

## REVIEW REPLY

### **Anonymous Referee #1**

The paper describes the development of flood loss models on the basis of a remarkable dataset of observed flood losses. This dataset was used to develop different kinds of loss models and to validate these models, as well as other models available in the literature. In general, the paper presents an interesting study. The novelty lays in the approach for developing a new approach for multi-variable flood loss models. However, while reading the paper, some questions arose. With some explanations added, the paper will be of interest for the flood loss modelling community.

We would like to sincerely thank the Anonymous Referee #1 for his positive review and input, which helps us significantly in improving the presentation of our study.

The main and principle question that arises is, if the random forest approach is sensitive to heteroscedasticity in the data. As figure 10 shows, the deviations from the observed data vary with magnitude. It is highly recommended to test the data for heteroscedasticity and to tackle with this issue in the development of the models if necessary. It would be of interest how the residuals are distributed.

We thank the Reviewer for pointing this aspect out, which we missed to properly address in the original version of the manuscript. We will deepen the analyses about heteroscedasticity, performing the tests suggested, in order to improve the quality and the robustness of our multi-variable model, and we will examine which measure can be taken to correct it where appropriate.

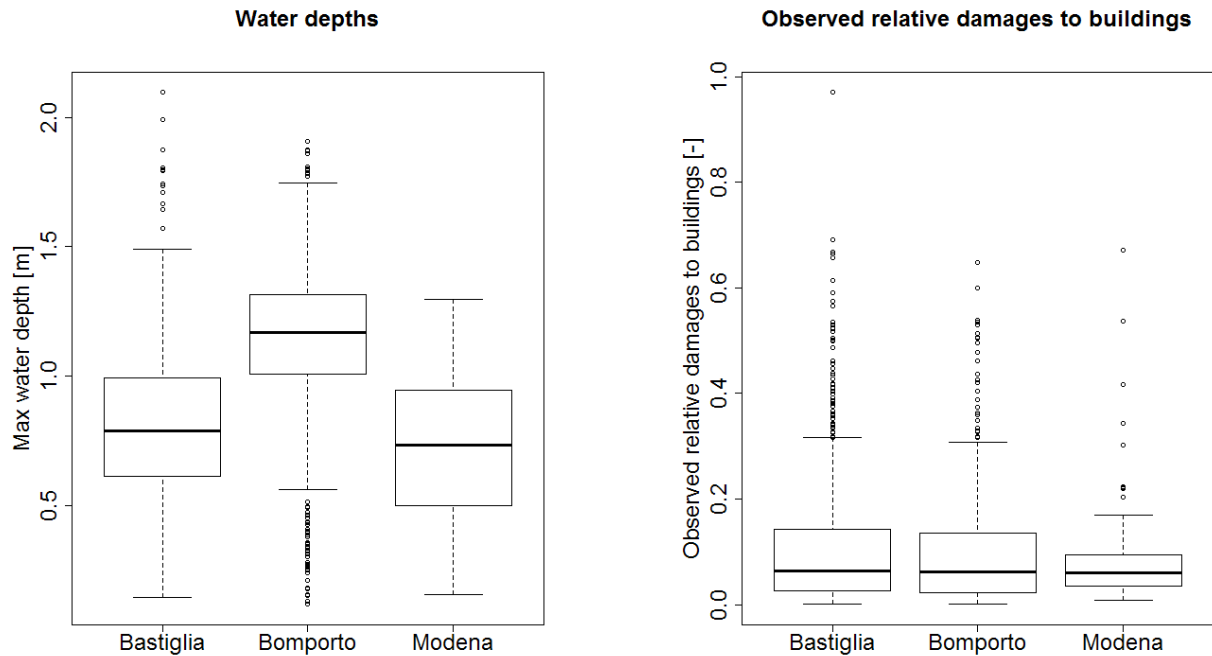
Furthermore, as in the introduction is stated, the slope of the floodplain is very regular. Thus, flow depths vary only in case of backwater effects of hydraulic obstacles. Hence, flow velocities in this relatively homogeneous case study may not be considered as independent variables (dependent on flow depth). I don't know how the flood model used for the analysis computes velocity and flow depth. Anyway, they are interlinked through the model used. However, this is a hypothesis and the contrary should be demonstrated.

For tree based models no assumption about independence of variables is needed. Anyway, as we are looking at the maxima both in case of water depth and velocity, they commonly refer to different time steps. Thus, we think it is not a problem if the descriptors show some degree of correlation. We will add a short explaining comment to the text.

While looking at Fig. 11, a question arises if both cases Bastiglia and Bomporto do have relatively homogeneous flow depths inside of their samples but differs remarkably between both. This may lead to an overrepresentation of a certain flow depth interval and hampers the transferability of a model calibrated on one case study to the other case study. Figure 1 strengthens this observation, although the

flow depths are not visible below the clustered points. I recommend showing a box plot of the flow depths at the single buildings for both case studies.

We agree with the Reviewer. In fact, water depths in Bastiglia are lower than in Bomporto, although the distributions of the observed damages are quite similar (as you can see in the box plots below). We agree that this is worth specifying it in the discussion of the results on model transferability.



The authors are asked to assess the reliability of the flood loss estimations (in monetary terms) by the home owners immediate after the flood event. I suspect that all home owners have the competency for estimating the damages to their buildings as professionals have (insurance experts and craftsmen commissioned to restore the building). The authors should describe how these estimations were “verified for authenticity” by the administration. If this verification was made following a reliable approach, the refunded value should be used for the analyses and not the estimations.

We will improve the clarity of the section where we explain our choice to consider as observed losses the damages as claimed by citizens in Form B, instead of the refunds. Due to the specific and strict compensation criteria (i.e. not all damage is compensated) the refunded amounts differ from the “actual” damage.

Another weak point is the use of the market value for the estimation of the building’s values. It is not described, if this value comprises the cost of the land too.

The study assesses flood damages to buildings, in particular to their structural part and their contents. The use of the economic value of the structural part of the building, that doesn’t take into account the land cost, is therefore congruent with the goal of the analysis. This will be clarified in the text.

Furthermore, it is not documented if this value is given for the area of the building footprint or for the living space that should be multiplied by the number of floors.

Only the first floor of each building has been considered, being the maximum water depth lower than 2.5 m. This will be better explained in the revised manuscript.

The comparison between different flood loss models should consider the used base value for assets.

We do not completely agree with this suggestion, because the models use damages, that are relativized based on each different context, therefore they are comparable to each other.

It would be of interest which approach the authors followed for the geolocation of the loss data.

Thanks, we will improve and detail its description in the revised manuscript.

p. 8, ln. 19: is the size of 1 to 200 m for element length or area of the element?

The cited size refers to the length of the triangular elements of the computational mesh, we will clarify it in the revised text.

p-11, ln. 26 chapter 4.2.1. It is not defined what “best performance” means here.

It refers in particular to the Root Mean Square Error, it will be clarified in the revised manuscript.

Results section. The model structure of the multi-variate model, i.e. the outcomes of the random forest analysis, should be described. Which parameter with which weights have been identified and structure the prediction model. In its present form, the reproducibility it is not given. One solution could be to adapt Fig. 4 and insert the resulting model structure.

Thank you for your suggestion. Unfortunately, the structure of a Random Forest (RF) is difficult to describe. A RF consists of 500 bootstrap replica of each record of the dataset with one tree grown for each replica. RF are black-boxes and it is not possible to report each tree including details about all splits. We will show examples of built trees (perhaps in an Appendix), i.e. adapting Fig. 4. Additionally, we will use the appendix to detail the algorithm.

p. 14, ln. 28-29. In addition to the comparison of the predicted losses with observed ones, it would be of interest splitting the dataset stochastically. Together with the comparison between both calibration datasets with the opposite case study data, the conclusion of the transferability could be grounded more reliably. A sensitivity test of the SMV model should be done.

The Random Forest algorithm includes a stochastic splitting of the data by using bootstrap replica of the dataset to learn the individual trees of the forest. The predictions of these trees are aggregated to a common prediction. A sensitivity test of variables included in the SMV model is done in terms of the analysis of variable importance (cf. Figure 6), with higher importance values for more sensitive variables.

p.17, ln.1-5. There is a conflict between text and figure 11. In the text, the grey dots are described as observations. In the figure, no blue dots are visible as mentioned in the text.

Thanks, this error will be corrected (grey dots refer to the estimation of relative loss using the MV models).

p.17, ln.10. “in the sake of brevity”. This can be shown in the appendix

Good suggestion, we will keep it in consideration and add it in the appendix of the revised manuscript.

p.18, ln. 16. What is “Sec. 8”?

It will be corrected (Sec. 5.1).

Fig. 1: The authors are asked to explain why they mapped only flow depths >10 cm. Are the analyses based on the full range of flow depths or are flow depths >10 cm generally omitted throughout the study?

In order to take the uncertainties of hydrodynamic modelling into account, we regarded as flooded only those areas with simulated water depths above 10 cm. This will be better explained in the revised manuscript, also providing the reader with references.