Greenhouse Gas Retrievals for the CO2M mission using the FOCAL method: First Performance Estimates
by S. Noël et al.

MS No.: amt-2023-194

# Reply to Referee 1

We thank the referee for the review and the comments. They will be considered in the revised version of the paper. In the following, the original reviewer comments are given in *italics*, our answer in normal font.

*The authors present performance estimates for their retrieval algorithm, FOCAL, applied to the upcoming CO2M mission. The performance estimates are based upon more detailed radiative transfer simulations using SCIATRAN, and are focused on the core retrieval algorithm biases rather than instrumental, spectroscopic, or meteorological. The authors estimate that FOCAL in it's current form meets (just!) the extremely stringent XCO2 and XCH4 requirements that have been placed on CO2M over land. Overall the paper is well written, provides useful information, and is well suited for AMT. However I have concerns about how representative the bias correction method is when being trained on the exact truth. More justification and information on the bias correction should be included, and conclusions updated accordingly. Specific comments follow below.*

We will add more information about the bias correction and update the conclusions, see also following answers.

## General Comments

1. *I recognize the need for bias correction, however I think what is done here may go beyond what is typical and requires at minimum more justification. The authors train their bias correction scheme against actual true values, which will not be known for CO2M. No results of the bias correction are shown, except it is clear that it likely has an impact on their final uncertainty quantification for large-scale fluxes based upon Figure 11 showing the best performance for the one month of data (April) that the bias correction is trained on. I would like to see more detail on what effect the bias correction is having, as well as what parameters it has deemed important. After all since this is entirely simulated data the results of the bias correction should be fully explainable. I realize the authors state that their results "do not consider any systematic errors in meteorology" but what this means in terms of the bias correction is missing.*

   The bias correction is indeed important to meet the requirements (this is essentially common standard for greenhouse gas retrieval algorithms).

   We use the a-priori model data as truth for the bias correction, because this is also planned to be done for real measurements data. Since the a-priori data are also the truth for the simulated data we agree with the referee that we do not consider here additional errors due to the limited knowledge of the truth. However, these errors strongly depend on the quality of the a-priori data which can only be determined by e.g. validation based on real data. Therefore we decided not to include this additional uncertainty. We will emphasise this in the updated manuscript.

   We will also add more information about the post-processing (filtering and bias correction), e.g. on which parameters it is based, and we will add a map of the bias correction.

   Note that the uncertainty quantification for large-scale fluxes shown in Figure 11 is not only based on the April 2015 data. The figure shows the results for all months of the year, i.e. also for data not used for training. We therefore consider our conclusions to be valid (under the mentioned assumptions).

The impact of systematic errors in meteorology on the bias correction cannot be reliably quantified, because this depends on the quality of the input meteorological data. In principle, errors in meteorology affect the data quality, but currently we assume that these data are on average correct and therefore their uncertainties for specific conditions should not have a major impact on the bias correction, especially because meteorological data are (intentionally) not used as parameters / features.

In this context, we would like to emphasise that FOCAL has been applied to real satellite data using similar bias correction as used and described in this paper, where we write in section 3.1:

"Applications to OCO-2, GOSAT and GOSAT-2 have shown that FOCAL is fast and produces accurate results. For example, the spatio-temporal bias of the FOCAL XCO2 product derived from TCCON comparisons is (after bias correction) in the order of 0.6 ppm for OCO-2 (Reuter and Hilker, 2022), and 0.6 (1.1) ppm for GOSAT (GOSAT-2) (Noël et al., 2022)."

This gives confidence that meaningful results for CO2M are presented in this publication although all results are based only on simulations and despite the fact that not all error sources are addressed in this publication.

2. *In some places I find the conclusions of the authors a little optimistic. For example the last sentence of the manuscript is "However, the current results give good confidence that the FOCAL-CO2M retrieval is able to fulfil the product quality requirements of the CO2M mission." But my interpretation of the results is if you neglect things not taken into account in the simulation (polarization, full sphericity, 2d/3d effects, among others), assume you have perfect spectroscopy and input ancillary data, assume you have a perfect instrument, and perform bias correction using perfect truth values, then you barely meet the requirements. This is not a knock on the retrieval method, the CO2M requirements are very stringent, but it is hard for me to have a takeaway other than that it is unlikely FOCAL will meet the requirements when applied to real CO2M data in its current form.*

   It is true that our simulations of the spectra do not consider all possible physical processes. However, it is not possible to include all of these in the simulations, because even if the radiative transfer model would be able to consider this the required input data on e.g. 2d/3d cloud or aerosol distribution is not available. We will mention this in the updated paper.

   Regarding polarisation:
   Some test data calculated by RAL (see reference in the paper) are available with/without considering polarisation in the radiative transfer. We have tested the impact of polarisation on our retrieval using these data, which is for XCO2 in the order of 0.1 ppm before bias correction. Therefore we think that neglecting polarisation in the SCIATRAN runs can be justified.

   Regarding sphericity:
   Both SCIATRAN and FOCAL could consider (at least pseudo) sphericity; however, in case of CO2M (normally nadir looking with a swath of about 200 km) spherical geometry has no major impact for solar zenith angles less than 75 deg and is therefore neglected.

   The main intention of the manuscript is to show that the FOCAL-CO2M retrieval method is in principle able to fulfil the requirements. We cannot finally state now that we will fulfil the requirements for real data. Especially for large scale fluxes we are indeed on the edge, but this is not the primary objective of CO2M. We will try to clarify this more in the updated paper.

   Nevertheless, we think that the current results – under the mentioned assumptions / limitations – show that we are on a good way (which is why we chose the "good confidence" formulation). For the revised version of the paper we removed "good". The modified sentence now reads as follows:

   "However, the current results give confidence that the FOCAL-CO2M retrieval algorithm will be able to generate products meeting the product quality requirements of the CO2M mission."

# Specific Comments

1. *p.3 l.57: I realize that a full description of the three algorithms is beyond the scope of this paper,*

*and that they can be found in the corresponding citations, but a brief statement here about their differences would be helpful for the reader.*

We will add some more information about the algorithms.

2. *p.3 l.78: "The SCIATRAN calculations are more complex than the FOCAL forward model. For example, they consider surface BRDF (bidirectional reflectance distribution function) effects, different aerosol types and distributions as well as clouds." This needs some further elaboration. While SCIATRAN itself can be much more complicated than the FOCAL forward model, it may not be configured that way. I see later on that polarization is neglected in the SCIATRAN calculations for example. I assume other features of SCIATRAN such as line of sight sphericity are not included even though they are supported as well.*

   We will include some more information about the SCIATRAN calculations.

   See also our response above regarding polarisation and sphericity.

3. *p.5 l.143: "we filter out all cloudy data" Surely not every scene with any clouds are filtered out or else there would be no purpose in including clouds in the SCIATRAN runs?*

   We consider clouds in the SCIATRAN calculations, but for the performance tests used in this study we only consider completely cloud-free soundings. We will clarify this in the text.

4. *p.6 l.171: "based on a set of training and test data" Most algorithms now recommend splitting data into three groups: Train, test, and validation. The reasoning being that you can overfit to your "test" data by tweaking your model parameters, or even what model you are using.*

   The bias correction is determined only from the training data. The test data are used to check if the performance on this data set is similar to the performance for the training set. Both training and test data are taken from the April 2015 results. In addition, we apply later on the bias correction to the complete 2015 data (with good results, as shown in the paper), so the additional 11 months can be considered as a validation data set.

5. *p.6 l.183: "bias correction does not consider any additional errors resulting from systematic differences between the estimated meteorological conditions and the actual atmosphere" Could this be included in some way? I see from previous publications that for real data truth is taken from a generated database. Is there no error estimate for this database that could be artificially included here?*

   As described above, it is not possible to include systematic differences between the estimated meteorological conditions and the actual atmosphere in the bias correction, because this would require a knowledge of specific systematic errors of the meteorological data, which are usually unknown (otherwise there would be a correction for them). If – as we assume – the meteorological data are on average correct, statistical (random-like) uncertainties in the data would have no impact on the bias correction.

   Previous applications of the FOCAL algorithm to e.g. GOSAT and GOSAT-2 use indeed as "true" XCO2 and XCH4 for the bias correction a database (e.g. derived from the SLIM climatology). We could also use the SLIM climatology for CO2M, but we assume that the actual CAMS model data which will be provided during the mission are a better estimate for the true XCO2 and XCH4. This is why we use the CAMS data as truth in this study.

   The bias correction will always try to reproduce the values given as truth. Random errors in the assumed truth are less relevant in this context, but any systematic error in the assumed true XCO2 and XCH4 will result in a corresponding error of the retrieved products. The latter error can only be quantified by comparisons with independent data (e.g. ground based measurements). There are error estimates for SLIM but these are not applicable to the operational CAMS data and therefore could only be used as a rough indication for possible additional errors due to the uncertainties of the "truth".

6. *p.7 l.204: ". . . cloud-free data . . . " As before, do you mean cloud VOD is identically 0? Or are some thin clouds included?*

   As mentioned above, for the performance verification we use only completely cloud-free data.

7. *Section 4.1: For the plume specifically we are probably getting down to the scale where 2d/3d effects are important. I don't expect the authors to quantify these effects since I know SCIATRAN is not capable of doing so and it is a large amount of work, but I was surprised to see no mention of the effect*

   We will mention that 2d/3d effects are not explicitly considered in the radiative transfer simulations.

8. *Section 4.3: Does the bias correction use the FOCAL retrieved optical depth in a significant way? I assume the idea would be to use the MAP aerosol parameters in the bias correction instead. Could you not test this directly by using the true aerosol OD + estimated errors of the MAP retrieval in the bias correction?*

   The derived parameters of the scattering layer (incl. optical depth) are used in the bias correction (will be shown in updated paper, see above). These parameters correlate quite well with aerosol. We did some tests using the true AOD as possible additional parameter for the bias correction, but this did not significantly change the results.

   There is not much information available about the quality of the retrieved MAP aerosol parameters, so we were not able to use this, especially since we would need this MAP L2 output consistent with our input data.

   However, the inclusion of MAP data for e.g. bias correction is foreseen in the algorithm, but this can only be tested with real data.

# Technical Corrections

1. *p.2 l.45: "welling" → "upwelling"*

   Corrected.

2. *p.2 l.53: I would recommend using "better" here instead of "higher", since "higher spatial resolution" can be ambiguous.*

   Changed.

3. *p.9 l.255: "is not much sensitive" Reword*

   Done, new text:
   "To show that the sensitivity of the retrieval to the choice of the a-priori is low, ..."

4. *p.9 l.257: "an" → "any"*

   Changed.