**Response to Anonymous Referee #1**

We thank Referee #1 for reading the manuscript carefully and providing thoughtful and constructive comments. The comments are noted with RC, our responses with AC, and the intended additions or changes in the manuscript are underlined.

General comments

**RC1.1** This study evaluates two approaches for simulating over one century-long daily streamflow at different Upper Rhône River catchment locations. The analysis examines the potential of atmospheric reanalyses simulations (ERA-20C), conceptual glacier-hydrological model and two downscaling approaches for providing temporal variations of river discharges, low flow sequences and flood events. The results indicate that bias correction is crucial for both precipitation and air temperature and both downscaling models. While the observed multi-scale variations of discharges (daily, seasonal and inter-annual) are well reproduced, the results for low flows are less satisfactory.

The manuscript is within the scope of the journal and has a good structure. Still, the novel scientific contribution can be better formulated. The Introduction says that the comparison and evaluation of different downscaling approaches have been the subject of multiple previous studies, so it is not clear what are still the research gaps and how this manuscript contributes to some novel findings/knowledge. The manuscript will also benefit from some more detailed justification of why such one-century runs are needed and/or beneficial.

> **AC1.1** Thank you for this comment. This point was also highlighted by Referee #3 (RC3.1) and Referee #4 (RC4.1).
>
> As mentioned in the introduction, a large number of previous studies described downscaling approaches to adapt climate model outputs used to force hydrological models. Several studies highlighted that the choice of the downscaling method can strongly modulate the hydrological regime simulated with the same model (Wood et al., 2004; Quintana Seguí et al., 2010; Chen et al., 2011).
>
> Realistic hydrological simulations based on a specific downscaling approach and fine model calibration at the basin-scale are often used for relatively short hydrological simulations, lasting from one year to one decade (e.g. Habets et al., 1999; Boscarello et al., 2014), while longer periods are required to study trends in precipitation (Ménégoz et al., 2020) and river flows (Brönnimann et al., 2022).
>
> Bonnet et al. (2017) simulated the water cycle at the scale of the entire French hydrological system over the 20th century, a period long enough to discuss decadal variability and long-term trends. However, their model shows local deficiencies, particularly in mountainous areas where the snow cover plays a crucial role, and where the dams - not considered in their study - have significantly affected the river flows since the 1950s.
>
> The novelty of our study relies on three combined features: (i) we used a fine model calibration of the hydrological model to accurately simulate snowmelt, including the impacts of the dams at the scale of the Upper Rhône River catchment; (ii) we applied this model over a long enough period - the entire 20th century - to allow a complete evaluation of the model to simulate the

hydrological regime in terms of mean, variability, long-term trend and extremes; (iii) we compared two configurations, one based on statistical downscaling and the other one using dynamical downscaling, highlighting the need to accurately simulate the temperature and precipitation lapse rates. Overall, we confirm the need for an additional bias correction of atmospheric variables after the downscaling step and before the application of the hydrological model.

The novelty of our study will be deeply described in the revised version of the article.

A discussion of the value of the proposed approach with some alternative approaches (e.g. stochastic weather generator) will be useful as well.

Thank you for this comment. Yes, we could use a stochastic weather generator (WGEN). We have currently developed a multi-site WGEN for the simulation of long time series (> 10,000 years) of weather scenarios for the large Aare catchment in Switzerland (GWEX). The entire simulation framework is described in the article by Viviroli et al, 2022 published last year in HESS.

However, WGENs are developed to generate plausible weather sequences. They do not aim to reproduce the spatio-temporal variations in weather observed in a given area. The main reason for this is that they are not forced by variations in large-scale atmospheric information. Unlike WGENs, downscaling models are. This is an important advantage of downscaling models when scenarios have to be generated for other climate contexts, where changes in the circulation and state of the atmosphere are likely to lead to changes in local weather conditions.

We agree that this point could be clarified in the manuscript, although it is not a major issue here. If lack of space is not an issue, we will clarify this point.

The proposed modelling chains are quite complex, so a rigorous description of the methods is challenging, resulting in difficulty in reproducing the proposed experiments exactly.

We recognize that the modeling chains are quite complex. These chains are quite common when weather conditions are produced from downscaling models for climate change impact studies. We tried to give as much detail as possible, and where this was not possible, we provided references to articles in which the methods are described in more detail. We hope the reader will feel comfortable with the actual details of the description.

For ease of understanding, we will include in the revised version of the article, at the beginning of the section devoted to the methods, a synthetic summary of the model chains.

**RC1.2** My second suggestion is to add more process-related interpretations of results and findings. For example, to discuss and present which processes lead to low flows and floods in the study area and to demonstrate that the suggested chains represent well such runoff generation processes in the (sub) catchments. For example, I'm not sure whether individual correction of the bias for air temperature and precipitation does not result in some artificial combination which can affect, e.g. low flow simulations of the model. If the low flows occur in winter, the impact of air temperature bias correction will be more important than precipitation corrections or vice versa. So it will be interesting to provide more process interpretations in the results.

**AC1.2** Thank you for this very interesting suggestion. Yes, you are right, for the upstream part of the Upper Rhône River catchment, low flows occur mainly in winter, under the effect of low temperatures. So, for precipitation, the lapse rate issue has little impact on low flows (see Fig. 13c). For temperatures, it conversely has a significant impact (see Fig. 12d). For low flows, it is therefore unlikely that the bias correction for precipitation and temperature results in an artificial combination that affects low flow simulations. This is clear for high elevations sub-basins (see Fig. 12d and 13c for Rhône@Porte-du-Scex). It is also observed in the lower part of the Upper Rhône River catchment, but this is less evident.

With regard to its influence on floods, a similar discussion can be proposed for the precipitation lapse rate. Increased precipitation leads to increased snowfall in winter, followed by increased snowmelt and runoff in spring. Snowmelt floods are more or less proportional. The impact on autumn discharges is different. Larger precipitation directly leads to larger floods during this season. The increase in flood intensity is expected to be much larger than the increase in precipitation amount, due to the non-linear runoff generation process that increases with precipitation intensity.

For floods, the influence of the temperature lapse rate is much less evident. In this region, the largest floods often occur in autumn, due to the large amounts of precipitation. In autumn, however, the temperatures can be low enough for precipitation in high elevation areas to fall as snow (as this was the case with the 2000 flood in the Wallis canton, see Hingray et al., 2010), leading to "reduced floods" compared to floods that would have occurred with the full amount of precipitation as rain. A wrong temperature lapse rate in the model will definitively determine the elevation of the 1°C isotherm (elevation of the snowfall/rainfall limit) and therefore the "effective catchment area" for these events. In MAR, the temperature lapse rate is overestimated, resulting in too cold weather in high elevations areas. When MAR is corrected for mean temperature only (and not for the lapse), this leads to a reduction in the intensity of autumn flood (Fig. 12d).

<u>We will add a comment on these points in the text.</u>

**RC1.3** My third comment is on the validation of the results. It will be interesting to see how the procedures work in some independent (validation) time periods.

**AC1.3** Thank you for this important comment. It would obviously be important to assess it.

For basins for which we have concomitant time series of weather and discharge observations, not perturbed by dams, a classical split-sample test would be easy to perform. This has been done in many other works in this mountainous context. This work has been carried out for the GSM-SOCONT model by Schaefli et al. (2005) for 3 sub-catchments of the Upper Rhône River basin. This shows that the calibration is solid and robust.

For all other basins, this split-sample test is not possible. This is because all data of all the period have to be considered for the signature-based calibration. We are convinced that the signature-based calibration is also reasonable. For catchments for which a classical calibration was possible, we split the period in two. We applied the signature-based calibration on the discharge signatures of period P1 with weather data from P1. We then simulated the discharge time series of period P2. The simulated time series remain in good agreement with the observed flows. The NSE coefficients are logically lower, but the difference is quite small.

We will add a few comments on this.

I wonder what is the relative contribution of each individual member of the proposed chain on the final results.

The main objective of our article is to assess the ability of two downscaling models, SCAMP and MAR, to produce hydrologically relevant precipitation and temperature scenarios from large-scale atmospheric information only. For this evaluation, we compare hydrological scenarios simulated from weather scenarios with reference hydrological scenarios also obtained by simulation, but from observed weather.

In other words, to evaluate weather scenarios in relation to observed weather, we use a filter, the hydrological filter of the considered Upper Rhône River basin. The interest of this filter is that it takes the hydrological point of view, which is a very demanding point of view due to the highly non-linear nature of many hydrological processes.

We therefore do not seek to estimate the relative contribution of each individual member of the simulation chain to the final results.

Specific comments

**RC1.4** Abstract: Please make the description of the results consistent with the conclusions. In the abstracts is written: "… the hydrological situations of low frequency (low flow sequences and flood events) are reasonably well reproduced." However in the conclusions is written: "The results for low flow activity are less satisfactory."

**AC1.4** Thank you for this comment. We will modify the abstract accordingly.

**RC1.5** Figure captions. The abbreviations used in the figures are not always explained in the captions, so it is sometimes difficult to understand them (without a detailed reading of the text).

**AC1.5** The following abbreviations will be explained in the captions: URR, P, T, MAP, MAT, Q, obs, ref, MAM, JJA, SON, DJF, BC, MAR, MAR-BC, SCAMP, SCAMP-BC, GSM-SOCONT.

**References**

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.

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.

Boscarello, L., Ravazzani, G., Rabuffetti, D., and Mancini, M.: Integrating glaciers raster-based modelling in large catchments hydrological balance: the Rhone case study, Hydrol. Process., 28, 496–508, https://doi.org/10.1002/hyp.9588, 2014.

Chen, J., Brissette, F. P., and Leconte, R.: Uncertainty of downscaling method in quantifying the impact of climate change on hydrology, J. Hydrol., 401, 190–202, https://doi.org/10.1016/j.jhydrol.2011.02.020, 2011.

Habets, F., Etchevers, P., Golaz, C., Leblois, E., Ledoux, E., Martin, E., Noilhan, J., and Ottlé, C.: Simulation of the water budget and the river flows of the Rhone basin, J. Geophys. Res.-Atmos., 104, 31 145–31 172, https://doi.org/10.1029/1999JD901008, 1999.

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.

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.

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.

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.

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.

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.