

Manuscript ID: egusphere-2026-3382

Manuscript: Evaluation of the GOSAT/GOSAT-2 XCO₂ proxy XCH₄ product by comparison with TCCON observations over China

Dear Reviewer,

Thank you very much for your constructive comments and suggestions. We have carefully revised our manuscript point by point following your comments, which have substantially contributed to strengthening the manuscript. We expect that the revised manuscript meets the publication requirements of ACP. Below, the reviewers' original comments are shown in blue, our responses are presented in black, and the corresponding revisions made in the manuscript are highlighted in red. All page and line numbers cited in our responses refer to the clean version of the revised manuscript.

Response to comments from reviewer 1

General Comments

The manuscript should provide GOSAT and GOSAT-2 difference in sampling pattern, number of successful retrievals in more detail.

Response:

We agree that the differences in the sampling strategies of GOSAT and GOSAT-2 and their effects on the availability of successful retrievals should be described more clearly. We have therefore added a comparison of their sampling characteristics in Sect. 2.1. We also added published quantitative information on the effect of the GOSAT-2 intelligent pointing function on clear-sky retrievals.

Revision (Sect. 2.1):

“GOSAT/TANSO-FTS performs discrete grid observations and revisits the same nominal observation points every 3 days, with a three-point cross-track sampling pattern used for routine observations since August 2010 (Japan Aerospace Exploration Agency et al., 2011). In contrast, GOSAT-2/TANSO-FTS-2 has a 6-day revisit cycle and employs a fully programmable pointing system, allowing more flexible spatial sampling. Its intelligent pointing function can redirect the line of sight toward nearby clearer areas when clouds are detected around the scheduled footprint, increasing the number of clear-sky retrievals over land by approximately 1.8 times relative to standard pointing.” (Page 4, Line 140 ~ 145)

In addition, the authors should check the observation condition, particularly the use of intelligent pointing function of GOSAT-2. Usually, intelligent pointing is not applied allied to TCCON validation sites in order to repeatedly target the same footprint.

Response:

We have further clarified the observation conditions of GOSAT-2 at validation sites. As described by Suto et al. (2021), intelligent pointing is switched off over calibration/validation (Cal/Val) sites and user-specified sites in current GOSAT-2 operations to enable repeated observations of the same target locations. Therefore, the

description of intelligent pointing in Sect. 2.1 has been revised to distinguish its general capability from the observation strategy applied at validation sites.

Revision (Sect. 2.1):

“However, intelligent pointing is switched off over calibration/validation (Cal/Val) sites and user-specified sites in current GOSAT-2 operations to enable repeated observations of the same target locations.” (Page 4, Line 145 ~147)

Specific Comments

Comment 1.1:

Abstract, lines 22 and 23

Description of “more accurate” and “better performance” are unclear.

Response:

We have revised the description by explicitly referring to the statistical metrics used in this study, including bias, RMSE, correlation coefficient, and the percentage of retrievals meeting the GCOS requirements.

Revision (Abstract):

“The comparison shows lower bias and RMSE of GOSAT products at Xianghe than at Hefei, while GOSAT-2 products are similar at both sites, with lower bias, RMSE, and standard deviation compared with the GOSAT products and approximately 51% of retrievals meeting GCOS requirements.” (Page 1, Line 27 ~ 30)

Comment 1.2:

Abstract, Lin 26

“due to bias in raw XCH₄” is not clear.

Response:

We have revised this sentence to clarify that the overestimation of the XCH₄/XCO₂ ratio is mainly associated with systematic positive biases in the satellite-retrieved XCH₄ products used in the proxy calculation.

Revision (Abstract):

“Furthermore, the overestimation of the ratio implies that the fundamental assumption of the XCO₂ proxy method is not fulfilled, primarily due to systematic positive biases in the satellite-retrieved XCH₄ used in the proxy calculation.” (Page 1, Line 32 ~ 33)

Comment 1.3:

Page 4 lines 130 and 132

“The spectral resolution in the SWIR of approximately 0.2 cm⁻¹ is better than that of TANSO-FTS, is not accurate. The value is the spectral sampling interval. The spectral resolution and optical throughput of GOSAT and GOSAT-2 are approximately the same.

Signal-to-noise ratio of GOSAT-2 is slightly better.

The mission requirements for GOSAT launched in 2009 and GOSAT-2 launched in 2018 are defined differently. The authors should refer to Suto, et al, AMT 2014 for GOSAT-2 rather than Kuze et al., 2009.

Response:

We have revised this statement accordingly and clarified that the spectral resolution (expressed as full width at half maximum, FWHM) and optical throughput of TANSO-FTS-2 are comparable to those of TANSO-FTS onboard GOSAT. The value of approximately 0.2 cm^{-1} represents the spectral sampling interval rather than the spectral resolution.

In addition, we agree that the improvement in the signal-to-noise ratio of GOSAT-2 relative to GOSAT should not be overstated. The corresponding description has therefore been revised to indicate that the signal-to-noise ratio is only slightly improved.

Furthermore, the references related to GOSAT-2 instrument characteristics have been updated by replacing the inappropriate reference to Kuze et al. (2009) with more relevant references, including Suto et al. (2021) and Imasu et al. (2023).

Revision (Sect. 2.1):

“GOSAT-2 was developed to continue and enhance the global greenhouse gas observation capability of GOSAT, with improvements in sensor design and observation efficiency and a slightly better signal-to-noise ratio (Suto et al., 2021; Imasu et al., 2023). Its core payload is TANSO-FTS-2, with an expanded spectral coverage of five bands, adding the capability to detect carbon monoxide (CO). The SWIR bands have a spectral sampling interval of 0.2 cm^{-1} , with FWHM values better than 0.4 cm^{-1} for Band 1 and 0.27 cm^{-1} for Bands 2 and 3. The mission was designed with target uncertainties of 0.5 ppm for CO₂ and 5 ppb for CH₄ over a 500 km spatial mesh and with a 1-month observational period (Imasu et al., 2023; Eoportal, 2024).” (Page 4, Line 133 ~ 139)

Comment 1.4:

Page 6, Line 205

OCO-2 does not have the CH₄ absorption band. Is the term “OCO-2 Proxy” correct?

Response:

Thank you for pointing this out. We agree that OCO-2 itself does not contain a CH₄ absorption band. However, the term “OCO-2 Proxy” refers to the heritage of the University of Leicester retrieval framework rather than to the satellite measurement used for the CH₄ retrieval. According to the C3S documentation, OCPR is defined as the “OCO-2 Proxy (PR) algorithm (used by Univ. Leicester).” The same documentation further states that the University of Leicester retrieval system is based on the original Orbiting Carbon Observatory (OCO) Full-Physics retrieval algorithm, which was subsequently modified for application to GOSAT spectra. The CH₄ OCPR product is then obtained by applying the proxy technique to this Full-Physics retrieval framework. Therefore, although the algorithm name contains “OCO-2,” the CH₄ product evaluated here is retrieved from GOSAT spectra rather than from OCO-2 observations. We have retained the term “The OCPR (OCO-2 Proxy) algorithm”

and supplemented the description of the retrieval principle in Sect. 2.3.1 to clarify the origin and application of the algorithm.

Revision (Sect. 2.3.1):

“The algorithm is based on the University of Leicester retrieval framework originally developed for the Orbiting Carbon Observatory (OCO) and OCO-2 missions and subsequently adapted to process GOSAT SWIR spectra as part of the Copernicus Climate Change Service (C3S) greenhouse gas suite (Boesch et al., 2024; Boesch and Di Noia, 2023).” (Page 4, Line 217 ~ 220)

Comment 1.5:

Page 13, Table 3

Observation period of GOSAT is much longer than that of GOSAT-2. Does the table 3 data use the same period for both satellites?

In addition, the geolocation matching criteria for example the maximum allowable distance between the satellite footprint center and the TCCON site should be described in more detail.

Response 1:

Regarding the observation period, GOSAT and GOSAT-2 have different operational lifetimes, and therefore the available observation periods of their retrieval products are not identical. The temporal coverage of the four Proxy XCH₄ products used in this study is provided in Table 2. The statistics presented in Table 3 were calculated using all available retrievals within the corresponding product periods after applying the official quality screening criteria. We have recalculated the statistical metrics for GOSAT SRPR and GOSAT-2 SRPR using only the overlapping time periods at the Hefei and Xianghe sites. The new results have been added as Figure 6 and Table 4. In addition, the related discussion in the manuscript has been revised to clarify that the comparison between GOSAT SRPR and GOSAT-2 SRPR is based on the same time range. The revised text now reports the updated bias, RMSE, SD, correlation coefficient, and GCOS compliance statistics for both products. We have also added an overall summary of the performance of the three GOSAT products (OCPR, SRPR, and FOCAL), highlighting their different performance characteristics across the evaluated metrics and the two sites.

Revision 1 (Sect. 3.1):

“Overall, the performance of the three GOSAT products is comparable overall, although their performance characteristics vary across different metrics and between the two sites, with no single product consistently outperforming the others at both sites. OCPR shows relatively high correlations but a pronounced positive bias at Hefei, while FOCAL has the highest proportion of data meeting the GCOS requirement at both sites but relatively low correlations. SRPR generally shows intermediate performance among the three GOSAT products across the evaluated metrics.” (Page 12, Line 386 ~ 390)

“To further assess the performance differences between GOSAT and GOSAT-2 based on the same SRPR product, Fig. 6 and Table 4 compares GOSAT SRPR with GOSAT-2 SRPR over the common time range at each site. The performance of the GOSAT-2 SRPR is better than GOSAT SRPR, with the lower RMSE (9.7–10.6 ppb), lower SD (9.0–10.3 ppb) and the higher correlation ($R \approx 0.9$) at both sites. In Hefei, 75.6% of the data points fall

within the GCOS target; however, a systematic overestimation is observed at Hefei, with 20.7% of the data pairs above the GCOS requirements and a positive bias of + 3.7 ppb. At the Xianghe site 69.9% of the data points meet the GCOS requirements, but 20.6% fall below this range, resulting in a negative bias of - 2.5 ppb.” (Page 13, Line 402 ~ 408)

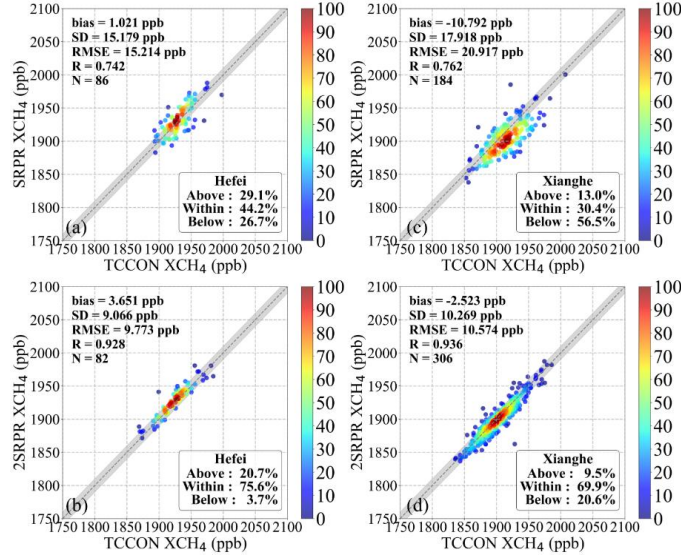


Figure 6. As Fig. 5, but for GOSAT SRPR and GOSAT-2 SRPR over their common observation period at the Hefei (a) (b) and Xianghe (c) (d) sites.

Table 4. As Table 3, but for GOSAT SRPR and GOSAT-2 SRPR over their common observation period.

Product \ Metric	GOSAT SRPR		GOSAT-2 SRPR	
	Hefei	Xianghe	Hefei	Xianghe
Bias (ppb)	1.021	-10.792	2.823	-2.337
RMSE (ppb)	15.214	20.917	8.501	9.553
R	0.742	0.762	0.918	0.933
SD (ppb)	15.179	17.918	8.019	9.263
N	86	184	67	185
Above (%)	29.1	13.0	19.4	8.6
GCOS: 10ppb	Within (%)	44.2	77.6	71.4
	Below (%)	26.7	3.0	20.0

Response 2:

Regarding the geolocation matching criteria, we have clarified the geolocation matching criterion in the revised manuscript. Satellite observations were selected based on the center coordinates of the satellite footprints, requiring both the latitude and longitude differences from the TCCON site to be within $\pm 2.5^\circ$. This corresponds to a $5^\circ \times 5^\circ$ latitude–longitude window rather than a fixed radial-distance threshold.

Revision 2 (Sect. 3, revised manuscript):

“GOSAT and GOSAT-2 observations were selected when the satellite footprint centers were within $\pm 2.5^\circ$ in both latitude and longitude of each TCCON station, and for each overpass a spatial mean was calculated for all valid pixels within the matching area” (Page 10, Line 325 ~ 327)

Technical Corrections

(1) Reference

Add Suto et al., Atmospheric Measurement Techniques 14, 2013–2039, 2021

Response 1:

We have added the reference of Suto et al. (2021) to the revised manuscript and cited it in the relevant discussion of the GOSAT-2 instrument and observation strategy.

Revision (Reference):

Suto, H., Kataoka, F., Kikuchi, N., Knuteson, R. O., Butz, A., Haun, M., Buijs, H., Shiomi, K., Imai, H., and Kuze, A.: Thermal and near-infrared sensor for carbon observation Fourier transform spectrometer-2 (TANSO-FTS-2) on the Greenhouse gases Observing SATellite-2 (GOSAT-2) during its first year in orbit, *Atmospheric Measurement Techniques*, 14, 2013–2039, 2021.

Response to comments from reviewer 2

Comment 2.1:

Page 1, Line 31

“impact on the Earth’climate (Hansen and Sato, 2001)” should be “impact on the Earth’s climate (Hansen and Sato, 2001)”.

Response:

We have corrected “Earth’climate” to “Earth’s climate” in the revised manuscript. (Page 1, Line 38)

Comment 2.2:

Page 4, Line 148–150

The authors should cite the appropriate updated TCCON data version documentation and include information on the accuracy of TCCON measurements.

Response:

TCCON released the updated GGG2020.1 data version in June 2026. Following this update, we updated the TCCON observations used in this study to the latest GGG2020.1 release. We have also revised the manuscript to specify the data release and to provide updated quantitative information on TCCON measurement

performance. The single-measurement precision is approximately 0.15% (~0.6 ppm) for XCO₂ and 0.2% (~4 ppb) for XCH₄ (Toon et al., 2009). In addition, for the GGG2020 retrieval framework, the estimated error budgets are below 0.16% (~0.6 ppm) for XCO₂ and below 0.4% (~7 ppb) for XCH₄ for solar zenith angles below 82° (Laughner et al., 2024). The corresponding references have been added to the revised manuscript. The Data Availability section has also been updated accordingly to provide the GGG2020.1 data version and the corresponding dataset DOIs for the Hefei and Xianghe sites.

Revision (Sect. 2.2):

“Ground-based FTS observations achieve exceptionally high precision; the single measurement precision for XCO₂ is approximately 0.15% (~0.6 ppm) and for XCH₄ it is 0.2% (~4 ppb) (Toon et al., 2009). For the GGG2020 retrieval framework underlying the current data release, the estimated error budgets are below 0.16% (~0.6 ppm) for XCO₂ and below 0.4% (~7 ppb) for XCH₄ for solar zenith angles below 82° (Laughner et al., 2024). These high-precision observational data serve as a reference for the validation of satellite-retrieved greenhouse gas products such as CO₂ and CH₄. The TCCON observations used in this study were obtained from the GGG2020.1 data release, a minor update to the GGG2020 product released in June 2026 (TCCON, 2026)” (Page 4, Line 153 ~ 159)

Revision (Data availability):

“The TCCON data are available at <https://tccondata.org/>. The specific datasets are available via their respective DOIs: <https://doi.org/10.14291/tccon.ggg2020p1.hefei01.R1> (Hefei) and <https://doi.org/10.14291/tccon.ggg2020p1.xianghe01.R0> (Xianghe).” (Page 29, Line 764 ~ 766)

Comment 2.3:

Page 9, Line 286

“the frequency of high-AOD observations”: Fig 3 shows only few and disparate periods with AOD observations at both site and only a very short overlapping period in 2023. Hence it is awkward to speak about “frequency” of observations. Rather, it seems that the AOD at Hefei is higher during most of the available measurements. In addition, the AOD has been reported to decrease over China. Why would tracking failures occur at during high AOD conditions at Xianghe and not at Hefei?

Response:

The expression “frequency of high-AOD observations” has been revised because the AOD records at the two sites cover only limited and largely non-overlapping periods. The revised manuscript therefore describes only the observed feature that AOD at Hefei is generally higher than at Xianghe during most of the available measurement periods.

Regarding the lower apparent occurrence of very high AOD at Xianghe, Eck et al. (2018) investigated several AERONET sites in Northeast Asia, including Xianghe, and demonstrated that high fine-mode AOD events can be affected by AERONET cloud screening. Specifically, for Xianghe, four of five days with daily mean fine-mode AOD at 500 nm above 2.0 were removed by Version 2 Level 2 cloud screening. Their multiyear

comparison further showed that, for Xianghe days with AOD > 2, 46% were screened in Version 2 and 17% in Version 3. These results indicate that some very high-AOD events at Xianghe may be underrepresented in quality-assured AERONET data because of cloud screening associated with extensive cloud cover. Therefore, the previous explanation based on sun-photometer tracking failures has been replaced by this cloud-screening effect.

Revision (Sect. 2.4):

“However, during most of the available measurement periods shown in Fig. 3, AOD values at the Xianghe site are generally lower than those at the Hefei site. At Xianghe, high-AOD pollution events are often associated with extensive cloud cover, which can reduce the availability of quality-assured AERONET observations. Consequently, some very high-AOD observations may be excluded by AERONET cloud screening, leading to an underrepresentation of these events in the available data (Eck et al., 2018).” (Page 9, Line 292 ~ 296)

Comment 2.4:

Page 14, Line 419-429

The manuscript analyzes seasonal variations of product biases. However, seasonal comparisons may be affected by uneven availability of observations among seasons and thus influence the derived seasonal statistics. Was this accounted for by the authors.

Response:

We have explicitly accounted for the uneven seasonal availability of observations in the revised seasonal analysis. The number of valid satellite–TCCON matchups (n) for each product and season has now been added to Fig. 8, allowing the sampling differences among seasons to be directly assessed. The revised discussion also incorporates the effect of sample availability on the derived statistics. In particular, at the Hefei site, the number of observations in summer is relatively small for all four products (n = 12–32), and the larger error bars observed in some summer XCH₄ bins, especially at high concentrations, are partly associated with this limited sample size. This sampling information has therefore been incorporated directly into the interpretation of the seasonal bias variations rather than treating the seasonal statistics independently of data availability.

In addition, the GOSAT-2 SRPR dataset has been updated and extended through the end of 2025, and the TCCON observations have also been updated to the latest data release. The available matchup samples and their seasonal distributions have been recalculated accordingly.

Revision (Sect. 3.1):

“The number of valid matchups (n) for each product and season is also shown in the figure. The sample availability varies among seasons, particularly at the Hefei site, where the number of observations in summer is relatively small for all four products (n = 12 ~ 32). The data range varies across seasons, reflecting the higher concentrations in summer and autumn than in winter and spring; notably, the mentioned large variations occur during summer.” (Page 14, Line 436 ~ 439)

“The relatively small summer sample size may partly contribute to the larger variability represented by these error bars.” (Page 15, Line 443 ~ 444)

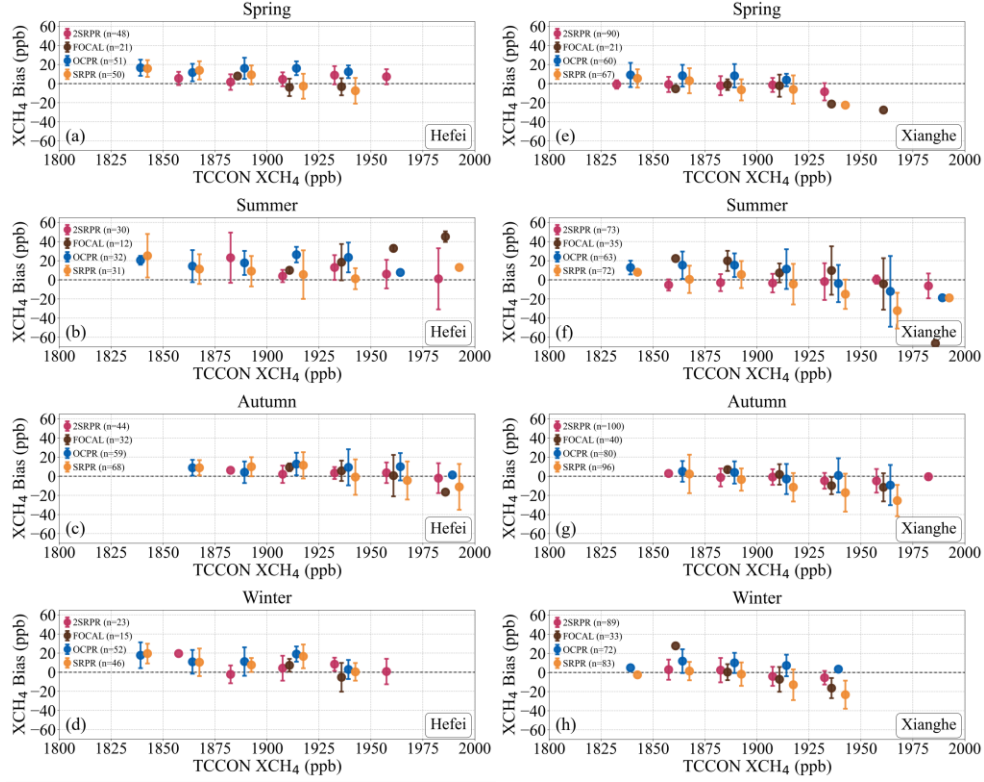


Figure 8. As Fig. 7, but broken down by season, for the Hefei (a-d) and Xianghe (e-h) sites. In spring the data are averaged over March, April and May (a, e), in summer over June, July and August (b, f), in autumn over September, October and November (c, g) and in winter over December, January and February (d, h). The values of n in the legends denote the number of valid Satellite – TCCON matchups used for each seasonal analysis.

Comment 2.5:

Page 18, Line 524

The terms “raw XCH_4 ” and “raw XCO_2 ” require clarification. The authors should define these variables and explain their roles in the proxy XCH_4 retrieval process, as they are used to interpret the observed biases in proxy products.

Response:

We have clarified the definitions and roles of $XCH_{4, \text{raw}}$ and $XCO_{2, \text{raw}}$ in the revised manuscript. Following Eq. (1), these variables are now defined as the directly retrieved column-averaged dry-air mole fractions of CH_4 and CO_2 , respectively, which are used to construct the $XCH_{4, \text{raw}} / XCO_{2, \text{raw}}$ proxy ratio. In the subsequent bias analysis, we also explicitly relate these quantities to the terms “raw XCH_4 ” and “raw XCO_2 ” used in the text and clarify that they are analyzed separately to investigate the origin of biases in the proxy ratio and the final Proxy XCH_4 product. The corresponding descriptions have been added to both the retrieval-method section and the results discussion.

Revision (Sect. 2.3):

$$XCH_4(\text{Proxy}) = \frac{XCH_{4, \text{raw}}}{XCO_{2, \text{raw}}} \times XCO_2(\text{model}) \quad (1)$$

where XCO_2 (model) represents the independent XCO_2 estimate derived from atmospheric transport models. This model-based value serves as a reliable reference to scale the retrieved ratio into absolute XCH_4 , while ensuring robustness against scattering interference. $XCH_{4,raw}$ and $XCO_{2,raw}$ represent the directly retrieved column-averaged dry-air mole fractions of CH_4 and CO_2 , respectively. (Page 6, Line 208 ~ 209)

Revision (Sect. 3.3):

“Consequently, we conducted a separate analysis on the $XCH_{4,raw}$ and $XCO_{2,raw}$ values (referred to hereafter as raw XCH_4 and raw XCO_2 , respectively), the directly retrieved quantities used to form the proxy ratio in Eq. (1);” (Page 18, Line 542 ~ 543)

Comment 2.6:

Page 23, Line 701

The conclusion that GOSAT-2 SRPR outperforms other products may be somewhat overstated, since the validation periods of GOSAT and GOSAT-2 are not fully consistent, in particular due to the shorter period of observations of GOSAT-2 and the changed conditions and higher CH_4 concentrations. Likewise, the FOCAL time series are much shorter (Fig 3) how does that influences comparison with other algorithms and the statical metrics? Please discuss.

Response:

We have revised the comparison to account for the different temporal coverage of the products. Because the full validation periods of GOSAT and GOSAT-2 are not identical, differences in atmospheric conditions, XCH_4 concentration ranges, and temporal sampling may influence the derived statistics. To reduce this effect, we recalculated the validation metrics for GOSAT SRPR and GOSAT-2 SRPR using only their common observation periods at the Hefei and Xianghe sites. The corresponding results have been added as Table 4, and the related discussion has been revised accordingly. The common-period comparison still shows better overall statistical performance of GOSAT-2 SRPR relative to GOSAT SRPR. Therefore, the original conclusion that GOSAT-2 SRPR “outperforms the other products” has been moderated and is now restricted to its performance within the common validation period.

We have also expanded the discussion of the shorter GOSAT FOCAL Proxy time series. Its limited temporal coverage results in fewer valid satellite–TCCON matchups and a more restricted sampling of atmospheric and XCH_4 conditions. Consequently, its statistical metrics are more sensitive to individual observations and to the specific conditions sampled during the available period, and random variability is less effectively reduced through averaging. Therefore, the FOCAL statistics are interpreted with lower statistical robustness than those derived from the longer time series, and this limitation has been explicitly stated in the revised manuscript.

Revision (Conclusions):

“In terms of site-specific validation, the performance of the GOSAT products (OCPR, SRPR, FOCAL Proxy) is better at Xianghe, whereas the GOSAT-2 SRPR product exhibits comparable performance across both sites. GOSAT OCPR shows the most severe systematic overestimation at the Hefei site. GOSAT SRPR exhibits

opposite bias signs at the two sites. Although the GOSAT FOCAL Proxy performs relatively well in the low-to-medium concentration range (>50% meeting GCOS requirements at both sites), it is constrained by a limited sample size, leading to larger error bars at high concentrations.” (Page 22, Line 667 ~ 671)

Comment 2.7:

Page 24, Line 709-712

The validation is based on only two TCCON sites in China, which represent different atmospheric conditions but cannot fully characterize the performance of satellite products across the entire country. The authors should acknowledge this limitation and further discuss the implications of the results for future satellite-based methane monitoring applications in China.

Response:

We have revised the Conclusions to explicitly acknowledge the limited spatial representativeness of the present validation. The revised text now states that, although the Hefei and Xianghe TCCON sites represent different atmospheric conditions, two sites cannot fully characterize the performance of satellite products across China. We have also expanded the discussion of the implications for future satellite-based methane monitoring in China. Specifically, the site-dependent biases identified in this study suggest that satellite XCH₄ products may exhibit regionally varying performance, highlighting the need for broader multi-site validation and region-specific bias correction. In addition, we note that future applications should further improve the spatiotemporal resolution and regional representation of the XCO₂ prior model, particularly in high-emission regions, to support more reliable and refined methane monitoring in China.

Revision (Conclusions):

“Overall, GOSAT-2 SRPR provides the highest accuracy and correlation under stringent spatial matching conditions, within the validation conditions examined in this study, indicating its potential for refined regional methane monitoring in China.” (Page 23, Line 693 ~ 694)

“It should also be noted that the present validation is limited to the Hefei and Xianghe TCCON sites. Although the two sites represent different atmospheric conditions, they cannot fully characterize the performance of satellite products across China. The site-dependent biases identified in this study indicate that satellite XCH₄ products may exhibit regionally varying performance. Therefore, broader multi-site validation and region-specific bias correction are needed to further improve satellite-based methane monitoring in China. Since the accuracy of Proxy products is inherently constrained by the XCO₂ model accuracy, future applications should prioritize differential bias correction schemes across different climate zones and improve the spatiotemporal resolution and local emission representation of models in high-emission regions, thereby providing more reliable data support for refined methane monitoring in China.” (Page 23, Line 702 ~ 709)

3 Additional Revisions Made by the Authors

3.1 Update of the GOSAT-2 SRPR and TCCON Datasets

During the revision process, updated versions of both the GOSAT-2 SRPR and TCCON datasets became available. The TCCON GGG2020.1 data release became available in June 2026, and the updated GOSAT-2 SRPR version 2.2.0 was released on 7 September 2026. To ensure that the revised manuscript is based on the most up-to-date observations, we updated both datasets to their latest available versions. The update of the TCCON dataset has already been described in detail in our response to Comment 2.2 and is only briefly mentioned here for completeness. Accordingly, all figures and statistical results involving the GOSAT-2 SRPR and TCCON data were regenerated and updated, and the corresponding analyses and discussions were slightly adjusted where necessary to reflect the updated results. The product versions, data sources, and access information have also been revised accordingly in the Data Availability section.

For GOSAT-2 SRPR, the previous release was affected by the limited temporal availability of the XCO₂ model, which required extrapolation from previous years for the later period. This issue has been corrected in version 2.2.0. After updating the dataset and recalculating the validation results, GOSAT-2 SRPR shows better overall statistical performance than GOSAT SRPR under both the $\pm 2.5^\circ$ and $\pm 5^\circ$ matching windows. Accordingly, the related discussion of the sensitivity to the spatial matching range has also been revised.

Following the update of the GOSAT-2 SRPR and TCCON datasets, the results shown in Fig. 14 and the corresponding analysis of the relationship between XCH₄ bias and AOD also changed substantially compared with the original manuscript, particularly at the Xianghe site. Therefore, the previous descriptions of distinct upward/downward curvature and specific AOD values at which the bias reached a maximum were removed, and the corresponding analysis was rewritten based on the updated results.

Revision 1 (Sect. 3.1):

“This discussion shows that the performance of GOSAT-2 SRPR is overall better than that of the other three products, with lower bias, RMSE and SD. When the matching radius is expanded from $\pm 2.5^\circ$ to $\pm 5^\circ$ (Appendix B), the statistical metrics of the products show varying degrees of change (Fig. B2). Nevertheless, GOSAT-2 SRPR still exhibits the best overall performance among the four products, indicating that the relative performance of the products is broadly robust to the choice of matching radius.” (Page 13, Line 409 ~ 414)

Revision 2 (Sect. 4):

“At the Hefei site, the mean XCH₄ bias is positive over the whole AOD range for all four products. For AOD below approximately 0.4, the biases generally remain within about 0–15 ppb and show a slight overall decreasing tendency with increasing AOD. For AOD > 0.5, the number of retrieval data points is sparse (mainly single retrieval results in each AOD bin and only for GOSAT OCPR and SRPR) with a monotonously increasing bias; the biases under these high-AOD conditions are predominantly positive, with some values exceeding 20 ppb. At the Xianghe site, all products are predominantly negatively biased, with the biases generally tending to become more negative as AOD increases. The GOSAT OCPR bias is predominantly positive, with the mean bias decreasing from approximately +5 to +10 ppb under low-AOD conditions to values close to zero as AOD increases. GOSAT SRPR shows the most pronounced negative bias, which becomes increasingly negative with increasing AOD and decreases to below –20 ppb at an AOD of approximately 0.4. GOSAT-2 SRPR exhibits relatively small biases, with values close to zero at low AOD and decreasing slightly to approximately –5 to –10 ppb as

AOD increases; at $AOD > 0.4$, the negative bias becomes smaller again and shifts toward zero. GOSAT FOCAL proxy is also pre-dominantly negatively biased and, despite some fluctuations, shows an overall tendency toward more negative values with increasing AOD.” (Page 13, Line 567 ~ 578)

Revision 3 (Conclusions):

“The influence of aerosols on the retrieval results should be eliminated by using the CO₂ Proxy method. However, the evaluation of the influence of AOD (at 1640 nm) shows the occurrence of an AOD-dependent bias at both sites, but with distinctly different effects. At the Hefei site, the XCH₄ biases are predominantly positive and decrease slightly with increasing AOD below 0.4. At the Xianghe site, OCPR remains positively biased but approaches zero as AOD increases, whereas the other products show negative biases that generally become more pronounced with increasing AOD. Furthermore, the GOSAT-2 SRPR products show high stability under low aerosol concentrations at both sites.” (Page 13, Line 687 ~ 692)

Revision 3 (Data availability, revised manuscript):

“The GOSAT-2 SRPR XCH₄ (v2.2.0) dataset was obtained from the SRON Scientific GOSAT-2 Level 2 Data Products for XCH₄ archive on Zenodo (<https://zenodo.org/records/22637354>).” (Page 29, Line 761 ~ 763)

3.2 Restructuring and Revision of Sect. 5.1

During the revision of Sect. 5.1, we further examined the presentation and interpretation of the relative-bias analysis. The original two figures in this section were based on the same relative-bias time series and contained essentially the same information, but organized in different ways: one grouped the results by bias component (Proxy XCH₄, model XCO₂, and the retrieved XCH₄/XCO₂ ratio), whereas the other grouped the same quantities by retrieval algorithm. To avoid redundancy and improve the clarity of the analysis, Figure 14 in the original manuscript was removed, while Figure 15 was retained and remains Figure 15 in the revised manuscript. which allows the relationships among the three bias terms to be examined directly for each retrieval algorithm.

In addition, following the updates of the GOSAT-2 SRPR and TCCON datasets described above, the relative-bias time series changed substantially. We therefore re-calculated the corresponding results and extensively revised both the analysis and the conclusions in Sect. 5.1. The updated results more clearly reflect the first-order relationship expressed in Eq. (12), in which the relative bias of the final Proxy XCH₄ is related to the combined relative biases of model XCO₂ and the retrieved XCH₄/XCO₂ ratio.

Results analysis, and corresponding conclusions are provided below.

Revision 1(Sect. 5.1):

“For the GOSAT OCPR product (Fig. 15a), the ratio bias generally fluctuates around zero, whereas the model XCO₂ bias becomes predominantly positive after approximately 2020. During the same period, the final XCH₄ bias is also mainly positive. The short-term variations in XCH₄ generally resemble those of the ratio bias, while the model XCO₂ bias varies more smoothly.

For the GOSAT SRPR algorithms (Fig. 15b), the model XCO₂ bias is generally negative, whereas the ratio bias remains positive during most of the time series. Between 2018 and 2022, the two components have opposite

signs and partially compensate for each other, resulting in a final XCH_4 bias that remains close to zero. From late 2022 onward, the negative model XCO_2 bias increases markedly, reaching approximately -1.0% to -1.5% , while the ratio bias remains positive but smaller in magnitude. Consequently, the final XCH_4 bias also becomes markedly negative during this period. This pronounced decrease is attributed to the limited temporal availability of the model data.

For GOSAT-2 SRPR (Fig. 15c), the biases of both components are generally smaller in magnitude than those of GOSAT SRPR. The model XCO_2 bias is mainly negative, whereas the ratio bias is generally close to zero or moderately positive. During 2021–2022, the biases of both components are relatively small, and the final XCH_4 bias are therefore also close to zero. In 2023, the negative model XCO_2 bias increases, accompanied by a negative bias in the final XCH_4 . Around 2024, the positive ratio bias increases, while the magnitude of the negative final XCH_4 bias decreases. Toward the end of the time series, all three relative biases gradually approach zero.

Overall, the comparison shows that the negative bias at the Xianghe site around 2023 is mainly associated with the substantially negative bias in model XCO_2 during the same period. Overall, the final Proxy XCH_4 bias is related not only to the magnitudes of the biases in the retrieved $\text{XCH}_4/\text{XCO}_2$ ratio and model XCO_2 , but also to whether the two components compensate or reinforce each other. A small final bias may therefore arise either from small biases in both components or from cancellation between biases of opposite signs.” (Page 20, Line 603 ~ 622)

Revision 1(Conclusions):

“Component decomposition shows that the final Proxy XCH_4 bias results from the combined effects of biases in the retrieved $\text{XCH}_4/\text{XCO}_2$ ratio and the XCO_2 model. A small final XCH_4 bias may arise from compensation between component biases of opposite signs. Therefore, product performance depends not only on the magnitude of the bias in each individual component, but also on the temporal stability of the error balance between the two components.” (Page 22, Line 676 ~ 679)