Atmos. Chem. Phys. Discuss., 12, C12370–C12388, 2013
www.atmos-chem-phys-discuss.net/12/C12370/2013/


Atmospheric Chemistry and Physics Discussions

# *Interactive comment on* "Comparison of improved Aura Tropospheric Emission Spectrometer (TES) $CO_2$ with HIPPO and SGP aircraft profile measurements" *by* S. S. Kulawik et al.

**S. S. Kulawik et al.**

susan.kulawik@jpl.nasa.gov

Received and published: 7 February 2013

—— Overview of responses and updates ——

Thank you to the reviewers for their helpful comments and suggestions. We have responded to each comment below.

Because the TES official v5 processing recently finished, the analysis was updated from the TES prototype CO2 product to the official TES v5 CO2 product. The results are similar but figures and tables have the updated comparisons. The HIPPO data were updated to the latest version of that dataset. There was no update to the SGP data other than additional time points.



Because the analysis was done on the TES CO2 product, rather than specially processed data, we were able to look at longitude shifts and box sizes for the SGP (land) data as suggested by referee #1. Analysis with specially processed data using constant prior was kept as the specially processed prototype results.

We found and fixed an indexing bug in the tacking on of AIRS or CONTRAIL data above SGP data which improves the effect of using AIRS or CONTRAIL. There was also a minor update to the interpolation of the aircraft data onto the TES pressure grid which had little effect on the results.

In response to referee #1, we have placed more emphasis on results at all pressures, particularly the addition of curtain plot comparisons in Fig 5 and Fig 11, and the addition of near-surface results to all tables.

Also in response to referee #1, we have characterized biases better spatially and temporally. Studies of the yearly increase indicate a time-dependent bias of -0.2 ppm/year which was corrected. Additionally we found a bias for 2010 and later (due to changes in TES calibration). Finally, we compared to CarbonTracker (CT2011) to characterize biases within a larger context and found a spatial bias pattern which is persistent in time, shown in Fig. 7d and 8.

The following co-authors were added, which are the data contributors from HIPPO: Rodrigo Jimenez, Sunyoung Park, Greg Santoni, Bruce Daube, Jasna Pittman, Britt Stephens, and Eric Kort.

—— End of overview ——

Referee #1 -It is not clear in the paper that most of the results and comparisons correspond to TES CO2 511hPa level. So, I have difficulties with the term 'profile' in the title and the abstract. It is not clear in this paper, (as the authors say §4.3) how the improved TES CO2 product can capture CO2 vertical structure. In particular conclusion of §4.2.1 shows that TES CO2 is not really sensitive to CO2 above 5 km.

We restructured the paper to emphasize the results at all pressure levels and have completed an OSSE to assess the effects of the full profile. We added curtain plot comparisons showing all pressures in Fig 5 and Fig 11, and added near-surface results to all tables. The indexing bug that was fixed for the combined AIRS+SGP now shows a significant improvement from 1.10 to 0.80 ppm standard deviation over using SGP alone. The new OSSE section (3.5.1) states, "Previous flux estimates using TES CO2 utilized observations at 511 hPa over the ocean (Nassar et al., 2011). We use an observing system simulation experiment (OSSE) to assess the information added by the full profile and by land observations. Using a similar OSSE to that which was used in Kulawik et al. (2010) we found an increase in 1.4 DOF when including TES land results and 0.1 or less DOF change when including all levels versus just the 511 hPa level (at about 5.5 km). The small increase when including all levels is because the averaging kernel row is very similar for all TES pressure levels as seen in Fig. 2. Even though little information seems to be added by using the full profile, 3-D var assimilation of TES profiles (described in Kuai et al. (2013)) compared better to validation data than a single level assimilation (unpublished work)."

-It would have been very interesting not to give only global rms error and bias of TES CO2 products but also to provide latitudinal and seasonal variation of these important parameters for model assimilation. I guess the main problem is the number of collocated measurements, but even if error increases their variations can provide useful information on TES CO2 products. You should try to plot it.

Since HIPPO-1, -2 and -3 cover 3 months in the Pacific, it was hard to draw general conclusions about season and latitude. We note the persistent low values at about +-15 more prominently. In response to this comment, we compared to the CarbonTracker assimilation system (CT2011), in Fig. 7d and Fig 8, finding location-dependent biases that are persistent in time and have developed a spatial-based correction which is now in the TES CO2 Lite products.

- Time and latitude best averaging of TES products is not clear and discussed enough,



it appears sometimes in contradiction with authors average choices.

We have updated the analysis to analyze HIPPO and SGP consistently, both averaging over +-5 degrees latitude, 10 degrees longitude, and 1 month.

Specific comments - P6284 Abstract : Two main information are missing in the abstract : the latitudinal range of comparisons (1st sentence),

We added text to clarify in the abstract, “We compare to ocean profiles from the first three Hiaper Pole-to-Pole Observations (HIPPO) campaigns between 40S and 40N with measurements between the surface and 14 km and find that TES CO2 estimates capture the seasonal and latitudinal gradients observed by HIPPO CO2 measurements. ”

- and the pressure/altitude of quoted errors in the sentence L12-15.

We added the following descriptor to clarify that TES is sensitive in the mid-Troposphere, “Validation requires validation data sources spanning the range of TES CO2 sensitivity, which is approximately 2.5 to 12 km with peak sensitivity at about 5 km and the range of TES observations in latitude (40S to 40N) and time (2005-2011) . “ The analysis has been updated to look at a range of TES pressure levels.

- L11 : Even if TES CO2 products capture well seasonal cycle and latitudinal variation, you should mention that this paper provide only global bias and error and do not give information about error and bias variations with season and latitude.

We added wording to the abstract to clarify this, “These comparisons are used to characterize the bias in the TES CO2 estimates and to assess whether calculated and actual uncertainties and sensitivities are consistent for these specific locations and time periods.”. We have also updated the analysis to include comparisons to CarbonTracker to assess spatially-dependent biases. Comparisons to SGP do not show any repeating seasonally-dependent biases but we have found and corrected a time-dependent bias trend of -0.2 ppm/year.

- P.6285. L16-L30: Is this paragraph necessary ? It might be shorter, in particular from L26 to L4 P6286.

This paragraph describes the steps previously taken to correct TES in order to be useful for assimilation and it puts the biases and errors found in this paper into the context of what is needed for scientific analysis.

- P6287 : What is the SGP measurement accuracy ? You should mention it.

We added this information, ". . .with a precision of $\pm$0.2 ppm (Biraud et al., 2012)." The errors from non-coincidence of TES will be much larger than this.

- P6287 : L13 Why this sentence inside brackets ? It is an important point. You should insert that TES spatio-temporal range impact is discussed at §4.1.2.

We rewrote this section to more carefully describe the updated averaging: " We select all HIPPO measurements for each campaign within a 10° latitude band of a TES observation. If the HIPPO measurements are separated by more than 30 days or 20° longitude, they are split into two groups and each group is averaged. TES measurements for the same latitude range, $\pm$10° longitude from the HIPPO average longitude, and $\pm$15 days from the HIPPO mean time are averaged for comparison. The impact of varying the coincidence criteria for time, latitude, and longitude is discussed in section 4.3. "

- P6287 : L25 : meaning of flag ' ..P ' ?

We changed the wording to "Measurements with good quality were used (flags '..P')". The description of the quality flags can be found in the SGP documentation; it seems sufficient to say that we screened based on quality and give the criteria that were used.

- 6289 : L27 : Are you sure that extending v2 band do not increase TES CO2 sensitivity to water vapor uncertainties?

Extending the v2 band should increase the accuracy of H2O, which in turn should



make H2O less of an interferent for CO2. However, the low values seen in the HIPPO comparisons at +-15 degrees latitude seem to be correlated to patterns seen in H2O. The predicted and actual errors both are better than the previous version but there is still improvement that can be made.

- 6294 : L11 . Averaging kernel is a matrix (A), AK corresponds to row of A each row corresponding to a TES level. You should be more precise.

I was not familiar with the convention that AK is a row of the A matrix. I changed all instances of AK to "A" or "averaging kernel".

- 6295 : Why do you use a different latitudinal range (4_ for HIPPO and 5_ for SGP )? Is there a specific reason ?

It was to allow more distinct HIPPO comparisons. However, this was updated to using +-5 degrees latitude to be consistent.

- 6299 : L19 : Looking at table 5b, for latitude, actual error decrease from 4_ (medium criteria) to 6_ (loose criteria), from you should mention it even if bias increase. You must discuss this point in term of best latitude range for TES average.

This table and text was updated with v5 results. We do note now that all comparisons improve for even loose coincidence criteria.

- 6299 :§4.1.3: Do you obtain such results on longitudinal shift with SGP comparisons ?

The prototype did not have the data processed for the longitudinal shifts. The paper has been updated to use the recently completed v5 dataset for all comparisons. The comparisons now include longitudinal shift at SGP.

- 6300 : L12: From conclusion of §4.1.2, the best time average is 14 days, so, why do you average TES over 1 month for SGP comparisons ?

1 month is about +-15 days. The +-14 days was just +-1 day different. It was a lot

easier to think of 1 month, "April", "June" as opposed to breaking the year up into 14 day increments. However, +-15 days is now used for HIPPO as well.

- 6301 : L10 : The conclusion is : missing validation data above 5 km create at least a mean bias uncertainty of 0.3 ppm. You should add ' for TES CO2 for 511 hPa level'. Can you compare this value (0.3 ppm) with the mean difference between the 3 profiles (above 5km) you use to create the 'true profile' ?

Added " in the mid-troposphere" to the following sentence to 4.2.1, " Extension of the SGP aircraft data with AIRS CO2 values changes the bias from 0.21 to 0.13 ppm, and improves the standard deviation from 1.10 to 0.80 ppm in the mid-troposphere and changes the bias from -0.1 to -0.3 ppm and improves the standard deviation from 1.00 to 0.96 ppm near the surface. The use of CONTRAIL aircraft data to extend the SGP profile improves the standard deviation from 1.10 to 0.82 and increases the bias slightly from 0.21 to 0.45 ppm in the mid-troposphere. " We also calculated the differences in 'true' at 10 km, "The difference between extending the SGP profile versus AIRS is -0.2 +- 1.4 ppm (AIRS higher) and versus CONTRAIL is 0.5 +- 1.5 ppm at 10 km (CONTRAIL lower)."

- 6303 : §4.3 is not so clear, the conclusion seems to be : TES CO2 for 511 hPa level

I think the reviewer's sentence was cut off.

- 6304 : L22 : You mention that "averaging TES within 5_ latitude, 10_ longitude and 14 days gives the best results" , you should add that is true only for ocean retrievals (not been tested for SGP over land), then you didn't test 5_ latitude but only 4_ and 6_ and it is not clear that the best range is between 4_ and 6_ (see previous remark). A more general comment about your conclusion : over land you use 1 month time average, so the number of TES measurements averaged is more important than over ocean for HIPPO comparisons (14 days average), so it is difficult to compare land and ocean rms and bias values of TES CO2 products. It can be very intersting to discuss the difference between the TES CO2 product quality over land and over ocean.



We added SGP into the analysis of coincidence criteria. We agree land and ocean coincidence criteria are likely different, however it seems that averaging is needed for TES regardless of location.

- 6316 - 6317 : Legend of table 5 and table 6. Can you precise what is the "mean" : it is not clear why the number 'n' is not the sum of others 'n' ?

The word "mean" doesn't make sense there. I took it out, thanks. The number 'n' is the average number of TES observations for each comparison. This is important because the error should scale as the inverse of the square root of the number of observations. The caption now reads, " Calculated correlations ("corr"), predicted ("pred") and actual ("actl") errors and biases (in ppm) when averaging within 5°, 10°, and 20° longitude for each of the datasets at 900 hPa near the surface ("surf") and 500 hPa in the mid-troposphere ("trop"). n is the # of TES observations averaged per comparison (important because the error should scale as the inverse of the square root of the number of observations per comparison if errors are uncorrelated and the measurements have the same true). The correlation and errors improve with increasing box size indicating that quasi-random, rather than systematic, errors dominate."

It seems that you make an average of the results of each flight, why don't you recalculate parameters using the 3 campaign together ? The result should be quite different as the number "n" will be higher.

The number "n", defined as the # of observations averaged for each comparison, would not be higher, so the error would not decrease by including all campaigns together, unfortunately.

- 6317-63199 : Actl error : why do you use a ratio ? I think that actual errror in ppm would be easier for interpretation.

We think that it is important to assess the predicted versus actual errors. But since it will likely be confusing to the reader we changed it to actual error instead of error ratio.

Technical comments - 6287, L9 : WMO signification ?

We spelled out "World Meteorological Organization (WMO)". This organization sets standards that atmospheric measurements can be traced to. E.g. if I validate vs. flasks which are validated to 0.2 ppm by WMO, my measurements can be traced to a WMO standard.

- 6291 : L8 : In the text use logarithm instead of “log” (just my opinion, to be confirmed with editor). Check it everywhere in the text.

We changed log to logarithm in all the text descriptions, but keep log() in equations.

- 6291 : L8-L9 : you should write : where Xest , Xtrue and Xa are the logarithm of the estimate, the true state and the a priori constraint vector, A is the averaging kernel (sensitivity of the estimate to the true state), G is the gain : : :.You should remove ’=dXest/dXtrue’ from the text.

Changed.

- 6291 : L15 , replace “the retrieved parameter is x = log (VMR)” by “the retrieved parameter, x, is the logarithm of the gas volume mixing ration (VMR)”

Changed.

- 6294 : You should replace “ reduced by 1/sqrt(40)” by ‘reduced by a factor square root of 40’. Generally, you should try to avoid to insert in the text ’1/sqrt(# measure)’ : To be confirmed by the editors.

Changed

- 6297 : L16 : Error in the sentence ’ comparisons nd predicted ’, what is nd ??

This was a typo. The text for Fig 6 has been updated with the new data version.

- 6299 : L11. sect “4.1.3” instead of “4.1.2”

Changed



- 6299 : L20 : You should remove a dot just after ’improve..’

Changed

- 6299 : L 21 : A “n” is missing in longitudinal (title 4.1.3)

Changed

- 6313 : A ’ End ’ is missing in the last row of “step 2”

It’s in the original word document. I’ll watch for this in the updated proofs.

- 6315 : Errors Unit are missing in table 4

Units added in the top of the table.

- 6316 : Errors and bias Unit are missing in Table 5

Units added to the caption: " Calculated correlations ("corr"), predicted ("pred") and actual ("actl") errors and biases (in ppm) when averaging within 5°, 10°, and 20° longitude for each of the datasets at 900 hPa near the surface ("surf") and 500 hPa in the mid-Troposphere ("trop")."

- 6316 : a bracket is missing line 9 of the legend.

Updated

- 6317 : Errors and bias Unit are missing in Table 6

Table 6 has been changed. Units are mentioned in the caption, "Similar to Table 5a, comparisons in ppm between TES and validation data with and without spatial bias correction (see section 3.6). "

- 6317 : In the last sentence you say “latitude” instead of “longitude”.

Updated. Thanks.

- 6318 : You should mention that column signification correspond to the previous figure.

I don't follow this suggestion. However, note that Table 7 previously on page 6318 has been folded into table 4.

- 6320 : Problem with left panel x-axis in figure 1

Updated

- 6321 : Problem with middle and right panels x-axis in figure 2

Updated.

- 6326 : L6 of the legend, replace "Eq. 6d, " by "Eq. 6 (d),"

Updated. It should just be Eq. 6.

REFEREE #2 - The authors present a thorough validation of a new TES CO2 product with several aircraft in-situ and AIRS satellite observations. The most problematic issue is the somewhat arbitrary correlation criterion for comparisons with the HIPPO campaign. The time window is large but statistics on the actual time match quality or e.g. back trajectory analyses are not provided. I think some more justification beyond a better fit between TES and HIPPO is needed. For SGP comparisons, monthly mean comparisons seemed to be fine. It is not clear why this was not also done for HIPPO

We have updated the comparisons to average HIPPO to average over the same latitude, longitude, and time range as for SGP, which is 10 degrees latitude, 20 degrees longitude, and 30 days time.

- Another point is that since this article is an update of a previous data version, it should be more clearly pointed out what is new and what is essentially the same as in Kuwalik et al. 2010. This mainly affects Section 3.

Section 3.1.1 lists the "Updates from the previous version" and in section 3.2 the statement was added, "The error analysis through Eq. 4 is shortened version from Kulawik et al. (2009). ". For Eq. 5, added the reference, "... (Boxe et al., 2010 Eq. 11)"



General comments: - the title only mentions HIPPO and SGP. However, you also use CONTRAIL and AIRS for your validation.

The main comparisons are with SGP and HIPPO, using CONTRAIL and AIRS to fill in the data gap in the true state at SGP above 5 km. AIRS, CONTRAIL, and Carbontracker (version CT2011) are now prominently mentioned in the abstract.

- TES matching criteria: 10 deg latitude, 4 deg longitude, 14 (!) days. That sounds like a very broad range. Can you justify especially the large temporal match criterion? In 14 days, an air parcel would be transported over a much larger geographical range than your lat/lon criterion.

We have explored variations in the coincidence criteria in Table 5b and also added temperature-based coincidence criteria which follow Wunch et al. (2011) and consider a wider longitude bin and narrower time window. Results seem to mainly scale with the # of profiles that are included.

- logarithmic retrieval, p. 6291: I understand that log retrievals are typically used to handle large dynamic range of a species (e.g. H2O) or to force positive values for small values with comparatively large noise. It is not obvious to me why this would be optimal for CO2. Please explain why a log scale retrieval was chosen (or provide a reference if it is explained elsewhere).

The TES retrieval code is set up to do retrievals of trace gases in log(). We added some explanation in section 3.2 as to why log(). " Note that for TES, all parameters besides temperature and emissivity are retrieved in log(), so that the retrieved parameter, x, is the logarithm of the gas volume mixing ratio relative to dry air (VMR). In TES processing, all trace gases are retrieved in log(VMR) because many of the trace gases measured vary logarithmically. "

- througout the text, the term "target" is used without a clear definition.

Changed "target" to "observation", e.g. changed to " A comparison of TES and CT2011

for southern hemisphere ocean observations between 20S and 40S found a similar trend"

- Section 3.2: it is not clear to me if this is new or a summary of the procedure that has also been used before.

The statement was added, "The error analysis through Eq. 4 is shortened version from Kulawik et al. (2009). ". For Eq. 5, added the reference, "... (Boxe et al., 2010 Eq. 11)"

- target-averaging: I am not convinced that averaging over some 40 targets will reduce biases in the 1-2

We agree that averaging should not have any effect on biasing. It seems like the reviewers comment was cut off but there is no longer improvement in biases with averaging.

- the term volume mixing ratio (VMR) is used but it should probably rather be dry-air mole fraction (DMF). For species in the ppm region, the difference between the two is very small. However, for abundant species like N2 DMF(N2)=0.8 while VMR(N2)=4!

Added a statement defining VMR "...is the logarithm of the gas volume mixing ratio relative to dry air (VMR)". I am used to VMR being defined as you define DMF above but we only work in trace gases.

- many of the figures are hard to read and look pixeled even if enlarged by 400

Updated the way the figures are saved where possible.

Specific comments:

- Introduction: I miss information on the temporal range for which TES data is available.

Added statement in section 2.1, " TES global survey observations were consistently taken from late 2004 through June, 2011"

- p. 6286, l. 8: providing some more details on the vertical width of the TES averaging kernel would be useful here. Also, please provide approximate altitude in meters with



the pressure values. This makes it easier to compare this with the flight altitude range mentioned earlier.

Added in altitude most places pressure is measured. In Fig 2 caption added, " The FWHM pressures, where the averaging kernel has half the peak value, occur at 750 and 215 hPa (2.5 and 12 km)."

- p. 6286, l. 21: better write "660 cmôĂĂĂ1 to 2260 cmôĂĂĂ1". BTW: the TES L1B data product quality description (V2) lists 650 cmôĂĂĂ1 as the lower wavenumber range and suggests not using data within 30 cmôĂĂĂ1 of the spectral range boundaries of the individual filters.

Thank you for noticing this. In response to this, the TES project has updated the outdated quality description to list the filter ranges in the TES L1B data product. For 1B2, the range is now listed as 950 -1130.

- p. 6286, l. 25: Table 1 lists 500 obs/day instead of 2000-3000. Why the discrepancy?

Section 2.1 now explains this better, "In 2006, TES averaged 1570 "global survey" observations per day. Of these, 743 per day are between 40S and 45N, and 505 per day have cloud < 0.5 OD and are of good quality. ".

- p. 6287, l. 7-9: again, pressure + altitude and more information on the vertical range of the TES averaging kernel would be useful here.

Added information regarding pressure and altitude relationship, " The profiles are measured between 0.3 km and 9 km ($\sim$307 hPa) with some extending up to 14 km ($\sim$151 hPa), covering a large fraction of the TES vertical sensitivity " Added text, " For comparison, the TES mid-tropospheric averaging kernel has a full-width-half-maximum range of 2.5 to 12 km. "

- p. 6287, l. 11-13: the sentence about the matching criteria should not be in brackets.

Updated this section. It now reads, "We select all HIPPO measurements for each

campaign within a 10 degree latitude band. If the HIPPO measurements are separated by more than 30 days or 20 degrees longitude, they are split into two groups. Then each group is averaged. TES measurements for the same latitude range, +-10 degrees from the HIPPO average longitude, and +-15 days from the HIPPO mean time are averaged for comparison. The impact of varying the coincidence criteria for time, latitude, and longitude is discussed in section 4.3. "

- p. 6287, l. 13: strange term: "HIPPO-identified"?

Updated wording to, " We use the profiles identified by the HIPPO team and the CO2.X field, based on 1s data median-filtered to 10s."

- p. 6287, l. 14: Please explain CO2-QCLS and CO2-OMS

These are different CO2 sensors on this mission. Text is updated to, " The CO2.X field is primarily derived from the quantum cascade laser spectrometer (CO2-QCLS) measurement with calibration gaps filled by measurements from the Observations of the Middle Stratosphere (CO2-OMS) instrument. For description of these sensors, see Wofsy (2011) and documentation online (http://www.eol.ucar.edu/projects/hippo and http://hippo.ornl.gov). "

- p. 6287, l. 17: "at levels of 2 ppm or more"? Do you mean more than 2 ppm difference between flask sample values and continuous in-situ measurements?

Yes. Updated text to clarify, " Note we do not use CO2 profiles from HIPPO-1 flights 8-11, when these 2 CO2 instruments received a small fraction of air contaminated by the aircraft cabin. The contaminated measurements showed more than 2 ppm altitude-dependent differences from flask data and a third in situ measurement. Flight 7 CO2.X data have been altitude-adjusted to match the flask data and correct for a small contamination effect of less than 1 ppm. Changes to the aircraft sampling system were made after HIPPO-1 and no contamination was detected thereafter in the reported data."



- p. 6287, l. 18: please provide flight altitudes consistently, either in m or km.

Updated all to km.

- p. 6289, l. 1: please also provide wavenumber values for nu2 and the laser bands for readers not familiar with Kulawik et al. 2010. How many laser bands are there? If too many, please add a table.

Updated this to read, "The laser bands are located between 900 and 1100 cm-1, in a relatively transparent region of the spectrum. We use the two bands centered at 960 cm-1 and 1080 cm-1."

- p. 6289, l. 8: what is "cloud pressure"?

Changed wording to, "cloud optical depth and height"

- p. 6289, l. 12: "biased even more" means with a larger low bias than the 511 hPa level?

Yes. Unclear wording updated to, " The 511 hPa result (at about 5.5 km) is biased low by about 6 ppm, with the surface result tending to be biased more than 6 ppm and the 133 hPa result tending to be biased less than 6 ppm."

- p. 6291, Eq. 3: The S matrices in Eq. 3 are not explicitly defined, so it is hard to follow the transition from Eq. 2 to Eq. 3. If you are using the same steps and notation of Connor et al. 2008, please say so.

Additional direction added for the reader, "For more details on the derivation and terms in Eq. 3, see Connor et al. Section 4.1 or Kulawik et al., section 3.3."

- p. 6292, l. 19: how would spectroscopic/calibration errors contribute to random error?

Errors, particularly line width errors, can be picked up by different retrieval components and can alias into both a constant offset and a variable error. Inconsistent calibration between bands can result in a variable error depending on variations in sensitivity.

Calibration errors can similarly alias into both a constant offset and a variable error. If I knew the nature of the error I could characterize its effect.

- p. 6292, l. 22: I think the validation with the aircraft profile should be a separate subsection.

Created a new subsection, "3.3 Comparisons to aircraft profile data" p. 6292, l. 24: I am confused with the different pressure grids used so far (I counted three so far: 5-level, 65-level, Table 3). Could you please add a table that lists them or explain them somewhere in Section 2.

The different grids are now mentioned in Section 3.1.1 and 3.2, "TES uses 65 pressures for the radiative transfer pressure grid. The retrieved parameters for CO2 are on a reduced set of pressures, e.g instead of retrieving 65 CO2 values, we retrieve 5 in Step 1 (see pressure list in Section 3.1.1) and 14 pressures in Step 2 (see list in Table 3)."

- p. 6293, l. 14: in what sense is your retrieval "non-linear"? Do you run several iterations?

Yes. Text added to section 3.2 to clarify, " The predicted sensitivity and retrieval non-linearity can be validated, as described in Kulawik et al. (2008, 2010), by running non-linear retrievals using two different a priori vectors, and , resulting in the iterative, non-linear retrievals, and , respectively."

- p. 6294, l. 6: Please use math symbols instead of sqrt().

Updated.

- p. 6294, l. 13: I disagree with "This AK shows the potential for resolving CO2 at different pressure levels once the spectroscopy is addressed". The step 1 AK may allow something like two seperate altitude retrievals. However, the kernel rows for step 2 basically overlap completely with a peak at the same altitude.



Averaging kernels always overlap for nadir-viewing instruments. Updated text to state, "A shows the potential for partially resolving CO2 at different pressure levels."

- p. 6295, l. 14: please use consistent naming for the HIPPO campaigns (not "HIPPO 3", "HIPPO-3" etc.

Updated all to HIPPO-X.

- p. 6304, l. 11, 15, 18: please fix at least 3 misspellings of "sensitivity" on this page.

Fixed. Thanks.

- p. 6320, Fig. 1: the bottom right plot misses the tick marks for pressure, the pressure axis should be labeled with units

Updated.

- p. 6321, Fig. 2: the small text on the figure is hard to read and looks pixeled.

The small text on the figures is also included in tables and text.

- p. 6322, Fig. 3: the SGP plot looks very pixeled, the latitud labels are unreadable.

Updated plot.

- p. 6323, Fig. 4: the right panels might go into a separate figure to be better readable. What is the black dashed line on the right panels? p. 6324, Fig. 5: there are no (a), (b), or (c) labels on the actual panels.

The right side of Fig 4 is now in Fig. 6. Labeled all subpanels with (a)-(f). Added a description of the black dashed line, " Plots versus latitude for the 511 hPa pressure level showing the TES value with error bars (red), HIPPO at the same pressure level (black), HIPPO with the TES observation operator applied (blue), and TES prior and initial guess (green)."

- p. 6325, Fig. 6: the statistical info on the panels is very interesting but hard to read. You might want to put is just to the right of each panel.

Added a statement in the Fig 6 caption, " The statistical information for panels (d)-(f) is listed in Table 6."

- p. 6326, Fig. 7: sorry, I did not understand from the figure caption what the difference between the orange and red TES data is.

This is now Fig. 9. The caption wording has been updated. It now reads, " Panels (a) and (b) show aircraft data with and without the TES averaging kernel applied, in green and orange respectively..."

---

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 6283, 2012.