Interactive comment on "Estimating flood damage in Italy: empirical vs expert-based modelling approach" *by* Mattia Amadio et al.

Mattia Amadio et al.

mattia.amadio@cmcc.it

Received and published: 28 January 2019

We would like to thank Reviewer 2 for the in-depth review of our paper. He helped us to imprive and clarify key points of the analysis and to make the discussion more valauble.

1. Currently the data-driven models developed in this study have been produced with data points from the same event it is validated on, hence no model transfer of the data-driven models is included. In practice a model transfer from one event to another is always required for flood risk studies, it would therefore be fairer to always train the models on 2 events and validate it on the third event. Such an approach is also carried out in Schröter et al., (2014) and Wagenaar et al., (2018) and both studies show that

multi-variable models typically have more difficulties in such a transfer setting.

Thank you for this very important comment. What the Reviewer suggests is definitely a valuable alternative for independent model validation. However, in case of adopting the suggested approach, one must consider that the results would depend on the selection of the calibrating events, since the available events are inevitably different in terms of data amount and quality. On the contrary, merging all the data and selecting two thirds in a Monte Carlo framework overtakes the problem of selecting one out of 3 available events. We believe this approach might increase the utility of the collected records and the statistical significance of the trained models. We added to 3.4 (page 9, Line 9): "Trained models share the same sampling approach for validation: the observation dataset is split in three parts, where two thirds are used to train the model and one third for validation. This process is iterated 1,000 times, scrabbling the data and resampling the training set at each cycle. The output takes the mean of all iterations and provides a curve which represents the empirical damage relationship for the three events. This cross-validation approach has been previously employed in Hasanzadeh Nafari et al. (2017) and in Seifert et al. (2010) in order to optimise the statistic utility of the collected sample. "

2. I think the data-driven UVMs wouldn't perform so well in a transfer setting because the main advantage of MVMs seems their transferability (Wagenaar et al., 2018). In the current setup this advantage of MVMs isn't used. Also if the model setup is changed some discussion is required on how significant the model transfer is between the events and whether a MVM is required or whether the events are so similar that a UVM would do.

As specified in 3.4 (now improved), all trained models share the same scrambling-and-resampling iterative approach. Changing the training approach for the UVM would mean to change it also for the MVMs in order for the comparison to remain mean-ingful. The advantage of MVMs is that they consider location-specific indicators and more hazard variables in addition to water depth; by feeding the MVMs with these

event-specific data (10 variables), while UVM only consider water depth, we are in fact assessing the added value of MVM and thus their transferability potential. See also the previous comment on that.

3. For the wider applicability of the results of this research some more discussion is required on to what extend the good performing literature models are tailored to the specific flood event and setting. These expert-based models seem to be made for Italy and for similar flood events to the one seen in this study. Are these models for example also made for the same region, did the developers have access to the damage data of these events or did they carry out surveys in the region? Point here is to help the reader identify when you can take a model from the literature and when you can't and for this we need more information about the good performing literature models.

Thank you for pointing out this. More details have been added to the description of literature models and the source of their data. Also, additional explanations have been added in the discussion section.

4. The abstract currently mostly summarizes the method, as a reader I would be very curious about the findings (what works better). Could you summarize these in the abstract.

Thank you, we updated the abstracts with more details about the findings.

5. Page 2 line 16-18: Can you clarify this sentence, it is unclear and seems very crucial for the story so I wouldn't want to look up the references to get this clarification.

The sentence has been rewritten as: "Synthetic models, on the other hand, are based on "what-if analyses", relying on expert-based knowledge in order to generalise the relation between the magnitude of a hazard event and the resulting damage estimate. That means, synthetic models have a higher level of standardisation and thus are better suited for both temporal and spatial transferability."

6. Page 3, line 32: You mention 1000 flood events in 45 years, that seems way too

much, what do you mean here by the word "events"?

Correct observation, the number of events refer to the AVI catalogue from CNR and in their records there are more than 1,000 unique event codes, however some of them refer to the same date. We then aggregated events in the same date and corrected the number to 300 events.

7. Page 6, line 27: You choose to use relative flood damages rather than absolute flood damages. This is a common choice, but I think not an obvious one, can you motivate this decision?

Added: "We chose to measure impacts in relative terms so as to make them easier to compare through different times (inflation effect) and places (different currencies)."

8. Section 3.2 introduction: Nice overview on UVMs and MVMs but I think this needs something on the transferability advantage of MVMs (see above).

Improved the intro: "[...] other parameters may influence the flood damage process, [...] a large number of other non-hazard factors can be significantly different from one place to another [...] Multivariable models (MVMs) can account for such additional factors and thus are able to adapt the damage estimate to the characteristics of a specific event and location. Therefore, they may be better-suited to describe the complexity of the flood-damage process for transferability purpose."

9. Section 3.2.1: Can you make a heading for each literature model.

Sub-chapters have been split differently to improve readibility.

10. Section 3.2.1: Huizinga got his damage curves from the literature also, could you reference to the study that Huizinga got his damage function from.

That is quite a long list of studies that have been averaged, none of which related to Italy; for this reason, we prefer to keep it shorter.

11. Page 9, line 2: Change "observation" in "observed"

Changed "observation datasets" into "observed records".

12. The Random Forests and ANN both have all sorts of tuning parameters. Like number of neurons (ANN), minimum number of observations per leaf (RF), learning rate (ANN) and more. Could you describe how you determined these settings?

Unless specified, RF and ANN run on default parameters. We added the minimum number of observations per leaf in RF (5). We also added to ANN: "The learning rate is controlled by coefficient μ: when μ is very small, the training process approximates the Gauss-Newton optimization algorithm (i.e. fast learning, low stability), while when μ is very large, the training process resembles the steepest descent algorithm (i.e. slow learning, high stability). The value of μ starts as 1 and is updated during each training epoch. In case a training epoch is successful in reducing the SSE in the output layer, then μ is reduced by half; otherwise, the value of μ is increased by a factor of two and a new training attempt is performed." The number of neurons in ANN is already specified: "the initial number of hidden neurons per hidden layer is approximated as two-thirds of the summation of the number of neurons in the previous and next layers".

13. On page 11, from line 20. You describe something about the setup of the study. I think this should be somewhere else in the manuscript as this probably applies to all data-driven models (that would be most fair to do this the same for all data-driven models). If not why did you do that differently for the other models?

The referenced setup is specifically related to the ANN model; we explained better the training procedure that is shared among the trained models (3.4, pg 9 line 11:) "All these models share the same sampling approach for training and validation: the observation dataset is split in three parts, where two thirds are used to train the model and one third for validation. This process is iterated 1,000 times, scrabbling the data and resampling the training set at each cycle."

14. Sometimes you use the word "water velocity", sometimes "flow velocity" and sometimes "water flow velocity", I think commonly the word "flow velocity" is used. Can you

unify this throughout the paper.

Yes, thank you, we now use "Flow velocity" consistently.

15. Page 16, line 14. Not all these citations fit a root function to data they just all have damage curves that have the shape of a root function. So please rephrase the sentence before the citation (message can be the same).

Thanks for this comment. The sentence has been rephrased as the following: "Our findings confirm previous results indicating that the curve shape described by the root function is the most adequate to describe the flood damage process".

16. In this study a limited number of variables was available for the MVMs. If more variables had been available the models might have performed better. Can you make this point somewhere.

Added to discussion: "We can't exclude that the performances of MVMs would benefit from the inclusion of additional predictive variables, such as those related to the early warning system and precaution measures, or social vulnerability; however, the availability of such information is limited for our case study."

17. References: Schröter, K., Kreibich, H., Vogel, K., Riggelsen, C., Scherbaum, F. and Merz, B. (2014), How useful are complex flood damage models? Water Resour. Res. 50, 3378–3395. doi:10.1002/2013WR014396, 2014. Wagenaar, D., Lüdtke, S., Schröter, K., Bouwer, L., Kreibich, H., 2018. Regional and Temporal Transferability of Multivariable Flood Damage Models. Water Resources Research. Volume 54, Issue 1. https://doi.org/10.1029/2017WR022233

Interesting articles, thank you, these have been added to the discussion.

Please also note the supplement to this comment:
https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2018-324/nhess-2018-324-AC1-supplement.pdf