Review of "Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements"

by Wu et al.
Atmos. Meas. Tech. Disc., amt-2017-415

This paper is describes the relatively straightforward application of the RemoTeC retrieval algorithm to data from the Orbiting Carbon Observatory-2 (OCO-2). RemoTeC was written to retrieve CO_2 and CH_4 column-average concentrations from the Greenhouse Gases Observing Satellite (GOSAT), and has been improved and validated over the years, as described in a number of publications. The authors have applied this mature algorithm to OCO-2 data to retrieve CO_2 (OCO-2 does not have the CH_4 band that GOSAT possesses), and find that after a few slight modifications, the error statistics of OCO-2 retrievals vs. ground truth data compare favorably both to the operational OCO-2 product as well as to RemoTeC retrievals of CO_2 from GOSAT.

General Comments

The paper is useful in that it shows that the RemoTeC algorithm can be successfully applied to OCO-2, though it is relatively dry and offers few new physical insights into sources of error/bias in the OCO-2 measurements. However, is it worthwhile piece of work, and I recommend publication after making some minor revisions.

My only main comment on the paper has to do with the filtering and bias correction, for which the bottom-line recipes are given. Some more information would be welcome. For instance, what other parameters were investigated for bias correction or filtering, such as the 1/(size parameter) variable used in GOSAT bias correction (Guerlet et al, 2013b)? Was the ω_s parameter of Guerlet et al. (2013b) found not to be useful for OCO-2, even though it was for GOSAT? A figure similar to that of Figure 11 in Guerlet et al. (2013b) would be very useful here to see how similar/different GOSAT vs. OCO-2 retrieval biases are. Also, how stringent were your filters overall – did you filter out 10% L2-processed soundings, 50%, etc? How was this different over land and ocean? A throughput map would be useful.

Specific Comments (P=page, L=line)
P2, L24: "XCO2 retrievals with this level of accuracy [<1%] can provide valuable information on…sources and sinks…" No! 1% = 4 ppm. We know that regional biases even 1 ppm (0.25%) in XCO2 are too large (Chevallier et al, 2014). Please modify or remove this statement.

P3, L2: "by aerosols and cirrus." Do water clouds not have any effect on scattering? I suggest changing this statement to "by aerosols and clouds."

P3, 1st paragraph. The authors describe a number of XCO2 retrieval algorithms but this list is certainly not exhaustive. There are the BESD and FOCAL retrievals from

M. Reuter, the TanSAT retrieval from D. Yang, and various versions of the PPDF retrieval of Oshchepkov and Bril. You should either cite these or make clear that you are not exhaustively listing all available retrievals.

P4, L22: "barometric law". Do you not mean the hypsometric equation (which combines the ideal gas law with the hydrostatic equation). They may be equivalent, I'm not entirely sure. But usually in this context, it is referred to as the hypsometric equation.

P4, L23: Are your priors adjusted for the secular growth rate of CO2 (since you just say you use CT from 2013)? Seems like you should, or you could probably introduce an artificial trend in your retrievals.

P5, L20: Does $\mathbf{S_y}$ include any estimate of forward model error, as you previously implied it might (P5, L2). Similar are the noise estimates taken from the OCO-2 suggested formulation, or do you calculate your own noise estimates somehow?

P6, L9-10: Please discuss whether the per-band radiance offsets were needed for GOSAT. My understanding is that they were needed for band 1, but not the other two bands. You could instead bring this up in section 4.3 as well, but I think it's important to contrast this need for the offsets in OCO-2 vs. that of GOSAT. For instance – I was thinking that maybe you needed them because your retrieval doesn't explicitly retrieve cirrus, so it would be difficult to retrieval soundings with a cirrus layer overlying a thin aerosol layer, which is a pretty common situation. Retrieval the 3 per-band offsets would be a pretty easy way to fake it. But that would likely then also be needed for GOSAT. Some discussion on this would be useful.

P7, L14: χ_2 is the symbol usually referred to as the total chi-squared. What you show is much more similar to the "reduced chi-squared", which is the total chi-squared divided by the number of degrees of freedom (# channels - # retrieved parameters). You really are giving the mean chi-squared per channel. You should make this clear, and that a value around unity would indicate a fit that is in line with the noise. Values consistently higher than unity mean there are the systematic errors in the forward model that are not able to be fitted away.

P8, top: In the discussion of using the SWIR-1 chi-squared as a bias correction parameter, it would be nice to lengthen this discussion. Does SWIR-2 chi-squared perform similarly? Other parameters? Mention that r=0.2 means that 0.04 percent of the variance is explained (or it will reduce the standard deviation by about 2%). Why do you include the offset "d" parameter when you already include a global bias correction? They would be directly related to each other. Does this multiplicative formula (equation 4) work better than an additive equation? Finally, it would be valuable if you could speculate on *why* this parameter seems to be correlated with the bias over land. And perhaps on why it is NOT correlated with the bias over water. Are the chi-squared values much lower over water? Finally, what is the

spatial distribution of this parameter? Is it highly scattered or does it seem to remove coherent regional biases?

P9, L10: Some comments on why the effect of the bias correction is largest for those 3 stations would be welcome. It seems like it should be substantial for all over-land stations, unless the chi-squared values were just worse for those stations. My guess is that your chi-squared is going to be correlated with SNR or surface albedo, and brighter surfaces will have larger corrections. If you plotted the mean correction on a map, this would probably become obvious.

Section 4.3 As mentioned before, contrast with the offset approach for GOSAT. Is the behavior of the fitted radiance offsets similar over land and ocean? How correlated are the fitted offsets for the 3 bands? (either in absolute terms, or relative to the mean radiance in their respective bands) If they are highly correlated, or not, that would give you a clue what they are correcting for (either cirrus, as I hypothesized earlier, or some instrument effect that is particular to OCO-2, and perhaps not GOSAT).

P11, top: Are the GOSAT vs. OCO-2 error statistics vs. TCCON similar for both land and ocean soundings?

Figures: in many of the figures, the font sizes make reading some of the text difficult (axis labels, bias numbers , TCCON site names, etc). Please try to make them bigger to increase legibility.

Technical comments
P4, various: spectral samplings → spectral samples
P5, L10: "radiative transfer model Hasekamp…" → "radiative transfer model (Hasekamp…"
P10, L8: proportional mis-spelled