I would like to thank the authors for the substantial effort to revise the paper and to reply to my comments. Unfortunately, I find that several of the points I raised were not adequately addressed. I would like to insist on the following:

We have addressed each specific remark individually to the best of our ability. We feel that some comments might reflect subjective interpretations, and in a few cases, the reviewer's expectations were not entirely clear to us. Where possible, we have nevertheless clarified the corresponding sections. Several of the points of this new review had already been considered either in our previous responses or in the revised version of the manuscript.

Taking all comments into account, we have made every effort to further improve the clarity of the manuscript and to refine descriptions.

- My overall impression is that the authors should make substantial revision of scientific language and make an effort for clearer organisation of the paper.

We revised the structure of the manuscript by separating more clearly the sections related to the physical evaluation of the storm events and their associated impact (intensity with the cost and frequency with the number of claims). We also reorganise the introduction to clarify the context and objectives of the research.

- Regarding the methods, I still do not fully understand the whole approach (see comments below).

We rephrased some sections of the method section, and we also responded individually to all the comments raised below.

- Finally, I have two general but critical comments:

A. The paper demonstrates results about the costliest storms, including famous cases, but we gain almost no knowledge about the storm characteristics that produce the damages. For instance, we gain no insights into claims' distance from storms' center, storms' intensity and wind speed at claims' location, seasonality of the clusters and storms, what kind of cyclones form the clusters, whether these clusters are multi-centered cyclones, or distinct storms (like in the case of Lothar). Figure 11 seems to summarize all new knowledge, but it only generally quantifies impacts because of clusters. Still, apart from the obvious (clusters are important for impacts), there is no information that someone could use in a subsequent study (see comment below on Fig. 11).

We have added more detailed explanations, including storm density and the distribution of storm intensity (minimal sea level pressure, maximal vorticity, and duration), to better physically characterise the storm events that result in damage. A more detailed investigation of the types of storms that cause damage and their inclusion in clusters is indeed an interesting question; however, it could constitute a study on its own. In the present manuscript, we chose to focus on quantifying the impact of such events rather than on their detailed physical characteristics. Nevertheless, we included this aspect in the discussion section, as a potential avenue for future research.

B. The paper insists too much on the 96h claims window and Generali's terms. This is fine because it connects to the claims dataset, which seems to be very important in linking cyclones with economic impacts. However, this dataset is not open access, and the 96h clause is Generali-specific (maybe I am wrong). I am aware that the method is adjustable to other temporal claim windows, but this would make the paper only attractive to other insurances with their own contract specificities and, of course, their own proprietary datasets.

You are right that the reinsurance event definition window is commonly set to 72 hours, although in practice it can be renegotiated between the insurer and the reinsurer depending on contractual arrangements. To clarify this point, we have rephrased the introduction to better explain the implications of the chosen temporal window and how it affects the interpretation of event clustering.

We also agree that the 96-hour window used in our study may be specific to Generali. However, this choice ensures consistency with the impact data available to us. To address your concern, we have added a sensitivity analysis in the appendix exploring different temporal windows. This analysis shows that the main results remain qualitatively similar when varying the window length, which supports our decision to retain the 96-hour definition for clustering in the core analysis.

Regarding data availability, the dataset cannot be made publicly accessible because it is the property of Generali France. That said, we stress that Generali has a large and geographically well-distributed exposure across France. This broad portfolio coverage makes it likely that the patterns and relationships identified in our study are representative of those that would be found for other insurers operating in the French market.

Combining these two critical comments together, I believe that the broader impact of the paper is rather weak, especially for the scientific community. I am thus inclined to recommend the rejection of the paper, except if the authors are willing to provide a more detailed association of claims to storms' physical characteristics so that results are useful to bridge the scientific community with the industry.

Specific comments

All line numbering refers to the manuscript with tracked changes.

1. Presentation

I am afraid I have to repeat the first comment of my original review. While the authors adequately addressed every single comment that I had in my first review, in my view the manuscript suffers from use of plain language and lack of focus on the narrative. The whole

manuscript demands important editing rather than addressing single phrases. I provide several examples below.

We have revised the structure of the manuscript and clarified several sections that may have been ambiguous. However, we point out that it is impossible to respond to comments that are non-specific or based primarily on a personal appreciation.

Line 23: automated scheme -> cyclone tracking method. Please be consistent throughout the manuscript on this terminology.

We rephrase this to “Tracking methods” throughout the manuscript.

Line 30: I appreciate this discussion on cyclone tracking schemes, but it seems a bit disconnected from your motivation and inadequate for the beginning of the article. You could move it to the methods (?).

We reorganised the introduction and moved this part to the Sec.2.1 Storm Data.

Line 34: I am not sure what makes SSI a “robust framework”. If I am not mistaken, SSI is a proxy for high-impact windstorms, not a framework. But even so, I am not sure what makes it “robust” here. Please elaborate, rephrase, or put into context.

“framework” is indeed not the most appropriate word. We revised the introduction in the manuscript.

Line 36: “restricted vision” is informal language. Anyway, SSI is anyway irrelevant to the existance or not of a storm.

We rephrase the whole introduction in the revised manuscript.

Line 37: A jet stream can be an atmospheric feature or a large-scale atmospheric condition. “Background dynamical environment” seems a bit uncommon, if not odd.

We rephrased this as “Large-scale atmospheric conditions”

Lines 40–41: I think that the severity of impacts only depends—from the hazard perspective—on the wind speed intensity. This is irrelevant to the underlying physical conditions. Please rephrase. Actually, the next and last phrase of the paragraph seems to be adequate.

In our proposed method, we use only surface wind speed as an indicator of hazard intensity when associating impacts. However, we aim to motivate the need to consider the full storm track, rather than relying solely on surface wind speed, in order to better characterise each event. We have clarified this point in the revised introduction.

Lines 58–59: Is the reference to Mailler et al. suggesting that “This metric was introduced based on the observation that storm occurrences do not follow a Poisson distribution with constant intensity”? If not, please provide a reference, or else move the citation to the end of this phrase.

The dispersion metric was introduced to answer the following questions raised by (Mailier et al. 2006) : *"How can the serial clustering of extratropical cyclones be quantified?"*.
This motivation for using this metric, measuring the dispersion which respect to a poisson process, is detailed in section 2.a by (Mailier et al. 2006)

"A useful statistical tool to model the succession of events such as cyclone occurrences at a particular location, is the point process [..]. The simplest hypothesis that can be formulated is that cyclones occur in a completely random fashion (i.e., the occurrence of one cyclone at any moment is independent of previous occurrences). In this case the occasional time aggregation of extratropical cyclones is only due to chance. The simplest model that describes complete serial randomness is the one-dimensional Poisson process with constant intensity (or rate), also known as the simple homogeneous Poisson process. A more complete description of this process and its properties is given in the appendix. However, cyclones affect the ambient baroclinicity of the storm track [...], so it is reasonable to think of possible dependencies that may cause the cyclones to occur either in clusters, or at more regular intervals."

We moved the reference to the end of the sentence and elaborated on what we meant by "Poisson distribution with constant intensity"

Line 66: The article has many phrases like this one which linguistically seem to make sense but, at least to me, I feel that I do not gain much insight into the issue. First of all, "threat" is informal language; second, we do not understand if extreme cyclones are clustered or are part of clustering; and third, "pronounced" is not really giving much qualitative or quantitative information.

We rephrased it to specify the issue. We strived to remove all unnecessary adjectives and adverbs.

Lines 66–70: Awkward use of parentheses.

We change the parentheses with commas.

Lines 75–77: These two phrases seem paradoxical. The first one reads: "claims falling outside an event definition are either not reimbursed.." and the next phrase reads: "..event definitions are typically based on claim date". If an event is defined by the claim, then how can a claim fall outside of an event?

We rephrased this section to make it clearer.

Lines 72–85: My impression is that in crucial parts of the text, the authors use plain language; therefore, I miss some important aspects of the rationale and motivation of this study. This new paragraph has many examples of inaccurate text. For instance, in the first phrase of the paragraph we read that insurances are interested in the "representation of hazard". I guess "hazard" is some kind of a measure for wind intensity. Then, to what does the term "event", which is introduced right after, refer? Does it refer to the date when high wind speeds caused some damages, or to a cyclone itself? If it is the former (and this is why an "event" is defined by the claim), then why do they "lack physical coherence"? What does

it mean that events are “adjusted to maximise.. reimbursable claims”? Then it is mentioned “consistent hazard-based definition”: consistent with what? And so on.

We rephrased this section to improve clarity. According to the definition given by the United Nations for Disaster Risk Reduction (Definition: Hazard | UNDRR 2007), a hazard is *“A process, phenomenon or human activity that may cause loss of life, injury or other health impacts, property damage, social and economic disruption or environmental degradation.”.* In the lines mentioned, the term “*hazard”* thus refers to the storm event itself rather than to wind intensity alone. When observed damage is attributed to a storm hazard, it is consequently associated with a specific reinsurance process. We clarify this in the revised introduction.

Line 81: This phrase seems to suggest that without linking high wind speed to physical drivers (the storms, in our case), it is not possible to anticipate potential impacts in the future. I guess that a 200-year return period of impacts due to windstorms can be estimated by multi-year simulations of wind speed alone

Estimating return levels based on wind speed alone is indeed a valid approach for quantifying wind-related hazard, and long simulations of wind fields can be used to derive, for example, a 200-year return period of extreme wind intensity. However, this corresponds to a hazard-based perspective rather than a storm-based or impact-based one.

In an insurance context, return periods should be defined with respect to physical storm events and their associated losses, not solely local wind exceedances. Distinguishing successive storm events is therefore essential to attribute impacts correctly and to assign appropriate return periods to individual events.

The proposed association of claims with windstorm events directly affects how rare-event return levels (e.g. 200-year losses) are defined. As discussed throughout the manuscript, reported claim dates are subject to delays and uncertainty, which can bias loss aggregation and, in turn, return-level estimates when based on claim timing alone. Moreover, simple temporal aggregation of wind or claims over fixed multi-day windows may merge distinct storms or split single events, particularly during clustered periods.

A physically consistent, event-based attribution, linking claims to identified storm systems rather than to wind speed alone, is therefore required to isolate individual storm impacts and to obtain meaningful estimates of storm-related return periods.

Lines 83–85: This phrase is cryptic to me. “Hazard representation” in what? In atmospheric modelling? I guess that “accuracy” refers to the minimization of high wind speed bias (provided that the hazard is high wind speed). If this is the case, then how does bias correction strengthen the “physical consistency of event definition”? What do “efficiency”, “reliability”, and “processes” refer to here? Why is wind speed accuracy relevant to this study?

We have rephrased this section in the revised version of the manuscript. What we emphasise is that, from a reinsurance perspective, the aggregation of claims should correspond to a single hazard event. However, aggregating claims within a fixed temporal window does not, in itself, ensure a physically meaningful characterisation of the underlying

hazard. Correcting claim dates and associating claims with specific storm tracks, therefore, provides the necessary physical basis for this aggregation.

Lines 86–97: This whole paragraph comes too late. It gives the motivation and reasoning of this study. All the above paragraphs make sense to me after I read this one.

We reorganised and rephrased the introduction to ensure a better flow.

Line 90: I am not sure why non-inclusion of storms' physical characteristics in risk calculation is presented as a limitation. I think that it is simply irrelevant to the risk since hazard is only defined by wind speed.

We have rephrased this section. While hazard intensity is often measured by surface wind speed, a storm should be considered as a full event, defined by its track rather than individual wind observations alone. Considering only wind speed captures only the wind-related risk, not the complete storm hazard, which may also include factors such as high precipitation. Moreover, relying solely on daily maximum wind speed makes it difficult to distinguish between individual storms occurring in close succession, as different storms can each produce high wind speeds over different areas on the same day.

Line 101: I am not really sure that C3S was designed primarily for reinsurance purposes.

The windstorm indicators derived from ERA5 reanalysis, which correspond to the reference you mentioned, were primarily targeting the insurance industry, as stated in the description of the dataset: "The primary users include the insurance sector, reinsurers and insurance industry service providers in response to their requirements for a catalogue of historic windstorm events within Europe.". We made this clearer in the revised introduction.

Line 107: What kind of biases? What is the limitation that is covered by this study?

Here, we meant that storm events cannot be defined using only the damage dataset, as this can be biased by the exposure. For example, if Generali only had contracts in the south of France 10 years ago, we would not see the impact of storms on the North of France. The other limitation of the damage dataset is the claim dates, which can be different from the actual damage date. With the approach developed in this paper, the damage dates are corrected by the association to a storm event, for which the impact date is certain.

Line 110: "Meteorology" is not a sector.

We rephrased this section.

Lines 111–114: Most of the content of this phrase is not relevant to a "meteorological standpoint" but to risk calculations (i.e., "consequences", "disentangling impacts", "conditions contributing to damages").

We have rephrased this section to highlight the added value of the paper from a meteorological perspective.

2. Methods

Line 137: Please change the reference according to the dataset used. C3S is the service hosting them.

We checked that the C3S service provides this reference for citation:

“Copernicus C3S: Windstorm tracks and footprints derived from reanalysis over Europe between 1940 to present, https://doi.org/10.24381/BF1F06A9, 2025.“

We do not know alternatives to recommendations from CDS to cite their datasets.

Line 140: Sect. C -> Appendix C

We modified this in the revised manuscript.

Line 142: If I got it right, the time of dstorm is not necessarily related to impacts, so “impacts date” is not a good way to name the date that a storm reaches 7.5°W. Please revise.

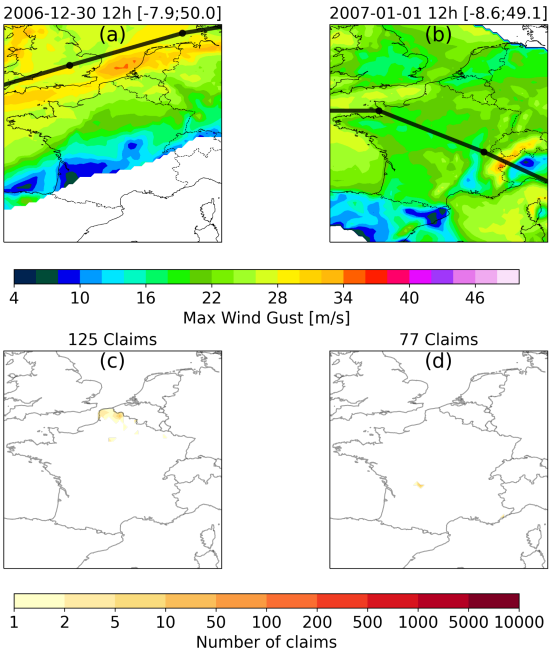
We revised this to “storm occurrence date”, consistent with the phrasing of (Flynn et al. 2025).

Line 140: What is meant by “well-developed”? Please revise the whole manuscript to avoid plain language and give physical content to the terms you use.

We revised this by removing the “well-developed” adjective.

Line 152: I am not sure I understand this ±12h window around cyclones’ center. Does this mean that the gust footprint is defined as the maximum gusts within 1300 km of each track point and within ±12h? That would mean that for some storms there will be an overlap sharing the same maximum values, right? For instance, in Fig. 1 the track points of the two storms are really close to each other (temporally and spatially) and probably they share the same maximum gusts. How often does it happen for at least two storms to share the same wind-gust maximum?

Correct: this temporal window is applied on top of the spatial mask.
Below, you can find the footprints and the associated claims of the two storms you mentioned.



Additional examples can be found in the first lines of Figs. 7, 8. In these figures, we can see that the highest windgust are well separated between the storm events.

Line 157: Relaxed relative to what? Apologies if I missed it, but I did not find the criteria used to retain or reject a track.

The criteria for rejecting tracks are described in the first two paragraphs of Sec. 2.1.
Our selected tracks should last at least 24h, against 48h in (Lockwood et al. 2022; Priestley et al. 2024). Our tracks don't have a minimal length, while (Lockwood et al. 2022; Priestley et al. 2024) require a minimal length of 1000km. Finally, there is no minimal threshold of maximal vorticity in our set of tracks, while this one is set to $10^{-5}s^{-1}$ in (Lockwood et al. 2022; Priestley et al. 2024). We compared our criteria to these specific papers, as our set of tracks was provided by Matthew Priestley (As stated in the Acknowledgement section).

Line 178: are -> area?

Yes, this is a typo. Thank you for spotting it.

Lines 175–180: I am not sure I understand the rationale here. If we consider that a gust footprint is calculated within 1300 km, this is because we believe that important impacts can take place within this radius. However, line 176 suggests that within 700 km one might get more intense winds. Is there a relationship between wind speed and radius? If so, please elaborate. Thereafter, lines 176–179 seem to suggest that this radius is chosen to tune the ratio between individual storms and clusters, while the 4-day window is chosen according to Generali's contracts. Is there an actual physical reason for the 4-day window and the 700 km radius? If yes, please elaborate. If not, I strongly suggest changing your approach so that clustering is based on physical criteria. Then, you can check how many storm clusters might match the claims.

Storm clustering is a broad concept, and as explained at the beginning of this paragraph, there is no unique way to quantify it (Dacre et Pinto 2020). It reflects the over-dispersion of storm counts, which can result from the close succession of several storms. As discussed in the introduction, clustering can be defined over long periods, such as an entire winter season, or over shorter periods of a few days or a week, depending on the objective. In this manuscript, we focus on clustering over short periods, which aligns with the definition of a reinsurance event window. For Generali, this window is set to 96 hours, which corresponds to the 4-day window used in our analysis. We have clarified this point in the revised introduction.

The choice of the radius is intended to ensure that storms within a potential cluster have a common area of impact. While there is no strict physical relationship between wind intensity and this radius, it is chosen to capture storms that could jointly contribute to damage. To address the reviewer's concern, we have added a sensitivity analysis examining the effect of both the temporal window and the radius on our results.

I appreciate the inclusion of Fig. 2, but I still find it hard to understand. First of all, Fig. 2 refers to an example which comes much later in the text (Section 4.2). Then, line 229 starts the description of the figure from the second histogram. Trying to figure out the first histogram, I deduce from it that there is a considerable number of claims. However, on the right-hand side it reads "1 Claim". Actually, I am not sure I understand the length of the red lines. In the caption it reads "storm events and their associated number of claims". Does this mean that the length of a red line corresponds to the number of claims from a single storm? I actually thought that the example here is to associate claims per storm. If so, then the red vertical lines already show this information (?!). Also, I do not understand how red lines are placed between the grey bars. I guess that a red line is placed according to the "dstorm" date? If so, this is not clear. Also, if this is the case, then I do not understand the comment about "good alignment" in line 231. If the claims are declared with a lag-time difference with respect to dstorm, then a "good alignment" might be just a matter of coincidence, right? Finally, I guess that the claims shown in the histogram correspond to the whole of France(?). If so, is it correct to assume that part of the histogram is due to claims in eastern France while the dstorm has to do with the storm being located over the Atlantic Ocean? Actually, if a storm moves, then instead of a red line there should be a "red area", since while the storm moves towards the east, it captures within its radius more and more regions, right?

I do not understand the last phrase of the caption.

Schematic -> Example

Plots -> histograms?

We revised the flowchart and its caption in the revised manuscript.

In principle, one could define a storm impact area and assign a more precise damage date, or even an hour of occurrence, based on storm displacement. However, the displacement of the maximum wind gust is already accounted for in the definition of the storm footprint. We therefore assume that using d_{storm} the corrected damage date provides a reasonable and physically consistent approximation.

Line 232: Storms are mentioned here as “large-scale events” but previously it is argued that the methods are tuned to retain small-scale systems.

We rephrased this. In this line, “large-scale events” make reference to the impact a storm can have on France. In Sect 2.1, “small-scale systems” refers to the spatial scale.

Line 234: Is there a reference or a basis for the assumption that a storm will result in at least 50 claims?

We rephrase this sentence and remove the mention of the 50 claims. The number of 50 claims comes from the observation of the bar plots of Fig. 2. Looking at the count of claims per day, we see that between 17 and 21 January 2009, claims are declared every day, but no more than 50 claims are declared per day. As described in Sec.2.3, Generali holds approximately 1millions policies. 50 claims over 10^6 policies corresponds to a claim occurrence ratio ($\frac{number\ of\ claims}{number\ of\ policies}$) of 10^{-5}, which is quite small if we compare to the claim ratio from (Prahl et al. 2015). The orange curve in Figure 6.A (Prahl et al. 2015), shows the gust dependence of the claim ratio.

Line 241: storm date -> dstorm

We change this according to the new notation.

Line 243: “The performances of the association are evaluated with three metrics, comparing the identified storms to the claim’s local maxima. The local maxima are identified by peaks over the time series of claim count gathering at least 10 claims.” This is an example of phrasing that gives me a hard time understanding the methods. How is the storm defined so that it is compared to a local maximum of claims? What is meant by peaks in the time series of claims? Are time series of claims for all France? I guess only for claims overlapping with the 1300 km radius. If different, please revise.

We have rephrased this section in the revised manuscript. Claims are aggregated over France, producing a daily time series of claim counts. Local maxima in this series, days with elevated numbers of claims, are identified and serve as targets for the association method. The objective is to associate as many storms as there are local maxima, while ensuring that storm dates are as close as possible to the dates of these maxima.

Lines 244–245: I do not understand these phrases. What is meant by major physical event? Is the physical event the storm? If so, this phrase suggests that there are also secondary or minor storms, and the tuning should capture the “major” one. Why is this assumed and not shown?

We do not make any distinction between the identified storms, and we have rephrased the manuscript to clarify this point. The only assumption made is that local maxima in the daily claim distribution correspond to storm events. These maxima are therefore used as targets in the association procedure. The proposed metrics quantify the similarity between the storm events and the distribution of local maxima by comparing (i) the number of detected storms with the number of local maxima, and (ii) the occurrence dates of storms with the dates of

the corresponding maxima. We rephrase this in the revised manuscript to clarify what is meant by “major” storms.

The paragraph continues with the definition of the metrics, but actually none is really clear to me (especially the last one in line 251).

We rephrased the definition of the last metric. Each of them is also defined in mathematical terms in the following paragraph.

Bottom line: how and why is the date that a storm is close to the reference longitude of 7.5°W important to define impacts? It seems completely arbitrary. Why is there no spatial criterion explicitly mentioned in this section? Do claim attributions with spatial criteria with respect to the storms’ area of influence? In fact, I am not really sure that the authors replied to my relevant comment in the previous review that reads “I actually failed to understand several concepts in section 3.1... better understand their necessity and use.”

I also do not understand the rationale of the reply to my question in the previous review: “In your methodological approach,... Could you please clarify.” So regarding the claims, a person might be mistaken by ±1 day on the date of impacts. So would it be reasonable (and much simpler) to simply consider that if an area within 1300 km from a storm center overlaps with a claim’s location within ±1 day, it could be attributed to that storm? In fact, the metrics used in the paper are not adequately explained, the conceptual diagram is quite difficult to understand, and I still lack examples or a good explanation about how a storm being 3 days before reaching the longitude of 7.5°W can be related to claims in France.

The core objective of this study is to robustly associate wind-related insurance claims with storm events in situations where both claim dates are uncertain and storms occur in close temporal succession.

Most existing studies rely primarily on matching claim dates with storm dates (Cusack 2023; Fonseca Cerda et al. 2024; Jaison et al. 2024; Welker, Röösli, et Bresch 2021). However, as we show in this manuscript, this approach is insufficient: claim dates may be reported with errors of several days, and clustered storms can overlap strongly in time. In the revised manuscript, we explicitly discuss the magnitude of $|d_{storm} - d_{claim}|$ and its implications. For example, in the case of storms Lothar and Martin, claims were reported up to two days before and after each storm, while the storms themselves were separated by only one day. In such situations, a simple ±1 day window combined with a fixed impact radius (e.g. 1300 km) is not sufficient to reliably separate impacts between events.

To address this, we separate the temporal and spatial criteria, as described in Sect. 3.1 and illustrated in Fig. 2.The temporal criterion is applied first to identify storms that could plausibly be responsible for claims. The longitude of 7.5°W is used solely to define a consistent reference date d_{storm} : it corresponds to the westernmost ERA5 grid point offshore of France, avoiding land effects, while remaining close to the French territory (the westernmost point of metropolitan France is at 5.05°W). This choice is therefore practical rather than arbitrary. Using another longitude (e.g. 5°W or 0°W) would not change the associations themselves, but would simply shift the optimal temporal window parameters

X_a and X_b.

Importantly, the temporal window around d_{storm} cannot be fixed a priori, because it depends on the reporting uncertainty of the claim dataset. While other studies have used fixed windows (e.g. ±5 days), we instead optimise X_a and X_b to match the characteristics of our data.

Spatial constraints are then introduced in a second step, not omitted. Storm footprints are defined using a 1300 km radius around storm track points. For each claim, we retain only storms that generate a local wind gust at the claim location; claims outside the footprint receive no gust value and are excluded. When multiple storms affect the same location, the claim is attributed to the storm producing the maximum local wind gust. This step is essential to discriminate impacts within clustered storms, something that a simple distance-and-date overlap criterion cannot guarantee.

In rare cases where identical gust values are attributed to multiple storms, claims are assigned to the storm with the earliest d_{storm}. Finally, storms associated with fewer than n_{claims} are filtered out, ensuring that only events with coherent and significant impacts are retained.

Overall, the association depends on the parameters X_b, X_a and n_{claims}, which are optimised for our dataset and hazard. While these values are data-specific, the framework itself is transferable. The revised manuscript clarifies the definition, role, and interpretation of these metrics and includes an explicit discussion of the distribution of $|d_{claim} - d_{storm}|$ for the final associated events.

Section 3.3, and especially lines 305–310: I understand that the 40 costliest storms are consistent with the strongest ones in a Meteo-France open public dissemination (Meteo-France, 2023), but this only validates that major storms are part of your dataset. This is certainly good, but your method should be expected to include the costliest storms in your datasets anyway (else it would be worrisome). I think the interest here is to see all the weaker storms and the role of clusters in enhancing impacts. Unfortunately, so far we gain no particular insight into the representativity of the captured storms from the whole methodology, or into the weaknesses and advantages of your methods regarding clustered storms. We also gain no insight into the role of smaller storms in producing damages. I also have a bit of a hard time understanding Fig. 6. Actually, “global damage statistics of storms”, “distribution of the total cost per storm, the mean cost per claim and the number of claims per storm”, and “aggregating policy-based costs at the storm level” are not explained. So far, the methods were done to associate Generali’s claims to storms, but apparently things get a bit perplexed with two additional monetary metrics (“policy-based costs” and “cost per claim”).

We have added further information characterising the full set of storms associated with impacts, with a clear distinction between clustered and individual storms. In particular, we now discuss the distributions of minimum mean sea level pressure, maximum vorticity, and storm duration for both individual and clustered impacting storms, and compare them with

the full set of storm trajectories. These elements provide a more complete physical characterisation of the events and offer additional insight into storms that contribute to damage, including those occurring in clustered situations.

Regarding the monetary metrics, we agree that their introduction was not sufficiently clear. We have revised the manuscript to define and explain these quantities earlier, in Section 2.3. In particular, we now clarify the distinction between total cost per storm, mean cost per claim, and the number of claims per storm, which are all aggregated at the storm level. We also revised the figure (now Figure9 in the revised manuscript) to explicitly distinguish between individual and clustered storms. This should make both the interpretation of the results and the associated discussion more transparent.

Section 4

Basically, Sections 4.1 and 4.2 focus on known cases that have been addressed thoroughly in the past. Using known cases is good to support the methodological approach, as it is done here. Thereafter, however, there is just a small section (4.3) to provide new insights into storms' relationship to impacts, which is rather weak. In fact, results in this section are summarized by Figure 11. This is a rather interesting figure, but it is hard to make use of it for any purpose other than quantifying the importance of clusters. As of now, we do not know the representativity of clusters in claims with respect to individual storms. We have no specific knowledge of the physical characteristics of the storms that produce the damages. We do not know, for instance, whether a loss rank 3 or 4 has a small percentage just because a very intense storms are only marginally affecting France. We also do not have knowledge about the lag dates between the events, the wind intensities that produce the damages, and whether, for instance, claims are collocated in clusters (at least this would provide some insight into recovery).

We introduced a more detailed analysis of the physical characteristics of the individual and clustered impacting storms.

Line 405: Awkward phrasing. Pretty much all extratropical storms are formed due to wave activity driven by the jet. Please rephrase or remove.

What we meant here is that the intense upper-level jet as well as its location, triggered the explosive deepening of storm Lothar. We rephrased this accordingly.

Lines 421–422: This phrase is not really clear. I am sorry for the naive question, but if I got it right, any person will claim insurance within 4 days from the damages. Then, methodologically, a specific storm can be associated to a claim if the claim takes place within +4 days from the first time that the claim's location is within 1300 km radius of this specific storm. Probably this comment connects to the methodological approach, but I think that this is a much simpler and accurate approach than the one in Section 2.

We have removed the reference to "vulnerability curves", which was indeed misplaced in this context.
In practice, however, insurance claims are not necessarily filed within four days of the damaging winds themselves. The observation of damage may be delayed, and the reporting of the claim can occur several days after the storm event. As a result, relying solely on the

claim date would allow a single claim to fall within the temporal window of multiple successive storms, particularly during clustered periods.
Consequently, a simple rule associating claims to any storm occurring within a fixed number of days is not sufficient to uniquely attribute damages. This limitation motivates the methodology described in Section~2, which explicitly accounts for storm proximity, wind intensity, and the possibility of overlapping storm influences. We have clarified this point in the methodological section to better explain why a more detailed association strategy is required.

Line 638: There are no experiments in this paper. Please revise.

We revised the author contribution.