Review of "XCO2 retrieval for GOSAT and GOSAT-2 based on the FOCAL algorithm" by Noel et al.

This paper describes applying the FOCAL retrieval of Max Reuter et al. (2017), originally developed for OCO-2, to GOSAT and GOSAT-2 data products. This is novel in and of itself, but particularly so for its application to GOSAT-2. The GOSAT XCO2 retrievals show good statistics when compared to TCCON as well as other established GOSAT XCO2 retrieval algorithms. Other gases are also retrieved (e.g., CH4) but are not discussed in this work. The FOCAL algorithm is particularly fast and therefore may serve as a candidate retrieval algorithm for the upcoming European CO2M mission.

General Comments

Overall, I found this paper useful and interesting, and will serve as an important reference. The subject matter is important, the layout of the paper is logical, the reasoning sound, and the results are generally laid out well. However, there are a number of problems that need to be addressed. While details of the retrieval, filtering, and bias correction were presented in a straightforward way, it was quite dry with little learned. Especially in the part about the random forest filter, which was used for both filtering and bias correction, but with little attempt on the part of the authors to explain the relevance of the features identified. The same goes for the prefilters, where it appeared that thresholds were drawn somewhat out of thin air for some of the parameters. It would have been useful if the authors had shown even a couple example plots of some of the prefilters and how thresholds were determined.

There were 25 figures in this paper, and in my opinion, many more than are useful, especially some of the earlier plots. I suggest the authors try to remove some panels in some plots, or some plots altogether, to show *representative* plots. For instance, all the noise model coefficients are given in Tables 6 & 7. Therefore, the authors can reduce Figs 3-6 to probably a single 2 or 4 panel plot (e.g., Fit Windows 2 & 3 for both GOSAT and GOSAT-2, P-polarization only). The same goes for Figs 9-12 (a single one would do) and Figs 13-16 (again, a single one would do, and not all bands are necessary). Plots are in the paper to explain findings, not to exhaustively present ever detail of the study, especially if some plots or features of plots are never discussed in the main body of the paper.

Finally, it appeared that many important previous works by other authors are never referenced, or included in the reference section but never cited in the main body. In general, referencing needs to be much improved in this work. Therefore, I recommend publication of this manuscript after a major revision to fix the issues with the burdensome # of plots and problems with referencing, as well as addressing all the specific concerns raised below.

Specific Comments

Section 2.3: This is a unique approach to a truth database to my knowledge – it needs more information (plots, etc) on how big this contiguous regions are / how much the TCCON data are expanded through this approach. A map of a month or a season of data density would fulfill

this, and I think be very interesting for readers. Otherwise, it's not clear how much this really expands over just using TCCON directly.

Section 2.3: Secondly, you say the requirement for contiguous regions, but you never say how close the ak-corrected CT value at the TCCON location & time has to agree with TCCON itself. Is that also 0.75 ppm? You imply this but never say – please correct this.

Section 3.1: Your terms "cloud albedo" and "water vapor path" are neither. These terms already have definitions in use by the community, and they are not how you define them. I suggest you rename "cloud albedo" to "effective albedo" or "effective scene albedo". Note you *will* screen out some bright desert scenes with your albedo filter, though probably not many. It looks like your 1.98 μm filter is doing most of the work. Regarding "water vapour path", it's nothing of the sort. It's more like an SNR_{wv} (wv="water vapour"), or $SNR_{1.93}$ (since this band is roughly at 1.93 μm). Low SNR_{wv}= clear, high SNR_{wv} = cirrus present. So please rename it to something else.

Section 3.2 – Please MOTIVATE why you use both polarisations separately. Do you believe you obtain more information than if you averaged them together, or do you believe you cannot accurately average because certain instrument properties (such as ILS) are different for the two polarisations, and they themselves cannot be averaged together?

Section 3.2.1 Near line 263, you talk about the "NIR", but early in the paper you refer to ALL the bands you use as "SWIR". I realize most scientists label the O2A band as NIR and everything past 1 micron as SWIR. Can you please go through the paper and ensure consistency between NIR and SWIR labels throughout?

Section 3.2.1 – Way too many plots, as I said in the general comments. As a rule of thumb, try not to overwhelm readers with a bunch of plots that all look essentially the same. Each panel of each plot should contribute to the story you are telling.

Section 3.3.1 – In general, your "basic filter" through the RSR filters (I'm looking at your figures 1-2 for this information) really does seem basic for GOSAT, as it filters out only 8 percent of the data (35.0%-->27.2%), and most of that comes from convergence. However, for GOSAT-2 not only do twice as many soundings fail to converge as for GOSAT, but the window 5 RSR also accounts for many failed soundings (5% for GOSAT-2, versus 0.3% for GOSAT, if I am counting right). Can you please comment on why this may be happening for GOSAT-2? Window 5 is the methane band I think. You may wish to split things out separately as land versus ocean – you may find very different behaviors for the two categories. In any event, please devote a few words in this section as to why this is happening. And please do say how differently the filters act on land vs. ocean.

Actually, looking at this further, I think it is the "broadband oscillation" in the fit residuals you mention for GOSAT-2 that may be causing the problem. Are those oscillations really correlated

with retrieved XCO2 quality? If not, you may wish to loosen that constraint for GOSAT-2, to save more soundings.

Secton 3.3.2 Near line 295, please also reference Mandrake et al (2013, AMT, "Semi-autonomous sounding selection for OCO-2), who did something similar for OCO-2.

Section 3.3.2, near line 310. I'm nearly certain that for water, SAA, VAA, SZA, VZA will be correlated with latitude. Because the orbit is sun-synchronous and you're looking to the glint spot over water, I'm willing to bet that any machine learning algorithm or even a simple correlation analysis can probably figure out where you are based on those quantities (or even only one or two of them). I suggest you be exceedingly careful in including those quantities. Please include a comment to this affect in the paper.

Section 3.3.2 – can you state how many training soundings total there were for GOSAT and GOSAT-2, for each of land and water? I wonder if your training set is general enough to avoid over-fitting. Also, please define "Relevance" as you use it in Figure 7 & 8.

Section 3.3.3 – This community did XCO2 bias correction long before OCO-2. Can you please reference earlier works on the subject? (The earliest I know of is Wunch et al., 2011, ACP "A method for evaluating bias…"; I believe there are similar references for GOSAT for the UoL retrieval, the NIES retrieval, and the RemoTeC retrieval).

Are you really using 10 parameters in your bias correction? This is way more than most groups usually use (which is typically 1-4; as I remember, Reuter et al. (2017) didn't use any in their OCO-2/FOCAL paper). Be careful – there could almost certainly be overfitting here. So my comment is 10 parameters simply doesn't seem to be justified based on past experience and the published methods of nearly all other retrievals for the last 10 years. Therefore, your using 10 parameters requires more justification than simply "this is what came out of the random forest algorithm".

Section 4, nearly line 378. Just a comment. The higher XCO2 variability over land has long been seen. I highly doubt this is due solely to surface variability. I think it is also caused by different scattering pathways that are not present over water. In particular, photons scattered downward by the atmosphere can be reflected off the surface back into the beam accepted by the sensor; this mechanism doesn't happen over water, so there are more ways for atmospheric scattering to degrade a retrieval. But that's mainly just a hypothesis.

Section 5 – page 13. Please include appropriate references for each of these algorithms here. Also you say for the validation of the GOSAT and GOSAT-2 FOCAL products, but really these comparisons to the other products are just for GOSAT only. You may wish to state upfront here that the vast majority of the presented validation is only for GOSAT. Only subsection 5.3 mentions GOSAT-2, and it only appears in a single validation figure (25). In fact, this paper is really begging for some basic comparison plots of GOSAT and GOSAT-2 to TCCON, to see how

well your algorithm works on GOSAT-2 as compared to GOSAT. Can you please add something to that effect?

Line 438 – I do not understand this statement about a "bias anomaly". Please be more clear about what you did here. Did you subtract some kind of mean bias with respect to TCCON from each algorithm? Please don't! Or if you did, you have to state somewhere what number you subtracted off each algorithm. If ACOS is high by 1 ppm relative to TCCON and you simply subtracted that off before making plots, it's critical to state that somewhere. It would be much better simply to NOT subtract off that bias, unless you can throughly justify why you did.

Section 5.2 end, L445 – Even if you don't have "sufficient data" for full seasonal cycle fits for all GOSAT-2 data vs. TCCON, you've got enough to make some basic plots. Please do so – the community is really interested in them. If not, there isn't a lot of point in including GOSAT-2 in this paper at all.

References: It looks like you have way more references in the Reference section of the paper, than you actually reference in the main body of the paper. A rule of papers: you MUST cite each reference in your references section somewhere in the main body of the paper. Please make sure this is the case.

Technical/Grammatical Comments

L47: Tansat, GOSAT, and OCO-2/3 instruments → The Tansat, GOSAT, and OCO-2/3 instruments

L280: XCO2 error is ambiguous. Suggest you change this to "XCO2 posterior uncertainty" or something more clear that it is the posterior error estimate from the OE itself, and not some error as compared to TCCON or something.

L292: Remove the word "exemplary". This isn't really an example, I assume this is a full indication of what is happening.