Reply to anonymous referee#2

We thank referee#2 for the careful examination of the submitted manuscript, the useful remarks made with respect to content and organization of the paper. Both referee#1 and referee#2 suggested several similar improvements. As we worked through the reviews sequentially (starting with remarks of referee#1) the document with change tracking assigns most of the changes to referee#1 (red), but we assume that referee#2 will nevertheless acknowledges that we did not neglect his/her recommendations. We comment in detail below.

We have to point out an important technical error in our original data analysis of the low-resolution spectra: contrary to the claim made in the manuscript, a set of atmospheric a-priori profiles different from that used by TCCON was used. We corrected this flaw in the revision and redid the complete analysis for all species. Therefore, the O_2, CH_4, and XCO data derived from the low-resolution spectra all slightly changed, and the figures and derived numbers have been updated accordingly. This correction essentially removed the 2% calibration offset in XCO derived from the prototype.

Although the paper well describes the new instrumental set-up, results do not support the conclusions. The results need to be enriched by more quantitative and thorough comparisons between the prototype instruments and other reference sensors. Descriptions of comparisons should be strengthened: instead of comparing time series, I would recommend using 1:1 (or scatter plots type) plots for more clarity and quantifications of the results.

We have added a second panel to Figs. 8, 9, and 11 which provides correlation plots between daily mean results from the prototype as function of the reference measurement. We decided to base this statistical analysis on daily mean values, because the measurements were in general not synchronized and integration times differed.

The request for adding comparisons with other reference sensors seems somewhat excessive to us for achieving the purpose we are aiming at in this work: an empirical demonstration that the novel low-resolution XCO prototype spectrometer is performing sufficiently well to be useful for source assignment and satellite validation activities. The TCCON is a reliable source of XCO (if very low solar elevation angles are avoided) and has been calibrated against other in-situ reference sensors in the framework of many campaigns at different TCCON sites. Due to the fact that it uses the same remote sensing geometry as the device under test (therefore probing the same air mass), that it delivers a sufficient number of data points within a measurement day, and that it offers at least similar sensitivity characteristics, the TCCON seems to us the ultimate practicable reference source for the desired task. However, we have in the revised version of the work included mid-infrared XCO observations, because these support our assumption that the remaining discrepancies between the prototype results and the TCCON data are due to a mismatch in vertical sensitivities.

Section 5 should be improved and organized in sub-sections for more clarity.

This overly long section has been split in several paragraphs.

Why only using 6 days for the comparisons? Is it statistically adequate? Please comment.

We have included an additional measurement day in mid of March, 2016, during which the prototype was also operated together with the other mobile spectrometer used as a reference and the TCCON spectrometer. The XCO validation is therefore now based on seven days, during which both the prototype and the TCCON spectrometer performed observations, together these span a 5 month interval. The XCO value in March was significantly higher than the other measurements taken in autumn, and is also compatible with the assumption of a stable calibration factor of the prototype with respect to the TCCON. The statistical confidence of the prototype XCO calibration slope is in the order of 0.5% (1 σ), which seems adequate, given that the variability of the atmospheric XCO is considerable. In order to put the results of our calibration in perspective, we have added the following statement: "The demonstrated performance of the prototype indicates that the design is well suited for source attribution (Wunch et al., 2009, detected intraday XCO enhancements of up to 30 % in the Los Angeles Basin) and satellite validation (the TROPOMI accuracy and precision targets are 15% and 10%, respectively).".

The other important aspect requiring empirical verification is that the modification preserves the characteristics of the primary channel. This check is based on fourteen measurement days (seven days before and six days after the modification was performed – during one day after modification of the prototype, the unmodified EM27/SUN used as reference was not operated). For proving that no significant change of the primary channel characteristics occurred, a single day before and a single day after the modification of the prototype would already be sufficient. The fact, that we have used several days allows demonstrating the calibration stability of the O_2 column to a very high level of confidence. It is demonstrated on the 0.02% level (Figure 8, lower panel). We would already be fully satisfied with a consistency on the 0.1% level. The TCCON claim of achieving a station-to-station consistency of about 0.1% in XCO_2 may serve as a relevant comparison benchmark (which is in fact still less ambitious than a demonstration of the O_2 column consistency on the 0.1% level, as we describe in detail below).

Use scatter plots (or other) for Fig 8, 9, 11, and 12, instead of tome series. Why the TCCON data are not used as a reference for such a validation. These data should be added to the comparison plots. Why not comparing XCO2 as well?

We have added additional panels to Figures 8, 9, and 11. Figure 12 only intends to suggest in an illustrative manner that the prototype spectrometer can detect variabilities below the 1% level, we therefore have not added a statistical analysis in this case.

Concerning XCO_2, comparisons between the EM27/SUN spectrometer and the TCCON have already been performed (Gisi et al., 2012; Frey et al., 2015). Here, we are only interested in demonstrating that the characteristics of the primary channel remain unaffected by the addition of the second channel. Comparing the O_2 columns is a much more sensitive test than investigating XCO_2, because the latter involves the rationing of the CO_2 column by the O_2 column. Most instrumental problems affect the target species column and the O_2 column to a similar degree and therefore tend to compensate at least partially in, e.g. XCO_2. For illustration, Hase et al., 2013, showed that an error in instrumental line shape affects the O_2 column four times as strong as it does affect XCO_2. Indeed, this tendency of compensation of errors was an important design consideration of the TCCON. In order to clarify our motivation of using the O_2 column, the manuscript contains the following remark "Note that a comparison of oxygen columns is a very sensitive test, as many instrumental errors tend to cancel out in the final column averaged abundances of the target species.".

Avoid the word "excellent agreement" in the actual comparisons.

Thanks, we have changed the wording accordingly.

- What is the precision of the prototype XCO, XCH4, and XCO2? How these precision compared to standard EM27/SUN?

There is no physical reason to expect a change in precision of XCH_4 and XCO_2. From the instrumental viewpoint the primary channel, which delivers XCH_4 and XCO_2, is still the identical setup. The small signal loss due to the decoupling of the secondary beam is compensated by slightly increasing the diameter of the adjustable iris which defines the aperture stop. The precision of XCO can be estimated from differences of successive measurements (0.35 ppb). We have added this information in the text.

- In figure 11, the agreements for last 3 days are different than the others days. Could you explain?

In the lower panel of Fig. 11, the data points for Nov. 11, and Nov. 16 are located below the calibration line by about 1.5%. The vertical sensitivity of the low resolution spectrometer and the TCCON measurement are not identical and this might trigger smoothing errors of this size if the a-priori profile was a poor approximation of the actual atmospheric state during this period. Before, when we erroneously used a different set of a-priori profiles in our analysis, we came up with a systematic smoothing error component of about 2.2%.

We have added a discussion of vertical sensitivities and the smoothing error: "The vertical sensitivity of the low resolution spectrometer differs somewhat from the TCCON sensitivity. This is a consequence of differing spectral resolutions, and introduces a smoothing error component in the comparison. Sensitivities for both the low-resolution spectrometer and TCCON are shown in the work by Hedelius et al., 2016, indicating that the CO sensitivity of the low-resolution spectrometer is even superior to that of the TCCON. In this study - focused on an instrumental validation - we minimized the systematic smoothing error contribution by using the same atmospheric CO a-priori profiles as the TCCON in our spectral analysis. An in-depth quantification of the smoothing error has been performed by Kiel et al., 2015b. The authors compared TCCON results with data derived from mid-infrared spectra (with a comparable degree of mismatch of vertical sensitivities) and found smoothing error contributions of a few per cent due to different a-priori profile choices and vertical sensitivities. We have included mid-infrared observations of XCO in the lower panel of Figure 11, because these results support our explanation of the remaining discrepancies being generated by different vertical sensitivities of the TCCON and the prototype measurements: the mid-infrared observations of XCO tend to be low compared to TCCON when the prototype results are low and high when the prototype results are high."

Figures 5 and 6 can be combined together.

Done.

Table 1 and 2 should be re-organized in one Table.

Done.

Figure 7, could you add the interfering species in the window?

We have added an additional figure (now Figure 6) which indicates all relevant absorbers in this spectral window. The spectral scene is shown at high spectral resolution, as the region is so crowded with absorption features that it is difficult to discern all the individual contributions by eye in a low-resolution spectrum.

References

Frey, M., F. Hase, T. Blumenstock, J. Groß, M. Kiel, G. Mengistu Tsidu, K. Schäfer, M. Kumar Sha, and J. Orphal: Calibration and instrumental line shape characterization of a set of portable FTIR spectrometers for detecting greenhouse gas emissions, Atmos. Meas. Tech., 8, 3047-3057, doi:10.5194/amt-8-3047-2015, 2015.

Gisi, M., Hase, F., Dohe, S., Blumenstock, T., Simon, A., and Keens, A.: XCO_2-measurements with a tabletop FTS using solar absorption spectroscopy, Atmos. Meas. Tech., 5, 2969-2980, doi:10.5194/amt-5-2969-2012, 2012.

Hase, F., B. J. Drouin, C. M. Roehl, G. C. Toon, P. O. Wennberg, D. Wunch, T. Blumenstock, F. Desmet, D. G. Feist, P. Heikkinen, M. De Mazière, M. Rettinger, J. Robinson, M. Schneider, V. Sherlock, R. Sussmann, Y. Té, T. Warneke, and C. Weinzierl: Calibration of sealed HCl cells used for TCCON instrumental line shape monitoring, Atmos. Meas. Tech., 6, 3527-3537, doi:10.5194/amt-6-3527-2013, 2013.