Thank you for your additional comments. Line numbers in our responses refer to the most recent revision of the manuscript. Italics denote reviewer comments quoted directly from the review; note that line numbers within reviewer comments refer to the previous manuscript. Grey boxes contain excerpts of the manuscript where changes were made, with blue indicating added text and ~~red indicating deleted text.~~

Response to Comments

*"**L43** This has been mentioned multiple times in the previous review, but I still feel obliged to ask. The authors have noted multiple times that different measures of fractal dimensions in this paper are meant to "characterize statistical relationships between scales rather than statistics at any particular style" (line 43), but what exactly is it that it is supposed to measure? It has to be a measure of something, and despite this being mentioned in both reviews, the paper still is not quite clear about it.*
The authors seem to oscillate between it being an "objective alternative to a subjective classification scheme" (line 347), and a measure that can be used to evaluate "the accuracy of atmospheric numerical simulations" (line 349). I am not sure if any of these claims are justified by the work presented in this paper. For example, as mentioned in the other review, it has to vary based on some objective state of the cloud field, which I believe should be some turbulent statistics of cloud regions. The paper shows that there is a consistent measure across different scales, but what if it does not change over different atmospheric conditions?
From the response from the authors, these claims seem to be more or less speculative. If that is the case, I believe it should be made clear that they are not what is being presented in this paper. "

We previously emphasized that physical explanations are proposed elsewhere but are out of scope here. Throughout the manuscript, several geometric and mathematical explanations are already provided. Mathematical definitions are provided at line 110 for the individual fractal dimension, 210 for the ensemble fractal dimension, 230 for the power-law distribution exponent for perimeters, 249 for the box dimension, and 265 for the correlation dimension. These mathematical definitions are the most precise explanations of what the metrics are actually measuring, but we also provide more qualitative explanations, designed to give a reader intuition, for what the metrics measure:

- L114 and Fig. 2: "Visually, self-similarity implies that the shape of an object such as that shown in Fig. 2 repeats at many different spatial scales."
- L139: "On an intuitive level, any given cloud in a cloud field can be considered to have two independent geometric properties: size and shape. By shape, we mean the roughness or complexity of the cloud boundary, which is perhaps the most visually notable aspect that differentiates clouds, such as those shown in Fig. 1. To quantify cloud shape,"
- L219: "the ensemble fractal dimension D_e applies when "lumping all the islands' coastlines together."
- L232: "a field with a large number of small clouds relative to large clouds would have a larger value of P and a higher value of β"
- Fig. 7 shows the visual interpretation of D_e
- Figs. 8 and 9 show visually how the box and correlation dimensions are measured
- L337: "the individual fractal dimension D_i that characterizes the geometric complexity of individual clouds, and an ensemble fractal dimension D_e that captures how the cloud field organizes hierarchically into structures spanning a wide range of sizes and shapes"

As for our suggestion that fractal dimensions could both be objective alternatives to classification schemes and also used to evaluate numerical simulations, the two claims are compatible with each other, so there is no "oscillation". As already cited in the manuscript, the fractal dimensions have been used for both use cases by other studies, so neither use case is speculative. L60: "fractal dimensions and size distribution exponents have previously been used to classify individual cloud types such as cirrus and cumulus

(Batista-Tomás et al., 2016), cloud field types such as sugar and gravel (Janssens et al., 2021), and compare simulated and observed clouds (Siebesma and Jonker, 2000; Christensen and Driver, 2021; Raghunathan et al., 2025)".

We already make clear that it is *not* the goal to use the metrics to evaluate simulations or meteorological conditions, but to simply refine the methodology by which the metrics, which are already in use, are computed:

- L57: "The aim here is to introduce simple and objective metrics for defining the geometries of clouds and cloud fields. ... Not all of the metrics we discuss are novel, as various fractal dimensions and size distribution exponents have previously been used"
- L75: "We also aim to examine and improve the methodologies by which the fractal dimensions are calculated in order to give future model intercomparison or cloud classification studies better tools."

To make the conclusions more explicit on this point, we modified L350:

> Calculation of the ensemble fractal dimension D_e bypasses these issues and may offer a more objective alternative to a subjective classification scheme such as sugar, gravel, fish, and flowers (Stevens et al., 2020), particularly if calculated as a correlation integral. The ease of its calculation using the correlation integral in satellite imagery makes it well-suited for ~~evaluating~~ future evaluations of the accuracy of atmospheric numerical simulations or for comparing regional meteorology.

As for the reviewer's question "what if it does not change over different atmospheric conditions?", fractal dimensions have already been found to vary as cited at L60 and quoted above. But even if they did not vary between meteorological conditions, they may still vary between simulations and observations, in which case they would still be useful as a metric. The Christensen and Driver (2021) paper cited at L60 is a prime example of such a use case.

*"**L191** I am assuming that the slopes have been obtained by least-squares linear regression, with the uncertainty being standard error estimate?*
1. I am asking because I can only guess how these numbers were obtained, especially the uncertainty in the slope of the linear fit, other than that they were obtained from "a simple linear fit" (line 132), and a different set of slopes were obtained "using a least-squared linear regression" (line 248). It may be obvious to the authors, but I would suggest, however simple it may be, being explicit about the method being used.
2. The errors associated with the slopes are also too small – especially the ones in Figure 6, where the standard error is merely 0.1%. If those are standard errors of the parameter in the linear model, they may be correct in a sense, but are not helpful. "

We agree that the fitting method should have been stated explicitly. We have added a clarifying sentence at the end of Sect. 2 (L106):

> Throughout, power-law exponents are computed using ordinary least-squares linear regressions to logarithmically transformed variables, and all reported uncertainties represent twice the standard error of the fitted slope (i.e. a 95% confidence interval).

Regarding the standard errors being "too small": we have updated the perimeter-area fits to compute a fit not to individual cloud perimeters, but to mean cloud perimeter in each logarithmically-spaced area bin. This is more consistent with the stated methodology in the Fig. 5 caption and should have been done this way from the start. The uncertainties are now representative of the scatter between bins rather than between individual clouds, consistent with the text description.

*"**L192** It is not clear to me why the authors claim that "filled fractal dimension does not display any statistically significant scale dependence", because I can tell that it does, only slightly less significantly. Given the small errors associated with these slopes, I would say that going from 1.36 to 1.5 is only*

marginally less scale dependent than going from 1.41 to 1.7 when the standard errors for the parameters of the linear regression are less than 1%, as shown in Figure 5. This also shows that the parameter-area relationship is not linear, which is what I mentioned in my last review; I suspect that all the linear relationships assumed in this paper are only marginally linear, and I suggest that the authors look at the plot of the residuals against the linear fit to prove otherwise (not just Figure 5, but for other linear fits as well). "

We agree that describing the filled fractal dimension as having "no statistically significant scale dependence" overstated the case. We have changed L193 to read:

> By contrast, the filled fractal dimension ~~does not display any statistically significant scale dependence.~~displays substantially less scale dependence, ranging only from 1.36 ± 0.01 to 1.55 ± 0.13.

The point of Fig. 5 is that filling cloud holes substantially reduces scale dependence, not that it eliminates it entirely.
Regarding residual plots: for the perimeter-area relationship, Fig. 5 already quantifies the departure from a single power law more directly than residuals would, by computing separate slopes across four decades of cloud area. For the other fits, we note that interpreting residual plots is itself subjective – whether a pattern in the residuals constitutes a meaningful departure from linearity or simply reflects noise is a matter of judgment. We do not think such plots would resolve any open question in the paper.

*"**L269** One could quantify the effect of ignoring the clouds extending beyond the domain by taking a (slightly) smaller sub-domain and calculating the correlation integral with and without those clouds. I believe it would be a useful exercise to confirm the claims from DeWitt and Garret (2024) about biases due to truncated clouds. "*

Line 269 in the prior manuscript discusses the correlation dimension, not the size distribution as DeWitt and Garrett (2024) considered, so that prior study is not relevant here and we do not suggest that it is. For the correlation dimension, we do not ignore truncated clouds.

*"**L280** I do admit that my wording was pretty confusing in the last review, so let me rephrase the question: in calculating the box dimension, I have found that (albeit on a slightly different scale) the number of boxes N is highly dependent on the choice of the box size and the location of the boxes (i.e. Where the boxes lie on the cloud perimeter). The uncertainty in the linear fit was more dependent on these choices than the internal variability of the data. There is no discussion on how the box dimension was obtained, which was what I meant to ask in the last review. "*

Section 4.2, specifically L245-L255 contains a description of how the box dimension was obtained. There is also Fig. 8 and its caption which provides a visual explanation of how the box dimension is calculated. The reviewer seems to imply that the fact that "the number of boxes N is highly dependent on the choice of the box size" implies some kind of subjectivity, but this is exactly the relationship that defines the box dimension in the first place (Eqn. 8).

*"**L282** From the response, the box dimension shown in Figure 10 is supposedly close to 1 for the "smallest two points". I'd like to see the actual fit; given the very small standard errors on the slopes, I expect the slopes will be very close to 1. "*

The table below shows D_{box} calculated using only the two and three smallest box sizes. With only two points there is no uncertainty on the slope; with three points the values are near 1 but with uncertainties of order ± 0.2, offering little useful information. This is expected: at small box sizes D_{box} approaches the Euclidean dimension of a single pixel, which is precisely why these points are excluded from the fit, as is standard practice for box-counting methods.

R	D_{box} (2 smallest)	D_{box} (3 smallest)
0.10	0.84	1.07 ± 0.26
0.15	0.81	1.04 ± 0.26
0.20	0.80	1.02 ± 0.25
0.25	0.79	1.00 ± 0.24
0.30	0.79	0.99 ± 0.23
0.35	0.78	0.97 ± 0.22

*"**L298** Based on the response from the authors and previous works, filtering out truncated clouds end up ignoring a large portion of the clouds.*
1. The authors acknowledge that a significant number of clouds will be filtered out; I would like to see it mentioned in the actual paper. And no, because a different set of data is being considered for a different analysis, I think it is necessary to be explicit about the effect of manually truncating the histogram.
2. I understand that the 50% filter was recommended by the other study, but I find it no better than the "subjective choice" mentioned in that paper. Why not ignore 100% clouds that are truncated? Isn't this virtually the same as introducing a subjective cut-off to the cloud size distribution, to perform a "simple linear fit" to a distribution that is only marginally linear?
3. Even with filtering, it does appear that there is a bit of a departure from linearity in Figure 12. What do the residuals look like, especially when the offsets are removed? Is there an underlying trend? "

We have added a note to the manuscript stating the number of clouds removed, as requested (L303):

> This method ensures that β is calculated from only the unbiased portion of the distribution, i.e. the portion that is not dominated by large clouds that extend beyond the measurement domain. This filtering removes approximately 0.06% to 0.4% of all clouds depending on threshold.

The analysis done by DeWitt and Garrett 2024 already extensively investigated this exact problem, including an analysis of a cloud dataset similar to ours and also a very different dataset of percolation lattices. In that paper we considered including truncated clouds, excluding them, and which filtering threshold to use, eventually determining 50% was adequate as convergence in the measured exponents was achieved near that threshold. This full analysis went far beyond a simple "subjective choice" as the bulk of the paper was spent providing experimental evidence for the suggested methodology.
Given that extensive past analysis, we believe citing and building upon past work is adequate here.
Regarding residual plots for Figure 12, we refer to our response to the L192 comment above: interpreting residual plots is itself subjective, and we do not think they would resolve any open question.

*"**L317** I think one should compare D_e and D_{box} at the same threshold. So I am not sure if "all three methods [...] point to a value of $D_e \approx 1.7$". The authors are being intentionally ambiguous here, as the only way to get a value of 1.7 is by averaging all fractal dimensions at all thresholds. I think this has to be made clear, not that somehow the fractal dimensions converge to a single value, as they seemingly do not. I would say that it would be more interesting if the fractal dimensions vary across the smaller sub- domains or over time, for example, which would better align with the suggested goal of this work. "*

We do compare D_e and D_{box} at the same threshold, in Table 1.
It is clear from the line quoted by the reviewer (L317 in prior version, 319 in revised manuscript) that we are not attempting to claim "the fractal dimensions converge to a single value." That line instead states "Although D_c and D_{box} decrease with increasing threshold, ...".
Also consider L311: "Generally, the fractal dimension decreases with threshold R ..."
We have made this change at L319 to make this more clear:

> As shown in Table 1, all three methods for calculating the ensemble fractal dimension – the box dimension, correlation dimension, and βD_i – point to a value ~~of $D_e \approx 1.7$~~roughly between 1.6 and 1.8, with a possible exception of the box dimension at higher reflectivity thresholds. Although D_c and D_{box}

decrease with increasing threshold, the product βD_i displays less of a trend due to the relatively weak dependence of both β and D_i on R. More importantly, whichever method is used to calculate D_e, the values are substantially higher than those for the individual fractal dimension D_i, which has a value of approximately 1.4 (Fig. 6).

References

Batista-Tomás, A., Díaz, O., Batista-Leyva, A., and Altshuler, E.: Classification and dynamics of tropical clouds by their fractal dimension, Quarterly Journal of the Royal Meteorological Society, 142, 983–988, 2016.

Christensen, H. M. and Driver, O. G. A.: The Fractal Nature of Clouds in Global Storm-Resolving Models, Geophysical Research Letters, 48, e2021GL095 746, https://doi.org/https://doi.org/10.1029/2021GL095746, 2021.

DeWitt, T. D. and Garrett, T. J.: Finite domains cause bias in measured and modeled distributions of cloud sizes, Atmospheric Chemistry and Physics, 24, 8457–8472, https://doi.org/10.5194/acp-24-8457-2024, 2024.

Janssens, M., Vilà-Guerau de Arellano, J., Scheffer, M., Antonissen, C., Siebesma, A. P., and Glassmeier, F.: Cloud Patterns in the Trades Have Four Interpretable Dimensions, Geophysical Research Letters, 48, e2020GL091 001, https://doi.org/https://doi.org/10.1029/2020GL091001, 2021.

Raghunathan, G. N., Blossey, P., Boeing, S., Denby, L., Ghazayel, S., Heus, T., Kazil, J., and Neggers, R.: Flower-Type Organized Trade-Wind Cumulus: A Multi-Day Lagrangian Large Eddy Simulation Intercomparison Study, Journal of Advances in Modeling Earth Systems, 17, e2024MS004 864, https://doi.org/https://doi.org/10.1029/2024MS004864, e2024MS004864 2024MS004864, 2025.

Siebesma, A. P. and Jonker, H. J. J.: Anomalous Scaling of Cumulus Cloud Boundaries, Phys. Rev. Lett., 85, 214–217, https://doi.org/10.1103/PhysRevLett.85.214, 2000.