Intercomparison of open-path trace gas measurements with two dual-frequency-comb spectrometers

Eleanor M. Waxman¹, Kevin C. Cossel¹, Gar-Wing Truong¹, Fabrizio R. Giorgetta¹, William C. Swann¹, Sean Coburn², Robert J. Wright², Gregory B. Rieker², Ian Coddington¹, and Nathan R. Newbury¹

¹Physical Measurement Laboratory, National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80305, USA
²Precision Laser Diagnostics Laboratory, University of Colorado Boulder, Boulder, CO 80309, USA

Correspondence to: Eleanor M. Waxman (eleanor.waxman@nist.gov)

Received: 8 March 2017 – Discussion started: 14 March 2017
Revised: 28 June 2017 – Accepted: 3 August 2017 – Published: 11 September 2017

Abstract. We present the first quantitative intercomparison between two open-path dual-comb spectroscopy (DCS) instruments which were operated across adjacent 2 km open-air paths over a 2-week period. We used DCS to measure the atmospheric absorption spectrum in the near infrared from 6023 to 6376 cm⁻¹ (1568 to 1660 nm), corresponding to a 355 cm⁻¹ bandwidth, at 0.0067 cm⁻¹ sample spacing. The measured absorption spectra agree with each other to within 5 × 10⁻⁴ in absorbance without any external calibration of either instrument. The absorption spectra are fit to retrieve path-integrated concentrations for carbon dioxide (CO₂), methane (CH₄), water (H₂O), and deuterated water (HDO). The retrieved dry mole fractions agree to 0.14 % (0.57 ppm) for CO₂, 0.35 % (7 ppb) for CH₄, and 0.40 % (36 ppm) for H₂O at ~ 30 s integration time over the 2-week measurement campaign, which included 24 °C outdoor temperature variations and periods of strong atmospheric turbulence. This agreement is at least an order of magnitude better than conventional active-source open-path instrument intercomparisons and is particularly relevant to future regional measurements on sampling towers is complicated by local wind patterns, local sources, and mixing within the planetary boundary layer (PBL), especially at night (Lauvaux et al., 2008, 2012; Ciais et al., 2010). Similarly, total-column measurements are particularly useful for sub-continental to global-scale measurements; however they are sensitive to atmospheric transport errors within the PBL. (Lauvaux and Davis, 2014), are affected by clouds and aerosols, are primarily limited to daytime measurements, and lack either the revisit rates or mobility for regional flux measurements. Horizontal integrated path measurements are complementary to point sensors and satellites: they cover spatial scales from one to tens of kilometers, provide measurements on second-to-minute timescales with portable instruments, and are thus appropriate for regional studies. Active-source open-path sensors such as open-path Fourier transform spectroscopy (FTS), differential optical absorption spectroscopy (DOAS), differential lidar (DIAL), or tunable diode laser absorption spectroscopy (TDLAS) are often used for these measurements and can retrieve path-averaged concentrations but typically with 10 % or greater uncertainties (EPA, 2017, and references therein). Recently, open-path dual-comb spec-

1 Introduction

Quantitative determination of greenhouse gas fluxes over a variety of temporal and spatial scales is necessary for characterizing source strength and intermittency and for future emissions monitoring, reporting, and verification. To this end, techniques exist to measure greenhouse gas concentrations on a variety of length scales, each of which has advantages and disadvantages. Point sensors provide valuable information about local sources, but their use for continuous regional measurements on sampling towers is complicated by local wind patterns, local sources, and mixing within the planetary boundary layer (PBL), especially at night (Lauvaux et al., 2008, 2012; Ciais et al., 2010). Similarly, total-column measurements are particularly useful for sub-continental to global-scale measurements; however they are sensitive to atmospheric transport errors within the PBL. (Lauvaux and Davis, 2014), are affected by clouds and aerosols, are primarily limited to daytime measurements, and lack either the revisit rates or mobility for regional flux measurements. Horizontal integrated path measurements are complementary to point sensors and satellites: they cover spatial scales from one to tens of kilometers, provide measurements on second-to-minute timescales with portable instruments, and are thus appropriate for regional studies. Active-source open-path sensors such as open-path Fourier transform spectroscopy (FTS), differential optical absorption spectroscopy (DOAS), differential lidar (DIAL), or tunable diode laser absorption spectroscopy (TDLAS) are often used for these measurements and can retrieve path-averaged concentrations but typically with 10 % or greater uncertainties (EPA, 2017, and references therein). Recently, open-path dual-comb spec-
troscopy (DCS) has emerged as a new technique that could potentially provide precise, accurate, continuous regional measurements of the mole fractions of CO$_2$, CH$_4$, H$_2$O, and HDO over kilometer-scale open paths (Rieker et al., 2014), thereby providing a new open-path sensing capability that falls between point sensing and total-column measurements.

Here we demonstrate that open-path DCS can indeed yield dry mole fractions over open-air paths with a high level of intercomparability, over long periods of time, and with sufficient precision to track variations in the ambient levels from local sources and sinks. Two completely independent open-path DCS instruments are operated over neighboring open-air paths during a 2-week measurement campaign. Although both DCS instruments use fully stabilized frequency combs, they are portable (Truong et al., 2016) and are operated nearly continuously during both day and night through laboratory temperature variations from 17 to 25°C, strong atmospheric turbulence, and outdoor air temperature variations from 4.6 to 28.9°C. The retrieved dry mole fractions for the two DCS instruments agree to better than 0.57 ppm$^1$ (0.14%) for CO$_2$ and 7.0 ppb (0.35%) for CH$_4$. This agreement is achieved without any “bias correction” or calibration of either instrument for absolute wavelength or for absolute concentration. Instead, it is a direct consequence of the negligible instrument line shape and precise frequency calibration of the DCS instruments, which leads to measured atmospheric absorption spectra that are identical to below 10$^{-3}$ and as low as 2.5 × 10$^{-4}$ (limited by the instrument noise level). The measured path-averaged CO$_2$ precision over a 2 km path is 0.90 ppm in 30 s, improving to 0.24 ppm in 5 min. For CH$_4$, the precision is 9.6 ppb in 30 s, improving to 2.1 ppb in 5 min. We also compare the DCS retrievals to a cavity ring-down spectroscopy (CRDS) point sensor located near the path that has been tied to the World Meteorological Organization (WMO) manometric scale (for CO$_2$) and gravimetric scale (for CH$_4$) through calibration with WMO-traceable gases. The agreement is within 3.4 ppm and 17 ppb for CO$_2$ and CH$_4$, respectively, limited by differences in the sampling volume and by the spectral database used to analyze the DCS transmission spectra.

Similar intercomparison measurements between conventional active open-path sensors are rare but have shown agreement of typically 1–20% (Thoma et al., 2005; Hak et al., 2005; Smith et al., 2011; Shao et al., 2013; Conde et al., 2014; Reiche et al., 2014; Thalman et al., 2015). Here, we find agreement between two DCS instruments that is an order of magnitude better and is comparable to that achieved with highly calibrated, state-of-the-art solar-looking FTS systems that retrieve vertical column measurements (Messerschmidt et al., 2011; Frey et al., 2015; Hedelius et al., 2016); however, open-path DCS does not require instrument-specific calibrations (e.g., of the instrument line shape) and provides a very different capability by retrieving the dry mole fractions across regional, kilometer-scale paths over day and night on a mobile platform. Moreover, as the agreement between open-path DCS instruments is below the level of natural background fluctuations, future measurements can facilitate accurate inverse modeling to identify sources and sinks of carbon emission over regions. As an initial demonstration, we discuss the observed diurnal variations from this 2-week measurement campaign in the final section of the paper.

2 Technique

2.1 Dual-comb spectroscopy

A frequency comb is a laser pulsed at a very precise repetition rate of $f_r$ (Cundiff and Ye, 2003; Hall, 2006; Hansen, 2006). Because the pulse rate is so precisely controlled, this creates a spectrum consisting of very narrow, evenly spaced modes called comb teeth. Dual-frequency-comb spectroscopy combines two of these combs with very slightly different pulse repetition rates that differ by $\Delta f_r$ and sends the light through the sample and on to a detector (see Fig. 1a) (Schiller, 2002; Schliesser et al., 2005; Coddington et al., 2008, 2016; Ideguchi, 2017). It is also possible to transmit only a single comb through the sample to measure both dispersion and absorbance (Giorgetta et al., 2015). The basic technique of dual-comb spectroscopy is illustrated in Fig. 1 and described in more detail in the literature (Schiller, 2002; Schliesser et al., 2005; Coddington et al., 2008, 2016).

A DCS system can be thought of as a high-resolution Fourier transform spectrometer but has a number of attributes that distinguish it from conventional horizontal open-path FTS systems and other open-path instruments, which could lead to higher-performance atmospheric trace gas monitoring. A compact, mobile DCS system such as this one has no moving parts, dense point spacing (200 MHz or 0.0067 cm$^{-1}$ in this work), effectively no instrument line shape, and a calibration-free wavelength axis as described in Rieker et al. (2014) and Truong et al. (2016). As a result, it over-samples the 5 GHz wide (0.15 cm$^{-1}$) pressure-broadened gas lines of carbon dioxide, methane, water, and other small molecules without distortion, which should suppress any instrument-specific systematics and allow comparison of DCS data between instruments and over time. Specifically relevant to open-path measurements, the comb output is a diffraction-limited eye-safe laser beam and thus can support much longer distances than typical open-path FTS systems; here we demonstrate 2 km round-trip measurements, but we have unpublished data for up to 11.6 km round-trip. Finally, unlike swept laser systems, DCS measures all wavelengths at once rather than sequentially and is therefore much more immune to turbulence effects as described in Rieker et al. (2014).

---

$^1$We use dry mole fraction for carbon dioxide and methane, denoted respectively as XCO$_2$ in units of ppm, which are micromoles of CO$_2$ per mole of dry air, and XCH$_4$ in units of ppb, which are nanomoles of CH$_4$ per mole of dry air.
2.2 Overview of experiment

Figure 2 provides an overview of our experiment. Two DCS instruments measured the atmospheric absorption across a 2 km round-trip open path that extended from the top of a building at the National Institute of Standards and Technology (NIST) Boulder campus to a pair of retroreflector located on a nearby hill. Both DCS instruments were based on a similar overall design and used self-referenced, stabilized frequency combs (Sinclair et al., 2015), but one was built by a team at NIST and the other by a team at the University of Colorado; they are hereafter referred to as DCS A and DCS B, respectively. As outlined below, the two instruments differed in their exact design and physical parameters. Nevertheless, no instrument-specific calibration or bias offset was applied to either system. The acquired atmospheric absorption spectra were fit to retrieve the column density of CO$_2$, CH$_4$, and H$_2$O (as well as HDO and $^{13}$CO$_2$ at lower precision) along with the path-averaged temperature from the CO$_2$ spectrum. From these data, combined with the measured atmospheric pressure and the path length (measured via time-of-flight laser ranging), we retrieved the path-averaged dry mole fractions as a function of time, which are compared between DCS instruments and to a nearby CRDS point sensor.

2.3 Dual-comb spectrometer

Figure 3a shows a simplified schematic of both DCS set-ups. Briefly, each DCS system used two mutually coherent self-referenced erbium-doped fiber frequency combs based on the design of Sinclair et al. (2015) with nominal repetition rates $f_r$ and the difference in repetition rates $\Delta f_r$ given in Table 1. Mutual optical coherence between the combs is enforced by phase-locking an optical tooth of each to a common continuous-wave (cw) laser and the carrier-envelope offset frequency of each to a common quartz microwave oscillator. Absolute frequency accuracy is then enforced by a bootstrapped approach that effectively locks the common cw laser to the same quartz microwave oscillator (Truong et al., 2016). The result is sub-hertz mutual coherence, $\sim 120$ kHz ($4 \times 10^{-6}$ cm$^{-1}$) absolute line widths, and $1.1$ MHz ($3.6 \times 10^{-5}$ cm$^{-1}$) absolute frequency accuracy (Truong et al., 2016). The direct output of the combs is spectrally broadened in highly nonlinear fiber to cover 7140–5710 cm$^{-1}$ (1.4–1.75 $\mu$m) and then filtered to isolate the spectral region of interest from 6023 to 6376 cm$^{-1}$ (1568 to 1660 nm).

The combined light from both combs is transmitted via single-mode fiber to a telescope, where it is launched to a retroreflector. The returning signal is collected onto an amplified, 100 MHz bandwidth InGaAs photodetector and digitized at a sampling rate $f_s$. We acquire a single interferogram at a period of $1/\Delta f_r$ or 1.6 ms for DCS A; 100 such interferograms are directly summed in real time on a field-programmable gate array (FPGA). These are transferred to a
Figure 2. Setup for the open-path dual-comb spectrometer (DCS) comparison at the NIST Boulder, CO campus. The main components for DCS A and DCS B are housed in a rooftop laboratory, including the frequency combs, telescope, receiver, and processor. For each DCS, the combined comb light is launched from a telescope; travels \( \sim 1 \) km through the atmosphere to a retroreflector; and returns to the telescope, where it is collected, detected, and processed. A separate cavity ring-down point sensor (CRDS) is located nearby with an inlet on a 30 m tower that is located \( \sim 160 \) m from the nearest point of the free-space DCS paths.

Table 1. Specifications of the two DCS systems. HCC: hollow corner cube.

<table>
<thead>
<tr>
<th></th>
<th>DCS A</th>
<th>DCS B</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design Details</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comb 1 repetition rate ( (f_r) )</td>
<td>( \sim 200 \text{MHz} )</td>
<td>( \sim 204 \text{MHz} )</td>
</tr>
<tr>
<td>Difference in repetition rate ( (\Delta f_r) )</td>
<td>624 Hz</td>
<td>870 Hz</td>
</tr>
<tr>
<td>Spectral filtering</td>
<td>Before combining combs</td>
<td>After combining combs</td>
</tr>
<tr>
<td>Booster amplifier</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Average power launched</td>
<td>4 mW</td>
<td>1.5 mW</td>
</tr>
<tr>
<td>Filtered spectral output</td>
<td>6376 to 6023 cm(^{-1})</td>
<td>6359 to 6003 cm(^{-1})</td>
</tr>
<tr>
<td>Telescope design</td>
<td>Home-built 3 in. diameter off-axis telescope</td>
<td>Modified commercial 6 in. diameter Ritchey–Chrétien telescope</td>
</tr>
<tr>
<td>Retroreflector</td>
<td>2.5 in. HCC, 5 arcsec</td>
<td>5 in. HCC, 5 arcsec</td>
</tr>
<tr>
<td>Round-trip path length</td>
<td>1950.17 m</td>
<td>1963.67 m</td>
</tr>
<tr>
<td>Typical averaging time</td>
<td>32 s</td>
<td>28 s</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Performance Metrics</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>30 s precision</td>
<td>0.90 ppm XCO(_2), 9.6 ppb XCH(_4)</td>
<td>2.15 ppm XCO(_2), 11.5 ppb XCH(_4)</td>
</tr>
<tr>
<td>5 min precision</td>
<td>0.24 ppm XCO(_2), 2.1 ppb XCH(_4)</td>
<td>0.60 ppm XCO(_2), 3.2 ppb XCH(_4)</td>
</tr>
</tbody>
</table>

computer where they are carrier-phase-corrected and further summed over an acquisition time of \( \sim 30 \text{ s} \). These summed interferograms are then Fourier transformed and scaled, using the known optical frequency comb tooth positions, to generate a transmission spectrum (e.g., Fig. 4a) spanning 355 cm\(^{-1}\) (\( > 10 \text{ THz} \)) with a point spacing of 0.0067 cm\(^{-1}\). As shown in Fig. 1a, the effective line shape for each sampled point is well approximated as two closely separated delta functions located at the known optical frequencies of the two comb lines that are heterodyned to produce the measured radio frequency (RF) signal (e.g., consider the solid and dashed yellow optical comb teeth that lead to the single solid yellow RF comb tooth.) The separation of the two delta functions (comb teeth) is negligible compared to the \( \sim 5 \text{ GHz} \) wide absorption lines but can be exactly incorporated into the spectral model.

The exact optical layout of DCS A is given in Truong et al. (2016). While following the same basic design, DCS B differs in several technical details. These include a slightly different output spectrum; slightly different comb tooth spacings and offset frequency; minor differences in the reference cw laser and its locking scheme; and different amplifier design, launched and received powers, and telescope design. Some of these differences are laid out in Fig. 3, Table 1, and Sect. 2.4 below.

We have found that the use of stabilized, phase-coherent frequency combs is a necessary but not sufficient prerequisite to reaching sub-percent agreement in retrieved gas concent
trations. It is critical that the spectrally filtered comb output does not include stray unfiltered light. Similarly, any stray reflections from the telescope that can “short-circuit” the atmospheric path must be avoided. As with FTS systems, nonlinearities are problematic. In the optical domain, nonlinearities can arise when the combs are combined in fiber with high optical power. These are minimized for DCS A by filtering the light, which decreases the peak power, before combining the combs. For DCS B the combs do not have booster amplifiers and thus have significantly lower power. Nonlinearities in the photodetection can also occur (Zolot et al., 2013); in laboratory tests with a CO reference cell, we verified no bias in retrieved concentration as a function of received power up to 300 µW, which is a factor of 2 higher than the maximum power for the open-path data. It was also critical to match the interferogram amplitude to the full dynamic range of the analog-to-digital converters (ADCs) to avoid effective nonlinearities in the digitization process.

### 2.4 Launch–receive telescope

The two telescope systems are shown in Fig. 3a. Due to the large spectral bandwidth, reflective optics are preferred to minimize chromatic dispersion. For DCS A, the launch–receive system was based on a bi-directional off-axis parabolic telescope with a 3 in. aperture, while for DCS B it was based on a 6 in. aperture Ritchey–Chrétien (RC) telescope with the light launched separately from behind the secondary mirror. In both cases, the launched beam diameter

---

**Figure 3.** (a) Configuration of DCS A and DCS B, both of which are based on fully self-referenced fiber-laser frequency combs. See text and Table 1 for details. DCS A includes booster amplifiers for higher launched optical power than DCS B. (b) Average optical return power for DCS A (blue) and DCS B (red) measured at the detector over about 2.5 days. The horizontal black line shows the approximate minimum power for useable SNR (15 µW). Right-hand side: the normalized power fluctuations for DCS A over 100 ms. The fluctuations are from turbulence and have a characteristic timescale much longer than the acquisition time for a single DCS spectrum, shown by the thickness of the vertical grey bar, and therefore do not lead to overall distortions in the spectrum. RC: Ritchey–Chrétien; HNLF: highly nonlinear fiber; HCC: hollow corner cube retroreflector; BS: beam splitter.
Figure 4. Raw spectra from DCS A (blue) and DCS B (red). (b) Corresponding baseline-corrected absorption spectra averaged for a 3 h period. The spectra overlap completely on this scale, so the DCS B absorbance has been flipped about zero. Inset: expanded view of several CO$_2$ lines. (c) Difference between the absorption spectra from DCS A and DCS B. The difference is shown both before (black trace) and after (grey trace) removing an etalon structure and agrees to better than $5 \times 10^{-4}$ after the etalon is removed. Inset: expanded view. (d) Residuals from a fit of the DCS A spectrum to HITRAN 2008. In general, the residuals are lower noise than the difference spectrum because of the higher signal-to-noise ratio of the DCS A than DCS B, but there are clear structures present near absorption lines due to imperfect line shapes of the spectral database.
Figure 5. Concentration retrievals from DCS A (blue dots) and DCS B (red lines) for HDO (ppm), H2O (%), dry CH4, dry CO2, and dry 13CO2 over 2 weeks at 30 s intervals. Excellent agreement is observed between both systems for all species, though it is clear that over this path length 13CO2 does not provide a strong enough signal to retrieve reliably. Highlighted section: 6 h, well-mixed period over which Allan deviations (Fig. 8) are calculated. Missing data are primarily due to telescope misalignment and, less frequently, loss of phase lock by one of the combs.

absorbance, and $v$ is the average optical frequency of the two participating comb teeth (e.g., Fig. 1b). We fit the natural logarithm of the transmission spectra, $-\ln[S(v)] = -\ln[I_0(v)] + A(v)$, where the first term is represented by a piecewise polynomial and the second by an absorption spectrum calculated from a spectral database with floated concentrations of 12CO2, 13CO2, 12CH4, 13CH4, H2O, and HDO. For a spectral database we use the high-resolution transmission molecular absorption database (HITRAN) 2008 (Rothman et al., 2009) and Voigt line shapes as this generates a consistent set of line parameters across our conditions and gases. The fit is performed in three steps: first, we fit the polynomial (typically seventh order) over small windows (typically 100 GHz or 3.33 cm$^{-1}$) and include the expected absorbance from relevant gas absorption lines. These polynomials are then stitched together to generate the overall polynomial baseline, which is removed from the measured spectrum to find $A(v)$. We then fit only the 30013 ← 00001 CO2 band in order to retrieve the path-averaged temperature. Finally, $A(v)$ is then refit over the entire spectral window by floating the gas concentrations at the retrieved path-averaged temperature. The retrieved path-averaged concentrations are converted to wet mole fractions by normalizing to the total number density of air molecules, which is calculated from the fitted (or separately measured) air temperature combined with the atmospheric pressure, as measured by a sensor co-located with the CRDS sensor and corrected for the altitude difference. Finally, wet CO2, 13CO2, and CH4 are converted to dry values (XCO2, X13CO2, XCH4) using $XS = S/(1 - c_{H_2O})$, where $XS$ is the dry species mixing ratio, $S$ is the retrieved wet species mixing ratio, and $c_{H_2O}$ is the retrieved H2O mole fraction.

3 Intercomparison results and discussion

3.1 Atmospheric spectrum comparison

Figure 4a shows the overall raw DCS transmission spectra from the two instruments averaged for a 3 h period. They differ significantly because of the different comb intensity profiles, $I_0(v)$. However, after the polynomial baseline fit discussed above is applied, the resulting 3 h averaged absorption spectra are nearly identical as shown in Fig. 4b. The inset of Fig. 4b shows the data sampling points (spaced at ∼200 MHz) across several absorption lines with widths of 5 GHz (0.2 cm$^{-1}$), indicating we have sufficient optical resolution to over-sample the lines. The difference of the absorption spectra, shown as the black line in Fig. 4c, has a standard deviation of 9 × 10$^{-4}$ with no observable structure at absorption lines. This difference is dominated by an etalon on the off-axis telescope used with DCS A. After manually fitting out the etalon structure, the remaining difference between DCS A and DCS B is attributed to measurement noise. DCS A has higher return power (see Fig. 3b), and the measure-
ment noise is primarily from relative-intensity noise (RIN) on the comb light. This RIN is mainly white but has a small peak near 14 MHz, which is mapped to \( \sim 6290 \text{ cm}^{-1} \) in the optical domain, leading to the observed noise increase in that spectral region. DCS B has lower return power, and the measurement noise is from the detector. Nevertheless, the two spectra agree to better than \( 5 \times 10^{-4} \) over the full spectral region (with the exception of the 7 cm\(^{-1}\) section at 6290 cm\(^{-1}\)) and better than \( 2.5 \times 10^{-4} \) over the region near 6100 cm\(^{-1}\), where both DCS systems have significant returned optical power. This very high level of agreement between the two spectra shows that there are no instrumental line shapes or detector nonlinearity effects distorting the observed spectral line shapes; otherwise, structure would be observed in the difference. Thus, the two DCS instruments measure the same comb-tooth-resolved atmospheric absorbance spectrum.

Figure 4d shows the residuals after fitting the absorption lines in the DCS A spectrum to HITRAN 2008 and removing the etalon. The higher signal-to-noise ratio (SNR) of the DCS A yields an even lower broadband noise than the difference spectrum, but there are clear residuals near spectral lines attributable to incorrect line shapes/parameters in the HITRAN 2008 database. Nevertheless, the overall magnitude of the residuals is very small in comparison to the spectral absorption.

3.2 Comparison of retrieved mole fractions from DCS A and DCS B

From the fitted concentrations, we retrieve the mole fractions as outlined in Sect. 2.5. The retrieved time series for \( X_{\text{CO}_2} \), \( X_{\text{CH}_4} \), \( X_{\text{H}_2\text{O}} \), and HDO are given in Fig. 5 at \( \sim 30 \text{ s} \) intervals. Gaps in the data are due to either telescope misalignment (primarily on the 6 in. RC telescope due to the lower-quality gimbal system) or, more rarely, a loss of phase lock of one of the four frequency combs. Excellent agreement is observed between both systems for all retrieved concentrations. Figure 6 shows the concentration differences, which exhibit a high-frequency white noise consistent with the quadrature sum of the DCS precisions given in Sect. 3.3. In addition, the differences show a slow wander about zero indicating slowly changing, small offsets between the two DCS instruments. \( \text{CH}_4 \) also shows a small negative offset for the second week of the campaign. A Gaussian curve approximates the distribution of the differences over the full 2 weeks reasonably well and is shown in Fig. 7. At 32 s averaging times, the mean and width of the distributions are \( \Delta X_{\text{CO}_2} = 0.57 \pm 2.4 \text{ ppm} \), \( \Delta X_{\text{CH}_4} = -7.0 \pm 16 \text{ ppb} \), \( \Delta X_{\text{H}_2\text{O}} = 36 \pm 90 \text{ ppm} \), and \( \Delta X_{\text{HDO}} = 0.10 \pm 0.30 \text{ ppm} \). These widths decrease to 1.5 ppm, 12 ppb, 66 ppm, and 0.17 ppm, respectively, for 5 min averaging times. These mean values correspond to a relative offset of 0.14 % \( \text{CO}_2 \), \(-0.35 % \text{ CH}_4 \), and 0.4 % \( \text{H}_2\text{O} \) and are close to the WMO compatibility standards of 0.1 ppm for \( \text{CO}_2 \) and 2 ppb for \( \text{CH}_4 \) (Tans and Zellweger, 2015). We emphasize the agreement here is achieved over a 2-week period despite outdoor temperature variations of 4.6 to 28.9 °C, DCS instrument ambient temperature variations from 17 to 25 °C, 10 to 90 % relative humidity fluctuations, and large turbulence-induced return power fluctuations.

Table 2 summarizes the systematic uncertainties of the DCS systems in terms of instrument-specific systematics in the top of the table and model-dependent uncertainties common to both instruments in the bottom part of the table. We
Table 2. List of systematic uncertainties. See discussion in Sect. 3.2 and 3.4 for more details. Upper half of table: instrument-specific systematics. Lower half of the table: model-dependent systematics common to both instruments. The final row is the estimated added uncertainty from the water correction, which is dominated by the nominal ∼ 10 % line strength uncertainty of the spectral database. (The fit uncertainty for the retrieved water concentration is much lower at 0.0065 %.)

<table>
<thead>
<tr>
<th>Systematic source [effect]</th>
<th>Effect on retrieved (\text{XCO}_2)</th>
<th>Effect on retrieved (\text{XCH}_4)</th>
<th>Effect on retrieved (\text{H}_2\text{O})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitting procedure [initial guess, baseline polynomial order, and window size]</td>
<td>0.07 ppm</td>
<td>1.4 ppb</td>
<td>4 ppm</td>
</tr>
<tr>
<td>RF detection and processing [RF reflections, ADC nonlinearities]</td>
<td>0.16 ppm</td>
<td>0.34 ppb</td>
<td>1.0 ppm</td>
</tr>
<tr>
<td>Telescope system [scattered light, polarization effects]</td>
<td>0.45 ppm</td>
<td>1.5 ppb</td>
<td>56 ppm</td>
</tr>
<tr>
<td>Spectral database [line strengths in HITRAN 2008]</td>
<td>1−2 %</td>
<td>10−20 %</td>
<td>5−10 %</td>
</tr>
<tr>
<td>Temperature path inhomogeneities [if &lt; 10 °C across path]</td>
<td>0.024 ppm</td>
<td>0.36 ppb</td>
<td>3.52 ppm</td>
</tr>
<tr>
<td>Path-averaged temperature [for 0.5 °C uncertainty]</td>
<td>0.64 ppm</td>
<td>2.9 ppb</td>
<td>8.6 ppm</td>
</tr>
<tr>
<td>Water correction [∼ 10 % line strength uncertainty]</td>
<td>0.4 ppm</td>
<td>2 ppb</td>
<td>n/a</td>
</tr>
</tbody>
</table>

discuss the instrument-specific uncertainties below and the model-dependent uncertainties in Sect. 3.5 in the context of the comparison with the point sensor.

To explore the source of the small systematic offsets between the DCS retrievals, we have performed a number of control comparisons. In the processing, we have varied the initial concentration guess in the fit with negligible effect. We have also varied the polynomial baseline fit by adjusting the window size from 100 to 150 GHz and polynomial order from seventh to ninth order and again found negligible variations of 0.02 % for \(\text{CO}_2\) (< 0.07 ppm), 0.07 % for \(\text{CH}_4\) (< 1.4 ppb), and 0.05 % (∼ 4 ppm) for \(\text{H}_2\text{O}\). In laboratory tests, we verified that the two DCS instruments retrieve the same \(\text{CO}_2\) concentrations to within 0.04 % for 8450 ppm of \(\text{CO}_2\) in a 30 m multipass cell (roughly mimicking the total absorption over the open path). In open-path tests, we have separated effects of the detection/acquisition system and optical system. First, the detected DCS A return signal was split and directed to the two separate data acquisition systems. The two processed signals yielded small differences of 0.16 ppm \(\text{CO}_2\), 0.34 ppb \(\text{CH}_4\), and 1.0 ppm \(\text{H}_2\text{O}\), presumably due to residual nonlinearities and reflections in the RF system and digitization. Second, the outgoing DCS A comb light was split and directed to the two different telescopes and acquisition systems. These two processed signals yielded larger differences of 0.45 ppm \(\text{CO}_2\), 1.5 ppb \(\text{CH}_4\), and 56 ppm \(\text{H}_2\text{O}\), possibly due to scattered light or polarization dependences in the launch and receive optical systems. Finally, residual phase noise between the two combs in a single DCS system can cause small biases in the retrieved concentrations, but these should be well below 0.1 % in this configuration

Figure 8 shows the precision versus averaging time (determined using the modified Allan deviation) based on the scatter across a 6 h period over which the \(\text{CO}_2\) and \(\text{CH}_4\) concentrations are reasonably flat, shown as the highlighted part of Fig. 5. (The Allan deviation for \(\text{H}_2\text{O}\) is not calculated because the atmospheric \(\text{H}_2\text{O}\) concentration varies significantly over this time period.) Under perfectly stable concentrations and white instrument noise, the precision should improve as the square root of averaging time, indicated as a grey line in Fig. 8. Initially, the Allan deviations do follow this slope, but the atmospheric concentrations, especially of \(\text{CO}_2\), vary over this 6 h period, and the Allan deviations reach a floor at ∼ 1000 s. As in Chen et al. (2016) it is also useful to plot the Allan deviation of the difference in retrieved concentration between the instruments, e.g., \(\Delta\text{XCO}_2\) and \(\Delta\text{XCH}_4\) of Fig. 6. This removes the atmospheric variability from the Allan deviation and provides information on the relative stability of the two instruments. As it includes contributions from both DCS instruments, it lies above both individual Allan deviations but similarly reaches a floor at ∼ 1000 s, indicating the floor of the individual Allan deviations is likely dominated by instrument rather than atmospheric variability for these data. We note a similar floor is found in Truong et al. (2016) for static laboratory cell data, where it was attributed to the existence of an etalon.
The precision at 30 s and 5 min averaging time is given at the bottom of Table 1. DCS A has superior CO\(_2\) precision because it has higher received optical comb power in that spectral region, whereas the DCS instruments have similar received power in the CH\(_4\) spectral region and therefore similar CH\(_4\) precisions. Regardless, the precision of either instrument is sufficiently high to measure the characteristic atmospheric fluctuations of these gases on tens-of-seconds timescales.

### 3.4 Comparison of open-path DCS to a CRDS

A commercial cavity ring-down point sensor, Picarro Model 1301\(^2\) (Crosson, 2008), was also located along the path as shown in Fig. 2. Its inlet was 30 m above ground on a radio tower, approximately 160 m perpendicular to the DCS beam path. Figure 9 compares the DCS A and CRDS (smoothed to 32 s resolution) time series. In general, their overall shapes agree well, with both systems tracking ~40 ppm variations in XCO\(_2\), 200 ppb variations in XCH\(_4\), and 1 % variations in H\(_2\)O over days. Nevertheless, there are clear discrepancies in terms of both short-duration spikes and a long-term overall offset between the DCS and CRDS time series.

The short-duration spikes are present in the CRDS time series and presumably arise from the very different spatial sampling of the two instruments. The DCS system measures the integrated column over 1 km (one way), while the CRDS is a point sensor and therefore much more sensitive to local sources. For example, a 1 m\(^3\) volume of air containing 500 ppm of CO\(_2\) from a vehicle driving under the sampling line will result in a sharp spike in the CRDS data as the air mass passes the sampling inlet. However, that same air mass will result in only a 0.025 % or 0.1 ppm increase in the DCS path-averaged concentration (assuming a

\(^2\)The use of trade names is necessary to specify the experimental results and does not imply endorsement by the National Institute of Standards and Technology.

---

**Figure 7.** Statistical distributions of the differences between DCS A and DCS B for dry CO\(_2\), dry CH\(_4\), H\(_2\)O, and HDO from Fig. 6. Histograms are shown in black with a fit to Gaussian curves in red. These data are for ~30 s intervals; the widths are approximately halved if the data is averaged to 5 min intervals.

**Figure 8.** Precision (Allan deviation) versus averaging time, \(\tau\), for CH\(_4\) and CO\(_2\) for DCS A (blue) and DCS B (red) over a 2 km path for the time period highlighted in Fig. 5. The improved precision of the DCS A results from higher signal-to-noise ratio of the measured spectra. The DCS B precision is slightly better than previously published precisions from Rieker et al. (2014). The grey line illustrates the slope expected for white noise. For DCS A, at averaging times from 30 to 1000 s, the precision roughly follows ~40 ppb/\(\sqrt{\tau}\) for CH\(_4\) and ~4 ppm/\(\sqrt{\tau}\) for CO\(_2\) (grey lines), where \(\tau\) is in seconds, before reaching a floor near 1000 s. Also shown are the Allan deviations for the difference between DCS A and DCS B from Fig. 6 (gold). These differential Allan deviations also reach a floor at ~1000 s, suggesting the floor in the individual Allan deviations are due to instrument rather than atmospheric variations.
400 ppm background). These spikes in the CRDS time series are damped here by the 32 s smoothing but are occasionally evident, especially during the second week. The general scarcity of such events does suggest that the air over the open path is usually fairly well mixed.

The long-term overall offset between the CRDS and DCS data is a consequence of their very different calibrations. The CRDS is tied to the WMO scale for CO$_2$ and CH$_4$ by directly injecting known dry WMO-calibrated CO$_2$ / CH$_4$ mixtures at different trace gas concentrations and different water vapor concentrations into its temperature- and pressure-controlled sampling cavity. This instrument was calibrated shortly after the measurement campaign and should thus have an absolute uncertainty close to that of the WMO-scale uncertainties of ∼0.07 ppm for CO$_2$ (Zhao and Tans, 2006) and ∼1.5 ppb for CH$_4$ (Dlugokencky et al., 2005).

In contrast, the DCS has no instrument-specific calibration but relies completely on a fit to a spectral database to extract the gas concentrations from the measured absorbance across a wide range of ambient pressures and temperatures. Here, we use HITRAN 2008, which has $^{12}$CO$_2$ line strength uncertainties of 1–2%, $^{13}$CH$_4$ line strength uncertainties of 10–20%, and H$_2$O line strength uncertainties of 5–10% (Rothman et al., 2009), leading to a poorer absolute calibration than the WMO-calibrated point sensor. From the data in Fig. 9, the differences between the CRDS and DCS data across the 2-week period are $-3.4 \pm 3.4$ ppm CO$_2$, $17 \pm 15$ ppb CH$_4$, and $580 \pm 462$ ppm H$_2$O at 5 min averaging. These correspond to relative offsets of $-0.85$ % for CO$_2$, 0.94 % for CH$_4$, and 6.9 % for H$_2$O, well within the stated uncertainties of HITRAN 2008. In previous DCS measurements, we found slightly different offsets, specifically 1.78 % for CO$_2$, 0.20 % for CH$_4$, and 1.74 % for H$_2$O in Rieker et al. (2014) and ∼1 % for CO$_2$ in Giorgetta et al. (2015). However, these previous data covered much shorter time spans, used an older CRDS point sensor calibration, and may have included small systematic offsets in the DCS systems due to technical issues discussed in Sect. 3.2.

This basic discrepancy between retrievals based on line shape parameters from a spectral database and manometric and gravimetric calibrations (WMO standard) is not unique to DCS. Several studies have calibrated the Total Carbon Column Observing Network (TCCON) retrievals against WMO-based instruments (Wunch et al., 2010; Messerschmidt et al., 2011; Geibel et al., 2012; Tanaka et al., 2012). Although TCCON is not a solely HITRAN-based analysis (Wunch et al., 2011), a correction factor of 0.9898 for CO$_2$, 0.9765 for CH$_4$, and 1.0183 for H$_2$O (Wunch et al., 2010) is needed to bring the overall TCCON retrievals into agreement with the WMO-based data. Additionally, theoretical calculations by Zak et al. (2016) found an approximately 0.5 % difference between CO$_2$ line parameters from HITRAN 2012 and their density functional theory calculations and an additional 0.5 % difference between the calculations and new measurements by Devi et al. (2016) in the 1.6 µm region. Certainly this discrepancy between retrievals from HITRAN-
and WMO-calibrated instruments is not fundamental, and further experimental work should lead to improved spectral database parameters and much better agreement. As noted in earlier work on CO$_2$, it will be important both to establish the correct line strengths and to account for complex line shapes and line mixing (e.g., Devi et al., 2007, 2016; Thompson et al., 2012; Bui et al., 2014; Long et al., 2015). A direct comparison of the open-path DCS spectra acquired here and laboratory DCS spectra acquired for WMO-calibrated gas samples can contribute to these future improvements and is planned. Finally, we emphasize that because the DCS instruments record the atmospheric absorption without instrument distortions, as spectral models improve, past open-path spectra can be refit with reduced uncertainty.

An accurate path-averaged air temperature is also important to avoid systematic offsets. Unlike vertical total-column measurements through the entire atmosphere, kilometer-scale open horizontal paths should have relatively low temperature inhomogeneities of around a few degrees Celsius, and thus the use of a single “path-averaged” temperature in the fit is sufficient for accurate retrievals. We verified this through a sensitivity study comparing retrievals for simulated spectra with temperature gradients up to 10°C over the path; the resulting bias was below 0.03 ppm CO$_2$ (0.007 %) and 0.4 ppb CH$_4$ (0.022 %), as shown in Table 2. On the other hand, any error in the path-averaged temperature can bias the mole fractions through two effects. First, the retrieved path-averaged concentration will vary weakly with temperature because of temperature-dependent line parameters. Second and dominantly, the final mole fraction calculation requires normalization by the air density. Here, this density is calculated from the ideal gas law using the measured air pressure and path-averaged temperature. Therefore, a fractional error in temperature leads to a corresponding fractional error in mole fraction. For example, a 0.15 % uncertainty in mole fraction requires 0.5°C uncertainty in the path-averaged air temperature. (See Table 2.) We verified that this simple linear relationship is valid up to a temperature uncertainty of 10°C in a sensitivity study. From the discussion in Appendix A, the use of a point temperature sensor near the end of the open path is clearly insufficient to achieve <$0.5°C$ uncertainty at many times of the day. Instead, for the data here, we have used the fitted path-averaged temperature, as discussed in Sect. 2.5. The approach effectively relies not only on the spectral database but, in this case, also on the variation in the Boltzmann distribution of the J-level population with temperature. In Table 2, we have taken a hopefully conservative uncertainty of 0.5°C for the path-averaged temperature, but more work is needed to establish the true uncertainty from these retrieved values. We note the fractional uncertainty in the measured atmospheric pressure from the sen-

![Figure 10](https://example.com/figure10.png)

Figure 10. Diurnal cycles for wind speed, wind direction, XCH$_4$, and XCO$_2$. Data from each day in Fig. 5 are overplotted in grey along with the hourly mean (blue) and median (red) values. Uncertainty bars on the median values span the 75th quantile and 25th quantile.
sor or altitude-based pressure changes across the optical path was below 0.36 %.

Finally, the calculation of the dry mole fraction requires an accurate removal of the water concentration. We do retrieve the water concentration with a high precision from the fits. As shown in Table 2, the dominant uncertainty in the water concentration is again the line strengths from the spectral database.

4 Diurnal cycles and source analysis

The 2 weeks of open-path data are analyzed for diurnal cycles, as shown in Fig. 10, with the intent of an initial understanding of CO$_2$ and CH$_4$ sources. For this analysis, the wind speed and wind direction were taken from the NCAR Mesa weather data (ftp://ftp.eol.ucar.edu/pub/archive/weather/mesa/), while the gas concentrations are from DCS A.

4.1 Carbon dioxide

As expected, the median of the diurnal cycle for CO$_2$ shows a peak in the early to mid-morning from commuter traffic, after which the CO$_2$ concentration decreases as the boundary layer rises. It remains approximately steady throughout the afternoon, decreases to a minimum between 19:00 and 20:00 local time, and then increases slightly overnight as the boundary layer collapses. We hypothesize that the afternoon behavior is due to the change in wind direction. Often overnight and through early morning the wind blows from the west to southwest, which brings in cleaner background air from the mountains bordering Boulder. However, in late morning the predominant wind direction shifts to the east and southeast, possibly bringing in higher CO$_2$ concentrations from the Denver metropolitan area – which lies approximately 30 km to the southeast of Boulder – over the course of the afternoon. Typically, the evening wind shifts back to out of the west, once again bringing in the cleaner mountain air and with it a decrease in CO$_2$ concentration.

4.2 Methane

Methane has a significantly weaker diurnal cycle than carbon dioxide, which is consistent with a species that lacks significant diurnally varying local sources. Rather, its concentration follows expected variations in the boundary layer height; the concentration increases overnight into the early morning as the boundary layer collapses, and then it decreases during the late morning through afternoon as the boundary layer rises again. The largest likely methane source near Boulder is local oil/gas fields, but these typically lie to the northeast, while the wind directions are generally out of the west to southeast. It is also possible that the methane comes from leaking natural gas infrastructure within the city.

5 Conclusions

Here we provide the first quantitative comparison of open-path dual-comb spectroscopy instruments. The dual-comb spectrometers were based on fully phase-coherent and stabilized fiber frequency combs and operated nearly continuously over a 2-week period. We performed these measurements over adjacent 2 km round-trip paths to measure concentrations of dry CO$_2$, dry CH$_4$, H$_2$O, and HD$_2$. The measured atmospheric absorbance spectra agree to better than 10$^{-3}$. Correspondingly, we find excellent agreement between the retrieved concentrations from the two instruments without the need for instrument calibration: over 2 weeks of near-continuous measurements, the retrieved CO$_2$ concentrations agree to better than 0.14 % (0.57 ppm), CH$_4$ concentrations agree to better than 0.35 % (7.0 ppb), and H$_2$O concentrations agree to better than 0.4 % (36 ppm). These values are very close to the WMO compatibility standards. The remaining disagreement is likely due to scattered stray light, polarization dependencies, and residual comb phase noise. We further compare the DCS measurements to a cavity ring-down point sensor located along our path. The measured CO$_2$ mole fraction agrees to within 0.85 %, the dry CH$_4$ mole fraction to within 0.94 %, and H$_2$O mole fraction to within 6.9 %. However, this CRDS point sensor is directly calibrated to the WMO scale for CO$_2$ and CH$_4$, while the DCS results are based on HITRAN 2008; we attribute the disagreement in CO$_2$ and CH$_4$ to inaccurate line parameters in the HITRAN database. (Most of the water discrepancy is attributed to the imperfect absolute water calibration of the CRDS point sensor.) Further improvements to the spectral database should reduce these discrepancies. Finally, this open-path DCS can exploit even broader spectrum combs up to 2.3 µm and down to 1.1 µm (Zolot et al., 2012; Okubo et al., 2015), which would enable measurements of similar quality for $^{13}$CO$_2$, NH$_3$, N$_2$O, and O$_2$. These results make open-path DCS a promising new system for greenhouse gas flux measurements from distributed sources.

Data availability. As per NIST regulations, all data are archived at NIST and available upon request.
Appendix A: Temperature studies

As described in Sect. 2.5, we extract the path-averaged temperature directly from a fit to the $30013 \rightarrow 00001$ overtone band of CO$_2$. We perform this fit on 5 min averages, rather than 32 s averages, under the assumption the temperature changes are still slow at that timescale. This path-averaged temperature is then used in a subsequent fit over the full spectral region to extract the column densities and finally the mole fractions. We use a common temperature (from DCS A fit) to analyze both data sets in order to separate out instrument effects from the temperature, but the fitted temperatures between instruments show less than 0.25°C bias.

Figure A1 compares this fitted path-averaged temperature from DCS A to three point sensors, two of which are located on the rooftop near the telescope launch point, and one is located ∼2.2 km away at an altitude ∼200 m above the overall open path. As shown in Fig. A1, the two rooftop temperature sensors located near the telescope agree well with each other but do not agree with the fitted path-averaged temperature. Moreover, that disagreement has a distinct diurnal character, supporting the argument it arises from a real temperature gradient. In contrast, the path-averaged temperature does often agree well with the temperature measured by the third temperature point sensor located at an altitude similar to or higher than the open path on the NCAR Mesa building (ftp://ftp.eol.ucar.edu/pub/archive/weather/mesa/). These data indicate that the point sensor located at the telescope site is not a good proxy for the path-averaged temperature; instead, the fitted path-averaged temperature should be used for the concentration fits because of temperature gradients. Note that the temperature gradients themselves do not lead to appreciable errors in the retrieved mole fractions if the correct path-averaged temperature is used (see Table 2 and Sect. 3.4).

Figure A1. The fitted path-averaged temperature over 2 weeks at 5 min intervals (red) compared to the measured air temperature from a rooftop anemometer located near the telescope (blue), a second thermistor temperature sensor also located on the roof but 100 m away (black), and a third rooftop temperature ∼2.2 km away at the NCAR Mesa facility (gold). Top panel: the difference between the two rooftop temperatures (black) agrees to within 1°C, but the difference between these rooftop sensors and fitted path-averaged temperature (red) shows larger 2–4°C diurnal differences, indicating it is not sufficient to measure the temperature at one “end point” of the open path. In fact, the path-averaged temperature agrees better with the more distant but higher-elevation temperature sensor located at the NCAR Mesa facility.
Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank Tim Newburger and Kathryn McKain for assistance with the CRDS calibrations, Terry Bullet and Jon Koffer for assistance in setting up the CRDS sampling, Ryan Thalman for detailed long-path instrument correlation data, and Adam J. Fleischer and Anna Karion for assistance with the manuscript. This work was funded by the Defense Advanced Research Program Agency DSO SCOUT program, ARPA-E MONITOR program under award number DE-AR0000539, and James Whetstone and the NIST Greenhouse Gas and Climate Science Initiative. Eleanor M. Waxman and Kevin C. Cossel are supported by National Research Council postdoctoral fellowships.

References


