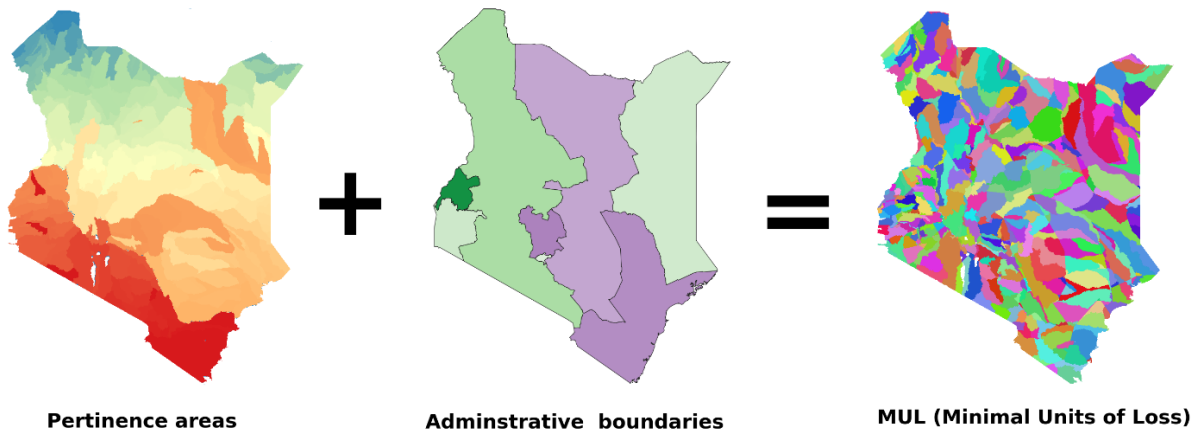
Figure 18 Example of computation of the Minimal Units of Loss map combining the Pertinence Areas (cut out with the national borders) with the administrative boundaries. Example for Kenya.

Once this subdivision is complete, the loss estimate procedure will require the computation of the loss for each MUL at each of the return periods for which the flood hazard maps are available, yielding to a Table of Loss for all the possible combinations of return period and geographical location (MUL) inside the area of interest (e.g., a country). The synthetic flood scenarios previously generated are then employed to compute the corresponding losses by combining the different data from this table, i.e. combining the losses of the return periods present in the single synthetic scenario in the MUL in which they occur, and then combining them in order to obtain the total loss information for the categories of interest (e.g. the different types of exposure) and with the desired territorial aggregation (e.g. Admin1 boundaries).

In this way the maximum possible detail of the loss analysis is achieved with a fraction of the computational effort that would be required with a massive direct loss computation for all the synthetic scenarios generated.

23

References

Alfieri, L., Lorini, V., Hirpa, F. A., Harrigan, S., Zsoter, E., Prudhomme, C., and Salamon, P.: A global streamflow reanalysis for 1980–2018, Journal of Hydrology X, 6, 100049, https://doi.org/10.1016/j.hydroa.2019.100049, 2020.

Alfieri, L., Libertino, A., Campo, L., Dottori, F., Gabellani, S., Ghizzoni, T., Masoero, A., Rossi, L., Rudari, R., Testa, N., Trasforini, E., Amdihun, A., Ouma, J., Rossi, L., Tramblay, Y., Wu, H., and Massabò, M.: Impact-based flood forecasting in the Greater Horn of Africa, EGUsphere, 1–19, https://doi.org/10.5194/egusphere-2023-804, 2023.

Andreadis, K. M., Schumann, G. J.-P., and Pavelsky, T.: A simple global river bankfull width and depth database, Water Resources Research, 49, 7164–7168, https://doi.org/10.1002/wrcr.20440, 2013.

Bates, P. D., Neal, J., Sampson, C., Smith, A., and Trigg, M.: Progress toward hyperresolution models of global flood hazard, in: Risk modeling for hazards and disasters, Elsevier, 211–232, 2018.

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Miralles, D. G., McVicar, T. R., Schellekens, J., and Bruijnzeel, L. A.: Global-scale regionalization of hydrologic model parameters, Water Resources Research, 52, 3599–3622, https://doi.org/10.1002/2015WR018247, 2016.

Cenci, L., Laiolo, P., Gabellani, S., Campo, L., Silvestro, F., Delogu, F., Boni, G., and Rudari, R.: Assimilation of H-SAF Soil Moisture Products for Flash Flood Early Warning Systems. Case Study: Mediterranean Catchments, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9, 5634–5646, https://doi.org/10.1109/JSTARS.2016.2598475, 2016.

Corral, C., Berenguer, M., Sempere-Torres, D., Poletti, L., Silvestro, F., and Rebora, N.: Comparison of two early warning systems for regional flash flood hazard forecasting, Journal of Hydrology, 572, 603–619, https://doi.org/10.1016/j.jhydrol.2019.03.026, 2019.

Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Lange, S., Müller Schmied, H., Hersbach, H., and Buontempo, C.: WFDE5: bias-adjusted ERA5 reanalysis data for impact studies, Earth System Science Data, 12, 2097–2120, https://doi.org/10.5194/essd-12-2097-2020, 2020.

Desai, B., Maskrey, A., Peduzzi, P., De Bono, A., and Herold, C.: Making development sustainable: the future of disaster risk management, global assessment report on disaster risk reduction, 2015.

Do, H. X., Gudmundsson, L., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM) – Part 1: The production of a daily streamflow archive and metadata, Earth System Science Data, 10, 765–785, https://doi.org/10.5194/essd-10-765-2018, 2018.

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A., and Feyen, L.: Development and evaluation of a framework for global flood hazard mapping, Advances in Water Resources, 94, 87–102, https://doi.org/10.1016/j.advwatres.2016.05.002, 2016.

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

Fan, J., Rosenfeld, D., Yang, Y., Zhao, C., Leung, L. R., and Li, Z.: Substantial contribution of anthropogenic air pollution to catastrophic floods in Southwest China, Geophysical Research Letters, 42, 6066–6075, https://doi.org/10.1002/2015GL064479, 2015.

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

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

Hagos, S. M., Leung, L. R., Yoon, J.-H., Lu, J., and Gao, Y.: A projection of changes in landfalling atmospheric river frequency and extreme precipitation over western North America from the Large Ensemble CESM simulations, Geophysical Research Letters, 43, 1357–1363, https://doi.org/10.1002/2015GL067392, 2016.

Harrigan, S., Zsoter, E., Alfieri, L., Prudhomme, C., Salamon, P., Wetterhall, F., Barnard, C., Cloke, H., and Pappenberger, F.: GloFAS-ERA5 operational global river discharge reanalysis 1979–present, Earth System Science Data, 12, 2043–2060, https://doi.org/10.5194/essd-12-2043-2020, 2020.

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.: SoilGrids250m: Global gridded soil information based on machine learning, PLOS ONE, 12, e0169748, https://doi.org/10.1371/journal.pone.0169748, 2017.

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., Lange, S., Hari, C., Reyer, C. P. O., Conrad, O., Zimmermann, N. E., and Frieler, K.: CHELSA-W5E5: Daily 1 km meteorological forcing data for climate impact studies, Earth System Science Data Discussions, 1–28, https://doi.org/10.5194/essd-2022-367, 2022.

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.

Laiolo, P., Gabellani, S., Rebora, N., Rudari, R., Ferraris, L., Ratto, S., Stevenin, H., and Cauduro, M.: Validation of the Flood-PROOFS probabilistic forecasting system, Hydrological Processes, n/a-n/a, https://doi.org/10.1002/hyp.9888, 2013.

Lange, S.: ISIMIP3BASD v2.5.0, 2021.

Lange, S. and Büchner, M.: ISIMIP3b bias-adjusted atmospheric climate input data (v1.1), 2021.

Lehner, B., Liermann, C. R., Revenga, C., Vörösmarty, C., Fekete, B., Crouzet, P., Döll, P., Endejan, M., Frenken, K., Magome, J., Nilsson, C., Robertson, J. C., Rödel, R., Sindorf, N., and Wisser, D.: High-resolution mapping of the world's reservoirs and dams for sustainable river-flow management, Frontiers in Ecology and the Environment, 9, 494–502, https://doi.org/10.1890/100125, 2011.

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.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O.: Estimating the volume and age of water stored in global lakes using a geo-statistical approach, Nat Commun, 7, 13603, https://doi.org/10.1038/ncomms13603, 2016.

Mulligan, M., Lehner, B., Zarfl, C., Thieme, M., Beames, P., Soesbergen, A. van, Higgins, J., Januchowski-Hartley, S. R., Brauman, K. A., Felice, L. D., Wen, Q., Leaniz, C. G. de, Belletti, B., Mandle, L., Yang, X., Wang, J., and Mazany-Wright, N.: Global Dam Watch: curated data and tools for management and decision making, Environ. Res.: Infrastruct. Sustain., 1, 033003, https://doi.org/10.1088/2634-4505/ac333a, 2021.

O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., Mathur, R., and Vuuren, D. P. van: A new scenario framework for climate change research: the concept of shared socioeconomic pathways, Climatic Change, 122, 387–400, https://doi.org/10.1007/s10584-013-0905-2, 2014.

Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H.: Deriving global flood hazard maps of fluvial floods through a physical model cascade, Hydrol. Earth Syst. Sci. Discuss, 9, 6615–6647, 2012.

Rudari, R., Silvestro, F., Campo, L., Rebora, N., Boni, G., and Herold, C.: Improvement of the global flood model for the GAR 2015, United Nations Office for Disaster Risk Reduction (UNISDR), Centro Internazionale in Monitoraggio Ambientale (CIMA), UNEP GRID-Arendal (GRID-Arendal): Geneva, Switzerland, 69, 2015.

Sampson, C., Smith, A., Bates, P., Neal, J., Alfieri, L., and Freer, J.: A high-resolution global flood hazard model, Water Resour. Res., 51, 7358–7381, https://doi.org/10.1002/2015WR016954, 2015.

Shirazi, M. A. and Boersma, L.: A Unifying Quantitative Analysis of Soil Texture, Soil Science Society of America Journal, 48, 142–147, https://doi.org/10.2136/sssaj1984.03615995004800010026x, 1984.

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.

Silvestro, F., Gabellani, S., Rudari, R., Delogu, F., Laiolo, P., and Boni, G.: Uncertainty reduction and parameter estimation of a distributed hydrological model with ground and remote-sensing data, Hydrology and Earth System Sciences, 19, 1727–1751, 2015.

Trigg, M. A., Michaelides, K., Neal, J. C., and Bates, P. D.: Surface water connectivity dynamics of a large scale extreme flood, Journal of Hydrology, 505, 138–149, https://doi.org/10.1016/j.jhydrol.2013.09.035, 2013.

UNISDR: Sendai framework for disaster risk reduction 2015–2030, United Nations Office for Disaster Risk Reduction: Geneva, Switzerland, 2015.

Ward, P. J., Jongman, B., Salamon, P., Simpson, A., Bates, P., Groeve, T. D., Muis, S., Perez, E. C. de, Rudari, R., Trigg, M. A., and Winsemius, H. C.: Usefulness and limitations of global flood risk models, Nature Climate Change, 5, 712–715, https://doi.org/10.1038/nclimate2742, 2015.

Wi, S., Yang, Y. C. E., Steinschneider, S., Khalil, A., and Brown, C. M.: Calibration approaches for distributed hydrologic models in poorly gaged basins: implication for streamflow projections under climate change, Hydrology and Earth System Sciences, 19, 857–876, https://doi.org/10.5194/hess-19-857-2015, 2015.

Winsemius, H. C., Van Beek, L. P. H., Jongman, B., Ward, P. J., and Bouwman, A.: A framework for global river flood risk assessments, Hydrol. Earth Syst. Sci., 17, 1871–1892, https://doi.org/10.5194/hess-17-1871-2013, 2013.

Wojtkiewicz, R., Rollins, J., and Foley, V.: US flood insurance-the NFIP and beyond, 2013.

Yamazaki, D., Kanae, S., Kim, H., and Oki, T.: A physically based description of floodplain inundation dynamics in a global river routing model, Water Resources Research, 47, W04501, 2011.

Yamazaki, D., Ikeshima, D., Sosa, J., Bates, P. D., Allen, G. H., and Pavelsky, T. M.: MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset, Water Resources Research, 55, 5053–5073, https://doi.org/10.1029/2019WR024873, 2019.