

1           **Improving GCM-based decadal ocean carbon flux**  
2           **predictions using observationally-constrained statistical**  
3           **models**

4                           **Gooya, P. <sup>1</sup>, Swart, N. C. <sup>1</sup>, Landschützer, P. <sup>2</sup>**

5                           <sup>1</sup>Canadian Centre for Climate Modeling and Analysis (CCCma)  
6   <sup>2</sup>Flanders Marine Institute (VLIZ)

7           **Key Points:**

- 8           • We use observationally trained statistical models to obtain decadal predictions of  
9           ocean carbon flux from initialized GCM-based predictors.  
10          • The hybrid GCM-statistical ocean carbon flux predictions show improved skill over  
11          hindcast predictions from the GCM's biogeochemical models.  
12          • The hybrid models are used to make decadal predictions for the ocean-atmosphere  
13          carbon flux over the decade ending in 2029.

---

Corresponding author: Parsa Gooya, [parsa.g76@gmail.com](mailto:parsa.g76@gmail.com)

## Abstract

Initialized climate model simulations have proven skillful for near-term predictability of the key physical climate variables. By comparison, predictions of biogeochemical fields like ocean carbon flux, are still emerging. Initial studies indicate skillful predictions are possible for lead-times up to six years at global scale for some CMIP6 models. However, unlike core physical variables, biogeochemical variables are not directly initialized in existing decadal prediction systems, and extensive empirical parametrization of ocean-biogeochemistry in Earth System Models introduces a significant source of uncertainty. Here we propose a new approach for improving the skill of decadal ocean carbon flux predictions using observationally-constrained statistical models, as alternatives to the ocean-biogeochemistry models. We use observations to train multi-linear and neural-network models to predict the ocean carbon flux. To account for observational uncertainties, we train using six different observational estimates of the flux. We then apply these trained statistical models using input predictors from the Canadian Earth System Model (CanESM5) decadal prediction system to produce new decadal predictions. Our hybrid GCM-statistical approach significantly improves prediction skill, relative to the raw CanESM5 hindcast predictions over 1990-2019. Our hybrid-model skill is also larger than that obtained by any available CMIP6 model. Using bias-corrected CanESM5 predictors, we make forecasts for ocean carbon flux over 2020-2029. Both statistical models predict increases in the ocean carbon flux larger than the changes predicted from CanESM5 forecasts. Our work highlights the ability to improve decadal ocean carbon flux predictions by using observationally-trained statistical models together with robust input predictors from GCM-based decadal predictions.

## Plain Language Summary

Using initialized Earth system model simulations for near term predictions of ocean biogeochemical variables is an emerging field of research. In particular, near term predictability of ocean carbon flux is central to efforts for planing and limiting climate change. Unlike physical variables whose predictability have been established, these simulations are only indirectly initialized and rely on heavily parameterized ocean biogeochemistry models. Here, we propose a new approach to acquire decadal predictions of air-sea carbon flux as alternatives to those based on ocean biogeochemistry models. Our methodology combines the explanatory power of statistical models that have widely been used for gap filling purposes for informing full coverage ocean carbon flux data products, and well established predictability skill of key physical predictors. We provide hybrid GCM-statistical ocean carbon flux hindcasts using predictors from CanESM5 and doing so, show that we can beat all CMIP6 decadal prediction system hindcast skills. We use our models to provide near future hybrid model forecast for ocean carbon flux. Our results shows the potential for improving predictability skill of ocean carbon sink by combining GCMs and observationally trained statistical models.

## 1 Introduction

The ocean accounts for sequestering nearly 25% percent of human CO<sub>2</sub> emissions annually (Hauck et al., 2020; Friedlingstein et al., 2022, 2020), playing a key role in mitigating climate change. Future changes in the ocean carbon flux are of direct relevance to climate change science (Friedlingstein et al., 2022) and policy making related to climate and emissions targets. Ocean carbon uptake has increased substantially over the past several decades in response to human induced increases in atmospheric CO<sub>2</sub> concentrations (Gooya et al., 2023; Rodgers et al., 2020; Lovenduski et al., 2016; McKinley et al., 2016; Wang et al., 2016). However, there is also substantial internal variability in the magnitude of the flux on seasonal to decadal time scales both regionally and globally (Landschützer et al., 2016; McKinley et al., 2017; Gruber et al., 2019; McKin-

64 ley et al., 2020). Decadal scale variability of ocean carbon flux is believed to be driven  
65 largely by variability in external forcing (McKinley et al., 2020), and specifically, the de-  
66 viations of atmospheric growth of CO<sub>2</sub> from the long term trend but also changes in cir-  
67 culation (DeVries et al., 2019; Keppler & Landschützer, 2019). Higher frequency inter-  
68 annual variability is largely attributable to modes of climate variability such as ENSO  
69 on global scale and other modes of high latitude variability on regional scales (McKinley  
70 et al., 2017). Predicting future variations in the ocean carbon sink on inter-annual to decadal  
71 time scales in the face of these multiple drivers is therefore challenging.

72 Decadal predictions, such as those made under the Decadal Climate Prediction Project  
73 (DCCP) are produced by Global Climate Models (GCMs) that are that are initialized  
74 with observations and also driven by external forcing (Kirtman et al., 2013). Predictive  
75 skill of key physical climate variables from such simulations have been well established  
76 in the literature (Boer et al., 2016). However, near term predictability of the ocean car-  
77 bon flux and other biogeochemical variables have only become possible with the recent  
78 advent of Earth System Models (ESMs) (Meehl et al., 2021) and are still at their infancy.  
79 Previous studies have shown potential predictability of the ocean carbon flux for up to  
80 7 years (Li et al., 2019; Séférian et al., 2018) and actual skill versus observation based  
81 estimates for 2-6 years based on different ESMs (Li et al., 2019; Ilyina et al., 2021). How-  
82 ever, ESM simulations are subject to biases, drifts (Kharin et al., 2012) and exhibit a  
83 wide range of prediction skill globally and regionally (Ilyina et al., 2021). Predictions  
84 of ocean carbon flux using ESMs are especially challenging given that ocean biogeochem-  
85 ical variables are not directly initialized in current decadal prediction systems (Sospedra-  
86 Alfonso et al., 2021), and that the ocean biogeochemical models themselves are heavily  
87 parameterized using empirical parameterizations (Christian et al., 2022).

88 Here we propose using observationally-trained statistical models forced by predic-  
89 tors from GCM/ESM-based decadal predictions, as an alternative to using the raw pre-  
90 dictions of ocean carbon flux obtained from the ESMs ocean biogeochemistry models.  
91 It is well established that the surface ocean partial pressure of CO<sub>2</sub>, and by extension  
92 the surface carbon flux, is closely related to physical predictors, such as sea-surface tem-  
93 perature and salinity, atmospheric CO<sub>2</sub> concentration and wind speed. These empiri-  
94 cal relationships are widely exploited in the observational community to infill sparse di-  
95 rect observations of the ocean carbonate system (e.g., Surface Ocean CO<sub>2</sub> Atlas, SOCAT),  
96 using indirect but more widely sampled physical variables (Landschützer et al., 2016).  
97 It is also common to post-process raw GCM results to produce more skillful predictions,  
98 for example through bias correction (Kharin et al., 2012). Our proposal is a logical ex-  
99 tension of these two established practises that combines the explanatory power that sta-  
100 tistical models learn from the relationships between observational predictors, and the es-  
101 tablished prediction skill of the process based physical models. Our principal goal is to  
102 establish a methodology that allows us to improve near-term predictions of the ocean  
103 carbon sink over and above the skill obtained from raw ESM predictions.

104 We begin by introducing the methodology and our statistical models of choice in  
105 Section 2. In section 3 we evaluate observational uncertainties and the performance of  
106 our statistical models when forced by observation based predictors. In section 4, we ap-  
107 ply the observationally trained statistical models to physical predictors from CanESM5  
108 simulations, and evaluate the skill of this hybrid approach relative to the raw CanESM5  
109 predictions over the hindcast period of 1990 to 2019. We go on to provide forecasts for  
110 ocean carbon flux over the decade 2019 to 2029 in section 5. We conclude by reflecting  
111 on how our approach could be improved and expanded on in future work.

## 112 2 Materials and Methods

### 113 2.1 Surface CO<sub>2</sub> flux data

114 For observations of the atmosphere-ocean CO<sub>2</sub> flux we use the SeaFlux Ocean carbon  
 115 sink ensemble product (Gregor & Fay, 2021). SeaFlux contains an ensemble of flux  
 116 estimates, based on six global observation-based mapping products for surface ocean partial  
 117 pressure of CO<sub>2</sub> ( $pCO_2$ ), and wind speeds from ERA5. The six products include three  
 118 neural-network-derived products (CMEMS-FFNN, MPI-SOMFFN, NIES-FNN), a mixed  
 119 layer scheme product (JENA-MLS), a multiple linear regression (JMA-MLR), and a machine  
 120 learning ensemble (CSIR-ML6) (Fay et al., 2021). We also use the mean across the  
 121 products, which we refer to as SF-MEAN. Given the sparseness of actual  $pCO_2$  measurements,  
 122 using the ensemble of products allows us to quantify uncertainties associated with  
 123 the data infilling and mapping techniques, and avoids overfitting to a single product.

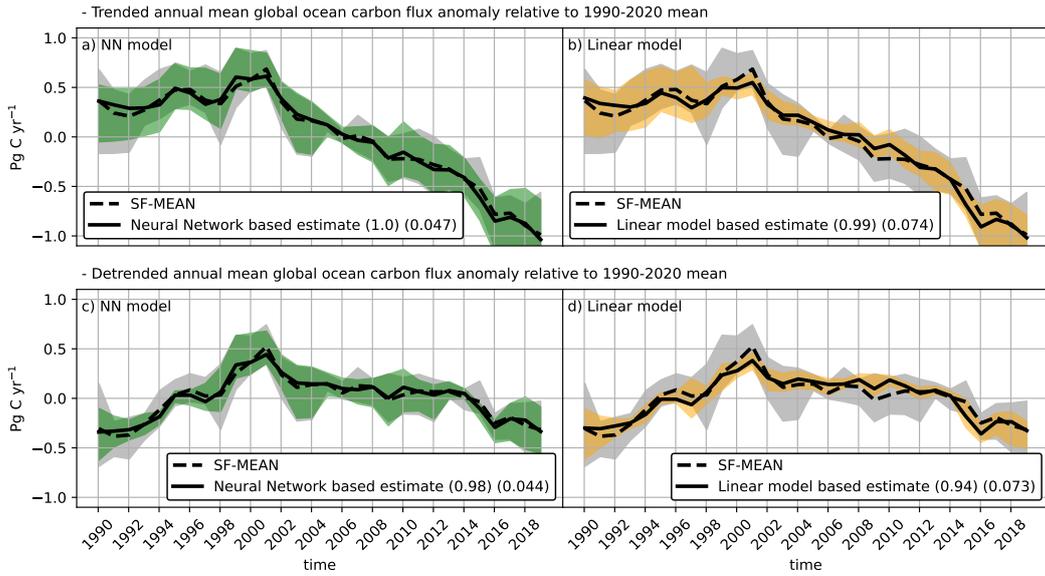
124 All six SeaFlux products show strong agreement in the long term (trended) changes  
 125 in ocean carbon flux (not shown here). Comparing linearly detrended versions of the SeaFlux  
 126 products shows cross correlation coefficients between them ranging from 0.47 to 0.95 (Fig.  
 127 S1). The MPI-SOM-FFN and JENA-MLS are least correlated with others. The lower  
 128 correlation skills for the two show that there are variabilities specific to these products  
 129 that are not common to other datasets, and known biases linked to data sparsity (Gloege  
 130 et al., 2021; Hauck et al., 2023). The averaged SF-MEAN contains signals common to  
 131 all of the products, and we use this as the most reliable estimate moving forward.

### 132 2.2 Statistical models and observed predictors

133 For each individual SeaFlux input dataset and SF-MEAN, we train a multi-linear  
 134 regression model and a neural network (NN) model to predict the surface atmosphere  
 135 ocean carbon flux, using three observation-based physical predictors - sea surface tem-  
 136 perature (SST), sea surface salinity (SSS), surface wind speed (sfcWind), one biological  
 137 predictor -surface chlorophyll concentrations (CHL), as well as atmospheric CO<sub>2</sub> con-  
 138 centrations ( $xCO_2$ ) (table S1). These are mainly physical predictors for which full cov-  
 139 erage observational products are available and are believed to drive the variability in ocean  
 140 carbon flux (Landschützer et al., 2016). Linear models are trained for each grid cell on  
 141 a standard one degree grid, while the NN models are trained over 16 biomes (Landschützer  
 142 et al., 2016), as explained further in SI (Sect. S1.1). By combining these biomes, we can  
 143 produce spatially resolved maps of the surface CO<sub>2</sub> flux, given the set of five input pre-  
 144 dictors at any point. In total that gives us 14 sets of models (7 set of linear models, and  
 145 7 NN models, one for each SeaFlux target predictand) that are later used to make hind-  
 146 casts and forecasts using modelled predictors from CanESM5. We have chosen to illus-  
 147 trate our approach using the linear and NN models, which have different structures and  
 148 levels of complexity, as illustrative examples. However, alternative models and predic-  
 149 tor variables could be used.

### 150 2.3 Decadal predictions using GCM base predictors

151 To make predictions the five predictors from Table S1 are obtained from CanESM5  
 152 simulations (Swart et al., 2019; Sospedra-Alfonso et al., 2021). We use a range of sim-  
 153 ulations, including standard free running CMIP6 historical simulations (Eyring et al.,  
 154 2016), as well as assimilation and hindcast and forecast runs (Boer et al., 2016). In as-  
 155 similation runs, CanESM5 is nudged towards observations for key physical variables (Sospedra-  
 156 Alfonso et al., 2021). For historical, hindcast and forecast simulations, the five predic-  
 157 tors are bias corrected to the same observational predictors used for training the mod-  
 158 els following the approach of (Kharin et al., 2012). This bias correction adjusts the mean  
 159 and trend of the predictors to be consistent with observations. These CanESM5 predic-  
 160 tors are fed to the each of the 14 statistical model sets mentioned above to produce hy-



**Figure 1.** Time series of the global ocean CO<sub>2</sub> flux anomalies for the (a) NN model (left panel) and (b) linear model (right panel) reconstruction using observational predictors. The black lines shows reconstruction using models that are trained on mean of SeaFlux products (SF-MEAN; solid) as well the mean product itself (dashed). The shadings represent the range estimates from the six different SeaFlux products (grey) and from NN and linear models reconstructions (green and orange). The numbers in the legends are correlation coefficients between the solid black lines and dashed black lines (first number) and root mean square error of the two time series (second number). (c) and (d) are same as (a) and (b) but are linearly detrended.

161 brid predictions of surface ocean CO<sub>2</sub> flux. For hindcasts and forecasts, predictions are  
 162 made for lead years 1 to 10. To test significance of prediction skill differences, we use a  
 163 1000 iteration bootstrap to test of (Goddard et al., 2013).

### 164 3 Evaluation of statistical models

165 In this section, we consider the performance of the statistical models trained on the  
 166 SeaFlux ensemble and using observed predictors, for predicting the global mean surface  
 167 carbon flux as defined by SF-MEAN (Fig. 1). When trained on SF-MEAN, both the NN  
 168 and linear models can accurately reconstruct the changes of the SF-MEAN ( $r > 0.9$ ),  
 169 indicating that the statistical models are able to capture the majority of the variance  
 170 in the global mean surface flux. The NN model shows higher skill in reconstructing SF-  
 171 MEAN relative to the linear model, reflected in higher correlations and lower root mean  
 172 square error (Fig. 1). Similarly, both linear and NN models are able to successfully re-  
 173 produce individual SeaFlux products on which they are trained (Fig. S2), with the NN  
 174 models again achieving tighter fits than the linear models. The orange and green shad-  
 175 ing in Fig. 1 represents the spread across models trained on individual SeaFlux prod-  
 176 ucts. These models are still able to successfully reproduce SF-MEAN, which gives an in-  
 177 dication of their generalizability. The smaller spread for the linear models (Fig. 1b, or-  
 178 ange shading), suggests they may be more generalizable (i.e. successful in predicting data  
 179 they were not trained on) than the NN models. We further explore the idea of gener-  
 180 alizability when using model-derived predictors in the following section.

## 4 Applying statistical models to physical predictors from the ESM

### 4.1 Assimilation run

The CanESM5 assimilation run is relaxed towards the observed physical state of the system, which forces physical variables, including our input predictors, to be close to observations. However, the detrended CO<sub>2</sub> flux from the CanESM5 biogeochemical component is not in good agreement with observations (Fig. 2 bottom row). We have identified issue in the model derived CO<sub>2</sub> flux, including seasonality that is out of phase with observations (not shown here), and it appears that the data ingestion in the assimilation run degrades the biogeochemical models performance. Indeed, previous results have shown that atmosphere-ocean CO<sub>2</sub> flux predictability is low in CanESM5, and particularly poor in the early lead years immediately following the assimilation run (Ilyina et al., 2021). A major goal of our effort is to see whether by replacing the CanESM5 biogeochemical model derived flux with one computed based on the statistical models leads to improvement.

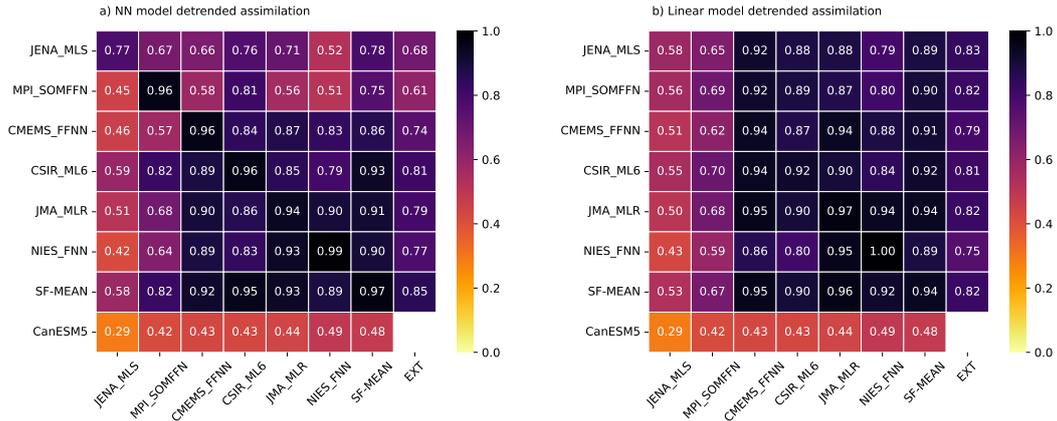
We use the linear and NN models previously trained using observed predictors, and for each of the six individual SeaFlux products and SF-MEAN as predictands (for a total of 14 model sets). We then extract the five input predictors from the (ensemble mean of 10) CanESM5 DCPD assimilation runs, apply the statistical models on these GCM-based predictors, and compare their skill against the original SeaFlux observational products (Fig 2).

The statistical models forced by CanESM5 assimilation predictors obtain similar skills in reproducing the individual SeaFlux products to the skills of the reconstructions that used predictors from observations (compare Fig. 2 and supplementary Fig. S2). This is a somewhat expected result given that assimilation runs assimilate physical predictors and are very close to the observations, but nonetheless it is first step in applying the models on data on which they were not directly trained. For both the linear and NN statistical models, the skill is in all cases is significantly higher than than skill of the raw CanESM5 CO<sub>2</sub> flux. These results indicate that statistical models trained on observations can usefully be applied to GCM-derived predictors. By using this approach we are able to avoid biases in the CanESM5 biogeochemical model by combining the observationally constrained statistical models with the directly initialized physical predictors from CanESM.

We compute the cross-correlation matrix for statistical models trained on one SeaFlux product in reproducing all the other five product and SF-MEAN (Fig. 2). This allows us to assess the impacts of observational uncertainty, and the potential consequences of overfitting statistical models to a single observational product. As expected, the statistical models are most skillful in reproducing the product on which they were trained (diagonal in Fig. 2). Correlation in reproducing other products can be lower than 0.5. The extent to which a model trained on one observational product can be generalized to others is measured with the mean of scores versus all other observational data products (mean of rows excluding the diagonal values as indicated in Fig. 2 EXT column). Overall, the linear models have larger extendibility scores, while the NN models produce better fits for the products on which they were trained. Our results illustrate that care should be taken in tightly fitting statistical models to a single observation based CO<sub>2</sub> flux product, as uncertainties exist. Moving forward, we will use statistical model trained on the SF-MEAN product as the best estimate. Based on the encouraging success so far, in the next section we will apply our approach to decadal predictions.

### 4.2 Prediction skill of CO<sub>2</sub> flux over the hindcast period

Hindcasts are ESM simulations that use the observationally constrained assimilation simulation as initial conditions, and which are then run freely under standard CMIP6

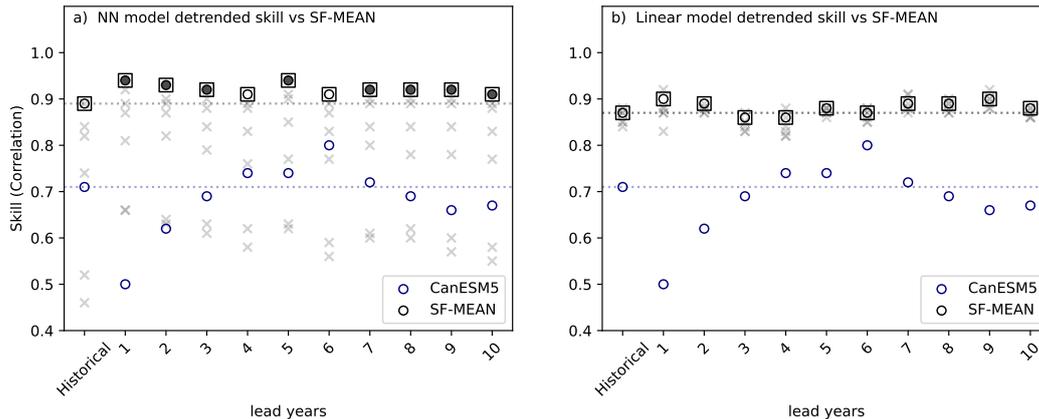


**Figure 2.** (a) Correlation matrix for the detrended global ocean carbon flux anomaly. The y axis indicate the product on which the NN model is trained and the x axis shows the data products against which the skill is evaluated. The EXT column measures the mean of skills excluding the diagonal element for each row. (b) Same as (a) but for the Linear model.

external forcings for ten years (Boer et al., 2016). Generally, as lead years increase (i.e. number of year since initialization) the hindcasts simulations lose memory of initialization and drift towards the preferred state of the model (historical simulations). However, raw CanESM5 ocean carbon flux DCPD scores show a decrease in the skill after initialization in hindcast compared to the historical free runs (Ilyina et al., 2021). This is not the expected result of initialization and indicates possible discrepancies with interactions between initialization and the CanESM5 biogeochemical decadal prediction system (initialization "shocks").

As an alternative to the biogeochemical model flux, we apply our SF-MEAN trained statistical models on predictors from the CanESM5 hindcast simulations over the period 1990 to 2019. The hindcast skill from both the linear and NN model when trained and evaluated against SF-MEAN are significantly larger than raw CanESM5 skills, with NN yielding slightly better scores (Fig. 3). Both statistical models show increase in skill after initialization, as expected and seen in physical predictors, and a gradual drop with lead time. As an even more stringent test, we compare the skill of the statistical models driven by CanESM5 predictors against the skill from all other available CMIP6 models that participated in DCPD. The NN model skill is higher than that shown by any raw CMIP6 model, when evaluated against SF-MEAN (Fig. S3) over 1990-2017 that is the period common to all models. Linear model score are higher than all CMIP6 models on all lead years except lead year 3 where CESM1 (Danabasoglu, 2019) yields slightly larger score (Fig. S3). These results clearly show the potential of our approach for improving the decadal CO<sub>2</sub> flux prediction skills.

To this point we have considered the absolute skill in predicting global mean surface CO<sub>2</sub> flux. An important concept in decadal prediction is the relative contribution to the absolute skill that is provided by the initialization. To assess whether initialization has added additional value to the predictions, the hindcast simulation skill can be compared to that found in standard, non-initialized CMIP6 historical simulations (Fig. 3). For the linear statistical models, hindcast skills are close to the corresponding historical skill, and do not show statistically significant improvement. That is, the linear model scores do not show significant added skill due to initialization. For the NN model, the hindcast skills are significantly larger than the historical skills at least for the first three years, based on a bootstrapping test (Fig. 3). This is the range where tempera-

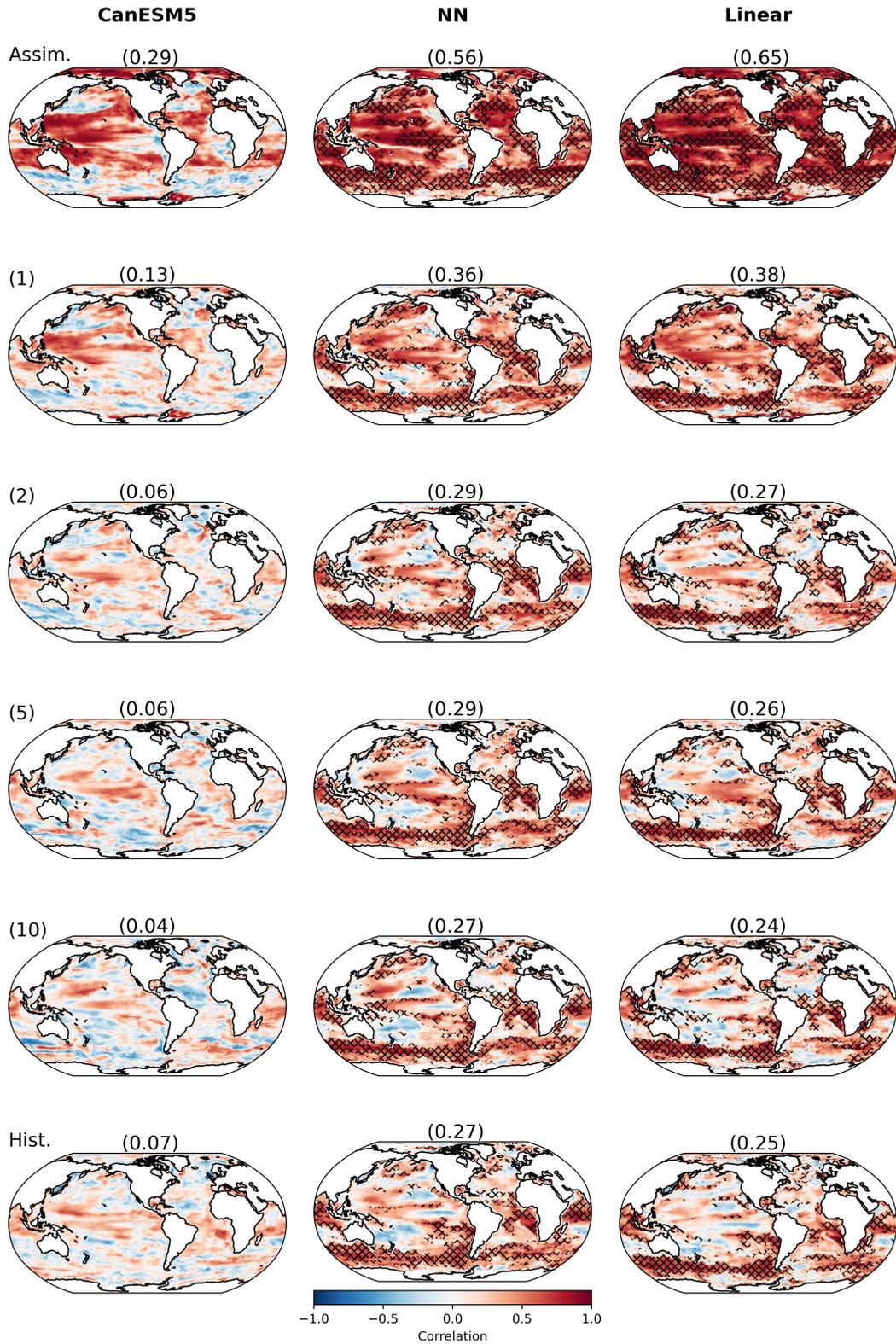


**Figure 3.** (a) Detrended global ocean carbon flux skills versus SF-MEAN for raw CanESM5 model (blue dots) and NN model trained on the SF-MEAN using bias corrected historical/hindcast predictors from CanESM5 (black dots). The scores that are statistically better than the raw CanESM5 score based on 1000 iteration bootstrap tests are shown with black boxes and the lead years where scores are significantly better than the corresponding historical score are filled. The grey marks in the background show scores from models trained on individual SeaFlux products versus the SF-MEAN. (b) Same as (a) but for the linear model.

263 ture variations largely control short term predictability of ocean carbon sink (Li et al.,  
 264 2019). The NN hindcast scores are not significantly better than historical for lead years  
 265 4 to 6, but show re-emergence of significance afterward. NN models consistently show  
 266 better fits to the dataset used for training them (Fig S2), but are also more subject to  
 267 overfitting than the linear models (Fig. 2). While more work is needed to understand  
 268 difference in model structure, our results show that initialization does add value to pre-  
 269 dictions made with the NN models (see also Fig. S4).

270 Both the hindcasts and historical run used observed atmospheric CO<sub>2</sub> concentra-  
 271 tions (as do our statistical models, as an input predictor). We expect that skills estimated  
 272 from the hindcast are higher than those achievable in true forecasts, because in true fore-  
 273 casts the atmospheric CO<sub>2</sub> concentration will not be known. It is not just the background  
 274 rate of increase that is relevant, but deviations in the growth rate of atmospheric CO<sub>2</sub>  
 275 are also known to be a key driver of decadal scale variability in the ocean CO<sub>2</sub> sink (McKinley  
 276 et al., 2020). This is an issue common to any DCP-style hindcast. Regardless, the im-  
 277 proved skill that the statistical models driven by CanESM5 based predictors show over  
 278 and above CanESM5 or other raw CMIP6 DCP model hindcast skills encourages us  
 279 to apply our methods to making future predictions in the following section. First how-  
 280 ever, we turn to considering the spatial pattern of skill over the hindcast period.

281 We compare spatially resolved temporal correlations between SF-MEAN, the CanESM5  
 282 raw biogeochemical model, and the two statistical models for the historical, assimilation  
 283 and lead years 1 to 10 of the hindcast experiments. Both the NN and linear models show  
 284 large correlations for the detrended flux over the majority of global ocean, when driven  
 285 by predictors from the CanESM5 assimilation run (Fig. 4). Compared to the raw flux  
 286 from the CanESM5 assimilation run, the statistical models significantly improve skill over  
 287 more than 55% of the global ocean (56% for NN and 65% for linear). The linear model  
 288 shows better average grid scale correlation compared to the NN model for assimilation  
 289 and lead year one hindcast. This is most likely due to the high grid scale training res-  
 290 olution of the linear model as opposed to biome scale resolution of the NN model (see



**Figure 4.** Grid wise correlation for the anomalous detrended ocean carbon flux versus SF-using assimilation, historical as well as lead years 1, 2, 5 10 predictors from CanESM5. The first column shows raw CanESM5 model skills, while the second and third columns show the NN and linear model based simulations. Hatches show regions where there is an statistically significant improvement in skill using a 1000 iteration bootstrap test compared to the raw CanESM5 results. The numbers on top of each panel are global mean of correlations.

291 supplements). Notably, the linear models has improved skill regionally, while the skill  
 292 of the globally integrated sink is better from the NN model. On longer hindcasts lead  
 293 yaers, the mean grid scale skill for the linear models drop faster than NN model and NN  
 294 model beats the linear model with small offsets and more percentage of grid cell (not shown  
 295 here) with significantly improved skills.

296 The regions that show significant improvements relative to raw CanESM5 model  
 297 include but are not limited to the highly active regions for the sink (Gooya et al., 2023)  
 298 which makes them important for both the flux magnitude and uncertainty. These are  
 299 regions where the largest sink is concentrated in smallest ocean surface area and where  
 300 internal and model uncertainty tend to be largest. Specifically, significant improvements  
 301 over the Southern Ocean is the common feature to all simulations. The Southern Ocean  
 302 is of key importance for ocean carbon sink (Gruber et al., 2019) where the models dis-  
 303 agree most (Gooya et al., 2023; Frölicher et al., 2015). In the hindcast simulations, skills  
 304 decrease with lead year, approaching the corresponding historical simulation skill on longer  
 305 lead times ( $>7$ ), as expected. For all lead years there is significant improvement beyond  
 306 the raw CanESM5 results regionally over more than 30% of the global ocean (hatched  
 307 areas in Fig. 4). Our results offer a potential pathway to better quantification of ocean  
 308 carbon sink predictions both regionally and globally.

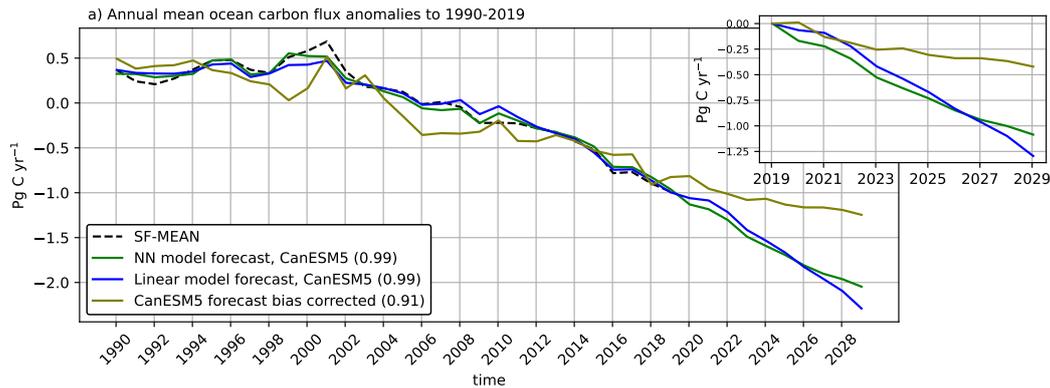
## 309 5 Hybrid forecast of the 2020-2029 ocean carbon sink

310 The ultimate purpose of decadal prediction systems is to provide forecasts of the  
 311 short term future evolution of the climate system, including the ocean carbon flux. In  
 312 this section, we use the statistical models trained on the SF-MEAN, and evaluated over  
 313 the hindcast period, to make predictions for the near term evolution of ocean carbon flux.  
 314 We extract ensemble means of our five predictors from CanESM5 DCPD forecasts for  
 315 the period 2019-2029, and bias correct them according to lead time following (Kharin  
 316 et al., 2012). We apply the statistical models on these predictors, and include the atmo-  
 317 spheric concentration of  $\text{CO}_2$  from SSP245 (Eyring et al., 2016), which is the same pro-  
 318 cedure applied to the hindcasts in the previous section.

319 Both NN model and linear model based forecasts predict that ocean carbon sink  
 320 is going to grow with a faster than linear rate over the next decade under the SSP245  
 321 scenario (Fig. 5). The linear model predicts slower rate of increase until 2022 compared  
 322 to the NN model, and an accelerated increase after to nearly  $1.29 \text{ pgC yr}^{-1}$  relative to  
 323 2019 by 2029. The rate of change in the linear model is consistent with the rate of change  
 324 of the atmospheric  $\text{CO}_2$  concentrations under the SSP245 scenario. The NN model pre-  
 325 dicta a more steady yet faster than linear increase of approximately  $1.09 \text{ pgC yr}^{-1}$  in global  
 326 ocean carbon sink relative to 2019. Both models are in close agreement regarding decadal  
 327 scale changes in the flux and predict larger changes compared to the bias corrected flux  
 328 from the CanESM5 biogeochemical component. The fact that the results show are largely  
 329 consistent between the two statistical models over 1990-2019 as well as the future fore-  
 330 cast globally and regionally (Fig. S5), increases our confidence in the results. Based on  
 331 the skill demonstrated in the hindcasts, we assert that our hybrid statistical-GCM pre-  
 332 dictions represent a more reliable estimate of future changes in the ocean carbon flux than  
 333 the raw model predictions.

## 334 6 Discussion and conclusions

335 We have proposed a methodology to improve the decadal predictability of the ocean  
 336 carbon flux by using statistical models as alternatives to the ocean biogeochemistry com-  
 337 ponents of decadal prediction systems. Through their training, the statistical models en-  
 338 code the relationships between physical predictors and the surface carbon flux found in  
 339 observations. Predictions are made by applying these observationally trained statisti-  
 340 cal models on (largely) physical predictors obtained from the GCM-based decadal pre-



**Figure 5.** Global ocean carbon flux decadal forecast based on bias corrected CanESM5 (olive), NN model (green), and linear model (blue) trained on SF-MEAN. The dashed black line shows SF-MEAN over the period of 1990-2019. The Forecasts show assimilation runs over this period and forecast initialized in 2019 after. The subplot shows anomalies relative to the 2019 ocean carbon flux on each product and shows the predicted changes until 2029 from different estimates. All global timeseries are scaled based on the spatial coverage of the SF-MEAN to account for differences in coverage.

341 diction systems. Unlike biogeochemical variables, the physical variables are directly ini-  
 342 tialized in current prediction systems, have a more established track record of skill, and  
 343 are based on less heavily parameterized processes than ocean biogeochemistry. In prin-  
 344 cipal, our approach can be thought of as an extension of traditional bias correction (Kharin  
 345 et al., 2012). Statistical bias correction schemes using linear/NN algorithms have pre-  
 346 viously been used for physical parameters in decadal prediction system (citation). Un-  
 347 like those, our approach does not use the same variable that is being bias corrected. In-  
 348 stead, it relies primarily on key physical predictors whose predictability have been well  
 349 evaluated.

350 We have demonstrated that in hindcasts, our hybrid statistical-GCM system im-  
 351 proves prediction skill for the surface ocean carbon flux relative to the ocean biogeochem-  
 352 ical model, both in the global mean, and regionally over broad areas of the ocean. In-  
 353 deed, for the global mean flux, our hybrid skills based on CanESM5 predictors beat all  
 354 available CMIP6 DCP6 models. Globally, the NN model can retain the memory of ini-  
 355 tialization of the predictors at least up to lead year three after initialization.

356 We have demonstrated our approach using two examples of observationally con-  
 357 strained statistical models of different complexities; a linear and a neural network model.  
 358 The two statistical models used here have different structures and use different combi-  
 359 nations of predictors. Both statistical models are able to reconstruct observed  $\text{CO}_2$  fluxes  
 360 when forced by observed predictors, and both perform well in hindcast evaluations driven  
 361 by CanESM5-based predictors (i.e. beating the skill of the raw CanESM5 flux). In gen-  
 362 eral, the NN model was able to achieve higher correlations when trained and evaluated  
 363 against a given surface flux product, but the linear model showed more "generalizabil-  
 364 ity" across products. In addition, while the linear model was quite robust to changes in  
 365 structure (predictors), the NN model was quite sensitive to changes in the number of pre-  
 366 dictors or neurons used. This shows the need for carefully adjusting such complex mod-  
 367 els and validation against other such models to avoid possible overfitting and to make  
 368 reliable estimates.

369 We emphasize that the two statistical models we have used are just examples of  
 370 our more general approach of applying observationally trained statistical models to GCM  
 371 predictors. Our method is not limited to the choice of ESM, observation based product,  
 372 or to the choice of the alternative model. Future work should test the ability of differ-  
 373 ent types of statistical models to improve upon our results, and could draw upon the large  
 374 body of work in developing empirical relationships for the purposes of infilling sparse pCO<sub>2</sub>  
 375 observations (Fay et al., 2021). Currently, CanESM5 is the only model with sufficient  
 376 number of simulations publicly available for 10-year hindcasts and forecast for all of the  
 377 required predictors. More robust estimates of the future changes of ocean carbon sink  
 378 would be possible with multimodel averages of predictors, since such multi-model pre-  
 379 dictions are generally more skillful (Tebaldi & Knutti, 2007). We also note that our ap-  
 380 proach is not limited to surface ocean carbon flux, but could also be applied to other bio-  
 381 geochemical predictors, or even less certain physical variables that could benefit from  
 382 exploiting empirical relationships based on well predicted quantities such as SST.

383 Based on the demonstrated skill of our hybrid approach in hindcasts, we have made  
 384 forecasts of the near term evolution of ocean carbon flux using both the linear and NN  
 385 models under ssp245 scenario. Both hybrid statistical models show consistent changes  
 386 over the period of 2019-2029 with faster than linear increase in the sink that are larger  
 387 than bias corrected CanESM5 forecasts. This information about predicted future changes  
 388 in the ocean carbon sink might be useful to climate science and policy effort, for exam-  
 389 ple the assessment of the global carbon budget (Friedlingstein et al., 2022). Moving for-  
 390 ward we encourage further research into improving decadal predictions by optimally ex-  
 391 ploiting all available observational information, and data science techniques, in conjunc-  
 392 tion with traditional GCM based predictions.

## 393 Open Research

394 The SeaFlux observation based ensemble is available publicly at [https://zenodo](https://zenodo.org/record/5482547)  
 395 [.org/record/5482547](https://zenodo.org/record/5482547). All model data used in this study are part of the World Climate  
 396 Research Programme’s (WCRP) 6th Coupled Model Intercomparison Project (CMIP6)  
 397 and open-access through Earth System Grid Federation (ESGF) repositories. Observa-  
 398 tional predictors used for training the statistical models are available through institu-  
 399 tional public repositories as cited in the Supplements. All other inquiries should be di-  
 400 rected to P. Gooya.

## 401 Acknowledgement

402 We thank CCCma seasonal to decadal (S2D) prediction team for their helpful in-  
 403 sights on this project. Specially, we thank Reinel Sospedra-Alfonso for his comments and  
 404 helpful suggestions during the development of this study and on a draft of the paper.

## 405 References

- 406 Boer, G. J., Smith, D. M., Cassou, C., Doblas-Reyes, F., Danabasoglu, G., Kirt-  
 407 man, B., . . . Eade, R. (2016, October). The Decadal Climate Predic-  
 408 tion Project (DCPP) contribution to CMIP6. *Geoscientific Model De-*  
 409 *velopment*, 9(10), 3751–3777. Retrieved 2023-04-19, from [https://](https://gmd.copernicus.org/articles/9/3751/2016/)  
 410 [gmd.copernicus.org/articles/9/3751/2016/](https://gmd.copernicus.org/articles/9/3751/2016/) (Publisher: Copernicus  
 411 GmbH) doi: 10.5194/gmd-9-3751-2016
- 412 Christian, J. R., Denman, K. L., Hayashida, H., Holdsworth, A. M., Lee, W. G.,  
 413 Riche, O. G. J., . . . Swart, N. C. (2022). Ocean biogeochemistry in  
 414 the canadian earth system model version 5.0.3: Canesm5 and canesm5-  
 415 canoe. *Geoscientific Model Development*, 15(11), 4393–4424. Retrieved  
 416 from <https://gmd.copernicus.org/articles/15/4393/2022/> doi:

- 417 10.5194/gmd-15-4393-2022
- 418 Danabasoglu, G. (2019). *NCAR CESM1-1-CAM5-CMIP5 model output prepared for*
- 419 *CMIP6 DCP6*. Earth System Grid Federation. doi: {10.22033/ESGF/CMIP6
- 420 .11542}
- 421 DeVries, T., Quéré, C. L., Andrews, O., Berthet, S., Hauck, J., Ilyina, T., ...
- 422 Séférian, R. (2019). Decadal trends in the ocean carbon sink. *Proceed-*
- 423 *ings of the National Academy of Sciences*, 116(24), 11646-11651. Retrieved
- 424 from <https://www.pnas.org/doi/abs/10.1073/pnas.1900371116> doi:
- 425 10.1073/pnas.1900371116
- 426 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., &
- 427 Taylor, K. E. (2016, May). Overview of the Coupled Model Intercomparison
- 428 Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific*
- 429 *Model Development*, 9(5), 1937-1958. Retrieved 2022-05-25, from [https://](https://gmd.copernicus.org/articles/9/1937/2016/gmd-9-1937-2016.html)
- 430 [gmd.copernicus.org/articles/9/1937/2016/gmd-9-1937-2016.html](https://gmd.copernicus.org/articles/9/1937/2016/gmd-9-1937-2016.html) (Pub-
- 431 lisher: Copernicus GmbH) doi: 10.5194/gmd-9-1937-2016
- 432 Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M.,
- 433 ... Zeng, J. (2021). Seaflux: harmonization of air-sea CO<sub>2</sub> fluxes from sur-
- 434 face pCO<sub>2</sub> data products using a standardized approach. *Earth System Science*
- 435 *Data*, 13(10), 4693-4710. Retrieved from [https://essd.copernicus.org/](https://essd.copernicus.org/articles/13/4693/2021/)
- 436 [articles/13/4693/2021/](https://essd.copernicus.org/articles/13/4693/2021/) doi: 10.5194/essd-13-4693-2021
- 437 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck,
- 438 J., ... Zheng, B. (2022, November). Global Carbon Budget 2022. *Earth*
- 439 *System Science Data*, 14(11), 4811-4900. Retrieved 2023-04-19, from
- 440 <https://essd.copernicus.org/articles/14/4811/2022/> (Publisher:
- 441 Copernicus GmbH) doi: 10.5194/essd-14-4811-2022
- 442 Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen,
- 443 A., ... Zaehle, S. (2020). Global carbon budget 2020. *Earth System Science*
- 444 *Data*, 12(4), 3269-3340. Retrieved from [https://essd.copernicus.org/](https://essd.copernicus.org/articles/12/3269/2020/)
- 445 [articles/12/3269/2020/](https://essd.copernicus.org/articles/12/3269/2020/) doi: 10.5194/essd-12-3269-2020
- 446 Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P.,
- 447 & Winton, M. (2015, January). Dominance of the Southern Ocean
- 448 in Anthropogenic Carbon and Heat Uptake in CMIP5 Models. *Jour-*
- 449 *nal of Climate*, 28(2), 862-886. Retrieved 2022-12-06, from [https://](https://journals.ametsoc.org/view/journals/clim/28/2/jcli-d-14-00117.1.xml)
- 450 [journals.ametsoc.org/view/journals/clim/28/2/jcli-d-14-00117.1.xml](https://journals.ametsoc.org/view/journals/clim/28/2/jcli-d-14-00117.1.xml)
- 451 (Publisher: American Meteorological Society Section: Journal of Climate) doi:
- 452 10.1175/JCLI-D-14-00117.1
- 453 Gloege, L., McKinley, G. A., Landschützer, P., Fay, A. R., Frölicher, T. L., Fyfe,
- 454 J. C., ... Takano, Y. (2021). Quantifying Errors in Observationally
- 455 Based Estimates of Ocean Carbon Sink Variability. *Global Biogeochem-*
- 456 *ical Cycles*, 35(4), e2020GB006788. Retrieved from [https://agupubs](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GB006788)
- 457 [.onlinelibrary.wiley.com/doi/abs/10.1029/2020GB006788](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GB006788) (eprint:
- 458 <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020GB006788>) doi:
- 459 <https://doi.org/10.1029/2020GB006788>
- 460 Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., ... Del-
- 461 worth, T. (2013). A verification framework for interannual-to-decadal
- 462 predictions experiments. *Climate Dynamics*, 40(1), 245-272. Retrieved
- 463 from <https://doi.org/10.1007/s00382-012-1481-2> doi: 10.1007/
- 464 s00382-012-1481-2
- 465 Gooya, P., Swart, N. C., & Hamme, R. C. (2023). Time-varying changes and un-
- 466 certainties in the cmip6 ocean carbon sink from global to local scale. *Earth*
- 467 *System Dynamics*, 14(2), 383-398. Retrieved from [https://esd.copernicus](https://esd.copernicus.org/articles/14/383/2023/)
- 468 [.org/articles/14/383/2023/](https://esd.copernicus.org/articles/14/383/2023/) doi: 10.5194/esd-14-383-2023
- 469 Gregor, L., & Fay, A. (2021, July). *SeaFlux: harmonised sea-air CO<sub>2</sub> fluxes from*
- 470 *surface pCO<sub>2</sub> data products using a standardised approach*. Zenodo. Retrieved
- 471 from <https://doi.org/10.5281/zenodo.5482547> (version 2021.04.03: in-

- 472 cludes globally integrated fluxes including all wind and pco2 products. version  
 473 2021.04.02: removed some files that remained from v2021.02 that I didn't  
 474 delete in v2021.04 version 2021.04 now extends the variables to calculate fluxes  
 475 from 1982-2020. The comparison period for fluxes is now limited to 1990-2019  
 476 (30 years). The area contains coastal fraction coverage. A missing strip along  
 477 the longitude 179.5°E is filled in. Negative values of pCO<sub>2</sub> are limited to 50  
 478 µatm (primarily affects JENA in Hudson Bay). version 2021.02 calibrates kw  
 479 with 14-C bomb estimated global average kw ( $16.5 \pm 3.2$  cm/hr) where the  
 480 average is calculated without ice weighting (v2021.01 included ice weighting).  
 481 Further, the date range of the data has been increased from 1982 to 2020  
 482 where possible (not possible for the scaling factor and pCO<sub>2</sub> product.) doi:  
 483 10.5281/zenodo.5482547
- 484 Gruber, N., Clement, D., Carter, B. R., Feely, R. A., Heuven, S. v., Hoppema,  
 485 M., ... Wanninkhof, R. (2019). The oceanic sink for anthropogenic  
 486 CO<sub>2</sub> from 1994 to 2007. *Science*, *363*(6432), 1193–1199. Retrieved  
 487 from <https://www.science.org/doi/abs/10.1126/science.aau5153>  
 488 (\_eprint: <https://www.science.org/doi/pdf/10.1126/science.aau5153>) doi:  
 489 10.1126/science.aau5153
- 490 Hauck, J., Nissen, C., Landschützer, P., Rödenbeck, C., Bushinsky, S., & Olsen,  
 491 A. (2023). Sparse observations induce large biases in estimates of the  
 492 global ocean carbon sink: an ocean model subsampling experi-  
 493 ment. *Philosophical Transactions of the Royal Society A: Mathematical,*  
 494 *Physical and Engineering Sciences*, *381*(2249), 20220063. Retrieved from  
 495 <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2022.0063>  
 496 doi: 10.1098/rsta.2022.0063
- 497 Hauck, J., Zeising, M., Le Quéré, C., Gruber, N., Bakker, D. C. E., Bopp, L., ...  
 498 Séférian, R. (2020). Consistency and Challenges in the Ocean Carbon Sink  
 499 Estimate for the Global Carbon Budget. *Frontiers in Marine Science*, *7*. Re-  
 500 trieved 2022-12-06, from <https://www.frontiersin.org/articles/10.3389/fmars.2020.571720>
- 501 Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., ...  
 502 Yeager, S. (2021). Predictable Variations of the Carbon Sinks and  
 503 Atmospheric CO<sub>2</sub> Growth in a Multi-Model Framework. *Geophysical*  
 504 *Research Letters*, *48*(6), e2020GL090695. Retrieved 2023-04-18, from  
 505 <https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL090695>  
 506 (\_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL090695>)  
 507 doi: 10.1029/2020GL090695
- 508
- 509 Keppler, L., & Landschützer, P. (2019). Regional wind variability modulates the  
 510 southern ocean carbon sink. *Sci Rep*, *9*, 7384. doi: 10.1038/s41598-019-43826-  
 511 y
- 512 Kharin, V. V., Boer, G. J., Merryfield, W. J., Scinocca, J. F., & Lee, W.-S.  
 513 (2012). Statistical adjustment of decadal predictions in a changing cli-  
 514 mate. *Geophysical Research Letters*, *39*(19). Retrieved from [https://](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052647)  
 515 [agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052647](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052647) doi:  
 516 <https://doi.org/10.1029/2012GL052647>
- 517 Kirtman, B., B.S, P., A.J, A., J., B., Bojariu, R., Camilloni, I., ... Hawkins, E.  
 518 (2013, 01). Near-term climate change: Projections and predictability..
- 519 Landschützer, P., Gruber, N., & Bakker, D. C. E. (2016). Decadal varia-  
 520 tions and trends of the global ocean carbon sink. *Global Biogeochem-*  
 521 *ical Cycles*, *30*(10), 1396–1417. Retrieved from [https://agupubs](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GB005359)  
 522 [.onlinelibrary.wiley.com/doi/abs/10.1002/2015GB005359](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015GB005359) (\_eprint:  
 523 <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2015GB005359>) doi:  
 524 <https://doi.org/10.1002/2015GB005359>
- 525 Li, H., Ilyina, T., Müller, W. A., & Landschützer, P. (2019). Predicting the  
 526 variable ocean carbon sink. *Science Advances*, *5*(4), eaav6471. Retrieved

- 527 from <https://www.science.org/doi/abs/10.1126/sciadv.aav6471>  
 528 (\_eprint: <https://www.science.org/doi/pdf/10.1126/sciadv.aav6471>) doi:  
 529 10.1126/sciadv.aav6471
- 530 Lovenduski, N. S., McKinley, G. A., Fay, A. R., Lindsay, K., & Long, M. C.  
 531 (2016). Partitioning uncertainty in ocean carbon uptake projections: In-  
 532 ternal variability, emission scenario, and model structure. *Global Bio-*  
 533 *geochemical Cycles*, 30(9), 1276-1287. Retrieved from <https://agupubs>  
 534 [.onlinelibrary.wiley.com/doi/abs/10.1002/2016GB005426](https://doi.org/10.1002/2016GB005426) doi:  
 535 <https://doi.org/10.1002/2016GB005426>
- 536 McKinley, G. A., Fay, A. R., Eddebar, Y. A., Gloege, L., & Lovenduski, N. S.  
 537 (2020). External Forcing Explains Recent Decadal Variability of the Ocean  
 538 Carbon Sink. *AGU Advances*, 1(2), e2019AV000149. Retrieved 2023-04-18,  
 539 from <https://onlinelibrary.wiley.com/doi/abs/10.1029/2019AV000149>  
 540 (\_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019AV000149>) doi:  
 541 10.1029/2019AV000149
- 542 McKinley, G. A., Fay, A. R., Lovenduski, N. S., & Pilcher, D. J. (2017). Natural  
 543 variability and anthropogenic trends in the ocean carbon sink. *Annual Review*  
 544 *of Marine Science*, 9(1), 125-150. Retrieved from [https://doi.org/10.1146/](https://doi.org/10.1146/annurev-marine-010816-060529)  
 545 [annurev-marine-010816-060529](https://doi.org/10.1146/annurev-marine-010816-060529) (PMID: 27620831) doi: 10.1146/annurev-  
 546 -marine-010816-060529
- 547 McKinley, G. A., Pilcher, D. J., Fay, A. R., Lindsay, K., Long, M. C., & Lovenduski,  
 548 N. S. (2016, Feb 01). Timescales for detection of trends in the ocean carbon  
 549 sink. *Nature*, 530(7591), 469-472. Retrieved from [https://doi.org/10.1038/](https://doi.org/10.1038/nature16958)  
 550 [nature16958](https://doi.org/10.1038/nature16958) doi: 10.1038/nature16958
- 551 Meehl, G. A., Richter, J. H., Teng, H., Capotondi, A., Cobb, K., Doblas-Reyes, F.,  
 552 ... Xie, S.-P. (2021, May). Initialized Earth System prediction from sub-  
 553 seasonal to decadal timescales. *Nature Reviews Earth & Environment*, 2(5),  
 554 340-357. Retrieved from <https://doi.org/10.1038/s43017-021-00155-x>  
 555 doi: 10.1038/s43017-021-00155-x
- 556 Rodgers, K. B., Schlunegger, S., Slater, R. D., Ishii, M., Frölicher, T. L., Toyama,  
 557 K., ... Fassbender, A. J. (2020). Reemergence of anthropogenic carbon into  
 558 the ocean's mixed layer strongly amplifies transient climate sensitivity. *Geo-*  
 559 *physical Research Letters*, 47(18), e2020GL089275. Retrieved from [https://](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL089275)  
 560 [agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020GL089275](https://doi.org/10.1029/2020GL089275)  
 561 (e2020GL089275 2020GL089275) doi: <https://doi.org/10.1029/2020GL089275>
- 562 Sospedra-Alfonso, R., Merryfield, W. J., Boer, G. J., Kharin, V. V., Lee, W.-S.,  
 563 Seiler, C., & Christian, J. R. (2021). Decadal climate predictions with the  
 564 Canadian Earth System Model version 5 (CanESM5). *Geoscientific Model De-*  
 565 *velopment*, 14(11), 6863-6891. Retrieved from [https://gmd.copernicus.org/](https://gmd.copernicus.org/articles/14/6863/2021/)  
 566 [articles/14/6863/2021/](https://doi.org/10.5194/gmd-14-6863-2021) doi: 10.5194/gmd-14-6863-2021
- 567 Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett,  
 568 N. P., ... Winter, B. (2019, November). The Canadian Earth System  
 569 Model version 5 (CanESM5.0.3). *Geoscientific Model Development*, 12(11),  
 570 4823-4873. Retrieved 2022-05-25, from [https://gmd.copernicus.org/](https://gmd.copernicus.org/articles/12/4823/2019/)  
 571 [articles/12/4823/2019/](https://doi.org/10.5194/gmd-12-4823-2019) (Publisher: Copernicus GmbH) doi: 10.5194/  
 572 gmd-12-4823-2019
- 573 Séférian, R., Berthet, S., & Chevallier, M. (2018). Assessing the decadal pre-  
 574 dictability of land and ocean carbon uptake. *Geophysical Research Let-*  
 575 *ters*, 45(5), 2455-2466. Retrieved from [https://agupubs.onlinelibrary](https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL076092)  
 576 [.wiley.com/doi/abs/10.1002/2017GL076092](https://doi.org/10.1002/2017GL076092) doi: [https://doi.org/10.1002/](https://doi.org/10.1002/2017GL076092)  
 577 2017GL076092
- 578 Tebaldi, C., & Knutti, R. (2007). The use of the multi-model ensemble in proba-  
 579 bilistic climate projections. *Philosophical Transactions of the Royal Society*  
 580 *A: Mathematical, Physical and Engineering Sciences*, 365(1857), 2053-2075.  
 581 Retrieved from <https://royalsocietypublishing.org/doi/abs/10.1098/>

582        `rsta.2007.2076` doi: 10.1098/rsta.2007.2076  
583        Wang, L., Huang, J., Luo, Y., & Zhao, Z. (2016, Nov 28). Narrowing the spread  
584        in cmip5 model projections of air-sea co2 fluxes. *Scientific Reports*, 6(1),  
585        37548. Retrieved from <https://doi.org/10.1038/srep37548> doi:  
586        10.1038/srep37548