

# High resolution variability of the ocean carbon sink

Luke Gregor<sup>1</sup>, Jamie Shutler<sup>2</sup>, and Nicolas Gruber<sup>1</sup>

<sup>1</sup>Environmental Physics, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Zürich,  
Switzerland

<sup>2</sup>Centre for Geography and Environmental Science, University of Exeter, Penryn, Cornwall, UK

## Key Points:

- Surface ocean concentration and air-sea fluxes of carbon dioxide are estimated at 8-daily, quarter degree resolution with a neural network.
- Variability at subseasonal timescales contributes substantially to the total variability of the ocean carbon sink.
- The high-resolution data provide novel observational insights into regional and ephemeral processes, such as upwelling and hurricanes.

---

Corresponding author: Luke Gregor, [luke.gregor@usys.ethz.ch](mailto:luke.gregor@usys.ethz.ch)

**Abstract**

Measurements of the surface ocean fugacity of carbon dioxide ( $f\text{CO}_2$ ) provide an important constraint on the global ocean carbon sink, yet the gap filling products developed so far to cope with the sparse observations are relatively coarse ( $1^\circ \times 1^\circ$  by 1 month). Here, we overcome this limitation by using the newly developed surface Ocean Carbon dioxide Neural Network (OceanCarbNN) method to estimate surface ocean  $f\text{CO}_2$  and the associated air sea  $\text{CO}_2$  fluxes ( $F\text{CO}_2$ ) globally at a resolution of 8-daily by  $0.25^\circ \times 0.25^\circ$  (8D) over the period 1982 through 2022. Globally, the method reconstructs  $f\text{CO}_2$  with accuracy similar to that of low-resolution methods ( $\sim 19 \mu\text{atm}$ ), but improves it in the coastal ocean. Although global ocean  $\text{CO}_2$  uptake differs little, the 8D product captures 15% more variance in  $F\text{CO}_2$ . Most of this increase comes from the better-represented subseasonal scale variability, which is largely driven by the better resolved variability of the winds, but also contributed to by the better resolved  $f\text{CO}_2$ . The high-resolution  $f\text{CO}_2$  is also able to capture the signal of short-lived regional events such as coastal upwelling and hurricanes. For example, the 8D product reveals that  $f\text{CO}_2$  was at least  $25 \mu\text{atm}$  lower in the wake of Hurricane Maria (2017), the result of a complex interplay between the decrease in temperature, the entrainment of carbon-rich waters, and an increase in primary production. By providing new insights into the role of higher frequency variations of the ocean carbon sink and the underlying processes, the 8D product fills an important gap.

**Plain Language Summary**

The ocean is important for the climate, as it takes up about a quarter of the carbon dioxide ( $\text{CO}_2$ ) we release into the atmosphere. To determine this carbon sink, we measure the levels of carbon dioxide in the surface of the ocean. However, these measurements are limited to where ships measure  $\text{CO}_2$ , leaving gaps in our understanding. To fill in these gaps, statistical methods are used, but previous approaches lack fine-scale detail. We overcome this limitation with a new neural network approach that estimates  $\text{CO}_2$  with more detail, with estimates every 8 days, at 25 km, compared to previous 100 km monthly estimates. Globally, our method is as accurate as the previous methods, being slightly more accurate in coastal areas. Although the total amount of carbon the ocean absorbs globally remains consistent, our results show more variability. Our method also detects short-lived local events, such as coastal upwelling and hurricanes. For example,

45 our results show that after Hurricane Maria in 2017, the ocean surface had lower car-  
46 bon dioxide. Overall, our detailed results give us new information about the small-scale  
47 changes in the ocean carbon sink and helps us fill in a gap in our understanding of sur-  
48 face ocean CO<sub>2</sub>.

## 1 Introduction

The global ocean is playing a pivotal role in limiting global warming by having absorbed approximately 25% of the anthropogenic carbon dioxide ( $\text{CO}_2$ ) emissions over the past 250 years (Sabine et al., 2004; Gruber, Clement, et al., 2019; Müller et al., 2023; Khatiwala et al., 2013). Although this uptake fraction has remained remarkably stable over time (Friedlingstein et al., 2022), the magnitude of the ocean carbon sink has varied substantially around this trend (Landschützer et al., 2015, 2016; Gruber, Landschützer, & Lovenduski, 2019; DeVries et al., 2023; McKinley et al., 2020; Rödenbeck et al., 2022; Bennington et al., 2022). The strongest evidence supporting this variability comes from observations of the surface ocean  $\text{CO}_2$  concentration (generally expressed in terms of its fugacity,  $f\text{CO}_2$ ), from which the air-sea  $\text{CO}_2$  flux ( $F\text{CO}_2$ ) can be inferred (R. H. Wanninkhof, 2014; Fay et al., 2021). Since the  $f\text{CO}_2$  observations are sparse in time and space, gap filling techniques are required to map them to the time-and-space continuous product needed for estimating the strength of the global ocean carbon sink over time (Rödenbeck et al., 2015; Fay et al., 2021). Most commonly, statistical or machine learning techniques are used to fill this gap (Telszewski et al., 2009; Landschützer et al., 2013; Rödenbeck et al., 2015; Gregor et al., 2019; Iida et al., 2021; Chau et al., 2022; Gloege et al., 2022), although also geospatial and data assimilation-type methods are used (Rödenbeck et al., 2022; Bennington et al., 2022). Owing to their global nature and good temporal coverage, satellite observations have proven to be key enablers for all gap filling methods (Shutler et al., 2020, 2024).

Typically, these  $f\text{CO}_2$  gap-filled products are produced at a monthly  $1^\circ \times 1^\circ$  resolution (henceforth denoted by 1M). This resolution has proven to be sufficient to constrain global ocean uptake of  $\text{CO}_2$  and its trends and variations (Landschützer et al., 2016; Gruber et al., 2023; DeVries et al., 2023; Gloege et al., 2021) as well as the seasonal variations and their changes over time (Landschützer et al., 2018; Rodgers et al., 2023). As a result, these products have played an important role in global assessments, especially those of the Global Carbon Budget (GCB) (Friedlingstein et al., 2022, 2023; Hauck et al., 2023). But there is a large amount of variability that current  $f\text{CO}_2$ -products cannot capture, especially at the regional and subseasonal scale. Modeling studies and observations at these scales have regularly revealed  $f\text{CO}_2$  variations that exceed those seen in the 1M products (Yu et al., 2020; Nicholson et al., 2022; Turi et al., 2014; Arruda et al., 2015; Friederich et al., 2008; Resplandy et al., 2024). These variations are driven by

82 finer-scale temporal features such as storms and upwelling events, but also by finer-scale  
83 spatial features, such as those associated with mesoscale circulation or strong fronts. In  
84 addition, high-frequency variations in surface winds contribute to high frequency vari-  
85 ations in the air-sea CO<sub>2</sub> fluxes as well (Whitt et al., 2019).

86 Our current ability to constrain such high-resolution  $F\text{CO}_2$  and  $f\text{CO}_2$  variability  
87 from observations is very limited. This is, in no small part, a consequence of the segmented  
88 way the ocean CO<sub>2</sub> system is currently sampled (R. Wanninkhof et al., 2019). On the  
89 one hand, we have underway CO<sub>2</sub> measurements from ships that give a good perspec-  
90 tive of the variability in space along a limited number of survey lines (Jones et al., 2012).  
91 On the other hand, we have time-series observations from a few sites that provide de-  
92 tailed temporal information (Bates et al., 2014). But we rarely have observations that  
93 cover both time and space in a synoptic manner.

94 The few observation-based studies clearly point to the scale of the challenge. Re-  
95 garding high frequency temporal variability, a glider-based study in the Southern Ocean  
96 found that mid-latitude cyclones can induce variability up to 20  $\mu\text{atm}$  within a frequency  
97 range of 1 to 10 days (Nicholson et al., 2022). Another glider-based study suggested that  
98  $p\text{CO}_2$  measurements need to be taken every three days to constrain uncertainty in dy-  
99 namically variable regions (Monteiro et al., 2015). Beyond the Southern Ocean, intense  
100 sporadic events like tropical cyclones and hurricanes can cause  $p\text{CO}_2$  fluctuations as large  
101 as 50  $\mu\text{atm}$  within a two-day window (Yu et al., 2020; Bates et al., 1998; Koch et al., 2009).  
102 Mooring-based studies from various locations have also revealed variations of more than  
103 50  $\mu\text{atm}$  within days to weeks (Leinweber et al., 2009; Sutton et al., 2014, 2017; Pardo  
104 et al., 2019; Torres et al., 2021).

105 Regarding high resolution spatial variability, gradients exceeding several tens of  $\mu\text{atm}$   
106 over tens of kilometers are regularly encountered along ship-tracks, especially in dynamic  
107 regions such as the Southern Ocean and boundary current regions. This leads to short  
108 spatial autocorrelation length scales (Jones et al., 2012). While the length scales for  $F\text{CO}_2$   
109 and  $f\text{CO}_2$  are typically around 100 km or more in open ocean gyre regions, they decrease  
110 to less than 50 km in these dynamic regions (Murphy et al., 2001; Jones et al., 2012).  
111 These findings are corroborated by drifting buoy and saildrone-based studies in the North-  
112 eastern Atlantic and Southern Ocean, which have reported spatial gradients of  $p\text{CO}_2$  on

113 the order of 10  $\mu\text{atm}$  over 20 km (Boutin et al., 2008; Merlivat et al., 2009; Sutton et al.,  
114 2021).

115 Better resolving the fine-scale variations of the air-sea  $\text{CO}_2$  fluxes matters for mul-  
116 tiple reasons. First, it permits to better resolve a number of key processes that govern  
117 the ocean uptake of  $\text{CO}_2$ , providing novel insights into how the ocean carbon sink func-  
118 tions. Second, it permits us to assess the role of potential aliasing effects that stem from  
119 missing variability, thereby potentially aliasing our estimate of the global ocean carbon  
120 uptake (Koch et al., 2009). Third, such high-resolution products can also provide crit-  
121 ical constraints for assessing the impact of natural or man-made perturbations, such as  
122 those associated with marine heatwaves (Mignot et al., 2022) or the purposeful release  
123 of alkaline substances to enhance the oceanic uptake of atmospheric  $\text{CO}_2$  (González &  
124 Ilyina, 2016; Lenton et al., 2018).

125 First attempts to cover global finer-scale variability than 1M in gap-filled  $f\text{CO}_2$   
126 products were undertaken by Rödenbeck et al. (2014) for the CarboScope Mixed-Layer  
127 Scheme (CarboScope-MLS) and by Chau et al. (2024) for the CMEMS-FFNN product.  
128 The CarboScope-MLS, while having higher temporal resolution (daily), suffers from its  
129 coarse spatial resolution of  $> 2^\circ$ . CMEMS-FFNN offers superior spatial resolution, but  
130 resolves  $f\text{CO}_2$  only at monthly resolution. These limitations are, in part, technological  
131 in nature. For example, the commonly used Surface Ocean  $\text{CO}_2$  Atlas (SOCAT) provides  
132 a monthly  $1^\circ \times 1^\circ$  resolution gridded product alongside the ungridded cruise tracks (Sabine  
133 et al., 2013). Similarly, remote sensing and reanalysis products are often available at monthly  
134 resolutions in addition to daily files. There are, however, a few studies that have esti-  
135 mated high resolution  $f\text{CO}_2$  ( $\sim 5$  km) at regional scales, e.g., the Gulf of Mexico (Chen  
136 et al., 2019) and South China Sea (Song et al., 2023). The regional constraint of these  
137 studies allows for simpler machine learning architecture (e.g., no clustering) that are able  
138 to predict  $f\text{CO}_2$  with greater fidelity than the global approaches.

139 In this study, we aim to bridge the “high-resolution gap” in current  $f\text{CO}_2$  prod-  
140 ucts by generating estimates at an 8-daily,  $0.25^\circ \times 0.25^\circ$  (henceforth referred to as 8D)  
141 resolution for both  $f\text{CO}_2$  and  $F\text{CO}_2$ . To achieve this, we introduce a novel machine-learning  
142 approach called the Ocean Carbon Neural Network (OceanCarbNN). OceanCarbNN builds  
143 on several lower-resolution previous approaches, but reaches an unprecedented resolu-  
144 tion in time and space.

145 The paper is organized as follows: First, we outline the datasets and methodology  
146 that underlie our innovative technique. Following this, we rigorously evaluate the model's  
147 output. We then explore the implications of high-resolution  $p\text{CO}_2$  data on  $F\text{CO}_2$  vari-  
148 ability across different temporal scales. Finally, we assess the local and global impact of  
149 these high-resolution estimates, including a case study focused on the influence of a hur-  
150 ricane.

## 151 2 Methods

### 152 2.1 The Ocean Carbon Neural Network method

#### 153 2.1.1 Design elements of OceanCarbNN

154 The Ocean Carbon Neural Network (OceanCarbNN) method is a classical machine  
155 learning-based regression approach to map the sparsely observed  $f\text{CO}_2$  data to the global  
156 surface ocean (Rödenbeck et al., 2015). To achieve the mapping at the 8D target res-  
157 olution, while maintaining robustness and scalability, several design elements are imple-  
158 mented, most of which build on the ideas of other studies.

159 The first design element concerns the target variable. We first subtract atmospheric  
160  $\text{CO}_2$  in order to remove most of the trend in the target variable (Ma et al., 2023). There-  
161 after, following Bennington et al. (2022), we remove the temperature effect from oceanic  
162  $f\text{CO}_2$  leaving the non-thermal, i.e., chemically driven, part of the signal. These two trans-  
163 formations capture the impact of two well-understood and quantifiable drivers, such that  
164 the machine learning part is focused on the variability imparted on  $f\text{CO}_2$  through the  
165 other drivers, such as biology and mixing (Sarmiento & Gruber, 2006).

166 The second design element is that we use a prior estimate of an 8-daily climatol-  
167 ogy of  $f\text{CO}_2$  as a predictor. This is inspired by the works of Landschützer et al. (2015)  
168 and Denvil-Sommer et al. (2019), both of which include information on the  $f\text{CO}_2$  sea-  
169 sonal cycle, the dominant mode of variability. The first study does this through cluster-  
170 ing based on a predefined monthly climatology of  $f\text{CO}_2$ , and the second by removing  
171 the monthly climatological signal from  $f\text{CO}_2$  before training their neural network.

172 The third design element is that we combine decision-trees and neural-networks.  
173 For the estimation of the climatological seasonal cycle, we use Gradient Boosted Deci-  
174 sion Trees (GBDT), thus taking advantage of the low bias nature of tree-based approaches.

175 For the estimation of the time-varying  $f\text{CO}_2$ , we use a feed-forward neural network method.  
 176 Theoretically, the differentiable nature of FFNNs better captures the relationships (i.e.,  
 177 gradients) between the  $f\text{CO}_2$  and its drivers (Holder & Gnanadesikan, 2021).

178 The fourth and final design element is that we include the rate of change of the drivers  
 179 as predictors. More specifically, the difference between the current and previous time step  
 180 is used for variables such as temperature and chlorophyll-a. This gradient adds additional  
 181 information about the rates of change and further improves the stability of the  $f\text{CO}_2$   
 182 predictions between time steps.

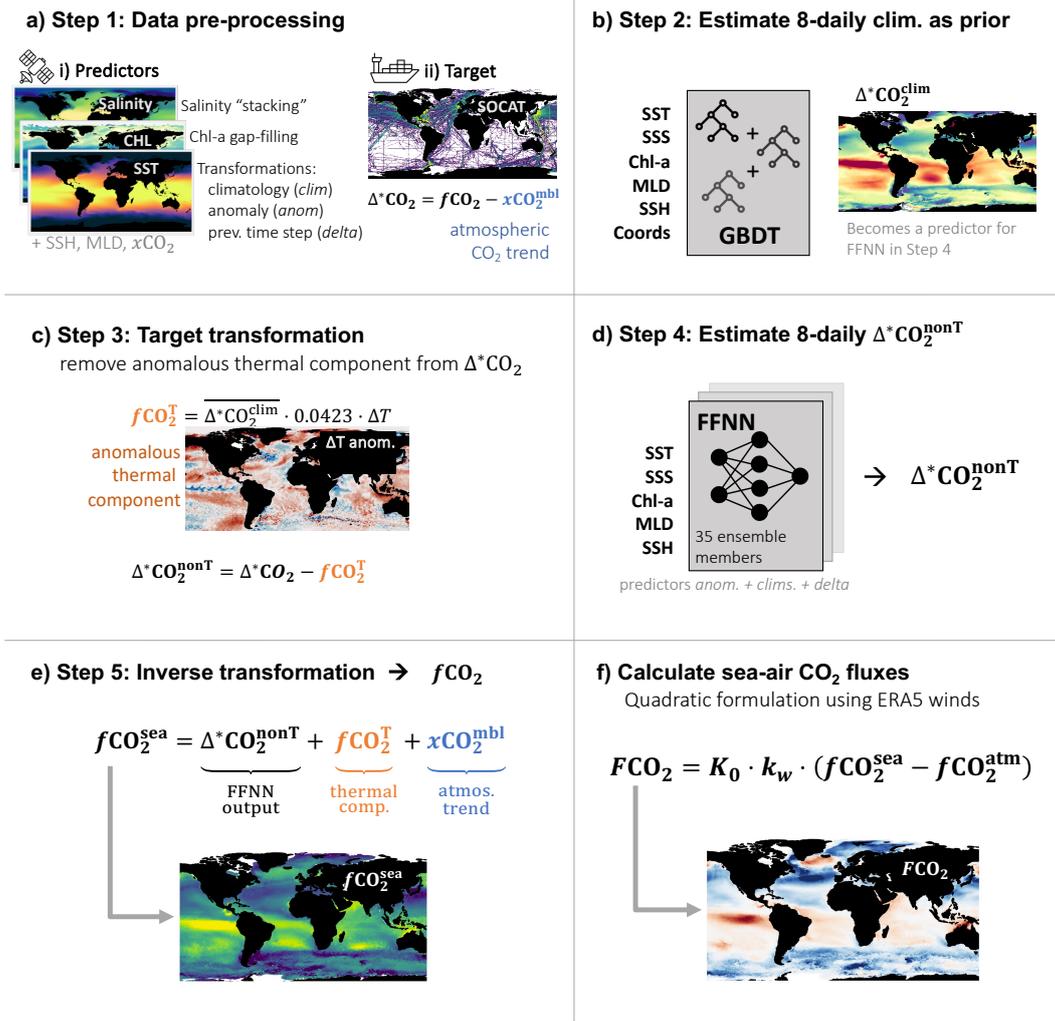
183 Overall, this set of design choices improves the stability and robustness of our es-  
 184 timates and improves the computational efficiency (compared to other methods), which  
 185 is important given that the amount of data increased by a factor of  $\sim 61$  by going from  
 186 the typical 1M resolution to 8D.

187 Figure 1 summarizes the 6 steps involved in the estimation of  $f\text{CO}_2$  and  $F\text{CO}_2$  by  
 188 OceanCarbNN: 1) Data preparation and preprocessing, including the detrending of the  
 189 target variable by subtracting the atmospheric  $\text{CO}_2$  mixing ratio to form the variable  
 190  $\Delta^*\text{CO}_2$ . 2) Machine learning part 1: Estimation of the 8-daily climatology of  $\Delta^*\text{CO}_2$ ;  
 191 3) Removal of the thermal component from  $\Delta^*\text{CO}_2$ ; 4) Machine learning part 2: Esti-  
 192 mation of the time-variable non-thermal target  $\Delta^*\text{CO}_2^{\text{non}T}$ ; 5) Reverse transformation  
 193 of  $\Delta^*\text{CO}_2^{\text{non}T}$  to obtain the time-variable  $f\text{CO}_2$  field; 6) Estimation of the air-sea  $\text{CO}_2$   
 194 fluxes,  $F\text{CO}_2$ . The methods section, in large, describes these steps in greater detail. Fur-  
 195 thermore, we describe the decomposition of the air-sea  $\text{CO}_2$  fluxes into different tempo-  
 196 ral modes of variability.

### 197 **2.1.2 Step 1: Data preparation and preprocessing**

The target variable is  $f\text{CO}_2$  from the ungridded SOCAT v2022 cruise track dataset  
 (D. C. Bakker et al., 2016; D. C. E. Bakker et al., 2023). The data are first binned to  
 8-daily by  $0.25^\circ \times 0.25^\circ$  resolution. No weighting is applied during the binning, mean-  
 ing that the data is analogous to the unweighted  $f\text{CO}_2$  data from the gridded SOCAT  
 product. We then detrend the data using the  $\text{CO}_2$  concentration in the marine bound-  
 ary layer ( $x\text{CO}_2^{\text{mbl}}$ ) from Dlugokencky et al. (2021):

$$\Delta^*\text{CO}_2 = f\text{CO}_2^{\text{SOCAT}} - x\text{CO}_2^{\text{mbl}} \quad (1)$$



**Figure 1.** Diagram showing the 6 steps in the OceanCarbNN approach. (a): Step 1: Data preparation and pre-processing (Figure 2, Section 2.1.2). Notably, we remove the atmospheric  $CO_2$  trend from SOCAT  $fCO_2$  by removing the mole fraction of  $CO_2$  at the marine boundary layer ( $xCO_2^{mb1}$ ). This yields the variable  $\Delta^*CO_2$ . (b) Step 2: An 8-daily climatology of  $\Delta^*CO_2$  is estimated and used as a prior in step-4 (Section 2.1.3). We use Gradient Boosted Decision Trees (GBDT). (c) Step 3: The target variable is transformed by removing the anomalous thermal component,  $fCO_2^T$  (after Bennington et al., 2022) (Section 2.1.5). The resulting non-thermal component,  $\Delta^*CO_2^{nonT}$ , is used as the target for the next step. (d) Step 4: An ensemble of 35 FFNNs predicts  $\Delta^*CO_2^{nonT}$  at 8D resolution (Section 2.1.5). (e) Step 5: Reverse transformation: The thermal component ( $fCO_2^T$ ) and atmospheric trend ( $xCO_2^{mb1}$ ) are added back to arrive at the OceanCarbNN estimate of  $fCO_2$ . (f) Step 6: Calculate the sea-air  $CO_2$  fluxes with the output from the previous step (Section 2.1.7).

198 For the detrending, we use  $x\text{CO}_2$  rather than atmospheric  $f\text{CO}_2$  in order to avoid  
 199 any unwanted impact of variations in atmospheric pressure and vapor pressure.

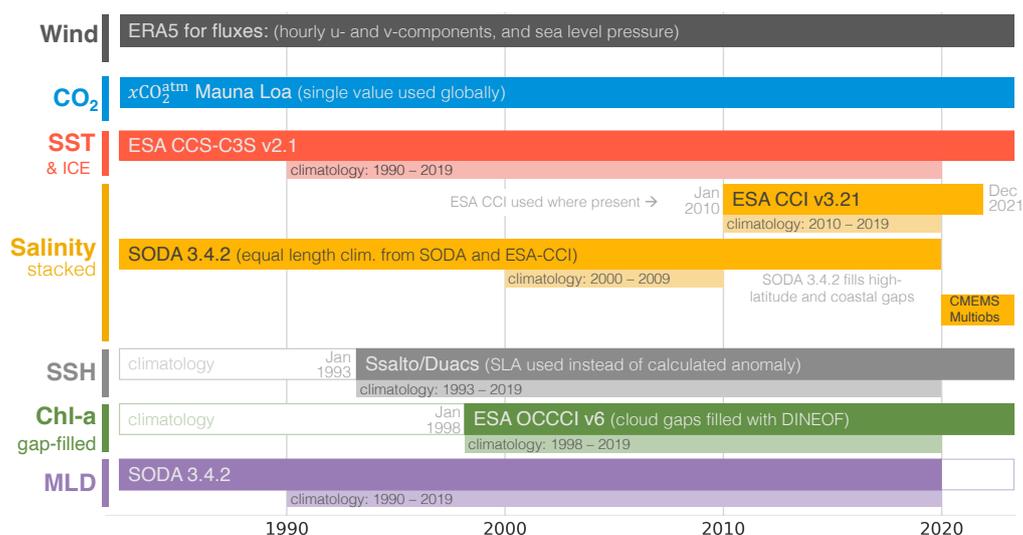
200 An important consideration for the selection of the predictor variable was the avail-  
 201 ability of these predictors at the 8D target resolution. This required some gap filling, and  
 202 some merging of different data sets. The final set is summarized in Figure 2 and con-  
 203 sists of the following predictors: atmospheric  $\text{CO}_2$  concentration at Mauna Loa ( $x\text{CO}_2^{\text{atm}}$ ;  
 204 Tans & Keeling, 2023), sea surface temperature (SST; Merchant et al., 2019), salinity  
 205 (Boutin et al., 2018; Carton et al., 2018; Droghei et al., 2016), sea surface height (SSH;  
 206 Taburet et al., 2019), chlorophyll-a (CHL; Sathyendranath et al., 2023), and mixed-layer  
 207 depth (MLD; Carton et al., 2018). This set of predictors is rather similar to that used  
 208 by most other gap-filling methods (Rödenbeck et al., 2015), except for the use of SSH,  
 209 the addition of the temporal derivatives of the drivers and prioritising the use of climate  
 210 data records.

211 Pre-processing and transformation (Figure 1a) of these variables can be separated  
 212 into: 1) gridding to 8-daily by  $0.25^\circ \times 0.25^\circ$ , 2) variable stacking for salinity and gap-  
 213 filling for CHL, 3) separation of anomalies from climatologies (*clim*, *anom*), 4) calculat-  
 214 ing the difference between the current and previous time step (*delta*) which is used as  
 215 a predictor. For the full description of how these data were prepared, see Supplemen-  
 216 tary Information Text S1.

### 217 **2.1.3 Step 2: Estimation of 8-daily climatology**

218 Landschützer et al. (2015) and Denvil-Sommer et al. (2019) have shown that the  
 219 climatological seasonal cycle of  $f\text{CO}_2$  is a powerful predictor when filling the gaps. This  
 220 is because to zeroth order, the seasonal cycle of  $f\text{CO}_2$  is relatively stationary, such that  
 221 using the climatological mean value for a given month is a good first guess of that month's  
 222 value. Both aforementioned studies used the climatology of Takahashi et al. (2009) and  
 223 interpolated it to the  $1^\circ$  resolution of their product. Here, we estimate this climatology  
 224 ourselves, especially since we require it at 8D resolution.

225 We estimate the quarter-degree 8-daily climatology of  $\Delta^*\text{CO}_2$  using Gradient Boosted  
 226 Decision Trees. We use climatological predictors along with transformations of the time  
 227 and space coordinates (Figure 2) — see Supplementary Information Text S1. The es-  
 228 timated  $\Delta^*\text{CO}_2^{\text{clim}}$  is smoothed with a rolling mean with a two-month by  $0.75^\circ$  window



**Figure 2.** A Gantt chart showing the use of various datasets in the OceanCarbNN. Each variable is represented by a color, with each product having its own horizontal bar and the thinner, lighter shaded bars represent the periods over which the climatology was calculated for that variable. In the case of salinity, where products overlap in time, the high-latitude and coastal gaps of the ESA-CCI product are filled by the CMEMS-Multiobs product, with ESA-CCI always taking priority. For a full list of these products, see Table S1 in the Supplementary Information.

229 (seven 8-day time steps) to avoid overfitting. We choose a large temporal window, but  
 230 narrow spatial window to preserve spatial gradients, particularly in the coastal regions.

#### 231 **2.1.4 Step 3: Target transformation**

Predictor variables for the second machine learning step include all variables as anomalies, together with their 8-daily climatologies, and the differences between the current and previous time step. The target,  $\Delta^* \text{CO}_2^{\text{nonT}}$ , is estimated by removing the thermal component of  $f\text{CO}_2$  (Figure 1c), after Bennington et al. (2022):

$$\Delta^* \text{CO}_2^{\text{nonT}} = \Delta^* \text{CO}_2 - \underbrace{\overline{f\text{CO}_2^{\text{clim}}} \cdot 0.0423 \cdot \Delta T}_{\text{thermal component}}, \quad (2)$$

232 where,  $\overline{f\text{CO}_2^{\text{clim}}}$  is the long-term mean estimate of the fugacity of  $\text{CO}_2$ , and  $\Delta T$  is the  
 233 temperature anomaly (as in Bennington et al., 2022), i.e., the anomaly relative to the  
 234 long-term mean sea-surface temperature.

#### 235 **2.1.5 Step 4: Estimation of $\Delta^* \text{CO}_2^{\text{nonT}}$**

236 We use a Feed-Forward Neural Network to estimate  $\Delta^* \text{CO}_2^{\text{nonT}}$ . This choice is based  
 237 on the work of Courtois et al. (2023), who demonstrate that such networks (that are not  
 238 too wide or deep) are able to extrapolate beyond the training observations. In addition,  
 239 neural networks were shown to be better able to capture the true relationship between  
 240 a target variable and its predictors (Holder & Gnanadesikan, 2021).

241 The collocated ship-track and predictor data are split into training, testing, and  
 242 validation subsets. We use the same approach as Bennington et al. (2022) where every  
 243 seventh month is considered a test (or validation) month. The validation subset uses the  
 244 same seven-month split, but with a three-month offset. The validation split is used to  
 245 avoid overfitting of the FFNN by stopping training when no improvement is observed.  
 246 Test data are not used during training of the FFNN and are used to assess performance.  
 247 Given the structure of our splits, seven train-test-validation splits can be created by start-  
 248 ing on a different month in 1982 (January through July), meaning that we can fully re-  
 249 construct the SOCAT cruise tracks with only test splits (Figure S1).

250 We use an ensemble of FFNNs to predict  $\Delta^* \text{CO}_2$ , with five FFNNs per train-test-  
 251 validation split. Together with the 7 splits, this results in a total of 35 ensemble mem-  
 252 bers. We use TensorFlow and Keras to construct our neural network ensemble (Abadi

253 et al., 2015). Each ensemble member has two hidden layers with 64 and 32 neurons with  
 254 ReLu activation (see Supplementary Information Text S2).

### 255 **2.1.6 Step 5: Retransformation**

256 The predicted output is first transformed back to  $\Delta^*CO_2$  from  $\Delta^*CO_2^{\text{nonT}}$  using  
 257 the thermal component of Eq. 2. The  $\Delta^*CO_2$  variable is in turn further transformed back  
 258 to  $fCO_2$  using Eq. 1. All 35 ensemble members are then averaged for the estimates of  
 259  $fCO_2$  and the standard deviation of the ensemble is calculated ( $\sigma_{\text{ens}}$ ).

### 260 **2.1.7 Step 6: Calculation of Fluxes**

We calculate sea-air  $CO_2$  fluxes ( $FCO_2$ ) using the bulk formulation:

$$FCO_2 = K_0 \cdot k_w \cdot (fCO_2 - fCO_2^{\text{atm}}) \cdot (1 - ice), \quad (3)$$

261 where  $K_0$  is the solubility of  $CO_2$  in seawater from Weiss (1974),  $k_w$  the gas transfer ve-  
 262 locity (R. H. Wanninkhof, 2014; Sarmiento & Gruber, 2006),  $fCO_2$  is the surface ocean  
 263  $CO_2$  fugacity predicted by OceanCarbNN,  $fCO_2^{\text{atm}}$  is the atmospheric marine bound-  
 264 ary layer  $CO_2$  fugacity, and  $ice$  is the sea-ice fraction from the temperature product shown  
 265 in Figure 2. The sign convention is that a positive flux in (3) is upward, i.e., indicating  
 266 outgassing of oceanic  $CO_2$ .

Atmospheric  $fCO_2$  is calculated from the dry air mixing ratio of atmospheric  $CO_2$   
 in the marine boundary layer from NOAA, i.e.,  $xCO_2^{\text{mb1}}$  (Dlugokencky et al., 2021) :

$$fCO_2^{\text{atm}} = xCO_2^{\text{mb1}} \times (P_{\text{atm}} - p_{H_2O}) \times \text{virial factor}, \quad (4)$$

where  $P_{\text{atm}}$  is the mean sea-level pressure from ERA5 (Hersbach et al., 2020),  $p_{H_2O}$  is  
 the partial pressure of water vapor based on Weiss and Price (1980), and the *virial factor*  
 accounts for the non-ideal behavior of  $CO_2$  (Weiss, 1974). For  $k_w$ , we use the quadratic  
 formulation of the sea-air  $CO_2$  fluxes from R. H. Wanninkhof (2014) scaled for ERA5  
 winds:

$$k_w = 0.271 \cdot U_{10}^2 \cdot \left( \frac{Sc}{660} \right)^{-1}, \quad (5)$$

267 where  $U_{10}^2$  is the second moment of the wind speed, and  $Sc$  is the Schmidt number for  
 268  $CO_2$  (Jähne et al., 1987). The second moment of the wind speed is calculated from hourly  
 269 ERA5 data using  $u^2 + v^2$ , where  $u$  and  $v$  are wind vectors.  $U_{10}^2$  is then regridded in the  
 270 time dimension to match the resolution of our output. The coefficient of gas transfer (0.271)

271 was obtained by ensuring that the global mean gas transfer coefficient for the period 1990  
 272 and 2019 and for the ice-free ocean matches the constraint of  $k_w = 16.5 \pm 3.2 \text{ cm hr}^{-1}$   
 273 (Sweeney et al., 2007; Naegler, 2009; Fay et al., 2021).

## 274 2.2 Decomposition of fluxes

275 We decompose the temporal variability of the air-sea  $\text{CO}_2$  fluxes into three modes  
 276 by integrating over the frequencies in the Fourier domain (after Gu et al., 2023): sub-  
 277 seasonal,  $< 3$  month frequencies; seasonal, 3 to  $\sim 15$  month frequencies; and interannual,  
 278  $> 15$  month frequencies. In some cases, we further separate the interannual variability  
 279 into sub-decadal (15 months to 8 years) and decadal ( $> 8$  years) variability. To simplify  
 280 the Fourier decomposition, we give all time steps the same length, i.e.,  $\frac{365}{46}$  days for the  
 281 high-resolution 8D product and  $\frac{365}{12}$  days for the 1M low-resolution product.

To identify the drivers of the air-sea  $\text{CO}_2$  flux variability, we apply a Reynolds de-  
 composition ( $y = \bar{y} + y'$ ) to Eq. 3. We thereby combine  $k_w$  and  $K_0$  as the gas trans-  
 fer coefficient  $k_x = K_0 \cdot k_w$ , in order to focus on the role of the wind variability (Doney  
 et al., 2009) – the temperature dependencies in  $k_w$  and  $K_0$  account for  $< 1\%$  of the vari-  
 ability of  $k_x$  (Woolf et al., 2016). We also look only at the role of the air-sea difference  
 in  $f\text{CO}_2$ , i.e.,  $\Delta f\text{CO}_2$ , since the variability of atmospheric  $f\text{CO}_2$  is much smaller than  
 that of the oceanic  $f\text{CO}_2$ . Furthermore, we neglect the role of sea-ice variations. With  
 these simplifications we decompose the air-sea  $\text{CO}_2$  flux,  $FCO_2$ , as follows:

$$FCO_2 = \underbrace{\overline{k_x \cdot \Delta f\text{CO}_2}}_{\text{mean state}} + \underbrace{k'_x \overline{\Delta f\text{CO}_2}}_{\text{wind variability}} + \underbrace{\overline{k_x \Delta f\text{CO}'_2}}_{f\text{CO}_2 \text{ variability}} + \underbrace{k'_x \cdot \Delta f\text{CO}'_2}_{\text{cross-term}}. \quad (6)$$

Since we are interested in the variability of  $FCO_2$  we do not have to consider the  
 first term where the long-term averages are taken for both variables. We calculate the  
 variability as representations of the variance ( $\sigma^2$ ). If we represent each of the terms in  
 Eq. 6 as  $a$ ,  $b$ , and  $c$  respectively, the total variance is represented by:

$$\begin{aligned} \sigma^2(a + b + c) &= \sigma^2(a) + \sigma^2(b) + \sigma^2(c) \\ &+ \underbrace{2 \cdot \text{cov}(a, b) + 2 \cdot \text{cov}(a, c) + 2 \cdot \text{cov}(b, c)}_{\text{covariances} \approx \text{residual}}. \end{aligned} \quad (7)$$

Given that we calculate the variance from the frequency domain, we calculate the  
 covariance for the different modes of variability as the residual of the total and summed

variances. The fraction contribution by each term is computed as:

$$\frac{\sigma^2(\text{wind})}{\sigma^2(\text{total})}, \frac{\sigma^2(f\text{CO}_2)}{\sigma^2(\text{total})}, \frac{\sigma^2(\text{cross term})}{\sigma^2(\text{total})}, \frac{\text{covariances}}{\sigma^2(\text{total})}. \quad (8)$$

282 Note that since the covariances can be negative, the fractional contribution can be  
283 negative, too.

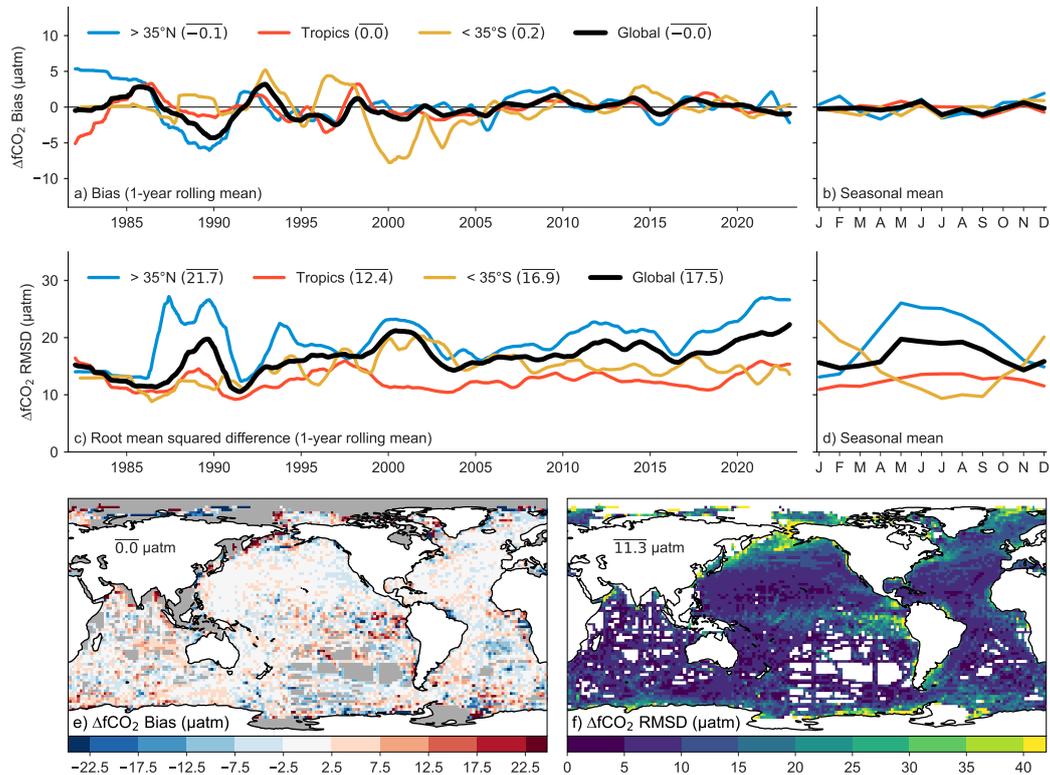
### 284 **3 Evaluation and assessment**

285 We first evaluate the 8D  $f\text{CO}_2$  product estimated by OceanCarbNN by determin-  
286 ing the large-scale offsets against the test and training data, and by comparing these off-  
287 sets against the ensemble spread that we use as an estimate of the prediction uncertainty.  
288 We then investigate in what way the high-resolution product is able to capture finer-scale  
289 structures in time and space. To this end, we focus on high-frequency observations from  
290 open ocean mooring stations (Sutton et al., 2019). We also assess how well the high-resolution  
291 estimates can track high-resolution spatial features observed along cruise tracks.

#### 292 **3.1 Uncertainties**

293 To determine the bias and root mean squared differences (RMSD) of the 8D prod-  
294 uct against the SOCAT data, we rely only on predictions that have not been used to train  
295 the subset of ensemble members. Given that we use an ensemble of results where the start-  
296 ing month changes (1-7), we have a complete representation of the SOCAT dataset (Figure  
297 S1; Gregor et al., 2019; Bennington et al., 2022).

298 The unweighted bias and RMSD for the  $f\text{CO}_2$  estimated by OceanCarbNN are low  
299 at a global scale at  $-0.06 \mu\text{atm}$  and  $19.2 \mu\text{atm}$ , respectively (Table S2), where a negative  
300 bias indicates that OceanCarbNN underestimates  $f\text{CO}_2$  relative to SOCAT. The open  
301 (coastal) ocean has a bias of  $0.07 \mu\text{atm}$  ( $-0.23 \mu\text{atm}$ ) and an RMSD of  $13.0 \mu\text{atm}$  ( $25.4$   
302  $\mu\text{atm}$ ) (Figure 3, Table S2), where the coastal ocean is defined as the ocean region within  
303  $300 \text{ km}$  from the coast or the  $1000 \text{ m}$  isobath (Laruelle et al., 2017). Given the spatial  
304 and seasonal inhomogeneity in the SOCAT observations, biases and RMSDs are also as-  
305 sessed over time and space (Figure 1 with weighted averages). We assess both metrics  
306 for three latitude bands: the high northern (blue lines;  $> 35^\circ\text{N}$ ) and southern latitudes  
307 (yellow lines;  $> 35^\circ\text{S}$ ) and the bounded lower latitudes (red).



**Figure 3.** Metrics using test data estimates from OceanCarbNN. Time series on the left show latitudinally averaged biases (a) and root mean squared difference (RMSD in c) with a one-year rolling mean applied. The legend also contains the time-averaged mean values (in  $\mu\text{atm}$ ) weighted by number of samples. The smaller time series figures on the right show seasonally averaged biases (b) and RMSD (d). The maps below show test bias (e) and RMSD (f) with the values showing the spatial average of the respective metrics weighted by area.

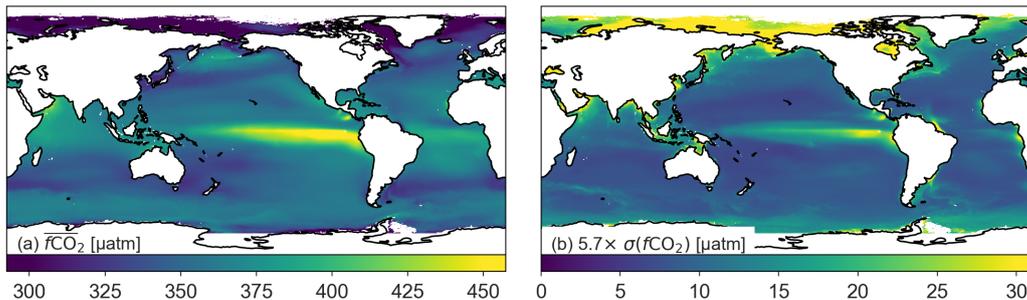
Biases are larger and more variable before 2000 when data are more sparse (Figure 3a). The spatial separation also reveals that biases compensate meridionally, resulting in lower global biases due to aggregation. There is no seasonality in the biases in any of the latitude bands. Spatially (Figure 3e), biases are mostly low ( $|bias| < 2.5 \mu\text{atm}$ ) except for the coastal regions and a few open ocean regions, e.g., the southern Indian Ocean, where there is a positive bias (red in Figure 3e).

In the northern high latitudes, the RMSD increases slightly over time ( $\sim 15$  to  $>20 \mu\text{atm}$ ; blue line in Figure 3c). This is likely due to the increase in the variance of the training data over time, particularly in the coastal ocean, where there is typically higher biogeochemical complexity (Figure S2). The RMSD for the tropics and southern regions remains constant ( $\sim 15 \mu\text{atm}$ ) throughout the forty-year period, with no significant slope. On average, RMSDs are larger in the summer months compared to the winter months ( $\sim 25 \mu\text{atm}$  vs.  $\sim 15 \mu\text{atm}$ ) of both the northern and southern high latitudes (blue and yellow lines in Figure 3d). The seasonality of the RMSD in the low-latitudes (red line in Figure 3d) is lower by comparison, but has a slight bias to the Northern Hemisphere summer, likely due to the Northern Hemisphere sampling bias.

At a global scale, the spatial distribution of the RMSD resembles the ensemble standard deviation ( $\sigma_{\text{ens}}$ ) with a spatial correlation  $r^2$  of 0.65 for the time averaged maps. This suggests that the ensemble spread is a good indicator of the spatial structure of the quality of the estimated  $f\text{CO}_2$ . This relationship was exploited by Chau et al. (2024) who scaled  $\sigma_{\text{ens}}$  to match the RMSD as closely as possible. Here, we follow the same idea to obtain a global map of uncertainty of our 8D product, but we simply multiply  $\sigma_{\text{ens}}$  with a factor of 5.7, i.e., the global mean ratio of  $\frac{\overline{\text{RMSD}}}{\sigma_{\text{ens}}} = \frac{11.3}{2.0}$ .

### 3.2 Assessing spatial and temporal scales of variability

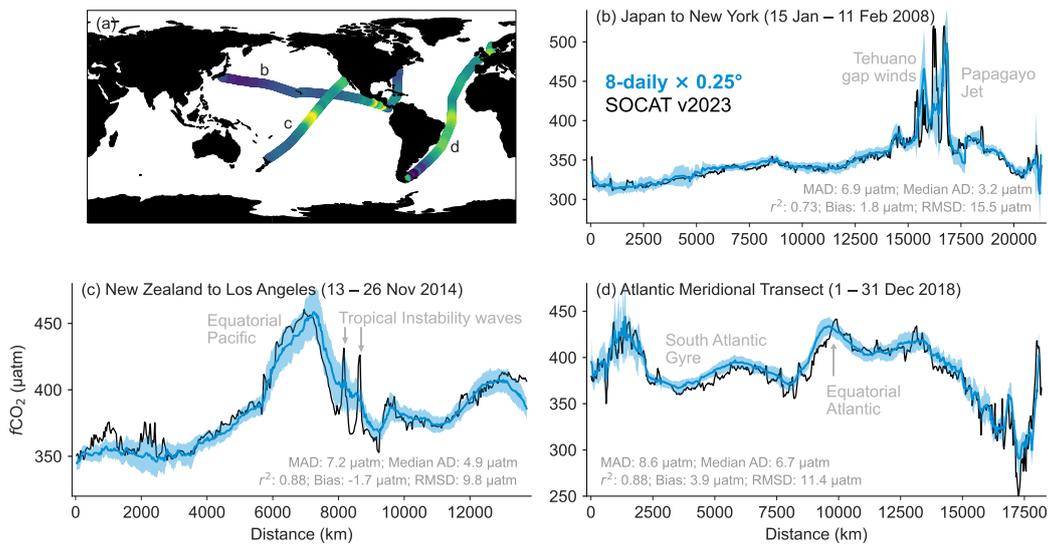
The benefit of estimating  $f\text{CO}_2$  at high resolution becomes most obvious when comparing the estimated product against the raw observations in SOCAT. The majority of the SOCAT data set consists of ship cruise tracks. In Figure 5, we compare three representative cruise tracks in the Atlantic and Pacific with the OceanCarbNN test-subset of  $f\text{CO}_2$ , where the test-subset is a subset of our output that has not been trained with the data that it is estimating. Along all three tracks, the large-scale variability of the SOCAT  $f\text{CO}_2$  observations is well represented. In open ocean regions, the RMSD scores



**Figure 4.** (a) The long-term mean of  $f\text{CO}_2$  for 1982 to 2022 and (b) the scaled standard deviation ( $5.7\sigma$ ) of the ensemble members of OceanCarbNN.

339 are actually substantially lower than the global average. An exception to this occurs in  
 340 poorly sampled regions, such as the South Atlantic gyre (Figure 5d, where the regional  
 341 biases can be  $> |5| \mu\text{atm}$ ). Most important, however, is the fact that the high-resolution  
 342 product captures a substantial fraction of the fine-scale structures of the observed  $f\text{CO}_2$ .  
 343 For example, the 8D product properly represents the strong outgassing signals associ-  
 344 ated with the upwelling driven by the gap winds off the coast of Central America (Fig-  
 345 ure 5a). The impact of these ephemeral winds, called Tehuano and Papagayo jets, on  
 346 the surface ocean  $\text{CO}_2$  system on ocean biogeochemistry has been well documented by  
 347 in situ measurements (Chapa-Balcorta et al., 2015), but is generally not well captured  
 348 by the global 1M products. Further, the 8D product properly captures the magnitude  
 349 of the equatorial upwelling in both the Atlantic and Pacific (Figure 5c,d), as well as the  
 350 strong variations off the coast of Europe (Figure 5d). At the same time, some deficien-  
 351 cies also emerge. The specific structure of the upwelling signal associated with the gap  
 352 winds is missed, as is the exact location of the equatorial upwellings offset. Furthermore,  
 353 the signals of two tropical instability waves in the equatorial Pacific (Figure 5c) are com-  
 354 pletely missed by the 8D product. Finally, the neural network cannot capture some ex-  
 355 tremely sharp gradients observed near New Zealand (Figure 5c).

356 Similar successes and limitations of the 8D product can be identified when com-  
 357 paring it with high-frequency observations from long-term moorings. Within SOCAT,  
 358 we identified five open-ocean mooring programs that have measured  $f\text{CO}_2$  at sub-daily  
 359 frequency for multiple years (Sutton et al., 2019; D. C. Bakker et al., 2016). These are  
 360 located in the Pacific and the Pacific sector of the Southern Ocean (Figure 6). The ob-  
 361 served  $f\text{CO}_2$  is well represented by the  $f\text{CO}_2$  estimated by OceanCarbNN at locations



**Figure 5.** Comparison of the observed  $f\text{CO}_2$  along three selected cruise tracks contained within the SOCAT database (black lines) with the estimated  $f\text{CO}_2$  from the 8D OceanCarbNN test-subset (blue lines). The blue envelope shows the scaled ensemble standard deviation ( $5.7\sigma_{\text{ens}}$ ). (a) Locations of the cruise tracks with the colors indicating the measured  $f\text{CO}_2$ . (b) Transect from Japan to New York occupied between Jan 15, 2008 and Feb 11, 2008. (c) Transect from New Zealand to Los Angeles occupied between Nov 13, 2014 and Nov 26, 2014. (d) Transect between Southampton and Punta Arenas occupied between Dec 1, 2018 and Dec 31, 2018 as part of the Atlantic Meridional Transect program. The distance along the x-axes of the cruise tracks (b, c, d) is plotted with zero being the most western point, regardless of the actual direction of travel.

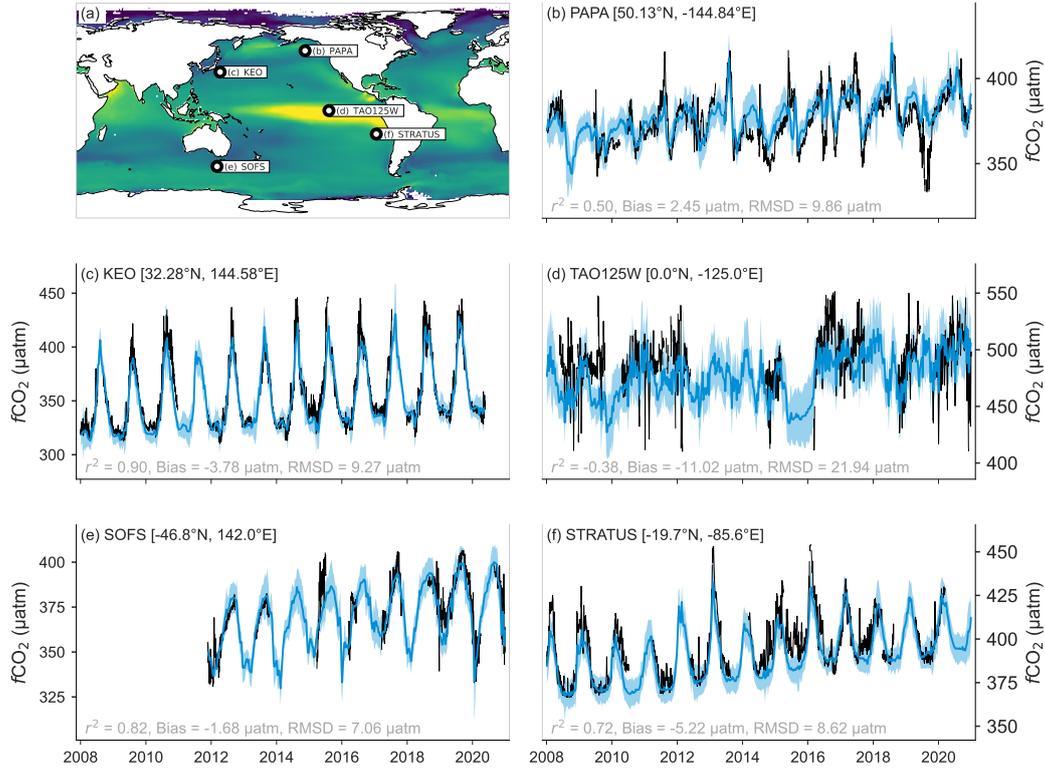
362 with a large seasonal cycle. The 8D product captures 90% of the variability observed at  
363 the KEO mooring off the coast of Japan (Figure 6a), 82% at the SOFS mooring south  
364 of Tasmania (Figure 6e) and 71% at the STRATUS mooring site off the coast of South  
365 America (Figure 6f). Also, shorter-term variations are generally well captured at these  
366 three sites. The performance of the OceanCarbNN 8D product is somewhat weaker at  
367 the other two sites. While the  $f\text{CO}_2$  estimates at the PAPA mooring location (Figure  
368 6b) are able to represent most of the seasonal variability, some extremes are not captured.  
369 This leads to only 50% of the observed variability being captured by the OceanCarbNN  
370 8D product. In the equatorial Pacific (TAO125W in Figure 6d) the OceanCarbNN 8D  
371 product appears to miss a substantial fraction of the high-frequency variability. Much  
372 of this variability is caused by equatorial instability waves. This mismatch is consistent  
373 with the mismatches seen in the Equatorial Pacific when analyzing the cruise line data  
374 (Figure 5c).

375 In summary, the 8D product is able to resolve simultaneously large and fine-scale  
376 structures in the observed  $f\text{CO}_2$ , with a few important exceptions. One example are very  
377 short-lived excursions, such as those associated with tropical instability waves. This is  
378 largely a consequence of these features having a propagation speed of around  $30 \text{ km day}^{-1}$   
379 (Legeckis, 1977), such that an 8-day resolution  $\sim 25 \text{ km}$  product is insufficient to correctly  
380 capture their dynamics.

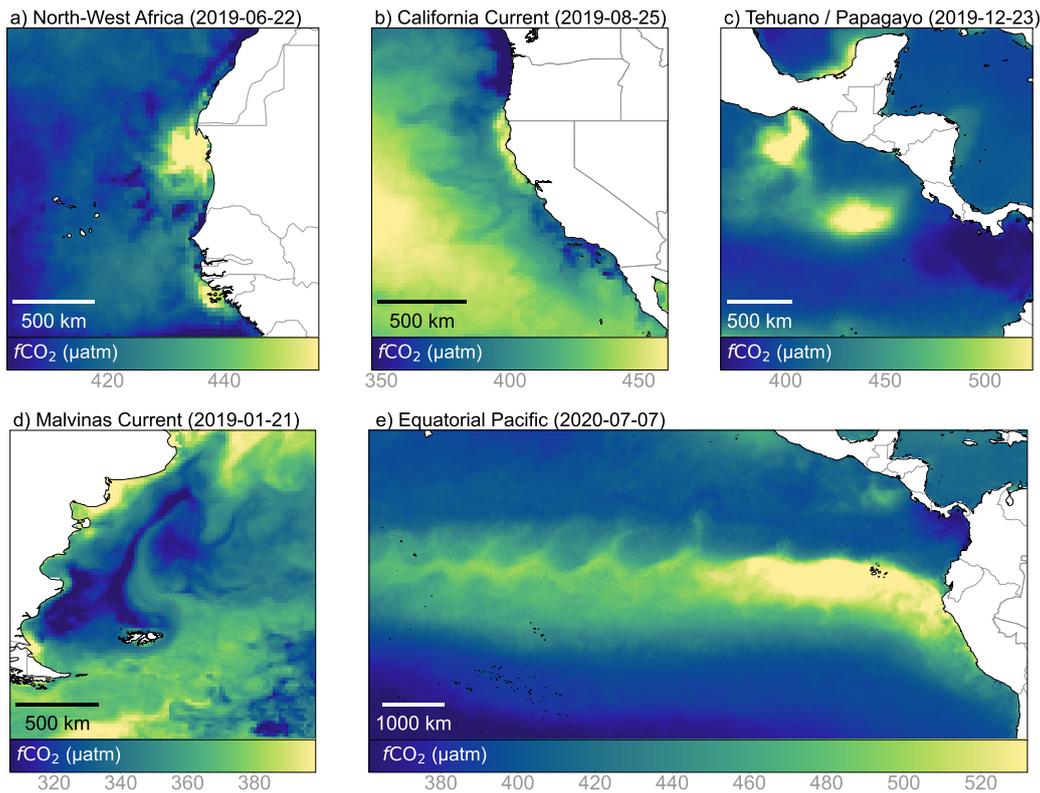
## 381 4 Patterns and variability of $f\text{CO}_2$

### 382 4.1 Representation of high-resolution features

383 Recognizing some shortcomings of our high-resolution mapped  $f\text{CO}_2$  product, it  
384 is nevertheless instructive to visualize its strengths in representing fine-scale features pre-  
385 viously not seen in gap filled  $f\text{CO}_2$  products. In Figure 7 we depict 5 snapshots from  
386 different regions of the global ocean (Supplementary Information Video S1). The 8D es-  
387 timates are able to represent a lot of fine-scale spatial variability in  $f\text{CO}_2$  that hitherto  
388 could not be seen in the 1M products. For example, the 8D estimates depict important  
389  $f\text{CO}_2$  details in the eastern boundary upwelling regions off northwestern Africa (Figure  
390 7a) and off the U.S. West Coast (Figure 7b). Of special note are the high  $f\text{CO}_2$  values  
391 tagging the coasts, reflecting recently upwelled waters, and the rapid offshore decrease  
392 of  $f\text{CO}_2$  owing primarily to strong biological drawdown. Also, the filamentous features



**Figure 6.** Mooring stations from Sutton et al. (2019) where  $f\text{CO}_2$  was measured. Blue lines show the OceanCarbNN test-subset of  $\Delta f\text{CO}_2$ , with the blue envelope showing the scaled ensemble standard deviation ( $5.7\sigma_{\text{ens}}$ ). Black lines show the measured mooring data resampled to a daily resolution. Model metrics are shown in gray in the bottom of each plot. Note that these estimates are not used to train the model. The map (c) shows the location of the moorings. Mooring locations roughly match the subplot locations.



**Figure 7.**  $f\text{CO}_2$  for different regions with the period represented shown by the date in brackets. (a) North-West African coastline, (b) The North-West coastline of the USA where the cal, (c) the West coast of Central America where the Tehuano gap winds, and Papagayo Jet winds occur, (d) the Malvinas (Falkland) current off the southeast coast of Argentina, (e) the equatorial Pacific from 165°W to 75°W.

393 of these low  $f\text{CO}_2$  waters are clearly visible. These structures correspond very well to  
 394 detailed regional observations and modeling studies (Turi et al., 2014; Lachkar & Gruber,  
 395 2013; Friederich et al., 2002).

396 Similarly, off the West coast of Central America, the 8D estimates reveal the spa-  
 397 tial extent of the upwelling driven maxima in  $f\text{CO}_2$  downstream of the Papagayo and  
 398 Tehuantepec gaps (Figure 7c). Despite the ephemeral nature of these gap winds, which  
 399 last for hours to several days (Romero-Centeno et al., 2003; Liang et al., 2009), their oceanic  
 400 signatures remain for long enough to be picked up well by our 8 daily product. The spa-  
 401 tial structure of these mountain gap wind features in  $f\text{CO}_2$  are consistent with what is  
 402 known from in-situ observations (see comparison to the cruise data above (Figure 5a)

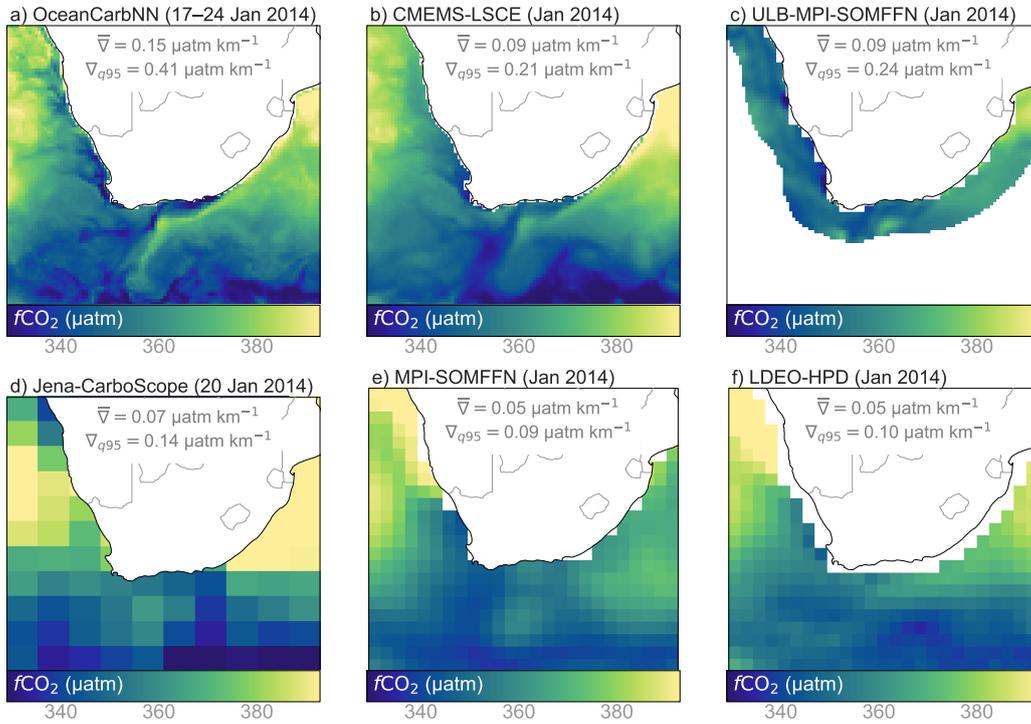
403 and the work by e.g., Chapa-Balcorta et al. (2015). Still, the spatial mismatches we iden-  
 404 tified in comparison with the cruise line data suggest that the 8-daily resolution is not  
 405 entirely sufficient to fully capture these ephemeral events.

406 The high spatial resolution is also able to resolve the very dynamic structure of the  
 407  $f\text{CO}_2$  in the Malvinas Current region (Figure 7d). In this region, strong biological pro-  
 408 ductivity over the Patagonian shelf interacts in a complex manner with the mixing of  
 409 very different waters masses, i.e., the mixing of the warm southward-flowing Brazil Cur-  
 410 rent (BC) and the cold northward flowing Malvinas Current (MC) (Arruda et al., 2015).  
 411 The meandering features have been seen in modeling studies (Arruda et al., 2015) and  
 412 the strong gradients created by the mixing are also regularly captured in the raw SO-  
 413 CAT cruise data (see also Figure 5d).

414 And finally, in the equatorial Pacific, the 8D  $f\text{CO}_2$  product begins to resolve the  
 415 tropical instability waves (Figure 7e). Their shapes are, however, a little distorted and  
 416 overly smoothed, which is as expected given their rapid propagation speeds (Legeckis,  
 417 1977). This mismatch has already been seen in the cruise line data (see also Figure 5c)  
 418 and also the mooring data from the TAO125W site (Figure 6d).

#### 419 **4.2 Comparison with other $f\text{CO}_2$ products**

420 The improvement of the OceanCarbNN 8D  $f\text{CO}_2$  estimates is also evident when  
 421 comparing it to several other  $f\text{CO}_2$ -products (Figure 8), namely, CMEMS-LSCE (1M  
 422 by  $0.25^\circ$  Chau et al., 2024), ULB-MPI-SOMFFN (1M by  $0.25^\circ$  Roobaert et al., 2023),  
 423 Jena-CarboScope by (daily by  $2.0^\circ$  Rödenbeck et al., 2014), MPI-SOMFFN by (1M by  
 424  $1^\circ$  Landschützer et al., 2016), and LDEO-HPD by (1M by  $1^\circ$  Bennington et al., 2022).  
 425 Apparent in the comparison is that the fine-scale gradients of the 8-daily OceanCarbNN  
 426 estimates are sharper compared to the other approaches. In this scenario, we find that  
 427 the mean gradients ( $\overline{\|\nabla f\text{CO}_2\|}$ ) of OceanCarbNN  $f\text{CO}_2$  are 60% stronger than the two  
 428 other high-resolution products (Figure 8b,c), and more than three times those of the monthly  
 429 by  $1^\circ$  resolution products (Figure 8e,f). The gradient increases to  $0.41 \mu\text{atm km}^{-1}$  when  
 430 considering the 95<sup>th</sup> percentile (representing dynamic regions) or  $\sim 8 \mu\text{atm}$  over 20 km,  
 431 thus approaching the sharp gradients recorded in observational studies (i.e.,  $\sim 10 \mu\text{atm}$   
 432 over 20 km Sutton et al., 2021).



**Figure 8.** A comparison of  $f\text{CO}_2$  from different  $f\text{CO}_2$ -products around Southern Africa, which includes the Agulhas Current and Benguela upwelling system. (a) OceanCarbNN  $f\text{CO}_2$  (this study) for the 8-day period 17 to 24 January 2014 (8-daily by  $0.25^\circ$ ). (b) CMEMS-LSCE (monthly by  $0.25^\circ$ ) by Chau et al. (2024), (c) ULB-MPI-SOMFFN (monthly by  $0.25^\circ$ , coastal only) by Roobaert et al. (2023), (d) Jena-CarboScope (daily by  $2^\circ$ ) by Rödenbeck et al. (2022), (e) MPI-SOMFFN (monthly by  $1^\circ$ ) by Landschützer et al. (2016), and (f) LDEO-HPD (monthly by  $1^\circ$ ) by Bennington et al. (2022). The metric  $\bar{\nabla}$  represents the average horizontal gradient of  $f\text{CO}_2$  for the plotted region and  $\nabla_{q95}$  represents its 95<sup>th</sup> percentile.

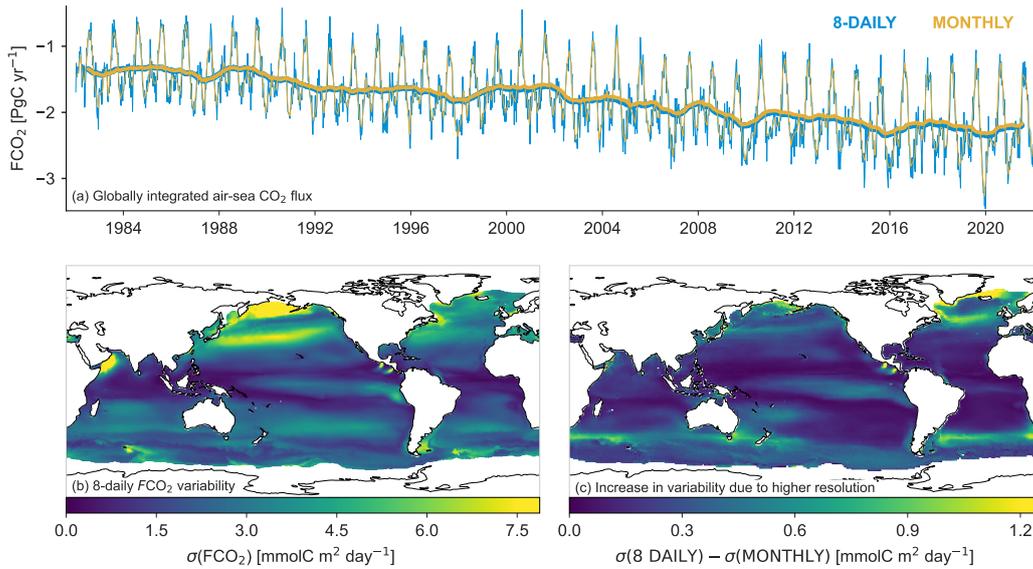
433 In summary, the OceanCarbNN method captures fine-scale spatial variability of  
434  $f\text{CO}_2$  with some skill. Though, there are still some high-frequency features (e.g., trop-  
435 ical instability waves) that are not well captured, even at the 8D resolution. However,  
436 our method is able to capture realistic variability that other methods tend to underes-  
437 timate.

## 438 5 Variability of the ocean carbon sink

### 439 5.1 Mean, trend, and variability of $F\text{CO}_2$

440 The globally integrated flux  $F\text{CO}_2$  inferred from the OceanCarbNN-8D product  
441 increases from  $1.51 \text{ PgC yr}^{-1}$  in 1990 to around  $-2.31 \text{ PgC yr}^{-1}$  in 2019, corresponding  
442 to a mean trend of  $-0.26 \text{ PgC yr}^{-1} \text{ decade}^{-1}$  over these 3 decades (Figure 9a). Over this  
443 period (1990–2019) this gives a mean global uptake of  $-1.89 \text{ PgC yr}^{-1}$ . The trend and  
444 mean uptake compares very favorably with previous estimates based on of gap filled  $f\text{CO}_2$   
445 products (DeVries et al., 2023). Concretely, it falls within the  $1\sigma$ -bounds of the SeaFlux  
446 ensemble of 6  $f\text{CO}_2$  products of  $-1.92 \pm 0.20 \text{ PgC yr}^{-1}$  calculated with ERA5 winds  
447 (Table 4 in Fay et al., 2021). Accounting for the outgassing of natural carbon associated  
448 with the balance between river input and burial, i.e., the so-called steady-state river out-  
449 gassing flux of about  $0.65 \text{ PgC yr}^{-1}$  (Regnier et al., 2022), the OceanCarbNN-8D prod-  
450 uct gives a total sink for anthropogenic  $\text{CO}_2$  of  $-2.31 \text{ PgC yr}^{-1}$  for 1990-1999, of  $-2.50$   
451  $\text{PgC yr}^{-1}$  for 2000-2009, and of  $-2.82 \text{ PgC yr}^{-1}$  for 2010-2019, consistent with the cur-  
452 rent best estimates of the magnitude and change of the ocean carbon sink (Gruber et  
453 al., 2023; DeVries et al., 2023).

454 The globally integrated flux varies substantially around these mean uptakes, with  
455 the seasonal cycle contributing the most variance (Figure 9a). There is also clear evi-  
456 dence of an increase in the magnitude of the seasonal cycle over time, confirming pre-  
457 vious findings based on theory, models, and observations (Landschützer et al., 2018; Rodgers  
458 et al., 2023). In addition, clear subseasonal variations are visible in the global timeseries,  
459 as well as interannual to decadal variations. Regarding the magnitude of the interannual  
460 to decadal variability, the OceanCarbNN-8D-based flux estimates tend to be on the lower  
461 end of the spectrum compared to other gap filled  $f\text{CO}_2$  products (DeVries et al., 2023).  
462 Still, it shows a clear stalling of the trend toward stronger uptake in the 1990s, and an



**Figure 9.** (a) Globally integrated sea-air CO<sub>2</sub> fluxes ( $FCO_2$ ) for 8-daily (8D, blue) and monthly (1M, yellow) estimates, where the latter is calculated by resampling the inputs of Eq. 3 to a lower monthly  $\times 1.0^\circ$  resolution and then calculating  $FCO_2$ . (b) The standard deviation ( $\sigma$ ) of  $FCO_2$  for the 8D estimates. (c) The difference between (b) and the 1M (monthly  $\times 1.0^\circ$  resolution).

463 acceleration thereafter (Landschützer et al., 2015; DeVries et al., 2019; Gruber, Clement,  
464 et al., 2019; Gruber et al., 2023).

465 To assess the impact of the higher resolution on our estimates of the variability of  
466 the ocean carbon sink, we contrast our 8D estimate with a monthly by  $1^\circ \times 1^\circ$  (1M) es-  
467 timate we obtain by averaging  $fCO_2$  from OceanCarbNN and all other inputs to Eq. 3  
468 to 1M (Figure 9a). Even though the 1M results are based on the averages of the 8D es-  
469 timates, the global fluxes of the two products are not identical on longer timescales, i.e.,  
470 the 1M product has a marginally smaller mean (1990-2019) uptake of  $-1.87 \text{ PgC yr}^{-1}$ .  
471 This small difference is likely due to small co-variances between wind and  $\Delta fCO_2$ , whose  
472 magnitude is scale-dependent (see below). As expected, much larger differences occur  
473 on subseasonal timescales, where the 8D product reveals higher crests and deeper troughs  
474 than the 1M product. This results in higher temporal variance, which is best analyzed  
475 spatially.

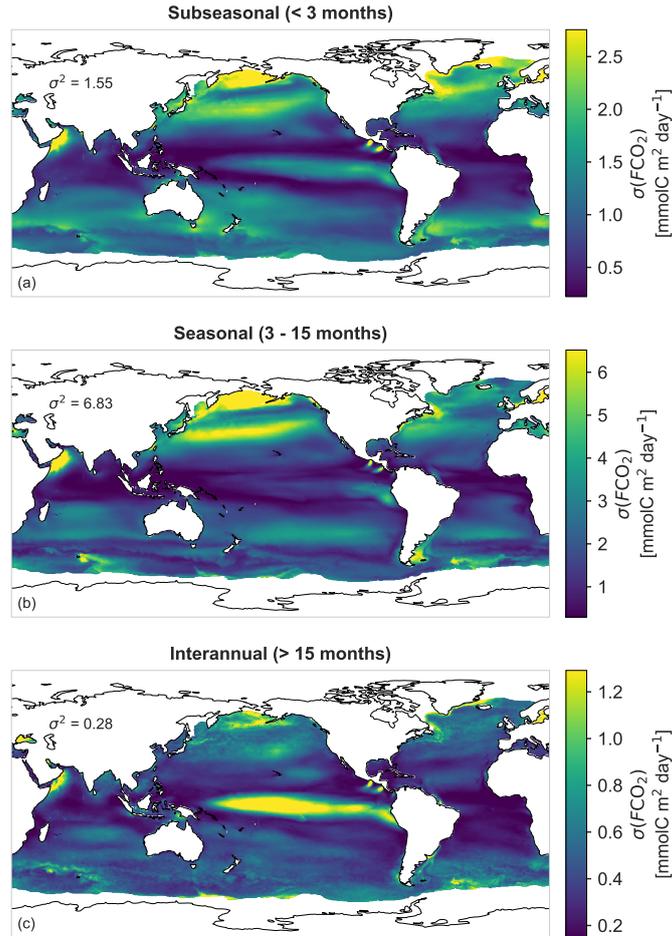
476 The map of the total standard deviation of the air-sea CO<sub>2</sub> flux  $FCO_2$  reveals strong  
 477 spatial differences ranging from near zero to more than 7 mmol C m<sup>-2</sup> day<sup>-1</sup> (Figure  
 478 9b). The northern mid- to high-latitudes have the largest variability (> 5 mmolC m<sup>-2</sup>  
 479 day<sup>-1</sup>), particularly in the Pacific basin. Some island and coastal regions (e.g., Kerguelen  
 480 Plateau and Oman upwelling regions) have similarly large variability. The low-latitude  
 481 regions have low variability (< 1 mmolC m<sup>-2</sup> day<sup>-1</sup>), except for the central and east-  
 482 ern tropical Pacific. In the global mean, the temporal standard deviation of the 8D prod-  
 483 uct amounts to  $\bar{\sigma} = 2.45$  mmolC m<sup>-2</sup> day<sup>-1</sup>. This is about 10% more than the global  
 484 mean temporal standard deviation  $\bar{\sigma}$  of the 1M product (2.20 mmolC m<sup>-2</sup> day<sup>-1</sup>).

485 While the global increase in the captured variability of the 8D product is modest,  
 486 regionally, the gain in variability of the air-sea CO<sub>2</sub> fluxes can be much more pronounced  
 487 (Figure 9c). The increase in variability is most notable along a band (~40°S) in the Sub-  
 488 Antarctic Zone (SAZ) in the Southern Ocean, where the 8D product increases the vari-  
 489 ability relative to the 1M product by more than 50% in some parts. This corresponds  
 490 well with the results from (Monteiro et al., 2015). The northern high-latitude Atlantic  
 491 Ocean exhibits a particularly large increase in variability ( $\gtrsim 1.5$  mmolC m<sup>-2</sup> day<sup>-1</sup>).  
 492 However, the relative increase is smaller when compared with the SAZ (< 40%).

## 493 5.2 Temporal decomposition of variability of $FCO_2$

494 As seen already in the global timeseries, the total variability of  $FCO_2$  is dominated  
 495 by the seasonal mode (10b), i.e., by variability at timescales between 3 and 15 months.  
 496 The standard deviation of the flux on this timescale goes up to more than 6 mmolC m<sup>-2</sup>  
 497 day<sup>-1</sup>, with highest values found in the North Pacific. In contrast, the tropical regions  
 498 exhibit nearly no variability on seasonal timescales. Globally, the standard deviation on  
 499 seasonal timescales amounts to  $\sigma = 2.61$  mmolC m<sup>-2</sup> day<sup>-1</sup>. The difference between  
 500  $\sigma(1M)$  and  $\sigma(8D)$  is small (0.03 mmolC m<sup>-2</sup> day<sup>-1</sup>), indicating that, at the seasonal scale,  
 501 monthly data captures most of the variability.

502 The subseasonal mode (< 3 months) is the next largest contributor to  $FCO_2$  vari-  
 503 ability ( $\sigma = 1.24$  mmolC m<sup>-2</sup> day<sup>-1</sup>, Figure 10a). The contribution to variability on these  
 504 shorter timescales is dominated by the mid-latitudes and the equatorial Pacific, with some  
 505 regions exceeding 2.5 mmolC m<sup>-2</sup> day<sup>-1</sup>. To first order, the spatial distribution of the  
 506 subseasonal variability is similar to that of the seasonal variability (Pearson  $R = \sim 0.69$ ).



**Figure 10.** Maps of the standard deviation of the seasonal (a), sub-seasonal (b), and interannual (c) components of 8-daily by  $0.25^\circ FCO_2$ . Values on the plots show the area weighted mean of the variance in  $(\text{mmolC m}^{-2} \text{ day}^{-1})^2$ . See Tables S3, S4, and S5 for low and high-resolution variance, standard deviation, and percentages, respectively.

507 An important difference is the strong sub-seasonal contribution in the Oman upwelling  
 508 system and in the Coral Sea and Tasman Sea regions east of Australia. As expected, the  
 509 gain in information by going from 1M to 8D is the largest in this sub-seasonal mode. Glob-  
 510 ally, the standard deviation increases from the 1M product nearly by a factor of 3 (see  
 511 Tables S3 and S4).

512 Finally, variability on time-scales longer than 15 months (interannual mode) con-  
 513 tributes the least to the overall standard deviation (Figure 10c). The global mean tem-  
 514 poral standard deviation amounts to  $0.52 \text{ mmolC m}^{-2} \text{ day}^{-1}$ , for both the 8D and 1M  
 515 products. Here, the interannual variability of the equatorial Pacific driven by El-Niño—Southern

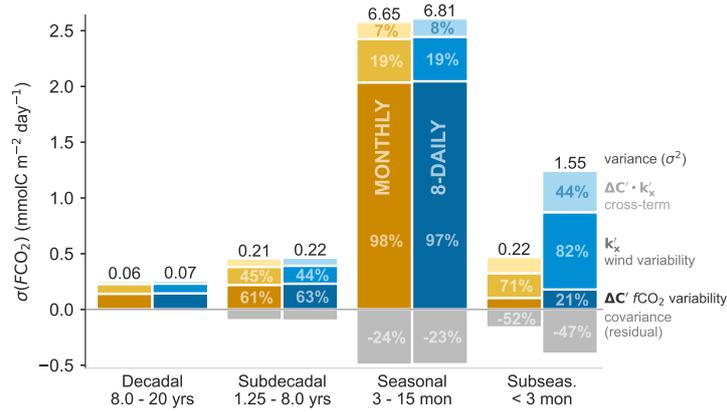
516 Oscillation (ENSO) is the most dominant feature, but also the higher latitudes contribute  
517 substantially, with several regions having standard deviations of more than 0.8 mmolC  
518  $\text{m}^{-2} \text{day}^{-1}$ .

519 This attribution of the total variability of the air-sea  $\text{CO}_2$  fluxes to three different  
520 modes of variability is qualitatively similar to that undertaken by (Gu et al., 2023). Per-  
521 haps unsurprisingly, there is good agreement between our studies with regard to the par-  
522 titioning into subseasonal, seasonal, and interannual modes of variability. They found  
523 that the total energy of seasonal variance was an order of magnitude larger than for the  
524 subseasonal and interannual modes. This aligns with our findings, where we find an even  
525 stronger relative contribution of the subseasonal mode to the total variance, particularly  
526 at the 1M resolution.

### 527 **5.3 Drivers of the variability of $f\text{CO}_2$**

528 The Reynolds decomposition of the air-sea  $\text{CO}_2$  fluxes (Eq. 6) permits us to iden-  
529 tify the main drivers for each mode of variability (Figure 11). The seasonal variability  
530 is dominated by changes in  $\Delta f\text{CO}_2$ , contributing 97% to the total seasonal changes. This  
531 is also the case for the longer modes of variability, with  $\Delta f\text{CO}_2$  contributing  $\sim 60\%$  to  
532 either the sub-decadal and decadal of modes variability. Wind variability, expressed in  
533 the variations in the gas transfer coefficient, matters as well, especially for the sub-decadal  
534 modes, where it contributes more than 40%. The cross-term contributions tend to be  
535 relatively unimportant for the seasonal and longer modes, contributing less than 10%  
536 to the overall variability. The co-variances between the different Reynolds terms (see Eq.  
537 7) are negative, thus contributing negatively to the overall variability. This is likely a  
538 result of negative correlations between the mean state and the variability, e.g., high wind  
539 speed regions/times tend to be regions/times with low variations in the air-sea differ-  
540 ence in  $f\text{CO}_2$ . This offsetting effect is particularly strong on the seasonal timescale, where  
541 it offsets the variability by more than 20%. Given the small difference between the 8D  
542 and 1M products on the seasonal and longer timescales, the results of the Reynolds de-  
543 composition are also nearly the same for these timescales.

544 This is not the case for the sub-seasonal mode of variability (Figure 11). And more  
545 interestingly, the large increase in the subseasonal mode of variability between 1M and  
546 8D is due to all components of the Reynold's decomposition (Table S3). Still, the largest



**Figure 11.** The contribution of  $\Delta fCO_2'$  (dark shading), wind and temperature ( $k'_x$ , medium shading), and the cross-term ( $\Delta fCO_2' \cdot k'_x$ , light shading) to each temporal mode of variability for 8-daily (blue) and monthly (yellow)  $FCO_2$ . The covariances are shown in light gray. The height of the bars (positive only) shows the total standard deviation ( $\sigma$ ) and the black numbers above each bar show the total variance ( $\sigma^2$ ). The total height of the bar (positive and negative) represents the sum of the variances without the covariance term. The percentage contributions of  $k'_x$ ,  $\Delta fCO_2'$ , cross-term, and covariances are with respect to the total variance. Thus, multiplying  $\sigma^2$  with the percentage is approximately the variance for that component (within the uncertainty of rounding). See Tables S3, S4, and S5 for low and high-resolution variance, standard deviation, and percentages, respectively.

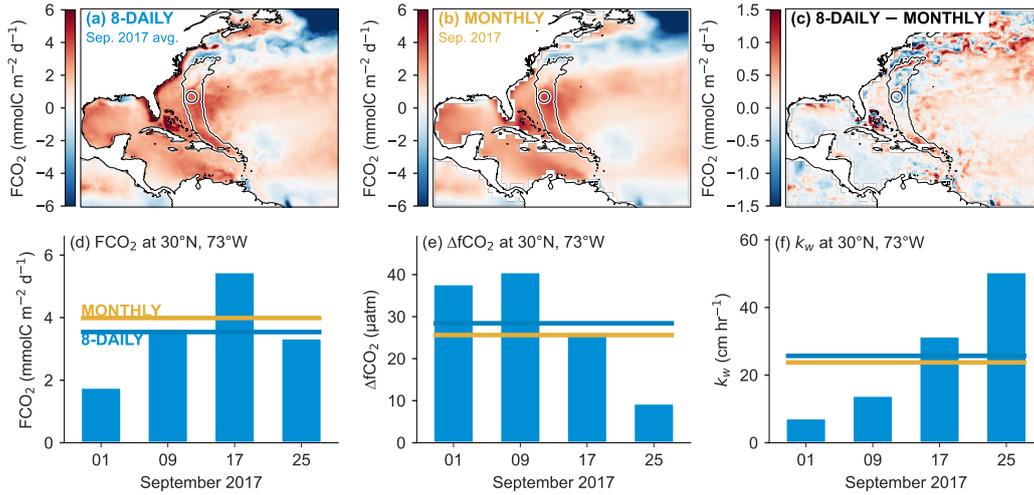
547 gain comes from the enhanced resolution of the wind variability as expressed in the con-  
 548 tribution of  $k'_x$ , which increases nearly threefold from 0.40 mmolC m<sup>-2</sup> day<sup>-1</sup> to 1.12  
 549 mmolC m<sup>-2</sup> day<sup>-1</sup> (1M to 8D). The cross-term variability at the subseasonal mode also  
 550 increases significantly from 0.32 mmolC m<sup>-2</sup> day<sup>-1</sup> to 0.82 mmolC m<sup>-2</sup> day<sup>-1</sup>. This in-  
 551 dicates that there is an increase in the interaction between  $k_x$  and  $f\text{CO}_2$  in the high-resolution  
 552 fluxes. Mechanistically this makes sense, since wind-driven upwelling (captured by  $k_x$ )  
 553 can lead to rapid changes in  $\Delta f\text{CO}_2$ , thus resulting in larger  $F\text{CO}_2$ . This is also expressed  
 554 in the larger covariance term that increases by a similar magnitude (negatively) to the  
 555 cross-term 11. The increase in  $\sigma(F\text{CO}_2)$  due to high-resolution  $\Delta f\text{CO}_2$  is 0.48 mmolC  
 556 m<sup>-2</sup> day<sup>-1</sup> — smaller than the other subseasonal increases from 1M to 8D, but still larger  
 557 than for any other mode (Table S3). Particularly striking too is the very strong nega-  
 558 tive contribution of the co-variance term, which reduces the overall variability by around  
 559 50%.

560 The results of the Reynold decomposition for the seasonal and longer timescales  
 561 are similar to those of Gu et al. (2023). The main consistency is that, on seasonal time-  
 562 scales, variations in  $k_x$ , i.e., wind dominate, while at longer timescales, the  $\Delta f\text{CO}_2$  term  
 563 dominates. More specifically, Gu et al. (2023) attributed 66% of subseasonal variabil-  
 564 ity (< 3 months) to the wind component, while we attribute 71% at the 1M-scale (com-  
 565 parable to their study).

566 In summary, improved representation of sub-seasonal variability requires high-frequency  
 567 information on both wind and  $\Delta f\text{CO}_2$ , and perhaps even more importantly, excellent  
 568 information about their co-variances.

## 569 **6 Case study: Hurricane Maria**

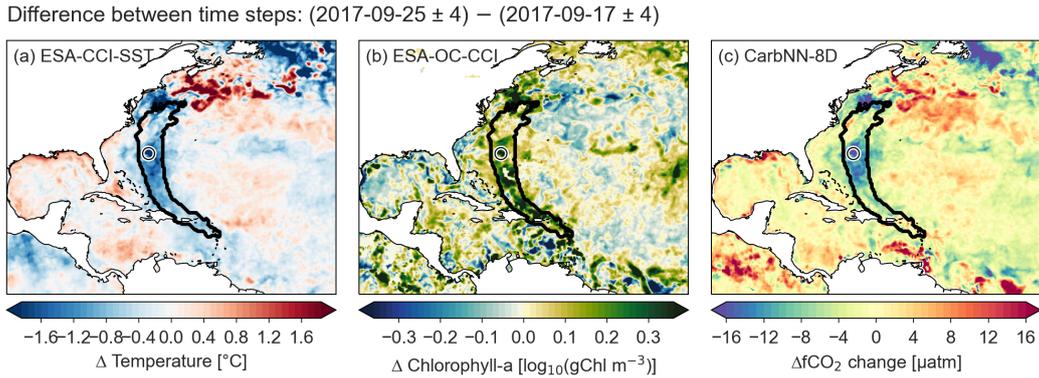
570 The importance of co-analyzing wind variations with changes in surface ocean bio-  
 571 geochemistry and hence surface ocean  $f\text{CO}_2$  is best shown through an example, for which  
 572 we use Hurricane Maria as a case study. Hurricane Maria occurred in September 2017  
 573 and belongs to the 10 most intense Atlantic hurricanes ever recorded. It made history  
 574 as it made first landfall in Puerto Rico before it turned northward and plowed through  
 575 the northwestern North Atlantic. Hurricane Maria was most intense from the 17<sup>th</sup> to 27<sup>th</sup>  
 576 of September, maintaining hurricane status throughout this period (shown by the black  
 577 contours in Figure 12a-c). Thereafter, the system moved eastward. Here, we investigate



**Figure 12.** The mean of the 8-daily (a) and monthly (b)  $FCO_2$  for September 2017. The contour line shows the path of Hurricane Maria from 17<sup>th</sup> to 27<sup>th</sup> September, where hourly winds exceeded  $20 \text{ m s}^{-1}$ . (c) the difference between the 8-daily mean and monthly values. (d) a bar plot of  $FCO_2$  for the four time steps in September 2017 for the location indicated by the circle in the maps (a-c). (e) and (f) represent the same, but for  $\Delta fCO_2$  and  $k_w$  respectively. The  $FCO_2$  shown in (d) can be thought of as the product of the corresponding bars in (e) and (f) which are scaled by the solubility (not shown). Slight differences between averages (horizontal lines in e,f) are due to the marker not being at exactly the same location for high and low-resolution estimates.

578 the local-scale impacts of the increase in variance by assessing the  $FCO_2$  during the pas-  
 579 sage of Maria, which represents a short-lasting extreme event that is missed in monthly  
 580 reconstructions.

581 At first glance, the spatial distribution of the September 2017 mean of the 8D and  
 582 1M  $FCO_2$  look similar (Figure 12a,b). However, the difference between the two mean  
 583 states (Figure 12c) shows that outgassing was in fact less intense along the northern part  
 584 of the hurricane path for the 8D  $FCO_2$  relative to 1M. To better understand the lower-  
 585 than-expected outgassing, we plot the temporal evolution of the fluxes,  $\Delta fCO_2$  and  $k_w$   
 586 at a point location ( $29^\circ\text{N}$ ,  $72^\circ\text{W}$ ) for each of the 8-daily time steps in September 2017  
 587 (Figure 12d-f). The location exhibits one of the strongest differences between the 8D and  
 588 1M  $FCO_2$  for the spatio-temporal domain.



**Figure 13.** Maps showing the difference in (a) sea surface temperature, (b) chlorophyll-a, (c)  $\Delta f\text{CO}_2$  before and after Hurricane Maria in 2017. The two time periods are in September 13 – 20, and September 21 – 28. The black contour line shows the path of Hurricane Maria, defined by wind speeds  $> 20 \text{ m s}^{-1}$ .

589 Importantly, we can also show that the response in Figure 12 is not just due to the  
 590 intensification of the wind, but also due to changes in  $f\text{CO}_2$  (Figure 12e,f). OceanCar-  
 591 bNN is able to capture a decrease in  $\Delta f\text{CO}_2$  from the period centered on the 17<sup>th</sup> to the  
 592 following period on the 25<sup>th</sup> of September, particularly between  $\sim 25^{\circ}\text{N}$  to  $\sim 30^{\circ}\text{N}$  (Fig-  
 593 ure 13c). The decrease ( $> |15| \mu\text{atm}$ ) co-occurs with a reduction in sea surface temper-  
 594 ature and an increase in chlorophyll-a (Figure 13a,b), a relationship which has been pre-  
 595 viously observed (Babin et al., 2004; Reul et al., 2021).

596 Mechanistically, the decrease in  $f\text{CO}_2$  is consistent with past studies that have found  
 597 tropical cyclones to cool the surface ocean (decreasing  $f\text{CO}_2$ ), but also inducing mix-  
 598 ing that entrains carbon-rich waters, thus increasing  $f\text{CO}_2$  (Yu et al., 2020). This is fol-  
 599 lowed by an increase in primary productivity (i.e., a reduction in  $f\text{CO}_2$  Babin et al., 2004;  
 600 Lévy et al., 2012; Yu et al., 2020). For our point location, SST decreased by  $2.6^{\circ}\text{C}$  be-  
 601 tween the two periods (Figure 12a), which would lead to a reduction in  $f\text{CO}_2$  of  $\sim 42$   
 602  $\mu\text{atm}$ . However, OceanCarbNN predicts a reduction of  $24 \mu\text{atm}$ , leaving an excess of  $18$   
 603  $\mu\text{atm}$ , which we attribute to the entrainment of DIC-rich waters. However, an increase  
 604 in chlorophyll-a would result in a further decrease in  $f\text{CO}_2$ , meaning that the contribu-  
 605 tion of entrainment could be even larger; however, this contribution cannot be empiri-  
 606 cally determined.

607 While OceanCarbNN captures the impacts of Hurricane Maria on  $F\text{CO}_2$  and  $f\text{CO}_2$ ,  
 608 the magnitude of the event is likely underestimated. For example, observation-based stud-  
 609 ies found that high-velocity winds increased outgassing by  $> 30 \text{ mmolC m}^{-2} \text{ day}^{-1}$  for  
 610  $\sim 24$  hrs (Yu et al., 2020; Ye et al., 2020), compared to the  $2 \text{ mmolC m}^{-2} \text{ day}^{-1}$  increase  
 611 observed from 9 to 17 September (Figure 12d). The 8D resolution averages out the short-  
 612 lasting spikes ( $\sim 24$  hrs). Furthermore, satellite and reanalysis products may underes-  
 613 timate the spikes; for example, ERA5 underestimates extreme wind conditions by be-  
 614 tween 5–10% (relative to satellites Campos et al., 2022). Future work could address this  
 615 by investigating the influence of using 8-daily  $\Delta f\text{CO}_2$  with hourly, daily, 8-daily and monthly  
 616  $k_w$  using an extreme wind speed specific wind speed dataset (e.g., [https://www.maxss](https://www.maxss.org)  
 617 [.org](https://www.maxss.org)).

## 618 7 Discussion

619 We provide here for the first time a global-scale gap filled  $f\text{CO}_2$  product at an 8  
 620 daily and  $0.25^\circ \times 0.25^\circ$  resolution, from which we can compute the air-sea  $\text{CO}_2$  fluxes  
 621 at the same unprecedented resolution. But what are we gaining from this increase in res-  
 622 olution in terms of quality and what are we learning from this in terms of processes? Next,  
 623 we discuss these two questions in turn.

### 624 7.1 Does higher resolution reduce uncertainty in the mapping of $f\text{CO}_2$ ?

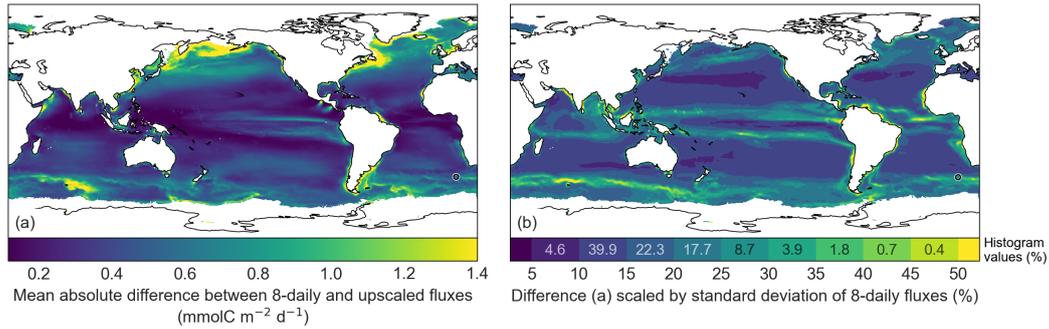
625 In their work investigating a range of gap filled  $f\text{CO}_2$  products at 1M resolution,  
 626 Gregor et al. (2019) found that all products had very similar RMSD estimates ranging  
 627 between 15–20  $\mu\text{atm}$ . They proposed that this is a methodological wall beyond which  
 628 it would be difficult to progress. Our question is thus, can higher resolution get us over  
 629 the wall?

630 In the open ocean, we find little to no reduction in the uncertainties, i.e., the RMSD  
 631 of our 8D product in the open ocean ( $\sim 13 \mu\text{atm}$ ) is similar to that of the previous ap-  
 632 proaches at 1M resolution (Figure 3f). Further, relatively large biases still occur in some  
 633 poorly-sampled regions (Figures 5 and 6). In the coastal ocean there is a some gain, with  
 634 OceanCarbNN having a lower RMSD (25.4  $\mu\text{atm}$ ) relative to OceanSODA-ETHZ (27  $\mu\text{atm}$ ),  
 635 a monthly by  $1^\circ p\text{CO}_2$ -product (Gregor & Gruber, 2021). An improvement in the coastal  
 636 ocean RMSD was also reported by Chau et al. (2024) in the CMEMS-FFNN approach

637 (28.5  $\mu\text{atm}$  to 27.6  $\mu\text{atm}$ ) by increasing the spatial resolution from  $1^\circ$  to  $0.25^\circ$  (though  
638 not the temporal resolution). The slight reduction in coastal RMSD in both approaches  
639 is likely due to the fact that sharper gradients and fine-scale features are better repre-  
640 sented in high-resolution estimates (e.g., Figure 8). In other words, there is better match-  
641 up between the gridded  $f\text{CO}_2$  observations and the gridded predictors, also called the  
642 representation uncertainty by (Gregor & Gruber, 2021).

643 Thus, while there is an improvement in the random uncertainties of the gap-filled  
644  $f\text{CO}_2$  product at 8D resolution, this is not a breakthrough. Some further reduction may  
645 be achieved by going to even higher resolution, especially in time, as this would permit  
646 to capture the more ephemeral or faster developing features, such as the tropical insta-  
647 bility waves and short-lived upwelling events. The big challenge here is the global-scale  
648 availability of predictor variables. The continuing development of global ocean reanal-  
649 yses that also increasingly incorporate ocean color (e.g., Green Mercator) or commer-  
650 cial satellite observations at daily and 10 m spatial resolutions (Shutler et al., 2024) may  
651 soon provide such very high frequency predictors.

652 At the same time, we may be at the limit of the capability of the current gener-  
653 ation of gap-filling methods that use simple architecture. With the current architecture,  
654 where each sample is independent in space and time, gap-filling methods need to solve  
655 two problems: 1) the basin-scale problem, where high accuracy of the inferred  $f\text{CO}_2$  is  
656 absolutely essential for constraining the ocean carbon sink, and 2) a fine-scale problem,  
657 where drivers can cause rapid and short spatial scales variations in  $f\text{CO}_2$  that need to  
658 be tracked with high precision. The challenge for the statistical methods is not simple,  
659 since similar changes in the drivers can have opposite effects on  $f\text{CO}_2$ . A typical exam-  
660 ple occurs during El Niño phases in the eastern equatorial Pacific. Warm phases in this  
661 region typically go together with low  $f\text{CO}_2$  as a result of the cessation of upwelling (Feely  
662 et al., 1999). At the same time, a local heatwave would still push up  $f\text{CO}_2$ . One option  
663 to explore is the use of multiscale methods that use information about the neighboring  
664 observations  $f\text{CO}_2$  or features (in space and time), such as those used for weather fore-  
665 casting (e.g., GraphCast and FourCastNet; Lam et al., 2023; Pathak et al., 2022).



**Figure 14.** (a) The mean absolute difference (MAD) between  $FCO_2^{8D}$  and  $FCO_2^{upscaled}$ , where the latter is calculated with  $k_x^{8D}$  and  $\Delta fCO_2^{1M}$  upscaled to 8D. Thus, (a) shows the impact of using high resolution  $\Delta fCO_2$ , even if a high resolution gas transfer velocity is used. (b) shows the MAD from (a) scaled to the standard deviation of the high resolution fluxes (Figure 9b). The values in the color bar represent the percentage cover that the histogram bin occupies in (b).

## 7.2 What do we learn from high resolution $FCO_2$ ?

Our results show that there are distinct benefits of high resolution predictions of  $fCO_2$  and air-sea  $CO_2$  flux. First, the high-resolution predictions of  $fCO_2$  reduce the uncertainty of the estimates relative to past low-resolution estimates in the coastal ocean due to improved match-ups between the in-situ target data (i.e., SOCAT) and the remote predictors (Section 7.1). And, the higher resolution captures more variability of  $fCO_2$  and air-sea  $CO_2$  fluxes (Figure 11). However, we also show that the majority of the gained variability, particularly at the subseasonal scale, is due to an increase in the temporal resolution of wind represented by  $k_x$ .

Since winds drive the majority of the increase in variability, a key question is whether a similar result can be achieved by “upscaling” the fluxes. The upscaled fluxes ( $FCO_2^{upscaled}$ ) are calculated using low-resolution  $fCO_2$  and high-resolution  $k_x$ . The difference between  $FCO_2^{8D}$  and  $FCO_2^{upscaled}$  (Figure 14a) quantifies the amount of information missed when relying solely on  $FCO_2^{upscaled}$ . In large parts of the ocean, particularly the gyre regions, we find minimal differences between  $FCO_2^{8D}$  and  $FCO_2^{upscaled}$  ( $< 0.4 \text{ mmolC m}^{-2} \text{ d}^{-1}$ , as shown in Figure 14a). However, in dynamic regions of the ocean, the differences can be substantial ( $> 1 \text{ mmolC m}^{-2} \text{ d}^{-1}$ ).

683           Scaling these results by the standard deviation of  $FCO_2^{8D}$  indicates the local im-  
684           portance of using the high resolution  $FCO_2$  (Figure 14b). The Antarctic Circumpolar  
685           Current (ACC) is the most striking feature, where significant additional variability is added  
686           due to the high resolution. Interestingly, this aligns with an observation-based study by  
687           in the Atlantic sector of the ACC that suggested a 3-daily sampling frequency of  $fCO_2$   
688           is required to constrain fluxes to a 10% uncertainty threshold (location indicated by the  
689           marker in Figure 14; Monteiro et al., 2015). This also holds true for the eastern bound-  
690           ary upwelling regions, and the continental shelf regions in the high latitude Pacific and  
691           Atlantic. Thus, it is in regions where there is the combination of high spatio-temporal  
692           variability and high wind speeds that drivers the large differences between  $FCO_2^{8D}$  and  
693            $FCO_2^{upscaled}$ . However, it is also important to recognize that more sporadic events, such  
694           as Hurricane Maria, are also captured, but they may contribute little to the large-scale  
695           variability due to their short-lived nature (Lévy et al., 2012). While a step in the right  
696           direction, we also have to note that the magnitude of the variability is likely underes-  
697           timated at the 8-daily resolution, e.g., Hurricane Maria (Lévy et al., 2012; Yu et al., 2020;  
698           Ye et al., 2020).

699           Despite the increases in local-scale variability, we do not see major differences be-  
700           tween the mean air-sea  $CO_2$  fluxes at 8-daily and monthly-resolutions at global and in-  
701           terannual scales (Figures 9a and 11). Thus, if the aim is to constrain  $fCO_2$  at a large  
702           scale, e.g., for the Global Carbon Budget (Friedlingstein et al., 2022), our analysis sug-  
703           gests that there is little reason to use 8-daily by  $0.25^\circ$  estimates. Importantly, increas-  
704           ing the resolution of  $fCO_2$  does not solve current unanswered problems and questions.  
705           For example, the growing divide in globally integrated  $fCO_2$  between the  $fCO_2$ -products  
706           and Global Ocean Biogeochemical Models (GOBMs) over the last decade (2010–2022).  
707           Several studies have pinned this divergence on the overestimation of decadal variabil-  
708           ity by the  $fCO_2$ -products (Gloege et al., 2021; Hauck et al., 2023), though the evidence  
709           is not conclusive, and it may be GOBMs that underestimate decadal variability (Mayot  
710           et al., 2023). However, the fact remains that  $fCO_2$ -products suffer from observational  
711           paucity and sampling biases (Rödenbeck et al., 2015; Ritter et al., 2017; Gloege et al.,  
712           2021; Hauck et al., 2023), a problem that the high-resolution  $FOC_2$  estimates presented  
713           here cannot solve.

## 714 8 Caveats

715 There are a number of specific caveats and challenges that need to be considered  
 716 in our novel 8D product. The first concerns the use of the NOAA marine boundary layer  
 717 product for the atmospheric dry air mixing ratio ( $x\text{CO}_2^{\text{mb1}}$ ) for the computation of the  
 718 air-sea  $\text{CO}_2$  difference. The second concerns the use of a stacked salinity product in or-  
 719 der to produce a high resolution product spanning four decades.

720 The use of the NOAA  $x\text{CO}_2^{\text{mb1}}$  product most likely will underestimate the oceanic  
 721  $\text{CO}_2$  uptake in regions downwind of the high anthropogenic  $\text{CO}_2$  emission regions. This  
 722 is because the NOAA  $x\text{CO}_2^{\text{mb1}}$  product is constructed from marine stations primarily lo-  
 723 cated in the Pacific Ocean, maximally away from any emissions. This contrasts with many  
 724 regions in the North Atlantic, in the western Pacific, and close to the continents that are  
 725 downstream of the major emitters and thus have substantially higher  $x\text{CO}_2$  than sug-  
 726 gested by the marine boundary layer product (Leinweber et al., 2009). (Palter et al., 2023)  
 727 recently pointed out that this effect can be quite substantial in the downwind regions,  
 728 and suggested that this effect needs to be included in regional assessments. At the same  
 729 time, they pointed out that this effect is negligible when the global ocean uptake is con-  
 730 sidered. Still, as we are pushing the oceanic  $f\text{CO}_2$  to higher resolution, we should also  
 731 pay more attention to the spatio-temporal variations in atmospheric  $\text{CO}_2$ .

732 The second caveat concerns the difficulties associated with the estimation of  $f\text{CO}_2$   
 733 over a four decade period, given limitations in the availability of the predictors of choice  
 734 (Figure 2). For chlorophyll and sea surface height, this means that climatologies are used  
 735 for the periods where there is no coverage (prior to 1998 and 1993 respectively). Given  
 736 that climatologies have smoother fields, the  $f\text{CO}_2$  estimates reflect this smoother input  
 737 data. In the case of salinity, three different products are used over the four decade pe-  
 738 riod (Figures 2, S3a).

739 The difference in variability between the salinity products can be large for some  
 740 regions (e.g., Figure S3e). For the majority of the open ocean, the difference between  
 741 salinity products does not make a significant difference, since salinity is a weak driver  
 742 of  $\Delta f\text{CO}_2$  ( $< |5| \mu\text{atm PSU}^{-1}$ ; Figure S3a). However, in the equatorial Pacific, particu-  
 743 larly in the western part of the basin, the sensitivity of  $\Delta f\text{CO}_2$  to salinity is large ( $>$   
 744  $20 \mu\text{atm PSU}^{-1}$ ). This means that salinity anomalies typically drive a change of more  
 745 than  $\sim 10 \mu\text{atm}$  in the western equatorial Pacific for the SODA v3.4.2 salinity (Figure

746 S3b). For the ESA-CCI salinity, the salinity climate data record product, salinity drives  
747 an average change of  $\sim 6 \mu\text{atm}$ .

748 We stress that this only impacts the variability of  $\Delta f\text{CO}_2$  since the mean state is  
749 captured by the 8-daily climatology of salinity. However, it does mean that variability  
750 of  $f\text{CO}_2$  in the western equatorial Pacific, may be overestimated in the period 1982-1992  
751 (SODA salinity) and underestimated from 1993-2009 (CMEMS-Multiobs). This applies  
752 particularly to the large-scale changes in salinity driven by El Niño. While this is a lesser  
753 problem for  $\Delta f\text{CO}_2$  predictions, this will have a much stronger influence on machine learn-  
754 ing estimates of total alkalinity (e.g., Gregor & Gruber, 2021)

## 755 9 Conclusions and next steps

756 In this study, we present the first 8-daily by  $0.25^\circ \times 0.25^\circ$  estimates of air-sea  $\text{CO}_2$   
757 fluxes at a global scale over four decades. The high-resolution p $\text{CO}_2$ -product, Ocean-  
758 CarbNN, is able to capture significantly more variability in air-sea  $\text{CO}_2$  flux, with the  
759 majority of the increased variability is driven by higher resolution winds, rather than  $f\text{CO}_2$ .  
760 However, we also show that high resolution  $f\text{CO}_2$  is important in regions with high vari-  
761 ability. This includes short-lived high intensity events such as upwelling events and hur-  
762 ricanes.

763 Following the approach of (Gregor & Gruber, 2021), the  $f\text{CO}_2$  estimates from the  
764 OceanCarbNN approach can be combined with high resolution estimates of total alka-  
765 linity to estimate high resolution ocean acidification parameters that could be used to  
766 better understand ocean acidification extremes, for example (Gruber et al., 2021; Desmet  
767 et al., 2022; Burger et al., 2020). Understanding such extreme events from historical data  
768 is important so that the drivers of extremes can be better characterized (Gruber et al.,  
769 2021). This then also allows us to understand current extreme events, such as the North  
770 Atlantic marine heatwave in 2023, that could drive anomalous changes in ocean acid-  
771 ification. This requires near real-time capability of machine learning approaches, which  
772 is technically quite feasible. However, the current release cycle of SOCAT means that  
773 near-real time estimates would be predicting up to 1.5 years beyond the target data (D. C. Bakker  
774 et al., 2016). The impact of predicting beyond the target data needs to be investigated  
775 before results can be used.

776 Another area of improvement lies in the temporal resolution of our air-sea CO<sub>2</sub> fluxes.  
 777 The psuedo-daily estimates of air-sea CO<sub>2</sub> fluxes could quite simply be estimated by up-  
 778 scaling  $\Delta f\text{CO}_2$  to daily resolution while using daily estimates of  $k_w$ . This could be fur-  
 779 ther improved by using daily predictors for those that are available (e.g., SST, and SSH)  
 780 alongside upscaled predictors for those that are not. However, the gain from 8-daily to  
 781 daily  $f\text{CO}_2$  at the current spatial resolution (0.25°) is unlikely to be as large as the im-  
 782 provement from monthly to 8-daily (Monteiro et al., 2015).

783 Finally, the field of machine learning is developing at an unprecedented rate. New  
 784 approaches such as Fourier neural operators used in Nvidia’s FourcastNet could be in-  
 785 corporated to better capture fine-scale variability of  $f\text{CO}_2$ . However, it is unclear if even  
 786 these approaches would be able to reduce the uncertainties beyond “the wall” that sits  
 787 between 15 and 20  $\mu\text{atm}$ .

## 788 10 Open Research

789 The OceanCarbNN dataset of  $\Delta f\text{CO}_2$  and  $F\text{CO}_2$  produced and used throughout  
 790 this study are available at [https://data.up.ethz.ch/shared/ESA-OHOA/OceanSODA](https://data.up.ethz.ch/shared/ESA-OHOA/OceanSODA_ETHZ_HR-v2023.01-full_carbsys/)  
 791 [\\_ETHZ\\_HR-v2023.01-full\\_carbsys/](https://data.up.ethz.ch/shared/ESA-OHOA/OceanSODA_ETHZ_HR-v2023.01-full_carbsys/) (this will change to a repository with a DOI once  
 792 publication is accepted). Code to create the OceanCarbNN data is hosted on [https://](https://gitlab.ethz.ch/oceansoda/oceancarbnn)  
 793 [gitlab.ethz.ch/oceansoda/oceancarbnn](https://gitlab.ethz.ch/oceansoda/oceancarbnn) and code for the study analysis and figures  
 794 are hosted at [https://gitlab.ethz.ch/oceansoda/gbc-gregor-et-al-high-res-variability](https://gitlab.ethz.ch/oceansoda/gbc-gregor-et-al-high-res-variability-fco2)  
 795 [-fco2](https://gitlab.ethz.ch/oceansoda/gbc-gregor-et-al-high-res-variability-fco2) (A DOI will be given to the code repositories once the publication is accepted).  
 796 All data used to create the abovementioned dataset are at least open-access under aca-  
 797 demic license and are listed here. SOCAT v2023 data was downloaded from [https://](https://socat.info/socat_files/v2023/SOCATv2023.tsv.zip)  
 798 [socat.info/socat\\_files/v2023/SOCATv2023.tsv.zip](https://socat.info/socat_files/v2023/SOCATv2023.tsv.zip) (D. C. Bakker et al., 2016; D. C. E. Bakker  
 799 et al., 2023). Sea surface temperature is from <https://doi.org/10.48670/moi-00169>  
 800 (Good et al., 2020). ERA5 data (wind and sea level pressure) are from [https://doi.org/](https://doi.org/10.24381/cds.adbb2d47)  
 801 [10.24381/cds.adbb2d47](https://doi.org/10.24381/cds.adbb2d47) (Hersbach et al., 2020, 2023). Salinity from 2010-2020 is from  
 802 <https://catalogue.ceda.ac.uk/uuid/fad2e982a59d44788eda09e3c67ed7d5> (Boutin  
 803 et al., 2021). Salinity and mixed layer depth from SODA v3.4.2 were downloaded from  
 804 <https://dsrs.atmos.umd.edu/DATA/soda3.4.2/REGRIDED/ocean/> (Carton et al., 2018).  
 805 Salinity after 2021 was downloaded from <https://doi.org/10.48670/moi-00051> (Droghei  
 806 et al., 2016). Chlorophyll-a data can be found at <https://www.oceancolour.org/> (Sathyendranath

807 et al., 2023). We used the reprocessed sea surface height from [https://doi.org/10.48670/](https://doi.org/10.48670/moi-00148)  
 808 [moi-00148](https://doi.org/10.48670/moi-00148) (see acknowledgements).

## 809 References

- 810 Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Zheng, X.  
 811 (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*.  
 812 Retrieved from <https://www.tensorflow.org/> (Software available from  
 813 [tensorflow.org](https://www.tensorflow.org/))
- 814 Alvera-Azcárate, A., Barth, A., Sirjacobs, D., Lenartz, F., & Beckers, J. M. (2011).  
 815 Data interpolating empirical orthogonal functions (DINEOF): A tool for geo-  
 816 physical data analyses. *Mediterranean Marine Science*, *12*(3), 5–11. doi:  
 817 10.12681/mms.64
- 818 Arruda, R., Calil, P. H. R., Bianchi, A. A., Doney, S. C., Gruber, N., Lima, I.,  
 819 & Turi, G. (2015). Air-sea CO<sub>2</sub> fluxes and the controls on ocean surface  
 820 pCO<sub>2</sub> seasonal variability in the coastal and open-ocean southwestern At-  
 821 lantic Ocean: A modeling study. *Biogeosciences*, *12*(19), 5793–5809. doi:  
 822 10.5194/bg-12-5793-2015
- 823 Babin, S. M., Carton, J. A., Dickey, T. D., & Wiggert, J. D. (2004). Satellite ev-  
 824 idence of hurricane-induced phytoplankton blooms in an oceanic desert. *Jour-  
 825 nal of Geophysical Research: Oceans*, *109*(C3). doi: 10.1029/2003JC001938
- 826 Bakker, D. C., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., ... Xu,  
 827 S. (2016). A multi-decade record of high-quality fCO<sub>2</sub> data in version 3 of  
 828 the Surface Ocean CO<sub>2</sub> Atlas (SOCAT). *Earth System Science Data*, *8*(2),  
 829 383–413. doi: 10.5194/essd-8-383-2016
- 830 Bakker, D. C. E., Alin, S. R., Bates, N., Becker, M., Feely, R. A., Gkritzalis, T.,  
 831 ... Wimart-Rousseau, C. (2023). *Surface ocean CO<sub>2</sub> atlas database version*  
 832 *2023 (SOCATv2023) (NCEI accession 0278913)*. NOAA National Centers for  
 833 Environmental Information. doi: 10.25921/r7xa-bt92
- 834 Bates, N. R., Astor, Y. M., Church, M. J., Currie, K., Dore, J. E., González-Dávila,  
 835 M., ... Santana-Casiano, J. M. (2014). A time-series view of changing sur-  
 836 face ocean chemistry due to ocean uptake of anthropogenic CO<sub>2</sub> and ocean  
 837 acidification. *Oceanography*, *27*(1), 126–141. doi: 10.5670/oceanog.2014.16
- 838 Bates, N. R., Takahashi, T., Chipman, D. W., & Knap, A. H. (1998). Variability

- 839 of pCO<sub>2</sub> on diel to seasonal timescales in the Sargasso Sea near Bermuda.  
 840 *Journal of Geophysical Research: Oceans*, 103(C8), 15567–15585. doi:  
 841 10.1029/98JC00247
- 842 Bennington, V., Galjanic, T., & McKinley, G. A. (2022). Explicit Physical Knowl-  
 843 edge in Machine Learning for Ocean Carbon Flux Reconstruction: The pCO<sub>2</sub>-  
 844 Residual Method. *Journal of Advances in Modeling Earth Systems*, 14(10).  
 845 doi: 10.1029/2021MS002960
- 846 Boutin, J., Merlivat, L., Hénoq, C., Martin, N., & Sallée, J. B. (2008). Air-sea  
 847 CO<sub>2</sub> flux variability in frontal regions of the Southern Ocean from CARbon In-  
 848 terface OCEan Atmosphere drifters. *Limnology and Oceanography*, 53(5part2),  
 849 2062–2079. doi: 10.4319/lo.2008.53.5\part\2.2062
- 850 Boutin, J., Vergely, J. L., Marchand, S., D’Amico, F., Hasson, A., Kolodziejczyk,  
 851 N., . . . Vialard, J. (2018). New SMOS Sea Surface Salinity with reduced  
 852 systematic errors and improved variability. *Remote Sensing of Environment*,  
 853 214(August 2017), 115–134. doi: 10.1016/j.rse.2018.05.022
- 854 Boutin, J., Vergely, J.-L., Reul, N., Catany, R., Koehler, J., Martin, A., . . . Don-  
 855 lon, C. (2021). *ESA Sea Surface Salinity Climate Change Initiative*  
 856 *(Sea\_Surface\_Salinity\_cci): Weekly sea surface salinity product, v03.21, for*  
 857 *2010 to 2020*. NERC EDS Centre for Environmental Data Analysis. Retrieved  
 858 Accessed on 21-02-2023, from \url{https://catalogue.ceda.ac.uk/uuid/  
 859 fad2e982a59d44788eda09e3c67ed7d5}
- 860 Burger, F. A., John, J. G., & Frölicher, T. L. (2020). Increase in ocean acidity  
 861 variability and extremes under increasing atmospheric CO<sub>2</sub>. *Biogeosciences*,  
 862 17(18), 4633–4662. doi: 10.5194/bg-17-4633-2020
- 863 Burt, W. J., Thomas, H., Miller, L. A., Granskog, M. A., Papakyriakou, T. N., &  
 864 Pengelly, L. (2016). Inorganic carbon cycling and biogeochemical processes in  
 865 an Arctic inland sea (Hudson Bay). *Biogeosciences*, 13(16), 4659–4671. doi:  
 866 10.5194/BG-13-4659-2016
- 867 Campos, R. M., Gramscianinov, C. B., de Camargo, R., & da Silva Dias, P. L.  
 868 (2022). Assessment and Calibration of ERA5 Severe Winds in the At-  
 869 lantic Ocean Using Satellite Data. *Remote Sensing*, 14(19), 4918. doi:  
 870 10.3390/rs14194918
- 871 Carton, J. A., Chepurin, G. A., & Chen, L. (2018). SODA3: A new ocean climate

- 872 reanalysis. *Journal of Climate*, *31*(17), 6967–6983. doi: 10.1175/JCLI-D-17  
873 -0149.1
- 874 Chapa-Balcorta, C., Hernandez-Ayon, J. M., Durazo, R., Beier, E., Alin, S. R.,  
875 & López-Pérez, A. (2015). Influence of post-Tehuano oceanographic pro-  
876 cesses in the dynamics of the CO<sub>2</sub> system in the Gulf of Tehuantepec, Mex-  
877 ico. *Journal of Geophysical Research: Oceans*, *120*(12), 7752–7770. doi:  
878 10.1002/2015JC011249
- 879 Chau, T. T. T., Gehlen, M., & Chevallier, F. (2022). A seamless ensemble-  
880 based reconstruction of surface ocean pCO<sub>2</sub> and air-sea CO<sub>2</sub> fluxes over  
881 the global coastal and open oceans. *Biogeosciences*, *19*(4), 1087–1109. doi:  
882 10.5194/bg-19-1087-2022
- 883 Chau, T.-T.-T., Gehlen, M., Metzl, N., & Chevallier, F. (2024). CMEMS-LSCE: A  
884 global, 0.25°, monthly reconstruction of the surface ocean carbonate system.  
885 *Earth System Science Data*, *16*(1), 121–160. doi: 10.5194/essd-16-121-2024
- 886 Chen, S., Hu, C., Barnes, B. B., Wanninkhof, R., Cai, W.-J., Barbero, L., & Pierrot,  
887 D. (2019). A machine learning approach to estimate surface ocean pCO<sub>2</sub> from  
888 satellite measurements. *Remote Sensing of Environment*, *228*, 203–226. doi:  
889 10.1016/j.rse.2019.04.019
- 890 Courtois, A., Morel, J. M., & Arias, P. (2023). Can neural networks extrapolate?  
891 Discussion of a theorem by Pedro Domingos. *Revista de la Real Academia de*  
892 *Ciencias Exactas, Físicas y Naturales - Serie A: Matemáticas*, *117*(2), 1–26.  
893 doi: 10.1007/s13398-023-01411-z
- 894 Denvil-Sommer, A., Gehlen, M., Vrac, M., & Mejia, C. (2019). LSCE-FFNN-v1: A  
895 two-step neural network model for the reconstruction of surface ocean p CO  
896 2 over the global ocean. *Geoscientific Model Development*, *12*(5), 2091–2105.  
897 doi: 10.5194/gmd-12-2091-2019
- 898 Desmet, F., Gruber, N., Köhn, E. E., Münnich, M., & Vogt, M. (2022). Tracking the  
899 Space-Time Evolution of Ocean Acidification Extremes in the California Cur-  
900 rent System and Northeast Pacific. *Journal of Geophysical Research: Oceans*,  
901 *127*(5), e2021JC018159. doi: 10.1029/2021JC018159
- 902 DeVries, T., Le Quéré, C., Andrews, O., Berthet, S., Hauck, J., Ilyina, T., . . .  
903 Séférian, R. (2019). Decadal trends in the ocean carbon sink. *Pro-*  
904 *ceedings of the National Academy of Sciences*, *116*(24), 201900371. doi:

- 905 10.1073/pnas.1900371116
- 906 DeVries, T., Yamamoto, K., Wanninkhof, R., Gruber, N., Hauck, J., Müller, J. D.,  
907 ... Zeng, J. (2023). Magnitude, trends, and variability of the global ocean  
908 carbon sink from 1985-2018. *Global Biogeochemical Cycles*, e2023GB007780.  
909 doi: 10.1029/2023GB007780
- 910 Dlugokencky, E., Thoning, K., Lan, X., & Tans, P. (2021). *NOAA Greenhouse Gas*  
911 *Reference from Atmospheric Carbon Dioxide Dry Air Mole Fractions from the*  
912 *NOAA GML Carbon Cycle Cooperative Global Air Sampling Network.*
- 913 Doney, S. C., Lima, I., Feely, R. A., Glover, D. M., Lindsay, K., Mahowald, N., ...  
914 Wanninkhof, R. H. (2009). Mechanisms governing interannual variability in  
915 upper-ocean inorganic carbon system and air-sea CO<sub>2</sub> fluxes: Physical cli-  
916 mate and atmospheric dust. *Deep Sea Research Part II: Topical Studies in*  
917 *Oceanography*, 56(8-10), 640–655. doi: 10.1016/j.dsr2.2008.12.006
- 918 Droghei, R., Buongiorno Nardelli, B., & Santoleri, R. (2016). Combining in situ  
919 and satellite observations to retrieve salinity and density at the ocean sur-  
920 face. *Journal of Atmospheric and Oceanic Technology*, 33(6), 1211–1223. doi:  
921 10.1175/JTECH-D-15-0194.1
- 922 Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M.,  
923 ... Zeng, J. (2021). SeaFlux: Harmonization of air-sea CO<sub>2</sub> fluxes from sur-  
924 face pCO<sub>2</sub> data products using a standardized approach. *Earth System Science*  
925 *Data*, 13(10), 4693–4710. doi: 10.5194/essd-13-4693-2021
- 926 Fay, A. R., & McKinley, G. A. (2014). Global open-ocean biomes: Mean and tempo-  
927 ral variability. *Earth System Science Data*, 6(2), 273–284. doi: 10.5194/essd-6  
928 -273-2014
- 929 Feely, R. A., Wanninkhof, R., Takahashi, T., & Tans, P. (1999). Influence of El Niño  
930 on the equatorial Pacific contribution to atmospheric CO<sub>2</sub> accumulation. *Na-*  
931 *ture*, 398(6728), 597–601. doi: 10.1038/19273
- 932 Friederich, G. E., Ledesma, J., Ulloa, O., & Chavez, F. P. (2008). Air-sea carbon  
933 dioxide fluxes in the coastal southeastern tropical Pacific. *Progress in Oceanog-*  
934 *raphy*, 79(2), 156–166. doi: 10.1016/j.pocean.2008.10.001
- 935 Friederich, G. E., Walz, P. M., Burczynski, M. G., & Chavez, F. P. (2002). Inor-  
936 ganic carbon in the central California upwelling system during the 1997-1999  
937 El Niño-La Niña event. *Progress in Oceanography*, 54(1-4), 185–203. doi:

- 938 10.1016/S0079-6611(02)00049-6
- 939 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Bakker, D. C. E.,  
940 Hauck, J., ... Zheng, B. (2023). Global Carbon Budget 2023. *Earth System*  
941 *Science Data*, 15(12), 5301–5369. doi: 10.5194/essd-15-5301-2023
- 942 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck,  
943 J., ... Zheng, B. (2022). Global Carbon Budget 2022. *Earth System Science*  
944 *Data*, 14(11), 4811–4900. doi: 10.5194/essd-14-4811-2022
- 945 Fukushima, K. (1969). Visual Feature Extraction by a Multilayered Network of Ana-  
946 log Threshold Elements. *IEEE Transactions on Systems Science and Cybernet-*  
947 *ics*, 5(4), 322–333. doi: 10.1109/TSSC.1969.300225
- 948 Gade, K. (2010). A Non-singular Horizontal Position Representation. *Journal of*  
949 *Navigation*, 63(03), 395–417. doi: 10.1017/S0373463309990415
- 950 Gloege, L., McKinley, G. A., Landschützer, P., Fay, A. R., Frölicher, T. L., Fyfe,  
951 J. C., ... Takano, Y. (2021). Quantifying Errors in Observationally Based  
952 Estimates of Ocean Carbon Sink Variability. *Global Biogeochemical Cycles*,  
953 35(4), 1–14. doi: 10.1029/2020GB006788
- 954 Gloege, L., Yan, M., Zheng, T., & McKinley, G. A. (2022). Improved Quantification  
955 of Ocean Carbon Uptake by Using Machine Learning to Merge Global Models  
956 and pCO<sub>2</sub> Data. *Journal of Advances in Modeling Earth Systems*, 14(2), 1–19.  
957 doi: 10.1029/2021MS002620
- 958 González, M. F., & Ilyina, T. (2016). Impacts of artificial ocean alkalization on the  
959 carbon cycle and climate in Earth system simulations. *Geophysical Research*  
960 *Letters*, 43(12), 6493–6502. doi: 10.1002/2016GL068576
- 961 Good, S. A., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., ... Wors-  
962 fold, M. (2020). The current configuration of the OSTIA system for opera-  
963 tional production of foundation sea surface temperature and ice concentration  
964 analyses. *Remote Sensing*, 12(4), 1–20. doi: 10.3390/rs12040720
- 965 Goyet, C., Bradshaw, A. L., & Brewer, P. G. (1991). The carbonate system in  
966 the Black Sea. *Deep Sea Research Part A. Oceanographic Research Papers*,  
967 38(Suppl. 2A), S1049-S1068. doi: 10.1016/S0198-0149(10)80023-8
- 968 Gregor, L., & Gruber, N. (2021). OceanSODA-ETHZ: A global gridded data  
969 set of the surface ocean carbonate system for seasonal to decadal studies  
970 of ocean acidification. *Earth System Science Data*, 13(2), 777–808. doi:

- 971 10.5194/essd-13-777-2021
- 972 Gregor, L., Lebehot, A. D., Kok, S., & Scheel Monteiro, P. M. (2019). A com-  
 973 parative assessment of the uncertainties of global surface ocean CO<sub>2</sub> esti-  
 974 mates using a machine-learning ensemble (CSIR-ML6 version 2019a) - have  
 975 we hit the wall? *Geoscientific Model Development*, *12*(12), 5113–5136. doi:  
 976 10.5194/gmd-12-5113-2019
- 977 Gruber, N., Bakker, D. C. E., DeVries, T., Gregor, L., Hauck, J., Landschützer,  
 978 P., ... Müller, J. D. (2023). Trends and variability in the ocean car-  
 979 bon sink. *Nature Reviews Earth & Environment*, *4*(2), 119–134. doi:  
 980 10.1038/s43017-022-00381-x
- 981 Gruber, N., Boyd, P. W., Frölicher, T. L., & Vogt, M. (2021). Biogeochemical ex-  
 982 tremes and compound events in the ocean. *Nature*, *600*(7889), 395–407. doi:  
 983 10.1038/s41586-021-03981-7
- 984 Gruber, N., Clement, D., Carter, B. R., Feely, R. A., van Heuven, S., Hoppema, M.,  
 985 ... Wanninkhof, R. H. (2019). The oceanic sink for anthropogenic CO<sub>2</sub> from  
 986 1994 to 2007. *Science*, *363*(6432), 1193–1199. doi: 10.1126/science.aau5153
- 987 Gruber, N., Landschützer, P., & Lovenduski, N. S. (2019). The Variable Southern  
 988 Ocean Carbon Sink. *Annual Review of Marine Science*, *11*(1), 159–186. doi:  
 989 10.1146/annurev-marine-121916-063407
- 990 Gu, Y., Katul, G. G., & Cassar, N. (2023). Multiscale Temporal Variability of the  
 991 Global Air-Sea CO<sub>2</sub> Flux Anomaly. *Journal of Geophysical Research: Biogeo-*  
 992 *sciences*, *128*(6). doi: 10.1029/2022JG006934
- 993 Hauck, J., Nissen, C., Landschützer, P., Rödenbeck, C., Bushinsky, S., & Olsen,  
 994 A. (2023). Sparse observations induce large biases in estimates of the global  
 995 ocean CO<sub>2</sub> sink: An ocean model subsampling experiment. *Philosophical*  
 996 *Transactions of the Royal Society A: Mathematical, Physical and Engineering*  
 997 *Sciences*, *381*(2249). doi: 10.1098/rsta.2022.0063
- 998 Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J.,  
 999 ... Thépaut, J.-N. (2023). *Era5 hourly data on single levels from 1940 to*  
 1000 *present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS).  
 1001 doi: 10.24381/cds.adbb2d47
- 1002 Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J.,  
 1003 ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of*

- 1004            *the Royal Meteorological Society*, 146(730), 1999–2049. doi: 10.1002/qj.3803
- 1005    Holder, C., & Gnanadesikan, A.    (2021).    Can machine learning extract the mech-  
1006            anisms controlling phytoplankton growth from large-scale observations?-A  
1007            proof-of-concept study.    *Biogeosciences*, 18(6), 1941–1970.    doi: 10.5194/  
1008            BG-18-1941-2021
- 1009    Iida, Y., Takatani, Y., Kojima, A., & Ishii, M.    (2021).    Global trends of ocean  
1010            CO<sub>2</sub> sink and ocean acidification: An observation-based reconstruction of sur-  
1011            face ocean inorganic carbon variables.    *Journal of Oceanography*, 77(2).    doi:  
1012            10.1007/s10872-020-00571-5
- 1013    Jähne, B., Heinz, G., & Dietrich, W.    (1987).    Measurement of the diffusion coef-  
1014            ficients of sparingly soluble gases in water.    *Journal of Geophysical Research*,  
1015            92(C10), 10767. doi: 10.1029/JC092iC10p10767
- 1016    Jones, S. D., Le Quéré, C., & Rdenbeck, C.    (2012).    Autocorrelation characteristics  
1017            of surface ocean pCO<sub>2</sub> and air-sea CO<sub>2</sub> fluxes.    *Global Biogeochemical Cycles*,  
1018            26(2), 1–12. doi: 10.1029/2010GB004017
- 1019    Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., . . . Liu, T.-Y.    (2017).  
1020            LightGBM: A highly efficient gradient boosting decision tree. In *Proceedings of*  
1021            *the 31st International Conference on Neural Information Processing Systems*  
1022            (pp. 3149–3157). Red Hook, NY, USA: Curran Associates Inc.
- 1023    Khatiwala, S., Tanhua, T., Mikaloff Fletcher, S. E., Gerber, M., Doney, S. C.,  
1024            Graven, H. D., . . . Sabine, C. L. (2013). Global ocean storage of anthropogenic  
1025            carbon. *Biogeosciences*, 10(4), 2169–2191. doi: 10.5194/bg-10-2169-2013
- 1026    Koch, J., McKinley, G. A., Bennington, V., & Ullman, D.    (2009).    Do hurri-  
1027            canes cause significant interannual variability in the air-sea CO<sub>2</sub> flux of  
1028            the subtropical North Atlantic?    *Geophysical Research Letters*, 36(7).    doi:  
1029            10.1029/2009GL037553
- 1030    Lachkar, Z., & Gruber, N.    (2013).    Response of biological production and air-  
1031            sea CO<sub>2</sub> fluxes to upwelling intensification in the California and Canary  
1032            Current Systems.    *Journal of Marine Systems*, 109–110, 149–160.    doi:  
1033            10.1016/j.jmarsys.2012.04.003
- 1034    Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet,  
1035            F., . . . Battaglia, P.    (2023).    Learning skillful medium-range global weather  
1036            forecasting. *Science*, 382(6677), 1416–1421. doi: 10.1126/science.adi2336

- 1037 Landschützer, P., Gruber, N., & Bakker, D. C. (2016). Decadal variations and  
1038 trends of the global ocean carbon sink. *Global Biogeochemical Cycles*, *30*(10),  
1039 1396–1417. doi: 10.1002/2015GB005359
- 1040 Landschützer, P., Gruber, N., Bakker, D. C., Schuster, U., Nakaoka, S., Payne,  
1041 M. R., ... Zeng, J. (2013). A neural network-based estimate of the seasonal  
1042 to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences*,  
1043 *10*(11), 7793–7815. doi: 10.5194/bg-10-7793-2013
- 1044 Landschützer, P., Gruber, N., Bakker, D. C., Stemmler, I., & Six, K. D. (2018).  
1045 Strengthening seasonal marine CO<sub>2</sub> variations due to increasing atmo-  
1046 spheric CO<sub>2</sub>. *Nature Climate Change*, *8*(February), 146–150. doi:  
1047 10.1038/s41558-017-0057-x
- 1048 Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C.,  
1049 van Heuven, S., ... Wanninkhof, R. H. (2015). The reinvigoration of  
1050 the Southern Ocean carbon sink. *Science*, *349*(6253), 1221–1224. doi:  
1051 10.1126/science.aab2620
- 1052 Laruelle, G. G., Landschützer, P., Gruber, N., Ti, J. L., Delille, B., Regnier, P. A.,  
1053 & Tison, J.-L. (2017). Global high-resolution monthly pCO<sub>2</sub> climatology for  
1054 the coastal ocean derived from neural network interpolation. *Biogeosciences*,  
1055 *14*(19), 4545–4561. doi: 10.5194/bg-14-4545-2017
- 1056 Legeckis, R. (1977). Long Waves in the Eastern Equatorial Pacific Ocean: A View  
1057 from a Geostationary Satellite. *Science*, *197*(4309), 1179–1181. doi: 10.1126/  
1058 science.197.4309.1179
- 1059 Leinweber, A., Gruber, N., Frenzel, H., Friederich, G. E., & Chavez, F. P. (2009).  
1060 Diurnal carbon cycling in the surface ocean and lower atmosphere of  
1061 Santa Monica Bay, California. *Geophysical Research Letters*, *36*(8). doi:  
1062 10.1029/2008GL037018
- 1063 Lenton, A., Matear, R. J., Keller, D. P., Scott, V., & Vaughan, N. E. (2018). As-  
1064 sessing carbon dioxide removal through global and regional ocean alkalization  
1065 under high and low emission pathways. *Earth System Dynamics*, *9*(2), 339–  
1066 357. doi: 10.5194/esd-9-339-2018
- 1067 Lévy, M., Lengaigne, M., Bopp, L., Vincent, E. M., Madec, G., Ethé, C., ...  
1068 Sarma, V. V. (2012). Contribution of tropical cyclones to the air-sea  
1069 CO<sub>2</sub> flux: A global view. *Global Biogeochemical Cycles*, *26*(2). doi:

- 1070 10.1029/2011GB004145
- 1071 Liang, J.-H., McWilliams, J. C., & Gruber, N. (2009). High-frequency response of  
1072 the ocean to mountain gap winds in the northeastern tropical Pacific. *Journal*  
1073 *of Geophysical Research: Oceans*, *114*(C12). doi: 10.1029/2009JC005370
- 1074 Ma, D., Gregor, L., & Gruber, N. (2023). Four Decades of Trends and Drivers of  
1075 Global Surface Ocean Acidification. *Global Biogeochemical Cycles*, *37*(7). doi:  
1076 10.1029/2023GB007765
- 1077 Mayot, N., Le Quéré, C., Rödenbeck, C., Bernardello, R., Bopp, L., Djeutchouang,  
1078 L. M., ... Zeng, J. (2023). Climate-driven variability of the Southern Ocean  
1079 CO<sub>2</sub> sink. *Philosophical Transactions of the Royal Society A: Mathematical,*  
1080 *Physical and Engineering Sciences*, *381*(2249). doi: 10.1098/rsta.2022.0055
- 1081 McKinley, G. A., Fay, A. R., Eddebbbar, Y. A., Gloege, L., & Lovenduski, N. S.  
1082 (2020). External Forcing Explains Recent Decadal Variability of the  
1083 Ocean Carbon Sink. *AGU Advances*, *1*(2), e2019AV000149. doi: 10.1029/  
1084 2019AV000149
- 1085 Merchant, C. J., Embury, O., Bulgin, C. E., Block, T., Corlett, G. K., Fiedler, E.,  
1086 ... Donlon, C. J. (2019). Satellite-based time-series of sea-surface temper-  
1087 ature since 1981 for climate applications. *Scientific data*, *6*(1), 223. doi:  
1088 10.1038/s41597-019-0236-x
- 1089 Merlivat, L., Gonzalez Davila, M., Caniaux, G., Boutin, J., & Reverdin, G. (2009).  
1090 Mesoscale and diel to monthly variability of CO<sub>2</sub> and carbon fluxes at the  
1091 ocean surface in the northeastern Atlantic. *Journal of Geophysical Research:*  
1092 *Oceans*, *114*(C3). doi: 10.1029/2007JC004657
- 1093 Mignot, A., von Schuckmann, K., Landschützer, P., Gasparin, F., van Gennip, S.,  
1094 Perruche, C., ... Amm, T. (2022). Decrease in air-sea CO<sub>2</sub> fluxes caused  
1095 by persistent marine heatwaves. *Nature Communications*, *13*(1), 4300. doi:  
1096 10.1038/s41467-022-31983-0
- 1097 Monteiro, P. M. S., Gregor, L., Lévy, M., Maenner, S., Sabine, C. L., & Swart, S.  
1098 (2015). Intraseasonal variability linked to sampling alias in air-sea CO<sub>2</sub> fluxes  
1099 in the Southern Ocean. *Geophysical Research Letters*, *42*(20), 8507–8514. doi:  
1100 10.1002/2015GL066009
- 1101 Müller, J. D., Gruber, N., Carter, B., Feely, R., Ishii, M., Lange, N., ... Zhu,  
1102 D. (2023). Decadal Trends in the Oceanic Storage of Anthropogenic

- 1103 Carbon From 1994 to 2014. *AGU Advances*, 4(4), e2023AV000875. doi:  
1104 10.1029/2023AV000875
- 1105 Müller, J. D., Schneider, B., & Rehder, G. (2016). Long-term alkalinity trends in the  
1106 Baltic Sea and their implications for CO<sub>2</sub>-induced acidification. *Limnology and*  
1107 *Oceanography*, 61(6), 1984–2002. doi: 10.1002/lno.10349
- 1108 Murphy, P. P., Nojiri, Y., Harrison, D. E., & Larkin, N. K. (2001). Scales of spatial  
1109 variability for surface ocean pCO<sub>2</sub> in the Gulf of Alaska and Bering Sea: To-  
1110 ward a sampling strategy. *Geophysical Research Letters*, 28(6), 1047–1050. doi:  
1111 10.1029/2000GL012375
- 1112 Naegler, T. (2009). Reconciliation of excess 14C-constrained global CO<sub>2</sub> piston  
1113 velocity estimates. *Tellus, Series B: Chemical and Physical Meteorology*, 61  
1114 B(2), 372–384. doi: 10.1111/j.1600-0889.2008.00408.x
- 1115 Nicholson, S. A., Whitt, D. B., Fer, I., du Plessis, M. D., Lebéhot, A. D., Swart,  
1116 S., ... Monteiro, P. M. (2022). Storms drive outgassing of CO<sub>2</sub> in the  
1117 subpolar Southern Ocean. *Nature Communications*, 13(1), 1–12. doi:  
1118 10.1038/s41467-021-27780-w
- 1119 Palter, J. B., Nickford, S., & Mu, L. (2023). Ocean Carbon Dioxide Uptake in the  
1120 Tailpipe of Industrialized Continents. *Geophysical Research Letters*, 50(21),  
1121 e2023GL104822. doi: 10.1029/2023GL104822
- 1122 Pardo, P. C., Tilbrook, B., van Ooijen, E., Passmore, A., Neill, C., Jansen, P., ...  
1123 Trull, T. W. (2019). Surface ocean carbon dioxide variability in South Pacific  
1124 boundary currents and Subantarctic waters. *Scientific Reports*, 9(1), 7592. doi:  
1125 10.1038/s41598-019-44109-2
- 1126 Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mar-  
1127 dani, M., ... Anandkumar, A. (2022). *FourCastNet: A Global Data-driven*  
1128 *High-resolution Weather Model using Adaptive Fourier Neural Operators* (No.  
1129 arXiv:2202.11214). arXiv. doi: 10.48550/arXiv.2202.11214
- 1130 Regnier, P. A., Resplandy, L., Najjar, R. G., & Ciais, P. (2022). The land-to-ocean  
1131 loops of the global carbon cycle. *Nature 2022 603:7901*, 603(7901), 401–410.  
1132 doi: 10.1038/s41586-021-04339-9
- 1133 Resplandy, L., Hogikyan, A., Müller, J. D., Najjar, R. G., Bange, H. W., Bianchi,  
1134 D., ... Regnier, P. (2024). A Synthesis of Global Coastal Ocean Green-  
1135 house Gas Fluxes. *Global Biogeochemical Cycles*, 38(1), e2023GB007803. doi:

- 1136 10.1029/2023GB007803
- 1137 Reul, N., Chapron, B., Grodsky, S. A., Guimbard, S., Kudryavtsev, V., Foltz, G. R.,  
1138 & Balaguru, K. (2021). Satellite Observations of the Sea Surface Salin-  
1139 ity Response to Tropical Cyclones. *Geophysical Research Letters*, *48*(1),  
1140 e2020GL091478. doi: 10.1029/2020GL091478
- 1141 Ritter, R., Landschützer, P., Gruber, N., Fay, A. R., Iida, Y., Jones, S. D., ... Zeng,  
1142 J. (2017). Observation-Based Trends of the Southern Ocean Carbon Sink. *Geo-*  
1143 *physical Research Letters*, *44*(24), 12,339–12,348. doi: 10.1002/2017GL074837
- 1144 Rödenbeck, C., Bakker, D. C., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S. D.,  
1145 ... Zeng, J. (2015). Data-based estimates of the ocean carbon sink variability  
1146 - first results of the surface ocean pco2 mapping intercomparison (SOCOM).  
1147 *Biogeosciences*, *12*(23), 7251–7278. doi: 10.5194/bg-12-7251-2015
- 1148 Rödenbeck, C., Bakker, D. C., Metzl, N., Olsen, A., Sabine, C. L., Cassar, N.,  
1149 ... Heimann, M. (2014). Interannual sea–air CO<sub>2</sub> flux variability from an  
1150 observation-driven ocean mixed-layer scheme. *Biogeosciences*, *11*(17), 4599–  
1151 4613. doi: 10.5194/bg-11-4599-2014
- 1152 Rödenbeck, C., DeVries, T., Hauck, J., Le Quéré, C., & Keeling, R. F. (2022). Data-  
1153 based estimates of interannual sea–air CO<sub>2</sub> flux variations 1957–2020 and their  
1154 relation to environmental drivers. *Biogeosciences*, *19*(10), 2627–2652. doi:  
1155 10.5194/bg-19-2627-2022
- 1156 Rodgers, K. B., Schwinger, J., Fassbender, A. J., Landschützer, P., Yamaguchi, R.,  
1157 Frenzel, H., ... Velo, A. (2023). Seasonal Variability of the Surface Ocean Car-  
1158 bon Cycle: A Synthesis. *Global Biogeochemical Cycles*, *37*(9), e2023GB007798.  
1159 doi: 10.1029/2023GB007798
- 1160 Romero-Centeno, R., Zavala-Hidalgo, J., Gallegos, A., & O'Brien, J. J. (2003). Isth-  
1161 mus of Tehuantepec Wind Climatology and ENSO Signal. *Journal of Climate*,  
1162 *16*(15), 2628–2639. doi: 10.1175/1520-0442(2003)016<2628:IOTWCA>2.0.CO;  
1163 2
- 1164 Roobaert, A., Regnier, P., Landschützer, P., & Laruelle, G. G. (2023). A novel sea  
1165 surface pCO<sub>2</sub>-product for the global coastal ocean resolving trends over the  
1166 1982–2020 period. *Earth System Science Data Discussions*, 1–32. doi:  
1167 10.5194/essd-2023-228
- 1168 Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., ...

- 1169 Rios, A. F. (2004). The Oceanic Sink for Anthropogenic CO<sub>2</sub>. *Science*,  
1170 305(5682), 367–371. doi: 10.1126/science.1097403
- 1171 Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C., Pfeil, B., Olsen, A., ...  
1172 Yoshikawa-Inoue, H. (2013). Surface Ocean CO<sub>2</sub> Atlas (SOCAT) grid-  
1173 ded data products. *Earth System Science Data*, 5(1), 145–153. doi:  
1174 10.5194/essd-5-145-2013
- 1175 Sarmiento, J. L., & Gruber, N. (2006). *Ocean Biogeochemical Dynamics*. Princeton,  
1176 NJ: Princeton University Press.
- 1177 Sathyendranath, S., Jackson, T., Brockmann, C., Brotas, V., Calton, B.,  
1178 Chuprin, A., ... Platt, T. (2023). *ESA Ocean Colour Climate*  
1179 *Change Initiative (Ocean.Colour\_cci): Version 6.0, 4km resolution data*.  
1180 NERC EDS Centre for Environmental Data Analysis. doi: 10.5285/  
1181 5011D22AAE5A4671B0CBC7D05C56C4F0
- 1182 Shutler, J. D., Gruber, N., Findlay, H. S., Land, P. E., Gregor, L., Holding, T., ...  
1183 Gaultier, L. (2024). The increasing importance of satellite observations to  
1184 assess the ocean carbon sink and ocean acidification. *Earth-Science Reviews*,  
1185 104682. doi: 10.1016/j.earscirev.2024.104682
- 1186 Shutler, J. D., Wanninkhof, R. H., Nightingale, P. D., Woolf, D. K., Bakker, D. C.,  
1187 Watson, A. J., ... Donlon, C. J. (2020). Satellites will address critical sci-  
1188 ence priorities for quantifying ocean carbon. *Frontiers in Ecology and the*  
1189 *Environment*, 18(1), 27–35. doi: 10.1002/fee.2129
- 1190 Song, Z., Yu, S., Bai, Y., Guo, X., He, X., Zhai, W., & Dai, M. (2023). Construction  
1191 of a High Spatiotemporal Resolution Dataset of Satellite-Derived pCO<sub>2</sub> and  
1192 Air–Sea CO<sub>2</sub> Flux in the South China Sea (2003–2019). *IEEE Transactions on*  
1193 *Geoscience and Remote Sensing*, 61, 1–15. doi: 10.1109/TGRS.2023.3306389
- 1194 Sutton, A. J., Feely, R. A., Maenner-Jones, S., Musielwicz, S., Osborne, J., Dietrich,  
1195 C., ... Weller, R. A. (2019). Autonomous seawater pCO<sub>2</sub> and pH time se-  
1196 ries from 40 surface buoys and the emergence of anthropogenic trends. *Earth*  
1197 *System Science Data*, 11(1), 421–439. doi: 10.5194/essd-11-421-2019
- 1198 Sutton, A. J., Sabine, C. L., Maenner-Jones, S., Lawrence-Slavas, N., Meinig, C.,  
1199 Feely, R. A., ... Kozyr, A. (2014). A high-frequency atmospheric and seawater  
1200 pCO<sub>2</sub> data set from 14 open-ocean sites using a moored autonomous system.  
1201 *Earth System Science Data*, 6(2), 353–366. doi: 10.5194/essd-6-353-2014

- 1202 Sutton, A. J., Wanninkhof, R. H., Sabine, C. L., Feely, R. A., Cronin, M. F., &  
 1203 Weller, R. A. (2017). Variability and trends in surface seawater  $p\text{CO}_2$  and  
 1204  $\text{CO}_2$  flux in the Pacific Ocean. *Geophysical Research Letters*, *44*(11), 5627–  
 1205 5636. doi: 10.1002/2017GL073814
- 1206 Sutton, A. J., Williams, N. L., & Tilbrook, B. (2021). Constraining Southern  
 1207 Ocean  $\text{CO}_2$  Flux Uncertainty Using Uncrewed Surface Vehicle Observations.  
 1208 *Geophysical Research Letters*, *48*(3), 1–9. doi: 10.1029/2020GL091748
- 1209 Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G. A., Sarmiento,  
 1210 J. L., & Wanninkhof, R. H. (2007). Constraining global air-sea gas exchange  
 1211 for  $\text{CO}_2$  with recent bomb  $^{14}\text{C}$  measurements. *Global Biogeochemical Cycles*,  
 1212 *21*(2), 1–10. doi: 10.1029/2006GB002784
- 1213 Taburet, G., Sanchez-Roman, A., Ballarotta, M., Pujol, M.-I., Legeais, J.-F.,  
 1214 Fournier, F., . . . Dibarboure, G. (2019). DUACS DT2018: 25 years of re-  
 1215 processed sea level altimetry products. *Ocean Science*, *15*(5), 1207–1224. doi:  
 1216 10.5194/os-15-1207-2019
- 1217 Takahashi, T., Sutherland, S. C., Wanninkhof, R. H., Sweeney, C., Feely, R. A.,  
 1218 Chipman, D. W., . . . de Baar, H. J. (2009). Climatological mean and  
 1219 decadal change in surface ocean  $p\text{CO}_2$ , and net sea–air  $\text{CO}_2$  flux over the  
 1220 global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*,  
 1221 *56*(8-10), 554–577. doi: 10.1016/j.dsr2.2008.12.009
- 1222 Tans, P. P., & Keeling, R. F. (2023). *Global Monitoring Laboratory - Carbon Cycle*  
 1223 *Greenhouse Gases*.
- 1224 Telszewski, M., Chazottes, A., Schuster, U., Watson, A. J., Moulin, C., Bakker,  
 1225 D. C. E., . . . Wanninkhof, R. (2009). Estimating the monthly  $p\text{CO}_2$  distri-  
 1226 bution in the North Atlantic using a self-organizing neural network. *Biogeo-*  
 1227 *sciences*, *6*(8), 1405–1421. doi: 10.5194/bg-6-1405-2009
- 1228 Torres, O., Kwiatkowski, L., Sutton, A. J., Dorey, N., & Orr, J. C. (2021). Char-  
 1229 acterizing Mean and Extreme Diurnal Variability of Ocean  $\text{CO}_2$  System  
 1230 Variables Across Marine Environments. *Geophysical Research Letters*, *48*(5),  
 1231 1–12. doi: 10.1029/2020GL090228
- 1232 Turi, G., Lachkar, Z., & Gruber, N. (2014). Spatiotemporal variability and  
 1233 drivers of  $p\text{CO}_2$  and air–sea  $\text{CO}_2$  fluxes in the California Current System:  
 1234 An eddy-resolving modeling study. *Biogeochemistry*, *11*(3), 671–690. doi:

- 1235 10.5194/bg-11-671-2014
- 1236 Wanninkhof, R., Pickers, P. A., Omar, A. M., Sutton, A., Murata, A., Olsen, A., ...  
1237 Schuster, U. (2019). A Surface Ocean CO<sub>2</sub> Reference Network, SOCONET  
1238 and Associated Marine Boundary Layer CO<sub>2</sub> Measurements. *Frontiers in*  
1239 *Marine Science*, 6.
- 1240 Wanninkhof, R. H. (2014). Relationship between wind speed and gas exchange over  
1241 the ocean revisited. *Limnology and Oceanography: Methods*, 12(6), 351–362.  
1242 doi: 10.4319/lom.2014.12.351
- 1243 Weiss, R. (1974). Carbon dioxide in water and seawater: The solubility of a non-  
1244 ideal gas. *Marine Chemistry*, 2(3), 203–215. doi: 10.1016/0304-4203(74)90015  
1245 -2
- 1246 Weiss, R., & Price, B. (1980). Nitrous oxide solubility in water and seawater. *Ma-*  
1247 *rine Chemistry*, 8(4), 347–359. doi: 10.1016/0304-4203(80)90024-9
- 1248 Whitt, D. B., Nicholson, S. A., & Carranza, M. M. (2019). Global Impacts of Sub-  
1249 seasonal (<60 Day) Wind Variability on Ocean Surface Stress, Buoyancy Flux,  
1250 and Mixed Layer Depth. *Journal of Geophysical Research: Oceans*, 124(12),  
1251 8798–8831. doi: 10.1029/2019JC015166
- 1252 Woolf, D. K., Land, P. E., Shutler, J. D., Goddijn-Murphy, L., & Donlon, C. J.  
1253 (2016). On the calculation of air-sea fluxes of CO<sub>2</sub> in the presence of temper-  
1254 ature and salinity gradients. *Journal of Geophysical Research: Oceans*, 121(2),  
1255 1229–1248. doi: 10.1002/2015JC011427
- 1256 Ye, H., Morozov, E., Tang, D., Wang, S., Liu, Y., Li, Y., & Tang, S. (2020). Varia-  
1257 tion of pCO<sub>2</sub> concentrations induced by tropical cyclones "wind-Pump" in the  
1258 middlelatitude surface oceans: A comparative study. *PLoS ONE*, 15(3). doi:  
1259 10.1371/journal.pone.0226189
- 1260 Yu, P., Wang, Z. A., Churchill, J., Zheng, M., Pan, J., Bai, Y., & Liang, C. (2020).  
1261 Effects of Typhoons on Surface Seawater pCO<sub>2</sub> and Air-Sea CO<sub>2</sub> Fluxes in the  
1262 Northern South China Sea. *Journal of Geophysical Research: Oceans*, 125(8).  
1263 doi: 10.1029/2020JC016258

## 1264 Acknowledgments

1265 We are indebted to the large number of scientists who have measured the surface ocean  
1266 *f*CO<sub>2</sub> over the last few decades, and to the programs and agencies that enabled or funded

1267 these observations. We are also very grateful to all those who have assembled and qual-  
1268 ity controlled these observations into readily usable products. A special thanks to the  
1269 Surface Ocean CO<sub>2</sub> Atlas (SOCAT) team under leadership of Dorothee Bakker. SOCAT  
1270 is an international effort, endorsed by the International Ocean Carbon Coordination Project  
1271 (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS) and the Integrated Ma-  
1272 rine Biosphere Research (IMBeR) program, to deliver a uniformly quality-controlled sur-  
1273 face ocean CO<sub>2</sub> database. We are also grateful to the many other programs that have  
1274 collected and processed marine observations in the last few decades, enabling the devel-  
1275 opment of novel statistical techniques for interpreting the CO<sub>2</sub> observations. The work  
1276 of LG, JS, and NG was supported by the ESA OceanHealth-OA project (contract num-  
1277 ber 4000137603/22/I-DT). JS has received further support by the ESA MAXSS project  
1278 (4000132954/20/I-NB). The Ssalto/Duacs altimeter products were produced and distributed  
1279 by the Copernicus Marine and Environment Monitoring Service (CMEMS) (<http://www.marine.copernicus.eu>).