

## ***Interactive comment on “Beyond the stage-damage function: Estimating the economic damage on residential buildings from storm surges” by Lea Skraep Svenningsen et al.***

### **Anonymous Referee #1**

Received and published: 25 March 2020

### **Brief summary**

The authors investigate the economic flood damage to private households caused by a storm surge event in Denmark based on an extensive data set provided by various danish official institutions. The data set contains a broad range of variables including building damage, flood experience, emergency measures and social vulnerability indicators. The characteristics of the storm surge are estimated by flood models. The authors develop three regression models based on this data set using three different variable sets. Based on the explanatory power of these models the authors conclude that the influence of the inundation depth is more than halved when using other ex-

C1

planatory variables additionally. Another finding is that multi-variable models give more accurate results than simple models. The authors also found variables representing social conditions or level of emergency measures in an area to be important.

The topic of the manuscript is interesting and certainly fits to the Journal's scope. From my point of view, the main novelty of the manuscript is the inclusion of national emergency actions in the models. The manuscript is mostly clearly written. However, it could benefit from a better structure of text body (e.g. by the inclusion of more paragraphs). Some methods and motivations need to be addressed more clearly. At the moment the manuscript lacks some additional figures (e.g. correlation matrices, distribution of the variables, maps of the residuals). All results which are essential for the main points of the study should be presented within the main text. Some of the conclusions are currently not sufficiently supported by the results. A more detailed validation of models and some adaptations (e.g. the use of the relative damage in the models) is required and could help to support the main points or even facilitate to draw some additional conclusions.

C2

## Broad comments

- In general the readability of the text would benefit from the use of more paragraphs and better structuring of the text body. Numbers between one and twelve should be written as words, all greater than twelve as a number (e.g. in lines 95 and 181).
- The introduction could use a review on recent advancements on machine learning methods used in estimated flood damage. Also the question why you used a different approach should be dealt with as you mention the choice of your approach as a benefit in the conclusion. This could potentially strengthen your point.
- I guess most of the identified variables have been included in other damage models before (also in papers you already cite as e.g. Merz et al. 2013 or Schröter et al. 2014). To my knowledge the inclusion of national emergency actions has not been done before and is indeed a novel point.
- Usually the relative damage (absolute damage divided by absolute asset values) is used to make different buildings comparable with each other. The use of the absolute damage in your case could also explain that age and size have a significant influence on the damage costs.
- Correlation matrices could help to get an idea of the multicollinearity.
- I think the distributions of the used variables could be presented a bit more detailed (e.g. by the means of violin plots).
- Why no detailed cross-validation? You could even make validation with spatial transfer of the models between the different municipalities.

C3

- Moran's I should be explained in the method section. You could also write a bit more about the motivation as not every reader might be familiar with this. In addition, you could also plot the residuals spatially to check for patterns and also as a histogram to check whether they are normally distributed or not. If not the Moran's I might have a limited meaning due to a "biased" mean.
- Results (such as on the autocorrelation) which are essential to the conclusion should also be shown in the main manuscript not only in the appendix.
- With your approach you assume the same regression parameters for all locations. Could a geographically weighted regression be more useful in case of spatially autocorrelated residuals? Especially, since you expect spatial differences.

## Specific comments

- line 106: I guess this is known for a much longer time. Maybe this source (doi:10.1111/j.1752-1688.1975.tb00689.x) is more appropriate.
- line 145: remove one "A"
- section 2.1.1: Does every household have a flood insurance?
- Section 3: Is the heating correlated with the age?
- Line 272-275: Is this really the case? The increasing effect on the damage costs are only distributed on more variables in model 2 and 3. I guess, if you would predict damage costs the predictions would not differ that much.
- Line 297: This is not sufficiently supported by your results. You have not investigated the accuracy of the models in terms of predictions yet.

C4

