Impact of different weather generator scenarios on extreme flood estimates in Switzerland

Eleni Kritidou, Martina Kauzlaric, Maria Staudinger, Guillaume Evin, Benoit Hingray, Marc Vis, Jan Seibert, Daniel Viviroli

Stochastic Environmental Research and Risk Assessment (2025) 39:847–866
https://doi.org/10.1007/s00477-024-02843-8

ORIGINAL PAPER

Impact of different weather generator scenarios on extreme flood estimates in Switzerland

Eleni Kritidou[1] · Martina Kauzlaric[2] · Maria Staudinger[1] · Guillaume Evin[3] · Benoit Hingray[3] · Marc Vis[1] · Jan Seibert[1] · Daniel Viviroli[1]

Accepted: 10 October 2024 / Published online: 11 February 2025

Abstract
The estimation of extreme floods using long continuous simulations is linked to uncertainties which are inherent in different components of the modeling chain. The main objective of this study was to investigate the role of precipitation input data from a weather generator for extreme flood estimates. A hydrometeorological modeling chain consisting of a multi-site weather generator (GWEX) at an hourly time scale, a rainfall-runoff model (HBV) and a hydrologic routing model (RS Minerve), was implemented, using different parameterizations of GWEX. While the sensitivity to the altered precipitation inputs was not uniform across the selected catchments due to their different physiographic characteristics, we found that the uncertainty of flood estimates increased with increasing return period. In addition, the flood peaks were strongly affected when a bootstrapping of precipitation was performed and to a lesser extent when weather types (WT) were used to condition the parameters of GWEX. However, the latter seemed to reduce the spread of the uncertainty both in generated precipitation and simulated floods. Therefore, results suggested that precipitation inputs strongly contribute to the uncertainties of extreme floods. Accounting for uncertainty information enhances the usefulness of long continuous simulations and is essential as a context for applications including hydraulic engineering, spatial planning and safety assessments.

Keywords Uncertainty estimation · Continuous simulations · Extreme floods · Weather generator

1 Introduction

Floods are high-impact natural hazards that cause loss of human lives, environmental deterioration, and damage to infrastructure. To develop effective flood management strategies and preparedness measures, reliable flood estimates are necessary. Traditionally, floods and their impacts have been studied through statistical techniques that rely on historical observations of streamflow and, therefore, on observed hydrographs. However, the relatively short length of the observations limits the robustness of flood estimates from common extrapolation techniques, especially when focusing on rare floods. In addition, missing data and assumptions about the antecedent catchment conditions lead to insufficient representations of the underlying flood processes (Boughton and Droop 2003; Winter et al. 2019; Viviroli et al. 2022; Staudinger et al. 2024). Therefore, estimates of extreme floods are highly uncertain.

To overcome these limitations, a combination of stochastic weather generators (WGs) with hydrological models is often adopted (Grimaldi et al. 2013; Ullrich et al. 2021), known as continuous simulation (CS) (Lamb et al. 2016). Using CS has the advantage that assumptions related to the antecedent catchment conditions like soil moisture state, snowpack, as well as storage levels of lakes and reservoirs can be omitted (Blazkova and Beven 2002; Camici et al. 2011; Viviroli et al. 2022). The reason behind this is that very long synthetic precipitation time series serve as input to rainfall-runoff models to generate long time series of streamflow, pairing a wide range of antecedent conditions

✉ Eleni Kritidou
eleni.kritidou@uzh.ch

[1] Department of Geography, University of Zurich, Zurich, Switzerland

[2] Institute of Geography and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

[3] Univ. Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP, IGE, Grenoble, France

and precipitation event characteristics. From these time series of streamflow, flood statistics are extracted. Other advantages of CS include its applicability to ungauged catchments (Staudinger et al. 2024) and the ability to link the flood estimates to the underlying physical processes as space-time interactions are well captured (Falter et al. 2015).

Although CS has been proven a valuable tool for flood estimation (Blazkova and Beven 2002; Okoli et al. 2019; Pathiraja et al. 2012) and has numerous advantages, it is computationally expensive (Viviroli et al. 2022) and characterized by high complexity, which often leads to limited applicability in practical context (Grimaldi et al. 2022; Lamb et al. 2016; Rogger et al. 2012; Winter et al. 2019). According to Grimaldi et al. (2022) this is mainly attributed to challenges related to the WGs. Specifically, the authors highlight the difficulty first, of finding appropriate WGs for CS approaches, despite the general availability of models, and second, of effectively implementing WGs due to their complex conceptual basis and theoretical background. To this end, the selection of appropriate WGs is subject to many challenges and is influenced by the context in which they are employed.

WGs were first introduced in the 1980s (Sohrabi and Brissette 2021) for hydrological applications (Todorovic and Woolhiser, 1975; Ailliot et al. 2015), to reproduce the statistical characteristics of the observational records, usually precipitation and temperature. Details on different WG models can be found in Langousis and Kaleris (2014), Ailliot et al. (2015) as well as Yin and Chen (2020). While WGs have been adopted in various fields, a large body of work pertains to their use for the estimation of floods and their frequency (Blazkova and Beven 2004; Calver et al. 2009; Falter et al. 2015; Leander, 2005; Okoli et al. 2019; Winter et al. 2019; Viviroli et al. 2022; Staudinger et al. 2024). Researchers have also used WGs in climate impact studies (Khalili et al., 2011; Kilsby et al. 2007; Peleg et al. 2017; Steinschneider et al. 2019; Ayele et al. 2024; Hawkins & Sutton, 2011; Moraga et al. 2021, 2022; Najibi et al. 2024; Sharafati et al. 2020.; Yuan et al. 2021) as well as for downscaling purposes (Fatichi et al. 2011; Keller et al. 2015, 2017) or long-term streamflow forecasting (Hwang et al. 2011; Sohrabi and Brissette 2021).

A systematic review of flood frequency analyses using weather generators can be found in Boughton and Droop (2003). For example, a combination of a multi-site weather generator with a hydrological model has been used for flood estimation by Winter et al. (2019). Specifically, the authors compared the continuous simulation approach with flood frequency analysis based on discharge records as well as a design storm approach for 16 Austrian catchments. They concluded that the fully continuous modeling approach gives robust results compared to the design storm approach and captures a wide range of runoff volumes. Similarly, Viviroli et al. (2022) combined a multi-site WG with the HBV hydrological model to estimate rare floods in the Aare River basin. It is noteworthy that in their study, hydrological routing is used to implement simplified representations of main river channels, floodplains, lakes and reservoirs. The results indicate that overall, CS performs well even for the estimation of low-probability events, despite uncertainties related to the extrapolation from observations and the choice of the weather generator.

According to Chen et al. (2016) the effectiveness of a WG should be assessed through catchment-scale hydrological modeling. This is because precipitation inputs play a critical role in hydrological modeling and forecasting (Bárdossy and Das 2008) and contribute to a certain degree to the total uncertainty of the modeled results. In many climate change impact studies, where regional climate models together with WG and hydrological models are implemented, the sources of uncertainty are separated into scenario uncertainty, model uncertainty and stochastic uncertainty (Hawkins and Sutton 2009, 2011; Moraga et al., 2022) Similarly, in CS approaches model uncertainty and stochastic uncertainty are fundamental for understanding the hydrological response and the reliability of flood estimates. Therefore, many researchers focus on understanding the role of WGs through different approaches. For example, Li et al. (2013) thoroughly examined the ability of six different precipitation distribution models to simulate extreme events in two Canadian catchments. Specifically, the models performed differently on a daily scale, with three-parameter precipitation models showing the best performance both for simulating extreme precipitation and driving the hydrological model. Similarly, Breinl (2016) tested three WGs of different complexity coupled with a lumped hydrological model for two Alpine catchments in Austria and France. From the hydrological outputs, he concluded that statistical approaches and long CS performed similarly at a daily scale. Legrand et al. (2024) assessed the ability of a multisite WG to produce hydrologically relevant weather scenarios for the upper Rhône, an alpine river where the interplay between weather variables in both space and time is determinant. They found that the simulated discharges are in good agreement with the reference ones, provided that the weather scenarios are bias-free for precipitation and temperature. These studies show that WGs have successfully been integrated in hydrological studies with CS and highlight the importance of understanding their impact on the overall robustness of the approach.Generally, a large number of studies have discussed the critical role of uncertainties in long CS (Blazkova and Beven 2002, 2004; Arnaud et al. 2017; Winter et al. 2019; Beneyto, Aranda, et al., 2023; Beneyto et al. 2023; Vesely et al. 2019; Viviroli et al. 2022; Legrand et al.

2024) however, only a few of these studies have sufficiently explored and dealt with the quantification of uncertainties in different parts of the modeling chain. For instance, Arnaud et al. (2017) employed an experimental approach to estimate the uncertainties of extreme flood quantiles using an event-based method. Specifically, the authors used a simulation-based flood frequency analysis approach called SHYREG to estimate uncertainties from the rainfall generator and the hydrological model both independently and combined. When the rainfall generator was taken as the only source of (parameter) uncertainty, the authors found that the 2-years, 10-years, 100-years and 1000-years confidence interval amplitude of discharge was, in most of the cases, two times higher than those of rainfall. Another finding of this study was that the uncertainties of the method were inversely related to the basin size and the length of the observational records used for model calibration. A recent study by Beneyto et al. (2023b) focused on the uncertainties from the WG only and analyzed how different precipitation regimes, climate extremality, or catchment characteristics influence the uncertainty of the flood quantile estimates. The study showed that the short length of observations is a major source of uncertainty in estimating high quantiles. It also highlighted the impact of the ξ parameter, which controls the tail of the precipitation distribution, to the overall WG's performance and to the uncertainties of discharge quantiles.

Given the above, the literature pertaining to CS suggested that a key source of uncertainty is the one related to WG parameters estimation, which has shown to be higher when the length of the observational records used is limited. Equally important is the uncertainty due to the stochastic generation process. Conducting multiple runs to generate precipitation provides different time series, resulting in varying discharge characteristics. The stochastic nature of the generation process is usually less evident when first-order statistical characteristics (such as mean precipitation or the probability of wet days) are examined, while it becomes more critical when focusing on extreme events where higher-order statistics are used (e.g., maximum length of wet spells or the coefficient of variation).

However, none of these studies has focused on the estimation of uncertainties from the WG for very rare events (i.e., return periods > 1,000 years), as these applications are computationally demanding and are often applied in a small number of regions and at small timescales.

In this study we employ a framework of CS with a hydrometeorological modeling chain, recently applied by Viviroli et al. (2022), to unravel the uncertainties of simulated extremes due to the WG (GWEX). Although every part of this modeling chain has been thoroughly assessed and compared with other approaches (e.g., a second independent WG, PMP-PMF (probable maximum precipitation–probable maximum flood) estimates, flood enveloping curves), a comprehensive framework to fully quantify uncertainties was not applied due to the considerable computational demand (Viviroli et al. 2022). Following these footsteps, we attempt to expand this approach towards evaluating the impact of GWEX to the uncertainties of the extreme flood estimates for several large Swiss catchments (>450 km²) with a wide range of physiographic characteristics. We do so by conducting two experiments: (a) we apply a bootstrapping of precipitation observations from which we get an ensemble of 10 members which are used to parameterized GWEX. The Extended Generalized Pareto Distribution (EGPD) which is used in GWEX has a heavy tail which includes inherent uncertainties and can lead to highly variable simulated precipitation intensities if modified. With the 10 ensemble members we succeeded different GWEX parameterizations that have different distribution tails. Accounting for the sensitivity of the tail behavior is crucial when modeling floods with large return periods (Papalexiou et al. 2013). Then, we (b) condition specific parameters of GWEX to weather types. Ultimately, we use the generated precipitation from (a) and (b) to run hydrological simulations and compare with the reference scenarios, when GWEX was parameterized based on the observations. This wide range of precipitation intensities allows us to assess the uncertainties of extreme flood estimates due to the GWEX.

2 Study catchments and data

We considered nine Swiss catchments that cover a wide range of physiographic characteristics. Switzerland is characterized by a complex topography with two mountain ranges (Alps and Jura) and the Central Plateau (Horton et al., 2022). The Alps are influenced by different climate contexts, often with a marked difference between northern and southern part (Schmocker-Fackel and Naef 2010a). The study catchments are located in different regions and characterized by high variability of meteorological and geomorphological conditions, with small-scale climate patterns (Schmocker-Fackel and Naef 2010a). There are considerable differences in precipitation and temperature, which in the Alps, according to Schmocker-Fackel and Naef (2010b), can be pronounced even in valleys that are close to each other and between valleys and mountain ridges. As a result, flood events of different magnitude and characteristics are expected. A classification scheme by Diezig and Weingartner (2007) and Sikorska et al. (2015) distinguishes six major flood types based on data from nine Swiss catchments. Those are flash floods, short-rainfall floods, long-rainfall floods, rain-on-snow floods, snowmelt floods and glacier-melt floods. With our catchment selection, we

Table 1 Properties of the study catchments sorted by decreasing area

River	Gauging site	Area (km^2)	Mean elevation (m a.s.l.)	Glacierisation (%)	Regime type[a]
Aare	Bern, Schönau	2965	1584	7.65	ng
Aare	Thun	2459	1739	9.22	ng
Thur	Andelfingen	1702	773	0	ps
Maggia	Locarno, Solduno	927	1534	0.47	npm
Sarine	Broc, Château d'en bas	636	1501	0.49	ndt
Inn	S-Chanf	616	2460	7.89	bgn
Thur	Jonschwil, Mühlau	493	1026	0	nppa
Kander	Hondrich	493	1846	6.69	ng
Kleine Emme	Emmen	478	1058	0	nppa

[a]Runoff regime type according to Weingartner and Aschwanden (1992), ng: nivo-glaciaire, ps: pluvial superieur, npm: nivo-pluvial méridional, ndt: nival de transition, bgn: b-glacio-nival, nppa: nivo-pluvial pre-alpin

covered most of these flood types. Besides this, we considered the catchment size as another key factor for hydrological response times and flood generation mechanisms. The area of the selected catchments varied from 478 km^2 to almost 3,000 km^2, with mean elevations from 773 m a.s.l. to 2,460 m a.s.l. (see Table 1 for details).

Furthermore, for 4 out of the 9 catchments there was a need for hydrological routing because they are regulated, namely for the Aare (at Bern and at Thun), Inn (at S-Chanf), Thur (at Thur Adelfingen) and Maggia (at Locarno) river basins (Fig. 1 with green color). It must be noted that many catchments in Switzerland are subject to human-induced changes that strongly influence the natural flow. In our case, this especially concerns the Aare and Maggia river basins. In the Aare river basin, there are two pre-alpine, artificially managed, natural lakes (lakes Briez and Thun) (Viviroli et al. 2022) and a relatively flat floodplain (Zischg et al. 2018). Besides this, several hydropower dams located in the headwaters of the Aare River play a critical role in altering streamflow (Wüest et al. 2007.; Viviroli et al. 2022). In the Maggia river basin, several operating rules related to reservoirs strongly influence the natural flow conditions. Specifically, there are eight hydropower reservoirs (Sambuco, Zött, Robiei, Peccia, Cavagnoli, Sfundau, Naret, Palagnedra) in the uppermost part of the Maggia catchment. Those are considered in the hydrological routing model, which allows us to include the corresponding regulation schemes for our estimates. Overall, the Aare, Inn, Thur and Maggia river basins are subdivided into 18, 8, 5 and 4 sub-catchments, respectively, which have an area between 18.7 and 1085 km². The remaining 4 catchments Kander, Kleine Emme, Thur (Jonschwil), and Sarine were not divided into sub-catchments. Rather, each was treated as one single unit by the hydrological model.

For this application, precipitation and temperature data were obtained from weather stations in Switzerland and neighboring countries. The Swiss station network is operated by MeteoSwiss, whereas stations in the neighboring countries are operated by several national and regional agencies. The meteorological datasets for precipitation comprised 1176 stations, of which 500 are located within Switzerland. The remaining 673 stations operate in the neighboring countries Austria, Italy, France and Germany. The observations were available in daily resolution for the period 1930–2019. The period 1930–1989 included only daily continuous records, whereas for 1990–2019 hourly records are available. The hourly records were used to disaggregate precipitation from daily to hourly as in Viviroli et al. (2022). Temperature observations were available in daily resolution at 26 stations (1930–2019) and in hourly resolution at 65 stations (1990–2019). For the calibration and evaluation of the hydrological model, hourly discharge observations were available for the stations operating at the study catchments from the Federal Office for the Environment. The record

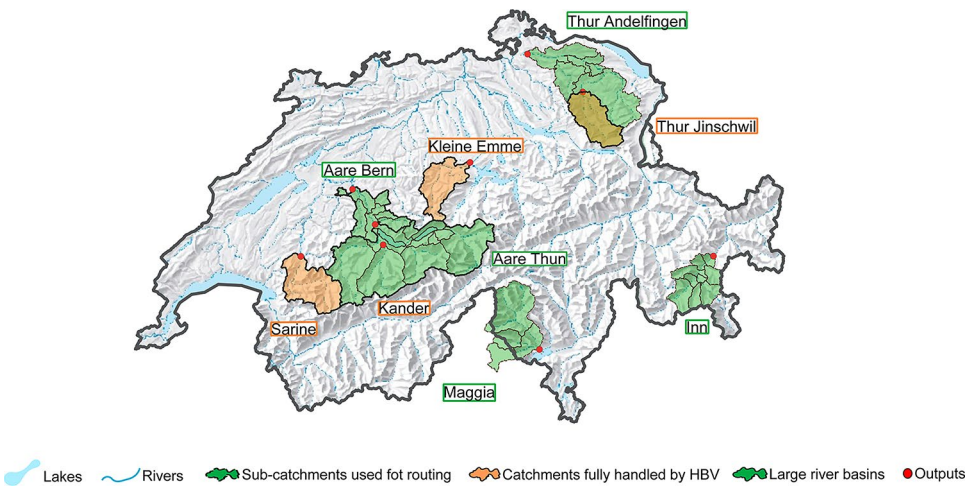

Fig. 1 Map of Switzerland with the four large river basins, from which green color depicts the sub-catchments used for routing and orange color the selected large catchments that are only modeled with HBV

length ranges between 6 and 96 years for the period 1923–2019 and the median record length is 45 years.

3 Methods

The proposed framework relies on a hydrometeorological modeling chain with long CS and has previously been used in the EXAR (Extreme flood events on the River Aare) project for hazard assessment (Andres et al., 2021; Viviroli et al. 2022). The approach consists of three main steps: (1) stochastic generation of weather scenarios (2) hydrological modeling (3) hydrological routing (Fig. 2). Within this framework, we want to understand how flood estimates depend on the parametrization of the WG. To do so, we conduct (a) an analysis of the precipitation scenarios at the stations (b) a bootstrapping experiment of precipitation observations used to generate mean areal precipitation, (c) a weather-typed conditioning of specific GWEX parameters and finally (d) a correlation analysis to unravel the importance of different factors to the uncertainty of flood estimates. Thus, in the following, we describe in detail the weather generator, the hydrological model, the hydrologic routing and then the different experiments considered to assess the effect of the WG parametrization on flood estimates.

3.1 Weather generator

GWEX is a stand-alone WG, first introduced by Evin et al. in 2018 and further developed and adapted since then. Here, it is considered a key component firstly because of its suitability for extreme precipitation events and secondly because it has been previously used as a basis for hydrological modeling and hazard assessment in the Aare river basin (Viviroli et al. 2022). It is a multi-site, two-part stochastic weather generator, which provides long simulations of precipitation and temperature preserving the statistical characteristics of the observations. In this study, GWEX relies on daily observed weather from 1930 to 2019 to produce outputs for all study catchments. Following the structure of Wilks (1998), precipitation occurrence at a single site is simulated by a two-state Markov-chain. The spatial structure of occurrences is modeled by an unobserved Gaussian process. To adequately simulate the wide range of positive (non-zero) precipitation intensities, the marginal distribution is based on the extended Generalized Pareto (GP) Type III distribution (E-GPD) (Naveau et al. 2016). Generally, GP distributions have been widely used in hydrological applications (for details see Langousis et al. 2016). EGPD was first introduced by Papastathopoulos and Tawn (2013) due to its suitability to describe the behavior of the upper tail, linked with high rainfall intensities. Thus, the full spectrum of precipitation intensities is modeled without the need for threshold selection (Evin et al. 2018; Haruna et al., 2021; Naveau et al. 2016). The cumulative distribution (CDF) is given by

$$F(x) = P(X \leq x) = G\left[\mathrm{H}_{\xi}\left(\frac{x}{\xi}\right)\right] \quad (1)$$

where X is the random variable representing the positive daily precipitation intensity according to the EGPD distribution. The transformation function G is the CDF that provides a smooth transition between both upper and lower tail of the distribution, which is compliant with the extreme value theory and

$$H_{\xi}\left(\frac{x}{\sigma}\right) = \begin{cases} 1-\left(1+\xi\frac{x}{\sigma}\right)_{+}^{-1/\xi} & if\ \xi \neq 0 \\ 1-\exp\left(-\frac{x}{\sigma}\right) & if\ \xi = 0 \end{cases} \quad (2)$$

with $\alpha_{+}=\max(\alpha,0)$. Here, the function G is defined as:

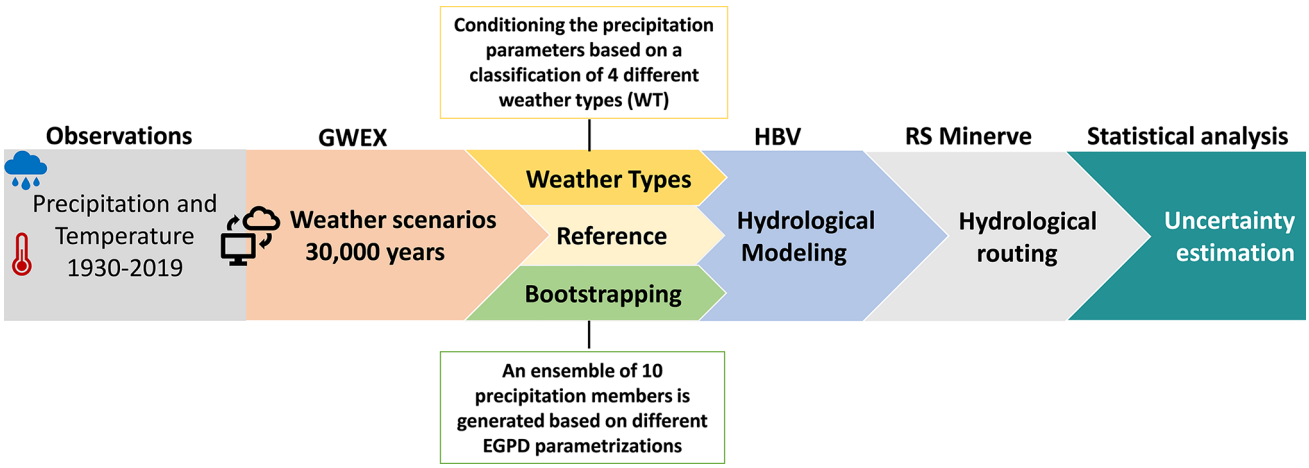

Fig. 2 Schematic illustration of the framework of long continuous simulations and the experiments conducted

$$G(v) = v^{\kappa} \tag{3}$$

Therefore, the model is given by:

$$F(x) = \left[H_{\xi}\left(\frac{x}{\sigma}\right)\right]^{\kappa} \tag{4}$$

The model has a three-parameter set {κ, ξ, σ}, estimated for every station: $\kappa > 0$ controlling the lower tail, $\xi \geq 0$ controlling the upper tail and $\sigma > 0$ being the scale parameter. The shape parameter ξ is estimated based on advanced regionalization methods (Evin et al. 2016; Haruna et al., 2021). The estimation of the remaining parameters of the marginal distribution relies on the maximum likelihood method. The spatial and temporal dependencies of the precipitation intensities are modeled simultaneously by a first-order multivariate autoregressive model. To account for the seasonality of precipitation intensity parameters are estimated based on a 3-month window. Thus, we have a set of parameters for every month, leading to a total number of twelve sets of parameters. Furthermore, the at-site tail dependence of extremes, as well as the innovations in the generation process, are described by a Student copula. The outputs of the weather generator are finally disaggregated to hourly scale based on hourly observations from 1990 to 2019 using the method of analogs (or method of fragments). These analogs are the days that are similar in terms of season and class of intensity (Evin et al. 2018, 2019; Staudinger et al. 2024; Viviroli et al. 2022). Thus, a selection of analog days is made by comparing the daily precipitation field of each day in the target simulation period with the daily precipitation field in the period 1990–2019 (for which we have hourly historical at-site records). A precipitation field consists of several stations within a specific radius so that the spatial and temporal structures are preserved. Using a distance criterion for comparison root mean square error, the 100 best analogs are retained for each day. Then, for every day in the simulation period the hourly values of the best analog day are used. Finally, the length of the modeled time series used in this study is 30,000 years at an hourly resolution. Further details about GWEX and the disaggregation can be found in Evin et al. (2018, 2019).

3.2 Hydrological modeling

HBV is a semi-distributed precipitation-runoff model for continuous runoff simulations (Bergström 1992; Seibert and Bergström 2022). The main reasons why HBV was used in this study were that (a) it has extensively been used in flood studies and proven to be a suitable modeling tool in Switzerland (Horton et al., 2021; Viviroli et al. 2022), (b) it well represents the variability of hydrometeorological regimes found in the alpine environment (Horton et al., 2021) and the complex snow processes, (c) it can efficiently be used for long continuous simulations (Viviroli et al. 2022) and (d) it does not require an excessive amount of inputs.

Here, the model was applied at an hourly time step, using the HBV-96 parametrization, which introduces a non-linear response function in the upper zone (Lindström et al. 1997). The sub-catchments are separated into different elevation zones (Seibert and Vis 2012). Long time series of mean areal precipitation (MAP) and mean areal temperature (MAT) from the WG serve as inputs to simulate catchment runoff. To represent the underlying physical processes, the standard version of the model consists of four major routines:

- The snow routine, where a degree-day method is used to simulate snow accumulation and snowmelt.
- The soil moisture routine, where the split between actual evaporation and groundwater recharge depends on the continuously simulated water storage.
- The groundwater routine, where simulated runoff is a function of groundwater storage.
- The routing routine where the streamflow routing in the catchment is based on a triangular weighting function.

For glaciated catchments, a glacier routine has recently been integrated (Seibert et al. 2018). In our study, the routine was used for catchments with a glacier coverage of 1% or higher. The four main routines entail 11 parameters that were calibrated while the glacier routine adds another 2 parameters. We use a 10-year warm-up period to initialize the model's storages (Seibert and Vis 2012).

The model calibration was based on a genetic algorithm (Seibert 2000) where the first step is to generate an optimized parameter set by selecting and recombining initial, randomly selected parameter sets. We accounted for parameter uncertainty by calibrating every sub-catchment 100 times. From this ensemble of 100 parameter sets, three parameter sets were identified that represent a low, intermediate and high range of simulated floods, respectively, based on a percentile approach first introduced by Sikorska-Senoner et al. (2020) and then applied by Viviroli et al. (2022). The three representative parameter sets were used to run the full modeling chain for the reference scenarios and the two experiments, but due to the high computational cost of the simulations, only the median parameter set was used for further comparisons. To assess the performance of the parameter sets, the Kling-Gupta efficiency with extra weight on higher flows (50–80% percentiles) was used. Lastly, for the regionalization of 25 ungauged sub-catchments, 33 calibrated sub-catchments were used as donors for parameter transfer. Following Kauzlaric et al. (2021), the calibrated sub-catchments were clustered using the runoff regime as a

discriminant and two donor sub-catchments were selected for every ungauged catchment minimizing the Euclidean distance in the sub-catchments' attribute space and transferring 50 randomly selected parameter sets from each donor.

3.3 Hydrological routing

RS Minerve is a semi-distributed model that simulates the free surface run-off and its propagation (García Hernández et al. 2020). Like HBV, RS Minerve has been widely used in Switzerland because of its suitability to include not only hydrological processes but also represent complex hydraulic networks (Horton et al., 2022; Viviroli et al. 2022; Sharma et al., 2021). For example, it has been implemented in the Valais for real-time flood forecasting (García Hernández et al. 2014) as well as for comparison with a long short-term memory model (Frank et al. 2023). The main function performed here was to route discharge according to the configuration of the river network and the flow time estimated by the aid of hydraulic simulations, as well as monitor the flow at relevant nodes of the main river reaches. In our study, RS Minerve was further used to implement lake retention and (where relevant) lake regulation, reservoir operation and floodplain retention.

3.4 Experimental set-up

Our experimental set-up focuses on the weather generator which is the first component of the modeling chain. In this study, GWEX was used to generate precipitation time series at hourly resolution and multiple sites. Specifically, it was first parameterized based on the observational records available for the whole 1930–2019 period to generate 30 precipitation scenarios of 1,000 years. Those are the reference scenarios. To explore the impact of GWEX on flood estimates, it was subjected to different experiments (described in detail in Sect. 3.4.1 and 3.4.2) while the other two components (HBV and RS Minerve) remained unchanged. This experimental framework allows us to identify only the uncertainty from different parameterizations of the weather generator and understand how precipitation uncertainty translates in the discharge quantiles. The reference scenarios serve as a basis for the comparison with both experiments. To do so, we used criteria such as the standard deviation (Eq. 5), coefficient of variation (Eq. 6), confidence interval amplitude (Eq. 7) as well as the relative change of means (Eq. 8) for both precipitation and discharge maxima values:

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{\mathrm{N}}} \tag{5}$$

where N is the size of the population, x_i each value of the population (here precipitation or discharge) and μ the population mean and.

$$CV = \frac{\sigma}{\mu} \tag{6}$$

$$CIA = \frac{q_{95\%} - q_{5\%}}{q_{50\%}} \tag{7}$$

where $q_{95\%}$, $q_{5\%}$ and $q_{50\%}$ the corresponding quantiles of discharge and precipitation.

$$\mathrm{RC} = \frac{y_{experiment} - y_{reference}}{y_{reference}} \times 100 \tag{8}$$

where y is the mean precipitation or discharge for a given return period.

3.4.1 Experiment 1

The first experiment relied on a resampling of precipitation observations with replacement (bootstrapping) to produce the database used for the estimation of GWEX parameters. The approach has been widely used in hydrological studies. The bootstrapping was applied by splitting the precipitation time series into blocks of 7 days which are then randomly resampled. The reason behind the selection of a 7-day period is the importance of preserving the spatial and temporal correlation of the values while randomly resampling. As a result, for every station, we have 1,000 resamples with equal likelihood to occur. To keep the computational cost manageable, an ensemble of 10 resamples was randomly selected for each station. The dates assigned to every bootstrapped sample were identical for each station. Then, long time series of MAP were computed for each of the ensemble members, as described in Sect. 3.1. While the methodology for the generation of MAP remains the same, the parameters of the EGPD distribution that models the marginal distribution of positive precipitation intensities at the gauges vary from member to member. To this end, GWEX was fitted and run with those 10 sets of new EGPD parameters. In this way, we use equiprobable scenarios to assess the uncertainty of the simulated flood peaks in response to the stochastic precipitation variability, which according to Paschalis et al. (2014) is a strong advantage compared to studies where only observations are used. Therefore, simulations were performed for every ensemble member and areas with large precipitation differences were found.

Table 2 Description of the CAP9 WT as given in Weusthoff (2011), grouped into four types as used in this study (CAP4)

WT1	[4] East indifferent [5] High Pressure over the Alps [8] High Pressure over Central Europe
WT2	[1] NorthEast, indifferent [6] North, cyclonic
WT3	[2] West-SouthWest, cyclonic [3] Westerly flow over Northern Europe
WT4	[7] West-SouthWest, cyclonic [9] Westerly flow over Southern Europe, cyclonic

3.4.2 Experiment 2

Generally, WGs can be designed to be conditioned on WT, meaning that the parameters are estimated separately for specific weather patterns or types. WGs often model precipitation based on wet and dry days. Thus, every day belongs to a given WT. Since precipitation is governed by various atmospheric processes and can be highly spatially variable, conditioning the parameters of GWEX to the WT can enhance the physical interpretation by discerning recurrent dynamical patterns for every region. In addition, it can shed light on the different characteristics of observed precipitation as well as the spatiotemporal correlations between sites (Stucki et al. 2012).

For this second experiment, 4 WTs were considered. The 4 WTs were derived from the CAP9 weather-type classification introduced by MeteoSwiss (Weusthoff 2011). CAP9 stems from a principal component analysis of ERA40 reanalysis (Uppala et al. 2005) based on mean sea level pressure and includes the following WTs: (1) NorthEast, indifferent (2) West-SouthWest, cyclonic, flat pressure (3) Westerly flow over Northern Europe (4) East, indifferent (5) High Pressure over the Alps (6) North, cyclonic, (7) West-SouthWest, cyclonic (8) Pressure over Central Europe (9) Westerly flow over Southern Europe, cyclonic. In the present work, 4 WTs were considered instead of the 9 original ones to reduce the computational demand and ensure robust fitting. The 4 WTs have been defined by grouping these 9 WTs (as summarized in Table 1) thanks to hierarchical clustering analysis.

In the present work, not all GWEX parameters were conditional on WT. Conditional parameters are those controlling at-site occurrence processes and multisite occurrence as well as intensity dependency. For their estimation all days of the observational period are first assigned to a weather type based on information about the atmospheric conditions. GWEX parameters for a given WT are then estimated from the observations (of both precipitation and temperature) of all days available to this WT. Parameters are also estimated on a monthly basis. For every region, there are thus in total 48 parameter sets estimated, corresponding to 4 weather types for every month (Table 2).

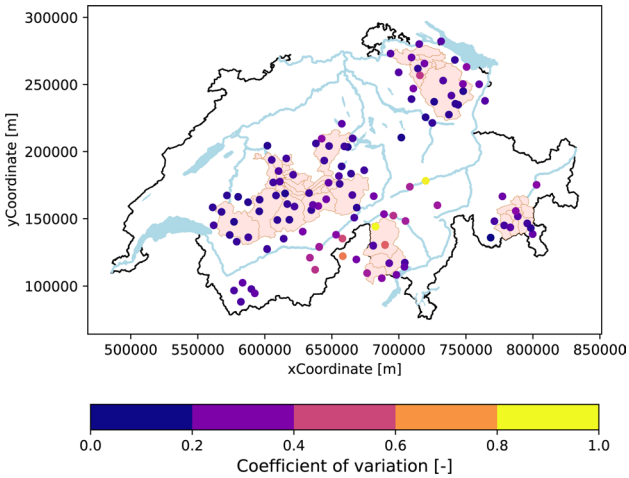

Fig. 3 Coefficient of variation of the maximum simulated hourly precipitation for the 10 ensemble members with different parameterizations. The light red areas indicate the four large river basins (Aare-Bern, Inn, Maggia, Thur) with routing using RS Minerve. Coordinates are given in the Swiss projection system (CH1903/LV03)

4 Results

4.1 Station-based analysis

For the individual stations, we analyzed the maximum hourly precipitation from the bootstrapped scenarios (10 ensemble members). The coefficient of variation of the 10 values obtained from the 10 ensemble members was calculated for the highest precipitation event observed. As shown in Fig. 3, higher values occurred in the Maggia region and lower values in the Aare, Inn and Thur. Specifically, in the Maggia region, most of the coefficients of variation of precipitation were higher than 0.4, and two of them exceeded the maximum calculated value which equals 1. While for the remaining three regions the number of stations taken into account for the interpolation to mean areal precipitation is higher, only one station showed a coefficient of variation between 0.4 and 0.6 (Thur region). This finding indicates that higher uncertainties in the precipitation inputs occurred in the Maggia region. This clear difference is due to a combination of reasons. First, the highest precipitation amounts were found in the observations in this region. High, as well as extremely high precipitation values in Ticino could be explained by the climatology of the region which is affected by the warm air masses which flow from south to the north and experience orographic uplift in the Alps. In combination with the less dense station network, several challenges arise in estimating heavy precipitation and therefore the estimates are subject to considerable uncertainties (Frei and Fukutome 2022). Thus, the robust estimation of the parameters of GWEX becomes challenging as well and this translates

to high coefficients of variation for the high precipitation events.

4.2 Experiment 1: bootstrapping - EGPD parameterization

In this paragraph, we analyze the results of the hydrometeorological modeling chain for our test catchments. As has been found in previous studies with CS, varying the parameters of different components can affect the flood estimates in a very wide range. In our first experiment, we used 10 different ensemble members of the mean areal precipitation aiming to assess their impact on the hydrological outputs. The annual maximum flood (AMF) has been calculated for every catchment and plotted as one exceedance curve (of 30,000 years), in a Gumbel plot, with 11 colored lines representing the 10 bootstrapped ensemble members (Fig. 4) and the reference member. From Fig. 4 two key findings emerged. First, a larger spread among the 10 ensemble members occurs for higher return periods for all catchments, indicating increasing uncertainty. Second, the impact of different GWEX parameterizations on the AMF is highly variable among the test catchments due to the different physiographic characteristics. For example, the Maggia catchment stands out as particularly sensitive to changes in precipitation scenarios as the spread of the exceedance curves begins widening already in the range of 100 to 1,000 years return period. Specifically, beyond the 100-year return period there is an exponential increase in the simulated floods. These outcomes for the Maggia catchment are in line with our prior station-based analysis, where extreme precipitation events were found in many stations of the region.

Unlike the Maggia, there are catchments like the Aare (both at Bern and at Thun) that show less sensitivity to precipitation inputs. In this case, the 10 ensemble members provide a relatively small range of uncertainty for estimates corresponding to return periods of 10,000 years. This could be attributed to the presence of two regulated lakes, Lake Thun and Lake Brienz, that can to some extent act as a buffer to high precipitation events, attenuating flood discharge. Thus, the dispersion between the ensemble members is not evident even for higher return periods, and the overall estimates could be considered robust. Little sensitivity to variations has also been found for the Kander, which is a partly glaciated catchment. Since it is a pre-alpine catchment, characterized by a glacio-nival regime, rain-on-snow events are also expected. Moreover, snow cover contributes to the runoff generation mechanisms and might be the reason why

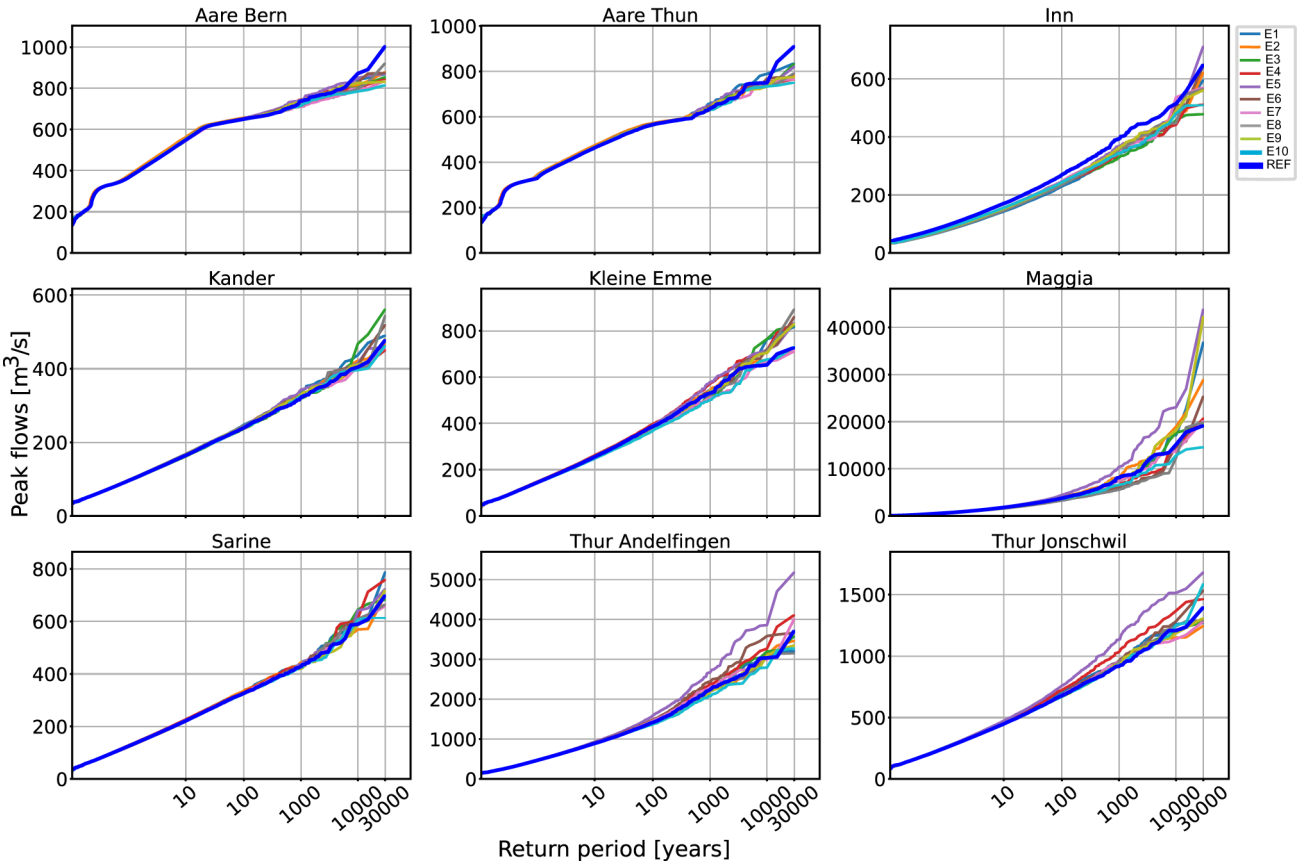

Fig. 4 Exceedance plots of AMF from 30'000 years of continuous simulation for the 9 large test catchments. The simulated flood peaks are based on the ensemble of 10 bootstrapped estimates (E1–E10) and the reference scenarios of the GWEX (REF.), shown as 11 lines with different colors

precipitation changes did not strongly affect the simulated extreme floods and uncertainty remained relatively small even for high return periods. The coefficients of variation for different selected return periods in our test catchments are displayed in Fig. 5. From the heatmap, higher return periods show larger coefficients of variation, indicating that the flood estimates are spread out over a wider range, thus reflecting higher uncertainty.

In an extension of the presented exceedance plots, we further split the 30,000 years of precipitation generated into 30 scenarios of 1,000 years. Thus, we focused on flood estimates of a maximum return period of 1,000 years using the 30 estimates that correspond to a given return period and aim to understand the statistical correlation between the precipitation changes and the simulated peaks. Results of this uncertainty analysis are shown for three example catchments in Fig. 6. The violin plots illustrate the distribution of the relative differences in mean precipitation (upper panels) and discharge (lower panels) for three selected return periods: 100 years, 300 years, and 1,000 years. The plots show that for the 100-year return period the relative differences are generally lower. The values are concentrated around the mean and the shape of the violin is approaching the normal distribution. The opposite behavior is depicted for the 1,000-year return period, where the differences are comparably higher, and the spread is larger. In this case, the violins are pointier. Thus, the lower return periods were linked to smaller differences in mean return level estimates and smaller spreads. In addition, the violin plots show that 100-year and 300-year estimates are more robust since there are no big outliers compared to the 1,000-year estimates. The fact that the same pattern occurs for both precipitation and discharge provides evidence that the variations of precipitation have a critical impact on the simulated flood estimates and suggests that precipitation is the dominant driver of differences for selected catchments.

From the results we conclude that there is an amplification of the relative changes from precipitation to discharge, which are more profound for Thur and Maggia and smaller for Kleine Emme. For instance, for the Thur catchment positive relative changes corresponding to the 300-years return period can reach up to 20%. This is translated to relative changes in discharge as high as 35% for the same return period. The same pattern is shown for the Maggia. There the relative changes for the 300-year return period range between −30% and 20% for precipitation and are translated to a range of -50–60% in simulated discharge. It is worth mentioning that the relative changes of discharge are almost 2.5 times higher than those of precipitation. Specifically, for precipitation relative changes are between −70% and 60%, which is then amplified to between −120% and 150% in simulated discharge. The results from Kleine Emme lead to similar conclusions.

The findings are in line with further analyses conducted for this experiment, where the relative changes between the reference scenarios and the ensemble members were computed. Figure 7 shows the variability of the ensemble members for precipitation (upper panel) and discharge (lower panel) for the corresponding return periods, for one example catchment (Maggia). From the boxplots, the most obvious finding that emerges is that the 1,000-year return period estimates exhibit the highest number of outliers. The results confirm that the higher the variations in precipitation input, the more prominent the variations in simulated discharge. Therefore, the less frequent events are associated with a higher degree of uncertainty. Interestingly, there is a similarity in the pattern of relative changes of annual maximum precipitation (AMP) and annual maximum flood (AMF). This means that when high relative changes occurred in

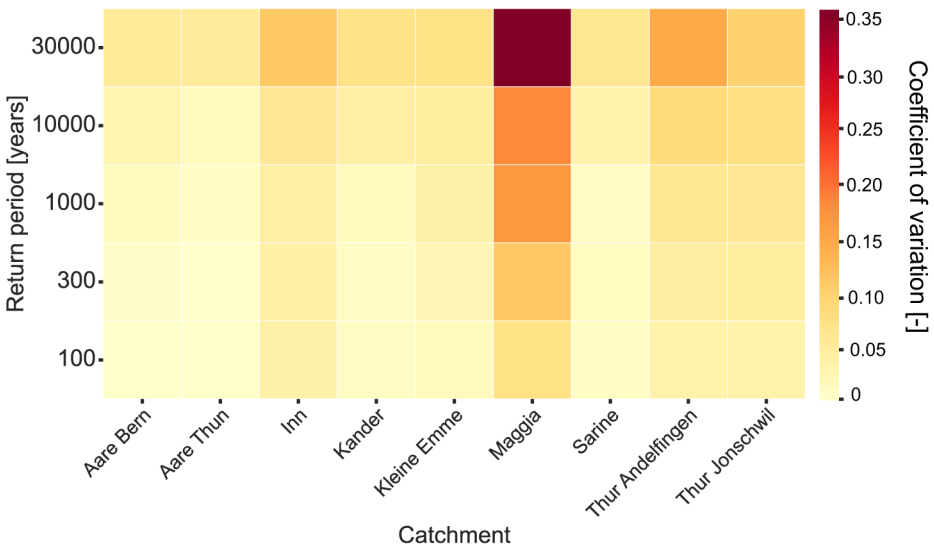

Fig. 5 Heatmap of coefficients of variation (a measure for the range of flood estimates) for 9 large test catchments for return periods of 100, 300, 1,000, 10,000 and 30,000 years. The coefficients of variation are derived from the ensemble of the 10 different bootstrapped members (and including the reference scenarios)

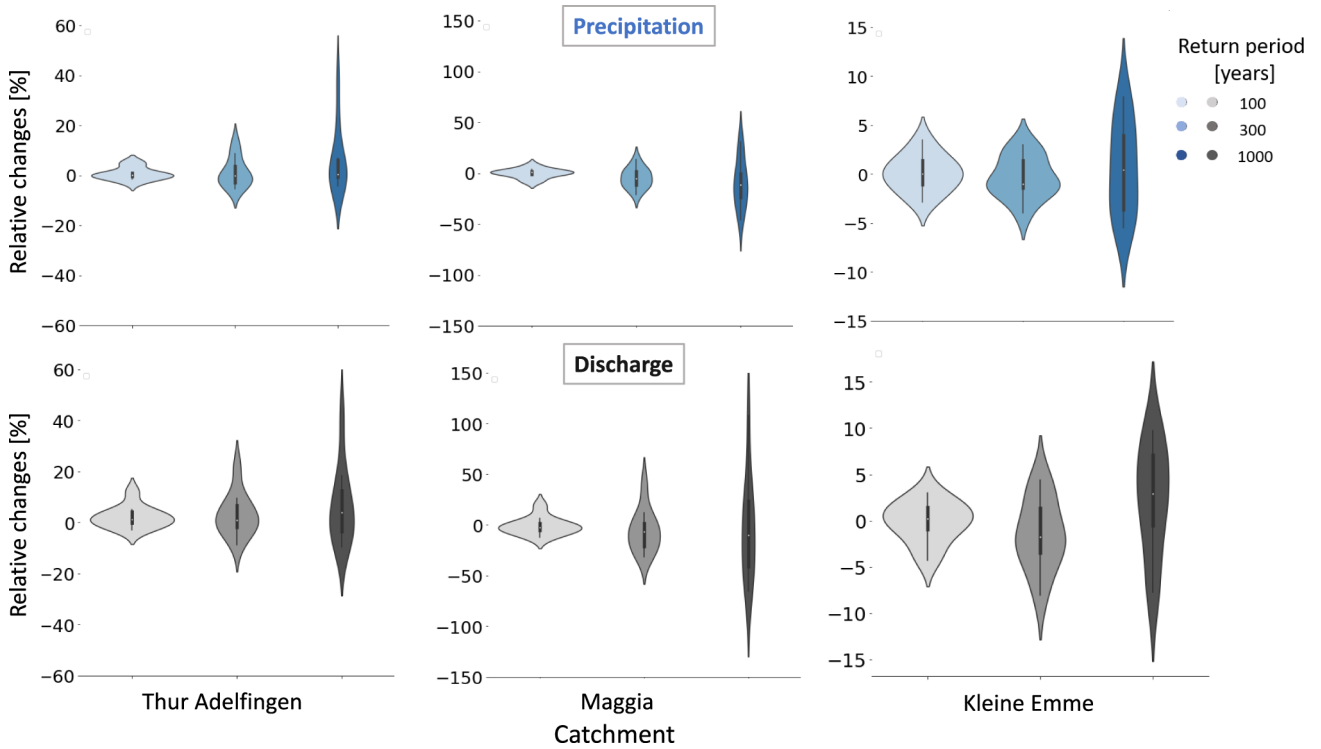

Fig. 6 Violin plots of the changes in mean precipitation (upper panels) and discharge (lower panels) between the reference scenarios and the 10 ensemble members. The mean has been estimated from the 30 scenarios of 1,000 years. The differences are shown for 100-year, 300-year and 1,000-year return periods for the Thur (Andelfingen), Maggia and Kleine Emme catchments

precipitation for a given ensemble member, the very member demonstrated also high relative changes in AMF. This general tendency is found in all our study catchments (not shown), but the order of magnitude of the estimated changes depends on the physiographic characteristics of the catchments. Thus, for some catchments, the changes in precipitation were found to be striking and associated with notable impact on the simulated extremes, while for others these variations were not strong enough or did not provide substantial differences in simulated extremes as shown in Fig. 4.

4.3 Experiment 2: weather-type conditioning (WTC)

In the second experiment we utilized a classification based on different WT to parameterize the weather generator. Again, a differentiated response of the individual catchments was found in the results. Figure 8 presents Q-Q plots for three example catchments based on the AMFs simulated by the weather-typed conditioned GWEX (y-axis) and the reference scenarios GWEX scenarios (x-axis). The quantiles of the weather-type conditioned scenarios tend to diverge markedly from the 1:1 line for the Maggia catchment, whereas for the Thur and Kleine Emme catchments, the deviation is less pronounced. Specifically, for the Thur only small differences from the reference scenarios' quantiles occur for values lower than 1,500 m^3/s, which approximately corresponds to a return period of 100 years. Above that point, the divergence is stronger, and the weather-type conditioned AMF tends to be slightly lower. For the Kleine Emme, the results indicate that the distribution of AMF derived using the reference member is similar to the distribution derived from the weather-type conditioned version of GWEX. Also, differences are limited for most of the other study catchments. On the contrary, a notable impact was found for the Maggia catchment. This could be attributed to a combination of factors that affect the estimation of GWEX parameters. First, the density of the station network is smaller compared to the other study regions and at the same time precipitation observations show a large number of extremes. Then, the classification of the data based on weather patterns and atmospheric conditions means that the number of days used for the estimation of GWEX parameters is reduced as a result of the grouping and that affects the statistical characteristics of simulated precipitation.

Similarly, as presented in the previous section, we split the simulated discharge into 30 scenarios of 1,000 years. The same example catchments are used to illustrate the difference between confidence intervals associated with the reference and the weather-type conditioned scenarios. Figure 9 shows the 99% confidence intervals (upper panels) of AMF. Generally, an asymptotic behavior towards extreme

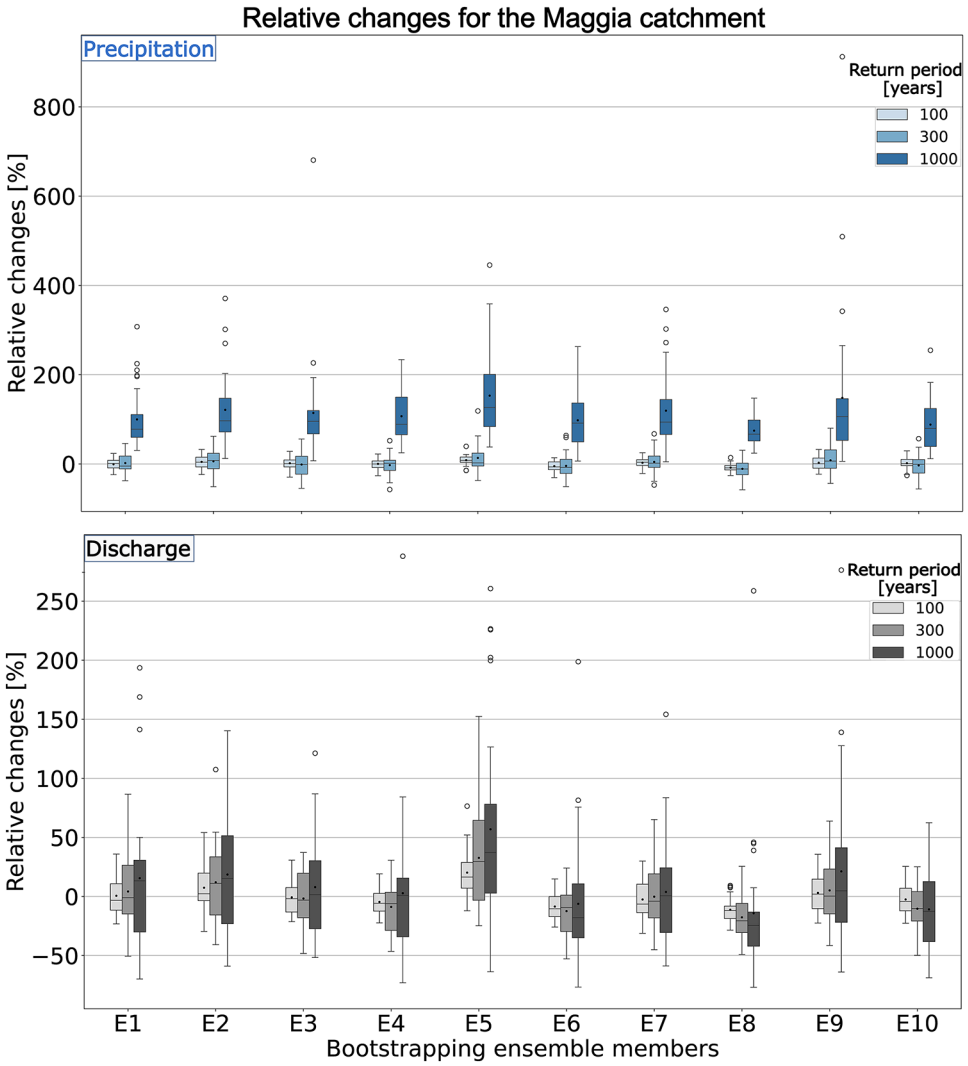

Fig. 7 Boxplots of relative changes of annual maximum precipitation (AMP) (upper panel) and AMF (lower panel) between the reference scenarios and the bootstrapped ensemble members: scenarios for 100-year, 300-year and 1,000-year return periods, for the Maggia catchment

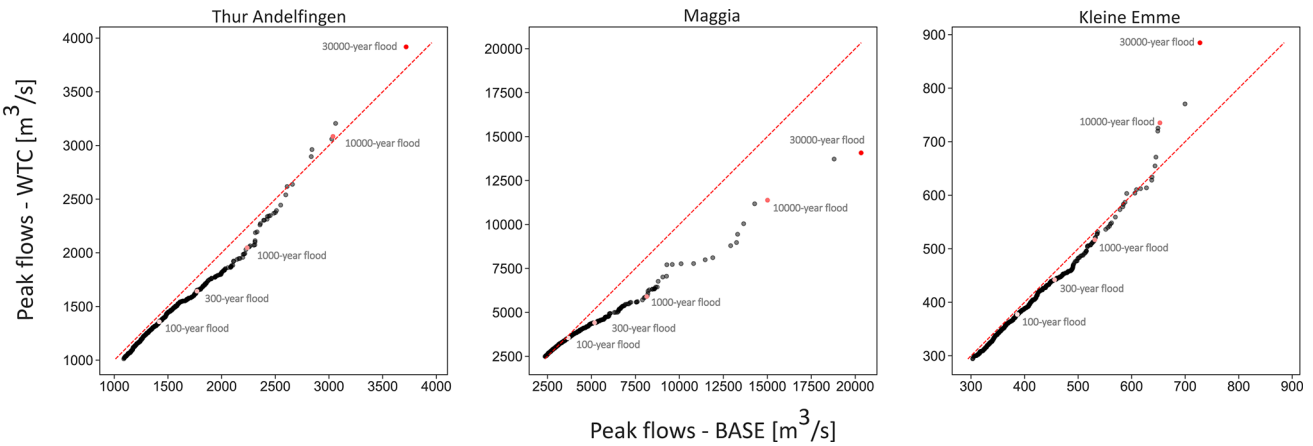

Fig. 8 Q–Q plots of the simulated AMFs based on 30,000 years from the reference scenarios on the horizontal axis and simulated AMF derived from the weather-type conditioning experiment on the vertical axis

frequencies can be seen. The simulated peaks are compared with extrapolations from the Federal Office for the Environment FOEN (Baumgartner et al. 2013) and show relatively good agreement for Thur and Maggia. However, for the Kleine Emme, there is a systematic underestimation of flood peaks both from the reference scenarios and the weather-type version. This is most likely caused by the challenges related to the generation of synthetic precipitation time series with GWEX in the comparatively small catchment of the Kleine Emme (478 km^2). Small-scale weather patterns like thunderstorms play a key role in this catchment and are most likely not captured adequately by the current weather generator which is designed for large catchments. To quantify the information on confidence intervals and better understand the impact of the weather type classification on the simulated peaks, the ratio between a 90% confidence interval and the median values is used as a criterion (confidence interval amplitude). The confidence interval amplitudes shown in the lower panels of Fig. 9 indicate a reduced uncertainty of the simulated flood peaks for the Maggia for all return periods when the weather types were used, compared to the reference scenarios. For the Thur and Kleine Emme there are no marked differences for the 100-year and 300-year return periods while for the 1,000-year return period, there is a small increase when the weather-type classification is applied, which might be attributed to a more realistic representation of some extreme precipitation events and corresponding floods.

Then, we compared AMP and AMF and calculated the relative differences. There was a clear variation in the relative differences for all example catchments (Fig. 10). While the relative changes in precipitation and discharge seem to be of a similar order of magnitude, for higher return periods they are increasing and characterized by more outliers. Most of the boxplots for the different return periods indicate a negative median, suggesting that the weather-type conditioning overall constrains the upper bound of extremes for our study catchments. However, positive relative changes up to 25% occur for precipitation values of Kleine Emme, probably attributed to the challenges related to the generation of precipitation scenarios for this catchment as mentioned further above. Nonetheless, weather-type conditioning seems to have a beneficial effect, as the data are grouped based on physical processes and therefore a wider range of events may be well represented. Again, the results suggest that the relative changes in AMF, although not very strong, follow the changes observed in AMP for all return periods.

5 Discussion

5.1 Extreme peak flows

Uncertainties in simulated extreme floods displayed pronounced heterogeneity among the test catchments. We found that certain catchments showed distinctly higher uncertainty than others. Most estimations from CS are in good agreement with observations, occasionally showing slightly higher or lower values, which was to be expected due to the different physiographic characteristics and availability of observational data. Estimates of extreme floods still include considerable uncertainties, for which we unravel the effect of precipitation variability. Notably, the tail of the precipitation distribution has a key role in the frequency and

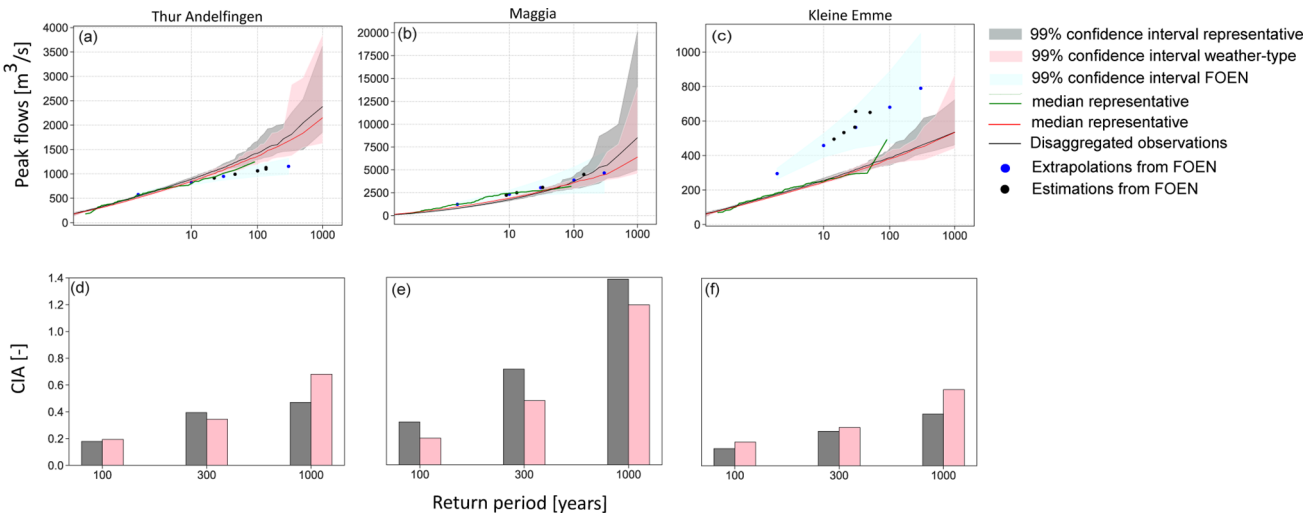

Fig. 9 Exceedance plots of the 99% confidence interval of the AMF from 30,000 years of CS for (**a**) Thur (Andelfingen) (**b**) Maggia and (**c**) Kleine Emme. Grey and pink colors correspond to results from the reference and weather-type conditioned scenarios, respectively. Green corresponds to simulation with disaggregated meteorology. Black points indicate the five highest observed peak flows (Baumgartner et al. 2013). Blue points are the extrapolation of peak flow records according to Federal Office of Environment (FOEN) with confidence interval (Baumgartner et al. 2013). The lower panels (**d**, **e**, **f**) show the comparison of the confidence interval amplitude (CIA) criterion

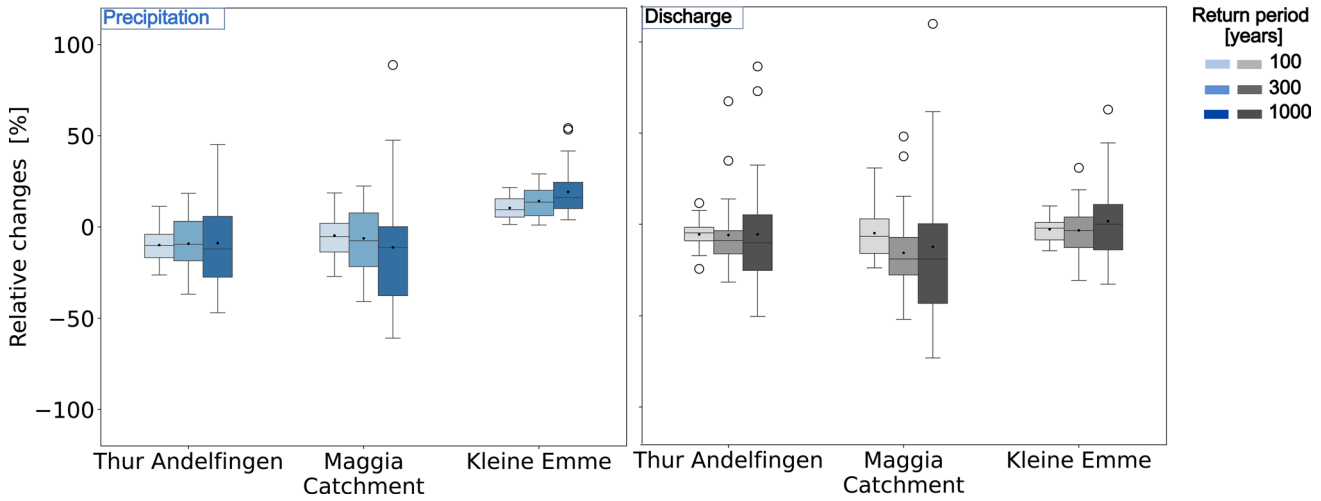

Fig. 10 Boxplots of relative changes of AMP (left panel) and AMF (right panel) between the reference and weather-type conditioned scenarios for 100-year, 300-year, and 1,000-year return periods, for 3 study catchments

magnitude of extreme events (Papalexiou et al. 2013) and has been suggested to govern the distribution of flood peaks. In a recent study by MacDonald et al. (2024), a stochastic weather generator was used to produce long time series of precipitation further used as input to a hydrological model. The authors used rainfall observations from a station in Germany and varied the ξ parameter, which describes the heaviness of the tail of the distribution, while keeping the scale and lower shape parameters as fitted. In this way, the authors studied the effect of different degrees of extremality on the flood peak distribution. The main conclusion of their study is that beyond a certain return period the tail of the rainfall distribution is the dominant factor of the tail of the flood peak distribution. In our approach, we account for the effect of precipitation distribution with our first experiment, where a bootstrapping approach is applied. Varying only the ξ parameter might be interesting because of its critical role on the tail behavior, but in this study all three EGPD parameters {κ, ξ, σ} have been varied together to achieve a robust fitting at each station. To this end, sampling precipitation with the bootstrapping approach helps us identify different possible degrees of extremality and at the same time a wide range of uncertainty in the synthetic precipitation time series, which is then reflected in flood estimates.

In Ticino, we simulated exceptionally high peak discharges, explained by the high generated precipitation intensities. Studies by Frei et al. (2000) and Jasper et al. (2002) have highlighted the challenges related to the remarkably high precipitation amounts in the given region and that even small errors in precipitation amounts can lead to important errors in the simulated runoff. In addition, Haruna et al. (2021) compared different regionalization methods to improve at-site estimates of daily precipitation for all stations in Switzerland based on different criteria. The authors found the highest 100-year precipitation estimates at the stations of Ticino, for all seasons. Since these exceptional events are the basis for the parameterization of the weather generator for the reference scenarios as well as for the different experiments applied, the uncertainties that are inherent propagate to the flood estimates. The underestimation of peak discharge found for the Kleine Emme is probably due to the relatively small basin size which is related to small-scale precipitation patterns that are hard to model using the current set-up of GWEX. Our results are in agreement with Paschalis et al. (2014), who used a weather generator in combination with a hydrological model to understand the influence of small-scale space-time rainfall variability on flood response. The authors found that Kleine Emme's flood response is considerably linked with small-scale rainfall variability. However, overall, it should be kept in mind that a direct comparison of long CS estimates with observations is not feasible for events of very low probability due to the limited length of the observation records.

5.2 Limitations

One limitation to be considered in our first experiment is the number of ensemble members. Generally, bootstrapping methods rely on random sampling with replacement to reproduce many samples that help to compute confidence intervals. A large sample of bootstrapped members (for example 1,000) would lead to statistically robust results and would probably give us a larger range of unexplored events. However, we used an ensemble of 10 members of synthetic precipitation only due to the high computational demand that our methodology requires. Furthermore, regarding the classification based on WT, one limitation is that only the parameters controlling the temporal and

inter-site correlations of precipitation occurrence and intensity are conditioned. Conditioning the EGPD parameters to the WT would improve our understanding of its effect, but so far it has been challenging to implement. The reason is that conditioning parameters to weather types means that we group the values and thus reduce the amount of information available to extract statistics. Therefore, some groups may include large observed values to which the parameters are very sensitive. Probably, longer time series would help to obtain robust estimates. Lastly, long CS are computationally expensive due to the combination of different models of different complexity.

5.3 Uncertainties

CSs are subject to different sources of uncertainties since they combine different methodological approaches and models (Grimaldi et al. 2013; Winter et al. 2019). In this study, we focused on selected uncertainties related to the WG. A more comprehensive framework for computing the uncertainties would be difficult due to computational constraints and the involvement of many parameters that would exponentially increase the complexity. However, the suggested framework focuses on elements that are considered sensitive and were expected to lead to marked changes.

Regarding the WG, previous studies have shown that it is a stand-alone application able to reproduce precipitation extremes in large river basins (Evin et al. 2019; Haruna et al., 2021; Viviroli et al. 2022). Here, we quantify itscontribution to the total uncertainty of the flood estimates based on two experiments dealing with parameter uncertainty. In a recent study, Moraga et al. (2022)combined regional climate model outputs with a weather generator and a distributed hydrological model to quantify the uncertainties in streamflow projections. They found that for the Kleine Emme and Thur catchments, extreme high flows are characterized by very large uncertainty due to stochastic uncertainty and which would be persistent even if perfect climate models were used. These findings support our findings for the Kleine Emme, where the GWEX realizations are not in good agreement with the disaggregated observations due to the small scale precipitation patterns that are likely not well captured. This highlights the importance of stochasticity in smaller scales (Addor et al. 2014; Fatichi et al. 2011; Moraga et al., 2022; Peleg et al. 2017) It should be mentioned here that the short length of the observational data as well as the variability in length between the stations in a region can affect the reliability of the simulated precipitation. In addition, precipitation measurement errors are inherent and cannot be avoided. Furthermore, realistic temporal disaggregation from daily to hourly values can be particularly challenging in regions such as the Ticino due to the specific characteristics of precipitation and measurement series and therefore can also influence the results shifting these towards higher or lower estimates (see Maloku et al. 2023). Additionally, the selection of the length of the scenarios is critical when estimating uncertainties. In this study, we estimated the uncertainties first by using one time series of 30,000 years and then by using the same time series but split into 30 scenarios of 1,000 years. This means that floods that were initially assigned to high return levels were then distributed to smaller return levels and thus, the uncertainty assigned to these return levels was increased. This is discussed in detail and visually depicted in Blazkova et al. (2017), where 100,000 years long time series are plotted on a Gumbel plot for different combinations of scenarios. While the effect of this sampling uncertainty is inherent in our results, this is also an assumption that comes with the proposed methodology of long CS and cannot be avoided. Other important uncertainties stem from hydrological modeling. For example, the choice of the hydrological model itself plays a key role as well as the structural uncertainty and parameter uncertainty (Viviroli et al. 2009; Arnaud et al. 2017; Sikorska and Renard 2017; Winter et al. 2019; Sikorska-Senoner et al. 2020; Horton et al., 2022; Viviroli et al. 2022). Here, we have accounted for hydrological parameter uncertainty with the three representative parameter sets (see Sect. 3.2) that present a prediction interval (Sikorska-Senoner et al. 2020; Staudinger et al. 2024; Viviroli et al. 2022) and that are used to run the routing model. The differences in these sets was mostly evident for high return levels. Here, we present the results of the median parameter set, as this seems sufficient for focusing on the uncertainty due to the WG inputs. However, we acknowledge that parameter uncertainty can be constrained due to the equifinality (Beven 2006). In addition, uncertainty in the hydrological model may come from structural uncertainties. A previous study by Addor et al. (2014) implemented a modeling chain to quantify uncertainties in future hydrological regimes in Swiss catchments and showed that the uncertainty in hydrological models can be the dominant source of uncertainty for Alpine catchments, where snow and ice melt processes are crucial. In contrast, uncertainties for lower elevation catchments were mainly driven by climate model uncertainty and natural climate variability. Thus, this is to be expected in our study as the test catchments have different physiographic characteristics and have shown different responses to the changes in precipitation inputs.

5.4 Explanatory factors

Various factors influence the uncertainties that were previously discussed and thus, the confidence interval amplitude (CIA) of the extreme flood estimates (Fig. 11). For the sake

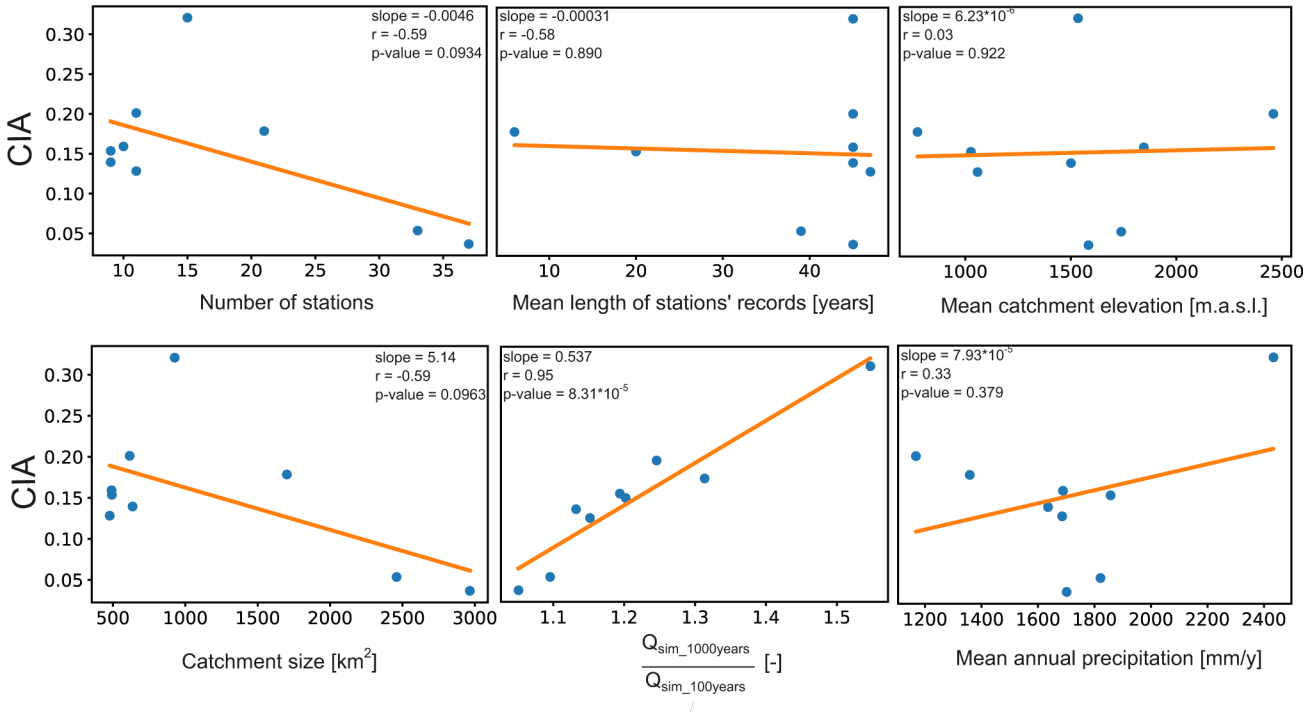

Fig. 11 Correlation scatter plots for six selected explanatory factors that contribute to the confidence interval amplitude (CIA) of the reference scenarios, calculated for the 1,000-year return period

of simplicity, we examined the CIA calculated based on the reference scenarios, for a return period of 1,000 years for all test catchments. As expected, the higher the number of stations available, the higher the amount of information related to precipitation, and thus more robust estimates were given. Similarly, we calculated the mean length of all station records per catchment and found a negative correlation with the uncertainty of the flood estimates. This was in line with our findings and explains a well-documented challenge in hydrology (Winter et al. 2019; Viviroli et al. 2022; Beneyto et al. 2023b; MacDonald et al. 2024), which links the limitations in the reliability of extreme flood estimates with the length of observational records. Furthermore, we considered the catchment size as another critical explanatory factor of flood variability, as it has been widely suggested in the literature (see e.g. Perdios and Langousis 2020), but also for the uncertainty of the flood estimates. The results demonstrated a clear negative correlation between CIA and catchment size. This finding is explained by the fact that the WG has been developed with a focus on large catchments, where small-scale precipitation patterns are not as important in the flood generation mechanism compared to smaller catchments (Arnaud et al. 2017). In addition, in larger catchments a bigger number of precipitation stations are considered in the generation process which leads to more robust precipitation estimates. Notably, the ratio of discharge corresponding to 1,000 years return period to 100 years respectively is significantly positively correlated with the CIA. This result shows that uncertainties are higher for catchments that show a strong increase of flood peaks with increasing return period. Moreover, we found no appreciable correlation between CIA and the mean catchment elevation. While mean annual precipitation is positively related to CIA, the p-value does not indicate a statistically significant correlation. Altogether, these results suggest that there is an association between CIA and most of the corresponding factors shown in Fig. 11, which highlights the variety of different sources that contribute to the uncertainties of our hydrometeorological modeling chain.

6 Conclusions

In this study we proposed a framework for extreme flood estimation based on long CSs. We focused on the uncertainties in meteorological input data since it is one of the most important sources of uncertainties in flood estimation. We used synthetic precipitation scenarios from a WG parameterized on the observations as a benchmark scenario. We first investigated the uncertainties induced by the variability of precipitation (bootstrapping experiment) and then compared the reference scenarios with the weather-typed conditioned scenarios. In this way, we explored a plethora of unobserved but plausible floods that are critical for flood risk management. One of the main conclusions of our study is that uncertainties in precipitation and flood estimates tend

to rise with increasing return period. While this is not surprising as a fact, we were able to gain quantitative insight into the uncertainties and their causes. Regarding the bootstrapping experiment, results demonstrated that most of our test catchments proved to be sensitive when different WG parameterizations were implemented. Thus, uncertainties in AMF were following the patterns of uncertainties of AMP. Often, the uncertainties due to precipitation were amplified after applying the hydrological model and hydrologic routing. This highlights the importance of reliable precipitation input data when focusing on extreme floods, since highly uncertain precipitation can lead to even more uncertain flood estimates. In addition, the uncertainties were lower for a higher number of stations available in the generation process and for larger catchment sizes. Despite not using a specific sensitivity or uncertainty framework, the comparison of the results itself helps identify uncertainties from the WG and unravel the sensitivity to selected elements. These insights are very valuable for interpretation of the results. Future work will target the impact of the hydrological model structure on the extreme flood estimates. The proposed methodology could be applied to other regions as well, however with due care regarding the scale of catchments considered and the factors examined here that contribute to the robustness and reliability of the results. The results of this analysis add important context to studies related to hazard assessment, safety analyses, and hydraulic engineering projects.

Acknowledgements We thank Alexandre Mas for his contributions related to the generation of the scenarios for the bootstrapping and the weather type experiments.

Author contributions B.H., D.V. and G.E. conceptualized the study. B.H. and G.E. conceived and provided the data for the bootstrapping and the weather type experiments. E.K. ran the simulations with support from M.K., M.S. and M.V. E.K. analyzed and visualized the data, carried out the investigation, and drafted the paper. All co-authors reviewed and edited the paper.

Funding Open access funding provided by University of Zurich. This research has been funded by the Federal Office for the Environment FOEN and the Swiss Federal Office of Energy SFOE as a part of the project “Extreme Floods in Switzerland” (EXCH).

Data availability Data generated for this study can be obtained from the first author upon reasonable request.

Declarations

Competing interests The authors declare no competing interests.

References

Addor N, Rössler O, Köplin N, Huss M, Weingartner R, Seibert J (2014) Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. Water Resour Res 50(10):7541–7562. https://doi.org/10.1002/2014WR015549

Ailliot P, Allard D, Monbet V, Naveau P (2015) Stochastic weather generators: an overview of weather type models. J De La Société Française De Statistique 156(1):101–113. http://www.sfds.asso.fr/journal

Arnaud P, Cantet P, Odry J (2017) Uncertainties of flood frequency estimation approaches based on continuous simulation using data resampling. J Hydrol 554:360–369. https://doi.org/10.1016/j.jhydrol.2017.09.011

Ayele BG, Mengistu TG, Woldemariam AD (2024) Evaluation of weather generator tools to estimate climate conditions in different agro ecological zones of North Shewa, Ethiopia. Discover Sustain 5(1). https://doi.org/10.1007/s43621-024-00330-2

Badoux A, Hegg C (2021) Grundlagen Extremhochwasser Aare: Hauptbericht Projekt EXAR. Methodik und Resultate. WSL Berichte, 104, Birmensdorf. https://www.wsl.ch/en/publications/extreme-flood-events-on-the-river-aare

Bárdossy A, Das T (2008) Influence of rainfall observation network on model calibration and application. Hydrol Earth Syst Sci 12:77–89. https://doi.org/10.5194/hess-12-77-2008

Baumgartner E, Boldi M-O, Kann C, Schick S (2013) Hochwasserstatistik am BAFU – Diskussion eines neuen methodensets. Wasser Energ Luft 105:103–110

Beneyto C, Aranda JÁ, Francés F (2023) Exploring the uncertainty of weather generators’ extreme estimates in different practical available information scenarios. Hydrol Sci J 68(9):1203–1212. https://doi.org/10.1080/02626667.2023.2208754

Beneyto C, Vignes G, Aranda JÁ, Francés F (2023b) Sample uncertainty analysis of daily flood quantiles using a weather generator. Water 2023 15:3489. https://doi.org/10.3390/w15193489

Bergström S, Institute (1992) (SMHI)/Sveriges Meteorologiska och Hydrologiska Institut (SMHI)

Beven K (2006) A manifesto for the equifinality thesis. J Hydrol 320(1–2):18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Blazkova S, Beven K (2002) Flood frequency estimation by continuous simulation for a catchment treated as ungauged (with uncertainty). Water Resour Res 38(8). https://doi.org/10.1029/2001WR000500

Blazkova S, Beven K (2004) Flood frequency estimation by continuous simulation of subcatchment rainfalls and discharges with the aim of improving dam safety assessment in a large basin in the Czech Republic. J Hydrol 292(1–4):153–172. https://doi.org/10.1016/j.jhydrol.2003.12.025

Blazkova SD, Blazek VD, Jansky B (2017) Continuous simulation for computing design hydrographs for water structures. Hydrol Process 31(13):2320–2329. https://doi.org/10.1002/hyp.11204

Boughton W, Droop O (2003) Continuous simulation for design flood estimation - a review. Environ Model Softw 18(4):309–318. https://doi.org/10.1016/S1364-8152(03)00004-5

Breinl K (2016) Driving a lumped hydrological model with precipitation output from weather generators of different complexity. Hydrol Sci J 61(8):1395–1414. https://doi.org/10.1080/02626667.2015.1036755

Brunner MI, Swain DL, Wood RR, Willkofer F, Done JM, Gilleland E, Ludwig R (2021) An extremeness threshold determines the regional response of floods to changes in rainfall extremes. Commun Earth Environ 2(1). https://doi.org/10.1038/s43247-021-00248-x

Calver A, Stewart E, Goodsell G (2009) Comparative analysis of statistical and catchment modelling approaches to river flood frequency estimation. J Flood Risk Manag 2(1):24–31. https://doi.org/10.1111/j.1753-318X.2009.01018.x

Camici S, Tarpanelli A, Brocca L, Melone F, Moramarco T (2011) Design soil moisture estimation by comparing continuous and storm-based rainfall-runoff modeling. Water Resour Res 47(5). https://doi.org/10.1029/2010WR009298

Chen J, Brissette FP, Zhang XJ (2016) Hydrological modeling using a multisite stochastic weather generator. J Hydrol Eng 21(2). https://doi.org/10.1061/(asce)he.1943-5584.0001288

Evin G, Blanchet J, Paquet E, Garavaglia F, Penot D (2016) A regional model for extreme rainfall based on weather patterns subsampling. J Hydrol 541(B):1185–1198. https://doi.org/10.1016/j.jhydrol.2016.08.024

Evin G, Favre A-C, Hingray B, Hingray B (2018) Stochastic generation of multi-site daily precipitation focusing on extreme events. Hydrol Earth Syst Sci 22(1):655–672. https://doi.org/10.5194/hess-22-655-2018

Evin G, Favre AC, Hingray B (2019) Stochastic generators of multi-site daily temperature: comparison of performances in various applications. Theor Appl Climatol 135(3–4):811–824. https://doi.org/10.1007/s00704-018-2404-x

Falter D, Schröter K, Dung NV, Vorogushyn S, Kreibich H, Hundecha Y, Apel H, Merz B (2015) Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. J Hydrol 524:182–193. https://doi.org/10.1016/j.jhydrol.2015.02.021

Fatichi S, Ivanov VY, Caporali E (2011) Simulation of future climate scenarios with a weather generator. Adv Water Resour 34(4):448–467. https://doi.org/10.1016/j.advwatres.2010.12.013

Frank C, Rußwurm M, Fluixa-Sanmartin J, Tuia D (2023) Short-term runoff forecasting in an alpine catchment with a long short-term memory neural network. Front Water. 5https://doi.org/10.3389/frwa.2023.1126310

Frei C, Fukutome S (2022) Extreme point precipitation. Hydrological Atlas of Switzerland. B: Water in the atmosphere https://hydromaps.ch/#en/8/46.830/8.190/bl_hds--precip_24h_2a$4/NULL

Frei C, Davies HC, Gurtz J, Schär C (2000) Climate dynamics and extreme precipitation and flood events in Central Europe. Integr Assess 1:281–299. https://doi.org/10.1023/A:1018983226334

García Hernández A, Claude J, Paredes Arquiola B, Roquier, Boillat J-L (2014) Swiss competences in river engineering and restoration: Special Session on Swiss Competences in River Engineering and Restoration of the Seventh International Conference on Fluvial Hydraulics (River Flow 2014), EPFL Lausanne, Switzerland, 5 September 2014

García Hernández J, Foehn A, Fluixá-Sanmartín J, Roquier B, Brauchli T, Arquiola P, J., De Cesare G (2020) RS MINERVE-Technical manual, v2.25.

Grimaldi S, Petroselli A, Arcangeletti E, Nardi F (2013) Flood mapping in ungauged basins using fully continuous hydrologic-hydraulic modeling. J Hydrol 487:39–47. https://doi.org/10.1016/j.jhydrol.2013.02.023

Grimaldi S, Volpi E, Langousis A, Papalexiou M, De Luca SL, Piscopia D, Nerantzaki R, Papacharalampous SD, G., Petroselli A (2022) Continuous hydrologic modelling for small and ungauged basins: a comparison of eight rainfall models for sub-daily runoff simulations. J Hydrol 610:127866. https://doi.org/10.1016/j.jhydrol.2022.127866

Haruna A, Blanchet J, Favre A (2022) Performance-based comparison of regionalization methods to improve the at-site estimates of daily precipitation. Hydrol Earth Syst Sci 26:2797–2811. https://doi.org/10.5194/hess-26-2797-2022

Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. Bull Am Meteorol Soc 90(8):1095–1107. https://doi.org/10.1175/2009BAMS2607.1

Hawkins E, Sutton R The potential to narrow uncertainty in projections of regional precipitation change. Climate Dynamics, 37(1), 407–418., Horton P, Schaefli B, Kauzlaric M (2011) (2022). Why do we have so many different hydrological models? A review based on the case of Switzerland. Wiley Interdisciplinary Reviews: Water, 9(1), e1574. https://doi.org/10.1002/wat2.1574

Hwang Y, Clark MP, Rajagopalan B (2011) Use of daily precipitation uncertainties in streamflow simulation and forecast. Stoch Env Res Risk Assess 25(7):957–972. https://doi.org/10.1007/s00477-011-0460-1

Jasper K, Gurtz J, Lang H (2002) Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. J Hydrol 269(3–4):40–52. https://doi.org/10.1016/S0022-1694(02)00138-5

Keller DE, Fischer AM, Frei C, Liniger MA, Appenzeller C, Knutti R (2015) Implementation and validation of a Wilks-type multi-site daily precipitation generator over a typical Alpine river catchment. Hydrol Earth Syst Sci 19(5):2163–2177. https://doi.org/10.5194/hess-19-2163-2015

Keller DE, Fischer AM, Liniger MA, Appenzeller C, Knutti R (2017) Testing a weather generator for downscaling climate change projections over Switzerland. Int J Climatol 37(2):928–942. https://doi.org/10.1002/joc.4750

Khalili M, Leconte R, Brissette F (2006) On the use of multi site generated meteorological input data for realistic hydrological modeling in the context of climate change impact studies. In: 2006 IEEE EIC Climate Change Conference, Ottawa, ON, Canada, 2006, pp. 1–7, https://doi.org/10.1109/EICCCC.2006.277261

Kilsby CG, Jones PD, Burton A, Ford AC, Fowler HJ, Harpham C, James P, Smith A, Wilby RL (2007) A daily weather generator for use in climate change studies. Environ Model Softw 22(12):1705–1719. https://doi.org/10.1016/j.envsoft.2007.02.005

Lamb R, Faulkner D, Wass P, Cameron D (2016) Have applications of continuous rainfall-runoff simulation realised the vision for process-based flood frequency analysis? Hydrol Process 30(14):2463–2481. https://doi.org/10.1002/hyp.10882

Langousis A, Kaleris V (2014) Statistical framework to simulate daily rainfall series conditional on upper-air predictor variables. Water Resour Res 50(5):3907–3932. https://doi.org/10.1002/2013WR014936

Langousis A, Mamalakis A, Puliga M, Deidda R (2016) Threshold detection for the generalized Pareto distribution: review of representative methods and application to the NOAA NCDC daily rainfall database. Water Resour Res 52(4):2659–2681. https://doi.org/10.1002/2015WR018502

Leander R, Buishand A, Aalders P, de Wit M (2005) Estimation of extreme floods of the River Meuse using a stochastic weather generator and a rainfall-runoff model. Hydrol Sci J 50(6):1089–1103. https://doi.org/10.1623/hysj.2005.50.6.1089

Legrand C, Hingray B, Wilhelm B, Ménégoz M (2024) Assessing downscaling methods to simulate hydrologically relevant weather scenarios from a global atmospheric reanalysis: case study of the upper Rhône River (1902–2009). Hydrol Earth Syst Sci 28(9):2139–2166. https://doi.org/10.5194/hess-28-2139-2024

Li Z, Brissette F, Chen J (2013) Finding the most appropriate precipitation probability distribution for stochastic weather generation

and hydrological modelling in nordic watersheds. Hydrol Process 27(25):3718–3729. https://doi.org/10.1002/hyp.9499

Lindström G, Johansson B, Persson M, Gardelin M, Bergström S (1997) Development and test of the distributed HBV-96 hydrological model. J Hydrol 201(1–4):272–288. https://doi.org/10.1016/S0022-1694(97)00041-3

MacDonald E, Merz B, Guse B, Nguyen VD, Guan X, Vorogushyn S (2024) What controls the tail behaviour of flood series: rainfall or runoff generation? Hydrol Earth Syst Sci 28(4):833–850. https://doi.org/10.5194/hess-28-833-2024

Maloku K, Hingray B, Evin G (2023) Accounting for precipitation asymmetry in a multiplicative random cascade disaggregation model. Hydrol Earth Syst Sci 27(20):3643–3661. https://doi.org/10.5194/hess-27-3643-2023

Moraga JS, Peleg N, Fatichi S, Molnar P, Burlando P (2021) Revealing the impacts of climate change on mountainous catchments through high-resolution modelling. J Hydrol 603. https://doi.org/10.1016/j.jhydrol.2021.126806

Moraga JS, Peleg N, Molnar P, Fatichi S, Burlando P, Müller-Thomy H, Sikorska-Senoner AE (20222019) Uncertainty in high-resolution hydrological projections: partitioning the influence of climate models and natural climate variability. Hydrol ProcessesHydrological Sci J 36(10).14531471https://doi.org/10.1002/hyp.14695

Najibi N, Perez AJ, Arnold W, Schwarz A, Maendly R, Steinschneider S (2024) A statewide, weather-regime based stochastic weather generator for process-based bottom-up climate risk assessments in California – Part I: model evaluation. Clim Serv 34. https://doi.org/10.1016/j.cliser.2024.100489

Naveau P, Huser R, Ribereau P, Hannart A (2016) Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection. Water Resour Res 52(4):2753–2769. https://doi.org/10.1002/2015WR018552

Okoli K, Mazzoleni M, Breinl K, Di Baldassarre G (2019) A systematic comparison of statistical and hydrological methods for design flood estimation. Hydrol Res 50(6):1665–1678. https://doi.org/10.2166/nh.2019.188

Papalexiou SM, Koutsoyiannis D, Makropoulos C (2013) How extreme is extreme? An assessment of daily rainfall distribution tails. Hydrol Earth Syst Sci 17(2):851–862. https://doi.org/10.5194/hess-17-851-2013

Papastathopoulos I, Tawn JA (2013) Extended generalised Pareto models for tail estimation. J Stat Plann Inference 143(1):131–143. https://doi.org/10.1016/j.jspi.2012.07.001

Paschalis A, Fatichi S, Molnar P, Rimkus S, Burlando P (2014) On the effects of small scale space–time variability of rainfall on basin flood response. J Hydrol 514:313–327. https://doi.org/10.1016/j.jhydrol.2014.04.014

Pathiraja S, Westra S, Sharma A (2012) Why continuous simulation? The role of antecedent moisture in design flood estimation. Water Resour Res 48(6). https://doi.org/10.1029/2011WR010997

Peleg N, Fatichi S, Paschalis A, Molnar P, Burlando P (2017) An advanced stochastic weather generator for simulating 2-D high-resolution climate variables. J Adv Model Earth Syst 9(3):1595–1627. https://doi.org/10.1002/2016MS000854

Perdios A, Langousis A (2020) Revisiting the statistical scaling of annual discharge maxima at daily resolution with respect to the basin size in the light of rainfall climatology. Water 12(2):610. https://doi.org/10.3390/w12020610

Rogger M, Pirkl H, Viglione A, Komma J, Kohl B, Kirnbauer R, Merz R, Blschl G (2012) Step changes in the flood frequency curve: process controls. Water Resour Res 48:1–15. https://doi.org/10.1029/2011WR011187

Schmocker-Fackel P, Naef F (2010a) Changes in flood frequencies in Switzerland since 1500. Hydrol Earth Syst Sci 14(8):1581–1594. https://doi.org/10.5194/hess-14-1581-2010

Schmocker-Fackel P, Naef F (2010b) More frequent flooding? Changes in flood frequency in Switzerland since 1850. J Hydrol 381(1–2):1–8. https://doi.org/10.1016/j.jhydrol.2009.09.022

Seibert J (2000) Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. Hydrol Earth Syst Sci 4:215–224. https://doi.org/10.5194/hess-4-215-2000

Seibert J, Bergström S (2022) A retrospective on hydrological modelling based on half a century with the HBV model. Hydrol Earth Syst Sci 26(5):1371–1388. https://doi.org/10.5194/hess-2021-542

Seibert J, Vis MJP (2012) Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. Hydrol Earth Syst Sci 16(9):3315–3325. https://doi.org/10.5194/hess-16-3315-2012

Seibert J, Vis M, Kohn I, Weiler M, Stahl K (2018) Technical note: representing glacier geometry changes in a semi-distributed hydrological model. Hydrol Earth Syst Sci 22(4):2211–2224.

Sharafati A, Pezeshki E, Shahid S, Motta D, Porto P Quantification and uncertainty of the impact of climate change on river discharge and sediment yield in the Dehbar river basin in Iran. J Soil Sediments 20(7):2977–2996. https://doi.org/10.1007/s11368-020-02632-0

Jordan V, Manso J, P., De Cesare G (2020) (2021). Development of a semi-distributed hydrological model for glaciated punatshangchu basin in Bhutan. J Appl Eng Technol Manag 1(1):1–13. https://doi.org/10.54417/jaetm.v1i1.19

Sikorska AE, Renard B (2017) Calibrating a hydrological model in stage space to account for rating curve uncertainties: general framework and key challenges. Adv Water Resour 105:51–66. https://doi.org/10.1016/j.advwatres.2017.04.011

Sikorska AE, Viviroli D, Seibert J (2015) Flood-type classification in mountainous catchments using crisp and fuzzy decision trees. Water Resour Res 51(10):7959–7976. https://doi.org/10.1002/2015WR017326

Sikorska-Senoner AE, Schaefli B, Seibert J (2020) Downsizing parameter ensembles for simulations of rare floods. Nat Hazards Earth Syst Sci 20(12):3521–3549. https://doi.org/10.5194/nhess-20-3521-2020

Sohrabi S, Brissette FP (2021) Evaluation of a stochastic weather generator for long-term ensemble streamflow forecasts. Hydrol Sci J 66(3):474–487. https://doi.org/10.1080/02626667.2021.1873343

Staudinger M, Kauzlaric M, Mas A, Evin G, Hingray B, Viviroli D (2024) The role of antecedent conditions in translating precipitation events into extreme floods at catchment scale and in a large basin context. https://doi.org/10.5194/egusphere-2024-909. EGUsphere [preprint]

Steinschneider S, Ray P, Rahat SH, Kucharski J (2019) A Weather-Regime-based stochastic Weather Generator for Climate Vulnerability Assessments of Water Systems in the Western United States. Water Resour Res 55(8):6923–6945. https://doi.org/10.1029/2018WR024446

Stucki P, Rickli R, Brönnimann S, Martius O, Wanner H, Grebner D, Luterbacher J (2012) Weather patterns and hydro-climatological precursors of extreme floods in Switzerland since 1868. Meteorol Z 21(6):531–550. https://doi.org/10.1127/0941-2948/2012/368

Ullrich SL, Hegnauer M, Nguyen DV, Merz B, Kwadijk J, Vorogushyn S (2021) Comparative evaluation of two types of stochastic weather generators for synthetic precipitation in the Rhine basin. J Hydrol 601:126544. https://doi.org/10.1016/j.jhydrol.2021.126544

Uppala SM, KÅllberg PW, Simmons AJ, Andrae U, Bechtold VDC, Fiorino M, Gibson JK, Haseler J, Hernandez A, Kelly GA, Li X, Onogi K, Saarinen S, Sokka N, Allan RP, Andersson E, Arpe K, Balmaseda MA, Beljaars ACM, Berg LVD, Bidlot J, Bormann N, Caires S, Chevallier F, Dethof A, Dragosavac M, Fisher M, Fuentes M, Hagemann S, Hólm E, Hoskins BJ, Isaksen L,

Janssen PAEM, Jenne R, Mcnally AP, Mahfouf J--F, Morcrette J--J, Rayner NA, Saunders RW, Simon P, Sterl A, Trenberth KE, Untch A, Vasiljevic D, Viterbo P, Woollen J (2005) The ERA-40 re-analysis. Q J R Meteorol Soc 131(612):2961–3012. https://doi.org/10.1256/qj.04.176

Vesely FM, Paleari L, Movedi E, Bellocchi G, Confalonieri R (2019) Quantifying uncertainty due to Stochastic Weather generators in Climate Change Impact studies. Sci Rep 9(1):9258. https://doi.org/10.1038/s41598-019-45745-4

Viviroli D, Mittelbach H, Gurtz J, Weingartner R (2009) Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland - Part II: parameter regionalisation and flood estimation results. J Hydrol 377(1–2):208–225. https://doi.org/10.1016/j.jhydrol.2009.08.022

Viviroli D, Sikorska-Senoner AE, Evin G, Staudinger M, Kauzlaric M, Chardon J, Favre AC, Hingray B, Nicolet G, Raynaud D, Seibert J, Weingartner R, Whealton C (2022) Comprehensive space-time hydrometeorological simulations for estimating very rare floods at multiple sites in a large river basin. Nat Hazards Earth Syst Sci 22(9):2891–2920. https://doi.org/10.5194/nhess-22-2891-2022

Weusthoff T (2011) Weather Type classification at MeteoSwiss - introduction of new automatic classification schemes. Arbeitsberichte Der MeteoSchweiz 235:46

Wilks DS (1998) Multisite generalization of a daily stochastic precipitation generation model. J Hydrol 210(1–4):178–191. https://doi.org/10.1016/S0022-1694(98)00186-3

Winter B, Schneeberger K, Dung NV, Huttenlau M, Achleitner S, Stötter J, Merz B, Vorogushyn S (2019) A continuous modelling approach for design flood estimation on sub-daily time scale. Hydrol Sci J 64(5):539–554. https://doi.org/10.1080/02626667.2019.1593419

Wüest A, Zeh M, Ackerman JD (2007) Lake Brienz Project: an interdisciplinary catchment-to-lake study. Aquat Sci 69(2):173–178. https://doi.org/10.1007/s00027-007-0016-0

Yin S, Chen D Weather generators. In: Oxford Research Encyclopedia of Climate Science, Oxford University Press, Oxford. https://doi.org/10.1093/acrefore/9780190228620.013.768, 2020.

Zischg AP, Felder G, Weingartner R, Quinn N, Coxon G, Neal J, Freer J, Bates P (2018) Effects of variability in probable maximum precipitation patterns on flood losses. Hydrol Earth Syst Sci 22(5):2759–2773. https://doi.org/10.5194/hess-22-2759-2018