Authors' responses in red italics

RC2

Review on "The climatology and nature of warm-season convective cells in cold-frontal environments over Germany" by Pacey et al.

This study presents a climatology of convective cells associated with cold fronts in a front relative coordinate frame and compares these to convective cells occurring in non-cold frontal environments. Convective cells are shown to be much more frequent on cold front days than non- cold front days and the most likely location for convective cells to develop on cold front days is found to be 350 – 400 km ahead of the front. Overall, the manuscript is clear, well written and the results are supported by evidence. However, I have three major concerns regarding this manuscript.

We thank the reviewer for taking the time to review the manuscript and for their constructive feedback. The revised manuscript will include several changes based on the reviewer's comments which are outlined in more detail below under each specific comment.

The first major comment concerns how well this study fits the scope of this journal and the broader context of the results. In the manuscript the link to actual hazards is weak and little emphasis is given to this aspect. Lightning is considered but relatively briefly. It should be clearer how the results of this study inform about meteorological hazards. This study also only focuses on Germany, a choice which is motivated by the availability of radar data. While the authors do state that the study should be expanded to all of Europe, the current manuscript may be of limited interest to readers from other places than Germany. At a minimum the authors should attempt to address the question of how do these results apply to elsewhere in the world? Do they only apply over continental areas for example?

The study focuses on convective cells which are known to be linked to hazards such as lightning, wind, rain and hail. From Figure 10d we see that most cells are indeed associated with lightning; around 80% of cells in the warm-sector. Furthermore, in the final section of the paper (Section 3.4) we go into extensive detail regarding the links to hazards e.g., intensity of cells and mesocyclone frequency in the cold-frontal framework. We will however add additional discussion in the introduction about the motivation of this work being related to convective hazards. We will also emphasise in the conclusion that this work improves understanding of convective hazard climatology.

Regarding the application to other parts of Europe we found that lightning frequency has the same distribution around the 700 hPa front in a larger European domain (see Figure R1). This gives us an early indication that the results could apply on a broader European scale but since radar data (especially doppler wind velocities which are needed for mesocyclone detection) is not readily available on a European scale it is not feasible to reproduce the entire results on a European scale at this time. We will include some discussion regarding our preliminary results on lightning frequency in a larger European domain in the new manuscript.

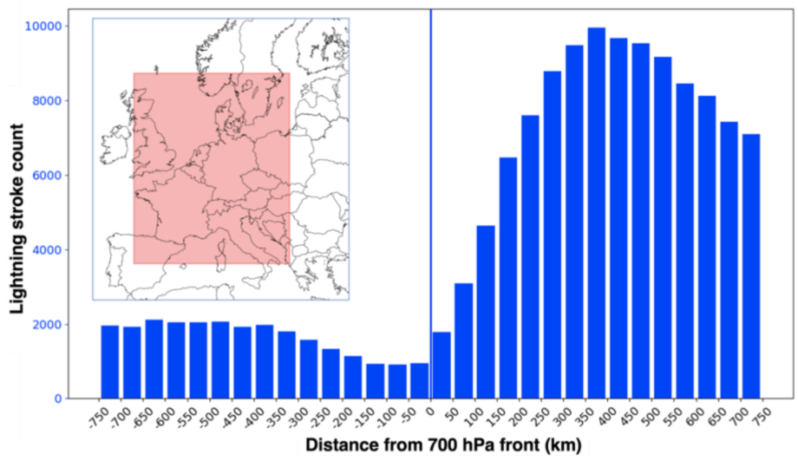

Figure R1: Lightning stroke count depending on the distance from the 700 hPa front between 2007–2016 (April–September). Lightning data were provided by the Met Office, which uses an arrival time difference network (ATDnet) to detect lightning strokes (Met Office, 2020).

The second major comment regards the assumption that the surface front is 300 km ahead of the 700-hPa front. This is an oversimplification and likely is not accurate in many cases. Specific points:

1. It is stated that this assumption is based on ERA5 data, but this is not presented – it should certainly be shown even if only as supporting material.
2. It is stated that the surface convergence zone is 300 km ahead of the 700-hPa front. In some cold frontal cases the wind shift (i.e. convergence zone) is not co-located with the temperature gradient so it may not be the thermal gradient which is 300 km ahead of the 700-hPa front. It is also inconsistent to use convergence to locate the surface front but a thermal gradient for the 700-hPa front.
3. This simple approach does not consider kata-cold fronts in which the front appears to slope forward with height as the cold air aloft has overrun the

surface front. Kata cold fronts can certainly trigger elevated convection and these fronts should be considered separately.

We thank the reviewer for their feedback regarding the surface front location relative to the 700 hPa front. The assumption is not only based on the mean surface convergence in ERA5 but also on knowledge of cold frontal slopes (1:100). We do not claim the surface front is always located exactly 300 km ahead of the 700 hPa front rather this is where the convergence is highest climatologically speaking. For this paper we are primarily interested in cold-frontal convective cell climatology. We are not focusing on specific case studies where, as the reviewer rightly mentions, the surface front location relative to the 700 hPa front may vary.

1. *We will include the climatological convergence at different height levels in the 700 hPa framework in the updated manuscript as the reviewer suggests (see Figure R2). The climatological value of CAPE, surface dewpoints and surface shortwave radiation are also overlayed.*

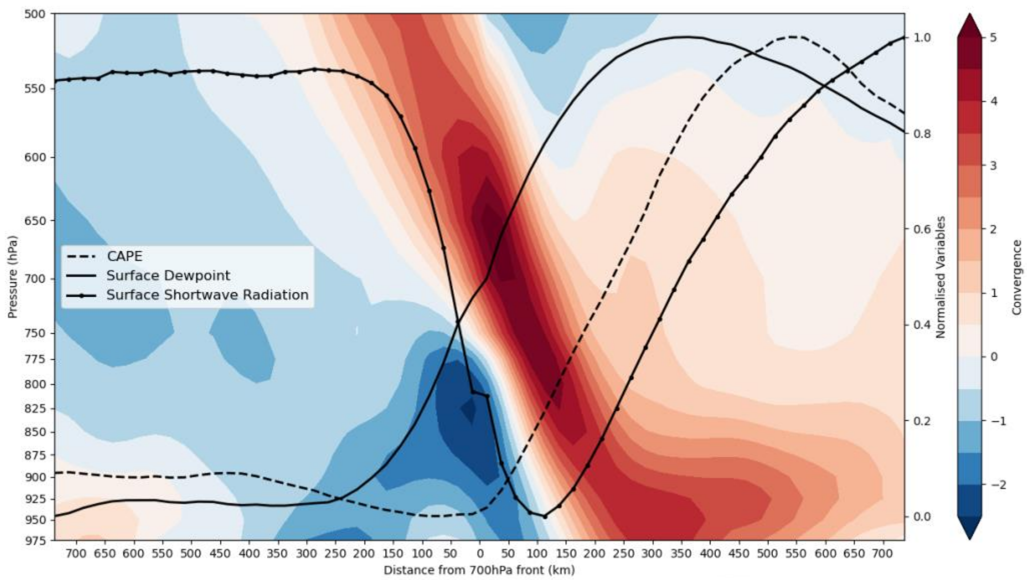

Figure R2: Climatological convergence between 975 hPa and 500 hPa (shaded), MUCAPE (dashed line), surface dewpoint (straight line), surface shortwave radiation (straight dotted line). Excluding convergence, the climatological values of the variables are normalised between 0 and 1 so the distribution around the front can be compared. Variables were derived in ERA5 data.

2. *Since the surface convergence is relevant for convective initiation, we believe this is an appropriate reference point. Furthermore, the surface temperature gradient is*

typically not as well-defined, especially in the warm-season. For this reason, automatic front detection methods are usually applied above the boundary layer.

3. *The thermodynamic gradient of kata fronts still backs with height as with an ana front (Moore and Smith, 1989, their Figure 1), thus the approach holds for both types of fronts. Furthermore, if these upper-level humidity fronts were being detected using our cold-front detection methods we would expect to find a much larger cell frequency behind the 700 hPa front. Since Kata fronts typically evolve from Ana fronts, it is not trivial to separate such fronts on a case-by-case basis. It may however be interesting for future work to investigate the effects of the overrunning upper-level dry intrusion on the resulting convection when a kata cold front is present. It is expected this increases the strength of the capping inversion thus inhibiting the premature release of convection which is relevant for severe convection. In the revised manuscript we will mention that we do not explicitly consider ana and kata cold fronts separately, but our results suggest we do not detect many of these upper-level humidity fronts in our algorithm.*

The third major comment concerns the clustering presented in section 3.3.3. Specific issues here are:

1. The manuscript lacks details on exactly how the clustering was done (e.g. was any normalisation on the input features performed?).
2. The choice of the number of clusters appears subjective whereas the silhouette score and elbow plots could be used to better justify the final number of clusters.
3. The number of clusters (30) is too large to be of practical use to e.g., forecasters
4. The justification for removing 6 of these 30 clusters is not clear and it appears that the clustering has identified 6 clusters which are not physically consistent – strongly suggesting that the clustering has not been performed in an optimal manner.
5. The only outcome of the clustering that is presented is the number of cells and the location of the front. It would be helpful for forecasters to also see additional meteorological variables associated with each of these clusters, for example, the MSLP and equivalent potential temperature.

1. *We thank the reviewer for their feedback on the clustering. We agree that further details would be useful in this section. For the reviewer's information, no normalisation was carried out on the input data as it is binary, each grid point is either 1 for a front grid point or 0 for a non-front grid point. Such information will be added to the revised manuscript for clarity.*

2. *The reviewer is right that the elbow method or silhouette score can be used to select the cluster number in a more objective manner. However, no clear elbow was identified using the elbow method and no optimal silhouette score (see Figure R3). In this case user expertise was used instead of the metrics. Further justification of the cluster number choice is discussed in the responses below. We will add a sentence explaining that the elbow method and silhouette scores were considered before selecting the cluster number and include Figure R3 below in the Appendix.*

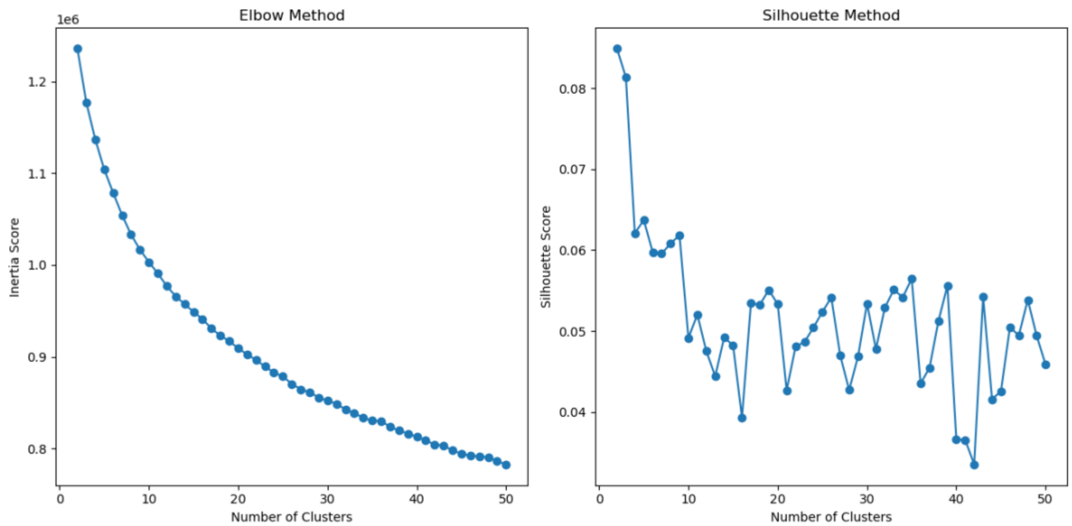

Figure R3: Elbow method (left) and silhouette score method (right) applied for cluster numbers between 2 and 50.

3. *We do not think that the cluster size is too large, forecasters are aware that cold fronts have a variety of orientations and different geographical locations. For example, forecasters at the German Weather Service (DWD) are used to identifying 29 different synoptic weather patterns (Wapler and James, 2014, their table 1). We carried out extensive testing between 10 and 50 clusters finding that lower cluster numbers resulted in fronts with different features being grouped in the same cluster. Larger cluster numbers on the other hand had less variance in each cluster but the results become less interpretable. Ultimately, a compromise somewhere must be found.*
4. *6 clusters were removed due to high variance within those clusters. This however does not indicate the clustering has not been performed in an optimal manner. When using a higher cluster number such as 50 these 6 clusters would have been split into separate clusters with less variance in the cluster, but as mentioned at 50 clusters it is hard for the reader to interpret such results.*

   5. The primary focus of this analysis is to see how the orientation/position of the front is linked to the cell detections. A composite of the equivalent potential temperature *field for each cluster will only show a boundary where the front is located, thus not adding any additional information.*

Below I also list some minor comments which would certainly improve the manuscript:

Minor comments

1. Line 41 – 42. Please expand this sentence to make it clearer. It needs to be stated that this is due to the frontal surface sloping rearwards with height. *On L39 we already mention the cold-front slope, we will amend this line to read "Due to the rearward slope of cold fronts, there is no concrete…". Thank you for the suggestion.*
2. Line 62, Question 3. This could be written in a manner so it can stand alone and does not need a reader to refer back to Q1 and Q2. This would likely make it clearer and easier to understand. *We think this would lead to a rather long research question. The reader would have just read questions 1 and 2 so it will be clear what is meant by question 3, we prefer to avoid unnecessary repetition in this case.*
3. Line 125. Figure 1. Could the domain where fronts are identified in be marked on this figure? *The domain showed in Figure 1 is the domain in which fronts were detected. We will add a sentence on this in the figure caption in the revised manuscript so it is clear. Thank you for the suggestion.*
4. Section 2.1.1. The criteria used to identify the fronts are quite large so will only identify quite strong fronts in terms of the thermal gradient. Do the results depend on these thresholds, and in particular, do the fronts still hold if weaker fronts are also considered? If not, it should be stressed more clearly that these results only apply to strong cold fronts. *We will emphasise in the method section (section 2.3) that we are focusing on cases where a single large-scale cold front is present in western/central Europe, and that smaller-scale and weaker fronts (below 6K / 100km) are not explicitly considered. Thank you for the comment.*
5. Section 2.3. How were these four examples selected and how representative of the whole data set are they? They look like quite standard fronts, so I am wondering if the method works well with more complicated or less uniform fronts. *The examples were selected to show fronts with different orientations and with cells in different locations relative to the front. The clustering analysis shows the primary frontal types contained in the dataset. The study does indeed focus on standard fronts where only one large-scale cold front is present over*

western/central Europe. We will include this point after the "timesteps containing two or more cold fronts" comment.

6. Section 2.3 / method. "Timesteps containing two or more cold front lines in the domain were omitted". Since timesteps with two fronts present are omitted, this means that this method only works for a small area and could not be expanded to e.g., European scale (even if the radar data was available). This is a notable limitation of this method which should be clearly highlighted, or the method improved to allow two or more fronts to be present at the same timestep. *In the event there are two large-scale fronts (L~1000 km) present it is not trivial to decide which front to use for the cell-distance calculations. Selecting the largest front for example may induce a bias as the smaller front may also have some influence on the resulting convection. As mentioned in the last comment we will remark in the revised manuscript that we focus on cases with a single large-scale (L~1000 km) cold front present over western/central Europe.*
7. Related to the point above, neglecting timesteps with two or more fronts present means that double fronts will be automatically ignored which may add a systematic bias to the results. *We believe the opposite is true as when two large-scale fronts are present it is not clear which front should be selected without inducing a systematic bias.*
8. Line 176. There is a typo here "in In Figure 2a…" *Thank you, this will be corrected.*
9. Section 3.1. There are a lot of numbers, percentages especially at the start of this section and it is hard to read. Many of these numbers etc, are in Table 1, but Table 1 is not referred to much here. It would help a reader to refer to Table 1 more. Furthermore, this section may be clearer if the number of cold front days was discussed first, and this information was added to table 1. *We agree that it would be useful to refer to Table 1 earlier in the section, thank you for the suggestion. The second line can be removed as the information is already contained in the table, thus making the text shorter.*

10.Line 199, should this be a comma before All rather than a period? *We feel including this in the same sentence would make it too lengthy. Instead, we will revise it to read "Such synoptic patterns are likely…"*

11.Line 210. Is the surface convergence influenced by land use, coastlines, topography etc.? *ERA5 at ~25 km resolution may struggle to resolve such features and when averaging across several thousand timesteps such effects will be filtered out leaving the surface convergence due to the front.*

12.Line 218 and elsewhere after this the phrase "pre-700-frontal environment…" is used. *We will use pre-700hPa-frontal and post-700hPa-frontal throughout in the revised manuscript for consistency, thank you for bringing this to our attention.*

Should hPa be added after 700 here? *We will use the notions of pre-700hPa-frontal and post-700hPa-frontal throughout.*

13.Line 272 – 274. How does this spatial climatology of convective cells relate to the spatial climatology of fronts as shown in Figure 1? Adding a sentence to relate these aspects would be helpful for a reader. *North-west Germany has the largest fraction of cell days on cold-frontal cell days which could indeed be linked to the high frequency of fronts seen in this region. We will add discussion regarding this in section 3.3.2 of the revised manuscript. Thank you for the comment.*

References

Moore, J. T., and K. F. Smith, 1989: Diagnosis of Anafronts and Katafronts. Wea. Forecasting, ***4****, 61–72, https://doi.org/10.1175/1520-0434(1989)004<0061:DOAAK>2.0.CO;2.*

Met Office, 2020: Lightning strike location data. Met Office, 3 pp., https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/data/adtnet_data_sheet.pdf.

Wapler, K. and James, P, 2014: Thunderstorm occurrence and characteristics in Central Europe under different synoptic conditions, Atmospheric Research, 158, https://doi.org/10.1016/j.atmosres.2014.07.011.