Nat. Hazards Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/nhess-2017-442-RC2, 2018


Natural Hazards and Earth System Sciences
Discussions

# *Interactive comment on* "On the role of building value models for flood risk analysis" *by* Veronika Röthlisberger et al.

**Anonymous Referee #2**

Received and published: 26 February 2018

The manuscript "On the role of building value models for flood risk analysis" describes five different models to estimate the exposure value. It compares the exposure value estimates with insurance data about exposure values and then compares the models too each other. It concludes that models based on land-use types and models based on uniform building values don't perform well compared to models that take into account building volume. The study seems to be carried out well and this exact topic is to my knowledge not published about before. Reading the paper raised however several questions and requests for clarification (see below).

More important points:

- The title is interesting, I like it. However I feel like the connection between the content of the paper and the title isn't very good yet. The role of building value models in flood



risk analysis gets very little attention compared to building value models in general. Could you either come up with a new title or reframe the abstract, introduction and conclusion a bit so the link between the paper and the title is better.

- On Page 9, line 20 you state that attached buildings are counted as one building, this is quite different from other models that I know. Maybe this is why you find the extreme values so important? I can imagine uniform models perform really bad because of this. This also makes the building stock of a particular area very important for the model performance with a uniform value. Maybe discuss this assumption behind the paper a bit better because it could make the results not applicable to other areas.

- If I understand it correctly you first determined which model performs best for the places that have good data available. Then you compare the other models to this best "benchmark" model for the whole country. Why do you take this two-step approach? Why not just compare everything to the available insurance data points and make your judgment on this?

- How accurate is the insurance data that the benchmark model was picked on? Are these values also based on a model or are these expert estimates? If expert estimates, do these experts have some valuation model that they apply? My worry is that the insurance data has artificial relationships in it (based on their valuation model) and that this study is just recreating the valuation model of the insurance companies. In that case the model that is closest to the currently applied insurance models performs best. Can you please discuss to what extend this is a possibility?

- The conclusion states that M1 and M3 underestimate the exposure values. This sounds like a bias in the model. Is this bias not simply a problem with the parameter values rather than a problem with the model itself? If you would just increase the parameters values wouldn't that get rid of the entire bias? Or I think I probably misunderstood what you meant by this sentence so please try to explain this a bit better.

- In this paper a clear separation is made between exposure and vulnerability. This is

common in the literature. However, exposure values can also be combined with the vulnerability in flood risk analysis, for example in an absolute damage functions. This has the advantage that you only need one model rather than an exposure value model and then a model to estimate the damage fraction to be multiplied by this exposure value. An example of this setup is Wagenaar et al. (2017) which used Machine Learning methods and damage data to directly estimate the damage. These estimates are based on both hazard and exposure characteristics. Implicitly such models therefore also include an exposure valuation model. Could you discuss the benefits of the two step approach taken in this study (exposure value separated from vulnerability)?

Minor points:

- Page 1, line 24. Please explain how risk management focuses on extreme exposure values. I'm a bit skeptical about this so please convince me.

- Page 5, line 28. Please explain what a hexagon is in this context.

- Page 12, line 19. Please explain homoscedastic.

- Figure 1 and the data selection section (2.4.2) are currently difficult to understand. Please start with why data selection is required and then lead the reader along figure 1 explaining at every step why each action is carried out.

- Establish a clear definition of building value in the introduction. You mention it is replacement value later on but maybe move that to the introduction. Also a good definition of replacement value is important and useful (can also be a reference). This seems quite relevant for this paper.

- In section 2.2 please explain that in your definition of exposure only buildings are included that can actually flood. I also know definitions of exposure in which any building is included.

- In 2.4.1 a whole list of abbreviations is introduced at once. This is difficult to follow and makes the text a bit of a puzzle (especially on the first read). So I would choose



not to use abbreviations in this case.

- Page 6, line 12-18. Please explain better why this approach is required. I don't get why you need the entire benchmarking next to the validation on insurance data.

- Page 12, line 30. MEA is I think MAE

Reference:

Wagenaar, D., de Jong, J., and Bouwer, L. M.: Multi-variable flood damage modelling with limited data using supervised learning approaches, Nat. Hazards Earth Syst. Sci., 17, 1683-1696, https://doi.org/10.5194/nhess-17-1683-2017, 2017.

---

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., https://doi.org/10.5194/nhess-2017-442, 2017.