

Carbon Sinks and Variations of $p\text{CO}_2$ in the Southern Ocean From 1998 to 2018 Based on a Deep Learning Approach

Yanjun Wang, Xiaofeng Li[✉], Fellow, IEEE, Jinming Song, Xuegang Li[✉], Guorong Zhong, and Bin Zhang[✉], Member, IEEE

Abstract—The Southern Ocean comprises 25% of the global ocean surface area, accounts for nearly half of the total carbon sink of the global oceans, and is a place that significantly reduces the impacts of anthropogenic CO_2 emissions. Due to the sparsity of observational data, the changes in Southern Ocean carbon sinks over time remain uncertain. In this study, we integrated correlation analysis and a feedforward neural network to improve the accuracy of carbon flux estimations in the Southern Ocean. Based on observation data from 1998–2018, we reconstructed the Southern Ocean's $p\text{CO}_2$ grid data during this period. The root-mean-square error obtained by fitting the observation data was $8.86 \mu\text{atm}$, indicating that the results were better than those of the two primary statistically based models in the Surface Ocean $p\text{CO}_2$ mapping intercomparison. The results also showed that the Southern Ocean's capacity to act as a carbon sink has gradually increased since 2000; it reduced during 2010–2013 but increased significantly after that. The Southern Ocean's seasonality is characterized by minimum carbon uptake in winter due to increased upwelling; this is followed by a rapid increase toward maximum uptake in summer, which is mainly biologically driven. There is an apparent double-ring structure in the Southern Ocean, as noted in other studies. This study confirms that the inner ring ($50^\circ\text{--}70^\circ\text{S}$) is a carbon source area gradually transforming into a carbon sink, while the outer ring ($35^\circ\text{--}50^\circ\text{S}$) continues to serve as a carbon sink.

Index Terms—Carbon sink, feedforward neural network (FFNN), machine learning, $p\text{CO}_2$, Southern Ocean.

I. INTRODUCTION

IN THIS study, the Southern Ocean is defined as the area south of 35°S , and is the only ocean that is not divided by continents [1]. Since the Industrial Revolution, the global ocean

has absorbed $\sim 30\%$ of anthropogenic carbon dioxide emissions [2]. The ocean plays a crucial role in absorbing atmospheric CO_2 ; without this absorption, earth's surface temperature would be much higher [3], [4]. Past observational data and model estimates have indicated that the Southern Ocean is responsible for absorbing up to half of all anthropogenic CO_2 emissions [2], [5]. Therefore, the Southern Ocean is inextricably linked to the global carbon cycle and climate change [6], [7]. Although the Southern Ocean is essential, due to a lack of observational data, past studies in this area have not provided accurate calculations of the trends in CO_2 changes, let alone an understanding of the driving factors of such changes [8].

Traditional estimates, such as atmospheric inversions approaches, were limited by the sparse atmospheric CO_2 measurements [9], [10]. Some spatial and temporal interpolation based on empirical relationships between CO_2 and proxy variables [11], [12] were largely focused on the relatively observation-rich regions [13]. The emergence of biogeochemical models temporarily provided more credible solutions for predicting CO_2 fluxes in this region [14], but this model simulation relied on the correctness of the process simulation structure [15].

In recent years, neural network algorithms have been widely employed to reconstruct surface $p\text{CO}_2$ data. Compared with statistical interpolations, neural network approaches are not limited by sparse observations and can adequately represent the interannual variability [15]. Gregor *et al.* [16] used a support vector machine (SVM) and random forest (RF) to reconstruct the $p\text{CO}_2$ data of the Southern Ocean. The resultant root-mean-square errors (RMSEs) were 16.45 and $24.04 \mu\text{atm}$. Meanwhile, Landschützer *et al.* [17] merged a self-organizing map (SOM) and feedforward neural network (FFNN) to construct the SOM-FFNN method to reconstruct the $p\text{CO}_2$ data of the Southern Ocean. The inputs include sea surface temperature (SST), sea surface salinity (SSS), the depth of the mixed layer (MLD), chlorophyll concentration (CHL), and other parameters. Their results indicated that the Southern Ocean carbon sink stagnated or even decreased between 1980 and 2000, but then gradually recovered its original strength in ~ 2002 . Both data products showed good interannual and seasonal cyclical changes, but the SOM-FFNN algorithm offers better performance than the machine learning algorithm (SVM and RF). Denvil-Sommer *et al.* [18] employed the Laboratory of Climate and Environmental Sciences (LSCE)-FFNN method to reconstruct global

Manuscript received November 11, 2020; revised January 7, 2021; accepted March 12, 2021. Date of publication March 17, 2021; date of current version April 7, 2021. This work was supported in part by the National Key R&D Program of China under Grant 2017YFA0603201, in part by the 13th Five-Year Informatization Plan of the Chinese Academy of Sciences, Construction of Scientific Data Center System, under Grant XXH-13514, and in part by Big Earth Data Science Engineering Project XDA19060104. (Corresponding author: Xiaofeng Li.)

Yanjun Wang, Xiaofeng Li, and Bin Zhang are with the Department of Marine Science Data Center, Institute of Oceanology, and the Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao 266071, China (e-mail: yjwang@qdio.ac.cn; xiaofeng.li@ieee.org; zhangbin@qdio.ac.cn).

Jinming Song, Xuegang Li, and Guorong Zhong are with the 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 (e-mail: jmsong@qdio.ac.cn; xiaofeng.li@ieee.org; zhongguorong18@mails.ucas.ac.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3066552

$p\text{CO}_2$ data, which maintained consistency with observational results. However, compared with the observed data, the Southern Ocean's reconstructed data have a larger error than other regions with more *in situ* observations.

In this study, using the Surface Ocean CO_2 ATLAS (SOCAT V.6) data of the Southern Ocean from 1998 to 2018, we applied the correlation analysis (CA)–FFNN method to reconstruct the monthly and $1^\circ \times 1^\circ$ $p\text{CO}_2$ data of the Southern Ocean. FFNN have advantages over other algorithms because it produces data more stable in sparse regions [15] and interpolates the sparse $p\text{CO}_2$ data with small bias [19]. This method is primarily divided into two steps. First, the correlation index of each parameter is calculated and sorted. Second, using the parameters with relatively high correlation coefficients as the input variables of the FFNN, the correlation matrix between $p\text{CO}_2$ data and other observed variables is constructed, and correlation parameters are used to reconstruct the $p\text{CO}_2$ data in the blank area of the Southern Ocean. This method facilitates data analysis and reconstruction for specific regions and improves the current situation wherein sites with limited observation data have higher RMSE values. Therefore, this approach may be generalizable for reconstructing regional data. Finally, we analyzed the seasonal, interannual, and interdecadal variations of $p\text{CO}_2$ in the Southern Ocean.

II. DATA AND METHODS

A. Data

The parameters used in the CA of the $p\text{CO}_2$ data included SST, SST anomaly (SSTA), SSS, and SSS anomaly (SSSA); these parameter data were all from the gridded dataset of Global Ocean Heat Content Change [20], while anomaly data were obtained by subtracting the average data values from the climatic state data of each month. Chlorophyll concentrations (Chl-a) were based on satellite remote sensing data from the European Space Agency's Glob Colour Project, and MLD data were from the French Research Institute for Exploitation of the Sea. The u - and v - components of the wind field at 10 m 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 are monthly averages over $0.5^\circ \times 0.5^\circ$.

In this study, the measured $p\text{CO}_2$ data used for FFNN training and validation were converted from the SOCAT grid $f\text{CO}_2$ dataset. The conversion relationship between $f\text{CO}_2$ and $p\text{CO}_2$ is as follows [21]:

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

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

$$T_{\text{subskin}} = \text{SST} + 0.17 \quad (2)$$

$$B \left(\frac{\text{m}^3}{\text{mol}} \right) = (-1636.75 + 12.0408\text{SST} - 3.27957$$

$$\times 10^{-2}\text{SST}^2 + 3.16528 \times 10^{-5}\text{SST}^3) \times 10^{-6} \quad (3)$$

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

The partial pressure of atmospheric CO_2 at each grid point was calculated by the following formula [7]:

$$p\text{CO}_{2a} = x\text{CO}_2 [P_{\text{eq}} - \text{VP}(\text{H}_2\text{O})] \quad (5)$$

where $x\text{CO}_2$ is the dry air mixing ratio of atmospheric CO_2 . The relevant data were collected from the United States National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratories marine boundary layer reference. Additionally, P_{eq} is the pressure at equilibrium, and $\text{VP}(\text{H}_2\text{O})$ is the steam of seawater at a given temperature [22]

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

where is the T_{subskin} is subskin temperature.

To reduce the computational difficulty and influence of excessively large datasets on neural network learning, we normalized all data by

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

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 SeaWiFS launch in 1997, our study period is between 1998 and 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.

B. FFNN Construction

The data collected and the corresponding $p\text{CO}_2$ observation data were used as the inputs for CA, and a covariance matrix was constructed using (8) and (9) as follows, to calculate the correlation coefficient, as shown in Fig. 1:

$$\text{Cov } X, Y = E[(X - u_x)(Y - u_y)] \quad (8)$$

$$\rho = \frac{\text{Cov}(X, Y)}{\beta_x \beta_y}. \quad (9)$$

Here, u is the mean of the value, β is standard deviation of the value, $\text{Cov}(X, Y)$ is the calculated covariance matrix, and ρ is the correlation coefficient. Through CA, the parameters with correlation coefficients > 0.1 were used as the input parameters, considering the relevance of chemical effects between SST and $p\text{CO}_2$, the SST should be used as input parameter [23]. After correlation analysis, the selected parameters were the SST, SSSA, MLD, CHL, the u -component (U) of the sea level wind field, and the partial pressure of atmospheric CO_2 ($p\text{CO}_{2a}$). The established correlation equations between $p\text{CO}_2$ and other parameters are summarized in

$$p\text{CO}_2 = f\left(\begin{matrix} \text{SST}, \text{SSSA}, \text{CHL}, \text{MLD}, \\ U, a\text{CO}_2, \text{Lon}, \text{Lat} \end{matrix}\right). \quad (10)$$

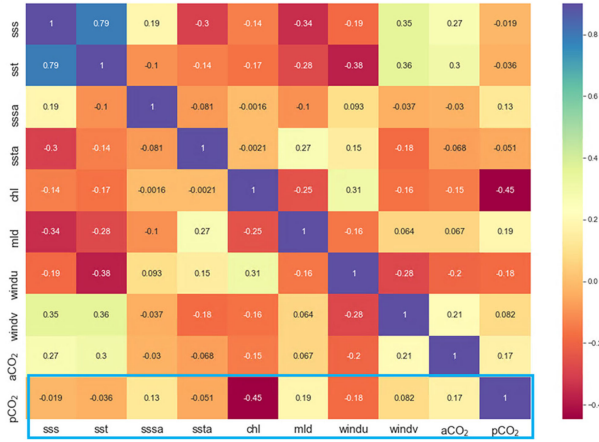


Fig. 1. Correlation coefficient matrix. Each colored box is the correlation coefficient value of the x-axis parameter and the y-axis parameter. Inside the blue box is $p\text{CO}_2$ correlation coefficient value with another parameter.

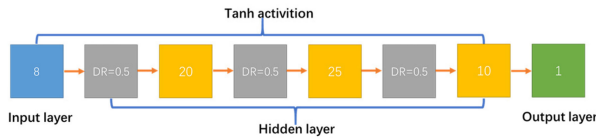


Fig. 2. Structure of the FFNN. Abbreviation: DR: Dropout rate.

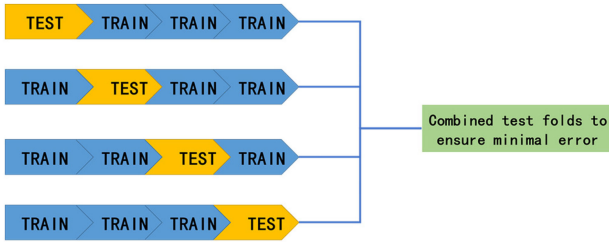


Fig. 3. k -fold cross validation, in this study divided into fourfold, in each fold, 25% data for test and others for training to build the optimal neural network. The yellow shape is test data and blue shape is train data.

The FFNN was used to construct a nonlinear regression model. Although the output data of an FFNN improve and become more accurate as the number of layers and neurons in the FFNN increases, the model's size also depends on the amount of data used for model training. Since the amount of observational data available for the Southern Ocean is significantly less than that for other regions, we constructed a relatively simple FFNN, 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 number of each layer's input tensor is marked on the figure. The grey square is the dropout layer, and the dropout rate is 0.5.

The k -fold cross validation is used to determine the hyperparameters of the neural network (see Fig. 3). The Southern Ocean data were divided into 25%/75% portions used for testing/training sets. The trained neural network contained eight layers, and the middle layer had three fully connected hidden layers. Simultaneously, to prevent the FFNN from overfitting, we added three dropout layers and gave each layer's dropout ratio 0.5. Through many tests and detailed analyses, the hyperbolic tangent (Tanh) was selected as the activation function of the

neuron, and the mean-squared error (MSE) as the loss function

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

where observed_i is the observation data, and predicted_i is the data predicted by the FFNN model using RMSProp as the optimization function [24].

By adjusting the adaptive learning rate ρ , the amount of information obtained was controlled. The CA-FFNN was then formed by combining a factor analysis of the main influencing factors and the FFNN.

C. Computation of Sea–Air CO_2 Fluxes

The formula for calculating the CO_2 flux at the air–sea interface is [25]

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

where a is the solubility of CO_2 in seawater ($\text{mol kg}^{-1} \text{ atm}^{-1}$), calculated by Weiss [21]

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

In (12), a_{subskin} is calculated by the subskin temperature, a_{skin} is calculated by the skin temperature. $f\text{CO}_{2w}$ is the fugacity of subskin seawater CO_2 , $f\text{CO}_{2a}$ is the fugacity of atmospheric CO_2 , and K is the exchange rate, which is usually considered as a function of wind speed [26]

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

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

$$\begin{aligned} Sc = & 2073.1 - 125.62 \times T_{\text{subskin}} + 3.6276 \times T_{\text{subskin}}^2 \\ & - 0.043219 \times T_{\text{subskin}}^3 \end{aligned} \quad (15)$$

where U is the monthly mean wind speed (m/s) at 10 m height from the cross-calibrated multiplatform 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 [7], [27], [28]. Based on an average wind speed of 6.38 ms^{-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.

D. Evaluation

As limited observational data exist for the Southern Ocean, and the dataset used for validation will be very small, the division of this dataset will result in substantial variance in the RMSE and mean-absolute error (MAE). To ensure reliable

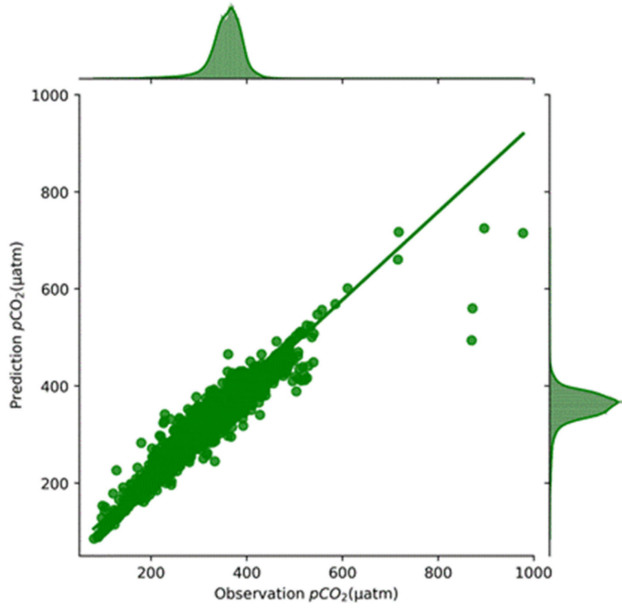


Fig. 4. Fitting of prediction and observation data.

TABLE I
COMPARISON OF THE ERRORS OF DIFFERENT ALGORITHMS

Algorithm	RMSE	MAE
CA-FFNN	8.86	5.01
LSCE-FFNN[18]	17.40	11.92
SOM-FFNN[19]	12.24	7.36

model validation, we used 100% of the data for model training, testing, and validation, and to continuously optimize both the neural network model and the internal weight. The final neural network was used to predict the area of the existing observations. The calculated RMSE is 8.86, while MAE is 5.01.

Fig. 4 shows that the predicted values are very close to the observed values and $R^2 = 0.93$. In Table I, SOM-FFNN merged a SOM and FFNN, 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.

III. RESULTS AND DISCUSSION

A. Seasonal Variation in Southern Ocean Sea Surface pCO_2

Based on our new data, we find a clear pCO_2 seasonal variation in the Southern Ocean. This observation is highly consistent with the seasonal changes observed in the reconstructed data products of other studies [29]–[31]. The seasonal mean amplitude of surface pCO_2 was $13.02 \mu\text{atm}$; our products have a similar seasonal trend compared with a station-observed actual data in the Southern Ocean [32], the pCO_2 reaching its maximum in winter, and becoming smaller in summer (see Fig. 5), and the seasonal changes of pCO_2 in the Southern Ocean resulted from the combined effects of biological and physical factors [33]. In winter, due to the strengthening of the Southern Ocean wind field, as shown in Fig. 6, the Ekman transport caused by the

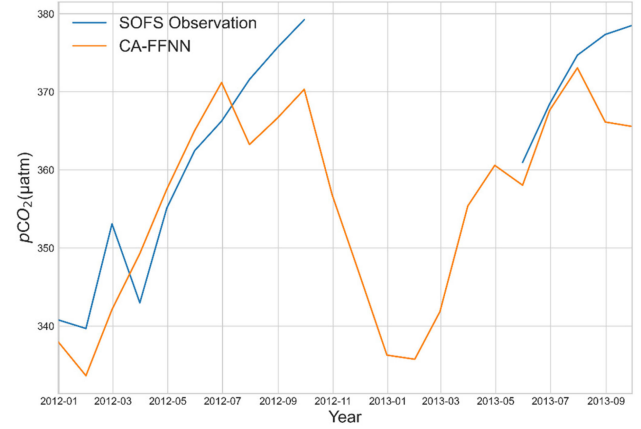
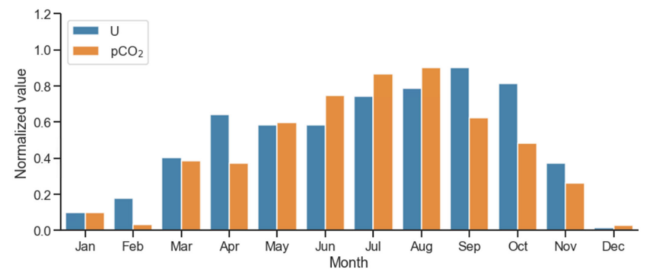
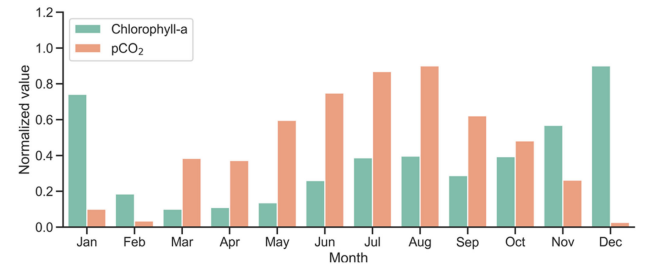


Fig. 5. Product data and actual value in SOFS.

Fig. 6. Normalized mean monthly U -component of wind and pCO_2 from 1998–2018.Fig. 7. Normalized mean monthly CHL and pCO_2 from 1998–2018.

wind field also intensifies [34], [35], strengthening upwelling and improving the efficiency of the biological pump.

The transfer of dissolved inorganic carbon from the bottom layer to the surface causes the pCO_2 of the surface layer to increase during this period, with upwelling playing a leading role. As sea ice melts in summer, primary productivity is enhanced, the Chl-a concentration increases, as shown in Fig. 7, and the ability of organisms to consume CO_2 is gradually restored [36], resulting in a decrease in surface CO_2 concentrations. Biological factors dominate this period.

B. Annual Variation in Southern Ocean Sea Surface pCO_2

From the calculated results of the model, the mean pCO_2 of the surface waters of the Southern Ocean increased from $351.88 \mu\text{atm}$ in 1998 to $372.65 \mu\text{atm}$ in 2018—a total increase of $20.77 \mu\text{atm}$ in 21 years and an annual mean increase of $0.99 \mu\text{atm}(\text{yr})^{-1}$. This high growth rate has been maintained, as shown in Fig. 8.

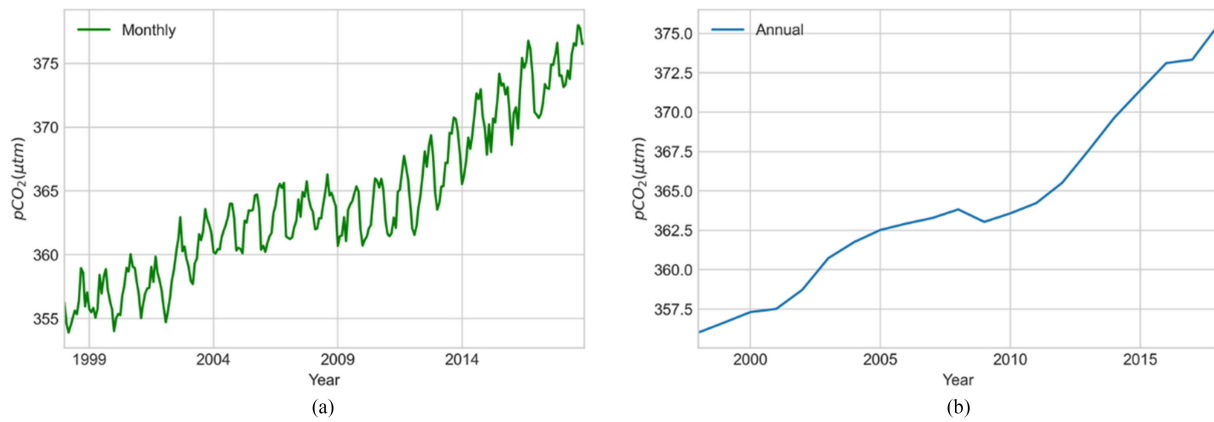


Fig. 8. (a) Monthly and (b) annual variations (μatm) in the $p\text{CO}_2$ of the Southern Ocean from 1998 to 2018.

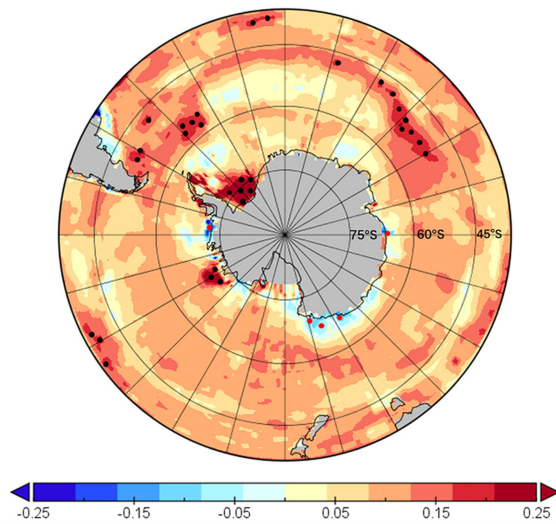


Fig. 9. Rate of change in the $p\text{CO}_2$ in $\mu\text{atm}(\text{yr})^{-1}$ of the surface Southern Ocean.

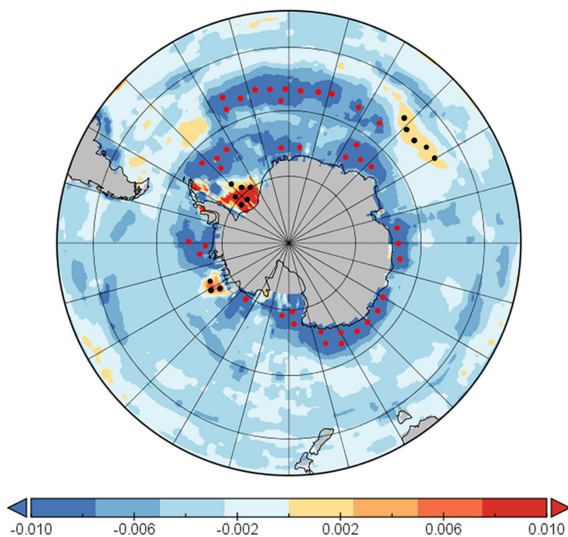


Fig. 10. Rate of change in the $\Delta f\text{CO}_2$ of the Southern Ocean, carbon sink when $\Delta f\text{CO}_2 < 0$, and carbon source when $\Delta f\text{CO}_2 > 0$.

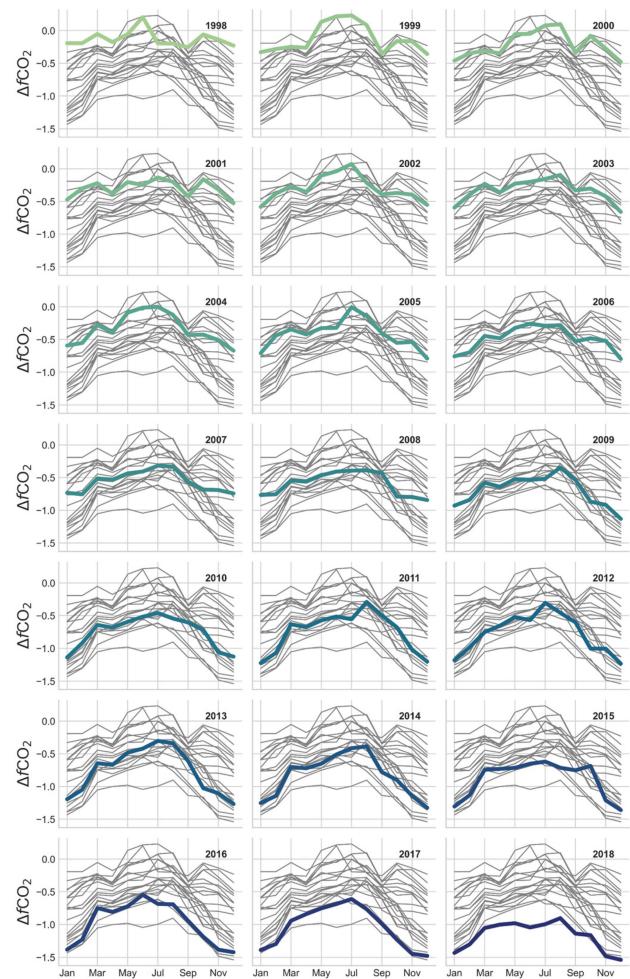


Fig. 11. Seasonal and annual variation in the $\Delta f\text{CO}_2$ of the Southern Ocean (μatm). The grey lines are the variations in other years and the colored lines represent the variation for each studied year (1998–2018).

By calculating the rate of change of every grid point in the Southern Ocean, the overall $p\text{CO}_2$ was found to be gradually increasing, as shown in Fig. 9. In terms of its regional distribution, a relatively high growth rate has been maintained from 35–55°S.

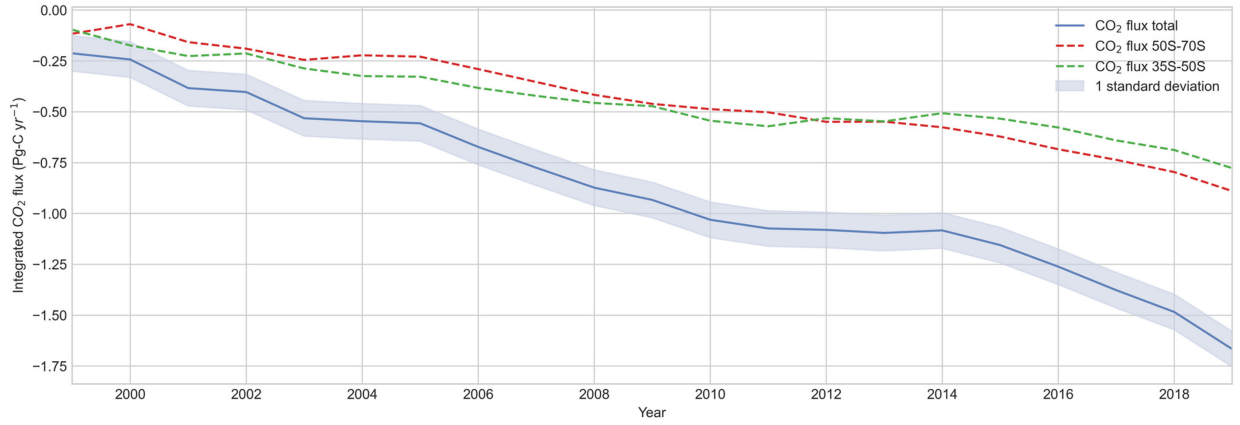


Fig. 12. Trends in CO₂ flux of the Southern Ocean from 1998–2018.

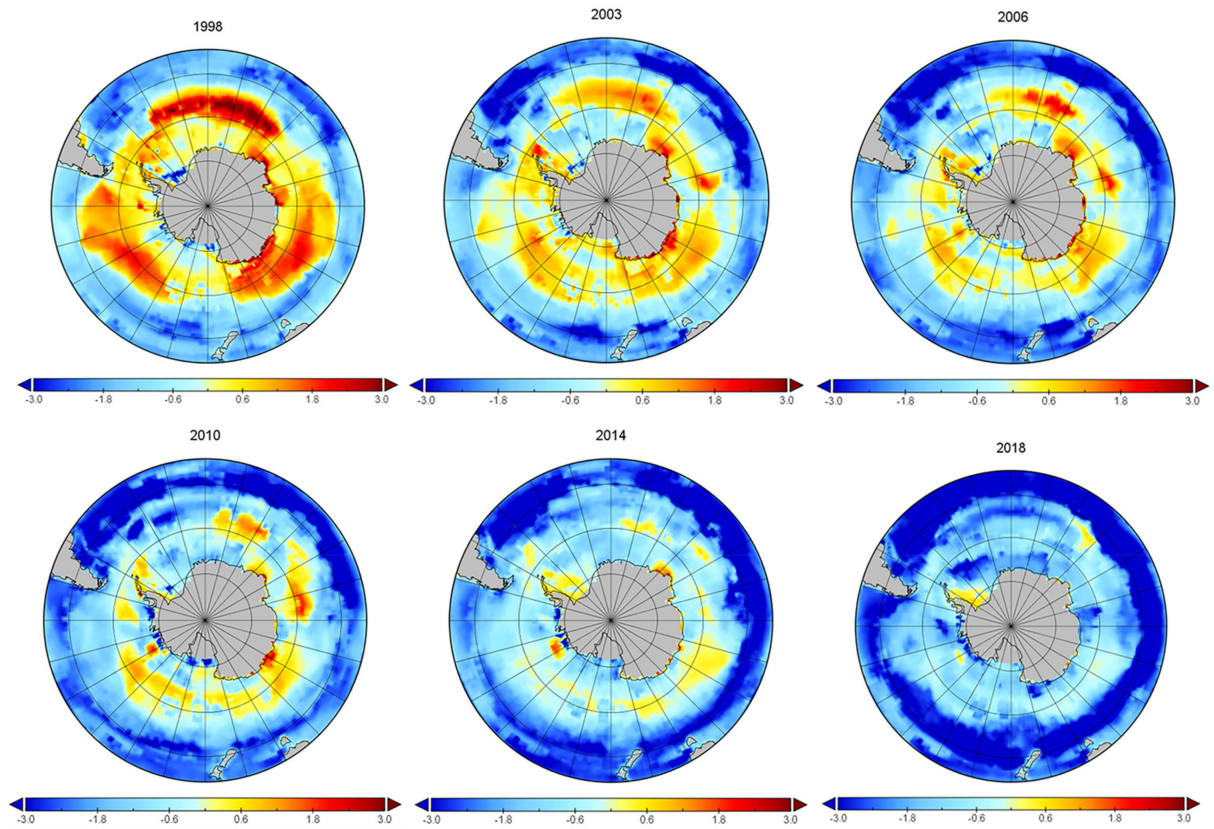


Fig. 13. Mean sea surface CO₂ fluxes (PgC) in 1998, 2003, 2006, 2010, 2014, and 2018 in the Southern Ocean.

The overall rate is relatively uniform. Previous studies have also shown that the $p\text{CO}_2$ of the Southern Ocean has maintained a strong growth momentum since 2002 [37], and our model captures this phenomenon well.

C. Variability in Sea–Air CO₂ Flux

Regarding the rate of change of Δ/CO_2 over space, the Southern Ocean is developing into a carbon sink. The red/black dots in Fig. 10 represent Δ/CO_2 regions toward positive/negative trends with high change rate. Based on the $p\text{CO}_2$ distribution of

the Southern Ocean since 1998, the inner ring region (50–70°S) gradually changed from a carbon source to a carbon sink, while the outer ring (35–50°S) has been a carbon sink, and the sinking has become more strengthened. By reconstructing the Southern Ocean, the CA–FFNN-based model was able to fully reproduce the temporal and spatial variations in the CO₂ flux and maintain consistency with other models concerning the evolution of intensity [38]. Using (12) to calculate the CO₂ flux, the Southern Ocean’s carbon sink was found to have changed substantially over the past two decades. The absorption of surface CO₂ in the Southern Ocean also displayed prominent

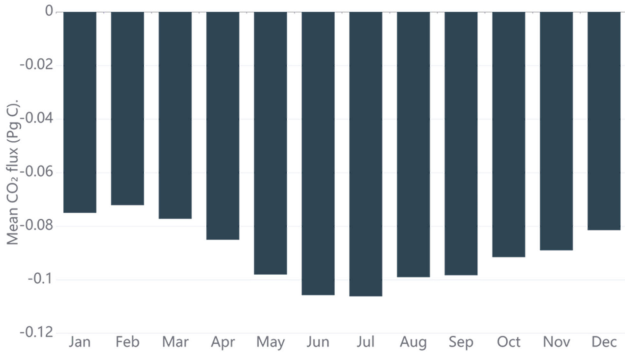


Fig. 14. Mean CO_2 flux of each month in $35\text{--}50^\circ\text{S}$ (PgC). The blue columns indicate that the Southern Ocean was a carbon sink.

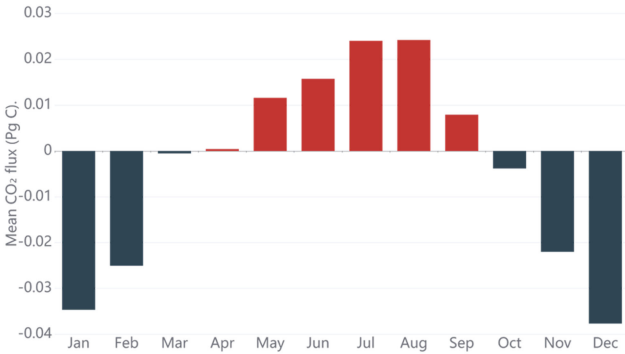


Fig. 15. Mean CO_2 flux of each month in $50\text{--}70^\circ\text{S}$ (PgC). The red columns indicate that the Southern Ocean was a carbon source, and the black columns indicate that it was a carbon sink.

seasonal characteristics, with the weakest absorption at the end of winter and the strongest in early summer (see Fig. 11). In this region, carbon sinks were saturated in the 1990s and restored to their original strengths in the early 21st century [39]. The data products could reconstruct this recovery in intensity and capture seasonal and regional changes in flux.

In terms of interannual changes, the carbon sink of the Southern Ocean increased from $-0.21 \text{ PgC}(\text{yr})^{-1}$ in 1998 to $-1.67 \text{ PgC}(\text{yr})^{-1}$ in 2018.

One standard deviation was used as an indicator of error

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

where x_i is the actual value, \bar{x} is the mean value of x , n is number of data, and the error range was within $\pm 0.087 \text{ PgC}(\text{yr})^{-1}$.

We found that the carbon sinks in this region have strengthened since 2000, except for when their development gradually slowed from 2010 to 2013; since then, they have recovered to their previous strengths. As shown in Fig. 12, this phenomenon was observed across different models [40].

The last period of stagnation occurred in the 1990s and there is much evidence to suggest that the stagnation during this period was closely related to the changes in the southern annular mode (SAM) [41]. However, the modeled stagnation was not strongly correlated with the SAM. Stability during this period was mainly due to the weakening of the carbon sink intensity from $35\text{--}50^\circ\text{S}$. Changes in this region have also been attributed to

the barometric asymmetry of the zonal waves 3 (ZW3) model [42]. Due to the lack of observational data, it is difficult for traditional models that rely on observations to capture such large interannual changes.

From the analysis of changes in the regional CO_2 flux, an obvious double-ring structure was observed in the fluxes of the Southern Ocean, as shown in Fig. 13, which is not always a carbon sink. The outer ring is the main carbon sink, at $35\text{--}50^\circ\text{S}$, and undertakes most CO_2 absorption. The outer ring is the main carbon sink, at 35.50°S , and undertakes most CO_2 absorption (Fig. 14). 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. 15. Based on our analysis of the interannual changes, this area is still a carbon sink.

However, this ring structure is gradually disappearing. Fig. 16 shows that, from 1998 to 2018, the carbon source in the inner ring area was disappearing each year, and the carbon sink in the outer ring was strengthening. As shown in Fig. 13, most Southern Ocean regions become carbon sinking regions, because the Δ/CO_2 in the Southern Ocean decreases significantly since 1998.

IV. CONCLUSION

In this study, we propose a deep-learning-based method for reconstructing $p\text{CO}_2$ data in the Southern Ocean that is generalizable for reconstructing regional data. The procedure is carried out in two steps. First, we collected the parameters that may affect changes in $p\text{CO}_2$ from the literature and experimental data and constructed a covariance matrix for all parameters. The correlation coefficients were then calculated, and the parameters with higher correlations and having an effect on the process change of $p\text{CO}_2$ were retained as the input FFNN in the second step after continuous and iterative calculation and optimization, the final model was constructed and used to reconstruct the $p\text{CO}_2$ data of the Southern Ocean with a monthly temporal resolution and a spatial resolution of $1^\circ \times 1^\circ$. First of all, we screen the parameters that affected $p\text{CO}_2$ changes according to the environmental characteristics of the Southern Ocean. Second, because the neural network has the advantage of interpolation in sparse regions, we build a new model based on the Southern Ocean's environmental data and neural network methods. Finally, in the Southern Ocean, the models have a better fitting than other models; compared with the measured data, the model maintained a favorable RMSE of $8.86 \mu\text{atm}$.

Analyses of the reconstructed results showed that the surface layer $p\text{CO}_2$ in the Southern Ocean changes seasonally and has increased yearly since 1998, such that the carbon sink is also increasing. The regional changes in flux exhibited a double-ring structure from $35\text{--}50^\circ\text{S}$, which is the main carbon sink area; seasonal carbon sources and sinks alternately appeared south of 50°S . Although our results are consistent with previous studies' results, due to the limited number of observational data in the Southern Ocean, the reconstructed surface $p\text{CO}_2$ products still need continuous validation. With the expanding observational granularity of the Southern Ocean, the algorithm will be further improved.

ACKNOWLEDGMENT

The $p\text{CO}_2$ data were provided by SOCAT (<https://www.socat.info/>), the SST data and SSS data were provided by gridded dataset of Global Ocean Heat Content Change (<http://159.226.119.60/cheng/>), the Chl-a data were provided by European Space Agency's GlobColour Project (<https://hermes.acri.fr/>), the MLD data were provided by French Research Institute for Exploitation of the Sea (<http://www.ifremer.fr/>), the wind data were provided by the European Centre for Medium-Range Weather Forecasts (<https://www.ecmwf.int/>), and the $x\text{CO}_2$ data were provided by United States National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratories marine boundary layer reference (<https://www.esrl.noaa.gov/>).

REFERENCES

- [1] A. K. Morrison, T. L. Frölicher, and J. L. Sarmiento, "Upwelling in the Southern Ocean," *Phys. Today*, vol. 68, no. 1, pp. 27–32, Dec. 2014, doi: [10.1063/PT.3.2654](https://doi.org/10.1063/PT.3.2654).
- [2] S. P. Khatriwala *et al.*, "Global ocean storage of anthropogenic carbon," *Biogeosciences*, vol. 9, pp. 8931–8988, 2012, doi: [10.3929/ethz-b-000062871](https://doi.org/10.3929/ethz-b-000062871).
- [3] I. Y. Fung, S. C. Doney, K. Lindsay, and J. John, "Evolution of carbon sinks in a changing climate," in *Proc. Nat. Acad. Sci.*, vol. 102, no. 32, pp. 11201–11206, Aug. 2005, doi: [10.1073/pnas.0504949102](https://doi.org/10.1073/pnas.0504949102).
- [4] A. A. Lacis, G. A. Schmidt, D. Rind, and R. A. Ruedy, "Atmospheric CO_2 : Principal control knob governing earth's temperature," *Science*, vol. 330, no. 6002, pp. 356–359, Oct. 2010, doi: [10.1126/science.1190653](https://doi.org/10.1126/science.1190653).
- [5] P. Friedlingstein *et al.*, "Global carbon budget 2019," *Earth Syst. Sci. Data*, vol. 11, no. 4, pp. 1783–1838, Dec. 2019, doi: [10.5194/essd-11-1783-2019](https://doi.org/10.5194/essd-11-1783-2019).
- [6] K. Caldeira and P. B. Duffy, "The role of the Southern Ocean in uptake and storage of anthropogenic carbon dioxide," *Science*, vol. 287, no. 5453, pp. 620–622, Jan. 2000, doi: [10.1126/science.287.5453.620](https://doi.org/10.1126/science.287.5453.620).
- [7] P. Landschützer, N. Gruber, and D. C. E. Bakker, "Decadal variations and trends of the global ocean carbon sink," *Global Biogeochem. Cycles*, vol. 30, no. 10, pp. 1396–1417, 2016, doi: [10.1002/2015GB005359](https://doi.org/10.1002/2015GB005359).
- [8] J. D. Majkut, B. R. Carter, T. L. Frölicher, C. O. Dufour, K. B. Rodgers, and J. L. Sarmiento, "An observing system simulation for Southern Ocean carbon dioxide uptake," *Philos. Trans. Roy. Soc. Math. Phys. Eng. Sci.*, vol. 372, no. 2019, Jul. 2014, Art. no. 20130046, doi: [10.1098/rsta.2013.0046](https://doi.org/10.1098/rsta.2013.0046).
- [9] P. Peylin *et al.*, "Multiple constraints on regional CO_2 flux variations over land and oceans," *Global Biogeochem. Cycles*, vol. 19, no. 1, 2005, doi: [10.1029/2003GB002214](https://doi.org/10.1029/2003GB002214).
- [10] C. Rödenbeck, "Estimating CO_2 sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport," *Max-Planck-Institut für Biogeochemie*, Jena, vol. 2005, no. 06, 2010.
- [11] C. E. Cosca *et al.*, "Seasonal and interannual CO_2 fluxes for the central and eastern equatorial Pacific Ocean as determined from $f\text{CO}_2$ -SST relationships," *J. Geophys. Res. Oceans*, vol. 108, no. C8, 2003, doi: [10.1029/2000jc000677](https://doi.org/10.1029/2000jc000677).
- [12] G.-H. Park *et al.*, "Variability of global net sea-air CO_2 fluxes over the last three decades using empirical relationships," *Tellus B Chem. Phys. Meteorol.*, vol. 62, no. 5, pp. 352–368, Jan. 2010, doi: [10.1111/j.1600-0889.2010.00498.x](https://doi.org/10.1111/j.1600-0889.2010.00498.x).
- [13] S. D. Jones, C. L. Quéré, C. Rödenbeck, A. C. Manning, and A. Olsen, "A statistical gap-filling method to interpolate global monthly surface ocean carbon dioxide data," *J. Adv. Model. Earth Syst.*, vol. 7, no. 4, pp. 1554–1575, 2015, doi: [10.1002/2014MS000416](https://doi.org/10.1002/2014MS000416).
- [14] N. P. Mongwe, M. Vichi, and P. M. S. Monteiro, "The seasonal cycle of CO_2 and CO_2 fluxes in the Southern Ocean: Diagnosing anomalies in CMIP5 Earth system models," *Biogeosciences*, vol. 15, no. 9, pp. 2851–2872, May 2018, doi: [10.5194/bg-15-2851-2018](https://doi.org/10.5194/bg-15-2851-2018).
- [15] D. C. E. Bakker *et al.*, "Data-based estimates of the ocean carbon sink variability—First results of the Surface Ocean $p\text{CO}_2$ mapping inter-comparison (SOCOM)," *Biogeosciences*, vol. 12, no. 23, pp. 7251–7251, Dec. 2015.
- [16] L. Gregor, S. Kok, and P. M. S. Monteiro, "Empirical methods for the estimation of Southern Ocean CO_2 : Support vector and random forest regression," *Biogeosciences*, vol. 14, no. 23, pp. 5551–5569, Dec. 2017, doi: [10.5194/bg-14-5551-2017](https://doi.org/10.5194/bg-14-5551-2017).
- [17] P. Landschützer *et al.*, "A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink," *Biogeosciences*, vol. 10, no. 11, pp. 7793–7815, 2013, doi: [10.3929/ethz-b-000078204](https://doi.org/10.3929/ethz-b-000078204).
- [18] A. Denvil-Sommer, M. Gehlen, M. Vrac, and C. Mejia, "LSCE-FFNN-v1: A two-step neural network model for the reconstruction of surface ocean PCO_2 over the global ocean," *Geosci. Model Develop.*, vol. 12, no. 5, pp. 2091–2105, May 2019, doi: [10.5194/gmd-12-2091-2019](https://doi.org/10.5194/gmd-12-2091-2019).
- [19] P. Landschützer *et al.*, "The reinvigoration of the Southern Ocean carbon sink," *Science*, vol. 349, no. 6253, pp. 1221–1224, Sep. 2015, doi: [10.1126/science.aab2620](https://doi.org/10.1126/science.aab2620).
- [20] L. Cheng *et al.*, "Improved estimates of changes in upper ocean salinity and the hydrological cycle," *J. Climate*, vol. 33, pp. 10357–10381, 2020, doi: [10.1175/JCLI-D-20-0366.1](https://doi.org/10.1175/JCLI-D-20-0366.1).
- [21] A. Körtzinger, "Determination of carbon dioxide partial pressure ($p\text{CO}_2$)," in *Methods of Seawater Analysis*. Hoboken, NJ, USA: Wiley, 2007, pp. 149–158.
- [22] *Vapor Pressure of Water: Calculator | Definition | Formulas*. Accessed: Jan. 3, 2021. [Online]. Available: <https://www.omnicalculator.com/chemistry/vapour-pressure-of-water>
- [23] L. Pinho, C. M. Duarte, H. Marotta, and A. Enrich-Prast, "Temperature dependence of the relationship between PCO_2 and dissolved organic carbon in lakes," *Biogeosciences*, vol. 13, no. 3, pp. 865–871, Feb. 2016, doi: [10.5194/bg-13-865-2016](https://doi.org/10.5194/bg-13-865-2016).
- [24] S. Ruder, "An overview of gradient descent optimization algorithms," 2017, *arXiv:1609.04747*.
- [25] A. J. Watson *et al.*, "Revised estimates of ocean-atmosphere CO_2 flux are consistent with ocean carbon inventory," *Nature Commun.*, vol. 11, Sep. 2020, Art. no. 4422, doi: [10.1038/s41467-020-18203-3](https://doi.org/10.1038/s41467-020-18203-3).
- [26] R. Wanninkhof, "Relationship between wind speed and gas exchange over the ocean," *J. Geophys. Res. Oceans*, vol. 97, no. C5, pp. 7373–7382, 1992, doi: [10.1029/92JC00188](https://doi.org/10.1029/92JC00188).
- [27] R. Wanninkhof *et al.*, "Global ocean carbon uptake: Magnitude, variability and trends," *Biogeosciences*, vol. 10, no. 3, pp. 1983–2000, Mar. 2013, doi: [10.3929/ethz-b-000065788](https://doi.org/10.3929/ethz-b-000065788).
- [28] T. Takahashi *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. II, Topical Stud. Oceanogr.*, vol. 56, no. 8, pp. 554–577, Apr. 2009, doi: [10.1016/j.dsr2.2008.12.009](https://doi.org/10.1016/j.dsr2.2008.12.009).
- [29] T. Takahashi *et al.*, "The changing carbon cycle in the Southern Ocean," *Oceanography*, vol. 25, no. 3, pp. 26–37, 2012, doi: [10.5670/oceanog.2012.71](https://doi.org/10.5670/oceanog.2012.71).
- [30] N. Metzl, C. Brunet, A. Jabaud-Jan, A. Poisson, and B. Schauer, "Summer and winter air-sea CO_2 fluxes in the Southern Ocean," *Deep Sea Res. I, Oceanogr. Res. Papers*, vol. 53, no. 9, pp. 1548–1563, Sep. 2006, doi: [10.1016/j.dsr.2006.07.006](https://doi.org/10.1016/j.dsr.2006.07.006).
- [31] S. J. Thomalla, N. Fauchereau, S. Swart, and P. M. S. Monteiro, "Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean," *Biogeosciences*, vol. 8, no. 10, pp. 2849–2866, Oct. 2011, doi: [10.5194/bg-8-2849-2011](https://doi.org/10.5194/bg-8-2849-2011).
- [32] P. Landschützer, N. Gruber, D. C. E. Bakker, I. Stemmler, and K. D. Six, "Strengthening seasonal marine CO_2 variations due to increasing atmospheric CO_2 ," *Nat. Climate Change*, vol. 8, pp. 146–150, Feb. 2018, doi: [10.1038/s41558-017-0057-x](https://doi.org/10.1038/s41558-017-0057-x).
- [33] E. H. Shadwick, T. W. Trull, B. Tilbrook, A. J. Sutton, E. Schulz, and C. L. Sabine, "Seasonality of biological and physical controls on surface ocean CO_2 from hourly observations at the Southern Ocean time series site south of Australia," *Global Biogeochem. Cycles*, vol. 29, no. 2, pp. 223–238, 2015, doi: [10.1002/2014GB004906](https://doi.org/10.1002/2014GB004906).
- [34] R. F. Anderson *et al.*, "Wind-driven upwelling in the Southern Ocean and the deglacial rise in atmospheric CO_2 ," *Science*, vol. 323, no. 5920, pp. 1443–1448, Mar. 2009, doi: [10.1126/science.1167441](https://doi.org/10.1126/science.1167441).
- [35] N. Gruber, P. Landschützer, and N. S. Lovenduski, "The variable Southern Ocean carbon sink," *Annu. Rev. Mar. Sci.*, vol. 11, no. 1, pp. 159–186, 2019, doi: [10.1146/annurev-marine-121916-063407](https://doi.org/10.1146/annurev-marine-121916-063407).
- [36] S. Takao *et al.*, "Effects of phytoplankton community composition and productivity on sea surface PCO_2 variations in the Southern Ocean," *Deep Sea Res. I, Oceanogr. Res. Papers*, vol. 160, Jun. 2020, Art. no. 103263, doi: [10.1016/j.dsr.2020.103263](https://doi.org/10.1016/j.dsr.2020.103263).
- [37] A. Silvano, "Changes in the Southern Ocean," *Nat. Geosci.*, vol. 13, pp. 4–5, Jan. 2020, doi: [10.1038/s41561-019-0516-2](https://doi.org/10.1038/s41561-019-0516-2).

- [38] R. Ritter *et al.*, “Observation-based trends of the Southern Ocean carbon sink,” *Geophys. Res. Lett.*, vol. 44, no. 24, pp. 12339–12348, 2017, doi: [10.1002/2017GL074837](https://doi.org/10.1002/2017GL074837).
- [39] T. L. Frölicher, J. L. Sarmiento, D. J. Paynter, J. P. Dunne, J. P. Krasting, and M. Winton, “Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models,” *J. Clim.*, vol. 28, no. 2, pp. 862–886, Jan. 2015, doi: [10.1175/JCLI-D-14-00117.1](https://doi.org/10.1175/JCLI-D-14-00117.1).
- [40] L. Gregor, A. D. Lebehoh, S. Kok, and P. M. Scheel Monteiro, “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.*, vol. 12, no. 12, pp. 5113–5136, Dec. 2019, doi: [10.5194/gmd-12-5113-2019](https://doi.org/10.5194/gmd-12-5113-2019).
- [41] L. Gregor, S. Kok, and P. M. S. Monteiro, “Interannual drivers of the seasonal cycle of CO_2 in the Southern Ocean,” *Biogeosciences*, vol. 15, no. 8, pp. 2361–2378, Apr. 2018, doi: [10.5194/bg-15-2361-2018](https://doi.org/10.5194/bg-15-2361-2018).
- [42] L. Keppler and P. Landschützer, “Regional wind variability modulates the Southern Ocean carbon sink,” *Sci. Rep.*, vol. 9, 2019, Art. no. 7384, doi: [10.1038/s41598-019-43826-y](https://doi.org/10.1038/s41598-019-43826-y).



Yanjun Wang received the B.S. degree in computer science and technology from PLA Artillery College, Hefei, China, in 2010, and the master's degree in computer system structure from the Harbin University of Science and Technology, Harbin, China, in 2013.

Since 2013, he has been an Engineer at Marine Science Data Center, Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China. His current research interests include ocean data visualization, big data analysis, and artificial intelligence applications.



Xiaofeng Li (Fellow, IEEE) received the B.S. degree in optical engineering from Zhejiang University, Hangzhou, China, in 1985, the M.S. degree in physical oceanography from the First Institute of Oceanography, Qingdao, China, in 1992, and the Ph.D. degree in physical oceanography from North Carolina State University, Raleigh, NC, USA, in 1997.

From 1997 to 2019, he was with the National Environmental Satellite, Data, and Information Service, National Ocean and Atmospheric Administration, Washington, DC, USA. His research interests

include synthetic aperture radar (SAR) applications in oceanography and marine meteorology, artificial intelligence oceanography, big data, and satellite image processing.

Dr. Li is an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the *International Journal of Remote Sensing*. He is an Editorial Board Member of the *International Journal of Digital Earth*, *Big Earth Data*, and the *Journal of Oceanology and Limnology*. He is also the Executive Editor-in-Chief of the *Journal of Remote Sensing* (a Science Partner Journal).



Jinming Song received the B.S. degree in chemistry from Jilin University, Changchun, China, in 1985, and the M.S. degree in chemical oceanography and the Ph.D. degree in marine geochemistry from the Institute of Oceanology, Chinese Academy of Sciences (IOCAS), Qingdao, China, in 1988 and 1998, respectively.

Since 1998, he has been a Senior Research Professor and Tutor of Ph.D., IOCAS. He has authored or coauthored more than 450 scientific papers and nine monographs. His research interests include marine

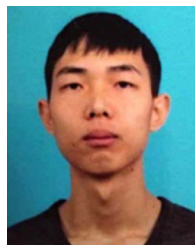
eco-environmental evolution, marine biogeochemical processes, and marine big data.

Dr. Song is an Associate Editor-in-Chief of the *Oceanologia et Limnologia Sinica*, the *Acta Oceanologica Sinica*, the *Acta Ecologica Sinica*, the *Journal of Tropical Oceanography*, etc.



Xuegang Li received the B.S. and M.S. degrees in geochemistry from the Changchun College of Geology, Changchun, China, in 1992 and 1996, respectively, and the Ph.D. degree in marine chemistry from the Graduate University of Chinese Academy of Sciences, Qingdao, China, in 2006.

Since 2006, he has been with the Institute of Oceanology, Chinese Academy of Science, Qingdao, China, where he is currently a Professor. His research interests include marine biogeochemistry cycle, acidification, and eutrophication.



Guorong Zhong received the B.S. degree in marine science from Dalian Maritime University, Dalian, China, in 2018. He is currently working toward the M.S. degree and Ph.D. degree in marine chemistry from the Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China.



Bin Zhang (Member, IEEE) received the B.S. degree in marine technology (ocean remote sensing and data processing) from the Tianjin University of Science and Technology, Tianjin, China, in 2009, and the Ph.D. degrees in marine geology from the Ocean University of China, Qingdao, China, in 2015.

From 2015 to 2017, he was a Postdoctoral Researcher with the Institute of Oceanology, Chinese Academy of Sciences, Qingdao, China, where he was an Engineer, from 2017 to 2019, and has been a Senior Engineer, since 2019. His current research interests

include ocean data quality control, big data analysis and mining, and artificial intelligence applications.