Answers to the comments of “Anonymous Referee 1”

First of all, we would like to thank you for the valuable feedback and comments.

In the following we will reply to your questions and comments point-by-point. Your questions and comments are given (sometimes a bit shortened) in bold font, and our answers in normal font. Extracts of the paper are added in italic font.

Note that two errors were discovered by us in the meanwhile which also have been corrected:

1. In Table 1 and Table A2 there were wrong numbers. Fortunately, only the numbers in the manuscript were erroneous but not the ones used for the calculation of the final result. Hence, nothing changed in Table 4 due to these corrections.
2. We discovered an error in the Excel table used to calculate the final corrections factors for the Tsukuba time-corrected data. This error is corrected now, causing that the relative deviation for the XCO_2 TK-LR-tcorr data is now similar to the XCO_2 TK-LR data.
   This caused the values for “TK t-corr -44s” to be changed in Table 4 and Table A2:
   a. Table 4:
      XCO_2: 0.16401 ± 0.00830 − 0.06318 → 0.11387 ± 0.00829 − 0.06314
      XCH_4: −0.1115 ± 0.00873 + 0.06690 → −0.18343 ± 0.00871 + 0.06685
      XCO: 1.46537 ± 0.0487 + 0.05470 → 1.16653 ± 0.04870 + 0.05454
   b. Table A2:
      XCO_2: 0.99836 ± 0.00008 + 0.00063 → 0.99886 ± 0.00008 + 0.00063
      XCH_4: 1.00112 ± 0.00009 − 0.00067 → 1.00185 ± 0.00009 − 0.00067
      XCO: 0.98556 ± 0.00047 − 0.00053 → 0.98847 ± 0.00047 − 0.00053

   This also caused changes in Figure 16 (which is the visualization of Table 4) and the main text in Section 7.3:

General Comments

How often should Travel Standard visits be conducted?

- This is a good question and the ideal answer would be "as often as possible".
- More differentiated, it would be useful to re-visit sites after any major instrumental intervention, or as in case of the timing error in Tsukuba, after a recognized problem is solved (or reprocess the data if it is solved software-sided). Drifts or sudden changes in the XAIR time series of a TCCON site turn out to be good indicators for the occurrence of instrumental changes that would make another site visit desirable.
- However, as one can see, the TS activities are quite an effort: The study described in this article ran for 15 month and we visited 3 sites. This long-time span of in average 5 month per site is caused by a) the time needed for the shipment and customs procedures b) the time needed at each site to collect sufficient data and c) the time for maintaining and assuring the proper characterization of the TS at KIT. Assuming this as an average speed, it would take around 12 years to visit all TCCON sites, which is too long. To speed this up, there are several approaches.
    - The rigorous solution would be to use several EM27/SUN being as TSs which are all based in Karlsruhe. Already, the use of a pair of spectrometers sharing the same shelter already would speed up the procedure. Hence, it is possible to have one EM27/SUN at a

campaign with the shelter, and the other is simultaneously collecting side-by-side measurements in Karlsruhe without a shelter.
    - Furthermore, it would be possible to combine several sites (e.g. visit more than one sites in the US before returning to KA). However, operating the TS en route without re-characterization in between would increase the risk of drifts and reduce the success chances of prolonged campaigns.
    - Also, possible would be to use other EM27/SUNs to compare TCCON sites close by to each other (e.g. all sites in Japan) and only visit one of the sites with the TS.
    - Furthermore, the approach of cascading down the global inter-continental calibration achieved by the TS presented here looks very promising. The TS would then visit a certain TCCON site in a region and the occasion of this encounter would be used to collocate further COCCON spectrometers operated by groups of this region during the TS visit. These spectrometers are used subsequently to spread out the results to further TCCON sites nearby.
- These thoughts are now included in the outlook of the paper:

  "*To make use of the valuable insights provided by the TS it would be desirable to visit TCCON sites regularly. However, the TS activities are quite some effort as it can be seen in this study. Continuing with the same speed would take around 10 years to visit all the TCCON sites (~ 3 per year). To speed this up different approaches are possible: The most direct one, which is already planned, would be to use several closely monitored EM27/SUNs to be used as TS in parallel sharing the same enclosure. This helps to increase the frequency of campaigns as one of the EM27/SUN spectrometers can be sent to a campaign whereas in parallel the other can perform side-by-side measurements in KA. Also, it would be possible to visit several sites between two calibration stops at KA. However, this would reduce the accuracy as the TS is less closely monitored. Another approach would be to visit one TCCON site per country and transfer its level to surrounding sites by using other EM27/SUN, which of course, must be monitored closely, too.*"

Ultimately, we need to tie all the measurements back to the WMO scale. If the KIT TCCON instrument is going to be the reference, should aircraft overflights or AirCore launches be conducted more regularly at the site to keep that instrument tightly related to the WMO scale? If, for logistical reasons, the KIT TCCON site isn't ideal for overflights/launches, could another site be used for this purpose and the Travel Standard relate that TCCON instrument to the KIT instrument?

- We agree that it is of great importance to tie the TCCON to the WMO scale. However, the idea of the TS is not to improve the absolute calibration of the TCCON but to investigate and increase the consistency within the TCCON. Specifically, we try to reduce station-to-station biases across TCCON caused by instrumental imperfections of real-world spectrometers.

- To make this intention clearer we added the following to the introduction:

  We changed "In order to produce reliable reference data, it is important to ensure that the network as a whole is accurately tied to the World Meteorological Organization's (WMO) trace gas scale (Hall et al., 2021, Dlugokencky et al., 2005), and that the network has minimal station-to-station biases."
  to
  "To produce reliable reference data, two things have to be considered. The first item is to ensure that the network as a whole is accurately tied to the World Meteorological Organization's (WMO) trace gas scale (Hall et al. (2021), Dlugokencky et al. (2005). The second is to minimize station-to-station biases across the network due to non-nominal behavior of the spectrometer."

  Furthermore, we added (the addition is underlined, the italic text is given for context):

"*In this work an additional method of further enhancing the TCCON's quality management is presented and applied. It is based on a portable EM27/SUN FTIR spectrometer operated in the framework of the Collaborative Carbon Column Observing Network (COCCON) (Frey et al. (2019)) which will be used as a traveling standard. This activity aims directly at the improvement of the site-to-site consistency.*"

Specific Comments

1. **Define XGas earlier in the paper.**
    - Is now defined in the introduction.
2. **Line 23: Explain what kind of pressure is meant by "pressure analysis"**
    - Added "An important auxiliary value for FTIR retrievals is the surface pressure. Using the pressure sensor onboard the TS, the surface pressure measurements at each site are also compared. The surface pressure analysis reveals..." at line 23.
3. **Line 49/50: Provide reference for WMO trace gas scale.**
    - Added "Hall et al. 2021", https://doi.org/10.5194/amt-14-3015-2021 for CO2
    - Added "Dlugokencky et al. 2005": https://doi.org/10.1029/2005JD006035
4. **Lines 128 - 133: No additional info here than in the Intro. Trim the Intro.**
    - Deleted line 56 and following in intro: "However, the collection of such a profile data set is laborious, expensive and the number of available in-situ profiles is too small for detecting minor biases of individual TCCON sites. Moreover, TCCON sites located in populated regions with severe flight restrictions are particularly difficult to address with this strategy."
5. **Adding a subscript or similar to mark, that XAIR from PROFFAST and GGG are different.**
    - Added "To make this clear we add the subscript "GGG" to the XAIR labels to indicate that we are using the standard GGG XAIR values and the inverted PROFFAST XAIR values." to the main text.
    - Added XAIR_GGG to the labels of the figures and an explanation to the caption.
6. **Line 172-178: Redundant text earlier to this section Remove for brevity.**
    - These lines were intended to be a summary. They have changed to: "In summary, we believe that the COCCON-TS for the TCCON presented in this paper is a valuable complement to the methods presented above"
7. **Line 153 - 155: Provide units for all quantities.**
    - When possible units and quantities were added.
8. **Line 212 - 216: Description of the logging events is not clear.**
    - Reformulated these lines, added times of the logging events
9. **Line 218-219: Sentence is not necessary. Remove for brevity.**
    - Has been removed.
10. **Line 309: "It is assumed that SN37 is constant. How is this assumption justified?**
    - This is justified by a long-term analysis of the TCCON-KA and SN37 data as shown in Alberti et al. 2022: https://doi.org/10.5194/amt-15-2433-2022, Figure 20.
    - Added "This assumption is justified by a long-term analysis of the reference EM27/SUN spectrometer (SN37) with the TCCON-Karlsruhe data as shown in Alberti et al. 2022, Figure 20."

11. **Lines 342-349: This correction approach effectively assumes the dependence on SZA is 100% in the Travel Standard instrument. How does the uncertainty in this assumption propagate into the calculation of absolute uncertainty of the CO measurements?**

    - For all gases we use the side-by-side measurements before and after each campaign to derive an upper threshold for drifts of the XGas results during the campaign.
    - Hence, the smaller the differences of the TS to the reference before and after the campaign, the smaller the uncertainty.
    - With the CO correction, we try to match the COCCON SN37 reference data as closely as possible. Any deviation from this (i.e. an inaccuracy of the CO-correction) is captured by the difference of the bias-compensation K_SN39^SN37 CO before and after the campaign. This difference is then propagated into the uncertainty of the final result.
    - Hence, the used method of deriving the calibration uncertainties implicitly includes the uncertainty of the CO correction.

12. **Lines 353 - 354: Why is TCCON KA not using own p Sensor?**

    - This is for historical reasons but offers several advantages:
        - Several comparisons with pressure sensors operated at KIT have proven that the DWD-sensors data agree with test-measurements at KIT (see also the comparison in this paper.)
        - The DWD-sensor is part of the operational weather service network and therefore monitored closely and the data are quality checked. Hence, we assume that the surface pressure data are of reliable quality.
        - Furthermore, we do not have to worry about maintaining and calibration of the sensor.

13. **Fig 8 Add to caption that TK-LR GGG only plotted for XCO.**

    - Added in the text: "Note that in Figure8 the GGG values are only plotted for XCO."
    - Added in the caption: "Note that for XCO the TK-LR data is also processed with GGG2020 and plotted using black triangles."

14. **General: Change caption and explanation of colors + shapes used in the plots**

    - Changed caption and text for Fig 1.
    - Caption Figure 5: Changed to "blue triangle shaped markers".
    - Caption Figure 8: Changed colors and added markers.
    - Caption Figure 9: Changed marker and colors.
    - Caption Figure 10: Changed colors
    - Adapted the text referring to Figure 10.
    - Section 6: Changed description of Figure 12.
    - Line 434: Yellow x-shaped markers --> black triangles
    - Lines 448, 449, 464: Colors in Figures do not match --> Now, they do
    - Line 512: WG-HR are red not green --> adapted to new colors.