On Path Length, Beam Divergence, and Retroreflector Array Size in Open-Path FTIR Spectroscopy

Cameron E. N. Power[1,2], Aldona Wiacek[1,2]

[1] Department of Astronomy and Physics, Saint Mary's University, Halifax, B3H 3C3, Canada
5   [2] Department of Environmental Science, Saint Mary's University, Halifax, B3H 3C3, Canada

Correspondence to: Cameron Power (cameron.power@smu.ca)

Abstract. Open-Path Fourier Transform InfraRed (OP-FTIR) spectroscopy is an established technique used to measure boundary layer trace gas concentrations, consisting (in this work) of a spectrometer with an active mid-IR source coupled to a single transmitting and receiving telescope, and a cube-corner retroreflector array separated from the spectrometer and

10   telescope by an atmospheric path. The detection limit is directly proportional to the optical path length in the atmosphere, which controls target gas spectral absorption depth; however, open-path beam divergence can lead to overfilling of the distant retroreflector array for one-way paths greater than ~300 m (details depend on specifics of spectrometer and telescope optics, plus array size), resulting in decreased returning radiation at the detector. In this case, the absorption signature of the target gas increases, but the signal to noise ratio of the recorded spectrum decreases. We present the results of theoretical spectral

15   simulations for formaldehyde (HCHO) that show how path length, interfering water concentration, and HCHO target concentration affect the expected differential absorption spectrum of the HCHO target. We demonstrate that two-way path lengths > ~300 m are necessary for robust HCHO spectral signatures (at typical random plus systematic noise levels). Next, we present the results of two field experiments where the retroreflector array area was increased to collect a larger fraction of returning radiation, at two-way path lengths ranging from 50 m to 1300 m. We demonstrate that the larger retroreflector array

20   resulted in a smaller decrease in the signal-to-noise ratio as a function of measurement path, ~1.5 m[-1] for the larger array as compared to ~3.6 m[-1] for the smaller array. Finally, we perform retrievals of HCHO concentrations from spectra collected at the same field site and path length in Halifax Harbour during 2018 and 2021, with a smaller and a larger retroreflector array, respectively. We demonstrate that retrievals based on larger retroreflector array spectra exhibit higher precision (average standard deviation of 0.352 ppb for 2021 and 0.678 ppb for 2018 in hourly formaldehyde data bins over two days), even though

25   systematic errors remain in the fitted spectra, due to water vapour. Where systematic fitting errors in interfering species are significant, a longer path may not be optimal for a given target gas, leading instead to biased retrievals; moreover, at very long pathlengths signal-to-noise ratio decreases with increasing water vapour conditions due to broadband spectrum signal reduction effects in water-saturated regions. We discuss factors to consider in the choice of path length and retroreflector array size in open-path FTIR spectroscopy, which must be made with care.

1 Introduction

1.1 OP-FTIR technique applications and implementation

Open-Path Fourier Transform InfraRed (OP-FTIR) spectroscopy is an established technique used to measure boundary layer trace gas concentrations, historically in fenceline monitoring of industrial emissions (e.g., Schill et al., 2022), near traffic emission sources (e.g., You et al., 2017), and in biomass burning events (e.g., Paton-Walsh et al., 2014). It has recently been used to monitor greenhouse gas emissions from agricultural fields (e.g., Flesch et al., 2016; Lin et al., 2019; Bai et al., 2019) and in urban settings (Byrne et al., 2020), also during the COVID-19 pandemic (You et al., 2021). Additionally, OP-FTIR has been applied in measurements of marine shipping port emissions (Wiacek et al., 2018), air-sea greenhouse gas exchange (Wiacek et al., 2021; Hellmich, 2022) and volcanic emissions (e.g., Pfeffer et al., 2018).

Most details of the system used in this study have been previously described (Wiacek et al., 2018) and here we only recount the general features of a monostatic OP-FTIR configuration, which consists of a spectrometer (containing an active infrared source, the interferometer, and an infrared detector), transfer optics, telescope, and a retroreflector array separated from the spectrometer by a one-way measurement path of (typically) a few hundred meters (the optical path travelled by the light is double the measurement path). Beyond this distance, the returning signal diminishes and measurement quality is degraded (see Section 1.3). The source generates mid-infrared light that is collimated and modulated in the spectrometer, then expanded by the telescope and sent toward the retroreflector array; the expanded beam passes through the atmosphere along the measurement path until it reaches the retroreflector, where it is reflected back towards the telescope and spectrometer for detection as a time-domain interferogram. In this configuration, unmodulated stray light from the thermally emitting atmosphere is easily rejected as a constant offset at the detector and the time-domain interferogram is subsequently Fourier-transformed into an atmospheric absorption spectrum.

Underlying trace gas concentrations are derived by iteratively fitting a simulated spectrum to a measured spectrum until the residual is minimized; the details of our retrieval process have been described elsewhere (Wiacek et al., 2018, and references therein). The retrieval result is a path-average trace gas concentration, which means that measurements are representative of the full atmospheric path, as opposed to a point, while retaining high precision. This is an inherent advantage of the open-path system in ambient atmospheric measurements: it provides a more representative sampling of the broader measurement location and minimizes the effect of localized emissions.

1.2 Retroreflector array overview and modifications

Retroreflector arrays are constructed from multiple smaller cube-corner retroreflector elements, which can have various coatings depending on the wavelength of light the system utilizes. In the case of infrared light, the most commonly used coatings are gold and silver, due to their high reflectivity in the infrared spectral region (Bennett, 1965). Each cube-corner

retroreflector reflects incoming light back toward the source in a slightly translated (~6 cm) but parallel path. This is accomplished by combining three flat mirrors orthogonally (Fig. 1, left), which results in the incoming beam being reflected twice to achieve a direction change of 180°, i.e., a retro-reflection (Fig. 1, right). Typically, commercially available arrays consist of either 30 or 60 cube-corner elements for either 12" or 24" array sizes.

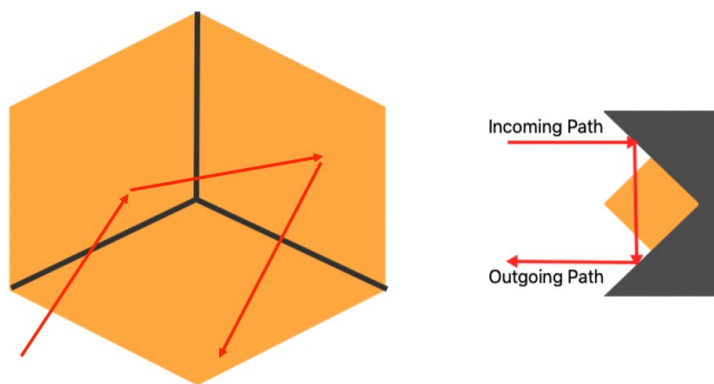



Figure 1: (Left) single cube-corner retroreflector, front view. (Right) single cube-corner retroreflector, side view, with example incoming beam and two reflections shown by red arrows.

70

To construct two larger than off-the-shelf retroreflector arrays for our work, 60 new cube-corner retroreflectors were purchased in 2020, each with a 63.5 mm outer diameter and nominal angle of acceptance of 30°. The beam deviation tolerance from parallel (after complete retro-reflection) is 20 arcsec (0.10 mrad). Each cube-corner retroreflector element was gold-coated (including a protective dielectric coating) to achieve 97% IR reflectivity. The pristine cube corners arrived at Saint Mary's University pre-mounted onto two custom array panels (30 new cubes on each panel, as required by the scientific investigation described in Wiacek et al., 2021) to which older cube corners (63 mm diameter, bare-gold) were added from two existing 24" off-the-shelf retroreflector arrays. The smaller and older retroreflector array is shown in Fig. 2 (left), while the larger retroreflector array constructed using the older cube corners (degraded somewhat from ~180 days of cumulative field use since 2015) and the 30 pristine cube corners (representing a 50% increase in area) is shown in Fig. 2 (right). Due to the relatively high cost of pristine cube corners (~$300 USD each in 2020), the larger array could not be constructed out of entirely newly sourced cube corners. Note that the difference in cube substrate design resulted in a gap in close-packing the larger array (new cubes have a glass-substrate while old cubes have a metal substrate).

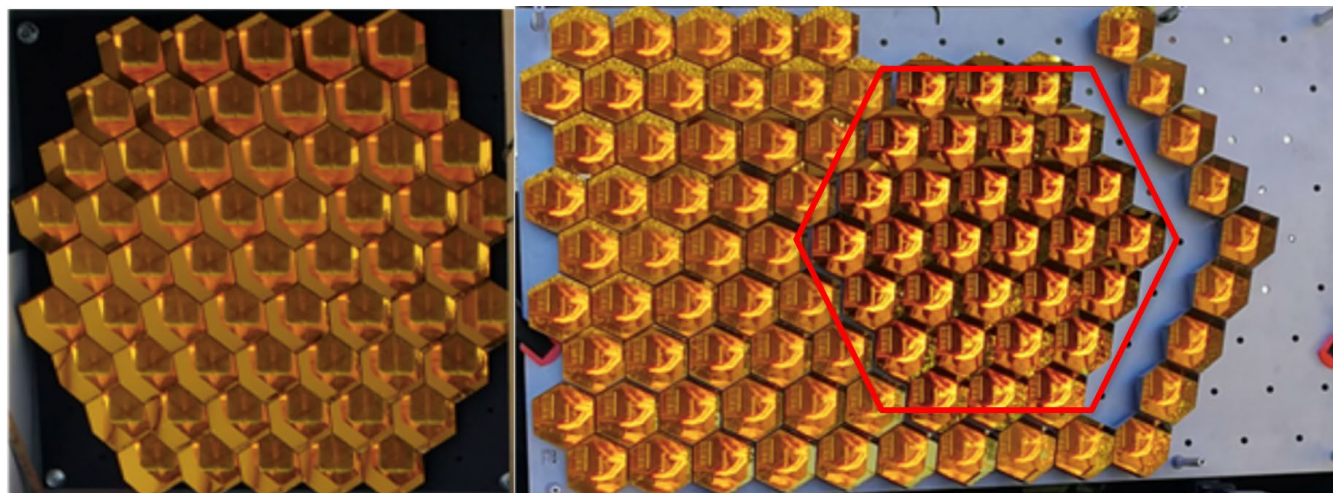



Figure 2: (Left) pristine bare-gold retroreflector array (59 elements) acquired off-the-shelf in 2015. (Right) larger retroreflector array constructed in 2020 from 30 new and 59 used cube corners (new elements enclosed by red hexagon). The old and new cubes are not identical in substrate design and so could not be perfectly close-packed, resulting in a gap (right).

1.3 Path length, beam divergence and retroreflector array overfilling

A critical variable in OP-FTIR spectroscopy is the optical path length, as it inversely correlates with the detection limit of the system; a longer path implies greater absorption and thus lower detection limits (Griffiths and de Haseth, 2007). At sufficiently long path lengths, however, beam divergence due to the necessarily imperfect beam collimation inside the spectrometer causes a decrease in returning signal due to overfilling of the distant retroreflector array (Fig. 3). In this simplified schematic of a mono-static arrangement, the co-located source (not shown), spectrometer (not shown) and telescope (blue cylinder) are on the left, while the retroreflector array (24" or 48") is on the right. The bi-directional red arrows show the IR radiation travelling along the open path, with a greatly exaggerated beam divergence. The larger array returns all – or at least *more* – divergent radiation from farther away towards the detector, depending on system and setup details, whereas the smaller array would be overfilled at that same distance. The spectrometer (3 mm aperture, 69 mm focal length) plus telescope (9:1 reducing) used in this experiment produces a 12" collimated beam, with an effective beam divergence of ~1 mrad observed in the open path in field measurements, resulting in retroreflector array overfilling at a one-way separation greater than ~300 m for our 24" retroreflector array. Finally, we note that array size cannot be increased *ad infinitum* in practice because as the beam hits the array at increasingly off-axis angles the cube corners begin to shade themselves and one another, leading to diminishing signal returns. This should be possible to mitigate in principle by a non-flat, e.g., a concave array, however construction complexity and cube-corner cost are prohibitive in practice.

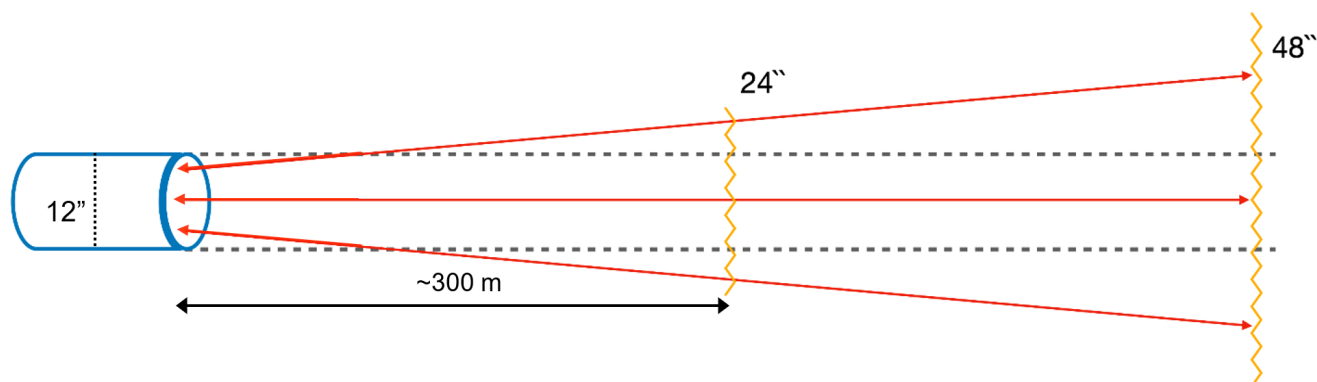



Figure 3: OP-FTIR beam divergence (exaggerated) after leaving a 12" diameter telescope, leading to overfilling of a 24" retroreflector array in our system configuration (see text for details).

1.4 Overview of study

The remainder of this paper is structured as follows. Section 2 outlines our experimental design, which first reviews the theoretical expectations for the absorption of formaldehyde (HCHO) in Section 2.1. Formaldehyde was chosen as the focus of our analysis because it is an important but challenging-to-measure ambient atmospheric constituent that is not reported by the ground-based National Air Pollutant Surveillance (NAPS) network in Canada. Nevertheless, it is detectable from space, e.g., by the TROPOMI instrument (October 2017 launch, 5.5 x 3.5 km^2 resolution after Aug 2019) and now by the geostationary TEMPO instrument (April 2023 launch, 2 x 4.75 km^2 resolution at centre of field of regard). TEMPO is expected to operationally produce hourly total columns of HCHO by October 2023. Further motivating the pursuit of more precise and diurnal surface HCHO products from OP-FTIR is the fact that surface HCHO concentrations have been derived from space-based total column measurements with the help of models (Zhu et al., 2017), but this is challenging work – both at the surface and from space (e.g., Vigouroux et al., 2020). Next, Section 2.2 describes two field experiments where path length was systematically increased and signal strength was analysed under conditions of different retroreflector array size. Section 2.3 discusses the setup of retrievals of formaldehyde concentration from field spectra recorded at the same very long path (and field site) but with different retroreflector array sizes (at different times). To the best of our knowledge, this is the first report of explicit tests of retroreflector array size in open-path measurements that vary path length, considering both theoretical expectations and field results, the latter incorporating systematic issues that arise in very long open-path experiments. Section 3 presents the results corresponding to Section 2, i.e., for spectral simulations, for increasing path in field experiments, and for increasing retroreflector array size in field experiments. Finally, Section 4 closes with a summary and key conclusions.

2 Experimental Design

2.1 Theoretical spectral simulations

130   The detection of trace gases by OP-FTIR depends on many variables besides the target gas concentration, such as the optical path length, interfering gas concentrations, environmental conditions (p, T, precipitation, which causes IR beam extinction), and instrumental parameters. Simulations were conducted for the target gas formaldehyde to determine the theoretical minimum path required for detection at 1 ppb (a baseline atmospheric concentration). The path was increased from 50 to 1500 m while all other parameters listed in Table 1 were held constant, which resulted in stronger absorptions by both target and

135   interfering gases. The increase in the target gas was isolated by calculating the differential absorption due to the target gas as the difference between a spectrum with the target and interfering gases, and another simulated spectrum with only the interfering gases (also using Table 1 input parameters, except setting the target concentration to 0 ppb). We performed these simulations using the forward model in the Multiple Atmospheric Layer Transmission (MALT) NLLS retrieval suite (Griffith, 1996; Griffith et al., 2012), which uses the HITRAN spectral database (Rothman et al., 2013) for line parameters.

140

Table 1: List of input parameters used in spectral simulations.

Target gas	formaldehyde (1 ppb, 0 ppb)
Spectral Window	2700 cm^{-1} - 2900 cm^{-1}
Pressure	1013.25 mb
Temperature	15 °C
Path length	50 m - 1500 m
Path step increase	50 m
Interfering gases	water (1%), methane (2 ppm), nitrous oxide (330 ppb)

145 (Temperature)

2.2 Impact of increasing path in field experiments on measured signal levels

150

To empirically determine the point at which effective beam divergence results in overfilling of a retroreflector for a particular system, an experiment can be conducted where open-path spectral measurements are made at successively increasing path lengths. This experiment also gives an indication of the overall effectiveness of the system at increasing path lengths for detecting a particular trace gas in a given spectral window while incorporating the practical effects of the increasing absorptions

155   due to any interfering species (usually water vapour). In May 2015, such an experiment was conducted with a smaller retroreflector array constructed of, at the time, 59 pristine bare-gold cube-corner retroreflectors. The retroreflector array was progressively moved further from the spectrometer starting at 50 m and increasing to a separation of 450 m in 100-m increments. This resulted in two-way separations of 100-900 m, where, as already mentioned, a two-way separation is double the physical separation between the retroreflector array and the spectrometer, representing the distance travelled by the beam

160 (optical path). At each increment, spectroscopic measurements were made and the return signal level was recorded to assess the effects of beam divergence. In October 2020, this experiment was repeated with the newly constructed larger (89 element) retroreflector array. We used separations from 50 m to 1575 m (two-way) in approximately 200 m increments, with the goal of assessing the benefits of the larger retroreflector array area, specifically with regard to overfilling.

165 At each location and path, we followed the manufacturer's recommended iterative procedure to focus the telescope on the retroreflector array (first a z-adjustment of the secondary mirror (in-out from primary mirror) followed by an x-y adjustment of the telescope (left-right, up-down), repeated at least once to maximize IR signal levels). The OP-FTIR system used was a Bruker Open Path System, comprising an internal IR source, 12" Schmidt-Cassegrain telescope, and a broadband IR detector responsive between 700 cm^{-1} and 6000 cm^{-1}. Data acquisition was handled by the spectroscopic software OPUS, developed by
170 Bruker, operating at a spectral resolution of 0.5 cm^{-1} (Wiacek et al., 2018).

2.3 Impact of increased retroreflector array area on retrieved concentrations of HCHO

To further analyze the effect of the larger retroreflector array area on measurements, formaldehyde retrievals were performed on OP-FTIR absorption spectra recorded at the same location in the summer of 2018 (using the smaller retroreflector array) and the summer of 2021 (using the larger retroreflector array). These measurements were conducted across Halifax harbour,
175 with a two-way measurement path of ~1120 m (Wiacek et al., 2021). As discussed above, the data acquisition of these measurements was conducted using the proprietary software from Bruker, OPUS RS, at a resolution of 0.5 cm^{-1}, then apodized with the Norton-Beer 'medium' function (Norton and Beer, 1976). Trace gas retrieval were performed with the Multiple Atmospheric Layer Transmission forward model (Griffith, 1996) and a non-linear least squares iterative fitting routine. Spectra used as inputs to retrievals were filtered for IR intensity (at 2100 cm^{-1}) greater than 0.15 arbitrary units. These filtering criteria
180 exclude poor weather conditions such as fog or heavy rain, which are the main cause of poor retrieval results, as well as occasional poor fits due to, e.g., poor pressure and temperature inputs. Spectral retrievals were performed between 2745 cm^{-1} and 2800 cm^{-1}, with interfering species H_2O, CH_4 and N_2O, which all have absorptions at this path length. Pressure and temperature data was obtained on an hourly basis from an Environment and Climate Change Canada meteorological station 7 km away. Other fitted parameters included the background transmission (a first order polynomial) and instrumental parameters
185 of phase and shift. Instrumental field of view and effective apodization parameters were held constant, as in the work of Paton-Walsh et al. (2014). Retrieval results were filtered for root mean square (RMS) residual values less than 0.01 (1% mean fitting error)

3 Results

3.1 Theoretical spectral simulations

190   Fig. 4 shows the effect of increasing the optical path length (which starts at 50 m and increases up to 1500 m in 50-m increments) on the combined absorption spectrum of the target and interfering gases. The peak absorption is ~55%, corresponding to water features at approximately 2720 cm^{-1}. Absorption increases with increasing path length and so at shorter two-way path lengths (< ~300 m) the water features reach only ~15% absorption. This means that detection limits are theoretically better at longer path lengths (if noise is constant), with the precise minimum two-way path different for each

195   individual target gas (in the absence of any interfering species). Yet the effect of interfering gases cannot be ignored in practice and the simulated differential absorption of only the target gas is next shown in Fig. 5.

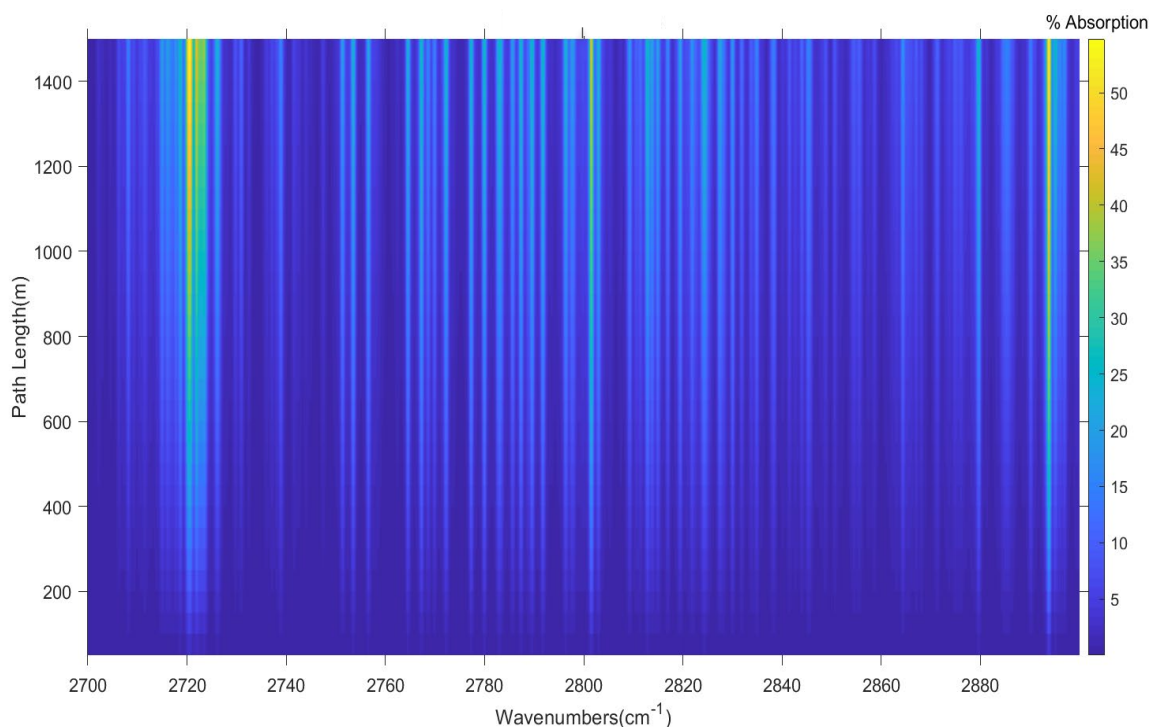



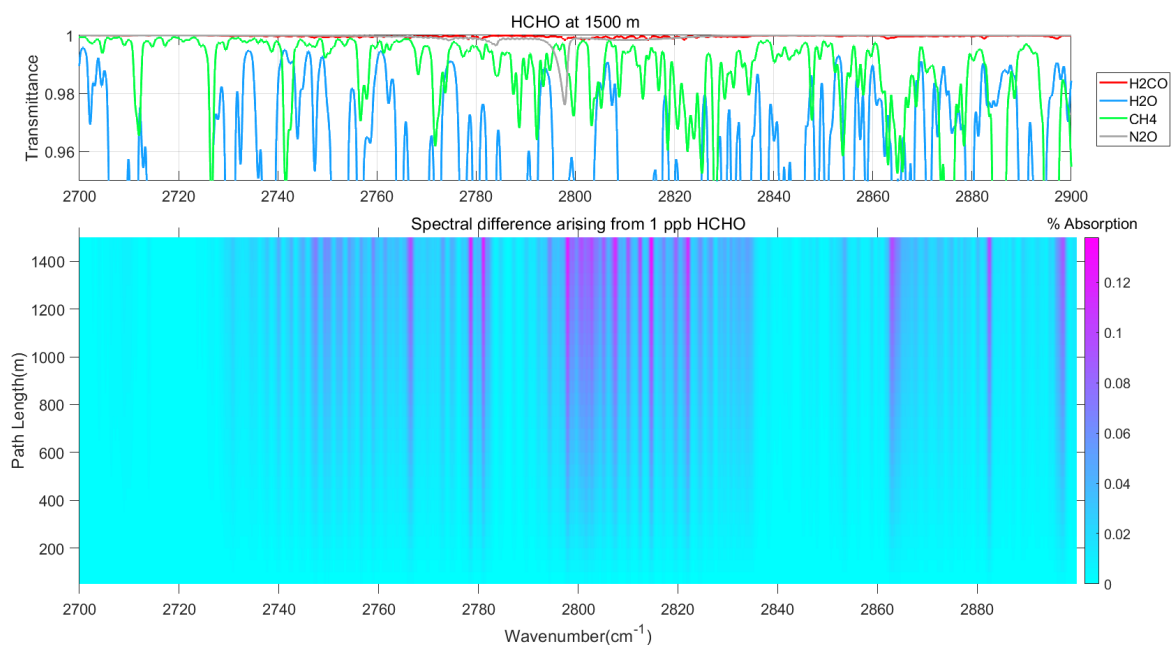
Figure 4: Percent absorption due to 1 ppb formaldehyde and interfering species (see Table 1) as a function of varying optical path.

200

In Fig. 5, the transmittance spectrum is cropped to see the spectral features of formaldehyde, which only reach a minimum transmittance of 0.9985 at maximum path difference. The peak differential absorption formaldehyde features (brightest pink bands) occur mainly between 2780 cm^{-1} to 2820 cm^{-1}. We see that short two-way path lengths (< ~300 m) provide an approximate differential percent absorption due to HCHO alone of < 0.06%, which is nevertheless technically large enough

205   for formaldehyde to be detected in the presence of interfering gases because the random spectral noise is ~3x lower than this

value at these separations (i.e., ~0.015%). The differential percent absorption of formaldehyde alone reaches a maximum of just 0.12% at the maximum path length of 1500 m, which is still ~5x higher than the random noise at these separations (i.e., ~0.023%). This means that at longer path lengths the interfering gases are not obscuring the target gas, i.e., the pink differential absorption bands intensify toward the maximum simulated path difference. However, it must be noted that this analysis

210 considers random noise only, and that systematic noise due to, e.g., spectroscopic parameter errors is always an issue in spectral fit residuals. For robust detection to be possible, the individual yet spectrally correlated features from the target gas should at least be comparable to the RMS noise levels of the retrieval, which include both random and systematic components.

215 **Figure 5: (Bottom panel) differential absorption due to 1 ppb formaldehyde in the presence of interfering gases (see Table 1) as a function of varying path. (Top panel) corresponding individual transmittance spectra for target and interfering gases (at the maximum path length of the bottom panel).**

3.2 Impact of increasing path in field experiments on measured signal levels

220 In the 2015 experiment ("Franklyn Street", smaller retro), it was determined that between 150 m and 250 m one-way separations between the spectrometer and retroreflector array (two-way optical paths of 300 m and 500 m), the retroreflector became overfilled and the IR signal intensity was progressively reduced at increasing paths up to a maximum of 450 m one-way (Fig. 6). In the 2020 experiment ("Otter Lake", larger retro), the signal began decreasing noticeably also between two-

way optical paths of 300–500 m but the rate of decrease in 2020 was half the rate of the 2015 experiment, as shown by the
225   slopes of best-fit lines in Fig. 6.

Apart from the expected slower signal decrease for a larger array, the arbitrary signal level maxima are 25% lower at Otter
Lake (~0.69, 100 m) as compared to Franklyn Street (~0.92, 100 m) despite the 50% larger retro array. The cause for this
decrease in signal levels is likely a combination of factors, including the higher number of pristine cubes in 2015 and a decrease
230   in the Globar™ infrared source intensity by ~10% between 2015 and 2020.

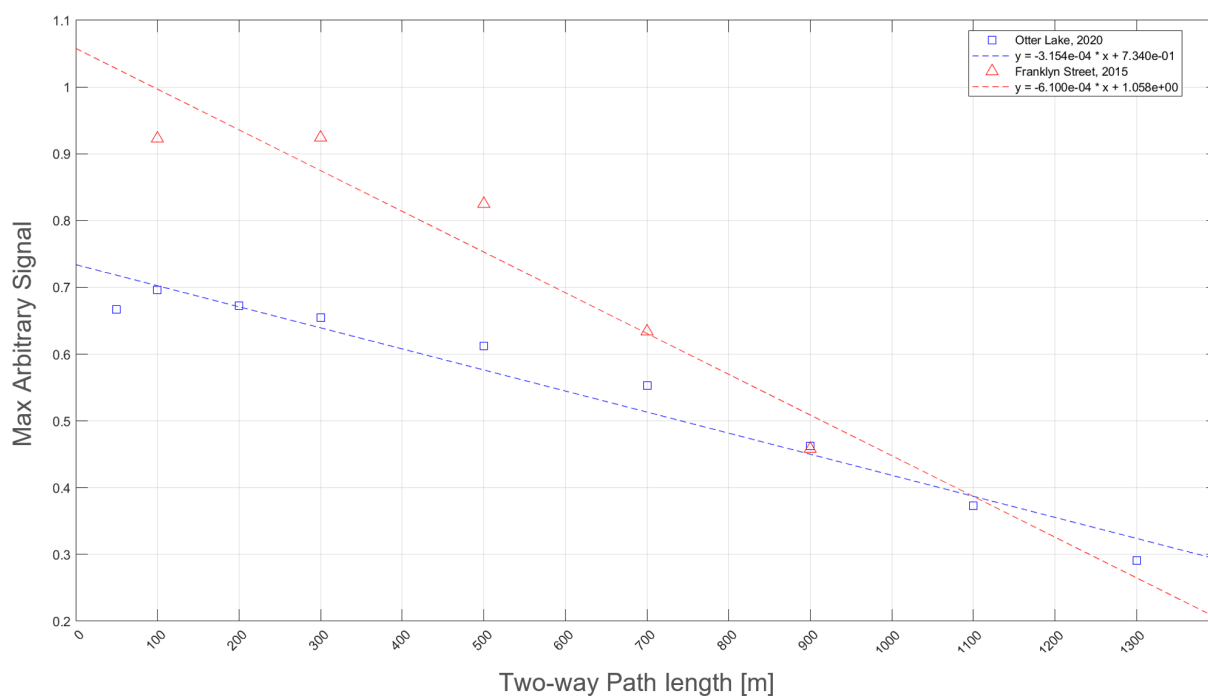



Figure 6: Maximum arbitrary IR signal intensity at 2500 cm⁻¹ for varying optical path lengths at Franklyn Street (May 27, 2015, 59 cube-corner retroreflector) and Otter Lake (October 16, 2020, 89 cube-corner retroreflector), together with lines of best fit.

235

To determine the instrumental noise in each experiment, the standard deviation was calculated after de-trending the raw spectra
in the region of 7640 cm⁻¹ to 7740 cm⁻¹ (outside of instrumental response). The de-trending removed any change in the mean
signal of the region, thus making the signal in the region vary about zero so that a true random noise value could be calculated.
The spectral region used to estimate instrumental noise is shown in Fig. 7 (top) for the Otter Lake experiment. The signal
240   remaining after de-trending has random noise characteristics and, upon closer examination (not shown), there are no correlated
spectral absorption features in this spectral region between spectra at different measurement paths (as expected). Note that the
same method of noise calculation was also applied to the Franklin Street data in the same region, but this is not shown. This
noise was used to calculate the signal-to-noise ratio (SNR) for both experiments using the maximum spectral signal at each

245   path length and dividing by the instrument noise. As expected, the longer paths, where the retroreflector is overfilled and the return signal is being lost, exhibit lower SNR values. The SNR values determined at both measurement locations are also shown in Fig. 7 (bottom), which shows that the SNR values at Otter Lake decrease at half the rate of the Franklyn Street values, based on the slopes of the lines of best fit (~ -1.5 m^{-1} vs. ~ -3.6 m^{-1}). This results in SNR values at Otter Lake being greater at longer path lengths, greater than ~ 800 m.

250

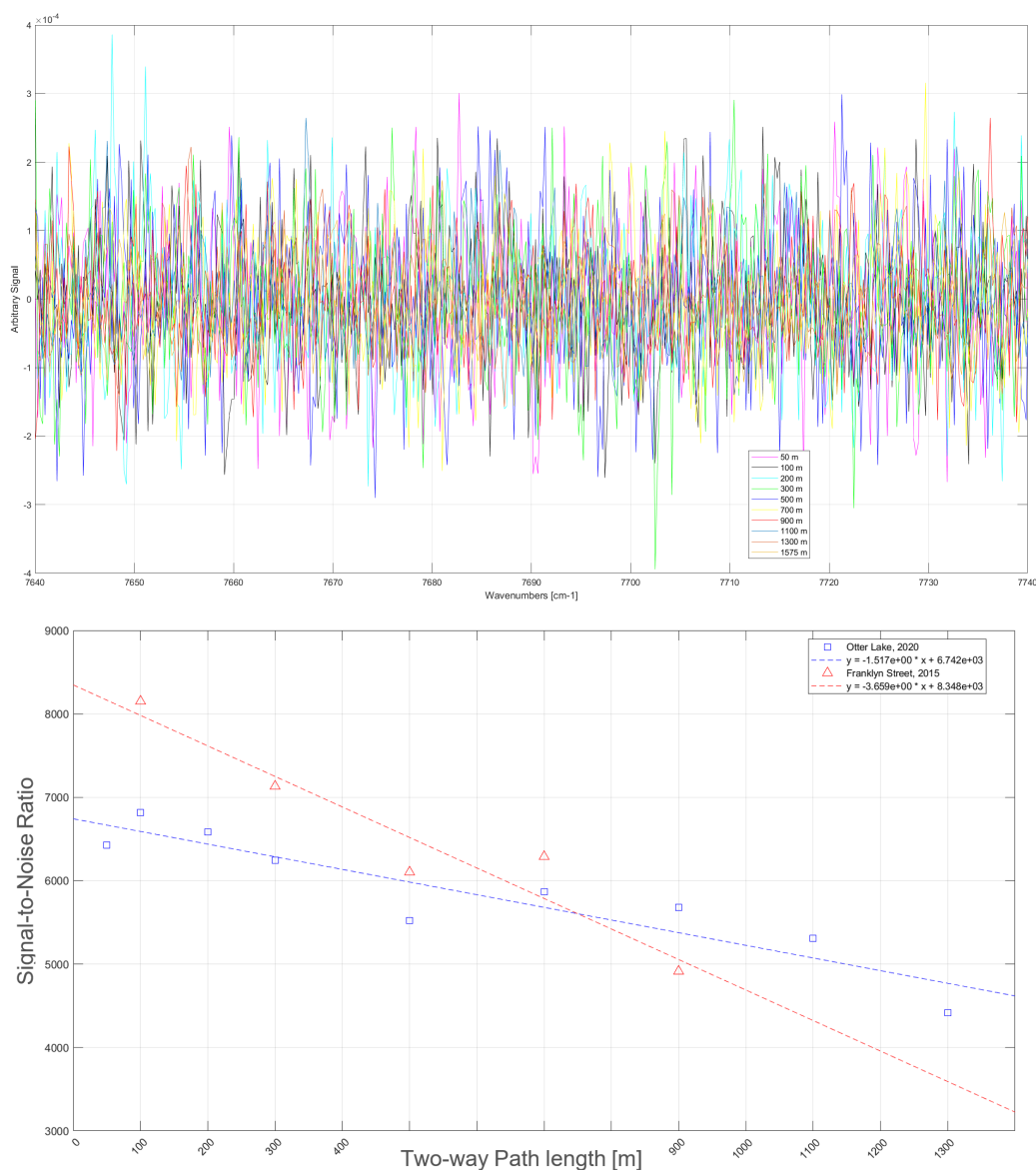



Figure 7: (Top) De-trended spectral signatures in a region outside of instrumental response at Otter Lake, which correspond to noise values. (Bottom) SNR values at Franklyn Street and Otter Lake, with lines of best fit.

3.3 Impact of increased retroreflector array area on retrieved concentrations of HCHO

Before analysing the impact of the lager retroreflector array used in 2021 on HCHO retrievals, the impact of different acquisition durations must also be accounted for. This is because while any increases in retrieved concentration precision are influenced by a larger retroreflector array (and higher SNR), they are also influenced by the different acquisition times used in each campaign (1 minute in 2018 and 2 minutes in 2021), with 2021 data thus having reduced spectral noise by $\sqrt{2}$ since $SNR \propto t^{1/2}$ (Griffiths and de Haseth, 2007). We account for the effect of the different acquisition times and isolate the effect of the larger retroreflector on the measured spectra using spectra from a two-month period of each dataset, starting in March of 2021 and August of 2018. The sampling geometry for each of these experiments was identical, with the larger retroreflector in 2021 being the only notable difference. These months represent the period when the infrared intensity was at its peak for each field campaign, before retroreflector wear and tear set in. In Fig. 8, the SNR, calculated using the same method as in Fig. 7, is plotted for the two-month IR peak in 2018 and in 2021, along with lines of best fit over an arbitrary time index. These are average daily SNR values, filtered to remove measurements with low IR intensity due to poor weather. Additionally, the SNR values from 2021, adjusted for the longer acquisition time (scaled down by a factor of $\sqrt{2}$), are also plotted with a line of best fit. The mean SNR value in 2021 is ~3x higher than the mean SNR value in 2018 (~ 5400 vs. ~ 1950), while the adjusted 2021 mean SNR is approximately double the mean SNR of 2018 (~ 3800 vs. ~ 1950). This shows that even when accounting for the longer acquisition times in 2021, the quality of measurements is still doubled by the larger retroreflectors.

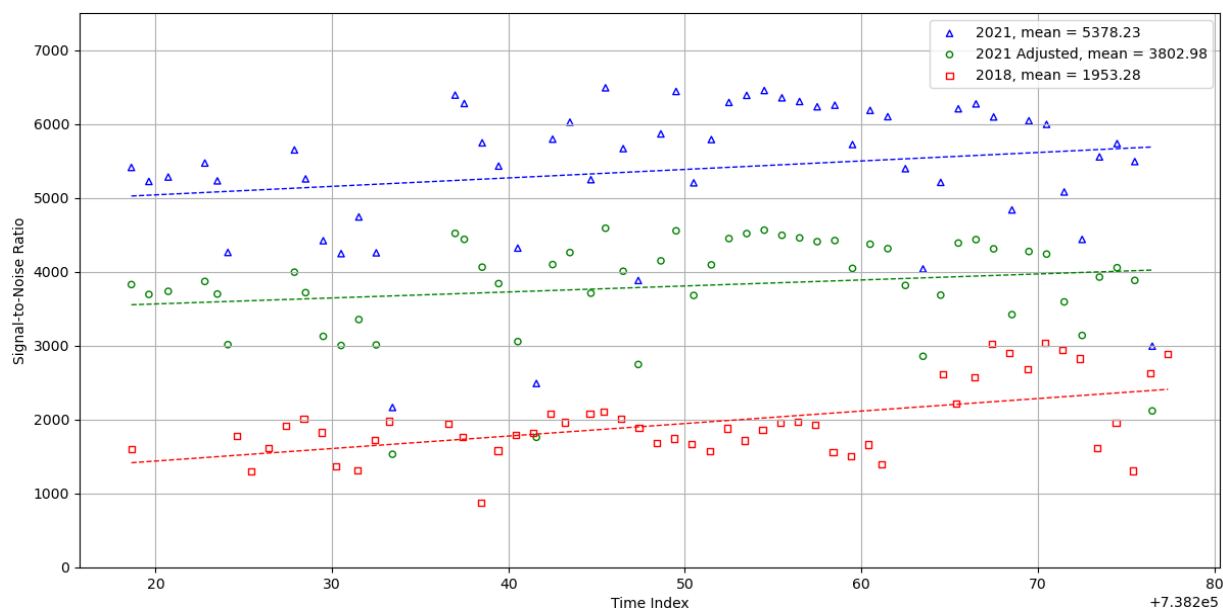



Figure 8: Daily mean signal-to-noise ratio in two-months of spectra in 2021, 2018, and "adjusted 2021" (see text for details), with lines of best fit shown.

275

Now we discuss the practical (retrieval) effect of a larger retroreflector array on the retrievals of formaldehyde from absorption spectra recorded with the same sampling geometry. The field experiments in 2018 (smaller array) and in 2021 (larger array) were both conducted over a 1120 m one-way path length, however, the data was gathered during largely non-overlapping seasons. To determine the best comparison timeframe between 2018 and 2021, we looked for conditions of similar specific

280 humidity due to the pervasive effects of water vapour in infrared spectra and on retrieved HCHO. Specific humidity was calculated using the relative humidity, temperature, and pressure (sourced from the Environment Canada station in Shearwater, ~7 km from the field location, and also used in the retrieval of trace gas concentrations). Additionally, we were interested in time periods with similar temperatures and thus the two-week period from the 1st to the 14th of September in 2018 and the 1st to the 14th June in 2021 was selected.

285

Fig. 9 shows the formaldehyde concentration retrieved for September 2018 and June 2021, with colour showing the infrared intensity at 2100 cm^{-1}. In 2021 both the acquisition time was longer (2 vs. 1 minutes or 480 vs. 240 co-adds at 4 Hz) and the sampling strategy was different in that 2021 observations were made by alternatively pointing to two different retroreflector arrays. This means that 2021 observations yield four times fewer spectra in each hour, which can be seen clearly in Fig. 9.

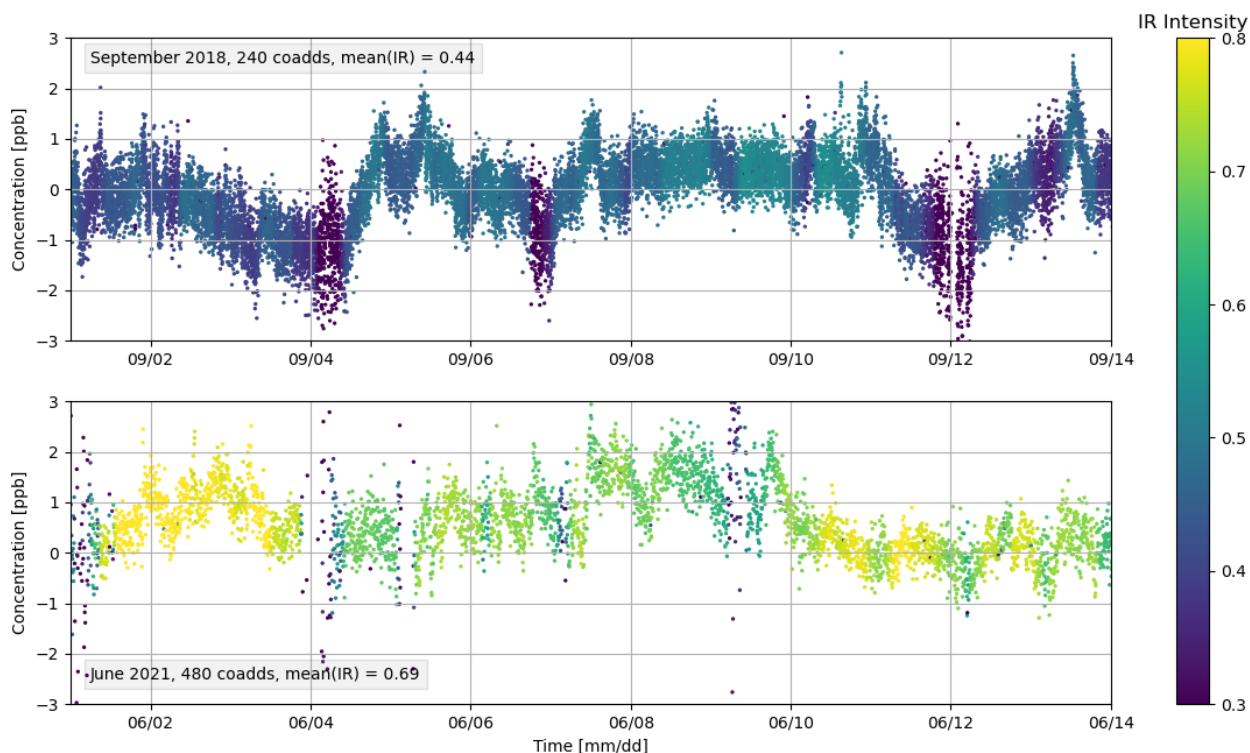
290 Infrared intensity is, on average, higher in 2021 than in 2018 by ~50% (~0.69 vs. ~ 0.44) due to the larger retroreflector resulting in greater returning radiation, and a higher SNR. Correspondingly, the scatter (precision) of the 2021 formaldehyde retrievals is smaller (better) as compared to the scatter of the 2018 retrievals, which is quantified next in Fig. 10.

295

Figure 9: Retrieved formaldehyde [ppb] for (top) September 2018 and (bottom) June 2021, where colour represents spectral infrared intensity in both time periods. 2018 measurements used the same field configuration but smaller retroreflectors than 2021 measurements. The acquisition length and sampling rate were different in each year (see text for details). See text for temporal sampling details.

300

Fig. 10 shows hourly mean retrieved HCHO concentrations (as boxplots) on September 12-13 in 2018 and on June 12-13 in 2021, chosen as days free from obvious ship emissions in our open path, which is traversed by ships. An hourly time frame was selected to minimize the impact of physical processes, such as daily photochemical variations or small and nearby HCHO emissions, on the spread of the hourly means, i.e., so that it corresponds more closely to the true measurement precision. The 1st and 3rd quartiles of the 2021 data are closer to each other when compared to the 2018 data, meaning that the majority of the data points fall within a narrower range (average interquartile range of 0.471 ppb for 2021 and 1.048 ppb for 2018). Additionally, the greater precision of the 2021 dataset is reflected in the smaller average hourly standard deviation over the two-day span (0.352 ppb versus 0.678 ppb), even though there are four times fewer spectra in each hour in 2021 and one may expect worse precision on account of this sampling difference. Due to a lower precision, there is no discernible diurnal pattern in 2018, while higher precision 2021 data shows some of the expected concentration minima during the night (0000 h - 0700 h) and maxima during the day (1000 h – 1800 h) over the two days examined. Lastly, 2021 data exhibits fewer negative retrieved HCHO concentrations over a two-day period, which we discuss next.

305

310

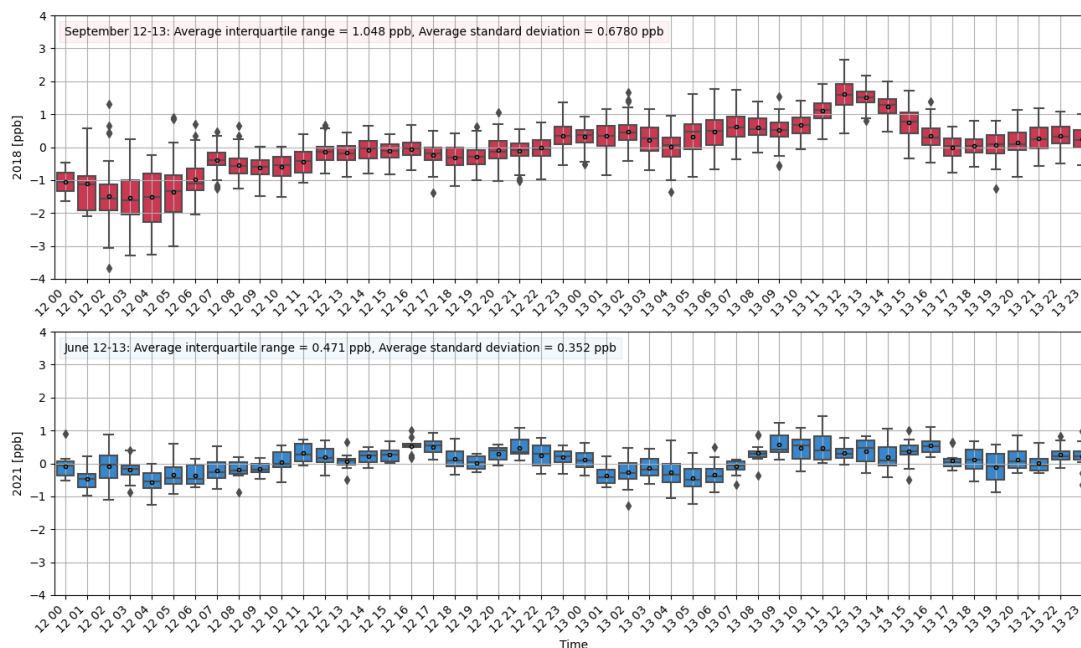
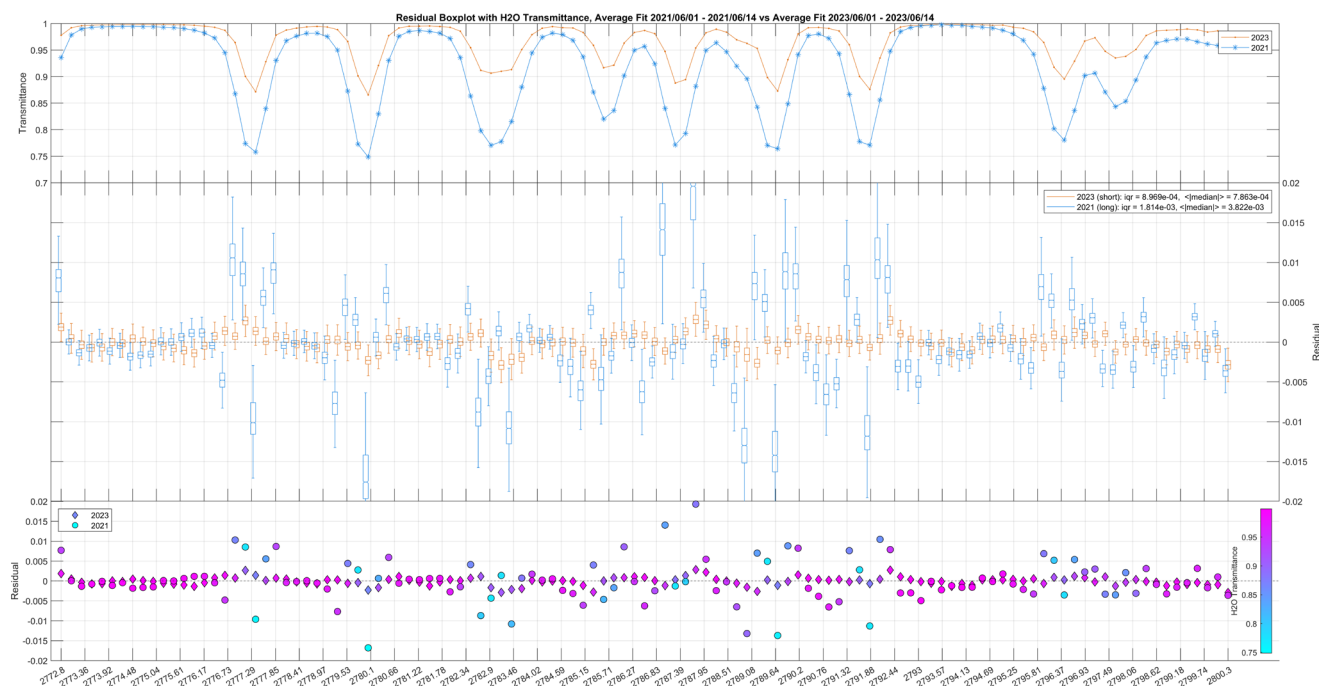



Figure 10: Formaldehyde hourly mean concentrations [ppb] from (top) September 12 - 13, 2018 and (bottom) June 12 – 13, 2021. Horizontal lines represent median concentrations, circles represent mean concentrations, upper and lower box edges represent 75th and 25th concentration percentiles ("q3" and "q1", respectively), upper and lower whiskers represent maximum and minimum concentration limits (defined as +/-2.7σ (approx. 99.3 percentile)), and grey diamonds are concentration outliers.

3.3.1 Impact of H2O on retrieved concentrations of HCHO

As discussed in previous sections (1.3), longer measurement paths allow for greater absorption of trace gases. This results in a lower detection limit for a low-abundance target gas such as formaldehyde, however, this is also coupled with an increase in the absorption of any interfering gases, such as water vapour. Greater absorption by water has an impact on the retrieval process via both systematic and random errors, which is something we noticed in analysing data from 2023 (not discussed up to now) that contained far fewer negative formaldehyde concentrations despite a lower spectral SNR and shorter path (i.e., lower formaldehyde absorption signal). First, a high absorption due to water increases the systematic misfitting to the measured spectrum. Fig. 11 shows the average H_2O transmittance (top), spectral residuals as boxplots (middle), and spectral residuals coloured by H_2O transmittance (bottom) for a two-week period in June of 2021 and again in June of 2023; for clarity, only half of the retrieval window is shown (2772.8 – 2800.3 cm^{-1}). The measurements in 2023 were recorded over a short two-way path of 440 m *vs.* 1120 m in 2021 and 2018. (This short path is consistent with the minimum two-way path for ~3x noise level detection of formaldehyde that we identified in simulations in Section 3.1.) The H_2O transmittance in 2021 and 2023 (top panel) shows a similar structure as expected, however, the transmittance depth in 2021 is consistently double the transmittance depth in 2023, overwhelmingly due to water vapour absorption features. The average spectral residuals (as boxplots for a two-

week period, middle panel) show the increase in systematic misfitting at the longer path. The average of the (absolute value) median residuals at each point in the spectral window for the shorter 2023 path is five times smaller than that of the longer

335   2021 path (7.863e-4 and 3.822e-3, respectively), i.e., the spectral residual for the shorter path varies less about zero and is lower overall. Additionally, the average interquartile range of the short-path dataset is half that of the long-path dataset (8.969e-4 and 1.814e-3, respectively), meaning that there is less variance in the middle half of the short-path dataset residuals.

340   **Figure 11: All plot panels are based on data from June 1 - 14, 2021 and June 1 - 14, 2023 (Top) Average H_2O transmittance spectrum for 2021 (blue) and 2023 (orange). (Middle) Spectral residuals from 2021 (blue) and 2023 (orange). Boxplot features as in Figure 10. The legend gives the mean value of the (absolute value) medians and the mean interquartile range in the spectral region shown. (Bottom) Average spectral residuals for 2021 (circles) and 2023 (diamonds), with colour indicating H_2O transmittance.**

345   In the case of a low absorption gas like formaldehyde, a high abundance of water can result in a retrieved concentration that is negative, which is particularly pronounced during the summer months. Fig. 12 shows the retrieved formaldehyde concentration as a function of retrieved water vapour concentration (as boxplots) for 2021 and 2023, binned into 0.25% H_2O increments. In both datasets the majority of measurements have a retrieved water concentration of 1.0% or less (~70% in 2021 and ~80 % in 2023); in these retrievals the median and mean retrieved formaldehyde concentration is positive. When the retrieved water

350   concentration exceeds 1.0%, the 2021 median and mean retrieved formaldehyde concentrations become negative, while in 2023 they remain positive. The mean SNR in the 2021 dataset is higher than the corresponding mean SNR in the 2023 dataset for each box because in 2021 the retroreflector array was both bigger and comprised pristine cubes. Nevertheless, ~40% of the formaldehyde concentrations retrieved in the 2021 dataset are negative, while less than ~2% are negative in the 2023 dataset.

This is very likely a result of the lower H2O transmittance in the 2021 dataset compared to the 2023 dataset (Fig. 10, top) and
355 the associated higher systematic misfitting of these water features (Fig. 10, middle) as the well-known spectroscopic database
uncertainties are amplified under the long path. Lastly, we discuss how water impacts random errors. At the longer path used
in this study (2021, blue, bold text insert) the broadband spectral effect of water vapour is large enough to systematically
reduce the SNR as water vapour abundance increases, whereas at the shorter path (2023, orange, bold text insert) SNR is
largely independent of water vapour. This is another mechanism by which pathlength reduces SNR, thereby increasing random
360 errors, in addition to by overfilling of the retroreflector array.

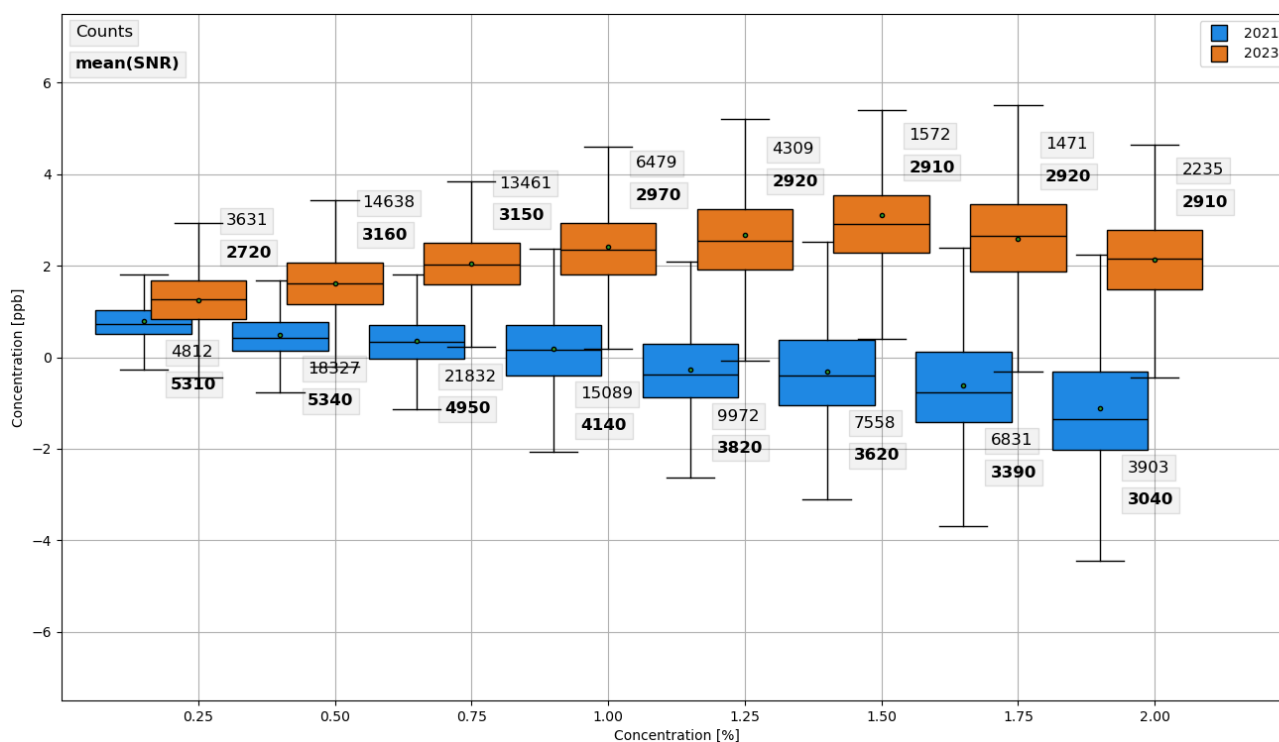



**Figure 12: Retrieved formaldehyde [ppb] plotted as a function of retrieved water vapour [%] from 2021 (blue) and 2023 (orange),
with 2021 data offset for clarity. Text inserts give the samples per box on the top and the mean SNR per box on the bottom (in bold).**
365 **Boxplot features as in Figure 10.**

4 Summary and Conclusions

Open-Path Fourier Transform InfraRed (OP-FTIR) spectroscopy is an established technique used to measure boundary layer
trace gas concentrations. Diurnal measurements are possible at high temporal frequency, making OP-FTIR data suitable for
the validation of space-based measurements of, e.g., daytime formaldehyde by the TEMPO geosynchronous instrument or day
370 and night ammonia by the CrIS sun-synchronous instrument. A critical variable in OP-FTIR spectroscopy is the optical path
length, as it inversely correlates with the detection limit of the system, where a longer path implies greater absorption and thus

lower detection limits. At sufficiently long path lengths, however, beam divergence due to the necessarily imperfect beam collimation inside the spectrometer causes a decrease in returning signal due to overfilling of the distant retroreflector array. To the best of our knowledge, this is the first report of explicit tests of retroreflector array size in open-path measurements that vary path length, considering both theoretical expectations and field results, the latter incorporating systematic issues that arise in very long open-path experiments.

We performed spectral simulations for formaldehyde wherein the path was increased from 50 m to 1500 m while all other simulations parameters remained constant. Increased path resulted in stronger absorptions by both target and interfering gases, as expected, with the differential absorption due to the target gas maximized away from its interfering species between 2730 cm^{-1} and 2830 cm^{-1}. Based on consideration of absorption strength vs random error levels alone (spectral SNR) short path lengths (< ~300 m, two-way) result in an approximate differential percent absorption due to HCHO alone of < 0.06%, which allows formaldehyde detection at ~3x the noise level in our system. At 1500 m the differential formaldehyde signal reaches just 0.12% due to transmittance nonlinearity, which is still ~5x higher than our system's random noise at these separations, however, at such long path lengths systematic errors in water vapour spectroscopy play an important role in formaldehyde retrieval results.

In 2015 and in 2018, we made OP-FTIR field measurements at increasing path lengths using a smaller retroreflector array and a 50% larger retroreflector array, respectively. From these field experiments, we determined that the SNR values from the measurements made with the larger retroreflector array decreased as a function of array-spectrometer separation at less than half the rate of the experiment with the smaller retroreflector. This directly correlated to the behaviour of the maximum spectral signal in the two experiments, as the maximum signal of the larger retroreflector experiment decreased at approximately half the rate of the smaller retroreflector experiment, while the noise remained constant.

In 2018 and in 2021, we made OP-FTIR field measurements at the same long path and location, but with a smaller and 50% larger retroreflector array, respectively. After accounting for the longer spectral acquisition time in 2021, we found the mean daily average SNR in a two-month period corresponding to the highest IR intensity for both experiments (March 2021 and August 2018) to be greater by a factor of ~2 in 2021 data using the larger retroreflector array.

Lastly, we compared retrieved formaldehyde concentration from the same measurement configuration during the first two weeks of September 2018 and the first two weeks of June 2021 based on similar environmental conditions (water vapour and temperature). Given the higher SNR in 2021, the scatter (precision) of the 2021 formaldehyde retrievals is smaller (better) as compared to the scatter of the 2018 retrievals. In examining the hourly mean formaldehyde concentration statistics over two days, we find an average interquartile range of 0.471 ppb for 2021 and 1.048 ppb for 2018; the average hourly standard deviation over the two-day span was 0.352 ppb for 2021 and 0.678 for 2018 ppb, even though there are four times fewer

spectra in each hour in 2021 and one may expect worse precision on account of this sampling difference. Due to the higher precision, the expected diurnal formaldehyde pattern can be discerned in 2021, but not in 2018, and 2021 data exhibits fewer negative retrieved HCHO concentrations than 2018 data over a two-day period. We also found that a 2023 short-path dataset with lower spectral SNR and lower formaldehyde absorption signal outperformed the 2021 dataset with far fewer negative

410 retrieved formaldehyde concentrations (2% vs. 40%). By analysing the spectral fit residuals we attribute this to spectral misfitting of strong water vapour absorptions in 2021, as the well-known spectroscopic database uncertainties are amplified under the very long path conditions (a systematic error). Finally, we document a second mechanism by which pathlength reduces SNR (associated with random errors), in addition to by overfilling of the retroreflector array: at very long pathlengths SNR decreases with increasing water vapour conditions due to broadband spectrum signal reduction effects in water-saturated

415 regions.

In summary, the choice of path length and retroreflector array size in open-path FTIR spectroscopy must be made with care. Longer paths increase target gas absorption (lowering detection limits) and larger retroreflector arrays improve the SNR of the spectra by increasing the return signal (improving retrieved concentration precision), but there are limitations to both. Open-

420 path beam divergence of a given system determines the size of arrays needed to reduce overfilling at a given separation (driven by the detection limits of a particular target gas), but array size is constrained by prohibitive cost and the self-shading of array elements (if it was to remain flat). Pathlength is constrained indirectly by 1) the accuracy of the spectroscopic parameters of water vapour, which may lead to systematic fitting errors and negative target gas concentrations, and 2) the reduced SNR due to broadband water vapour absorption. Nevertheless, there is an optimum array size and path combination to find in each

425 specific observational environment and application, as explored in this work.

Author contributions. AW a conceived of and designed the study, with input from CP. Both authors performed data acquisition and retrievals, examined and interpreted the results and prepared the manuscript.

430 *Competing interests.* The contact author has declared that none of the authors has any competing interests.

Acknowledgements. The authors thank members of the Wiacek Atmospheric Research Group for assistance with OP-FTIR field deployment in 2018 and 2021 (Taylor Gray, Morgan Mitchell, Ian Ashpole) as well as Taylor Gray for developing the initial version of the differential optical absorption simulation codes.

435

Financial support. This research has been supported by the Canadian Space Agency (grant no. 16SUASMPTN), the Canadian National Sciences and Engineering Research Council through the Discovery Grants Programme (grant no. RGPIN-2014-03888, RGPIN-2022-05225), and Saint Mary's University.

References

440   Bai, M., Suter, H., Lam, S. K., Flesch, T. K., and Chen, D.: Comparison of slant open-path flux gradient and static closed chamber techniques to measure soil N2O Emissions, Atmospheric Measurement Techniques, 12(2), 1095–1102, https://doi.org/10.5194/amt-12-1095-2019, 2019.

Bennett, J. M., and Ashley, E. J.: Infrared reflectance and emittance of silver and gold evaporated in ultrahigh vacuum, Applied
445   Optics, 4(2), 221, https://doi.org/10.1364/ao.4.000221, 1965.

Byrne, B., Strong, K., Colebatch, O., You, Y., Wunch, D., Ars, S., Jones, D. B., Fogal, P., Mittermeier, R. L., Worthy, D., and Griffith, D. W. T.: Monitoring urban greenhouse gases using open-path Fourier transform spectroscopy, Atmosphere-Ocean, 58(1), 25–45, https://doi.org/10.1080/07055900.2019.1698407, 2020.

450

Flesch, T. K., Baron, V. S., Wilson, J. D., Griffith, D. W. T., Basarab, J. A., and Carlson, P. J.: Agricultural gas emissions during the spring thaw: Applying a new measurement technique, Agricultural and Forest Meteorology, 221, 111–121, https://doi.org/10.1016/j.agrformet.2016.02.010, 2016.

455   Griffith, D. W., Deutscher, N. M., Caldow, C., Kettlewell, G., Riggenbach, M., and Hammer, S.: A Fourier transform infrared trace gas and isotope analyser for Atmospheric Applications, Atmospheric Measurement Techniques, 5(10), 2481–2498, https://doi.org/10.5194/amt-5-2481-2012, 2012.

Griffiths, P. R., and de Haseth, J. A.: Fourier transform infrared spectrometry (2nd ed.), Wiley-Interscience, 2007.

460

Griffith, D.: Synthetic calibration and quantitative analysis of gas-phase FT-IR spectra, Applied Spectroscopy, 50(1), 59–70, https://doi.org/10.1366/000370296390662, 1996.

Hellmich, M.: Ocean-air N2O trace gas fluxes in Halifax Harbour derived from open-path FTIR measurements, MSc thesis,
465   Saint Mary's University, Canada,133 pp., 2022.

Lin, C., Grant, R., Heber, A., and Johnston, C.: Application of open-path fourier transform infrared spectroscopy (op-FTIR) to measure greenhouse gas concentrations from Agricultural Fields, Atmospheric Measurement Techniques, 12(6), 3403–3415, https://doi.org/10.5194/amt-12-3403-2019, 2019.

470

Norton, R. H., and Beer R.: New apodizing functions for Fourier spectrometry, J. Opt. Soc. Am., 66, 259-264, 1976.

Paton-Walsh, C., Smith, T. E. L., Young, E. L., Griffith, D. W. T., and Guérette, É.-A.: New emission factors for Australian vegetation fires measured using open-path Fourier transform infrared spectroscopy – Part 1: Methods and Australian temperate forest fires, Atmos. Chem. Phys., 14, 11313–11333, https://doi.org/10.5194/acp-14-11313-2014, 2014.

Pfeffer, M., Bergsson, B., Barsotti, S., Stefánsdóttir, G., Galle, B., Arellano, S., Conde, V., Donovan, A., Ilyinskaya, E., Burton, M., Aiuppa, A., Whitty, R., Simmons, I., Arason, Þ., Jónasdóttir, E., Keller, N., Yeo, R., Arngrímsson, H., Jóhannsson, Þ., Butwin, M., Askew, R., Dumont, S., von Löwis, S., Ingvarsson, Þ., La Spina, A., Thomas, H., Prata,F., Grassa, F., Giudice, G., Stefánsson, A., Marzano, F., Montopoli M., and Mereu L. : Ground-based measurements of the 2014–2015 Holuhraun volcanic cloud (Iceland), Geosciences, 8(1), 29, https://doi.org/10.3390/geosciences8010029, 2018.

Power, C., Hanley, J., Hellmich, M., and Wiacek, A.: On Path Length, Beam Divergence, and Retroreflector Size in Open-Path FTIR Spectroscopy, American Geophysical Union Fall Meeting (2020), San Francisco, Virtual, United States, https://agu.confex.com/agu/fm20/meetingapp.cgi/Paper/767625, 2020.

Rothman, L.S., Gordon, I.E., Babikov, Y., Barbe, A., Benner, D., Bernath, P.F., Birk, M., Bizzocchi, L., Boudon, V., Brown, L.R., Campargue, A., Chance, K., Cohen, E.A., Coudert, L.H., Devi, V.M., Drouin, B.J., Fayt, A., Flaud, J.-M., Gamache, R.R., Harrison, J.J., Hartmann, J.-M., Hill, C., Hodges, J.T., Jacquemart, D., Jolly, A., Lamouroux, J., Le Roy, R.J., Li, G., Long, D.A., Lyulin, O.M., Mackie, C.J., Massie, S.T., Mikhailenko, S., Müller, H.S.P., Naumenko, O.V., Nikitin, A.V., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E.R., Richard, C., Smith, M.A.H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G.C. Tyuterev, Vl.G., and Wagner, G. : The HITRAN 2012 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 130, 4-50, https://doi.org/10.1016/j.jqsrt.2013.07.002, 2013.

Schill, S. R., McEwan, R. S., Moffet, R. C., Marrero, J. E., MacDonald, C. P., and Winegar, E. D.: Real-World Application of Open-Path UV-DOAS, TDL, and FT-IR Spectroscopy for Air Quality Monitoring at Industrial Facilities, Spectroscopy Outside The Laboratory, 37(11), 18-22, https://doi.org/10.56530/spectroscopy.qz5173x6, 2022.

Wiacek, A., Li, L., Tobin, K., and Mitchell, M.: Characterization of trace gas emissions at an intermediate port, Atmospheric Chemistry and Physics, 18(19), 13787–13812, https://doi.org/10.5194/acp-18-13787-2018, 2018.

Wiacek, A., Hellmich, M., and Flesch, T.: Application of Open-Path Fourier Transform Infrared (OP-FTIR) Spectroscopy to Air-Sea Greenhouse Gas Exchange, OSA Optical Sensors and Sensing Congress 2021, Paper JW4D.1, https://doi.org/10.1364/AIS.2021.JW4D.1, 2021.

You, Y., Byrne, B., Colebatch, O., Mittermeier, R. L., Vogel, F., and Strong, K.: Quantifying the impact of the COVID-19 pandemic restrictions on CO, CO2, and CH4 in downtown Toronto using open-path fourier transform spectroscopy, Atmosphere, 12(7), 848, https://doi.org/10.3390/atmos12070848, 2021.

510

You, Y., Staebler, R. M., Moussa, S. G., Su, Y., Munoz, T., Stroud, C., Zhang, J., and Moran, M. D.: Long-path measurements of pollutants and micrometeorology over Highway 401 in Toronto, Atmos. Chem. Phys., 17, 14119–14143, https://doi.org/10.5194/acp-17-14119-2017, 2017.

515   Vigouroux, C., Langerock, B., Bauer Aquino, C. A., Blumenstock, T., Cheng, Z., De Mazière, M., De Smedt, I., Grutter, M., Hannigan, J. W., Jones, N., Kivi, R., Loyola, D., Lutsch, E., Mahieu, E., Makarova, M., Metzger, J.-M., Morino, I., Murata, I., Nagahama, T., Notholt, J., Ortega, I., Palm, M., Pinardi, G., Röhling, A., Smale, D., Stremme, W., Strong, K., Sussmann, R., Té, Y., van Roozendael, M., Wang, P., and Winkler, H.: TROPOMI–Sentinel-5 Precursor formaldehyde validation using an extensive network of ground-based Fourier-transform infrared stations, Atmos. Meas. Tech., 13, 3751–3767, 520   https://doi.org/10.5194/amt-13-3751-2020, 2020.

Zhu, L., Jacob, D. J., Keutsch, F. N., Mickley, L. J., Scheffe, R., Strum, M., González Abad, G. Chance, K., Yang, K., Rappenglück, B., Millet, D. B., Baasandorj, M., Jaeglé, L., and Shah, V.: Formaldehyde (HCHO) As a Hazardous Air Pollutant: Mapping Surface Air Concentrations from Satellite and Inferring Cancer Risks in the United States, Environmental 525   Science & Technology, 51 (10), 5650-5657, https://doi.org/10.1021/acs.est.7b01356, 2017.