REVIEWER'S REPORT

I have now read the article "Building-scale flood loss estimation through enhanced vulnerability pattern characterization: application to an urban flood in Milano, Italy". In my opinion, the article needs a major revision before it can be accepted for publication. The main issues related to this article are:

GENERAL COMMENTS

-**Lack of focus**: the focus of the paper is weak. It is not clear if the authors plan to focus on loss estimation, enhance vulnerability assessment, focus on the use of indicators or the use of vulnerability/damage curves and why the study is important and for who.

-**Theoretical background:** It is not clear what the authors understand with the term "vulnerability". Generally, the reference to the theoretical background is very weak. I see the following connection:

Source+pathway estimation= flood modelling (susceptibility mapping), Receptor = exposure, Consequence = loss

But then what is vulnerability? Is it receptor+consequence? Is it the monetary loss? Or a pre-existing condition based on the building characteristics? It is not clear. And what about risk? Throughout the manuscript, the term vulnerability is used in different ways.

-**Clarity of the description of the methodology:** The methodology is not clear. A methodological workflow figure would certainly help. Additional and detailed information is needed regarding the choice of indicators, their scoring and weighting. Discussion and justification are needed in issues such as e.g. the use of absolute monetary loss.

-Important **publications on the topic are not mentioned** and cited. A **short literature** review and the presentation of existing gaps that are going to be filled are absolutely necessary. The authors have to show what has been available until now, and what are the gaps that they are filling with their research. This is not clear in the manuscript.

- A connection between **results and possible end-users** and their needs is entirely missing.

-The results are described but **not adequately interpreted.**

-The **limitations** part is very short and weak. What were the assumptions and limitations of the study and (based on them) what could be the **future research development**?

-Most of the **figures** are of poor quality and the legends cannot be read. The choice of colour for different categories is not representative.

-The **title** has to be reconsidered after the completion of the review. What do the authors mean with "enhanced" vulnerability pattern characterization?

SPECIFIC COMMENTS:

Comments on the text:

Line 37-40: Reference is missing (Vogel and O'Brian, 2004).

Line 46: "how vulnerability is generated". Is vulnerability generated? Or is it an existing condition? Please check the general issues above regarding the theoretical background of the study, definition and understanding of vulnerability.

Line 58: the authors have to refer at this point to existing publications and authors that have worked intensively with indicator-based methods for floods or flash floods and similar phenomena (see references below):

Balica, S. F., Douben, N. & Wright, N. G. Flood vulnerability indices at varying spatial scales. *Water Sciences and Technology* **60.10**,2571–2580, https://doi.org/10.2166/wst.2009.183 (2009).

Kappes, M.S.; Papathoma-Köhle, M. & M. Keiler, 2012. Assessing physical vulnerability for multi-hazards using an indicator-based methodology. Applied Geography, 32, 2, 577-590. https://doi.org/10.1016/j.apgeog.2011.07.002

Papathoma-Köhle, M.; Schlögl, M.; Fuchs, S. 2019. Vulnerability indicators for natural hazards: an innovative selection and weighting approach. *Scientific Reports,* 9, 15026. https://doi.org/10.1038/s41598-019-50257-2

Line 62: Akbas (2009): please check again. Akbas does NOT use an indicator-based approach he has developed a vulnerability curve.

Lines 76-77: "source, pathway and receptor refer to the physical process". Why does "receptor" refer to the physical process? Later on, it is stated that the receptor "refers to the entities that may be damaged by the hazards". Please revise.

Line 179: why weight "0" for buildings with basement and not the other way round? It seems that the weights are low for buildings that are more vulnerable and this needs to be clarified and justified. This is obvious also in Table 1.

Table 1: The same comment as above. Why new buildings have more weight than old ones? They should be less vulnerable or not? Who decides what is "bad"? Who decides the building status? What are the criteria? Is this expert judgement and who are the experts? The authors themselves? A panel of external experts (number and background)? Please indicate metrics. The height of the building is in meters, right?

Moreover, I think that with the word "weight" the authors mean "score". The "weight" should show the importance and hierarchy of the different indicators. For example, if the material of the building is more important than the height then the weight of the material should be higher. The score would show the individual score that would be assigned to a building for a particular indicator (e.g. for the indicator "building material" the score for wooden buildings is 1). Furthermore, a general discussion about the weight is missing. What kind of weighting is there in the literature? Is "equal weighting" acceptable? How can we acquire weights? (e.g.

statistical or participatory methods). The authors seem to follow an indicator-based approach. What are the weights of the indicators? If the authors have decided not to use any, why not?

Lines 175-177: So vulnerability classes equal a relative market value? Once more it is not clear what vulnerability is for the authors. Is the market value better to use in the assessment than the reconstruction value and why? The market value can be influenced by many factors such as view, the safety of the neighbourhood, access to infrastructure, touristic attraction, etc.

Lines 191-194: is only the scenario of 2014 taken into consideration or also the 10, 100, 500-year flood? Not clear...

Line 202: reference to a different method (curves) is rather confusing but using indicators to inform damage curves is very interesting. On the positive side, if this is really the case, I propose to emphasize it more and make it clear from the beginning of the paper.

Line 214-215: more info and detail are required. In most vulnerability or damage curves the "degree of loss" is used. What exactly is the case in the present study and why?

Line 216: "by building footprint"? what happens with multi-storey buildings?

Lines 245-262: I find this part very difficult to follow. Again, here, a methodological workflow would be helpful.

Line 253: Who decides these classes (High, Low, Medium etc.) and under what criteria? Who is the receptor of these results (and the potential end-user of the study?) What are their needs and how are these covered by the results?

Lines 271-273: How do the authors interpret these results?

Lines 280-283: How do the authors interpret these results?

Line 284: "buildings falling within 'Other roads and associated land'" How do buildings belong to a category named "roads".

Line 290-291: "making a distinction between elements with and without basement" why is this distinction necessary?

Line 371: In which way was the understanding of vulnerability improved?

Line 377-378: The authors should elaborate more here. Reference to existing curves is also missing. This is a good source of the recent advances in the field:

Fuchs, S.; Keiler, M.; Ortlepp, R.; Schinke, R.; Papathoma-Köhle, M. 2019. Recent advances in vulnerability assessment for the built environment exposed to torrential hazards: Challenges and the way forward. *Journal of Hydrology,* 575: 587-595. https://doi.org/10.1016/j.jhydrol.2019.05.067

Line 377-383: The limitations part of the paper is very short. All the assumptions and limitations have to be acknowledged.

Conclusions: there is a focus on vulnerability here whereas the paper seemed to have as an aim the calculation of loss. Furthermore, there is no discussion about possible uses and end-users of the presented research.

Comments on the Figures:

Figure 5: Attention: in the map legend it is NISR instead of NIRS.

Figure 7: Why this choice of colours? For example, high buildings should be less vulnerable and experience a lower degree of loss (red colour) and the low buildings that should be more vulnerable are green.

Figure 7d: what is the difference between apartment, residential and attached/semi-detached? Are they not all residential?

Figure 7f: Building values are uniform? Is it because they are the market values? So, they have nothing to do with the physical vulnerability in my opinion.

Figure 8: Similar comment about the choice of colours for buildings with and without a basement.

Figure 8: what are "other buildings" and "other vulnerability class"?

Figure 9: The difference between the three maps is minimal. Is it because it is only a smaller part of a larger inundation map? How do the authors explain this similarity?

Figures 10, 11: damage curves with absolute damage make no sense to me. What kind of information do they deliver that can be useful? Give more information about the making of the curves. Which regression model did the authors use? (Weibull distribution, logistic?). How do the authors interpret the fact that after a point the curves are steady? Why do they not extend beyond a specific water depth? What is the role of the number of floors? And most importantly: what are the potential uses and end-users of these curves?