

Referee 1

Summary: The manuscript analyses ensembles of simulations from five general circulation models covering the period 1950-present, driven solely by changes in anthropogenic aerosol (AA) forcings. The goal is to identify the AA's signature in the economic losses from high winds across Europe. These simulations are compared with control simulations, with constant external forcing set at preindustrial levels.

The main conclusion is that the AA signal in wind-related losses is clear, although the exact mechanisms are hampered by high model spread.

Recommendation: In my opinion, the tesseract question is interesting, and the results could be useful, despite the large model spreads. However, I also see several aspects in the manuscript that need clear improvement. My main concerns include the statistical analysis and several paragraphs concerning the description of internal and externally forced variability and the role of the AMOC, which, in my opinion, are quite unclear.

- The author thanks Referee 1 for feedback on the manuscript. Their suggestions have improved the quality of the manuscript, in particular the descriptions of multidecadal forcings in the revised Introduction, and the statistical methods which have been gathered into a new subsection in the revised manuscript.
- Though it may be of little consequence here, the author would like to clarify that the climate model dataset studied in this manuscript differs from that described above: it contains ensembles of simulations from six climate models covering the 1850-2014 period.

Main points

1) Basically, the only statistical analysis conducted to identify the impact of AA on wind-related losses is a t-test to assess changes in the mean between two periods. I have several issues with this approach:

- The AA is not constant in time, and so the impact of AA on wind-related losses should display a comparable time evolution. I am aware that the complexity of the mechanisms involved, including the purported impact of AA on the AMOC and its feedback on storminess may involve some temporal lag, but from the model time series shown in Figure 2b only one model shows a time evolution can be roughly compared to the time evolution of the short-wave downwelling radiative forcing shown, for instance, in Hassan et al. The SWR displays a very clear maximum in the last 2 decades of the 20th century, but the time series of losses shown in this manuscript (Figure 2b) does not show, by far, this type of behaviour, with the possible exception of the model CanESM5.

Therefore, I think a more sophisticated statistical methodology should be applied to identify the AA signal in wind-related losses. A stationary t-test is not enough

- The reviewer describes how the AA forcing of storm losses in individual model-mean time-series of loss anomalies (Fig. 2b of manuscript) is *not* consistent with the multimodel ensemble mean SWR over the sub-polar North Atlantic region shown in Hassan et al. (2021; hereafter H21). This lack of consistency leads the reviewer to conclude that the modelled loss anomalies are not due to AA forcing, and that the highly significant forcing reported by t-tests are generally misleading, which in turn leads to their request for a more sophisticated statistical test.
- The hypothesis that underpins this conclusion is that the two quantities compared by the reviewer – model-mean European storm loss anomalies in Fig. 2b and multimodel-mean shortwave radiative

forcing in the northernmost Atlantic – should be consistent. However, these two quantities are fundamentally different in the following respects:

- H21 presented SWR anomalies over a limited area of the northernmost North Atlantic, and such a quantity has an uncertain relation to European storm losses, since anomalies in other parts of the climate system will exert influence on European storminess
 - individual model-mean signals in Fig. 2b contain larger sampling variability than the multi-model means used in H21, therefore, internal climate variability has smaller amplitude in the multimodel-mean used in H21
 - H21 and this study use results from different sets of climate models
- Therefore, the comparison of data in Fig. 2b with the diagnostic in H21 is not a suitable approach to measuring the reliability of the modelled AA forcing of windstorm losses
 - Nevertheless, a review of the two types of statistical testing used in the manuscript has been conducted, prompted by the reviewer's interest in this aspect of the study.
 - The multimodel mean signal (Fig. 2a) was assessed statistically using its standard error. It was found that the multimodel loss anomaly is more than ten standard errors above zero (Fig. 2a), indicating very high confidence that the multimodel mean loss anomaly due to AA forcing is significantly different from zero. A more sophisticated test will not materially change this result.
 - This study has assessed the amplitude of the mean responses from each of the six models (Fig. 2b) using a t-test which was designed to provide robust results:
 - Consideration of the size of sample: a compromise is required between using larger numbers of years for robustness, versus inclusion of years with weaker AA forcing hence smaller signals and less robust results. The original manuscript used the 40-year period 1970-2009, and this is lengthened to 1960-2009 in the revised manuscript, to include strong AA-forcing in the 1960s (e.g. Lund et al., 2019; Needham and Randall, 2023).
 - The t-test contains time information implicitly: it is focused on the historical period of maximum AA forcing.
 - The reviewer suggests a more sophisticated test involving time evolution of AA forcing. Such a test includes a model of the temporal character of anomalies as the basis of testing. Given the complex time evolution of signals in the climate system (noted by the reviewer), it is not clear whether a more sophisticated test would be more accurate than a t-test which implicitly includes time information via the use of the 1960-2009 period corresponding to maximum AA forcing.
 - The manuscript has been revised in response to the reviewer's questions. The revisions include a general discussion of the considerations used in the design of the t-test, to ensure its results were as robust as possible, as well as a specific change to the sample of years (1970-2009 increased to 1960-2009). These changes were included in the new subsection 2.3, and have improved the quality of the manuscript. The author is grateful for the reviewer's feedback on statistical testing.

2) The way to conduct the t-test is unclear. The data description states that the time series of wind-related losses has been smoothed by a -20year Butterworth filter 'to highlight the decadal variations in the results' (line 125). Does this mean that the t-test has been applied to the smoothed data? Also, the ensuing paragraph seems to indicate that the standard deviations of the series, needed to estimate the 'separation' of the mean values, have been calculated using the smoothed values. If this is the case, then the calculation is not correct, because the number of degrees of freedom (the sqrt(N) by which the standard deviation is divided) is not the number of time steps, but much less. Otherwise, one could artificially inflate the precision of the mean estimate, but smoothing the original data will. This needs to be clearly explained.

- Thanks for highlighting how the description of statistical testing is not sufficient to prevent such a concern about analysis methods. The manuscript has been revised to provide a more detailed description of the statistical tests (standard errors, and t-tests) in the new subsection 2.3. Specifically, the second paragraph of subsection 2.3 explicitly states the use of unfiltered annual means in standard error calculations: ‘... is computed using the entire timeseries of unfiltered annual means from each model’s Control integration...’. The final paragraph of the new subsection 2.3 describes the t-test, and also explicitly states the use of unfiltered data: ‘... unfiltered ensemble-mean loss anomalies for each year in the 1960-2009 period...’.

Particular points

3) *The explanation of what constitutes internal variability is not correct and should be thoroughly revised. For instance, the text states that ‘ Climate models consistently produce slower climate variations in the North Atlantic sector in the absence of all external forcings, which is referred to as internal variability. This sentence is misleading. First, internal climate variations are produced by models forced by constant forcing (not the absence of forcing); secondly, internal variability occurs at all time scales (from second-scale turbulence to millennial-scale ocean currents.*

‘This multidecadal driver can be broken down into two different components.’.

Internal variability is not a ‘driver’, and this sentence may lead to confounding internal variability with external climate drivers.

‘The first concerns the Atlantic Meridional Overturning Circulation (AMOC)’.

The AMOC is a term that describes a three-dimensional climate system. It is not ‘internal’ or ‘external’ variability per se. Actually, the manuscript later describes the impact of an external forcing (AA) on the AMOC. This paragraph is really not well-structured and sounds superficial.

‘ The second component is atmosphere-based, and consists of shorter-timescale extremes of sufficient magnitude to alter multidecadal averages of storminess’.

Again, the atmospheric variability may be internal or externally forced, at all time scales, not only the slow variations. Although high-frequency internal variability is also reflected in low-frequency variability (through sampling variability), some atmospheric processes generate intrinsic low-frequency internal variability.

Next, the manuscript cites EL Niño as an example of atmospheric internal variability, whereas it is very well known that ENSO arises through the coupling of the Tropical Ocean with the Tropical atmosphere.

There are many other aspects in this discretion that need revision. I really hate to be harsh, but this part of the introduction, in my opinion, needs a complete revision.

- The author thanks Referee 2 for highlighting potential for confusion in the discussion of internal variability. Their recommendation to revise the entire paragraph has been followed, and the points raised have been addressed, including:
 1. Defining internal variability in terms of *constant* external forcing
 2. Describing how storm variations emerge from internal variability (replacing the phrasing where internal variability *drives* storminess, to avoid confusion with external forcings which may drive internal variability)
 3. AMOC, QBO and ENSO are described as *modes* of internal variability

In addition, this part of the introduction is really not necessary for the manuscript's goal. I interpret the author as suggesting that AA may directly affect the atmospheric circulation, and that the AMOC, in turn,

modifies the meridional heat transport and thus the atmospheric circulation. Whereas the direct path is very rapid, the AMOC path may display a delay of several years, due to the sluggish response of the deeper ocean. But this is only my interpretation of what the author wish to say after reading the whole manuscript. The issue of internal or externally forced variability is, in my opinion, just steering the flow of the text from the relevant messages, even if it were properly phrased.

- The discussion of internal climate system variability is included in the Introduction to meet the aim of providing information on all the identified causes of multidecadal variations of European windstorms. This aim of the Introduction has been retained, because (a) wider context is generally considered good practice for an Introduction, and more specifically, such context helps to understand why later results on AA-forced signals do not fully explain full observed anomalies, and (b) managers of weather risk are a key part of the target audience in this study of windstorm losses, and a relatively concise overview of all contributory factors to multidecadal storminess is useful to people who oversee a broad remit of hazards and regions.
- The second sentence of the second paragraph in the Introduction has been replaced with two sentences, to highlight how the following four paragraphs are intended to provide an overview of the contributors to multidecadal storm anomalies, before the main study focuses on one specific forcing: 'The following four paragraphs briefly describe the evidence linking each of these time-varying forcings to multidecadal anomalies in storminess. These summaries provide wider context, prior to conducting a more in-depth analysis of one specific multidecadal forcing in this study.'

4) 'Observed multidecadal changes in European storm losses align with the AA-forced signals'

Is there a reference for this? I really doubt that this sentence, as written, can be correct, since wind-related losses would be primarily affected by GDP and/or population growth. Does the author mean 'normalised' losses? Beyond socio-economic factors, is there really evidence that AA is the main physical factor affecting wind-related losses, given the large internal variability of the atmosphere? If the sentence were correct, it would imply that atmospheric variations are almost entirely driven by aerosol forcing, which is not true. This sentence really requires a very solid backup and very careful phrasing (what does 'align' mean exactly?).

- The reviewer states that the sentence '*Observed multidecadal changes in European storm losses align with the AA-forced signals*' implies that '*atmospheric variations are almost entirely driven by aerosol forcing*'.
- The implication stated by the reviewer was a surprise, and definitely not the intended message. To prevent such a misunderstanding, the statement in the original manuscript '*Observed multidecadal changes in European storm losses align with the AA-forced signals*' has been replaced by the following three sentences in the revised manuscript: 'There is evidence to suggest the link from AA to the mid-latitude wind climate holds for European storm damages too. Cusack (2023) developed a reconstruction of windstorm losses from the mid-20th century to the present-day, with economic effects removed, and its documented multidecadal variations in European windstorm losses co-vary with temporal changes in AA forcing (e.g. Lund, et al., 2019). In summary, the findings from several recent studies indicate AA forcing played some role in the rising windstorm losses over the last few decades of the 20th century, and its subsequent decline this century.'

5) In brief, sulphates have produced the largest radiative forcing in the industrial period, growing from a relatively small amount at the start of the 20th century to peaks in the 1980s and 1990s'

*I guess the author means the largest *changes* in short-wave anthropogenic forcings. The largest radiative forcing is by far the sun. This is an example of inaccurate writing that can be found elsewhere in the manuscript.*

- Thanks. This inaccuracy is corrected in the revised manuscript: ‘... sulphates have caused the largest multidecadal anomalies of shortwave radiative forcing in the industrial period...’, and the author conducted a wider search to find and correct similar mistakes.

6) The contribution of socioeconomic factors to the wind-related losses is taken into account by multiplying the physical (wind) factors by the population in a particular grid-cell (equation on page 4, equations are not numbered!). It seems the population is considered constant over time (?). This would be quite unrealistic, as population and GDP growth over time would, I think, increase the exposure to wind extremes and thus dominate any trends in wind-driven losses. It would be rather easy to include population or GDP growth. Is there a reason not to do so? This point should at least be discussed.

- A new paragraph was inserted in the revised version of subsection 2.2 to explain why socio-economic causes of loss trends are removed from the analysis: ‘While populations evolve over time, we follow the common practice of using population data fixed to a common year when applying Eq. (2). This ensures all changes in modelled losses are caused by changes in storm wind climate, in isolation from socio-economic factors that drive changes in nominal losses.’
- Equations are numbered in the revised manuscript.