communications earth & environment

Article

https://doi.org/10.1038/s43247-024-01566-6

The Southern Ocean carbon sink has been overestimated in the past three decades

Guorong Zhong[1,2,3,4], Xuegang Li[1,2,3,4] ✉, Jinming Song[1,2,3,4] ✉, Fan Wang[1,3,4], Baoxiao Qu[1,3,4], Yanjun Wang[1,4], Bin Zhang[1,4], Jun Ma[1,3,4], Huamao Yuan[1,2,3,4], Liqin Duan[1,2,3,4], Qidong Wang[1,3,4], Jianwei Xing[1,3,4] & Jiajia Dai[1,3,4]

Employing machine learning methods for mapping surface ocean pCO_2 has reduced the uncertainty in estimating sea-air CO_2 flux. However, a general discrepancy exists between the Southern Ocean carbon sinks derived from pCO_2 products and those from biogeochemistry models. Here, by performing a boosting ensemble learning feed-forward neural networks method, we have identified an underestimation of the surface Southern Ocean pCO_2 due to notably uneven density of pCO_2 measurements between summer and winter, which resulted in about 16% overestimating of Southern Ocean carbon sink over the past three decades. In particular, the Southern Ocean carbon sink since 2010 was notably overestimated by approximately 29%. This overestimation can be mitigated by a winter correction in algorithms, with the average Southern Ocean carbon sink during 1992-2021 corrected to -0.87 PgC yr^{-1} from the original -1.01 PgC yr^{-1}. Furthermore, the most notable underestimation of surface ocean pCO_2 mainly occurred in regions south of 60°S and was hiding under ice cover. As the surface ocean pCO_2 under sea ice coverage in the winter is much higher than the atmosphere, if sea ice melts completely, there could be a further reduction of about 0.14 PgC yr^{-1} in the Southern Ocean carbon sink.

The increasing concentration of atmospheric CO_2 since the onset of the industrial era has been affecting the natural climate due to the greenhouse effect. This effect is partially mitigated by the global ocean CO_2 uptakes, which account for about one-quarter of the anthropogenic CO_2 emissions[1]. Natural climate variability and anthropogenic climate change also feedback to influence the sea–air CO_2 exchange[2]. It is essential to quantify the global ocean carbon sink and its temporal variability to understand further the response of the carbon cycle to future global change. The surface ocean partial pressure of CO_2 (pCO_2) measurements from the SOCAT dataset were widely used and mapped into continuous gridded data to estimate the sea–air CO_2 flux[3]. Due to a lower spatial decorrelation length scale of hundreds of kilometers in the surface ocean than that of thousands of kilometers in the atmosphere[4], surface ocean pCO_2 has more notable spatial variability than atmospheric pCO_2. Considerable variability and sparse measurements of surface ocean pCO_2 indicate insufficient observations to estimate CO_2 flux in most ocean areas directly. Great uncertainty in carbon sink estimation arises from sparse and uneven pCO_2 measurements, the gas transfer velocity, and the cool skin effect[5–7]. Recent application of machine learning algorithms in pCO_2 mapping methods increased data availability and further reduced the uncertainty in pCO_2-based carbon sink estimates[8–11]. The average net global ocean carbon sink during the last three decades was documented as -1.40 to -2.45 PgC yr^{-1}(see refs. 7,12–15). The differences between results were caused by differences in algorithms, division of global biogeochemical provinces, and selection of pCO_2 predictors. The accuracy of pCO_2 mapping based on machine learning methods remains to be improved, especially in polar regions with sparser pCO_2 measurements.

The Southern Ocean south of 35°S was a strong carbon sink and has contributed to about 40% of global ocean anthropogenic CO_2 uptakes from 1870 to 1995[1,16–18]. Changes in the Southern Ocean carbon sink strongly affect the global ocean CO_2 uptake. However, the Southern Ocean carbon sink estimated by pCO_2-based machine learning methods was about 0.4 PgC yr^{-1} stronger than the result from global ocean biogeochemical models since 2012[19,20]. A notable seasonal variability of surface ocean pCO_2 was reported in the Southern Ocean, mainly south of 50°S, with high pCO_2 levels and carbon sources observed in winter[13,21]. The strongly seasonally uneven surface ocean pCO_2 measurements with missing winter observations may result in an overestimation of the Southern Ocean carbon sink from pCO_2 products compared to the in situ observations[2,19,22,23]. Besides supplying more measurements from sailboats or floats[24], whether the

[1]Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China. [2]University of Chinese Academy of Sciences, Beijing, China. [3]Laoshan Laboratory, Qingdao, China. [4]Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China. ✉e-mail: lixuegang@qdio.ac.cn; jmsong@qdio.ac.cn

overestimation in pCO_2-based machine learning methods compared to the in situ observations can be solved by improving algorithms is worth investigating. Thus, we reestimated the Southern Ocean carbon sink using a boosting ensemble learning feed-forward neural networks (BEL FFNNs) method and investigated the influence of seasonally uneven SOCAT pCO_2 measurements on the pCO_2 mapping and carbon sink estimate.

Results

Influence of uneven measurements on the Southern Ocean pCO_2 mapping

Various machine learning methods were applied in the surface ocean pCO_2 mapping and CO_2 flux estimating[8,9,11,13,14,25], where a majority of methods are based on non-linear relationship fitting between SOCAT fCO_2 measurements, or converted pCO_2, and environmental variables[3]. However, the number of SOCAT measurements is uneven between different seasons in the Southern Ocean. The SOCAT measurements in the Southern Ocean are concentrated in summer with low surface ocean pCO_2 (Supplementary Fig. 1a), with the number of high-pCO_2 winter measurements only about one-fifth of that in summer. In most Southern Ocean areas, the SOCAT measurements covered less than four unique months from 1992 to 2021 (Supplementary Fig. 1b). Meanwhile, high surface ocean pCO_2 was observed during the winter in the Southern Ocean, according to the research based on either pCO_2 measurements or gridded products. The lack of measurements during high surface ocean pCO_2 seasons in most Southern Ocean areas may notably influence the non-linear relationship fitting and pCO_2 mapping, particularly in the months from June to September with the sparsest measurements. The seasonally unbalanced distribution of measurements may be a potential reason for the higher pCO_2 predicting error of the Southern Ocean than the neighboring areas in previous research[9,11,13].

To evaluate the influence of seasonal-uneven SOCAT measurements on pCO_2 mapping, the RMSE and bias from May to September were compared between different validation groups, with the only difference in training strategy (Table 1). The comparison of pCO_2 predicting error between different training strategies reveals a substantial influence of uneven measurements on the accuracy of machine learning pCO_2 predicting method. Training neural networks with SOCAT measurements only from April to October instead of all months resulted in a notable decrease of 0.5–1.4 μatm in RMSE. This decrease in RMSE caused by the change of training strategies was even more effective than the decrease of less than 0.4 μatm caused by the improvement of the FFNN structure (BEL FFNNs in Table 1). This is because the quantity of pCO_2 measurements is notably imbalanced among different seasons, with data in winter much less than in summer. T When training with such seasonal imbalanced data, the neural network tends to perform better in data-rich summer, while the performance is notably worse in data-sparse winter. This effect of imbalanced data can be mitigated through re-balancing data distribution and using re-balancing design in the loss function or learning algorithm of neural networks[26]. Training the neural networks with partial winter measurements is a data distribution re-balancing method, as the number of measurements is less unbalanced after the data-rich summer was removed from the training set. As a result, in the areas south of 50°S, the BEL FFNNs and the individual FFNN trained with sectional winter measurements suggested a notably lower RMSE during winter than those trained with all-month measurements. The bias during May–September between predicted pCO_2 and SOCAT measurements of more than −3 μatm was notably different from the range of −1 to ~1 μatm in other months, indicating a notable underestimation of surface seawater pCO_2 in the areas south of 50°S. In particular, the pCO_2 from May to September in the area south of 60°S, as predicted by BEL FFNNs using measurements from all months, was underestimated by an average of 5.77 μatm. Meanwhile, the pCO_2 predicted by the individual FFNN in the Southern Ocean south of 60°S was also notably lower than SOCAT data in winter due to missing winter measurements. When training with measurements only from April to October, the BEL FFNNs with uncorrected predictors reached the lowest winter RMSE in south of 60°S, and the bias was only −1.38 μatm. Similarly, the predicted pCO_2 by the individual FFNN was only 1.47 μatm lower than measurements on average, indicating a notable improvement in the underestimation of pCO_2 during winter. In the 50–60°S region, training BEL FFNNs and the individual FFNN with sectional winter measurements can also reduce the predicting bias in winter. By training BEL FFNNs only with measurements from April to September, the winter RMSE in the 50–60°S area was the lowest of uncorrected predictors among different training strategies, with a bias of only −0.36 μatm.

In contrast, there is no notable underestimation of winter surface ocean pCO_2 in the Southern Ocean between 35 and 50°S. The RMSE by training BEL FFNNs with sectional winter measurements was larger than that by training with all measurements, and the improvement of bias was also not observed when using sectional winter measurements. On the other hand, corrected pCO_2 predictors may better reflect the drivers of surface ocean pCO_2 in the Southern Ocean during winter. By using the month as a predictor, the correction of pCO_2 predictors can also effectively mitigate the underestimation of winter pCO_2 in the Southern Ocean. Simply changing pCO_2 predictors without correcting the training period, the RMSE of BEL FFNNs with weighted predictors during winter in the 50–60°S region decreases to 11.49 μatm, and the bias reduces to −1.08 μatm compared to BEL FFNNs with original predictors (see predictors listed in Supplementary Table 1). The same decrease in RMSE was also observed in the areas south of 60°S. Using both correction methods simultaneously, the RMSE can be minimized to 10.93 μatm and 13.29 μatm in the 50–60°S and regions south of 60°S, respectively. The bias also fell within an acceptable range of −1 to 1 μatm, close to the bias level in other months without notable underestimation or overestimation. In addition, we also test the pCO_2 RMSE and bias of traditional regression methods for comparison, including a multiple linear regression (MLR) and a multiple non-linear regression (MNLR). As expected, the traditional regression methods are more susceptible to seasonal-uneven measurements, showing higher RMSE and more severe underestimation of Southern Ocean pCO_2. In particular, the MLR using measurements from all months resulted in a winter RMSE of 34.02 μatm in the region south of 60°S, and output pCO_2 values lower than the real measurements by an average of 17.29 μatm. This result is barely acceptable, and it also explains why traditional regression methods in previous research were generally limited to specific seasons. Although the MNLR performs better than the MLR, its RMSE was still much higher than that of the FFNN methods, and the MNLR output pCO_2 during winter was also notably lower than measurements in regions south of 50°S. Similar to the treatment of the training period of the two FFNN methods, a lesser underestimation of Southern Ocean pCO_2 in winter was found when using only partial winter measurements for regression. In the 35–50°S region, the RMSE of traditional regression methods was still higher than the two FFNN methods, and the influence of seasonal-uneven measurements was not remarkable.

The BEL FFNNs and individual FFNN both showed an underestimated surface seawater pCO_2 in the Southern Ocean south of 50°S in winter, indicating that the underestimation of surface seawater pCO_2 was not caused by the structure of FFNN but rather by the seasonally uneven pCO_2 measurements. Training networks with sectional winter measurements and correction of pCO_2 predictors can mitigate the underestimation of surface Southern Ocean pCO_2 in winter. Considering that the BEL FFNNs have a lower RMSE compared to the individual FFNN, the BEL FFNNs using corrected pCO_2 predictors and training with measurements only from April to October have better accuracy of the pCO_2 mapping in the Southern Ocean during winter.

With the correction of the training period and pCO_2 predictors, the bias of predicted pCO_2 from May to September was notably smaller than the uncorrected result (Fig. 1a). In July, the pCO_2 predicted by the BEL FFNNs was notably lower than SOCAT measurements, with a considerable bias of 6.6 μatm. In contrast, the bias from October to April was generally within the range of −1 to 1 μatm, indicating not remarkable overestimation or underestimation of surface seawater pCO_2 in the Southern Ocean. With the winter correction, the bias from May to September decreased notably to near zero. Even in the most biased July, the bias of corrected BEL FFNNs fell to

Table 1 | Comparison of pCO_2 predicting error in the Southern Ocean during May–September among different methods and regression periods

Regression method	Regression period	35–50°S		50–60°S		S of 60°S	
		RMSE (μatm)	Bias (μatm)	RMSE (μatm)	Bias (μatm)	RMSE (μatm)	Bias (μatm)
BEL FFNNs	May–Sep	12.01	+0.61	11.97	+0.22	14.76	+0.33
	Apr–Sep	10.70	+0.38	11.88	−0.36	14.87	−0.90
	May–Oct	11.39	−0.36	12.07	−0.66	14.50	−0.07
	Apr–Oct	11.10	−0.68	12.09	−1.16	14.20	−1.38
	Mar–Nov	11.32	−1.26	12.51	−2.13	15.50	−4.47
	All months	**9.86**	**−0.39**	13.27	−3.13	16.09	−5.77
BEL FFNNs with corrected predictors	May–Sep	12.24	+0.74	11.93	+0.17	14.44	+1.17
	Apr–Sep	10.76	+0.30	11.26	+0.09	13.49	−0.06
	May–Oct	12.81	−0.28	11.44	+0.04	14.27	−0.20
	Apr–Oct	11.83	−0.79	**10.93**	**−0.25**	**13.29**	**−0.74**
	Mar–Nov	11.69	−1.49	13.07	−0.46	14.07	+0.02
	All months	9.88	−0.29	11.49	−1.08	14.61	−0.90
Individual FFNN	May–Sep	12.05	+0.50	11.99	+0.07	15.15	0.36
	Apr–Sep	10.58	+0.30	12.02	−0.52	15.04	−1.14
	May–Oct	11.43	−0.59	12.21	−0.75	14.92	−0.32
	Apr–Oct	11.19	−0.83	12.29	−1.49	14.86	−1.47
	Mar–Nov	11.64	−1.43	12.82	−2.56	15.76	−3.95
	All months	9.95	−0.34	13.31	−3.33	17.14	−5.38
Individual FFNN with corrected predictors	May–Sep	12.57	+0.50	12.30	+0.37	14.54	+0.83
	Apr–Sep	10.72	+0.27	11.45	+0.14	13.79	−0.27
	May–Oct	13.09	−0.45	11.52	−0.07	14.46	−0.23
	Apr–Oct	11.35	−0.92	11.06	−0.09	13.63	−0.27
	Mar–Nov	12.02	−1.55	11.55	−0.09	17.86	+0.50
	All months	9.95	−0.28	12.06	−1.04	15.85	−0.05
MLR with corrected predictors	May–Sep	16.34	−2.33	13.67	−1.71	21.12	+1.09
	Apr–Sep	15.48	−0.67	14.51	−2.29	22.21	−3.00
	May–Oct	16.98	−4.55	16.29	−4.83	23.46	−0.89
	Apr–Oct	15.87	−2.93	16.39	−5.56	24.79	−6.69
	Mar–Nov	16.09	−2.53	16.75	−6.22	29.21	−13.40
	All months	17.70	−2.30	19.01	−6.11	34.02	−17.29
MNLR with corrected predictors	May–Sep	14.37	−0.83	12.17	−0.62	17.04	+2.41
	Apr–Sep	13.51	+0.67	11.20	−0.95	17.70	−2.86
	May–Oct	14.55	−4.06	13.18	−3.47	17.70	+2.41
	Apr–Oct	14.35	−3.45	13.53	−3.76	20.21	−6.64
	Mar–Nov	14.21	−1.72	12.72	−4.70	18.81	−8.01
	All months	15.02	+0.90	13.92	−5.35	25.68	−13.41

BEL FFNNs: boosting ensemble learning based on three FFNNs constructed in this work; Individual FFNN: one FFNN with the same structure used in the BEL; *MLR* multiple linear regression; *MNLR* multiple non-linear regression, see Supplementary Note 3; Regression Period: a period of SOCAT samples used for training neural networks or performing classic regression to predict pCO_2 during May–September; RMSE and bias: calculated from the difference between predicted pCO_2 and SOCAT measurements during May–September based on the K-fold cross-validation method. Corrected predictors: pCO_2 predictors selected by a stepwise BEL FFNNs algorithm based on increasing weightings of winter measurements, see Table 1. Bold numbers: the regression period and method with the lowest RMSE adopted in the final pCO_2 product construction.

only −0.3 μatm, substantially mitigating the underestimation of winter surface ocean pCO_2 in the Southern Ocean. The bias at different latitudes reveals that the underestimation of surface seawater pCO_2 in the Southern Ocean due to seasonally uneven measurements becomes more remarkable at higher latitudes (Fig. 1b). In the region south of 50°S, the uncorrected average deviation is negative, reaching approximately −20.92 μatm at around 68.5°S. The difference in the effect of the winter correction may be related to the density of measurements, as the decrease in bias was more notable in sparsely sampled high-latitude areas.

Compared to the observation from the Southern Ocean Flux station (SOFS, 142.0°E, 46.8°S)[27,28], the pCO_2 values from May to September from different methods were lower due to the lack of SOCAT winter data for training (Fig. 2). During years that winter SOCAT data are available, such as 2012, 2013, and 2018, the pCO_2 values from different methods were close to the observations from the SOFS time series station[8,9,13,14,29–31]. The surface ocean pCO_2 of BEL FFNNs product after correction in winter was about 10 μatm higher than the uncorrected BEL FFNNs results. It was much closer to the time series observation, suggesting a better accuracy of corrected BEL FFNNs pCO_2 than the uncorrected results. Both the validation based on the SOCAT dataset and the validation based on time-series observations from the SOFS station suggest that correction of the training period and pCO_2 predictors can effectively mitigate the underestimation due to seasonally

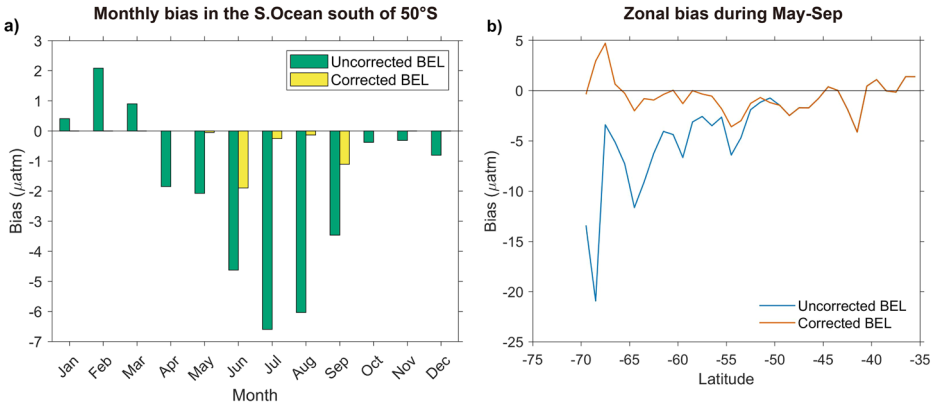

Fig. 1 | Distribution of bias between predicted pCO_2 and SOCAT measurements in the Southern Ocean south of 35°S. a Monthly bias in the Southern Ocean south of 50°S (predicted pCO_2 minus SOCAT measurements). **b** Zonal bias during May–Sep. Uncorrected BEL: based on training sample of all seasons. Corrected BEL: based on training sample only from April to October. BEL boosting ensemble learning FFNNs used in this study.

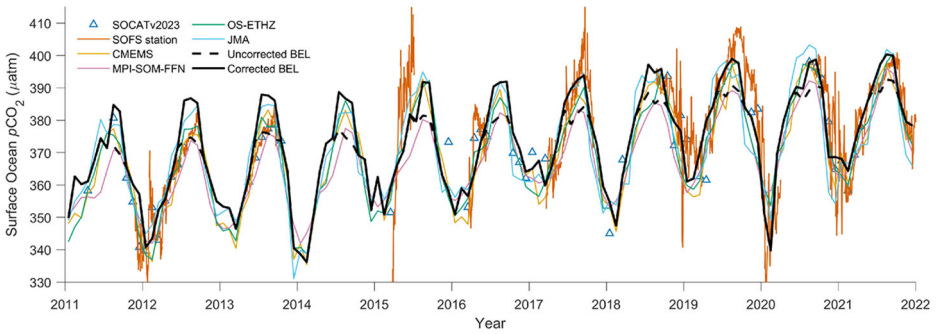

Fig. 2 | Comparison between corrected and uncorrected ensemble learning pCO_2 product in the SOFS station at 142.0°E, 46.8°S. CMEMS: the Copernicus Marine Environment Monitoring Service product from refs. 9,29. MPI-SOM-FFNN: the Self-Organizing Map-Feed-Forward Network product from refs. 13,30. OS-ETHZ: the Satellite Oceanographic Datasets for Acidification project product from refs. 8,31. JMA: the Japan Meteorological Agency product from ref. 14. Uncorrected BEL: boosting ensemble learning FFNNs based on training samples of all seasons. Corrected BEL: pCO_2 during May–September were predicted based on corrected predictors and training samples only from April to October.

uneven measurements. In addition, the evaluation based on the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) dataset also suggested that winter-corrected BEL FFNNs pCO_2 was better consistent with the in situ float data (Supplementary Fig. 2). Therefore, the final pCO_2 product constructed in this study consists of pCO_2 data from October to April based on all measurements and pCO_2 data from May to September based on corrected pCO_2 predictors and measurements only from April to October.

Overestimated Southern Ocean carbon sink due to biased pCO_2 mapping

The validation based on SOCAT measurements and SOFS time series observations reveals that the pCO_2 products constructed using FFNN with the entire monthly pCO_2 measurements from the SOCAT dataset may notably underestimate the winter surface ocean pCO_2 in the Southern Ocean south of 50°S. Due to upwelling derived from the wind driving a strong surface divergence through the Ekman transport[32], the surface seawater pCO_2 in the winter of the Southern Ocean is notably higher than in the summer, with strong carbon source regions in winter[13,25,32]. The previous studies may have underestimated the strength of carbon sources in the winter of the Southern Ocean, leading to an overestimation of the overall carbon sink intensity in the Southern Ocean. Our results demonstrate that the variations in surface seawater pCO_2 before and after the winter correction notably impact the pCO_2 difference and CO_2 flux across the air-sea interface (Fig. 3). The surface seawater pCO_2 in the Southern Ocean south of 50°S is higher than atmospheric pCO_2 from May to September, and the pCO_2 difference after correction became larger, particularly in the region south of 60°S. However, due to the influence of sea ice coverage, the pCO_2 flux in the area south of 60°S is nearly zero from May to September, and the difference in CO_2 flux between uncorrected and corrected BEL FFNNs was not remarkable. In the 35–50°S area with relatively more measurements, the seasonal variation pattern of pCO_2 differs from that south of 50°S, and there is almost no change in the pCO_2 difference and CO_2 flux between uncorrected and corrected BEL FFNNs. Overall, although the south of 60°S shows the most considerable change in winter ΔpCO_2 before and after correction, the underestimation of surface seawater pCO_2 in the 50–60°S area unaffected by sea ice coverage is the main reason for the overestimation of the carbon sink intensity in the Southern Ocean. The corrected average Southern Ocean carbon sink from May to September is -0.58 PgC yr^{-1}, decreasing by 0.34 PgC compared to the uncorrected results.

Over the past 30 years, the corrected average Southern Ocean carbon sink was -0.87 ± 0.16 PgC yr^{-1}, which is ~0.14 PgC yr^{-1} lower than before the correction, suggesting an overestimation of about 16%. The overestimation of the carbon sink intensity in the Southern Ocean is mainly observed after 2010, with a decrease in the decadal average carbon sink from -1.20 PgC yr^{-1} to -0.93 PgC yr^{-1} after correction. This indicates that the seasonally uneven measurements led to an overestimation of the Southern Ocean carbon sink by ~29% compared to the corrected intensity during this period (Fig. 4). Although the corrected Southern Ocean carbon sink was lower than uncorrected results in the 1990s, the variability pattern was similar before and after correction. Since 2001, the Southern Ocean carbon sink has generally strengthened, but the strengthening rate is relatively slower after the winter correction. The variability of the Southern Ocean carbon sink from our corrected BEL product was highly consistent with previous research based on models or observations, in which the Southern Ocean carbon sink receded substantially in the 1990s, reaching a trough at the beginning of the 21st century[33,34], and subsequently restrengthened to full intensity[18,35,36]. Compared to previous products, our estimation of the corrected Southern Ocean carbon sink shows a similar intensity in the 1990s and the lowest intensity since 2003. However, research based on SOCCOM buoy data also suggested a notably weaker Southern Ocean carbon sink, challenging existing results from pCO_2 products[2], although the float pCO_2 data calculated indirectly from pH and alkalinity seems to be overestimated in organic-rich freshwaters[37]. Notably, there was almost no difference

between the uncorrected and corrected carbon sink from 1999 to 2001, when the Southern Ocean carbon sink was at its weakest point in the past three decades. The relatively denser measurements in the SH winter than in other decades may be one important reason. Around 2000, the SOCAT winter measurements were close to half of the measurements in summer. Therefore, the influence of seasonally uneven measurements is relatively minor.

Both carbon sinks before and after the winter correction consistently show a rapid weakening of the Southern Ocean carbon sink during the 1990s. The corrected Southern Ocean carbon sink in this work weakened from -0.99 ± 0.15 Pg C yr^{-1} in 1992 to -0.68 ± 0.13 Pg C yr^{-1} in 2001, and then strengthened back to -1.13 ± 0.14 Pg C yr^{-1} until 2021. Such notable interannual changes were also found in other research based on observations covering the past two decades[8,9,13,38,39]. The contribution of the

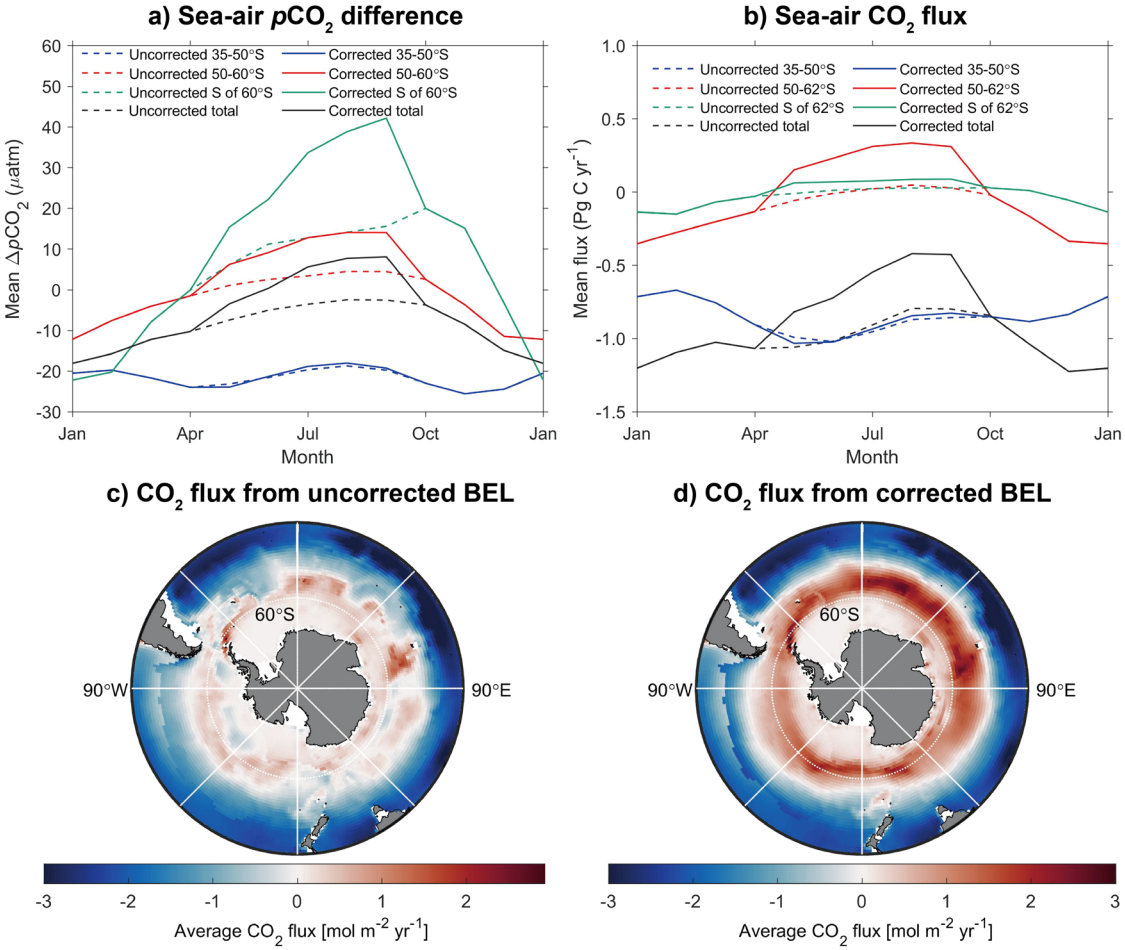

Fig. 3 | Distribution of average sea–air pCO_2 difference and CO_2 flux from May to September during 1992–2021. a Sea–air ΔpCO_2 in different regions: surface ocean pCO_2 minus atmospheric pCO_2. **b** Sea–air CO_2 flux in different regions. **c** distribution of CO_2 flux from uncorrected BEL FFNNs product. **d** distribution of CO_2 flux from corrected BEL FFNNs product. Uncorrected BEL FFNNs: constructed from SOCAT measurements of all months. Corrected BEL FFNNs: data from May to September constructed based on corrected predictors and only SOCAT measurements during Apr–Oct.

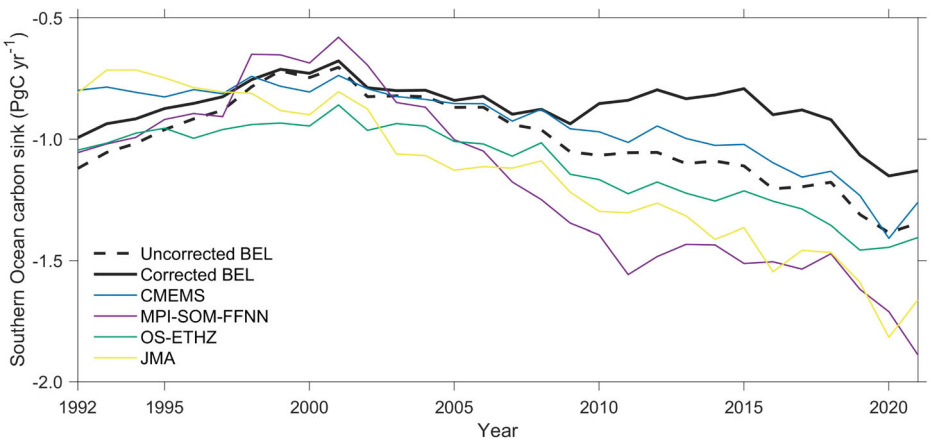

Fig. 4 | Interannual variability of the Southern Ocean carbon sink from uncorrected and corrected ensemble learning method and other products. CMEMS: the Copernicus Marine Environment Monitoring Service product from refs. 9,29. MPI-SOM-FFNN: the Self-Organizing Map-Feed-Forward Network product from refs. 13,30. OS-ETHZ: the Satellite Oceanographic Datasets for Acidification project product from refs. 8,31. JMA: the Japan Meteorological Agency product from ref. 14. Uncorrected BEL: boosting ensemble learning based on training samples of all seasons. Corrected BEL: pCO_2 during May–September were predicted based on training samples only from April to September.

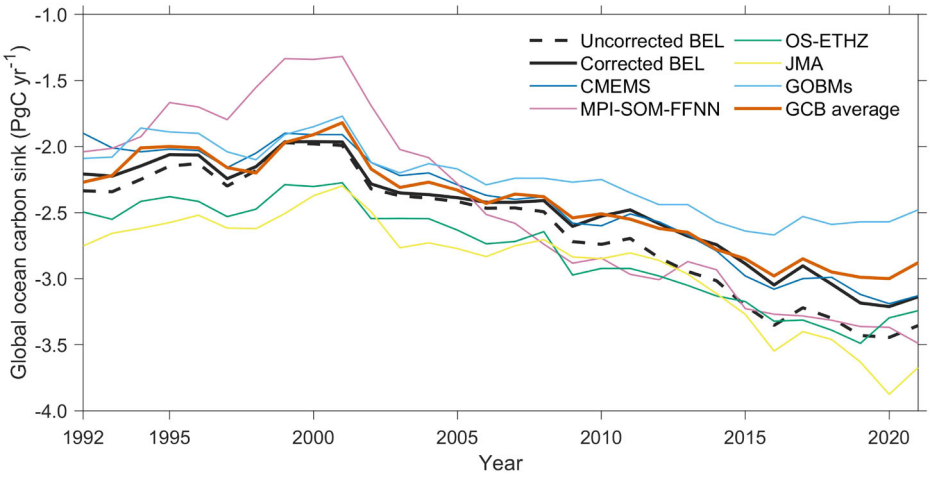

Fig. 5 | Global ocean carbon sink over the past three decades after the Southern Ocean correction. GOBMs: average results of 10 global ocean biogeochemical models (https://doi.org/10.18160/GCP-2022, ref. 19). GCB average: average results of 10 global ocean biogeochemical models and 7 pCO_2 products in the Global Carbon Budget 2022[19]. CMEMS: the Copernicus Marine Environment Monitoring Service product from refs. 9,29. MPI-SOM-FFNN: the Self-Organizing Map-Feed-Forward Network product from refs. 13,30. OS-ETHZ: the Satellite Oceanographic Datasets for Acidification project product from refs. 8,31. JMA: the Japan Meteorological Agency product from ref. 14. Uncorrected BEL: boosting ensemble learning FFNNs based on training sample of all seasons. Corrected BEL: Southern Ocean pCO_2 during May–September were corrected for uneven measurements.

Southern Ocean south of 35% on the global ocean CO_2 uptakes decreased from ~63% in 1992 to 45% in 2021. The weakening of the Southern Ocean carbon sink in the 1990s was thought to be caused by the strengthening of the upper-ocean overturning circulation and CO_2 release in source areas and the weakening of CO_2 uptake in sink areas due to a southward shift of westerlies[18,28,40–42]. The upwelling of the Southern Ocean increased by ~40% due to enhanced wind-driven circulation[43]. However, the weakening of the upwelling since the beginning of the 21st century led to the reinvigoration of carbon sink[18]. Research based on an idealized upper ocean box model also suggested that the slowed growth rate of atmospheric pCO_2 and the global sea surface temperature response to the 1991 eruption of Mt Pinatubo volcanic were two external forces to explain the global-scale reduction in the ocean carbon sink, and the reinvigoration of carbon sink was driven by the acceleration of atmospheric pCO_2 growth after 2001[44].

After the winter correction for seasonally uneven measurements in the Southern Ocean, the global ocean carbon sink estimated from the Stepwise FFNN product and corrected Southern Ocean pCO_2 data was relatively lower than other pCO_2 products (Fig. 5). However, our estimates are more consistent with the average results from the Global Carbon Budget study, based on 10 global ocean biogeochemistry models and 7 pCO_2 products[19]. The global ocean carbon sink estimated from previous pCO_2 products was notably stronger than the result from biogeochemical models, and the discrepancy mainly occurred in the Southern Ocean carbon[19]. The corrected Southern Ocean carbon sink decreased the discrepancy with model results, indicating that previous pCO_2 products using the SOCAT dataset may also experience an overestimation of the Southern Ocean carbon sink due to seasonally uneven measurements.

Influence of sea ice cover on the Southern Ocean carbon sink

The remarkable differences in the Southern Ocean carbon sink before and after correction were only observed in the last decade. However, the pCO_2 difference across the interface after the winter correction was much more notable, particularly in the Southern Ocean south of 60°S, where the surface seawater pCO_2 is much higher than atmospheric pCO_2 (Fig. 6). The vertical mixing and biological activity were reported as primary controlling factors of surface ocean pCO_2 in continent shelf areas, leading to more considerable uncertainty in CO_2 flux estimate compared to the open ocean[45–47]. However, CO_2 exchange between the seawater and the atmosphere in the Antarctic shelf is impeded due to the extensive sea ice coverage in most areas south of 60°S. As a result, despite the high surface seawater pCO_2 in this region, the CO_2 release to the atmosphere is limited, and the carbon source intensity is close to zero. The sea ice coverage in the Southern Ocean south of 60°S also eliminates the influence of seasonally uneven pCO_2 measurements. Therefore, although the pCO_2 difference across the interface was more remarkable after the winter correction in areas south of 60°S, the carbon source intensity and its difference before and after correction remain close to zero. Recent research has reported that the melting of sea ice in the Arctic Ocean exposes more sea surface, serving as one of the essential factors of rapid acidification in the Arctic Ocean[48]. Similarly, in the Amundsen and Bellingshausen Seas of the Southern Ocean, which are characterized by warm water intrusion from the open ocean, the highest basal ice shelf melting rates have been observed[49–51]. The Antarctic shelf ocean warming accelerated by increasing El Niño variability was hastening the ice shelf/sheet melt[52]. Unlike the Arctic Ocean, the surface ocean pCO_2 under sea ice coverage in the winter Southern Ocean was much higher than in the atmosphere. If the sea ice completely melts, a remarkable amount of CO_2 will be directly released into the atmosphere through the exposed sea surface. Furthermore, sea ice melting can indirectly impact the surface ocean pCO_2 and carbon sink intensity in the Southern Ocean through various pathways, such as reducing sea surface temperature and altering convective overturning rates[53].

Assuming complete sea ice melt and neglecting indirect factors based on the recent sea–air pCO_2 difference, the changes in simulated carbon sink intensity vary between summer and winter periods in the Southern Ocean (Fig. 7). During the summer period in the Southern Ocean, when sea ice coverage is limited, and the surface seawater pCO_2 in the covered areas is lower than atmospheric pCO_2, it is assumed that the complete melting of the currently covered sea ice would have little impact on the summer carbon sink intensity. However, during the winter, when sea ice coverage is extensive, and the surface seawater pCO_2 in the covered areas is much higher than atmospheric pCO_2, the complete melting of the currently covered sea ice would result in the release of CO_2 from the exposed surface Southern Ocean at an average rate of 0.28 PgC yr^{-1} during winter. This would weaken the role of the Southern Ocean in the global ocean CO_2 uptakes and the role of the global ocean in buffering the rise in atmospheric CO_2 concentration. In addition, unlike the relatively stable increasing trend in carbon sink intensity during summer over the past 20 years, the winter carbon sink intensity in the Southern Ocean has shown notable fluctuations in the last decade but with no remarkable long-term trend. Based on the current data, considering the presence of sea ice-covered carbon sink areas in spring and autumn, the complete disappearance of sea ice coverage would lead to an average reduction of 0.14 PgC yr^{-1} in the overall annual CO_2 absorption in the Southern Ocean. The magnitude of this reduction depends on the pCO_2 values of seawater covered by sea ice, and it is still uncertain how this will change in the future. However, it can be anticipated that sea ice melting will slow down the rate of carbon sink enhancement in the Southern Ocean for a considerable period until the continuously rising atmospheric pCO_2 exceeds the surface ocean pCO_2 beneath the winter sea ice.

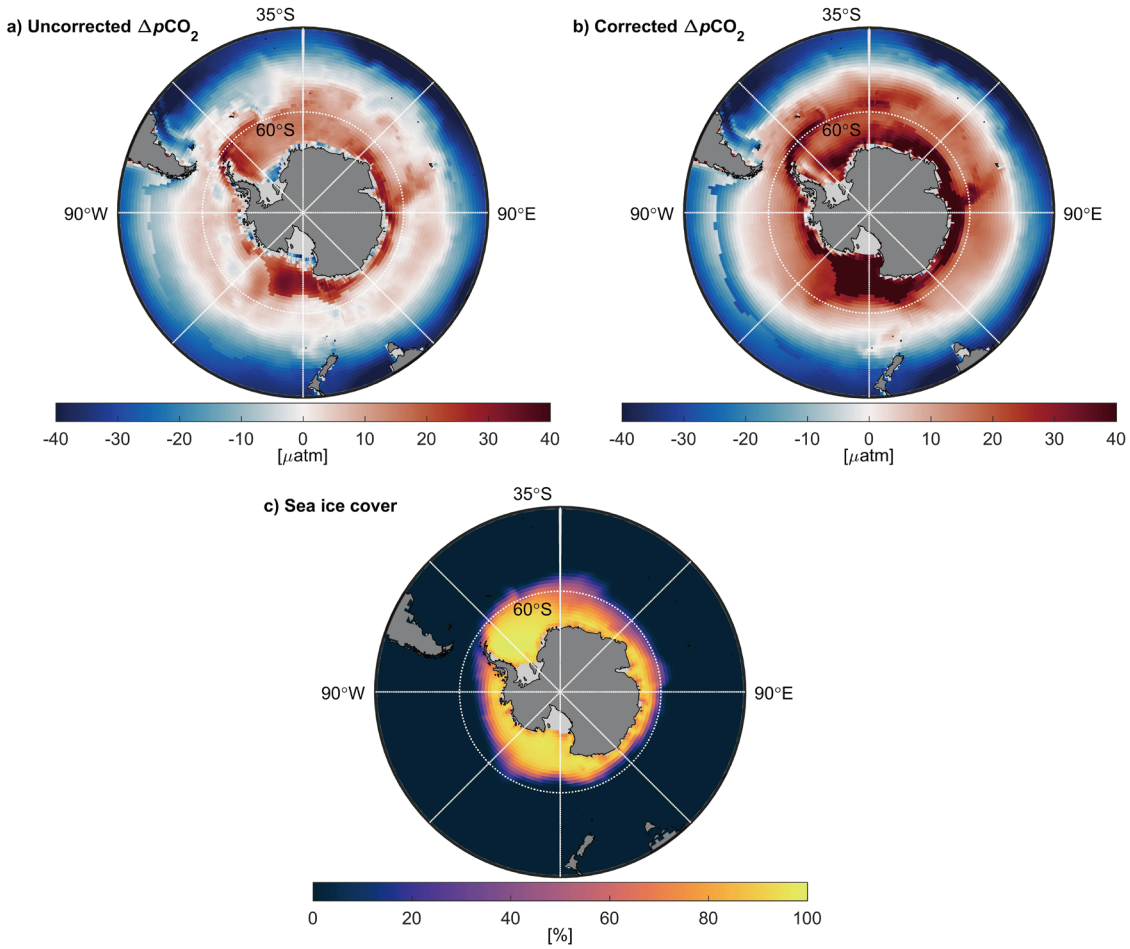

Fig. 6 | Distribution of average sea–air ΔpCO_2 and sea ice coverage during May–Sep in the Southern Ocean. a ΔpCO_2 calculated from uncorrected BEL product. **b** ΔpCO_2 calculated from corrected BEL product. **c** Sea ice coverage from ERA5 product[57,58]. ΔpCO_2 surface ocean pCO_2 minus atmospheric pCO_2.

Data and methods

pCO_2 mapping and winter correction

The surface ocean pCO_2 converted from the Surface Ocean CO_2 Atlas version 2023 (SOCAT v2023) dataset[3,54,55] was used for pCO_2 mapping by fitting the non-linear relationship between pCO_2 and environmental variables in Table 2. The SOCAT dataset includes quality-controlled global observations of in situ surface ocean fugacity of carbon dioxide (fCO_2), sea surface temperature, and salinity on ships, moorings, autonomous and drifting surface platforms for the global oceans and coastal seas from 1957 to 2023. This dataset is provided as a synthesis version and a gridded version, with an estimated fCO_2 accuracy of better than 5 μatm. The gridded fCO_2 was converted to pCO_2 using in situ sea surface temperature and atmospheric pressure[56], and then the converted pCO_2 was used in training neural networks:

$$pCO_2 = fCO_2 \cdot \exp\left(P_{\mathrm{atm}}^{\mathrm{surf}} \frac{B + 2 \cdot \delta}{R \cdot T}\right)^{-1} \quad (1)$$

where $P_{\mathrm{atm}}^{\mathrm{surf}}$ is the atmospheric pressure using ERA5 sea level pressure product[57,58], B and δ are viral coefficients calculated from sea surface temperature[59], R is the gas constant and T is the absolute temperature.

The relationship fitting was based on a boosting ensemble learning feed-forward neural networks (BEL FFNNs) consisting of three FFNNs. The first FFNN (FFNN I in Fig. 8) outputs will be used as a pCO_2 predictor in the second and the last FFNN (FFNN II and FFNN III in Fig. 8). Compared to shallower neural networks, neural networks with more hidden layers require far fewer neurons to achieve the same predicting error. Therefore, we used multiple hidden layers for all FFNNs with 10 neurons in each hidden layer, and adjusted the number of hidden layers to achieve an optimal FFNN size based on changes in predictor errors. The average of several FFNN outputs with changing initial states was taken as the final pCO_2 prediction value. The pCO_2 predictors used in this work were selected by the updated stepwise FFNN algorithm based on the variation of pCO_2 predicting error caused by each predictor[11], as described in the section "Selection and correction of pCO_2 predictors". A winter correction was carried out by changing pCO_2 predictors and the temporal period of measurements used for training networks.

To eliminate the influence of FFNN structure on correction validity, the performances of the individual FFNN and BEL FFNNs were compared under different training strategies: one based on SOCAT pCO_2 measurements of all months and others based on sectional winter measurements. The predicted pCO_2, root mean square error (RMSE), and bias with different training strategies and pCO_2 predictors were compared to evaluate the influence of seasonally uneven pCO_2 measurements and to determine which training strategy will be used. The predictor error was calculated using a K-fold cross-validation method, where the pCO_2 measurements were divided into four groups by year, and each one was predicted by the other three groups[11,60]. Then, the results were further compared to the observations from the SOFS station (142.0°E, 46.8°S) time series stations for validation[27,28]. The final pCO_2 product includes two types of data: (1) the pCO_2 from October to April based on SOCAT measurements of all seasons, and (2) the pCO_2 from May to September based on sectional SOCAT winter measurements and corrected predictors.

Selection and correction of pCO_2 predictors

The pCO_2 predictors input into the FFNN reflect the drivers of surface ocean pCO_2 and its variability. When changing the input pCO_2 predictors, both the FFNN predicted pCO_2 value and the predicting error notably change, and these changes can be even greater than those caused by altering the FFNN structure. However, the environmental factors driving pCO_2 and its variability notably differ among different regions. The surface ocean pCO_2 is largely affected by the upwelling and biological drawdown in the Antarctic region, and is affected by meridional overturning circulation in the subantarctic region[32,43]. This means that using different predictors in different latitude regions can better reflect the regional influencing factors of pCO_2 and its variability. To find the best combination of pCO_2 predictors in different regions, we have proposed a Stepwise FFNN algorithm in previous work, where the changes in predicting error are fed back to the selection of input pCO_2 predictors[11]. This algorithm allows for the objective selection of pCO_2 predictors in different regions that result in the lowest pCO_2 predicting error. The procedure of the Stepwise FFNN algorithm is determining pCO_2 predictors one by one until no further reduction in predicting error is achieved by either adding or removing any predictors. Specifically, the first pCO_2 predictor is selected by comparing predicting errors when individually using each collected environmental variable (listed in Supplementary Table 1) as input of the FFNN. The variable with the lowest error is determined as the first pCO_2 predictor, which is also the predictor that has the greatest impact on the distribution or variability of regional surface ocean pCO_2. Subsequently, leaving the first predictor unchanged, the predicting errors are compared when using each environmental variable as the second input of the FFNN. The environmental variable with the lowest error is determined as the second pCO_2 predictor. In the same way, new predictors are continuously determined one after another, until the predicting error no longer continues to decrease regardless of which one variable is added as a pCO_2 predictor. Meanwhile, whenever a new predictor is determined, the algorithm also tests if the predicting error will decrease when sequentially removing each determined predictor, in order to eliminate co-correlation and prevent overfitting. For example, when the fourth predictor is determined, the model tests the change in predicting error by individually removing each one from the previously determined three predictors. If the error decreases after removing a previously determined predictor, this predictor is highly correlated with other determined predictors. By adding and removing variables in the input of the FFNN one by one in this way, the algorithm ultimately identifies a set of pCO_2 predictors that minimize the pCO_2 predicting error. In this work, the single FFNN structure used in the previous Stepwise FFNN algorithm has been replaced with a structure of ensemble learning FFNNs with stronger fitting capabilities (see Fig. 8), referred to as the Stepwise BEL algorithm.

However, the Stepwise BEL algorithm relies on predicting errors for determination of pCO_2 predictors, and the number of SOCAT pCO_2 measurements in the Southern Ocean during the winter season is much

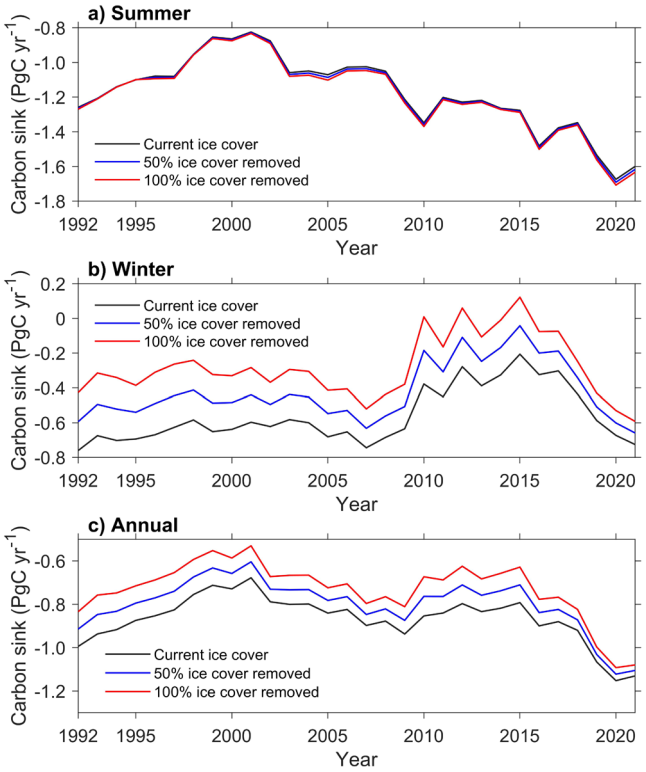

Fig. 7 | Southern Ocean carbon sink on different scenarios of sea ice melt. a CO_2 flux from December to February in each year. **b** CO_2 flux from June to August in each year. **c** Annual CO_2 flux. Current ice cover: ice coverage data from the ERA5 product[57,58]. 50% ice cover removed: assuming that 50% of current ice cover melts. In total, 100% ice cover removed: assuming that all ice cover melts.

Table 2 | Winter correction of pCO_2 predictors

Area		pCO_2 predictors
35–50°S	Pacific sector	SST, sin(Longitude), xCO_2, Latitude, SSS, Photosynthetically Available Radiation, Chlorophyll, Mixed layer depth, cos(Longitude), Mixed layer depth $_{anom}$, Remote sensing reflectance at 531 nm and 555 nm
	Indian sector	SST, Total absorption at 645 nm, Number of months since January 1992, Mixed layer depth, SSS, W velocity of ocean currents at 105 m, Surface pressure, Total absorption at 678 nm, W velocity of ocean currents at 195 m, Total backscattering at 667 nm, Nitrate, Total absorption at 555 nm, Mixed layer depth $_{anom}$, Particulate organic carbon, DIC, W velocity of ocean currents at 65 m, Remote sensing reflectance at 488 nm, Total backscattering at 443 nm
	Atlantic sector	Latitude, SSS, Dry air mixing ratio of atmospheric CO_2, Particulate organic carbon, Total backscattering at 488 nm, Mixed layer depth, Diffuse attenuation coefficient, Total backscattering at 412 nm, Sea surface height, cos(Longitude), SST, Remote sensing reflectance at 460 nm, Total backscattering at 547 nm, Bathymetry, Total absorption at 678 nm, Total backscattering at 469 nm, Remote sensing reflectance at 678 nm
50–60°S	Uncorrected	SSS_{anom}, SST, Mixed layer depth, Dry air mixing ratio of atmospheric $CO_{2\ anom}$, Bathymetry, Sea surface height $_{anom}$, W velocity of ocean currents at 105 m, DIC, Dissolved oxygen, Nitrate
50–60°S	Corrected	Dry air mixing ratio of atmospheric CO_2, Mixed layer depth, SST, DIC, month, SSS, Bathymetry, Latitude, W velocity of ocean currents at 105 m, Dissolved oxygen, W velocity of ocean currents at 5 m, Dry air mixing ratio of atmospheric $CO_{2\ anom}$, Mixed layer depth $_{anom}$
S of 60°S	Uncorrected	DIC, Bathymetry, SSS, Alkalinity, cos(Longitude), SST, Sea surface height $_{anom}$, W velocity of ocean currents at 195 m, 5 m, and 65 m, SSS_{anom}
S of 60°S	Corrected	Bathymetry, SSS, Alkalinity, SST, month, W velocity of ocean currents at 65 m, 105 m, and 195 m, Dissolved oxygen, cos (Longitude), Sea surface height, Latitude, SSS_{anom}

Uncorrected predictors were selected by a Stepwise BEL algorithm updated from ref. 11, corrected predictors were selected using the same algorithm but increasing the weighting of winter SOCAT measurements; The sort order of pCO_2 predictors indicated a relative contribution on decreasing predicting errors. The predictors denoted by subscript "anom" represent the monthly anomaly obtained by subtracting the monthly climatology. Data sources of used products are listed in Supplementary Table 1.

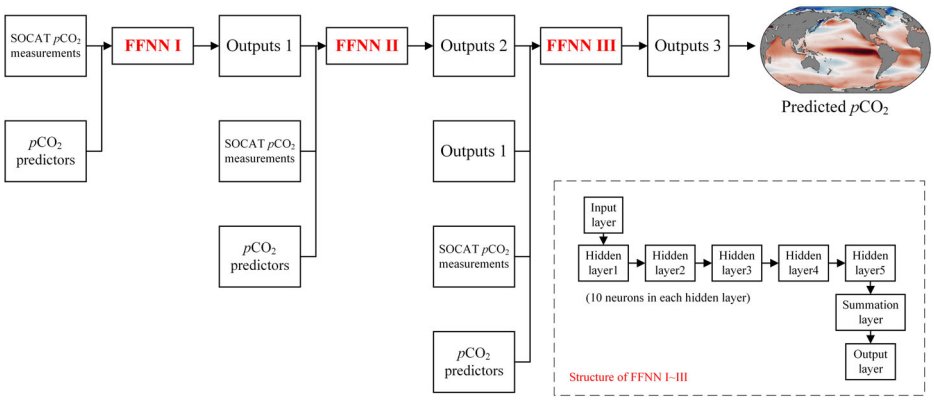

Fig. 8 | FFNN I~III: feed-forward neural network. pCO_2 predictors: selected environmental variables in Table 2. Outputs: pCO_2 values predicted by corresponding FFNN.

lower than in the summer, leading to a lower weighting on winter predicting errors compared to summer in the determination. As a result, the selected predictors are more reflective of factors influencing pCO_2 distribution in the summer (such as biological drawdown) while neglecting those in the winter (such as enhanced vertical mixing). Therefore, we increased the weighting of winter data to be nearly equal to that of summer, to carry out a winter correction of pCO_2 predictors in the Southern Ocean:

$$\mathrm{RMSE} = \sqrt{\frac{3 * \sum (\Delta pCO_{2\mathrm{May-Sep}})^2 + \sum (\Delta pCO_{2\mathrm{Oct-Apr}})^2}{3 * N_{\mathrm{May-Sep}} + N_{\mathrm{Oct-Apr}}}} \quad (2)$$

where the ΔpCO_2 was the difference between predicted pCO_2 and SOCAT pCO_2 measurements, and N was the number of monthly SOCAT measurements ($3*N_{\mathrm{May\text{-}Sep}} \approx N_{\mathrm{Oct\text{-}Apr}}$). Based on a self-organization map method, the Southern Ocean was divided into different regions according to the similarity of pCO_2 drivers, including two belt regions and three sectors connecting to major basins[11]. Therefore, the selection of pCO_2 predictors and reconstruction of pCO_2 in this work was based on three latitude areas: 35–50°S, 50–60°S, and south of 60°S (Table 2). Since there were no observed effects of uneven seasonal distribution on the neural network training in the 35–50°S region connecting the major basins, the correction of pCO_2 predictors is only conducted in the area south of 50°S.

CO_2 flux estimate

The sea–air CO_2 flux was estimated based on the pCO_2 difference across the interface[4,7]:

$$F = k \cdot (a_{\mathrm{subskin}} \cdot pCO_{2\mathrm{w}} - a_{\mathrm{skin}} \cdot pCO_{2\mathrm{atm}}) \quad (3)$$

where pCO_{2w} represents surface ocean pCO_2 and $pCO_{2\ atm}$ represents atmospheric pCO_2. The pCO_{2atm} was calculated from the xCO_2 of the NOAA Greenhouse Gas Marine Boundary Layer Reference product[61] and sea level pressure from the ERA5 monthly averaged data[57,58], with the water vapor correction[62]. a_{skin} and $a_{subskin}$ are the solubility of CO_2 at the skin and subskin layers[5], calculated from temperature and salinity[59]. k is the CO_2 transfer velocity as a function of wind speed[63]:

$$k = \Gamma(660/Sc)^{0.5}U^2 \quad (4)$$

where Sc is the Schmid number of CO_2 in seawater, and U is the average wind speed using the ERA5 product[57,58]. The transfer velocity was scaled by the scale factor (Γ) 0.27 for ERA5 wind products to match the ^{14}C constraint[64].

Uncertainty

The uncertainty of sea–air CO_2 flux estimate includes mainly three parts: the uncertainty of transfer velocity k, the cool skin impact, and the uncertainty in the surface ocean pCO_2 reconstruction. The uncertainty of transfer velocity k was related to the wind product and considered about 5–30%[6,21,65], and here we used 10%. Recent research suggested an underestimate of 0.35 PgC yr^{-1} in the global ocean carbon sink caused by the cool skin impact[6]. The uncertainty caused by the temperature and salinity gradient was considered 3% and 1.7% after the subskin correction, respectively[5,7]. The last uncertainty term came from the reconstruction of gridded surface ocean pCO_2 data, including the uncertainty of the pCO_2 measurement, averaging to 1° × 1° grids, and the pCO_2 interpolation. Thus, the total uncertainty in the pCO_2 reconstruction was calculated on average[66], where the measurement uncertainty σ(meas) was about 2–5 µatm[4,67], which was lower than the others and can be neglected[68]. The uncertainty of averaging to 1° × 1° grids, σ(grid), used 5 µatm from the previous research[54]. For the pCO_2 interpolation uncertainty σ(map), we used the predicting error of 7–25 µatm in different regions[11]. The uncertainty in each area was calculated as the following[68]:

$$\sigma(<pCO_2>)^2 = \frac{\sigma(\mathrm{grid})^2}{N_{\mathrm{eff}}(\mathrm{grid})} + \frac{\sigma(\mathrm{map})^2}{N_{\mathrm{eff}}(\mathrm{map})} \quad (5)$$

The $\sigma(<pCO_2>)$ calculated from the pCO_2 interpolation uncertainty ranges from 1.7 to 6.6 µatm in each region. Based on the average CO_2 transfer velocity of 0.07 mol C m^{-2} yr^{-1} in the Southern Ocean, the uncertainty $\sigma(pCO_2)$ caused by the pCO_2 interpolation errors in different regions range from ±0.05 to ±0.10 PgC yr^{-1}. The total uncertainty of pCO_2 interpolation estimated by the sum of squares of $\sigma(pCO_2)$ in each province was ±0.13 PgC yr^{-1}, corresponding to roughly 15% of the average Southern Ocean carbon sink estimated below. Thus, combining the uncertainties stemming from transfer velocity, cool skin influences, and pCO_2 interpolation, the final uncertainty was ±18.4% ($= \sqrt{10\%^2 + 3\%^2 + 1.7\%^2 + 15\%^2}$), using the square root of the sum squares propagation, corresponding to ± 0.16 PgC yr^{-1} (1σ).

Data availability

The dataset of 1° × 1° gridded surface ocean pCO_2 and sea–air CO_2 flux is available from the Institute of Oceanology of the Chinese Academy of Sciences Marine Science Data Center (https://doi.org/10.12157/iocas.2021.0022). Source data for each figure is available from the Figshare repository (https://doi.org/10.6084/m9.figshare.26066269).

Code availability

The MATLAB codes of the Stepwise BEL algorithm (version 2023.2) are available from the GitHub repository (https://github.com/GuorongZhong/Stepwise-BEL-FFNN-code-for-MATLAB.git).

Received: 12 October 2023; Accepted: 15 July 2024;
Published online: 24 July 2024

References

1. Sabine, C. L. et al. The oceanic sink for anthropogenic CO_2. *Science* **305**, 367–371 (2004).
2. Bushinsky, S. M. et al. Reassessing Southern Ocean air-sea CO_2 flux estimates with the addition of biogeochemical float observations. *Glob. Biogeochem. Cycle* **33**, 1370–1388 (2019).
3. Bakker, D. C. et al. A multi-decade record of high-quality fCO_2 data in version 3 of the Surface Ocean CO2 Atlas (SOCAT). *Earth Syst. Sci. Data* **8**, 383–413 (2016).
4. Wanninkhof, R. et al. Global ocean carbon uptake: magnitude, variability and trends. *Biogeosciences* **10**, 1983–2000 (2013).
5. Woolf, D. K., Land, P. E., Shutler, J. D., Goddijn-Murphy, L. M. & Donlon, C. J. On the calculation of air-sea fluxes of CO_2 in the presence of temperature and salinity gradients. *J. Geophys. Res. Oceans* **121**, 1229–1248 (2016).
6. Woolf, D. K. et al. Key uncertainties in the recent air-sea flux of CO_2. *Glob. Biogeochem. Cycle* **33**, 1548–1563 (2019).
7. Watson, A. J. et al. Revised estimates of ocean-atmosphere CO_2 flux are consistent with ocean carbon inventory. *Nat. Commun.* **11**, 4422 (2020).
8. Gregor, L. & Gruber, N. OceanSODA-ETHZ: a global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification. *Earth Syst. Sci. Data* **13**, 777–808 (2021).
9. Chau, T. T. T., Gehlen, M. & Chevallier, F. A seamless ensemble-based reconstruction of surface ocean pCO_2 and air–sea CO_2 fluxes over the global coastal and open oceans. *Biogeosciences* **19**, 1087–1109 (2022).
10. Gloege, L., Yan, M., Zheng, T. & McKinley, G. A. Improved quantification of ocean carbon uptake by using machine learning to merge global models and pCO_2 data. *J. Adv. Model. Earth Syst.* **14**, e2021MS002620 (2022).
11. Zhong, G. et al. Reconstruction of global surface ocean pCO_2 using region-specific predictors based on a stepwise FFNN regression algorithm. *Biogeosciences* **19**, 845–859 (2022).
12. Zeng, J. et al. A global surface ocean fCO_2 climatology based on a feed-forward neural network. *J. Atmos. Ocean. Technol.* **31**, 1838–1849 (2014).
13. Landschützer, P., Gruber, N. & Bakker, D. C. Decadal variations and trends of the global ocean carbon sink. *Glob. Biogeochem. Cycle* **30**, 1396–1417 (2016).
14. Iida, Y., Takatani, Y., Kojima, A. & Ishii, M. Global trends of ocean CO_2 sink and ocean acidification: an observation-based reconstruction of surface ocean inorganic carbon variables. *J. Oceanogr.* **77**, 323–358 (2021).
15. Rödenbeck, C., DeVries, T., Hauck, J., Le Quéré, C. & Keeling, R. F. Data-based estimates of interannual sea–air CO_2 flux variations 1957–2020 and their relation to environmental drivers. *Biogeosciences* **19**, 2627–2652 (2022).
16. Mikaloff Fletcher, S. E. et al. Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean. *Glob. Biogeochem. Cycle* **20**, GB2002 (2006).
17. Frölicher, T. L. et al. Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *J. Clim.* **28**, 862–886 (2015).
18. Landschützer, P. et al. The reinvigoration of the Southern Ocean carbon sink. *Science* **349**, 1221–1224 (2015).
19. Friedlingstein, P. et al. Global carbon budget 2022. *Earth Syst. Sci. Data* **14**, 4811–4900 (2022).
20. Mayot, N. et al. Climate-driven variability of the Southern Ocean CO_2 sink. *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.* **381**, 20220055 (2023).
21. Takahashi, T. et al. Climatological mean and decadal change in surface ocean pCO_2, and net sea–air CO_2 flux over the global oceans. *Deep-Sea Res. Part I Oceanogr. Res. Pap.* **56**, 2075–2076 (2009).
22. Hauck, J. et al. Consistency and challenges in the ocean carbon sink estimate for the global carbon budget. *Front. Mar. Sci.* **7**, 571720 (2020).
23. Gloege, L. et al. Quantifying errors in observationally based estimates of ocean carbon sink variability. *Glob. Biogeochem. Cycle* **35**, e2020GB006788 (2021).
24. Landschützer, P., Tanhua, T., Behncke, J. & Keppler, L. Sailing through the southern seas of air–sea CO_2 flux uncertainty. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **381**, 20220064 (2023).
25. Wang, Y. et al. Carbon sinks and variations of pCO_2 in the Southern Ocean from 1998 to 2018 based on a deep learning approach. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **14**, 3495–3503 (2021).
26. Kang, B., et al. Decoupling representation and classifier for long-tailed recognition. *8th International Conference on Learning Representations, ICLR 2020* https://doi.org/10.48550/arXiv.1910.09217 (2020).
27. Sutton, A. J. et al. Autonomous seawater pCO_2 and pH time series from 40 surface buoys and the emergence of anthropogenic trends. *Earth Syst. Sci. Data* **11**, 421–439 (2019).
28. Sutton, A. J. et al. High-resolution ocean and atmosphere pCO_2 time-series measurements from mooring SOFS_142E_46S in the Indian Ocean (NCEI Accession 0118546). *NOAA National Centers for Environmental Information* https://doi.org/10.3334/cdiac/otg.tsm_sofs_142w_46s (2014).
29. European Union-Copernicus Marine Service (CMEMS). Global Ocean Surface Carbon. *Marine Data Store (MDS)* https://doi.org/10.48670/MOI-00047 (2019).
30. Landschützer, P. et al. An observation-based global monthly gridded sea surface pCO2 and air-sea CO2 flux product from 1982 onward and its monthly climatology (NCEI Accession 0160558). Version 7.7. *NOAA National Centers for Environmental Information* https://doi.org/10.7289/v5z899n6 (2017).
31. Gregor, L. & Gruber, N. OceanSODA-ETHZ: a global gridded data set of the surface ocean carbonate system for seasonal to decadal studies of ocean acidification (v2022) (NCEI Accession 0220059). *NOAA National Centers for Environmental Information* https://doi.org/10.25921/m5wx-ja34 (2020).
32. Gruber, N., Landschützer, P. & Lovenduski, N. S. The variable Southern Ocean carbon sink. *Annu. Rev. Mar. Sci.* **11**, 159–186 (2019).
33. Le Quéré, C. et al. Saturation of the Southern Ocean CO_2 sink due to recent climate change. *Science* **316**, 1735–1738 (2007).
34. Lovenduski, N. S., Gruber, N. & Doney, S. C. Toward a mechanistic understanding of the decadal trends in the Southern Ocean carbon sink. *Glob. Biogeochem. Cycle* **22**, GB3016 (2008).
35. Gregor, L., Kok, S. & Monteiro, P. M. Interannual drivers of the seasonal cycle of CO_2 in the Southern Ocean. *Biogeosciences* **15**, 2361–2378 (2018).
36. Munro, D. R. et al. Recent evidence for a strengthening CO_2 sink in the Southern Ocean from carbonate system measurements in the Drake Passage (2002–2015). *Geophys. Res. Lett.* **42**, 7623–7630 (2015).
37. Abril, G. et al. Large overestimation of pCO_2 calculated from pH and alkalinity in acidic, organic-rich freshwaters. *Biogeosciences* **12**, 67–78 (2015).
38. Rödenbeck, C. et al. Interannual sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme. *Biogeosciences* **11**, 4599–4613 (2014).
39. Ritter, R. et al. Observation-based trends of the Southern Ocean carbon sink. *Geophys. Res. Lett.* **44**, 12–339 (2017).
40. Gillett, N. P. & Thompson, D. W. Simulation of recent Southern Hemisphere climate change. *Science* **302**, 273–275 (2003).
41. Miller, R. L., Schmidt, G. A. & Shindell, D. T. Forced annular variations in the 20th century intergovernmental panel on climate change fourth assessment report models. *J. Geophys. Res.-Atmos.* **111**, D18101 (2006).

42. Rödenbeck, C. et al. Data-based estimates of the ocean carbon sink variability–first results of the Surface Ocean pCO_2 Mapping intercomparison (SOCOM). *Biogeosciences* **12**, 7251–7278 (2015).
43. DeVries, T., Holzer, M. & Primeau, F. Recent increase in oceanic carbon uptake driven by weaker upper-ocean overturning. *Nature* **542**, 215–218 (2017).
44. McKinley, G. A., Fay, A. R., Eddebbar, Y. A., Gloege, L. & Lovenduski, N. S. External forcing explains recent decadal variability of the ocean carbon sink. *AGU Adv.* **1**, e2019AV000149 (2020).
45. Qu, B., Song, J., Yuan, H., Li, X., & Li, N. Air-sea CO_2 exchange process in the southern Yellow Sea in April of 2011, and June, July, October of 2012. *Cont. Shelf Res*. **80**, 8–19 (2014).
46. Laruelle, G. G. et al. Global high-resolution monthly pCO_2 climatology for the coastal ocean derived from neural network interpolation. *Biogeosciences* **14**, 4545–4561 (2017).
47. Song, J. et al. Carbon sinks/sources in the Yellow and East China Seas —Air-sea interface exchange, dissolution in seawater, and burial in sediments. *Sci. China Earth Sci.* **61**, 1583–1593 (2018).
48. Qi, D. et al. Climate change drives rapid decadal acidification in the Arctic Ocean from 1994 to 2020. *Science* **377**, 1544–1550 (2022).
49. Jacobs, S. S., Jenkins, A., Giulivi, C. F. & Dutrieux, P. Stronger ocean circulation and increased melting under Pine Island Glacier ice shelf. *Nat. Geosci.* **4**, 519–523 (2011).
50. Nakayama, Y., Schröder, M. & Hellmer, H. H. From circumpolar deep water to the glacial meltwater plume on the eastern Amundsen Shelf. *Deep-Sea Res. Part I Oceanogr. Res. Pap.* **77**, 50–62 (2013).
51. Hellmer, H. H., Kauker, F., Timmermann, R. & Hattermann, T. The fate of the southern Weddell Sea continental shelf in a warming climate. *J. Clim.* **30**, 4337–4350 (2017).
52. Cai, W. et al. Antarctic shelf ocean warming and sea ice melt affected by projected El Niño changes. *Nat. Clim. Chang.* **13**, 235–239 (2023).
53. Merino, N. et al. Antarctic icebergs melt over the Southern Ocean: climatology and impact on sea ice. *Ocean Model.* **104**, 99–110 (2016).
54. Sabine, C. L. et al. Surface Ocean CO_2 Atlas (SOCAT) gridded data products. *Earth Syst. Sci. Data* **5**, 145–153 (2013).
55. Bakker, D. C. E., et al. Surface Ocean CO2 Atlas Database Version 2023 (SOCATv2023) (NCEI Accession 0278913), monthly gridded product. *NOAA National Centers for Environmental Information* https://doi.org/10.25921/r7xa-bt92 (2023).
56. Landschützer, P. et al. A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences* **10**, 7793–7815 (2013).
57. Hersbach, H. et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **146**, 1999–2049 (2020).
58. Hersbach, H. et al. ERA5 monthly averaged data on single levels from 1940 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* https://doi.org/10.24381/cds.f17050d7 (2023).
59. Weiss, R. Carbon dioxide in water and seawater: the solubility of a non-ideal gas. *Mar. Chem.* **2**, 203–215 (1974).
60. Gregor, L., Lebehot, A. D., Kok, S. & Scheel Monteiro, P. M. A comparative assessment of the uncertainties of global surface ocean CO_2 estimates using a machine-learning ensemble (CSIR-ML6 version 2019a)–have we hit the wall? *Geosci. Model Dev.* **12**, 5113–5136 (2019).
61. Lan, X., Tans, P., Thoning, K. & NOAA Global Monitoring Laboratory. NOAA greenhouse gas marine boundary layer reference—CO_2 data sets. *NOAA GML* https://doi.org/10.15138/DVNP-F961 (2023).
62. Dickson, A. G., Sabine, C. L. & Christian, J. R. Determination of pCO_2 in air that is in equilibrium with a discrete sample of sea water. In *Guide to Best Practices for Ocean CO_2 Measurements* (eds Dickson, A. G. et al.) 7–8 (North Pacific Marine Science Organization, 2007).
63. Wanninkhof, R. Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res. Oceans* **97**, 7373–7382 (1992).
64. Sweeney, C. et al. Constraining global air-sea gas exchange for CO_2 with recent bomb ^{14}C measurements. *Glob. Biogeochem. Cycle* **21**, GB2015 (2007).
65. Ho, D. T. et al. Toward a universal relationship between wind speed and gas exchange: Gas transfer velocities measured with $^3He/SF_6$ during the Southern Ocean Gas Exchange Experiment. *J. Geophys. Res. Oceans* **116**, C00F04 (2011).
66. Wang, G., Dai, M., Shen, S. S., Bai, Y. & Xu, Y. Quantifying uncertainty sources in the gridded data of sea surface CO_2 partial pressure. *J. Geophys. Res. Oceans* **119**, 5181–5189 (2014).
67. Pfeil, B. et al. A uniform, quality controlled Surface Ocean CO_2 Atlas (SOCAT). *Earth Syst. Sci. Data* **5**, 125–143 (2013).
68. Landschützer, P., Gruber, N., Bakker, D. C. & Schuster, U. Recent variability of the global ocean carbon sink. *Glob. Biogeochem. Cycle* **28**, 927–949 (2014).

Acknowledgements

This research has been supported by the National Natural Science Foundation of China (grant nos.42176200), Laoshan Laboratory (LSKJ202204001, LSKJ202205001), National Key Research and Development Program (2022YFC3104305), Shandong Province and Yantai City Talent Programs. We thank the data support of the Marine Science Data Center and Public Technical Service Center, Institute of Oceanology, Chinese Academy of Sciences. We thank SOCAT for sharing the pCO_2 observation data. The Surface Ocean CO_2 Atlas (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 CO_2 database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for their contributions to SOCAT.

Author contributions

B.Q., Y.W., and B.Z. collected the environmental data products collection. J.M., Q.W., and J.X. synthesized the data product. G.Z. designed the method, and H.Y. improved it. G.Z., J.D., and L.D. performed the data construction. G.Z. wrote the manuscript draft and X.L., J.S., and F.W. contributed to the writing of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-024-01566-6.

Correspondence and requests for materials should be addressed to Xuegang Li or Jinming Song.

Peer review information *Communications Earth & Environment* thanks Gian Navarra and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: José Luis Iriarte Machuca, Alireza Bahadori, and Clare Davis. A peer review file is available.