Reconstruction of $p\mathrm{CO}_2$ Data in the Southern Ocean Based on Feedforward Neural Network

Yanjun Wang, Xiaofeng Li, Jinming Song, Xuegang Li, Guorong Zhong, and Bin Zhang

1 Introduction

The ocean plays a vital role in regulating global climate change, About ~30% of total emissions since the pre-industrial period has been stored in the ocean, However, about 50% of the oceanic uptake of anthropogenic carbon takes place in the Southern Ocean. It dominates the global heat and carbon dioxide absorption, therefore, many scientists regard the Southern Ocean as the main research region. The "Southern Ocean" ($< 35\,^{\circ}$S) was proposed by scientists around 2000 and was determined to be the fifth largest ocean in the world. It is the only ocean that completely surrounds the earth but is not divided by continents. It has important differences from ocean currents in the Pacific, Indian and Atlantic oceans—Antarctic Circumpolar Current (ACC). Moreover, the Southern Ocean is also an important region for global carbon absorption and release. Before industrial time, due to the influence of upwelling in the Southern Ocean, it has become a major carbon source region [6]. With the influence of human activities, the atmospheric pressure gradient shifted and turned into a carbon sink region. In the following section, We use the SOCAT dataset to build a Feedfoward neural network (FFNN), based on this network we reconstruct the Southern Ocean $p\mathrm{CO}_2$ data and calculate the CO_2 flux changes in the region, compare with other method, Our algorithm is compared with two neural network algorithms and has a smaller root mean square error.

Y. Wang · B. Zhang
Department of Marine Science Data Center, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China

X. Li (✉)
CAS Key Laboratory of Ocean Circulation and Waves, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China
e-mail: lixf@qdio.ac.cn

J. Song · X. Li · G. Zhong
Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, and the Center for Ocean Mega-Science,Chinese Academy of Sciences, Qingdao 266071, China

X. Li and F. Wang (eds.), *Artificial Intelligence Oceanography*,
https://doi.org/10.1007/978-981-19-6375-9_9

1.1 Observations of pCO_2 in Southern Ocean

Many data of the carbonate system can only be obtained by in-situ measurement. Due to the harsh environment of the Southern Ocean, the data collection is lacking. For sea surface data, through the continuous efforts of the scientists, the Surface Ocean CO_2 Atlas [13] has complies and quality control of ship data, fixed-point observation data, and drifting buoy data to formed a relatively complete observation data set (Fig. 1). This data set contains the pCO_2 data which can be used to calculate the sea-air carbon dioxide flux. We will use this database as the truth value to construct our neural network and reconstruct the pCO_2 gridded data of the entire Southern Ocean.

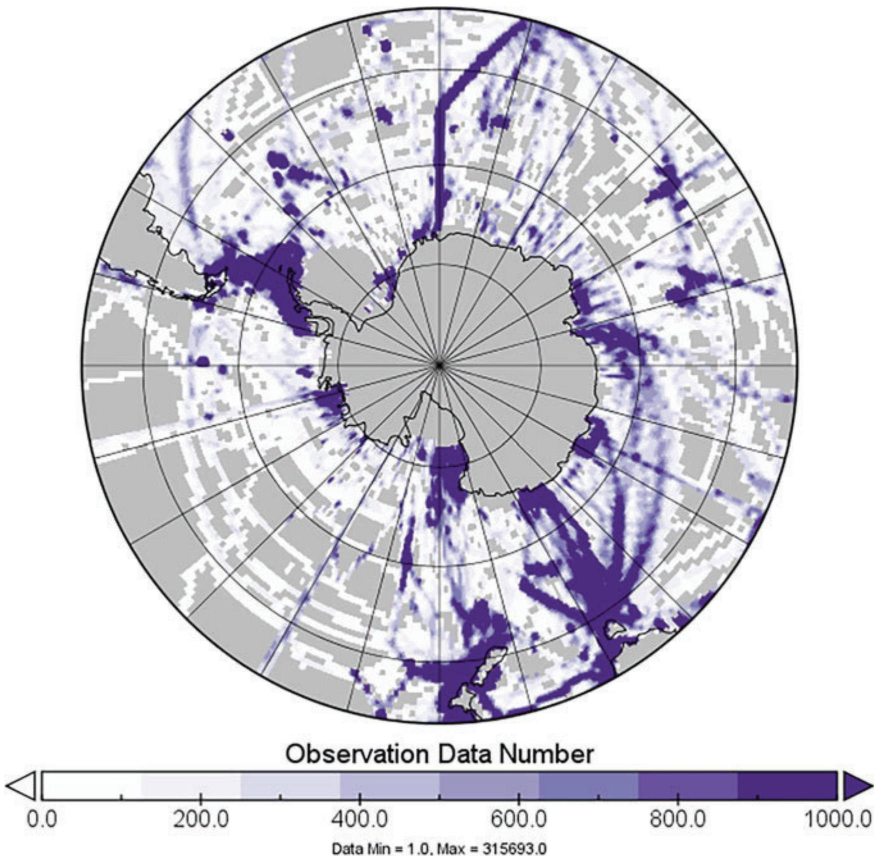

Fig. 1 1998-2018 SOCAT data observation heat map

1.2 Comparison of Reconstruction $\mathrm{p}CO_2$ Data

The results obtained by some traditional atmospheric inversions algorithms are greatly affected by the amount of observational data [17, 20]. Some spatial and temporal interpolations are based on empirical relationships between carbon dioxide and alternative variables, and are mainly concentrated in areas with relatively rich observations.

Neural network approaches have been frequently used in the reconstruction of surface $p\mathrm{CO}_2$ in recent years. To recreate the $p\mathrm{CO}_2$ data of the Southern Ocean, Gregor et al. employed a support vector machine (SVM) and a random forest (RF). The root-mean-square errors (RMSEs) were 16.45 μatm and 24.04 μatm, respectively. Meanwhile, Landschutzer et al. [11] created the SOM-FFNN method by combining a self-organizing map (SOM) with a feedforward neural network (FFNN) to recreate $p\mathrm{CO}_2$ data from the Southern Ocean. Sea surface temperature (SST), sea surface salinity (SSS), Mixed Layer Depth (MLD), chlorophyll concentration (CHL), and other metrics are used as inputs. The study shows that during the period 1980-2000, the Southern Ocean carbon sink has remained stagnant or even weakened, and continued to increase after 2002. Both data products showed good interannual and seasonal cyclical changes, but compare with the traditional machine learning algorithm (SVM and RF), SOM–FFNN show better performance. Denvil-Sommer et al. [3] employed the Laboratory of Climate and Environmental Sciences (LSCE)–FFNN method to reconstruct global $p\mathrm{CO}_2$ data, which maintained consistency with observational results. However, compared with the observed data, the Southern Ocean's reconstructed data has a larger error than other regions with more in situ observations.

In this chapter, we use the Surface Ocean CO_2 ATLAS (SOCAT V.6) data from 1998 to 2018 in the Southern Ocean, we applied the (CA)–FFNN method to reconstruct the monthly and $1^\circ \times 1^\circ$ $p\mathrm{CO}_2$ data of the Southern Ocean. Due to FFNN produces more stable data in sparse areas [20], and interpolates the data with small deviation [12], we use this method to reconstruct the Southern Ocean regional data. The procedure is separated into two parts. First, each parameter's correlation index is calculated and arranged. Second, the $p\mathrm{CO}_2$ data in the southern ocean blank area was interpolated using a relational model employing parameters with reasonably strong correlation coefficients as input variables of the FFNN. The current scenario, in which stations with less observation data have larger RMSE values, is improved by this strategy. As a result, this method might be used to recreate regional data. Finally, we looked at $p\mathrm{CO}_2$ fluctuations in the Southern Ocean on a seasonal, interannual, and interdecadal scale.

2 Data and Methods

2.1 Data

The parameters used in the CA method included SST , SST anomaly (SSTA), SSS, and SSS anomaly (SSSA); these parameters were all from the gridded dataset of Global Ocean Heat Content Change [2], while anomaly data were obtained by subtracting the average data values from the climatic state data of each month. Chlorophyll concentration (Chl-a) were based on satellite remote sensing data from the European Space Agency's Global Color Project, while MLD data were obtained from the French Institute of Marine Development. The u- and v- components of the wind field at 10 meters above sea level (a.s.l.) were taken from the European Centre for Medium-Range Weather Forecasts. All these data except MLD are monthly averages over a $1^\circ \times 1^\circ$ Lat/Lon box. MLD data is monthly averages over $0.5^\circ \times 0.5^\circ$.

In this chapter, we convert the $f\mathrm{CO}_2$ data in the SOCAT data set to $p\mathrm{CO}_2$ data as the training set and test set of FFNN. Transformation relationship between $f\mathrm{CO}_2$ and $p\mathrm{CO}_2$ is as follows [10]:

$$f\mathrm{CO}_2 = p\mathrm{CO}_2 \cdot \exp\left(p \cdot \frac{B + 2\delta}{R \times T_{\mathrm{subskin}}}\right) \tag{1}$$

where p is the atmospheric pressure (Pa), R is the gas constant (8.314 J K^{-1} mol^{-1}), SST is the sea surface temperature (K), $\mathrm{T}_{subskin}$ is the subskin temperature and B and δ are the correction coefficients, which are calculated as:

$$T_{\mathrm{subskin}} = \mathrm{SST} + 0.17 \tag{2}$$

$$B\left(\frac{m^3}{\mathrm{mol}}\right) = (-1636.75 + 12.0408SST - 3.27957 \times 10^{-2}\mathrm{SST}^2 + 3.16528 \times 10^{-5}\mathrm{SST}^3) \times 10^{-6} \tag{3}$$

$$\delta\left(\frac{m^3}{\mathrm{mol}}\right) = (57.7 - 0.118T_{\mathrm{subskin}}) \times 10^{-6} \tag{4}$$

The partial pressure of atmospheric CO_2 was calculated by the following formula [14]:

$$\mathrm{pCO}_{2a} = \mathrm{xCO}_2\left[P_{\mathrm{eq}} - \mathrm{VP}\,(H_2O)\right] \tag{5}$$

where xCO_2 is the dry air mixing ratio of atmospheric CO_2. The relevant data are collected from the reference data of marine boundary layer in the Earth System Research Laboratory of the National Oceanic and Atmospheric Administration (NOAA). Additionally, P_{eq} is the pressure at equilibrium, and VP (H_2O) is the steam of seawater at a given temperature [8]

$$\mathrm{VP} = 0.61121 \times e^{\left(18.678 - \frac{T_{\mathrm{subskin}}}{234.5}\right) \times \frac{T_{\mathrm{subskin}}}{257.14 + T_{\mathrm{subskin}}}} \tag{6}$$

where the T_{subskin} is subskin temperature.

In order to reduce the complexity of calculation of too large data set on neural network learning, we use Eq. 7 to normalize all data.

$$x^{'} = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{7}$$

where x is actual value, min(x) is the minimum value of x, max(x) is the maximum value of x.

Since the Chl-a data in this study did not include relevant records before the launch of SeaWiFS in 1997, our research period was from 1998 to 2018. The spatial resolution of all parameter data was $1^\circ \times 1^\circ$. Longitude (Lon) and latitude (Lat) are in 360° and 180° coordinate systems, and trigonometric conversion functions were used to ensure continuity and normalization.

2.2 *Nonlinear Neural Network Model for the* pCO_2 *Reconstruction in the Southern Ocean*

We use Equations 8 and 9 to calculate the correlation coefficient, and build a covariance matrix between $p\mathrm{CO}_2$ and other collected data, as shown in Fig. 3.

$$\mathrm{Cov}\ (, Y) = E\left[(X - u_x)\left(Y - u_y\right)\right] \tag{8}$$

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\beta_x \beta_y} \tag{9}$$

where u is the mean of the value, β is standard deviation of the value, Cov(X, Y) is the calculated covariance matrix, and ρ is the correlation coefficient.

We use the parameters with correlation coefficients > 0.1 as the input parameters, considering the relevance of chemical effects between SST and $p\mathrm{CO}_2$ [18], We still use SST as an input parameter. After correlation analysis, the selected parameters were the SST, SSSA, MLD, CHL, the u-component (U) of the sea surface wind field, and the partial pressure of atmospheric CO_2 ($p\mathrm{CO}_{2a}$). The established correlation equations between $p\mathrm{CO}_2$ and the main parameters are summarized in Eq. 10.

$$p\mathrm{CO}_2 = f\begin{pmatrix}\mathrm{SST, SSSA, CHL, MLD,}\\ U, \mathrm{aCO}_2, \mathrm{Lon, Lat}\end{pmatrix} \tag{10}$$

A nonlinear regression model was built using the FFNN. Although an FFNN's output data improves and becomes more accurate as the number of layers and neurons

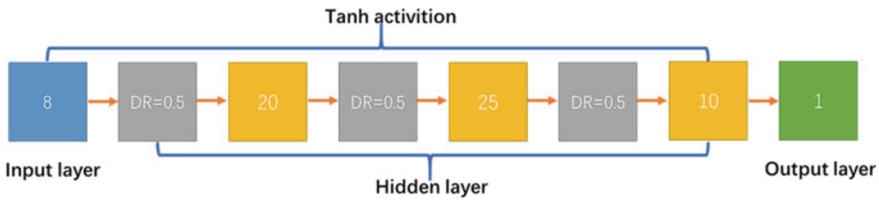

Fig. 2 The Structure of our FFNN, The gray square is the dropout layer and dropout rate is 0.5, blue square is input layer, yellow square is hidden layer, green square is output layer

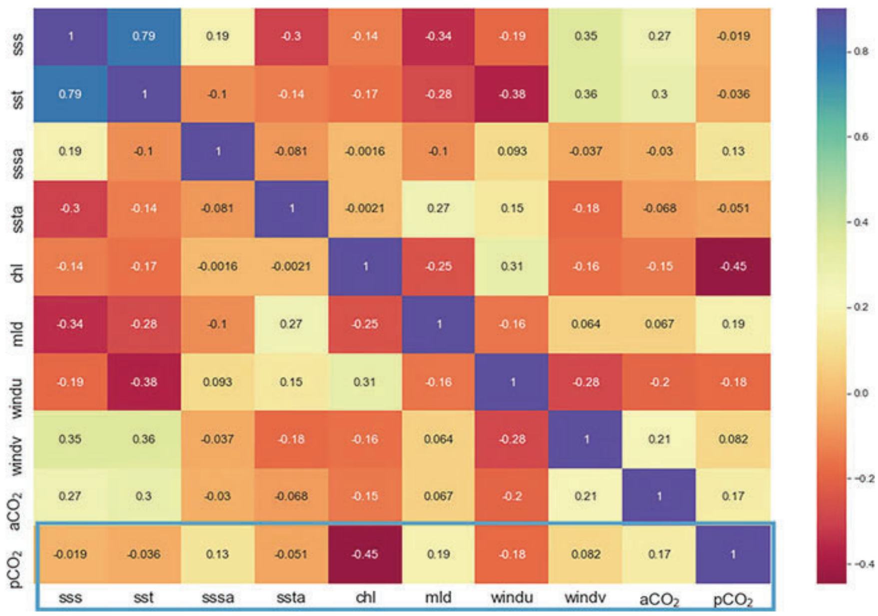

Fig. 3 Matrix of correlation coefficients. The correlation coefficient value of the x-axis and y-axis parameters is represented by each colored box. The value of the pCO_2 correlation coefficient with other parameters is contained within the blue box

in the FFNN grows, the model's size is also determined by the amount of data utilized for model training. Because there is less observational data for the Southern Ocean than for other regions, we built a simple FFNN structure, the neutral network structure of which is shown in Fig. 2. The final model at Step 2 has eight layers (six hidden layers), and the numbers on the figure represent the size of the tensor input to each layer. A gray square represents the dropout layer, and the dropout rate is 0.5. The hyperparameters of the neural network were determined using k-fold cross-validation (Fig. 4).

The data were divided into 75%/25% portions used for training/testing sets. The neural network consists of eight layers, and the middle layer had six completely connected hidden layers, we added three dropout layers and gave each layer's dropout ratio 0.5 to prevent the FFNN from overfitting. Through many tests and detailed

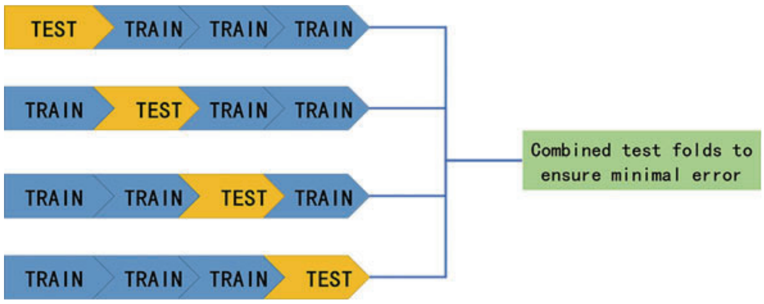

Fig. 4 *k*-Fold cross-validation, which was divided into four folds in this study, with 25% data for testing and the rest for training to create the best neural network. The yellow shape represents test data, whereas the blue shape represents train data

analyses, the hyperbolic tangent (Tanh) was selected as the activation function of the neuron, and the using the mean squared error (MSE) as the loss function:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\text{observed}_i - \text{predicted}_i\right)^2 \tag{11}$$

where observed_i is the observation data, and predicted_i is the data predicted by the FFNN model, and we using RMSProp as the optimization function [21].

In order to control the amount of information, we adjusted the adaptive learning rate. The CA–FFNN was then formed by combining a main factor analysis and based on the parameters, we build a FFNN structure and get a nonlinear regression model through training.

2.3 Calculation of Carbon Dioxide Flux in the Southern Ocean

The formula for calculating the carbon dioxide flux at the air-sea interface is [29] :

$$F = K \cdot \Delta f\text{CO}_2 = K \cdot (a_{\text{subskin}}\text{fCO}_{2w} - a_{\text{skin}}\text{fCO}_{2a}) \tag{12}$$

where a is the solubility of CO_2 in seawater (mol kg^{-1} atm^{-1}), calculated by Weiss [10]:

$$\begin{aligned}\ln a &= -60.2409 + 93.4517\left(\frac{100}{T_{\text{subskin}}}\right) 23.3585 \times \ln\left(\frac{T_{\text{subskin}}}{100}\right) S \\ &\times \left[0.023517 - 0.023656 \times \left(\frac{T_{\text{subskin}}}{100}\right) + 0.0047036 \times \left(\frac{T_{\text{subskin}}}{100}\right)^2\right]\end{aligned} \tag{13}$$

In Equation 12, $a_{subskin}$ is calculated by the subskin temperature, a_{skin} is calculated by the skin temperature. fCO_{2w} is the fugacity of subskin seawater CO_2, fCO_{2w} is the fugacity of subskin seawater CO_2, fCO_{2a} is the fugacity of atmospheric CO_2, and K is the exchange rate, which is usually considered as a function of wind speed .

$$K = \Gamma(660/\mathrm{Sc})^{0.5}U^2 \tag{14}$$

Here, Sc is the Schmidt number of CO_2 in seawater at a given T_{subskin} temperature, such that:

$$Sc = 2073.1 - 125.62 \times T_{\mathrm{subskin}} + 3.6276 \times T^2_{\mathrm{subskin}} - 0.043219 \times T^3_{\mathrm{subskin}} \tag{15}$$

where U is the monthly mean wind speed (m/s) at 10 m height from the cross-calibrated multi-platform ocean surface wind vector analysis product and Γ is the scale factor which was evaluated based on different wind speed products (e.g., 0.39, 0.251, 0.31, etc.) and have been used in other studies [14, 24, 28]. Based on an average wind speed of 6.38 m s^{-1} in the ECMWF product the scale factor of 0.31 was used to reach a global mean transfer velocity of 16 cm h^{-1}, consistent with the new radiocarbon-based constraints.

2.4 Evaluation

Due to the limited observation data in the Southern Ocean, the data set used for verification will be very small, so the segmentation of the data set will lead to huge differences between RMSE and mean absolute error. In order to ensure reliable model verification, we used 100% data to train, test and verify the model, and continuously optimized the neural network model and the internal weight. Finally, the neural network was used to predict the observed area. RMSE is calculated to be 8.86 μatm, while MAE is 5.01.

Figure 5 shows that the predicted values are very close to the observed values and R^2= 0.93. In Table 1, we list the RMSE and MAE between the results of different algorithms and the actual values. SOM-FFNN merged a self-organizing map (SOM) and feedforward neural network, and the RMSE is 12.24. LSCE-FFNN employed the Laboratory of Climate and Environmental Sciences, and the RMSE is 17.40. We conclude that the CA-FFNN-based models outperform both the SOM–FFNN and LSCE–FFNN.

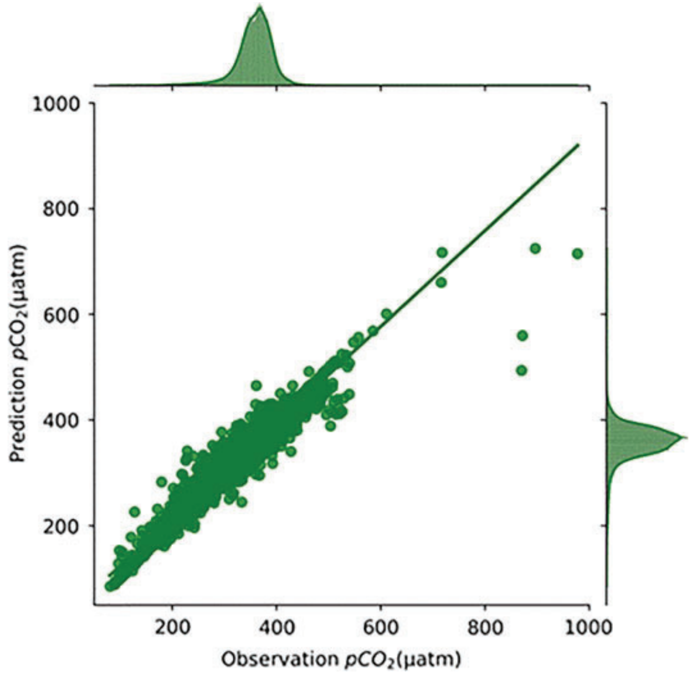

Fig. 5 Scatter fit of product data and observation data with same station

Table 1 Comparison of our Algorithms' Errors to LSCE-FFNN and SOM-FFNN

Artificial Intelligence Algorithm	RMSE	MAE
FFNN for Southern Ocean	8.86	5.01
LSCE–FFNN [3]	17.40	11.92
SOM–FFNN [12]	12.24	7.36

3 Results and Discussion

3.1 Seasonal Variation in Southern Ocean Sea Surface $\mathbf{p}CO_2$

According to the new dataset, the $p\mathrm{CO}_2$ data changes periodically with the seasons. This result is consistent with the seasonal changes in other studies [16, 25, 27]. The seasonal mean amplitude of ocean surface $p\mathrm{CO}_2$ in the southern ocean was 13.02 μatm and our data has similar seasonal variation characteristics compared with the observational data of the Southern Ocean [15], the $p\mathrm{CO}_2$ reaching its minimum in summer, and increase in winter (Fig. 6), and driven by both biological and physical factors, $p\mathrm{CO}_2$ in the Southern Ocean shows obvious seasonal changes [22], In winter,

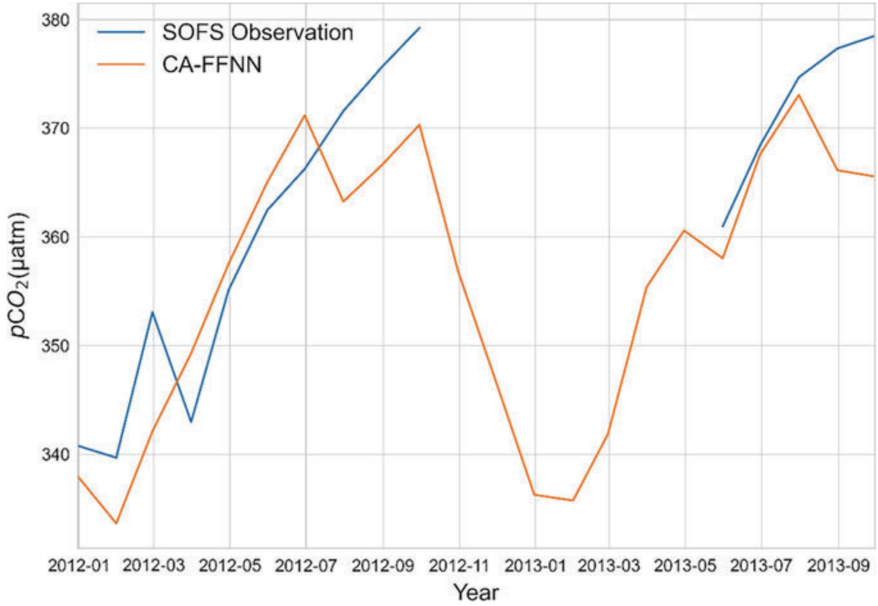

Fig. 6 In SOFS, the product data and real value

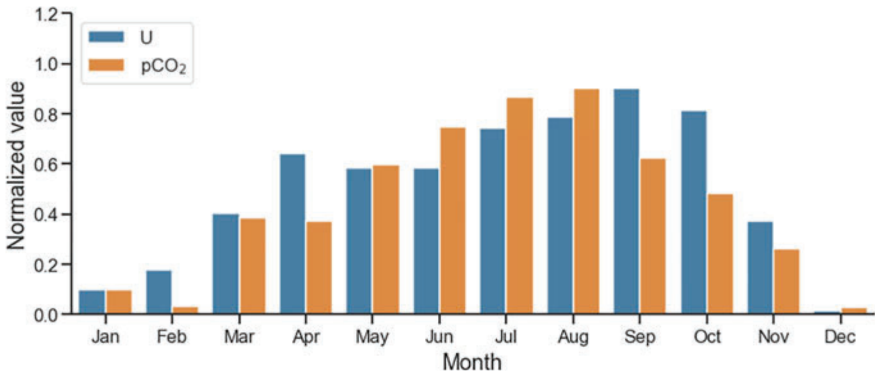

Fig. 7 From 1998 to 2018, the normalized mean monthly U-component of wind and pCO_2 was calculated

due to the enhancement of the wind field in the Southern Ocean, as shown in the Fig. 7, the Ekman transport caused by the wind field also intensifies [1, 7], strengthening upwelling and improving the efficiency of the biological pump.

The dissolved inorganic carbon in the bottom layer migrates to the surface layer under the influence of the upwelling, making the surface pCO_2 increase continuously. With the melting of sea ice in the Southern Ocean in summer, marine primary productivity gradually recovers, the Chl-a concentration increases, as shown in Fig. 8,

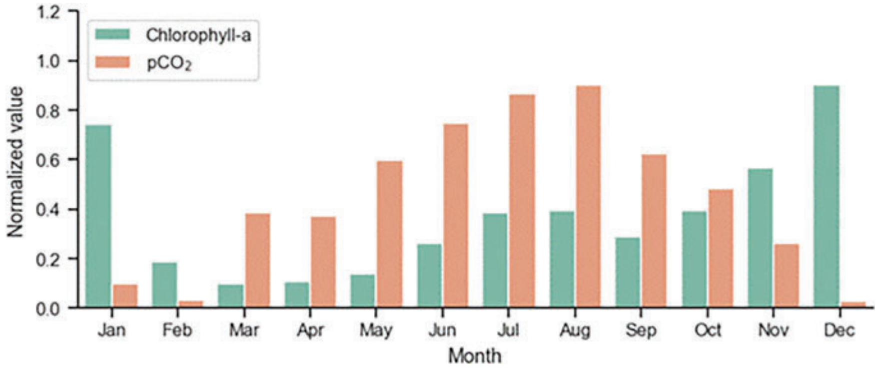

Fig. 8 From 1998 to 2018, the average monthly CHL and $p\mathrm{CO}_2$ data were normalized

and CO_2 in sea water is absorbed through photosynthesis [26], which lead to surface $p\mathrm{CO}_2$ decrease. This period is mainly due to biological factors.

3.2 Annual Variation in Southern Ocean Sea Surface pCO_2

Analyzing the inter-annual change of the reconstructed $p\mathrm{CO}_2$ data from 1998 to 2018, the mean surface $p\mathrm{CO}_2$ of the Southern Ocean increased from 351.88 μatm to 372.65 μatm—a total increase of 20.77 μatm in 21 years and an annual mean increase of 0.99 μatm/yr. As shown in Fig. 9, the Southern Ocean $p\mathrm{CO}_2$ has maintained a high growth rate.

By calculating the linear rate of change in the Southern Ocean spatial region over a 21-year period, it is found that the $p\mathrm{CO}_2$ in most areas is gradually increasing, as shown in Fig. 10. The growth rate around 35 °to 55 ° is faster than other regions.

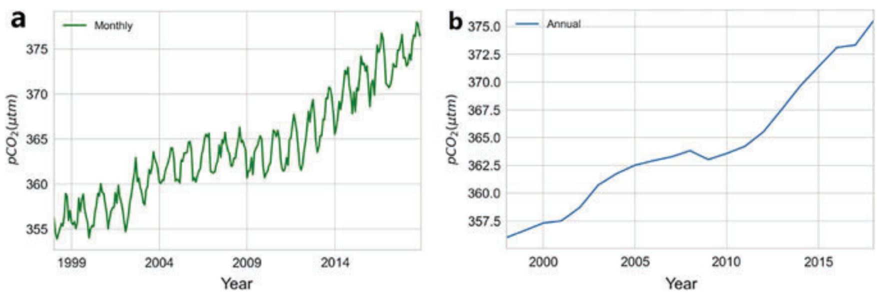

Fig. 9 **a** is monthly fluctuations in the Southern Ocean's $p\mathrm{CO}_2$ (atm) from 1998 to 2018; **b** is yearly fluctuations in the Southern Ocean's $p\mathrm{CO}_2$ (atm) from 1998 to 2018

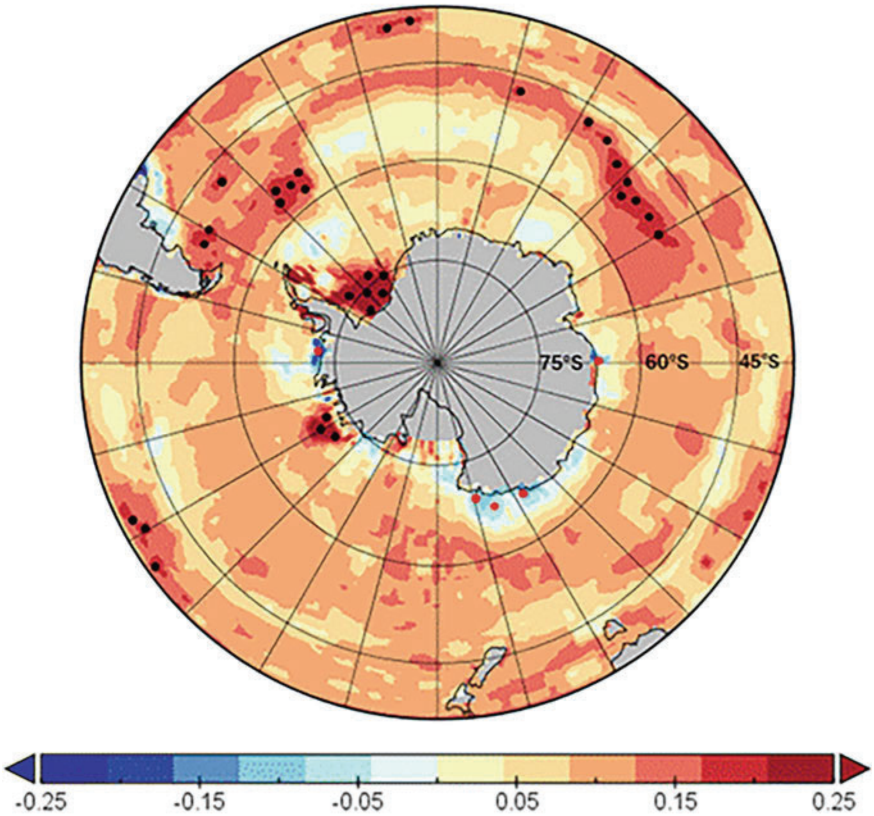

Fig. 10 The rate of change in the surface Southern Ocean's pCO_2 (μatm yr^{-1}) concentration

Since 2002, many study results have shown that pCO_2 in the Southern Ocean has maintained a high growth rate [23], and our data also shows this trend.

3.3 *Variability in Sea—AirCO_2 Flux*

As for the rate of change of ΔfCO_2, Most of the Southern Ocean is transforming into a carbon sink area. The black/red dots in Fig. 11 represent ΔfCO_2 regions toward positive/negative trends with high change rate. According to the distribution of pCO_2 in the Southern Ocean since 1998, the status of inner ring ($50 - 70\,°$S) as a carbon source is changing, while the outer ring ($35 - 50\,°$S) has always maintained a strong carbon sink state, and there is no tendency to weaken. The changes of CO_2 flux in the Southern Ocean calculated by our model are consistency with other models for the evolution of intensity [19]. Using Eq. 12 to calculate the CO_2 flux, the Southern

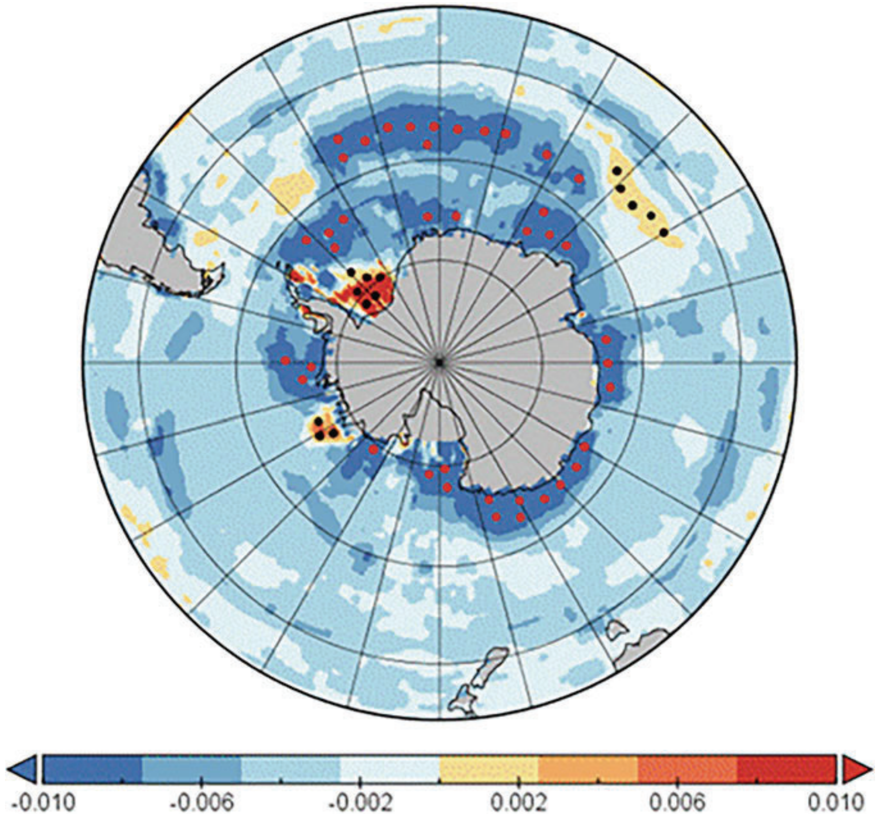

Fig. 11 Carbon sink when $\Delta fCO_2 < 0$, carbon source when $\Delta fCO_2 > 0$, Rate of change in the ΔfCO_2 of the Southern Ocean

Ocean's CO_2 flux was found to have changed substantially over the past two decades. The ΔfCO_2 in the Southern Ocean also changes regularly with the seasons, with the strongest in early summer and get the weakest at the end of winter (Fig. 13). Many studies have shown that in early 1990s, the Southern Ocean was saturated with carbon and regained its vitality at the beginning of the 21st century [4]. The data products reproduces the strong increase of carbon sinks in the Southern Ocean since the 21st century (Fig. 14).

In terms of interannual changes, the carbon sink of the Southern Ocean increased from -0.21 Pg C yr^{-1} in 1998 to -1.67 Pg C yr^{-1} in 2018.

One standard deviation was used as an indicator of error:

$$\sigma_n = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n^2}} \tag{16}$$

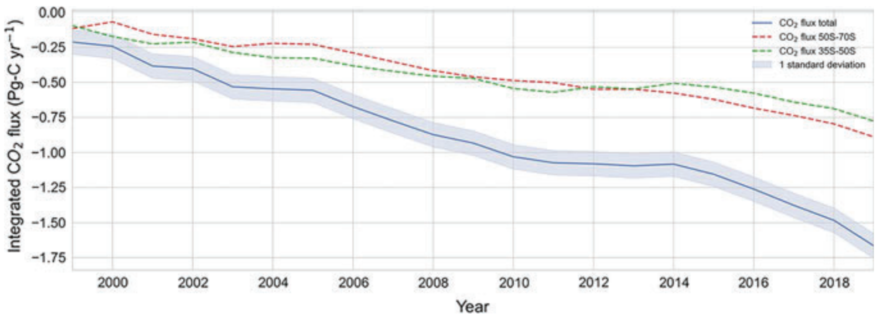

Fig. 12 CO_2 flux trends in of the Southern Ocean from 1998 to 2018

where x_i is the actual value, $\overline{x}$ is the mean value of x, n is number of data, and the error range was within $\pm$ 0.0.087 Pg C yr^{-1}.

We found that the carbon sinks in the Southern Ocean did not always maintain a trend of rapid growth. During 2010-2013, the carbon sinks stagnated. As shown in Fig. 12, we found the similar phenomenon in many other reconstructed data [5].Many studies have shown that changes in the Southern Annular Mode (SAM) led to the stagnation of carbon sinks in the 1990s [5]. However, the stagnation was not strongly correlated with the SAM. Stability during this period was mainly due to the weakening of the carbon sink intensity from $35 - 50\,^{\circ}$S.Changes in this region have also been attributed to the barometric asymmetry of the Zontal Waves 3 (ZW3) model [9]. As for models that rely on observational data, it is difficult to capture such large and subtle inter-annual changes.

As shown in Fig. 16, there is an obvious double-ring structure before 2010, which is not always a carbon sink. The inner ring ($50 - 70\,^{\circ}$S), change with the seasons. In April, May, June, July, August, and September, the region serves as a carbon source, emitting CO_2 into the atmosphere. In October, November, December, January, February, and March, it absorbs CO_2, as shown in Fig. 14. The outer ring ($35 - 50\,^{\circ}$S) is the main carbon sink region (Fig. 15), and undertakes most CO_2 absorption. From the perspective of the inter-annual changes in the entire region, the Southern Ocean carbon dioxide flux changes to carbon sinks.

However, with the increase of carbon sink in the outer ring and the weakening of the carbon source in the inner ring, after 2010 this ring structure is gradually disappearing. As shown in Fig. 16, most Southern Ocean regions become carbon sinking regions, because the ΔfCO_2 in the Southern Ocean decrease significantly since 1998.

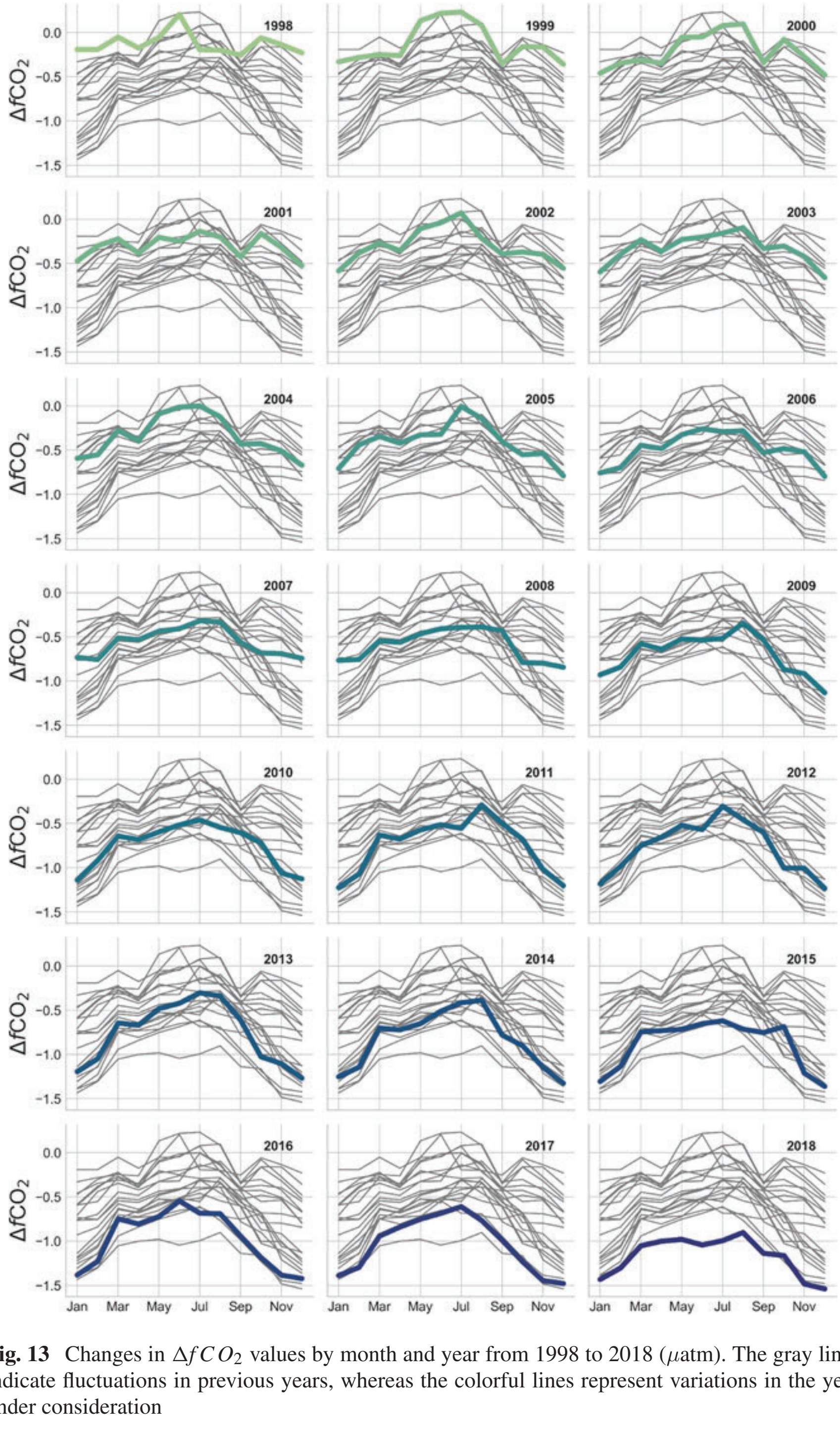

Fig. 13 Changes in ΔfCO_2 values by month and year from 1998 to 2018 (μatm). The gray lines indicate fluctuations in previous years, whereas the colorful lines represent variations in the year under consideration

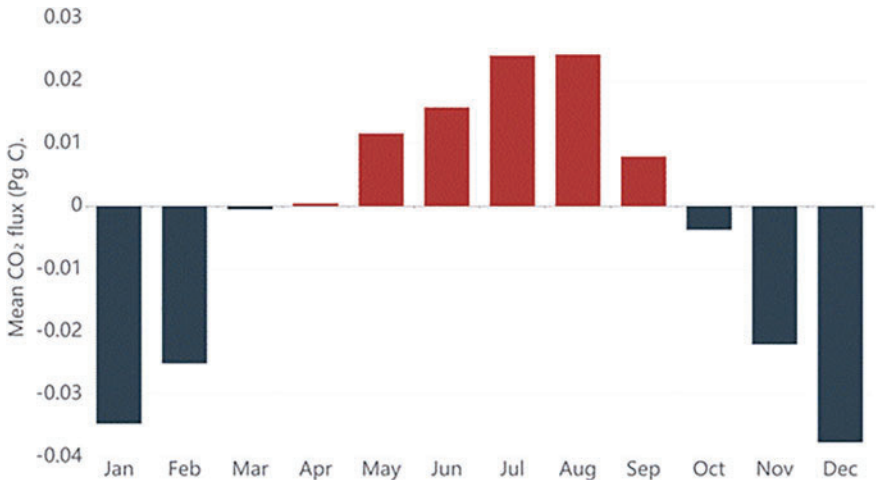

Fig. 14 Each month's average CO_2 flux in $50-70\,^{\circ}$S (Pg C). The Southern Ocean was a carbon supply in the red columns, whereas the Southern Ocean was a carbon sink in the black columns

Fig. 15 Each month's average CO_2 flux in $35-50\,^{\circ}$S (Pg C). The Southern Ocean was a carbon sink, as indicated by the blue columns

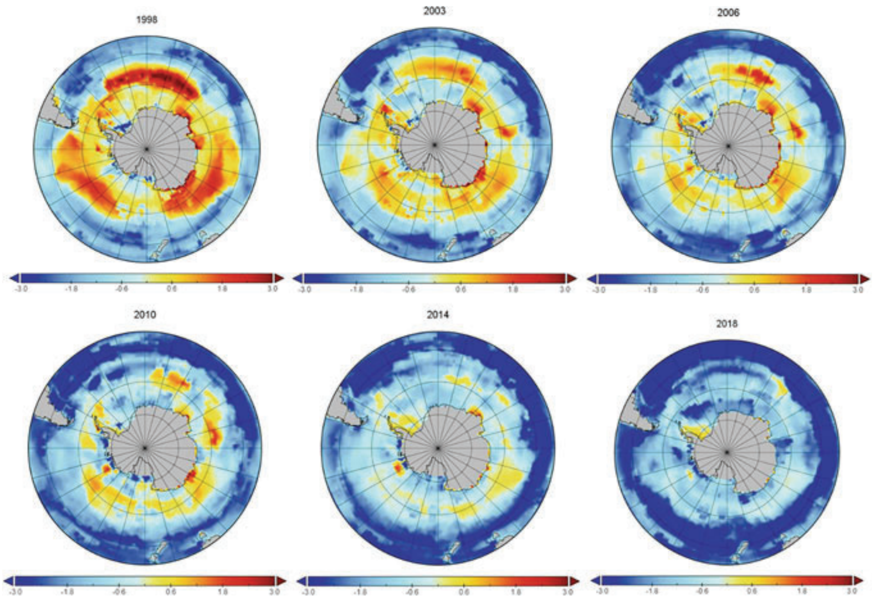

Fig. 16 In the Southern Ocean, mean sea surface CO_2 fluxes (Pg C) were measured in 1998, 2003, 2006, 2010, 2014, and 2018

4 Conclusion

In this chapter, we propose a feedforward neural network for reconstructing $p\mathrm{CO}_2$ data in the Southern Ocean that is generalizable for reconstructing regional data. The reconstruction process consists of two steps. First, we collect all parameters that may have impact on $p\mathrm{CO}_2$ from the literature and experimental data and obtain the covariance matrix of the variables by calculation. The parameters with higher correlation coefficient values and an effect on the process change of $p\mathrm{CO}_2$ were kept as inputs FFNN, and the final model was constructed and used to reconstruct the $p\mathrm{CO}_2$ data of the Southern Ocean with a monthly temporal resolution and a spatial resolution of $1^\circ \times 1^\circ$ in the second step after continuous and iterative calculation and optimization.

First of all, we find the key parameters that affect $p\mathrm{CO}_2$ in the Southern Ocean changes. Secondly, use the advantages of neural network technology to interpolate in the data sparse area, and build a new model by filtering parameters. Finally, in the Southern Ocean, we compare the new data with the measured data and get the root mean square error with 8.86 μatm which is better than the data reconstructed from global data.

The results of our reconstruction demonstrate that $p\mathrm{CO}_2$ in the Southern Ocean's surface layer varies seasonally and has risen since 2000. It did, however, reach a halt from 2010 and 2013, after which it resumed its upward trend. In the Southern Ocean, carbon dioxide flux is distributed in a double ring shape. The primary carbon sink region is $35 - 50\,^{\circ}$S; south of $50\,^{\circ}$S, seasonal carbon sources and sinks alternated. Despite the fact that our findings are consistent with earlier studies, the reconstructed surface pCO_2 products require ongoing verification. Our model will improve as the frequency and range of observations in the Southern Ocean increase.

References

1. Anderson RF, Ali S, Bradtmiller LI, Nielsen SHH, Fleisher MQ, Anderson BE, Burckle LH (2009) Wind-driven upwelling in the southern ocean and the deglacial rise in atmospheric co2. Sci 323(5920):1443–1448. https://doi.org/10.1126/science.1167441
2. Cheng L, Trenberth KE, Gruber N, Mann ME, Abraham JP, Fasullo J, Guancheng L, Zhu J (2020) Improved estimates of changes in upper ocean salinity and the water cycle. 2020:GC095–03
3. Denvil-Sommer A, Gehlen M, Vrac M, Mejia C (2019) LSCE-FFNN-v1: A two-step neural network model for the reconstruction of surface ocean pCO_2 over the global ocean. Geosci Model Dev 12(5):2091–2105. https://doi.org/10.5194/gmd-12-2091-2019
4. Frölicher TL, Sarmiento JL, Paynter DJ, Dunne JP, Krasting JP, Winton M (2015) Dominance of the southern ocean in anthropogenic carbon and heat uptake in cmip5 models. J Clim 28(2):862–886. https://doi.org/10.1175/JCLI-D-14-00117.1
5. Gregor L, Lebehot AD, Kok S, Scheel Monteiro PM (2019) 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(12):5113–5136. https://doi.org/10.5194/gmd-12-5113-2019
6. Gruber N, Gloor M, Mikaloff Fletcher SE, Doney SC, Dutkiewicz S, Follows MJ, Gerber M, Jacobson AR, Joos F, Lindsay K, Menemenlis D, Mouchet A, Müller SA, Sarmiento JL, Takahashi T (2009) Oceanic sources, sinks, and transport of atmospheric CO2. Glob Biogeochem Cycles 23(1).https://doi.org/10.1029/2008GB003349
7. Gruber N, Landschützer P, Lovenduski NS (2019) The variable southern ocean carbon sink. Annu Rev Mar Sci 11(1):159–186. https://doi.org/10.1146/annurev-marine-121916-063407
8. Junzeng X, Qi W, Shizhang P, Yanmei Y (2012) Error of saturation vapor pressure calculated by different formulas and its effect on calculation of reference evapotranspiration in high latitude cold region. Procedia Eng 28:43–48. https://doi.org/10.1016/j.proeng.2012.01.680
9. Keppler L, Landschützer P (2019) Regional wind variability modulates the southern ocean carbon sink. Sci Rep 9(1):7384. https://doi.org/10.1038/s41598-019-43826-y
10. Körtzinger A (1999) Determination of carbon dioxide partial pressure (p(CO2)). Methods Seawater Anal, John Wiley & Sons Ltd, chap 9:149–158. https://doi.org/10.1002/9783527613984.ch9
11. Landschützer P, Gruber N, Bakker DCE, Schuster U, Nakaoka S, Payne MR, Sasse TP, Zeng J (2013) A neural network-based estimate of the seasonal to inter-annual variability of the atlantic ocean carbon sink. Biogeosciences 10(11):7793–7815. https://doi.org/10.5194/bg-10-7793-2013
12. Landschützer P, Gruber N, Haumann FA, Rödenbeck C, Bakker DCE, van Heuven S, Hoppema M, Metzl N, Sweeney C, Takahashi T, Tilbrook B, Wanninkhof R (2015) The reinvigoration of the southern ocean carbon sink. Sci 349(6253):1221–1224. https://doi.org/10.1126/science.aab2620
13. Landschützer P, Gruber N, Bakker DCE (2016) Decadal variations and trends of the global ocean carbon sink. Glob Biogeochem Cycles 30(10):1396–1417. https://doi.org/10.1002/2015GB005359

14. Landschützer P, Gruber N, Bakker DCE (2016) Decadal variations and trends of the global ocean carbon sink. Glob Biogeochem Cycles 30(10):1396–1417. https://doi.org/10.1002/2015GB005359
15. Landschützer P, Gruber N, Bakker DCE, Stemmler I, Six KD (2018) Strengthening seasonal marine CO2 variations due to increasing atmospheric CO2. Nat Clim Chang 8(2):146–150. https://doi.org/10.1038/s41558-017-0057-x
16. Metzl N, Brunet C, Jabaud-Jan A, Poisson A, Schauer B (2006) Summer and winter air- sea CO2 fluxes in the southern ocean. Deep Sea Res Part I: Ocean Res Pap 53(9):1548–1563. https://doi.org/10.1016/j.dsr.2006.07.006
17. Peylin P, Bousquet P, Le Quéré C, Sitch S, Friedlingstein P, McKinley G, Gruber N, Rayner P, Ciais P (2005) Multiple constraints on regional CO2 flux variations over land and oceans. Glob Biogeochem Cycles 19(1). https://doi.org/10.1029/2003GB002214
18. Pinho L, Duarte CM, Marotta H, Enrich-Prast A (2016) Temperature dependence of the relationship between pCO_2 and dissolved organic carbon in lakes. Biogeosciences 13(3):865–871. https://doi.org/10.5194/bg-13-865-2016
19. Ritter R, Landschützer P, Gruber N, Fay AR, Iida Y, Jones S, Nakaoka S, Park GH, Peylin P, Rödenbeck C, Rodgers KB, Shutler JD, Zeng J (2017) Observation-based trends of the Southern Ocean Carbon Sink. Geophysical Research Letters 44(24):12,339–12,348. https://doi.org/10.1002/2017GL074837
20. Rödenbeck C, Bakker DCE, Gruber N, Iida Y, Jacobson AR, Jones S, Landschützer P, Metzl N, Nakaoka S, Olsen A, Park GH, Peylin P, Rodgers KB, Sasse TP, Schuster U, Shutler JD, Valsala V, Wanninkhof R, Zeng J (2015) Data-based estimates of the ocean carbon sink variability - first results of the Surface Ocean pCO_2 mapping intercomparison (SOCOM). Biogeosciences 12(23):7251–7278. https://doi.org/10.5194/bg-12-7251-2015
21. Ruder S (2017) An overview of gradient descent optimization algorithms. arXiv:1609.04747 [cs] 1609.04747
22. Shadwick EH, Trull TW, Tilbrook B, Sutton AJ, Schulz E, Sabine CL (2015) Seasonality of biological and physical controls on surface ocean CO2 from hourly observations at the southern ocean time series site south of australia. Glob Biogeochem Cycles 29(2):223–238. https://doi.org/10.1002/2014GB004906
23. Silvano A (2020) Changes in the southern ocean. Nat Geosci 13(1):4–5. https://doi.org/10.1038/s41561-019-0516-2
24. ...Takahashi T, Sutherland SC, Wanninkhof R, Sweeney C, Feely RA, Chipman DW, Hales B, Friederich G, Chavez F, Sabine C, Watson A, Bakker DCE, Schuster U, Metzl N, Yoshikawa-Inoue H, Ishii M, Midorikawa T, Nojiri Y, Körtzinger A, Steinhoff T, Hoppema M, Olafsson J, Arnarson TS, Tilbrook B, Johannessen T, Olsen A, Bellerby R, Wong CS, Delille B, Bates NR, de Baar HJW (2009) Climatological mean and decadal change in surface ocean pCO2, and net sea- air CO2 flux over the global oceans. Deep Sea Res Part II: Top Stud Ocean 56(8):554–577. https://doi.org/10.1016/j.dsr2.2008.12.009
25. Takahashi T, Sweeney C, Hales B, Chipman DW, Newberger T, Goddard JG, Iannuzzi RA, Sutherland SC (2012) The changing carbon cycle in the southern ocean. Oceanography 25(3):26–37
26. Takao S, Nakaoka SI, Hashihama F, Shimada K, Yoshikawa-Inoue H, Hirawake T, Kanda J, Hashida G, Suzuki K (2020) Effects of phytoplankton community composition and productivity on sea surface pCO2 variations in the Southern Ocean. Deep Sea Res Part I: Ocean Res Pap 160:103263. https://doi.org/10.1016/j.dsr.2020.103263
27. Thomalla SJ, Fauchereau N, Swart S, Monteiro PMS (2011) Regional scale characteristics of the seasonal cycle of chlorophyll in the southern ocean. Biogeosciences 8(10):2849–2866. https://doi.org/10.5194/bg-8-2849-2011

28. Wanninkhof R, Castle RD, Shannahoff J (2016) A multi-decade record of high-quality fCO2 data in version 3 of the Surface Ocean CO2 Atlas (SOCAT). https://doi.org/10.1594/PANGAEA.851406
29. Watson AJ, Schuster U, Shutler JD, Holding T, Ashton IGC, Landschützer P, Woolf DK, Goddijn-Murphy L (2020) Revised estimates of ocean-atmosphere CO2 flux are consistent with ocean carbon inventory. Nat Commun 11(1):4422. https://doi.org/10.1038/s41467-020-18203-3