~~Simulating one century~~ Assessing downscaling methods to simulate hydrologically relevant weather scenarios from a global atmospheric reanalysis: case study of the Upper Rhône River (1902-2009)~~of river discharges, low flow sequences and flood events of an alpine river from large-scale atmospheric information~~

Caroline Legrand[1], Benoît Hingray[1], Bruno Wilhelm[1,†], and Martin Ménégoz[1]

[1]Univ. Grenoble Alpes, CNRS, INRAE, IRD, Grenoble INP*, IGE, 38000 Grenoble, France
[1]*Institute of Engineering and Management Univ. Grenoble Alpes
[†]deceased, 5 April 2022

Correspondence: Caroline Legrand (caroline.legrand@univ-grenoble-alpes.fr)

Abstract. We assess the ability of two ~~typical simulation~~ modeling chains to reproduce, over the last century (1902-2009) and from large-scale atmospheric information only, the temporal variations ~~of~~ in river discharges, low flow sequences and flood events, observed at different locations of the Upper Rhône River ~~(URR)~~ catchment, an alpine river straddling France and Switzerland $(10,900\,\mathrm{km}^2)$. The two ~~simulation~~ modeling chains are made up of a downscaling model, either statistical (SCAMP) or dynamical (MAR), and the glacio-hydrological model GSM-SOCONT. Both downscaling models, forced by atmospheric information from the ~~ERA-20C atmospheric reanalysis~~ global atmospheric reanalysis ERA-20C, provide time series of daily scenarios of precipitation and temperature used as ~~input~~ inputs to the hydrological model. With hydrological regimes ranging from highly glaciated ones in its upper part to mixed ones dominated by snow and rain downstream, the ~~URR~~ Upper Rhône River catchment is ideal to evaluate the different ~~simulation chains~~ downscaling models in contrasting and demanding hydro-meteorological configurations where the interplay between weather variables, both in space and time, is determinant. Whatever the river sub-basin considered, the simulated discharges are in good agreement with the reference ones, provided that the weather scenarios are bias-corrected. The observed multi-scale variations ~~of~~ in discharges (daily, seasonal and interannual) are well reproduced~~and the hydrological situationsof low frequency (low~~. The low frequency hydrological situations, such as low flow sequences and ~~flood events~~annual discharge maxima (used as flood proxy indicators) are reasonably well reproduced. The variations in flood activity over the last century are also rather well reproduced. The results for low flow activity are less satisfactory. Bias correction is crucial for both precipitation and temperature and both downscaling models. For the dynamical one, a bias correction is also essential to get realistic daily temperature lapse rates. Uncorrected scenarios lead to irrelevant hydrological simulations, especially for the ~~sub-catchments~~ sub-basins at high elevation, mainly due to irrelevant snowpack dynamic simulations. The simulations also highlight the difficulty to simulate precipitation dependency to elevation over mountainous areas.

1 Introduction

Climate change is expected to exacerbate flood hazard through an intensification of the hydrological cycle, which will likely alter the magnitude, frequency, and/or seasonality of floods (Blöschl et al., 2017). Another concern is that of future low flows and drought situations, which are also expected to be more frequent, longer and more intense (Ruosteenoja et al., 2018; Masson-Delmotte et al., 2021). However, projecting the possible evolution of hydrological extremes at the ~~catchment scale~~ catchment-scale is still challenging, and although a large number of works has been developed for a number of rivers worldwide, considerable uncertainty about possible future changes remains, both for changes in the intensity and frequency of extreme events (e.g. Kundzewicz et al., 2016; Roudier et al., 2016; Vidal et al., 2016; Di Sante et al., 2021; Evin et al., 2021; Lemaitre-Basset et al., 2021).

Hydrological scenarios required for climate change impact studies are ~~actually~~ commonly obtained by simulation with hydrological models from ensembles of projected meteorological scenarios. To allow for a relevant impact assessment, meteorological scenarios have to fulfill some constraints imposed by the strong non-linearity and the high spatial and temporal variability of hydrological processes (e.g. strong ~~dependence~~ dependency of temperature, radiative fluxes, precipitation, etc, on elevation and aspect in mountainous environments). For instance, the meteorological scenarios have to be bias-corrected (e.g. with respect to space and seasonality) and to have rather high spatial and temporal resolution (Lafaysse et al., 2014). Because such requirements are not fulfilled by General Circulation Models (GCMs), meteorological scenarios are classically obtained with downscaling models, either dynamical or statistical.

Dynamical Downscaling Models (DDMs) are Regional Climate Models (RCMs) nested within ~~the~~ a GCM to generate fine resolution climate information (Giorgi and Mearns, 1991). They solve the full equations of mass, energy, and momentum conservation laws in the atmosphere to account for the physical interactions of land-atmosphere processes with consideration of the heterogeneity of topography, soil, vegetation, and climate variables in a region or a catchment. Over the past three decades, RCMs have been widely used in a number of studies for hydrological purposes (e.g. ~~?Arnell et al., 2003; ?; Leander et al., 2008; Leung and~~ Arnell et al., 2003; Leander et al., 2008; Leung and Qian, 2009; see Tapiador et al., 2020 for a complete review).

On the other hand, Statistical Downscaling Models (SDMs) are based on empirical relationships identified from observations between large-scale atmospheric variables (or predictors) and local weather variables (or predictands) (Von Storch et al., 1993). Weather scenarios derived from SDMs are produced from time series of predictors extracted from large-scale atmospheric outputs of climate models. SDMs have been widely used to i) generate local weather scenarios for past or future climates from GCM experiments (e.g. Wilby et al., 1999; Hanssen-Bauer et al., 2005; Boé et al., 2007; Lafaysse et al., 2014; Dayon et al., 2015), ii) produce local weather forecasts from large-scale weather forecasts of numerical weather prediction models ~~(e.g. Obled et al., 2002; Gangopadhyay et al., 2005; Marty et al., 2012; ?)~~(e.g. Obled et al., 2002; Gangopadhyay et al., 2005; Marty et al. , and iii) produce reconstructions of past weather conditions from observations and global atmospheric reanalysis data ~~(e.g. Wilby and Quinn~~ (e.g. Wilby and Quinn, 2013; Kuentz et al., 2015; Bonnet et al., 2020; Devers et al., 2021).

A general discussion of the pros-and-cons of selecting one downscaling approach over the other has been the subject of multiple manuscripts (see Fowler et al., 2007 or Maraun et al., 2010 for a review). DDMs have the advantage of being formulated

on physical principles, but their simulations are computationally intensive and ~~time consuming~~ time-consuming, with a resolution that is generally still to coarse to catch local processes. ~~Their~~ They are obviously not free of limitations and their outputs often need to be corrected for impact studies. In contrast, SDMs are very popular because of their computational efficiency and ease of use. However, they may sometimes miss some important interactions and/or correlations between meteorological variables, both in space and time.

A ~~prerequisite for~~ large number of studies have set up modeling chains to simulate river discharges in response to large-scale atmospheric trajectories over specific periods. As shown by Wood et al. (2004) and Quintana Seguí et al. (2010), for instance, the choice of the downscaling approach can strongly influence the simulation results. Not all downscaling approaches are necessarily relevant for the targeted simulations, but impact-oriented assessments can guide the model selection. For climate impact analyses ~~is that the modeling chain considered for the simulations is~~ focusing on hydrology, for instance, the modeling chains have to be able to reproduce in a relevant way the multi-scale hydrological variations that result from the large-scale atmospheric trajectories of the considered period (e.g. Lafaysse et al., 2014).

In this work, we assess and compare the ability of two ~~typical simulations~~ modeling chains to reproduce, ~~over the last century (1902-2009) and~~ from large-scale atmospheric information only, the observed temporal variations ~~of discharges , especially those of low flows and flood events,~~ in discharges in the Upper Rhône River (URR) catchment. ~~We additionally assess the necessity and the added value of a bias correction step in the elaboration of weather scenarios. The two~~ The modeling chains are made up of ~~i) the atmospheric reanalysis ERA-20C (Poli et al., 2016), ii) either the statistical downscaling model SCAMP (Raynaud et al., 2020) or the dynamical downscaling model MAR (Gallée and Schayes, 1994), and iii)~~ a downscaling model, either statistical (SCAMP; Raynaud et al., 2020) or dynamical (MAR; Gallée and Schayes, 1994), and the glacio-hydrological model GSM-SOCONT (Schaefli et al., 2005). Both downscaling models are forced by the global atmospheric reanalysis ERA-20C (Poli et al., 2016).

We combine three innovative features. i) We evaluate the modeling chains in contrasting and demanding hydro-meteorological configurations where the interplay between weather variables, both in space and time, is determinant. The URR catchment ~~is~~, a mesoscale alpine catchment straddling ~~Switzerland and France . With~~ France and Switzerland, indeed presents a number of different hydrological regimes, ranging from highly glaciated ones in its upper part to mixed ones dominated by snow and rain downstream ~~, the URR catchment offers a great opportunity to evaluate the different downscaling approaches in contrasting and demanding hydro-meteorological configurations where the interplay between weather variables, both in space and time, is determinant~~. ii) We evaluate the modeling chains over the entire 20th century, a period long enough to assess the ability of the modeling chains to reproduce daily variations in observed discharges, low frequency events (low flow sequences and annual discharge maxima), and variations in low flow and flood activities. iii) For both downscaling models, we evaluate the need for additional bias correction prior to application of the hydrological model for precipitation, temperature and temperature lapse rates.

The paper is structured as follows. Section ~~2~~ 2 describes the study area and Sect. ~~3~~ 3 the data. Section ~~4~~ 4 presents the components of the different ~~simulation~~ modeling chains. A meteorological and hydrological assessment is carried out in Sect. ~~5.~~
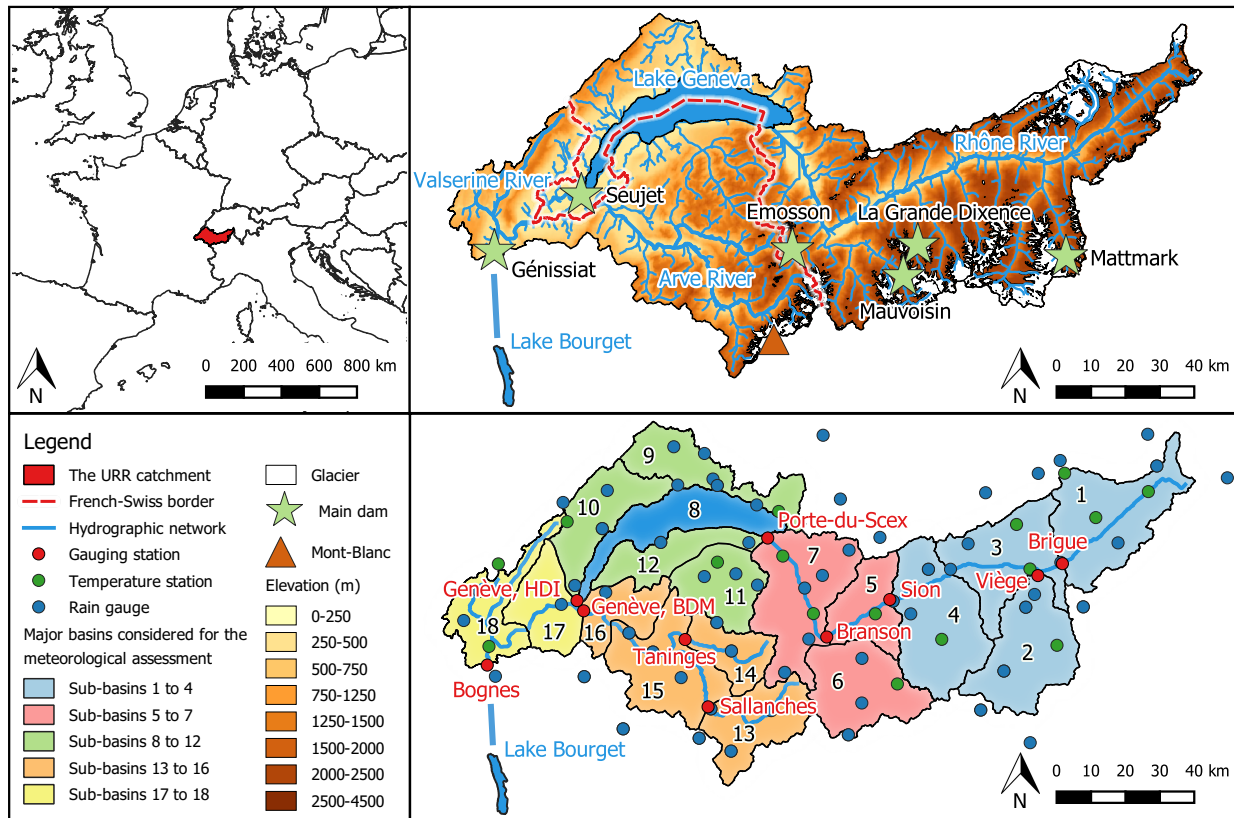
Figure 1. Study area. (**top**) Localisation and hydrological characteristics of the Upper Rhône River (URR) catchment. (**bottom**) Locations of the different weather and gauging stations. Genève, HDI: Genève, Halle-de-l'Ile. Genève, BDM: Genève, Bout-du-Monde.

, 5. Results are discussed in Sect. 6. 6. Finally, Sect. 7 7 sums up the main ~~outputs~~ results of this study and outlines future

90 ~~research lines~~ lines of research.

2   Study area

The Upper Rhône River (URR) catchment $(10,900\,\mathrm{km}^2)$ covers the southwestern part of the Swiss Alps and a part of the northern French Alps (Fig. 1). The altitude of this catchment ranges from 300 m to above 4,800 m at the top of the Mont Blanc. The presence of steep slopes ~~make~~ makes this area particularly prone to natural hazards such as landslides, floods and

95 avalanches, which are strongly connected to meteorological conditions (Beniston, 2006; Raymond et al., 2019).

In this region, the climate is continental and the temporal variability of precipitation and temperature is high. Mean annual precipitation ranges from 600 mm (in some parts of Wallis canton, Switzerland) to 1100 mm (Chamonix, France). It rains from 30 % to 45 % of the days, with an annual maximum daily precipitation reaching locally 45 mm day^{-1} to 105 mm day^{-1} on

average (Isotta et al., 2014). Glaciated areas covered 17 % of the catchment upstream to Lake Geneva and 10 % of the whole
100 ~~catchment in 2019~~ URR catchment in 2015 (GLIMS, 2015).

The gauging station of Bognes (Rhône@Bognes hereafter) records daily mean discharges at the outlet of the URR catchment. It is located at Injoux-Génissiat (France), 46 km downstream of the confluence of the Rhône and the Arve Rivers. Before the confluence of the two rivers, the hydrological regime of the URR catchment is glacio-nival: the important seasonal snowpack dynamics, combined with the late summer contribution of glacier melt~~result~~, results in a strong seasonality of flow. Flood
105 events are observed in spring due to snowmelt and in late summer/fall as a ~~results~~ result of events with large to very large rainfall amounts (e.g. the so-called "Binn-Simplon" situations; OFEG, 2002). The URR hydrological regimes turns to pluvio-nival downstream, with the successive contributions of lower elevation areas, especially that of the Arve River. Flood events from the Arve River mainly result from extreme rainfall events in fall (e.g. the so-called "retour d'Est events"; Metzger, 2023).

Upstream to Lake Geneva, the URR hydrological regime is considered to be natural until the 1950s, before the construction
110 of several large seasonal water reservoirs, mainly used for hydroelectricity production and flood protection. The reservoirs ~~typically~~ store the snow and glacier melt inflows from high elevation areas in spring and summer for hydroelectricity production in winter. The total storage capacity of all reservoirs in the catchment is 1200×10^6 m^3 or roughly 20 % of the annual catchment precipitation. Since the 1950s, the URR hydrological regime is thus significantly altered, with a reduced seasonality (higher river discharges in winter, smaller ones in spring/summer) and reduced flood discharges in summer and fall (Hingray et al.,
115 2014). Downstream, Lake Geneva, which is the largest natural water reservoir in Western Europe (580 km^2), has a significant natural buffer effect on river flows. Its influence on flows ~~was even~~ has even been exacerbated since 1884 with the construction of a downstream regulation weir for flood protection and hydroelectricity production (Grandjean, 1990).

~~Study area. (**top**) Localisation and hydrological characteristics of the URR catchment. (**bottom**) Locations of the different weather and gauging stations.~~

120 ## 3 Data

3.1 Meteorological and hydrological data

The density of weather stations covering the URR catchment is rather high. Daily meteorological variables are available from January 1, 1961 to December 31, 2015 for 62 rain gauges and 39 temperature stations in the catchment and in the areas border-ing it. The period for which observed daily discharges are available depends on the gauging station. It covers up to one century
125 for the URR at ~~Rhône~~Porte-du-Scex, the outlet of the river to Lake Geneva (Fig. 2). Note that for some ~~basins~~sub-basins, there is no observed meteorological data concomitant with "natural" discharge data. Weather and hydrological data were provided by i) the Office Fédéral de l'EnVironnement (OFEV) for the Swiss part, and ii) Météo France and the Banque Hydro for the French part. The observed level of Lake Geneva ~~is also available from~~ has also been available from the OFEV since 1886 at daily time step.
130 ~~Data used. Periods for which daily variables are available are shown in green. Hatchings indicate periods for which the hydrological regime of the URR catchment is significantly altered by dams.~~

		1900		1961	2010	2015

Figure 2. Data used. Periods for which daily variables are available are shown in green. Hatchings indicate periods for which the hydrological regime of the Upper Rhône River (URR) catchment is significantly altered by dams. Rhône@Genève, HDI: Rhône@Genève, Halle-de-l'Ile. Arve@Genève, BDM: Arve@Genève, Bout-du-Monde. MAR: dynamical downscaling model. SCAMP: statistical downscaling model.

3.2 Atmospheric data

The ~~ERA-20C reanalysis~~ global atmospheric reanalysis ERA-20C from the European Centre for Medium-Range Weather Forecasts (ECMWF; Poli et al., 2016) is used i) as boundary conditions for the ~~RCM MARexperiments~~dynamical downscaling model MAR, and ii) directly as input for the statistical downscaling model SCAMP. This reanalysis provides 6-hourly data over the 1900-2010 period at a 1.25° spatial resolution. The URR catchment is covered by 8 ERA-20C grid points.

4 Methods

4.1 Modeling chains

The modeling chains developed in this study are made up of i) a downscaling model, either dynamical (MAR) or statistical (SCAMP), to generate time series scenarios of regional weather from the global atmospheric reanalysis ERA-20C, and ii) the

glacio-hydrological model GSM-SOCONT, to simulate the corresponding discharge time series scenarios at different gauging stations in the URR catchment.

In this work, we also assess the need for a bias correction of the downscaled weather scenarios prior to their use for hydrological simulations. In some of the configurations considered in the following, a bias correction step is additionally included in the modeling chains. This applies to precipitation, temperature and temperature lapse rates.

The different models considered in the simulation chains are described below, in Sect. 4.2 for the hydrological model, in Sect. 4.3 for the dynamical and statistical downscaling models, and in Sect. 4.4 for the bias correction model. Section 4.5 sums up the different experiments carried out with these modeling chains.

4.2 Hydrological model

4.2.1 The GSM-SOCONT model

The discharge simulations ~~have been~~ were performed with a semi-distributed and daily configuration of GSM-SOCONT (Glacier and SnowMelt SOil CONTribution model; Schaefli et al., 2005), a bucket-type model that uses time series of ~~Mean Areal Precipitation (MAP) and Temperature (MAT)~~ mean areal precipitation and temperature as inputs for each hydrological unit ~~.~~ (see Fig. 15 in Supplementary Materials). The URR catchment is ~~distributed over~~ divided into 18 sub-basins~~, following Obled et al. (2009) with an additional criteria considering the availability of discharge observations~~. They were selected so that they are roughly the same size, and that a gauging station is located at the outlet of a sub-basin wherever possible (Obled et al., 2009). The ice-covered and the ice-free parts of each sub-basin, extracted from the Global Land Ice Measurements from Space (GLIMS, 2015), are considered separately and divided into 500 m elevation bands.

For each elevation band, further referred to as Relatively Homogeneous Hydrological Unit (RHHU), ~~a daily potential evapotranspiration time series is derived from temperature. The regional temperature-evapotranspiration relationship used for this was estimated on a monthly basis from the CRU (Climate Research Unit) potential evapotranspiration and temperature estimates produced for the 1900-2010 period (Harris et al., 2014).~~

For each RHHU, ~~daily MAP and MAT~~ daily mean areal precipitation and temperature are estimated from neighboring weather stations using the Thiessen's weighting method and a regional and ~~time varying elevation-temperature~~ time-varying temperature-elevation relationship (Hingray et al., 2010). ~~MAT~~ Mean areal temperature is estimated for the mean RHHU elevation. ~~MAP~~ Mean areal precipitation is supposed to be solid if ~~MAT~~ mean areal temperature is smaller than a critical temperature Tc_1, liquid if ~~MAT~~ mean areal temperature is higher than a critical temperature Tc_2, and mixed otherwise. The Tc_1 and Tc_2 values are fixed to 0 °C and 2 °C following Hingray et al. (2010) and Froidurot et al. (2014).

For each RHHU, the snowpack temporal evolution is computed based on ~~MAP and MAT~~ mean areal precipitation and temperature time series. Each simulation day, solid and/or liquid precipitation are added to the snowpack storage and/or to its liquid water content respectively. The liquid water content is also increased by snowmelt if any and the snowpack is increased by refreezing water if any. Snowmelt (or refreezing water) is estimated with a degree-day model from the positive (respectively negative) temperature degrees of the day. When the liquid water content reaches the snowpack retention capacity, an "equivalent

rainfall" is released. The retention capacity is assumed to be proportional to the snowpack water equivalent following Kuchment and Gelfan (1996). For glaciated RHHUs, ice melt can additionally occur, namely when the glacier surface is free of snow. Ice melt is also estimated with a degree-day model.

For glaciated RHHUs, the rainfall and meltwater-runoff transformation is completed through two linear reservoirs, one for ice melt and one for the rainfall or equivalent rainfall. For non-glaciated RHHUs, the rainfall or equivalent rainfall is separated into infiltration and effective rainfall. Infiltration feeds a non-linear soil reservoir which produces deep infiltration and sub-surface flow. Infiltration (resp. actual evapotranspiration) is estimated from equivalent rainfall (resp. potential evapotranspiration) and simulated soil moisture. Effective rainfall feeds a linear overland reservoir which produces direct runoff.

Potential evapotranspiration (PET) estimates were derived from PET estimates of the Climate Research Unit (CRU) produced at a 0.5° spatial resolution from a variant of the Penman-Monteith formula (Harris et al., 2014). In order to derive PET estimates for high elevation RHHUs, PET was assumed to be a linear function of temperature T. The PET-T relationship was estimated for the region on a monthly basis from the CRU PET and T estimates produced for the 1900-2010 period.

The discharge simulated at the outlet of each sub-basin is the sum of the different discharge components produced by the different RHHUs from the glaciated and non-glaciated areas of the sub-basin. Discharges simulated for the ~~18~~ sub-basins are ~~routed within the hydrological network producing the discharge~~ summed to produce those at each gauging station and at the catchment outlet. The daily time step used in the simulations and the rather small size of the catchment area allow to disregard the routing of discharges through the river network.

Seven parameters have to be estimated for each ice-free RHHU (e.g. snowmelt degree-day factor, storage capacity and recession coefficient of reservoirs). Three additional parameters have to be estimated for glaciated RHHUs (ice melt degree-day factor, and recession coefficient of ice and snow reservoirs).

4.2.2 Hydrological model calibration

In most cases (as in here), hydrological models simulate the natural behavior of the considered catchments. One frequent approach for estimating the model parameters in such a configuration requires a time period with concomitant observations of weather and "natural" discharges, i.e. concomitant observations of weather and discharges not altered by anthropic activities or waterworks. For most sub-basins of the URR catchment, there is no such period. Daily weather observations are mainly available from the 1960s (here for period ~~P1~~ P* = 1961-2015) whereas their hydrological behavior was natural until the 1950s only, strongly altered after by ~~the~~ dams.

The method used for ~~parameter~~ parameters estimation was thus adapted to the ~~catchment~~ sub-basins data configuration, depending i) on the "perturbation" level of the ~~catchment~~ sub-basins hydrological behavior during ~~period P1~~ the period P*, and ii) on the availability of flow observation records for the period prior the 1950s. For gauged ~~catchments~~ sub-basins for which the hydrological behavior can be considered as natural (or at least not significantly altered) over ~~period P1~~ the period P* (or at least over a significant sub-period of ~~P1~~ P*), parameters were estimated with a classical hydrological calibration approach,

minimizing an objective function estimated from simulated and observed discharges time series on the same period, namely:

$$F_{natural} = 1 - NSE_{chrono} \tag{1}$$

where NSE_{chrono} is the Nash-Sutcliffe Efficiency criterion (Nash and Sutcliffe, 1970) obtained from day-to-day deviations between observed and simulated daily discharges.

210   For gauged ~~catchments~~ sub-basins for which the hydrological behavior is significantly altered over ~~P1~~ P* and for which "natural" flow observations are available prior to 1950, parameters were estimated based on hydrological signatures (Sivapalan et al., 2003; Winsemius et al., 2009). In the present case, parameters ~~are~~ were calibrated so that simulated signatures reproduce at best observed ones but observed and simulated signatures come from different periods following Hingray et al. (2010) (e.g. 1961-2015 and 1922-1963 respectively for the Viège ~~basin). We thus assume that the weather regimes and the~~

215   ~~natural hydrological behavior of the catchment have not significantly changed over the last century, which seems a reasonable assumption to make in first approximation. Signatures~~ sub-basin).

The signatures considered here are the interannual daily regime (366 values) and the statistical distribution of the annual discharge maxima. The objective function is a combination of the ~~NSE~~ NSE criterion and the Kolmogorov-Smirnov distance d_{KS}, applied respectively between simulated and observed signatures:

220   $$F_{altered} = \underline{0,5*}0.5\times(1 - NSE_{regime}) + \underline{0,5*}0.5\times\frac{d_{KS}}{\max(Q_{obs_{KS}}, Q_{sim_{KS}})} \tag{2}$$

where $d_{KS} = |Q_{obs_{KS}} - Q_{sim_{KS}}|$ is the Kolmogorov-Smirnov distance with $Q_{obs_{KS}}$ and $Q_{sim_{KS}}$ the corresponding discharge percentiles in the observed and simulated distributions.

~~The objective functions values obtained for~~ For ungauged sub-basins~~calibrated from signatures are given in Supplementary Materials (Table ??). The results of this~~, parameters were obtained via regionalization, following the methodological recommendations

225   of Bárdossy (2007) and Viviroli et al. (2009). All the ungauged sub-basins located upstream of a given gauging station were calibrated at the same time and forced to share the same parameters set. The discharge time series used for the calibration is the time series simulated with this multiple sub-basins configuration at the downstream gauging station, where observations are available. The ungauged sub-basins can be seen in Fig. 1. The sub-basins grouped together for the calibration of their parameters are also listed in Table S2 in Supplementary Materials.

230   In practice, whatever the calibration configuration and the calibration objective function, we used the automatic calibration algorithm DDS (Dynamically Dimensioned Search; Tolson and Shoemaker, 2007). The objective functions values and the results of the classical and signature-based ~~calibration are illustrated with the results obtained for the Viège basin in Fig. ??.~~ calibrations are given for each sub-basin in Table S2, Fig. S3 and Fig. S4 respectively in Supplementary Materials.

~~Finally, for ungauged basins, parameters were obtained via regionalization from neighboring gauged basins, following the~~

235   ~~methodological recommendations of Bárdossy (2007) and Viviroli et al. (2009). In practice, whatever the calibration configuration and~~ Depending on available data, we thus considered different calibration approaches. Because of this data context, it is not always easy to assess the relevance of the calibration. For sub-basins where concomitant weather and "natural" discharge data are available over a sufficiently long period, this assessment can be made with a classical split-sample test. The split-sample

tests carried out by Schaefli et al. (2005) for three URR sub-basins show that the classical calibration is efficient and robust in this context.

For sub-basins for which the hydrological behavior is significantly altered by dams, the ~~calibration objective function, we used the automatic calibration algorithm DDS (Dynamically Dimensioned Search; Tolson and Shoemaker, 2007).~~ split-sample test is not possible, as the data of the entire period have to be considered for the signature-based calibration. The analyses described below nevertheless suggest that the signature-based calibration remains efficient and robust in our context.

To assess this, we recalibrated the parameters of the four URR sub-basins with a natural (or at least not significantly altered) hydrological regime, using only the hydrological signatures. In a first analysis, the signatures were estimated using the same data (period P0) than those considered for the classical calibration (see Fig. S5 in Supplementary Materials). The simulated time series remain in good agreement with the observed ones. The NSE coefficients are logically lower than those obtained with the classical calibration approach, but the differences are quite small (see Fig. S6 in Supplementary Materials).

In a second analysis, we carried out a similar split-sample test for the signature-based calibration. To do this, we split the period into two (sub-periods P1 and P2), we recalibrated the parameters using only the hydrological signatures of period P1 with the weather data from P1 and, with this set of parameters, we simulated over period P2 the discharge time series from the weather data of P2. The simulated time series remain in good agreement with those observed. They are also in good agreement with those obtained with a signature-based calibration when signatures are derived from the entire period P0. The NSE coefficients are logically lower, but the differences are quite small (see Fig. S7 in Supplementary Materials).

4.2.3 Modeling ~~of~~ the behavior of Lake Geneva

An ad hoc model was developed to simulate the influence of Lake Geneva on river flows. The lake has a natural buffer effect on flows, additionally altered with its regulation. From the 1990s, the outflows ~~first~~ follow the main regulation objectives of the 1997 settlement (CERCG, 1997), namely i) a target water level in the lake for each calendar day, ii) the environmental low flow to be satisfied downstream, and iii) the discharge to not exceed except during periods of high water in the lake. The day-to-day variations ~~of~~ in lake outflows are additionally driven by the hydroelectric production required from the plant at the lake outlet to satisfy a part of the regional electricity demand.

By lack of appropriate data, we here only accounted for the natural buffer effect of the lake and for the main regulation objectives mentioned above. The natural buffer effect ~~was~~ is simply modeled with the classical three equations system needed to simulate the behavior of unregulated reservoirs, namely a water balance, a ~~storage~~ water level-storage and a reservoir flow equation (Hingray et al., 2014). Here, a simple linear water level-storage relationship is considered and the outflow that would be obtained without regulation is assumed to be proportional to the volume stored in the reservoir. For the water balance, changes in storage are obtained each day from reservoir inflows (direct precipitation, upstream and lateral basins flows) and losses (evaporation, lake outflow), where evaporation ~~was~~ is estimated with the Rohwer's equation (Rohwer, 1931).

4.3 Downscaling models

4.3.1 The dynamical downscaling model MAR

The Modèle Atmosphérique Régional (MAR; Gallée, 1995; Gallée and Schayes, 1994; Gallée et al., 1996) is a hydrostatic primitive equation model for regional atmospheric simulations. It includes a detailed scheme of clouds microphysics with six prognostic equations for specific humidity, cloud droplet concentration, cloud ice crystals (concentration and number), concentration of precipitating snow particles and rain drops. The convective adjustment is parameterized according to Bechtold et al. (2001). MAR is coupled to the one-dimensional land surface scheme SISVAT (Soil Ice Snow Vegetation Atmosphere Transfer; De Ridder and Schayes, 1997; Gallée et al., 2001) that includes a snow multilayer scheme (Brun et al., 1992; Gallée and Duynkerke, 1997).

First designed for polar regions, especially Antarctica and Greenland (e.g. Gallée et al., 1996), MAR was also applied over other regions worldwide, such as midlatitudes areas (e.g. Wyard et al., 2017; Doutreloup et al., 2019), mountainous areas ~~(e.g. Himalaya, Ménégoz et al., 2013)~~ (e.g. Himalaya; Ménégoz et al., 2013) and West Africa (e.g. Chagnaud et al., 2020). For the present study, we ~~use~~ used the MAR simulation forced by the ~~ERA-20C reanalysis over~~ global atmospheric reanalysis ERA-20C over the 1902-2009 period and produced at a 7 km resolution over the European Alps (Ménégoz et al., 2020; Beaumet et al., 2021). The URR catchment is covered by 281 MAR grid points.

4.3.2 The statistical downscaling model SCAMP

SCAMP (Sequential Constructive Atmospheric Analogues for Multivariate weather Predictions; Raynaud et al., 2020) is a statistical downscaling model based on atmospheric analogues (Lorenz, 1969). It assumes that similar large-scale atmospheric configurations lead to similar local or regional weather situations (e.g. Obled et al., 2002; Chardon et al., 2014). The simulation process is partly stochastic: for each simulation day, one of the 30-nearest analogues is randomly selected and used as weather scenario for this day. SCAMP can thus be used to generate multiple weather scenarios. In the present work, SCAMP was used to generate 30 time series of daily spatial weather scenarios for the 1902-2009 period from ~~ERA-20C reanalysis~~ the global atmospheric reanalysis ERA-20C outputs. The simulation process is described in the following and summarized in Fig. 3.

(1) For each day of the simulation period, the 30-nearest atmospheric analogue days are identified from candidate days available in the archive period (1961-2009 in the present case). The candidate days are the days found within a 61-days calendar window centered on the target day. A two-step analogue selection is considered. The 100-nearest analogues in terms of large-scale atmospheric circulation are first identified. The selection criterion is the Teweles-Wobus Score (Teweles and Wobus, 1954) applied to the daily geopotential heights at 1000 and 500 hPa. It quantifies the similarity between fields from their shapes, informing thus on the origin of air masses. The 30-nearest analogs are sub-selected from ~~those~~ these 100-nearest analogs, based on small-scale atmospheric features following Raynaud et al. (2020) (namely 600 hPa vertical velocities and large-scale temperature at 2 m from September to May, large-scale precipitation otherwise).

(2) Following Chardon et al. (2016), SCAMP is used to generate ~~MAP and MAT~~ Mean Areal Precipitation (MAP) and Mean Areal Temperature (MAT) values for the whole URR catchment. To generate weather scenarios with values different from
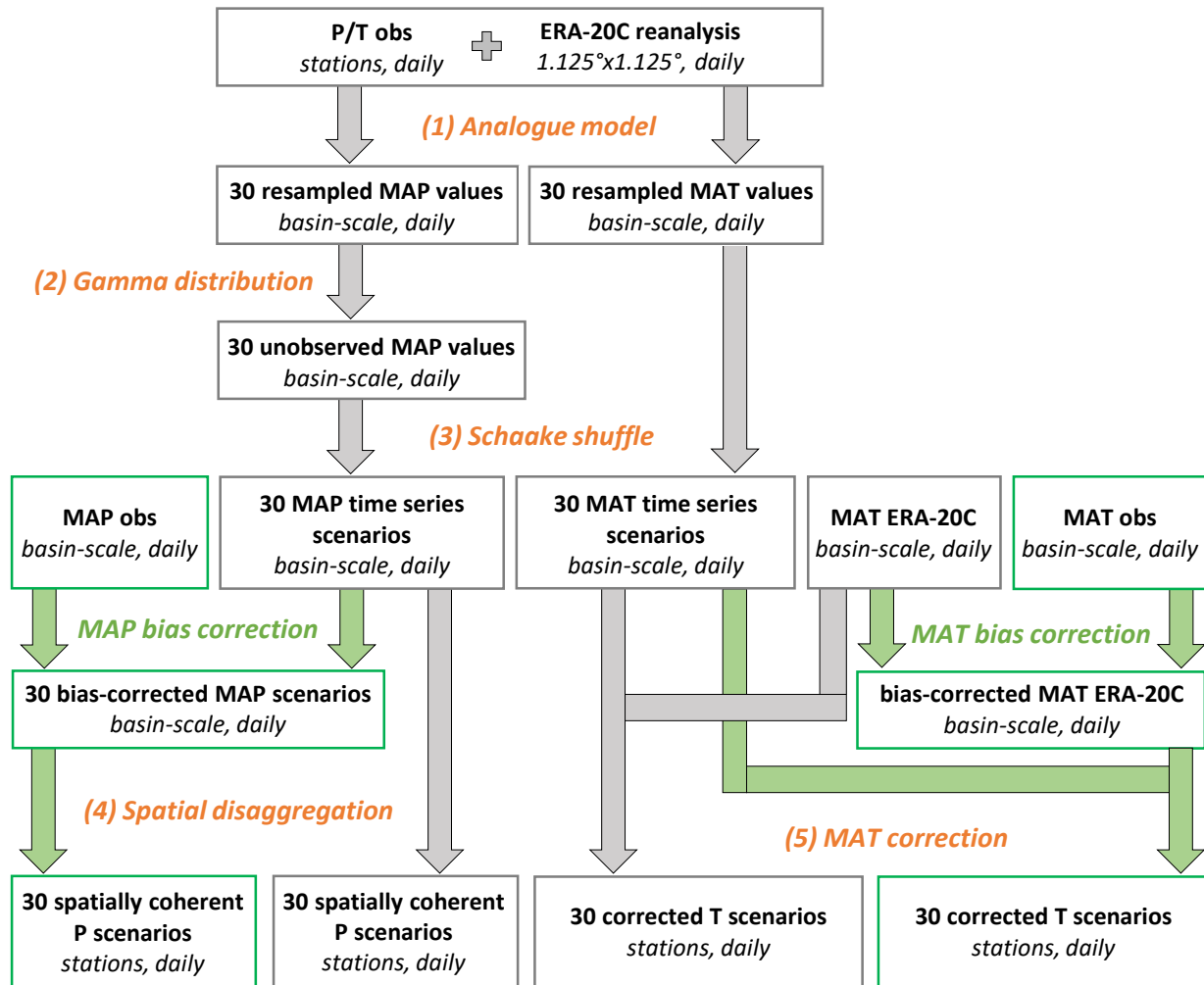
Figure 3. Scheme of the statistical downscaling model SCAMP. (**orange**) Components used in the SCAMP simulation. (**green**) Additional components used in the SCAMP bias-corrected simulation. (**bold**) Outputs obtained after each step. (**italics**) Spatial and temporal resolutions. P: Precipitation. T: Temperature. MAP: Mean Areal Precipitation. MAT: Mean Areal Temperature.

observations, SCAMP extends the analogue method with a random generation process from a day-to-day adjusted statistical distribution (Chardon et al., 2018; Raynaud et al., 2020). For precipitation, a Gamma distribution is fitted each day to the 30 regional MAP values obtained for the whole URR catchment from the analogues. The distribution is then used to generate a new sample of 30 regional MAP values, thus unobserved (and possibly higher than observations). This adaptation was shown to improve the simulation of maximum regional MAP accumulations (Raynaud et al., 2020, Fig. 8 therein).

(3) 30 regional MAP and MAT time series scenarios are produced from the 30 regional MAP and MAT values generated each day. To improve the temporal consistency between consecutive days in each time series, the 30 regional MAP and MAT

310 scenarios obtained for each day are paired with ~~those of previous days~~ the 30 scenarios of the previous day with a Schaake shuffle reordering approach (Clark et al., 2004; Raynaud et al., 2020).

(4) Following Mezghani and Hingray (2009) and Viviroli et al. (2022), the regional MAP time series scenarios are finally disaggregated with a non-parametric method of fragment to produce spatial MAP time series scenarios for the URR catchment (with one ~~MAP~~ mean areal precipitation value for each RHHU of the hydrological model). In practice, the regional MAP
315 scenario of each day is spatially disaggregated using the spatial pattern of an analog day for which weather observations are available.

(5) The large-scale atmospheric features considered for the identification of the analogues are not as informative for temperatures. The temperature additive correction method of Kuentz et al. (2015) was thus used to make each day the regional MAT time series scenarios coherent with the regional MAT ~~configuration of the ERA-20C reanalysis~~ time series of the global
320 atmospheric reanalysis ERA-20C. For each prediction day and each scenario, the correction factor is the difference between the regional MAT value of the ERA-20C reanalysis and the regional MAT value of the scenario. The correction factor is applied to correct the temperature of all stations for this scenario. For instance, if the daily regional MAT value of the scenario is $2\,°C$ warmer than the regional MAT value of the ERA-20C reanalysis, then all local temperatures of the scenario are lowered by $2\,°C$.

325 **4.4 Bias correction**

~~Examples of bias correction for regional MAP. MAR: sub-basins 1 to 4, January. SCAMP: whole catchment, August. Control period: 1961-2009. (a) empirical cumulative distribution functions (ecdf). (b) Correction functions. For SCAMP, the grey and green bands represent the confidence interval at 90 % level. The median scenarios are indicated by the black and green solid lines.~~

330 ~~As shown later,~~ As shown in Sect. 5.1, the simulated weather scenarios from ~~both MAR and~~ the dynamical downscaling model MAR and the statistical downscaling model SCAMP can be significantly different from the observations. Corrected precipitation and temperature scenarios ~~were~~ are also considered in the following. A quantile mapping bias correction (BC hereafter) was used for both variables and both models (Déqué, 2007). The different regional MAP and MAT time series were corrected using the observed regional MAP and MAT time series as references. As commonly done, the correction is additive
335 for MAT and multiplicative for MAP, and one BC function was estimated for each calendar month respectively.

~~To keep the small-scale variability between the different scenarios produced by SCAMP, one single BC function is considered for the 30 time series. An~~ For both models, an analytical BC function ~~is~~ was used to ease interpolations of correction factors estimated between empirical percentiles (a degree 4 polynomial ~~is~~ was fitted to the empirical correction factors). To avoid non relevant extrapolations of correction factors for extreme ~~precipitation~~ MAP values non simulated in the control period, the cor-
340 rection value was ~~also~~ bounded. It was bounded for the highest ~~precipitation~~ MAP percentiles (95 to 100 percentiles) to their mean empirical correction value. For ~~temperature~~MAT, the corrections for the lowest (0 to 5) and for the highest percentiles (95 to 100) were similarly set to their mean empirical correction values respectively.
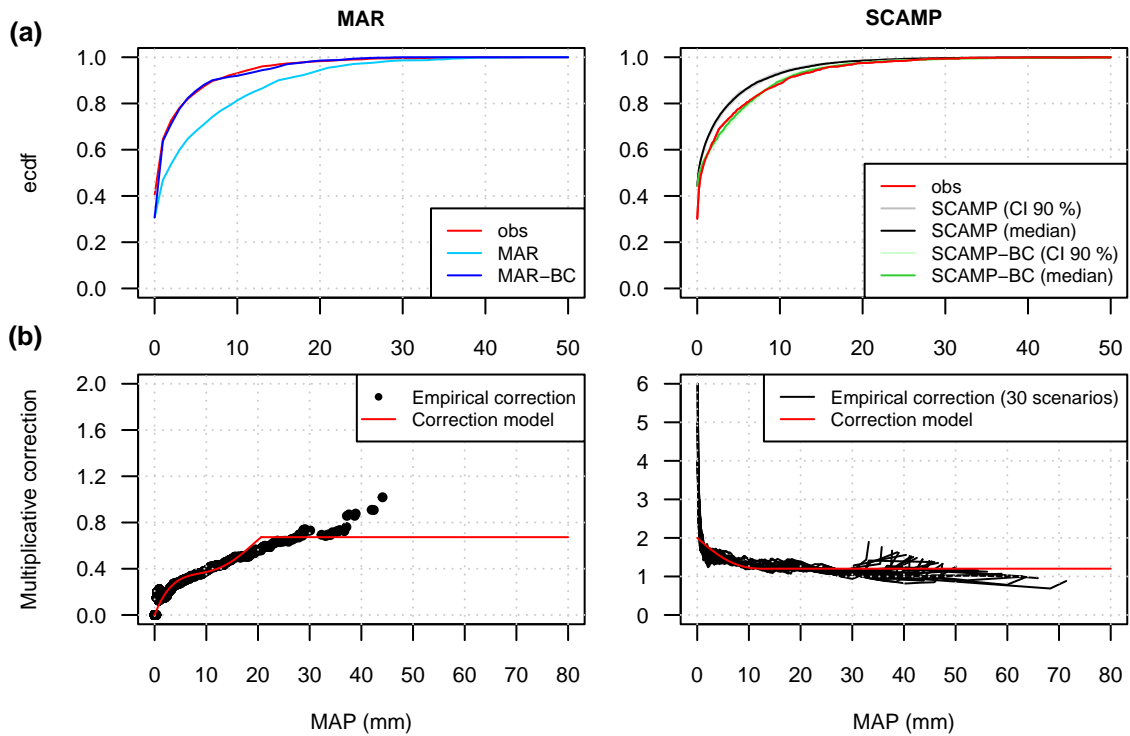
Figure 4. Examples of bias correction for Mean Areal Precipitation (MAP). MAR (dynamical downscaling model): sub-basins 1 to 4, January. SCAMP (statistical downscaling model): Upper Rhône River (URR) catchment, August. Control period: 1961-2009. (**a**) empirical cumulative distribution functions (ecdf). (**b**) Correction functions. For SCAMP, the grey and green bands represent the confidence intervals at 90 % level. The median scenarios are indicated by the black and green solid lines.

For SCAMP, the MAP and MAT corrections ~~can~~ could only be performed at the scale of the whole ~~catchment.~~ URR catchment. To keep the small-scale variability between the different scenarios produced by SCAMP, one single BC function was considered for the 30 time series. For MAR, ~~a specific BC was~~ specific BCs were applied for each of the ~~5~~ five major sub-basins ~~of the system~~ (~~shown in~~ Fig. 1~~)~~. Examples of BC functions are shown for MAP in Fig. 4.

Note that for the ~~dynamical downscaling~~ MAR simulations, a BC of the temperature lapse ~~rate~~ rates was also neces-sary. ~~This point is further~~ For hydrological simulations, temperature estimates are required for each elevation band. In the present work, they are obtained by interpolation from temperature values available at neighboring locations, using a regional temperature-elevation relationship. For the MAR simulations, this relationship is assumed to be linear and its slope, the so-called "lapse rate", is estimated for each time step from simulated temperatures using all MAR grid cells in the URR catchment.

As illustrated by its statistical distribution in Fig. 5b, the lapse rate estimated from MAR simulations varies from one day to another, as does the lapse rate estimated from observations. However, it is on average higher than that of the observations,
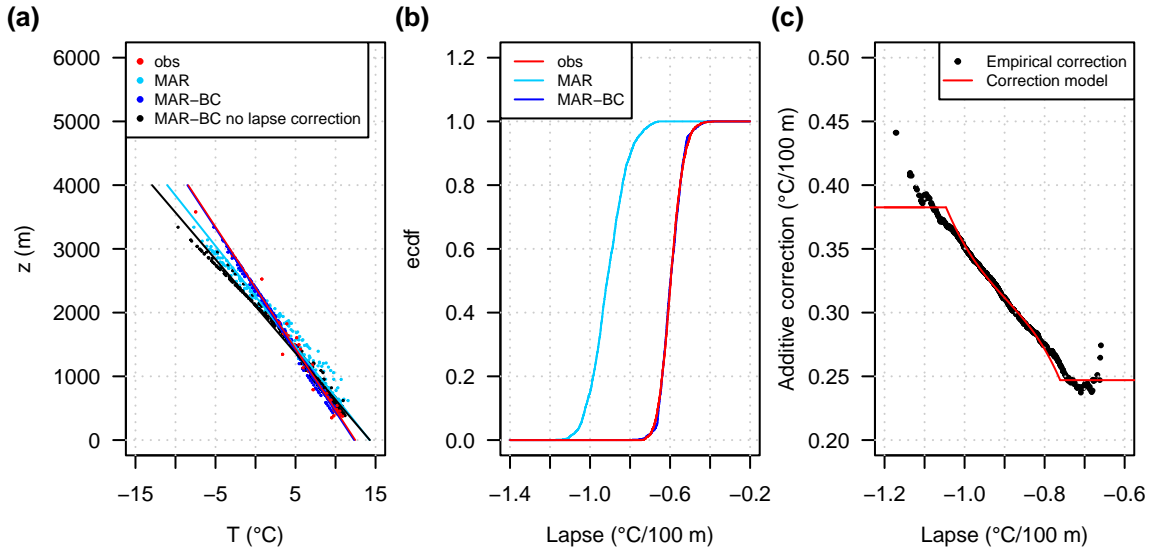
Figure 5. Illustration of the bias in the temperature-elevation relationship in the simulations of the dynamical downscaling model MAR. (**a**) Mean interannual relationship (and corresponding linear regressions) at the catchment-scale from observed temperatures (obs, red), raw MAR simulations (MAR, cyan), MAR simulations corrected for bias in mean temperature and lapse rate (MAR-BC, blue) and MAR simulations corrected for bias in mean temperature only (MAR-BC no lapse correction, black). (**b**) and (**c**) Examples of empirical cumulative distribution functions (ecdf) and correction model for temperature lapse rate (sub-basins 1 to 4, July). Note that the blue and red lines overlap in Fig. (**a**) and (**b**).

355 as also illustrated in Fig. 5a with the mean temperature-elevation relationships estimated from both data sets. The bias in the lapse rate, which likely results from a warm bias in the lower atmospheric layers of the model, is quite large, of the order of $0.3\,°\mathrm{C}\,/\,100$ m (Fig. 5b).

The bias in temperature in dynamical downscaling models is recognized and corrected for a long time. To the best of our knowledge, the bias in the lapse rate is not. A common approach to use model temperatures for hydrological simulations is

360 to identify the BC function for a given reference altitude and to use this function to correct model temperatures for all other elevations. In this process, however, the lapse rates remain unchanged and the corrected MAT may still present residual biases for all elevations different from the reference one. This is the case for the present work. For the mountainous context considered here, as discussed in Sect. ~~6.2~~6.2, this has important implications on the simulated hydrology.

For MAR temperatures, we thus considered a two-part quantile mapping correction function that takes into account both

365 biases, that of the MAT value simulated for a given reference elevation and that of the lapse rate. The corrected MAT value for a given elevation z and a given time t was obtained as follows:

$$\mathrm{MAT}_{\text{MAR-BC}}(z,t) = \mathrm{MAT}_{\text{MAR-BC}}(z_{ref},t) + \mathrm{lapse}_{\text{MAR-BC}}(t) \times (z - z_{ref}) \tag{3}$$

where $\text{MAT}_{\text{MAR-BC}}(z_{ref}, t)$ is the bias-corrected MAT value at the reference elevation z_{ref} (the correction depends on the percentile of the MAT value at t) and $\text{lapse}_{\text{MAR-BC}}(t)$ is the bias-corrected lapse rate value at t (the correction depends on the percentile of the lapse rate value estimated from MAR outputs at t).

In practice, we chose the mean elevation of the URR catchment (i.e. 1525 m) as reference elevation. The MAT and lapse rate corrections were carried out independently at the monthly scale and for the five major basins shown in Fig. 1. The results presented in the following for the MAR simulations were obtained with this two-part BC. The influence of the lapse rate correction is discussed in Sect. 6.2.

Note that for the SCAMP simulations, a BC of the temperature lapse rates was not necessary. In SCAMP simulations, the regional temperature-elevation relationship is estimated each day from the SCAMP temperatures at the observation stations. For each simulation day, as these temperatures are derived from the temperatures of an observed analog day, the relationship of the scenario always corresponds to an observed one. The temperature-elevation relationship and its variations over time are therefore coherent with the observations.

4.5 Experimental setup

Four experiments are considered in the present work. They are summarized in Fig. ~~5~~6. Weather scenarios are produced with either the ~~MAR model or the SCAMP model~~ dynamical downscaling model MAR or the statistical downscaling model SCAMP from large-scale atmospheric information (ERA-20C data). Weather scenarios obtained with each downscaling model are used to force the GSM-SOCONT model and simulate hydrological scenarios. Weather scenarios are first used with their raw values and then with their bias-corrected values. In the following, the simulations with raw weather scenarios are referred to as "MAR" and "SCAMP" simulations. ~~Bias-corrected~~ The bias-corrected simulations are referred to as "MAR-BC" and "SCAMP-BC" simulations. Both weather and hydrological scenarios can be compared to their counterpart references (observed or simulated as explained below). Among other things, we will compare the ability of each hydro-meteorological ~~simulation~~ modeling chain to reproduce the ~~day-to-day temporal variations of~~ temporal variations in both weather variables and discharges (Table 1).

As many sub-basins have altered hydrological regimes, the "hydrological reference" used for the comparison is the discharge time series obtained via hydrological simulation with the ~~"observed weather " as input~~observed weather variables as inputs. For some upstream sub-basins, ~~which~~ for which the hydrological behavior can be considered as ~~roughly~~ natural, the evaluation could also rely on a comparison with discharge observations. We however ~~choose~~ chose to use the simulated reference. This first makes the evaluation homogeneous for all URR sub-basins. This additionally allows to ~~only focus~~ focus only on the ability of the downscaling chains to simulate hydrologically relevant weather scenarios. In other words, this allows to not distort the evaluation by intrinsic errors introduced by the hydrological model. This point is further discussed in Sect. 6.4.

A multi-scale evaluation of simulations ~~has been~~ is carried out. We ~~here present~~ present here the meteorological evaluations carried out at the scale of ~~the 5~~ two of the five major sub-basins shown in Fig. 1 and hydrological evaluations at ~~4~~ four illustrative gauging stations. Rhône@Porte-du-Scex $\left(\text{gauged area of } 5,390\,\text{km}^2\right)$ is the outlet to Lake Geneva of the upstream part of the URR catchment. Rhône@Genève, Halle-de-l'Ile $\left(7,945\,\text{km}^2\right)$ is the outlet of Lake Geneva. Arve@Genève, Bout-du-Monde
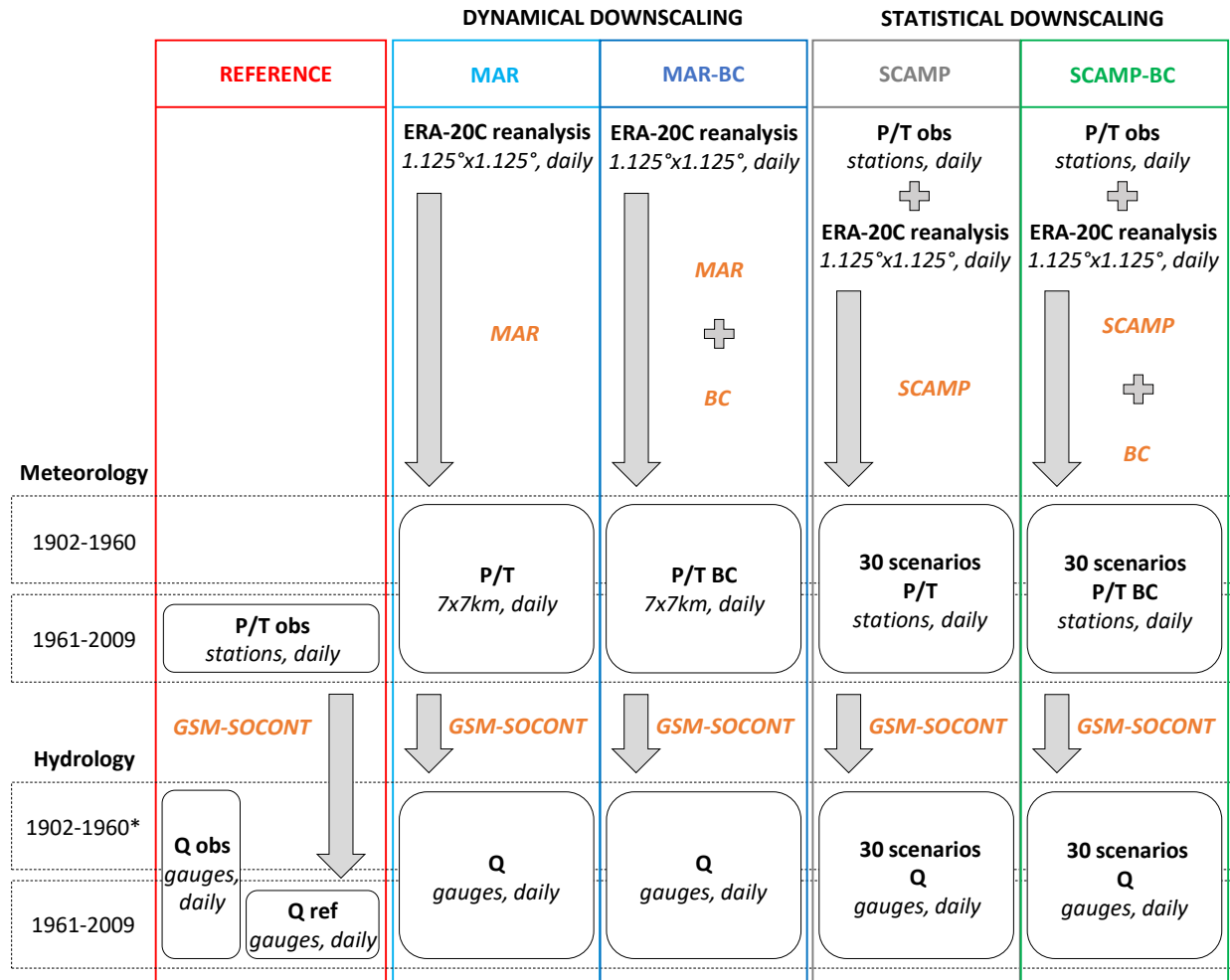
Figure 6. Summary of the different experiments. (**orange**) Models used: MAR (dynamical downscaling model), SCAMP (statistical downscaling model), BC (bias correction model), GSM-SOCONT (hydrological model). (**bold**) Outputs obtained after each step. (*italics*) Spatial and temporal resolutions. (*****) Time period depending on the considered gauge. P/T obs: observed precipitation/temperature. P/T: simulated precipitation/temperature. P/T BC: simulated bias-corrected precipitation/temperature. Q obs: observed discharge. Q ref: simulated discharge from observed weather variables. Q: simulated discharge.

is on the Arve River before its confluence with the Rhône River $(1,990\,\mathrm{km}^2)$, and Rhône@Bognes $(10,900\,\mathrm{km}^2)$ is the ~~most downstream location of the system~~outlet of the URR catchment.

Table 1. Summary of the different hydro-meteorological variables simulated and assessed in this study. MAP: Mean Areal Precipitation. MAT: Mean Areal Temperature. URR: Upper Rhône River.

Assessment objective	Variables	Spatial resolution	Period
Year-to-year variations in weather variables	Annual MAP time series	Two major sub-basins	1902-2009
	Annual MAT time series	Two major sub-basins	1902-2009
Seasonality of weather variables	Seasonal cycle of monthly MAP	Two major sub-basins	1961-2009
	Seasonal cycle of monthly MAT	Two major sub-basins	1961-2009
Multi-scale variations in discharges	Daily discharge time series	Four gauging stations	1981-1983 as example
	Mean monthly discharge time series	Four gauging stations	1961-2009
	Mean annual discharge time series	URR catchment	1902-2009
Hydrological extremes	Low flows	Four gauging stations	1961-2009
	Annual discharge maxima	Four gauging stations	1961-2009
	Low flow activity	URR catchment	Three 30-year sub-periods
	Flood activity	URR catchment	Three 30-year sub-periods

5 Results

5.1 Weather

Simulated weather variables are compared to observations in Fig. ~~6 and 7~~ 7 and 8 for both downscaling models. Figure ~~6~~ 7 shows observed and simulated year-to-year variations ~~of~~ in annual MAP and MAT and Fig. ~~7~~ 8 the seasonal cycles of monthly MAP and MAT. Results are presented for sub-basins in the upper part of the URR catchment (sub-basins 1 to 4, Fig. ~~6 and 7~~ 7 and 8, left) and for sub-basins downstream and around Lake Geneva (sub-basins 8 to 12, Fig. ~~6 and 7~~ 7 and 8, right). For the statistical downscaling model SCAMP, both figures present the dispersion between the 30 annual MAP values obtained from the 30 time series scenarios (grey and green bands). Note that the dispersion is rather large for precipitation (e.g. up to 600 mm for annual MAP, Fig. ~~6~~7), illustrating the important uncertainty in the large-scale-small-scale relationship for this variable in this region.

For both downscaling models, simulated year-to-year variations ~~of regional~~ in annual MAP and MAT are in good agreement with observed ones, whatever the area considered (Fig. ~~6~~7). The positive trend in temperature starting in 1980 is also adequately reproduced, resulting from the global combination of the warming related to anthropogenic greenhouse gases and the reduced anthropogenic aerosol cooling (Reid et al., 2016), especially pronounced over Europe (Nabat et al., 2014). However, the mean simulated variables can be rather different ~~than~~ from the observed ones. For ~~MAR~~ the dynamical downscaling model MAR, for instance (cyan lines), simulated annual ~~MAP~~ MAPs are 12 % higher than observations for sub-basins around the lake and 58 % higher for sub-basins in the upper part of the URR catchment. For the statistical downscaling model SCAMP (grey bands and black lines), the differences, although much smaller, are still significant~~: depending~~. Depending on the SCAMP scenario,
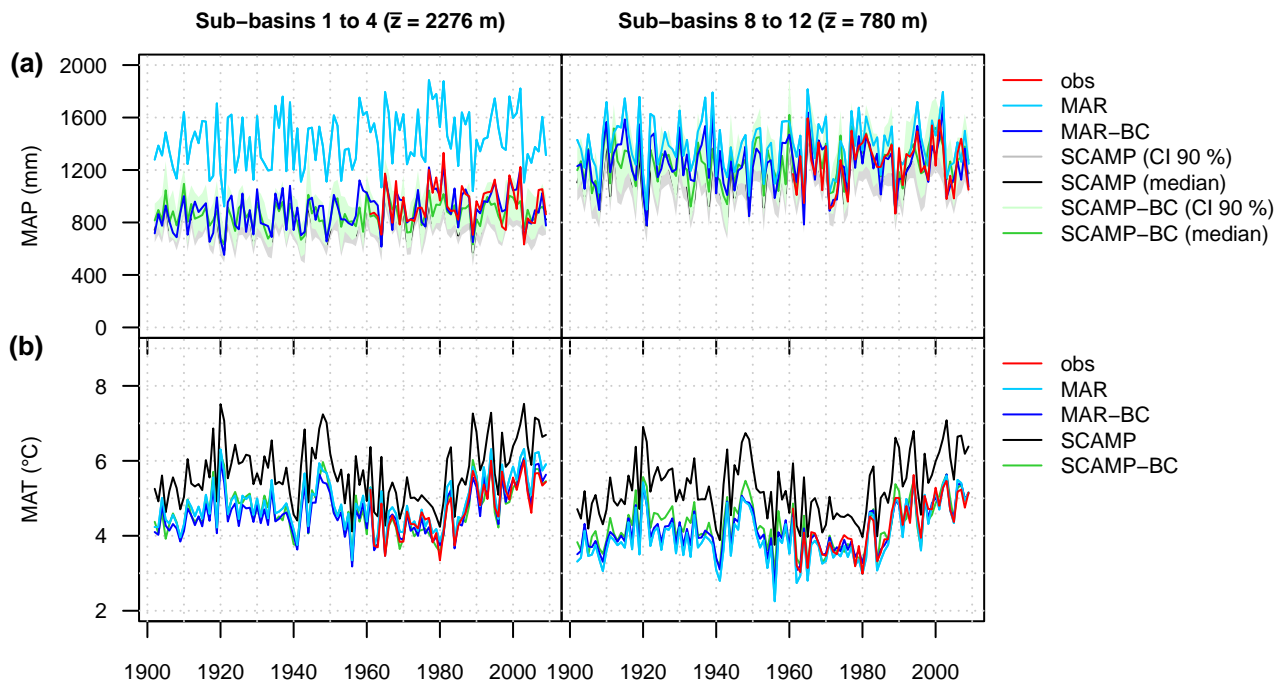
Figure 7. Time series of annual (**a**) ~~Mean Areal Precipitation~~ (MAP) and (**b**) Mean Areal Temperature (MAT) for two major basins (1902-2009). The average elevation of each major basin is indicated in brackets. For MAP, the grey and green bands represent the confidence ~~interval~~ intervals at 90 % level. The median scenarios are indicated by the black and green solid lines. For MAT, the 30 SCAMP time series scenarios are identical and ~~both~~ correspond to the raw or to the ~~BC annual~~ bias-corrected time series of the ERA-20C reanalysis. obs: observed weather. MAR/MAR-BC: raw/bias-corrected weather scenario produced with the dynamical downscaling model MAR. SCAMP/SCAMP-BC: raw/bias-corrected weather scenarios produced with the statistical downscaling model SCAMP.

simulated annual ~~MAP~~ MAPs are 2 % to 8 % smaller than observations around the lake and 10 % to 15 % smaller for the upper part of the URR catchment. Some differences are also obtained for annual ~~MAT~~MATs. They are small for MAR (0.1 °C around the lake, 0.3 °C in the upper part of the URR catchment). For SCAMP, the simulations are roughly 1 °C warmer for the whole area (black lines).

425   As shown in Fig. ~~7,~~ 8, the deviations from observations can vary a lot from one season to another. In the upper part of the URR catchment, precipitation amounts simulated with MAR are similar to observations in summer but much larger in fall and spring (resp. +56 % and +71 %) and even larger in winter (+110 %). Around the lake, simulated precipitations are almost similar to observations in fall (+8 %) and summer (-7 %) but are again significantly larger in winter and spring (+24 %). For SCAMP scenarios, differences are almost exclusively found in summer~~when simulated MAP~~, when simulated MAPs

430 are 15 % to 30 % smaller than ~~the~~ observations for the whole area. A significant seasonality of the deviations is also found for ~~MAT~~MATs. For MAR, the difference is -0.7 °C in spring and up to +1.1 °C in winter. For SCAMP, the seasonality is even
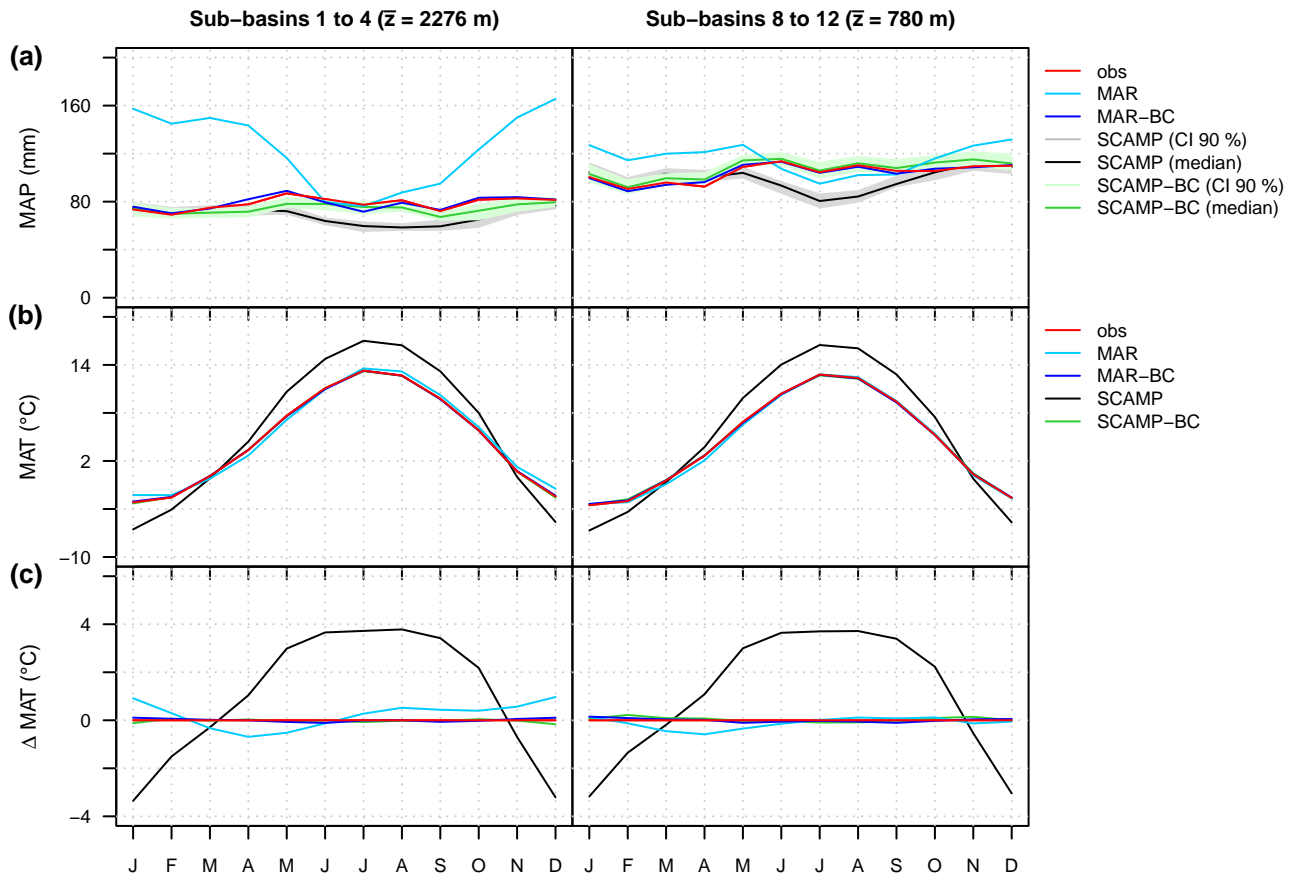
Figure 8. Seasonal cycles of (**a**) Mean Areal Precipitation (MAP) and (**b**, **c**) Mean Areal Temperature (MAT) for two major basins (1961-2009). The average elevation of each major basin is indicated in brackets. For MAP, the grey and green bands represent the confidence ~~interval~~ intervals at 90 % level. The median scenarios are indicated by the black and green solid lines. For MAT, the 30 SCAMP seasonal cycle scenarios are identical and ~~both~~ correspond to the raw or to the ~~BC~~ bias-corrected seasonal cycle of the ERA-20C reanalysis. obs: observed weather. MAR/MAR-BC: raw/bias-corrected weather scenario produced with the dynamical downscaling model MAR. SCAMP/SCAMP-BC: raw/bias-corrected weather scenarios produced with the statistical downscaling model SCAMP.

larger: simulations are up to 3.7 °C warmer than observations in summer and~~conversely~~, conversely, up to 2.7 °C colder in winter.

For both models, the differences mentioned above are significantly reduced and sometimes vanish completely after BC (blue lines for MAR-BC and green bands/solid lines for SCAMP-BC). This is the case for annual variables (Fig. ~~6~~7) but also for monthly variables (Fig. ~~7~~8) as the BC is performed on a monthly basis. For SCAMP, as the BC is performed at the catchment-scale, some biases can remain at the ~~sub-catchment~~ sub-basin scale. For annual ~~MAP~~ MAPs, for instance, simulations are still

3 % to 9 % smaller than observations in the upper part of the URR catchment, and 0 % to 6 % larger around the lake, depending on the scenario (Fig. ~~6~~7).

440 ## 5.2 Discharge seasonality and variations

~~Time series of daily discharges at (**a**) Rhône@Porte-du-Scex, (**b**) Rhône@Genève, Halle-de-l'Ile, (**c**) Arve@Genève, Bout-du-Monde, (**d**) Rhône@Bognes for the 1981-1983 period. (**e**) Time series of daily level of Lake Geneva for the same period. The green bands represent the confidence interval at 90 % level. The median scenarios are indicated by the green solid lines.~~

~~Time series of daily discharges at (**a**) Rhône@Porte-du-Scex, (**b**) Rhône@Genève, Halle-de-l'Ile, (**c**) Arve@Genève, Bout-du-Monde,~~
445 ~~(**d**) Rhône@Bognes for the 1981-1983 period. (**e**) Time series of daily level of Lake Geneva for the same period. The grey bands represent the confidence interval at 90 % level. The median scenarios are indicated by the black solid lines.~~

The discharges obtained via hydrological simulation from weather scenarios are compared with their reference counterparts, i.e. with the discharges obtained via hydrological simulation from observed weather variables. The comparison ~~was~~ is applied to time series of discharges at daily, monthly and annual resolutions, and to time series of characteristic discharge variables
450 (i.e. minimum monthly discharge observed each year and annual maximum daily discharge). ~~Results presented here only cover parts of our comparisons.~~ Note that for SCAMP and SCAMP-BC simulations, for which 30 time series scenarios are simulated, the value of the reference discharge variable considered for a given time is compared to the 30 values obtained for that time from the 30 scenarios respectively.

At a daily time step, the discharge time series of the MAR-BC and SCAMP-BC simulations are in good agreement with
455 the reference ones~~as shown with~~, as shown by the results obtained for the four illustrative gauging stations in Fig. ~~8~~9. The agreement is even larger for time series of mean monthly discharges (Fig. ~~10~~11, left). The large seasonality of flows~~but also the temporal variations(day-to-day variations, monthly variations) of flows is~~, but also their daily and monthly temporal variations, are well reproduced, especially for the Rhône River upstream of Lake Geneva. These results are obtained for almost all gauging stations, even ~~for stations~~ those located downstream of Lake Geneva, despite the significant influence of the lake regulation on
460 flows and the rather crude regulation model used for its representation (Fig. ~~8~~9e).

For both downscaling models, the BC of weather scenarios significantly improves the simulations (Fig. ~~8~~9 versus Fig. ~~9).~~ ~~BC is actually~~ 10, Table 2). The BC is required for both precipitation and temperature variables. This is especially visible ~~with in the~~ results obtained for the upstream URR catchment. At Rhône@Porte-du-Scex (Fig. ~~9~~10a), the discharge variations are rather well reproduced with raw weather scenarios but the discharges are overestimated with SCAMP and underestimated with
465 MAR during the spring and summer. These deviations may be surprising, as raw SCAMP and MAR precipitation simulations are biased toward not enough summer precipitation in SCAMP and~~conversely~~, conversely, much larger winter, spring and fall precipitation in MAR (Fig. ~~7~~8). They actually derive from temperature scenarios ~~which~~ that are too warm in SCAMP in summer and not warm enough in MAR. This point is further discussed in Sect. ~~6.2~~6.2.
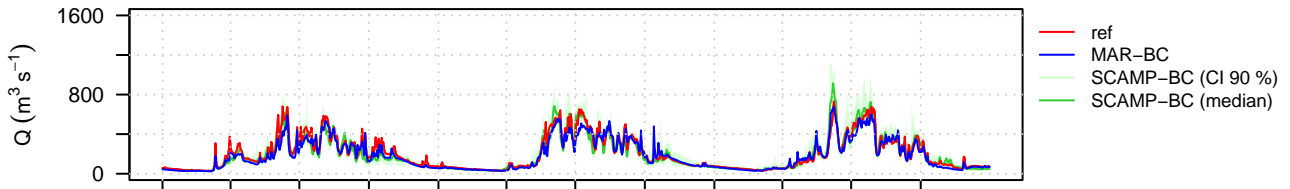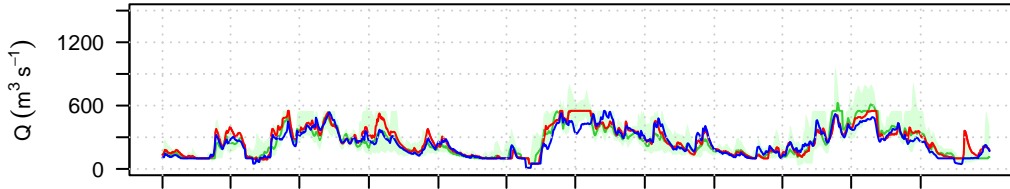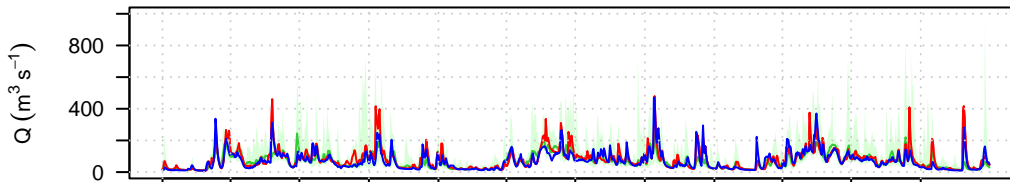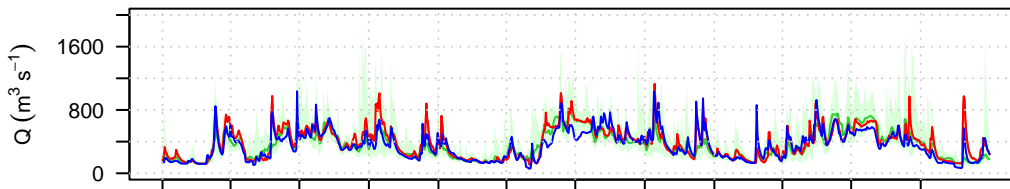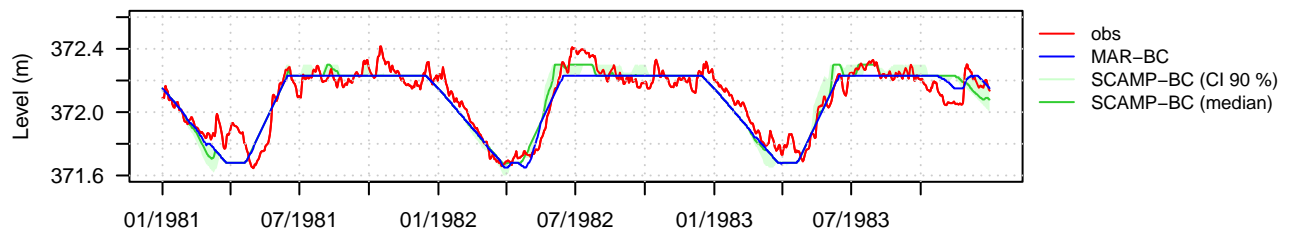
Figure 9. Time series of daily discharges at (**a**) Rhône@Porte-du-Scex, (**b**) Rhône@Genève, Halle-de-l'Ile, (**c**) Arve@Genève, Bout-du-Monde, and (**d**) Rhône@Bognes for the 1981-1983 period. (**e**) Time series of daily level of Lake Geneva for the same period. The green bands represent the confidence intervals at 90 % level. The median scenarios are indicated by the green solid lines. ref: simulated discharge from observed weather variables. MAR-BC: simulated discharge from the bias-corrected weather scenario produced with the dynamical downscaling model MAR. SCAMP-BC: simulated discharge from the bias-corrected weather scenarios produced with the statistical downscaling model SCAMP. obs: observed level of Lake Geneva.
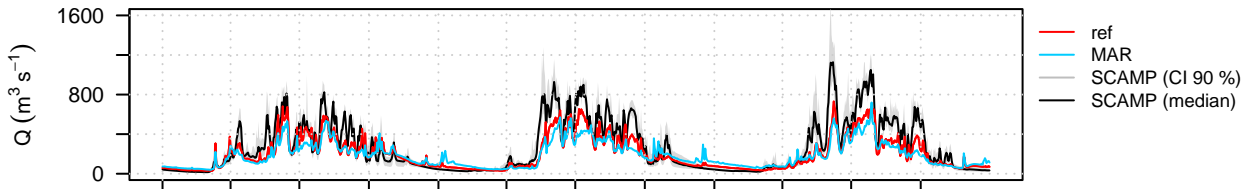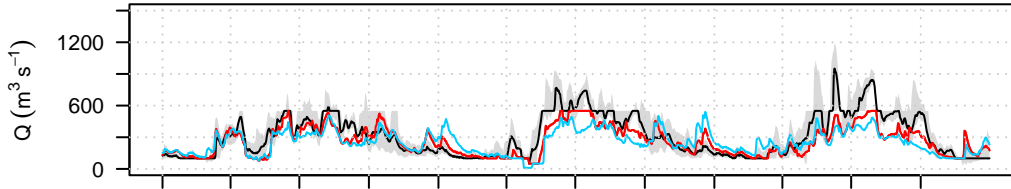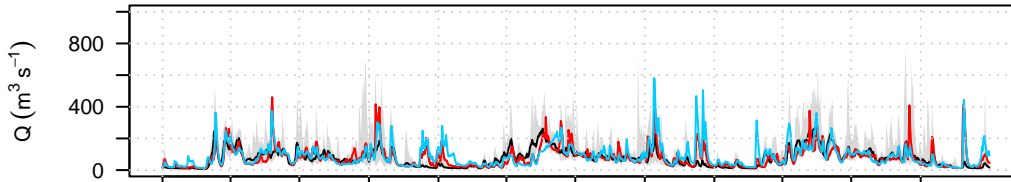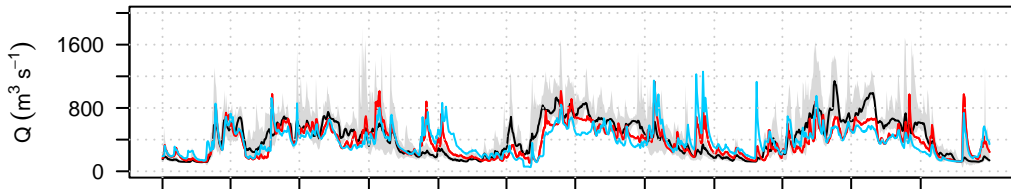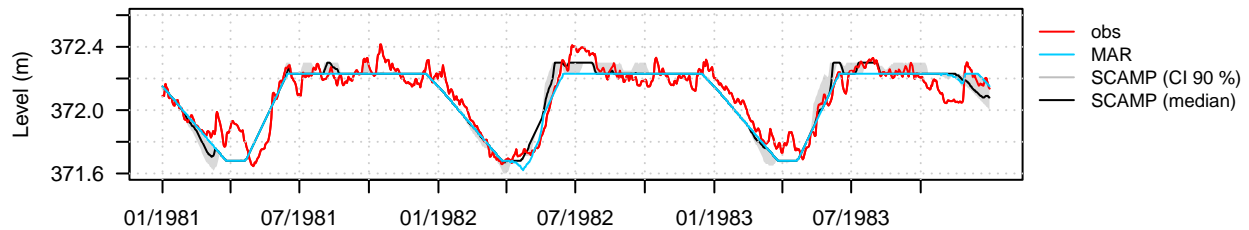
Figure 10. ~~Scatter plots~~ Time series of ~~mean monthly~~ daily discharges ~~, low flows and flood events~~ at (**a**) Rhône@Porte-du-Scex, (**b**) Rhône@Genève, Halle-de-l'Ile, (**c**) Arve@Genève, Bout-du-Monde, and (**d**) Rhône@Bognes for the ~~1961-2009~~ 1981-1983 period. ~~See text~~ (**e**) Time series of daily level of Lake Geneva for ~~definitions~~ the same period. The grey bands represent the confidence intervals at 90 % level. The median scenarios are indicated by the black solid lines. ref: simulated discharge from observed weather variables. MAR: simulated discharge from the raw weather scenario produced with the dynamical downscaling model MAR. SCAMP: simulated discharge from the raw weather scenarios produced with the statistical downscaling model SCAMP. obs: observed level of ~~low flows and flood events~~Lake Geneva.

Table 2. Performance assessment of the modeling chains for the 1961-2009 period. For the dynamical downscaling model (MAR/MAR-BC), the NSE coefficient is calculated from the reference and the simulated discharge time series at each gauging station. For the statistical downscaling model (SCAMP/SCAMP-BC), the statistical index CRPSS (Continuous Ranked Probability Skill Score; Chardon et al., 2014) is used for the evaluation. For each gauging station, the CRPSS compares the probabilistic predictions of discharge obtained for each simulation day from i) the climatological distribution and ii) the 30 simulated time series scenarios. For a perfect model: NSE = 1 and CRPSS = 1. For a model worse than climatology: NSE < 0 and CRPSS < 0.

	MAR	MAR-BC	SCAMP	SCAMP-BC
Rhône@Porte-du-Scex	0.84	0.90	0.10	0.68
Rhône@Genève, Halle-de-l'Ile	0.71	0.83	0.11	0.63
Arve@Genève, Bout-du-Monde	0.33	0.55	0.23	0.43
Rhône@Bognes	0.60	0.74	0.14	0.55

5.3 Flood events and low flow sequences

The hydrological relevance of simulated weather scenarios is further evaluated with simulations of low flow sequences and flood events. ~~Figure 10~~ Note that the annual discharge maxima are used as flood proxy indicators. Figure 11 (middle, right) presents scatter plots of simulated and reference values of annual monthly low flows and annual ~~daily floods~~ discharge maxima obtained for the 49 years of the 1961-2009 period. The same figure with results for the first half of the century is given in ~~Supplementary Materials (Fig. S3).~~ Fig. S8 in Supplementary Materials. For low flows, the month with the lowest monthly flow in the reference discharge series is identified for each year of the period. The 49 lowest monthly low flows are compared to their simulated counterparts for the same months. For ~~flood events~~ annual discharge maxima, a similar comparison is made: for each year, the day with the highest reference daily flow is identified, and the corresponding discharge is compared to the maximum daily discharges obtained around that day in the simulations~~(the~~. The maximum discharge considered for the comparison in the simulation is identified from a 7-day window centered on the reference day~~)~~.

For all gauging stations, the MAR-BC simulation leads to very satisfactory results for low flows. By contrast, ~~flood events~~ the annual discharge maxima tend to be underestimated (at Rhône@Porte-du-Scex, Rhône@Genève, Halle-de-l'Ile and Rhône@Bognes) or poorly simulated (at Arve@Genève, Bout-du Monde). For the SCAMP-BC simulations, the low flows and ~~flood events~~ the annual discharge maxima are characterized by a very large inter-scenario variability, making the interpretation of these results more difficult. For low flows, the medians of simulated scenarios are close to the reference ones at all gauging stations. For ~~flood events~~ annual discharge maxima, the medians are also close to the reference values at Rhône@Porte-du-Scex. They are~~however~~, however, off center downwards ~~at Rhône@Genève, Halle-de-l'Ile and at the two stations~~ on the Rhône River ~~downstream of Lake Geneva~~ for the other three stations. The least well reproduced ~~flood events~~ annual discharge maxima are those of the Arve basin and those at Rhône@Bognes. Conversely to the upper part of the URR catchment, ~~these floods~~ the
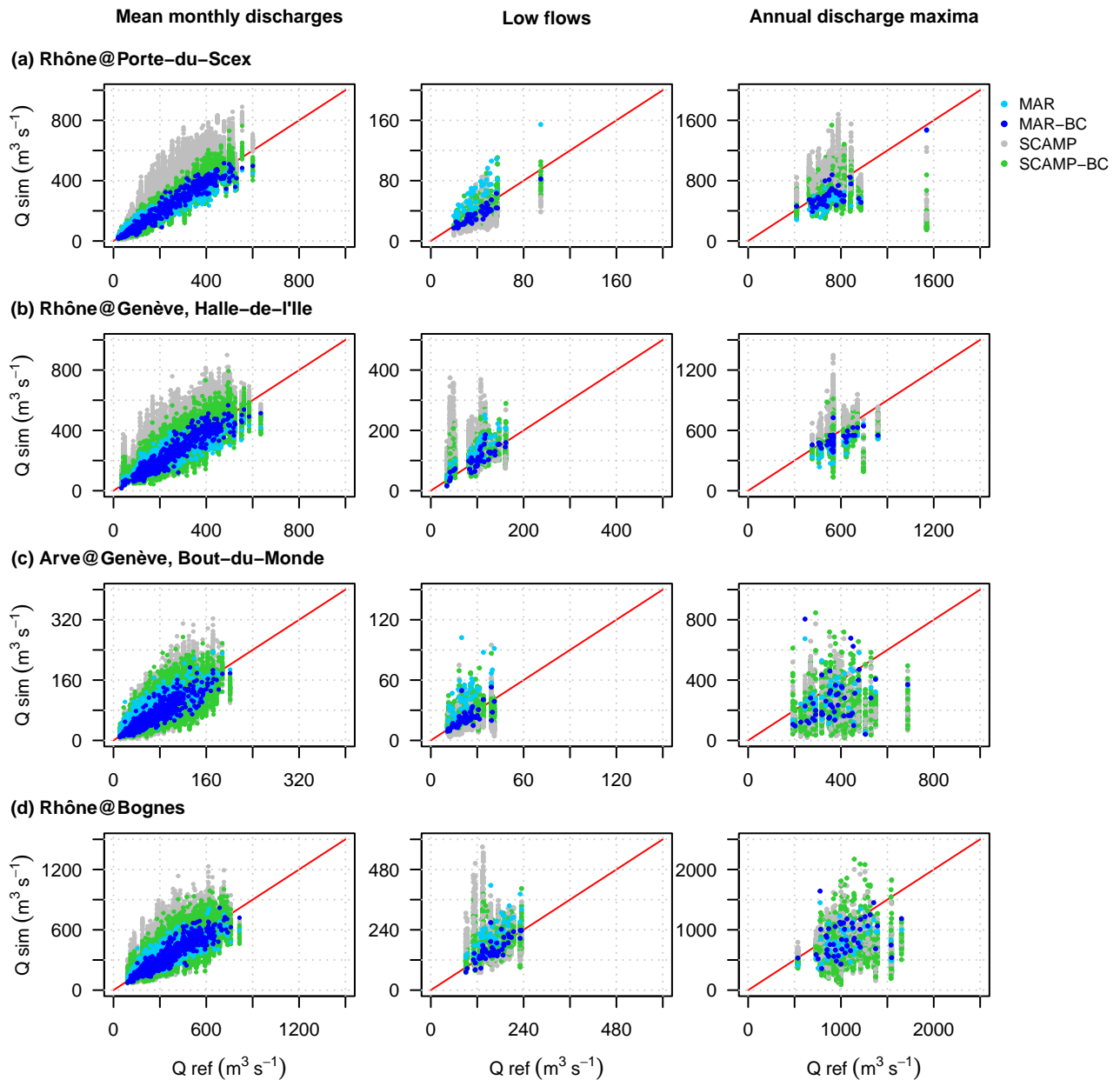
Figure 11. Scatter plots of mean monthly discharges, low flows and annual discharge maxima at (**a**) Rhône@Porte-du-Scex, (**b**) Rhône@Genève, Halle-de-l'Ile, (**c**) Arve@Genève, Bout-du-Monde, and (**d**) Rhône@Bognes for the 1961-2009 period. Q ref: simulated discharge from observed weather variables. MAR/MAR-BC: simulated discharge from the raw/bias-corrected weather scenario produced with the dynamical downscaling model MAR. SCAMP/SCAMP-BC: simulated discharge from the raw/bias-corrected weather scenarios produced with the statistical downscaling model SCAMP.

490 annual discharge maxima occur mainly in late summer and fall (see Fig. ~~S4~~ S9 in Supplementary Materials). The MAR-BC and SCAMP-BC simulations may thus fail to reproduce the large rainfall amounts in that season, especially convective events, known to generate the largest fall floods in this area.

~~(**a**) Flood activity and (**b**) low flow activity at Rhône@Bognes for three 30-year sub-periods: 1920-1949, 1950-1979 and 1980-2009. See text for definitions of flood activity and low flow activity. (**c**) Mean annual discharge at Rhône@Bognes for~~
495 ~~the period 1902-2009. The grey and green bands represent the confidence interval at 90 % level. The median scenarios are indicated by the black and green solid lines.~~ For flood and low flow events, the added value of a BC for weather scenarios is again important (Fig. 11). Depending on the hydrological variable considered, the added value of a temperature BC is not necessarily equivalent to that of a precipitation BC. In the upstream parts of the URR catchment, for instance, low flows occur mainly in winter, due to the low temperatures in this season. The quality of simulated winter low flows thus depends to a
500 great extent on the quality of temperature scenarios, and much less on the quality of precipitation scenarios. Conversely to precipitation corrections, temperatures corrections therefore lead to a significant improvement in low flow simulations. This is illustrated by the additional analyses presented in the discussion (Sect. 6.2).

5.4 Mean annual discharges, flood and low flow activities

The ~~simulation~~ modeling chains considered previously are ~~typically~~ commonly used for hydroclimatic projections with large-
505 scale atmospheric ~~outputs of GCM~~ GCM outputs as forcing variables. The potential impact of climate change on hydrology is often assessed considering changes in the statistical characteristics of hydrological regimes ~~, for instance~~ (e.g. changes in seasonality and year-to-year variability, changes in flood and low flow activities~~, etc~~). We therefore also assess the ability of the considered ~~simulation~~ modeling chains to simulate the ~~observed variations of different hydrological regime features~~ "reference" variations in different characteristics of the URR hydrological regime over the last ~~100 years.~~ century (1920-2009).
510 In this section, as no better reference time series is available, the "references" are observed discharges for the 1920-1960 period and simulated discharges from observed weather variables for the 1961-2009 period. The results are therefore to be interpreted with caution.

~~We first consider variations of~~

The simulated year-to-year variations in mean annual discharges are first compared to the "reference" ones over the 1920-2009
515 period (Fig. 12 c, d). Due to the large interannual variability of discharges, it would be rather difficult to assess the ability of the modeling chains to catch any long-term trends in this variable. However, the year-to-year variations are well reproduced in both timing and amplitude, especially for the second half of the century. Recall especially that the hydrological model does not alter the comparison during this sub-period, as the two time series compared are obtained by hydrological simulation.

We then consider the variations in flood and low flow activities of the URR catchment. Flood activity is ~~here defined as the~~
520 ~~number of flood events within a~~ defined here as the average number of daily discharges per 30-year period ~~exceeding a given~~ above a given daily discharge threshold. Low flow activity is similarly defined as the average number of months ~~within a~~ per 30-year period ~~for~~ during which the mean monthly discharge is below a given discharge threshold. The thresholds retained are the reference discharge values exceeded ~~in average one time per year on~~ on average once a year over the entire 90-year
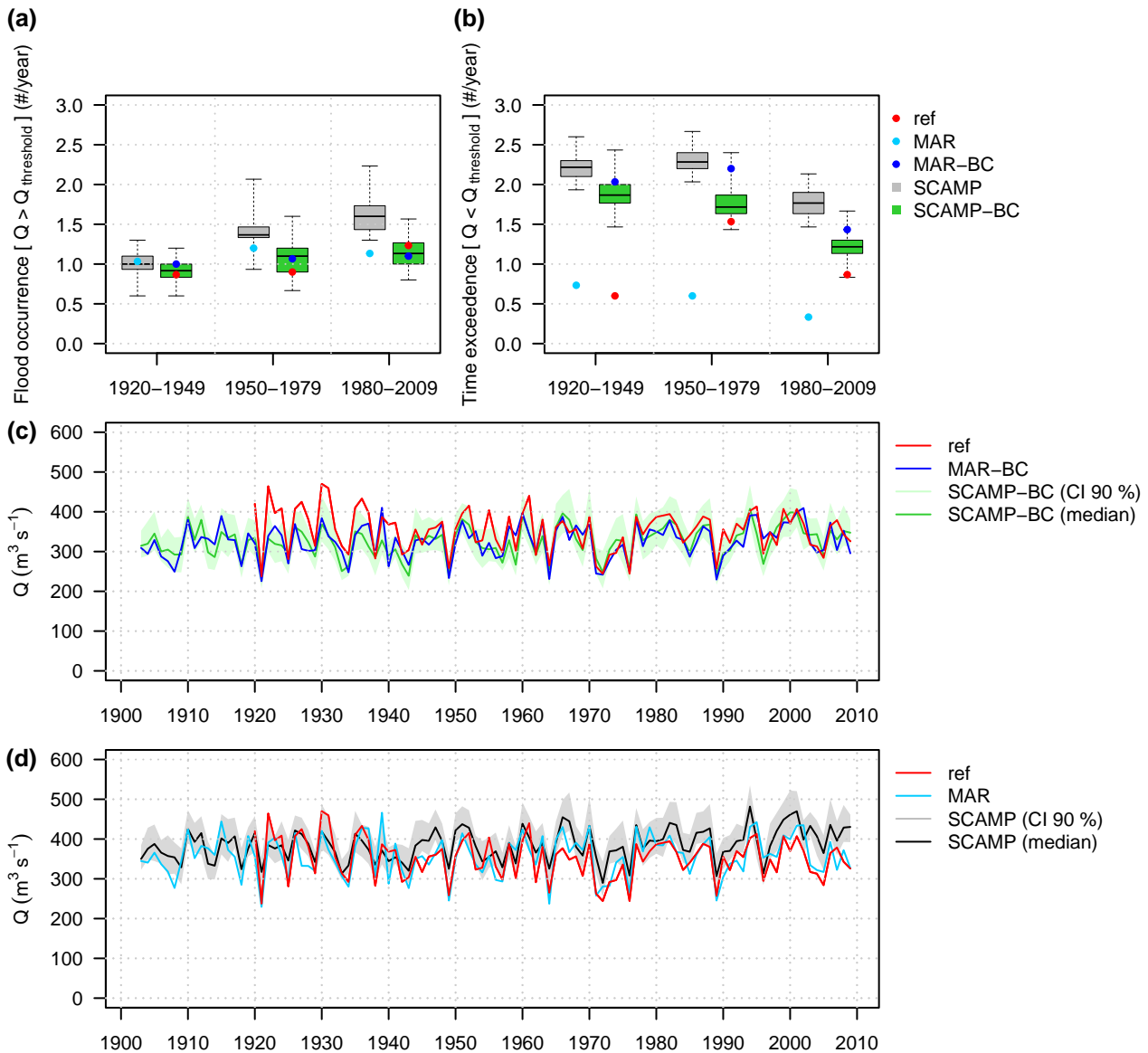
Figure 12. (a) Flood activity and (b) low flow activity at Rhône@Bognes for three 30-year sub-periods: 1920-1949, 1950-1979 and 1980-2009. See text for definitions of flood activity and low flow activity. (c) Mean annual discharge time series at Rhône@Bognes for the 1902-2009 period simulated with MAR-BC and SCAMP-BC. (d) The same for simulations with MAR and SCAMP. The grey and green bands represent the confidence intervals at 90 % level. The median scenarios are indicated by the black and green solid lines. As no better reference time series is available, note that the "references" are observed discharges for the 1920-1960 period and simulated discharges from observed weather variables for the 1961-2009 period. MAR/MAR-BC: simulated discharge from the raw/bias-corrected weather scenario produced with the dynamical downscaling model MAR. SCAMP/SCAMP-BC: simulated discharge from the raw/bias-corrected weather scenarios produced with the statistical downscaling model SCAMP.

simulation period (1920-2009~~.~~). For flood activity, the threshold was calculated by considering the 90 largest daily discharges over the 90-year period. If the threshold was exceeded several times during five consecutive days, only the date of the maximum discharge was retained. For low flow activity, the threshold corresponds to the 10th percentile, i.e. the 90 lowest values of mean monthly reference discharges over the 90-year period. At Rhône@Bognes, the thresholds are respectively 962 $\mathrm{m^3\ s^{-1}}$ for floods and ~~162~~167 $\mathrm{m^3\ s^{-1}}$ for low flows. The flood and low flow activities, estimated for each of the three 30-year sub-periods 1920-1949, 1950-1979 and 1980-2009, are presented in Fig. ~~11~~12 (a, b).

~~The~~ In both BC simulations, the number of flood events exceeding the threshold ~~in the MAR-BC and SCAMP-BC simulations as well as the variations of~~ and the variations in flood activity from one sub-period to another are in good agreement with the reference ones. The observed ~~flood activity increase is well~~ increase in flood activity is rather well reproduced ~~in the simulations~~ (Fig. ~~11~~12a). The results for low flow activity are less satisfactory ~~, especially for the 1920-1949 sub-period~~ (Fig. ~~11~~12b). ~~The number of low flow sequences~~ On the one hand, the number of mean monthly discharges below the threshold ~~in both simulations is~~ indeed ~~twice~~ is, whatever the sub-period, significantly higher than that of the reference. ~~This suggests a limitation of both downscaling models to simulate long persistent dry sequences.~~

~~Simulated year-to-year variations of mean annual discharges are next compared to the "reference" ones over the 1920-2009 period (Fig. 11c). The "references" are observed discharges for the 1920-1960 period and simulated discharges from observed weather for the 1961-2009 period. Because of the large interannual variability of discharges, it would be rather difficult to assess the ability of the chains to catch possible long term trends of this variable. However, the year-to-year variations are well reproduced in both timing and~~ amplitude especially for the second half of the century (recall especially that the hydrological model does not alter the comparison during this~~ It is more than three times higher for the 1920-1949 sub-period~~as both time series compared are obtained by hydrological simulation).~~. This overestimation is mainly due to longer winter low flow durations (not shown). On the other hand, the variations in low flow activity from one sub-period to another are only partially reproduced. While the small decrease observed between the last two sub-periods is rather well captured, the large increase between the first two sub-periods is fully missed. The reasons for this are unclear. The large differences obtained with the raw and bias-corrected downscaling simulations suggest that some limitations remain in the bias-corrected weather scenarios (e.g. too low simulated winter temperatures). The large differences in the reference low flow activity obtained between the sub-periods also suggest an issue with the stationarity assumption (e.g. hydrological signatures) and/or with the low temporal homogeneity of the data set considered in this work (e.g. the reference discharge time series made from observed/simulated discharges for the beginning/end of the period). These points will be worth further investigations in future works.

6 Discussion

All in all, and as already illustrated in ~~some previous works (e.g. Boé et al., 2007; Kuentz et al., 2015; Caillouet et al., 2017)~~a number of previous works (e.g. Boé et al., 2007; Kuentz et al., 2015; Bonnet et al., 2017; Caillouet et al., 2017; Weber et al., 2021), hydrologically relevant weather scenarios (or reconstructions) can be achieved with either statistical or dynamical downscal-

ing models from large-scale atmospheric information only. As ~~illustrated here also~~also illustrated here, this may require some preliminary ~~BC~~ BCs to atmospheric model outputs.

As discussed in the following, the need for corrections can be attributed to some limitations of the models. It may be also attributed to the quality of the available "observations". We will discuss issues related to reanalysis data and to lapse rates for both temperature and precipitation. We will also discuss some issues related to the hydrological model considered in the modeling chains.

6.1 Reanalysis data

The ~~ERA-20C reanalysis considered for~~ global atmospheric reanalysis ERA-20C considered in this study is used as pseudo-observation of the ~~atmospheric~~ state and dynamics of the atmosphere over a large spatial area covering the European domain. ~~It is however produced by assimilation of~~ As it is produced by assimilating only sea level pressure and wind measurements~~only~~, ERA-20C data are not, however, free of limitations. The quality of the geopotential at 500 hPa and of ~~the~~ other large-scale variables (~~like~~ such as vertical velocities at 600 ~~vertical velocities~~ hPa, temperature at 2 m, large-scale precipitation~~, etc~~) may be ~~therefore rather low for some variables~~ rather low and may impact the skill of both downscaling models.

This is ~~for instance the case~~for the ~~case, for instance, for the~~ regional MAT ~~value~~ time series used to force the SCAMP time series scenarios. The large bias in the regional MAT SCAMP scenarios highlighted in Fig. ~~6 and 7~~ 7 and 8 directly derives from the bias ~~of~~in the regional MAT of the ERA-20C reanalysis over the considered domain. Similar limitations were reported by Bonnet et al. (2017), who had to correct the biases of a downscaled version of the ERA-20C reanalysis to simulate realistic mean river flows in France.

As shown by Horton and Brönnimann (2019), using a reanalysis assimilating more data~~like~~, such as the recent ERA5 reanalysis (Hersbach et al., 2020), could lead to more relevant weather scenarios. Such reanalyses were not used in the present work~~because~~, as they generally cover a much shorter ~~time period (about~~ period (around 60 years) preventing the simulation and evaluation of hydro-meteorological scenarios over a century.

6.2 Temperature lapse rate

~~For hydrological simulations, MAT estimates are required for each elevation band. They are obtained via interpolation from temperature values available at neighboring stations using a regional elevation-temperature relationship.~~

~~For SCAMP simulations, this relationship is estimated each day from the SCAMP temperatures at the stations. For each simulation day, as these temperatures are derived from the temperatures of an observed analog day, the relationship of the scenario always corresponds to an observed one. The elevation-temperature relationship and its variations with time are thus coherent with the observations.~~

~~For MAR simulations, this relationship was assumed to be linear and its slope, the so-called "lapse rate", was estimated for each time step from simulated temperatures using all MAR grid cells in the URR catchment. As illustrated in Fig. 12a~~As shown in Sect. 4.4 (Fig. 5a), the lapse rates estimated from MAR ~~temperatures are in average stronger~~ simulations are on average

higher than those estimated from ~~temperature observations. This likely results from a warm bias in the lower atmospheric layers of the model.~~

~~(a) Elevation-temperature relationship at the catchment-scale. Each point corresponds to the mean temperature over the period 1961-2009 observed at each station (n = 39) or simulated at each grid cell of the model MAR (n = 281). (b) and (c) Examples of empirical cumulative distribution functions (ecdf) and correction model for temperature lapse rate (sub-basins 1 to 4, July). (d) Time series of daily discharges at Rhône@Porte-du-Scex for the 1981-1983 period.~~

observations. For the mountainous context considered here, this bias has important implications on the simulated hydrology. A ~~stronger~~ higher lapse rate leads to lower temperatures than those observed for high elevation bands ~~especially~~ in particular (where few observations are available) and vice versa for low elevation bands. This ~~makes logically~~ logically makes the simulated snowpack dynamics significantly different from the observed one. Lower temperatures lead to more frequent solid precipitation, more snow accumulation, less ~~spring snowmelt , presence of a~~ snowmelt in spring, snow cover for longer ~~duration~~ durations and over larger areas, etc. For the highest elevation bands, lower temperatures can even lead to a ~~perennial simulated~~ simulation of a perennial snowpack, preventing any ~~ice melt simulation~~ simulation of ice melt (which can only occur in the model for elevation bands where the glacier is free of snow). All of this results in poorly simulated hydrological regimes. As already suggested in Sect. 5.4, a biased lapse rate is, ~~with,~~ for instance, ~~a period of~~ expected to give a poor simulation of winter low flow characteristics. A higher lapse rate is also expected to lead to delayed snowmelt floods and to low ~~flow in winter that is too long, a delayed snowmelt flood, too low~~ flows in late summer that are not sustained as they should ~~be~~ by ice melt~~, etc~~ ~~(Fig. 12d). Similarly, flood regimes are expected to be poorly simulated as the fraction of liquid precipitation available for flood generation is probably mis-estimated.~~ .

~~The bias in temperature simulated with RCMs is recognised and corrected for a long time. A common approach to use model temperatures for hydrological simulations is to identify the BC function for a given reference altitude z_{ref} and to use this function to correct model temperatures for all other elevations. In this process however, the model lapse rates remain unchanged and the corrected MAT may still present residual biases for all elevations different from the reference one. This is the case for~~ If less evident, a biased lapse rate may also give a poor simulation of flood regimes. In the URR catchment, the largest floods often occur in fall due to large precipitation amounts. In fall, the temperatures can be low enough for precipitation to fall as snow in high elevation areas. Such situations lead to "reduced floods" compared to floods that would have occurred if all the ~~present work.~~

~~We thus considered here a two-part quantile mapping correction function of MAR temperatures that accounts for both biases, that of the MAT value simulated for a reference elevation and that of the lapse rate. The corrected MAT value for a given elevation z and a given time t is obtained as follows:~~

$$\text{MAT}_{\text{MAR-BC}}(z,t) = \text{MAT}_{\text{MAR-BC}}(z_{ref},t) + \text{lapse}_{\text{MAR-BC}}(t) \times (z - z_{ref})$$

~~where $\text{MAT}_{\text{MAR-BC}}(z_{ref},t)$ is the~~ precipitation had fallen as rain. This situation was observed, for instance, during the flood of October 15, 2000, making the flood damage in the region much smaller than expected (Hingray et al., 2010). The temperature lapse rate is determinant, as it defines the elevation of the snowfall/rainfall limit and therefore the "effective area" of the
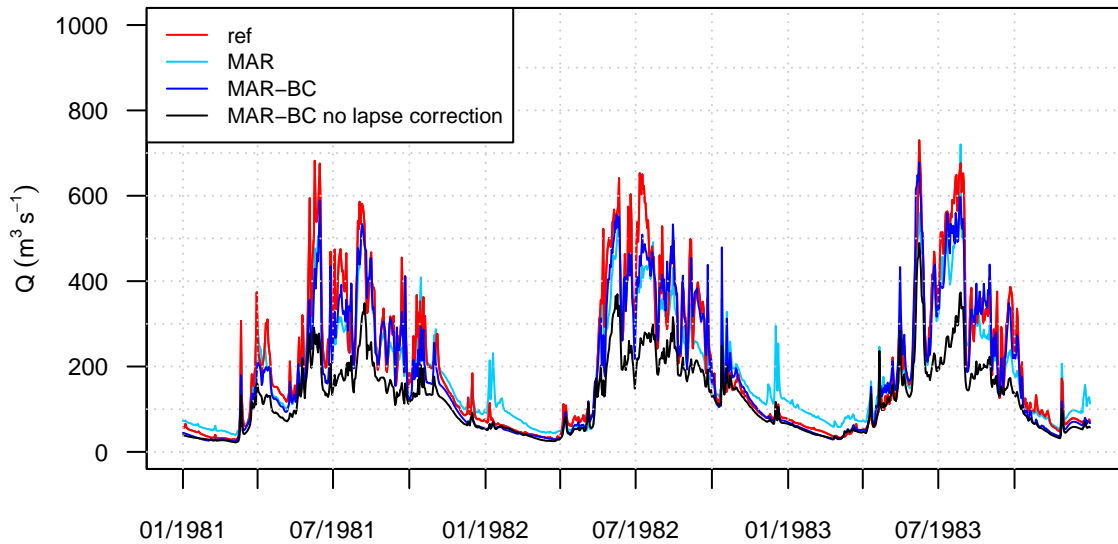
Figure 13. Influence on simulated hydrology of the bias in temperature lapse rate for the dynamical downscaling model MAR. Illustration with the time series of daily discharges at Rhône@Porte-du-Scex for the 1981-1983 period. ref: simulated discharge from observed weather variables. MAR: simulated discharge from the raw weather scenario. MAR-BC: simulated discharge from the bias-corrected weather scenario considering both the bias correction of the Mean Areal Temperature (MAT) for a given reference elevation and the bias correction of the temperature lapse rate. MAR-BC no lapse correction: simulated discharge from the bias-corrected weather scenario considering only the bias correction of the MAT for a given reference elevation.

catchment for these events. The overestimated lapse rate in the MAR model, which results in too cold weather in high elevations areas, is thus expected to lead to lower intensity floods in fall. This is illustrated by the differences between the two bias-corrected ~~MAT value at the reference altitude z_{ref} (the correction depends on the percentile of the MAT value at t) and~~

625 ~~lapse$_{MAR-BC}(t)$ is the bias-corrected lapse rate value at t (the correction depends on the percentile of the lapse rate value estimated from MAR outputs at t).~~

~~In practice, we chose the mean altitude of the URR catchment (i.e. 1525) as reference altitude. The MAT and lapse rate corrections were carried out independently at the monthly scale and for the 5 major basins shown in Fig. 1. Examples of ecdf and correction functions for MAR lapse rates are shown in Fig. 12 (b, c).~~

630 ~~As illustrated~~ experiments produced with MAR in Fig. ~~12d, the~~ 13. When MAR is not corrected for the lapse rate, the intensity of fall floods is significantly lower than when it is.

The added value of the temperature lapse rate correction for hydrological simulations is ~~definitively large~~thus clearly significant, due to the direct effects of temperature on snow dynamics ~~. To~~ (snow/rain repartition, snowpack evolution). To the best of our knowledge, the issue of the temperature lapse rate ~~issue was not given a lot of~~ has not received much attention

635 in the past, but it ~~likely should~~ should probably receive more, at least in areas covering large elevation ranges and where highly

non linear non-linear behaviors with respect to temperature have to be simulated. These results should also lead scientists to integrate this issue when evaluating dynamical downscaling models, and to consider appropriate BC approaches before using model outputs to force impact models.

6.3 Orographic precipitation enhancement

640 A similar, but different, issue arises for precipitation. As mentioned abovein Sect. 5.1, significant differences are obtained between MAP MAPs estimated from downscaled simulations and from observations. This obviously translates directly into significant differences in the simulated hydrology. However, developing the development of relevant MAP estimates is always still a challenge in hydrometeorology, especially hydro-meteorology, particularly in mountainous areas (e.g. Ruelland, 2020).

In the present work, the MAP study, the MAPs estimated for the reference hydrological simulation were are obtained from
645 station observations using the Thiessen's weighting method. For the reasons mentioned in the following, no precipitation-elevation relationship was considered in this estimation. However, the annual precipitation generally increases with elevation in the region. Such a This dependency is rather clear from observations. It is actually also found in the MAR simulations, even if although the simulated precipitation lapse rates may overestimate the true ones (Ménégoz et al., 2020).

The "no precipitation lapse rate" assumption retained for our simulations is thus therefore not really valid. To illustrate
650 the influence of this assumption, we carried out auxiliary "reference" simulations using a constant but elevation-dependant adjustment factor for precipitation (see e.g. Viviroli et al., 2022)(e.g. Viviroli et al., 2022). In practice, all precipitation data of a given station are multiplied by a constant value, depending on the difference between the elevation of the station and that of the target hydrological unit, before application of the Thiessen's weighting process. Two "with precipitation lapse rate" experiments were producedperformed. The adjustment factors were obtained with assuming a linear increase assumption of 5
655 % and 10 % respectively per 100 m of elevation.

As shown in Fig. 1314, accounting for a precipitation lapse rate has contrasting effects depending on the area considered. It significantly increases the annual MAP estimates for high elevation sub-basins (Fig. 1314a, left), but has almost no influence on MAP estimates for low elevation sub-basins (Fig. 1314a, right). This reflects the typical under-representation of high elevation stations in the region. All in all, accounting for a precipitation lapse rate significantly reduces the differences between annual
660 amounts from observations and MAR simulations.

For hydrological simulations, however, the precipitation-elevation dependency is not trivial to take into account in a relevant way. As shown by Ruelland (2020), if while the orographic enhancement can be clearly identified from annual and seasonal means, it is no more obvious evident at the event-scale, and for any given event, the spatial pattern of precipitation generally depends on where the precipitation event first occurs. This is likely the reason why the precipitation lapse rate is significantly
665 lower in summer than in winter in the region (Fig. 1314b), due to much more frequent convective events (e.g. Ménégoz et al., 2020). According to Bárdossy and Pegram (2013), on a daily scale, the orographic effects generally contribute a small part to precipitation variationvariations. In many cases, the orographic enhancement obtained for long accumulation durations is actually often related to more frequent precipitation in high elevations rather than more intense precipitation. Although not verifiable from available observations, this is probably also the case in the region.
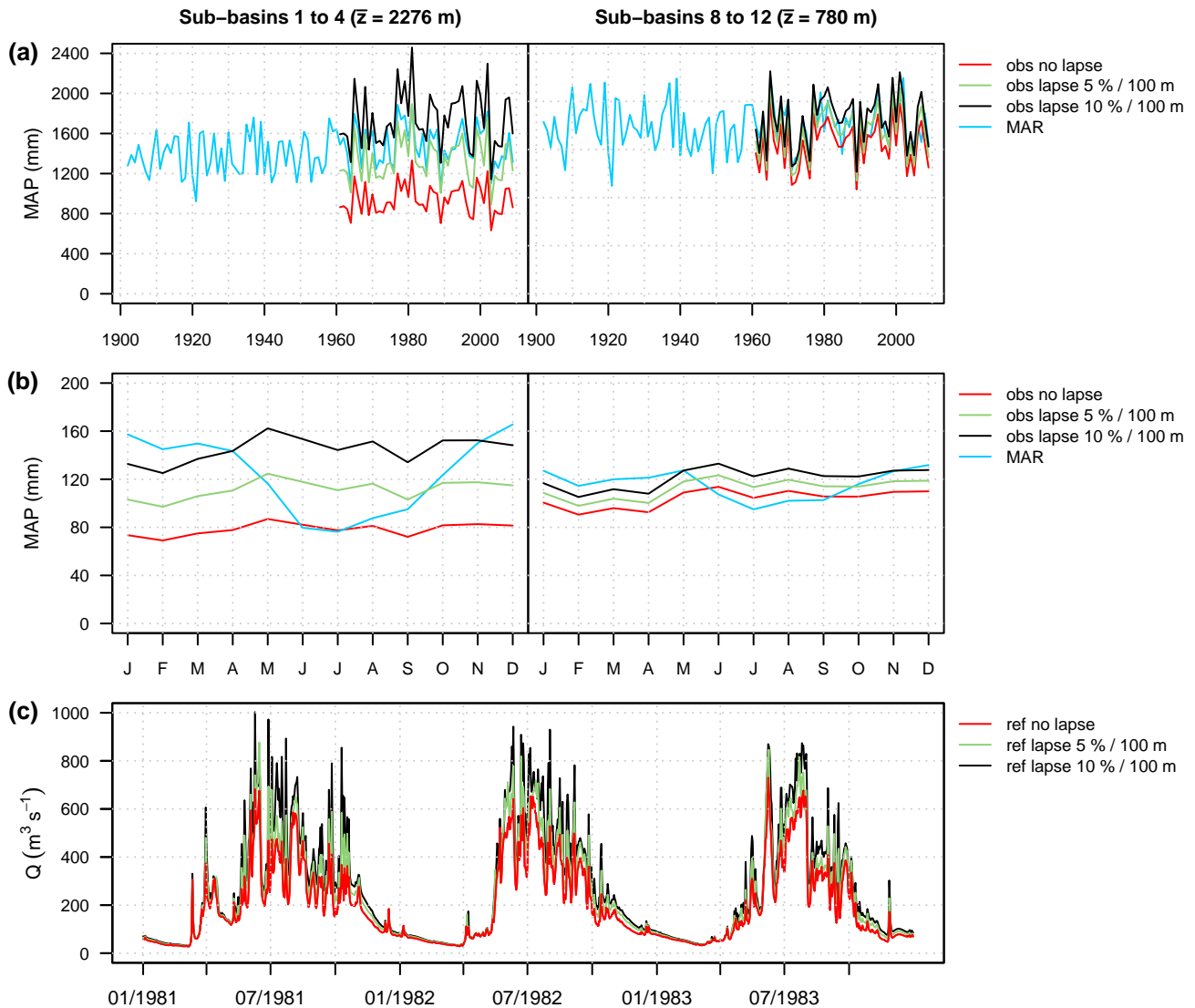
Figure 14. ~~Influence of the precipitation lapse rate.~~ (**a**) Time series of annual Mean Areal Precipitation (MAP) (1902-2009) and (**b**) seasonal cycles of MAP (1961-2009) for two major basins. The average elevation of each major basin is indicated in brackets. obs: observed MAP considering different precipitation lapse rates. MAR: raw MAP scenario produced with the dynamical downscaling model MAR. (**c**) Time series of daily discharges at Rhône@Porte-du-Scex for the 1981-1983 period. ref: simulated discharge from observed weather variables considering different precipitation lapse rates.

670   Although often determinant, orographic effects cannot be ~~accounted for~~ taken into account by a simple and constant adjustment factor, but for practical reasons and lack of better knowledge, this is the usual practice in hydrological modeling. Using a

constant adjustment factor for all time steps of a given period ~~however, is~~ is, however, likely to have significant ~~and detrimental~~ implications on hydrological simulations. This ~~is particularly likely to amplify in an unrealistic way~~ may amplify the largest precipitation events in an unrealistic way, resulting in turn ~~in~~ to huge and unrealistic floods (Hingray et al., 2010). For the

675 URR catchment, hydrological simulations with a precipitation lapse rate produce significantly larger floods (Fig. ~~13~~14c). With a lapse rate of 5 % / 100 m, the ~~maximum annual floods~~ annual discharge maxima of the 1961-2009 period are increased by 3 to 106 % (+25 % on average). With a lapse rate of 10 % / 100 m, the increase is even larger (+56 % on average). It is up to 200 % for the largest flood ever observed in the area (~~15 October~~ October 15, 2000) (not shown).

  The precipitation-elevation dependency was disregarded in the present work, to avoid such unrealistic simulations. This is

680 not really satisfactory, especially regarding the annual water budget of sub-basins. The latter will be mis-represented as a result of the overall dry bias in the precipitation input. Without a better knowledge of how precipitation amounts depend on elevation at the event-scale, disregarding the elevation dependency ~~is probably the best compromise option to choose~~ was found to be an acceptable compromise option if relevant flood events have to be simulated.

  Note that in our simulations, the dry bias ~~in the precipitation input is probably corrected via an adjusted parameterization~~

685 ~~of the hydrological model , allowing for simulated evapotranspiration losseslower than the actual ones and allowing in turn~~ ~~for the simulation of reasonable discharge time series (parameterization adjustments to correct for imperfect input data are~~ ~~expected to be rather common in the calibration of hydrological models~~is probably compensated during hydrological parameter optimization. Parameter optimization generally has the effect of forcing the model to close the water balance. Reservoir-based conceptual models, such as the one used here, can thus compensate for missing precipitation by reducing the evapotranspiration

690 losses. This problem is well known, but rarely discussed. An example is the work of Minville et al. (2014).

6.4 Impact-oriented evaluation of downscaled weather scenarios

In this work, we assessed and compared the ability of two downscaling models to simulate hydrologically relevant weather scenarios. The hydrological model considered here for this assessment is obviously not free of limitations.

  In the model, as mentioned in the previous section, the influence of orography on precipitation was disregarded to avoid the

695 generation of irrelevant floods. This representation is obviously far from satisfactory, and more relevant representations, at the event-scale especially, would be worth considering in future years, when available.

  The mean areal precipitation and temperature were estimated for each spatial unit of the model using the Thiessen's weighting method. This is a rather crude method, and other methods may provide better estimates (e.g. inverse distance weighting or kriging). However, they have also important limitations in mountainous environments (e.g. difficulty to account for the influence of

700 topography on weather spatial patterns at the event-scale). The Thiessen's weighting method was here retained for its simplicity and its ability to take into account the time-varying temperature-elevation relationship in a straightforward way.

  The hydrological model relies also on assumptions, potentially crude. For instance, the characteristics of the sub-basins are assumed not to have changed over the last century. For the glacier cover, this assumption does not really hold. According to Huss (2011), however, the contribution of glacier melt in the region was relatively stable over the 20th century, with a similar

705 glacier contribution during the periods 1961-1990 and 1908-2008. The glacier retreat strengthened over the period 1988-2008,

but the corresponding increase in the glacier contribution to the URR discharges was found to be rather limited (13% in August). While not fully satisfactory, the assumption of a constant glacier therefore seems reasonable.

The signature-based calibration of the model, used for sub-basins with altered discharge data, is also not optimal. Parameters were calibrated so that simulated signatures reproduce at best observed ones, but observed and simulated signatures come from different periods. We thus implicitly assume that the weather regimes and the natural hydrological behavior of the URR sub-basins have not changed significantly over the last century. Both assumptions seem to be reasonable in first approximation. Indeed, over the last century, no significant trends in precipitation have been observed in the region (Masson and Frei, 2016) and the hydrological regimes of the natural sub-basins have remained almost unchanged (see Fig. S10 in Supplementary Materials).

The hydrological model is thus not optimal for a number of reasons. However, this is not expected to influence the main results of our study. Indeed, the model is mainly used here as a complex and non-linear filter to assess the ability of the downscaling chains to simulate hydrologically relevant weather scenarios. In this impact-oriented assessment context, the hydrological model can be imperfect, as it can also be used to produce the hydrological reference against which the hydrological scenarios will be compared. In the context of the URR catchment, the model is a powerful tool to achieve an impact-oriented assessment of downscaled weather scenarios in contrasting and demanding hydro-meteorological configurations, where the interplay between weather variables, both in space and time, is determinant.

7 Conclusion

In ~~the present work~~this study, two hydro-meteorological modeling chains ~~are~~ were used to simulate the past variations ~~of~~ in discharge at several stations of the Upper Rhône River ~~basin~~catchment, a mesoscale ~~basin~~ catchment in the western Alps. ~~Discharges are~~ The discharges were simulated with the glacio-hydrological GSM-SOCONT model using weather scenarios downscaled with the MAR and SCAMP models from ~~ERA-20C data~~ the data of the global atmospheric reanalysis ERA-20C. MAR is a dynamical downscaling model, SCAMP a statistical one providing an ensemble of downscaled scenarios.

~~For both~~ The originality of this study is threefold. i) We evaluated the modeling chains in contrasting and demanding hydro-meteorological configurations where the interplay between weather variables, both in space and time, is determinant. ii) The simulations cover the entire 20th century, a period long enough to assess the ability of the modeling chains to reproduce daily variations in observed discharges, low frequency events, and variations in low flow and flood activities. iii) For both downscaling models, we evaluated the need for additional bias correction of the weather scenarios, including that of temperature lapse rates. This framework allowed to highlight important criteria to be met for the simulation of relevant hydrological scenarios.

For both modeling chains, given the difficult modeling context of the ~~basin~~ URR catchment (e.g. unknown effects of the large upstream dams and regulation of Lake Geneva, scarcity of observations for model calibration) and the fact that the weather scenarios are only produced from large-scale atmospheric information, the simulated discharges are globally in good agreement with the reference ones. For the 1961-2009 period, the multi-scale variations ~~of~~ in reference discharges (daily, seasonal and

interannual) are well reproduced. To some extent, the simulations also reproduce the low flow sequences and ~~annual floods~~ the annual discharge maxima quite well. For the first half of the century, the agreement with the reference discharges is lower (but still reasonable), likely due to lower data quality (ERA-20C and discharges data) and/or to some modeling assumptions and choices (e.g. signature-based calibration, stationarity assumption). Nevertheless, both modeling chains are able to accurately reproduce the variations in flood activity over the last ~~100 years~~ century. The results for low flow activity are less satisfactory.

Both modeling chains are likely to be appropriate for the generation of relevant regional weather scenarios for different climate contexts, from outputs of ad hoc GCM experiments. Thanks to its much lower computational cost, the SCAMP model is to be favored when large ~~ensemble~~ ensembles of climate simulations have to be downscaled. The statistical nature of SCAMP also allows to account for the uncertainty in the downscaling relationship. The ensemble of weather scenarios generated by SCAMP for any large-scale scenario allows thus to simulate and account for the small-scale internal variability of weather. This is another advantage over MAR, allowing a more robust assessment of possible low-frequency changes in hydro-meteorological regimes ~~(e.g. Lafaysse et al., 2014; Raynaud et al., 2020)~~(e.g. Lafaysse et al., 2014). In a future study, we will force the SCAMP/GSM-SOCONT chain with a CMIP6-PMIP4 paleosimulation ensemble (Jungclaus et al., 2017; Kageyama et al., 2018) to assess the variations ~~of~~ in hydro-meteorological regimes of the ~~URR~~ Upper Rhône River catchment over the last millenium. We also expect to confront the simulated variations in flood activity with those obtained in previous works from the sediments archives of Lake Bourget ~~(Jenny et al., 2014; Wilhelm et al., 2022)~~(Jenny et al., 2014; Evin et al., 2019; Wilhelm et al., 2022). Moreover, this could also corroborate the influence of warming and atmospheric circulation changes on multi-decadal flood activity over the last two centuries, recently highlighted by Brönnimann et al. (2022) for different European rivers.

As already shown in previous works, the hydrological behavior of river basins can be simulated from large-scale atmospheric information only. In ~~the present work~~this study, we also ~~show~~ showed that the simulation of ~~relevant hydrological scenarios requires a BC~~ hydrologically relevant weather scenarios required a bias correction of the downscaled weather scenarios. The highly non-linear behavior of hydrological systems does ~~actually~~ not bear biased weather. This was made evident here for temperature, due to the highly non-linear thermal sensitivity of snow variables. If it seeks to be relevant, ~~BC~~ the bias correction step is not necessarily ~~simple (Switanek et al., 2022; ?)~~straightforward (Maraun, 2016; Switanek et al., 2022). For temperature, ~~BC is typically~~ the bias correction is commonly applied for the temperature of reference stations. For the ~~URR catchment, it~~ Upper Rhône River catchment, the bias correction was also needed for the temperature lapse rate simulated by the dynamical downscaling model MAR. The temperature lapse rate correction was determinant to avoid irrelevant simulations of the snowpack dynamics at high elevations, and consequently of the hydrology.

~~BC is~~ The bias correction was also required for precipitation. Significant differences were also found between ~~reference and~~ the reference and the downscaled precipitation, particularly for MAR in winter. ~~BC~~ The bias correction of downscaled precipitation was thus also applied so that the statistical distribution of precipitation scenarios fits that of reference. But the quality of the reference is questionable. It was indeed developed disregarding the significant precipitation-elevation relationship in the area, to avoid the simulation of irrelevant flood events. The much larger precipitation amounts simulated by MAR for high elevation sub-basins are thus not necessarily irrelevant, ~~despite~~ considering the snowfall undercatch issues (Kochendorfer et al., 2017). ~~A~~ An interesting perspective of this work would be to recalibrate the hydrological model using MAR data that provide

these larger precipitation amounts at high elevations, to study their physical realism. Overall, a better understanding of the precipitation-elevation relationship in mountainous areas, especially its likely variations in time and ~~dependence~~ dependency on event types, would improve hydro-meteorological analyses and simulations and make them more relevant. This would benefit from more observations in high altitude areas, a critical issue pointed out for a long time (e.g. Hingray et al., 2012). Despite their multiple limitations in mountainous areas, radar data could also provide valuable insights into orographic drivers of precipitation (e.g. Germann et al., 2022).

If downscaling models clearly need to be refined in the future, scientists should also consider improving ~~BC~~ the bias correction methods for such challenging configurations. They should also consider the possibility of better understanding and accounting for ~~event-scale~~ the precipitation orographic enhancement ~~in the models~~at the event-scale. Dynamical downscaling models such as MAR are likely promising tools for such analyses, but their value for areas with marked relief should be better estimated.

Author contributions. This study is part of CL's PhD thesis. BW and BH supervised the PhD. All authors contributed to the design of the study and the analysis framework. CL performed the simulations and produced the figures presented in this study. CL, MM and BH contributed to the writing of the document and to the editing of the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This work was supported by the French National Research Agency in the framework of the "Investissements d'avenir" program (ANR-15-IDEX-02). A part of the simulations could be performed from the Grenoble Alpes Research Data and Computing Infrastructure GRICAD (https://gricad.univ-grenoble-alpes.fr/, last access: 20 February 2023). We would ~~especially~~ like to thank Mondher Chekki for ~~its~~ his precious technical and informatic ~~help~~support, and Julien Beaumet for providing the outputs of the MAR simulation. We also thank the editor and four anonymous reviewers for their constructive comments, which helped us to improve the manuscript.

References

Arnell, N. W., Hudson, D. A., and Jones, R. G.: Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa, J. Geophys. Res., 108, 4519, https://doi.org/10.1029/2002JD002782, 2003.

Beaumet, J., Ménégoz, M., Morin, S., Gallée, H., Fettweis, X., Six, D., Vincent, C., Wilhelm, B., and Anquetin, S.: Twentieth century temperature and snow cover changes in the French Alps, Reg. Environ. Change, 21, 114, https://doi.org/10.1007/s10113-021-01830-x, 2021.

Bechtold, P., Bazile, E., Guichard, F., Mascart, P., and Richard, E.: A mass-flux convection scheme for regional and global models, Q. J. Roy. Meteorol. Soc., 127, 869–886, https://doi.org/10.1002/qj.49712757309, 2001.

Beniston, M.: Mountain Weather and Climate: A General Overview and a Focus on Climatic Change in the Alps, Hydrobiologia, 562, 3–16, https://doi.org/10.1007/s10750-005-1802-0, 2006.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B., Aronica, G. T., Bilibashi, A., Bonacci, O., Borga, M., Čanjevac, I., Castellarin, A., Chirico, G. B., Claps, P., Fiala, K., Frolova, N., Gorbachova, L., Gül, A., Hannaford, J., Harrigan, S., Kireeva, M., Kiss, A., Kjeldsen, T. R., Kohnová, S., Koskela, J. J., Ledvinka, O., Macdonald, N., Mavrova-Guirguinova, M., Mediero, L., Merz, R., Molnar, P., Montanari, A., Murphy, C., Osuch, M., Ovcharuk, V., Radevski, I., Rogger, M., Salinas, J. L., Sauquet, E., Šraj, M., Szolgay, J., Viglione, A., Volpi, E., Wilson, D., Zaimi, K., and Živković, N.: Changing climate shifts timing of European floods, Science, 357, 588–590, https://doi.org/10.1126/science.aan2506, 2017.

Bonnet, R., Boé, J., Dayon, G., and Martin, E.: Twentieth-Century Hydrometeorological Reconstructions to Study the Multidecadal Variations of the Water Cycle Over France, Water Resour. Res., 53, 8366–8382, https://doi.org/10.1002/2017WR020596, 2017.

Bonnet, R., Boé, J., and Habets, F.: Influence of multidecadal variability on high and low flows: the case of the Seine basin, Hydrol. Earth Syst. Sci., 24, 1611–1631, https://doi.org/10.5194/hess-24-1611-2020, 2020.

Boé, J., Terray, L., Habets, F., and Martin, E.: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies, Int. J. Climatol., 27, 1643–1655, https://doi.org/10.1002/joc.1602, 2007.

Brun, E., David, P., Sudul, M., and Brunot, G.: A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting, J. Glaciol., 38, 13–22, https://doi.org/10.3189/S0022143000009552, 1992.

Brönnimann, S., Stucki, P., Franke, J., Valler, V., Brugnara, Y., Hand, R., Slivinski, L. C., Compo, G. P., Sardeshmukh, P. D., Lang, M., and Schaefli, B.: Influence of warming and atmospheric circulation changes on multidecadal European flood variability, Clim. Past, 18, 919–933, https://doi.org/10.5194/cp-18-919-2022, 2022.

Bárdossy, A.: Calibration of hydrological model parameters for ungauged catchments, Hydrol. Earth Syst. Sci., 11, 703–710, https://doi.org/10.5194/hess-11-703-2007, 2007.

Bárdossy, A. and Pegram, G.: Interpolation of precipitation under topographic influence at different time scales: Interpolation of Precipitation, Water Resour. Res., 49, 4545–4565, https://doi.org/10.1002/wrcr.20307, 2013.

Caillouet, L., Vidal, J.-P., Sauquet, E., Devers, A., and Graff, B.: Ensemble reconstruction of spatio-temporal extreme low-flow events in France since 1871, Hydrol. Earth Syst. Sci., 21, 2923–2951, https://doi.org/10.5194/hess-21-2923-2017, 2017.

CERCG: Règlement sur la manoeuvre de l'ouvrage de régularisation du niveau du lac Léman à Genève, Arrêté du 17 septembre 1997, Genève, Suisse, 1997.

Chagnaud, G., Gallée, H., Lebel, T., Panthou, G., and Vischel, T.: A Boundary Forcing Sensitivity Analysis of the West African Monsoon Simulated by the Modèle Atmosphérique Régional, Atmosphere, 11, 191, https://doi.org/10.3390/atmos11020191, 2020.

Chardon, J., Hingray, B., Favre, A.-C., Autin, P., Gailhard, J., Zin, I., and Obled, C.: Spatial Similarity and Transferability of Analog Dates for Precipitation Downscaling over France, J. Climate, 27, 5056–5074, https://doi.org/10.1175/JCLI-D-13-00464.1, 2014.

Chardon, J., Favre, A.-C., and Hingray, B.: Effects of Spatial Aggregation on the Accuracy of Statistically Downscaled Precipitation Predictions, J. Hydrometeorol., 17, 1561–1578, https://doi.org/10.1175/JHM-D-15-0031.1, 2016.

835 Chardon, J., Hingray, B., and Favre, A.-C.: An adaptive two-stage analog/regression model for probabilistic prediction of small-scale precipitation in France, Hydrol. Earth Syst. Sci., 22, 265–286, https://doi.org/10.5194/hess-22-265-2018, 2018.

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., and Wilby, R.: The Schaake Shuffle: A Method for Reconstructing Space-Time Variability in Forecasted Precipitation and Temperature Fields, J. Hydrometeorol., 5, 243–262, https://doi.org/10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2, 2004.

840 Dayon, G., Boe, J., and Martin, E.: Transferability in the future climate of a statistical downscaling method for precipitation in France, J. Geophys. Res.-Atmos., 120, 1023–1043, https://doi.org/10.1002/2014JD022236, 2015.

De Ridder, K. and Schayes, G.: The IAGL Land Surface Model, J. Appl. Meteorol., 36, 167–182, https://doi.org/10.1175/1520-0450(1997)036<0167:TILSM>2.0.CO;2, 1997.

Devers, A., Vidal, J.-P., Lauvernet, C., and Vannier, O.: FYRE Climate: A high-resolution reanalysis of daily precipitation and temperature
845 in France from 1871 to 2012, Clim. Past, 17, 1857–1879, https://doi.org/10.5194/cp-17-1857-2021, 2021.

Di Sante, F., Coppola, E., and Giorgi, F.: Projections of river floods in Europe using EURO-CORDEX, CMIP5 and CMIP6 simulations, Int. J. Climatol., 41, 3203–3221, https://doi.org/10.1002/joc.7014, 2021.

Doutreloup, S., Wyard, C., Amory, C., Kittel, C., Erpicum, M., and Fettweis, X.: Sensitivity to Convective Schemes on Precipitation Simulated by the Regional Climate Model MAR over Belgium (1987–2017), Atmosphere, 10, 34, https://doi.org/10.3390/atmos10010034,
850 2019.

Déqué, M.: Frequency of precipitation and temperature extremes over France in an anthropogenic scenario: Model results and statistical correction according to observed values, Global Planet. Change, 57, 16–26, https://doi.org/10.1016/j.gloplacha.2006.11.030, 2007.

Evin, G., Wilhelm, B., and Jenny, J.-P.: Flood hazard assessment of the Rhône River revisited with reconstructed discharges from lake sediments, Global Planet. Change, 172, 114–123, https://doi.org/10.1016/j.gloplacha.2018.09.010, 2019.

855 Evin, G., Somot, S., and Hingray, B.: Balanced estimate and uncertainty assessment of European climate change using the large EURO-CORDEX regional climate model ensemble, Earth Syst. Dynam., 12, 1543–1569, https://doi.org/10.5194/esd-12-1543-2021, 2021.

Fowler, H. J., Blenkinsop, S., and Tebaldi, C.: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling, Int. J. Climatol., 27, 1547–1578, https://doi.org/10.1002/joc.1556, 2007.

Froidurot, S., Zin, I., Hingray, B., and Gautheron, A.: Sensitivity of Precipitation Phase over the Swiss Alps to Different Meteorological
860 Variables, J. Hydrometeorol., 15, 685–696, https://doi.org/10.1175/JHM-D-13-073.1, 2014.

Gallée, H.: Simulation of the Mesocyclonic Activity in the Ross Sea, Antarctica, Mon. Weather Rev., 123, 2051–2069, https://doi.org/10.1175/1520-0493(1995)123<2051:SOTMAI>2.0.CO;2, 1995.

Gallée, H. and Duynkerke, P. G.: Air-snow interactions and the surface energy and mass balance over the melting zone of west Greenland during the Greenland Ice Margin Experiment, J. Geophys. Res.-Atmos., 102, 13 813–13 824, https://doi.org/10.1029/96JD03358, 1997.

865 Gallée, H. and Schayes, G.: Development of a 3-Dimensional Meso-Gamma Primitive Equation Model - Katabatic Winds Simulation in the Area of Terra-Nova Bay, Antarctica, Mon. Weather Rev., 122, 671–685, https://doi.org/10.1175/1520-0493(1994)122<0671:DOATDM>2.0.CO;2, 1994.

Gallée, H., Pettré, P., and Schayes, G.: Sudden Cessation of Katabatic Winds in Adélie Land, Antarctica, J. Appl. Meteorol., 35, 1142–1152, https://doi.org/10.1175/1520-0450(1996)035<1142:SCOKWI>2.0.CO;2, 1996.

870 Gallée, H., Guyomarc'h, G., and Brun, E.: Impact Of Snow Drift On The Antarctic Ice Sheet Surface Mass Balance: Possible Sensitivity To Snow-Surface Properties, Bound.-Lay. Meteorol., 99, 1–19, https://doi.org/10.1023/A:1018776422809, 2001.

Gangopadhyay, S., Clark, M., and Rajagopalan, B.: Statistical downscaling using K-nearest neighbors, Water Resour. Res., 41, W02 024, https://doi.org/10.1029/2004WR003444, 2005.

Germann, U., Boscacci, M., Clementi, L., Gabella, M., Hering, A., Sartori, M., Sideris, I. V., and Calpini, B.: Weather Radar in Complex 875 Orography, Remote Sens., 14, 503, https://doi.org/10.3390/rs14030503, 2022.

Giorgi, F. and Mearns, L. O.: Approaches to the simulation of regional climate change: A review, Rev. Geophys., 29, 191–216, https://doi.org/10.1029/90RG02636, 1991.

GLIMS: Glacier Database, Boulder Colorado, USA, NASA, National Snow and Ice Data Center, https://doi.org/10.7265/N5V98602, last access: 12 September 2020, 2015.

880 Grandjean, P.: La régularisation du lac Léman, Hydrology in Mountainous Regions, I - Hydrological Measurements: the Water Cycle, Proceedings of two Lausanne Symposia, IAHS, 193, 769–776, 1990.

Hanssen-Bauer, I., Achberger, C., Benestad, R. E., Chen, D., and Førland, E. J.: Statistical downscaling of climate scenarios over Scandinavia, Climate Res., 29, 255–268, https://doi.org/10.3354/cr029255, 2005.

Harris, I., Jones, P., Osborn, T., and Lister, D.: Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset, 885 Int. J. Climatol., 34, 623–642, https://doi.org/10.1002/joc.3711, 2014.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., 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., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, 890 J.: The ERA5 global reanalysis, Q. J. Roy. Meteorol. Soc., 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hingray, B., Schaefli, B., Mezghani, A., and Hamdi, Y.: Signature-based model calibration for hydrological prediction in mesoscale Alpine catchments, Hydrolog. Sci. J., 55, 1002–1016, https://doi.org/10.1080/02626667.2010.505572, 2010.

Hingray, B., Dedieu, J.-P., Lebel, T., Obled, C., Sicart, J.-E., Six, D., Vincent, C., Wagnon, P., and Zin, I.: Observations glaciologiques et hydrométéorologiques en zone de montagne : quelques problématiques et perspectives actuelles, La Houille Blanche, 2, 5–11, 895 https://doi.org/10.1051/lhb/2012009, 2012.

Hingray, B., Picouet, C., and Musy, A.: Hydrology A Science for Engineers, CRC Press, 592 pp., ISBN 9781466590595, 2014.

Horton, P. and Brönnimann, S.: Impact of global atmospheric reanalyses on statistical precipitation downscaling, Clim. Dynam., 52, 5189–5211, https://doi.org/10.1007/s00382-018-4442-6, 2019.

Huss, M.: Present and future contribution of glacier storage change to runoff from macroscale drainage basins in Europe, Water Resour. Res., 900 47, W07 511, https://doi.org/10.1029/2010WR010299, 2011.

Isotta, F. A., Frei, C., Weilguni, V., Tadic, M. P., Lassegues, 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 Vertacnik, G.: The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, Int. J. Climatol., 34, 1657–1675, https://doi.org/10.1002/joc.3794, 2014.

905   Jenny, J.-P., Wilhelm, B., Arnaud, F., Sabatier, P., Giguet Covex, C., Mélo, A., Fanget, B., Malet, E., Ployon, E., and Perga, M. E.: A 4D sedimentological approach to reconstructing the flood frequency and intensity of the Rhône River (Lake Bourget, NW European Alps), J. Paleolimnol., 51, 469–483, https://doi.org/10.1007/s10933-014-9768-4, 2014.

Jungclaus, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., Egorova, T., Evans, M., González-Rouco, J. F., Goosse, H., Hurtt, G. C., Joos, F., Kaplan, J. O., Khodri, M., Klein Goldewijk, K., Krivova, N., LeGrande, A. N., Lorenz, S. J., Luterbacher, J., Man,

910   W., Maycock, A. C., Meinshausen, M., Moberg, A., Muscheler, R., Nehrbass-Ahles, C., Otto-Bliesner, B. I., Phipps, S. J., Pongratz, J., Rozanov, E., Schmidt, G. A., Schmidt, H., Schmutz, W., Schurer, A., Shapiro, A. I., Sigl, M., Smerdon, J. E., Solanki, S. K., Timmreck, C., Toohey, M., Usoskin, I. G., Wagner, S., Wu, C.-J., Yeo, K. L., Zanchettin, D., Zhang, Q., and Zorita, E.: The PMIP4 contribution to CMIP6 - Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 past1000 simulations, Geosci. Model Dev., 10, 4005–4033, https://doi.org/10.5194/gmd-10-4005-2017, 2017.

915   Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclaus, J. H., Otto-Bliesner, B. L., Peterschmitt, J.-Y., Abe-Ouchi, A., Albani, S., Bartlein, P. J., Brierley, C., Crucifix, M., Dolan, A., Fernandez-Donado, L., Fischer, H., Hopcroft, P. O., Ivanovic, R. F., Lambert, F., Lunt, D. J., Mahowald, N. M., Peltier, W. R., Phipps, S. J., Roche, D. M., Schmidt, G. A., Tarasov, L., Valdes, P. J., Zhang, Q., and Zhou, T.: The PMIP4 contribution to CMIP6 - Part 1: Overview and over-arching analysis plan, Geosci. Model Dev., 11, 1033–1057, https://doi.org/10.5194/gmd-11-1033-2018, 2018.

920   Kochendorfer, J., Rasmussen, R., Wolff, M., Baker, B., Hall, M. E., Meyers, T., Landolt, S., Jachcik, A., Isaksen, K., Brækkan, R., and Leeper, R.: The quantification and correction of wind-induced precipitation measurement errors, Hydrol. Earth Syst. Sci., 21, 1973–1989, https://doi.org/10.5194/hess-21-1973-2017, 2017.

Kuchment, L. and Gelfan, A.: The determination of the snowmelt rate and the meltwater outflow from a snowpack for modelling river runoff generation, J. Hydrol., 179, 23–36, https://doi.org/10.1016/0022-1694(95)02878-1, 1996.

925   Kuentz, A., Mathevet, T., Gailhard, J., and Hingray, B.: Building long-term and high spatio-temporal resolution precipitation and air temperature reanalyses by mixing local observations and global atmospheric reanalyses: the ANATEM model, Hydrol. Earth Syst. Sci., 19, 2717–2736, https://doi.org/10.5194/hess-19-2717-2015, 2015.

Kundzewicz, Z. W., Krysanova, V., Dankers, R., Hirabayashi, Y., Kanae, S., Hattermann, F. F., Huang, S., Milly, P. C. D., Stoffel, M., Driessen, P. P. J., Matczak, P., Quevauviller, P., and Schellnhuber, H.-J.: Differences in flood hazard projections in Europe - their causes

930   and consequences for decision making, Hydrolog. Sci. J., 62, 1–14, https://doi.org/10.1080/02626667.2016.1241398, 2016.

Lafaysse, M., Hingray, B., Mezghani, A., Gailhard, J., and Terray, L.: Internal variability and model uncertainty components in future hydrometeorological projections: The Alpine Durance basin, Water Resour. Res., 50, 3317–3341, https://doi.org/10.1002/2013WR014897, 2014.

Leander, R., Buishand, T. A., van den Hurk, B. J. J. M., and de Wit, M. J. M.: Estimated changes in flood quantiles of the river Meuse from

935   resampling of regional climate model output, J. Hydrol., 351, 331–343, https://doi.org/10.1016/j.jhydrol.2007.12.020, 2008.

Lemaitre-Basset, T., Collet, L., Thirel, G., Parajka, J., Evin, G., and Hingray, B.: Climate change impact and uncertainty analysis on hydrological extremes in a French Mediterranean catchment, Hydrolog. Sci. J., 66, 888–903, https://doi.org/10.1080/02626667.2021.1895437, 2021.

Leung, L. R. and Qian, Y.: Atmospheric rivers induced heavy precipitation and flooding in the western US simulated by the WRF regional

940   climate model, Geophys. Res. Lett., 36, L03 820, https://doi.org/10.1029/2008GL036445, 2009.

Lorenz, E. N.: Atmospheric Predictability as Revealed by Naturally Occurring Analogues, J. Atmos. Sci., 26, 636–646, https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2, 1969.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Curr. Clim. Change Rep., 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

945   Maraun, D., Wetterhall, F., Ireson, A. M., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, RG3003, https://doi.org/10.1029/2009RG000314, 2010.

Marty, R., Zin, I., Obled, C., Bontron, G., and Djerboua, A.: Toward Real-Time Daily PQPF by an Analog Sorting Approach: Application to

950   Flash-Flood Catchments, J. Appl. Meteorol. Clim., 51, 505–520, https://doi.org/10.1175/JAMC-D-11-011.1, 2012.

Masson, D. and Frei, C.: Long-term variations and trends of mesoscale precipitation in the Alps: recalculation and update for 1901-2008, Int. J. Climatol., 36, 492–500, https://doi.org/10.1002/joc.4343, 2016.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T. K., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B.: Cli-

955   mate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2391 pp., https://doi.org/10.1017/9781009157896, 2021.

Metzger, A.: Retour d'est - La géochronique du temps qu'il fait, La Géographie, 1588, 64–65, https://doi.org/10.3917/geo.1588.0064, 2023.

Mezghani, A. and Hingray, B.: A combined downscaling-disaggregation weather generator for stochastic generation of multisite hourly

960   weather variables over complex terrain: Development and multi-scale validation for the Upper Rhone River basin, J. Hydrol., 377, 245–260, https://doi.org/10.1016/j.jhydrol.2009.08.033, 2009.

Minville, M., Cartier, D., Guay, C., Leclaire, L., Audet, C., Le Digabel, S., and Merleau, J.: Improving process representation in conceptual hydrological model calibration using climate simulations, Water Resour. Res., 50, 5044–5073, https://doi.org/10.1002/2013WR013857, 2014.

965   Ménégoz, M., Gallée, H., and Jacobi, H. W.: Precipitation and snow cover in the Himalaya: from reanalysis to regional climate simulations, Hydrol. Earth Syst. Sci., 17, 3921–3936, https://doi.org/10.5194/hess-17-3921-2013, 2013.

Ménégoz, M., Valla, E., Jourdain, N. C., Blanchet, J., Beaumet, J., Wilhelm, B., Gallée, H., Fettweis, X., Morin, S., and Anquetin, S.: Contrasting seasonal changes in total and intense precipitation in the European Alps from 1903 to 2010, Hydrol. Earth Syst. Sci., 24, 5355–5377, https://doi.org/10.5194/hess-24-5355-2020, 2020.

970   Nabat, P., Somot, S., Mallet, M., Sanchez-Lorenzo, A., and Wild, M.: Contribution of anthropogenic sulfate aerosols to the changing Euro-Mediterranean climate since 1980, Geophys. Res. Lett., 41, 5605–5611, https://doi.org/10.1002/2014GL060798, 2014.

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

Obled, C., Bontron, G., and Garçon, R.: Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues

975   sorting approach, Atmos. Res., 63, 303–324, https://doi.org/10.1016/S0169-8095(02)00038-8, 2002.

Obled, C., Zin, I., and Hingray, B.: Choix des pas de temps et d'espace pour des modélisations parcimonieuses en hydrologie des crues, La Houille Blanche, 5, 81–87, https://doi.org/10.1051/lhb/2009059, 2009.

OFEG: Les crues 2000 - Analyse des événements/cas exemplaires, Rapports de l'OFEG, Série Eaux, 2, Office Fédéral des Eaux et de la Géologie, Berne, Suisse, 2002.

980   Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J.-N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An Atmospheric Reanalysis of the Twentieth Century, J. Climate, 29, 4083–4097, https://doi.org/10.1175/JCLI-D-15-0556.1, 2016.

Quintana Seguí, P., Ribes, A., Martin, E., Habets, F., and Boé, J.: Comparison of three downscaling methods in simulating the impact of climate change on the hydrology of Mediterranean basins, J. Hydrol., 383, 111–124, https://doi.org/10.1016/j.jhydrol.2009.09.050, 2010.

985   Raymond, F., Wilhelm, B., and Anquetin, S.: Is Precipitation the Main Trigger of Medium-Magnitude Floods in Large Alpine Catchments?, Water, 11, 2507, https://doi.org/10.3390/w11122507, 2019.

Raynaud, D., Hingray, B., Evin, G., Favre, A.-C., and Chardon, J.: Assessment of meteorological extremes using a synoptic weather generator and a downscaling model based on analogues, Hydrol. Earth Syst. Sci., 24, 4339–4352, https://doi.org/10.5194/hess-24-4339-2020, 2020.

Reid, P. C., Hari, R. E., Beaugrand, G., Livingstone, D. M., Marty, C., Straile, D., Barichivich, J., Goberville, E., Adrian, R., Aono, Y.,
990   Brown, R., Foster, J., Groisman, P., Hélaouët, P., Hsu, H., Kirby, R., Knight, J., Kraberg, A., Li, J., Lo, T., Myneni, R. B., North, R. P., Pounds, J. A., Sparks, T., Stübi, R., Tian, Y., Wiltshire, K. H., Xiao, D., and Zhu, Z.: Global impacts of the 1980s regime shift, Global Change Biol., 22, 682–703, https://doi.org/10.1111/gcb.13106, 2016.

Rohwer, C.: Evaporation from free water surfaces, Technical Bulletin, 271, United States, Department of Agriculture, https://doi.org/DOI: 10.22004/ag.econ.163103, 1931.

995   Roudier, P., Andersson, J. C. M., Donnelly, C., Feyen, L., Greuell, W., and Ludwig, F.: Projections of future floods and hydrological droughts in Europe under a +2°C global warming, Clim. Change, 135, 341–355, https://doi.org/10.1007/s10584-015-1570-4, 2016.

Ruelland, D.: Should altitudinal gradients of temperature and precipitation inputs be inferred from key parameters in snow-hydrological models?, Hydrol. Earth Syst. Sci., 24, 2609–2632, https://doi.org/10.5194/hess-24-2609-2020, 2020.

Ruosteenoja, K., Markkanen, T., Venäläinen, A., Räisänen, P., and Peltola, H.: Seasonal soil moisture and drought occurrence in Europe in
1000   CMIP5 projections for the 21st century, Clim. Dynam., 50, 1177–1192, https://doi.org/10.1007/s00382-017-3671-4, 2018.

Schaefli, B., Hingray, B., Niggli, M., and Musy, A.: A conceptual glacio-hydrological model for high mountainous catchments, Hydrol. Earth Syst. Sci., 9, 95–109, https://doi.org/10.5194/hess-9-95-2005, 2005.

Sivapalan, M., Blöschl, G., Zhang, L., and Vertessy, R.: Downward approach to hydrological prediction, Hydrol. Process., 17, 2101–2111, https://doi.org/10.1002/hyp.1425, 2003.

1005   Switanek, M., Maraun, D., and Bevacqua, E.: Stochastic downscaling of gridded precipitation to spatially coherent subgrid precipitation fields using a transformed Gaussian model, Int. J. Climatol., 42, 6126–6147, https://doi.org/10.1002/joc.7581, 2022.

Tapiador, F. J., Navarro, A., Moreno, R., Sánchez, J. L., and García-Ortega, E.: Regional climate models: 30 years of dynamical downscaling, Atmos. Res., 235, 104 785, https://doi.org/10.1016/j.atmosres.2019.104785, 2020.

Teweles, S. and Wobus, H. B.: Verification of Prognostic Charts, B. Am. Meteorol. Soc., 35, 455–463, https://doi.org/10.1175/1520-0477-
1010   35.10.455, 1954.

Tolson, B. A. and Shoemaker, C. A.: Dynamically dimensioned search algorithm for computationally efficient watershed model calibration, Water Resour. Res., 43, W01 413, https://doi.org/10.1029/2005WR004723, 2007.

Vidal, J.-P., Hingray, B., Magand, C., Sauquet, E., and Ducharne, A.: Hierarchy of climate and hydrological uncertainties in transient low-flow projections, Hydrol. Earth Syst. Sci., 20, 3651–3672, https://doi.org/10.5194/hess-20-3651-2016, 2016.

1015   Viviroli, D., Mittelbach, H., Gurtz, J., and Weingartner, R.: Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland - Part II: Parameter regionalisation and flood estimation results, J. Hydrol., 377, 208–225, https://doi.org/10.1016/j.jhydrol.2009.08.022, 2009.

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

Von Storch, H., Zorita, E., and Cubasch, U.: Downscaling of Global Climate-Change Estimates to Regional Scales - an Application to Iberian Rainfall in Wintertime, J. Climate, 6, 1161–1171, https://doi.org/10.1175/1520-0442(1993)006<1161:DOGCCE>2.0.CO;2, 1993.

Wagner, P. D., Fiener, P., Wilken, F., Kumar, S., and Schneider, K.: Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions, J. Hydrol., 464-465, 388–400, https://doi.org/10.1016/j.jhydrol.2012.07.026, 2012.

Weber, M., Koch, F., Bernhardt, M., and Schulz, K.: The evaluation of the potential of global data products for snow hydrological modelling in ungauged high-alpine catchments, Hydrol. Earth Syst. Sci., 25, 2869–2894, https://doi.org/10.5194/hess-25-2869-2021, 2021.

Wilby, R., Hay, L., and Leavesley, G.: A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan River basin, Colorado, J. Hydrol., 225, 67–91, https://doi.org/10.1016/S0022-1694(99)00136-5, 1999.

Wilby, R. L. and Quinn, N. W.: Reconstructing multi-decadal variations in fluvial flood risk using atmospheric circulation patterns, J. Hydrol., 487, 109–121, https://doi.org/10.1016/j.jhydrol.2013.02.038, 2013.

Wilhelm, B., Rapuc, W., Amann, B., Anselmetti, F. S., Arnaud, F., Blanchet, J., Brauer, A., Czymzik, M., Giguet-Covex, C., Gilli, A., Glur, L., Grosjean, M., Irmler, R., Nicolle, M., Sabatier, P., Swierczynski, T., and Wirth, S. B.: Impact of warmer climate periods on flood hazard in the European Alps, Nat. Geosci., 15, 118–123, https://doi.org/10.1038/s41561-021-00878-y, 2022.

Winsemius, H. C., Schaefli, B., Montanari, A., and Savenije, H. H. G.: On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information, Water Resour. Res., 45, W12 422, https://doi.org/10.1029/2009WR007706, 2009.

Wood, A. W., Leung, L. R., Sridhar, V., and Lettenmaier, D. P.: Hydrologic Implications of Dynamical and Statistical Approaches to Downscaling Climate Model Outputs, Clim. Change, 62, 189–216, https://doi.org/10.1023/B:CLIM.0000013685.99609.9e, 2004.

Wyard, C., Scholzen, C., Fettweis, X., Van Campenhout, J., and François, L.: Decrease in climatic conditions favouring floods in the southeast of Belgium over 1959-2010 using the regional climate model MAR, Int. J. Climatol., 37, 2782–2796, https://doi.org/10.1002/joc.4879, 2017.
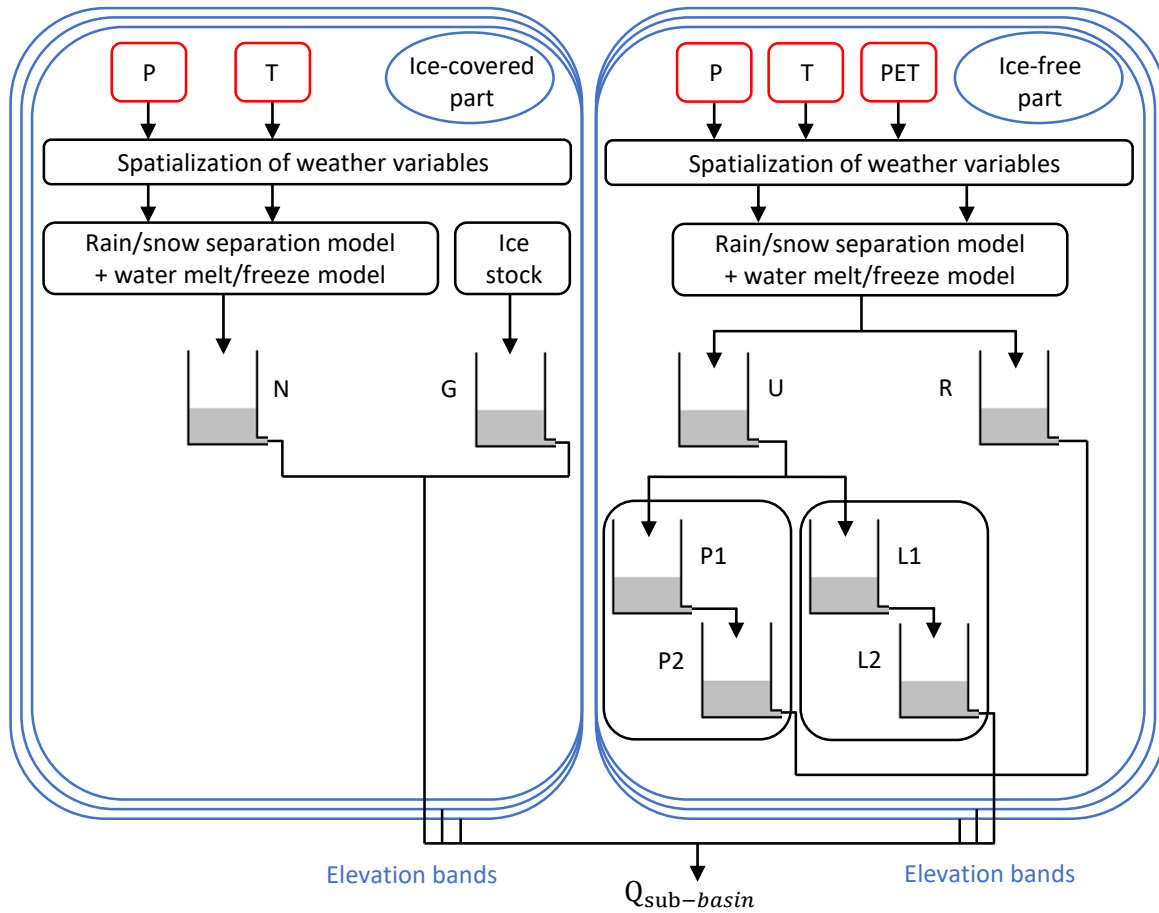
Figure 15. Scheme of the glacio-hydrological model GSM-SOCONT. The model uses daily data on precipitation (P), mean temperature (T) and potential evapotranspiration (PET) as inputs. A snow reservoir N represents the stock of snow that evolves through the processes of snow accumulation, melting and refreezing. The ice reservoir G models ice melt. Infiltration and evapotranspiration losses are simulated with a superficial reservoir U. River flow from runoff (by infiltration capacity excess) is simulated with a non-linear surface reservoir R. A Nash-like cascade of two linear reservoirs (L1 and L2) is used to simulate sub-superficial storage and release. A Nash-like cascade of two linear reservoirs (P1 and P2) is used to simulate deep groundwater storage and release.

14S1

Table S2. Calibration characteristics for each sub-basin and values of the components of the objective functions used for the automatic calibration of the hydrological model. NSE_{chrono} and NSE_{regime} correspond to the Nash-Sutcliffe Efficiency criterion applied to the discharges time series and the interannual daily regimes respectively. d_{KS} is the Kolmogorov-Smirnov distance, i.e. the maximum difference between the simulated and observed distributions of the annual discharge maxima. Fd_{KS} refers to the second part of the objective function $F_{altered}$ in Eq. (2). Rhône@Genève, HDI: Rhône@Genève, Halle-de-l'Ile. Rhône@Genève, BDM: Rhône@Genève, Bout-du-Monde.

Sub-basins	Gauging station	Calibration type	Calibration period(s)	NSE_{chrono}	NSE_{regime}	d_{KS}	Fd_{KS}
1	Rhône@Brigue	Chronological	1965-2015 for obs/sim	0.82	-	-	-
2	Rhône@Viège	Climatological	1922-1963 for obs 1961-2015 for sim	-	0.99	48	0.17
3, 4	Rhône@Sion	Climatological	1916-1956 for obs 1961-2015 for sim	-	0.98	163	0.19
5	Rhône@Branson	Climatological	1941-1956 for obs 1961-2015 for sim	-	0.33	45	0.32
6, 7	Rhône@Porte-du-Scex	Climatological	1941-1956 for obs 1961-2015 for sim	-	0.93	35	0.12
8, 9, 10, 11, 12	Rhône@Genève, HDI	Chronological	1961-1970 for obs/sim	0.72	-	-	-
13	Arve@Sallanches	Chronological	1965-2015 for obs/sim	0.67	-	-	-
14	Arve@Taninges	Chronological	1961-2015 for obs/sim	0.65	-	-	-
15, 16	Arve@Genève, BDM	Chronological	1965-2015 for obs/sim	0.44	-	-	-
17, 18	Rhône@Bognes	Climatological	1923-1947 for obs 1961-2015 for sim	-	0.33	62	0.10

1045 Example of calibration of a sub-basin with an altered hydrological regime (Viège sub-basin). Observation period: 1922–1963. Simulation period: 1961-2015. (**a**) Interannual daily regime. (**b**) Gumbel plot for annual discharge maxima. The x-axis is the reduced variate for the given return period , i.e. $u = -\ln(-\ln(1 - 1/T))$. Dashed lines correspond to 90 % confidence bounds of the Gumbel distribution estimated on observed data.
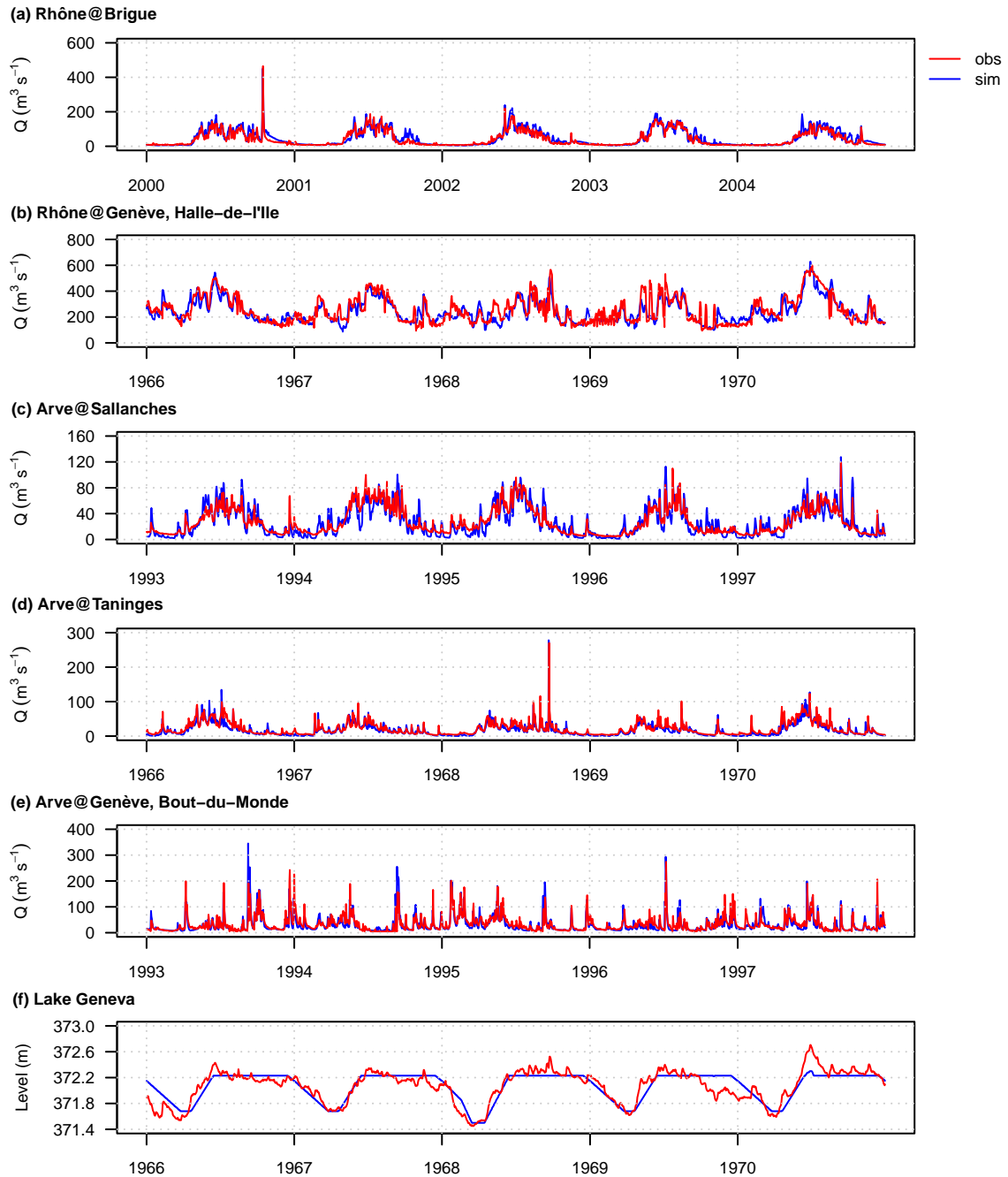
Figure S3. ~~Scatter plots~~ Classical calibration of ~~mean monthly discharges, low flows and flood events at~~ sub-basins with a natural (~~a~~or at least not significantly altered) hydrological regime. **(a)** Rhône@~~Porte-du-Scex~~ Brigue (~~1905-1956~~1965-2015)~~, (b)~~ . **(b)** Rhône@Genève, Halle-de-l'Ile (~~1923-1956~~1961-1970)~~.~~ . **(c)** Arve@Sallanches (~~c~~1965-2015). **(d)** Arve@Taninges (1961-2015). **(e)** Arve@~~Arve~~Genève, Bout-du-Monde (~~1904-1960~~1965-2015)~~,~~ . **(f)** Example of Lake Geneva levels observed and ~~(d)~~ simulated over the same period as discharges at Rhône@~~Bognes (1920-1947)~~Genève, Halle-de-l'Ile. The calibration criterion is the NSE coefficient calculated from the observed and simulated time series of discharge at the sub-basin outlet.
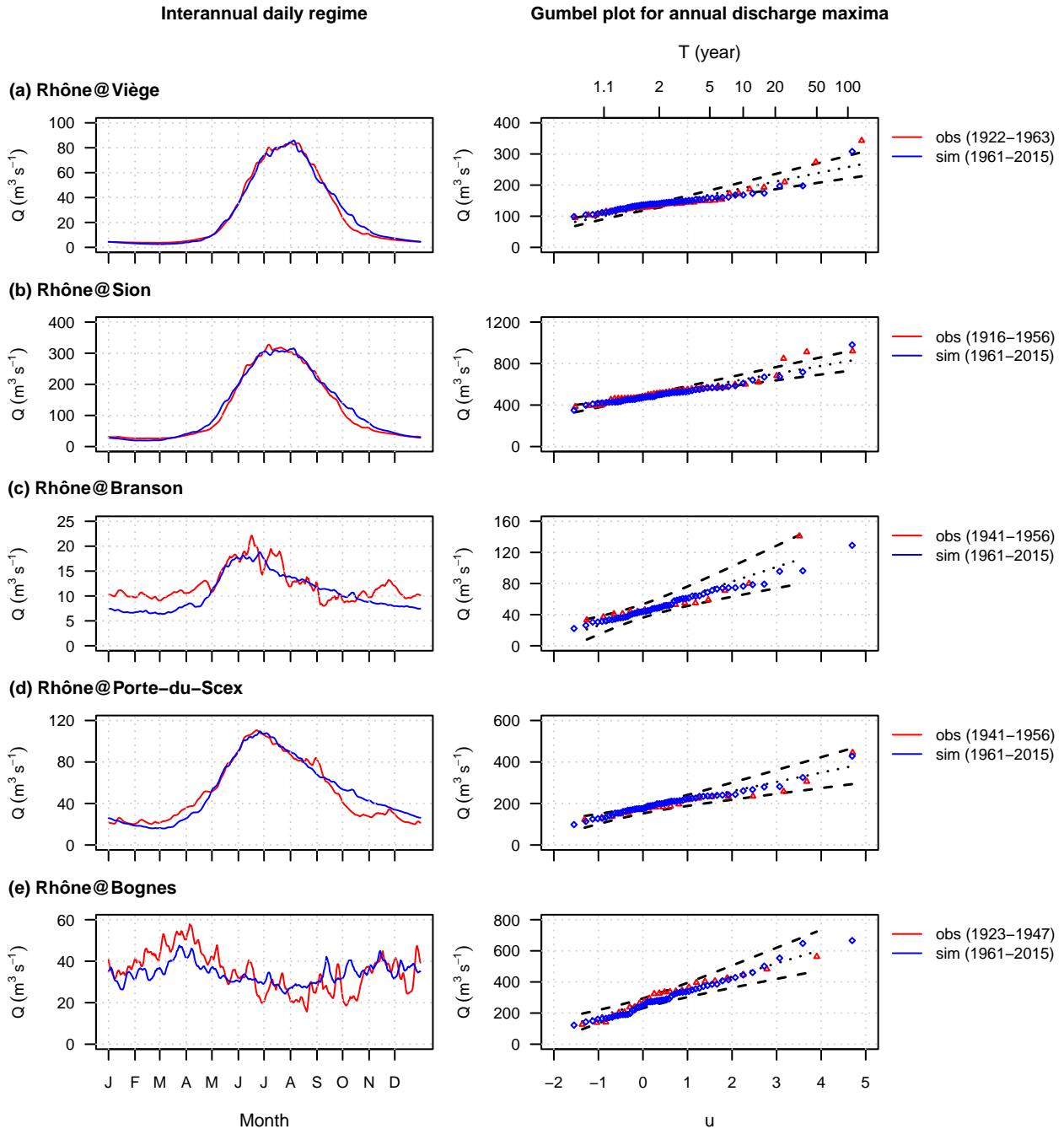
Figure S4. ~~Scatter plots~~ Signatures-based calibration of ~~flood events by season at (a)~~ sub-basins with an altered hydrological regime. **(a)** Rhône@~~Porte-du-Scex, (b)~~ Viège. **(b)** Rhône@~~Genève, Halle-de-l'Ile, (c) Arve~~Sion. **(c)** Rhône@~~Genève, Bout-du-Monde, and (d)~~ Branson. **(d)** Rhône@~~Porte-du-Scex. (e) Rhône~~Bognes. **(left)** Interannual daily regime. **(right)** Gumbel plot for annual discharge maxima. The x-axis is the ~~1961-2009~~ reduced variate u for the given return period T, i.e. $u = -\ln(-\ln(1 - 1/T))$. Dashed lines correspond to 90 % confidence bounds of the Gumbel distribution estimated on observed data. The calibration criteria are the NSE coefficient between simulated and observed regimes and the Kolmogorov-Smirnov coefficient between the observed and the simulated distributions of annual discharge maxima.
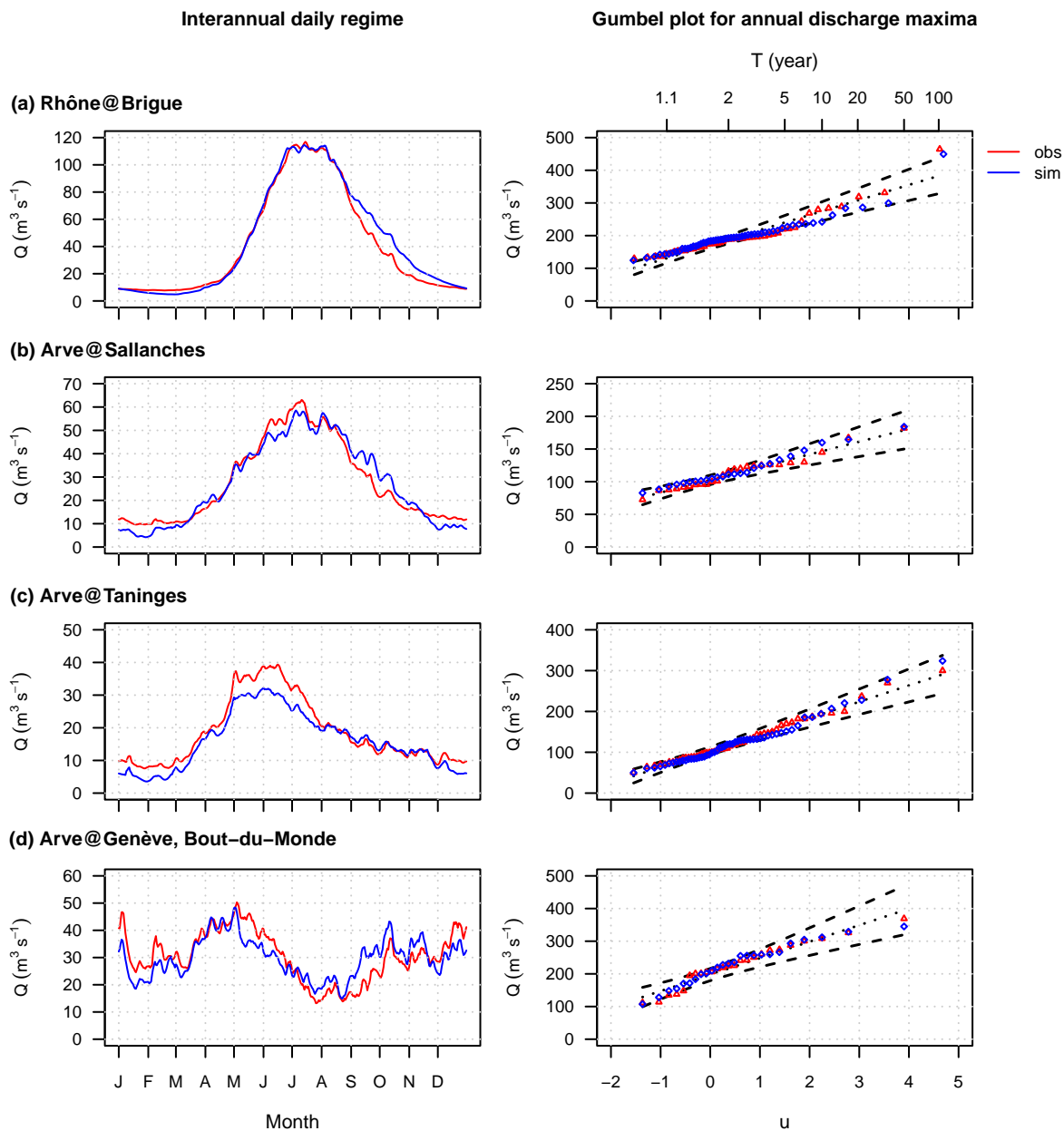
Figure S5. Signatures-based calibration of sub-basins with a natural (or at least not significantly altered) hydrological regime. **(a)** Rhône@Brigue (1965-2015). **(b)** Arve@Sallanches (1965-2015). **(c)** Arve@Taninges (1961-2015). **(d)** Arve@Genève, Bout-du-Monde (1965-2015). **(left)** Interannual daily regime. (**right**) Gumbel plot for annual discharge maxima. The x-axis is the reduced variate u for the given return period T, i.e. $u = -\ln(-\ln(1 - 1/T))$. Dashed lines correspond to 90 % confidence bounds of the Gumbel distribution estimated on observed data. The calibration criteria are the NSE coefficient between simulated and observed regimes and the Kolmogorov-Smirnov coefficient between the observed and the simulated distributions of annual discharge maxima.
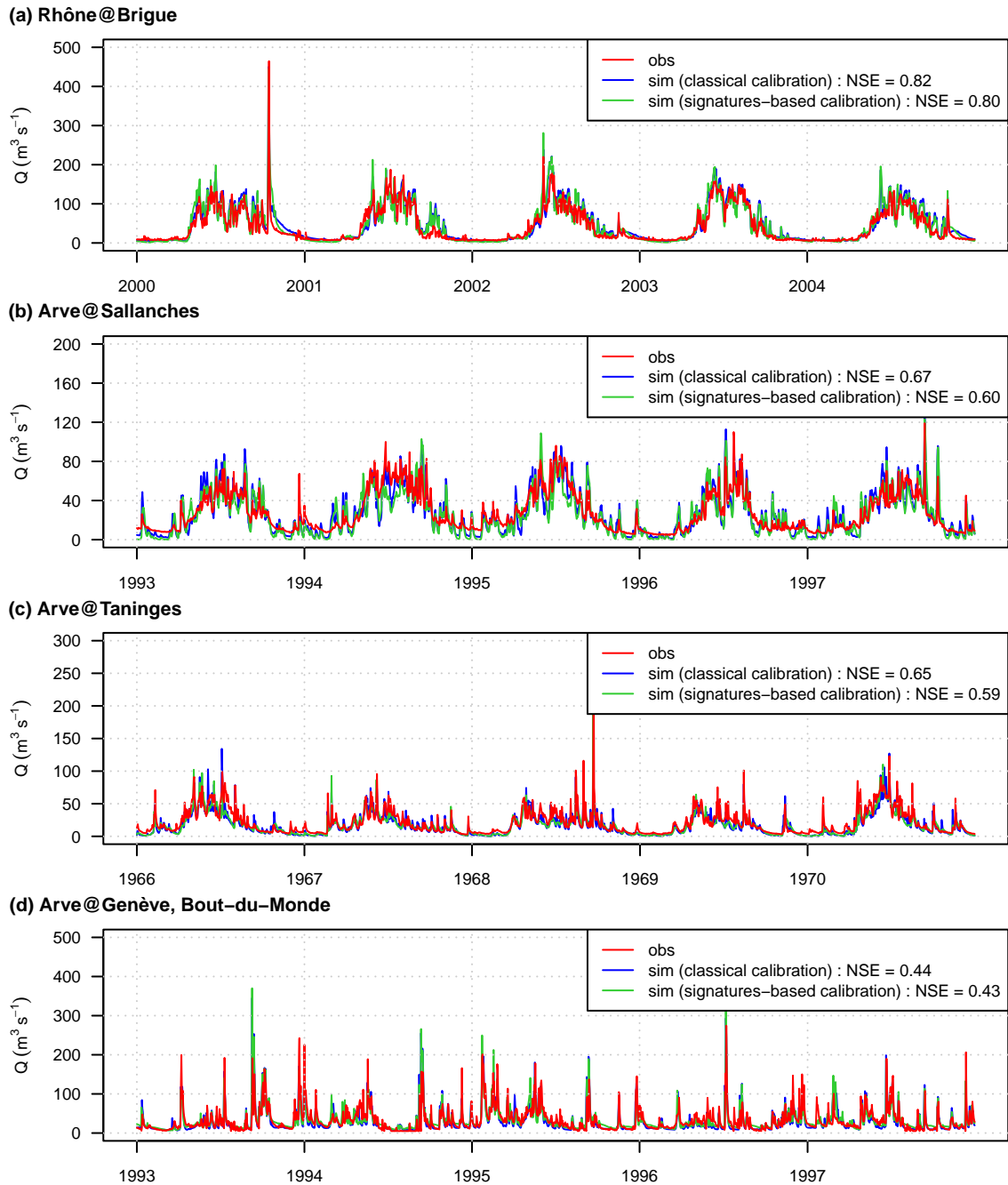
Figure S6. Comparison of the two types of calibration for sub-basins with a natural (or at least not significantly altered) hydrological regime. The NSE coefficients were calculated using discharges data from the same period for both types of calibration: classical calibration and signatures-based calibration. **(a)** Rhône@Brigue (1965-2015). **(b)** Arve@Sallanches (1965-2015). **(c)** Arve@Taninges (1961-2015). **(d)** Arve@Genève, Bout-du-Monde (1965-2015).
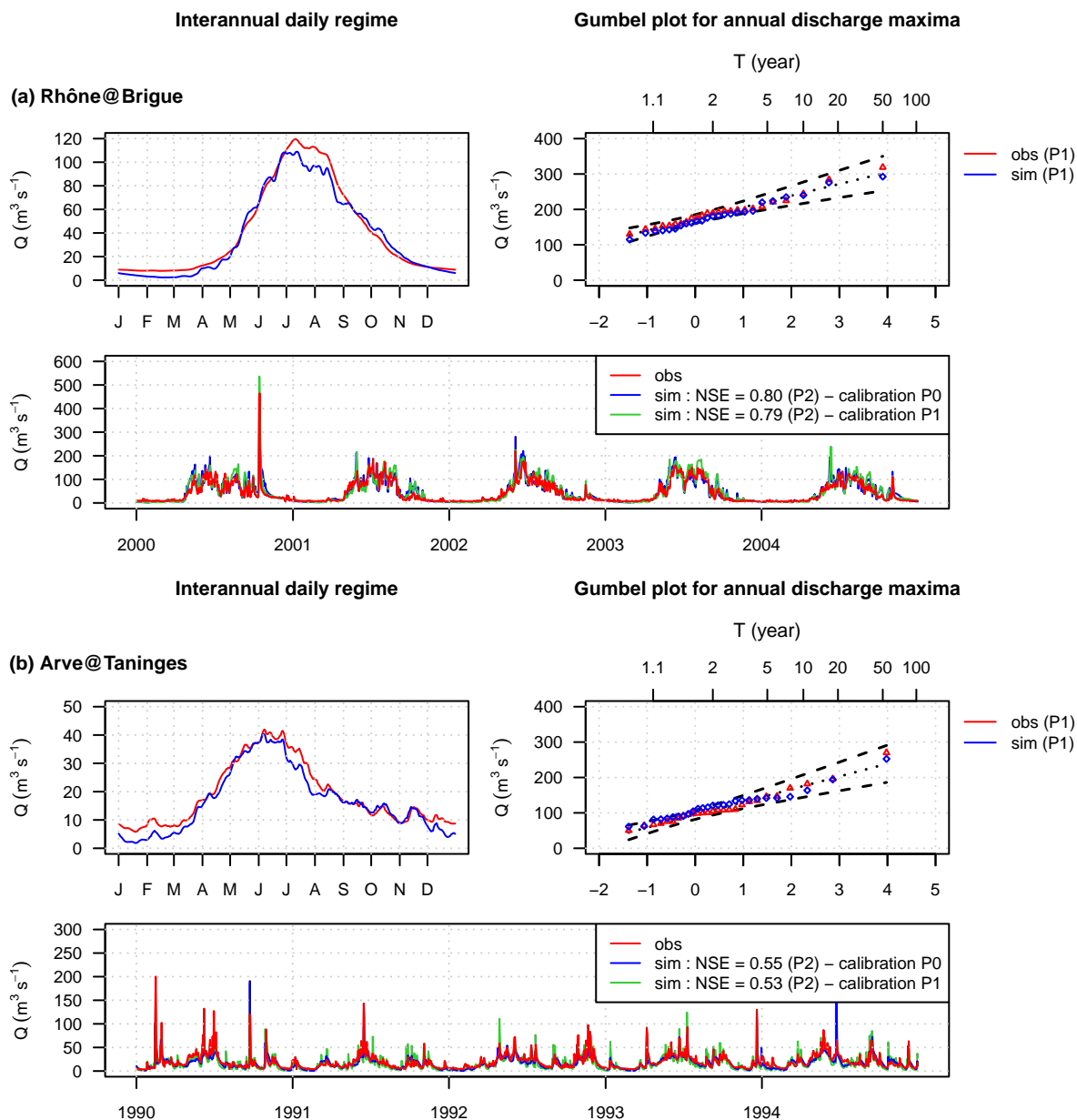
Figure S7. Calibration-validation procedure applied to two sub-basins with a natural (or at least not significantly altered) hydrological regime. The signature-based calibration was applied to the signatures of period P1 with weather data from P1. The signatures are the interannual daily regime and the statistical distribution of the annual discharge maxima. The discharge time series obtained with this signature-based calibration is shown along with those observed and those obtained from calibration based on the signatures of the entire period P0. The NSE coefficients were calculated from the observed and simulated time series of discharge over the period P2 from the weather data of P2. **(a)** Rhône@Brigue: P0 = 1965-2015, P1 = 1965-1989, P2 = 1991-2015. **(b)** Arve@Taninges: P0 = 1961-2015, P1 = 1961-1987, P2 = 1989-2015.
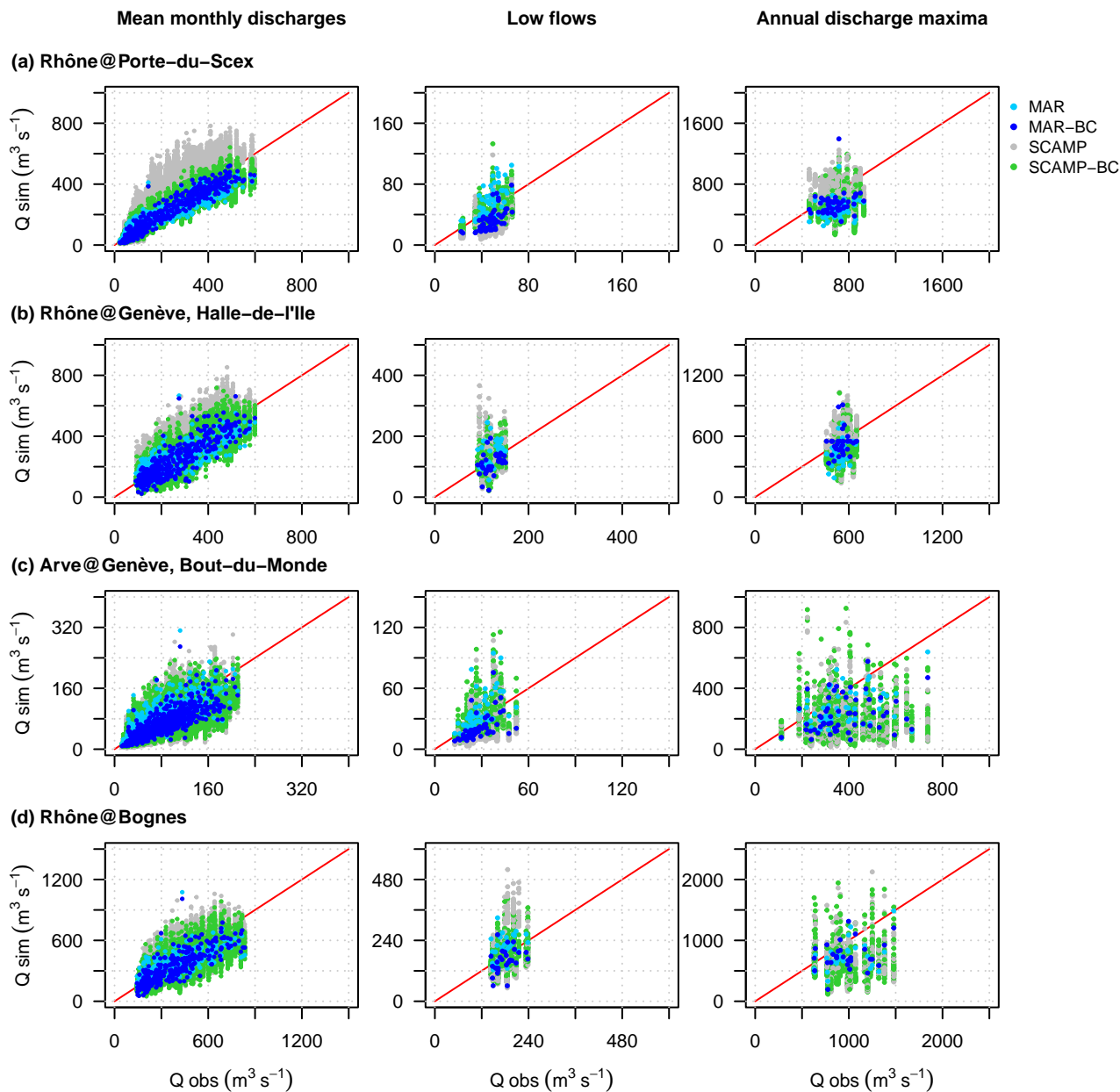
Figure S8. Scatter plots of mean monthly discharges, low flows and annual discharge maxima at (**a**) Rhône@Porte-du-Scex (1905-1956), (**b**) Rhône@Genève, Halle-de-l'Ile (1923-1956), (**c**) Arve@Genève, Bout-du-Monde (1904-1960), and (**d**) Rhône@Bognes (1920-1947). Q obs: observed discharge. MAR/MAR-BC: simulated discharge from the raw/bias-corrected weather scenario produced with the dynamical downscaling model MAR. SCAMP/SCAMP-BC: simulated discharge from the raw/bias-corrected weather scenarios produced with the statistical downscaling model SCAMP.
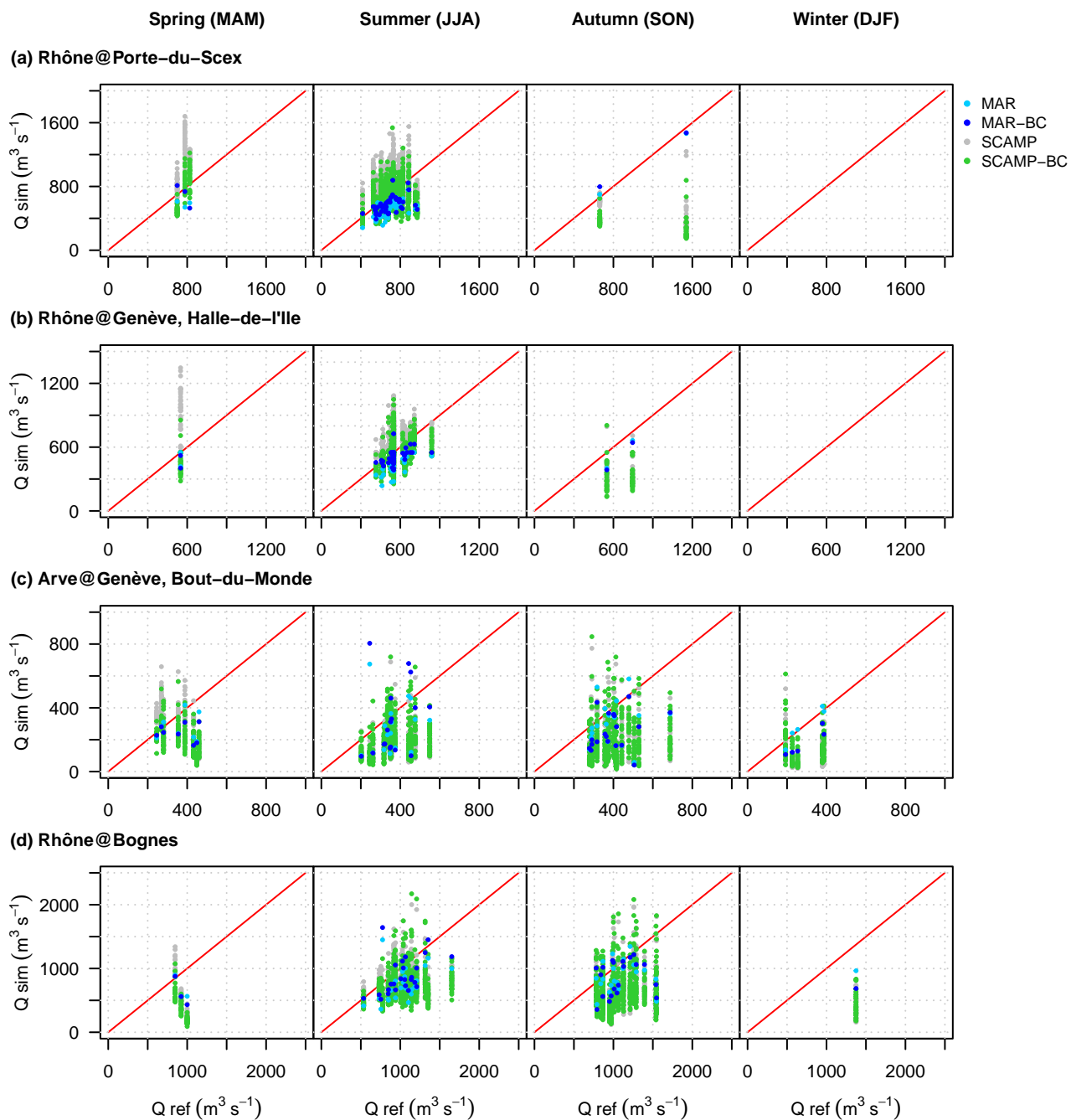
Figure S9. Scatter plots of annual discharge maxima by season at (**a**) Rhône@Porte-du-Scex, (**b**) Rhône@Genève, Halle-de-l'Ile, (**c**) Arve@Genève, Bout-du-Monde, and (**d**) Rhône@Bognes for the 1961-2009 period. MAM: March-April-May. JJA: June-July-August. SON: September-October-November. DJF: December-January-February. Q ref: simulated discharge from observed weather variables. MAR/MAR-BC: simulated discharge from the raw/bias-corrected weather scenario produced with the dynamical downscaling model MAR. SCAMP/SCAMP-BC: simulated discharge from the raw/bias-corrected weather scenarios produced with the statistical downscaling model SCAMP.
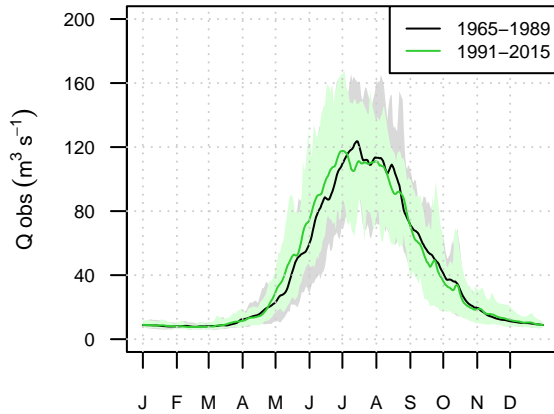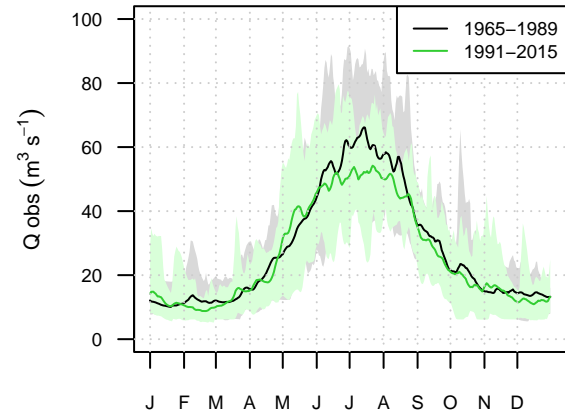
Figure S10. Assessment of the stationarity assumption of the natural hydrological behavior of the Upper Rhône River catchment over the last century. For sub-basins with a natural (or at least not significantly altered) hydrological regime, we split the period of available observed discharge data into two periods. The mean interannual daily regimes are indicated by the black and green solid lines. The grey and green bands represent the confidence intervals at 90 % level. **(a)** Rhône@Brigue. **(b)** Arve@Sallanches. **(c)** Arve@Taninges. **(d)** Arve@Genève, Bout-du-Monde.