Response to Reviewers

Note: Only one reviewer had comments on the revised manuscript. Reviewer comments in *italics*, authors' responses in regular font.

1. Wetting of the temperature probe. The authors do not demonstrate that the temperature probe did not get wet in the project GoMACCS. Saying there's no evidence of jumps at the base of Stratocumulus clouds is not proof that it does not get wet when going into a cumulus cloud which has greater liquid water content. I asked to see simple time series of liquid water content, updraft and temperature. Several examples should be shown and a thorough response given.

Response: We previously chose to use Sc because the properties of well-mixed boundary layer are well defined and easily measured. We do agree that the fact that it's not the same project is not ideal. We have now done an analysis for the cumulus clouds in this project. Here's our methodology:

a) For every cloud penetration of at least 4 seconds (200 m in length), we looked at the temperature (T) for the 5 seconds prior to entering the cloud, and for the 5 s after leaving the cloud. If there's wetting, then we would expect a bias in these T measurements. We excluded any penetrations where the entry or exit isn't cleanly separated from other clouds, i.e. where there's a small gap between clouds, it's difficult to say that the thermometer was dry prior to entering the second cloud.

b) For each penetration, we compute the difference in temperature at symmetric times prior to entry and after departure, i.e. we compute Delta_T(j,i) = T(t_entry - i) - T(t_exit + i) for i = 1 to 5 seconds and j is the penetration index.

c) We find that if we analyze only the first second before entry and the first second after departure, i.e. Delta_T(j,1), the mean is 0.18 K, with a standard deviation of 0.35 K, i.e. a temperature difference of zero is well within one sigma of the mean, so we see no statistical evidence for wetting.

d) We repeat the analysis in (b) using the population of values Delta_T(j,i) for i = 1 to 2, i = 1 to 3, i = 1 to 4, and i = 1 to 5 seconds, i.e. with wider and wider symmetric time intervals. No matter the choice of time intervals, the mean is always smaller than the standard deviation by a ratio of between 1.5 and 2. Again, we see no statistical evidence for wetting.

e) If we look at mean Delta_T(j,i) over all penetrations as a function of i, one might expect that if wetting is occurring, then mean Delta_T will tend towards smaller values as i increases and the thermometer dries. However, no such trend is obvious. The order for mean Delta_T from largest to smallest values is i = 2, 4, 1, 5, 3, with 2 and 4 very similar in magnitude, 1 clearly smaller, and 5 and 3 smaller again but very similar. These results also exhibit no statistical evidence for wetting.

We conclude that there is no statistical support for the hypothesis that the T measurements are contaminated by wetting. We have replaced some text in the manuscript on this subject in Section 2.1 with the following text to reflect this new analysis: "Further analysis was done for this study to confirm this result. We see no statistically significant bias in clear-air temperature for time intervals ranging in duration from 1 to 5 s prior to entering a cloud and after leaving the same cloud."

2. Below cloud air. The air below cloud base can be substantially warmer than the air in the surrounding environment. As the authors mention there can be a difference of at least 2 kJ/kg, which is signficant in the calculation of sa (e.g Fig 3).

That is a good point. This would bias our quantitative results but it is unlikely to change the qualitative results. We analyzed the mixed-layer data to get proper values and found that if we used the mean of the 90th to 99th percentile of the temperature values (rather than the overall mean), the temperature would be 0.5 K higher. This leads to a MSE bias of 0.5 kJ/kg. For a sense of how important it is, the sensitivity tests (Section 3.2.2) shift the MSE profiles by 400 m, which correspond to a change in MSE of entrained air of around 3 to 4 kJ/kg depending on the day. Since the latter sensitivity tests didn't greatly change the qualititative results, this mixed-layer air property bias, which is nearly an order of magnitude lower, won't either. We have added the following text in Section 3.2.2 to address this point:

"Another assumption made in this method is that the mixed-layer air comprising the cloud is an unbiased sample. Updrafts, however, are biased towards warmer temperatures so T_a is likely biased slightly low. If we were to use only the top 10% of temperature values in calculating MSE, the temperature bias estimated from the data is about 0.5 K, which leads to a $s_a$bias of 0.5 kJ/kg. This is much smaller than the change in MSE caused by shifting the entrainment levels by 400 m, which is around 3 to 4 kJ/kg depending on the day. Thus, we do not expect the mixed-layer air bias in MSE to substantively change the results."

3. Detrainment layers. The study of Raymond et al (1991) discussed the issues concerning detrainment layers – i.e. they occur at particular altitudes according to the environmental thermodynamic conditions in different days. E.g. cloud top near an inversion. It is more likely to miss that layer with an aircraft that is penetrating clouds than to encounter it! The conclusions (and abstract) should be qualified.

We agree with this point, and the the following text is already in the manuscript to address this issue (end of Section 3.1.1):

"Alternately, as noted earlier, previous studies [e.g. Carpenter et al., 1998a] have inferred that detrainment occurs at specific levels within clouds. Because we only sampled one level of each cloud, we may not have been sampling at the level that detrainment was occurring."

We have also added this qualification to the conclusions (last sentence of the first paragraph).

4. The method should ideally eventually (not in this paper) be compared with model results and with results from an aircraft performing boxes around the cloud.

Agreed, particularly with the modeling aspect. The aircraft box method has large uncertainties (primarily with the assumption that the control volume is at steady state with respect to moisture and energy), at least as formulated in the past. Perhaps there are better ways to do this but at the moment it's difficult to endorse with confidence. We have added the following sentence into the conclusion to emphasize this point: "Testing this methodology in high resolution models is an important future step to gain further confidence in these results."

5. Precipitation. While it is likely that the aerosol concentration was high on most days and it is also likely that the high concentration prevented the production of precipitation, the authors should present the criterion used to exclude clouds with precipitation size drops present.

Fair enough. We see very few drops larger than 50 um diameter near cloud base. The corresponding cloud base rain rates are much lower than 0.01 mm/day (which from other projects we know we can measure easily with this suite of instruments). Exactly how low is tough to quantify since counting statistics are too poor for a good quantification.