https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

Open Access Earth System Science Data Discussions

HERA: a high-resolution pan-European hydrological reanalysis (1950-2020)

Aloïs Tilloy[1], Dominik Paprotny[2], Stefania Grimaldi[1], Goncalo Gomes[1], Alessandra Bianchi[1], Stefan Lange[2], Hylke Beck[3], Luc Feyen[1]

[1]European Commission, Joint Research Centre, Italy
[2]Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, Potsdam, Germany
[3]King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

Correspondence: Aloïs Tilloy (alois.tilloy@ec.europa.eu)

Abstract

Since 1950, European rivers have been put under increasing pressure by anthropogenic activities, resulting in changes in climate, land cover, soil properties and channel morphologies. These evolving environmental conditions can translate into changes in hydrological conditions. The availability of consistent estimates of river flow at global and continental level is a necessity to assess and attribute changes in the hydrological cycle. To overcome limitations posed by observations (incomplete records, inhomogeneous spatial coverage), we simulate river discharge for Europe for the period 1950 – 2020 using a state-of-the-art hydrological modelling approach. We use the new European set up of the LISFLOOD model, running at 1 arcminute (≈1.8 km) with six-hourly time steps. The hydrological model is forced by climate reanalysis data (ERA5-land) bias-corrected and downscaled to the model resolution with weather observations. The model also ingests 72 surface fields maps representing catchment morphology, vegetation, soil properties, land use, water demand, lakes and reservoirs. Inputs related to human activities are evolving through time to emulate changes in society. The resulting Hydrological European ReAnalysis (HERA), provides six-hourly river discharge for 282 521 river pixels with upstream area > 100km^2. We assess its skill using 2901 river gauging stations distributed across Europe. Overall, HERA delivers satisfying results, with a general weak underestimation of observed mean discharge and flow variability. We find that the performance of HERA increases through time between 1950 and 2020. The fine spatial and temporal resolution result in an enhanced performance compared to other reanalysis for small-to-medium-scale catchments (100-10 000 km^2), with degraded performance remaining for small catchments. HERA is the first long-term, high-resolution hydrological reanalysis for Europe. Despite its limitations, it enables the analysis of hydrological dynamics related to extremes, human influences, and climate change at a continental scale while keeping local relevance. It also creates the opportunity to study these dynamics in ungauged catchments across Europe.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

Introduction

In the last century, Europe has experienced a growth in its population, economy and urban area (Li *et al.*, 2021; Paprotny and Mengel, 2023). Recent decades also witnessed a rapid rise in global air temperature, attributable to anthropogenic activities (IPCC, 2023). These evolving conditions have significantly changed flows in European streams and rivers, leading to multiple challenges for hydrological sciences, related, for example, to long term variability, climate change, extremes or human alterations of the water cycle (Blöschl *et al.*, 2019b). To understand the impacts of these changes, hydrologists need consistent, reliable and long hydrological series. Observations, despite continuous improvements (Blöschl *et al.*, 2019a; Ekolu *et al.*, 2022), are still lacking at high enough spatial density across Europe and are often uncertain and discontinuous. One option to overcome these limitations is to rely on a suit of models (climate, hydrological, land use) to simulate past hydrological conditions and interpret changing dynamics in the hydrological cycle in connection with rapidly changing human systems (e.g., Richards and Gutierrez-Arellano, 2022). This article introduces the Hydrological European ReAnalysis (HERA) for the period 1950-2020, providing consistent estimates of river flow for European rivers at unprecedented spatial and temporal resolution.

Hydrological models are essential tools when it comes to understand and characterise processes related to the water cycle (e.g., flood and drought forecasting). In the past three decades, there has been efforts in developing models that are able to simulate hydrological processes at large scale (continental to global scale). A myriad of these Global Hydrological Models (GHMs), differing in their conceptualization, now exist (Beck *et al.*, 2017; Sood and Smakhtin, 2015; Kauffeldt *et al.*, 2016; Prudhomme *et al.*, 2011). The nature of GHMs implies that they are usually run at coarse spatial resolution (0.5º), limiting their relevance for local and regional water resource problems (Sood and Smakhtin, 2015). Nonetheless, the development of GHMs has been fuelled by continuous improvements in remote sensing technologies and processing power (Yang *et al.*, 2021). Remote sensing technologies now provide high resolution input for hydrological models such as land use and vegetation properties. The advancements in computational capabilities and performances of the models also allows to refine both the spatial and temporal scale of hydrological models, providing more accurate representation of surface and subsurface processes and reducing modelling uncertainties (Wood *et al.*, 2011).

A key limitation preventing to simulate past river flows has been the availability of meteorological inputs for hydrological models. Among potential inputs, climate reanalysis offers several advantages: temporal coverage (typically several decades), numerous, homogeneous environmental variables, and spatial resolution. Reanalysis data are outputs of climate models calibrated on observed data worldwide

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

(Brönnimann *et al.*, 2018). Here we use ERA5-land, the land component of ERA5. A main advantage of ERA5-Land compared to ERA5 is its horizontal resolution, which is 9 km globally, compared to 31 km in ERA5. This enhanced resolution is obtained by downscaling meteorological variables from ERA5. The temporal resolution is hourly as in ERA5. Reanalysis data are obtained from short-term model forecasts and can be affected by forecast errors; they are not observations (Pfahl and Wernli, 2012). Furthermore, variables produced in ERA5 are averages over grid cells. This averaging combined with the relatively coarse resolution of ERA5/ERA5-land often smooths local extremes (Donat *et al.*, 2014, Tilloy *et al.*, 2022). To tackle this issue, we downscale and bias-correct ERA5-land with a gridded observational dataset, EMO-1 (**Section 2.2**).

In the context of the European Flood Awareness System (EFAS), an operational system for European flood monitoring and forecasting (https://www.efas.eu), there has been efforts in recent years to develop more detailed surface fields (e.g., land use, vegetation) (Choulga *et al.*, 2023a) and observational climate inputs (Thiemig *et al.*, 2022) at a spatial resolution of 1 arcminute (1',0.0167º, typically 1.5–3 km² over Europe). These developments come along with improvements on the LISFLOOD hydrological model underpinning EFAS. LISFLOOD is a spatially distributed grid-based hydrological and channel routing model which was initially developed for flood forecasting and flood risk assessment. However, it is also able to model effects of land use change, climate change and river regulation measures (Burek *et al.*, 2013) and has been used in a wide range of hydrological applications, such as mapping population under water stress in relation to how much water is reserved for the environmental (Vanham *et al.*, 2021) and projecting droughts in view of climate change (Cammalleri *et al.*, 2020a). It was also used in the generation of the GLOFAS hydrological reanalysis (Harrigan *et al.*, 2020).

This article brings together improvements from diverse fields (i.e., remote sensing, climate modelling, machine learning, hydrology) to generate a state-of-the-art hydrological reanalysis of Europe for a domain covering EU27 countries, UK, Switzerland, Iceland, Norway and the Balkan countries (Serbia, Montenegro, Bosnia-Herzegovina, Kosovo, North Macedonia and Albania) over the past 70 years. HERA aims to reproduce as accurately as possible the evolution of the hydrological landscape of Europe by using the latest development of LISFLOOD (Improved processing speed, spatial and temporal resolutions), also used in the generation of the latest EFAS v5.0 reanalysis (1991-2022) (Decremer *et al.*, 2023) (**Section 2.1**). Climate inputs are derived from ERA5-land, bias corrected and downscaled to 1 arcminute to improve the representation of extremes (**Section 2.2**). We generated dynamic socioeconomic inputs (water demand, land use and reservoir maps) to capture the effect of human activities on the water cycle (**Section 2.3**). These developments make this simulation the first long-term Pan-European hydrological reanalysis taking

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

into accounts the evolving socioeconomic conditions of the continent since 1950. In **Section 3**, we assess the performance of HERA against observations from 2901 river gauges in Europe.

Method

The modelling framework developed to generate the HERA dataset is presented in a flowchart in **Figure 1**. The framework is organized around the LISFLOOD hydrological model, which is used to simulate river discharge. For this run, we use calibrated parameters for the European setting of LISFLOOD, developed by ECMWF in the context of the EFAS-5 calibration (CEMS-Flood online documentation, 2023). We fist introduce the LISFLOOD hydrological model and its calibration procedure (**Section 2.1**). **Figure 1** also displays the main input of LISFLOOD: high-resolution climate inputs (**Section 2.2**), state-of-the-art static maps (**Section 2.3.1**), dynamic socioeconomic maps (**Section 2.3**).

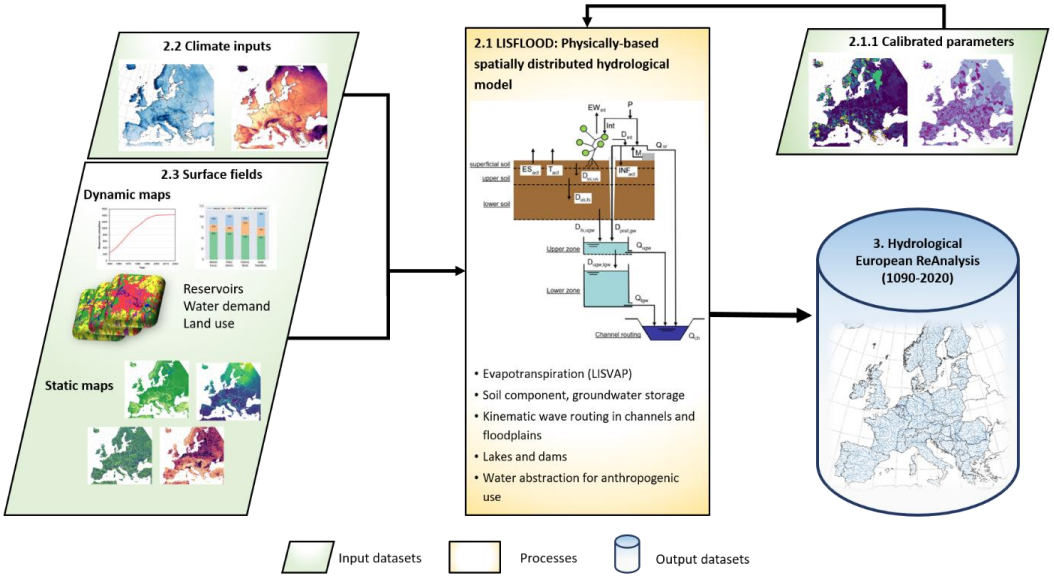


Figure 1: Flowchart of the framework employed in the generation of HERA. Numbers relate to the section in which each component of the framework is presented.

1.1 Hydrological modelling

The LISFLOOD model

Here we simulate sub-daily continuous streamflow time series over Europe (**Figure 1**) by means of the LISFLOOD model. LISFLOOD (Burek *et al.*, 2013; Van der Knijff *et al.*, 2010) is a spatially distributed, semi-physical rainfall-runoff model combined with a routing module for river channels (Dottori *et al.*, 2022). The model has been developed by the JRC (Joint Research Centre) since the late 1990s. The model

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

is used operationally for large-scale flood forecasting in the European Flood Awareness System (EFAS, efas.eu) and the Global Flood Awareness System (GLOFAS, globalfloods.eu). However, LISFLOOD has also been used in drought monitoring (Cammalleri *et al.*, 2020b, 2017), to assess the effect of flood adaptation measures, environmental flow protection, or climate change (Burek *et al.*, 2013; Mentaschi *et al.*, 2020; Vanham *et al.*, 2022). Since 2019, the model is open source and available on github along with a set of auxiliary tools (https://github.com/ec-jrc/lisflood-code). LISFLOOD is composed of the following main components:

- 3 soil layers (superficial, upper, lower) for water balance modelling;
- sub-models for the simulation of groundwater and subsurface flow (using 2 parallel interconnected reservoirs);
- a sub-model for the routing of surface runoff to the nearest river channel;
- a sub-model for the routing of channel flow.

Other processes such as snow melt, infiltration, interception of rainfall, leaf drainage, evaporation and water uptake by vegetation, surface runoff, and exchange of soil moisture between soil layers are also simulated by the model (LISFLOOD OS online documentation, 2023). LISFLOOD is also able to model lakes and reservoirs.

In this work, we use the latest version of LISFLOOD (v4.1.2, January 2023), which includes upgrades compared to previous versions in the hydrological routines and improvements in the modelling of water abstraction for anthropogenic use. Moreover, LISFLOOD v4.1.2 benefits of improvements in the management of large inputs and in the computational performances. **Figure 2** displays the domain on which the hydrological simulation is performed, which includes Europe, North Africa and Middle East. However, the data was retained for a smaller domain, comprising 42 European countries, excluding non-EU countries of the former Soviet Union, countries in North Africa and Middle East, and Turkey, resulting in the same domain as the HANZE study (Paprotny and Mengel, 2023; Paprotny *et al.*, 2023). We run the model using a 1' grid, which is also used in EFAS-5 (Decremer *et al.*, 2023). The temporal resolution of the simulation is 6-hourly, which is the standard for EFAS since 2020. Due to the size and spatial resolution of our domain combined with the 6-hourly time-steps, we divide the simulations in 71 yearly chunks starting on 3 January 1950 following a 71-year pre-run. Due to rapidly evolving socioeconomic conditions in catchments of Europe, we change the input socioeconomic maps with the start of every new calendar year of the simulation (**Section 2.4**). This differs from the standard EFAS settings, which assume static land use and reservoir network, and only varies the water demand values. We retain river pixels with an upstream area greater than 100 km^2, resulting in simulations in the 282 521 river pixels displayed in **Figure 2**.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

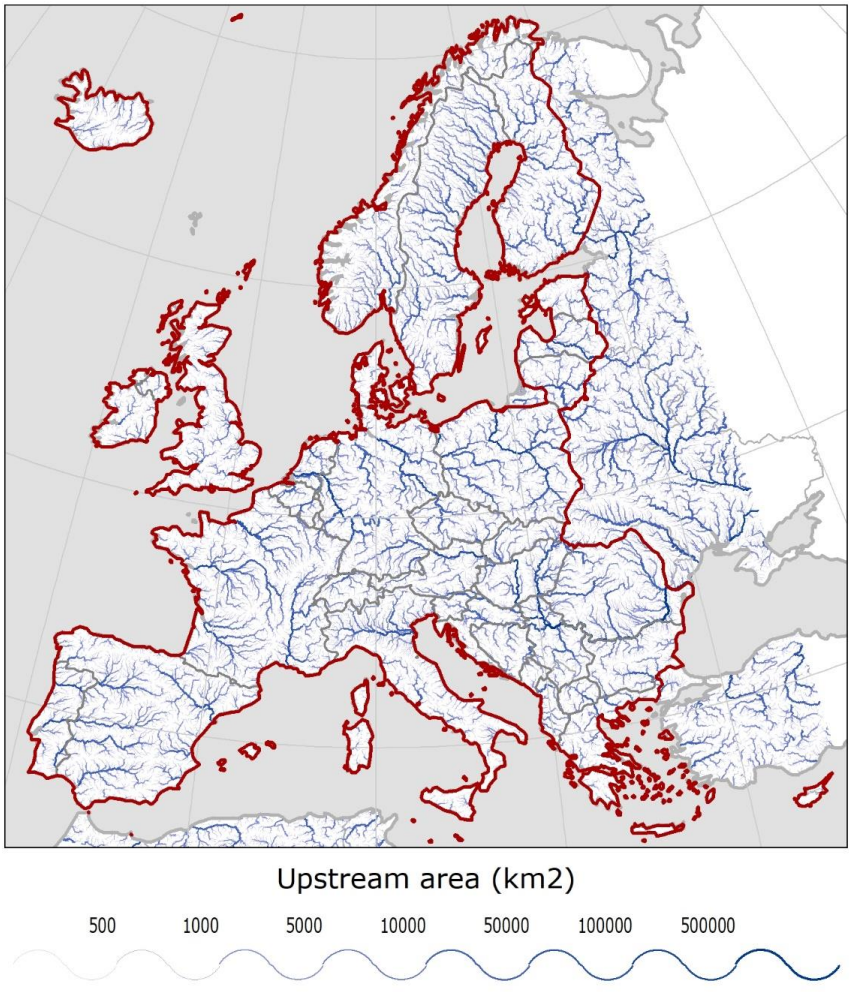


Figure 2 River network (rivers with an upstream area > 100km^2) on which discharge data has been generated. The HERA domain (in which data is provided) is confined by the red bordered area.

Model Calibration

In this work, we also take advantage of the new EFAS-5 calibration that was completed in December 2022 by ECMWF. The calibration was performed using the EMO-1 meteorological dataset (Thiemig *et al.*, 2022) over the period 1990-2021. Daily and/or six-hourly data discharge at a total of 1903 stations, identified through a selection process based on several criteria (CEMS-Flood online documentation, 2023) were used to calibrate the LISFLOOD model over Europe. The calibration was performed at catchment level, with catchments of the 1903 selected stations entailing 69.6% of the HERA domain, a map showing the extent of calibrated and uncalibrated catchments is provided in **Supplementary Figure S1**. The calibration has

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

been performed on 14 parameters that influence the modelling of snow melt, water infiltration into the soil, surface water flow, groundwater flow, lakes and reservoirs dynamics. These parameter were allowed to vary within physically realistic ranges. A list of the calibration parameters is provided in **Supplementary Table S1**.

Coastal and endorheic catchments with drainage area smaller than 150 km^2, representing 6.5% of the HERA domain, are modelled with default parameter values. Parameter values for other ungauged catchments were estimated by parameter regionalisation. These catchments are mostly located near the coastlines, southern Italy and Greece, and represent 23.9% of the HERA domain. The parameter regionalization consists in transferring parameter values (except the ones linked to reservoirs and lakes) from a calibrated catchment to an ungauged catchment. Catchments are matched according to climatic and geographical similarities (Beck *et al.*, 2016). For more information on the calibration of EFA-5, we refer to the online documentation of the Copernicus Emergency Management Service for floods (CEMS-Flood online documentation, 2023).

Climate inputs: Bias-adjusted climate reanalysis data

To force the hydrological model LISFLOOD, we use a modified version of the climate reanalysis dataset ERA5-land (Muñoz-Sabater *et al.*, 2021). The main steps involved in the preparation of the climate inputs are summarized in **Figure 3**. The following variables are retrieved from ERA5-land at hourly temporal resolution for 1950-2020:

- Total precipitation (tp)
- Mean temperature (ta)
- Mean zonal and meridional wind speed (u, v)
- Mean dew point temperature (td)
- Total surface solar radiation downwards (ssrd)

Precipitation and temperature data were aggregated to 6-hourly resolution, and the other variables to daily resolution (**Figure 3**). Minimum and maximum daily temperature were also calculated, while dew point temperature was converted into relative humidity and actual vapour pressure.

ERA5-Land data is too coarse for the grid resolution used in LISFLOOD (**section 2.1**), which requires meteorological data with a 1’ resolution. To downscale ERA5-Land data from 0.1° = 6’ to 1’, we therefore performed statistical downscaling and bias adjustment using ISIMIP3BASD v3.0.0 (Lange 2019, Lange 2022, Frieler *et al.* 2024). The ISIMIP3BASD method was initially developed for phase 3 of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) and aims to provide robust bias adjustment of

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

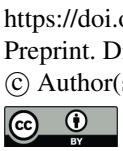
extreme values, preservation of trends across quantiles, and a clearer separation of bias adjustment and statistical downscaling compared to its predecessors (Lange, 2019). Particularly the robust bias adjustment of extremes is relevant in the context of this analysis. We use the new EMO-1 gridded observational dataset (1' version of EMO-5, Thiemig *et al.* 2022) developed for the operational EFAS-5 as the high-resolution reference dataset. EMO-1 covers the period 1990−2020 and has also been used directly as climate inputs for the calibration of LISFLOOD. We use 1990−2020, where both datasets overlap, as the training period for the algorithm. . The trained algorithm is then applied to ERA5-Land to produce high-resolution data for both the training period and 1950−1989, where high-resolution data comparable to EMO-1 are not available. The resulting climate data consistently cover 1950−2020. The ISIMIP3BASD method is applied on the following variables:

- daily mean near-surface relative humidity (hurs), obtained from actual vapor pressure (vp),
- daily and 6-hourly total precipitation (pr),
- daily total surface downwelling shortwave radiation (rsds),
- daily mean near-surface wind speed (ws),
- daily and 6-hourly mean near-surface air temperature (tas),
- diurnal near-surface air temperature range (tasrange = tasmax − tasmin),
- diurnal near-surface air temperature skewness (tasskew = (tas − tasmin)/tasrange).

Here, tasmin and tasmax are the daily near-surface air temperature minimum and maximum, respectively.

Version 3.0.0 of ISIMIP3BASD differs technically from version 2.5.0 that was used to produce the climate forcing data for phase 3b of the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP3b, Frieler *et al.* 2024), yet both versions produce the same results and we apply version 3.0.0 using the same climate variable-specific parameter settings as for the ISIMIP3b data production (Lange 2021, Frieler *et al.* 2024). ISIMIP3BASD was initially designed to be applied to daily data. In order to bias-adjust and statistically downscale sub-daily (6-hourly pr and tas) data, we apply ISIMIP3BASD independently to every 6-hour period of the day and merge the results in post-processing. The bias adjustment part of ISIMIP3BASD is a trend-preserving quantile mapping method. Parametric quantile mapping is applied to pr, sfcwind, tas, and tasrange. Non-parametric quantile mapping is applied to hurs, rsds, and tasskew. The bias adjustment is done at the spatial resolution of ERA5-Land, 6', using spatially aggregated EMO-1 data. Data resulting from the bias-adjustment are then statistically downscaled to 1' spatial resolution by using the statistical relationships between EMO-1 data at 1' and 6' (**Figure 3**). The downscaling method is conservative in the sense that the 1' output data would be identical to the 6' input data if we spatially aggregated the former back up to 6' resolution.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024



Finally, potential evapotranspiration (et_0), potential open-water evapotranspiration (e_0) and potential bare soil evapotranspiration (es_0) are computed with bias-adjusted and downscaled data at pixel level using an approach based on the Penman-Monteith equation with the LISVAP model (LISVAP online documentation, 2023).

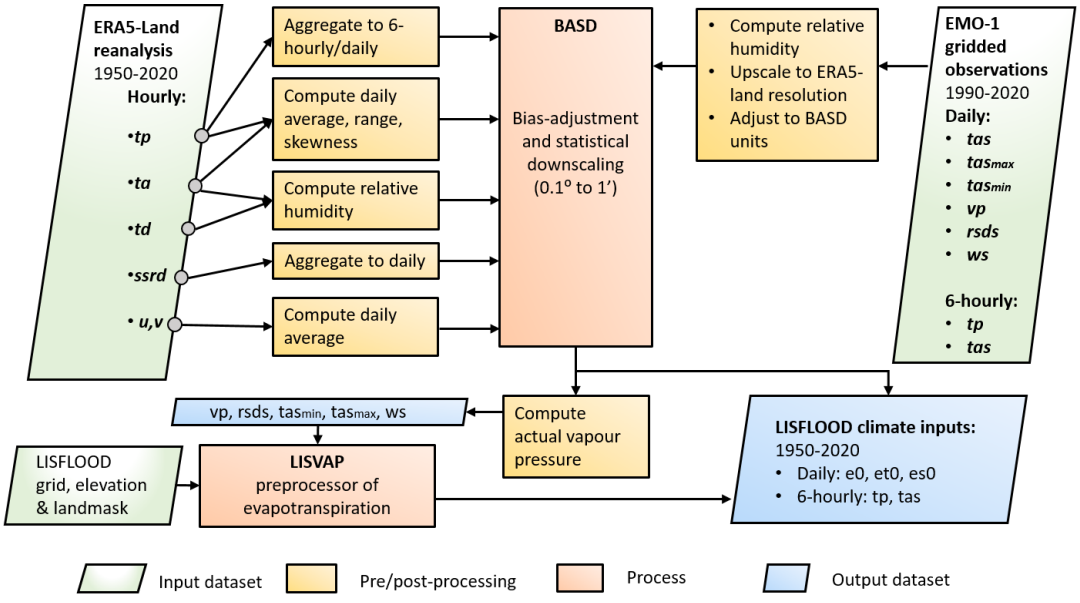


Figure 3: Climate inputs pre-processing scheme, including temporal aggregation, bias-adjustment, statistical downscaling and processing of evapotranspiration.

Surface field maps

To accurately represent hydrological processes, LISFLOOD requires a set of surface fields maps. Depending on the model set-up, LISFLOOD can ingest up to 108 surface fields divided in six categories:

(i) Catchment morphology and river networks
(ii) Vegetation cover types and properties
(iii) Soil properties
(iv) Land use
(v) Water demand
(vi) Lake and reservoir information

The first three categories, thereafter referred to as static maps, are directly taken from the CEMS_SurfaceFields_2022 open-source dataset of the Copernicus Emergency Management Service, developed for the European domain at 1 arc min resolution, which can be found in the JRC Data Catalogue (Choulga *et al.*, 2023b). The three last categories are derived from CEMS_SurfaceFields_2022 and modified to take into account socioeconomic changes (hereafter referred as dynamic socioeconomic maps).

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

This section briefly presents each of the previously enumerated map categories, with an emphasis on dynamic socioeconomic maps, which are original to this work.

Static maps

Static maps include surface fields of morphology and channel shapes (14 maps), vegetation properties (18 maps) and soil properties (29 maps).

Morphology and river network information are directly used for the computation of snow melting, temperature scaling, river routing and open water evapotranspiration. Morphologic information is derived from elevation and includes elevation gradient, standard deviation of elevation within grid, Manning's roughness coefficient. Maps representing channel shapes and river networks provide information on grid cell area (which vary with latitude as the grid projection is WGS84), local drainage direction, upstream area and channel dimensions. All morphology and river network maps are derived from the Multi-Error-Removed Improved-Terrain Digital Elevation Model v.1.0.3 (MERIT DEM) (Yamazaki *et al.*, 2019) and the Catchment-based Macro-scale Floodplain (CaMa-Flood) Global River Hydrodynamics Model v4.0 maps (Yamazaki, 2023).

Vegetation cover types and property maps are involved in the computation of precipitation interception, evaporation, transpiration, surface runoff and root water uptake. These properties are described though four variables: crop coefficients (transpiration), crop groups (water uptake), manning roughness (surface runoff) and leaf area index (interception and evaporation). Each of these variable are mapped for three different land cover types: forest, irrigated and other. Maps of planting and harvesting days for rice, which has specific water demands, are also available. Vegetation properties are derived from several data sources including the Copernicus Global Land Service (CGLS) Leaf Area Index (LAI) 1km (Copernicus, 2021), the Spatial Production Allocation Model (SPAM) – Global Spatially-Disaggregated Crop Production Statistics Data for 2010 (Yu *et al.*, 2020; International Food Policy Research Institute, 2019). and the Food and Agriculture Organisation (FAO) of the United Nations Irrigation and Drainage Paper No.56 (Allen *et al.*, 1998).

Soil properties refer to physical characteristics of the soil and aim to describe the water dynamics through a vertical soil profile. In LISFLOOD, the soil profile is composed of three layers: superficial (0 – 5cm) upper (5 – varying (30 – 50) cm) and lower soil layer. For each layer, variable representing soil hydraulic properties (e.g., soil moisture content, pore size index) are provided. Similarly to vegetation properties maps, variables are mapped for different land cover type, here forest and other. Soil properties are computed

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

from a global dataset, the International Soil Reference and Information Centre (ISRIC) SoilGrids250m global gridded soil 386 information release 2017 (Hengl *et al.*, 2014). based on more than 150 000 observation sites and covariate data.

A table summarizing all the static and dynamic surface fields maps used in this piece of work is provided in **Supplement Table S2**. For more details on the surface fields maps, their production and input datasets used ,we refer to (Choulga *et al.*, 2023a)

Dynamic land use

LISFLOOD includes six classes of land use as inputs: rice, other irrigated land, forest, sealed surfaces, open water, and other (non-irrigated agriculture, non-forest natural, pervious artificial), which are mostly based on CLC-Refined 2006 dataset by Batista e Silva *et al.*, (2013) in the default setting. Interception, evapotranspiration, infiltration, and surface runoff respond differently to each land use types. With the aim to better represent complex rainfall-runoff processes, LISFLOOD accounts for the sub-grid variability in land use. Therefore, the spatial distribution of each land use class is defined as a percentage of the whole represented area of a given pixel (LISFLOOD OS online documentation, 2023). We modify here the grid cell fractions of each land use class using HANZE-Exposure land use maps at 100 m resolution (Paprotny and Mengel, 2023) for 42 countries in the study area. In the remaining part of the domain, we use coarser, 5' resolution maps from HYDE 3.2 (Klein Goldewijk *et al.*, 2017) to modify the 2006 values. The temporal evolution of land area of each class is displayed in **Figure 5.a**. There has been a strong increase in sealed surfaces (+40%), while for the other relevant land use classes the changes are less than 10%, with increases in land occupied by irrigated agriculture (except rice), water surface (due to reservoir construction) and forests.

Dynamic water abstraction

Human water use, representing water withdrawal from the natural environment (e.g., rivers, reservoirs, groundwater) for human needs, is grouped in four main sectors: livestock, domestic, manufacturing industry, and energy production. Within LISFLOOD, water use is supplied by surface water bodies and groundwater depending on the sector (Choulga *et al.*, 2023a). To derive monthly historic sectoral water withdrawal maps, we followed the methodology of Huang *et al.*, (2018) and used by the Food and Agriculture Organization (FAO) AQUASTAT sectoral water withdrawal data (FAO, 2023) as a starting point. These data were subsequently spatially and temporally disaggregated using a variety of datasets. These include the Global Human Settlement Layer (Schiavina *et al.*, 2019; Florczyk *et al.*, 2019) for population estimates, the Global Change Analysis Model (GCAM; Calvin *et al.*, 2019) for regional water

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

withdrawal and electricity consumption, and the Gridded Livestock of the World (GLW; Gilbert *et al.*, 2018) for livestock distribution. Additional datasets included the Multi-Source Weather (MSWX; Beck *et al.*, 2022) for air temperature data, United States Geological Survey (USGS) water withdrawal estimates, and Vassolo and Döll (2005) industrial and thermoelectric withdrawal maps. More information on water demand and input datasets used is provided in Choulga *et al.* (2023a).

We extrapolated the water withdrawal maps to the period 1950-1978 using annual gridded 0.5 degree data from ISIMIP 3a (Frieler *et al.*, 2024; Wada *et al.*, 2016) that were downscaled to 1' resolution using historical population data from HANZE (Paprotny and Mengel, 2023) and HYDE 3.2 (Klein Goldewijk *et al.*, 2017) for other parts of the domain. Intra-annual (monthly) cycling of water use in the energy and domestic sectors was estimated for 1950–1978 using the same approach as for 1979–2020, informed by temperature data from our input meteorological dataset (**section 2.3.1**). Livestock water use is assumed constant before 1979. Water demand and use for irrigation is computed directly by the hydrological model based on land use data and available water.The evolution of water use by sectors between 1950 and 2020 is displayed in **Figure5.c**.

Dynamic reservoir maps

Maps on reservoirs contain location and identifier of reservoirs and are linked to tables containing information on storage capacity, construction year and a set of values associated to reservoir rules of operations in LISFLOOD. Normal reservoir outflow rates were also adjusted through the model calibration (**Section 2.1.2**). The year of construction for each reservoir was taken from the EFAS reservoirs database, HANZE (Paprotny and Mengel, 2023), Global Reservoir and Dam Database (GRanD) v1.3 (Lehner *et al.*, 2011), or additional manual research for reservoirs not covered by the three datasets. The reservoir maps are updated every year by adding newly built reservoirs. **Figure4.b** shows the evolution of the number of reservoirs in Europe during the period 1950 – 2020. The number of reservoirs in the model increased six-fold from 244 in 1950 to 1419 in 2020, though few were built since the late 1980s. Total water use peaked in 1990 after more than doubling since the 1950s, before declining due to a drop in demand from manufacturing and energy sectors. Nonetheless, there are usually much stronger trends at country or catchment levels.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

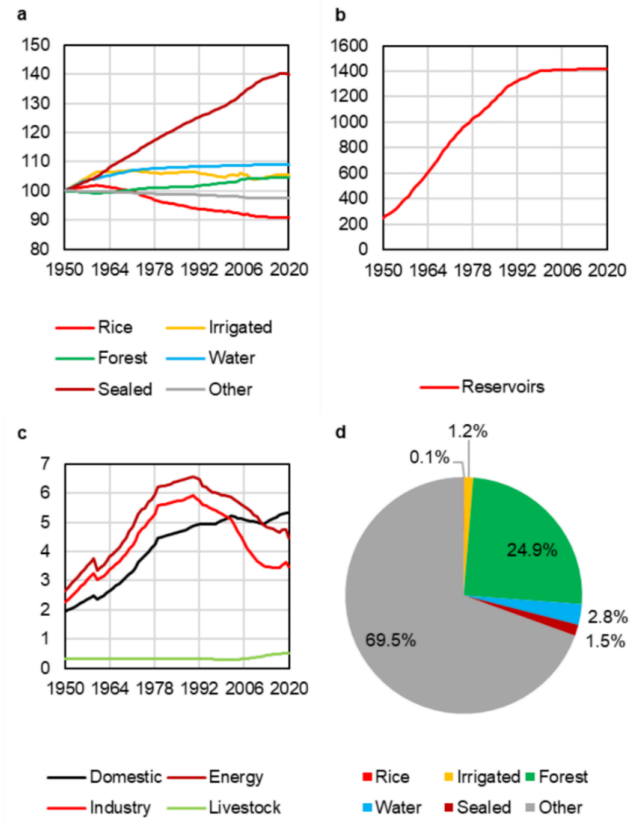


Figure 4: Variation in socioeconomic inputs in the hydrological model, average for the entire EFAS domain: (a) land area by use category, 1950=100, (b) number of existing reservoirs, (c) water demand by sector in mm per grid cell per year, (d) shares of land use between the different classes in 2020.

Results

Technical validation

To evaluate our hydrological reanalysis, we use daily river discharge observations from 3442 stations across Europe. The data were obtained for 60% from the Global Runoff Data Centre (GRDC) and for 40% from national public datasets of France, Norway, Poland, Spain, Sweden and the United Kingdom. Furthermore, this dataset was compiled independently from one used in the EFAS calibration (**Section 2.1.2**). The stations' record duration vary between 1 year and 71 years. The selection of stations used for validation is based on several criteria:

- Spatial matching: to link stations to its corresponding river pixel, we scan the nine modelled pixels around the location of the river gauges, retaining the one with the closest simulated mean discharge

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

(Q_{mean}) to the observed one. A total of 398 stations did not match with river pixels, mostly due to their upstream area being lower than 100km^2. For a more accurate spatial matching, we use, when available LISFLOOD coordinates from the EFAS calibration (1026 stations).

- Mean discharge comparison: For some stations, the ratio between observed and simulated Q_{mean} were suspicious. This could be due to an erroneous spatial match (i.e., matching of a river with a station on a tributary). As uncertainty grows with smaller streams, we decide to remove with a suspicious Q_{mean} ratio (r_{Qmean} >6 or r_{Qmean} >3 if $Q_{mean,obs}$>10 m^3/s) (103 stations)
- Based on model skill (KGE, See below), we identify stations with low skill, and remove the ones where the distance between the station and its corresponding pixel exceeds 2.5 km (25 stations). With this process, we may remove valid spatial matches, however, we put an emphasis on eliminating faulty matches, as these could bias the conclusion of the validation.
- Finally, a manual check was performed on 74 stations with KGE<-0.41, removing 24 more stations due to wrong spatial match, erroneous station location, and doubtful observations.

This procedure resulted in the selection of 2901 river stations across Europe, with an upstream area ranging from 100 to 785,421 km^2. Among these stations, more than half (1758) have an upstream area of less than 1000km^2 and a fifth (579) has an upstream area of less than 200km^2.

Performance at the daily scale is assessed using the modified Kling-Gupta efficiency metric (KGE',Gupta *et al.*, 2009; Kling *et al.*, 2012). KGE' is used as the standard performance metric in EFAS and GLOFAS (Harrigan *et al.*, 2020; Cammalleri *et al.*, 2020b), but also in other hydrological model assessments (Lin *et al.*, 2019; Harrigan *et al.*, 2020; Beck *et al.*, 2017) and it is composed of three components: correlation, bias errors, and variability errors:

$$KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (1)$$

$$\beta = \frac{\mu_s}{\mu_o} \quad (2)$$

$$\gamma = \frac{\sigma_s/\mu_s}{\sigma_o/\mu_o} \quad (3)$$

where r is the Pearson correlation coefficient between simulated (s) and observed (o) flow, β is the bias ratio, γ is the variability ratio, μ the mean discharge, and σ the discharge standard deviation. KGE' and its three components are dimensionless with an optimal value on 1. It is important to note here that KGE' values should not be interpreted like the more traditional Nash-Sutcliff efficiency (NSE, Nash and Sutcliffe, 1970). Indeed, for KGE' the mean flow benchmark has a value of $KGE' = 1 - \sqrt{2} = -0.41$. Any value above -0.41 therefore exceeds the benchmark (Knoben *et al.*, 2019). In **Section 3.1.1**, we assess model

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

390 performance across space, time (1950-2020) and catchment size, in order to identify strengths and weaknesses of HERA.

Despite covering many aspects of the performance of hydrological models, KGE' mainly focuses on mean values. As this dataset also aims to be used for long term analysis of hydrological extremes (high and low
395 flows), we also look at how well extreme high and low quantiles are reproduced, and we compare these performances with the one of median flows. We also assess how well the timing and seasonality on annual maxima and minima are reproduced in HERA.

Hydrological performance

400 We quantify here the overall performance of HERA in terms of KGE' as well as the decomposition of this indicator into its three components: correlation, bias and variability. **Figure 5** displays the distribution of KGE' and its three components across the 2901 stations retained. Among these stations, 2811 (97%) have a KGE'>-0.41, meaning the reanalysis is skilful for these stations (**Figure 5.a**). The median KGE' across all catchments is 0.54 while the mean is 0.42, although this value varies widely across catchments (**Figure**
405 **5.a**, **Figure 6.a**). The mean correlation value is relatively high (0.69) with 89% of the stations having r>0.5 (**Figure 5.b**). From **Figure 5.c** and **Figure 5.d**, we can observe that there is a tendency to slightly underestimate flows (mean bias ratio = -8.3%) and flow variability (mean variability ratio= -14.7%). The bias ranges between 0.8 – 1.2 (0.5 – 1.5) in 51% (89%) of the river gauges, which is considered as very good for hydrological reanalysis (Harrigan *et al.*, 2020; Alfieri *et al.*, 2020; Lin *et al.*, 2019; Yang *et al.*,
410 2021).

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

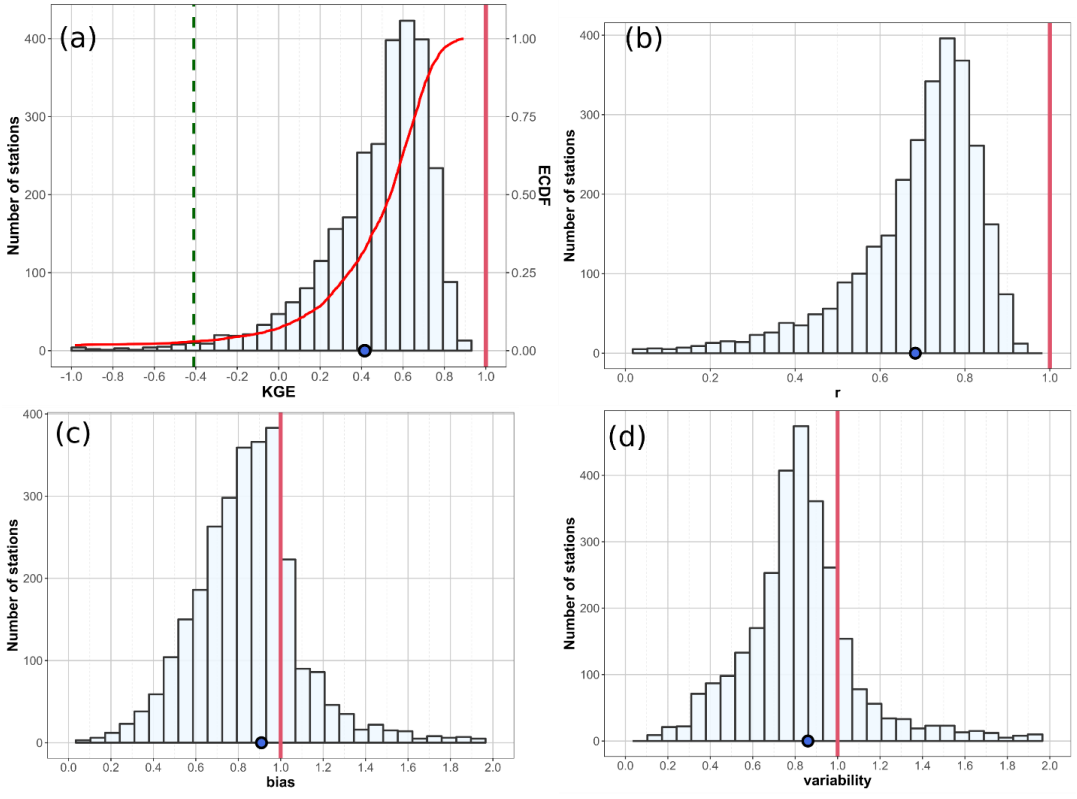


Figure 5: HER hydrological skill for the 2901 selected stations in terms of (a) KGE' and its three components: (b) pearson (r) correlation, (c), bias ratio, (d) variability ratio. The red vertical line represents the ideal value and the blue dot represent the mean for all stations.

Figure 6 shows the spatial performance of the model in terms on KGE' and its components. The highest skill can be observed in central and north-western Europe. The vast majority of stations in UK, Germany, France, Austria, Switzerland (which together account for 46% of all 2901 stations) exhibit a good (>0.5) to very good (>0.75) KGE'. On the other hand, performance is relatively poor in Spain, Cyprus, Southern Italy, Scandinavia and the Baltic countries. Factors that can explain the poor performances in southern Europe include the strong influence of lakes and reservoirs (**Figure 7.c**) and their deviation from the temperate climate of central Europe, where best performances are found. In northern Europe and mountainous areas, a lower performance can be attributed to the influence of snowmelt on river flow. Worst performing catchments are mainly driven by negative 1.2es (**Figure 6.c**) in Spain and Scandinavia. **Figure 6.d** highlights the variability ratio of simulated over observed flow. Overall, our reanalysis exhibits lower variability than observations, with 82% of the catchments having a variability ration lower than 1.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

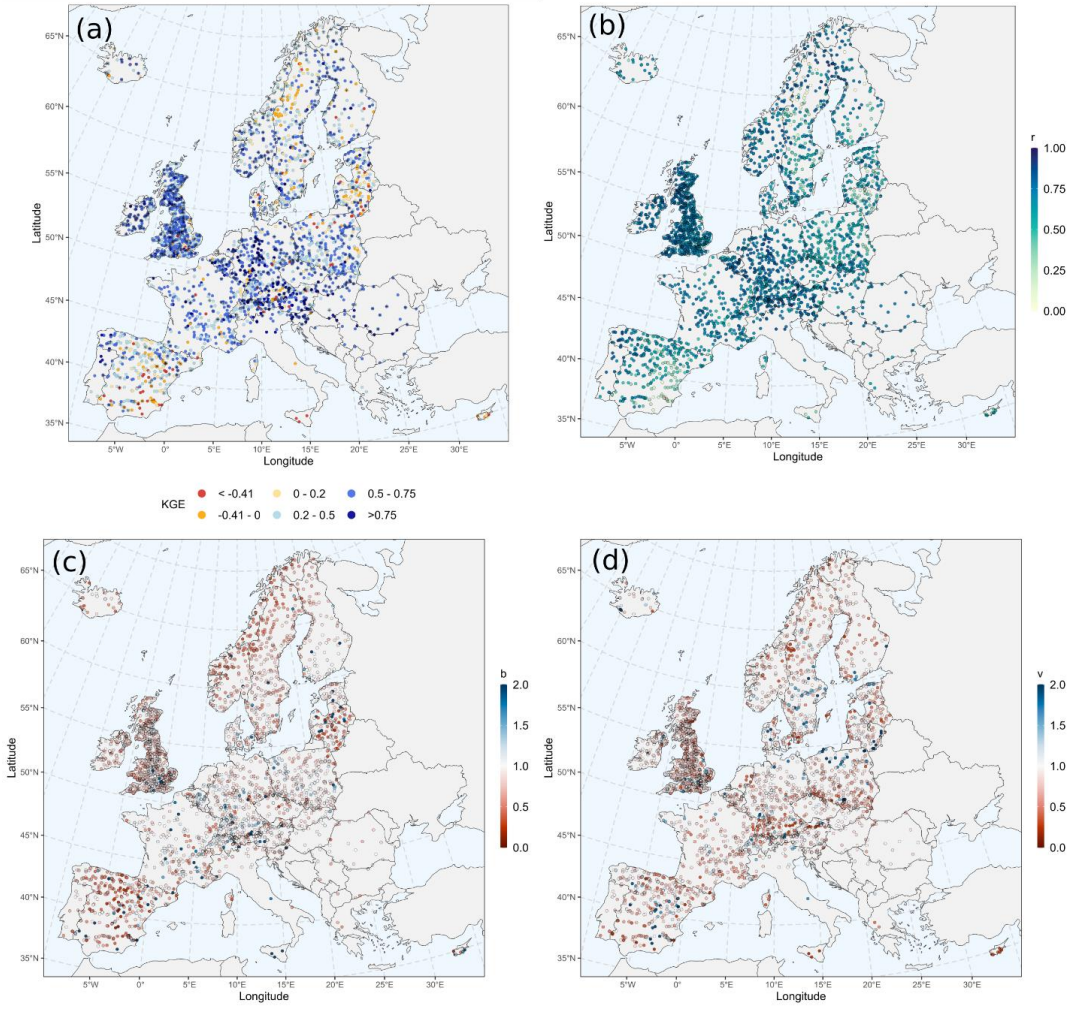


Figure 6: Maps of spatial skill of HER for the 2901 selected stations in terms of (a) KGE' and its three components: (b) pearson (r) correlation, (c), bias ratio, (d) variability ratio.

KGE' values and its three components are in line with previous LISFLOOD applications at global (Alfieri *et al.*, 2020; Hirpa *et al.*, 2018) and European level (Zajac *et al.*, 2013; CEMS-Flood online documentation, 2023). However, the improved resolution of the present reanalysis allows comparison with observations in smaller catchments compared to global studies (mean upstream area = 7615 km^2), where model performance is typically lower (Harrigan *et al.*, 2020). **Figure 7** breaks down the performance of the reanalysis according to decades (**Figure 7.a**), catchment area (**Figure 7.b**) and reservoir impact (**Figure 7.c**).

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

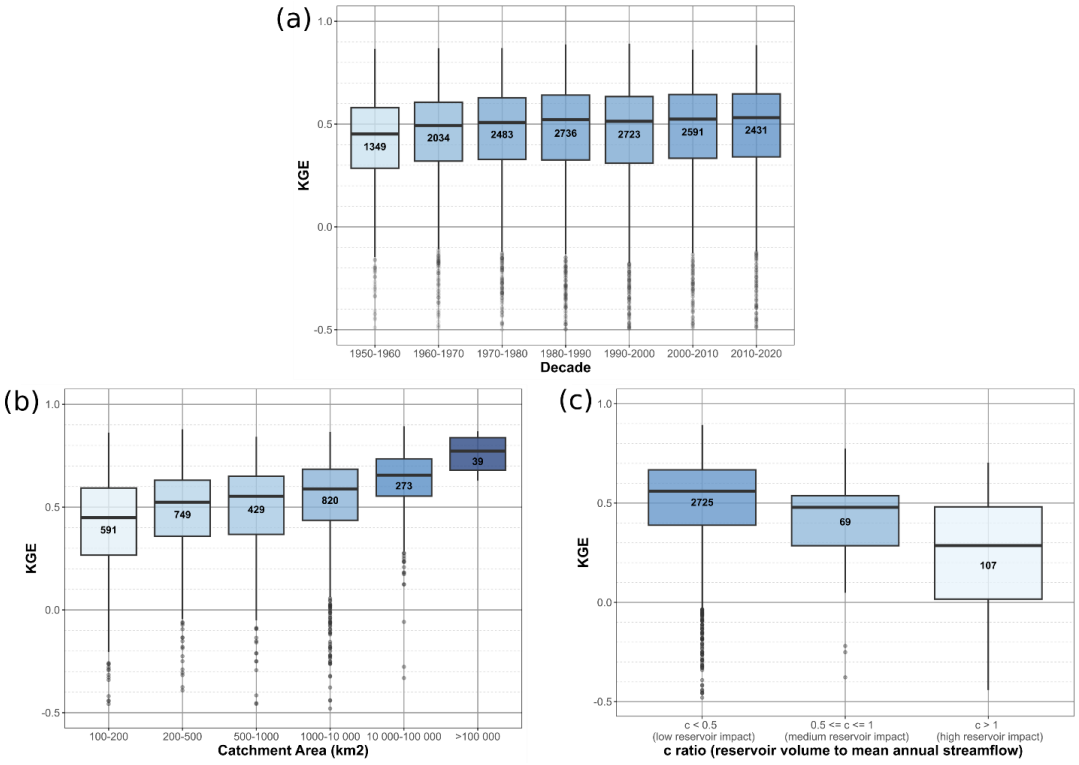


Figure 7: Boxplot of HER KGE' according to different classifications of the 2901 river stations used in the validation, (a) time, (b) catchment area and (c) reservoir impacts. Numbers inside boxplot represent the amount of river gauges for each category, while the colour of the boxplot represent the median performance of the group from low (light blue) to high (dark blue).

Overall, the skill of HERA shows a slight increase through time, from a median KGE' value of 0.44 (IQR 0.25 – 0.57) during the 1950s to 0.52 (IQR 0.30 – 0.64) in the 2010s. Skill increases between 1950 and 1980 and then stabilizes from 1980 to 2020, though the results are influenced by changes in gauge data availability over time. It also could be driven by improved climate inputs. **Figure 7.b** unravel unsurprising yet interesting patterns. Model skill increases with catchment size, from a median KGE' of 0.44 (IQR 0.25 – 0.59) for very small catchments to 0.77 (IQR 0.68 – 0.84) for the 39 catchments with an upstream area greater than 100,000km^2. Such patterns have already been observed at global scales (Harrigan *et al.*, 2020). It is important to note that a majority of stations used in this validation (61 %) have an upstream area below 1000km^2, the median upstream area of the 2901 stations is 597 km^2, which is half of the median upstream area of the 1903 stations used in the calibration of EFAS-5 (CEMS-Flood online documentation, 2023).

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

Finally, we divide stations according to reservoir impacts. From the 1420 reservoirs active in 2020 (which represent the maximum amount over the considered time window), we estimate the impact of reservoirs on streamflow at pixel level. This is done by computing the ratio (c [-]) of reservoir volume to mean discharge proposed by (Nilsson *et al.*, 2005) at every pixel. The ratio has been computed with the accuflux function from PCRaster and compares the upstream cumulative reservoir capacity [m^3] and the cell-specific annual volume of annual streamflow [m^3] (Zajac *et al.*, 2017). This ratio varies between 0 and 1608 downstream of embalse Finisterre in central Spain. Most of the river pixels highly impacted by reservoirs are found in southern Europe, more in particular in Spain and Bulgaria. **Figure 7.c** highlights the influence of reservoirs on the skill of the reanalysis. River cells affected (medium and high, c >0.5) only represent 6% of stations and river pixels in the domain (**Figure 1**). Median skill is the lowest for highly impacted (c > 1) stations, with a median KGE' of 0.24 where low impacted stations have a median KGE' of 0.55. This highlights the difficulty of large-scale hydrological models such as LISFLOOD to accurately simulate reservoir outflows (Zajac *et al.*, 2017).

Reproduction of extremes

HERA can be used to assess changes in several aspects of European river flows, including average discharges, flow regimes, seasonality (**Section 3.1.2**). The efforts made in refining both spatial and temporal scale (high-resolution, bias corrected) also aim to make HERA suitable for the analysis of extremes. It is notorious that large scale hydrological models forced by climate reanalysis often fail to reproduce extreme hydrological event characteristics, for example flood magnitudes tend to be typically underestimated (Brunner *et al.*, 2021b; McClean *et al.*, 2023). We analyse here how well HERA reproduces different daily flow quantiles (q05, median, q95) through the Person correlation coefficient (*r*) and the coefficient of determination (R^2) (**Figure 8**). The ability of the reanalysis to capture annual maxima/minima and their seasonality is also assessed (**Figure 9**).

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

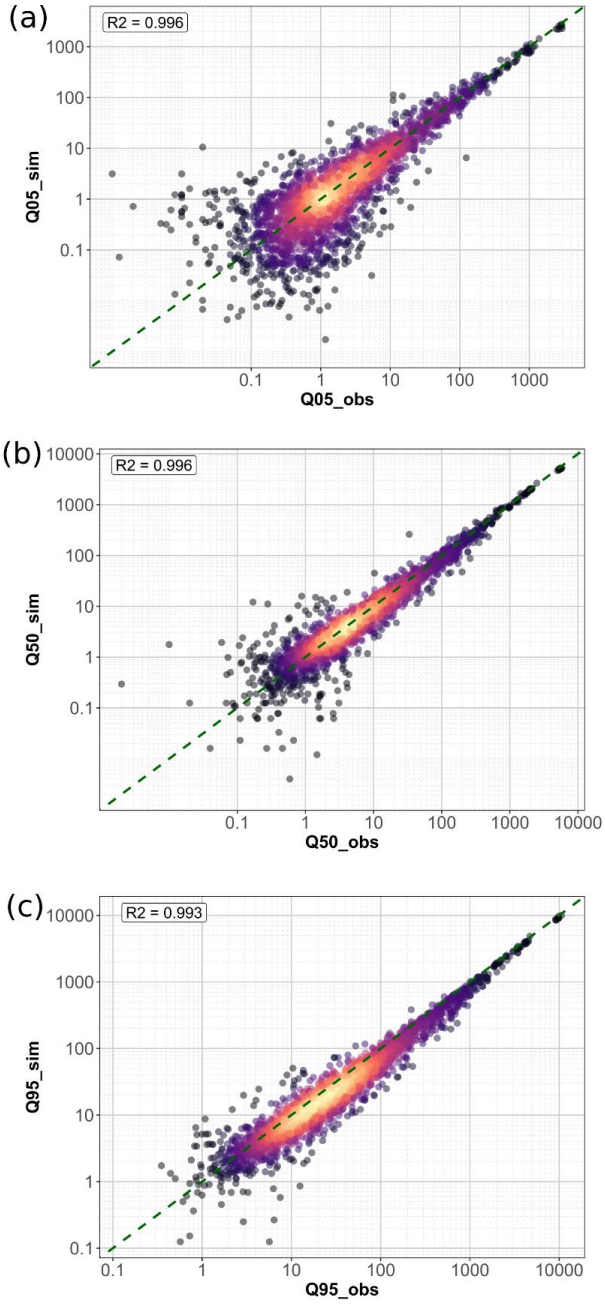


Figure 8: Scatterplot of observed versus simulated daily river flow quantiles [$m^3.s^{-1}$]: (a) 5% quantile, (b) median (q50), (c) 95% quantile (q95) for the 2901 River gauges.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

Figure 8 displays scatter plots of observed versus simulated quantiles. Each point represents one of the 2901 stations. From **Figure 8**, we observe that low (5% quantile: Q_{05}) and median (Q_{50}) flows, are generally well represented with a $R^2 = 0.996$ in both cases (**Figure 8.a** and **Figure 8.b**), especially for larger discharge values. However, despite this general good agreement, there is a more pronounced deviation of simulated values from observations for lower flow values, expressed by a higher dispersion for Q_{05}. These deviations can be attributed to errors or biases in our climate inputs (McClean *et al.*, 2023), in the hydrological model (Feyen and Dankers, 2009), but also to errors in flow measurements, especially for Q_{05} (Despax, 2016; Tomkins, 2014) and anthropogenic impacts on low and median flow regimes (Brunner, 2021) that are not accurately represented in the model (see **Figure 7.c**). The number of stations with large deviations in the reproduction of high flow statistics (Q_{95}) is minor compared to Q_{05} and Q_{50}. Nonetheless, despite a relatively high R^2 (0.993), there is a general underestimation bias in the simulated values (**Figure 8.c**). As mentioned above, such underestimation biases are common for large scale hydrological models. Similarly to low and median flows, errors in low flow statistics can be due to biases and smoothing of extremes in our climate inputs and errors in hydrological modelling. Furthermore, it has been shown that using KGE' for calibration purposes can result in an underestimation of peak flows (Brunner *et al.*, 2021b). Uncertainty associated to flow measurements also play a major role for high flows, as rivers discharge are usually not directly measured during floods (Despax, 2016). Finally, both spatial and temporal resolutions of our model can affect its ability to reproduce high flows, particularly for flash floods in small catchments.

We also assessed the ability of the reanalysis in the timing of annual maxima and minima events as well as their overall seasonality. **Figure 9** displays the errors in mean day of occurrence computed with circular statistics following the method explained in (Berghuijs *et al.*, 2019). We observe that the median error in the mean day is very close to zero for both maxima (IQR=-11 – 11) and minima (IQR = -25 – 25), but with more dispersion for drought compared to floods. This can be linked to different drivers and characteristics (Brunner *et al.*, 2021), with droughts being slow onset events lasting longer than flood events (Van Loon, 2015). Despite some outliers (mostly located in north-eastern and south-eastern Europe, not displayed here), seasonal errors are reasonable for both flood and drought events.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

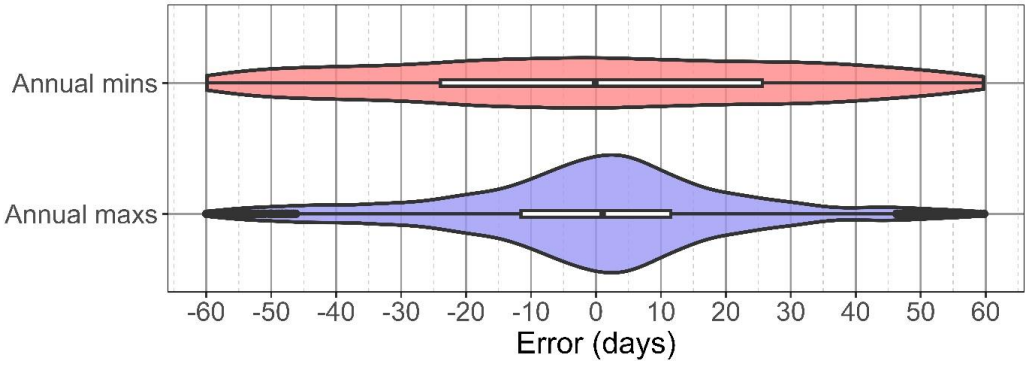


Figure 9: Violin plot of error in mean day of occurrence of annual maxima/minima computed with circular statistics. Inside each violin plots, boxplots display the median, 1st and 3rd quantiles.

Usage notes

The HERA datasets brings together developments from its inputs to the hydrological modelling that have been described in **Section 2** and quantified in **Section 3.1**. Despite still being a relatively short period compared to human history on earth, the 70 years of homogeneous modelled river flow provided by HERA captures the most intense period of climate and socioeconomic change, often called the Anthropocene, and offers multiple research opportunities:

- Assessment of long-term trends in European river regimes
- Provision of benchmark data for "data poor" areas
- Generation of catalogues of flood and drought events
- Identification of spatial and temporal correlations between European catchments
- Identification of changes in hydrological extremes characteristics (frequency, magnitude, timing)
- Combination with other data products for compound hazard analysis
- Provision of scenarios for flood inundation simulations

In this section, we briefly present a simple usage of the data, addressing changes in regime for diverse rivers across Europe (**Figure 10**).

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

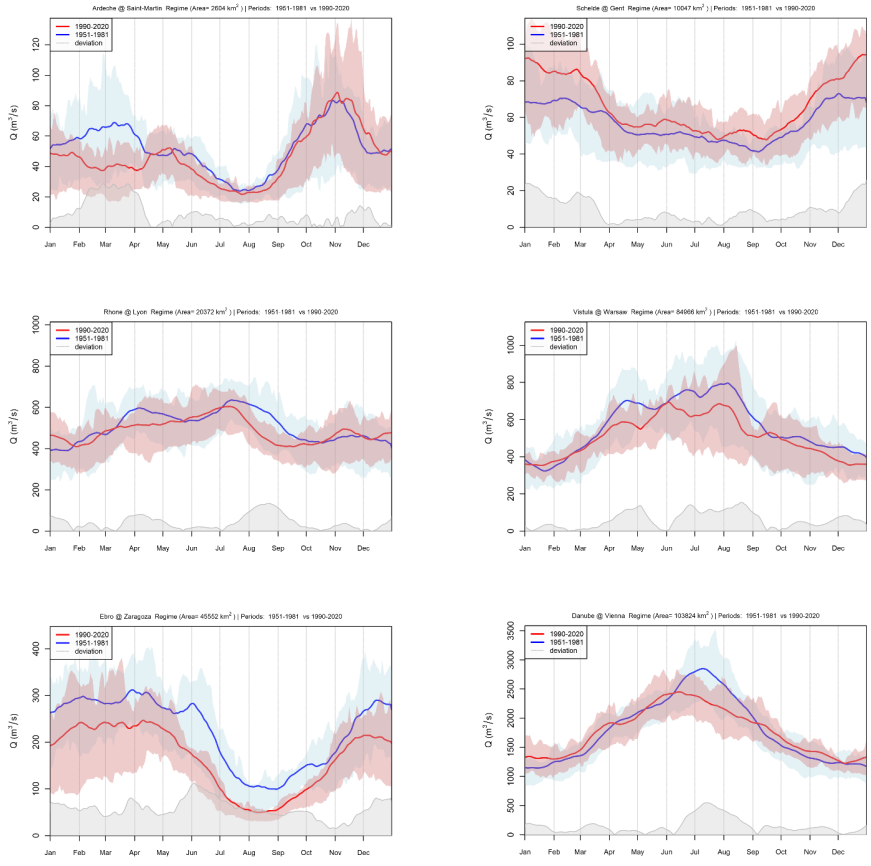


Figure 10: Changes in hydrological regimes between the first (1951-1981) and the last (1990-2020) 30-years periods for diverse European rivers

Figure 10 displays hydrological regimes, here represented as the mean of a 30-days' average moving window over a given period, for six European rivers. These rivers differs in terms of hydrological regimes, with three main regimes represented:

- Mediterranean pluvial regime for the Ardèche, with its recognisable high flows in autumn.
- Pluvial or oceanic regime for the Schelde in Ghent and the Ebro in Zaragoza
- Nival regime for the upper Rhone in Lyon, the Danube in Vienna and the Vistula in Warsaw.

These six rivers also vary in terms catchment area, geographic location (France, Austria, Poland, Belgium, Spain) climates present in their upstream areas (Mediterranean, Continental, Oceanic, Alpine) and geomorphic conditions. In **Figure 10**, two regimes are displayed for each river, in blue for the period 1951 – 1981 (first 30 years without 1950, which is impacted by the spin-up of LISFLOOD (LISFLOOD OS online documentation, 2023), and in red for the second for the last 30 years of the simulation, 1980 – 2020.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

By comparing the two regimes, one can observe diverging patterns of changes among these rivers. For instance, the Schelde and the Ebro, both pluvial rivers see opposite patterns of change, the Schelde sees an increase of its average discharge throughout the year, while the Ebro has a constant water deficit. For the upper Rhone and Danube, which are influenced by snowmelt in their upper catchments, we see lower and earlier flow peaks in spring and summer. The Vistula sees an overall increase in flow throughout the year. Finally, the Ardèche has reduced flow throughout the year, with a notable decrease in winter. The timing of the autumn peak seems to have slightly shifted towards earlier dates, a feature which was also unravelled in a recent study on trends in Mediterranean floods (Tramblay *et al.*, 2023).

Previous studies, focusing on different metrics and periods already highlighted comparable diverging patterns between different European regions (Peña-Angulo *et al.*, 2022; Stahl *et al.*, 2012; Blöschl *et al.*, 2019a).

Discussion

Recent developments in diverse scientific fields including climate modelling, hydrological modelling and remote sensing made the generation of high resolution reanalysis products possible. As a result, the HERA dataset brings discharge data for all European rivers with upstream area larger than 100 km^2. With its refined spatial and temporal resolution, HERA captures dynamics in extremes events (floods and drought), reveals spatial heterogeneities and simulates anthropogenic water use with more accuracy than any other hydrological reanalysis products (Harrigan *et al.*, 2020; Schellekens *et al.*, 2017). Calibrating hydrological models can significantly improve river flow simulation (Beck *et al.*, 2017; Kauffeldt *et al.*, 2016). Around 93.5% of the HERA domain received specific calibration parameters through the calibration process (**Section 2.1.2**) or parameter regionalization (Beck *et al.*, 2016). This is a very high calibration coverage for a GHM, which are not systematically calibrated (Beck *et al.*, 2017), that can be explained by the relatively high coverage in river gauging stations in Europe.

It is difficult to compare HERA with two recent hydrological reanalysis: GLOFAS-ERA5 (Harrigan *et al.*, 2020) and GRFR (Yang *et al.*, 2021) for several reasons: the two aforementioned dataset are global, have larger spatial resolution (respectively 0.25° and 0.05°), shorter temporal coverage and do not account for anthropogenic land and water use changes. Conversely, HERA shares a great amount of features with the EFAS v5.0 reanalysis (Decremer *et al.*, 2023), with comparable performances. Nonetheless, EFAS v5.0 only covers the period 1990 – 2022 and assumes static socioeconomic conditions (land use, water abstraction, reservoirs). Similarly to other aforementioned hydrological reanalysis, HERA exhibits reduced

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

performance in cold and semi-arid catchments. This can be related to model errors in modelling snow processes or the underestimation of precipitation in northern latitudes (Beck *et al.*, 2017). Semi-arid environments are notoriously challenging areas for hydrological models due to their strong interannual variability in precipitation and runoff (Cantoni *et al.*, 2022). GHMs, including LISFLOOD, tend to poorly represent runoff in small-to-medium size catchments (10-10 000km^2) (Harrigan *et al.*, 2020; Sood and Smakhtin, 2015), and nearly 90% of the catchments used in the validation of HERA (**Section 3.1**) are small-to-medium size catchments. The drop in performance with smaller catchment area in HERA remains, however, moderate compared to the GLOFAS-ERA5 global hydrological reanalysis (Harrigan *et al.*, 2020). The presence of reservoirs also has an impact on reanalysis performances. While including reservoirs in the hydrological modelling has a positive impact on LISFLOOD performances (Zajac *et al.*, 2017), high uncertainty remains on the operating rules specific to each reservoir or operators. Moreover, the 1422 reservoirs used to generate HERA most likely only represent a fraction (largest ones) of all reservoirs currently in operation in the modelled domain (Speckhann *et al.*, 2021).

HERA is generated through hydrological modelling, which brings a suite of uncertainties. Uncertainties in hydrological modelling can be divided into four categories: (i) model inputs, (ii) model structure, (iii) parameter values and (iv) observed data. It remains challenging to quantify these uncertainties, however, the quality of inputs, and more in particular climate inputs is often referred as an important factor of uncertainty (Beck *et al.*, 2017; Sood and Smakhtin, 2015). The improvement of overall modelling performances through time highlighted in **Figure 7.a** could therefore be related to improving climate inputs. Indeed, the quality of climate reanalysis are influenced by the availability of observations, which in ERA5-land become more sparse and inhomogeneous as we go back further in time (Hersbach *et al.*, 2020; Muñoz-Sabater *et al.*, 2021). Despite efforts in bias correcting and downscaling our climate inputs, it seems that on average, HERA slightly underestimates river discharges, with a more pronounced bias for high flows. These biases can be related to an underestimation of precipitation in our climate inputs, in particular for extreme events (McClean *et al.*, 2023; Mahto and Mishra, 2019).It has also been shown that uncertainties in precipitations can result in even higher uncertainties in runoff in (semi-)arid catchments (Beck *et al.*, 2016; Sood and Smakhtin, 2015). Uncertainties inherent to model structures have long been overshadowed (Beck *et al.*, 2017). Nonetheless, the large impact of model selection on streamflow and trend estimates is now increasingly acknowledged (Karlsson *et al.*, 2016; Clark *et al.*, 2016). Other uncertainties car arise from surface field maps (**Section 2.3**) and calibration parameters, even if these latter are reduced by the model calibration procedure. Ultimately, uncertainties remain even in the data we consider as “ground truth”: flow measurements. These uncertainties are related to the way river discharge is measured (instruments and

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

605 rating curves). And with sparser gauging and more complex hydraulic conditions for high and low flows, uncertainty rises (Despax, 2016).

Data availability

We make available the HERA hydrological reanalysis along with it climate and dynamic socioeconomic
610 inputs through the JRC data catalogue at http://data.europa.eu/89h/a605a675-9444-4017-8b34-d66be5b18c95 (Tilloy *et al.*, 2024). **Table 1** provides a brief description of the dataset and **Table 2** gives a general overview of the content of the dataset.

Table 1: Description of the HERA dataset

DATASET DESCRIPTION	
Data type	Gridded
Projection	WGS 1984 – EPSG 4326
Spatial coverage	EU27, UK, Switzerland, Iceland, Norway, Serbia, Montenegro, Bosnia-Herzegovina, Kosovo, North Macedonia, Albania
Temporal coverage	03-01-1950 to 31-12-2020
Temporal resolution	Six-hourly data
File format	netcdf

We want to stress here that even if the hydrological reanalysis starts on January 3rd 1950, discharge values
615 in the first months may be inaccurate due the lower zone's spin up (Burek *et al.*, 2013), especially in dry catchments. We therefore do not recommend to use data for the beginning of 1950.

The dataset consists three distinct folders that are described here and in **Table 2**:

- Climate inputs: folder containing the climate forcings use to run the LISFLOOD hydrological
620 model. Out of the five variables provided, three are at daily temporal resolution, potential evapotranspiration, potential evaporation and potential evaporation from bare soil (obtain with LISVAP, LISVAP online documentation, 2023), while two have a six-hourly time step, precipitation and temperature. The spatial resolution of the climate inputs is 1'. The files are in netcdf format with one file per year per variable for a total of 355 files (2.3 TB of data).
625 - Socioeconomic inputs: folder containing the dynamic surface fields maps (**Section 2.3**), divided into three categories: land use, reservoirs and water demand. The land use subfolder contains 426 yearly files (4.6 GB) of land use fraction maps for each six land use classes. The reservoir subfolder hosts 71 yearly files (3.6 GB) of reservoir location and identifier, reservoirs are added/discarded

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

from the simulation every year according to their construction/destruction data. Finally, the water demand subfolders contains four files (3.9 GB) representing water demand for the considered sectors (**Section 2.3.3**). Each file contains monthly maps of water abstraction for a given sector. All socioeconomic inputs are provided in the netcdf format.

- River discharge: this folder contains the flagship of the present dataset, yearly river discharge netcdf files at six-hourly time step for all European river. The initial output of LISFLOOD was filtered to retain solely rivers with an upstream area greater than 100 km^2. This results in continuous streamflow estimates for 282 521 river pixels and significantly reduces data size (2.3 GB per file, 166 GB total).

All data from this dataset share the same projection (WGS 84) grid and spatial resolution (1'). Static surface fields maps were directly retrieved from the LISFLOOD static and parameter maps for Europe dataset (Choulga *et al.*, 2023b), which were developed in the context of the new EFAS deployment (Decremer *et al.*, 2023). It is important to note that HERA simulates discharge on a slightly smaller domain than the original EFAS domain, the mask used for HERA is also provided in the dataset.

Table 2: List of inputs and outputs of LISFLOOD provided in the HERA database (link here).

Subfolder	File	Resolutions	Variable/content	Unit
	area_hera_01min.nc	1'	*mask of the hera domain*	
climate_inputs/ e0	e0_yyyy.nc	1', daily	*potential evaporation computed with lisvap from downscaled and bias-corrected actual vapour pressure, solar radiations, min/max daily temperature and 10m wind speed.com*	mm.d^{-1}
climate_inputs/ et0	et0_yyyy.nc	1', daily	*potential evapotranspiration computed with lisvap from downscaled and bias-corrected actual vapour pressure, solar radiations, min/max daily temperature and 10m wind speed.com*	mm.d^{-1}
climate_inputs/ es0	es_yyyy.nc	1', daily	*potential evaporation from bare soil computed with lisvap from downscaled and bias-corrected actual vapour pressure, solar radiations, min/max daily temperature and 10m wind speed.*	mm.d^{-1}
climate_inputs/ pr6	pr6_yyyy.nc	1', six-hourly	*downscaled and bias-corrected six-hourly precipitation*	mm.d^{-1}

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

climate_inputs/tp6	ta6_yyyy.nc	1', six-hourly	*downscaled and bias-corrected six-hourly average temperature*	°c
socioeconomic_maps/landuse	fracforest_european_01min_yyyy.nc	1', yearly	*fraction of pixel area covered by evergreen and deciduous needle leaf and broad leaf tree areas*	
socioeconomic_maps/landuse	fracsealed_european_01min_yyyy.nc	1', yearly	*fraction of pixel area covered by urban areas, characterizing the human impact on the environment*	
socioeconomic_maps/landuse	fracirrigated_european_01min_yyyy.nc	1', yearly	*fraction of pixel area covered by irrigated areas of all possible crops excluding rice*	
socioeconomic_maps/landuse	fracwater_european_01min_yyyy.nc	1', yearly	*fraction of pixel area covered by rivers, freshwater and saline lakes, ponds and other permanent water bodies over the continents*	
socioeconomic_maps/landuse	fracrice_european_01min_yyyy.nc	1', yearly	*fraction of pixel area covered by irrigated areas of rice*	
socioeconomic_maps/landuse	fracother_european_01min_yyyy	1', yearly	*fraction of pixel area covered by agricultural areas, non-forested natural area, pervious surface of urban areas*	
socioeconomic_maps/reservoirs	res_european_01min_yyyy.nc	1', yearly	*location and identifier of each reservoir*	
socioeconomic_maps/water_demand	dom_1950_2020.nc	1', monthly	*daily supply of water volume for indoor and outdoor household purposes and for all the uses that are connected to the municipal system (e.g., water used by shops, schools, and public buildings)*	$mm.d^{-1}$
socioeconomic_maps/water_demand	ene_1950_2020.nc	1', monthly	*daily supply of water volume for fabricating, processing, washing and sanitation, cooling or transporting a product, incorporating water into a product*	$mm.d^{-1}$
socioeconomic_maps/water_demand	ind_1950_2020.nc	1', monthly	*daily supply of water volume for the cooling of thermoelectric and nuclear power plant*	$mm.d^{-1}$
socioeconomic_maps/water_demand	liv_1950_2020.nc	1', monthly	*daily supply of water volume for domestic animal need*	$mm.d^{-1}$
river_discharge	dis.herayyyy.nc	1', six-hourly	*river discharge for river pixels with upstream area>100km^2.*	$m^3.s^{-1}$

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

645

Conclusion

Despite the limitations discussed above, HERA successfully delivers a state-of-the-art, high-resolution, long-term hydrological reanalysis for Europe in the form of homogeneous river flow data generated with the LISFLOOD model. While covering a much longer period than the EFAS v5.0 reanalysis, HERA benefits from the development and calibration associated with this latter, which represent steps forward compared to previous EFAS hydrological reanalysis. To our knowledge, no other hydrological reanalysis currently provides discharge data for Europe at similar scales with a similar temporal coverage. Due to its extensive period, the datasets is particularly suited for the assessment of long-term trends of several hydrological signatures. The unprecedented spatial resolution allows for a robust analysis of small-to-medium catchments at continental scale. One of the main objectives of HERA is to advance the reproduction of extreme events, notably by the mean of the sub-daily temporal resolution and bias corrected climate input. On average, both magnitude and seasonality of extremes are well simulated, even if more work is needed to fully assess HERA performances in reproducing extremes, in particular for high flows. The inclusion of dynamic socioeconomic conditions further provides a more realistic reanalysis of river flows in European catchments. The modelling framework developed here further forms a basis for creating alternative (counterfactual) time series of river discharges where climatic or socioeconomic conditions can be kept static rather than dynamic, enabling attribution of changes in hydrological regimes across Europe (Kreibich *et al.*, 2019; Sauer *et al.*, 2021; Scussolini *et al.*, 2023).

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

Code availability

The lisflood model is open source and available on github along with a set of auxiliary tools (https://github.com/ec-jrc/lisflood-code). A sample of the settings file used to generate HERA discharge data with lisflood is provided along with the Python and R codes used to validate the simulation and generate the figure displayed in this manuscript are available on github (https://github.com/Alowis/HERA)

Supplement

Author contribution

Aloïs Tilloy: conceptualization, data curation, formal analysis, software, writing – original draft preparation. **Dominik Paprotny**: conceptualization, methodology, formal analysis, writing – original draft preparation. **Stefania Grimaldi**: methodology, software, supervision. **Goncalo Gomes**: software. **Alessandra Bianchi**: visualization. **Stefan Lange**: conceptualization, methodology, writing – reviewing and editing. **Hylke Beck**: conceptualization, methodology, writing – reviewing and editing. **Luc Feyen**: conceptualization, methodology, supervision, writing – reviewing and editing.

Competing interests

The authors declare that they have no conflict of interest.

Financial support

Dominik Paprotny was supported by the German Research Foundation (DFG) through project "Decomposition of flood losses by environmental and economic drivers" (FloodDrivers), grant no. 449175973.

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.

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: FAO Irrigation and Drainage Paper, https://www.researchgate.net/publication/284300773_FAO_Irrigation_and_drainage_paper_No_56, last access: 11 January 2024, 1998.

Batista e Silva, F., Lavalle, C., and Koomen, E.: A procedure to obtain a refined European land use/cover map, Journal of Land Use Science, 8, 255–283, https://doi.org/10.1080/1747423X.2012.667450, 2013.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

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.

Beck, H. E., van Dijk, A. I. J. M., de Roo, A., Dutra, E., Fink, G., Orth, R., and Schellekens, J.: Global evaluation of runoff from 10 state-of-the-art hydrological models, Hydrology and Earth System Sciences, 21, 2881–2903, https://doi.org/10.5194/hess-21-2881-2017, 2017.

Beck, H. E., Dijk, A. I. J. M. van, Larraondo, P. R., McVicar, T. R., Pan, M., Dutra, E., and Miralles, D. G.: 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, E710–E732, https://doi.org/10.1175/BAMS-D-21-0145.1, 2022.

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., and Kirchner, J. W.: The Relative Importance of Different Flood-Generating Mechanisms Across Europe, Water Resources Research, 55, 4582–4593, https://doi.org/10.1029/2019WR024841, 2019.

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., Lun, D., Arheimer, B., Aronica, G. T., Bilibashi, A., Boháč, M., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Frolova, N., Ganora, D., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T. R., Kohnová, S., Koskela, J. J., Ledvinka, O., Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Salinas, J. L., Sauquet, E., Šraj, M., Szolgay, J., Volpi, E., Wilson, D., Zaimi, K., and Živković, N.: Changing climate both increases and decreases European river floods, Nature, 573, 108–111, https://doi.org/10.1038/s41586-019-1495-6, 2019a.

Blöschl, G., Bierkens, M. F. P., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J. W., McDonnell, J. J., Savenije, H. H. G., Sivapalan, M., Stumpp, C., Toth, E., Volpi, E., Carr, G., Lupton, C., Salinas, J., Széles, B., Viglione, A., Aksoy, H., Allen, S. T., Amin, A., Andréassian, V., Arheimer, B., Aryal, S. K., Baker, V., Bardsley, E., Barendrecht, M. H., Bartosova, A., Batelaan, O., Berghuijs, W. R., Beven, K., Blume, T., Bogaard, T., Borges De Amorim, P., Böttcher, M. E., Boulet, G., Breinl, K., Brilly, M., Brocca, L., Buytaert, W., Castellarin, A., Castelletti, A., Chen, X., Chen, Y., Chen, Y., Chifflard, P., Claps, P., Clark, M. P., Collins, A. L., *et al.*: Twenty-three unsolved problems in hydrology (UPH) – a community perspective, Hydrological Sciences Journal, 64, 1141–1158, https://doi.org/10.1080/02626667.2019.1620507, 2019b.

Brönnimann, S., Allan, R., Atkinson, C., Buizza, R., Bulygina, O., Dahlgren, P., Dee, D., Dunn, R., Gomes, P., John, V. O., Jourdain, S., Haimberger, L., Hersbsbach, H., Kennedy, J., Poli, P., Pulliainen, J., Rayner, N., Saunders, R., Schulz, J., Sterin, A., Stickler, A., Titchner, H., Valente, M. A., Ventura, C., and Wilkinson, C.: Observations for reanalyses, B. Am. Meteorol. Soc., 99, 18511866, https://doi.org/10.1175/BAMS-D-17-0229.1, 2018.

Brunner, M. I.: Reservoir regulation affects droughts and floods at local and regional scales, Environ. Res. Lett., 16, 124016, https://doi.org/10.1088/1748-9326/ac36f6, 2021.

Brunner, M. I., Slater, L., Tallaksen, L. M., and Clark, M.: Challenges in modeling and predicting floods and droughts: A review, WIREs Water, 8, https://doi.org/10.1002/wat2.1520, 2021a.

Brunner, M. I., Melsen, L. A., Wood, A. W., Rakovec, O., Mizukami, N., Knoben, W. J. M., and Clark, M. P.: Flood spatial coherence, triggers, and performance in hydrological simulations: large-sample

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

evaluation of four streamflow-calibrated models, Hydrology and Earth System Sciences, 25, 105–119, https://doi.org/10.5194/hess-25-105-2021, 2021b.

Burek, P., van der Knijff, J., and De Roo, A.: LISFLOOD - Distributed Water Balance and Flood Simulation Model - Revised User Manual 2013, https://publications.jrc.ec.europa.eu/repository/handle/JRC78917, last access: 12 January 2024, 2013.

Calvin, K., Patel, P., Clarke, L., Asrar, G., Bond-Lamberty, B., Cui, R. Y., Di Vittorio, A., Dorheim, K., Edmonds, J., Hartin, C., Hejazi, M., Horowitz, R., Iyer, G., Kyle, P., Kim, S., Link, R., McJeon, H., Smith, S. J., Snyder, A., Waldhoff, S., and Wise, M.: GCAM v5.1: representing the linkages between energy, water, land, climate, and economic systems, Geoscientific Model Development, 12, 677–698, https://doi.org/10.5194/gmd-12-677-2019, 2019.

Cammalleri, C., Vogt, J., and Salamon, P.: Development of an operational low-flow index for hydrological drought monitoring over Europe, Hydrological Sciences Journal, 62, 346–358, https://doi.org/10.1080/02626667.2016.1240869, 2017.

Cammalleri, C., Gustavo Naumann, Naumann, G., Mentaschi, L., Bisselink, B., Gelati, E., de Roo, A., and Feyen, L.: Diverging hydrological drought traits over Europe with global warming, Hydrology and Earth System Sciences Discussions, 24, 5919–5935, https://doi.org/10.5194/hess-2020-93, 2020a.

Cammalleri, C., P. Barbosa, Barbosa, P., Barbosa, P., and Vogt, J.: Evaluating simulated daily discharge for operational hydrological drought monitoring in the Global Drought Observatory (GDO), Hydrological Sciences Journal-journal Des Sciences Hydrologiques, 65, 1316–1325, https://doi.org/10.1080/02626667.2020.1747623, 2020b.

Cantoni, E., Tramblay, Y., Grimaldi, S., Salamon, P., Dakhlaoui, H., Dezetter, A., and Thiemig, V.: Hydrological performance of the ERA5 reanalysis for flood modeling in Tunisia with the LISFLOOD and GR4J models, Journal of Hydrology: Regional Studies, 42, 101169, https://doi.org/10.1016/j.ejrh.2022.101169, 2022.

Choulga, M., Moschini, F., Mazzetti, C., Grimaldi, S., Disperati, J., Beck, H., Salamon, P., and Prudhomme, C.: Technical note: Surface fields for global environmental modelling, https://doi.org/10.5194/egusphere-2023-1306, 21 August 2023a.

Choulga, M. M., Beck, H., Moschini, F., Mazzetti, C., Grimaldi, S., Disperati, J., Salamon, P., and Prudhomme, C.: LISFLOOD static and parameter maps for Europe, [dataset] Joint Research Centre Data Catalogue, 2023b.

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., Fowler, H. J., Prudhomme, C., Arnold, J. R., and Brekke, L. D.: Characterizing Uncertainty of the Hydrologic Impacts of Climate Change, Curr Clim Change Rep, 2, 55–64, https://doi.org/10.1007/s40641-016-0034-x, 2016.

CEMS-Flood online documentation: https://confluence.ecmwf.int/display/CEMS/CEMS-Flood, last access: 14 December 2023.

Copernicus: Copernicus Global Land Service - LAI, [dataset] https://land.copernicus.eu/global/products/lai, 2021.

Decremer, D., Mazzetti, C., Carton, C., Gomes, G., Russo, C., Ramos, A., Grimaldi, S., Disperati, J., Ziese, M., Garcia Sanchez, R., Jacobson, T., Salamon, P., and Prudhomme, C.: EFAS v5.0 hydrological reanalysis, [dataset] Joint Research Centre Data Catalogue, 2023.

Despax, A.: Incertitude des mesures de débit des cours d'eau au courantomètre. Amélioration des méthodes analytiques et apports des essais interlaboratoires, These de doctorat, Université Grenoble Alpes (ComUE), 2016.

Donat, M. G., Sillmann, J., Wild, S., Alexander, L. V., Lippmann, T., and Zwiers, F. W.: Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis datasets, J. Climate, 27, 5019–5035, https://doi.org/10.1175/JCLI-D-13-00405.1, 2014.

Ekolu, J., Dieppois, B., Sidibe, M., Eden, J. M., Tramblay, Y., Villarini, G., Peña-Angulo, D., Mahé, G., Paturel, J.-E., Onyutha, C., and van de Wiel, M.: Long-term variability in hydrological droughts and floods in sub-Saharan Africa: New perspectives from a 65-year daily streamflow dataset, Journal of Hydrology, 613, 128359, https://doi.org/10.1016/j.jhydrol.2022.128359, 2022.

LISFLOOD OS online documentation: https://ec-jrc.github.io/lisflood/, last access: 15 December 2023.

LISVAP online documentation: https://ec-jrc.github.io/lisflood-lisvap/, last access: 15 December 2023.

FAO: AQUASTAT, [dataset] https://www.fao.org/aquastat/en/, 2023.

Feyen, L. and Dankers, R.: Impact of global warming on streamflow drought in Europe, J. Geophys. Res., 114, D17116, https://doi.org/10.1029/2008JD011438, 2009.

Florczyk, A. J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Melchiorri, M., Pesaresi, M., Politis, P., Schiavina, M., Sabo, F., and Zanchetta, L.: GHSL Data Package 2019 - Technical report by the Joint Research Centre (JRC), European Union, 38 pp., https://doi.org/10.2760/0726, 2019.

Frieler, K., Volkholz, J., Lange, S., Schewe, J., Mengel, M., del Rocío Rivas López, M., Otto, C., Reyer, C. P. O., Karger, D. N., Malle, J. T., Treu, S., Menz, C., Blanchard, J. L., Harrison, C. S., Petrik, C. M., Eddy, T. D., Ortega-Cisneros, K., Novaglio, C., Rousseau, Y., Watson, R. A., Stock, C., Liu, X., Heneghan, R., Tittensor, D., Maury, O., Büchner, M., Vogt, T., Wang, T., Sun, F., Sauer, I. J., Koch, J., Vanderkelen, I., Jägermeyr, J., Müller, C., Rabin, S., Klar, J., Vega del Valle, I. D., Lasslop, G., Chadburn, S., Burke, E., Gallego-Sala, A., Smith, N., Chang, J., Hantson, S., Burton, C., Gädeke, A., Li, F., Gosling, S. N., Müller Schmied, H., *et al.*: Scenario setup and forcing data for impact model evaluation and impact attribution within the third round of the Inter-Sectoral Model Intercomparison Project (ISIMIP3a), Geoscientific Model Development, 17, 1–51, https://doi.org/10.5194/gmd-17-1-2024, 2024.

Gilbert, M., Nicolas, G., Cinardi, G., Van Boeckel, T. P., Vanwambeke, S. O., Wint, G. R. W., and Robinson, T. P.: Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010, Sci Data, 5, 180227, https://doi.org/10.1038/sdata.2018.227, 2018.

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.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

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, MacMillan, R. A., Batjes, N. H., Heuvelink, G. B. M., Ribeiro, E., Samuel-Rosa, A., Kempen, B., Leenaars, J. G. B., Walsh, M. G., and Gonzalez, M. R.: SoilGrids1km — Global Soil Information Based on Automated Mapping, PLOS ONE, 9, e105992, https://doi.org/10.1371/journal.pone.0105992, 2014.

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.

Hirpa, F. A., Salamon, P., Beck, H. E., Lorini, V., Alfieri, L., Zsoter, E., and Dadson, S. J.: Calibration of the Global Flood Awareness System (GloFAS) using daily streamflow data, Journal of Hydrology, 566, 595–606, https://doi.org/10.1016/j.jhydrol.2018.09.052, 2018.

Huang, Z., Hejazi, M., Li, X., Tang, Q., Vernon, C., Leng, G., Liu, Y., Döll, P., Eisner, S., Gerten, D., Hanasaki, N., and Wada, Y.: Reconstruction of global gridded monthly sectoral water withdrawals for 1971–2010 and analysis of their spatiotemporal patterns, Hydrology and Earth System Sciences, 22, 2117–2133, https://doi.org/10.5194/hess-22-2117-2018, 2018.

International Food Policy Research Institute: Global Spatially-Disaggregated Crop Production Statistics Data for 2010 Version 2.0, [dataset] https://mapspam.info/data/, 2019.

IPCC: Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, 1st ed., Cambridge University Press, https://doi.org/10.1017/9781009157896, 2023.

Karlsson, I. B., Sonnenborg, T. O., Refsgaard, J. C., Trolle, D., Børgesen, C. D., Olesen, J. E., Jeppesen, E., and Jensen, K. H.: Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change, Journal of Hydrology, 535, 301–317, https://doi.org/10.1016/j.jhydrol.2016.01.069, 2016.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., and Thielen, J.: Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level, Environmental Modelling & Software, 75, 68–76, https://doi.org/10.1016/j.envsoft.2015.09.009, 2016.

Klein Goldewijk, K., Beusen, A., Doelman, J., and Stehfest, E.: Anthropogenic land use estimates for the Holocene – HYDE 3.2, Earth System Science Data, 9, 927–953, https://doi.org/10.5194/essd-9-927-2017, 2017.

Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, Journal of Hydrology, 424–425, 264–277, https://doi.org/10.1016/j.jhydrol.2012.01.011, 2012.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

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, 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Kreibich, H., Blauhut, V., Aerts, J. C. J. H., Bouwer, L. M., Van Lanen, H. A. J., Mejia, A., Mens, M., and Van Loon, A. F.: How to improve attribution of changes in drought and flood impacts, Hydrological Sciences Journal, 64, 1–18, https://doi.org/10.1080/02626667.2018.1558367, 2019.

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.

Li, X., Zhou, Y., Hejazi, M., Wise, M., Vernon, C., Iyer, G., and Chen, W.: Global urban growth between 1870 and 2100 from integrated high resolution mapped data and urban dynamic modeling, Communications Earth & Environment, 2, 1–10, https://doi.org/10.1038/s43247-021-00273-w, 2021.

Lin, P., Pan, M., Beck, H. E., Yang, Y., Yamazaki, D., Frasson, R., David, C. H., Durand, M., Pavelsky, T. M., Allen, G. H., Gleason, C. J., and Wood, E. F.: Global Reconstruction of Naturalized River Flows at 2.94 Million Reaches, Water Resources Research, 55, 6499–6516, https://doi.org/10.1029/2019WR025287, 2019.

Mahto, S. S. and Mishra, V.: Does ERA-5 Outperform Other Reanalysis Products for Hydrologic Applications in India?, Journal of Geophysical Research: Atmospheres, 124, 9423–9441, https://doi.org/10.1029/2019JD031155, 2019.

McClean, F., Dawson, R., and Kilsby, C.: Intercomparison of global reanalysis precipitation for flood risk modelling, Hydrology and Earth System Sciences, 27, 331–347, https://doi.org/10.5194/hess-27-331-2023, 2023.

Mentaschi, L., Alfieri, L., Dottori, F., Cammalleri, C., Bisselink, B., Roo, A. D., and Feyen, L.: Independence of future changes of river runoff in Europe from the pathway to global warming, Climate, 8, https://doi.org/10.3390/cli8020022, 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., and Thépaut, J.-N.: ERA5-Land: a state-of-the-art global reanalysis dataset for land applications, Earth Syst. Sci. Data, 13, 4349–4383, https://doi.org/10.5194/essd-13-4349-2021, 2021.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, Journal of Hydrology, 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Nilsson, C., Reidy, C. A., Dynesius, M., and Revenga, C.: Fragmentation and Flow Regulation of the World's Large River Systems, Science, 308, 405–408, https://doi.org/10.1126/science.1107887, 2005.

Paprotny, D. and Mengel, M.: Population, land use and economic exposure estimates for Europe at 100 m resolution from 1870 to 2020, Sci Data, 10, 372, https://doi.org/10.1038/s41597-023-02282-0, 2023.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

Paprotny, D., Terefenko, P., and Śledziowski, J.: An improved database of flood impacts in Europe, 1870–2020: HANZE v2.1, Earth System Science Data Discussions, 1–37, https://doi.org/10.5194/essd-2023-321, 2023.

Peña-Angulo, D., Vicente-Serrano, S. M., Domínguez-Castro, F., Lorenzo-Lacruz, J., Murphy, C., Hannaford, J., Allan, R. P., Tramblay, Y., Reig-Gracia, F., and El Kenawy, A.: The Complex and Spatially Diverse Patterns of Hydrological Droughts Across Europe, Water Resources Research, 58, e2022WR031976, https://doi.org/10.1029/2022WR031976, 2022.

Pfahl, S. and Wernli, H.: Quantifying the relevance of cyclones for precipitation extremes, J. Climate, 25, 6770–6780, https://doi.org/10.1175/JCLI-D-11-00705.1, 2012.

Prudhomme, C., Parry, S., Hannaford, J., Clark, D. B., Hagemann, S., and Voss, F.: How Well Do Large-Scale Models Reproduce Regional Hydrological Extremes in Europe?, Journal of Hydrometeorology, 12, 1181–1204, https://doi.org/10.1175/2011JHM1387.1, 2011.

Richards, N. and Gutierrez-Arellano, C.: Effects of community-based water management decisions at catchment scale, an interdisciplinary approach: the case of the Great Ruaha River Catchment, Tanzania, Water Practice and Technology, 17, 598–611, https://doi.org/10.2166/wpt.2022.010, 2022.

Sauer, I. J., Reese, R., Otto, C., Geiger, T., Willner, S. N., Guillod, B. P., Bresch, D. N., and Frieler, K.: Climate signals in river flood damages emerge under sound regional disaggregation, Nat Commun, 12, 2128, https://doi.org/10.1038/s41467-021-22153-9, 2021.

Schellekens, J., Dutra, E., Weiland, F. S., Minvielle, M., Calvet, J.-C., Decharme, B., Eisner, S., Fink, G., Flörke, M., Peßenteiner, S., van Beek, R., Polcher, J., Beck, H., Orth, R., Calton, B., Burke, S., Dorigo, W., Weedon, G. P., and Delft, H.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, 2017.

Schiavina, M., Freire, S., and MacManus, K.: GHS-POP R2019A - GHS population grid multitemporal (1975-1990-2000-2015) (R2019A), [dataset] https://ghsl.jrc.ec.europa.eu/ghs_pop2019.php, 2019.

Scussolini, P., Luu, L. N., Philip, S., Berghuijs, W. R., Eilander, D., Aerts, J. C. J. H., Kew, S. F., van Oldenborgh, G. J., Toonen, W. H. J., Volkholz, J., and Coumou, D.: Challenges in the attribution of river flood events, WIREs Climate Change, n/a, e874, https://doi.org/10.1002/wcc.874, 2023.

Sood, A. and Smakhtin, V.: Global hydrological models: a review, Hydrological Sciences Journal, 60, 549–565, https://doi.org/10.1080/02626667.2014.950580, 2015.

Speckhann, G. A., Kreibich, H., and Merz, B.: Inventory of dams in Germany, Earth System Science Data, 13, 731–740, https://doi.org/10.5194/essd-13-731-2021, 2021.

Stahl, K., Tallaksen, L. M., Hannaford, J., and van Lanen, H. A. J.: Filling the white space on maps of European runoff trends: estimates from a multi-model ensemble, Hydrol. Earth Syst. Sci., 16, 2035–2047, https://doi.org/10.5194/hess-16-2035-2012, 2012.

Thiemig, V., Gomes, G. N., Skøien, J. O., Ziese, M., Rauthe-Schöch, A., Rustemeier, E., Rehfeldt, K., Walawender, J. P., Kolbe, C., Pichon, D., Schweim, C., and Salamon, P.: EMO-5: a high-resolution multi-variable gridded meteorological dataset for Europe, Earth System Science Data, 14, 3249–3272, https://doi.org/10.5194/essd-14-3249-2022, 2022.

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

Tilloy, A., Malamud, B. D., and Joly-Laugel, A.: A methodology for the spatiotemporal identification of compound hazards: wind and precipitation extremes in Great Britain (1979–2019), Earth System Dynamics, 13, 993–1020, https://doi.org/10.5194/esd-13-993-2022, 2022.

Tilloy, Alois; Paprotny, Dominik; Feyen, Luc; Grimaldi, Stefania; Gomes, Goncalo; Beck, Hylke; Lange, Stefan; Bianchi, Alessandra : HERA: a high-resolution pan-European hydrological reanalysis (1950-2020). European Commission, Joint Research Centre (JRC) [dataset] PID: http://data.europa.eu/89h/a605a675-9444-4017-8b34-d66be5b18c95, 2024:

Tomkins, K. M.: Uncertainty in streamflow rating curves: methods, controls and consequences, Hydrological Processes, 28, 464–481, https://doi.org/10.1002/hyp.9567, 2014.

Tramblay, Y., Arnaud, P., Artigue, G., Lang, M., Paquet, E., Neppel, L., and Sauquet, E.: Changes in Mediterranean flood processes and seasonality, Hydrology and Earth System Sciences, 27, 2973–2987, https://doi.org/10.5194/hess-27-2973-2023, 2023.

Van Loon, A.: Hydrological drought explained, Wiley Interdisciplinary Reviews: Water, 2, 359–392, https://doi.org/10.1002/wat2.1085, 2015.

Vanham, D., Alfieri, L., Flörke, M., Grimaldi, S., Lorini, V., Roo, A. de, and Feyen, L.: The number of people exposed to water stress in relation to how much water is reserved for the environment: a global modelling study, The Lancet Planetary Health, 5, e766–e774, https://doi.org/10.1016/S2542-5196(21)00234-5, 2021.

Vanham, D., Alfieri, L., and Feyen, L.: National water shortage for low to high environmental flow protection, Sci Rep, 12, 3037, https://doi.org/10.1038/s41598-022-06978-y, 2022.

Vassolo, S. and Döll, P.: Global-scale gridded estimates of thermoelectric power and manufacturing water use, Water Resources Research, 41, https://doi.org/10.1029/2004WR003360, 2005.

Wada, Y., Flörke, M., Hanasaki, N., Eisner, S., Fischer, G., Tramberend, S., Satoh, Y., van Vliet, M. T. H., Yillia, P., Ringler, C., Burek, P., and Wiberg, D.: Modeling global water use for the 21st century: the Water Futures and Solutions (WFaS) initiative and its approaches, Geoscientific Model Development, 9, 175–222, https://doi.org/10.5194/gmd-9-175-2016, 2016.

Wood, E. F., Roundy, J. K., Troy, T. J., van Beek, L. P. H., Bierkens, M. F. P., Blyth, E., de Roo, A., Döll, P., Ek, M., Famiglietti, J., Gochis, D., van de Giesen, N., Houser, P., Jaffé, P. R., Kollet, S., Lehner, B., Lettenmaier, D. P., Peters-Lidard, C., Sivapalan, M., Sheffield, J., Wade, A., and Whitehead, P.: Hyperresolution global land surface modeling: Meeting a grand challenge for monitoring Earth's terrestrial water, Water Resources Research, 47, 2010WR010090, https://doi.org/10.1029/2010WR010090, 2011.

Yamazaki, D.: CaMa-Flood, [software] http://hydro.iis.u-tokyo.ac.jp/~yamadai/cama-flood/index.html, 2023.

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.

Yang, Y., Pan, M., Lin, P., Beck, H. E., Zeng, Z., Yamazaki, D., David, C. H., Lu, H., Yang, K., Hong, Y., and Wood, E. F.: Global Reach-Level 3-Hourly River Flood Reanalysis (1980–2019), Bulletin of the

https://doi.org/10.5194/essd-2024-41
Preprint. Discussion started: 19 February 2024

American Meteorological Society, 102, E2086–E2105, https://doi.org/10.1175/BAMS-D-20-0057.1, 2021.

Yu, Q., You, L., Wood-Sichra, U., Ru, Y., Joglekar, A. K. B., Fritz, S., Xiong, W., Lu, M., Wu, W., and Yang, P.: A cultivated planet in 2010 – Part 2: The global gridded agricultural-production maps, Earth System Science Data, 12, 3545–3572, https://doi.org/10.5194/essd-12-3545-2020, 2020.

Zajac, Z., Zambrano-Bigiarini, M., Salamon, P., Burek, P., Gentile, A., and Bianchi, A.: Calibration of the LISFLOOD hydrological model for Europe, https://www.efas.eu/sites/default/files/Manuals/JRC87717_efas_calibration_report%20_final_9_jan_2014.pdf, last access: 15 January 2024, 2013.

Zajac, Z., Revilla-Romero, B., Salamon, P., Burek, P., Hirpa, F. A., and Beck, H.: The impact of lake and reservoir parameterization on global streamflow simulation, Journal of Hydrology, 548, 552–568, https://doi.org/10.1016/j.jhydrol.2017.03.022, 2017.