Review of the manuscript «A framework for modeling clustering in natural hazard catastrophe risk management and the implications for re/insurance loss perspectives» by S. Khare, A. Bonazzi, C. Mitas and S. Jewson.

General comments

The authors present a novel methodology to model clustering of natural perils and its impacts for (re-) insurance loss calculation. The so-called "Super-Cluster" methodology has clear advantages over other, simpler approaches. The authors present the theoretical aspects in great detail. Very interestingly, they use European winter storms as a concrete example to illustrate their modelling approach. They also illustrate the potential difference in the (re-) insurance price modelled by using their new approach compared to the more commonly used Poisson approach. The manuscript is a valuable contribution to the discussion of the modelling of clustered processes in the insurance sector.

The manuscript is generally very well written. The authors very clearly describe their work and their assumptions. The Figures support their argumentation and conclusions.

Although the manuscript uses European winter storms as an example it hardly tackles any physical questions. Still, it has the potential to foster ongoing discussions in the field of risk assessment. I therefore recommend to improve the issues raised below and to publish the manuscript in NHESS.

Specific comments

SC1: The authors mention (only at the very end of section 3) that one problem of modelling clustered natural perils is that historical data is limited to about 50 years, and that this is clearly a very short record to parameterize clustered behaviour. They also mention that the use of dynamical models is restricted because of systematic biases in these models. I basically agree with the authors' view, but suggest emphasizing these limitations earlier in the manuscript. In addition, Pinto et al. (2013) recently provided a study on European winter storm clustering using GCM runs. Discussing their results with those in the manuscript would certainly beneficial.

SC2: I have one issue with the role the North Atlantic Oscillation should play in the context, and how the authors use it in their argumentation. The authors show simple graphics (with some basic statistics) for the relation between NAO and (severe) winter storms. However, from what they show, I found it not convincing that NAO (on a yearly basis) should impact the clustering of winter storms. In fact, anomalies in NAO can be seen simply as a consequence of more or less (or stronger/weaker) extra-tropical cyclones in the European region rather than a modulating factor. The authors should elaborate more on that issue. They might also discuss more existing literature on clustering of European winter storms and its physical basis (e.g. Vitolo et al. 2009, Pinto et al. 2009, Pinto et al. 2013). Furthermore, the authors do not mention how the NAO index they use is calculated in detail.

SC3: Regarding the parameterization in the European winter storm example: How sensitive are the results on the exact choice of thresholds and parameters? Is it a coincidence that the over-dispersion of the model exactly matches the one from the historical data set, or is it tuned to do so? Is there an analytical process to come up with the exact parameters presented, or is this rather a calibration process until the model fits with the modellers' understanding?

SC4: In terms of robustness of the statistics presented in the manuscript: All results are based on annual statistics (e.g. the relation between NAO and the number/severity of storms). Are the results and conclusions robust if looked at seasonal (e.g. September to April) statistics rather than annual statistics? What happens if the 1990 and/or 1999 cluster is excluded from the analysis? How much influenced are the results by these exceptional years?

SC5: (p. 5246) There is more published literature on clustering of European winter storm. Please include and discuss it.

SC6: (p. 5251) “we ignore seasonality and consider loss statistics based on annual time scales”: As mentioned before an easy-to-implement test of robustness could be to analyse the data not only on an annual but also a seasonal basis.

SC7: (p. 5252-5253) Are the last two paragraphs of section 2.1 really needed? From my understanding, the lack of correlation between events in a Poisson simulation is a generally known fact.

SC8: (p. 5253) “if over the course of one year the atmosphere tends to be in a strong positive phase of the NAO”: In my understanding the NAO is not really stable beyond much more than the synoptic time scale... see also comments in SC2 above.

SC9: (p. 5258) Item 5 in the list: It would help if that point would be mentioned earlier. I fully understand and agree to the point that the authors do not want to explain the physical mechanisms behind clustering, but rather seek for an efficient way to parameterize that phenomenon. However, it would help to mention that clearly right from the beginning.

SC10: (p. 5260) What does CRESTA mean and stand for? What are CRESTA zones?

SC11: (p. 5260) Please describe in detail how the NAO index is calculated. For example, is it derived from station observations, or from EOF analysis?

SC12: (p. 5261, l5-10) Are the relations in Figs. 3 and 4 statistically significant? Again, I think the argument that the NAO is the source of clustering behaviour is a bit shaky.

SC13: (p. 5262, l2) Exceedance probability of the maximum SSI value over the 39 years: With m=39 (as you mention above) and n=39 you could actually mention the exact value, correct?

SC14: (p. 5262, l10) “Particularly for short return periods around 10 years”: If I understood the graph correctly it is rather below 10 years, right? At 10 years, the clustered and the Poisson model have already nearly converged.

SC15: (p. 5262, l20) The description of Figure 5 and the Figure itself do not match. The OEP of the historical data is shown in red, and not in green as mentioned in the text, right? Furthermore, the caption of Fig. 5 mentions dashed lines which are not visible, correct? Please check consistency and correct if necessary throughout the manuscript, including Figures and Captions.

SC16: (p. 5262, l24-29) This is actually the “clustering step” mentioned in section 2.2, p. 5254, under 1)? Essentially, you define only one Super-Cluster. Maybe you could add the naming convention you introduced in section 2 (Cluster “K” etc.) here as well, just to make the correspondence between the example and the theory very clear. How realistic is it to only use one Super-Cluster? In your Super-

Cluster model, do you assume the mean annual rate is simply 1/39 years for all historical events? Is there any objective reason to choose an SSI threshold of 2.5?

SC17 (p. 5264, l10): The fact that such a year exists in the historical data itself is not convincing evidence that a 1000 year return period is unreasonable. Maybe it is more appropriate if the authors clearly express that this is their personal view (event if perfectly well backed by their modelling experience)? In addition, comparing with other published literature that came to the same conclusion might improve the argumentation.

SC18 (p. 5264, l17-22): It might be useful to mention Pinto et al. (2013) here since they provide some guidance from a dynamical model perspective.

SC19 (p. 5265, l1): You could more emphasize the impact in contract prices. The reader can deduce from the text and Figures 6 and 7 how the price would change from the changes in mean loss and change in standard deviations. However, it is not explicitly mentioned in the text for the four examples (high and low layer, zero and infinite reinstatements).

SC20 (p. 5266, l22): You mean Fig. 7 instead of 6, right?

SC21 (p. 5267 l10 to p. 5268 l4): This part could be shortened considerably. Although the two explanations why clustering impacts the standard deviation make sense, one (or a shortened combination of both) is sufficient.

SC22 (p. 5270, l20-24): Is this also due to the fact that building a timeline simulation from only 135 (historical) events might simply not add enough variation? How would that problem be tackled in reality?

SC23 (p. 5271, l3-5): See above; the reasoning why large-scale atmospheric oscillation should dictate frequency behaviour of storms is not very solid.

SC24 (p. 5271, l12): In the winter storm example, the use of a clustering algorithm is not shown or described acutally (see also SC16).

SC25 (p. 5272, l23-29): Repetition from just the paragraph before.

SC26 (p. 5275, l12): What does “EEF” stand for?

References:

Pinto, J.G., et al. (2013): Serial clustering of extratropical cyclones over the North Atlantic and Europe under recent and future climate conditions, Journal of Geophysical Research: Atmospheres 118 (22), 12476–12485. DOI: 10.1002/2013JD020564.

Pinto, J.G., et al. (2009): Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO, Climate Dynamics 32 (5), 711-737.

Vitolo, R., et al. (2009): Serial clustering of intense European storms, Meteorologische Zeitschrift 18 (4), 411-424.

Technical corrections

p. 5249, l.5-6: …North Atlantic hurricanes, European windstorms, …

p. 5249, l.18: It was rather January/February 1990.

p. 5249, l.26: within the context of natural hazard catastrophe modelling.

p. 5250, l.16: …starting point for practitioners.

p. 5252, l.28: Repetition of “is to”.

p. 5253, l.11: “cross-event correlation”

p. 5253, l.14: North Atlantic Oscillation

p. 5256, l.20: explored in Sects. 3 and 4

p. 5260, l.19: The use of “running average” instead of “rolling mean” might be more common.

p. 5262, l.19: …is depicted by the blue line…

p. 5262, l.28-29: …which we found to be…

p. 5263, l.17: …with no event or only one event are assigned zero SSI…

p. 5267, l.17: … annual aggregate loss, the fact that there is no impact on the mean loss makes sense…

p. 5276, l.9: …the above ratio is less than one…

p. 5278, l.22: Kossin et al. are not referenced in the main body of the manuscript.

p. 5279: Figure 1: I do not understand the dotted line on the lower right of the Figure. Caption: Include reference to section 4 after “attachment point” and “exhaustion point”. The meaning of these terms is not familiar.

p. 5280: Figure 2, caption: Rather use “running average” instead of “rolling mean”? Check the color-coding. Red represents storms and black NAO, right?

p. 5283: Figure 5: Let the legend overlap with the figure (e.g. place it into the lower left corner of the axis) and you save some space for the figure. Caption: Check the color coding! That was quite confusing!

p. 5284: Figure 6 and 7: Maybe you could add to the caption that Figure 6 refers to the lower layer and Figure 7 to the higher layer?