Replies to comments made by reviewers #1 and #2 for "The GeoCarb greenhouse gas retrieval algorithm: Simulations and sensitivity to sources of uncertainty" (amt-2023-17).

Reviwer #1

The authors of this manuscript would like extend our gratitude to Reviewer #1 for the helpful comments which we hope have all been addressed to the reviewer's satisfaction.

Reviewer's summary

This paper describes the algorithm for retrievals of column averaged concentrations of CO2 (XCO2), CH4 (XCH4), and CO (XCO) from GeoCarb. The authors provided an overview of the GeoCarb instrument and some example measurement (scanning) modes. They also described the level 2 retrieval system that is largely based on the ACOS algorithm for OCO-2/3 XCO2 retrievals, including pre-processors for aerosol/cloud filtering, a full physics optimal estimation algorithm, and a post-processing step for data filtering and bias correction. Through several sensitivity tests using simulated radiance data, the authors demonstrated that the anticipated GeoCarb retrieval errors could meet the mission precision requirements for all three trace gases, although some potential uncertainty sources (e.g., stray light) were not considered in those tests. Based on the test results, the authors suggested that errors in XCO2 and XCH4 retrievals were dominated by systematic errors, whereas noise-driven errors were more important for XCO. Overall, this is a comprehensive simulation study on the retrieval performance of a (now unfortunately canceled) future greenhouse gas satellite mission. The paper is well organized, the discussion is detailed, and the retrieval methodology is well established in previous studies. The topic should be of interest to the readers of Atmos. Meas. Tech. I would recommend that the paper be accepted for publications after mostly minor revisions. However, I do feel that the presentation quality of the paper needs to be improved âthere are many typos and grammatical errors that can be corrected if the authors can carefully proofread the manuscript.

Reply: We have addressed the typos and errors pointed out by the reviewer and those pointed out by the second reviewer and have performed a careful proofreading of the manuscript.

Specific comments from the reviewer along with author replies:

Abstract: perhaps add the caveat that several other potential uncertainty sources are not included in the present study.

Reply: A sentence was added to the abstract to address this.

Page 2, line 1: please elaborate what you mean by âsecondary changes to the Earthâs climateâ.

Reply: This is covered in the IPCC, 2021, cited in the text but we added a generic elaboration: "... resulting in secondary changes to the Earth's climate including changes in weather and surface processes """

Page 2, line 9: how do you define the accuracy of spatial and temporal scales?

Reply: We think that it is implied in the sentence with "enough to resolve sources and sinks". Especially with the

word "resolve". But we changed "scales" to "resolutions" in an attempt to clarify.

Page 3, Line 14: one may argue that a constellation of LEO satellites can provide some temporal resolution â I wouldnât say it is ânot possible with polar orbitersâ.

Reply: Changed "characteristics in a manner that is not possible with polar orbiters." -> 'characteristics that can be more difficult with polar orbiters.'

Page 4, line 30: the precision, as defined in the mission requirements, appears to be different from the simulation results shown later in the paper. Can the authors comment on the differences in the definition and how that may affect the interpretation of the results?

Reply: We are not exactly clear as to what reviewer is referring to. The precision requirements in table 1 are the requirements defined for the mission. In the results we show that those precision requirements are met for all experiments after filtering.

Figure 1: please check the unit for radiances.

Reply: The units have been fixed. Added "/sr"

Page 12, line 8: some of the spectral ranges listed here are less than twice the FWHM? Is there concern that this may lead to some inaccuracies in the convolved spectrum?

Reply: The units were wrong. Changed "um" -> "cm^-1"

Section 3.3: can the authors comment on the a priori profiles for relatively short-lived CO (and to a lesser extent, CH4)? Will the a prioir profiles adequately represent those over major source regions such as major cities?

Reply: The reviewers believe that this kind of discussion is out of the scope of this paper. Details of the trace gas prior profiles are in Laughner et al. (2023).

Table 4, for SIF mean, the a priori is 0 in the table, but Page 15, line 7 states that the priori SIF mean is from GASBAG. Please clarify. Also please check for consistency between the text and the table.

Reply: We thank the reviewer for noticing this inconsistency. In fact, for this study, the prior SIF was set to zero as in the table. We have made the appropriate changes to the text to reflect that.

Section 3.5: does the A-band preprocessor algorithm produce SIF retrievals for OCO missions? If thatâs the case, what is the reason behind using GASBAG for official SIF retrievals for GeoCarb?

Reply: The A-band preprocessor does not produce SIF retrievals for OCO missions. It simply does what is described in the paper. The reviewer maybe thinking of the IMAP-DOAS preprocessor used for the OCO missions (Frankenberg et al. 2012a, Frankenberg et al., 2012a). GASBAG is a complete replacement for SIF and single band retrievals (Somkuti et al., 2021) which is significantly more flexible and was planned to be used instead of IMAP-DOAS for the GeoCarb mission.

Page 17, lines 21-24: this sentence is hard to understand.

Reply: We admit that there was a typo in that sentence ("greater that 5\%" should be "greater than 5\%") and, even a missing period in the next sentence! We believe it is clear now and hopefully the typo fixes resolves the reviewer's comment.

Section 4.1: can the authors comment on the impact of different SSPs (anticipated vs. simulated) on retrieval throughput and retrieval errors?

Reply: Different SSPs will affect solar and sensor geometry which have affects on retrieval results by ground location. The currently chosen SSP was the optimal choice for GeoCarb at the time. Of course sliding eastward or westward would have negative affects on retrievals in the west and east, respectively. Ultimately we think the optimal choice of 87 degrees west was the best for the study as unpredictable changes in the exact SSP were expected to happen.

Eq. 16: please check for accuracy.

Reply: The missing cosine function on relative azimuth has been added.

Page 20, line 14: figure 4 shows AMF greater than 4.2?

Reply: This is intentional to show values outside the AMF filter threshold of 4.2.

Table 6: the table states that Runs # 10 and 11 have all perturbations in 4-9 but page 26 states that target shift (Run #7) is excluded in the kitchen sink tests. Please clarify.

Reply: We thank the reviewer for noticing this inconsistency. The table has been fixed.

Page 25, line 19: is âLS scaling factorâ the same as âILS stretchâ in Table 4?

Reply: Yes. We changed all references to "ILS stretch" to "ILS scaling factor" to be clear except when we are differentiating between a positive scale (a stretch) or a negative scale (a squash).

Eq. 19: please check for accuracy.

Reply: We have checked and believe the equation is correct.

Page 27, line 1: is Xgas the retrieved column or profile? Please clarify.

Reply: Xgas is the retrieved column. This has been clarified.

Page 27, line 2: the forward model for the simulator uses 61 levels in the atmosphere but here the interpolation is to 72 levels â please clarify.

Reply: This was a oversight by the main author for which I apologize. In fact, truth is provided at 25 levels. What happens is that the scene generation code takes as input meteorology at the 91 ECMWF levels, outputs that meteorology to the so called Met file used for L2FP, ABP, and GASBAG, and constructs 25 level meteorology, along with 25 level trace gas, aerosol, and cloud profiles, and outputs those as the scene input to the L1B simulator on which RT is performed. The text has been modified to reflect this.

Figure 5: it is a bit confusing â from page 23 Iâm under the impression that there are ~33000 soundings. Yet the number of raw soundings is far smaller at ~12000 for the aerosol/cloud free case. Does this mean that the pre-filtering is excluding a large number of aerosol/cloud free scenes as well?

Reply: The total number of soundings that GeoCarb would nominally obtain per scan block is what is listed in table 5. Then the paragraph at (page 20, line 16) describes the down sampling which results in 55248 soundings from all scan blocks and all seasons. Then after filtering out soundings over ocean 33111 soundings remain. Finally, the preprocessing results in 12520 soundings remain.

Figure 7: can the authors comment more on the spatial pattern of the errors and how that may affect, for example, flux estimates? The text describing the figure (page 30, lines 4-8) is also somewhat inconsistent with the figure. For example, CH4 appears to be showing a positive bias at high latitudes in the map, but the text points to a negative bias.

Reply: We believe that commenting on how the spatial errors will affect flux is out of the scope of the paper. The simulator aerosol scheme is described in the paragraph starting at (page 21, line 27) and it is stated that there are no assumptions of aerosol/cloud spatial correlation, only that a wide range of aerosol/cloud conditions are sampled. The "negative" bias at high latitudes mentioned in the text is a mistake. This has been changed to "positive".

Figure 8: the figure is informative but very complex. Perhaps it can be simplified. The standard deviation and mean after bias correction are not discussed much in the text.

Reply: We admit that the figure is a bit complex but a similar figure has been used in the past with a paper related to OCO-2. We think that the figure is rather clever in presenting a lot of information in one place and presents the point we are trying to make about post-process filtering.

Table 7: with dust AOD as a filter, is there concern that this may lead to sampling bias for certain dust-laden areas â since dust distribution is quite inhomogeneous?

Reply: We agree that dust distribution is quite inhomogeneous but we believe the dust AOD retrieval in L2FP is robust enough to correlate with this spatial variability. So, we are assuming that this filter will affect only dust-laden areas in which the dust AOD is too large to make a useful retrieval.

Page 36 line 15: could you please update the reference for Crowell et al. 2021 in the reference list?

Reply: Fixed.

Page 37, line 3: the last sentence of the paragraph needs to be re-written.

Reply: The sentence has been removed since soundings over water were not operationally planned for the late GeoCarb mission.

Section 5.7: please explain what âhand-tunedâ means here.

Reply: We think that the nature of the filter building has been discussed in detail along with the fact that it is difficult to make them perfect. "hand-tuned" was meant to say that some experiments have specific filters "tuned" for them. The sentence was meant to say that results could be improved with the ever evolving filtering construction. We decided to just remove the sentence because this is already made clear in the paper.

Table 9: please indicate how the component error is calculated.

Reply: We indicated this.
Replies to comments made by reviewer #2 of "The GeoCarb greenhouse gas retrieval algorithm: Simulations and sensitivity to sources of uncertainty" (amt-2023-17).

The authors of this manuscript would like extend our gratitude to Reviewer #2 for providing an extremely thorough review. We hope that all of the reviewer's comments have been addressed to the reviewer's satisfaction.

Reviwer #2

Reviewer's summary

Review of âThe GeoCarb greenhouse gas retrieval algorithm: Simulations and sensitivity to sources of uncertaintyâ by McGarragh et al., amt-2023-17

This paper presents a study for the GeoCarb mission, in which the authors describe the Level-2 retrieval algorithm for XCO, XCO2 and XCH4 and they conduct simulations to study error budgets for those three different trace gas retrievals in an effort to determine which sources of uncertainty play dominant roles in the total Level-2 error budget. This is achieved by running the Level-2 algorithm on synthetic radiance spectra generated with full knowledge of the atmospheric/radiative âtruthâ in the forward simulation. The inversion of these L1 radiances is then carried out with perturbations introduced into L2-algorithm input data like spectroscopic databases, meteorological priors; or perturbations added to the L1 spectra (e.g. from an instrument noise model). The results suggest that GeoCarb would have been able to meet retrieval precision requirements after bias correction.

The manuscript fits into the scope of AMT, the methods are valid and the results support the conclusions. However, the paper currently reads more like an internal, technical report and less like a scientific publication: it is very long, does not clearly emphasize what the novel concepts/results are - as compared to the existing literature on GeoCarb and the ACOS algorithm - and typos and imprecise language occur frequently throughout the text. While quality and clarity of the overall presentation must be improved, I also have some general questions about the methodological approach and the results.

General Comments from the reviewer along with author replies:

G1: This paper is quite similar to Polonsky et al. (2014), OâBrien et al. (2015) and OâDell et al. (2012). As of now, a reader has to go to those articles to find out how they are different with respect to the methods and the results. The authors need to outline more clearly the novel aspects of their manuscript with respect to previous literature and make clear what new science the community can learn here. The main message currently seems to be that GeoCarb would have been able to meet mission requirements. This had already been confirmed by Polonsky et al. (2014). The authors need to explain what the added value of this publication would be.

Reply: We added two paragraphs to the introduction to outline the similarities and differences.

G2: The paper is too long. Shorten by at least 5 pages to make this manuscript more accessible to readers. Here are some starting points for reducing the length: Much of chapter 3 can be found in previous works related to GeoCarb and OCO-2 and that content may be condensed. Similarly, the polarization experiment is a repeat of a previous study and it does not yield new conclusions. It can be left out of this manuscript.

Reply: We agree that the paper is *long*. We shortened the length of the Inversion, L2 forward model, and L1B simulation sections citing past papers. We also hope that this addresses the comment from the reviewer that the manuscript "reads more like an internal, technical report".

G3: The bias correction (section 5.1.2) plays a very important role in this work, but as far as I can tell, the actual formalism is never introduced. How does the bias correction work?

Reply: We are confused by this comment. The bias correction is discussed in section 3.6 with references therein.

G4: I find the discussion of the baseline experiment (Section 5.1) not satisfying from a scientific point of view. This is the simplest simulation experiment you carry out and you have knowledge of âthe truthâ. Biases that remain after bias correction are not explained and understood. You speculate about the roots causes for these results (âmay be dueâ, âlikely dueâ), but the reader cannot really understand what is going on here. What is the influence of the bias correction on the observed biases? It could be helpful to show a version of Figure 7, but for the clear-sky results (Figure 5). Is poor cloud filtering really the cause for the XCO bias over the Amazon? Wouldnât it then be worth to build a cloud filter that works better? You mention altitude dependent biases, but from section 3.1 I understood that altitude knowledge is perfect in this experiment. Please study your baseline experiment more carefully to convince readers that you understand what is driving biases. A more thorough discussion of the retrievals performed on the clear-sky ensemble of spectra may be illustrative.

Reply: We believe this is out-of-scope for this paper. In fact, what you are suggesting is a topic for one or more papers. The bias correction has been presented in the provided references. This paper is about formalizing the algorithm for a mission and providing some sensitivity results related to the instrument and the algorithm.

G5: The manuscript is written as if the GeoCarb mission had not been defunded last year. The authors should clearly indicate that the proposed instrument has been cancelled, unfortunately.

Reply: This has been indicated at the end of section 1. But, in fact, the instrument is being delivered to NASA, in full, late 2023 or early 2024, with the hope that it will eventually be adopted in a future mission proposal.

Specific Comments from the reviewer along with author replies:

The abstract should focus more on what the novel concepts or results of this study are. The first paragraph should be condensed.

Reply: We left the length of the first paragraph unchanged. The paper serves two purposes: 1) To introduce the GeoCarb instrument and associated L2 algorithms for the GeoCarb mission. And, 2) To provide the sensitivity study presented. The first paragraph addresses the first purpose.

Abstract, Line 4: XCO2, XCH4, XCO are introduced as column integrated amounts here, while on page 3, line 20 they are referred to as column integrated concentrations. I assume that in both cases the authors are referring to dry-air column averaged mole fractions. Please clarify in the abstract and throughout the text.

Reply: We made this consistent by using "column-integrated dry-air mole fractions" in the first few cases and in the rest of the paper as "column-integrated fractions".

Abstract, Line 6: âhyperspectralâ may be confusing for some readers, because it is often used to refer to coarse spectral resolution measurements with broad spectral coverage. Why not include the nominal FWHM of the different bands in the abstract to be more specific? âHyperspectralâ on Page 2, line 32 may be confusing in the same way.

Reply: We understand that the spectral resolution can vary significantly in so-called "hyperspectral" instruments but don't believe that including numerical details in the two cases mentioned is relevant for those cases. The spectral resolution of the instrument is clearly stated in section 2.

Abstract, Line 7: OCO-3 is not a polar orbiting instrument

Reply: Removed "polar".

Abstract, Line 13: ârelativelyâ is not very precise language, please elaborate.

Reply: "a relatively comprehensive error budget" -> "an error budget"

Page 2, Line 2: Why not more recent IPCC report as reference?

Reply: The 6th assessment has been cited instead.

Page 2, Line 25: remove reference OâBrien et al., 1998: it is unrelated to greenhouse gas surface fluxes and the context of this sentence.

Reply: Removed.

Page 2, Lines 29-31: MOPITT does not use any of the spectral bands described in the previous sentence and it

also makes use of thermal spectra; not just NIR/SWIR spectra. Please rewrite and consider using Deeter et al., 2003 (https://doi.org/10.1029/2002JD003186) as a reference instead of/in addition to Deeter et al., 2004.

Reply: This has been rewritten by putting the MOPITT reference in the previous paragraph and the reference has been changed as suggested.

Page 2, Line 34: perhaps update the GOSAT-2 reference with a more recent publication, e.g. Suto et al., 2021 - https://doi.org/10.5194/amt-14-2013-2021

Reply: Reference updated to the suggested reference.

Page 3, Lines 1-3: add Sentinel-5 to list of future missions.

Reply: Sentinel-5 has been added.

Page 3, Lines 29-31: Reword so readers wonât assume Bril et al and Butz et al are modifications of WFM-DOAS.

Reply: "modifications" -> "algorithms"

Page 3, Line 34 - Page 4 Line 1: Not all of these algorithms are optimal estimation retrieval algorithms, for instance, some are based on Phillips-Tikhonov regularisation -> adjust wording.

Reply: The wording has been adjusted to be more general.

Page 4, Line 8: what does ârelatively comprehensiveâ mean?

Reply: "a relatively comprehensive error budget" -> "an error budget"

Table 1: Where are these requirements from? Please provide details in the text.

Reply: We were aware of this. Unfortunately, there is no publicly available document with the requirements to reference. They are in the original proposal and in documents owned by the instrument builder but both are not publicly available. In reality, this paper is the public source of those requirements.

Page 8, Lines 6, 8: detector persistence is not an optical âaberrationâ. Rephrase this also on page 22, Line 19 and page 42, line 19.

Reply: These sentences have been rephrased.

Page 8, second paragraph: You state that you are aware of existing instrument issues related to keystone/smile, stray light, etc.. The reader is left to wonder what extent these problems might have. You are proposing conclusions on GeoCarbâs potential performance later on in the paper without taking these effects into account,

while these effects and the ability to correct for them, may definitely impact your results. Please add more context.

Reply: We believe that this is out-of-scope for this paper. In fact, it was planned that separate papers would be (and may still be) written on these effects and therefore more detailed context was left to these future publications once mitigation methods had/have been developed.

Page 10, Line 15: Why are you using an outdated HITRAN version? Same question for page 26, second paragraph.

Reply: HITRAN version 2016 is what was used the produce the absorption coefficient tables at the time. We do not believe this is a problem since in an experiment that simulates radiances the spectroscopy is assumed perfect. It is true that updated abortion coefficients may add features that provide addition information but we do not think that this will affect the outcome of the experiments in the paper. Of course updated spectroscopy is very important for retrieving from real data.

Page 10, Line 29/30: Why are you not using the 2020 version of the solar spectrum?

Reply: Same case as in the HITRAN comment. It is what was used and we do not think it will make a significant difference.

Page 11, Lines 18-19: At this point in the GeoCarb mission, have the Stokes coefficients been determined? If so, what are they? OâBrien et al. (2014) mention that calibration efforts were in planning when their paper came out.

Reply: At this point in the GeoCarb mission the Stokes coefficients have not been formally determined by measurement although the coefficients used are updated relative to OâBrien et al. (2014) to reflect instrument design changes. It is expected that final coefficients will be different than the ones used for the paper but that this will not significantly affect the results of the sensitivity experiments.

Page 11, Lines 19-22: to support this sentence it would be great to present the coefficients here.

Reply: We believe that the coefficients used, being that they are not final, are not worth presenting for this study. The coefficients will change after the formal characterization of instrument calibration. Of course these values will be important for interpretation of real data and they will be distributed to users in the L1B files.

Page 12, Lines 18-20: Why did you chose to use the mean noise values, instead of the actual noise levels across the spectral ranges (they appear to be known)? What does âroughly quadraticallyâ mean? Be specific.

Reply: The exact variation of noise with wavelength was unknown at the time and will not be know until instrument calibration is completed although mean noise coefficients can be predetermined well.

"roughly quadratically" means that if plotted the noise coefficients would look like they vary quadratically with wavelength but it won't be exactly quadratic (which is why they are provided to the user in tabular form in the L1B files.)

Again, this is a theoretical study. Everything going in is assumed exact, except for the perturbations of course. It is true that the shape of the noise with wavelength could have an affect on the sensitivity to different spectral features, but we do not believe that it will have a significant affect on the outcome of the study.

Table 3: Why are these values different from the ones in Table 2 of Polonsky et al. 2014?

Reply: These values are more recent estimates.

Page 12, Lines 13-15: The reference Kumer et al. (2013) only deals with one band relevant to GeoCarb, namely the CH4 band. Add references and context so readers can understand if the noise model is based on measurements, and which ones.

Reply: The sentence was generalized and the outdated references removed. It is now clear that the final noise model will be based on measurements and the current model is based on theory. Unfortunately there is no publicly available update-to-date reference for the model.

Page 13, Lines 33-34: Please add a reference or explain why you expect the ILS stretch factor to mitigate effects from scene heterogeneity. I thought scene heterogeneity would mostly introduce asymmetry into the ILS. This would not be corrected for in your set up.

Reply: Although not published and not shown in this study, it has been observed by the authors that the ILS stretch factor mitigates some of the scene inhomogeneity effects. This effect was intended to be addressed in an entire other publication on scene inhomogeneity that may or may not be published after the cancellation or a possible reboot.

Page 13, Line 12: what does ânearly identicalâ mean?

Reply: The sentence has been reworded to hopefully make it more clear.

Page 14/Table 4: Clearly indicate changes with respect to previous algorithm versions here.

Reply: We added to the caption the most recent reference that describes the algorithm on which the current algorithm is based.

Page 16, Lines 6-7: How did you determine the minimum Chi2 value possible and how did you chose the dP and Chi2 thresholds for cloud filtering?

Reply: These details are given in Taylor et al. (2012). This paragraph, along with the paragraph on GASBAG are really here just given as brief summaries since they are important for the overall L2FP pipeline.

Page 16, 3.5.2: What is the advantage of using the GASBAG retrieved gas ratios, instead of retrieving these yourself?

Reply: Simply put: speed. GASBAG makes simplifications, for example for scattering, that make it *significantly* faster than L2FP. As an aerosol/cloud preprocessor this is important especially given the amount of data that GeoCarb would have had to (or may in the future have to) deal with and the amount that can be filtered out avoiding running the much slower L2FP on soundings that will not produce scientifically useful retrievals.

In addition, the plan was to run GASBAG for SIF retrievals anyway so this makes the pipeline more efficient.

This is the same approach taken for OCO-2/OCO-3 with IMAP-DAOS.

Page 16, Lines 18-19: What are the thresholds used for the gas ratios?

Reply: The ratios have been added.

Page 19, Line 7: Why 78.06 degrees as upper limit? The plane-parallel assumptions breaks down at 70 degrees; solar zenith angle could/should be a filter criterion.

Reply: Breaks down at 70 degrees? This depends on a lot.

In this case 78.06 degrees is just the maximum satellite zenith angle encountered in the set of scan blocks. Sure, the plane-parallel assumption may start to break down at around 70 degrees but we do use a pseudo spherical outgoing beam approximation and it is still possible to make usable retrievals at satellite zenith angles larger than 70 degrees.

Filter variables are picked in order of decreasing effect. Regarding retrievals with satellite zenith angles to produce a usable retrieval, what is happening is that the effects of these large angles are indirectly showing affects in some of the first 12 filter variables and therefore get filtered out.

Page 21, Line 21: properly cite and acknowledge the carbon tracker work: https://gml.noaa.gov/ccgg/carbontracker/CT2017/citation.php

Reply: The cite has been updated.

Page 21, Line 21: Explain why you chose to use Carbon Tracker instead of GGG2020/GEOS to create the trace gas profiles in the simulator. For some simulator inputs you use the same information as what you use for inputs into the retrievals and for some you use different inputs.

Reply: Indeed there are some differences in the simulation inputs compared to the priors used in the L2FP retrieval. This is inherent in the retrieval algorithm itself. Other differences are in the aerosol and surface priors/parametrizations. We believe that the differences that exist will have a negligible affect on the outcome of the experiments performed for the paper.

Page 21, Line 23: Which ECMWF model are you using? Please add a citation.

Reply: Added citation.

Page 21, Line 27: Add citation or link with which readers can access the CALIPSO data product you used; not just the algorithm design paper.

Reply: Added citation.

Page 22, Line 8: Which of the three kernels did you actually use?

Reply: All three. They are meant to be used together as described in the references cited. The MODIS MCD43B product provides weights for each kernel.

Page 22, line 15: âsolution used in the simulator is essentially the same toâ¦â is it the same or not? Be specific.

Reply: "essentially the same" means exactly what it means, and implies "not important".

RT solutions vary mathematically. But, matrix operator methods such as doubling/adding or eigen matrix/adding, eigen matrix BVP techniques such as the so-called discrete ordinate method, or, in the CSU simulator's case, the successive order of interaction (SOI) technique, will all produce results with negligible differences relative to the other forward model error sources important to these retrievals. In this case, the RT solution differences between the simulator and L2FP are not relevant and not worth discussion.

This has been unchanged.

Page 22, Line 25: what is a âreasonably realistic error budgetâ?

Reply: "a reasonably realistic error budget" -> "an error budget for GeoCarb given the uncertainties investigated"

Page 22, Line 33: Why do you use GASBAG to filter for clouds instead of L2 variables?

Reply: This was covered by the reply to the "Page 16, 3.5.2" comment.

Page 23, Lines 20-21: On pg. 10, line 27 you write that you use the quadratic term. Please clarify.

Reply: The is referring to the quadratic term in the BRDF variation with wavelength discussed in the L2FP algorithm description. We left this unchanged.

Page 23, Lines 21-22: Are the L1B spectra modelled with spectral baselines that are just linear functions of wavelength? If yes: I think this is too simplistic. Please discuss in the text.

Reply: We assume the reviewer is referring to the linear variation of Rayleigh/aerosol/cloud/surface properties with wavelength. We really don't think that using more complex functions will affect the experimental results. Although, using more complex functions with wavelength has recently been the focus of some research (unpublished) in the community.

Page 23, Lines 32-33: How big are the differences between the prior/truth CO2/CH4/CO?

Reply: The differences are not significant in this case and would not be worth showing. The L1B simulations try to produce realistic radiances. The priors to L2FP are not *assumed inputs* to the algorithm. They are a starting point to solving for these variables, which, as shown are sufficient to retrieve the Xgas parameters to within instrument requirements. It is true that a better guess of priors can maybe improve the retrieval results a bit but the algorithm itself would be deficient if it could not use generic profiles like these.

No changes made here.

Page 23, Line 33: What are the âsubtle differencesâ in the radiative transfer? Be specific.

Reply: This has been addressed in the reply for the comment made for "Page 22, line 15". The assumption here is that the RT has been worked out. There is over half a century of atmospheric radiative transfer literature. The appropriate references to the mathematical solutions employed for the L1B simulations and the L2FP retrievals have been given with references and references therein. The numerical differences are insignificant in the experimental results of this study.

Page 25, Lines 24-25: The reader does not know how this experiment differs specifically from OâBrien et al. (2015). Please elaborate what the differences are and why this experiment needs to be repeated. Remove ââ¦and GeoCarb instrument modelâ, as the noise model is irrelevant here, since the experiment does not consider the radiance noise from the instrument model.

Reply: We believe the difference is made clear in the sentence:

1) Updated simulation/retrieval framework.
2) Updated GeoCarb instrument model.

The instrument model does not just include radiance noise. This is clear from section: 3.2.1.

Page 25, Line 29: If I understand correctly from page 29, the simulation runs on pixels the size of roughly 50x50 km^2? Why do you shift by 1 km? Please explain in the text. Maybe mention the pixel size here.

Reply: As stated in section 2: The footprint size used in the experimental setup is actually 2.7 x 5.4 km. It is true that it has changed a bit with the evolution of the instrument but we do not believe that this will affect the outcome of this experiment. We have taken the reviewers advice and have reminded the reader of the footprint size here.

Page 26, Lines 4-6: The reader may not necessarily know what the actual differences are between the two models.

Reply: We are assuming that reader is aware that different models, particularly models from different organizations do differ, sometimes significantly.

Page 26, Line 12: âT-widthâ - do you mean Temperature-dependence?

Reply: Yes. We have changed "T-width" to "temperature-dependence".

Page 26, Line 13: âspeed dependenceâ - which line profile are you fitting?

Reply: There isn't just one line profile fitted to. References to the spectroscopic computations are given at page 10, line 13. We think further detail is out-of-scope and if the reader is interested they should seek out the references given.

Page 27, Line 17: âoptimalâ: rephrase. Optimal averaging kernels would be 1 everywhere.

Reply: "optimal" was meant to refer to the increased sensitivity at the surface in regards to the GeoCarb mission objectives. The sentence has been rephrased.

Page 28, Line 1-2: Please add a citation to your statement that CO mixing ratio typically increases with height. I am not sure that is accurate (see e.g. www.atmos-chem-phys.org/acp/5/2901/).

Reply: This was poorly worded. What happens is that the sensitivity of the retrieval algorithm to CO increases relative to that of the other two gases with height. It is unknown to the authors why this is. It could be due to differences in the dependence of the spectroscopy with pressure and temperature of CO relative to the other two gases.

We rephrased the sentence to be more general and less speculative.

Page 28, Line 30: It would be worthwhile putting these numbers in the context to previous simulation studies using the ACOS algorithm, for instance OâDell et al., AMT, 2012. They also did this baseline experiment and found the precision for OCO-2 XCO2 to be 0.25 ppm. Why do you find larger standard deviation in the present experiment? Does the bias go the other way now? Why? What geophysical variables impact the baseline experiment (surface pressure?)? Additionally, Xi et al., AMT, 2015, doi:10.5194/amt-8-4817-2015 should be cited and discussed for reference.

Reply: We do agree that this could be useful to some readers but this would lengthen the manuscript more than it is. We feel that the paper should be considered an updated look at and a relatively current status of the GeoCarb mission in 2023 (even though it is questionable that it will fly). OâDell et al. (2012 and 2018) are different in that they don't have fourth band and don't retrieve the other two gases.

We definitely agree that Xi et al. (2015) should have been referenced. This has been done in the introduction.

Page 30, Lines 2-3: I agree that the bias is âcuriousâ. Instead of speculating about the cause of the bias it would be valuable to actually confirm where this bias comes from. This sentence currently may give the impression that you observed a large bias, but did not bother investigating further, since the bias correction somehow helps you out.

Reply: Confirming where the bias is from is out-of-scope for this paper and subject for another paper.

Page 30, Line 6: Please remove the word âmildâ.

Reply: We replaced "mild" with "small". We think that it is clear that when using terms describing relative differences they are based against the mission requirements.

Page 30, Figure 7: How are data aggregated here (spatially and temporally)? Please explain in the text.

Reply: We think that it is made clear how the data is aggregated in section 4.2.

Page 30, Line 15: When using the H2O ratio from GASBAG for cloud and aerosol screening here: What is the threshold here vs. in the prescreening, section 3.5?

Reply: The caption for figure 8 contains the sentence: "The filter thresholds are shown as vertical dashed lines."

Prescreening is very different and much more liberal than the post-process filtering as explained in the text. The thresholds will be different and are not comparable.

Page 30, Line 19: âerrorsâ: these are not âerrorsâ, but low values in P_hat reflect the true actual shorter light path. Rephrase.

Reply: Changed "errors" to "effects".

Page 31, Line 3: âincreasedâ over what?

Reply: Relative to smaller particles. Reworded.

Page 31, Line 5: what are ârelatively largeâ ice crystals?

Reply: It just was just meant to mean that ice crystals are larger relative to other aerosols. We removed "relatively large".

Page 31, Line 5: Is there a significant difference in scattering efficiency for ice particles between the strong CO2 band and the 2.3 um band? Add a reference.

Reply: We generalized this statement.

Page 31, Lines 10-13: Since you know the truth: be specific and quantify the error in prior surface pressure in high altitude locations where bias is observed.

Reply: We believe that going in to this detail is out-of-scope for the paper. We are making suggestions about known problems that, related GHG retrievals, can constitute papers in their own right.

Page 31, Line 10-13: What is the pre-screening signal threshold to flag ice/snow?

Reply: There is no "pre-screening signal threshold to flag ice/snow". We rephrased a bit here.

Page 31, Lines 13-14: Does the ILS fit in other bands also show this behaviour? What does the ILS fit do in the spectroscopy experiment?

Reply: What is apparent is the ILS fit is not as important as a filter variable in the other bands, as what is made clear, the filter variables are selected by decreasing importance. The ILS fit is not significant in the spectroscopy experiment.

Page 31, Line 17: Why italic and when is an experiment âunusualâ?

Reply: Again, in this entire process post-filtering is difficult and comes down to somewhat scientifically subjective decisions by the product providers to provide scientifically useful data. But, the product before post post-processing would be available for user if need be.

We removed the italics but kept "usually" because it is made clear in the paper that some experiments share filters and some have their own.

Page 31, Line 21: When is a filter âsufficientâ?

Reply: We think it is clear that the choice of filtering thresholds is subjective and has been explained in the cited reference. It is also mentioned (and cited) that more systematic methods of filtering are a subject of current research.

Page 32, Figure 8: dP-from-ABP plot: Why does light path enhancement, as seen here, lead to negative XCO2 bias - just as light path shortening? This seems unphysical to me. Please double check your results and explain in the text. The same can also be observed in the dP-from-L2 plot. In contrast, the CO2-grad-delta plot shows the behaviour that I would expect, where light path enhancement and shortening lead to opposite signs in bias.

Reply: We agree that increased photon path length would lead to more CO2 absorption but that is not what we are presenting in those figures. It is simply the effect of aerosols/clouds on the retrieval manifesting itself in surface pressure error which is being used as a filter threshold variable.

Page 33, Table 7: it appears that the last three variables in the bias correction have no significant impact on the performance of the correction. Why did you choose to include them anyway?

Reply: Showing those variables in the table is simply a way of showing that the importance of the variables decreases to the point that, as the reviewer pointed out, they become insignificant.

Page 33, Table 7: Why did you choose not to include albedo as a bias correction variable? How helpful would that be? Please discuss.

Reply: Page 33, Table 7 lists post-process filters, not bias correction variables.

As far as albedo for filtering? It has been stated that we are using the most influential variables in decreasing order of influence. And, in a previous reply to the reviewer we stated that effects from variables can show up in other variables. It is true that albedo will have effects on the retrievals but these effects may manifest themselves in other variables. There are references that explain the bias correction in greater detail.

Page 33, Lines 10-11: Why such large underestimation of the uncertainty for XCH4? -> weaker spectral signatures?

Reply: Simply put: we don't know, and we are not going to speculate on it. This kind of retrieval of all three gases simultaneously from these bands is relatively new and further analysis of the errors is subject to further research.

Page 33, Lines 11-13: Please discuss why XCO errors are mostly noise driven?

Reply: To remind the reviewer, there a lot of variables involved, so there are a lot of reasons this can happen but we do think that it is primarily because the signal to XCO is lower.

A comment has been added for this.

Section 5.1.3: Please put your conclusions here into the context of previous work. i.e. is it a novel conclusion that XCH4, XCO2 are mostly affected by systematic errors and XCO more by noise? Is this generally true for satellite remote sensing missions targeting these gases?

Reply: This is our conclusion based on the experiments. Any more analysis is out-of-scope for this particular paper and perhaps could be the subject of another paper. Previous research? There really isn't much involving simultaneous retrievals of all three of these gases with state vectors as complex as the one we use.

Page 35, Lines 2-3: How was this new filter built?

Reply: The same way the other filters were built which is summarized in a previous section with further details in the references therein.

Page 35, Section 5.3: Have you tried fitting a radiometric offset for this experiment?

Reply: No we did not. Or, a per band radiometric scaling is another option.

Page 35, Section 5.3: What is the mechanism by which the bias correction is able to correct for errors caused by this experiment, given that the correction does not include albedo.

Reply: It is not that the bias correction is *not able* to use surface reflectance as a variable. As stated before the variables that are used for the bias correction (and the filtering) are the most influential variables. Remember, that the effects from surface reflectance will manifest themselves in other variables. There are references that explain the bias correction in greater detail.

Page 37, Line 1: âlargely consistentâ -> what are the differences? Please explain.

Reply: This just means that there are no significant differences. So the updated, wavelength dependent, polarization model does not introduce any significant differences.

Page 37, Line 6: What are the âthree other daysâ?

Reply: As mentioned previously the pointing test was performed only for the day in March 21, 2016. The other three days June 21, 2016; September 21, 2016; and December 21, 2016; where not included in this experiment.

Page 38, section 5.7: Why does XCH4 raw respond differently to this experiment than XCO2? Qualitatively it appears from Fig. 15 that the meteorological perturbation impacts these two column quantities in very different ways.

Reply: Good question! We will leave this for further research. We think that we made clear that XCH4 has

large biases that we don't understand. The good news is that our bias correction does a good job of correcting those errors. Perhaps the bias correction would be a good start for further research.

Page 39, Line 5: âAll three target gas species still meet our mission requirementsâ Add âafter bias correction.â

Reply: "requirements" -> "requirements after filtering and bias correction"

Page 39, Line 8: What does âroughly consistentâ mean?

Reply: With errors that are comparable with baseline errors given mission precision requirements.

Page 41, Line 4-5: What are âreal errorsâ and âgood quality fractionsâ?

Reply: "real errors" -> "baseline errors"

"good quality fraction" -> "fractions"

Page 41, Line 14: âRetrievals of [â¦] are drivenâ¦â -> âErrors of [â¦] retrievals are drivenâ¦â

Reply: "Retrievals" -> "Retrieval errors"

Page 42, Line 1: see above: âRetrievalsâ -> âErrorsâ.

Reply: "Retrievals" -> "Retrieval errors"

Page 42, Line 2: âGeoCarb will do amazingly wellâ extremely colloquial language. Rephrase.

Reply: "GeoCarb will do amazingly well" -> "GeoCarb will do better than expected"

Page 42, Line 3: âThe retrievals of XCO2 , XCH4 , and XCO meet the mission precision requirements for all error sources, alone and in combination.â -> âThe bias corrected retrievals of â¦.â

Reply: "The retrievals of" -> "The filtered and bias corrected retrievals of"

Technical Corrections from the reviewer along with author replies:

Abstract, Line 2: ââ¦objectives of GeoCarb areâ¦â -> ââ¦objectives of GeoCarb wereâ¦â

Reply: Changed.

Abstract, Line 8: Geostationary -> geostationary

Reply: Changed.

Abstract, Line 8: âGeoCarb will beâ¦â -> GeoCarb would have beenâ¦

Reply: Changed.

Abstract, Line 12: âL2 algorithmsâ: several algorithms?

Reply: We did not change this. The paper is definitely focused on L2FP but the other algorithms briefly touched on are also considered L2.

Abstract, Line 15: ââ¦in addition to,â¦â -> ââ¦in addition toâ¦â

Reply: Changed.

Abstract, Line 20: remove âlike SIFâ so that readers do not think that SIF retrieval errors are investigated here.

Reply: Removed.

Page 2, Line 3: ââ¦sources including,â¦â -> ââ¦sources includingâ¦â

Reply: Changed.

Page 2, Line 25: âIt turns out that there is more signal relative to greenhouse gas surface fluxesâ¦â: adapt wording of this sentence, because this sentence reads as if the fluxes are directly inferred from spectral measurements.

Reply: The sentence has been reworded.

Page 2, Line 30: ââ¦to measure of COâ¦â -> remove of.

Reply: Changed.

Page 3, Line 1: verb is missing in sentence

Reply: Fixed.

Page 3, Lines 11-17: Change wording to reflect that GeoCarb has been defunded. (âThe mission was a collaborationâ¦â)

Reply: Fixed.

Page 3, Line 20: concentrations -> dry-air mole fractions?

Reply: This has been fixed throughout the paper.

Page 3, Line 33: âThis is especially in the caseâ add true

Reply: Added.

Page 4, Line 6: why italic font?

Reply: It was an emphasis but we removed it.

Page 4, Line 8-9: ââ¦have been been previous discussedâ¦â -> ââ¦have been previously discussedâ¦â

Reply: Changed.

Page 4, Line 31: move citation to end of sentence.

Reply: Fixed.

Table 2: correct âCH4$_4$â in the last row

Reply: Fixed.

Page 5, Line 13: âGeoCarb is the first plannedâ¦â -> âGeoCarb was the first plannedâ¦â

Reply: Changed.

Caption of Figure 1: âSample spectrumsâ -> âSimulated spectraâ; add â.â at end of sentence

Reply: Changed.

Page 8, Line 21: add â.â at end of sentence

Reply: Fixed.

Page 11, Equation 9: Add reference to the respective equation in OâBrien et al. (2014) and double check if it should be (Qcos2phi - Usin2phi) or (Qcos2phi + Usin2phi).

Reply: Reference to OâBrien et al. (2015) is already given in two places in that section. The linear wavelength dependence is the difference in those equations relative to OâBrien et al. (2015).

The sign in the equation has been changed to be consistent with OâBrien et al. (2015).

Page 16, Line 8: to thick -> too thick

Reply: Changed.

Page 17, Line 14: \citet{Keely} -> \citep{Keely}

Reply: Changed.

Page 18, Line 7: CSU acronym has not been introduced before; spell out.

Reply: Fixed.

Page 20, Line 19: ââ¦GeoCarb will ultimatelyâ¦â -> ââ¦GeoCarb would have beenâ¦â

Reply: "GeoCarb will ultimately sample" -> "GeoCarb would have sampled"

Page 21, Line 32: exists -> exist

Reply: Changed.

Page 22, Line 8: isotopic -> isotropic

Reply: Changed.

Page 22, Line 18: spectrums -> spectra

Reply: Changed.

Page 22, Line 31: âL1Bâ -> âL2â ?

Reply: L1B is correct here.

Page 23, Line 7: ââ¦GeoCarb willâ¦â -> ââ¦GeoCarb was not meant toâ¦â

Reply: Changed.

Page 23, Line 19: âFor aerosol. the sameâ¦â remove â.â

Reply: Fixed.

Page 23, Line 22: âincludedâ -> âincludeâ

Reply: Fixed.

Page 24, Line 5: âsimulator is ranâ -> âsimulator is runâ

Reply: We intended to use this tense here.

Page 25, Line 17: \citet{MDRA} -> citep{MDRA}

Reply: Changed.

Page 26, Eq 19: u_{gas,a} has not been introduced

Reply: That is because it should be u_{gas,ap} which has been fixed.

Page 27, Line 14: ââ¦ for GeoCarbâ¦â -> ââ¦for our GeoCarb simulationâ¦â

Reply: Changed.

Page 28, Line 1: âCOâ -> âXCOâ

Reply: Changed.

Page 28, Line 9: âcertain amount soundingsâ -> âcertain amount of soundingsâ

Reply: Changed.

Page 28, Line 10: ââ¦due to muchâ¦â -> ââ¦due to too muchâ¦â?

Reply: Changed.

Figures 5-7, 10-18: there is an issue with the rendering of text in all of these (labels, etc.). Please update.

Reply: We are not sure what the issues the reviewer is referring to.

Page 31, Line 18: âdo toâ -> âdue toâ

Reply: Changed.

Page 31, Lines 21-22: âwhereâ -> âwereâ

Reply: Changed.

Page 31, Line 26: XCO2 subscript contains zero, not O.

Reply: Fixed.

Page 31, Line 27: ââ¦operational OCO-2 XCO2 retrievalâ¦â -> ââ¦operational OCO-2 XCO2 bias correctionâ¦â

Reply: Changed.

Page 32, Figure 8: dP-from-L2 plot: missing negative sign on axis label

Reply: We see the negative sign on the access labels. Maybe this is a problem with the reviewers PDF viewer.

Page 32, Figure 8: CO2-grad-delta plot: missing negative sign on axis label

Reply: We see the negative sign on the access labels. Maybe this is a problem with the reviewers PDF viewer.

Page 33, Table 7: add units to mu, sigma, RMS

Reply: To keep the table at its current size we put the units in the caption.

Table 8: explain abbreviation âDU=wat+SSâ etc in caption.

Reply: This has been added.

Table 8, Table 9: change spelling of XCO, XCO2, XCH4 to X$_CO$, X$_CO_2$, X$_CH_4$ to be consistent with the rest of the manuscript.

Reply: We believe it is consistent: Text uses: $X_{\mrm{CO}_{2}}$ whereas tables and figures use XCO$_{2}}$

Although, table 8 had some type setting issues that were fixed.

Figure 9, caption: âShow areâ -> âShown areâ

Reply: Changed.

Figure 9, caption: âand the with noise resultsâ -> reword

Reply: Changed.

Page 35, Line 6: ââ¦retrievals ofâ¦are driven byâ¦â -> ââ¦errors ofâ¦are driven byâ¦â

Reply: We kept this unchanged as we think it is clear what is being expressed.

Page 35, Line 16: âsensitive to the overallâ¦calibrationâ -> âsensitive to an offset in the overallâ¦calibrationâ

Reply: Changed.

Page 36, Lines 6-7: Rephrase: there is very little CO2 in bands 1 and 4.

Reply: We are simply pointing out that fact.

Page 39, Line 8: âppm. 3.6â¦â remove â.â

Reply: Fixed.

Page 42, Line 18: âinstrument affectsâ -> âinstrument effectsâ

Reply: Changed.

Page 42, Lines 18, 20: âaffectsâ -> âeffectsâ

Reply: Changed.

References

Baker et al. 2010: one comma too many

Reply: Fixed.

OâBrien et al. 2009: add url or doi.

Reply: A url or doi unfortunately don't exist.

GeoCarb MDRA: add url or doi.

Reply: These unfortunately don't exist as this document is proprietary.