

Reviewer 1

The manuscript “*The influence of lakes and reservoirs on flood peaks at hourly vs. daily timescale in Switzerland*” (i) analyzes the effect of water bodies (i.e., reservoirs and lakes) on flood peak at both hourly and daily timescale (i.e., the ratio daily/subdaily) in Switzerland, (ii) identifies parameters/catchment features that exhibit a significant effect on the D/H ratio, and (iii) concludes, if I correctly interpret the main message, that the ratio is particularly high (i.e., tends to one, hence daily and subdaily flood peak tends to converge) when the contributing water body area exceeds about 60% of the catchment area. Hence, if only the catchment area is considered to assess the ratio D/H, which is usually 0.8 for catchments larger than 5000 km², there is a risk of underestimating this ratio in smaller basins that are strongly influenced by water bodies (at least for the small return periods considered here).

Overall, I find the general idea of assessing the water bodies effect on the daily/subdaily ratio quite interesting. In my view, the study suggests that for gauges highly influenced by waterbodies, flood frequency analysis on daily peak flows could, in some cases, provide a reasonable approximation of instantaneous peak flows (IPFs), even in relatively small catchments.

However, I have some concerns regarding the framework setting, the general writing, and, most importantly, the discussion and interpretation of the findings.

In particular, I think that the framework setting, especially the inclusion of different catchment descriptors in the random forest model, is not sufficiently discussed from a physical perspective. A deeper interpretation of the results would be necessary. Catchment area and contributing area remain the main parameters discussed, although two additional descriptors are included in the model (biogeographical region and geological permeability). Their role is not clearly interpreted in the discussion, and they seem to disappear from the narrative once the model results are presented.

In addition, I believe there is room for improvement in the writing, which is sometimes repetitive and slightly dense.

Finally, I do not find that the results and conclusions are fully clarified in terms of their implications. For example, it is not entirely clear what is meant by “large-sample analyses”, which could refer to several different methodological frameworks (e.g., regional flood frequency analysis, IPF estimation, attribution studies). I recommend that the authors be more explicit and precise about the practical message and utility of their findings.

Overall, I recommend publication of the paper, since the topic is highly relevant and I appreciate the general idea, but only after a **major revision**, where I would expect improvements in the clarity of the text and a deeper interpretation of the findings, especially regarding the role of the selected catchment descriptors and the implications

for large-sample studies

Reply: *Thank you very much for your constructive review and acknowledging the value of our work. Based on your comments, we have made the following major adjustments: (1) improved the framing of our work by aligning research questions, results, and conclusions; (2) included all of the predictors of the random forest model in the discussion; (3) carefully edited the text to improve the reading flow; and (4) clarified the practical implications of our work with respect to flood estimation.*

Below, I provide specific comments and technical corrections.

(lines 3 to 6): Could you please rephrase this sentence (especially points (1) and (2))?

Since you refer to “approaches”, the points would be clearer if presented as “(1) by comparing upstream and downstream gauges over four local case studies, and (2) ...”;

Reply: *Thank you for highlighting the need to rephrase this sentence. We rewrote it in the following way: ‘Water bodies such as lakes and reservoirs can play a crucial role in reducing flood peaks, at both daily and hourly time resolution. While previous research has relied on large-samples of catchments to estimate the differences between IPF and daily mean flows, it remains unassessed how the effect of standing water bodies influences the ratio between daily mean and instantaneous flow peaks downstream of water bodies, that is to which degree the attenuation effect of reservoirs on flood peaks is scale dependent. In this study we aim to address this research gap by asking how water bodies influence the ratio between IPF and daily mean flows and therefore the attenuation effect on daily vs. instantaneous peak flows. To address this question, we follow two approaches that complement and reinforce each other: a case study approach that strengthens process understanding and a large-sample analysis that allows us to draw generalizable conclusions.’*

(line 6): I suggest using “hourly peak discharge” instead of “hourly flows”, which would be more precise;

Reply: *We replaced ‘hourly flows’ by ‘peak discharge’ in line 6 and throughout the manuscript.*

(lines 10 to 12): What exactly do you mean by “... that it should be considered in large-sample analyses using suitable metrics”? The term “large-sample analysis” is quite general. It would be useful to specify which type of large-sample study is concerned and how your results would concretely affect such analyses;

Reply: *We clarified that the term ‘large-sample’ is used in contrast to case studies for which detailed information on water body influence may be available: ‘This suggests that water body influence should also be considered in large-sample studies that go beyond case studies for which the exact influence may be known and quantified using suitable metrics such as the fraction of the total catchment above the water body.’ We furthermore clarified the implications of our results: ‘This implies that daily flood peaks are a good proxy for hourly peaks in regulated catchments in contrast to natural ones.’*

(Line 17): “However, information on IPFs is often not available and IPFs can differ substantially from daily flows.” I think this is essentially the core motivation of the study and should be emphasized more strongly, also by highlighting that subdaily data are often less available than daily data;

Reply: *This statement is indeed crucial for the motivation of this study and we emphasized it even more by rewriting the first part of the introduction: ‘Floods can develop rapidly at sub-daily time scales if they are triggered by intense convective storms, especially in small catchments (e.g., Blöschl, 2022). Therefore, information on maximum flows occurring at a sub-daily time scale, so called instantaneous peak flows (IPFs), is crucial for the design of flood protection infrastructure. Still, floods are often described using daily data because sub-daily data are less frequently available than daily data. Because information on IPFs is often not available, considerable research efforts have been put into estimating IPFs from more commonly available daily flow data (e.g., Fuller, 1914; Ding et al., 2015, 2016; Ding and Haberlandt, 2017; Ellis and Gray, 1966; Bartens et al., 2024; Fill and Steiner, 2003). Such estimates are derived by relating the ratio between daily mean flows and IPFs, which is referred to as the peak ratio, to catchment characteristics that are also available in ungauged basins (Ding et al., 2015).’*

(line 31): Consider “varying depending on reservoir characteristics” instead of “between reservoirs”;

Reply: *We rephrased the sentence as suggested.*

(line 67): “(2) to which degree can the consideration of water bodies improve the analysis of flood flows in large-sample analyses”

Could you clarify this research question? How exactly do you expect your analysis of the daily/subdaily peak ratio under water body influence to improve large-sample analyses? Which specific aspect of large-sample flood studies would benefit from this? This point should be clarified (also in lines 73–74);

Reply: *Thank you for highlighting this inconsistency in the framing of the paper. We removed this research question to improve the alignment of questions, results, and conclusions.*

(line 68): ‘first’ instead of ‘First’;

Reply: *We made the suggested adjustment.*

(lines 80-81): This part seems partially repetitive;

Reply: *We combined the two sentences into one to avoid redundancy: ‘We selected four case studies with streamflow gauges both up- and downstream of the water body to demonstrate how daily and sub-daily flows are dampened by water bodies locally.’*

(lines 89-90): I do not fully understand the configuration of the Walensee case study. From Figure 1b it appears that both gauge stations are located downstream of the lake, or that the “upstream” gauge drains a different catchment. Could you clarify this hydrological configuration? Also, considering the 77% increase in catchment area

between the two gauges, how does this affect the interpretation of the attenuation results?

Reply: *Thank you for pointing out the need to clarify the hydrological configuration of the Walensee case study. We added the following note to the text: ‘Please note that the Linth (draining the area of Canton of Glarus) comes from the South, flows into the Walensee and then leaves the Walensee to the West. As a consequence, the upstream catchment is very narrow and located close to the upstream station (southern gauge).’ The 77% increase in catchment area between the two gauges means that there are several tributaries that may also influence flood peaks in addition to the water body. We highlighted this in the following way: ‘This means that this case study does not allow for a clean separation between the influence of the water body and tributaries on the flood peaks.’*

(lines 93-94): I do not fully understand this sentence, specifically why those early periods were excluded from the samples. Please clarify the reasoning;

Reply: *Thank you for highlighting the need for clarification. We added the following explanation: ‘We trimmed both the up- and downstream time series to the period after the construction of the reservoir where necessary such that we could directly compare natural flood peaks upstream of the reservoir with their regulated counterparts downstream.’*

(line 125): When more than one water body is located upstream of a gauge station, how was the contributing area evaluated?

Reply: *We clarified that ‘Then, we calculate the total area which lies above the most downstream water body in case it is located on the main stem or the most downstream water bodies in the tributaries if they are located in the tributaries and divide it by the river catchment area.’*

(lines 224 to 226): Here you describe the random forest model used to identify the most relevant catchment descriptors for explaining the D/H ratio. You identify contributing area percentage and catchment area as the most important descriptors, together with biogeographical region and geological permeability. Regarding the latter two variables, why is no physical interpretation provided? In the discussion, these descriptors are not considered anymore, although they were selected as relevant by the model. I think a deeper interpretation of their role is necessary, especially in terms of hydrological processes controlling the daily/hourly peak relationship. Also, the analysis is conducted only for the 10-year return period, correct? What do you expect would happen for higher return periods, for which the attenuation effect of water bodies is usually smaller? This limitation should be discussed more explicitly.

Reply: *Thank you for highlighting the need to come back to the other predictors in the discussion. We added the following statement to the discussion: ‘While catchment area and contributing area percentage are the most important predictors of the D/H ratio, other predictors such as the biogeographical region, a metric for catchment similarity in*

terms of climate, geology, and vegetation, and other catchment characteristics such as permeability, an indicator for flashiness, only have limited predictive power (Fig. 5)' Yes, the main part of the analyses focuses on the 10-yearly flood peak because this estimate can be reliably derived given the rather small sample at hand. We clarified this in the methods section: 'We focus on 10-yearly extreme flows since these represent a good compromise between regular floods (e.g., 2-yearly) and more extreme floods (e.g., 30-yearly) and because they can still be estimated reliably given the rather small sample at hand, which would not be the case for very rare floods with longer return periods.' We also clarified this focus in the results section: 'To investigate how important these two factors are relative to each other for explaining the D/H ratio related to the 10-yearly flood peak...'

(line 236): Should this refer to Fig. 6a instead of Fig. 6b?

Reply: *Yes, we adjusted the reference accordingly.*

(lines 273-275): "Therefore we conclude that catchment area is a crucial factor influencing the peak ratio – unless much of the catchment area lies above water bodies." If I understood correctly, you are highlighting that catchment area generally controls the D/H ratio, but when a large portion of the catchment is drained by water bodies, this control is overridden and the ratio approaches unity due to strong attenuation of hourly peaks. I suggest reformulating this conclusion more explicitly, as it represents one of the key messages of the manuscript.

Reply: *We have rewritten parts of the discussion section to better highlight our key finding and its implications for flood estimation: 'Even small catchments could have similar daily and sub-daily flood peaks if they are strongly influenced by water bodies... These findings suggest that daily flood peaks are a good proxy for hourly flood peaks in strongly regulated catchments. This means that in such catchments, instantaneous peak flows may be estimated from daily peak discharge in case sub-daily data is not available.'*

Reviewer 2

I appreciated the manuscript, particularly the clarity of the writing and the way the text guides the reader through the rationale and objectives of the work. The study addresses an interesting and relevant topic, namely the influence of water bodies on flood peaks at hourly versus daily timescales. Overall, the manuscript is well structured and readable, and the topic is timely given the increasing availability of sub daily hydrological data.

Reply: *Thank you very much for assessing our manuscript and highlighting the relevance of our work.*

I found that the motivation of the study would benefit from being emphasized more

clearly in the introductory section. In particular, the availability of instantaneous peak flow (IPFs) information appears to represent a central element underlying the relevance of the work, yet this importance is not always highlighted as strongly as it could be. Strengthening the discussion around the limitations in the availability of IPF data, especially in comparison with daily observations, would help clarify why the present analysis is needed and what gap it is intended to address. Expanding on the challenges related to subdaily data availability would also reinforce the timeliness of the study.

Reply: *Thank you for stressing the need to better highlight the motivation for our study in the introduction. We rewrote the first part of the introduction to better highlight the relevance of information on IPFs and that such information is often lacking: ‘Floods can develop rapidly at sub-daily time scales if they are triggered by intense convective storms, especially in small catchments (e.g., Blöschl, 2022). Therefore, information on maximum flows occurring at a sub-daily time scale, so called instantaneous peak flows (IPFs), is crucial for the design of flood protection infrastructure. Still, floods are often described using daily data because sub-daily data are less frequently available than daily data. Because information on IPFs is often not available, considerable research efforts have been put into estimating IPFs from more commonly available daily flow data (e.g., Fuller, 1914; Ding et al., 2015, 2016; Ding and Haberlandt, 2017; Ellis and Gray, 1966; Bartens et al., 2024; Fill and Steiner, 2003). Such estimates are derived by relating the ratio between daily mean flows and IPFs, which is referred to as the peak ratio, to catchment characteristics that are also available in ungauged basins (Ding et al., 2015).’*

In addition, the conceptual basis for comparing daily and hourly flood peaks could be further strengthened by making the underlying hypothesis more explicit. You stated that in many basin-scale studies, water bodies are not systematically incorporated when estimating peak flows, particularly when transferring information across temporal resolutions. This can obscure the physical mechanisms controlling the relationship between daily and subdaily peaks. Your study recall that attenuation effects are scale-dependent, with a disproportionately stronger damping at the hourly (and sub-hourly) scale, which in turn reduces the contrast between daily and hourly peaks downstream of reservoirs. Explicitly framing the study around this scale-dependent attenuation would help clarify the novelty of the work and better justify the comparison at its core.

Reply: *Thank you very much for these suggestions on how to strengthen the framing of our study. We rewrote parts of the introduction to explicitly frame the study around the scale-dependent attenuation effect of reservoirs and lakes: ‘Both lakes and reservoirs can better buffer intense and short flood events than longer-lasting floods, because the flow volume of these events is usually limited (Xiong et al. 2019). This suggests that the attenuation effect of reservoirs on flood peaks may be scale dependent, which may reduce the contrast between daily and hourly flood peaks downstream of reservoirs.’ And ‘While previous research has relied on large-samples of catchments to estimate the differences between IPF and daily mean flows, it remains unassessed how the effect of*

standing water bodies influences the ratio between daily mean and instantaneous flow peaks downstream of water bodies, that is to which degree the attenuation effect of reservoirs on flood peaks is scale dependent. In this study we aim to address this research gap by asking how water bodies influence the ratio between IPF and daily mean flows and therefore the attenuation effect on daily vs. instantaneous peak flows.'

A further point that could be improve is the clarity of the two methodological approaches and their respective roles. While the distinction between local case studies and the large-sample Swiss dataset is valuable, their complementary contributions are not fully highlighted. It would strengthen the abstract to briefly clarify how these approaches reinforce each other (e.g., process understanding from case studies versus generalization from the large-sample analysis).

Reply: *Thank you for this concrete suggestion on how to better highlight the value of the two complementary approaches, which we know emphasize in the introduction: 'To address this question, we follow two approaches that complement and reinforce each other: a case study approach that strengthens process understanding and a large-sample analysis that allows us to draw generalizable conclusions. Using the case study approach, we first analyse the attenuation of sub-daily vs. daily peak flows by lakes and reservoirs at four locations in Central Europe, for which streamflow data are available up- and downstream of water bodies. This approach allows us to illustrate the differences in the dampening effect of water bodies on flood peak quantiles derived from datasets with different temporal resolutions. Using the large-sample approach, we second explore the impact of the catchment fraction influenced by a water body on the ratio between daily and hourly flood quantiles across Switzerland. This allows us to demonstrate that the results derived for the four case studies are generalizable to other catchments.'*

From my perspective, also the discussion could be strengthened by more clearly linking the modelling choices to hydrological processes. The random forest model incorporates several catchment descriptors, yet their role is not consistently interpreted in a physically meaningful way throughout the manuscript. Providing a more explicit process-based explanation of the results would enhance the overall contribution of the study. In this context, it would be helpful to revisit all included predictors when discussing the outcomes, rather than focusing primarily on catchment and contributing area. Descriptors such as geological permeability and biogeographical region are potentially informative, but their influence is not clearly articulated and they are largely absent from the interpretation of the model results. A more balanced discussion that explains how each variable contributes, and why, would improve coherence and reinforce the connection between the data-driven analysis and the underlying hydrological mechanisms.

Reply: *Thank you for your suggestion on how to strengthen the discussion. We added a*

more balanced discussion of the results presented in Figure 5, expanding beyond the role of contributing area percentage in shaping the D/H ratio: ‘While catchment area and contributing area percentage are the most important predictors of the D/H ratio, other predictors such as the biogeographical region, a metric for catchment similarity in terms of climate, geology, and vegetation, and other catchment characteristics such as permeability and an indicator for flashiness, only have limited predictive power (Fig. 5)’

Specific comments

Given that the study is largely descriptive, the interpretation of the results would benefit from additional supporting information or analyses, as detailed below.

- Title

The geographical scope of the study is generally clear in the manuscript, but the title could better reflect the full spatial extent of the analyzed catchments. Since the study includes basins beyond a single national context, aligning the title with the broader geographical coverage would improve consistency between the framing of the work and its actual domain. Since some of the analyzed catchments are located in Germany, I suggest reconsidering the title.

Reply: *Thank you for highlighting this inconsistency. We widened the geographical scope of the title to: ‘The influence of lakes and reservoirs on estimated flood peaks at hourly vs. daily timescales in the Alps’.*

- Return periods and additional analyses

Given the descriptive nature of the study, I suggest adding graphs for at least one additional return period (e.g., 25-year floods). This would allow Figures 3 and 4 to include additional curves and help highlight the influence of return period. Alternatively, the authors should explain why additional return periods were not considered and clarify why the analysis stops at the selected values.

Reply: *We justified the focus on the 10-year flood in the large-sample analysis as follows: ‘We focus on 10-yearly extreme flows since these represent a good compromise between regular floods (e.g., 2-yearly) and more extreme floods (e.g., 30-yearly) and because they can still be estimated reliably given the rather small sample at hand, which would not be the case for very rare floods with longer return periods.’*

- Robustness of flood estimates

In the sentence “Additionally, we can estimate these 10-yearly floods fairly robustly, given that the streamflow records for most stations exceed 10 years by a large margin,” I suggest being more specific about what “a large margin” means. Please provide

quantitative information on record length (e.g., median, minimum, or range of years).

Reply: *We specified that ‘80 % of catchments have at least 40 years of data’.*

- Figure 2

I understand that the distinction between upstream and downstream stations is conveyed through transparency differences. However, this is not immediately clear. I suggest increasing the transparency contrast to make the distinction more visually evident.

Reply: *We clarified the meaning of the two different transparency levels in the figure caption: ‘...with upstream and downstream catchments highlighted in saturated and transparent colors, respectively.’*

- Lines 205–208

I suggest explicitly recalling the relevant figures after each description of the effects of water bodies on high flows, to facilitate interpretation and improve readability.

Reply: *We added appropriate figure references to each of these statements: ‘(1) sub-daily and daily flows become more similar to each other (Fig. 2); (2) sub-daily flows decrease more strongly than daily flows (Fig. 3a) and (3) absolute extreme flows decrease (Fig. 2).’*

- Figure 4

It would be helpful to visualize the relationship between the x and y variables more clearly, for example by adding the regression line and reporting the R^2 value directly in the figure. This would allow readers to better assess the strength and magnitude of the relationship.

Reply: *We refrained from adding a regression line because what we wanted to highlight in this figure is the role of reservoir influence (contributing area fraction) rather than catchment area.*

- Line 235 (definitions)

Some parts of the discussion appear repetitive, and multiple terms are used to refer to similar concepts. The term “non-water-body-influenced” is introduced here. Is this equivalent to catchments with a contributing area percentage below 60%? The sentence currently combines several definitions (“weakly correlated,” “contributing area percentage below 60%,” and “non-water-body-influenced”), making it difficult to follow. I suggest reformulating this section to clearly define each category and their relationships.

Reply: *Thank you for highlighting the need for clarification. We reformulated to: ‘There are substantial differences in the D/H ratio between weakly water-body-influenced (contributing area percentage below 60%) and strongly water-body-influenced catchments (contributing area percentage above 60%).’ We avoid using the term non-water-body influenced in this context to avoid confusion.*

- Lines 239-240

It is not entirely clear why the model is trained on weakly influenced catchments and then applied to predict the D/H ratio in strongly influenced ones (lines 239–240). The rationale behind this choice should be more explicitly stated. As it stands, the reasoning is only implicit, leaving some ambiguity about the intended purpose of this modelling strategy.

I suggest clarifying whether the objective is to establish a baseline relationship under near-natural conditions and then assess deviations induced by strong water body influence, or whether there is another justification. Providing a clear explanation of this choice, along with its implications and limitations, would improve the transparency of the methodology and help the reader better understand the interpretation of the results.

Reply: *We clarified the rationale behind this choice both in the methods and results section as follows: ‘We train this model on the weakly influenced catchments and attempt to predict the D/H ratio of the strongly influenced catchments to highlight the importance of including water-body influence information to explain the D/H ratio in regulated catchments.’*

- General readability

In several places, sentence readability could be improved where multiple definitions are introduced simultaneously (lines 235-245). Shorter sentences and clearer structure would help readability.

Reply: *We rewrote the section highlighted as described above and went through the entire manuscript again to improve and clarify sentence structure.*

Final recommendation

Overall, the manuscript addresses an important topic and is clearly written. However, based on the points raised above, particularly regarding the need for clearer terminology, stronger emphasis on the study motivation, and additional supporting analyses, I recommend major revisions before the manuscript can be considered for publication.

Reviewer 3

I liked reading this paper which is nicely and clearly written. The results are not surprising but anyway useful in providing some quantitative statements about the effects of the influence of reservoirs/lakes (and area) on the ratio between hourly and daily peaks. I have just some comments that I hope will be useful to improve the paper through a minor revision. More generally I think that:

Reply: *Thank you very much for acknowledging the value of our work and your suggestions on how to further improve the manuscript.*

- Since the ratio between daily and sub-daily peaks does not refer to individual events, but to estimated quantiles, this should be stated clearly in the title/abstract of the paper. Why hasn't the analysis focused on individual events? With the data available this would have been possible and, in principle, would have allowed an even clearer attribution of the dampening to the reservoir/lakes, I think. Some clarification on this choice is needed.

Reply: *Thank you for highlighting the need to justify this choice. We indeed compared estimated flood quantiles rather than individual peaks. We added the following justification for this choice to the methods section:*

'To compare daily and sub-daily flood peaks, we work with flood estimates, that is flood quantiles estimated for specific return periods. We used such estimates instead of focusing on individual events as flood events detected in one time series are not necessarily detected in the other series.'

Furthermore, we adjusted both the title and abstract to better reflect this choice by highlighting that we are talking about 'estimated flood peaks'.

- The methodology section should provide some more details on the methods and the reasons for choosing them. For example I have not understood why the analysis of sensitivity of peak flows to the catchment area is performed through a log-linear model "without an intercept" (see detailed comments below). Also, it is unclear to me how Figure 5b and Figure 6 are obtained. Some math (not too much) would maybe help.

Reply: *Thank you for highlighting the need to clarify some methodological choices like the choice of a linear model without an intercept, the quantification of increase in node purity (shown in Figure 5b), and the partial prediction of the peak ratio (shown in Figure 6). We specified these choices in the methods section as follows:*

- *intercept: 'We chose a linear model with no intercept to force the regression line through the origin (0,0) because a catchment with a catchment area of 0 should also have a 10-yearly flow of 0.'*

- *increase in node impurity: 'To identify the most important characteristics, we used two metrics derived from the random forest model to quantify variable importance: The increase in mean squared error (% INC MSE) and the increase in node purity. [...]. The increase in node purity reflects a characteristic's importance in improving the decision tree's structure by quantifying the total decrease in impurity as quantified by the Gini*

index from splits using a characteristic (Liaw and Wiener 2002).'

- partial prediction of the peak ratio: 'Additionally, we use partial dependence plots (Friedman 2001 and Greenwell 2016) to estimate the marginal effect of the most important variables for predicting the D/H ratio. These plots help visualize the relationship between a subset of the predictors and the response while accounting for the average effect of the other predictors in the model. In doing so, they reveal the marginal relationship between the predicted variable and each predictor and can be used to assess the type of relation across the value range of each variable. We use these partial dependence plots to identify when the contributing area percentage has a strong influence on the D/H ratio.'

- As also stressed by another Reviewer, while the analyses done to compare the “area” and the “contributing area percentage” effects are certainly valuable and most important, it is strange that the other attributes, which from Figure 5 appear not that less-important, are then neglected in the analyses and discussions.

Reply: *We added the following justification for our focus on the two most important variables to the results section: 'Our further analyses focus on the influence of the two most important characteristics, namely, contributing area percentage and catchment area, because they seem to be the most important predictors of the D/H ratio related to the 10-yearly flood peak.' Furthermore, we expanded the discussion on the other two predictors considered: 'While catchment area and contributing area percentage are the most important predictors of the D/H ratio, other predictors such as the biogeographical region, a metric for catchment similarity in terms of climate, geology, and vegetation, and other catchment characteristics such as permeability, an indicator for flashiness, only have limited predictive power (Fig. 5). This highlights that existing estimation approaches deriving instantaneous peak flows from daily data using catchment area and other catchment characteristics can profit from additionally including information on water body influence.'*

Detailed comments:

Lines 85-90: The Walensee does not have a small difference in contributing areas between the upstream and downstream gauges. Why then not including other Swiss cases in the Case-Study analysis?

Reply: *While we would have loved to include additional or alternative Swiss case studies, we did not find any other cases fulfilling the criterion of having up- and downstream gauges with sub-daily streamflow records.*

Line 119: I don't understand the sentence “we excluded lakes which are part of other lake catchments”. Even the following example does not clarify the point.

Reply: *Thank you for highlighting the need for clarification. We rewrote the sentence to 'Since we are interested in the total lake-influenced catchment area, we only included the most downstream lake within a catchment.'*

Line 144: The ratio between daily and sub-daily peaks does not refer to individual events, but to estimated quantiles. This was not so clear in the previous part of the paper. Maybe the title or, at least, the abstract should include this information. Besides, it would be interesting to know in how many years the daily maxima and the subdaily maxima do not correspond to the same event.

Thank you for highlighting the need to highlight the focus on estimated quantiles. We indeed compared estimated flood quantiles rather than individual peaks. We added the following justification for this choice to the methods section:

‘To compare daily and sub-daily flood peaks, we work with flood estimates, that is flood quantiles estimated for specific return periods. We used such estimates instead of focusing on individual events as flood events detected in one time series are not necessarily detected in the other series.’

Furthermore, we adjusted both the title and abstract to better reflect this choice.

Line 160: How is “impurity” defined?

Reply: *We specified that ‘impurity’ was defined using the Gini index: ‘The increase in node purity reflects a characteristic’s importance in improving the decision tree’s structure by quantifying the total decrease in impurity as quantified by the Gini index from splits using a characteristic (Liaw and Wiener 2002).’*

Line 162: From the description here, I cannot grasp how the “partial dependence plots” work.

Reply: *Thank you for highlighting the need for clarification. PDPs help visualize the relationship between a subset of the predictors and the response while accounting for the average effect of the other predictors in the model. We specified this in the manuscript.*

Line 174: Why is it necessary to fit log-linear models without an intercept? Is the equation then $Q = A^{\beta}$? Why should Q be equal to 1 when the area is 1? The property declared in the following section, “The slope shows by how many % points extreme flows increase when the catchment size increases by 1 %”, also holds for the model $Q = \alpha A^{\beta}$, i.e., a log-linear model with the intercept. Maybe adding the equation would be beneficial. I am not a fan of a too mathematical treatment in hydrology papers but sometimes math helps the reader.

Reply: *By setting the intercept to 0, we force the model through the origin (0,0), thereby assuming that a catchment with 0 km² area should also have 0 flow. We justified this choice in the manuscript as follows: ‘We chose a linear model with no intercept to force the regression line through the origin (0,0) because a catchment with a catchment area of 0 should also have a 10-yearly flow of 0. The slope of this regression model shows by how many % points extreme flows increase when the catchment size increases by 1 %.’*

Figure 2: Is there a reason for scaling logarithmically the y-axis (flow estimates)?

Besides, data could be added here with a plotting position formula.

Reply: *We logarithmically scaled the y-axis to limit its vertical extent. We used a theoretical instead of an empirical distribution to derive flow estimates for specific return periods.*

Figure 4: Maybe 4 points corresponding to the downstream gauges in the Case-Study analysis could be added here as a reference. This would demonstrate that the two analyses are consistent (even though not geographically).

Reply: *Thank you for this suggestion. While we think that this is a nice idea, we refrained from doing this because only one of the case study catchments is located in Switzerland (see Figure 1 in the manuscript).*

Lines 222-225: This result is interesting. Even though not dominant, also the biogeo_region and permeability attributes seem important. Why are they neglected in the following analyses?

Reply: *We focused the subsequent analysis on the two most important predictors (contributing area percentage and catchment area) in order to disentangle the influence of reservoirs from the one of catchment area.*

Figure 6: As said above, I do not understand how these diagrams are obtained.

Reply: *Thank you for highlighting the need for clarification. PDPs help visualize the relationship between a subset of the predictors and the response while accounting for the average effect of the other predictors in the model. We specified this in the manuscript. We used the R-package pdp by Greenwell to construct the plots.*

Lines 239-245: I wonder if your experimental setup is “fair”. In Figure 7 the brown line is calibrated and the light-blue/green one is in extrapolation (if I understand it well).

Wouldn't be better to calibrate the model on some no-water-body influenced catchments and then apply it to both non-influenced and influenced catchments not used in calibration? I am confident that the result would be similar but not affected by the calibration issue. Or, if the Authors want to demonstrate that the “contributing area percentage” is necessary, wouldn't it be better to use all catchments for calibrating the model without that attribute? Probably, in that case, also the brown distribution would be biased but in the overestimation range.

Reply: *It is correct that the random forest model used to predict the results shown in Figure 7 was trained on weakly water-body influenced catchments only. What we are trying to show is that doing so will allow us to predict the D/H ratio in these catchments, while it will fail to do so in regulated catchments. This highlights that information on water body influence is crucial to predict the D/H ratio in strongly regulated catchments. We clarified the purpose of this analysis in the methods section of the manuscript: ‘This analysis aims to highlight the importance of water-body influence information for explaining the D/H ratio in regulated catchments.’ Calibrating the model on a subsample of weakly water-body influenced catchments only would have substantially reduced the amount of data available for model fitting, which is why we would like to avoid this.*

Random forest models avoid overfitting by measuring prediction error on an out-of-bag sample that is sampled randomly in each iteration of the model fitting process.

Figure 8: Indeed from the figure I assume that the model used is $Q = A^{\beta}$ (the lines meet at the point (1, 1)). It is not clear to me why this should be preferred to $Q = \alpha A^{\beta}$. Some clarification is needed in the method part.

Reply: *As explained above, we force the model through the origin (0,0) by setting the intercept to 0, thereby assuming that a catchment with 0 km² area should also have 0 flow. We justified this choice in the manuscript as follows: ‘We chose a linear model with no intercept to force the regression line through the origin (0,0) because a catchment with a catchment area of 0 should also have a 10-yearly flow of 0. The slope of this regression model shows by how many % points extreme flows increase when the catchment size increases by 1 %.’*