Atmospheric
Measurement
Techniques

Open Access

Discussions

AMTD

Interactive
comment

Interactive comment on "Comparison of CO_2 from NOAA Carbon Tracker reanalysis model and satellites over Africa" *by* Anteneh Getachew Mengistu and Gizaw Mengistu Tsidu

D. Baker (Referee)

david.f.baker@noaa.gov

Received and published: 10 August 2018

In this manuscript, CO_2 mixing ratio fields over the continent of Africa taken from the CarbonTracker-2016 (CT2016) CO_2 flux inversion system are compared to CO_2 measurements obtained from the GOSAT and OCO-2 satellites. Though it is not said explicitly, this CO_2 comparison is presumably done using column averages (commonly refered to as "X_{CO2}"), either using a straight pressure weighting or by sampling the CT model using the same vertical averaging kernel and prior CO_2 vertical profile that the satellite retrieval uses on the true atmosphere. The CT2016 system takes a set

of somewhat-realistic surface CO_2 fluxes (fossil fuel burning, land use change, wild-fires, photosynthesis and respiration from the land biosphere, air-sea fluxes) as a first guess, runs them forward through the TM5 off-line atmospheric tracer transport model to obtain CO_2 mixing ratios in the interior of the atmosphere, samples these at the time and place of surface in situ data, and uses the measurement differences to improve the initial flux estimates using a fixed-lag ensemble Kalman smoother as the inversion method. CT2016 thus can be thought of as a model of CO_2 in the atmosphere that has been forced to agree with surface CO_2 measurements, or alternatively as a glorified interpolation between the in situ measurements. The difference between CT2016 and the satellite data is mainly viewed here as a deficiency in the CT model as compared to the more accurate satellite data. Explanations for the differences are suggested in terms of likely flux errors in the CT2016 model.

Comparisons of this sort are of value, but in my view are better viewed from a different perspective. First, if one is interested in understanding the processes driving CO_2 fluxes in Africa (fires, photosynthesis, respiration from plants or soils, etc), one might do better to look at the results of an inversion of the satellite CO_2 data compared to the fluxes obtained from a system like CT2016 that inverts in situ CO_2 data. The reason for this is that the column averages compared here have much information in the upper part of the column that blows in from outside the bounds of Africa and reflect the effect of fluxes from around the globe; thus, column CO_2 differences compared over Africa may not reflect the impact of local fluxes, but also of far-field fluxes that have blown in over Africa. The inversion systems, which model atmospheric transport, are supposed to sort this out and apportion the fluxes to the right locations – they are the tool that I would use if understanding fluxes is the goal.

The direct comparisons of column CO_2 do have value, but one should not assume that the model is wrong and the satellite measurements right, as has been done here for the most part. The retrieval of CO_2 from satellite data is beset with a variety of challenges (scattering due to clouds and aerosols, instrument issues, unknown spec-

Interactive comment

troscopy, surface characterization issues, etc.) that often result in systematic errors that, at the moment, are usually removed in a separate step after the retrieval. A key component of this bias correction process is the comparison to independent data, such as column-averaged CO_2 measurements from the TCCON network, for example (using comparisons similar to those in this manuscript, but with the model being replaced by the TCCON data). However, comparison to models that have assimilated in situ data (like CT2016) have also been used in this bias correction step. It has been found that the systematic errors in the satellite retrievals are generally larger than those from an ensemble of models, so that the models may be used to help find errors in the satellite retrievals (instead of vice versa, as is done in this manuscript). Hopefully that situation will change as satellite retrieval schemes improve, but at the moment it is not a good assumption that most of the difference between model and satellite measurements can be attributed to model error. Really, there are errors in both model and measurement and one should quantify both with appropriate uncertainties. The authors do provide satellite CO_2 retrieval uncertainties here, but these do not capture the impact of systematic errors in the retrievals and are overly optimistic – one cannot assume that because they are small, the majority of the model-measurement difference must be due to model error.

In my view, the model-measurement comparisons done here are most valuable for understanding errors in the satellite retrievals. Africa is a continent with wide extremes in surface type (desert vs. rainforest vs savannah) and aerosol loading. These conditions have a strong impact on the satellite retrievals, so a study of this sort over Africa can tell much about how these systematic errors vary geographically. I would encourage the authors to consider whether their comparisons might be better viewed through that lens.

In general, the English punctuation and usage in this manuscript need extensive editing, which I have not attempted to do here. The authors need to clarify certain aspects of their method, most notably what CO_2 quantity they are plotting on their figures (pre-

sumably column-averaged CO_2) and how they do their vertical averaging. Measurement errors of 2 to 5 ppm are not small or reasonable – they are show-stoppers that prevent the satellite data from being useful for inferring surface CO_2 fluxes (since these measurement errors feed through to similar errors in the flux results). Some of the statistical metrics presented here are unconventional and are difficult to understand, and do not add much to the presentation – I suggest below that those sections of the document be removed.

Detailed comments follow below:

page 1, Line 1: "poor global coverage and resolution of satellite observations": Usually the satellites are thought to have good coverage over the globe, at least when compared to the in situ data.

Abstract: Here, the fidelity of the CarbonTracker model is being tested by comparing its a posteriori 3-D CO2 fields to CO_2 measured from satellites and by the TCCON network of ground-based sun-viewing spectrometers. That would make sense if the accuracy of those measurements was thought to be better than that of the model. But what if the reverse was true? Then it would make more sense to check the accuracy of the measurements by comparing them to the (more accurate) model. In the early days of satellite CO_2 data (and we are still in the early days, really) it was thought that that latter situation was actually the case, and models were used to correct the satellite data. Given that, some discussion of the assumptions underlying the comparison of CarbonTracker to the satellite and TCCON measurements would be useful in this paper before moving on to the comparison.

p1 L15: "relative accuracies of 1.22 and 1.95 ppm were found between the model and the two data sets"; these were judged to be "reasonably good". The in situ CO_2 data have relative accuracies an order of magnitude better than this (0.1 to 0.2 ppm) – why then should measurements an order of magnitude less accurate be considered "good"? Some guidelines concerning what is considered a good error versus what is considered

a bad error ought to be given to support this assertion. The density and coverage of a data type might also factor into this assessment, as well as the importance of random versus systematic errors.

p1 L17-18: "...probability of detection ranges from 0.6 to 1 and critical success index ranges from 0.4 to 1..." These statistics may not be familiar to the general reader. Maybe say "(see main text for definitions)" when using these in the abstract?

p1 L20-21: "GOSAT and OCO-2 X_{CO2} are lower than that of CT2016 by upto 4 ppm over North Africa (10 N -35 N) whereas it exceeds CT2016 X_{CO2} by 3 ppm over Equatorial Africa (10 S-10 N)."

It seems like the sign of this is wrong. Also, these satellite data are raw or bias-corrected?

p1 L25: "In these cases, the model overestimates the local emissions and underestimate CO_2 loss." Here, it is assumed that the satellite data are correct and the model is wrong. But what if the reverse was actually true? In reality, both model and satellite are wrong, to differing degrees that are quantified using the metric of uncertainty. You should give reasons why you think the uncertainty on the satellite measurements is lower than the uncertainty on the modeled CO_2, to defend why you think the model is wrong and the satellites right. The satellites have retrieval biases and random errors that could easily be larger than the modeling errors in CO_2 (errors due to transport, errors in the in situ data, inversion assumptions, etc).

p2 L25-28: In introducing the TCCON data, you should note that the TCCON network measures a column-integrated CO_2 mixing ratio; yes, it is "ground-based", since the sensor sits on the ground, but the measurement is a column average (unlike the ground-based in situ data) and that is why it is useful for validating the satellite data, which are also column averages.

p3 L3: "Other studies have revealed that significant improvement in estimation of

weekly and monthly CO_2 fluxes can be achieved subject to CO_2 retrieval error of less than 4 ppm" This may be true if the errors are purely random – with enough low-precision data of this sort, the errors may average down to something lower and more useful. However, if this is accompanied by systematic errors of similar magnitude, these will not average out. Therefore some discussion of random versus systematic error must be added when giving out these error quantities.

p3 L21: "V02.XX" – you need to complete the version numbers for two products on this line

p4 L3-4: "These findings suggest that it is important to assess the accuracy and uncertainty of X_{CO2} from models with respect to observations" Again, the authors point to the differences between satellite data and models, then leap to the assumption that the models are responsible for most of those differences when the measurements might actually be more responsible. Some discussion of which is more error-prone (model or measurement) is needed.

p4 L27: "CT2016 and CT2017 respectively": you should not refer to the 2016 near-real-time CT as "CT2017", as that term is reserved for the subsequent release of the full data span (the release that would have been put out in 2017 if the releases were put out when indicated; the release that finishes at the end of 2016). The near-real-time releases are different from the standard releases in that they bring out a subset of the full data set, without the usual quality assurances, for early use by modelers. Even if that is only a label for use in this manuscript, it will confuse people.

p5 L1: Which version of ACOS GOSAT retrievals did you use in this study?

p5 L2-3: "GOSAT is the world's first spacecraft to measure the concentrations of carbon dioxide and methane, the two major greenhouse gases, from space." Not true: many satellites/instruments measured both in the thermal IR. Also SCIAMACHY, which came out much earlier, measured both in the near-IR. What you could say is that GOSAT was the first spacecraft dedicated solely to measuring CO_2 and CH_4.

p5 L15: When discussing the OCO-2 data, you must say what version of retrieval you are using: version 7? Also, you should indicate whether you are looking at raw X_{CO2} values or bias-corrected values. If bias corrected, you should indicate whether you used the additional "s31" bias correction, designed to correct albedo-related errors not caught in the initial bias correction. This extra bias correction could be particularly important over Arica, with its large differences between desert and tropical rainforest. Also, you should discuss whether you sample your model with the OCO-2 averaging kernel and prior CO_2 profile when comparing to the OCO-2 retrievals.

p5 L21: (Methods): the satellites estimate both a profile of CO_2 on 20 levels, as well as a vertical weighted average, "X_{CO2}", computed from these. Which do you compare to? If the latter, you should discuss the vertical weighting kernel for X_{CO2} and how you sample your model to account for this. This methods section seems to be the correct place to discuss this.

p6 L19: "Fig. 1 shows the five-years average of CT2016 (Fig. 1a) and GOSAT (Fig. 1b) X_{CO2} distribution." It would have been more useful to do the 5-year average over a true 5-year span, rather than the "April 2009 to June 2014" span that you used. As it stands, you have an extra 3 month span (April to June) that is unbalanced by the remainder of an annual cycle. As a result, this three month span throws off what otherwise would be the average of five full annual cycles. If flux is positive north of the equator during April to June (which it probably is, this being right at the end of the burn season there) then you tend to see positive values there in Figure 1a and 1b. This is due to the seasonality of the flux rather than the annual mean value. It might be easier to discuss annual mean fluxes in the text if these plots reflected averages over pure annual cycles.

p7 L3: "respiration from forest in the region is overestimated" Why do you think that this is not due to photosynthesis being underestimated?

p7 last line: "standard deviation of 1.31 ppm indicating better consistency and less potential outliers." Better than what?

p8 L3: change "require" to "permit" Also, if the difference can be up to 1.5 degrees on either side, then we are talking about a box of 3 degrees on a side in both latitude and longitude. On p5 L29 you should then change the box from being 1.5 degrees long and wide to 3 degrees long and wide.

Fig 3 caption: please provide units (ppm)

p8 L19: A correlation should be unitless – remove "ppm"

Figure 3: It is not clear to me why Figures 3a and 1c are different – they are both showing the difference (or bias) between CT2016 and GOSAT on an annual mean basis. So why aren't they identical?

p8 L20: Good – here you are examining whether some of the difference might be due to the satellite data rather than the model.

Section 3.2: The statistics presented here are abstruse and require the reader to go back to the other references not only for what the statistics mean, but even what the acronyms stand for. It is not clear what the message is that these statistics convey. The only point I gleaned from the discussion was that the annual mean bias between model and data is higher where there are fewer data to average over – that is a point that could be made with Figure 1 alone. I would suggest deleting this section altogether.

p13 L9, p14 L11-14: Here you are suggesting that local fluxes are responsible for the changes in CO_2 seen over Africa, but it is quite possible that a large part of these changes are due to fluxes elsewhere, over different continents, blowing downwind over Africa. Without inversion results linking concentrations to the fluxes that caused them, your argument will remain weak using this approach.

p25 L16: "Satellite observations are able to capture the CO_2 concentration truly and objectively." I don't understand how you can say this based on the results that you have shown. You have shown large differences between CarbonTracker and the satel-lite data, on the order of 5 ppm on a seasonal and regional basis. This could be due

to satellite retrieval problems in addition to the model problems that have been empha-
sized here. You haven't shown any data or comparisons that suggest the satellite data
are better than the model, really. Maybe you could compare GOSAT and OCO-2 across
the short time period that they overlap (late 2014, 2015) and if they agree closely then
suggest they are both getting the right answer. From the plots you have showed in this
manuscript, however, it looks like there are significant differences between OCO-2 and
GOSAT... suggesting that the satellite errors are significant.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-84, 2018.