

Comment on **essd-2022-215**

Mengqi Zhang

Community comment on "Carbon dioxide cover: carbon dioxide column concentration seamlessly distributed globally during 2009–2020" by Haowei Zhang et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-215-CC1>, 2022

In this paper, a spatiotemporal interpolation method is developed, and a data product with full spatiotemporal coverage is generated by using the XCO₂ data of GOSAT. I'm not particularly aware of the article category for ESSD, but compared to other research papers in ESSD, it's more suitable for technical description articles, at least at this stage. The article has some obvious scientific errors and inappropriate knowledge descriptions, the following points should be considered to improve the quality of the article, especially some major errors. As community comments, I believe these comments will increase the understanding of carbon monitoring satellite data assimilation and improve this research.

Although we are very concerned about the carbon cycle and the spatiotemporal distribution of CO₂, for atmospheric inversion models, sparse data observations are sufficient to obtain carbon fluxes. **NOTE I'm not denying that we don't need a spatially seamlessly CO₂ distribution**, but the introduction should explain why we need a spatiotemporally seamlessly CO₂, such as calculating global averages, analyzing seasonal changes.

P6. Line 160. **Equation 1 looks very strange and seems wrong.** What is the difference between coefficients c and e ? What is the difference between d and g ? The result needs to be checked carefully and even recalculated. In addition, adopting this method to construct time series would lead to significant drawbacks. I believe the authors may not understand the CO₂ growth rate, please see papers such as Buchwitz. et.al, 2018 and A. Chatterjee et.al, 2017 to understand the significance of CGR in reflecting vegetation, climate, etc. The method of presetting a function to fit will not be able to capture the real change in the CO₂ growth rate, and the function will be directly known after derivation. One of the reasons we want seamless data is to better calculate the global average, and thus the growth rate, that is, the net flux. Fitting with a fixed function would lose this.

In the abstract and the introduction, the authors claim that the amount of XCO₂ data is mainly affected by factors such as clouds and aerosols. This is wrong, it's actually a swath issue. The design of the width is related to the fluctuation amplitude of atmospheric CO₂ and optical inversion factors. Authors are advised to read the relevant literature and correct the description in this section. While this is not that important for this article, the readers need to understand the real background for this work.

P3, L65-83. This section overlaps with method descriptions in subsequent part, and it is

not appropriate to introduce too much about the methods of this study in the introduction.

P4, L98 "...the accuracy of the comparison between the GOSAT data product and the TCCON site was 0.56 ppm", it is not appropriate to use the "**accuracy**" word, it should be stated, such as standard deviation, bias, etc.

P4, L100, "... three-day temporal resolution. The time resolution of GOSAT-2 satellite is 6 days..." is inappropriate. The correct description is the revisit cycle/repeat cycle. Please differentiate these concepts.(temporal resolution,time resolution, repeat cycle,)

P4, L105. However, the OCO-2_L2_Lite_FP9r provides data locations that are gradually shifted over time by satellite observations. This sentence is difficult to understand. I think you should express that the orbits of the sub-satellite points are evenly distributed? Illustration may be needed.

P4, L112 column-averaged XCO2. This is wrong. And the full name of XCO2 is wrong, including the title, abstract and etc. I think it should be carefully checked the full text of the corresponding full scientific name.

P4. L107, "fixed location" should be more clear. L109. the six data channels are wrong. TANSO-FTS is a 4-band interferometer.

P4., L123. Please correct for **column-averaged abundances of CO2** expression. And the results showed that R2 was 0.9686, and RMSE was 1.3811 . Please indicate the source.

P5. 146 EBK theory or EBK method ? Please express it in a unified way. such as L150.

In Section 2.4, more indicators for accuracy evaluation should be added. Bias and standard deviation are necessary in the verification of XCO2.

P7. 195. It is not necessary to show the results of each year, only a few examples, such as some months or a specific year, are sufficient. From the analysis of the data, such as some seasonal changes, changes in CO2 growth, and spatial differences may be more meaningful.

Figure 6, As said at the beginning, of course we know that CO2 is rising, but its growth rate is more meaningful, and it is recommended to draw a related graph of the growth rate. If it does not reflect reasonable fluctuations, but a fully sinusoidal pattern, the study would be significantly flawed.

From the research point of view, averaging kernel and the prior profile should be considered in comparison with OCO-2. Although they may be ignored in some cases and not important on monthly validation where accuracy is not required, the article should mention it.

Figure 3,4 and 5. P_{XCO2} T_{XCO2} P_{XCO2} should be described uniformly (note the space). Other than that, I would suggest that it would be better to do time series validation on a monthly basis. Specifically, the horizontal axis is time, and the vertical axis is parameters such as error, which can also be filled with error distribution, which is more intuitive.

Figure 7~17. It looks like this resolution may be trapped in a highly smooth phenomenon, which means it may not really be 0.25 degrees. It is recommended to draw a detailed map of some regions to show that the method does have this good resolution and can capture reasonable and sufficient spatial gradient changes. It is also recommended to compare the results of models, such as CarbonTracker or the L4B model products of

GOSAT-NIES, to demonstrate the rationality of the results.

As mentioned above, in addition to the error evaluation in the time dimension, the error in the spatial dimension should also be evaluated to illustrate the reliability of the data.

I am not a Native English speaker. But I believe that the English of this article should be greatly improved, including scientific names of many nouns, and descriptions in scientific language.

Buchwitz, Michael, et al. "Computation and analysis of atmospheric carbon dioxide annual mean growth rates from satellite observations during 2003–2016." *Atmospheric Chemistry and Physics* 18.23 (2018): 17355-17370.

Chatterjee, A., et al. "Influence of El Niño on atmospheric CO₂ over the tropical Pacific Ocean: Findings from NASA's OCO-2 mission." *Science* 358.6360 (2017): eaam5776.