Dear authors,

Thank you for your revised manuscript. I have obtained one further review from one of the initial referees. From that and my own reading, we feel that you have addressed most of the points well, but the referee has some remaining concerns - particularly, about how the definition of scene area and orientation of a feature could lead to larger than expected errors. I am not sure if we have misunderstood something or whether the orientation with respect to the radar is indeed a potentially large issue. I have attached these comments (see pdf) and would appreciate if you would be able to respond to them and modify the manuscript as you feel appropriate. Other than these, I feel and the reviewer agrees that comments have been addressed well. Please let me know if you have any questions!

Best wishes,

Andrew

Dear Editor

Before we answer your and the reviewer questions, we would like to emphasize a few points. First of all, the first round of review provided a very interesting in-depth set of questions thanks to the appreciated efforts of two reviewers. The fact that, now, you think that there is a “potentially large issue” with our work is puzzling. We are surprised that you decided to call for major revisions after a round of long, yet minor revisions. The reviewer is using words like “critical”, “substantial”, etc. which we think are not relevant to the discussion at this stage.

First it is important to note that our work is not a house of cards; it will not collapse just because of one aspect. The only potential issue is the delineation of the scene area and the orientation of the shield when using space borne radar.

As emphasized in the original manuscript, the method does not only apply for such narrow swath scanning radars. Half of the analysis provided in the manuscript indeed concerns outputs of model simulations free from radar geometry. Our technique can be applied to various other satellite borne instruments, for instance from passive microwave observations of hydrometeors for which the swath width is much larger than the radar one as discussed for decades in the literature (e.g., Nesbit et al., 2003). So, if the geometry were an issue, it would not be a “*critical”* one for using the method. At best, it could be a limitation of the technique for these particular radar observations.

Also, the consistency between the results from the radar-based analysis or the model-based outputs rules out any “*large issue*” with the method.

The source of uncertainty pointed out by the reviewer can arise from specific alignment of the radar swath together with a specific aspect ratio of the convective cloud shield. As shown in our responses during the previous review round, the distribution of aspect ratio of the cloud shield is smoothly distributed across the range of possible values and is centered at 0.5. So only ~50% of the cases can qualify for alignment computations, and even fewer if very elongated systems are concerned. Therefore, in any case, this potential source of uncertainty cannot again be “*large*”.

This said, it is interesting to identify such a source of uncertainty and we thank the reviewer for pointing it out to us. The reviewer suggests that our estimations of F (and possibly of P) could be unrelated to the actual convective fraction and that such an uncertainty can impact our classification. In the following, we provide new computations confirming:

1) our estimates of the scene area are very consistent with the true cloud shield area
2) our estimates of F are very consistent with the true convective fraction
3) the uncertainty on the estimation of F does not impact significantly our classification
4) we also discuss the possible impact on estimating P and finally address the comments on the TOOCAN algorithm.

We think it is interesting to specify these few cases to warn the reader of a cautionary interpretation of the computation of F and P in some cases of the radar configuration. As suggested by the reviewer we have added a dedicated paragraph in the section “Limits of applicability” that reads:

L483: “In the case of the narrow satellite observations of the radar, the estimation of the shield area (and, by consequence, the F parameter) using an along track aligned rectangle yields a biased estimation of this area. Nevertheless, this 'algorithmic' variable correlates very strongly with the true area and does not overall impact our methodology. But, in the case of elongated systems (eccentricity < 0.3) with a specific alignment with respect to the along track, in the range of 30° to 45°, this effect could be significant. We underscore the need to use our technique with caution for these infrequent cases: combination of cloud shield morphology and relative observations configuration.”

Reviewer's comments

I appreciate the authors' efforts in addressing my previous comments and revising the manuscript based on the initial feedback. While most of the authors' responses satisfactorily address my concerns, some aspects regarding the robustness of the methodology remain questionable, despite the additional sensitivity analysis provided. In this analysis, the authors examine how much the classification changes when noise is applied to four parameters — the convective fraction (F), convective area (A), characteristic length (L), and percentile (P). This is a valuable addition, as the manuscript presents a methodological approach. However, the analysis only assesses the method's sensitivity to errors in these parameters, without evaluating the typical magnitude of such errors in practice.

We agree with the reviewer (and have clearly expressed in the manuscript) that we only assessed the method's sensitivity to errors/uncertainty in these parameters. We also have clearly stated that evaluating the magnitude of such errors in practice is not an easy task. Were new insights from new research provide such an evaluation, our uncertainty propagation technique could easily be replicated to incorporate such new information.

We show next that i) our estimate of the scene area is consistent with the true cloud shield area, ii) our estimate of F is consistent with the true convective fraction (Response 1). We implement a geometry-free estimation of F (that is not without its own limitations) that we use to elaborate more on the possible uncertainty on estimating F and propagate it through the algorithm to quantify its effect on the classification, which is small (Response 2). We do this after clarifying a little misunderstanding on the propagation technique and the 10% threshold mentioned in the sensitivity analysis.

Based on the authors' response regarding the definition of the scene area, I am concerned that typical errors in F (and possibly P) could exceed the 10% threshold mentioned in the sensitivity analysis.

There is a small misunderstanding about the interpretation of the 10% threshold. As described in the manuscript, we have propagated uncertainties spanning + or - 1%,5%,10% etc. The analysis revealed a value ~10% , above which the final classification is significantly impacted. Below this value the classification hardly changes. In propagating the 10% uncertainty in the F variable, ALL the cloud shields are perturbed in our computations using a uniform distribution. So, the 10% value has to be interpreted as a threshold for the whole set of cloud shields and cannot be compared directly, with for instance uncertainty that would impact a subset of the cloud shield population. Said otherwise, the effect of having 2% of the systems out of 50000 with a 100% uncertainty is very different (much less) than having a uniformly distributed 2% uncertainty on the 50000 shields.

I encourage the authors to either provide evidence that this is not the case or revise their methodology for defining the scene area. I also have follow-up (more minor) comments regarding the interpretation of P and the sensitivity to the tracking algorithm. Please see my detailed comments below

We discuss the impact of the area selection on P (and L) and clarify the misunderstanding about the interpretation of P (Response 3) before closing on the influence of the tracking algorithm comment (Response 4).

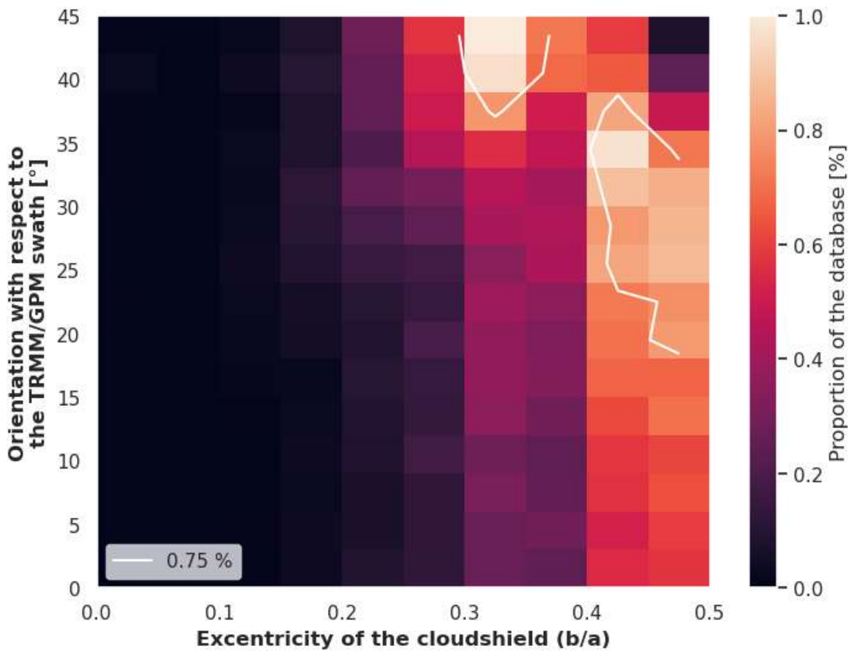
Scene area

I thank the authors for their detailed explanation of how the scene area is defined for each case. To my understanding, the scene area corresponds to the smallest rectangle that fully encloses the DCS (with one pixel of padding) and is aligned with the radar swath grid. I see two potential sources of bias in this definition: (1) the shape of the DCS, particularly its deviation from a rectangular form, and (2) its orientation relative to the grid.

Non-rectangular DCSs that are oriented at approximately 45° relative to the satellite grid will result in significantly larger scene areas than those that are rectangular and aligned with the grid. For the method to yield a meaningful classification of convective element arrangements within DCSs, such orientation- and shape-related biases are undesirable.

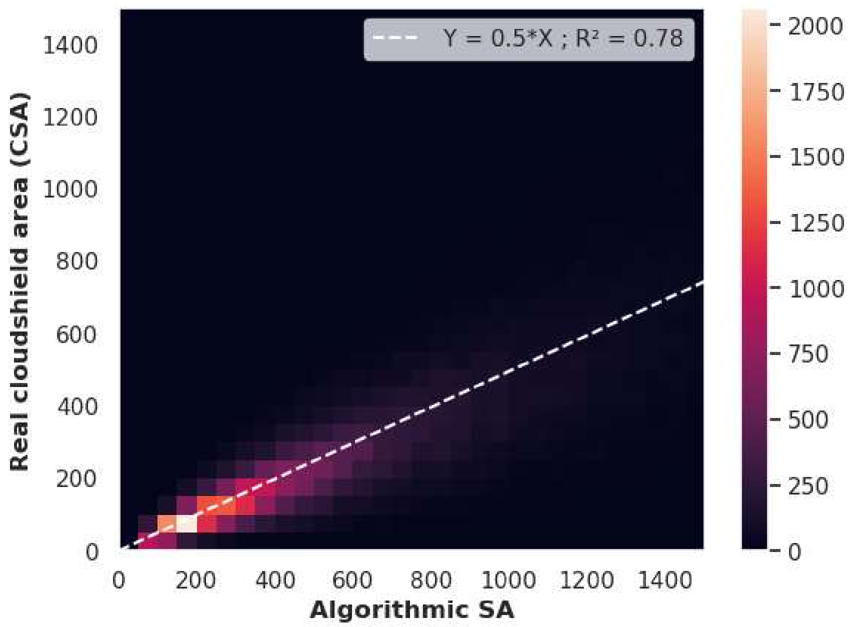
We first introduce the distribution of the cloud shield in the eccentricity-orientation phase space diagram. The eccentricity is the classical way to characterize the shape of the DCS and is readily available in the TOOCAN outputs. It is used in the following to estimate the departure from a rectangular shape as highlighted by the reviewer. The lower the eccentricity the more elongated the DCS; DCS with eccentricity above 0.5 can be considered as quasi-circular and DCS with an eccentricity below 0.2 can be thought of quasi-linear systems (e.g., *Liu, C., and E. Zipser, 2013: Regional variation of morphology of organized convection in the tropics and subtropics. J. Geophys. Res. Atmos., 118, 453–466, https://doi.org/10.1029/2012JD018409.*). Given the growing uncertainty in estimating the eccentricity and the angle for quasi-circular objects, we restrict ourselves to the eccentricity below 0.5 for which both parameters can be estimated with confidence. The (relative) orientation is computed as the angle between the major axis of the DCS and the TRMM/GPM swath (between 0 and 45°).

The figure R1 below shows the distribution of systems in this phase space. In the lower range of the eccentricity, the systems are observed with a wide range of orientation yet local maxima are found in the upper range of the orientation distribution. Hardly no systems are found with eccentricity below 0.15 (as already shown in the previous round of reviews). The case study put forth by the reviewer corresponds to the local maxima at 40-45° and eccentricity ~0.3.



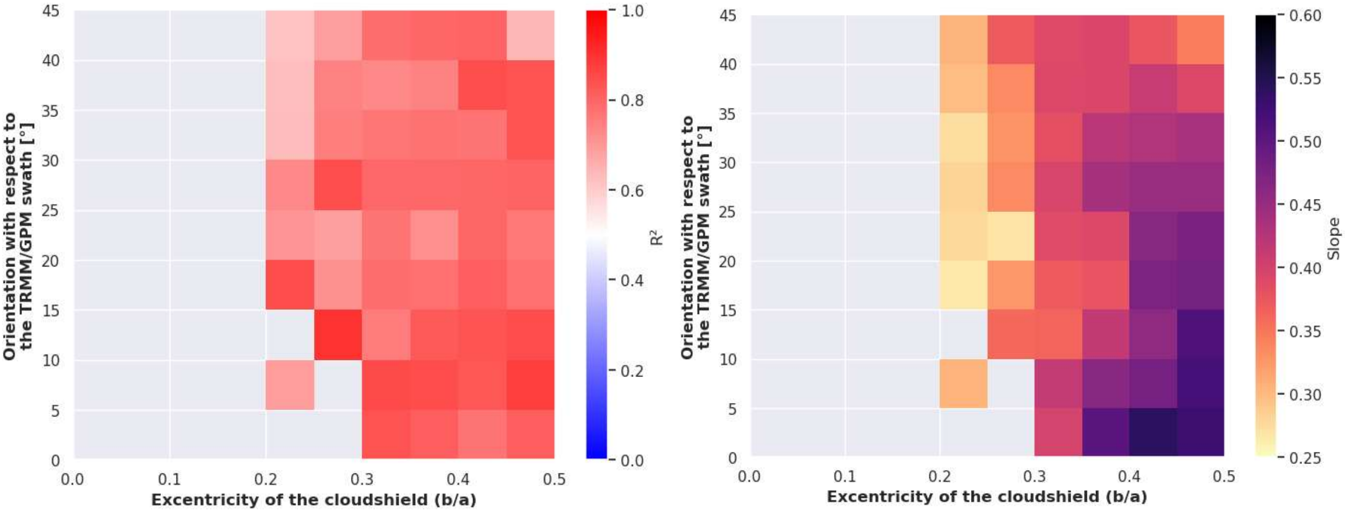
[Figure R1:] *2D histogram representing the joint distribution of the relative orientation angle with respect to the TRMM/GPM swath against the aspect ratio*

Before exploring the distribution of F or the impact on P we focus first on the distribution of the cloud shield area (CSA) and its estimate through scene area (SA) using our rectangle-based method. Figure R2 shows the diagram between the actual cloud shield (CSA) and the algorithm estimate (SA). There is a strong relationship between the 2 variables with a R^2 around 0.8. The regression slope is 0.5 indicating that the algorithmic CSA is roughly twice as large as the true A, which is expected.



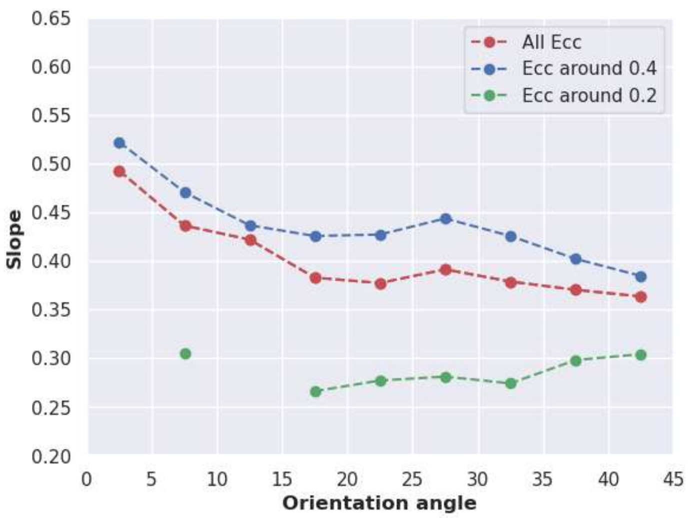
[Figure R2:] *2D histogram representing the joint distribution of the real cloud shield area (CSA) against the algorithmic scene area (SA)*

To rule out any specific orientation-geometry bias in the SA estimate with respect to the true CSA value, we have computed the linear correlation between the two across the phase space. The data have been selected in boxes of eccentricity and orientation and only the cases with at least 100 points are used in the computations. Figure R3.a indicates that in most of the cases, the R^2 is high without any region-specific departure.



<u>Figure R3:</u> 2D histogram representing the joint distribution of the aspect ratio against the relative orientation angle with respect to the TRMM/GPM swath for a) the linear regression coefficient R^2 and b) the slope of the linear regression fit between SA and CSA

As shown in Figure R3.b, the slope of the regression does reveal regime specific values which are investigated next more quantitatively. The figure below shows the regression slope for a given eccentricity, for a very elongated and an elongated system (Ecc =0.2 and 0.4) along with the mean of all Ecc values below 0.5.



<u>Figure R4:</u> Extraction of binned values of the slope (from Figure R3) against the orientation angle for three given Ecc (aspect ratio) values shown in colors.

For the very elongated case, the slope does not change much with the orientation angle. Similarly for angles above 15° where most of the cases are found, the less elongated system case shows not much change with the orientation angle. Below 15° there are few cases and a slight tendency for the slope to decrease with the angle.

The lack of a significant dependency of the regression slope with the orientation angle tells that for a given eccentricity, the bias of the estimation of SA shows no strong relationship with the orientation angle ruling out a possible large effect on the method.

While it is acceptable for F to be treated as an algorithmic variable representing the convective fraction, this remains valid only if there is a clear and consistent relationship between F and the actual convective fraction (i.e., larger F values should correspond to higher convective fractions).

We replicate below the same analysis for the F parameter. We plot below the joint distribution of the real convective fraction as a function of the algorithmic F for the full population of DCS.

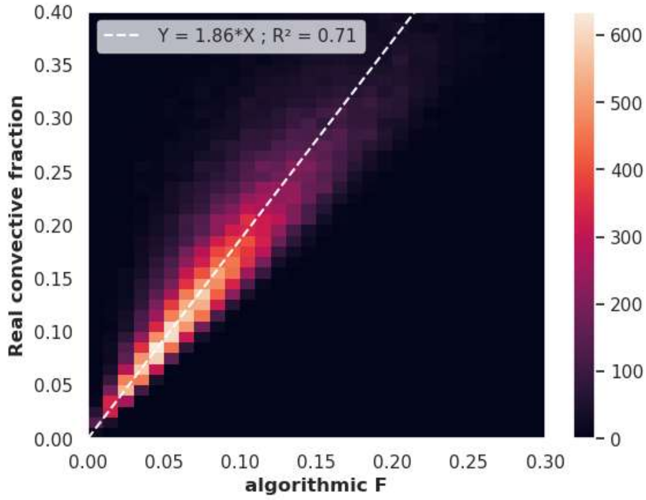


Figure R5: 2D histogram representing the joint distribution of the real convective fraction against the algorithmic convective fraction (F), the linear regression is shown with the white dashed line

This plot indicates a clear and consistent relationship between the real convective fraction as a function of the algorithmic F with a strong correlation of R^2=0.71. As anticipated the real convective fraction is mostly larger than the estimate, or equivalently the algorithmic area is mostly smaller than the true area.

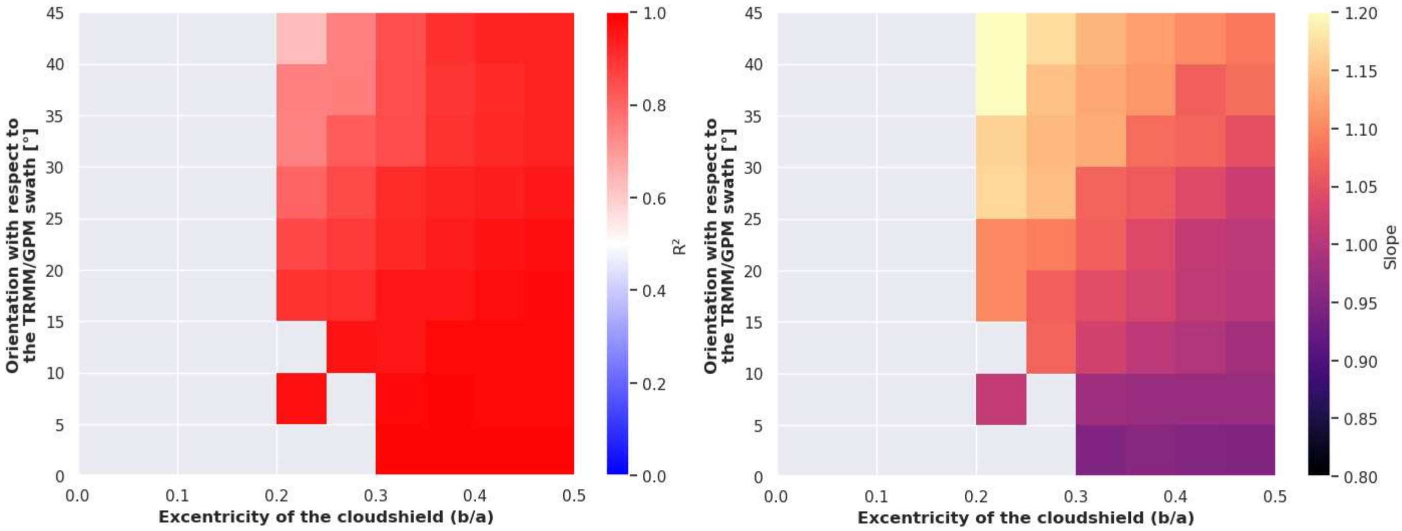
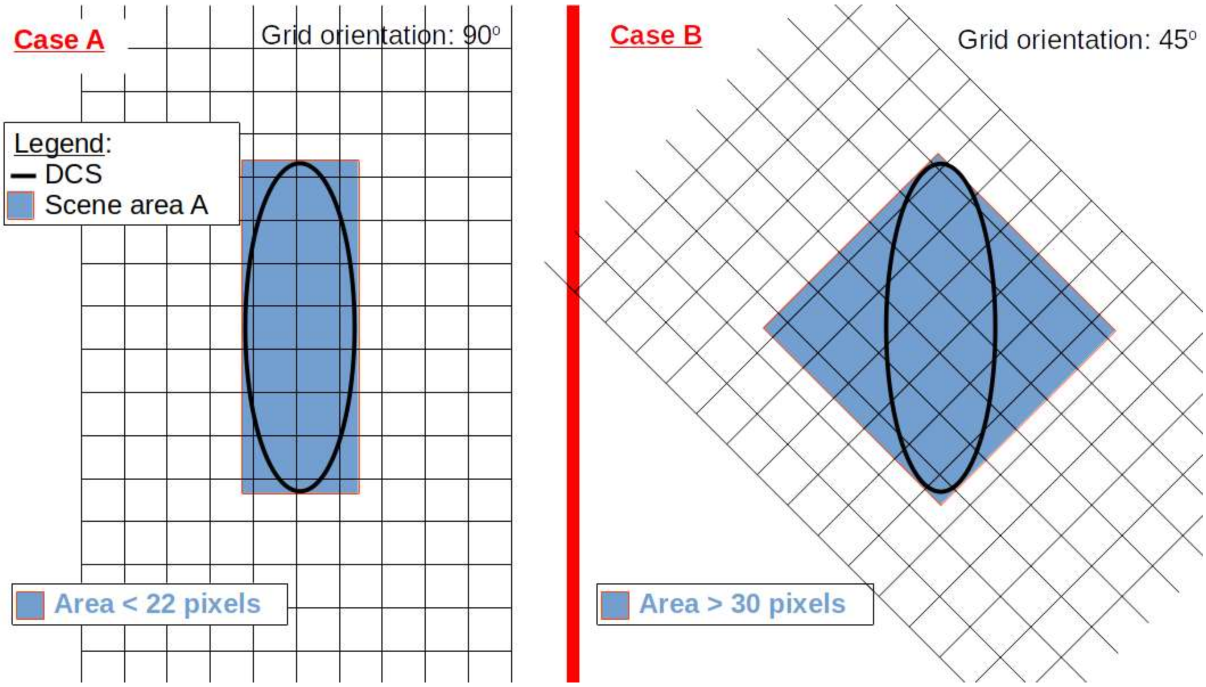


Figure R6: *2D histogram representing the joint distribution of the aspect ratio against the relative orientation angle with respect to the TRMM/GPM swath for a) the linear regression coefficient R^2 and b) the slope of the linear regression fit between the true convective fraction and the algorithmic convective fraction F*

To rule out any specific orientation-geometry bias in the F estimate with respect to the true F value, we have computed the linear correlation between the two across the phase space. The data have been selected in boxes of eccentricity and orientation and only the cases with at least 100 points are used in the computations. The figure above indicates that in most of the cases, the R^2 is rather high without any local bias. For the very rare cases with eccentricity of 0.2 and angle above 35 we observed a less strong correlation (although significant with R^2~0.6). The range of variation of the slope at a given eccentricity is small and following the lines of the arguments for SA, this rules out also in the case of F any large effect on the results.

Based on my understanding of the current methodology, I am concerned that this relationship may not hold given the current definition of scene area. To illustrate this point, consider a simplified example of an elliptical DCS (aspect ratio ≈ 1/3) observed by two radars with different grid orientations (cases A and B in the figure below). As it is the same DCS, the method should ideally yield similar convective fractions and classifications in both cases.



However, as shown, case B produces a significantly larger scene area than case A — a difference of approximately (30–22)/22 = 36%. This increase in scene area could lead to a decrease in the derived convective fraction by about 26%, potentially resulting in different classifications for the same physical system depending solely on its orientation relative to the radar grid (as suggested by the additional sensitivity analysis). Differences in DCS geometry would add on this orientation bias, making the convective fraction F less representative of the actual physical convective fraction.

This is opposed to assuming an algorithmic F. The two cases should not give similar values, they should give similar sensitivity across the population. Or formulated in another way, the uncertainty resulting from the estimation of F should not influence the results of the overall unsupervised classification but there is no need for various F estimates to be similar. In addition, as discussed in the introduction there is a little misunderstanding in interpreting the 10% threshold of our sensitivity analysis. Nevertheless, we have tried to explore alternative ways to compute F as one source of uncertainty in estimating F and in the next section we detail the effect on the overall classification.

A geometry-free F estimation is obtained by fitting the best rectangle using the cloud shield coordinates in a two-dimensional space free from the resolution. This is achieved by allowing rotation and non-integer width and height (in the pixel space). This permits the estimation free from the constraints of the radar grid. Therefore, it should be noted that finding the optimal rectangle this way is not always feasible, as one of the rectangle's dimensions will often exceed the radar swath width, creating 'information-free' areas. Nevertheless, keeping these limitations in mind, we have computed F this way for all the systems. The comparison between the two estimates is shown below in the form of a relative difference computed as (F_new-F_old)/F_old.

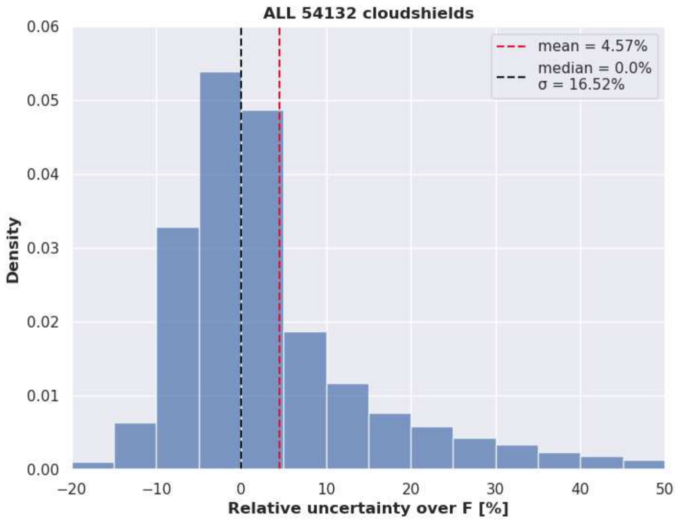


Figure R7: Histogram of the relative uncertainty over F between the two methods (in %). The black dashed line indicates the median of the distribution while the red one illustrates the mean

The median of the distribution is 0%, and the mean is 4.57%, with a standard deviation of 16.5%. The distribution is also asymmetric. The configuration mentioned by the referee (high aspect ratios and orientations close to 45°) are found on the positive tail of the distribution. We also note that sometimes negative relative errors occur, i.e., a decrease in F. There is therefore no systematic bias of one method over another, with a slight tail effect on the high values. The influence of the most extreme cases, such as those cited above, is indeed visible in the high values in the relative error, but it remains limited by the effects of numbers when we return to the total distribution of F. We note that unlike what we did in our error propagation technique, the distribution is here non uniform.

We therefore calculated the following metrics characterizing the robustness of the final classification in the same way, but using the actual distribution (no more uniform assumption) for the relative error injection calculated previously, and obtained the following results: silhouette score: 0.3206; ARI: 91.1%; proportion of mismatch: 2.96% (i.e., 1 602 scenes over 54132). By plotting these values on the sensitivity study graph shown in the previous answer, which is reproduced below, we obtain a value of ~5% equivalent injected error, which is well below the 10% threshold we had set as the limit (see cyan diamond in the following figure).

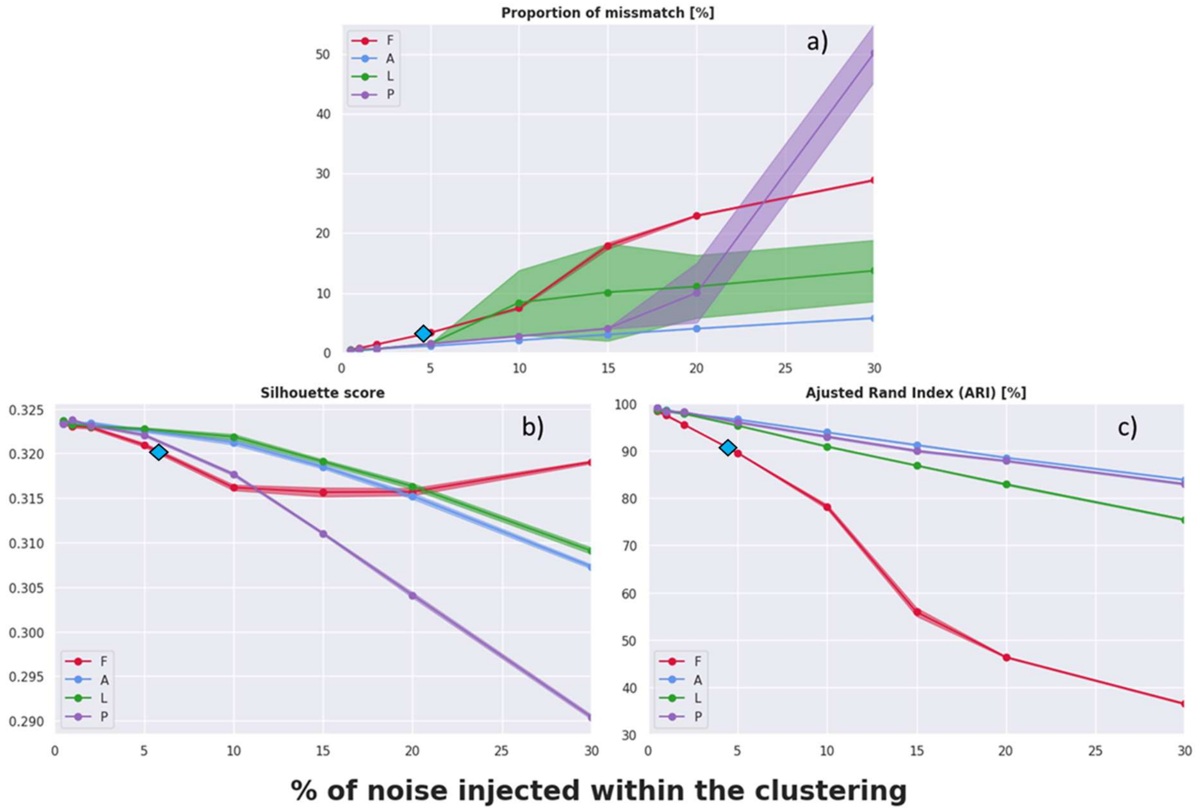


Figure R8: *Same as Figure B3 (see the previous round of responses). The blue squares indicate the corresponding values for each of the indicators, replaced on the F curve (in red).*

In conclusion, assuming an alternative way to compute the area is giving a reasonable uncertainty distribution (without being better or worse), then only a few percent of mismatches are observed, further ruling out, quantitatively, any large effect of this geometry-orientation problem.

In addition, I recommend that the authors include a clear description of how the scene area is defined in the manuscript, as this aspect is critical to the methodology and may substantially influence the classification results.

We modified the L374-375 as follows: "A scene is defined as the rectangular area that encompasses the cloud shield identified by the TOOCAN algorithm, as illustrated by the black contour in Figure 3 (left)." ⇒ "A scene is defined as the rectangle that most closely frames the outline of the cloud shield identified by the TOOCAN algorithm, as illustrated by the black contour in Figure 3 (left). For the TOOCAN-Radar dataset, this rectangle must be oriented in the direction of the satellite's orbit within the TRMM/GPM swath geometry. It is not possible to extend beyond this geometry because radar information is missing there. When the data is available, a margin of 1 pixel is also used at the edges."

Finally, the limitations associated with the definition of scene area should be discussed in the "Limits of applicability" section.

We have added one paragraph in the form of a cautionary note in the mentioned section to summarize the discussion. See the response to the editor at the beginning of the document.

Figures:

Some figures still lack panel labels (letters) for reference. This is particularly the case for Figure 14, where letters are mentioned in the caption but do not appear in the figure itself. For clarity and consistency, I recommend adding panel letters (e.g., a, b, c, …) to all relevant figures, as these are preferable to references such as “left” and “right.”

We have modified Fig. 14’s caption accordingly.

Large variations in scene area would also influence the P metric, which is likely to be larger for case B than case A, possibly exceeding the 15% threshold reported in the sensitivity analysis.

Uncertainty in SA could influence not only F, but also P and L. There are no obvious reasons for P to be larger in case B than in case A. Indeed, P depends on SA and A (note that A is the only parameter independent of geometry), as well as its decomposition into elementary structures and L. Therefore, it is not possible to make a direct estimate using the given example.

I encourage the authors to either (1) provide evidence that typical errors in F and P do not exceed the thresholds discussed in Figure B3, or (2) revise their method to reduce these biases.

Following the previous discussion, the same lines of arguments can be put forth for P as we did for SA and F and we think that the geometry-orientation cannot yield to a large issue in P. But unlike for these variables there is no straightforward way to redo the computations. Indeed, to remove the dependency on the angles would require reprojecting the data in the non-regular grid before reprocessing the archive, changing the decomposition and the scene generation without guarantee to be free of problems.

Beyond the fact that it is important to interpret the 4 variables together and not apart, as already argued, we would like to emphasize the covariation among the variables. The classification would be influenced by the covariation of the 4 variables together. Given the strong relationship between algorithmic SA and the true SA, there are no obvious reasons to imagine that covariation in the 4D space is disturbed by the geometry-orientation. So, it is difficult to anticipate the amount of misclassification.

Also like we did for the F explanation, the fact that there is a strong consistency between the model-based results, free of geometry-orientation artefact, and the radar case indicates that if any effect there is, it is not large.

<u>Interpretation of the percentile P:</u>

- L. 465-466 : “computing a percentile P, which quantifies the deviation of the real scene’s spatial arrangement from a random arrangement (Figure 9, Red arrow in Figure 6).”

That wording implicitly assumes that higher percentiles mean stronger deviations from random.
But that is not necessarily true, low percentiles can also represent strong but opposite direction deviations (more regular patterns instead of random). Please reformulate.

- L. 470-473 :

For clarity, I copy the Authors’ response to my comment on the previous round of revision.

“We don’t use the clustering notion to illustrate the meaning of the value of P. Besides, rarity is somehow a vague notion that can lead to ambiguous interpretations. We don’t mention it directly. Instead, we express the notion of how many times an event occurs within the total number of attempts in the context of the generated scenes, as the percentile computation. Besides, as described in the paper, our objective is broader than simply distinguishing between aggregated and clustered, as with metrics such as Iorg/Lorg, for example. Rather,we aim to quantify the deviation from randomness and the structuring of convective cores. Therefore, we believe that the original phrase is appropriate in this context."

I understand that the objective of the paper is broader than simply distinguishing between aggregated and clustered patterns, and the proposed method indeed achieves this, as indicated by the combination of P with the three other variables. However, my comment focuses specifically on the variable P.
The variable P measures the fraction of randomly generated scenes that have a lower L than the actual scene. Therefore, to my understanding, the variable P, as defined in the current version of the manuscript, is more of a measure of the degree of clustering expressed in a statistical sense (on a probabilistic scale) than a measure of the deviation from randomness. This is because both very low and very high values of P indicate strong deviation from randomness, as noted above. Conversely, as P increases from 0 to 1, the spatial arrangement transitions progressively from more regular to more clustered scenes. If the Authors would like to put more emphasis on the deviation from randomness than the degree of aggregation, I would suggest to use a modified metric, for example 1-2.min(P,1-P), that will more closely quantify this aspect.

- L.539-540 "On the other hand, the P distribution, indicating the probability of the characteristic length to differ from a random arrangement (Figure 12, d) is very much skewed to the highest values (around 1) with a very long tail. This indicates that only very few scenes can be associated with a randomly generated pattern."

The P distribution does not directly indicate a probability of differing from a random arrangement. First, very low P values are also indicative of deviation from randomness, as noted above. Second, such P-values (percentiles) only quantify the conditional probability of observing a value less extreme than the observed statistic (here L) under the null hypothesis (here, H0 = randomness (H1 = non-randomness)). Formally, they measure Prob(L <= Lobs | H0), which has no direct relation with P(H1|L<=Lobs) = 1 – Prob(H0| L<=Lobs).

We agree and modified the wording on the incriminated sentence:
"On the other hand, the P distribution is very much skewed to the highest values (around 1) with a very long tail (Figure 12, d). This indicates that most of the scenes exhibit a characteristic length that cannot be easily reproduced by random generation".

To illustrate, consider a bag containing many dice — some fair, some modified. You randomly draw one die, roll it, and obtain a six. Under the null hypothesis of a fair die, the probability of rolling a six is 1/6, and thus the P-value is relatively high (not indicative of deviation). However, this outcome alone tells us nothing about whether the die was fair or not. The P-value only quantifies how likely the observation is if the die were fair; it does not give the probability that the die is unfair given the observation.
Similarly, in your analysis, the P metric does not directly represent the probability of nonrandomness.

While the mapping between P and the probability of deviation from randomness may appear more meaningful when the reference (random) distribution is approximately bell-shaped, there is no one-to-one correspondence. The interpretation of P thus depends both on the shape of the null distribution and, in a Bayesian sense, on the prior probability of non-randomness, which can vary across cases.

- L. 574 : “very unlikely to be randomly spatially distributed with high L values”

This statement remains somewhat ambiguous. It could be misinterpreted as referring to the unlikelihood of observing L relative to all high values in a random distribution, rather than the intended meaning: that the observed spatial pattern is specifically unlikely under the null hypothesis and falls on the side of clustered rather than regular patterns.

Again, we do not share the referee’s vision around the interpretation of P. Figure R9 shows an example of a regularly spaced convective field where P has high value, illustrating that P alone is not particularly interpretable. It has to be compared to the 3 other key parameters to be understandable, as we state all along the manuscript. P is not comparable to an integrated metric like Iorg or Lorg, as we previously expressed in the manuscript and the previous round of response.

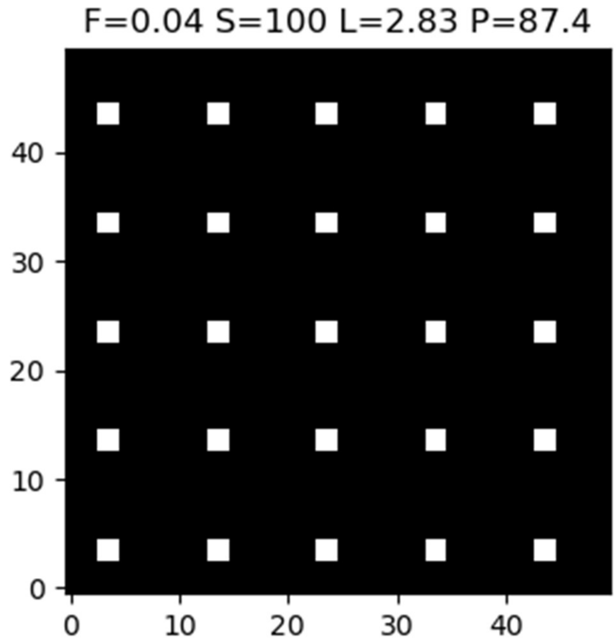


Figure R9: Idealised case of regularly spaced convective field. The 4 key parameters are computed and shown on the top of the figure, with specifically low L and high P

Sensitivity to the TOOCAN seed definition:

In the previous round of review, I shared my concern that the method used for detecting DCSs in the TOOCAN algorithm may restrain the diversity of spatial arrangements of DCCs, even when these DCCs are identified through radar measurements.

While the response provides a useful description of the TOOCAN procedure, it does not demonstrate that the IR-based identification of convective systems is fully independent of radar-based DCC detection (microwave measurements).

We do not understand this statement. We emphasized already that our study is TOOCAN-dependent and that we provide enough information for the interested reader to repeat our analysis using any other DCS identification algorithm.

Furthermore, it should be emphasized that any identification procedure (single threshold, Detect and Spread in 2D, or TOOCAN like approach) would identify the same features in the IR if the system is a well-defined, isolated large storm, like for instance the MCC of Maddox et al. (1980). For this case, the DCS delineation technique does not influence the detection much. TOOCAN departs from alternative techniques in the more complex cases where multiple systems do merge and split. It is also in this case that TOOCAN provides a more physically consistent depiction of the cloud shield than other techniques as shown in previous, already cited works. In particular the interested reader can find such examples based on idealized case studies described in Prein et al (2024).

Given that cloud-top temperature and convective precipitation intensity are often correlated (e.g., Arkin and Meissner, 1986),it seems surprising to me that the TOOCAN definition of DCS, which emerge from single convective seeds, would have no impact on the spatial arrangement of DCCs identified in radar.

We are very puzzled by this statement. The reviewer refers to a well-known work between accumulated surface precipitation and cold cloudiness area at the monthly and 250x250km scale to support his statement that at the instantaneous scale there is an impactful relationship between the TOOCAN seeds (cold IR temperature) and the deep convective cells (no surface precipitation).

The lack of correlation at short and small scales between IR cloud top, convection in column and surface precipitation has been well described in the literature and is the fundamental reason to use radar (or passive microwave) for surface precipitation estimate. For instance the seminal work of Yuter and Houze (YUTER, S. E., and R. A. HOUZE JR, 1998: The natural variability of precipitating clouds over the western Pacific warm pool. Q. J. R. Meteorol. Soc., 124, 53–99, https://doi.org/10.1256/smsqj.54503.) addressed this issue in a quantitative manner. More generally, the interested reader can find some recent updates on the satellite-based estimation technique and how the limitations of the IR can be overcome in various review articles. A good entry point could also be the joint GEWEX/IPWG assessment report on satellite-based precipitation estimates (Roca et al., 2021).

The irrelevance of the first statement “given that” implies that the syllogism used by the reviewer in this comment is actually not valid.

Since convective seeds in the TOOCAN algorithm are defined as very cold brightness temperature minima, and only one seed is retained per DCS, one would expect the radar identified DCCs to be spatially close to these IR-based seeds. This makes it difficult to conceive how the TOOCAN DCS definition could have no influence on the spatial distribution of radar-detected DCCs.

We suggest that there is possibly a misunderstanding induced by the loose terminology used in the IR based world and the radar world. Convective seeds refer to cloudiness with a cold top while deep convective cores correspond to the column. In particular the TOOCAN seeds are defined over 3-time steps (1h30 in the present case) and with a minimum extension of 625km2 (Fiolleau and Roca 2013, Roca et al 2017, the present manuscript) which is not the same as instantaneous, km-scale convective cores depicted by the TRMM radar.

A brief quantitative check (e.g., distance statistics between TOOCAN seeds and radar DCC centroids), or reference to an earlier validation study would help clarify this point.

The radar observations are available only at the overpass time and not every time step of the life cycle of the DCS. So, statistics between TOOCAN seeds and radar DCC centroids makes no actual statistical sense.

Reference:
Arkin, P. A., and B. N. Meisner, 1987: The Relationship between Large-Scale Convective Rainfall and Cold Cloud over the Western Hemisphere during 1982-84. Mon. Wea. Rev., 115, 51–74, https://doi.org/10.1175/1520-0493(1987)115<0051:TRBLSC>2.0.CO;2.