Answer letter

Dear editorial team, dear reviewers,

thank you for your constructive feedback. The comments are helpful to improve our manuscript.
Please find individual answers to the comments below, comments are italicised, answers in normal font. All comments that are not specifically answered here (mostly language improvements) do not require further explanation and will be incorporated in the manuscript directly. Comments about the definition of the cap and Fig. 2 are not directly addressed, as we decided to remove this analysis altogether, as the sample size of joint occurrence of the cap and strong positive temperature events is very small (see explanation below). The conclusion, that the cap cannot be the main reason for the high temperatures, still holds – because the co-occurrence is just so rare.
We will happily provide an updated version of the manuscript upon approval of this answer letter and thank the reviewers for their input.

Fiona Fix-Hewitt on behalf of the authors

Answers to Reviewer 1

"Extreme weather events" is here used for "positive temperature anomalies", "thunderstorms", and "dust". "Extreme weather events" is vague. This could be more precise. Also, the location is not specified. Therefore, the title could be shorter and more precise with a revision such as, "Occurrence of atmospheric deserts with positive temperature anomalies, thunderstorms, and dust in Europe".
We renamed the manuscript to "Co-occurrence of atmospheric deserts over Europe with strong positive temperature anomalies, thunderstorms, and dust events"

Only the last paragraph (of four) discusses the results of the study, with only one quantitative piece of information mentioned. More results should be presented in the abstract, with more quantitative information. The abstract does not say why this study is important. [...] Should the abstract be one paragraph? Most of the abstract describes the method. Only the last paragraph (five sentences) contains results, and those results are not particularly quantitative (for example, lines 15–18). Expanding the results portion of the abstract would improve balance.
In our opinion the abstract should have a sentence or a short paragraph for each of: Motivation, Objective+Scope, Methods, Results, Principal conclusions. Also see
Schultz, D. M. . (2009). Eloquent science: a practical guide to becoming a better writer, speaker, and atmospheric scientist. American Meteorological Society.
We refine the abstract to outline the importance and scope better, as well as highlighting relevant results. We do not think that stating more numbers in the abstract is helpful, as the results show spatial distributions, which are not easily summarised in one number, and the main message is the increase in the probabilities across the entire domain.

Language, punctuation and notation
We will do our best to improve language and punctuation. We will also adapt the notation of scientific variables ($T_{2\mathrm{m}}$, $AOD_{550\mathrm{nm}}$).

I have suggested breaking up section 5 into three separate sections. Some of the appendices should be integrated into the text.
We will break up section 5 and rearrange the description of the methods between the methods and the appendices to improve readability.

1. The introduction would benefit from clearer structure and organization. Consider the topics of the sentences [...].
We rewrite the introduction to take these suggestions (and those of the second reviewer) into account.

2. By the time the reader reaches the conclusion section, the main results are presented: Locations with atmospheric deserts are associated with anomalously high near-surface temperatures, a higher prevalence of lightning, and increased dust relative to times without atmospheric deserts. While these results are original, the manuscript would benefit from a clearer discussion of their implications. In particular, it would be helpful to explain how these findings advance understanding of EML-related processes and convection.
We will add more explanation, interpretation and context.

3. Lines 35–36 state the goal of the manuscript: "In this work we want to address the question whether ADs show a similar behaviour as EMLs and co-occur with positive temperature anomalies (PTAs), and whether thunderstorm formation is favoured in specific regions of the ADs." As written, this goal is not fully addressed. Although the manuscript examines the frequency of ADs in Europe, it does not assess whether this behaviour is similar to that of EMLs, as no direct comparison is performed for Europe, North America, or globally. Revising this statement so that it aligns more closely with the analyses actually performed would improve consistency.
We think you may have misunderstood this line. The behaviour of EMLs with respect to heat and thunderstorms is known (and described in this manuscript), so we want to investigate whether ADs **also**

behave like this, as they are a generalisation and can be expected to show similar behaviour. The 'different regions of the ADs' refers to regions **of** the AD, as in AD edge vs centre, not ADs in different regions. We rephrase this, to prevent similar misunderstandings: "The well-documented links between EMLs, high temperatures (Carlson and Ludlam, 1968; Cordeira et al., 2017), and convective initiation (e.g. Carlson and Ludlam, 1968; Keyser and Carlson, 1984; anicci and Warner, 1991b) motivate the investigation of whether these connections extend to ADs more generally [...]"

4. The presentation of the results is highly analytical in tone. Additional interpretation linking the statistical results to physical understanding of EMLs and convective processes would strengthen the manuscript.
We will expand the interpretations and physical context in the discussion to strengthen the physical understanding. The results should mainly report the outcome of the analysis.

5. The manuscript often assumes a level of background knowledge comparable to that of the authors. As a result, it can read as communication between specialists rather than as a presentation aimed at a broader readership. Providing more explanation of the results and their significance would help readers encountering these ideas for the first time.
We will add further context to address this comment.

8. Line 23: Is "advected" the most appropriate word here? Advection is defined as the transport of an atmospheric property by the velocity field. Referring to an atmospheric desert (an object) as advected may not be precise. Consider revising this wording consistently throughout the manuscript.
While you are correct that 'advection' in the physical sense is the transport of a property by the velocity field, we think that 'advection of an air mass' is commonly used in the field, so that the meaning is clear, especially because using different wording complicates the sentences sometimes, thereby reducing the readability rather than improving it.

12. The manuscript transitions from "heat waves" (e.g., line 29) to positive temperature anomalies (line 36). Although this may appear subtle, it is an important distinction. Heat waves typically imply longevity and extreme magnitude, whereas PTAs are instantaneous and may represent relatively small departures from the mean. It would be helpful to justify whether PTA is the most appropriate term for standardized anomalies of two and four standard deviations.
We use 'heat waves' only when referring to other literature that investigated them. For our analysis, we specifically chose 'strong positive temperature anomaly event' over 'heat event' and mentioned that 'events' in this study refer to individual hours to make the distinction between heat waves and the instantaneous hot events considered in this study clear (l. 172 in the preprint). If this was not clear enough yet, we make sure now to specifically say 'strong' every time to indicate we are only considering very unusual events and highlight the difference further.

13. Are acronyms such as AD, BL, BLH, and PTA necessary, or could some be removed to improve readability? Spelling out most, if not all, acronyms makes for less effort on the readers.
'AD' is used very often throughout the manuscript, and we want to establish the name (and acronym) for future use, which is why we think introducing this is useful. 'BL', 'CBL', and 'BLH' are widely used in the field and therefore considered to be known. 'PTA' is introduced so that it can be used in the mathematical expressions (P(PTA)) which would be illegible if the full expression was used, but in the text it is never used as acronym anyway.

18. Line 52: The first sentence is incomplete.
There was just a comma missing to make the sentence logical.

22. Taking the log of OR produces a symmetric field around zero, but loses the easy interpretability of the OR. Reconsider whether presenting the results as OR would be preferable.
You are right that the OR itself may be easier to interpret quantitatively. However, using the standardised logOR does not only make it centred at zero, but also makes the magnitude interpretable with respect to the sampling variability – which is especially important when the sample sizes are small. Thereby,

we can interpret the pattern across the domain (and across regions with different marginal probabilities) better. We consider this an important feature, so that we consciously chose this metric over OR.

24. Line 100: The basis for classifying a single flash as an "erroneous outlier" is unclear. A single flash could represent a valid event; further justification is needed.

You are correct that also single flashes can be a valid event. Blitzortung reliably captures widespread lightning activity, but isolated single-stroke detections are more prone to false positives (e.g., from noise or misclassification). Physically, genuine convective lightning activity is also expected to produce spatially clustered strokes rather than a single isolated detection, since active thunderstorms typically generate multiple flashes over neighbouring grid cells within a short time window. We therefore exclude grid cells where only a single flash was detected within a radius of one ERA5 grid cell. This filtering is conservative and consistent with our focus on widespread lightning activity associated with ADs. We will adapt the explanation in the manuscript accordingly.

43. The relatively small difference between median temperatures during AD and non-AD periods (Fig. B1) is somewhat unexpected. A similar issue arises in lines 236–238 concerning the British Isles (only 20% of AD cases occur with PTAs). Further discussion of the implications of these findings would be helpful.

The difference shown in that Figure is not of mean temperature, but of mean standardised temperature anomaly (which was stated in the text, but the figure missed the $'$). We clarify this in the text and correct the figure description from t2m to t2m$'$ or rather $T'_{2\mathrm{m}}$ now. A difference of 1 in the standardised anomaly is not small.

About the British Isles: There is no previous study about ADs and their impact, so there is no expectation of how often ADs should co-occur with high temperatures. In fact, the British Isles are among the regions where ADs co-occur most often with PTAs - but ADs are not that frequent. We do discuss, that this is interesting, as EML-related heat is discussed in several studies for the UK – although as far as we know, no statistical analyses exist. We will highlight this more in the text.

47. Line 231: The rationale for omitting these two summers should be explained.

We are **not** omitting these periods. We are just saying that they are indeed responsible for the pattern that we see (since it disappears if we **were** to omit them). We make this more clear in the text: "A striking feature in Fig. 1a is the increased probability of positive temperature anomalies in the western Mediterranean Sea and the Atlantic, especially close to the Iberian and French coasts. This pattern is likely linked to marine heat waves that occurred in the North Atlantic and Mediterranean during the study period (e.g. Juza et al., 2024; England et al., 2025). To test this, we repeated the analysis excluding the periods of the two major marine heat waves (summer 2022 and summer 2023) as a diagnostic check. This reduced the probability of strong positive anomalies around the British Isles and in the western Mediterranean, supporting the interpretation that these heat waves substantially contribute to the observed pattern. We emphasise that this exclusion was used only for diagnostic purposes and is not applied in the final analysis presented here. Removing these two summers would substantially shrink our already limited two-year dataset and would eliminate summer data entirely. Extending the time series to dilute the influence of these events is not feasible given data storage constraints."

58. Lines 321–322: "A layer is conditionally unstable when the potential temperature gradient partial$\theta/\partial z = (\theta_{BLH+2km}\theta_{BLH})/2$ km is less than 3.5 K km^{-1}." Conditional instability is when the saturated equivalent potential temperature decreases with height. Please explain the rationale for using potential temperature and for the threshold of 3.5 K km^{-1}. Provide citations to original source material.

You are right to mention this. We used θ as a proxy here, and the value of 3.5 K/km is an approximation for the mid troposphere. To be more precise, we now use the gradient of the saturated equivalent potential temperature: $\partial\theta_{Es}/\partial z < 0$. The results remain qualitatively the same.

59. Lines 328–329: "a conditionally unstable layer provided by the AD with modest CIN values underneath makes it possible for thunderstorms to occur even in the centre of an AD". This statement linking a conditionally unstable layer and modest CIN to thunderstorm occurrence appears primarily based on

parcel theory. Provide the meaning of this statement.
Thank you for raising this point. We agree that this statement is based on parcel theory, and we clarify its meaning. A conditionally unstable layer implies that, once a parcel is lifted past its level of free convection, it will continue to rise buoyantly, supporting deep convection. Modest CIN adds to this favourable setting by delaying convective initiation: it allows additional surface heating (and thus destabilisation) to occur before storms form, while still being weak enough not to suppress initiation altogether. The combination of these two ingredients – a conditionally unstable layer aloft and modest CIN – therefore creates conditions favourable for thunderstorm development, even in the centre of an AD. We will clarify this in the text.

60. Lines 331–332: "These results can be expected also approximately hold when taken over the whole day instead of daytime only." This claim is not supported by evidence presented here. Please clarify.
This is supported, as the patterns of standardised logOR are very similar for allday and only daytime conditions (Figs. 3e,g and 4a,e in the preprint) . Hence it can be expected that the allday results are dominated by daytime events and their explanation mostly holds for the allday result. We clarify this in the manuscript.

65. Line 398: AARs are not defined at the point where they are introduced.
They are defined one sentence later: "An interesting intersection with aerosol atmospheric rivers (AAR) and atmospheric rivers (ARs) emerges. ARs are narrow, elongated regions of elevated water vapour transport (Ralph et al., 2018). Chakraborty et al. (2021) apply this concept to narrow, elongated regions of aerosol transport and define AARs."

66. Line 408: Do you need both Summary and Conclusions in the title? Is one sufficient?
As both summary and conclusions are given in the section and they are not the same, we will keep both in the title.

74. In various figures, values of ln(OR) are plotted only for very large values (> 10 for Figs. 5 and C1 >5 for Figs. 1 and C2). In Figs. 2 and 3, ln(OR) is colored for all nonzero grid points. Is there a reason why ln(OR) shouldn't be plotted for all nonzero ln(OR) as in Figs. 2 and 3? Doing so would show more structure in the fields and be consistent throughout the manuscript.
We have thoroughly thought about these colourscales in these plots. As the value ranges are so different across the plots, using one colourscale on all of them makes it hard to make patterns visible on all of them. i.e. Figure 5 would be saturated everywhere if a colourscale was used that still makes patterns in the other figures visible. The visualisation of the patterns is the main purpose of these figures, rather than making individual values visible across figures. We therefore do not think the current colour scales are problematic. We will change the white in the centre of the colour scale to a light grey, to make it more visible.

Answers to Reviewer 2

Manuscript structure / clarity: the scientific question is not framed very clearly. This begins with the distinction between EMLs and ADs not being very clear, so the added benefit of analyzing ADs vs EMLs remains equally unclear. Consequently, the introduction and background come across as a rather broad introduction, that doesn't have a succinct scope. A lot of knowledge about the authors' prior work is assumed, but this should be avoided for a stand-alone manuscript. Referring to the work is of course necessary, but briefly summarizing the key points necessary to understand this work on its own is necessary.

Thank you for pointing this out. We rearrange the introduction so that the definitions and purpose become more clear. We make the references to the previous work more thorough and clear.

Methodological clarity: In some areas it is not quite clear what methodology was applied and for what reason (e.g. what kind of interpolation). This should be documented clearly throughout. The methods are also scattered through the methods section, a few additional definitions in the results (e.g. odds ratio or PTA) and additional methods in the appendix. Please consider consolidating this.

We rearrange the methods and appendices to make this more clear and understandable. The odds ratio and PTA were already explained in the methods and mentioned again in the respective results sections. We will make this more clear. To clarify: the method to calculate the standardised anomaly is defined in the methods generally, using an unspecified variable x, and this is applied to dust and temperature (which is where PTA results from). We will highlight this more.

Temporal component: While most data used appears to be hourly or 3-hourly, the time-collocation of the different phenomena is not entirely clear. Parts of the results refer to events or days, is there further aggregation performed beyond the hourly? If so, how? For the discussion of convective environments, are these at the same time or before the occurring lightning? As soon as convection initiates, the atmospheric profiles are no longer representative.

All data is hourly (the only dataset that is 3-hourly was interpolated to hourly, as mentioned). It is clearly stated that 'anomaly events' refers to individual hours (l. 172 in the preprint). We will make this even more clear. For Figure 4, a daily minimum of days with lightning meeting the condition is displayed, which is why we talk about 'those days'. We will reformulate this to make it more clear.

Sample size: Some information on the sample size of co-occurring events in the 2- year period would be helpful. While the probability of lightning, PTA or dust are shown, and the probability of AD / AD edge / AD center is shown in contours, their combination is not shown. This makes it hard to estimate by eye, how many events actually contribute to a combined occurrence and how representative such a sample may be.

Figure 3f: It looks as though single, highly organized convective events are dominating this figure, e.g. MCS in southern France. MCS are likely to be long-lived and persist in sub-ideal environments (e.g. through the night) and potentially become elevated in nature as well [...]. Highly organized systems have very high flash rates and their longevity would also cause the event to affect many hours of data. It would be helpful to have the contributing sample size somewhere (so the actual number/fraction of hours/days that have both lightning and AD presence)

Thank you for this important comment. We have carefully revisited the topic and decided to keep the figures in the manuscript as they are, because we are mainly interested in bringing out graphically the large-scale spatial patterns across the domain.

To complement the information from the figures in the main manuscript we have now added supplementary figures depicting the sample sizes (along with a short explanation in the main text). The supplementary figures are provided as Figures 1–3 in this answer letter and in the appendix of the manuscript.

For strong positive temperature anomaly and dust events, there are several hundred events co-occurring with ADs across most of continental Europe (for details see Figs. 1, 2), hence it is unlikely that the patterns of conditional probabilities and log-odds ratios are dominated by individual events. In particular the log-odds ratios are suitably standardised to reduce the influence of rare events with large variances.

However, for the much rarer events of lightning the joint occurrences together with ADs are naturally even more rare, in particular when restricting the co-occurrence to the AD centre only (see Figure 3). Hence, the manuscript now points out that the values in the individual cells of Figure 3 have to be interpreted much more carefully because they might be driven by a single long-lived or highly-organised system such as an MCS. An example is the isolated extremum over southern France in Figure 3f (of the preprint) that you pointed out.

Despite this limitation, we feel that Figure 3 of the preprint serves its main purpose, namely conveying the large-scale spatial patterns across the domain. In particular, the band of elevated log-odds ratios extending across the Iberian Peninsula and along the northern flank of the Alps (Figure 3g of the preprint, and similarly in panels c and e for continental Europe) is coherent over several hundred kilometres and hundreds of independent grid cells. Thus, this cannot be driven a single convective system.

We will revise the manuscript text to state this distinction explicitly, so that readers are guided to interpret the large-scale, spatially coherent pattern rather than individual grid cells or local extrema. Thanks again for pointing this out.

Removed analysis: We have removed the temperature-anomaly analysis under AD-induced capping inversions (Fig 2 of the preprint): the cap is itself rare, making joint occurrences too sparse for robust interpretation. This does not affect our conclusion – a capping inversion cannot be the primary driver of the widespread positive temperature anomalies in Fig. 1 of the preprint, precisely because it occurs too infrequently to explain a signal of that spatial extent and magnitude.

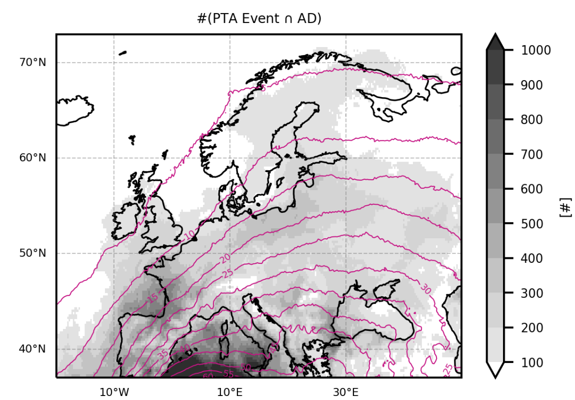

Figure 1: Joint occurrences of PTA event and AD. Joint occurrence numbers in shading. The probability of AD is denoted in contours.

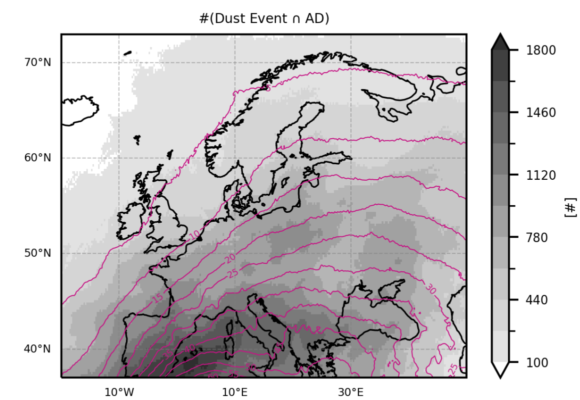

Figure 2: Joint occurrences of dust event and ADs. Joint occurrence numbers in shading. The probability of AD is denoted in contours.

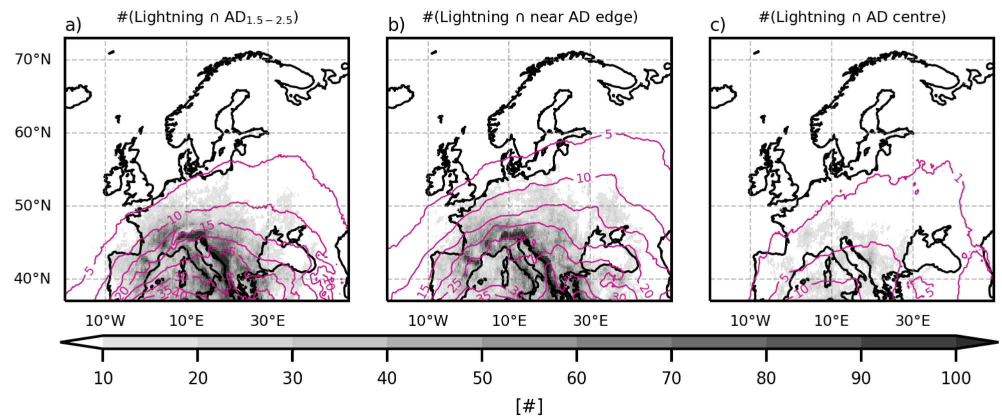

Figure 3: Joint occurrences of lightning **(a)** with AD between 1.5 and 2.5 km, **(b)** near AD edge , **(c)** in AD centre. Contours mark the probabilities of $AD_{1.5-2.5\mathrm{km}}$, AD edge, and AD centre, respectively, as in Fig 3 in the preprint.

Line 3: rather than introducing atmospheric deserts as generalized EMLs, I would prefer a succinct definition
We will incoporate this.

Line 32: I suspect instead of Feldmann et al., 2021; Feldmann et al., 2025 might have been meant here (as referred to later in the discussion of results); overall this list of references could be discussed in a more differentiated manner, rather than just listing all of them
Both Feldmann studies are relevant to this manuscript and this was on purpose. We revise the literature discussion to make it more differentiated and motivate the individual references.

Line 38: consider including Brennan and Wilhelm 2025 (this also should be referenced in the discussion of combined dust-thunderstorm events)
We include the reference in the discussion, to put our results and motivation in context and provide an outlook and some insight into joint events (i.e. dust and thunderstorms) which are not discussed in the manuscript. We will also extend the dust section in the introduction.

Line 732: Consider expanding this part explicitly with the expected effects on incoming solar radiation and changes to precipitation
We will expand the part about dust and its effects to give a more complete picture of the state of research.

Line 80: Why was this data period of 2 years chosen?
Due to the high computational demand of the method, we chose a 2-year period for a first assessment of ADs. At the time of the project start, this was the most recent 2-year period. This study uses the same period (rather than an updated one) for consistency with the previous work (Fix-Hewitt et al 2026).

Line 89: The domain listed here is different from the domain shown in the results figures
The domain listed here is the full domain used in the trajectory calculation. The interpretation of the results is confined to the region north of 37N, as Europe is the main focus. We will clarify this.

Line 95: Why Blitzortung in particular over more Europe-oriented networks? This study doesn't require a global network.
Blitzortung data is available to us as members of the network and will make a future expansion of the study area possible. Other European datasets do not cover the entire study area for this study either, i.e. see Fig 5 in:

Schulz, W., Diendorfer, G., Pedeboy, S., and Poelman, D. R.: The European lightning location system EUCLID – Part 1: Performance analysis and validation, Nat. Hazards Earth Syst. Sci., 16, 595–605, https://doi.org/10.5194/nhess-16-595-2016, 2016.

Our lightning analysis is not aimed at absolute flash-rate climatology or precise stroke-level accuracy, but at qualitative co-occurrence patterns with ADs over a large, heterogeneous domain. We aggregate lightning data to hourly resolution and a grid with 0.25° grid-spacing to match ERA5. For this purpose, the somewhat lower detection efficiency and location accuracy of Blitzortung relative to specialized regional networks is an acceptable trade-off for the broader, more homogeneous spatial coverage it provides.

Data section in general: consider adding a variable table with all abbreviations and units

There are only 3 variables used in this study: standardised anomalies of $AOD_{550\mathrm{nm}}$ and $T_{2\mathrm{m}}$, as well as lightning flashes. We do not consider a table necessary here.

Line 106: How is the data remapped and interpolated? With what method?

The data is interpolated bilinearly in space and linearly in time, using cdo's built-in functions. We add this information to the manuscript.

Line 115: a brief definition of the AD airmass is needed somewhere here. For this to be an independent paper, this cannot just be referred to in FH26

The definition of ADs is given earlier in the manuscript and their detection is described in lines 115-125. The entire paragraph is a summary of the method from FFH26, not only a reference.

We clarify as: "The calculation of the trajectories and detection of the AD – the air mass originating in the BL of desert, (semi-) arid and/or elevated source regions – follows Fix-Hewitt et al. (2026): Trajectories are calculated forward in time with the Lagrangian Trajectory Analysis Tool (LAGRANTO; Sprenger and Wernli, 2015). They are initiated hourly along a "curtain" marking the northern boundary of the source region in North Africa (see Fig. 1). The grid spacing of initialisation locations is 5 km in the horizontal and 10 hPa in the vertical between 1100 and 400 hPa. Only those trajectories are calculated forward that have a northward (Europe-ward) wind component and start below the interpolated boundary layer height BLH_i (linearly interpolated between daily maxima of the BLH, as a way to take into account former BLH air that is above the night-time BLH). Trajectories are integrated for 120 h. This results in approximately 108 million trajectories, of which 61 million pass 37° N and are therefore taken into account in this analysis. Trajectories are aggregated to grid boxes of 0.25°×0.25°×500 m. A grid box containing at least one trajectory is defined as an AD-cell, and the AD air mass consists of all AD-cells combined. If there is any AD-cell in a column, this is defined as an AD-column. For the statistical analysis in this study, "AD presence" refers to the presence of an AD-column, unless noted otherwise. For more details on the detection method refer to Fix-Hewitt et al. (2026)."

Line 166f: Where is the northerly curtain of initiation locations? This is rather ambiguous

We add a figure marking the curtain of initiation locations.

Line 141: How much data is discarded in the "in-between" area?

With the thresholds used in the manuscript, 64.4 million grid cells were classified as near-edge and 9.2 million as AD-centre, while 12.2 million cells were excluded from either category to maintain a clear distinction between the two regions. We add this to the manuscript.

Line 161: Please consider including a justification for the different periods – is it that beneficial to have a different period for t2m than dust?

Line 163: How is the detrending done? Moving window, linear trend, ...

We add: "For $T_{2\mathrm{m}}$ we use 1992–2021 as the reference period, to follow the standard 30-year climatological window, which yields more robust estimates than a shorter period (especially in the presence of a climate trend). Because a climate-change signal can be expected in $T_{2\mathrm{m}}$, the full $T_{2\mathrm{m}}$ time-series (01/1992–04/2024) is linearly detrended prior to anomaly calculation (using cdo, Schulzweida 2023). The dust reanalysis data is only available from 2003 onward, so we use 2003–2021. As no comparable long-term trend is expected in the dust time series, we do not detrend it and consider the shorter 2003–2021 baseline period sufficient."

Line 200: Using ' for an anomaly is rather unusual, often a delta is used.
We intentionally keep $'$, which is also often used, i.e. in the Reynolds decomposition $x' = x - \overline{x}$, which is similar to the nominator in Eq 1 in the manuscript. Δ can also be misinterpreted as a difference between one time step and another. The manuscript clearly explains the definition in Eq. 1, so we keep the notation.

Line 200: exceeds 2 standard deviations? 2 sigma?
As the variable is a standardised anomaly, it is correct to say that it exceeds 2. This is equivalent to saying the (non-standardised) anomaly exceeds 2 standard deviations.

Line 327ff: "on such days..." To me, throughout the manuscript, the temporal references become unclear. ERA5 and lightning data are used hourly, whereas dust is 3-hourly. When looking at convective environments, is this done the hour before the lightning? Or the maximum convective potential of the day? This sentence refers to full days, but I do not find a description of daily aggregation performed on the data.
All data is hourly (dust was interpolated to hourly, too). Lightning is attributed to the full hour prior to the recorded time stamp (i.e. timestamps are floored).
For Figure 4b: θ is used at the times of lightning (or rather up to an hour before, since lighting time stamps are floored).
For Figure 4c: this shows the daily minimum (given a flash meets the condition any time that day), exactly because once convection was initiated, the profile has changed.
Figure 4d shows CIN at the time of daily minimum $\partial\theta/\partial z$.
For the revised version of the manuscript we use a θ_{Es} criterion as mentioned above, and clarify the explanation.

Section 5.3.2: This section should include a discussion of Brennan and Wilhelm, 2025, as this explicitly investigates the relevance of Sahara dust outbreaks for hail occurrence in Europe.
We do not mention this reference in the dust section itself, where no reference is made to events that are dusty and stormy, but we will discuss joint occurrences in the final discussion then and mention the Brennan and Wilhelm reference there.

Appendix A: Consider rephrasing "some methods". Moreover, to me it is not logical which methods are back here vs the main body. Especially the odds ratio should at least be succinctly explained in the main body, with more details provided here.
We rearrange the methods and appendices to make it more understandable.

Appendix B: The supportive figures are not referenced in the manuscript. If they bring additional value, they should be mentioned. Otherwise, consider putting them in a supplement rather than the appendix.
The figures are referenced (although sometimes only through mentioning the respective appendix section):
A1 (l 131), A2-3 (l 134)
B1 (l 213), B2 (l 219 ff)
C1-2 (l 172)
However, we will remove Appendix C, as it does not yield much more information, and the use of extreme values reduces the sample size significantly. Figure B2 will be removed, as the discussion of the cap's influence will be changed in a future version of the manuscript as mentioned in the earlier answer about sample sizes.