Response to Referee's Comments

Referee 1: The authors use a neural network model to generate a $p\mathrm{CO_2}$ product for the global ocean using the SOCAT data, and combine these $p\mathrm{CO_2}$ estimates with a wind speed product to compute the $\mathrm{CO_2}$ flux. The ensemble model results compare well overall to the observations, and the carbon flux estimates are in-line with the literature. My main comments concern how novel these results are compared to the extensive literature on the topic, and the interpretation of some of the model statistics.

Authors: We would like to thank Referee 1 for constructive comments and suggestions on our study. Each point will be addressed in the following and the manuscript has been revised on this base. Throughout this document, the referee's comments are in bold.

1 There is a lot of previous literature using spatially and temporally sparse observations of surface $p\mathrm{CO_2}$ to generate global data products and provide estimates of ocean carbon uptake, some of which use very similar methods to those in this present manuscript. The authors cite this previous literature, but there's very little discussion of it. Consequently, I found it difficult to interpret how the present authors' methods and results are novel and differed from these previous studies. The motivation appears to be in lines 41-46, however, I don't follow how the previous literature did not incorporate "space-time varying uncertainty estimates"? It would also appear that the incorporation of the coasts is relatively new, though the authors then cite a few recent studies and declare that it's a closed gap? I suggest that the introduction needs to contain a much clearer description of how the methods used here compare to previous studies, and what is new about this analysis.

This study is made up of our efforts to reproduce and intensively analyse the spatially and temporally varying surface $p\mathrm{CO_2}$ fields, the air–sea $\mathrm{CO_2}$ fluxes, and their reconstruction uncertainties over the global ocean. We acknowledge previous studies pursuing the same target and are aware that the existing observation-based mapping methods (for instance proposed by Rödenbeck et al., 2013; Landschützer et al., 2016; Denvil-Sommer et al., 2019; Gregor et al., 2019; Watson et al., 2020) succeeded in obtaining a relatively low misfit between the reconstructed and observational data (see in Rödenbeck et al., 2015; Gregor et al., 2019; Friedlingstein et al., 2020). Although similar interpolation and machine learning approaches (e.g., clustering, classical regression, neural networks) and/or similar sets of predictor variables for $p\mathrm{CO_2}$ have been considered in the preceding literature, model design and implementation are still different (e.g., proportion of SOCAT data used in model fitting and evaluation). The present manuscript reflects our vision on the following key features.

i. <u>A design of an ensemble of numerous feed forward neural network (FFNN) models:</u>

   It is based on a Monte Carlo approach wherein each model is trained and validated on sub-samples randomly drawn from the monthly gridded SOCATv2020 data and available data of predictors. The ensemble size of 100 is considered in this study. Our proposed ensemble approach was developed at the Laboratoire des Sciences du Climat et de l'Environnement (LSCE) as both an extension and an improvement of the first version (LSCE-FFNN-v1, Denvil-Sommer et al., 2019). Quality assessments comparing these two model versions are documented in Chau et al. (2020). Besides, the proposed approach inherits strengths of the existing statistical models and further aims at reducing mapping uncertainties induced by, for instance, discrete boundaries in the two-step clustering-regression by Landschützer et al. (2016); Gregor et al. (2019) or the two-step FFNN-based reconstruction of $p\mathrm{CO_2}$ climatologies and anomalies by Denvil-Sommer et al. (2019). As described in the Method section (2.2) in the manuscript, each FFNN model follows a leave-p-out cross-validation approach, i.e., the exclusion of p gridded SOCAT data of the reconstructed month itself in model training and validation. This allows to reduce model over-fitting. In addition, it leaves more independent data for evaluation than previous approaches, results obtained by the proposed reconstruction are in line with the others though (see e.g., Friedlingstein et al., 2020, and references therein).

ii. <u>Quantification and evaluation of model best estimates (ensemble means) and uncertainties (ensemble spreads):</u>

There exists other ensemble-based methods, their concepts and principle objectives are nevertheless different. For example, Gregor et al. (2019) and Gregor and Gruber (2021) introduce machine-learning ensembles with a small ensemble size of different two-step clustering-regression models mapping surface $p\mathrm{CO}_2$ and propose the ensemble mean as their model best estimate. In a broader context, Rödenbeck et al. (2015) suggest an intercomparison of multiple mapping methods targeting the identification of common or distinguishable features of different mapping results. Hauck et al. (2020) and Friedlingstein et al. (2020) synthesize $p\mathrm{CO}_2$ mapping products and refer to an ensemble of their observation-based estimates of air-sea CO_2 fluxes as a benchmark to compare with the one derived from ocean biogeochemical models.

Up to recently, most of these studies have used misfits between the reconstructed and observational data (e.g., the root-mean-square deviation, RMSD) to evaluate product quality and infer uncertainty estimates of the reconstructed $p\mathrm{CO}_2$. Reconstruction errors of $p\mathrm{CO}_2$ are then propagated to get uncertainty estimates of the reconstructed CO_2 fluxes (see in Landschützer et al., 2014, for instance). By construction, such uncertainty estimates are restricted to oceanic regions and periods when observations are available (Lebehot et al., 2019; Hauck et al., 2020), and the uncertainty quantification of an averaged $p\mathrm{CO}_2$ or an integrated flux is under low confidence due to sparse data density. An advantage of our approach is that an ensemble of 100 model outputs of $p\mathrm{CO}_2$ and CO_2 fluxes is available at each $1^\circ \times 1^\circ$ ocean grid cell of the globe for each month in the period 1985–2019. The ensemble asset facilitates the quantification of model uncertainty of $p\mathrm{CO}_2$ and CO_2 fluxes averaged or integrated over space and time of interest (see for instance Figures 5 and 9 in the manuscript). This is expected to provide more robust estimates than the ones based on reconstruction errors.

iii. <u>Seamless analysis of the reconstructed data and uncertainty estimates over the open ocean and coastal zones</u>:
An in-depth analysis has been made and presented for both the open and coastal regions divided by latitude bands. Interpretations of good or poor reconstructions of surface $p\mathrm{CO}_2$ and air-sea CO_2 fluxes (e.g., data density and distribution, regional to local characteristics of $p\mathrm{CO}_2$ and its potential drivers, model design and resolution) and changes in spatial and seasonal variations of CO_2 fluxes are given. To strengthen our interpretation, we have shown both the temporal and spatial distribution of the reconstructed $p\mathrm{CO}_2$ and CO_2 fluxes fields, model-data misfits, model uncertainty, and linked these materials with their driving mechanisms suggested in previous literature. More importantly, we have made an intercomparison of model reconstruction ability between regions, identified oceanic sectors where the model does not fit the data well, and suggested further improvements on the data reconstruction based on the proposed space-time varying uncertainty fields.

With these points involved in the manuscript, we believe that our study is novel and statistics and keys findings therein would be useful contributions for the marine science community. However, we agree with the referee that the first version of the manuscript missed part of discussions on the comparison among the existing methods, and thus the novelty of this study was not bold. We consider this referee's feedback important and it has been taken into account in our revision. Precisely, we have reworked on the last two paragraphs in the Introduction section (Lines 37-55 of the manuscript). The two new paragraphs are reproduced below.

- <u>Previous text</u>: *Various data-based approaches including machine learning, classical regression, and mixed layer schemes have been proposed in Rödenbeck et al. (2013); Landschützer et al. (2016); Denvil-Sommer et al. (2019); Gregor et al. (2019); Watson et al. (2020) (see also other mapping methods in Rödenbeck et al., 2015) to infer gridded maps of surface ocean* $p\mathrm{CO}_2$ *from the sparse set of observations, targeting the improved reconstruction of spatially and temporally varying surface* $p\mathrm{CO}_2$ *fields and air–sea* CO_2 *fluxes over the global ocean and major basins. While these studies provide model bias and standard errors and use these statistics as model uncertainty estimates, none has so far analysed space-time varying uncertainty estimates, e.g., based on the model dispersion of a large set of realizations of* $p\mathrm{CO}_2$ *and air–sea flux estimates. Moreover, up to recently, most of these reconstructions did not cover the coastal ocean, a gap that has been closed by validated estimates of mean climatologies of* $p\mathrm{CO}_2$ *(Laruelle et al., 2017; Landschützer et al., 2020), air–sea flux density and the total coastal C sink (Laruelle et al., 2014).*

  In this work, we first propose a new inference strategy for this problem based on an ensemble of 100 neural network models mapping the monthly gridded SOCATv2020 data. The approach consists in reconstructing the monthly $p\mathrm{CO}_2$

fields and the contemporary air–sea fluxes over the period 1985–2019 on a spatial resolution of $1^\circ \times 1^\circ$. *Mean and standard deviation are computed from the ensembles of 100 model outputs. They are used to estimate the mean state and uncertainty of the carbon fields seamlessly for different time scales (e.g., monthly, yearly, and multi-decadal) and spatial scales (e.g., grid cells, sub-basins, and the global ocean). Based on the uncertainty estimates, we identify regions that should be prioritized in future observational programs and model development in order to reduce model errors and uncertainty. Potential drivers of the spatio-temporal distribution and the magnitude of air–sea* $\mathrm{CO_2}$ *fluxes are discussed with the aim to better attribute underlying processes and detect potential focus regions for further studies on the evolution of oceanic* $\mathrm{CO_2}$ *sources and sinks.*

- Revised text:

Various data-based approaches have been proposed to infer gridded maps of surface ocean $\mathrm{pCO_2}$ *from the sparse set of observational data. They have been successful in obtaining similarly low misfits between the reconstructed and evaluation data and reasonable estimates of air-sea* $\mathrm{CO_2}$ *fluxes (see in Rödenbeck et al., 2015; Gregor et al., 2019; Friedlingstein et al., 2020) although model design and implementation are quite different (e.g., proportion of SOCAT data used in model fitting and evaluation). Aside from data reconstruction built on a single model mapping* $\mathrm{pCO_2}$ *data with machine learning, classical regression, or mixed layer schemes (see Rödenbeck et al., 2013; Landschützer et al., 2016, for a few), ensemble-based approaches have recently emerged but with their own concepts and objectives. For example, Denvil-Sommer et al. (2019) designed a two-step reconstruction of* $\mathrm{pCO_2}$ *climatologies and anomalies based on five neural network models and selected the one that reproduced the* $\mathrm{pCO_2}$ *field with the smallest model-data misfit. Gregor et al. (2019) and Gregor and Gruber (2021) introduced machine-learning ensembles with six to sixteen different two-step clustering-regression models mapping surface* $\mathrm{pCO_2}$ *and suggest that the use of their ensemble mean is better than each member estimate. In a broader context, Rödenbeck et al. (2015) presented an intercomparison of fourteen mapping methods targeting the identification of common or distinguishable features of different products in long-term mean, regional and temporal variations. Hauck et al. (2020) and Friedlingstein et al. (2020) also synthesized* $\mathrm{pCO_2}$ *mapping products and take an ensemble of their observation-based estimates of air-sea* $\mathrm{CO_2}$ *fluxes as a benchmark to compare with the one derived from ocean biogeochemical models. Despite positive conclusions overall, statistical data reconstructions are still subject to further improvements. In Rödenbeck et al. (2015), Hauck et al. (2020), Bushinsky et al. (2019), and Denvil-Sommer et al. (2021), the authors explain that substantial extensions of surface ocean observational network systems are essential to better determine* $\mathrm{pCO_2}$ *and fluxes at finer scales and reduce mapping uncertainties. So far mapping uncertainties have been estimated by using misfits between the model outputs and SOCAT data (e.g., the root-mean-square deviation, RMSD). By construction, such uncertainty estimates are restricted to oceanic regions and periods when observations are available (Rödenbeck et al., 2015; Lebehot et al., 2019; Gregor et al., 2019) and the uncertainty quantification of an averaged* $\mathrm{pCO_2}$ *or an integrated flux over space and time of interest is under low confidence due to sparse data density. Furthermore, most of the previous mapping methods target* $\mathrm{pCO_2}$ *data and evaluate their estimates solely over the open ocean, with the coastal data excluded or not fully qualified. In Laruelle et al. (2014, 2017) the authors present spatial distribution of air–sea flux density and estimates of total coastal C sink while a recent study (Landschützer et al., 2020) limits their estimation to monthly climatologies of* $\mathrm{pCO_2}$ *over the global ocean including the coastal regions.*

In this work, we propose a new inference strategy for reconstructing the monthly $\mathrm{pCO_2}$ *fields and the contemporary air–sea fluxes over the period 1985–2019 with a spatial resolution of* $1^\circ \times 1^\circ$. *It is based on a Monte Carlo approach, an ensemble of 100 neural network models mapping sub-samples drawn from the monthly gridded SOCATv2020 data and available data of predictors. This ensemble approach was developed at the Laboratoire des Sciences du Climat et de l'Environnement (LSCE) as both an extension and an improvement of the first version (LSCE-FFNN-v1, Denvil-Sommer et al., 2019). In the following sections, we first present the ensemble of neural networks designed with the aim of leaving aside the issue of discrete boundaries in the existing two-step clustering-regressions (see further discussion in Gregor and Gruber, 2021) and reducing the mapping uncertainties induced by the two-step reconstruction of the* $\mathrm{pCO_2}$ *fields (Denvil-Sommer et al., 2019) or by an ensemble-based reconstruction with a small ensemble size. In addition, each FFNN model follows a leave-*p*-out cross-validation approach, i.e., the exclusion of* p *gridded SOCAT data of the* *reconstructed month itself in model training and validation. This allows to reduce model over-fitting and to leave much more independent data for model evaluation than the previous studies. Mean and standard deviation computed from the ensemble of 100 model outputs are defined as estimates of the mean state and uncertainty of the carbon fields. As one of the novel key findings of this study compared to the existing ones, we compute and analyze the estimates of* $p\mathrm{CO_2}$ *and air–sea fluxes, model errors, and model uncertainties for different time scales (e.g., monthly, yearly, and multi-decadal) and spatial scales (e.g., grid cells, sub-basins, and the global ocean). We then suggest the use of an indicator map built on the space-time varying uncertainty fields instead of model-data misfits for identifying regions that should be prioritized for future observational programs and model development in order to improve the data reconstruction. Last but not least, the model best estimates and uncertainty of* $p\mathrm{CO_2}$ *and air–sea fluxes are analysed seamlessly over the open ocean to the coastal zone. Potential drivers of the spatio-temporal distribution and the magnitude of open ocean and coastal* $\mathrm{CO_2}$ *fluxes are discussed with the aim to better identify underlying processes and to detect potential focus regions for further studies on the evolution of oceanic* $\mathrm{CO_2}$ *sources and sinks.*

2 I think the methods section is missing a few key details that will help support this manuscript.

- **First, it would be helpful for the authors to explain how to interpret and compare the RMSD and r^2 values for each region. The reason being, that these values are listed for each ocean region, but it's a little unclear what differences in these values between regions is saying about the model estimate. For example, I was surprised by how low the RMSD value for the Southern Ocean is (slightly lower than the global mean), despite the somewhat limited observational data and well documented, substantial inter-annual variability. However, the Southern Ocean does have a lower r^2 value, which the authors seem to attach a greater weight to in their interpretation.**

We have revised the manuscript and will add in Section Methods details of the statistics used in this study to facilitate the readers' interpretation of our results. In general, RMSD measures the model skill in terms of mean distance between model estimates and evaluation data while r^2 measures the proportion of data variation predicted by the model. RMSDs between the model and SOCAT gridded data over the Southern Ocean (open: 19.18 μatm, coastal: 35.73 μatm) are slightly higher [lower] than the global errors (open: 17.87 μatm, coastal: 35.86 μatm) for the open ocean [coastal zone], but the regional r^2 values (open: 0.62, coastal: 0.65) are lower than the global ones (open: 0.78, coastal: 0.70). The global scores involve the ones of all the regions, where the poorest reconstruction were found over the Arctic, subpolar, and coastal regions (Figures 3c, 5e, S2, and Table S2). Compared to other metrics such as model bias and r^2, RMSD takes another role as an outlier detector of model-data misfits which gives larger weights to such high errors over these regions. Yet, data sampling is limited over the Southern Ocean but similar circumstances appear over polar/subpolar and coastal regions. We have also learned that the interannual variability of $p\mathrm{CO_2}$ over the Southern Ocean is moderate compared to that over the Equatorial Pacific and polar/subpolar regions (see also in Rödenbeck et al., 2015; Denvil-Sommer et al., 2019). However, air-sea fluxes vary greatly over the Southern Ocean (SO), we also show that the SO RMSD between our fluxes and SOCAT-based estimates are larger than those of certain regions (Table S2).

The statistics (e.g., Bias, RMSD, r^2, and number of data grided from SOCAT observations) listed in Table S2 and scattered in Figure 3c for different open and coastal regions provide a general comparison of the reconstruction skill of the CMEMS-LSCE-FFNN model among the oceanic basins. Nevertheless, examining merely these numbers would not give us a robust assessment of the full story behind. As one of the contributions of this study compared to the existing publications, an intensive analysis has been made and presented in subsections 3.1.2-3.1.6 for both the open ocean and the coastal zones. Interpretations of key factors driving a good or poor reconstruction of surface $p\mathrm{CO_2}$ (e.g., data density and distribution, regional to local characteristics of $p\mathrm{CO_2}$ and its potential drivers, model design and resolution) are given. To strengthen our interpretations, we have shown both the temporal and spatial distribution of SOCAT data, model-data errors, model uncertainty (e.g., in Figures 3ab, 5e, S1-S3) and scattered them with their driving mechanisms suggested in the literature. Based on these materials, we have made an intercomparison of model reconstruction ability between regions, identified oceanic sectors where the model does not fit the data, and importantly we have suggested improvements on the data reconstruction.

- **Second, I'm a little confused by equation (2). Why is the equation for the mean squared deviation (MSD) shown when it's the root mean squared deviation (RMSD) which is calculated throughout the manuscript? Also, the text refers back to this equation for the definition of the σ misfit, but this definition is itself within the MSD equation and is not clearly labeled on its own.**

  For Equation 2, we will put the square root over the formula of MSD. The precise definition of σ_{misfit} will be given in the revised manuscript.

- **Lastly, I think the description of the wind speed product used should be included in the main text rather than the supplementary, considering that this will have a large impact on the overall flux numbers (which the authors do highlight in the results).**

  The wind speed product will be included in the main text.

3 I suggest re-working the 2nd paragraph of the abstract. This paragraph currently reads like a laundry list of different regions and where they fall in terms of largest total source/sink, largest flux density source/sink, along with coastal and open ocean qualifiers. This many iterations of "X is the greatest ..." makes it difficult to follow-along and is not particularly interesting (e.g. the equatorial Pacific as the strongest source of carbon to the atmosphere is not a surprising result). Instead, highlight some of the other key findings, like the increase in ocean carbon uptake over the 1985-2019 timeframe (right now the mean is just listed, but the change is highlighted in the conclusion).

The second paragraph of the abstract has been modified leading to changes in the full abstract as follows.

- Previous text: *We have estimated the air–sea $\mathrm{CO_2}$ fluxes ($fg\mathrm{CO_2}$) over the global ocean from the open sea to the continental shelves. Fluxes and associated uncertainty were computed from an ensemble-based reconstruction of $\mathrm{CO_2}$ sea surface partial pressure ($p\mathrm{CO_2}$) maps trained with observations from the Surface Ocean $\mathrm{CO_2}$ Atlas v2020 database. The ensemble mean (which is the best estimate provided by the approach) fits independent data well and a broad agreement between the spatial distribution of model-data differences and the ensemble standard deviations (which are our model uncertainty estimate) is seen. The space-time varying uncertainty fields identify oceanic regions where improvements in data reconstruction and extensions of the observational network are needed. Poor reconstructions of $p\mathrm{CO_2}$ are primarily found over the coasts and/or in regions with sparse observations, while $fg\mathrm{CO_2}$ estimates with largest uncertainty are observed over the open Southern Ocean ($44°S$ southward), the subpolar regions, the Indian gyre, and upwelling systems.*

  Our estimate of the global net sink for the period 1985–2019 is $1.643 \pm 0.125\ \mathrm{PgC\,yr^{-1}}$ including $0.150 \pm 0.010\ \mathrm{PgC\,yr^{-1}}$ for the coastal net sink. Results suggest that the open ocean Subtropical Pacific (between $18°N$–$49°N$) has the strongest $\mathrm{CO_2}$ sink ($0.485 \pm 0.014\ \mathrm{PgC\,yr^{-1}}$) among the basins of the world, followed by the open ocean sub-basins in the Southern hemisphere. The coastal Subpolar Atlantic (between $49°N$–$76°N$) is the most significant coastal net sink, amounting to one third of the total coastal uptake; the northern Pacific continental shelves (north of $18°N$) are the next contributors. The Equatorial Pacific (between $18°S$–$18°N$) is the predominant source emitting $0.523 \pm 0.016\ \mathrm{PgC\,yr^{-1}}$ of $\mathrm{CO_2}$ back to the atmosphere. Based on the mean flux density per unit area, the most intense $\mathrm{CO_2}$ drawdown is, however, observed over the Arctic ($76°N$ poleward) followed by the Subpolar Atlantic and Subtropical Pacific for both open ocean and coastal sectors. The mean efflux density over the Equatorial Pacific remains the highest, but similar densities can also be found along other strong upwelling systems in the equatorial band.

- Revised text: *We have estimated the air–sea $\mathrm{CO_2}$ fluxes ($fg\mathrm{CO_2}$) over the global ocean from the open sea to the coastal ocean. Fluxes and associated uncertainty are computed from an ensemble-based reconstruction of $\mathrm{CO_2}$ sea surface partial pressure ($p\mathrm{CO_2}$) maps trained with observations from the Surface Ocean $\mathrm{CO_2}$ Atlas v2020 database. The ensemble mean (which is the best estimate provided by the approach) fits independent data well and a broad agreement between the spatial distribution of model-data differences and the ensemble standard deviations (which is our model* *uncertainty estimate) is seen. The space-time varying uncertainty fields identify oceanic regions where improvements in data reconstruction and extensions of the observational network are needed. Poor reconstructions of* $p\mathrm{CO_2}$ *are primarily found over the coasts and/or in regions with sparse observations, while* $fg\mathrm{CO_2}$ *estimates with largest uncertainty are observed over the open Southern Ocean (*$44°S$ *southward), the subpolar regions, the Indian gyre, and upwelling systems.*

Our estimate of the global net sink for the period 1985–2019 is $1.643\pm0.125\ \mathrm{PgC\,yr^{-1}}$ *including* $0.150\pm0.010\ \mathrm{PgC\,yr^{-1}}$ *for the coastal net sink. Among oceanic basins, the open ocean Subtropical Pacific (*$18°$*N–*$49°$*N) and the coastal Subpolar Atlantic (*$49°$*N–*$76°$*N) are recognized as the strongest* $\mathrm{CO_2}$ *sinks contributing respectively* $0.485\pm0.014\ \mathrm{PgC\,yr^{-1}}$ *to the global ocean sink and one third to the total coastal uptake. Reconstruction results show significant changes in the global integration of* $\mathrm{CO_2}$ *fluxes exchanging through the air-sea interface. We compute a net flux of* $0.784\pm0.178\ \mathrm{PgC\,yr^{-1}}$ *in the year 1985 and an increase of the global ocean sink with a growth rate of* $+0.062\pm0.006\ \mathrm{PgC\,yr^{-2}}$. $\mathrm{CO_2}$ *absorption by the ocean was rather stable in the 1990s followed by an anomalous reduction in the years 1999–2001 and a strengthening uptake of* $2.301\pm0.126\ \mathrm{PgC\,yr^{-1}}$ *in the 2010s. The temporal standard deviation of the annual ocean uptake is* $0.526\pm0.022\ \mathrm{PgC\,yr^{-1}}$ *for the full period. The link between its large interannual to multi-year variations and the El Niño-Southern Oscillation climate variability is also reconfirmed.*

OTHER COMMENTS

4 Lines 37-40: Should all the manuscripts be separated with a comma rather than a semicolon? And why is the Rödenbeck et al. (2015) manuscript specifically highlighted as "other mapping methods"?

The comma is now used to separate the references. Rödenbeck et al. (2015) is cited in the manuscript as one of the studies which made an intercomparison between different observation-based mapping methods reconstructing ocean surface $p\mathrm{CO_2}$ and quantifying $\mathrm{CO_2}$ fluxes. However, we have changed the text in the Introduction (see our reply to Referee's comment **1**).

5 Line 139: How is the variability in "analytical equipment" accounted for here?

Thank you. The word "analytical equipment" was not appropriate in the context of the sentence in lines 139-141. We have rewritten it as follows:

- Previous text: *Variability in the number of cruises and analytical equipment induces measurement latitude and longitude offsets from the cell center, e.g., with an average of* $0.34°\pm0.14°$ *as reported in Sabine et al. (2013) which are not taken into account.*
- Revised text: *Variability in the location of cruises and instruments induces measurement latitude and longitude offsets from the cell center, e.g., with an average of* $0.34°\pm0.14°$ *as reported in Sabine et al. (2013) which are not taken into account.*

6 Figure 2: I suggest directly labeling each region in the figure with the abbreviated label (i.e. SpA for subpolar Atlantic) for clarity. Figure 5 and 8: The tick marks in the colorbar for these figures are relatively large and look like a negative sign, I'd suggest making them much smaller.

Figures 2, 5, and 8 have been modified following the Referee's suggestion. The label of Figure 2 is now changed from the numbers to the abbreviated names of 11 regions. The size of tick marks in the colorbar of Figures 5 and 8 is also reduced.

References

Bushinsky, S. M., Landschützer, P., Rödenbeck, C., Gray, A. R., Baker, D., Mazloff, M. R., Resplandy, L., Johnson, K. S., and Sarmiento, J. L.: Reassessing Southern Ocean air-sea CO_2 flux estimates with the addition of biogeochemical float observations, Global biogeochemical cycles, 33, 1370–1388, 2019.

Chau, T. T. T., Gehlen, M., and Chevallier, F.: QUALITY INFORMATION DOCUMENT for Global Ocean Surface Carbon Product MULTIOBS_GLO_BIO_CARBON_SURFACE_REP_015_008, Research report, Le Laboratoire des Sciences du Climat et de l'Environnement, https://resources.marine.copernicus.eu/product-detail/MULTIOBS_GLO_BIO_CARBON_SURFACE_REP_015_008/INFORMATION, 2020.

Denvil-Sommer, A., Gehlen, M., Vrac, M., and Mejia, C.: LSCE-FFNN-v1: a two-step neural network model for the reconstruction of surface ocean pCO_2 over the global ocean, Geoscientific Model Development, 12, 2091–2105, https://doi.org/10.5194/gmd-12-2091-2019, 2019.

Denvil-Sommer, A., Gehlen, M., and Vrac, M.: Observation System Simulation Experiments in the Atlantic Ocean for enhanced surface ocean pCO_2 reconstructions, Ocean Science Discussions, pp. 1–23, 2021.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S., Aragão, L. E. O. C., Arneth, A., Arora, V., Bates, N. R., Becker, M., Benoit-Cattin, A., Bittig, H. C., Bopp, L., Bultan, S., Chandra, N., Chevallier, F., Chini, L. P., Evans, W., Florentie, L., Forster, P. M., Gasser, T., Gehlen, M., Gilfillan, D., Gkritzalis, T., Gregor, L., Gruber, N., Harris, I., Hartung, K., Haverd, V., Houghton, R. A., Ilyina, T., Jain, A. K., Joetzjer, E., Kadono, K., Kato, E., Kitidis, V., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Liu, Z., Lombardozzi, D., Marland, G., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pierrot, D., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Smith, A. J. P., Sutton, A. J., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., van der Werf, G., Vuichard, N., Walker, A. P., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, X., and Zaehle, S.: Global Carbon Budget 2020, Earth System Science Data, 12, 3269–3340, https://doi.org/10.5194/essd-12-3269-2020, 2020.

Gregor, L. and Gruber, N.: OceanSODA-ETHZ: a global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification, Earth System Science Data, 13, 777–808, https://doi.org/10.5194/essd-13-777-2021, 2021.

Gregor, L., Lebehot, A. D., Kok, S., and Scheel Monteiro, P. M.: A comparative assessment of the uncertainties of global surface ocean CO_2 estimates using a machine-learning ensemble (CSIR-ML6 version 2019a) – have we hit the wall?, Geoscientific Model Development, 12, 5113–5136, https://doi.org/10.5194/gmd-12-5113-2019, 2019.

Hauck, J., Zeising, M., Le Quéré, C., Gruber, N., Bakker, D. C. E., Bopp, L., Chau, T. T. T., Gürses, z., Ilyina, T., Landschützer, P., Lenton, A., Resplandy, L., Rödenbeck, C., Schwinger, J., and Séférian, R.: Consistency and Challenges in the Ocean Carbon Sink Estimate for the Global Carbon Budget, Frontiers in Marine Science, 7, 852, https://doi.org/10.3389/fmars.2020.571720, 2020.

Landschützer, P., Gruber, N., Bakker, D., and Schuster, U.: Recent variability of the global ocean carbon sink, Global Biogeochemical Cycles, 28, 927–949, 2014.

Landschützer, P., Gruber, N., and Bakker, D. C.: Decadal variations and trends of the global ocean carbon sink, Global Biogeochemical Cycles, 30, 1396–1417, 2016.

Landschützer, P., Laruelle, G. G., Roobaert, A., and Regnier, P.: A uniform pCO_2 climatology combining open and coastal oceans, Earth System Science Data Discussions, pp. 1–30, 2020.

Laruelle, G. G., Lauerwald, R., Pfeil, B., and Regnier, P.: Regionalized global budget of the CO_2 exchange at the air-water interface in continental shelf seas, Global biogeochemical cycles, 28, 1199–1214, 2014.

Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J.-L., Delille, B., and Regnier, P.: Global high-resolution monthly pCO_2 climatology for the coastal ocean derived from neural network interpolation, Biogeosciences, 14, 4545–4561, 2017.

Lebehot, A. D., Halloran, P. R., Watson, A. J., McNeall, D., Ford, D. A., Landschützer, P., Lauvset, S. K., and Schuster, U.: Reconciling observation and model trends in North Atlantic surface CO_2, Global Biogeochemical Cycles, 33, 1204–1222, 2019.

Rödenbeck, C., Keeling, R. F., Bakker, D. C., Metzl, N., Olsen, A., Sabine, C., and Heimann, M.: Global surface-ocean pCO_2 and sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme, Ocean Science, 9, 193–216, 2013.

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., Landschützer, P., Metzl, N., Nakaoka, S., Olsen, A., Park, G.-H., Peylin, P., Rodgers, K. B., Sasse, T. P., Schuster, U., Shutler, J. D., Valsala, V., Wanninkhof, R., and Zeng, J.: Data-based estimates of the ocean carbon sink variability – first results of the Surface Ocean pCO_2 Mapping intercomparison (SOCOM), Biogeosciences, 12, 7251–7278, https://doi.org/10.5194/bg-12-7251-2015, 2015.

Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., Pfeil, B., Olsen, A., Metzl, N., Kozyr, A., Fassbender, A., Manke, A., Malczyk, J., Akl, J., Alin, S. R., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W.-J., Chavez, F. P., Chen, A., Cosca, C., Feely, R. A., González-Dávila, M., Goyet, C., Hardman-Mountford, N., Heinze, C., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Key, R. M., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A., Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G.-H., Paterson, K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Salisbury, J., Santana-Casiano, J. M., Sarma, V. V. S. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R., Skjelvan, I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K., Telszewski, M., Thomas, H., Tilbrook, B., Vandemark, D., Veness, T., Watson, A. J., Weiss, R., Wong, C. S., and Yoshikawa-Inoue, H.: Surface Ocean CO_2 Atlas (SOCAT) gridded data products, Earth System Science Data, 5, 145–153, https://doi.org/10.5194/essd-5-145-2013, 2013.

Watson, A. J., Schuster, U., Shutler, J. D., Holding, T., Ashton, I. G., Landschützer, P., Woolf, D. K., and Goddijn-Murphy, L.: Revised estimates of ocean-atmosphere CO_2 flux are consistent with ocean carbon inventory, Nature communications, 11, 1–6, 2020.