The authors would like to thank Referee#1 and Referee# for their valuable comments and suggestions on this manuscript. Below we have addressed the individual remarks by each reviewer. The Referee's comments and questions are blue, the authors' replies are formatted as plain text, and excerpts from the manuscript as well as changes to the manuscript are given in italics.

Reply to Anonymous Referee #1

The manuscript "Experiments with CO_2 -in-air reference gases in high-pressure aluminium cylinders" of M. Schibig et al. studies the stability of CO_2 in air mixtures at ambient mole fractions. The topic is relevant, since accurate and reproducible measurements with traceability to standard scales are needed to detect changes in regional sources and sinks of CO_2.

The manuscript is generally well written and concise. The chain of arguments is sound, and the topic is relevant for the scientific community. I therefore do recommend publication in AMT after addressing the following concerns.

General comments

The experiments carried out are clearly relevant for laboratory studies using large amount of standard gases within a short time period. However, this is mostly not the case for long-term monitoring of CO_2 and other research projects, where the gas of a standard cylinder is used over a much longer time period and in intervals. During period without use of standard gas, re-equilibration might take place, and the effect of changing CO_2 mole fraction during the use of a standard that was observed in this study will in many cases not happen or be much less pronounced in a normal measurements set-up for ambient CO_2.

We disagree with the referee. The adsorption/desorption effect is mainly pressure driven and not dependent on time. According to Langmuir's adsorption/desorption equation, the equilibrium between the molecules on the cylinder wall and in the air has to change with changing pressure.

Stability is much better for low flow conditions, but again, in reality, it might even be worse due to effects of the regulators. Especially in realistic measurement set-ups, this can be a problem, since only small aliquots of standard gases are used in longer time intervals, and the air is mostly sampled from the regulator and not directly from the cylinder. I recommend adding a few words on this issue.

In realistic measurement set-ups at an e.g. atmospheric measurement station, the main valve of a cylinder remains open and therefore the gas in the stem and the regulator can equilibrate with the gas in the cylinder body. Then, when a cylinder is measured e.g. to calibrate or as a target, the first few minutes are usually discarded to make sure the regulator is flushed properly and to avoid such effects. This is recommended good practice in every WMO/GAW report. If a cylinder is used only sporadic, the regulator should be flushed several times anyway. Therefore, we do not think that this makes a big difference if best practice recommendation for trace gas measurements are followed properly.

The study was carried out with dry air mixtures ($H_2O < 1$ µmol mol^{-1}). Residual water content might have a significant impact on the behavior of CO_2 absorption. Has the low water content be verified by measurements or other means?

The water content is ensured by the drying system NOAA uses to fill the calibration gas tanks at Niwot Ridge station and is continually measured (Meeco Waterboy, Meeco Inc., USA) during the filling procedure. For further information such as the setup and specifications of the filling station and the procedure see: https://www.esrl.noaa.gov/gmd/ccl/ccl.html and https://www.esrl.noaa.gov/gmd/ccl/airstandard.html.

What will be the effect of residual water, even if less than 1 µmol mol^{-1}? Could it be that differences in the residual water content explain at least partly the difference between individual cylinders or fillings?

At such low water contents, we do not expect any significant influence of the water on the CO_2 adsorption/desorption effects, because CO_2 molecules occupy only a fraction of the available active sites on the cylinder wall as shown in the paper. It would be different if there were more water in the cylinder, because the water molecules would compete with the CO_2 molecules for the available active sites and hinder CO_2 molecules to adsorb to the cylinder wall.

A very recent publication studies similar effects including the influence of water on the stability of gaseous reference materials (Brewer et al., 2018). Citation of this work should be made in the final AMT version of the paper.

The publication was added to the references.

SGS (Superior Gas Stability) cylinders are mentioned in the introduction and methods, but no results are shown in the paper. In the conclusions, they are mentioned again, saying that they behave in the same way as untreated cylinders. This should also be shown and discussed in the results. E.g. individual fits could be shown for SGS and untreated cylinders in a separate figure similar to Fig.6.

We added the following sentence to section 3.2:

"*The two SGS cylinders do not show a significantly different behavior, the form of the CO_2 enrichment with decreasing pressure as well as the amount is the same as for the normal cylinders within the given uncertainty (Fig. 6 b).*"

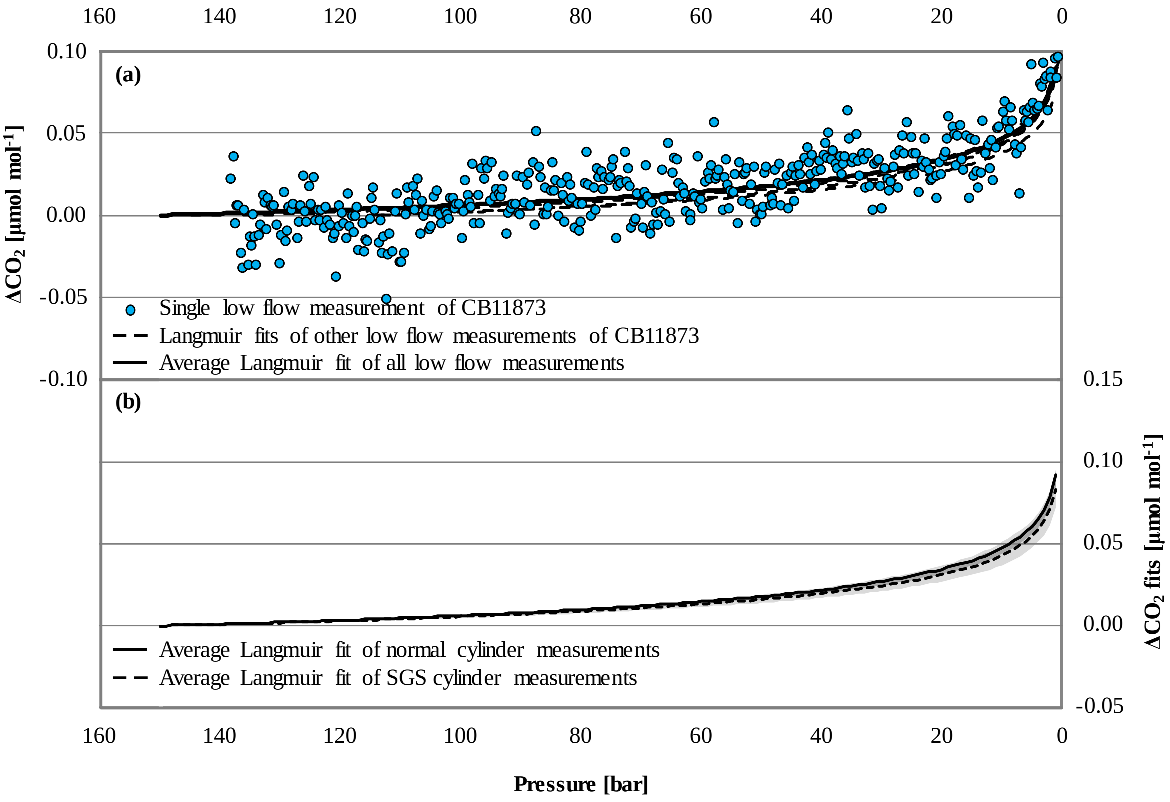
and we changed the sentence at page 10, line 34 from:

"*Two additional low flow measurements with horizontally positioned cylinders were done.*"

to:

"*Additionally, a low flow run with two horizontally positioned cylinders (one normal and one SGS cylinder) was done.*"

Additionally we expanded figure 6 that shows all low flow measurements of one cylinder (panel a) with a second panel that shows the average fit of all low flow measurements of the normal and SGS cylinders, respectively, with the uncertainties given as greyed area (panel b).



"Figure 6: a) The blue circles represent the CO_2 mole fraction measurement of a low flow experiment started on 17.10.2016 with CB11873 vertically positioned as a function of pressure, note the inverse pressure scale. The black dashed lines are the individual fits following the Langmuir model of the other low flow experiments done with CB11873 vertically positioned, the black solid line represents the average Langmuir fit using all low flow experiments with the cylinders vertically positioned. b) The black solid and dashed line correspond to the average Langmuir fit of all normal and SGS cylinder measurement, respectively, that were done under low flow conditions, the greyed area corresponds to the standard deviation of the averages. In order to plot all data in one plot, the corresponding ($CO_{2,ini}$ – $CO_{2,ad}$) was subtracted from the measurements and the fits in both panels."

In section 3.3, we replaced the first sentence

"In the high flow mode, eight complete drainings were done with cylinders vertically positioned…"

with

"In high flow mode, each of the six normal and the two SGS cylinders were drained once with cylinders vertically positioned…"

In section 4, the first sentence

“*The low flow measurements with cylinders vertically positioned show repeatedly comparable CO_2 enrichment with decreasing pressure, no matter which cylinder was measured.*“

was shortened and a second sentence was added, it reads now:

“*The low flow measurements with cylinders vertically positioned show repeatedly comparable CO_2 enrichment with decreasing pressure. Neither the normal nor the SGS cylinders showed any unique features with respect to CO_2 enrichment.*”

I don’t see much additional value of the experiments with heating and changing the orientation of the cylinders during venting. The results of the heating experiments were not consistent between different runs, and probably more experiments would be needed to get a clear picture. For example, the mole fraction change after the start of the heating shown in Fig. 15 is not significantly different from changes observed at higher pressures during the same run. The results of the experiments with changing orientation during the draining are also based on only one run for each experiment, and it is unsure if they can be reproduced. The paper could be shortened and would improve if only the results of the low and high flow experiments, including different cylinder orientation and treatment, are presented.

We disagree with the referee that the heating experiments and changing the position do not add any value to the manuscript. However, we agree that more experiments are needed, which is also stated in the manuscript. This is also why we would like to keep these experiments in the manuscript. These tests with no clear result might inspire other laboratories to do more experiments in this direction to solve this issue. However, we shortened the paragraph at page 17, line 22 from

“*With the three cylinders that were moved upside down, the picture is not very clear. Because of a logger failure , there are no temperature measurements on the first cylinder, which is why it will not be discussed here. The second and the third cylinders put upside down show a drop in the CO_2 mole fraction after they have been moved. While the second cylinder remains stable at the slightly lower CO_2 mole fraction, the CO_2 mole fraction of the last cylinder goes up first by about 0.07µmol mol^{-1} from 30 to 15 bar and falls back by roughly the same amount until the end of the experiment. The temperature measurements of the two cylinders look the same. The only small difference between the two is the pressure when they were turned, the second cylinder was turned at 32.2 bar, the third cylinder was turned at 27.8 bar. However, whether this caused the different behavior in the CO_2 measurements remains unclear. The results from the experiments with constant heating…*”

to

“*With the three cylinders that were moved upside down, the picture is not very clear. Also the results from the experiments with constant heating…*”

I further recommend re-writing the conclusions. Currently, they are difficult to read without the full context of the paper, and present results which are not mentioned previously (e.g. SGS cylinders). Furthermore, the statement ‘This opens the possibility to use a general correction function in case a calibration cylinder on a field station runs empty' should be made in the results section because it needs more careful discussion. Most likely, corrections will be associated with high uncertainties, since the calibration sequence at stations is different from your experiments in which the cylinders were emptied with a constant flow.

We changed the conclusion from:

"*The tested aluminum cylinders behaved always the same within uncertainties, the individual cylinders did not show distinct unique features. This is also true for the SGS cylinders, indicating no benefit in using these tanks in* CO_2 *measurements at ambient level. To describe the* CO_2 *enrichment in low flow settings, the Langmuir adsorption/desorption model using averaged coefficients is sufficient to describe the* CO_2 *enrichment effects in aluminum cylinders. This opens the possibility to use a general correction function in case a calibration cylinder on a field station runs empty. However, we still recommend changing calibration cylinders before the pressure drops below 30 bar in order to avoid the steepest part of the enrichment at the lowest pressures, and the corrections that add uncertainty to the measurements. At the same time the currently recommended threshold of 20 bar (WMO, 2016) is supported by measurements of this study. Using the low flow coefficients for the Langmuir model, a drop from 150 to 20 bar results in a* CO_2 *enrichment of about 0.036* μ*mol* mol^{-1}*, which is still well within the WMO compatibility goal between laboratories.*

In high flow settings additional thermal diffusion effects and Rayleigh fractionation come into play that overrule the simultaneously ongoing Langmuir adsorption/desorption. Depending on the positioning of the cylinder, CO_2 *can be increasing or decreasing with decreasing pressure. However, this might be only the case for systems with a steady high flow. If cylinders are decanted in quick bursts with enough time in between to allow them to equilibrate thermally, thermal fractionation should not be able to develop and only Langmuir adsorption/desorption effects have to be taken into account. Some of the observed effects remain unexplained because the measurements were inconsistent, or the behavior of air in the cylinder needs to be modelled explicitly. To answer these questions additional controlled experiments would be necessary. A further benefit could be gained by using a CRDS (cavity ring down spectroscopy) gas analyzer because it does not need to be calibrated as often as an NDIR analyzer and it could measure several gas species simultaneously.*"

to

"*Six 29.5 L Luxfer L6X® as well as two 29.5 L Luxfer L6X® SGS aluminum cylinders were used to investigate the stability of the* CO_2 *mole fraction of ambient level* CO_2*-in-air mixtures with decreasing pressure. In low flow settings (0.3 L* min^{-1}*), the Langmuir adsorption/desorption model using averaged coefficients is sufficient to describe the* CO_2 *enrichment. With this function, the* CO_2 *enrichment over a pressure range of 150 to 1 bar was calculated to be 0.090 ± 0.009* μ*mol* mol^{-1}*, where the given error corresponds to the standard deviation (1-sigma) of the fitted* CO_2 *enrichment of the individual cylinder drainings. The tested aluminum cylinders behaved always the same within uncertainties, the individual cylinders did not show distinct unique features. This is also true for the SGS cylinders, indicating no benefit in using these tanks for* CO_2 *measurements at ambient level. This opens* *the possibility to use a general correction function in case a calibration cylinder on a field station runs empty. However, we still recommend changing calibration cylinders before the pressure drops below 30 bar in order to avoid the steepest part of the enrichment at the lowest pressures, and the corrections that add uncertainty to the measurements. At the same time the currently recommended threshold of 20 bar (WMO, 2016) is supported by measurements of this study. Using the low flow coefficients for the Langmuir model, a drop from 150 to 30 bar results in a CO_2 enrichment of about 0.026 $\mu mol\ mol^{-1}$, whereas a drop from 150 to 20 bar yields a CO_2 enrichment of 0.34 $\mu mol\ mol^{-1}$.which is still well within the WMO compatibility goal between laboratories. By using bigger cylinders (e.g. 50 L) the surface to volume ratio becomes smaller compared to the 29.5 L cylinders used in this study, which might be beneficial in minimizing the CO_2 enrichment effect at lower pressures. We discourage the use of smaller cylinders as their surface-to-volume ratio increases. Approximating the top and bottom area of a cylinder as a disk perpendicular to the cylinder length (L), and assuming that the effective adsorption area remains the same where the cylinder diameter (R) has been compressed, the surface to volume scales as $(2\pi RL + 2\pi R^2)/(\pi R^2 L) = (L+R)/RL$. We expect a commonly used Luxfer N060 (internal volume 10.7 liter) to be worse by ~30%.*

In high flow settings (5.0 $L\ min^{-1}$), additional thermal diffusion effects and Rayleigh fractionation come into play that add to, or can overrule the simultaneously ongoing Langmuir adsorption/desorption. Depending on the positioning of the cylinder, CO_2 can be increasing or decreasing with decreasing pressure. We have demonstrated that these effects very likely do play a role, but before a satisfactory explanation can be attempted a considerable number of additional controlled experiments, as well as modeling of the flow and mixing in cylinders will be necessary. A further benefit could be gained by using a CRDS (cavity ring down spectroscopy) gas analyzer because it does not need to be calibrated as often as an NDIR analyzer and it could measure several gas species, such as CH_4 or CO, simultaneously."

Specific comments

Page 2, line 13: kilogram is a SI unit, despite the fact that it is still based on an artifact. It should be removed from the list of examples in parenthesis.

"*…, the kilogram or…*" was removed

Page 3, line 21ff: Add a short description of the performance of the analytical system (repeatability, drift etc.) here. This could be done by moving paragraph 3.1 to the method section.

We moved paragraph 3.1 to the method section and changed the paragraph and figure numbering accordingly.

Page 3, line 34: According to Fig. 1, C1 is repeatedly measured, not C2. Page 9, first lines of result section and section 3.1 would better fit in the method section.

Correct, changed to C1.

Page 10, section 3.2.: The low flow experiments are probably the most relevant for most users of standard gases. Beside the average of all cylinders, the result of only one (representative) standard is shown, while in total 38 experiments of the same type were made. It would be valuable to see the variation between different cylinders / fillings, which could be added in either an additional figure or Fig. 6 (e.g. individual fits for all experiments). I recommend to also show and discuss the similar behavior of SGS and untreated cylinders could here.

We added a second panel to Fig. 6, that shows the average fits of the low flow experiments of the normal and the SGS cylinders with the uncertainty given as greyed area (see also similar question above).

Figures 6 and 8: The y-axis shows ΔCO_2, and not CO_2 mole fraction, which needs to be corrected.

We replaced "*CO_2 mole fraction…*" with "*ΔCO_2…*" in Figure 6 and 8. The same is true for Figure 12, which was changed in the same manner on the authors' behalf.

Why do the measurements at higher pressures show a negative delta? Especially in Fig. 8 all ΔCO_2 as well as the fits at higher pressures are negative. Is this correct?

($X_{CO2,ini}$ and $X_{CO2,ad}$ correspond to $CO_{2,ini}$ and $CO_{2,ad}$, respectively, see also reply to Referee#2)

It is correct but might be misleading. For this figures, the term $CO_{2,ini}$ was subtracted from the Langmuir-equation (equation 1) and the remainder with the corresponding coefficients was plotted. $X_{CO2,ini}$ corresponds to the CO_2 mole fraction before adsorption to the walls occurs, therefore it is slightly higher than the CO_2 average value at the beginning of the measurements. This results in negative ΔCO_2 values in the beginning of the measurements. For clarity we shifted the values by adding $X_{CO2,ad}$, this shifts the fits upwards so that the initial value is exactly 0. This was done for Fig 6, 8, and 12. In all three captions we changed

"*...the corresponding $CO_{2,ini}$ was subtracted…*"

to

"*…the corresponding ($X_{CO2,ini} – X_{CO2,ad}$) was subtracted…*",

accordingly.

Technical corrections

Page 2, line 24: the latest available GGMT report is not cited (WMO, 2016). It should be added. The format of the citations of the WMO reports needs also to be changed.

The latest GGMT report was added and the format of the citations were changed.

Page 2, line 28: replace 'SI values' with 'SI traceable values'.

Done.

Page 2, line 35: Cite the latest GGMT report here.

Done.

Page 3, line 36: Change to ‘An additional full calibration was made at the end of each experiment’.

Done.

Page 6, line 25: Kitzis (2017) is missing in the references.

Correct, we added the reference.

References

Brewer, P. J., Brown, R. J. C., Resner, K. V., Hill-Pearce, R. E., Worton, D. R., Allen, N. D. C., Blakley, K. C., Benucci, D., and Ellison, M. R.: Influence of Pressure on the Composition of Gaseous Reference Materials, Analytical Chemistry, 90, 3490-3495, 2018.

WMO: 18th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2015), La Jolla, CA, USA, 13-17 September 2015, GAW Report No. 229, World Meteorological Organization, Geneva, Switzerland, 2016.

Corrections on the authors’ behalf:

We replaced all symbols for liter “l” with a capital L for better readability (also in the figures).

As mentioned above, the y-axis of Fig. 12 was changed from “*CO_2 mole fraction…*” to “*ΔCO_2…*” as requested by Referee#1 for Fig. 6 and Fig. 8.

Page 5 line 21: For consistency, we changed “*normal flow*” to “*low flow*”

Page 9, line 27: For clarity, we changed the first sentence from:

“*No general filtering was applied to the measured data.*”

To

“*No data selection was applied to the measured data.*”

Page 10, line 32 ff: We replaced the two sentences

“*The standard deviation of the average α is very low. If the measurement system’s repeatability as deduced from the target gas measurements is taken into account, a realistic error of α should be about four times bigger.*”

with

“*The given uncertainty range of 0.000004 corresponds to about 9.3 % of (1-0.999957). Considering the calculated CO_2 enrichment of 0.085 μmol mol^{-1}, 9.3 % equates to about*

0.008 µmol mol^{-1}, which is consistent with the measurement system's repeatability of 0.01 µmol mol^{-1} as deduced from the target gas measurements before the analyzer change."

Page 13, line 6: For better readability the two sentences were shortened from

"*If only the CO_2 measurements below 50 % of the cylinder's pressure are used to calculate α, then the average fractionation factor for the outflow becomes 1.00021 ± 0.00004, indicating an even stronger fractionation. When the Rayleigh fractionation with the stronger fractionation factor is only applied after the cylinder is half empty, when the temperature difference between the upper and the lower side of the cylinder body reaches its final value of 0.3 K, the average final depletion is -0.26 ± 0.07 µmol mol^{-1}.*"

to

"*If only the CO_2 measurements below 50 % of the cylinder's pressure are used to calculate α, then the average fractionation factor for the outflow becomes 1.00021 ± 0.00004, indicating an even stronger fractionation with a final average depletion of -0.26 ± 0.07 µmol mol^{-1}.*"

Page 15 line 7: "*vertically*" was replaced with "*horizontally*".

Page 16, line 24: For clarity reasons we changed the sentence from

"*This is also the moment when the temperature difference between…*"

to

"*This is also the reason when the measured temperature difference between…*"

Page 18, line 18:We replaced 0.036 µmol mol^{-1} with 0.034 µmol mol^{-1}

Page 24: For clarity reasons, the caption of figure 3 (now figure 2) was changed from

"*Flow schematic of the high flow inlet system. The sample gas enters on the left side at 5.0 l min^{-1}. A small aliquot of 0.3 l min^{-1} goes to the analyzer, the vast remainder of 4.7 l min^{-1} goes to the exhaust. The ratio between the gas going to the analyzer and the exhaust, respectively, can be adjusted by the needle valve on the exhaust side.*"

to

"*Flow schematic of the high flow inlet system. The sample gas enters on the right side at 5.0 L min^{-1}. A small aliquot of 0.3 L min^{-1} goes to the analyzer, the remainder of 4.7 L min^{-1} goes to the exhaust. The ratio between the gas going to the analyzer and the exhaust, respectively, is set by the dimensions of the inner and outer tube and can be adjusted by the needle valve on the exhaust side.*"

Page 28, line 4: "*vertically*" was replaced with "*horizontally*"

Page 34: We changed the colors of figure 13 for better readability and changed the caption accordingly.

Page 35: The x-axis title of figure 14 was corrected.

Page 37, line 5: “*vertically*” was replaced with “*horizontally*”, the x-axis title of figure 16 was corrected.