Anonymous Referee #1

Broadly I like this paper. The findings may be a bit specific to this work and I wonder about the scope for broader applicability but its a nice piece of work which solves a clear problem in working with loss data.

We thank the reviewer for the developed comments and the interesting feedback. The responses to each comment are in blue.

My more specific comments are therefore mostly minor:

17. This is a global figure? Seems very high for Europe, even if Economic rather than Insured loss.

Indeed, different estimates can be found. We will rephrase this sentence to:
"In Europe, they rank among the costliest natural hazards with billions of economic and insured losses per winter."

25-29. I Don't entirely understand what this bit is saying.

The SSI metrics is defined using the wind (or wind-gust) and a percentile of this field (usually 98th percentile). The simplicity of its formula makes it easy to compute over land at any given day. However, this metric does not consider the lifecycle of the storm. In particular the duration of the storm or its deepening ratio might be other important variables to consider when trying to understand storm damage. Also, the presence of consecutive storms, which cannot be integrated directly in the SSI, is likely to influence the damage. We will rephrase this section to make it clearer.

45-49. Seems well justified.

Figure 1 caption - Should 'different by' be 'defined by'?

Thank you for spotting this, we will change it.

120-130 - Interesting discussion of clustering metrics and comments on how a clear definition has not been fixed, or at least not used within the insurance industry. From a loss perspective is it useful to split two clusters affecting France at the same time into separate clusters if they do not overlap?

This depends on the reinsurance policies in place in the company. Usually it is more advantageous to gather the expenses over one event as this will correspond to one request to reinsurance; but in some cases, like Lothar and Martin (in Dec. 1999), when both storms resulted in important damage, splitting the two events was more advantageous. Consequently, the ability to split or gather the events depending on the reinsurance contracts in place represents a great advantage compared to the initial situation, where events cannot be distinguished and are automatically grouped.

155-159 Clearly a good dataset with locations recorded. Postcode level is surely fine for ERA5 resolution hazard.

Thank you. Yes: post-code level is already a good resolution when comparing with ERA5.

164 - Good observation and an important problem to address.

174 - I can believe this but can you give a bit more explanation/justification. It reads slightly like you argue that the loss data over-represents big storms then remove the little storms yourself. From figure 2a I interpret that lowering N leads to more claims dates than storm dates which is a different problem.

The loss data is indeed biased around major storms because of public perception and media covering. This means that when filling the claim date, the insured person might be more likely to fill a date corresponding to a well-known storm event. This is corrected by the use of storm footprints, defined on their tracks and wind gust value. Using this data, the claims are associated to the storm with the highest wind gusts, which is not necessary the storm which raised the most attention (in the case of consecutive storms for example).

Another potential source of error occurs when the insured person does not know the exact date of the damage (for example, during holidays; when s.he is away). In such cases, the damage can be filled over any day. We became aware of this limitation by spotting some days with less than 10 claims across the entire France. Given the spatial extent of winter storms, such low claim counts are highly unlikely. Based on this, we adjusted the data to group claims into “reasonably large events.” For example, Generali owns approximately 1million contracts over France, which represent 3% of the market share in France.

We will clarify this in the related section.

203 - 'Enables the capture'

We will fix this mistake.

Section 2 - Overall I find this interesting. Loss data is often focused around extreme events for vulnerability development and this study is an opportunity to work with a sufficiently complete loss history that the impact of smaller events is observed.

This is exactly the point we want to raise. In the case of winter storms the succession of smaller events can sometimes be as costly as the occurrence of a major storm.

246 - How typical Generali's portfolio is of the French market is crucial here. I expect the spread of risks cannot be shown but some commentary could be made as to whether the missing events were very focused on a particular region etc. Some commentary on the exposure comes in 271.

It is not entirely clear which storm Météo France refers to when mentioning the event of 16–17 December 2019. The associated damage is likely captured in the storms we identified on 20 and 21 December (see figure 3 of the original manuscript).

The January 2012 period, during which storm Andrea occurred, was marked by a succession of storms. Andrea mainly affected northern France, where Generali has exposure. In our database, three storms from January are linked to claims. The resulting damage is distributed across these events, which likely explains why Andrea does not appear among the 40 costliest storms.

Storm Calvann, which impacted western France on 2 January 2003, also appears in our database but is not listed among the costliest events. This is most likely because Generali has fewer policies in that region.

We will clarify these points in the revised version of the manuscript.

Fig 4 - Is there a relationship between total storm count and cluster counts? The grey bars corresponding to less clustered years appear to group in the 2000s which is also a period of low overall activity based on the red dots.
As the counts are made over a restricted period (the winter months), more storms over a fixed window means higher chances of clustering. From a statistical perspective, the storms are in fact

more likely to the be close to an existing storm date.
In fact, over the 2000s', Generali experienced few storms causing damage.

Section 4.1 seems like the most crucial results, using the dataset and methods discussed in much of the rest of the paper to draw conclusions.

We will move this section at the end of the discussion section (actual position of section 4.3) for clarity.

Fig 5. I would expect b) cost per claim to be similar, as observed. But do a) and c) depend on the definition of a storm? If your catalogue contains more small storms the total cost and total claims would be lower, as is observed.
This is one of the main results of the method. Including smaller storms result in lower values for the total cost and total number of claims per storms, compared to catalogs containing only intense events.

Fig 6. Could we see the spread of losses within clusters for different cluster lengths? For example in a 3 storm cluster is the second ranked loss still 26%?

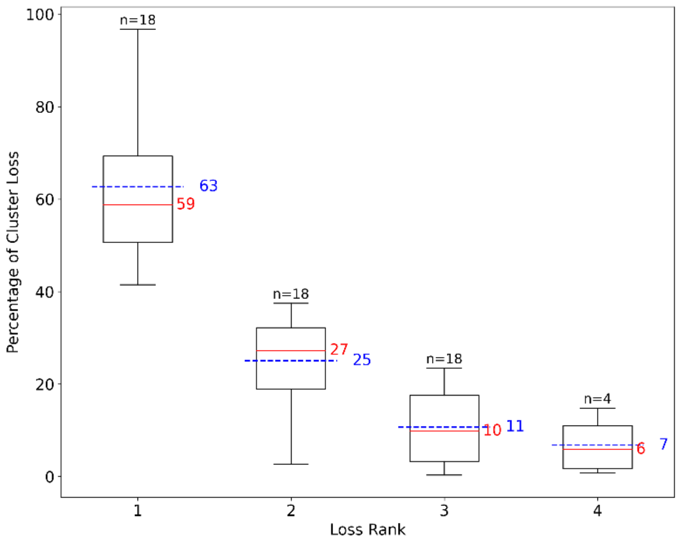

Figure 1
Distribution of the loss within the members of the cluster of at least 3 storms as a function of the loss rank. Thick red lines indicate the median, and blue dashed ones the mean.

We selected the clusters containing at least 3 impacting storms. For each storm member of a cluster, we compute its share of the loss as the total cost of the member divided by the total cost of the cluster. A loss rank of 1 corresponds to the costliest member of the cluster. Figure 1 above shows the percentage of the cluster loss as a function of the loss rank.

As we observe only 18 clusters of this type, statistics might be less robust. We can draw similar conclusions to the ones of the manuscript using the full set of clusters events (figure 6a of the manuscript). The 2nd costliest member of the cluster still represents, on average, 26% of the total losses of the cluster.

Section 4.2 shows some value to the method. Often losses from these storms are indistinguishable.

Fig 7/8 what is the resolution of the loss contour maps? The scales go to 6k and 5k respectively but within what area are 5k claims observed? Also there are some white dots that look like they hit Paris that don't seem to be explained.

The resolution is the same as the one of ERA5 (0.25°x0.25°). The claims, available at building level are aggregated over this grid. The white spot in Paris corresponded to an issue in the color bar ticks. The issue is fixed in the new version of the figures (see below).

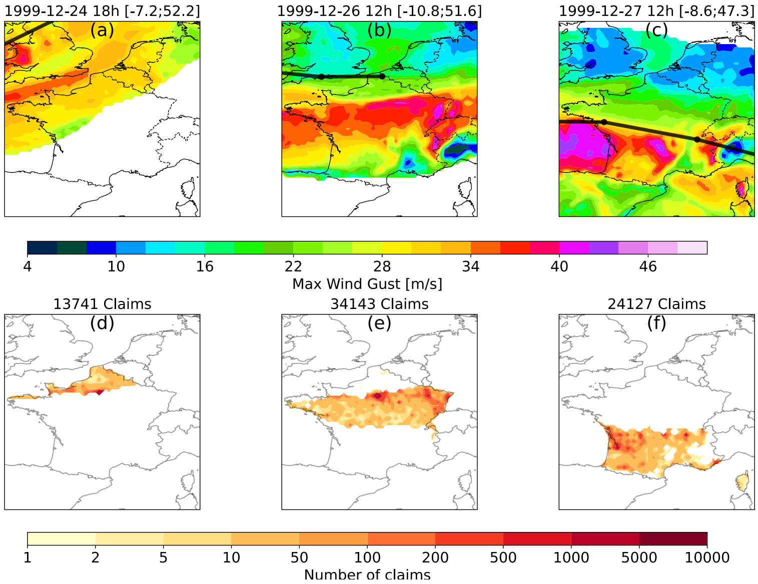

Figure 2 - Association of claims to storms. Each column contains the maximum wind gust speed and the number of claims associated with a storm in Dec. 1999. Panels (a) and (d) show maps obtained for storm Anatol ("1999-12-24 18h [-7.1;52.1]", (b) and (e) for storm Lothar ("1999-12-26 12h [-4.2;51.6]") and (c) and (f) for storm Martin ("1999-12-27 12h [-5.2;49.4]"). Panels (a), (b) and (c) show wind-gust footprints and the storm trajectories (thick black lines), panels (d), (e) and (f) show the spatial distribution of the number of claims associated with each storm, while titles indicate the total number of claims for the whole event.

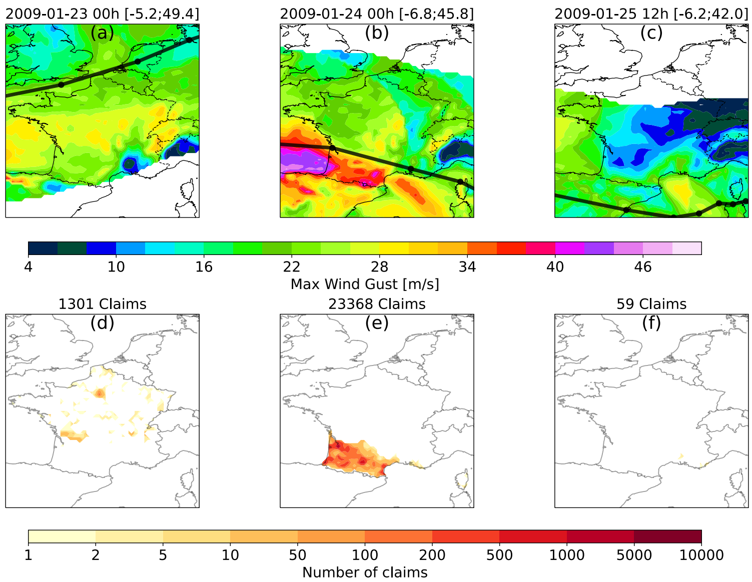

Figure 3- Same as 8 with storm A("2009-01-23 00h [-5.2;49.4]") in panels (a) and (d), storm Klaus ("2009-01-24 00h [-6.8;45.8]") in panels (b) and (e) and storm B "2009-01-25 12h [-6.2;42.0]" in panels (c) and (f).

365-368 Losses are often due to debris, hence consecutive storms may impact losses. This is very difficult for the insurance industry to study and datasets such as this may enable that.

Thank you for raising this. We will clarify this point in the text. The dataset could indeed identify the additional impact related to successive storm as it can isolate the impact of individual storms.

Section 4.3 is also interesting. As stated in 384 Klaus is often treated as a single event.

Thank you for the feedback.