

## Format Key:

*Reviewer comments in blue italics.*

Author responses in black.

### RC2: Anonymous Referee #2, 23 Aug 2024

#### *General comments:*

*This paper describes a nice study to look at the impact of compensating for increasingly long pathlengths with larger arrays of retro-reflectors with the aim of improving the detection limits and precision of measurements of formaldehyde at typical ambient concentrations of 1ppb in the study area.*

*I have marked this as a minor revision but to address my concerns some extra analysis as well as discussion is required.*

We thank RC2 for their thorough reading of our submission and their thoughtful comments, which have improved our manuscript.

*The paper in its current form lacks some important discussion points, such as:*

- 1. The data of formaldehyde retrievals include a large proportion of negative concentrations, but the paper does not discuss any possible retrieval strategies that could be applied to help fit the interfering gases and minimise these negative retrievals.*

We have included a discussion of some retrieval strategies and checks (further down) that can be applied in order to minimize the occurrence of negative retrieval values. While it is part of ongoing work, we think that a retrieval optimization is out of scope for this paper, which is focused on path length and retroreflector size, while only characterizing the precision and (negative) bias of retrieved values. Optimizing the retrieval to minimize negatives is a non-trivial task. To our knowledge, one can start by examining: updated spectroscopic database parameters for the most prevalent interferer (water, but also other interferers) including temperature and pressure dependencies; correlations between the target gas and other retrieved parameters (interfering species, instrumental parameters, continuum parameters); retrievals that also include temperature as a retrieval target; multi-step retrieval approaches where water can be retrieved first and then constrained in the HCHO window; right up to even switching retrieval algorithms altogether to approaches that are mathematically constrained to return only positive values.

2. *Regardless of whether a better retrieval strategy could be designed, the formaldehyde retrievals could be analysed to determine an actual limit of detection for each of the measurement set-ups described so that a time-series of measurements of ambient concentrations above this detection limit could be provided. This really is required (in my opinion) to make this manuscript publishable.*

The reviewer raises an interesting point that cuts to the heart of the strengths and weaknesses of FTIR spectroscopic concentration measurements: they are quite precise and can nicely measure relative changes (like diurnal variations or pollution plumes), but their accuracy may be off by ~5% (e.g., Smith et al., 2011, AMT, for relatively high abundance CO<sub>2</sub>, CO and CH<sub>4</sub>); in the case of ‘threshold’ gases like formaldehyde (~1 ppb level, heavily overlapped by interfering species), these accuracy (systematic) issues may cause dips into negative concentration values in retrieval algorithms that are not mathematically constrained to yield positive values (like ours). While negative concentrations are unphysical, they still convey information on relative change in the whole timeseries (positives and negatives). Moreover, limiting reported values to those above a theoretically determined threshold (using only *random* noise), does not report on the bias (*systematic* or accuracy) problem.

Given this context, we quantified how the spectral signature of formaldehyde compares to typical *random* noise values in Section 3.1 and used that to determine a minimum path. While we can easily use our random noise value and the spectral signature at a given path to calculate the minimum detectable concentration of formaldehyde (it will be better than the 1 ppb assumed in the simulations of Figures 4 and 5), the negative bias problems are arising from *systematic* noise introduced by water vapour while overlapping the formaldehyde features (Figures 11 and 12), and we have no easy way to quantify that (Smith et al., 2011 used reference gas cells in a laboratory setting). We agree that a time series of positive formaldehyde values is required for formaldehyde *process studies*, but we believe that it is not required in this paper, and that a presentation of negative values is in fact instructive to consider when making choices on path length and array size given beam divergence and interfering gases.

3. *The discussion of the differential absorption appears to be discussed only as a %, whereas by my understanding MALT will fit the absorption area (not depth). Simply using a higher spectral resolution would provide a greater absorption depth and probably improve the detection limit and precision. Whilst this may not be possible with the equipment owned by the authors – it should be included in the discussion (since it will*

*also benefit the problem of interfering gases by improving the selectivity as the absorption lines become separated at higher resolutions).*

The reviewer is correct that MALT fits area and not depth. The instrument is operating at its maximum ‘mid-range’ resolution of  $0.5 \text{ cm}^{-1}$ . There is a limit to how much one can resolve the rotational structure in the horizontal open-path context because of strong pressure-induced line broadening processes, particularly for large molecules. We noted these points in Section 2.2.

### ***Specific comments:***

*I provide some more specific comments for some different sections of the paper below:*

#### ***Abstract***

*Line 16: We demonstrate that two-way path lengths  $> \sim 300 \text{ m}$  are necessary for robust HCHO spectral signatures (at typical random plus systematic noise levels.” But this MUST also depend upon the ambient concentration!! Some discussion about typical concentrations of formaldehyde in different environments would help set this context.*

The reviewer is correct. We included the assumed concentration in the abstract and the range of ambient values in Section 2.1.

*Line 19: We demonstrate that the larger retroreflector array resulted in a smaller decrease in the signal-to-noise ratio as a function of measurement path,  $\sim 1.5 \text{ m}^{-1}$  for the larger array as compared to  $\sim 3.6 \text{ m}^{-1}$  for the smaller array” – Won t this depend upon the field of view of the individual spectrometer and telescope? If so then this is quite a specific detail that probably doesn t belong in an Abstract.*

Agreed. The Abstract and manuscript was revised significantly in this respect, also in response to RC1.

*Line 23 : (average standard deviation of 0.352 ppb for 2021 and 0.678 ppb for 2018 in hourly formaldehyde data bins over two days), - This is also detail that doesn t belong in an Abstract. And referring to the data by the different years is a strange choice – when it is the path length and cube-corner array size that it what matters not the date.*

*Quoting the standard deviation to 3 significant figures seems over the top.*

We replaced the reference to year with array size in the Abstract and we replaced the standard deviation values with a factor ( $\sim 2x$ ) for the precision increase.

## Experimental Design

*Line 130 (p, T, precipitation, which causes IR beam extinction), Use of parenthesis is confusing here as the latter clause refers only to precipitation not temperature or pressure. Rephrase to clarify?*

Revised to “(p, T, and IR beam extinguishing precipitation)”.

## Results

*Line 229: Could the detector efficiency and/or pre-amp and amplifier gains also have decreased as the equipment aged? What is the difference in signal to noise at a part of the spectrum near to the formaldehyde absorption? Some of these factors will decrease the signal but not necessarily impact the signal to noise?*

We have revised this text to include detector aging as well, and the influence of alignment quality of the array in a plane orthogonal to the beam. It is difficult to estimate signal and noise accurately near the formaldehyde absorption because of highly variable interfering gases (there are no reliably ‘free’ regions of the spectrum near formaldehyde); we have clarified our discussion of how we estimate SNR instead.

*Figure 7: I am trying to understand the decision to calculate noise from a part of the spectrum outside of the detector response. I am not sure if this measure of signal to noise is the same as calculating it at a point in the spectrum with no absorption features but close to the wavenumber region where the gas of interest absorbs. Does this produce the same S:N? (Sorry if I am being slow – it has been a long week!) Maybe clarify this point in the text in any case??*

We have clarified the description of this calculation in the manuscript (also in response to RC1) and we also provided additional details why we think this is an equivalent method to assess noise within the response to RC1 (our response to their comment L359f). It is not possible to reliably check that both methods produce the same result because the spectrum is rather congested with interfering gases everywhere (e.g., the top panel of Figure 5 shows that water is everywhere present in the formaldehyde window). Before we settled on our approach, we attempted to calculate SNR from several apparently ‘clean’ spectral locations (including ones with full signal saturation) and found slightly differing results in all cases, with residual spectral structure after de-trending (which is absent outside of the detector response). However, in response also to RC1, we have modified the manuscript discussion to acknowledge that there is ‘noise in the noise’ and we abandoned fitting any trend lines to the decreasing SNR – while simply noting that SNR appears to decrease at a slower rate with the larger array.

*As well as the spectral signal to noise as discussed in and around Figure 7, there is the retrieval signal to noise. i.e. the spectrum to spectrum retrieved values give some idea of precision in a stable atmosphere. Given the values shown in Figure 9, I expect a discussion of the LOD here in terms of concentrations 3 x the retrieved value noise” – to determine where you have a clear detection of formaldehyde in the atmosphere.*

We quantify the measurement precision based on the scatter of retrieved values in Figure 9 and report how this varies in hourly bins in Figure 10. (These are the average standard deviations that we removed from the Abstract and replaced with a factor of 2 improvement (decrease) for the larger array). These would correspond to an LOD of ~0.35 ppb (large array) and ~0.69 ppb (small array) – *absent any systematic errors*, as we discussed in our response to General Comment 2 (RC2).

*Line 306 and 308: Are 3 and 4 significant figures justified here?*

We aimed for 3 decimal places, so 3 or 4 significant figures, which is admittedly high. We have changed values to 2 decimal places (2 or 3 sig figs), which is easier for a reader to absorb and compare (though also arbitrary).

**Citation:** <https://doi.org/10.5194/amt-2024-97-RC2>