

OCEANOGRAPHY

Improved air-sea CO₂ flux estimates from sailboat measurementsJacqueline Behncke^{1,2*}, Tatiana Ilyina^{1,3,4}, Fatemeh Chegini³, Peter Landschützer^{1,5}

Despite their importance in the climate system, remote ocean regions and their ability to absorb anthropogenic carbon dioxide (CO₂) remain highly uncertain. To address this issue, citizen science initiatives, including sailboats, expand the observational network. Using observing system simulations and novel sailboat tracks, we demonstrate how integrating sailboat data improves estimates of ocean carbon uptake. While we underestimate the ocean carbon sink when mimicking real-world sampling, adding available sailboat data does not substantially improve reconstructions. Nevertheless, increased sampling reveals a stronger carbon sink, particularly between 40°S and 60°S. The improvement persists with hypothetical measurement uncertainties, but substantial differences arise depending on whether positive or negative biases are applied to the race track data. While we show that two additional circumnavigations already improve the ocean mean sink estimate, we further highlight the finding that the additional data remain insufficient to correct the overestimated CO₂ sink trend, calling for continuation of the ongoing data collection.

INTRODUCTION

The ocean plays a critical role in regulating Earth's climate by acting as a substantial carbon sink that annually absorbs over a quarter of anthropogenically emitted carbon dioxide (CO₂) from the atmosphere (1, 2). However, climate change alters the carbon uptake capacity of the ocean (1, 3), making the monitoring of the sea surface partial pressure of CO₂ (pCO₂) or the fugacity of CO₂ (fCO₂) indispensable in order to assess the impacts of climate change. Despite this need, a declining trend in observations in recent years and an imbalance in sampling efforts across the hemispheres persist (4, 5). While the Ship of Opportunity (SOOP) program has successfully contributed to systematic sampling in the Northern Hemisphere (6), the Southern Ocean remains more irregularly sampled (4, 5, 7, 8), leading to biases in the reconstruction of the air-sea CO₂ flux (9–12). Specifically, insufficient sampling leads to a 31% overestimation of the Southern Ocean decadal variability (9) and a strong overestimation of decadal trends (10). Considering the key role of remote ocean basins such as the Southern Ocean in ocean carbon and heat uptake (13–15), and the moderate success of gap-filling methods in further improving fCO₂ estimates (9, 10, 16), previously unexplored observational efforts have been undertaken to expand the observational network.

Sailboat races are emerging as a means to address observational gaps by providing a unique source of oceanographic data, particularly in remote ocean regions, as part of citizen science efforts (17, 18). Despite their predominant occurrence in the North Atlantic, regular repeating circumnavigation races (Vendée Globe, every 4 years; The Ocean Race, every 3 to 4 years) also provide valuable data in the Southern Ocean—both regions prone to uncertainty in the air-sea CO₂ flux (19, 20) and drivers on multiple timescales (21). Hence, sailboats complement the observational network consisting so far of research ships, voluntary observing ships, drifting and moored

buoys [e.g., (22)], gliders [e.g., (23–26)], biogeochemical floats [e.g., (27–29)], and Saildrones (30–32) and have been shown to significantly affect the air-sea CO₂ flux estimate, particularly between 40°S and 60°S during austral summer (17). Previous research (17) has shown that even when considering potential measurement uncertainties within the range of $\pm 5 \mu\text{atm}$ (1 atm = 101.325 kPa), the impact of new sailboat measurements on the air-sea CO₂ flux is still detectable, unlike when considering a systematic measurement bias (17). What we still lack, however, is information on whether sailboat measurements actually improve the air-sea CO₂ flux estimates, which has yet to be quantified. Here, we address this challenge.

Recent Observing System Simulation Experiments investigate the impact of different sampling strategies on carbon fluxes by using a model testbed as a benchmark to quantify the improvement (10, 27, 30, 33–36). These investigations typically involve subsampling an fCO₂ model testbed based on synthetic sampling schemes and reconstructing the air-sea CO₂ flux with neural network gap-filling methods [e.g., (37)] to compare against the model truth. However, previous studies often focus on optimal, yet not operational, sampling scenarios rather than feasible real-world sampling schemes such as our repeating circumnavigations. Thus, although studies show promising improvements linked to optimized sampling (10, 38), the implementation is often not feasible. The frequency of sailboat races allows us to quantify the extent to which realistic sampling by sailboats, particularly in the Southern Ocean, improves the estimate of the air-sea CO₂ flux, providing a path forward in improved monitoring of the Southern Ocean CO₂ uptake.

Here, we conduct an observation system simulation by subsampling the global ocean biogeochemical model HAMBURG Ocean Carbon Cycle (HAMOCC) coupled to the ocean general circulation model Max Planck Institute Ocean Model (MPIOM) that contributes to the Global Carbon Budget (1, 39–41) mimicking the present-day observations [www.socat.info; (4)] of the sea surface fCO₂ and apply the two-step neural network method self-organizing map-feed-forward neural network (SOM-FFN) (37, 42) to reconstruct different air-sea CO₂ flux estimates. We quantify potential improvements from our existing 161 days of sailboat data from several sailboats on the air-sea CO₂ flux estimate as well as the potential effect of more data by subsampling data from previous years using realistic

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works. Distributed under a Creative Commons Attribution License 4.0 (CC BY).

¹Max Planck Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany.

²International Max Planck Research School on Earth System Modelling, Bundesstrasse 53, 20146 Hamburg, Germany. ³Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany. ⁴Helmholtz-Zentrum Hereon, Max-Planck-Straße 1, 21502 Geesthacht, Germany. ⁵Flanders Marine Institute (VLIZ), Jacobsenstraat 1, 8400 Ostend, Belgium.

*Corresponding author. Email: jacqueline.behncke@mpimet.mpg.de

sailboat tracks (Fig. 1). Our analysis further explores the effect of random measurement uncertainties and systematic biases associated with the sailboat data. We investigate whether reconstructions continue to improve when additional, albeit biased, data are added, to assess whether the increase in quantity compensates for the lack of quality. Here, we show that continuous and long-term observing improves the mean air-sea CO_2 flux estimate; however, our analysis reveals that the reconstructed trend from the neural network method remains overestimated even after adding “3 circumnavigation” races, highlighting the need for multidecadal observing strategies.

RESULTS

In this study, we show how incorporating different synthetic sailboat data improves the reconstructions of a known-truth model field in

an observing system simulation experiment. We take the full hind-cast model fCO_2 field from MPIOM-HAMOCC and subsample it at times and locations where (based on a 1° -by- 1° grid location) historical and synthetic observations exist. Figure 1 presents the different pseudo-observation tracks used in this study to subsample the model field.

Using the SOM-FFN method, i.e., a two-step neural network method [see (37) for details], we then gap-fill these subsampled data to reconstruct the complete model fCO_2 field. Figure 1A presents the reconstructed global fCO_2 based on subsampled model data mimicking the real-world observations from Surface Ocean CO_2 Atlas (SOCAT) (“existing sailboat”). After deriving the air-sea CO_2 flux estimates from the subsampled and gap-filled fCO_2 field, our analysis proceeds in three steps: (i) assessing the neural network’s performance with present-day sampling against the model truth; (ii)

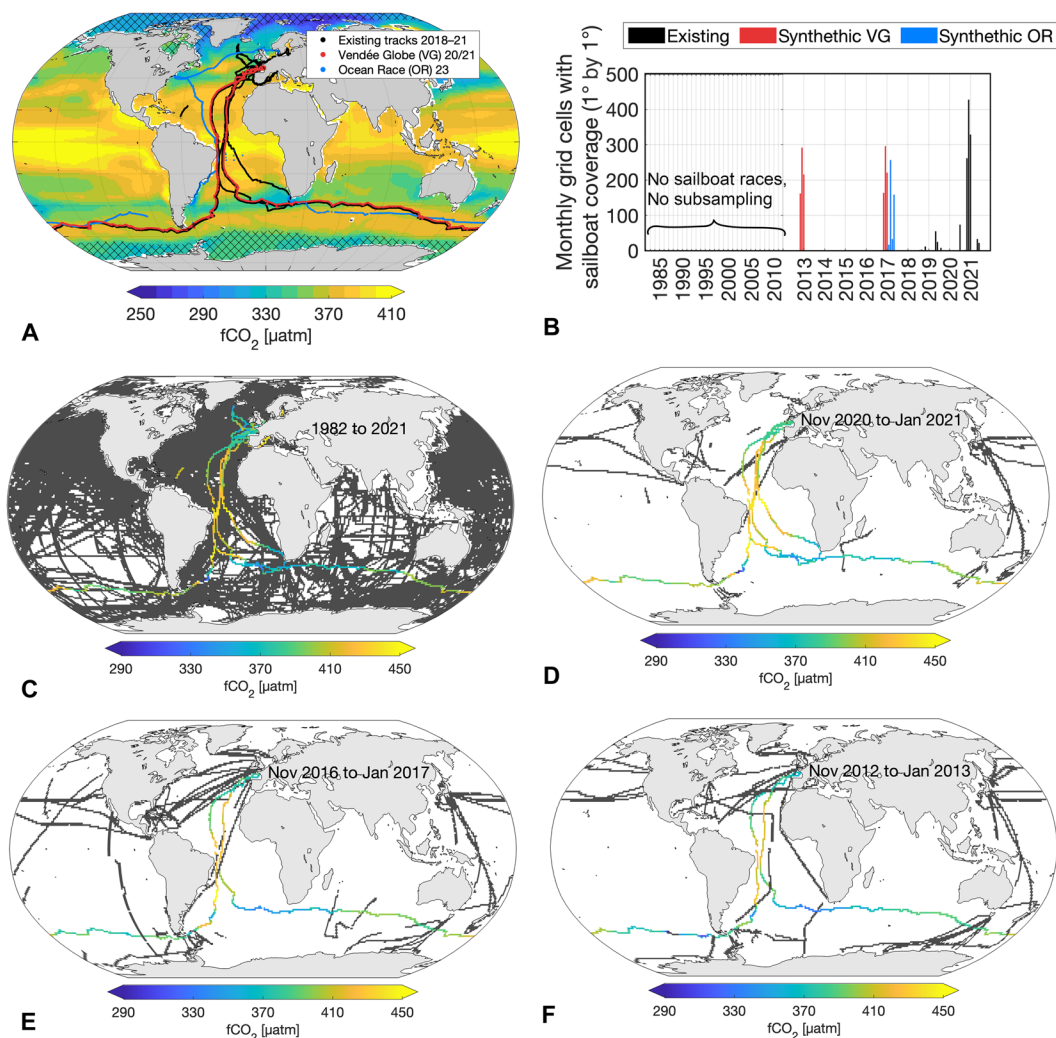


Fig. 1. Sailboat sampling schemes. (A) Reconstructed fCO_2 based on SOCAT sampling. The lines represent different sailboat tracks, which are used alongside the rest of SOCATv2022 tracks in the subsampling process [note: The Vendée Globe track (red) was plotted with a small offset from its original position to avoid overlap with the “existing sailboat” tracks, even though Vendée Globe 2020/21 is included in the “existing sailboat” tracks]. Hatched areas are regions with a climatological maximum sea-ice concentration greater than 50% and are excluded. (B) Sailboat data availability per month. Sampling scheme A “existing sailboat” = black data. Sampling scheme C “3 circumnav.” = black + red data. Sampling scheme D “2 diff. circumnav.” = black + blue data. Sampling scheme B “without sailboat” excludes all colored sailboat track data. (C to F) Tracks used to subsample the model during different time periods for the “3 circumnavigations” run. Gray lines indicate SOCAT tracks; colored lines represent subsampled sailboat fCO_2 tracks.

quantifying improvements in air-sea CO₂ flux estimates from adding different sailboat data tracks; and (iii) assessing the impact of measurement uncertainties and biases.

Performance of neural-network reconstruction

We evaluated the neural network method's performance in reconstructing fCO₂ (based on available present-day observation tracks from SOCATv2022 “existing sailboat”) by comparing the fCO₂ model truth with the “existing sailboat” fCO₂ estimate. By comparing fCO₂ data distributions, we observe a relatively strong agreement (fig. S1). The neural network successfully reconstructed fCO₂ values close to the model truth, with the exception of the polar north (fig. S1), i.e., a region that has been identified as erroneous in previous studies (39, 42) and is thus not further considered here (see Materials and Methods). Probability density functions for the original model fCO₂ and the model-subsampled neural-network-reconstructed fCO₂ show dissimilarities with Bhattacharyya distances (BDs) ranging from 0.00 to 0.14. The BD (43) measures the similarity between two probability distributions, with lower values indicating greater similarity. In this context, the distances ranging from 0.00 to 0.14 suggest a high degree of similarity between the distribution of the model fCO₂ and reconstructed fCO₂.

The fCO₂ distributions in the tropics show the highest agreement with relatively low BDs, suggesting a robust representation of the observed patterns, and almost identical means (0.23 to 1.11 μatm difference) between model and reconstruction followed by the fCO₂ distributions in the middle latitudes (fig. S1). In the Southern polar region, dissimilarities are more pronounced, featuring BD between 0.04 and 0.05 and higher offsets between mean fCO₂ values ranging from 0.92 to 3.78 μatm (fig. S1).

Both over- and underestimation of fCO₂ occur along the coastal ocean (Fig. 2), a region that is highly variable (44, 45). We find that the neural network in combination with the present-day sampling overestimates fCO₂ in the undersampled Southern Ocean by around 2 to 3 μatm (Fig. 2B and fig. S2). A similar overestimation of fCO₂ was found in (30) using a different reconstruction method and a large ensemble testbed of Earth System Models. In contrast, while similar in magnitude to our results, the sign of the mismatch is opposite to a study conducted in (10) using a different single hindcast model, illustrating the limitations of such an analysis to a single model. However, here, we are interested in potential improvements in space and time from adding measurement compared to the baseline. In addition, we do not test the effect of different gas transfer schemes, as our study solely focuses on the improvement in the fCO₂ from increased sailboat sampling. We find that additional sailboat data from “3 circumnavigations” already reduces the fCO₂ in the Southern Ocean, hence improving the reconstruction (fig. S3).

Behncke *et al.* (17) quantified the detectable change caused by the addition of existing sailboat data based on real-world observations. To validate this and determine the similarities between observation and model data, we applied the same method to our subsampled model data and compared reconstructions based on subsampled model data with those based on observations. The neural network demonstrated consistent but smaller signal detection in subsampled model data, mirroring observations closely, thus underscoring the method's realistic performance on both model and observation data (fig. S4).

The effect of the sailboat sampling scheme on the model reconstruction is more subtle, complicating detection as the relationships

between driver variables and fCO₂ are more consistent in the model setup, compared to the real world. Measured fCO₂ is subject to variability and uncertainties, such as measurement uncertainties and temporal averaging when producing gridded fields; fCO₂ measurements taken at one point in time are averaged per month, reducing accuracy. In addition, the effect in observations may be more pronounced because effects of hypothetical biogeochemical and physical processes and their variability that could be missing or unresolved in models might be captured in the real world (39, 46–48).

Effect of sailboat sampling on fCO₂ estimates

Adding data from the existing sailboat tracks does not substantially reduce the fCO₂ bias when averaged over the full time period 1982–2021, probably due to their small number and their late addition to the SOCAT database (Fig. 2, A and B). The global mean fCO₂ is overestimated by 0.99 μatm regardless of the inclusion of sailboat data and particularly in the Southern Ocean by around 3 μatm (Fig. 2, A and B, and fig. S2). Similarly, the median bias is 1.17 μatm with sailboat data and 1.16 μatm without it. The spatial bias pattern stays nearly identical (Fig. 2A). In contrast, adding sailboat fCO₂ data from “3 circumnavigation” tracks substantially reduces biases mainly in the North Atlantic and in the Southern Ocean, mirroring the data addition in these regions (Fig. 2, A to C). The fCO₂ estimate improves with the addition of “3 circumnavigations” particularly between 40°S and 60°S (Fig. 2 and fig. S2), similar to findings in (17). Due to the neural network's capacity to extrapolate over time and space, fCO₂ estimates in other regions, e.g., in the Indian Ocean and West Pacific, where we do not add new data, improve as well (Fig. 2, B and C). However, we also observe compensating effects, indicating a trade-off in other regions, e.g., the South Pacific (Fig. 2, B and C). While the neural network learns new features and minimizes errors, it also adopts inaccurate representations of processes due to overfitting tendencies, resulting in increased biases in, e.g., parts of the East Pacific, where the fCO₂ is underestimated. Globally, however, the median bias decreases to near zero, when adding data from “3 circumnavigation” tracks (Fig. 2).

Effect of sailboat sampling on air-sea CO₂ flux estimates

Using these fCO₂ reconstructions and model wind data, we compute the air-sea CO₂ flux estimates by applying a bulk gas transfer formulation (fig. S5) (37, 49). Figure 3 shows the latitudinal air-sea CO₂ flux estimates and the best guess closest to the model truth with the direction of improvement. Mirroring the spatial improvement pattern seen in the fCO₂ estimate (Fig. 2), the air-sea CO₂ flux estimate improves the most with the addition of “3 circumnavigations,” particularly between 40°S and 60°S (Fig. 3), similar to findings in (17). This results in a more negative air-sea CO₂ flux, indicating an enhanced global ocean carbon sink (Fig. 3). In contrast, we observe no improvement of the reconstruction between 20°S and 40°S (Fig. 3), where the model truth sink is weaker than estimated with more sailboat data.

In addition to spatial patterns, we evaluate the influence of added sailboat data on the temporal trajectory of the air-sea CO₂ flux (Fig. 4). Similar to the estimated fCO₂, the air-sea CO₂ flux does not show substantial improvement with the addition of existing sailboat tracks (see proximity of gray and blue lines in Fig. 4, A to E). However, the global mean air-sea CO₂ flux density bias slightly decreases from 0.06 to −0.02 mol C m^{−2} year^{−1}, and in the Southern Ocean, it improved from 0.10 to 0.00 mol C m^{−2} year^{−1} when data from two

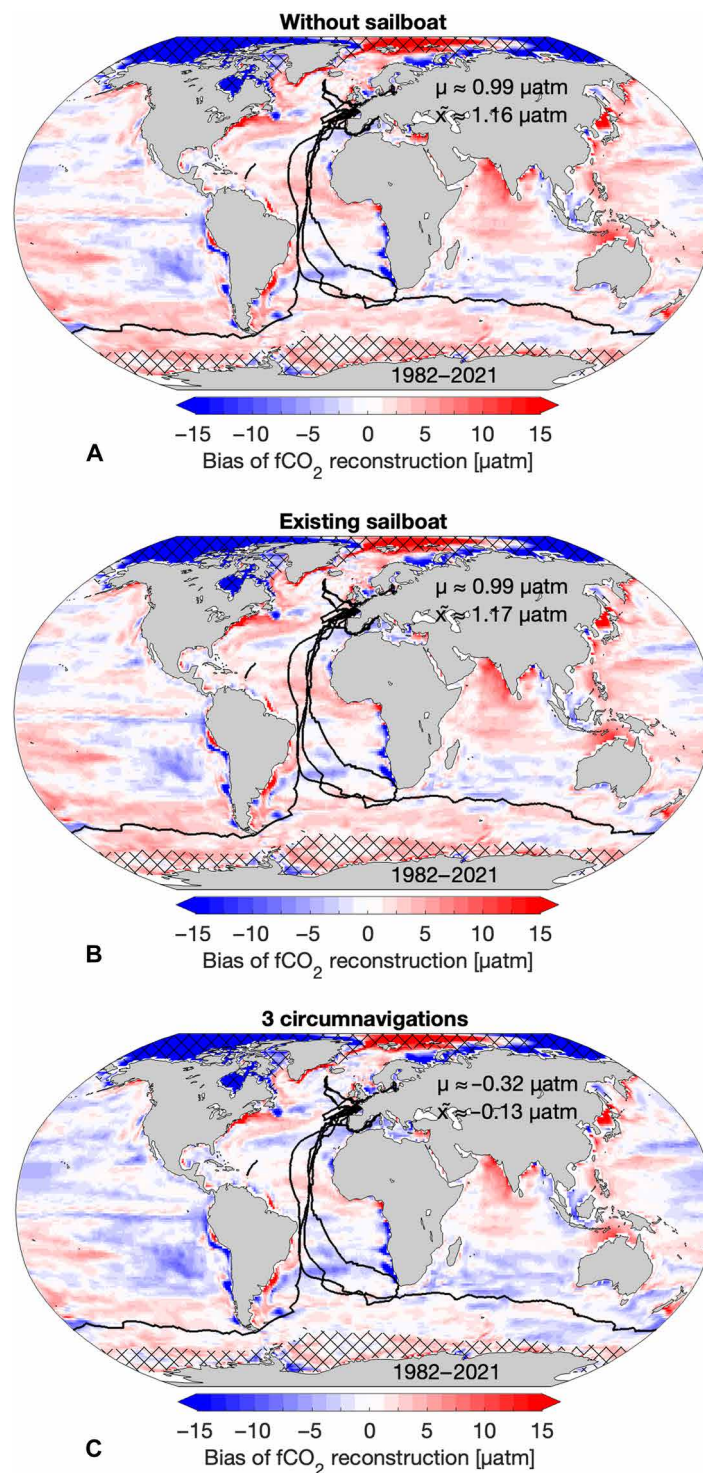


Fig. 2. Spatial fCO_2 bias and improvement patterns resulting from the integration of different sampling schemes. (A to C) Mean bias of subsampled and reconstructed fCO_2 averaged over 1982 until 2021 based on data subsampled after SOCAT (A) “without sailboat” track data B (B) including “existing sailboat” tracks A (C) including existing sailboat tracks and 2 additional circumnavigations (C or “3 circumnav.”). Black lines represent sailboat tracks. Hatched areas are regions with a climatological maximum sea-ice concentration greater than 50% and are excluded. μ represents the mean, while \bar{x} represents the median.

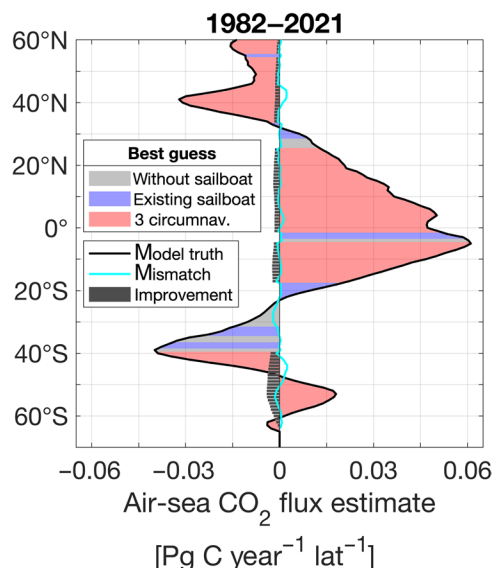


Fig. 3. Latitudinal air-sea CO₂ flux estimates for the period from 1982 to 2021.

The black line represents the model truth. The area under the black curve is shaded to depict the “best guess” estimate closest to the model truth, with the color corresponding to one of three scenarios: A (“existing sailboat”), B (“without sailboat”), or C (“3 circumnav.”). The cyan line highlights the persistent mismatch between the “best guess” estimate and the model truth. Gray bars illustrate the improvement in estimates from scenario B (“without sailboat”) to the “best guess.” For example, bars pointing in the negative direction indicate that the “best guess” reduced the estimate by that amount. A cyan line pointing to the left indicates that the estimate is still too negative, while a line pointing to the right indicates that the estimate is still too positive.

additional circumnavigation tracks (“3 circumnavigations”) were included (Fig. 4, D and E). Without these data, we observe that the annual air-sea CO₂ flux is almost consistently too positive, indicating excessive outgassing and/or insufficient uptake (see gray lines in Fig. 4). However, when data from “3 circumnavigation” tracks are added, the air-sea CO₂ flux estimate decreases by $\sim 0.1 \text{ mol C m}^{-2} \text{ year}^{-1}$. This adjustment brings the estimate closer to the model truth (Fig. 4) by reducing the excessive outgassing in the tropics (Fig. 4B) and increasing the insufficient uptake north and south of 30° before 2000 (Fig. 4, A and C). It also results in an exaggeration of the total carbon sink after 2000 on a global scale (Fig. 4, A to C and D) and in the Southern Ocean (Fig. 4, C and E) as a trade-off.

The inclusion of additional sailboat data not only affects the air-sea CO₂ flux where the data have been added, but the neural network extrapolates in the past, affecting the air-sea CO₂ reconstructions consistently over time. Even though no new data were added prior to 2012/2013, the data addition in the later years influenced the pre-2000 air-sea CO₂ flux estimate (Fig. 4, A to E), similar to results shown in (30), probably due to the sparse data during that period, as less than 1% of the monthly 1°-by-1° grid cells contain measurements (5). In contrast to the findings in (17, 27), where the effect of additional data on the air-sea CO₂ flux was primarily detected in the most recent years, attributed to the use of atmospheric CO₂ as a predictor with trend behavior, our study demonstrates improvement over the entire time series. This more generalized impact, similar to (30), is mainly attributed to the increased data volume spanning multiple years as well as methodological differences such as the different nature of data.

The larger dataset, covering, e.g., a decade starting in the “3 circumnavigation” run, offers broader temporal coverage. This allows the model to learn generalized patterns and make informed predictions further back in time, enhancing its ability to extrapolate beyond the immediate training period. In addition, the use of model data, instead of observations, provides consistent relationships between fCO₂ and predictors that enable the neural network to learn and generalize patterns more effectively.

In the Southern Ocean, the air-sea CO₂ flux density is improved before 2000 and in some of the better observed recent years when adding “3 circumnavigations”; however, it notably worsens between 2005 and 2012 as a trade-off for improving the full time series (Fig. 4, C and E and fig. S6, C and D). Globally, the estimate is most improved before 2000, but it also continues to show improvement afterward (Fig. 4, A to D, and fig. S6, A and B).

Our findings further reveal a remarkable behavior: While we observe a shift in the mean due to the addition of the measurements, the shape of the time series and, consequently, the air-sea CO₂ flux anomalies and trend (represented by dashed lines in Fig. 4, A to C) remains unchanged and continues to be overestimated compared to the model truth trend, particularly in the Southern Ocean, even with the addition of data from two additional circumnavigation tracks (Fig. 4).

Adding data from one additional but different circumnavigation track, such as from The Ocean Race 2023 instead of the Vendée Globe 2020/21 (Fig. 1), to the existing sailboat data improves the air-sea CO₂ flux estimates as effectively as adding data from two identical Vendée Globe 2020/21 tracks (see cyan and red lines in Fig. 4).

The effect of measurement uncertainties and biases on the air-sea CO₂ flux estimate

We explore the effect of two types of measurement errors applied to the subsampled sailboat fCO₂ data on the resulting air-sea CO₂ reconstructions: random measurement uncertainties, representing zero-mean fluctuations in the measurements, and systematic biases, representing constant offsets. Random uncertainties are assumed to be normally distributed around zero, while biases introduce a fixed, consistent deviation from the true value.

Specifically, we consider six scenarios (Fig. 5): (i) a low-end measurement uncertainty of $\pm 5 \mu\text{atm}$ (best-case scenario); (ii and iii) a constant positive or negative bias of $5 \mu\text{atm}$ (low-end systematic offset); (iv) a high-end measurement uncertainty of $\pm 10 \mu\text{atm}$; and (v and vi) a constant positive or negative bias of $10 \mu\text{atm}$ (high-end systematic offset). An uncertainty of $\pm 5 \mu\text{atm}$ is at the lower range of uncertainties observed in the field (50, 51), thus representing a “best-case” estimate, while $10 \mu\text{atm}$ is at the higher end with a systematic offset representing the “worst-case” scenario of limited calibration and maintenance as well as system limitations.

Sailboat data from “3 circumnavigation” tracks with an associated random measurement uncertainty ranging between $\pm 5 \mu\text{atm}$ still improve the air-sea CO₂ flux estimate, presenting a similar picture to unaffected reconstructions, both showing improved reconstructions in the data-sparse period before 2000 and in recent years (see the teal dashed line following the red line in Fig. 5, A to C). This is consistent with other studies examining the effect of random measurement uncertainties, showing that the addition of data can improve air-sea CO₂ flux estimates, even when those data are affected by high random measurement uncertainties (17, 52).

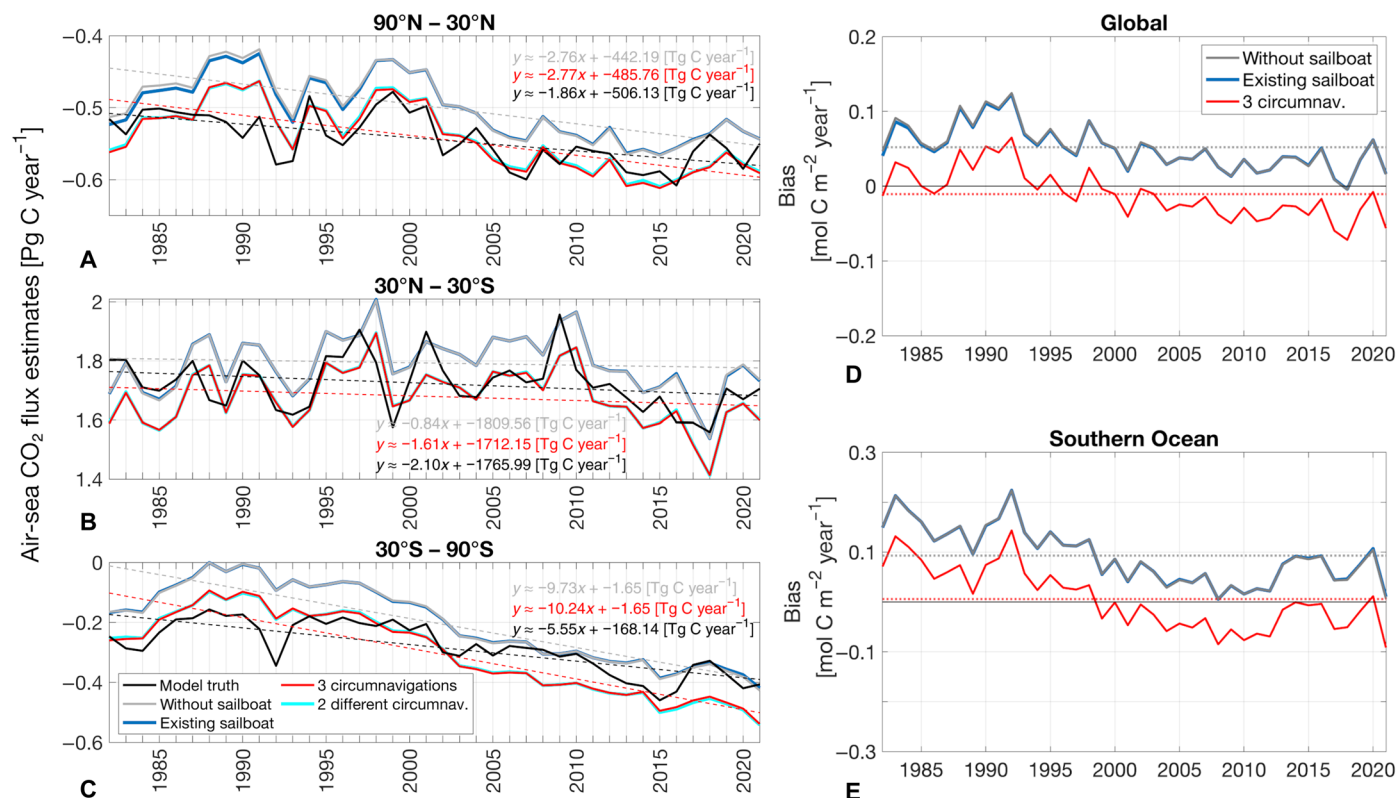


Fig. 4. The impact of adding sailboat data on air-sea CO₂ flux time series. (A to C) Air-sea CO₂ flux estimates with and without different sailboat tracks in different regions. Dashed lines show the long-term trend. (D and E) Time series of bias (global and in Southern Ocean) in reconstructed air-sea CO₂ flux density. Dotted lines represent the long-term mean bias. Note that the gray line, indicating the bias of the estimate based on “without sailboat,” and the blue line, indicating the bias of the estimate based on “existing sailboat,” are in close proximity to each other. Regions with a climatological maximum sea-ice concentration greater than 50% are excluded.

While adding a negative bias to the measurements worsens the reconstruction as expected (dotted lines in Fig. 5), introducing positive measurement offsets of either 5 or 10 μatm to the sailboat data improves the air-sea CO₂ flux reconstructions (dash-dotted lines in Fig. 5). The impact of the systematic offsets, both negative and positive, is most pronounced before 1990 in high latitudes (Fig. 5, A and C) and in recent years where biased sailboat data were added starting around 2015 (Fig. 5, A to C), indicating that the bias is not extrapolated in time, but in space. Positive biases lead to reduced CO₂ uptake in recent years, particularly in the high latitudes (dash-dotted lines in Fig. 5, A and C), and increased outgassing in the tropics (dash-dotted lines in Fig. 5B), whereas negative biases cause the opposite. Similar to (17), we also find that Southern Ocean air-sea CO₂ estimates are most sensitive to measurement errors, particularly measurement biases, with biased estimates diverging from the unbiased estimate around 2005 (Fig. 5C), while estimates in the northern high latitudes and tropics only visibly diverge around 2018/2019 (Fig. 5, A and B).

Notably, positive measurement biases counteract the overestimation of the carbon uptake in high latitudes and underestimation of outgassing in the tropics with “3 circumnavigations” in recent years, bringing the air-sea CO₂ flux estimates even closer to the model truth (see dash-dotted lines and red solid line compared to black line in Fig. 5, A to C). This shows that the impact of biased observations is more complex than one would expect at first sight. Negative offsets, on the other hand, in “3 circumnavigations” lead to an overestimation

of the ocean carbon uptake in high latitudes and underestimation of the outgassing in the tropics especially in recent years.

Focusing on the sailboat tracks alone, we find that both the high and low biased scenarios lead to similar errors, but, as expected, in opposite directions. A 10- μatm positive bias leads to a mean flux bias of $0.14 \text{ mol C m}^{-2} \text{ year}^{-1}$, and a negative measurement bias leads to a flux bias of $-0.16 \text{ mol C m}^{-2} \text{ year}^{-1}$. Therefore, the difference likely originates from other places, where the neural network extrapolates the information gained by sailboat tracks. This is visualized in fig. S7. While adding negative biases worsens the reconstruction, adding positive biases actually improves the reconstruction in the Southern Ocean outgassing region, the South Atlantic, and the Equatorial Pacific.

These results indicate that the neural network method used here (37) and potentially other machine learning methods are sensitive to systematic biases and their extrapolation. Biased measurements may compensate for missing information (in our case, high-end $f\text{CO}_2$) elsewhere, improving the reconstruction for the wrong reason (fig. S8). Therefore, even biased observations can improve reconstructions when data availability is limited and data distribution is skewed (fig. S8), as evidenced by the proximity of the dash-dotted and dotted lines to the solid red lines in Fig. 5; however, our analysis shows that this result is strongly linked to the underlying model tested, the machine learning algorithm, and the direction of the bias.

In conclusion, the improvement in the air-sea CO₂ flux reconstruction caused by the addition of data from “3 circumnavigations”

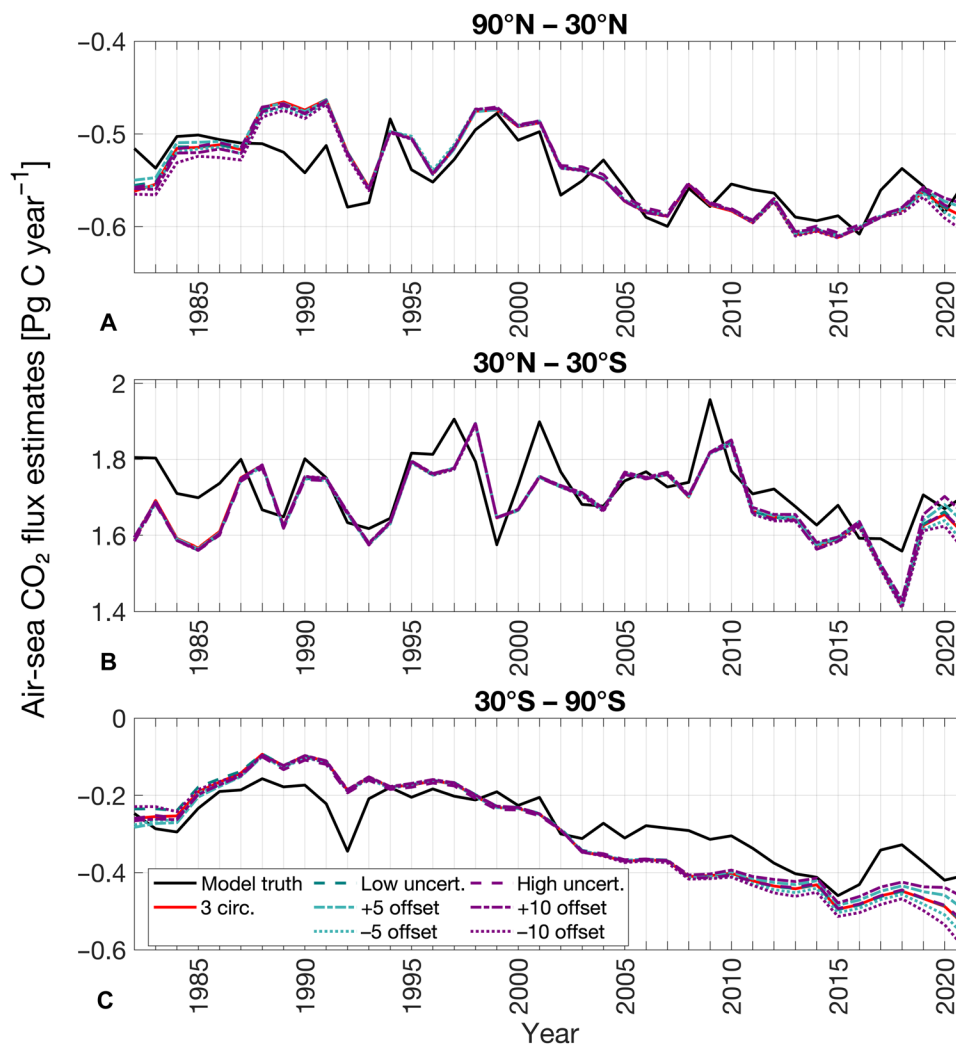


Fig. 5. The effect of measurement uncertainties and biases on the air-sea CO₂ flux estimate. (A to C) Air-sea CO₂ flux estimates based on the model truth (solid black lines), the “3 circumnavigation” sailboat sampling (solid red lines), and the same sampling with added measurement errors. Measurement errors are applied only to the subsampled sailboat fCO₂ data. Teal lines represent cases with low-end errors ($\pm 5 \mu\text{atm}$ uncertainty or $5 \mu\text{atm}$ bias), and purple lines represent high-end errors ($\pm 10 \mu\text{atm}$ uncertainty or $10 \mu\text{atm}$ bias). Random measurement uncertainties are shown as dashed lines (note: they appear on top of each other), while systematic biases are shown as dotted (negative bias) and dash-dotted (positive bias) lines.

persists even when the sailboat data contain uncertainties. Although biased data can still contribute to improved overall reconstructions, for the wrong reason through spatial extrapolation, the bias impact becomes substantial when observational constraints are weak: Underobserved regions and data-sparse periods are particularly sensitive. A positive measurement bias may coincidentally shift the reconstruction closer to the model truth. However, negative biases, especially those approaching the upper limit of what is expected from the measurement system, worsen the air-sea CO₂ flux estimates.

Unlike studies such as (5, 53), which applied bias corrections to the entire underlying observational dataset, our study introduces measurement errors only to a limited subset of sailboat data, similar to (17). We show that our method, likely similar to other machine learning approaches, is sensitive to biased predictors and targets; however, the neural network’s sensitivity depends on the extent of biased data, the measurement coverage, and the underlying model used as ground truth.

Although the addition of “3 (unbiased) circumnavigations” improves estimates across all decades by producing a near-constant shift in the reconstructed air-sea CO₂ flux through time (Fig. 4), the influence of measurement biases is largely confined to the period when these data were sampled (around 2009; Fig. 5). These seemingly contrasting responses arise from different mechanisms by which the neural network integrates information. Adding rare, unbiased, recurring observations enables the model to better represent the underlying statistical distribution of fCO₂ across the full period, thereby learning a more accurate mean relationship between predictors and target fCO₂ and improving the reconstruction even in earlier decades, resulting in a near-zero mean bias before 2011 in the fCO₂ density distribution (fig. S9).

In contrast, introducing systematic measurement biases does not alter the learned relationships several years back in time prior to when they were introduced, but directly affects the estimates where and when the biased data are present. Biased observations influence

the reconstruction over time, but the bias signal itself is not propagated back in time; it is mainly expressed during the years when the biased data are introduced, because, in our interpretation, the earlier decades are already well represented in the learned $f\text{CO}_2$ distribution (fig. S9). Consequently, a systematic bias in the data mainly influences recent years, where these rare observations are added (54).

DISCUSSION

Here, we quantify the improvement in the air-sea CO_2 flux estimate by adding sailboat observations in an observing system simulation experiment. We find that a commonly used neural network method, using the available real-world sampling scheme, underestimates the model-truth ocean carbon sink, which aligns with findings in (30). However, increased sampling with sailboats improves the global estimate and increases the ocean carbon sink from 1982 to the end of 2021, particularly in the North Atlantic and in the Southern Ocean between 40°S and 60°S . For instance, by adding data from “3 circumnavigation” tracks, the global mean air-sea CO_2 flux density bias decreases from 0.06 to -0.02 $\text{mol C m}^{-2} \text{ year}^{-1}$, and in the Southern Ocean from 0.10 to 0.00 $\text{mol C m}^{-2} \text{ year}^{-1}$.

Despite the observed shift in the mean air-sea CO_2 flux when adding more data, the overestimated air-sea CO_2 flux trend persists even with the addition of two additional circumnavigations. We conclude that data from two additional circumnavigations are insufficient to improve the trend, suggesting that the Global Carbon Budget model and data products are likely to continue diverging even with the incorporation of more circumnavigation data (1, 46). Additional data are needed to improve the trend. Currently, the trend in the reconstructions deviates from the model trend. By incorporating more input data over multiple decades that inherently follow the correct trend into the neural network, the output reconstructions are expected to align more closely with the model's true trend. Thus, we conclude that the inclusion of more sailboat data with trend-consistent information could improve the trend accuracy.

This underscores the need for implementing multidecadal observing strategies. Regular sailboat circumnavigations, such as The Ocean Race and the Vendée Globe, which alternate on a staggered 4-year cycle, present a viable approach. These events ensure a major round-the-world race approximately every 1 to 2 years, thereby providing consistent and comprehensive datasets to refine our understanding and improve the accuracy of reconstructions over extended temporal scales. However, sailboats cannot fully cover seasonal measurement gaps, especially in winter at high latitudes. This further underscores the need to support complementary platforms such as autonomous floats (27, 28) and Saildrones (32), which provide year-round data under challenging conditions. Combining these platforms ensures more complete coverage and improves long-term monitoring of air-sea CO_2 fluxes.

Previous studies have shown that adding data from sailboats and Argo floats to the observing system can significantly affect and improve $f\text{CO}_2$ and air-sea CO_2 flux reconstructions, even when the data contain measurement uncertainties, but not if these data are biased (17, 52). Here, we show that adding more data, even if they contain uncertainties, can improve air-sea CO_2 flux reconstructions and strengthen the overall observing system. This supports previous findings (55, 56) that data quantity can compensate for reduced data quality. However, biased data come with a trade-off. They can still

help improve overall reconstructions, but for the wrong reason. Compensating for the lack of high-end $f\text{CO}_2$ samples in the machine learning training data, biased measurements led to an improved reconstruction in high $f\text{CO}_2$ regions. As a result, biased data improve global reconstructions while they reduce accuracy regionally where the measurements are actually taken. Such improvements have to be taken with caution as they are likely machine learning- and model-dependent and may not be translatable into the real world, creating a false acceptance for measurement biases. Instead, we show in our study that to interpret reconstructions with biased measurements, it is essential to consider the underlying data density distribution, the model used for the testbed analysis as well as the machine learning background bias (relative to the model truth), and the direction (positive or negative) of any measurement bias before interpreting the machine learning extrapolation, and we highly recommend to consider these in future observing system design studies.

Although the finding that adding new measurements improves the air-sea CO_2 flux estimates probably holds irrespective of the baseline model choice, the observed magnitude of improvement from increased sailboat sampling is highly model-dependent and could vary with different models. We work under the assumption that HAMOCC represents a known truth for our experiment; however, the HAMOCC model is known to underestimate surface ocean $f\text{CO}_2$ in the Southern Ocean (39) and exhibits an amplified seasonal cycle (20). This could affect the results by making the improvements appear more pronounced or less consistent across seasons than they would with a model that accurately captures seasonal variations in $f\text{CO}_2$, potentially distorting the long-term trend. Furthermore, we only use one neural network method (37) on the model data, and different neural network methods might respond differently to data addition. We recommend more extensive analyses in the future including different neural network approaches and different model data, i.e., the products used in the Global Carbon Budget (1).

While we studied how underway $f\text{CO}_2$ measurements from sailboats affect and improve reconstructions, we have yet to identify the specific ocean features responsible for this effect. To determine why these changes and improvements occur, future research should focus on another added value of the data: the high-resolution nature of sailboat data, which resolves important small-scale ocean features driving variability in the air-sea CO_2 flux [e.g., (57)].

The global $f\text{CO}_2$ measurement coverage is biased toward the Northern Hemisphere, which leads to reconstruction errors; however, the addition of sailboat data provides better data coverage in the Southern Ocean with the potential to improve air-sea CO_2 flux reconstructions in the future. The frequency of the sailboat races meets the demand for more affordable and innovative platforms (16, 58–61). Considering that the Southern Ocean carbon uptake will increase (62, 63) and the observational coverage recently declined [www.socat.info; (5)], more observations are needed to closely monitor the ocean carbon sink over multiple decades, and sailboat races provide the opportunity to do that. We conclude that while the addition of sailboat data has the potential to improve air-sea CO_2 flux reconstructions, particularly in the Southern Ocean, further expansion of the sailboat-based observational network is essential to help minimize discrepancies in Global Carbon Budget sink estimates. Continued efforts to increase the volume and coverage of sailboat measurements will be crucial in refining our understanding of oceanic carbon dynamics.

MATERIALS AND METHODS

Datasets and sampling masks

In this study, our primary aim is to quantify the improvement in our air-sea CO₂ flux estimate caused by adding different sailboat data to our flux reconstructions. To achieve this, we use the model fCO₂ field as our starting point and apply different subsampling schemes (Fig. 1) to get pseudo-observations. Most of the races took place in the North Atlantic between 2018 and the end of 2021. However, the Antarctic circumnavigation race, Vendée Globe, primarily navigated the Southern Ocean and occurred in 2020/2021, repeating every 4 years. Another notable round-the-world race contributing substantial Southern Ocean data is The Ocean Race, held in 2023 and recurring approximately every 3 years. This study focuses on the period from 1982 to 2021, using The Ocean Race tracks solely to subsample earlier years.

Data X—"truth"

Ground truth fCO₂ values from the HAMOCC model coupled to the ocean general circulation model MPIOM, contributing to Global Carbon Budget simulation at a monthly 1°-by-1° resolution from 1982 until the end of 2021 (1, 39–41).

The HAMOCC model simulates the oceanic cycles of carbon in the global MPIOM. HAMOCC features biology and inorganic carbon chemistry processes in the water column and sediment (39, 64). Marine primary producers are represented by two state variables: bulk phytoplankton and diazotrophs. The growth of bulk phytoplankton is limited by temperature and light as well as by the availability of nutrients including nitrate, phosphate, and iron linked by constant Redfield ratios across organic compartments. The growth of nitrogen-fixing cyanobacteria is parameterized analogously to that of the bulk phytoplankton, but at a lower rate, and is extended by representing their buoyancy. Zooplankton grazes on bulk phytoplankton, producing particulate organic matter that enters the detritus pool. Opal particles or CaCO₃ is produced during detritus formation, depending on silicate availability. HAMOCC simulates the upper sediment by 12 biologically active layers and a burial layer to represent the dissolution and decomposition of inorganic and organic matter as well as the diffusion of pore water constituents (65). The HAMOCC model has been extensively evaluated in previous studies and successfully used for climate predictions and projections as well as simulating the past climate (39–41, 66–70).

Data A—"existing sailboat"

Subsampled model fCO₂ (X) generated using the SOCATv2022 sampling scheme, including sailboat data. We used underway fCO₂ measurement tracks from the International Monohull Open Class Association (IMOCA) 60-foot class sailboats *Seaexplorer-Yacht Club de Monaco* (until 2019 *Malizia*) and *Newrest-Art & Fenêtres* (now *Nexans-Art & Fenêtres*) during offshore sailing and training events from 2018 to 2021. Together, all sailboats collected 161 days of fCO₂ measurements until the end of 2021.

Data B—"without sailboat"

Subsampled model fCO₂ (X) generated using the SOCATv2022 sampling scheme, excluding sailboat data.

Data C—"3 circumnavigations"

Subsampled model fCO₂ (X) generated using the SOCATv2022 sampling scheme including existing sailboat data and two additional Antarctic circumnavigations (Vendée Globe; 2012/13, 2016/17, and 2020/21) (from which only the latter one is from the existing track and also included in data A; the track of the Vendée Globe 2020/21,

where fCO₂ measurements were taken, was used to subsample the model data in earlier years when Vendée Globe races occurred without fCO₂ measurements.)

Data D—"2 different circumnavigations"

Subsampled model fCO₂ (X) generated using the SOCATv2022 sampling scheme including sailboat data and two different round-the-world races (Vendée Globe 2020/21 and the preliminary The Ocean Race 2017/18) (for the latter one, we used existing tracks from 2023).

In the following, we added random measurement uncertainties and biases to the subsampled sailboat fCO₂ data in "3 circumnavigations":

Data E—"3 circumnav + low uncert"

Data C, but with an added random measurement uncertainty ranging from ±5 μatm (lower end of what is expected from the measurement device) to the data from sailboat tracks.

Data F—"3 circumnav + 5 offset"

Data C, but with an added +5-μatm measurement offset to the data from sailboat tracks.

Data G—"3 circumnav – 5 offset"

Data C, but with an added –5-μatm measurement offset to the data from sailboat tracks.

Data H—"3 circumnav + high uncert"

Data C, but with an added random measurement uncertainty ranging from ±10 μatm (higher end of what is expected from the measurement device) to the data from sailboat tracks.

Data I—"3 circumnav + 10 offset"

Data C, but with an added +10-μatm measurement offset to the data from sailboat tracks.

Data J—"3 circumnav – 10 offset"

Data C, but with an added –10-μatm measurement offset to the data from sailboat tracks.

The reconstruction of sea surface fCO₂ and air-sea CO₂ fluxes

The gaps in all subsampled sea surface fCO₂ maps (data A to J) were filled using the two-step neural network method called SOM-FFN (37). In the first step, an SOM classified the ocean into biogeochemical provinces based on common patterns in predictor variables, including HAMOCC sea-surface temperature, sea-surface salinity, mixed layer depth, and an fCO₂ climatology (1, 39–41). The second step involved an FFN establishing nonlinear relationships between predictors and subsampled fCO₂ data within each province. The source code for the neural network method can be found in (71). Predictors were HAMOCC sea-surface temperature, sea-surface salinity, mixed layer depth, atmospheric CO₂ concentration, and phytoplankton plus cyanobacteria biomass integrated over a depth of 37 m. This specific depth, limited to the upper ocean biology and excluding the deep chlorophyll maxima, was chosen due to the limitations of equivalently and previously used ocean color observations for case-1 water, which are restricted to the first optical depth below 40 m (72). The MPIOM-HAMOCC model output can be found in (73). Based on the reconstructed fCO₂ data A to J and model truth X, we computed flux estimates A to J and X by applying a bulk gas transfer formulation with a quadratic relationship between wind speed and transfer velocity to the reconstructed maps (A to J), as well as the model truth X (37, 49). The mean gas transfer was standardized to a global average rate of 16.5 cm hour⁻¹ (74). While the neural network reconstructed fCO₂ closely matches the model fCO₂,

the reconstructed air-sea CO₂ flux differs from the model flux. This discrepancy arises because of the bulk gas transfer parametrization with a standardized mean gas transfer velocity of 16.5 cm hour⁻¹, which is not ideal for the model fCO₂. Acknowledging that our primary focus is not on realistically estimating the air-sea CO₂ flux but rather on assessing the impact of different sampling schemes on the flux estimate, we decided to eliminate this uncertainty by applying the same parametrization on the model fCO₂ (49). This allowed us to establish a comparable “original” flux model truth for our analysis.

We focus both on the air-sea CO₂ flux and the fCO₂ because of the high sensitivity of air-sea CO₂ flux estimates to errors in fCO₂. In high wind regions, small fCO₂ errors can lead to large flux errors due to high gas transfer velocity, while the same fCO₂ error in low wind regions results in smaller flux errors. Furthermore, the direction of the fCO₂ error improvement (negative or positive) significantly affects the air-sea CO₂ flux estimate. A small negative improvement in fCO₂ can either decrease or increase the flux estimate depending on whether the ocean fCO₂ is greater or less than the atmospheric fCO₂, respectively. Similarly, a small positive improvement can increase or decrease the flux estimate based on the same conditions. By focusing on both estimates, we ensure a comprehensive assessment, minimizing the risk of overlooking critical factors that could skew the air-sea CO₂ flux estimates.

In this study, we define the air-sea CO₂ flux density as the instantaneous flux into or out of the ocean [in units of moles of carbon per square meter per year, (mol C m⁻² year⁻¹)], where positive flux indicates outgassing and negative flux indicates uptake (fig. S5), generally aligning its direction with fCO₂. In contrast, the air-sea CO₂ flux—as well as the ocean carbon sink when referring to a negative integrated signal—is quantified as the spatially integrated flux over a given area [in petagrams of carbon, (PgC)] and typically follows the opposite direction.

Statistical analyses

We exclude regions with a climatological maximum sea-ice concentration greater than 50% from most of our analysis as sparse observations (4), and the influence of seasonal changes in sea-ice coverage introduces high uncertainties into the neural network reconstruction in that region (11, 12, 45, 75) and due to the model’s less realistic representation of the high latitudes (39). The ice zone in the Southern Ocean coincides with the high uptake region south of ~60°S (fig. S5).

Neural network performance evaluation

To evaluate the performance of the neural network in reconstructing the HAMOCC fCO₂, we use probability density functions and calculate the BD (43), which measures the similarity between the probability density functions of the reconstruction and the original model fCO₂. Lower BD values indicate a higher degree of similarity between the probability density functions. We further compare the detectable signal in the air-sea CO₂ flux estimate caused by the addition of sailboat data in our subsampled model data to the signal in observation data from a previous study using a signal-to-noise detection described in (17).

Improvement quantification

We used a Monte Carlo approach, generating 10- to 40-member ensembles for each of the 10 subsampling scenarios. For scenarios A (“existing sailboat”), B (“without sailboat”), C (“3 circumnavigations”), D (“2 different circumnavigations”), E (“3 circumnavigations + low

measurement uncertainty”), and F (“3 circumnavigations + positive 5 μatm offset”), we generated 40 ensemble members each. For scenarios G (“3 circumnavigations – 5 μatm offset”), H (“3 circumnavigations + high measurement uncertainty”), I (“3 circumnavigations + 10 μatm offset”), and J (“3 circumnavigations – 10 μatm offset”), we generated 10 ensemble members each, as the standard deviation across 10 versus 40 ensembles differs only slightly and using fewer members allows for more computationally efficient analysis.

We generated the ensembles by varying training and validation dataset splits to enhance the reliability of our air-sea CO₂ flux estimates and identify potential random errors caused by the sensitivity of the neural-network approach to different subsets of the data. The ensemble mean served as the best estimate for each sampling scenario.

The bias is calculated as the mean of the reconstruction R (i.e., A, B, and C) minus the model truth (X): bias = mean(R) – mean(X) and measures the over- and underestimation in the reconstructions over different time periods. Values near zero not only indicate a good reconstruction but could also indicate that positive and negative differences cancel out. We calculated the mean bias as well as the median bias to minimize the impact of regional outliers caused by regions of high uncertainty (Fig. 2).

Supplementary Materials

This PDF file includes:

Figs. S1 to S9
References

REFERENCES

1. P. Friedlingstein, M. O’Sullivan, M. W. Jones, R. M. Andrew, J. Hauck, P. Landschützer, C. Le Quéré, H. Li, I. T. Luijckx, A. Olsen, G. P. Peters, W. Peters, J. Pongratz, C. Schwingshackl, S. Sitch, J. G. Canadell, P. Ciais, R. B. Jackson, S. R. Alin, A. Arneth, V. Arora, N. R. Bates, M. Becker, N. Bellouin, C. F. Berghoff, H. C. Bittig, L. Bopp, P. Cadule, K. Campbell, M. A. Chamberlain, N. Chandra, F. Chevallier, L. P. Chini, T. Colligan, J. Decayeux, L. M. Djeutchouang, X. Dou, C. Duran Rojas, K. Enyo, W. Evans, A. R. Fay, R. A. Feely, D. J. Ford, A. Foster, T. Gasser, M. Gehlen, T. Gkritzalis, G. Grassi, L. Gregor, N. Gruber, Ö. Gürses, I. Harris, M. Hefner, J. Heinke, G. C. Hurtt, Y. Iida, T. Ilyina, A. R. Jacobson, A. K. Jain, T. Jarniková, A. Jersild, F. Jiang, Z. Jin, E. Kato, R. F. Keeling, K. Klein Goldewijk, J. Knauer, J. I. Korsbakken, X. Lan, S. K. Lauvset, N. Lefèvre, Z. Liu, J. Liu, L. Ma, S. Maksyutov, G. Marland, N. Mayot, P. C. McGuire, N. Metz, N. M. Monacchi, E. J. Morgan, S.-I. Nakaoka, C. Neill, Y. Niwa, T. Nützel, L. Olivier, T. Ono, P. I. Palmer, D. Pierrot, Z. Qin, L. Resplandy, A. Roobaert, T. M. Rosan, C. Rödenbeck, J. Schwinger, T. L. Smallman, S. M. Smith, R. Sospedra-Alfonso, T. Steinhoff, Q. Sun, A. J. Sutton, R. Séférian, S. Takao, H. Tatebe, H. Tian, B. Tilbrook, O. Torres, E. Tourigny, H. Tsujino, F. Tubiello, G. van der Werf, R. Wanninkhof, X. Wang, D. Yang, X. Yang, Z. Yu, W. Yuan, X. Yue, S. Zaehle, N. Zeng, J. Zeng, Global carbon budget 2024. *Earth Syst. Sci. Data* **17**, 965–1039 (2025).
2. N. Gruber, D. C. E. Bakker, T. DeVries, L. Gregor, J. Hauck, P. Landschützer, G. A. McKinley, J. D. Müller, Trends and variability in the ocean carbon sink. *Nat. Rev. Earth Environ.* **4**, 119–134 (2023).
3. C. Le Quéré, C. Rödenbeck, E. T. Buitenhuis, T. J. Conway, R. Langenfelds, A. Gomez, C. Labuschagne, M. Ramonet, T. Nakazawa, N. Metz, N. Gillett, M. Heimann, Saturation of the Southern Ocean CO₂ sink due to recent climate change. *Science* **316**, 1735–1738 (2007).
4. D. C. E. Bakker, B. Pfeil, C. S. Landa, N. Metz, K. M. O’Brien, A. Olsen, K. Smith, C. Cosca, S. Harasawa, S. D. Jones, S. Nakaoka, Y. Nojiri, U. Schuster, T. Steinhoff, C. Sweeney, T. Takahashi, B. Tilbrook, C. Wada, R. Wanninkhof, S. R. Alin, C. F. Balestrini, L. Barbero, N. R. Bates, A. A. Bianchi, F. Bonou, J. Boutin, Y. Bozec, E. F. Burger, W.-J. Cai, R. D. Castle, L. Chen, M. Chierici, K. Currie, W. Evans, C. Featherstone, R. A. Feely, A. Fransson, C. Goyet, N. Greenwood, L. Gregor, S. Hankin, N. J. Hardman-Mountford, J. Harlay, J. Hauck, M. Hoppema, M. P. Humphreys, C. W. Hunt, B. Huss, J. S. P. Ibáñez, T. Johannessen, R. Keeling, V. Kitidis, A. Körtzinger, A. Kozyr, E. Krasakopoulou, A. Kuwata, P. Landschützer, S. K. Lauvset, N. Lefèvre, C. Lo Monaco, A. Manke, J. T. Mathis, L. Merlivat, F. J. Millero, P. M. S. Monteiro, D. R. Munro, A. Murata, T. Newberger, A. M. Omar, T. Ono, K. Paterson, D. Pearce, D. Pierrot, L. L. Robbins, S. Saito, J. Salisbury, R. Schlitzer, B. Schneider, R. Schweitzer, R. Sieger, I. Skjelvan, K. F. Sullivan, S. C. Sutherland, A. J. Sutton, K. Tadokoro,

- M. Telszewski, M. Tuma, S. M. A. C. van Heuven, D. Vandemark, B. Ward, A. J. Watson, S. Xu, A multi-decadal record of high-quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT). *Earth Syst. Sci. Data* **8**, 383–413 (2016).
5. Y. Dong, D. C. E. Bakker, P. Landschützer, Accuracy of ocean CO₂ uptake estimates at a risk by a reduction in the data collection. *Geophys. Res. Lett.* **51**, e2024GL108502 (2024).
 6. Z.-P. Jiang, J. Yuan, S. E. Hartman, W. Fan, Enhancing the observing capacity for the surface ocean by the use of volunteer observing ship. *Acta Oceanol. Sin.* **38**, 114–120 (2019).
 7. B. Pfeil, A. Olsen, D. C. E. Bakker, S. Hankin, H. Koyuk, A. Kozyr, J. Malczyk, A. Manke, N. Metzl, C. L. Sabine, J. Akl, S. R. Alin, N. Bates, R. G. J. Bellerby, A. Borges, J. Boutin, P. J. Brown, W.-J. Cai, F. P. Chavez, A. Chen, C. Cosca, A. J. Fassbender, R. A. Feely, M. González-Dávila, C. Goyet, B. Hales, N. Hardman-Mountford, C. Heinze, M. Hood, M. Hoppema, C. W. Hunt, D. Hydes, M. Ishii, T. Johannessen, S. D. Jones, R. M. Key, A. Körtzinger, P. Landschützer, S. K. Lauvset, N. Lefèvre, A. Lenton, A. Lourantou, L. Merlivat, T. Midorikawa, L. Mintrop, C. Miyazaki, A. Murata, A. Nakadate, Y. Nakano, S. Nakaoka, Y. Nojiri, A. M. Omar, X. A. Padin, G.-H. Park, K. Paterson, F. F. Perez, D. Pierrot, A. Poisson, A. F. Rios, J. M. Santana-Casiano, J. Salisbury, V. V. S. S. Sarma, R. Schlitzer, B. Schneider, U. Schuster, R. Sieger, I. Skjelvan, T. Steinhoff, T. Suzuki, T. Takahashi, K. Tedesco, M. Telszewski, H. Thomas, B. Tilbrook, J. Tjiputra, D. Vandemark, T. Venes, R. Wanninkhof, A. J. Watson, R. Weiss, C. S. Wong, H. Yoshikawa-Inoue, A. Uniform, Quality controlled Surface Ocean CO₂ Atlas (SOCAT). *Earth Syst. Sci. Data* **5**, 125–143 (2013).
 8. C. L. Sabine, S. Hankin, H. Koyuk, D. C. E. Bakker, B. Pfeil, A. Olsen, N. Metzl, A. Kozyr, A. Fassbender, A. Manke, J. Malczyk, J. Akl, S. R. Alin, R. G. J. Bellerby, A. Borges, J. Boutin, P. J. Brown, W.-J. Cai, F. P. Chavez, A. Chen, C. Cosca, R. A. Feely, M. González-Dávila, C. Goyet, N. Hardman-Mountford, C. Heinze, M. Hoppema, C. W. Hunt, D. Hydes, M. Ishii, T. Johannessen, R. M. Key, A. Körtzinger, P. Landschützer, S. K. Lauvset, N. Lefèvre, A. Lenton, A. Lourantou, L. Merlivat, T. Midorikawa, L. Mintrop, C. Miyazaki, A. Murata, A. Nakadate, Y. Nakano, S. Nakaoka, Y. Nojiri, A. M. Omar, X. A. Padin, G.-H. Park, K. Paterson, F. F. Perez, D. Pierrot, A. Poisson, A. F. Rios, J. M. Santana-Casiano, V. V. S. S. Sarma, R. Schlitzer, B. Schneider, U. Schuster, R. Sieger, I. Skjelvan, T. Steinhoff, T. Suzuki, T. Takahashi, K. Tedesco, M. Telszewski, H. Thomas, B. Tilbrook, D. Vandemark, T. Venes, A. J. Watson, R. Weiss, C. S. Wong, H. Yoshikawa-Inoue, Surface Ocean CO₂ Atlas (SOCAT) gridded data products. *Earth Syst. Sci. Data* **5**, 145–153 (2013).
 9. L. Gloege, G. A. McKinley, P. Landschützer, A. R. Fay, T. L. Frölicher, J. C. Fyfe, T. Ilyina, S. Jones, N. S. Lovenduski, K. B. Rodgers, S. Schlunegger, Y. Takano, Quantifying errors in observationally based estimates of ocean carbon sink variability. *Global Biogeochem. Cycles* **35**, e2020GB006788 (2021).
 10. J. Hauck, C. Nissen, P. Landschützer, C. Rödenbeck, S. Bushinsky, A. Olsen, Sparse observations induce large biases in estimates of the global ocean CO₂ sink: An ocean model subsampling experiment. *Philos. Trans. R. Soc. A* **381**, 20220063 (2023).
 11. A. Jersild, P. Landschützer, A spatially explicit uncertainty analysis of the air-sea CO₂ flux from observations. *Geophys. Res. Lett.* **51**, e2023GL106636 (2024).
 12. D. J. Ford, J. Blannin, J. Watts, A. J. Watson, P. Landschützer, A. Jersild, J. D. Shutler, A comprehensive analysis of air-sea CO₂ flux uncertainties constructed from surface ocean data products. *Global Biogeochem. Cycles* **38**, e2024GB008188 (2024).
 13. T. L. Frölicher, J. L. Sarmiento, D. J. Paynter, J. P. Dunne, J. P. Krasting, M. Winton, Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *J. Climate* **28**, 862–886 (2015).
 14. P. Landschützer, N. Gruber, D. C. E. Bakker, Decadal variations and trends of the global ocean carbon sink. *Global Biogeochem. Cycles* **30**, 1396–1417 (2016).
 15. R. G. Williams, A. J. S. Meijers, V. M. Roussenov, A. Katavouta, P. Ceppi, J. P. Rosser, P. Salvi, Asymmetries in the Southern Ocean contribution to global heat and carbon uptake. *Nat. Clim. Chang.* **14**, 823–831 (2024).
 16. L. Gregor, A. D. Lebehot, S. Kok, P. M. S. Monteiro, A comparative assessment of the uncertainties of global surface-ocean CO₂ estimates using a machine learning ensemble (CSIR-ML6 Version 2019a)—Have we hit the wall? *Geosci. Model Dev. Discuss.* **12**, 5113–5136 (2019).
 17. J. Behncke, P. Landschützer, T. Tanhua, A detectable change in the air-sea CO₂ flux estimate from sailboat measurements. *Sci. Rep.* **14**, 3345 (2024).
 18. P. Landschützer, T. Tanhua, J. Behncke, L. Keppler, Sailing through the southern seas of air-sea CO₂ flux uncertainty. *Philos. Trans. R. Soc. A* **381**, 20220064 (2023).
 19. F. F. Pérez, M. Becker, N. Goris, M. Gehlen, M. López-Mozos, J. Tjiputra, A. Olsen, J. D. Müller, I. E. Huertas, T. T. Chau, V. Cainzos, A. Velo, G. Benard, J. Hauck, N. Gruber, R. Wanninkhof, An assessment of CO₂ storage and sea-air fluxes for the Atlantic Ocean and Mediterranean sea between 1985 and 2018. *Global Biogeochem. Cycles* **38**, e2023GB007862 (2024).
 20. P. Rustogi, P. Landschützer, S. Brune, J. Baehr, The impact of seasonality on the annual air-sea carbon flux and its interannual variability. *NPJ Clim. Atmos. Sci.* **6**, 66 (2023).
 21. C. Ostle, P. Landschützer, M. Edwards, M. Johnson, S. Schmidtko, U. Schuster, A. J. Watson, C. Robinson, Multidecadal changes in biology influence the variability of the North Atlantic carbon sink. *Environ. Res. Lett.* **17**, 114056 (2022).
 22. X. Yang, C. A. Wynn-Edwards, P. G. Strutton, E. H. Shadwick, Drivers of air-sea CO₂ flux in the subantarctic zone revealed by time series observations. *Global Biogeochem. Cycles* **38**, e2023GB007766 (2024).
 23. F. P. Chavez, J. Sevadjan, C. Wahl, J. Friederich, G. E. Friederich, Measurements of pCO₂ and pH from an autonomous surface vehicle in a coastal upwelling system. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **151**, 137–146 (2018).
 24. T. Daniel, J. Manley, N. Trenaman, The wave glider: Enabling a new approach to persistent ocean observation and research. *Ocean Dyn.* **61**, 1509–1520 (2011).
 25. P. M. S. Monteiro, L. Gregor, M. Lévy, S. Maenner, C. L. Sabine, S. Swart, Intraseasonal variability linked to sampling alias in air-sea CO₂ fluxes in the Southern Ocean. *Geophys. Res. Lett.* **42**, 8507–8514 (2015).
 26. S.-A. Nicholson, D. B. Whitt, I. Fer, M. D. du Plessis, A. D. Lebehot, S. Swart, A. J. Sutton, P. M. S. Monteiro, Storms drive outgassing of CO₂ in the subpolar Southern Ocean. *Nat. Commun.* **13**, 158 (2022).
 27. S. M. Bushinsky, P. Landschützer, C. Rödenbeck, A. R. Gray, D. Baker, M. R. Mazloff, L. Resplandy, K. S. Johnson, J. L. Sarmiento, Reassessing Southern Ocean air-sea CO₂ flux estimates with the addition of biogeochemical float observations. *Global Biogeochem. Cycles* **33**, 1370–1388 (2019).
 28. A. R. Gray, K. S. Johnson, S. M. Bushinsky, S. C. Riser, J. L. Russell, L. D. Talley, R. Wanninkhof, N. L. Williams, J. L. Sarmiento, Autonomous biogeochemical floats detect significant carbon dioxide outgassing in the high-latitude Southern Ocean. *Geophys. Res. Lett.* **45**, 9049–9057 (2018).
 29. N. L. Williams, L. W. Juranek, R. A. Feely, K. S. Johnson, J. L. Sarmiento, L. D. Talley, A. G. Dickson, A. R. Gray, R. Wanninkhof, J. L. Russell, S. C. Riser, Y. Takeshita, Calculating surface ocean pCO₂ from biogeochemical Argo floats equipped with pH: An uncertainty analysis. *Global Biogeochem. Cycles* **31**, 591–604 (2017).
 30. T. H. Heimdal, G. A. McKinley, A. J. Sutton, A. R. Fay, L. Gloege, Assessing improvements in global ocean pCO₂ machine learning reconstructions with Southern Ocean autonomous sampling. *Biogeosciences* **21**, 2159–2176 (2024).
 31. S. Nickford, J. B. Palter, L. Mu, The importance of contemporaneous wind and pCO₂ measurements for regional air-sea CO₂ flux estimates. *J. Geophys. Res. Oceans* **129**, e2023JC020744 (2024).
 32. A. J. Sutton, N. L. Williams, B. Tilbrook, Constraining Southern Ocean CO₂ flux uncertainty using uncrewed surface vehicle observations. *Geophys. Res. Lett.* **48**, e2020GL091748 (2021).
 33. A. Denvil-Sommer, M. Gehlen, M. Vrac, Observation system simulation experiments in the Atlantic Ocean for enhanced surface ocean pCO₂ reconstructions. *Ocean Sci.* **17**, 1011–1030 (2021).
 34. L. M. Djeutchouang, N. Chang, L. Gregor, M. Vichi, P. M. S. Monteiro, The sensitivity of pCO₂ reconstructions to sampling scales across a Southern Ocean sub-domain: A semi-idealized ocean sampling simulation approach. *Biogeosciences* **19**, 4171–4195 (2022).
 35. S. Henson, K. Bisson, M. L. Hammond, A. Martin, C. Mouw, A. Yool, Effect of sampling bias on global estimates of ocean carbon export. *Environ. Res. Lett.* **19**, 024009 (2024).
 36. J. Yun, J. Liu, K. W. Bowman, L. Resplandy, D. Carroll, The potential of the orbiting carbon observatory-2 column CO₂ measurements to constrain air-sea CO₂ fluxes. *ESS Open Arch. eprints* **389**, 172252933.38918992 (2024).
 37. P. Landschützer, N. Gruber, D. C. E. Bakker, U. Schuster, S. Nakaoka, M. R. Payne, T. P. Sasse, J. Zeng, A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences* **10**, 7793–7815 (2013).
 38. Y. Wei, S. T. Gille, M. R. Mazloff, V. Tamsitt, S. Swart, D. Chen, L. Newman, Optimizing mooring placement to constrain Southern Ocean air-sea fluxes. *J. Atmos. Oceanic Tech.* **37**, 1365–1385 (2020).
 39. T. Ilyina, K. D. Six, J. Segsneider, E. Maier-Reimer, H. Li, I. Núñez-Riboni, Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations. *J. Adv. Model. Earth Syst.* **5**, 287–315 (2013).
 40. T. Mauritsen, J. Bader, T. Becker, J. Behrens, M. Bittner, R. Brokopf, V. Brovkin, M. Claussen, T. Crueger, M. Esch, I. Fast, S. Fiedler, D. Fläschner, V. Gayler, M. Giorgetta, D. S. Goll, H. Haak, S. Hagemann, C. Hedemann, C. Hohenegger, T. Ilyina, T. Jahns, D. Jimenez-de-la-Cuesta, J. Jungclaus, T. Kleinert, S. Kloster, D. Kracher, S. Kinne, D. Kleberg, G. Lasslop, L. Kornbluh, J. Marotzke, D. Matei, K. Meraner, U. Mikolajewicz, K. Modali, B. Möbis, W. A. Müller, J. E. M. S. Nabel, C. C. W. Nam, D. Notz, S.-S. Nyawira, H. Paulsen, K. Peters, R. Pincus, H. Pohlmann, J. Pongratz, M. Popp, T. J. Raddatz, S. Rast, R. Redler, C. H. Reick, T. Rohrschneider, V. Schemann, H. Schmidt, R. Schnur, U. Schulzweida, K. D. Six, L. Stein, I. Stemmler, B. Stevens, J.-S. von Storch, F. Tian, A. Voigt, P. Vrese, K.-H. Wieners, S. Wilkenskjaeld, A. Winkler, E. Roeckner, Developments in the

- MPI-M Earth system model version 1.2 (MPI-ESM1.2) and its response to increasing CO₂. *J. Adv. Model. Earth Syst.* **11**, 998–1038 (2019).
41. H. Paulsen, T. Ilyina, K. D. Six, I. Stemmler, Incorporating a prognostic representation of marine nitrogen fixers into the global ocean biogeochemical model HAMOCC. *J. Adv. Model. Earth Syst.* **9**, 438–464 (2017).
 42. P. Landschützer, G. G. Laruelle, A. Roobaert, P. Regnier, A uniform pCO₂ climatology combining open and coastal oceans. *Earth Syst. Sci. Data* **12**, 2537–2553 (2020).
 43. A. Bhattacharyya, On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.* **35**, 99–109 (1943).
 44. A. Roobaert, G. G. Laruelle, P. Landschützer, N. Gruber, L. Chou, P. Regnier, The spatiotemporal dynamics of the sources and sinks of CO₂ in the global coastal ocean. *Global Biogeochem. Cycles* **33**, 1693–1714 (2019).
 45. A. Roobaert, P. Regnier, P. Landschützer, G. G. Laruelle, A novel sea surface pCO₂-product for the global coastal ocean resolving trends over 1982–2020. *Earth Syst. Sci. Data* **16**, 421–441 (2024).
 46. T. DeVries, K. Yamamoto, R. Wanninkhof, N. Gruber, J. Hauck, J. D. Müller, L. Bopp, D. Carroll, B. Carter, T.-T.-T. Chau, S. C. Doney, M. Gehlen, L. Gloege, L. Gregor, S. Benson, J. H. Kim, Y. Iida, T. Ilyina, P. Landschützer, C. Le Quéré, D. Munro, C. Nissen, L. Patara, F. F. Pérez, L. Resplandy, K. B. Rodgers, J. Schwinger, R. Séférian, V. Sicardi, J. Terhaar, J. Triñanes, H. Tsujino, A. Watson, S. Yasunaka, J. Zeng, Magnitude, trends, and variability of the global ocean carbon sink from 1985 to 2018. *Glob. Biogeochem. Cycles* **37**, e2023GB007780 (2023).
 47. J. Hauck, M. Zeising, C. Le Quéré, N. Gruber, D. C. E. Bakker, L. Bopp, T. T. T. Chau, Ö. Gürses, T. Ilyina, P. Landschützer, A. Lenton, L. Resplandy, C. Rödenbeck, J. Schwinger, R. Séférian, Consistency and challenges in the ocean carbon sink estimate for the global carbon budget. *Front. Mar. Sci.* **7**, 571720 (2020).
 48. N. Mayot, C. Le Quéré, C. Rödenbeck, R. Bernardello, L. Bopp, L. M. Djeutchouang, M. Gehlen, L. Gregor, N. Gruber, J. Hauck, Y. Iida, T. Ilyina, R. F. Keeling, P. Landschützer, A. C. Manning, L. Patara, L. Resplandy, J. Schwinger, R. Séférian, A. J. Watson, R. M. Wright, J. Zeng, Climate-driven variability of the Southern Ocean CO₂ sink. *Philos. Trans. R. Soc. A* **381**, 20220055 (2023).
 49. R. Wanninkhof, Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res. Oceans* **97**, 7373–7382 (1992).
 50. R. Arruda, D. Atamanchuk, M. Cronin, T. Steinhoff, D. Wallace, At-sea intercomparison of three underway pCO₂ systems: Intercomparison of pCO₂ systems. *Limnol. Oceanogr. Methods* **18**, 63–76 (2020).
 51. L. Olivier, J. Boutin, G. Reverdin, N. Lefèvre, P. Landschützer, S. Speich, J. Karstensen, M. Labaste, C. Noisel, M. Ritschel, T. Steinhoff, R. Wanninkhof, Wintertime process study of the North Brazil Current rings reveals the region as a larger sink for CO₂ than expected. *Biogeosciences* **19**, 2969–2988 (2022).
 52. T. H. Heimdal, G. A. McKinley, The importance of adding unbiased Argo observations to the ocean carbon observing system. *Sci. Rep.* **14**, 19763 (2024).
 53. A. J. Watson, U. Schuster, J. D. Shutler, T. Holding, I. G. C. Ashton, P. Landschützer, D. K. Woolf, L. Goddijn-Murphy, Revised estimates of ocean-atmosphere CO₂ flux are consistent with ocean carbon inventory. *Nat. Commun.* **11**, 4422 (2020).
 54. A. R. Fay, T. H. Heimdal, V. Acquaviva, P. A. Shaum, G. A. McKinley, Sensitivity of ocean carbon sink estimates to rare observations. *Geophys. Res. Lett.* **52**, e2025GL117961 (2025).
 55. D. Hohensee, “Vergleich von pCO₂ Sensoren für die Oberfläche des Ozeans - Eine Wirtschaftlichkeitsuntersuchung,” thesis, Christian-Albrechts-Universität zu Kiel, Germany (2017).
 56. B. R. Carter, N. L. Williams, W. Evans, A. J. Fassbender, L. Barbero, C. Hauri, R. A. Feely, A. J. Sutton, Time of detection as a metric for prioritizing between climate observation quality, frequency, and duration. *Geophys. Res. Lett.* **46**, 3853–3861 (2019).
 57. A. R. Gray, The four-dimensional carbon cycle of the Southern Ocean. *Ann. Rev. Mar. Sci.* **16**, 163–190 (2024).
 58. A. E. R. Hassoun, T. Tanhua, I. Lips, E. Heslop, G. Petihakis, J. Karstensen, The European ocean observing community: Urgent gaps and recommendations to implement during the UN ocean decade. *Front. Mar. Sci.* **11**, 1394984 (2024).
 59. C. Meinig, E. F. Burger, N. Cohen, E. D. Cokelet, M. F. Cronin, J. N. Cross, S. de Halleux, R. Jenkins, A. T. Jessup, C. W. Mordy, N. Lawrence-Slavas, A. J. Sutton, D. Zhang, C. Zhang, Public-private partnerships to advance regional ocean-observing capabilities: A saildrone and NOAA-PMEL case study and future considerations to expand to global scale observing. *Front. Mar. Sci.* **6**, 448 (2019).
 60. D. Roemmich, M. H. Alford, H. Claude, K. Johnson, B. King, J. Moum, P. Oke, W. B. Owens, S. Pouliquen, S. Purkey, M. Scanderbeg, T. Suga, S. Wijffels, N. Zilberman, D. Bakker, M. Baringer, M. Belbeoch, H. C. Bittig, E. Boss, P. Caili, F. Carse, T. Carval, F. Chai, D. Ó. Conchubhair, F. d’Ortenzio, G. Dall’Omo, D. Desbruyeres, K. Fennel, I. Fer, R. Ferrari, G. Forget, H. Freeland, T. Fujiki, M. Gehlen, B. Greenan, R. Hallberg, T. Hibiya, S. Hosoda, S. Jayne, M. Jochum, G. C. Johnson, K. Kang, N. Kolodziejczyk, A. Körtzinger, P.-Y. L. Traon, Y.-D. Lenn, G. Maze, K. A. Mork, T. Morris, T. Nagai, J. Nash, A. N. Garabato, A. Olsen, R. R. Pattabhi, S. Prakash, S. Riser, C. Schmechtig, C. Schmid, E. Shroyer, A. Sterl, P. Sutton, L. Talley, T. Tanhua, V. Thierry, S. Thomalla, J. Toole, A. Troisi, T. W. Trull, J. Turton, P. J. Velez-Belchi, W. Walczowski, H. Wang, R. Wanninkhof, A. F. Waterhouse, S. Waterman, A. Watson, C. Wilson, A. P. S. Wong, J. Xu, I. Yasuda, On the future of Argo: A global, full-depth, multi-disciplinary array. *Front. Mar. Sci.* **6**, 439 (2019).
 61. C. Whitt, J. Pearlman, B. Polagye, F. Caimi, F. Muller-Karger, A. Copping, H. Spence, S. Madhusudhana, W. Kirkwood, L. Grosjean, B. M. Fiaz, S. Singh, S. Singh, D. Manalang, A. S. Gupta, A. Maguer, J. J. H. Buck, A. Marouchos, M. A. Atmanand, R. Venkatesan, V. Narayanaswamy, P. Testor, E. Douglas, S. de Halleux, S. J. Khalsa, Future vision for autonomous ocean observations. *Front. Mar. Sci.* **7**, 697 (2020).
 62. N. P. Mongwe, M. Vichi, P. M. S. Monteiro, The seasonal cycle of pCO₂ and CO₂ fluxes in the Southern Ocean: Diagnosing anomalies in CMIP5 Earth system models. *Biogeosciences* **15**, 2851–2872 (2018).
 63. Y. Silvy, T. L. Frölicher, J. Terhaar, F. Joos, F. A. Burger, F. Lacroix, M. Allen, R. Bernardello, L. Bopp, V. Brovkin, J. R. Buzan, P. Cadule, M. Dix, J. Dunne, P. Friedlingstein, G. Georgievski, T. Hajima, S. Jenkins, M. Kawamiya, N. Y. Kiang, V. Lapin, D. Lee, P. Lerner, N. Mengis, E. A. Monteiro, D. Paynter, G. P. Peters, A. Romanou, J. Schwinger, S. Sparrow, E. Stofferahn, T. Tjiputra, E. Tourigny, T. Ziehn, AERA-MIP: Emission pathways, remaining budgets and carbon cycle dynamics compatible with 1.5°C and 2°C global warming stabilization. *Earth Syst. Dynam.* **15**, 1591–1628 (2024).
 64. K. D. Six, E. Maier-Reimer, Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model. *Global Biogeochem. Cycles* **10**, 559–583 (1996).
 65. C. Heinze, E. Maier-Reimer, A. M. E. Winguth, D. Archer, A global oceanic sediment model for long-term climate studies. *Global Biogeochem. Cycles* **13**, 221–250 (1999).
 66. H. Li, T. Ilyina, T. Loughran, A. Spring, J. Pongratz, Reconstructions and predictions of the global carbon budget with an emission-driven Earth system model. *Earth Syst. Dynam.* **14**, 101–119 (2023).
 67. B. Liu, K. D. Six, T. Ilyina, Incorporating the stable carbon isotope ¹³C in the ocean biogeochemical component of the Max Planck Institute Earth System model. *Biogeosciences* **18**, 4389–4429 (2021).
 68. J. Maerz, K. D. Six, I. Stemmler, S. Ahmerkamp, T. Ilyina, Microstructure and composition of marine aggregates as co-determinants for vertical particulate organic carbon transfer in the global ocean. *Biogeosciences* **17**, 1765–1803 (2020).
 69. W. A. Müller, J. H. Jungclaus, T. Mauritsen, J. Baehr, M. Bittner, R. Budich, F. Bunzel, M. Esch, R. Ghosh, H. Haak, T. Ilyina, T. Kleine, L. Kornbluh, H. Li, K. Modali, D. Notz, H. Pohlmann, E. Roeckner, I. Stemmler, F. Tian, J. Marotzke, A higher-resolution version of the Max Planck Institute Earth System model (MPI-ESM1.2-HR). *J. Adv. Model. Earth Syst.* **10**, 1383–1413 (2018).
 70. D. M. Nielsen, F. Cegini, J. Maerz, S. Brune, M. Mathis, M. Dobrynin, J. Baehr, V. Brovkin, T. Ilyina, Reduced arctic ocean CO₂ uptake due to coastal permafrost erosion. *Nat. Clim. Chang.* **14**, 968–975 (2024).
 71. A. Jersild, SOMFFN, Version 2022 (2023); <https://github.com/aljersild/SOMFFNv2022>.
 72. B. Nababan, D. Ulfah, J. P. Panjaitan, Light propagation, coefficient attenuation, and the depth of one optical depth in different water types. *IOP Conf. Ser. Earth Environ. Sci.* **944**, 012047 (2021).
 73. Global Carbon Project, Supplemental Data of Global Carbon Budget 2022, Version 1.0 (2022); <https://doi.org/10.18160/gcp-2022>.
 74. T. Naegler, Reconciliation of excess ¹⁴C-constrained global CO₂ piston velocity estimates. *Tellus B* **61**, 372–384 (2022).
 75. P. J. Duke, R. C. Hamme, D. Ianson, P. Landschützer, N. C. Swart, P. A. Covert, High-resolution neural network demonstrates strong CO₂ source-sink juxtaposition in the coastal zone. *J. Geophys. Res. Oceans* **129**, e2024JC021134 (2024).
 76. Y.-H. Park, T. Park, T.-W. Kim, S.-H. Lee, C.-S. Hong, J.-H. Lee, M.-H. Rio, M.-I. Pujol, M. Ballarotta, I. Durand, C. Provost, Observations of the antarctic circumpolar current over the udintsev fracture zone, the narrowest choke point in the Southern Ocean. *J. Geophys. Res. Oceans* **124**, 4511–4528 (2019).

Acknowledgments: We thank the sailors, particularly B. Herrmann and F. Amedeo, and their teams, Team *Malizia* and *Nexans-Wewise*, for the dedicated involvement in collecting valuable data. The SOCAT is an international effort, endorsed by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS), and the Integrated Marine Biosphere Research (IMBER) program, to deliver a uniformly quality-controlled surface ocean CO₂ database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for contributions to SOCAT. **Funding:** Research Foundation Flanders (FWO) contract I001821N (VLIZ ICOS; P.L.). Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC 2037 “CLICCS—Climate, Climatic Change, and Society”—Project Number 390683824 (T.I.). EU Horizon 2020 research and innovation programme under grant agreement no. 101003536 (ESM2025—Earth System Models for the Future) (T.I.). EU Horizon Europe research and innovation programme under grant agreement no. 101188028 (TRICUSO) (P.L. and J.B.). International Max Planck Research School on Earth System Modelling

(IMPRS-ESM) (J.B.). EU Horizon Europe research and innovation programme under grant agreement no. 101137682 (AI4PEX) (P.L. and T.I.). **Author contributions:** Conceptualization: P.L., J.B., and T.I. Methodology: P.L., J.B., and T.I. Investigation: J.B. Visualization: J.B. Validation: F.C. and T.I. Software: F.C., J.B., and T.I. Supervision: P.L. and T.I. Resources: P.L. Funding acquisition: P.L. and T.I. Project administration: P.L. and T.I. Formal analysis: J.B. and T.I. Writing—original draft: J.B. and T.I. Writing—review and editing: J.B., P.L., T.I., and F.C. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The MPIOM-HAMOCC model output can be found in (73). Results from

the neural-network based SOM-FFN method are publicly available online (10.5281/zenodo.15131072). All data and code needed to evaluate and reproduce the results in the paper are present in the paper and/or the Supplementary Materials.

Submitted 20 May 2025
Accepted 9 December 2025
Published 9 January 2026
10.1126/sciadv.adz1502