We thank Irene Garcia Marti for her detailed comments.

Dear authors, I read your paper with curiosity. I think it is interesting for scientists to begin collaborating with the insurance sector, so that researchers and insurers can better understand the impacts of severe weather on different socio-economic sectors. I do think it is a timely topic, but I believe the manuscript requires quite some work to become clear and deliver its core messages effectively. Hence, my recommendation is a Major Revision, and I hope the comments in this document will be helpful. Good luck.

Major comments

In this section you can find comments in two categories: structural and data analysis. For the structure, I have the impression the paper could benefit from a clearer structure, with a better division between data descriptions and the methods, whereas for the analysis comments there are parts that remain unclear.

1) Structural comments

Introduction: The introduction requires some streamlining, since it intertwines motivating reasons to carry out such a study with lengthy descriptions of previous work. As a result, it is difficult to follow the storyline the authors wish to convey. For example, in L36-L51 you begin talking about the hazard, exposure, vulnerability framework, but this somehow becomes diluted in the rest of the paragraph. It might be helpful for readers to center the introduction about these three components of risk management using the risk propeller figure, so that the references to these multiple insurance companies and other articles are somehow anchored to this image. Then in L65-L68 the authors roughly describe the analysis that will be doing, which I find too detailed for an introduction, to then explain the paper structure, which jumps back to the general scope. Overall, I think this section requires streamlining and making sure the message the authors wish to convey is effectively delivered.

Answer: We thank the reviewer for her suggestion. We will streamline the introduction by shortening the paragraph on damage datasets and expanding the part on damage functions. This is also in line with the comments by the other reviewers.

Data and methods: I would recommend re-structuring this section. While reading, there are parts mixing data description with the methods, which interrupt the flow. For example, L102-L114 describe the ERA5 data (and other generalities) right after the equations for LI are presented. Then in L116 the flow is recovered. Same goes for the description of PERILS in L154-L163. On the one hand, in the introduction the authors mention a hazard-exposure-vulnerability schema. On the other hand, I have the impression that the hazard component is ERA5, the exposure is PERILS, and the vulnerability the data/curves from AON. So I would recommend restructuring this section in 2.1.1 - Hazard; 2.1.2 - Exposure; 2.1.3 - Vulnerability and then a 2.2 - Meteorological loss index and 2.3 - Catastrophe model that are thoroughly explained without data description intrusions.

Answer: We will improve the structure of section 2 by clearly separating the data description from the methods.

Summary and discussion: I find this section long and I am not sure what the main conclusions of this work are. Is there any way of separating the “more technical”discussion part from the “more abstract” conclusions? Overall, I do not see the “take home message”, or how does this relate with the two very concrete research questions posed in L60-L64. Also, how might the insurance sector be using the insights gained in this study?

Answer: We will improve the structure of the discussion, also considering the other reviewers' comments.

2) Data analysis comments

2.1 - Meteorological loss index

In L86 the text say "Losses are proportional to the wind power or the wind kinetic energy flux....". Perhaps softening or extending this description might be useful for a generic reader to comprehend the meaning and implications of this.

Answer: This is the standard explanation for justifying the cubic relationship used in the LI calculation (see e.g. Klawa & Ulbrich, 2003). We will try to formulate the sentence more clearly in the revised manuscript.

In L87-L88 you mention that "...only the 2% of wind gusts....cause damage". I am missing here some elaboration about what are the damages that you have in mind. Are we talking infrastructural damage? Agricultural damage? To public or private assets? Is personal propriety included here? If this is one of the four assumptions in the paper, I would expect to have a solid description of what is the meaning of "damage" for the authors in this work.

Answer: We mean private buildings, in line with the reasoning by Klawa & Ulbrich (2003). We will clarify this in the updated manuscript.

In L90 you mention "In the case that no insurance data, population density can be used as a proxy for the exposure component". Indeed, but then does it mean that you are focused in damage in cities, hence, roads, agriculture, or forestry damages are out of the study? Also, how frequently do you bump into records that have no insurance data associated? I think this study could benefit from some extra clarity on how much insurance data is available, as long as its contents. This might help at assessing whether population density is a matching candidate for the insurance data or requires combining it with other layers (e.g. land use, urban tree, urban morphology).

Answer: As mentioned in the previous reply, we focus only on private buildings at a scale above city level, and thus in an aggregated form (here: 0.25°x0.25°). We will clarify this in the revised manuscript.

Also, I wonder how the different spatial dimensions are accommodated in this analysis. For example, population density from CIESIN at 0.25deg is roughly 30km, but then how insurance data are aggregated? Per country? Per NUTS region? And how does this relate with the spatial resolution of ERA5, ERA5-Interim and the catastrophe model from Aon? I believe it would be useful to have a section discussing the harmonization of the spatial dimension, so that it is clearer what the two models receive as input.

Answer: Our study is the first to compare a full insurance windstorm model (which is not available publicly) to a simplified meteorological loss index. For proprietary reasons, we can only use the Aon model output at country-scale and in a normalized (and thereby anonymized) form. Therefore, we also aggregate the LI at country level and focus on a straightforward comparison of the two methods. We will explain this better in the revised manuscript.

4.1 - Windstorm loss

Here in Figure 5 some results are visualized in the geographic space. In this figure I have two comments. First, the results are presented in a per-country basis, but the analysis seems to have been carried out on pixels much smaller than the country surface. I wonder if results can be

presented using NUTS 2 regions or a spatial unit that is closer to the spatial dimension of the analysis. If results are aggregated for the sake of visualization, this would be understandable, but then I would expect a clearer description of the treatment of the spatial dimension throughout the manuscript. What is the resolution of the insurance data? How are all these harmonized? Second, the colorscale chosen in this figure might not be ideal to visually perceive differences. Perhaps a sequential colormap (with 3 colors) or a perceptually uniform sequential colormap (eg. Like viridis) might be a better choice to guide the reader to the differences you describe.

Answer: The reviewer is correct that the LI calculation is based on gridded data. However, as stated in the previous comment, the Aon model output is only available at country level. Therefore, we decided to aggregate all data to the same "spatial units" (= countries). In the revised manuscript, we will change the coloring of some figures to better highlight differences.

Also, I do not really understand how to interpret the Figures with the storm ranks. What helps the reader understand what is relevant?

Answer: The idea behind the Figures of the storm ranks was to compare the storm ranking between ERA5 (LI method) and Aon IF's Euro WS model. The results are straightforward to understand: If the R^2 is close to one, then there is a good agreement between the rankings computed with different methodologies. Lower R^2 means a higher disagreement between the two datasets. In the revised manuscript, we will additionally show Spearman's rank correlation coefficients at country level to improve the presentation of these results.

Minor comments

Answer: Thanks for the comments; we will consider them in the revised manuscript.

References

Klawa, M. and Ulbrich, U.: A model for the estimation of storm losses and the identification of severe winter storms in Germany, Nat. Hazards Earth Syst. Sci., 3, 725-732, https://doi.org/10.5194/nhess-3-725-2003, 2003