

Reply to Referee #2

Dear reviewer:

Thank you for reviewing our manuscript and providing constructive comments. Below, the comments are reproduced in black and our point-by-point answers are in blue. Each substantive response is followed by the corresponding revised text and/or figures and tables. Revised passages and figure captions are quoted from the revised manuscript and Supplement. Page and line numbers refer to the revised manuscript; S-numbered figures and tables refer to the revised Supplement.

Sincerely,

Yike Wang

On behalf of all co-authors

Wuhan University

General comment

This manuscript showed the validation results of XCO₂ from the DQ-1 satellite instrument, compared with other satellite XCO₂ and reanalysis data. All analyses look sincerely conducted, results are well organized, and writing quality seems fine. However, there are some questions and comments about suggested results as below. Final decision for this manuscript will be executed after checking authors' response and revised manuscript.

ANSWER: We thank the reviewer for the positive assessment and for the detailed questions. We have revised the manuscript and Supplement substantially, with particular attention to the Abstract motivation, satellite sampling differences, the OCO-3 operational interruption, the official GOSAT-2 bias correction, the regional Australia/tropical-Pacific diagnostics, ocean reflectance and wind-speed relationships, and the same-sample CAMS comparisons requested below.

Specific comments

Comment 1

Readers often read abstract first, sometimes only to see the main lesson of that paper. Thus the completeness of abstract is very important. In that context, the abstract of this manuscript should be improved more. What is DQ-1 ACDL? How does it relate to the IPDA? Why do we need to evaluate the DQ-1 ACDL data? Overall, the research motivation was not well transferred from this abstract.

ANSWER: We have rewritten the Abstract to explain what DQ-1 ACDL is, how it relates to IPDA, and why the active product needs to be evaluated against passive XCO₂ products. The opening now identifies ACDL as a spaceborne IPDA lidar that retrieves XCO₂ using active laser measurements, notes its nighttime/high-latitude sampling capability, and states that active and passive products still differ in sampling and effective column definition. This provides the motivation for the unified comparison framework before the main quantitative results are introduced.

Changes in the manuscript:

Abstract.

Revised text: “Abstract. Da Qi-1 (DQ-1) carries the Aerosol and Carbon Dioxide Lidar (ACDL), a spaceborne integrated path differential absorption (IPDA) lidar that retrieves the column-averaged dry-air mole fraction of carbon dioxide (XCO₂) using active laser measurements. Unlike passive near-infrared/shortwave-infrared sensors that rely on reflected sunlight, ACDL can observe at night and high latitudes, but its sampling and effective column definition differ from passive XCO₂ products. Evaluation is therefore needed before active and passive observations can be interpreted jointly. We compare DQ-1 ACDL with the Orbiting Carbon Observatory (OCO-2 and OCO-3) and Greenhouse gases Observing SATellite (GOSAT and GOSAT-2) missions from June 2022 to December 2024 using a unified framework that combines Copernicus Atmosphere Monitoring Service (CAMS)-based coherence assessment, sampling availability, daily 2° aggregation, spatiotemporal collocation, and correction for XCO₂ column-definition differences. Validation against the Total Carbon Column Observing Network (TCCON) gives $R = 0.92$, a mean bias of 0.21

ppm, and a root-mean-square error (RMSE) of 1.49 ppm for daytime DQ-1, comparable to GOSAT. Over coherent and well-sampled regions, DQ-1 is broadly consistent with OCO-2 and OCO-3, with mean differences generally below 0.5 ppm and similar spatial and meridional structures. Over ocean, the DQ-1–OCO differences vary with surface reflectance and wind speed, with different reflectance relationships for OCO-2 and OCO-3. Nighttime DQ-1 is also broadly consistent with CAMS-derived XCO₂. Regional analyses show complementary DQ-1 coverage where passive observations are limited. These results support DQ-1 ACDL as a quantitatively comparable and complementary component of the satellite XCO₂ observing system.”

Comment 2

In Figure 1, the number of available GOSAT-2 data number looks larger than that of GOSAT-1 data, which is OK because the GOSAT-2 is recent one. Then why the data number of OCO-3 is smaller than that of OCO-2? Additionally, the number of available OCO-3 data decreases over the tropical Pacific, but increases over the 40-50 °N in the North America and Asia. How to explain this? I did not see meaningful authors' statements about this issue.

This seems important in terms of sampling bias factor related to provided validation results.

ANSWER: We have expanded Appendix A1 to explain the spatial and temporal sampling differences. The original Fig. 1 is now Fig. A1, and Fig. A2 shows monthly availability. OCO-3 is carried by the ISS, whose precessing orbit restricts its coverage to approximately 52° S–52° N and concentrates sampling toward the orbital turning latitudes near 50°. This is consistent with the enhanced coverage over North America and Asia at 40–50° N. Its ocean-glint sampling is also more restricted by ISS pointing geometry and operations than that of OCO-2 (Taylor et al., 2023).

The lower total OCO-3 count additionally reflects the 2023–2024 operational interruption discussed in our response to Comment 5. Outside the interruption period, tropical-Pacific OCO-3 sampling remains approximately 51 % of OCO-2 sampling. Thus, the gap alone does not explain the spatial contrast. GOSAT-2 provides more valid

gridded observations than GOSAT in this analysis, but both remain much more sparsely sampled than the OCO products.

We now explicitly acknowledge the consequence for interpretation: same-date, same-grid, time-limited collocation reduces mismatch, but differing coverage still affects the number and representativeness of retained pairs. Here, a “valid observation” is one sensor–day–grid record meeting the sampling-length criterion, not one raw satellite sounding.

Changes in the manuscript:

Appendix A1, p. 31, lines 738–744.

Revised text: “In contrast, the precessing ISS orbit of OCO-3 restricts its coverage to approximately 52° S–52° N and produces a strongly nonuniform latitudinal sampling pattern. Sampling density increases toward the orbital turning latitudes near 50°, consistent with the enhanced OCO-3 sampling over parts of North America and Asia in the 40–50° N latitude band seen in Fig. A1 (Taylor et al., 2023). GOSAT and GOSAT-2 exhibit much sparser point-sampling patterns, with GOSAT-2 providing more valid gridded observations than GOSAT, although both remain substantially sparser than the OCO series.”

Appendix A1, p. 31, lines 746–751.

Revised text: “Compared with OCO-2, OCO-3 provides fewer valid observations over much of the ocean, including the tropical Pacific. This pattern is consistent with the lower sampling rate of OCO-3 in glint mode, which is constrained by the pointing geometry and operational requirements of the ISS platform (Taylor et al., 2023). During normal science operations, clouds, aerosols, solar-illumination conditions, and quality screening can further modulate the number of valid passive retrievals, but these retrieval-related effects are distinct from the sampling patterns imposed by orbital and viewing geometry.”

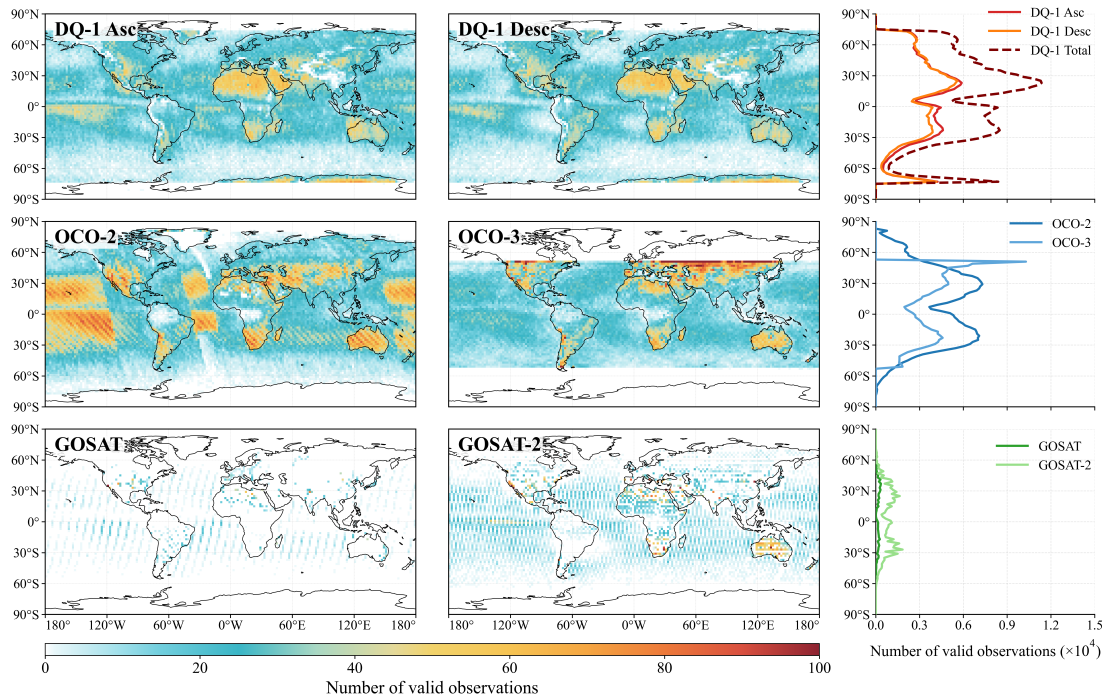


Figure A1. Spatial distributions and zonal totals of cumulative number of valid observations for the satellite XCO₂ products at 2° × 2° resolution.

Comment 3

Overall, GOSAT-2 XCO₂ overestimation looks obvious. What is the reason? I know that the main topic of this manuscript is DQ-1. Once the GOSAT-2 overestimation issue was suggested in this manuscript, however, authors need to explain its reason.

I only found a single statement about this issue in Line 397-398 as next "This systematic positive bias may be related to the fact that the GOSAT-2 SWIR L2 product used here has not been empirically bias corrected", however, I could not figure out the meaning of this statement.

ANSWER: We agree that the original statement was insufficient. We have now applied the official NIES land- and ocean-specific empirical bias corrections to each quality-controlled GOSAT-2 retrieval before daily 2° aggregation and collocation. GOSAT-2 V02.20 is used through February 2024 and V02.21 from March 2024 onward. The subsequent CAMS-based column-definition correction is applied only after collocation

and serves a different purpose: harmonizing the effective XCO₂ column definition rather than removing an empirical product bias.

For the same 880 GOSAT-2–TCCON collocations, the empirical correction changes the mean bias from +3.56 to +0.01 ppm and the RMSE from 3.98 to 1.54 ppm, while R changes only slightly from 0.91 to 0.92. Corrected GOSAT-2 is therefore used in the principal inter-satellite comparisons, while Fig. 3d and Tables S2–S3 retain the original values for transparency. The reduction in bias indicates that a substantial fraction of the original offset is associated with the systematic component represented by the official empirical correction. The correction improves product consistency but does not identify the physical origin of the original retrieval offset.

We also added an important qualification from the NIES release notes: the official correction coefficients were derived using TCCON GGG2020 as the reference. Accordingly, the corrected GOSAT-2–TCCON comparison is not fully independent of TCCON. This dependence is now stated in Sect. 4.2, while the DQ-1–TCCON comparison remains an independent observational validation of DQ-1.

Changes in the manuscript :

Sect. 2.1.2, p. 5, lines 137–147.

Revised text: “For GOSAT-2, the official land- and ocean-specific empirical bias corrections were applied to the individual XCO₂ retrievals following the V02.20 and V02.21 release notes (NIES GOSAT-2 Project, 2026a, b). The official correction coefficients were derived using TCCON GGG2020 as the reference. GOSAT-2 V02.20 was used through February 2024 and V02.21 from March 2024 onward. The empirical bias correction was applied to each quality-controlled individual retrieval before daily 2° aggregation and collocation, whereas the CAMS-based column-definition correction was subsequently applied to the collocated samples. The passive satellite products are provided as individual soundings and provide the XCO₂ observations for comparison with DQ-1. The corresponding quality-control criteria are summarized in Table 2, and the spatiotemporal collocation and column-definition correction are described in Sects. 2.2.2 and 2.2.3. Although all products report XCO₂, their effective column definitions are not identical because the passive products use product-specific prior profiles, pressure weights, and column averaging kernels, whereas

DQ-1 ACDL uses its weighting function and effective pressure layers. These differences are accounted for in Sect. 2.2.3.”

Sect. 3.2, pp. 16, lines 422–428.

Revised text: “For GOSAT-2, the uncorrected XCO₂ product shows a pronounced positive bias relative to TCCON, with $R = 0.91$, a mean bias of 3.56 ppm, and an RMSE of 3.98 ppm (gray symbols in Fig. 3d). Similar positive biases of GOSAT-2 XCO₂ relative to TCCON have also been reported in previous studies (Ji et al., 2024; Yang et al., 2025). After applying the official empirical bias correction to the pixel-level GOSAT-2 XCO₂ data, the mean bias decreases to 0.01 ppm and the RMSE to 1.54 ppm, while R changes only slightly to 0.92 (colored symbols in Fig. 3d). The substantial reductions in bias and RMSE indicate that the pronounced overestimation in the uncorrected product is largely associated with systematic biases addressed by the empirical correction. After correction, GOSAT-2 shows good agreement with TCCON.”

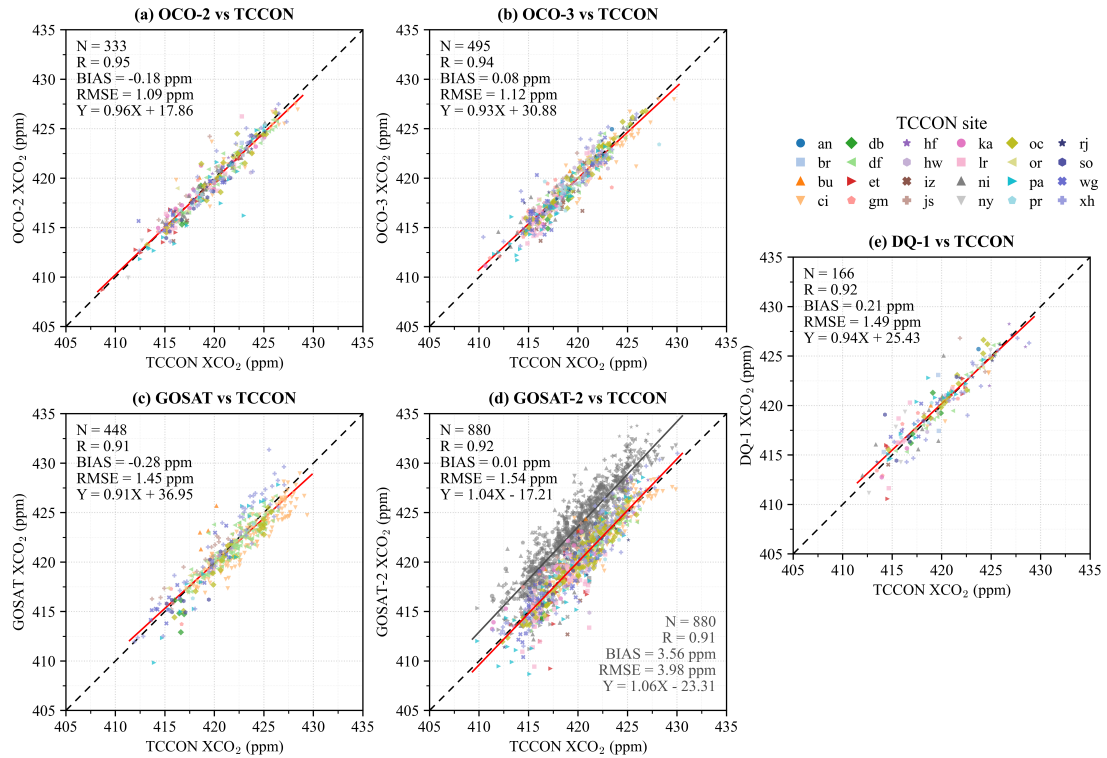


Figure 3. Comparison of satellite XCO₂ products with TCCON observations. (a–e) OCO-2, OCO-3, GOSAT, GOSAT-2, and DQ-1 XCO₂ compared with collocated TCCON XCO₂, respectively. Different colors and symbols denote individual TCCON sites. In panel (d), colored and gray symbols denote empirically bias-corrected and uncorrected GOSAT-2 XCO₂, respectively, with the corresponding linear least-squares fits shown by the red and

gray solid lines. Red solid lines in panels (a–c, e) indicate linear least-squares fits, and black dashed lines indicate the 1:1 relationship. The TCCON site abbreviations are defined in Table A2

Table S2. Satellite–TCCON validation statistics, site-level variability, and 95 % confidence intervals. The 95 % confidence intervals (CIs) of the pooled bias and RMSE are defined by the 2.5th and 97.5th percentiles of 20 000 TCCON-site cluster-bootstrap realizations. Interstation bias is the standard deviation of the site-specific mean biases, and site-mean RMSE is the arithmetic mean of the site-specific RMSE values, with equal weight assigned to each contributing site.

Product	N	Bias (95 % CI) (ppm)	Interstation bias (ppm)	RMSE (95 % CI) (ppm)	Site-mean RMSE (ppm)
OCO-2	333	−0.18 [−0.35, 0.00]	0.60	1.09 [0.89, 1.31]	1.04
OCO-3	495	0.08 [−0.10, 0.28]	0.56	1.12 [0.98, 1.27]	1.12
GOSAT	448	−0.28 [−0.72, 0.60]	1.40	1.45 [1.21, 1.85]	1.60
GOSAT-2 original	880	3.56 [3.24, 3.88]	1.19	3.98 [3.75, 4.27]	4.03
GOSAT-2 corrected	880	0.01 [−0.20, 0.26]	0.97	1.54 [1.33, 1.81]	1.70
DQ-1	166	0.21 [−0.08, 0.48]	0.73	1.49 [1.25, 1.74]	1.40

Comment 4

In Figure 5, the bias looks large over the land of Australia and tropical Pacific. What is the reason to explain this? There was authors' statement in Line 425 "This pattern suggests that the residual differences are more regional and sampling dependent than globally systematic". Again, I could not figure out the meaning of this statement. This statement is consistent with results and discussions in next paragraphs?

ANSWER: We agree that the original sentence stating that the residuals were ‘regional and sampling dependent’ was too vague and inferred more than Fig. 5 alone could establish. We have removed that interpretation. Revised Sect. 3.4.1 now reports the regional statistics first and deliberately postpones physical interpretation to the subsequent CAMS-, reflectance-, and wind-based diagnostics.

Over Australia, the mean DQ-1–OCO-2 and DQ-1–OCO-3 biases are +0.76 and +0.57 ppm, compared with +0.46 and +0.15 ppm over all selected land grid cells. The tropical Pacific shows a different pattern: DQ-1–OCO-2 has a regional mean bias of +0.58 ppm versus +0.51 ppm over all selected ocean grid cells, whereas DQ-1–OCO-3 has +0.30 ppm versus +0.40 ppm. Thus, an enhanced tropical-Pacific mean bias is not common to both OCO comparisons. In contrast, the mean grid-cell RMSE values over the tropical Pacific are 1.27 and 1.26 ppm, both 0.06 ppm above the corresponding all-ocean values of 1.21 and 1.20 ppm (Table S4).

The regional diagnostics are now interpreted separately. Over Australia, DQ-1 and the OCO products show opposite-signed mean departures from the common CAMS-derived reference. Over the tropical Pacific, the same-sample CAMS comparison shows substantially larger DQ-1 scatter (RMSE 1.09/1.08 ppm) than OCO-2/OCO-3 scatter (0.41/0.42 ppm). The reflectance and wind analyses then test surface-condition relationships without treating either factor as a demonstrated cause. This sequence removes the ambiguity of the original sentence and keeps the observation, diagnosis, and limitation distinct.

Changes in the manuscript:

Sect. 3.4.1, p. 20, lines 504–516.

Revised text: “Figure 5 further shows the spatial distributions of the mean DQ-1–OCO-2 and DQ-1–OCO-3 differences, together with the corresponding zonal-mean results. Spatially, most DQ-1–OCO-2 and DQ-1–OCO-3 differences are within approximately ± 1 ppm, with no clear global-scale systematic bias. Local positive differences occur mainly over several oceanic sampling bands, Australia, and some low- and middle-latitude land regions, whereas negative differences are scattered over North Africa, Eurasia, and parts of the tropical oceans. Although most grid-cell differences remain within approximately ± 1 ppm, several regional features are evident (Fig. 5). Over Australia, the mean DQ-1–OCO-2 and DQ-1–OCO-3 biases are +0.76 and +0.57 ppm, respectively, compared with +0.46 and +0.15 ppm over all selected land grid cells. The tropical Pacific shows a different pattern. For DQ-1–OCO-2, the regional mean bias is +0.58 ppm, slightly higher than the all-ocean value of +0.51 ppm, whereas for DQ-1–OCO-3 the regional mean bias is +0.30 ppm, lower than the corresponding all-ocean value of +0.40 ppm. Thus, an enhanced regional mean bias is not common to both tropical-Pacific comparisons. In contrast, the mean grid-cell RMSE values

are 1.27 and 1.26 ppm, respectively, both 0.06 ppm higher than the corresponding all-ocean values of 1.21 and 1.20 ppm (Table S4). These regional differences are examined further in Sects. 3.4.2 and 3.4.3 using CAMS-derived XCO₂, surface reflectance, and ocean wind-speed diagnostics.”

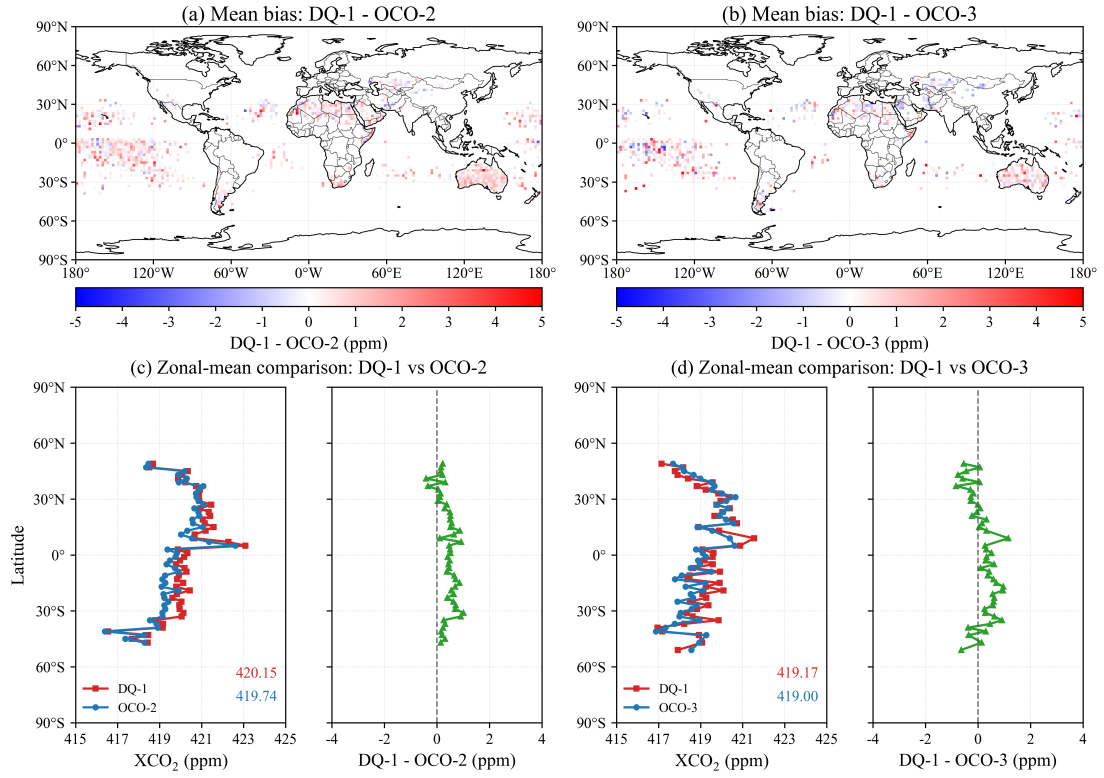


Figure 5. Spatial and zonal differences between DQ-1 and OCO-2/OCO-3 XCO₂ over the selected comparison regions. (a, b) Spatial distributions of the mean DQ-1–OCO-2 and DQ-1–OCO-3 differences, respectively. (c, d) Corresponding zonal-mean XCO₂ values and zonal-mean differences.

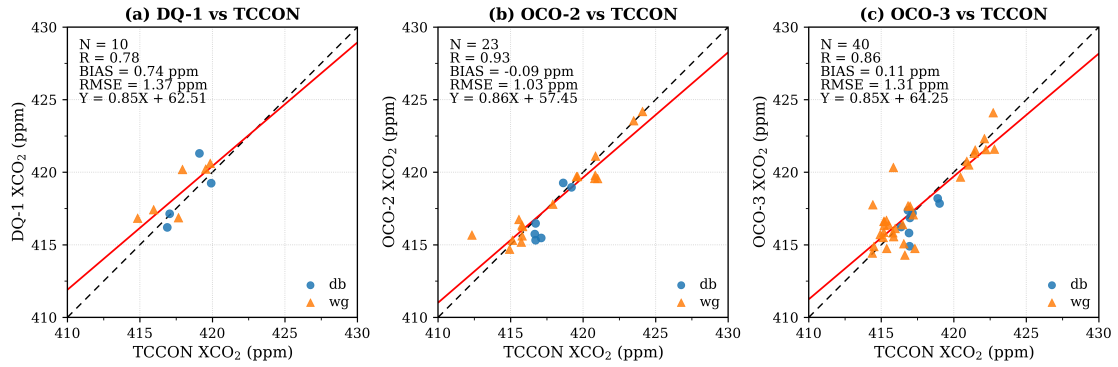


Figure S5. Scatterplot comparisons of DQ-1, OCO-2, and OCO-3 XCO₂ with TCCON observations at the Darwin (db) and Wollongong (wg) sites in Australia. Panels (a–c) show the comparisons for DQ-1, OCO-2, and OCO-3, respectively. The dashed black line denotes the 1:1 line, and the red line shows the linear regression fit.

Table S4. Regional and surface-type grid-cell statistics, MODIS Band 6 reflectance, and ERA5 10 m wind speed for the DQ-1–OCO-2 and DQ-1–OCO-3 comparisons.

Region	Pair	Mean grid-cell Band 6 reflectance	Mean 10 m wind speed (m s ⁻¹)	Mean grid-cell bias (ppm)	Median grid-cell bias (ppm)	Mean grid-cell RMSE (ppm)
Australia	DQ-1–OCO-2	0.3306	—	+0.76	+0.72	1.17
	DQ-1–OCO-3	0.3270	—	+0.57	+0.48	1.04
Tropical Pacific	DQ-1–OCO-2	0.0508	6.48	+0.58	+0.40	1.27
	DQ-1–OCO-3	0.0348	6.73	+0.30	+0.28	1.26
Sahara	DQ-1–OCO-2	0.5834	—	+0.51	+0.47	1.11
	DQ-1–OCO-3	0.5768	—	+0.22	+0.19	0.96
All selected land	DQ-1–OCO-2	0.3970	—	+0.46	+0.46	1.10
	DQ-1–OCO-3	0.4007	—	+0.15	+0.11	1.04
All selected ocean	DQ-1–OCO-2	0.0374	5.93	+0.51	+0.36	1.21
	DQ-1–OCO-3	0.0313	5.96	+0.40	+0.21	1.20

Note: Regional statistics were calculated by equally weighting the contributing $2^\circ \times 2^\circ$ grid cells. MODIS Band 6 reflectance was averaged over grid cells with valid same-day MYD09CMG.061 retrievals. ERA5 10 m wind speed is reported only for the Tropical Pacific and all selected ocean samples. Bias is defined as DQ-1 minus OCO.

Comment 5

Figure 6 (XCO₂ comparison between DQ-1 and OCO-2) looks similar Figure 7 (XCO₂ comparison between DQ-1 and OCO-2). In tropical Pacific, the validation data number in OCO-3 (Fig. 7) is smaller than those in OCO-2 (Fig. 6), probably due to the data number difference (Fig. 1). Based on Figure 2, I can notice that the data number difference between OCO-2 and OCO-3 suddenly becomes larger after August 2023 for a year.

What happen in this period? This malfunction of OCO-3 can make the data difference in tropical Pacific between OCO-2 and OCO-3? (Over Australia, RMSE data points look similar between OCO-2 and OCO-3). Is there any special issue of OCO-3 for only ocean surface monitoring?

ANSWER: The reduced OCO-3 availability in late 2023–mid-2024 reflects a mission-level operational interruption associated with temporary storage on the ISS; it is not evidence of an ocean-specific instrument malfunction. Appendix A1 now separates this interruption from normal-operation retrieval losses and from the persistent ocean-sampling constraints associated with ISS viewing and glint-pointing geometry (Taylor et al., 2023).

We also performed a common-operation test excluding October 2023–July 2024. Across the remaining 21 full months, OCO-2 and OCO-3 provide 17,127 and 8,734 valid tropical-Pacific observations, respectively, so OCO-3 still provides only about 51 % of the OCO-2 sampling outside the interruption. This confirms that the persistent sampling difference is not created solely by the gap.

More importantly, the tropical-Pacific residual statistics are essentially unchanged when the interruption period is excluded: DQ-1–OCO-2 bias/RMSE changes from 0.58/1.27 to 0.61/1.26 ppm, and DQ-1–OCO-3 changes from 0.30/1.26 to 0.30/1.28 ppm. Thus, the interruption reduces OCO-3 sampling availability but is not the main cause of the regional DQ-1–OCO-3 residual characteristics.

Changes in the manuscript:

Appendix A1, p. 32, lines 755–764.

Revised text: “The monthly statistics are shown in Fig. A2. The mean monthly numbers of valid observations for DQ-1 Asc and Desc are 0.77×10^4 and 0.71×10^4 , respectively, and the corresponding value for DQ-1 Total is approximately 1.48×10^4 . This value is higher than the corresponding OCO-2 value of 0.99×10^4 . OCO-3 has a mean monthly number of valid observations of approximately 0.61×10^4 but exhibits substantial month-to-month variability. Reduced sampling is evident in October and early November 2023, followed by a prolonged data gap until mid-July 2024. The gray-shaded region in Fig. A2 marks the OCO-3 data gap associated with its temporary storage on the ISS. This was a mission-level interruption rather than an ocean-specific instrument malfunction. During normal science operations, OCO-3 ocean sampling is additionally constrained by viewing and glint-pointing geometry and ISS operational requirements, while clouds, aerosols, solar-illumination conditions, and quality screening further affect retrieval availability (Taylor et al., 2023). The

prolonged data gap nevertheless contributed to the lower overall number of valid OCO-3 observations relative to OCO-2 over the study period.”

Appendix A1, pp. 32–33, lines 765–776.

Revised text: “To assess whether the reduced tropical-Pacific sampling was primarily caused by this interruption, we repeated the comparison using only common-operation full months, excluding October 2023–July 2024. During the remaining 21 months, OCO-2 and OCO-3 provided 17,127 and 8,734 valid observations, respectively, corresponding to monthly means of 815.6 and 415.9. OCO-3 sampling therefore remained approximately 51 % of OCO-2 sampling outside the interruption period, indicating that the interruption contributed to, but did not fully explain, the lower tropical-Pacific sampling. We further assessed whether the OCO-3 interruption affected the regional DQ-1–OCO comparison over the tropical Pacific by recalculating the bias and RMSE for the same 21 common-operation months. For DQ-1–OCO-2, the bias and RMSE changed from 0.58 and 1.27 ppm for the full period to 0.61 and 1.26 ppm, respectively; for DQ-1–OCO-3, the corresponding values changed from 0.30 and 1.26 ppm to 0.30 and 1.28 ppm. These small changes indicate that the OCO-3 interruption had little influence on the tropical-Pacific bias and RMSE. GOSAT-2 and GOSAT have much lower mean monthly numbers of valid observations, approximately 0.14×10^4 and 0.03×10^4 , respectively. Overall, these results highlight the complementarity of multi-satellite observations in spatial coverage and temporal continuity.”

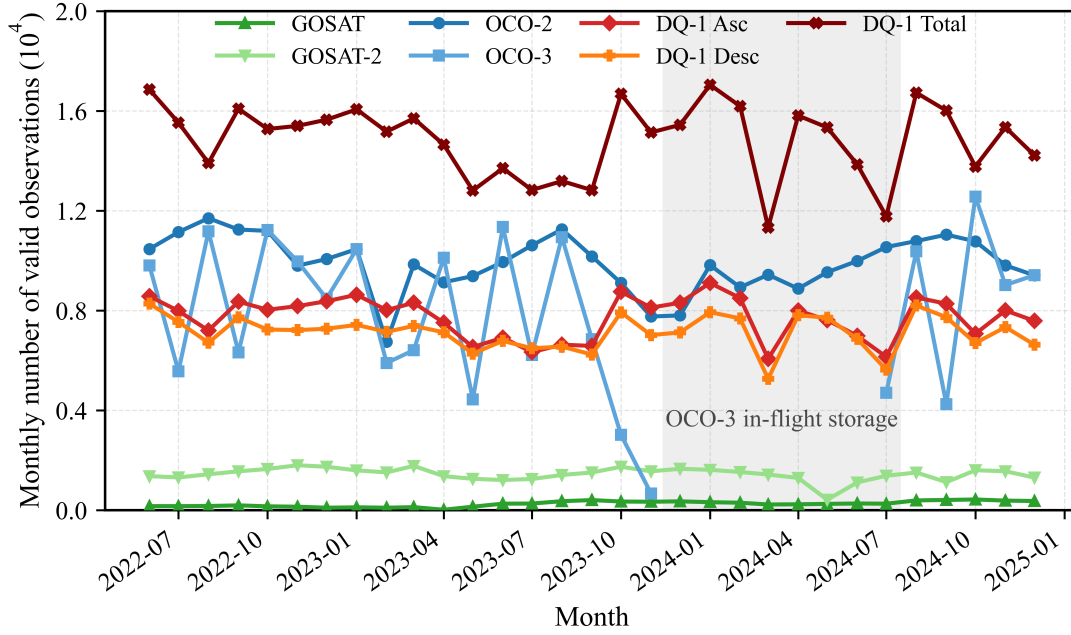


Figure A2. Monthly variations in the number of valid observations for the satellite XCO₂ products. The gray-shaded region indicates the OCO-3 data gap associated with its temporary storage on the ISS.

Comment 6

In Line 447, there is a statement as next "These patterns suggest that the spatial characteristics of DQ-1–OCO-2/OCO-3 bias and RMSE cannot be explained by surface reflectance alone, and may also be related to ocean glint observation conditions and the distribution of valid collocated samples".

How the ocean glint differs? Is there any clue that can be applied to high RMSE in tropical Pacific ocean? Since the bias is especially large over this tropical Pacific (different from other ocean area), authors need to add more detailed statement about this part, at least based on previous study results.

ANSWER: We expanded this discussion and added an explicit sea-surface diagnostic. The revised analysis distinguishes two issues. First, MODIS Band 6 reflectance is a contextual surface indicator rather than the actual OCO glint radiance or sun–sensor geometry. Over ocean, its relationship with the residual statistics is pronounced but pair dependent: DQ-1–OCO-2 bias and RMSE increase with reflectance ($R = +0.77$ and $+0.64$), whereas DQ-1–OCO-3 bias and RMSE decrease ($R = -0.77$ and -0.78). Thus,

reflectance is associated with the oceanic residuals, but not through a common response shared by OCO-2 and OCO-3.

Second, the ERA5 wind-speed diagnostic shows weak correlations at the individual 2° grid-cell level (bias $R = -0.16/-0.29$; RMSE $R = -0.03/-0.08$), but the 1 m s⁻¹ bin means decrease systematically with increasing wind speed (DQ-1-OCO-2: $R = -0.97/-0.81$ for bias/RMSE; DQ-1-OCO-3: $R = -0.94/-0.89$). The strong binned trends together with the large within-bin spread indicate a shift in the conditional mean with wind speed, while substantial grid-cell variability remains. This is why the revised text no longer characterizes the wind relationship simply as ‘weak.’

These diagnostics also argue against explaining the tropical-Pacific RMSE simply by stronger winds: mean wind speeds there are slightly higher than the selected-ocean averages (6.48/6.73 versus 5.93/5.96 m s⁻¹), yet the RMSE is also slightly higher. ERA5 10 m wind is only an indirect roughness proxy and does not reconstruct the instantaneous sun-sensor glint geometry. The same-sample CAMS comparison provides an additional clue: tropical-Pacific RMSE is 1.09/1.08 ppm for DQ-1 versus 0.41/0.42 ppm for OCO-2/OCO-3. We therefore state that the enhanced scatter is more evident on the DQ-1 side under these collocated conditions, without uniquely attributing it to glint, surface roughness, or a specific retrieval mechanism.

Changes in the manuscript :

Sect. 3.4.3, pp. 22–23, lines 546–584.

Revised text: “To further diagnose the influence of surface reflectance on the residuals, we calculate reflectance-binned statistics for the bias and RMSE of DQ-1-OCO-2 and DQ-1-OCO-3. For all samples, the linear relationship between bias and reflectance is weak, with correlation coefficients of approximately $R = -0.22$ for DQ-1-OCO-2 and $R = -0.24$ for DQ-1-OCO-3. This indicates that, when all samples are considered, reflectance is not the dominant factor controlling the mean bias. For samples over land, the DQ-1-OCO-2 bias increases slightly with reflectance, with $R = 0.42$, whereas the DQ-1-OCO-3 bias shows a weaker relationship with reflectance, with $R = -0.21$. For samples over ocean, the correlations are stronger, but the two comparison pairs show opposite trends: the DQ-1-OCO-2 bias increases with reflectance, with $R = 0.77$, whereas the DQ-1-OCO-3 bias decreases with reflectance, with $R = -0.77$. Thus, surface reflectance is clearly associated

with the oceanic bias statistics, but the direction of the association is pair dependent: positive for DQ-1–OCO-2 and negative for DQ-1–OCO-3.

RMSE shows a similar pattern. For all samples, the correlations between RMSE and reflectance are weak for both DQ-1–OCO-2 and DQ-1–OCO-3, with correlation coefficients of approximately $R = -0.24$ and $R = -0.35$, respectively. The correlations remain limited for samples over land, with $R = -0.17$ and $R = -0.42$, respectively. By contrast, the RMSE–reflectance relationships are more pronounced for samples over ocean, but again show opposite trends for the two comparison pairs: the RMSE of DQ-1–OCO-2 over ocean increases with reflectance, with $R = 0.64$, whereas that of DQ-1–OCO-3 decreases with reflectance, with $R = -0.78$. The corresponding RMSE relationships show the same pair dependence. Taken together, the ocean results indicate pronounced but opposite-signed reflectance associations rather than a common reflectance response across the two OCO comparisons.

The ERA5 wind-speed analysis reveals different behavior at the individual-grid-cell and binned levels (Fig. S6). At the individual 2° grid-cell scale, the correlations with wind speed are weak, with $R = -0.16$ and -0.29 for bias and $R = -0.03$ and -0.08 for RMSE for DQ-1–OCO-2 and DQ-1–OCO-3, respectively. After grouping the grid cells into 1 m s^{-1} wind-speed bins, however, both the bin-mean bias and RMSE decrease with increasing wind speed, with $R = -0.97$ and -0.81 for DQ-1–OCO-2 and $R = -0.94$ and -0.89 for DQ-1–OCO-3, respectively. The strong binned relationships, together with the large spread of individual grid cells within the bins, indicate a systematic change in the mean residual statistics with wind speed, while substantial grid-cell variability remains.

The tropical Pacific has slightly higher mean wind speeds than the selected ocean samples as a whole, but its mean grid-cell RMSE is also slightly higher (Table S4). Thus, the elevated tropical-Pacific RMSE cannot be explained simply by stronger near-surface winds. Moreover, ERA5 10 m wind speed is only an indirect indicator of sea-surface roughness and does not describe the instantaneous sun–sensor glint geometry of individual OCO observations. The present diagnostics therefore indicate that ocean-surface conditions are associated with the residual statistics, but they do not identify a single reflectance- or wind-related mechanism that explains the tropical-Pacific differences.”

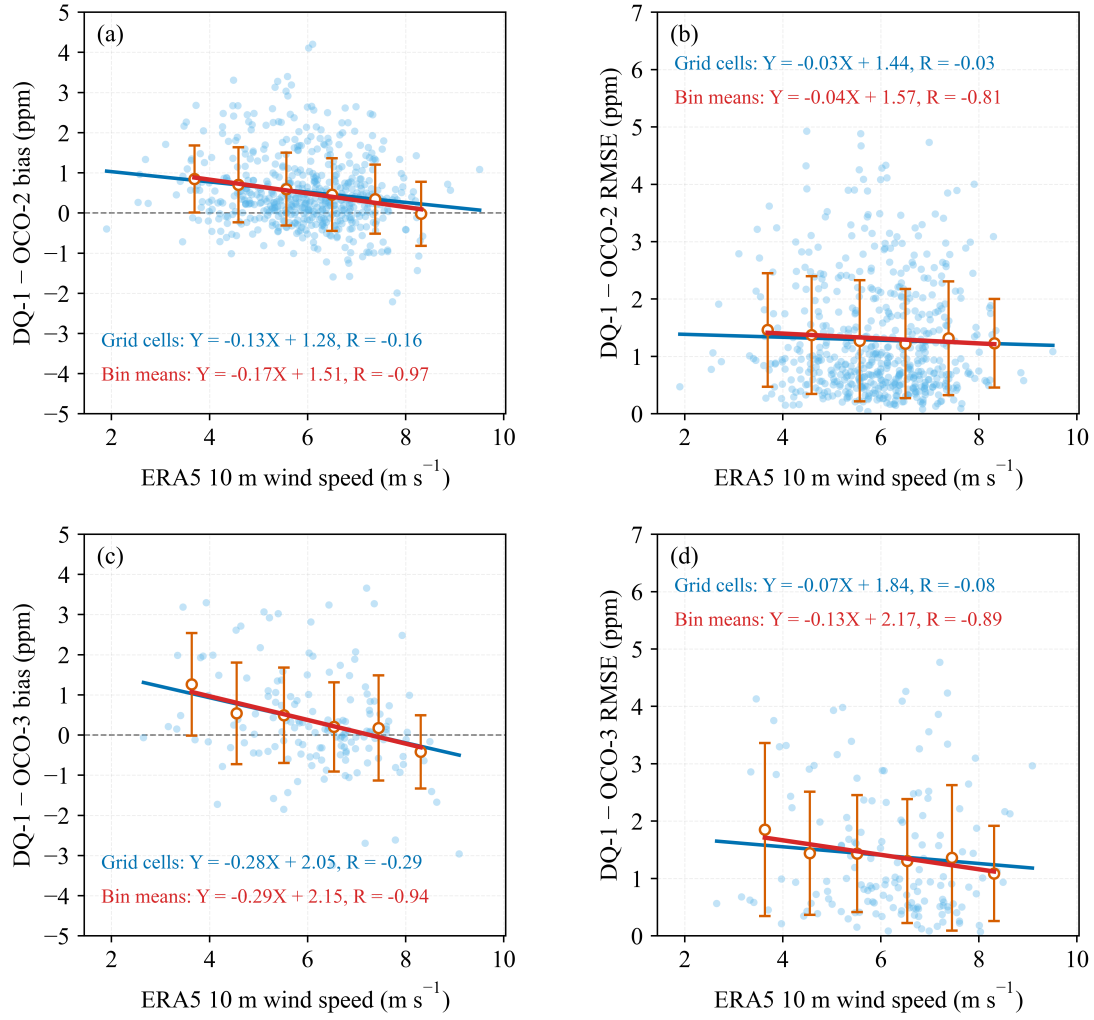


Figure S6. Dependence of oceanic DQ-1-OCO-2 and DQ-1-OCO-3 bias and RMSE on ERA5 10 m wind speed. Panels (a, b) show DQ-1-OCO-2 and panels (c, d) show DQ-1-OCO-3. Light blue dots denote individual oceanic 2° grid cells, while orange circles and error bars show 1 m s^{-1} wind-speed-bin means ± 1 standard deviation. Blue and red lines show linear least-squares fits to the individual grid-cell values and displayed bin means, respectively; the corresponding regression equations and Pearson correlation coefficients are annotated in each panel. Bias is defined as DQ-1 minus OCO.

Comment 7

Please cite figure number appropriately after important statements (especially in chapter 3.4). I cannot make good connection between written statements and suggested figures.

ANSWER: Revised. We have checked the cross-references throughout Sect. 3.4. Figure 5 accompanies the spatial/zonal residual discussion; Figs. S3–S5 and Table S5

accompany the CAMS and regional TCCON diagnostics; Figs. 6–7 accompany the reflectance analysis; and Fig. S6 and Table S4 accompany the wind-speed analysis. This makes each quantitative statement traceable to the corresponding figure or table.

Changes in the manuscript:

Sect. 3.4.1, p. 20, lines 504–516.

Revised text: “Figure 5 further shows the spatial distributions of the mean DQ-1–OCO-2 and DQ-1–OCO-3 differences, together with the corresponding zonal-mean results. Spatially, most DQ-1–OCO-2 and DQ-1–OCO-3 differences are within approximately ± 1 ppm, with no clear global-scale systematic bias. Local positive differences occur mainly over several oceanic sampling bands, Australia, and some low- and middle-latitude land regions, whereas negative differences are scattered over North Africa, Eurasia, and parts of the tropical oceans. Although most grid-cell differences remain within approximately ± 1 ppm, several regional features are evident (Fig. 5). Over Australia, the mean DQ-1–OCO-2 and DQ-1–OCO-3 biases are +0.76 and +0.57 ppm, respectively, compared with +0.46 and +0.15 ppm over all selected land grid cells. The tropical Pacific shows a different pattern. For DQ-1–OCO-2, the regional mean bias is +0.58 ppm, slightly higher than the all-ocean value of +0.51 ppm, whereas for DQ-1–OCO-3 the regional mean bias is +0.30 ppm, lower than the corresponding all-ocean value of +0.40 ppm. Thus, an enhanced regional mean bias is not common to both tropical-Pacific comparisons. In contrast, the mean grid-cell RMSE values are 1.27 and 1.26 ppm, respectively, both 0.06 ppm higher than the corresponding all-ocean values of 1.21 and 1.20 ppm (Table S4). These regional differences are examined further in Sects. 3.4.2 and 3.4.3 using CAMS-derived XCO₂, surface reflectance, and ocean wind-speed diagnostics.”

Comment 8

Figure 8 (DQ-1 vs. CAMS-derived DQ-1) is interesting. The RMSE is also high over Australia, Sahara desert, and tropical Pacific, similar to Figure 6 and 7. Then my question is how about the comparison results between OCO and CAMS-derived XCO₂? If they are closer, then DQ-1 research team can achieve some tips about bias factors.

ANSWER: We added the requested OCO–CAMS comparisons in Figs. S3–S4 and Table S5. DQ-1 and each OCO product are evaluated against their corresponding CAMS-derived XCO₂ on exactly the same DQ-1–OCO collocated sample set, which

controls sample-set differences when comparing the two products with the model reference.

The result is region dependent. Over the tropical Pacific, OCO-2/OCO-3 are substantially closer to CAMS in RMSE (0.41/0.42 ppm) than DQ-1 (1.09/1.08 ppm), so the enhanced scatter is more evident on the DQ-1 side under these collocated conditions. Over Australia, DQ-1 RMSEs (0.63/0.56 ppm) are lower than those of OCO-2/OCO-3 (0.80/0.69 ppm), and DQ-1 and OCO show opposite-signed mean departures from CAMS. Over the Sahara, all corresponding RMSEs are below 1 ppm. Therefore, the analysis does not support a uniform statement that OCO is closer to CAMS everywhere.

These additions are daytime diagnostics. The nighttime DQ-1–CAMS analysis is retained separately because passive NIR/SWIR instruments cannot provide an equivalent nighttime comparison. CAMS is used only as a common diagnostic reference rather than as observational truth, so closer agreement with CAMS does not uniquely assign an error to either satellite retrieval.

Changes in the manuscript:

Sect. 3.4.2, pp. 21–22, lines 530–544.

Revised text: “To further diagnose these regional differences, DQ-1 and OCO-2/OCO-3 were independently compared with CAMS-derived XCO₂ using exactly the same collocated samples (Figs. S3–S4 and Table S5). Over Australia, DQ-1 shows small positive biases of +0.22 and +0.23 ppm relative to CAMS-derived XCO₂ for the OCO-2 and OCO-3 collocations, respectively, whereas OCO-2 and OCO-3 show negative biases of –0.53 and –0.35 ppm. Thus, the positive DQ-1–OCO differences over Australia are accompanied by opposite-signed departures of DQ-1 and the OCO products from the common CAMS-derived reference. A site-based comparison with TCCON at Darwin and Wollongong provides an independent regional check and likewise shows a positive DQ-1 bias, although only 10 DQ-1 collocations are available (Fig. S5). Over the tropical Pacific, the DQ-1–CAMS-derived XCO₂ RMSE values are 1.09 and 1.08 ppm, compared with 0.41 and 0.42 ppm for OCO-2 and OCO-3, respectively. In contrast, over the Sahara, both DQ-1 and the OCO products show RMSE values below 1 ppm relative to CAMS-derived XCO₂, and the contrast between the two products is much smaller than that over the tropical Pacific (Table S5). Over the

tropical Pacific, the substantially larger DQ-1 RMSE relative to the common CAMS-derived reference suggests that the enhanced scatter is more pronounced in DQ-1 under these collocated conditions. This differs from Australia, where DQ-1 and the OCO products show opposite-signed mean departures from CAMS. CAMS is used here only as a common diagnostic reference rather than as observational truth; these comparisons therefore help identify relative product behavior but do not uniquely attribute retrieval errors.”

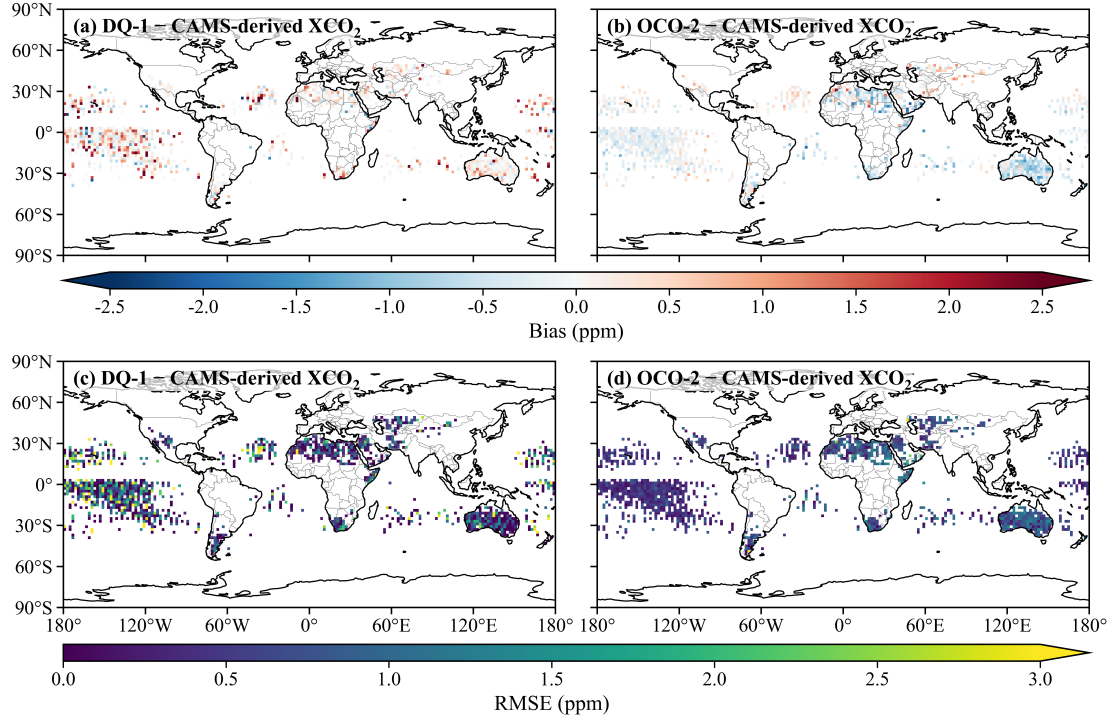


Figure S3. Spatial distributions of daytime bias and RMSE relative to CAMS-derived XCO_2 for DQ-1 and OCO-2, based on the same DQ-1–OCO-2 collocated samples. Panels (a, b) show the biases for DQ-1 and OCO-2, respectively, and panels (c, d) show the corresponding RMSEs. Bias is defined as satellite XCO_2 – CAMS-derived XCO_2 .

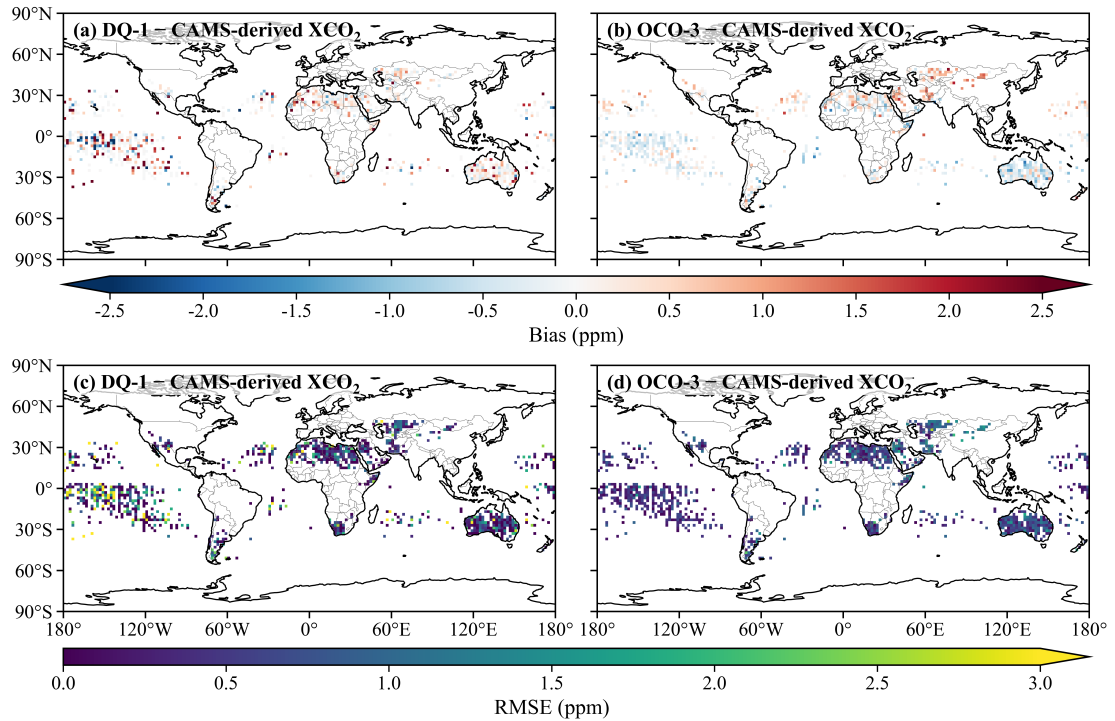


Figure S4. Spatial distributions of daytime bias and RMSE relative to CAMS-derived XCO_2 for DQ-1 and OCO-3, based on the same DQ-1–OCO-3 collocated samples. Panels (a, b) show the biases for DQ-1 and OCO-3, respectively, and panels (c, d) show the corresponding RMSEs. Bias is defined as satellite XCO_2 – CAMS-derived XCO_2 .

Table S5. Regional daytime statistics relative to CAMS-derived XCO₂ for the DQ-1-OCO-2 and DQ-1-OCO-3 comparisons.

Region	Pair	DQ-1-CAMS-derived	OCO-CAMS-derived	DQ-1-CAMS-derived	OCO-CAMS-derived
		XCO ₂	XCO ₂	XCO ₂	XCO ₂
		mean bias (ppm)	mean bias (ppm)	mean RMSE (ppm)	mean RMSE (ppm)
Australia	DQ-1-OCO-2	+0.22	-0.53	0.63	0.80
	DQ-1-OCO-3	+0.23	-0.35	0.56	0.69
Tropical Pacific	DQ-1-OCO-2	+0.37	-0.20	1.09	0.41
	DQ-1-OCO-3	+0.05	-0.25	1.08	0.42
Sahara	DQ-1-OCO-2	+0.16	-0.36	0.51	0.80
	DQ-1-OCO-3	+0.22	+0.00	0.60	0.62
All selected land	DQ-1-OCO-2	+0.16	-0.29	0.58	0.79
	DQ-1-OCO-3	+0.18	+0.03	0.60	0.71
All selected ocean	DQ-1-OCO-2	+0.37	-0.14	1.00	0.45
	DQ-1-OCO-3	+0.29	-0.11	0.96	0.43

Note: Values are equally weighted across contributing $2^\circ \times 2^\circ$ grid cells. CAMS-derived XCO₂ is used as a common diagnostic reference rather than as truth. Bias is defined as satellite XCO₂ minus CAMS-derived XCO₂.

Comment 9

Please add the site name in the caption of Figure 10 (there are many abbreviation in the legend, but I do now know what those are).

ANSWER: The original Fig. 10 is now Fig. 3. Its revised caption explicitly directs readers to Table A2, which gives the full name, abbreviation, coordinates, altitude, and valid observation period for every TCCON site. This makes the abbreviations identifiable without repeating the complete 24-site list in the caption.