Response to anonymous Referee #1

Referee comments appear in *italic*.

General Comments

This is an interesting and informative paper that complements existing cloud overlap statistics using higher resolution vertical profiles of cloud and total water content. It is well written and contains information that may be used to improve the treatments of radiation transfer in Global Climate Models (GCMs).

Scientific Comments

Isn't there a fundamental mismatch between the scales used in the paper and those that are appropriate for the radiation calculations in a GCM that uses an atmospheric time step of 30-minutes and calls its radiation code to begin each time step (CM2.1 for example)? Even though the geometric grid spacing of a GCM is 75-300 km, the relevant distance scale for overlap statistics would seem to be much smaller. Assume that the 30-minute temporal resolution for a GCM radiation calculation is coupled with a 10 m/s wind velocity. This results in a distance scale of 18-km and if the wind speed is elevated to 40 m/s the relevant distance scale is 72-km. These scales are smaller than the geometric grid spacing in most GCMs and the corresponding scales computed considered in the paper. It seems to me that the distance scale that should be used to compute the relevant cloud overlap statistics would vary according to the model time step for radiation transfer calculations in a given GCM (and the accompanying wind speed) rather than the geometric grid spacing and be applied over the entirety of the geometric grid to represent cloud overlap. I am interested in the author's opinion on the relationship between the GCM time step and the appropriate time scale for compilation of overlap statistics.

We don't think that a GCM's time step for dynamics or radiation should be related to our choice of 2-hr, 4-hr or 8-hr time scales to represent a specific spatial scale under the frozen turbulence hypothesis we have implicitly adopted. When radiation is called in a GCM at a particular gridcolumn, it is applied to an instantaneous snapshot of the cloud field within the spatial confines of the gridcolumn as it appears at the instance the call to the radiation routine is made. The cloud field at that instance is not the result of within-grid advection between atmospheric time steps; this kind of advection is simply not resolved. Actually, even gridpoint to gridpoint cloud advection is usually not accounted for in GCMs. The cloud field at a particular GCM gridcolumn instead results from the particular dynamical and thermodynamical conditions at the time the stratiform/convective parameterization is called. Obviously, our cloud fields made from a collection of 10 sec columns within 2-hr, 4-hr or 8-hr time intervals are not snapshots. We rather *force* them to have the appearance of snapshots by *assuming* that the cloudy columns advecting over the observing instrument remain unaltered during the time period chosen to define the analysis segment. By assuming that the atmospheric columns passing over the instrument are frozen in time for that period we essentially create a pseudo-instantaneous cloud field analogous to the one appearing within a GCM gridpoint when the cloud routines are called.

Perhaps another way to think of this issue is that a GCM has both unresolved spatial variability (sub-gridscale), and unresolved temporal variability (between time steps). We believe that the former is the more serious shortcoming of GCMs and what we should focus on. Our reasoning is that the spatial scales of cloud and convective events are often only a few km, much less than the typical GCM gridcolumn size, while the temporal scale of cloud development (which scales more with vertical velocity than horizontal velocity) is much closer to (or even larger than) a 30 minute time step. Granted, unresolved temporal variability may be important in some situations but it can probably be addressed by reducing the timestep, while unresolved spatial variability within the gridcolumn poses a more challenging problem.

Now, regarding pure advective effects, there seems to be another timescale, equal to the cloud scale divided by the wind speed. For example, 2 km divided by 10 m/s is 200 seconds or about 3 minutes. So in this sense, we are certainly not resolving the passage of individual cloud elements with 30 minute timesteps. But presumably there are many such cloud elements in the large GCM domain (with respect to cloud scale), so the significance of individual cloud advections is relatively unimportant.

How did you treat precipitation in the MICROBASE profiles? Precipitation may produce a noticeable cloud radar return, but have a smaller impact on the radiation transfer than cloud droplets.

We only use the Microbase cloud liquid and ice water condensates when the quality flags of the dataset indicate problem-free data. Microbase is based on the ARSCL product (Clothiaux et al. 2000) where information about the presence of surface precipitation is included. For those occurences, there is no cloud condensate retrievals by the Microbase algorithm. But it would certainly be naïve to claim that the Microbase retrievals are completely unaffected by the presence of precipitating particles somewhere within the atmospheric column. Precipitation will tend to make cloud alignments appear stronger than they really are (i.e., hydrometeors tend to correlate better than suspended cloud particles alone).

Suggestion for future work: It would be interesting to identify obvious frontal situations and convective situations and compute overlap statistics (conditional sampling). Such an analysis may help in the interpretation of the ensemble results.

This is indeed a good suggestion. There are also other options for compositing alphas and ranks, such as by water phase, whether there is or not cloud in the intervening separation distance, etc. Our work has not exhausted all research avenues and we hope that we'll have the opportunity to conduct more analysis both on the cloud fields themselves, as well as the associated radiation fields.

Technical Comments

Page 599, line 6: typo (condensate)

"condensate" is properly spelled.

Page 600, line 10 prefer "estimate the profiles" rather than "determine the profiles"

Changed as suggested.

What is a "valid cell"? Page 601 line 16: I am assuming that you mean that there could be instances when a MICROBASE retrieval is not available.

Yes, exactly. A valid cell is one where the LWC/IWC is either zero (clear) or a valid (i.e., non-missing) value greater than zero.

Page 602, last line; you should move the sentence regarding the definition of the rank correlation ahead in the discussion.

We did indeed move it ahead.

Page 604, line 5; stipulate radiative fluxes

Done (the reviewer apparently means page 603).

Page 607, line 4; "Apparently the low and high clouds of summer multi-layer cloud systems are more anticorrelated than in winter". I would expect this structure to be observed in the summer because there is a preponderance of distant active convection over the Rockies that produces cirrus anvils that advect over the SGP. These anvil cirrus are not associated with local convection leading to the observed anticorrelation.

Thank you for this information. However, we think the phenomena you discuss may produce more random (less anti-correlated) low and high clouds in the summer, which is the opposite of what we observe.

Page608, lines 20-30; The alpha histogram resembles a Weibull distribution (and probably others). Have you fitted this curve with some candidate distributions?

We have not. This is an interesting suggestion, but we have not yet given much thought to how such a fit can be used in practice. Perhaps in a GCM grid cell, a decorrelation length can be picked randomly according to such a distribution?

Page 610, line 21; "which tends to be dominated by stability"—I suggest: "tends to be more stratified".

Suggestion adopted.