Major Comments:

1. Data assimilation methodology
   a. We agree that the forecast error variance model used here does underestimate the Summertime errors over land (particularly where there is significant vegetation). The current model accounts for the latitudinal variation in forecast errors (which are generally higher in the Northern Hemisphere), but not the seasonal variation. This will be added in a future version of the system, using the global in-situ observation network. We have added clarification in the paper about this.
   b. We have added a description of the correlation length scales to the paper, which uses longer zonal length scales in the tropics. This approach has been validated for ozone, carbon monoxide and carbon dioxide assimilation.
   c. We have used a constant observation error standard deviation of 2 ppm in this work. While the actual measurement and retrieval errors are likely somewhat lower than this (and variable), there is also a contribution from representation error (which is more uniform). The large number of tuning runs with comparisons with in-situ data were done to verify this estimate (as well as the background error estimates). Because the background error estimates were born out by the in-situ observations, we are confident that the observation errors used are a reasonable estimate as well. We have added some text on this to the paper.

2. Discussion of Results
   a. We chose 6 representative NOAA surface flask sites with widely different geographical locations, terrain, and success of the assimilation (Not all of these cases show significant improvement). It was not that interesting to plot the comparisons from nearby sites (They look very similar). When plotting the bars in Figure 5, it becomes much easier to see the latitudinal trend in the errors for both the mean and random components when each location is at a different latitudinal location.
   b. We had chosen to use all of the data available from these 3 sources that overlap with the period of our assimilation. We prefer to present error statistics from the largest possible data set. But we did make a plot of the comparisons with a subset of the CMDL aircraft measurements for the February-May 2006 period. There are no big differences in the results, but we have included these plots in the paper. Note that one of the sites did not have sufficient observations for this period to
create meaningful statistics, so it is not included in the paper. We also changed the comparisons with the aircraft profiles to the period 2006, but there were not as many flights at the THD measurement site, so we changed this to BNE (in Nebraska) which also lies along the 40° N latitude that we are analyzing here. We have changed the text accordingly. We did misstate the dates of the INTEX-B campaign in the introduction, and it has been corrected to read February - May 2006.

(c) We have changed this figure to cover the period January-December 2006. The text has also been changed accordingly. And we have also included a comparison with TCCON column observations over Park Falls, Wisconsin.

(d) We have added some text to discuss this comparison further. It is possible that there remains a small negative bias in the AIRS retrievals, but the southern hemisphere errors remain much smaller than those in the Northern hemisphere. Note that the y axis scale is just 8 ppm for this figure, while the rest are between 20 and 25 ppm.

3. Description about carbon cycle

(a) This approach is already being used at the ECMWF. Please see Engelen, et al. (2009) and Chevalier et al. (2009), who use a two step method for flux inversion. The carbon dioxide analysis from the assimilation is used as the input to the flux inversion. We have changed the text to be sure that this existing approach is described, but have removed any statement that this assimilation would be appropriate for flux inversion, since we have not yet done so.

(b) We have included additional explanation about the surface flux forcing. The CO$_2$ in our model is in fact too high (not too low). The plots of $O - F$ are generally negative (we are using the standard data assimilation convention here, though these observations are not assimilated. The high bias is likely result of an underestimated carbon sink, and the fact that we have included biomass burning, in addition to the TRANSCOM fluxes, which are carbon neutral.

(c) We agree that this sentence was not clear it has been changed to say that where there is little seasonal cycle the model errors are much lower, and it is a challenge to improve the CO$_2$ field through assimilation.

(d) This comment has been removed.

4. Figures

(a) If log scale is used, it is very hard to see the difference between the two averaging kernels because most of the space is in the stratosphere. A linear scale is really the only way to plot this.

(b) We have removed Figure 6a, and agree that it is unclear.

Minor Comments
1. This reference has been dropped.

2. We have changed the wording here.

3. This has been changed.