

Overall comment:

The study aims at reviewing the methods, contents and datasets of dynamic vulnerability assessments to floods, while basing on the previous conceptualizations of vulnerability dynamics. While the study offers some interesting insights in terms of methodological development, it has two major drawbacks: 1) unjustified methodological choices, that led to a small and possibly very limited sample, and 2) novelty – while the study points out gaps, and claims to provide “roadmap for advancing more robust and dynamic flood vulnerability assessments”, it stops short of that and focuses mainly on reiterating what has or has not been done. In sum, the study could be worthy of publication if it a) fulfilled a proper systematic search and review strategy, which is in this case doable and warranted; b) reviewed an exhaustive sample of papers, c) produced a bit more interesting contribution beyond gaps.

Response: We want to thank the reviewer for the extensive review provided. We appreciate the time this reviewer took to offer their reflection on the paper, and we intend to use this review to sharpen the paper's scope and address concerns about its scientific relevance. As we understand, the main critiques relate to a mismatch between the planned scope of the article and the methodological choices that were made, along with the reviewer's confusion about the objective of the article stemming from an insufficient elaboration on said scope and the key terminology around vulnerability and dynamics. We understand the reviewer's concerns and have reflected on several options that we hope help to address them. Since sharing the first version of the rebuttal, the author-team has worked hard on revising the manuscript in response to the feedback by the reviewers. As a result, along with the revised manuscript, we are also sharing an updated version of the rebuttal. Changes to the initial response are highlighted in orange text color.

We acknowledge the ambiguity in the introduced concepts and categories used for the analysis. We intend to better contextualize this research by defining the term of vulnerability and articulating how we are using it in our review, along with the term ‘vulnerability dynamics’. Additionally, we use the 28 studies that formed the main body of the first version of this review to test the categories we apply in the study and justify the categorizations.

Second, we acknowledge that despite our initial claim, the research conducted cannot be called systematic. We discussed options to redo our search by including Scopus and/or Web of Science. However, the original search terms do not return sufficient substance on either search platform, and would necessitate an entirely different strategy for querying these databases and thus require starting from scratch. While we acknowledge that the paper requires significant revisions, we do not believe that re-structuring our entire literature search is necessary to address the reviewer's comments. We decided to complement the set of initially identified studies through a targeted search on scopus but still without the claim or purpose of conducting a systematic review. Instead, we propose that we would more clearly articulate that the purpose of this manuscript is to showcase methods that have been or could be used to assess dynamic vulnerability in the context of flood-related hazards and their impacts. More specifically, we investigate how conventional vulnerability assessment methods (e.g., social vulnerability indices, physics- and process-based models, and statistical and narrative-based methods) can be used or tailored to assess vulnerability dynamics. Instead of using a set of vulnerability categories (as done in the original version), we intend to reorganize the existing material to focus on the methods themselves. We specifically investigate what data are used, which (sub-)dimensions of vulnerability are considered, how these methods capture temporal dimensions of vulnerability dynamics, what type of dynamics are captured, and what key limitations are mentioned by the studies, if any. In restructuring the

manuscript to achieve this goal, we can leverage all of the papers we have already collected and reviewed as part of our initial submission.

Third, we intend to discuss whether there are any patterns visible regarding specific causes or situations that result in vulnerability dynamics. This should offer some evidence to reflect on possible ways to categorize vulnerability dynamics, also in the light of already proposed conventions (e.g. de Ruiter & Van Loon, 2023). We do not want to create one based on the evidence we collect, but hope it could be a starting point for future attempts to come up with a relevant categorization of different vulnerability dynamics. Lastly, we also want to articulate limitations of the methods introduced in the papers that we included in our review.

We believe such an analysis of the existing literature still has significance, even if it is not structured as a systematic review. As the reviewer noted, vulnerability research is vast and (partly) ambiguous. Vulnerability dynamics are a subject of interest at the moment, but there are limited insights into how to best assess vulnerability dynamics and their contribution to risk profiles or outcomes. With our review, we want to showcase the existing research in this area, focusing on what has been used/produced in relation to the flood (disaster) risk management community.

We are convinced that the data we have collected offers a promising starting point for providing relevant insights to the community. However, acknowledging the methodological limitations of our current approach, we apply a targeted literature review to identify relevant studies across different methodological frameworks. We intend to use a similar approach to what has been done by Ward et al. (2020; <https://doi.org/10.1016/j.wasec.2020.100070>) and di Angeli et al. (2022; <https://doi.org/10.1016/j.ijdr.2022.102829>), which are just two of many examples of influential research using targeted (non-systematic) literature reviews. With these changes to our analysis and intended scope, we offer a starting point for further investigations from within the flood risk management community, which can be linked to investigations from other hazard communities (e.g., fire, drought) in the future.

Below, we include line-by-line responses to the reviewer's comments.

Comments in more detail:

- 1) The gap that the study addresses could be articulated more clearly – p. 2 lines 45-50 – the references to vulnerability of what and “how” and “why” could be opened up. If the authors refer to the methods (“how?”) then Jurgilevich et al 2017 review covers that, in addition to what. I’m not sure what authors understand as to “why”.

Response: We agree with the reviewer that this scoping paragraph is not very clear. The main framing for the scoping paragraph is that there have been reviews that focus on specific methods to investigate pre-, post-event changes (Moreira et al. 2021) or investigate what subdimensions of social vulnerability are outcomes and pre-conditions for multi-hazard events (Drakes & Tate, 2022), but there is a lack of a general overview of the methods used to investigate vulnerability dynamics. This paper aims to showcase and discuss how different conventional vulnerability assessment methods can be used/tailored to assess dynamics in vulnerability. Per Method, we investigate what data (types) are used, which (sub-)dimensions of vulnerability are considered, what temporal coverage is possible (resolution and time span), what changes in vulnerability can be reported, to which causes/situations these methods are applied and what limitations are reported regarding the method.

- 2) It would be beneficial if the authors could explain as to why we need to understand/assess V dynamics from the perspective of multi, cascading and aggravating hazards. It is somewhat articulated that vulnerabilities can be “interacting”, but more tangible substantiating examples would be beneficial. Also, a lot of vulnerability indicators, drivers or dimensions are the same for several hazards (e.g., typical indicators such as age, income, housing type, education level are relevant to consider for floods as well as for heat-related events, storms and others), so why do we need to account for them separately in e.g. cases of consecutive events?

Response: We added an additional section which builds on the short introduction of the concept of vulnerability in the introduction and offer more substance on the pluralistic understanding of vulnerability to introduce and justify the terminology and focus we apply in this study (see response to comment Method 1).

- 3) Is the overall rationale that vulnerability is also driven by impact, so vulnerability is dynamic as a result of a hazard in addition to its own inherent dynamics? Isn't this what is called dynamic risk?

Response: We would not claim that the overall rationale is that vulnerability is driven by impacts alone; there are more causes that influence the vulnerability. Underlying socioeconomic developments, for example (e.g., how residents use left-over money to either invest in DRM or buy beautiful property at the seashore at risk of seasonal flooding), have effects on vulnerability as well. Similarly, interactions between hazards and respective hazard-related impact drivers can also influence the vulnerability (e.g., a person might be able to withstand a certain strong wind when walking on the sidewalk, or they could wade through knee-deep water, but they might not be able to withstand a combination of the two).

- 4) Line 55 – if the study points out gaps and provides synthesis – it is not a roadmap. The actual contribution of the paper stops at synthesizing gaps and advances

Response: We agree with the reviewer and have reframed the key outputs of this study following this substantive review process.

- 5) Main criticism concerns methods- I do not consider methodological choices of the authors justified enough not to adhere to the protocol of systematic review. The research question of the study warrants a systematic review, and the field is homogenous enough to pursue it (as previous reviews have done successfully). The search keywords (flood, vulnerability) position the sample firmly into risk and adaptation literature, thus there is little challenge of dealing with the definition of vulnerability from epidemiology for example. In this vein, the justification of following a semi-systematic review is weak.

Response: In line with the feedback from the other reviewer and in response to other comments by this reviewer (see our response to comment 6) for example), we decided to change the methodological approach and scope for this paper. Instead of doing a (semi-)systematic or comprehensive study, we aim for an showcase of how conventional vulnerability assessment methods (curves, indicators, process-based, statistical, qualitative) can be used/tailored to assess vulnerability dynamics.

We use the search query on Google Scholar (as done in the current version) to identify a starting set of papers that helped refine the key method categories to use, data type groups etc. We then use a

targeted literature review per method to investigate relevant studies that offer more insights into how these methods can be applied to assess vulnerability dynamics.

The approach will not be systematic but a showcase, using similar approaches to influential studies like Ward et al. (2020; <https://doi.org/10.1016/j.wasec.2020.100070>) or di Angeli et al. (2022; <https://doi.org/10.1016/j.ijdr.2022.102829>). We thus believe that the findings can still be relevant to the community.

- 6) Google Scholar is not an appropriate search engine as it is guided by algorithms and previous user history.

Response: We thank the reviewer for this crucial comment. It was an honest mistake to base the review on Google Scholar, as we were unaware of this limitation. It's also perhaps relevant to acknowledge that this work started as a spin-off from a larger paper effort, which aims to develop an overview of how vulnerability dynamics are considered in different hazard disciplines (flood hazard research as one of them). We've reflected quite a bit on how to move forward with this paper under review, as the applied search terms as used in this review would return no hits when applied to Scopus. Still, we agree with the critical feedback from the reviewer (see our previous response under 5)).

- 7) Furthermore, the authors have a very limited search sequence. For example, search for risk assessments could yield more suitable papers, as these which often contain the assessment of vulnerability. Additionally, vulnerability dynamics may be an established term in a very niche theme of multi-hazard research, but it is not a well-established term overall, and there are plenty of studies that do relevant things but do not call it vulnerability dynamics. Overall, the search sequence limits the sample in many ways.

Response: We are convinced that an analysis (refined in line with the other comments from the reviewers) of the methods we have identified in this search can offer relevant insights for researchers who want to further advance the research on vulnerability dynamics. We have doubts that completely different patterns would emerge from this additional/different set of publications. We believe that there are always blind spots and limits to any literature review. As a result, we are reluctant to disregard the effort the co-author team put into reviewing 900+ papers and start from scratch with an alternative set of search terms. We preliminarily tested a search query on Scopus that returns a manageable amount of papers for a review: TITLE-ABS-KEY ({vulnerability NEAR/3 socio?economic} OR {vulnerability NEAR/3 social} OR {vulnerability NEAR/3 physical} OR {vulnerability NEAR/3 assessment} OR (vulnerability NEAR/3 analysis) OR {coping capacity} OR {adaptive capacity} OR {preparedness}) AND TITLE-ABS-KEY (flood* OR inundation) AND TITLE-ABS-KEY (dynamic* OR compound* OR cascading OR {multi?hazard} OR {paired event?} OR {temporal change*} OR {spatio?temporal}). The returns show a similar pattern representing the pluralistic understanding of vulnerability.

- 8) The authors have a limiting set of inclusion criteria. For example, criteria 5 – the study has to adopt a definition of vulnerability from AR6 is too restricting, because IPCC SREX and AR5 have pretty much the same definition and conceptualization. Even if the conceptualizations in the studies are somewhat different stemming from the evolution of vulnerability concept (in IPCC AR4, 5 and 6) – it is still a doable task to appraise the literature and categorize according to the definitions adopted in this study (AR6).

Response: In line with our response to previous reviewer feedback (comment 2), we clarified the inclusion criteria as well. Amongst which is a wider inclusion of vulnerability definitions. This is in line with the practical review we did, as many studies used very different framings (e.g., talking about adaptive capacity and sensitivity) but were considered in this study. We hope that the clarified discussion of vulnerability and how we apply it in this study has improved this.

- 9) Lines 90-95 discuss why previous IPCC AR definitions have been excluded. I will challenge this rationale – previous studies can indeed be comparable, where vulnerability as per AR5/6 corresponds to adaptive capacity + sensitivity. Previous reviews have done it that way, and there is plenty of studies that point out the evolution of risk and vulnerability frameworks where AR4 vulnerability is comparable to AR5/6 risk, and AR5/6 vulnerability is adaptive capacity + sensitivity in later frameworks.

Response: We agree, as elaborated upon in our response to the previous comment.

- 10) The study has a very extensive section on limitations, mainly justifying their methodological choices and treating many of these as inherent review choices. The comments above highlight that many of these limitations could have been overcome and some choices are not justified enough.

Response: We hope that with the revised methodological approach and refined analysis, we were able to address some of the discussed limitations and offer valuable findings to the research community.