Response to comments from Referee 2
Black: Referee's comments; Blue: Authors' answers; Green: Changes in the manuscript

We thank the referee for the review and for providing useful feedback, which we consider in the revised version of the paper.

Referee:
A. General Comments

The advantage of the TROPOMI measurement is its capability to cover entire globe in a single day with higher spatial resolution. Readers are interested in its validation for the data with large satellite zenith angles and how accurate the fast L2 retrieval algorithm is. The present manuscript looks like a technical report. The research paper must be concise and needs analysis for root causes of bias. The manuscript includes many redundant portions, which must be shortened. The abstract and the conclusion are also too long. Major revision is needed.

Authors' response:
Our intention of including S5P CH_4 and CO validation results in one paper was to make use of the common description of the reference data sets and validation techniques description. We have provided possible reasons where large deviations in bias are observed.

The trace gas products from both satellite and ground-based remote sensing measurements provide the best estimate of the atmospheric state via a retrieval of the measured radiance spectra. However, there are several interfering parameters that are different for the two cases, which result in the respective uncertainties in their evaluation and comparison. Both methods, i.e., a direct comparison of satellite and ground-based reference data as well as a comparison with some corrections applied to one or the other data set, have their own advantages. In case of a direct comparison, we can get an estimate of the magnitude of differences due to some of the interfering parameters. In the latter case, we try to align the satellite and ground-based products as much as possible and then check for the differences. We believe, and also as pointed by Referee 1, that this information is relevant for the users of TROPOMI CH_4 and CO data and therefore we find it useful to include these different cases in the paper.

As suggested by the referee, we have removed one sub-plot from each of Fig. 8 and Fig. 16, both-sub plots from Fig.17 and Fig. 36 and adapted the discussions in the main text accordingly to reduce the length of the paper.

Referee:
I have following suggestions.

(1) Match up condition

As TORPOMI has higher sampling density and spatial resolution, stricter match up conditions can be applied than for existing instruments such as GOSAT and OCO-2. 100 km or 50 km is too long for the sites located near urban area such as Saga.

Authors' response:
We agree with the reviewer that the co-location criteria for TROPOMI can be stricter as compared to the GOSAT and OCO-2 validation studies.

We found typical examples of co-location criteria used for GOSAT and OCO-2 validation studies: Noël et al. (2021) used a maximum spatial distance of satellite measurements from TCCON station of 500 km and maximum time difference of 2 h for GOSAT and GOSAT-2 validation. Parker et al. (2020) used all GOSAT soundings within ± 5° of a TCCON site for GOSAT XCH_4 data validation. The co-location criterion used by Wunch et al. (2017) and O'Dell et al. (2018) for OCO-2 validation against TCCON requires the OCO-2 footprint to be within 2.5° latitude and 5.0° longitude of the TCCON station and the observations that occurred within 2 h of each other.

We have tested several co-location criteria and found that for CO, a co-location radius of 50 km gives robust results for the global stations in the networks (Sha et al., 2018; see extracted corresponding plot below - Figure 1).
However, when applying the same 50 km co-location radius we do not get robust statistics for CH_4. This is due to the fewer TROPOMI CH_4 pixels available, in comparison to CO, due to strict pixel filtering (clouds, SZA, …) and operational CH_4 pixels currently available only over land. As a result, we have relaxed the co-location radius to 100 km for CH_4 validation study.

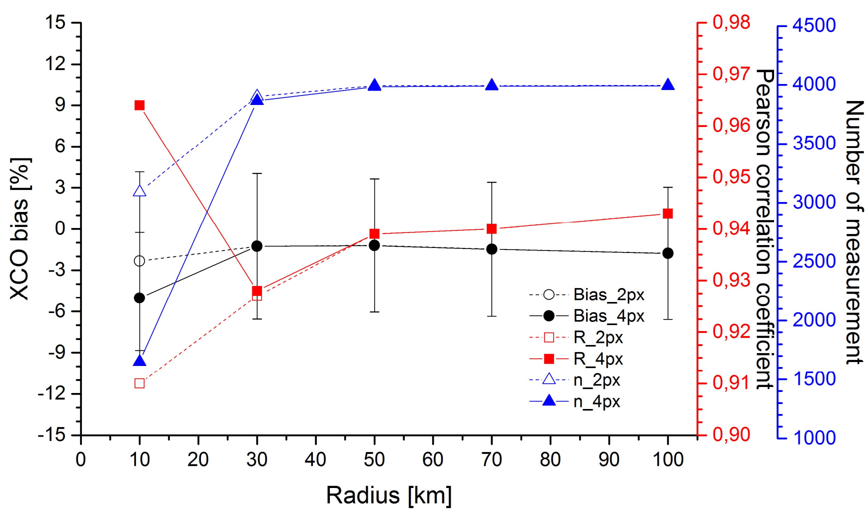

Figure 1: XCO bias for Lauder as a function of the coincidence criterion of radius. Solid shapes are for coincidence > 4 S5P pixels and empty shapes are for >2 S5P pixels.

We agree with the reviewer that a stricter match up condition based on site characteristic might be useful. However, our goal was to use the same strict co-location criterion, which is valid for all stations in the network. This is also the reason why we have implemented a stricter cone selection criterion where we follow the ground-based FTIR line-of-sight with a defined opening angle of the cone. This criterion is especially effective in pixel selection for stations located close to regions with high emission sources, where there are possible scenarios when the ground-based FTIR line-of-sight is not covering all pixels observed by the satellite using the circular co-location criterion (an example is given in Figure 26 of our discussion paper).

Referee:
(2) Summary table

There are several numbers of systematic errors in the abstract, the main text, and the conclusion. The summary table with numbers and conditions will help readers' understanding.

Authors' response:
Taking into account the comments of Referee 1, we have added the Standard deviation of all stations values along with the Mean of all stations to each validation settings in separate rows at the end of each table. In view of not increasing the length of the paper, we hope that these two columns at the end of each table will help the readers' understanding and to get a quick overview of the results for the respective settings.

Referee:
(3) NDAC

There are already plenty of match up data with TCCON. Explanation why NDAC data is additionally need, is required in more detail. Do authors need more data at high latitude stations with large solar angles?
Authors' response:
Although TCCON and the FTIR NDACC instruments use the same high-resolution FTIR spectrometer, the instruments are configured differently and the retrieval method is different. TCCON measures dry air averaged columns while NDACC measures vertical profiles with limited vertical resolution (typically 2 to 3 partial column are distinguished). Both networks have proven to be valuable in the comparison with satellite data and differences between the networks allow to better understand the subtleties in the comparison. Besides this, making use of the measurements from both networks allows a better coverage of different land types, geographic locations, atmospheric conditions, … .

Referee:
(4) Geometry dependency

Authors mentioned solar zenith angle dependency. TROPOMI has wide cross-track coverage. Is there also viewing angle dependency? Is the bias due to forward calculation error by the radiative transfer model used in the L2 retrieval?

Authors' response:
The relative bias for CO plotted as a function of the viewing zenith angle (VZA) is shown in the figure below (Figure 2). We do not observe any significant dependence of the bias on the VZA.

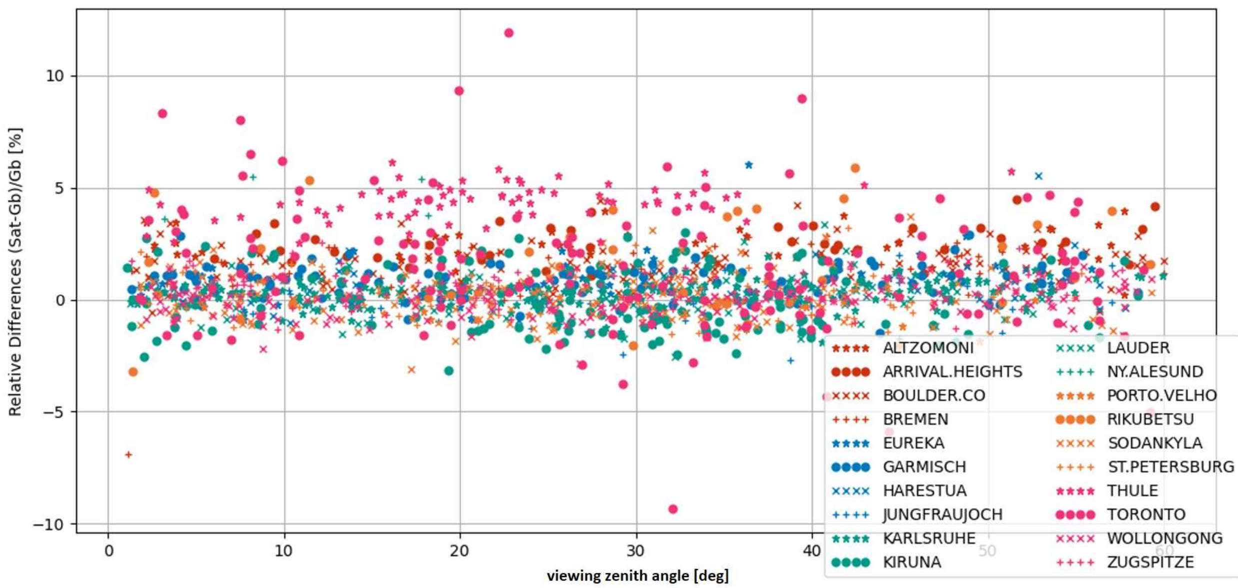

Figure 2: Relative biases for CO plotted as a function of the S5P measurement viewing zenith angles.

Specific retrieval conditions happen at specific solar zenith angles (SZA). The bias related to the surface conditions at specific sites happen during specific moments of the year (e.g. snow/ice and vortex at Sodankylä and East Trout Lake). This is not a problem of the forward model. The forward model (if so) would add a dependency on the relative azimuth angle, which is related to the BRDF effects, but not on SZA alone.

Referee:
Discussion on how to reduce bias such as SZA dependent and surface-albedo dependent ones will be useful for readers.

Authors' response:
The future S5P operational CH_4 bias correction will be done using only S5P CH_4 measurements (see Lorente et al., 2021; for details).

Referee:
B. Specific Comments

(1) Abstract, page 2, line 12
A brief explanation of "QA" in the abstract is needed.

Authors' response:
Referee 1 also had a comment on this line. As a result, we have modified the sentence and removed the QA value from it. The modified sentence is as follows:
"We found that the S5P standard and bias-corrected methane data over land surface for the recommended quality filtering fulfil the mission requirements of bias (systematic error) less than 1.5% and random error less than 1 %."

Similar modification is also done for the carbon monoxide statement.
"We found that the S5P carbon monoxide data over all surfaces for the recommended quality filtering in general fulfil the mission requirements of bias (systematic error) less than 15% and random error of <10%."

Referee:
(2) Page 2, line 16, Page 8 line 207, "Smoothing uncertainty", Page 8
It appears in the abstract. In the main text, it appears first in Page 8. Brief explanation will help readers' understanding.

Authors' response:
We have rephrased the sentence in the abstract to make it clearer at the first instance of the occurrence of smoothing uncertainty.
"The contribution of uncertainty due to smoothing at the individual stations was estimated and found to be dependent on the location."

Referee:
(3) Page 8, Line 209

Detailed explanations on TCCON site are not main topics of this paper. The information is available in the TCCON WIKI and not needed in the main text.

Authors' response:
We agree with the reviewer that the detailed review of all TCCON sites is not needed. This is the reason why no such explanation was given in section "2.2 Ground-based TCCON reference data set". However, we have provided a short description giving examples of a few sites and their corresponding features, which are useful for S5P CH_4 validation result analysis and interpretation. We find this information useful for the readers who are not familiar with the TCCON sites. For the same reason we have provided a short description of the NDACC stations giving examples of a few sites and their corresponding features. However, for NDACC no such wiki exists.

Referee:
(4) page 9, line 241, "a priori alignment"
The explanation will help readers' understanding.

Authors’ response:
 We have added an explanation to the sentence as:
“The a priori alignment, i.e. aligning the a priori to a common one, is done to compensate/correct its contribution to the smoothing equation (Rodgers and Connor, 2003).”

Referee:
 (5) Page 16, line 479, “sufficient number”
How many pixels are needed for robust statistics?

Authors’ response:
Our co-location pairs are considered valid only when more than four S5P pixels are remaining. This reduces the random uncertainty by at least a factor of two.

Referee:
C. Technical Corrections

(1) Page 40, Figure 2, Page 48. Figure 10, Page 74 Figure 36
There are too many colors to identify each site.

Authors’ response:
Following the reviewer’s suggestion to shorten the paper, we decided to remove Fig. 17 and Fig. 36 as a similar message can be conveyed by Fig. 18 and Fig. 37 respectively showing examples for a few sites and the monthly distribution of the relative biases.
As per Fig. 2 and Fig. 10, the goal is to present the full range of surface albedo covered by the reference ground-based networks used in our study and the overall view of the bias change and scatter as observed for the two operational S5P products (see corresponding explanation in section 4.1). We have improved the colour scale to make the sites distinguishable from one another.

References
Noël, S., Reuter, M., Buchwitz, M., Borchardt, J., Hilker, M., Bovensmann, H., Burrows, J. P., Di Noia, A., Suto, H., Yoshida, Y., Buschmann, M., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Podolske, J. R., Pollard, D. F., Sha, M. K., Shiomi, K., Sussmann, R., Té, Y., Velazco, V. A., and Warneke, T.: XCO2 retrieval for GOSAT and GOSAT-2 based on the FOCAL algorithm, Atmos. Meas. Tech., 14, 3837–3869, https://doi.org/10.5194/amt-14-3837-2021, 2021.

Parker, R. J., Webb, A., Boesch, H., Somkuti, P., Barrio Guillo, R., Di Noia, A., Kalaitzi, N., Anand, J. S., Bergamaschi, P., Chevallier, F., Palmer, P. I., Feng, L., Deutscher, N. M., Feist, D. G., Griffith, D. W. T., Hase, F., Kivi, R., Morino, I., Notholt, J., Oh, Y.-S., Ohyama, H., Petri, C., Pollard, D. F., Roehl, C., Sha, M. K., Shiomi, K., Strong, K., Sussmann, R., Té, Y., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: A decade of GOSAT Proxy satellite CH4 observations, Earth Syst. Sci. Data, 12, 3383–3412, https://doi.org/10.5194/essd-12-3383-2020, 2020.

Wunch, D., Wennberg, P. O., Osterman, G., Fisher, B., Naylor, B., Roehl, C. M., O'Dell, C., Mandrake, L., Viatte, C., Kiel, M., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Notholt, J., Warneke, T., Petri, C., De Maziere, M., Sha, M. K., Sussmann, R., Rettinger, M., Pollard, D., Robinson, J., Morino, I., Uchino, O., Hase, F., Blumenstock, T., Feist, D. G., Arnold, S. G., Strong, K., Mendonca, J., Kivi, R., Heikkinen, P., Iraci, L., Podolske, J., Hillyard, P. W., Kawakami, S., Dubey, M. K., Parker, H. A., Sepulveda, E., García, O. E., Te, Y., Jeseck, P., Gunson, M. R., Crisp, D., and Eldering, A.: Comparisons of the Orbiting Carbon Observatory-2 (OCO-2) XCO2 measurements with TCCON, Atmos. Meas. Tech., 10, 2209–2238, https://doi.org/10.5194/amt-10-2209-2017, 2017.

O'Dell, C. W., Eldering, A., Wennberg, P. O., Crisp, D., Gunson, M. R., Fisher, B., Frankenberg, C., Kiel, M., Lindqvist, H., Mandrake, L., Merrelli, A., Natraj, V., Nelson, R. R., Osterman, G. B., Payne, V. H., Taylor, T. E., Wunch, D., Drouin, B. J., Oyafuso, F., Chang, A., McDuffie, J., Smyth, M., Baker, D. F., Basu, S., Chevallier, F., Crowell, S. M. R., Feng, L., Palmer, P. I., Dubey, M., García, O. E., Griffith, D. W. T., Hase, F., Iraci, L. T., Kivi, R., Morino, I., Notholt, J., Ohyama, H., Petri, C., Roehl, C. M., Sha, M. K., Strong, K., Sussmann, R., Te, Y., Uchino, O., and Velazco, V. A.: Improved retrievals of carbon dioxide from Orbiting Carbon Observatory-2 with the version 8 ACOS algorithm, Atmos. Meas. Tech., 11, 6539–6576, https://doi.org/10.5194/amt-11-6539-2018, 2018.

Sha, M. K., Langerock, B., De Mazière, M., Dils, B., Feist, D. G., Sussmann, R., Hase, F., Schneider, M., Blumenstock, T., Notholt, J., Warneke, T., Petri, C., Kivi, R., Te, Y., Wennberg, P. O., Wunch, D., Iraci, L. T., Strong, K., Griffith, D. W. T., Deutscher, N. M., Velazco, V. A., Morino, I., Ohyama, H., Uchino, O., Shiomi, K., Goo, T. Y., Pollard, D. F., Borsdorff, T., Hu, H., Hasekamp, O. P., Landgraf, J., Roehl, C. M., Kiel, M., Toon, G., TCCON team, and NDACC team: VDAF CO validation results & ESA AO # 28603 project TCCON4S5P first CO validation results using TCCON and NDACC data, in: S5P First products release workshop, ESA-Esrin, Frascati (Rome), Italy, 25 – 26 June, https://nikal.eventsair.com/QuickEventWebsitePortal/sentinel-5p-first-product-release-workshop/sentinel-5p, 2018.

Lorente, A., Borsdorff, T., Butz, A., Hasekamp, O., aan de Brugh, J., Schneider, A., Wu, L., Hase, F., Kivi, R., Wunch, D., Pollard, D. F., Shiomi, K., Deutscher, N. M., Velazco, V. A., Roehl, C. M., Wennberg, P. O., Warneke, T., and Landgraf, J.: Methane retrieved from TROPOMI: improvement of the data product and validation of the first 2 years of measurements, Atmos. Meas. Tech., 14, 665–684, https://doi.org/10.5194/amt-14-665-2021, 2021.