

**Response to Reviewers' comments on the manuscript:**

***CMEMS-LSCE: A global 0.25-degree, monthly reconstruction of  
the surface ocean carbonate system***

T. T. T. Chau, M. Gehlen, N. Metzl, F. Chevallier

We would like to thank the two reviewers for reading the manuscript thoroughly and providing constructive feedback. Based on their comments, we have improved the manuscript. Detailed replies are given hereafter. Each section comprises replies for general comments and minor (or specific) comments. Replies to Reviewer 2 start at page 22. The text quoted from the initial manuscript is in blue and revisions are presented in green.

On behalf of the authors,  
Thi-Tuyet-Trang Chau

## Replies to comments by Reviewer 1

### General comments by Reviewer 1 (GC1)

**GC1.1.** *Chau et al. present a exclusive approach using discrete ocean surface data of pCO<sub>2</sub> and total alkalinity (TA) to obtain a new monthly reconstruction for the period 1985 to 2021 with 0.25° resolution of the marine carbonate system variables. The reconstruction is based on the use of a feed-forward neural network (FFNN) for pCO<sub>2</sub>. For TA they use locally interpolated alkalinity regression (LIAR). The reconstruction is based on the CMEMS (Copernicus Marine Environment Monitoring Service) product, which provides global reconstructions of sea surface temperature (SST) and surface salinity (SSS) for the same period, including chlorophyll and other physical variables such as sea surface height. The authors start from a previous work where they published a similar database made with a resolution of 1° where only pCO<sub>2</sub> has been reconstructed. Here they expand the resolution by increasing it to 0.25° with the inclusion of TA, and then using the thermodynamic equations of the marine carbonate system they obtain the variables: Dissolved Inorganic Carbon (DIC), pH and degree of saturation of aragonite and calcite. In this way a product is generated that can be used to evaluate the impact of ocean acidification by other users and stake-holders. The quality of the reconstruction is contrasted with values observed at a series of oceanic and other coastal time stations. The authors provide two databases, one with 1° resolution and the other with 0.25° resolution.*

*The motivation and idea behind the paper is not original in the sense that this has been done before on a seasonal climate scale, but instead, here, the authors exploit the potential of CMEMS to obtain a reconstruction of all carbonate system variables on a spatial scale that has not been achieved so far and that can certainly be very useful in the evaluation of biogeochemical models and for the study of ocean acidification and in coastal regions of higher variability.*

#### **Authors:**

We thank **Reviewer 1** for highlighting potential use cases of our new CMEMS-LSCE datasets of surface ocean carbonate variables at high resolution. The two key points that set this contribution apart from previous studies include (1) a model upgrade for monthly pCO<sub>2</sub> reconstructions in spatial resolution from 1° (Chau et al., 2022) to 0.25° (this study) and (2) an extension to provide high-resolution datasets of A<sub>T</sub>, DIC, pH, and calcium carbonate saturation states covering the 37-year period. We hope the CMEMS-LSCE data product will be explored in further analyses of fine-scale spatiotemporal variations in marine carbonate variables complementary to previous contributions.

%%%

**GC1.2.** *The article is well written and provides detailed information both in the formalization of the equations and in the graphical information that is extended in the figures and equations of the appendix. However, it does not develop a specific discussion section of this new database or a comparison with other climatologies of  $p\text{CO}_2$ , DIC, AT and pH that would allow us to see the benefits, improvements and qualities of the new product. The authors, instead, compare in the 'Conclusions and Discussion' section the acidification rates with other observational results of other authors.*

**Authors:**

We thank **Reviewer 1** for appreciating the manuscript presentation. Our evaluation strategy is based on gridded SOCAT, GLODAP, and various time series station data. As suggested by **Reviewer 2** (comment **GC2.2.**) to gain reliability for our data evaluation, in this revision, we have shown new results at additional 38 time series stations with  $p\text{CO}_2$  and pH measurements and 4 sites with  $A_T$  and DIC measurements. Our product assessment is now stretched across the tropics, the subpolar sector, and the Southern Ocean to complement our previous data evaluation over the subtropical regions (See Figure A1b and statistics added in Tables A3 and A4 in the revised manuscript). An intercomparison with 1°-climatological data reconstructions (e.g., Broullon et al., 2019, 2020; Keppler et al., 2020) seems to us too outside the scope of our study given the importance of underlying methodological choices in such intercomparisons: (1) the discrepancy in mapping methods, input data resource, and the ratio of training and validation datasets used in model fitting, (2) uncertainty from post-processing applied for some products (e.g., filtering, smoothing, calibration), (3) the normalization of different data covering periods, and (4) quality of evaluation (or reference) data; e.g., observation data paucity should be one major concern to evaluate seasonal cycle reconstructions. The effect of such methodological choices calls into question the interpretation of differences between products if the different data providers do not actively contribute with sensitivity studies.

%%

**GC1.3.** *The source of information for the pCO<sub>2</sub> reconstruction is the Surface Ocean CO<sub>2</sub> Atlas version 2022 (SOCATv2022, 1985-2021) observations of CO<sub>2</sub> fugacity (fCO<sub>2</sub>). This database provides not only fCO<sub>2</sub> but the data are REQUIRED to be accompanied by SST and SSS. The fCO<sub>2</sub> data cannot be used independently of the SST and SSS with which it has been reported, since the temperature in the observation of fCO<sub>2</sub> has a high impact on the fCO<sub>2</sub> value itself (a bias of 1°C generates a bias in pCO<sub>2</sub> of 4.2%, ~18 µatm). The development of the pCO<sub>2</sub> reconstruction expressed in equation (1) does not meet that requirement. The authors mix the SOCAT observations with the SST and SSS reconstructions of the CMEMS product. This generates important errors as they themselves show in the reconstruction at oceanic (Figure 7) and coastal (Figure 5) fixed stations. Similarly, with TA, the observations used in LIAR also use temperature and salinity in Global Ocean Data Analysis Project bottle data version 2.2022 (GLODAPv2.2022, Lauvset et al., 2022). GLODAPv2 does not report TA without temperature and salinity observations so neither should different data sources be mixed when applying the LIAR methodology as is done in equation 2. Therefore, methodologically, the manuscript is seriously flawed in its numerical approach. The process should be done in two stages, first obtaining a set of FFNNs trained with fCO<sub>2</sub>, SST and SSS with the SOCAT data (and additionally the variables already included in equation 1), and then projecting that FFNN onto CMEMS' own reconstructions of SST and SSS. The same is true for TA and the use of LIAR. At least the SSS used in equation 2 should include the GLODAP SSS and not the CMEMS SSS. Better is to include the GLODAP SST, also. Then the coefficients developed with LIAR are used on the CMEMS reconstruction. This would greatly improve the reliability of the algorithms by better reproducing both the oceanic and coastal time series, not to mention that the GLODAP reconstructions shown in Figure 8 will do so as well. All this allows us to have a better estimate of the quality of the obtained algorithms since we can apply them to both oceanic and coastal time series with their own predictors and validate these algorithms. As currently performed in the manuscript, this validation is strongly biased because the SST and SSS reconstructions of CMEMS on these series clearly disagree when comparing point data with monthly means as indicated in the manuscript itself in Figure A8. In addition, a simple linear regression of TA versus salinity would report a better fit than the LIAR model applied in the manuscript.*

*As shown above, the monthly reconstruction proposed by the authors would be strongly improved if the two-step process is applied. The current product shown has a very poor quality in terms of validity since its comparison with the fixed time-series station used shows very high RMSD values (Figure 5, Table 7 and A3).*

#### **Authors:**

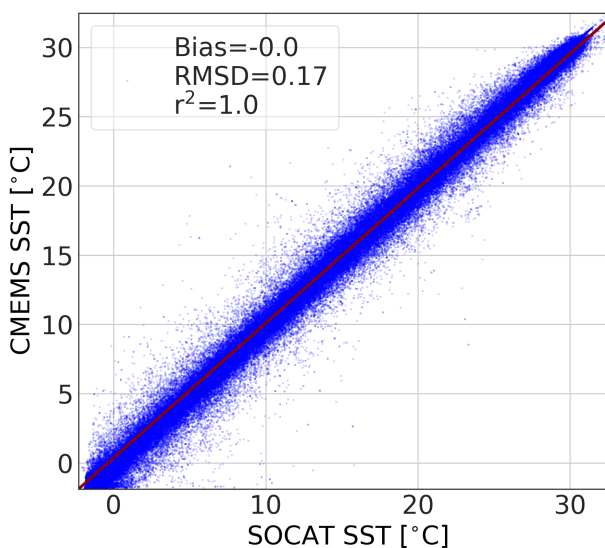
- **For pCO<sub>2</sub>:** Our FFNNs formally link a specific pCO<sub>2</sub> estimate (from the gridded SOCAT products) and specific estimates of environmental conditions (from the datasets listed in Table 1<sup>(1)</sup>), as expressed in Equation 1. There is no flaw in this approach that simply exploits the power of FFNNs (which are themselves non-linear regressors). The two-step approach proposed by the

---

<sup>1</sup> Table 2 in the initial manuscript (Table 1 has been removed as suggested by **Review 2**, comment **SC2.2**).

reviewer unnecessarily complicates the reconstruction process. In addition, how the second step (“*projecting that FFNN onto CMEMS’ own reconstructions of SST and SSS*”) can be made is not obvious: it seems to us that it would lose all the benefit gained by the first step.

Note that the bias between CMEMS SST and SOCAT data (or in situ observations) is relatively small (see Figures **GC1.3.** and **MC1.13.** in this document and also Figure A10 in the manuscript). Besides, the quality control has not been done for SOCAT salinity. There are some cruises in SOCAT with no salinity data and they have been replaced by SSS from the World Ocean Atlas (WOA) to recalculate CO2 fugacity (see Pfeil et al, 2013). SOCAT SSS would not be used in data reconstruction unless a critical quality control is performed. We have not changed the method.



*Figure GC1.3. Scatter plot of SOCAT SST and CMEMS SST gridded data over the global ocean in the period 1985-2021. The two datasets well fit to the bisector (red line) with no systematic bias, a RMSD of 0.17 and a  $r^2$  of 1.*

- **For A<sub>T</sub>: Reviewer 1’s comment (GC1.3.)** quoted below does not reflect precisely our method application. “Similarly, with TA, the observations used in LIAR also use temperature and salinity in Global Ocean Data Analysis Project bottle data version 2.2022 (GLODAPv2.2022, Lauvset et al., 2022). GLODAPv2 does not report TA without temperature and salinity observations so neither should different data sources be mixed when applying the LIAR methodology as is done in

equation 2. Therefore, methodologically, the manuscript is seriously flawed in its numerical approach.”

LIAR coefficients were estimated with GLODAPv2 data (Olsen et al., 2016) of SSS, SST,  $A_T$ ,... (Carter et al., 2018). In this study, we do not retrain LIAR models but use these existing coefficients to predict  $A_T$  with CMEMS SSS and SST (see Lines 257-260<sup>2</sup>) in the revision attached at the end of this document):

“*Locally interpolated alkalinity regression (LIAR; Carter et al., 2016, 2018) is an ensemble-based regression method developed for the global reconstruction of total alkalinity ( $A_T$ ). Regression coefficients were learned on GLODAPv2 data (Olsen et al., 2016) binned within regular windows of  $5^\circ \times 5^\circ$ . For prediction, the LIAR software interpolates between the regression coefficients to arbitrary resolutions specified by the users.*”

%%%%%%%%%

### **Minor comments by Reviewer 1 (MC1)**

Reviewer’s comments	Replies from Authors	Notes
<p><b>MC1.1. Line 18</b>  <i>“reconstructions with root-of-mean-square-deviation from observations less than 8%, 4%, and 1% relative to the global mean”</i> The relative percentage of RMSD over the mean is not a good parameter to evaluated the goodness of the results. For example, the accuracy of <math>A_T</math> is better than 0.1%, and <math>pCO_2</math> is similar. The percentages of</p>	<p>Lines 17-19 (abstract): we quote the full sentence below.  <i>“Product qualification with observation-based data confirms reliable reconstructions with root-of-mean-square-deviation from observations less than 8%, 4%, and 1% relative to the global mean of <math>pCO_2</math>, <math>A_T</math> (DIC), and <math>pH</math>.”</i>            We are surprised by this comment because normalizing the RMSD, for instance over the mean, may simplify the interpretation of this statistical quantity. In the quoted sentence of the abstract, it allows using the same metric (the relative amplitude of the error) for the four variables. We have not changed the sentence.</p>	

<sup>2</sup> Lines 231-234 in the initial manuscript

<p>RMSD reported are about two orders of magnitude higher.</p>	<p><b>Reviewer 1</b> mentions that “<i>the accuracy of <math>A_T</math> is better than 0.1%, and <math>pCO_2</math> is similar</i>”. These values likely correspond to analytical errors based on measurement quality controls at each station/location, e.g., 4 <math>\mu\text{mol kg}^{-1}</math> for GLODAPv2.2022 <math>A_T</math> (Lauvset et al, 2022) and 2-5 <math>\mu\text{atm}</math> for SOCATv2.2022 <math>pCO_2</math> (Bakker et al, 2022). Unsurprisingly, our reconstruction RMSD is substantially larger than analytical errors, which is only a minor source of uncertainty in the reconstruction process.</p> <p>Our reconstruction RMSD (e.g. 14.3 <math>\mu\text{atm}</math>, 22.1 <math>\mu\text{mol kg}^{-1}</math>, 22.7 <math>\mu\text{mol kg}^{-1}</math>, 0.022 for <math>pCO_2</math>, <math>A_T</math>, DIC, and pH) is in line with those reported in the previous studies (see our discussion in Lines 661-667* quoted below):</p> <p><i>“For instance, Iida et al. (2021) calculated <math>1\sigma</math>-uncertainty based on the median absolute deviation of regression model fits from open-ocean observations. Their approach yielded global <math>\sigma</math>-averages of 17.8 <math>\mu\text{atm}</math>, 11.5 <math>\mu\text{mol kg}^{-1}</math>, 0.018, and 0.110 for <math>pCO_2</math>, normalized DIC, pH, and <math>\Omega_{ar}</math>, respectively. In Gregor and Gruber (2021), the authors propagated the sum squared errors (global RMSD and measurement uncertainties) of <math>pCO_2</math> (15 <math>\mu\text{atm}</math>) and <math>A_T</math> (22 <math>\mu\text{mol kg}^{-1}</math>) obtaining global uncertainty estimates of 19 <math>\mu\text{mol kg}^{-1}</math> in DIC and 0.022 in pH.”</i></p>	<p>*Lines 615-619 in the initial manuscript</p>
<p><b>MC1.2. Line 20</b></p> <p>“<i>and 0.4% for pH</i>” It is a bit odd to report percentages of a logarithmic magnitude such as pH.</p>	<p>In the statistical sense, we consider pH as a variable similar to <math>pCO_2</math> and other carbonate system variables. All statistics are therefore reported with respect to the reconstructed variable. As explained in the previous comment (<b>MC1.1.</b>), with the intention of having a concise abstract, we choose to show the percentage of errors against the global mean value of each variable. It is noteworthy that percentages are also used in the scientific report <a href="#">SDG 14.3.1 (Table 1)</a>.</p>	

<p><b>MC1.3. Line 92</b></p> <p>The associated uncertainty reported in the article (<math>\sigma</math>) refers only to the uncertainty of the 100 replicate FFNNs, but they do not incorporate the uncertainty that each of the FFNNs has with respect to the SOCAT <math>p\text{CO}_2</math> values they are trying to replicate. The paper is only assessing a part of the uncertainty, by the way the smallest part and therefore not evaluating the ability of the FFNN set to reconstruct the input values.</p>	<p>The CMEMS-LSCE-FFNN 100-ensemble approach subsamples the gridded data of <math>p\text{CO}_2</math> and predictors to compose different training and test datasets, i.e., 100 training datasets for 100 FFNN models. In practice, it would allow to account for multiple sources of input data uncertainty from measurement errors, data sampling bias, data post-processing, etc, which have been poorly quantified in the input data products so far. In addition, the first layer of FFNNs is also initialized randomly at each of the 100 iterations. Therefore, our ensemble-based uncertainty includes the randomness in both subsampling datasets of <math>p\text{CO}_2</math> and predictors and in FFNN initialization. In Chau et al. (2022) (Section Methods), the authors described the ensemble approach comprehensively. This study extends the model by Chau et al. (2022) and thus recaps its principle.</p> <p>We modify the text in Lines 239-241* and add another one (in green) as follow for clarification:</p> <p><i>“After excluding the data in the reconstruction month, the data within the 3-month window are <b>randomly</b> separated into FFNN training and validation subsets with a ratio of 2 : 1. <b>The subsampling process is repeated for each 100 FFNN runs that results in 100 different datasets for model fitting.</b>”</i></p>	<p>*Lines 214-215 in the initial manuscript</p>
---	--	---



<p><b>MC1.4. Table 1, Table 2 and also Table 3</b> should include a value or an estimate of the uncertainty of each of the variables, either in their analytical determination or that which each product or reconstruction generates for each of the variables. This helps the reader to evaluate the quality of the reconstruction as a function of own error in the determination of each of the reconstructed variables.</p>	<p>Thank you. We have added the measurement errors with respect to each variable in Tables 1 and 2* if they are available from input data resources.</p>	<p>*Tables 2 and 3 in the initial manuscript (Table 1 has been removed as suggested by <b>Review 2</b>, comment SC2.2)</p>
<p><b>MC1.5. Line 126</b> Table 3 is cited before Table 2</p>	<p>We have revised the manuscript and cited Tables/Figures in order.</p>	
<p><b>MC1.6. Line 214</b> It is not sufficiently clear how to proceed with the reconstruction. It talks about excluding data in the month of reconstruction. Therefore, it would appear that for each month 100 FFNN reconstructions are performed. If this is correct, the RMSD for each month should be included in the figure or table of the SOCAT pCO<sub>2</sub> reconstruction since that data is not used in the month-specific reconstruction.</p>	<p>We quote Lines 238-241* from the revised manuscript for a straightforward response to <b>Reviewer 1</b> (modification in green corresponding to our reply to comment <b>MC.1.3.</b>):  <i>“The datasets of SOCAT fCO<sub>2</sub> and predictors are first reprocessed to match model fitting requirements (Sect. 2.1). After excluding the data in the reconstruction month, the data within the 3-month window are <b>randomly separated into FFNN training and validation subsets with a ratio of 2 : 1. The subsampling process is repeated for each 100 FFNN runs that results in 100 different datasets for model fitting. The excluded SOCATv2022 datasets are used in model evaluation.</b>”</i>  Here we specify the three independent datasets for FFNN training, validation, and evaluation. In the fitting phase of FFNN, we do not use SOCAT fCO<sub>2</sub> in the month specified for reconstruction to train and validate</p>	<p>*Lines 213-215 in the initial manuscript</p>

	FFNN models. In the reconstruction step, predictors data are available over the global ocean and FFNNs reconstruct $f\text{CO}_2$ for the target months. This exclusion strategy, called cross-validation, is widely used within machine learning approaches to avoid overfitting.	
<b>MC1.7. Line 249</b> Fig A7 is not cited in order.	Thank you. We have revised the manuscript and cited Tables/Figures in order.	
<b>MC1.8. Lines 285 and 210</b> How do you solve the discontinuities of the variable 'longitude' around the prime meridian $0^\circ$ . This is usually solved using the sine and cosine functions of longitude. Any reason for not doing so? Does this variable really bring any improvement in the FFNN?	To preserve the continuity of longitude at $0^\circ$ , we have applied both the sine and cosine functions to that coordinate. Hence, our global maps of carbonate variables (e.g. Figures 1, 6, 9) do not show discontinuity at the prime meridian. The sine is also used to transform latitude. Data transformation of predictor variables is explicitly presented in a sequence of preceding studies for the CMEMS-LSCE-FFNN model development (Denvil-Sommer et al 2021, Chau et al 2022). In the first manuscript version, we avoided repeating part of the data processing and model description from the previous studies. As the readers would concern, we have called back this information in the revision (Lines 162-164): <i>“The sine function is applied to convert latitude while both the sine and cosine are used to transform longitude to conserve their periodical behaviors.”</i>	
<b>MC1.9. Table 4</b> First of all, it should be pointed out that there is an excess of significant figures, not only in this table but throughout the text. Regarding the $p\text{CO}_2$ results, the	The manuscript describes and evaluates long-term datasets of multiple variables. A significant number of figures and tables corresponds to the presentation of many results of these variables. $p\text{CO}_2$ errors (e.g. Bias, RMSD) have been reported with 1-2 decimals in previous studies (Landschuter et al 2020; Denvil et al 2019, Gregor et al.	

<p>authors should remove all decimal places since analytically its precision is 2 <math>\mu\text{atm}</math> as described in the article. But more importantly, once the superfluous decimal places have been removed, what is observed is that there is practically no significant improvement between the product 'r025' and 'r100'.</p>	<p>2019, 2021). In this revision, we reduce the decimals from 2 to 1 for <math>p\text{CO}_2</math>, <math>A_T</math>, and DIC. The modification has been applied for Tables, Figures, and texts involving these variables. Note that 2-5 <math>\mu\text{atm}</math> reported in the manuscript represents the precision of measurement replications or analytical errors based on measurement quality control at each station/location (Sutton et al., 2019; Bakker et al., 2022).</p> <p>Table 3* shows a marginal improvement from r100 to r025 in terms of global evaluation metrics. For the open ocean, we expect to obtain similar skill scores for both FFNN models as the spatial autocorrelation of open-ocean <math>p\text{CO}_2</math> is estimated within <math>400\pm 250</math> km (Jones et al., 2012) and the SOCAT <math>1^\circ</math>-open-ocean dataset was used in model fitting. As also noted by Chau et al., (2022), <math>p\text{CO}_2</math> over the coastal ocean is characterized by high variability at small scales. For instance, <math>p\text{CO}_2</math> levels can vary with a horizontal gradient as large as 470 <math>\mu\text{atm}</math> over a distance of less than 0.5 km (Chavez et al., 2018; Feely et al., 2008). Probably, statistical models would need a spatial resolution much finer than <math>0.25^\circ</math> (25 km) and a temporal resolution higher than monthly in order to capture such high variability in surface ocean <math>p\text{CO}_2</math> present in observations (see also Bakker et al., 2016; Laruelle et al., 2017). In addition, measurement uncertainty of SOCAT gridded data due to undersampling is possibly one of the major sources of the irreducible model-data errors. Please refer to our reply to comment <b>MC1.11.</b> for a discussion on the benefits of the higher resolution.</p>	<p>*Tables 4 in the revised manuscript (Table 1 has been removed as suggested by <b>Review 2</b>, comment SC2.2)</p>
<p><b>MC1.10. Line 350</b> How is the regridding process performed? What type of interpolation is performed?</p>	<p>All the 3-dimensional datasets provided in this study have been saved as netCDF numerical files. To regrid these datasets, we use the Climate Data Operators (CDO) remapping operator, namely “remap”. CDO remap supports converting netCDF datasets from one horizontal grid to another. This operator</p>	

	<p>has been widely used in standard processing for numerical and statistical model outputs.</p> <p>We have revised the last sentence in Lines 374-377* to make it clear to the readers.</p> <p><i>“Table 43** also presents statistics for the monthly FFNN products of surface ocean pCO<sub>2</sub> at spatial resolutions of 0.25° (r025) and 1° (r100) together with their variants (r100 → 025 and r025 → 100). The latter are respectively extrapolation and interpolation versions of the original r100 and r025 datasets, i. e.,. We used the Climate Data Operators (CDO) remap operator to regrid FFNN model outputs (r100 and r025) <del>regridded</del> to a finer or coarser spatial resolution.”</i></p>	<p>*Lines 348-350 in the initial manuscript</p> <p>**Table 4 in the initial manuscript (Table 1 has been removed as suggested by <b>Review 2</b>, comment <b>SC2.2.</b>)</p>
<p><b>MC1.11. Line 354</b></p> <p><i>“The FFNN(r025) central to this study yields a lower RMSD and a higher correlation to the SOCAT data than the FFNN(r100→ 025)”</i>. Unfortunately, there is no significant difference between the two products. This statement is not correct. Line 393. It seems a very marginal the 2% improvement in pCO<sub>2</sub> reconstruction capability</p>	<p>Our statement is upheld even though the increment in global skill scores relative to a low to high spatial resolution is not large. Here we do not mention getting a significant improvement but still obtained higher scores in RMSD and r<sup>2</sup> when increasing the model spatial resolution. Please refer to Table 3* for verifying the statistics with respect to FFNN(r025) and FFNN(r100→r025) and our reply to comment <b>MC.1.9.</b> for analysis.</p> <p>Apart from Table 3*, benefits by increasing model spatial resolution from 1° to 0.25° are also demonstrated in Figures 2-4 with analyses shown in Lines 414-420**:</p> <p><i>“The two FFNN reconstructions (r025 and r100) share similarities in overall structures of pCO<sub>2</sub> over the coastal-open-ocean continuum (Figs. 2-4). However, the higher spatial resolution outperforms its lower resolution counterpart in reproducing fine-scale features of pCO<sub>2</sub> in the transition from nearshore regions to the adjacent open ocean. The increase in model spatial resolution translates into a greater spatial coverage of the continental shelves such as Labrador Sea, Northern Europe, and Sea of Japan (Fig. 3), and thus</i></p>	<p>*Table 4 in the initial manuscript (Table 1 has been removed as suggested by <b>Review 2</b>, comment <b>SC2.2.</b>)</p> <p>**Lines 387-394 in the initial manuscript</p>

	<p><i>an increase in the number of data over the coastal domain. The increase in spatial resolution allows a gain in prediction probability of pCO<sub>2</sub> variations on the order of roughly 2% over the Eastern Boundary Currents to 8% over the Western South Atlantic (Figs. 2-3b)."</i></p> <p>This study also points out temporal data sampling bias as a source of uncertainty that would highly constrain model reconstruction skills. Based on the assessment at station time series (Figure 5 and Table A3), we found that in situ observations have been sampled with low frequency and the bias of sampling date is about a week from a month center. With the low number of observations and high variability of pCO<sub>2</sub> (20.12 to 69.98 μatm) over these stations, it would not be statistically sufficient to refer to their temporal mean as a representative of monthly averages. A large model-data deviation would be retained even if we increase spatial resolution (see text in Lines 434-447*** for further analysis).</p>	<p>***Lines 403-414 in the initial manuscript</p>
<p><b>MC1.12. Line 375 and 393</b></p> <p>Line 375. The differences in RMSD between the regridded r100 and r025 products are very small, or even in some as in Canary Current System it is larger (strange?). There is no significant improvement in the coastal regions between the two products.</p>	<p>Thank you for pointing this out. We have revised Figures 2-4. In the previous version, we made a technical error in co-locating the two model outputs to coastal SOCAT grid cells so statistics were not precise enough. The revision slightly modifies RMSD and r<sup>2</sup> values over all regions but does change our conclusion.</p>	
<p><b>MC1.13. Lines 404-423</b></p> <p><i>“Analyzing the eight station time series, we have found that data have been sampled within a few days with an average offset of about a week from the month center. At</i></p>	<p>We have demonstrated the better performance of FFNNr025 in terms of intra-seasonal to interannual variability of coastal sites (Sutton et al., 2019). By increasing the model resolution by 16-fold, this study partly resolves the spatial sampling bias from pCO<sub>2</sub> observations (lower RMSD and higher r<sup>2</sup> for</p>	

these coastal sites, the temporal standard deviation from monthly averages of  $p\text{CO}_2$  ( $\sigma_i^{p\text{CO}_2}$ ) exceeds measurement errors (2  $\mu\text{atm}$ , Sutton et al., 2019).  $\sigma_i^{p\text{CO}_2}$  ranges from 20.12  $\mu\text{atm}$  at GREYREEEF to values as large as 65.6  $\mu\text{atm}$  at CAPEARAGO or 69.98  $\mu\text{atm}$  at FIRSTLANDING. The monthly average of  $p\text{CO}_2$  might not be adequately represented by discreet samples at sites with a large temporal standard deviation of  $p\text{CO}_2$ . The misfit between the monthly reconstruction and discreet observations is exacerbated in dynamical coastal environments and might explain in part the large RMSD of reconstructions of monthly coastal  $p\text{CO}_2$  (e.g., GREYREEEF: 38.34  $\mu\text{atm}$ , CAPEARAGO: 79.86  $\mu\text{atm}$ , FIRSTLANDING: 77.32  $\mu\text{atm}$ ) for the r025 reconstruction. The RMSD is mostly lower for the FFNN reconstruction at 0.25° resolution compared to the FFNN at 1° resolution by 2.11  $\mu\text{atm}$  (CCE2) to 23.32  $\mu\text{atm}$  (COASTALMS). Similarly, r2 increases between 7%-23% at higher resolution. Overall, seasonal to interannual variations of coastal-ocean  $p\text{CO}_2$  are better reproduced in the

the higher resolution) although large model-observation mismatches still persist. As replied to comments **MC1.9.** and **MC1.11.**, the sparsity of data samples (biases from observation locations to the grid cell center about 0.34° ± 0.14° as reported in Sabine et al., 2013) and high variability of coastal  $p\text{CO}_2$  (e.g., 470  $\mu\text{atm}$  in a distance of 0.5km; see in Chavez et al., 2018 and Feely et al., 2008) would draw the conclusion that much higher resolution or extensions of observing system are necessary to fully capture coastal  $p\text{CO}_2$ .

Temporal data sampling bias should be considered as a great source of uncertainty contributing to large model-observation mismatch even though model spatial resolution is getting finer. We illustrate this through Figure 5 with the corresponding analysis being in the paragraph (Lines 437-439\*) quoted by **Review 1.**

The key discussion we found is as follows

*“Analyzing the ~~eight~~ station time series, we have found that data have been sampled within a few days with an average offset of about a week from the month center. At these coastal sites, the temporal standard deviation from monthly averages of  $p\text{CO}_2$  ( $\sigma_i^{p\text{CO}_2}$ ) exceeds ~~measurement~~ analytical errors (2  $\mu\text{atm}$ , Sutton et al., 2019)”.*

With the low number of observations and high variability of  $p\text{CO}_2$  over these stations, it would not be statistically sufficient to refer to the temporal mean of instantaneous observations as a representation of monthly averages. We then provide evidence that the large values  $\sigma_i^{p\text{CO}_2}$  at time series stations (e.g., GREYREEEF: 20.12  $\mu\text{atm}$ , CAPEARAGO: 65.6  $\mu\text{atm}$ , FIRSTLANDING: 69.98  $\mu\text{atm}$ ) correspond to high RMSDs (e.g., GREYREEEF: 38.34  $\mu\text{atm}$ , CAPEARAGO: 79.86  $\mu\text{atm}$ , FIRSTLANDING: 77.32  $\mu\text{atm}$ ).

About the effect of model-observation bias of SST on the reconstruction skills of  $p\text{CO}_2$ , we refer to our reply to the general comment **GC1.3** above.

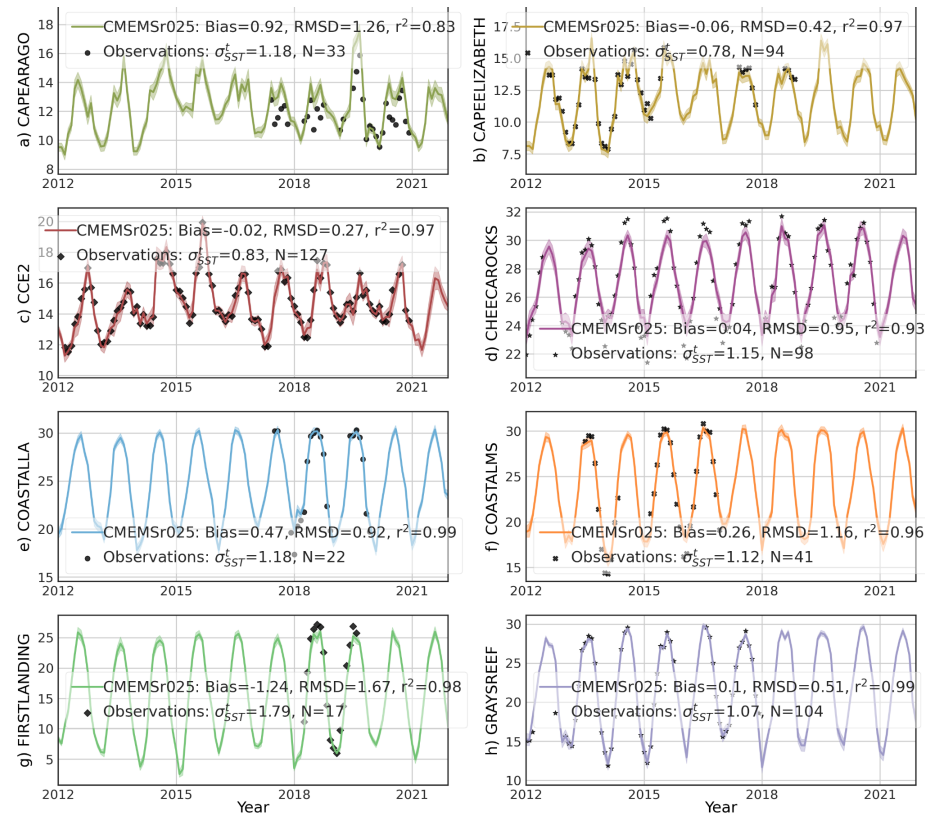
\*Lines 403-405 in the initial manuscript



reconstruction at 0.25° resolution (Fig. 5).”

Here, it becomes evident that comparing monthly reconstructions with point values in coastal areas of high variability results in very low predictive ability on the part of the product produced. As indicated in the general comment, this should be evaluated considering the variability of SST and SSS in the study area because in this way the biases that the CMEMS product has to reproduce point values from monthly mean values are being transferred to pCO<sub>2</sub>. The aforementioned increases in r<sup>2</sup> are relatively small if we consider the important biases involved, which in some products even increase as the resolution improves, as in FIRSTLANDING or CHEECAROCKS.

Also illustrated in Figure MC.1.13., the bias between SST<sup>CMEMS</sup> and in situ SST from Sutton et al’s time series is always lower than 0.5°C for many stations. A bias in SST would not be the dominant source of high reconstruction errors at these stations.



**Figure MC.1.13.** Time series of surface ocean SST (°C) at coastal observing stations<sup>\*\*</sup>: CMEMS SST estimate (curve), associated 1σ-uncertainty (envelope), and monthly average of in situ observations (point). CMEMS reanalysis data at 0.25° (r025) resolutions are co-located to in situ observations provided by Sutton et al. (2019). Statistics include number of months with observations (N), Bias, RMSD, and r<sup>2</sup>. σ<sub>i</sub><sup>SST</sup> stands for temporal standard deviation from monthly averages of SST observations.

<sup>\*\*</sup> See Table A2 and Fig. A1b in the manuscript.

<p><b>MC1.14. Lines 437-438</b></p> <p><i>"The largest model uncertainty (<math>\sigma &gt; 30 \mu\text{mol kg}^{-1}</math>) is computed nearshore and surrounding oceanic islands, a feature inherited from input uncertainty associated with the CMEMS salinity product (Fig. A8a)." This described here is very relevant. In fact, it would be necessary to show graphically the correlation between the uncertainty in TA and SSS in the CMEMS product in both the coastal and oceanic domains. Possibly it shows a very relevant correlation. A similar should be done with the uncertainties of <math>p\text{CO}_2</math> and SST in the CMEMS product.</i></p>	<p>Total alkalinity (<math>A_T</math>) is predominantly controlled by the processes that govern sea surface salinity (SSS) (Broecker and Peng, 1982; Millero et al, 1998). The typical relationship between these two variables is linear and can be estimated at a high precision (Lee et al, 2006; Carter et al, 2018; Broullon et al, 2019). From the statistical point of view, the distribution of <math>A_T</math> uncertainty is generally driven by SSS uncertainty: <math>A_T</math> uncertainty increases as SSS uncertainty increases (see Figure <b>MC1.17.</b>).</p> <p>To the contrary, <math>p\text{CO}_2</math> is characterized by multiple physical, biological, and chemical processes. Uncertainties from many input data products thus contribute to <math>p\text{CO}_2</math> uncertainty estimates. Drivers of <math>p\text{CO}_2</math> uncertainty are not analyzed as input uncertainty has not been fully quantified or published so far for many environmental variables.</p>	
<p><b>MC1.15. Lines 451-465</b></p> <p><i>"The reconstruction of AT distributions relies on LIAR coefficients fit with GLODAPv2 data (Olsen et al., 2016) covering the years before 2015. These data are also part of the latest version GLODAPv2.2022 (Lauvset et al., 2022). They do therefore not correspond to an independent dataset for the evaluation data of the CMEMS-LSCE reconstruction. To overcome this limitation, reconstructions of AT and DIC are</i></p>	<p>First of all, Lines 488-496* quoted by <b>Reviewer 1</b> describes the evaluation of our data product of <math>A_T</math> and DIC and time series of in situ observations. This complements the assessment with GLODAP data. As opposed to the interpretation by <b>Reviewer 1</b>, these lines do not contain any analysis about "how a large part of the discrepancies between the TA and DIC reconstruction is due to the discrepancies in SSS and SST of the CMEMS product and observations". But we have revised the following sentence to have a better sense (other modifications follow the revisions according comment <b>GC2.2.</b>)</p> <p><i>"To <del>overcome this limitation</del> accomplish a cross-validation, reconstructions of <math>A_T</math> and DIC are compared to observations for <del>Eulerian</del> eight time series</i></p>	<p>*Lines 451-465 in the initial manuscript</p>

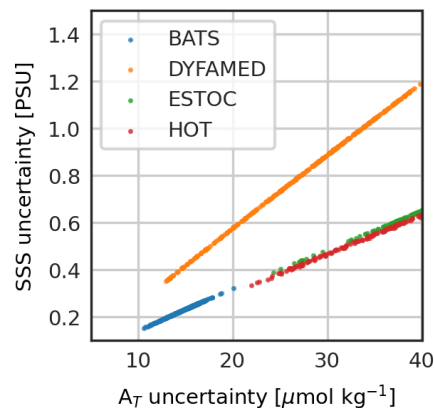


<p>compared to observations for Eulerian time series stations: BATS, DYFAMED, ESTOC, and HOT (see Table 3 and Fig. A1b for data sources and station locations). Figure 7 illustrates the comparison between monthly time series of AT and DIC extracted from the CMEMS-LSCE datasets and measurements at these long-term monitoring sites”. These lines and Figure 7 show again how a large part of the discrepancies between the TA and DIC reconstruction is due to the discrepancies in SSS and SST of the CMEMS product, indicating that the reconstruction is not well done. In the case of the DYFAMED station it is very noticeable and contrasts that other products such as climatologies like those cited in the article (Lauvset et al. 2016; Broullón et al. 2019) do not show bias as high as the reconstruction performed here.</p>	<p>stations: <i>AWIPEV, BATS, DYFAMED, ESTOC, and HOT, ICELAND, IRMINGER, and KERFIX</i> (see Table 3 and Fig. A1b for data sources and station locations).”</p> <p>In Figure 7 we illustrate both the relatively good and poor reconstructions at long-term time series of observations. Note that, Lauvet et al, (2016) and Broullon et al, (2019) provided climatologies of <math>A_T</math> and DIC and associated mapping errors. As the climatology is smoother than the monthly fields (with intra- to interannual variability) proposed in this study, the errors reported in the previous studies are evidently smaller than those presented here. Their magnitudes are not comparable. In addition, Lauvet et al, (2016) and Broullon et al, (2019) did not evaluate the reconstruction at DYFAMED as mentioned by <b>Reviewer 1</b>. Please kindly refer to our reply to <b>SC2.14</b>. for a profound analysis of high model-data mismatch of <math>A_T</math> at DYFAMED.</p>	
<p><b>MC1.16. Line 473</b>  <i>“The lowest prediction skill of temporal variability is obtained for ESTOC. Particularly, seasonality to multiyear variations in DIC are predicted at <math>r^2=0.47</math></i></p>	<p>We thank the reviewer for highlighting our good estimates of <math>A_T</math> and DIC at ESTOC in terms of RMSD. We skipped this element in the initial manuscript. We have added it in Lines 515-520*:</p>	<p>*Line 473-476 in the initial manuscript</p>

<p>for ESTOC compared to <math>r^2 &gt; 0.7</math> for BATS and HOT.” This is not correct. The regression coefficient is not the only criterion for assessing predictive ability. In this case the variability observed at ESTOC is lower than at BATS and HOT, so a lower <math>r^2</math> does not mean lower skill. In fact, the RMSD at is the lowest of all the stations evaluated in TA. In terms of DIC the three stations show similar RMSD.</p>	<p><i>“Despite showing good estimates of <math>A_T</math> and DIC in RMSD at ESTOC, temporal variability of observations are reconstructed at the lowest <math>r^2</math>.<del>The lowest prediction skill of temporal variability is obtained for ESTOC.</del> Particularly, seasonality to multi-year variations in DIC are predicted at <math>r^2 = 0.47</math> for ESTOC compared to <math>r^2 &gt; 0.7</math> for <b>AWIPEV, ICELAND, IRMINGER, BATS and HOT</b>. Over all the stations, the model underestimates temporal changes of <math>A_T</math> (Fig. 7a; BATS: <math>r^2 = 0.33</math>, DYFAMED: <math>r^2 = 0.12</math>, ESTOC: <math>r^2 = 0.03</math>, HOT: <math>r^2 = 0.32</math>) which can be attributed to the large discrepancy in variability between in situ measurements and the CMEMS time series of salinity (Fig. A10a; BATS: <math>r^2 = 0.33</math>, DYFAMED: <math>r^2 = 0.19</math>, ESTOC: <math>r^2 = 0.03</math>, HOT: <math>r^2 = 0.35</math>).”</i></p> <p>We indeed analyze both RMSD and <math>r^2</math> throughout the manuscript. To be precise, <math>r^2</math> is not ‘regression coefficient’ but the determination coefficient. This metric allows evaluating the model predictive ability in temporal variations of the variables of interest. As shown in Eq (11) in the manuscript, <math>r^2</math> is the model-data covariance normalized with the temporal variability of <math>A_T</math> and DIC reproduced by FFNN and observed at each station. Therefore, <math>r^2</math> values at BATS and HOTS are comparable to the one at ESTOC even though the two former stations show higher temporal variability of <math>A_T</math> and DIC. Our analysis in Lines 515-520* holds true.</p>	
<p><b>MC1.17. Line 478</b>  <i>“Model uncertainty (<math>1\sigma</math>-envelop) of monthly <math>A_T</math> and DIC estimates (Fig. 7a) is also inflated somewhat proportional to the CMEMS salinity product uncertainty (Fig. A10a).”</i> Evidently. A figure showing that would be useful. That is why including this product in the LIAR training phase for</p>	<p>We include here Figure MC2.17 showing the relationship of <math>A_T</math> and SSS uncertainty. It is indeed well-known that SSS is the dominant driver of <math>A_T</math> and same for their uncertainty. Please refer to our replies to comments <b>MC1.14.</b> and <b>GC.1.3.</b> for a further analysis.</p>	

TA does not help to obtain the best possible reconstruction.

**Figure MC2.17.** Scatter plot of monthly uncertainty values of a) CMEMS-LSCE  $A_T$  (Figure 7) against CMEMS SSS (Figure A10) at four stations.



**MC1.18. Line 527**

*“The reconstructed pH time series reproduce measurement variability with relatively high correlation,  $r^2$  in [0.21,0.69], that reinforces the reliability of CMEMS-LSCE pH”.*

It does not seem that the level of correlation obtained with this reconstruction is significant with such low levels of  $r^2$ . Additionally, the fact that there is no discussion in the article where these levels are compared with other products even if they are only climatic such as those of Takahashi et al. 2014, or others cited in the article for AT and DIC.

Similar to  $p\text{CO}_2$ , measurements of pH are subject to undersampling (see also our reply to comment **MC1.11.**). The data density is even lower for pH (see Figures 8 and A3 and Table A3). Remember that the comparison is between monthly mean estimates of pH and the mean of observations which are normally available within some days to a week. Temporal data sampling bias should be considered as a great source of uncertainty attributed to large model-observation mismatch (and thus moderate values of  $r^2$ ) even with a finer model spatial resolution. We show that through Table A3 and an analysis is given in the paragraph (Lines 558-576\*).

Please note that the published data cited by Reviewer 1 are climatology. Strictly speaking, statistical evaluations between the climatological product and in situ observation data (sampled with very low frequency) is not robust (see also our reply to the general comment **GC1.2.**)

\*Lines 514-530 in the initial manuscript

<p><b>MC1.19. Line 576</b></p> <p>“Conclusions and Discussion” It should be "Discussion and Conclusions" But on the other hand the discussion is made not in terms of the assessment of the quality of the reconstruction of the product but in terms of the results in terms of ocean acidification.</p>	<p>Thank you for your suggestion. We have changed Section “Conclusions and Discussion” to “Summary” as proposed by <b>Reviewer 2</b> (comment <b>SC2.16.</b>).</p>	
<p><b>MC1.20. Line 594</b></p> <p><i>“In comparison to CMEMS-LSCE at monthly and 1° resolutions (Chau et al., 2022b), the reconstructions over coastal areas are improved at higher resolution (Figs. 2-4).”</i> This is not demonstrated in the article. The reduction in RMSD between the two products is very small or marginal.</p>	<p>We modify the sentences below (Lines 640-643*) to better summarize our results:</p> <p><i>“In comparison to CMEMS-LSCE at monthly and 1° resolutions (Chau et al., 2022b), the reconstructions over coastal areas are improved at higher resolution—(Figs. 2-4).—Furthermore, †The monthly, 0.25° reconstruction outperforms its 1° counterpart in reproducing horizontal and temporal gradients of pCO<sub>2</sub> over a variety of oceanic regions as well as at nearshore time series stations (Figs. 2-5).”</i></p> <p>The improvement in terms of global metrics is marginal but we still gain advantages when increasing model spatial resolution from 1° to 0.25° (e.g. better capturing horizontal and temporal gradients). In this manuscript, we also try to address the question why increasing the spatial resolution by 16-fold does not impressively reduce model-observation discrepancy. Please kindly refer to our reply to comments <b>MC1.9.</b> and <b>MC1.11.</b> for a comprehensive explanation.</p>	<p>*Lines 594-597 in the initial manuscript</p>
<p><b>MC1.21. Line 609</b></p> <p>Line 609 “The spatial distribution of long-term mean 1<math>\sigma</math>-uncertainty estimates</p>	<p>The reviewer is correct, but, as discussed in Chau et al (2022b), the ensemble spread is a good proxy for the reconstruction uncertainty.</p>	

<p>(Figs. 1b, 6cd, and 9cd) indicates higher confidence levels for open-ocean estimates than over the coastal sector”. This is very unrepresentative of product quality since it represents there producibility of the 100 FFNN but does not evaluate the RMSD between input and reconstructed data.</p>		
<p><b>MC1.22. Table 7</b> Both pCO<sub>2</sub>, A<sub>T</sub> and DIC quantities should not have decimal places (mean, RMSD).</p>	<p>pCO<sub>2</sub>, A<sub>T</sub>, DIC, and their reconstruction errors/uncertainty have been reported with 1-2 decimals in the previous studies (Landschuter et al 2013, 2020; Denvil et al 2019, Gregor et al. 2019, 2021, Chau et al. 2022b). In this revision, we reduce the decimals from 2 to 1 for pCO<sub>2</sub>, A<sub>T</sub>, and DIC. The modification has been applied for Tables, Figures, and texts involving these variables.</p>	
<p><b>MC1.23. Line 655</b> No comparisons with other reconstructions like MODO-DIC of Keppler et al. 2020, or AT from Broullon et al. 2019 or Lee et al. 2006.</p>	<p>Please see our answer to comments <b>GC1.2.</b> and <b>MC1.15.</b></p>	

## Replies to comments by Reviewer 2

### General comments by Reviewer 2 (GC2)

**GC2.1.** *The authors reconstructed 0.25-degree monthly full carbonate system variables during the period 1985-2021 based on surface ocean observation data. Distributions of pCO<sub>2</sub> were reconstructed based on the machine learning method established by the authors (Trang-Chau et al. 2022) and those of TA were based on the LIAR method (Carter et al. 2016; 2018). While few reconstructions of full carbonate system variables are available at this moment, a comprehensive understanding of global surface ocean pH distributions is essential for monitoring ocean acidification, which is related to the SDG indicator 14.3.1. This study can enhance researches on the global carbon cycle as well as provide critical information to policymakers and stakeholders. I think this study has sufficient value to be published in this journal, but major concerns listed below should be addressed appropriately. I would like to encourage the authors to improve the study and revise the manuscript for better understanding.*

#### **Authors:**

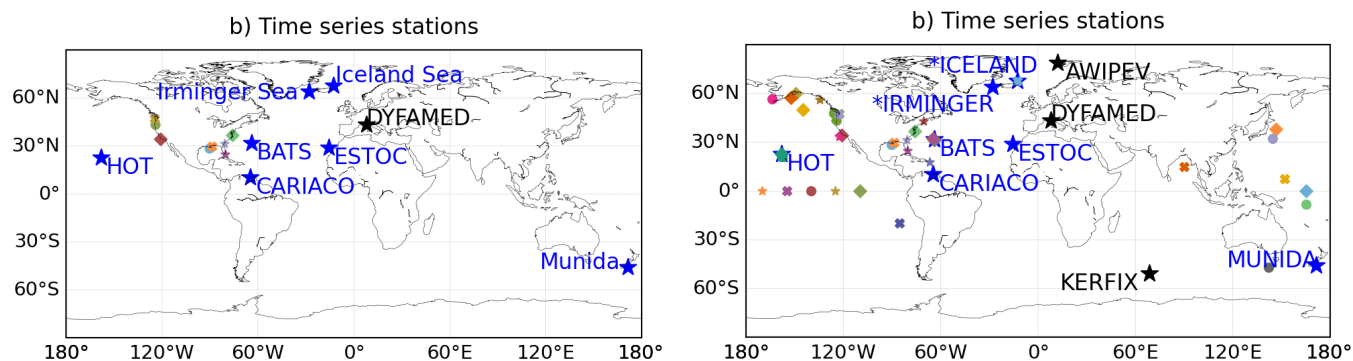
We are grateful for **Reviewer 2**'s positive evaluation and constructive comments which help us to improve our manuscript. Please kindly find our replies to address his/her concerns below.

%%%%%%%%%

**GC2.2.** *The concept of this study itself is not novel, and the assessment of uncertainty in the reconstructed fields and the validation of the method become important. The authors derived uncertainty distributions in reconstructed parameters from the spread of 100 model ensemble. They also demonstrated the validity of the method by comparing the result of this study with observation data that were not used for learning and those of time-series points. The time-series used in this study are biasedly located in the subtropical region, so comparing their data with the results of this study does not seem a good indicator of uncertainty. For validation of this method, the authors must take a comparison with other reconstruction(s) into account, if needed.*

#### **Authors:**

Further to the reviewer’s suggestion, we have added comparisons to 38 time series stations located outside the subtropics (the tropics, the subpolar sector, and the Southern Ocean). The results appear in Figure A1b, Tables A3 and A4. They confirm the reliability of CMEMS-LSCE datasets (see our analysis in Lines 421-447, 488-522, 558-576 in the revised manuscript attached at the end of this document).



**Figure GC2.2. Revised Figure A1b (right):** *b) Location of time series stations recording in situ observations used in data evaluation (Table 2): blue stars for ocean acidification (Bates et al., 2014), black stars for  $A_T$  and DIC (Metzl et Lo Monaco, 1998; Coppola et al., 2021; Gattuso et al., 2023), and other coloured scattered objects for  $pCO_2$  and pH (Sutton et al., 2019). Asterisk (\*) marks the two stations with  $A_T$  and DIC observations (Olafsson et al., 2010) available for assessments.*

Suggestions from the two reviewers about an intercomparison with other products are interesting but would bring us well outside the scope of our study if we take them carefully enough. Such intercomparison between data products deserve proper investigations on (1) the discrepancy in mapping methods, input data resource, and the ratio of training and validation datasets used in model fit, (2) uncertainty from post-processing applied for some products (e.g., filtering, smoothing, calibration), (3) the normalization of different data covering periods, and (4) quality of evaluation (or reference) data; e.g., data paucity should be one major concern to evaluate seasonal cycle reconstructions. For  $pCO_2$ , the evaluation of multiple products including CMEMS-LSCE at  $1^\circ$ , monthly resolutions (Chau et al., 2022) was done in the previous studies (Hauck et al., 2020; Gregor et al., 2021; Friedlingstein et al., 2022), and it is well confirmed that the quality of CMEMS-LSCE is in line with the others. This manuscript investigates an upgrade of multi-year reconstructions of  $pCO_2$  and other carbonate system variables by increasing spatial resolution from  $1^\circ$  to  $0.25^\circ$ , that has not been done in the previous studies.

%%%%%%%%%

**GC2.3.** *The authors use external SST and SSS instead of those incorporated in the datasets in the learning process. This seems unusual because the oceanographic condition represented by temperature and salinity considerably affects the ocean biogeochemistry in the observed area. If the authors think the use of external SST/SSS to be essential, they must demonstrate that the impact of differences between external SST/SSS and those in the datasets is negligible.*

**Authors:**

Our reconstructions require gridded SST and SSS datasets without any gaps, which is not available from SOCAT. We therefore use other data sources. Based on statistical assessments, the difference between CMEMS SST and SOCAT SST (or in situ observations) is relatively small (see Figures **GC1.3.** and **MC1.13.** in this document and also Figure A10 in the manuscript). Besides, no quality control has been done for SOCAT salinity. There are some cruises in SOCAT with no salinity data and SOCAT has alternatively used SSS from the World Ocean Atlas (WOA) to recalculate CO<sub>2</sub> fugacity (see Pfeil et al, 2013). We have not changed the method.

%%%%%%%%%

**GC2.4.** *In addition, the manuscript seems to contain unnecessary sentences and be lengthened. Shortening the manuscript will increase readability.*

**Authors:**

Thank you. We have shortened the manuscript, in particular based on the reviewer's specific comments.

%%%%%%%%%



### **Specific comments by Reviewer 2 (SC2)**

<b>Reviewer's comments</b>	<b>Replies from Authors</b>	<b>Notes</b>
<b>SC2.1. Introduction</b> This section seems too long and needs to be shortened.	We have revised the manuscript and removed part of the unnecessary sentences.	
<b>SC2.2. L71</b> not only “extrapolate” but also “interpolate”.	Thank you. “interpolate” was added.	
<b>SC2.3. L91 Table 1 and Appendix A</b> Table 1 only shows six carbonate system variables and is not necessary. Reference to them in the text is enough. In the same context, Appendix A is also unnecessary because it only contains general explanations of carbonate system variables as written in, e.g., Dickson et al. 2007.	We removed both Table 1 and Appendix A.	

<p><b>SC2.4. L95-131</b> All or a part of these explanations had better be transferred to the beginning of the “3 Reconstruction method” section.</p>	<p>These lines were modified accordingly.</p>	
<p><b>SC2.5. L135</b> Which were used, sea surface height anomaly (SLA) or sea surface dynamic height (SLA+MDT)? Please clarify.</p>	<p>As defined by data providers of CMEMS SSH, we revised the text as follows: “<i>sea surface height above geoid</i>”.</p>	
<p><b>SC2.6. L155</b> SOCAT’s full name was already mentioned in L66.</p>	<p>We replaced the full name with its abbreviation.</p>	
<p><b>SC2.7. L160-165</b> Using global 0.25 deg binned data derived from SOCAT cruise data is a usual way, even though using 1 deg binned data does not significantly affect the result.</p>	<p>Using a global 0.25° dataset gridded from SOCAT underway measurements for FFNN model training is indeed our ultimate goal. We have contacted SOCAT experts to investigate further how to grid measurements into 0.25° open-ocean datasets knowing that quality control of measurements is critical.</p>	
<p><b>SC2.8. L213</b> It should be clarified how you dealt with longitude and latitude parameters.</p>	<p>To preserve the continuity of longitude at 0°, we have applied the sine and cosine functions to that coordinate. Hence, our global maps of carbonate variables (e.g. Figures 1, 6, 9) do not show discontinuity at the prime meridian. The sine is also used to transform latitude. Data transformation of predictor variables is explicitly presented in a sequence of preceding studies</p>	

	<p>for the CMEMS-LSCE-FFNN model development (Denvil-Sommer et al., 2021, Chau et al., 2022). In the first manuscript version, we avoided repeating part of the data processing and model description from the previous studies. As the readers would concern, we have called back this information in the revision (Lines 162-164):</p> <p><i>“The sine function is applied to convert latitude while both the sine and cosine are used to transform longitude to conserve their periodical behaviors.”</i></p>	
<p><b>SC2.9. Table 4</b></p> <p>This table contains RMSDs and coefficients of determination, and the name “skill score” is not appropriate. RMSDs of r025 are not significantly different from those of r100 according to Table 4, and therefore the authors should not emphasize an improvement of the prediction skill. The results only show that a fine-scale reconstruction was achieved with no adverse effect.</p>	<p>We have changed “Skill scores” to “<b>Evaluation statistics</b>” in the Table caption.</p> <p>Table 3* shows a marginal improvement from r100 to r025 in terms of global evaluation metrics. For the open ocean, we expect to obtain similar skill scores for both FFNN models as the spatial autocorrelation of open-ocean <math>p\text{CO}_2</math> is estimated within <math>400\pm 250</math> km (Jones et al., 2012) and the same SOCAT <math>1^\circ</math>-open-ocean dataset was used in model fitting. As also reviewed in Chau et al., (2022), <math>p\text{CO}_2</math> over the coastal ocean is characterized with high variability at small scales. For instance, <math>p\text{CO}_2</math> levels can vary with a horizontal gradient as large as <math>470 \mu\text{atm}</math> over a distance of less than 0.5 km (Chavez et al., 2018; Feely et al., 2008). Statistical models would need a spatial resolution much finer than <math>0.25^\circ</math> (25 km) and a temporal resolution higher than monthly in order to capture such high variability in surface ocean <math>p\text{CO}_2</math> present in observations (see also Bakker et al., 2016; Laruelle et al., 2017) (see our discussion in Section Summary**).</p> <p>Apart from Table 3*, benefits of increasing the model spatial resolution from <math>1^\circ</math> to <math>0.25^\circ</math> are also demonstrated in Figures 2-4 with analyses shown in Lines 414-420***</p>	<p>*Table 4 in the initial manuscript (Table 1 has been removed as suggested by <b>Review 2</b>, comment <b>SC2.2.</b>)</p> <p>** Section Conclusion and Discussions in the initial manuscript</p>

	<p><i>“The two FFNN reconstructions (r025 and r100) share similarities in overall structures of pCO<sub>2</sub> over the coastal-open-ocean continuum (Figs. 2-4). However, the higher spatial resolution outperforms its lower resolution counterpart in reproducing fine-scale features of pCO<sub>2</sub> in the transition from nearshore regions to the adjacent open ocean. The increase in model spatial resolution translates into a greater spatial coverage of the continental shelves such as Labrador Sea, Northern Europe, and Sea of Japan (Fig. 3), and thus an increase in the number of data over the coastal domain. The increase in spatial resolution allows a gain in prediction probability of pCO<sub>2</sub> variations on the order of roughly 2% over the Eastern Boundary Currents to 8% over the Western South Atlantic (Figs. 2-3b).”</i></p> <p>This study also points out temporal data sampling bias as a source of uncertainty that would highly constrain model reconstruction skills. Based on the assessment at station time series (Figure 5 and Table A3), we found that in situ observations have been sampled with low frequency and the bias of sampling date is about a week from a month center. With the low number of observations and high variability of pCO<sub>2</sub> (20.12 to 69.98 μatm) over these stations, it would not be statistically sufficient to refer to their temporal mean as a representative of monthly averages. A large model-data deviation would be retained even if we increase spatial resolution (see text in Lines 421-447 for further analysis).</p>	<p>***Lines 387-394 in the initial manuscript</p>
<p><b>SC2.10. Fig 2-4</b> The results from the two methods, r100 and r025, have almost the same structure. Please explain the reason why the authors focused on the comparison of them.</p>	<p>As expressed in Line 384-393*, the motivation for a comparison between r100 and r025 in Figures 2-4 to show improvements of horizontal gradients in the higher resolution over different oceanic conditions. The three figures respectively present results in:</p>	<p>*Line 358-367 in the previous version</p>

	<ul style="list-style-type: none"> <li><input type="checkbox"/> permanent Eastern Boundary current upwelling systems with relatively high <math>p\text{CO}_2</math>,</li> <li><input type="checkbox"/> regions characterized by low <math>p\text{CO}_2</math> values driven by cold water temperatures and strong biological production,</li> <li><input type="checkbox"/> other regions either under the influence of strong river runoff or monsoon-driven upwelling.</li> </ul> <p>It is expected that the two models with different resolutions share the same large-scale structure described above. In Line 413-420** (below), we further analyse the benefits obtained with the higher resolution.</p> <p><i>“The two FFNN reconstructions (r025 and r100) share similarities in overall structures of <math>p\text{CO}_2</math> over the coastal-open-ocean continuum (Figs. 2-4). However, the higher spatial resolution outperforms its lower resolution counterpart in reproducing fine-scale features of <math>p\text{CO}_2</math> in the transition from nearshore regions to the adjacent open ocean. The increase in model spatial resolution translates into a greater spatial coverage of the continental shelves such as Labrador Sea, Northern Europe, and Sea of Japan (Fig. 3), and thus an increase in the number of data over the coastal domain.”</i></p>	<p>**Line 387-391 in the previous version</p>
<p><b>SC2.11. L374 Fig. 3</b></p> <p>RMSD for the Sea of Japan is suppressed by using data in the subtropical regions (Tsushima warm current area and Kuroshio area) which generally can be estimated more easily. The RMSD must be calculated from data restricted north of the subtropical front.</p>	<p>Thank you for the comment. However, we should use all the available SOCAT data over the Sea of Japan to evaluate RMSD setting the assessment consistent with the other coastal regions. The evaluation over specific sub-basins can be considered in a regional study of <math>p\text{CO}_2</math> variability.</p>	

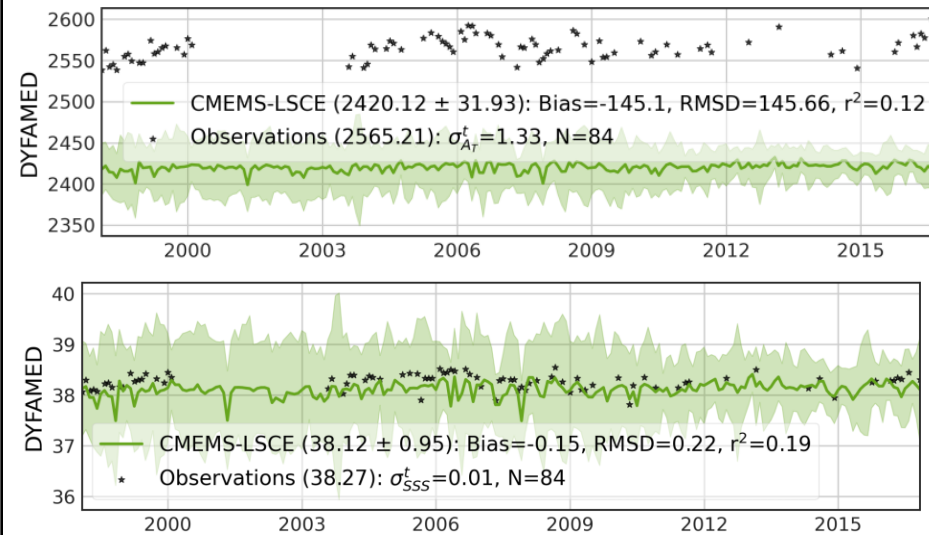
<p><b>SC2.12. L403-414</b></p> <p>The discrepancy between the estimated and observed pCO<sub>2</sub> not only originated from the timescale but also from the method itself. The method cannot express short-term phenomena inherently because it used external SST and SSS instead of those incorporated in the datasets in the learning process.</p>	<p>We agree with the reviewer. As explained above (replies to comments <b>GC1.3.</b> and <b>GC2.3.</b>), the use of gridded SST and SSS is a necessity for the reconstruction and there is no significant discrepancy between SOCAT (or in situ) data and CMEMS gridded data.</p>	
<p><b>SC2.13. 5.2 Total alkalinity and dissolved inorganic carbon</b></p> <p>In the method of this study, discrepancies in estimated and observed DIC were initially derived from pCO<sub>2</sub> and TA estimation and propagated via carbonate system calculations. The discussion on uncertainty should be written along with such a concept.</p>	<p>Thank you for this point. We have added the discussion in the revised manuscript (see Lines 471-474*):</p> <p><i>“<del>The largest model uncertainty</del> DIC uncertainty is computed through CO2SYS error propagation with reconstruction uncertainties of pCO<sub>2</sub> and A<sub>T</sub> set as inputs. The largest values <del>model uncertainty</del> (<math>\sigma &gt; 30 \mu\text{mol kg}^{-1}</math>) appear <del>is computed nearshore and surrounding oceanic islands</del> (Fig. 6d). A similar feature is found on the field of A<sub>T</sub> (Fig. 6c), <del>a feature inherited from input uncertainty associated with the CMEMS salinity product</del> (Fig. A8a).”</i></p>	<p>*Lines 437-438 in the initial manuscript</p>
<p><b>SC2.14. L467-469</b></p> <p>The authors attributed a large <math>\sigma</math> of DYFAMED estimates to a limited number of observations in the Mediterranean, but GLODAPv2 includes alkalinity measurement data in the Mediterranean. Schneider et al. 2007 successfully derived</p>	<p>There exists indeed a few observations over the surface Mediterranean Sea in GLODAPv2.2022 (Lauvset et al., 2022) used for model evaluation (see f.i. Figure 8 in the manuscript), and even lower data density in GLODAPv2 (Olsen et al., 2016) used for LIAR coefficients fits (Carter et al. 2018). As illustrated in Figure <b>SC2.14.</b>, the bias between CMEMS SSS and observations is tiny compared to that of LIAR A<sub>T</sub> and observations. The paucity of</p>	

the salinity-alkalinity relationship. The discrepancy in DYFAMED seems to be attributable to salinity discrepancy only.

GLODAPv2 data over the Mediterranean Sea and the distinction in AT-SSS relationship over this region from other basins would lead to a biased LIAR estimate of  $A_T$  ( $-145.1 \mu\text{mol kg}^{-1}$ ). Schneider et al., (2007) derived the estimates of  $A_T$ -SSS relationship by using local observations, but the estimated  $A_T$  is also subject to a large error range ( $\pm 114.94 \mu\text{mol kg}^{-1}$ ), see Eq 1 quoted below:

$$A_T = 73.7 (\pm 3.0) \cdot S - 285.7 (\pm 114.94) \mu\text{mol kg}^{-1}$$

Figure SC2.14. Monthly time series of  $A_T$  (Figure 7) and SSS (Figure A10) at DYFAMED.



We have added one sentence in Lines 508-510\* for clarification.

*“Although the bias between reanalysed SSS and observations (Fig. A10) is relatively small ( $-0.15 \mu\text{mol kg}^{-1}$ ), LIAR (Carter et al., 2018) was trained on GLODAPv2 (Olsen et al., 2016) including  $\alpha$ -only few observations in this area. The distinct relationship between alkalinity and salinity prevailing in*

\*Lines 467-469 in the initial manuscript

	<i>the Mediterranean Sea is likely not reproduced by LIAR leading to an underestimation of <math>A_T</math> and a systematic bias to DIC at DYFAMED (Fig. 7)."</i>	
<p><b>SC2.15. L539-540</b></p> <p>I think that SDG indicator 14.3.1, "Average marine acidity (pH) measured at agreed suite of representative sampling stations", is worth mentioning here. Global mean pH based on observation can be a proxy for the indicator. In addition, it is also valuable information that the global mean pH becomes 8.0 with one decimal place, not 8.1 often said.</p>	<p>Thank you. In lines 686-689, we have mentioned the SDG 14.3.1 indicator for ocean acidification.</p> <p><i>"The global maps of CMEMS-LSCE pH, <math>\Omega</math>, and their trend estimates would be potential indicators for ocean acidification along with the SDG 14.3.1 - "Average marine acidity (pH) measured at agreed suite of representative sampling stations" (<a href="https://sdgs.un.org/goals/goal14">https://sdgs.un.org/goals/goal14</a>: last access 31/07/2023)."</i></p> <p>However, the global mean CMEMS pH over 1985-2022 is about 8.082 (Table 6*). With one decimal, it becomes 8.1, the same value as reported previously.</p>	<p>*Table 7 in the initial manuscript (Table 1 has been removed as suggested by <b>Review 2</b>, comment <b>SC2.2</b>.)</p>
<p><b>SC2.16. 6. Conclusion and discussion</b></p> <p>This section had better be titled "Summary". It does not seem to include discussion.</p>	<p>Thank you. We have modified the section title as suggested.</p>	



## Other changes:

- Data repository: the following sentence is added at the end of the Introduction (Lines 136-138) to make data repository visible to the users (details of data access can be found in Section Data availability):  
*“The high-resolution data product described in this manuscript (netCDF format) can be accessed via repository under data DOI: 10.14768/a2f0891b-763a-49e9-af1b-78ed78b16982.”*
- Contribution of GLODAP for this study: we add one sentence in Acknowledgement.  
*“The Global Ocean Data Analysis Project (GLODAP, [www.glodap.info](http://www.glodap.info), last access: 21 August 2023) provides access to ocean surface-to-bottom quality controlled data of carbonate system variables collected through international cruises.”*
- Typo errors / references: they are corrected / updated in the revised manuscript.

## References

1. Bakker, D., Alin, S., Becker, M., Bittig, H., Castaño-Primo, R., Feely, R. A., Gritzalis, T., Kadono, K., Kozyr, A., Lauvset, S. K., Metzl, N., Munro, D., Nakaoka, S.-i., Nojiri, Y., O'Brien, K., Olsen, A., Pfeil, B., Pierrot, D., Steinhoff, T., Sullivan, K., Sutton, A., Sweeney, C., Tilbrook, B., Wada, C., Wanninkhof, R., Wranne, A. W., et al.: SOCAT version 2022 for quantification of ocean CO<sub>2</sub> uptake, [https://www.socat.info/wp-content/uploads/2022/06/2022\\_Poster\\_SOCATv2022\\_release.pdf](https://www.socat.info/wp-content/uploads/2022/06/2022_Poster_SOCATv2022_release.pdf), 2022.
2. Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibáñez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., and Xu, S.: A multi-decade record of high-quality f CO<sub>2</sub> data in version 3 of the Surface Ocean CO<sub>2</sub> Atlas (SOCAT), *Earth System Science Data*, 8, 383–413, <https://doi.org/10.5194/essd-8-383-2016>, 2016.
3. Broecker, W. S., & Peng, T. H.: *Tracers in the Sea* (Vol. 690). Palisades, New York: Lamont-Doherty Geological Observatory, Columbia University, 1982.
4. Broullón, D., Pérez, F. F., Velo, A., Hoppema, M., Olsen, A., Takahashi, T., Key, R. M., Tanhua, T., González-Dávila, M., Jeansson, E., et al.: A global monthly climatology of total alkalinity: a neural network approach, *Earth System Science Data*, 11, 1109–1127, 2019.
5. Broullón, D., Pérez, F. F., Velo, A., Hoppema, M., Olsen, A., Takahashi, T., Key, R. M., Tanhua, T., Santana-Casiano, J. M., and Kozyr, A.: A global monthly climatology of oceanic total dissolved inorganic carbon: a neural network approach, *Earth Syst. Sci. Data*, 12, 1725–1743, <https://doi.org/10.5194/essd-12-1725-2020>, 2020.
6. Carter et al. 2016, Locally interpolated alkalinity regression for global alkalinity estimation, *Limnol. Oceanogr. Methods*, 14, 268–277, <https://doi.org/https://doi.org/10.1002/lom3.10087>.
7. Carter et al. 2018, Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate, *Limnology and Oceanography: Methods*, 16, 119–131, <https://doi.org/https://doi.org/10.1002/lom3.10232>, 2018.
8. Chau, T. T. T., Gehlen, M., and Chevallier, F.: A seamless ensemble-based reconstruction of surface ocean pCO<sub>2</sub> and air–sea CO<sub>2</sub> fluxes over the global coastal and open oceans, *Biogeosciences*, 19, <https://doi.org/10.5194/bg-19-1087-2022>.
9. Chavez, F. P., Sevajjian, J., Wahl, C., Friederich, J., and Friederich, G. E.: Measurements of pCO<sub>2</sub> and pH from an autonomous surface vehicle in a coastal upwelling system, *Deep-Sea Res. Pt. 2*, 151, 137–146, <https://doi.org/10.1016/j.dsr2.2017.01.001>, 2018.
10. Coppola, L., Boutin, J., Gattuso, J.-P., Lefèvre, D., and Metzl, N.: The Carbonate System in the Ligurian Sea, chap. 4, pp. 79–103, John Wiley & Sons, Ltd, <https://doi.org/https://doi.org/10.1002/9781119706960.ch4>, 2020.
11. Coppola, L., Diamond, R. E., and Carval, T.: Dyfamed observatory data. SEANO [dataset], <https://doi.org/0.17882/43749>, 2021.
12. 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<sub>2</sub> over the global ocean, *Geoscientific Model Development*, 12, 2091–2105, <https://doi.org/10.5194/gmd-12-2091-2019>, 2019.

13. Dickson et al (Eds.) 2007, Guide to Best Practices for Ocean CO<sub>2</sub> Measurements. PICES Special Publication 3, 191 pp.
14. Feely, R. A., Sabine, C. L., Hernandez-Ayon, J. M., Janson, D., and Hales, B.: Evidence for upwelling of corrosive “acidified” water onto the continental shelf, *Science*, 320, 1490–1492, <https://doi.org/10.1126/science.1155676>, 2008.
15. Friedlingstein, P., O’Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijckx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneeth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, O., Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., McGrath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O’Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pocock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2022, *Earth System Science Data*, 14, 4811–4900, <https://doi.org/10.5194/essd-14-4811-2022>, 2022.
16. Gattuso, J.-P., Alliouane, S., and Fischer, P.: High-frequency, year-round time series of the carbonate chemistry in a high-Arctic fjord (Svalbard), *Earth Syst. Sci. Data*, 15, 2809–2825, <https://doi.org/10.5194/essd-15-2809-2023>, 2023.
17. Gregor, L., Lebehot, A. D., Kok, S., and Scheel Monteiro, P. M.: A comparative assessment of the uncertainties of global surface ocean CO<sub>2</sub> 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.
18. 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.
19. 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, *Front. Mar. Sci.*, 7, 852, <https://doi.org/10.3389/fmars.2020.571720>, 2020.
20. Intergovernmental Oceanographic Commission (2019), Indicator methodology for SDG 14.3.1: Indicator description; Metadata template; Data template; Metadata instructions. Paris, France, Intergovernmental Oceanographic Commission of UNESCO. DOI: <http://dx.doi.org/10.25607/OBP-655>
21. Iida, Y., Takatani, Y., Kojima, A., and Ishii, M.: Global trends of ocean CO<sub>2</sub> sink and ocean acidification: an observation-based reconstruction of surface ocean inorganic carbon variables, *Journal of Oceanography*, 77, 323–358, 2021.
22. Jones, S., Le Quéré, C., and Rödenbeck, C.: Autocorrelation characteristics of surface ocean pCO<sub>2</sub> and air-sea CO<sub>2</sub> fluxes, *Global biogeo-chemical cycles*, 26, 2012.
23. Keppler, L., Landschützer, P., Gruber, N., Lauvset, S. K., & Stemmler, I. (2020). Seasonal carbon dynamics in the near-global ocean. *Global Biogeochemical Cycles*, 34, e2020GB006571. <https://doi.org/10.1029/2020GB006571>
24. Landschützer, P., Laruelle, G. G., Roobaert, A., and Regnier, P.: A uniform pCO<sub>2</sub> climatology combining open and coastal oceans, *Earth System Science Data*, 12, 2537–2553, <https://doi.org/10.5194/essd-12-2537-2020>, 2020.
25. Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., Alin, S., Álvarez, M., Azetsu-Scott, K., Barbero, L., Becker, S., Brown, P. J., Carter, B. R., da Cunha, L. C., Feely, R. A., Hoppema, M., Humphreys, M. P., Ishii, M., Jeansson, E., Jiang, L.-Q., Jones, S. D., Lo Monaco, C., Murata, A., Müller, J.

- D., Pérez, F. F. References, Pfeil, B., Schirnick, C., Steinfeldt, R., Suzuki, T., Tilbrook, B., Ulfsbo, A., Velo, A., Woosley, R. J., and Key, R. M.: GLODAPv2.2022: the latest version of the global interior ocean biogeochemical data product, *Earth System Science Data*, 14, 5543–5572, <https://doi.org/10.5194/essd-14-5543-2022>, 2022.
26. Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Perez, F. F., Suzuki, T., and Watelet, S.: A new global interior ocean mapped climatology: the  $1^\circ \times 1^\circ$  GLODAP version 2, *Earth Syst. Sci. Data*, 8, 325–340, <https://doi.org/10.5194/essd-8-325-2016>, 2016.
  27. Laruelle, G. G., Landschützer, P., Gruber, N., Tison, J.-L., Delille, B., and Regnier, P.: Global high-resolution monthly pCO<sub>2</sub> climatology for the coastal ocean derived from neural network interpolation, *Biogeosciences*, 14, 4545–4561, 2017.
  28. Lee, K., Tong, L. T., Millero, F. J., Sabine, C. L., Dickson, A. G., Goyet, C., Park, G.-H., Wanninkhof, R., Feely, R. A., and Key, R. M.: Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans, *Geophysical research letters*, 33, 2006.
  29. Metzl, N. and Lo Monaco, C.: OISO - Ocean Indien Service d'Observation [dataset], <https://doi.org/10.18142/228>, 1998.
  30. Millero, F. J., Lee, K., & Roche, M.: Distribution of alkalinity in the surface waters of the major oceans. *Marine Chemistry*, 60(1-2), 111-130, 1998.
  31. Olsen, A., Key, R. M., Van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., et al.: The Global Ocean Data Analysis Project version 2 (GLODAPv2)—an internally consistent data product for the world ocean, *Earth System Science Data*, 8, 297–323, 2016.
  32. Olafsson, J., Olafsdottir, S., Benoit-Cattin, A., and Takahashi, T.: The Irminger Sea and the Iceland Sea time series measurements of sea water carbon and nutrient chemistry 1983–2008, *Earth System Science Data*, 2, 99–104, 2010.
  33. Pfeil, B., Olsen, A., Bakker, D. C., Hankin, S., Koyuk, H., Kozyr, A., Malczyk, J., Manke, A., Metzl, N., Sabine, C. L., et al.: A uniform, quality-controlled Surface Ocean CO<sub>2</sub> Atlas (SOCAT), *Earth System Science Data*, 5, 125–143, 2013.
  34. Rödenbeck, C., Keeling, R. F., Bakker, D. C., Metzl, N., Olsen, A., Sabine, C., and Heimann, M.: Global surface-ocean pCO<sub>2</sub> and sea–air CO<sub>2</sub> flux variability from an observation-driven ocean mixed-layer scheme, *Ocean Science*, 9, 193–216, 2013.
  35. 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<sub>2</sub> Atlas (SOCAT) gridded data products, *Earth Syst. Sci. Data*, 5, 145–153, <https://doi.org/10.5194/essd-5-145-2013>, 2013.
  36. Schneider et al. 2007, Alkalinity of the Mediterranean Sea, *Geophys. Res. Lett.*, 34, L15608, doi:10.1029/2006GL028842.
  37. 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<sub>2</sub> and pH time series from 40 surface buoys and the emergence of anthropogenic trends, *Earth System Science Data*, 11, 421–439, 2019.