

## **Response to Comments on the Manuscript (essd-2023-28):**

### **Seamless mapping of long-term (2010-2020) daily global XCO<sub>2</sub> and XCH<sub>4</sub> from GOSAT, OCO-2, and CAMS-EGG4 with a spatiotemporally self-supervised fusion method**

---

Dear Editors and Referees,

We would like to sincerely express our gratitude to you for your careful reading and constructive comments.

According to the comments, we have tried our best to improve the manuscript, and an item-by-item response follows. The modified parts have been highlighted in yellow color in the revised manuscript.

Once again, we are particularly grateful for your careful reading and constructive comments. Thanks very much for your time.

Best regards,

Qiangqiang Yuan

## Response to Comments of Referee #2:

### General comments:

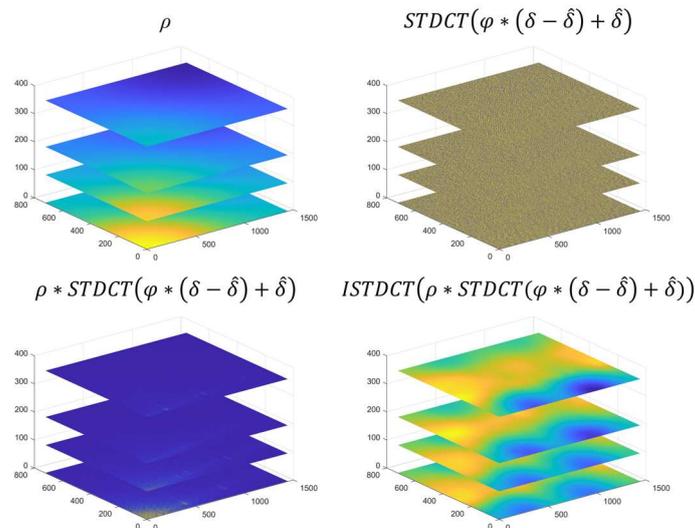
This study describes an effective approach to generate global long-term seamless XCO<sub>2</sub> and XCH<sub>4</sub> based on a self-supervised fusing method from OCO-2, GOSAT, and CAMS-EGG4. Generally, this paper is well organized and written, of which the methodology, validation techniques, and experiment results are reasonable. However, some details are unclear and several issues are required to be modified before this paper being published in ESSD. A major revision is recommended. Specific comments are listed as follows.

**Response:** We sincerely appreciate the referee for his/her comments and suggestions for improving the paper. An item-by-item response to the valuable comments raised by the referee follows. Thanks for your time.

### Major comments:

**Q2.1:** P9L202, Eq. (6): Is it possible to visualize intermediate results of STDCT? The visualization of intermediate results of STDCT can help understand this procedure.

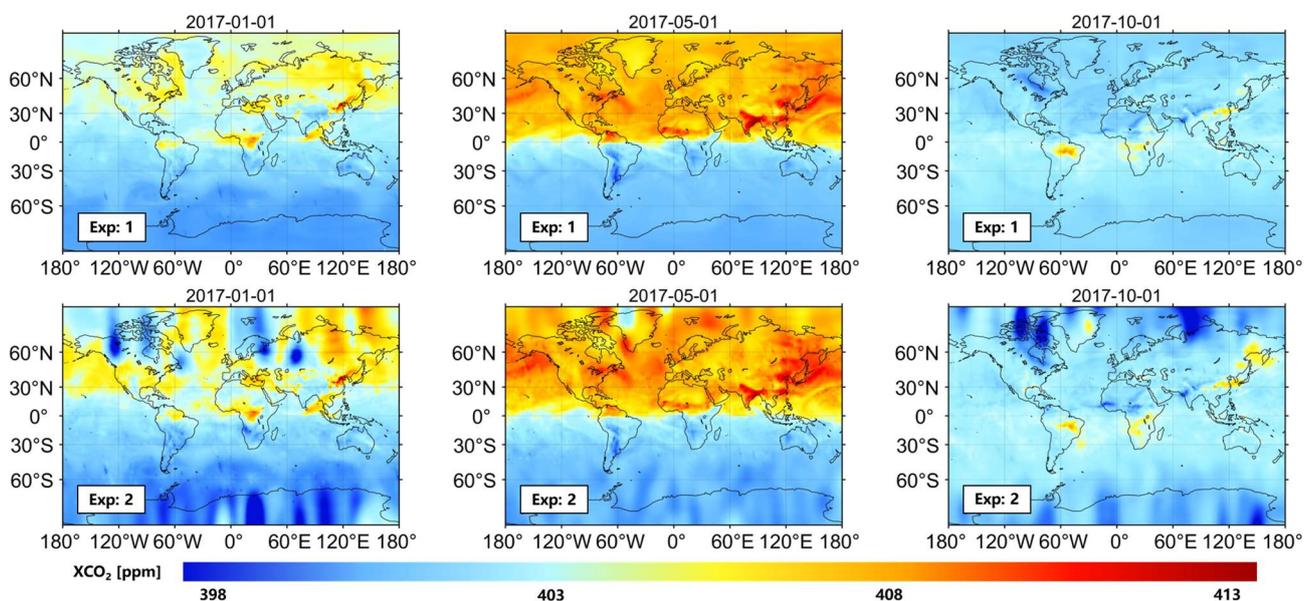
**Response:** Thank the referee for his/her comment. An example to visualize the intermediate results of *STDCT* (fusion with OCO-2 XCO<sub>2</sub>) at the 20th iteration in 2017 has been presented (see Figure r1).



**Figure r1.** Visualization of intermediate results of *STDCT* (fusion with OCO-2 XCO<sub>2</sub>) at the 20th iteration in 2017.

**Q2.2: P9L205, Eq. (7):** I notice that the power of the subitem in the denominator is 1, which is different from that in the given references, such as Garcia (2010).

**Response:** Thank the referee for his/her careful comment. The exponent (Exp) of the subitem in the denominator from Eq. (7) can be set as 1 or 2 (default) (Garcia, 2010), which controls the effect of smoothing. However, as shown in the example (see Figure r2), the fused results (such as with OCO-2 XCO<sub>2</sub>) could be over-smoothed in high-latitude regions (> 60°N or S) when the Exp is configured as 2. By contrast, setting this parameter as 1 will largely reduce the over-smoothing effect. Therefore, the Exp is considered 1 in our study to provide more reasonable results.



**Figure 2.** Daily fused XCO<sub>2</sub> using the exponents of (upper) 1 and (lower) 2 over the globe. Color ramp stands for the values of XCO<sub>2</sub>.

### Reference:

Garcia, D.: Robust smoothing of gridded data in one and higher dimensions with missing values, *Computational Statistics & Data Analysis*, 54, 1167–1178, <https://doi.org/10.1016/j.csda.2009.09.020>, 2010.

**Q2.3:** Please present the whole formula derivation processes from Eq. (5) to Eq. (6).

**Response:** Thank the referee for his/her comment. Since *STDCT* is based on a composition of one-dimensional DCTs along each dimension (Strang, 1999), the solution for one-dimensional DCT is given as an instruction as follows:

$$E(\hat{\delta}) = \|(\hat{\delta} - \delta)\|^2 + \varepsilon R(\hat{\delta}) \quad (\text{r1})$$

where  $\| \cdot \|$  signifies the Euclidean norm;  $\delta$  and  $\hat{\delta}$  are varying vectors;  $\varepsilon$  indicates a smoothing factor. A simple and straightforward approach to express the roughness ( $R$ ) is by using a second-order divided difference (Weinert, 2007, Whittaker, 1923) which yields, for a one-dimensional data array:

$$R(\hat{\delta}) = \|M\hat{\delta}\|^2 \quad (\text{r2})$$

where  $M$  is a tridiagonal square matrix defined by:

$$M_{i,i-1} = \frac{2}{s_{i-1}(s_{i-1} + s_i)}$$

$$M_{i,i} = \frac{-2}{s_{i-1}s_i}$$

$$M_{i-1,i} = \frac{2}{s_i(s_{i-1} + s_i)}$$

for  $2 \leq i \leq n-1$ , where  $n$  is the number of values in  $\hat{\delta}$ , and  $s_i$  represents the step between  $\hat{\delta}_i$  and  $\hat{\delta}_{i+1}$ . Assuming repeating border elements ( $\delta_0 = \delta_1$  and  $\delta_{n+1} = \delta_n$ ) gives:

$$-M_{1,1} = M_{1,2} = \frac{1}{s_1^2}$$

$$-M_{n,n-1} = -M_{n,n} = \frac{1}{s_{n-1}^2}$$

At present, combining Eq. (r1) and (r2) acquires:

$$(I_n + \varepsilon M^T M)\hat{\delta} = \delta \quad (\text{r3})$$

where  $I_n$  is the  $n \times n$  identity matrix and  $M^T$  stands for the transpose of  $M$ . Eq. (r3) can be further extended to multi-dimensional regularly gridded data using DCTs, which is much more complicated. More detailed descriptions for the extension of Eq. (r3) are provided in Buckley (1994) and Garcia (2010).

### Reference:

Buckley, M.J.: Fast computation of a discretized thin-plate smoothing spline for image data, *Biometrika*, 81, 247–258, <https://doi.org/10.1093/biomet/81.2.247>, 1994.

Garcia, D.: Robust smoothing of gridded data in one and higher dimensions with missing values, *Computational Statistics & Data Analysis*, 54, 1167–1178, <https://doi.org/10.1016/j.csda.2009.09.020>, 2010.

Strang, G.: The discrete cosine transform, SIAM Review, 41, 135–147, <https://www.jstor.org/stable/2653173>, 1999.

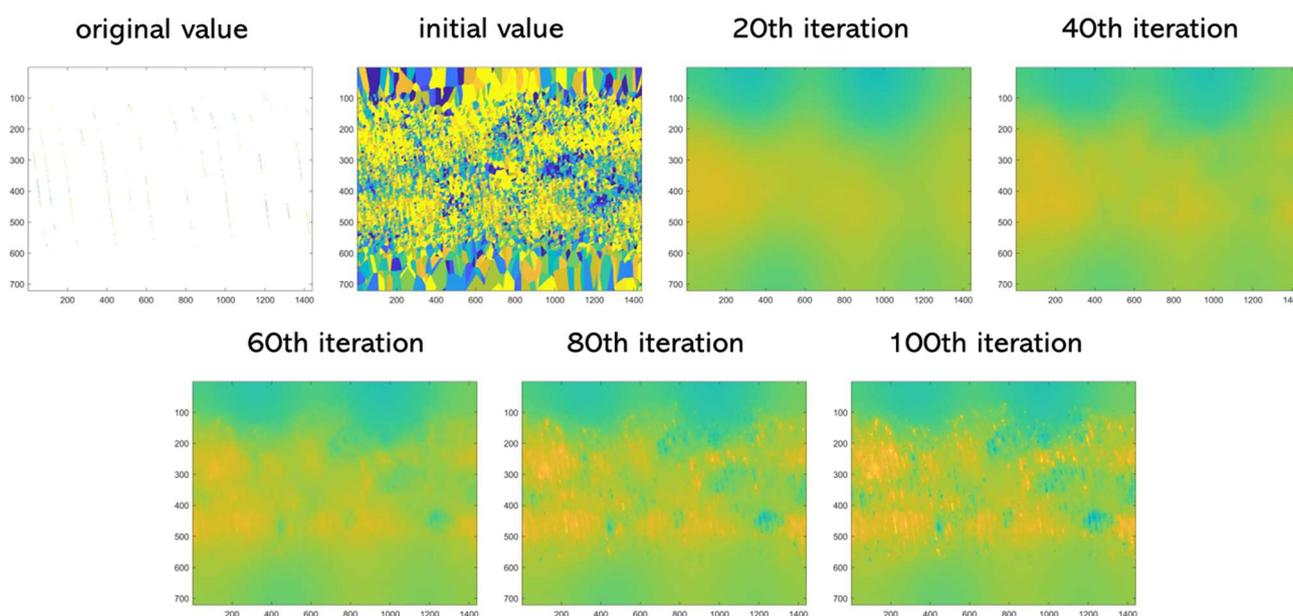
Weinert, H.L.: Efficient computation for Whittaker–Henderson smoothing, Computational Statistics & Data Analysis, 52, 959–974, <https://doi.org/10.1016/j.csda.2006.11.038>, 2007.

Whittaker, E.T.: On a new method of graduation, Proceedings of the Edinburgh Mathematical Society, 41, 62–75, <https://doi.org/10.1017/S0013091500077853>, 1923.

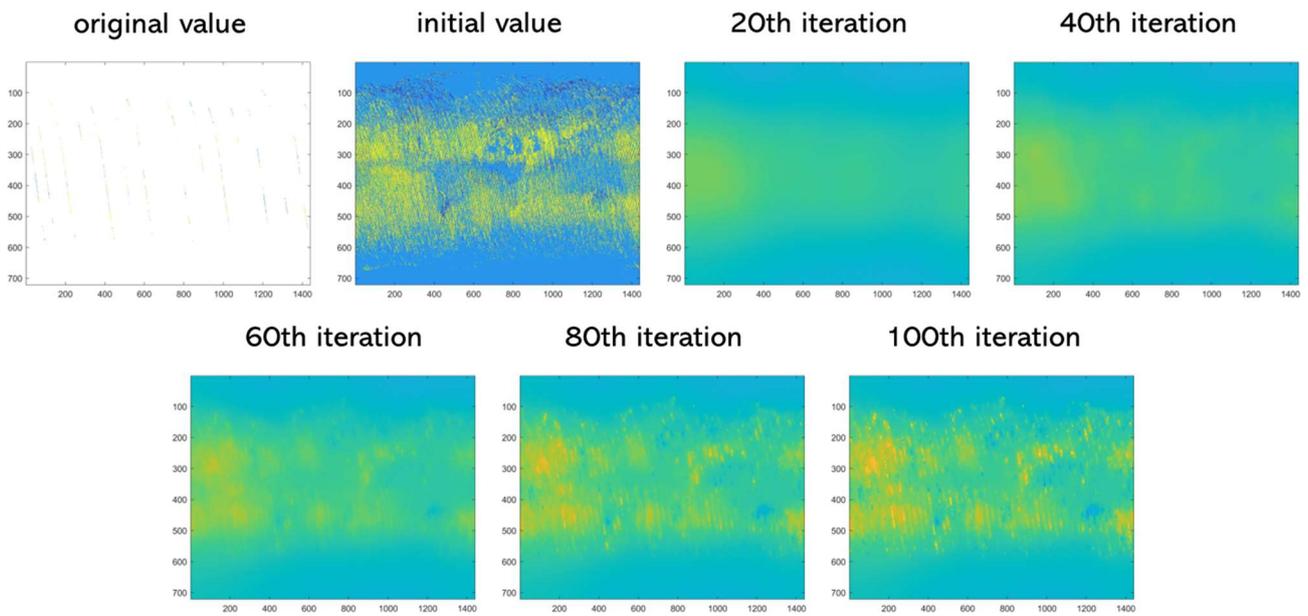
**Q2.4: P9L209: Different initializations of  $\hat{\delta}$  may lead to different final results. Please provide a brief discussion?**

**Response:** Thank the referee for his/her constructive comment. At first, we would like to apologize for that the description of the initializations for missing values in  $\hat{\delta}$  is misleading in our study, which has been revised. Due to the large sparsity of satellite XCO<sub>2</sub> and XCH<sub>4</sub>, different initializations of  $\hat{\delta}$  will cause non-negligible differences in fused results. Here, an example of three initializations is presented to tell the differences among them, defined as follows:

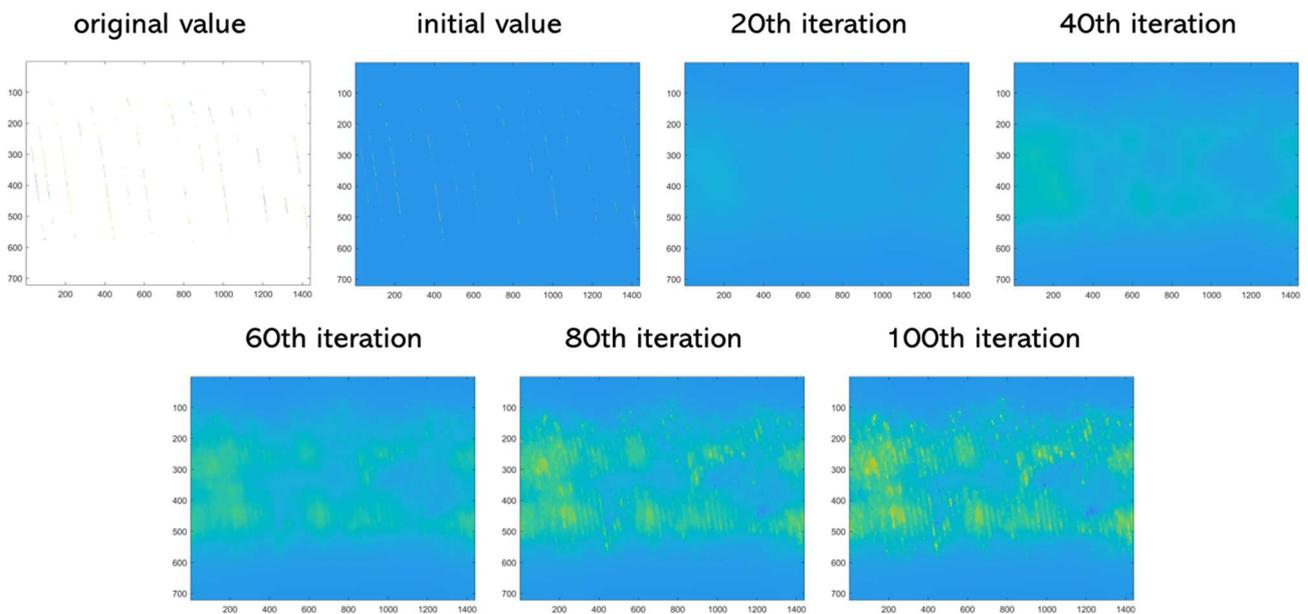
1. Type 1: spatiotemporal nearest neighbor interpolation (adopted in our study).
2. Type 2: temporal nearest neighbor interpolation.
3. Type 3: replacement with a constant value (e.g., 1).



**Figure r3.** Schematic diagram to show the iteration process of  $\delta$  (fusion with OCO-2 XCO<sub>2</sub>) via Type 1 initialization in 2017-04-10.

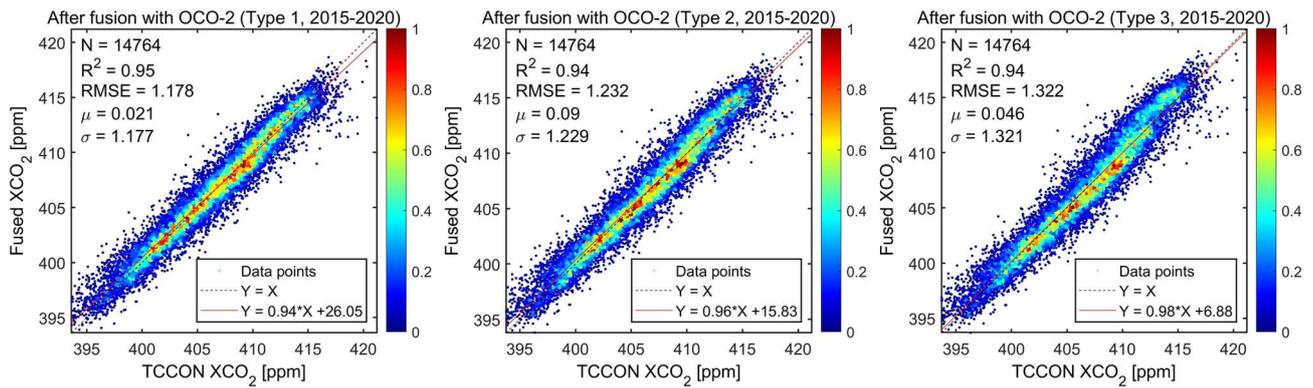


**Figure r4.** Schematic diagram to show the iteration process of  $\delta$  (fusion with OCO-2 XCO<sub>2</sub>) via Type 2 initialization in 2017-04-10.



**Figure r5.** Schematic diagram to show the iteration process of  $\delta$  (fusion with OCO-2 XCO<sub>2</sub>) via Type 3 initialization in 2017-04-10.

Figure r3-r5 demonstrate the iteration processes of  $\delta$  (fusion with OCO-2 XCO<sub>2</sub>) using three types of initializations in 2017-04-10. It is clear that different initializations of  $\delta$  could generate similar but different results. Figure r6 illustrates the in-situ validation results of the fused results with OCO-2 XCO<sub>2</sub> during 2015-2020 through three types of initializations. As observed, the fused XCO<sub>2</sub> using Type 1 initialization achieves the best performance, which signifies that more prior information in the initialization can improve the fusion accuracy.



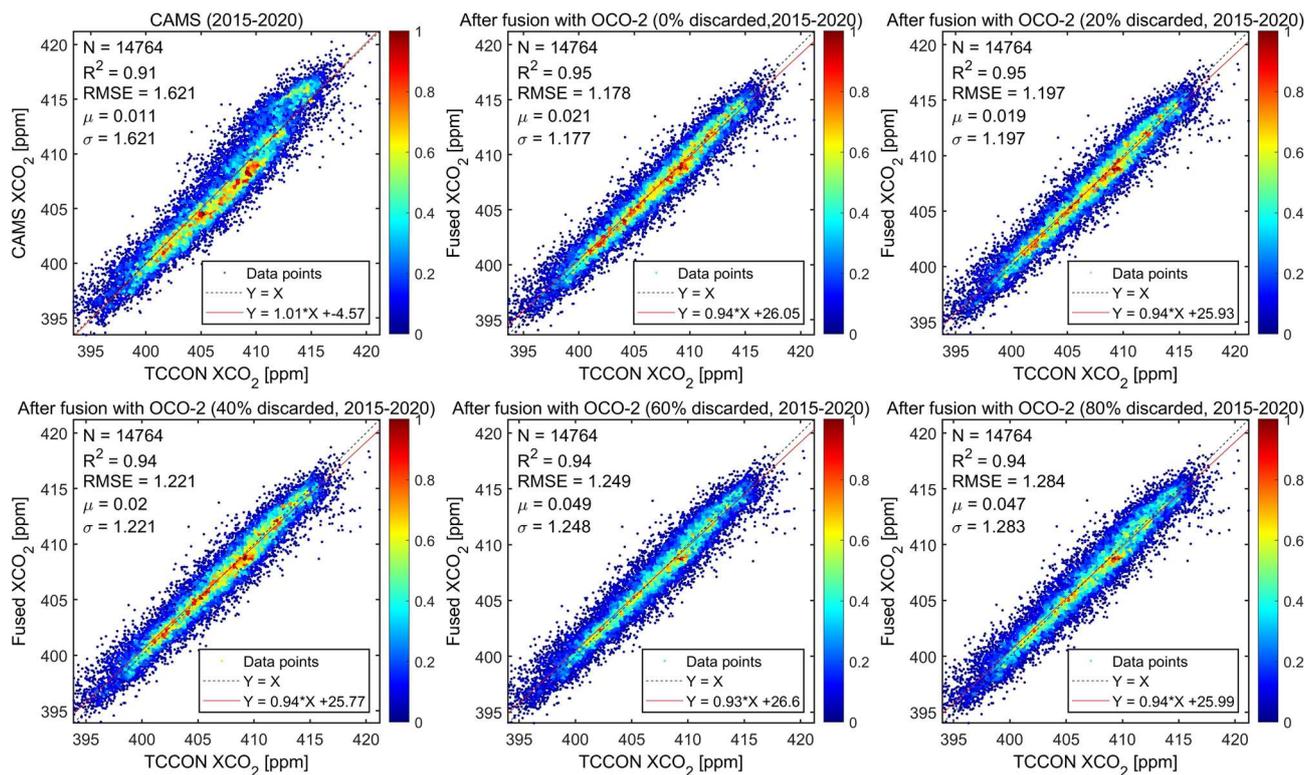
**Figure r6.** Density scatter-plots of the in-situ validation results for fused XCO<sub>2</sub> with OCO-2 using (left) Type 1, (middle) Type 2, and (right) Type 3 initialization from 2015 to 2020. Black dotted and red full lines stand for the 1:1 and fitted lines, respectively. Color ramps show the normalized densities of data points. X: TCCON data; Y: fused data. Unit: ppm for RMSE,  $\mu$ , and  $\sigma$ .

### The main revision is as follows:

It is worth noting that  $\hat{\delta}$  is initialized through the spatiotemporal nearest neighbor interpolation.

**Q2.5: Will the data completeness of XCO<sub>2</sub>/XCH<sub>4</sub> from OCO-2/GOSAT affect the accuracy of final fused results? More data should imply more usable information. Please provide a brief discussion.**

**Response:** Thank the referee for his/her crucial comment. An example to show the in-situ validation results of the fused results with OCO-2 XCO<sub>2</sub> (different data completeness) during 2015-2020 is depicted in Figure r7. As can be seen, the fusion with OCO-2 XCO<sub>2</sub> of less data completeness (20-80% discarded) can variously reduce the performance. However, the fused results still present better accuracy than that of CAMS-EGG4, which indicates the robustness of the proposed fusion method.



**Figure r7.** Density scatter-plots of the in-situ validation results for CAMS-EGG4 and fused XCO<sub>2</sub> with OCO-2 (different data completeness) from 2015 to 2020. Black dotted and red full lines stand for the 1:1 and fitted lines, respectively. Color ramps show the normalized densities of data points. X: TCCON data; Y: CAMS-EGG4/fused data. Unit: ppm for RMSE,  $\mu$ , and  $\sigma$ .

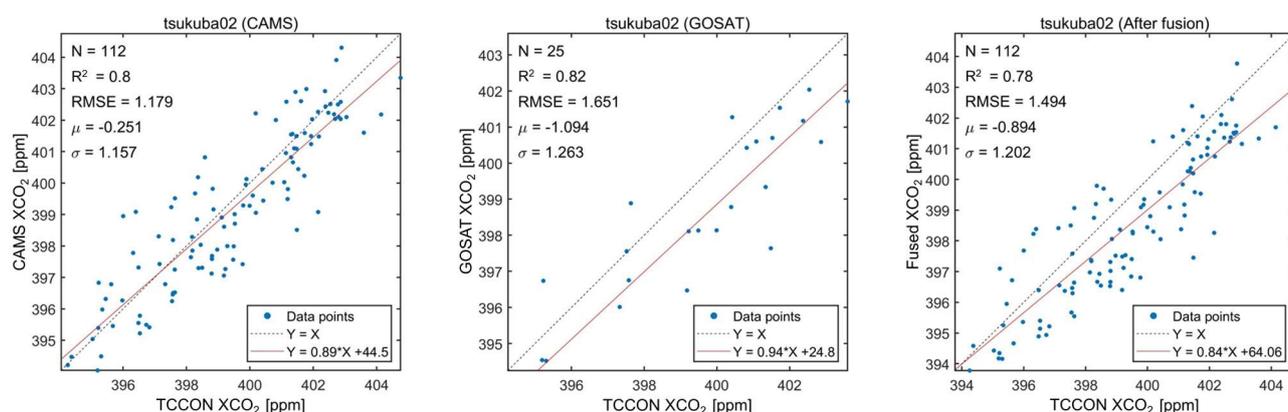
**Q2.6:** The data of XCO<sub>2</sub>/XCH<sub>4</sub> from OCO-2/GOSAT is extremely sparse in some regions, I wonder if the performance could be improved after fusion under this condition.

**Response:** Thank the referee for his/her crucial comment. Same to Q2.5 (see 80% discarded in Figure r7), the fused results with extremely sparse data present a decreased accuracy, which is still superior to that of CAMS-EGG4.

**Q2.7:** Table 3-5: The metrics of the individual in-situ validation do not exceed those of CAMS-EGG4 for a few stations after fusion. What could be the potential reasons? Please provide a further discussion.

**Response:** Thank the referee for his/her comment. The performance will reduce for a few stations after fusion, which is mainly affected by the poor quality of satellite XCO<sub>2</sub> and XCH<sub>4</sub>. For instance (see Figure r8), all the metrics (e.g.,  $R^2$ , RMSE, and  $\sigma$ ) of fused XCO<sub>2</sub> with GOSAT are worse than those

of CAMS-EGG4 on tsukuba02 from 2010 to 2014. This is likely attributed to the generally underestimated values of GOSAT XCO<sub>2</sub> (i.e.,  $\mu$ : -1.094 ppm).



**Figure r8.** Scatter-plots of the in-situ validation results for CAMS-EGG4 (left), GOSAT (middle), and fused XCO<sub>2</sub> (right) on tsukuba02 during 2010-2014. Black dotted and red full lines stand for the 1:1 and fitted lines, respectively. X: TCCON data; Y: CAMS-EGG4/GOSAT/fused data. Unit: ppm for RMSE,  $\mu$ , and  $\sigma$ .

### Minor comments:

**Q2.8: P5L121:** The authors did not consider the latest XCO<sub>2</sub> from OCO-3 for fusion. What is the reason?

**Response:** Thank the referee for his/her comment. The latest XCO<sub>2</sub> from OCO-3 presents a similar accuracy to that from OCO-2 with a shorter period (available after August 2019) (Taylor et al., 2023). As a result, the OCO-2 XCO<sub>2</sub> product is currently employed for fusion in this study. The XCO<sub>2</sub> from OCO-3 can be considered in our future works.

### Reference:

Taylor, T. E., O'Dell, C. W., Baker, D., Bruegge, C., Chang, A., Chapsky, L., Chatterjee, A., Cheng, C., Chevallier, F., Crisp, D., Dang, L., Drouin, B., Eldering, A., Feng, L., Fisher, B., Fu, D., Gunson, M., Hammerle, V., Keller, G. R., Kiel, M., Kuai, L., Kurosu, T., Lambert, A., Laughner, J., Lee, R., Liu, J., Mandrake, L., Marchetti, Y., McGarragh, G., Merrelli, A., Nelson, R. R., Osterman, G., Oyafuso, F., Palmer, P. I., Payne, V. H., Rosenberg, R., Somkuti, P., Spiers, G., To, C., Wennberg, P. O., Yu, S., and Zong, J.: Evaluating the consistency between OCO-2 and OCO-3 XCO<sub>2</sub> estimates derived from the NASA ACOS version 10 retrieval algorithm, AMTD, <https://doi.org/10.5194/amt-2022-329>, 2023.

**Q2.9: P5L130:** Similarly, the authors also did not adopt the popular XCO<sub>2</sub> from Carbon Tracker

**for fusion.**

**Response:** Thank the referee for his/her comment. The XCO<sub>2</sub> from the Carbon Tracker (<https://gml.noaa.gov/ccgg/carbontracker/>) is popular but performed at a coarse spatial resolution of 3°×2°. In addition, the Carbon Tracker only provides XCH<sub>4</sub> from 2000 to 2010. By contrast, the CAMS-EGG4 XCO<sub>2</sub> and XCH<sub>4</sub> products are more appropriate and adopted in this study. The XCO<sub>2</sub> and XCH<sub>4</sub> from the Carbon Tracker can be considered in our future works.

**Q2.10: The figures in the Supplement are too many to follow, which are unnecessary. Table 3-5 have summarized their metrics.**

**Response:** Thank the referee for his/her comment. The figures in the Supplement have been removed from the manuscript.

**Q2.11: Table 3-5: It's better to abbreviate “CAMS-EGG4” into “CAMS” instead of “CE”, which is consistent with other texts.**

**Response:** Thank the referee for his/her careful comment. “CAMS-EGG4” has been abbreviated into “CAMS” in Table 3-5.

**The main revision is as follows:**

**Table 3.** Metrics of the individual in-situ validation results for CAMS-EGG4, GOSAT, and fused XCO<sub>2</sub>. The best and second metrics are denoted with bold and underlined fonts. CAMS: CAMS-EGG4; AF: after fusion. Unit: ppm for RMSE and  $\sigma$ .

**Table 4.** Metrics of the individual in-situ validation results for CAMS-EGG4, OCO-2, and fused XCO<sub>2</sub>. The best and second metrics are denoted with bold and underlined fonts. CAMS: CAMS-EGG4; AF: after fusion. Unit: ppm for RMSE and  $\sigma$ .

**Table 5.** Metrics of the individual in-situ validation results for CAMS-EGG4, GOSAT, and fused XCH<sub>4</sub>. The best and second metrics are denoted with bold and underlined fonts. CAMS: CAMS-EGG4; AF: after fusion. Unit: ppb for RMSE and  $\sigma$ .

**Q2.12: Future works and limitations are missing in the Conclusions.**

**Response:** Thank the referee for his/her comment. Future works and limitations have been appended in the manuscript.

**The main revision is as follows:**

Overall, the developed fusion method generates high-quality full-coverage XCO<sub>2</sub> and XCH<sub>4</sub> datasets over the globe from 2010 to 2020. However, it only considers the global spatiotemporal knowledge of self-correlation in GOSAT and OCO-2 products without attention to local spatiotemporal information. Meanwhile, the spatial resolution and available period of fused results should be further enhanced, which are devised as 0.1° and more than 20 years (e.g., 2000-2020), respectively. To fix these issues, we will spare no effort to work on our future works.

**Q2.13: Is it feasible to acquire global seamless XCO<sub>2</sub> and XCH<sub>4</sub> only from OCO-2 and GOSAT based on the proposed method?**

**Response:** Thank the referee for his/her comment. Generating global seamless XCO<sub>2</sub> and XCH<sub>4</sub> only from OCO-2 and GOSAT is still a challenge due to their significant sparsity without any external data. At present, the proposed method merely can provide some over-smoothed results, which are required to be improved in our future works.

Last but not least, we gratefully thank the referee again for his/her significant comments and suggestions, which have greatly helped us to improve the technical quality and presentation of our manuscript.