

General comments:

This manuscript presents an intercomparison of DQ-1/ACDL XCO₂ with OCO-2, OCO-3, GOSAT, and GOSAT-2. The topic is relevant, and the comparison of active and passive XCO₂ observations is useful. My main concern is that several choices in sample selection and collocation are not sufficiently justified, which makes it difficult to assess the robustness of the reported agreement. I believe the paper makes a valuable contribution and is suitable for publication after revision.

Special comments:

1. The analysis is restricted to the top 10% of grid cells based on the coherence and sampling score. Why was the 90th-percentile threshold chosen? Because this selection favors relatively homogeneous, well-sampled regions, it would be useful to show whether the reported bias and RMSE change when a less restrictive threshold is used.
2. The comparison uses $2^\circ \times 2^\circ$ daily averages with a maximum time difference of 4 h. Please clarify the rationale for these choices, especially the 4 h threshold. A simple test using a smaller spatial or temporal window would help show how sensitive the results are to the collocation criteria.
3. CAMS is involved in the region screening, column-definition correction, and nighttime DQ-1 comparison. How large is the column-definition correction in practice? Showing the comparison statistics before and after this correction would make its influence on the results clearer.
4. GOSAT-2 shows a large positive bias relative to the other products, but the manuscript also notes that this product has not been empirically bias corrected. This makes the direct comparison with bias-corrected OCO products somewhat difficult to interpret. The discussion should clarify this distinction, or include bias-corrected GOSAT-2 results if available.
5. All satellite products are further restricted to 380–480 ppm after the product-specific quality screening. What is the reason for applying this additional range filter? Please also report the number or fraction of observations removed for each product.
6. The TCCON comparison provides an important independent evaluation, but more information on the collocations is needed. Please report the sample size for each satellite and indicate, if possible, whether a small number of stations dominates the pooled statistics. Some estimate of uncertainty for the reported bias and RMSE would also be useful.
7. Please clarify whether the daily grid-cell XCO₂ values are simple arithmetic means or uncertainty-weighted means.
8. The reduced OCO-3 coverage during the stowage/reconfiguration period should be separated from limitations caused by clouds, aerosols, or solar illumination.
9. The title “First comparison” seems somewhat strong, since previous studies cited in the manuscript have already compared or validated DQ-1 against other XCO₂ datasets. The main novelty appears to be the unified multi-satellite comparison framework.
10. Please check the manuscript for typographical errors.