



Comparison of continuous in situ CO₂ observations at Jungfraujoch using two different measurement techniques

M. F. Schibig¹, M. Steinbacher², B. Buchmann², I. T. van der Laan-Luijkx^{1,*}, S. van der Laan^{1,**}, S. Ranjan^{***}, and M. C. Leuenberger¹

¹Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

²Swiss Federal Laboratories for Materials Science and Technology (Empa), Dübendorf, Switzerland

* now at: Meteorology and Air Quality, Wageningen University, Wageningen, the Netherlands

** now at: Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK

*** formerly at: Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland

Correspondence to: M. C. Leuenberger (leuenberger@climate.unibe.ch)

Received: 30 May 2014 – Published in Atmos. Meas. Tech. Discuss.: 16 July 2014

Revised: 12 November 2014 – Accepted: 27 November 2014 – Published: 6 January 2015

Abstract. Since 2004, atmospheric carbon dioxide (CO₂) is being measured at the High Altitude Research Station Jungfraujoch by the division of Climate and Environmental Physics at the University of Bern (KUP) using a nondispersive infrared gas analyzer (NDIR) in combination with a paramagnetic O₂ analyzer. In January 2010, CO₂ measurements based on cavity ring-down spectroscopy (CRDS) as part of the Swiss National Air Pollution Monitoring Network were added by the Swiss Federal Laboratories for Materials Science and Technology (Empa). To ensure a smooth transition – a prerequisite when merging two data sets, e.g., for trend determinations – the two measurement systems run in parallel for several years. Such a long-term intercomparison also allows the identification of potential offsets between the two data sets and the collection of information about the compatibility of the two systems on different time scales.

A good agreement of the seasonality, short-term variations and, to a lesser extent mainly due to the short common period, trend calculations is observed. However, the comparison reveals some issues related to the stability of the calibration gases of the KUP system and their assigned CO₂ mole fraction. It is possible to adapt an improved calibration strategy based on standard gas determinations, which leads to better agreement between the two data sets. By excluding periods with technical problems and bad calibration gas

cylinders, the average hourly difference (CRDS – NDIR) of the two systems is $-0.03 \text{ ppm} \pm 0.25 \text{ ppm}$. Although the difference of the two data sets is in line with the compatibility goal of $\pm 0.1 \text{ ppm}$ of the World Meteorological Organization (WMO), the standard deviation is still too high. A significant part of this uncertainty originates from the necessity to switch the KUP system frequently (every 12 min) for 6 min from ambient air to a working gas in order to correct short-term variations of the O₂ measurement system. Allowing additional time for signal stabilization after switching the sample, an effective data coverage of only one-sixth for the KUP system is achieved while the Empa system has a nearly complete data coverage. Additionally, different internal volumes and flow rates may affect observed differences.

1 Introduction

Carbon dioxide (CO₂) is an important greenhouse gas which plays a major role for the radiative forcing of the atmosphere and therefore contributes significantly to the greenhouse effect (Arrhenius, 1896). Since natural as well as anthropogenic processes alter the mole fraction of CO₂ in the atmosphere, it is important to survey the mole fraction changes over a long time period to understand these pro-

cesses and draw conclusions about the involved CO₂ fluxes (Ramonet et al., 2010; Schimel et al., 2001). Additionally, in situ CO₂ measurements can be used in modeling studies as input parameters or to verify model output (Uglietti et al., 2011; Chevallier et al., 2010; Peters et al., 2010). Therefore it is crucial that the data sets are well calibrated and traced back to a common reference scale and that the measurement systems demonstrate high precision and accuracy. This guarantees minimal biases between different measurement sites and laboratories and improves the compatibility of the data sets. For CO₂ measurements, the World Meteorological Organization (WMO) recommends a compatibility of ± 0.1 ppm for the northern hemisphere (WMO, 2011).

High-altitude observatories such as the High Alpine Research Station Jungfraujoch (JFJ) are predestined to monitor the background mole fraction of trace gases like CO₂ because they are mostly in the free troposphere.

That is why the division of Climate and Environmental Physics of the Physics Institute, University of Bern (KUP) started measuring CO₂ with a non-dispersive infrared (NDIR) gas analyzer (Sick Maihak, Germany, model S710) embedded in a combined CO₂ / O₂ analyzing system at JFJ in late 2004. At the beginning of 2010, a cavity ring-down spectrometer (until July 2011 a Picarro Inc., USA, model G1301, afterwards a Picarro Inc., USA, model G2401) was installed by the Swiss Federal Laboratories for Materials Science and Technology (Empa) as part of the Swiss National Air Pollution Monitoring Network (NABEL), enabling a long-term performance of CO₂ measurements at JFJ by Empa. To ensure that both records could be merged smoothly, both systems ran in parallel for a significant period, not only to detect potential biases but also to provide information about the compatibility of the two different measurement systems as had been done at other stations, e.g., at Mauna Loa (Komhyr et al., 1989; Peterson et al., 1977). Because of the year-round manned infrastructure and excellent accessibility by train, it was possible to test the compatibility of the two systems in situ under real conditions at a background site featuring only limited atmospheric variation.

In this study we report the data of the KUP and the Empa in situ CO₂ measurement system from January 2010 until December 2012.

2 Methods

2.1 Sampling site

The High Altitude Research Station Jungfraujoch (JFJ) is located 7°59'20" E, 46°32'53" N on the northern ridge of the Swiss Alps. The station itself is on a mountain saddle between the mountains Mönch (4099 m a.s.l.) and Jungfrau (4158 m a.s.l.), at an altitude of 3580 m a.s.l. and is accessible by train the whole year round. Because of the high elevation, the station is above the planetary boundary layer most

of the time and receives predominantly air from the free troposphere. It is therefore an ideal location to measure the atmospheric background air of continental Europe (Henne et al., 2010; Zellweger et al., 2003; Baltensperger et al., 1997). Nevertheless, the station is sometimes influenced by polluted boundary layer air, especially during Föhn events (Zellweger et al., 2003) or during hot summer days, when air from the surrounding valley is thermally uplifted to JFJ (Zellweger et al., 2000; Baltensperger et al., 1997). A comprehensive in situ measurement program with more than 70 trace gases and a large suite of aerosol properties is run continuously at JFJ by an international consortium of research institutions. JFJ is also one of the currently 29 Global Atmosphere Watch (GAW) stations.

2.2 NDIR measurements (KUP)

The KUP CO₂ measurements are based on a combined system to monitor CO₂ and O₂ changes in the atmosphere. The ambient air enters through a strongly ventilated (600 m³ h⁻¹) common inlet on the observatory's roof to a manifold, which serves many air detectors, where an aliquot is drawn to the KUP system. The air is cryogenically dried to a dew point of -90°C (FC-100D21, FTS systems, USA). CO₂ is measured by a NDIR spectrometer (Maihak S710) with a working range of 350 to 450 ppm and a frequency of 1 Hz; O₂ is measured by a paramagnetic cell. To avoid influences caused by ambient air temperature and density fluctuations, the temperature as well as the pressure is stabilized. The box with the pressure and flow control system is embedded in a temperature-controlled box ($45 \pm 0.05^{\circ}\text{C}$); the precision of the pressure regulation is ± 0.05 mbar and the measurement cell of the NDIR analyzer is heated to $55 \pm 0.05^{\circ}\text{C}$. Measurements are done in a cyclic sequence of 18 h: each gas is measured for 6 min, with only the last 115 s of a 6 min period used for mole fraction determination to allow for signal stabilization after changing the sample source. At the beginning of each sequence, the system is calibrated with two reference gases (high and low span). A working gas is measured between two ambient air measurements to correct for short-term variations. Therefore, an 18 h measurement sequence looks as follows:

B – G – K – B – H – B – A – A – B – A – A – B –
 ... – A – A – B,

where B represents the working gas, G the high span CO₂/low span O₂, K the low span CO₂/high span O₂, H the target cylinder and A represents ambient air.

All measurements ending at a particular hour are used for the calculation of hourly mean CO₂ observations, which in our case includes six ambient observation values per hour. Cylinder measurements (see Table 1) with a known mole fraction show a precision better than 0.04 ppm for 1 h analysis. The precision is calculated as the square root of the sum

Table 1. KUP calibration cylinders and their measured CO₂ values on the NDIR and the CRDS system. These values were used to calculate the second degree polynomial to correct the nonlinearity of the NDIR system (See also Fig. 9).

Cylinder number	Used before as	CO ₂ Picarro [ppm]	CO ₂ Maihak [ppm]	ΔCO ₂ (Picarro – Maihak) [ppm]	ΔCO ₂ (Picarro – Maihak corrected) [ppm]	Pressure [bar]
LK542510	G1	431.99 ± 0.02	431.74 ± 0.04	0.25	0.05	22.1
LK94521	K5	360.57 ± 0.07	359.96 ± 0.05	0.61	−0.06	23.8
LK570931	B20	438.01 ± 0.02	437.73 ± 0.04	0.28	−0.01	21.9
LK542204	B21	385.09 ± 0.02	384.87 ± 0.04	0.22	0.04	26.8
LK3137	B22	384.99 ± 0.02	387.71 ± 0.05	0.28	0.10	16.1
LK3077	B24	418.68 ± 0.08	418.78 ± 0.03	−0.06	−0.12	14.1
LK542469	B26	432.06 ± 0.03	431.87 ± 0.02	0.19	−0.01	23.7
LK564848	B27	448.40 ± 0.03	447.83 ± 0.02	0.57	0.08	22.7
LK570916	B29	437.66 ± 0.07	437.43 ± 0.05	0.23	−0.06	185

of the squared standard deviations of the individual measurements. The CO₂ values are reported on the WMO X2007 scale. In February 2011, the Maihak analyzer was replaced with an identical Maihak device because of technical problems.

2.3 Cavity ring-down spectroscopy measurements (Empa)

From late December 2009 until July 2011, Empa was measuring CO₂ with a commercially available wavelength-scanned cavity ring-down spectrometer (Picarro Inc., USA; G1301) coupled to a custom-built calibration/drying unit. Initially, the sample air from the manifold was dried prior to analysis by means of a Nafion dryer to a dew point of < −30 °C. Along with CO₂, the instrument also measured CH₄ and H₂O at the same sample frequency of approximately 0.5 Hz. The H₂O measurements allowed for correction of the CO₂ mole fraction in case of interferences (i.e., dilution and pressure broadening) due to potential water moisture in the system. From August 2010 on, the Nafion dryer was short cut, no water vapor removal was used and CO₂ dry air mole fractions were determined after application of an empirical humidity correction. The empirically determined water vapor corrections for CRDS analyzers do not significantly change over time and are very similar for different analyzers. According to Rella et al. (2013), the error in the reported dry air CO₂ mole fractions when using different correction coefficients is < 0.1 ppm, up to a H₂O volume mixing ratio of 2.5 Vol-%. 2.5 Vol-% of H₂O at JFJ refers to a dew point of more than 14 °C, a value that was never reached at JFJ. The error can be further reduced when using the dedicated correction coefficients for the respective analyzer and doing repeated water vapor interference measurements, as is the case here.

After a system breakdown due to a faulty electronic board in July 2011, the G1301 analyzer was replaced in September 2011 by the newer G2401 model (Picarro Inc., USA) that

is also capable of monitoring CO mole fractions. Since then, the sample gas has been dried again for the beneficial effects with respect to the precision of the CO analysis.

Calibrations are performed every 46 h with two calibration gases (high and low span). In addition, a target gas is analyzed every 15 h to detect potential shorter-term instrument sensitivity changes. The mole fractions of the calibration gases are determined by the World Calibration Center for CH₄, CO₂, CO and O₃ at Empa (Table 3). Measurements of used KUP standard cylinders at JFJ show a precision better than 0.04 ppm for the 15 min analysis (Table 1). Similar to the KUP, CO₂ mole fractions are also reported on the WMO X2007 scale.

2.4 Calibration gases

The gas used to calibrate the KUP measurements is compressed outside air filled in steel cylinders delivered by Carbagas (Switzerland). Normally, each gas cylinder is first delivered to our laboratory at the University of Bern and measured for its CO₂ mole fraction and δO₂/N₂ and δAr/N₂ values by mass spectrometry and a combined Licor 7000 NDIR and Oxzilla (Sable Systems, USA) system (CO₂ and O₂). The measured values are calibrated with a subset of three standards from the WMO/GAW Central Calibration Laboratory (CCL) run by the National Oceanic and Atmosphere Administration (NOAA, Boulder, USA) and are used as assigned values of the KUP system at JFJ (see Table 2). The cylinders are shipped to JFJ and stored in the basement until their usage as calibration gases. A maintained stock of calibration gases at JFJ is necessary because the KUP system has to be calibrated frequently for O₂ measurements (Uglietti et al., 2008) and has a rather high gas consumption of up to 150 mL per minute.

The calibration gases used by Empa consist of compressed dry natural air in aluminium cylinders filled by themselves with a modified oil-free compressor (Rix Industries, USA). The mole fractions of the gases are determined by

Table 2. List of calibration gas cylinders used by the NDIR system in the period from 2010 to 2012: K, G, H and B correspond to the usage as low span, high span, target and working gas, respectively. The gases were measured at the University of Bern and calibrated with a suite of NOAA standard cylinders.

Cylinder number	Used as	From	To	Assigned CO ₂ [ppm] WMO X2007
LK94521	K5	10.03.2009	31.12.2010	360.31 ± 0.50
LK541816	K6	31.12.2010	06.11.2012	380.77 ± 0.25
LK542332	K7	06.11.2012	now	384.41 ± 0.07
LK542217	G2	23.07.2009	now	429.02 ± 0.25
LK542119	B23	11.12.2009	27.04.2010	379.90 ± 0.25
LK542171	H1	08.09.2010	now	440.16 ± 0.25
LK3077	B24	27.04.2010	07.09.2010	416.60 ± 0.10
LK541869	B25	07.09.2010	09.02.2011	416.79 ± 0.10
LK542469	B26	09.02.2011	17.06.2011	431.76 ± 0.25
LK564848	B27	17.06.2011	19.10.2011	448.04 ± 0.25
LK557112	B28	19.10.2011	23.2.2012	425.35 ± 0.05
LK570916	B29	23.02.2012	03.07.2012	437.55 ± 0.10
LK564850	B30	03.07.2012	05.10.2012	395.62 ± 0.01
LK542463	B31	5.10.2012	23.01.2013	390.34 ± 0.05

Table 3. List of calibration gas cylinders used for the CRDS system in the period from 2010 to 2012 as measured at the Empa laboratory.

Cylinder number	Used as	From	To	Assigned CO ₂ [ppm] WMO X2007
D621867 (E062)	Cal. gas 1	18.12.2010	now	394.11 ± 0.03
CA06290 (E063)	Cal. gas 2	18.12.2010	now	429.72 ± 0.02
CB555583	Target	18.12.2010	now	426.04 ± 0.04

the World Calibration Center for CH₄, CO₂, CO and O₃ at Empa through comparison with a suite of standards from the WMO/GAW CCL (see Table 3). After measuring the calibration gases, the cylinders are shipped to JFJ.

Lists of the calibration gases used by both systems during the period of intercomparison are included in Tables 2 and 3.

3 Results

3.1 Comparison of the two data sets

The CO₂ measurements of both systems for the period 2010–2012 are shown in Fig. 1. In general, the two data sets show a good agreement. There are two longer gaps in the CRDS record from 11 June to 16 September 2011 and 5 to 18 January 2012. The first gap was caused by technical problems of the Picarro, which was replaced by a new four-channel model. In January 2012 a hard-disk failure was responsible for the roughly 2-week break. The NDIR data set shows four longer gaps from 31 January to 16 February 2011, 21 to 28 December 2011, 14 to 25 April 2012 and 7 to 24 October 2012. The reason for these gaps in the NDIR data set was technical issues with the Maihak analyzer, eventually leading to a replacement of the analyzer with an identical one.

Furthermore, a power failure was responsible for the failure of a gas flow controller and, in April and October 2012, the dysfunction of the gas drying system. Throughout the record, small gaps are also present due to system crashes, failures in the regulation of the gas flows, etc. The remaining hourly mean values, where we have overlap of both systems (20 460 common hours), are in good agreement. The data sets show a similar seasonality as well as peaks caused by pollution events (e.g., advection of air masses from the Po basin in Italy or from northeastern Europe).

The average seasonality of the monthly mean values, spanning 36 months, is 10.01 ± 0.33 ppm and 10.05 ± 0.37 ppm for the CRDS and the NDIR data set, respectively. Due to missing data, one monthly value in each data set has to be interpolated. For the complete data sets, the CRDS data show an annual increase of 1.89 ± 0.01 ppm year⁻¹, whereas the NDIR data show a slope of 1.69 ± 0.01 ppm year⁻¹; the uncertainties correspond to the error of the linear increase fitted to the data. Selecting only data points where both systems have overlapping data points, the difference in the slopes decreases slightly as documented by values of the annual increase of CO₂ of 1.91 ppm year⁻¹ (CRDS) and 1.72 ppm year⁻¹ (NDIR). A correlation plot of the CRDS data set against the NDIR data set reveals a very high correlation with a slope of 0.991 ppm ppm⁻¹, an intercept

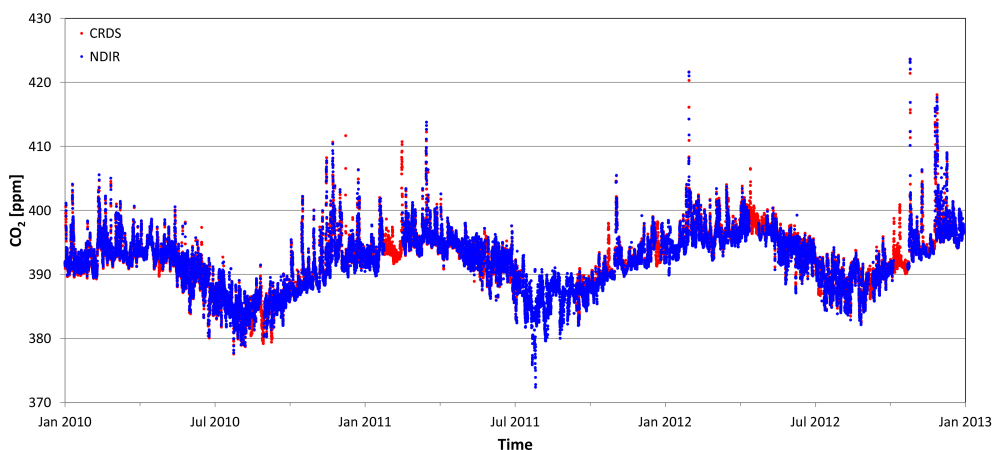


Figure 1. CO₂ measurements at JFJ of the CRDS system (red points) and the NDIR system (blue points) against time. The values represent the hourly averages of the two measurement systems.

3.662 ppm and a R^2 of 0.9909 (Fig. 2, grey points). It can be seen in Fig. 3 (grey points) that the differences of the two data sets are very stable over the 3-year period of the comparison. The average of the difference between the CRDS data set and the NDIR data set is 0.04 ± 0.40 ppm (Fig. 3, grey points); negative and positive differences equal out to almost zero. During periods of rapid CO₂ changes, the standard deviation of the differences is larger than during stable periods and significantly exceeds the precision of both instruments. This point can be emphasized by zooming in to higher-resolution data (115 s averages). For this exercise, the high-resolution CRDS data are aggregated to 115 s averages corresponding to the averaging intervals of the NDIR system. It can be clearly seen that the difference becomes noisier with fast changes (e.g., during the late afternoon in Figs. 4 and 5) and decreases again with more stable conditions (e.g., during the nighttime in Figs. 4 and 5).

The behavior of short-term variations is different for both systems as shown by a change of statistical characteristics when discarding values for which the change rate is above a certain threshold (Table 4). These changes are most probably due to different volumes and gas flow resulting in different residence times and leading to dispersion effects. For example, the NDIR system uses a water trap with a large volume, potentially dampening the CO₂ signal despite the higher flow rate compared to the rather small Nafion dryer of the CRDS system. The cutoff criteria of the change rate is set to 0.75 ppm h^{-1} and values with a higher change rate are discarded. Below 0.75 ppm h^{-1} there is still an improvement of the agreement, but the loss in data points is too high. Additionally, the standard error starts to increase again below a change rate of 0.75 ppm h^{-1} because many data points are omitted by this criteria. By setting the threshold to 0.75 ppm h^{-1} , roughly 20 % of the common data points are omitted. The correlation between the remaining data for both systems gets slightly closer to the ideal 1 : 1 function with a

slope of $0.996 \text{ ppm ppm}^{-1}$, an intercept of 1.5696 ppm and a R^2 of 0.9924 (Fig. 3, black points). Also, the differences between the CRDS and the NDIR data set are smaller; considering only the data points of the more stable conditions (Fig. 3, black points), the average of the difference remains almost 0, namely $0.05 \text{ ppm} \pm 0.32 \text{ ppm}$.

However, there are problematic features worth mentioning. By replacing the working gas B26 with B27 and B28 with B29, small shifts occur that are most likely caused by an inaccurate assignment of the CO₂ values of the working gases (Fig. 3). By excluding the periods of the cylinders B27 and B29, 11 038 common data points remain; the average of the differences, however, stays almost the same, namely -0.03 ppm but with a reduced standard deviation of 0.25 ppm.

In late 2010, the difference between the two data sets increased significantly because of technical issues (see Sect. 3.1). After replacing the analyzer, the difference was again very small and stable (Fig. 3). Furthermore, with the exchange of the Picarro in summer 2011, a small jump in the difference between the two systems occurred (Fig. 3). Since the replacement happened during a single working gas period of the NDIR system (B27), the offset was probably caused by the change of the Picarro analyzer.

3.2 Drift of calibration gases and their corrections

In the case of regular periodic calibrations, the pressure of a calibration gas cylinder decreases relatively linear over its lifetime. Because steel cylinders show pressure-dependent adsorption and desorption effects of gases such as CO₂ or H₂O (Langmuir, 1918), CO₂ continuously desorbs from the cylinder walls and increases the CO₂ mole fraction of the calibration gas during its usage. To avoid a large drift of the CO₂ mole fraction in calibration gases, Keeling et al. (1998) recommends using calibration gas cylinders only to a remaining

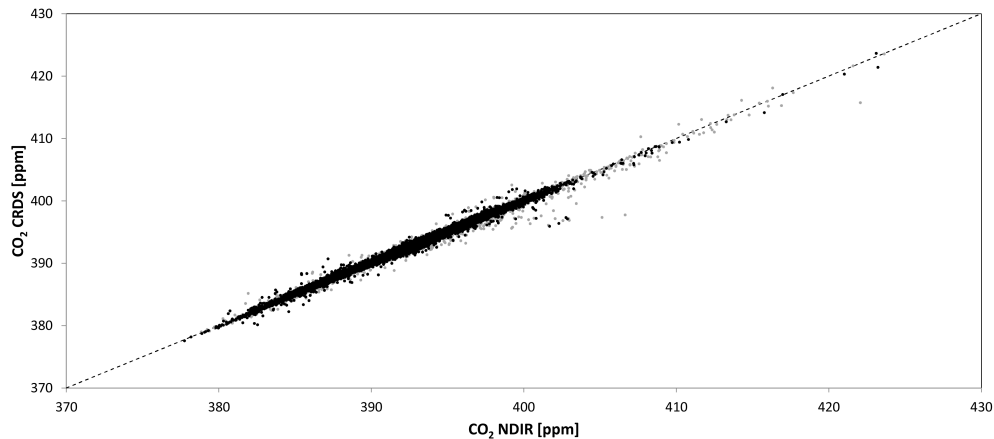


Figure 2. The hourly averages of the CRDS CO₂ measurements versus the hourly averages of the NDIR CO₂ measurements (grey points) and the CRDS CO₂ measurements versus the hourly averages of the NDIR CO₂ measurements without periods of rapid mole fraction changes (black points) over the whole period of comparison. The dashed diagonal represents the ideal 1 : 1 agreement. Considering all data points, the R^2 is 0.9909 with a slope of 0.9908 ppm ppm⁻¹ and an intercept of 3.6623 ppm. Excluding periods with rapid mole fraction changes, the agreement of the two data sets is slightly better with a R^2 of 0.9924, a slope of 0.9961 ppm ppm⁻¹ and an intercept 1.5696 ppm.

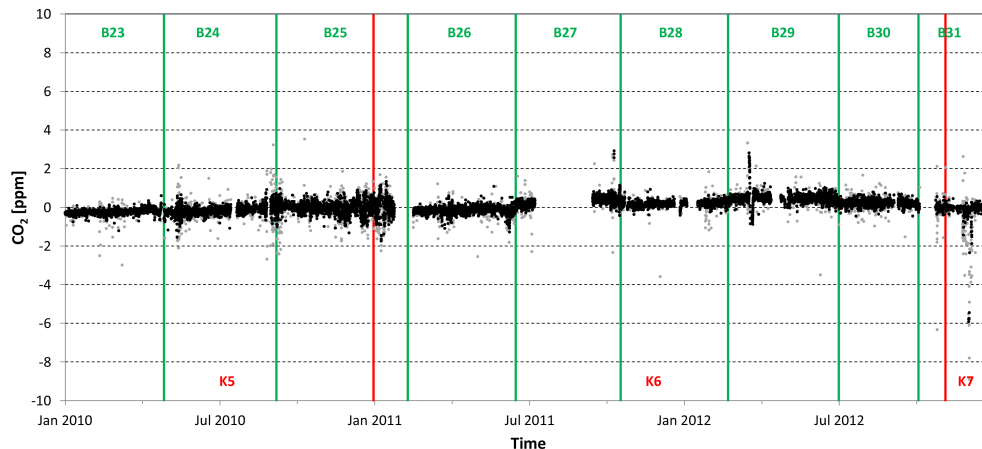


Figure 3. The difference of the CRDS – NDIR CO₂ measurements of all common hourly data points (grey points) and only during stable periods with a CO₂ change of less than 0.75 ppm h⁻¹ (black points) against time. The green bars represent changes of the working gases of the NDIR system with the green number indicating the according cylinder and the red bars indicate the changes of the low span with the according number, also in red. The high span was not replaced throughout the entire period.

pressure of 25 to 30 bar. The adsorption of CO₂ to the cylinder wall can be calculated according to Langmuir (Langmuir, 1918) and corrected for by the formula:

$$WG_{\text{corr}}(t) = WG_{\text{meas}}(t) - \left(\frac{a \cdot b \cdot p_0}{b \cdot p_0 + 1} - \frac{a \cdot b \cdot p(t)}{b \cdot p(t) + 1} \right), \quad (1)$$

where WG_{meas} corresponds to the measured CO₂ value of the working gas [ppm], a and b are constants (related to the Avogadro constant, the number of elementary spaces of the surface, the molecular mass and the temperature), p_0 is the start pressure and $p(t)$ is the pressure at the time of the measurement. This enrichment of CO₂ is more pronounced at lower pressures compared to high pressures, as seen in Eq. (1). For example, at 100 bar cylinder pressure the correction is only

about 0.05 ppm, whereas at 30 bar it is already up to 0.3 ppm (with $a = 5\,715\,797$ and $b = 568\,897$, a and b values were derived from measurements of cylinder B23). In steel cylinders, present water vapor would cause an even more pronounced desorption effect. However, all cylinders used here contain dry gas. In Leuenberger et al. (2014), an in-depth discussion of this adsorption/desorption influence of gases on steel and aluminium cylinders is presented, based on dedicated experimental results in our laboratory as well as in a climate chamber.

In April 2010 the working gas tank B23 ran completely empty due to a faulty pressure reader and therefore showed such an enrichment effect. Hence the CO₂ values of the NDIR calibrated with the uncorrected working gas are under-

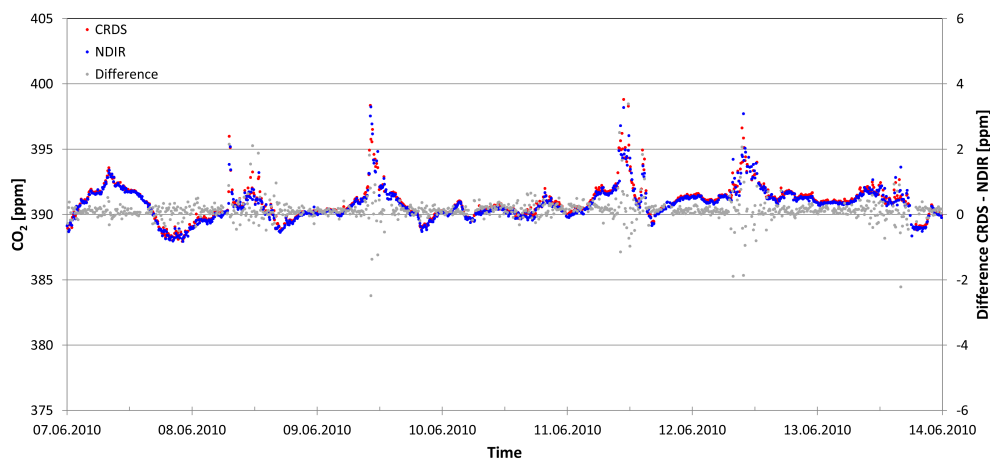


Figure 4. Time series of 115 s averaged CO₂ mole fractions measured with CRDS (red) and NDIR (blue) at JFJ (left-hand y axis scale) and the difference between the two systems against time (grey points) (right-hand y axis scale) for 1 week in June 2010.

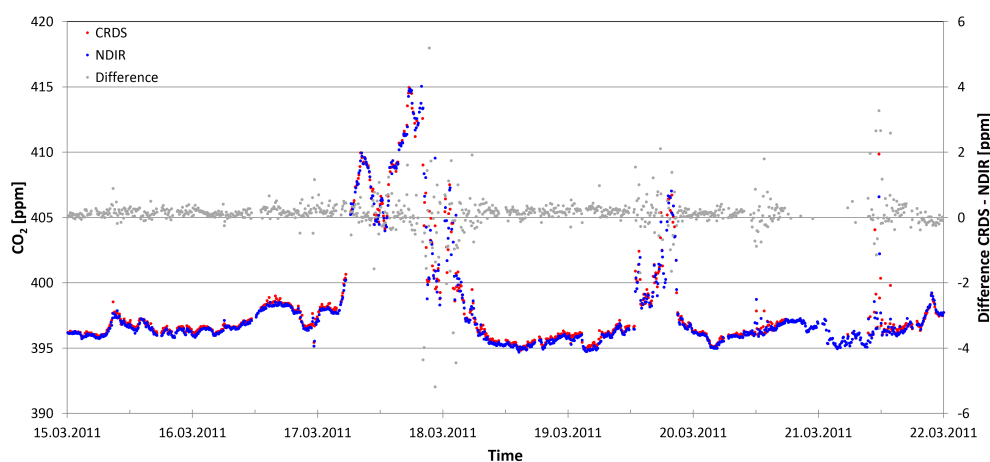


Figure 5. Time series of 115 s averaged CO₂ mole fractions measured with CRDS (red) and NDIR (blue) at JFJ (left-hand y axis scale) and the difference between the two systems against time (grey points) (right-hand y axis scale) for 1 week in March 2011.

estimated compared to the CRDS measurements. By recalculating measurements of the working gas using only the low and the high span (Fig. 6, black points), it is possible to approximate the evolution of the working gas CO₂ enrichment (Fig. 6, grey points) and correct for it (Fig. 7). The NDIR data are in much better agreement with the CRDS measurements after applying this correction. Comparing the uncorrected values of the B23 period to the corrected period, the R^2 of the NDIR values against the CRDS values increases from 0.89 to 0.99 (Fig. 8). The remaining working gases cylinders are recalculated in the same manner. Most of them show similar behavior as B23, even when they are replaced at a remaining pressure of around 25 bar. The only difference is the intensity of the enrichment, which is weaker than with B23 because they did not run completely empty. Note that for B23 we use time instead of pressure as the dependent variable because the pressure reader did not work properly

during this period. Although not preferred, this is possible because the pressure decreased linearly long-term.

3.3 Polynomial correction of the NDIR Maihak measurements

NDIR spectrometers are nonlinear in response and hence require a correction of the data – generally of the polynomial form. Ideally, the analyzer response function is implemented by the producer, mostly for a restricted mole fraction range. This is also the case for the Maihak system. The Maihak measurements at JFJ are done in differential mode; thus the reference cell is always flushed with gas of known CO₂ mole fraction and compared to the cell with the sample. To set the polynomial function to overcome the nonlinearity of the device, the mole fraction of the reference gas has to be entered into the analyzer's system.

Table 4. The statistics of different cutoff criteria of rapid CO₂ mole fraction changes and their influence on the number of points remaining, the average of the hourly differences of the NDIR and the CRDS system, the standard deviation of the differences and the according standard errors. The differences between the cutoff criteria are most probably caused by different volumes and flow rates of the systems and small time lags in the time stamps. The measurement precision of each system is about 10 times better than the standard deviation of the differences.

Cutoff criteria [ppm h ⁻¹]	Number of points	Average of differences [ppm]	Standard deviation [ppm]	Standard error [ppm]
All	19779	0.038	0.401	0.0028
< 2	19260	0.044	0.363	0.0026
< 1	17474	0.049	0.336	0.0025
< 0.75	16094	0.053	0.322	0.0025
< 0.5	13602	0.056	0.312	0.0027
< 0.25	8666	0.060	0.295	0.0032

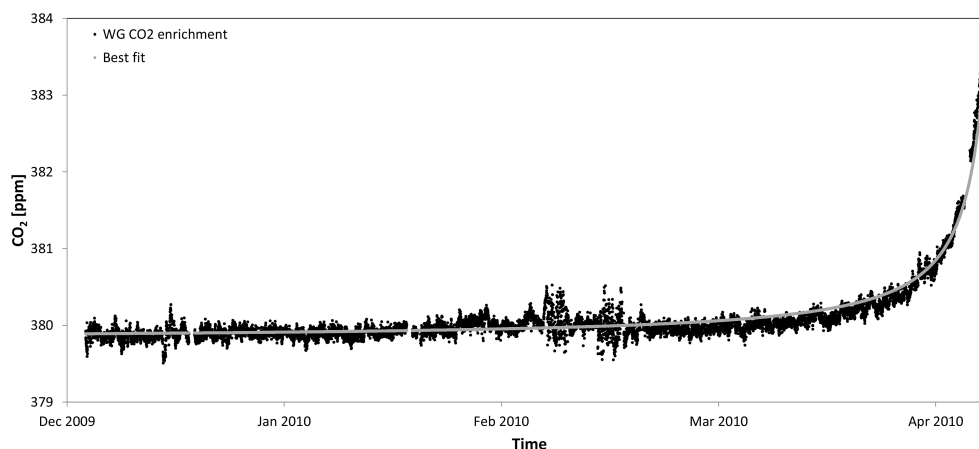


Figure 6. The CO₂ values of the calibration gas cylinder B23 calculated by using only the low and the high span cylinders against time (black points) and the calculated best fit function according to Langmuir (grey points) with $a = 5\,715\,797$ and $b = 568\,897$. The CO₂ values increase with time (corresponding to decreasing pressure). The best fit function was used to correct the measured working gas CO₂ values of the NDIR system.

In fall 2011, 11 used calibration gas cylinders of the KUP system were measured with the CRDS system and remeasured with the NDIR system. The remeasuring of these cylinders reveals that the polynomial function of the Maihak is insufficient. The second order polynomial of the difference between CRDS (Empa) – NDIR (KUP) as a function of CRDS (Empa) yields a R^2 of 0.86. A very similar polynomial dependence is found between the assigned values of NDIR (KUP) and the assigned values according to Table 2 when discarding two cylinders with low remaining pressure (Table 1). We prefer the first polynomial because the cylinders might have already faced desorption effects due to changing pressure between the assignment measurement in Bern (> 2 years before) and the remeasurement at JFJ, whereas the comparison with the CRDS took place within a few days. Since the cavity ring-down technique is linear in the considered mole fraction range, the factors of the polynomial are used to correct the measurement values of the NDIR system, removing this dependence completely (Fig. 9 and Table 1). The correction is only applied to the data measured by the

NDIR analyzer actually in use at JFJ. The data of the first NDIR analyzer, which had to be replaced in the beginning of 2011 due to technical problems, remain unaltered since the working points of the two NDIR instruments are different. There is also no possibility to correct the data of the older NDIR analyzer properly, because at the time there were no comparable measurements with a nearly linear measurement system. However, values in the center of this span, where most of the ambient CO₂ mole fractions were measured, should not have been affected strongly; roughly 75 % of all NDIR values are affected less than 0.1 ppm by the polynomial correction. The accuracy of the standard measurements is calculated as the square root of the sum of squared trueness and squared precision (Menditto et al., 2007) and estimated to be 0.08 ppm over the working range from 350 to 450 ppm based on the CRDS – NDIR comparison including both instrumental uncertainties.

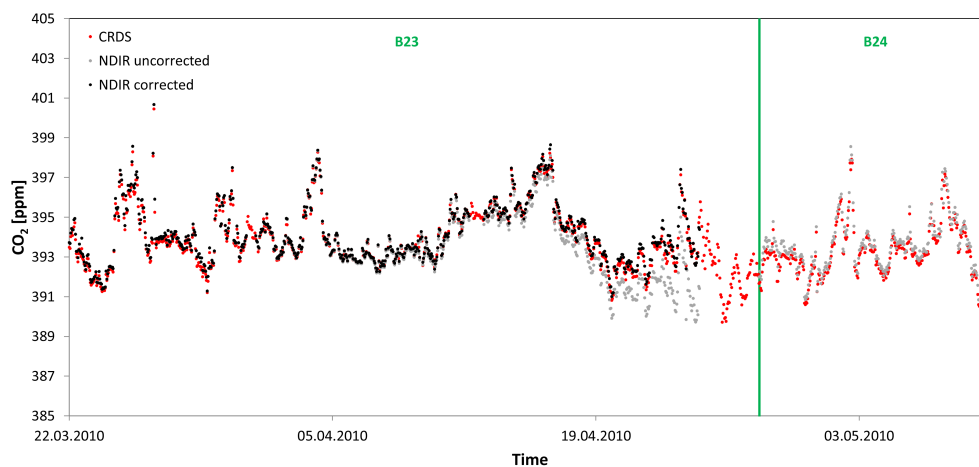


Figure 7. Time series of the CO₂ mole fractions determined with CRDS (red points) and NDIR (uncorrected in grey, after correction in black). The desorption-corrected NDIR CO₂ values show a much better agreement with the CRDS CO₂ values than the uncorrected, which are severely underestimated towards the end of the B23 period. After replacement of the working gas cylinder B23 with B24, the uncorrected NDIR CO₂ values show a much better agreement with the CRDS CO₂ values.

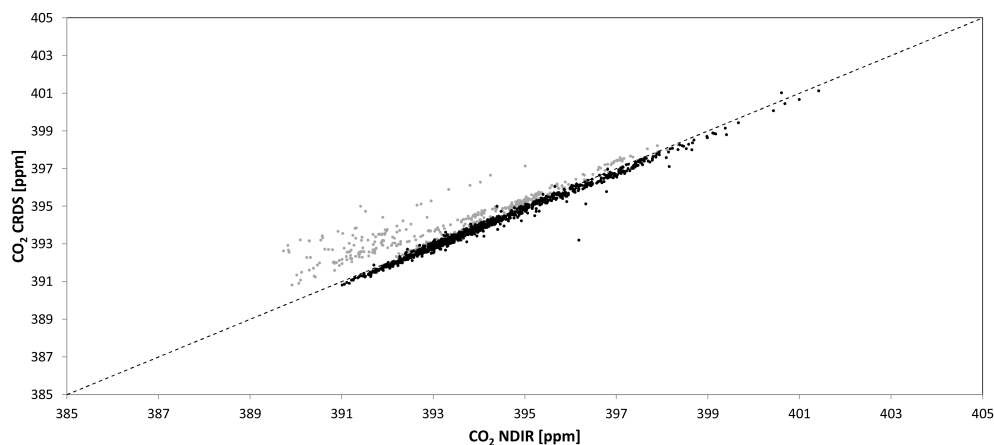


Figure 8. Correlation of ambient CO₂ mole fraction measured by CRDS versus the uncorrected CO₂ mole fraction of the NDIR instrument (grey points) and the desorption-corrected CO₂ NDIR mole fractions (black points), respectively during the period of the working gas cylinder B23. The diagonal represents the ideal 1 : 1 agreement. Due to the CO₂ enrichment in the working gas cylinder B23 the uncorrected KUP values seem to be much lower than the Empa values. By applying the correction, the values of the two systems are in a much better agreement.

4 Discussion

Long periods of overlapping measurements are helpful for assessing the compatibility of different analyzers. In this study we analyze two different data sets which were measured at the same station over 3 years (the simultaneous measurements are still ongoing) to determine the compatibility and quantify differences and biases. The compatibility is challenging since the Empa used the cavity ring-down technique (Picarro Inc.) and the KUP a non-dispersive infrared spectrometer (Maihak) to measure the CO₂ mole fraction of the air. Also, the calibration procedure, given the two techniques used, is different and therefore has to be investigated.

The seasonality based on monthly measured averages (10.01 ± 0.33 ppm and 10.05 ± 0.37 ppm for the CRDS and the NDIR data set) between the two data sets agree nicely and are also in good agreement with the results of flask measurements done at JFJ (10.54 ± 0.18 ppm) (van der Laan-Luijkx et al., 2013). The trend calculation (1.89 and 1.69 ppm year⁻¹) for the complete and common data periods deviates by 0.2 ppm year⁻¹. In the first half of this comparison, the difference of CRDS – NDIR is negative, indicating that the NDIR values are higher than the CRDS values. In the second half the differences are mainly positive, especially during the periods with the working gases B27 and B29. Additionally, changing the CRDS instrument caused a

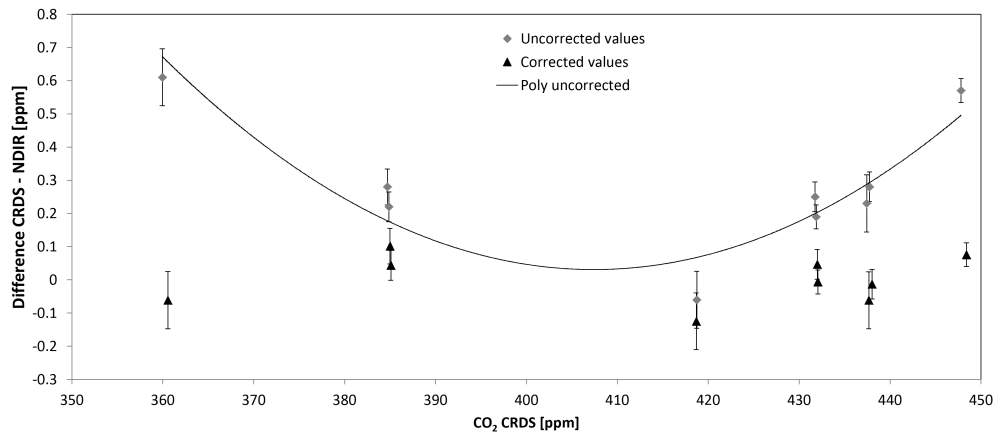


Figure 9. The difference between the Picarro (CRDS) – Maihak (NDIR) CO₂ values as a function of the CRDS values of old KUP calibration cylinders. The data show a polynomial dependency (grey diamonds); by correcting the data points with the according second degree polynomial (black line), the difference between the two systems varies slightly around zero (black triangles). The vertical error bars represent the combined standard deviation of both systems whereas the horizontal error bars represent only the standard deviation of the CRDS system. Because of the scale, the horizontal error bars are not visible.

small offset. This offset, together with the imperfectly assigned working gases, results in a slope of 0.2 ppm year⁻¹ for the differences and a slope difference of the same amount between the two CO₂ records. The annual CO₂ increase of the CRDS data set is in good agreement with the average global CO₂ trend of 1.97 ppm year⁻¹ for 2003–2012 (Tans, 2014), whereas the NDIR trend is a bit too low. Nevertheless, it is important to mention that for proper trend analysis the period of 3 years is too short. The CO₂ trend of the whole NDIR data set from the end of 2004 until the end of 2013 is 1.99 ppm year⁻¹, which is in much better agreement with the global trend of the same period of 2.03 ppm year⁻¹ (Tans, 2014).

Despite the fact that it is common knowledge to use well-assigned standard cylinder values for accurate measurements, we find larger offsets for two calibration cylinders, namely B27 and B29 used by the NDIR system. Therefore the calculated CO₂ values of the NDIR data set are a bit too low for these two periods. This effect would have been undetected without the comparison measurements, at least for the B27 cylinder: the offset is only around 0.2 ppm, which is quite small compared to a seasonality of roughly 10 ppm and even more so compared to the daily variations in extreme cases of up to 25 ppm. Experiences of ongoing international comparisons of traveling cylinders highlight similar issues (Zellweger et al., 2011). Therefore, a thorough check of assigned values of calibration and working cylinders is crucial.

Another problem concerning the calibration cylinders is their CO₂ mole fraction. Ideally the low span is lowest in CO₂ mole fraction, the working gas intermediate, preferably close to the expected value of the specimen, and the high span accordingly highest. However, in some periods (e.g., B26, B27, B29) the CO₂ mole fraction of the working gas is higher than, or very close to, the CO₂ mole fraction of the

high span (e.g., B28) and therefore not ideal for calibrating the measurements with a very high accuracy. However, because of logistic problems these cylinders have to be used as working gases.

Also, the stability of the CO₂ mole fraction in the calibration cylinders is very important. Adsorption/desorption effects occur when the pressure of steel cylinders drops under 20 to 30 bar (2 to 3 MPa) (Keeling et al., 1998), leading to an enrichment of CO₂ in the cylinder and therefore a CO₂ value different from the assigned. Hence, cylinders should not be used until they are empty. A recalculation of the working gas cylinders with a two-point calibration using the low and the high span show that the effect already starts at roughly 80 to 90 bar (8 to 9 MPa) and has to be corrected for. This finding is supported by Leuenberger et al. (2014). At higher pressures the effects are still very small, but with decreasing pressure the difference between the measured CO₂ mole fraction of the calibration gas and the assigned value increases, thus leading to a decrease in the accuracy of the calculated CO₂ values.

By applying the corrections mentioned above (see Sects. 3.2 and 3.3), discarding two periods of badly assigned working gases (B27 and B29) and excluding periods of fast CO₂ changes, the average difference of the hourly values decreases to $-0.03 \text{ ppm} \pm 0.25 \text{ ppm}$. Whereas the average value of almost zero is very good, the standard deviation is still high in comparison to the WMO goal of 0.1 ppm and the precision of each individual system, which is roughly five times better. The rather high standard deviation is probably caused by differing averaging intervals of the hourly values and dispersion of the air parcels in the two systems due to different volumes and flow rates of the air.

Both systems obtain the outside air from the same air inlet at Sphinx observatory at JFJ, but the air is led to the two

systems by separate tubing. Due to different flow rates and volumes of the two systems, the origin of an air parcel measured at a certain time is not necessarily exactly the same. For example, the volume of the NDIR drying unit leads to a residence time of 3 minutes, significantly longer than for the CRDS system. Therefore, NDIR system records slightly dampened ambient signals. Furthermore, it is possible that the according air parcels are shifted a little bit, thereby leading to small differences of the CO₂ values of the two systems. This could be the reason for the larger observed differences between the two data sets during fast CO₂ mole fraction changes of the outside air.

Overall, the CO₂ mole fraction measured by the two systems shows the same pattern not only during stable periods but also during extreme events and even in the high-resolution data. Most of the noisier periods can be traced to specific technical problems, whereas systematic shifts are most probably caused by badly assigned calibration gases or drifts of the CO₂ mole fraction in the cylinders due to adsorption/desorption effects. While it is possible to correct for the pressure-dependent drifts of the calibration gases mathematically, it is harder to correct for the shifts caused by badly assigned calibration gases since there are not a lot of independent comparison measurements that can be used for correction.

5 Conclusions

The CO₂ data sets of two different CO₂ measurement systems, a nondispersive infrared analyzer and a cavity ring-down spectrometer, running parallel at the High Altitude Research Station JFJ are compared. The evaluation of the two CO₂ records from the NDIR system of the KUP to the CRDS system of the Empa shows that the two systems are generally in good agreement, not only long-term but also down to the 1 Hz scale despite the fundamentally different measurement principles and gas handling systems. Therefore the two data sets can be used to complement each other. However, the comparison also reveals (i) adsorption/desorption effects in the calibration gas steel cylinders used by NDIR system, (ii) insufficient nonlinearity correction of the NDIR analyzer and (iii) periods of small biases because of badly assigned calibration gas cylinders. Adsorption effects can be corrected for by monolayer adsorption equation (Langmuir, 1918). Nonlinearity of the NDIR analyzer is constrained with a second order polynomial, resulting in better agreement between the two data sets, in particular for values strongly deviating from the average. More research in adsorption/desorption effects of calibration gas cylinders has been done at the KUP and is reported in Leuenberger et al. (2014).

Finally, we like to emphasize that using steel cylinders is not adequate for high-precision trace gas determinations due to gas adsorption/desorption. Therefore, we conclude that all laboratories using steel cylinders for calibration purposes

should switch to calibration gases stored in aluminium cylinders to minimize gas composition changes and allow an improved assignment of the cylinder.

It is helpful to continue this comparison exercise for another couple of years to show that once aluminium cylinders are in use for both systems, the compatibility will improve further.

Acknowledgements. This work was financially supported by the Swiss National Science Foundation (SNF-Project 200020_134641) and the Federal Office of Meteorology and Climatology MeteoSwiss in the framework of Swiss GCOS. Empa's CO₂ measurements are part of the Swiss National Air Pollution Monitoring network, a joint responsibility of Empa and the Swiss Federal Office for the Environment. We would like to thank the International Foundation High Altitude Research Stations Jungfraujoch and Gornergrat (HFSJG), especially the custodians Martin Fischer, Felix Seiler and Urs Oetz for changing the calibration gases cylinders and other maintenance work on the measurement systems. Additionally, the authors would like to thank Peter Nyfeler for his precious work and help in maintaining and repairing the systems in the laboratory in Bern and also at JFJ.

Edited by: C. Gerbig

References

- Arrhenius, S.: On the Influence of Carbonic Acid in the Air upon the Temperature of the Ground, *Philosophical Magazine and Journal of Science, Series 5*, 41, 237–276, 1896.
- Baltensperger, U., Gäggeler, H. W., Jost, D. T., Lugauer, M., Schwikowski, M., Weingartner, E., and Seibert, P.: Aerosol climatology at the high-alpine site Jungfraujoch, Switzerland, *J. Geophys. Res.*, 102, 19707–19715, 1997.
- Chevallier, F., Ciais, P., Conway, T. J., Aalto, T., Anderson, B. E., Bousquet, P., Brunke, E. G., Ciattaglia, L., Esaki, Y., Frohlich, M., Gomez, A., Gomez-Pelaez, A. J., Haszpra, L., Krummel, P. B., Langenfelds, R. L., Leuenberger, M., Machida, T., Maignan, F., Matsueda, H., Morgui, J. A., Mukai, H., Nakazawa, T., Peylin, P., Ramonet, M., Rivier, L., Sawa, Y., Schmidt, M., Steele, L. P., Vay, S. A., Vermeulen, A. T., Wofsy, S., and Worthy, D.: CO₂ surface fluxes at grid point scale estimated from a global 21 year re-analysis of atmospheric measurements, *J. Geophys. Res.-Atmos.*, 115, D21307, doi:10.1029/2010JD013887, 2010.
- Henne, S., Brunner, D., Folini, D., Solberg, S., Klausen, J., and Buchmann, B.: Assessment of parameters describing representativeness of air quality in-situ measurement sites, *Atmos. Chem. Phys.*, 10, 3561–3581, doi:10.5194/acp-10-3561-2010, 2010.
- Keeling, R. F., Manning, A. C., and McEvoy E. M.: Methods for measuring changes in atmospheric O₂ concentration and their application in southern hemisphere air, *J. Geophys. Res.*, 103, 3381–3397, 1998.
- Kohmyr, W. D., Harris, T. B., Waterman, L. S., Chin, J. F. S., and Thoning, K. W.: Atmospheric Carbon Dioxide at Mauna Loa Observatory – 1. NOAA Global Monitoring for Climatic Change Measurements With a Nondispersive Infrared Analyzer, 1974–1985, *J. Geophys. Res.*, 94, 8553–8547, 1989.

- Langmuir, I.: The Adsorption of Gases on Plane Surfaces of Glass, Mica and Platinum, *Journal of the American Chemical Society*, 40, 1361–1403, 1918.
- Leuenberger, M. C., Schibig, M. F., and Nyfeler, P.: Gas adsorption and desorption effects on cylinders and their importance for long-term gas records, *Atmos. Chem. Phys. Discuss.*, 14, 19293–19314, doi:10.5194/acpd-14-19293-2014, 2014.
- Menditto, A., Patriarca, M., and Magnusson, B.: Understanding the meaning of accuracy, trueness and precision, *Accred. Qual. Assur.*, 12, 45–47, 2007.
- Peters, W., Krol, M. C., van der Werf, G. R., Houweling, S., Jones, C. D., Hughes, J., Schaefer, K., Masarie, K. A., Jacobson, A. R., Miller, J. B., Cho, C. H., Ramonet, M., Schmidt, M., Ciattaglia, L., Apadula, F., Helta, D., Meinhardt, F., di Sarra, A. G., Piacentino, S., Sferlazzo, D., Aalto, T., Hatakka, J., Strom, J., Haszpra, L., Meijer, H. A. J., van der Laan, S., Neubert, R. E. M., Jordan, A., Rodo, X., Morgui, J. A., Vermeulen, A. T., Popa, E., Rozanski, K., Zimnoch, M., Manning, A. C., Leuenberger, M., Uglietti, C., Dolman, A. J., Ciais, P., Heimann, M., and Tans, P. P.: Seven years of recent European net terrestrial carbon dioxide exchange constrained by atmospheric observations, *Global Change Biol.*, 16, 1317–1337, 2010.
- Peterson, J. T., Kohmyr, W. D., Harris, T. B., and Chin, J. F. S.: NOAA carbon dioxide measurements at Mauna Loa Observatory, 1974–1976, *Geophys. Res. Lett.*, 4, 354–356, 1977.
- Ramonet, M., Ciais, P., Aalto, T., Aulagnier, C., Chevallier, F., Cipriano, D., Conway, T. J., Haszpra, L., Kazan, V., Meinhardt, F., Paris, J. D., Schmidt, M., Simmonds, P., Xueref-Remy, I., and Necki, J. N.: A recent build-up of atmospheric CO₂ over Europe. Part 1: observed signals and possible explanations, *Tellus Ser. B*, 62, 1–13, 2010.
- Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., Steinbacher, M., Sweeney, C., Wastine, B., and Zellweger, C.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air, *Atmos. Meas. Tech.*, 6, 837–860, doi:10.5194/amt-6-837-2013, 2013.
- Schimmel, D. S., House, J. I., Hibbard, K. A., Bousquet, P., Ciais, P., Peylin, P., Braswell, B. H., Apps, M. J., Baker, D., Bondeau, A., Canadell, J., Churkina, G., Cramer, W., Denning, A. S., Field, C. B., Friedlingstein, P., and Goodale, C.: Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems, *Nature*, 414, 169–172, 2001.
- Tans, P. P.: Annual Mean Global Carbon Dioxide Growth Rates, available at: <http://www.esrl.noaa.gov/gmd/ccgg/trends/>, last access: 27 May 2014.
- Uglietti, C., Leuenberger, M. C., and Valentino, F. L.: Comparison between real time and flask measurements of atmospheric O₂ and CO₂ performed at the High Altitude Research Station Jungfraujoch, Switzerland, *Sci. Tot. Environm.*, 391, 196–202, 2008.
- Uglietti, C., Leuenberger, M., and Brunner, D.: European source and sink areas of CO₂ retrieved from Lagrangian transport model interpretation of combined O₂ and CO₂ measurements at the high alpine research station Jungfraujoch, *Atmos. Chem. Phys.*, 11, 8017–8036, doi:10.5194/acp-11-8017-2011, 2011.
- van der Laan-Luijkx, I. T., van der Laan, S., Uglietti, C., Schibig, M. F., Neubert, R. E. M., Meijer, H. A. J., Brand, W. A., Jordan, A., Richter, J. M., Rothe, M., and Leuenberger, M. C.: Atmospheric CO₂, δ(O₂/N₂) and δ¹³CO₂ measurements at Jungfraujoch, Switzerland: results from a flask sampling intercomparison program, *Atmos. Meas. Tech.*, 6, 1805–1815, doi:10.5194/amt-6-1805-2013, 2013.
- WMO: 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques, WMO Report No. 194, World Meteorological Organization, Research Department, Atmospheric Research and Environment Branch, Geneva, Switzerland, 2011.
- Zellweger, C., Amman, M., Buchmann, B., Hofer, P., Lugauer, M., Rüttimann, R., Streit, N., Weingartner, E., and Baltensperger, U.: Summertime Noy speciation at the Jungfraujoch, 3580 m above sea level, Switzerland, *J. Geophys. Res.*, 105, 6655–6667, 2000.
- Zellweger, C., Forrer, J., Hofer, P., Nyeki, S., Schwarzenbach, B., Weingartner, E., Ammann, M., and Baltensperger, U.: Partitioning of reactive nitrogen (NO_y) and dependence on meteorological conditions in the lower free troposphere, *Atmos. Chem. Phys.*, 3, 779–796, doi:10.5194/acp-3-779-2003, 2003.
- Zellweger, C., Schwarzenbach, B., Steinbacher, M., and Buchmann, B.: System and performance audit of surface ozone, methane, carbon dioxide and carbon monoxide at the Global GAW station Cape Point South Africa, March 2011, WCC-Empa-Report 11/1, WMO, Geneva, Switzerland, 2011.