

**Assessment of a multi species in situ FTIR for precise  
atmospheric greenhouse gas observations**

**S. Hammer<sup>1</sup>, D. W. T. Griffith<sup>2</sup>, G. Konrad<sup>1</sup>, S. Vardag<sup>1</sup>, C. Caldow<sup>2</sup> and I. Levin<sup>1</sup>**

[1]{Institut für Umweltphysik, University of Heidelberg, Germany}

[2]{School of Chemistry, University of Wollongong, Australia}

Correspondence to: S. Hammer ([shammer@iup.uni-heidelberg.de](mailto:shammer@iup.uni-heidelberg.de))

## Abstract

We thoroughly evaluate the performance of a multi-species, in situ **Fourier Transform InfraRed (FTIR)** analyser with respect to high accuracy needs for greenhouse gas monitoring networks. The in situ FTIR analyser **is shown to** measure CO<sub>2</sub>, CO, CH<sub>4</sub> and N<sub>2</sub>O mole fractions continuously, all with better reproducibility than the inter-laboratory compatibility (ILC) goals, requested by the **World Meteorological Organization (WMO) for the Global Atmosphere Watch (GAW) program**. Simultaneously determined  $\delta^{13}\text{CO}_2$  reaches reproducibility as good as 0.03‰. Residual dependencies between the measured components and the thermodynamic properties of the sample, **(temperature, pressure and flow rate)** and the cross sensitivities among the sample constituents **are investigated and quantified**. We describe **an improved sample delivery and control system that minimises the pressure and flow rate variations making post processing corrections for those quantities non-essential**. Temperature disequilibrium effects resulting from the evacuation of the sample cell are quantified and **improved by the usage of a faster temperature sensor**. The instrument has proven to be linear for all **measured** components in the ambient **concentration** range. The temporal stability of the instrument **is characterised on different time scales**. Instrument drifts on a weekly time scale **are only observed for CH<sub>4</sub> (0.04 nmol mol<sup>-1</sup> day<sup>-1</sup>) and  $\delta^{13}\text{CO}_2$  (0.02‰ day<sup>-1</sup>)**. Based on 10 months of continuously collected quality control measures **the long term reproducibility of the instrument is estimated to  $\pm 0.016 \mu\text{mol mol}^{-1} \text{CO}_2$ ,  $\pm 0.03 \text{‰ } \delta^{13}\text{CO}_2$ ,  $\pm 0.14 \text{ nmol mol}^{-1} \text{CH}_4$ ,  $\pm 0.1 \text{ nmol mol}^{-1} \text{CO}$  and  $\pm 0.04 \text{ nmol mol}^{-1} \text{N}_2\text{O}$** . We propose that a calibration and quality control scheme with weekly calibrations of the instrument that is sufficient to reach WMO-GAW inter laboratory compatibility goals.

## 2 1 Introduction

3 The globally distributed in situ greenhouse gas (GHG) monitoring network is one of the  
4 mainstays of modern climate research. Only a few continuous atmospheric CO<sub>2</sub> records go  
5 back to the 1950s (Keeling et al., 1976), but nowadays many stations monitor nearly all long-  
6 lived GHGs with in situ instrumentation (Worthy, 2003; Messenger et al., 2008). The required  
7 accuracy and precision for measurements of the most important GHG species have been set to  
8 limits that allow extracting the required biogeochemical information from spatial differences  
9 **that** are needed for quantifying continental scale GHG fluxes and their inter-annual changes  
10 (WMO report No. 5, 1981, cited in Francey and Steele, 2003). Over the most recent decades,  
11 non-dispersive infrared (NDIR) analysis of CO<sub>2</sub> and gas chromatography (GC) of CO<sub>2</sub> and all  
12 other long-lived GHGs has been proven to provide this accuracy and precision; they have thus  
13 become standard techniques for GHG monitoring. Both techniques require special care,  
14 maintenance, frequent calibration and quality control measures to **guarantee** data quality; they  
15 are thus labour-intensive in their day-to-day operation. In recent years, optical techniques like  
16 Cavity Ring Down Spectroscopy (CRDS), Off-Axis Integrated Cavity Output Spectroscopy  
17 (OA-ICOS) or Fourier Transform Infrared Spectroscopy (FTIR) have reached similar or even  
18 better precisions than traditional GC systems. Contrary to the discrete samples measured with  
19 GC systems, these techniques offer real continuous data acquisition and are in general less  
20 labour intensive (Winderlich et al., 2010). Optical techniques can be divided into two  
21 fundamentally different methods: (1) Laser based methods and (2) broad band infrared  
22 spectroscopy. The major distinctive feature between them is the range of the recorded and  
23 evaluated absorption spectrum. Whereas laser-based instruments are tuned to a narrow  
24 absorption window with ideally little interference from other species, the FTIR scans a broad  
25 IR range, thereby offering the possibility to measure a large number of species  
26 simultaneously. The laser-based techniques like CRDS and OA-ICOS have been extensively  
27 and successfully studied by the GHG measurement community; some instruments have  
28 already found their way into today's observational networks (Winderlich et al., 2010).

29 **Another benefit of optical spectroscopy is the possibility of discriminating isotopologues, e.g.**  
30 **it allows for continuous measurement of  $\delta^{13}\text{CO}_2$ . The challenge of isotopologue-selective**  
31 **measurements is, however, to accurately determine the absolute sums of all individual**

isotopologues. To do so, either all relevant isotopologues have to be measured independently, as done by the in situ FTIR analyser, or assumptions on the mean relation between  $^{13}\text{CO}_2$  and  $^{12}\text{CO}_2$  in clean air have to be taken into account during evaluation and/or should be already incorporated in the calibration of the instrument (Chen et al., 2010).

So far, studies discussing the use of the in situ FTIR technique for GHG monitoring purposes are rare, although promising (Griffith et al., 2010). Therefore, detailed investigations of the potential and the possible shortcomings of this technique with respect to the accuracy goals set by WMO/GAW experts for clean background air monitoring should be performed before it is used widely within the monitoring community. In particular questions regarding long term stability, calibration frequency and cross sensitivity of different trace gases have to be addressed. In the present paper we report on experiments that were performed at the University Heidelberg, Institut für Umweltphysik (IUP) on instrument repeatability, parameter- and cross sensitivity, linearity and long term stability of an in situ FTIR analyser that was designed and built at the University of Wollongong, Australia (UoW) (Griffith et al., 2012). These data have been gathered in the IUP laboratory as well as in the course of the ICOS (Integrated Carbon Observation System, <http://www.icos-infrastructure.eu/>) Demonstration Experiment, where the instrument was run at two field stations in Europe. The UoW FTIR instrument is subject to an ongoing development process, and many findings of this paper have already led to improvements in newer instrument versions. Nonetheless, our findings are generally applicable to any in situ FTIR instrument and can therefore be used as guideline for in situ FTIR users in order to raise awareness for high-end accuracy applications.

In the first part of this paper (section 2) the initial instrumental setup and its subsequent modifications are introduced, along with a description of the spectroscopic retrieval technique and the standard operating conditions used. Section 3 discusses and quantifies the residual sensitivities of the measured mole fractions to the thermodynamic properties of the sample, such as temperature and pressure. The issue of inter-species cross sensitivities is discussed as well. This section is extended in Appendix A by an evaluation of the thermodynamic conditions in the cell after sample exchange. After characterising the instrument's sensitivities its response function in the ambient concentration range is investigated in section 4, followed by an exhaustive study on instrument stability and performance on short (weekly) and long

1 (monthly) time scales in section 5. This section also includes an empirical determination of  
2 the required calibration frequency. An overall error assessment is given in section 6. The  
3 paper concludes with a discussion of the general applicability of the in situ FTIR analyser for  
4 background greenhouse gas monitoring purposes, as well as a recommendation for further  
5 improvement of the instrument performance.

6

1

## 2    **2    Instrumental setup and sample handling**

3    Griffith et al. (2010, 2012) describe the improved in situ FTIR instrument used in the present  
4    work, which is based on the early version of the instrument as described by Esler et al.  
5    (2000a; 2000b) and findings of the current study. The essential parts of the in situ FTIR  
6    analyser and the modifications introduced in Heidelberg, based on the first year's findings,  
7    are presented here. The different hardware configurations of the instrument throughout this  
8    study are specified in section 2.1 below.

### 9    **2.1    Instrument components and sample handling**

10    The instrument consists of a commercially available FTIR interferometer (IRcube, Bruker  
11    Optics, Germany) and a 3.5 litre multi-pass cell with 24 m optical path length (PA-24,  
12    InfraredAnalysis, Anaheim, USA) (see Figure 1). To avoid artefacts, the transfer optics  
13    between the interferometer, the multi-pass cell and the interferometer housing itself are  
14    permanently flushed with high purity nitrogen (99.999%). The FTIR interferometer and the  
15    multi-pass cell are aligned via an optical bench situated in an actively temperature controlled  
16    enclosure. In its basic configuration the in situ FTIR analyser has four separate sample inlets  
17    and air is drawn through the instrument using an oil-free vacuum pump (model MV2NT,  
18    Vacuubrand, Germany) at the outlet of the instrument. More intake lines can be added via a  
19    multi-position valve (MWSD16 selection valve, Valco, USA) connected to any of the four  
20    standard inlets.

21    The in situ FTIR analyser has a built-in sample drying system consisting of a 24 inch Nafion  
22    dryer (Permapure, Toms River, NJ, USA) operated in counter-flow mode, followed by a  
23    chemical dryer filled with magnesium perchlorate ( $\text{Mg}(\text{ClO}_4)_2$ ) for uptake of residual water.  
24    The Nafion dryer is placed in a separate compartment of the instrument, along with tubing,  
25    valves, PC and power supplies. The drying system reaches a dew point of  $\approx -65^\circ\text{C}$  and can be  
26    by-passed if measurement of  $\text{H}_2\text{O}$  isotopologues is of interest. For more details on  $\text{H}_2\text{O}$   
27    isotopologue measurements refer to Parkes et al. (paper in preparation for AMT).

28

The modifications of the instrumental setup performed in Heidelberg are all related to sample handling as well as controlling and measuring sample conditions. In the following the different instrumental setups are introduced:

a) *Initial UoW setup*: The sample flow through the system is adjusted with a needle-valve (NV) and monitored using a flow meter (FM) mounted at the outlet of the cell (see Fig. 1, blue parts). The multi-pass cell is equipped with an in situ PT100 resistance temperature detector (RTD) and with a pressure sensor (HPM-760s, Teledyne Hastings, USA) to determine the thermodynamic sample properties. Both sensors are assumed to be linear and calibrated using a simple two point calibration.

b) *ICOS Demonstration Experiment setup*: For additional stabilisation of the sample flow and pressure in the cell, the ICOS Demonstration Experiment setup used an external add-on mass flow-controller (MFC) (2 slpm, MKS Instruments, USA) in between the outlet of the multi-port valve and the instrument air inlet (see Figure 1). Pressure and temperature sensors remain unchanged.

c) *Intermediate setup*: The Intermediate setup has one built-in mass flow-controller (Model 3660, Kofloc, Japan), replacing the needle-valve flow meter unit as well as the external MKS flow controller (compare Figure 1). The internal MFC is located at the outlet of the cell and can be operated in flow control mode or in pressure control mode, when combined with a software embedded feedback-loop to the pressure sensor. The RTD temperature sensor in the cell was also replaced by a faster responding J-type thermocouple. The thermocouple was centred in the cell to get a more representative temperature measurement. The temperature sensor of the enclosure temperature control was removed from the cell wall. This version was equivalent to that originally provided by Ecotech (Spectronus GHG analyser, Knoxfield, Australia) in the first commercially available versions of the analyser.

d) *IUP setup*: In addition to the Intermediate setup the sample pressure in the cell is controlled by an electronic pressure controller (EPC) (P-602CV EL-Press, Bronkhorst, The Netherlands) mounted at the cell inlet (see. Figure 1). This configuration is functionally equivalent to the current version described by Griffith et al. (2012) with both pressure and flow control and available commercially from Ecotech.

**In all configurations** the cell is operated at slight overpressure to increase signal-to-noise ratio and to ease leak detection. A diaphragm pump fitted with an EPDM membrane (N 86 KN.18,

KNF Neuberger, Germany) is used to pressurize ambient air up to 1800 hPa. Long term GC experience has shown that these pumps have negligible effects on the measured species. Nevertheless each pump is tested for contamination prior to its use.

## 2.2 Spectroscopic concentration retrieval

The spectral range of the IRcube is 1800-7500  $\text{cm}^{-1}$  with a  $1\text{cm}^{-1}$  resolution. The in situ FTIR analyser records and stores a broadband absorption spectrum from 1800-5000  $\text{cm}^{-1}$  for each measurement. The recorded spectra are analysed online by nonlinear least squares fitting of sections of the measured spectrum with a modelled spectrum calculated from the HITRAN database of absorption line parameters (Rothman et al., 2005). The theoretical spectrum is calculated by MALT (Multiple Atmospheric Layer Transmission) as described elsewhere (Griffith, 1996; Griffith et al., 2003; Griffith 2012). Three separate spectral regions are fitted for each spectrum: 2150-2320  $\text{cm}^{-1}$  for  $^{13}\text{CO}_2$ ,  $^{12}\text{CO}_2$ , CO and  $\text{N}_2\text{O}$ , 3001-3150  $\text{cm}^{-1}$  for  $\text{CH}_4$  and 3520-3775  $\text{cm}^{-1}$  for  $\text{CO}_2$  (all isotopologues) and residual  $\text{H}_2\text{O}$ . The spectral analysis determines the molar concentrations ( $c_i$  [ $\text{mol m}^{-3}$ ]) of each gas species. To convert molar concentrations into mole fractions ( $x_i$  [ $\text{mol mol}^{-1}$ ]), sample pressure and temperature need to be taken into account:

$$x_{i(\text{wet})} = c_i / (p/RT) \quad (1)$$

where  $p$  is the absolute cell pressure,  $T$  the absolute temperature,  $R$  the Universal Gas Constant and  $i$  the investigated species. Since residual water is determined from the spectrum for each sample, dry air mole fractions  $x_i$  can easily be derived from:

$$x_{i(\text{dry})} = x_{i(\text{wet})} / (1 - x_{\text{H}_2\text{O}}) \quad (2)$$

## 2.3 Standard Operating Conditions (SOC)

The Standard Operating Conditions (SOC) of the in situ FTIR comprises the instrument settings, measurement mode and interval as well as sample change-over strategies.

- Instrument settings: The enclosure temperature is set to 30.0°C. It is stable within  $\pm 0.06^\circ\text{C}$ , which leads to a stability of the cell-temperature of  $\pm 0.02^\circ\text{C}$  for moderately stable laboratory conditions of  $\pm 1^\circ\text{C}$ . All samples are dried using the built-in drying system (Fig. 1). The sample flow rate is set to  $1 \pm 0.02$  slpm. For the ICOS Demonstration Experiment setup the cell pressure was set via sample delivery pressure



and was kept at  $1100 \pm 8$  hPa. In the modified IUP setup, sample pressure and flow are controlled separately to better than  $\pm 0.1$  hPa and  $\pm 0.008$  slpm, respectively.

- **Measurement mode:** The in situ FTIR analyser offers static and dynamic measurement modes. In the static measurement mode, the cell is filled with the sample, sample flow is shut off and the sample is subsequently measured for a certain measurement interval. The dynamic mode measures the sample for a certain interval while it is continuously flushed through the cell. Since systematic differences between the two measurement modes were found (see. Sec. 3.5) we chose to flush both sample types, ambient air and air from high pressure cylinders, continuously through the cell with the same flow rate of  $1.00 \pm 0.02$  slpm. This approach is taken to assure comparability for both sample types.
- **Measurement interval:** The measurement interval is set to 3 min, in which a 2.5 min spectra collection period is followed by 0.5 min online analysis with the installed PC. This time interval was chosen as a compromise between instrument precision (increasing with averaging time, see Sec. 5.1) and smoothing-out natural variability in the ambient air, which itself blurs the averaged spectra. For example, in Heidelberg CO values can change by more than  $100 \text{ nmol mol}^{-1}$  within 30 min during rush-hour situations (Hammer et al., 2009). In addition, the 3 min measurement interval equals the approximate turn over time of the sample in the cell at 1 slpm flow rate.
- **Sample change-over strategy:** Under SOC each sample change-over, i.e. changing from ambient to cylinder measurements and vice versa, involves a two-step evacuation of the cell that is described in detail in Appendix A. Possible disadvantages of evacuating the cell, i.e. by disturbance of moisture or temperature equilibrium, will be discussed in Section 3.6 and Appendix A. In SOC each cylinder measurement is performed over 30 min, including the sample change-over, and requires a total gas volume of about 27 litres of air.

## 3 Residual sensitivities to sample properties and inter-species cross sensitivities

### 3.1 Origin of residual and cross sensitivities

#### 3.1.1 Residual sensitivities to sample properties

The line shapes of the investigated species are dependent on pressure- and Doppler-broadening and thus depend on sample properties like pressure and temperature. For both broadening effects temperature- and pressure-dependent line widths are tabulated in the HITRAN 2004 database (Rothman et al., 2005); they are considered by the FTIR spectra evaluation program MALT (Griffith, 1996), using the measured sample temperature and pressure. However, the line shape parameters themselves are subject to ongoing improvement: as an example, for the CO<sub>2</sub> line parameters, differences of up to a few percent are reported in recent studies (Long et al., 2011; Nakamichi et al., 2006). Small errors in the HITRAN parameters lead to systematic biases in the retrieved molar concentrations. We will refer to this error contribution as the *line shape error*.

Furthermore, the retrieved molar concentrations are biased by potential offsets in the measured sample temperature and pressure, since these quantities are directly used in MALT to select the tabulated line shape parameters. In the following, these introduced retrieval biases are referred to as the *introduced spectroscopic error*.

In addition to the *line shape and the introduced spectroscopic error*, the conversion from molar concentrations to mole fractions (see Eq. 1) constitutes another direct link to the measured sample properties and their precision. This direct link is established through the sample density (dependent on temperature and pressure) and is thus approximately the same for all species on a percentage basis. This emphasizes the importance of accurate sample temperature- and pressure measurements to minimize the *density error*. Accurate and temporally stable sensor calibration down to a level of 0.01% is thus important to reach the required accuracy and precision for the greenhouse gas measurements. Determining the average sample temperature is challenging since temperature is not homogeneous within the cell. In all current setups temperature is measured in one location only, assuming a constant

temperature distribution. However, the *true* temperature distribution in the cell depends on sample flow rate and injection (see Appendix A). The measured mole fractions may thus implicitly depend on the flow rate as well.

The combination of the line shape, the introduced spectroscopic and the density error are thus the underlying cause of “residual” pressure and temperature sensitivities. The term “residual” is used as the first order changes of these parameters are already accounted for and only the deviations between the measured and the *true* sample and spectroscopic properties cause these effects. Since all error contributions are coupled, it is experimentally not possible to disentangle the error contributions. However, in a synthetic MALT study it is possible to investigate the introduced spectroscopic and the density error. This approach will be discussed in Section 3.3 and 3.4.

### 3.1.2 Inter-species cross sensitivities

Apart from residual pressure, temperature and flow sensitivities, additional inter-species cross sensitivities exist, which are caused by overlapping spectral absorption regions of different trace species. Generally the MALT least squares fit is able to disentangle the contributions to absorption at each wave number and to attribute their shares to the different species. Nevertheless, the MALT algorithm, the measured spectra, and the HITRAN data are not perfect and small inter-species cross sensitivities remain. Since H<sub>2</sub>O absorption in the infrared region occurs at many different wavelengths, the inter-species sensitivity to residual water vapour is noticeable for all investigated species, apart from CH<sub>4</sub>. The second strongest absorber in ambient air is CO<sub>2</sub>. Since the <sup>13</sup>CO<sub>2</sub>, and N<sub>2</sub>O absorptions in the 2150-2320 cm<sup>-1</sup> region overlap with strong absorption of <sup>12</sup>CO<sub>2</sub>, measurable inter-species cross sensitivities to CO<sub>2</sub> also exist.

## 3.2 Sensitivity experiments

We carried out a series of dedicated experiments to detect and quantify the residual sensitivities to sample properties (pressure, temperature and flow) and the cross sensitivities (H<sub>2</sub>O and CO<sub>2</sub>) of the in situ FTIR analyser. If applicable, these measurements were used to define correction functions for each species. In all experiments the investigated sample property or species was systematically varied while all other parameters or species were kept as constant as possible. The experiments have been repeated several times over the course of

one year to investigate temporal stability of the sensitivities. For each test we used ambient Heidelberg air collected in high pressure cylinders with a diving compressor (model P3W, Bauer, Germany) and dried to a dew-point of approximately -40°C. The cylinders (40L L6X aluminium, Luxfer, UK) and pressure regulators (model 14A, Scott Specialty Gases, USA) have proven to be suitable for high precision GHG measurements by GC analysis (Hammer, 2008). The GHG concentrations in each test cylinder were checked for drifts by GC analysis before and after use.

To investigate the residual sensitivities against temperature, pressure and flow, the respective parameter was tuned at the FTIR analyser itself. Determining the inter-species cross sensitivities involved a custom-made mixing device, consisting of two mass flow controllers (MFC) and a scrubbing agent, either Ascarite® for CO<sub>2</sub>, or Mg(ClO<sub>4</sub>)<sub>2</sub> for H<sub>2</sub>O. The mixing device divides the sample stream into two branches, one of them containing the scrubbing agent. After scrubbing, both branches are re-combined and the mixed gas is injected into the in situ FTIR analyser.

A residual or cross sensitivity experiment determines the concentrations of a fixed sample with respect to at least four different settings of the investigated sensitivity parameter. For each setting we allowed sufficient time to re-establish equilibrium in the whole system, i.e. until the variability in CO<sub>2</sub> was on the order of the instrument's repeatability, and then held constant for at least 30 min. The averaged dry air mole fractions for each equilibrium setting were then used to quantify the sensitivity. In the following sections, the results of the experiments will be discussed in detail with respect to their significance and temporal stability. As a reference for the required precision to monitor natural variability at clean air background sites, we will refer to the inter-laboratory compatibility (ILC) goals as defined by the WMO-GAW expert group (WMO, 2011). The authors are aware that “precision” and “compatibility” are two different concepts. However, since to our knowledge no explicit precision goals for GHG measurements are defined, we use the inter-laboratory compatibility instead, which are as follows: 0.1 or 0.05 µmol mol<sup>-1</sup>, respectively for CO<sub>2</sub> in the Northern and Southern Hemispheres, 0.01 ‰ for δ<sup>13</sup>CO<sub>2</sub>, 2 nmol mol<sup>-1</sup> for CO and CH<sub>4</sub> and 0.1 nmol mol<sup>-1</sup> for N<sub>2</sub>O. A compilation of all residual and cross sensitivity parameters is given in Table 1.

### 3.3 Residual Pressure Sensitivity (RPS):

As the accuracy of the piezo-resistive pressure transducer is 0.25% and thus much larger than the required 0.01% significant calibration offsets can be expected. The magnitude of Residual Pressure Sensitivity (RPS) depends on the accuracy and calibration of the pressure sensor as well as errors in the Hitran pressure-dependent line widths and MALT model. Consequently, all five species, CO<sub>2</sub>,  $\delta^{13}\text{CO}_2$ , CO, CH<sub>4</sub> and N<sub>2</sub>O, show a significant residual sensitivity to cell pressure. Two different pressure ranges, one from 800 to 1200 hPa and a sub-range around the operating pressure from 1085 to 1115 hPa were tested. The observed RPSs were linear and compatible for both pressure ranges as displayed in Figure 2a for CO<sub>2</sub>. This allows the use of a linear correction function to account for the RPS. The slope of the correction function was determined by a weighted total least squares fit, accounting for errors in pressure and the investigated species mole fractions (Krystek and Anton, 2007). During the year of our investigations, nine RPS experiments were conducted for all species. The temporal evolution of the derived pressure sensitivity slopes for CO<sub>2</sub> is shown in Figure 2b. The observed slopes vary between 0.0078 and 0.0092 [ $\mu\text{mol mol}^{-1} \text{ hPa}^{-1}$ ]. The uncertainties of the slopes depend on the investigated pressure range. Since the RPS has proven to be linear, a larger pressure range is preferable to determine the sensitivity since the fit is more stable and the experiment is easier to conduct. Table 1 summarizes the averaged slopes and their 1 $\sigma$  standard deviation. For CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O the residual pressure sensitivity slopes were temporally stable within their uncertainties and had a standard deviation of 10 %.  $\delta^{13}\text{CO}_2$  and CO showed a slight temporal change leading to a standard deviation of 30 % for the correction functions for these two components.

Comparing the RPS on a percentage basis shows that the corrections factors for the different trace gases differ only by ca. 25%. This would be in accordance with substantial contribution from a common density error (see Sec. 3.1.1). Re-analysing the measured spectra with different pressures allowed investigation of the effect of the density error and the introduced spectroscopic error separately. CH<sub>4</sub> and N<sub>2</sub>O did not show spectroscopic error contributions, as both species do not have individually resolved lines. For CO<sub>2</sub> the effect of the density error and the introduced spectroscopic error are of similar magnitude whereas for CO the introduced spectroscopic error is three times larger. Since  $\delta^{13}\text{CO}_2$  is calculated as a ratio, the density effect and most of the introduced spectroscopic effect cancel.

To better judge the relevance of the residual sensitivities, an example based on typical ranges of the investigated parameter is given in Table 1 as well. Pressure variations of  $\pm 10$  hPa typical of the initial UoW hardware version translate to  $0.17 \mu\text{mol mol}^{-1} \text{CO}_2$  variations; they are thus larger than the inter-laboratory compatibility (ILC) target for  $\text{CO}_2$ . The same is true for  $\delta^{13}\text{CO}_2$ , and  $\text{N}_2\text{O}$  (see Table 1). All three components thus need to be corrected. The uncertainty introduced by the residual pressure correction is negligible for  $\text{CO}_2$  and  $\text{N}_2\text{O}$ . In case of  $\delta^{13}\text{CO}_2$  a temporal trend in the RPS leads to a larger uncertainty in the averaged correction factor. This error can be reduced by introducing a temporally changing RPS correction. Even assuming a constant RPS reduces the pressure-induced deviation by a factor of 2.5 compared to the uncorrected values. For  $\text{CO}$  and  $\text{CH}_4$  the RPS corrections are smaller than the required ILCs for the assumed 10 hPa pressure change. Still, for  $\text{CH}_4$  in contrast to  $\text{CO}$  the RPS is significant and its correction will improve the precision of the measurements. Apart from  $\delta^{13}\text{CO}_2$ , the stability of the correction parameters indicates that annual determination of the residual pressure sensitivity should be sufficient to take into account systematic long term drifts of the RPSs. The underlying cause of the drift in the  $\delta^{13}\text{CO}_2$  RPS correction parameter is not yet understood. Temporal drifts in the calibration of the pressure sensor would cause a systematic drift in the correction parameters of all species except  $\delta^{13}\text{CO}_2$ .

For the instrument versions after the ICOS demonstration setup the pressure is controlled to  $\ll 1$  hPa and pressure corrections become very small or negligible. In the IUP instrumental setup, a potential drift of the pressure sensor can additionally be controlled by the second pressure sensor in the additional EPC.

## 3.4 Temperature sensitivities

### 3.4.1 Residual Temperature Sensitivity (RTS)

Compared to the RPS, the Residual Temperature Sensitivity (RTS) has the additional complication that we do not measure the *true* mean sample temperature with one temperature sensor (see Appendix A). The RTD sensor used in the ICOS Demonstration Experiment setup was placed near the outlet of the cell whilst the thermocouple used from the Intermediate

1 **setup onwards** was located in the middle of the cell. The sample temperature distribution  
2 within the cell depends **on** at least three parameters, a) the set temperature of the enclosure, b)  
3 the sample temperature when entering the cell, c) the sample flow rate. In total, six dedicated  
4 RTS experiments were performed for all species. In order to investigate the RTS the  
5 temperature of the cell enclosure was varied either in its normal operational range, i.e.  
6 between 29.9°C and 30.1°C, or in a wider range from 29.7°C to 30.3°C. For each set  
7 temperature step we allowed enough time for the **spectrometer and the sample cell** to  
8 equilibrate. In our setups this took roughly 60 to 90 min, based on the CO<sub>2</sub> variability. For  
9 CO<sub>2</sub>, N<sub>2</sub>O and  $\delta^{13}\text{CO}_2$ , linear residual temperature sensitivities were observed in each  
10 experiment. Although each experiment showed good linear relations **for CO<sub>2</sub>**, the slopes  
11 varied substantially and even changed sign after six months. For N<sub>2</sub>O and  $\delta^{13}\text{CO}_2$  the RTS  
12 was stable within 30% for all experiments. The residual temperature sensitivity for CO and  
13 CH<sub>4</sub> was weak, not temporally stable and only for some experiments a distinct relation to cell  
14 temperature **was** observed. The averaged RTS slopes and their standard deviations are given  
15 in Table 1. For better **classification** of the results an example based on observed peak-to-peak  
16 temperature variability of 0.1°C is listed **in Table 1** as well. When considering the ILC targets,  
17 only the observed RTS for  $\delta^{13}\text{CO}_2$  is significant and needs to be corrected.

18 In order to investigate the effect of the initial sample temperature on the cell temperature we  
19 conducted an experiment where a cylinder was first measured at room temperature before the  
20 entire cylinder was cooled to zero degrees Celsius and measured again. The cell temperature  
21 as well as the measured mole fractions did, however, not change in this experiment. This  
22 implies that the **residence time of the** sample in the inlet and drying system, **i.e. the Nafion**  
23 **dryer**, is sufficient to compensate for **at least** a 30°C temperature difference **of the incoming**  
24 **sample**.

### 26 **3.4.2 Temperature Disequilibrium Sensitivity (TDS)**

27 **The Temperature Disequilibrium Sensitivity (TDS) described in this subsection is related to**  
28 **the RTD temperature sensor used in the UoW and the ICOS Demonstration Experiment setup.**  
29 **These findings resulted in a replacement of the RTD temperature sensor with a J-type**  
30 **thermocouple from the Intermediate setup onwards.**

Using the RTD sensor, stronger temperature sensitivities were present in all cylinder gas measurement records, implying that temperature is one of the key parameters to perform precise calibration measurements. In the course of the ICOS Demonstration Experiment the in situ FTIR analyser was set up in different laboratories under different environmental conditions, i.e. averaged laboratory temperatures ranged from 17 to 27°C. Although the enclosure temperature was stable at  $30.00 \pm 0.05^\circ\text{C}$  at all locations, cell temperature varied slightly and was anti-correlated with laboratory temperature. Figure 3 shows the deviations from the averaged mole fraction of the 24- to 48-hourly measured sub-target tank with respect to the cell temperature. The other simultaneously measured cylinders, two calibration gases and the regular target gas, show similar temperature dependencies. We will further refer to this effect as the Temperature Disequilibrium Sensitivity (TDS).

The observed TDSs for cylinder measurements are much larger than the RTSs; which were derived from the dedicated temperature experiments described in section 3.4.1. The slopes of the linear TDS and their uncertainties are summarized in Table 1 as well. For the typical temperature range of  $0.1^\circ\text{C}$  the effect is significantly larger than the ILC goals for  $\text{CO}_2$ ,  $\delta^{13}\text{CO}_2$  and  $\text{N}_2\text{O}$ . A possible cause for the difference between the two temperature sensitivity results may arise from different conditions of the instrument. While in the dedicated RTS experiments sample air was continuously flushed through the cell, each cylinder measurement, performed under standard operating conditions, comprises evacuation of the cell prior to the measurement in flow mode (Appendix A). The evacuation of the cell causes adiabatic cooling on the order of  $5^\circ\text{C}$ . Although this adiabatic cooling effect is compensated for during the re-filling of the cell, the slow response of the RTD sensor, caused by its large thermal mass, leads to a biased temperature measurement, even after the sample temperature has recovered. In the UoW and ICOS demonstration experiment setup the evacuation-induced temperature deviation was further amplified by the fact that the sensor of the active enclosure temperature control was mounted directly on the cell walls. The adiabatic cooling thus disturbed the temperature equilibrium in the enclosure, leading to a feedback-loop of the active temperature control.

To investigate if the TDS can be explained by a biased sample temperature measurement we examined the influence of a  $1^\circ\text{C}$  temperature bias on the retrieved concentrations of a synthetic absorption spectrum. This allows separating the impact of the density error from the introduced spectroscopic error. The line shape error does not show up in such a study. The



theoretical temperature bias correction parameters are listed in Table 1 and shown in Figure 3 as dashed black lines. The theoretical temperature sensitivity of an assumed temperature bias of 1°C agrees well with the observed TDS for all species. Only for N<sub>2</sub>O the theoretical temperature sensitivity is smaller than the observed TDS.

This result gives strong evidence for biased sample temperature measurements after an evacuation process. The TDS was the main reason to replace the RTD sensor with a faster responding thermocouple. Such thermocouples have become the standard temperature sensor from the Intermediate setup onwards. In addition, we moved the sensor of the active enclosure temperature control out of direct contact with the cell wall. With these measures, the TDS was no longer observed for cylinder measurements.

For all cylinder measurements performed with the ICOS Demonstration Experiment setup we use the TDS correction parameters derived from the sub-target gas measurements to correct all our standard and target gas measurements. The anti-correlated relation between room and cell temperature is most likely caused by an overcompensation of the active enclosure temperature control.

### 3.5 Flow rate sensitivity

Neither the spectroscopic nor the mole fraction determinations have a direct link to the sample flow rate through the cell. The flow rate has only an indirect effect through the temperature distribution in the cell. Thus we expect the flow rate sensitivity to be small compared to the temperature sensitivity. Nevertheless, the flow rate sensitivity was investigated in three dedicated experiments, mainly motivated from the observed difference between a constantly flushed and a closed-off cell. Apart from CO, the measured mole fractions are always higher in a closed off cell than in a constantly flushed cell. Under our standard operating conditions this effect can be as large as 0.25  $\mu\text{mol mol}^{-1}$  for CO<sub>2</sub>, 0.6 ‰ for  $\delta^{13}\text{CO}_2$ , -0.1  $\text{nmol mol}^{-1}$  for CO, 1.4  $\text{nmol mol}^{-1}$  for CH<sub>4</sub> and 0.28  $\text{nmol mol}^{-1}$  for N<sub>2</sub>O. We suppose that the observed difference between static and dynamic measurements is caused by different temperature distributions in the cell.

The observed flow sensitivity in the range between 0.8 and 1.2 slpm could be linearly approximated, however the linearity broke down when approaching zero flow. As expected, the flow sensitivities are small. In addition to this, the gas-flow through the cell was very stable due to the additional sample flow controller. Thus the effects of the flow sensitivity are negligible for all trace gases (see. Table 1). For  $\delta^{13}\text{CO}_2$  the flow sensitivity is larger than the ILC target; however, as the repeatability of  $\delta^{13}\text{CO}_2$  measurements is still on the order  $\pm 0.03$  ‰ this result is not crucial for the instrument performance. Nevertheless, we apply the flow rate sensitivity correction to our data.

### 3.6 H<sub>2</sub>O cross sensitivity

Measuring the H<sub>2</sub>O cross sensitivity is experimentally challenging and time consuming since it takes time until the moisture equilibrium between gas-phase and instrument surfaces is established. The drying cartridge in the mixing device acts as an additional resistance and it was not until our last experiments that we introduced a needle valve in the direct branch of the mixing device to counteract this resistance. Thus, for the first experiments the total resistance of the mixing device changed according to flow proportion in the course of the H<sub>2</sub>O cross sensitivity experiment. This lead to a variation in cell-pressure on the order of 10hPa during the experiment. In the following all results have been corrected for the residual pressure sensitivity before the H<sub>2</sub>O cross sensitivity was determined.

A further challenge is a species-dependent hysteresis effect in the H<sub>2</sub>O cross sensitivity. Figure 4 shows the H<sub>2</sub>O cross sensitivity for N<sub>2</sub>O and the very pronounced hysteresis effect. In contrast to the dedicated H<sub>2</sub>O cross sensitivity experiments performed later on, this test was conducted with gradually changing cell moisture. The H<sub>2</sub>O cross sensitivity is linear while humidifying the cell (from red to yellow in Figure 4), whereas a clear hysteresis appears during drying (from green to blue). The origin of the H<sub>2</sub>O hysteresis is believed to be related to surface effects in the instrument, but was not further investigated since such large moisture variations do not occur in our standard operation conditions. They may, however, become relevant, if also water and water isotopologues are to be measured with the instrument.

For the five dedicated H<sub>2</sub>O cross sensitivity experiments the investigated H<sub>2</sub>O range was restricted to moisture levels between 2 and 10  $\mu\text{mol mol}^{-1}$ , i.e. those which may occur during the lifetime (typically 2 months) of one  $\text{Mg}(\text{ClO}_4)_2$  drying cartridge. Individual H<sub>2</sub>O

concentrations were kept as constant as possible for each set point to minimise transient effects. Within the measurement precision, no significant H<sub>2</sub>O cross sensitivity was observed for N<sub>2</sub>O and δ<sup>13</sup>CO<sub>2</sub>. This is explained by the weak H<sub>2</sub>O absorption in the spectral region of N<sub>2</sub>O, and for δ<sup>13</sup>CO<sub>2</sub> the H<sub>2</sub>O cross sensitivity is low since all CO<sub>2</sub> isotopologues are subject to a similar H<sub>2</sub>O cross sensitivity in this H<sub>2</sub>O range. Although some experiments did show a clear relation between H<sub>2</sub>O level and CH<sub>4</sub> or CO mole fractions, no unequivocal cross sensitivity could be determined for these two trace gases. Only for CO<sub>2</sub> a more or less temporally stable H<sub>2</sub>O cross sensitivity was found with a 1σ standard deviation of ± 50%. The unstable results are most likely due to experimental problems. Precise adjustment and stability of moisture levels in the range between 2-10 μmol mol<sup>-1</sup> is difficult to achieve and in the initial set-up of the mixing device small H<sub>2</sub>O cross sensitivities have been superimposed by cell pressure variations. Further experiments with additional methods to stabilise cell pressure were needed and later on performed with the improved version of the in situ FTIR analyser. However, since the variations in the residual moisture level during standard operating conditions can be restricted to less than ± 2 μmol mol<sup>-1</sup> the H<sub>2</sub>O cross sensitivity is not very important. In Table 1 the H<sub>2</sub>O effect is shown in a calculation for a 2 μmol mol<sup>-1</sup> effect, which exemplifies that for all species the effect is smaller than the ILC target. Nevertheless for CO<sub>2</sub> the H<sub>2</sub>O cross sensitivity is larger than the instrument's repeatability and thus worth correcting. The CO<sub>2</sub> variability associated with the H<sub>2</sub>O cross sensitivity could at least be halved by applying the H<sub>2</sub>O correction.

### 3.7 CO<sub>2</sub> cross sensitivity

<sup>12</sup>CO<sub>2</sub> is the strongest absorber in the spectral range between 2150 and 2320 cm<sup>-1</sup> and its absorption ranges do partly overlap with those of <sup>13</sup>CO<sub>2</sub>, N<sub>2</sub>O and CO analysed in the same spectral range. Since MALT uses a broad spectral region for fitting, it is able to distinguish the different species; however certain cross sensitivities to <sup>12</sup>CO<sub>2</sub> remain for all mentioned species. For CH<sub>4</sub> the CO<sub>2</sub> cross sensitivity effect is negligible, since the CH<sub>4</sub> concentration is derived at 3001-3150 cm<sup>-1</sup> where no significant <sup>12</sup>CO<sub>2</sub> absorption occurs. In contrast to all previously discussed sensitivities, the influence of the CO<sub>2</sub> cross sensitivity cannot be reduced by minimising the variability of the causing agent. Thus, precise determination of the CO<sub>2</sub> cross sensitivity is vital.

To vary the CO<sub>2</sub> amount we used the mixing device described above with an Ascarite (Sigma Aldrich, USA) filled cartridge. Ascarite has shown to be suitable to remove CO<sub>2</sub> without altering the mole fractions of the other investigated greenhouse gases (Glatzel-Mattheier, 1997). The H<sub>2</sub>O that is produced by Ascarite during the CO<sub>2</sub> uptake is removed by the drying system. In order to use the results of the CO<sub>2</sub> cross sensitivity experiment for δ<sup>13</sup>CO<sub>2</sub> as well, it is crucial to remove the CO<sub>2</sub> entirely in the Ascarite branch of the mixing device to avoid isotope fractionation. This was verified by taking aliquot flask samples for each CO<sub>2</sub> level and analysing δ<sup>13</sup>CO<sub>2</sub> by mass spectrometry. For our experimental conditions with CO<sub>2</sub> amounts as high as 800 μmol mol<sup>-1</sup> and a flow rate of up to 1 slpm, 80g of Ascarite are sufficient to completely remove CO<sub>2</sub> from the Ascarite branch over the 10h duration of the experiment.

In total seven cross sensitivity experiments were conducted. The five initial experiments spanned approximately the ambient CO<sub>2</sub> range (340 to 440 μmol mol<sup>-1</sup>). The later experiment used spiked CO<sub>2</sub> concentrations in order to investigate a wider CO<sub>2</sub> concentration range. Figure 5 displays results for both ranges of the CO<sub>2</sub> cross sensitivity for the worst-case example of N<sub>2</sub>O. Both experiments are in accordance with each other, although they were performed several months apart. However, the wider CO<sub>2</sub> range reveals further details about the shape of the CO<sub>2</sub> cross sensitivity for N<sub>2</sub>O. The wide range of the CO<sub>2</sub> cross sensitivity can be adequately described by a cubic relation, with a correlation coefficient r<sup>2</sup>>0.99. The typical range of unpolluted ambient air (370 to 420 μmol mol<sup>-1</sup>) can be approximated linearly with less than 0.05 nmol mol<sup>-1</sup> N<sub>2</sub>O deviation. However, for larger CO<sub>2</sub> values the deviation between the two approximations becomes substantial. The sensitivity of N<sub>2</sub>O to CO<sub>2</sub> can be reduced to be negligible by selection of a narrower spectral range for N<sub>2</sub>O analysis excluding severe overlap with CO<sub>2</sub>, as described in Griffith et al. (2012).

The CO<sub>2</sub> cross sensitivities for CO and δ<sup>13</sup>CO<sub>2</sub> in the wider range of CO<sub>2</sub> are also not linear. Similarly to the case for N<sub>2</sub>O, the cross sensitivities for both species can be approximated linearly in the unpolluted CO<sub>2</sub> range. For CO the CO<sub>2</sub> cross sensitivity starts deviating significantly from a linear relation for CO<sub>2</sub> values above 500 μmol mol<sup>-1</sup>. In Table 1 the magnitude of the CO<sub>2</sub> cross sensitivities are summarised for the linearly approximated unpolluted CO<sub>2</sub> range. δ<sup>13</sup>CO<sub>2</sub> actually depends inversely on CO<sub>2</sub> and is dealt with explicitly in Griffith et al. (2012). For N<sub>2</sub>O and δ<sup>13</sup>CO<sub>2</sub> the effect of the CO<sub>2</sub> cross sensitivity is by far larger than the ILC targets, whereas for CO as well as for CH<sub>4</sub> the CO<sub>2</sub> cross sensitivity is

smaller. The Heidelberg FTIR data post-processing includes the non linear CO<sub>2</sub> cross sensitivity corrections for all species, also in the linear range.

#### 4 Instrument Response Function

As the raw *absolute* mole fraction determination of the in situ FTIR analyser differs from the internationally accepted WMO scales by up to a few percent, depending on species (Griffith et al., 2010, 2012), calibration of the FTIR analyser with internationally accepted standard reference material is necessary to achieve the required accuracy and comparability for ambient air monitoring. The shape of the Instrument Response Function (IRF) determines the number of required calibration standards. To determine the IRF of the in situ FTIR analyser we analysed the IUP-Heidelberg set of primary laboratory standards calibrated by the WMO Central Calibration Laboratories (CCL). We have 13 laboratory standards, which have been calibrated for CO<sub>2</sub> on the X2007 mole fraction scale. Seven of these cylinders have also been calibrated for N<sub>2</sub>O and five for CH<sub>4</sub> by NOAA/ESRL. NOAA/ESRL acts as the WMO-GAW Central Calibration Laboratory (CCL) for all these trace gases. For  $\delta^{13}\text{CO}_2$ , we received four reference standards that were calibrated on the VBDP/j-RAS06 scale by the Max-Planck Institute for Biogeochemistry (MPI-BGC) in Jena, being the WMO-GAW Central Calibration Laboratory for stable isotopes in CO<sub>2</sub>. The same cylinders were calibrated for the CO mole fractions by the MPI-BGC GasLab on the NOAA/ESRL 2004 scale. The calibrated range for each species is given in Figure 6 and its caption.

All FTIR measurements of these calibration cylinders were corrected for residual and cross sensitivities as described in the previous sections; measurements were performed on three consecutive days. Linear regression of the sensitivity-corrected FTIR mole fractions against the assigned cylinder reference values showed no significant curvature but have a significant non-zero intercept. In Figure 6 the residuals to the linear fit confirm that the assumption of linear instrument response functions is justified for all species. The error bars in Figure 6 depict the combined error of the FTIR measurement and the error of the calibration cylinder assigned mole fractions. The standard deviations of the residuals are 0.03  $\mu\text{mol mol}^{-1}$  for CO<sub>2</sub>, 0.04 ‰ for  $\delta^{13}\text{CO}_2$ , 0.4 nmol mol<sup>-1</sup> for CO, 0.4 nmol mol<sup>-1</sup> for CH<sub>4</sub> and 0.08 nmol mol<sup>-1</sup> for N<sub>2</sub>O. From these results we can conclude that the IRFs are linear over the investigated range for all tracers. However, the non-zero intercepts of the regression functions prevent a simple

one point calibration, leading to a minimum of two calibration standards for the FTIR. For high accuracy demands, as in atmospheric background monitoring programs, the usage of three calibration standards to define the IRF is advised. A three point calibration reduces the sensitivity to individual outliers and instrument noise in the calibration measurements.

## 5 Temporal stability of the FTIR analyser and calibration frequency

In order to assess the precision of the in situ FTIR analyser, measurement repeatability as well as reproducibility on different time scales has to be determined. In the following, we will use the terminology related to the GAW Glossary of QA/QC (Klausen and Scheel, 2007). The issue of measurement stability is vital for all long term monitoring efforts and determines the required calibration frequency to reach a certain level of reproducibility. In the following we will subdivide the stability requirement into two parts, short term and long term. ‘Short term’ handles drifts and other artefacts on a time scale from hours to days, while ‘long term’ covers weeks to months.

### 5.1 Short term stability

To quantify short term drifts of the FTIR analyser, a target gas was continuously flushed through the cell at the standard flow rate of 1 slpm for six days, comprising a weekend as well as weekdays, to cover different laboratory conditions. Absorption spectra with an averaging time of one minute were recorded and analysed offline to avoid down time. The respective time series of the dry air mole fractions, calculated according to Eq. 2, are shown in Figure 7 in grey. The sensitivity corrected results are shown as well (coloured symbols).

This short term stability test was conducted in September 2011 for the Intermediate setup. Thus, the sample properties in the cell remained fairly constant over the course of the experiment as expressed by their  $1\sigma$  standard deviations:  $\pm 0.01$  °C for cell temperature,  $\pm 0.01$  hPa for cell pressure,  $\pm 0.2$   $\mu\text{mol mol}^{-1}$  for moisture level, and  $\pm 0.01$  slpm for sample flow rate. The Heidelberg laboratory temperature oscillated between 23.2°C and 25.1°C for this period, however, the influence on cell temperature was small. The generally very stable sample properties in the cell lead to only small corrections of the residual sensitivities (Figure 7). Only for CO<sub>2</sub>, a clear correlation to cell temperature was found. The RTS corrections thus

1 improve the CO<sub>2</sub> standard deviation of the 1-minute measurements from 0.037 to 0.029 μmol  
2 mol<sup>-1</sup>, for all other species the sensitivity corrections are marginal.

3 The initial 6 hours of the test show a small settling-in effect for CO<sub>2</sub> as well as for CO. Most  
4 likely this can be attributed to adjustments of the pressure regulator of the cylinder and the  
5 cell parameters. δ<sup>13</sup>CO<sub>2</sub> and CH<sub>4</sub> show a slight trend over the six days, persisting even after  
6 the sensitivity correction with 0.02 ‰ day<sup>-1</sup> for δ<sup>13</sup>CO<sub>2</sub> and -0.04 nmol mol<sup>-1</sup> day<sup>-1</sup> for CH<sub>4</sub>.  
7 The drift in δ<sup>13</sup>CO<sub>2</sub> might be related to a fractionation effect in the fast emptying high pressure  
8 cylinder, however, this cannot be proven since no pre and post mass spectrometer  
9 measurements have been performed on this cylinder.

10 The same dataset from September 2011 (Figure 7) can be used to determine the repeatability  
11 of the FTIR measurements using Allan variance analysis (Werle et al., 1993). In Table 2, the  
12 1σ repeatability for 2.5 and 10 min averaging time is given. 2.5 min averaging time is used in  
13 standard operating conditions for the Heidelberg in situ FTIR (3 min measurements including  
14 30 sec online analysis time). The 10 min repeatability is given for reasons of comparability to  
15 the earlier results from Griffith et al. (2010, 2012). Similar to the findings by Griffith et al.  
16 (2010, 2012), the repeatability of all species except for CO<sub>2</sub> initially improves with the square  
17 root of averaging time, although in our experiments this is only valid for the first 40 min and  
18 not for 1-2 hours as found by Griffith et al. (2010, 2012). For longer integration intervals the  
19 repeatability still improves, but at a slightly lower rate. For sensitivity-corrected CO<sub>2</sub>, the  
20 repeatability improves steadily up to an averaging time of 15 hours, but only by one third of  
21 the square root of time.

22 We determined the reproducibility for measurements averaged over three minutes, as given in  
23 Table 2, by pooling three one minute spectra and calculated their 1σ standard deviation. This  
24 reproducibility includes any potential changes or arbitrary drifts in the spectrometer or any  
25 sensor over the investigated six days period. The comparison of the repeatability and  
26 reproducibility of the three minute averages in Table 2 emphasises the remarkably good short  
27 term stability of the in situ FTIR.

## 29 5.2 Long term stability of the in situ FTIR analyser

30 The long term stability of the instrument response function (IRF) determines the calibration  
31 frequency as well as the calibration strategy. If the IRF is absolutely stable in time,



interpolation between repeated calibrations does not improve measurement accuracy. This is due to the intrinsic uncertainty of each calibration measurement, which is then passed on to the measurements and increases their noise. If, on the other hand, the temporal changes of the IRF are larger than the instrument's repeatability, regular calibration improves the accuracy substantially. In the latter case the question concerning the required calibration frequency arises. In section 5.1 we showed that, compared to the noise, for CO<sub>2</sub>, CO and N<sub>2</sub>O no systematic change of residual pressure and temperature sensitivity-corrected data occurred over the timescale of up to six days. Here we will investigate the longer time scales, based on the data set we collected during the ICOS Demonstration Experiment. During these field campaigns, 24- and 48-hourly calibrations were performed with two standard cylinders, covering a suitable mole fraction range for all five components. In between the calibrations, a so-called target or surveillance tank was measured for quality control. In terms of sensitivity correction and calibration the target cylinder was analysed in a similar way as any unknown sample from a cylinder. In the next section we will revisit the influence of the residual and cross sensitivities discussed in section 3 on the example of the CO<sub>2</sub> long term target record.

### 5.2.1 Influence of the CO<sub>2</sub> residual and cross sensitivities on long term records

Figure 8 displays the deviations of the CO<sub>2</sub> target measurement from the mean mole fraction over the different evaluation stages from the raw FTIR measurements to the finally calibrated data. Each target gas value consists of the average and the standard deviation of five sequentially recorded three minute spectra. The dry-air mole fractions as calculated by MALT are shown in blue; they exhibit step changes as well as gradual changes. The 1 $\sigma$  standard deviation for the CO<sub>2</sub> target gas is 0.25  $\mu\text{mol mol}^{-1}$  with a peak to peak variability of 0.8  $\mu\text{mol mol}^{-1}$  over this five months measurement period. The prominent changes in the dry-air mole fraction can be related to changes in cell temperature and pressure, caused by changing laboratory conditions and/or a degradation of the flow controller, that was additionally installed during the ICOS Demonstration Experiment. Correction for residual and cross sensitivities determined in section 3 improves the reproducibility of the 24- to 48-hourly values as well as the 1 $\sigma$  scatter over the whole period by a factor of two to only 0.11  $\mu\text{mol mol}^{-1}$  (red dots in Figure 8). The residual and cross sensitivity correction takes care of most of the pronounced step changes, however some outliers are persistent (e.g. in mid of July) and can thus not be explained by a change in one of our investigated residual and cross



sensitivities. The variability in the sensitivity-corrected target gas measurements is still larger than the observed short term repeatability expressed by the error bars and investigated in section 3. In addition, the CO<sub>2</sub> variability still shows long term changes and outliers. Thus, applying regular, e.g. daily, calibration might help to reduce the CO<sub>2</sub> variability further. By using a linear interpolation between the 24- to 48-hourly calibration measurements to determine the IRF, the 1 $\sigma$  standard deviation of the CO<sub>2</sub> target measurements is reduced to 0.05  $\mu\text{mol mol}^{-1}$  as shown by black squares in Figure 8.

Although the calibrated measurements are generally satisfying the WMO compatibility goal of  $\pm 0.1 \mu\text{mol mol}^{-1}$  for CO<sub>2</sub> measurements in the Northern Hemisphere, individual measurement periods show much more scatter (e.g. in July). **Outliers in the target gas record (July 2011) are caused by bad calibration measurements.** The impact of a single calibration measurement depends largely on the strategy chosen to derive the IRF. We therefore investigated different calibration strategies such as a) averaged IRF, b) interpolated IRF between smoothed calibration measurements and c) interpolated IRF between neighbouring calibration measurements. Based on the repeatability of the target cylinder measurements, strategy c) yielded the best results, although its sensitivity to single bad calibration measurements is largest. Careful selection of calibration outliers is thus recommended to minimise artefacts in the time series.

## 5.2.2 Long term reproducibility for all species

Figure 9 comprises the sensitivity-corrected and calibrated target gas measurements for all components. In September 2011 the FTIR system was modified **to the Intermediate setup**. The long term reproducibility of the **ICOS Demonstration Experiment setup** can be judged based on the 1 $\sigma$  standard deviations up to September 2011: 0.06  $\mu\text{mol mol}^{-1}$  for CO<sub>2</sub>, 0.05 ‰ for  $\delta^{13}\text{CO}_2$ , 0.45  $\text{nmol mol}^{-1}$  for CO, 0.28  $\text{nmol mol}^{-1}$  for CH<sub>4</sub> and 0.1  $\text{nmol mol}^{-1}$  for N<sub>2</sub>O. **The 1 $\sigma$  standard deviations are thus** close to or within the ILC targets for all trace gases (WMO, 2011). For  $\delta^{13}\text{CO}_2$  the reproducibility of 0.05‰ is acceptable, keeping in mind that the in situ FTIR analyser is one of the first instruments delivering continuous  $\delta^{13}\text{CO}_2$  measurements. The WMO-GAW requested target compatibility **for  $\delta^{13}\text{CO}_2$**  is 0.01‰, and yet only met by very few mass spectrometer laboratories (Huang et al., 2011).

No significant drift was observed for any species. The small step change in CO is not explained by any of the investigated sensitivities. Re-calibration of the working standards as well as the target tank at the Max-Planck Institute for Biogeochemistry confirmed that neither the target cylinder nor the calibration cylinders did drift in any of the investigated species.

For CO<sub>2</sub> it is obvious that the performance of the in situ FTIR analyser declined after middle of June 2011. Checking the FTIR parameters revealed that from middle of June onwards the standard deviation of the sample pressure increased from better than  $\pm 2$  hPa to more than  $\pm 7$  hPa. This increase was most likely caused by a progressive degradation of the additionally installed external mass flow controller (blue MFC in Figure 1). In section 3.3 we have shown that the largest pressure sensitivity corrections are found for CO<sub>2</sub>; consequently, we observe the largest deviation between the short and long term reproducibility for CO<sub>2</sub>. The decrease in reproducibility for CO is caused by the step change on June 15<sup>th</sup> (compare Figure 9) and for N<sub>2</sub>O by several outliers.

The first red shaded area in Figure 9 marks the re-building of the instrument to the Intermediate setup. The reproducibility derived with this instrument configuration is given in Table 2. The largest improvement compared to the ICOS Demonstration Experiment setup, a factor of 2, was achieved for N<sub>2</sub>O, improving the in situ FTIR analyser precision to well below the ILC target. The performance for CO<sub>2</sub>,  $\delta^{13}\text{CO}_2$  and CH<sub>4</sub> remains at a comparable good level. The drift, which was observed for CO is caused by a drifting calibration gas cylinder being used during this time. Un-calibrated results suggest that the CO reproducibility improved as well.

To decouple sample pressure and flow, and to ease sample handling, an additional electronic pressure controller (EPC) (El-Press, Bronkhorst, The Netherlands) was installed at the inlet of the cell in the IUP setup. In Figure 9 the target gas measurements performed with this new configuration are shown after the second red shaded area. The reproducibility of the sensitivity-corrected and -calibrated target gas results is also summarised in Table 2. The introduction of the additional EPC improved the reproducibility of all components except for N<sub>2</sub>O by approximately a factor of two. The dashed line in Figure 9 in mid December 2011 indicates a change in our standard operating conditions. From there on we skipped the evacuation step during sample exchange and used the flushing approach (refer to Appendix A). This was done in order to investigate the benefits of not disrupting the temperature and moisture equilibrium in the cell by the evacuation. In Figure 9 no significant difference

between both sample change-over strategies is visible. Thus after two months we changed back to our SOC to save calibration gas and to shorten the calibration measurement interval. This finding depends on the temperature sensor type used. It will most likely not be valid for the slow responding RTD temperature sensor.

In the final IUP setup short and long term reproducibility are similar for all components (see Table 2). These results demonstrate that the long term stability of the in situ FTIR analyser is suitable for background air monitoring if measurements are calibrated and properly corrected for residual and cross sensitivities.

### 5.3 Calibration frequency

The target gas record can further be used to estimate the influence of the calibration frequency on the reproducibility. In Figure 9 we used the highest calibration frequency of 24- to 48-hourly calibrations with two standard gases. With the target gas data set we can investigate the influence of stepwise prolonged calibration intervals on the  $1\sigma$  reproducibility of the results. Since, especially for CO<sub>2</sub>, the performance of the FTIR analyser can be separated into two periods, stable and less stable conditions (compare gray shaded area in Fig. 9), we will evaluate the calibration frequency test for both stability conditions separately. The less stable period, without proper control of the additional MFC introduced in the ICOS Demonstration Experiment setup, the instrument configuration is comparable to the initial UoW setup.

In each panel of Figure 10 the relation between  $1\sigma$  reproducibility and calibration frequency is plotted. Closed symbols represent stable instrument conditions with a cell pressure variability of less than  $\pm 2$  hPa, open symbols show less stable conditions; here, the cell pressure variability was four times larger. Under stable instrument conditions the  $1\sigma$  target reproducibility of CO<sub>2</sub> was better than  $0.03 \mu\text{mol mol}^{-1}$  for 24- to 48-hourly calibrations. Prolonging the calibration frequency even to 2 or 3 months changes the reproducibility only slightly to  $0.05 \mu\text{mol mol}^{-1}$ . During less stable instrument conditions the reproducibility improves steadily with increasing calibration frequency, however, even 24- to 48-hourly calibrations do not result in the reproducibility of the stable conditions. It seems that under less stable conditions substantial variations occur on sub-daily time scales, for example during the campaign in OPE the laboratory temperature changed by ten degrees within one day.

The other trace species, apart from N<sub>2</sub>O, show a similar behaviour to that of CO<sub>2</sub>, although the difference between stable and less stable instrument conditions is not as pronounced. For

1 the highest calibration frequency the achieved reproducibility is on par or even better than the  
2 short term reproducibility determined in section 5.1. This is caused by the fact that the  
3 investigated target measurements as well as the calibration measurements are averaged values  
4 of five individual spectra reducing the scatter by a factor of  $\sqrt{5}$ . This implies that for a single 3  
5 min ambient air measurement the reproducibility is a factor of  $\sqrt{5}$  larger than the one shown  
6 for the target measurements. In contrast to the IRF of other gases, the IRF for N<sub>2</sub>O has two  
7 large step changes during the stable instrument conditions (not shown) leading to the inverse  
8 results in the calibration frequency investigation. The recorded sample and instrument  
9 properties give no evidence that could explain the step changes in the N<sub>2</sub>O IRF.

10 Increasing the calibration frequency from weekly to 24- or 48-hourly improves the  
11 reproducibility on average by 20% for all gases. Depending on the specific accuracy needs,  
12 weekly calibrations should be sufficient and beneficial for calibration gas consumption as  
13 well.

1

## 2    **6    Error assessment for ambient air measurement**

3    Different sources of error contribute to the total error of an individual ambient air  
4    measurement. These error contributions are partly systematic and partly random and can often  
5    not easily be separated. Obvious error contribution arises from: a) measurement repeatability,  
6    b) uncertainties in residual and cross sensitivity corrections, c) uncertainties introduced by the  
7    two-point calibration and d) uncertainty of the assigned mole fraction of the calibration  
8    cylinders themselves.

9    These error sources are not independent from each other. For example, the uncertainties  
10    originating from a) and b) (repeatability and the residual and cross sensitivity corrections)  
11    feed directly into the uncertainty introduced by the regular calibration. Thus, a combined  
12    uncertainty comprising a), b) and c) can be derived from the reproducibility of the target  
13    measurements as listed in Table 2. However, the target measurement reproducibility is  
14    derived from averaged values and has thus to be multiplied by a factor of 2. For ambient air  
15    measurements the uncertainties of the CO<sub>2</sub> cross sensitivities must be added, since this error  
16    component is not present in the target measurements. The total uncertainties are thus given in  
17    the last column of Table 2. The contribution of the uncertainties of the assigned values of the  
18    calibration cylinders is an additional, systematic error component **that** has to be accounted for  
19    separately.

20

## 21    **7    Conclusions and outlook**

22    The **modified** in situ FTIR spectrometer is well suited for GHG monitoring and fulfils the  
23    precision, accuracy and stability needs for CO<sub>2</sub>, CO, CH<sub>4</sub> and N<sub>2</sub>O measurements at  
24    background sites. The reproducibility for  $\delta^{13}\text{CO}_2$  does not formally fulfil the ILC targets set  
25    by WMO-GAW; however, 0.03‰ target reproducibility (0.07‰ single measurement) is still  
26    very good for a continuously measuring instrument. **The final IUP setup of the in situ FTIR is**  
27    **additionally equipped with an internal EPC and MFC to control sample pressure and flow**  
28    **rate. To improve the temperature measurement the RTD sensor was replaced by a**  
29    **thermocouple that is located in the middle of the cell to improve the representativeness of the**

1 measurement. These improvements lead to a factor of two better repeatability for all species.  
2 In the IUP setup the effects of the residual pressure sensitivity and flow rate-sensitivity are  
3 negligible. The same is true for the H<sub>2</sub>O cross sensitivity, since the residual moisture  
4 variations can be restricted to **only** a few  $\mu\text{mol mol}^{-1}$  of H<sub>2</sub>O. Reducing the causes of  
5 variability is always superior to any post-processing correction. However, for those  
6 parameters where this is not possible, e.g. CO<sub>2</sub> **cross sensitivity**, precise determination of these  
7 parameters is essential. Currently, **the** largest remaining issue **is** related to the ability of  
8 measuring the **true** sample temperature in the cell. **Resolving this temperature problem might**  
9 **allow for measuring calibration and target gas cylinders in static mode, which would reduce**  
10 **the gas consumption by a factor of five. Measuring the sample temperature at multiple**  
11 **locations could be a first step to improve the representativeness of the temperature**  
12 **measurement.** The small Reynolds number in the cell ( $\approx 10$ , **see Appendix A**) for our standard  
13 operating conditions constitutes a more general problem. Improved sample delivery to the cell  
14 **that** increases turbulence would help minimizing temperature gradients within the cell. We  
15 also recommend placing the Nafion drier inside the temperature controlled enclosure, since it  
16 is important for thermal sample pre-conditioning.

17 The instrument has proven to be acceptably linear for all components in the ambient  
18 concentration range (also for non background conditions), however, the instrument response  
19 functions have a non-zero offset for all components, implying the need of at least two  
20 calibration gases. The in situ FTIR analyser is sufficiently stable to run it with only weekly  
21 calibrations. In standard operation conditions, **almost** 30 litres of air are needed for one  
22 calibration measurement. Thus, a 50 litre cylinder pressurised to 20 MPa lasts for more than  
23 four years, even if 10% to 20% of the gas is **remaining in the cylinder** to avoid potential drifts  
24 **of components such as CO<sub>2</sub>** (Kitzis, 2009; Langenfelds et al., 2005). **Although the lifetime of**  
25 **a FTIR calibration gas cylinder is thus longer than that for classical GC systems (typical**  
26 **lifetime  $\approx 1$  year) it is shorter than the expected calibration gas cylinder lifetime for other**  
27 **optical techniques like Quantum cascade Lasers (QCL) or CRDS. For the laser techniques the**  
28 **calibration cylinder lifetime is expected to be on the order of decades, depending on the**  
29 **calibration scheme (e.g. Winderlich et al., 2010).** For conservative, high accuracy  
30 applications, one may want to have a full year of overlap between two calibration gas  
31 generations, resulting in an effective calibration gas lifetime of three years. For high accuracy  
32 applications we recommend **the use of** three calibration cylinders, spanning the range of  
33 expected ambient concentrations. The usage of three calibration gases instead of the

1 minimally needed two reduces the sensitivity to individual outliers and instrument noise. In  
2 any case, we recommend daily target or surveillance gas measurements for quality control.  
3 The lifetime of a daily target gas is, unfortunately, limited to nine months only, thus we  
4 recommend a second sub-target being measured on a two weekly basis only. The sub-target  
5 can then also be used to inter-connect two to three standard gas generations.

6 Comparing the in situ FTIR to other optical state-of-the-art greenhouse gas analysers like  
7 Cavity Ring Down Spectroscopy, Off-Axis Integrated Cavity Output Spectroscopy or  
8 Quantum Cascade Lasers shows that the performance of the in situ FTIR is comparable to all  
9 other techniques for all five species. Currently, no other technique is able to provide  
10 measurements of CO<sub>2</sub>,  $\delta^{13}\text{C}$ , CO, CH<sub>4</sub> and N<sub>2</sub>O with a single instrument. The gas  
11 consumption and the sample change-over time of all other techniques is, however smaller.  
12 This is beneficial for the conservation of calibration gas as well as fast sample exchanges e.g.  
13 at atmospheric tower sites with multiple inlet heights measured with one single instrument.  
14 Residual and cross sensitivities are a general issue for all optical GHG analysers and are  
15 subject to ongoing investigation (e.g. Rella et al., 2012, Zellweger et al., 2012, Vogel et al.,  
16 2012). Providing a comprehensive comparison of the different techniques for all species is,  
17 however, beyond the scope of this paper. A comparison study of recent CO measurement  
18 techniques can be found by Zellweger et al. (2012), for the performance of other species and  
19 techniques, please refer to individual papers e.g. Winderlich et al. (2010) for CO<sub>2</sub> and CH<sub>4</sub>,  
20 McManus et al. (2008) for N<sub>2</sub>O, and Vogel et al. (2012) for  $\delta^{13}\text{C}$ .

## Acknowledgements

Martin Riggenbach and Graham Kettlewell are gratefully acknowledged for building and improving the hardware and software of the in situ FTIR analyser and for many hours of sharing ideas with the authors. We are very grateful to Armin Jordan and Bert Steinberg for the preparation and careful calibration of the high pressure cylinders used during the ICOS Demonstration Experiment. We wish to thank Alex Vermeulen (ECN) and the team from KNMI for hosting us at Cabauw tower, and Sébastien Conil (Andra) and Marc Delmotte (LSCE) for the support at OPE. We thank Vanessa Sherlock and the two anonymous reviewers for their comments that helped to improve this paper substantially. This work has been funded by the European Commission under the ICOS Preparatory Phase Project (INFRA-2007-211574) and additionally supported by the “Global Networks” program of the University of Heidelberg, and International Science Linkages/Europe Fund grant (EF-090007) and Australian Research Council Discovery Grant (DP0879468) from the Australian Government.



## Appendix A: Sample exchange and thermodynamic equilibrium in the cell

Measurement accuracy depends on the instrument precision as well as on the complete exchange of the gas sample in the cell without memory effects. Thus, the cell has to be either evacuated or flushed for a sufficiently long time to completely remove the previous sample. We will briefly discuss both approaches:

a) Flushing: The mean exchange time  $\tau$  of a well-mixed volume is given by the ratio of volume to flow rate. The influence of the preceding sample decreases according to  $\exp(-t/\tau)$  where  $t$  is the flushing time. Thus, we need more than  $t = 8 \tau$  to reduce the memory effect to less than 0.03 %, which in case of a  $100 \mu\text{mol mol}^{-1}$  change in  $\text{CO}_2$  between the two samples results in a  $0.03 \mu\text{mol mol}^{-1}$  memory effect. With the given cell volume and the standard flow rate of 1 slpm, a flushing time of 28 min is thus theoretically required. An example of this change-over method is illustrated in Figure A1 (black squares). In this example the equilibrium value, which is defined as the average mole fraction measured from minute 30 to 45 is reached after 21 minutes, with a mole fraction difference between the two samples of  $140 \mu\text{mol mol}^{-1}$ .

b) Evacuation: To reach comparably small memory effects as under a), the cell and the inlet system need to be evacuated to  $<0.3 \text{ hPa}$ , which, with the available equipment, is not reached within 30 min. Thus, we decided to use a two step evacuation procedure: The cell is first evacuated to 10 hPa, then filled with the new sample to 500 hPa, evacuated a second time down to 10 hPa and subsequently filled to the desired cell pressure of 1100 hPa. This stepwise change-over is much faster than a one-step evacuation and ensures that the memory effect is less than 0.02 %. The complete sample change-over procedure requires 7 litres of gas and takes 8 minutes, including one minute stabilisation time after the final pressure and flow settings are reached. An example of the two-step evacuation procedure is also shown in Figure A1 (blue triangles).

Using the two-step evacuation procedure for sample exchange, close to equilibrium values are already observed 12 min after a sample change-over. As evacuation and re-filling affects the temperature of the cell, the first two measurements (taken in minutes: 6 to 9 and 9 to 12) have to be discarded.

1 However, exchanging the sample entirely, and thus avoiding memory effects, is not  
2 necessarily sufficient to avoid transient settling-in effects after sample change-over. The  
3 thermodynamic properties of the sample, like temperature and pressure, need to reach their  
4 equilibrium conditions as well. Sample temperature is especially crucial, since, in contrast to  
5 pressure, we cannot assume that temperature is homogeneously distributed within the cell.  
6 This may have several causes, a) different sample temperatures when entering the cell, b) the  
7 spectrometer heats one end of the cell causing a small but significant temperature gradient  
8 along the cell, c) the Reynolds number of the cell calculated for our SOC is only about 10;  
9 thus, mixing is far from turbulent, leading to a persistence of any temperature gradient. For a  
10 tube geometry the Reynolds number can be calculated as:  $Re = (u_m d)/\nu$ , with  $u_m$ : mean  
11 velocity of the gas ( $1 \cdot 10^{-3} \text{ m/s}$ ) in the cell,  $d$ : diameter of the cell (0.15m) and  $\nu$ : kinematic  
12 viscosity of air ( $1.5 \cdot 10^{-5} \text{ m}^2/\text{s}$ ).

13 The impact of thermodynamic disequilibrium is shown in Figure A1 by the red symbols.  
14 These measurements have been performed using the same cylinder, regulator and evacuation  
15 procedures as for the data indicated in blue, however, the cell was rotated by  $180^\circ$ . Turning  
16 the cell upside-down changed the position of the temperature sensor as well as that of the  $\frac{1}{4}$ "  
17 dip-tube that delivers the sample into the cell. For the measurements plotted in red the dip  
18 tube was at the top of the cell, and the temperature sensor was at the bottom. We can assume  
19 that the memory effect caused by incomplete sample exchange is similar for both positions,  
20 since the same two-step evacuation approach was used, but due to the different location of the  
21 sample inlet and/or the temperature sensor the sample apparently takes longer until it reaches  
22 its thermodynamic equilibrium. Since this settling-in effect is seen for all components, it  
23 seems likely that it might be related to temperature. The measured sample temperature  
24 between the two cell positions changed by  $0.5^\circ\text{C}$ ; higher temperatures were observed with the  
25 temperature sensor in the upper position. Thus we can conclude that temperature is not  
26 homogeneously distributed in the cell, and that we are not able to measure the true mean  
27 sample temperature with one temperature sensor only. However as long as the temperature  
28 distribution in the cell is stable under standard operation conditions and is reached for both,  
29 ambient air and cylinder measurements, the calibration of the instrument will compensate for  
30 the error in temperature measurement.



## References

- Chen, H., Winderlich, J., Gerbig, C., Hofer, A., Rella, C. W., Crosson, E. R., Van Pelt, A. D., Steinbach, J., Kolle, O., Beck, V., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Santoni, G. W., and Wofsy, S. C.: High-accuracy continuous airborne measurements of greenhouse gases ( $\text{CO}_2$  and  $\text{CH}_4$ ) using the cavity ring-down spectroscopy (CRDS) technique, *Atmos. Meas. Tech.*, 3, 375-386, 2010.
- Esler, M. B., Griffith, D. W. T., Wilson, S. R., and Steele, L. P.: Precision trace gas analysis by FT-IR spectroscopy 1. Simultaneous analysis of  $\text{CO}_2$ ,  $\text{CH}_4$ ,  $\text{N}_2\text{O}$  and CO in air, *Analytical Chemistry*, 72(1), 206-215, 2000a.
- Esler, M. B., Griffith, D. W. T., Wilson, S. R., and Steele, L. P.: Precision trace gas analysis by FT-IR spectroscopy 2. The  $^{13}\text{C}/^{12}\text{C}$  isotope ratio of  $\text{CO}_2$ , *Analytical Chemistry*, 72(1), 216-221, 2000b.
- Francey, R. J. and Steele L. P. Measuring atmospheric carbon dioxide - the calibration challenge, *Accreditation and Quality Assurance*, 8, 200-204, 2003.
- Glatzel-Mattheier, H.: Bilanzierung von  $\text{CH}_4$ -Emissionen in Deutschland anhand atmosphärischer Messungen in Heidelberg, Ph.D. thesis, University of Heidelberg, Germany, 129 pp., 1997.
- Griffith, D. W. T.: Synthetic calibration and quantitative analysis of gas phase infrared spectra, *Appl. Spectrosc.*, 50(1), 59-70, 1996.
- Griffith, D. W. T., Esler, M. B., Steele, L. P., and Reisinger, A.: Non-linear least squares: high precision quantitative analysis of gas phase FTIR spectra, in: 2nd International Conference on Advanced Vibrational Spectroscopy, Nottingham, 2003.
- Griffith, D. W. T., Deutscher, N. M., Krummel, P., Fraser, P., van der Schoot, M., and Allison, C.: The UOW FTIR trace gas analyser: Comparison with Loflo, AGAGE and tank measurements at Cape Grim and GASLAB, Baseline Atmospheric Program (Australia), 2010.
- Griffith, D. W. T., Deutscher, N. M., Caldow, C., Kettlewell, G., Rigggenback, M., and Hammer, S.: A Fourier Transform Infrared trace gas analyser for atmospheric applications, *Atmospheric Measurements Techniques*, *Atmos. Meas. Tech.*, 5, 2481-2498, doi:10.5194/amt-5-2481-2012, 2012.

- 1 Hammer, S.: Quantification of the regional H<sub>2</sub> sources and sinks inferred from atmospheric  
2 trace gas variability, Ph.D. thesis, University of Heidelberg, Germany, 121 pp., 2008.
- 3 Hammer, S., Vogel, F., Kaul, M., and Levin, I.: The H<sub>2</sub>/CO ratio of emissions from  
4 combustion sources: comparison of top-down with bottom-up measurements in the Rhine-  
5 Neckar region in south-west Germany, *Tellus*, 61B, doi:10.1111/j.1600-0889.2009.00418.x,  
6 2009.
- 7 Huang L., Chivulescu A., Allison C., Brailsford G., Brand W.A., Wendeberg M.,  
8 Bollenbacher A., Keeling R., Levin I., Sabasch M., Leuenberger M., Mukai H., Nakazawa T.,  
9 Aoki S., Neubert R., Aerts-Bijma A., Verkouteren M., White J., Vaughn B., Michel S., Zhou  
10 L., Liu L. X., A Report of  $\delta^{13}\text{C}$  &  $\delta^{18}\text{O}$  Measurements in NBS 19 and NBS 18 pure CO<sub>2</sub>:  
11 Uncertainty in Traceability of CO<sub>2</sub> Isotope Measurements, Proceedings of the 15th  
12 IAEA/WMO meeting of CO<sub>2</sub> experts, Jena, Sept. 2009, WMO-GAW Report 194, ed. A.  
13 Jordan and W.A. Brand 169-175, 2011.
- 14 Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl, C. A., Guenther, P. R.,  
15 Waterman, L. S., and Chin, J. F. S.: Atmospheric carbon dioxide variations at Mauna Loa  
16 Observatory, Hawaii. *Tellus*, 28(6), 538–551, 1976.
- 17 Kitzi, D.: Preparation and stability of standard reference air mixtures, available at  
18 <http://www.esrl.noaa.gov/gmd/cc1/airstandard.html> (last access: 1 Feb 2012), 2009.
- 19 Klausen, J., and Scheel, H.-E.: WMO/GAW Glossary of QA/QC-Related Terminology,  
20 available at <http://gaw.empa.ch/glossary.html>, 2007.
- 21 Krystek, M., Anton, M.: A weighted total least-squares algorithm for fitting a straight line,  
22 *Meas. Sci. Technol.* 18, 3438-3442, 2007.
- 23 Langenfelds, R. L., van der Schoot, M. V., Francey, R. J., and Steele, L. P.: Modification of  
24 air standard composition by diffusive and surface processes, *J. Geophys. Res.*, 110, D13307,  
25 doi:10.1029/2004JD005482, 2005.
- 26 Long, D., Bielska, K., Lisak, D., Havey, D. K., Okumura, M., Miller, C. E., and Hodges J.  
27 T.: The air-broadened, near-infrared CO<sub>2</sub> line shape in the spectrally isolated regime:  
28 Evidence of simultaneous Dicke narrowing and speed dependence, *J. Chem. Phys.*, 135,  
29 064308, doi:10.1063/1.3624527, 2011.

1 McManus, J. B.; Shorter, J. H.; Nelson, D. D.; Zahniser, M. S.; Glenn, D. E.; McGovern, R.  
2 M: Pulsed quantum cascade laser instrument with compact design for rapid, high sensitivity me-  
3 asurements of trace gases in air, *Applied Physics B*, Volume 92, Issue 3, pp.387-392.

4 Messenger, C., Schmidt, M., Ramonet, M., Bousquet, P., Simmonds, P., Manning, A., Kazan,  
5 V., Spain, G., Jennings, S.G., Chiaia, P.: Ten years of CO<sub>2</sub>, CH<sub>4</sub>, CO and N<sub>2</sub>O fluxes over  
6 Western Europe inferred from atmospheric measurements at Mace Head, Ireland, *Atmos.*  
7 *Chem. Phys. Discuss.*, 8, 1191–1237, 2008.

8 Nakamichi, S., Kawaguchi, Y., Fukuda, H., Enami, S., Hashimoto, S., Kawasaki, M.,  
9 Umekawa, T., Morino, I., Suto, H., and Inoue, G.: Buffer-gas pressure broadening for the (3  
10 00 1)III ← (0 0 0) band of CO<sub>2</sub> measured with continuous-wave cavity ring-down  
11 spectroscopy, *Phys. Chem. Chem. Phys.*, 8, 364-368, doi: 10.1039/B511772K, 2006.

12 Rella C. W., H. Chen, A. E. Andrews, A. Filges, C. Gerbig, J. Hatakka, A. Karion,  
13 N. L. Miles, S. J. Richardson, M. Steinbacher, C. Sweeney, B. Wastine, and C. Zellweger:  
14 High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid  
15 air, *Atmos. Meas. Tech. Discuss.*, 5, 5823-5888, 2012

16 Rothman, L. S., Jacquemart, D., Barbe, A., Chris Benner, D., Birk, M., Brown, L. R., Carleer,  
17 M. R., Chackerian Jr., C., Chance, K., Coudeth, L. H., Dana, V., Devi, V. M., Flaud, J.-M.,  
18 Gamache, R. R., Goldman, A., Hartmann, J.-M., Jucks, K. W., Maki, A. G., Mandin, J.-Y.,  
19 Massie, S. T., Orphal, J., Perrin, A., Rinsland, C. P., Smith, M. A. H., Tennyson, J.,  
20 Tolchenov, R. N., Toth, R. A., Vander Auwera, J., Varanasi, P., and Wagner, G.: The  
21 HITRAN 2004 molecular spectroscopic database, *Journal of Quantitative Spectroscopy &*  
22 *Radiative Transfer*, 96(2), 139-204, 2005.

23 Vogel F. R. , L. Huang, D. Ernst, L. Giroux, S. Racki, and D. E. J. Worthy: Evaluation of a  
24 cavity ring-down spectrometer for in situ observations of <sup>13</sup>CO<sub>2</sub> , *Atmos. Meas. Tech.*  
25 *Discuss.*, 5, 6037-6058, 2012

26 Werle, P., Muecke, R., and Slemr, F.: The limits of signal averaging in trace gas monitoring  
27 by tunable diode laser absorption spectroscopy (TDLAS), *Appl. Phys.*, B57, 131-139, 1993.

28 Winderlich, J., Chen, H., Höfer, A., Gerbig, C., Seifert, T., Kolle, O., Kaiser, C., Lavrič, J.V.,  
29 and Heimann, M.: Continuous low-maintenance CO<sub>2</sub>/CH<sub>4</sub>/H<sub>2</sub>O measurements at the Zotino  
30 Tall Tower Observatory (ZOTTO) in Central Siberia, *Atmos. Meas. Tech. Discuss.*, 3(2),  
31 1399-1437, 2010.

- 1 WMO: Report of the 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other  
2 Greenhouse Gases, and Related Tracers Measurement Techniques, 7–10 September 2009,  
3 GAW Report No. 194, WMO TD No. 1553, available at:  
4 [http://www.wmo.int/pages/prog/arep/gaw/ 5 documents/GAW 194 WMO TD No 1553 web](http://www.wmo.int/pages/prog/arep/gaw/5documents/GAW194WMO_TD_No1553_web_low_resol.pdf)  
5 [low resol.pdf](http://www.wmo.int/pages/prog/arep/gaw/5documents/GAW194WMO_TD_No1553_web_low_resol.pdf), last access: 2 Feb 2012, Jena, Germany, 2011.
- 6 WMO: Report of the WMO/UNEP//ICSU Meeting on Instruments, Standardization and  
7 Measurements Techniques for Atmospheric CO<sub>2</sub>, Geneva, 8-11 September 1981.
- 8 Worthy, D. E. J., Canadian Baseline Program: Summary of progress to 2002, Meteorological  
9 Service of Canada, Canada, 2003.
- 10 Zellweger, C., Steinbacher, M., and Buchmann, B.: Evaluation of new laser spectrometer  
11 techniques for in situ carbon monoxide measurements, *Atmos. Meas. Tech.*, 5, 2555-2567,  
12 doi:10.5194/amt-5-2555-2012, 2012
- 13

Table 1. Summary of all residual and cross sensitivity experiments for all species. For each investigated sensitivity (except for the flow rate sensitivity) three rows are given: The first row states the averaged slopes ( $dy/dx$ ) of the linear regression together with their  $1\sigma$  uncertainties. The second row gives an example of the magnitude of the sensitivity for a typical range of parameter variations. The last row categorizes the temporal stability of the sensitivities over the course of one year. In case of a temporally unstable sensitivity the largest sensitivity value ( $dy/dx$ ) is given as an upper estimate and in those cases no uncertainty estimates are tabulated. For the flow rate sensitivity an additional row is given stating the difference between a measurement performed at a flow rate of 1 slpm and a static measurement. For the  $CO_2$  cross sensitivity only the linear cross sensitivities for unpolluted  $CO_2$  levels are given. The temperature disequilibrium sensitivities (TDS) that were derived from cylinder measurements in the ICOS Demonstration Experiment setup are given as well. The last three rows summarize the effect of a  $1^\circ C$  temperature bias as result of theoretical MALT study. The correction factors are given independently for the density and spectroscopic error. The red colouration of some numbers is to highlight their significance relative to the WMO inter laboratory compatibility goals.



1 Table 1:

	CO <sub>2</sub>	error	δ <sup>13</sup> CO <sub>2</sub>	error	CO	error	CH <sub>4</sub>	error	N <sub>2</sub> O	error
	[μmol mol <sup>-1</sup> ]		[‰]		[nmol mol <sup>-1</sup> ]		[nmol mol <sup>-1</sup> ]		[nmol mol <sup>-1</sup> ]	
dx/dp: [unit hPa <sup>-1</sup> ]	0.0085	0.0004	0.005	0.002	0.006	0.002	0.031	0.003	0.007	0.001
Typical variation for: 20 hPa	0.17	0.01	0.10	0.04	0.12	0.04	0.62	0.06	0.14	0.02
temporal development	stable		linearly decreasing		stable		linearly decreasing		stable	
dx/dT [unit °C <sup>-1</sup> ]	<0.8	-	0.6	0.2	<1	-	<1.6	-	0.6	0.2
Typical variation for: 0.1°C	<0.08	-	0.06	0.02	0.10	-	0.16	-	0.06	0.02
temporal development	unstable		rel. stable		unstable		Unstable		rel. stable	
dx/dFlow [unit slpm <sup>-1</sup> ]	0.15	0.001	-0.9	0.6	<2	-	<4	-	<-0.8	-
Typical variation for: 0.03 slpm	0.00	0.00	-0.03	0.02	<0.1	-	<0.1	-	<-0.02	-
static vs. 1 slpm	0.25		0.60		0.10		1.40		0.28	
temporal development	stable		unstable		unstable		unstable		unstable	
dx/dH <sub>2</sub> O [unit ppm <sup>-1</sup> (H <sub>2</sub> O)]	0.04	0.02	-	-	<0.2	-	<0.2	-	-	-
Typical variation for: 2 μmol mol <sup>-1</sup> (H <sub>2</sub> O)	0.08	0.04	-	-	<0.4		<0.4	-	-	-
temporal development	stable				unstable		unstable			
dx/dCO <sub>2</sub> [unit μmol mol <sup>-1</sup> ]	-	-	0.006	0.0003	<0.015		-		0.008	0.0008
Typical variation for: 50 μmol mol <sup>-1</sup> (CO <sub>2</sub> )	-	-	0.3	0.02	0.75		-		0.4	0.04
temporal development			stable		unstable				stable	
dx/dT (TDS) [unit °C <sup>-1</sup> ]	2.07	0.05	4.1	0.1	-4.6	0.3	10.2	0.3	3.2	0.1
Typical variation for: 0.1°C	0.2	0.01	0.41	0.01	0.50	0.03	1.0	0.03	0.32	0.01
temporal development	stable		stable		stable		stable		stable	
dx/dT [unit °C <sup>-1</sup> ] 1°C temperature bias MALT experiment	2.0		4.45		-4.0		7.8		1.7	
dx/dT [unit °C <sup>-1</sup> ] related	1.3		0.0		0.3		5.0		1.0	

1	to the density error					
	$dx/dT$ [unit °C <sup>-1</sup> ] related to spectroscopic error	0.7	4.45	-4.3	2.8	0.7

1  
2 Table 2. Repeatability ( $1\sigma$ ) for 2.5 and 10 min averaging times deduced from Allan-Variance  
3 analysis. Reproducibility ( $1\sigma$ ) for 3 min measurements of a single tank over a six day period  
4 and target gas reproducibility over the course of several months (compare Figure 9). Total  
5 uncertainty for a single ambient air measurement (see section 6).

Species	Repeatability ( $1\sigma$ )		Reproducibility ( $1\sigma$ )			Total uncertainty incl. errors in sensitivities ( $1\sigma$ )
	2.5 min	10 min	Single tank 6 days 3 min	Target: Inter- mediate setup	Target: IUP setup	
CO <sub>2</sub> [ $\mu\text{mol mol}^{-1}$ ]	0.018	0.012	0.023	0.043	0.016	0.032
$\delta^{13}\text{CO}_2$ [‰]	0.05	0.03	0.06	0.062	0.034	0.07
CH <sub>4</sub> [ $\text{nmol mol}^{-1}$ ]	0.2	0.1	0.2	0.21	0.12	0.25
CO [ $\text{nmol mol}^{-1}$ ]	0.2	0.1	0.2	0.26	0.11	0.22
N <sub>2</sub> O [ $\text{nmol mol}^{-1}$ ]	0.07	0.04	0.07	0.042	0.042	0.084

6  
7

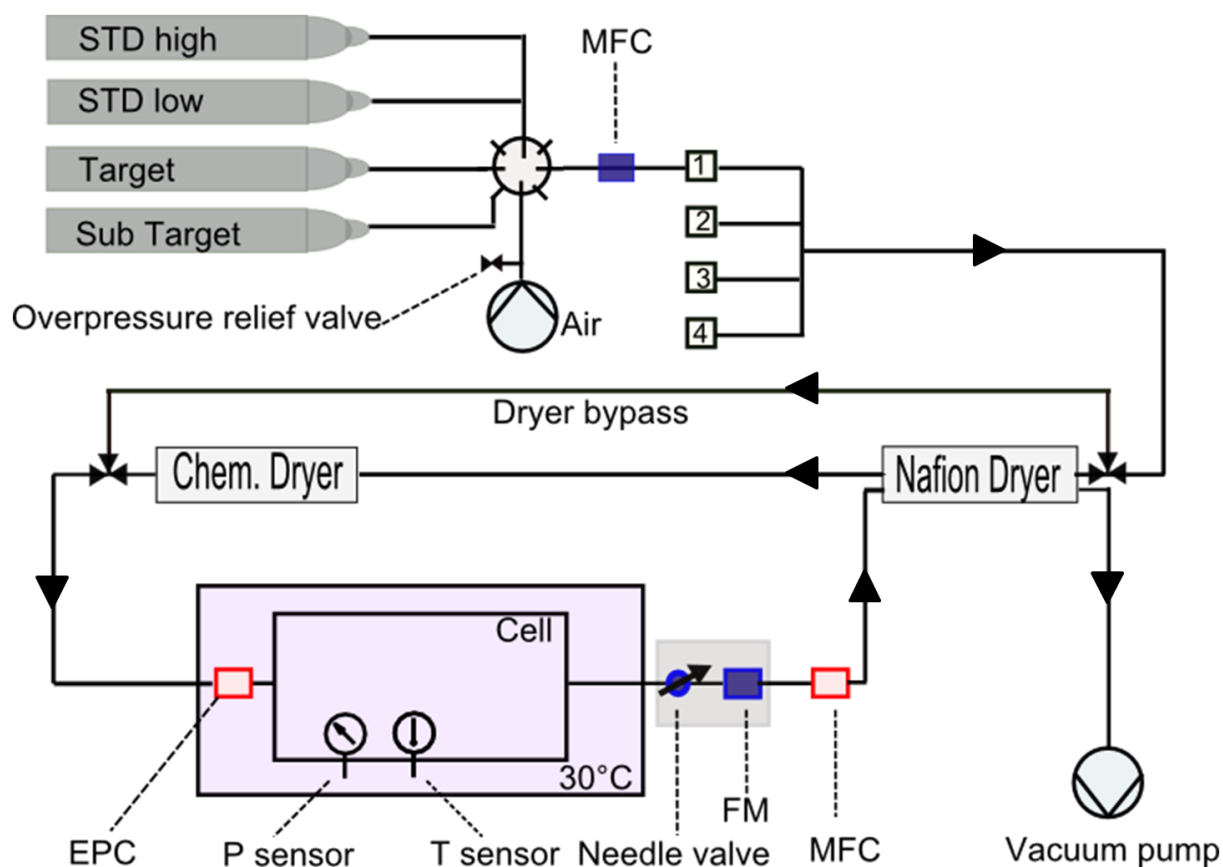


Figure 1. Schematic set-up of the in situ FTIR analyser. The blue parts have been replaced with the red parts in the final IUP set-up. The modifications include Mass Flow Controllers (MFC) and Electronic Pressure Controllers (EPC) and replace the original needle valve - Flow Meter (FM) unit.

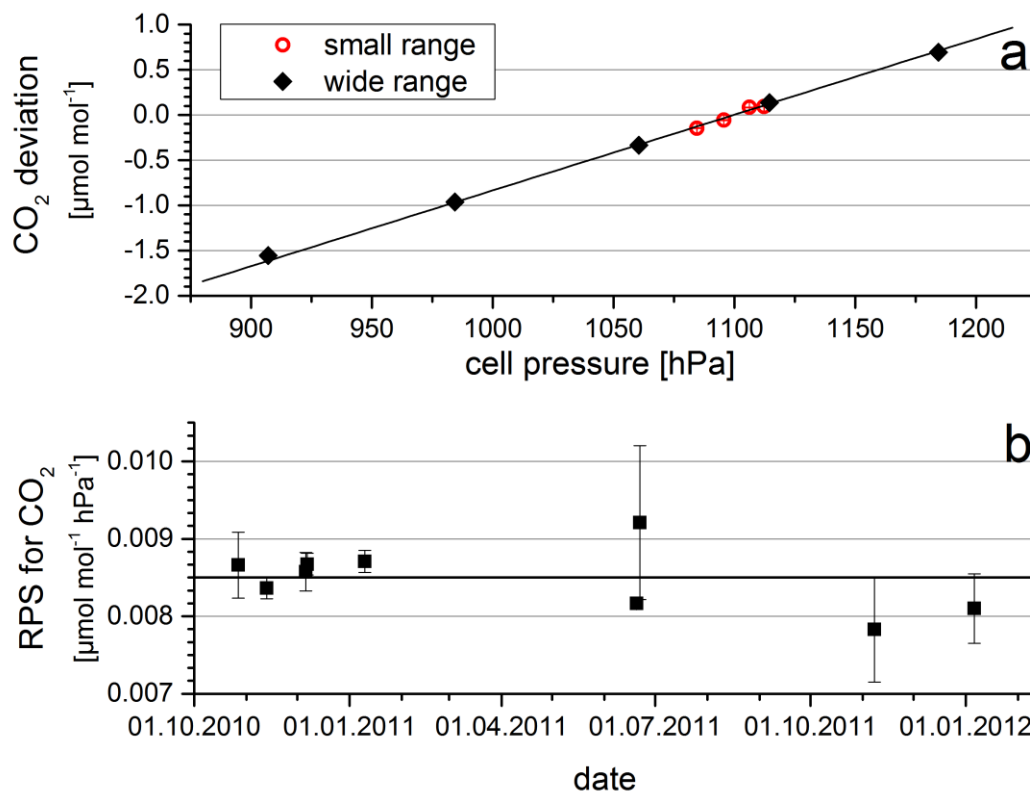
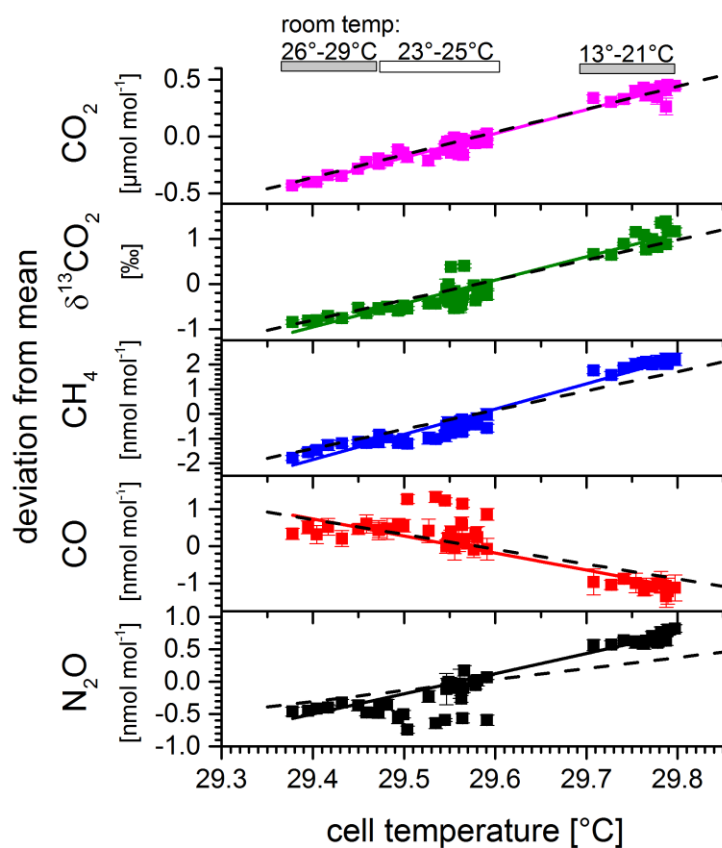


Figure 2. a) Residual pressure sensitivity (RPS) for CO<sub>2</sub> measured on 20<sup>th</sup> and 22<sup>nd</sup> of June, 2011. Small and wide pressure ranges are shown together with a common linear fit. Individual errors are smaller than the symbols. b) Temporal stability of the slope of the residual pressure sensitivity for CO<sub>2</sub>. The mole fractions of the used cylinders were all in the range of ambient concentrations in Heidelberg.

1



2

3 Figure 3. Temperature disequilibrium sensitivity (TDS) of the sub-target measurements,  
 4 shown as deviation from mean mole fractions. Values have been corrected for all residual  
 5 sensitivities except for residual temperature, **but not calibrated**. The three different  
 6 temperature ranges in the cell result from different laboratory temperatures at the remote **field**  
 7 stations and in the IUP laboratory **that** are indicated in the top panel. **The dashed lines**  
 8 **correspond to a TDS for a 1°C bias in sample temperature determination, based on the density**  
 9 **and the introduced-spectroscopic error only. Please refer to the text for more details.**

10

11

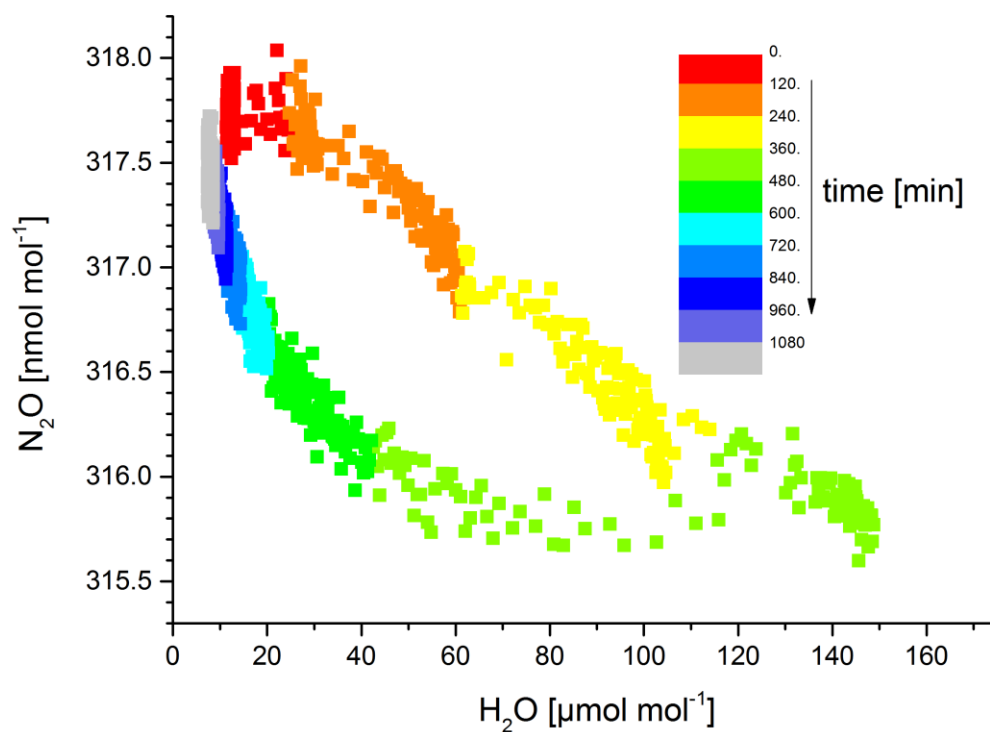


Figure 4.  $\text{H}_2\text{O}$  cross sensitivity for  $\text{N}_2\text{O}$ . In the course of this experiment the cell was first gradually humidified and dried afterwards. The colour code represents time since start of the experiment in minutes.

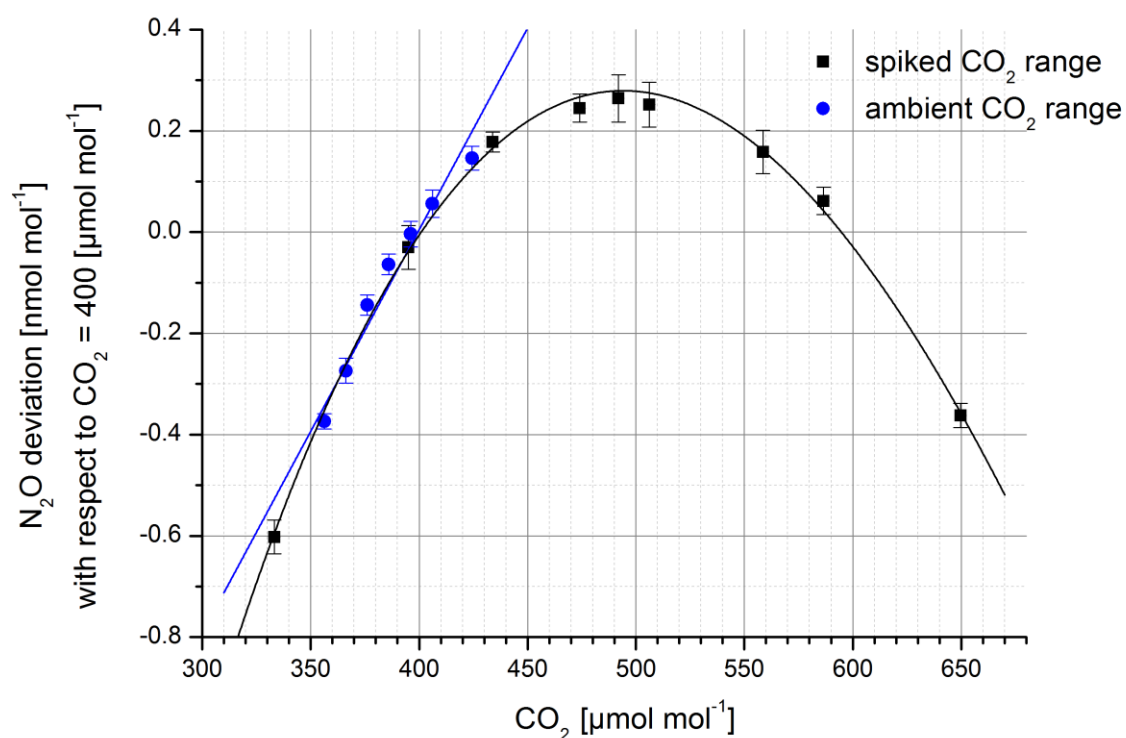
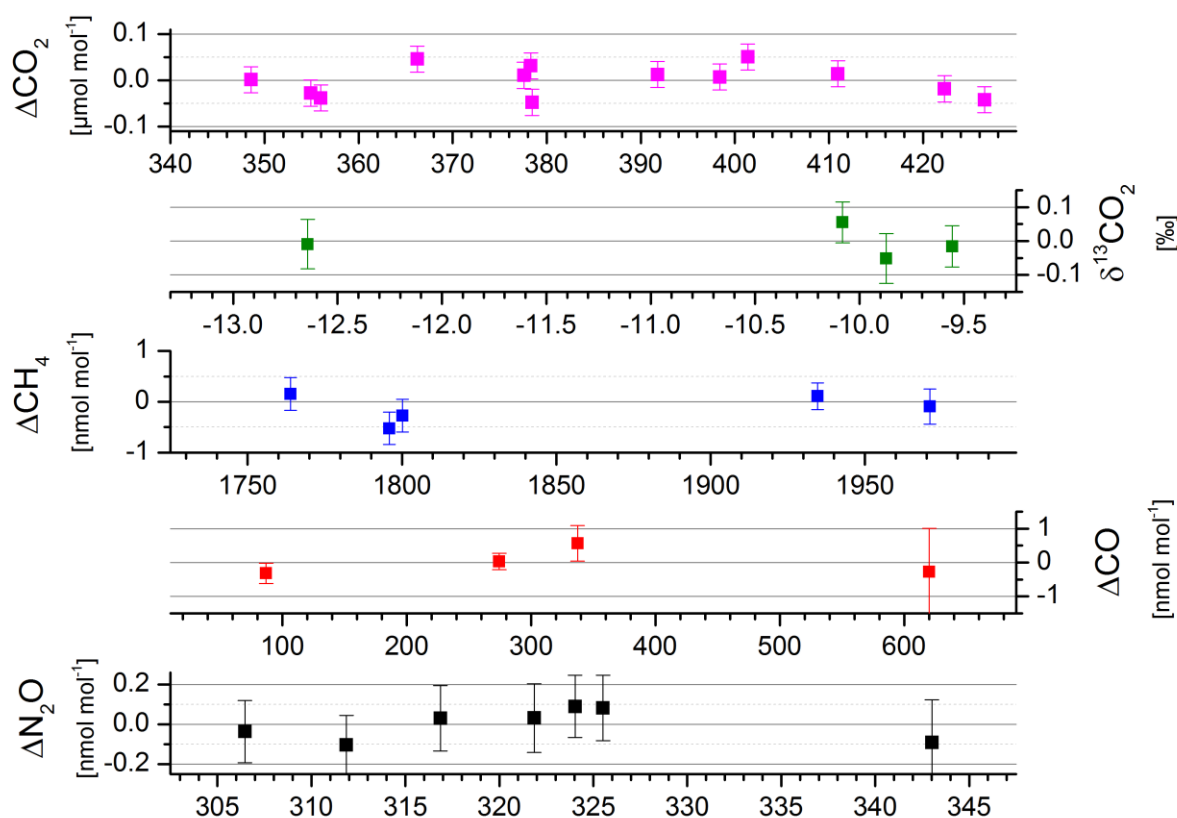


Figure 5. CO<sub>2</sub> cross sensitivity for N<sub>2</sub>O, shown as deviation to a reference CO<sub>2</sub> mole fraction of 400 μmol mol<sup>-1</sup>. In blue a CO<sub>2</sub> cross sensitivity experiment spanning the unpolluted CO<sub>2</sub> range is shown, together with its linear approximation. Results from a spiked CO<sub>2</sub> experiment and a cubic fit are displayed in black. Both experiments have been conducted with N<sub>2</sub>O mole fractions of about 311 nmol mol<sup>-1</sup>.



1



2

3

4 Figure 6. Linearity of the FTIR analyser. For each species the residual to the linear fit of the  
 5 externally assigned mole fractions against sensitivity-corrected FTIR mole fractions are  
 6 shown. The investigated mole fraction ranges are defined by the spread of the **secondary**  
 7 **laboratory cylinders used for the Heidelberg GC system** and are as follows:  $\text{CO}_2$ : 348 to 426  
 8  $\mu\text{mol mol}^{-1}$ ;  $\delta^{13}\text{CO}_2$ : -12.65 to -9.55‰;  $\text{CO}$ : 90 to 620  $\text{nmol mol}^{-1}$ ,  $\text{CH}_4$ : 1757 to 1970  $\text{nmol}$   
 9  $\text{mol}^{-1}$  and  $\text{N}_2\text{O}$ : 307 to 343  $\text{nmol mol}^{-1}$ .

10

11

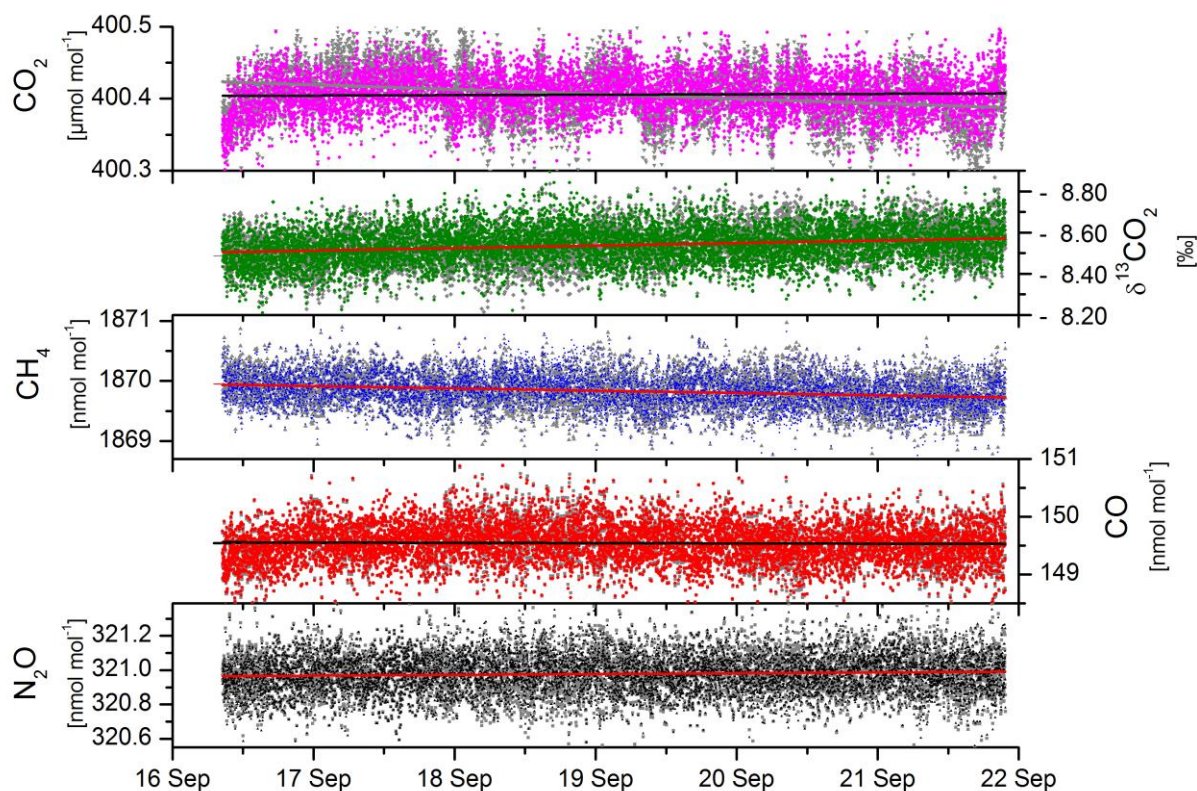
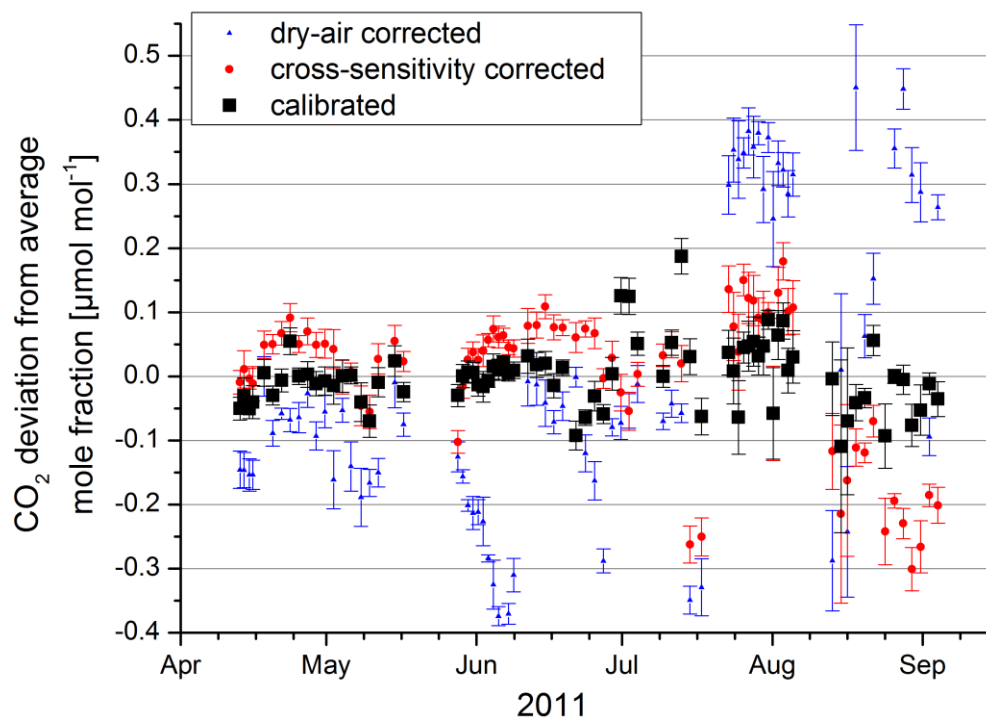


Figure 7. Short term stability of the FTIR components during September 2011. Minute-by-minute cylinder measurements over the course of six days. In grey the dry air mole fractions are given (but are only visible in the case of  $\text{CO}_2$ ). The coloured symbols show the sensitivity corrected values, with a linear fit to detect drifts.

1



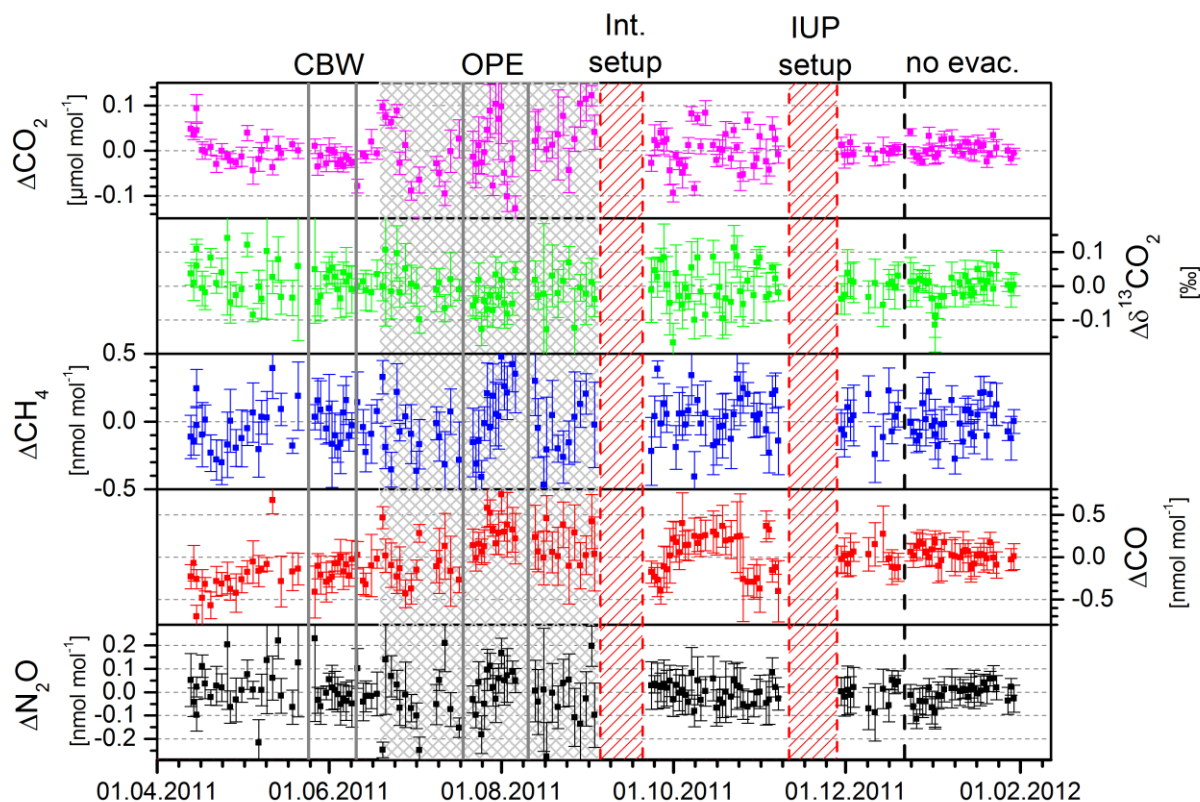
2

3

4 Figure 8. CO<sub>2</sub> target gas record in its different data processing stages.

5

1



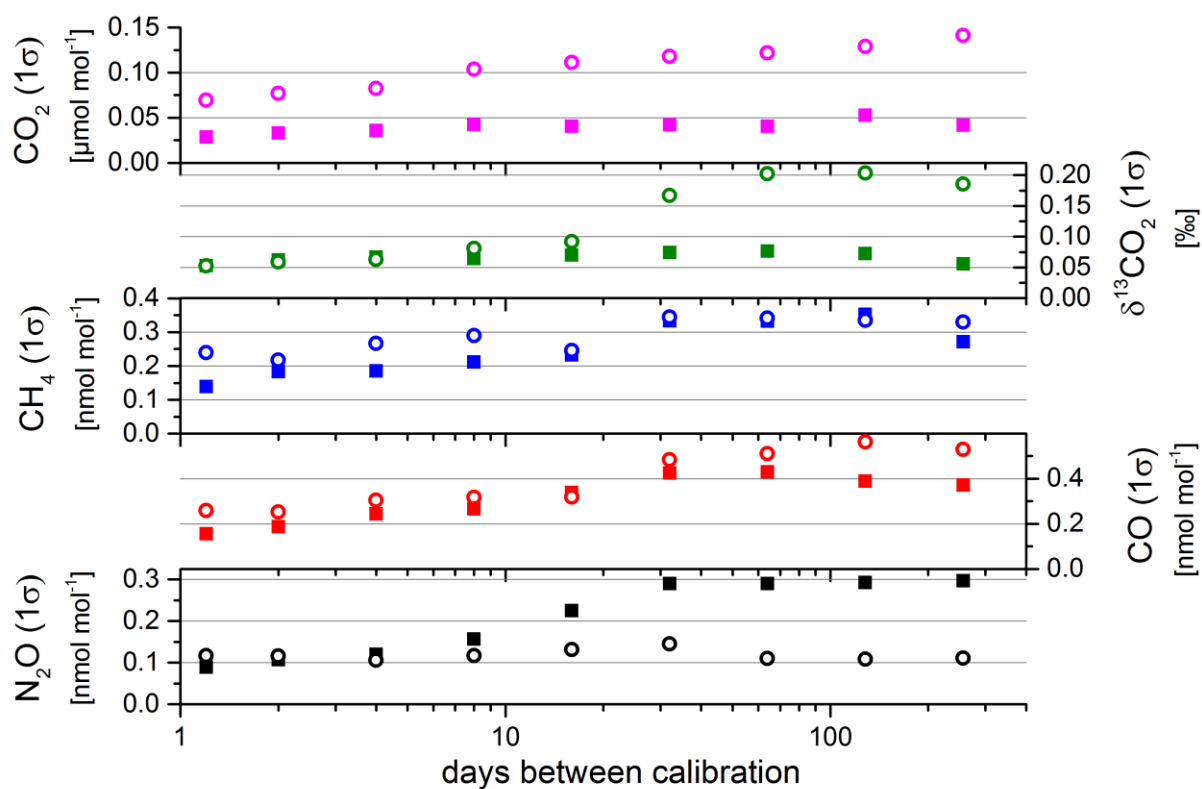
2

3

4 Figure 9. Long term stability of the FTIR analyser based on the 24- to 48-hourly target gas  
 5 measurements. Vertical gray lines denote location changes of the instrument to Cabauw the  
 6 Netherlands (CBW) and Houdelaincourt, France (OPE). The gray shaded area highlights a  
 7 period with less stable cell pressure. Red shaded areas mark instrument down times due to  
 8 modifications and/or laser failure. In September 2011 the internal mass flow controller and  
 9 end of November 2011 the electronic pressure controller (EPC) was installed. The dashed  
 10 black line indicates the change in standard operation conditions from then on no evacuation  
 11 was performed during sample exchange.

12

1



2

3

4 Figure 10. Calibration frequency test. 1σ standard deviation of the target gas measurements as  
 5 a function of days between calibrations. Open symbols represent less stable instrument  
 6 conditions with sample pressure variations (1σ) of more than ±7 hPa. Filled symbols relate to  
 7 stable conditions with sample pressure variations (1σ) of less than ±2 hPa.

8

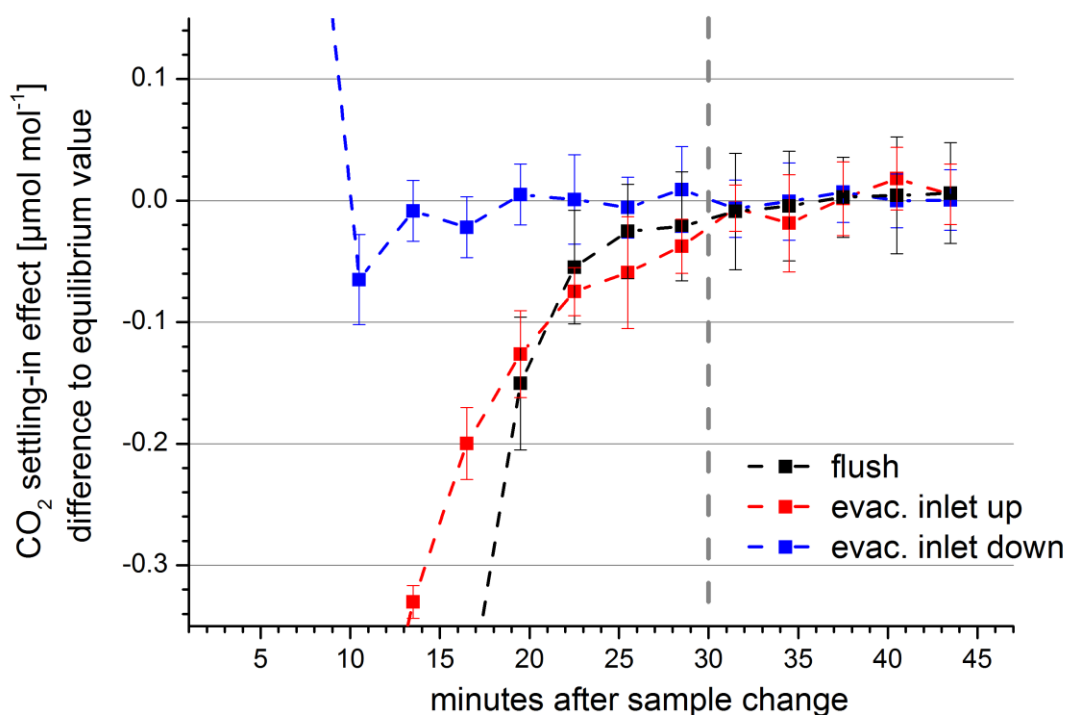


Figure A1. Comparison of the different sample change-over methods for CO<sub>2</sub>. The settling-in effects are shown as relative derivation from the equilibrium value defined by the average of the last five measurements. Sample exchange by flushing only is shown in black, sample exchange via evacuation, as used in **the standard operation conditions**, in blue with standard cell orientation and in red with the cell turned upside down. Measurements are shown centred in their 3 min interval.