Review of essd-2022-215

Zhang et al., "Carbon dioxide cover: carbon dioxide column concentration seamlessly distributed globally during 2009–2020"

This paper describes a gap-fill method for discrete XCO_2 data from the GOSAT satellite data product, and the resulting dataset, the Carbon Dioxide Coverage (CDC). The CDC contains monthly XCO_2 data with a spatial resolution of $0.25° \times 0.25°$ grid box, globally for the period 2009–2020. The authors further make efforts to validate the dataset with other XCO_2 data from TCCON and OCO-2. The CDC dataset has already been made available on the "figshare" website and is downloadable with a readme file. Such a statistical mapping dataset makes it possible to understand global XCO_2 features from discrete satellite data. However, I found some parts of the manuscript difficult to follow, and I think the authors need to elaborate more on the data, the method used, and on the method of evaluation. As the focus of the manuscript seems to be on the technical description of the method, I feel that this manuscript with major revision may be more suitable for a technical journal or a journal on data modeling or analysis.

General comments:

1.   The authors should clarify the reason more why they used XCO_2 from the GOSAT FTS SWIR Level 3 (L3) product and not the FTS SWIR Level 2 (L2) product, to create the CDC dataset. The GOSAT L3 product is processed dataset of XCO_2 in GOSAT L2 product that filled the gaps in the L2 XCO_2 data with a Kriging method. The L3 product contains monthly-averaged XCO_2 values on a 2.5 degree latitude $\times$ 2.5 degree longitude grid. The L3 product covers all grids on Earth, but any mesh with no L2 data within a circle of 500-km radius were masked in the product to remove estimated data with extremely high minimum mean square errors (Watanabe et al., 2015). So, I am wondering why the authors chose GOSAT L3 product in which a Kriging method was already applied. Usually, these kinds of statistical gap-fill methods are applied to XCO_2 in the GOSAT or OCO-2 L2 data products which contains retrieved XCO_2 at each satellite observation point from spectral data in the Level 1B product (e.g. Hammerling et al., 2012; Liu et al., 2012; Watanabe et al., 2015). It would be more effective/impressive to show the usefulness of the proposed method in this manuscript if the authors applied the method to L2 XCO_2 data as the L2 XCO_2 data is more sparse

than L3 XCO$_2$ data.

2. The authors emphasis that the proposed method with the EBK methods for spatial prediction and time variation of XCO$_2$ improve the dataset accuracy effectively (e.g. Abstract, L65, 90, 146). However, it seems inadequate to conclude so, only from written results of the evaluations in the current manuscript which shows the results from the CDC dataset only. For more detailed evaluation for the authors' intended conclusion, the CDC dataset should be compared to XCO$_2$ dataset with previous methods such as the ordinary Kriging method as well as the original GOSAT L3 XCO$_2$ against TCCON observations.

3. Overall, the manuscript is somewhat vague and suffers from numerous incorrect descriptions. The presentation style could certainly be improved through corrections and elaborating the procedure more so that readers can understand the data processing procedure used to create the CDC dataset. Please see specific comments below for detail.

Specific comments:

L1: It seems to me that the title does not summarize the contents. One cannot guess the contents is on observations, model simulations, data analysis or dataset on CO2 from the title. The title should be more specific.

L11, 43: XCO$_2$ is the column-averaged dry-air mole fraction of CO$_2$ and is not equal to simple CO$_2$ concentration. Please define XCO$_2$ correctly.

L13, 35, 44 etc.: I cannot understand what the authors mean by "sniffing" in "carbon sniffing satellites" or "sniffer" in "sniffer satellites".

L13: Please replace "Gosat-2" by "GOSAT-2". A proper noun such as satellite names should be referred correctly.

L18, 67: Please clarify "the raw GOSAT data". In the later section, the authors write that they used GOSAT L3 product which is not "raw" data but processed data.

L29: I cannot understand what the authors mean by "respectively" in this sentence.

L32: Did the authors use "R" here, not R^2? Please clarify why they use R only here, while R^2 is used elsewhere.

L63: Please clarify "the prior time profile information of XCO_2". Is this time series of observed XCO2 or model simulated XCO2 or something else that is prepared prior to the analysis?

L65: What are "the original data" and "the prior information" exactly in this case?

L72: I cannot find in the method section how the "populated data" is used in this study. Please describe the populated data in the methods section.

L84: The authors write the spatial coverage of CDC is "from 50°S to 50°N" in Line 76 and "55°S to 55°N" in Line 84. Which is correct for the CDC dataset?

L87: What is "the model" here? Please define.

L89, 97: Replace "Tansat" by "TanSat".

L98, 105: I think that "GOSAT_FTS_L3_V2.95" and "OCO-2_L2_Lite_FP9r" are not proper product names. The authors need to clarify exactly the dataset they used; the product name, version, data provider etc.

L99: I could not find the comparison results of "0.56 ppm" in Watanabe et al., 2015. "0.56 ppm" in Noel et al., 2021 is for retrieved XCO2 by the ACOS FOCAL algorithm, but not for Level 3 product used here. The authors should read papers carefully and refer correctly.

L100: Do the authors use GOSAT-2 product as well?

L105: Related to my general comment #1, it is unclear why the authors discuss Level 3 and Level 2 products in same manner as the purposes and properties of the products

are different. The product the authors used in this study is GOSAT Level 3 product which is processed XCO_2 data with fixed grids in monthly interval processed by a geostatistical mapping approach (a Kriging) based on Level 2 product. While the "OCO-2_L2_Lite_F99r" is Level 2 product, it is natural that data locations are shifted along the satellite orbit. If the GOSAT L3 data used here is from NIES, Japan, their GOSAT L2 product is also downloadable from their website.

L109: The GOSAT TANSO-FTS sensor has seven channels: six for SWIR three bands and one for TIR, and GOSAT also carries TANSO-CAI sensor (JAXA/NIES/MoE, "GOSAT/IBUKI Data Users Handbook", 2011). So, "the six data channels of the sensor" may be replaced by "six of the seven channels of the TANSO-FTS sensor".

L110: The sentence is not correct. GOSAT collect data even in polar regions during winter. Just the spectral data in such regions are not used in XCO_2 retrieval processing for the SWIR Level 2 product.

L112: Please replace http by https.

L123-124, 125: The numbers, 0.9686, 1.3811, and 0.7, suddenly appear without any explanation about the evaluation method. Since this section is on "Materials and Methods", the results should be written in the results section.

L127: Was this analysis performed using a particular software program and/or libraries? Please give software information in the method section so it can be referred by other investigators.

L136: The authors did not mention the data provider of Level 3 data used in this study. It is supposed from the version "V2.95" that the authors use NIES L3 product. If so, then the product is already monthly-averaged. Please explain how the authors perform monthly-averaged calculations on the monthly L3 product.

L135-145: Since the XCO_2 variations are largely different over land and oceans, it is expected that semi-variograms may also differ. Did the author examine semi-variogram models in this aspect?

L136: Please clarify the method of downscaling from 2.5° degrees in NIES L3 product to 0.25° in this study.

L153-154: Monthly L3 data is used for the spatial interpolation with the EBK theory to derive the monthly XCO2 dataset (i.e. monthly to monthly). Are the resulted monthly XCO_2 time series in the 0.25° resolution actually anomalous among adjacent times? Then, it would be useful to include a sample figure to show how the time curve parameter library overcome this issue.

L155: What is "a time profile parameter library"? Is it a data matrix of parameters for Eq. (1) at each point or a kind of software library? Please clarify.

L156: Is "each point" in the sentence specifically the original L3 grid points (2.5° mesh) or the CDC dataset points (0.25° mesh)?

L160: The CO_2 trend is not perfectly linear for time scale of 2000-2020. Does the linear fit (a + b*t) show enough performance for this period? Have the authors checked the residuals from the fitted curves?

L160: Is the equation (1) correct? The 3rd and 5th terms, and the 4th and 6th terms are the same, respectively.

L163: I cannot follow the idea by the authors to use TCA theory to integrate spatial and temporal attributes. By using the EBK theory in Section 2.3.1, the authors already obtain monthly XCO_2 at 0.25° mesh based on monthly XCO_2 in L3 product. Then it would be possible that the author can simply calculate time series at each 0.25° mesh based on Eq. (1) to fill time-gap in the original L3 product though it has only a few missing values. Does TCA theory have some advantages?

L167: I cannot understand why only b and c in Eq. (1) are representative and considered here. Please clarify.

L177: Did the authors examine other method in transfer learning and compared the results?

L184: Does "the model predictions" mean the CDC dataset? What is "the input data"?

L185: The authors chose the coefficient of determination (R^2) and the root mean square error (RMSE) to evaluate the proposed algorithm to create the CDC dataset. However, the absolute biases of the dataset against the observation from TCCON or OCO-2 cannot be evaluated by R^2 and RMSE. For example, predicted XCO_2 in some TCCON sites like BU and HF in Figure 3 show apparent biases against TCCON observations. The bias calculation should be considered for the evaluation in addition to the two indicators.

L191: Are definitions of P and R for R^2 (Eq. (4))correct? Usually, they are opposite, i.e. P for observation and R for predicted values.

L194: There are several months with no data in Figures 7 to 17. Since the authors already made temporal interpolation in sections 2.3.2 and 2.3.3, I expected that the CDC results have no missing data in temporal direction. Please explain this point.

L195: Table 3 appears before Tables 1 and 2. Please renumber the tables.

L197: Please write more detail about the method of comparison. TCCON observations are continuous observation in time, while GOSAT observations are made at about 13:00 local time. Hence XCO_2 in GOSAT L2 and L3 products are largely represent XCO_2 in 13:00 local time. Which time span in a day of TCCON observation did the authors use to make monthly averages? Also, spatially, how was the CDC dataset at 0.25 mesh compared to XCO2 at TCCON sites (e.g. use spatially interpolated values into TCCON site locations or just use the mesh value at where TCCON sites are located).

L199: Why did the authors choose numbers 0.95 for R^2 and 1.5 for RMSE for criteria?

L200: The sentence says "24 TCCON sites", while Table 1 has 23 TCCON sites. "24" might be "23"?

L203: It would be helpful to see the dataset features if the authors could provide figures of time series of XCO_2 from the CDC dataset, GOSAT L3 product, and TCCON observations at each TCCON sites.

L204: Li et al. (2022) produce 8-day XCO_2 dataset from OCO-2 L2 product and Zeng et al. (2014) used ACOS-GOSAT Level 2 XCO2 product. GOSAT L3 product used in this study is already monthly-mean and smoothed onto 2.5° grid and based on bias-corrected L2 product by TCCON observations. Thus, it is not possible to judge the proposed method, i.e. the CDC dataset, is more refined than the previous work. The authors need to consider various factors to compare their results with the previous work.

L210: What is "the passive inversion mode"? Please describe the word.

L215-216: Please describe the processing method in detail: what "bad data" is, the "statistics" the author used, spatial resolution for monthly-average, time-span for monthly-average etc.

L216, 218: Previously, the authors use R^2, but here R is used for the evaluation. What is the reason for changing the evaluation metrics?

L220: It seems that the slopes of the fitted lines in orange in Figure 4 do not fit to centers of high-density areas in red. Are the figures okay?

L221-222: I think only through this comparison, the accuracy and stability of the CDC dataset may not be justified.

L223: It is not clear why the authors conclude the deviations is due to "revisit period" of the satellites. Please clarify.

L224-225: The authors mention about GOSAT-2 in this section. Did the authors use GOSAT-2 data in addition to GOSAT L3 product?

L225: There are no "official algorithms" of OCO-2 and GOSAT-2. What do the authors mean by that?

L230: How many datasets with removed data did the authors prepare for the evaluation?

L240: Same question in L194. Doesn't the time interpolation used in this study cover the no data period in the original L3 product?

L243: The labels are useful for dataset users, but the title is misleading because evaluation of uncertainty itself is not mentioned in this section. By using Kriging methods, Hammerling et al. (2012) and Watanabe et al. (2015) and others estimated data uncertainties in each grid. Can the authors also estimate uncertainty of the CDC dataset?

L247: L3 product is not observations, but processed data.

L247, 249: Did the authors use GOSAT-2 product as well?

L252: The sentence says "uncertainty", but "uncertain" in Table3 and the CDC dataset (h5 files) on the website.

L253: The information on GOSAT L3 is missing in this section.

L258-259: These lines are about latitude range which is already mentioned elsewhere. Why did the authors write the same sentence again in this section?

L261-262: The last sentence about the compressed code seems to be on the CDC dataset. Maybe better to move the sentence in L257.

L265-278: The sentences are almost the same as the abstract. Do the authors have any other topics that they contend lastly?

L279: Again, the same question. Did the author use GOSAT-2 data too?

L288-289: The author TS is missing?

L288(References): Please check the format of the references. Some are not in ESSD style.

L445(Table 2): Which is "Nums" in Table 2: the number of original OCO-2 data or the

8

number of monthly-averaged data?

L445(Table 2): Is it right that the evaluation method is R, not R^2?

L470(Table 3): "Data type" column seems not well organized because "Matrix" and "Float"/"Int" are in different categories. "Scalar" may be paired with "Matrix". Scalar and Matrix can have data in Float or Integer format.

L485(Figure 1): It would be useful to show TCCON sites that are used in this study in different color in the figure because the CDC dataset does not cover high latitudes and some sites were not used for the evaluation.

L505(Figure 2): What are "the official website" and "Shp" in Figure 2? Please clarify.