Hyper-resolution large-scale hydrological modelling benefits from improved process representation in mountain regions

Joren Janzing[1,2,3], Niko Wanders[4], Marit van Tiel[5,6], Barry van Jaarsveld[4], Dirk N. Karger[6], and Manuela I. Brunner[1,2,3]

[1]WSL Institute for Snow and Avalanche Research SLF, Davos Dorf, Switzerland
[2]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
[3]Climate Change, Extremes and Natural Hazards in Alpine Regions Research Center CERC, Davos Dorf, Switzerland
[4]Department of Physical Geography, Utrecht University, Utrecht, the Netherlands
[5]Laboratory of Hydraulics, Hydrology and Glaciology, ETH Zurich, Zurich, Switzerland
[6]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, bâtiment ALPOLE, Sion, Switzerland

Correspondence: Joren Janzing (joren.janzing@slf.ch)

Abstract. Many of the world's major rivers originate in mountain regions and a large fraction of the global population relies on these regions for their water supply. The hydrological cycle of mountain regions and their dependent downstream regions are often studied using large-scale to global hydrological models (LHMs). The increasing spatial resolution of these models allows for improved representation of complex mountain topography, but existing model deficiencies in cold and high-elevation regions limit potential model performance gains. Such model performance gains might be realized by investing in a better representation of hydrological processes that are relevant in mountain regions such as snow-accumulation and -melt. However, how much improved process representation would increase LHM performance remains largely unquantified. Here, we set up the hyper-resolution 30 arcsec (approx. 1 km) global hydrological model PCR-GLOBWB 2.0 (PCRaster Global Water Balance) over the larger Alpine domain and implement several changes to make it better suited for representing hydrological processes in mountain regions. These changes include a.) the use of novel high-resolution meteorological forcing datasets; b.) an extended snow module based on a seasonally varying degree-day factor and an exponential melt function; c.) a regional calibration of the snow module against a snow reanalysis product; d.) a new integrated glacier module; and e.) an adjusted runoff partitioning scheme that increases the contributions to the fast runoff components in the soil. Our evaluation of the effect of these different adjustments on model performance for discharge shows that while the meteorological forcing has a major effect on discharge simulations, it results in a mixed pattern of performance gains and losses over the domain. In addition, the structural and parametric changes, i.e. the snow module modification, glacier representation and runoff partitioning, improve discharge simulations in mountain regions: the snow module modification leads to an improved representation of the snowmelt peak for high-elevation catchments, the glacier module supplies additional water to glacierized catchments, and runoff partitioning in the soil improves the representation of streamflow in flashy catchments at lower elevations. We use these insights to present a new setup of the large-scale and hyper-resolution PCR-GLOBWB 2.0 model that is better suited to study hydrological processes in and beyond mountain regions around the world.

1 Introduction

Mountain regions play a critical role in supplying water to almost 2 billion people living in downstream regions and are therefore often referred to as the "water towers" of the world (Viviroli et al., 2007; Immerzeel et al., 2010, 2020). Water storage in snow and glaciers or a lack thereof is particularly important for drought development and recovery: the snow drought in the Italian Alps in 2022 developed into a streamflow drought downstream and affected many communities in the Po Plain (Colombo et al., 2023), whereas Alpine glaciers provided surrounding rivers with surplus melt-water during the 2003 Central European Drought due to very negative mass balances (Van Tiel et al., 2023). Hydrological processes in mountain regions thus have an over-proportional footprint well beyond mountain ranges. Therefore, considering the driving hydrological processes in mountain regions such as snow accumulation or glacier melt is critical when studying hydro-systems in mountains and their dependent downstream regions.

Large-scale or global hydrological models (LHMs) are often used to study mountain hydro-systems and their dependent downstream areas (e.g. Viviroli et al., 2007, 2020; Khanal et al., 2021), but also to examine many other hydrological systems that exceed the scale of individual catchments. Example applications of such models include water resources (e.g. Dolan et al., 2021; Leijnse et al., 2024) and climate change impact assessments, such as those performed within the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP; Warszawski et al., 2014). However, the coarse spatial resolution of many large-scale models – often tens of kilometers – limits the usefulness of their output for policymakers, who are often interested in more detailed information. This scale gap triggered a call for hyper-resolution, kilometer-scale models that are applicable at continental to global scales (Wood et al., 2011; Bierkens et al., 2015). This call has been addressed by an increasing number of studies: proposed model solutions include the 1 km setup of the ParFlow model for the Contiguous United States (Yang et al., 2023) or the PCR-GLOBWB 2.0 model over Europe at a similar resolution (30 arcsec; Hoch et al., 2023), which Van Jaarsveld et al. (2025) recently used to perform a first global run at hyper-resolution.

At coarse spatial resolutions, there is ample evidence suggesting that LHMs do not accurately capture mountain processes. Recent evaluations found that global hydrological and land-surface models show particularly poor performance at high elevations (Heinicke et al., 2024) and in cold climates (Gädeke et al., 2020; Hou et al., 2023) compared to other regions. One of the underlying issues is the extreme heterogeneity in mountain regions in terms of e.g. topography, meteorology and soil types. Hyper-resolution models are expected to better represent this heterogeneity and could thus improve model performance in mountain catchments, e.g. for snow simulations (Malle et al., 2024). However, to realize such increased model performance, the processes at play also need to be represented and parameterized with sufficient detail and accuracy. Part of the reduced performance in mountain regions can indeed be attributed to issues with process representation, as indicated by misrepresentations of both the volume and timing of snowmelt peaks as well as poor performance in basins where glaciers are not represented by the models (Gädeke et al., 2020). Furthermore, neglecting glaciers or snow transport also leads to the formation of unrealistic "snow towers" at high elevations (Freudiger et al., 2017; Hoch et al., 2023). Several studies thus suggest that improving

cryospheric process representation should be a focus of further LHM development (Gädeke et al., 2020; Heinicke et al., 2024; Van Jaarsveld et al., 2025).

Many LHMs represent snow melt using a temperature-index model which relates melt to air temperature via a degree-day factor (DDF; Telteu et al., 2021). Even though full energy balance models are becoming more popular, especially at the catchment scale, temperature-index models remain widely used because they come with reduced computational demand, require minimal meteorological forcing, and are accurate when calibrated (e.g. Hock, 2003; Magnusson et al., 2015). However, LHMs often use very simplistic temperature-index schemes, e.g. by using only a time-constant DDF or by omitting calibration (e.g. Gosling and Arnell, 2011; Sutanudjaja et al., 2011; Müller Schmied et al., 2021; Stacke and Hagemann, 2021). Snow module comparisons at the catchment scale suggest that performance gains might be achieved by changes in the structure of the snow module (Girons Lopez et al., 2020).

Proper evaluation of such improvements hinges on the availability of high quality reference data to compare how snow is represented by different snow module structures. Evaluations of snow processes in LHMs are often performed against global products representing snow water equivalent (SWE) or snow cover fraction derived from satellite measurements or reanalyses (e.g. Schellekens et al., 2017; Gädeke et al., 2020; Van Jaarsveld et al., 2025). Of the two, SWE is hydrologically the most relevant in regions with seasonal snow cover, but SWE reanalysis products often have too coarse a spatial resolution (often 25 km or more) to be representative for mountain regions (Mortimer et al., 2020). In addition, the coarse model resolutions of LHMs themselves have prevented direct comparisons of SWE simulations with SWE estimates at individual snow measurement stations before the era of high-resolution modelling. Now, higher spatial resolutions of LHMs and new detailed regional SWE reanalysis products (e.g. Mott et al., 2023; Olefs et al., 2020) enable such direct SWE comparisons with the output of LHMs at a regional scale for mountain regions.

Whereas snow modules are present in most LHMs, many models have largely neglected glaciers (Gädeke et al., 2020; Telteu et al., 2021; Hanus et al., 2024). Glaciers can be an important additional water source during their melt season; on average, glacier storage change contribution to total runoff ranges from 4% in the river Danube at Ceatal Izmail in September to 25% for the river Rhône at Beaucaire in August (Huss, 2011), and such contributions can become especially important during drought (Van Tiel et al., 2021, 2023). Including glaciers in hydrological models can thus potentially improve discharge simulations (Wiersma et al., 2022; Hanus et al., 2024), although such improvements will be limited to the summer months and to regions with substantial glacier cover. Glaciers are represented in hydrological modelling in two ways, namely by a.) including an internal glacier module in the hydrological model ("integrated models") or b.) using the output from an external glacier model as the input to the hydrological model (i.e. coupling the models; "coupled models"). An example of a coupled model is the setup created by Hanus et al. (2024), who used output from the Open Global Glacier Model (OGGM; Maussion et al., 2019) as the input to the Community Water Model V1.08 (CWatM; Burek et al., 2020). Similarly, Wiersma et al. (2022) coupled the Global Glacier Evolution Model (GloGEM; Huss and Hock, 2015) to PCR-GLOBWB 2.0. An example of an integrated model at the catchment scale is HBV-light (*Hydrologiska Byråns Vattenavdelning*) (Seibert and Vis, 2012; Seibert et al., 2018a), which calculates glacier mass balance, area evolution and runoff internally. The external glacier models used in coupled model set ups generally have more detailed process representation or more detailed calibration than would be feasible for integrated

glacier modelling in LHMs. Still, we argue that integrated models can also have certain advantages over coupled models. First, integrated glacier modules are physically consistent with the surrounding model framework, which is not necessarily the case for externally coupled glacier models. For example, large-scale glacier models may use different precipitation correction factors for each individual glacier, which is inconsistent with the precipitation in non-glacierized gridcells. Furthermore, glacier geometries evolve over time and assumptions have to be made on how increases in the non-glacierized area are dealt with by the hydrological model (e.g. Hanus et al. (2024) assume the same relative glacier area change over all cells covered by a glacier, whereas in reality area changes mainly affect the glacier terminus). Second, integrated models can be more flexible: an integrated module can – in contrast to coupled models – be run simultaneously with the hydrological model and avoids the coupling steps related to transferring data between the models. This could make integrated models easy to use when forcing them with an ensemble of meteorological forcing datasets.

Aside from snow and glaciers, rivers in mountainous or hilly regions often respond rapidly to local rainfall events, leading to "flashy" discharge behaviour. These flashy responses are caused by the heavy precipitation, thin soils and steep slopes that characterize these regions and that make these regions susceptible to floods (Weingartner et al., 2003). Simplifications in the representation of soil processes and runoff production seem to limit LHM performance in flashier basins (Gharari et al., 2019). Generally, LHMs split the soil into a few layers that store and exchange water, but the exact details can vary significantly: each model has a different number of soil layers (e.g. CWatM: 3 layers (Burek et al., 2020); WaterGAP: 1 soil layer (Müller Schmied et al., 2021)) and these layers can have different thicknesses (e.g. CWatM: upper layer 5 cm thick (Burek et al., 2020), WaterGAP: soil is 0.1 up to 4 m (Telteu et al., 2021)). Interaction between these soil layers determines how water is partitioned over different runoff processes. Most LHMs are not locally calibrated (Telteu et al., 2021), relying instead on a standard parameterization which should avoid obscuring structural deficiencies (Refsgaard and Storm, 1996; Andréassian et al., 2012). Without representing additional soil processes, we hypothesize that hyper-resolution LHMs can already realize further performance gains by reconsidering standard parameterizations (Hoch et al., 2023). For example, on steeper slopes the contribution of near-surface runoff components is larger (Weingartner et al., 2003). Changing how water fluxes from the soil are partitioned across different processes that contribute to discharge (e.g. reduced groundwater recharge and increased saturation excess and interflow) could thus potentially capture more flashy behaviour. This could also improve the local relevance of these models, although their main focus will remain the larger catchments.

Nevertheless, any potential gains in model performance due to improved process representation or parameterization are constrained by the quality of the model input. It is thus important that such conditions are met first. Hydrological modelling is indeed sensitive to the meteorological forcing dataset used as input (e.g. Raimonet et al., 2017; Tang et al., 2023; Gebrechorkos et al., 2024). For hyper-resolution LHMs, the horizontal resolution of the meteorological forcing dataset is of particular importance as using too coarse meteorological forcing can severely reduce potential performance gains from moving towards hyper-resolution hydrological modelling (Hoch et al., 2023). High-resolution meteorological reanalysis products can be derived by downscaling coarser products by exploiting statistical, physical or heuristic relationships or by using dynamically generated regional reanalysis products. Both types of downscaled products often outperform coarser global reanalysis products, e.g. in representing precipitation (e.g. Karger et al., 2021b; Keller and Wahl, 2021). These products should also represent

temperature gradients with elevation in more detail, which can be important for snow modelling (Malle et al., 2024). However, a main difference between the two types of downscaling is that regional reanalysis products also explicitly represent higher resolution atmospheric dynamics. While Hoch et al. (2023) studied the effect of the spatial resolution of the meteorological forcing dataset (using statistical downscaling) on hyper-resolution LHM performance, it remains to be assessed how the exact procedure of deriving data at higher spatial resolutions influences model performance. Furthermore, despite improved resolutions, precipitation products in particular are known for large uncertainties over mountain regions (e.g. Isotta et al., 2015; Gampe and Ludwig, 2017; Bandhauer et al., 2022). It is thus important to assess how sensitive hydrological model performance in mountain regions is to the specific biases and large uncertainties of meteorological forcing datasets.

While large-scale hydrological simulations at higher spatial resolution have become feasible thanks to increasingly available computational resources, it is yet unclear by how much hydrological simulations can improve when combining such high-resolution models with the latest generation of meteorological forcing datasets and improved process representation. Therefore, here we aim to explore the effect of (1) using different meteorological datasets, (2) improving snow and glacier representations, and (3) changing runoff partitioning in the soil on discharge simulations in PCR-GLOBWB 2.0. We focus on the larger Alpine region, as an example of a mountain region that is normally implicitly simulated by LHMs. A similar set-up can in principle be applied at larger scales. While more detailed hydrological modelling approaches are available for the Alps at the national or catchment scale, the Alps are very rarely studied as a whole at a similarly high spatial resolution, which inhibits comparisons between different regions within the Alps. We hypothesize that hyper-resolution LHM performance for discharge in mountain regions will increase by (H1) using forcing products that include a representation of smaller-scale atmospheric dynamics compared to other forcing products, (H2) improving the representation of mountain hydrological processes, such as snow and ice melt; and (H3) reviewing and adjusting standard parameterizations. To test these hypotheses, we first assess how strongly discharge simulations are affected by the meteorological forcing chosen to drive the model. Second, we quantify the effect of structural changes in the model setup on model performance, namely by expanding the existing snow module and adding a new glacier module. Third, we study the effect of parameter changes on model performance by calibrating SWE against a detailed regional SWE reanalysis product with assimilated observations and changing parameters controlling the volumes of soil compartments.

2 Methods

2.1 Model setup and study outline

We use the PCR-GLOBWB 2.0 model (Sutanudjaja et al., 2018) in the 30-arcsec setup developed by Hoch et al. (2023) (approx. 1 km at the equator, 650 m in the longitudinal direction in the Alps). The model runs at a daily time step. PCR-GLOBWB 2.0 is a global hydrological model and contains different modules, which represent both natural processes related to vegetation, snow, soil, groundwater and river routing and anthropogenic processes such as human water use and irrigation. Here, we use a regional model setup (longitude: 3-18°; latitude: 43-51°) covering the Alps and the upstream parts of the catchments of four

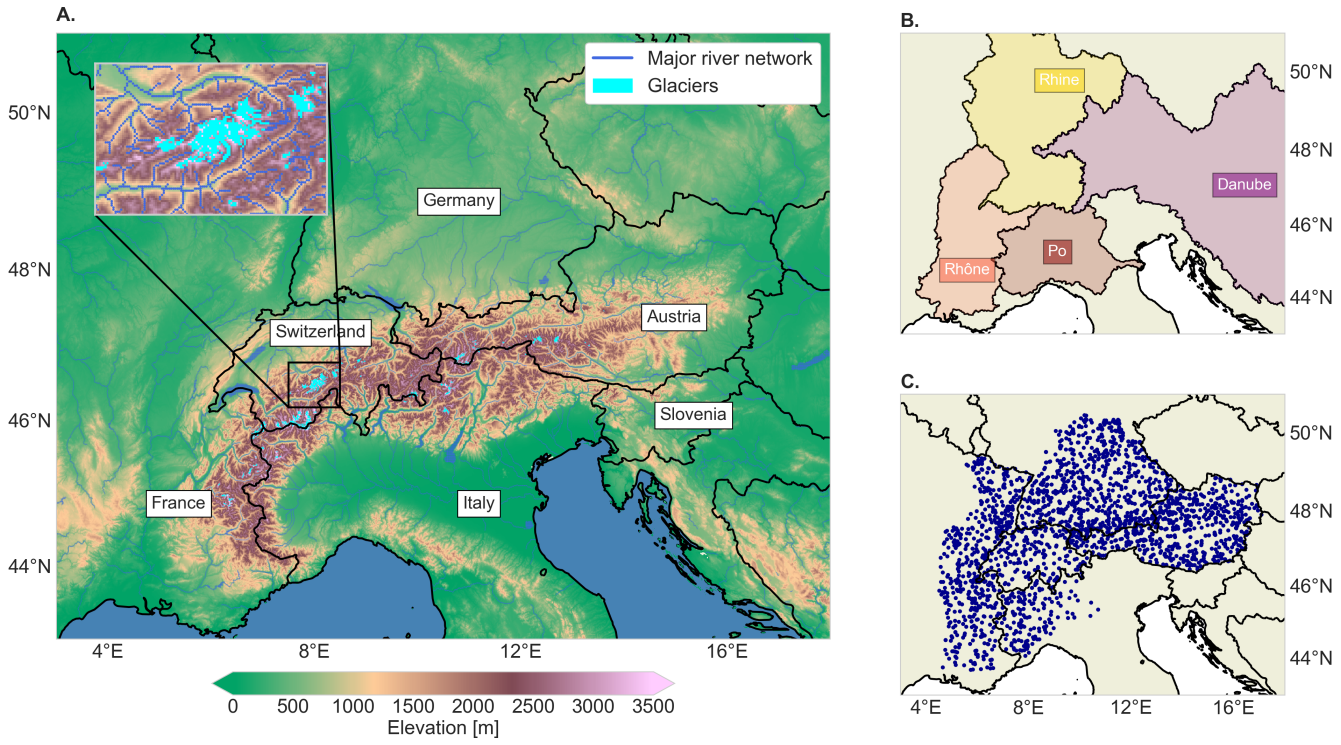

Figure 1. A. Overview of the model domain, highlighting the larger Alpine countries, major rivers, topography, and glaciers. The inset map shows the area around the Aletsch Glacier (the largest glacier in the Alps) in more detail. Elevations are derived from the upscaled MERIT Hydro DEM (Yamazaki et al., 2019). B. The larger river basins of the Alpine rivers, indicated with a slight shading. C. The discharge stations used to evaluate the model (see Section 2.2).

major Central European rivers (i.e. the Rhone, Rhine, Danube, and Po; Figure 1). We focus on the period 1990–2019, as all forcing datasets are available for this time period and initial glacier volumes are often only available for around the year 2000.

We implement the different forcing datasets and model changes in a step-by-step manner and thus perform several model runs. A schematic overview and further details on the sequential model runs performed in this study are provided in Figure 2 and Table 1, respectively.

2.2 Datasets

2.2.1 Model input

To quantify the sensitivity of hydrological models to the choice of the forcing dataset, we assess how discharge simulations vary under different meteorological forcing datasets. We focus on the input variables precipitation rate and near-surface air temperature as these are available for a wide range of potential meteorological datasets. Evaporation is then calculated within PCR-GLOBWB 2.0 using the method from Hamon (1963). For our comparison, we use the following meteorological datasets:

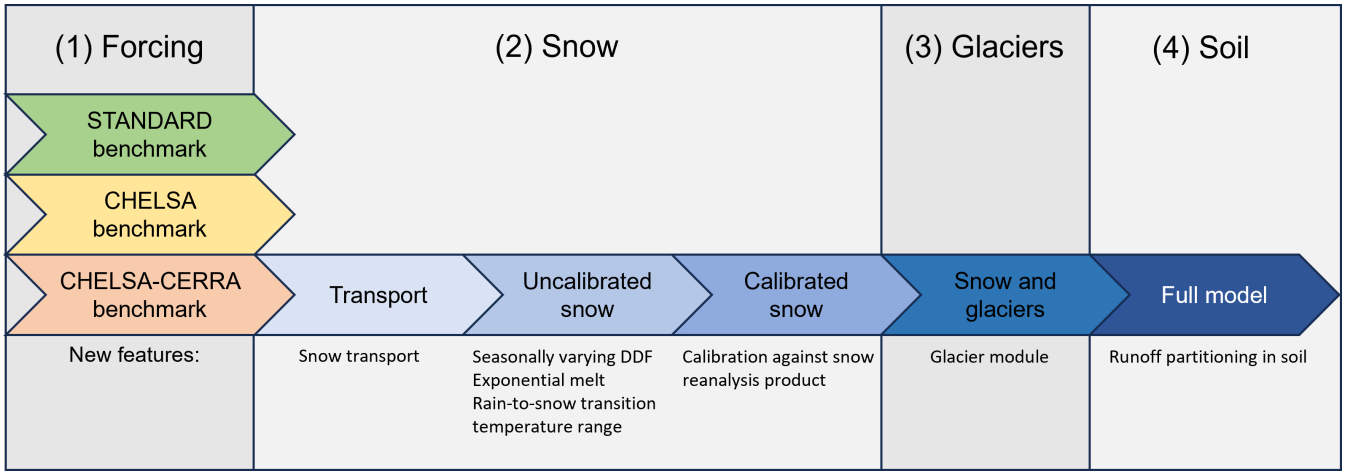

Figure 2. Overview of the different model runs performed in this study. (1) The standard model set-up with standard parameterization is our benchmark model and we run it with different forcing datasets. (2) Several changes are implemented in the snow module in different runs. (3) A new glacier module is added. (4) Runoff partitioning is adjusted, which is the final model change.

(1) "STANDARD" input for hyper-resolution PCR-GLOBWB 2.0, (2) Climatologies at High resolution for the Earth's Land Surface Areas v2.1 (CHELSA) and (3) CERRA-CHELSA, a mixed dataset with temperature from the Copernicus European Regional ReAnalysis (CERRA) further downscaled using the CHELSA algorithm and precipitation data directly from CERRA-Land.

The STANDARD forcing was created by Van Jaarsveld et al. (2025). They created an internal meteorological downscaling scheme in PCR-GLOBWB 2.0. This scheme uses coarser scale meteorological input, in this case the W5E5 v2.0 (WFDE5 over land merged with ERA5 over the ocean) dataset (Lange et al., 2021), which has a spatial resolution of 0.5 degrees. This coarser dataset is then downscaled to 30 arcsec spatial resolution using monthly climatologies from CHELSA-BIOCLIM+ (Climatologies at High resolution for the Earth's Land Surface Areas - bioclimatic variables plus) (Karger et al., 2017; Brun et al., 2022b). CHELSA (Karger et al., 2017, 2021a) is a downscaled reanalysis product based on the ERA5 product (Hersbach et al., 2020). CHELSA uses heuristic and physical relationships to downscale this forcing data to 30 arcsec spatial resolution. Downscaling is performed using topography, atmospheric lapse rates (for temperature; Karger et al. (2023)), spatial wind fields, and the height of the boundary layer (for precipitation; Karger et al. (2021b)). Finally, CERRA (Schimanke et al., 2021; Ridal et al., 2024) is a regional reanalysis product over Europe provided at 5.5 km spatial resolution (approx. 180 arcsec). CERRA-Land is the associated surface analysis (Verrelle et al., 2022), which includes also additional data-assimilation of precipitation observations. We decided to use precipitation directly from CERRA-Land, since this dataset is already near the effective resolution of precipitation and the terrain effect could be over-represented by downscaling the data further (Daly et al. (1997); Karger et al. (2021b) uses 3 km). In contrast, near-surface air temperature can be further downscaled to generate more accurate spatial melt patterns at 30 arcsec resolution. Therefore, we created a new "CHELSA-CERRA" temperature dataset, for which temperature was taken from the CERRA dataset and was downscaled using the topographical CHELSA v2.1 algorithm (Karger et al.,

Table 1. List of the sequential model runs performed within this study.

Run name	Snow	Glaciers	Soil	Forcing	Time period	Purpose
STANDARD benchmark	Standard, without snow transport	No	Standard	STANDARD	Evaluation II: 2010-2019	Comparing meteorological input; Benchmark
CHELSA benchmark	Standard, without snow transport	No	Standard	CHELSA	Evaluation II: 2010-2019	Comparing meteorological input
CERRA-CHELSA benchmark	Standard, without snow transport	No	Standard	CERRA-CHELSA	Evaluation II: 2010-2019	Comparing meteorological input; Benchmark
Transport	Standard, with snow transport	No	Standard	CERRA-CHELSA	Evaluation II: 2010-2019	Effect of snow transport
Uncalibrated snow	Updated, but uncalibrated	No	Standard	CERRA-CHELSA	Evaluation II: 2010-2019	Effect of model structure changes
Calibrated snow	Updated	No	Standard	CERRA-CHELSA	Evaluation II: 2010-2019	Effect of snow calibration
Snow and glaciers	Updated	Yes	Standard	CERRA-CHELSA	Evaluation II: 2010-2019	Effect of glaciers
Full model	Updated	Yes	Updated	CERRA-CHELSA	Full period: 1990-2019	Effect of runoff partitioning

2023). For simplicity, we refer to the combined meteorological product as "CERRA-CHELSA" in the remainder of this paper.

To implement a new glacier module, we used the locations of the existing glaciers and their initial ice thickness using the consensus estimate from Farinotti et al. (2019), which are representative for the year 2003 for most glaciers. To our knowledge, there are no estimates of glacier volumes available that go further back in time and that cover all glaciers on the global, continental or Alpine scale. This glacier volume dataset has a varying spatial resolution (max. 200 m) has previously also been used to initialize glaciers for the same study period by Hanus et al. (2024).

2.2.2 Evaluation and calibration data

Within this study, we use several reference datasets against which we compare model inputs (forcing) and outputs such as streamflow, SWE, and glacier changes (Table 2). As a reference meteorological dataset, we use the Alpine Gridded Precipitation Dataset (APGD; Isotta et al., 2014). This is a gridded product at 5 km spatial resolution covering the period 1971–2008 and is based on interpolated rain gauge data over the Alps. We choose this dataset as a reference, because unlike the other datasets

it is specifically created for the Alps, does not use reanalyses, and has been used as a reference meteorological dataset before (Isotta et al., 2015). However, the dataset was not corrected for undercatch (which can be tens of percentage points in the Alps (Sevruk, 1985)). Please note that we do not use this dataset as the forcing input, since LHMs are generally run at larger scales and therefore we used forcing products that are at least available at the continental scale.

Reference measurements of discharge in rivers around the Alps are taken from both national and regional agencies (sources listed in Table A1). We only select stations in the basins of the Po, Rhine, Rhone and Danube rivers, leaving us with 2167 stations in total for the period 1990–2019. Please note that for evaluations, we only use stations that have at least 3 years of data over the considered evaluation period (see Section 2.4). Since this dataset is an extended and updated version of the Large-sample hydro-meteorological dataset for the Alps (LHDA) from Schlemper et al. (2024), we refer to this dataset as LHDA+.

For SWE comparisons, we use a set of different SWE products. We use two 1 km-gridded SWE regional reanalysis products at daily resolution, namely 1) a product by the Operational Snow Hydrological Service (OSHD) over (hydrological) Switzerland (Mott et al., 2023), which is available for the period 1998–2022, and 2) a product over Austria created with the SNOWGRID model (Olefs et al., 2013, 2020) for the period 1961–2020. Furthermore, we use the spatially much coarser SWE output from two atmospheric reanalysis products, namely CERRA-Land (5.5 km; 1984–2021) (Verrelle et al., 2022) and ERA5-Land (0.1 degree resolution (approx. 9 km); 1950–present) (Copernicus Climate Change Service, 2019). Finally, we use data from 1047 Alpine measurement stations where SWE was inferred from snow depth (Fontrodona-Bach et al., 2023b). For the locations, we refer the reader to Figure S2 in the Supporting Information.

For glaciers, we use spatially explicit maps on glacier elevation changes derived from satellite observations at 100 m spatial resolution (Hugonnet et al., 2021a), that are then upscaled to match the 30 arcsec model resolution. Here, we refer to this dataset as the "Glacier elevation change maps" (GECM). We use maps from two ten-year periods, namely 2000–2009 and 2010–2019. We further use individual glacier mass balance measurements from World Glacier Monitoring Service (WGMS) (2023) and Glacier Monitoring Switzerland (GLAMOS) (2022). We only considered glaciers that have a minimum size of 3 km^2 (roughly 4-5 grid cells). Finally, we use glacier response time simulations from Zekollari et al. (2020) and glacier outlines and surface area estimates from Raup et al. (2007) and the GLIMS Consortium (2018) for further evaluation (see Section 2.4).

For soil moisture evaluation, we use the European Space Agency Climate Change Initiative (ESACCI) COMBINED soil moisture data v8.1 (Gruber et al., 2019; Dorigo et al., 2017; Preimesberger et al., 2021). The ESACCI dataset includes satellite observations of soil moisture in the top 5 cm of the soil at a resolution of 0.25 degrees.

2.2.3 Ancillary data

To analyze what explains spatial patterns in model performance, we needed additional information on catchments, climate or topography. For each of the catchments in LHDA+ the dataset includes a range of catchment characteristics. Catchment area and reservoirs (location and capacity) were both derived from government agencies or from existing databases (see Table A1). Although our reservoir database is more detailed than other large-scale databases such as the one by Lehner et al. (2005), it does not provide a complete overview of all reservoirs in the region. The fraction of the catchment covered by glaciers was

computed from the Randolph Glacier Inventory (RGI) 6.0 (Pfeffer et al., 2014; RGI Consortium, 2017). Snowfall fraction and potential evapotranspiration per catchment were calculated based on our own simulated model output. Finally, for analyses that use elevation, we use the Multi-Error-Removed Improved-Terrain Hydro Digital Elevation Model (MERIT Hydro DEM) (Yamazaki et al., 2019). The MERIT Hydro DEM was upscaled from its original 3 arcsec resolution to 30 arcsec resolution by Hoch et al. (2023), who used it as the default DEM of the 30 arcsec version of PCR-GLOBWB 2.0.

2.3 Model development

Based on the initial regional model setup introduced in Section 2.1, we further develop the representation of cryospheric and soil processes to improve discharge simulations in mountain regions. We aim to find a regionally valid setup that works well for a larger domain and that is thus not directly fine-tuned for individual catchments, in line with the philosophy of many global hydrological models. The next few sections describe the structural changes made to the PCR-GLOBWB 2.0 model. The parameters used in the equations and any fixed values are listed in Table S1 in the Supporting Information. The calibration strategy for specific parameters is then further outlined in Section 2.3.4.

2.3.1 Snow module

The existing version of PCR-GLOBWB 2.0 includes a snow module consisting of a temperature index approach with a constant DDF. A temperature-index model generally has the following form:

$$M = \text{DDF} \times (T - T_{\text{thresh}}), \tag{1}$$

where M represents the melt rate (m day^{-1}), DDF the degree-day factor (m°C^{-1}day^{-1}), T the daily average temperature (°C), and T_{thresh} the temperature threshold above which melt occurs. We build on this existing setup and expand it with elements of the snow model outlined in Magnusson et al. (2014), namely (1) a seasonally varying DDF, (2) exponential dependence on temperature, and (3) a rain-to-snowfall transition temperature range.

First, we replace the constant DDF with a seasonally varying one to capture the effects of changes in the solar declination throughout the year, following the approach outlined in Slater and Clark (2006):

$$\text{DDF} = \frac{\text{DDF}_{\text{max}} + \text{DDF}_{\text{min}}}{2} + \sin\left(\frac{k2\pi}{366}\right)\left(\frac{\text{DDF}_{\text{max}} - \text{DDF}_{\text{min}}}{2}\right). \tag{2}$$

For the Northern Hemisphere, DDF_{max} is the degree-day factor on 21st of June (summer solstice; m°C^{-1}day^{-1}) and DDF_{min} is the degree-day factor on 22nd of December (winter solstice; m°C^{-1}day^{-1}). k represents the day of the year since 21st of March (equinox; -). Second, we implement an exponential relationship between temperature and melt following Magnusson et al. (2014). This formulation makes the melt more sensitive to increasing temperatures than under the assumption of a linear relationship and allows for limited melt below the threshold temperature, to account for days when the average temperature is below the threshold temperature, but the maximum temperature surpasses it.

$$M = \text{DDF}\, m_{\text{m}} \left(\frac{T - T_{\text{thresh}}}{m_{\text{m}}} + ln\left(1 + \exp\left(-\frac{T - T_{\text{thresh}}}{m_{\text{m}}}\right)\right)\right), \tag{3}$$

Table 2. List of used datasets and their description.

Name	Variables	Spatial coverage	Temporal coverage	Spatial resolution	Temporal resolution	Reference
Meteorology						
STANDARD	Precipitation and temperature	Global	1979-2019	30 arcsec	Daily	Van Jaarsveld et al. (2025)
CHELSA (v2.1)	Precipitation and temperature	Global	1979-2019	30 arcsec	Daily	Karger et al. (2021a, b, 2023)
CERRA and CERRA-Land (for CERRA-CHELSA)	Precipitation and temperature	Europe	1984-2021	5.5 km (temp. downscaled to 30 arcsec)	Daily	Schimanke et al. (2021); Verrelle et al. (2022); Ridal et al. (2024)
APGD	Precipitation	Alps	1971-2008	5 km	Daily	Isotta et al. (2014); Isotta and Frei (2013)
Discharge						
LHDA+	Discharge, catchment attributes	Alps	1990-2020	Stations	Daily	based on Schlemper et al. (2024)
Snow						
OSHD	SWE	Switzerland	1998-2022	1 km	Daily	Mott et al. (2023); Mott (2023)
SNOWGRID	SWE	Austria	1961-2020	1 km	Daily	Olefs et al. (2020)
ERA5-Land	SWE	Global	1950-present	0.1 degree	Daily	Copernicus Climate Change Service (2019)
CERRA-Land	SWE	Europe	1984-2021	5.5 km	Daily	Verrelle et al. (2022)
NH-SWE	SWE	Northern Hemisphere	1950-2022	Stations	Daily	Fontrodona-Bach et al. (2023b)
Glaciers						
GECM	Elevation changes	Global	2000-2009; 2009-2019	100 m (upscaled to 30arcsec)	10 years	Hugonnet et al. (2021a)
WGMS	Mass balance	Global	Varying	Glaciers	Yearly	World Glacier Monitoring Service (WGMS) (2023)
GLAMOS	Mass balance	Switzerland	Varying	Glaciers	Yearly	Glacier Monitoring Switzerland (GLAMOS) (2022)
Consensus Estimate	Glacier volumes	Global	Varying, approx. 2003	Varies per glacier (max. 200 m)	-	Farinotti et al. (2019)
Response time simulations	Glacier volume	Alps	2018-2300	Glaciers	Yearly	Zekollari et al. (2020)
GLIMS	Glacier outlines and area estimates	Global	Varying			Raup et al. (2007); GLIMS Consortium (2018)
RGI	Glacier outlines and area estimates	Global	Varying			Pfeffer et al. (2014); RGI Consortium (2017)
Soil						
ESACCI	Soil moisture	Global	1978-2023	0.25 degrees	Daily	Gruber et al. (2019); Dorigo et al. (2017); Preimesberger et al. (2021); Dorigo et al. (2023)
Elevation						
MERIT Hydro DEM	Elevation	Global	-	3 arcsec (upscaled to 30 arcsec)	-	Yamazaki et al. (2019)

where m_{m} is a parameter controlling the transition between melt and no melt (°C) and was kept constant by Magnusson et al. (2014).

Third, we adapt the snowfall and rainfall partitioning to account for snow and rainfall coincidence by creating a temperature transition zone where rainfall smoothly changes into snowfall (Magnusson et al., 2014).

$$P_{\mathrm{snowfall}} = \frac{P}{1 + \exp\left(\frac{T - T_{\mathrm{snowfall}}}{m_{\mathrm{p}}}\right)}, \tag{4}$$

where P_{snowfall} represents precipitation falling as snow (m day^{-1}), P total precipitation (m day^{-1}), T daily average temperature (°C), T_{snowfall} temperature below which most precipitation falls as snow (°C), and the parameter m_p determines the range where snow and rainfall co-occur (°C) and was derived from snowfall observations by Magnusson et al. (2014). Aside from additions to the snow module, we also ignore refreezing in the snowpack, since previous analyses have shown that it did not improve simulations (Magnusson et al., 2014; Girons Lopez et al., 2020).

Furthermore, we also include the lateral snow transport scheme introduced by Van Jaarsveld et al. (2025) as a separate development step to better quantify its effect against the other development steps in the snow and glacier modules. Van Jaarsveld et al. (2025) implemented a lateral snow transport scheme based on Frey and Holzmann (2015) to avoid unrealistic snow accumulation at high elevations. This scheme transports part of the snow downhill based on the surface slope whenever the snowcover exceeds a SWE content of 0.625m. This threshold is based on values for the forest snow holding capacity and snow density in Frey and Holzmann (2015), using a similar approach as in CWatM (Burek et al., 2020). However, the snow that is transported should sometimes be part of glacier accumulation, which is why we here apply the lateral transport scheme only outside of glaciers. When we introduce the glacier module, we thus restrict the lateral snow transport and apply it only on non-glacierized areas. This means that snow can only be transported from a.) a non-glacierized cell to a non-glacierized cell and b.) from a non-glacierized cell to a glacierized cell. There is no snow transport from a glacierized cell to either a glacierized cell or a non-glacierized cell. When snow is transported onto a glacierized cell, it is becomes part of the snow cover on the glacier: it can thus reduce ice melt and become part of the glacier accumulation (Kuhn, 2003; Freudiger et al., 2017).

2.3.2 Glacier module

We introduce a new glacier module to PCR-GLOBWB 2.0. To create glacierized cells, we derive glacier geometries and volumes from ice thickness estimates by Farinotti et al. (2019). We then resample and regrid these volumes to the model raster of 30 arcsec, applying a correction factor to preserve the total ice volume of the glaciers. A cell is considered fully glacierized as soon as any ice is present in the gridcell, and there are no partially glacierized cells. The static part of the glacier scheme is based on Seibert et al. (2018a) and is schematically shown in Figure 3. The glacier consists of two parts: an ice reservoir and a water reservoir, representing water contained within the glacier. The glacier ice reservoir only decreases by melting when it is not covered by snow, following a simple temperature-index scheme (see Equation 1) using the DDF for snow multiplied with a correction factor (C_{ice}) to account for the lower albedo of the glacier ice surface (Seibert et al., 2018a). The glacier water reservoir increases in volume through the addition of glacier melt, snowmelt occurring on the glacier, and rain falling on the glacier during times when no snow is present. If snow cover is present, rainfall is added to the snow free water reservoir. The

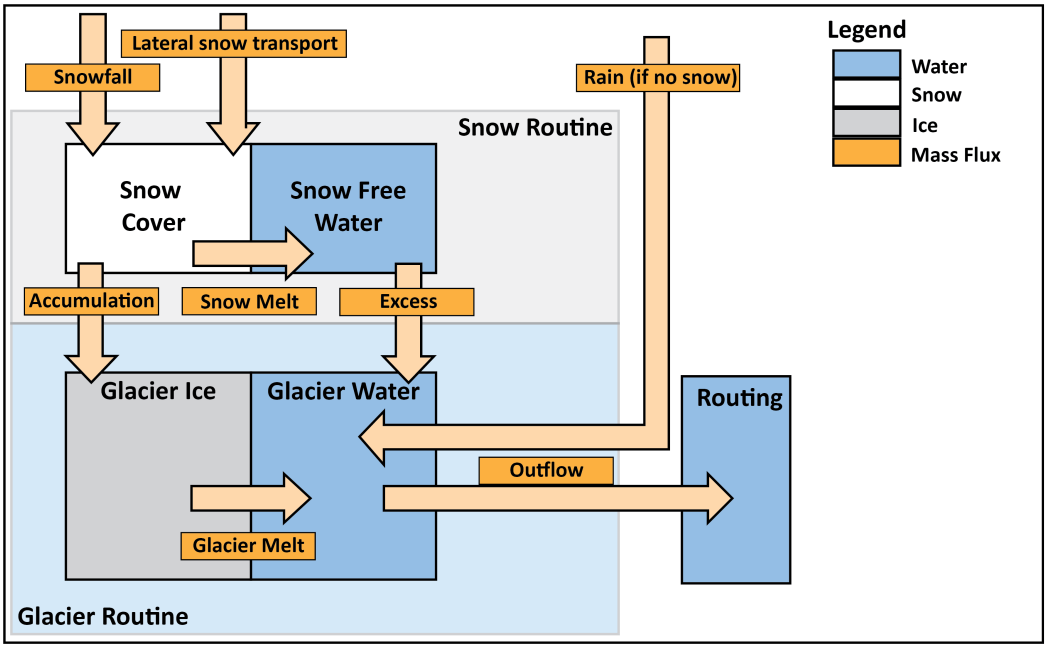

Figure 3. Schematic overview of the glacier module and its interaction with the snow module.

glacier water is then released from each individual glacier cell in the following way (Stahl et al., 2008):

$$Q = S(K_{\text{min}} + K_{\text{range}} * e^{-A_{\text{g}}\text{SWE}}), \tag{5}$$

where Q is the glacial water release (m/day), S the glacial water storage (m), and SWE the snow water equivalent on the glacier cell (m). K_{min}, K_{range} (day^{-1}) and A_{g} (m^{-1}) are additional parameters determining the rate of melt water release.

For glacier accumulation, every day a fraction (f_{acc}; day^{-1}) of the snow on the glaciers is converted to glacial ice and is added to the glacier ice reservoir (Seibert et al., 2018a).

Glacier geometries change over time depending on their mass balance, which affects the quantity of melt over time. To account for such temporal changes, we implement the empirical Δh-parameterization scheme from Huss et al. (2010), as implemented in HBV (Seibert et al., 2018a). As the glacier volumes are representative for around the year 2003, we only apply this scheme after the year 2000, keeping the glacier volume and area constant between 1990-2000 while still calculating ice melt and accumulation. The Δh-parameterization is based on the assumption that glacier mass loss leads to specific patterns of change in glacier ice surface elevation. Huss et al. (2010) identified specific parameterizations for glaciers of different sizes that relate changes in mass to surface elevation changes. Note that we do not apply the width-scaling applied by Seibert et al. (2018a), as we are running the model in a spatially-distributed way. Before running the model, we assign a glacier ID to each group of glacier cells that are part of an individual glaciers based on the RGI outlines. Then, we create offline maps of distributed glacier thickness, where each individual glacier loses a specific fraction of its mass (in steps of 1 percent mass loss) following Seibert et al. (2018a). During the hydrological model runs, we update the individual glacier geometries each year

after the 1st of September (i.e. beginning of the hydrological year used by the OSHD), by reading in the distributed glacier thickness from these offline maps based on the mass balance simulated for the current year. Mass changes do not necessarily occur in steps of 1 percent, leading to leftover mass or mass loss: for example, if the total mass loss of a glacier is 2.3 percent of the initial volume, we are left with 0.3 percent of leftover mass loss for this glacier. To address this, we distribute such leftover mass or mass loss evenly over the glacier area. Under the Δh-parameterization, glaciers can only grow to their original extent. If glaciers gain mass compared to their initial extent, we add the additional mass to the gridcell downstream of the glacier. A full description of the Δh-parameterization is provided in Appendix B.

2.3.3 Soil module

In the original model setup, each grid cell in PCR-GLOBWB 2.0 has two main soil layers and can contribute to river flow in three main ways: via direct runoff (infiltration or saturation excess), interflow, or groundwater contributions (Sutanudjaja et al., 2018). The first two components can be considered "fast" components, whereas the latter represents a "slow" component. Initial model runs performed with this standard setup suggested that the model produces too slow runoff responses and can not properly represent the fast components. Therefore, we introduced some measures to allow the model to produce more flashy runoff responses. Based on a sensitivity analysis, we decided to reduce the size of the top soil layer by making it half as thick everywhere, while keeping the total soil thickness constant. Since the depth of the upper soil layer in PCR-GLOBWB 2.0 is constant over the domain and is by default set to 30 cm (Bierkens and Van Beek, 2009), halving the thickness still corresponds to a thickness of 15 cm, which is in line with the range of thicknesses that other LHMs are able to resolve (Telteu et al., 2021). This simple change should better represent the behaviour of thin soils often found on steep slopes (Weingartner et al., 2003), and causes saturation excess and interflow to occur more rapidly, thus enabling faster runoff responses.

2.3.4 Calibration

PCR-GLOBWB 2.0 has not been calibrated and generally uses parameters derived from external datasets (Sutanudjaja et al., 2018). Still, we do calibrate the degree-day factors of snow and ice to increase regional applicability and to ensure realistic glacier geometry evolution. We calibrate on the specific process (e.g. on SWE) instead of discharge to avoid compensating for deficiencies in other processes with parameter calibration and to increase the stability of these parameters when temperature change (Sleziak et al., 2020). We try to find one parameter set that is regionally valid, i.e. constant over the entire domain. To save computational time, we perform the calibration of snow and glacier parameters offline, i.e. we run the snow and glacier component separately without running the rest of PCR-GLOBWB 2.0 and compare it to the reference datasets. We use Monte Carlo sampling as implemented in the SPOTPY python package (Houska et al., 2015) and manually test the realism of identified parameter sets. In contrast, model evaluation is performed using the full model run. The considered calibration period is 2000–2009 (see Section 2.4). Note that any other parameters not mentioned here remain fixed and their values can be found in Table S1 in the Supporting Information.

The updated snow module required the calibration of 2 parameters, namely DDF_{max} and DDF_{min}. As a reference dataset, we used a snow water equivalent reanalysis product over Switzerland (Mott et al., 2023). We chose this dataset because of its extensive data assimilation, its spatial continuity, its high quality which is unmet by products covering larger spatial domains (e.g. ERA5-Land or CERRA-Land), and because Switzerland covers diverse climatic regions. However, as every other SWE dataset, this product also comes with some uncertainties. Note that we do not calibrate over Austria, even though we have similar data available there as well, to be able to test the transferability of the snow scheme to other regions. To explicitly account for elevation-dependent melt patterns, we averaged SWE spatially over partially-overlapping elevation zones (500-1500 m, 1000-2000 m, 1500-2500 m, 2000-3000 m, 2500-3500 m and Switzerland) and maximized the average Nash-Sutcliffe Efficiency (NSE; Nash and Sutcliffe (1970)) across these elevation bands.

After the snow calibration, we calibrate the glacier module. This module requires the calibration of a melt correction factor C_{ice} and accumulation factor f_{acc}. We use satellite measurements of glacier elevation changes (Hugonnet et al., 2021a) as a reference dataset due to their coverage of all glaciers in the calibration domain (Switzerland for consistency with the snow calibration). We calculate the elevation changes between the beginning and end of the calibration period: We correct the model output for density differences between ice and water ($\rho_{\text{ice}} = 916.7$ kg m^{-3} ; $\rho_{\text{water}} = 1000$ kg m^{-3}), only focus on locations with larger elevation changes (>2 m) to avoid noise) and minimize the mean absolute error for the elevation changes.

2.4 Evaluation

The aim of our evaluation is to assess the effect of the model development both on simulated discharge and the representation of individual processes such as snow accumulation and melt. We thus evaluate the forcing datasets for precipitation and the model output for streamflow, SWE, glacier mass balance changes, glacier surface elevation evolution and soil moisture. We generally evaluate the daily simulations, except for the glaciers for which we use annual mass balances. For evaluation, we split the study period into three blocks with different annual mean air temperature characteristics: (1) the calibration period (2000–2009), (2) a colder evaluation period (Evaluation I, 1990–1999); and (3) a warmer evaluation period (Evaluation II, 2010–2019) (see Figure S1A in the Supporting Information). The main evaluation is performed over Evaluation period II, which is also the period considered if a specific time interval is not specified. In the Supporting Information (Section S2), we use both evaluation periods to assess the transferability of the new model set up to time periods with different mean temperatures within the framework of a differential split sample test (DSST; Klemeš, 1986; Seibert, 2003).

The evaluation of the meteorological forcing focuses on precipitation, which is more spatially heterogeneous than temperature. We try to analyze the differences between the datasets and their realism, acknowledging the large uncertainties around high-altitude precipitation. We resample all three forcing datasets (STANDARD, CHELSA and CERRA; see Section 2.2) to the grid points of the coarser reference APGD and calculate both the Pearson correlation and the absolute bias against the APGD.

We evaluate the performance of the daily discharge simulations by comparing simulated discharge against observed streamflow, i.e. the station data from the LHDA+. We match model grid cells to discharge stations by matching the catchment contours following Godet et al. (2024). We then evaluate discharge simulation performance by comparing the simulated to the observed

time series using the Kling-Gupta efficiency (KGE; Gupta et al., 2009), using only stations that have at least three years of data over the considered evaluation period (number of valid stations: Evaluation period I: 1676; Calibration period: 1869, Evaluation period II: 2105). The KGE is defined as:

$$\text{KGE} = 1 - \sqrt{(r-1)^2 + (\frac{\mu_{sim}}{\mu_{obs}} - 1)^2 + (\frac{\sigma_{sim}}{\sigma_{obs}} - 1)^2}, \tag{6}$$

where, r is the Pearson correlation coefficient between observations and simulations, μ_{obs} is the mean over the observations, μ_{sim} is the mean over the simulations, σ_{obs} is the standard deviation over the observations, and σ_{sim} is the standard deviation over the simulations. KGE scores above -0.41 indicate that model simulations improve performance compared to assuming a constant flow corresponding to the average of the observed discharge time series (Knoben et al., 2019). To assess whether and by how much a specific change in model structure improves model performance for discharge, we use the KGE skill score (KGESS) as used by Knoben et al. (2020) and Van Jaarsveld et al. (2025). The KGESS compares the KGE score of a model run with a new setup against a model benchmark:

$$\text{KGESS} = \frac{\text{KGE}_{\text{model}} - \text{KGE}_{\text{bench}}}{1 - \text{KGE}_{\text{bench}}}, \tag{7}$$

where, $\text{KGE}_{\text{model}}$ is the KGE score of the new model run and $\text{KGE}_{\text{bench}}$ is the KGE score of a benchmark model run, which represents an intermediate step within the model improvement chain.

The Alpine region contains many reservoirs for water regulation, especially for hydropower production (Lehner et al., 2005; Brunner and Naveau, 2023). Their presence and how they are represented can influence model performance (e.g. Hanasaki et al., 2006; Abeshu et al., 2023). Therefore, we investigate how model performance for discharge varies between natural and regulated catchments. As not each catchment with a reservoir might be strongly regulated, we define a catchment's degree of regulation by dividing its total reservoir capacity by its mean discharge (in units of time). We only consider a catchment to be regulated when this degree of regulation exceeds 0.1 year (see Figure S3 in the Supporting Information). Another way to separate regulated from natural catchments is the water balance signature (WB), which describes the normalized deviation from a closed water balance assuming no long-term storage effects (Salwey et al., 2023). WB is defined as:

$$\text{WB} = \frac{Q}{P} - (1 - \frac{E_{\text{P}}}{P}), \tag{8}$$

where Q is the (observed) averaged discharge (mm/day), P is the precipitation (mm/day), and E_{P} is the potential evapotranspiration (mm/day) averaged over the catchment and study period. Positive values of WB indicate that a catchment's discharge is higher than expected. Assuming that the meteorological components could be well-estimated, such positive values could suggest that a catchment gains more water than what comes in through precipitation. Negative values of WB indicate that a catchment loses more water than just the potential evapotranspiration. Salwey et al. (2023) have shown that this metric can be used to identify catchments affected by hydropower production and water transfers. In addition, we performed a small evaluation evaluation indicating that there indeed exists a relationship between WB and the degree of regulation (see Figure S3 in the Supporting Information). However, other factors such as errors in the meteorological forcing or additional water input

from glaciers due to imbalance can also lead to strong water balance deviations, which can also affect model performance with respect to discharge. We thus use WB as a general metric to study the effect of such deviations in the water balance on model performance. In addition, we also use WB as an indication for water transfers, hydropower production and other water balance deviations in combination with catchment-based information on reservoirs.

In addition to discharge, we evaluate model performance for SWE by calculating spatial averages over different elevation bands (0-1000 m, 1000-2000 m, 2000-3000 m or the entire country) over Switzerland. Additionally, we evaluate it over Austria, which has not been used for model calibration and can therefore provide insights on how well the model generalizes to other regions. Then, we compute the average seasonal SWE cycle over these zones and compare it to the seasonal SWE cycles derived from different snow reanalysis products (OSHD, SNOWGRID; see Table 2). Furthermore, we calculate the KGE for SWE by comparing modelled SWE and SWE inferred from observations at specific measurement stations (NH-SWE; see Table 2).

We evaluate glacier simulations both in terms of their mass balance and their geometry evolution. Simulated mass balances were summed over the hydrological year (starting in October to facilitate comparisons with the observations of WGMS and GLAMOS; see Table 2) and compared to observed time series as well as to observed average geodetic mass loss estimates derived from GECM. To evaluate glacier geometry evolution, we both visually compare spatial patterns of glacier surface elevation changes against observations (GECM; see Table 2) and quantitatively evaluate long-term glacier changes. It is difficult to evaluate the long-term response of simulated glaciers to climate forcing given the relatively short study period. To address this problem, we repeat an experiment by Zekollari et al. (2020), in which glaciers are continuously forced (for 300 years) with the modelled mean mass balance for the period before 2018. The glaciers respond to this forcing by changing their shapes and they stabilize when they are in balance with the applied forcing. We then calculate the e-folding time scale (i.e. the time interval after which glaciers still had $1/e$ of their initial volume) and compare this to the estimates of Zekollari et al. (2020).

We evaluate soil moisture against the ESACCI soil moisture dataset. This evaluation is more difficult than the evaluation of other variables, since the satellite data are not directly comparable to our model output: they measure soil moisture in the top 5 cm of the soil, whereas we model moisture in the top 15 cm or 30 cm (see Table 2). Furthermore, the resolution of the observations is much coarser than the one of the simulations. Therefore, we resample our simulations to the resolution of the satellite data using spatial averages and only select locations with at least 50 percent of valid daily data over Evaluation period II. We calculate the Spearman rank correlation coefficient between the daily model simulations and the satellite observations to assess model performance, because this metric should be applicable despite the different soil moisture depths.

3 Results

3.1 Effect of meteorological forcing

Precipitation in all three meteorological forcing datasets used in this comparison correlates well with gridded precipitation station data from the APGD (see Figure 4). Precipitation from CERRA-CHELSA shows generally higher correlations with observed precipitation than the precipitation of the STANDARD and CHELSA input datasets (compare Figures 4A and B with

C). The correlation between the three datasets and the reference dataset varies across the Alps and is especially low in the Po Plain. All meteorological datasets show a positive bias in precipitation over the Alps and a negative bias over the Apennines (see Figures 4D, E, F). Around the Alps, the STANDARD and CHELSA forcings show slightly positive biases, whereas CERRA-CHELSA shows a slightly negative bias. Overall, CERRA-CHELSA and CHELSA have a smaller precipitation bias than the STANDARD dataset (mean absolute bias: CERRA-CHELSA: 0.4 mm/day; CHELSA 0.5 mm/day; STANDARD: 3 mm/day).

The choice of meteorological forcing has a strong effect on simulated discharge, but the effect is not uniform across catchments. Using the STANDARD forcing, we see generally better model performance for discharge over the Alps than the surrounding areas (see Figure 4G). However, locally there is poor model performance for discharge in certain catchments in the western Alps, southern Switzerland and eastern Austria. Forcing the model with CERRA-CHELSA leads to improved performance in these regions, as well as in southern Germany (see Figure 4I). In contrast, CERRA-CHELSA leads to reductions in model performance for discharge in eastern France, parts of Switzerland and western Austria. Using CHELSA leads to a slight worsening of model performance for discharge compared to runs with STANDARD or CERRA-CHELSA forcing, except in parts of the Alps (see Figure 4H).

In summary, we find that discharge simulations generated with the CERRA-CHELSA and STANDARD forcing datasets are generally better than those generated with the CHELSA dataset (see Figure 4J). As the precipitation of the CERRA-CHELSA dataset aligns better with the reference precipitation dataset than the STANDARD dataset (see Figure 4A, B, D and E), we performed all further analyses with the CERRA-CHELSA dataset.

3.2 Snow representation

SWE representation benefits to some degree from the proposed adjustments of the snow module (see Figure 5). The introduction of the snow transport scheme only improved snow representation at the highest elevations (2000–3000 m), where unrealistic snow towers were a major issue (compare the Transport run with the CERRA-CHELSA benchmark run in Figure 5C and F). Here, the snow transport scheme ensures that the snow is redistributed to lower elevations, where it subsequently melts away. ERA5-Land and CERRA-Land also show very high SWE values suggesting that these models also suffer from unrealistic snow build-up. Without calibration, further structural changes to the snow module (i.e. the seasonally varying DDF, exponential temperature dependence, and a rain-to-snowfall transition temperature range) did not improve performance against observational SWE stations (see Figure 5G and H). Averaged over elevation zones, however, these changes lead to an improvement of the SWE representation at higher elevations (e.g. for elevation zone 2000-3000m, the KGE compared to the reference products increases from 0.85 to 0.89 over Switzerland and from 0.53 to 0.57 over Austria). At lower elevations these changes lead to too low melt rates over Austria (for elevation zone 1000-2000m, KGE increased from 0.87 to 0.88 over Switzerland, but dropped sharply from 0.64 to 0.46 over Austria; compare Uncalibrated snow and Transport runs in Figure 5A, B, D, and E). These differences between elevation zones suggest that the model structure leads to varying melt rates with elevation: this is to be expected, as the temperatures at which snow starts to melt are reached later in the year at higher elevations, so that the same temperatures are combined with different values of the time-varying DDF which produces differing melt rates. Calibrating the

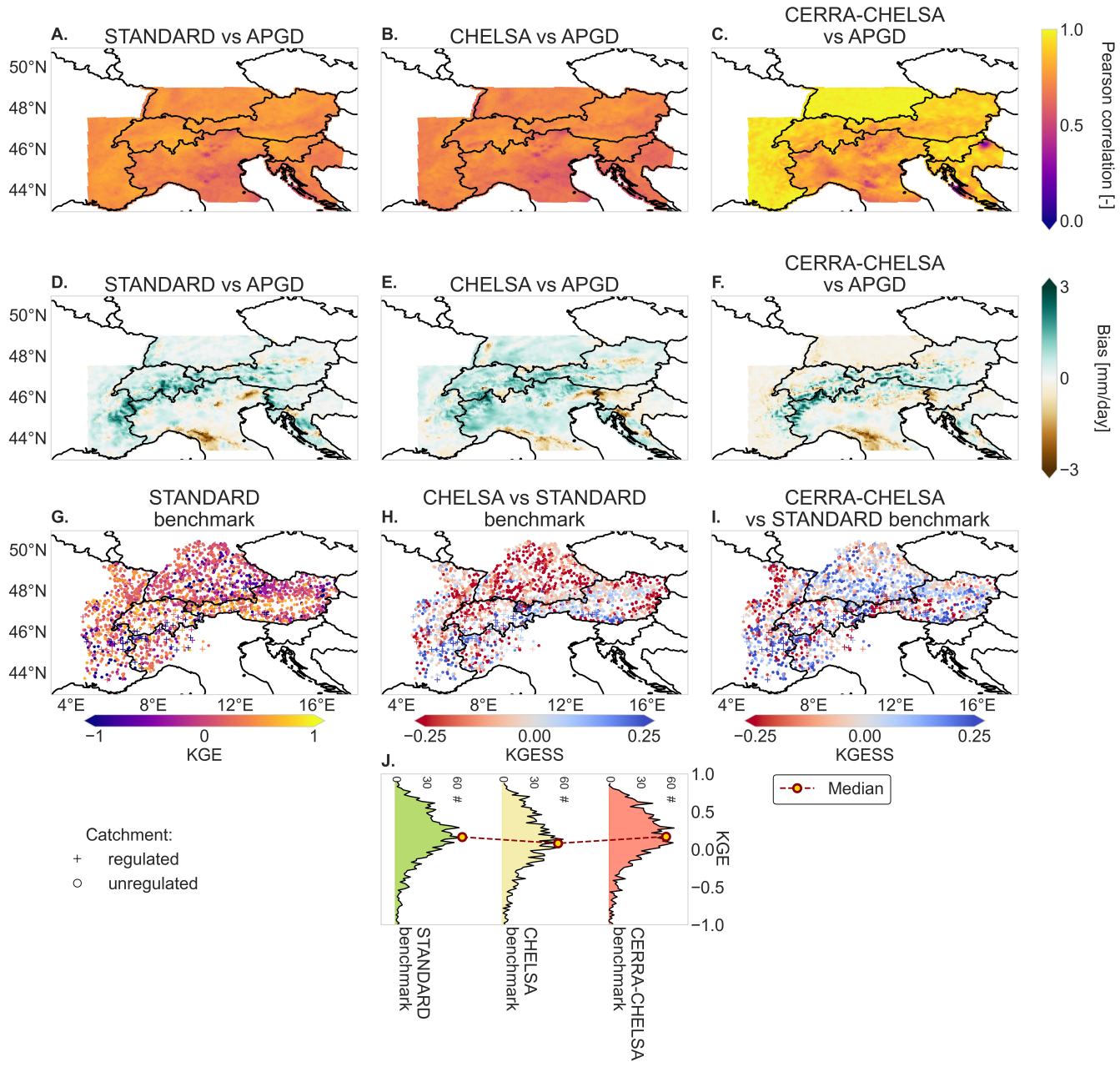

Figure 4. Differences in meteorological forcing datasets and the effect of their choice on model performance for discharge over the larger Alpine domain. The top row compares the correlation of precipitation in the (A) STANDARD, (B) CHELSA, and (C) CERRA-CHELSA datasets with observed precipitation (APGD, 2000–2008). The middle row shows the bias of mean daily precipitation against observed precipitation for the (D) STANDARD, (E) CHELSA, and (F) CERRA-CHELSA datasets. The bottom row compares discharge simulations generated with the different forcing datasets: (G) KGEs for the STANDARD benchmark run. KGESS for the CHELSA benchmark run (H) and the CERRA-CHELSA benchmark run (I) against the STANDARD benchmark run. (J) Ridge line plots showing the distribution of KGEs (in number of catchments) for the three benchmark runs. Note that roughly 6 percent of the stations have a KGE smaller than -1 and fall outside of the bounds in J.

DDF in the snow module against SWE has the best performance over Switzerland (KGE averaged over Switzerland against OSHD: 0.91), still with slightly too high melt rates over Switzerland (see Full model run in Figure 5A, B, and C). In Austria, the Calibrated snow run is, while slightly less accurate than in Switzerland, the most accurate of the presented model runs (KGE averaged over Austria against SNOWGRID: 0.80), even though no Austrian SWE data was used for calibration (see Figure 5D, E and F). This demonstrates that the snow module is generally transferable to other regions. The comparison of our model runs against SWE estimates at measurement stations (see Figures 5G and H) confirms that SWE simulations profit from the introduction of an improved snow transport scheme and calibration.

Changes in SWE performance are reflected in the performance of discharge simulations (see Figure 6). Snow transport improves discharge simulations in the highest parts of the Alps, but has hardly any effect outside of the mountains (see Figure 6B). Structural changes to the snow module lead to an improvement of discharge simulations in catchments at the highest elevations, but a worsening in catchments at lower elevations (see Figure 6C). Finally, calibrating SWE improves discharge representation in most of the catchments that worsened from the structural changes, with a few exceptions in the Alps (see Figure 6D). Figure 7A, B, and C illustrate that the changes in the snow module mainly improve model performance for discharge in catchments with high snowfall fractions, whereas in catchments with low snowfall fractions the changes are negligible or slightly negative. Still, some catchments with high snowfall fractions experience decreases in performance for discharge: these decreases mostly happen in the presence of reservoirs and/or negative values of WB (median KGESS for catchments with more than 30% snowfall compared to CERRA benchmark: regulated: 0.017; unregulated: 0.080).

3.3 Glacier representation

The new glacier module captures the general behaviour of glaciers. Both spatial patterns of glacier elevation changes and time-mean patters of mass balances are roughly reproduced (see Figure 8A, B, C, D, E, and F), even though the model shows biases for individual glaciers in mass balance (see Figure 8A, B and G) or retreat (see rapid retreat of Mer de Glace in Figure 8D). Generally, mean geodetic mass balances are underestimated (see Figure 8G). Figure 8H shows the results from the equilibrium-experiment based on Zekollari et al. (2020). Most glaciers reach equilibrium over time, although there is significant committed mass loss. Overall, we end up with around 40% committed mass loss in 2018, which was also found by Zekollari et al. (2020). Our modelling scheme thus captures the general behaviour of long-term glacier responses, with glacier retreat adjusting to a new steady-state condition. However, we see deviations for individual glaciers, with a larger glaciers generally retreating slightly less and smaller glaciers retreating slightly more than the reference dataset. In conclusion, our model evaluation shows that the new glacier module works reasonably well for the total or groups of alpine glaciers (see Figure 8G and H), while it can be significantly biased at the scale of individual or small glaciers (see Figure 8G, H).

The addition of the glacier module mainly improves discharge simulations in highly glacierized catchments (see Figure 6E and Figure 7F). The positive effect of glaciers is much less visible in catchments with a small glacier or snowfall fraction (see Figure 7E). Furthermore, our results indicate that in certain regions, especially around the heavily-regulated Rhône river in south-western Switzerland, discharge performance can decrease with the addition of glaciers (see Figure 6E). Such perfor-

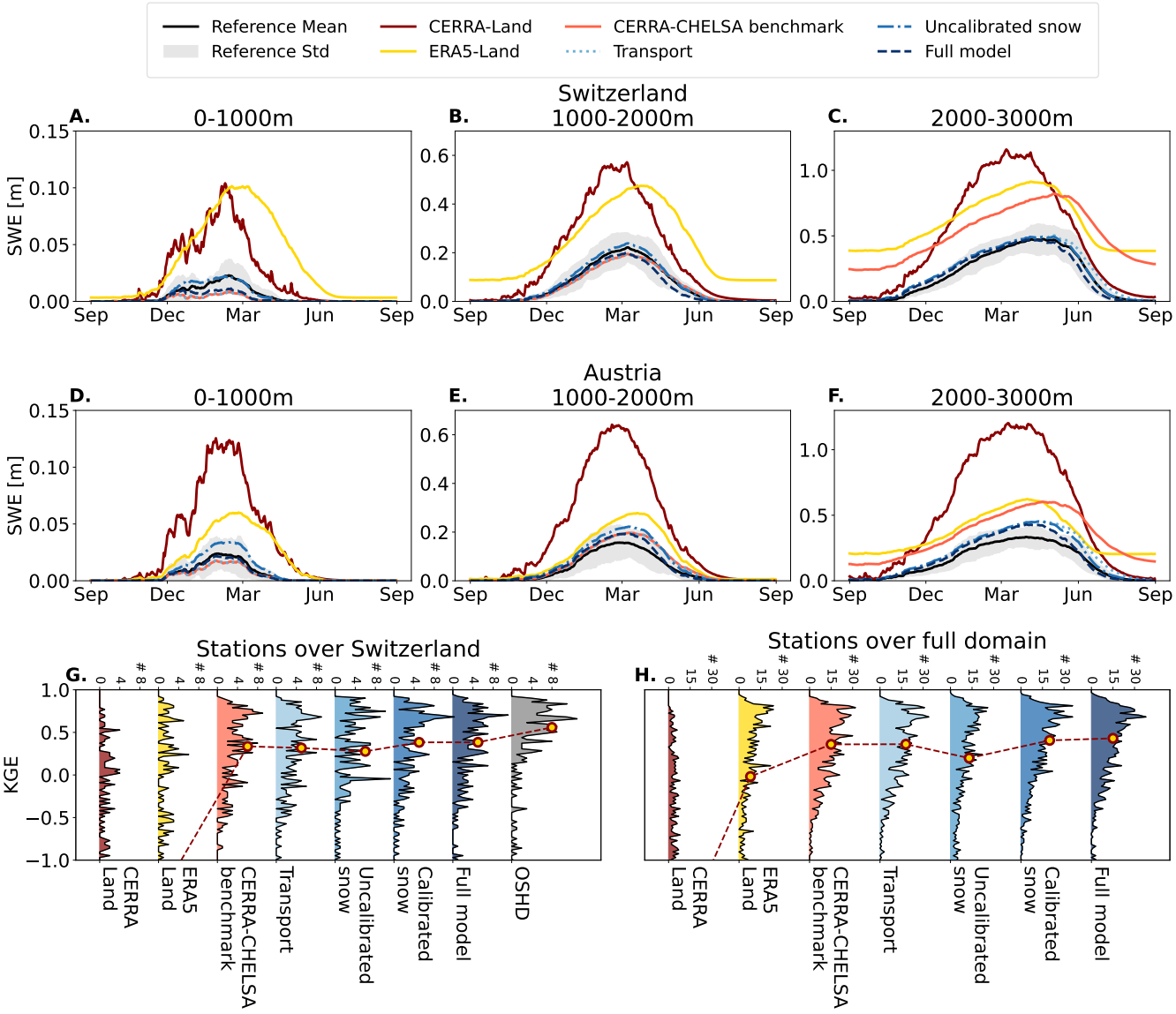

Figure 5. Snow representation for different elevation zones and regions. Average snow climatology over Switzerland (top row) and Austria (middle row) for grid cells at elevations between 0-1000 m (A and D), 1000-2000 m (B and E) and 2000-3000 m (C and F) derived from the gridded simulations and reanalysis products. The reference products are the OSHD reanalysis for Switzerland and the SNOWGRID product for Austria. Note that for simplicity we only show the Full model run instead of the Calibrated snow run, as these have the same snow module configuration. The bottom row shows KGEs of SWE time series at different measurement stations of Fontrodona-Bach et al. (2023b) over Switzerland (G; 251 stations) and the full Alpine domain (H; 1047 stations) for different model runs and reanalysis products. For G and H, note that roughly 3 to 10 per cent of stations have a KGE < -1 (>30 per cent for CERRA-Land and ERA5-Land).

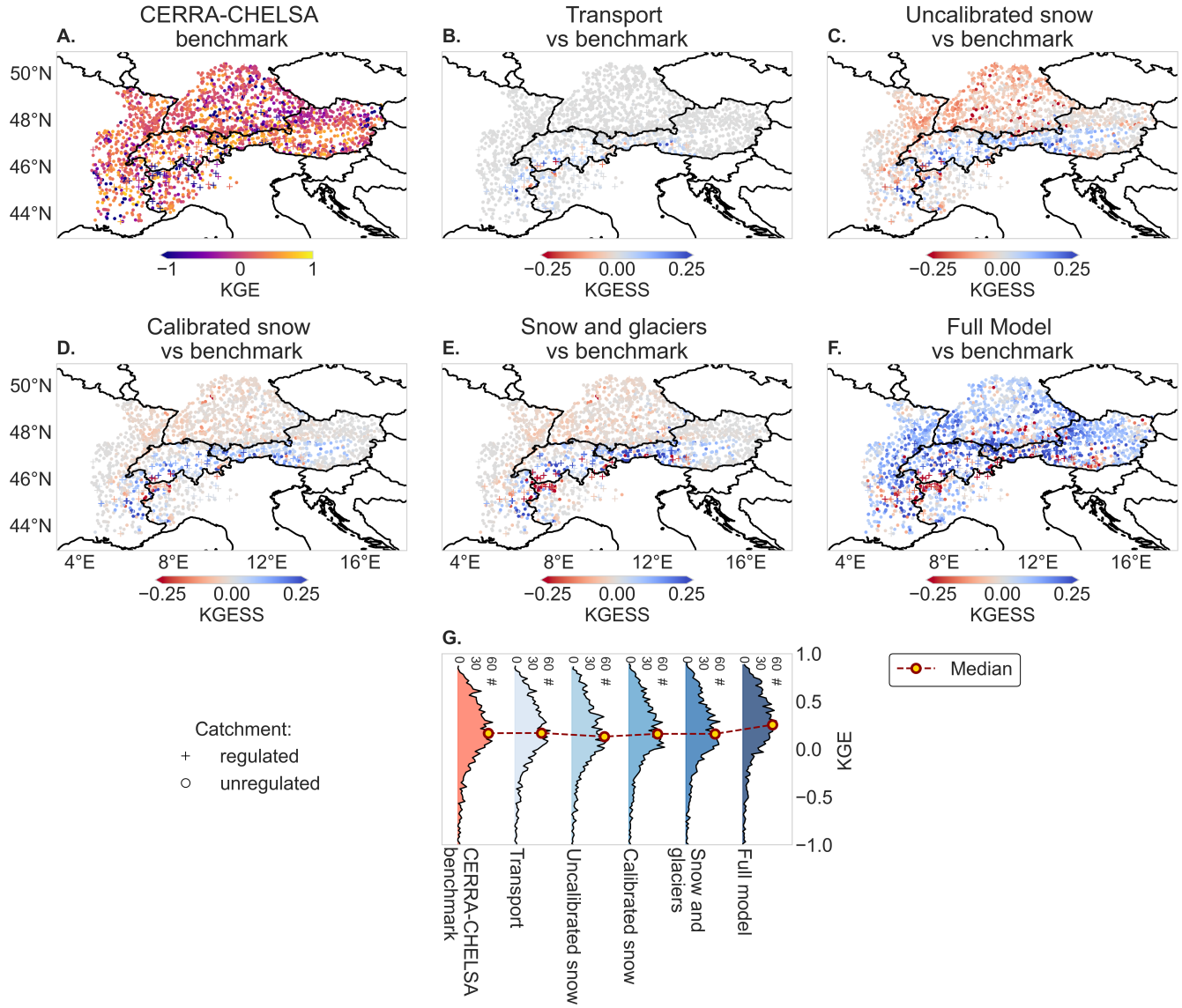

Figure 6. Model performance and its changes for discharge at the measurement stations for different adjustments in the snow routine. (A) Absolute KGEs for the CERRA-CHELSA benchmark run. Changes in performance (KGESS) with respect to the CERRA-CHELSA benchmark run for different model configurations: (B) Transport, (C) Uncalibrated snow, (D) Calibrated snow, (E) Snow and glacier and (F) Full model (including soil thickness change). (G) Distribution of KGE scores for the different model runs across catchments. Note that roughly 6 percent of stations have a KGE smaller than -1 and fall outside of the bounds.

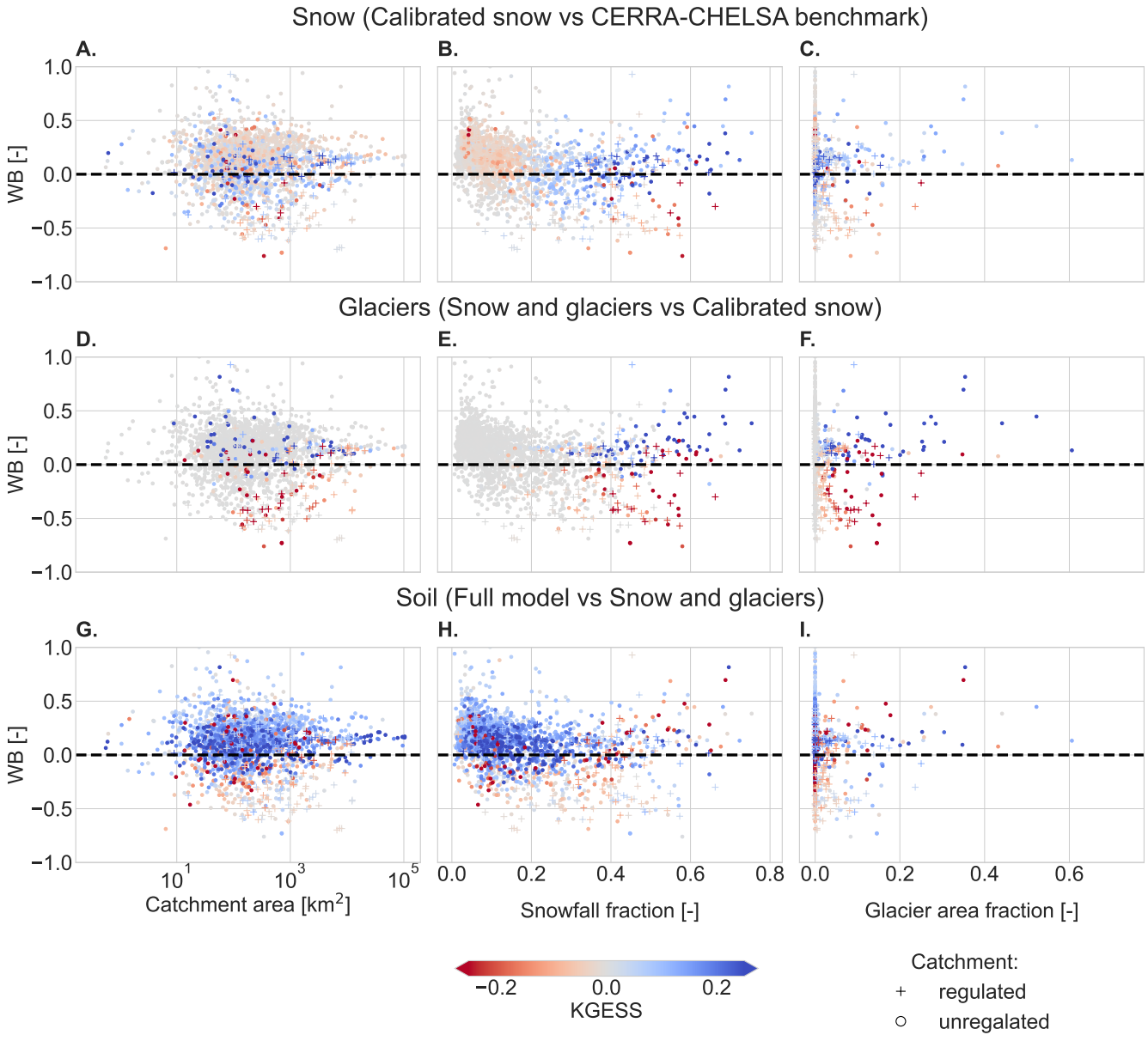

Figure 7. Model performance change for discharge (KGESS) at the measurement stations after introducing different adjustments to the model (indicated with the colours). The model performance changes are compared to catchment characteristics. Top row: effect of the combined snow changes on discharge simulations (difference between the Calibrated snow and the CERRA-CHELSA benchmark runs); middle row: effect of the introduction of glaciers (Snow and glaciers vs. Calibrated snow); and bottom row: effect of the changes made to the soil (Full model vs. Snow and glaciers). (A), (D) and (G) show the dependence of model performance changes for discharge on WB and catchment area; (B), (E) and (H) on WB and snowfall fraction; and (C), (F), and (I) on WB and glacier area fraction. Note that roughly 3 percent of the stations have a WB larger than 1 and fall outside of the figure bounds.

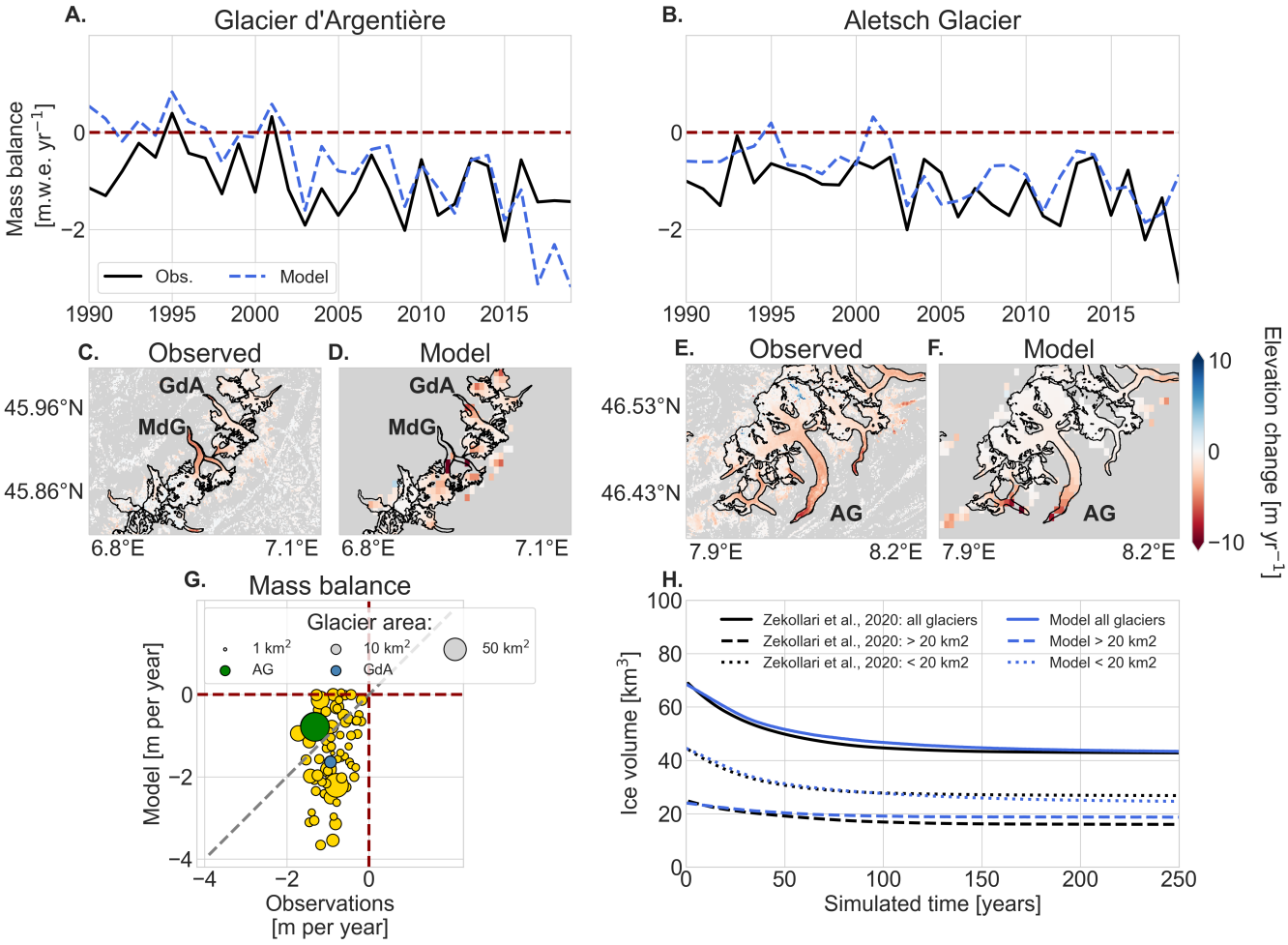

Figure 8. Evaluation of the glacier scheme newly integrated in the hydrological model. Top row shows annual time series of simulated against observed mass balance in meters water equivalent per year for Glacier d'Argentière (A) and Aletsch Glacier (B) (observations for Aletsch: Glacier Monitoring Switzerland (GLAMOS) (2022); Argentière: World Glacier Monitoring Service (WGMS) (2023)). Middle row shows the elevation change over 2010–2019 for the region around Glacier d'Argentière (C. observations; D. model) and around the Aletsch Glacier (E. observations; F. model). Observations from Hugonnet et al. (2021a), outlines from GLIMS Consortium (2018). (G) Average elevation change (2010–2019) of glaciers against observations from Hugonnet et al. (2021a). (H) Comparison of the evolution of glacier volume over time under this continuous forcing with their mean mass balance from 1990-2018 to modelled estimates from Zekollari et al. (2020). AG: Aletsch Glaciers; GdA: Glacier d'Argentière.

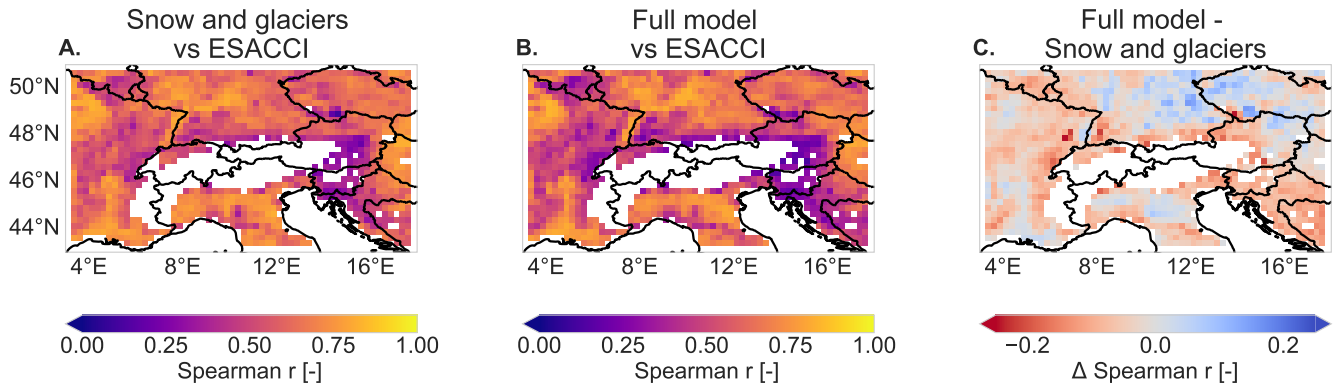

Figure 9. Comparison of soil moisture simulations against the ESACCI satellite observations. We show the Spearman correlation between the Snow and glacier (A) and Full model runs (B) against the ESACCI observations. (C) Difference in Spearman correlation coefficients between the Full model and the Snow and glaciers run. White areas indicate regions which had too many days without data (more than 50% of the time).

mance decrease generally occurs in catchments with a negative WB or reservoir regulation (see Figure 7F; median KGESS in catchments with >5% glacier cover compared to Calibrated snow: regulated -0.19, unregulated: 0.15).

3.4 Soil partitioning

Model performance for soil moisture varies over the domain (see Figure 9A and B), with generally higher performance in flatter low-elevation areas (such as the Rhone Valley, the Po Plain or the Rhine Valley) and lower performance over hilly or mountain areas. Implementing the soil changes has a mixed effect on performance: generally, it improves performance in areas where the model was already performing well (i.e. the lower flatlands) and worsens performance in locations with lower model performance (see Figure 9C).

Over the entire domain, the changes made to the soil module, i.e. to the runoff partitioning, generally increase model performance for discharge (see Figures 6F). The improvements in discharge performance are strongest in catchments with lower snowfall fractions (<0.3; see Figure 7H) and are generally independent of the catchment area (see Figure 7G). Discharge performance increases are stronger in unregulated catchments (see Figure 7G, H, and I; median KGESS compared to Snow and glaciers: regulated: 0.03, unregulated: 0.09).

3.5 Evaluation of new model setup

Our new model setup, which includes updated snow, glacier and soil modules, leads to general performance increases in streamflow simulations compared to the existing PCR-GLOBWB 2.0 setup, with performance depending on catchment characteristics (see Figure 10 and the summary provided in Table 3 and Figure 11). Catchment area and WB are major controls of absolute model performance for discharge in terms of KGE, which is highest in large and natural catchments, in which the wa-

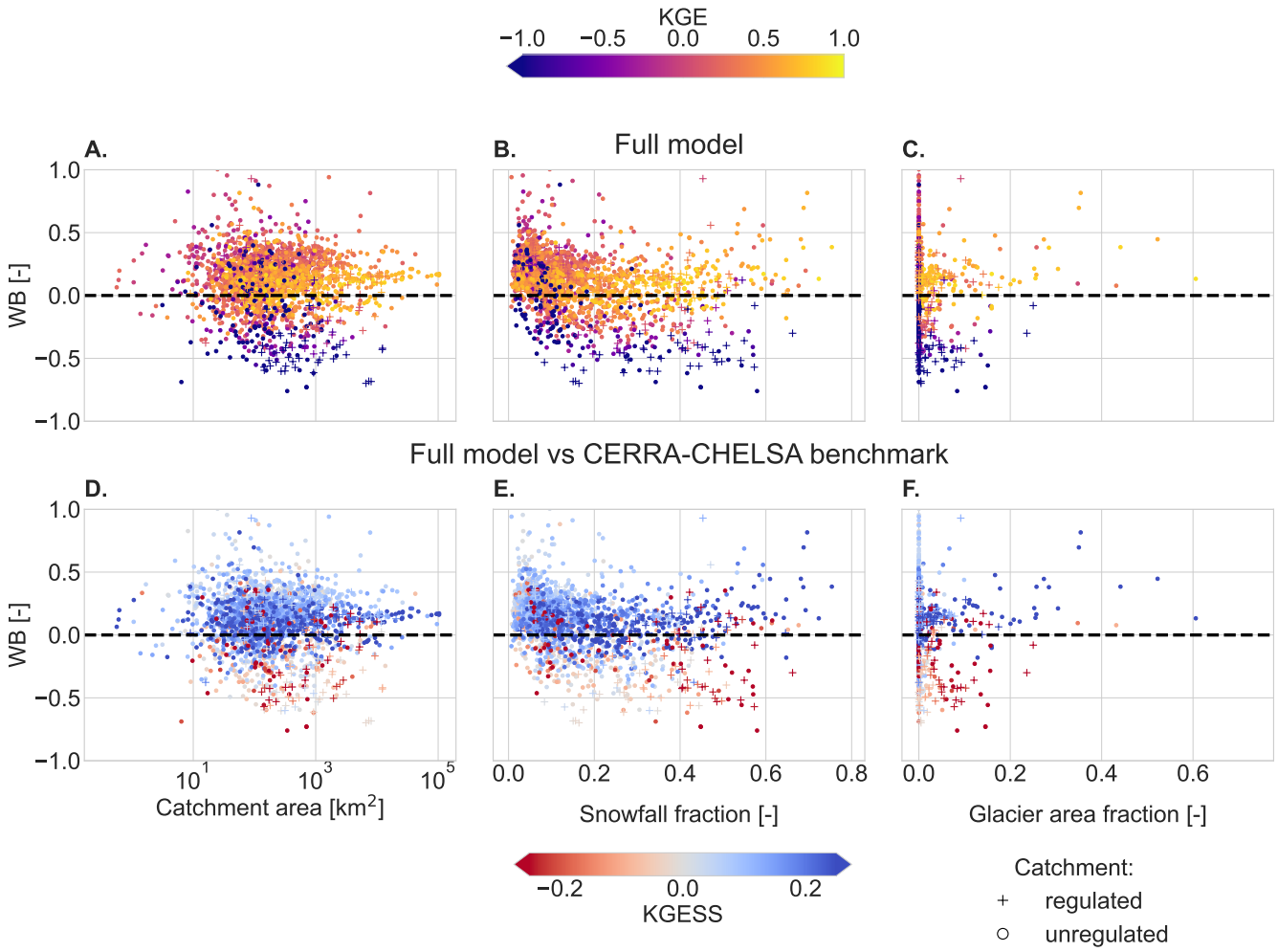

Figure 10. Model performance of the Full model run for discharge (top row, showing KGE) and the total model performance changes for discharge compared to the CERRA-CHELSA benchmark run (bottom row, showing KGESS) in the evaluation catchments in relation to different catchment characteristics: WB and catchment area (A, D), WB and snowfall fraction (B, E), and WB and glacier area fraction (C, F). Note that roughly 3 percent of the stations have a WB larger than 1 and fall outside of the figure bounds.

530 ter balance is nearly closed (see Figure 10A). Similarly, the model performs well in snow covered and glacierized catchments, where the model additions lead to a substantial improvement in model performance for discharge (see Figure 10B, C, E, and F). In contrast, the model adjustments can decrease performance in catchments with a negative WB (see Figure 10D, E, and F; median KGESS of Full Model against CERRA Benchmark: positive WB: 0.09, negative WB = 0.02).

Table 3. Summary of the effect of the different model implementations on discharge simulations. The symbols have the following meaning: ↑: improvement, ↑↑: large improvement, ↓: worsening, ↓↓: large worsening, -: no effect, ∼: substantial but mixed effect.

Catchment type	Meteorological forcing	Snow module	Glacier module	Soil partitioning
Small catchments	∼	-	-	↑
Large catchments	∼	-	-	↑
Natural rainfall dominated catchments	∼	-	-	↑↑
Natural snowfall dominated catchments	∼	↑	↑	∼
Natural glacierized catchments	∼	↑	↑↑	-
Regulated catchments	∼	↓	↓	-

4 Discussion

4.1 Evaluation and recommendations for further model development

Our new model setup generally led to increased model performance compared to the old setup: structural and parameter changes applied to the snow, glacier and soil modules improved both SWE and discharge simulations (Figures 6, 7, 5, and 10). This highlights the importance of improving process representation in hyper-resolution modelling efforts.

Updating the runoff partitioning in the soil leads to a clear improvement in the simulation of discharge in natural catchments (Figures 6F and 7G, H, and I; column "Soil partitioning" in Table 3), with rainfall-dominated catchments (≈ less than 30% snowfall) benefiting the most (Figure 7G, H). This major increase in model performance is caused by a modest change in the soil parameters, supporting the suggestion that the move to hyper-resolution requires careful review of parameterizations in LHMs (Hoch et al., 2023). Still, absolute performance in smaller rainfall-dominated catchments remains relatively limited (Figure 10A and B). In addition, soil moisture representation did not improve in regions that already had poor performance (Figure 9C and D). Please note that the reference product used for computing model errors can have its own biases (Dorigo et al., 2015, 2017) and has a much coarser spatial resolution than our model, so error estimates might not be entirely representative. Our results therefore suggest that there is a need for more representative ways to compare soil moisture simulations and observations. Still, the representation of soil moisture and fast discharge responses needs to be further improved if LHMs are supposed to be applicable at smaller spatial scales. These further improvements can come from two directions: first, soil heterogeneity could be included in models in more detail. For example, Van Jaarsveld et al. (2025) suggested that sub-grid scale land cover variability could still be important in hyper-resolution modelling and we hypothesize that this would also improve the representation of spatial variations in discharge behaviour. However, including this information would come at the cost of increased computational demands. Second, more explicit consideration of hill-slope processes such as preferential flow in the model structure (e.g. Rahman and Rosolem, 2017; Gharari et al., 2019; Fan et al., 2019) could lead to further improvements in the simulation of flashy runoff responses.

Discharge simulations benefit from the structural changes made to the snow module mostly in snow-dominated catchments (Figures 6C and 7E; column "Snow module" in Table 3). Similarly, Girons Lopez et al. (2020) found for the catchment model HBV that exponential melt combined with seasonally varying DDFs improved discharge simulations, whereas the rain-to-snow transitions improved SWE representation but led to slightly poorer results for discharge. However, whereas our discharge simulations in most snow-dominated mountain catchments improved due to these structural changes, we noted a slight decrease in discharge performance at lower elevations where snow contributions are less important. One reason for why the structural changes to the snow module are better suited for catchments in mountainous terrain might be differences in dominant snow processes between high and low elevations, such as the frequency of rain-on-snow events. Since Magnusson et al. (2014) built their snow model for alpine Switzerland, they might have prioritized representing melt patterns at higher elevations with thick snow cover over patterns in flatter terrain with only limited snowfall. For example, our scheme does not explicitly include melt due to liquid precipitation. Another reason for this slight decrease in discharge performance at lower elevations could be related to our choice of regionally-averaged DDFs, since in reality these DDFs show smaller-scale variability in space (e.g. with aspect, albedo, elevation, landcover, vegetation,...) (e.g. Kuusisto, 1980; Rango and Martinec, 1995; Hock, 2003; Ismail et al., 2023). Such variability is not accounted for in our model since we focused on regionally valid parameterizations instead of local solutions. However, ignoring this variability could lead to biases in SWE representation in specific locations. More elaborate snow module formulations, such as parameterizations that include aspect (e.g. Immerzeel et al., 2012) or radiation (e.g. Hock, 1999), could increase our ability to capture more detailed spatial melt patterns. Such approaches can become feasible now that higher spatial resolutions can resolve slopes and represent vegetation cover and land use in a more detailed way. However, they come at the cost of increased model complexity and a larger number of input variables. In any case, the slight decrease in discharge performance at lower elevations was reduced via the calibration of modelled SWE against a regional SWE reanalysis product, leaving us with a good overall discharge and SWE simulation performance (compare Figure 6C and D). The improved discharge representation after calibration highlights the performance gains that can be achieved by including more regional data into LHMs. However, regional calibration is only possible due to the comparably high quality of observational and reanalysis data in Switzerland. Many other regions around the globe continue to face a lack of observations of water balance components (Wilby, 2019), which challenges accurate regional calibration. Furthermore, the slightly reduced performance in SWE representation over Austria (which was excluded from calibration) compared to Switzerland already indicates that parameters can vary between regions (compare Figure 5D, E, F with Figure 5A, B, C). Highly-detailed calibration in one region might give a false sense of accuracy when applying the set-up outside of the calibration region. Regionally valid datasets must thus be chosen with care.

Adding a glacier module led to general improvements in discharge simulations in glacierized catchments in the Alps in catchments with a near-zero or positive value for WB (i.e. more observed discharge than is "expected"; see Figure 7F; column "Glacier module" in Table 3). The effect of glaciers on general discharge performance further downstream remains limited (see Figure 6E), although this effect might be larger for certain months or seasons (Wiersma et al., 2022). Aside from discharge, glacier mass balances and spatial patterns in elevation changes are also reasonably well represented (see Figure 8). Still, we

note that glaciers can show significant biases in the mass balance (see Figure 8G) and in the responses (see Figure 8H). Differences in long-term responses compared to previous experiments in which ice dynamics were incorporated might be partly explained by the Δh-parameterization, which ignores potential delays in glacier response to mass changes (Seibert et al., 2018a). Further biases in both mass balance estimates and responses are likely related to the relatively coarse spatial resolution of our model compared to the size of individual glaciers, which makes it more difficult to accurately describe patterns in melt or snow accumulation for small glaciers consisting of only a few grid cells. Melt representation of individual glaciers could be improved by resolving glaciers at higher spatial resolution than the resolution simulated here, for example by including elevation zones (Seibert et al., 2018a). Resolving glaciers at higher spatial resolution could also lead to even more realistic glacier retreat, since the Δh-parameterization was originally designed for higher spatial resolutions (Huss et al., 2010) and is likely less accurate at the coarser 30-arcsec model grid. Further gains in glacier representation could be realized by improving glacier accumulation estimates by continuing developing the snow component, or by improving spatial melt patterns by varying the glacier DDF in space (analogously to what we suggested for snow).

While the improvements in process representation generally lead to an increase in model performance for discharge, there are regions where the performance of discharge simulations decreases after implementing the structural and parametric changes (see Figure 10D, E, and F). Our analysis shows that such performance decreases are common in catchments with negative values of WB, which can be related to issues with the meteorological forcing, glacier melt estimates or with the representation of water abstractions or (hydropower) reservoirs in the model. Indeed, the strongest negative performance changes after the introduction of snow and glaciers often occur in catchments with negative WB or regulation (see Figure 7B and F; row "Regulated catchments" in Table 3). In the Alps, significant glacier or snow melt occurs above hydropower reservoirs (e.g. in Switzerland, 4% of all hydropower is related to glacier mass loss (Schaefli et al., 2019)). Hydropower severely changes flow seasonality (Arheimer et al., 2017), essentially decoupling observed streamflow from snow and ice melt. Accurate representation of river regulation and hydropower reservoirs is thus important, but LHMs appear to have difficulties with modelling streamflow in these regulated catchments (e.g. Veldkamp et al., 2018; Tu et al., 2024). These issues might become even more apparent at hyper-resolution, because a.) small rivers are now represented and might be affected by reservoirs that are not represented in the model, and b.) reservoir schemes in LHMs were developed to represent regulation behaviour on a coarse grid (e.g. they mimic the combined effect of all reservoirs within a 50 by 50 km grid cell) and likely need an update when moving to higher spatial resolutions (Shin et al., 2019). However, even if more reservoirs were introduced to the model, accurate discharge modelling in regulated systems remains challenging due to limited data on operation strategies and regulations (Turner and Voisin, 2022). Thus, both improving the representation of human water management in models and collecting new data should remain an active area of research.

Besides model structure and parameterization, the meteorological reanalysis forcing has a major influence on model performance for discharge (see Figure 4G, H, and I; column "Meteorological forcing" in Table 3). This implies that improvements in simulation performance that can potentially be achieved by structural changes are bounded by the quality of the meteorological forcing dataset, especially in a non-calibrated model set-up. Replacing the STANDARD meteorological forcing product with CERRA-CHELSA led to an overall increase in correlation with the reference dataset (compare Figure 4A and C), but

showed mixed performance changes for discharge over the model domain (compare Figure 4G and I). This apparent inconsistency could be related to uncertainties affecting our precipitation evaluation and the LHM. Our precipitation evaluation is likely affected by uncertainties in the reference APGD. Although the APGD is based on a particularly dense network of direct observations, it still has uncertainties related to precipitation undercatch (for which it was not corrected), and in regions with less dense station coverage (Isotta et al., 2015). Our evaluation of the precipitation products should therefore be interpreted as an indication of consistency with interpolated surface measurements but not necessarily as an indication of absolute performance. Collecting new, better, and more observations of meteorological variables over mountainous regions should thus remain a priority. The limited performance for discharge in small flashy catchments, as was pointed out before, might further prevent the direct translation from more accurate precipitation input to improved discharge performance. The ever-present uncertainty in forcing datasets and their evaluation also highlights the advantage of the calibration strategy we proposed and applied here, which focuses on processes like snow or glaciers rather than on discharge. Calibrated models can compensate for meteorological uncertainty and model deficiencies (Refsgaard and Storm, 1996). Including the representation of processes such as snow or glaciers into the calibration procedure avoids some parameter equifinality and preserves internal relationships (e.g. Duethmann et al., 2014; Finger et al., 2015), which facilitates identifying where models need improvements. Finally, as presented in the Methods section, the STANDARD and CHELSA meteorological forcing datasets were derived using statistical or physical/heuristic downscaling, whereas the CERRA-CHELSA dataset is based on a dynamically generated regional reanalysis product. None of these methods used to create the hyper-resolution forcing datasets led to general improvements in simulation performance across the full domain. Regional differences between the datasets, observational uncertainty and deficiencies in hydrological model structure did have a pronounced effect on model performance for discharge.

4.2 Applicability and limitations

Our new model setup generally shows a better representation of discharge than the existing hyper-resolution version of PCR-GLOBWB 2.0 (see Figure 10 D, E, and F). Absolute discharge performance is generally satisfactory, especially at larger spatial scales and in natural rivers (see Figure 10 A, B, and C) and spatial patterns can be resolved in high detail (see Figure 11 A and B). Applications of such hyper-resolution LHMs should thus make use of these strengths by studying spatial patterns at larger spatial scales. Still, the model shows poor discharge simulation performance in some catchments, especially in smaller catchments at lower elevations and in rivers with a negative WB. Users should thus be aware of such limitations when looking at individual catchments. While the hyper-resolution simulations might not be completely accurate in certain places, they can still be valuable in helping to complete the spatial picture and enable studying spatial variability (Seibert et al., 2018b). This is especially true in regions where no observational data are available. There, these LHMs provide a first order estimate of river discharge, without requiring local calibration.

In addition to discharge, LHMs also provide information on other variables such as SWE. Although there exist many local or national snow reanalysis products that are more accurate than our SWE simulations, semi-global products are often unrepresentative for mountain regions, since their coarse-resolution makes it difficult to represent complex terrain (Mortimer et al., 2020; Mudryk et al., 2025). Our analysis shows that our higher resolution model setup outperforms coarser reanalysis products such

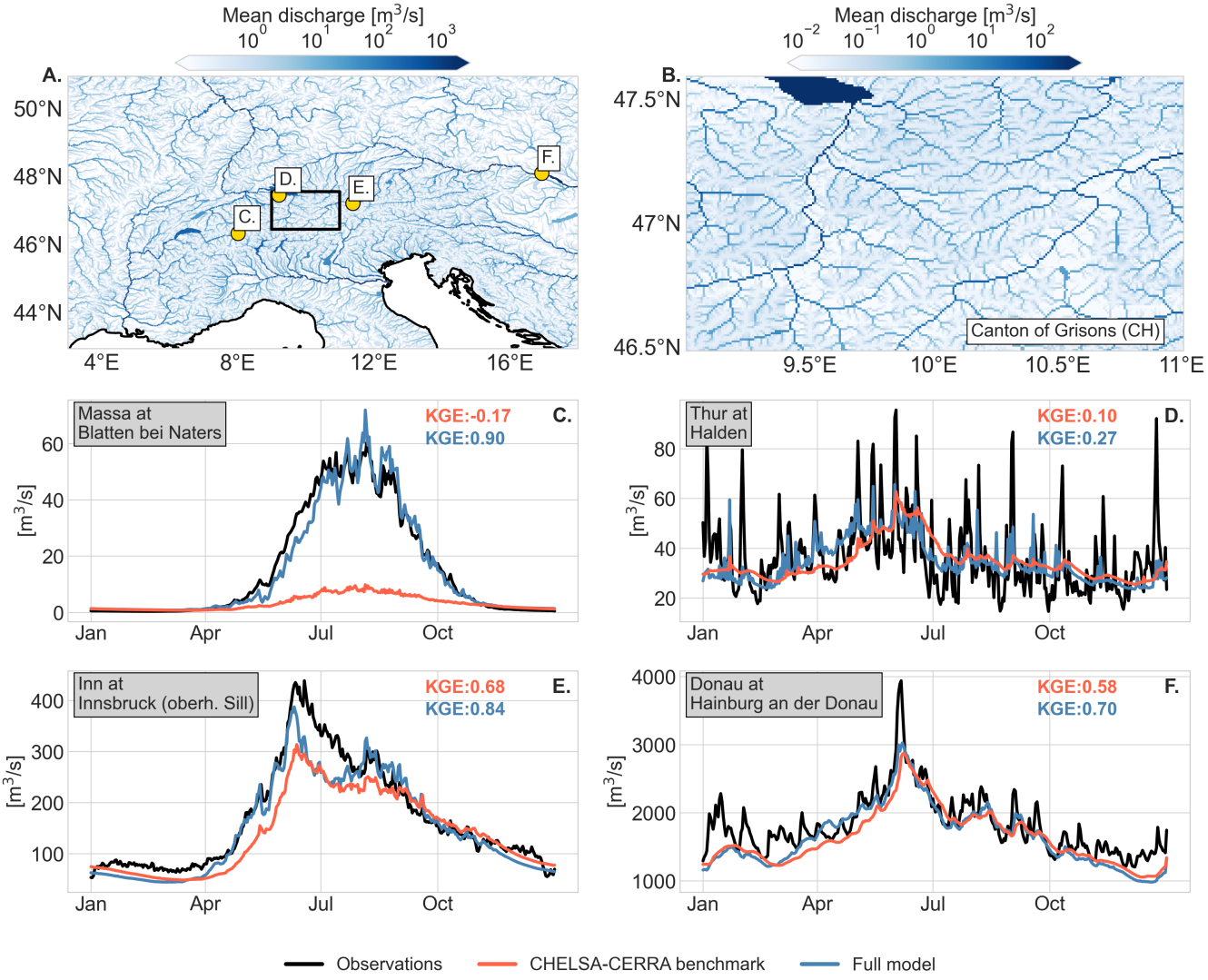

Figure 11. Illustration of the discharge simulations of the model with examples from hydrologically diverse river stations. (A) Average discharge (1990–2019) over the model domain. (B.) Average discharge (1990–2019) over the Swiss canton of Grisons (inset shown in A. in black). The bottom two rows show climatologies (2010–2019) of observed vs. simulated discharge for the glacier-dominated Massa at Blatten (C), the rain-dominated Thur at Halden (D), snow-dominated Inn at Innsbruck (E), and the larger river Danube at Hainburg (F).

as CERRA-Land and ERA5-Land, even without calibration (Figure 5), but does not quite reach the performance of national-scale reanalysis products such as the OSHD product (Figure 5G). Hyper-resolution LHMs could thus provide additional SWE products that bridge the quality gap between national and global datasets.

How suitable is the model for climate impact assessments – a typical application that hydrological models? While many larger-scale climate change impact assessments still use temperature-index based approaches, including studies focusing on snow (e.g. Kraaijenbrink et al., 2021; Yang et al., 2022) or glaciers (e.g. Kraaijenbrink et al., 2017; Van Tiel et al., 2023; Hanus et al., 2024), there is discussion about the robustness of degree-day approaches under climatic change (Carletti et al., 2022). A short evaluation, which explicitly addressed model transferability, shows that model performance for discharge and SWE remains mostly consistent over the warmer and colder evaluation periods compared to the reference period, although the temperature changes over the study period are limited (Figure S1 in the Supporting Information). Furthermore, we specifically aimed for increased robustness of parameters by focusing on SWE or glacier melt instead of streamflow during calibration. Sleziak et al. (2020) showed that the stability of hydrological model parameters such as DDFs across periods with different climate conditions can be increased by emphasizing snow representation during calibration. All of this suggests a reduced sensitivity of our model to increases in temperature. Still, the general caveats of the model remain applicable when extrapolating into the future. For example, the glacier module shows a realistic dynamic response to climate over larger regions, but not necessarily for each individual glacier (Figure 8). Users should be mindful of this uncertainty in individual glacier responses to climate change. Over time, continued glacier retreat will reduce glacier melt contributions to discharge, since globally most glaciers have either already surpassed their peak in glacier melt contributions or will do so in the coming decades (Huss and Hock, 2018). This will lead rivers to shift towards more snow-dominated regimes (Farinotti et al., 2012). Further into the future (i.e. near the end of the century) the exact response of glaciers will thus be less important.

In summary, the presented model setup adds value to regional scale hydrological studies, focusing on general patterns and on larger rivers which are nearly natural to moderately regulated. Caution should be used when interested in individual, small or heavily-regulated rivers, where catchment-scale models likely show better performance. Still, in regions with limited data availability, large scale model runs with global parameterization might remain the best option available even for such small rivers.

5 Conclusions

Hydrological processes in mountain regions control the supply of water to dependent downstream regions and are therefore important well beyond the fringe of mountains. An accurate representation of these processes in LHMs is thus crucial, especially now that these models are moving towards higher spatial resolutions. In this paper, we proposed different model adjustments to the well-known and frequently used PCR-GLOBWB 2.0 model in its high-resolution version to make large-scale models better suited for applications in mountain regions. We studied how meteorological forcing and an improved representation of snow, glaciers and soil affect discharge in the larger Alpine region. Based on our model evaluation, we conclude that:

- Meteorological forcing is uncertain over Alpine regions and different forcing datasets lead to major differences in hydrological model performance for discharge. These performance differences vary spatially between different forcing datasets.

- Discharge simulations forced by a reanalysis product using high-resolution atmospheric dynamics at 5.5 km (CERRA-CHELSA) did not consistently outperform simulations with the other forcing products which use coarser atmospheric dynamics at 31 km (STANDARD and CHELSA; Hypothesis 1 is not supported).

- An improved representation of snow and glaciers improves SWE and discharge simulations in high Alpine catchments where natural flow conditions dominate (Hypothesis 2 is supported).

- The introduction of better runoff partitioning in the soil leads to an improvement of discharge simulations in smaller rain-dominated catchments by increasing their flashy response to rainfall peaks. However, these catchments still show overall weak model performance for discharge, suggesting that there is more to gain from improved routing in the soil (Hypothesis 3 is supported).

- A major control limiting model performance for discharge in Alpine catchments are strong deviations from a closed water balance. Such deviations are sometimes related to hydropower plants or water transfers that are not accurately represented in the current model setup, which should be a priority of future model development.

- Finally, we present a new model setup with an improved representation of hydrological processes relevant in alpine regions, which is well suited to study regional and larger-scale streamflow and snow patterns in and around mountain regions. This new setup can be used to help quantify water resources and to study how these are impacted by human water use or climate in the Alpine region or around the world.

Code and data availability.

Code

The existing version of PCR-GLOBWB 2.0 is available on https://github.com/UU-Hydro/PCR-GLOBWB_model. The new model changes will be made available as a separate branch on https://gitlab.ethz.ch/gjanzing/PCR-GLOBWB_model/-/tree/alpine_model?ref_type=heads.

Data

We used meteorological data from several different sources. The downscaled STANDARD forcing over the model domain from (Van Jaarsveld et al., 2025) will be uploaded to the repository EnviDat upon acceptance of this manuscript. The source dataset W5E5v2.0 is available at (Lange et al., 2021) and the CHELSA-BIOCLIM+ climatologies are available at Brun et al. (2022a). Forcing from the CHELSA v2.1 can be downloaded from Karger et al. (2021a). This dataset on ENVIDAT will be updated to cover the full study period. Temperature from CERRA can be downloaded from Schimanke et al. (2021), whereas

the precipitation of CERRA-Land can be retrieved from Verrelle et al. (2022). The downscaled CHELSA-CERRA temperature data will be made available on EnviDat upon acceptance of this manuscript. Finally, precipitation from the APGD is available at Isotta and Frei (2013).

The MERIT Hydro DEM is available from Yamazaki et al. (2019). For our model domain, the upscaled version will be provided on EnviDat upon acceptance of this manuscript together with STANDARD forcing and PCR-GLOBWB 2.0 input. The upscaled MERIT Hydro DEM is also available at (Verkaik and Sutanudjaja, 2024).

Data on discharge, catchment area and reservoirs can be requested from the agencies or downloaded from the sources listed in Table A1. The DEM used for catchment delineation is the COPERNICUS DEM, which is available at European Space Agency and Airbus (2022).

The SWE OSHD dataset for Switzerland is available at Mott (2023) and the SNOWGRID dataset for Austria can be downloaded from GeoSphere Austria (2022). Further snow reanalysis products used are ERA5-Land, available at Copernicus Climate Change Service (2019), and CERRA-Land, available at Verrelle et al. (2022). Finally, estimates of SWE from station data can be downloaded from Fontrodona-Bach et al. (2023a).

Glacier volumes are available from Farinotti (2019). Remotely-sensed glacier elevation changes can be downloaded from Hugonnet et al. (2021b). Data on mass balances of individual glaciers can be retrieved from World Glacier Monitoring Service (WGMS) (2023) and from Glacier Monitoring Switzerland (GLAMOS) (2022). Glacier stability simulations were kindly provided by Harry Zekollari and are available upon reasonable request from the authors of Zekollari et al. (2020). Glacier outlines and glacier outlines and areas were retrieved from GLIMS Consortium (2018). Glacier percentage per catchment can be calculated from the Randolph Glacier Inventory (Pfeffer et al., 2014) and using the catchment outlines in Table A1.

Soil moisture data (v8.1) from the ESACCI can be downloaded from Dorigo et al. (2023).

Appendix A: Sources for streamflow, catchment shapes and reservoirs

Table A1. Agencies and databases as sources for data on streamflow, catchment shapes and reservoirs. 1: DEM used is the COPERNICUS DEM (European Space Agency and Airbus, 2022).

Country	Data Type	Sources	Link
Austria	Streamflow	Austrian Ministry of Agriculture, Forestry, Regions and Water Management	https://ehyd.gv.at/
	Catchment Shapes	Large-Sample Data for Hydrology and Environmental Sciences for Central Europe (Klingler et al., 2021)	
	Reservoirs	Austrian Ministry of Agriculture, Forestry, Regions and Water Management; Simmler (1961); Partl (1977)	https://www.bml.gv.at/
France	Streamflow	Ministry of the Environment, Sustainable Development and Energy (Banque HYDRO)	https: //www.hydro.eaufrance.fr/
	Catchment Shapes	delineated from DEM[1]	
	Reservoirs	Comité Français des Barrages et Réservoirs	https:www.barrages-cfbr.eu
Germany	Streamflow	Bavarian State Office for the Environment and the State Institute for the Environment Baden-Württemberg	https://www.lfu.bayern.de and https://www.lubw. baden-wuerttemberg.de
	Catchment Shapes	State Institute for the Environment Baden-Württemberg, delineated from DEM[1]	https://www.lubw. baden-wuerttemberg.de
	Reservoirs	Speckhann et al. (2020, 2021)	
Italy	Streamflow	Regional Environmental Agencies from Lombardia, Aosta, and Piemonte	https://www.arpalombardia.it; http://presidi2.regione.vda.it/ str_dataview_download and http://www.arpa.piemonte.it
	Catchment Shapes	delineated from DEM[1]	
	Reservoirs	AQUASTAT Geo-referenced Database on Dams; the Italian Ministry of Infrastructure; OpenStreetMap	https://www.fao.org/aquastat/ en/databases/dams; http://dati.mit.gov.it/catalog/ dataset/grandi-dighe-italiane
Switzerland	Streamflow	Federal Office for the Environment	https://www.bafu.admin.ch
	Catchment Shapes	Federal Office for the Environment	https://www.bafu.admin.ch
	Reservoirs	Federal Office for the Environment	https://www.bafu.admin.ch

Appendix B: Δ h-parameterization

The following detailed description of the Δ h-parameterization is based on Seibert et al. (2018a) and Huss et al. (2010).

Each glacier has a minimum surface elevation E_{min} and a maximum surface elevation E_{max} (both in meters above sea level). This topographic range of the glacier surface can be split into N elevation zones (in our case into 20 steps, with E_{min} and E_{max} rounded to the nearest multiple of 50 m). Each elevation zone E_i with $1 <= i <= N$ is then normalized to :

$$E_{i,\text{norm}} = \frac{E_{\text{max}} - E_i}{E_{\text{max}} - E_{\text{min}}} \tag{B1}$$

The surface elevation of a glacier responds to glacier mass loss in a specific way that depends on its surface elevation. Huss et al. (2010) provided an empirical relationship between the normalized elevation zone $E_{i,\text{norm}}$ and the normalized (unitless) change in water equivalent of the glacier ice within this elevation zone $\Delta h_{i,\text{norm}}$;

$$\Delta h_{i,\text{norm}} = (E_{i,\text{norm}} + a)^{\gamma} + b(E_{i,\text{norm}} + a) + c, \tag{B2}$$

with a, b, c, and γ as empirical coefficients. Huss et al. (2010) provided three different sets of values for these empirical coefficients based on the initial surface area of the glacier.

Next, we need to couple this theoretical unitless ice thickness change to the actual observed mass loss. This is done by means of a scaling factor f_S. This scaling factor is the ratio between the total mass loss over the glacier ΔM (units of m water equivalent) and the integrated normalized change in surface elevation, scaled by the surface area of the elevation zone A_i (as a fraction of the total glacier area).

$$f_S = \frac{\Delta M}{\sum_{i=1}^{N} A_i * \Delta h_{i,\text{norm}}} \tag{B3}$$

Now, we can compute the new water equivalent for each elevation zone after a certain amount of mass loss.

$$h_{i,k+1} = h_{i,k} + f_S \Delta h_{i,\text{norm}}, \tag{B4}$$

with $k = 0$ as the initial glacier profile and h_i is the ice thickness (units of m water equivalent) in each cell in that specific elevation zone. Huss et al. (2010) restricted surface elevation lowering at the edge of the glacier (where $h < 10$ m). Here, we do not apply this because our grid cells are much coarser than those they used and the ice thickness is thus almost never that thin. Finally, we make sure that cells can not have less than 0 ice thickness. Any leftover mass loss is distributed over the rest of the glacier. Note that we apply this procedure to all glaciers independent of size. However, this scheme will hardly affect glaciers consisting of only a few cells, which are mostly governed by the mass balance per grid cell. For larger glaciers, this scheme becomes more important.

Author contributions. JJ, NW, MvT and MIB contributed to the conceptualization of the study. JJ, BvJ and DNK worked on software and data curation. JJ performed the formal analyses and visualizations. JJ wrote the original draft, with all authors contributing to reviewing and editing the final draft. MIB and NW were involved in supervision and MIB acquired the funding for this project.

Competing interests. Manuela Brunner and Niko Wanders are editors with HESS. The authors declare no further competing interests.

Acknowledgements. We thank the Swiss National Science Foundation (SNSF) for funding this study through project 'Predicting floods and droughts under global change' PZ00P2_201818 (granted to MIB). We want to thank Christoph Schlemper and Jonas Götte for helping with getting access to data. We further thank Tobias Jonas for his valuable suggestions with regard to the snow module and thank Harry Zekollari for kindly providing us with the reference data for the glacier stability simulations.

References

Abeshu, G. W., Tian, F., Wild, T., Zhao, M., Turner, S., Chowdhury, A. F. M. K., Vernon, C. R., Hu, H., Zhuang, Y., Hejazi, M., and Li, H.-Y.: Enhancing the representation of water management in global hydrological models, Geoscientific Model Development, 16, 5449–5472, https://doi.org/10.5194/gmd-16-5449-2023, 2023.

Andréassian, V., Le Moine, N., Perrin, C., Ramos, M.-H., Oudin, L., Mathevet, T., Lerat, J., and Berthet, L.: All that glitters is not gold: the case of calibrating hydrological models, Hydrological Processes, 26, 2206–2210, https://doi.org/10.1002/hyp.9264, 2012.

Arheimer, B., Donnelly, C., and Lindström, G.: Regulation of snow-fed rivers affects flow regimes more than climate change, Nature Communications, 8, 62, https://doi.org/10.1038/s41467-017-00092-8, 2017.

Bandhauer, M., Isotta, F., Lakatos, M., Lussana, C., Båserud, L., Izsák, B., Szentes, O., Tveito, O. E., and Frei, C.: Evaluation of daily precipitation analyses in E-OBS (v19.0e) and ERA5 by comparison to regional high-resolution datasets in European regions, International Journal of Climatology, 42, 727–747, https://doi.org/10.1002/joc.7269, 2022.

Bierkens, M. F. P. and Van Beek, L. P. H.: The global hydrological model PCR-GLOBWB: conceptualization, parameterization and verification, Tech. rep., Department of Physical Geography, Utrecht University, Utrecht, The Netherlands, https://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf, 2009.

Bierkens, M. F. P., Bell, V. A., Burek, P., Chaney, N., Condon, L. E., David, C. H., de Roo, A., Döll, P., Drost, N., Famiglietti, J. S., Flörke, M., Gochis, D. J., Houser, P., Hut, R., Keune, J., Kollet, S., Maxwell, R. M., Reager, J. T., Samaniego, L., Sudicky, E., Sutanudjaja, E. H., van de Giesen, N., Winsemius, H., and Wood, E. F.: Hyper-resolution global hydrological modelling: what is next?, Hydrological Processes, 29, 310–320, https://doi.org/10.1002/hyp.10391, 2015.

Brun, P., Zimmermann, N. E., Hari, C., Pellissier, L., and Karger, D. N.: CHELSA-BIOCLIM+ A novel set of global climate-related predictors at kilometre-resolution, EnviDat [data set], https://doi.org/10.16904/ENVIDAT.332, 2022a.

Brun, P., Zimmermann, N. E., Hari, C., Pellissier, L., and Karger, D. N.: Global climate-related predictors at kilometer resolution for the past and future, Earth System Science Data, 14, 5573–5603, https://doi.org/10.5194/essd-14-5573-2022, 2022b.

Brunner, M. I. and Naveau, P.: Spatial variability in Alpine reservoir regulation: deriving reservoir operations from streamflow using generalized additive models, Hydrology and Earth System Sciences, 27, 673–687, https://doi.org/10.5194/hess-27-673-2023, 2023.

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

Carletti, F., Michel, A., Casale, F., Burri, A., Bocchiola, D., Bavay, M., and Lehning, M.: A comparison of hydrological models with different level of complexity in Alpine regions in the context of climate change, Hydrology and Earth System Sciences, 26, 3447–3475, https://doi.org/10.5194/hess-26-3447-2022, 2022.

Colombo, N., Guyennon, N., Valt, M., Salerno, F., Godone, D., Cianfarra, P., Freppaz, M., Maugeri, M., Manara, V., Acquaotta, F., Petrangeli, A. B., and Romano, E.: Unprecedented snow-drought conditions in the Italian Alps during the early 2020s, Environmental Research Letters, 18, 074 014, https://doi.org/10.1088/1748-9326/acdb88, 2023.

Copernicus Climate Change Service: ERA5-Land hourly data from 1950 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set], https://doi.org/10.24381/CDS.E2161BAC, 2019.

Daly, C., Taylor, G., and Gibson, W.: The PRISM approach to mapping precipitation and temperature, in: Proceedings of the 10th AMS Conference on Applied Climatology, pp. 10–12, Reno, United States, 1997.

Dolan, F., Lamontagne, J., Link, R., Hejazi, M., Reed, P., and Edmonds, J.: Evaluating the economic impact of water scarcity in a changing world, Nature Communications, 12, 1915, https://doi.org/10.1038/s41467-021-22194-0, 2021.

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., De Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., Van Der Schalie, R., Seneviratne, S. I., Smolander, T., and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sensing of Environment, 203, 185–215, https://doi.org/10.1016/j.rse.2017.07.001, 2017.

Dorigo, W., Preimesberger, W., Hahn, S., Van Der Schalie, R., De Jeu, R., Kidd, R., Rodríguez-Fernández, N., Hirschi, M., Stradiotti, P., Frederikse, T., Gruber, A., and Madelon, R.: ESA Soil Moisture Climate Change Initiative (Soil_Moisture_cci): COMBINED product, Version 08.1, NERC EDS Centre for Environmental Data Analysis, https://catalogue.ceda.ac.uk/uuid/6f99cdb86a9e4d3da2d47c79612c00a2/, 2023.

Dorigo, W. A., Gruber, A., De Jeu, R. A. M., Wagner, W., Stacke, T., Loew, A., Albergel, C., Brocca, L., Chung, D., Parinussa, R. M., and Kidd, R.: Evaluation of the ESA CCI soil moisture product using ground-based observations, Remote Sensing of Environment, 162, 380–395, https://doi.org/10.1016/j.rse.2014.07.023, 2015.

Duethmann, D., Peters, J., Blume, T., Vorogushyn, S., and Güntner, A.: The value of satellite-derived snow cover images for calibrating a hydrological model in snow-dominated catchments in Central Asia, Water Resources Research, 50, 2002–2021, https://doi.org/10.1002/2013WR014382, 2014.

European Space Agency and Airbus: Copernicus DEM, Copernicus, https://doi.org/10.5270/ESA-c5d3d65, 2022.

Fan, Y., Clark, M., Lawrence, D. M., Swenson, S., Band, L. E., Brantley, S. L., Brooks, P. D., Dietrich, W. E., Flores, A., Grant, G., Kirchner, J. W., Mackay, D. S., McDonnell, J. J., Milly, P. C. D., Sullivan, P. L., Tague, C., Ajami, H., Chaney, N., Hartmann, A., Hazenberg, P., McNamara, J., Pelletier, J., Perket, J., Rouholahnejad-Freund, E., Wagener, T., Zeng, X., Beighley, E., Buzan, J., Huang, M., Livneh, B., Mohanty, B. P., Nijssen, B., Safeeq, M., Shen, C., van Verseveld, W., Volk, J., and Yamazaki, D.: Hillslope hydrology in global change research and earth system modeling, Water Resources Research, 55, 1737–1772, https://doi.org/10.1029/2018WR023903, 2019.

Farinotti, D.: A consensus estimate for the ice thickness distribution of all glaciers on Earth - dataset, ETH Zurich [data set], https://doi.org/10.3929/ETHZ-B-000315707, 2019.

Farinotti, D., Usselmann, S., Huss, M., Bauder, A., and Funk, M.: Runoff evolution in the Swiss Alps: projections for selected high-alpine catchments based on ENSEMBLES scenarios, Hydrological Processes, 26, 1909–1924, https://doi.org/10.1002/hyp.8276, 2012.

Farinotti, D., Huss, M., Fürst, J. J., Landmann, J., Machguth, H., Maussion, F., and Pandit, A.: A consensus estimate for the ice thickness distribution of all glaciers on Earth, Nature Geoscience, 12, 168–173, https://doi.org/10.1038/s41561-019-0300-3, 2019.

Finger, D., Vis, M., Huss, M., and Seibert, J.: The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments, Water Resources Research, 51, 1939–1958, https://doi.org/10.1002/2014WR015712, 2015.

Fontrodona-Bach, A., Schaefli, B., Woods, R., Teuling, A. J., and Larsen, J. R.: NH-SWE: Northern Hemisphere Snow Water Equivalent dataset based on in-situ snow depth time series and the regionalisation of the ΔSNOW model, Zenodo [data set], https://doi.org/10.5281/ZENODO.7565252, 2023a.

Fontrodona-Bach, A., Schaefli, B., Woods, R., Teuling, A. J., and Larsen, J. R.: NH-SWE: Northern Hemisphere Snow Water Equivalent dataset based on in situ snow depth time series, Earth System Science Data, 15, 2577–2599, https://doi.org/10.5194/essd-15-2577-2023, 2023b.

Freudiger, D., Kohn, I., Seibert, J., Stahl, K., and Weiler, M.: Snow redistribution for the hydrological modeling of alpine catchments: Snow redistribution for hydrological modeling, Wiley Interdisciplinary Reviews: Water, 4, e1232, https://doi.org/10.1002/wat2.1232, 2017.

Frey, S. and Holzmann, H.: A conceptual, distributed snow redistribution model, Hydrology and Earth System Sciences, 19, 4517–4530, https://doi.org/10.5194/hess-19-4517-2015, 2015.

Gampe, D. and Ludwig, R.: Evaluation of gridded precipitation data products for hydrological applications in complex topography, Hydrology, 4, 53, https://doi.org/10.3390/hydrology4040053, 2017.

Gebrechorkos, S. H., Leyland, J., Dadson, S. J., Cohen, S., Slater, L., Wortmann, M., Ashworth, P. J., Bennett, G. L., Boothroyd, R., Cloke, H., Delorme, P., Griffith, H., Hardy, R., Hawker, L., McLelland, S., Neal, J., Nicholas, A., Tatem, A. J., Vahidi, E., Liu, Y., Sheffield, J., Parsons, D. R., and Darby, S. E.: Global-scale evaluation of precipitation datasets for hydrological modelling, Hydrology and Earth System Sciences, 28, 3099–3118, https://doi.org/10.5194/hess-28-3099-2024, 2024.

GeoSphere Austria: SNOWGRID Klima v2.1, GeoSphere Austria [data set], https://doi.org/10.60669/FSXX-6977, 2022.

Gharari, S., Clark, M. P., Mizukami, N., Wong, J. S., Pietroniro, A., and Wheater, H. S.: Improving the representation of subsurface water movement in land models, Journal of Hydrometeorology, 20, 2401–2418, https://doi.org/10.1175/JHM-D-19-0108.1, 2019.

Girons Lopez, M., Vis, M. J. P., Jenicek, M., Griessinger, N., and Seibert, J.: Assessing the degree of detail of temperature-based snow routines for runoff modelling in mountainous areas in central Europe, Hydrology and Earth System Sciences, 24, 4441–4461, https://doi.org/10.5194/hess-24-4441-2020, 2020.

Glacier Monitoring Switzerland (GLAMOS): Swiss glacier mass balance (release 2022), GLAMOS - Glacier Monitoring Switzerland [data set], https://doi.org/10.18750/MASSBALANCE.2022.R2022, 2022.

GLIMS Consortium: GLIMS Glacier Database, Version 1, National Snow and Ice Data Center [data set], https://doi.org/10.7265/N5V98602, 2018.

Godet, J., Gaume, E., Javelle, P., Nicolle, P., and Payrastre, O.: Technical note: Comparing three different methods for allocating river points to coarse-resolution hydrological modelling grid cells, Hydrology and Earth System Sciences, 28, 1403–1413, https://doi.org/10.5194/hess-28-1403-2024, 2024.

Gosling, S. N. and Arnell, N. W.: Simulating current global river runoff with a global hydrological model: model revisions, validation, and sensitivity analysis, Hydrological Processes, 25, 1129–1145, https://doi.org/10.1002/hyp.7727, 2011.

Gruber, A., Scanlon, T., Van Der Schalie, R., Wagner, W., and Dorigo, W.: Evolution of the ESA CCI Soil Moisture climate data records and their underlying merging methodology, Earth System Science Data, 11, 717–739, https://doi.org/10.5194/essd-11-717-2019, 2019.

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.

Gädeke, A., Krysanova, V., Aryal, A., Chang, J., Grillakis, M., Hanasaki, N., Koutroulis, A., Pokhrel, Y., Satoh, Y., Schaphoff, S., Müller Schmied, H., Stacke, T., Tang, Q., Wada, Y., and Thonicke, K.: Performance evaluation of global hydrological models in six large Pan-Arctic watersheds, Climatic Change, 163, 1329–1351, https://doi.org/10.1007/s10584-020-02892-2, 2020.

Hamon, W. R.: Estimating potential evapotranspiration, Transactions of the American Society of Civil Engineers, 128, 324–338, https://doi.org/10.1061/TACEAT.0008673, 1963.

Hanasaki, N., Kanae, S., and Oki, T.: A reservoir operation scheme for global river routing models, Journal of Hydrology, 327, 22–41, https://doi.org/10.1016/j.jhydrol.2005.11.011, 2006.

Hanus, S., Schuster, L., Burek, P., Maussion, F., Wada, Y., and Viviroli, D.: Coupling a large-scale glacier and hydrological model (OGGM v1.5.3 and CWatM V1.08) – towards an improved representation of mountain water resources in global assessments, Geoscientific Model Development, 17, 5123–5144, https://doi.org/10.5194/gmd-17-5123-2024, 2024.

Heinicke, S., Volkholz, J., Schewe, J., Gosling, S. N., Schmied, H. M., Zimmermann, S., Mengel, M., Sauer, I. J., Burek, P., Chang, J., Kou-Giesbrecht, S., Grillakis, M., Guillaumot, L., Hanasaki, N., Koutroulis, A., Otta, K., Qi, W., Satoh, Y., Stacke, T., Yokohata, T., and Frieler, K.: Global hydrological models continue to overestimate river discharge, Environmental Research Letters, 19, 074 005, https://doi.org/10.1088/1748-9326/ad52b0, 2024.

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.

Hoch, J. M., Sutanudjaja, E. H., Wanders, N., van Beek, R. L. P. H., and Bierkens, M. F. P.: Hyper-resolution PCR-GLOBWB: opportunities and challenges from refining model spatial resolution to 1 km over the European continent, Hydrology and Earth System Sciences, 27, 1383–1401, https://doi.org/10.5194/hess-27-1383-2023, 2023.

Hock, R.: A distributed temperature-index ice- and snowmelt model including potential direct solar radiation, Journal of Glaciology, 45, 101–111, https://doi.org/10.3189/S0022143000003087, 1999.

Hock, R.: Temperature index melt modelling in mountain areas, Journal of Hydrology, 282, 104–115, https://doi.org/10.1016/S0022-1694(03)00257-9, 2003.

Hou, Y., Guo, H., Yang, Y., and Liu, W.: Global evaluation of runoff simulation from climate, hydrological and land surface models, Water Resources Research, 59, https://doi.org/10.1029/2021WR031817, 2023.

Houska, T., Kraft, P., Chamorro-Chavez, A., and Breuer, L.: SPOTting model parameters using a ready-made python package, PLoS ONE, 10, e0145 180, https://doi.org/10.1371/journal.pone.0145180, 2015.

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., and Kääb, A.: Accelerated global glacier mass loss in the early twenty-first century, Nature, 592, 726–731, https://doi.org/10.1038/s41586-021-03436-z, 2021a.

Hugonnet, R., McNabb, R., Berthier, E., Menounos, B., Nuth, C., Girod, L., Farinotti, D., Huss, M., Dussaillant, I., Brun, F., and Kääb, A.: Accelerated global glacier mass loss in the early twenty-first century - Dataset., Theia [data set], https://doi.org/10.6096/13, 2021b.

Huss, M.: Present and future contribution of glacier storage change to runoff from macroscale drainage basins in Europe, Water Resources Research, 47, https://doi.org/10.1029/2010WR010299, 2011.

Huss, M. and Hock, R.: A new model for global glacier change and sea-level rise, Frontiers in Earth Science, 3, https://www.frontiersin.org/articles/10.3389/feart.2015.00054, 2015.

Huss, M. and Hock, R.: Global-scale hydrological response to future glacier mass loss, Nature Climate Change, 8, 135–140, https://doi.org/10.1038/s41558-017-0049-x, 2018.

Huss, M., Jouvet, G., Farinotti, D., and Bauder, A.: Future high-mountain hydrology: a new parameterization of glacier retreat, Hydrology and Earth System Sciences, 14, 815–829, https://doi.org/10.5194/hess-14-815-2010, 2010.

Immerzeel, W. W., van Beek, L. P. H., and Bierkens, M. F. P.: Climate change will affect the Asian water towers, Science, 328, 1382–1385, https://doi.org/10.1126/science.1183188, 2010.

Immerzeel, W. W., van Beek, L. P. H., Konz, M., Shrestha, A. B., and Bierkens, M. F. P.: Hydrological response to climate change in a glacierized catchment in the Himalayas, Climatic Change, 110, 721–736, https://doi.org/10.1007/s10584-011-0143-4, 2012.

Immerzeel, W. W., Lutz, A. F., Andrade, M., Bahl, A., Biemans, H., Bolch, T., Hyde, S., Brumby, S., Davies, B. J., Elmore, A. C., Emmer, A., Feng, M., Fernández, A., Haritashya, U., Kargel, J. S., Koppes, M., Kraaijenbrink, P. D. A., Kulkarni, A. V., Mayewski, P. A., Nepal, S., Pacheco, P., Painter, T. H., Pellicciotti, F., Rajaram, H., Rupper, S., Sinisalo, A., Shrestha, A. B., Viviroli, D., Wada, Y., Xiao, C., Yao, T., and Baillie, J. E. M.: Importance and vulnerability of the world's water towers, Nature, 577, 364–369, https://doi.org/10.1038/s41586-019-1822-y, 2020.

Ismail, M. F., Bogacki, W., Disse, M., Schäfer, M., and Kirschbauer, L.: Estimating degree-day factors of snow based on energy flux components, The Cryosphere, 17, 211–231, https://doi.org/10.5194/tc-17-211-2023, 2023.

Isotta, F. and Frei, C.: APGD: Alpine precipitation grid dataset, Federal Office of Meteorology and Climatology MeteoSwiss [data set], https://doi.org/10.18751/CLIMATE/GRIDDATA/APGD/1.0, 2013.

Isotta, F. A., Frei, C., Weilguni, V., Perčec Tadić, M., Lassègues, P., Rudolf, B., Pavan, V., Cacciamani, C., Antolini, G., Ratto, S. M., Munari, M., Micheletti, S., Bonati, V., Lussana, C., Ronchi, C., Panettieri, E., Marigo, G., and Vertačnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, International Journal of Climatology, 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

Isotta, F. A., Vogel, R., and Frei, C.: Evaluation of European regional reanalyses and downscalings for precipitation in the Alpine region, Meteorologische Zeitschrift, 24, 15–37, https://doi.org/10.1127/metz/2014/0584, 2015.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., and Kessler, M.: Climatologies at high resolution for the earth's land surface areas, Scientific Data, 4, 170 122, https://doi.org/10.1038/sdata.2017.122, 2017.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., and Kessler, M.: Climatologies at high resolution for the earth's land surface areas, EnviDat, https://doi.org/10.16904/ENVIDAT.228, 2021a.

Karger, D. N., Wilson, A. M., Mahony, C., Zimmermann, N. E., and Jetz, W.: Global daily 1 km land surface precipitation based on cloud cover-informed downscaling, Scientific Data, 8, 307, https://doi.org/10.1038/s41597-021-01084-6, 2021b.

Karger, D. N., Lange, S., Hari, C., Reyer, C. P. O., Conrad, O., Zimmermann, N. E., and Frieler, K.: CHELSA-W5E5: daily 1 km meteorological forcing data for climate impact studies, Earth System Science Data, 15, 2445–2464, https://doi.org/10.5194/essd-15-2445-2023, 2023.

Keller, J. D. and Wahl, S.: Representation of climate in reanalyses: an intercomparison for Europe and North America, Journal of Climate, 34, 1667–1684, https://doi.org/10.1175/JCLI-D-20-0609.1, 2021.

Khanal, S., Lutz, A., Kraaijenbrink, P. D. A., van den Hurk, B., Yao, T., and Immerzeel, W. W.: Variable 21st century climate change response for rivers in High Mountain Asia at seasonal to decadal time scales, Water Resources Research, 57, e2020WR029 266, https://doi.org/10.1029/2020WR029266, 2021.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrological Sciences Journal, 31, 13–24, https://doi.org/10.1080/02626668609491024, 1986.

Klingler, C., Kratzert, F., Schulz, K., and Herrnegger, M.: LamaH-CE: LArge-SaMple DAta for Hydrology and Environmental Sciences for Central Europe – files, Zenodo [data set], https://doi.org/10.5281/ZENODO.4525244, 2021.

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.

Knoben, W. J. M., Freer, J. E., Peel, M. C., Fowler, K. J. A., and Woods, R. A.: A brief analysis of conceptual model structure uncertainty using 36 models and 559 catchments, Water Resources Research, 56, e2019WR025 975, https://doi.org/10.1029/2019WR025975, 2020.

Kraaijenbrink, P. D. A., Bierkens, M. F. P., Lutz, A. F., and Immerzeel, W. W.: Impact of a global temperature rise of 1.5 degrees Celsius on Asia's glaciers, Nature, 549, 257–260, https://doi.org/10.1038/nature23878, 2017.

Kraaijenbrink, P. D. A., Stigter, E. E., Yao, T., and Immerzeel, W. W.: Climate change decisive for Asia's snow meltwater supply, Nature Climate Change, 11, 591–597, https://doi.org/10.1038/s41558-021-01074-x, 2021.

Kuhn, M.: Redistribution of snow and glacier mass balance from a hydrometeorological model, Journal of Hydrology, 282, 95–103, https://doi.org/10.1016/S0022-1694(03)00256-7, 2003.

Kuusisto, E.: On the Values and Variability of Degree-Day Melting Factor in Finland, Hydrology Research, 11, 235–242, https://doi.org/10.2166/nh.1980.0011, 1980.

Lange, S., Menz, C., Gleixner, S., Cucchi, M., Weedon, G. P., Amici, A., Bellouin, N., Müller Schmied, H., Hersbach, H., Buontempo, C., and Cagnazzo, C.: WFDE5 over land merged with ERA5 over the ocean (W5E5 v2.0), ISIMIP Repository [data set], https://doi.org/10.48364/ISIMIP.342217, 2021.

Lehner, B., Czisch, G., and Vassolo, S.: The impact of global change on the hydropower potential of Europe: a model-based analysis, Energy Policy, 33, 839–855, https://doi.org/10.1016/j.enpol.2003.10.018, 2005.

Leijnse, M., Bierkens, M. F. P., Gommans, K. H. M., Lin, D., Tait, A., and Wanders, N.: Key drivers and pressures of global water scarcity hotspots, Environmental Research Letters, 19, 054 035, https://doi.org/10.1088/1748-9326/ad3c54, 2024.

Magnusson, J., Gustafsson, D., Hüsler, F., and Jonas, T.: Assimilation of point SWE data into a distributed snow cover model comparing two contrasting methods, Water Resources Research, 50, 7816–7835, https://doi.org/10.1002/2014WR015302, 2014.

Magnusson, J., Wever, N., Essery, R., Helbig, N., Winstral, A., and Jonas, T.: Evaluating snow models with varying process representations for hydrological applications, Water Resources Research, 51, 2707–2723, https://doi.org/10.1002/2014WR016498, 2015.

Malle, J. T., Mazzotti, G., Karger, D. N., and Jonas, T.: Regionally optimized high-resolution input datasets enhance the representation of snow cover in CLM5, Earth System Dynamics, 15, 1073–1115, https://doi.org/10.5194/esd-15-1073-2024, 2024.

Maussion, F., Butenko, A., Champollion, N., Dusch, M., Eis, J., Fourteau, K., Gregor, P., Jarosch, A. H., Landmann, J., Oesterle, F., Recinos, B., Rothenpieler, T., Vlug, A., Wild, C. T., and Marzeion, B.: The Open Global Glacier Model (OGGM) v1.1, Geoscientific Model Development, 12, 909–931, https://doi.org/10.5194/gmd-12-909-2019, 2019.

Mortimer, C., Mudryk, L., Derksen, C., Luojus, K., Brown, R., Kelly, R., and Tedesco, M.: Evaluation of long-term Northern Hemisphere snow water equivalent products, The Cryosphere, 14, 1579–1594, https://doi.org/10.5194/tc-14-1579-2020, 2020.

Mott, R.: Climatological snow data since 1998, OSHD, EnviDat [data set], https://doi.org/10.16904/ENVIDAT.401, 2023.

Mott, R., Winstral, A., Cluzet, B., Helbig, N., Magnusson, J., Mazzotti, G., Quéno, L., Schirmer, M., Webster, C., and Jonas, T.: Operational snow-hydrological modeling for Switzerland, Frontiers in Earth Science, 11, https://www.frontiersin.org/articles/10.3389/feart.2023.1228158, 2023.

Mudryk, L., Mortimer, C., Derksen, C., Elias Chereque, A., and Kushner, P.: Benchmarking of snow water equivalent (SWE) products based on outcomes of the SnowPEx+ Intercomparison Project, The Cryosphere, 19, 201–218, https://doi.org/10.5194/tc-19-201-2025, 2025.

Müller Schmied, H., Cáceres, D., Eisner, S., Flörke, M., Herbert, C., Niemann, C., Peiris, T. A., Popat, E., Portmann, F. T., Reinecke, R., Schumacher, M., Shadkam, S., Telteu, C.-E., Trautmann, T., and Döll, P.: The global water resources and use model WaterGAP v2.2d: model description and evaluation, Geoscientific Model Development, 14, 1037–1079, https://doi.org/10.5194/gmd-14-1037-2021, 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.

Olefs, M., Schöner, W., Suklitsch, M., Wittmann, C., Niedermoser, B., Neururer, A., and Wurzer, A.: SNOWGRID – A new operational snow cover model in Austria, in: Proceedings of the International Snow Science Workshop 2013, pp. 38–45, Grenoble, France, https://www.semanticscholar.org/paper/SNOWGRID-%E2%80%93-A-New-Operational-Snow-Cover-Model-in-Olefs-Sch%C3%B6ner/7577e333b866a0edd86d39dfdac21ce8a86e5b67, 2013.

Olefs, M., Koch, R., Schöner, W., and Marke, T.: Changes in snow depth, snow cover duration, and potential snowmaking conditions in Austria, 1961–2020—a model based approach, Atmosphere, 11, 1330, https://doi.org/10.3390/atmos11121330, 2020.

Partl, R.: Statistik 1977 der großen Talsperren und Flußstauwerke Österreichs, Österreichischer Wasserwirtschaftsverband, ISBN 978-0-387-81461-2, 1977.

Pfeffer, W. T., Arendt, A. A., Bliss, A., Bolch, T., Cogley, J. G., Gardner, A. S., Hagen, J.-O., Hock, R., Kaser, G., Kienholz, C., Miles, E. S., Moholdt, G., Mölg, N., Paul, F., Radić, V., Rastner, P., Raup, B. H., Rich, J., Sharp, M. J., and Consortium, T. R.: The Randolph Glacier Inventory: a globally complete inventory of glaciers, Journal of Glaciology, 60, 537–552, https://doi.org/10.3189/2014JoG13J176, 2014.

Preimesberger, W., Scanlon, T., Su, C.-H., Gruber, A., and Dorigo, W.: Homogenization of structural breaks in the Global ESA CCI Soil Moisture Multisatellite Climate Data Record, IEEE Transactions on Geoscience and Remote Sensing, 59, 2845–2862, https://doi.org/10.1109/TGRS.2020.3012896, 2021.

Rahman, M. and Rosolem, R.: Towards a simple representation of chalk hydrology in land surface modelling, Hydrology and Earth System Sciences, 21, 459–471, https://doi.org/10.5194/hess-21-459-2017, 2017.

Raimonet, M., Oudin, L., Thieu, V., Silvestre, M., Vautard, R., Rabouille, C., and Moigne, P. L.: Evaluation of gridded meteorological datasets for hydrological modeling, Journal of Hydrometeorology, 18, 3027–3041, https://doi.org/10.1175/JHM-D-17-0018.1, 2017.

Rango, A. and Martinec, J.: Revisiting the Degree-Day Method for Snowmelt Computations, JAWRA Journal of the American Water Resources Association, 31, 657–669, https://doi.org/10.1111/j.1752-1688.1995.tb03392.x, 1995.

Raup, B., Racoviteanu, A., Khalsa, S. J. S., Helm, C., Armstrong, R., and Arnaud, Y.: The GLIMS geospatial glacier database: A new tool for studying glacier change, Global and Planetary Change, 56, 101–110, https://doi.org/10.1016/j.gloplacha.2006.07.018, 2007.

Refsgaard, J. C. and Storm, B.: Construction, calibration and validation of hydrological models, in: Distributed hydrological modelling, edited by Abbott, M. B. and Refsgaard, J. C., pp. 41–54, Springer Netherlands, Dordrecht, ISBN 978-94-009-0257-2, https://doi.org/10.1007/978-94-009-0257-2_3, 1996.

RGI Consortium: Randolph Glacier Inventory - a dataset of global glacier outlines, Version 6, National Snow and Ice Data Center, https://doi.org/10.7265/4M1F-GD79, 2017.

Ridal, M., Bazile, E., Le Moigne, P., Randriamampianina, R., Schimanke, S., Andrae, U., Berggren, L., Brousseau, P., Dahlgren, P., Edvinsson, L., El-Said, A., Glinton, M., Hagelin, S., Hopsch, S., Isaksson, L., Medeiros, P., Olsson, E., Unden, P., and Wang, Z. Q.: CERRA, the Copernicus European Regional Reanalysis system, Quarterly Journal of the Royal Meteorological Society, n/a, https://doi.org/10.1002/qj.4764, 2024.

Salwey, S., Coxon, G., Pianosi, F., Singer, M. B., and Hutton, C.: National-scale detection of reservoir impacts through hydrological signatures, Water Resources Research, 59, e2022WR033 893, https://doi.org/10.1029/2022WR033893, 2023.

Schaefli, B., Manso, P., Fischer, M., Huss, M., and Farinotti, D.: The role of glacier retreat for Swiss hydropower production, Renewable Energy, 132, 615–627, https://doi.org/10.1016/j.renene.2018.07.104, 2019.

Schellekens, J., Dutra, E., Martínez-de la Torre, A., Balsamo, G., van Dijk, A., Sperna Weiland, F., 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., and Weedon, G. P.: A global water resources ensemble of hydrological models: the eartH2Observe Tier-1 dataset, Earth System Science Data, 9, 389–413, https://doi.org/10.5194/essd-9-389-2017, 2017.

Schimanke, S., Ridal, M., Moigne, P. L., Berggren, L., Undén, P., Randriamampianina, Roger, R., Andrea, U., Bazile, E., Bertelsen, A., and Brousseau, P.: CERRA sub-daily regional reanalysis data for Europe on single levels from 1984 to present, ECMWF [data set], https://doi.org/10.24381/CDS.622A565A, 2021.

Schlemper, C., Götte, J., and Brunner, M.: Large-sample hydro-meteorological dataset for the Alps, HydroShare [data set], https://doi.org/10.4211/hs.f1c12fa6c5be4c61a5ec617fa62e13d6, 2024.

Seibert, J.: Reliability of model predictions outside calibration conditions: Paper presented at the Nordic Hydrological Conference (Røros, Norway 4-7 August 2002), Hydrology Research, 34, 477–492, https://doi.org/10.2166/nh.2003.0019, 2003.

Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrology and Earth System Sciences, 16, 3315–3325, https://doi.org/10.5194/hess-16-3315-2012, 2012.

Seibert, J., Vis, M. J. P., Kohn, I., Weiler, M., and Stahl, K.: Technical note: Representing glacier geometry changes in a semi-distributed hydrological model, Hydrology and Earth System Sciences, 22, 2211–2224, https://doi.org/10.5194/hess-22-2211-2018, 2018a.

Seibert, J., Vis, M. J. P., Lewis, E., and van Meerveld, H.: Upper and lower benchmarks in hydrological modelling, Hydrological Processes, 32, 1120–1125, https://doi.org/10.1002/hyp.11476, 2018b.

Sevruk, B.: Correction of precipitation measurements: Swiss experience, in: Papers presented at the workshop on the correction of precipitation measurements, pp. 187–196, Zurich, Switzerland, https://library.wmo.int/records/item/41746-papers-presented-at-the-workshop-on-the-correction-of-precipitation-measurements?offset=58, 1985.

Shin, S., Pokhrel, Y., and Miguez-Macho, G.: High-resolution modeling of reservoir release and storage dynamics at the continental scale, Water Resources Research, 55, 787–810, https://doi.org/10.1029/2018WR023025, 2019.

Simmler, H.: Die Talsperren Österreichs: Statistik 1961, Springer-Verlag, ISBN 978-3-7091-5546-2, 1961.

Slater, A. G. and Clark, M. P.: Snow data assimilation via an ensemble Kalman filter, Journal of Hydrometeorology, 7, 478–493, https://doi.org/10.1175/JHM505.1, 2006.

Sleziak, P., Szolgay, J., Hlavčová, K., Danko, M., and Parajka, J.: The effect of the snow weighting on the temporal stability of hydrologic model efficiency and parameters, Journal of Hydrology, 583, 124 639, https://doi.org/10.1016/j.jhydrol.2020.124639, 2020.

Speckhann, G., Kreibich, H., and Merz, B.: Inventory of dams in Germany, GFZ Data Services [data set], https://doi.org/10.5880/GFZ.4.4.2020.005, 2020.

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.

Stacke, T. and Hagemann, S.: HydroPy (v1.0): a new global hydrology model written in Python, Geoscientific Model Development, 14, 7795–7816, https://doi.org/10.5194/gmd-14-7795-2021, 2021.

Stahl, K., Moore, R. D., Shea, J. M., Hutchinson, D., and Cannon, A. J.: Coupled modelling of glacier and streamflow response to future climate scenarios: modelling of glacier and streamflow, Water Resources Research, 44, https://doi.org/10.1029/2007WR005956, 2008.

Sutanudjaja, E. H., van Beek, L. P. H., de Jong, S. M., van Geer, F. C., and Bierkens, M. F. P.: Large-scale groundwater modeling using global datasets: a test case for the Rhine-Meuse basin, Hydrology and Earth System Sciences, 15, 2913–2935, https://doi.org/10.5194/hess-15-2913-2011, 2011.

Sutanudjaja, E. H., van Beek, R., Wanders, N., Wada, Y., Bosmans, J. H. C., Drost, N., van der Ent, R. J., de Graaf, I. E. M., Hoch, J. M., de Jong, K., Karssenberg, D., López López, P., Peßenteiner, S., Schmitz, O., Straatsma, M. W., Vannametee, E., Wisser, D., and Bierkens, M. F. P.: PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model, Geoscientific Model Development, 11, 2429–2453, https://doi.org/10.5194/gmd-11-2429-2018, 2018.

Tang, G., Clark, M. P., Knoben, W. J. M., Liu, H., Gharari, S., Arnal, L., Beck, H. E., Wood, A. W., Newman, A. J., and Papalexiou, S. M.: The impact of meteorological forcing uncertainty on hydrological modeling: a global analysis of cryosphere basins, Water Resources Research, 59, e2022WR033 767, https://doi.org/10.1029/2022WR033767, 2023.

Telteu, C.-E., Müller Schmied, H., Thiery, W., Leng, G., Burek, P., Liu, X., Boulange, J. E. S., Andersen, L. S., Grillakis, M., Gosling, S. N., Satoh, Y., Rakovec, O., Stacke, T., Chang, J., Wanders, N., Shah, H. L., Trautmann, T., Mao, G., Hanasaki, N., Koutroulis, A., Pokhrel, Y., Samaniego, L., Wada, Y., Mishra, V., Liu, J., Döll, P., Zhao, F., Gädeke, A., Rabin, S. S., and Herz, F.: Understanding each other's models: an introduction and a standard representation of 16 global water models to support intercomparison, improvement, and communication, Geoscientific Model Development, 14, 3843–3878, https://doi.org/10.5194/gmd-14-3843-2021, 2021.

Tu, T., Wang, J., Zhao, G., Zhao, T., and Dong, X.: Scaling from global to regional river flow with global hydrological models: Choice matters, Journal of Hydrology, 633, 130 960, https://doi.org/10.1016/j.jhydrol.2024.130960, 2024.

Turner, S. W. D. and Voisin, N.: Simulation of hydropower at subcontinental to global scales: a state-of-the-art review, Environmental Research Letters, 17, 023 002, https://doi.org/10.1088/1748-9326/ac4e38, 2022.

Van Jaarsveld, B., Wanders, N., Sutanudjaja, E. H., Hoch, J., Droppers, B., Janzing, J., Van Beek, R. L. P. H., and Bierkens, M. F. P.: A first attempt to model global hydrology at hyper-resolution, Earth System Dynamics, 16, 29–54, https://doi.org/10.5194/esd-16-29-2025, 2025.

Van Tiel, M., Van Loon, A. F., Seibert, J., and Stahl, K.: Hydrological response to warm and dry weather: do glaciers compensate?, Hydrology and Earth System Sciences, 25, 3245–3265, https://doi.org/10.5194/hess-25-3245-2021, 2021.

Van Tiel, M., Weiler, M., Freudiger, D., Moretti, G., Kohn, I., Gerlinger, K., and Stahl, K.: Melting alpine water towers aggravate downstream low flows: a stress-test storyline approach, Earth's Future, 11, e2022EF003 408, https://doi.org/10.1029/2022EF003408, 2023.

Veldkamp, T. I. E., Zhao, F., Ward, P. J., Moel, H. d., Aerts, J. C. J. H., Schmied, H. M., Portmann, F. T., Masaki, Y., Pokhrel, Y., Liu, X., Satoh, Y., Gerten, D., Gosling, S. N., Zaherpour, J., and Wada, Y.: Human impact parameterizations in global hydrological models improve estimates of monthly discharges and hydrological extremes: a multi-model validation study, Environmental Research Letters, 13, 055 008, https://doi.org/10.1088/1748-9326/aab96f, 2018.

Verkaik, J. and Sutanudjaja, E. H.: Data supplement to the GLOBGM v1.0: a parallel implementation of a 30 arcsec PCR-GLOBWB-MODFLOW global-scale groundwater model, Utrecht University, https://doi.org/10.24416/UU01-44L775, 2024.

Verrelle, A., Glinton, M., Bazile, E., Le Moigne, P., Randriamampianina, Roger, R., Ridal, M., Berggren, L., Undén, P., Schimanke, S., Mladek, R., and Soci, C.: CERRA-Land sub-daily regional reanalysis data for Europe from 1984 to present, ECMWF, https://doi.org/10.24381/CDS.A7F3CD0B, 2022.

Viviroli, D., Dürr, H. H., Messerli, B., Meybeck, M., and Weingartner, R.: Mountains of the world, water towers for humanity: Typology, mapping, and global significance, Water Resources Research, 43, https://doi.org/10.1029/2006WR005653, 2007.

Viviroli, D., Kummu, M., Meybeck, M., Kallio, M., and Wada, Y.: Increasing dependence of lowland populations on mountain water resources, Nature Sustainability, 3, 917–928, https://doi.org/10.1038/s41893-020-0559-9, 2020.

Warszawski, L., Frieler, K., Huber, V., Piontek, F., Serdeczny, O., and Schewe, J.: The Inter-Sectoral Impact Model Intercomparison Project (ISI–MIP): Project framework, Proceedings of the National Academy of Sciences, 111, 3228–3232, https://doi.org/10.1073/pnas.1312330110, 2014.

Weingartner, R., Barben, M., and Spreafico, M.: Floods in mountain areas—an overview based on examples from Switzerland, Journal of Hydrology, 282, 10–24, https://doi.org/10.1016/S0022-1694(03)00249-X, 2003.

Wiersma, P., Aerts, J., Zekollari, H., Hrachowitz, M., Drost, N., Huss, M., Sutanudjaja, E. H., and Hut, R.: Coupling a global glacier model to a global hydrological model prevents underestimation of glacier runoff, Hydrology and Earth System Sciences, 26, 5971–5986, https://doi.org/10.5194/hess-26-5971-2022, 2022.

Wilby, R. L.: A global hydrology research agenda fit for the 2030s, Hydrology Research, 50, 1464–1480, https://doi.org/10.2166/nh.2019.100, 2019.

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: OPINION, Water Resources Research, 47, https://doi.org/10.1029/2010WR010090, 2011.

World Glacier Monitoring Service (WGMS): Fluctuations of Glaciers Database, World Glacier Monitoring Service (WGMS), https://doi.org/10.5904/WGMS-FOG-2023-09, 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, C., Tijerina-Kreuzer, D. T., Tran, H. V., Condon, L. E., and Maxwell, R. M.: A high-resolution, 3D groundwater-surface water simulation of the contiguous US: Advances in the integrated ParFlow CONUS 2.0 modeling platform, Journal of Hydrology, 626, 130 294, https://doi.org/10.1016/j.jhydrol.2023.130294, 2023.

Yang, Y., Chen, R., Liu, G., Liu, Z., and Wang, X.: Trends and variability in snowmelt in China under climate change, Hydrology and Earth System Sciences, 26, 305–329, https://doi.org/10.5194/hess-26-305-2022, 2022.

Zekollari, H., Huss, M., and Farinotti, D.: On the imbalance and response time of glaciers in the European Alps, Geophysical Research Letters, 47, e2019GL085 578, https://doi.org/10.1029/2019GL085578, 2020.