

Interactive comment on “Residential flood loss estimated from Bayesian multilevel models” by Guilherme S. Mohor et al.

Nivedita Sairam (Referee)

nivedita@gfz-potsdam.de

Received and published: 17 January 2021

The paper presents a Bayesian multi-level model to estimate residential flood losses. The study is meticulously undertaken and the manuscript is very clear. The authors infer that flood source/type is an important aspect that causes variability in damage model parameters. This is determined based on significant differences in damage model parameters/hyper-parameters across flood types such as levee breach, riverine, surface and groundwater flooding. Additionally, the study claims that the model with predictors - water depth, building area, contamination, duration, precaution, insurance uptake and flood experience performs the best.

Please address the following comments:

C1

Abstract: 1. Line 11: 'may complicate' lacks clarity. Please rephrase to directly mention the impact of the developed modelling approach

Introduction: 2. Line 31: Abbreviation - BMM is not used anywhere in the paper 3. Lines 36-45: This paragraph about data could be moved to the section 2.1 – Data

Data and Methods: 4. In table 2, Please also provide the split of data samples across events and regions for each flood type. The different flood types may be relevant event characteristics when split across regions. 5. In the model selection, please explicitly explain what 'little gain' means? From the values, I understand that elpd diff > 4 when adding each variable is considered as significant improvement. Is it correct? ([https://avehtari.github.io/modelselection/CV-FAQ.html#15_How_to_interpret_in_Standard_error_\(SE\)_of_elpd_difference_\(elpd_diff\)](https://avehtari.github.io/modelselection/CV-FAQ.html#15_How_to_interpret_in_Standard_error_(SE)_of_elpd_difference_(elpd_diff))) Also, please refer to STAN/BRMS forums for considerations while choosing models based on LOO-ELPD differences. Studies commonly consider elpd differences above 2SE as a significant improvement (e.g. <https://iopscience.iop.org/article/10.1088/1748-9326/ab4937>). Here, the STAN developers recommend 4SE as a safe threshold (<https://discourse.mc-stan.org/t/loo-comparison-in-reference-to-standard-error/4009/2>). However, since it is a measure of balance between bias and variance, I would like to leave it to the authors' discretion. 6. Please provide (in main manuscript) the ELPD differences and SE across the four model types – single level, flood type, event and region for fit6+11 (From SI, table S3). Are there significant differences? Which of the models perform the best?

Discussion: 7. Please present some discussion on the model diagnosis (section 3.2) 8. The concept of transferability is only discussed in lines 260-263. Please provide further analysis or information regarding how the model addresses the transferability challenge.

SI: 9. Table S3 – rephrase Year to Event 10. Please mention the corresponding SI tables in the respective sections (main manuscript). Where are tables S4 and S5 rele-

C2

vant?

Interactive comment on Nat. Hazards Earth Syst. Sci. Discuss., <https://doi.org/10.5194/nhess-2020-388>, 2020.