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 will make 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 is 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).

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.

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.

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 indundation 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. However, the sections could certainly be streamlined.

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.

Line 112: "A hazard is defined …" Should it not be "We define a hazard …"? Otherwise please give a reference.
Yes, will be changed.

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.

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. We will try to get the original data and figure form the authors of the two publications to improve our figure, but the problems rely unfortunately in the publication in the past.

Line 135: Why do you use the unit m³ / (m * s) and not just use the reduced unit m²/s? Where do the extra meters come from? Same for Fig 4c).

We used this unit to make it clear to the reader, the is 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. But we could derive it once and then use 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-

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, will be extended.

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.

More details form various example of already existing heavy rainfall hazard maps in the state of Baden-Württemberg can be provided to make this more concrete.

Line 184-187: "To ensure …." This could go to the discussion.

Yes, will be changed.

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.
Good suggestions – we will do.

Additionally I wonder, if the thresholds are a bit low: A 2 km² 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 km²? 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 in it several stakeholder workshops – but we will put a bit more emphasize on the decisions and why it was developed like presented in the paper.

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.
We are in the process of preparing a publication for the model – hopefully it will be available as pre-print beginning of September 2025. Meanwhile we will provide a link for a conference contribution: https://uni-freiburg.de/unr-hydro/wp-content/uploads/sites/94/406-TdH-2025-Leistert-Haensler-Schmit-Steinbrich-Weiler_dinA4-300dpi.pdf

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?
We will add a brief description about the Radar product in section 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 will adapt 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 will remove the label and explain the box.
g) "resulting extent PFHA" should be labeled with f)
We will adapt this

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. We will include a section on the accumulation threshold earlier when we explain the concept of PFHA and PFI. 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.

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. See also above. We will discuss this more in detail.

Line 269: Please provide an English translation for the non-german readers. Yes, will be provided

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).
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 then we had recorded pluvial flood events. We will make this clearer in the manuscript.

Lne 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 will adapt the section.

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 will adapt the section pointing out the need for reliable model simulations.

But still, the PFI - based on its aggregation concept 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.
Yes, will be 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 will move the description of the 'Hinweiskarte Starkregengefahren' as well as the example from Switzerland to the introduction but keep the comparison with the PFI in the discussion

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.
We will reformulate the sentence.

Technical comments:
Thank you for pointing them out. We will correct 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.