Reviewer #1

We thank the reviewer for their time and suggestions, which helped to improve the manuscript. Below, we include a point-by-point response to the reviewer's comments. Their comments are in black; our responses follow in blue.

General Comments

December-January Linkage: It's interesting that this combination was the only pair found to yield a notable statistical linkage. This appears to be at odds with other studies applying alternative approaches such as nudged hindcast experiments for attribution purposes. For example, Williams et al. (2025) shows a robust linkage between stratospheric conditions in January 2022 and the notably stormy February 2022 affecting the UK and northern Europe. Although only performed for one such case study, many other examples exist within the ERA5 record in which late winter storminess might conceivably be linked to the stratosphere (e.g., February 2020).

Could the result the authors find therefore perhaps reflect some limitation of the ACE2 model in failing to capture such relationship, maybe owing to the training algorithm or the number of years of training data considered? It would be good for the authors to discuss their results within the context of earlier published literature using such alternative approaches.

Response: The reviewer brings up a good point regarding late winter predictability. However, there is a fundamental difference between our methodology and the examples mentioned. Our research focuses on overall predictability rather than specific events. While studies such as Williams et al. (2025) and events such as February 2020 show clear linkages in those particular years, our goal was to identify a consistent statistical rule that holds across the multi-decadal record. The December–January period emerged as the only pair with enough statistical weight to be identified as a general trend in the ACE2 model.

The absence of a more robust February signal in our results does not mean those connections do not exist in individual years. Therefore, while individual events in other months may be predictable in specific years, our results suggest this does not hold as a general rule across the entire period.

However, as we considered this a good point, we will update the Limitations and future work section to clarify that, while ACE2 captures the most consistent signals, it may not capture the predictability of specific cases that deviate from the long-term average.

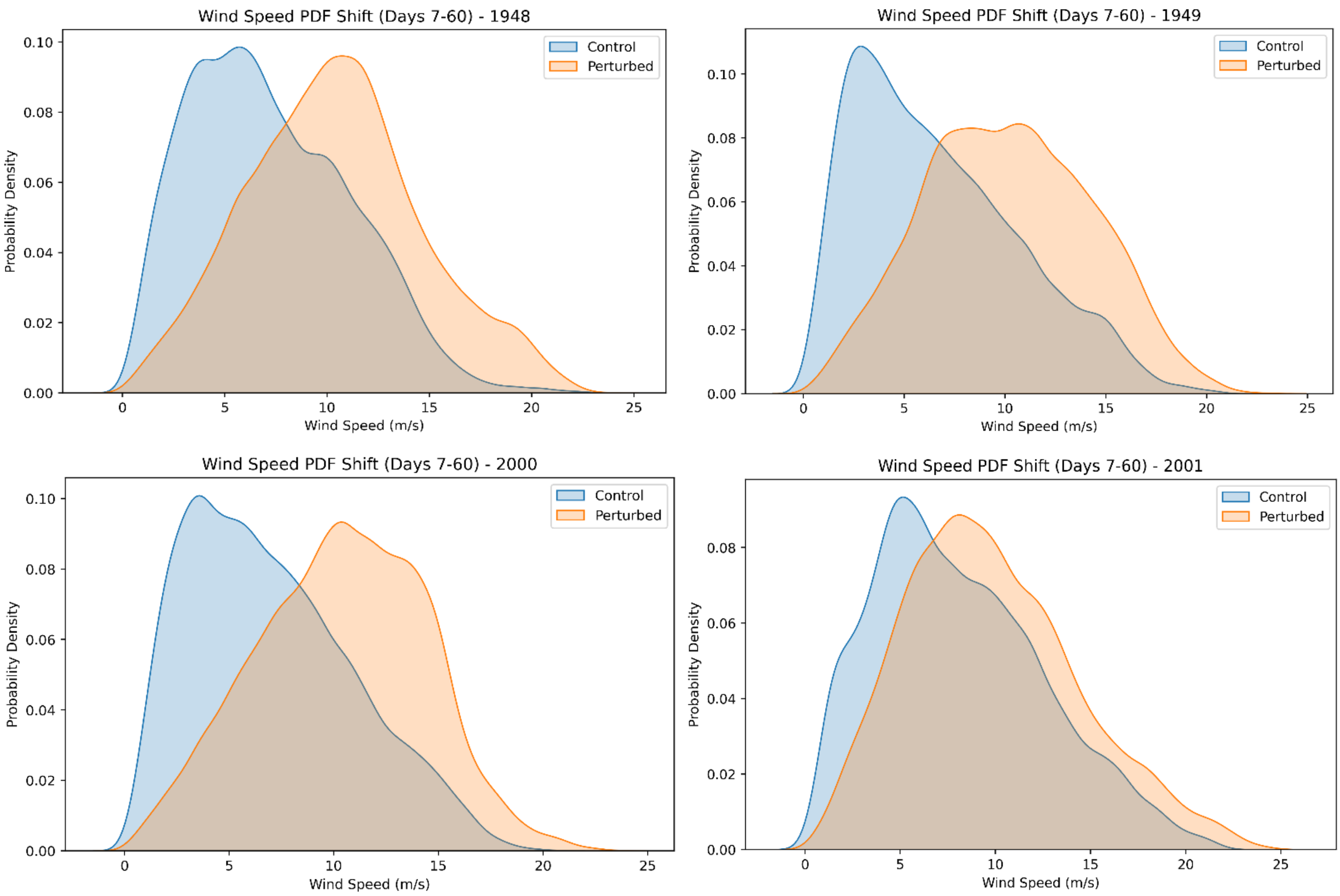
Figure 6: Whilst this Figure is interesting, I think it would be interesting for the authors to probe further the output from these modelling experiments. I would be interested to know how shifted the top 10th percentile of the distribution is in the perturbed relative to the control case. Is there a shift in the PDF of all the days between perturbed and control?

Again, by taking multiple grid cells into consideration, such evaluation becomes more powerful due to the increased sample size, and the result can be deemed more robust. So, I strongly encourage the authors to consider multiple grid cells, particularly if considering such analysis which would help increase the importance of the study.

Using days 7-60, the difference in the number of days above 90th percentile according to the control case, between perturbed and control, could also be assessed. What is the risk enhancement of a 90th percentile or above day due to the effective strengthening of the lower stratospheric polar vortex? Spatial maps could be an option for presenting this.

Response: We agree with the reviewer on this point. As we understand it, this comment has three key points. First, the reviewer is interested in whether there's a shift in the PDF between the perturbed and controlled runs. Here, we present figures showing the results of these modelling experiments, evaluated using PDFs of surface wind speeds in the North Sea region (Fig. 1). As requested, we considered multiple grid cells (lon, lat = 14, 10) to make the evaluation more robust. In all investigated years, we observe a distinct rightward shift in the PDF for the perturbed run relative to the control. Notably, the tail of the distribution (the top 10th percentile) exhibits a significant amount of increased density in the upper part of the tail.

In summary, these plots show a shift in the distribution of wind speeds over the North Sea, indicating that a stronger polar vortex consistently leads to higher average wind speeds and a significant increase in the frequency of extreme events. Specifically, they demonstrate that the "tail" of the distribution (the windiest days) becomes more populated, directly quantifying the enhanced risk requested by the reviewer. We will include these plots in the revised version.



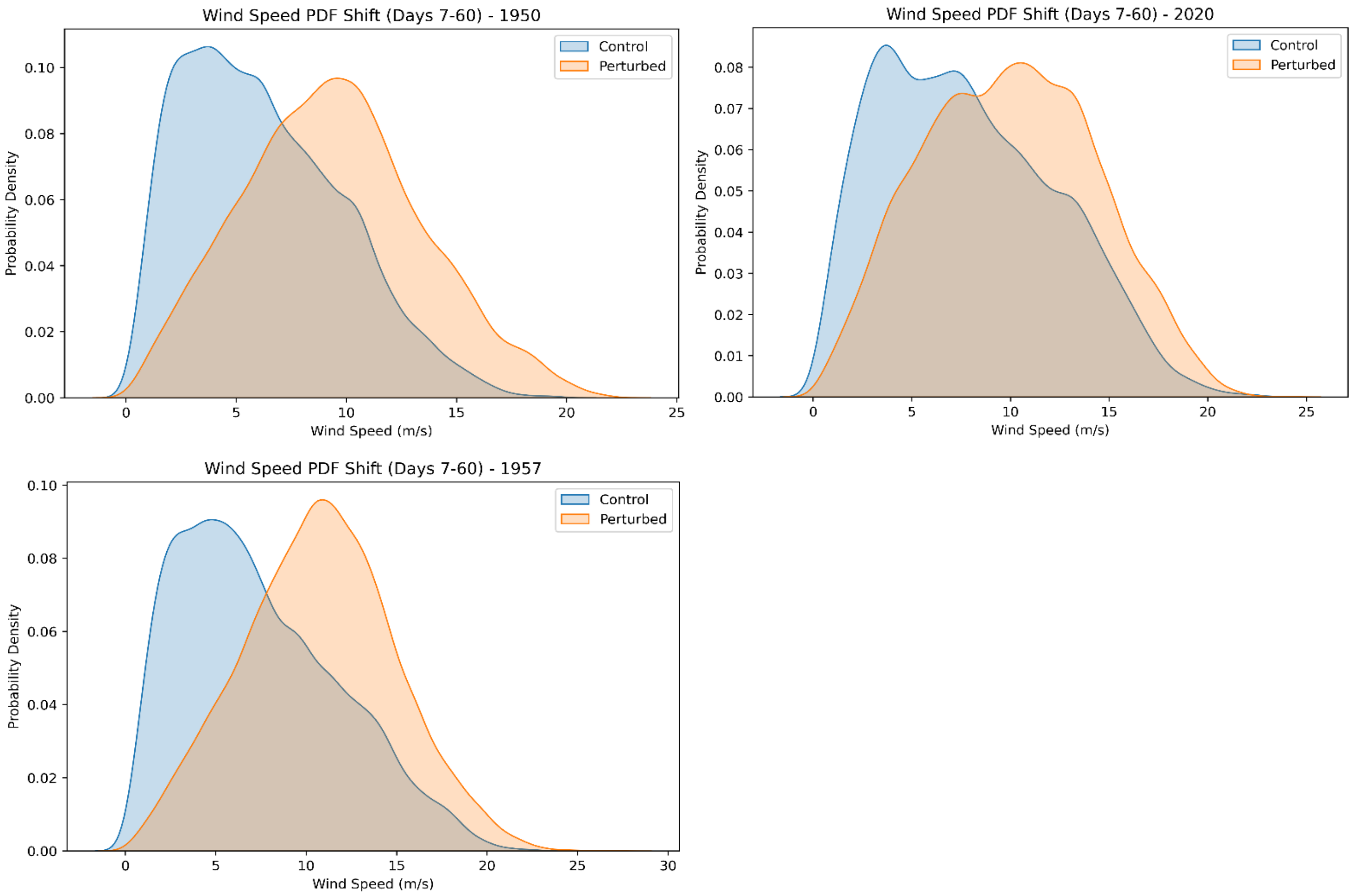


Figure 1: Probability density functions (PDFs) of 10m wind speed in the North Sea for days 7–60. Comparison between the controlled (blue) and perturbed (orange) initial conditions across seven ensemble years

Second, the reviewer recommends comparing the number of days exceeding the controlled 90th percentile between perturbed and controlled experiments for days 7–60. This would help quantify the increased risk associated with the strengthening of the lower-stratospheric polar vortex, possibly illustrated using spatial maps.

This plot (Fig. 2) represents how a stratospheric perturbation alters the frequency of extreme wind events by defining extremes using the control climate, counting exceedances in the perturbed climate, and mapping the resulting percentage-point change. The approach isolates changes in the tail of the wind speed distribution, providing a physically interpretable and statistically robust measure of extreme event risk rather than mean-state changes. We observe that, in most regions of the North Sea, the frequency of extreme days is increasing. Moreover, the frequency of extreme wind days is increasing, peaking at 2%-6% along the coastlines of Denmark, northern Germany, and southern Norway. This suggests that a stable polar vortex affects the storm track or enhances storm intensity specifically within this region. While most of the regions show an increase, parts of the eastern UK coast show a slight decrease or a neutral response.

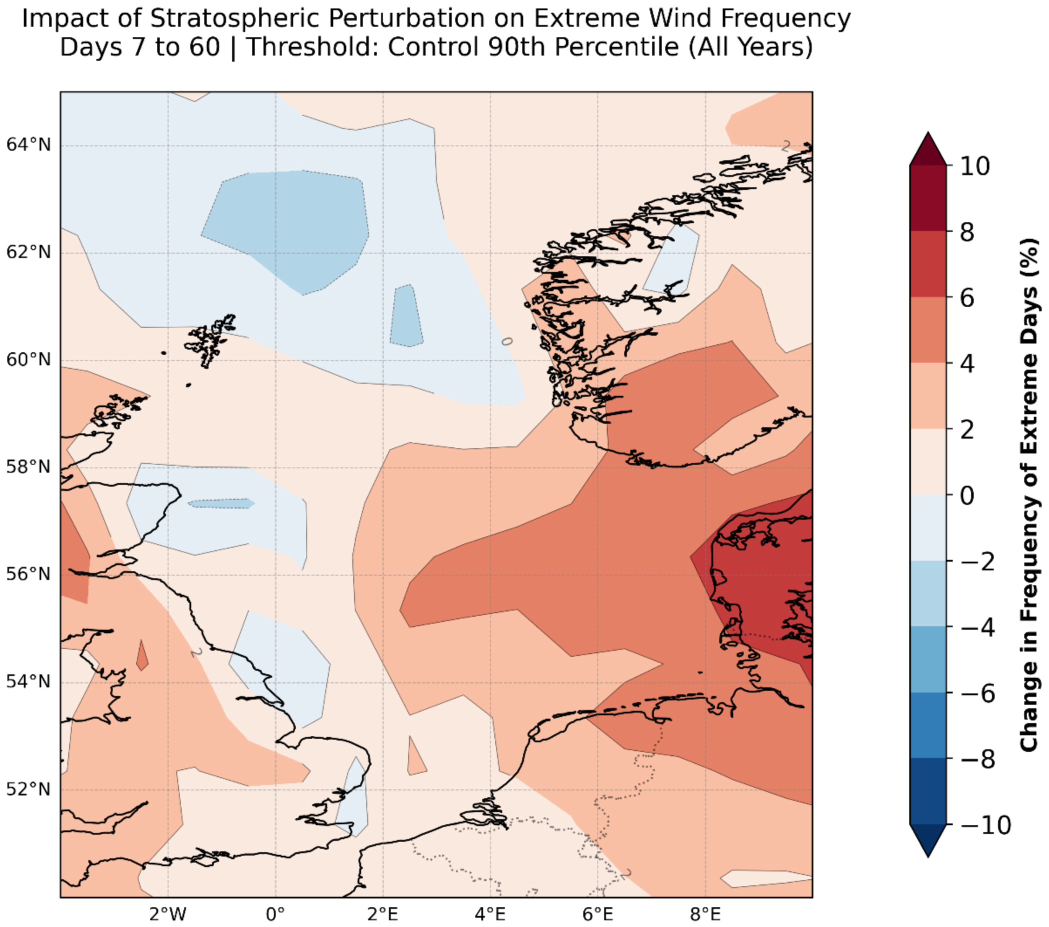


Figure 2: Impact of Stratospheric Perturbation on the Occurrence of Extreme Wind Events (Years: 1948, 1949, 1950, 1957, 2000, 2001, 2020) | Threshold: Control 90th Percentile wind speed

This Table estimates how much the risk of extreme (top-10%) surface wind events increases or decreases under a perturbed experiment, relative to a control climate, averaged over a specific European region. We observe that the risk of extreme surface wind events increases in all the selected years.

Year	Increase %
1948	4.09
1949	3.77
1950	3.42
2000	4.08
2001	3.71
2020	4.04
1957	4.15

Table 1: Risk Ratios for 90th percentile wind speed events in the North Sea region across seven ensemble members

ACE2 Experiments: If the experiments are fast and easy to run, would it have made sense to consider other initialisation dates besides December 1st? A lagged ensemble of dates could have been (or perhaps be considered in a future study) by maybe taking several of the initialisation dates either side of 1 December.

However, in doing so, the each initialisation date (ensemble member) cannot be considered fully independent?

Response: As mentioned in the abstract, sensitivity emulations initialised at the start of November did not show a meaningful response in the ensuing month. In line with Reviewer 2's recommendation, these results will be presented in the revised manuscript.

But otherwise, this suggestion opens a compelling path for follow-up research, as our current study focuses specifically on the seasonal dynamics between December and January. Because of that focus, we intentionally applied the initialisation mask to the December 1st fields, after the statistical analysis showed no useful predictability for early or late winter overall. This does not necessarily mean that individual late-winter events are highly predictable, as the reviewer notes in other comments, but the data suggest this is not the rule. In principle, the reviewer's comment would lead us to fine-tune the dates of maximum predictability throughout the winter season, but we believe this would substantially expand the manuscript. Exploring the broader implications across the weeks in December and January would certainly be a valuable next step, and we take this as a valuable comment for our next research.

Specific Comments

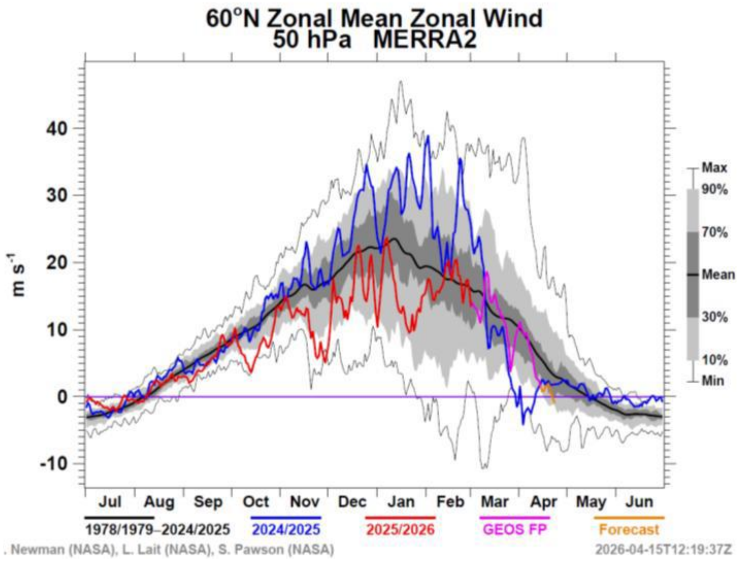
L11-12: If emulated mean January surface wind speeds increase following modified 'stratospheric initial conditions on 1st December', what exactly is this in response to? Decrease in 70 hPa air temperature and/or increase (decrease) in zonal (meridional) wind?

Or is this just an upper bound concerning possible surface wind speed increase that could result? I think it should be clarified.

Response: The initial conditions are changed using spatially resolved anomaly patterns. These patterns were derived from the composite analysis explained in the main body of the manuscript. We will change the sentence in the abstract.

L23-24: The intensity of the polar vortex most commonly attains maximum strength in January but interannual variability is largest from January to March (e.g., below).

So, I'm not sure this sentence is quite correct. Downward coupling of stratospheric anomalies to the surface tends to maximise in mid-to late-winter in accordance with larger swings in the strength of the polar vortex?



Response: L23-24 was: "This higher skill is likely linked to stronger stratosphere–troposphere coupling between November and January, as polar vortex anomalies develop and begin to descend toward the surface."

We will change the sentence to "This higher skill is likely linked to stronger stratosphere–troposphere coupling that peaks in late December and extends until mid-February, as polar vortex anomalies develop and begin to descend toward the surface."

L54-55: 'Scaife et al. (2014) used a new long-range forecast system…' → Can it be considered that 'new' given the study is now over 10 years old?

Response: It will be rephrased.

L72-76: Citations missing here. I suggest the authors expand more upon the complexity of the role of ENSO on European winter weather, including its evolution from late autumn into late winter/early spring.

For instance, Thornton et al. (2022) which discusses the importance of an early winter teleconnection via a tropospheric pathway that project strongly onto the East Atlantic Pattern, whilst the stratospheric teleconnection pathway is relatively inactive (Ineson and Scaife, 2009). The role of ENSO changes over winter (in part due to the increasing role of the stratospheric pathway), which projects more strongly onto the NAO compared with the EAP into winter.

Response: We thank the reviewer for this insight. We will expand the introduction a bit according to these suggestions, keeping in mind that the focus of our study does not lie on the ENSO-stratospheric teleconnections. This point is, however, also related to the discussion in the manuscript of the role of large-scale SSTs on North Sea storminess, and to the fact that storminess in the North Sea does not show any persistence at monthly timescales, i.e. January and February storminess are statistically independent. Although this does not invalidate the ENSO-storminess teleconnections, it does indicate that the ENSO impact on storminess specifically (not on the mean seasonal circulation) may be just too weak.

L73-74: This may be a true result according to reanalyses such as ERA5, but it has been shown that this result might be affected by a limited sample of years. See Ineson et al. (2024) for example, where they provide evidence using hindcast experiments that this finding might be due to chance.

With a bigger sample of years, a difference in SSW frequency between El Niño and La Niña states would perhaps emerge, in accordance with the known ENSO impact on the boreal wintertime polar vortex strength.

Response: We can expand this section by adding the reviewer's suggested reference (Ineson et al. 2024). This citation notes that while reanalysis datasets such as ERA5 show a higher frequency of SSWs during both ENSO phases than under neutral conditions, recent large-ensemble modelling suggests this could be an artefact of the limited observational record. Specifically, these models suggest that a more linear relationship might emerge with a larger sample size.

L112-114: Here, other such approaches could be mentioned, including stratospheric nudging protocols, such as SNAPSI, for assessment of predictability and attribution (Hitchcock et al., 2022).

Response: As also per the comment of reviewer #2, we will include a discussion of the SNAPSI experiments.

L116-120: The content of some but not all sections are explicitly outlined here. What can the reader expect in the rest of Section 4 and Section 5?

Response: We will revise this paragraph to include more detailed information on each section and what the reader can expect.

L133-134: Some further information of the 90th percentile definition would be beneficial. For instance, is this estimate daily evolving, or calculated using all days across the month? If daily evolving, did the authors experiment in anyway by applying moving average or filters to optimise this? If not, was it tested that the difference between the 90th percentile from day to day is somewhat gradual in its evolution (if it jumps about a lot from day to day then this would suggest artefacts due to sampling could be influential)?

Although I appreciate the results are unlikely sensitive to the exact choice picked, given the relatively large sample size considered?

Response: We thank the reviewer for the request for clarification. The 90th percentile was calculated using all daily wind speed values for the specific calendar month across the entire study period (all January days from 1940 to 2024). This provides a single, stationary threshold for that month, avoiding day-to-day jumps or sampling artefacts. Given the relatively large sample size from the ERA5 record for a full month, this estimate is statistically robust and reflects the general storminess threshold for that time of year. We will add this information to this paragraph.

L135-138: I am curious as to why the authors decide to evaluate one grid cell only and not consider multiple grid cells across a wider area. Was this opted for due to computational limitations? It is not clear what the justification is for doing this.

Response: We thank the reviewer for this point. The choice of a single grid cell is justified by the high spatial coherence of the storminess signal in this region. As shown in Figure 2 (manuscript), the Pearson correlation coefficients between the chosen cell and the wider North Sea basin are remarkably high (r >= 0.7 over a large area). This indicates that the chosen location acts as a reliable centre of action for regional wind variability.

However, we chose to examine multiple grid cells near the coast to analyse stratospheric temperature and zonal wind anomalies. By creating these composites, we wanted to see whether the polar vortex is more stable in December months that lead to stormy Januaries. We added two more grid cells covering the UK and German coasts because these areas are near land, where the greatest societal and economic impacts occur, reflecting the concerns raised by the author.

As for the UK coast, we chose the grid near Hull, a port city in East Yorkshire, England. The second grid is chosen along the German coast, near Bremerhaven (Fig. 3).

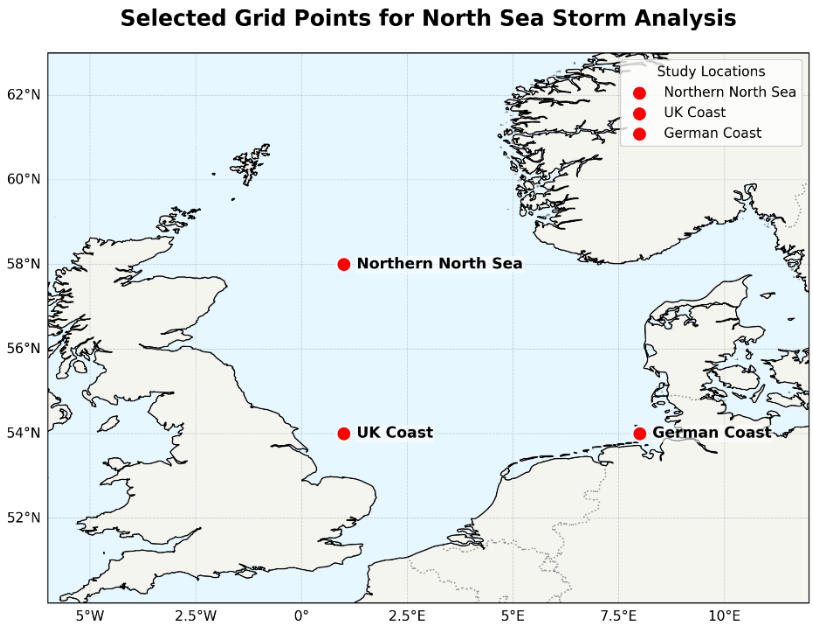


Figure 3: Selected grid points for North Sea storm analysis. The Northern North Sea is the grid we have already explored in the manuscript; the UK and German Coasts are the new grid cells that we chose for the revised version.

After computing the January storms for both newly selected grids, we found that the five stormiest Januarys were associated with a comparatively colder and windier (westerly) polar stratosphere in December (Figures 5-8). This observation clearly supports our main hypothesis for this research: that the stormiest Januaries are preceded by a stable polar vortex. We will add these figures to the manuscript and provide brief descriptions.

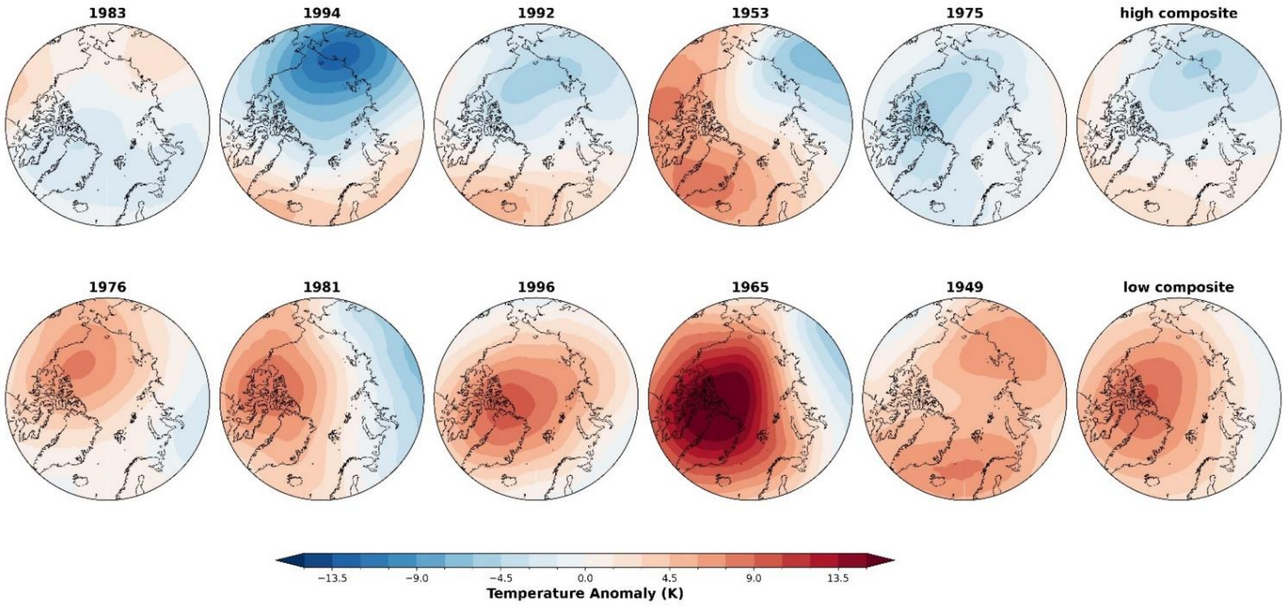


Figure 4: Temporal mean of air temperature at 70 hPa in the Decembers prior to Januaries with a high (upper row) and low (lower row) number of storms **(UK Coast)**

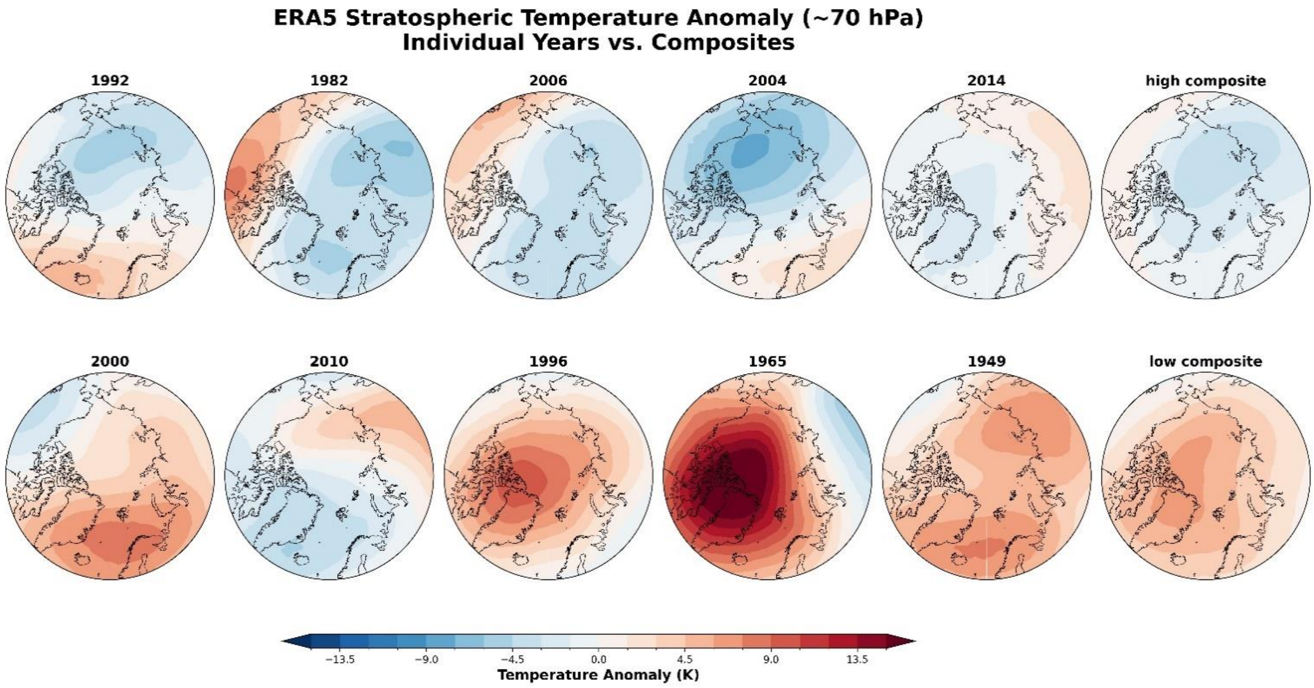


Figure 5: Temporal mean of air temperature at 70 hPa in the Decembers prior to Januaries with a high (upper row) and low (lower row) number of storms **(German Coast)**

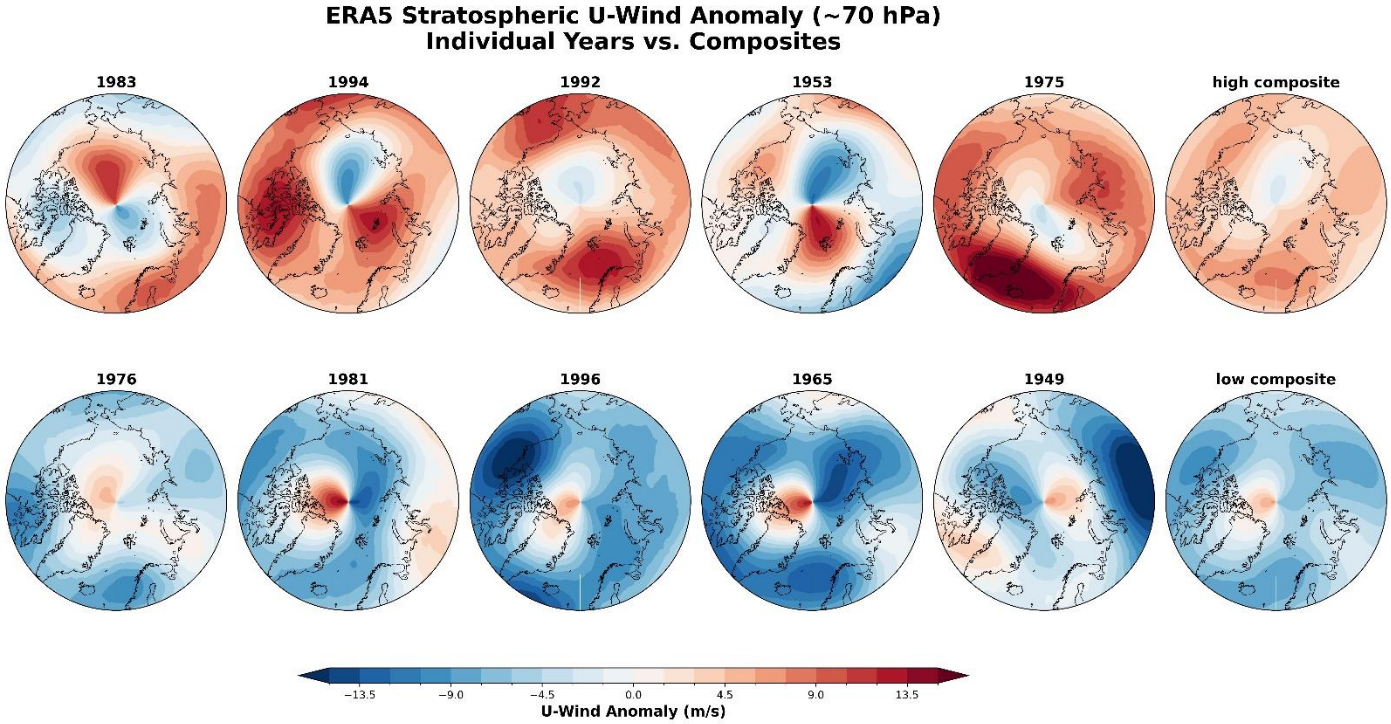


Figure 6: Temporal mean of zonal wind speed at 70 hPa in the Decembers prior to Januaries with a high (upper row) and low (lower row) number of storms **(UK Coast)**

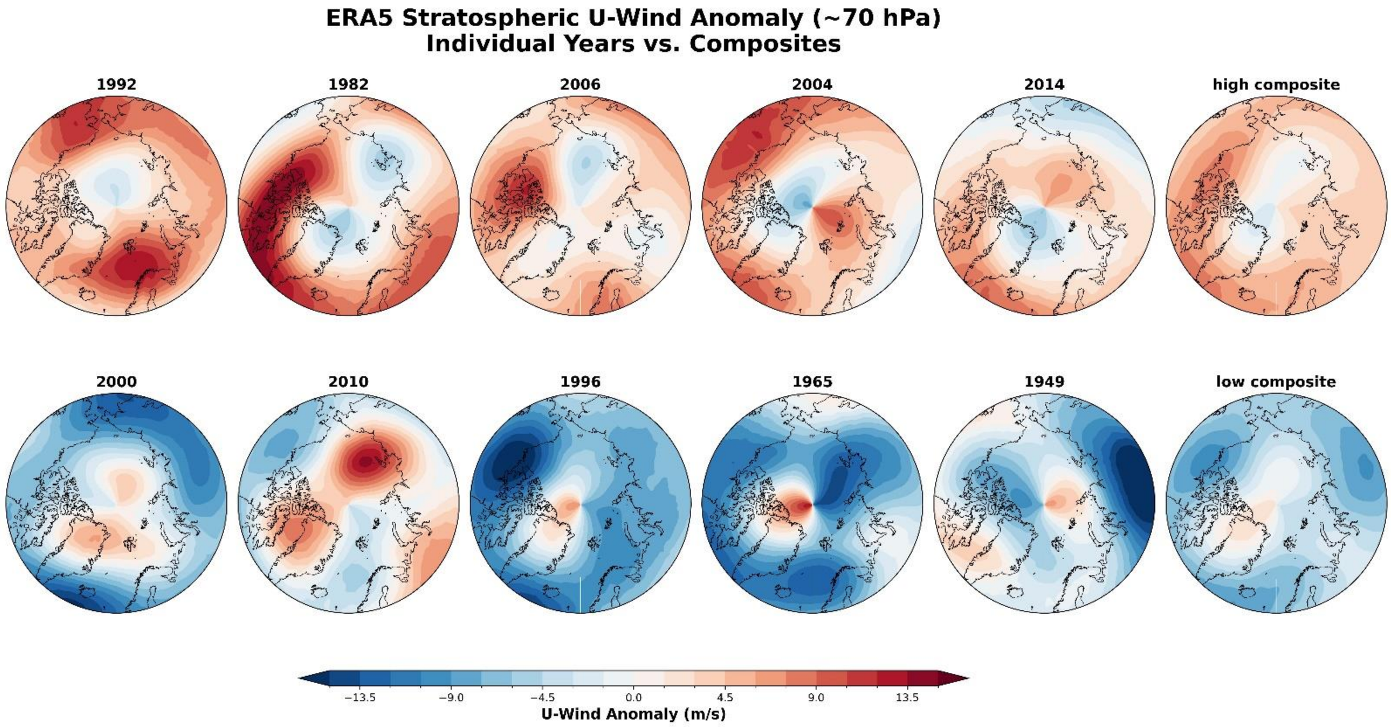


Figure 7: Temporal mean of zonal wind at 70 hPa in the Decembers prior to Januaries with a high (upper row) and low (lower row) number of storms **(German Coast)**

L147-149: Shown in Fig. 1b? Mentioning 'global map' confused me at first.

I have a slight concern that the short 'coherence length' implies the result reported in this study is not necessarily applicable to the whole North Sea basin but rather only a limited central northern part of the North Sea? Or is this unfair?

Response: We will change this part, and we hope that the previous response answered this concern.

L161: '…configurations of SSWs…' → What is meant by this? Different configurations of vortex morphology or perhaps certain types of events (e.g., split v displacement or non-downward/downward propagating etc.)?

Response: By "configurations", we referred to the vortex morphology (split vs displacement events). As noted by *White et al. 2021,* among the two types of SSWs: displacements and splits, vortex split events are often linked to more persistent and robust surface pressure anomalies compared to displacements. We will update the manuscript to clarify that the regional coherence of storminess likely depends on these specific stratospheric states rather than a generic intensification of polar vortex dynamics.

L163-164: It would be good for the authors to expand on this. Dependency on prior weather regimes driven by intraseasonal modes of influence such as the MJO? MJO is just one example but a good one to mention in this context as studies suggest stratosphere-troposphere coupling is strongly MJO dependent (e.g., Yadav et al., 2024).

Response: The reviewer brings up a very good point. We evaluated the MJO using the paper the reviewer suggested (Yadav et al., 2024) and examined the timeline of MJO occurrence. The MJO affects the NAO through both tropospheric and stratospheric pathways. Enhanced convection over the tropical Indian Ocean (Phases 2-3) increases the chance of a positive NAO via a tropospheric Rossby wave response and a strengthened stratospheric polar vortex, whereas convection over the western Pacific (Phases 6–7) enhances upward planetary wave propagation, weakens the polar vortex, and favours a negative NAO. The NAO response is stronger and more persistent when the stratospheric pathway is activated, particularly following strong MJO Phase 6–7 events associated with sudden stratospheric warmings.

However, we think that the state or phase speed of the MJO cannot explain the deviant behaviour of North Sea storms in winter 2001-2002. The causality chain indicated in the papers mentioned in the previous paragraph indicates that the MJO can affect the polar vortex and the state of the NAO. Particularly storming convection activity in the Indian Ocean (phase 3 of the MJO) is associated with a stronger NAO (Henserson et al., 2016). This is what happened in the boreal winter 2001-2002. The NAO index in December 2001 was mildly negative but strongly swung to positive in January and February 2002. Also, December 2001 witnessed SSW, but this warming did not lead to a weaker vortex in January. A possible explanation for the behaviour of the polar vortex is indeed the disturbance of the intense MJO during January and February 2002. However, the very strong NAO in January 2002 did not lead to a very stormy January in the North Sea, and so the chain of events to the surface climate breaks here. The reviewers' suggestion naturally leads to a follow-up study in which the December polar vortex and the January convection in the Indian Ocean are simultaneously distorted. The lag between the MJO and the NAO does not seem particularly useful for seasonal prediction in general. However, in the case of 2001-2002, the NAO remained persistently and strongly in the positive phase in January and February, so there may be a case where the reviewers' suggestion naturally leads to a follow-up study in which the December polar vortex and the January convection in the Indian Ocean do not seem to be a possibility.

We will expand this discussion and add the revised manuscript.

L171-173: I think it would be good to expand by acknowledging however that the spatial pattern of the anomalies differs markedly between across each of the top 5 stormiest and no-storm Januaries.

Response: We will expand this sentence.

Figure 3 caption: A good idea to state that these anomalies correspond to high and low storm number Januaries for the region of the North Sea of focus in this study.

Response: We will change the caption.

L176-178: Figs. 4 and 5 are for both zonal and meridional wind but only zonal is mentioned?

It seems that the meridional wind component anomaly pattern is less important in terms of difference between most and less stormiest Januarys, which is not surprising. Although it is interesting a negative region tends to emerge in the North Atlantic sector (anomalous northerly flow and greater tendency towards cyclonic pattern over Scandinavia)?

Response: We think that the reviewer means the positive (southerly) meridional wind anomalies seen in the composites over the North Atlantic-Barents Sea. Indeed, this is also a persistent pattern in almost all cases shown in this figure. We will also add a brief description of the meridional wind.

L179: 'We also include the December monthly mean of geopotential height anomalies at 200 hPa…' → I don't see it anywhere. Did the authors mean to include it?

Response: We did not include the geopotential height anomalies. We will delete this paragraph about geopotential height.

L223: '…ensemble members…' → How many were considered?

Response: We think this sentence is incorrect, so we will change it. For each year, there was only one perturbed and one controlled run, and we added the anomaly mask from December 2015 to the December 1st fields of the perturbed runs.

L230-235: Again, I'm left wondering how many members are considered for each of the "control" and "perturbed" ensemble. Is it 6 for each, taking the Decembers shown in Figures 3-5 on the bottom row? There is only 5 for such cases shown, however.

Or were the initial conditions for each December from ERA5 also perturbed (if so, it is not explained)?

Response: We clarify that the term *ensemble* was used incorrectly in the original manuscript. Our study does not utilise a multi-member ensemble. Instead, we performed pairs of deterministic simulations: a control run and a perturbed run.

For each year, the control run is initialised with unmodified ERA5 data on December 1st and integrated for 60 days. The perturbed run uses the same December 1st initial conditions, but with the addition of stratospheric anomalies derived from December 2015, a period chosen for its representative stable polar vortex configuration. We will revise the text to replace "ensemble" with "control run" and "perturbed run," and explicitly detail this initialisation process to avoid further confusion.

L267: '…years with missing data…' → What missing data? Surely ERA5 is continuous for the predictor variables considered over the full record (1940-2024)?

Response: We wrote this sentence incorrectly. There was no missing data. We will change this sentence.

L278: '…1948, 1949, 1950, 2000, and 2020…' → How were these 6 Decembers (also including 2001) decided upon for analysis? I note that they are different apart from 1950, as the identified five least stormy Januaries, shown in the bottom row of Figures 3-5.

Its not clear to me why different Decembers were used to those identified earlier. The authors need to add more information to explain their reasoning.

Response: We acknowledge that the selection criteria and the presentation of these years were not sufficiently clear. These specific years were selected as initial conditions because they preceded "relatively non-stormy" Januaries, which we define here as months with only 1–2 storm events. This allows us to test whether imposing stratospheric anomalies can trigger or increase surface wind speeds during quiet periods.

We also thank the reviewer for spotting the inconsistency in our figures. The inclusion of 1950 was an error, as January 1951 had zero storms and did not meet the 1–2-storm criteria. We have removed the 1950 figure and replaced it with 1957, which preceded January 1958 with exactly one storm. It is also worth noting that 1957 shows a moderate increase in wind speed from 0.5 to 1.5 ms^{-1}. We will update this section to explicitly state this selection logic and ensure the text aligns with the revised figures.

L282: 'December 2001' → In what way is this a slightly different situation to the other five? More storms occurred in January 2002? Needs explaining.

Unless the authors mean to highlight this case as an example of which the perturbation applied to a December preceding a January largely absent of storms just so happens to result in a decrease in mean surface wind (perhaps pointing to the role of internal variability which can still occasionally result in the opposite-signed response?). In any case, it's unclear to me.

Response: See also our response to point L163-164. We will expand on this point. Indeed, the reviewer's interpretation is correct, and the year 2001 does not fit with the view that stratospheric temperatures are a good predictor of storminess. This is not entirely surprising, since it is unlikely that a single factor can explain 100% of the variability, and we also wanted to show a case where it does not work. Regarding the reviewer's previous comment, we also examined the Madden-Julian oscillation and, depending on the results, include it as a possible modulating mechanism of stratosphere-troposphere coupling.

This point is also related to the representativeness of our North Sea grid cell. All three North Sea grid cells in January show an average number of storms, although an SSW in December nudges the circulation towards fewer storms. According to Yadav et al. (2024), a slow MJO episode from late October 2001 to early January 2002 led to an SSW on 30 December 2001, but subsequent strong convection over the Indian Ocean could have nudged the NAO towards a strong positive value in January and February. Additionally, data from NOAA CSL: Stratospheric Modelling & Analysis: SSWC include a table of major midwinter SSWs compiled by Dr A.H. Butler. We identified two SSWs in December 2001 and February 2002. Our interpretation is that the SSW and the MJO had opposing effects that winter, leading to an unremarkable number of storms in January, despite the rather strong NAO index during those months. We will add a summarised version of this explanation in the manuscript.

L283-284: 'We only see a modest increase, mainly south of Norway in the North Sea.' → But more widely a decrease of ~0.5 to 1 m s-1?

Response: We will revise the manuscript to provide a more balanced description of the spatial anomalies, acknowledging the widespread decrease of approximately 0.5-1 ms^{-1}.

L285: 'There are also other contributing factors, but they are not addressed in this study.' → This sentence reads as though the authors know what these were. I think its not so straightforward to necessarily know what these were, but examples of suspected factors could be highlighted (e.g., did an active phase of the MJO occur prior to January 2002 that would counteract the role of the strong polar vortex?).

Response: No, we were not referring to any particular factors. We were just arguing that stratospheric temperature cannot fully explain storminess variability, and that storminess variability in the North Sea seems disconnected from storminess in other regions, so there must be other factors, and those factors will probably need to be regional. We will discuss the possible role of the MJO (see previous comments).

Figure 6: Why is there a difference in domain considered compared to Fig. 2? I think the rationale needs explaining.

I'm not sure how lacking in storm activity the subsequent Januaries are apart from following December 1950 (which was earlier stated as containing zero throughout the month). If alternative Decembers to those shown in Figs. 3-5 were picked, it would be good to state the counts of storms for these (are they non-zero, albeit a low number?).

Perhaps the biggest wind speed increase is therefore an unsurprising result in the perturbed case following Dec 1950 which had zero storms?

Response: Figure 6 lacks geographical coordinates, so we will remake it to show the whole North Sea. We will state the observed storm counts for January's, as shown in Figure 6. We will also remove the 1950 case and replace it with 1957 (see previous comment L278).

Figure 7: Would it not be meaningful to display what years 0 to 25 actually correspond to on the x-axis?

Response: We will change the x-axis.

Figures 7-8: It is explained that Fig. 8 is for the test set of data (30 %). Given the correlation is the same for both, I assume Fig. 7 is also for this same test set of data. It is not explained for Fig. 7.

Response: Yes, they are the same test set. We understand how it can seem unclear, and so, we will explain it a bit more.

L305-309: It would be helpful to describe each row in turn or re-order the rows in Fig. 10 in accordance with the text.

Response: We will fix that.

Figure 10: I would suggest inverting the colour scale to match the direction of the PC loading according to its physical representation. In other words, strong vortex favours negative 200 hPa GH (blue), positive U at 70 hPa (red) and negative TA at 70 hPa (blue).

Response: Very good point. The colour scale will be reversed.

L311-312: I suggest making it clearer that this is a typical response over northern Europe, but not necessarily for central and southern Europe. For instance, Afargan-Gerstman et al. (2024) showed increased storm activity following SSWs here, using subseasonal reforecast experiments (e.g., their Fig. 3).

However, the footprint of strongest wind speeds tends to occur on the equatorward side of the main storm track. So, following weak polar vortex (or SSW) events, regions such as Iberia are at increased risk of surface wind extremes (e.g., Ayarzagüena et al., 2018; Utrabo-Carazo et al., 2024).

Response: We will add this part about central and southern Europe.

Table 2: The correlation values should always be closer to 1 or -1 than R2. The values only seem to match up for Nov-Dec and Dec-Jan? The correlation for Nov-Dec is not expressed as a range. Needs fixing.

Response: We will fix it.

L328-330: I'm not sure this reflects the finding in Kidston et al. (2015). I would argue their Fig 1a implies the most pronounced surface impacts occurs in January-February (i.e., towards late winter). This fits with the relatively inactive stratospheric pathway in early winter reported upon, at least associated with ENSO (Ineson and Scaife, 2009).

Response: We think that the reviewer is to some extent correct and that our interpretation of Fig 1a in Kidstone et al. (2015) needs to be reformulated, but it is not in contradiction with our results. Kidstone et al. (2015) show that the persistence of stratospheric circulation anomalies peaks in mid-January and extends into mid-February (with persistence of about 40 days). We find, however, that stratospheric temperatures in January are, on average, not a good predictor of surface storminess in February. Our results, therefore, require an additional step to the persistence of circulation, namely, the impact of those circulation anomalies on the surface. Our analysis, if correct, shows that this second step is missing.

We will reformulate this paragraph more carefully to address this question, which is subtler than we initially thought.

L371: It would be good to elucidate other likely factors. Were there any particular drivers relevant during the early 2001/2002 winter that might have favoured more quiescent conditions in the North Sea region? The importance of internal variability, which is inherently large, should also be stressed.

Response: See also the responses to points L163-164 and L282. We looked into the role of the Madden-Julian Oscillation, as suggested by the reviewer in another comment, and examined storminess behaviour in other North Sea grid cells, as this may provide hints about other possible factors. The other grid cells (UK Coast and German Coast) selected to validate the one-grid-cell approach also show low storm counts (1 storm in UK Coast, 3 in German) in January 2002.

Technical Remarks

L33: '...perhaps easier...' → '...more skilful...' perhaps?

L62: 'SSW...' → 'SSWs...'

L69-70: 'That study...' → 'Kidston et al. (2015)...'?

L71: '...making an accurate representation of the stratosphere essential...' → 'making an accurate representation of the stratosphere also essential...'

L74: 'SSW...' → 'SSWs...'

L140: 'e.g.' → Suggest this and rest of sentence is put in parentheses for readability.

L159: 'SSW...' → 'SSWs...'

Figure 10 caption: Suggest reordering mention of each variable in turn down the page (i.e., GH, U and TA).

L316: '...introduction 1' → Delete '1'.

L363: '?' → Should there be a missing citation that needs adding here? If not, delete.

Response to all technical remarks: Will be corrected.