Review of "Regime-based Aerosol-Cloud Interactions from CALIPSO-MODIS and the Energy Exascale Earth System Model version 2 (E3SMv2) over the Eastern North Atlantic"

Summary

This work studies aerosol-cloud interactions in the Eastern North Atlantic region, comparing satellite retrievals to a GCM model output, for a large set of days, to elucidate the relationship between micro- and macrophysical cloud properties: cloud droplet number concentration, liquid water path, and boundary-layer extinction coefficient. They analyze the relationships between these variables seasonally, comparing simulations to satellite data, and the main novelty of the study is that they also analyze the behavior for 4 meteorological regimes that are found using clustering techniques on ERA5 reanalysis data. This regime clustering gives new insights by separating natural covariability and clarifying one of the relationships. The paper is well written, and the discussion is very detailed and provides a full understanding of the studied system and its physical processes. I mostly have minor comments regarding some methods, and about how to better summarize and provide ideas to modelers based on their discussion.

Minor comments

- I suggest highlighting the novelties of the paper in the abstract, introduction, and summary. In particular, I am not sure if the novelty is only the analysis based on regimes, or if the seasonal analysis is also novel? Or is this particular model and satellite product comparison new?

- The discussion in every Section is very thorough, but many hypotheses point to modeling biases or ideas for model improvements, which are not the main scientific contribution of this work. It would be nice to assess if these hypotheses are true by confirming some diagnostics on the resulting model parameters. Another thing that could be done is to order these recommendations and try to assess which model improvements are more likely or feasible.

Line by line comments

- L17 Clustering was performed on satellite or simulation data? Or both?

- L18 Maybe explain the 4 regimes before they start appearing

- L161 Are there comparisons for other cloud types?

- L195 What is the value of that coarse vertical resolution?

- L213 Time formatting: Should it be 1 p.m or 13:00?

- L214 Was the date also a variable?

- L226 So the DEC was used after optimizing the k-means clustering? Or was it also tested for different k values?

- L246 "followed by fall", "lowest during winter"?

- L278 Is this index computed from the data? Is it a fit with confidence interval?

- L355-359 This sentence is a bit confusing

- Fig. 5: Composites mean that these are based on the mean values of each cluster? Or are these the centroids?

- L401 Details were already given in the previous Section

- L407 Is there a reason why the regime order does not follow the expected trough-ridge transition?

- L426 I think it is important to report the number of events and percentage for each regime in the main manuscript, for statistical significance. Now that I see the supplementary information, maybe it is worth cautioning the readers that regime 4 had the lowest amount of information

- L442 "are listed"

- Fig. 9: The median sigma values were selected for each regime or for the entire dataset?

- L720 I suggest mentioning the four regimes