

## Authors' responses to the comments of Reviewer #1

We appreciate your review and comments on our manuscript, "[The Pluvial Flood Index \(PFI\): a new instrument for evaluating flash flood hazards and facilitating real-time warning](#)". Your feedback is valuable to us, and we made the recommended revisions accordingly. We provide detailed responses to each of your comments below.

### General comments

The authors introduce a novel index, the pluvial flood index (PFI), designed to assess and communicate the hazard potential of an area with respect to pluvial flooding. The PFI depends on the results hydrological and hydrodynamic simulations. It increases with the fraction of a reference area where thresholds for at least one of the following variables are exceeded: inundation depth, flow velocity, or specific surface runoff. The thresholds are chosen to represent safety for pedestrians and cars. Finally, the fraction is classified into four classes from "low" to "very large" hazard. This design is based on the idea that the index should be easily communicated to the public. The authors suggest to use the PFI for hazard forecasting and for the creation of hazard maps. Especially in times of climate change it is highly important to improve disaster risk management in regard to pluvial floods and I agree with the necessity to improve existing concepts. However, in my opinion, the manuscript sometimes fails to submit to the reader the distinction of the novel aspect of the PFI and the needed models which are technically exchangeable and already existing. PFI and the underlying models can be viewed completely separately. The novel aspect of this study is solely the use of three thresholds and the moving circular window as a reference area.

[We agree that part of the publication deals with the definition of the PFI and its calculations and definition of the thresholds. However, another very relevant aspect of the paper is the definition of the pluvial flood hazard areas \(PFHA\), a combination of the three most relevant hazards \(water depth, velocity and specific discharge\). Based on the valuable comments of you and the second reviewer we substantially changed and refined our manuscript. We hope that this becomes now clearer.](#)

Generally I wonder, if safety for pedestrians and cars should be the main indicator for pluvial flood hazard, because another main aspect of flood hazard is the damage on houses and infrastructure, which the authors do not mention and discuss.

[Damage of houses and infrastructure is certainly another highly relevant flood hazard, however, the loss of lives when pedestrians are being downed or captured in vehicles should have first priority. However, the defined thresholds for water depth and velocity match well the thresholds for damage data of buildings that has been derived and analyzed in a parallel project \(e.g. Singh et al., \(2025\) for private households or Guntu et al., \(2025\) for commercial buildings. We made these aspects clearer in the revised version](#)

Some parts of the manuscript describe accurately the fundamental hydrological processes that have to be represented in the models for a sound hazard estimation of pluvial floods, while other relevant aspects of the computation PFI lack some explanation. The PFI is sensitive to the parameters of the chosen circular buffer radius and the accumulation threshold. This part is missing in the "Discussion". The parameter "accumulation threshold" is never explained.

We addressed these points further below in the specific comments.

The manuscript is well written, the structure could be improved in some parts. After addressing following questions and points of concern I recommend this manuscript for publication.

The feedback to the other points is given in the specific comments below

I will refer to these previous points in the following detailed comments:

### **Specific comments**

Line 53: You are suggesting the PFI as part of forecast model chain. I assume that the bottleneck are the underlying hydrodynamic models. Are there hydrodynamic models that run as part of a forecast chain on a federal state level? How long is the lead time? Could you clarify this?

At the moment this is not the case, at least in Germany. But several studies have shown the potential of directly merging hydrological and hydrodynamic model to improve forecasting of pluvial and fluvial floods and due to the rapid improvements of calculation times of hydrodynamic models (e.g. GPU), the lead time to run these models for specific areas where the hydrological forecast exceeds a certain threshold of overland flow, the lead times should be in the order of couple of minutes. We included this aspect more prominent now in the reformulated introduction

Line 70-105 ff, Relevant Processes: This part accurately describes the processes that a model needs to represent to compute the two variables needed for the computation of the PFI (flow velocity and inundation depth). As the models are technically interchangeable and already existent, I wonder if such a detailed description of processes is necessary. Maybe this section could be shortened.

We wanted to make sure that all readers, independently if they are hydrologists or not, are on the same level. Following the remarks of reviewer 2 on this, we actually had to extend the section on the discussion of the role of interflow. In order to improve the structure, we moved the whole discussion in a separate section on relevant processes controlling the generation of pluvial floods. With this, it should also become clearer that these aspects are not part of the PFHA/PFI but a prerequisite, in order to be able that a meaningful pluvial flood hazard analysis can be conducted.

Line 101 ff.: Does the model you use consider geomorphological processes? Please clarify. No, it does not as nearly all models used for pluvial floods. But we wanted to address the general need for such models. However, we made it more clear now, that also AccRo does not include these processes.

Line 112: "A hazard is defined ..." Should it not be "We define a hazard ..."?

Otherwise please give a reference.

Yes, definitely. We changed it.

Line 113: Why did you base the PFI solely on safety for pedestrians and cars? In situation where vehicles float I could imagine graver impacts such as flooded houses and basements with a possible threat to life. Or are these two measures solely an indicator for the overall hazard? Furthermore, I wonder if there could even be a situation where cars are swimming while pedestrians and houses are still safe? Why this differentiation and focus on cars? Otherwise this section could be also shortened.

See also above. The PFHA are defined in order to include both measures, the water depth is more sensitive for floating of vehicles, the velocity and specific discharge more related to safety of pedestrians. We will make this clear in a revised version. To accurately predict the flooding of houses requires detailed knowledge about the structure and building of the houses (windows, garage, etc), which is usually not available at the larger scale. Hence, these details cannot be included in the PFHA but the overall defined threshold match well the damages of houses for observed pluvial events. This should become clearer now in the revised version of the manuscript.

Figure 2: These two figures violate almost all principles for scientific presentations: the two figures are too small, the font size should be increased, they are not color blind friendly and the legend for the different symbols is missing, the grid is plotted over the annotations. A lot of information is missing in the caption (e.g. yellow box). This figure opens more questions than it answers and is containing information that is eventually not needed. Generally, this Figure is more confusing than helping, I think. Since you decided to just use to fixed values as thresholds for the PFI, you could think about dropping this low-quality figure entirely.

We absolutely agree. However, we build the figure based on already existing figures (Martínez-Gomariz et al. 2016; 2018), which already violate the principles. Therefore, we tried to improve the figures by replacing some labels and increasing the fonts, but this was not done for all, as the figure is not available as a vector file. To overcome this we rebuild the figure and also removed some details less relevant for the PFI context

Line 135: Why do you use the unit  $\text{m}^3 / (\text{m} * \text{s})$  and not just use the reduced unit  $\text{m}^2/\text{s}$ ? Where do the extra meters come from? Same for Fig 4c).

We used this unit to make it clear to the reader, that it refers to specific discharge that has to be divided by the cross-sectional width. Knowing my students, I would assume that they are unable to understand the reduced units directly, so I thought it is helpful. We now derived it once and then used the reduced units afterwards.

Line 161: You are saying that radar QPE can be used as an input, but at the same time you refer previously to the small spatial scale of the triggering precipitation events. In my opinion, the model results and the index can only be meaningful if high resolution rainfall data is used as input, which is up to now only possible with weather radar because of the low density of the rain gauge network. This should be pointed out explicitly.

Maybe we should have defined QPE in this context. We were referring to weather radar products that are improved with rain gauge data and high-resolution weather models (resulting in spatial resolutions of 500-1000m and spatial resolution of 5 min) as it is already operationally in use at meteoswiss or currently under development and the German

Weather Service and probably at many more. To make this clearer we added a data section in the methods chapter where we describe the radar based QPE in more detail.

Line 165: “weighted by radius” This becomes clear in Fig. 3c) but you should explain this a little more here by adding something like “weighted by the radius of the circular buffer to focus more on the areas close to the buffer center (see Fig. 3c)” or something similar.

Yes, we extended it.

Line 178: “based on experiences” Can you give some examples or insights from your experience here? My gut feeling also agrees with the proposed values but maybe you should just delete this statement or give some more concrete examples. See following comment to line 194.

As proposed, we deleted this statement in the revised version.

Line 184-187: “To ensure ....” This could go to the discussion. Yes, we changed it.

Line 194: This refers to my general comment, that some essential aspects of the PFI are not explained. I agree, that defining thresholds is always, at least a bit, arbitrary. However, “defined as accurately as possible” is not a sufficient description. What was the workflow when you defined it as accurately as possible?

Maybe you could already explain here why you chose these thresholds instead of just doing it at the beginning of the “Discussion” section.

We followed the suggestion and moved the evaluation study of the PFI thresholds now to the PFI definition section.

Additionally I wonder, if the thresholds are a bit low: A 2 km<sup>2</sup> circular buffer has a diameter of ~ 1.6 km. On the edges areas would be classified as hazardous because of a PFHA more than a kilometer away. In regard to communication and safety it is surely better to pick a larger than a too small area of the circular buffer to conceal uncertainties, which can be difficult to communicate. What was your motivation behind deciding for 2 km<sup>2</sup>? Why not simply use the PFHA?

The motivation was to find a compromise between your thoughts and the thoughts of reviewer #2 suggesting a much smaller area to avoid showing higher PFI on hillslopes. In general, the presented approach to visualize the PFHA is one possibility to do so and we will make it clearer, that there are certainly several other approaches, as for example directly mapping the PFHA (as we also do – see example Figure 4). However, the PFHA cannot be seen when mapping large areas, so either the PFI could be calculated based on other units (catchments, political boundaries, etc.), but as discussed, this would always change the PFI as the related area is different. Therefore, we developed our approach in close cooperation with many potential users and optimized it in several stakeholder workshops – we added now more thought on why it was developed as presented. We hope it becomes clearer now.

Line 222: AccRo: You are mentioning this model various times but the reference Leistert et al. 2025 is insufficient. I could not find anything. Can you give a DOI, or a URL? Same in line 340.

Meanwhile a preprint of the model description is available and referenced in the manuscript.

Line 232: Radar QPE: Why is there no description about the product in the “Data” section? The rainfall product has a major impact on the outcome of your hazard assessment. To my knowledge there is no radar product with a spatial resolution of 250m in Germany. Or is it a composite product?

As stated above we added a brief description about the Radar product in section 3 (previously 2). We used the data from the Kachelmann Group, which provides QPE in 250m x 250m resolution for recent years (since 2020)

Fig. 4: a) The fonts of the rain gauge labels is a bit too small.

We adapted the font size

e) Unit of the specific surface runoff. See comment to line 135.

Subplot top right: remove label “Fig 4”, little grey box not explained in caption but much later in text.

We added a more explicit label to explain the box and also included it in the capture

g) “resulting extent PFHA” should be labeled with f)

We adapted this - and also a few other wrong assignments of letters.

Line 259: You mention the accumulation threshold here for the first time. Do you refer to the accumulation threshold your underlying model uses to define streams and rivers? Figure 6 shows that the PFI is sensitive to this parameter. Please clarify what you mean with accumulation threshold and discuss why you chose the values you chose.

Thank you for pointing this out. The idea behind the threshold is that when using design storm events as input (in our case a 100yr event), we have widespread heavy rainfall input which would lead to unrealistically high flooding in the downstream areas. In reality, convective rainfall events have a certain spatial extent only, somehow limiting the amount of water available for pluvial flooding. In order to mimic this, a maximum accumulation area is introduced after which it is assumed that the capacity of the river is large enough to capture the accumulated flow. We added this now when describing the design storm data in the new Data section.

Line 261: The PFI is using thresholds that are aimed at the safety of humans and cars so it is not surprising that damages on infrastructure are not as well reflected by the PFI. However, this is a major impact and the focus on cars and pedestrians slightly diminishes the informativeness of the PFI. Maybe you should also discuss this. This is related to a previous comment and is also tackled by our answer which we give above.

Line 269: Please provide an English translation for the non-german readers.

Yes, we provided a translation to it

Line 270 ff.: Because of the small extent of the rainfall events which can trigger a pluvial flood, I wonder how well these small and rare events are represented in the underlying statistics of a 100-yr event. Your approach using events from CatRaRE seems to make much

more sense. Maybe you could discuss this, because the rainfall input is, besides the PFI, the other major factor for the creation of a flood hazard map.

Thank you for pointing this out, so we can make this clearer. Actually, there are two different approaches we aim for in the application of the PFI. First is that we calculate the PFI for real events – also to show the potential for forecasting. In this case we would use radar QPE (like they are reflected in CatRaRe). Second case is the general susceptibility for pluvial floods. In this case we need to work with design rainfall events. The 100yr design rainfall event we use is based on the extreme value analysis of 1 hour precipitation events recorded at stations and then spatially interpolated. In this case the data is available for the full state of BaWü at a spatial resolution of 1x1km and was the standard data for the pluvial risk management in the state of BaWü in the first version. Alternatively, one could use the KOSTRA dataset (5x5 km resolution). We also included these aspects in the newly introduced data section. The comparison with CatRaRe in the figure was mainly to highlight the fact that there are certainly much more heavy rainfall events all over the place than we had recorded pluvial flood events. We tried to make this clearer now in the manuscript.

Line 321 ff: “The PFI is more advanced, as it ...” I would change this sentence because the PFI could be calculated with any kind of velocity or inundation maps, regardless of the quality of these maps. Then the PFI would not be very reliable and would not encompass all relevant aspects. Here a clear distinction should be made: the PFI is one thing, the underlying models are something else. You can still compute the PFI based on unsuited models and then it will have little informative value.

You are certainly right. We adapted the section thereby highlighting that a sound data basis is needed to allow for a meaningful flood analysis.

Line 335: “The PFI can serve as a fundamental approach...” Relating to the previous point I find this again a bit “high-pitched”. The foundation for a large-scale warning system would be a reliable model, not the PFI. The PFI is “only” aggregating the grids resulting from the models.

Again, you are certainly right. We adapted the section now stating that the PFI can serve as key information in order to provide tailored information. Based on its aggregation concept, the PFI can provide a large-scale pluvial flood warning not yet present.

Line 349: Please translate German terms to English. Usually you would give a reference and add the URL in the references.

We changed it accordingly

Line 348-359: This description of the “State-of-the-art” belongs to the introduction in my opinion. Line 359: Please describe in more depth, which accumulation threshold you are referring to.

We moved the description of the ‘Hinweiskarte Starkregengefahren’ as well as the example from Switzerland to the introduction: Nevertheless both concepts have in our understanding some substantial shortcomings, we but keep the comparison with the PFI in the discussion – making it more clear by adding a Figure depicting the impact on the PFI if one would follow

the concept suggested by the ‘Hinweiskarte Starkregengefahren’ as well as the example from Switzerland.

Line 382: “The PFI considers all hydrological and hydraulic factors...” Again, I think that this formulation is misleading and refer to my previous comments to line 321 and 335.

Yes, we reformulated the sentence.

Technical comments:

Thank you for pointing them out. We Corrected them

Line 148: “capture the hazards of a flash floods”

Fig. 5: b) should be “E/extent”

Line 402: Reference “Apel ...” occurs twice.

Line 420: Reference format error.

Line 437: Reference format error.

## Authors’ responses to the comments of Reviewer #2

We appreciate your review and comments on our manuscript, “**The Pluvial Flood Index (PFI): a new instrument for evaluating flash flood hazards and facilitating real-time warning**”. Your feedback is valuable to us, and we made the recommended revisions accordingly. We provide detailed responses to each of your comments below.

This article introduces a novel approach to assessing and mapping pluvial (flash) flood risk, referred to as the Pluvial Flood Index (PFI). The proposed method is demonstrated through two case studies: first, a hindcast of a flash flood that occurred on June 2, 2024, in the Wieslauf catchment in Germany; and second, the generation of flood hazard maps for a region of approximately 1,000 km<sup>2</sup> in the southwest of Baden-Württemberg. In the latter case, the reliability of the PFI maps is evaluated using a database of flash flood incidents that caused damage to infrastructure and roads. This database, originally compiled by the state authorities of Baden-Württemberg and extended by the authors, covers events from 1995 to 2024.

The proposed approach, along with its underlying motivation and evaluation, may initially appear appealing; however, upon closer examination, it raises several concerns.

First, the method does not actually generate new information but rather involves a post-processing of standard runoff maps produced using a combination of a rainfall-runoff model and a 2D hydraulic model. This post-processing is carried out in two steps. In the first step, a Boolean variable called PFHA (Pluvial Flood Hazard Area) is assigned to each pixel based on predefined flood depth and velocity thresholds, categorizing them as either low or high hazard. In the second step, a spatial smoothing technique is applied to calculate, for each pixel, a Pluvial Flood Index (PFI), which reflects the proportion of high-hazard pixels within a circular buffer surrounding that pixel. Four proportion ranges are defined, corresponding to four levels of PFI. However, this spatial smoothing—central to the PFI concept—is not adequately justified, and the choice of the circular buffer size (2 km<sup>2</sup>) remains unexplained.

It is understandable that the initial motivation behind the development of the PFI method was to enhance the readability of flood risk maps at a large spatial scale. In the two case studies presented, the “Pluvial Flood Hazard Areas” (PFHA) are mostly confined to riverbeds and broader floodplains, making them difficult to discern on large-scale maps. To address this, the proposed PFI method uses spatial smoothing to create broader, more visually prominent zones around clusters of high-hazard PFHA pixels. While this improves map legibility, it is crucial to recognize that spatial smoothing is a digital artifact. The resulting PFI levels are not solely the product of hydrological and hydraulic modeling—they are shaped by additional processing choices and assumptions, and should therefore not be overinterpreted.

Unfortunately, the authors appear to let their enthusiasm override critical assessment. One must ask: is there any sound reason to believe that hillslopes located in the surroundings of inundated floodplains are more exposed to pluvial flooding than hillslopes located elsewhere? A careless reading of the PFI map might suggest so, but such a conclusion lacks physical basis. Proximity to a floodplain does not inherently increase a hillslope’s vulnerability to pluvial flooding.

In short, the PFI method should be presented for what it truly is: a straightforward graphical tool designed to highlight clusters of high flood hazard at the regional scale, not a physically grounded index of flood exposure.

Thank you for this critical assessment of the general PFHA and PFI methodology.

We agree that the PFHA and PFI could be considered as a simple post processing of the output of a hydrological and 2D-hydraulic model chain and could be applied to any rasterized, large-scale data of maximum water depth, maximum flow velocity and maximum specific discharge – if these variables are available. We reformulated the manuscript at several spots to make clearer what actually defines PFHA/PFI regions but also on what is required to get a meaningful estimation of pluvial flood hazards

With respect of the location of the PFHA we added a detailed comparison with fluvial flood estimates in the discussion section. As we originally did not include the location of rivers in Fig 5a, it looks like that PFHA are only confined to riverbeds and broader floodplains, actually already defined as flood prone areas by fluvial flood estimates. But this is for many regions not the case, but the PFHA/PFI covers areas aside of river beds as well. We will come back to this in more detail at the end of our replies.

One point we made clearer in the revised version of the manuscript is the spatial scale the PFI aims at. We now included a section on this directly at the beginning of the methods section. We further restructured the method section so that the actual target of the PFI should now become more clear.

Regarding the underlying hydrological and hydraulic processes, we still think that this is relevant to be included in the manuscript, since a meaningful PFHA and related PFI (for both, operational forecasting and identifying general pluvial flood susceptibility) requires the sound modelling of hydrological and hydraulic process. As already discussed, these processes are often not represented in current developments (e.g. ‘Hinweiskarte Starkregen’) implemented at state or even federal level. However, to make it clearer that this is not a direct part of the PFI but a prerequisite, we moved the respective section outside of the PFI definition section into a new chapter 2.

Second, the evaluation procedure is insufficiently described and critically discussed. In Figure 6, past damaging flood events are depicted as dots. However, flood events have spatial extents— especially those occurring in catchments with upstream drainage areas of 10 to 20 km<sup>2</sup>. How was the representative location of each flood event determined for the evaluation? It appears likely that the dots indicate the locations where the most significant damages occurred, typically along riverbeds and, in many cases, within flood-prone plains. If this is the case, then the apparent skill of the flood risk assessment may not stem from the PFI method itself, but rather from the underlying rainfall-runoff and hydraulic simulation models that generate the base maps.

We agree that it would be beneficial to have better damage data to verify the PFI. However, there is hardly any information available in the case of pluvial flood events at the scale of interest. For a few villages, we have some more explicit spatial information, but this is typically related only to the location of emergency calls during the flooding. The database which is available for the state of BW and included in Fig. 6 is more or less a binary data set, indicating if in a certain community there was a damage following a heavy rainfall event or not. Hence the dots simply represent communities where there was a damage. We made this clearer in the revised version of the manuscript. Since the quality of the PFHA/PFI depends on the quality of the models used to estimate the parameters (maximum water depth, maximum flow velocity and maximum specific discharge) to define hazard areas, it is simply impossible to disentangle the quality of the PFI from the quality of the underlying model data. We made this clearer in the revised version.

For the Wieslauf example, we collected some data on maximum inundation – but of course only for a few locations and after a few days after the catastrophic event. Although this data does not at all allow to identify all spots affected by the pluvial flood, we use it to verify that PFHA/PFI structures somehow are reflecting inundated areas. We added a respective section on the chapter describing the flash flood event of 2024. We also included a Figure in the supplementary material showing simulated max water depth as well as PFHA and PFI distribution for the two hotspot regions for which we have ‘observed’ data available.

Moreover, although the article is ostensibly focused on pluvial flooding, the evaluation is confined to areas with catchment sizes larger than 10 to 20 km<sup>2</sup>, which more closely align with riverine or flash flood events rather than true pluvial floods. This choice contradicts the stated objective of the manuscript. What proportion of the DWD-CatRaRe database was excluded due to this threshold? It would be helpful to visualize all flood events from the database in Figure 6, using a distinct color to differentiate those included in the evaluation. Additionally, the relationship between the recorded heavy precipitation events (Figure 6a) and the damaging flood events (Figure 6b) is unclear—some commentary on their correspondence is needed.

It seems that this is a misunderstanding which probably arises from the fact that we did not describe the accumulation threshold and its motivation in sufficient detail. The PFI aims at large scale, however, in order to reflect the spatial scale of convective precipitation events and the typical catchment areas related to pluvial floods, we limit the accumulation area of generated flows to 10km<sup>2</sup> and 20km<sup>2</sup>, respectively. In the revised version we described the need for the thresholds in detail in the model and data section of the method chapter. We hope this becomes clearer now, So we do not consider any catchment areas larger than 10 to 20km<sup>2</sup>, depending on the chosen threshold. Hence we also did not exclude any convective, short and small scale events from CatRaRe, but the opposite – we excluded all events with a duration of more than 4h, which can be assumed to be rather of frontal nature. We also did not exclude events from the ‘damage’ database, but used all data that was reported – definitely knowing that it is incomplete.

The comparison between heavy precipitation events (Figure 6a) and the damaging flood events (Figure 6b) was included to show that there were many more heavy rain events (actually more or less all over the study region as can be expected) compared to the damage events (which of course are somehow linked to the valleys or hillslopes – however not to all valleys) – so basically this shows the added value of a PFI forecast with respect to a solely precipitation forecast. We reformulated the section so that this hopefully becomes clearer.

Further concerns arise in Section 2.1, “Relevant Processes.” The section implies that pluvial runoff and flash floods are primarily caused by direct overland flow due to either saturation or infiltration excess. This is an outdated and overly simplistic view that has long been challenged in hydrological literature. Decades of research into hillslope processes and flash flood generation have demonstrated the complexity and spatial-temporal variability of runoff mechanisms, including the often-dominant role of subsurface flows (interflows). Such simplification is not appropriate for publication in an international journal and should be reconsidered. Since detailed discussion of rainfall-runoff processes is not central to the manuscript’s main contribution, it may be best to remove or significantly revise this section.

We agree that interflow can play an important role but mainly for events with longer duration. For short events under typically dries soil moisture conditions, interflow is quite unlikely in our regions, which we can show with a larger body of observational data. Therefore, we wanted to make the case of the relevance of saturation or infiltration excess overland flow for the generation of pluvial floods. Numerous sprinkling experiments revealed the importance of excess infiltration in the case of short-term high-intensity heavy rainfall events representing the main cause for pluvial flooding (e.g. Ries, F., Kirn, L., & Weiler, M. (2020). Runoff reaction from extreme rainfall events on natural hillslopes: a data set from 132 large-scale sprinkling experiments in south-western Germany. *Earth System Science Data*, 12(1), 245-255).

As mentioned previously we think that the correct representation of the dominant runoff generation processes is key for a meaningful estimation of the PFI. Hence, we suggest to leave the section in but as a separate chapter.

In conclusion, the article presents a potentially useful approach. However, the PFI method should not be overstated. It should be honestly described for what it is: a simple graphical tool aimed at highlighting clusters of flood hazard on regional maps. If the primary focus is on pluvial flooding, then there is no justification for excluding events in small upstream catchments—especially when past records of such events exist. While including these areas may lead to less favorable evaluation outcomes, such results would be highly relevant and valuable for the community working on pluvial flood risk mapping. If the current threshold is retained, then the manuscript should clearly state that its focus lies with riverine or flash floods rather than pluvial floods.

We restructured the manuscript to make it clearer what the PFI method is based on and for what it designed, but also what the prerequisites are to have a sound estimate of pluvial floods.

We also reformulate section 3.2 (4.2 in the revised version) so it becomes clearer that we did not exclude any small upstream catchments but that the focus is explicitly on including them.

Finally, the discussion and analysis need to be further developed. It appears that the PFI and PFHA indices primarily identify flood-prone river plains. If this is indeed the case, it should be explicitly acknowledged and critically considered.

We disagree that PFHA are primarily linked to river plains (see also response above), but some valley (with and without rivers!) are certainly included, which of course makes sense, since water usually accumulates not at hill tops. However, if one examines the spatial PFI patterns in Fig 6 closely, one can see that not all river-plains are characterized per se as high hazard areas but that regional differences due to runoff generation processes as well as the 2d hydrodynamic flow accumulation are present. If a simple topographical index would be used, the regional differences would be much less pronounced than in the case of the PFI.

To make this clearer we added a comparison between PFHA and areas flooded by an HQ100 flood event (see Figure 7 as well as Fig S2 in the supplementary material) in the manuscript. HQ100 areas are typically confined to major river plains. So if the PFHA/PFI would primarily identify flood prone river plains, PFHA should mainly occur within the HQ100 zones. However, this is actually not the case, since our analysis shows that the major portion of the PFHA is actually outside of HQ100 areas. This can be explained by the fact that the PFHA is mainly designed for the smaller valleys at the higher elevated parts of catchments, while hq100 areas usually are more confined to the flood plains in the lower parts of the catchments.

We also would have completely different results if we would not consider the runoff generation processes but simply use rainfall estimates as input for the 2d-hydrodynamic flow accumulation – which as already discussed, are still represented in many current developments (e.g. ‘Hinweiskarte Starkregen’). Since this was also pointed out by Reviewer 1, we included an analysis on the impacts of ignoring hydrological processes on the PFI (see Figure 8), demonstrating the importance of including the hydrological processes in the data preparation for pluvial flood analysis. We think that this finding is also a justification for leaving the section on the hydrological processes in the manuscript.