Atmos. Meas. Tech. Discuss., 5, 7293–7322, 2012
www.atmos-meas-tech-discuss.net/5/7293/2012/
doi:10.5194/amtd-5-7293-2012

Atmospheric Measurement Techniques Discussions

Atmospheric CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ measurements at Jungfraujoch, Switzerland: results from a flask sampling intercomparison program

I. T. van der Laan-Luijkx[1,2,*], S. van der Laan[1], C. Uglietti[1,], M. F. Schibig[1], R. E. M. Neubert[2], H. A. J. Meijer[2], W. A. Brand[3], A. Jordan[3], J. M. Richter[3], M. Rothe[3], and M. C. Leuenberger[1]**

[1]Climate and Environmental Physics, Physics Institute and Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
[2]Centre for Isotope Research, University of Groningen, Groningen, The Netherlands
[3]Max Planck Institute for Biogeochemistry, Jena, Germany
[*]currently at: Meteorology and Air Quality, Wageningen University, Wageningen, The Netherlands
[**]currently at: Byrd Polar Research Center, Ohio State University, Columbus, Ohio, USA

Received: 22 July 2012 – Accepted: 28 August 2012 – Published: 26 September 2012

Correspondence to: I. T. van der Laan-Luijkx (ivanderlaan@climate.unibe.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

AMTD
5, 7293–7322, 2012

Abstract

We present results from an intercomparison program of CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ measurements from atmospheric flask samples. Flask samples are collected on a bi-weekly basis at the High Altitude Research Station Jungfraujoch in Switzerland for three European laboratories: the University of Bern, Switzerland, the University of Groningen, the Netherlands and the Max Planck Institute for Biogeochemistry in Jena, Germany. Almost 4 yr of measurements of CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ are compared in this paper to assess the measurement compatibility of the three laboratories. While the average difference for the CO_2 measurements between the laboratories in Bern and Jena meets the required compatibility goal as defined by the World Meteorological Organisation, the standard deviation of the average differences between all laboratories is not within the required goal. However, the obtained annual trend and seasonalities are the same within their estimated uncertainties. For $\delta(O_2/N_2)$ significant differences are observed between the three laboratories. The comparison for $\delta^{13}CO_2$ yields the least compatible results and the required goals are not met between the three laboratories. Our study shows the importance of regular intercomparison exercises to identify potential biases between laboratories and the need to improve the quality of atmospheric measurements.

1 Introduction

Atmospheric measurements of greenhouse gases and related tracers are important for studies on the global carbon cycle and climate change research. The carbon cycle includes all processes involving the exchange of CO_2 between the atmosphere, oceans and terrestrial biosphere. $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ measurements[1] offer

[1]Throughout this paper, we follow the terminology recommendation from Coplen (2011). The term $\delta^{13}CO_2$ is used to denote $\delta^{13}C$ of CO_2 in air on the VPDB scale.

AMTD
5, 7293–7322, 2012

additional information on the exchange of CO_2 between the different reservoirs (Battle et al., 2000; Ciais et al., 1995; Keeling et al., 1993, 2011). Modelling studies use the atmospheric measurements from many globally spread locations to estimate carbon fluxes, which are subsequently used in climate models to understand and predict climate change. One of the major challenges in this field is to minimize the measurement uncertainties and especially to minimize the biases between laboratories and measurement locations. A bias between measurement stations can cause a large difference in the estimated carbon fluxes. For example, the data assimilation system CarbonTracker (Peters et al., 2007) yields considerably different results for the estimated surface fluxes if a constant bias is (artificially) introduced in the measurements of one single observation site. A linear relationship was found between the measurement bias introduced at one station and the obtained surface fluxes. This relationship is found to be 68 Tg C yr^{-1} for each 1 ppm of bias introduced in the CO_2 measurement record (Masarie et al., 2011).

To emphasize the importance of the quality of atmospheric measurements, the World Meteorological Organisation (WMO) has defined goals for the measurement compatibility of different atmospheric species. The goals are defined based on the required data quality for the use in e.g. inversion studies or the interpretation of large scale atmospheric data measured by different laboratories. The defined goals for CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ are ±0.1 ppm (0.05 ppm in the Southern Hemisphere), ±2 per meg and ±0.01 ‰, respectively (WMO, 2011). Within a single laboratory, this goal for CO_2 is reached by most laboratories with the present-day instrumentation. For $\delta^{13}CO_2$, the goal is not reached within all laboratories, as it is difficult to reach with currently available techniques. $\delta(O_2/N_2)$ measurements are in general very challenging. The absolute atmospheric variations of O_2 are in the same order as for CO_2, because they are stoichiometrically related. However, they have to be detected against a very high background of 21 % (e.g. Keeling, 1988), compared to the CO_2 background of about 0.04 %. The required goal for the precision of $\delta(O_2/N_2)$ measurements of 2 per meg corresponds to a relative precision of about 0.0002 % and is currently not yet reached

AMTD
5, 7293–7322, 2012

by the laboratories able to perform high-precision $\delta(O_2/N_2)$ measurements. The consistency for $\delta(O_2/N_2)$ measurements between any two laboratories is at the moment not better than ±5 per meg. While an international scale for $\delta(O_2/N_2)$ measurements is not yet available, most laboratories use the scale provided by the Scripps Institution of Oceanography, United States (SIO). This scale is also used in this paper.

To improve the quality of atmospheric measurements and to verify that measurements at different locations, by different laboratories, are not biased by the used sampling methods, materials, analytical techniques and calibration strategies and scales, intercomparison programs between different laboratories have been started (e.g. Manning et al., 2009; Masarie et al., 2001; WMO, 2011). These programs are used to assess the compatibility between laboratories and measurement locations. In these programs, either real air samples or sets of cylinders containing different concentrations are used. Specific intercomparison projects of in-situ observations by different laboratories are rare. This "super-site" approach requires that flasks are filled with air at the same time and location using the individual sampling protocols of different laboratories and that the flask measurements are performed in the different laboratories. Especially for $\delta(O_2/N_2)$ measurements, there are limited studies on this kind of compatibility. The first "super-site" intercomparison program for $\delta(O_2/N_2)$ measurements was started in 1991 at Cape Grim, Tasmania, Australia by three laboratories: the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia, the University of Rhode Island, United States and SIO (Battle et al., 2006; Langenfelds et al., 1999). The main global intercomparison program for $\delta(O_2/N_2)$ measurements is the Global Oxygen Laboratories Link Ultra-precise Measurements (Gollum) program, in which sets of 3 cylinders are shipped around the world that are measured in the 11 laboratories currently able to perform high precision $\delta(O_2/N_2)$ measurements (http://gollum.uea.ac.uk). Furthermore, another "super-site" intercomparison program is on-going at Alert, Canada, including $\delta(O_2/N_2)$ analyses by SIO and the Max Planck Institute for Biogeochemistry in Jena, Germany (MPI).

AMTD
5, 7293–7322, 2012

In 2007, three European laboratories have started a new intercomparison project at the High Altitude Research Station Jungfraujoch in Switzerland. Flasks are filled on a bi-weekly basis for the laboratories of the University of Bern, Switzerland (UBE), the University of Groningen, the Netherlands (RUG) and MPI. For each laboratory, the flasks filled at Jungfraujoch are identical to the flasks these laboratories use for their own respective field stations. This has yielded unique datasets for the comparison of three different atmospheric species by three laboratories.

This paper first describes the sampling location, sampling procedures, and measurement techniques. Subsequently the results of the measurements of CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ are presented and discussed.

2 Methods

2.1 Sampling location

The High Altitude Research Station Jungfraujoch is located at 7°59′20″ E and 46°32′53″ N in the Swiss Alps. It is situated at an altitude of 3580 m a.s.l. on a mountain saddle between the mountains Jungfrau and Mönch (http://www.ifjungo.ch). Due to its high elevation the station is most of the time situated above the planetary boundary layer and the air is mainly influenced by the free troposphere, representing atmospheric background conditions of continental Europe. A flask sampling program has been started on site in 2000 by the University of Bern, initially on a bi-weekly basis, and later on the frequency was increased to weekly sampling. The sampling program has been extended with the additional bi-weekly sampling for the other two laboratories in this intercomparison program in December 2007. The flask filling usually takes place on (Friday) mornings around 07:00 a.m. LT to make sure that the samples represent clean background air and to minimize the influence of uplifted air masses from the boundary layer (Uglietti et al., 2008).

AMTD
5, 7293–7322, 2012

2.2 Flask types

For this intercomparison program, glass flasks are filled every 2 weeks with ambient air at Jungfraujoch for the three participating laboratories. Each laboratory uses its proprietary flasks with slightly different designs. The UBE flasks are 1 l glass flasks with two valves each placed at one end of the flask. The flasks are fitted with glass valves from Louwers (Hapert, the Netherlands) with Viton O-rings. The RUG glass flasks have identical valves, but the design is different in that the valves are situated on the same side of the flask. One of the valves is assigned to be the inlet of the flask. On this side a dip tube is placed inside the flask which is connected to the inlet, so that the air always flushes the entire flask. The volume of the RUG flasks is 2.5 l. The MPI flasks are 1 l glass flasks with two valves, one on each end of the flask. The valves have seals made of Kel-F (PCTFE). More details about the flasks, valves and seals are presented by Sturm et al. (2004) and Rothe et al. (2005).

2.3 Flask sampling

Since the end of 2007, flasks are filled every 2 weeks using dedicated flask sampling units. In the intermediate weeks, flasks are filled for UBE only. For this paper, we have included flasks filled between December 2007 and August 2011, which amounts to 96 different sampling dates. Flasks are filled in pairs for both UBE and RUG, and in triplicates for MPI. The design of the flask sampling system has been changed during the course of intercomparison project. Before March 2009, all flasks were connected in series in the following order: MPI – UBE – RUG, using a single pump. From March 2009 onwards, two parallel filling setups are used: the MPI flasks are filled using a dedicated pump and the UBE and RUG flasks are using a common pump (KNF Neuberger) to fill the flasks in series. Prior to sampling, the air is dried using U-shaped glass tubes filled with anhydrous magnesium perchlorate ($Mg(ClO_4)_2$) and sealed with glass wool plugs. Dedicated intake lines are used for the flask filling, which consist of 15 m PVC tubing connected to the sampling units with Decabon tubing. To completely flush the entire

AMTD
5, 7293–7322, 2012

volumes, the flasks are flushed for about 30 min using a flow of about 2–3 l min^{-1}. The flasks are flushed and filled to a pressure of 1600 hPa for MPI and 950 hPa for UBE and RUG, while the average air pressure at Jungfraujoch is about 650 hPa. After the filling procedure, the flasks are transported back to the respective laboratories. For UBE and RUG this is done in batches of multiple flasks, leading to a storage time of the flasks at Jungfraujoch in the order of a couple of weeks. The difference between the pressure in the flasks and the local air pressure (also during the waiting time in the laboratories) can affect the concentrations of the air in the flasks, especially the $\delta(O_2/N_2)$ values, by permeation through the o-rings used to seal the flasks. This effect was studied by Sturm et al. (2004) and leads to an increased difficulty to meet the compatibility goals for $\delta(O_2/N_2)$.

2.4 Measurement techniques

After the filling procedure at Jungfraujoch, the flasks are measured in their respective laboratories. For the CO_2 measurements, the method used at UBE is different from the methods used at both RUG and MPI. At RUG and MPI the CO_2 concentration is measured using a Hewlett-Packard Gas Chromatograph (GC), model 6890, comparable to the setup described by Worthy et al. (2003) and van der Laan et al. (2009). More details are presented by Sirignano et al. (2010) for RUG and Jordan and Brand (2003) for MPI. In Bern, the CO_2 concentration is measured simultaneously with the $\delta(O_2/N_2)$ values using mass spectrometry. In this case, the CO_2 is also measured as the ratio of CO_2 to N_2 and the obtained δ-value is converted to a CO_2 concentration using the known CO_2 concentration of the machine reference gas. A correction factor is applied to correct for the N_2O background value produced in the ion source due to sample nitrogen and oxygen reactions. More details about this method are presented by Leuenberger et al. (2000b).

The $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ measurements are performed in all three laboratories using mass spectrometry. For $\delta(O_2/N_2)$ dual inlet isotope ratio mass spectrometers

AMTD
5, 7293–7322, 2012

(DI-IRMS) are used in a similar manner as described by Bender et al. (1994). UBE and MPI use a Finnigan MAT DELTA plus XL/XP from Thermo Electron (Bremen, Germany) and RUG uses a Micromass Optima (Micromass, now Elementar Manchester, UK). More details about the specific measurements in each laboratory are described by Leuenberger et al. (2000a) for UBE, van der Laan-Luijkx et al. (2010) for RUG and Brand (2005) for MPI.

$\delta^{13}CO_2$ is measured as the last of the three species presented in this paper, since the CO_2 is first extracted from the air sample before the analysis takes place. At UBE a Finnigan MAT DELTA XL mass spectrometer is combined with a GC column. CO_2 is extracted online from the air sample with liquid nitrogen and the column is used to separate N_2O from the CO_2. At RUG, a second Micromass Optima is used. The CO_2 is extracted from the air sample with liquid air, and a correction is applied for the co-trapped N_2O. At MPI a Finnigan MAT mass spectrometer is used in combination with the custom developed BGC-AirTrap to separate CO_2 from the air sample. More details are described by Sturm et al. (2006) for UBE, Sirignano et al. (2004) for RUG and Werner et al. (2001) for MPI.

3 Results

3.1 CO_2

For intercomparing CO_2 abundance measurements at the different laboratories, results from 96 filling dates have been included in the analysis. For some dates not all 3 laboratories have valid flask results, due to e.g. logistical problems, measurement issues or leaking flasks. Flask results that were obviously influenced by measurements problems or leakages have been removed from the data set. For each laboratory, the resulting amount of sampling dates with valid results for the CO_2 concentrations are: 90 for UBE, 84 for RUG and 82 for MPI. For UBE, on 80 dates 2 flasks have been used to obtain an average value, for 10 dates there was only 1 valid flask. For RUG, we included 75

AMTD
5, 7293–7322, 2012

values based on the average of 2 flasks and 9 are measurements of a single flask. For MPI, 64 values are averages of 3 flasks, 16 are averages of 2 flasks and for 2 sampling dates only 1 flask was included. For the sampling dates with more than 1 valid flask, the average standard errors in the mean of the duplicate or triplicate flasks are 0.05 ppm for UBE, 0.06 ppm for RUG and 0.06 ppm for MPI (see also Table 1). This is well within the WMO goal for compatibility of 0.1 ppm.

Figure 1 shows the results for the CO_2 measurements of the flasks sampled at Jungfraujoch. As indicated above, these values represent average data of 2 or 3 flasks, or the single value of sampling dates with only 1 valid flask sample. The fits shown in the figure are linear trends and double harmonic seasonal components and do not include those points that are considered outliers of the fit, based on a 2.7 sigma exclusive filter of the residuals. This filter excludes 4 values for UBE, 3 for RUG and 3 for MPI. From the figure it is clear that the flasks from the three laboratories follow the same trend as well as seasonality. In some cases, all three laboratories show a value far away from the fit, but the three data points are close together. These data represent e.g. local or nearby pollution events. There are also sampling dates with large differences between the values obtained by one laboratory compared to the other two, most likely due to e.g. measurement issues or small flask leakages.

Figure 2 shows the differences between each set of two laboratories. The figure includes also an indication of the mean differences. The average values of the differences and their standard deviations are shown in Table 2. The difference between the measurements of UBE and MPI is the smallest. This is true for both the absolute value of the difference as well as the standard deviation of the average difference, which is smaller than for the other two comparisons. The RUG values are slightly lower than the values from the other two laboratories. Although the mean difference between UBE and MPI of 0.08 ppm is within the WMO compatibility goal, the majority of the calculated differences is outside of this range. If we start from the obtained average difference values, and then apply the 0.1 ppm accepted deviations, only 34 % of the UBE-MPI differences are within these limits. For UBE-RUG this is 21 % and for MPI-RUG this

AMTD
5, 7293–7322, 2012

is 24 %. We therefore conclude that these flask measurements do not yet meet the required compatibility goals for CO_2.

As stated in Sect. 2.3, the sampling setup has been changed in March 2009. Before this date, the flasks from all three laboratories were sampled in series. After that, the flasks from MPI are sampled parallel to those of UBE and RUG, which are sampled in series. The results for the average differences between the laboratories before and after this change are included in Table 2. From these values it is clear that the standard deviations of the average values increase from the first to the second period. The higher standard deviation could imply that the new sampling procedure has introduced a larger difference between the laboratories. However, from the results shown in Fig. 2, it is not clear that the bigger difference is introduced directly in March 2009. Larger variations are visible in the periods summer/autumn 2009 as well as between June 2010 and February 2011. From this data we cannot assign a bias between the results from the three laboratories due to the changed setup.

The fits and derived fit parameters for annual trends and seasonality for the individual data series from each laboratory are shown in Table 3. The average annual trend obtained from the data sets are 1.76 ± 0.17 ppm yr^{-1} for UBE, 1.94 ± 0.18 ppm yr^{-1} for RUG and 1.83 ± 0.17 ppm yr^{-1} for MPI. Within their estimated uncertainty ranges these values correspond well to each other. The average of these values is 1.85 ± 0.09 ppm yr^{-1}. For the seasonal amplitudes, the three results also agree within their error bars, although the UBE result is, on the edge of significance, lower than the other two. The average value for the amplitudes is 10.54 ± 0.18 ppm, representing low seasonal variations as expected for the high altitude continental background station Jungfraujoch. Seasonalities at other European sampling locations are more pronounced due to local and regional influences of the biosphere and fossil fuel combustion (e.g. Thompson et al., 2009; van der Laan et al., 2010).

AMTD
5, 7293–7322, 2012

3.2 $\delta(O_2/N_2)$

The amount of samples included in the analysis for $\delta(O_2/N_2)$ is 86 for UBE, 87 for RUG and 79 for MPI. For UBE, 79 values are averages of 2 flasks and for 7 sampling dates only 1 valid flask was used. For RUG there are 74 averages of 2 flasks and 13 single flask measurements. For MPI 48 values are averages of 3 flasks, 23 are averages of 2 flasks and 8 are single flask values. The standard errors of the $\delta(O_2/N_2)$ values obtained from the averages of 2 or 3 flasks are shown in Table 1. These are 6 per meg for UBE, 8 per meg for RUG and 3 per meg for MPI. Comparing this to the required WMO goal for compatibility, we conclude that none of our three laboratories meets the required accuracy needed to reach the compatibility goal of 2 per meg. The WMO states in its report that the goal of 2 per meg is not yet reached and that the compatibility between any two laboratories is not yet better than 5 per meg. The internal reproducibility for our flask samples is below 5 per meg only for MPI, the other two laboratories do not yet meet this range.

Figure 3 shows the results for the $\delta(O_2/N_2)$ values of the atmospheric samples for the three laboratories. The error bars indicated in the figure are the standard errors of the mean of the results of 2 or 3 flasks. Values that represent only a single flask are not assigned an error bar. Using the 2.7 sigma residuals filter as described in Sect. 3.1, 1 value is rejected for UBE, 4 for RUG and 4 for MPI. The figure shows a large variability between the $\delta(O_2/N_2)$ values for the three laboratories. Samples that represent local pollution events, as seen in Fig. 1 for CO_2 are not recognisable as such for $\delta(O_2/N_2)$, due to the higher variability in the data sets.

Figure 4 shows the differences between each set of two laboratories. The average values for the differences are indicated in the figure and included in Table 2. In Fig. 3 it is visible that the $\delta(O_2/N_2)$ values for UBE are significantly lower than the values of the other two laboratories. This offset can probably be explained by a problem with the scale definition for UBE. The average difference between MPI and RUG is −3 per meg, whereas for UBE-RUG it is −33 per meg and for UBE-MPI it is −31 per meg. Also

AMTD
5, 7293–7322, 2012

the standard deviation of the average difference is larger for the comparisons to UBE than between MPI and RUG. For MPI-RUG the average difference is within 5 per meg, however, only 16 % of the values are within 5 per meg from the average difference. For UBE-RUG, 13 % are within 5 per meg of the average difference and for UBE-MPI this number is 18 %. For $\delta(O_2/N_2)$, significant improvements of the sampling procedures, the storage of the flasks and the measurements in the three laboratories are needed to meet the WMO goals. Table 2 also includes the difference between the measurements of the samples collected before and after March 2009. The obtained values do not show a change based on the modification in the setup.

The indicated fits for the 3 data sets in Fig. 3 are quite different from each other. The obtained parameters for each laboratory are given in Table 3. The data sets cover almost four years, which is a short time to obtain robust values for the long term annual trend, considering the large variability in the data sets. The seasonalities of the fits should be comparable between the three laboratories based on this time period. The large variability of the $\delta(O_2/N_2)$ data does however lead to significant differences. The quality of the obtained fits and estimates for the trend and seasonal amplitudes are significantly different for each laboratory. The correlation coefficients R^2 are 0.58 for UBE, 0.73 for RUG and 0.87 for MPI. The obtained values for the annual decrease rates differ significantly as well. Especially for the UBE data, the trend estimate is unrealistic, due to the high variability of the data set. Since the focus of this study is the comparison between the measurements of different laboratories, we have included most of our data in our analysis. However, if this data set would be used for trend analysis, a stronger filtering strategy could be applied. If a 1.9 sigma exclusive filter would be used, instead of the used 2.7 sigma filter (see Sect. 3.1), the trend estimate for UBE would become more robust at: -21 ± 2 per meg yr^{-1} (with $R^2 = 0.81$). For RUG and MPI the trend estimates are already more robust (given the higher initial R^2-values), and removing more data points does not alter the trends estimates that much. For the seasonality, the obtained values for the amplitude compare well between MPI and RUG, 85 ± 4 per meg and 84.1 ± 2.2 per meg respectively. This value is, as for CO_2, lower than at other

AMTD
5, 7293–7322, 2012

stations which are in the European atmospheric boundary layer (e.g. Kozlova et al., 2008; Popa et al., 2010; Thompson et al., 2009; van der Laan-Luijkx et al., 2010). The value for Jungfraujoch represents a signal of a background station influenced mostly by the free troposphere.

3.3 $\delta^{13}CO_2$

For the analysis of $\delta^{13}CO_2$ we have included 88 values for UBE, 82 for RUG and 67 for MPI. For UBE, 75 are averages of the values of two flasks and 13 are single flask measurements. For RUG, 53 are averages of two values and 29 are single flasks. For MPI 53 values are averages of three flasks, 10 are averages of 2 flasks and 4 values are single values. The standard errors of the averages for the duplicate and triplicate samples are 0.08 ‰ UBE, 0.07 ‰ RUG, 0.009 ‰ MPI (see Table 1). The WMO compatibility goal of 0.01 ‰ is only met by MPI, the other two laboratories are far above the prescribed goal.

Figure 5 shows the results for the $\delta^{13}CO_2$ measurements from flasks sampled at Jungfraujoch. The standard errors of the averaged values are indicated as error bars. For single flask values no error bar is included in the figure. Filtering the data using the method described above, removes 3 values for UBE, 5 for RUG and 1 for MPI. The figure shows the seasonality in the $\delta^{13}CO_2$ signal as well as a small decreasing trend. The decrease rate is not clearly visible due to the short time scale. The results from the three laboratories follow the same pattern. The fits shown in the figure are linear trends and single harmonic seasonal components.

Figure 6 shows the differences between the laboratories. The average differences are close together as seen in the figure as well as in Table 2. However, the variability for each comparison is quite large. The average differences are -0.03 ± 0.04 ‰ for UBE-RUG, -0.02 ± 0.03 ‰ for UBE-MPI and -0.02 ± 0.03 ‰ for MPI-RUG. This result makes clear that the WMO goal for $\delta^{13}CO_2$ is not met between any of the three laboratories. The percentage of measurements within the WMO goal of 0.01 ‰ from the obtained average differences are 4 % for UBE-RUG, 9 % for UBE-MPI and 10 % MPI-RUG. The

AMTD
5, 7293–7322, 2012

compatibility of the $\delta^{13}CO_2$ results from the flasks sampled at Jungfraujoch for our three laboratories, therefore should be taken as a strong indicator for the limited possibility of interpreting the data series presented here. Table 2 also includes the values obtained before and after March 2009. These values do not show a difference due to the change in the setup.

The obtained parameters for the trend and seasonality are presented in Table 3. The results from the three laboratories do not compare well with each other within their estimated uncertainties. The trend estimates for UBE and RUG of -0.081 ± 0.018 ‰ yr^{-1} and -0.069 ± 0.015 ‰ yr^{-1} are much too high compared to the estimate obtained for MPI of -0.016 ± 0.014 ‰ yr^{-1}. The latter is in good agreement with the trend from the GLOBALVIEW-CO2C13 dataset, which is also -0.02 ‰ yr^{-1} for our latitude (GLOBALVIEW-CO2C13, 2009). The fact that the intra-laboratory precision of MPI is much better than the other two laboratories (see Table 1), enables this better trend estimate on the relatively short time-scale of four years. The other two laboratories would need a longer data record to obtain a valid trend estimate. For UBE, additional flasks are sampled at Jungfraujoch and data from these flasks is available for the entire period 2000–2012. The obtained trend from the complete UBE record is estimated at -0.013 ± 0.004 ‰ yr^{-1}, much closer to the trend estimate from MPI. The average seasonal amplitude for the three laboratories is 0.51 ± 0.07 ‰, which is lower than obtained from other European stations, e.g. the obtained seasonal amplitude from the GLOBALVIEW-CO2C13 dataset for $\delta^{13}CO_2$ for our latitude is 0.7 ‰ (GLOBALVIEW-CO2C13, 2009), indicating again that Jungfraujoch is less influenced by regional and local emissions.

4 Discussion, conclusion and recommendations

The study presented in this paper covers a long-term comparison of measurements of in-situ sampled flasks for CO_2, $\delta(O_2/N_2)$ as well as $\delta^{13}CO_2$. Intercomparison programs are important to document the interlaboratory compatibility, to indicate the need

AMTD
5, 7293–7322, 2012

of improvement and to detect measurement problems in specific laboratories. Global intercomparison programs are however rare, since they are time-consuming and costly. The global programs that do exist are primarily focussed on the intercomparison of cylinder measurements or flask filled under laboratory conditions. In intercomparison programs such as the Cucumber project (Manning et al., 2009) cylinders are shipped between laboratories to compare their measurements. The cylinders are therefore measured in each laboratory about once per year. The compatibility between laboratories under laboratory conditions is quite different from a field study, since biases can be introduced not only by the measurements, but also in the sampling procedure or the storage in the different flasks. Our study has shown that our three laboratories do not yet meet the required WMO goals for compatibility for the presented flask sampling program. However, it is important to note that Jungfraujoch is a very challenging measurement location, especially for $\delta(O_2/N_2)$ due to its low air pressure and that the compatibility between the laboratories is better when based on e.g. the Cucumber program. The quality of flask sample data is very relevant, as flask sampling is used at many sampling locations because this is easier to achieve at remote locations and it enables multiple sampling locations in terms of cost effectiveness compared to continuous measurements. Flask samples are therefore widely used in carbon cycle studies. Further efforts should be made to increase internal reproducibility of laboratories as well as the compatibility between laboratories for this sampling method. Our intercomparison program is therefore an important tool to assess inconsistencies, which is the first step to be able to minimize them.

Especially for $\delta(O_2/N_2)$ measurements, intercomparison programs are rare. The desired high precision and accuracy for $\delta(O_2/N_2)$ measurements is reached by only a few laboratories. $\delta(O_2/N_2)$ is difficult to measure, therefore more collaboration and intercomparisons are needed to establish better compatibility between laboratories. Combined trend analysis of CO_2 and $\delta(O_2/N_2)$ is an important tool to study the global oceanic CO_2 uptake. Differences in obtained CO_2 and $\delta(O_2/N_2)$ trends between laboratories can therefore have a large impact on these estimates. The global oceanic CO_2

AMTD
5, 7293–7322, 2012

uptake is for example estimated by Manning and Keeling (2006) and van der Laan-Luijkx et al. (2010), who found $2.2 \pm 0.6\,\mathrm{PgC\,yr^{-1}}$ and $1.8 \pm 0.8\,\mathrm{PgC\,yr^{-1}}$ respectively. Using the same approach as van der Laan-Luijkx et al. (2010), we obtain from our data the following estimates for the global oceanic CO_2 uptake: $6.4 \pm 1.7\,\mathrm{PgC\,yr^{-1}}$ for UBE ($3.0 \pm 1.2\,\mathrm{PgC\,yr^{-1}}$ when using the more strict data filtering as described in Sect. 3.2), $3.6 \pm 1.4\,\mathrm{PgC\,yr^{-1}}$ for RUG and $1.5 \pm 1.0\,\mathrm{PgC\,yr^{-1}}$ for MPI. These large differences are mainly caused by the large differences in the $\delta(O_2/N_2)$ trend between the three laboratories (see Table 3). These values are based on only short time series, and can therefore be significantly improved by extending the data series. Longer time series are therefore necessary before these estimates can be used in a study to obtain the global oceanic CO_2 uptake. However, our estimates do show that differences between measurements of different laboratories can have a large impact on global carbon cycle estimates and therefore reflect that the ambitious WMO compatibility goals have a scientific justification. Laboratories should continue to improve their measurement precision and accuracy and continue to assess them in regular intercomparison programs.

Acknowledgements. The authors would like to thank the custodians of the High Altitude Research Station Jungfraujoch, specifically Martin Fischer and Felix Seiler, for their efforts in filling the flasks on a bi-weekly basis, maintaining the flask sampling units and the laboratory and for taking care of the flask logistics. The authors furthermore thank Peter Nyfeler for his work in the UBE laboratory to maintain the mass spectrometers and organising the flask measurements. The following staff members from RUG are acknowledged for their respective contributions to the flask logistics, measurements and data analysis: Bert Kers, Jan Schut, Henk Jansen, Anita Aerts-Bijma, Janette Spriensma and Hans Roeloffzen. The authors also thank Bert Steinberg from MPI for his contributions to the flask measurements.

AMTD
5, 7293–7322, 2012

References

Battle, M., Bender, M. L., Tans, P. P., White, J. W. C., Ellis, J. T., Conway, T., and Francey, R. J.: Global carbon sinks and their variability inferred from atmospheric O_2 and $\delta^{13}C$, Science, 287, 2467–2470, 2000.

Battle, M., Fletcher, S. E. M., Bender, M. L., Keeling, R. F., Manning, A. C., Gruber, N., Tans, P. P., Hendricks, M. B., Ho, D. T., Simonds, C., Mika, R., and Paplawsky, B.: Atmospheric potential oxygen: New observations and their implications for some atmospheric and oceanic models, Global Biogeochem. Cy., 20, GB1010, doi:10.1029/2005gb002534, 2006.

Bender, M. L., Tans, P. P., Ellis, J. T., Orchardo, J., and Habfast, K.: A High-Precision Isotope Ratio Mass-Spectrometry Method for Measuring the O_2/N_2 Ratio of Air, Geochim. Cosmochim. Ac., 58, 4751–4758, 1994.

Brand, W. A.: O_2/N_2 Storage Aspects and Open Split Mass Spectrometric Determination, in: Proceedings of the 12th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques, WMO/TD-No. 1275, edited by: Worthy, D. and Huang, L., WMO, Geneva, 2005.

Ciais, P., Tans, P. P., White, J. W. C., Trolier, M., Francey, R. J., Berry, J. A., Randall, D. R., Sellers, P. J., Collatz, J. G., and Schimel, D. S.: Partitioning of Ocean and Land Uptake of CO_2 as Inferred by $\delta^{13}C$ Measurements from the Noaa Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, J. Geophys. Res.-Atmos., 100, 5051–5070, 1995.

Coplen, T. B.: Guidelines and recommended terms for expression of stable-isotope-ratio and gas-ratio measurement results, Rapid Commun. Mass Sp., 25, 2538–2560, 2011.

GLOBALVIEW-CO2C13: Cooperative Atmospheric Data Integration Project – $\delta^{13}C$ of Carbon Dioxide, CD-ROM, NOAA ESRL, Boulder, Colorado, available at: ftp://ftp.cmdl.noaa.gov/ccg/co2c13/GLOBALVIEW/, path:ccg/co2c13/GLOBALVIEW, 2009.

Jordan, A. and Brand, W. A.: Technical Report: MPI-BGC, Germany, in: Proceedings of the 11th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques, WMO/TD-No. 1138, edited by: Toru, S. and Kazuto, S., WMO, Geneva, 2003.

Keeling, C. D., Piper, S. C., Whorf, T. P., and Keeling, R. F.: Evolution of natural and anthropogenic fluxes of atmospheric CO_2 from 1957 to 2003, Tellus B, 63, 1–22, 2011.

AMTD
5, 7293–7322, 2012

Keeling, R. F.: Measuring Correlations between Atmospheric Oxygen and Carbon-Dioxide Mole Fractions – a Preliminary-Study in Urban Air, J. Atmos. Chem., 7, 153–176, doi:10.1007/BF00048044, 1988.

Keeling, R. F., Najjar, R. P., Bender, M. L., and Tans, P. P.: What Atmospheric Oxygen Measurements Can Tell Us About the Global Carbon-Cycle, Global Biogeochem. Cy., 7, 37–67, 1993.

Kozlova, E. A., Manning, A. C., Kisilyakhov, Y., Seifert, T., and Heimann, M.: Seasonal, synoptic, and diurnal-scale variability of biogeochemical trace gases and O_2 from a 300-m tall tower in central Siberia, Global Biogeochem. Cy., 22, GB4020, doi:10.1029/2008GB003209, 2008.

Langenfelds, R. L., Francey, R. J., Steele, L. P., Keeling, R. F., Bender, M. L., Battle, M., and Budd, W. F.: Measurements of O_2/N_2 ratio from the Cape Grim air archive and three independent flask sampling programs, in: Baseline Atmospheric Program Australia 1996, edited by: Gras, J. L., Derek, N., Tindale, N. W., and Dick, A. L., Bureau of Meteorology and CSIRO Atmospheric Research, Melbourne, Australia, 1999.

Leuenberger, M., Nyfeler, P., Moret, H. P., Sturm, P., and Huber, C.: A new gas inlet system for an isotope ratio mass spectrometer improves reproducibility, Rapid Commun. Mass Sp., 14, 1543–1551, 2000a.

Leuenberger, M., Nyfeler, P., Moret, H. P., Sturm, P., Indermuhle, A., and Schwander, J.: CO_2 concentration measurements on air samples by mass spectrometry, Rapid Commun. Mass Sp., 14, 1552–1557, 2000b.

Manning, A. C. and Keeling, R. F.: Global oceanic and land biotic carbon sinks from the Scripps atmospheric oxygen flask sampling network, Tellus B, 58, 95–116, 2006.

Manning, A. C., Jordan, A., Levin, I., Schmidt, M., Neubert, R. E. M., Etchells, A., Steinberg, B., Ciais, P., Aalto, T., Apadula, F., Brand, W. A., Delmotte, M., Giorgio di Sarra, A., Hall, B., Haszpra, L., Huang, L., Kitzis, D., van der Laan, S., Langenfelds, R. L., Leuenberger, M., Lindroth, A., Machida, T., Meinhardt, F., Moncrieff, J., Morguí, J. A., Necki, J., Patecki, M., Popa, E., Ries, L., Rozanski, K., Santaguida, R., Steele, L. P., Strom, J., Tohjima, Y., Thompson, R. L., Vermeulen, A. T., Vogel, F., and Worthy, D. E.: Final report on CarboEurope "Cucumber" intercomparison programme, 2009.

Masarie, K. A., Langenfelds, R. L., Allison, C. E., Conway, T. J., Dlugokencky, E. J., Francey, R. J., Novelli, P. C., Steele, L. P., Tans, P. P., Vaughn, B., and White, J. W. C.: NOAA/CSIRO Flask Air Intercomparison Experiment: A strategy for directly assessing consistency among

AMTD
5, 7293–7322, 2012

atmospheric measurements made by independent laboratories, J. Geophys. Res.-Atmos., 106, 20445–20464, 2001.

Masarie, K. A., Petron, G., Andrews, A., Bruhwiler, L., Conway, T. J., Jacobson, A. R., Miller, J. B., Tans, P. P., Worthy, D. E., and Peters, W.: Impact of CO_2 measurement bias on CarbonTracker surface flux estimates, J. Geophys. Res.-Atmos., 116, D17305, doi:10.1029/2011JD016270, 2011.

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., Miller, J. B., Bruhwiler, L. M. P., Petron, G., Hirsch, A. I., Worthy, D. E. J., van der Werf, G. R., Randerson, J. T., Wennberg, P. O., Krol, M. C., and Tans, P. P.: An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker, P. Natl. Acad. Sci. USA, 104, 18925–18930, 2007.

Popa, M. E., Gloor, M., Manning, A. C., Jordan, A., Schultz, U., Haensel, F., Seifert, T., and Heimann, M.: Measurements of greenhouse gases and related tracers at Bialystok tall tower station in Poland, Atmos. Meas. Tech., 3, 407–427, doi:10.5194/amt-3-407-2010, 2010.

Rothe, M., Jordan, A., and Brand, W. A.: Trace gases, $\delta^{13}C$ and $\delta^{18}O$ of CO_2-in-air samples: Storage in glass flasks using PCTFE seals and other effects, in: Proceedings of the 12th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques, WMO/TD-No. 1275, edited by: Worthy, D. and Huang, L., WMO, Geneva, 2005.

Sirignano, C., Neubert, R. E. M., and Meijer, H. A. J.: N_2O influence on isotopic measurements of atmospheric CO_2, Rapid Commun. Mass Sp., 18, 1839–1846, 2004.

Sirignano, C., Neubert, R. E. M., Rödenbeck, C., and Meijer, H. A. J.: Atmospheric oxygen and carbon dioxide observations from two European coastal stations 2000–2005: continental influence, trend changes and APO climatology, Atmos. Chem. Phys., 10, 1599–1615, doi:10.5194/acp-10-1599-2010, 2010.

Sturm, P., Leuenberger, M., Sirignano, C., Neubert, R. E. M., Meijer, H. A. J., Langenfelds, R., Brand, W. A., and Tohjima, Y.: Permeation of atmospheric gases through polymer O-rings used in flasks for air sampling, J Geophys. Res.-Atmos., 109, D04309, doi:10.1029/2003JD004073, 2004.

Sturm, P., Leuenberger, M., Valentino, F. L., Lehmann, B., and Ihly, B.: Measurements of CO_2, its stable isotopes, O_2/N_2, and ^{222}Rn at Bern, Switzerland, Atmos. Chem. Phys., 6, 1991–2004, doi:10.5194/acp-6-1991-2006, 2006.

AMTD
5, 7293–7322, 2012

Thompson, R. L., Manning, A. C., Gloor, E., Schultz, U., Seifert, T., Hänsel, F., Jordan, A., and Heimann, M.: In-situ measurements of oxygen, carbon monoxide and greenhouse gases from Ochsenkopf tall tower in Germany, Atmos. Meas. Tech., 2, 573–591, doi:10.5194/amt-2-573-2009, 2009.

Uglietti, C., Leuenberger, M., and Valentino, F. L.: Comparison between real time and flask measurements of atmospheric O_2 and CO_2 performed at the High Altitude Research Station Jungfraujoch, Switzerland, Sci. Total Environ., 391, 196–202, 2008.

van der Laan, S., Neubert, R. E. M., and Meijer, H. A. J.: A single gas chromatograph for accurate atmospheric mixing ratio measurements of CO_2, CH_4, N_2O, SF_6 and CO, Atmos. Meas. Tech., 2, 549–559, doi:10.5194/amt-2-549-2009, 2009.

van der Laan, S., Karstens, U., Neubert, R. E. M., van der Laan-Luijkx, I. T., and Meijer, H. A. J.: Observation-based estimates of fossil fuel-derived CO_2 emissions in the Netherlands using Delta ^{14}C, CO and ^{222}Radon, Tellus B, 62, 389–402, 2010.

van der Laan-Luijkx, I. T., Karstens, U., Steinbach, J., Gerbig, C., Sirignano, C., Neubert, R. E. M., van der Laan, S., and Meijer, H. A. J.: CO_2, $\delta O_2/N_2$ and APO: observations from the Lutjewad, Mace Head and F3 platform flask sampling network, Atmos. Chem. Phys., 10, 10691–10704, doi:10.5194/acp-10-10691-2010, 2010.

Werner, R. A., Rothe, M., and Brand, W. A.: Extraction of CO_2 from air samples for isotopic analysis and limits to ultra high precision delta ^{18}O determination in CO_2 gas, Rapid Commun. Mass Sp., 15, 2152–2167, 2001.

WMO: 15th WMO/IAEA Meeting of Experts on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques, WMO/TD-No. 1553, WMO, Geneva, 2011.

Worthy, D. E. J., Platt, A., Kessler, R., Ernst, M., and Racki, S.: Measurement Procedures and Data Quality, Canadian Baseline Program; Summary of progress to 2002, Meteorological Service of Canada, 97–120, 2003.

AMTD
5, 7293–7322, 2012

Table 1. Average standard errors in the mean of the duplicate or triplicate flasks for the CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ measurements from each of the three laboratories.

	UBE	RUG	MPI
CO_2 (ppm)	0.05	0.06	0.06
$\delta(O_2/N_2)$ (per meg)	6	8	3
$\delta^{13}CO_2$ (‰)	0.08	0.07	0.009

AMTD
5, 7293–7322, 2012

Table 2. Average CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ differences between each set of two laboratories and their standard errors in the mean. Also given are the standard deviations. The results are given for the entire data set as well as for the two sub-periods: before March 2009 (part 1) and after March 2009 (part 2).

	UBE – RUG		UBE – MPI		MPI – RUG	
	average	stdev	average	stdev	average	stdev
CO_2 (ppm)	0.20 ± 0.06	0.6	0.08 ± 0.05	0.4	0.14 ± 0.06	0.5
CO_2 (part 1)	0.18 ± 0.06	0.3	0.21 ± 0.05	0.3	0.042 ± 0.07	0.3
CO_2 (part 2)	0.21 ± 0.09	0.7	0.01 ± 0.07	0.5	0.19 ± 0.08	0.6
$\delta(O_2/N_2)$ (per meg)	-33 ± 4	40	-31 ± 4	30	-3 ± 3	26
$\delta(O_2/N_2)$ (part 1)	-33 ± 6	30	-14 ± 6	30	-16 ± 4	20
$\delta(O_2/N_2)$ (part 2)	-37 ± 5	40	-38 ± 4	30	1 ± 4	27
$\delta^{13}CO_2$ (‰)	-0.03 ± 0.04	0.3	-0.02 ± 0.03	0.22	-0.02 ± 0.03	0.20
$\delta^{13}CO_2$ (part 1)	-0.06 ± 0.05	0.25	0.00 ± 0.07	0.20	-0.13 ± 0.04	0.10
$\delta^{13}CO_2$ (part 2)	-0.02 ± 0.05	0.3	-0.02 ± 0.03	0.23	-0.00 ± 0.03	0.21

AMTD
5, 7293–7322, 2012

Table 3. CO_2, $\delta(O_2/N_2)$ and $\delta^{13}CO_2$ trends and seasonal amplitudes based on the fit of the data sets from each laboratory: UBE, RUG and MPI. The used fit is a linear combination of a linear trend and a double (for CO_2) or single (for $\delta(O_2/N_2)$ and $\delta^{13}CO_2$) harmonic seasonal component.

	UBE	RUG	MPI
Trend CO_2 (ppm yr^{-1})	1.76 ± 0.17	1.94 ± 0.18	1.83 ± 0.17
Amplitude CO_2 (ppm)	10.3 ± 0.3	10.6 ± 0.4	10.7 ± 0.3
Trend $\delta(O_2/N_2)$ (per meg yr^{-1})	−29[a] ± 3	−23 ± 3	−17.3 ± 1.5
Amplitude $\delta(O_2/N_2)$ (per meg)	69[a] ± 5	85 ± 4	84.1 ± 2.2
Trend $\delta^{13}CO_2$ (‰ yr^{-1})	−0.081[b] ± 0.018	−0.069 ± 0.015	−0.016 ± 0.014
Amplitude $\delta^{13}CO_2$ (‰)	0.592 ± 0.028	0.455 ± 0.022	0.485 ± 0.018

[a] More realistic values are obtained when a stronger filter is applied to the data: −21 ± 2 per meg yr^{-1} and 73 ± 3 per meg for the linear trend and seasonal amplitude respectively. [b] The trend estimate based on the complete record available for UBE between 2000 and 2012 is: −0.013 ± 0.004 ‰.

AMTD
5, 7293–7322, 2012

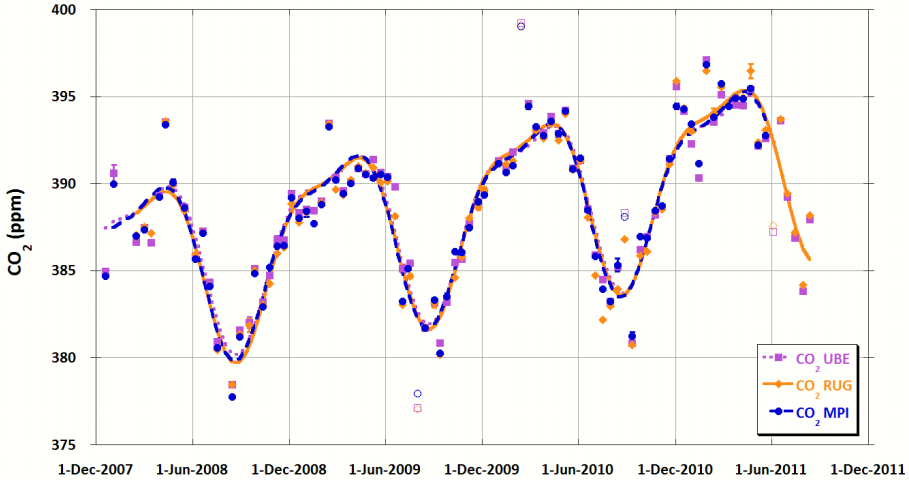

Fig. 1. CO_2 concentration at Jungfraujoch, Switzerland from flask samples measured by three laboratories: University of Bern (UBE) (pink squares), University of Groningen (RUG) (orange diamonds) and Max Planck Institute in Jena (MPI) (blue circles). The values are the averages of 1, 2 or 3 flasks. The fits through the data are linear trends and double harmonic seasonal. Open symbols represent those values that are outliers to the fit of the individual data set. The error bars represent the standard error of the average value of 2 or 3 flasks. For single flask measurements error bars are not shown.

AMTD
5, 7293–7322, 2012

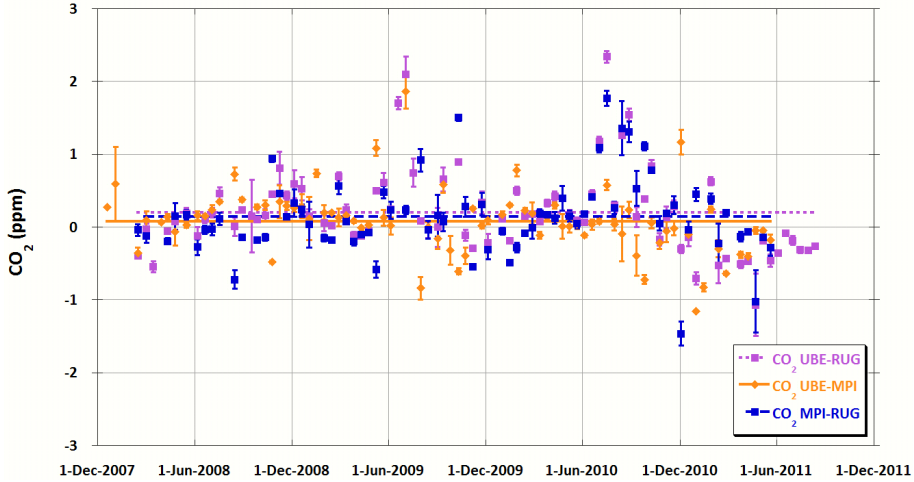

Fig. 2. Differences of the CO_2 concentration measured by each set of two laboratories. Also indicated are the average differences. These are: 0.20 ppm for UBE-RUG, 0.08 ppm for UBE-MPI and 0.14 ppm for MPI-RUG. The error bars represent the quadratically added standard errors of the measurements of the two laboratories.

AMTD
5, 7293–7322, 2012

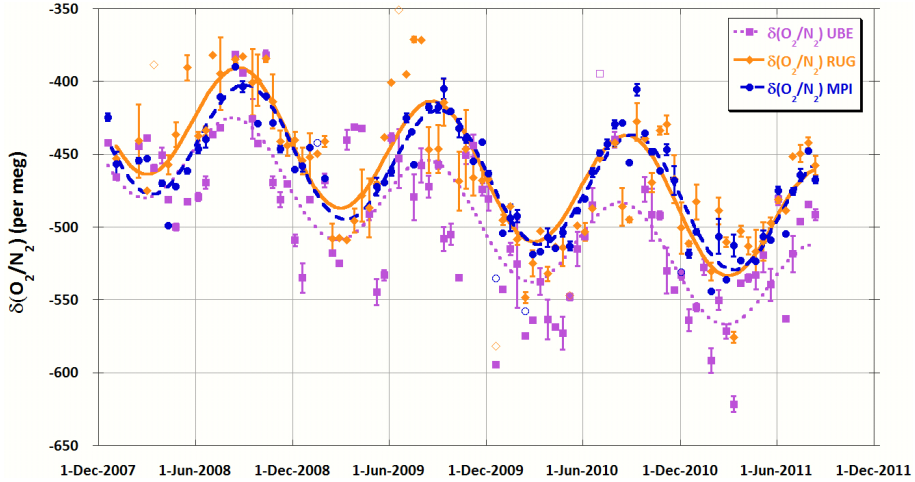

Fig. 3. $\delta(O_2/N_2)$ observations from Jungfraujoch, Switzerland from flask samples measured by three laboratories: UBE (pink squares), RUG (orange diamonds) and MPI (blue circles). The values are the averages of 1, 2 or 3 flasks. The fits through the data are linear trends and single harmonic seasonal components. Open symbols represent those values that are outliers to the fit of the individual data set. The error bars represent the standard error of the average value of 2 or 3 flasks. For single flask measurements error bars are not shown.

AMTD
5, 7293–7322, 2012

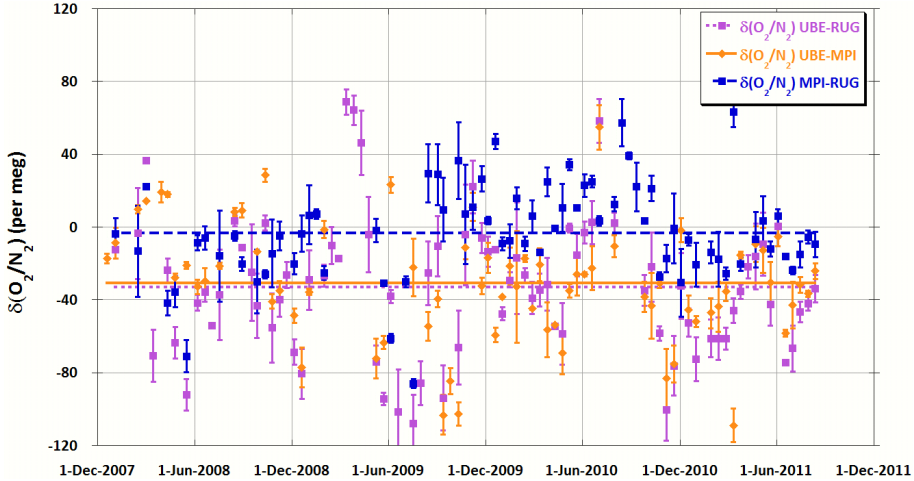

Fig. 4. Differences of the $\delta(O_2/N_2)$ values measured by each set of two laboratories. Also indicated are the average differences. These are: −33 per meg for UBE-RUG, −31 per meg for UBE-MPI and −3 per meg for MPI-RUG. The error bars represent the quadratically added standard errors of the measurements of the two laboratories.

AMTD
5, 7293–7322, 2012

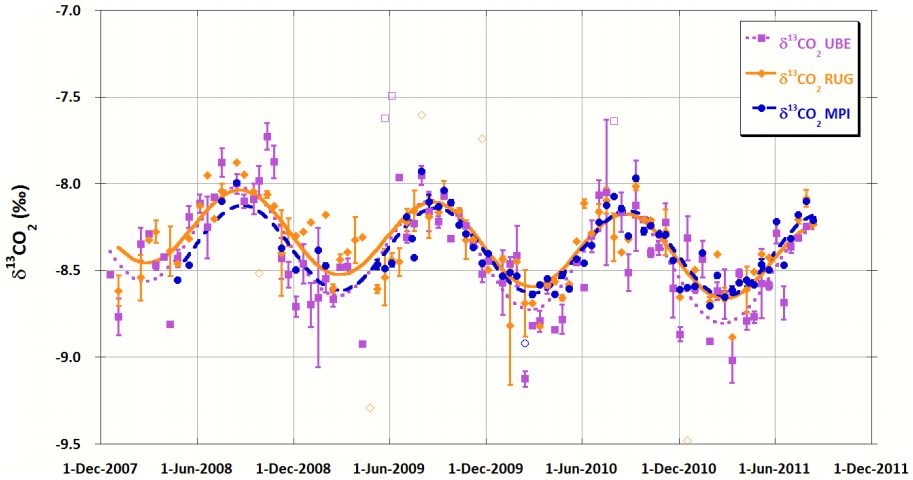

Fig. 5. $\delta^{13}CO_2$ observations from Jungfraujoch, Switzerland from flask samples measured by three laboratories: UBE (pink squares), RUG (orange diamonds) and MPI (blue circles). The values are the averages of 1, 2 or 3 flasks. The fits through the data are linear trend and single harmonic seasonal components. Open symbols represent those values that are outliers to the fit of the individual data set. The error bars represent the standard error of the average value of 2 or 3 flasks. For single flask measurements error bars are not shown.

AMTD
5, 7293–7322, 2012

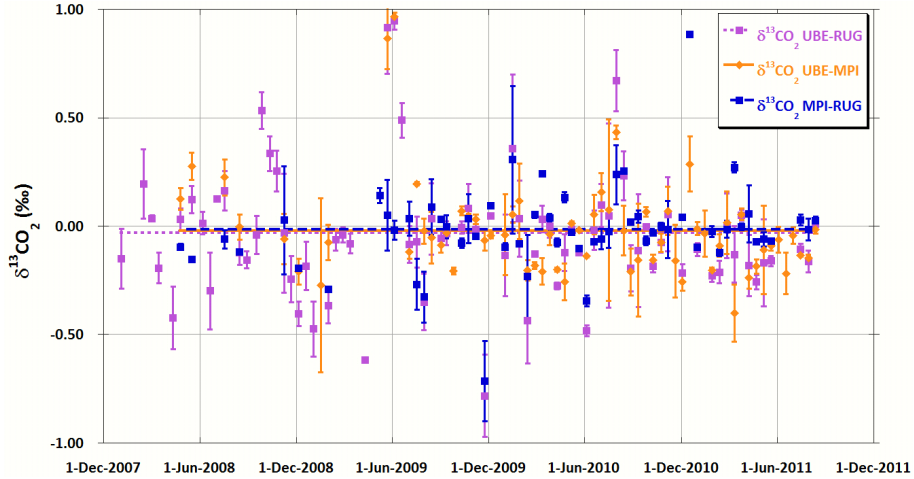

Fig. 6. Differences of the $\delta^{13}CO_2$ values measured by each set of two laboratories. Also indicated are the average differences. These are: −0.03 ‰ for UBE-RUG, −0.02 ‰ for UBE-MPI and −0.02 ‰ for MPI-RUG. The error bars represent the quadratically added standard errors of the measurements of the two laboratories.

AMTD
5, 7293–7322, 2012