

Reply to Anonymous Referee #1

Reviewers' comments are given in *italics* while the replies are given in roman.

Received and published: 26 August 2014

This manuscript describes the details about CO₂ intercomparison measured by NDIR and CRDS at Jungfraujoch. The manuscript contains a lot of information for in-situ CO₂ observations. But more detail explanations are needed in some important sections. I recommend the publication of the manuscript in AMT after the major revisions.

General comments

In KUP system, there is no description how to calculate the CO₂ mole fraction of sample gases by using the assigned values of working gas, low span and high span. In other words, how to transfer the WMO X2007 scale to sample gases?

The CO₂ mole fraction of the KUP system is calculated as follows:

$$XCO_2 = CO_{2,raw} * Slope_{G,K} + B_{assigned} - (B_{raw} * Slope_{G,K})$$

where XCO₂ is the calculated CO₂ mole fraction of the sample, CO_{2,raw} is the signal from the NDIR instrument when measuring a sample, Slope_{G,K} is the slope calculated from the high span (G) and the low span (K), B_{assigned} is the assigned value of the working gas (B) and B_{raw} is the signal from the NDIR instrument when measuring the working gas. The CO₂ values are reported on the WMO X2007 scale based on assigned values of the reference gases which were calibrated in Bern against three NOAA calibration tanks with different CO₂ concentrations (see page 7059, line 6).

The manuscript concludes that most of the reasons for discrepancy between NDIR and CRDS measurements are due to NDIR system such as (i) adsorption/desorption effects in the calibration gas, (ii) insufficient nonlinearity correction of the NDIR and (iii) badly assigned calibration gas cylinders. Also (iv) recommendation to use aluminium cylinder is mentioned in conclusion. I think these are not enough for conclusion section of this manuscript, since (i) is deeply discussed in other paper (Leuenberger et al., 2014) and (ii), (iii) and (iv) are not new findings. In the conclusion section, it is better to describe more new or useful information found in this intercomparison.

This study was done to show whether the two data sets can be linked smoothly or not, which is not self-evident due to the fundamentally different measurement principles and gas handling systems.

Page 7068, line 2 to line 13 was changed to:

“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 Jungfraujoch were compared. The evaluation of the two CO₂ records from the NDIR system of the KUP to the CRDS system of the Empa showed that the two systems are generally in good agreement, not only on the 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 could be used to complement each other. However, the comparison revealed also (i) adsorption/desorption effects in the calibration gas steel cylinders used by NDIR system and (ii) insufficient nonlinearity correction of the NDIR analyzer (iii) periods of small biases because of imperfectly assigned calibration gas cylinders. Adsorption effects can be corrected for by monolayer adsorption equation (Langmuir, 1918, Leuenberger, 2014). Nonlinearity of the NDIR analyzer was constrained with a 2nd order polynomial, resulting in better agreement between the two data sets, in particular for values strongly deviating from the average.”

Page 7057, line 23-24, It is hard to conclude “a precision better than 0.04 ppm” from Table 1.

The value in the manuscript is the average precision of the measurements in Table 1 (column 4), which is 0.037 ppm. This number is not to be confused with the accuracy of the measurement system which is 0.08 ppm (in the AMTD manuscript a wrong value of 0.12 ppm was mentioned which is changed to 0.08 ppm) calculated by the square root of the sum of squared trueness and precision (Menditto et al., *Accred. Qual. Assur.*, 2006, 12, 45, reference is now included in the manuscript too). The value given on page 7065, line 6 corresponds to the mean absolute deviation of the comparison measurements (see new table 1, column 6). A sentence about the way we calculated these values was added at line 7075 (precision) and line 7065 (accuracy).

Page 7058, line 9-11, According to Rella et al. (2013), water vapor collection can be possible in case of lower H₂O level. Had the observed H₂O mixing ratio been lower enough from August 2010 to July 2011? Even if the empirical humidity correction was smaller enough, it is worth to describe the amount of correction and its uncertainty.

“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) (see e.g. Fig. 14 there), 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 Jungfraujoch refers to a dew point of more than 14 °C, a value that is never reached at Jungfraujoch. The error can be even reduced when using the dedicated correction coefficients for the respective analyzer and doing repeated water vapor interference measurements, as it was the case here.” was added on page 7058 line 11 after the last sentence.

In return the sentence

“The water vapour correction functions for both instruments were determined by humidifying a dry reference gas by passing it over a water droplet in a coiled tubing (see Rella et al., 2013 for details).”

at page 7058, line 16-18 was deleted.

Page 7060, line 19-21, Why one month values are needed for calculating seasonality? Hourly average data are better for this purpose.

Hourly values show more details but for the seasonality they are not very well suited because they are too noisy. Seasonality based on the maximum minus the minimum of hourly values of the same year can lead to nonrepresentative results due to extreme values caused by e.g. synoptic events, thermal up lift of air from the surrounding valleys etc. Therefore you have to smooth the measurements with a running mean, with harmonic functions, spline functions etc., but in any case a cut off period is introduced in the calculation. In the manuscript the cut off period was chosen to be a month because the nature of the seasonality is still preserved but the short term variations are removed. Besides that, monthly means are very common in meteorology and climatology and are often used to report data, e.g. CO₂ values at Mauna Loa, Cape Grim etc .

Page 7062, line 14-15, “the offset was probably caused by the change of the Picarro analyzer”. Are there any other reasons such as drying system for sample gas?

In August 2010, the drying system of the CRDS was short cut and the dry air mole fractions were calculated based on an empirically determined correction function. Since this short cut didn't show an effect on the comparison between the two systems the reapplication of the dryer shouldn't have an influence in reverse.

Page 7063, line 21-23, Fig. 8 is not necessary, since the information from Fig. 8 is limited.

We like to keep this figure, because together with the figures 6 and 7 it documents the improvement achieved by the adsorption/desorption correction on the showcase of the cylinder B23 and therefore these three figures form an entity. In figure 8 it is clearly visible how the points, which are not “in line”, are corrected by the Langmuir function.

Page 7063, line 23-24, “The remaining working gases cylinders were recalculated in the same manner, too.” The “large offsets” in B27 and B29 are supposed to be found by this recalculation using low and high span gases. Is it possible to correct these “large offsets”?

It is hard to correct for these offsets because we have no independent information/measurements on these cylinders. The main problem is that B27 and B29 have CO₂ (assigned) mole fractions of 448.04 ppm and 437.55 ppm, whereas the low and the high span of this period have (assigned) CO₂ mole fractions of 429.02 ppm and 380.77 ppm. So the two cylinders are much higher than the high span. These two cylinders (B27 and B29) were thought to be the next high spans, but because of logistic problems we had to use them as working gases instead. By recalculating B27 and B29 with the low and the high span the difference between the calculated and the assigned value of the cylinders B27 and B29 is 0.1 ± 0.14 ppm and 0.48 ± 0.10 ppm, respectively. But as mentioned above, they are far beyond the span.

Page 7064, line 12-24. It is hard for me to accept these sentences and Fig. 9. It is well known that relation between output signals of NDIR and CO₂ mole fractions shows nonlinear. But in Fig. 9, both CRDS (Empa) and NDIR (KUP) are CO₂ mole fractions calibrated by other standard gases. Even if the analyzer has substantial nonlinearity, calibrated values remain no such effect when they use suitable numbers and suitable range of standard gases. Therefore I suppose the numbers and/or range of standard gases were not adequate for measuring CO₂ values by NDIR. I recommend to describe the method of measuring eleven used calibration gas cylinders of the KUP system more precisely.

The system uses only three calibration standards, a low span, a high span and a working gas. The low and the high span are used to determine the slope of 18 h intervals whereas the working gas is used to calculate the offset and short term variations. The high span had a CO₂ mole fraction of 429.02 ppm (only one high span was used in this period), the low spans had a CO₂ mole fraction of 360.31 ppm, 380.77 ppm and 384.41 ppm. The working gas is normally between the low and the high span around the values of the outside air. The eleven cylinders mentioned on page 7064 in line 12-24 of the AMTD were either already used calibration gas cylinders (ten) with a remaining pressure of around 20 bar except two cylinders which were almost empty and one unused cylinder, which have been re-analysed on Jungfrauoch and at the Bern laboratory. Before the cylinders were shipped back, they were measured on the two systems to get more information about the compatibility of the CRDS and the NDIR and not for calibration reasons.

Page 7065, line 20-26, I agree that one of the reason for discrepancy of the trend is due to the short observation period. But I think the contribution of assigned values for the working gases is not significant. Because the difference of the CRDS-NDIR CO₂ measurements plotted in Fig. 3 shows no such trend (0.2 ppm year⁻¹). The error of fitting process assumed to contribute the trend discrepancy.

It is striking that with each change of a working gas the difference between the two systems changed stepwise, especially in the second part of the comparison. Therefore the offsets between the two systems are most likely caused by the working gases. Especially in the second half of this comparison the NDIR values seem to be underestimated (because the difference Picarro – Maihak becomes bigger) which gives a slope to the differences of the two systems which is indeed 0.18 ppm y^{-1} and therefore is also in the values from which the slopes were calculated. (See also reply to comments of Referee #2.)

Page 7066, line 13-20, When we use a nonlinear analyzer, calculated CO₂ mole fractions depend on the range of standard gases. Have you confirmed that the influence of different CO₂ mole fraction in working gases is small enough?

According to the manufacturer, the analyzer is calibrated to a range from 350 ppm to 450 ppm, so the used calibration gases are within this range, the working span was added to the manuscript. Before the NDIR measurements were started at the end of 2004, the Maihak was tested in the laboratory with different standard gases. The Maihak showed a good response to the different CO₂ mole fractions and was decided to be suitable for measuring ambient CO₂ mole fractions. However, normally the used working gases have a CO₂ mole fraction of roughly 380-410ppm which is around the mole fraction of outside air and where the influence of the nonlinearity is still quite small, but with inapt working gases like B27 which had an assigned CO₂ concentration of 448.04 the influence is already quite big and probably adds to the problem of the imperfectly assigned CO₂ value.

Page 7067, line 11-20, The large volume of the drying unit leads to both damping and delaying effects. The delaying time of three minutes is not negligible when you pick up the 115 s data from CRDS measurement in Figs. 4 and 5. The difference between the two systems expected to be improved by considering this delaying time.

The delay time depends on the flow rate of the system which in turn is dependent on the outside pressure and varies, so the 3 minutes are not a constant delay and therefore difficult to correct for systematically. Shifting the time of one of the measurements minimized the differences only minimal. Therefore the combination of the time shift caused by different volumes in combination with the dampening and mixing of the air masses due to the relatively large volume in the water trap seems to be responsible for the small differences in the 1 Hz data. The hourly averages which are our usual output product from the system shouldn't be much affected, especially during stable periods because increasing and decreasing CO₂ mole fractions should cancel each other on longer time scales.

Page 7067, line 26-29, It is possible to make a correction of working gases by comparing low and high span gases or re-calibration in University of Bern after used in Jungfraujoch.

The calculation of the working gases with the low and the high span was done for the correction of the adsorption/desorption effects. This calculation has often a relatively high standard deviation. It is therefore hard to argue that the calculated values are better than the assigned ones since the standard deviations are overlapping. The cylinders were also remeasured when they came back to Bern, but they were low in pressure and showed strong CO₂ enrichment effects.

Page 7081, Figure 6, The differences between black points and grey points look small before Apr. 2010. But it is not easy to confirm in steep slope periods. Are differences small enough?

The differences CO₂ enrichment – Best fit were calculated. The average difference before April 2010 was -0.02 ppm ± 0.10 ppm, afterwards the difference was -0.04 ppm ± 0.13 ppm. Since the two differences are the same within the uncertainties, they should be small enough.

Technical corrections

Page 7058, line 23-24, "Table 1" should be changed to "Table 3".

"Table 3" was added at Page 7058, line 23 at the end of the sentence.

Table 1 on line 23-24 is correct, the comment doesn't refer to the calibration gases of the CRDS system but to measurements done at JFJ with several used calibration cylinders listed in table 1. But

to make it more clear page 7058, line 22-23 was changed from “Measurements of standard cylinders showed...” to “Measurements of used KUP standard cylinders at JFJ showed...”

Page 7060, line 21, “do” to “to”?

Correct, we changed “do” to “to”

Page 7061, line 25, “0.75 ppm” to “0.75 ppm h⁻¹”.

Correct, “0.75 ppm” was changed to “0.75 ppm per hour”

Page 7061, line 27, “0.996 ppm” to “0.996 ppm/ppm”.

Correct, “0.996 ppm” was changed to “0.996 ppm / ppm”

Page 7061, line 27, “1.5696 ppm/ppm” to “1.5696 ppm”.

Correct, “1.5696 ppm / ppm” was changed to “1.5696 ppm”

Page 7062, line 20, “Langmuir, 1918” cannot be found in References.

Langmuir (1918) was added to the references.

Page 7072, Table 1., The value of LK94521 measured by CO₂ Maihak seems to be “359.96” rather than “459.96”.

Page 7072, line 20: Correct, “459.96” was changed to “359.96”

Reply to Anonymous Referee #2

Reviewers’ comments are given in italics while the replies are given in roman.

Received and published: 15 September 2014

The paper is describing three years (2010-2012) of in-situ CO₂ measurements at Jungfraujoch Station performed by two types of analyzers (NDIR and CRDS). The description of the measurements and comparison of both time series is clearly explained, but would need additional details in some parts. I do recommend the manuscript for publication in AMT after some revisions proposed listed below. General remark: the averaged difference between NDIR and CRDS is 0.04 +/-0.4 ppm, or a little bit higher depending on the cutoff used to reject highly variable data from the time series. However this difference is variable over time and I would like to see more analysis of this variability. For example in Spring 2012 it seems that the CRDS systematically higher than NDIR by about 0.2 ppm during a couple of months.

This period coincides with the usage of the cylinder B29 which was most probably imperfectly assigned.

Could you provide the difference on a monthly basis? Also there were some changes in the setup of the measurement (change of the NDIR sensor in 2011, and of the CRDS one in 2011; moving from dry to wet air for the CRDS measurements). It would be interesting to evaluate if the CO₂ differences are similar or not for the periods corresponding to these different setup.

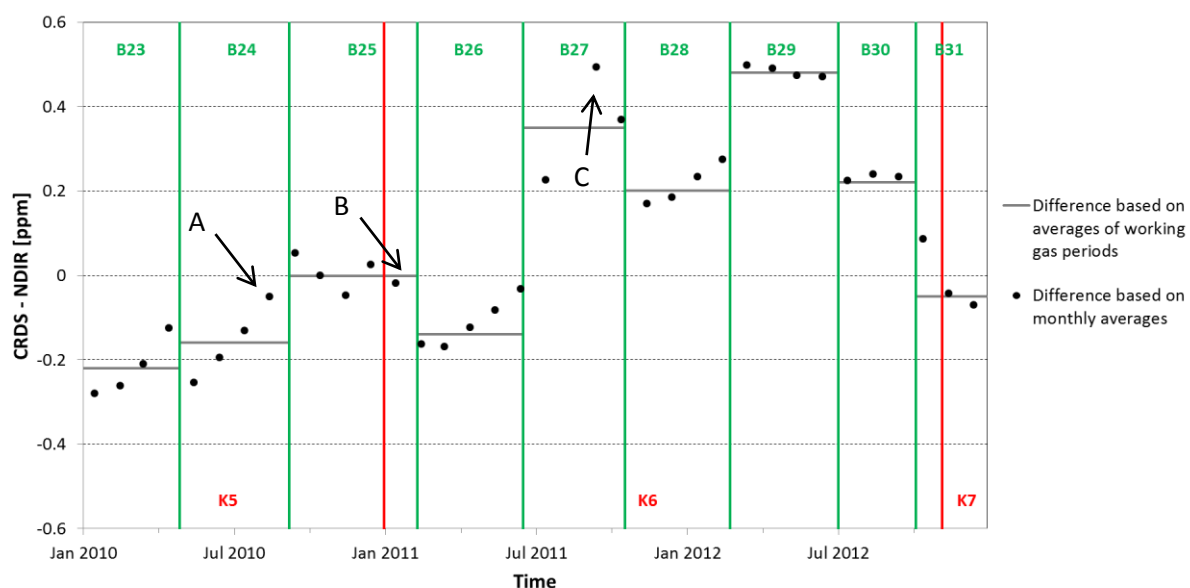


Fig. 1: Monthly averages of CO₂ differences measured with CRDS and NDIR. The green lines define the periods of the NDIR working gas cylinders (green numbers), the red lines define the periods of the NDIR low span cylinders (red numbers). The black points are the monthly averages of the differences between the CRDS and the NDIR, the grey lines represent the averages of the differences based on the working gas periods. A: Nafion dryer was bypassed; B: NDIR analyzer was replaced; C: CRDS analyzer was replaced and Nafion drying system was reactivated

The changes of the working gas cylinders can be clearly seen in the differences based on the averages of the working gas periods as well as in the differences based on monthly averages especially from spring 2011 on. Until July 2011, the measurements of the NDIR system tend to be a bit higher than the CRDS measurements, afterwards they are lower, which adds a slope to the differences of 0.2 ppm y⁻¹.

Bypassing the Nafion dryer in summer 2010 doesn't seem to have an influence on the comparability of the two datasets. The influence of the exchange of the NDIR analyzer in February 2011 is hard to assign because at the same time the working gas had to be changed but the NDIR measurements seem to be higher by roughly 0.15 ppm after the change. It looks as if the change of the CRDS analyzer and reapplication of the drying system raised the CRDS measurements by roughly 0.15 ppm (difference first point – last point in period B27, the second point wasn't used because there are only a few measurements at the end of the month). Because the bypassing of the Nafion dryer in summer 2010 didn't have an effect on the differences, the reapplication shouldn't have an effect too. Therefore the change in the differences is most probably caused by the new Picarro instrument and imperfectly assigned working gas cylinders.

7054 (25): Allowing additional time?

"Allowing additionally" was changed to "Allowing additional time..."

7057 (10): Could you precise the level of temperature/pressure regulation you have for the NDIR instrument?

The laboratory at Jungfraujoch where the gas control system and the NDIR are running is air conditioned to 20 ± 1.5 °C, the precision of the pressure regulation is ± 0.05 mbar, the temperature regulation in the box with the paramagnetic cell, flow controller, pressure controller etc. has a precision of ± 0.02 °C and the Maihak measurement cell is heated to 55 ± 0.05 °C. These numbers were added to the manuscript at page 7057, line 8.

7058 (20): you mention the measurement of a target gas to detect short term changes, but you don't present any result from this gas. How do you use this information? Tables 2 and 3 provide the assigned values of the calibration tanks for NDIR and CRDS. You should precise in the legend that the values are assigned by central lab at Univ. of Bern and EMPA respectively. These intermediate steps between NOAA and the stations may introduce a bias. Have you performed any cross-comparison between the two central labs?

At Table 2 the sentence "The gases were measured at the University of Bern and calibrated with a suite of NOAA standard cylinders." was added to the caption. At Table 3 "...as measured at the Empa laboratory." was added at the end of the caption.

There was one direct comparison between the EMPA and the KUP laboratories within the 5th WMO Round Robin exercise. The submitted results were:

	CA08182	CA08186	CC86203
NOAA	371.78 ± 0.00 ppm	399.90 ± 0.02 ppm	384.10 ± 0.02 ppm
Empa	371.76 ± 0.06 ppm	399.96 ± 0.06 ppm	384.15 ± 0.07 ppm
KUP	371.63 ± 0.02 ppm	399.86 ± 0.01 ppm	383.99 ± 0.01 ppm

Since recently Empa takes also part in the "Cucumber intercomparison programme" which will allow for further direct comparison of the two laboratories (see <http://cucumbers.uea.ac.uk> for more information).

7064: after the correction of the non-linearity of the NDIR you should reduce the bias of the target gas (H1, 440ppm) compared to its assigned WMO value by about 0.35 ppm. Is it what you observe?

The assigned CO₂ mole fraction the target cylinder is 440.16 ppm. The average of the uncorrected measurements is 439.66 ± 0.27 ppm, the average of the corrected measurements is 439.88 ± 0.27 ppm so the bias becomes smaller by 0.22 ppm. By excluding periods of the most probably imperfectly assigned working gas cylinders B27 and B29 the average of the uncorrected and corrected measurements is 439.70 ± 0.21 ppm and 440.00 ± 0.19 ppm, respectively, meaning that the bias of the measurement towards the assigned value becomes smaller by 0.30 ppm.

7065 (6): how do you estimate the 0.12 ppm accuracy of the standard measurements?

The value corresponds to the mean of the absolute deviation of the comparison measurements. It will be changed to 0.08 ppm which was calculated with the square root of the sum of the squared standard deviation of the polynomial corrected differences (CRDS – NDIR) and the squared precision.

7065 (22): I am not fully convinced by the attribution of the trend differences (0.2 ppm) to the assigned values of the working gases. Could you elaborate this discussion?

Page 7065, line 22 was changed from:

"The trend calculation (1.89 ppm year⁻¹ and 1.69 ppm year⁻¹) for the complete and the common data periods deviate by 0.2 ppm year⁻¹ which we mainly attribute to uncertainties in the assigned values for the working gases."

to

"The trend calculation (1.89 ppm year⁻¹ and 1.69 ppm year⁻¹) for the complete and the common data periods deviate by 0.2 ppm year⁻¹. In the first half of this comparison, the difference 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, the change of the CRDS instrument caused a small offset which together with the imperfectly assigned working gasses added a slope of 0.2 ppm year⁻¹ to the differences. Since the change of these differences is stepwise and coincides with the changes of the working gases and the change of the CRDS instrument, the slope in the differences and therefore also the differences in the slopes of the two data sets is most probably caused by the imperfectly assigned working gases and the change of the CRDS instrument in the second half of the comparison.”

7067: ‘independent comparison measurements’: again I am surprised that you don’t mention the use of the target gas which should be there to evaluate problems with the calibration scale.

Events such as the changing of the working gases can be observed in the target gas measurements but to our understanding, the target gas should only be used to indicate problems, not to correct the measurements.

Changes on the author’s behalf:

Page 7058, line 22-23: “Measurements of standard cylinders showed a precision better...” was changed to “Measurements of used KUP standard cylinders at JFJ showed a precision better...”

Page 7062, line 14: “KUP system” was replaced with “NDIR system” for consistency.

Page 7066, line 5: First “KUP” was replaced with “NDIR system”, second “KUP” on line 5 was replaced with “NDIR” for consistency.

Page 7066, line 6: “high” was wrong and therefore changed to “low”.

Page 7070, line 3-4: Reference was updated from “Leuenberger, M. C., Schibig, M. F., and Nyfeler P.: Gas adsorption and desorption effects on cylinders influence long-term gas records, submitted to ACPD 2014.” to “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, 2014.”