Atmos. Chem. Phys. Discuss., 10, C180–C185, 2010
www.atmos-chem-phys-discuss.net/10/C180/2010/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Cluster analysis of midlatitude oceanic cloud regimes – Part 1: Mean cloud and meteorological properties" *by* N. D. Gordon and J. R. Norris

Anonymous Referee #1

Received and published: 19 February 2010

Cluster analysis of midlatitude oceanic cloud regimes – Part 1: Mean cloud and meteorological properties

General Comments

Having carefully read this manuscript and thought about it for some time, and also reading the companion paper, I feel that it is an interesting contribution to our picture of midlatitude clouds over oceans. There are some issues, primarily in the presentation, that should be addressed (see Specific Comments), but in general I found the methods and analysis to be interesting and reasonable, though not always totally convincing. Additional revision of the text (and maybe figures) could strengthen the arguments and clear up some confusion.

Specific Comments

These comments are divided basically into the methodological parts of the paper versus the analysis that follows the clustering.

First, I think there are a couple of open questions about the data and methods used here. The data being used makes sense, but some choices should be supported. The most obvious example is the choice to use NCEP reanalysis rather than ERA-40 and/or ERA-Interim. It would be good to include a statement justifying (rationalizing?) this decision, especially since the differences between NCEP and ERA-40 are well-known (e.g., Trenberth et al 2005 10.1007/s00382-005-0017-4, Bromwich et al 2007 doi:10.1029/2006JD007859) and could influence the "meteorology" in the clusters. Second, it would be helpful to remind readers about possible deficiencies of the Zhang et al. 2004 radiative fluxes; how do they compare with other estimates (Trenberth et al 2009 DOI:10.1175/2008BAMS2634.1 report potentially large errors in the surface fluxes)? Neither of these explanations need to be very detailed, but it would help ease the skeptical reader's mind. On page 1564, lines 9-13, the authors assert that the reason for using midlatitudes because there is synoptic forcing of clouds there and because the vertical motion in the reanalysis is best constrained. Is the statement that the the vertical motion is observationally constrained more in midlatitudes correct? Especially over the oceans (and in winter)? This may be the case, but it isn't clearly true, so if there is a simple explanation or reference it would be helpful.

There are also a couple of questions that remain about the clustering analysis. One basic question about this application of the clustering algorithm is whether using the CTP-tau histograms as "nine-type arrays" assumes that the 9 bins are independent? Since the clusters are based on the Euclidean distance of each new point from the clusters, it seems like correlation between bins would constrain the resultant clusters (e.g., you'd always get high-cloud and low-cloud clusters if the high bins and low bins are anticorrelated). So I guess the questions are, (1) is there much correlation among bins, and (2) does it matter? Also pertaining to the clustering are the issues of the two obvious subjective choices involved: the number of clusters and determining the convergence. Regarding the number of clusters, the authors are clear that this is a subjective choice, and essentially say that seven clusters represents the sweet spot for this analysis. This is fine, but then the second choice comes into play, which is deciding which 7-cluster set to use. Page 1568, lines 10-12, says that the cluster set with the smallest total variance was used, which makes sense, but in the rest of that paragraph the authors tell us that the other cluster sets trade a frontal cluster for a low or middle level cloud cluster. While there is the assurance that it doesn't matter, since these other clusters look like intermediate cloud types, it is a little hard to reconcile this with the choice of 7 clusters rather than, say, 6. Considering the similarity of the two frontal clusters, one could easily imagine losing one of these, with the "weak frontal" cluster being an intermediate type of cloud between the "Deep Altostratus" and "Strong Frontal" clusters. This may not be the case, but this should be made more clear in the text. Is the minimum total variance criteria a good one? Is there a substantial difference in total variance? How would this all play out if 6 clusters were chosen? This hopefully isn't a matter of re-doing the analysis with 6 clusters, but rather a suggestion to be more clear about the consequences of these two choices and showing that the final decision makes sense compared to the alternatives.

There are also a few issues that could be more clearly addressed in the analysis. The tables show mean values, which is useful to get a sense for what these clouds are like. The figures that show profiles only show the anomalies, which are also useful, especially for the authors' purposes, but I did find myself wishing for the mean profiles as well. This is especially true for the clusters that show near-average behavior, like the most common cluster ("small cumulus"). In a similar vein, the authors put a lot of weight on the "distinct large-scale dynamical conditions" found within each cluster, but there's a degree of doubt about how distinct the conditions really are. It would be nice to see some indication of whether the composite profiles are really distinct from

ACPD

10, C180–C185, 2010

each other; for example, error bars could be added at each pressure level showing the standard deviation within each cluster. This is, in my estimation, the one major point that needs to be addressed. If the meteorological conditions within the clusters are not substantially different from each other, then they do not warrant the names assigned to them, and the conclusions of the paper are substantially weakened. On the other hand, if the meteorological conditions are truly distinct, then this really strengthens the argument for cluster analysis being an important diagnostic tool.

The composite spatial distributions in Figure 7 are not very well discussed in the text. The patterns listed do seem to be present in the figure. The interpretation makes sense, but the main pattern the figure shows is that dynamically forced clouds (strong and weak frontal and cirrus) occur near the dynamical forcing, while the less forced clouds (cumulus, stratiform) don't occur near the dynamical forcing. The value of this figure is a little lost on me; what happens if the composite is centered on small cumulus?

The spatial maps are interesting to look at. Figure 8 is useful for understanding the naming conventions of the clusters, the relationships between cluster meteorological conditions, and even the composite picture of Figure 7. It's surprising that Figure 8 was left until the last two pages of text. Perhaps more of a "technical comment," but it should be mentioned that the maps appear to have data shown poleward of 50 degrees. There also appear to be some graphical artifacts along this poleward edge of the domain (white blips here and there). It would also be easier to see details if the maps only extended to the poleward edge of the domain; there's no use wasting space on Antarctica and Greenland. Figures 11 and 12 are also interesting, but I think the DJF season also needs to be included, even if the conditions are near average. An alternative might be to show maps of the phase and amplitude of the seasonal cycle using something like an "Evans plot" (e.g., http://www.ncl.ucar.edu/Applications/evans.shtml).

Finally, there's also the question of the domain used, and whether any of the results here are sensitive to the domain. Based on Figure 8, it would seem that moving

ACPD

10, C180–C185, 2010

poleward would reduce the frequency of the low, stratiform cluster that appears to be strongly influenced by the California stratocumulus (and other Sc regions to a lesser extent), and also strongly impact the frequency of the “small cumulus” cluster. Meanwhile, moving equatorward would reduce the occurrence of “deep altostratus.” Would changing the domain also change the number of clusters to use, and would similar clusters be expected to emerge?

Technical Corrections

page 1562, line 23: “to due to” -> “due to”

page 1565, line 17: “CTP-t” -> “CTP-\tau”

page 1566, line 15-16: Readers probably don’t care if it is technical or not, just say “(see Table 1 of Norris, 1998a)”

page 1567, line 23: “greatest absence of” -> “least”

page 1571, line 7: “e.g.” -> “i.e.”

page 1571, line 15: “Fig 8a” is referenced, but the figures panels are not actually labelled. This is repeated several times, so labels should probably be added to the figure panels.

page 1574, line 24: “constraining” -> “keeping” Or some other change of words to make this sentence more clear.

Table 3 & Figures 1-8: The clusters are named in Tables 1 and 2, but not in Table 3 or the figures. It would be easier to keep things straight if the names were applied to the figures.

Figure 1: There are tiny tickmarks on the top and bottom of each panel. Yes, they are tiny, but it would be worth the effort to delete them. Also, why repeat the colorbar 7 times?

ACPD

10, C180–C185, 2010

Interactive Comment

Figures 2-6: All these profile figures are so related to each other, it might be worth combining them into one big figure. It could be 7 columns by 5 rows (or the other way). The individual panels might be smaller, but that could be made up by getting rid of extraneous labels and making the space around each panel a little smaller. The information would still be visible, and comparisons would be easier. It would be a good use of Tufte's "small multiples" concept. Just a suggestion, obviously.

Interactive comment on Atmos. Chem. Phys. Discuss., 10, 1559, 2010.