Atmospheric
Measurement
Techniques

Assessment of errors and biases in retrievals of X_{CO_2}, X_{CH_4}, X_{CO}, and X_{N_2O} from a $0.5\,cm^{-1}$ resolution solar-viewing spectrometer

Jacob K. Hedelius[1], Camille Viatte[2], Debra Wunch[2,a], Coleen M. Roehl[2], Geoffrey C. Toon[3,2], Jia Chen[4,b], Taylor Jones[4], Steven C. Wofsy[4], Jonathan E. Franklin[5,c], Harrison Parker[6], Manvendra K. Dubey[6], and Paul O. Wennberg[2]

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA, USA
[2]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[3]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
[4]School of Engineering and Applied Sciences and Department of Earth and Planetary Sciences,
Harvard University, Cambridge, MA, USA
[5]Department of Physics & Atmospheric Science, Dalhousie University, Halifax, Nova Scotia, Canada
[6]Earth and Environmental Sciences, Los Alamos National Laboratory, Los Alamos, NM, USA
[a]now at: Department of Physics, University of Toronto, Toronto, Ontario, Canada
[b]now at: Electrical and Computer Engineering, Technische Universität München, Munich, Germany
[c]now at: School of Engineering and Applied Sciences and Department of Earth and Planetary Sciences,
Harvard University, Cambridge, MA, USA

Correspondence to: Jacob K. Hedelius (jhedeliu@caltech.edu)

Abstract. Bruker™ EM27/SUN instruments are commercial mobile solar-viewing near-IR spectrometers. They show promise for expanding the global density of atmospheric column measurements of greenhouse gases and are being marketed for such applications. They have been shown to measure the same variations of atmospheric gases within a day as the high-resolution spectrometers of the Total Carbon Column Observing Network (TCCON). However, there is little known about the long-term precision and uncertainty budgets of EM27/SUN measurements. In this study, which includes a comparison of 186 measurement days spanning 11 months, we note that atmospheric variations of X_{gas} within a single day are well captured by these low-resolution instruments, but over several months, the measurements drift noticeably. We present comparisons between EM27/SUN instruments and the TCCON using GGG as the retrieval algorithm. In addition, we perform several tests to evaluate the robustness of the performance and determine the largest sources of errors from these spectrometers. We include comparisons of X_{CO_2}, X_{CH_4}, X_{CO}, and X_{N_2O}. Specifically we note EM27/SUN biases for January 2015 of 0.03, 0.75, −0.12, and 2.43 % for X_{CO_2}, X_{CH_4}, X_{CO}, and X_{N_2O} respectively, with 1σ running precisions of 0.08 and 0.06 % for X_{CO_2} and X_{CH_4} from measurements in Pasadena. We also identify significant error caused by nonlinear sensitivity when using an extended spectral range detector used to measure CO and N_2O.

1 Introduction

Measurements of atmospheric mixing ratios of greenhouse gases (GHGs), including CO_2 and CH_4, are needed to aid in estimating fluxes and flux changes, and to ensure international treaties to reduce emissions are fulfilled. The Total Carbon Column Observing Network (TCCON) makes daytime column measurements of these gases. The Orbiting Carbon Observatory2 (OCO-2) and Greenhouse Gases Observing Satellite (GOSAT) missions enable column GHG measurements with global coverage. These GHG monitoring satellites make measurements at one time of day and, therefore, lack the temporal resolution that a dedicated ground site provides.

Due to cost, lack of infrastructure, and stringent network requirements, there are limited ground sites on a global scale; e.g., there are no TCCON sites currently in operation in continental Africa, South America, or central Asia (Wunch et al., 2015), and there currently is no urban area with more than one TCCON site. Cheaper, portable, solar-viewing Fourier transform spectrometers (FTSs) can make contributions in these settings provided they have long-term precision. The Bruker Optics™ EM27/SUN, with the "SUN" indicating a built-in solar tracker, is a transportable FTS that may supplement global GHG measurements made by current networks (Gisi et al., 2012). This unit is small and stable enough to easily be transported for field campaign measurements, including measurements at multiple locations in 1 day. Column-averaged dry-air mole fractions (DMFs) of gases (X_{gas}) are retrieved from the EM27/SUN measurement, like the TCCON. X_{gas} is calculated from (Wunch et al., 2010):

$$X_{gas} = \frac{column_{gas}}{column_{dry\ air}} = 0.2095 \frac{column_{gas}}{column_{O_2}}, \qquad (1)$$

where the 0.2095 factor is the fraction of dry air that is oxygen.

Retrieved X_{gas} has been compared with a co-located TCCON site in Karlsruhe, Germany, in past work for 26 days of X_{CO_2} retrievals from one EM27/SUN instrument (Gisi et al., 2012), and 6 days of both X_{CO_2} and X_{CH_4} retrievals from five EM27/SUN instruments (Frey et al., 2015).

Operators of these instruments have different end goals to better understand the carbon cycle. X_{CO_2} and X_{CH_4} retrievals from these instruments have been compared with satellite measurements in areas without a TCCON site (Klappenbach et al., 2015) as well as with satellite measurements in highly polluted areas (Shiomi et al., 2015). Emission flux estimates from the Berlin area ($< 30 \times 30$ km^2) were made by combining upwind/downwind measurements from five spectrometers and were compared with a simulation (Hase et al., 2015). Chen et al. (2016) have assessed gradient strengths around a large dairy farm ($\sim 100\,000$ cows) in Chino, California ($< 12 \times 12$ km^2), using measurements from upwind/downwind spectrometers. Weather Research and Forecast Large-Eddy Simulations (WRF-LES, 4 km resolution) were used in combination with four simultaneous measurements to estimate fluxes from specific grid boxes in a subregion of the Chino dairy farm area, which is within a larger urban area (Viatte et al., 2016).

The column measurements used in these studies provide some advantages over in situ measurements, including less sensitivity to vertical exchange, surface dynamics, and small-scale emissions (McKain et al., 2012), which are difficult to model. Though column measurements can depend on mixed layer height in highly polluted areas, generally, column measurements depend primarily on regional-scale meteorology, and regional fluxes (Wunch et al., 2011b; McKain et al., 2012). For example, Lindenmaier et al. (2014) used observations from a single TCCON site to verify 1 day of emissions from coal power plants of about 2000 MW each at ~ 4 and 12 km away. Because of their large spatial sensitivity, column measurements are well suited for estimation of net emissions, model comparison, and satellite validation. A single site has been used to estimate Los Angeles, California (L.A.), emissions based on a sufficiently accurate emissions inventory and the observation that X_{gas} anomalies within L.A. are highly correlated (Wunch et al., 2009, 2016). Generally though, a single column measurement site is insufficient to estimate emissions from an entire urban region (Kort et al., 2013). However, multiple column measurements can be combined to characterize part or all of an urban area (Hase et al., 2015; Chen et al., 2016; Viatte et al., 2016).

The main goal of this work is to quantitatively evaluate the robustness of EM27/SUN retrievals over a long period of time. This is accomplished by comparing retrievals from the EM27/SUN with a co-located standard (TCCON site) at Caltech, in Pasadena, California, United States. TCCON spectrometers make the same type of measurements (direct solar near-infrared) at high spectral resolution. Here we report X_{CO_2}, X_{CH_4}, X_{CO}, and X_{N_2O} comparison measurements from an EM27/SUN. The X_{CO} and X_{N_2O} measurements were made possible by a detector with an extended spectral range provided by Bruker™. The EM27/SUN X_{CO_2} and X_{CH_4} to TCCON comparison is the longest to date, 186 measurement days spanning 11 months. In part of January 2015, an additional three EM27/SUN instruments were at Caltech for 9 to 12 days of X_{CO_2} and X_{CH_4} comparisons to assess their relative biases. In Sect. 2 we briefly describe differences in instruments and the data acquisition process. In Sect. 3 we describe the retrieval software. In Sect. 4 we describe the inherent properties of EM27/SUNs such as instrument line shapes (ILSs), frequency shifts, ghosts, detector linearity, and external mirror degradation. Section 5 focuses on biases and sounding precision of different gases compared with the TCCON. Section 6 describes sources of instrumental error. We conclude with general recommendations of tests to perform on any new type of direct solar near-infrared (IR) instrument used to retrieve abundances of atmospheric constituents.

2 Instrumentation

2.1 TCCON IFS 125HR

All TCCON sites employ the high-resolution Bruker Optics™ IR FT spectrometer (IFS) 125HR that has been described in detail elsewhere (Washenfelder et al., 2006; Wunch et al., 2011b). For the Caltech TCCON site (34.1362° N, 118.1269° W, 237 m a.s.l.), the IFS 125HR uses an extended InGaAs (indium gallium arsenide) detector, covering 3800–11 000 cm^{-1} for detection and retrieval of all gases relevant to this study (O$_2$, CO$_2$, CH$_4$, CO, and N$_2$O). Figure 1 has example spectra from IFS 125HR and
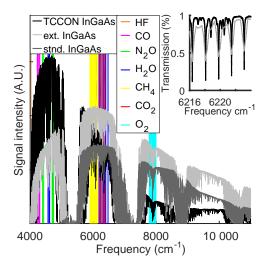
Figure 1. Example of scaled spectra from three different detector types, with retrieval windows highlighted. The spectrum from the EM27/SUN extended InGaAs detector was scaled 10 times more than the spectrum from the standard InGaAs detector.

EM27/SUN instruments, with the spectral regions where individual gases are retrieved highlighted. Oxygen (O_2) abundance is useful in calculating the DMF because it represents the column of dry air and is combined with the column of the gas of interest to yield the DMF (Wunch et al., 2010).

The Caltech IFS 125HR uses a resolution of approximately 0.02 cm^{-1} (with a maximum optical path difference (MOPD) of 45 cm). It takes about 170 s to complete one forward/backward scan pair. TCCON sites have single sounding 2σ uncertainties of 0.8 ppm (X_{CO_2}), 7 ppb (X_{CH_4}), 4 ppb (X_{CO}), and 3 ppb (X_{N_2O}) (Wunch et al., 2010). TCCON data are tied to the World Meteorological Organization (WMO) in situ trace gas measurement scale through extensive comparisons with in situ profiles obtained from aircraft and balloon flights. We use the TCCON as a standard against which to compare the EM27/SUN instruments. TCCON data from this study are publicly available from the Carbon Dioxide Information Analysis Center (Wennberg et al., 2014).

2.2 Caltech EM27/SUN

EM27/SUN spectrometers have been described elsewhere (Gisi et al., 2012; Frey et al., 2015; Klappenbach et al., 2015) so we focus on differences in setup and acquisition here. The standard EM27/SUN configuration uses an InGaAs detector sensitive to the spectral range spanning 5500–12 000 cm^{-1}, which permits detection of O_2, CO_2, CH_4, and H_2O (Frey et al., 2015). For this study, the Caltech EM27/SUN was delivered with an extended-band InGaAs detector sensitive to 4000–12 000 cm^{-1}, which allowed for additional measurements of CO and N_2O (Fig. 1). All EM27/SUN spectrometers used in this study (Sects. 2.2, 2.3) used the typical MOPD of 1.8 cm, corresponding to a spectral reso-

lution of 0.5 cm^{-1}. Interferograms (ifgs) were acquired in direct-current-coupled mode to allow post-acquisition low-pass filtering of brightness fluctuations to reduce the impact of variable aerosol and cloud cover effects (Keppel-Aleks et al., 2007). Ghosts were reduced as data were acquired by employing the interpolated sampling option provided by Bruker™ (see also Sect. 4.3). A 10 KHz laser fringe rate is used to reduce scanner velocity deviations, and each forward/backward scan took 11.6 s, or 5.8 s per individual measurement.

To be more consistent with the TCCON measurements, no spectrum averaging or interferogram apodization was applied before retrieving DMFs. We recommend averaging only after retrievals if disc storage and processor speeds are sufficient, so spurious data can be filtered. To test the pre- vs. post-averaging effect we used 9 retrieval days with 26 000 forward/backward measurements and used Bruker™ OPUS software to create spectra from ifgs. We compared retrievals from using five combined backward/forward measurements averaged pre-retrieval with those averaged post-retrieval. We also compared combined forward/backward measurements using a medium Norton–Beer apodization with those using no special apodization. Results are in Table 1 and suggest that different averaging methods cause only small inconsistencies, under ~ 0.02 % for X_{CO_2} and X_{CH_4}.

The EM27/SUN was placed within 5 m of the Caltech TCCON solar tracker mirrors on the roof of the Linde+Robinson building (Hale, 1935). Measurements started on 2 June 2014 and, for this study, we include 186 measurement days that end on 4 May 2015. About 800 000 individual EM27/SUN measurements and 40 000 individual TCCON measurements were acquired over this period. Of these, about 580 000 and 15 000 were considered coincident and were not screened out by our quality control filters (QCFs). Our QCFs were conservative, and they required signal > 30 (Sect. 4.4), solar zenith angle (SZA) < 82°, 370 ppm < X_{CO_2} < 430 ppm, $X_{CO_2,error}$ < 5 ppm, $X_{CO,error}$ < 20 ppb, and $X_{CH_4,error}$ < 0.1 ppm. Other users may consider stricter QCFs. After averaging data into 10 min bins, there were about 6500 binned comparison points.

2.3 LANL and Harvard EM27/SUN instruments

Three additional EM27/SUN instruments were compared with the Caltech TCCON site in January 2015 – one owned by Los Alamos National Laboratory (LANL) and two owned by Harvard University (HU). To be consistent, all the acquisition and retrieval settings were the same as for the Caltech EM27/SUN. As opposed to the Caltech EM27/SUN (also abbr. cn), the LANL (abbr. pl) and HU instruments (abbr. ha and hb) used the original InGaAs detector type sensitive over 5500–12 000 cm^{-1} (Frey et al., 2015). The LANL instrument, however, has a different high-pass filter, allowing it to measure up to 14 500 cm^{-1}. This different filter is neither beneficial nor disadvantageous to this instrument as no

Table 1. Pre-averaging and apodization effects on EM27/SUN retrievals.

% error	X_{CO_2} Md.	σ	X_{CH_4} Md.	σ	X_{H_2O} Md.	σ	X_{CO} Md.	σ	X_{N_2O} Md.	σ
5 fwd/bwd pre-avgd.[a,b]	<0.01	0.01	−0.02	0.01	0.36	0.13	<0.01	0.15	0.30	0.12
NB med. apodz.[a,c]	0.29	0.09	−0.07	0.10	0.35	0.23	−1.01	0.58	−1.36	0.55

Measurement compared over 1–10 July 2014. Md denotes the median. NB denotes the medium Norton–Beer apodization. [a] As compared to retrievals from 1 fwd/bwd averaged non-apodized measurement averaged over same time post-retrieval. [b] Same apodization as standard. [c] Same pre-averaging as standard.

gas column amounts are retrieved in that region. The LANL instrument was first used in January 2014 and has been compared with multiple TCCON sites in the United States, including sites at Four Corners, LANL, NASA Armstrong, Lamont, Park Falls, and multiple Caltech comparisons (Parker et al., 2015). The HU instruments have been operational since May 2014 and were compared against each other at Harvard before traveling over 4100 km to Caltech. As noted by Gisi et al. (2012) and Chen et al. (2016), the ILS of these instruments is remarkably stable considering the long distances they traveled.

3 Retrieval software

SFIT (Pougatchev et al., 1995), PROFFIT ("PROFile fit", Hase et al., 2004), and GGG (Wunch et al., 2015) are the three widely used retrieval algorithms to fit direct solar spectra and obtain column abundances of atmospheric gases. PROFFIT is maintained by the Karlsruhe Institute of Technology (KIT) and has been used to obtain DMFs from EM27/SUN instruments as well as NDACC-IRWG sites (Gisi et al., 2012; Frey et al., 2015; Hase et al., 2015). GGG is maintained by the Jet Propulsion Laboratory (JPL) and has been used to obtain DMFs from other low-resolution instrument measurements (e.g., an IFS 66, see Petri et al., 2012), in addition to being used to retrieve DMFs from the MkIV spectrometer in balloon-borne measurements (Toon, 1991) and for the Atmospheric Trace Molecule Spectroscopy Experiment (ATMOS) flown on the space shuttle (Irion et al., 2002). GGG is the retrieval algorithm used by the TCCON (Wunch et al., 2011b). We chose to use GGG for our analysis because (1) we want to be consistent with the TCCON for comparison and (2) the GGG software suite containing GFIT is open-source allowing us to adapt routines if needed. We used the GGG2014 version for retrievals (Wunch et al., 2015).

All retrievals used the same pTz and H_2O modeled profiles as well as the same a priori profiles (Wunch et al., 2015). We also used the same meteorological surface data for retrievals from all five instruments. All retrievals also used the same 0.2 hPa surface pressure offset. This offset was determined by comparing measurements from the standard barometer with a calibrated Paroscientific Inc. 765–16B

Barometric Pressure Standard that has a stated accuracy of better than 0.1 hPa.

3.1 Interferogram-to-spectrum – double-sided

TCCON uses an interferogram-to-spectrum subroutine part of GGG to perform fast Fourier transforms (FFTs) to create spectra from ifgs (Wunch et al., 2015). Though the Bruker™ OPUS software used to operate the spectrometer can also perform FFTs, we again chose to use GGG to maintain consistency. A developmental version of GGG was used, which was adapted to also allow FFT processing on EM27/SUN interferograms. GGG splits a raw forward/backward ifg into two different double-sided ifgs which are then FFTed to yield two spectra. GGG also corrects source brightness fluctuations (Keppel-Aleks et al., 2007).

3.2 EM27/SUN GGG and interferogram processing suite (EGI)

To make GGG retrievals simpler for new EM27/SUN users, an add-in software suite (EGI) was developed at Caltech to create correctly formatted input files. This suite is open-source and can be obtained through correspondence to the email address listed. EGI can be run using MATLAB or Python. EGI runs in UNIX, Mac OS, and Linux environments and runs GGG on multiple processors. EGI centralizes settings for paths to read and write files, it coordinates separately acquired ground weather station and GPS data with EM27/SUN ifgs, and it optimizes processing order. It also provides some ancillary calculations such as a spectral signal-to-noise ratio (SNR) calculation. EGI provides a simple way to turn on and off saving of ancillary retrieval files (i.e., spectral fits and averaging kernels). EGI can run for instruments employing one or two detectors, such as the type described by Hase et al. (2016). Like the GGG software suite, EGI also includes benchmark spectra acquired under different conditions to run simple tests on. EGI is automated, reducing the learning time as well as the amount of user time needed to retrieve DMFs. After an initial setup, EGI will run from ifgs to retrieved X_{gas} with two commands. On a computer with 1400 MHz processors the code takes ∼ 30 s per CPU to process each interferogram from the EM27/SUN extended InGaAs detector.

Table 2. ILS of EM27/SUN instruments.

Instrument num – ID	9 January 2015 ME, PE (mrad)[a]	28 January 2015 ME, PE (mrad)
Caltech (42 – cn)	0.986, *4.88*	0.979[b], *3.58*
LANL (34 – pl)	0.999, *−1.34*	
Harvard 1 (45 – ha)	0.973[c], *−1.99*	
Harvard 2 (46 – hb)	0.991[c], *4.18*	0.991, *4.00*

Missing values indicate ILS not characterized on that day. [a] Phase error values are italicized. [b] After realigning this instrument the ME was as high as 0.994. [c] As reported by Chen et al. (2016).

4 Instrument characterizations and performance

4.1 Instrument line shape

Knowledge of the instrument line shape (ILS), or the observed shape of a spectral line from a monochromatic input, is crucial in assessing instrument performance and avoiding unknown biases in retrievals. Two parameters in the the LIN-EFIT algorithm (Hase et al., 1999) are used to characterize the ILS in relation to an ideal instrument, namely the modulation efficiency (ME) and phase error (PE). ME and PE both describe the interferogram and vary with OPD (Hase et al., 1999; Frey et al., 2015). PE is the angle between the real and imaginary parts of the FT of the ILS (Wunch et al., 2007). PE has an ideal value of 0 radians, and indicates the degree of asymmetry in spectral lines. ME is a measure of the normalized observed interferogram signal compared with that of a nominal instrument with an ideal value of 1 (unitless) (Hase, 2012). At maximum OPD (MOPD), an ME < 1 causes a broadening of the measured spectral lines, while an ME > 1 at MOPD causes a narrowing. The ILS can be calculated by analyzing absorption lines measured through a low-pressure gas cell, and varies with OPD (Hase et al., 1999). Here, we use only single ME and PE values at the MOPD (Frey et al., 2015) to describe the ILS. We characterized the ILS for the EM27/SUN instruments using the method described elsewhere (Frey et al., 2015; Klappenbach et al., 2015). This method is able to characterize ME to within 0.15 % using the LINEFIT algorithm (Hase et al., 1999), with supplemental MATLAB scripts for automation purposes (Chen et al., 2016). ILS can affect retrieved column values. We note that the ME at MOPD of the cn and ha instruments in Table 2 are significantly lower than those reported by KIT on campus of ∼ 0.997 (Frey et al., 2015), and post-campaign of ∼ 0.996 (Klappenbach et al., 2015).

For this study, the ILS is used to help explain biases, to demonstrate the stability of the instruments, and gives insight into how well the EM27/SUN instruments are aligned and their optical aberrations. Though GGG2014 retrievals do not account for non-ideal ILS, future versions of GGG will. For the current study, we assume that ILS impacts using PROF-FIT will be similar to impacts using GGG. This assumption
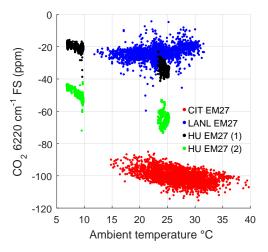

Figure 2. Frequency shifts (FS) of all four instruments vary with temperature because the lasers are not frequency-stabilized. FS for the CO$_2$ 6220 cm^{-1} window are shown. FS of the Caltech (CIT) instrument are far from zero, so an empirical correction is made to correct the sample spacing number. Only every 300th CIT point and every 20th LANL point is plotted for clarity. HU EM27 1 and 2 are also referred to as ha and hb respectively by Chen et al. (2016).

will need to be tested when GGG also can account for a non-ideal ILS. Because future GGG retrievals will be revised using historical ILS measurements, a need remains to monitor the ILS both for future retrievals and as an indicator if realignment is necessary.

4.2 Frequency shifts

EM27/SUN units contain a HeNe 633 nm (15 798 cm^{-1}) metrology laser to sample the IR signal accurately as a function of the OPD. The laser is not frequency-stabilized (Gisi et al., 2012). This causes apparent spectral frequency to change with temperature as is shown in Fig. 2. Frequency shifts are affected by changes in the input laser wavenumber, laser alignment, and IR beam alignment. The input laser wavenumber will affect the spacing between spectral points. Since the frequency shift is furthest from zero for the Caltech EM27/SUN (on order of −100 ppm, in red Fig. 2), the spectral spacing is empirically corrected in the EGI suite based on the CO$_2$ 6220 cm^{-1} window frequency shifts. This made little difference for the primary gases of interest affecting X$_{CO_2}$ by 0.015 % and X$_{CH_4}$ by −0.005 %, though it did affect X$_{H_2O}$ by 4 %.

4.3 Ghosts

Ghosts are artificial spectral features linked to the aliasing of true spectral lines that arise in FTS spectra (Learner et al., 1996). The InGaAs detectors are optically sensitive at wavenumbers greater than half the HeNe metrology laser frequency (7899 cm^{-1}). To fulfill the Nyquist criterion and

prevent aliasing, the IR interferogram is sampled twice each laser interferogram cycle, on the rising and falling edge. However, if the laser sampling is asymmetric – for example from a faulty electronics board – aliasing can still occur, folded across the half laser frequency (Messerschmidt et al., 2010). Because the asymmetry is typically small, the aliased signal, or ghost spectrum, is small compared with the true spectrum (Dohe et al., 2013; Wunch et al., 2015).

In EM27/SUN instruments the laser sampling error (LSE) can be minimized as data are collected by employing the interpolated sampling option provided by Bruker™. This resampling mode uses only the rising edge of the laser interferogram and assumes constant velocity in between the rising edges to interpolate the sampling (Gisi, 2014). We use a narrow band-pass filter ($3\,\mathrm{dB}$ band width 5820–$6150\,\mathrm{cm}^{-1}$) in the Caltech EM27/SUN to test for LSE ghosts at $9800\,\mathrm{cm}^{-1}$. The ghost to parent ratio is 1.73×10^{-4} at a $10\,\mathrm{kHz}$ acquisition rate without the interpolated sampling activated. This ghost is eliminated with the interpolated sampling turned on. In actual solar tests, turning the interpolated sampling on and off had no noticeable effect on the DMF retrievals for the Caltech EM27/SUN; however this may not hold true for all instruments. The LSE ghost also disappeared at an acquisition frequency of $20\,\mathrm{kHz}$, and returned at higher acquisition frequencies. We opted for the recommended $10\,\mathrm{kHz}$ acquisition rate with the interpolated sampling on for all EM27/SUNs in this analysis because other instruments may be more significantly affected by LSE ghosts. A double-frequency ghost remains at $\sim 11\,900\,\mathrm{cm}^{-1}$ from radiation passing through the interferometer twice that is much larger than the LSE ghost, but is not in a region that will affect retrievals.

4.4 Mirror degradation and detector linearity

Solar tracking mirrors provided with the EM27/SUN instruments are gold with a protective coating. Gold is used because of its excellent reflectance in the near-IR and low reflectance in the visible region (Bennett and Ashley, 1965), which allows a high signal while reducing excess heating of the field stop and other optics. Through extended tests, we noted the first two mirrors (gold on plated aluminum, with a coating) degrade over time, with an e-folding degradation time of ~ 90 days as is shown in Fig. 3. Arbitrary units (AUs) for signal are the maximum ordinate values of the unmodified interferograms multiplied by 6450. The AUs of signal happen to be close to the spectral SNR – a scaling factor of 1.3 applied to the arbitrary signal has an R^2 of 0.63 relative to the SNR. Cleaning helped restore some signal, but never to the original values. The mirror change may not have restored full signal because the rest of the optics were not cleaned at the time of the mirror change. Below the blue 150 AU line in Fig. 3 the fitted O_2 root mean square (rms) as a percentage of the continuum level dropped 26 times faster with signal intensity than above it. The instrument did come with an extra
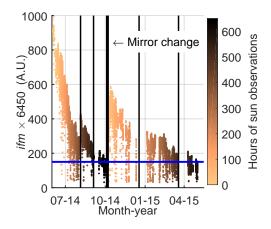

Figure 3. Interferograms from EM27/SUN instruments are negative, with the most negative ordinate values at ZPD and saturation occurring at -1. Here the interferogram maximums (ifm) refer to the maximum (least negative) ordinate values of the raw interferograms. They were normalized so the maximum is 1000 and are plotted with time showing the loss of signal. These values are affected by clouds, which are the cause for much of the scatter. They are also affected by SZA which explains some apparent intermediate increases. Only every 50th point is plotted for clarity. Mirror cleaning (thin black lines) helped restore some signal, but never to original values. The 150 AU line is in blue.

set of mirrors, but because mirrors are consumable parts, it adds recurring cost and effort to maintain these instruments long-term. After 1 year of use, the third mirror (gold coated glass) still remains completely intact. Feist et al. (2016) had success using steel mirrors under the very harsh conditions at the Ascension Island TCCON site, though at a cost of 35 % reflectivity per mirror. The JPL TCCON sites near Caltech noted no degradation on the external gold mirrors over more than 1 year of measurements. The lack of degradation on the third external mirror and the JPL TCCON mirrors is likely due to differences in how the mirrors were manufactured, including how the gold is applied to the substrate and the coatings used. Mirror degradation has likely not been a widely reported problem for most of the EM27/SUN community, perhaps because these instruments typically are stored indoors and only used for a few days for campaigns (for example, Frey et al., 2015). However, this problem may affect mirrors on other EM27/SUN instruments when mirrors are exposed outside for extended periods of time.

With signal loss, we would anticipate that gas measurements would become noisier but remain unbiased. However, with time, the Caltech EM27/SUN X_{CO_2} and X_{CH_4} DMFs decreased relative to the TCCON DMFs as mirror reflectance decreased, and X_{CO_2} and X_{CH_4} increased when the mirrors were replaced. The TCCON IFS 125HR InGaAs detectors are already known to be sufficiently linear that no correction is required (Wunch et al., 2011b). We also performed a simple test repeatedly adding mesh screens in front of the en-
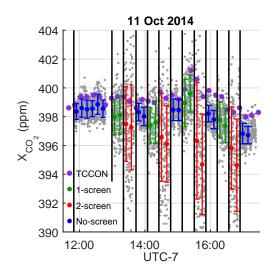
Figure 4. X_{CO_2} retrievals on 11 October 2014 when mesh screens were repeatedly moved in front of and away from the EM27/SUN (with extended InGaAs detector) entrance window. Gray points are all EM27/SUN measurements. Large points are 10 min averages. Error bars are 1σ. This test was performed a few days after the mirrors were replaced.

trance window to filter some of the light. In these tests X_{CO_2} and X_{CH_4} changed on order of 3 and 0.01 ppm respectively when using the extended InGaAs detector in the presence of filters transmitting $\sim 25\,\%$ of the light. Figure 4 shows results from this test on X_{CO_2}; results from X_{CH_4} are similar. This provides strong evidence that the extended InGaAs detector is nonlinear. We repeated the test using the standard InGaAs detector, and changes in X_{CH_4} and X_{CO_2} biases were of the order of 10 times smaller and could be attributed to scattering off the mesh screen placed in front of the entrance window. Figure 5 shows the difference between the EM27/SUN and TCCON X_{CO_2} and X_{CH_4} as the total signal changed. After the mirrors were changed, the relative difference actually went up with some signal loss before decreasing again, for reasons we do not understand.

Detector nonlinearity in FTS instruments can be corrected in the ifgs post-acquisition in two ways. The first option deals with artifacts around the ZPD (zero path difference point) and is already included in GGG (Keppel-Aleks et al., 2007). When the ifg is smoothed, a nonlinear detector exhibits a dip around the ZPD which can be used to diagnose and reduce detector nonlinearity effects. EM27/SUN measurements are too noisy to properly characterize or detect this dip and so this correction is insufficient. The other option is to compare detector response with radiance from a controlled external light source, such as a blackbody, with very accurate radiation flux measurements (on order of $0.01\,\%$) (Thompson and Chen, 1994). By characterizing the response to the true flux as it is varied, the detector can be characterized and ifgs can be appropriately scaled and corrected. However, this requires

extremely controlled precise measurements, as all nonlinearity is likely less than $1\,\%$, so measurements must be more precise than $1\,\%$.

An option to prevent nonlinearity from interfering with measurements is to only use the detector over its linear range by sufficiently attenuating the incoming sunlight. However, the SNR is already low so we opted against this method. Ultimately, we purchased the non-extended InGaAs detector at the loss of CO, and N_2O for future measurements for the Caltech instrument. For the historical field measurements we use a bias correction to match the TCCON for the nearest comparison days. The nonlinearity has nearly an equal effect for short times, but has a larger variation on multi-monthly scales as the mirrors degrade. In future measurements we recommend against using these extended InGaAs detectors. Addition of band-pass filters or use of different detectors will be necessary to provide high-quality measurements of CO, CO_2, and CH_4 (Hase et al., 2016).

The data shown in Fig. 5 were divided into bins based on the signal intensity and were separated before and after the mirror change. Within each bin the relationship was treated as approximately linear. Fits using fewer than 10 points or with correlation coefficients less than 0.1 were discarded. The change with half signal was calculated. The analysis was repeated for 10 bins and again for 20 bins. The weighted mean change in X_{CO_2} for halving the signal is $-1.43\,\text{ppm}$ in agreement with the mesh tests or

$$\Delta X_{CO_2}\left(\text{ppm}^{-1}\right) = 2.06\ln\left(S/S_0\right), \qquad (2)$$

where S and S_0 are the final and initial signals respectively. This relationship holds for S and S_0 in the middle $80\,\%$. For a similar methane analysis the mean change for half signal is $-7.25\,\text{ppb}$ or

$$\Delta X_{CH_4}\left(\text{ppb}^{-1}\right) = 10.5\ln\left(S/S_0\right). \qquad (3)$$

5 Comparisons with X_{gas}

GGG2014 includes an air-mass-dependent correction factor derived for TCCON X_{gas} measurements. The air mass correction factor for each gas is calculated using data obtained at a variety of relatively clean sites as described by Wunch et al. (2011b). We expect that the air mass dependence, which is due primarily to spectroscopic uncertainties, should be common for the same type of measurement. Parker et al. (2015) noted that the average EM27/SUN factors are similar compared to the TCCON for X_{CO_2} at three clean sites in the United States. The X_{CH_4} β factor was different (-0.0077 EM27/SUN, 0.0053 TCCON) but when applied here it worsened the R^2 and standard deviation of the comparisons. This could be because the air mass dependence of X_{CH_4} may not be solely from spectroscopic issues. Hence, we used the same air-mass-dependent correction factors as the TCCON.
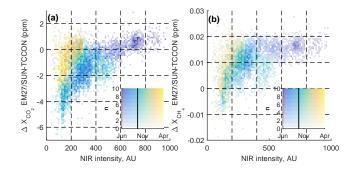
Figure 5. (a) The X_{CO_2} retrieved from the EM27/SUN compared to TCCON decreased with signal intensity for the first set of mirrors. In October the mirrors were changed, which caused the retrieved X_{CO_2} to increase. The inset is the legend for the average date and number of points in the histogram bins. **(b)** X_{CH_4} retrieved from the EM27/SUN compared with TCCON.

Figure 6. Top row: averaging kernels from the Caltech EM27/SUN instrument. Bottom row: averaging kernels from the TCCON.

To compare measurements between the TCCON and the EM27/SUN instruments, data were first averaged into 10 min bins to reduce the variance of binned differences (Chen et al., 2016). The median of the X_{CO_2} differences between sequential time bins is smallest (around 0.26 ppm) for 10 min bins over the entire ∼ 11 month time period. Less averaging is more affected by noise, and more averaging starts to include instrument drift and true atmospheric variations. Averages were weighted using retrieval errors $\hat{x}_{\mathrm{err}}$ as in Eq. (4):

$$\hat{\bar{x}} = \frac{\sum\limits_{i}\hat{x}_i\hat{x}_{i,\mathrm{err}}^{-2}}{\sum\limits_{i}\hat{x}_{i,\mathrm{err}}^{-2}}, \tag{4}$$

where $\hat{x}_i$ is the retrieved value from the ith measurement in a bin, and $\hat{\bar{x}}$ is the bin average.

5.1 Averaging kernels

When comparing retrieved X_{gas} measurements (also denoted $\hat{c}$) from different remote sensing instruments, differences in their averaging kernels (AKs or $\boldsymbol{a}_i$, where i represents an instrument indicator number) and a priori profiles must be taken into account, using for example, the methods described by Rodgers and Connor (2003). Wunch et al. (2011a) compared GOSAT and TCCON total column DMFs using this method. Because GGG scales a priori profiles rather than retrieving the full profile, these AKs are vectors (i.e., column averaging kernels) rather than matrices.

Averaging kernels depend on several factors including how strong the lines are in the retrieval windows, and viewing geometry (e.g., SZA for solar-viewing instruments). Because the TCCON IFS 125HR and EM27/SUN instruments have different spectral resolutions, the apparent absorption strengths are different and so are the averaging kernels. Averaging kernels for a gas differ for each microwindow. We
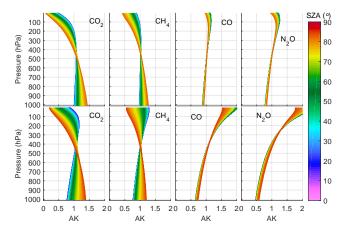
combined AKs of a given gas from different microwindows using an unweighted average. Averaging kernels for the Caltech EM27/SUN for the GGG retrieval windows are shown in Fig. 6. Averaging kernels from the other EM27/SUN instruments are similar. TCCON averaging kernels have been discussed by Wunch et al. (2011b) and are shown on the bottom row in Fig. 6. As a numerical example, for X_{CO_2} measured at 50° SZA and 900 hPa using GGG, the AK is 1.10 for EM27/SUN instruments and 0.93 for TCCON instruments. This means EM27/SUN instruments are slightly more sensitive to a change in CO_2 near the surface relative to TCCON instruments. More importantly, they have the opposite sensitivity to an error in the a priori volume mixing ratio (VMR) profile at 900 hPa.

In our particular case, reducing the smoothing error using Eq. (A13) from Wunch (2011a) and using the a priori as the comparison ensemble changes little as the effect of the differences in averaging kernels from the top of the atmosphere tends to cancel out the effect of differences at the bottom. TCCON and EM27/SUN a priori profiles were the same in this comparison. However, we need to consider that the a priori profiles used in the retrieval are not representative of a highly polluted place, such as Pasadena, which is located in the same air basin as Los Angeles. Because differences in column measurements compared to background or a priori profiles occur primarily because of differences at the surface we can adjust retrievals for one instrument taking into account this knowledge using

$$\hat{c}_1 = \frac{a_{1,s}}{a_{2,s}}\left[\hat{c}_2 - c_\mathrm{a}\right] + c_\mathrm{a}. \tag{5}$$

Definitions of the terms in, as well as a discussion of assumptions needed to obtain Eq. (5) are in Appendix A. We applied Eq. (5) to the X_{CO_2} and X_{CH_4} retrievals.

In summary, to compare biases between two instruments, we account for diurnal dependences, then average data into comparable time bins, and take into account our prior knowl-
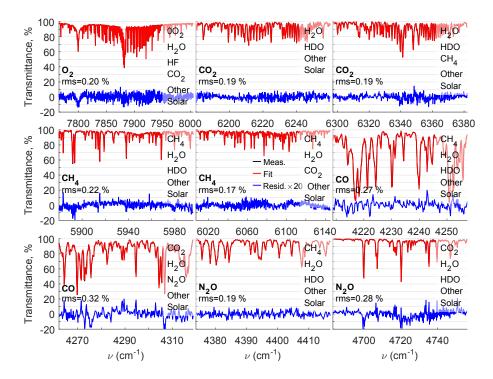
Figure 7. Example spectral fits and residuals ($\times$ 20) from several of the retrieval windows from 28 June 2014, 10:55:16 (UTC$-$8), SZA $=$ 17.1°. The primary species of interest and root mean square (rms) of the residuals are listed on the left. Other species fit in the window are listed on the right.

edge of the atmospheric profile and differences in averaging kernels.

5.2 Full comparisons of X$_{gas}$ from extended-band InGaAs detector with a TCCON site

Gisi et al. (2012) noted that measurements taken within the first 30 min of moving the instrument to the roof and turning it on needed to be filtered out because of high scatter while waiting for the instrument to operate stably. We did not observe a similar requirement for our data. This could be because our instruments were not subjected to such fast temperature changes. It could also be because the laser frequency shift, which changes with temperature, does not seem to significantly impact our retrievals.

Examples of spectral fits from several of the retrieval windows are shown in Fig. 7 for a single spectrum. These are not necessarily representative of all the conditions under which the 800 000 spectra were acquired. The residuals are larger than those reported by Gisi et al. (2012) and Frey et al. (2015) because of the lower SNR from spectra recorded using the extended InGaAs detector.

The full time series (186 days) of the difference between the Caltech EM27/SUN and TCCON measurements is shown in Fig. 8. From this figure we see that X$_{CO_2}$ and X$_{CH_4}$ are the gases most affected by the mirror change in October 2014 (by about 3 ppm and 12 ppb respectively). For all

gases, scatter of retrieved X$_{gas}$ increases as signal decreases. Figure 9 shows the retrieved X$_{CO_2}$ and X$_{CH_4}$ from all four EM27/SUN instruments for 9–12 days in January 2015 plotted against those from TCCON. We report biases for January 2015 as scaling factors to approximate to the TCCON, or scaling factors compared to 1. Biases were calculated using a linear least squares fit forced through the origin. A summary of the biases for all gases as compared to the TCCON is provided in Table 3.

5.3 X$_{CO_2}$

We note a smaller bias in X$_{CO_2}$ with respect to the TCCON (+0.03 %, see Table 3) compared to previous EM27/SUN studies (Gisi et al., 2012; Frey et al., 2015; Klappenbach et al., 2015). These previous studies retrieved X$_{gas}$ from EM27/SUN spectra using PROFFIT. When compared with the TCCON X$_{CO_2}$ retrievals, Gisi et al. (2012) noted a +0.12 % bias, Frey et al. (2015) noted a +0.49 % bias, and Klappenbach et al. (2015) noted a +0.43 % bias. Reasons for these differences could be from (1) spectroscopy differences between PROFFIT and GGG2014 used for EM27/SUN X$_{gas}$ retrievals, (2) because Gisi et al. (2012) used an earlier version of GGG for TCCON retrievals, and (3) because Frey et al. (2015) and Klappenbach et al. (2015) applied empirical corrections before comparing with the TCCON. In this

Table 3. EM27/SUN to Caltech TCCON biases.

	Caltech, January	LANL, January	Harvard 1	Harvard 2	weighted % bias
n	285	241	187	164	
n days	12	12	9	9	
X_{CO_2}	0.9999 (*0.16*)	1.0006 (*0.14*)	1.0009 (*0.15*)	0.9998 (*0.15*)	0.03
X_{CH_4}	1.0069 (*0.19*)	1.0066 (*0.20*)	1.0103 (*0.14*)	1.0066 (*0.14*)	0.75
X_{H_2O}	0.9840 (*1.27*)	0.9791 (*1.44*)	0.9886 (*1.12*)	0.9791 (*1.01*)	−1.73
X_{CO}	0.9988 (*2.30*)				−0.12
X_{N_2O}	1.0243 (*0.42*)				2.43

Italicized values in parentheses are percent standard deviations as compared to the TCCON over the dataset for January 2015.
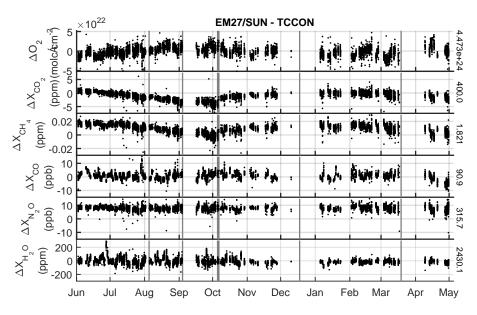

Figure 8. Full time series of EM27/SUN measurements as compared to TCCON from June 2014 to May 2015. Thin vertical gray lines represent mirror cleaning. The thick line represents the mirror change. To the right are TCCON means over time to get a sense of percent deviations.

section, we investigate two possible causes of bias: spectral resolution and instrument line shape.

Following Gisi et al. (2012), we attempted to determine whether the cause of the bias is due to the difference in spectral resolution between the EM27/SUN and TCCON instruments. Petri et al. (2012) also considered resolution bias in their study using a $0.11\,\text{cm}^{-1}$ resolution instrument and an older version of GGG. They did not report a bias in X_{CO_2} retrievals, but noted that X_{CO_2} decreased by $\sim 0.12\,\%$ as interferograms were truncated to obtain spectra with resolutions of 0.02 to $0.5\,\text{cm}^{-1}$. Most of the change occurred as the resolution changed from 0.1 to $0.5\,\text{cm}^{-1}$ (see Fig. 11 therein). In contrast, Gisi et al. (2012) noted a 0.13 % increase in X_{CO_2} as the resolution changed from 0.02 to $0.5\,\text{cm}^{-1}$ in PROFFIT. Here we find a $0.08\,\% \pm 0.16\,\%$ (1σ) decrease in X_{CO_2} when the resolution is decreased from 0.02 to $0.49\,\text{cm}^{-1}$ in GGG, though part of this change would be offset by considering the differences in averaging kernels.

Previous studies noted an increase in X_{CO_2} of 0.15 % for a 1 % increase in modulation efficiency at max OPD (Gisi et al., 2012; Frey et al., 2015). Using PROFFIT we performed a similar test for spectra taken under various conditions at various times of day and obtained a similar result of a $0.10\,\% \pm 0.02\,\%$ (1σ) increase in X_{CO_2} for a 1 % increase in ME at the MOPD. For this study we assume that impacts of the ILS on retrievals will be similar in GGG and PROFFIT. Though we report a single value, there is an air mass dependence of $\sim 0.05\,\%$ increase in EM27/SUN PROFFIT retrievals for a 1 % increase in ME and air mass change of 1.

For instruments using the standard InGaAs detectors, the X_{CO_2} 10 min running 1σ precision is 0.075 % [0.034 to 0.18 %, 95 % CI]. The wide confidence interval (CI) is from a combination of atmospheric variability being aliased into the running standard deviation as well as different SNRs among instruments. The spectral SNRs for measurements using this detector were in the range 1000–5000 and their pre-
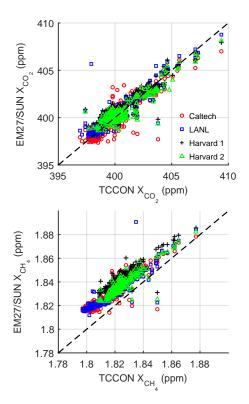
Figure 9. Retrieved EM27/SUN measurements (10 min averaging) as compared to the TCCON from January 2015. This provides a visual representation of the data – offset and scatter of data between X$_{gas}$ from different instrument types – in Table 3. The black dashed line is the 1 : 1 line.

cision for X$_{CO_2}$ retrievals was only weakly correlated with $1/\sqrt{SNR}$. Chen et al. (2016) found that the 1σ X$_{CO_2}$ precision among 10 min binned EM27/SUN$_a$-EM27/SUN$_b$ differences is 0.01 %. These data were acquired in a way that about 67 spectra were acquired every 10 min, and because two instruments were used, the single sounding precision is $\sim 0.01\% \times \sqrt{67/2} \approx 0.058\%$, which falls in our measured running 1σ precision range. Comparing to the TCCON, Gisi et al. (2012) reported that the 1σ daily precision is 0.08 %. The extended InGaAs detector naturally has a lower spectral SNR, in the range 100–1000, with a median of 400 over the full time series. Most of the variation in the SNR is due to loss of mirror reflectivity, but even with non-degraded gold mirrors, it is ~ 5 times lower because of the different detector. The median running 1σ precision over the full time series is 0.26 % for the X$_{CO_2}$ product from the extended InGaAs detector. Because the SNR changed with time due to loss of mirror reflectivity, so did the precision. The correlation between $1/\sqrt{SNR}$ and running 1σ X$_{CO_2}$ precision was strong ($R^2 = 0.75$) for retrievals from this detector and followed

$$\sigma_{XCO_2} = 0.17 + \frac{8.4}{\sqrt{SNR - 57}}. \qquad (6)$$

An additional study we have not performed that could help in reducing bias would be to omit all or part of a CO$_2$ window with strong water lines. Because of the low resolution of these spectrometers (see inset Fig. 1), water lines and CO$_2$ lines often overlap. This can lead to inaccurate retrievals despite a good overall fit because H$_2$O and CO$_2$ can both be wrong, but in compensating ways. Reducing the size of a window would reduce precision but would decrease water and temperature sensitivity. This adjustment could also be performed for CH$_4$, which is retrieved over three windows in GGG.

5.4 X$_{CH_4}$

The EM27/SUN X$_{CH_4}$ retrievals are 0.75 % higher than those of TCCON (see Table 3). In previous work, high biases of 0.47 % for a 0.11 cm^{-1} instrument (Petri et al., 2012), and 0.49 % (Frey et al., 2015) and 1.87 % (Klappenbach et al., 2015) for EM27/SUNs, were noted. Petri et al. (2012) attributed most (0.26 %) of their bias to differences in resolution and noted for a single day that the bias increased as resolution decreased. In our simulations we find a 0.28 % ± 0.20 % (1σ) increase in X$_{CH_4}$ when the resolution is reduced from 0.02 to 0.49 cm^{-1}. Using PROFFIT the impact of a 1 % decrease in ME is a 0.15 % ± 0.01 % (1σ) increase in X$_{CH_4}$. Again, although we report a single value there is an air mass dependence of about a 0.12 % decrease in X$_{CH_4}$ using PROFFIT retrievals for an air mass change of 1, and a 1 % decrease in ME. Resolution and ME combined account for only half of the observed methane bias. Petri et al. (2012) suggested improper dry air mixing ratio and pT profiles, or spectroscopy as sources of error. Improper surface pressure, error in the calculated Observer-Sun Doppler Stretch (OSDS) due to pointing errors coupled with solar rotation, or error in the assumed field of view (FOV) may also contribute to the bias (see Sect. 6).

Chen et al. (2016) found that the 1σ X$_{CH_4}$ precision among 10 min binned EM27/SUN$_a$-EM27/SUN$_b$ differences is 0.01 %, which is equivalent to a single sounding 1σ precision of $\sim 0.058\%$. Using the same method as for X$_{CO_2}$, the X$_{CH_4}$ running 1σ precision from instruments using the standard InGaAs detectors is 0.057 % [0.037 to 0.25 %, 95 % CI], in agreement with Chen et al. (2016). The median running 1σ precision for X$_{CH_4}$ from instruments using the extended InGaAs detector is 0.33 %. X$_{CH_4}$ precision from the extended InGaAs measurements is also correlated with $1/\sqrt{SNR}$.

5.5 X$_{CO}$ and X$_{N_2O}$

X$_{N_2O}$ and X$_{CO}$ were also measured using an EM27/SUN spectrometer in this study. Hase et al. (2016) have also reported on X$_{CO}$ measurements using an EM27/SUN modified to include a second InGaAs detector with optical filters. Column CO measurements are desirable because CO is a tracer of combustion. Here these measurements were made possible

because the extended detector is sensitive to the region 4200–4800 cm^{-1}, which contains useful windows where N_2O and CO molecules absorb IR radiation. Both the X_{CO} and X_{N_2O} retrievals are highly sensitive to changes in the modeled temperature profile. The nonlinearity of the detector had a less pronounced effect on X_{CO} and X_{N_2O} retrievals than it had on X_{CO_2} and X_{CH_4} retrievals (Fig. 8). X_{CO} and X_{N_2O} also have poorer precision than X_{CO_2} and X_{CH_4}, so any nonlinearity effect could be less than the noise. The 4200–4800 cm^{-1} spectral region is also affected differently from the nonlinearity than the 5000–7000 cm^{-1} region where column CH_4 and CO_2 are retrieved; the continuum levels changed more for the latter region. This may also explain in part why there is no noticeable change in X_{CO} and X_{N_2O} with signal. For X_{CO} the median 1σ precision is 3.7 %. In our simulations reducing the spectral resolution from the TCCON (0.02 cm^{-1}) to near the EM27/SUN (~ 0.5 cm^{-1}), X_{CO} decreases 2.5 % $\pm$ 4.2 % (1σ) in low-resolution spectra, and at Caltech this change varies with time.

In general, as is seen in Fig. 8, X_{N_2O} retrievals were highly scattered and had a large offset from TCCON. In our simulations, reducing the resolution from TCCON (0.02 cm^{-1}) to EM27/SUN (0.5 cm^{-1}) decreased X_{N_2O} by 1.5 % $\pm$ 0.6 % (1σ). Retrievals from the 4430 cm^{-1} window were low (~ 6 %), while the 4719 and 4395 cm^{-1} regions were biased slightly high (~ 1 %). The retrievals from the 4719 cm^{-1} region additionally had some long-term trends for reasons we do not understand. For X_{N_2O} the median 1σ precision is 1.9 %.

5.6 X_{H_2O}

Because of the significantly lower spectral resolution of the EM27/SUN spectrometers, the spectral band widths for the H_2O retrievals were increased as compared to the standard TCCON approach (Wunch et al., 2010). For lower resolution spectra, the H_2O lines appear much broader and the observed transmittance is much lower at the edges of standard TC-CON spectral window. Thus, the spectral ranges of the low-resolution windows were expanded. Some of the standard TCCON windows used to retrieve H_2O had too few spectral points from the low-resolution instrument for good fits and were omitted. When expanding the windows, we ensured that no lines were admitted that made the effective ground-state energy E'' greater than ~ 400 cm^{-1}. This reduces the temperature sensitivity to the modeled temperature profiles. As with the TCCON windows, we tried to keep a wide range of H_2O line strengths to accommodate large seasonal and site-to-site variations of the H_2O column. Windows were kept as wide as possible without encountering large spectral fitting residuals.

For X_{H_2O}, we find a median 1σ precision of 1.9 % from the instrument using the extended InGaAs detector. For instruments using the standard-InGaAs detectors, the X_{H_2O} 1σ precision is 0.81 % [0.36 to 2.12 %, 95 % CI].
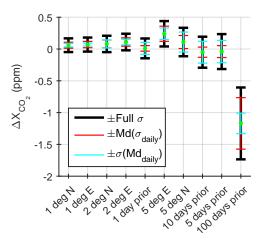

Figure 10. Standard deviations and biases from using wrong model pTz and H_2O profiles as compared to using the standard option for time and location. Tests are in order of increasing full σ. Red represents intraday variability. Cyan represents interday variability.

6 Sensitivity tests on retrievals

As with the TCCON, EM27/SUN retrievals require modeled atmospheric pressure, temperature, altitude (pTz), and water profiles (Wunch et al., 2015). Here atmospheric profiles are generated from the NCEP/NCAR 2.5° reanalysis product (Kalnay et al., 1996) by interpolating to the correct location at local noon of the desired day. These profiles also include the tropopause height which is used to vertically shift a priori profiles, as tropopause height can significantly affect column DMFs such as X_{CH_4} and X_{HF} (Saad et al., 2014). Selecting a profile for an incorrect location or day could lead to errors.

We ran test retrievals for the July 2014 period with incorrect profile information derived separately at latitudes north (1, 2, and 5°) and longitudes west (1, 2, and 5°) of our observation site, and well as from profiles derived 1, 5, 10, and 100 days prior to the measurement dates. In general, the profiles generated from a more distant location in space and time caused larger retrieval errors. For X_{CH_4} and X_{CO}, the main variability from the standard retrievals was in daily offsets (standard deviation of daily medians $\sigma \left(\text{Md}_{\text{daily}} \right)$) which had values of 3 and 4 ppb respectively for the 100 day prior model. The medians of daily standard deviations $\text{Md} \left(\sigma_{\text{daily}} \right)$ were 0.5 ppb for both X_{CH_4} and X_{CO} for the 100 day prior model. X_{N_2O} and X_{H_2O} also had more errors from $\sigma \left(\text{Md}_{\text{daily}} \right)$, except for profiles within 2°, which more strongly affected diurnal variability $\text{Md} \left(\sigma_{\text{daily}} \right)$. For these two species, the 100 day prior model $\sigma \left(\text{Md}_{\text{daily}} \right)$ were 2 ppb and 50 ppm and $\text{Md} \left(\sigma_{\text{daily}} \right)$ were 1 ppb and 20 ppm respectively. These values are shown for X_{CO_2} in Fig. 10 for all tested models. The 100 day prior model had $\sigma \left(\text{Md}_{\text{daily}} \right) = 0.16$ ppm and $\text{Md} \left(\sigma_{\text{daily}} \right) = 0.4$ ppm, as well as a 1.2 ppm bias when using these models for X_{CO_2}.

Table 4. Meteorological sensitivity tests on EM27/SUN retrievals.

	X_{CO_2}		X_{CH_4}		X_{CO}		X_{N_2O}	
Error	Offset	Daily	Offset	Daily	Offset	Daily	Offset	Daily
+1 hPa surf	0.032	0.004	0.036	0.010	0.10	0.14	0.06	0.18
+10 K (surf – 700 hPa)	0.257	0.076	−0.006	0.036	10.1	1.2	0.53	0.23

Errors expressed as percentages. Daily is the median of the daily standard deviations, $\text{Md}\left(\sigma_{\text{daily}}\right)$.

Table 5. Perturbations used in uncertainty budget.

Perturbation	Magnitude
ap[a] volume mixing ratio (VMR)	downshift by 1 km[b]
ap temperature	+1 K all altitudes
ap pressure	+1 hPa all altitudes
Pointing offset (po)	increased by 0.05°
Surface pressure	+1 hPa
Calculated OSDS[c]	+2 ppm
Field of view (FOV)	+7 %

See also Fig. 11. [a] ap denotes a priori. [b] ap VMRs were shifted independently. For X_{H_2O} and X_{HDO}, concentrations were decreased by 50 % at all levels. [c] OSDS = observer sun Doppler stretch.
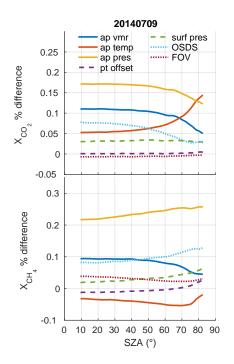

Figure 11. Uncertainty budget for EM27/SUN instruments using GGG2014. See Table 5 for magnitudes of perturbations.

Various user, instrumental, and measurement errors can reduce the accuracy and precision of retrievals. GGG uses retrieved O_2 column amount with the average DMF of O_2 (0.2095) to calculate the dry pressure column of air. However, to calculate the O_2 absorption coefficients, GGG takes into account the surface pressure, which can lead to measurement inaccuracies if the wrong surface pressure is used. Wunch et al. (2011b) reported a 0.04 % X_{CO_2} bias for a +1 hPa surface pressure offset in the TCCON. Similarly, we find a 0.032 % X_{CO_2} bias per +1 hPa surface pressure offset, with a 0.004 % σ variation on average throughout a day. Because the pressure offset affects O_2 retrievals, the other species are also affected (Table 4). X_{CO} may be particularly affected by a pressure bias because such a large fraction of the column CO is near the surface.

Using the same July 2014 dataset used to test the sensitivity of the retrievals to error in the pTz profile and surface pressure, we further estimated the sensitivity to error in the temperature in the lower atmosphere (surface – 700 hPa). GGG uses a single temperature profile per day that represents the local-noon temperatures, and the surface temperature is extracted from that profile. Such temperature error can arise in particular at the beginning and end of the day when the temperature is typically cooler than at noon. Here we derived the sensitivity of the retrievals to a +10 K error in the lower atmosphere (Table 4). X_{CO} has a significantly larger bias than the other species, likely because water absorption lines are the strongest spectral features in the CO retrieval window and water absorption lines are highly sensitive to changes in temperature. Water lines are also much stronger than N_2O lines in the N_2O windows. These tests suggest that offsets under 1 hPa and 1 K would cause small (~ 0.1 ppm) biases on X_{CO_2}, but a 4 K difference in near-surface (ground – 700 hPa) temperature could cause ~ 0.4 ppm bias in X_{CO_2} which is larger than our reported 1σ precision. For other studies using multiple spectrometers and multiple meteorological measurements for X_{gas} retrievals, we recommend cross-comparing meteorological measurements to eliminate bias – preferably to a standard.

Finally, we perform a sensitivity study following the methodology of Wunch et al. (2015). The magnitudes of the applied perturbations are in Table 5. The results of this uncertainty budget study are presented for a day for X_{CO_2} and X_{CH_4} in Fig. 11. We do not include a sum in quadrature because we do not have an exhaustive list of sources of uncertainty. This uncertainty budget indicates that the low-resolution instruments are especially sensitive to biases in a priori pressures and a priori volume mixing ratio (VMR) profiles. Some of these errors may partially account for the unexplained long-term drifts we noted compared to TC-

Table 6. Tests for assessing biases and sensitivities of solar-viewing, remote sensing instruments.

Assessment	Test/observation	Type	Accepted correction[a]	Root cause	Similar instr. effect	EM27/SUN test
Incoming radiation attenuation effect	Gray filter after solar tracker & before interferometer	M	Recom'd replace detector. Alt. empirical	Detector nonlinearity	Consistent for same detectors	Sect. 4.4
ILS	Measure with low-p gas cell (preferred), stable laser, or ambient air (least recom'd)	M	Retrievals with non-ideal ILS	Instrument misalignment; in-built	Potentially large differences	Gisi et al. (2012); Chen et al. (2014); Frey et al. (2015); Sects. 4.1 (measured), 5.3, 5.4 (impacts)
	Adjust FOV (if ILS is measured but not accounted for in retrieval)	RIA	Not recom'd			
Ghost to parent ratio	Use blackbody source & narrow band filter post-interferometer	M		Laser mis-sampling	Likely similar, potentially large diffs	Gisi et al. (2012); Frey et al. (2015); Sect. 4.3
Ghost effects	Measurements with & without ghost correction (e.g., XSM, or ifg resampling before FFT)[b]	M or RIA	Recom'd interpol. during acq or post-resampling	Laser mis-sampling	Likely similar, potentially large diffs	Sect. 4.3
Frequency shifts	Changes or large 0 offset	O & RIA	Input spectral spacing	Improper laser wavenumber, misalignment of laser or NIR beam	Shifts differ, effect similar	Sect. 4.2
Solar gas stretch	Changes or large 0 offset	O & RIA	OSDS	Poor spectral fits of solar lines; SE or res.	Similar for same detector & res.	Sect. 6
Spectral fitting windows	Width, locations	RIA		Instrument resolution requires adaptation	Same for similar res. (widths) & detector (locations)	Gisi et al. (2012); Sect. 5.7 (H_2O) Sect. 5.3 (discussion)
Averaging kernels	Used when comparing with a different instrument type	O	Rodgers and Connor (2003) and prior info.	Diff. sensitivity at atmos. layers from differing resolutions[b] & VG	Same for similar res., microwindows & VG	Sect. 5.1
SZA artifacts	Multi-day measurements in clean location	O	Empirical[a] (Wunch et al. 2011b)	ILS, or SE	See ILS entry	Frey et al. (2015); Parker et al. (2015)
Long-term artifacts	Preferred co-location with accepted measurements (e.g., TCCON)	O		Various (e.g., instrument settling, changing alignment, other)	May widely differ	Herein – for extended InGaAs only
Region/zone dependence	Co-location with spatially distributed accepted measurements	O/M		A priori insufficiencies	Likely similar	Parker et al. (2015)
Surface pressure effects	Manually adjust pressure inputs.	RIA	Accurate barometer pres. calibr.	Poor calculation of O_2 column, directly or by poor fitting	Similar effects for similar resolutions	Sect. 6
pTz & H_2O model profile sensitivity	Adjust modeled meteorological profiles	RIA	Improve met. profiles	Non-representative pTz+H_2O profile	Similar effects for similar resolutions	Sect. 6
A priori VMR surface sensitivity	Adjust a priori VMR near surface	RIA	Improve a priori profiles; reduce effect with AKs	Non-representative VMR profile (e.g., polluted mixed layer)	Similar effects for similar res. & true VMR profile	Parker et al. (2015)
Opt. avg. time	Allan type plot; e.g., Chen et al. (2016)	O	Empirical	SNR & true atmospheric variation	Depends on SNR & location	Chen et al. (2016) Sect. 5
Resolution effects	Truncate high-resolution ifg	RIA	Apply offset	Inst. res.	Similar for all solar-viewing insts.	Gisi et al. (2012); Petri et al. (2012) Sects. 5.3–5.6
Uncertainty budget for current fitting algorithm	Various, test on each new algorithm (Wunch et al. 2015)	RIA	Informative	Various	Similar effects for similar resolutions	Sect. 6

M denotes measurement (setups/adjustments required before acquisition), RIA denotes retrieval input adjustment (post-data acquisition, pre-retrieval), O denotes observation post-retrieval (may require prior planning of locations of measurements or longer term measurements), SE denotes spectroscopy errors, VG denotes viewing geometry, res denotes resolution.
[a] Though empirical corrections are occasionally accepted, it is always recommended to correct the underlying problem(s) if possible.
[b] XSM is Bruker™ code for interpolation during acquisition.
[c] GGG can provide ifg resampling if two detectors are on instrument. Note that the preferred correction is always of the root cause.
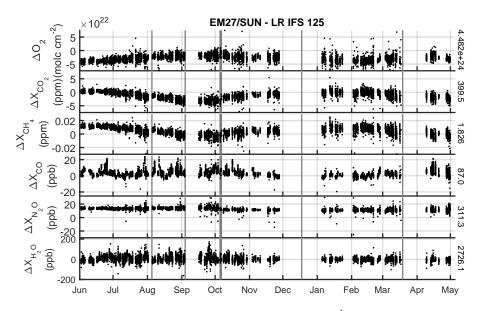
Figure 12. Time series comparison of EM27/SUN retrievals to retrievals from the $0.5\,\mathrm{cm}^{-1}$ resolution IFS 125HR spectra.

CON that are unrelated to signal (e.g., Fig. 8, October–November 2014). For example, surface pressure and calculated Observer-Sun Doppler Stretch (OSDS) were correlated with EM27/SUN to TCCON X_{CO_2} differences in the long-term measurement. However, there was no apparent trend in the spectral residuals from fitting solar lines as the OSDS changed so these correlations may not indicate cause.

Differences in X_{gas} between different instruments are due to a combination of differences in resolution, and real instrumental imperfections and instabilities. To attempt to distinguish between resolution causing differences (e.g., by limitations in the forward model) or instrumental issues, we repeat the test performed by Gisi et al. (2012, Fig. 11 therein) of truncating IFS 125HR interferograms for the full time series. Results are shown in Fig. 12. Mean values for X_{CO_2} are slightly lower because of differences from retrievals on spectra of different resolutions, as described in Sect. 5.3. When comparing 10 min averaged TCCON data with lower resolution IFS 125HR retrievals we note monthly standard deviations on order of 0.15 % for X_{CO_2} and X_{CH_4}. This suggests the standard deviations of comparing retrievals from the EM27/SUN with the TCCON (Table 3) on these timescales are close to the current precision limits for directly comparing X_{CO_2} and X_{CH_4} retrieved from spectra of these different resolutions. Results in Fig. 12 are slightly more scattered than in Fig. 8 and have different offsets. The data still show an increase in X_{CO_2} and X_{CH_4} in October–November 2014 for reasons we do not understand, and unfortunately we have no ILS characterizations over this period.

Long-term drifts may or may not affect instruments employing the standard InGaAs detector and may be eliminated by future retrieval updates. They may also arise in part from how the comparison was made, e.g., the assumptions

to derive Eq. (A4) may not be valid for CH_4 and N_2O. As a follow-up study, brief 5–6 day comparisons using a standard InGaAs detector were made for the months of August, September, and November 2015. Scaling factors varied from 0.99905 to 1.00001 for X_{CO_2} and from 1.01228 to 1.00893 for X_{CH_4}, with larger day-to-day variability. Long-term (1 year or more) comparisons of these instruments employing the standard-InGaAs detector are needed before claims of long-term accuracy can be made or the full magnitude of drift can be quantized. Errors that could lead to drifts likely would be correlated amongst all EM27/SUN instruments, so the comparison would need to be against a standard such as the TCCON. Future studies may also benefit from comparing results using different retrieval algorithms, as the magnitude of errors that may lead to drifts in X_{gas} may vary among algorithms. Meanwhile, operators have already found many purposeful ways to use these instruments that require only short-term (about 1 month) precision without any assumptions about precision for longer time periods (for example Hase et al., 2015; Chen et al., 2016; Viatte et al., 2016). Studies using these spectrometers independently longer term can also be performed depending on the degree of precision required. Limits on precision described herein are likely to only improve in future work.

7   Conclusions

Despite the challenge associated with the extended InGaAs detector and mirror degradation, the EM27/SUN instruments perform well on short timescales with 1σ running 10 min precisions of 0.075 % for X_{CO_2} and 0.057 % for X_{CH_4} retrieved from measurements using the standard InGaAs detectors. These instruments perform well in terms of mobility and

stability, maintaining alignment despite frequent movement and jostling – an ideal characteristic of mobile FTS instruments. Measurements from the standard detector are precise enough to be used for campaigns of up to a few months and to provide useful supplementary X_{gas} measurements to established networks like TCCON. However, we recommend regular – 6 months to 1 year depending on use – comparison with established measurements (e.g., a TCCON site) to account for long-term drift. The frequency of comparison with established measurements may need to be reevaluated when more long-term comparison data become available. Simultaneous use of several EM27/SUN instruments may also help characterize drift. We also recommend regular – about monthly depending on use – ILS characterization. Our experience also suggests that use of the extended InGaAs detector without limiting the spectral band-pass in the EM27/SUN is incompatible with X_{CO_2} and X_{CH_4} retrievals that are precise long-term.

In general, we recommend all new ground-based, solar-viewing, remote sensing FTS instruments to undergo some or all tests listed in Table 6 to evaluate their performance. We also recommend comparisons of retrieval outputs to those of existing instrumentation (e.g., TCCON or NDACC-IRWG). These tests assume that one of the three widely used and accepted retrieval algorithms (GGG, PROFFIT, and SFIT), known to provide accurate spectral fitting, is used. New retrieval algorithms should be subjected to additional comparisons with currently accepted algorithms. Some of the results of these tests will be similar across all instruments of a given type, and so do not need to be repeated if they have been performed on another instrument elsewhere.

Appendix A: Assumptions and limitations in the AK correction

To derive Eq. (5), we begin with Eq. (22) in Rodgers and Connor (2003):

$$\hat{c}_i = c_\mathrm{a} + \sum_\mathrm{k} h_\mathrm{k} a_{i,\mathrm{k}} \cdot \left(x_{\mathrm{t,k}} - x_{\mathrm{a,k}}\right) + \epsilon_i. \tag{A1}$$

To include the pressure-weighting function $\boldsymbol{h}$ (Connor et al., 2008), we have used summation notation. The "hat" represents a retrieved value, c represents a column (scalar) value, and ϵ is the error. Subscript i is for a particular instrument, subscript a represents the a priori, subscript k is for a particular atmospheric layer, and subscript t represents the true atmosphere. The vectors $\boldsymbol{a}$ and $\boldsymbol{x}$ represent the column averaging kernel and atmospheric VMR profile respectively. This equation is derived from Eq. (1) in Rodgers and Connor (2003) using a Taylor series expansion about the a priori profile, and assuming linearity about it.

To compare retrievals from remote sounding instruments, a comparison profile (also called the comparison ensemble mean, denoted x_c) is used. Here, we have used the daily a priori profiles, which were the same for all instruments, as the comparison profiles. We note, however, that the comparison profiles should describe the real atmosphere as far as possible (Rodgers, 2000). Though the a priori profile has a drawdown in CO_2 from the biosphere near the surface, the real atmosphere in Pasadena is polluted near the surface. Thus this choice of comparison profiles is not ideal in our situation.

If we ignore retrieval error Eq. (A1), and further assume that $\boldsymbol{x}_\mathrm{t} = \boldsymbol{x}_\mathrm{a}$ except at the surface, it can be rewritten as

$$\frac{1}{a_{i,\mathrm{s}}} \left(\hat{c}_i - c_\mathrm{a}\right) = h_\mathrm{s} \left(x_{\mathrm{t,s}} - x_{\mathrm{a,s}}\right), \tag{A2}$$

where the subscript s represents a surface value. If we are comparing measurements from two different instruments, $i = 1$ and $i = 2$, in the same location, $x_{\mathrm{t,s}}$ and h_s are the same. Because the a priori profiles are also the same,

$$\frac{1}{a_{1,\mathrm{s}}} \left(\hat{c}_1 - c_\mathrm{a}\right) = \frac{1}{a_{2,\mathrm{s}}} \left(\hat{c}_2 - c_\mathrm{a}\right), \tag{A3}$$

which can be rewritten as

$$\hat{c}_1 = \frac{a_{1,\mathrm{s}}}{a_{2,\mathrm{s}}} \left(\hat{c}_2 - c_\mathrm{a}\right) + c_\mathrm{a}. \tag{A4}$$

Even in the absence of error, retrievals from instruments with different averaging kernels will still differ.

We adjust the EM27/SUN X_{CO_2} and X_{CH_4} retrievals using Eq. (A4) before comparison with the TCCON, which adjusts X_{CO_2} by up to ~ 1.2 ppm and X_{CH_4} by up to ~ 8 ppb. Future work could improve on this methodology using a better comparison ensemble or more representative a priori profiles for retrievals from measurements in Pasadena. This correction is not applied to X_{H_2O} because the AKs vary more among spectra because of larger variations in absorption strengths. It is also not applied to X_{CO} and X_{N_2O} because using $\boldsymbol{x}_\mathrm{t} = \boldsymbol{x}_\mathrm{a}$ is too poor of an assumption and makes the comparison worse between the TCCON and EM27/SUN retrievals in terms of R^2.

Competing interests

The authors declare that they have no conflict of interest.

References

Bennett, J. M. and Ashley, E. J.: Infrared Reflectance and Emittance of Silver and Gold Evaporated in Ultrahigh Vacuum, Appl. Opt., 4, 221, doi:10.1364/AO.4.000221, 1965.

Chen, J., Samra, J., Gottlieb, E., Budney, J., Daube, C., Daube, B., Hase, F., Gerbig, C., Chance, K., and Wofsy, S.: Boston Column Network: Compact Solar-Tracking Spectrometers and Differential Column Measurements, in: American Geophysical Union Fall Meeting, San Francisco, California, 15–19 December, Abstract ID: A53L-3381, doi:10.13140/RG.2.1.2284.1361, 2014.

Chen, J., Viatte, C., Hedelius, J. K., Jones, T., Franklin, J. E., Parker, H., Gottlieb, E. W., Wennberg, P. O., Dubey, M. K., and Wofsy, S. C.: Differential column measurements using compact solar-tracking spectrometers, Atmos. Chem. Phys., 16, 8479–8498, doi:10.5194/acp-16-8479-2016, 2016.

Connor, B., Boesch, H., Toon, G., Sen, B., Miller, C., and Crisp, D.: Orbiting Carbon Observatory: Inverse method and prospective error analysis, J. Geophys. Res., 113, D05305, doi:10.1029/2006JD008336, 2008.

Dohe, S., Sherlock, V., Hase, F., Gisi, M., Robinson, J., Sepúlveda, E., Schneider, M., and Blumenstock, T.: A method to correct sampling ghosts in historic near-infrared Fourier transform spectrometer (FTS) measurements, Atmos. Meas. Tech., 6, 1981–1992, doi:10.5194/amt-6-1981-2013, 2013.

Feist, D. G., Arnold, S. G., Hase, F., and Ponge, D.: Rugged optical mirrors for Fourier transform spectrometers operated in harsh environments, Atmos. Meas. Tech., 9, 2381–2391, doi:10.5194/amt-9-2381-2016, 2016.

Frey, M., Hase, F., Blumenstock, T., Groß, J., Kiel, M., Mengistu Tsidu, G., Schäfer, K., Sha, M. K., and Orphal, J.: Calibration and instrumental line shape characterization of a set of portable FTIR spectrometers for detecting greenhouse gas emissions, Atmos. Meas. Tech., 8, 3047–3057, doi:10.5194/amt-8-3047-2015, 2015.

Gisi, M.: EM27/SUN, in: Annual Joint NDACC-IRWG & TCCON Meeting, Bad Sulza, Germany, May 12–14, available at: http://www.acom.ucar.edu/irwg/IRWG_2014_presentations/Wednesday_PM/Gisi_Bruker_EN27.pdf (last access: 4 May 2016), 2014.

Gisi, M., Hase, F., Dohe, S., Blumenstock, T., Simon, A., and Keens, A.: XCO$_2$-measurements with a tabletop FTS using solar absorption spectroscopy, Atmos. Meas. Tech., 5, 2969–2980, doi:10.5194/amt-5-2969-2012, 2012.

Hale, G. E.: The Astrophysical Observatory of the California Institute of Technology, Astrophys. J., 82, 111–139, doi:10.1086/143663, 1935.

Hase, F.: Improved instrumental line shape monitoring for the ground-based, high-resolution FTIR spectrometers of the Network for the Detection of Atmospheric Composition Change, Atmos. Meas. Tech., 5, 603–610, doi:10.5194/amt-5-603-2012, 2012.

Hase, F., Blumenstock, T., and Paton-Walsh, C.: Analysis of the instrumental line shape of high-resolution fourier transform IR spectrometers with gas cell measurements and new retrieval software., Appl. Opt., 38, 3417–3422, doi:10.1364/AO.38.003417, 1999.

Hase, F., Hannigan, J., Coffey, M. T., Goldman, A., Hopfner, M., Jones, N. B., Risland, C. P., and Wood, S. W.: Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements, J. Quant. Spectrosc. Ra., 87, 25–52, doi:10.1016/j.jqsrt.2003.12.008, 2004.

Hase, F., Frey, M., Blumenstock, T., Groß, J., Kiel, M., Kohlhepp, R., Mengistu Tsidu, G., Schäfer, K., Sha, M. K., and Orphal, J.: Application of portable FTIR spectrometers for detecting greenhouse gas emissions of the major city Berlin, Atmos. Meas. Tech., 8, 3059–3068, doi:10.5194/amt-8-3059-2015, 2015.

Hase, F., Frey, M., Kiel, M., Blumenstock, T., Harig, R., Keens, A., and Orphal, J.: Addition of a channel for XCO observations to a portable FTIR spectrometer for greenhouse gas measurements, Atmos. Meas. Tech., 9, 2303–2313, doi:10.5194/amt-9-2303-2016, 2016.

Irion, F. W., Gunson, M. R., Toon, G. C., Chang, A. Y., Eldering, A., Mahieu, E., Manney, G. L., Michelsen, H. a, Moyer, E. J., Newchurch, M. J., Osterman, G. B., Rinsland, C. P., Salawitch, R. J., Sen, B., Yung, Y. L., and Zander, R.: Atmospheric Trace Molecule Spectroscopy (ATMOS) Experiment Version 3 data retrievals., Appl. Opt., 41), 6968–6979, doi:10.1364/AO.41.006968, 2002.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, S., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bull. Am. Meteorol. Soc., 77, 437–471,

doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Keppel-Aleks, G., Toon, G. C., Wennberg, P. O., and Deutscher, N. M.: Reducing the impact of source brightness fluctuations on spectra obtained by Fourier-transform spectrometry., Appl. Opt., 46, 4774–4779, doi:10.1364/AO.46.004774, 2007.

Klappenbach, F., Bertleff, M., Kostinek, J., Hase, F., Blumenstock, T., Agusti-Panareda, A., Razinger, M., and Butz, A.: Accurate mobile remote sensing of XCO$_2$ and XCH$_4$ latitudinal transects from aboard a research vessel, Atmos. Meas. Tech., 8, 5023–5038, doi:10.5194/amt-8-5023-2015, 2015.

Kort, E. A., Angevine, W. M., Duren, R., and Miller, C. E.: Surface observations for monitoring urban fossil fuel CO$_2$ emissions: Minimum site location requirements for the Los Angeles megacity, J. Geophys. Res. Atmos., 118, 1–8, doi:10.1002/jgrd.50135, 2013.

Learner, R. C. M., Thorne, A. P., and Brault, J. W.: Ghosts and artifacts in Fourier-transform spectrometry, Appl. Opt., 35, 2947–2954, doi:10.1364/AO.35.002947, 1996.

Lindenmaier, R., Dubey, M. K., Henderson, B. G., Butterfield, Z. T., Herman, J. R., Rohn, T., and Lee, S.-H. Multiscale observations of CO$_2$, ^{13}CO$_2$, and pollutants at Four Corners for emission verification and attribution. Proceedings of the National Academy of Sciences of the United States of America, 111, 8386-8391, doi:10.1073/pnas.1321883111, 2014.

McKain, K., Wofsy, S. C., Nehrkorn, T., Eluszkiewicz, J., Ehleringer, J. R., and Stephens, B. B.: Assessment of ground-based atmospheric observations for verification of greenhouse gas emissions from an urban region, Proc. Natl. Acad. Sci. USA, 109, 8423–8428, doi:10.1073/pnas.1116645109, 2012.

Messerschmidt, J., Macatangay, R., Notholt, J., Petri, C., Warneke, T., and Weinzierl, C.: Side by side measurements of CO2 by ground-based Fourier transform spectrometry (FTS), Tellus, Ser. B Chem. Phys. Meteorol., 62, 749–758, doi:10.1111/j.1600-0889.2010.00491.x, 2010.

Parker, H. A., Hedelius, J., Viatte, C., Wunch, D., Wennberg, P. O., Chen, J., Wofsy, S., Jones, T., Franklin, J., Dubey, M. K., Roehl, C. M., Podolske, J. R., Hillyard, P. W., and Iraci, L. T.: Compact Solar Spectrometer Column CO$_2$, and CH$_4$ Observations: Performance Evaluation at Multiple North American TCCON Sites, in: American Geophysical Union Fall Meeting, San Francisco, California, 14–18 December, Abstract ID: GC11B-1030, 2015.

Petri, C., Warneke, T., Jones, N., Ridder, T., Messerschmidt, J., Weinzierl, T., Geibel, M., and Notholt, J.: Remote sensing of CO$_2$ and CH$_4$ using solar absorption spectrometry with a low resolution spectrometer, Atmos. Meas. Tech., 5, 1627–1635, doi:10.5194/amt-5-1627-2012, 2012.

Pougatchev, N. S., Connor, B. J., and Rinsland, C. P.: Infrared measurements of the ozone vertical distribution above Kitt Peak, J. Geophys. Res., 100, 16689–16697, doi:10.1029/95JD01296, 1995.

Rodgers, C.: Inverse methods for atmospheric sounding: Theory and practice, World Scientific, Singapore, 256 pp., doi:10.1142/3171, 2000.

Rodgers, C. D. and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108, 4116, doi:10.1029/2002JD002299, 2003.

Saad, K. M., Wunch, D., Toon, G. C., Bernath, P., Boone, C., Connor, B., Deutscher, N. M., Griffith, D. W. T., Kivi, R., Notholt, J., Roehl, C., Schneider, M., Sherlock, V., and Wennberg, P. O.: Derivation of tropospheric methane from TCCON CH$_4$ and HF total column observations, Atmos. Meas. Tech., 7, 2907–2918, doi:10.5194/amt-7-2907-2014, 2014.

Shiomi, K., Kuze, A., Kawakami, S., Kataoka, F., Hedelius, J., Viatte, C., Wennberg, P., Wunch, D., Roehl, C., Leifer, I., Tanaka, T., Iraci, L., Bruegge, C., and Schwander, F.: GOSAT CO$_2$ and CH$_4$ validation activity with a portable FTS at Pasadena, Chino, and Railroad Valley, in: American Geophysical Union Fall Meeting, San Francisco, California, 14–18 December, Abstract ID: A41I-0162, 2015.

Thompson, A. and Chen, H. M.: Beamcon III, a linearity measurement instrument for optical detectors, J. Res. Natl. Inst. Stand. Technol., 99, 751–755, doi:10.6028/jres.099.067, 1994.

Toon, G. C.: The JPL MkIV interferometer, Opt. Photonics News, 2, 19, doi:10.1364/OPN.2.10.000019, 1991.

Viatte, C., Lauvaux, T., Hedelius, J. K., Parker, H., Chen, J., Jones, T., Franklin, J. E., Deng, A. J., Gaudet, B., Verhulst, K., Duren, R., Wunch, D., Roehl, C., Dubey, M. K., Wofsy, S., and Wennberg, P. O.: Methane emissions from dairies in the Los Angeles Basin, Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-281, in review, 2016.

Washenfelder, R. A., Toon, G. C., Blavier, J. F., Yang, Z., Allen, N. T., Wennberg, P. O., Vay, S. A., Matross, D. M., and Daube, B. C.: Carbon dioxide column abundances at the Wisconsin Tall Tower site, J. Geophys. Res. Atmos., 111, 1–11, doi:10.1029/2006JD007154, 2006.

Wennberg, P. O., Wunch, D., Roehl, C., Blavier, J.-F., Toon, G. C., and Allen, N.: TCCON data from California Institute of Technology, Pasadena, California, USA, Release GGG2014R1., doi:10.14291/tccon.ggg2014.pasadena01.R1/1182415, 2014.

Wunch, D., Taylor, J. R., Fu, D., Bernath, P., Drummond, J. R., Midwinter, C., Strong, K., and Walker, K. A.: Simultaneous ground-based observations of O$_3$, HCl, N$_2$O, and CH$_4$ over Toronto, Canada by three Fourier transform spectrometers with different resolutions, Atmos. Chem. Phys., 7, 1275–1292, doi:10.5194/acp-7-1275-2007, 2007.

Wunch, D., Wennberg, P. O., Toon, G. C., Keppel-Aleks, G., and Yavin, Y. G.: Emissions of greenhouse gases from a North American megacity, Geophys. Res. Lett., 36, L15810, doi:10.1029/2009GL039825, 2009.

Wunch, D., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Stephens, B. B., Fischer, M. L., Uchino, O., Abshire, J. B., Bernath, P., Biraud, S. C., Blavier, J.-F. L., Boone, C., Bowman, K. P., Browell, E. V., Campos, T., Connor, B. J., Daube, B. C., Deutscher, N. M., Diao, M., Elkins, J. W., Gerbig, C., Gottlieb, E., Griffith, D. W. T., Hurst, D. F., Jiménez, R., Keppel-Aleks, G., Kort, E. A., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C. M., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M. A.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, Atmos. Meas. Tech., 3, 1351–1362, doi:10.5194/amt-3-1351-2010, 2010.

Wunch, D., Wennberg, P. O., Toon, G. C., Connor, B. J., Fisher, B., Osterman, G. B., Frankenberg, C., Mandrake, L., O'Dell, C., Ahonen, P., Biraud, S. C., Castano, R., Cressie, N., Crisp, D., Deutscher, N. M., Eldering, A., Fisher, M. L., Griffith, D. W. T., Gunson, M., Heikkinen, P., Keppel-Aleks, G., Kyrö, E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt,

J., Miller, C. E., Morino, I., Notholt, J., Oyafuso, F. A., Ret-
tinger, M., Robinson, J., Roehl, C. M., Salawitch, R. J., Sher-
lock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D. R.,
Uchino, O., Warneke, T., and Wofsy, S. C.: A method for eval-
uating bias in global measurements of CO$_2$ total columns from
space, Atmos. Chem. Phys., 11, 12317–12337, doi:10.5194/acp-
11-12317-2011, 2011a.

Wunch, D., Toon, G. C., Blavier, J.-F. L., Washenfelder, R.
A., Notholt, J., Connor, B. J., Griffith, D. W. T., Sherlock,
V., and Wennberg, P. O.: The Total Carbon Column Ob-
serving Network, Philos. Trans. R. Soc. A, 369, 2087–2112,
doi:10.1098/rsta.2010.0240, 2011b.

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu,
C., Feist, D. G., and Wennberg, P. O.: The Total Carbon
Column Observing Network's GGG2014 Data Version, 43,
doi:10.14291/tccon.ggg2014.documentation.R0/1221662, 2015.

Wunch, D., Toon, G. C., Hedelius, J., Vizenor, N., Roehl, C. M.,
Saad, K. M., Blavier, J.-F. L., Blake, D. R., and Wennberg,
P. O.: Quantifying the Loss of Processed Natural Gas Within
California's South Coast Air Basin Using Long-term Measure-
ments of Ethane and Methane, Atmos. Chem. Phys. Discuss.,
doi:10.5194/acp-2016-359, in review, 2016.