Interactive comment on “Feasibility of reconstructing the basin–scale sea surface partial pressure of carbon dioxide from sparse in situ observations over the South China Sea” by Guizhi Wang et al.

Anonymous Referee #2
Received and published: 7 November 2020

The authors have used a remote-sensing-based pCO_2 field to derive EOFs, fit those EOFs to in situ pCO_2 observations collected over almost 2 decades, and then used the scaled EOFs to estimate the full surface pCO_2 record in the South China Sea.

It is a very interesting paper and many parts of it are clearly communicated, but it is also incomplete. The method validation and uncertainty quantification are missing. These should be an entire section of the paper and not just an added sentence or two, so the paper should be returned to the authors for major revisions.

Response: Thank you for bringing up the issues of validation and uncertainty quantification. We first addressed the cross validation issue. The regression-based reconstruction is often valid when outliers are not present. Our reconstruction follows this approach. Nonetheless, we agree with you and have conducted a cross-validation check of our reconstruction. The maximum RMSE of our cross-validation is 5.22 μatm, which occurred in 2006 when there were only 25 grid boxes with in situ pCO_2 data and which had the largest spatial standard deviation, 49.40 μatm, among the 13 years under consideration. This accuracy is very good compared to the spatial standard deviation of the in situ data in the same year. The temporal standard deviation of the reconstructed data is in the range of 2.12- 6.60 μatm. The cross-validation RMSEs are in the range of 2.43-5.22 μatm. We thus conclude that the reliability of our reconstruction is well supported by the cross-validation result. We will include our cross-validation method and result in the revised paper.

Second, we addressed the uncertainty issue. We made grid-by-grid comparisons between the observed pCO_2 and reconstructed pCO_2 in two ways. One is comparison with observed underway data (see Figure R1) and the other is comparison with pCO_2 calculated from observed DIC and total alkalinity around Station SEATS (18 °N, 116 °E) (see Figure R2). The RMSE between the reconstructed data and the observed underway data is in the range from 0.01-31.67 μatm (see Table R1). The difference between the reconstructed data and the observed data around Station SEATS ranges from -7 to 10 μatm with the relative error within 2.1%. Both comparisons will be provided as other ways of validation in the revised paper.

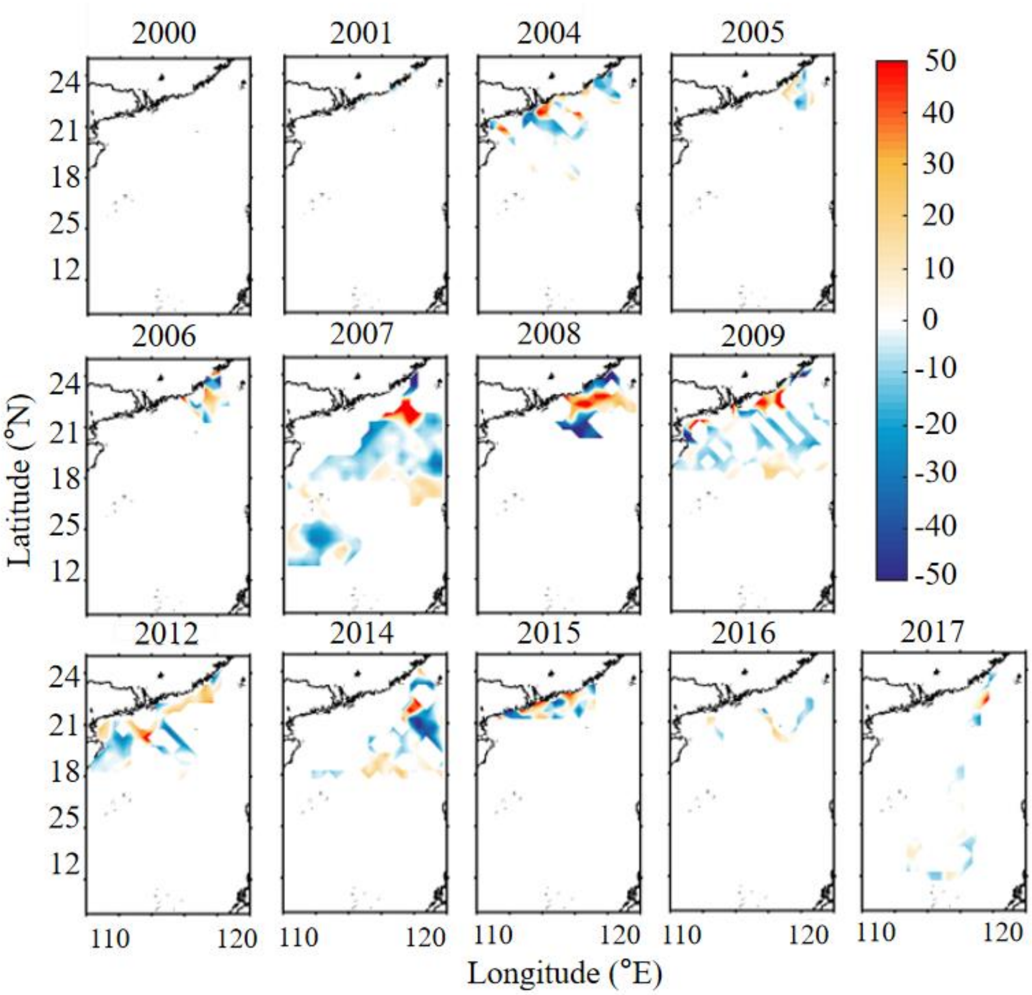

Figure R1: Difference between the reconstructed pCO_2 and the observed underway pCO_2 in years of 2000, 2001, 2004-2009, 2012, 2014-2017 (unit: μatm).

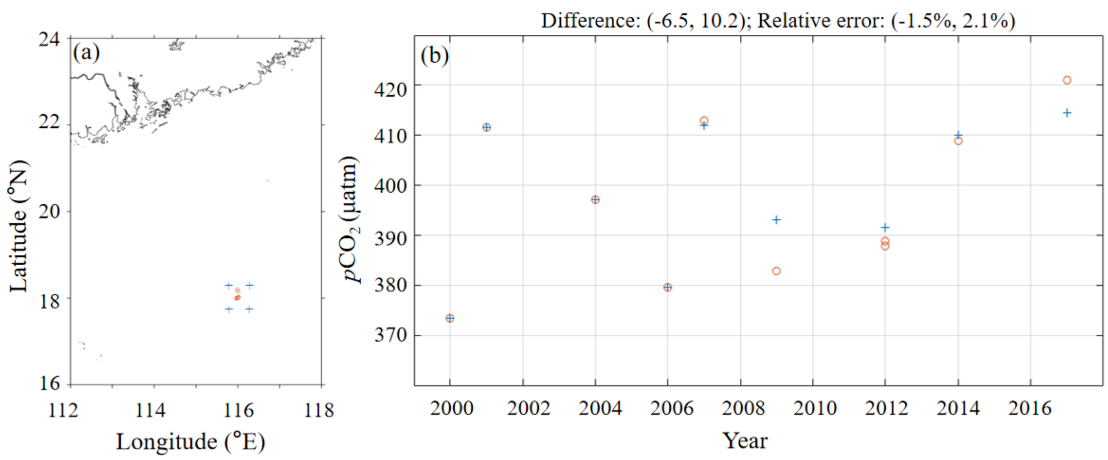

Figure R2: The comparison between the pCO_2 calculated from the observed total alkalinity and DIC and those from our reconstruction around Station SEATS (18 °N, 116 °E). (a) Locations of the observation stations and the grids on which the reconstructed pCO_2 was selected for comparison. The red circles are observation stations and the blue pluses represent the centers of the reconstruction grids, (b) The comparison between the pCO_2 data calculated from the observed DIC and total alkalinity and those from our reconstruction. The red circles represent the pCO_2 calculated from the observed total alkalinity and DIC and the blue pluses represent the reconstructed data. The difference is pCO_{2R} - pCO_{2C}, where pCO_{2R} is the reconstructed pCO_2 and pCO_{2C} is calculated from the observations, and the relative error is $(pCO_{2R}-pCO_{2C})/pCO_{2C}\times 100\%$.

Table R1: The RMSE between the reconstructed and the observed underway pCO_2 data ($RMSE_{RC}$) and between the remote-sensing derived pCO_2 estimates and the observed underway pCO_2 data ($RMSE_{RS}$)(unit: μatm).

Year	2000	2001	2004	2005	2006	2007	2008	2009	2012	2014	2015	2016	2017
$RMSE_{RC}$	0.01	7.27	19.72	16.28	31.67	16.50	26.14	20.41	15.48	18.82	27.83	13.04	12.76
$RMSE_{RS}$	12.82	20.15	47.94	65.69	88.97	25.12	43.79	36.80	30.73	24.20	NaN	NaN	NaN

It is a bit unclear whether this paper is presenting new data along with a new method or just a new method. There are two cruises in 2005 and 2006 with a reference given as "this paper" and, if these data sets are truly being published for the first time in this paper, then the manuscript should highlight that there are new data in the abstract. This would raise the value of this paper if there are indeed new data being made available along with the analysis. I might have missed the text that explained this.

Response: This paper is presenting new data along with a new method. In the revised data the new data in 2005 and 2006 will be highlighted in the abstract and in the main text. In addition, pCO_2 data from literature and calculated from our alkalinity and DIC around a basin station, SEATS (18 °N, 116 °E), will be presented to compare with the reconstructed pCO_2 as a way of validation in the revised paper.

The part of the paper that deals with the pCO_2 mapping approach is not yet complete because the authors have not assessed the uncertainties of their approach. I recommend one or two exercises. First, the approach should be repeated after removing some of the in situ pCO_2 measurements. Each cruise should be removed, one at a time. After removing a cruise, the analysis should be conducted using only the remaining data. Then the withheld cruise can be used to quantify how good of a job the mapping procedure does at reconstructing the withheld cruise. This should be repeated for every cruise in the dataset to get bulk statistics. If there is only one cruise worth of data in each year, then (I believe this reconstruction wouldn't work and instead) large swaths of latitude/longitude should be removed from the cruises and the remaining data should be used to reconstruct the missing data. This will allow the errors in the approach to be quantified. Second, if a model is available for the South China Sea that has pCO_2, then the model can also have the Bai et al. 2015 approach applied, be subsampled where the cruise measurements are, and be analyzed in the same way proposed here. This will reveal both the point-by-point reconstruction errors and allow the uncertainties for the overall pCO_2 average estimate, for example, to be quantified. Currently, the validation is left as an unsupported statement that the results look about right, which is insufficient for publication of a paper describing a quantitative method.

Response: Again, you have suggested a cross-validation procedure. As aforementioned, we have conducted a leave-one-out cross-validation study: Withholding a grid box datum, making the reconstruction using the remaining in situ data, and computing the difference between the withheld datum and the reconstructed datum at the same grid box. This is done for every grid box with in situ data for each year. The final cross-validation result is output as RMSE. The maximum RMSE is 5.22 μatm, which occurred in 2006, and the minimum is 2.43 μatm, which occurred in 2017. The year 2017 has 77 in situ data grid boxes. The spatial standard deviation of the data in 2017 is 17.55 μatm. Compared to the 2006 data described earlier, a more accurate reconstruction for 2017 is expected because of more grid boxes with in situ data and smaller spatial variability. This is supported by the cross-validation.

Your second suggestion can be mathematically proven, because the cross-validation RMSE of reconstruction from the sub-sample of the Bai et al. (2015) complete data is only the truncation error, which is equal to zero or very close to be zero. The reason is that the EOFs computed from Bai et al. (2015) data form a complete basis for the same data. Thus, the original data field can be exactly represented by a linear span of the EOFs.

There are other smaller problems that should also be addressed if a revised version of the paper is submitted:

1. The model should not be used in any region where there is no fitting data. This includes most of the South China Sea south of ~12.5 N.

Response: This is exactly the point that shows the power of the spectral optimal gridding (SOG) method using EOFs, in contrast to the traditional optimal interpolation method, such as kriging and inverse distance weighting. EOFs are a diagonalized representation of the covariance of climate dynamics, and thus are providing a consistency constraint of the $p\mathrm{CO_2}$ field. This allows us to use a small number of grid boxes with in situ data to interpolate and extrapolate to the entire region. Of course, the reconstruction accuracy is better when more observed data are available.

2. There should be an assessment of how good of a job the Bai et al. approach does at reproducing the in situ observations in a RMSE and bias sense. The estimates from this approach should be compared to the measurements from the data sets that are used here (and that Bai et al. did not use to design their routine). If the Bai et al. approach gives a different average $p\mathrm{CO_2}$ than the in situ measurements, then the climatology created from the remote sensing product should not be used to generate the Standardized Anomalies of Obs. Data (as indicated in Figure 1). I believe an independent climatology would then be needed. Otherwise, a significant average bias would have to be compensated by a large average value for one or more EOFs. In a best-case scenario, that would be EOF 1, but if, for example, the observations were mostly found in the dark blue patch of Figure 6c then the resulting reconstruction would be problematic. It seems likely that a large average value of EOF 3, which is highly variable spatially, would then be fit to the measurements to fix a homogenous bias between the in situ and remote-sensing records. This is just one example of the kinds of problems that could occur if the Bai reconstruction doesn't adequately resolve the mean or the variability. If nothing else, the Bai et al. validation should be discussed in this paper.

It would also be interesting to see how this approach compares to competing approaches, for example a neural network that relates the in situ $p\mathrm{CO_2}$ measurements to seawater property values that can be measured using remote sensing. This approach is more commonly used in global reconstructions. The Bai et al. approach is another clear competing approach.

Response: We did a grid-by-grid assessment of the remote-sensing derived $p\mathrm{CO_2}$ using the observed underway $p\mathrm{CO_2}$ as shown in Figure R3. In addition, a comparison between the observed

pCO_2 and reconstructed pCO_2 is provided in Figure R1. Both comparisons will be provided as validations in the revised paper. The RMSEs of the two comparisons are shown in Table R1 and will be provided in the revised paper. Furthermore, pCO_2 data from literature and calculated from our alkalinity and DIC around a basin station, SEATS (18 °N, 116 °E), were compared with the reconstructed pCO_2 as another way of validation as shown in Figure R2 and will be presented in the revised paper.

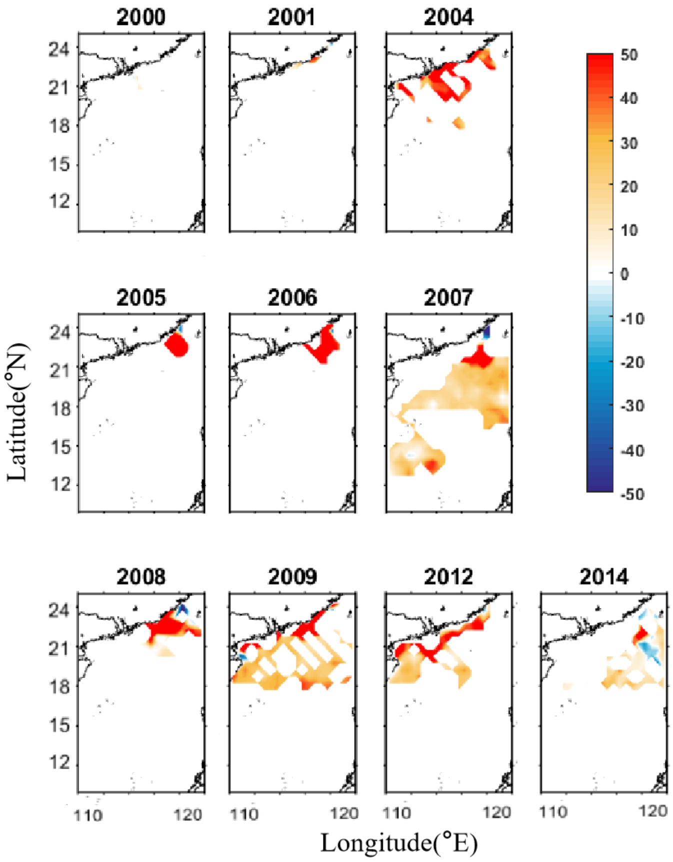

Figure R3: The difference between the remote-sensing derived pCO_2 estimates and the observed underway pCO_2 data in years 2000, 2001, 2004-2009, 2012 and 2014 (unit: μatm).

With regard to EOFs and independent climatology, the mathematical theory is like Fourier expansion of orthogonal polynomials, which can be sine functions, Legendre polynomials, and any set of eigenfunctions of a self-adjoint operator. Thus, EOFs form a complete basis for a data field although they may be different when using different anomalies computed from different climatologies and standard deviations. With different anomalies, variances may be re-distributed to different EOFs due to the different anomaly calculation methods. EOF rotation may help to reorganize certain variances into some specific EOF modes, and hence to provide an explanation of climate dynamics. However, this EOF rotation is not needed for the purpose of reconstruction as long as our EOFs form a complete basis. This completeness is guaranteed by the SVD algorithm for computing our EOFs here.

As for reconstruction using a neural network approach, the data produced by Jo et al. (2012) show an overall RMSE of 32.59-44.52 in summer pCO_2 reconstruction as validated using the observed underway data in the northern South China Sea, which overlaps with the RMSE of our reconstruction.

Specific comments:

15: consider deleting "capacity"

Response: The suggestion will be taken in the revised paper.

23: "The reconstructions always agree with observations." Delete or quantify this statement. The agreement is not absolute.

Response: This statement in the revised paper will be changed to "The RMSE between the reconstructed data and the observed underway data is in the range from 0.01-31.67 μatm and the difference between the reconstructed data and those calculated from observations around Station SEATS ranges from -7 to 10 μatm with the relative error within 2.1%, both of which indicate a good agreement of our reconstruction with observations."

28: The ocean

Response: The suggestion will be taken in the revised paper.

36: The sea-air CO_2 flux is the negative of the ocean carbon uptake, so this sentence is partially tautological.

Response: In the revised paper "helps quantify the oceanic carbon uptake capacity" will be deleted.

37: This sentence has several language errors. It also needs to be better-quantified or referenced. What is the decorrelation length scale for pCO_2 generally? How much of the ocean is constrained by those measurements alone without the newly proposed spatio-temporal mapping techniques? Mostly, I think a reference should be added to this sentence that supports this statement.

Response: The language errors of the sentence will be eliminated in the revised paper. The paper Bakker et al. (2016) will be added to the reference list. This paper is the most recent published compilation of measured pCO_2.

57: References needed for RS pCO_2 here.

Response: The suggestion is taken. Bai et al. (2015) will be added here in the revised paper.

Figure 1. What is meant by "standard deviations"? Standard deviations of grid values, or deviations of values within each grid cell?

Response: It is the temporal standard deviation of the RS pCO_2 values on each grid box.

Figure 2. Consider changing this map to a 2 dimensional histogram showing number summers with measurements (probably with colored bins).

Response: This map will be changed to the following Figure R4 in the revised paper.

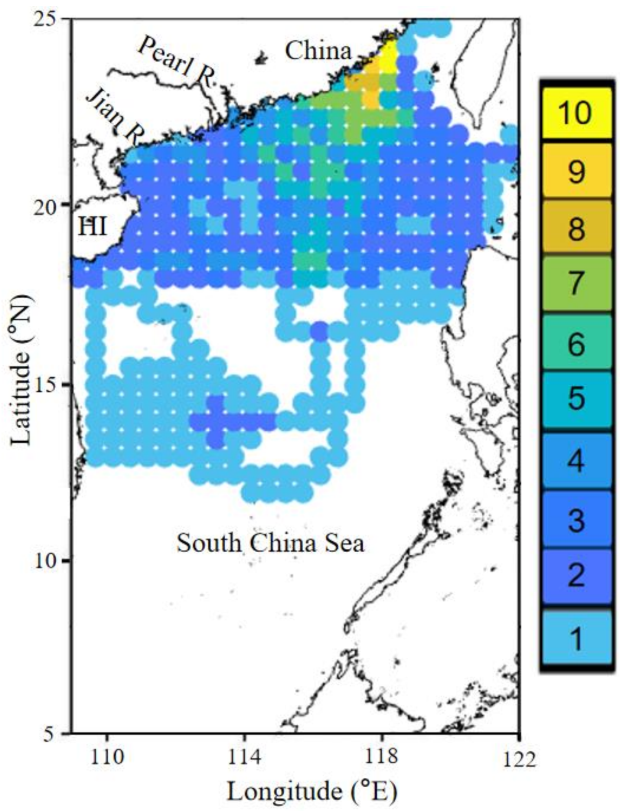

Figure R4: The number of summers with underway sea surface pCO_2 observations in the SCS in the period of 2000-2017. HI represents Hainan Island, Jian. R. is the Jianjiang River, and Pearl R. represents the Pearl River.

109: These estimates were... change "data" to "estimates" in this section since pCO_2 is not measured. In figure 1 as well.

Response: The changes will be made here and in Figure 1 in the revised paper.

138: Where is this symbol used?

Response: The symbol $\langle\cdot\rangle$ is only used once for expected value in Eq. (2) in the paper. In the revised version, we will replace the symbol $\langle\cdot\rangle$ by E[ ], which is more commonly used in statistics and science, while $\langle\cdot\rangle$ is a symbol commonly used in the field of theoretical physics.

141: how many EOFs were used? Say here.

Response: The suggestion will be taken in the revised paper. Eight EOFs were used.

182: This only shows the fields. One must compare this field to other figures to get an indication of how well the reconstruction performs. A plot showing differences between observed and reconstructed values is required.

Response: We have taken your suggestion and produced two figures showing differences between observed and reconstructed values. See Figures R1 and R2.

187: It is not enough to say "we fit the model to the data, so it fits the data." Statistics of goodness-of-fit should be presented. Furthermore, demonstrating that the method works requires withholding several cruises worth of data from the training data set and then using those cruises to verify that the method reconstructs the withheld data. Statistics and plots are required to quantify how well the reconstruction does.

Response: Again, this is a cross-validation issue discussed earlier. We will include our results of cross-validation and uncertainty quantification in the revised paper as shown in our response to the previous comments.

195: What is meant by reasonable?

Response: In the revised paper the RMSE, 7.27 μatm, will be provided here, which indicates that the reconstruction appears reasonable.

205: it is unclear what is meant by “the large spatial gradient of in situ data.”

Response: Here it means the large spatial variation of in situ data. The “gradient” is changed to “variation” in the revision.

214: 2.383 is given to excessive precision. An attempt should be made to quantify the uncertainty and the data should be reported to the appropriate precision.

Response: The suggestion is taken. The rate will be given as 2.4±0.8 μatm/yr in the revised paper.

229: why would a higher rate be expected in a marginal sea? I would argue that 2.4 uatm/year is completely within expectations of the atmospheric pCO_2 trend over this time period given the large uncertainties in this approach and the likely natural variability in surface ocean pCO_2.

Response: Considering the uncertainty in the rate is 0.8 μatm/yr, we agree with the reviewer that our rate is consistent with the trend shown at Station HOT in the Pacific. In the revised paper, this statement will be included and the sentence about “a higher rate be expected in a marginal sea” will be deleted.

References

Bai, Y., Cai, W.-J., He, X., Zhai, W. D., Pan, D., Dai, M., and Yu, P.: A mechanistic semi-analytical method for remotely sensing sea surface pCO2 in river-dominated coastal oceans: a case study from the East China Sea, J. Geophys. Res. 120, 2331–2349, doi:10.1002/2014JC010632, 2015.

Bakker, D. C. E., Pfeil, B. C., Landa, S., Metzl, N., M. O'Brien, K., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W. J., Castle, R. D., Chen, L. Q., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C., Huss, B., Ibanhez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Kortzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschutzer, P., Lauvset, S. K., Lefevre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., R. Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S., Vandemark, D., Ward, B., Watson, A. J., and Xu,

S. Q.: A multi-decade record of high-quality fCO_2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT), Earth System Science Data, 8, 383–413, doi:10.5194/essd-8-383-2016, 2016.

Jo, Y. H., Dai, M. H., Zhai, W. D., Yan, X. H., and Shang, S. L.: On the variations of sea surface pCO_2 in the northern South China Sea: A remote sensing based neural network approach, J. Geophys. Res., 117, C08022, doi:10.1029/2011JC007745, 2012.