Response to Referee's Comments

Meike Becker (Referee 3): The authors present an estimate of global air-sea CO_2 fluxes based on interpolating gridded SOCAT pCO_2 data. They use an ensemble of 100 feed-forward neural network models (FFNN) and sea surface height, sea surface temperature, sea surface salinity, mixed layer depth, chlorophyll-a, atmospheric mole fractions, a pCO_2 climatology and position data as drivers. They present an uncertainty analysis based on their ensemble spread as a semi-independent parameter, which is better than many available air-sea flux products. However, there are a few things that should be improved.

Authors: We would like to thank Meike Becker (Referee 3) for her positive feedback and suggestions. We respond to each point as follows. Throughout this document, the referee's comments are in bold.

1 One point that needs improvement is the description of the driving data, where some important information is missing. The driving data that were used are not available for the full period for which the authors present flux maps. How did you deal with that? Did you use a climatology for CHL, SST, MLD etc. before the early/mid 90s? If so, what was this based on? This information is crucial for interpreting interannual variability prior to the mid-90s.

The information of all data used in our reconstruction has been presented in Table S1 (Supplementary material). As shown in this table, monthly data of sea surface temperature (SST) and atmospheric mole fractions (xCO_2) which are possibly key drivers to trends and interannual variability of the pCO_2 field are available over the full period (1985-2019). It is also noted at the end of Table S1: ***For some data unavailable before 1998, climatologies based on all available data were used as predictors. Exceptionally, predictors for SSH before 1993 are climatologies plus a linear trend in order to retain the overall response to the global warming. MLD before 1992 was taken as the average MLD between 1992 and 1997.* We, however, agree with the referee to make this information clearer to the readers. This information will be added in the main text.

2 Another thing that I want to point out, is the inconsistent and partly misleading use of the terms 'observations', 'sample' and 'data'. The authors base their product on a gridded version of the SOCAT data set (monthly, 1x1). In order to avoid confusion, the term 'observations' should be reserved for data that has been retrieved from field work, in the case the original pCO_2 measurements in the SOCAT database. The gridded version contains monthly, 1x1 averages of these pCO_2 measurements. When the authors write about 'X observations' in a certain region, they actually mean 'grid boxes with observations'. Please make sure that this becomes clearer. In line 195 for example, the authors write 50 to 220 samples per year'. Here the authors should specify that they mean 'grid boxes with data' as the reader easily can assume that there were only 50-220 pCO_2 observations every year.

Thank you for this suggestion. We will use precisely the terms 'observations', 'sample' and 'data' in the revised version of the manuscript.

3 I also want to comment on Figure S1. Here the authors show the coverage of the gridded SOCAT product and its variability where they mention pCO_2 individuals'. I don't understand if this means the original SOCAT pCO_2 observations (i.e. a measure of how well the gid box mean represents the actual conditions), or the pCO_2 of the gridded version (showing the variability within the gridded product).

In the (b) and (c) subplots of Figure S1 we show maximal variability of pCO₂ individuals within a grid cell, i.e.,

$$\max_{t} \{ p \text{CO}_{2,tij}^{\max} - p \text{CO}_{2,tij}^{\min} \}$$

where t, ij indicates time and space indices. $pCO_{2,tij}^{max}$ and $pCO_{2,tij}^{min}$ were converted from the corresponding values of CO_2 fugacity observations which are available in the monthly gridded SOCATv2020 database. We think that the term ' pCO_2

individuals' is correct in this sense. We will make consistent use of the terms of 'observations' and 'gridded data' throughout the main text, and reword the legend of Figure S1 to avoid any ambiguity.

4 I understand that the authors used the subocean divisions from RECCAP 1. This of course increases the comparability to the results of RECCAP 1, but also this makes the results difficult to interpret. Using a biome scheme such as used in RECCAP 2 (e.g. after Fay and McKinley (2014)) would have led to a clearer separation of regions with similar characteristics, and thus increased the interpretability. I also miss a discussion of how this product performs in comparison to other global air-sea CO₂ flux products.

We are on the same page with Referee 3 that using biomes proposed by Fay and McKinley (2014) would provide a better interpretation of small-scale characteristics of pCO_2 and air-sea CO_2 fluxes. We have also used the biome mask for further studies on their trends, seasonal cycles, and spatial and interannual variability. However, geometries of the biomes (e.g., their boundaries) would complicate the evaluation of the CMEMS-LSCE-FFNN model estimates and uncertainty, e.g., the comparison between the model outputs and sparse SOCAT data, and the analysis of results obtained for the open ocean and the coastal regions. Regarding the scope of this study, we have chosen to use the subocean divisions with latitude bands. Consequently, the CMEMS-LSCE-FFNN estimates of regional air-sea CO_2 fluxes have been compared to the ones presented in RECCAP1.

Each of the observation-based reconstruction methods for pCO_2 and CO_2 fluxes has both strengths and weaknesses, we have revised the Introduction (Lines 37-55), the new paragraphs (see italic text below) better interpret these terms and discuss the performance of the CMEMS-LSCE-FFNN approach compared to the others. In-depth intercomparisons amongst different model-based and/or observation-based products are presented in previous works including Rödenbeck et al. (2015); Denvil-Sommer et al. (2019); Hauck et al. (2020) are beyond the scope of this study.

Revised text:

Various data-based approaches have been proposed to infer gridded maps of surface ocean 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 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 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 pCO_2 climatologies and anomalies based on five neural network models and selected the one that reproduced the 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 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 pCO_2 mapping products and take 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. 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 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 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 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 pCO_2 over the global ocean including the coastal regions.

In this work, we propose a new inference strategy for reconstructing the monthly 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 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 pCO_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 pCO_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 CO₂ 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 CO_2 sources and sinks.

5 Minor suggestions

- L 40: *p*CO₂ was not introduced as an abbreviation.
- L 59: Tr is not described.
- L 86: change to: p is the number of grid cells with observations.
- Figure 3: The yellow bars in panel c) are very difficult to read, especially the first one.
- Figure 4a/b: You show the number of observation (or grid cells with observations) per year. Please change that.

We have taken them into account in this revision.

6 Additionally change STD to s. Go through the manuscript and make sure, that you use consistent terminology.

STD will be changed to σ for a consistent use of its notation.

7 L 140, add temporal offsets from the cell center. In many regions this will be the dominant one, especially during the productive season.

There is no temporal offsets provided in the SOCAT database.

8 L 182: Be aware that Laruelle et al. (2017) and Landschützer et al. (2020) are climatologies.

Landschützer et al. (2020) reconstructed monthly open and coastal pCO_2 based on two approaches: one proposed in Landschützer et al. (2016) using SOCAT data at $1^{\circ} \times 1^{\circ}$ resolution for the open ocean over 1982-2016 and another proposed in Laruelle et al. (2017) using SOCAT coastal data at $0.25^{\circ} \times 0.25^{\circ}$ resolution over 1998-2015. The monthly reconstructions were evaluated with SOCAT data over the common period 1998-2015 and the long-term mean and seasonal climatologies of pCO_2 shown in Landschützer et al. (2020) were created from those fields.

In lines 179-182 in the manuscript, we have written: For the 1998–2015 period, the CMEMS-LSCE-FFNN model scored an RMSD of 35.84 µatm, larger than the coastal reconstruction error of 26.8 µatm by Landschützer et al. (2020). The latter unified data for the same period from two conceptually equivalent reconstruction models, one covering the open ocean (Landschützer et al., 2016) and one targeting the coastal ocean (Laruelle et al., 2017).

9 L 307: Please round the uncertainties to 2 significant digits (or less if it seems unrealistically low) and the measured value to the same number of digits, for example 2.336 ± 0.104 to 2.34 ± 0.10 . Please do so for all uncertainties in the manuscript.

We have rounded estimates of air-sea fluxes and uncertainties to 3 digits since some of them become 0 with less than 3 digits; for instance, fluxes and uncertainty estimates over coastal regions (see in Table 2).

10 L 328-330: Please correct this. Primary production and respiration have usually only a very small influence on alkalinity (if we neglect anerobic remineralization processes for the moment): primary production increases alkalinity, while remineralization processes reduce alkalinity

This will be corrected.

11 L 332: Another important influence factor in coastal regions is the inflow of terrestrial POC, e.g. in the southern North Sea, leading to the release of CO₂ to the atmosphere.

We will consider to add this in the manuscript.

12 L 376-377: Are these really the dominant factors? After you argumentation for why the open ocean region is neutral (vertical convection brings up old, DIC rich water which balances the influx during summer) I would expect the absence of this deep mixing in coastal, shallow regions to be one of the major reasons why the coastal regions are a larger sink than the open ocean.

As shown in Bates (2006), Arrigo et al. (2010), and Ishii et al. (2014), surface DIC concentration is higher over the open, deep basins than the shallow coastal shelf seas of the subpolar Pacific, particularly induced by deep mixing during winter/spring. Also in this period, the coastal sector is covered by seasonal sea-ice resulting in a neutral region of air-sea fluxes while the open sea-ice free ocean (e.g., the southern Bering Sea) acts as a strong source of CO_2 . During spring/summer, high CO_2 uptake is found in coastal shelf seas influenced by river freshwater (e.g., Beaufort Sea, Arrigo et al., 2010) or by high biological production and sea-ice melt-water (e.g., Bering–Chukchi Shelves and the Gulf of Alaska, Yasunaka et al., 2016).

Thank you for correction. The text in lines 375-377 of the manuscript is revised with a broader context as follows. <u>Previous text:</u>

The enhanced uptake of CO_2 by the coastal ocean compared to the open ocean results from melt water discharge and high primary production over the shelves of the Chukchi and Bering Seas and the Gulf of Alaska in the spring/summer Yasunaka et al. (2016).

Revised text:



Figure RP3. Yearly global integrated air–sea flux estimates derived from the CMEMS-LSCE-FFNN ensemble (mean \pm uncertainty) for 1985–2019. Multivariate El Niño-Southern Oscillation Index (MEI; Wolter and Timlin, 1993, https://psl.noaa.gov/enso/mei/, last access: December 2020) is used to generally indicate a link between variations, e.g. Yearly uptake - Trend , in the CMEMS-LSCE-FFNN sink estimate and the ENSO climate mode (El Niño: MEI > 0.5, La Niña: MEI < -0.5, Neutral: otherwise).

The annual uptake of CO_2 by the coastal shelf seas is much higher than that compared to the open, deep basins as a result of a lower surface DIC concentration induced by winter/spring mixing in the shallower areas and the restriction of seasonal seaice on air-sea CO_2 exchanges (Bates, 2006; Arrigo et al., 2010; Ishii et al., 2014). Thus, the coastal sector acts as a neutral region of CO_2 fluxes in winter (Fig. 8). During spring and summer, a substantial amount of CO_2 is also absorbed in the coastal shelf seas influenced by river freshwater (e.g., Beaufort Sea) or by high biological production and sea-ice melt-water (e.g., Bering–Chukchi shelves and the Gulf of Alaska) (Arrigo et al., 2010; Yasunaka et al., 2016).

13 To be honest, I can't really see from this figure that it covaries with the ENSO mode. As I see it the flux increases equally often during La Nina as during El Nino. It would be more interesting to see a comparison of the interannual variability with other air-sea flux products.

The covariate between the ENSO events and the temporal variability of the global carbon sink has been covered by its increasing long-term trend in Figure 9 in the manuscript. We have added another curve whose values are ticked on the right y-axis of the same figure. This curve stands for the yearly flux variability, i.e., the yearly ocean uptake estimate after removing its long-term trend. See Figure RP3 as the revised version of Figure 9.

In-depth intercomparisons amongst different model-based and/or observation-based products are presented in previous works including Rödenbeck et al. (2015); Denvil-Sommer et al. (2019); Hauck et al. (2020). As far as we know, none of these studies shows a comparison of the covariate of the interannual variability of the flux products and the ENSO events. This point raised by the referee is interesting and will be considered in our future studies.

References

- Arrigo, K. R., Pabi, S., van Dijken, G. L., and Maslowski, W.: Air-sea flux of CO₂ in the Arctic Ocean, 1998–2003, Journal of Geophysical Research: Biogeosciences, 115, 2010.
- Bates, N. R.: Air-sea CO2 fluxes and the continental shelf pump of carbon in the Chukchi Sea adjacent to the Arctic Ocean, Journal of Geophysical Research: Oceans, 111, 2006.
- 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 CO₂ flux estimates with the addition of biogeochemical float observations, Global biogeochemical cycles, 33, 1370–1388, 2019.
- 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 pCO₂ 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 pCO₂ reconstructions, Ocean Science Discussions, pp. 1–23, 2021.
- Fay, A. R. and McKinley, G. A.: Global open-ocean biomes: mean and temporal variability, Earth System Science Data, 6, 273–284, https://doi.org/10.5194/essd-6-273-2014, 2014.
- 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 CO₂ 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.
- Ishii, M., Feely, R. A., Rodgers, K. B., Park, G.-H., Wanninkhof, R., Sasano, D., Sugimoto, H., Cosca, C. E., Nakaoka, S., Telszewski, M., Nojiri, Y., Mikaloff Fletcher, S. E., Niwa, Y., Patra, P. K., Valsala, V., Nakano, H., Lima, I., Doney, S. C., Buitenhuis, E. T., Aumont, O., Dunne, J. P., Lenton, A., and Takahashi, T.: Air–sea CO₂ flux in the Pacific Ocean for the period 1990–2009, Biogeosciences, 11, 709–734, https://doi.org/10.5194/bg-11-709-2014, 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 pCO₂ 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 CO₂ 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 pCO₂ 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 CO₂, 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 pCO₂ and sea–air CO₂ 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 pCO₂ Mapping intercomparison (SOCOM), Biogeosciences, 12, 7251–7278, https://doi.org/10.5194/bg-12-7251-2015, 2015.

- Wolter, K. and Timlin, M. S.: Monitoring ENSO in COADS with a Seasonally Adjusted Principal, in: Proc. of the 17th Climate Diagnostics Workshop, Norman, OK, NOAA/NMC/CAC, NSSL, Oklahoma Clim. Survey, CIMMS and the School of Meteor., Univ. of Oklahoma, 52, vol. 57, 1993.
- Yasunaka, S., Murata, A., Watanabe, E., Chierici, M., Fransson, A., van Heuven, S., Hoppema, M., Ishii, M., Johannessen, T., Kosugi, N., Lauvset, S. K., Mathis, J. T., Nishino, S., Omar, A. M., Olsen, A., Sasano, D., Takahashi, T., and Wanninkhof, R.: Mapping of the air–sea CO2 flux in the Arctic Ocean and its adjacent seas: Basin-wide distribution and seasonal to interannual variability, Polar Science, 10, 323–334, https://doi.org/10.1016/j.polar.2016.03.006, iSAR-4/ICARPIII, Science Symposium of ASSW2015, 2016.