Response to referee comments

Manuscript: Lightning2EarthCARE: Collocated geostationary lightning and EarthCARE observations of electrically active storms

We would like to thank both referees for their time and effort in reviewing our manuscript and for their valuable comments and suggestions, which helped us clarify several parts of the manuscript. We have considered all the comments and revised the manuscript accordingly. Below, we provide a point-by-point response to each comment and describe the changes made to the manuscript.

Referee 1:

Comment: Line 85: "In this study, we use lightning groups to retain fine spatial structure while ensuring a consistent level of detail across"

I suggest rewrite this sentence. This part "to retain fine spatial structure while ensuring a consistent level ..." is very non scientific. If you are interesting in a better spatial location why not use the events. The events are more associate ... maybe "more fine spatial structure" ... of the flash area. Despite de GLM and LI having different resolution, considering the algorithm of flash identification, we should expect less correspondence between the number of events and groups, however, the number of flashes should be equivalent. Also, commonly the thunderstorms severity is associate with measures of flash rate, note group rate. Why not include the flash information in this subset?

Response: *The choice of lightning groups is mainly motivated by the need for more accurate spatial and temporal collocation with EarthCARE. Groups are also more directly related to observed optical emissions, whereas identification of flashes depends more strongly on the clustering thresholds. We also agree that flash information is valuable, and therefore we preserve the flash association of each group in our dataset, although it is not directly used in the present analysis. We changed the text to:*

"In this study, we use lightning groups for collocation with EarthCARE observations. Compared to individual events, groups provide a more consistent representation of lightning activity across both instruments, as differences in pixel size make event-level observations less directly comparable. Groups also better preserve the spatial and temporal structure of lightning activity compared to flashes, allowing for more precise matching with EarthCARE. Using flash centroid locations alone would reduce the spatial information available for collocation, particularly for flashes extending over large areas. Nevertheless, we retain information about the parent flash associated with each group for possible future use (Sect. 3.1)."

Comment: Line 155: “to aggregate lightning groups into physically meaningful entities” Try change the “ physically meaningful entities” to a text more scientific. Ex. to aggregate lightning groups in cluster …

Response: *We have revised the sentence and changed it to:*

“to aggregate lightning groups into clusters representing individual storm systems”

Comment: Line 170: Fig 2 description is missing. In the figure I think that is missing some legend. There is just legend for red cluster. Another colors is not lighting clusters?

Response: *We agree that the description was not sufficient, so we changed the legend of Fig. 2 as well as the captions:*

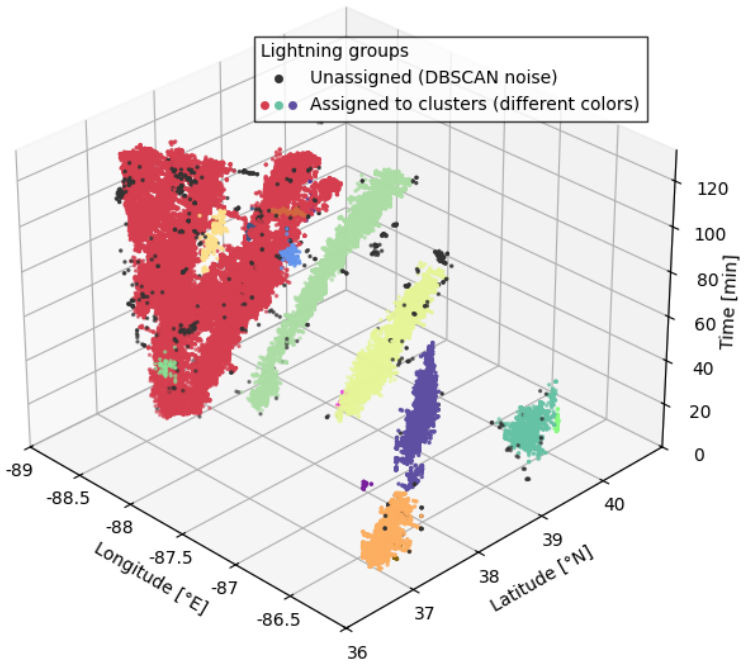


Figure 2. Spatiotemporal distribution of lightning groups observed on 2 August 2024 along EarthCARE orbit number-frame 01017B. Colors indicate distinct clusters identified using a DBSCAN algorithm in latitude–longitude–time space in the first clustering stage. Black points denote unassigned groups (DBSCAN noise). The largest cluster (red) forms a single DBSCAN cluster that splits into two branches with time, potentially reflecting the splitting of a thunderstorm.

Comment: Line 185: “Figure 2 illustrates this issue, showing a case in which one of the identified clusters corresponds to a splitting supercell” I suggest explain better this sentence, What cluster is splitting? I cannot see this supercell in figure 2 directly.

Response: *We have revised the text to describe the observed branching of the largest lightning cluster more clearly. Specifically, we clarify that the cluster remains a single DBSCAN cluster while separating into two branches over time, with one branch propagating in a different direction. Although this behavior may indicate supercell splitting, it was not independently confirmed using ground-based radar observations. We have therefore removed the identification of the storm as a supercell and instead refer to possible storm splitting. The text has been revised as follows:*

"Figure 2 illustrates this issue, showing a case in which the largest lightning cluster (red) remains a single DBSCAN cluster but separates into two branches as time progresses. Both branches initially originate from the same region before diverging in different directions, possibly reflecting the splitting of a thunderstorm. At the time of the EarthCARE overpass (approximately 60 min on the time axis), the two branches could potentially be identified as separate storm clusters. However, because the initial DBSCAN clustering is performed over the entire time window, they remain connected and are identified as a single cluster."

Comment: Lines 255–270: What is "storm-level" I could not understand this therm, this idea. I this that is necessary a better explanation about what is "storm-level".

"It provides storm-level records derived from along-track CPR–lightning associations.";
"....indicating that the EarthCARE CPR directly overpassed the storm."

"The CPR sample with the maximum lightning activity defines the lightning storm peak location and time. A land–coast–water surface classification is assigned from the fraction of CPR samples intersecting land within the storm footprint. Lightning statistics for the overpassed storm are computed relative to the peak time, including total lightning counts within ±2.5 min and the corresponding mean lightning location, as well as counts within both ±2.5 min and ±2.5 km of the CPR track. Temporal storm evolution is characterized at the parent-cluster level using per-minute lightning group counts and mean lightning positions within a ±1 hour window around the CPR overpass, allowing estimates of storm duration and travel distance." Is not clear to me the definition of "storm-level". Ex. "evolution is characterized" ... what this represent, evolution is characterized ... how???

Response: *Thank you for pointing this out. We have revised Sections 2.4 and 3.3 to clarify the meaning of storm-level records and the characterization of lightning storm evolution. In particular, we now explain that storm-level refers to the aggregation of lightning information for individual lightning clusters associated with the EarthCARE CPR track, in contrast to the along-track dataset, which provides lightning counts for individual CPR samples. We have also clarified that temporal evolution refers to changes in lightning activity and its spatial distribution, described using per-minute lightning group counts and mean lightning positions within a ±1 hour window around the CPR overpass. The changes to the text are as follows:*

Section 2.4

"The third dataset, lightning storm catalogue (Dataset C in Fig. 3), is generated from the strict along-track lightning counts. Unlike Dataset B, which provides information for individual CPR samples, lightning storm catalogue summarizes lightning activity for each lightning cluster intersected by the EarthCARE nadir track. The CPR sample with the maximum lightning activity is used to define the storm peak time and location. Summary lightning statistics and temporal evolution information are then computed relative to this overpass reference.

For simplicity, the three datasets are also referred to as frame-, track-, and storm-level datasets, representing lightning activity in the broader region surrounding the EarthCARE overpass, lightning counts associated with individual CPR samples, and summary information for individual lightning clusters associated with the CPR track, respectively. A detailed description of the resulting data records, including their file organization and key variables, is provided in Sect. 3."

Section 3.3

"The lightning storm catalogue is intended to support storm-scale analyses of lightning activity in relation to EarthCARE observations and to facilitate comparisons across storms, surface types, and observing conditions. In contrast to the track-level dataset, which provides lightning counts for individual CPR samples, the storm catalogue aggregates this information for individual lightning clusters intersected by the EarthCARE nadir track. Each storm-level record therefore represents a lightning cluster associated with the CPR overpass. A cluster is included in the catalogue when at least one CPR sample has a non-zero lightning_count_2p5 associated with that cluster."

"The temporal evolution of the lightning activity is described at the parent-cluster level using per-minute lightning group counts and mean lightning positions within a ±1 hour window around the CPR overpass. These quantities provide information on changes in lightning activity and its spatial distribution over time, allowing estimates of storm duration and travel distance."

Comment: Lines 465: Conclusions. I suggest include in conclusion some new perspective that Lightning2EarthCARE is difference in relation to another satellites subsets, for example, in relation to Precipitation Features dataset. Highlight the characteristic of lighting activity life cycle state evolution (1h after and before), giving some aspect of the life cycle of convection. Maybe some elements more general in chapter can bee move to conclusion.

I suggest close your conclusion with some connection with the literature and not just with the repository.

Response: *Thank you for this suggestion. We have changed the final paragraph of the conclusions to better highlight the contribution of Lightning2EarthCARE, mainly its ability to place the observed vertical cloud information in the context of the storm's electrical development. The final paragraph is as follows:*

"Building on previous studies linking lightning activity to strong convective updrafts and ice mass (Deierling and Petersen, 2008), as well as recent investigations of convective dynamics using

EarthCARE's first spaceborne Doppler measurements (Galfione et al., 2025), Lightning2EarthCARE facilitates further research on the relationship between lightning activity and the dynamics and microphysics of convective storms. The dataset brings together lightning-producing storms identified along the EarthCARE track and establishes links between their vertical cloud observations and the temporal evolution of lightning activity monitored by geostationary satellites. This makes it possible to provide information not only on lightning counts at the time of the overpass but also on whether lightning activity is increasing, decreasing, or remaining relatively stable, placing the observed vertical cloud structure in the context of the storm's electrical development. As the EarthCARE record grows, extending the catalogue will increase the number of sampled storms and support more robust analyses across different cluster types and lightning development stages, including studies of convection-lightning coupling, model evaluation, and integration with other space- or ground-based observations. In this way, Lightning2EarthCARE provides a basis for improving our understanding of how lightning activity relates to the internal structure and evolution of convective storms."

Referee 2:

Comment: Figure 1 caption: I think this is the first time MSI is presented in the paper. Please write the full text for MSI the first time it appears.

Response: *Part of Figure 1 caption was changed to:*
"(b) Zoomed-in section of an EarthCARE multi-spectral imager (MSI) frame D at mid-latitudes."

Comment: Line 109: It would be helpful to see an example of the aerosol product available from EarthCARE. Maybe another panel could be supplied in Figure 9 and 10

Response: *We added panels to Figures 9 and 10 showing an example of ATLID Level 2 product, including the particle backscatter coefficient, with clouds and aerosols distinguished using lidar-based simple classification. We also added a paragraph describing these observations.*

The changes to the text, figures, and captions are as follows:
"The ATLID observations provide complementary information on the uppermost cloud layers, as the lidar signal is rapidly attenuated within optically thick convective clouds. However, compared with CPR, ATLID reveals finer structures near cloud top, including thin ice clouds. In cloud-free regions, ATLID also detects aerosol layers, as evident around the smaller convective cell in Fig. 9b, where aerosols are distinguished from clouds using lidar-based simple classification."

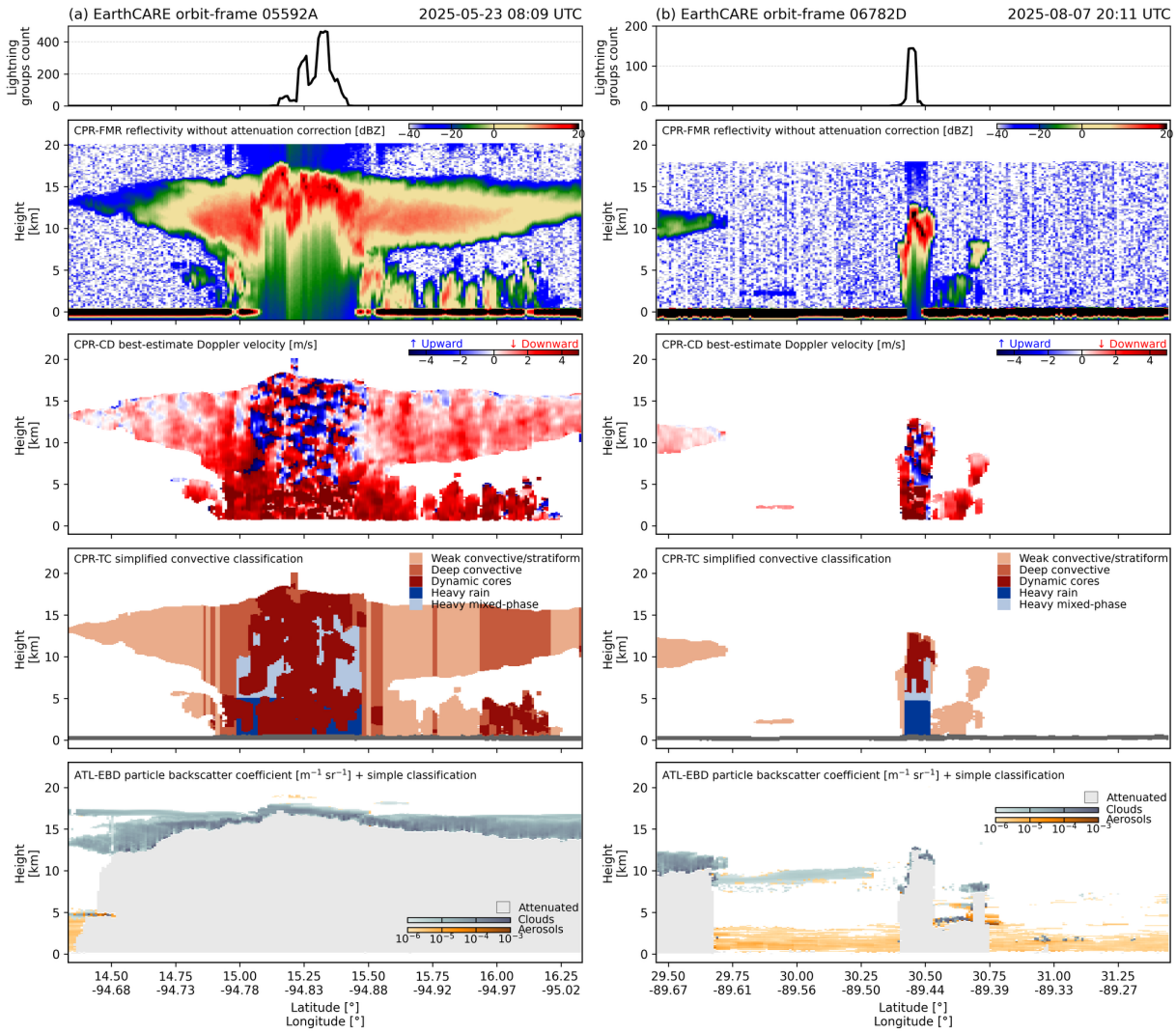


Figure 9. Comparison of GOES-GLM lightning activity and EarthCARE CPR and ATLID observations for two selected cases, shown in columns (a) and (b). The upper panel in each example shows the lightning group count along the EarthCARE track, based on lightning detections within ±2.5 min and ±2.5 km of each vertical sample. The following panels show CPR reflectivity, Doppler velocity, simplified convective classification, and ATLID particle backscatter coefficient, with clouds and aerosols distinguished using lidar-based simple classification.

The changes are identical for Figure 10 as well.

Comment: Figure 6: What are the green and purple shaded areas (B,D and F,H)? It is not explained in the text or the caption.

Response: *The shaded areas represent the number of storms sampled in the corresponding EarthCARE frames, which are grouped as A+E, B+D, and F+H because they cover the same*

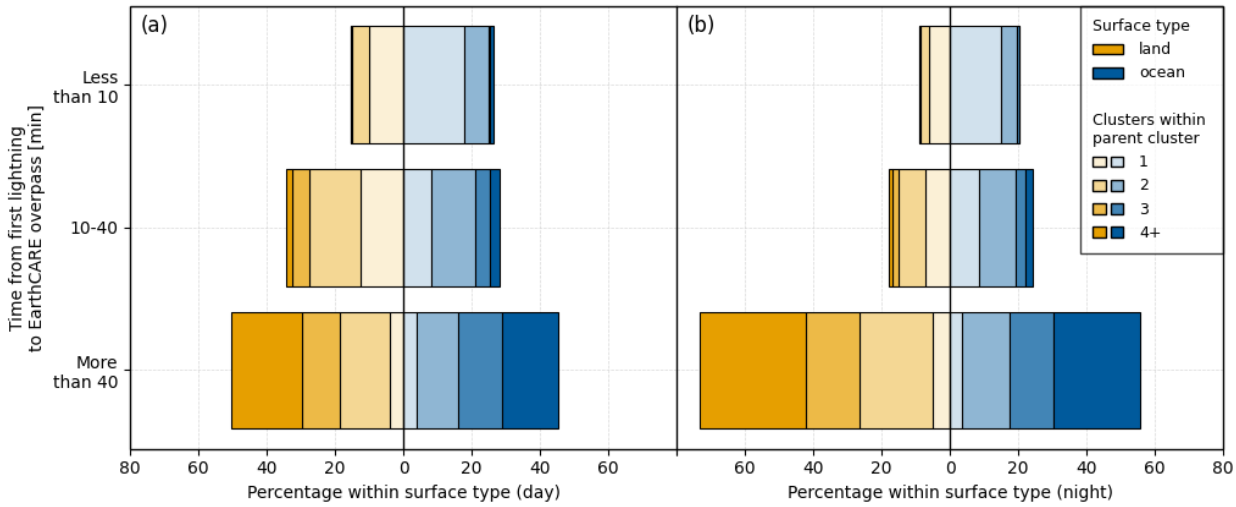
latitudinal ranges. This grouping is already indicated in the legend of Fig. 6. However, we have added further explanation in the main text and included a reference to Fig. 1, which shows the latitudinal ranges of the individual frames. The text has been changed as follows:

"The spatial distribution shows a seasonal meridional shift, with enhanced sampling moving northward and southward in phase with the seasonal displacement of warmer conditions and convective activity. This is illustrated, for example, by the higher number of storms sampled in EarthCARE frames B and D during the northern hemisphere summer months (Fig. 6)."

"Figure 6. Daily number of lightning storms in the catalogue, shown as a 10-day rolling mean and grouped by EarthCARE frames (see Fig. 1 for the latitudinal ranges of the frames)."

Comment: Figure 7: The signs on the y-axis < and > are confusing. I would rather write "less than -10" and "greater than -40" if that is the meaning.

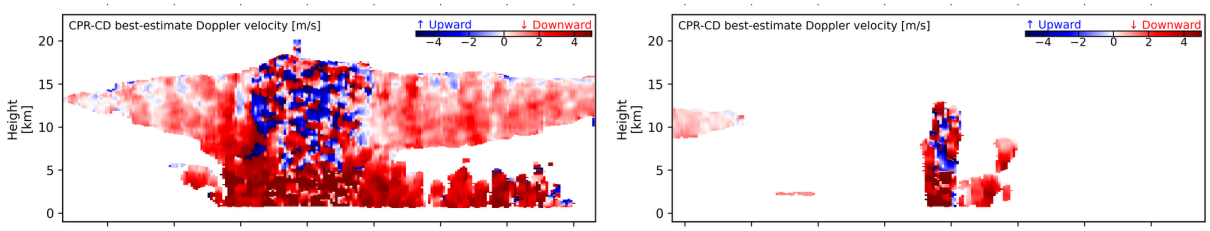
Response: *We have changed the y-axis title, tick labels and caption of Fig. 7 for better clarity:*



"Figure 7: Percentage distribution of lightning clusters across early, intermediate, and mature lightning stages for land and ocean cases during (a) daytime and (b) nighttime. The stages are defined by the time elapsed between the first lightning occurrence in the parent cluster and the EarthCARE overpass: less than 10 min (early), 10–40 min (intermediate), and more than 40 min (mature). Shading indicates the number of child clusters associated with each parent cluster at the time of the overpass."

Comment: Figure 9: In the velocity plots, are negative values updrafts or downdrafts? Please clarify in the text or the caption.

Response: *Negative values represent updrafts. We have added an explanatory label to the legend in Doppler velocity panels in Figure 9 and 10 and also clarified it in the text:*



"In the examples shown in Figs. 9 and 10, negative Doppler velocities (shown in blue) correspond to apparent upward motion but cannot necessarily be interpreted as updrafts due to possible velocity aliasing. Therefore, the patchy Doppler velocity structure is more indicative of the location of convective cores than of individual updrafts and downdrafts."