

Reply to editor and reviewers

Dear editor and reviewers,

We sincerely thank you for your valuable and insightful comments, which have greatly helped improve the clarity and quality of our manuscript. In the response, we have carefully revised the relevant sections. We believe these revisions have strengthened the manuscript. Below, we provide detailed responses to each of the reviewers' comments.

#RC1

The paper investigates the microphysical properties of global precipitation systems (PSs) using GPM Dual-frequency Precipitation Radar (DPR) data from 2018–2022. An objective k-means clustering approach (with PCA for dimensionality reduction) is applied to classify precipitation systems into eight distinct types:

- Non-extreme PSs: high-latitude shallow, subtropical shallow, moderate, deep.
- Extreme PSs: extreme deep, strong, extreme strong, marine extreme.

Key findings:

- Continental PSs generally have larger mean drop diameters (D_m) than oceanic PSs, while oceanic PSs have higher normalized intercept parameters (N_w).
- Extreme PSs show balanced raindrop breakup and coalescence, while shallow PSs are dominated by coalescence.
- Clear land–ocean contrasts and latitudinal variations are found in microphysical structures.
- Diurnal and seasonal cycles differ by PS type, with continental systems peaking in the afternoon/summer, while shallow oceanic PSs peak at night.

The manuscript is generally well written and presents novel results. However, it would benefit from a few clarifications and structural improvements. I therefore recommend a minor revision addressing the following points:

Reply: We sincerely thank for the positive evaluation and comments. Accordingly, we have revised the manuscript to improve the logical flow of sections, clarified key methodological details, and ensured that all figures and tables are clearly labeled and referenced.

- The methodology and results section would benefit from a clearer and more logical narrative. At present, the text shifts between clustering, physical interpretation, and microphysical discussion in a way that can confuse the reader. A more transparent structure would be to explicitly present the workflow as follows:

1. Application of clustering algorithm

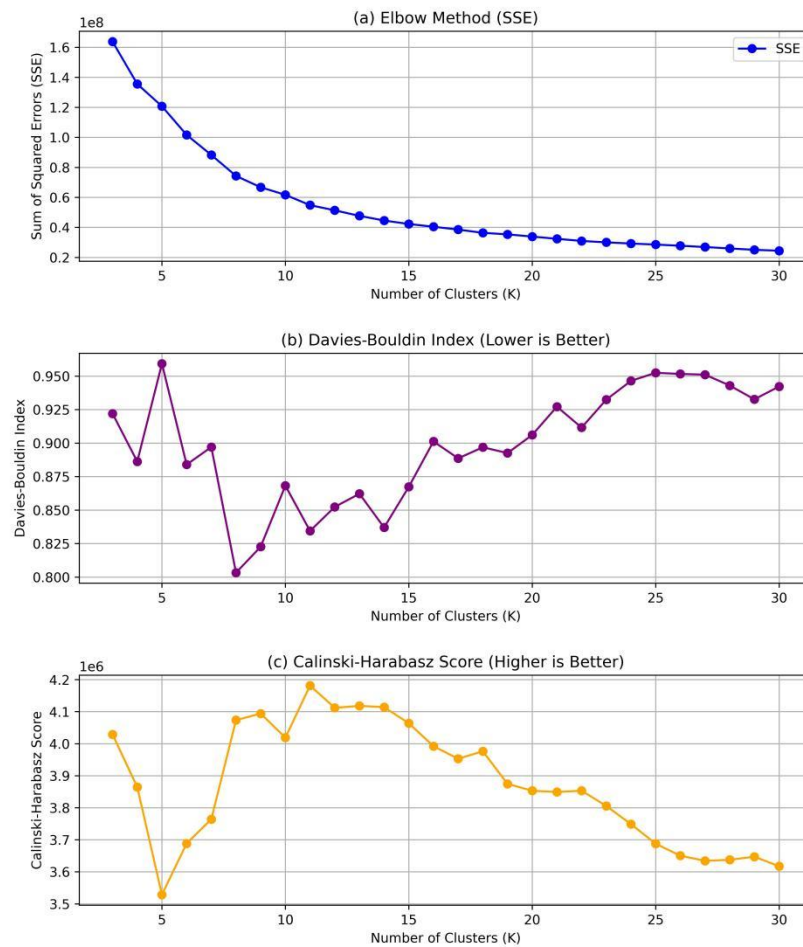
State clearly that the clustering was applied to the precipitation feature (PF) database, which contains ~9 million PFs identified from GPM DPR data. PF input variables include: precipitation rate, radar echo top heights, drop size distribution (D_m , N_w), convective/stratiform fractions, and spatial metrics.

Reply: We revised the sections 2.3 to include the sample counts and input features. It is worth noting that the parameters are not from the Precipitation Feature (PF) database of Liu et al. (2008), because the PF dataset does not include the profiles of D_m and N_w with 176 levels. Therefore, we reproduced the PS dataset to better derive the required parameters. We referred the new dataset as PS rather than PF to distinguish it from the original PF dataset.

2. Selection of optimal number of clusters

Explain that the Davis–Bouldin index and Calinski–Harabasz score were evaluated for different values of k , and the minimum DB index at $k=8$ was taken as evidence that eight classes provided the best compromise between compactness and separation. Clarify if the Elbow Method applied to the Within-Cluster Sum of Squares suggests the same number of classes. This will justify the choice of $k=8$.

Reply: Thank you for this valuable comment. We have also applied the Elbow Method to evaluate the Within-Cluster Sum of Squares (SSE). However, as shown in Reply-Figure 1, the SSE decreases continuously with increasing numbers of clusters, without exhibiting a clear inflection point. Compared with the Davis – Bouldin index (DBI) and Calinski – Harabasz score (CHS), the Elbow Method provides less conclusive guidance for determining the optimal number of clusters in our case. Therefore, we relied primarily on the DBI and CHS, both of which consistently indicated that $k=8$ offers the best compromise between compactness and separation.



Reply-Figure 1. Changes in SSE error, DB index and CH score with the number of clusters from 2 to 20.

3. Characterization of each cluster

Emphasize that each cluster is then characterized by its distinctive features, such as mean Dm or cloud top height... Based on these distinguishing characteristics, the clusters are named descriptively (e.g., “shallow,” “deep,” “extreme strong,” “marine extreme”). This step should be made explicit, because the current version sometimes reads as if the naming were imposed rather than derived.

Reply: We thank the reviewer for this very helpful suggestion. In the revised manuscript, we have moved the description of the naming of the eight clusters to Section 3.1 and clarified the basis for assigning these names, improving the readability and flow of the text.

4. Emergent spatial and temporal patterns

Only after the clusters are defined should the manuscript show that these objectively derived groups exhibit coherent spatial distributions (e.g., shallow clusters dominating high latitudes, extreme clusters in the tropics). This is an important and interesting result: the clustering, based purely on precipitation properties, also reflects geophysical organisation in space, suggesting that the classification captures physically meaningful regimes.

Reply: Thank you for your valuable comment. We have revised the manuscript and provided a more detailed explanation in Section 3.1. In this section, we clarified both the naming rules of the clusters and their spatial distributions.

- The methodology should also clarify how the stratiform, convective, land, and marine percentages are computed. Are these:

- (a) computed for each PF individually and then averaged across all PFs in a cluster, or
- (b) computed directly from the total number of pixels across all PFs in a cluster?

Reply: In this study, the percentages of stratiform, convective, land, and marine pixels were first calculated individually for each PS and then averaged across all PSs within a cluster, rather than computed from the total number of pixels across all PSs. We have clarified this point in the corresponding section of the manuscript.

- Extreme events context: The discussion of extreme precipitation systems should be better grounded in previous literature, particularly Zipser et al. (2006), Ni et al. (2017), and Bang and Cecil (2021), which provide benchmarks for extreme convective systems observed by satellites.

Reply: Thank you for this valuable suggestion. We have revised Section 3.1 to strengthen the discussion of extreme convective systems by incorporating relevant findings from previous literature.

- Logical inconsistency (line 287): The argument is circular. Earlier, weak updrafts were inferred from low 40 dBZ echo heights. Later, the rapid decrease in reflectivity with height is attributed again to weak updrafts. This leads to a logical loop and should be clarified.

Reply: Thank you for pointing it. We have clarified the description to avoid circular reasoning. The observations of lower 40 dBZ echo tops and the rapid decrease in reflectivity are now presented as indicators of weaker updrafts, rather than using one to justify the other.

- Balanced processes would result in no change in the reflectivity and D_m .

Reply: Thank you for this helpful comment. We have added this clarification in the revised manuscript, noting that the balanced processes would result in no change in reflectivity and D_m .

- Phase-change influence (line 303): The observed change in D_m likely corresponds to changes in precipitation phase across the melting layer. Please reference Mroz et al. (2024). Additional CFADs of Z , D_m , and N_w as a function of temperature or height

relative to the freezing level would strengthen the analysis and could be added as supplementary material.

Reply: Thank you for raising this valuable point. We fully agree that the rapid changes correspond to changes in precipitation phase across the melting layer because of the changes of effective dielectric constant, particle size etc. We added discussion about this point in the section 3.3 and conclusion section. Nevertheless, as we did not include the ERA5 temperature profiles in the Precipitation System dataset, it would take too much time to redownload the DPR dataset and reproduce the dataset. Therefore, we did not plot the CFAD in the temperature coordinate. We will carefully address this issue in future work.

- Algorithm-induced correlations (section 3.4): The observed correlations between D_m and N_w may be artifacts of retrieval assumptions, since the GPM algorithm enforces a correlation between D_m and precipitation rate (see Chase et al., 2020). This must be acknowledged explicitly.

Reply: Thank you for your valuable comment. We have acknowledged in the revised manuscript (Section 3.4) that the observed correlations between D_m and N_w may be partly influenced by algorithm-induced assumptions.

- Extreme precipitation rates: DPR is not well-suited for quantifying extreme rain rates because Ku/Ka frequencies are strongly affected by attenuation and multiple scattering in heavy precipitation. Values above $\sim 100 \text{ mm h}^{-1}$ should be treated with caution and interpreted in light of retrieval limitations (see Battaglia et al.).

Reply: We have addressed this point in the Conclusion section of the manuscript. Since the study does not further discuss precipitation rates beyond the quantifications presented in Table 1, we note the limitations of DPR in measuring extreme rain rates, particularly above $\sim 100 \text{ mm h}^{-1}$, due to attenuation and multiple scattering.

- Language and Clarity Issues: Many sentences exceed 40 words; shortening them would improve readability. Descriptions of Figs. 7–8 are overly detailed in-text.

Reply: We have revised the manuscript to improve language clarity by shortening long sentences and reducing overly detailed descriptions in Figs. 7–8. Specifically, we have shortened sentences exceeding 40 words and reduced the in-text descriptions of Figs. 7–8.

#RC 2

Major comments:

1. The stated purpose of the analysis is to evaluate distributions of D_m and N_w within precipitation systems that are classified globally using machine learning. A significant flaw in the methodology is that both of these parameters are included as features in the classification algorithms. Thus, the classification of the types is partly (with unknown, possibly varying weight) determined by the parameters that are to be evaluated. The parameters differ amongst the classifications because it was predetermined to be so.

Reply: Thank you for this valuable comment. As mentioned in the last paragraph of the introduction section, this study includes two purposes. The first is to objectively classify

global PSs and the second is to analyze the microphysical characteristics of different PSs. In literature, most papers used subjective methods to recognize various precipitation types, such as use storm height to define deep convection, use precipitation rate to define extreme precipitation etc. Using single threshold is limited because the precipitation systems are quite complex. Therefore, we choose to classify the PS with as much parameters as possible and the k-means clustering was conducted using a combination of multiple precipitation characteristics, including microphysical parameters (D_m and N_w), structural parameters (H_{top} , MAXHT20/30/40), and macroscopic properties (precipitation area, near-surface precipitation rate, and the convective/stratiform fraction) etc. The subsequent analysis of D_m and N_w is intended to characterize the overall microphysical features of the already-defined PS types, rather than to provide an independent statistical test of $D_m - N_w$ differences among classes.

This approach is popular used in literature. For example, Ryu et al. (2021) used the D_m and N_w to objectively classify three types of heavy rainfall and then discussed the microphysical properties of three types of heavy rainfall. Overall, as D_m and N_w are key descriptors of precipitation microphysics, they are naturally included in both the clustering framework and the physical interpretation of the resulting PS types.

2. To be clear, I am not a radar meteorologist. Generally, though, I believe radar retrievals of microphysical properties, e.g., D_m , would require attention to the fact that the coefficients in the DeltaZe power law relations are not necessarily constant. This would be particularly true in cold clouds where ice particles exhibit various sizes, yes, but also vary in density, habit, and orientation. Cold clouds are definitely sampled here. The methods used to calculate the microphysical parameters presented is largely omitted from this manuscript, as is any justification or explanation of their global applicability. How do you know the differences presented are not artifacts of the retrieval and what are the limitations of the global product?

Reply: We thank the reviewer for the insightful comment. We would like to clarify that the microphysical parameters used in this study, including D_m and Z_e , are from the GPM DPR satellite products rather than calculated through our own algorithms. The focus of this study is on applying these existing products to classify and analyze global precipitation systems, rather than developing or validating retrieval algorithms. Besides, amounts of validation work of the GPM DPR products have been widely conducted globally (Adirosi et al., 2021; Chase et al., 2020; Gatlin et al., 2020; Huang et al., 2021; Peinó et al., 2024; Seela et al., 2024). These researches have confirmed that research using these products are reliable.

Similar works using the DPR products are commonly used in the literature; for example, Li et al. (2024) and Wen et al. (2023) also analyzed microphysical parameters from DPR or other satellite products without providing in-depth explanations of the retrieval methodology. In response to the reviewer's concern, we have revised the data description section to explicitly reference the relevant algorithm documentation and literature that describe the retrieval of these parameters in the GPM DPR products.

3. The application of the unsupervised classification is overly complicated. First, there is the PCA. The only reason given to use this is to reduce dimensionality for feature selection supplied to the k-means algorithm. For cloud research, it would be more conventional and straight-forward to reduce the height dimension using an integrating variable, like total water path. There might be some particular reason for using PCs, but none is actually given.

Then, once the clusters are identified, they are (subjectively?) named in such a way that implies they could easily be distinguished with a simple threshold. So why not just apply a geographical and cloud top height threshold to isolate a series of recognizable systems rather than expecting an unsupervised classification to reproduce them?

You might even lose something with the unsupervised approach if cloud types you would want to contrast are actually quite similar. An example of this that appears in the paper are land vs marine systems, which are subset after the fact because the clustering (largely) did not differentiate them.

Reply: We appreciate the thoughtful comments. First, the application of PCA is necessary in this study. The PSs are characterized by multiple variables with vertical structures, and directly including all height-resolved profiles would lead to a very high-dimensional feature space, which would substantially reduce the robustness and efficiency of the k-means algorithm. PCA provides an objective way to retain the dominant modes of vertical variability while filtering out redundant information. Unlike vertically integrated variables (e.g., total water path), which aggregate information and may obscure vertical contrasts, PCA preserves key vertical structure signals that are physically relevant to precipitation processes. Meanwhile, total water path is physically meaningful. In contrast, vertically integrated $Z_e/D_m/N_w$ are rarely used in literature because it is difficult to explain their physical meanings.

Second, the reviewer suggests that simple geographical or cloud-top-height thresholds could be used to isolate recognizable precipitation systems. However, no universally accepted global classification scheme for precipitation systems based on fixed thresholds currently exists. Precious research usually focused on specific precipitation types with one or two precipitation features. However, precipitation characteristics are influenced by multiple interacting factors, including environmental moisture, thermodynamic structure, circulation regime, and topography, all of which vary substantially across regions. A threshold-based approach would therefore introduce strong subjectivity and may fail to capture transitional or mixed regimes that do not conform to predefined criteria.

In contrast, the unsupervised clustering approach integrates multiple precipitation features and objectively obtains various types of precipitation systems. The naming of clusters is performed after clustering, based on their emergent physical characteristics, which does not influence the clustering itself. Meanwhile, the clusters are named with reference to previous literature. While some clusters can indeed be broadly interpreted using simple descriptors (e.g., “shallow” or “deep”), this does not imply that such PSs could be

classified using a single threshold.

We are not sure what the reviewer specifically means by the “*You might even lose something with the unsupervised approach if cloud types you would want to contrast are actually quite similar.*”. We would like to clarify that the unsupervised clustering approach in our study does indeed distinguish between continental and marine precipitation systems based on the combined precipitation-related features. As we have known, the discrepancies of precipitation between land and ocean is the one of the most significant characteristics of global precipitation distributions. For example, the hailstorms rarely occur over ocean because the middle level of marine convection tend to be significantly weaker than those over land. The input features of clustering algorithm do not include land-ocean information. However, it separates continental and marine precipitation system well, which proves that the clustering algorithm is reliable.

Minor comments:

L54: define terms

Reply: Defined.

L102: So many acronyms. Are all these PR PS MS NS HS FS SSE DB CH CFAD etc really needed? Some like FS are never even defined. Some are defined in the abstract but not elsewhere. Some are defined multiple times and may not be necessary (like PS). Some are defined and hardly used (like CH).

Reply: We appreciate this valuable comment. In the revised manuscript, we have carefully reviewed all acronyms and simplified their usage. For acronyms that are rarely used, we have spelled them out in full spell throughout the manuscript.

L115-116: “those contained the widely” revise.

Reply: Revised.

L135: It’s unclear if PSs are individual synoptic systems, cloud systems, or just samples that may have included multiple samples of a given cloud.

Reply: This definition is provided in the first paragraph of Section 2.2. Following the reviewer’s comment, we have further clarified the definition of PS in the revised manuscript to avoid confusion regarding the physical meaning of PS.

Sect. 2.3: More details on how PCA was used to augment k-means would be helpful. All parameters or only Z_e were put through PCA? How many PCs? How was time-height handled?

Reply: We have revised the section 2.3 to clearly clarify the use of PCA. PCA was applied to the vertical profiles of three parameters: Z_e , D_m , and N_w . The purpose of PCA was to reduce the vertical dimension, compressing the original 176 height levels into a single representative component for each variable. Only one principal component was retained for each profile, as it explains the dominant variance of the

vertical structure and serves as a compact descriptor for clustering. The PCA was applied once, independently for each parameter, and no temporal dimension was involved in the PCA procedure. Time was not treated as an input dimension; only the vertical (height) dimension was reduced.

L140-141: some repetition in these statements

Reply: Revised.

L185-187: Is “extreme” really the right word for the four classes that are described as such? Those clusters are infrequent, but that is not the same as extreme, even if some may have some extreme characteristic amongst or within features. This way of isolating extreme isn’t as statistically tractable as analyses of distributions of the cloud water parameters. So it seems like we need to understand extreme in this context means, and why we should be interested in it.

Reply: The term “extreme” is widely used in the atmospheric science community. Generally, an “extreme” event is defined based on its rarity. However, the percentile thresholds used to define extremes vary depending on the context. In the research of extreme precipitation, thresholds such as 1%, 0.1% or even 0.01% percentile are commonly employed. For example, Hamada et al. (2015) selected the uppermost 0.1% of Precipitation Features (PF) to define extreme rainfall events and extreme convective events. In this study, the four types of extreme PS account for 0.39%, 0.22%, 0.023%, and 0.0105% of all PSs. Therefore, it is reasonable to define these low-occurrence PS types as extreme PS.

L212: “where is dominated” a word is missing or something here

Reply: Corrected.

There are lots of unsourced and highly generalized statements (e.g., L253-254, 265-266)

Reply: The statements at L253–254 have been removed in the revised manuscript. The statements at L265–266 are now properly supported with relevant references.

L258: extra space

Reply: Deleted.

Reference

Adirosi, E., Montopoli, M., Bracci, A., Porcù, F., Capozzi, V., Annella, C., Budillon, G., Bucchignani, E., Zollo, A. L., Cazzuli, O., Camisani, G., Bechini, R., Cremonini, R., Antonini, A., Ortolani, A., and Baldini, L.: Validation of GPM rainfall and drop size distribution products through disdrometers in Italy, *Remote Sensing*, 13, 2081, <https://doi.org/10.3390/rs13112081>, 2021.

Chase, R. J., Nesbitt, S. W., and McFarquhar, G. M.: Evaluation of the microphysical assumptions

within GPM-DPR using ground-based observations of rain and snow, *Atmosphere*, 11, 619, <https://doi.org/10.3390/atmos11060619>, 2020.

Gatlin, P. N., Petersen, W. A., Pippitt, J. L., Berendes, T. A., Wolff, D. B., and Tokay, A.: The GPM validation network and evaluation of satellite-based retrievals of the rain drop size distribution, *Atmosphere*, 11, 1010, <https://doi.org/10.3390/atmos11091010>, 2020.

Hamada, A., Takayabu, Y. N., Liu, C., and Zipser, E. J.: Weak linkage between the heaviest rainfall and tallest storms, *Nat Commun*, 6, 6213, <https://doi.org/10.1038/ncomms7213>, 2015.

Huang, H., Zhao, K., Fu, P., Chen, H., Chen, G., and Zhang, Y.: Validation of precipitation measurements from the Dual-frequency Precipitation Radar Onboard the GPM Core Observatory using a polarimetric radar in South China, *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–16, <https://doi.org/10.1109/TGRS.2021.3118601>, 2021.

Li, D., Qi, Y., and Li, H.: Statistical characteristics of convective and stratiform precipitation during the rainy season over South China based on GPM-DPR observations, *Atmospheric Research*, 301, 107267, <https://doi.org/10.1016/j.atmosres.2024.107267>, 2024.

Liu, C., Zipser, E. J., Cecil, D. J., Nesbitt, S. W., and Sherwood, S.: A cloud and precipitation feature database from nine years of TRMM observations, *Journal of Applied Meteorology and Climatology*, 47, 2712–2728, <https://doi.org/10.1175/2008JAMC1890.1>, 2008.

Peinó, E., Bech, J., Polls, F., Udina, M., Petracca, M., Adirosi, E., Gonzalez, S., and Boudevillain, B.: Validation of GPM DPR Rainfall and Drop Size Distributions Using Disdrometer Observations in the Western Mediterranean, *Remote Sensing*, 16, 2594, <https://doi.org/10.3390/rs16142594>, 2024.

Ryu, J., Song, H.-J., Sohn, B.-J., and Liu, C.: Global distribution of three types of drop size distribution representing heavy rainfall from GPM/DPR measurements, *Geophysical Research Letters*, 48, e2020GL090871, <https://doi.org/10.1029/2020GL090871>, 2021.

Seela, B. K., Janapati, J., Lin, P.-L., Chen-Hau, L., and Huang, M.-Q.: Evaluation of GPM DPR Rain Parameters with North Taiwan Disdrometers, *Journal of Hydrometeorology*, 25, 47–64, <https://doi.org/10.1175/JHM-D-23-0027.1>, 2024.

Wen, L., Chen, G., Yang, C., Zhang, H., and Fu, Z.: Seasonal variations in precipitation microphysics over East China based on GPM DPR observations, *Atmospheric Research*, 293, 106933, <https://doi.org/10.1016/j.atmosres.2023.106933>, 2023.