Response to anonymous referee #2

Overall assessment:
This manuscript pursues a worthy goal: to quantify the effect of often arbitrarily chosen precipitation lapse rates on hydrological models. However there are important problems of structure and the results are overly justified in ways that go beyond what the data show, often discussing and interpreting what might very well be in fact modelling artifacts. I believe the paper could be made much more concise and readable by trimming down the physical interpretation of the lapse rates that are actually not physics, but properties of the models used.

We thank referee #2 for reading the manuscript carefully and providing thoughtful and constructive comments. Below, our responses are presented in blue and the referee comments in black (bold).

General comments:

1. **What I see in the results is that larger PLRs result in more Q, and more extremes. However this can be explained by the fact that most rain gauges are below the average catchment altitude, therefore positive lapse rates will result in more water than measured at the rain gauge (especially when the nearest-neighbor Thiessen method is used). In a way, in this context the positive lapse rates are "creating" water, especially since the model calibration is kept constant. This alone could explain the correlation of Q with PLR and also the correlation of maximum flows with PLRs. As such, I would need to be convinced that the correlations found are not artifacts of the experimental setting.**

   The effect mentioned above can particularly affect summer floods, therefore it may be a possibility to focus the discussion on the reverse effect observed with winter floods.

   Thank you for this comment. In our methodology, increasing the lapse rates leads to an increase in precipitation (for most catchments as shown in Fig. 3a), which in turn serves as input to the hydrological model. As the referee points out, the characteristics of the station network are important for interpreting the results of our experiment. We have explicitly discussed and acknowledged this aspect at several points in the manuscript. Therefore, we do not consider the correlation to be an artifact; rather, it reflects a relationship that we have identified and discussed in Section 5.3:

- "A shift towards more winter floods occurred when the lapse rate was reduced below 5%, whereas higher lapse rates led to more summer floods (Fig. 7 and Fig. S2)."
- "Spring and summer floods are associated with a combination of high precipitation intensities and snowmelt. In our experimental setup, increasing the precipitation amounts fed into the hydrological model and redistributing them across elevation zones with a higher internal lapse rate resulted in greater snow accumulation at higher elevations. This snowpack is subsequently retained and melts during the summer months, with higher lapse rates leading to a larger snowmelt contribution (Fig. 9). "
- "Overall, both flood frequency and magnitude increase linearly with PLR." Additionally, with our experimental set up we try to understand the changes in flood mechanisms which explained in detail in the last paragraph of section 5.4 where we argue that "Reducing the lapse rate weakens this seasonality, whereas increasing the lapse rate does not substantially alter it." and that "For example, catchments dominated by summer floods and exhibiting strong seasonality show greater dependence between ensemble members than catchments dominated by winter floods (Brunner and Slater, 2022). In our application, increasing the lapse rates increased the frequency and magnitude of summer floods which are characterized by large return periods, while retaining their seasonality pattern."

2. **I feel that the discussions and interpretations are overreaching. In p.16 (results), or in the discussion section, the manuscript discusses processes to a large extent. However, it is only based on model runs, therefore the processes are only those that have been implemented in HBV, which is imperfect as all models.**

   The manuscript should focusing more on the uncertainties of the hydrological model, since it is the main target of the research. The way it is structured and the way things are explained though make it seem that the general effect of lapse rates in floods is analyzed based on physical understanding. This deviates the main target.

   While section 5.5.3 acknowledges these limitations, the fact remains that it makes many of the discussions less relevant.

   Thank you for this thoughtful comment. The HBV model used in this study is a widely applied hydrological model that represents hydrological processes commonly considered in the modelling of Alpine catchments. The limitations are known and are discussed in the manuscript. In p.16 we are largely describing the findings on the sensitivity coefficient. We have revised the section to avoid the interpretation of results, although we find that is was only done to a small extent here.
   We acknowledge that some of the interpretations may give the impression that the simulated responses can be directly attributed to physical processes. The interpretations are, of course, conditional on the process representation in the model. Model parameters are known to be related to seasonally varying processes (i.e., snowmelt) but also to the hydrological conditions examined (here floods). We believe that the interpretation of the results is largely attributed to model processes that are well represented in most hydrological models, keeping in mind that the sensitivity or dominance of certain parameters depends on several factors, such as spatial and temporal resolution, hydrological components and catchment characteristics.
   It should be mentioned that many of the citations we use for the interpretation of results are based on simulated results. A the same time, we ensure that reader is aware to that the results are subject to model uncertainty by mentioning that "These observations are consistent with prior modelling and empirical studies showing that snowpack accumulation…" and later that "Sensitivity coefficients derived from hydrological simulations inherently reflect model complexities (Zhang et al., 2022)."
   However, to address this concern, we have revised the manuscript to work out more clearly that the results are model dependent and to avoid interpretations that go beyond what can be inferred from the simulations. Specifically:

- We changed l452-454 (lines refer to the submitted document) from "This can be explained by the dynamics of snow accumulation and melt processes in catchments at higher elevations (Barrera et al., 2020; Weiler et al., 2025) and the strong role of precipitation in snowpack variability above 1400 m a.s.l. (±200m) in the mountainous regions of Switzerland (Morán-Tejeda et al., 2013)." to ". This can be explained by the critical role of antecedent catchment conditions, such as soil moisture and snow processes, when simulating floods in Alpine catchments (Staudinger et al., 2025)."
- Later we elaborated in l464-465 to remind the reader the associated uncertainty of processes representations. "The variability of snowpack is, however, less straightforward to detect in the snowmelt component because of the varying lag time between the snowfall and its release from the snowpack, which ultimately contributes to streamflow (Schaefli et al., 2013), as well as the challenges to accurately represent these complex processes in the hydrological model."
- Later, in l481 we explain that "In summer, the lower elasticities observed for high-elevation catchments could be linked to higher soil moisture (Sankarasubramanian et al., 2001), which generally increases with elevation in Switzerland (up to 2000 m a.s.l.) (Pellet and Hauck, 2016), as well as to the key role of snow

processes (Sankarasubramanian et al., 2001). " We have replaced the references with a more targeted one and now reads as " Additionally, the importance of soil moisture parameters of the hydrological model increases with elevation (Herzog et al., 2026) which can saturate the catchment and lead to direct runoff." In this recent modeling study, Herzog et al. (2026) use the fully distributed, physically based hydrological model WaSiM to examine the elevation-dependent seasonal and diurnal parameter sensitivities. They found that parameter sensitivities related to soil moisture showed marked seasonality and elevation dependent differences in an Alpine catchment. They specifically mention *"The results show, that overall soil moisture volumes show only little variation on the seasonal scale, but pronounced elevation gradients. The moisture gradient goes from moderate values in low elevations, to a peak in sub-basin 3 with a mean elevation of around 2300 m. Soil moisture volumes decrease from here towards higher elevations, where they begin to show some seasonality with higher moisture volumes during summer and drier conditions during winter. In contrast to the overall moisture volumes, the parameter sensitivities show a more pronounced seasonality.".*

- We deleted l468-469: "Additionally, wet antecedent conditions as maximum soil moisture (Pellet and Hauck, 2016), often coincide with snowmelt (Harpold and Molotch, 2015)." which was referring to processes and we added an explanation earlier in the paragraph based on a modeling study that shows the important of these antecedent conditions. We added a line at the beginning of the section because the link with antecedent conditions fits quite well "This can be explained by the critical role of antecedent catchment conditions, such as soil moisture and snow processes, when simulating floods in Alpine catchments (Staudinger et al., 2025).". We simplified this paragraph explaining the winter sensitivities so that easier to follow.
- We changed l491 from "Additionally, snow protects the glacier from melting" to "Additionally, in the HBV model snow protects the glacier from melting because it does not allow the glacier to melt as soon as snow is present." so that it is clear to the reader that this due to the model conceptualization.

Lastly, the goal of the research is to explore how a single, often overlooked component of the modelling chain – the specification of lapse rates – affects flood frequency estimates. Our study does not aim to assess all sources of uncertainty associated with the hydrological model, or to address the structural limitations of conceptual, lumped hydrological models like HBV. Uncertainties related to the hydrological model and its parameterization have been examined in previous work (see Kritidou et al. (2026)). Rather, we have deliberately limited the scope of this study to isolate and investigate the effects of lapse rates.

3. **The same comment extends to the stochastic rainfall generator model, which also has limits (even though GWEX is very well established). For instance, a lack of seasonal shift in the very rare floods (1000 years) might be a feature of the precipitation generator rather than a question of lapse rates. What would really help here is a discussion and analysis of the results of the GWEX model, because currently the reader does not see the output rainfall amounts, their seasonality, etc.**
To push the argument further, the sensitivity analysis does not enforce physical consistency when it allows PLR_WG and PLR_HBV to be different. While this is fine as these are "statistical" parameters that represent modelling practices, they cannot be discussed in physical terms.

We used seasonality as an indicator of changes in floods based on changes in the lapse rates. This does not imply that the weather generator fails to capture the seasonality of extreme precipitation. Instead, it means that the seasonality of large floods is not substantially affected by the increase of mean areal precipitation which is caused by the increase of the lapse rate in the weather generator. Some of these floods are typical summer floods in mountainous catchments caused by large precipitation amounts. An increase in the mean areal precipitation used as input to the hydrological model does not change the dynamics of these large events. We also clarify this in line 371-373 (submitted doc) "Consequently, more

frequent floods appear to be more sensitive to the changes in PLRs, whereas the mechanisms generating large floods appear to be less affected." As well as in l546-550 "Both in the coupled change of PLR_GWEX and PLR_HBV and the isolated change of PLR_HBV, the CIA decreased with increasing PLR and return period. This may reflect changes in flood generation mechanisms. As discussed earlier, a reduction in PLR leads to a shift towards winter floods. These events are typically smaller in magnitude as precipitation input was altered and are therefore associated with smaller return periods. In contrast, large floods occur mainly in summer and correspond to higher return periods. Thus, adjustments to the lapse rate primarily affect flood events with smaller return periods, while larger events the flood-generating processes are less impacted."

The precipitation amounts are given as relative changes compared to the 5% lapse rate in Fig. 3a, allowing the reader to understand how much the annual precipitation changes when the PLR is varied (from 5% to 0%, 2.5%, 7.5% and 10%). To ensure that readers understand some important information regarding the precipitation amounts from the beginning, we have added some additional plots in the Supplementary material and we refer to it in Section 3.2 ("The simulated precipitation amounts and relevant statistics for each PLR_WG can be found on Fig. S1 of the Supplementary material."). Thus, we provide a first description of the PLRs to the reader, before showing more advanced results.

More information on the seasonality and characteristics of the weather generator can be found in the papers focusing on the development of the model (Evin et al., 2018, 2019), which also assess its suitability. We refer to these papers frequently in the manuscript.

4. **The design of the figures should be revisited. Many figures are too sketchy. There are important inconsistencies in the size of plots and axes.**
   We have revised various figures in the manuscript to ensure consistency.

5. **There are shortcomings in the structure of the paper. Notably**

- **some of the data that are presented in the results sections**

  We have now moved the table of section 4.1 to section 2.2 as it is related to precipitation station network characteristics.

- **some methods are presented in the results section**

  We deleted the following sentences (lines refer to the submitted document) as they might give the impression that methodological steps are presented in the results:

- L258: "As a first step in the analysis, we calculated the number of stations at different elevation bands and the mean length of the observations based on the stations considered for each catchment."
- First sentence of section 4.1 and the first part of the first sentence of section 4.3. ("To understand how sensitive catchments and runoff components are to changes in lapse rates in both the weather generator 315 and the hydrological model,…")
- l315-317,l "To understand how sensitive catchments and runoff components are to changes in lapse rates in both the weather generator and the hydrological model, we computed the sensitivity of different runoff components to precipitation changes induced by varying lapse rates for different seasons."
- l352-353, "To further understand this change in flood seasonality, we calculated the flood seasonality ratio (RF) from the AMFs simulated with all PLRs and then estimated the relative change (ΔRF) between the 5% (baseline) and the other PLRs".

- l396, “We then calculated uncertainty indicators to understand how the choice of PLR affects flood events across different return periods.”.
- Paragraph 1 of section 4.2. We used some relevant information in section 3.3 of the methodology section where the motivation for using the threshold elevation is introduced.
- l540 “We grouped the annual maxima from all PLRs by return period and quantified the corresponding uncertainties relative to the baseline lapse rate (5%).”.

Additionally, we rephrased in l365-367 from “To further understand the effect of PLRs on flood generation mechanisms, we calculated the mean coefficient of variation (CV) of flood seasonality per return period. First, the CV was calculated for each catchment based on the number of events per month, return period, and lapse rate” to “To further understand the effect of PLRs on flood generation mechanisms, we used the mean coefficient of variation (CV) of flood seasonality per return period (Fig. 8).”.

We moved L293-296, “We further estimated the area below the threshold elevation (Fig 4c), which represents the percentage of the catchment where precipitation is simulated as rain. Values above 50% indicate that the threshold elevation exceeds the mean catchment elevation and that precipitation is predominantly simulated as rain in most catchment elevation zones. Conversely, values below 50% indicate that precipitation is simulated as snow in most elevation zones.” to the methodology, just after the presentation of the threshold elevation.

We moved l414-418 to the methodology section as they explain how the RI is estimated and what the motivation of using it is.

Some discussion elements are provided in the results section

To address this comment:

- We revised l330-335 from “The glacier melt runoff sensitivity coefficients were negative for almost all catchments and seasons, indicating that an increase in precipitation leads to a decrease in glacier melt runoff, whereas a decrease in precipitation leads to an increase in glacier melt runoff. Although the correlation with elevation is stronger in winter and spring, the results show an inverse correlation between small (<5%) and large PLRs (>5%). However, it is difficult to interpret the relationship between sensitivity coefficient and elevation, as the number of data points available is limited and the correlation is not statistically significant.” to “The glacier melt runoff sensitivity coefficients were negative for almost all catchments and seasons. Although the correlation with elevation is stronger in winter and spring, the results show an inverse correlation between small (<5%) and large PLRs (>5%). “ to avoid the interpretation of the results which was already in the discussion sections.
- We deleted the part of the last sentence of section 4.1 “…either because these catchments were still well represented by stations or because only a few high-elevation stations contributed to the estimates for small catchments” as it provides some interpretation already mentioned in the discussion section 5.1.
- We deleted the part of l404: “indicating that the choice of the PLR has limited influence on the uncertainty of large floods.”
- We moved l542-545 “A key strength of our approach is that we sample flood events derived from very long synthetic time series, which allows us to examine very rare extremes (up to 1000-year events) for each experiment that are not captured in the relatively short observed time series. Additionally, this comparative analysis allows us to better understand how certain 545 choices in hydrological modelling studies influence runoff simulations and specifically the extremes” to the first paragraph of the conclusions.

Specific comments:

l.24: one could argue that convection is exactly the uplift of air masses or the reason behind the uplift rather than the result, so maybe use more exact language

We have adjusted the sentence accordingly: "They represent an empirical relationship that reflects atmospheric processes such as the orographic uplift of moist air masses, which promotes convection and condensation.".

l.94-95: what is the reason of including this paragraph since in the current approach the PLR is also linear, positive and constant

Thank you for this comment. We have revised it to make the motivation of our research clearer. We changed the sentence to "Even in large sample studies where catchments with considerably different physiographic characteristics are studied, a single value for the lapse rate is often prescribed (Clerc-Schwarzenbach et al., 2024; Clerc-Schwarzenbach and Do Nascimento, 2026; Hegdahl et al., 2021; Parajka et al., 2005). However, the impact of this modelling decision is overlooked."

l.99-100: In some catchments, the station elevation might be close to the average elevation, while in other catchments not. Does this create a bias as some catchments are corrected more than others?

Yes, for some catchments the station elevation differs considerably from the mean catchment elevation. This bias is presented in Fig. 3a of the submitted manuscript and discussed in detail in Section 3.1. Additionally, we highlight this in the discussion and specifically in lines 441-442.

l.102: The unit should be specified in % per 100m. Also, what about negative lapse rates?

We have corrected the unit in the text. Negative lapse rates can indeed occur, and the precipitation-elevation relationship can vary substantially in both space and time (Benoit et al., 2024). However, our study focuses on the effects of different positive lapse rates representing long-term precipitation-elevation relationships, as implemented in bucket-type models such as HBV. Although negative lapse rates may occur under specific conditions, their relevance to long-term average conditions is less straightforward. We therefore restrict our experiments to positive lapse rates.

l.133-134: maybe it would be beneficial to include a reasoning on why these numbers and intervals were chosen, since only 5% is explained below. Also, Why is this range chosen? A reference would be needed to justify it.

Thank you for this suggestion. Previous studies using either precipitation products or modelling approaches have reported lapse rates within a range relevant to our study. For example, Finger et al. (2015) used the HBV in a multiple data set calibration approach to estimate runoff composition in different catchments in Switzerland, showing that the lapse rate parameter varied from 5 to 15% per 100m. Ruelland (2020) used an inverse snow-hydrological modelling approach by calibrating local gradients and found mean rates of 30% km^{-1} in two different models. Ménégoz et al. (2020) found rates of 33% km^{-1} for winter and 38% km^{-1} for summer in the European Alps. Farinotti et al. (2012) found rates of 0.00–0.07% m-1 over 9 high-alpine Swiss catchments, with a mean value of 0.035% m^{-1}. These rates apply to long-term climatology, not to single events, and could therefore serve as a basis for comparison with our study. We have added this information in section 3.1.

Section 2 is very short. It could be beefed up by moving here some of the material in figure 3b and the associated text, which is largely about data description. Also, don't see 27 catchments on figure 1.

Following the referee's suggestion, we have now moved section 4.1 to data description. There are 29 catchments. There might have been some confusion because three large catchments (Thur, Maggia, Inn) were modeled individually in HBV, as were their sub-catchments. We highlighted the outlines of the three large catchments on the map below (number 27, 28, 29) and exchanged the old figure for this map. We have also changed the color scheme slightly to make the distinction of the sub-catchments in the three large catchments clearer. Lastly, we have numbered all catchments.

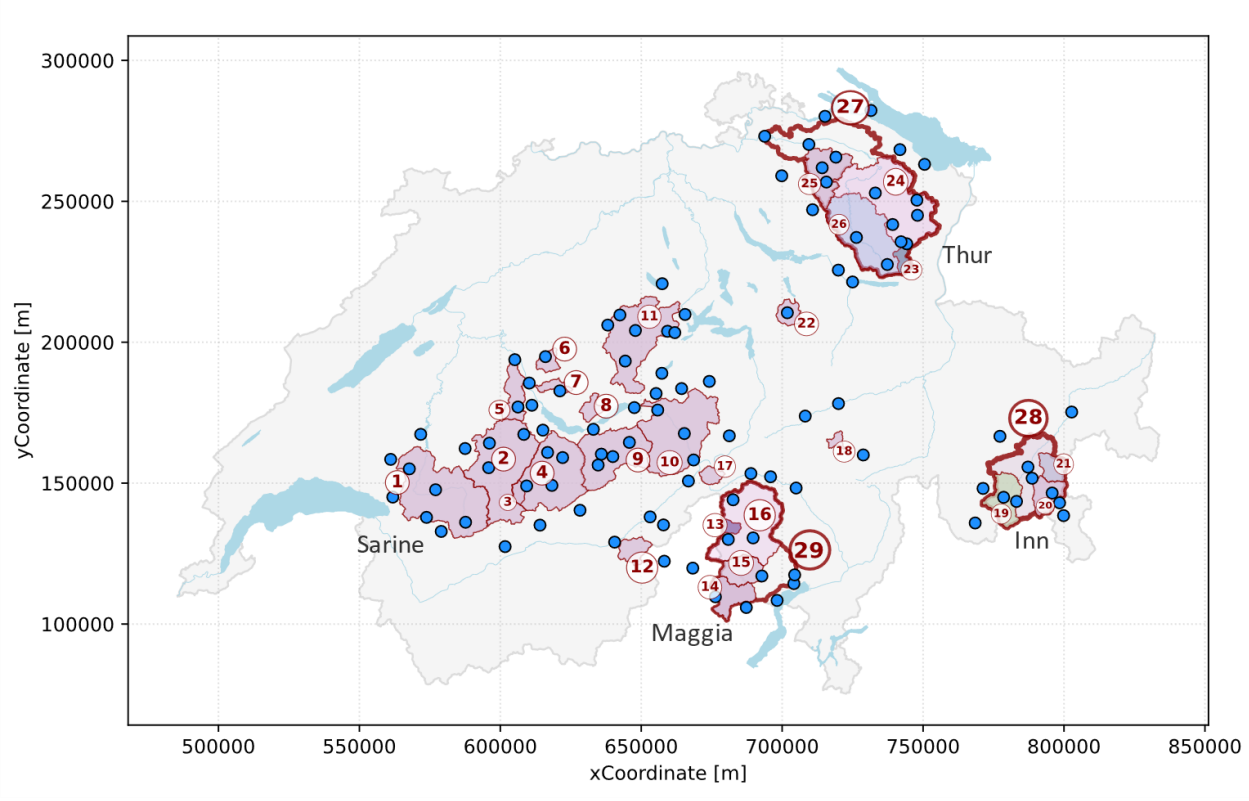


Figure 1 Map of Switzerland with study catchments in purple. The blue dots represent the precipitation gauging station network while the grey areas highlight the exemplary catchments. Relief map: Swiss Federal Office of Topography.

Section 3.1: the text is not very clear in this section. Maybe provide a schematic representation of the workflow and different sensitivities strategies.

We have updated the section as well as the schematic so that the sensitivity strategies are more clearly presented.

l.143: some more background on these papers could be useful since they seem to be a prequel to this work

We added some relevant information to these lines.

l.180-186: since there is a large number of parameters with possibly complex multivariate interactions, it is hard to assess the effect of using reducing this to only 3 representative parameters sets (min, med, max). Can a sensitivity study be shown, maybe on a single catchment, to reassure me that these 3 parameters sets are representative of the entire ensemble?

For each catchment, we derived 100 plausible parameter sets by calibrating the HBV model 100 times using a genetic algorithm (Seibert, 2001). The three selected parameter sets correspond to the median,

the 90th percentile and the 10th percentile of the ensemble of derived exceedance curves (Sikorska-Senoner et al., 2020). Generally, these three selected parameter sets adequately represent the range of the ensemble and are therefore considered representative for the full ensemble. To illustrate the representativity of the chosen parameter sets with regard to annual maximum floods for these 100 parameter sets, we show the results of a single catchment below. In the plot, red, yellow and blue lines represent the upper, intermediate and lower parameter set.

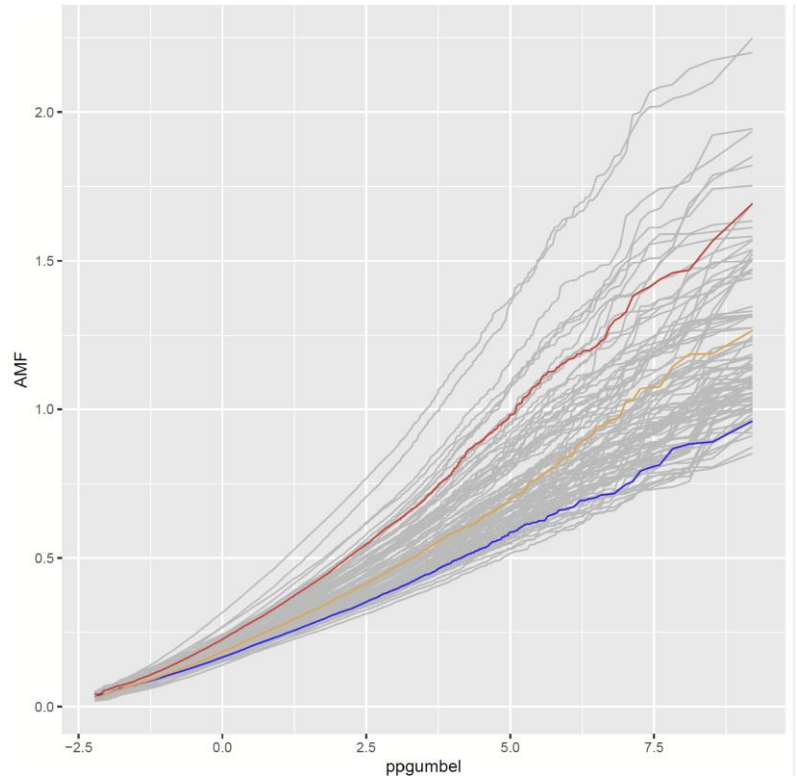
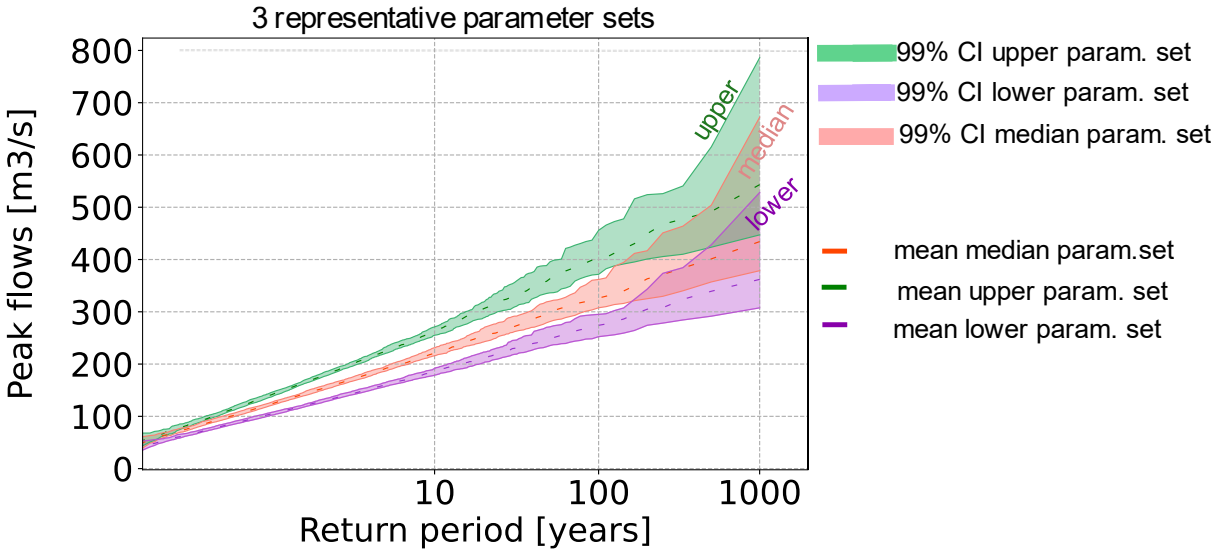


Figure 2 Annual maximum simulated runoff (mm/h) plotted with the Gumbel reduced variate for the 100 parameter sets . The red, yellow and blue lines represent the upper, intermediate and lower parameter set.

In addition, the plot below shows an example of the simulated floods for these 3 representative parameter sets for one catchment.



The logic of the simulations carried out is much better explained in the first sentences of section 3.4 than in section 3.1. I think it should be explained only once, and with a schematic, as mentioned above.
Following the referee's suggestion, we have now merged Section 3.4 with 3.1 and made the necessary adjustments for clarity.

l.214: What is the rationale of this metric for flood seasonality, as others could be envisioned.
We have moved the sentence from l.415 in the methodology to make the motivation clearer. Additionally, we added an introductory sentence explaining that seasonality is considered an important characteristic of simulated floods which helps understand the influence of PLR.

l.222: Here I understand the term elevation zones as what is commonly called elevation bands. It would be useful to specify that these terms are equivalent.
Yes, elevation zones refer to elevation bands. We have retained the term "elevation zones" because it is the terminology most used in the context of the HBV model. However, to clarify the terminology and ensure accessibility for readers more familiar with the term "elevation bands," we have added the latter in parentheses alongside "elevation zones" in line 102.

l.240: this description is a bit redundant and can be clarified
We believe that the description is detailed and clear enough for a reader. We have however improved all the equations in this section. In combination with your comment further below that you could not found the equation for the sensitivity analysis, we believe that is because of the presentation of the metrics in this section. Thus, we revised this section so that it presents more clearly the equations and reorder them to follow the sequence of the presented results.

There should be a notation for the different lapse rates in eq. 6
We have added the notation and revised all equations accordingly (following reviewer's 3 comments).

Figure 2 is useful, but I have the following comments:

- **It should come at the beginning of section 3 rather than at the end**

- **It should provide more details of the sensitivity workflow strategy with both PLRs**

- **The picture on the left, while funny, is not very informative and could be removed to save space**

- **Overall, the graphical design could be improved**

Following the referee's suggestions, we have moved the plot to the beginning of section 3. We also removed the picture on the left and made the sensitivity steps clearer.

Figure 3 shows that rain gauges are largely located in the lower portion of catchments, and that lapse rates are mostly extrapolated to higher altitudes. This should be explicitly discussed.

We agree, and we also explicitly discussed this in many parts of the paper. Especially:

- in section 4.2 we mentioned that "Changes in mean annual precipitation show a clear linear relationship with ΔElevation: catchments with a greater difference between mean catchment elevation and mean station elevation have larger changes in annual precipitation, reaching values of up to ±40%. "
- in section 5.1 we mentioned "The linear correlation between ΔElevation (mean catchment elevation minus the mean Thiessen-weighted station elevation) and differences in mean annual precipitation from the weather generator (calculated for 0%, 2.5%, 7.5%, and 10% relative to the baseline 5%) highlights the influence of the meteorological network on the generation of synthetic time series for each catchment. In the weather generator, the lapse rate is applied to every station to adjust precipitation to the mean catchment elevation. When the mean catchment elevation deviates substantially from the elevations represented by the station network, the lapse rate adjustment amplifies precipitation discrepancies. "

Figure 3: The labels PLR000 etc are not very clear. Either rename or explain in legend

We added a short explanation in the figure caption.

l.357: Where does this change in mean annual maximum precipitation comes from? It seems that there is even a change of 80% in some cases, so it might be beneficial to explain how is this change is imposed in the model

In GWEX, the lapse rate is used to adjust precipitation for the elevation difference between stations and the mean catchment elevation prior to spatial interpolation to mean areal precipitation using Thiessen weights. Using the different PLRs, we obtain five different MAP synthetic time series, which serve as inputs to the hydrological model. For each PLR, we calculated the annual maximum. This is illustrated in Section 3.2 ("Overall, GWEX generated a 30,000-year time series (i.e., 30 scenarios of 1,000 years each) for each station on an hourly time scale. Precipitation at each station is individually treated with a lapse rate (PLR_WG) to adjust the data from station elevation to mean catchment elevation. Then, the adjusted precipitation station data are interpolated with Thiessen weights to determine mean areal precipitation (MAP). Here, we vary PLR_WG in a reasonable range (0%–10%) to explore its impact on runoff simulation.")

We calculated the differences between the annual maximum values obtained with the 5% and those obtained with the other PLRs. Then, we estimated the mean annual maximum values used in this plot. However, we identified an inadvertent duplication of data points in the original plot, which resulted in more points being displayed than intended. This plotting error has been corrected, and Figure 7e has been updated accordingly. The differences now go up to 40%. This correction does not affect the correlation results or the main conclusions of the study.

In response to the reviewer's comment, we have added a sentence at the beginning of section "Indicators for the influence of PLRs": "For our analysis, we have estimated the mean differences of mean annual

and annual maximum precipitation between the 5% PLR_WG and the other PLR_WG values.". This clarifies that we estimate these differences and then use them in the analysis.

Section 4.1 mixes results and data description, some of it could be moved to the data description section.

Thank you for this comment. We understand that some of this information might fit well in the data description. Accordingly, we moved the table showing the number of stations across elevation classes to the data description, as it provides an overview of the dataset. However, we retained Figs. 3a and 3b, as we consider these plots to present results rather than merely describe the data. We considered moving them to the data description section, but that would require the presentation of PLR_WG (which is part of the methods) prior to that, which would make the structure of the manuscript more complicated.

Section 4.2: At this stage, it is not clear for the reader why the threshold elevation is introduced and its relation with the PLR. This should have been explained beforehand, when the entire methodology is explained.

The motivation for using the threshold elevation is first introduced in the last paragraph of section 3.3 in the methodology section, together with equation 1, which explains how the threshold elevation is derived. However, part of the motivation is also presented in section 4.2. In response to the referee's previous comment about part of the methods being in the results, we have moved part of paragraph 1 of section 4.2 to the methodology where the motivation for using and the explanation of threshold elevation is presented for the first time. We revised the paragraph to avoid repetitions with the material that was presented before.

Figure 4 is graphically imbalanced with some very small and some very large plots. In c), it is unclear why the individual samples needed in the whiskers plot. Axes limits are inconsistent.

Following the referee's suggestions, we have revised the figure to address the imbalance. However, we kept the dots, as these are useful for illustrating the extremes outside the whiskers.

Section 4.3: This sensitivity analysis is presented as results, but was not described in the methods section. Therefore, the reader is confused because there seems to be a discrepancy between what has been described in the methods and the content of the results. It is unclear how the sensitivity coefficient epsilon is calculated. Frustratingly, I could not find it in any of the equations given. I am unfortunately confused and not able to follow the discussions on positively or negatively correlated sensitivity, since sensitivity is not clearly defined.

We thank the reviewer for this comment. The sensitivity coefficient (epsilon) is defined in Equation (5) in Section 3.5, along with a detailed description and relevant references. However, to improve the logical flow and facilitate the reader's understanding, we have reorganized the presentation of the indicators in Section 3.5, introducing the sensitivity coefficient first, as it is the first indicator discussed in the Results section (Section 4.3). We hope that this reorganization makes the link between the methodological framework and the subsequent presentation of the results clearer.

The discussion of figure 8 should revolve on the effect of varying lapse rates, however these are not clearly discussed. Instead, the discussion is about the processes underlying floods. To understand, the reader has to go back to the text describing equation 6. Graphically, the colors do not convey a meaning.

Our discussion of the processes underlying flood generation is intended to illustrate how changes in precipitation lapse rates (PLRs) affect the simulated hydrological processes, with different responses

observed between high- and low-elevation catchments. In particular, our results show that high-elevation catchments exhibit strong seasonality, which tends to be preserved even when the PLRs are varied and precipitation inputs increase. This suggests that the hydrological response of these catchments is less sensitive to changes in precipitation induced by varying the PLRs.

We have revised the text to clarify the link between the effects of varying PLRs and the resulting changes in simulated processes. Specifically, we changed the sentence in l551 (submitted document) to "This finding aligns with the lower sensitivity of high-elevation catchments to changes in precipitation caused by the change in PLRs." In this paragraph, we also attribute our results with results from the modeling study of Brunner and Slater (2022), that support that the catchments with strong seasonality show greater dependence among ensemble members.

Regarding the colors, we changed the plot so that the boxplots have the same color, as they differ only in that they represent a different return period (x axis).

I find figure 9 redundant with the second column of figure 5

The second column of Figure 5 presents the sensitivity coefficient (epsilon), which quantifies the percentage change in the snow component in response to a 1% change in precipitation due to the change in PLR. This analysis focuses on seasonal variations in the snow component. In contrast, Figure 9 examines the contribution of snowmelt to annual maximum runoff and how this contribution changes with varying precipitation lapse rates (PLRs). To obtain this contribution, we calculated the ratio of simulated snowmelt to simulated annual maximum runoff for each catchment. Then we calculated the mean monthly values for all catchments. This approach allows us to assess how changes in PLRs affect the average contribution of snowmelt to flood peaks on a monthly scale. Therefore, although both figures address the snow component, they provide complementary information and are not redundant.

l.415 and below: The explaination of the rationale for RI should come when RI is explained, in the methods section, rather than in the results

We have moved the l.415-419 in the methodology to make the motivation clearer.

In figure 11, I don't see a clean and interpretable dependence of the relationship with PLR

We agree that there is some scatter in Fig. 11 that makes the relationship between the redistribution index (RI) and the mean absolute relative change in AMFs less clear.

The strength of the positive correlation generally increases from PLR000 to PLR075, as indicated by the Spearman correlation coefficients (ρ = 0.515, 0.645, and 0.861, respectively), before decreasing slightly for PLR100 (ρ = 0.816). The scatter observed at smaller lapse rates can be explained by the fact that, for PLR000, precipitation is not redistributed to different catchment elevation zones. Consequently, the influence of the RI on differences in simulated AMFs is less pronounced and contributes to a weaker relationship.

l. 456-458: I disagree with this statement: usually higher-elevation catchments tend to have thinner (or no) soils, less alluvial aquifers, and therefore can store less groundwater than lower-altitude catchments.

Thank you for pointing this out. A study from Staudinger et al. (2017) found that high-elevation catchments are associated with large dynamic and mobile storage estimates with snow and groundwater contributing considerably to that. Specifically, the authors found that with increasing elevation the relative contribution to the dynamic catchment storage was similar for groundwater storage while it

increased for snow and decreased for soil. These findings highlight that the relationship between elevation and catchment storage capacity is complex and depends on the characteristics of individual catchments. A generalization based solely on elevation may be a simplification and therefore we have revised the statement accordingly. We have now revised the statement in l456-458 to "In addition to seasonal snow storage, alpine and pre-Alpine catchments may contain important subsurface water stores (Cochand et al., 2019; Somers and McKenzie, 2020; Staudinger et al., 2017; Weiler et al., 2025), although they may have thinner soils and lack of alluvial aquifers."

l.491: the insulating affect of snow on glaciers is highly nonlinear and unlikely to be fully represented in HBV

Thank you for this comment. We acknowledge that this is a highly complex process, which is simplified in typical hydrological models. However, we mentioned this in the manuscript because, in the HBV model, it is conceptualized such that there is no glacier melt when snow is present. We tried to clarify that by changing the sentence to "Additionally, in the HBV model snow protects the glacier from melting because it does not allow the glacier to melt as soon as snow is present. "

l. 514: I believe that this pattern is not convincingly explained. Since with higher lapse rates we get more snow in most catchments, since the mean elevation of most of them is above the station altitude. Is it an albedo or local circulation/temperature effect? Or could it be that with lower lapse rates we get more precipitation in elevations below the station, which is mostly rain rather than snow, and that contributes to direct floods caused by the rainfall amount rather than the snowmelt? (This would contrast Figure 9 though)

The amount of snow we get in the catchment is indeed related to the difference between mean catchment elevation and mean station elevation, as this difference increases or decreases mean areal precipitation. However, whether precipitation falls as rain or snow depends on the threshold elevation, as explained in section 4.2. As mentioned in a previous comment by the reviewer, perhaps the motivation for the threshold elevation was insufficiently clear, but we have now attempted to clarify it. In Section 4.3 (lines 299-302), we explain that with a 0% lapse rate there is no redistribution of precipitation across elevations. Therefore, more precipitation falls in lower elevation zones than with larger lapse rates, which produces snowmelt earlier in the season rather than snow accumulation and snowpack development during the winter season. Fig. 9 confirms that by showing that the contribution of snow to the peak flows is higher in winter for small PLRs. Snow accumulation at high elevations is lower for small PLRs than for higher PLRs. Snow falling at lower elevations, and specifically below threshold elevation is indeed rain and not snow and can lead to direct floods. The changes in flood mechanisms that explain this pattern are discussed in detail in the following paragraphs of section 5.3. This pattern is supported by different figures and discussed in detail throughout the document, and we hope that the revisions will make this even clearer.

l.532: I am not sure I get the explanation. Is it implied that in that case it is easier for the snow to start melting in higher elevations already in the winter? Is such a claim that intuitive or self-justified?

Thanks for this question. The statement may not be sufficiently explained. In line 531-533 we mentioned "This snowpack is subsequently retained and melts during the summer months, with higher lapse rates leading to a larger snowmelt contribution (Fig. 9). This also explains why the 0% PLR produces most low-magnitude floods in winter, whereas in summer both flood frequency and magnitude increase linearly with PLR."

We mean that a 0% PLR in the weather generator and the hydrological model leads to an overall smaller amount of precipitation, and no redistribution to higher elevations is applied. That leads to lower precipitation amounts in higher elevations compared to other PLRs. As a result, we observed more low-magnitude floods which come from snow melting at the lower elevations. Therefore, we did not intend

to state that snow starts melting in high elevation already in winter but rather that less snow accumulates at higher elevations because precipitation is not redistributed to the higher elevations. We have therefore rephrased the sentence to avoid misinterpretation and now it reads: "*This snowpack is subsequently retained and melts during the summer months, with higher lapse rates leading to a larger snowmelt contribution (Fig. 9) and overall both flood frequency and magnitude increase linearly with PLR. Additionally, the opposite is true for the 0% PLR which produces most low-magnitude floods in winter because precipitation falls either as rain, or, in case of snow, melts earlier in the season*"

Minor comments/typos:

l.249: word events written twice

l.266: add a reference to figure 3a

l.283: in -> on

caption of figure 4: s -> is

l.550-551: rephrase

l.655: rephrase

Thank you for the minor comments. We have addressed them all in the manuscript.