Reviewer 2

The manuscript discusses how the risk for damage from hail has changed in the past decades in Europe. It addresses a problem that is not widely studied and is well in the scope of the journal. To showcase that damage from hail has increased, the authors present results from hazard-based studies as well as studies and catalogues looking at insured losses. To study the causes behind this trend, the authors use DAMIP models from CMIP6 to study the effect of different forcing mechanisms on northern Mediterranean sea temperatures which have been shown to be linked with hail-producing storms. They find that sea temperature in the Mediterranean has had a mostly linear response to forcings. They show that from 1850 until 1970 the cooling effect from anthropogenic aerosols dominated the overall trend of Mediterranean temperatures, after which warming effect due to increased greenhouse gases took over and has caused rapid warming since, leading to an increased hail risk in Europe.

We are very grateful to Reviewer 2 for providing comments which improve the manuscript significantly.

General comments:

Overall, the manuscript is well written and structured clearly .The literature review is conducted thoroughly and the synthesis in the introduction nicely shows how damage from hail has increased in recent decades in parts of Europe. The DAMIP models are cleverly used to attribute different trends in temperatures to different forcings. While this analysis is valid, it could be extended to better substantiate its link to hailstorm damage or risk. I have two main points to highlight:

1. *While the link between certain variables and the occurrence of hailstorms is established in previous studies, the analysis presented in the manuscript using DAMIP models does not convincingly address the trends in hailstorm damage directly or through these links. To better justify the claims made in the manuscript, the association of surface temperatures (sea or near-surface) to the variables mentioned in the introduction, such as CAPE and near-surface moisture or wet-bulb temperature, should be explained more in detail or demonstrated in the analysis. This would highlight more in accordance with previous literature how hailstorm damage trends are linked with trends of other variables.*

   The revised manuscript contains a more detailed explanation of how Mediterranean temperatures are linked to hail damage. An overview is given in the Introduction: how warming waters humidify the air (largely based on your comment 5 below) in lines 26 to 28, then followed by the association of Mediterranean Sea temperatures with hailstorms in lines 30 to 33. Section 4 of the revised text gives fuller details on the connection from Mediterranean moisture to hail trends, in lines 171-193, including several references to past studies showing a direct influence of the Mediterranean moistening low-level air in past major hailstorms. The revised manuscript contains much more evidence of the established link from Mediterranean temperatures to hail damage.

2. *The extent of the analysed area is quite small to claim that conclusions can be drawn for the whole of Europe.*

   We agree that the Mediterranean area in the original analysis was quite small, and now study an expanded Mediterranean area depicted in Figure 1 of the revised manuscript.

   The mechanism with which conditions in northern Mediterranean affect hailstorm occurrence elsewhere in Europe should be highlighted more than through the mention of a reference.

   The weather conditions conducive to large hail over mainland Europe are summarised in lines 174-177 of the revised manuscript (specifically, a trough to the west of where hailstorms develop, and how this trough acts to draw Mediterranean air northwards, as described in Kunz et al., 2020).

   In addition or alternatively, some statistics from previous studies of how often a hailstorm (elsewhere) in Europe is influenced by high dew points in the Mediterranean could be included.

   The revised manuscript contains our best knowledge of the statistics, in the paragraph at lines 171-180. There are no comprehensive statistics, but Kunz et al. (2020) describe the Mediterranean moisture source in their typical hailstorm case in western Europe, and we also cite research into very severe events in recent decades, which all drew moisture from the Mediterranean.

The revised text also discusses how hail trends at all locations depends on other factors, in addition to Mediterranean warming, in Section 4 (lines 190-192), and in the final paragraph of the Conclusions.

Moreover, the claim that hailstorm damage can be assessed for whole Europe in this way, is valid only if a significant majority of damage from hail occurs in the mentioned areas in central and western Europe. If this is the case, it should be somehow shown in the manuscript. Otherwise, I suggest the title (and elsewhere in the text) to be more focused on a specific area of Europe.

Lines 171-180 of the revised manuscript provide the evidence that the Mediterranean has a major influence on severe hail in the higher risk areas of central and southern Europe. Specifically, five references are provided containing evidence of how the Mediterranean is a key source of low-level moisture for damaging hail in the higher risk areas of Europe. We have revised the manuscript to clarify the study concerns higher risk parts of Europe, rather than every location of the continent.

Specific comments/questions:

1. *Line 13: Last sentence in the abstract is not factually incorrect but could be rephrased to make it not sound like it is the trends that are warming seas, rather than the forcing. For example, "Given the current trends in anthropogenic forcing, seas are expected to continue warming..."*

The final sentence of the Abstract has been rephrased to clarify how Anthropogenic forcings are warming the Mediterranean.

2. *Line 54: During which time period does the increase of 1 to 1.5% p.a. occur?*

This is clarified in lines 215-216 of the revised manuscript. Your question highlights the different time periods over which trends are computed by various studies. An extra sentence has been added on how the Rädler et al. (2018) trends refer to a time period which is most closely aligned with the Mediterranean warming beginning around 1980.

3. *Line 59: Does Figure 1a show annual losses for automobiles for combined wind and hail perils or only hail, which comprise the vast majority?*

The auto losses are for the combined wind+hail perils. This is clarified in the revised manuscript, lines 236-237.

4. *Line 64: I am unfortunately not very familiar with insurance terminology. How does loss cost relate to total loss? Does it increase with increasing losses?*

The text was revised to define loss cost and provide additional explanation, in lines 96-100 of new subsection 2.2.

5. *Line 79 and 87: Previous studies have found rising low-level humidity and wet-bulb temperature to be linked with an increased hail risk. Is the thermodynamic effect here that SST and near-surface air increase as much, which through increased evaporation retains relative humidity values? Thus wet-bulb temperature rises along with specific humidity?*

Yes, thanks, we include this detailed explanation in the revised text, lines 27-29.

We would add one minor point here, that the wet-bulb temperature depends on air temperature as well as specific humidity, as shown in a psychrometric chart. This means the wet-bulb temperature is not simply a measure of air humidity, though it remains quite widely used by researchers because it is well observed. Note that the revised manuscript makes no reference to wet-bulb temperatures.

6. *Line 108: Although a link to CMIP documentation is provided, Table 1 could include more details about the models used, for example model resolution.*

Three extra columns of information per model have been added to Table 1 in the revised manuscript, including the resolution of atmosphere and ocean models.

7. *Line 111: What is the effect of using near-surface temperature from the model but sea surface temperature from observations? Are SSTs prescribed in the DAMIP models? Is the effect inconsequential because only anomalies are considered?*

The effect is inconsequential because anomalies of SST and near-surface air temperature are similar for the large spatial region (Figure 1) and multidecadal scales analysed here, considering their coupling via heat fluxes. They are often used interchangeably in long historical datasets (e.g. Morice et al., 2012; https://doi.org/10.1029/2011JD017187). More specifically, Rubino et al. (2020; https://doi.org/10.1038/s41598-020-64167-1) showed their anomalies were very similar at large space and long timescales (their Figure 4).

Coupled climate models are used in DAMIP experiments: their SSTs are simulated by the ocean model.

8. *Line 113: What is the purpose of using data from different time periods for observed and modelled values? Is it to highlight the small anomalies at the start of the period and on the other hand the increased anomaly projected to continue after the end of modelled data? I realise that methodologically it has little influence since the anomalies are calculated for the common period.*

Yes, model data from 1850-1869 highlight small impacts from external forcings simulated by DAMIP models in the early period, while the observed data after 2014 show the continued warming, especially in the 2020s. The graphic can provide this extra information, in addition to clearly showing the observed/model overlap period for validation.

9. *Line 120: What are the boundaries of this region based on? Are the modelled near-surface temperatures considered only over the shaded region in Figure 2, i.e. over land? This could be specified in the figure caption or text.*

The revised manuscript analyses a slightly larger region shown in the new Figure 1. The shaded region denotes the area of the Mediterranean Sea considered in this analysis, and this is added to the revised figure caption. Its definition is based on the area of the Mediterranean adjacent to Europe.

10. *Line 139: What is the cause of the reduced amplitude of the multidecadal oscillation? Is due to smoothing from using a multi-model and ensemble mean?*

The answer depends on the relative contributions from internal climate variability and external forcing toward the observed multidecadal oscillation. There are two main possibilities:

i. There was a substantial contribution from internal climate variability. In this case, one would expect the DAMIP multimodel mean to simulate a reduced amplitude of this oscillation, since model ensemble members are initialised with different phases of internal variability which would tend to cancel in the multimodel mean.
ii. Relatively little contribution from internal climate variability. In this case, the reduced amplitude of the multidecadal oscillation from DAMIP models may be due to the external forcings applied to the models not specified accurately, or if the simulation of the climate impact from the forcing is not accurate. There is evidence to suggest this case is plausible: the cooler period starts in the late 19th century and persists during a period with three major climate-changing eruptions [Krakatoa in 1883; Santa Maria in 1902, Novarupta in 1912], then the regions warms during the subsequent quieter volcanic period. Models are known to be sensitive to the specification of the volcanic forcing (e.g. Toohey et al., 2014), and it is also known that model simulations of the climate impacts of volcanoes have weaknesses (e.g. Driscoll et al., 2012; Zanchettin et al., 2022). We have no evidence that this multidecadal oscillation is caused by volcanic activity, and merely note it is possible, though the reviewer's suggestion of smoothing from using the ensemble mean is a simpler explanation.)

To our knowledge, there have been no published results on the relative roles of internal climate variability and external forcing toward the observed multidecadal oscillation, and it's not clear if the simpler explanation (i.e. the use of the multimodel mean) is correct.

The above level of detail is too much for our intended Brief Communication, and would distract attention away from the focus on the recent exceptional warming. It is an interesting topic for a future study.

11. Line 143: Does this refer to statistical significance in the amplitude of the multidecadal oscillation in Hist? If so, based on which test?

Yes, the model Control simulations (no external forcing, internal variability only) indicate that internal variability produces multidecadal oscillations with expected amplitude of 0.046 K for a 70-member ensemble mean. In sharp contrast, the simulated amplitude of the oscillation in the Hist ensemble mean is about 0.5 K. Therefore, the signal in Hist is many times larger than the standard deviation due to internal climate variability.

12. *Line 169: Are the more references to substantiate the claim about anthropogenic aerosols driving the multidecadal changes such as the profound peak in European windstorm damages? I would be interested to see the given reference but it is yet to be published.*

The contribution of anthropogenic aerosols (AA) to the historic peak in storminess in the late 20th century is specifically discussed by Cusack and Cox (2025). Hassan et al. (2021) give a more general description of the effects of AA on mid-latitude winds. Note that this point about AA forcing past multidecadal variations in other perils is removed from the revised manuscript.

13. *Line 183: Given the relatively good agreement between models, can you comment on why especially in the most disagreeing models (4 and 6) it looks like the deviation from the multi-model mean increases with time in Figure 3b? Is this an effect of the multi-model mean diverging from the climatological mean, i.e. larger forcing introducing more spread between the models?*

It appears that their relative deviations (i.e. anomaly / multimodel mean signal) do not grow in time. Climate models can contain different sensitivities to applied forcings, and as these forcings increase in time, these different sensitivities will lead to larger differences in impacts.

It is interesting to note how model 4 has the largest sensitivity to both GHG and Aero, while model 6 has the smallest sensitivity to both. It would require expert analysis to gain insight into why different models produce different impacts on the Mediterranean Sea temperatures for a given external forcing, and it is not clear whether the necessary diagnostics are available. Such an investigation is beyond the scope of the study.

14. *Line 199: Although the overall trend between 1850 and 1970 is a cooling one, to my ear the way this sentence is phrased makes it sound like there was continuous cooling throughout the period.*

The phrasing has been adjusted to “The Aero effect dominated up to the late 1970s to produce a slightly cool period on average,…” in lines 267-268 of the revised manuscript.

Technical comments/typos:

1. *Line 45: I believe here and elsewhere “Radler” in citation should be “Rädler”.*

Thanks, this has been corrected throughout the manuscript.

2. *Line 142: Symbol for standard deviation does not show properly.*

Thanks, it’s unclear what went wrong in original word document, and it has been fixed in the revised text.

3. *Line 158: The word “then” is repeated in the sentence.*

Fixed in revised manuscript.

References used in replies to Reviewer 2

Driscoll, S., A. Bozzo, L. J. Gray, A. Robock, and G. Stenchikov 2012: Coupled Model Intercomparison Project 5 (CMIP5) simulations of climate following volcanic eruptions, J. Geophys. Res., 117, D17105, https://doi.org/10.1029/2012JD017607.

Hassan, T., R. J. Allen, W. Liu, and C. A. Randles, 2021: Anthropogenic aerosol forcing of the Atlantic meridional overturning circulation and the associated mechanisms in CMIP6 models. Atmos. Chem. Phys., 21, 5821–5846, https://doi.org/10.5194/acp-21-5821-2021.

Toohey M., Krüger K., Bittner M., Timmreck C., Schmidt H. 2014: The impact of volcanic aerosol on the Northern Hemisphere stratospheric polar vortex: mechanisms and sensitivity to forcing structure, Atmos. Chem. Phys., 14, 13063–13079, https://doi.org/10.5194/acp-14-13063-2014.

Zanchettin D., Timmreck C., Khodri M., Schmidt A., Toohey M., et al. 2022: Effects of forcing differences and initial conditions on inter-model agreement in the VolMIP volc-pinatubo-full experiment, Geosci. Model Dev., 15, 2265–2292, https://doi.org/10.5194/gmd-15-2265-2022.