Specific comments and technical corrections of "Stocktaking of methods for assessing dynamic vulnerability in the context of flood hazard research"

<u>Overview</u>

This article aims to systematically review methods used to assess dynamic vulnerability in the context of floods. The authors also state that they compiled their findings about the drivers and effects of vulnerability dynamics in a dataset (but do not make this available to the reviewer as far as I can tell). The concept of dynamic vulnerability is important for many use cases in both research and practice. Unfortunately, the manuscript is unorganized, often vague, and many of the claims are imprecise. There are signs in the writing that the authors did not finish revising and editing the paper (e.g., an incomplete sentence L180-181 among other sloppy instances). These are limitations that a careful revision could overcome if the editor was forgiving. However, the review is not systematic and that is the main potential value of the manuscript. To overcome this challenge, the review needs a more rigorous sampling strategy, more distinctive and clear conceptual classifications, and an insights-driven synthesis approach.

<u>Section Comments</u>

Abstract

- Difficult to parse because of vague language choices and a lack of examples. For example, the authors don't define "dynamic vulnerability." They explain that dynamic vulnerability is driven by a wide range of characteristics and that it is critical to understand flood risk.
- The abstract should define key concepts, such as dynamic vulnerability. Perhaps the authors feel that the term is self-evident or widely understood, but it is ambiguous to me and their sample of 28 studies suggests it is not a widespread concept. The second paragraph of the Intro suggests that the authors recognize vulnerability is a pluralistic concept and needs a definition.
- The authors should consider whether their four categories of vulnerability dynamics are concepts that can be grouped together (the events categories probably can be, but "underlying dynamics" seems like a different conceptual entity) and whether it is appropriate to call them "vulnerability dynamics."
- It is difficult to know what this study is about by reading the abstract.
- It seems unfair to highlight in the abstract that a key methodological gap is the limited integration for multi-hazard. The title and beginning of the abstract explicitly frames the study "in the context of flood hazard research."

Introduction

- Imprecise writing. A few early examples:
    - What do the authors mean by "robust protective infrastructure?"
    - The first paragraph of the introduction talks about the *number of flood events reported* in a database. The following sentences are about "this increase." Do the authors really want to focus on reported events in a database? It seems like with the reference to vulnerability (also the central framing of the paper) they mean to talk about an increase in flood impacts over time?
- L41 – "These challenges are evident in assessments of…" -> how? It would be helpful to have an illustration of how the challenges are evident instead of just saying they are.
- L46 – I don't know what the authors mean by the "what" of assessments or the "why" and "how." Do they mean vulnerability in "what," "how" and "why?" Please be more specific. The following examples do not clarify the what/why/how distinctions. I'm not sure how to place the subsequent examples in the context of the preceding sentences. This is the scoping paragraph of the paper. If the reader does not understand these points, they will struggle to understand the purpose or contributions of this paper. The title and abstract make the review seem focused on methods of accounting for or measuring dynamic vulnerability, but this scoping paragraph seems disconnected from that framing.
- The authors cite Simpson et al., 2021 but make no mention of their proposed "response" dimension that has an implicit dynamic component to it. How can the authors reconcile the "response" dimension to complex risk assessment in that study with their "dynamic vulnerability" framing here? I'm sure that is possible to do, but it is not done here. I think it is necessary to do because these definitions are fickle and the article presents itself as a review, so should aim to be comprehensive and to clarify muddled, overlapping, and imprecise concepts in the field.
- L52 – "a comprehensive overview of approaches for assessing dynamic flood vulnerability" in reviewing only 28 studies? This is not a good sign. I'm worried the authors are too attached to this "dynamic vulnerability" language choice instead of studying the underlying phenomena and processes required to develop a rich sample and understanding. Many fields like economics and sociology study things that can fit into "vulnerability" but may not use that language choice. One reason is that, as the authors indicate, vulnerability is multifaceted and pluralistic so some fields like to be more specific about what they are studying than using a vague catch-all term. Another reason is that there

is increasing attention in some fields to avoiding stigmatizing language such as "vulnerability" - https://www.nih.gov/about-nih/what-we-do/science-health-public-trust/perspectives/writing-respectfully-person-first-identity-first-language.

Methods

- Section 2.1: I'm not convinced these are "types" of "categories" of dynamic vulnerability. A "type" or "category" of vulnerability might be "financial vulnerability." The choice to call these concepts "categories of dynamic vulnerability" is a major concern for me. I understand that the authors base this on a previous paper, but that does not mean it is a useful conceptual device for classifying approaches. Figure 1 makes panels 1-4 seem like the response of vulnerability to different exogenous events (sometimes flood-related, sometimes not). Not sure that characterizing the sequence of events which the icons/colors represent) is a "category" of dynamic vulnerability because there can be so much heterogeneity in the vulnerability to response to any event or sequence of events. It's such a limited representation of the dynamics surrounding agent decision-making that it's distracting. For example, if a person loses their job, they are most likely more susceptible to some forms of harm from flooding than they were before. If a person gets a raise, they are potentially less susceptible to some forms of harm from flooding than they were before. For instance, they could directly cope with economic flood damage or invest in protective infrastructure. But they could also increase their spending and take on more credit, possibly having a lower budget or financial capacity to deal with harms associated with flooding. In summary, I think "vulnerability dynamics" can refer to the change in vulnerability surrounding any of the individual exogenous events or an entire span of events. I don't see how the number and type of events are different categories of vulnerability dynamics. They seem more like different realizations of exogenous events that different people/organizations/institutions may respond differently to.
- More on Figure 1 - I think it is attractive and promising, but the icons on the top are not self-evident, there is no link to the definitions of the categories and the 4 panels (is it supposed to be 1:1?), I'm not sure what is the purpose of the magnifying glasses, and the lack of a legend makes aspects uninterpretable. I think part of the problem is how abstract the concepts of vulnerability and dynamic vulnerability still are at this point of the manuscript. The Figure 1 might work better if it sticks to a specific example, like a structure's susceptibility to damage for a fixed level of flooding, over different sequences of shocks. It might work to have a few rows with different examples, demonstrating how vulnerability

dynamics look different for different types of actors and different sequences of events/shocks.

- The concept of vulnerability is so abstract here and it would be very helpful to have more concrete examples. Otherwise, it is ambiguous and distracting to keep seeing references like "financial vulnerability." What does this mean? I think the reader needs guidance on a specific example of a vulnerability and how it can change over time. On the financial vulnerability reference, I have a lot of confusion. Referring to Thomson et al. 2023, they write "[Our] approach utilizes a series of … to improve understanding of how systemic financial risk could arise from flood impacts to residential properties. As such, this work illustrates a more nuanced approach to evaluating flood-induced financial vulnerabilities." I know this is not the paper under review, but the present study characterizes Thomson as assessing "financial vulnerability" so I wanted to raise the relevant context. Is systemic financial risk a type of flood-induced financial vulnerability? Perhaps, but it's not self-evident and needs much more development and spoon-feeding as a concept. What is the element at risk? What are the specific social, economic, and physical characteristics of the element at risk that characterize "financial vulnerability?" Why isn't "financial vulnerability" a vulnerability category? Are we just talking about income?

- I don't see the four categories of dynamic vulnerability as a useful division of concepts. The examples for co-occurring event dynamics and underlying dynamics categories seem very similar. For example, a pandemic lockdown can induce an economic crisis. I think the authors need a better classification scheme or much more justification for the one proposed here. On L59 they say "we identify four categories of dynamic vulnerability essential for flood risk assessment (Figure 1)" but don't say *how* they identified those categories.

- The authors claim the review is systematic (L81) but it's not clear how to determine what a successful systematic review "of dynamic flood vulnerability" looks like. Is it models of dynamic flood vulnerability? Is it drivers? Is it data sources for evaluating? The authors may feel that my subsequent comments miss the mark because I focus on things outside their scope, but the issue is that they don't define their scope clearly in the Introduction. "Approaches for assessing dynamic flood vulnerability" is general and vague, yet only 28 studies in the sample? And I see that none of the studies that immediately come to my mind (anchored in the Thomson example thinking about "financial vulnerability" which I think is really just income) are cited in this paper, like:
  - Gallagher, Justin, and Daniel Hartley. 2017. "Household Finance after a Natural Disaster: The Case of Hurricane Katrina." American Economic Journal: Economic Policy 9 (3): 199–228.

- o Deryugina, Tatyana, Laura Kawano, and Steven Levitt. 2018. "The Economic Impact of Hurricane Katrina on Its Victims: Evidence from Individual Tax Returns." American Economic Journal: Applied Economics 10 (2): 202–33.
  - o Deryugina, Tatyana. 2017. "The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance." American Economic Journal: Economic Policy 9 (3): 168–98.
  - o Or even Kousky, C., Palim, M., & Pan, Y. (2020). Flood damage and mortgage credit risk: A case study of Hurricane Harvey. Journal of Housing Research, 29(sup1), S86-S120 which is a motivating reference for Thomson et al. (2023).
- Also on the systematic claim, I don't understand how investments in flood-risk reduction are out of scope of a review focused on dynamic vulnerability. These investments often occur in the context of post-disaster intergovernmental transfers and disaster aid. For example, see Deryugina, Tatyana. 2017. "The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance." American Economic Journal: Economic Policy 9 (3): 168–98 and Davlasheridze, Meri, Karen Fisher-Vanden, and H. Allen Klaiber. "The effects of adaptation measures on hurricane induced property losses: Which FEMA investments have the highest returns?." Journal of Environmental Economics and Management 81 (2017): 93-114. This is also a particularly interesting dynamic because of connections to gentrification. Areas that become safer may become more desirable to live, attracting investment and increasing overall exposure (and potentially risk!) while also potentially displacing people with lower incomes.
- I am surprised to see that Google Scholar was used for obtaining an initial sample. A recent article in PlosOne recommends in its abstract that "whilst Google Scholar can find much grey literature and specific, known studies, it should not be used alone for systematic review searches. Rather, it forms a powerful addition to other traditional search methods." See Haddaway, Neal Robert, et al. "The role of Google Scholar in evidence reviews and its applicability to grey literature searching." PloS one 10.9 (2015): e0138237. One of the most important reasons not to use Google Scholar is that it is much more difficult to reproduce and replicate a search than using a more typical database for a systematic review such as Scopus or Web of Science. The large number of not peer-reviewed studies in their initial sample and studies excluded based on title and abstract make the use of this search engine questionable as well.
- In addition to the use of Google Scholar, I find several aspects of the article sample process flawed:
  - o Criterion iii requires that the article be freely accessible to the reviewers. Was interlibrary loan not available to the authors? This can unnecessarily

limit the sample size and with authors from so many institutions it is surprising that this is a limiting factor for including some studies. Dropping 35 studies for this reason seems

- o Criterion v requires that the study adopt a definition of vulnerability consistent with the IPCC or UNDRR. This seems to substantially limit the authors' ability to systematically review dynamic vulnerability and how it is studied in research. It seems it would be more effective to search for research that focuses on *specific* drivers of dynamic vulnerability (like recovery, changes in income, adaptation funding, infrastructure projects, etc.,). This criterion seems to induce a selection bias for disciplines that utilize vocabulary preferred by the authors. I think the review would have more useful insights if it expanded its search to include studies that investigate

- o Criterion vi requires that the study address one of the vulnerability dynamics identified in Figure 1. Seeing that these dynamics were *a priori* determined by the authors, instead of informed by a review of potentially relevant literature, is concerning and demands that the authors more transparently define *how* they identified these categories. They do not do this in the beginning of section 2.1.

- o Criterion vii requires that the study provide details allowing replication. This is a major limiting factor. What degree of published studies adequately provide details allowing replication? Few studies in the area of flood risk assessment even make their data available, a requirement for replication.

- o Criterion viii requires that the study is a case study. Why? What is an example of a study that is not a case study and was excluded for that reason?

- All of the criteria for screening need a justification, and many of them lack one.
- How many of the studies in the final sample of 28 come from the additional papers added by collaborators?
- On Line 98 they say they categorize methods into five groups but then proceed to only define four groups.

Results

- I am skeptical about their Google Scholar search. The first study mentioned in the results section is Phifer et al., 1988, which must be added by one of the authors because it has no mention of "vulnerability assessment" or "vulnerability analysis" and these are requirements for search terms. If the first relevant study

does not have these search terms, why do the authors think these are reliable search terms to guide a "systematic" search? The authors would likely better serve their review goal if they started from their known examples of studies and worked backwards to identify search terms.

- The Phifer et al., 1988 article raises another issue – is this article about vulnerability or impacts? The title is "The *Impact* [emphasis mine] of Natural Disaster on the Health of Older Adults: A Multiwave Prospective Study" and in the abstract states "The present study examined the *impact* [emphasis mine] of flood exposure on the physical health of this sample..." Revising the IPCC definition, "vulnerability refers to the social, economic, and physical characteristics of an element at risk that make it susceptible to harm in the event of exposure to a hazard." I suppose the element at risk here is the health of older adults. In that case, how is the study's focus not about the element at risk of health of older adults? One could argue that declining health can make one more susceptible to future declines in health. This raises the concern that the authors of the review are not clear about vulnerability "in what" and the review can complicate the concept of vulnerability in the reader's mind instead of clarifying and enriching it.

- The relatively high proportion of studies in Europe raises concerns about the sample of studies as well, given the high representation of European authors. Given the content and focus of the Phifer et al., 1988 article, I can think of a number of health and economic studies that are appropriate for the authors to include (such as those I cited earlier). I think the authors would better serve their "systematic" review goals if they used a snowball sampling approach from the studies they knew, looked at citations/references for relevant studies to include, and then worked backwards from this sample to identify appropriate screening terms for a broader review of relevant articles.

- The idea that only two studies assess vulnerability dynamics due to single flood events is wrong. I know this because one of the studies they describe in the single event section investigates the impact of the 2015 Mozambique flood on household consumption and poverty levels, using a difference-in-difference approach "to quantify changes in vulnerability attributable to the flood event." If income, poverty, welfare, etc., are all forms of vulnerability (this was not clear to me as there was no description before on how one can measure vulnerability), then this review is missing a sizable number of studies in economics that use a wide variety of methods in causal inference to evaluate the impact of floods on households. In addition to the studies I cited before, there are a number of studies that evaluate how disasters and other exogenous events (like risk communication efforts) affect flood insurance take-up and flood-risk reduction investments. The review misses key studies from authors such as Kunreuther, Kousky, Czajkowski, Botzen, Bubeck, and Gallagher (and these are authors I

know without dynamic vulnerability being a specific area of focus for me – surely they miss many more key studies and authors). For example:

- o Bubeck, P., Berghäuser, L., Hudson, P., & Thieken, A. H. (2020). Using panel data to understand the dynamics of human behavior in response to flooding. Risk Analysis, 40(11), 2340-2359.
- o Bubeck, P., Botzen, W. J., Kreibich, H., & Aerts, J. C. (2013). Detailed insights into the influence of flood-coping appraisals on mitigation behaviour. Global environmental change, 23(5), 1327-1338.
- o Botzen, W. W., Kunreuther, H., Czajkowski, J., & de Moel, H. (2019). Adoption of individual flood damage mitigation measures in New York City: An extension of protection motivation theory. Risk analysis, 39(10), 2143-2159.
- o Gallagher, J. (2014). Learning about an infrequent event: Evidence from flood insurance take-up in the United States. American Economic Journal: Applied Economics, 206-233
- o Kousky, C. (2017). Disasters as learning experiences or disasters as policy opportunities? Examining flood insurance purchases after hurricanes. Risk analysis, 37(3), 517-530

- The results section describes studies and characterizes their features (e.g., in Figure 4) in ways that do not convey insights. They hardly break down Figure 4 in the results text. Part of the problem is that the review does not state clear research questions so some of the results seem non sequitur. For example, starting on L154, the authors discuss the timing of data collection. They discuss this for 4 sentences. This is in a paragraph about "the time intervals between the consecutive events, the number of flood events considered, and the duration of the analyses." With this paragraph's stated focus, why is it "important to note that Kreibech et al. (2017, 2023) use a literature review considering various reports and publications from different years to assess vulnerability dynamics, while we took the date of the scientific publication to determine the time lag between events and data collection?" This long discussion seems irrelevant. What does any of this have to do with vulnerability dynamics? What does it have to do with methods for assessing vulnerability dynamics?

- L179: "The studies mentioned above capture the vulnerability dynamics regarding different elements at risk and vulnerability dimensions. While Phifer et al. (1988) focus on human health and well-being, taking into account demographic, economic, and health dimensions of vulnerability," -> that's the end of the sentence! This is unacceptable for a submission to such an esteemed journal.

- L196 – this paragraph's focus on elements at risk and contributing vulnerability dimensions is shallow and vague. It would be much more helpful for the authors

to explicitly state what the vulnerability dimensions are and what the impact of floods are on these dimensions. For example, what does the following mean? "Albulescu and Armas (2024) use augmented impact chains to express the effects of hazard impacts and risk mitigation measures on vulnerability without focusing on a certain element at risk"

- Section 3.5 is titled "dataset for vulnerability dynamics" but then the text on 235 starts talking about findings from studies. Very unorganized and confusing. It's nice to see some discussion of the findings from these studies, but it is presented in an unorganized way and is hard to follow. The separation into "categories" of consecutive event dynamics, co-occurring hazard dynamics, single-event dynamics, and underlying vulnerability detracts from the flow of the discussion. This separation forces distinctions that move relevant findings apart from each other. Review articles should synthesize in ways that produce insights, not just produce a list of thematically linked findings and leave it to the reader to connect the dots.

Discussion

- I don't understand how the indicator-based approaches, which the authors state entirely avoid investigating how vulnerability indicators change because of flood events, can be included in this review on vulnerability dynamics. This again supports the idea that the authors' categories of vulnerability dynamics are not helpful for synthesizing vulnerability dynamics. In fact, their categories are more like hazard dynamics (these are mostly hazard realizations over time). Vulnerability dynamics suggest attention to changes in vulnerability dimensions over time.
- The discussion on "curve methods" L334 is underwhelming. The discussion does not acknowledge a wide literature on flood damage functions and substantial uncertainty in relating flood and structure characteristics to economic damages. There is also a random seeming reference to "storage tanks" in L341. What exactly do the authors mean that a damage curve is static? I think they mean that most damage curves take generally static characteristics of a structure as input, such as number of stories, foundation type, and first-floor elevation, and produce percent damage estimates for possible inundation depths relative to first floor. But that's not all damage functions, and even these characteristics can change over time. For example, consider the following study that uses building quality as an input variable, a structure characteristic that may change between flooding at time t1 and time t2 if the damage is large enough: Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F., & Merz, B. (2014). How useful are complex flood damage models? Water Resources Research, 50(4), 3378-3395.

This topic is complex and would benefit from more thoughtful consideration of relevant literature.

- Why "agent-based modeling" on L346 but "Agent-based Model" on L349?
- How can the authors claim that "no peer-reviewed study seems to apply [agent-based modeling] to the assessment of dynamics?" L346. For example, see Taberna, A., Filatova, T., Hadjimichael, A., & Noll, B. (2023). Uncertainty in boundedly rational household adaptation to environmental shocks. Proceedings of the National Academy of Sciences, 121(29), e2215675120. I am sure there are many more relevant agent-based modeling studies.
- The discussion on causality is welcome, but underwhelming. The authors miss many economic studies that use methods in causal inference to identify the effects of floods and other exogenous events on features the authors defined elsewhere as vulnerability (see my many references to economics papers above). The main challenge is the transferability of these findings to the many analytical contexts in which projections of vulnerability dynamics can help inform risk assessments and decisions. Such a discussion is extremely important and would be a very nice contribution to the literature.
- I am confused that the authors highlight that they excluded a study by Bryant et al., 2022 because Bryant et al. frame changes as adjustments to hazard rather than vulnerability dynamics. Making a choice based on language rather than whether the object of study is the same is a poor way to do the kind of synthesis this "systematic review" aims for. Committing to strict language choices for a new idea of "vulnerability dynamics" – which is not widespread or standardized (like the way the IPCC is a helpful standard definition for "vulnerability") – is clearly a limiting factor of this review. There are clearly many studies that evaluate changes in vulnerability over time. Leaving these out because of language choices – as opposed to connecting relevant fields and literatures to create a richer concept of dynamic vulnerability (if that's even the appropriate catch-all for what the authors want to synthesize) – is a failure to be systematic.
- I appreciate their discussion on limitations, but I disagree with the dismissal of these limitations as a normal part of any review and disagree that their process "ensured that we captured a broad and representative sample of studies relevant to dynamic flood vulnerability assessment." (L414) I think the authors miss too many relevant references to consider the review systematic or helpful. The authors are clearly aware that terms such as "panel" and "longitudinal" would be helpful and don't satisfyingly justify why they stick to jargon in their field of (multi-) risk. The topic of study is important and for publication in an interdisciplinary journal such as this one, it is necessary that the authors take efforts to increase representation of studies outside of their own small community.