Format Key

Reviewer comments in blue italics

Author responses in black

Reviewer 1: Anonymous reviewer #1, 09 Jan 2025

Checklist for reviewers: 1) Good 2) Good 3) Good

Manuscript should be: Accepted subject to technical corrections

Suggestions for revision or reasons for rejection:

In my opinion, the authors addressed all relevant points which were raised in the first round of discussions. While some of the details of how these points were addressed might be debatable, I do not see any major issues with the manuscript. Since these debatable details do not challenge the core message, findings, and results of the manuscript, I would be happy with some technical corrections or, if the handling editors considers this more appropriate, even a publication as is. Below, I provide a few comments, which do not need to be addressed necessarily and a few technical corrections for the authors to consider and I would be happy to leave all further discussions of the paper to the wider scientific community.

We thank Reviewer 1 for taking the time to read and provide valued feedback on this manuscript.

L. 170: I might be misreading this, but it is of course not the software which is operating at its maximum capacity – Bruker also uses this software for their high resolution instruments. The instrument simply operates at its (hardware limited) maximum optical path difference.

Indeed. We tweaked the wording to make it more clear that resolution is hardware-limited.

L. 180: Here maybe not necessary, but I generally like to also cite the 1977 erratum to the 1976 paper (appeared in J. Opt. Soc. Am., Vol. 67, No.3, March 1977, page 419) since the table with the quoted parameters change (which is typically the main reason somebody looks this up these days). Since this is often forgotten, even some textbooks have the table with the wrong values.

An interesting point. A reference to the erratum has been added to the manuscript.

L. 189: It just occurred to me when reading this: overfilling should in principle also reduce the (effective) field of view (FOV) and with that the instrument line function if it is limited by the FOV. Especially in cases with strong interfering absorbers (like here) a slightly wrong ILS might cause such issues as the "negative" amounts of HCHO. But it is likely difficult to judge if this actually happens.

Indeed. As part of ongoing work, the effect of fitting or fixing the FOV in the retrieval continues to be investigated, including correlations between HCHO, FOV and other fitted parameters.

L. 240: It could have been interesting to leave the actually distance when overfitting occurs a free fitting parameter. But maybe you had to little data constraining your fit.

Indeed. Fitting without constraining the overfilling distance was unsuccessful in our sparse dataset.

L. 301: Technically it should be a factor of sqrt(2)≈1.4 not 1.5.

This paragraph is a bit dense because both array size and acquisition time were changed between experiments. The factor of sqrt(2) is a result of the change in acquisition time (applies to the SNR), while the factor of 1.5 results from increasing the retroreflector area by 50% (applies to the IR signal). We tweaked the wording to stress this better.

L. 302ff: I would not think that the signal level is that sensitive to the orthogonal alignment of the retro array – that is why we use them after all. The water vapor hypothesis sounds more reasonable to me, but difficult to judge with the provided data. But I think it is also okay if this question is not fully resolved.

We also thought that the orthogonal alignment of the retro array would not be a critical factor, but in our later experiment we discovered that it mattered a lot more than we expected (probably in connection with the very long separation). Since at least two things were changing at once (alignment precision, water vapour), it is not possible to separate them from this dataset.

Reviewer 2: Anonymous reviewer #2, 17 Jan 2025

Checklist for reviewers: 1) Fair 2) Good 3) Good

Manuscript should be: Accepted subject to minor revisions

Suggestions for revision or reasons for rejection:

The study summarizes how the retrieval of formaldehyde (HCHO) measured by Open-Path Fourier Transform Infrared spectroscopy is influenced by different path lengths and retro reflector array sizes. Generally, the study can be divided into three parts, first of all spectral simulations for different open-path lengths were performed to determine the theoretical minimum path length for the detection of 1 ppb of formaldehyde given all other influencing variables are constant. Second, the impact of the retro reflector on the signal strength received at different path lengths has been analysed using data from two field campaigns. Finally, the retrieved formaldehyde concentrations for the two different array sizes at a fixed path length are evaluated, again using data from two field campaigns.

We thank Reviewer 2 for taking the time to thoroughly read and comment on the manuscript during this process.

General comments:

The manuscript is well structured and flows smoothly, making it easy to follow most of the time. Personally, I would suggest to add a table where the experiments throughout the years are listed and an indication of which of the two retro arrays and open path lengths were used, so the readers have central section to look up the details. Maybe also add the array size to the labels of the respective Figures or to the caption of each Figure.

This is a good suggestion and a table has been added to section 3.3 with a summary of the key acquisition data for each field experiment in this section. Figure captions were also updated to include information about array size (as 'large' or 'small').

Unfortunately, the measurements spanned over several years and the results are influenced not only by specifically introduced changes, but also by aging of the instruments IR light source and of the retro reflectors. Especially the improvement in Signal-to-Noise while using a larger reflector array could have been better highlighted by a more methodical experimental setup e.g., by changing the array size within one measurement campaign every day by covering parts of it for a few days. I acknowledge, that this will be time

consuming and might not be feasible, but I think this would make the results easier to compare.

We agree that multiple factors (including wear and tear on individual retro cubes) changing together is not desirable for strict experimental control. Indeed, it is not feasible to repeat the extended field work, which was designed with a scientific (and not technical) purpose in mind. We do have one very short set of spectra (minutes) with a part of the array covered and (as expected) the array with all cubes exposed leads to higher SNR spectra.

Nevertheless, the manuscript nicely shows how a larger reflector array can beneficial for larger open path lengths.

Thanks. The one-day experiments with variable path and increasing overfilling are most controlled in this regard.

Specific comments:

Lines 274ff: This contradicts the information presented in Figure 7 at the bottom, where only one datapoint of the "Otter lake" experiment shows a better SNR when the retro array is overfilled. Even though it is mentioned that the SNR is influenced by non-constant noise and that interpretation should be approached with caution. Maybe add an additional third panel to Figure 7, where the SNRs are normalized by the mean or median of the SNR of the underfilled retro array for each respective experiment. This would show the relative decrease of SNR caused by different path lengths and should indicate a slower decrease for the larger retro reflector array.

We highlighted the relative decrease of SNR in the original submitted manuscript, including lines of best fit, but these were removed in the revised manuscript because of the uncertainty due to non-constant noise. Dividing by the mean or median is an interesting suggestion, but each experiment has a different range of SNR values, hence showing such a normalization of the SNR could be misleading in its own way. In the end, having the reader infer (i.e., 'eyeball') the slower SNR decrease with a larger array (Otter Lake) seems like the most non-complicating presentation of this dataset.

Line 461ff: Would it be possible to add a small section about what you found the best compromise in your study?

A sentence has been added to the very end of the manuscript to summarize our findings on the optimum OPL for HCHO in our coastal environment.

Technical corrections:

The array sizes in Figure 1 and in the text are given in inches and I am not sure whether this fits with the AMT guidelines, which strictly state SI or SI derived units should be used.
Figure 8: The description still contains a half sentence of an earlier manuscript version, this should be removed.

All instances of inches have been replaced with centimeters. Figure 8 caption fixed.

Reviewer 3: Paton-Walsh, Clare, 18 Jan 2025

Checklist for reviewers: 1) Good 2) Good 3) Good

Manuscript should be: Accepted subject to technical corrections

Suggestions for revision or reasons for rejection:

I recommend that the article is accepted with a very minor tweak - which is to add a sentence or phrase at around line 15 of the Abstract stating upfront that formaldehyde is used as an example of a trace gas that has relatively weak absorption features at ambient concentrations and so is a challenge to retireve from OP-FTIR and therefore is sensitive to changes in the instrumental performance such as the path-length, beam divergence and array size.

We thank Dr. Paton-Walsh for her feedback for the improvement of this manuscript. A sentence has been added to the abstract of the manuscript to reflect that HCHO also serves as a proxy for any low abundance trace gas targeted for an OP- FTIR retrieval.