Review of 'Identification of high-wind features within extratropical cyclones using a probabilistic random forest - Part 2: Climatology' by Eisenstein et al. submitted to Weather and Climate Dynamics
10.5194/wcd-2023-10

General comments:

This is an interesting study describing the results of applying a methodology previously published by the authors to develop a 19-extended-winters climatology of mesoscale wind features in cyclones over Western and Central Europe. The topic is very suitable for this journal. The research is novel in the generation of this type of climatology over this region, and in the method used to obtain the climatology. I have a few questions about the details of the methods (some of the description is confusing, at least to me) and about the robustness of the results. The work would also benefit from a more detailed comparison with previous literature to draw out the novelties of the findings.
The paper is well presented apart from some small language glitches (not all of which I've likely commented on in my review).

Thank you for the thorough review of our manuscript and many valuable comments that help to improve our paper. We have been carefully considering each comment, which are addressed in blue below. Text changes and more details will be provided in the Author's response.

Major specific comments:

L105 Here it is stated that the focus of this study is on cyclones occurring in Western and Central Europe. This geographical focus of the study should be included in the paper title.
Thank you for raising this. We add "over Europe" at the end of the title.

Introduction Given the focus of this paper on the different sources of strong low-level winds, it might be helpful to include a diagram showing where these are located within a classic cyclone as a first figure. The different features are explained in the text but it may be difficult for less knowledgeable readers to visualise these. Related to this comment, some more detail on how the features are defined in probabilistic characterisation would be helpful. I realise that these details are described in the part 1 of this paper but it should be possible for readers to follow part 2 without having first read part 1.
We understand the concern but given the easy and open access of Part 1 we do not think it is necessary to reuse its Figure 1 in Part 2. However, we specifically refer to Figure 1 of Part 1 in the Introduction now and add a few sentences about the distinction of the features in Section 2.3

L73 Here you refer to the Earl et al. study. One interesting result from that study is the different of the prevelence of the features changes depending on whether those generating the top 1% or 0.1% of daily maximum wind gusts are considered. Did you consider changing your windspeed threshold to look at how the relative importance of the features changes?
This is indeed an interesting result to add. Furthermore, we add pie diagrams with higher wind thresholds and a corresponding discussion to Section 3.1 showing an increase of CJ and CFC occurrence with higher thresholds (see response to Section 3.1 below).

L203 "the proportion of NF is considerably lower in COSMO-REA6 data" – this statement is true but "no feature" strong winds still account for about 30% of the strong winds. Please can you say more about what these strong winds are likely to be caused by and whether they are likely to be linked to cyclones? How often are these NF strong wind regions close to the tracked cyclones (all strong wind regions are within 15 degrees of a cyclone (L167), but are the NF regions relatively more likely to be further away (e.g. beyond 7.5 degrees)?
NF winds are commonly associated with weaker winds (see our response to your next comment and the added Fig. FR2) and, as already mentioned in the preprint, NF detects parts of the CCBa. Looking at a relative version of the system-relative composite (number of occurrences divided by the number of windy conditions → Figure FR1), the contours of NF shift to the north-eastern quadrant,

outside the area of the other features, which show less change. We include this figure in the Appendix.
Furthermore, Part 1 showed that for some cases the spatial independence leads to uncertainty in the area where the introduced features occur, resulting in NF being the most probable features (see Section 7.4 of Part 1 for more details) or the area between double fronts shows uncertainty (see Section 7.1). This is not commonly the case; however, it might explain the overlap in the southern quadrants to some extent.

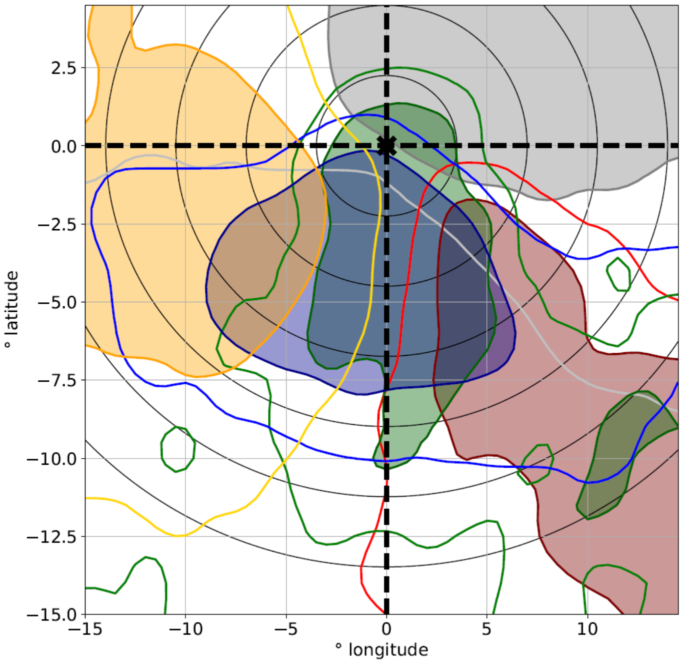

Figure FR1: As Fig. 6 but divided by the number of windy conditions.

Section 3.1 Here the pie charts showing the relative frequency of the different strong wind features are discussed with the pie charts showing percentage occurrence values to 1 decimal place. It would be useful to have some sense of how robust these relative percentages are to small changes in the datasets used and hence how meaningful the interpretations are. For example, what happens if some of the years are omitted from the datasets used, the definition of strong winds is changed slightly, or the domain of interest (the blue box in Fig. 2) is shifted slightly?
This is a valid point, and we had a further look into this. First, we add pie diagrams for higher wind thresholds to the main text (Figure FR2) showing that the proportion of NF decreases significantly, while especially the proportions of CJ and CFC increase consistently to Earl et al. (2017). Furthermore, three sets of ten randomly chosen winter seasons show only slight changes in proportions (Figure FR3), which is expected given the relatively small fluctuations shown in Fig. 4. A shift in the spatial domain would yield different proportions as can be seen in the Earth-relative plot in the preprint (Fig. 5).

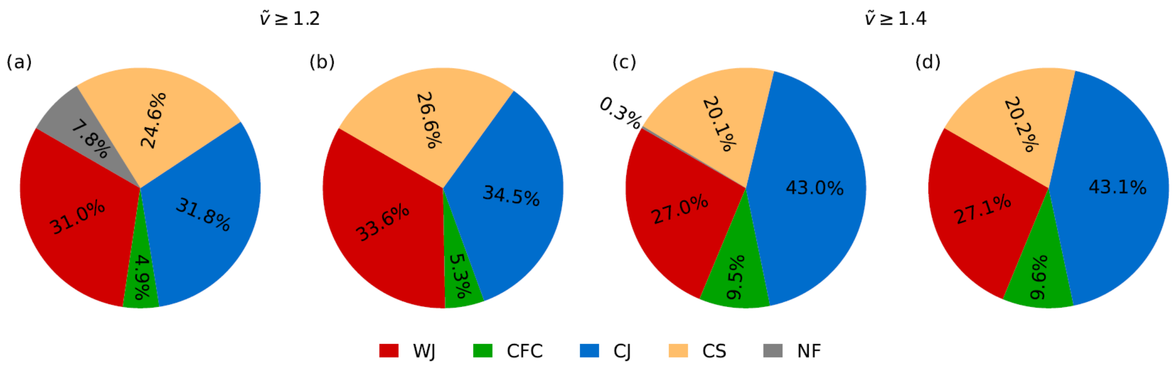

Figure FR2: As Fig. 2 m,o (COSMO-REA6 all; MAXP) but for grid points with (a),(b) $\tilde{v} \geq 1.2$ and (c),(d) $\tilde{v} \geq 1.4$.

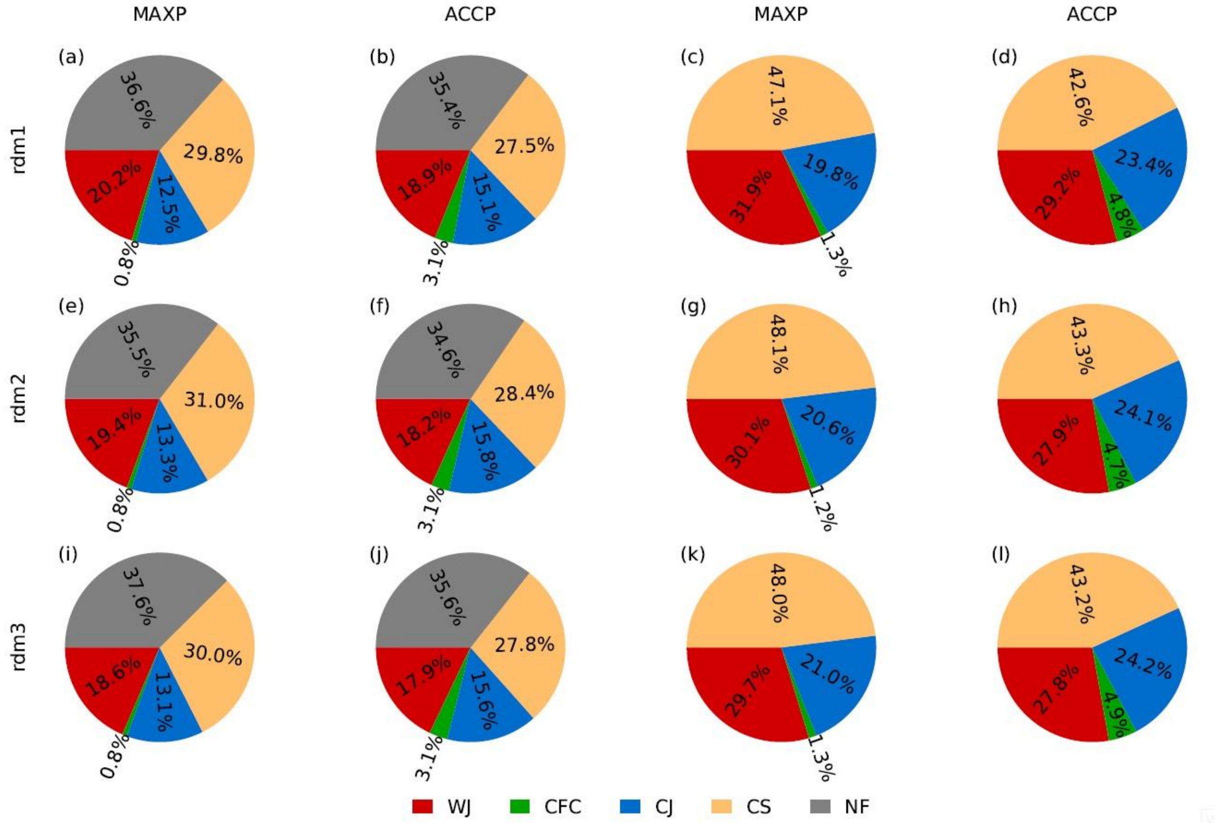

Figure FR3: As Fig. 2 m--o (COSMO-REA6 all) but for three different sets of 10 randomly chosen winter seasons.

L228 For calculation of the seasonal and inter-annual evolution of the wind features, the strong winds need to be present at at least 100 grid points over the domain. The model grid is a latitude-longitude grid with a spacing of 0.0625 degrees. Hence, the size of the grid boxes (in km2) changes with latitude. Is this change taken into account?

The choice of 100 grid points is subjective and was chosen to remove noise while keeping the rather small-scale CFC feature. This could have been based on an area threshold, too, but given that our data are on a lat-lon grid, we decided to use that. See further discussion below (comment to L271).

Existing literature While I don't know of other papers that have produced long term climatologies of mesoscale wind features, there are some other papers that the authors should consider referencing. Rivière et al. (2014, https://rmets.onlinelibrary.wiley.com/doi/10.1002/qj.2412) shows the evolution from a dominant warm to dominant cold conveyor belt jet (see fig. 10 particularly) and Dacre et al. (2012, https://journals.ametsoc.org/view/journals/bams/93/10/bams-d-11-00164.1.xml) describes the generation of a cyclone "atlas" which includes analysis of the warm and cold conveyor belts. I also suggest attempting a more quantitative comparison with the existing literature that is cited. The term "mostly consistent" is used in the conclusions – can you provide a deeper comparison? How do the cyclones over the Western and Central Europe region analysed here compare with those over the North Atlantic or UK? It would also be helpful to contrast the descriptions of the cold and warm jets here with some much older literature. For example, for me the classic paper on cyclone structure is Browning and Roberts (1994, https://rmets.onlinelibrary.wiley.com/doi/epdf/10.1002/qj.49712052006). See also Browning (2005, https://rmets.onlinelibrary.wiley.com/doi/10.1256/qj.03.201). This deeper analysis would help the reader understand what new knowledge is generated by this research. Currently this seems rather modest as it is detailed in just two sentences in the conclusions as follows: "The large number of storms investigated helped revealing the large variability in the location of CFC in a system-relative framework. Other differences to the literature include the time of occurrence of the CJ already several hours before the time of maximum depth in contrast to Hewson and Neu (2015)."

Thank you for the suggested references. A more thorough comparison with literature will be implemented in the introduction and conclusions but also throughout the paper when relevant.

Minor specific comments:

L2 I suggest replacing ”high winds” with ”strong winds” as used on the previous line as ”high” could refer to the altitude of the winds rather than their strength. Please also consider other places where ”high” or ”highest” is used.

While we understand this concern, we think the usage of “high” synonymously to “strong” is clear in the paper. We change this in some places but decided to not change this throughout the paper to be consistent to Part 1 (and the title).

L3 ”strong cold-sector winds” are not a ”dynamical feature” (unlike the other mesoscale features considered).

Thank you for drawing attention to this. We rephrase this sentence.

L35 Here it is stated that cold frontal convection is more common in cyclones following the Norwegian cyclone model. Can you point to evidence supporting this statement? Although cyclones following the Shapiro-Keyser conceptual model undergo frontal fracture, the cold front can be stronger than the warm front in idealised lifecycle simulations of these cyclones with the converse true for Norwegian-type cyclones (e.g., see Fig. 3 of Shapiro et al. 1999: https://link.springer.com/chapter/10.1007/978-1-935704-09-6 14).

Several Shapiro-Keyser cyclones in the recent years showed a less active cold front (e.g., Egon 2017, Xavier 2017, Friederike 2018). This is, however, not directly linked to a “weaker” cold front. We realise that we cannot provide statistical evidence such that we rephrase this statement to clarify this.

L55 Regarding tracking cyclones over the Mediterranean, you might be interested in this recent paper: https://wcd.copernicus.org/preprints/wcd-2022-63/

Thank you for the recommendation.

L59 Probably worth pointing out that the calculation of Laurila et al. considers the wind extremes using monthly values (i.e., the extreme wind factor is the monthly 98th percentile/monthly mean)

Thank you for pointing this out. We clarify this in the text.

L66 In this paragraph you should point out that the studies referred to consider different regions e.g., the Parton et al. and Earl et al. studies only consider the UK and so may not be representative of cyclones in general.

This is done accordingly.

L107 Here it states that ”January 2001 to December 2019” data is used whereas at the start of this section the observational dataset is described as extending to mid-2020.

Although the data set is available until mid-2020, we do not include 2020 to get the same number (19) for each month and being as close to the COSMO-REA6 data set as possible. We clarify this in the text. See also next comment.

L114 ”we concentrate here on October 2000 to March 2019, i.e., a minor shift of three months”. The model data starts 3 months earlier than the observational data but ends 9 months earlier. The reason this doesn’t make any difference is because only extended winter data is considered - this could be made clearer.

We apologise for the confusion. The COSMO-REA6 data set is actually available from 1995 to mid-2019, while the observations are available from 2001 to mid-2020. To make the data sets as comparable as possible, we include 19 extended winter seasons with the same number of months, however, a slight shift of three months. We clarify this in the text.

L114 Can you please state the number of data points used to calculate the 98th percentile at each location and date. I think it’s 210 (21 days of data for each of 10 years) but it would be good to know if I’m correct. The 98th percentile winds at any grid point will be strongly dependent on whether a localised strong wind region happens to cross that grid point on one of these 210 days. How does the value of this 98th percentile (or the lower "windy" threshold) vary spatially for a sample time snapshot? Also, is it correct that the wind speed threshold at a given time and place will be different in absolute terms for the model and observational data because the normalisation is calculated separately for the model and observational data?

210 data points is correct for COSMO-REA6. For Part 1 we tested the computation with +-30 days for an exemplary month but could not see substantial differences. To save computational costs we decided to keep +-10 days instead.

Figure FR4 shows a snapshot for COSMO-REA6 (left) and overlayed with observations (right). While only including 10 years – instead of 20 as done for the observational data set – might give exceptionally strong events more weight, the comparison of COSMO-REA6 to observations shows only slight differences, which are probably due to the different data sets instead of single events. The biggest differences occur at mountain peaks as to be expected.

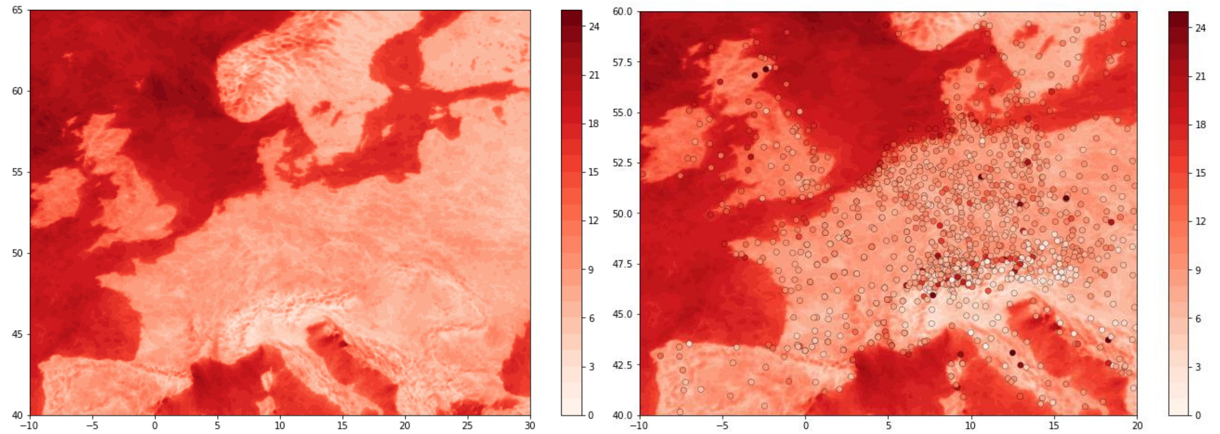

Figure FR4: 98th percentile of wind speed in m/s for 01 January, 0 UTC. COSMO-REA6 (left) and COSMO-REA6 overlayed with observations (right).

L129 Here it states "Here, we apply RAMEFI to station observations and COSMO-REA6 data under windy conditions during the extended winter months, regardless of whether a storm occurred or not. However, we later filter the output for cyclone occurrence as discussed in Sect. 2.4." Please can you clarify if any of the results shown are not filtered for cyclone occurrence? I wondered if the filtering for cyclones might only be applied in Sections 3.2 and 3.3 or even just in Section 3.3.

We apologise for the confusion. All results are filtered by cyclone and area to be consistent throughout the paper.

Fig. 2 Can you provide more details on how the track density is calculated? This field is very smooth in the plot so I suspect that there is some averaging occurring (typically track densities are plotted as the number of tracks within a certain area, such as a 5 degree spherical cap, of each grid point). I don't think it's possible that there are up to about 3 storms with their MSLP centre located at a single grid point each winter. Also, it would be more helpful to readers for the track density per year to be plotted rather than the number over the 19 years.

The track density calculation considers a radius of 750km. We calculate the sum of previously identified cyclone tracks in any grid points that are at least once within the domain. The track density is then the total accumulated number of tracks divided by the given time period to get the track density per year. This will be clarified in the text and caption accordingly.

Fig. 3 The domain used for analysis of the observational data is, I think, smaller than that used for the model data (see L105 and L118) with the latter extending further east and north. However, a further domain is specified in L169. If these domains are different then how does this difference affect the comparison between the results from the observations and model output shown in this figure? What would the figure look like if the same domains were used?

First of all, we found a mistake in the manuscript. The analysed zonal extent for COSMO-REA6 is 10°W to 25°E (not 30°E). Figure FR5 shows pie diagrams for observations (top row) and COSMO-

REA6 land grid points (lower row) for the same domain (10°W to 20°E, 45°N to 60°N) showing just slight differences to the original pie diagrams in Fig. 3.

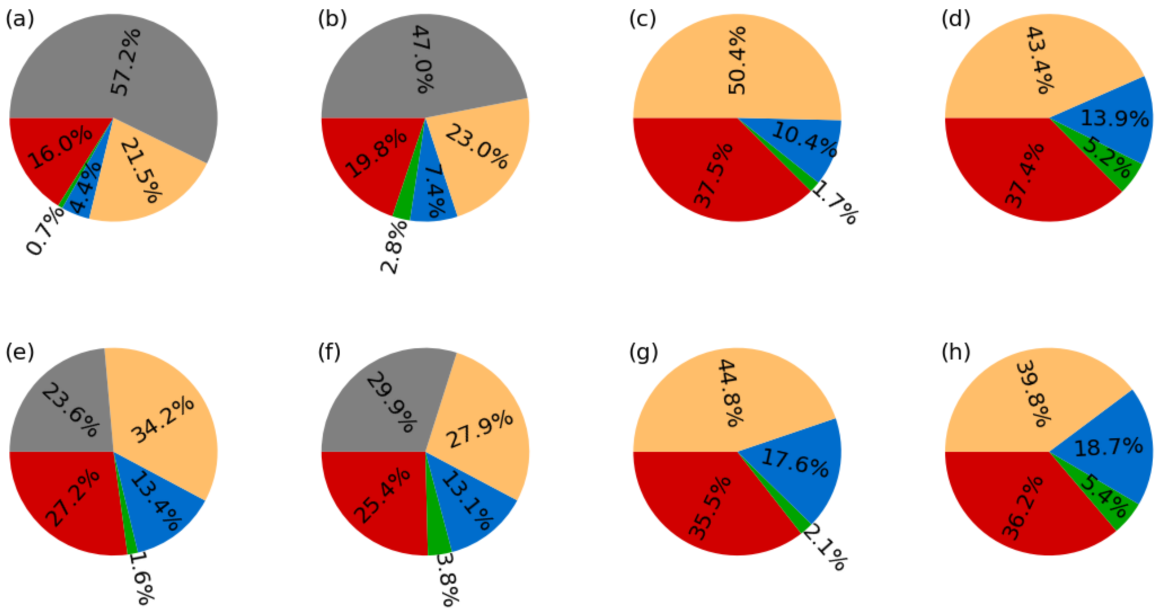

Figure FR5: As Fig 2a—h for observations (top row) and COSMO-REA6 grid points over land (bottom row) for a consistent area of 10°W to 20°E, 45°N to 60°N. The first and third column show MAXP, while the second and fourth column show ACCP.

Fig. 4 Does this plot consider a feature as occurring if it has at least 100 grid points with any percentage of likelihood at each grid point or only if it is the most probable feature at a given grid point? What is the definition of a stormy time step (is it that any feature is present at at least 100 grid points?).
It is considered when it is the most probable feature. We added this to the text.
The definition of stormy time steps is written in l. 171f but will be clarified.

L242 "However, the recent two decades w[h]ere characterised by an unusual number of storms in October," - please provide evidence for this statement.
This statement arised from a submitted paper (Mömken et al. 2023). Looking at insurance data and SSI (at least over Germany), we found almost no noteworthy winter storms in November in our time period while we found several in October (e.g., Christian/St. Jude's Day Storm 2013, Xavier 2017, Herwart 2017). We rephrased this statement clarifying this.

L243 I'm confused by the statement "given the lower threshold of $\tilde{v}$ - leads to a larger number of stormy time steps in the 19 years". $\tilde{v}$ is the wind speed normalised by it's 98th percentile value on the day of the year (and time of day and location) over the period 2005-2015 and stormy timesteps are those in which the windspeed exceeds 80% of $\tilde{v}$. Given this normalisation I'd expect the number of stormy days not to vary too much between months (as that is what the normalisation is trying to achieve). I suspect I've misunderstood something here - is it that the location specific nature of the windspeed threshold simply means that it picks out all stormy events? Please can you explain?
If the threshold was the 80th percentile of v, the number of exceedances would be the same over the ocean and over land. However, we consider 80% of the 98th percentile. As seen in Figure A1, v_{98} is about twice as high for ocean grid points and the area between 80% of v_{98} is larger for ocean grid points compared to land. As written in Appendix A, ocean grid points experience windy conditions around 45% more often than land grid points. Therefore, a cyclone might be considered a "storm" over the ocean, but not once it reaches land.
We realise that "lower" is misleading here and we will rephrase this sentence.

L247 Here the frequency of the cold jet is linked to that of cold air outbreaks. Please can you explain why these two phenomena could be expected to be linked?
The two phenomena are not directly linked and our statement is indeed misleading. We will rephrase and clarify this.

L271 The definition of "windy" is different here to that in L167. L167 defines it as strong winds "in the vicinity of 15 degrees of the cyclone centre," which implies a circular region around the cyclone whereas here a box with ±15 degrees in the zonal direction and -15 to 5 degrees in meridional direction is stated. Which is correct? Also, how is the reducing zonal extent (in km) with latitude accounted for (e.g., 30 degrees in the zonal direction varies from 2330 to 1390 km between 45 and 65 degrees North)?
We apologise for the confusion. For the whole climatology, we considered ±15° in the zonal direction and -15 to 5° in meridional direction as commented above.
As mentioned above, the changing distance in km for different latitudes was not considered, as our data is on a lat-lon grid. This is similar to other papers using cyclone-relative analysis (e.g., Sinclair et al., 2020 (https://wcd.copernicus.org/articles/1/1/2020/), Priestley et al., 2022 (https://wcd.copernicus.org/articles/3/337/2022/)). We do not anticipate that a change to area-thresholds would change the overall statements in our discussion. It would be desirable to test this effect, but the computational costs to do this for the whole climatology are quite high.

299 Here it is stated "As orography can induce convection, CFC might be detected without the occurrence of a cold front or even cyclone". I thought only windy features were identified close to a cyclone were considered though (see above comment and also the comment relating to line 129).
This is indeed misleading. What we wanted to say is that wind speeds not associated with one of the features but for example associated with a strong pressure gradient hitting orography can induce convection. For example, this commonly happens over the Norwegian mountains.

Fig. 6 I think a constant latitude has been assumed here (of 50 degrees) in the calculation of the distance circles as they appear circular. If the variation of zonal distance with latitude had been considered the distance circles would not be circular. This assumption should be stated in the caption. Please also add to the caption that the grey shading in (a) is for the NF winds.
Both will be clarified in the caption.

Fig. 6a How is the location of the NF winds close to the cyclone centre consistent with Fig. 5b which shows NF winds are typically found to the south and north of the main storm track?
The histograms in Fig. 5 are divided by the number of windy conditions. This is not done in Fig. 6a. Doing so leads to NF mostly occurring in the north-eastern quadrant as discussed above (see Fig. FR1).

L373 "RR values over 1m s−1". "RR" isn't defined in this paper but based on the earlier paper by the authors I guessed that it stood for precipitation amount. However, the units are then incorrect. Please define RR.
Precipitation amount *RR* has been introduced in Section 2.1 (l99). However, we apologise for the confusion regarding the units. These are of course mm h^{-1}.

Fig. 8 What heights are all the fields considered at?
We added the height in Section 2 (Data and methods): mean sea level pressure (*p*), 2m air temperature (*T*), wind speed at 10m (*v*), wind direction at 10m (*d*) and precipitation amount (*RR*).

L383 Here it is stated that it is not surprising that the highest gust ratio is found for the CFC features because convection is associated with high instability and turbulence. The grid spacing of the model is about 6 km. Does it use a convection scheme or is convection represented explicitly? How are the gusts diagnosed in the model? Related to this point, it would be interesting to know how the climatology of model gusts shown in Section 4 compares to a similar climatology generated from the observations given the likely parametrization dependence of the model-derived gusts. Have you considered this?
The model uses the mass flux scheme from Tiedtke (1989) for subgrid-scale convection. A note on this will be added in Section 2.2.
Unfortunately, the observational data set we use does not include wind gusts, such that a comparison is not possible.

L425 What is the evidence from your analysis that the CFC feature requires a convective trigger? The term ”trigger” is also vague: do you include both forced ascent and ascent due to the release of convective instability?
This statement indeed includes poor choices of words. It will be rephrased and clarified.

Technical errors:
Thank you for making us aware of these errors. All will be changed accordingly.

L34 ”causing” → ”it causes”.

L44 ”As CJ” → ”As the CJ”

L51 ”perception on” → ”perception of”

L74 ”convection-induces” → ”convection-induced”

L85 ”northeastern” → ”northeast”

L106, L170, L190 ”less than” → ”fewer than” (n.b. you should use ”fewer than” if you can count the objects e.g., fewer bottles of lemonade but less lemonade). Please also check elsewhere in the paper.

L110 ”Ger- man” → ”German”

L156 Remove full stop before start of 1st sentence.

L168 ”in altitudes above 800m consistent to Part 1” → ”at altitudes above 800m consistent with Part 1”

L197 ”perscpective” - spelling.

L232 ”any” → ”every”

L239 ”amount” → ”number” (because the number of stormy timesteps can be counted).

L241 ”consistent to Feser et al.” → ”consistent with Feser et al.”

L242 ”where” → ”were”

L244 ”corresponds” → ”correspond”

L301 ”of CJ” → ”of the”

L306 add ”this” before ”suggests”

L326 ”(Gentile and Gray, 2023)” → ”Gentile and Gray (2023)”.

L330 ”particular” → ”particularly”

L446 ”helped revealing” → ”helped to reveal”

L450 ”larger” → ”longer”