Reviewer comments appear in *italics*.

General comments

This paper is generally well written. It follows previous literature examining the effect of cloud vertical overlap and horizontal inhomogeneity on top of atmosphere cloud radiative effects (CREs). This is done using diagnostic radiative transfer calculations in two versions of the GEOS-5 model with the difference being the cloud scheme used in the model. This allows the authors to examine the sensitivity of the CRE to assumed cloud structure to the underlying simulated cloud field, which is rather different between the two versions of the model.

1. The discussion in Section 5.1, and associated figures, would benefit from the inclusion of the change in cloud fraction and the change in the overcast cloud fluxes (Eq. 12). The current discussion is somewhat qualitative and could be made more quantitative by using the actual changes in the terms that make up CRE.

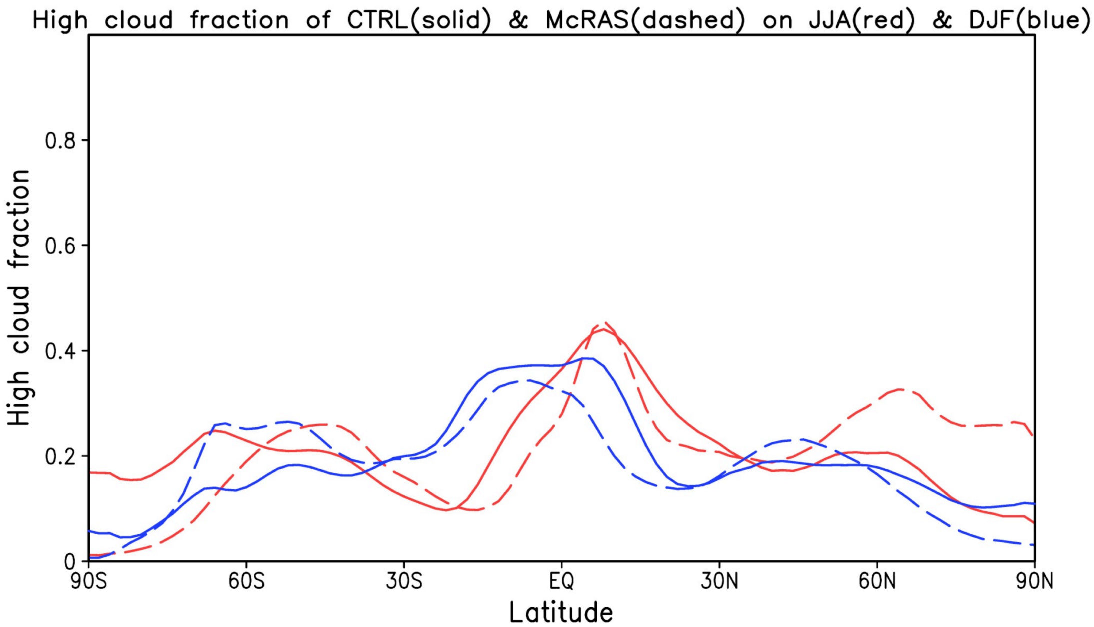
Some discussion along these lines can be found in section 5.3. While it would certainly be interesting to show the changes of F_{ovc} and C_{tot}, this is not possible because the model does not provide F_{ovc} diagnostics. As mentioned in p. 12301, eq. (12b) which shows CRE in terms F_{ovc} and C_{tot} is only used for qualitative interpretation, while eq. (12a) is actually used by the AGCM to diagnose CRE.

It would also be interesting to have some indication of the change in the horizontal inhomogeneity of the vertically integrated water path or optical thickness for all experiments. Although not defined for each layer in the homogeneous cases there can be some variability in the vertical integrals for these cases which affect the radiation.

Yes, it is true that even for horizontally homogeneous layers, fractional cloudiness and overlap creates inhomogeneous distributions of integrated water path or optical thickness. We recognize that while we discuss some of the effects of changing rank correlation decorrelation lengths in terms of integrated water path distributions, the discussion is qualitative (p. 12305). Alas, during our runs neither did we save the cloudy subcolumns of the generator that would allow us to calculate WP/tau variability a posteriori, nor did we calculate and save such a statistic "on the fly". To do so at this stage would require re-running all experiments from scratch something we unfortunately cannot afford to do.

2. For the latitude and time varying decorrelation lengths do you have a sense if the shift in the simulations, e.g., deep convection around the tropics, follows a similar timing and position around the equator in its shift poleward.

This is a good point. Obviously, the temporal shift of convection in the model will not match the seasonal variation of decorrelation lengths. Since it'd be logistically very hard to track the model's convection and adjust the decorrelation length accordingly, we have to retain the approach we used. We take some comfort in the fact that the model seems to have a seasonally varying convection (for both cloud schemes, more clear in McRAS-AC). See figure below, high cloud fraction seasonal mean from Exp. 1.



3. At the end of Section 5.1 you point out that both the longwave and shortwave CRE can not both be tuned to observations using just horizontal homogeneity and vertical overlap. Could you expand this discussion somewhat? You seem to be talking about global means but are regions, potentially dominated by particular types of clouds, in which it is possible.

It's a bit more subtle than this. What we actually said was that tuning cannot be achieved if one CRE is overestimated and the other underestimated. This is because either overlap or inhomogeneity assumptions change both CREs in the same direction. Whether this is local or global does not matter. But it is indeed true that regionally, one will encounter cases where both CREs are overestimated or underestimated; in these cases tuning via overlap and/or introduction of inhomogeneity can potentially work.

4. How does the magnitude and sign of the sensitivities in CRE compare to previously published results?

We believe that the most closely related paper to ours is that of Shonk and Hogan (2010). In section 6 we compared some of their results to ours. Another candidate paper for comparisons is Räisänen et al. (2004). Upon closer inspection, however, it is hard to do a comparison. Their MRO and FGen experiments are similar to our Exp. 1 and Exp. 4. However, their zonal plots of Fig. 7 do not show the differences between these two experiments, but the differences from the ICA "truth", so their curves should be subtracted. Also their global statistics of Table 1 (again bias and rmse are against ICA) are shown in terms of fluxes, not CREs.

5. The zonal mean cloud amount in Fig. 9 are significantly different. Could you comment on the realism of both? Although the focus of this paper is on the sensitivity to different model configurations if the cloud structure are completely unrealistic then the lack of *sensitivity in CRE for a particular version of the model is a symptom of poorly simulated clouds.*

The zonally-averaged profiles of cloud fraction are indeed quite different. Unfortunately, there aren't many observation-based datasets to compare with. Perhaps one of the few exceptions is the cloud fraction vertical distribution dataset CALIPSO-GOCCP. See Fig. 4 http://www.agu.org/journals/jd/jd1005/2009JD012251/. Unfortunately, this figure shows two seasonal means and not an annual mean; moreover, we have not yet explored this dataset enough to decide whether a comparison is appropriate only after GEOS-5 output has been transformed via the CALIPSO simulator of the COSP simulator (http://cfmip.metoffice.com/COSP.html). Nevertherless, a casual comparison seems to suggest that McRAS-AC may be overestimating cloud fractions throughout the entire depth of the midlatitude troposphere, something that probably plays a major role in the smaller sensitivity of this cloud scheme to vertical cloud overlap. If cloud fraction realism is assessed using total cloud fractions, one can also refer to this paper http://www.geosci-model-dev-discuss.net/5/1381/2012/gmdd-5-1381-2012.html. In the upper left panel of Fig. 4, annual zonal total cloud fractions are shown for the same GCM and two cloud schemes used here. The modeled total cloud fraction curves are not exactly the same as the ones shown here because the max-ran overlap implementation of the radiation algorithm used in the GMDD paper simulations is different than the max-ran implementation of RRTMG's cloud generator used here. Still, on this figure alone, one would conclude that McRAS-AC's total cloud fraction is more realistic than that of CTL.

On the bigger issue now of whether it is appropriate to test sensitivity of CRE to overlap if simulated clouds are considered unrealistic by other measures, we recognize that it is a fair point. But these kind of tests are nearly impossible to conduct using observations, so one has to resort to models and accept all weaknesses that come along. Upon further reflection, our paper essentially addresses this very question, whether it is worth being concerned about overlap for certain cloud schemes when other unrealistic aspects of the cloud simulation need to be addressed first. Clearly, a cloud scheme that produces overcast and clear skies more frequently than reality, probably underestimates the effects of overlap, but the effort in this case should not be directed towards overlap realism, but rather towards having better profiles of cloud fraction. Currently, however, we don't have enough observational evidence to support a conclusion that weaker sensitivity of CRE to overlap is necessarily a poor result.

Specific comments

Abstract, line 24: "atmospheric" -> "atmospheric layers"

While this error existed in the Word version of the manuscript we submitted it is not on the PDF version of the discussion paper published.