frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 18 May 2023
DOI 10.3389/fmars.2023.1181095

OPEN ACCESS

EDITED BY
Haiyong Zheng,
Ocean University of China, China

REVIEWED BY
Wei-Jen Huang,
National Sun Yat-sen University, Taiwan
Abhra Chanda,
Jadavpur University, India

*CORRESPONDENCE
Chunli Liu
chunliliu@sdu.edu.cn

RECEIVED 07 March 2023
ACCEPTED 02 May 2023
PUBLISHED 18 May 2023

CITATION
Li W, Liu C, Zhai W, Liu H and Ma W (2023) Remote sensing and machine learning method to support sea surface pCO_2 estimation in the Yellow Sea. *Front. Mar. Sci.* 10:1181095. doi: 10.3389/fmars.2023.1181095

Remote sensing and machine learning method to support sea surface pCO_2 estimation in the Yellow Sea

Wei Li[1], Chunli Liu[1*], Weidong Zhai[2], Huizeng Liu[3] and Wenjuan Ma[1]

[1]Marine College, Shandong University, Weihai, China, [2]Frontier Research Center, Southern Marine Science and Engineering Guangdong Laboratory, Zhuhai, China, [3]Institute for Advanced Study, Shenzhen University, Shenzhen, China

With global climate changing, the carbon dioxide (CO_2) absorption rates increased in marginal seas. Due to the limited availability of *in-situ* spatial and temporal distribution data, the current status of the sea surface carbon dioxide partial pressure (pCO_2) in the Yellow Sea is unclear. Therefore, a pCO_2 model based on a random forest algorithm has been developed, which was trained and tested using 14 cruise data sets from 2011 to 2019, and remote sensing satellite sea surface temperature, chlorophyll concentration, diffuse attenuation of downwelling irradiance, and in-situ salinity were used as the input variables. The seasonal and interannual variations of modeled pCO_2 were discussed from January 2003 and December 2021 in the Yellow Sea. The results showed that the model developed for this study performed well, with a root mean square difference (RMSD) of 43 µatm and a coefficient of determination (R^2) of 0.67. Moreover, modeled pCO_2 increased at a rate of 0.36 µatm year^{-1} (R^2 = 0.27, $p < 0.05$) in the YS, which is much slower than the rate of atmospheric pCO_2 (pCO_2^{air}) rise. The reason behind it needs further investigation. Compared with pCO_2 from other datasets, the pCO_2 derived from the RF model exhibited greater consistency with the in-situ pCO_2 (RMSD = 55 µatm). In general, the RF model has significant improvement over the previous models and the global data sets.

KEYWORDS

machine learning, random forest, remote sensing, the Yellow Sea, pCO_2

1 Introduction

The rapid growth of fossil fuel usage and industry has increased the atmospheric carbon dioxide (CO_2) concentration by approximately 40% since the Industrial Revolution (Landschützer et al., 2014; Friedlingstein et al., 2022). Global oceans absorb 30% of the CO_2 released by industry and human activities and they are a significant sink for atmospheric CO_2. Coastal seas cover 7% of the oceanic surface area but the sea-air exchange carbon fluxes (FCO_2) comprise approximately 25–50% of the global oceans (Laruelle et al., 2018), and thus they play important roles in absorbing atmospheric CO_2 (Dai et al., 2022). Due to the effects of the complex physical environment and biological activities, great errors occur in estimations of FCO_2 in coastal seas (Landschützer et al., 2018; Mignot et al., 2022). Therefore, estimating sea surface carbon dioxide partial pressure (pCO_2) accurately for coastal seas is critical for precisely estimating the global FCO_2 (Laruelle et al., 2018).

In general, pCO_2 is regulated by thermodynamic effects, biogeochemical effects, mixing effects, and air–sea exchange effects (Liu et al., 2019; Ye et al., 2022). Some environmental variables can characterize these four effects. In particular, the sea surface temperature (SST, °C) directly reflects thermodynamic effects, while the chlorophyll concentration (Chl, mg m^{-3}) and diffuse attenuation of downwelling irradiance (Kd, m^{-1}) can indicate biogeochemical effects on the surface pCO_2. In addition, the SST, salinity (SSS, psu), and mixed layer depth (MLD, m) are closely related to mixing effects, and the wind speed can characterize the sea–air exchange process (Gu et al., 2021).

Due to their unique advantage in terms of high spatiotemporal resolution, satellite approaches are efficient for observing pCO_2. In previous studies, both semi-analytical (Hales et al., 2012; Bai et al., 2015; Chen et al., 2017) and empirical approaches (Lohrenz et al., 2010; Tao et al., 2012; Qin et al., 2014; Chen et al., 2016; Chen et al., 2019; Fu et al., 2020) were used to estimate the sea surface pCO_2. Many studies have used satellite data to estimate the sea surface pCO_2, but recent studies also examined and compared the capability of semi-analytical and empirical algorithms for estimating the coastal pCO_2 (Chen et al., 2017; Chen et al., 2019). However, the high spatiotemporal variability and diversity of pCO_2, the inaccuracy of satellite data, and limited availability of *in-situ* pCO_2 data from coastal seas make it challenging to establish a model of pCO_2. Several efforts have been made to construct various algorithms or models, but the satellite-derived pCO_2 in coastal seas generally has higher uncertainty than that for open seas, and the root mean square difference (RMSD) can be as high as 90 µatm (Chen et al., 2019).

The Yellow Sea (YS) is an important coastal sea in the west Pacific Ocean. The pCO_2 in the YS has considerable seasonal variations and an unbalanced spatial distribution (Wang and Zhai, 2021). For example, extremely high pCO_2 values have been observed during the summer in the center of the YS, whereas extremely low pCO_2 values have been observed in the southwestern YS (Qu et al., 2014; Zhai, 2018). Since the 1980s, many studies have investigated carbonate, pCO_2, and FCO_2 in the YS (Xue et al., 2011; Qu et al., 2014; Zhai et al., 2014; Zhai, 2018; Choi et al., 2019; Deng et al., 2021). However, accurately quantifying pCO_2 and FCO_2 in the YS remains a challenge. In particular, Wang and Zhai (2021) indicated that the YS is a carbon sink and FCO_2 is about −0.5 ± 1.9 mol m^{-2} $year^{-1}$, whereas Qu et al. (2014) suggested that the YS is a carbon source. In addition, the physical and biological conditions in coastal seas have changed due to rapid climate change. For example, SST and Chl have increased (Liu et al., 2021; Lu et al., 2021). These variations will have influenced the changes in the sea surface pCO_2. Indeed, recent studies showed that the CO_2 absorption rates increased in some coastal seas (Li and Zhai, 2019; Xiong et al., 2020). To the best of our knowledge, no previous studies have quantified the long-term trend in the carbon absorption capacity of the YS due to the lack of *in-situ* pCO_2 data over the entire YS. Thus, in order to accurately quantify the pCO_2 in the YS and understand the response of the pCO_2 to global climate change, we developed an inversion model of pCO_2 in the YS in the present study. Two previous remote sensing studies investigated the pCO_2 in the YS (Tao et al., 2012; Qin et al., 2014), and both used *in-situ* SST and Chl data to establish multiple polynomial regression (MPR) models. This modeling method is simple but the errors are large. Therefore, in the present study, we aimed: (1) to develop machine learning models for accurately deriving pCO_2 from satellite remote sensing data; and (2) to analyze the long-term trend in the pCO_2 during 2003–2021 in the YS.

2 Materials and methods

2.1 Study area

The YS is a semi-enclosed shelf shallow sea (29.5°N–40.5°N, 118.5°E–126.5°E) located west of the Liaodong Peninsula and east of the Korean Peninsula (Figure 1). The mean water depth is 44 m (Liu et al., 2009). The areas and depths of the North Yellow Sea (NYS) and South Yellow Sea (SYS) are 70 × 10^3 km^2 and 38 m, and

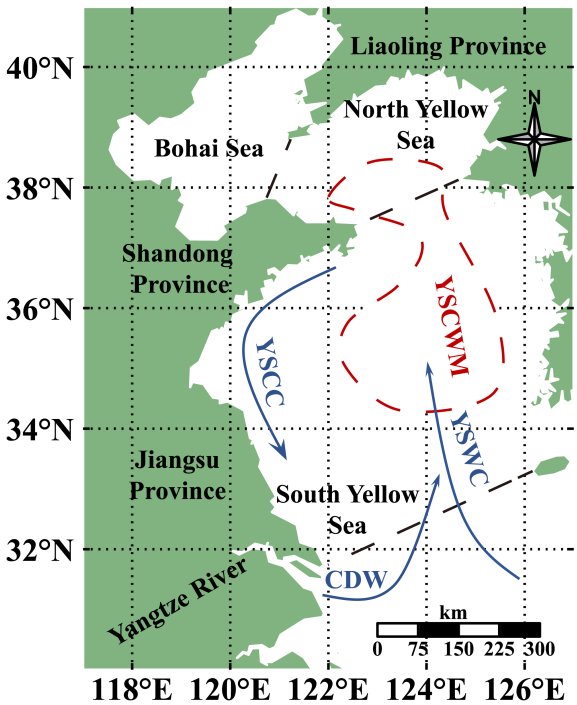

FIGURE 1
Chart of the study region. The three black dashed lines represent the boundaries between the North Yellow Sea (NYS) and Bohai Sea, the NYS and South Yellow Sea (SYS), and the SYS and East China Sea (ECS).

300×10^3 km^2 and 44 m, respectively. The climate and ocean circulations exhibit strong seasonality due to the effect of the East Asian Monsoon (Ding et al., 2018). In the winter, the YS is mainly influenced by the Yellow Sea Warm Current (YSWC) and the Yellow Sea Coastal Current. The Yellow Sea Warm Current invades the YS from south to north, and brings warm ocean water to the YS, which makes some regions into carbon sources in the YS (Xue et al., 2011). In the summer, the central YS is occupied by the Yellow Sea Cold Water Mass (YSCWM) and there is a strong thermocline above the YSCWM. In addition, the northeastern extension of the Changjiang Dilution Water (CDW) carries a considerable amount of nutrients to the west of the SYS, and this region sustains high phytoplankton production, thereby leading to lower pCO_2 values (Qu et al., 2014). Overall, the YS current is an important factor that affects pCO_2. A previous study showed that the coastal currents in the YS have strengthened in recent years (Liu S, et al., 2023), which may affect the interannual variation in the pCO_2 in the YS.

The YS is surrounded by rapidly developing economic regions, and the rapid development of mariculture has caused severe environmental problems, such as phytoplankton blooms and changes in ocean acidification. Therefore, the carbon cycle process in the YS is managed by both the coastal hydrodynamics and human activities (Choi et al., 2019).

2.2 Data sets

We collected fugacity of CO_2 (fCO_2) data from 14 cruises conducted between 2011 and 2019, which homogenously covered the entire annual cycle (Table 1). Data were derived from four cruises conducted in 2019 by Yu et al. (2022), and data collected from 10 other cruises by Wang and Zhai (2021).

fCO_2 was conversed into pCO_2 using the following formula (1):

$$fCO_2 = pCO_2 \cdot \exp\left(p \cdot \frac{B + 2\sigma}{RT}\right) \tag{1}$$

where p is the total pressure (Pa), R is a gas constant (8.314 J K^{-1} mol^{-1}), T is the absolute temperature of the sea surface (K), and B and σ are rectification coefficients, which are calculated with formulas (2) and (3).

$$B = (-1636.75 + 12.0408 \times T - 3.27957 \times 10^{-2}T^2 + 3.16528 \times 10^{-5}T^3) \times 10^{-6} \tag{2}$$

$$\sigma = (57.7 - 0.118T) \times 10^{-6} \tag{3}$$

The inverse model of pCO_2 in the YS was established with Chl, SST, SSS, and Kd as input variables. In addition, Julday (Jday, or day of year) was selected as an input to highlight the periodical changes in pCO_2 (Lefevre et al., 2005; Signorini et al., 2013). Chl and Kd, SST, and SSS were used to represent biochemical, thermodynamic, and mixing effects on the sea surface pCO_2, respectively. Level 3 8-days and monthly SST (°C), Chl (mg m^{-3}), and Kd (m^{-1}) data sets were obtained from Moderate Resolution Imaging Spectroradiometer (MODIS)-Aqua for January 2003 and December 2021 (https://oceancolor.gsfc.nasa.gov/) at a spatial resolution of 4 km. SSS data observed directly by ocean color sensor satellites are not available, so *in-situ* SSS data were used to develop the model in this study. The HYbrid Coordinate Ocean Model (HYCOM) SSS data set (monthly products with a 4-km resolution) was selected to derive maps of the sea surface pCO_2 (available from: https://www.hycom.org/). In addition, the gridded atmospheric pCO_2 (pCO_2^{air}) data set (daily, with a spatial resolution of 2° × 2.5°) provided by Rödenbeck et al. (2013) was used (available from: http://www.bgc-jena.mpg.de/SOCOM/).

Due to the influence of cloud cover, sensor technology, atmospheric correction algorithms, and other factors, satellite remote sensing data have a high missing rate in time and space. Therefore, satellite data were interpolated using Data Interpolating Empirical Orthogonal Functions (DINEOF) to obtain more matching pairs. A pixel located at 122°E and 33.2°N was selected to verify the rationality of the reconstructed data. The reconstructions agreed with the original data and complemented the missing data well (Figure 2).

Satellite data were matched with *in-situ* data according to (Le et al., 2019). Briefly, a time window of ± 8 days was applied between the *in-situ* and satellite-derived data. In addition, in order to filter sensor and algorithm noise, the median of a 3 × 3-pixel box was focused on every sample point. If the coefficient of variation for the effective pixels in the 3 × 3-pixel box was ≤ 0.4, the extracted data were used to develop the model together with the *in-situ* data. Finally, we obtained 638 matched pairs from 14 cruises (Figure 3).

2.3 Model training and testing, and model selection

The 638 matched pairs were split into training and test data sets in a stratified random manner, where they accounted for 80% and 20% of the pairs, respectively. Histograms showing the

TABLE 1 Comparison of two empirical modeling approaches. .

Approach	RMSD (μatm)	R^2	MAE (μatm)	MAPE
PSO-SVR	43	0.63	35	9%
	54	0.44	40	11%
RF	**34**	**0.82**	**24**	**6%**
	43	**0.67**	**32**	**8%**

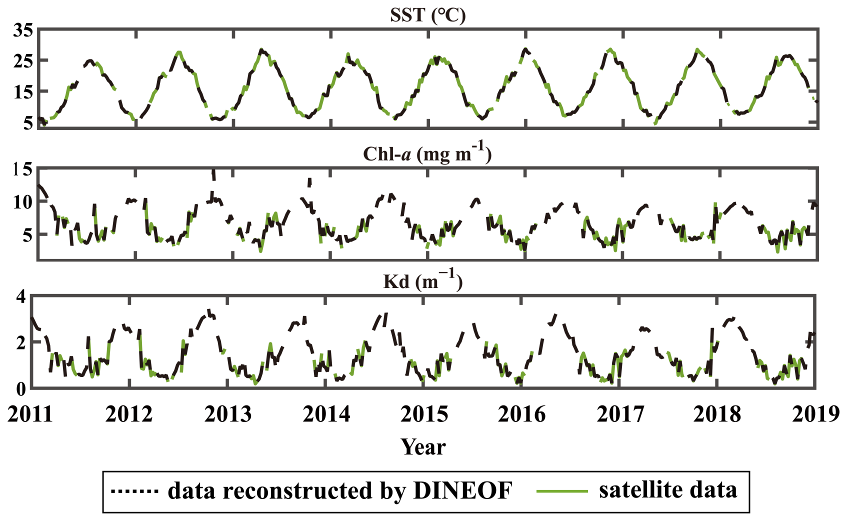

FIGURE 2
Comparison of reconstructed and original data.

distributions of the sample points in the training and test data sets are presented in Figure 4. Evaluation indicators comprising the RMSD, coefficient of determination (R^2), mean absolute error (MAE), and mean absolute percentage error (MAPE) were employed to quantify the reliability of the pCO_2 model.

Two machine learning algorithms comprising Random Forest (RF) and particle swarm optimization-support vector regression (PSO-SVR) were used to develop sea surface pCO_2 models because of their high generalizability for nonlinear relationships (Mountrakis et al., 2011). The inversion model was established using identical data sets. The algorithm was determined as formula (4).

$$\begin{aligned} pCO_2 &= f(\text{input variables}) \\ &= f(\text{SST, Kd, SSS, Chl}, \cos(2\pi(\text{Julday} - \gamma)/365) \end{aligned} \tag{4}$$

The value of γ was optimized iteratively (0 to 365) until the RMSD reached a minimum value.

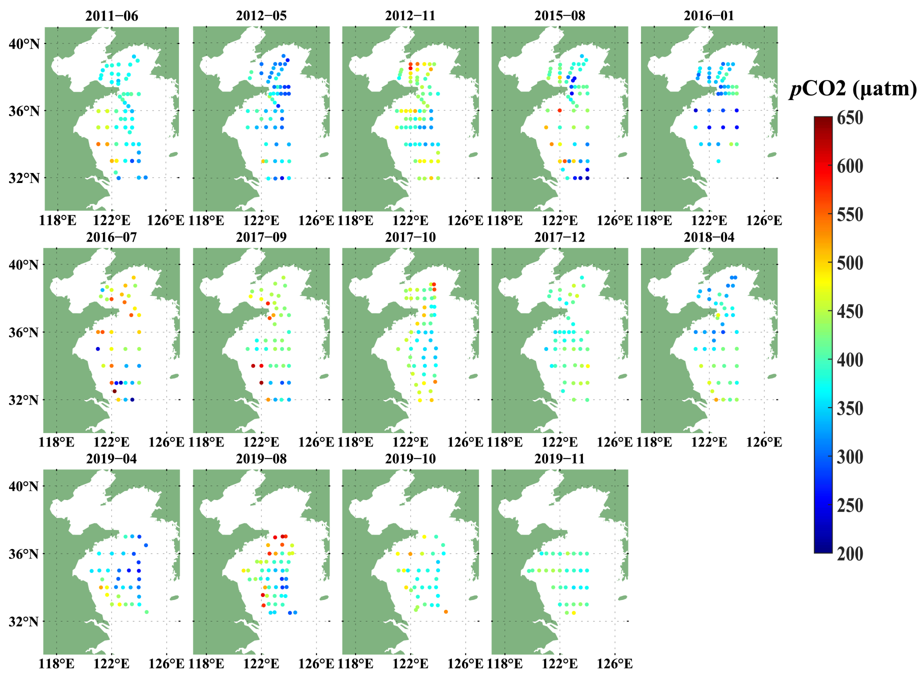

FIGURE 3
Spatial distribution of 638 matched pairs.

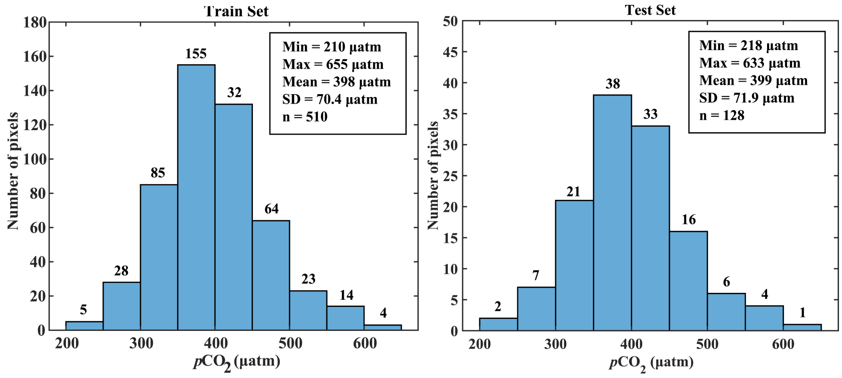

FIGURE 4
Histograms showing the distributions of the sample points in the training and test data sets.

2.4 Random forest

The RF consists of multiple decision trees, where the structure of a single decision tree is based on a group of training data (Breiman, 2001). In RF, a bootstrap strategy is used to conduct resampling from the original data sets to produce multiple subgroups. The structure regression trees are then obtained for every subgroup, and the final output is the mean of the outputs of all regression trees.

RF model development (Figure 5) requires the determination of three customized parameters: the number of randomly selected variables for constructing the tree (mtry), the minimum number of terminal nodes for each tree (node size), and the number of trees (ntree) (Sun et al., 2016).

The node size was set to 5 because this is a common value for regression models (Sun et al., 2016). The grid search method was used to determine the RF parameters ntree and mtry (Figure 6). The optimal values were determined with the minimal RMSD, and 4 and 200 were selected as the best mtry and ntree values, respectively, for the RF model.

2.5 Model sensitivity to input variables

Sensitivity analysis was conducted to assess the sensitivity of the model to the inherent uncertainties in SST, SSS, Chl, and Kd. The original pCO_2 (using the original inputs) was compared with the new pCO_2 (using inputs with extra added uncertainties) derived from the same RF model to identify the model's sensitivity to the uncertainty in these inputs. Only one input variable was changed in each analysis and the remaining variables were kept the same. Statistical parameters comprising the mean bias (MB), mean ratio (MR), RMSD, and R^2 were applied to quantify the sensitivity.

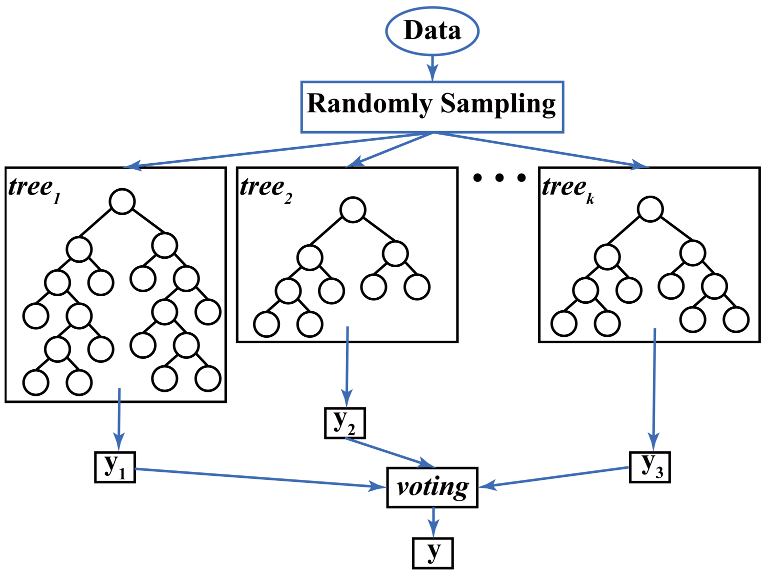

FIGURE 5
General Random Forest model development process.

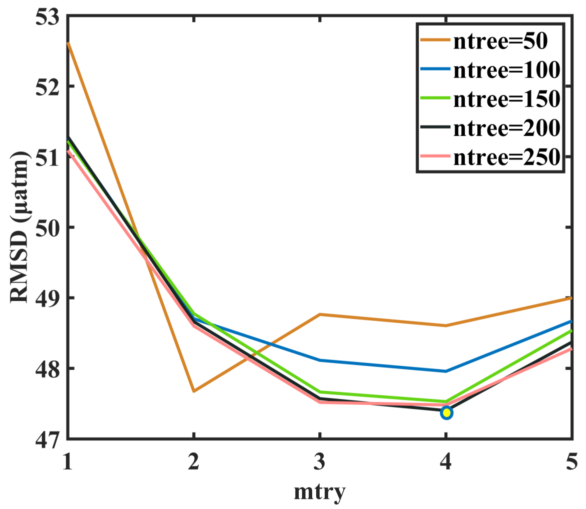

FIGURE 6
Influence of mtry and ntree on RMSD.

The uncertainties of environmental variables were determined by referring to published studies. In particular, the uncertainty of remote sensing SST is ≤ 1°C (Hao et al., 2017), the uncertainty of HYCOM SSS is about 0.5 when SSS is more than 32, the uncertainty of HYCOM SSS is about 3 when SSS is less than 32 (Jang et al., 2022), and the uncertainties of Chl and Kd are 32% and 48%, respectively (Cui et al., 2014). Thus, we used ± 1°C, ± 1, ± 30%, and ± 45% as the uncertainties of SST, SSS, Chl, and Kd, respectively.

3 Results

3.1 Model performance

Table 1 shows that RF outperformed PSO-SVR. The R^2 and RMSD values were 0.82 and 34 μatm, and 0.67 and 43 μatm for the model training and test data sets, respectively.

The sea surface pCO$_2$ predicted by the RF model was slightly underestimated when the sea surface pCO$_2$ was larger than 500 μatm, and slightly overestimated when pCO$_2$ was smaller than 300 μatm (Figure 7). The pCO$_2$ values estimated by the model varied in the range of 250–550 μatm, with some larger than 550 μatm and lower than 250 μatm. A histogram showing the residuals (modeled pCO$_2$ minus field pCO$_2$) is presented in Figure 7, which demonstrates that 82.45% of the residuals were within the interval of ± 50, i.e., the observed 50 μatm pCO$_2$ standard deviation.

3.2 Model sensitivity

Statistically, when a bias of +1°C was applied to the SST input, the RF model overestimated the sea surface pCO$_2$ slightly (RMSD = 10 μatm, R^2 = 0.96, MB = 3 μatm), and when a bias of −1°C was applied to the SST input, the RF model underestimated the sea surface pCO$_2$ slightly (R^2 = 0.96, RMSD = 10 μatm, MB = −2 μatm) (Figure 8). These results suggest that pCO$_2$ increased with SST, and vice versa, which is consistent with the relationship between temperature and pCO$_2$ in thermodynamics.

Compared with the SST, the RF pCO$_2$ model was more sensitive to the uncertainties in SSS. Moreover, the RF model was more sensitive to lower SSS values, where a change of −1 in SSS resulted in a substantial decrease in the predicted pCO$_2$. In particular, with input +1 uncertainty in SSS, the RF pCO$_2$ model tended to overestimate the sea surface pCO$_2$ (R^2 = 0.83, RMSD = 20 μatm, and MB = 5 μatm) and with input −1 uncertainty in SSS, the RF pCO$_2$ model tended to greatly underestimate the sea surface pCO$_2$ (R^2 = 0.73, RMSD = 30 μatm, and MB = −16 μatm).

Similar to SST, the RF pCO$_2$ model exhibited minor sensitivity to Chl. When all data were used in the calculations with +30% uncertainties added, the RF model slightly overestimated pCO$_2$ (R^2 = 0.96, RMSD = 10 μatm, and MB = 2 μatm). With input −30% uncertainties in Chl, the RF model slightly underestimated pCO$_2$ (R^2 = 0.95, RMSD = 11 μatm, and MB = −3 μatm). Similarly, the RF pCO$_2$ model was insensitive to Kd. With +45% and −45% uncertainties added in Kd, the new pCO$_2$ was not very different from the original pCO$_2$. In particular, with a bias of +45% uncertainty added to Kd, the RF slightly overestimated the surface pCO$_2$ (R^2 = 0.93, RMSD = 16 μatm, and MB = 9 μatm), and with a bias of −45% uncertainty added, the RF pCO$_2$ model slightly underestimated the pCO$_2$ (R^2 = 0.89, RMSD = 18 μatm, and MB= −8 μatm).

The sensitivity of the RF model was different according to the uncertainty in each environment variable, but the differences introduced by each variable were generally within the range of the uncertainty of the model itself.

3.3 Seasonal and interannual variations in pCO$_2$ in the YS

The RF model was applied to monthly MODIS and HYCOM data for the period between January 2003 and December 2021 to generate monthly climatological maps and determine the annual trend in pCO$_2$ in the YS (Figure 9).

Spatially, due to the effects of the hydrology environment and terrestrial organic matter, the pCO$_2$ values tended to decrease from the nearshore to central areas, and the highest pCO$_2$ values were observed in the SYS. Seasonally, there were apparent variations in pCO$_2$ throughout the YS (Figure 9). Statistically, the average sea surface pCO$_2$ values were 377 ± 7 μatm, 430 ± 6 μatm, 426 ± 11 μatm, and 378 ± 10 μatm in the spring, summer, autumn, and winter, respectively. In addition to these seasonal patterns, more complex variations were found in the spring and autumn (Figure S1). In most years, pCO$_2$ decreased in March because of phytoplankton blooms, and increased in September or November because of the collapsing seasonal stratification.

The annual mean sea surface pCO$_2$ values were extracted to explore the interannual variation. The results showed that the surface pCO$_2$ values in the YS increased between 2003 and 2021 at a rate of 0.36 μatm year^{-1} (R^2 = 0.27, p< 0.05, N = 19) (Figure 10).

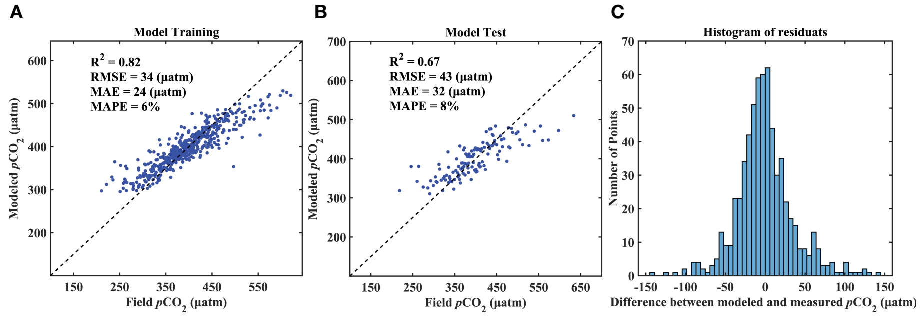

FIGURE 7
Performance evaluation for RF using **(A)** training and **(B)** test data sets; and **(C)** histogram of residuals.

According to the model sensitivity analysis results in section 3.2, when a bias of +1°C was applied to the SST input, the RF model overestimated pCO_2 by 10 μatm. The annual rate of change in the SST determined by the remote sensing products was 0.039°C year^{-1} (Figure S2). Therefore, increasing the SST approximately led to an increase in the pCO_2 at a rate of 0.39 μatm year^{-1} in the YS. The pCO_2 in the YS has increased in the past 19 years, but its rate of increase was lower than that for pCO_2^{air} (with a rate of 2.31 μatm year^{-1}; R^2 = 0.99, p< 0.01, N = 19) in the same period (Figure S3). Therefore, the ΔpCO_2 (sea surface pCO_2– pCO_2^{air}) exhibited a remarkable decreasing trend with a rate of −1.95 μatm year^{-1} (R^2 = 0.92, p< 0.01, N = 19).

Moreover, the spatial trends in pCO_2 were obtained by calculating the trend for each grid in pCO_2 (Figure 10B). In general, pCO_2 increased in most regions of the YS, with a range from 0 to 2.78 μatm year^{-1} from 2003 to 2021. Decreasing trends were also found in some regions. For example, pCO_2 decreased in the NYS and the runoff area of the Changjiang River. These results indicate that the NYS and runoff area of the Changjiang River have more substantial carbon absorption capacities. Both pCO_2 and Chl tended to decrease in the runoff area of the Changjiang River (Figures 10B, S4). Therefore, the decrease in the transportation of terrestrial organic matter might be the main reason for the decrease in pCO_2 in this area, which might alleviate the seasonal hypoxia phenomenon.

4 Discussion

4.1 Evaluation based on comparisons with field observations of sea surface pCO_2

Two algorithms were tested to establish models for estimating pCO_2. The best RMSD and R^2 values for the model were 43 μatm and 0.67 in the YS, respectively (Figure 7). The accuracy of four data sets were evaluated by comparing with field observations of sea surface pCO_2. The resolutions, names of the four data sets, and comparisons of the results are shown in Table 2.

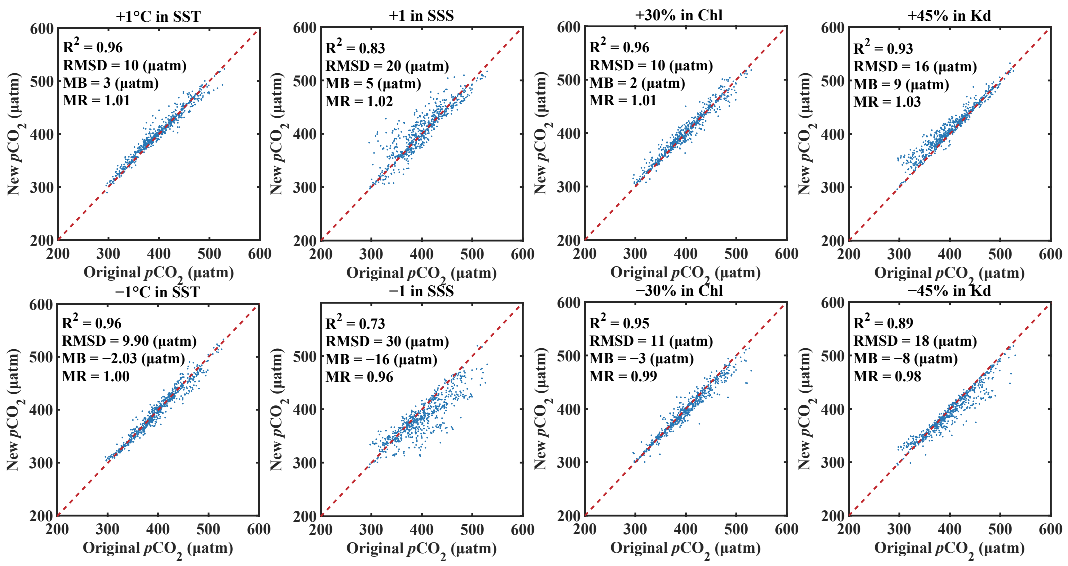

FIGURE 8
Sensitivity of RF model to the uncertainties in SST, SSS, Chl, and Kd.

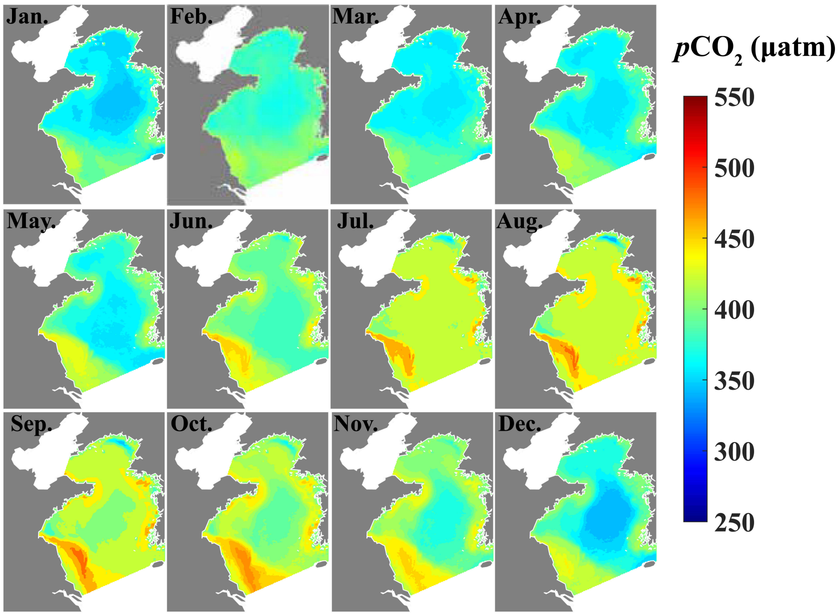

FIGURE 9
Monthly climatological maps of pCO_2 in the YS from January 2003 to December 2021.

Figure 11 shows scatter diagrams to compare the results. The pCO_2 derived from the RF model exhibited greater consistency (RMSD = 55 μatm) with the *in-situ* pCO_2 than CSIR-ML6 (RMSD = 71 μatm), MPI-SOMFNN (RMSD = 82 μatm), and $SatCO_2$ (RMSD = 119 μatm). The significant underestimation of the field pCO_2 by $SatCO_2$ was predictable because the algorithm was originally developed for the ECS and it may not be applicable to the YS. Significant differences between the global pCO_2 products and *in-situ* data in coastal seas were expected (Landschützer et al., 2020). Moreover, CSIR and ML6 were not effective at matching the pCO_2 in the YS, as shown by the number of scatter points in Figure 11. The comparison of four products showed that the RF model was the optimal method for estimating pCO_2 in the YS because the root mean square difference was less than those with the other three products (CSIR-ML6, MPI-SOMFNN, and $SatCO_2$). Understanding the variations in pCO_2 can provide greater insights into the response of the carbon absorption capacity to climate change in the YS. Erroneous estimates may be obtained in coastal seas if global pCO_2 products are used, which might affect quantification of the longer-term trends in global carbon budgets.

4.2 Satellite estimation of pCO_2 in coastal seas

Due to its unique advantage in terms of high spatiotemporal resolution, satellite remote sensing is an effective method for observing the sea surface pCO_2. Table 2 lists some inversion models for pCO_2 in coastal seas. The maximum RMSD for these models was 45.19 μatm. Tao et al. (2012) and Qin et al. (2014)

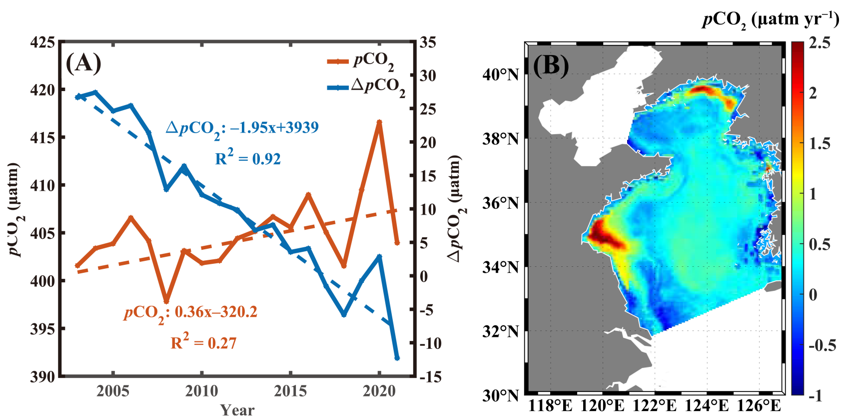

FIGURE 10
(A) Long-term trends in regional average pCO_2 and ΔpCO_2 ($pCO_2 - pCO_2^{air}$); and **(B)** spatial trends in pCO_2 during 2003–2021.

TABLE 2 Published models based on remote sensing of sea surface pCO_2 and global pCO_2 products.

Reference	Model or data set	Study area	Spatial resolution/Model inputs	RMSD (H atm)
Gregor et al. (2019)	CSIR-ML6	Yellow Sea	1° x1°	71
Landschützer et al. (2016)	MPI-SOMFNN	Yellow Sea	1° x1°	82
Bai et al. (2015)	$SatCO_2$	Yellow Sea	1.6 km	119
this study	RF	Yellow Sea	4 km	55
Parard et al. (2014)	SOM	Baltic Sea	SST. Chl. CDOM, NPP, MLD. Jday	35
Tao et al. (2012)	MPR	Yellow Sea and Bohai Sea	SST. Chl	31.74
Qin et al. (2014)	MPR	Yellow Sea	SST. Chl	16.68–21.46
Chen et al. (2016)	MNR	West Florida Shelf	SST. Kd. Chl. Iday	<11.79
Liu J, et al. (2023)	MNR	East China Sea	SST. SSS, Chl. Jday, LAT. LON	3.73-45.19

SOM, Self Organizing Map; MNR, Multi-variate Nonlinear Regression; NPP, Net Primary Production; CDOM, Colored Dissolved Organic Matter; LAT, Latitude; LON, Longitude.

established pCO_2 estimation models based on MPR using the *in-situ* SST and Chl, and the RMSD values for the two models were 15.82 −31.7 and 16.68–21.46, respectively, and both were less than 43. The error was small for the two models, mainly because the *in-situ* data used for modeling were mostly located in the YS center, with few data located in the nearshore area. The MPR-based inversion model was developed using the same training data sets employed in the present study, and the error was much larger than 43 µatm. Overall, the error was acceptable for the RF model developed in this study. The RMSD of the model for estimating the surface pCO_2 in the YS was higher than that in other marginal seas due to the following three reasons. (1) The uncertainty of satellite data and field pCO_2. In the YS, the error of satellite remote sensing Kd and Chl data can reach 48%, and 32%, respectively (Cui et al., 2014). Moreover, the pCO_2 data used in this study were converted from fCO_2, and fCO_2 was estimated using the dissolved inorganic carbon and total alkalinity. The uncertainty in the pCO_2 obtained by using this method is ± 5%, which is larger compared with ± 1% using directly measured pCO_2 data (Wang and Zhai, 2021). (2) The hydrological complexity of the YS environment leads to a wide range of sea

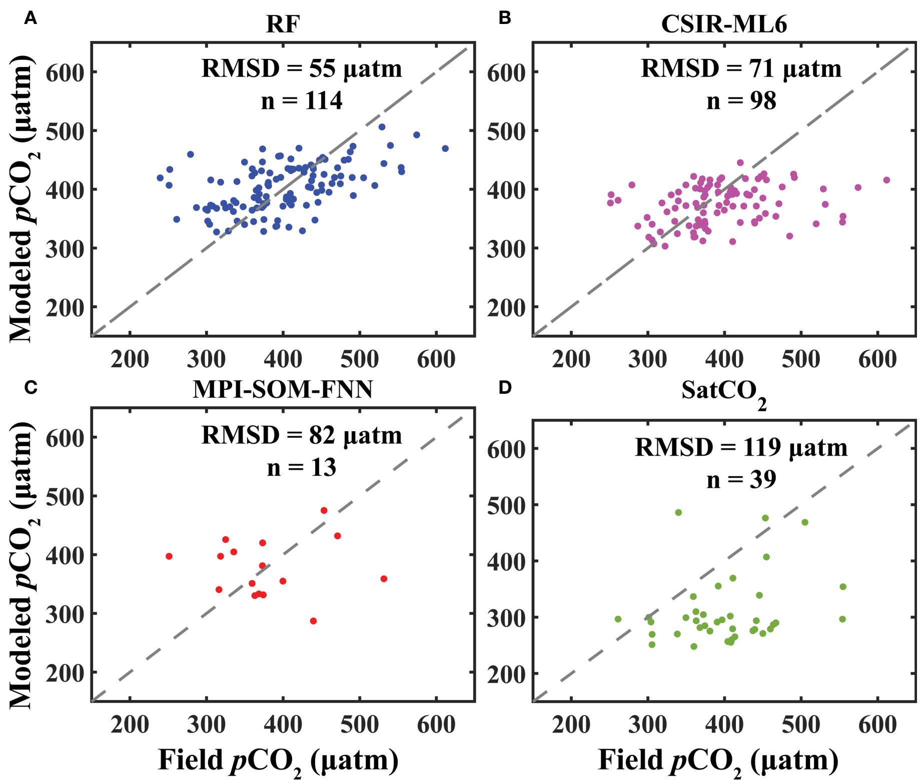

FIGURE 11
Scatter plots of pCO_2 obtained from **(A)** RF model, **(B)** CSIR-ML6, and **(C)** MPI-SOMFNN; and **(D)** $SatCO_2$ against the field pCO_2 in the test set.

surface pCO_2 changes. In particular, the magnitude of the change in pCO_2 in the YS is 450 µatm (Figure 3), but only about 350 µatm in the Gulf of Mexico (Fu et al., 2020) and the Gulf of Maine (Signorini et al., 2013). The performance of the model constructed for the YS was similar to that of a model for the Baltic Sea (RMSD = 47.48 µatm, R^2 = 0.63) (Zhang et al., 2021), where pCO_2 ranged from 100 −600 µatm. (3) Importantly, the RF model needed to include all of the processes from 2011 to 2019. These three reasons explain why estimating pCO_2 is very difficult in the YS compared with other marginal seas, and thus the error is large.

4.3 Advantages and limitations of RF model

The comparisons of the models based on the two algorithms showed that the RF algorithm was advantageous for inverting the sea surface pCO_2 in the YS (Table 1; Figure 11), and the uncertainty was less than 50 µatm. However, the RF model still has some problems.

First, in the eastern YS, the seasonal variation in the pCO_2 obtained from the RF model differed compared with the *in-situ* pCO_2. Choi et al. (2019) found that pCO_2 tended to increase from May to February in the Southeastern YS. However, the maximum pCO_2 obtained by RF inversion was in August (Figure 9). Wang and Zhai (2021) divided the YS region west of 124°E into four regions and analyzed the seasonal variations in the pCO_2. They found that the maximum values in the four regions occurred in July, September, or October, with none in February. Due to the effect of hydrodynamics and other factors, the seasonal patterns in the pCO_2 differ greatly in the eastern YS and western YS. Therefore, the differences in the seasonal variations in pCO_2 may be explained by only using *in-situ* data for the area located west of 124°E for modeling, and thus the model was unable to fully identify the pCO_2 control process.

Second, using the RF model to compute the interannual trends in the pCO_2 could introduce uncertainties. The homogenously collected cruise data covered the whole annual period (Table 3). The variation in pCO_2 was influenced by physical and biogeochemical processes in the sea, and the increase in atmospheric CO_2 (Xue et al., 2016). However, the parameters (SST, Chl, Kd, and SSS) used in this study could only characterize the physical and biogeochemical processes in the sea. If changes in pCO_2 caused by increases in the atmospheric CO_2 could not be captured implicitly by one or more of the four parameters (SST, SSS, Chl, and Kd), uncertainties would be introduced when computing the interannual trend in the pCO_2 (Chen et al., 2019). The long-term trend of SST in the YS was influenced by regional climate change (Park et al., 2015), that is to say, the change of SST included the change of atmospheric CO_2 internally and implicitly, therefore, the increase in the SST appeared to can capture the effects of increasing atmospheric CO_2 on the pCO_2, the interannual trend was still credible to some extent.

Third, in the present study, RF performed poorly at simulating data from both ends of the data sets (underestimation for high values and overestimation for low values) (Figure 7), which may be explained as follows. First, due to the features of the algorithm itself, RF averages the results for all regression trees. The underestimation of extreme values and overestimation of small values appears to be a common problem for RF regression models (Čeh et al., 2018; Zimmerman et al., 2018; Wolfensberger et al., 2021). Second, the training data sets contained very few extreme pCO_2 values and they were underrepresented in the RF model, thereby leading to a more mean-biased output from the RF model.

In general, the problems with the RF model described above were caused by the unbalanced distributions of the modeling data sets. The number of extreme pCO_2 values (>550 µatm or<250 µatm) was relatively small in the field measurements (only 4.7%) but it did not seem to affect the interannual variation in the pCO_2. However, extreme pCO_2 is an influential component of the carbon cycle and it has significant impacts on the health of marine ecosystems. Therefore, it is very necessary to accurately estimate the extreme pCO_2. The crucial limitation of RF model is that its ability to estimate new pCO_2 is limited by the range of the training data set. That mean it can not estimate the pCO_2 beyond the range of the training data set (no extrapolation). Therefore, a better RF model may be developed by using a data set with

TABLE 3 Cruises and statistics for SST, SSS, and sea surface pCO_2 measurements used for model training and test (mean ± standard deviation).

Season	Time	SST (°C)	SSS	pCO_2 (µatm)	Number of observations
Spring	2012–05 2018–04 2019–04	10.4 ± 2.9	32.1 ± 0.8	361 ± 58	133
Summer	2011–06 2015–08 2016–07 2019–08	23.0 ± 3.7	31.1 ± 1.1	410 ± 88	204
Autumn	2012–11 2017–09 2017–10 2019–10 2019–11	19.3 ± 3.7	31.5 ± 0.5	425 ± 58	231
Winter	2016–01 2017–12	8.6 ± 3.1	32.2 ± 0.3	373 ± 51	92
average/Total samples	—	17.2 ± 6.6	31.6 ± 0.9	400 ± 73	660

a wider range of variation, which can improve the reproducibility of the RF model for extreme values. Therefore, we suggest that the modeling data set need to include all pCO_2 values that can be matched to the satellite data, some extreme values in the *in-situ* data sets should not be arbitrarily deleted (excluding the low and high values caused by measurement errors).

5 Conclusions

In this study, we constructed a RF model of the YS with SST, SSS, Chl, Kd, and Julday as the inputs. The RF model performed well at estimating pCO_2, with an RMSD of 43 µatm and R^2 of 0.67. The RF model was applied to satellite data from between 2003 and 2021 to obtain a 19-year time sequence of pCO_2 in the YS. Spatially, except for the eastern YS, the spatial pCO_2 distributions derived by the RF model matched with the *in-situ* data. According to the interannual changes, the sea surface pCO_2 increased in most regions of the YS, but there were differences among the regions, with decreased trends in the pCO_2 in the NYS and the runoff area of the Changjiang River, which appears to contrast with the background global warming and increasing atmospheric CO_2 concentration. The present study is the first to using machine learning methods to estimate the pCO_2, and also the first to determine the long-term trend in the pCO_2 in the YS. Future research should focus on obtaining balanced *in-situ* pCO_2 data and coupling the RF model with a mechanistic model to develop more accurate pCO_2 models. In addition, the reasons for the increasing trend in the pCO_2 in the YS should be explored.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

WL: Methodology, Software, Writing-original draft. CL: Conceptualization, Resources, Writing-review & editing. WZ: Investigation, Writing-review & editing. HL: Software, Writing-review & editing. WM: Formal analysis. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the following research grants: the National Natural Science Foundation of China-Shandong joint fund (U1806203), Shandong Provincial Natural Science Foundation (ZR2020MD098), Shandong Universities Interdisciplinary Research and Innovation Team of Young Scholars (2020QNQT20).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2023.1181095/full#supplementary-material

References

Bai, Y., Cai, W. J., He, X. Q., Zhai, W. D., Pan, D. L., Dai, M. H., et al. (2015). A mechanistic semi-analytical method for remotely sensing sea surface pCO_2 in river-dominated coastal oceans: a case study from the East China Sea. *J. Geophys. Res. Ocean.* 120, 2331–2349. doi: 10.1002/2014JC010632

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Čeh, M., Kilibarda, M., Lisec, A., and Bajat, B. (2018). Estimating the performance of random forest versus multiple regression for predicting prices of the apartments. *ISPRS. Int. J. Geo-Inf.* 7, 168. doi: 10.3390/ijgi7050168

Chen, S. L., Hu, C. M., Barnes, B. B., Wanninkhof, R., Cai, W. J., Barbero, L., et al. (2019). A machine learning approach to estimate surface ocean pCO_2 from satellite measurements. *Remote Sens. Environ.* 228, 203–226. doi: 10.1016/j.rse.2019.04.019

Chen, S. L., Hu, C. M., Byrne, R. H., Robbins, L. L., and Yang, B. (2016). Remote estimation of surface pCO_2 on the West Florida shelf. *Cont. Shelf. Res.* 128, 10–25. doi: 10.1016/j.csr.2016.09.004

Chen, S. L., Hu, C. M., Cai, W. J., and Yang, B. (2017). Estimating surface pCO_2 in the northern gulf of Mexico: which remote sensing model to use? *Cont. Shelf. Res.* 151, 94–110. doi: 10.1016/j.csr.2017.10.013

Choi, Y., Kim, D., Cho, S., and Kim, T. W. (2019). Southeastern yellow Sea as a sink for atmospheric carbon dioxide. *Mar. pollut. Bull.* 149, 110550. doi: 10.1016/j.marpolbul.2019.110550

Cui, T. W., Zhang, J., Tang, J. W., Sathyendranath, S., Groom, S., Ma, Y., et al. (2014). Assessment of satellite ocean color products of MERIS, MODIS and SeaWiFS along the East China coast (in the yellow Sea and East China Sea). *ISPRS-J. Photogramm. Remote Sens.* 87, 137–151. doi: 10.1016/j.isprsjprs.2013.10.013

Dai, M. H., Su, J. Z., Zhao, Y. Y., Hofmann, E. E., Cao, Z. M., Cai, W. J., et al. (2022). Carbon fluxes in the coastal ocean: synthesis, boundary processes, and future trends. *Annu. Rev. Earth Planet. Sci.* 50, 593–626. doi: 10.1146/annurev-earth-032320-090746

Deng, X., Zhang, G. L., Xin, M., Liu, C. Y., and Cai, W. J. (2021). Carbonate chemistry variability in the southern yellow Sea and East China Sea during spring of 2017 and summer of 2018. *Sci. Total. Environ.* 779, 146376. doi: 10.1016/j.scitotenv.2021.146376

Ding, Y., Bao, X. W., Yao, Z. G., Song, D. H., Song, J., Gao, J., et al. (2018). Effect of coastal-trapped waves on the synoptic variations of the yellow Sea warm current during winter. *Cont. Shelf. Res.* 167, 14–31. doi: 10.1016/j.csr.2018.08.003

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., et al. (2022). Global carbon budget 2021. *Earth Syst. Sci. Data* 14, 1917–2005. doi: 10.5194/essd-14-1917-2022

Fu, Z. Y., Hu, L. S., Chen, Z. D., Zhang, F., Shi, Z., Hu, B. F., et al. (2020). Estimating spatial and temporal variation in ocean surface $p\mathrm{CO_2}$ in the gulf of Mexico using remote sensing and machine learning techniques. *Sci. Total. Environ.* 745, 140965. doi: 10.1016/j.scitotenv.2020.140965

Gregor, L., Lebehot, A. D., Kok, S., and Scheel Monteiro, P. M. (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, 5113–5136. doi: 10.5194/gmd-12-5113-2019

Gu, Y. Y., Katul, G. G., and Cassar, N. (2021). The intensifying role of high wind speeds on air-sea carbon dioxide exchange. *Geophys. Res. Lett.* 48, e2020GL090713. doi: 10.1029/2020GL090713

Hales, B., Strutton, P. G., Saraceno, M., Letelier, R., Takahashi, T., Feely, R., et al. (2012). Satellite-based prediction of $p\mathrm{CO_2}$ in coastal waters of the eastern north pacific. *Prog. Oceanogr.* 103, 1–15. doi: 10.1016/j.pocean.2012.03.001

Hao, Y. L., Cui, T. W., Singh, V. P., Zhang, J., Yu, R. H., and Zhang, Z. L. (2017). Validation of MODIS Sea surface temperature product in the coastal waters of the yellow Sea. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* 10, 1667–1680. doi: 10.1109/JSTARS.2017.2651951

Jang, E., Kim, Y. J., Im, J., Park, Y.-G., and Sung, T. (2022). Global sea surface salinity via the synergistic use of SMAP satellite and HYCOM data based on machine learning. *Remote Sens. Environ.* 273, 112980. doi: 10.1016/j.rse.2022.112980

Landschützer, P., Gruber, N., and Bakker, D. C. (2016). Decadal variations and trends of the global ocean carbon sink. *Glob. Biogeochem. Cycle* 30, 1396–1417. doi: 10.1002/2015GB005359

Landschützer, P., Gruber, N., Bakker, D. C. E., and Schuster, U. (2014). Recent variability of the global ocean carbon sink. *Glob. Biogeochem. Cycle* 28, 927–949. doi: 10.1002/2014GB004853

Landschützer, P., Gruber, N., Bakker, D. C. E., Stemmler, I., and Six, K. D. (2018). Strengthening seasonal marine CO_2 variations due to increasing atmospheric CO_2 nature climate change. *Nat. Clim. Chang.* 8, 146–150. doi: 10.1038/s41558-017-0057-x

Landschützer, P., Laruelle, G. G., Roobaert, A., and Regnier, P. (2020). A uniform $p\mathrm{CO_2}$ climatology combining open and coastal oceans. *Earth Syst. Sci. Data* 12, 2537–2553. doi: 10.5194/essd-12-2537-2020

Laruelle, G. G., Cai, W. J., Hu, X. P., Gruber, N., Mackenzie, F. T., and Regnier, P. (2018). Continental shelves as a variable but increasing global sink for atmospheric carbon dioxide. *Nat. Commun.* 9, 1–11. doi: 10.1038/s41467-017-02738-z

Le, C. F., Gao, Y. Y., Cai, W. J., Lehrter, J. C., Bai, Y., and Jiang, Z. P. (2019). Estimating summer sea surface $p\mathrm{CO_2}$ on a river-dominated continental shelf using a satellite-based semi-mechanistic model. *Remote Sens. Environ.* 225, 115–126. doi: 10.1016/j.rse.2019.02.023

Lefevre, N., Watson, A. J., and Watson, A. R. (2005). A comparison of multiple regression and neural network techniques for mapping *in situ* $p\mathrm{CO_2}$ data. *Tellus. Ser. B-Chem. Phys. Meteorol.* 57, 375–384. doi: 10.3402/tellusb.v57i5.16565

Li, C. L., and Zhai, W. D. (2019). Decomposing monthly declines in subsurface-water pH and aragonite saturation state from spring to autumn in the north yellow Sea. *Cont. Shelf. Res.* 185, 37–50. doi: 10.1016/j.csr.2018.11.003

Liu, J., Bellerby, R. G., Zhu, Q., and Ge, J. Z. (2023). Estimation of sea surface $p\mathrm{CO_2}$ and air–sea CO_2 flux in the East China Sea using *in-situ* and satellite data over the period 2000–2016. *Cont. Shelf. Res.* 254, 104879. doi: 10.1016/j.csr.2022.104879

Liu, Q., Dong, X., Chen, J. S., Guo, X. H., Zhang, Z. R., Xu, Y., et al. (2019). Diurnal to interannual variability of sea surface $p\mathrm{CO_2}$ and its controls in a turbid tidal-driven nearshore system in the vicinity of the East China Sea based on buoy observations. *Mar. Chem.* 216, 103690. doi: 10.1016/j.marchem.2019.103690

Liu, S. C., Luo, Z. P., Wang, Y. W., Rao, Q. R., Zhang, X. S., Yu, B., et al. (2023). Interannual variation in winter thermal front to the east of the Shandong peninsula in the yellow Sea. *J. Sea. Res.* 193, 102370. doi: 10.1016/j.seares.2023.102370

Liu, Z. Y., Wei, H., Lozovatsky, I., and Fernando, H. (2009). Late summer stratification, internal waves, and turbulence in the yellow Sea. *J. Mar. Syst.* 77, 459–472. doi: 10.1016/j.jmarsys.2008.11.001

Liu, J. L., Xia, J., Zhuang, M. M., Zhang, J. H., Sun, Y. Q., Tong, Y. C., et al. (2021). Golden seaweed tides accumulated in pyropia aquaculture areas are becoming a normal phenomenon in the yellow Sea of China. *Sci. Total. Environ.* 774, 145726. doi: 10.1016/j.scitotenv.2021.145726

Lohrenz, S. E., Cai, W. J., Chen, F. Z., Chen, X. G., and Tuel, M. (2010). Seasonal variability in air-sea fluxes of CO_2 in a river-influenced coastal margin. *J. Geophys. Res. Ocean.* 115 (C10). doi: 10.1029/2009jc005608

Lu, X. L., Liu, C. L., Niu, Y., and Yu, S. X. (2021). Long-term and regional variability of phytoplankton biomass and its physical oceanographic parameters in the yellow Sea, China. *Estuar. Coast. Shelf. Sci.* 260, 107497. doi: 10.1016/j.ecss.2021.107497

Mignot, A., von Schuckmann, K., Landschützer, P., Gasparin, F., van Gennip, S., Perruche, C., et al. (2022). Decrease in air-sea CO_2 fluxes caused by persistent marine heatwaves. *Nat. Commun.* 13, 4300. doi: 10.1038/s41467-022-31983-0

Mountrakis, G., Im, J., and Ogole, C. (2011). Support vector machines in remote sensing: a review. *ISPRS-J. Photogramm. Remote Sens.* 66, 247–259. doi: 10.1016/j.isprsjprs.2010.11.001

Parard, G., Charantonis, A., and Rutgerson, A. (2014). Remote sensing algorithm for sea surface CO_2 in the Baltic Sea. *Biogeosci. Discuss.* 11, 12255–12294. doi: 10.5194/bgd-11-12255-2014

Park, K.-A., Lee, E.-Y., Chang, E., and Hong, S. (2015). Spatial and temporal variability of sea surface temperature and warming trends in the yellow Sea. *J. Mar. Syst.* 143, 24–38. doi: 10.1016/j.jmarsys.2014.10.013

Qin, B. Y., Tao, Z., Li, Z. W., and Yang, X. F. (2014). Seasonal changes and controlling factors of sea surface $\mathrm{pCO_2}$ in the yellow Sea. *IOP. Conf. Ser.: Earth Environ. Sci.* 17, 012025. doi: 10.1088/1755–1315/17/1/012025

Qu, B. X., Song, J. M., Yuan, H. M., Li, X. G., and Li, N. (2014). 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. doi: 10.1016/j.csr.2014.02.001

Rödenbeck, C., Keeling, R. F., Bakker, D. C., Metzl, N., Olsen, A., Sabine, C., et al. (2013). Global surface-ocean $p\mathrm{CO_2}$ and sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme. *Ocean. Sci.* 9, 193–216. doi: 10.5194/os-9-193-2013

Signorini, S. R., Mannino, A., Najjar, R. G.Jr., Friedrichs, M. A. M., Cai, W. J., Salisbury, J., et al. (2013). Surface ocean $p\mathrm{CO_2}$ seasonality and sea-air CO_2 flux estimates for the north American east coast. *J. Geophys. Res. Ocean.* 118, 5439–5460. doi: 10.1002/jgrc.20369

Sun, H. W., Gui, D. W., Yan, B. W., Liu, Y., Liao, W. H., Zhu, Y., et al. (2016). Assessing the potential of random forest method for estimating solar radiation using air pollution index. *Energy Conv. Manage.* 119, 121–129. doi: 10.1016/j.enconman.2016.04.051

Tao, Z., Qin, B. Y., Li, Z. W., and Yang, X. F. (2012). Satellite observations of the partial pressure of carbon dioxide in the surface water of the huanghai Sea and the bohai Sea. *Acta Oceanol. Sin.* 31, 67–73. doi: 10.1007/s13131-012-0207-y

Wang, S. Y., and Zhai, W. D. (2021). Regional differences in seasonal variation of air–sea CO_2 exchange in the yellow Sea. *Cont. Shelf. Res.* 218, 104393. doi: 10.1016/j.csr.2021.104393

Wolfensberger, D., Gabella, M., Boscacci, M., Germann, U., and Berne, A. (2021). RainForest: a random forest algorithm for quantitative precipitation estimation over Switzerland. *Atmospheric. Measurement. Techniques.* 14, 3169–3193. doi: 10.5194/amt-14-3169-2021

Xiong, T. Q., Wei, Q. S., Zhai, W. D., Li, C. L., Wang, S. Y., Zhang, Y. X., et al. (2020). Comparing subsurface seasonal deoxygenation and acidification in the yellow Sea and northern East China Sea along the north-to-South latitude gradient. *Front. Mar. Sci.* 7, 686. doi: 10.3389/fmars.2020.00686

Xue, L., Cai, W. J., Hu, X. P., Sabine, C., Jones, S., Sutton, A. J., et al. (2016). Sea Surface carbon dioxide at the Georgia time series site, (2006–2007): air–sea flux and controlling processes. *Prog. Oceanogr.* 140, 14–26. doi: 10.1016/j.pocean.2015.09.008

Xue, L., Zhang, L., Cai, W.-J., and Jiang, L.-Q. (2011). Air-sea CO2 fluxes in the southern yellow Sea: an examination of the continental shelf pump hypothesis. *Cont. Shelf. Res.* 31, 1904–1914. doi: 10.1016/j.csr.2011.09.002

Ye, H. J., Tang, S. L., and Morozov, E. (2022). Variability in Sea surface $p\mathrm{CO_2}$ and controlling factors in the bay of Bengal based on buoy observations at 15°N, 90°E. *J. Geophys. Res. Ocean.* 127, e2022JC018477. doi: 10.1029/2022JC018477

Yu, S. Q., Xiong, T. Q., and Zhai, W. D. (2022). Quasi-synchronous accumulation of apparent oxygen utilization and inorganic carbon in the south yellow Sea cold water mass from spring to autumn: the acidification effect and roles of community metabolic processes, water mixing and spring thermal state. *Front. Mar. Sci.* 9, 858871. doi: 10.3389/fmars.2022.858871

Zhai, W. D. (2018). Exploring seasonal acidification in the yellow Sea. *Sci. China Earth Sci.* 61, 647–658. doi: 10.1007/s11430-017-9151-4

Zhai, W. D., Zheng, N., Huo, C., Xu, Y., Zhao, H. D., Li, Y. W., et al. (2014). Subsurface pH and carbonate saturation state of aragonite on the Chinese side of the north yellow Sea: seasonal variations and controls. *Biogeosciences* 11, 1103–1123. doi: 10.5194/bg-11-1103-2014

Zhang, S. P., Rutgersson, A., Philipson, P., and Wallin, M. B. (2021). Remote sensing supported Sea surface $\mathrm{pCO_2}$ estimation and variable analysis in the Baltic Sea. *Remote Sens.* 13, 259. doi: 10.3390/rs13020259

Zimmerman, N., Presto, A. A., Kumar, S. P., Gu, J., Hauryliuk, A., Robinson, E. S., et al. (2018). A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. *Atmos. Meas. Tech.* 11, 291–313. doi: 10.5194/amt-11-291-2018