

Review 'Adaptability and transferability of flood loss functions in residential areas'

Dear authors, dear editors,

It was my pleasure to review this excellent paper. It is a very well-written manuscript, both in terms of content and language use. The results shown are interesting and important for a range of scholars working on flood damage modelling.

Nonetheless, there are a number of structural, content and methodological points of improvement that should be addressed before the paper can be published. I list my comments here by major and minor points.

Major comments

1. In my opinion, some more effort could be put in the fitting of the results into the broader scientific field. The paper rightfully refers to a range of literature on flood damage model comparison and uncertainty evaluation. It is for example very similar to the paper by De Moel and Aerts (2011) and Jongman et al. (2012), who also compare different flood damage models amongst each other and with empirical case-study data. Right now, these studies are only mentioned in the introduction. It would be good if the authors could try to embrace this more in the entire paper: what are you doing differently than existing studies? What is the added value of these results? How do the uncertainty results compare?
2. The implications of the results could be emphasized better in the abstract, results section and conclusions. Right now the conclusion that 'more attention should be paid to flood loss assessments' by using 'more loss data' is in my view not strong enough. The results can be used to give more specific implications of this study: *can* we apply stage-damage functions to other areas? If yes, how can the results found here be used to make this application better, i.e. how do we transfer and adapt models? If your results show that we should not use the models in different areas, how will we move forward?
3. It is not always quite clear what the goal of the study is. I understand that you try to assess whether models give accurate results when applied in different geographies than they are designed for. It seems to me that an important part of this assessment is in fact uncertainty assessment: the differences in the models are in fact the result of parameter definition (steepness of function etc), which partly result from the geography of design but for a large part from pure uncertainty in these parameters. This uncertainty assessment is addressed specifically in various parts of the paper, for example in figure 6. But then in other paragraphs, for example page 3509, you state that 'the focus of this study was not to evaluate the uncertainty of flood risk curves'. I understand that full uncertainty assessment is not possible, but attributing all differences to 'model plausibility' (i.e. regional focus) is also not realistic. Please discuss this point further.
4. The model comparison is not quite clear.

The first reason is, that you extensively discuss the specific model FLEMO_{AT} (~2 pages), but only marginally discuss the other 3 main models applied here (one paragraph). You don't fully describe what these other models are based on: are the curves designed for a certain house type or are they general? Can they be compared this easily with the extended FLEMO model, or should we take certain things in consideration when doing this?

Second, you make model combinations that are not explained clearly on forehand. In the results section (p. 3508, line 16), the reader suddenly sees himself confronted with 57 model

combinations. Maybe you can explain this in the methodology and already mention it in the introduction?

A third point is the use of asset values. You seem to compare stage-damage functions of different models, but use the same asset values in the comparison. It is important to discuss that damage models are developed as a combination of depth-damage functions and corresponding asset values. While the ICPR model has quite steep damage functions, the asset values linked to these relative function are very low compared to other models, which is why the model generally gives an underestimation of losses. By taking only loss functions and not using model-specific asset values, you lose part of the comparability. This is an important element of your study, that should be discussed throughout the methods and results.

Finally, connected with the previous comment: in the paper (e.g. p.3507) you discuss uncertainty due to asset values. Somehow you use a range of asset values, but it is not made clear enough in the methods, result discussion and conclusions what this uncertainty is. Please elaborate further.

5. I feel the comparison with empirical data is not always discussed into enough depth. Since you compare the results and empirical losses directly, you have to 100% sure that everything that is included in the modelled losses is also in the reported losses, and the other way around: e.g. direct losses, indirect losses, structural damage, content damage, the damage due to contamination etc. Are these fully consistent? It would be great if you could discuss this further.

Also, the discussion on causality of the results is still not wide enough. You suggest on several points that the FLEMO model is better because it includes contamination (e.g. page 3503) but it is not discussed in enough depth how this shows from the results. Together with my comments on the use of asset values, your conclusions are therefore not always compelling. I would suggest to improve this by adding more discussion on the model characteristics and the *reasons* for comparative differences.

6. I miss a clear overview figure or table that shows the reader easily how the results of the different models compare, and which is 'better'. Right now a table is included with yes/no as to whether the estimates are in the significance interval. It would be great if this could be extended to a more continuous scale, which could support the discussion on the model discussion.
7. The section titled 'conclusion' is currently more a summary. The entire first paragraph and parts of the subsequent paragraphs describe again what you did. In my opinion it would be much better for the paper if it would get a real conclusion in which the results are put in context and the implications are made clear, without summarizing the methodology.

Smaller comments

1. As mentioned previously: a quick introduction of the comparison method in the introduction would be useful. This should include mention of the models that will be compared, and the fact that you only look at depth-damage functions and not asset values.
2. Page 3488: explain better what the depth-damage functions are based on, i.e. what they represent: repair/replacement costs of structure, content, contamination, etc: this is not always the same in all models. Also mention how they are linked to asset values, that are very different in all models.

3. Section 2.1: is there any information on sectoral losses (e.g. residential vs commercial)? This would be interesting information for the comparison of models.
4. Section 2.2.1: some parts of this section link to the results of your hazard validation. Maybe you can be more concise here, and move some of this section to section 3.1, especially the last part of the paragraph?
5. Section 2.3.2: it was not directly clear here that you discuss 'general' damage models (very briefly) and a localized specific damage model (FLEMO_{AT}). This difference can be made more clear.
6. You use 'surveys_GR', 'surveys_BY' throughout the paper. These seem technical variable names you use in the modelling. It is not reader friendly and not directly clear what the difference is and how this can be interpreted. In my view it would be better to describe these two different samples in words (e.g. refer to 'the Bavarian sample') and discuss how this changes the results
7. Page 3503, second paragraph: here conclusions are drawn about which model is better, and about inclusion of contamination: however, it seems this is done by comparing the models mutually, not with empirical data. How can you say which one is better and draw causal relationships at this point?
8. Page 3504, second paragraph: similar to previous comment, it seems to me that these conclusions should be drawn *after* comparison with observed losses (next page), right?
9. All figure captions can be extended. Right now most of the figures cannot be understood by the reader without going back to the methodology and/or result description. It would be better if you could provide a longer caption that fully explains the content of the figures.
10. Figure 4 has way too much information and is not informative as it is. I would suggest splitting it up in different figures, or choosing another way to represent such a vast number of graphs.
11. Abstract line 21: 'loss assessment in *the* future'
12. Line 7 p. 3491: 'mainly located' – be more explicit
13. Line 17 p. 3491: remove 'between 1971 and 2006'
14. P. 3506, line 23: explain what '..estimate the reported loss well' means.