Reply to the review of Anonymous Referee #1

The authors would like to thank Anonymous Referee for the valuable comments. In the following, referee's comments are given in bold and author's responses in plain text. Suggested new text is quoted in italics together with page and line numbers.

This paper described the testing of various materials in an aluminum chamber designed such that various materials can be inserted in the chamber and tested for adsorption effects. This paper is a follow-on from a previous paper that described the testing chamber and analytical challenges (10.5194/amt-2019-197). The paper is well-written and contains informative figures. For many of the materials tested, desorption of CO_2, CH_4, and CO were minimal, which is good news for the measurement community. However, it is not exactly clear how the results of these experiments relate to atmospheric trace gas analysis due to significant differences between this work and how compressed gas standards are used in practice.

We understand the concern of our reviewer. The idea behind the setup and the material experiments is the following: compared to a standard cylinder (e.g. 29.5 L or 50 L) the surface to volume ratio of the small cylinders is bigger, and one can further increase the surface area by inserting materials. By testing materials at various pressure and temperature ranges, problematic materials can be identified. The following assumption lies behind these experiments: if the material has adsorbed significant amount of gas while filling the cylinder, this should be desorbed towards the end of the experiments controlled by desorption. The controlled experiments and the numerous materials tested in this study showed even under increased surface areas and lower than normally employed pressure ranges, surface effects for most materials were minimal, meaning that currently used materials are well-suited for their applications in atmospheric measurement community.

The following paragraphs will be added in discussion:

"The presented setup enabled the investigation of surface effects under "extreme" conditions which favored adsorption/desorption. Compared to common usage in the atmospheric measurement and gas metrology communities, our study has differed in cylinder size, geometric surface to volume ratios, pressure and temperature ranges. Previous studies (Leuenberger et al. 2015, Brewer et al. 2018, Schibig et al. 2018) investigating surface effects in compressed gas cylinders have used (50 L, 10 L, or 29.5 L) cylinders. The geometric surface of the small (5 L) aluminum cylinder used in this study is 0.18 m^2, which results in a surface to volume ratio of 35.7 for the unloaded cylinder. Compared to 29.5 L Luxfer cylinders, the small cylinders are estimated to be more prone to adsorption by 40 %. Inserting material blocks into the aluminum cylinder further increased the surface area. Therefore, the setup allows to test materials under increased surface to volume ratios in which the surface effects should be stronger and dominant. However, despite our efforts of increasing the surface material effects were minor.
In addition to the properties of the materials, pressure and temperature play a role on surface effects. The following assumption lies behind the pressure experiments: if the material has adsorbed a significant amount of gas while filling the cylinder, this should be desorbed towards the end of the experiments controlled by desorption. The onset of the desorption for all tested materials except Dursan and partly DLC was observed well-below atmospheric pressures.
Increasing temperature is expected to facilitate desorption by providing the required energy to desorb the gas molecules from the surface and mix into gas phase. On the contrary, cooling the cylinder and its content favor adsorption and it is expected that this results in a decrease in the measured amount fraction."

General comments: You mention that you did not subtract the background, or blank (empty) run from the experiments. In figure 5, it would seem that the empty run shows much the same signals as the materials tested, which I suppose is why you state that only Dursan showed significant difference. I think this needs further explanation related to the other Satar et al 2019 paper (10.5194/amt-2019-197).

We thank our reviewer for pointing this out. In order to clarify this point, the following will be added on page 9, at line 10 and the corresponding paragraph will be rearranged:

"For CO_2 measurements, the amount fraction differences for all materials except Dursan were less than 0.17 µmol mol^{-1}, with slight differences among the various loadings. Of this difference, 0.05 µmol mol^{-1} is related to the blank cylinder (background effect). The blank cylinder corresponded to the "14 bar after heating" case presented in Satar et al. (2019). More information on the blank cylinder and its filling history is provided in the above-mentioned publication. It is also crucial to consider that during all material block experiments, glass pieces were also present in the small measurement chamber. When the material runs were compared to the experiments with glass, except the DLC loading, the remaining differences were in the order of 0.02 µmol mol^{-1}, which corresponded to the 5 minute-standard deviation of the measured data. Moreover, during the evacuation of the measurement chamber with the DLC loading, a slightly increasing trend of -0.004 µmol mol^{-1} bar^{-1} was observed. For the steel loaded cylinder, the experiments where a multiport valve was upstream of the analyzer showed slightly more variation both for final amount fractions and during the pressure run. Whereas, the mass flow controller employed did not have a significant effect on the CO_2 measurements. "

On page 12, you mention that you measured the "blank" or background at the beginning at the end of the experiments, and the "blank" shown in figure 6 is the mean. How for this impact the conclusions? Do you know if the background changed smoothly over time, or abruptly as a results of adding Dursan or DLC?

Our concern about the repeated heating cycles on page 12 line 6, was the change of the background effect of the measurement chamber only due to heating. Inserting any of the materials should not have changed the background effect of the cylinder. During the experiments the blank cylinder has evolved to be better after re-use and refilling with the same gas. Therefore, we don't think that the inserted materials can result in irreversible effects.

For clarity the sentence on page 12 line 6 is modified to:
"This presumably resulted in a change of the background effect over the course of the presented analysis in the range of 0.04 µmol mol^{-1}."

Further, I think you can better explain why you chose the maximum deviation (page 7, line 9), given that making measurements at sub-ambient pressure introduces complications, as described in the other Satar et al 2019 paper.

In our opinion, the maximum deviation is the most suitable parameter, since we are interested in the maximum possible effect which might occur related to the desorption process. Moreover, all data was processed using the same criterion cancelling out the effects originating from the analyzer and enable to focus on the differences between the materials.

The following will be added on page 7, line 9:
"Maximal difference was chosen to highlight the maximum possible effect related to desorption. Since all data was processed using the same criterion, the effects originating from the analyzer is cancelled out and we focus on the differences between the materials."

How do these results compare with others studies (Leuenberger et al 2015; Schibig et al 2018) that suggest that a Langmuir isotherm can be used to model the physical adsorption? Your tests seem to show a much steeper increase at the lowest pressures. Is the abrupt increase at the lowest pressure related to desorption, or is it complicated by analysis at low pressure? I realize the analysis is discussed in the first Satar et al 2019 paper, but I think some important aspects need to be repeated here.

We understand the view of our reviewer that this point needs further explanation, we have already provided more information on this comparison in the revised version of the Satar et al 2019 paper and its supplementary material. However, it is crucial to note that the aluminum cylinder without loading presented in this study corresponds to the case “14 bar aluminum after heating” and showed a maximal deviation from the initial amount fraction as low as 0.05 µmol mol^{-1} making the modelling of this increase extremely difficult. For the material loadings, adsorbed amounts should be distributed between the blank cylinders and the material blocks. However as previously explained, subtracting the maximal amount fraction difference of the glass loaded cylinder from the material loaded cylinder resulted in amount fractions in the order of the standard deviation of the measured data. Therefore, for the majority of the materials fitting the Langmuir isotherm is not reasonable.

Since it not common to use a cylinder of gas down to less than 10% of the fill pressure, what would these results say about that practice?

These results highlight that the currently used materials are non-problematic at various temperature and pressure ranges, which is good news for the atmospheric measurements community and gas applications. The experiments were conducted under “extreme” conditions to understand possible surface effects to its full extent. Please see the explanation on the first paragraph of the replies.

I am also concerned about how to interpret the results with respect to materials used by the scientific community. The inside of an aluminum cylinder, for example, does not typically look like the outside. Presumably the process of manufacturing a cylinder (using a ram rod) alters the surface roughness of the inside, and may deposit trace elements on the internal surface. While you tested the same aluminum alloy used for cylinders, the tested materials might not be representative of actual cylinders. Have you tested pieces of material cut from aluminum or steel cylinders? Maybe it doesn’t matter since the results don’t show significant desorption, but this should at least be discussed.

We thank our reviewer for pointing this out. We have not tested materials cut from aluminum or steel cylinders. This can be done within the scope of another study. Regarding other tested materials, commercial coatings aiming to provide an inert surface undergo specific cleaning and chemical or physical vapor deposition of processes, which should be consistent regardless of the coated piece. Concerning copper and brass, these materials might be used in regulators or as seals, and not commonly used for cylinders. Unfortunately, one individual study with a new setup is not capable of answering of all questions about materials employed in the atmospheric measurement community. The study was a first step in understanding surface effects by increasing surface areas, for a significant number of materials under various pressure and temperature ranges with a sufficient number of replicates.

The following will be added into discussion:

“Testing pieces cut from the aluminum and steel cylinders commonly used in the community would be a valuable addition to enable direct comparison between the commonly used cylinder materials and the produced material blocks at low pressures and high temperatures. “

Specific Comments

P1, L5: suggest: "For this study we used small-volume chambers designed to be used for adsorption studies". – text modified accordingly

P3, L30: suggest: "A mother cylinder was...." – text modified accordingly

Figure 4: Minor point: I'm not sure of the significance of the box plots based on only 3 data points. I guess you are assuming normal distributions and assigning quartiles based on the standard deviation.

We thank our reviewer for this comment. In the box plots R uses the "fivenum" function which is the descriptive statistics for minimum, 1st quartile, median, 3rd quartile and the maximum. In case of 3 data points, the 1st quartile is the mean of the minimum value and the median, whereas the 3rd quartile is the mean of the median and the maximum value. Therefore, the whiskers extend to the minimum and the maximum values, respectively. Since in Fig. 4 and Fig. 7, all data points are shown in addition to the box plots, using box plots for the figures is non-problematic.

On page 7, at line 10, we will add the following:

"The median is denoted by the horizontal line, whereas the mean is shown by the square. Since for most cases only 3 replicates are present, the 1st quartile is the mean of the minimum and the median, whereas the 3rd quartile is the mean of the median and the maximum value. For clarity, data points used for the box plots are also shown and they are denoted by the black points. "

We have also rearranged Fig.4 to better distinguish the mean (denoted by the squares) from the results of the replicates (denoted by the black points)

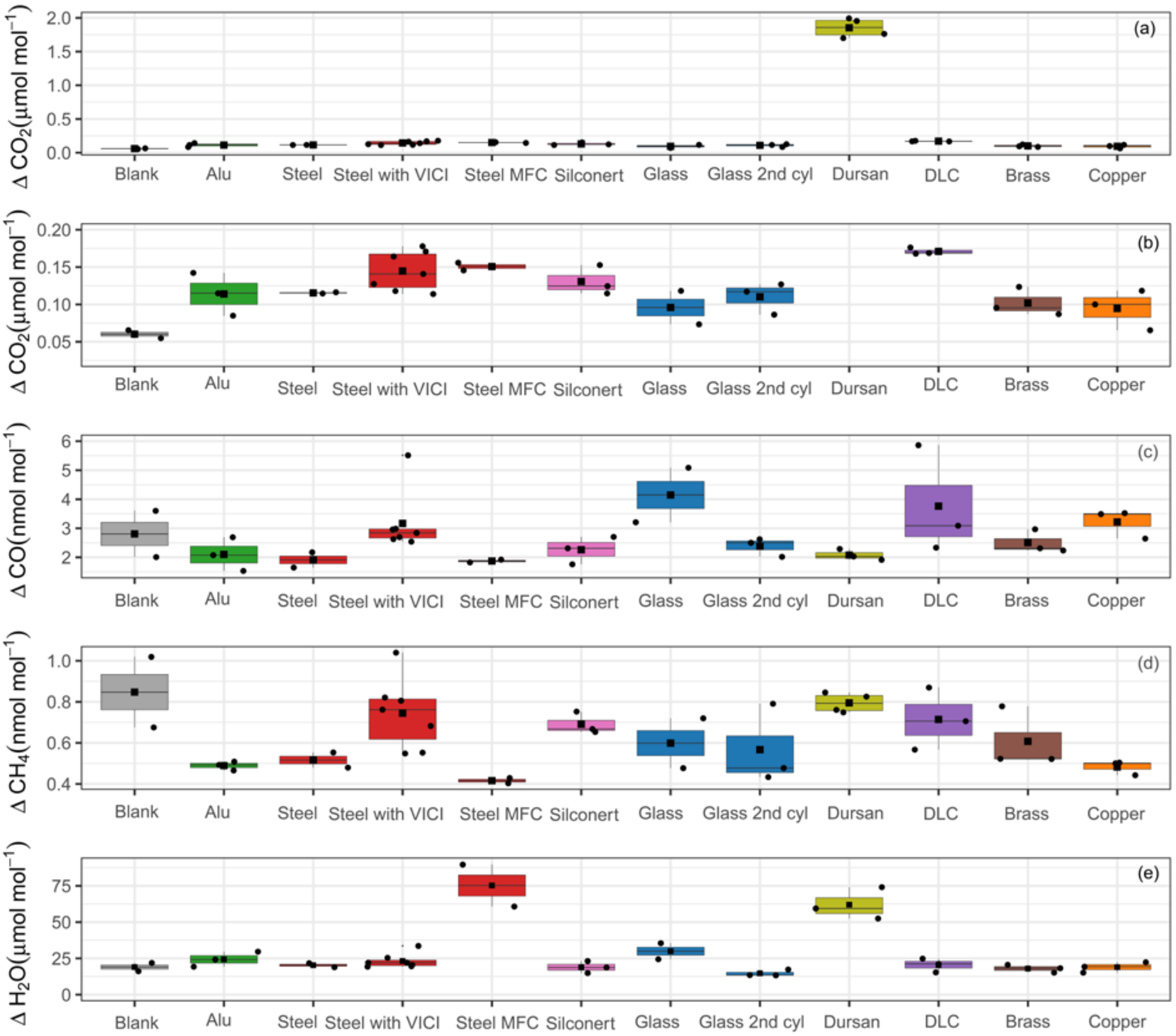

Figure 4. Box plots for all materials for the species (a) CO_2, (b) zoom-in for CO_2, (c) CO, (d) CH_4 and (e) H_2O. y-axes show the maximal amount fraction difference relative to the initial amount fraction. Horizontal lines in each box plot shows the median, whereas the square in the center of the box is the mean of the maximal amount fractions of the replicates.