Response to Referee's Comments

Authors: This document is a point-by-point reply to all referees' comments on our manuscript entitled "*A seamless ensemble-based reconstruction of surface ocean $p\mathrm{CO}_2$ and air–sea CO_2 fluxes over the global coastal and open oceans*". We split the document into Sections 1, 2, and 3. Throughout this document, the referees' comments are in bold and the revised text included in the manuscript is in italic. The manuscript has been revised corresponding to this point-by-point reply.

1 #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.

We would like to thank Referee 1 for constructive comments and suggestions on our study.

1.1 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 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 gridded SOCAT 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. A design of an ensemble of numerous feed forward neural network (FFNN) models:

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. Quantification and evaluation of model best estimates (ensemble means) and uncertainties (ensemble spreads):

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 pCO_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 pCO_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 observation–based data (e.g., the root-mean-square deviation, RMSD) to evaluate product quality and infer uncertainty estimates of the reconstructed pCO_2. Reconstruction errors of pCO_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 pCO_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 pCO_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 pCO_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. Seamless analysis of the reconstructed data and uncertainty estimates over the open ocean and coastal zones:
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 pCO_2 and air-sea CO_2 fluxes (e.g., data density and distribution, regional to local characteristics of pCO_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 pCO_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 first manuscript). The new paragraphs are produced below.

Lines 35-85 in the revised manuscript:
Various data-based approaches have been proposed to infer gridded maps of surface ocean pCO_2 from the sparse set of observation–based 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 (e.g., Rödenbeck et al., 2013; Landschützer et al., 2016; Iida et al., 2021), 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 $p\mathrm{CO}_2$ 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 $p\mathrm{CO}_2$ mapping products and took 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 $p\mathrm{CO}_2$ 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 $p\mathrm{CO}_2$ or an integrated flux over space and time of interest is under low confidence due to sparse data density. Also, most of the aforementioned mapping methods target $p\mathrm{CO}_2$ data and estimate air-sea fluxes solely over the open ocean, with the coastal data excluded or not fully qualified. In Laruelle et al. (2014), the authors present spatial distributions of air-sea flux density and estimates of the total coastal C sink inferred from spatial integration methods on coastal SOCAT data. Laruelle et al. (2017) adapted the two-step neural network approach described in Landschützer et al. (2016) to the coastal ocean $p\mathrm{CO}_2$. The coastal and open ocean products were combined into a single reconstruction to yield a global monthly climatology of $p\mathrm{CO}_2$ presented in Landschützer et al. (2020). Notwithstanding these advances, a global reconstruction and its uncertainty assessment of monthly varying coastal surface ocean $p\mathrm{CO}_2$ and air-sea fluxes are still missing.

In this work, we propose a new inference strategy for reconstructing the monthly $p\mathrm{CO}_2$ 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 $p\mathrm{CO}_2$ 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 in future observational programs and model development in order to improve 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 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 CO_2 sources and sinks.

1.2 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 observation–based 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. 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.

We have revised the manuscript and added in Section Methods details of the statistics used in this study to facilitate the readers' interpretation of our results (see the new Section 2.4 reproduced below). 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. 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 similar to 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 heretofore publications, an intensive analysis of the data reconstruction has been made and presented in Sections 3.1.2 and 3.1.3 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 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.

The precise definitions of σ_{misfit} and the root mean squared deviation (RMSD) are given in the revised manuscript. We have rewritten Section 2.4 (Statistics). The new Section 2.4 is as follows.

Lines 148-173 in the revised manuscript:

The mean (μ) and standard deviation (σ) of the 100-member ensembles of $p\mathrm{CO}_2$ and $fg\mathrm{CO}_2$ are respectively chosen as their best estimate and the associated uncertainty. Unless stated otherwise, a model best estimate and its uncertainty computed at each desired space-time resolution are denoted by $\mu_{\mathrm{ensemble}} \pm \sigma_{\mathrm{ensemble}}$, where

$$\mu_{\mathrm{ensemble}} = \frac{\sum_{i=1}^{i=100} p\mathrm{CO}_2^{\mathrm{Reconstruction}(i)}}{100}, \quad \sigma_{\mathrm{ensemble}} = \sqrt{\frac{\sum_{i=1}^{i=100} \left(p\mathrm{CO}_2^{\mathrm{Reconstruction}(i)} - \mu_{\mathrm{ensemble}}\right)^2}{100}}, \tag{2}$$

and $p\mathrm{CO}_2^{\mathrm{Reconstruction}(i)}$ is one of the 100 members of the reconstructed $p\mathrm{CO}_2$ fields. Similar definitions are applied for $fg\mathrm{CO}_2$. The units of air-sea flux estimates is $\mathrm{molC\,m^{-2}yr^{-1}}$ for a flux density and converted to $\mathrm{PgC\,yr^{-1}}$ for an integral over a region or the global ocean.

Model robustness of the reconstructed $p\mathrm{CO}_2$ fields is evaluated on the gridded SOCAT data and in situ observations (Sutton et al., 2019). The evaluation data is denoted as $p\mathrm{CO}_2^{\mathrm{Observation}}$ in the following formulas. Standard statistics include the coefficient of determination (r^2), misfit mean (model bias) and misfit standard deviation,

$$\mu_{\mathrm{misfit}} = \frac{\sum_{j=1}^{j=N} \mathrm{d}p\mathrm{CO}_2^j}{N}, \quad \sigma_{\mathrm{misfit}} = \sqrt{\frac{\sum_{j=1}^{j=N} \left(\mathrm{d}p\mathrm{CO}_2^j - \mu_{\mathrm{misfit}}\right)^2}{N}}, \tag{3}$$

and the root-mean-square deviation (RMSD)

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{j=1}^{j=N} \left(\mathrm{d}p\mathrm{CO}_2^j\right)^2}{N}}, \tag{4}$$

where $\mathrm{d}p\mathrm{CO}_2^j = p\mathrm{CO}_2^{\mathrm{Reconstruction}}[j] - p\mathrm{CO}_2^{\mathrm{Observation}}[j]$, and N is a number of evaluation data. All these scores are computed for different coastal and open regions from the scale of grid cells to the global scale.

Generally, RMSD measures the reconstruction 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. Compared to other metrics such as mean absolute bias and r^2, RMSD takes another role, an outlier detector, which gives larger weights to high model–data misfits. Note that r^2, μ_{misfit}, σ_{misfit}, and RMSD reflect the model performance with respect to evaluation data, while $\sigma_{\mathrm{ensemble}}$ measures the stability of the model best estimate μ_{ensemble}. Nevertheless, these different statistics should consistently reflect the skill of the model reconstruction, e.g., depending on the density and distribution of data sampling.

In the next section, both the temporal and spatial distributions of gridded SOCAT data and in situ observations, model–data errors, model best estimates and uncertainties are shown. An intensive analysis is presented for both the open ocean and the coastal zones. We then interpret key factors leading to a good or poor reconstruction of surface $p\mathrm{CO}_2$ and $fg\mathrm{CO}_2$, e.g., SOCAT data density and distribution, model design and resolution, regional to local characteristics of $p\mathrm{CO}_2$ and $fg\mathrm{CO}_2$, and their potential driving mechanisms.

- **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 is now added in the main text.

  Lines 91-92 in the revised manuscript:
  k is the gas transfer velocity computed as a function of the 10-meter ERA5 wind speed (Hersbach et al., 2020) following Wanninkhof (2014) and its coefficient is scaled to match a global mean transfer velocity of 16.5 $\mathrm{cm\ h^{-1}}$ (Naegler, 2009).

1.3 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. A new version of the full abstract is reproduced below.

Modified abstract:
We have estimated global air–sea CO_2 fluxes ($fg\mathrm{CO}_2$) from the open ocean to coastal seas. Fluxes and associated uncertainty are computed from an ensemble-based reconstruction of CO_2 sea surface partial pressure ($p\mathrm{CO}_2$) maps trained with

gridded data from the Surface Ocean 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 deviation (which is our model uncertainty estimate) is seen. Ensemble-based uncertainty estimates are denoted by $\pm 1\sigma$. 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 pCO_2 are primarily found over the coasts and/or in regions with sparse observations, while $fgCO_2$ 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. Among the ocean basins, the subtropical Pacific ($18°N$–$49°N$) and the subpolar Atlantic ($49°N$–$76°N$) appear respectively to be the strongest CO_2 sinks for the open ocean and the coastal ocean. Based on mean flux density per unit area, the most intense 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. Reconstruction results also show significant changes in the global annual integral of all open- and coastal-ocean CO_2 fluxes with a growth rate of $+0.062 \pm 0.006\ \mathrm{PgC\,yr^{-2}}$ and a temporal standard deviation of $0.526 \pm 0.022\ \mathrm{PgC\,yr^{-1}}$ over the 35-year period. The link between its large interannual to multi-year variations and the El Niño-Southern Oscillation climate variability is reconfirmed.

OTHER COMMENTS

1.4 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 if they are part of the sentence. 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 pCO_2 and quantifying CO_2 fluxes. However, we have changed the text in the Introduction (see our reply to Referee's comment **1.1**).

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

The word "analytical equipment" was not appropriate in the context of the sentence in Lines 139-141 of the first manuscript. Thank you for pointing it out. We have rewritten this sentence as follows. Temporal sampling bias is also a source of uncertainty, it is now added in this sentence as suggested by Referee 3 (comment **3.7**).

Lines 197-200 in the revised manuscript:
Variability in the sampling time and location of cruises and instruments induces temporal sampling bias (e.g., towards some days in a month and/or the summer months at high latitudes) and latitude and longitude offsets from the cell center (e.g., with an average of $0.34° \pm 0.14°$ as reported in Sabine et al., 2013) which are not taken into account.

1.6 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.

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 all the figures is also reduced.

---

2 #REFEREE 2: At present, there are many data products in marine physics, such as temperature and salinity products, but there are few data products in marine chemistry. I support the publication of more marine chemistry data products.

We thank Referee 2 for his/her interest in marine chemistry data products and comments/suggestions on our study.

2.1 The author reconstructed surface ocean $p\mathrm{CO_2}$ based on FFNN with region divided by latitudes and similar predictors with previous researches was used, which is not novel.

With the proposed ensemble-based mapping method, statistics, and keys findings presented in the manuscript, we believe that our study is novel and would be a valuable contribution for the marine science community. However, we admit 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 easy to interpret. We consider this Referee's feedback important and have revised the manuscript in such a way that our three main contributions (see Lines 30-72 in this document and our reply to comment **2.5** below) are elaborated and highlighted. We added relevant information to the last paragraphs in Section Introduction (Lines 35-85 in the revised manuscript). Note that oceanic regions divided by latitude bands are only used for the analysis of our results, FFNN models themselves do not follow oceanic regions or biomes in clustering before training as opposed to Landschützer et al. (2016) and Gregor et al. (2019). The Referee's comment **"The author reconstructed surface ocean $p\mathrm{CO_2}$ based on FFNN with region divided by latitudes"** seems to be misleading.

2.2 The reconstruction of $p\mathrm{CO_2}$ and sea-air $\mathrm{CO_2}$ flux over global coastal oceans are interesting works but the author needs to do much more works on the validation of coastal results. Because a standard deviation of 41.79 μatm between $p\mathrm{CO_2}$ results and SOCAT observations possibly leads to opposite results in the estimate of coastal $\mathrm{CO_2}$ flux.

After the introduction of our new ensemble-based approach, the current manuscript indeed presents numerous results and an intense analysis for evaluating our global reconstruction of monthly $p\mathrm{CO_2}$ and fluxes from the open ocean to the coastal zone. The coastal-ocean reconstruction is compared with monthly gridded SOCAT data (not used in our model fitting) and with the open-ocean reconstruction. We compute and analyze the estimates of coastal $p\mathrm{CO_2}$ and air–sea fluxes, their 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) (see Figures 4-8 and Table 2 in the main text and more in the Supplementary). This is one of the novel contributions of this study which complement to the existing ones focusing on analyzing the spatial distribution and/or a monthly climatology of their coastal estimates (see our interpretation in the Introduction in the revised manuscript).

Referee 2 explains that **a standard deviation of 41.79 μatm between $p\mathrm{CO_2}$ results and SOCAT observations possibly leads to opposite results in the estimate of coastal $\mathrm{CO_2}$ flux**. This seems to be misleading. Indeed, we have written in the manuscript:
Lines 184-188 in the revised manuscript (Lines 126-130 in the first manuscript):
The reconstructed $p\mathrm{CO_2}$ field matches SOCAT data well: both are normally distributed with the same mean of $361.3\ \mu\mathrm{atm}$ (Fig. 3a) and a high agreement for all percentiles (Fig. 3b) is seen. The slight under- or overestimation at high and low percentiles implies that the model is slightly biased towards the mean value, as is expected when predictor variables do not fully explain predictand variables in the training dataset. This reduced variability is also reflected in the difference between the data standard deviation based on SOCAT $p\mathrm{CO_2}$ ($41.79\ \mu\mathrm{atm}$) and the one based on CMEMS-LSCE-FFNN ($36.30\ \mu\mathrm{atm}$).
In this context, $41.79\ \mu\mathrm{atm}$ is the standard deviation of SOCAT data itself, it is not the standard deviation of differences between the reconstructed and SOCAT data.

However, in this revision, we added a new subsection (3.1.3 Time series stations) including model evaluation on data sampled at in situ stations (Sutton et al., 2019). This would facilitate for the readers qualifying our product (see this new subsection in Lines 360-384 of the revised manuscript; text below). Consequently, we reorganise Section 3.1 (Evaluation) as follows:

- Section 3.1.1 Global ocean remains the same as in the first manuscript.
- Section 3.1.2 Ocean basins comprises the model evaluation for regions in the Arctic, Atlantic, Pacific, Indian Ocean, Southern Ocean.
- Section 3.1.3 Time series stations (new in this revision) includes the model evaluation on both open and coastal data sampled at in situ stations (Sutton et al., 2019).

As part of this new section, our coastal-ocean reconstruction is evaluated on data sampled at the time series stations (Figures S5). Despite less skill than the ocean-ocean reconstructions (Figures S6 and S7), our coastal-ocean reconstructions are rather compatible with observation–based pCO_2 data (Figure S8). In general, all reconstructed time series cover the full period 1985-2019 and would therefore be useful for estimating long-term means, trends, and variations of CO_2 surface partial pressure and ultimately the corresponding air-sea fluxes. observation–based data are still sparse and mostly distributed over the past two decades (see also the data density in Figures S1, S3, and S4 in the Supplementary), densifying observation networks is in priority to provide a better validation of both coastal- and open-ocean data reconstructions.

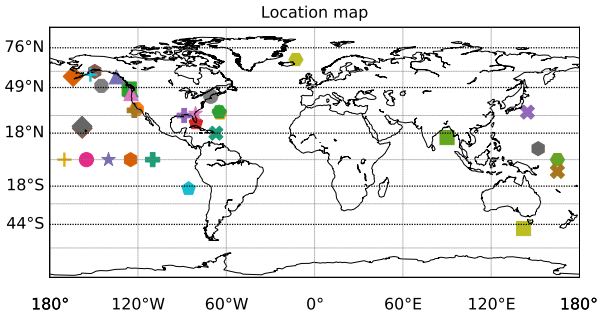

Figure S5. Location map of in situ measurements of ocean surface pCO_2 (Sutton et al., 2019).

Lines 360-384 in the revised manuscript (Section 3.1.3 Time series stations):

CMEMS-FFNN-LSCE estimates of pCO_2 are now compared with moored pCO_2 time series provided by Sutton et al. (2019). This data product comprises pCO_2 measurements collected from a wide range of oceanic regions since 2004 (Figs. S5–S8). Most of the stations were established in the North Atlantic and the North and Equatorial Pacific, one site is in the IO and another in the SO. Approximately one third of Sutton et al. (2019) sites belong to the coastal seas and shelves (Fig. S8). Table S3 details the information of the moored pCO_2 time series.

Observation–based data used for model–data comparison (black points in Figs S6–S8) are monthly averages of pCO_2 measurements at each site. This interpolation results in monthly time series with a number of data N between 9 (NH10) and 98 (WHOTS). The ensemble mean μ_{ensemble} and ensemble spread σ_{ensemble} (Eq. 2) are computed from the CMEMS-LSCE-FFNN ensemble of model outputs at the four nearest model grid boxes of each location. Results confirm a reasonably good reconstruction of the proposed approach. The model best estimates (coloured thick lines) characterise pCO_2 trends and variations of in situ data well and the model ensembles almost catch the observation–based data in their 99% confidence interval (light shaded envelop). Over 90% of the time series stations, the model estimation obtains a moderate to high coefficient of determination r^2 with a linear model–data correlation r larger than 0.5 (e.g., BTM: 0.98, CRESCENTREEF: 0.92, HOGREEF: 0.84, SOFS: 0.79, TAO110W: 0.75, WHOTS: 0.73). Mean bias μ_{misfit} (Eq. 3) and RMSD (Eq. 4) are relatively low compared to mean pCO_2 values of the time series stations.

Half of the open-ocean reconstructions have model errors less than 20 μatm and even less than 10 μatm at KEO, PAPA, SOLS, STRATUS, and WHOTS (Figs S6 and S7). Despite less skill than the open-ocean reconstructions, the coastal-ocean reconstructions are quite compatible with the in situ data (Fig. S8). Most of RMSDs remain lower than 20% of the mean pCO_2 values of coastal time series (e.g., CCE2: 36.53 μatm, ICELAND: 12.26 μatm, M2: 36.58 μatm). For some other stations

in the US west coast and the oceanic regimes of coral reef, the estimates differ from the observation–based data in terms of magnitude of pCO_2 (e.g., CRIMP2, LA PARGUERA) and/or of its seasonal cycle (e.g., CHABA, CHEECAROCKS, SEAK).

The reconstructed time series cover the full period 1985-2019 while observation–based data are still sparse and almost distributed over the past two decades (Figs. S6-S8). The CMEMS-LSCE-FFNN time series would be useful for estimating and assessing long-term means, trends, and variations of CO_2 surface partial pressure and the corresponding air-sea fluxes.

2.3 The CHL data used was only from 1992 to 2019, and was not available in the Arctic and the Southern Ocean in winter, the details about how the reconstruction was carried out when CHL was not available should be declared clearly in the method section.

This information is now added in the Method section (see the italic text below). To be precise, climatologies based on all available CHL data (1998-2019) were used as predictors for data unavailable before 1998. We set CHL approximately to 0 mg m^{-3} over the Arctic and the Southern Ocean in winter when no data are available (e.g., Landschützer et al., 2016; Denvil-Sommer et al., 2019; Gregor et al., 2019).

Lines 107-109 in the revised manuscript:
CHL was set approximately to 0 $mg\ m^{-3}$ over the Arctic and the Southern Ocean winter when no data is available. In case of data unavailable before 1998, climatologies based on all available data were used as predictors.

2.4 The subskin temperature correction (Watson et al., 2020) should be considered in the estimate of sea-air CO_2 flux.

Watson et al. (2020) proposed a double correction to the SOCAT data and to the computation of the CO_2 flux in order to remove some aliasing caused by the temperature vertical gradient within the marine boundary layer. However, the Watson et al. (2020) adjustment, if applied here, would add roughly $0.9\ PgCyr^{-1}$ to the global ocean sink estimate based on observations. The adjusted ocean sink estimate would thus surpass the land sink and result in a large carbon budget imbalance. More evidence of their genericity is needed before applying skin and subskin corrections (Friedlingstein et al., 2020).

2.5 The author should reconsider the topic of this manuscript. If the author want focus on the CO_2 flux of global open oceans, additional work was necessary rather than only discussing spatial distribution or interannual variability, because the reconstruction method in this manuscript and the results was not novel. If the author want focus on the CO_2 flux of global coastal oceans, which was still a research gap, much more works are needed to make the result convince.

As mentioned in our response to Referee's comments **1.1** and **2.1**, the manuscript presents three main contributions:

i. Our data reconstruction is based on a new model design - an ensemble of 100 neural network models.

ii. We quantify and evaluate model best estimates and uncertainties based on the ensemble asset. For the first time, the space-time varying uncertainty estimates (see for instance the Introduction - Lines 35-85 - and Figures 5 and 9 in the revised manuscript) derived from the ensemble of model outputs are presented and analysed. We promote the use of the proposed uncertainty fields which would be more informative than the RMSD-based uncertainty fields (e.g., proposed in Landschützer et al., 2014) in identifying oceanic sectors where further improvements on the data reconstruction will be needed.

iii. More importantly, the manuscript presents a seamless analysis of the reconstructed data and uncertainty estimates over the open ocean and coastal zones. Furthermore, the open ocean estimates are considered as references for the coastal data assessment.

We believe that the ensemble-based approach and analysis therein are novel and the Introduction section has been changed to better highlight these contributions. Also, we have added a new section (3.1.3) presenting an evaluation of the reconstructed

data on independent in situ observations in the revised version of the manuscript. Precisely, we propose to add Figures S5-S8, Table S3, and an interpretation of these results (see also in our reply to Referee's comment **2.2**).

2.6 Line 76: "An ensemble of 100 FFNNs was used to reconstruct monthly $p\mathrm{CO_2}$ fields……", How are these 100 models built? Why did you do that? How are the results of 100 models selected? Line 84-85: "The random extraction and the FFNN training were repeated 100 times so that 100 versions of the monthly FFNNs have been obtained", Why is it the "100 times"? How is the "100 times" determined? Does it converge after 100 iterations?

The description of the construction of the ensemble approach is already given in the manuscript as follows (Figure 1 is used to illustrate a neural network model mapping the target $p\mathrm{CO_2}$ and predictor variables).

Lines 117-120 in the revised manuscript (Lines 81-84 in the first manuscript):
To reconstruct the $p\mathrm{CO_2}$ fields over the global ocean for each target month over the 1985–2019 period, all the available SOCAT data and the co-located predictors have been collected for the month before and the month after the target month. We randomly extracted two thirds of each one of these datasets to make training datasets for the FFNNs, leaving the remaining third to be corresponding test datasets. The FFNNs were then trained for each target month.

Our ensemble approach comprises multiple network models, each trained and validated on resampled data of SOCAT $p\mathrm{CO_2}$ and predictors. We added Figure S2 in the Supplementary document and a new paragraph in (text below) in this revision of our manuscript, explaining the reasons we have selected 100 model runs for our study.

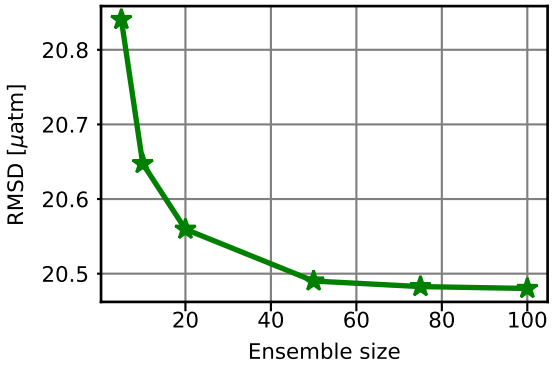

Figure S2. RMSD between a best estimate (ensemble mean) and SOCAT data of ocean surface $p\mathrm{CO_2}$ with respect to the ensemble size in $\{5,\ 10,\ 20,\ 50,\ 75,\ 100\}$.

Lines 124-134 in the revised manuscript:
The random extraction and the FFNN training were repeated 100 times so that 100 versions of the monthly FFNNs have been obtained. Note that our ensemble approach belongs to the classes of bootstraping and Monte Carlo methods in statistics. Theoretically, the number of samples or the ensemble size must be substantially large to get a convergence. However, it was demonstrated in the literature (e.g., Goodhue et al., 2012; Efron et al., 2015) that with the ensemble size of 50 the model estimation is likely stable and with the ensemble size over 100 the improvement in standard errors between model outputs and evaluation data is negligible. Fig. S2 shows an illustration of the reconstruction skill with respect to the ensemble size S. For each ensemble of S model outputs of $p\mathrm{CO_2}$ ($S \in \{5,\ 10,\ 20,\ 50,\ 75,\ 100\}$), the root-mean-square deviation (RMSD) is computed between the ensemble mean (our best model estimate) and SOCAT data over the period 1985-2019. As seen in this figure, the reconstruction starts to stabilize with $S = 50$. In this study, we have exploited a large but realistic amount of computing resources to run an ensemble of $S = 100$ neural network models.

2.7 Table 2, What is the meaning of two numbers in the rightmost column of Table 2, for example: 0.07 ± 0.04, 0.30 ± 0.13

The numbers without brackets (e.g., 0.07 ± 0.04, 0.30 ± 0.13) in the rightmost column of Table 2 in the manuscript refer to the estimates derived from observation–based methods. In the caption of Table 2, we wrote: *In column 'RECCAP1', values in parentheses are the 'best' estimates proposed by RECCAP1 studies, the others are the estimates computed with different methods using $p\mathrm{CO}_2$ observations*. More information is now added to the caption of Table 2.

Modified caption of Table 2:

Yearly mean of contemporary air–sea CO_2 fluxes ($\mathrm{PgC\,yr^{-1}}$) integrated over the global ocean and 11 RECCAP1 regions. Mean estimate and uncertainty ($\mu_{\mathrm{ensemble}} \pm \sigma_{\mathrm{ensemble}}$) of the CMEMS-LSCE-FFNN approach is shown for the coast (C), the open ocean (O), and the total area (T). For a comparison, estimates derived from RECCAP1 (Canadell et al., 2011; Schuster et al., 2013; Ishii et al., 2014; Sarma et al., 2013; Lenton et al., 2013; Wanninkhof et al., 2013) are provided. In column 'RECCAP1', values in parentheses are the 'best' estimates proposed by RECCAP1 studies which were derived from averages or medians of estimates based on the $p\mathrm{CO}_2$ climatology or $p\mathrm{CO}_2$ diagnostic model, and/or the atmospheric and ocean inversions, and GOBM models. The 'RECCAP1' values out of parentheses are the estimates derived from different methods mapping observation–based data of $p\mathrm{CO}_2$. With an exception for the global estimate (Wanninkhof et al., 2013), those of the RECCAP1 sub-basins are available only for the open ocean.*

2.8 Line 444-446: "The global open ocean uptake obtained in this study of 1.344 ± 0.111 PgC $\mathrm{yr^{-1}}$ lies between the observation based net sink estimate by Wanninkhof et al. (2013) (1.18 PgC $\mathrm{yr^{-1}}$) and the global sum of regional best estimates given in Table 2 (1.8 PgC $\mathrm{yr^{-1}}$)". In table 2, I can't find the value of 1.8 PgC $\mathrm{yr^{-1}}$

It means that 1.8 PgC yr^{-1} is the sum of all the 'best' estimates (between brackets) given in Table 2.

2.9 Line 462-463: The discrepancy is possibly due to an overestimation of Arctic $p\mathrm{CO_2}$ by the CMEMS-LSCE-FFNN (see in Sect. 3.1.2) and to the lack of estimates over a large portion of the seasonally sea–ice covered regions. This sentence means that the data in the Arctic are not accurate at present. So the data in the Arctic is not suitable for use at present.

Results shown in Sect. 3.1.2 in the manuscript confirm that the model reconstruction of $p\mathrm{CO}_2$ over the Arctic does not fit SOCAT data well and is much more uncertain than for other oceanic regions. The factors behind the poor estimates of Arctic $p\mathrm{CO}_2$ have been further discussed in the Discussion section (Lines 460-469 in the first manuscript, Lines 555-564 in the revised manuscript). Despite the need for further improvements, our analysis fairly documents the current status and discusses the way forward.

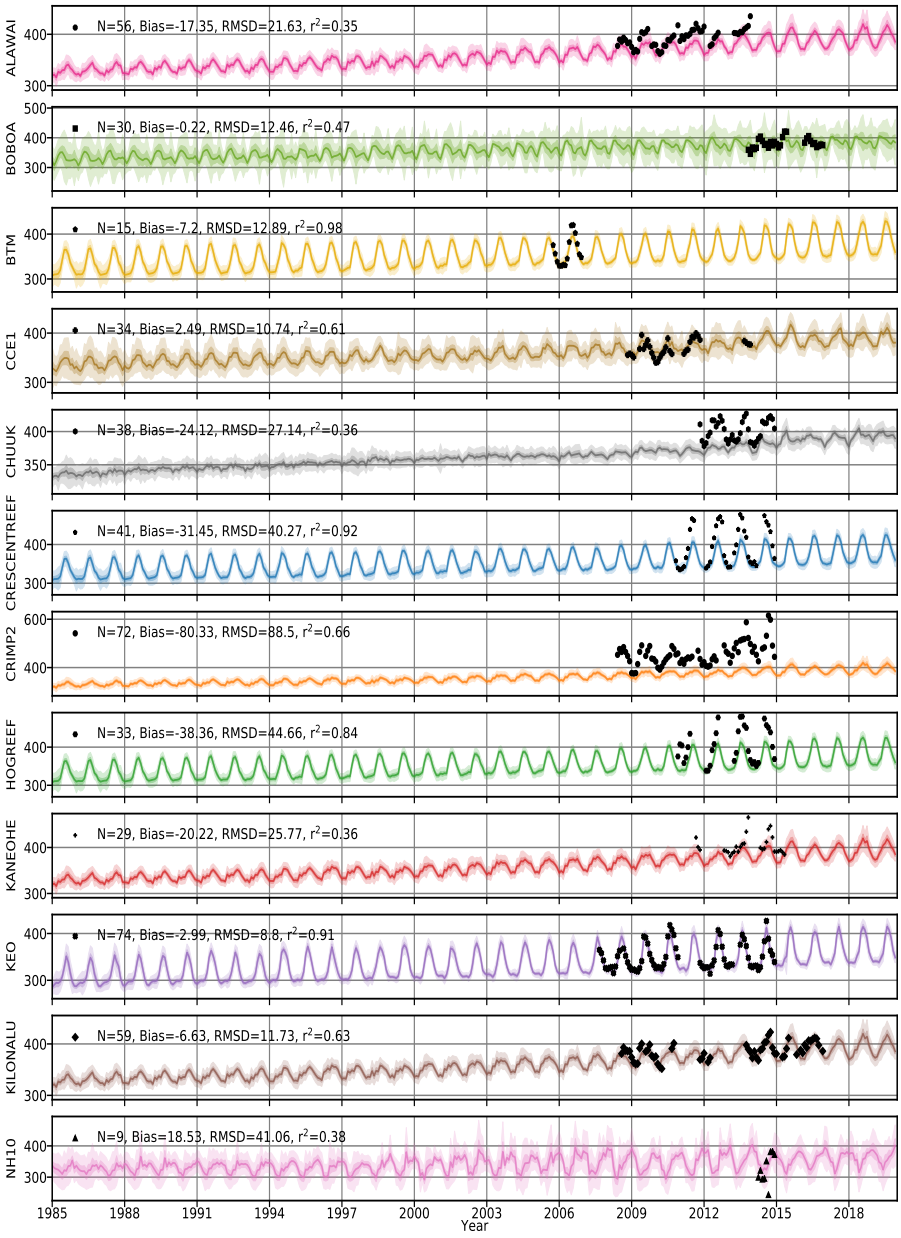

Figure S6. Time series of open ocean surface pCO_2 at different stations - part 1 (see station locations in Fig. S5 and Table S3). Evaluation data are monthly averages of measurements at each station (Sutton et al., 2019). The ensemble mean μ_{ensemble} and ensemble spread σ_{ensemble} (Eq. 2) are computed from reconstructed data at the four nearest neighbors of that location. Number of grid boxes with observations N, model bias μ_{misfit} (Eq. 3), RMSD (Eq. 4), and model–data correlation r^2 have been computed on these monthly interpolated data. In each subplot, dots stand for observation–based data and the coloured line with shaded areas stand for the mean and uncertainty envelops computed from the CMEMS-LSCE-FFNN 100-member ensemble (dark: 68% confidence interval, i.e. $\mu_{\text{ensemble}} \pm \sigma_{\text{ensemble}}$; light: 99% confidence interval, i.e. $\mu_{\text{ensemble}} \pm 3\sigma_{\text{ensemble}}$).

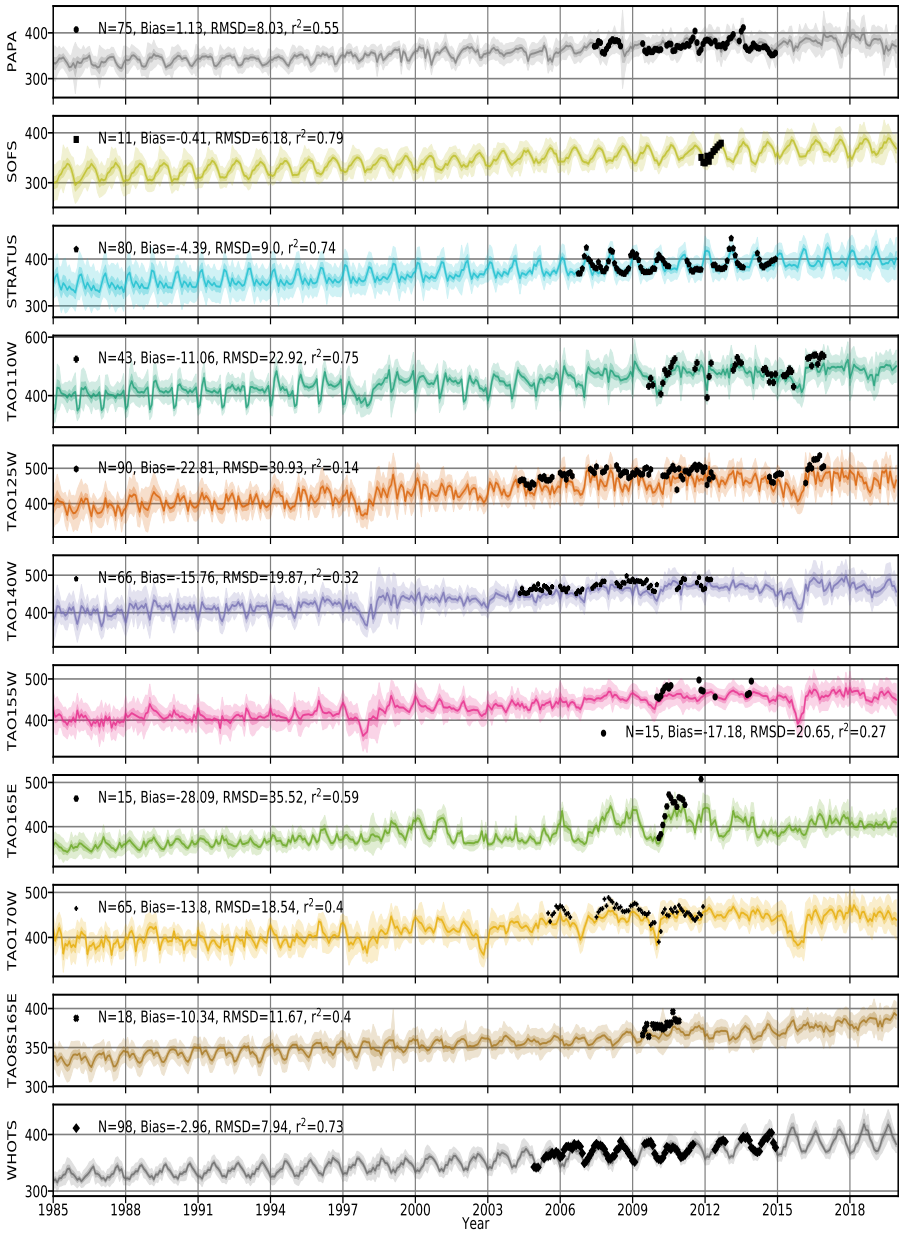

Figure S7. Time series of open ocean surface pCO_2 at different stations - part 2 (see station locations in Fig. S5 and Table S3). Evaluation data are monthly averages of measurements at each station (Sutton et al., 2019). The ensemble mean μ_{ensemble} and ensemble spread σ_{ensemble} (Eq. 2) are computed from reconstructed data at the four nearest neighbors of that location. Number of grid boxes with observations N, model bias μ_{misfit} (Eq. 3), RMSD (Eq. 4), and model–data correlation r^2 have been computed on these monthly interpolated data. In each subplot, dots stand for observation–based data and the coloured line with shaded areas stand for the mean and uncertainty envelops computed from the CMEMS-LSCE-FFNN 100-member ensemble (dark: 68% confidence interval, i.e. $\mu_{\text{ensemble}} \pm \sigma_{\text{ensemble}}$; light: 99% confidence interval, i.e. $\mu_{\text{ensemble}} \pm 3\sigma_{\text{ensemble}}$).

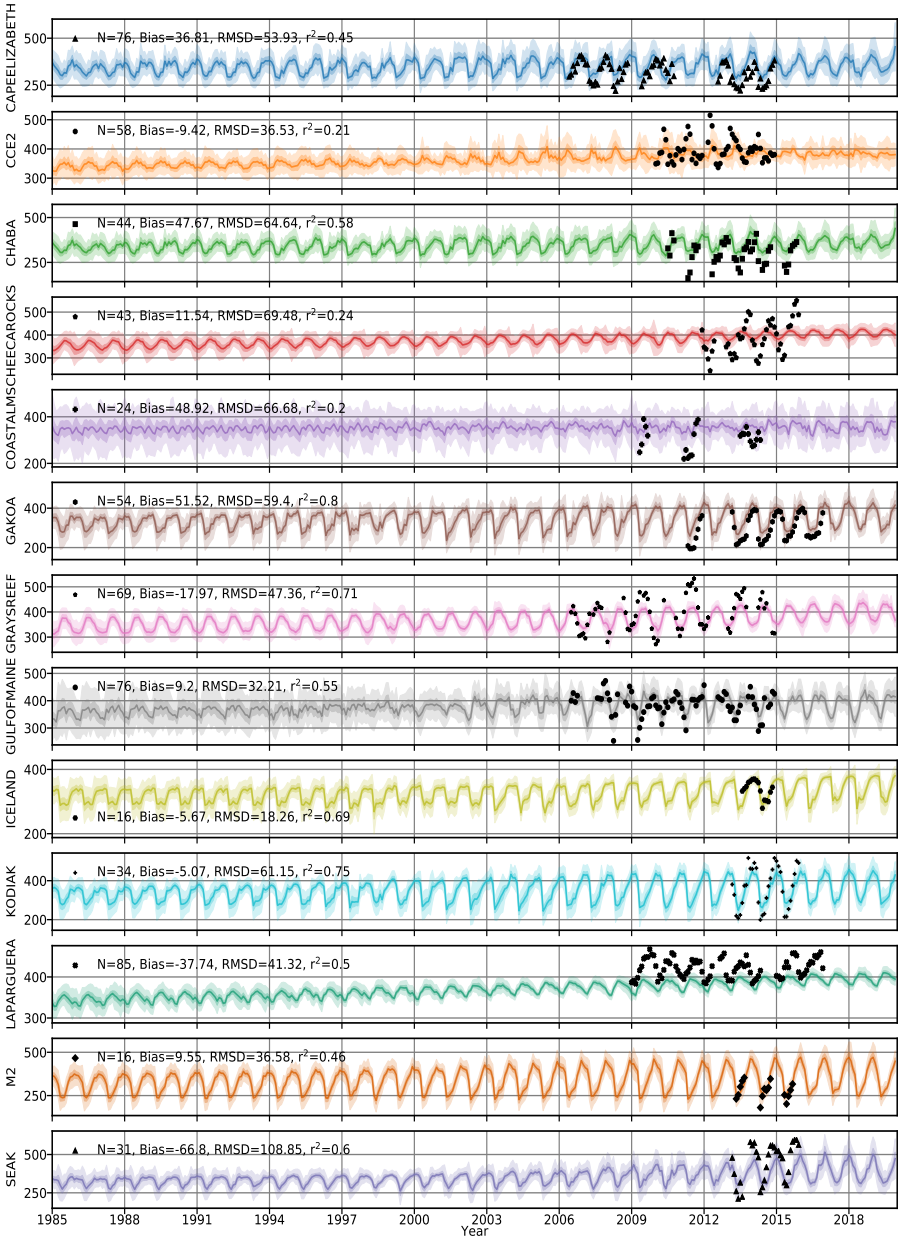

Figure S8. Time series of coastal ocean surface pCO_2 at different stations (see station locations in Fig. S5). Evaluation data are monthly averages of measurements at each station (Sutton et al., 2019). The ensemble mean μ_{ensemble} and ensemble spread σ_{ensemble} (Eq. 2) are computed from reconstructed data at the four nearest neighbors of that location. Number of grid boxes with observations N, model bias μ_{misfit} (Eq. 3), RMSD (Eq. 4), and model–data correlation r^2 have been computed on these monthly interpolated data. In each subplot, dots stand for observation–based data and the coloured line with shaded areas stand for the mean and uncertainty envelops computed from the CMEMS-LSCE-FFNN 100-member ensemble (dark: 68% confidence interval, i.e. $\mu_{\text{ensemble}} \pm \sigma_{\text{ensemble}}$; light: 99% confidence interval, i.e. $\mu_{\text{ensemble}} \pm 3\sigma_{\text{ensemble}}$).

3 #REFEREE 3 (Meike Becker): 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.

We would like to thank Meike Becker (Referee 3) for her positive feedback and suggestions.

3.1 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 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 is now added to the main text.

Lines 107-111 in the revised manuscript:
CHL was set approximately to 0 $\mathrm{mg\ m^{-3}}$ *over the Arctic and the Southern Ocean winter when no data is available. In case of data unavailable before 1998, climatologies based on all available data were used as predictors. Exceptionally, predictors for SSH before 1993 were 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.*

3.2 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 highlighting this. We agree that the wording has to be carefully selected. We went through the manuscript and refer now to "observation–based data" rather than "observations" (see for instance our correction in Lines 71-82 in this document).

3.3 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 pCO_2 individuals within a grid cell, i.e.,

$$\max_t\{pCO_{2,tij}^{\max} - pCO_{2,tij}^{\min}\}$$

where t, ij indicates time and space indices. $p\mathrm{CO}_{2,tij}^{\max}$ and $p\mathrm{CO}_{2,tij}^{\min}$ 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 '$p\mathrm{CO}_2$ individuals' is correct in this sense. We have made a consistent use of the terms of 'observations' and 'gridded data' throughout the main text, and reworded the caption of Figure S1 to avoid any ambiguity.

Modified caption of Figure S1:

(a) Spatial distribution of monthly gridded SOCATv2020 data. (b,c) Maximal variability of $p\mathrm{CO}_2$ individual data within a $1^\circ \times 1^\circ$-grid box (μatm), i.e. $\max_t\{p\mathrm{CO}_{2,tij}^{\max} - p\mathrm{CO}_{2,tij}^{\min}\}$, where t and ij indicate time and space indices. $p\mathrm{CO}_{2,tij}^{\max}$ and $p\mathrm{CO}_{2,tij}^{\min}$ were converted from the corresponding values of CO_2 fugacity observations in the monthly gridded SOCATv2020 database. Fig. S1c shows the distribution of the variability larger than the 80%-quantile.

3.4 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 CO_2 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 $p\mathrm{CO}_2$ and air-sea CO_2 fluxes. 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 $p\mathrm{CO}_2$ and CO_2 fluxes has both strengths and weaknesses. We have revised the Introduction (Lines 37-55 in the first manuscript) to better highlight the differences between the present approach and previous ones, as well as to emphasize its novelty. See new paragraphs of the Introduction in Lines 35-85 in the revised manuscript. 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), and Hauck et al. (2020) are beyond the scope of this study.

3.5 Minor suggestions

- **L 40: $p\mathrm{CO}_2$ 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 took these suggestions into account in this revision. Particularly for Tr, it was defined in Eq (1) in the first manuscript with respect to $Tr = kL(1 - f_{\mathrm{ice}})$. We deleted Tr from Eq. (1) since it is not referred elsewhere throughout the text. In the revised manuscript, Eq. (1) is as follows

$$\begin{aligned} fg\mathrm{CO}_2 &= kL(1 - f_{\mathrm{ice}})\ \Delta p\mathrm{CO}_2 \\ &= kL(1 - f_{\mathrm{ice}})\left(p\mathrm{CO}_2^{\mathrm{atm}} - p\mathrm{CO}_2\right). \end{aligned} \tag{1}$$

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

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

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

Temporal offsets of data sampling is now mentioned in the sentence in Lines 139-141 of the first manuscript. The revised text is below.

Lines 197-200 in the revised manuscript:
Variability in the sampling time and location of cruises and instruments induces temporal sampling bias (e.g., towards some days in a month and/or the summer months at high latitudes) and latitude and longitude offsets from the cell center (e.g., with an average of $0.34^\circ \pm 0.14^\circ$ *as reported in Sabine et al., 2013) which are not taken into account.*

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

Landschützer et al. (2020) present a global ocean $p\mathrm{CO}_2$ climatology product combining two individual data reconstructions: one for the open ocean proposed by Landschützer et al. (2016) using SOCAT data at a $1^\circ \times 1^\circ$ resolution over 1982-2016 and another for the coastal ocean proposed by Laruelle et al. (2017) using coastal SOCAT data at a $0.25^\circ \times 0.25^\circ$ resolution over 1998-2015. As stated in these two studies, their reconstructed data are available for each month. In Landschützer et al. (2020) (Table 1), the authors reported RMSDs (RMSEs in their study) of each of the open-ocean and coastal-ocean reconstructions computed with respect to monthly gridded SOCATv5 data over the common period 1998-2015. In Section 3.3, they then present a comparison between their seasonal climatology product and a seasonal climatology computed from SOCAT data but are aware that this assessment could not be robust due to temporal sampling bias of SOCAT observations.

In Lines 179-185 in the first manuscript, we cited 26.8 μatm in Landschützer et al. (2020) (Table 1) which is the RMSD of the coastal reconstruction by Laruelle et al. (2017). Even though the CMEMS-LSCE-FFNN RMSD (35.84 μatm) reported here were computed using Eq. (4) over the same period as in the previous study, RMSDs of the two approaches are quite different. Our study uses the MARCATS mask proposed by Laruelle et al. (2013) leading to a smaller coastal area and a different number of evaluation data than those in Laruelle et al. (2017) and Landschützer et al. (2020). These two studies use the coastal mask defined within $400\ \mathrm{km}$ distance from the sea shore. In addition, the leave-p-out cross-validation used in CMEMS-LSCE-FFNN model fitting permits to leave much more independent data for model evaluation than the previous approaches (see Sections Introduction and Methods in the revised manuscript).

We rewrote Lines 179-185 in the first manuscript to better highlight differences in model errors between these coastal reconstructions.

Lines 238-246 in the revised manuscript:
For the 1998–2015 period, the CMEMS-LSCE-FFNN approach scored an RMSD of 35.84 μatm while a recent coastal reconstruction by Landschützer et al. (2020) obtained an error of 26.8 μatm (see their Table 1). The latter presents a global ocean $p\mathrm{CO}_2$ *climatology product by unifying data over the same period from two conceptually equivalent reconstruction models: one covering the open ocean at a* $1^\circ \times 1^\circ$ *resolution (Landschützer et al., 2016) and one targeting the coastal ocean at a* $0.25^\circ \times 0.25^\circ$ *resolution (Laruelle et al., 2017). These heretofore reconstructions cover the coastal region with a broader definition (*$400\ \mathrm{km}$ *distance from the sea shore) than the MARCATS mask used in this study leading to the differences in characteristics and numbers of evaluation data of* $p\mathrm{CO}_2$*. In addition, the CMEMS-LSCE-FFNN model was designed with the leave-p-out cross-validation approach excluding much more independent data from monthly model fitting for model evaluation than in the previous models.*

3.9 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).

3.10 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

We have corrected it. Alkalinity was removed from the discussion in Lines 328-330 of the first manuscript. Below is the modified text.

Lines 419-422 in the revised manuscript:
High wind speeds also strengthen vertical mixing, a process supplying dissolved inorganic carbon (DIC) and nutrients to the surface ocean. During the spring and summer months, a vigorous biological activity (Sigman and Hain, 2012) counteracts the warming induced decrease in CO_2 *solubility and increase in* pCO_2 *by drawing down DIC (Feely et al., 2001).*

3.11 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 CO_2 to the atmosphere.

Thank you. The impact of the inflow of terrestrial nutrients on air-sea CO_2 exchanges over the coastal SpA is now discussed in our revised manuscript.

Lines 424-427 in the revised manuscript:
This contrasts with other coastal regions (e.g., southern North Sea and Baltic Sea) where the respiration of terrestrial particulate organic carbon from river run-off contributes to making these areas a strong seasonal source of CO_2 *(Borgesa and Gypensb, 2010; Becker et al., 2021).*

3.12 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 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. In spring/summer time, high CO_2 uptake is found in coastal shelf seas influenced by sea-ice retreat and high biological production, e.g. Chukchi and Gulf of Alaska (Yasunaka et al., 2016, 2018), and/or by dilution of sea waters from river freshwater with low salinity and DIC concentration, e.g., Beaufort Sea, Laptev, and East Siberia Seas (Arrigo et al., 2010).

Thank you for correction. We have rephrased our argument with a broader context as follows.

Lines 469-475 in the revised manuscript:
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 seas of the SpP, particularly induced by deep mixing during winter/spring. Over the same period, seasonal sea-ice also restricts gas exchanging, the coastal sector thus acts as a neutral region of CO_2 *fluxes (Fig. 8). During spring and summer, a substantial amount of* CO_2 *is also absorbed in the coastal shelf seas influenced by high biological production in large ice-free areas (e.g., Chukchi and Gulf of Alaska), and/or by dilution of sea waters from river freshwater*

with low salinity and DIC concentration (e.g., Beaufort, Laptev, and East Siberia Seas) (Arrigo et al., 2010; Yasunaka et al., 2016, 2018).

3.13 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 first 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 the revised version of Figure 9 below.

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). This study further focuses on presenting the three main points: (1) an introduction of our novel ensemble-based approach for reconstructing the global monthly pCO_2 fields and air–sea fluxes from the open ocean to coastal regions, (2) quantification and evaluation of model best estimates and uncertainties based on the ensemble asset for different time scales (e.g., monthly, yearly, and multi-decadal) and spatial scales (e.g., grid cells, sub-basins, and the global ocean), (3) a seamless analysis of the reconstructed data and uncertainty estimates over the open ocean and coastal zones. These key points are now also elaborated in the Introduction (see Lines 35-85 in the revised manuscript). Regarding the scope of this study and intense analysis presented in the current manuscript, we suggest to retain Figure 9 as it is revised as follows. However, this point raised by the referee is interesting and will be considered in our future studies.

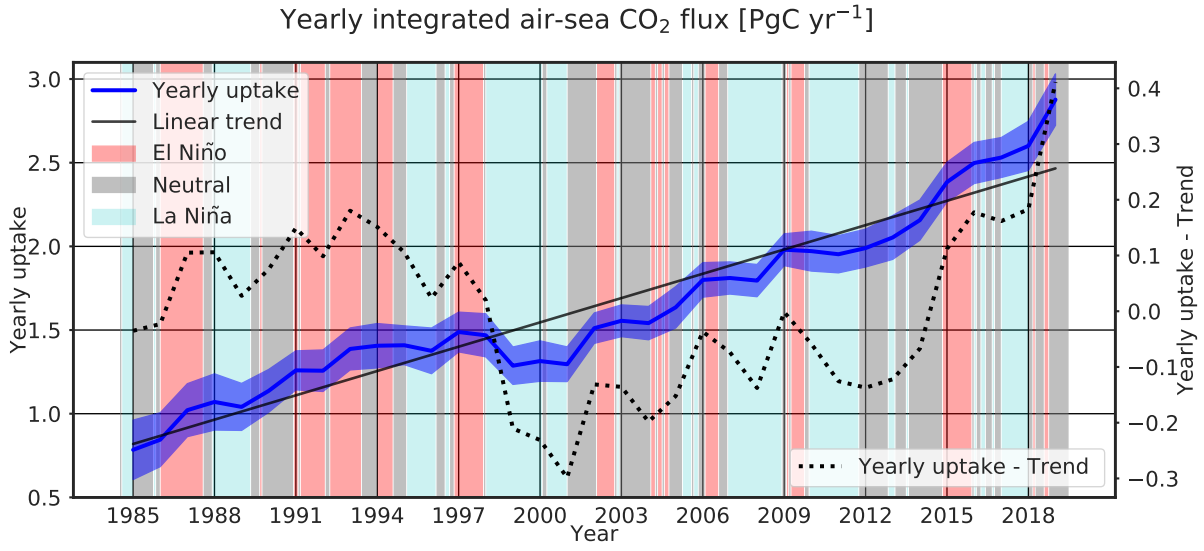

Figure 9. Yearly global integrated air–sea flux estimates derived from the CMEMS-LSCE-FFNN ensemble (mean ± uncertainty) for 1985–2019. Multivariate El Niño-Southern Oscillation Index (MEI; Wolter and Timlin, 1993, https://psl.noaa.gov/enso/mei/, last access: December 2020) is used to generally indicate 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).

** In addition to the comments addressed above, we also account editorial suggestions from our colleague Nicolas Metzl. We would like to thank him for his comments and supports to improve our manuscript. Aside from some points matching with the ones suggested by the three referees, the following items make other changes in this revision from the first manuscript:

- The color in Figures 5 (a,b), 8, and S9 (a,b) is reset: blue for CO_2 sink and red for CO_2 sources.
- The study of Yasunaka et al. (2018) is cited when discussing CO_2 fluxes over some sectors in the Arctic and subpolar regions.
- "Fair data used statement for SOCAT" is added in Acknowledgements.

References

Arrigo, K. R., Pabi, S., van Dijken, G. L., and Maslowski, W.: Air-sea flux of CO_2 in the Arctic Ocean, 1998–2003, Journal of Geophysical Research: Biogeosciences, 115, 2010.

Bates, N. R.: Air-sea CO_2 fluxes and the continental shelf pump of carbon in the Chukchi Sea adjacent to the Arctic Ocean, Journal of Geophysical Research: Oceans, 111, 2006.

Becker, M., Olsen, A., Landschützer, P., Omar, A., Rehder, G., Rödenbeck, C., and Skjelvan, I.: The northern European shelf as an increasing net sink for CO_2, Biogeosciences, 18, 1127–1147, https://doi.org/10.5194/bg-18-1127-2021, 2021.

Borgesa, A. V. and Gypensb, N.: Carbonate chemistry in the coastal zone responds more strongly to eutrophication than ocean acidification, Limnology and Oceanography, 55, 346–353, https://doi.org/https://doi.org/10.4319/lo.2010.55.1.0346, 2010.

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.

Canadell, J. G., Ciais, P., Gurney, K., Le Quéré, C., Piao, S., Raupach, M. R., and Sabine, C. L.: An international effort to quantify regional carbon fluxes, Eos, Transactions American Geophysical Union, 92, 81–82, 2011.

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, 17, 1011–1030, https://doi.org/10.5194/os-17-1011-2021, 2021.

Efron, B., Rogosa, D., and Tibshirani, R.: Resampling methods of estimation, 2015.

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.

Feely, R. A., Sabine, C. L., Takahashi, T., and Wanninkhof, R.: Uptake and storage of Carbon Dioxide in the ocean: the global CO_2 survey, Oceanography, 14, 18–32, https://doi.org/10.5670/oceanog.2001.03, 2001.

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.

Goodhue, D. L., Lewis, W., and Thompson, R.: Does PLS have advantages for small sample size or non-normal data?, MIS quarterly, pp. 981–1001, 2012.

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.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Vil-

laume, S., and Thépaut, J.-N.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/https://doi.org/10.1002/qj.3803, 2020.

Iida, Y., Takatani, Y., Kojima, A., and Ishii, M.: Global trends of ocean CO_2 sink and ocean acidification: an observation–based reconstruction of surface ocean inorganic carbon variables, J. Oceanogr., 77, 323–358, 2021.

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 CO_2 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., 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., Dürr, H., Lauerwald, R., Hartmann, J., Slomp, C., Goossens, N., and Regnier, P.: Global multi-scale segmentation of continental and coastal waters from the watersheds to the continental margins, Hydrology and Earth System Sciences, 17, 2029–2051, 2013.

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.

Lenton, A., Tilbrook, B., Law, R. M., Bakker, D., Doney, S. C., Gruber, N., Ishii, M., Hoppema, M., Lovenduski, N. S., Matear, R. J., McNeil, B. I., Metzl, N., Mikaloff Fletcher, S. E., Monteiro, P. M. S., Rödenbeck, C., Sweeney, C., and Takahashi, T.: Sea–air CO_2 fluxes in the Southern Ocean for the period 1990-2009, Biogeosciences, 10, 4037–4054, https://doi.org/10.5194/bg-10-4037-2013, 2013.

Naegler, T.: Reconciliation of excess 14C-constrained global CO_2 piston velocity estimates, Tellus B: Chemical and Physical Meteorology, 61, 372–384, 2009.

Olafsson, J., Olafsdottir, S. R., Takahashi, T., Danielsen, M., and Arnarson, T. S.: Enhancement of the North Atlantic CO_2 sink by Arctic Waters, Biogeosciences, 18, 1689–1701, 2021.

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.

Sarma, V. V. S. S., Lenton, A., Law, R. M., Metzl, N., Patra, P. K., Doney, S., Lima, I. D., Dlugokencky, E., Ramonet, M., and Valsala, V.: Sea–air CO_2 fluxes in the Indian Ocean between 1990 and 2009, Biogeosciences, 10, 7035–7052, https://doi.org/10.5194/bg-10-7035-2013, 2013.

Schuster, U., McKinley, G. A., Bates, N., Chevallier, F., Doney, S. C., Fay, A. R., González-Dávila, M., Gruber, N., Jones, S., Krijnen, J., Landschützer, P., Lefèvre, N., Manizza, M., Mathis, J., Metzl, N., Olsen, A., Rios, A. F., Rödenbeck, C., Santana-Casiano, J. M., Takahashi, T., Wanninkhof, R., and Watson, A. J.: An assessment of the Atlantic and Arctic sea–air CO_2 fluxes, 1990–2009, Biogeosciences, 10, 607–627, https://doi.org/10.5194/bg-10-607-2013, 2013.

Sigman, D. M. and Hain, M. P.: The biological productivity of the ocean, Nature Education Knowledge, 3, 1–16, 2012.

Sutton, A. J., Feely, R. A., Maenner-Jones, S., Musielwicz, S., Osborne, J., Dietrich, C., Monacci, N., Cross, J., Bott, R., Kozyr, A., et al.: Autonomous seawater pCO_2 and pH time series from 40 surface buoys and the emergence of anthropogenic trends, Earth System Science Data, 11, 421–439, 2019.

Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean revisited, Limnology and Oceanography: Methods, 12, 351–362, 2014.

Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., Gruber, N., Doney, S. C., McKinley, G. A., Lenton, A., Le Quéré, C., Heinze, C., Schwinger, J., Graven, H., and Khatiwala, S.: Global ocean carbon uptake: magnitude, variability and trends, Biogeosciences, 10, 1983–2000, https://doi.org/10.5194/bg-10-1983-2013, 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.

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/https://doi.org/10.1016/j.polar.2016.03.006, iSAR-4/ICARPIII, Science Symposium of ASSW2015, 2016.

Yasunaka, S., Siswanto, E., Olsen, A., Hoppema, M., Watanabe, E., Fransson, A., Chierici, M., Murata, A., Lauvset, S. K., Wanninkhof, R., Takahashi, T., Kosugi, N., Omar, A. M., van Heuven, S., and Mathis, J. T.: Arctic Ocean CO_2 uptake: an improved multiyear estimate of the air–sea CO_2 flux incorporating chlorophyll a concentrations, Biogeosciences, 15, 1643–1661, https://doi.org/10.5194/bg-15-1643-2018, 2018.