Atmospheric
Measurement
Techniques

Building the COllaborative Carbon Column Observing Network (COCCON): long-term stability and ensemble performance of the EM27/SUN Fourier transform spectrometer

Matthias Frey[1], Mahesh K. Sha[2,a], Frank Hase[1], Matthäus Kiel[3,a], Thomas Blumenstock[1], Roland Harig[4],
Gregor Surawicz[4], Nicholas M. Deutscher[5], Kei Shiomi[6], Jonathan E. Franklin[7], Hartmut Bösch[8,9], Jia Chen[10],
Michel Grutter[11], Hirofumi Ohyama[12], Youwen Sun[13], André Butz[14,b], Gizaw Mengistu Tsidu[15], Dragos Ene[16],
Debra Wunch[17], Zhensong Cao[13], Omaira Garcia[18], Michel Ramonet[19], Felix Vogel[20], and Johannes Orphal[1]

[1]Institute of Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[2]Royal Belgian Institute for Space Aeronomy, Brussels, Belgium
[3]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA
[4]Bruker Optics GmbH, Ettlingen, Germany
[5]Centre for Atmospheric Chemistry, School of Earth, Atmosphere and Life Sciences, Faculty of Science, Medicine and
Health, University of Wollongong, Wollongong, NSW, Australia
[6]Japan Aerospace Exploration Agency, Tsukuba, Japan
[7]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA
[8]Department of Physics and Astronomy, University of Leicester, Leicester, UK
[9]National Centre for Earth Observation (NCEO), University of Leicester, Leicester, UK
[10]Environmental Sensing and Modeling, Technische Universität München, Munich, Germany
[11]Centro de Ciencias de la Atmósfera, Universidad National Autónoma de México, Mexico City, Mexico
[12]Center for Global Environmental Research, National Institute for Environmental Studies, Tsukuba, Japan
[13]Anhui Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Hefei, China
[14]Institut für Umweltphysik, Universität Heidelberg, Germany
[15]Department of Earth and Environmental Sciences, Botswana International University of Science and Technology,
Gaborone, Botswana
[16]National Institute for Research and Development in Optoelectronics (INOE), Magurele, Romania
[17]Department of Physics, University of Toronto, Toronto, Canada
[18]Izaña Atmospheric Research Centre (IARC), Meteorological State Agency of Spain (AEMET), Tenerife, Spain
[19]Laboratoire des sciences du climat et de l'environnement, Gif-sur-Yvette, France
[20]Climate Research Division, Environment and Climate Change Canada, Toronto, Canada
[a]formerly at: Institute of Meteorology and Climate Research (IMK-ASF), Karlsruhe Institute of Technology (KIT),
Karlsruhe, Germany
[b]formerly at: Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt e.V.,
Oberpfaffenhofen, Germany

Correspondence: Matthias Frey (m.frey@kit.edu)

Abstract. In a 3.5-year long study, the long-term performance of a mobile, solar absorption Bruker EM27/SUN spectrometer, used for greenhouse gas observations, is checked with respect to a co-located reference Bruker IFS 125HR spectrometer, which is part of the Total Carbon Column Observing Network (TCCON). We find that the EM27/SUN is stable on timescales of several years; the drift per year between the EM27/SUN and the official TCCON product is 0.02 ppmv for XCO_2 and 0.9 ppbv for XCH_4, which is within the 1σ precision of the comparison, 0.6 ppmv for XCO_2 and 4.3 ppbv for XCH_4. The bias between the two data sets is 3.9 ppmv for XCO_2 and 13.0 ppbv for XCH_4. In order to avoid sensitivity-dependent artifacts, the EM27/SUN is also compared to a truncated IFS 125HR data set derived from full-resolution TCCON interferograms. The drift is 0.02 ppmv for XCO_2 and 0.2 ppbv for XCH_4 per year, with 1σ precisions of 0.4 ppmv for XCO_2 and 1.4 ppbv for XCH_4, respectively. The bias between the two data sets is 0.6 ppmv for XCO_2 and 0.5 ppbv for XCH_4. With the presented long-term stability, the EM27/SUN qualifies as an useful supplement to the existing TCCON network in remote areas. To achieve consistent performance, such an extension requires careful testing of any spectrometers involved by application of common quality assurance measures. One major aim of the COllaborative Carbon Column Observing Network (COCCON) infrastructure is to provide these services to all EM27/SUN operators. In the framework of COCCON development, the performance of an ensemble of 30 EM27/SUN spectrometers was tested and found to be very uniform, enhanced by the centralized inspection performed at the Karlsruhe Institute of Technology prior to deployment. Taking into account measured instrumental line shape parameters for each spectrometer, the resulting average bias across the ensemble with respect to the reference EM27/SUN used in the long-term study in XCO_2 is 0.20 ppmv, while it is 0.8 ppbv for XCH_4. The average standard deviation of the ensemble is 0.13 ppmv for XCO_2 and 0.6 ppbv for XCH_4. In addition to the robust metric based on absolute differences, we calculate the standard deviation among the empirical calibration factors. The resulting 2σ uncertainty is 0.6 ppmv for XCO_2 and 2.2 ppbv for XCH_4. As indicated by the executed long-term study on one device presented here, the remaining empirical calibration factor deduced for each individual instrument can be assumed constant over time. Therefore the application of these empirical factors is expected to further improve the EM27/SUN network conformity beyond the scatter among the empirical calibration factors reported above.

1 Introduction

Precise measurements of atmospheric abundances of greenhouse gases (GHGs), especially carbon dioxide (CO_2) and methane (CH_4), are of utmost importance for the estimation of emission strengths and flux changes (Olsen and Randerson, 2004). Furthermore, these measurements offer the prospect of being usable for the evaluation of emission reductions as specified by international treaties, e.g., the Paris COP21 agreement (https://unfccc.int/resource/docs/2015/cop21/eng/l09r01.pdf, last access: 4 March 2019). The Total Carbon Column Observing Network (TCCON) (Wunch et al., 2011) measures total columns of CO_2 and CH_4 with reference quality. TCCON achieves a calibration accuracy with a 1σ error of 0.2 ppmv for XCO_2 and 2 ppbv for XCH_4 and a total uncertainty budget of below 1 ppmv for XCO_2 and below 5 ppbv for XCH_4, respectively (Wunch et al., 2010, 2015). However, the instruments used by this network are rather expensive and need large infrastructure to be set up and expert maintenance, which has to be performed on site. Therefore TCCON stations have sparse global coverage, especially in Africa, South America and large parts of Asia (Wunch et al., 2015). Current satellites like the Orbiting Carbon Observatory-2 (OCO-2) (Frankenberg et al., 2015) and the Greenhouse Gases Observing Satellite (GOSAT) (Morino et al., 2011) on the other hand offer global coverage. Nonetheless, they suffer from coarse temporal resolution (the repeat cycle of OCO-2 is 16 days), and in the case of GOSAT from sparse spatial sampling as well as limited precision of a single measurement. These limitations mostly inhibit a straightforward estimation of the emission strength of localized sources of CO_2 and CH_4 like cities, landfills, swamps or fracking and mining areas from satellite observations. Recently OCO-2 data were used for estimating the source strength of power plants (Nassar et al., 2017) and urban emissions (Ye et al., 2017). However, this can only be done for power plants and urban areas that lie directly under the OCO-2 overpass locations. TCCON stations are also the primary validation for OCO-2 (https://ocov2.jpl.nasa.gov/files/ocov2/OCO-2_SciValPlan_111005_ver1_0_revA_final_signed1.pdf; last access: 4 March 2019), and validating the satellite observations at different locations is critical for the validation effort (Wunch et al., 2017).

The previously described Bruker EM27/SUN portable FTIR spectrometer (Gisi et al., 2011; Frey et al., 2015; Hedelius et al., 2016) is a promising instrument to overcome the above-mentioned shortcomings as it is a mobile, reliable, easy-to-deploy and low-cost supplement to the Bruker IFS 125HR spectrometer used in the TCCON network. So far the EM27/SUN was mainly used in campaigns for the quantification of local sinks and sources (Hase et al., 2015; Chen et al., 2016). In this work the long-term performance of the EM27/SUN with respect to a reference high-resolution TCCON instrument is investigated. Additionally, the ensemble

performance of several EM27/SUN spectrometers is tested. During 2014–2018, 30 EM27/SUN were tested at the Karlsruhe Institute of Technology (KIT) before being shipped to the customers. Several instruments that were distributed before this calibration routine at KIT was established were upgraded with a second channel for CO observations at Bruker Optics™ and after this also checked at KIT. This results in a unique data set as all EM27/SUN are directly compared to a reference EM27/SUN, continuously operated at KIT, as well as a co-located TCCON instrument. From this data set an EM27/SUN network precision and accuracy can be estimated.

The COllaborative Carbon Column Observing Network (COCCON) is intended to be a lasting framework for creating and maintaining a greenhouse gas-observing network based on common instrumental standards and data analysis procedures. Currently, about 18 working groups operating EM27/SUN spectrometers are contributing. We expect that COCCON will become an important supplement of TCCON, as the logistic requirements are low and the spectrometers are easy to operate. It will increase the global density of column-averaged greenhouse gas observations and, due to the fact that the spectrometers are portable, will especially contribute to the quantification of local sources.

2 Methodology

2.1 TCCON data set

As part of the TCCON, the Karlsruhe Institute of Technology (KIT) operates a high-resolution ground-based spectrometer at KIT, Campus North (CN) near Karlsruhe (49.100° N, 8.439° E, 112 m a.s.l.). Standard TCCON instruments have been described in great detail elsewhere (Washenfelder et al., 2006; Wunch et al., 2011). The Karlsruhe instrument, in the following called HR125, is the first demonstration of synchronized recordings of TCCON near-infrared (NIR) and NDACC mid-infrared (MIR) spectra using a dedicated dichroic beamsplitter (BS) arrangement (Optics Balzers Jena GmbH, Germany) with a cut-off wavenumber of 5250 cm^{-1}. It uses an InGaAs (indium–gallium–arsenide) detector in conjunction with an InSb (indium–antimonide) detector; details can be found in Kiel et al. (2016b). By the TCCON measurements, the relevant wavenumber region 4000–11000 cm^{-1}, corresponding to wavelengths λ between 0.9 and 2.5 μm, is covered so that, among other species, O_2, CO_2, CH_4, CO and H_2O can be retrieved. A figure showing the spectral range of TCCON and the EM27/SUN can be found in Hedelius et al. (2016), Fig. 1. The TCCON measurements were chosen as reference measurements because these gases are also measured by the EM27/SUN spectrometer. For TCCON measurements in the NIR the HR125 records single-sided interferograms with a resolution of 0.014 cm^{-1} ($\Delta\lambda = 3.5$ pm) or 0.0075 cm^{-1} ($\Delta\lambda = 1.9$ pm), correspond-

ing to a maximum optical path difference (MOPD) of 64 and 120 cm. The recording time for a typical measurement consisting of two forward and two backward scans is 212 and 388 s, respectively. The applied scanner velocity is 20 kHz. TCCON site Karlsruhe participated in the Infrastructure for Measurement of the European Carbon Cycle (IMECC) aircraft campaign (Messerschmidt et al., 2011; Geibel et al., 2012). The spectrometer has been used for calibrating all gas cells used by TCCON for instrumental line shape (ILS) monitoring (Hase et al., 2013).

TCCON data processing is performed using the GGG Suite software package (Wunch et al., 2011). In this study, the current release version, GGG 2014, is used (Wunch et al., 2015). The software package includes a pre-processor correcting for solar brightness fluctuations (Keppel-Aleks et al., 2007) and performing a fast Fourier transform including a phase error correction routine to convert recorded interferograms into solar absorption spectra. Note that forward and backward scans are split by the preprocessing software and analyzed separately. The central part of the software package is nonlinear least-squares retrieval algorithm GFIT. It performs a scaling retrieval with respect to an a priori profile, and then integrates the scaled profile over height to calculate the total column of the gas of interest. The software package additionally uses meteorological data from the National Center for Environmental Protection and National Center for Atmospheric Research (NCEP/NCAR) (Kalnay et al., 1996) and provides daily a priori gas profiles. TCCON converts the retrieved total column abundances VC$_{gas}$ of the measured gases into column-averaged dry air mole fractions (DMFs), where the DMF of a gas is denoted as $X_{gas} = \frac{VC_{gas}}{VC_{O_2}} \times 0.2095$. In this representation several errors cancel out that affect both the target gas and O_2. However, residual bias with respect to in situ measurements still persists, as well as a residual spurious dependence of retrieval results on the apparent airmass. Therefore the GGG suite also includes a post-processing routine applying an empirical airmass-dependent correction factor (ADCF) and airmass-independent correction factor (AICF). The AICF is deduced from comparisons with in situ instrumentation on aircrafts (Wunch et al., 2010).

2.2 HR125 low-resolution data set

In addition to the afore-mentioned TCCON data product, a second data product from the HR125 will be used in this work, in the following called HR125 LR. For this product the raw interferograms are first truncated to the resolution of the EM27/SUN, 0.5 cm^{-1}. At 0.5 cm^{-1} resolution, the ILS of the HR125 is expected to be nearly nominal. However, to avoid any systematic bias of the HR125 LR data with respect to the EM27/SUN results, the same procedure for ILS determination from H_2O signatures in open path lab air spectra was applied and the resulting ILS parameters adopted for the trace gas analysis. The analysis procedure will be explained in detail in Sect. 2.3; the retrieval software used for this data set is

PROFFIT Version 9.6 (Hase et al., 2004). The reason for the construction of this HR125 LR data set is that with this approach the analysis for the two instruments can be performed in exactly the same way. The resolution is harmonized; the averaging kernels for a given airmass are nearly identical. Differences between the EM27/SUN and the HR125 LR data set can then be attributed to instrumental features alone and do not need to be disentangled from retrieval software, resolution and airmass dependency differences. Note that for the low-resolution data set, forward and backward scans are averaged and then analyzed, whereas they are analyzed separately for the TCCON data set. Therefore the number of coincident measurements with the EM27/SUN data set compared to the TCCON data set is lower.

2.3　EM27/SUN data set

The EM27/SUN spectrometer, which was developed by KIT in collaboration with Bruker Optics™, is utilized for the acquisition of solar spectra. The instrument has been described in great detail in Gisi et al. (2012); in the following a short overview is given. The central part of this Fourier transform spectrometer (FTS) is a RockSolid™ pendulum interferometer with two cube corner mirrors and a CaF_2 beamsplitter. The EM27/SUN routinely records double-sided interferograms; the compensated BS design minimizes the curvature in the phase spectrum. This setup achieves high stability against thermal influences and vibrations. The retroreflectors are gimbal-mounted, which results in frictionless and wear-free movement. In this aspect the EM27/SUN is more stable than the HR125 high-resolution FTS, which suffers from wear because of the use of friction bearings on the moving retroreflector. Over time this leads to shear misalignment and requires regular realignment (Hase, 2012). The gimbal-mounted retroreflectors move a geometrical distance of 0.45 cm, leading to an optical path difference of 1.8 cm which corresponds to a spectral resolution of $0.5 \, \mathrm{cm}^{-1}$.

In a first pre-processing step, a solar brightness fluctuation correction is performed similarly to Keppel-Aleks et al. (2007). Furthermore, the recorded interferograms are Fourier transformed using the Norton–Beer medium apodization function (Davis et al., 2010). This apodization is useful for reducing sidelobes around the spectral lines, an undesired feature in low-resolution spectra, which would complicate the further analysis. A quality control, which filters interferograms with intensity fluctuations above 10 % and intensities below 10 % of the maximal signal range, is also applied.

In this work, spectra were analyzed utilizing PROFFIT Version 9.6, a nonlinear least-squares spectral fitting algorithm, which gives the user the opportunity to provide the measured ILS as an input parameter, an option chosen for this study (Hase et al., 2004). This code is in wide use and has been thoroughly tested in the past for the HR125 as well as the EM27/SUN, e.g., Schneider and Hase (2009), Sepúlveda et al. (2012), Kiel et al. (2016a), and Chen et al.
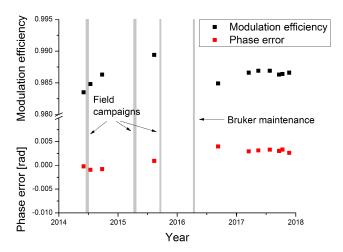
(2016). Due to the low resolution of the EM27/SUN, the atmospheric spectra were fitted by scaling of a priori trace gas profiles, although PROFFIT has the ability to perform a full profile retrieval (Dohe, 2013). As the source of the a priori profiles, the TCCON daily profiles introduced in Sect. 2.1 are utilized to be consistent with the TCCON analysis. Also for the daily temperature and pressure profiles, the approach from TCCON was adopted, using NCEP model data together with on-site ground pressure data from a meteorological tall tower (http://www.imk.kit.edu/messmast/; last access: 4 March 2019).

For the evaluation of the O_2 column the 7765–8005 cm^{-1} spectral region is used, which is also applied in the TCCON analysis (Wunch et al., 2010). For CO_2 we combine the two spectral windows used by TCCON into one larger window ranging from 6173 to 6390 cm^{-1}. CH_4 is evaluated in the 5897–6145 cm^{-1} spectral domain. For H_2O the 8353–8463 cm^{-1} region is used. This differs from TCCON, which deploys several narrow spectral windows, a strategy which is more in line with high-resolution spectral observations. For consistency reasons, and to reference the results to the WMO scale, the EM27/SUN retrieval also performs a postprocessing. The AICFs from TCCON are adopted, and similarly to Wunch et al. (2010), an airmass dependency correction is performed, although other numerical values for the correction parameters are used. Details can be found in Frey et al. (2015) and Klappenbach et al. (2015).

3　Long-term performance

3.1　ILS analysis

Accurate knowledge of the real ILS of a spectrometer is extremely important because errors in the ILS lead to systematic errors in the trace gas retrieval. For this reason regular ILS measurements were performed from the beginning of this study 4 years ago to detect possible misalignments and alignment drifts. The source of a de-adjustment is mostly mechanical shock, due to, e.g., impacts or vibrations especially due to transportation of the instruments. For the analysis of the measured data, version 14.5 of retrieval software LINEFIT (Hase et al., 1999; Hase, 2012) is used. Due to the fact that the EM27/SUN is equipped with a circular field stop aperture, the ILS is nearly nominal. Therefore, to keep the treatment concise, we use the simple two-parameter ILS model offered by LINEFIT. A detailed description of the ILS analysis is given in Frey et al. (2015). The time series of the ILS measurements is shown in Fig. 1; the modulation efficiency (ME) at maximum optical path difference (MOPD) ranges between 0.9835 and 0.9896, with a mean value of 0.9862 and a standard deviation of 0.0015. The phase error is close to zero for the whole time series, with a mean value of 0.0019 ± 0.0018. This modulation efficiency is significantly different from nominal, which is surprising, as great care was

Figure 1. ILS time series of the reference EM27/SUN. Results for modulation efficiency and phase error were obtained with LINE-FIT 14.5. The mean value of the modulation efficiency is 0.9862 with a standard deviation of 0.0015. For the phase error an average value of 0.0019 ± 0.0018 is observed. As can be seen from the closely spaced measurements in 2017, there is no seasonality in the ILS values. Grey areas denote periods of transportation of the instrument.

taken to align the instrument. Therefore open path measurements were also performed for the HR125 at a resolution of $0.5 \, \text{cm}^{-1}$ to investigate whether this method shows a bias. For this small optical path difference, the alignment of the HR125 should be very close to nominal. However, the LINEFIT analysis shows a ME of 0.9824 at MOPD. From this result it is concluded that this method shows an overall low bias of around 1.5 %–2 %, probably due to a slight underestimate of the pressure-broadening parameters of H_2O in the selected spectral region.

There is no overall trend apparent in the time series; the remaining differences in the modulation efficiency are probably due to the remaining uncertainty of the measurement technique. As is indicated by the more frequent measurements in 2017, there is also no seasonality in the results of the open path measurements. It should be noted that the measurement routine was refined in the course of this work. In particular, in the beginning (2014) it was assumed that the inside of the EM27/SUN is free of water vapor, so the instrument was not vented during the lamp measurements. However, sensitivity studies as presented in Frey et al. (2015) revealed that the influence of the water vapor column inside the spectrometer can not always be neglected. After this discovery the instrument was vented during the open path measurements. This is why the 2014 calculations show larger scatter, as here the amount of water vapor inside the spectrometer is not known. For this analysis it was assumed that also for the 2014 measurements the total pressure inside the spectrometer is the same as of the surrounding air, which is a sensible assumption as the spectrometer is not evacuated. This also

Table 1. Estimated ME uncertainties for various error sources.

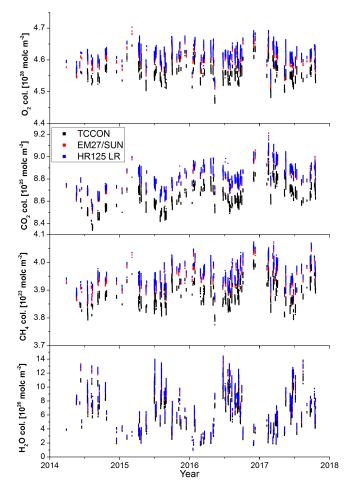| Error source	Uncertainty	Propagation on ME
Temperature	$\pm 0.8 \, \text{K}$	$\pm 0.16 \, \%$
Total pressure	$\pm 3 \, \text{mbar}$	$\pm 0.19 \, \%$
Distance	$\pm 5 \, \text{cm}$	$\pm 0.04 \, \%$
Partial pressure H_2O	$\pm 0.5 \, \text{mbar}$	$\pm 0.13 \, \%$
Measurement noise		$\pm 0.05 \, \%$
Total		$\pm 0.29 \, \%$

explains why the deviations become smaller in 2017. A further test to verify the stability of the instrument is the X_{air} parameter, which is the surface pressure divided by the measured column of air. This test will be shown in Sect. 3.3.

The grey lines in Fig. 1 denote transportation of the spectrometer over longer distances for field campaigns in Berlin (northeastern Germany), Oldenburg (northern Germany) and Paris (France) and for maintenance at Bruker Optics. Note that no realignment of the interferometer was performed during this maintenance. Only the reference HeNe laser was exchanged due to sampling instabilities during interferogram recordings. More specifically, the laser wavelength was unstable, resulting in a corruption of parts of the measured spectra. Later in 2016 and 2017 this instrument was not used for campaigns since it has been chosen as the reference EM27/SUN for comparison measurements next to the HR125 spectrometer in order to take measurements at Karlsruhe as continuously as possible. The instrument was not realigned during the whole comparison study.

An error estimation for the open path measurements is given in Table 1. For the temperature and pressure error, the stated accuracies of the data logger manufacturer were used. For the other potential error sources reasonable estimates were made. The total error, given by the root-squares sum of the individual errors, is 0.29 % in ME amplitude, consisting of several errors of approximately the same magnitude.

3.2 Total column time series

In this section the total column measurements from the EM27/SUN are compared to the reference HR125 spectrometer. For the measurements, the EM27/SUN was moved to a terrace on the top floor of the IMK-ASF, building 435 KIT CN (49.094° N, 8.436° E; 133 m a.s.l.) on a daily basis if weather conditions were favorable. The spectrometer was moved from the lab on the fourth floor to the roof terrace on the seventh floor, thus being exposed to mechanical stress. The instrument was coarsely oriented north, without effort for levelling. If further orientation was needed, the spectrometer was manually rotated so that the solar beam was centered onto the entrance window. The CamTracker program was then able to track the sun. The spectrometer was operated at ambient temperatures. During summer, the spectrometer

Figure 2. Total column time series for O_2, CO_2, CH_4 and H_2O measured at KIT in Karlsruhe from March 2014 until October 2017. The number of interferograms and recording time for the different data types are the following: TCCON: 2 IFGs, 114 s; EM27/SUN: 10 IFGs, 58 s; HR125 LR: 4 IFGs, 152 s. Only coincident measurement points (within 1 min) are depicted.

heated up to temperatures above 40 °C. In order to protect the electronics from the heat, a sun cover for the EM27/SUN was built, which reduced the temperatures inside the spectrometer by about 10 °C. In winter the temperatures were as low as −4 °C at the start of measurements. Double-sided interferograms with $0.5\,\mathrm{cm}^{-1}$ resolution were recorded. With 10 scans and a scanner velocity of 10 kHz, one measurement takes about 58 s. For precise time recording, a GPS receiver was used.

The full time series from March 2014 to November 2017 is shown in Fig. 2 for the three data sets. For better visibility only coincident data points measured within 1 min between EM27/SUN and the other data sets are shown. There are 8349 paired measurements between EM27/SUN and TCCON and 4624 between EM27/SUN and HR125 LR; in total there are 50 550 EM27/SUN and 25 361 TCCON measurements.

All gases show a pronounced seasonal cycle, where the variability in water vapor is strongest with values below 1×10^{26} molec. m^{-2} in winter and up to 14×10^{26} molec. m^{-2} in summer. Furthermore, the seasonal cycle of water vapor is shifted with respect to the other species. Another feature seen is that there is an offset in the EM27/SUN (red squares) and HR125 LR (blue squares) total column data with respect to the TCCON data (black squares). The occurrence of a systematic bias when reducing the spectral resolution has been observed by several investigators (Petri et al., 2012; Gisi et al., 2012). The observed offset between EM27/SUN and HR125 LR measurements is smaller. The remaining difference can be attributed to the different measurement heights of the HR125 (112 m) and EM27/SUN (133 m). For a quantitative analysis we do not utilize the total column measurements, but rather use the X_{Gas}, as in this representation systematic errors, e.g., ILS errors, timing errors, tracking errors and nonlinearities, mostly cancel out. Furthermore, the height dependence largely cancels out in this representation. The comparison will be presented in the following sections.

First, a sensitivity study is provided demonstrating the effect of changes in the ILS on the gas retrieval. For this 1 h of measurements around solar noon on 1 August 2016 and 15 February 2017, corresponding to solar elevation angles (SEAs) of 60 and 30°, were analyzed with artificially altered ILS values. The results are shown in Table 2. An increase of 1 % in the modulation efficiency leads to a decrease of 0.35 % (0.37 %) in the retrieved O_2 column, 0.31 % (0.31 %) in H_2O, 0.26 % (0.28 %) in CH_4 and 0.50 % (0.57 %) in CO_2 for an SEA of 60° (30°). So the change in the retrieved total column is not alike, but a unique characteristic of each species, and also slightly airmass-dependent. As the decrease in the CO_2 column is larger than the decrease in the O_2 column, XCO_2 decreases with an increasing ME, 0.16 % (0.19 %) for 1 % ILS increase, whereas XCH_4 increases 0.10 % (0.09 %). This is opposed to prior studies (Gisi et al., 2012; Hedelius et al., 2016) reporting an increase in XCO_2 and decrease in XCH_4 for an increase in the modulation efficiency, albeit in agreement with the findings from Hase et al. (2013) for the HR125 spectrometer, reporting that a change in the modulation efficiency results in a larger relative decrease in the CO_2 column than in the O_2 column.

3.3 X_{air}

In this section the column-averaged amount of dry air (X_{air}) is investigated. This quantity is a sensitive test of the stability of a spectrometer because for X_{air} there is no compensation of possible instrumental problems, in contrast to the DMFs, where errors can partially cancel out. X_{air} compares the measured oxygen column (VC_{O_2}) with surface pressure measurements (P_{S}):

$$X_{\mathrm{air}} = \frac{0.2095}{\mathrm{VC}_{O_2} \cdot \overline{\mu}} \cdot \left(\frac{P_{\mathrm{S}}}{g} - \mathrm{VC}_{H_2O} \cdot \mu_{H_2O} \right), \tag{1}$$

Table 2. Sensitivity study on the effect of ILS changes on the retrieval of the total gas columns. Depicted are hourly pooled data on 1 August 2016 and 15 February 2017 around solar noon, corresponding to solar elevation angles of 60 and 30°. The resulting ILS dependency of XCO_2 is -0.16% and -0.19% for 60 and 30° SEA, for a 1% ME increase. XCH_4 increases by 0.10% (0.09%).

ME	O_2 (10^{28} molec. m^{-2})	H_2O (10^{26} molec. m^{-2})	CH_4 (10^{23} molec. m^{-2})	CO_2 (10^{25} molec. m^{-2})
August 2016				
0.99	4.6097	7.4551	3.9457	8.7321
1.00	4.5936	7.4323	3.9356	8.6879
February 2017				
0.99	4.6718	3.7746	4.0261	9.0968
1.00	4.6545	3.7628	4.0148	9.0455

Here $\overline{\mu}$ and μ_{H_2O} denote the molecular masses of dry air and water vapor, respectively, g is the column-averaged gravitational acceleration and VC_{H_2O} is the total column of water vapor. The correction with VC_{H_2O} is necessary as the surface pressure instruments measure the pressure of the total air column, including water vapor. For an ideal measurement and retrieval with accurate O_2 and H_2O spectroscopy, as well as accurate surface pressure, X_{air} would be 1. However, due to insufficiencies in the oxygen spectroscopy, this value is not obtained. For TCCON measurements X_{air} is typically ~ 0.98 (Wunch et al., 2015). For the EM27/SUN prior studies showed a factor of ~ 0.97 (Frey et al., 2015; Hase et al., 2015; Klappenbach et al., 2015). Large deviations ($\sim 1\%$) from these values indicate severe problems, e.g., errors with the surface pressure, pointing errors, timing errors or changes in the optical alignment of the instrument. As mentioned in Sect. 3.1, here X_{air} is used to check whether the small changes in the modulation efficiency indicated by the open path measurements are due to actual alterations in the alignment of the EM27/SUN or due to the residual uncertainty of the calibration method.
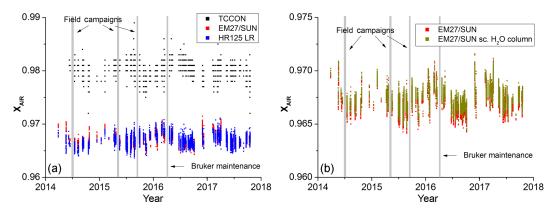
Panel (a) of Fig. 3 shows the X_{air} time series of TCCON, the EM27/SUN and HR125 LR. For clarity, only coincident data points that were measured within 1 min between the different data sets are shown. Grey areas denote periods where the EM27/SUN was moved over long distances for campaigns or maintenance. The absolute values of X_{air} differ for the data sets, with 0.9805 ± 0.0012 for TCCON, 0.9669 ± 0.0010 for the EM27/SUN and 0.9670 ± 0.0011 for HR125 LR. The difference between the EM27/SUN and the HR125 LR is within 1σ precision. The difference between the EM27/SUN and TCCON data set, which is commonly observed as previously noted, is a consequence of the different resolution together with the different retrieval algorithm (Gisi et al., 2012). It can be seen that all data sets exhibit a seasonal variability, which is more prominent in the TCCON data, as can also be seen from the higher standard deviation. From this higher variability it can be concluded that the airmass dependency in the official TCCON O_2 retrieval is

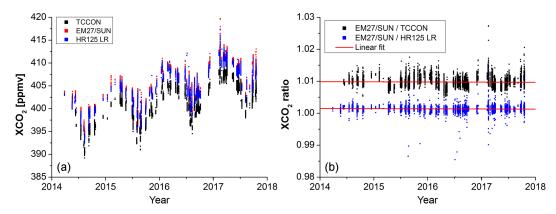**Table 3.** XCO_2 biases between the EM27/SUN and HR125 data sets.

XCO_2 ratio	No. of coincidences	Mean (1σ)	Yearly trend in the ratio
EM27/TCCON	8349	1.0098 (0.0015)	-5×10^{-5}
EM27/HR125 LR	4624	1.0014 (0.0011)	-5×10^{-5}

higher than for the PROFFIT retrieval on reduced-resolution TCCON measurements, a finding also observed by Gisi et al. (2012) between the TCCON retrieval and the PROFFIT retrieval at full resolution. For the PROFFIT retrieval, it is suspected that part of the variability stems from insufficiencies in the utilized HITRAN 2008 H_2O linelist. It was reported by Tallis et al. (2011) that in the 8000–9200 cm^{-1} region, line intensities are low by up to 20% compared to other wavenumber regions. This in return will lead to a systematic overestimation of the water column, which also affects X_{air}. To test the sensitivity of X_{air} with respect to the measured H_2O column, in panel (b) of Fig. 3 the original EM27/SUN time series is compared to a data set where the H_2O column is artificially reduced by 20%. This approach is further justified by a study from the Romanian National Institute for Research and Development in Optoelectronics (INOE) conducted in 2017, where we compared total column amounts of water vapor from an EM27/SUN and a radiometer. We found that the EM27/SUN values were systematically higher by 20%. And indeed, the standard deviation, which is here used as a measure of the seasonal variability, of the modified time series (0.0009) is lower when compared to the original time series (0.0010).

There are no obvious steps and there is no significant drift between the EM27/SUN and the HR125 LR data sets, so that it can be concluded that the EM27/SUN is stable during the complete course of the over 3-year long comparison, and differences seen in the modulation efficiency are introduced by the remaining uncertainty in the calibration method.

Figure 3. (a) shows the X_{air} time series measured at KIT in Karlsruhe for the TCCON, EM27/SUN and HR125 LR data sets. For clarity, only coincident measurements (within 1 min) of the data sets are plotted. Grey areas denote periods where the EM27/SUN was moved over long distances. **(b)** shows a comparison of the original EM27/SUN time series with a modified version, where a scaling factor of 0.8 was applied to the H_2O total column.
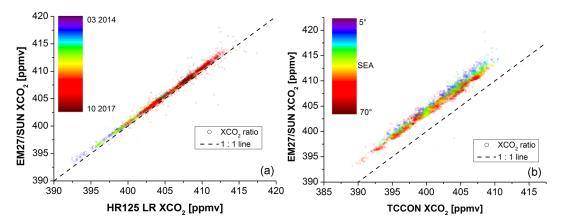

Figure 4. (a) shows the XCO_2 time series measured at KIT in Karlsruhe for the three data sets from March 2014 to October 2017. For clarity, only coincident measurements (within 1 min) of the data sets are plotted. **(b)** shows the XCO_2 ratio between the EM27/SUN and the two HR125 data sets. A linear fit was applied to investigate a possible trend in the ratios.
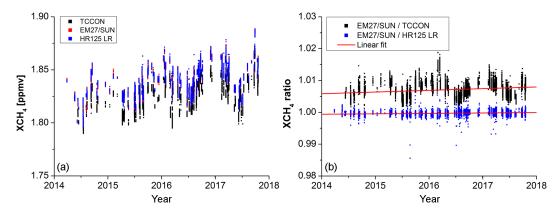
3.4 XCO$_2$

In Fig. 4 XCO_2 time series of the three data sets are shown together with the offsets between the data sets. The general characteristics of the data sets are similar. The yearly increase in XCO_2 due to anthropogenic emissions of about 2 ppmv can be seen as well as the seasonal cycle with a decrease in XCO_2 of approximately 10 ppmv during summer due to photosynthesis, characteristic of mid-latitude stations. Despite these agreements in the general trend, there are also differences between the data sets. Relative to the TCCON data the EM27/SUN and the HR125 LR data sets are biased high (0.98 % and 0.84 %, respectively). The scaling factors are calculated by taking the mean of all individual coincident point ratios (EM27/SUN/TCCON and EM27/SUN/HR125 LR). Together with these ratios a standard deviation is also derived; see Table 3. A high bias was also observed by Gisi et al. (2012); Frey et al. (2015), albeit with smaller absolute differences. This is due to the fact that (1) in the Gisi et al. pa-

per the TCCON data were retrieved with an earlier version of GFIT (GGG2012) and (2) after the publication of the Frey et al. paper the Karlsruhe TCCON data were reprocessed with a customized GFIT retrieval accounting for baseline variations (Kiel et al., 2016b). The offset between EM27/SUN and TC-CON shows a seasonal variability. The reasons for this are mainly the differences in airmass correction, averaging kernels and retrieval algorithm. These effects have been investigated before (Gisi et al., 2012; Frey et al., 2015; Klappenbach et al., 2015; Hedelius et al., 2017; Kiel et al., 2016a). The averaging kernels of the EM27/SUN have been previously presented and compared to TCCON in a study by Hedelius et al. (2016).

It has to be noted that the level of uncertainty for XCO_2 is significantly higher between COCCON and TCCON compared to the internal EM27/SUN consistency. According to Table 3, a current calibration uncertainty with respect to TC-CON of 0.6 ppmv is estimated.

Figure 5. **(a)** shows the XCO$_2$ comparison between EM27/SUN and HR125 LR. The colorbar denotes the date of the measurement; the dashed line is the 1 : 1 line. In **(b)** the comparison with TCCON is shown. Note that here the colorbar shows the solar elevation angle.

Figure 6. **(a)** shows the XCH$_4$ time series measured at KIT in Karlsruhe for the three data sets from March 2014 to October 2017. For clarity, only coincident measurements (within 1 min) of the data sets are plotted. **(b)** shows the XCH$_4$ ratio between the EM27/SUN and the two HR125 data sets. A linear fit was applied to investigate a possible trend in the ratios.

Table 4. XCH$_4$ biases between the EM27/SUN and HR125 data sets.

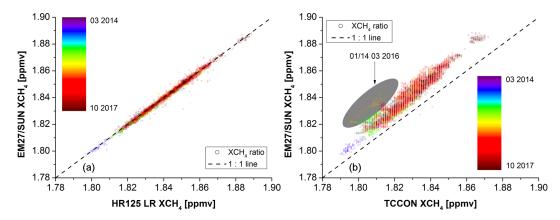| XCH$_4$ ratio	No. of coincidences	Mean (1σ)	Yearly trend in the ratio
EM27/TCCON	8349	1.0072 (0.0024)	0.0005
EM27/HR125 LR	4624	0.9997 (0.0008)	0.0001

For the long-term stability of the EM27/SUN the focus lies on the comparison with the HR125 LR data set, where the above-mentioned differences cancel out. There is a small offset between the two data sets, resulting in a calibration factor of 1.0014, which is constant over time in the analyzed time period. To test this assumption a linear fit was applied to the XCO$_2$ ratios; see panel (b) of Fig. 4. In Table 3 the slope coefficient is depicted. For both comparisons the yearly trend in the ratio is well within the 1σ precision (0.44 ppmv) of the data set. In absolute numbers the slope per year is $\approx$

-0.02 ppmv for both ratios, or a drift smaller than 0.1 ppmv over the whole comparison period of around 3.5 years.
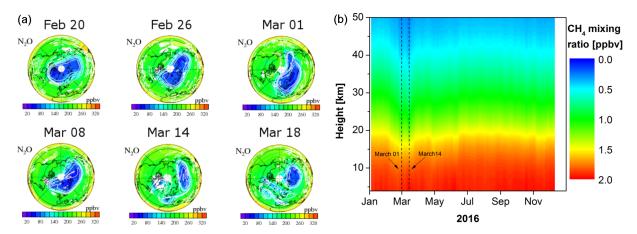
Figure 5 shows the data sets in a different representation. In panel (a) the EM27/SUN is compared to the HR125 LR; the colorbar indicates the date of measurement and the dashed line is the 1 : 1 line. It can be seen that there is no trend in the data apart from the overall increase in time due to anthropogenic emissions. In panel (b) the EM27/SUN is compared to the TCCON data set; the colorbar shows the SEA. This representation is chosen so that the remaining air-mass dependency of the ratio can be seen. It is also interesting to note that omitting the TCCON AICF for our analysis would move the data set significantly closer to the 1 : 1 line. The scaling factor would change from 1.0098 to 0.9995. As this finding is not true for XCH$_4$ and is probably coincidental, we maintain the AICF.

3.5 XCH$_4$

Figure 6 shows the XCH$_4$ time series of the different data sets. As for XCO$_2$, the general features are in agreement

Figure 7. (a) shows the XCH$_4$ comparison between EM27/SUN and HR125 LR. The colorbar denotes the date of the measurement; the dashed line is the 1 : 1 line. In **(b)** the comparison with TCCON is shown. The shaded area encloses measurements from 1 and 14 March 2016. For these days the ratio is significantly different with respect to the remaining data set (see text for discussion).

Figure 8. In **(a)** N$_2$O MLS data from the Aura satellite are shown as a tracer for the position of the polar vortex for several days in February and March 2016. Data and plots courtesy of the NASA science team (https://mls.jpl.nasa.gov/, last access: 4 March 2019). **(b)** shows CH$_4$ mixing ratios from NDACC FTIR station Jungfraujoch in Switzerland, downloaded from the NDACC archive (http://www.ndaccdemo.org/stations/jungfraujoch-switzerland/, last access: 4 March 2019). For dates with no measurements the data have been interpolated using a weighted average. Dotted lines depict 1 and 14 March 2016. For these dates, the XCH$_4$ data significantly differ from the remaining data set.

for all data sets. There is a slight annual increase of about 10 ppbv. Also, there is a seasonal cycle with a variability of ≈ 30 ppbv; however, compared to XCO$_2$ the interannual seasonality strength and phase vary significantly between the years due to the many different variable sinks and sources of methane, e.g., Dlugokencky et al. (1997). The differences between the data sets largely resemble the differences observed for XCO$_2$. The bias between EM27/SUN and TCCON is 0.72 %; see Table 4. This bias is close to the bias observed by Hedelius et al. (2016), 0.75 %, where they used the GGG software package for the analysis of EM27/SUN spectra. Although a single bias is reported, as was observed for XCO$_2$ the offset is not constant, but rather shows a seasonality. The calibration uncertainty between COCCON and TCCON is estimated to amount to 5 ppbv for XCH$_4$; see Table 4. The retrievals between EM27/SUN and HR125 LR agree within

1σ precision (0.9997 ± 0.0008). Panel (a) of Fig. 7 shows the ratio between EM27/SUN and HR125 LR color-coded with the observation date. As for XCO$_2$, no trend is apparent. An explicit linear fit to the XCH$_4$ ratio produces a slope coefficient of 0.0001, 1 order of magnitude smaller than the 1σ precision of the ratio (0.0008).

An interesting feature is observed in the ratio between EM27/SUN and TCCON data sets; see panel (b) of Fig. 7. In general the pattern is similar to that of XCO$_2$, with a slight dependence on the SEA. The ratio in the figure is color-coded with the date of observation rather than the SEA. It can be seen that for 1 and 14 March 2016 (shaded area in Fig. 7), the XCH$_4$ ratio significantly differs from the other observations. Previous work by Ostler et al. (2014) has shown that stratospheric intrusion, caused for example by the subsidence of the polar vortex, has a different effect on MIR

and NIR retrievals, even when using the same a priori profile. This is due to the differing sensitivity of the retrievals with respect to altitude. Therefore, differences between the true atmospheric profile and the assumed a priori profiles on these days could cause the differences seen. This effect will also lead to larger differences between EM27/SUN and TCCON XCH_4 because of the different impact on the retrieved columns due to differing sensitivities. A spread of the polar vortex to mid-latitudes could lead to significantly altered CH_4 profiles compared to the a priori profiles, explaining the observed differences in the XCH_4 ratio.

Figure 8a shows N_2O data from the Microwave Limb Sounder (MLS) on the Aura satellite for several days in February and March 2016 on the 490 K potential temperature level, corresponding to a height of approximately 18 km. N_2O is chosen because it serves as a tracer for the position of the polar vortex. Indeed, it seems that beginning in March 2016 the polar vortex stretches out to mid-latitudes. To further test this hypothesis, in Fig. 8b independent NDACC CH_4 profiles from the Jungfraujoch station in 2016 are shown. The station is situated approximately 270 km south of Karlsruhe with a station height of 3580 m. For dates without measurements, the data were interpolated using a weighted average. The dotted black lines denote 1 and 14 March 2016, the dates on which the XCH_4 ratio between EM27/SUN and TCCON shows an anomaly. The changed profile shape during that period is clearly visible. As this station is south of Karlsruhe, it is expected that also for Karlsruhe the CH_4 profile will show considerable downwelling, explaining the observed anomaly in the XCH_4 ratio.

4 Ensemble performance

Having investigated the long-term stability of the EM27/SUN with respect to a reference spectrometer in the previous section, here the level of agreement of an ensemble of EM27/SUN spectrometers is presented. The procedure is the same as for the comparison between the reference EM27/SUN and the HR125. First, the ILS is analyzed, followed by calibration factors for XCO_2 and XCH_4.

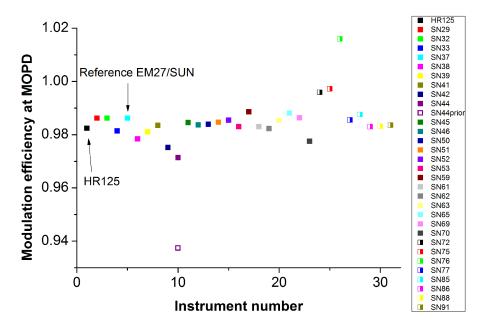### 4.1 ILS measurements and instrumental examination

The measurement of the ILS is a valuable diagnostic for detecting misalignments of spectrometers. Differences in the ILS of the EM27/SUN spectrometers due to misalignment can lead to biases in the data products between the instruments. Here the spread of ILS values of all EM27/SUN spectrometers that were checked at KIT in the past 4 years is estimated. Numerical values are given in Table 5; the results are shown in Fig. 9. The black square denotes an ILS measurement of the HR125 spectrometer, also with 1.8 cm MOPD. This test was done to check for an absolute offset of

Table 5. Summary of the modulation efficiencies at MOPD and phase errors for all EM27/SUN calibrated in Karlsruhe. "ref" denotes the reference EM27/SUN and "prior" denotes an ILS measurement with instrument SN44 prior to calibration at KIT.

Instrument SN	ME at MOPD	Phase error (rad)
29	0.9862	0.0014
32	0.9862	0.0034
33	0.9814	−0.0017
37 (ref)	0.9862	0.0019
38	0.9784	0.0009
39	0.9811	−0.0005
41	0.9835	0.0001
42	0.9752	0.0039
44	0.9714	−0.0019
44 (prior)	0.9374	−0.0074
45	0.9845	0.0034
46	0.9837	0.0024
50	0.9839	0.0023
51	0.9847	0.0017
52	0.9854	0.0048
53	0.9830	0.0025
59	0.9886	0.0029
61	0.9830	0.0013
62	0.9823	0.0053
63	0.9853	0.0011
65	0.9881	0.0024
69	0.9863	0.0030
70	0.9775	0.0056
72	0.9959	0.0030
75	0.9972	0.0041
76	1.0160	0.0007
77	0.9855	0.0016
85	0.9876	0.0025
86	0.9830	0.0031
88	0.9832	0.0007
91	0.9836	0.0021

our method. The HR125 would be expected to show an ideal ILS for short optical path differences, but a value of 0.9824 was obtained. From this measurement it is concluded that our method shows an absolute offset and that values between 0.98 and 0.99 are desired.

In general, the agreement between the 30 tested EM27/SUN is good, with an ensemble mean of 0.9851 ± 0.0078, which does not differ significantly from the value obtained for the HR125, but there are exceptions. Instrument SN 44 was checked at KIT only after an upgrade with the second channel was performed at Bruker Optics. Before realignment, the instrument showed a very low ME value of 0.9374. A realignment of the instrument enhanced the ME to 0.9714. This is still significantly low compared to the EM27/SUN ensemble mean, but the difference was drastically reduced. The second instrument showing strong deviations from the ensemble mean is SN76 with an ILS of 1.0160, the only instrument showing overmodulation. The ILS was even higher

Figure 9. Modulation efficiencies at MOPD for all EM27/SUN tested in Karlsruhe. For SN44 prior, ILS measurements were taken before an alignment check and subsequent realignment of the instrument. For comparison reasons, an ILS measurement for the HR125 was also performed.

(1.0350) when the first ILS measurements were performed. Due to our findings, the manufacturer exchanged the beamsplitter, which reduced the overmodulation, but it partly remained. In the meantime it was recognized that the cause of the error was the manufacturer during assembly of the instrument forgetting to insert the foreseen spacer to achieve the correct detector position with respect to the beamsplitter. The beamsplitter is coated, and the coating is applied on both sides of the beamsplitter over half the surface area. If the optical axis of the detector element coincides with the transition region of the two coating areas, detrimental effects occur. For this reason the detector element needs to be raised with respect to the interferometer. This problem occurred for instrument SN 77, but there it was diagnosed and corrected by KIT (ILS before lifting: 1.0340; ILS after correction: 0.9855).

The above-mentioned problems show the benefit of the calibration routine at KIT. Imperfections from nonideal alignments were diagnosed and corrected. Also, other detrimental effects, e.g., double-passing, channeling, nonlinearity issues, solar tracker problems, inaccurate positioning of the second detector, or camera issues, were corrected or minimized for a number of instruments. Finally, we checked whether the linear interpolation method suppressing sampling ghosts was activated.

4.2 XCO₂ and XCH₄ comparison measurements

After checking the alignment and performing lamp measurements, side-by-side solar calibration measurements were performed on the terrace on top of the KIT-IMK office build-ing with each spectrometer with respect to the reference EM27/SUN and also a co-located HR125 spectrometer. Calibration measurements started in June 2014 and are ongoing, if new spectrometers arrive for testing. The aim is to have at least 1 day of comparison measurements so that the spectrometers can be scaled to TCCON via the reference EM27/SUN. TCCON is extensively compared to measurements on the WMO scale. Dates of the comparison measurements for the different spectrometers as well as number of coincident measurements are given in Table 6. On 21 January 2016, our reference spectrometer suffered from laser sampling errors after approximately 1 h of measurements. Therefore the number of coincident measurements for SN62 and 63 that were checked on this date are sparse. A typical calibration day is depicted in Fig. 10.

The calibration factors and standard deviations for all instruments with respect to the reference spectrometer are also depicted in Table 6. Calibration factors and standard deviations were obtained using the methods described in Sect. 3.4. The calibration factors are close to nominal for all species and instruments. For XCO_2 the ensemble mean is high compared to the reference EM27/SUN, with a mean calibration factor of 0.9993 and a standard deviation of 0.0007. In Fig. 11 histograms of the calibration factor distributions are depicted for XCO_2, XCH_4, and O_2, respectively. The histograms are not conspicuous.

Applying the mean calibration factor to all calculated calibration factors centers the data around the ensemble mean. As an estimate for the spread of the calibration factors $\frac{1}{n}\Sigma|X\text{Gas factor}-1|$, we arrive at an average bias between

Table 6. Calibration factors for XCO_2, XCH_4, and O_2 for all investigated instruments with respect to the reference EM27/SUN spectrometer (SN37) as well as calibration dates and number of coincident measurements. Values in brackets denote percent standard deviations.

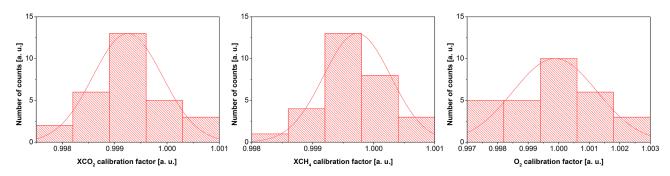| Instr. SN	Dates	No. co.	XCO_2 factor	XCH_4 factor	O_2 factor
29	140 606, 140 718	490	1.0004 (0.02)	0.9997 (0.03)	1.0008 (0.03)
32	150 414–150 422	1548	0.9997 (0.03)	0.9997 (0.03)	1.0004 (0.03)
33	170 807, 170 815	339	0.9991 (0.03)	0.9994 (0.04)	1.0009 (0.05)
38	150 410–150 421, 160 121	1609	0.9989 (0.03)	0.9997 (0.04)	0.9988 (0.04)
39	140 717, 150 414, 150 415	1210	0.9992 (0.04)	0.9994 (0.04)	1.0003 (0.04)
41	140 717, 150 414–150 422	1877	0.9999 (0.03)	1.0002 (0.03)	0.9991 (0.03)
42	160 730, 160 801	368	0.9978 (0.04)	1.0003 (0.04)	0.9975 (0.03)
44	170 227	286	0.9979 (0.03)	0.9984 (0.03)	0.9985 (0.03)
45	170 807, 170 815	382	0.9995 (0.03)	0.9991 (0.04)	1.0008 (0.02)
46	170 808, 170 815	503	0.9993 (0.03)	0.9994 (0.03)	1.0003 (0.03)
50	150 421, 150 422	699	0.9999 (0.03)	0.9995 (0.03)	0.9995 (0.03)
51	160 126, 160 129	256	0.9995 (0.03)	0.9993 (0.03)	1.0007 (0.05)
52	150 421, 150 422	727	0.9990 (0.04)	0.9998 (0.05)	1.0002 (0.05)
53	150 421, 150 422	729	0.9987 (0.03)	1.0001 (0.03)	0.9992 (0.04)
59	160 318	273	0.9998 (0.03)	0.9991 (0.03)	1.0019 (0.04)
61	151 002, 170 713	618	0.9993 (0.03)	0.9996 (0.04)	1.0000 (0.04)
62	160 121	18	0.9988 (0.04)	0.9990 (0.02)	1.0002 (0.02)
63	160 121	15	1.0003 (0.05)	1.0001 (0.05)	1.0002 (0.07)
65	160 511	234	1.0005 (0.04)	0.9998 (0.05)	1.0020 (0.03)
69	160 908, 170 713	636	0.9994 (0.03)	0.9993 (0.03)	1.0008 (0.03)
70	160 831, 160 906	522	0.9985 (0.02)	1.0005 (0.03)	0.9978 (0.03)
72	170 215, 170 216	433	0.9994 (0.05)	1.0001 (0.03)	0.9999 (0.04)
75	170 516, 170 517	852	0.9993 (0.03)	0.9991 (0.03)	1.0018 (0.05)
76	170 608	365	0.9991 (0.04)	0.9997 (0.04)	1.0026 (0.06)
77	170 927	389	0.9999 (0.03)	0.9997 (0.03)	1.0001 (0.04)
85	180 213, 180 214	371	0.9993 (0.03)	1.0003 (0.03)	0.9990 (0.03)
86	180 213, 180 214	464	0.9986 (0.03)	1.0002 (0.03)	0.9975 (0.05)
88	180 314	154	0.9990 (0.03)	1.0008 (0.03)	0.9982 (0.03)
91	180 228	148	0.9985 (0.03)	1.0008 (0.03)	0.9977 (0.04)

the instruments of 0.20 ppmv. From Table 6 we can also calculate an average standard deviation $\frac{1}{n}\Sigma|\sigma|$ of 0.13 ppmv. For XCH_4 the ensemble mean is closer to the reference EM27/SUN (0.9997 ± 0.0006) as compared to XCO_2. From this results an average bias of 0.8 ppbv. The average standard deviation is 0.6 ppbv. These values are comparable to results obtained in a study from Hedelius et al. (2017). They checked the intercomparability of the four United States TC-CON sites using an EM27/SUN as a traveling standard. They report average biases of 0.11 ppmv for XCO_2 and 1.2 ppbv for XCH_4; for the average standard deviations they obtain 0.34 ppmv (XCO_2) and 1.8 ppbv (XCH_4). It has to be noted that for the Hedelius et al. (2017) study only data within ± 2 h of local noon were taken into account, whereas here no constraints regarding the time of measurement were applied. As another sensitive test the O_2 total column calibration factors are given. In contrast to XCO_2 and XCH_4, there is no canceling of errors in this quantity. The ensemble mean is slightly high compared to the reference EM27/SUN (0.9999 ± 0.0014). The average bias is 0.11 % O_2 with an average standard deviation of 0.04 % O_2.

Note that for our setup this average bias is a worst case scenario. The bias only applies if no calibration factor is used in the subsequent analysis. The strength of this calibration routine is that the computed calibration factors can be used, thereby significantly lowering the bias between different EM27/SUN spectrometers. The remaining bias is then given by the long-term drift of the individual instrument (see Sect. 3.4 and 3.5) and sudden alignment drifts due to mechanical strain from, e.g., transport and campaign use. To estimate this drift, we utilize the calibration factors before and after the Berlin campaign performed in 2014. There the drifts between five instruments were below 0.005 % XCO_2 and 0.035 % XCH_4 (Frey et al., 2015).

Ideally, we would expect identical calibration factors as we took the real ILS of the instruments into account. As this is not the case, we investigate whether the remaining differences can be attributed to the uncertainties of the open path measurements, which are summarized in Table 1. The results are incorporated into Fig. 12. Panel (a) shows the correlation between O_2 and XCO_2 calibration factors. Black squares denote the empirical calibration factors derived from

Figure 10. Calibration measurements performed on 14 April 2015 on top of the KIT-IMK office building north of Karlsruhe.

Figure 11. Histograms of the empirical XCO_2, XCH_4, and O_2 calibration factors for the different instruments with respect to the reference EM27/SUN. The red line overlying the histograms is a fit of a Gaussian function to the histogram. For the histograms, calibration measurements of 29 instruments were accumulated.

the side-by-side measurements. The red squares show calculated calibration factors based on the ME uncertainty budget. The dashed red line is a linear fit through the calculated factors. About half the measured empirical factors are within the bounds of the factors derived from the ME error budget. Furthermore, the slopes of the calculated and empirical factors are in good agreement, confirming that the ME uncertainty is contributing to the uncertainty of the calibration factors. The other contributions for this uncertainty are due to a superposition of various small device-specific imperfections. Panel (b) of Fig. 12 shows the correlation between O_2 and XCH_4 calibration factors. The findings mentioned above for the O_2 and XCO_2 correlation also hold true here.

5 Conclusions and outlook

Based on a long-term intercomparison of column-averaged greenhouse gas abundances measured with an EM27/SUN FTIR spectrometer and with a co-located 125HR spectrometer, respectively, we conclude that the EM27/SUN offers highly stable instrument characteristics on timescales of several years. The drifts on shorter timescales reported by Hedelius et al. (2016) were probably exclusively – as conjectured by the authors of the study – due to a deviation from the instrumental design as originally recommended. The application of a wideband detector suffering from nonlinearity together with steadily decreasing signal levels due to ageing of the tracker mirrors seem to be the reason for the observed drifts.

The favorable instrument stability which is preserved even during transport events and operation under ambient conditions suggests that the EM27/SUN spectrometer is well suited for campaign use and long-term deployment at very remote locations as a supplement of the TCCON. A deployment at remote sites is further facilitated by the recent development of an automated enclosure for the EM27/SUN, which enables unattended remote operation (Heinle and Chen, 2018; Dietrich and Chen, 2018). An annual to biannual check of the instrument performance by performing a side-by-side intercomparison with a TCCON spectrometer seems adequate for quality monitoring. To separate out instrumental drifts from atmospheric signals, the addition of
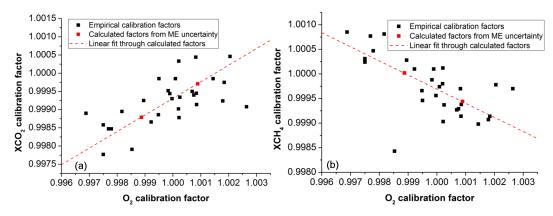
Figure 12. Correlation of O_2 calibration factors and XCO_2 **(a)** as well as XCH_4 **(b)** calibration factors. Black squares show the empirical calibration factors from the side-by-side measurements, red squares show calculated factors derived from the total ME uncertainty shown in Table 1, and the dashed red line is a linear fit through the calculated factors. The slope of empirical and calculated factors is in good agreement.

low-resolution spectra derived from the TCCON measurements is highly useful, because in this kind of comparison, the smoothing error and any possible resolution-dependent biases of the analysis software cancel out. The ensemble performance of 30 EM27/SUN spectrometers turns out to be very uniform, supported by a centralized acceptance inspection performed at KIT before the spectrometers are deployed. When using the empirical ILS parameters derived for each spectrometer, the scatter in XCO_2 amounts to 0.13 ppmv, while it is 0.6 ppbv for XCH_4. The standard deviation of the oxygen columns is 0.04 %. We expect that the conformity of measurement results will be even better than indicated by this scatter, if the remaining empirical calibration factors are taken into account. These empirical calibration factors are likely composed of several small device-specific error contributions; a major contribution was identified to stem from the uncertainty of the ILS measurements.

Continuation and further development of the COCCON activities seem highly desirable for achieving the optimal performance of the growing EM27/SUN spectrometer network. The implemented pre-deployment procedures of testing, optimizing, and calibrating each device – executed by experts at a central facility – help to ensure consistent results from EM27/SUN spectrometers operated in any part of the world. This approach is corroborated by the proven excellent long-term stability of instrumental characteristics, and the proven high degree of stability under thermal and mechanical burdens as they occur during transport. In order to maintain the reliability of the EM27/SUN spectrometers, we suggest investigators send the instrument to KIT for a biennial inspection. The EM27/SUN spectrometer does not require continuous expert maintenance and it is very simple to operate; we therefore expect that many investigators world-wide who are not keen on becoming FTIR experts will be attracted by this measurement device, operating it as a side activity. Current COCCON work supported by ESA in the framework

of the COCCON PROCEEDS project will result in an easy-to-handle preprocessing tool optimized for the EM27/SUN spectrometer. This tool will generate quality-checked spectra from raw interferograms, which then are forwarded to a central data analysis facility. A demonstration setup of the central facility will be part of COCCON PROCEEDS. When finally implemented on an operational level, the facility will remove the whole burden of the quantitative trace gas analysis from the operator and ensure the consistency of the trace gas analysis chain to the utmost degree. Furthermore, it will enable a timely reanalysis of all submitted spectra after upgrades of the retrieval procedures and minimize the risk of data loss if operators for some reason are stopping their activity. Finally, this centralized facility will serve as a unique contact point for the data users.

References

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, https://doi.org/10.5194/acp-16-8479-2016, 2016.

Davis, S., Abrams, M. C., and Brault, J. W.: Fourier Transform Spectrometry, Academic press, 190–198, 2010.

Dietrich, F. and Chen, J.: Portable Automated Enclosure for a Spectrometer Measuring Greenhouse Gases, Geophys. Res. Abstracts, 20, EGU2018-16281-1, https://doi.org/10.13140/RG.2.2.11591.14248, 2018.

Dlugokencky, E. J., Masarie, K. A., Tans, P. P., Conway, T. J., and Xiong, X.: Is the amplitude of the methane seasonal cycle changing?, Atmos. Environ., 31, 21–26, https://doi.org/10.1016/S1352-2310(96)00174-4, 1997.

Dohe, S.: Measurements of atmospheric CO₂ columns using ground-based FTIR spectra, Ph.D. thesis, Karlsruhe Institute of Technology, 2013.

Frankenberg, C., Pollock, R., Lee, R. A. M., Rosenberg, R., Blavier, J.-F., Crisp, D., O'Dell, C. W., Osterman, G. B., Roehl, C., Wennberg, P. O., and Wunch, D.: The Orbiting Carbon Observatory (OCO-2): spectrometer performance evaluation using pre-launch direct sun measurements, Atmos. Meas. Tech., 8, 301–313, https://doi.org/10.5194/amt-8-301-2015, 2015.

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, https://doi.org/10.5194/amt-8-3047-2015, 2015.

Geibel, M. C., Messerschmidt, J., Gerbig, C., Blumenstock, T., Chen, H., Hase, F., Kolle, O., Lavric, J. V., Notholt, J., Palm, M., Rettinger, M., Schmidt, M., Sussmann, R., Warneke, T., and Feist, D. G.: Calibration of column-averaged CH₄ over European TCCON FTS sites with airborne in-situ measurements, Atmos. Chem. Phys., 12, 8763–8775, https://doi.org/10.5194/acp-12-8763-2012, 2012.

Gisi, M., Hase, F., Dohe, S., and Blumenstock, T.: Camtracker: a new camera controlled high precision solar tracker system for FTIR-spectrometers, Atmos. Meas. Tech., 4, 47–54, https://doi.org/10.5194/amt-4-47-2011, 2011.

Gisi, M., Hase, F., Dohe, S., Blumenstock, T., Simon, A., and Keens, A.: XCO₂-measurements with a tabletop FTS using solar absorption spectroscopy, Atmos. Meas. Tech., 5, 2969–2980, https://doi.org/10.5194/amt-5-2969-2012, 2012.

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, https://doi.org/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, https://doi.org/10.1364/AO.38.003417, 1999.

Hase, F., Hannigan, J. T., Coffey, M., Goldman, A., Höpfner, M., Jones, N., P. Rinsland, C., and Wood, S.: Intercomparison of retrieval codes used for the analysis of high-resolution, ground-based FTIR measurements, J. Quant. Spectrosc. Ra., 87, 25–52, 2004.

Hase, F., Drouin, B. J., Roehl, C. M., Toon, G. C., Wennberg, P. O., Wunch, D., Blumenstock, T., Desmet, F., Feist, D. G., Heikkinen, P., De Mazière, M., Rettinger, M., Robinson, J., Schneider, M., Sherlock, V., Sussmann, R., Té, Y., Warneke, T., and Weinzierl, C.: Calibration of sealed HCl cells used for TCCON instrumental line shape monitoring, Atmos. Meas. Tech., 6, 3527–3537, https://doi.org/10.5194/amt-6-3527-2013, 2013.

Hase, F., Blumenstock, T., Dohe, S., Gross, J., and Kiel, M.: TCCON data from Karlsruhe (DE), ReleaseGGG2014R1, TCCON data archive, hosted by CaltechDATA, https://doi.org/10.14291/tccon.ggg2014.karlsruhe01.R1/1182416 (last access: 4 March 2019), 2014.

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, https://doi.org/10.5194/amt-8-3059-2015, 2015.

Hedelius, J. K., Viatte, C., Wunch, D., Roehl, C. M., Toon, G. C., Chen, J., Jones, T., Wofsy, S. C., Franklin, J. E., Parker, H., Dubey, M. K., and Wennberg, P. O.: Assessment of errors and biases in retrievals of XCO$_2$, XCH$_4$, XCO, and XN$_2$O from a 0.5 cm^{-1} resolution solar-viewing spectrometer, Atmos. Meas. Tech., 9, 3527–3546, https://doi.org/10.5194/amt-9-3527-2016, 2016.

Hedelius, J. K., Parker, H., Wunch, D., Roehl, C. M., Viatte, C., Newman, S., Toon, G. C., Podolske, J. R., Hillyard, P. W., Iraci, L. T., Dubey, M. K., and Wennberg, P. O.: Intercomparability of XCO$_2$ and XCH$_4$ from the United States TCCON sites, Atmos. Meas. Tech., 10, 1481–1493, https://doi.org/10.5194/amt-10-1481-2017, 2017

Heinle, L. and Chen, J.: Automated enclosure and protection system for compact solar-tracking spectrometers, Atmos. Meas. Tech., 11, 2173–2185, https://doi.org/10.5194/amt-11-2173-2018, 2018.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bull. Am. Meteorol. Soc., 77, 437–471, https://doi.org/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, https://doi.org/10.1364/AO.46.004774, 2007.

Kiel, M., Hase, F., Blumenstock, T., and Kirner, O.: Comparison of XCO abundances from the Total Carbon Column Observing Network and the Network for the Detection of Atmospheric Composition Change measured in Karlsruhe, Atmos. Meas. Tech., 9, 2223–2239, https://doi.org/10.5194/amt-9-2223-2016, 2016a.

Kiel, M., Wunch, D., Wennberg, P. O., Toon, G. C., Hase, F., and Blumenstock, T.: Improved retrieval of gas abundances from near-infrared solar FTIR spectra measured at the Karlsruhe TCCON station, Atmos. Meas. Tech., 9, 669–682, https://doi.org/10.5194/amt-9-669-2016, 2016b.

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, https://doi.org/10.5194/amt-8-5023-2015, 2015.

Messerschmidt, J., Geibel, M. C., Blumenstock, T., Chen, H., Deutscher, N. M., Engel, A., Feist, D. G., Gerbig, C., Gisi, M., Hase, F., Katrynski, K., Kolle, O., Lavrič, J. V., Notholt, J., Palm, M., Ramonet, M., Rettinger, M., Schmidt, M., Sussmann, R., Toon, G. C., Truong, F., Warneke, T., Wennberg, P. O., Wunch, D., and Xueref-Remy, I.: Calibration of TCCON column-averaged CO$_2$: the first aircraft campaign over European TCCON sites, Atmos. Chem. Phys., 11, 10765–10777, https://doi.org/10.5194/acp-11-10765-2011, 2011.

Morino, I., Uchino, O., Inoue, M., Yoshida, Y., Yokota, T., Wennberg, P. O., Toon, G. C., Wunch, D., Roehl, C. M., Notholt, J., Warneke, T., Messerschmidt, J., Griffith, D. W. T., Deutscher, N. M., Sherlock, V., Connor, B., Robinson, J., Sussmann, R., and

Rettinger, M.: Preliminary validation of column-averaged volume mixing ratios of carbon dioxide and methane retrieved from GOSAT short-wavelength infrared spectra, Atmos. Meas. Tech., 4, 1061–1076, https://doi.org/10.5194/amt-4-1061-2011, 2011.

Nassar, R., Hill, T. G., McLinden, C. A., Wunch, D., Jones, D. B. A., and Crisp, D.: Quantifying CO$_2$ Emissions From Individual Power Plants From Space, Geophys. Res. Lett., 44, 10045–10053, https://doi.org/10.1002/2017GL074702, 2017.

Olsen, S. C. and Randerson, J. T.: Differences between surface and column atmospheric CO$_2$ and implications for carbon cycle research, J. Geophys. Res.-Atmos., 109, D02301, https://doi.org/10.1029/2003JD003968, 2004.

Ostler, A., Sussmann, R., Rettinger, M., Deutscher, N. M., Dohe, S., Hase, F., Jones, N., Palm, M., and Sinnhuber, B.-M.: Multistation intercomparison of column-averaged methane from NDACC and TCCON: impact of dynamical variability, Atmos. Meas. Tech., 7, 4081–4101, https://doi.org/10.5194/amt-7-4081-2014, 2014.

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, https://doi.org/10.5194/amt-5-1627-2012, 2012.

Schneider, M. and Hase, F.: Ground-based FTIR water vapour profile analyses, Atmos. Meas. Tech., 2, 609–619, https://doi.org/10.5194/amt-2-609-2009, 2009.

Sepúlveda, E., Schneider, M., Hase, F., García, O. E., Gomez-Pelaez, A., Dohe, S., Blumenstock, T., and Guerra, J. C.: Long-term validation of tropospheric column-averaged CH$_4$ mole fractions obtained by mid-infrared ground-based FTIR spectrometry, Atmos. Meas. Tech., 5, 1425–1441, https://doi.org/10.5194/amt-5-1425-2012, 2012.

Tallis, L., Coleman, M., Gardiner, T., Ptashnik, I., and Shine, K.: Assessment of the consistency of H$_2$O line intensities over the near-infrared using sun-pointing ground-based Fourier transform spectroscopy, J. Quant. Spectrosc. Ra., 112, 2268–2280, https://doi.org/10.1016/j.jqsrt.2011.06.007, 2011.

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, D22305, https://doi.org/10.1029/2006JD007154, 2006.

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, https://doi.org/10.5194/amt-3-1351-2010, 2010.

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.: The Total Carbon Column Observing Network, Philos. Trans. Ser. A, 369, 2087–2112, https://doi.org/10.1098/rsta.2010.0240, 2011.

Wunch, D., Toon, G. C., Sherlock, V., Deutscher, N. M., Liu, X., Feist, D. G., and Wennberg, P. O.: The Total Carbon Column Observing Network's GGG2014 Data Version, 10.14291/TCCON.GGG2014.DOCUMENTATION.R0, 2015.

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) X_{CO_2} measurements with TCCON, Atmos. Meas. Tech., 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.

Ye, X., Lauvaux, T., Kort, E. A., Oda, T., Feng, S., Lin, J. C., Yang, E., and Wu, D.: Constraining fossil fuel CO_2 emissions from urban area using OCO-2 observations of total column CO_2, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1022, 2017.