We thank all reviewers for their constructive comments, which helped to improve the paper. Below, we address all comments point-by-point.

#reviewer2

This paper is describes the relatively straightforward application of the RemoTeC retrieval algorithm to data from the Orbiting Carbon Observatory-2 (OCO-2). RemoTeC was written to retrieve CO_2 and CH_4 column-average concentrations from the Greenhouse Gases Observing Satellite (GOSAT), and has been improved and validated over the years, as described in a number of publications. The authors have applied this mature algorithm to OCO-2 data to retrieve CO_2 (OCO-2 does not have the CH_4 band that GOSAT possesses), and find that after a few slight modifications, the error statistics of OCO-2 retrievals vs. ground truth data compare favorably both to the operational OCO-2 product as well as to RemoTeC retrievals of CO2 from GOSAT.

C-General Comments

The paper is useful in that it shows that the RemoTeC algorithm can be successfully applied to OCO-2, though it is relatively dry and offers few new physical insights into sources of error/bias in the OCO-2 measurements. However, is it worthwhile piece of work, and I recommend publication after making some minor revisions. My only main comment on the paper has to do with the filtering and bias correction, for which the bottom-line recipes are given. Some more information would be welcome. For instance, what other parameters were investigated for bias correction or filtering, such as the 1/(size parameter) variable used in GOSAT bias correction (Guerlet et al, 2013b)? Was the ω_s parameter of Guerlet et al. (2013b) found not to be useful for OCO-2, even though it was for GOSAT? A figure similar to that of Figure 11 in Guerlet et al. (2013b) would be very useful here to see how similar/different GOSAT vs. OCO-2 retrieval biases are. Also, how stringent were your filters overall –did you filter out 10% L2-processed soundings, 50%, etc? How was this different over land and ocean? A throughput map would be useful.

R-The correlations of XCO2 difference with other parameters as used for GOSAT are now shown in Figure S1 in attachment. We tried the potential bias correction parameter (aerosol size, reff) as previously used by GOSAT retrievals. However, as can be seen in the bottom right panel of **Fig.S1** in 'Supporting Information' (SI) there is no clear dependency between this parameter.

For the overall throughput, we add in the paper "The overall L2-processed throughput is around 15%. When estimated separately, the percentages are 15.8%, 14.0% and 16.0% for target, land and ocean soundings, respectively."

C1- P2, L24: "XCO2 retrievals with this level of accuracy [<1%] can provide valuable information on...sources and sinks..." No! 1% = 4 ppm. We know that regional biases even 1 ppm (0.25%) in XCO2 are too large (Chevallier et al, 2014). Please modify or remove this statement.

R1-Modified. The statement now becomes " The XCO2 derived from GOSAT has an accuracy in the order of a few tents of a percent. XCO2 retrievals with this level of accuracy can provide valuable information on the variation of CO2 ."

C2-P3, L2: "by aerosols and cirrus." Do water clouds not have any effect on scattering? I suggest changing this statement to "by aerosols and clouds."

R2-modified.

C3-P3, 1 st paragraph. The authors describe a number of XCO2 retrieval algorithms but this list is certainly not exhaustive. There are the BESD and FOCAL retrievals from M. Reuter, the TanSAT retrieval from D.

Yang, and various versions of the PPDF retrieval of Oshchepkov and Bril. You should either cite these or make clear that you are not exhaustively listing all available retrievals.

R3- Indeed, we are not trying to list all available retrieval algorithms so we use the word 'including'.

C4- P4, L22: "barometric law". Do you not mean the hypsometric equation (which combines the ideal gas law with the hydrostatic equation). They may be equivalent, I'm not entirely sure. But usually in this context, it is referred to as the hypsometric equation.
R4- Modified.

P4, L23: Are your priors adjusted for the secular growth rate of CO2 (since you just say you use CT from 2013)? Seems like you should, or you could probably introduce an artificial trend in your retrievals.
R5- No, for now, the growth rate is not considered in the prior. It is important to note that the retrieval results are hardly affected by the prior, see Figure 1.

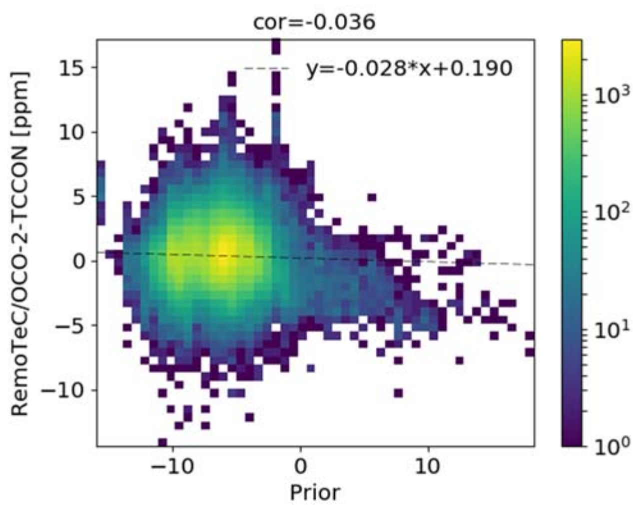

Figure 1. Dependency between XCO_2 difference and prior variation (unit ppm).

P5, L20: Does Sy include any estimate of forward model error, as you previously implied it might (P5, L2). Similar are the noise estimates taken from the OCO-2 suggested formulation, or do you calculate your own noise estimates somehow?
R6-Sy only include noise estimates from the OCO-2 suggested formulation. Modified to be clear.

P6, L9-10: Please discuss whether the per-band radiance offsets were needed for GOSAT. My understanding is that they were needed for band 1, but not the other two bands. You could instead bring this up in section 4.3 as well, but I think it's important to contrast this need for the offsets in OCO-2 vs. that of GOSAT. For instance – I was thinking that maybe you needed them because your retrieval doesn't explicitly retrieve cirrus, so it would be difficult to retrieval soundings with a cirrus layer overlying a thin aerosol layer, which is a pretty common situation. Retrieval the 3 per-band offsets would be a pretty easy way to fake it. But that would likely then also be needed for GOSAT. Some discussion on this would be useful.
R7- For GOSAT the intensity offset fitting is only needed for the O2A band but not for the SWIR bands. This suggests that the offset corrects for an instrumental artefact rather than a retrieval artefact. The state vector differences between OCO-2 and GOSAT are now added in 4.3 section.

P7, L14: χ2 is the symbol usually referred to as the total chi-squared. What you show is much more similar to the "reduced chi-squared", which is the total chi- squared divided by the number of degrees of freedom (# channels - # retrieved parameters). You really are giving the mean chi-squared per channel. You should make this clear, and that a value around unity would indicate a fit that is in line with the noise. Values consistently higher than unity mean there are the systematic errors in the forward model that are not able to be fitted away.
R8- Indeed, χ2 is the "reduced chi-squared ". Modified accordingly in the paper.

P8, top: In the discussion of using the SWIR-1 chi-squared as a bias correction parameter, it would be nice to lengthen this discussion. Does SWIR-2 chi-squared perform similarly? Other parameters? Mention that r=0.2 means that 0.04 percent of the variance is explained (or it will reduce the standard deviation by about 2%). Why do you include the offset "d" parameter when you already include a global bias correction? They would be directly related to each other. Does this multiplicative formula (equation 4) work better than an additive equation? Finally, it would be valuable if you could speculate on why this parameter seems to be correlated with the bias over land. And perhaps on why it is NOT correlated with the bias over water. Are the chi-squared values much lower over water? Finally, what is the spatial distribution of this parameter? Is it highly scattered or does it seem to remove coherent regional biases?
R9- For bias correction, using SWIR-2 chi-squared gives similar results as using SWIR-1 chi-squared when we look at some overall statistics like station-to-station bias, or standard deviation. Using other parameters like those listed in Figure S1, the performance is different and can increase the station-to-station bias. Indeed, the goal of "d" in the bias correction is to correct a global bias. We tried to keep the bias correction purely multiplicative, since the leading scaling term would just link the spectroscopic calibration to the in-situ calibration. The performance is more or less the same with an additive equation.

The reason why the SWIR-1 chi-squared is highly correlated with the bias over land but not with that over ocean is probably related to the fact that high chi2 values over land are often related to bright surfaces. We know that retrieving aerosol over bright land surface is challenging. What we also see here is that SWIR-1 chi-squared is highly correlated (cor is around 0.75) with land surface albedo. However, using albedo directly to do the bias correction can NOT achieve similar performance and will make some statistics worse, for example seasonal variations. So, apart from bright surfaces, the bias correction with chi2 corrects XCO_2 retrievals for cases where the forward model is less capable of fitting the measurements. By doing the correction, as we can see in the paper, we can reduce regional biases since the station-to-staion bias become less. For ocean glint, as we mentioned, aerosols play a less important role and mainly act as an extinction layer. Thus, we cannot see similar feature with land retrievals.

P9, L10: Some comments on why the effect of the bias correction is largest for those 3 stations would be welcome. It seems like it should be substantial for all over-land stations, unless the chi-squared values were just worse for those stations. My guess is that your chi-squared is going to be correlated with SNR or surface albedo, and brighter surfaces will have larger corrections. If you plotted the mean correction on a map, this would probably become obvious.
R10- Indeed, the chi-squared is highly correlated (correlation coefficient is around 0.75) with surface albedo. This could be partly attributed to aerosols since it is difficult to account scattering effects of aerosols over bright surfaces. We added a comment with the reason for the large correction effect for the 3 stations, " This happens due to that the goodness of fit is highly correlated with surface albedo and thus make the corrections apparently to regions with large albedos.".

Section 4.3 As mentioned before, contrast with the offset approach for GOSAT. Is the behavior of the fitted radiance offsets similar over land and ocean? How correlated are the fitted offsets for the 3 bands? (either in absolute terms, or relative to the mean radiance in their respective bands) If they are highly correlated, or not, that would give you a clue what they are correcting for (either cirrus, as I hypothesized earlier, or some instrument effect that is particular to OCO-2, and perhaps not GOSAT).
R11- The fitted radiance offsets are similar in retrievals over land and ocean. The intensity offsets for the 3 bands are moderately correlated with each other (around 0.35).

P11, top: Are the GOSAT vs. OCO-2 error statistics vs. TCCON similar for both land and ocean soundings?
R12- It is difficult to compare OCO-2 ocean retrievals to that of GOSAT because the collocated TCCON sites are quite limited (4 stations).

Figures: in many of the figures, the font sizes make reading some of the text difficult (axis labels, bias numbers , TCCON site names, etc). Please try to make them bigger to increase legibility.
R13-Thanks for the suggestion. We updated this in the revised manuscript.

Technical comments
P4, various: spectral samplings à spectral samples
P5, L10: “radiative transfer model Hasekamp...” à “radiative transfer model (Hasekamp...”
P10, L8: proportional mis-spelled
R14- modified.

Supporting Information for"Carbon dioxide retrieval from OCO-2 satellite observations using the RemoTeC algorithm and validation with TCCON measurements"

Lianghai Wu1, Otto Hasekamp1, Haili Hu1, Jochen Landgraf1, Andre Butz2,3, Joost aan de Brugh1, Ilse Aben1, Dave F. Pollard4, David W. T. Griffith5, Dietrich G. Feist6, Dmitry Koshelev7, Frank Hase8, Geoffrey C. Toon9, Hirofumi Ohyama10, Isamu Morino10, Justus Notholt 11, Kei Shiomi12, Laura Iraci13, Matthias Schneider15, Martine de Maziére 14, Ralf Sussmann15, Rigel Kivi16, Thorsten Warneke11, Tae-Young GOO17, and Yao Té7

[1]SRON Netherlands Institute for Space Research, Utrecht, The Netherlands
[2]Institute of Atmospheric Physics, Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR), Wessling-Oberpfaffenhofen, Germany
[3]Meteorologisches Institut, Ludwig-Maximilians-Universität (LMU), Munich, Germany
4National Institute of Water and Atmospheric Research Ltd (NIWA), Lauder, New Zealand
5University of Wollongong, Wollongong, Australia
6Max Planck Institute for Biogeochemistry, Jena, Germany
7LERMA-IPSL, Sorbonne Universités, UPMC Univ Paris 06, CNRS, Observatoire de Paris, PSL Research University, 75005,
Paris, France
8Karlsruhe Institute of Technology (KIT), IMK-ASF, Karlsruhe, Germany
9Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA
10National Institute for Environmental Studies (NIES), Tsukuba, Japan
11University of Bremen, Bremen, Germany
12Japan Aerospace Exploration Agency, Tsukuba, Japan
13NASA Ames Research Center, Moffett Field, CA, USA
14Royal Belgian Institute for Space Aeronomy, Brussels, Belgium
15Karlsruhe Institute of Technology (KIT), Institute of Meteorology and Climate Research (IMK-IFU), Garmisch Partenkirchen, Germany
16Finnish Meteorological Institute, Sodankylä, Finland
17National Institute of Meteorological Research, Seoul, Republic of Korea

Correspondence to: Lianghai Wu (l.wu@sron.nl)

1 Overview

Here, we provide additional information about: Fig.S1: Error on XCO2 retrievals as a function of six parameters: air mass, water column, blended albedo, mean signal in O2 A-band, aerosol ratio and aerosol size parameter (reff); Fig.S2: same as Fig. S1 but after bias correction; Fig.S3:Validation of individual XCO2 retrieved from OCO-2 measurements after bias correction; Fig.S4-Fig.S6: Validation of overpass averaged retrievals with TCCON before bias correction for targer, land and ocean soundings, respectively; Fig.S7-Fig.S9: The dependence of the bias on latitude before bias correction for targer, land and ocean soundings, respectively.

2 Content of this file

(1). Figures S1 to S9.

(2). Table S1

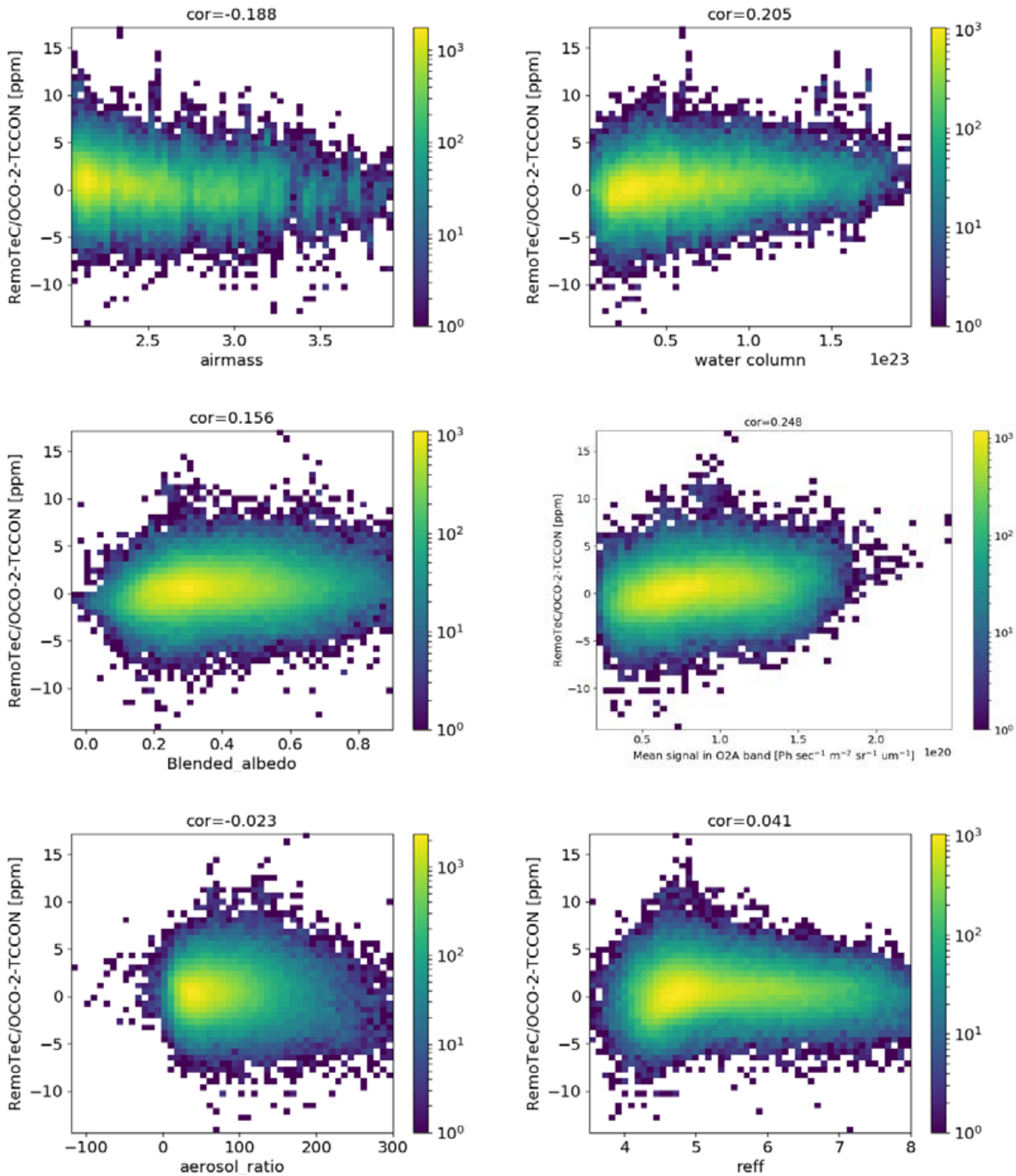

Figure S1. Error on XCO2 retrievals as a function of six parameters: air mass, water column, blended albedo, mean signal in O2 A-band, aerosol ratio and aerosol size parameter (reff). Different colors represent the frequency of point occurrence.

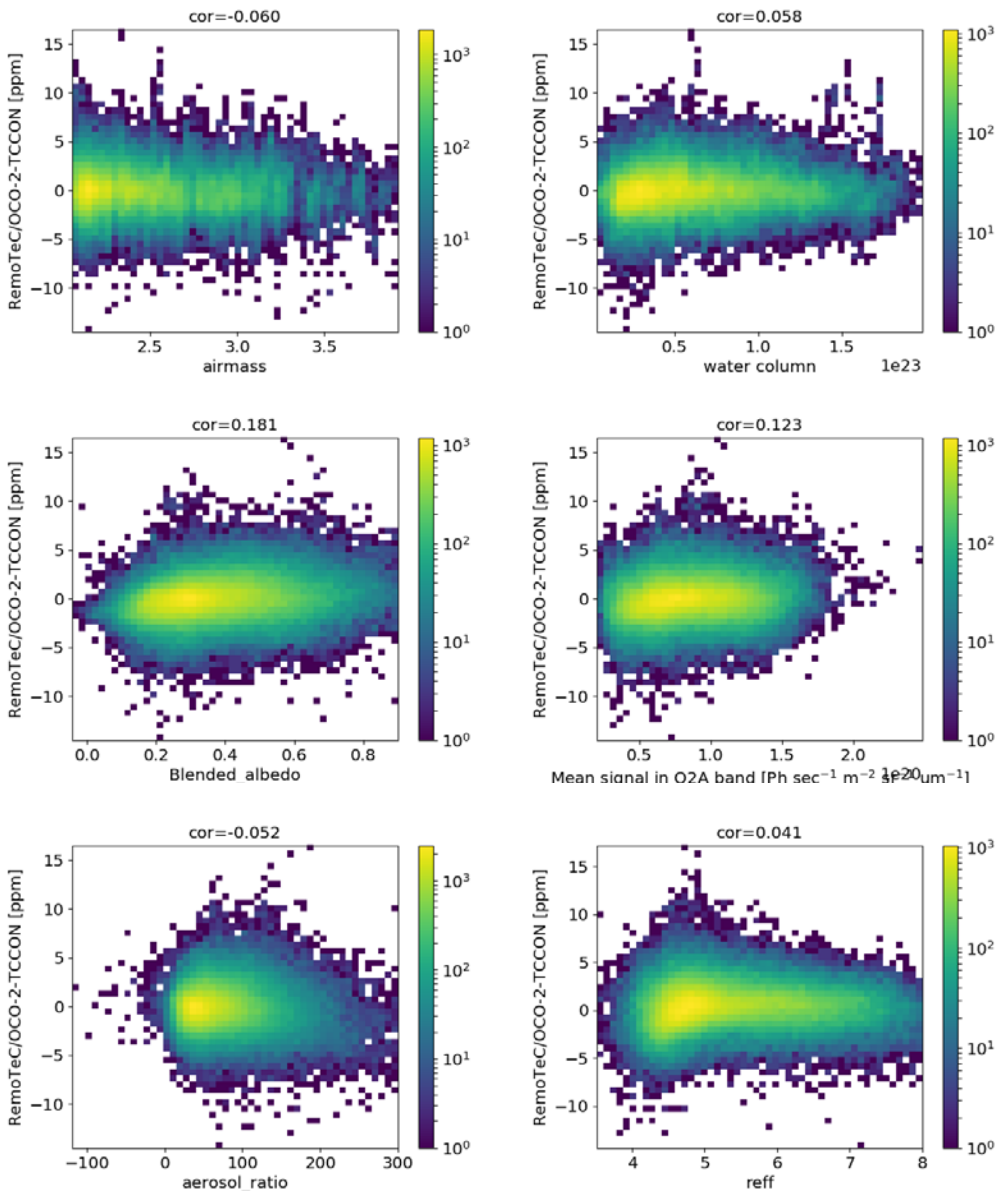

Figure S2.Same as Figure S1 after bias correction.

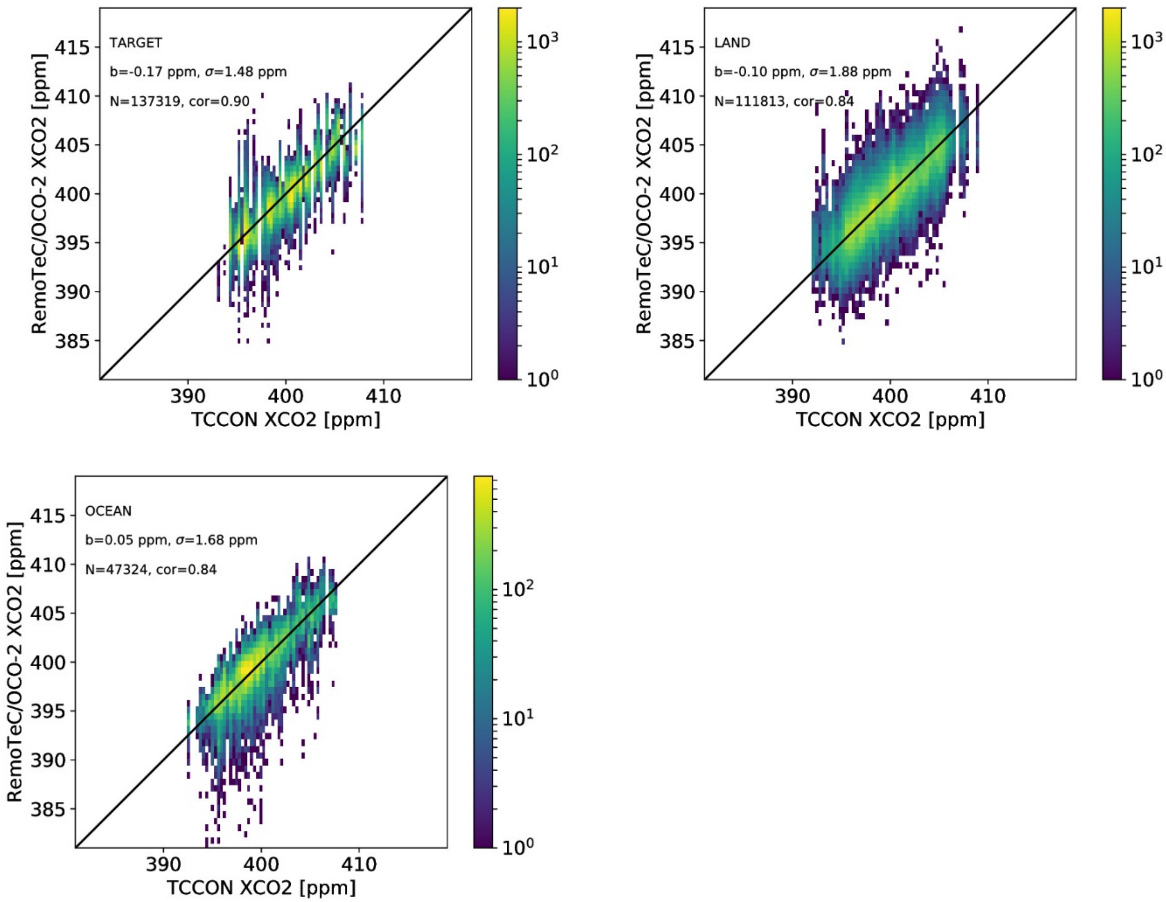

Figure S3. Validation of individual XCO2 retrieved from OCO-2 measurements with collocated TCCON data after bias correction.

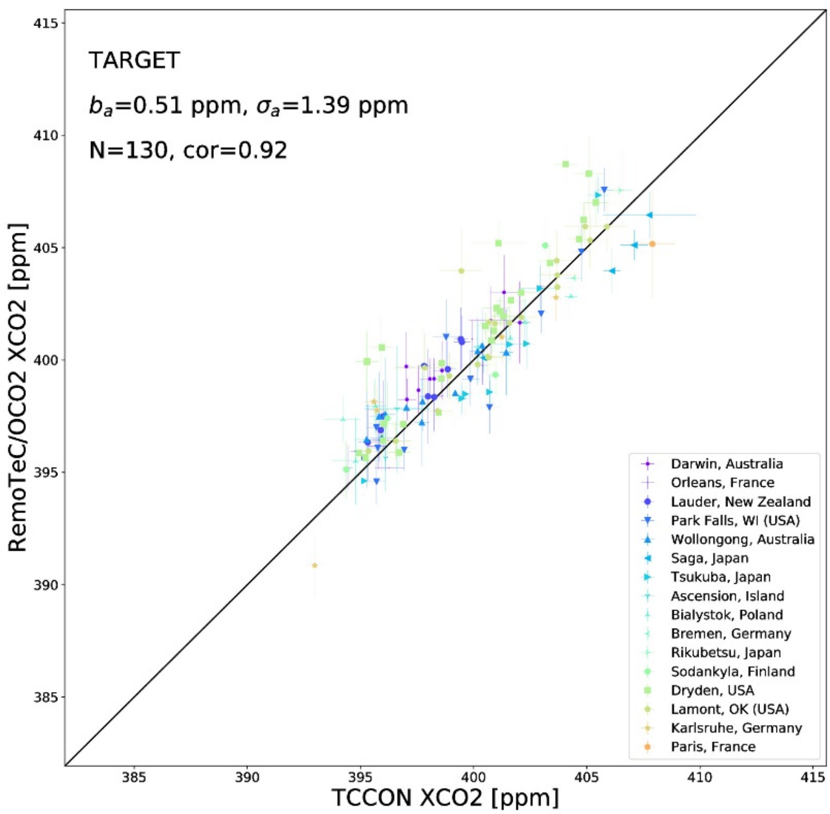

Figure S4. Validation of averaged XCO2 retrieved from OCO-2 target measurements with collocated TCCON data before bias correction. The standard deviation of individual TCCON data and that of RemoTeC/OCO-2 retrievals are presented with error bars. The bias (b_a), standard deviation (σ_a), number of points (N), the Pearson correlation coefficient (cor) and one-to-one line are included.

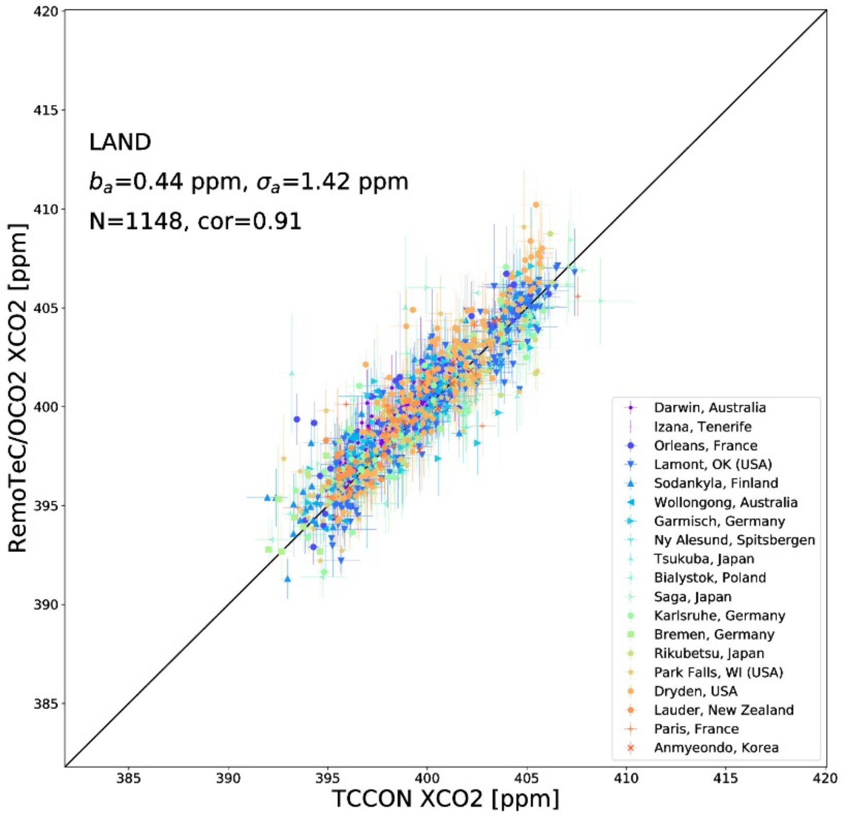

Figure S5. Same as Fig. S4, but for OCO-2 land type measurements obtained under nadir and glint modes.

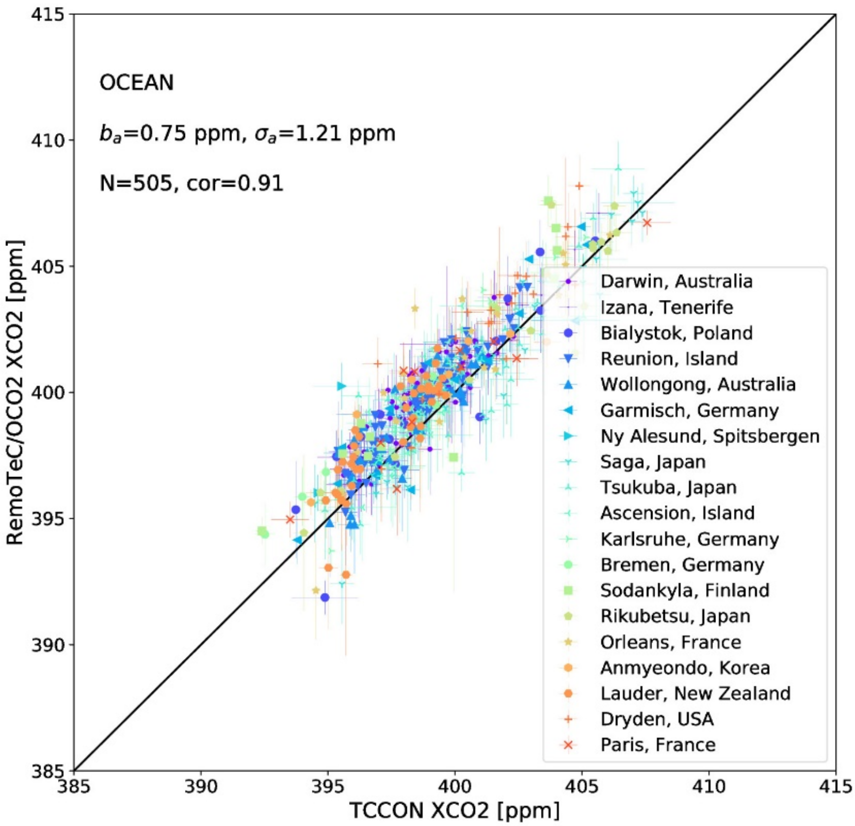

Figure S6. Same as Fig. S4, but for OCO-2 ocean type measurements obtained under glint mode.

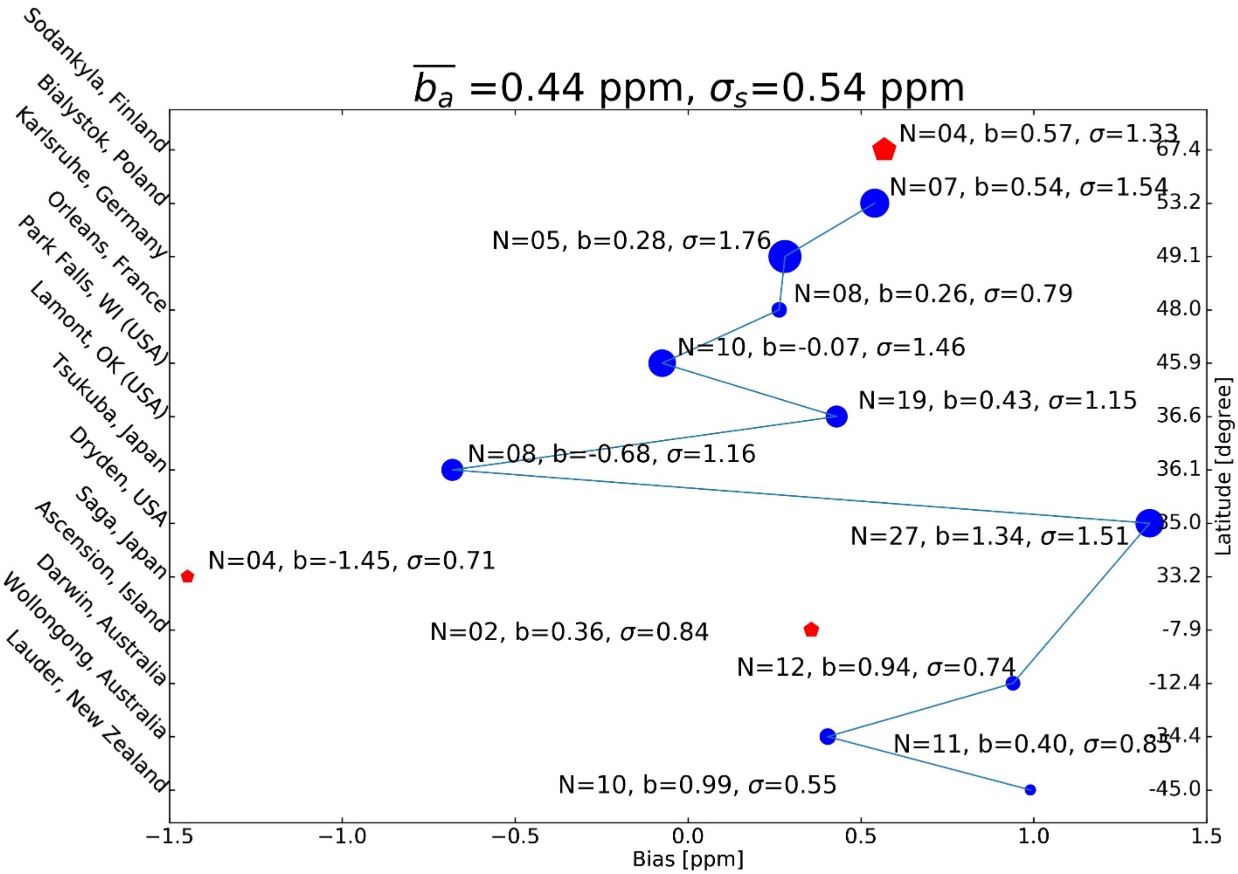

Figure S7. The dependence of the bias between RemoTeC/OCO-2 target XCO_2 retrievals coincident with TCCON data on the latitude of each station. Shown are the averaged results before bias correction. Stations with less than 5 collocation points (marked with red pentagon) should be interpreted with care and are therefore excluded from the calculation of the derived parameters including mean bias (b_a) and the station-to-station variability (σ_s). The size of each dot represents the standard deviation of the difference at each station.

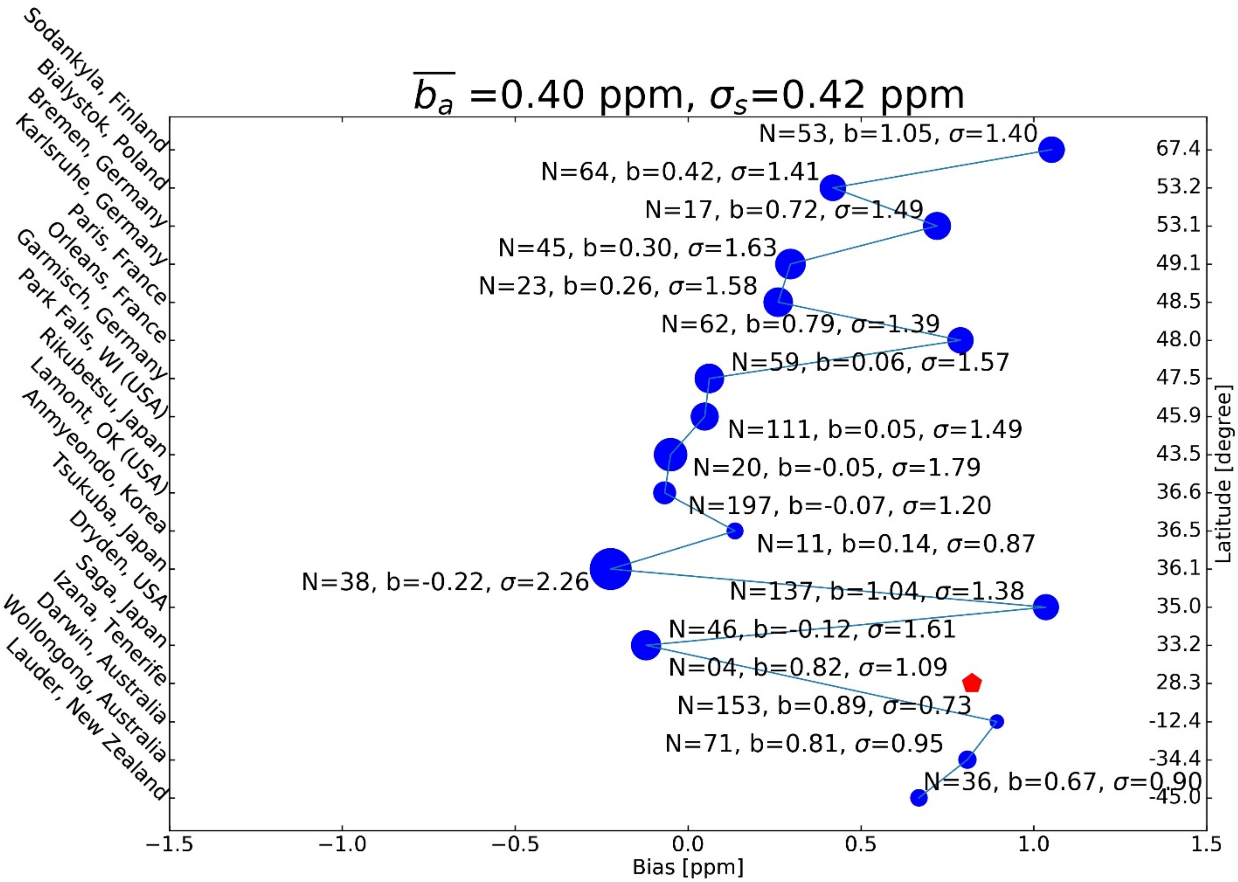

Figure S8. Same as Fig. S7, but for OCO-2 land type measurements obtained under nadir and glint modes.

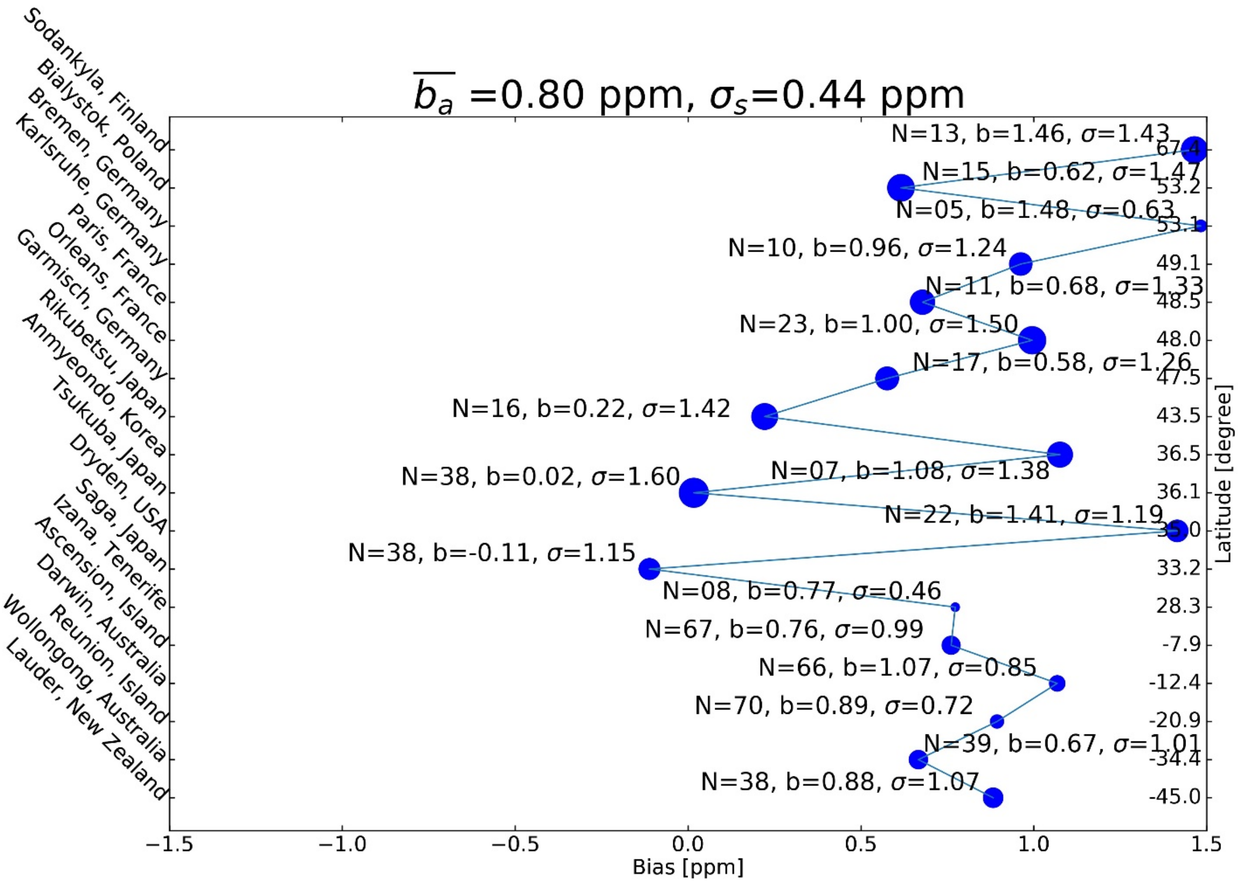

Figure S9. Same as Fig. S7, but for OCO-2 ocean type measurements obtained under glint mode.

Table S1. Correlation coefficients between XCO2 difference between TCCON and OCO-2 retrievals with filtering parameters listed in Table 2 in the paper.

parameters	Correlation before bias correction	Correlation after bias correction
sza	-0.19	-0.05
vza	-0.07	-0.04
χ^2	0.20	0.00
χ^2_{NIR}	0.18	0.09
χ^2_{SWIR1}	0.20	-0.04
χ^2_{SWIR2}	0.15	-0.05
Blended albedo	0.16	0.18
sev	0.07	0.10
α^s	0.04	0.01
$\tau_{0.765}$	-0.05	-0.02
Aerosol ratio parameter	-0.02	-0.05
water column	0.20	0.06
Ioff1	-0.21	-0.15
Ioff2	0.05	0.10
Ioff3	-0.11	-0.09