The authors would like to thank the referee for the kind review and constructive comments. All the comments will be addressed in blue, between each question within the review text.

The authors present a simulation-based assessment of CO2 and CH4 column retrievals for a novel nanosatellite constellation concept called SCARBO. The constellation would involve 20+ small satellites each carrying a Fabry-Perot interferometer (NanoCarb) for CH4 and CO2 and a multi-angle polarimeter (SPEXone) for aerosols. The authors present extensive error analysis with a focus on aerosol-related retrieval errors and the ability of the SPEXone auxiliary instrument to mitigate those errors. They use for this purpose standard OSSE methods including Rodgers (2000) optimal estimation techniques, and show that their NanoCarb-SPEXone instrument design should be capable of delivering high-precision XCO2 and XCH4 retrievals. The paper is well written and a good fit for AMT. I recommend acceptance for publication subject to the following comments and questions.

Specific comments

L. 36: Please clarify what is meant by "on small areas".

We mean that urban areas, that concentrate 70% of fossil fuel related CO2 emission (Duren and Miller, 2012), are very small compared to the total continental surface, 0.5-0.6% of ice-free land surface, as reported in the latest IPCC WGIII report (https://report.ipcc.ch/ar6wg3/pdf/IPCC_AR6_WGIII_FinalDraft_Chapter08.pdf): 653,365 (Liu et al., 2020) km2 in 2015 compared to 130,000,000 km2 (IPCC, 2022) of ice-free land surface.

We adapted the introduction sentence to better reflect this idea (revised manuscript line 37).

L. 40: Can you provide a reference for 2x2 km2 resolution being fine enough to resolve point sources? Of what magnitude? TROPOMI can resolve only extreme methane point sources at similar (5.5x7 km2) resolution (e.g., Pandey et al., 2019).

2x2 km2 is fine enough for all sources ≥10 MtCO2/year as per all the literature preparing CO2M mission: e.g. Bovensmann et al. (2010), Kuhlmann et al. (2019), etc. Higher spatial resolution is required to resolve smaller emission rates: e.g. 50x50 m2 for point sources ≥1 MtCO2/year (Strandgren et al., 2020).

We added the emission rate related to the 2x2 km2 spatial resolution, as well as Kuhlmann et al. (2019) reference in the introduction (revised manuscript line 42).

L. 41: Extensive recent work has shown that plumes observed by imaging spectrometers do not look Gaussian. For example Cusworth et al. (2021) use an integrated mass enhancement method to quantify CO2 emissions from individual power plants observed by the PRISMA satellite instrument, and the TROPOMI team and others have used a variety of methods to quantify CH4 plume emission rates (eg, Pandey et al., 2019) at km-scale resolution, but Gaussian plume modeling seems poorly suited to the problem.

Indeed, there are many different methods relying on plume images. We adapted the formulation and added relevant references, including Varon et al. (2018) that notably compared three of these different approaches (revised manuscript lines 43-44).

L. 66: "requirements for operational top-down monitoring of anthropogenic GHG emissions" – What are these requirements and what does "operational" CO2/CH4 monitoring mean?

Those requirements are: (1) imager instrument to resolve strong point source emission plumes (2) XCO2 precision better than 1 ppm, XCH4 precision better than 10 ppb (3) high revisit frequency, daily if possible. They were discussed for the design of the CarbonSat and then CO2M concepts, and identified as key to an European operational monitoring of fossil fuel CO2 emissions (Ciais et al., 2014, report by the EC: doi/10.2788/52148).

We adapted the text to reflect these details (see revised manuscript lines 69-71).

L. 79: "geophysical parameters necessary to retrieve XCO2 and XCH4" – can you say what these parameters are or point to them in the text?

Those are some of the geophysical parameter involved in the Full Physics retrieval of XCO2 and XCH4. They include CO2, CH4, O2, H2O and aerosols, which have been taken into account for the OPDs optimisation that defined the current design of NanoCarb used for this study.

Surface albedo must also be taken into account and has been included in the state vector, as well as temperature which has not been included in the state vector here, as per current NanoCarb design hypotheses (see review #1 answers).

In order to lighten the already lengthy introduction, we just refer to the section where those aspects are presented (revised manuscript line 86).

L. 80-81: "2.3 x 2.3 km2 spatial resolution, enabling to detect emission plumes from megacities and hotspots (e.g. > 10 Mt CO2 yr-1 power plants)" – You seem to use "hotspot" and "point source" interchangeably. Megacities are examples of hotspots and power plants examples of point sources. For point sources, where does the 10 Mt/y threshold come from?

We corrected this sentence (revised manuscript line 87).

The 10 MtCO2/yr with a 2x2 km2 is the threshold that was determined for CO2M definition, based on resolution and XCO2 precision experiments. The SCARBO precision and resolution objectives were chosen to be close to those of CO2M.

References cited before in the text (e.g. Kuhlmann et al, 2019) already give details and reference papers regarding this threshold.

L. 86: Please describe CO2M.

The CO2M acronym is already defined at line 77 (revised manuscript) and the very complete Mission Requirement Document is cited right after: Meijer and Team, 2019.

L. 95: "scattering error-critical atmospheric and observational parameters" – Are these the "geophysical parameters" you mentioned before? It would be helpful to list these out somewhere in the introduction if not too lengthy or point to them in the text.

Those are not entirely the same: the 'geophysical parameters' mentioned before are parameters that influence radiative transfer in the shortwave infrared, whereas these 'scattering error-critical atmospheric and observational parameters' explore geophysical parameters (surface albedo model, aerosol layer optical depths and height) and an observational one (solar zenith angle).

Indeed, the following paragraph that announces the structure of the article, in the pre-print, did not mention these 'scattering-error-critical atmospheric and observational parameters'. It is fixed in the revised manuscript (revised manuscript lines 114-115).

Ray Nassar and Dan Cusworth's works about satellite monitoring of CO2 emissions from power plants should be cited somewhere. Same for TROPOMI methane plume papers (Pandey et al. and others) since there has been a lot of recent work on these topics.

Recent articles by Ray Nassar et al. (2021), Cusworth et al. (2021) and Pendey et al. (2019) have been added as references when mentioning plume emission rate methods earlier in the introduction (see revised manuscript lines 43-44).

Can you explain why SCARBO uses an FP rather than grating? Is it about financial cost, instrument size/weight, something else? Also please cite other instruments/concepts that use FP - eg Jervis et al. (2021).

The FP technology used by NanoCarb, for SCARBO, has been chosen to propose a very compact instrument, of a volume comprised in a few cm^3 compared to several tens of litres for conventional instruments. Gratings permit to reach high levels of performance, but the bulk silicon Fabry-Perot enables to build a smaller – and simpler – instrument.

Reference to GHGSat FP technology (Jervis et al., 2021) has been added when mentioning small satellite concepts (revised manuscript line 74-75).

L. 165: Not clear what "0.003" means, is it an error (1 or 2 sigma)?

0.003 comprises both systematic and random errors, and is used for the SPEXone performance assessment, thus equivalent to 1 sigma, see Hasekamp et al. (2019).

We added a parenthesis to better underline this (revised manuscript line 174).

L. 173: What is an "emission clump"? This terminology is non-standard.

Emission clumps are defined in Wang et al. (2019) paper: 'In this study, we characterize area and point fossil fuel CO2 emitting sources which generate coherent XCO2 plumes that may be observed from space. We characterize these emitting sources around the globe and they are referred to as "emission clumps" hereafter.'

We adapted the manuscript text to reflect this definition (see revised manuscript lines 181-182).

L. 234: Are you referring to the instrument temperature?

We are referring to atmospheric temperature; the revised manuscript has been adapted to provide this explanation (see revised manuscript lines 244).

Table 2: How conservative is the 4 hPa (0.4%) error for surface pressure? It seems quite small.

This is the uncertainty used in the ACOS algorithm that produces the official OCO-2 XCO2 product. See ATBD p 43: https://docserver.gesdisc.eosdis.nasa.gov/public/project/OCO/OCO_L2_ATBD.pdf

The same surface pressure a priori uncertainty was used for the 5AI inverse scheme paper Dogniaux et al. (2021).

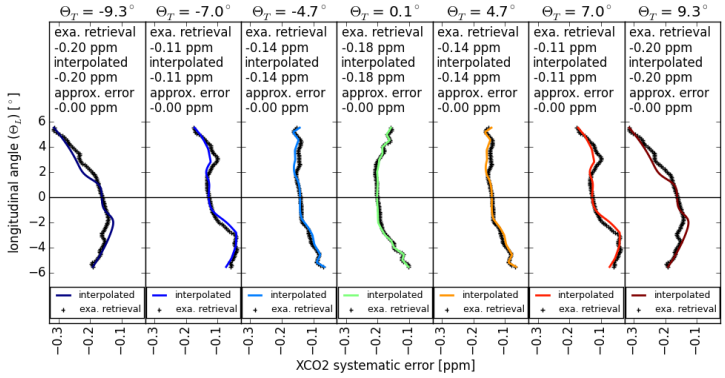
L. 271: I would suggest pointing to the appendix here because I initially wondered if the "combination" of single-pixel measurements was through averaging or something more.

We moved the reference to the Appendix there (see revised manuscript lines 282).

L. 289-290: "Errors arising from the interpolation have been assessed and are negligible (not shown)" – What is the magnitude of the error?

For seven θ_T transversal positions within the swath (θ_T = -9.3°, -7.0°, -4.7°, 0.1°, +4.7°, +7.0°, +9.3°), for a situation corresponding to ALB=VEG, SZA=50°, CLH=2km, COD=0.15 and FOD=0.08, we evaluated the interpolation approximation errors between combined exactly retrieved L2 results and combined interpolated L2 results. If interpolation approximation errors can be seen between pixel-wise exactly retrieved results (black crosses in Fig. R2.1 and Fig. R2.2) and pixel-wise interpolated results (coloured lines in Fig. R2.1 and Fig. R2.2), the interpolation approximation errors become small to even negligible when pixel-wise results are combined (see Fig. R2.1 and Fig. R2.2): up to a maximum of 0.01 ppm for XCO2 systematic and random errors, up to a maximum of 0.05 ppb for XCH4 systematic and random errors.

We purposefully did not include this discussion in order to lighten the content of the paper, so we adjusted the revised manuscript just by adding a parenthesis explaining the magnitude of these errors (revised manuscript lines 301-302).



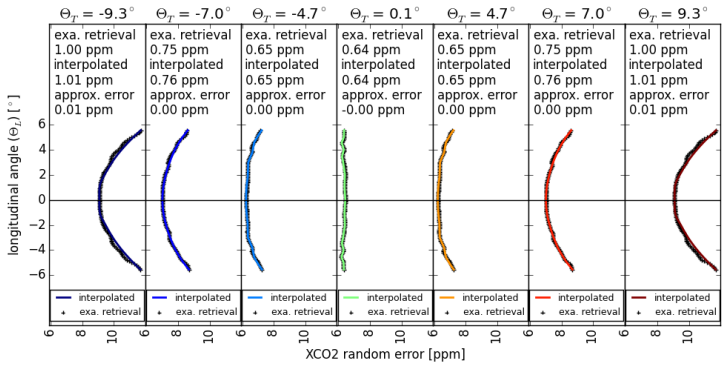
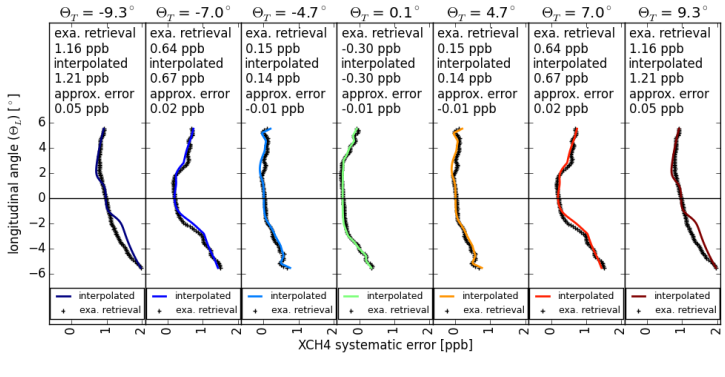


Figure R2.1 Comparison of pixel-wise interpolated (coloured full lines) and exactly retrieved (black crosses) XCO2 systematic (top) and random (bottom) errors for 7 different transversal positions θ_T.



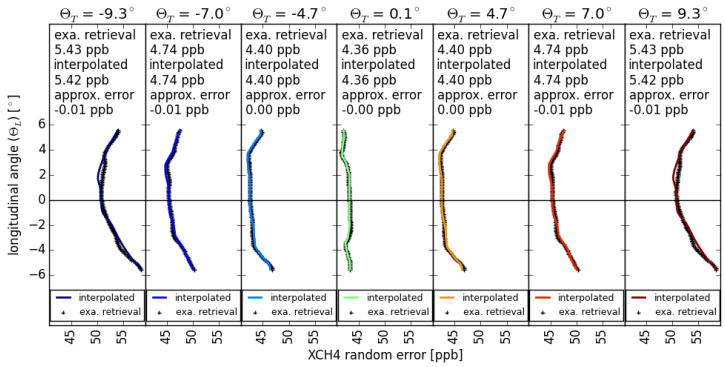


Figure R2.2 Comparison of pixel-wise interpolated (coloured full lines) and exactly retrieved (black crosses) XCH4 systematic (top) and random (bottom) errors for 7 different transversal positions θ_T.

L. 297-300: CH4 falls off rapidly above the troposphere, so why use a uniform vertical profile? Is the impact of this unrealistic profile on the retrieval small enough to be neglected?

As it is the very first performance assessment of the – still maturing – NanoCarb concept realised by using a complete inverse radiative transfer scheme, we chose simplified CO2 and CH4 profiles. However, we agree with both referees that realistic CO2 and CH4 profiles must eventually be considered for a final SCARBO evaluation study, that could be a full OSSE and that would demonstrate its maturity to fly and accomplish its mission.

We adapted the paragraph describing the a priori atmosphere to include these previous comments (revised manuscript line 313-315).

Besides, several aspects of the study conducted here argue in favour of a small impact of this hypothesis:

- we only consider scattering-related a priori misknowledge, the CH4 profile is supposed to be a priori perfectly known, thus no smoothing error can arise from a unrealistic CH4 profile in the retrievals
- XCH4 column averaging kernels show a small sensitivity to the top of the atmosphere (AKs values of about 0.5, cf Fig. 8), arguing for little CH4 information coming from the top of the atmosphere.

Fig. 7: Is the much lower DOFs for FOD in the with-SPEX scenario merely due to using a much lower prior error for FOD compared to no-SPEX?

The much lower FOD DOFs in the with-SPEX case is – according to our understanding – the result of (1) the tight constraint brought by much lower a priori error brought by SPEXone (2) the already existing low information content regarding this mode of aerosols.

In other words, the information brought by SPEXone is so large compared to the one available in NanoCarb measurements (see larger but still low no-SPEX FOD DOFs) that only the a priori information contributes to the estimation of fine mode aerosols in the with-SPEX case. This is illustrated by the very low FOD DOFs in the with-SPEX case.

L. 424-425: Do the albedo DOFs actually increase? They seem to be equal to 1 in both scenarios.

Yes, they do. Albedo DOFs are slightly below 1, but it is not distinguishable at this scale. However, we recognize that this increase between no-SPEX and with-SPEX is very small and not distinguishable.

We adapted the manuscript to stress that, in the case of albedo, the increase is very small (revised manuscript line 443-444).

L. 437-439: This seems odd since the albedo DOFs look to be almost or exactly 1.0.

Albedo DOFs are indeed not exactly 1.0 (undistinguisable). We adapted the revised manuscript (revised manuscript line 419).

Another way to look at this estimation entanglement is to examine covariance coefficients between CO2 scaling factor and albedo parameters in the Optimal Estimation a posteriori covariance matrix. Divided by the standard deviations of these variables, it yields correlation coefficients. For example, for NanoCarb, we find a correlation coefficient of 0.72 between CO2 scaling factor and B2 albedo. By comparison, this correlation coefficient is 0.04 if run a similar exercise with an OCO-2 measurement.

Fig. 8: It's not clear how you compute column averaging kernels when your state vector doesn't include a vertical column but rather a single scaling factor for each gas. If the state vector included CH4 and CO2 at different vertical levels then you would obtain A = dxhat/dx giving the AK for each vertical layer. How do you get column averaging kernels when optimizing just a scaling factor?

We make use of a complementary way to compute the AK matrix: A = GK. The 5AI inverse scheme stores the atmospheric layer-wise GHG jacobians and uses them to compute a vertical column averaging kernel. Those are obtained with the following equation:

$$(a_{GHG})_j = G_{GHG}\, K_{GHG,j} x_{a,GHG,j} / h_j$$

with $(a_{GHG})_j$, the column averaging kernel of the j-th atmospheric layer for a given GHG, $G_{GHG}\, K_{GHG,j}$, the [1,m]x[m,1] matrix product between the GHG scaling factor related line of gain matrix G_{GHG}, and the stored j-th atmospheric layer GHG jacobian $K_{GHG,j}$ (with m being the length of the measure vector), $x_{a,GHG,j}$ the a priori

GHG concentration for the j-th atmospheric layer and h_j the pressure weighting function of the j-th atmospheric layer.

Fig. 8: Also the column averaging kernels look quite smooth, what is the vertical resolution here?

The vertical resolution can be obtained by analysing the averaging kernel row peak widths, as per Rodgers (2000). While the equation written above enables to compute the column averaging kernel – through the gain matrix row dedicated to GHG scaling factor – the layer-wise averaging kernels are not available.

A simple way to try to analyse those is to realise a small performance test that considers a GHG profiles instead of scaling factors in the state vector. Fig R2.3 shows the layer-wise averaging kernels that we obtain for CO2, considering the a priori covariance matrix used by ACOS, the official OCO-2 algorithm. As seen on Fig R2.3, no width can be properly identified, which is consistent with the below-unity GHG degrees of freedom for NanoCarb measurements.

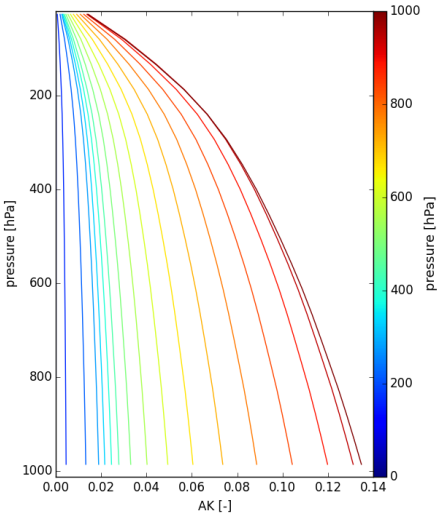


Figure R2.3 Layer-wise CO2 averaging kernel obtained when retrieving a CO2 profile from NanoCarb measurements.

Several plots show regions where retrievals did not converge satisfactorily but I cannot find in the manuscript what method you use for the optimization. Is it Newton, Levenberg-Marquardt, something else?

We used Levenberg-Marquardt optimization method. We modified the revised manuscript to add this information (revised manuscript line 217-218).

Fig. 13: Why is there striping in this figure? Because of the loss of precision with increasing transversal position?

Indeed, the stripping is caused by the increase of random error with the transversal position within the swath. We modified the revised manuscript to add this explanation (revised manuscript lines 674-675).

Technical corrections

L. 39: “large-swath”

We corrected it (revised manuscript line 41).

L. 51: “best fits”

We corrected it (revised manuscript line 54).

L. 151: “radiance spectrum”

We corrected it (revised manuscript line 159).

L. 164: “at 50 spectral band” seems like a typo?

This is not a typo.

L. 178: “compromises”

We corrected it (revised manuscript lines 187-188).

L. 276: “fasten” doesn’t seem like the right word here. Do you mean “speed up” or something similar?

We corrected it (revised manuscript lines 287).

L. 422: “in these situations” typo

We corrected it (revised manuscript lines 440).

L. 612: “then” typo

We corrected it (revised manuscript lines 631).

L. 637: “mentioned” typo

We corrected it (revised manuscript lines 656).

Consider changing “scattering error-critical” to “scattering-error-critical” everywhere. If I understand correctly it’s meant to be a compound adjective and might be clearer with two –'s.

We made this change.

References

Pandey, S., Gautam, R., Houweling, S., van der Gon, H. D., Sadavarte, P., Borsdorff, T., Hasekamp, O., Landgraf, J., Tol, P., van Kempen, T., Hoogeveen, R., van Hees, R., Hamburg, S. P., Maasakkers, J. D., and Aben, I.: Satellite observations reveal extreme methane leakage from a natural gas well blowout, P. Natl. Acad. Sci. USA, 116, 26376–26381, https://doi.org/10.1073/pnas.1908712116, 2019.

D.H. Cusworth, R.M. Duren, A.K. Thorpe, M.L. Eastwood, R.O. Green, P.E. Dennison, C. Frankenberg, J.W. Heckler, G.P. Asner, C.E. Miller: Quantifying global power plant carbon dioxide emissions with imaging spectroscopy, AGU Adv., 2 (2) (2021), https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020AV000350

Jervis, D., McKeever, J., Durak, B. O. A., Sloan, J. J., Gains, D., Varon, D. J., Ramier, A., Strupler, M., and Tarrant, E.: The GHGSat-D imaging spectrometer, Atmos. Meas. Tech., 14, 2127–2140, https://doi.org/10.5194/amt-14-2127-2021, 2021.

References

Liu, X., Huang, Y., Xu, X. et al. High-spatiotemporal-resolution mapping of global urban change from 1985 to 2015. Nat Sustain 3, 564–570 (2020). https://doi.org/10.1038/s41893-020-0521-x

Bovensmann, H., Buchwitz, M., Burrows, J. P., Reuter, M., Krings, T., Gerilowski, K., Schneising, O., Heymann, J., Tretner, A. and Erzinger, J.: A remote sensing technique for global monitoring of power plant CO2 emissions from space and related applications, Atmos. Meas. Tech., 3(4), 781–811, doi:10.5194/amt-3-781-2010, 2010.

Kuhlmann, G., Broquet, G., Marshall, J., Clément, V., Löscher, A., Meijer, Y. and Brunner, D.: Detectability of CO2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO2 Monitoring (CO2M) mission, Atmos. Meas. Tech., 12(12), 6695–6719, doi:10.5194/amt-12-6695-2019, 2019.

Strandgren, J., Krutz, D., Wilzewski, J., Paproth, C., Sebastian, I., Gurney, K. R., Liang, J., Roiger, A. and Butz, A.: Towards spaceborne monitoring of localized CO2 emissions: an instrument concept and first performance assessment, Atmos. Meas. Tech., 13(6), 2887–2904, doi:10.5194/amt-13-2887-2020, 2020.

Ciais, P., Dolman, A. J., Bombelli, A., Duren, R., Peregon, A., Rayner, P. J., Miller, C., Gobron, N., Kinderman, G., Marland, G., Gruber, N., Chevallier, F., Andres, R. J., Balsamo, G., Bopp, L., Bréon, F.-M., Broquet, G., Dargaville, R., Battin, T. J., Borges, A., Bovensmann, H., Buchwitz, M., Butler, J., Canadell, J. G., Cook, R. B., DeFries, R., Engelen, R., Gurney, K. R., Heinze, C., Heimann, M., Held, A., Henry, M., Law, B., Luyssaert, S., Miller, J., Moriyama, T., Moulin, C., Myneni, R. B., Nussli, C., Obersteiner, M., Ojima, D., Pan, Y., Paris, J.-D., Piao, S. L., Poulter, B., Plummer, S., Quegan, S., Raymond, P., Reichstein, M., Rivier, L., Sabine, C., Schimel, D., Tarasova, O., Valentini, R., Wang, R., van der Werf, G., Wickland, D., Williams, M., and Zehner, C.: Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system, Biogeosciences, 11, 3547–3602, https://doi.org/10.5194/bg-11-3547-2014, 2014.

Wang, Y., Ciais, P., Broquet, G., Bréon, F.-M., Oda, T., Lespinas, F., Meijer, Y., Loescher, A., Janssens-Maenhout, G., Zheng, B., Xu, H., Tao, S., Gurney, K. R., Roest, G., Santaren, D. and Su, Y.: A global map of emission clumps for future monitoring of fossil fuel CO2 emissions from space, Earth Syst. Sci. Data, 11(2), 687–703, doi:10.5194/essd-11-687-2019, 2019.

Dogniaux, M., Crevoisier, C., Armante, R., Capelle, V., Delahaye, T., Cassé, V., De Mazière, M., Deutscher, N. M., Feist, D. G., Garcia, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Pollard, D. F., Roehl, C. M., Shiomi, K., Strong, K., Té, Y., Velazco, V. A. and Warneke, T.: The Adaptable 4A Inversion (5AI): description and CO2 retrievals from Orbiting Carbon Observatory-2 (OCO-2) observations, Atmos. Meas. Tech., 14(6), 4689–4706, doi:10.5194/amt-14-4689-2021, 2021.

Hasekamp, O. P., Fu, G., Rusli, S. P., Wu, L., Di Noia, A., aan de Brugh, J., Landgraf, J., Martijn Smit, J., Rietjens, J. and van Amerongen, A.: Aerosol measurements by SPEXone on the NASA PACE mission: expected retrieval capabilities, J. Quant. Spectrosc. Radiat. Transf., 227, 170–184, doi:https://doi.org/10.1016/j.jqsrt.2019.02.006, 2019.