We thank the referees for their thoughtful and insightful comments which have helped us improve this manuscript.

In light of several converging feedbacks, we propose changing the title as follows: “A method for characterizing the spatial organization of deep convective cores in deep convective systems’ cloud shield using idealized elementary convective structure decomposition”, putting emphasis on the assumption of elementary convective segmentation with circular structures of varying radii.

Another major change is that the convective area S has been replaced by the letter A to align more closely with the term 'area' rather than 'surface'. In our response, we have kept S consistent with the comments made, but all occurrences have been modified within the manuscript.

We have also updated some of the figures as requested, along with adjustments in the text. Two additional figures have been added, the first one in the “Results” section (new Figure 15), and the last one in the Annex B section (Figure B3). In addition, we have reinforced the sensitivity study in Annex B by conducting a Monte Carlo-style experiment where we propagate error within the clustering process, in order to evaluate the stability of the method with respect to each of the four key parameters. Text has been amended in tracked changes.

Below we provide point-by-point responses to each comment.

<u>RC1 :</u>

This paper introduces a novel method to classify and study the spatial organization and geometry of convective systems as well as their distribution of convective cores. Four key metrics are introduced based on infrared brightness temperatures in geostationary satellite observations and outgoing longwave radiation in km-scale numerical simulations (with a delta x of 3km) of idealized cloud scenes. The metrics give a more holistic description of convective organization, including the size and relative fraction of convective cores as well as their spatial arrangement and randomness. The metrics are discussed well in the context of already existing metrics for convective organization and the authors make an argument that convective organization cannot be well described when relying on one single parameter instead of combining different aspects of the spatial structure. Overall, the paper is well-structured, well-written and is a valuable contribution to allow for a more sophisticated analysis of convective processes. While I do not think that any additional analyses are needed, I think there are still a few parts in the paper that need clarifications, in particular because the main point of the paper is to introduce a new technique that should be straightforward to reproduce. I recommend the paper for minor revisions.

We would like to thank the referee for the positive and constructive feedback. We will address all the points in the following. Please see below our responses to the individual comments, with the original comments included in black and answers in blue.

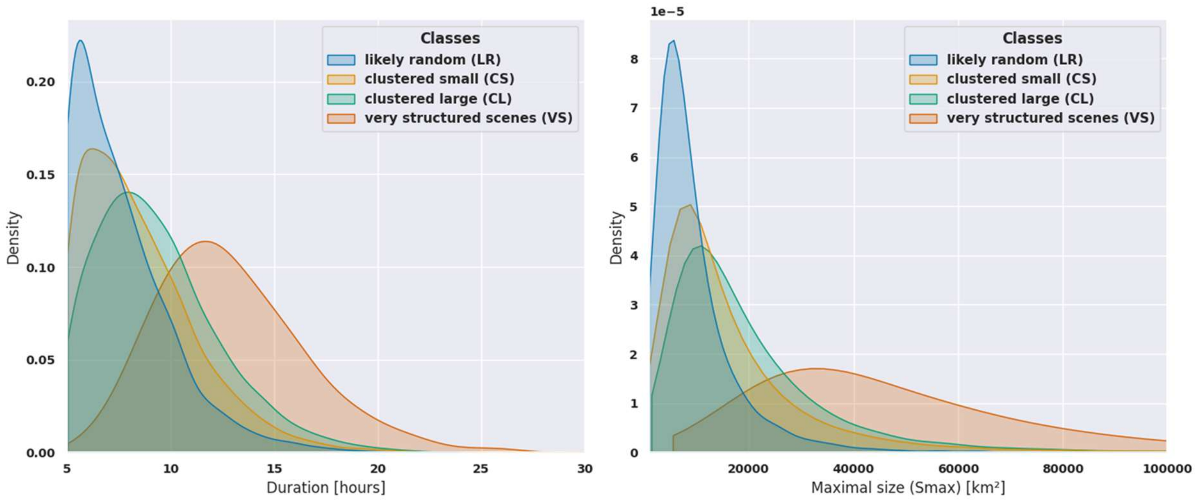
General comments:

Physical processes

Acknowledging that this paper focuses on the introduction and development of a new method, it is not expected that the conclusions center around new findings about convective organization. However, I think the paper would benefit from explaining how the introduced metrics and the classes that are based upon these four metrics, could be used to identify certain processes. Right now, the differences in the distributions (e.g. in Fig. 13-14) are not really discussed in the context of which underlying weather systems and processes would produce the different signatures in convective organization.

We would like to point out that the main purpose of this paper is not to explore internal processes, but rather to constrain the impacts of the organization of convection in relation to the morphological properties of thunderstorms.

We have added a figure in the results section showing for each of the classes the distributions of two key morphological parameters of the DCS life cycle: Duration (lifetime) and maximum extension of the associated anvil cloud (Smax). The results are discussed in a paragraph added at the end of the section.



<u>Figure 15:</u> Empirical probability density function (estimated via KDE) for each class from the TOOCAN-radar dataset. Left: Duration (lifetime in hours), Right: Maximum extension of the associated anvil cloud along the life cycle of the DCS.

This said a direct way to explore the different processes that would produce different organization of convection would be to extend a recent study we published (Roca et al., 2025) where the dynamical and thermodynamical environment in which the system morphology evolves has been analyzed. For instance, it would be interesting to link the F and C parameters to CAPE and precipitable water while the L and P might be more related to the vertical shear of the horizontal wind. Such efforts are underway and will be part of another study.

In addition, it would be useful to name a few examples of how this method can be applied in weather and climate research.

We have added the following paragraph to the conclusion section:

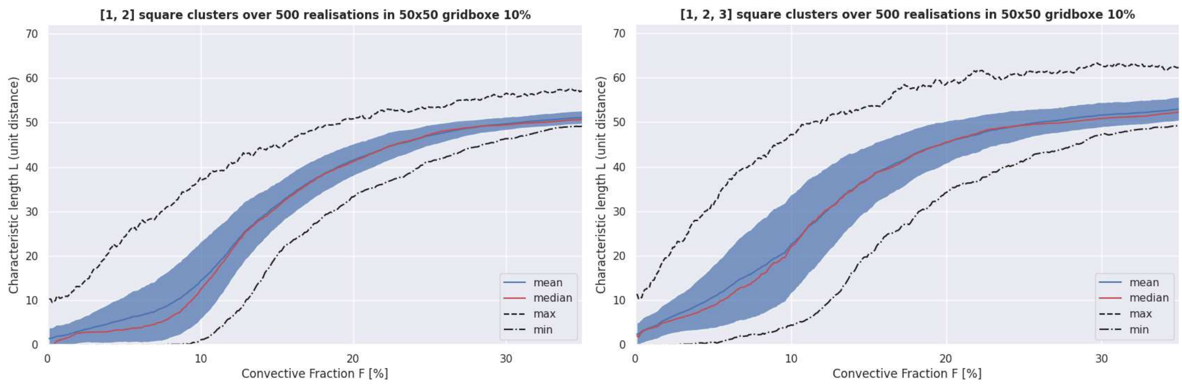
“Examples of the potential applications of this method in the realm of weather and climate research include severe weather and climate modelling. Establishing a statistical relationship between flooding (or strong winds events) and the internal organization of the system would help nowcasting of these events by completing the ingredients list (Doswell et al., 1996). Recent intercomparison exercises (Prein et al., 2024; Feng et al., 2025) have revealed biases between the latest generation of km scale global models and the satellite observed cloud shield morphology. The interpretation of these biases would benefit from applying our method to these simulations to relate the limitations of the simulated morphology to that of the convective organization as shown here.”

Decomposition into elementary structure

It appears to me that the decomposition into the elementary structures that is based on the square-like nature of the convective core has implications for the characteristic length L. Can you explain in more detail if this assumption is a limitation for the range of L values that you get?

This point was also raised by the second reviewer. First, we have changed the title to highlight the use of this strong hypothesis as stated before by adding “using idealized elementary convective structure decomposition”. Next, we assume that the referee requests details on the influence of the elementary decomposition on the P parameter rather than the L parameter, since the L parameter is independent of it. Only the L distribution of the generated scenes may be influenced by the assumptions of the decomposition.

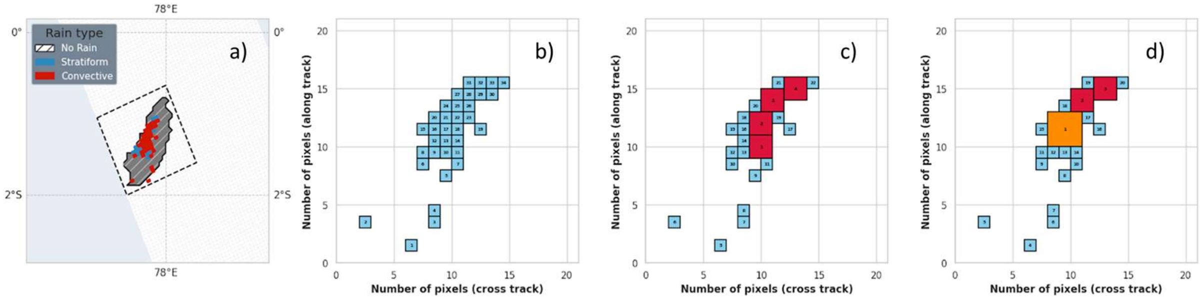
Firstly, autocorrelation is sensitive to dominant structures. Using structures composed of only one pixel is often not discriminating for low fractions (F). Figure 10 (left) shows the range of variation of L for a given F in a 50x50 grid with randomly positioned 1x1 structures. For low F values, the characteristic length values are all confined until F exceeds a certain threshold. The same is true when 2x2 structures are also permitted, with the range of F expanding more rapidly. The same pattern continues up to 5x5 structures (Figure 10, right). See Figure below that complements the Figure 10.



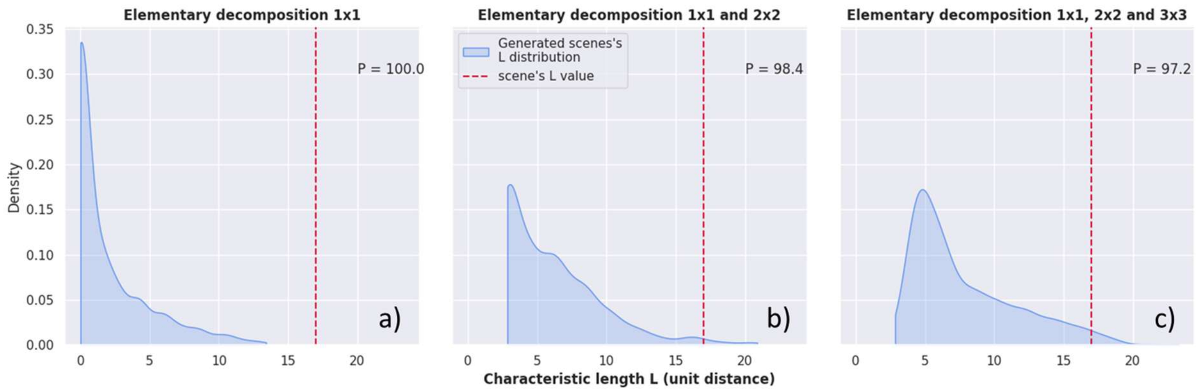
Same as figure 10 with Left: only square clusters of size 1x1 and 2x2 (MaxSquareSize = 2), Left: only square clusters of size 1x1 and 2x2 and 3x3 (MaxSquareSize = 3).

Besides, it is impossible to break down all scenes using only structures that are 2x2 or larger. Almost always, structures that are 1x1 must be used in addition.

The elementary decomposition impacts the calculation of L distribution for generated scenes, with the majority of the influence coming from the segmentation of dominant structures (larger squares or larger objects if a different decomposition is used). In fact, the more roughly the structures are cut up, the more sensitive L is to the main structures and their respective arrangements. Conversely, by cutting everything into finer elements, for example into 1x1 and/or 2x2 objects, which is also always possible, L will have a more confined range of values and will therefore struggle to express itself in a sufficiently broad distribution as a key parameter for the rest of the characterization process. We therefore have a much greater impact on P than on L itself here, which we have quantified in the example using the figure 3:



Same as figure 3 (left and right only) with decomposition only with b) square clusters of size 1x1, c) size 1x1 and 2x2 (MaxSquareSize = 2), d) size 1x1, 2x2 and 3x3 (MaxSquareSize = 3)



Same as figure 9 (left only) with decomposition only with a) square clusters of size 1x1, b) size 1x1 and 2x2 (MaxSquareSize = 2), c) size 1x1, 2x2 and 3x3 (MaxSquareSize = 3).

When decomposing with either 1x1 or 1x1 and 2x2 structures, P takes on higher values, suggesting that the organisation is less random. The L distribution shifts to higher values when using larger structures. In other words, the inclusion of large elements augments the probability for the final structure to occur from pure randomness compared to not including larger structures.

Differences between observations and model simulations

In the very beginning of the paper, the definition of “convective” vs. “stratiform” is introduced for models and observations. I am wondering how much of the differences that you find between the datasets actually go back to how you define “convective” in the model dataset. It makes sense to base this definition on physical processes such as updrafts, but I am afraid that, as you mention yourself, this would lead to significantly different regions than the purely radar-based signature in the satellite observations.

There is some ambiguity as to the definition of a convective column. In order to assess the sensitivity of our method to this ambiguity we have used two datasets each with a different definition. But we are not elaborating on the differences between the model and the observation or between the definitions.

In addition to that, I think that the paper would benefit from a more thorough discussion of differences between models and observations, as an example for how to apply the introduced method (i.e. that it can help to validate models that partly resolve convective processes).

While we are definitely interested in the suggestion, the focus of this paper is primarily on the impact of convective organisation on the cloud shield of DCSs. Therefore, the aim of this study is not to validate or compare models with satellite observations as such a comparison is not straightforward and would require a dedicated effort beyond the scope of the present paper.

Figure labels

The font size in all figures need to be increased. In some figures, such as, Fig. 3, the labels are barely readable.

Done, most of the figures/labels have been enlarged

Detailed comments

21: Variables -> metrics ? (variables sounds more like specific atmospheric fields whereas metrics specifies that you introduce and suggest a measure)

Yes, we decided to modify all the occurrences of “variables” by “parameter”.

27: *Idealized* kilo-meterscale numerical simulation

Done.

54 - 67: When reviewing the existing literature on convective organization indices, could you go a bit more into detail on what variables these indices are usually based on?

Yes, we added this sentence to give more specific context:

⇒ “Other existing metrics that are more object oriented such as COP, ABCOP and ROME (White et al. (2018), Jin et al. (2022), Retsch et al. (2020)) consider different convection attributes, like the object's area, mean area of objects, relative or geometric distance between objects, in addition to their center within a specific grid.”

1. 104: remove “…”

Done.

L,. 134 ff: When introducing the radar collocations, could you clarify which variables you are working with - is it only the retrieved rain rates or other retrieved quantities? Do you leverage the radar reflectivity values at all?

Here it is only the precipitation type classification that is at stake. Other variables available will be used in future works. All the details can be found in the following references (and the ones that were already cited):

Le, M. and V. Chandrasekar, 2013a: Precipitation Type Classification Method for DualFrequency Precipitation Radar (DPR) Onboard the GPM, IEEE Trans. Geosci. Remote Sens., 51(3), 1784–1790.

Awaka, J., T. Iguchi, and K. Okamoto, 2009: TRMM PR standard algorithm 2A23 and its performance on bright band detection, J. Meteor. Soc. Japan, 87A, 31–52.

We precise a bit as follow, adding the two references cited above:

⇒ "For our analysis, we focus on the classification of precipitation type into stratiform, convective and other rain type, retrieved from the 2A23 algorithm (Awaka et al. 2009, Le and Chandrasekar 2013, Awaka et al., 2016, 2021; Chen et al., 2025)."

144, 146: colocation -> collocation

Done

179: Is this classification based on radar reflectivity thresholds?

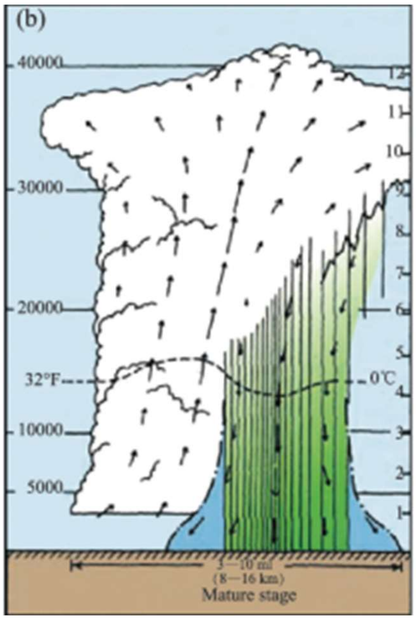
The classification process for 2A23 is very complex, as described in details in its ATBD: https://gpm.nasa.gov/sites/default/files/document_files/ATBD_DPR_201811_with_Appendix3b_0.pdf

We modified the incriminated part as follows: ⇒ "Within the radar swath, precipitation is classified as stratiform, convective, or non-precipitating, following the precipitation type from the 2A23 algorithm introduced before, and excluding the "other" rain type. Convective echoes are highlighted in red in Fig. 1."

1. 236: It makes sense that a grid cell can be classified as convective either when it contains heavy convective rainfall or when there are updrafts that indicate strong vertical motion (which may happen prior to and in a different grid cell than the heavy precipitation). However, it is not clear to me why a grid cell below the downdraft threshold would also be classified as convective. While strong downdrafts are part of the convective system as a whole, the processes are quite different than in updraft regions, so I would like to better understand what we gain from having strong updrafts and strong downdrafts in the same category. Would it make sense to leverage the information of hydrometeors in the model?

We thank the referee for pointing these out. We have not properly introduced the method of Marinescu and we shortcut way too much the methodology in the original manuscript.

Indeed the two thresholds correspond to two regions of the column and consider the absolute value of the vertical velocity (Marinescu et al., 2016). So it either can be updrafts or downdrafts. The fact that both can be classified as "convective" can be understood as being the two sides of a convective cell with a given life cycle as shown on the following figure of a mature convective cell (Markowski and Richardson, 2010). See also the marinescu et al paper and references therein.



A mature convective cell. from Markowski, P., and Y. Richardson, 2010: Mesoscale Meteorology in Midlatitudes. Wiley, 1–407 pp.

The original sentence "be also classified as convective if the vertical velocity above the 0°C isotherm exceeds the updraft threshold or falls below the downdraft threshold." as been modified in to: "be also classified as convective if the absolute value of vertical velocity updraft (or downdraft) above (below) the 0°C isotherm exceed 5 m/s (3m/s)"

Table 1 also is modified and now reads:

Convective Precipitation threshold (mm/h)	O°C isotherm height (m)	Cloud top height (m)	stratiform Precipitation threshold (mm/h)	Cirriform OLR threshold (Wm-2)	Updraft threshold above the 0°C isotherm (m/s)	Downdrafts threshold below the 0°C isotherm(m/s)
10	3000	6000	0.04	172	5	-3

Note that the algorithm was originally correctly implemented though.

1. 286: Can you explain the “high-resolution precipitation ground based core and updrafts/downdrafts joint occurrence analysis”?

The two mentioned papers (Moroda et al.; Lamer et al., 2023) report case study analyses of high resolution vertically resolved simultaneous measurements of both the dynamics and the precipitation within storms using Phased Array Weather Radars. These analyses reveal the interplay between the hydrometeor distribution and the vertical velocity, and emphasize the strong but complex relationship that takes place at small space and time scales. Moroda et al. build a conceptual model of the lifecycle of the precipitation cores and the updraft trajectory in the precipitation cell. Their analysis is based on a joint occurrence analysis of precipitation and updrafts/downdrafts.

To clarify our purpose the incriminated sentence is modified and now reads “precipitation core and updrafts/downdrafts joint occurrence analysis from high-resolution ground based observations.”

1. 282- 284: I understand the assumption that the hydrometeor distribution is the result of small-scale convective dynamics and cloud microphysics, and that, in other words, we can see convective cores as aggregates of these small-scale processes.

   But where does the assumption of a square-like geometry come from? Could it not be that these small-scale aggregates are, for instance, elongated like in squall lines, or is there evidence that convective cores have a circle- or square-like geometry with the same distance to all its surroundings?

The circular assumption relies on a vertically erected plume model which is a useful idealization of deep convection dynamics. In this simple case, the shape of the cross section of the plume is indeed circular. Yet if the plume is slanted (due to wind shear), the cross section of the plume would look like an ellipse. The circular assumption (which translates into square pixels) should be understood as an idealized, first order approximation. To better convey this dimension of our work the title of the paper has been modified.

Fig. 4: Is the core size given in the number of pixels?

⇒ Figure 4 caption was modified : “Size distribution of the convective cores Left: TOOCAN-radar dataset. Right: TOOCAN-RCE dataset. The Houze-like detection is shown in blue and corresponds to the rounded integer value of the square root of the cores' area, while our elementary decomposition is shown in red and corresponds to the size of the square cores.“

Fig. 5: Add in the figure caption that this is for the satellite data. Are those the composites over all tracked DCSs?

We modified it as follows in the text ⇒ “Figure 5 shows a comparison of the mean vertical profile of reflectivity for the present decomposition and that of Houze-like cores, using 8 years of collocated radar measurements to allow for robust statistics.”

And in the figure caption: “Figure 5: Reflectivity profiles (composites using 8 years of collocated radar measurements) of convective cores for Left: our elementary decomposition, Right: Houze-like shared edges decomposition, with different core sizes shown in color”.

1. 657: A new method for what?

The incriminated sentence was reformulated: “In this study, a new method is introduced to quantify and characterize spatial arrangement of convective areas within a specific grid encompassing the instantaneous cloudshield of a DCS.”

1. 727: Please check the DOI of your dataset. I could not access it.

We checked the DOI of our datasets and it seems to work from our side:

TOOCAN dataset: https://doi.org/10.14768/1be7fd53-8b81-416e-90d5-002b36b30cf8

IR GEO-ring dataset: https://doi.org/10.14768/93f138f5-a553-4691-96ed-952fd32d2fc3

1. 344: What is meant by this sentence: “As a consequence, unlike regular gridded data, the method is required to be shield-specific.” I am confused by the statement that the method cannot work for regular gridded data.

Thank you. We have modified the incriminated sentence into : “As a consequence, unlike regular gridded data of fixed size, the method requires defining a selected area that corresponds to the cloudshield dimensions”.

1. 371: It is not clear what the difference is between the "distribution of convection across the scene” and its “spatial arrangement”. As I understand it the first two metrics describe the area and amount of convection (“how much?”) and the last two methods focus more on the geometry and where within the cloud shield convection is present?

We modified it as follows: “The first two parameters describe the amount of convection within the scene, using both absolute and relative metrics.”

Fig. 11: The difference in the maximal area between the satellite data and the model simulation appears to be quite significant. Is this a consequence of the effective resolution of the datasets, and does this indicate that the model physics cannot reproduce the large cloud shields that we can observe? In addition, the duration of the DCs seems similar, which is interesting because the spatial and temporal scales should also be linked or not? I think it would be useful to discuss this point in more detail.

It is not possible to make a one-to-one comparison between the real world and idealized long-channel RCE (no Coriolis) simulations, for which the scales are not directly comparable with reality. So, it is not a problem of the model physics, but a problem of the numerical experiment set-up.

1. 544: I think it would be helpful to remind the reader what the four key variables are in the figure caption?

Done in figure caption.

L. 546: If I understood it correctly, K-means clustering is used to produce the four classes in Fig. 13 and these classes are based on the multivariate coherence of the four metrics.

Yes it’s totally correct.

It is, however, not explained in detail how the PDFs of each metric relate to the respective class (from random to organized). For instance, Fig. 13 a) shows quite distinct distributions for F between class 0 and 1 although these are the classes closer related to each other. I do not expect to go into the details of all possible combinations of the four metrics, but it would be decent to describe a little bit more how convective systems with substantially different distributions for F and P can still be more alike each other when they are similar in terms of L and S.

Yes indeed. This fact emerges from the inherent differences between class 0 and the rest, as class 0 corresponds to a more likely random distribution of the cores, while the three others have higher values in the P distribution. The classification from random to organized is based on the P distribution (and secondary on the A and L variables for which the order is the same in terms of means and modes of distribution), that is more and more close to 1 as the class number increases. Class 0 and Class 1 are not much closer to each other than any other Class, it is simply a distance between points in a 4-dimensional space that we are trying to interpret more physically.

⇒ We included a modification in the dedicated paragraph to express this point: “As P increases (and secondary L and A too), so does the level of organization.”