Interactive comment on “Improved Retrievals of Carbon Dioxide from the Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm” by Christopher W. O’Dell et al.

J. Marshall (Referee)
marshall@bgc-jena.mpg.de

Received and published: 26 September 2018

This paper lays out the substantial work that has gone into the ACOS retrieval algorithm over the last several years, focussing on the adaptations in the version 7 and 8 builds tailored to OCO-2 data. It is quite an expansive paper as it contains information about several generations of the algorithm as applied to two different instruments, and marks the first peer-reviewed update of the full retrieval since 2012. The authors have done a good job of coherently structuring this information however, making it quite easy to follow. The manuscript is well-written and clear, and will provide a much-needed reference for the community of users already actively exploiting the resultant XCO₂ data.
The paper is optimally suited for publication in AMT, and is appropriate for publication once the authors have addressed a very few minor comments.

The main scientific questions that I was left with after reading the paper were related to the surface pressure differences. This difference between the retrieved and prior surface pressure leads to the dominant term in the empirical bias correction, and has quite a lot of structure, as is seen in Figure 6. The estimated uncertainty on the prior surface pressure was first decreased from 4 hPa to 1 hPa in B3.3, then increased to 2hPa in GOSAT B7.3, and then again to 4 hPa in OCO-2 B7. The discussion on P41 goes into some more detail on this, discussing some hypotheses about what the source of this difference could be, such as errors in the (temperature dependence of the) oxygen absorption cross section.

This plausible explanation, combined with the fact that the option of tightening the constraint on the prior is discussed, suggests that you are relatively certain that the 4 hPa uncertainty in the prior surface pressure is artificially exaggerated to allow for flexibility to account for this undefined bias source. This is implied in the discussion on P20, which suggests that a 1 hPa uncertainty is likely more realistic for the majority of scenes. This is again obliquely discussed at the end of Section 3, when discussing the differences between the surface pressure between the ECMWF and GEOS5 prior, which show only a difference of about 0.6 hPa, but this could also result from these two models being similarly biased and/or assimilating the same data.

It would be good to see the overinflation of the surface pressure prior uncertainty explicitly stated in the P41 discussion, along with a "best guess" estimate of what the true uncertainty in the prior surface pressure is. At present the information is all there, but the reader has to collect the information from several locations and piece it together. This point is likely relevant for other missions and retrievals as well, and can help users better interpret the data. Some further discussion into how you plan to tackle this identified problem, with or without improved O\textsubscript{2} spectroscopy, would be a welcome addition to Section 6.
Typos/style comments:
P2, L26: missing reference
P3, L3: as compared to: I would suggest changing this to "based on comparisons to" or even "when compared to".
P5, L10: missing reference
P5, L12: close parentheses after 3.1
P7, L3: a minor semantic point, but I would suggest "traits" rather than "behaviors"
P10, L11: You should define what ATBD stands for, or just use "as described in Crisp et al. (2010)."

Figure 3: The resolution could be better, will likely get picked up during editing.
P12, L8: evidence of similar → evidence of a similar

Figure 5 caption: The reference to the "operational retrieval" is confusing - which version is meant? I assume V7, but this should be clear.
P13, L32: Out → Our

There was some patchiness and inconsistency in the writing in the section describing ABSCO. Sometimes they were referred to as v5.0 or v4.0 (e.g. P14, L26; P15, L7), but usually just as 4.0, 4.2, or 5.0. On P15, L6, the second "the" should be removed. The paragraph starting at line 6 on P15 should perhaps be reworked, with the information incorporated in the two preceding paragraphs, which cover much of the same terrain (e.g. spectra at multiple temperatures in v5.0). Or perhaps I'm confused by this paragraph in general: "Because these spectra currently do not enable evaluation of intensities at accuracies greater than around 1

P17, L25: becomes → because
P20, L5: I would change "mean" to "the mean"
P20, L25: add comma after "correction"

P21, L13: missing citation

P21, L29: remove first "and"

P23, L1: Wasn’t the potential temperature colocation scheme developed by Keppel-Aleks et al. (2011)? It doesn’t appear in Wunch et al. (2011), or at least not in the Wunch paper you’re citing there.

P23, L11: remove the first "were"

P23, table: missing reference for Sodankyla

P25, caption: $4\times4^\circ \rightarrow 4^\circ \times 4^\circ$

P26, L9: averaging-kernel corrected $\rightarrow$ averaging-kernel-corrected

P27, L6: data was $\rightarrow$ data were

P27, L31: I guess that the P in IDP already contains preprocessor. I had to go look it up again to be sure though, which raises the question: is it really worth defining this TLA when it’s only used twice in the main text? It could be defined separately in the caption of Table A1.

P27, L33: significant scattering present: I found this a bit awkward. Perhaps remove "present", or instead refer to "the presence of significant scattering in the atmosphere".

P28: In the caption it says the quality flags are applied cumulatively from top to bottom and left to right. Based on the fraction that passes each flag it appears left to right happens first, and then top to bottom. Please correct/clarify.

P30, L8-10: I’m a bit confused about how this pointing error in the instrument will be corrected in the next version. The next version of the instrument? Or the next version of the algorithm? If it’s an instrument error, surely the algorithm will just be better taking it into account in the next version. Please clarify.
P30, L17: reason for this variable to be: Maybe better as "reason why this variable is"?

P30, L32: as has been noted (Butz et al., 2013). → as has been noted by Butz et al. (2013).

P33, caption: circles shown mean → circles show mean

P34, L14: is stated → as stated

P35, L16: line → in line

P38, Figure 17: There seems to be a large interhemispheric gradient associated with the correction to the prior, on the order of 3 ppm (in panel f). Could more information about the prior be given, besides the fact that it is the same as that used by TCCON?

P39, L15: "ocean was slightly lower than land": Please restate with all the necessary (semi-implicit) information, e.g. "the ocean values were slightly lower..." or similar.

P39, L15: yieled → yielded

P40, caption: "Overpass-mean level validation": I found this a bit confusing. What is "level" here?

P43, L9: over-pass → overpass

P44, L1: data that is → data that are