Response to Referee 4

We would like to thank the referee for the time and effort put into reviewing the manuscript. This response (R) carefully addresses the comments (C). Where applicable, changes are proposed to the manuscript accordingly.

C: According to my understanding, authors reach a very important conclusion. In data-scarce regions, where no "local" information is available on building vulnerability, the use of OSM derived spatial measures to learn multi-variable models gives comparable performance to alternative multi-variables models (which use comprehensive and detailed information about preparedness, socio-economic status and other aspects of building vulnerability), and better performance than models based only on water depth. Second, although the use of OSM does not resolve at all the problem of transferability (i.e. models remain strongly context specific and can be transferred only to regions with similar building geometric features than the calibration one), it supports transferability, by guarantee consistency in input variables between the implementation and the derivation context. Such considerations are explicitly written in the abstract, but they are implicit/hidden in the text; on the contrary, they should be highlighted to improve the usability of paper results.

R: We agree with the reviewer, and we will emphasise these outcomes in the conclusions as suggested in the track changes manuscript.

C: Given the high correlation among OSM derived spatial measures and the impossibility to rank the weight of such variables in shaping damage, the fact that OSM derived models perform similar than other MV models could be due, on the one hand, on the fact that spatial measures serve as proxy variables for other vulnerability parameters not considered by the model (as suggested by authors), but also on the fact that damage mostly depends on extensive variables, like area and perimeter, while the role of other variables (both geometric or socio-economic/technical) is negligible. From this perspective, it could be useful to evaluate also the performance of a model based only on three/two variables: water depth, area (and perimeter)

R: We support this suggestion. The performance of models using different numbers of input variables derived from building geometries is anlaysed in section 4.1 of the paper and the results are reported in Figure 5 and Table 4, e.g. the model using 2 variables is based on water depth and Perimeter-area ratio. For clarification, we suggest to include these details to the text of section 4.1 and also give these details in the caption of Table 4.

C: Could model uncertainty also be linked to differences in building geometries derived from present OSM data and building geometries at the time of the event?

R: This is an interesting point. OSM Data have been retrieved in 2017 and the surveys of affected households have been conducted in the period from 2002 to 2014. In theory, some of the buildings may have changed in terms of geometry by for instance retrofitting, demolition and new construction. However, we consider that this only applies to a minor number of buildings. Instead, other sources of uncertainty play a more important role. For instance, the uncertainty related to limited detail of geo-location information which affects the matching of telephone survey data to OSM building objects. This has been addressed with the benchmarking model Brm.

C: Table 2 and Table A1 do not help in understanding the meaning of the different spatial mesures. Meaning of extreme values should be included in Table 2, at least

R: We think the key information to understand the meaning of the different spatial measures is given in Table 2 and the table with geometry examples in Annex A1. However, some details may be hidden in the text of Table 2 and not directly apparent to the reader. As we have been asked in the previous round of reviews to remove the description of spatial measures from the text, we suggest changing the structure of Table 2. We propose to include an additional column, which gives information about the ranges of the different variables.