The presented study aimed to produce monthly sea surface pCO_2 maps for the South China Sea (SCS). Given SCS is a typical temperate/subtropical marginal sea, the pCO_2 sea surface maps for this waters is necessary for understanding the CO2 flux in temperate marginal sea and even global CO2 flux. From this perspective, the study and the data it present is very meaningful. However, the manuscript still have some major flaws which do not advise me to give a yes to publishing it in its current status.

[Response]: We appreciate that the reviewer valued our study. Our point-by-point responses are listed below.

Major comments

1. The manuscript was about a dataset generation, but from the abstract and the last section, what kind of data was used as input for the method was missing.

[Responds]: Accepted, and we will add the information of input data in our revision. Note that data input includes remote sensing derived data (sea surface salinity, sea surface temperature, chlorophyll), the spatial pattern of pCO_2 calculated by Empirical Orthogonal Function, atmospheric CO_2, and time labels (month).

2. As I understand EOF was an important part of the method used for pCO_2 maps generation, but in the entire section of methods, no paragraph or sentence was about EOF

[Responds]: The reviewer is correct that EOF was used to obtain the main spatiotemporal pattern of the RS derived pCO_2 and then as features in our reconstructed model. Following suggestions, we will add the information of EOF as follows "The EOF reflects the spatial commonality of variables shown in the time series, which is widely used to calculate spatial patterns of climate variability (e.g. Levitus et al., 2005; Dye et al., 2020; McMonigal and Larson, 2022). Typically, the spatial commonality of variables, also named EOF modes, are found by computing the eigenvalues and eigenvectors of a spatially weighted anomaly covariance matrix of a field. Each EOF mode'?s corresponding variance represents its degree of interpretation of spatial pattern of the variable.".

3. The language of the manuscript still need some efforts. The current version contains too many redundant phrases and sentences without clear meaning and very difficult to read through and get the logical flow. Readers expect concise and precise expression in an academic paper. and there are some grammar mistake and fuzzy expression.

[Responds]: We will pay special attention on the presentation during our revisions.

4. the range of legend in nearly all the map figures were too large and cannot show the spatial gradient of pCO_2 distribution, e.g, figure 6, 8, 11, 12,13.

[Responds]: Accepted. We will adjust the range of colorbar in figures as follows (Figure R1-R7).

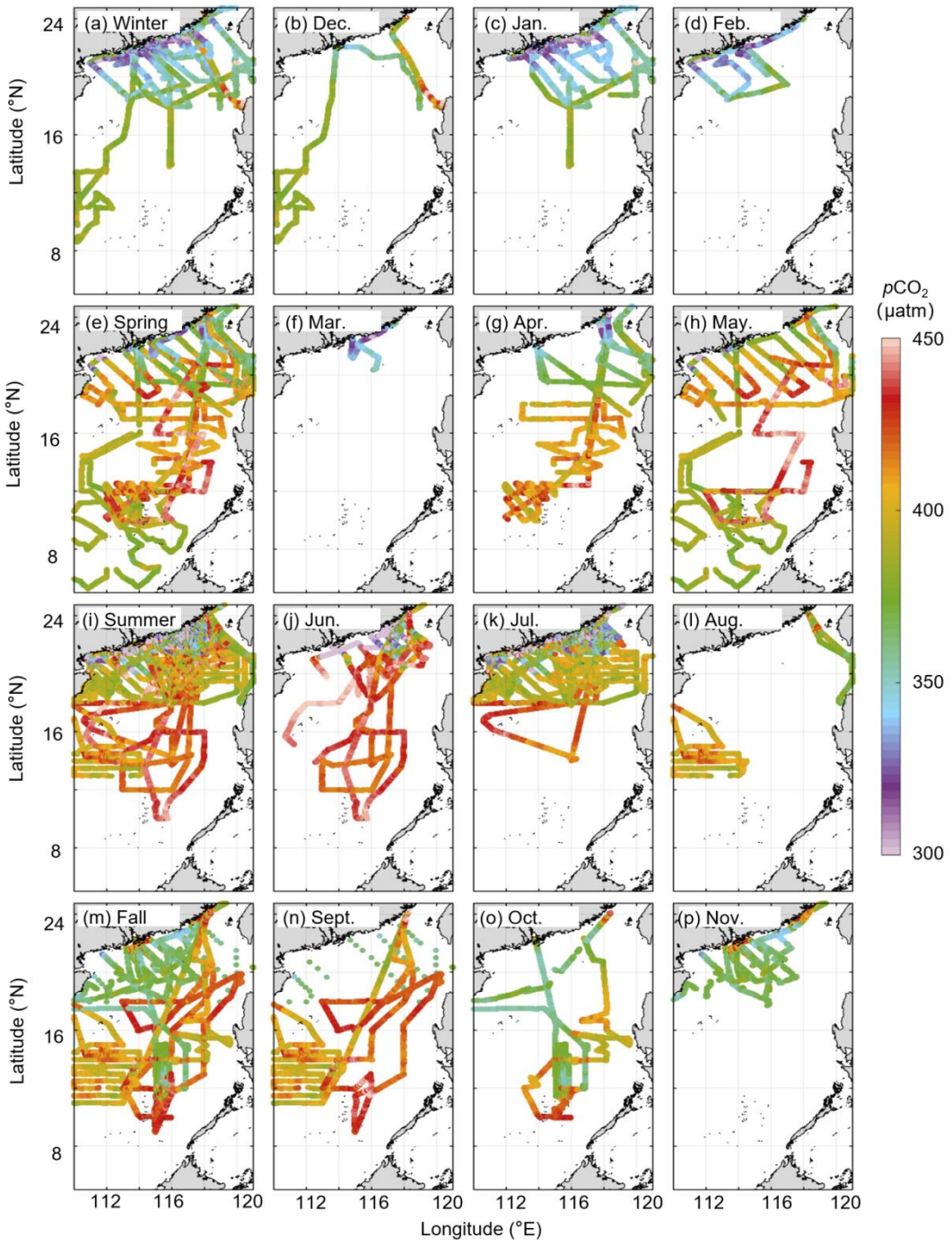

Figure R1. Seasonal and monthly sea surface pCO_2 fields in the South China Sea. The data sources can be found in Table 1 (a. winter; b. December; c. January; d. February; e. Spring; f. March; g. April; h. May; i. Summer; j. June; k. July; l. August; m. Fall; n. September; o. October; p. November).

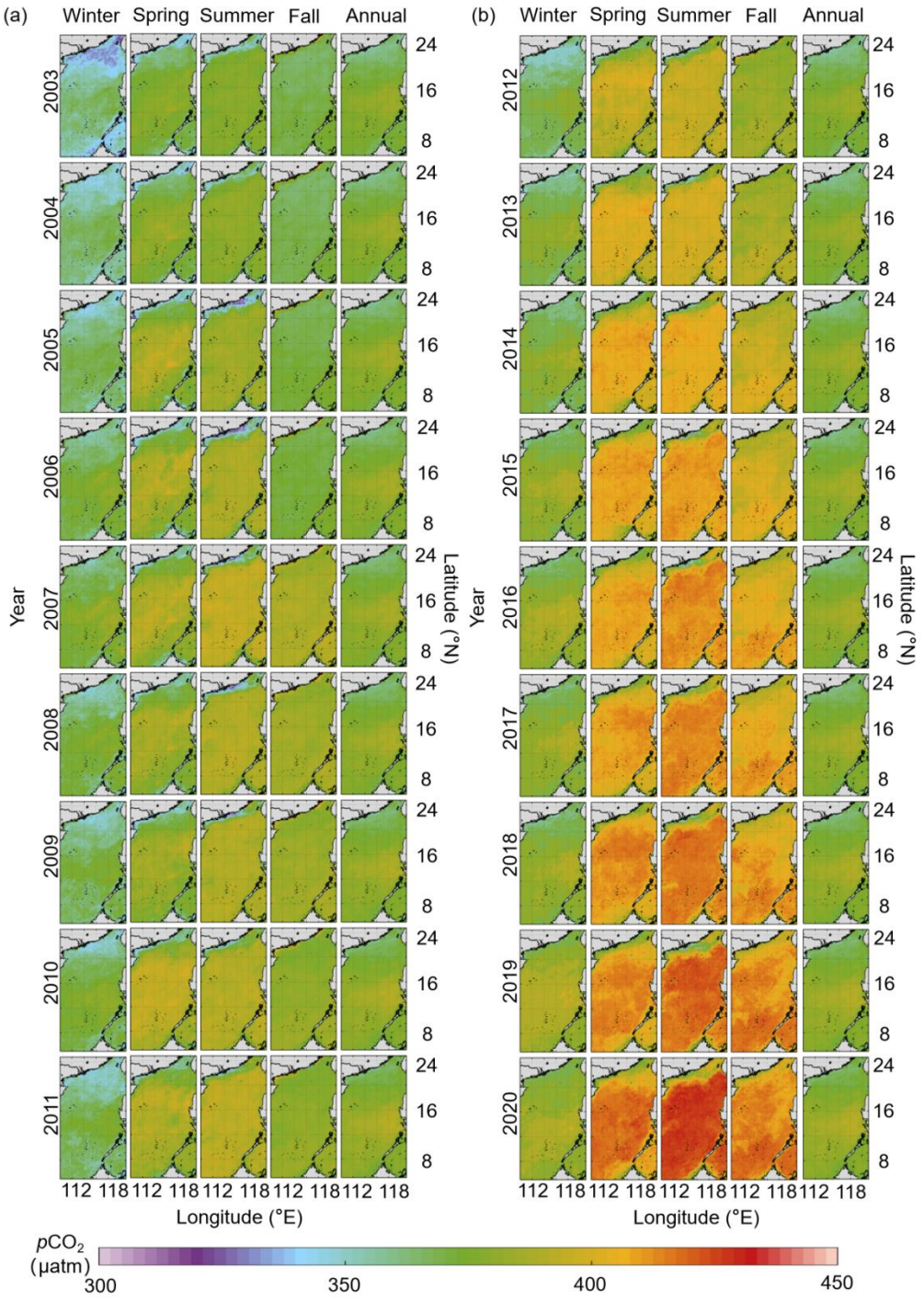

Figure R2. Reconstructed seasonal and annual sea surface pCO_2 fields in the South China Sea during the period 2003 to 2020 (a, 2003-2011; b, 2012-2020).

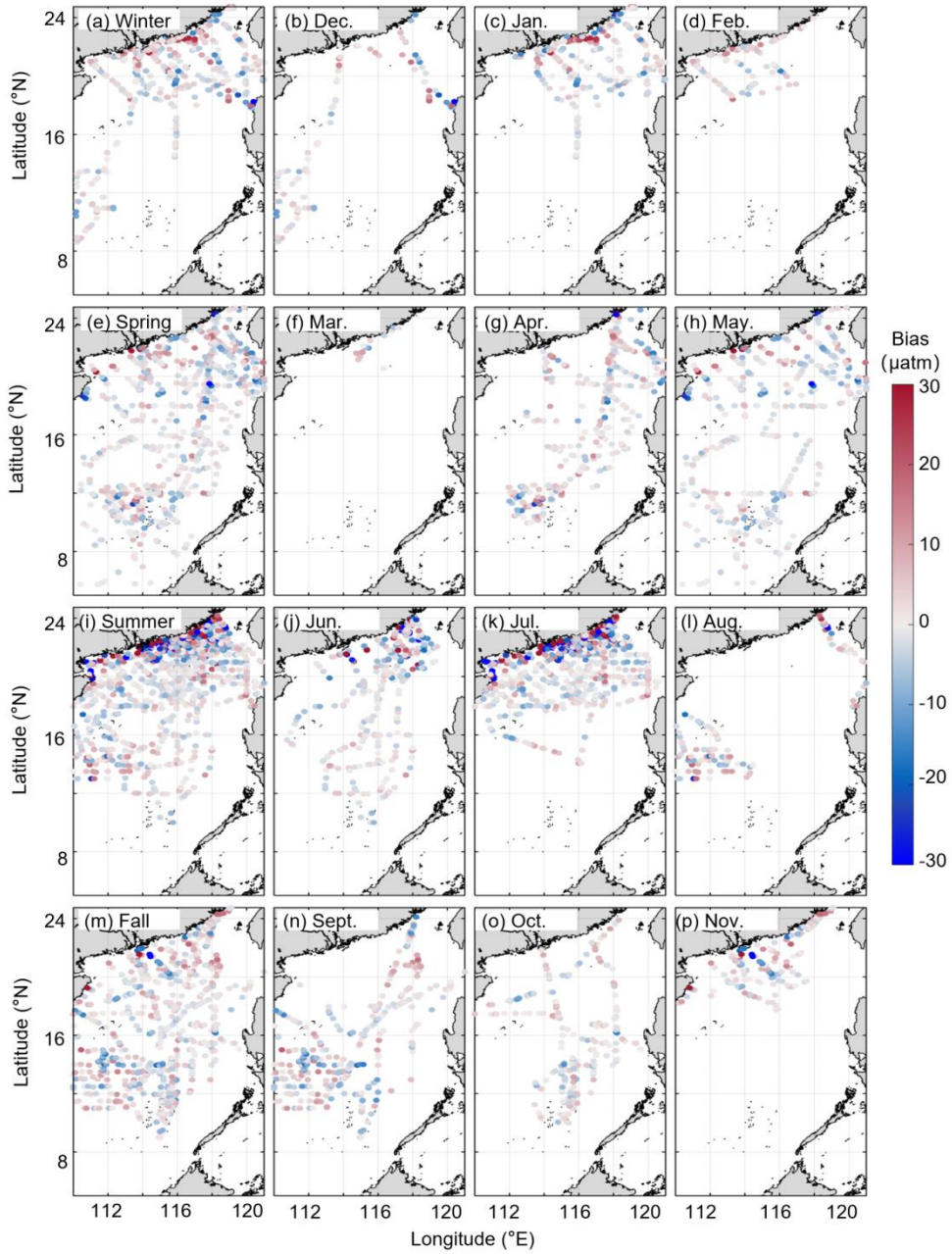

Figure R3. Differences between the seasonal and monthly reconstructed pCO_2 and the in situ pCO_2 data for the test set (a. winter; b. December; c. January; d. February; e. Spring; f. March; g. April; h. May; i. Summer; j. June; k. July; l. August; m. Fall; n. September; o. October; p. November). .

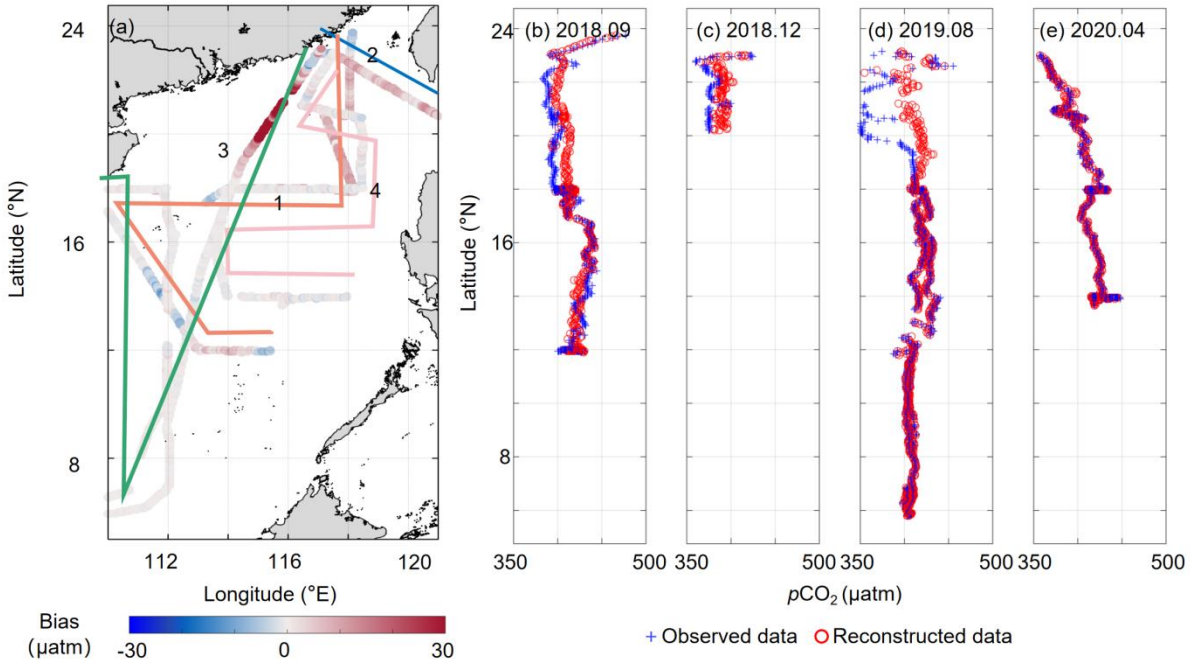

Figure R4. Difference between the reconstructed pCO_2 data and four independently in situ datasets during the four seasons. In (a), the numbers 1–4 represent September (2018.9, b), December 2018 (2018.12, c), August 2019 (2019.8, d), and April 2020 (2020.4, e), respectively.

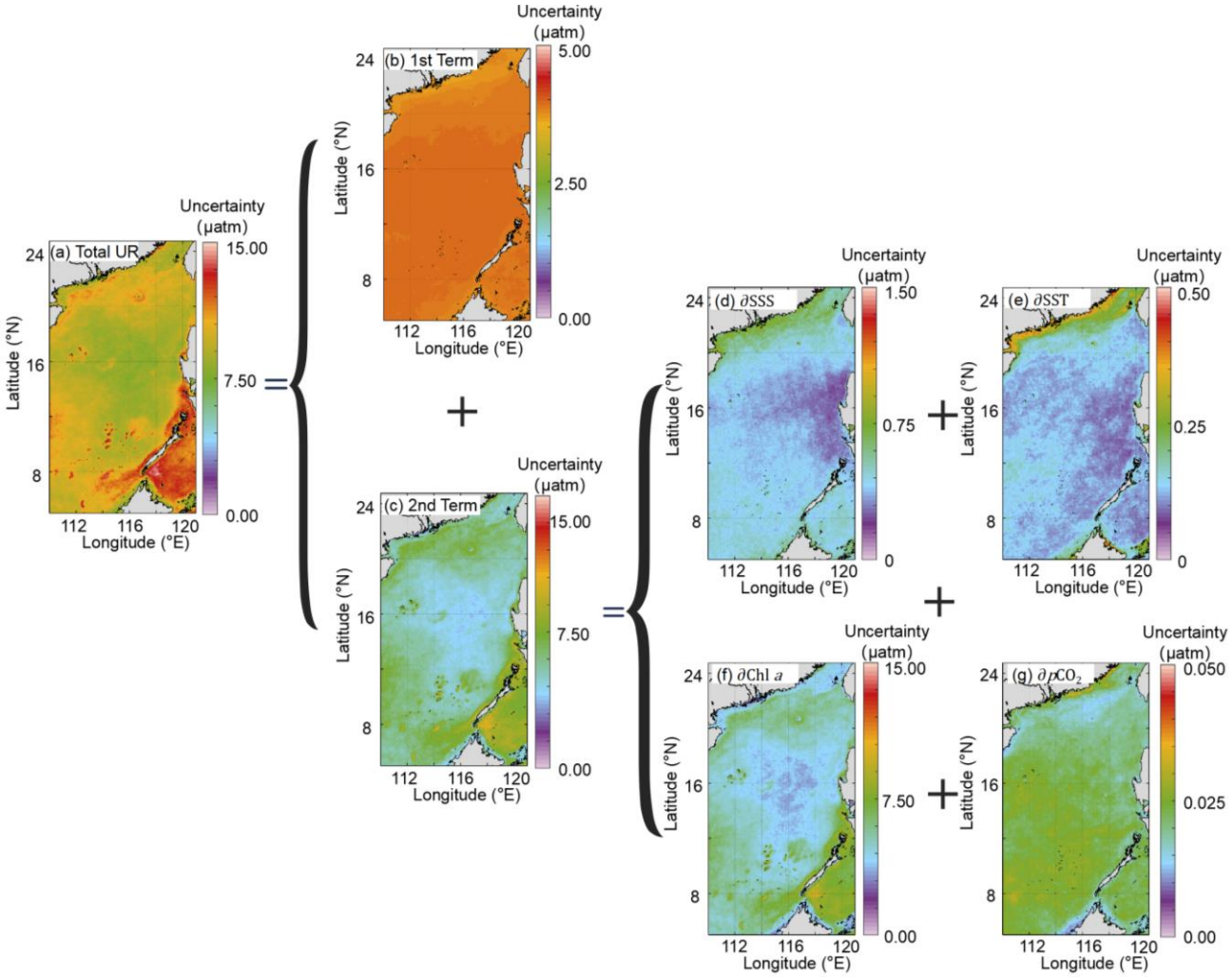

Figure R5. Uncertainties of the reconstructed pCO_2 fields (a, Total uncertainty in Equation 6; b. the first term of Equation 6; c. the second term of Equation 6; d stands for the $(\frac{\partial pCO2}{\partial SSS})dSSS$ in the the second term of Equation 6; e stands for the $(\frac{\partial pCO2}{\partial SST})dSST$ in the the second term of Equation 6; f stands for the $(\frac{\partial pCO2}{\partial Chl\ a})dChl\ a$ in the the second term of Equation 6; g stands for the $(\frac{\partial pCO2}{\partial RS\_derived\_pCO2})dRS\_derived\_pCO2$ in the the second term of Equation 6.

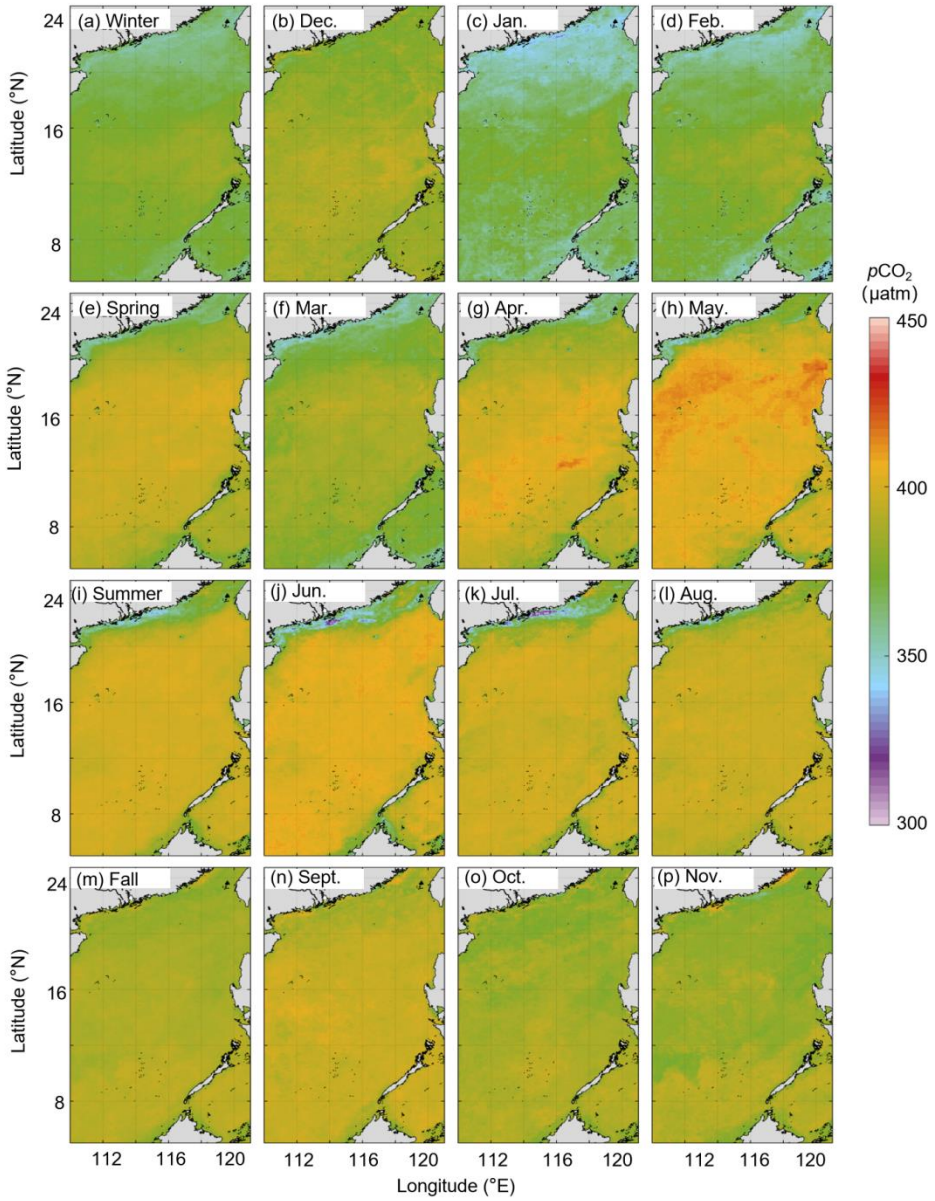

Figure R6. Long-term (2003–2020) seasonal and monthly average pCO_2 field (unit: µatm) (a. winter; b. December; c. January; d. February; e. Spring; f. March; g. April; h. May; i. Summer; j. June; k. July; l. August; m. Fall; n. September; o. October; p. November).

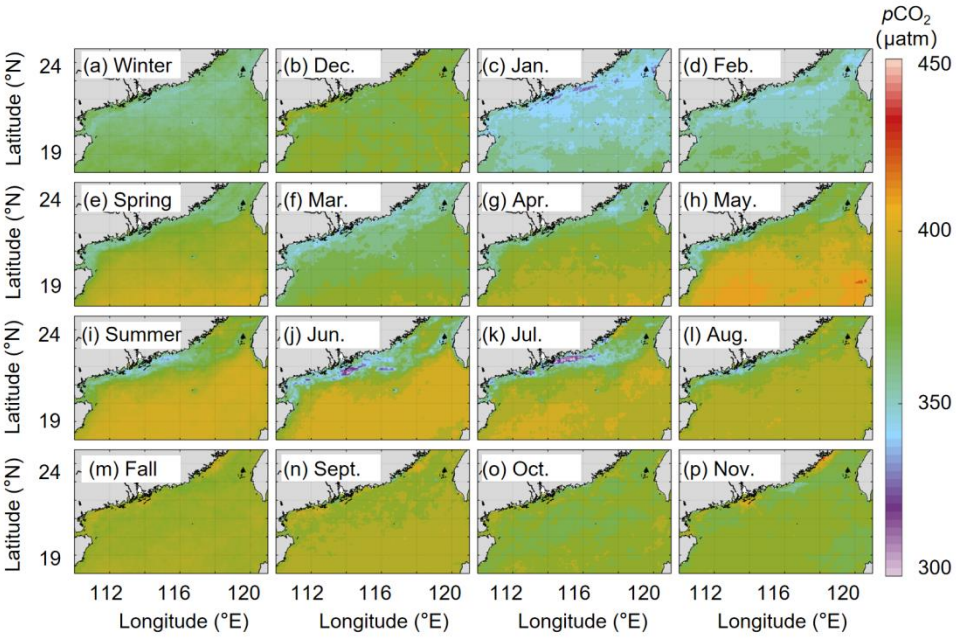

Figure R7. Long-term (2003–2020) seasonal and monthly averaged pCO_2 field in the region north of 18°N (unit: μatm) (a. winter; b. December; c. January; d. February; e. Spring; f. March; g. April; h. May; i. Summer; j. June; k. July; l. August; m. Fall; n. September; o. October; p. November).

5. what is the intention of including figure 4, if it is the quality of the remote sensing based pCO_2 maps included for further pCO_2 maps derivation, should the authors just need to include the information from the data distributor?

[Responds]: The reviewer is right that Figure 4 showed the quality of the RS-derived pCO_2 data. Following suggestions, we will remove this figure.

6. the study site section(2.1) should just serve the question "why mapping pCO_2 in SCS is important?", no other information is needed here.

[Responds]: We thank the reviewer for the comment. In the revision the importance of mapping pCO_2 in SCS will be added to section 2.1. The spatial distribution of pCO_2 is largely controlled by water mass missing and exchanges, thus, we retain in the introduction to the surface ocean circulation and water mass exchanges in the South China Sea in this section.

7. be consistent with the terminology, sometimes it is "in-situ", but "observational data" and "observed data" were present many times.

[Responds]: Accepted. We will unify the 'in-situ'/'observational data'/'observed data' to 'in situ data'.

8. in the abstract (line 12-14,), the importance of mapping pCO_2 in SCS should be addressed before presenting the method, generated data and its quality.

[Responds]: Accepted. Before presenting our method, we will add the following information "The South China Sea (SCS) is the largest marginal sea of the North Pacific Ocean, and mapping sea surface pCO_2 of this region is essential to better understand the spatiotemporal modes of CO_2 fluxes in marginal seas. In addition, we contend that the SCS is one of the most studied marginal seas in terms of carbon cycle in the world, which could thus be a model system for marginal sea carbon research" to show the importance of mapping pCO_2 in the SCS.

9. part of the input pCO_2 data of the presented study is from unpublished study (line 158), meaning not peer-reviewed.

[Responds]: We used two unpublished datasets in this paper. One of them is sea surface pCO_2 in China seas (0-42°N, 105-132°E) over 2003-2019 with a spatial resolution of 1 km and temporal resolution of a month (Bai et al., unpublished, line 158). This is the second version of pCO_2 in China seas. The first version was published on the SatCO2 website (http://www.satco2.com/index.php?m=content&c=index&a=show&catid=317&id=188) based on Bai et al. (2015). And this second version data can be cited as follows "Yu, S., Song, Z., Bai, Y., and He, X.: Remote Sensing based Sea Surface partial pressure of CO2 (pCO_2) in China Seas (2003-2019) (2.0). Zenodo, 2022. https://doi.org/10.5281/zenodo.7372479". Another dataset is the SSS data produced by 'Wang et al (in press)' in line 212. This paper has been accepted by Remote Sensing and its DOI number will be added in the revision as "Wang, Z., Wang, G., Guo, X., Hu, J., and Dai, M. Reconstruction of High-Resolution Sea Surface Salinity over 2003–2020 in the South China Sea Using the Machine Learning Algorithm LightGBM Model. Remote. Sens., 14, 6147, 2022. https://doi.org/10.3390/rs14236147 ".

Thus, in the revision, we will update the information accordingly.

10, line 308: Figure 7, validating the model output with the model training data gives no useful information, suggest removing this part

[Responds]: Accepted. We will only keep the results of test sets in Figure 7 as follows (Figure R8).

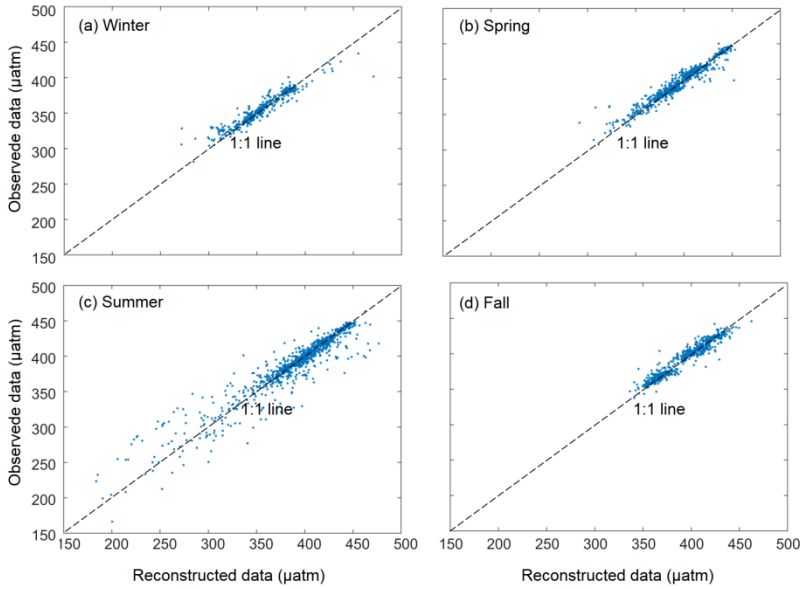

Figure R8. Comparison between the monthly reconstructed and the in situ pCO2 values for Tesing set (monthly results were overlaid to the four seasons: (a) Winter: Dec., Jan., Feb.; (b) Spring: Mar., Apr., May; (c) Summer: Jun., Jul., Aug.; (d) Fall: Sept., Oct., Nov.).

Minor comments

line 15-17"Using a machine learning-based method facilitated by empirical orthogonal function (EOF).... between 2003 and 2020" should specifically mention what kind of data was used for the methods input.

[Responds]: Accepted. Please refer to our response to Major Comment # 1 as of above.

linse 17- 20 "We validate our reconstruction with three independent testing datasets where,.... northern basin of the SCS." how independent are the three data set?

[Responds]: We validate our reconstruction with three independent testing datasets which are not involved model training. We will add this information in our revision.

line 22 "our reconstructions and observed data" grammar mistake.

[Responds]: Accepted. In the revision, we will rewrite this sentence as follows “The root-mean-square error (RMSE) between our reconstructed data and in situ data in TEST.1 averaged to ~10 μatm”

Line 27-28 "we present a new method to assess the uncertainty that includes the bias from the reconstruction and its sensitivity to the features,... quantifies the spatial distribution patterns of uncertainty." then the assessment method should be concisely introduced here. in addition, given this is a data presentation paper, the newly developed method should not in the highlight, unless it is a method presentation paper.

[Responds]: In the revision, we will rewrite this sentence as follows "we assess the uncertainty that includes the bias from the reconstruction and its sensitivity to the features.".

line 19 "that our reconstruction is effectively captures the main features of both the" ,check the grammar.

[Responds]: We apologize for the mistake. In the revision, we will rewrite this sentence as follows “our reconstruction effectively captures the main features of sea surface pCO_2 distributions in the SCS in both the spatial and temporal patterns”.

line 38,, "22–26%", I assume it should be 22%–26%.

[Responds]: The reviewer is correct, and we will make the crrection in the revision.

line 54-55: ":The former typically use statistical interpolations and regression methods" does not fit with the neighouring sentence, rewrite it or delete it.

[Responds]: Accepted. In the revision, we will delete this sentence.

line 61- 63 ,"However, because of the complex and dynamic nature of biogeochemical and physical processes in coastal areas, characterization of sea surface pCO_2 and subsequently the air-sea CO2 fluxes both in time and space in marginal seas remains challenging", this sentence is too strong and undermines the motivation of presented study, rewrite it,

[Responds]: Accepted. In the revision, we will rewrite this sentence as follows “Thus, machine learning has been widely used for reconstructing sea surface pCO_2 for the global ocean; however, it still remains challenging to extend this method to marginal seas”.

line 67: "clear need", what kind of need is clear need? a need can be strong, urgent, but not clear, need itself is a clear expression,

[Responds]: Accepted. In the revision, we will change this sentence to “Therefore, there is a strong need to achieve sea surface pCO_2 coverage in the SCS with a highest spatiotemporal resolution”.

line 73: "(sea surface temperature, SST; chlorophyll a, Chl a),", pay attention to journal requirements on abbreviation

[Responds]: Accepted. In the revision, we will rewrite this sentence as follows “Zhu et al. (2009) presented an empirical approach to estimate sea surface pCO_2 in the northern SCS in summer using satellite-derived data, including sea surface temperature (SST) and chlorophyll *a* (Chl *a*), ...”.

line 74: "underway "pay attention to the usage of underway, it is ambiguous in the manuscript.

[Responds]: Accepted. In the revision, we will change this to “in situ data” throughout the manuscript.

line 82, "the whole China Sea", where is the China Sea? do you mean all the seas in China's territory?

[Responds]: We referred to South China Sea, East China Sea, Yellow Sea, and Bohai Sea (99 - 122°E & 0 - 24°N). In the revision, we will add more details accordingly.

line 84: "(reported in Wang et al., 2021).", pay attention to the format of the reference citation
[Responds]: Accepted. We will make this correction.

line 84: "Bai et al. (unpublished) subsequently", if the work is not publised, then it should not be cited or discussed, as it is not peer-reviewed.
[Responds]: This dataset is an updated version based on Bai et al. (2015). Please refer to our response to Major Comment # 9 as of above. In the revision, we will update this citation to "Yu et al. (2022)".

line 94-96: include the input data here.
[Responds]: Accepted. In the revision, we will add some details of input data as follows "and selecting the remote sensing derived data (sea surface salinity, sea surface temperature, chlorophyll), the spatial patterns of pCO_2 calculated by Empirical Orthogonal Function, atmospheric CO_2, and time labels (month) as input data".

line 137-138 : there is no asterisk in the table and the meaning of the asterisk led note is not clear.
[Responds]: Accepted. In the revision, we will modify this Table as follows (Table R1).

Table R1. Summary of the seasonal in situ data of sea surface pCO_2 in the South China Sea for the period 2003-2020 used in this study.

Season	Spring			Summer		
	March	April	May	June	July	August
Cruise time	2004.03	2005.04 2008.04 2009.04 2012.04 2020.04*	2004.05 2011.05 2014.05 2020.05*	2006.06 2016.06 2017.06* 2019.06* 2020.06*	2004.07 2005.07 2007.07 2008.07 2009.07 2012.07 2015.07* 2019.07*	2007.08 2008.08 2019.08*
Season	Fall				Winter	
	September	October	November	December	January	February
Cruise time	2004.09 2007.09 2008.09 2020.09*	2003.10 2006.10	2006.11 2010.11	2006.12	2009.01 2010.01 2018.01	2004.02 2006.02
Data source	Li et al. (2020) *This study					

line 144 "Figure 3 shows the spatial and temporal distributions of surface water pCO_2.", the spatial distribution of in-situ measurements or data from other source?
[Responds]: Figure 3 shows the spatial distribution of in-situ measurements. In the revision, we will add more details as follows: "Figure 3 shows the spatial and temporal distributions of sea surface water pCO_2 of in situ measurements." .

line 157: how the remote sensing-derived pCO_2 data were derived? which methods, what is the quality? and output from unpublished study should not be used.

[Responds]: This dataset is an updated version based on Bai et al. (2015). Please refer to our response to Major Comment # 9 as of above. In the revision, we will change this citation to "Yu et al. (2022)", and will add more details of this dataset as follows "Yu et al. (2022) subsequently used a non-linear regression to develop a retrieval algorithm for seawater pCO_2 in the China Sea, and the satellite-derived pCO_2 data from 2003-2018 were provided by the $SatCO_2$ platform (www.SatCO2.com). In the retrieval algorithm of Yu et al. (2022), the input parameters are sea surface temperature, chlorophyll-a concentration, remote sensing reflectance of three bands (Rrs412, 443, 488 nm), the temperature anomaly in longitude direction, and the theoretical thermodynamic background pCO_2 under corresponding SST. Although the root mean squared errors (RMSE) for the pCO_2 were relatively large (21.1 μatm), it successfully showed the spatial distribution of the pCO_2 in China Seas (Yu et al., 2022)." .

line 184-187: "Wang et al. (in preparation) found a relatively high differential between the....observed data", meaning of this super long sentence is not clear.

[Responds]: Accepted. In the revision, we will modify this sentence as follows "For the sea surface salinity (SSS) data, Wang et al. (2022) found relatively large difference between the different open source SSS databases (i.e., multi-satellite fusion data from https://podaac.jpl.nasa.gov/; model data from https://climatedataguide.ucar.edu/; multidimensional covariance model data from https://resources.marine.copernicus.eu/) and the in situ SSS data." .

line 198 "pCO_2 filling method of", should explain the filling method here!

[Responds]: Accepted. In the revision, we will modify this sentence as follows "Secondly, we used the pCO_2 filling method according to Fay et al. (2021) to fill the missing pCO_2 measurements with the RS pCO_2 data, and this filling method can be found in section 3.1." because that the pCO_2 filling method would be explained in section 3.1.

line 201: "pCO_2 reconstruction model" pCO_2 reconstruction was used many times in the manuscript, but sea surface pCO_2 is not something one can reconstruct, it is a properties or variable of of the sea water, one can measure it ,describe it, retrieve its distribution, but not reconstruct pCO_2 itself. So, please pay attention to the verb usage.

[Responds]: Accepted. In the revision, we will change "pCO_2 reconstruction model" to "pCO_2 retrieve algorithm"

References

Dye, A. W., Rastogi, B., Clemesha, R. E. S., Kim, J. B., Samelson, R. M., Still, C. J., & Williams, A. P.: Spatial patterns and trends of summertime low cloudiness for the Pacific Northwest, 1996–2017. Geophysical Research Letters, 47, e2020GL088121, 2020. https://doi.org/10.1029/2020GL088121

Levitus, S., Antonov, J. I., Boyer, T. P., Garcia, H. E., and Locarnini, R. A.: EOF analysis of upper ocean heat content, 1956–2003, Geophys. Res. Lett., 32, L18607, 2005. doi:10.1029/2005GL023606.

McMonigal, K., & Larson, S. M.: ENSO explains the link between Indian Ocean dipole and Meridional Ocean heat transport. Geophysical Research Letters, 49, e2021GL095796, 2022. https://doi.org/10.1029/2021GL095796

Wang, Z., Wang, G., Guo, X., Hu, J., and Dai, M. Reconstruction of High-Resolution Sea Surface Salinity over 2003–2020 in the South China Sea Using the Machine Learning Algorithm LightGBM Model. Remote. Sens., 14, 6147, 2022. https://doi.org/10.3390/rs14236147.

Yu, S., Song, Z., Bai, Y., and He, X.: Remote Sensing based Sea Surface partial pressure of CO2 (pCO2) in China Seas (2003-2019) (2.0). Zenodo, 2022. https://doi.org/10.5281/zenodo.7372479.