

Interactive comment on “Lower-tropospheric CO₂ from near-infrared ACOS-GOSAT observations” by Susan S. Kulawik et al.

Anonymous Referee #3

Received and published: 4 November 2016

This work separates the information in the CO₂ data from the GOSAT satellite into partial columns of CO₂ below approximately 2.5 km elevation ("LMT") and above 2.5 km ("U"); the sensitivity of each necessarily bleeds into the other part of the column, due to the slim vertical information content of the measurements. Bias corrections are computed for both LMT and U by comparing each partial column to aircraft data, sampled in the vertical to be consistent with the sensitivity of each GOSAT partial column. The accuracy of these bias-corrected LMT and U values are then assessed by comparing them to a) aircraft measurements (including some not used in the bias corrections), b) in situ CO₂ measurements at surface sites located far away from the influence of continental air, and (more qualitatively) c) to CO and fire count measurements from the MOPITT satellite over the tropics.

In general, the computed lower tropospheric CO₂ partial columns (LMT) compare quite

C1

well to the independent surface CO₂ data at oceanic sites. The LMT patterns over Africa and South America compare quite well, qualitatively, to patterns of MOPITT CO at the surface as well as to MODIS fire counts; this would be expected if CO₂ produced by fires were to be the dominant source of CO₂ variability in the tropics. The U patterns also are similar to the seasonal outflow off those continents seen in MOPITT CO at 5 km. Finally, the bias-corrected LMT and U products agree well with aircraft CO₂ profile data, when they are sampled with the appropriate GOSAT averaging kernels. Since the aircraft data themselves were what the bias corrections were computed from in the first place, this is perhaps to be expected, to some extent. The real question is how well the bias-corrected LMT and U values would compare to columns computed from aircraft data not used in the bias correction calculations. This is difficult to assess here: the bias correction and validation steps used different sets of aircraft profile data that overlapped each other by about 50% (Line 1384) and which were chosen using two different coincidence criteria. It would have been better to keep the aircraft profiles used in the bias correction separate from those used in the validation.

Further complicating the validation step over land, the criterion used for time coincidence was +/- one week: in other words, aircraft measurements seeing air with a complete synoptic weather cycle or more of CO₂ differences were compared to the GOSAT data. This was done to increase the number of comparisons available, but the effect of this was to make it difficult to separate the impact of measurement biases from true CO₂ variability over land. This "co-location error" was assessed by sampling the analyzed CO₂ fields from CarbonTracker at the same GOSAT - aircraft location differences, but it is not completely clear from the text how these co-location corrections were used (subtracted off on a shot-by-shot basis or just statistically at the end). The overall bias correction applied to the LMT data seems to vary by about 1.7 ppm between using a 48-hour time coincidence criterion and a 1-hour criterion (Lines 1352-1357), suggesting that the uncertainty in the bias corrections is indeed large.

The main conclusion I take away from this manuscript is that the GOSAT data contain

C2

a lot of useful information about lower tropospheric CO₂ variability that is mostly lost when packaged in the form of XCO₂ that most modelers have used to infer fluxes from up to now. A second conclusion I take is that it is not particularly easy to validate the lower-tropospheric part of the column with aircraft data, due to the sparseness of the coincidences. Overall, the authors have done a very nice job with this work. I have made some suggestions below for clarifying the text. In particular, I think the description of the values presented in the tables is currently quite confusing and could be clarified by better descriptions in the captions.

Detailed comments:

L 25: maybe call them "partial column mixing ratios"? The partial column amount could be in other units...

L 30: rather than saying "errors", could you be more precise and say "root mean square errors" (or whatever is appropriate for the statistic you are using)? We say "errors" colloquially, but it would be better to be more precise in the published work.

L 40: I would say "separated better"

L 46: Again, "uncertainty" would be better than "error" here and elsewhere in the manuscript

L 61: "model assimilation" – do you mean "data assimilation"?

L 81: Stephens et al investigated the separation of the extra-tropical northern land fluxes against the sum of the tropical land fluxes and the extra-tropical southern land fluxes, not just the southern hemisphere land fluxes, as is currently stated. Please reword this.

L 122: For clarity, I would suggest rewording this to: "ACOS-GOSAT v3.5 XCO₂ values from the Lite product with quality flag of 0 are used, along with the full CO₂ profile...". Also, immediately after, it is not clear whether you mean the averaging kernel matrix or vector by "profile averaging kernel" – maybe you could word that more clearly.

C3

L 130-136: It might be worth noting that the aircraft data have errors themselves, but that you are considering them to be small compared to the other errors in the problem.

L 134: by "Tropopause" do you mean "tropopause height"? If so, please reword (and no caps).

L 140-141: Ok, here you define ESRL – it would be better earlier. Ditto for "ObsPack". You should mention that these are CO₂ measurements you are looking at.

L 145-148: This screening approach for the remote ocean data initially had me confused. I understood why "nighttime" data would not be compared to the early afternoon satellite data, but why would you throw out "marine" (which attempts to capture data from the ocean, which should be less prone to diurnal variability than the continental air"? And why would you accept "continental" air without ensuring that it is taken at approximately the same time of day as the satellite data (early afternoon)? And finally, if you accept "allvalid" data, it would seem necessary to check the time stamp on the data to ensure that it is taken at approximately the same early afternoon local time that the satellite data are taken, to avoid potential biases due to diurnal variability of CO₂ (especially for air coming off of any nearby continents). Looking at Figure 2, I see that all the sites you have selected are well off the continents, so maybe that is the answer: you are probably not seeing much diurnal variability coming off the continents. Why do you throw away "marine" air, though?

L 170: "15 CO₂-OMS" is this correct? (What is the "15" for?)

L 204: "the bottom 2 pressure levels" – of what, the retrieval? If so, what pressure thickness do they go up to?

L 210-211: Could you give a reference for the fire map work of Desclotres?

L 252-255: It would be interesting if you could come up with a figure that showed how these degrees of freedom split nicely at 2.5 km above the surface. Not required, but would be interesting...

C4

L 258-262: If this is how LMT and U are defined (on a pressure grid), it would be good to note here explicitly that the 2.5 km definition that you use elsewhere in the document is an average definition, and that the split height goes up or down depending on surface pressure.

L 272: replace "26" with " $n_{interf}=26$ ", so that the size of n_{interf} is defined.

L 274: "size $n_{interf} \times n_{CO2}$ ": this refers to the size of A_{xy} , not $A_{xy}(y-y_a)$, as stated.

L 280: why wouldn't the full-layer values in this vector be $1/19 = 0.0526316$? If so, the rounding to 0.52 is incorrect.

L 294-295: Hopefully more than two significant digits for these quantities were used in the actual calculations. For the first value in h_u , I calculate 0.034483, so the 0.035 that is given isn't rounded correctly.

L 304: Why wouldn't $f_{LMT} = 9/38 = 0.236842$? That is 1% different from what you have.

L 305: Why wouldn't $f_U = 29/38 = 0.7631579$?

L 309: You've left off the epsilon at the end of equation 5a.

L 312: Since you have defined your pressure weighting functions, h, as row vectors instead of column vectors (that is how I interpret 2a-c, since there are no commas between the values), you should not have a transpose symbol on h_{XCO2} in the equation on this line. Equations 5a and 6a are correct as written.

L 316: Since the h's are row vectors, you want to indicate one row vector over another one here: to me, the comma between them doesn't achieve that. Though it is convoluted, $h = [h_{LMT}^T, h_U^T]^T$ would be correct. Or you could put the equation on more than one line, with h_{LMT} on top of h_u in matrix h.

L 321-322: The equations involving h need to be fixed here, too.

C5

L322-323: This is the first that we have heard of the noise vector having a dimension related to the dimension of the measured spectra – put that earlier after equation (1)?

L 338: If the h vectors are indeed row vectors as you defined them, then equation (7a) is correct, but the transpose symbols in (7b) are incorrect – you should have a transpose on the right-most h in each term and not on the left-most. That will get you to the 2x2 matrices you want to end up with.

L 342: Predicted errors are mentioned here and given in Table 1, but you have not yet explained how you predict these errors. Is this what is described later in Section 3.2?

L 347-350: Is this information that the prior uncertainties for LMT and U are 34 and 9 ppm given in the tables somewhere, or is this the only place you say it? If so, instead of saying "note that", you should say "We assumed a priori uncertainties of 34 and 9 ppm on LMT and U in the calculations that follow" or something like that, so that the reader knows and isn't looking for that information somewhere else.

L 354: Again, the transpose symbols should be switched here, too, if the h's really are row vectors.

L 367: It would be better to change "corr" to some symbol ("c"?) so that people don't think it is some subscript that didn't get subscripted properly...

L 377-378: I think you want these equations to apply on an element-by-element basis. As written, the equations don't mean anything, since the h's are all vectors. (i.e. you can't divide a vector by a vector) Re-write these on an element by element basis.

L 380: It would be good to show where 2.5 km would lie in terms of pressure on Figure 3, on average.

L 398: Section 3.2 in general: I think this could use a bit more explaining to help the reader understand what you are doing. If I understand correctly, you are assuming that the CO₂ profile measured by aircraft at SGP is the true CO₂ profile, then you plug this in as x_{true} in equations 5 and 6 to get the LMT and U that GOSAT would have seen

C6

there, by assuming the GOSAT averaging kernels and priors in the equations. And you assume the measurement error and interference terms are zero? Also, this approach gives what you would expect at the SGP site, not generally, correct? It would be good to explain this for the reader. Also, how high do the SGP profiles go, and what do you do to get the CO₂ profile above the top of the flight?

L 429: Could you please describe this flat prior some more: is it flat in space as well as time, or does it vary by latitude?

L 475: Section 4.2: You should explicitly state your key assumption here, namely that the aircraft measurements are unbiased and have a small measurement error compared to the errors in the GOSAT profiles.

L 489: It looks like you forgot to add a reference between the parentheses here – maybe an O'Dell bias correction reference?

L 506: The GOSAT and aircraft data are compared if they are within a week on either side? So clearly we are not getting synoptic scale variability – that is noise.

L 529-530: Since Figure 3 only gives pressures rather than altitudes on the vertical axis, the reader cannot, in fact, see the sensitivities that you mention, without pretty good knowledge of how altitude lines up with pressure. Perhaps you could add some rough altitudes on Figure 3 to help her/him out?

L 577: "represents the standard deviation of all the sites" – not clear what this means: the standard deviation of the site means, or the standard deviation of the measurements across all sites rolled together (without accounting for the number of measurements at each site)

Tables 2 & 3: "Ocean Surface" data are not aircraft – maybe change the title to reflect that? It would be a good idea in the caption to say what data is in each column, referring to the text if necessary for details.

Section 5.1: This description of how the biases and standard deviations across sites

C7

are calculated is quite confusing. I didn't see any description of how the 15 observations used to get the averages in Table 3 are selected, for example. Could you maybe give an example of how the different quantities are all calculated? Or else describe it better? Also, are these all GOSAT-minus-aircraft quantities being calculated? (I want to get the sign right on the biases.)

L 612: I read a value of -0.2 ppm on Table 2, as opposed to the -0.3 ppm given here in the text.

L 620-621: it is the variability in the bias, not the overall bias itself, that decreases, right?

L 624: I read 0.7 ppm on Table 2 for GOSAT LMT ocean, not 0.5 ppm as you say here.

L 624: I read 0.1 ppm on Table 2 for GOSAT U ocean, not 0.2 ppm as you say here.

L 625-627: "The LMT location dependent bias is no worse than the XCO₂ location dependent bias" I don't understand this statement: on Table 2, it looks like this co-location bias (on the top line) is several times larger for LMT than for U or XCO₂, over both land and ocean.

L 630-633: "The variability of the LMT ct_ct bias is 0.7 ppm, and when the 5 sites with ct_ct co-location error larger than 0.5 ppm are taken out, the GOSAT LMT bias variability drops to 0.7 ppm." Did you mean to put in a lower number for the second 0.7 ppm?

L 634-635: "Taking out the top 4 GOSAT bias outliers results in a GOSAT bias variability": which column type are you discussing here? Full XCO₂, or LMT, or U?

L 645-646: "The surface ocean has 1.0 ppm co-location error, also including the vertical co-location." I do not understand why the ESRL ocean comparison does not show biases at least this large, because it should reflect these same vertical bias errors that are measured using the CT co-location differences. Why is that?

C8

L 646-648: "The AJAX comparisons, which are primarily from GOSAT underflights, has a co-location error half that of the ESRL land matches, which have coincidence criteria of 7 days, and 3-5 degrees." It might be useful to discuss the implications of this. It suggests that the +/- 7 day colocation criterion is matching up GOSAT columns with aircraft columns at all different phases of synoptic scale variability over mid-latitude land regions. Thus most of this standard deviation is due to true variability in column CO₂ due to weather systems moving by, rather than to any sort of measurement issues. It would be interesting to do a sensitivity study looking at only the closest co-locations in time/space to see how much the mis-match drops in those cases (even if there are only a few scenes to look at).

L 662: Please state how the 15 scenes in each n=15 average were selected. Were they all at the same site? All contiguous in time?

L 696-698: "is approximated by the standard deviation of the CarbonTracker model at the validation location and time and the model at the satellite observation location and time" You mean to say that you calculate the standard deviation of the difference of the CarbonTracker values at the two different times/locations, correct? This could maybe be worded better to get that across.

L 738: Is the "location-dependent bias" the same thing as the correlated error, a_o ?

L 740-741: "In this paper, we find this variability to be 1.0 ppm." Where on Tables 2-4 does this 1.0 ppm number come from? Is it for XCO₂ or LMT or U?

L 776: "all error terms multiplied by approximately *0.6": It is not clear which error terms (predicted or calculated) you are multiplying by the extra factor, or why you chose 0.6 to multiply them by. What is the basis for this?

L 778: "Over ocean, the error correlation is the same, but the multiplication factor is " Text to end the sentence is missing here.

L 781: "one location": from the text, I thought that multiple locations were captured in

C9

Table 3. Why do you say that only one location was used? Do you mean to say that the validation location is fixed and the GOSAT data change around it, within the colocation criteria? But that there are more than one validation site?

L 797-803: It is not clear to me whether the co-location errors have been removed from the n=1 GOSAT errors in Table 3 or not. Maybe this could be mentioned here again (if it was already mentioned somewhere else) to remind the reader. It seems like the colocation errors ought to be removed from these numbers, since that error source is artificial.

L 824-826: "At the two sites where aircraft and TCCON are jointly observed, SGP in Oklahoma and LEF in Wisconsin, XCO₂ agrees with TCCON rather than the aircraft." This really cannot be seen in the figure.

Figure 7, caption: The order of the sites in the figures is described incorrectly in the first sentence. Reword to: "Seasonal cycle at 5 sites arranged from west to east (a-e) and north to south (f-j)" Reword the last sentence to: "The amplitude of the LMT prior is consistently too large for (i-j)..."

L 833-834: "There is also a shift to later in the seasonal cycle minimum going either west to east". From the figure, it seems this should read "east to west".

L 843: "...there could be a mismatch in air mass." To be clear, you are mainly worried about a mismatch in the vertical, right? That might be worth mentioning. You have mismatches in the horizontal, too, due to the +/- one week matching criterion, but these would be more of an issue over land than here over the ocean, well away from the continents.

L 849-852: "In Table 4, the correlated error for surface sites is higher than for ocean aircraft comparisons (1.0 ppm vs. 0.3 ppm, respectively), and the mean bias is also higher (0.7 ppm vs. 0.1 ppm, respectively)." I am unable to find these values in Table 4. The first pair of numbers (1.0 vs. 0.3) seems to be related to co-location errors, not

C10

correlation errors. I can't find the second pair anywhere.

Figure 8: What does "RET2" refer to?

L 891-893: "We look at the $\Delta\text{CO}/\Delta\text{CO}_2$ emission ratio in May and August to check the enhancements seen in LMT relative to MOPITT in these two months has a similar ratio is seen both months." This sentence is not clear. Please reword and correct the grammar.

L 888-916: This whole discussion of emissions ratios is very hard to follow. First, you ought to state somewhere in this section that you are comparing CO_2 from GOSAT to CO from MOPITT and to fire counts because you want to check whether fires in this part of the world might be responsible to the CO_2 patterns you are seeing (yes, perhaps obvious, but you ought to state this up front to introduce why you are looking at this other data – there are other processes that can cause variability in near-surface CO_2 over the tropical land that are not fire-related). Second, the emissions ratios you mention on lines 900-902 do not seem to correspond to anything in Table 5: why don't they? Why bother including a Table 5 if you don't seem to be talking about the same numbers? If the emissions ratios are calculated somehow from what you have in Table 5, you need to say how you calculate them.

L 926: "Measurements CO_2 and 4 km and 1 km are performed..." This is not clear – please reword.

L 939-946: This should be "Table 6a-c"

L 972: "in July for CAR, SGP, and SCA...". There is disagreement at the site in Nebraska as well that seems to be due to co-location error.

L 1338-1339: "As the bias correction for simulated OCO-2 data is very similar factor" – this wording is not clear – please correct.

L 1340: "the constraint" – it is not clear what constraint is being referred to.

C11

Table A1: It is not clear what the column headings "ocean bias correction" and "land bias correction" refer to: are these the coefficients on each term that are solved for? If so, some other wording would be better. If not, and if these are some sort of average bias correction across all the data points, then it would be good to know what the average value of each parameter in the fit (e.g. `albedo_2` or `co2_grad_delta`) was that corresponded to the bias correction.

Table A3a and A3b: in the captions of these tables, you should indicate which quantity (XCO_2 , LMT, or U) the bias that you are discussing pertains to.

Editorial comments:

L 28: replace "over" with "upon"?

L 36: say either "a CO/CO_2 emission factor" or " CO/CO_2 emission factors"

L 85: add a comma before "which"

L 89: replace "which" with "that"

L 107: Reword to "Measurements of CO_2 vertical profiles from aircraft, which extend from the surface to..."

L 109: Reword to "The second dataset that is used is CO_2 measurements from remote surface flask sites"

L 111: Reword to "...assuming that CO_2 mixing ratios in the lower 0-2.5 km are well mixed..."

Figure 2 caption, first line: Reword to: "...ESRL aircraft profiles (orange), which occur over both land..."

Figure 2 caption, third line: replace "campaign" with "data", to make this parallel with the other 4 entries.

L 120: rather than "shortwave", why not say "near-infrared" to be consistent with what

C12

you say elsewhere in the manuscript?

L 122: You should spell out what "ACOS" stands for.

L 122: capitalize "Lite"?

L 127: replace "sufficiently" with "sufficient"

L 129: replace with NOAA/ESRL. Also, you need to define what "ESRL" stands for. This would be a good place to do so. Maybe reword to "Aircraft and ocean measurements taken by NOAA's Earth System Research Laboratory (ESRL) are obtained from an ObsPack product..."

L 171-172: for clarity, I would suggest rewording "with additional screening for the new products" to ", with additional screening performed in our analysis here," (i.e. it wasn't clear to me what "new products" referred to)

L 185: change "which" to "that"

L 218: a second right paren should be added at the end of the line

L 226: Reword to "... contain information that allow each measurement to be split into 2 or 3 vertical columns"?

L 249: Add "to" after "chosen"?

L 277: change "contributes" to "contribution"

L 282: correct to: "...bottom level is half that of the other levels"

L 403: Add a "?" after 2nd question. Maybe add an "and" before the 3rd question.

L 440: "variation * airfraction" When I first looked at this, I couldn't figure out what the asterisk was for. Maybe write out "times" or take out the space on either side of the "*"?

L 501: fix: "the a"

C13

L 587-590: What is the implication that both co-location biases are -0.3 ppm then? That their total is -0.6 ppm and thus that the actual bias with respect to HIPPO and ESRL ocean data is actually now +0.3 ppm? Or that the the -0.3 ppm HIPPO and ESRL bias is actually removed entirely by the co-location bias? (i.e. are the two additive?)

L 605: "Appendix B, table b": I see a Table B1 – is this what you are referring to?

L 647: Change "has" to "have"

L 684-685: There are right parentheses missing in these two equations.

L 734: fix "XCO₂"

L 795: add "the idea" before "that"?

Interactive comment on Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-720, 2016.

C14