Makoto Saito

Response to Reviewer #1

We thank the reviewer for the careful evaluation of our manuscript and for the positive comments regarding the value of the GOSAT-2 Level 2 and Level 4 products. We appreciate the constructive suggestions, which helped us improve the clarity and completeness of the manuscript. All comments have been carefully considered and addressed as described below.

Reviewer #1

This study by Saito et al. provides a high-level overview and evaluation of the GOSAT-2 XCH4 product and the fluxes inferred from its use. GOSAT-2 is shown to produce significantly more successful observations than GOSAT-1, owing in part to its intelligent pointing. GOSAT-2 and GOSAT-1 XCH4 observations, in aggregate, are consistent, though the GOSAT-1 bias correction appears to drive a latitude-dependent bias between the two instruments. The GOSAT-2 observations are ingested into a previously described inversion system, with the posterior methane fluxes being remarkably different than the prior spatially (South America, Arabian Peninsula) and temporally (Northern Hemisphere seasonality). Generally speaking, the inversion using GOSAT-2 observations is consistent with the inversion using the GOSAT-1 observations, and both differ from the inversion using in situ observations, particularly close to the Equator. This paper will serve as a nice reference for GOSAT-2, and I recommend its publication with minor revisions described below, all at the discretion of the authors.

Minor comments

While reading the methods, I was interested in 1) how the pre- and post-screening works, and 2) the spectral range used for retrievals. Just a sentence on each would be helpful.

Response

We thank the reviewer for this suggestion. We revised Section 2.2.3 to provide a brief description of the pre-screening and post-screening procedures used in the G2L2 retrieval algorithm. The revised text now clarifies that observations are first screened to remove soundings affected by instrumental quality issues and possible cloud or aerosol contamination prior to the retrieval, whereas an additional quality screening based on retrieval diagnostics and convergence criteria is applied after the retrieval. We also expanded the description of the TANSO-FTS-2 instrument and added a summary of the spectral coverage used for the retrieval in Table 1.

Revision

Section 2.2.3 now states:

“Observations are first screened prior to the retrieval to remove soundings with instrumental quality issues or indications of cloud and aerosol contamination. After the retrieval, an additional quality screening based on convergence diagnostics and retrieval quality criteria is applied, and only successful retrievals are included in the final product.”

In addition, Table 1 and Section 2.1.3 summarize the wavelength ranges of the TANSO-FTS-2 spectral bands.

Line 256: you write “repeated soundings over the same location within short time intervals were counted only once per day.” Could you clarify what you mean by this and when repeat soundings would occur?

Response

We thank the reviewer for pointing out that the original description was unclear. We added an explanation that repeated soundings mainly arise from targeted observation operations. In these operations, the same location may be observed multiple times to increase the probability of obtaining successful cloud-free retrievals. To avoid overcounting observations when comparing observation statistics between GOSAT and GOSAT-2, repeated soundings within short time intervals were counted only once per day.

Revision

Additional explanatory text was added in Section 3.1.

For lines 291–293, could you add a sentence explaining how the South Atlantic Anomaly would impact the retrievals, and only for GOSAT-2 (as well as a reference if possible)?

Response

We thank the reviewer for this suggestion. We agree that the reduced data coverage over central-eastern South America deserves clarification. The underlying mechanism has not yet been fully identified. However, reduced radiance levels in TANSO-FTS-2 Band 1 have been observed in this region, and a possible influence of the South Atlantic Anomaly has been suggested within the mission team.

To avoid overstating the current level of understanding, the revised manuscript now explains that the reduced coverage is associated with observed Band 1 radiance degradation and that a possible influence of the South Atlantic Anomaly has been suggested, while the causal mechanism remains under investigation.

## Technical corrections

### Line 17: would water not be the most abundant reactive greenhouse gas?

Response

We thank the reviewer for pointing this out. We agree that the original wording could be misleading because water vapor is the most abundant greenhouse gas in the atmosphere. We revised the sentence to avoid this ambiguity and to better emphasize the role of methane in anthropogenic climate forcing.

### Line 20: “force” to “forcer”

Response

We thank the reviewer for identifying this typographical error. The text has been corrected.

### Line 23: remove “to guide”

Response

We thank the reviewer for the suggestion. The sentence has been revised accordingly.

### Line 42: might Frankenberg et al. (2005) be a better reference here?

Response

We thank the reviewer for the recommendation. Following the reviewer's suggestion, we reexamined the relevant SCIAMACHY literature. We concluded that Frankenberg et al. (2006) is the most appropriate reference for this statement because it focuses on satellite-derived global methane distributions from SCIAMACHY observations. The citation has therefore been revised accordingly.

### Figure 4: maybe a red-white-blue color map would be easier to read for the “diff” plot.

Response

We thank the reviewer for this suggestion. We considered adopting a red–white–blue diverging colormap for the difference panels. However, we retained the original color scale because the current representation allows regions with very small positive or negative differences to remain visible. This helps distinguish locations where both G1L2 and G2L2 observations were available and could be

directly compared from locations without sufficient collocated observations.

We therefore judged that the original color scale better serves the purpose of illustrating both the sign of the differences and the spatial extent of the comparison dataset.

Response to Reviewer #2

We thank the reviewer for the thorough evaluation of our manuscript and for recognizing the potential value of the GOSAT-2 Level 2 and Level 4 $CH_4$ products. We appreciate the constructive comments regarding uncertainty characterization, methodological transparency, interpretation, and data-product documentation.

In response, we substantially revised the manuscript. Major revisions include (1) expansion of the inversion methodology to provide a more self-contained description of the G2L4 framework, (2) clarification of the assumptions underlying the observation and prior error settings, (3) addition of a dedicated discussion on limitations and interpretation of the G2L4 product, including the current absence of posterior uncertainty and information-content diagnostics, (4) clarification of the interpretation of comparisons among G2L4, G1L2-based, and SURF+AIR-based inversions, and (5) addition of Appendix A (Tables A1–A3) describing the structure and contents of the released G2L4 dataset.

Detailed responses to each comment are provided below.

Reviewer #2

General assessment

This manuscript presents a clear and useful description of the GOSAT-2 Level 2 $XCH_4$ product and the new GOSAT-2 Level 4 $CH_4$ flux product. The topic is well suited to ESSD, and the dataset is potentially valuable for the methane-cycle and atmospheric-inversion communities. The manuscript is generally well organized, and I particularly appreciate the comparison among GOSAT-2-only, GOSAT-based, and surface/aircraft-constrained inversions. This comparison is a useful way to evaluate the contribution of GOSAT-2 observations to regional and global $CH_4$ flux estimates.

I recommend publication with major revision. My concerns are primarily about documentation, uncertainty characterization, and interpretation rather than about the fundamental value of the dataset. In particular, because this is an ESSD data paper, the manuscript should provide enough information for readers and future users to understand the effective quality, uncertainty, limitations, and appropriate use of the Level 4 product without relying too heavily on external or forthcoming technical documents. I consider the manuscript suitable for publication after these issues are addressed.

Major comments

Provide stronger uncertainty and information-content characterization of the Level 4 flux product

The manuscript presents posterior $CH_4$ fluxes and posterior-minus-prior differences, but it does not

provide enough quantitative information for users to assess where the G2L4 product is strongly constrained by GOSAT-2 observations and where it remains prior-dominated. This is particularly important because the manuscript itself notes that large posterior adjustments do not necessarily imply strong observational constraint; they may also arise in regions with weak sampling, persistent cloud cover, retrieval bias, or large prior uncertainty.

I ask the authors to add quantitative diagnostics of information content or uncertainty for the G2L4 product. Ideally, this would include posterior uncertainty, posterior/prior uncertainty reduction, degrees of freedom for signal, aggregated averaging kernels, or an equivalent observation-impact (or flux inversion performance) diagnostic. However, I recognize that full-resolution posterior covariance diagnostics may be difficult for a global 1° monthly 4D-Var system. Therefore, regionally aggregated diagnostics, e.g., over the 42 land regions used later in the manuscript and summarized by year or season, would be sufficient and very useful.

If formal uncertainty or DFS diagnostics are not available, the manuscript should state this clearly and provide an alternative diagnostic or a more explicit qualitative description of which regions are likely to be observation-dominated versus prior-dominated.

Response

We thank the reviewer for this important and constructive comment. We agree that quantitative diagnostics describing uncertainty, information content, and observational constraint would be highly valuable for users of the G2L4 product, particularly for interpreting the degree to which posterior flux estimates are constrained by GOSAT-2 observations rather than by prior assumptions.

However, the current operational G2L4 system is based on a high-dimensional global 4D-Var framework and does not compute or archive posterior covariance matrices, inversion averaging kernels, degrees of freedom for signal (DFS), or related information-content diagnostics. Consequently, formal posterior uncertainty estimates, uncertainty-reduction metrics, or observation-impact diagnostics are not available for the current release. Generating and storing such diagnostics for the full operational system would require substantial additional computational resources and is beyond the scope of the present product version.

To address the reviewer's concern, we thoroughly revised the manuscript to clarify these limitations and to prevent potential over-interpretation of the inversion results. Specifically, we added a new subsection (Section 4.3, “Limitations and interpretation of the G2L4 product”), which explicitly states that posterior uncertainty estimates and information-content diagnostics are not currently provided. The new section also clarifies that posterior-minus-prior flux differences should be interpreted as inversion adjustments rather than as direct measures of uncertainty reduction or

observational constraint.

In addition, we revised several statements throughout the manuscript to adopt a more cautious interpretation of posterior flux adjustments. Expressions implying strong observational constraint were softened, particularly in the Abstract, Discussion, and Conclusions. We also added Appendix A (Table A3), which explicitly documents information that is not included in the current G2L4 release, including posterior uncertainties, posterior covariance matrices, inversion averaging kernels, and DFS diagnostics.

We believe these revisions substantially improve the transparency of the dataset and provide users with a clearer understanding of the current capabilities and limitations of the G2L4 product.

Justify the use of a uniform 20 ppb observation prior error covariance

Section 2.3.4 states that each assimilated GOSAT-2 $XCH_4$ retrieval is assigned a uniform 20 ppb observation error. This is a consequential assumption because it controls the relative weighting of observations in the inversion. However, retrieval quality likely varies with local footprint characteristics.

Please provide a clearer justification for the 20 ppb value. Is this intended to represent an average error, a conservative upper bound, or a value inherited from a previous inversion configuration? If possible, please report G2L2 retrieval-error statistics, even in a compact form, such as a histogram or statistics stratified by land/ocean, latitude band, or surface type. If the authors retain the uniform value, please briefly discuss how this assumption may affect regional flux estimates and whether regions with poorer retrieval quality could be over-weighted.

Response

We thank the reviewer for this important comment. We agree that the specification of the observation error covariance is an important component of the inversion system because it governs the relative weighting of the assimilated observations.

The uniform 20 ppb observation error used in this study is not a new assumption introduced specifically for the G2L4 product. Rather, it follows the configuration adopted in the previously developed NISMON 4D-Var inversion framework and was retained to ensure consistency with earlier GOSAT-based methane flux inversions.

The 20 ppb value should be interpreted as an effective observation error used within the inversion system rather than as the retrieval uncertainty of an individual G2L2 sounding. It is intended to account for not only retrieval errors but also representation errors and other sources of

mismatch between observations and the model.

We agree that retrieval quality may vary geographically and among individual observations. However, the current operational G2L4 system does not employ observation-specific error variances and instead applies a uniform observation error treatment globally. To clarify this point, we revised Section 2.3.4 and added discussion regarding the rationale for the 20 ppb value and the limitations of this simplifying assumption.

Revision

Section 2.3.4 was revised to clarify the rationale for the adopted 20 ppb observation error and to explicitly state that the value represents the effective observation uncertainty used by the inversion system rather than the uncertainty of an individual retrieval.

Make the inversion methodology more self-contained

The manuscript refers readers to Niwa et al. and the G2L4 ATBD for several important details of the inversion system. While this is traceable, the present manuscript should include enough methodological detail for readers to evaluate the main assumptions behind the released data product.

Please add a compact, self-contained description of the inversion cost function, control vector, observation term, prior term, and prior covariance construction. In particular, please state the spatial and temporal correlation assumptions or localization scales used in the prior covariance matrix. Section 2.3.4 states that prior covariances for rice paddies, wetlands, and soil oxidation are derived from a 120-year VISIT ensemble with spatial localization, but the manuscript does not state the localization scale or whether month-to-month temporal correlations are imposed. These details are needed for readers to understand the effective resolution and degree of smoothing of the nominal 1° monthly product.

Response

We thank the reviewer for this valuable suggestion. To make the manuscript more self-contained, we substantially expanded Sections 2.3.1-2.3.6. The revised manuscript now includes an explicit description of the Bayesian cost function, observation and prior terms, control variables, prior flux construction, observation-error settings, and optimization framework.

In addition, we clarified the treatment of prior covariance matrices. For rice-paddy, wetland, and soil-oxidation fluxes, prior covariances are derived from a 120-year VISIT ensemble. The manuscript now explicitly states that spatial localization is applied using a Gaussian localization function (Niwa et al., 2022) to suppress unrealistic long-range correlations, and that no explicit

month-to-month temporal correlations are imposed.

We believe these additions substantially improve the transparency of the inversion framework and provide readers with a clearer understanding of the principal smoothing assumptions used in the released G2L4 product.

Address the bias-correction asymmetry between the G2L2 and G1L2 comparisons

A central comparison in the manuscript is between the G2L4 inversion and an analogous inversion using G1L2. However, the G2L2 product used for G2L4 is not bias-corrected, whereas the G1L2 V03.05 product used for comparison is bias-corrected. This asymmetry complicates interpretation of regional differences between the two satellite-based inversions.

This issue is especially important because the manuscript itself shows that the latitudinal structure of G2L2–G1L2 differences is sensitive to the presence or absence of the G1L2 bias correction; the latitudinal trend largely disappears when uncorrected G1L2 V03.00 is used. Therefore, regional differences between the G2L4 and G1L2-based inversions cannot be attributed solely to differences in GOSAT versus GOSAT-2 sampling, coverage, or retrieval characteristics.

Ideally, the authors would repeat the G1L2-based inversion using uncorrected G1L2 V03.00 for a cleaner like-for-like comparison. If this is not feasible, please explicitly qualify the interpretation wherever regional differences between G2L4 and the G1L2-based inversion are discussed. This would help readers understand how to interpret the differences in XCH4 and fluxes.

Response

We thank the reviewer for this important comment. We agree that the comparison between G2L4 and the G1L2-based inversion is affected by the different treatment of bias correction. The G2L2 product used for G2L4 is not bias corrected, whereas the G1L2 V03.05 product used in the comparison includes a bias-correction procedure.

As noted by the reviewer, our comparison of G1L2 V03.05 with the uncorrected G1L2 V03.00 product shows that a substantial part of the latitudinal structure in the G2L2–G1L2 differences can be attributed to the G1L2 bias-correction scheme. We therefore agree that regional differences between the G2L4 and G1L2-based inversions cannot be attributed solely to differences in sampling, coverage, or retrieval characteristics.

Repeating the full G1L2-based inversion using G1L2 V03.00 was not feasible within the scope of the present revision. Instead, we revised the manuscript to explicitly qualify the interpretation of differences between the two satellite-based inversions.

Revision

The discussion of G1L2–G2L2 differences in Section 3.3 was expanded. The manuscript now explicitly states that part of the observed spatial differences likely arises from differences in the bias-correction strategy rather than from the underlying satellite observations themselves.

In addition, Section 4.1 was revised to clarify that differences between the G2L4 and G1L2-based inversions may reflect not only sampling and retrieval characteristics, but also differences in retrieval algorithms and product post-processing, including the G1L2 bias-correction scheme.

These revisions were intended to ensure that readers do not interpret the differences between the two inversion products as being solely attributable to differences between GOSAT and GOSAT-2 observations.

Clarify the design and interpretation of the three inversion experiments

The comparison among G2L4, the G1L2-based inversion, and the SURF+AIR-based inversion is one of the manuscript's strengths. However, the rationale for the three-experiment design should be stated more explicitly. Please clarify what scientific question each comparison is intended to answer. For example: Does G2L4 reproduce the global $CH_4$ budget? Does the choice of satellite product alter regional flux attribution? Where do satellite-only constraints diverge from surface/aircraft constraints?

In addition, please clearly identify any differences among the experiments that could affect interpretation. Assumptions, priors, and constraints of each flux inversion in this manuscript can be put in a table to provide readers a concise way to understand the different flux inversion assumptions. For example, the manuscript notes that some prior components, including biomass-burning emissions, differ between the G2L4 setup and the Niwa et al. SURF+AIR system. If priors differ, then regional discrepancies cannot be attributed cleanly to the observational dataset alone. This caveat should be stated clearly in Section 4.1 and wherever regional G2L4 versus SURF+AIR differences are interpreted. This request will greatly help future readers interpret how the dataset advances the methane flux inversions.

Response

We thank the reviewer for this helpful comment. We agree that the scientific rationale for the three inversion experiments should be stated more explicitly. To address this point, we revised Section 4.1 to clarify the purpose of each comparison. Specifically, the comparison with the G1L2-based

inversion is intended to assess the influence of differences between satellite retrieval products, whereas the comparison with the SURF+AIR-based inversion is intended to identify differences between satellite-only and in-situ observational constraints. The overall comparison also serves as an independent consistency check of the G2L4 product.

We also revised the discussion to clarify that differences among inversions may arise not only from observational constraints but also from differences in prior assumptions, including biomass-burning emissions. Accordingly, regional differences should not be attributed solely to the observational datasets.

Revision

Section 4.1 was revised to clarify the scientific purpose of the three inversion experiments and to provide additional discussion of differences in observational constraints and prior configurations among the compared inversions.

Add normalized coverage or retrieval-success diagnostics

Figures 1 and 2 show raw counts of successful $XCH_4$ retrievals, which are useful. However, raw counts conflate several effects: orbital sampling, targeting strategy, attempted soundings, cloud screening, surface type, and retrieval success. A normalized metric would help users distinguish regions where few observations were attempted from regions where observations were attempted but frequently failed screening.

Please consider adding a figure or panel showing retrieval-success fraction by latitude band or region, ideally normalized by the number of attempted soundings or by the orbit-geometry-limited maximum. If attempted-sounding counts are not readily available, please discuss this limitation and consider adding an alternative normalized coverage metric.

Response

We thank the reviewer for this valuable suggestion. We agree that normalized coverage metrics can be useful for distinguishing the effects of orbital sampling, screening, and retrieval success. However, the G2L2 product provides only observations that successfully passed the pre-screening, retrieval, and post-screening procedures. Information on attempted soundings and the specific screening stage at which an observation was rejected is not distributed as part of the standard G2L2 product and was therefore not available for the present analysis.

In principle, such information may be recoverable from internal processing records, but retrieval-success fractions cannot be consistently derived from the publicly released dataset.

Furthermore, the GOSAT and GOSAT-2 observing strategies are designed to provide nearly uniform sampling along the satellite orbit. Consequently, normalization by the number of attempted observations would mainly affect regions where cloud contamination and screening losses are large, while the broad-scale spatial patterns shown in Figures 1 and 2 would be expected to remain qualitatively similar.

To address the reviewer's concern, we added text to clarify that the observation counts shown in Figures 1 and 2 represent successful retrievals only and therefore reflect the combined effects of sampling strategy, cloud screening, surface conditions, and retrieval performance.

Be more cautious in interpreting regional and sector-dominated flux adjustments

The manuscript includes a useful discussion of posterior adjustments in anthropogenic-source-dominated regions. However, the inversion optimizes net surface $CH_4$ flux and does not directly isolate sector-specific emissions. Therefore, sectoral interpretation should be framed carefully.

Please clarify that the sectoral interpretation is based on prior sectoral dominance or regional source composition, rather than on a sector-resolved inversion. For regional adjustments such as the large negative adjustment in East Asia, please provide posterior uncertainty estimates if available. If formal posterior uncertainties are not available, please state this clearly and frame these adjustments as qualitative or diagnostic rather than as robust sector-specific corrections.

Response

We thank the reviewer for this important comment. We agree that the G2L4 inversion optimizes net surface $CH_4$ fluxes and does not directly resolve individual source sectors. We also agree that interpretations of posterior adjustments in anthropogenic-source-dominated regions should therefore be presented with appropriate caution.

To address this concern, we revised Section 4.2 to adopt more cautious language throughout the discussion of sector-dominated regions. In particular, we changed the section title from “Capability of satellite observations for evaluating sectoral emissions” to “Insights into anthropogenic-dominated regions from satellite observations” in order to avoid implying that the inversion directly quantifies sector-specific emissions.

We also revised several statements that could be interpreted as evidence of strong observational constraints or direct sectoral attribution. Expressions such as “constrain” were replaced with more cautious wording such as “provide useful observational information”, and the discussion was revised to emphasize qualitative interpretation of regional posterior adjustments rather than direct sector-specific corrections.

As noted in response to Major Comment 1, formal posterior uncertainty estimates are not available in the current operational G2L4 framework. This limitation is now explicitly discussed in the newly added Section 4.3. Accordingly, regional posterior adjustments are interpreted as qualitative or diagnostic indications of inferred net flux changes rather than as robust sector-specific emission estimates.

Revision

Section 4.2 was revised to use more cautious language when discussing anthropogenic-source-dominated regions. For example, the text now states: "Figure 9 illustrates that the G2L4 product provides useful observational information for anthropogenic-dominated regions, offering new insight into regional $CH_4$ flux patterns." rather than wording that could imply direct sectoral attribution or strong observational constraint.

Improve data-product documentation and reusability

For reuse, the manuscript dataset should clearly document variable names, units, dimensions, time conventions, missing values, prior and posterior fields, posterior-minus-prior fields, uncertainty fields if available in tabular form added to the appendix. This should be added also added to the Zenodo archive with a README explaining which netcdf files correspond to which dataset and time period. The manuscript should also have recommended citation for the dataset as well.

Response

We thank the reviewer for this helpful suggestion. We agree that clear documentation is essential for ensuring the long-term usability and reproducibility of the G2L4 dataset.

To improve the documentation of the released data product, we added Appendix A describing the structure and contents of the G2L4 $CH_4$ product. The appendix summarizes the principal variables included in the dataset, associated units, dimensions, file conventions, and information that is not currently included in the operational release.

Specifically, we added:

- Table A1, describing the variables contained in the G2L4 product;
- Table A2, summarizing data conventions and product specifications;
- Table A3, documenting information that is not currently provided, including posterior uncertainty estimates, posterior covariance matrices, inversion averaging kernels, and DFS diagnostics.

Regarding external documentation, the G2L4 product is distributed through the official GOSAT-2

project archive together with accompanying documentation, including the Algorithm Theoretical Basis Document (ATBD), release notes, and product format descriptions. These documents provide additional information on the structure and usage of the distributed NetCDF files.

Revision

A new Appendix A was added:

- Table A1: variables included in the G2L4 product;
- Table A2: data conventions and product specifications;
- Table A3: information not included in the current release.

Additional comments

Section 2.3.3: Please clarify how EDGAR v6.0 anthropogenic emissions are handled for 2019–2022. Are anthropogenic prior emissions held fixed at the final EDGAR year, extrapolated, or treated in some other way? This matters because any real anthropogenic emission trend not represented in the prior could appear as a posterior correction.

Response

We thank the reviewer for this important comment. We agree that the treatment of anthropogenic prior emissions beyond the final year available from EDGAR should be clearly documented.

To clarify this point, we revised Section 2.3.3 to explicitly state that anthropogenic emissions from EDGAR v6.0 and VISIT-derived biospheric fluxes were held fixed after their final available release year. Therefore, these prior flux components do not contain interannual variability during the study period. The manuscript now also clarifies that interannual changes in posterior flux estimates primarily reflect information introduced through the assimilated observations rather than variability in these prior flux components.

Please make figure captions more self-contained. For maps and regional comparisons of posterior–prior differences, please specify whether values are annual means, annual totals, monthly means, or flux densities, and state how missing data or masks are handled.

Response

We thank the reviewer for this helpful suggestion. We reviewed all figure captions and revised them where necessary to make the displayed quantities more explicit.

In particular, Figure 5 was revised to clarify that the mapped values represent annual $CH_4$ flux densities for 2021 and their posterior-minus-prior differences, together with the corresponding

units. We also reviewed the remaining figure captions to ensure that averaging periods and quantities are clearly described.

The manuscript has some awkward phrasing. For example, phrases such as “to guide to constrain atmospheric concentrations” and “expanded observational capacities of GOSAT-2 retrieval to its predecessor” should be revised for clarity.

Response

We thank the reviewer for the careful reading of the manuscript. The examples identified by the reviewer were revised in the revised manuscript. In addition, we reviewed the manuscript throughout and revised a number of sentences to improve clarity, readability, and consistency of expression. Particular attention was given to descriptions of observational constraints and flux interpretation, for which more precise and cautious wording was adopted throughout the manuscript.

Revision

Several sentences were revised for clarity and readability, including the examples identified by the reviewer. Additional editorial revisions were made throughout the manuscript to improve consistency and presentation.

Summary

This is a valuable manuscript describing a potentially important GOSAT-2 methane flux data product. I support publication after revision. The main changes needed are to make the inversion assumptions more transparent, better characterize uncertainty and observational constraint, address bias-correction and prior-related confounds in the intercomparison, and improve documentation for future users of the dataset.

Response

We thank the reviewer for the positive assessment of the manuscript and for the detailed and constructive comments. In response, we revised the manuscript to provide a more self-contained description of the inversion framework, clarified limitations related to uncertainty characterization and observational constraints, addressed bias-correction and prior-related caveats in the intercomparison, and substantially improved documentation of the released G2L4 dataset.

We believe these revisions have significantly improved the transparency, interpretability, and usability of both the manuscript and the dataset.