remote sensing

MDPI

Article

Mid-long term prediction of surface seawater organic carbon in the southern South China Sea based on multi-applicability CNN-LSTM prediction model

Na Liu[1†], Kuncheng Zhang[1,2†], Jing Yu[1,3*], Shaoyang Chen[4], Hao Zheng[5]

[1] Institute of Marine Development of Ocean University of China, Qingdao 266100, China
[2] College of Marxism, Ocean University of China, Qingdao 266100, China
[3] College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao 266100, China
[4] School of Marine Science and Technology, Tianjin University, Tianjin 300072, China
[5] College of Environmental Science and Engineering, Ocean University of China, Qingdao 266100, China
† These authors contributed equally to this work.
* Correspondence: by6801@ouc.edu.cn

Abstract: The organic carbon pool is a crucial component of the ocean carbon cycle. The study of organic carbon distribution and interannual variability in the land-sea interface can contribute to understanding the global ocean carbon cycle and ecological effects in the context of the Anthropocene and help achieve Sustainable Development Goals (SDGs). At present, there has been a certain amount of research on the source and flux of carbon in the process of the ocean carbon cycle, but the prediction of marine carbon is still in its infancy. In this paper, a CNN-LSTM deep learning model that takes into account spatio-temporal features was used to make a 5-year mid-long term rolling prediction of particulate organic carbon (POC) and yellow matter (CDOM) using MODIS Level 2 semimonthly synthetic data from the official website of NASA from January 2002 to June 2020. The model uses chlorophyll-a data to adjust the parameters. The results showed that the model could also be applied to the mid-long term rolling prediction of POC and CDOM. The model was capable of accurately predict POC and CDOM over periods of three and two years, respectively (R>0.5). Meanwhile, the 5-year trends of the predicted and actual values were verified using the least squares method and the Mann-Kendall trend test. The results showed that the predicted and actual values of sea surface POC in 2015-2020 show an overall upward trend; the predicted and actual values of sea surface CDOM in the same period showed an overall downward trend. The mid-long term prediction of surface seawater organic carbon in the southern South China Sea helps humans explore the regional characteristics of organic carbon in the coral reef waters of the South China Sea, study the changing trend of surface seawater organic carbon.

Keywords: Southern SCS; Organic carbon; CNN-LSTM model; Mid-long term rolling prediction; Interannual trends; Carbon cycle

Citation: To be added by editorial staff during production.

Academic Editor: Firstname Lastname

Received: date
Revised: date
Accepted: date
Published: date

Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Particulate organic carbon (POC) and dissolved organic carbon (DOC) are carbon compounds found in the land-ocean interface that originate from the decomposition of living organisms. They have been extensively studied by researchers such as Siegenthaler et al. (1993) and Wu et al. (2015), who have recognized their significance in understanding carbon cycling. Additionally, POC and DOC play crucial roles in marine ecosystems, as highlighted by studies conducted by Mopper et al. (1991), Hansell et al. (2001), and Kowalczuk et al. (2010). The concentration of POC, when compared to the total organic carbon (TOC), typically represents a relatively small portion, averaging around 4% (Cai et al., 2015; Wu et al., 2015). The flux of POC holds significant importance as it serves as a key

indicator of the effectiveness of the oceanic biological pump, as emphasized by Shih et al. (2020). It is widely accepted that new products in the upper ocean is typically offset by the sinking of POC, according to studies conducted by Eppley and Peterson (1979) and Pan and Wong (2015). The POC:Chl-a ratio is generally considered a valuable indicator of phytoplankton communities' ecological and physiological conditions (Lee et al., 2020). Approximately 97% of organic carbon in the ocean is found in dissolved form as dissolved organic matter (DOM), as indicated by studies by Hansell and Carlson (1998) and Kowalczuk et al. (2010). In both freshwater and marine systems, DOC represents one of the largest reservoirs of cycling organic matter on Earth (Cole et al., 2007; Butman and Raymond, 2011; Ren et al., 2016). Recent research has brought attention to the potential underestimation of the role of DOC in the ocean's carbon cycle (Carlson et al., 1994; Dunne et al., 2007). Studies conducted by Hansell and Carlson (1998), Ducklow et al. (2001), and Pan and Wong (2015) suggest that the export of DOC from the surface ocean to deeper depths may be comparable to the export of POC. In subtropical regions, recent studies have revealed that the proportion of organic carbon exports attributable to the transport of DOC may exceed 50% of the total carbon export (Abell et al., 2000; Doval and Hansell, 2000; Emerson, 2014).

The marginal sea covers only 7% of the ocean's surface area, but its contribution to global ocean primary productivity, nutrient remineralization, and burial of calcium carbonate and organic carbon are much higher in proportion (Walsh, 1991; Cai, 2011), making the marginal sea a key area for global carbon cycle studies. Gaining insights into and effectively managing the mechanisms and processes associated with POC and DOC in marginal seas are of utmost importance for comprehending the global carbon cycle (Kowalczuk et al., 2010; Sexton et al., 2011). The South China Sea (SCS) is a vast tropical marginal sea located within the Pacific Ocean, as indicated by research conducted by Chen et al. (2001) and Wong et al. (2007). The SCS has been identified as a significant source of atmospheric CO_2, as emphasized by studies conducted by Chen et al. (2001), Dai et al. (2009), and Dai et al. (2013a). Its significance in the global carbon cycle is underscored by research conducted by Wang et al. (2017). High levels of organic carbon content are observed throughout the SCS (Wang et al., 2017). The flux and degradation of organic carbon play a significant role in shaping microbial communities within the ocean and have far-reaching implications for the carbon cycle in the SCS (Wu et al., 2015). Our understanding of carbon cycling in coral reef systems in the SCS is currently limited due to the lack of research on the distribution and seasonal variations of organic carbon in this region, as emphasized by Zhu et al. (2022). In the face of increasingly serious ocean acidification, coral reef bleaching is severe, and nearly two-thirds of coral reefs are at risk of extinction (Schleyer et al., 2008; Savini et al., 2014). Further investigations are needed to fill this knowledge gap and enhance our understanding of the carbon cycle in coral reef systems in the SCS.

Previous research on POC in the Chinese offshore region has predominantly concentrated on investigating its sources, distribution patterns, and composition (Shen et al., 2020). Some studies have also attempted to explore the dynamics of POC further (Liu et al., 2019). There are relatively few studies on the spatial prediction of POC at home and abroad, mainly focusing on the prediction of POC flux (Marwick et al., 2015). There are also some attempts to predict POC content (Van Pinxteren et al., 2017), but they focus on predicting POC in seabed sediments. A study conducted by Diesing et al. (2017) involved the use of a random forest model to predict the amount of organic carbon present in the surface sediments of the Northwestern European continental shelf. Seiter et al. (2004) conducted a study to predict the total organic carbon content at the top layer of deep-sea sediments globally. They employed a combination of qualitative and quantitative statistical methods with a spatial resolution of 1º×1º. Regarding long-term predictions, Dzierzbicka-Glowacka et al. (2011) conducted a study focusing on long-term predictions of POC concentration in the southern Baltic Sea. They utilized a one-dimensional particulate organic carbon model (1D POC) and analyzed POC concentration trends from 1965

to 1998. Based on these observations, the researchers generated scenarios for POC concentration in the year 2050. However, the study focused on generating future POC concentration scenarios using a modeling approach based on historical trends, but did not provide a direct validation of these predictions against observed data.

In recent years, there has been a notable increase in research efforts dedicated to investigating marine DOM, as highlighted by studies conducted by Kowalczuk et al. (2010) and Yang et al. (2020). The Joint Global Ocean Flux Study (JGOFS) has made significant strides in unraveling the temporal and spatial variations of DOC in the marine system. Through extensive research, as emphasized by Jeandel et al. (2015), JGOFS has successfully identified and characterized the sources and sinks of DOC across different time scales. For many years, researchers have been actively engaged in studying the sources and fluxes of DOC, as evidenced by studies conducted by Aitkenhead et al. (2007), Fu et al. (2013), and Birkel et al. (2014). Additionally, scholars have been striving to develop methods for predicting DOC concentration based on its absorbance characteristics (Banoub, 1973; Stewart and Wetzel, 1981; Marhaba et al., 2003; Carter et al., 2012). This line of research aims to establish a relationship between the optical properties of water and DOC concentration. In the past, the surface distribution of DOC in the ocean has been assessed through the analysis of discrete water samples collected during individual research cruises, as outlined by Hansell and Carlson (2001). However, this method provides a limited spatial and temporal coverage, as it relies on data from a single cruise (Mannino et al., 2008). To overcome the limitations of traditional sampling methods, researchers have suggested the use of satellite remote sensing to estimate DOC concentrations in the ocean. Specifically, the retrieval of yellow matter, known as Chromophoric Dissolved Organic Matter (CDOM), from satellite data has been proposed as a potential method to achieve higher spatial and temporal resolution in monitoring DOC levels (Hoge et al., 1995; Mannino et al., 2008; Kowalczuk et al., 2010; Asmala et al., 2012; Tehrani et al., 2013). This approach, as proposed by Ferrari (2000), leverages the spectral characteristics of CDOM to estimate DOC concentrations over large areas. By utilizing satellite remote sensing, it is possible to obtain frequent and extensive measurements of CDOM, which can then be used as a proxy for DOC concentrations, allowing for a more comprehensive and dynamic assessment of DOC distribution in the ocean. Despite certain limitations highlighted by Mannino et al. (2008), Nelson and Siegel (2013), and Hestir et al. (2015), the utilization of the DOC-CDOM model for estimating DOC concentrations has demonstrated robustness through the testing with independent datasets and methods. It is important to note that there are significant seasonal and spatial variations in the DOC-CDOM relationship, as indicated by studies conducted by Kowalczuk et al. (2013), Catalá et al. (2016), and Iuculano et al. (2019). However, despite these variations, the overall performance of the algorithms used in the model has been found to be reliable, with an estimated uncertainty ranging from 4% to 10% (Mannino et al., 2008). These findings indicate the potential of employing satellite-based algorithms to provide reasonably accurate estimates of DOC concentrations, despite the inherent challenges associated with seasonal and spatial differences. This method is a quick and inexpensive alternative to direct analytical methods for measuring DOC concentrations (Spencer et al., 2007; Asmala et al., 2012). The use of satellite remote sensing to estimate DOC concentrations offers a rapid and cost-effective alternative to traditional analytical methods (Spencer et al., 2007; Asmala et al., 2012). Coastal samples have exhibited strong correlations in the prediction of DOC concentrations ($0.64<R^2<0.82$) (Massicotte et al., 2017). Remote sensing-based algorithms utilizing aCDOM data have shown promise in retrieving DOC concentrations for the SCS during the summer and winter seasons. Studies by Siegel et al. (2002), Johannessen et al. (2003), Mannino et al. (2008), and Li et al. (2019) have demonstrated the feasibility of estimating DOC levels using remote sensing data during the months of August and January. Mannino et al. derived DOC from aCDOM through the relationship between aCDOM and DOC and established the pattern and intra-annual variation of DOC distribution in the SCS region, representing the first validated algorithm for offshore aCDOM,

DOC, and CDOM spectral slope (Mannino et al., 2008). A more comprehensive approach for estimating DOC concentrations, even in highly photobleached seawater, was proposed by Fichot and Benner (2011). Their method involves utilizing CDOM coefficients at specific wavelengths, such as 275 nm and 295 nm, to estimate DOC levels. Unlike traditional methods that may struggle to accurately measure DOC in highly photobleached seawater, this approach offers a more reliable means of quantifying DOC concentrations. CDOM, which consists of intricate compounds, is recognized as a valuable indicator of water quality (Li et al., 2018). It serves as a proxy for assessing the level of organic pollution in water to some extent (Shao and Wang, 2020; Li et al., 2021). Seasonal and interannual variations in CDOM play a pivotal role in the marine organic carbon cycle, as highlighted by Ma et al. (2011). These fluctuations provide valuable insights into the temporal dynamics of organic carbon in marine ecosystems. By studying the changes in CDOM concentrations and properties over different time scales, researchers can gain a better understanding of the factors influencing organic carbon distribution, transport, and transformation in the marine environment. This knowledge contributes to our comprehension of the overall carbon cycling processes and their impacts on ecosystem functioning. Consequently, studying the seasonal to interannual variation of CDOM is crucial for unraveling the complexities of the marine organic carbon cycle.

While considerable efforts have been made to study and quantify organic carbon distribution in the global oceans (Gardner et al., 2006; Jeandel et al., 2015), our understanding of organic carbon dynamics in tropical oligotrophic marginal reef waters remains limited (Fichot and Benner, 2011; Lorenzoni et al., 2011). Filling this gap is crucial for gaining a holistic understanding of the global carbon cycle and the intricate interactions between organic carbon and reef ecosystems. This study used an improved CNN-LSTM model that used chlorophyll-a (Chl-a) data for training and learning to extract features and find optimal parameters (Liu et al., 2022). The CNN-LSTM model based on Chl-a data directly predicts POC and CDOM over five years and explores the distribution and interannual trend of sea surface organic carbon in the Liyue Tan area of Nansha.

2. Materials and methods

2.1. Data processing

There are two giant atolls in the SCS, one of which is Liyue Tan. Liyue Tan is the largest underwater atoll in Nansha, 9514 km^2. There are no islands and sandbanks on the reef. At the southwest corner of Liyue Tan are Antang Jiao, Huoxing Jiao, Gongzhen Jiao, Liyue Nanjiao, and Houteng Jiao. Houteng Jiao can be exposed at low tide. There are Xiongnan Jiao in the east and Dayuan Tan in the west. The Liyue Tan has good pre-Cenozoic oil and gas resource potential (Hu et al., 2011). Liyue Tan is situated within the Nansha Islands, located in the southern region of the SCS. The Nansha Islands are driven by terrain and monsoon with frequent mesoscale eddies (Chu et al., 1998). The SST in the Nansha sea area remains above 22 °C all year round; salinity levels in this region are typically above 33, and the concentration of Chl-a is deficient (<0.3 mg/m^3) (Pan et al., 2012; Pan et al., 2013; Ho et al., 2015), and the DOC concentration in the oligotrophic area is about 60-80 μM (Callahan et al., 2004). The hydrological conditions in the Liyue Tan area are mainly influenced by mixing with salt water, warm, and low-nutrient open SCS. The average surface water temperature of Liyue Tan was significantly lower than that of Nansha, which may be related to the upwelling of low-temperature bottom water caused by the northeastward transport of the west coast of Nansha (Dai et al., 2013b). The marginal sea has large spatial and temporal variability and complex physical and biogeochemical processes. The South China Sea (SCS) is a tropical marginal sea characterized by its semi-enclosed nature. The upper circulation in the SCS exhibits distinct patterns depending on the season. During winter, the circulation is primarily cyclonic and is driven by the influence of the East Asian monsoon, but is cyclonic in the north and anticyclonic in the south in summer (Qu et al., 2009). Near northern Liyue Tan is the high-abundance

area of phytoplankton in the Nansha Islands. Its average abundance is about three times that in Nansha Islands (Dai et al., 2013b). The sea surface pCO_2 in the SCS varies from 37.8 to 57.1 Pa, which is higher than the atmospheric pCO_2, showing as the source of atmospheric CO_2. In some sea areas, the calcification of coral reefs resulted in highly high surface pCO_2 in the Nansha Islands (Yan et al., 2011). The particular geographical location, complex topographic conditions, and hydrological environment of the SCS and the input of major rivers such as the Pearl River and Mekong River, together with the frequent occurrence of meteorological and hydrological events such as typhoons and mesoscale eddies (Sun et al., 2014), have led to the complex carbon cycle characteristics of the SCS.

Total organic carbon (TOC) includes POC and DOC. The study selected POC and CDOM for predictive analysis. Nansha Islands is far from the mainland, and the actual measurement and space station observation require a lot of workforce and material resources. Therefore, utilizing remote sensing technologies, the research collected long-term, large-scale data. Satellite ocean color products are usually calibrated based on existing observations to provide more comprehensive coverage. The experiment used data from the NASA Ocean Color website (https://oceancolor.gsfc.nasa.gov/). The downloaded data are satellite time series. To ensure a more comprehensive dataset with minimal missing values, the study utilized semi-monthly MODIS sensor synthetic product data from January 2002 to June 2020. The study utilized satellite data from the Terra and Aqua satellites, which provide accurate measurements with a spatial resolution of 1 km. The selected dataset specifically covered the Liyue Tan area in the southern part of the SCS, bounded by latitudes 11.1°N to 12.1°N and longitudes 115.8°E to 117.5°E (Figure 1). The study reanalyzed the data and expanded it 20 times according to the divided grid to produce a grid product with a 20 km 20 km spatial resolution. A single grid's value is replaced with the mean of 400 original data points.

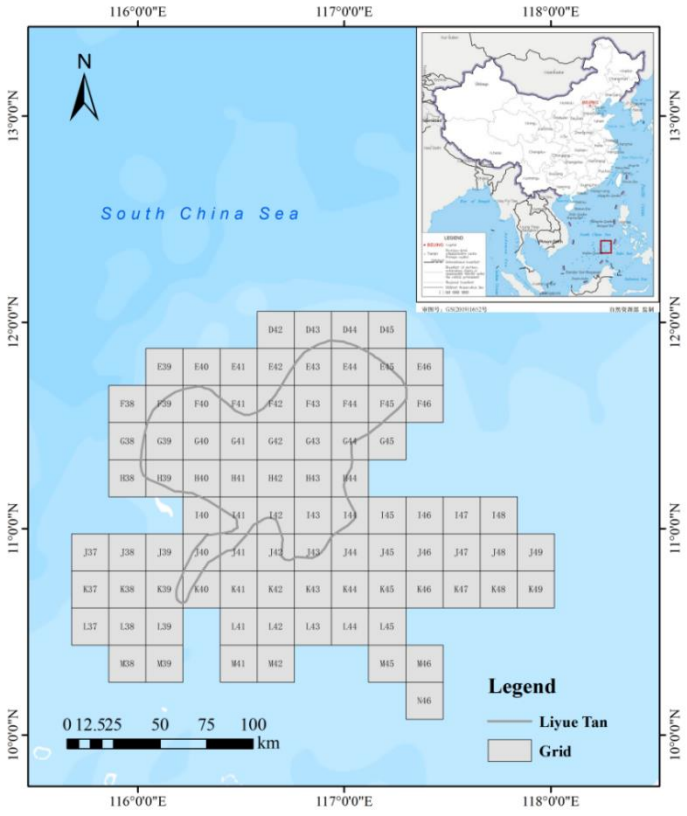

Figure 1. Study area.

POC (mg/m³) data can be obtained directly by downloading MODIS Level 2 product data. The POC algorithm (Stramski et al., 2008) adopted by the NASA website can better

invert the global oceanic POC surface concentration and obtain remote sensing products with high spatial and temporal resolution, which provides a good data source for POC studies. Cui et al. (2017) verified and corrected the monthly average POC remote sensing products of MODIS / AQUA satellite released by NASA and found that the POC remote sensing products had an excellent linear relationship with the measured data in the northern SCS (R^2 = 0.72). Due to the highly complex biochemical composition of CDOM, there is no standard substance so far, so the concentration of CDOM in seawater cannot be directly determined. Studies typically characterize the relative concentration of CDOM in seawater by measuring its spectral absorption coefficient (Kim et al., 2022), typically using a 400 nm or 440 nm absorption coefficient to represent its concentration broadly. Different optical properties of water bodies lead to different inversion algorithms for CDOM in different water bodies. To quantify the concentration of CDOM in this study, the absorption coefficient at 440 nm was utilized. This parameter was derived through a reflectance inversion process applied to MODIS satellite data, specifically bands 8, 9, and 10. The inversion model refers to the CDOM inversion model (R^2=0.78) obtained by regression analysis of MODIS remote sensing data in the Ledong sea area of the SCS using the Tassan model by Zhang et al. (2017), which is a CDOM inversion model using Sea WiFS data established by Tassan (1994) in the Bay of Naples based on relevant information. The vast majority of the sea area of Ledong meets one class's national water quality standard. In this paper, the water quality of the Liyue Tan sea area in the southern part of the SCS is high and is also of first-class water quality. The inversion equation of CDOM in this study is as follows:

$$\lg[a(440)] = -2.68 - 2.28 \times \lg\left[\left(b_8 \div b_{10}\right) \times \left(b_9\right)^{0.66}\right] \quad (1)$$

Among them, $a(440)$ is the absorption coefficient of CDOM at 440 nm, b_8, b_9 and b_{10} are the reflectance of MODIS data in 8, 9 and 10 bands, respectively.

Cloud coverage in the South China Sea is notably high, especially during the summer months, with cloud coverage reaching around 80%, which may result in sensor data loss and anomalies. Consequently, to preserve the integrity of the data, this study removed outliers outside the quartile range and used spline interpolation for missing values. To ensure the model's convergence speed and prevent gradient changes, the original data is preprocessed using Min-Max normalization before being fed into the model. In this study, Min-Max normalization was applied to the data using the scikit-learn machine learning preprocessing library. This normalization technique rescaled the data to a range between 0 and 1, ensuring that all values were evenly distributed within this range.

To ensure reliable model training and evaluation, the dataset was divided into three sets: training, validation, and test sets. In terms of the temporal dimension, the dataset encompassed 444 data points spanning from January 2002 to June 2020. The training set consisted of 210 data points spanning from January 2002 to September 2010. The validation set included 114 data points covering the period from September 2010 to June 2015. Lastly, the test set comprised 120 data points from June 2015 to June 2020.

2.2. CNN-LSTM model

In this work, CNN and LSTM are combined to form the CNN-LSTM model. CNN has the benefit of recognizing spatial information characteristics and extracting multidimensional picture attributes. Moreover, it still maintains the benefits of LSTM analyzing time-series data. CNN extracts the intricate characteristics of the indicators in the CNN-LSTM model. The LSTM model is then used to predict based on these depth characteristics (Liu et al., 2022).

The experiment used the CNN-LSTM prediction model of Chl-a, improved in the author's previous study. The dropout rate in the model was set to 0.001. The model selected the LSTM unified initial Tanh activation function. The LSTM layer was initialized

with 64 hidden neural units, the kernel layer utilized the glorot_uniform kernel initializer, the Zeros method was used for weight initialization, and hard_sigmoid was the activation function for the loop steps. The CNN-LSTM model was set with a learning rate of 0.001, a batch size of 8, and an epoch of 150. During execution, the model used the MSE loss function and Adam optimization.

3. Results

3.1. Correlation analysis of indicators

Pearson correlation analysis was conducted among Chl-a, POC, and CDOM. Pearson correlation coefficients are shown in Figure 2.

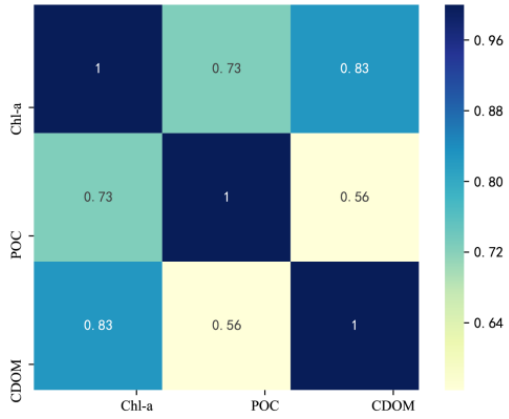

Figure 2. Correlation of indicators.

As seen in Figure 2, the Pearson correlation of Chl-a with POC was 0.73, and the Pearson correlation of Chl-a with CDOM reached 0.83. Chl-a was positively correlated with both POC and CDOM, and the correlation was high.

3.2. Parameter optimization of CNN-LSTM

3.2.1. Spatial Correlation

For this experiment, the grid that has a strong association with the target grid is chosen to extract features and predict the organic carbon in the target area. Therefore, it is essential to verify the grid's spatial correlation before predicting. To represent grid correlation, calculate mutual information entropy (Equation S1). The mutual information entropy between the D42 grid and other grids is computed using the D42 grid region as an example (Figure S1, Figure S2).

Grid D42 shows the highest association with D43 and E42 for both POC and CDOM, as can be shown in Figs. S1 and Figs. S2. The mutual information entropy of grid D42 of POC with D43 and E42 is 0.620 and 0.595, and the mutual information entropy of grid D42 of CDOM with D43 and E42 is 0.955 and 0.807, respectively. The D42 grid is adjacent to D43 and E42. The analysis shows that there is potential for a strong correlation between adjacent grids of POC and CDOM.

On all chosen grids in the study area, mutual information entropy studies were carried out. (Figure 3). Each grid area was shown to have a significant correlation with the region around it. The mutual information entropy of the POC grid is smaller than that of the CDOM, indicating that the POC in the Liyue Tan area has a larger grid area different than the CDOM.

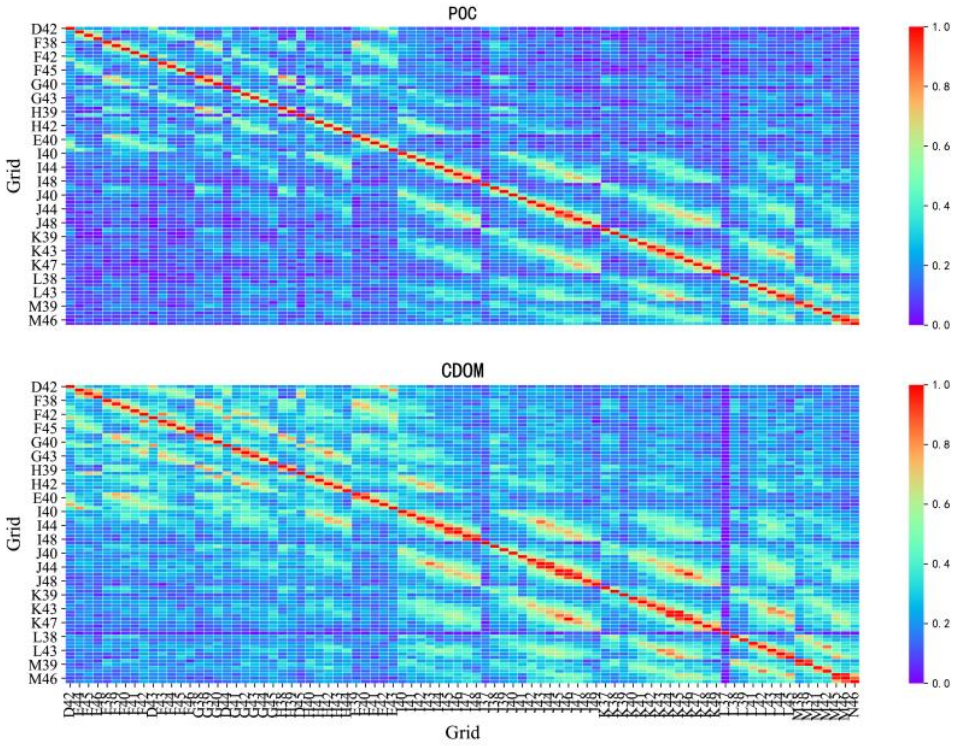

Figure 3. Grid mutual information heat map.

3.2.2. Training process

In predicting the target grid, the two grids with the most extensive mutual information with the target grid are selected and input to the model together with the target grid to predict the target grid. Since POC and CDOM are highly correlated with Chl-a, the authors attempted to predict POC and CDOM directly using the improved Chl-a prediction model from the authors' previous study (Liu et al., 2022). The (D42-D43-E42) combination is used in this study to predict the grid D42 series. The loss function is calculated during the model training. The epoch of POC is set to 300. After 150 rounds, the training and test sets' loss curves are smooth. The epoch of CDOM is set to 500. After 150 rounds, the training and test sets' loss curves start to progressively converge (Figure S3, Figure S4).

3.3. Single-step prediction

The D42 grid's POC and CDOM sequences were predicted using the CNN-LSTM model, and a 95% confidence interval was calculated. The input for prediction is the three-dimensional organic carbon data from grids D42, D43, and E42. The following data from the target grid is predicted utilizing 30-time steps of the data moving forward one step at a time in the single-step prediction procedure.

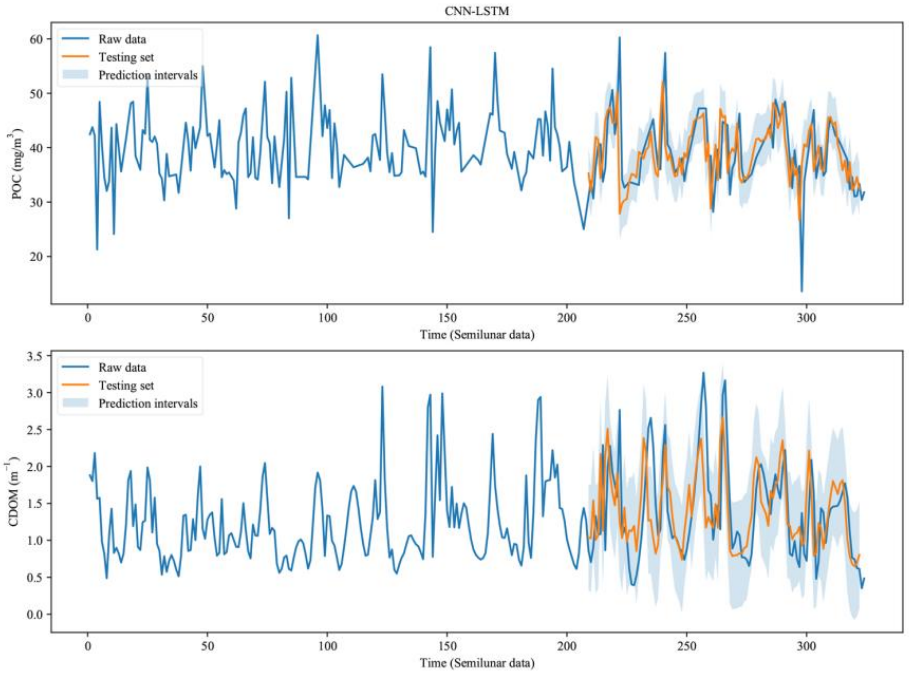

Figure 4. Actual and predicted values for single-step prediction.

As seen in Figure 4, for predicting POC and CDOM, the predicted values fit the actual value curve ideally with narrow confidence intervals.

The MAE (Equation S3) and MSE (Equation S2) of the prediction were calculated, and the results are shown in Table 1. the MAE was 1.68 mg/m^3, and the MSE was 5.84 mg/m^3 for POC prediction. The MAE was 0.272 m^{-1}, and the MSE was 0.136 m^{-1} for CDOM prediction.

Table 1. Prediction error of POC and CDOM.

	POC (mg/m^3)	**CDOM (m^{-1})**
MAE	1.68	0.272
MSE	5.84	0.136

The prediction effect of POC and CDOM is ideal. That shows that the Chl-a prediction model can be used for both Chl-a prediction and POC and CDOM prediction. The reason may be that POC and CDOM are highly correlated with Chl-a, so the three indicators can be predicted by a model.

3.4. Long-term prediction and interannual trend analysis of organic carbon

3.4.1. Mid-long term prediction of POC

During the rolling prediction procedure, the model uses a sliding window approach where predicted data is added to the sample to update it. The model then uses a sequence of 72 time steps (equivalent to 3 years) of organic carbon data to predict the subsequent organic carbon values. This iterative process continues, with each prediction being incorporated into the input sequence for the next prediction. The study predicted POC sequences for all gridded areas for up to five years in the long term and removed grids where predictions were not ideal due to the presence of reefs (e.g., J41-J43, K38-K41). The pearson correlation coefficient (Equation S4) between the predicted and actual values is shown in Figure 5.

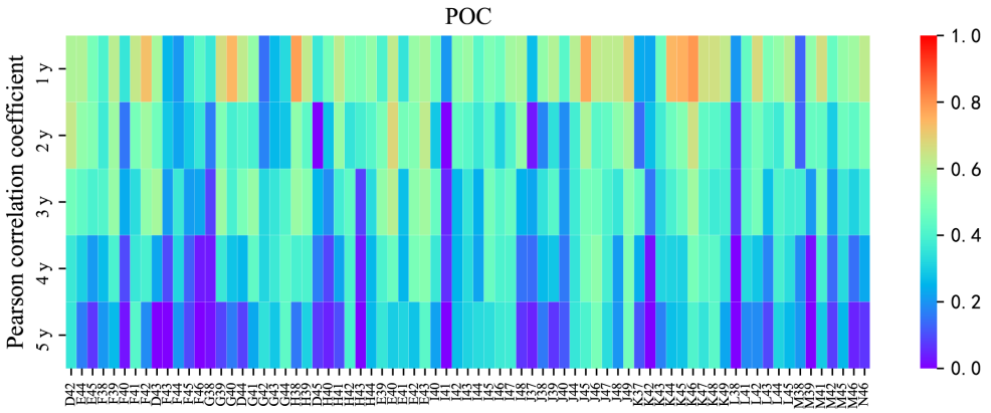

Figure 5. Pearson correlation coefficient of 5-year long-term prediction of POC series in all grid regions.

Figure 5 demonstrates that for the prediction of POC, as the prediction time increases, the prediction accuracy of each grid area of the data set gradually decreases.

The grid averages of MAE, MSE, and Pearson correlation coefficient of POC long-term prediction were calculated in order to more properly statistics the accuracy of long-term prediction (Table S1).

As the prediction period lengthens, the average MAE and MSE of the prediction region increase, indicating a higher level of prediction uncertainty. Additionally, the Pearson correlation coefficient, which measures the correlation relationship between predicted and actual POC values, gradually decreases. These findings suggest that the accuracy of long-term POC prediction decreases over time. Nevertheless, the results of the long-term POC prediction still provide support for the suitability and effectiveness of the Chl-a CNN-LSTM long-term rolling prediction model. Despite the increasing prediction errors, the model demonstrates its capability to generate POC predictions over an extended period. When estimating the 5-year POC concentration for the entire sea region, the grid's Pearson correlation coefficient is almost always lowered to less than 0.4, the predicted value differs significantly from the actual value. The prediction results show that the Pearson correlation coefficient of the 1-year prediction is high (>0.5). The CNN-LSTM model can accurately predict POC for up to one year in the Liyue Tan area in the southern part of the SCS (Table S1).

3.4.2. Mid-long term prediction of CDOM

The study used CNN-LSTM models to make long-term predictions of CDOM sequences for all grid regions for up to five years and removed grids (such as K40-K42) where predictions were not ideal due to the presence of reefs. The correlation between the actual and predicted values is shown in Figure 6.

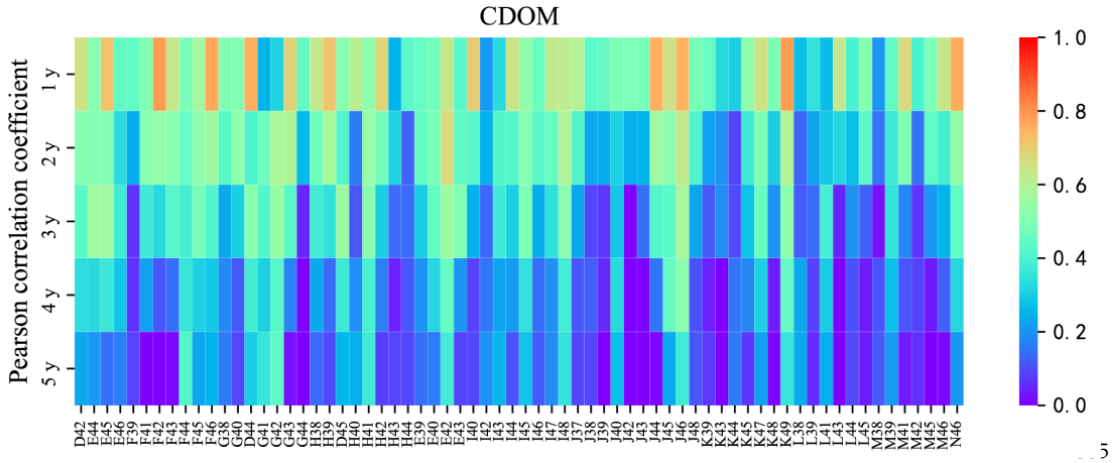

Figure 6. Pearson correlation coefficient of 5-year long-term prediction of CDOM sequences in all grid regions.

Figure 6 demonstrates that for the prediction of CDOM, the overall effect of long-term prediction is similar to that of POC. Each grid area's prediction accuracy gradually declines as prediction time goes up. The prediction accuracy of CDOM is generally worse than that of POC.

In order to provide a more comprehensive assessment of the CNN-LSTM long-term prediction model for CDOM, the average values of MAE, MSE, and Pearson correlation coefficient were calculated for each grid in the study area. This approach allows for a more accurate evaluation of the model's performance across the entire spatial domain (Table S2).

As the prediction period lengthens, there is a consistent trend observed in the average MAE and MSE within the prediction region. These error metrics tend to increase over time, indicating that the accuracy of the predictions decreases with longer prediction periods. In contrast, the Pearson correlation coefficient, which measures the linear relationship between predicted and actual values, exhibits a gradual decline as the prediction period lengthens. This suggests that the model's ability to capture the true underlying patterns and variations in the CDOM data diminishes over time. The results of the long-term CDOM prediction provide evidence for the effectiveness of the Chl-a CNN-LSTM model in accurately forecasting CDOM levels over an extended period. From the standpoint of the entire sea region, the Pearson correlation coefficient of the grid is virtually all reduced to less than 0.4 when predicting the three-year CDOM concentration, the predicted value is considerably different from the actual value. The prediction level of 1-year was similar to that of POC, and the Pearson correlation coefficient was high (0.5). The CNN-LSTM model can also predict CDOM in the southern SCS for up to one year (Table S2).

3.5. Trend analysis

The monotonous trend of the POC and CDOM time series was examined in this study using the least squares method and the M-K trend test method. The M-K test statistic Z was determined by dividing the time series into a subset of annual units (Equation S7). The probability that the trend exists was defined as P and the significance level (α) was chosen at 0.05. The study area's 5-year long-term predicted CDOM values and actual values were analyzed using statistical methods such as the least squares method and the M-K trend test.

Table 2. Interannual trend analysis.

Index	Trend test		Actual value	Predicted value
POC	M-Ktest	Z	1.16	0.160

CDOM	Least squares	P	0.320	0.250
		k	1.84e-2	3.28e-3
		b	39.1	41.6
	M-K test	Z	0.452	-0.376
		P	0.608	0.234
	Least squares	k	1.31e-3	3.69e-4
		b	1.34	1.43

The M-K test revealed a slightly rising trend in the actual and predicted values of POC in the overall area from 2015 to 2020, although it was not statistically significant ($P < 1-\alpha$). The least squares test also confirmed the upward trend of POC in the study area.

The trend analysis of the CDOM grid study area showed the interannual variation of CDOM in the study region showed a minor decrease trend from 2015 to 2020, albeit this trend was not statistically significant ($P < 1-\alpha$). The least squares test also confirmed the downward trend of CDOM in the study area.

4. Discussion

The single-step prediction experiment and mid-long term prediction of POC and CDOM show that the CNN-LSTM prediction model for Chl-a is also applicable to POC and CDOM. The reason may be that Chl-a has a high correlation with POC and CDOM (Figure 2). Therefore, the CNN-LSTM model trained by Chl-a data fitting also applies to POC and CDOM in the same region. Hung et al. (2007) also verified the high correlation between POC and Chl-a in the northern SCS ($R = 0.69$), which is close to the results of this study. The high correlation of Chl-a with POC and CDOM indicated that phytoplankton contributed more to POC and CDOM. The positive correlation observed between Chl-a and CDOM can be attributed to several factors. These include the presence of high primary productivity, significant particle influx from the surface, and relatively low levels of photodegradation within the water column. These combined conditions contribute to the accumulation of CDOM in the water (Wang et al., 2017). According to the findings of Wang et al. (2021), it is proposed that the in situ primary production, along with subsequent microbial transformation processes, play a crucial role in shaping the dynamics of DOM.

The correlation between Chl-a and CDOM was higher than that of POC (Figure 2), and the spatial grid correlation of CDOM was also higher (Figure S1, Figure S2). However, the prediction accuracy of CDOM is generally lower than that of POC. The reason may be that there may be errors in the direct application of the CDOM inversion model in the Ledong sea area of the SCS (Zhang et al., 2017). It may also be due to the apparent seasonal variation of POC. The concentration of POC is low in spring and summer, and the lowest value generally occurs in May. The concentration of POC reaches the highest in winter, and the concentration of POC in autumn begins to rise rapidly until the peak in winter. However, CDOM has no apparent seasonality.

Previous studies have found that, like Chl-a, POC is significantly affected by monsoon and temperature. More excellent primary production, variations in monsoon intensity, and frequent internal solid waves in the SCS facilitated POC's downward and lateral flow (Alford et al., 2015; Ma et al., 2017), resulting in higher POC concentrations. The POC concentration in the surface seawater of Liyue Tan in the SCS is between 20-60 mg/m^3 (Figure 4), consistent with the previous research results (Huang et al., 2007; Liu et al., 2007; Cai et al., 2015). When the temperature rises, both Chl-a and POC show a downward trend. When the temperature falls, both Chl-a and POC show a significant upward trend (Cui et al., 2017). In addition, the high correlation between POC and Chl-a indicates that POC is sensitive to biological changes and is greatly affected by organisms. The increase in phytoplankton biomass and primary production will result in increased zooplankton biomass and upper-middle layer debris concentration, as well as more zooplankton consumers, including fish (Dzierzbicka-Glowacka et al., 2011), increasing POC content.

CDOM has no apparent seasonality and has the lowest concentration in summer (Li et al., 2019). CDOM in seawater can be divided into two categories according to their sources: in situ degradation of organisms in the ocean and dissolved organic matter in terrestrial soil brought into the ocean through river runoff (Zhu et al., 2022). Compared with nearshore and estuarine areas, the Nansha sea area is not influenced by land-based sources and belongs to Class I water bodies with low CDOM content (Nelson et al., 1998). The CDOM content in the Liyue Tan sea area ranges from 0.5 m^{-1} to 3 m^{-1} (Figure 4). The majority of DOM found in the upper ocean is susceptible to biodegradation processes, with varying turnover times that can span from minutes to decades. Apart from biodegradation, DOM is also influenced by photochemical reactions occurring in the presence of light. These photochemical processes lead to the photobleaching of CDOM and the direct breakdown of biologically unstable or refractory compounds (Moran and Zepp, 1997; Sharpless and Blough, 2014). CDOM is mostly formed in-situ by biological production in the marginal seas of the SCS, and it is eliminated via photochemical degradation and microbial consumption (Nelson and Siegel, 2013). The dominant component of CDOM in the Nansha sea area is humus, as observed by Siegel et al. (2002). However, it is worth noting that the high seas DOM pool in this region primarily consists of colorless DOM rather than light-absorbing CDOM. These components are resistant to microbial degradation but susceptible to solar radiation. Photodegradation processes, influenced by solar radiation, have a significant impact on the abundance and distribution of CDOM in the surface layer of the SCS (Helms et al., 2013). In summer, the absorption of CDOM in sea surface water decreases rapidly due to intense photobleaching (Vodacek et al., 1997). The amount and composition of CDOM in the upper layer of the SCS are influenced by various factors, including lateral and vertical mixing processes. These mixing processes, such as mesoscale eddies and frontal zones, play a crucial role in altering the biogeochemical variables and controlling the distribution of CDOM. Wang et al. (2017) suggest that the interaction between different water flows, mediated by mixing processes, can significantly impact the quantity and characteristics of CDOM in the upper layer of the SCS. In winter, surface water mixed with upwelling brings new nutrients and CDOM from the deep.

The study found that the presence of islands and reefs and whether the grid arrangement is regular may affect the prediction accuracy. The mid-long term prediction accuracy of some grids (such as F43-F46) in the study area of POC is low, and the Pearson correlation coefficient of 2 years is almost below 0.4 (Figure 5). The reason may be that the existence of islands and reefs has an impact on the prediction of POC. When excluding part of the grid where the reefs exist, the study found that the Pearson correlation coefficient of the three-year POC value prediction is above 0.5 for mid-long term POC prediction in the open ocean. That indicates that the CNN-LSTM model can predict POC in the open ocean for up to three years. The M and N grid areas' ability to predict outcomes is poor. The reason is that the irregular shape of the grid area leads to low mutual information entropy between the target grid and the surrounding grid. The input of the CNN-LSTM model is no longer a high correlation spatial data, thereby reducing the prediction accuracy.

The same as the POC index, the mid-long term prediction accuracy of the regional partial grid (J37-J48, K39-K43) of CDOM is low. The Pearson correlation coefficient of the two-year prediction is almost below 0.4 (Figure 6). The reason may be due to the influence of Antang Jiao, Huoxing Jiao, Gongzhen Jiao, Liyue Nanjiao, and Houteng Jiao on CDOM prediction. For the sea area where the reefs exist, the Pearson correlation coefficient of 1-year prediction in more grid areas is above 0.5, indicating that the 1-year prediction of the reef sea area is ideal. Excluding some grids where islands and reefs exist, the study found that the CDOM long-term prediction in the open ocean is worse than that of POC. The long-term prediction of POC in the open ocean can reach three years, while the Pearson correlation coefficient of the 3-year prediction of CDOM is almost below 0.4. The Pearson correlation coefficient of the 2-year prediction of CDOM is mostly above 0.5. Thus, the CNN-LSTM model can predict CDOM in the open ocean for up to two years. The grid prediction accuracy of numbers L, H, and M is low, possibly due to the irregular shape of

these regions and the low correlation between the predicted target grid and the surrounding grid. When the CNN-LSTM model predicts, the data with a significant difference from the target grid is added as the spatial input, which affects the prediction effect. Therefore, organic carbon prediction should divide the study area into more regular grid areas.

The M-K trend and the least squares test showed that POC showed a slight upward trend from 2015 to 2020 (Table 2). Previous studies showed that the primary productivity on the surface of sea samples was positively correlated with POC (Eppley and Peterson, 1979), and the concentration of Chl-a in the Liyue Tan area also showed an upward trend (Liu et al., 2022), which verified the upward trend of POC. CDOM shows a slight downward trend from 2015-2020 (Table 2). The changing trend of CDOM is opposite to that of POC. Studies have pointed out that the decomposition of POC is the primary source of DOC release (Boss et al., 2001). It is speculated that the opposite change trend of CDOM and POC may be related to the conversion between the two. The M-K trend test P value is relatively low because the data is from remote sensing. Due to the lack of measured data validation, there is a deviation between the actual value and the obtained value. Additionally, there are a lot of missing values in the remote sensing data, and the interpolated sequence has an inaccuracy when compared to the actual value.

The variation trend of the single grid predicted value is different from the actual value, indicating that the grid's mid-long term prediction is unstable and has an error. That is most likely because the prediction's accuracy will decrease year over year as predicted data is added to the series to predict the subsequent value in the rolling prediction for 2015-2020. Additionally, the prediction accuracy of a single grid area is also influenced by elements like human activity trajectories and coral reefs. Oil and gas resources are abundant in Liyue Tan, and oil and gas development operations in neighboring nations may have an impact on the anticipated level of organic carbon. Greater human development activities may have an impact on changes in organic carbon, making the model inapplicable or impairing its ability to predict the future. The presence of reefs in the Liyue Tan region may also have an impact on the distribution of organic carbon among the various grid areas. A single grid area's mid-long term projection contains errors. However, the long-term prediction of organic carbon is in line with the reality for the entire region. In light of this, the study concludes that the model provides better performance in large-scale prediction than in a single grid region. This finding suggests the model has the potential to be utilized to a broader range of areas, including additional SCS seas and even international seas.

5. Conclusion

The prediction of carbon cycle elements at home and abroad is in the initial stage. This study predicted the organic carbon in tropical oligotrophic marginal reef seas for the first time. The results of single-step prediction experiments of POC and CDOM show that the CNN-LSTM prediction model of Chl-a is also suitable for POC and CDOM, which indicates that the CNN-LSTM model has the potential to predict more elements. The Pearson correlation coefficients (R) of regional POC and CDOM rolling prediction for one year were 0.514 and 0.524, respectively. From the general study area, the CNN-LSTM model can predict POC and CDOM in the Nansha sea area for up to one year. The study found that compared with the prediction of organic carbon in open oceans, the prediction performance of reefs is relatively poor. The CNN-LSTM model can predict POC in open seas for up to 3 years and CDOM for up to 2 years. The mid-long term prediction of organic carbon in coral reef waters of the southern SCS can help us understand the carbon cycle and the changes of ecological effects in coral reef areas in the context of climate change.

The study found that the POC concentration on the surface of the sea area of Liyue Tan showed an upward trend and the CDOM concentration showed a downward trend from 2015 to 2020. Previous studies have pointed out that the decomposition of POC is the primary source of DOC released (Boss et al., 2001). It is speculated that the opposite trend of CDOM and POC may be related to the conversion between the two. Predicting

organic carbon trends in coral reef areas is conducive to exploring the carbon conversion mechanism in the ocean carbon cycle. The predicted trend of organic carbon is consistent with the actual trend, which verifies the accuracy of the multi-year trend prediction of the CNN-LSTM model. The trend test for large-scale regions is more favorable, indicating that the model may be applicable to large-scale sea areas and has the potential to predict global organic carbon.

Acknowledgments: The work is supported by the National Social Science Foundation of China (No. 20VHQ002).

References

Abell, J., Emerson, S., Renaud, P., 2000, Distribution of TOP, TON, and TOC in the North Pacific subtropical gyre: Implications for nutrient supply in the surface ocean and remineralization in the upper thermocline, J. Mar. Res., 58, 203-222.

Aitkenhead, M.J., Aitkenhead-Peterson, J.A., McDowell, W.H., Smart, R.P., Cresser, M.S., 2007. Modelling DOC export from watersheds in Scotland using neural networks. Comput Geosci 33(3), 423-436.

Alford, M.H., Peacock, T., MacKinnon, J.A., Nash, J.D., Buijsman, M.C., Centuroni, L.R., Chao, S.Y., Chang, M.H., Farmer, D.M., Fringer, O.B., Fu, K.H., Gallacher, P.C., Graber, H.C., Helfrich, K.R., Jachec, S.M., Jackson, C.R., Klymak, J.M., Ko, D.S., Jan, S., Johnston, T.M.S., Legg, S., Lee, I.H., Lien, R.C., Mercier, M.J., Moum, J.N., Musgrave, R., Park, J.H., Pickering, A.I., Pinkel, R., Rainville, L., Ramp, S.R., Rudnick, D.L., Sarkar, S., Scotti, A., Simmons, H.L., St Laurent, L.C., Venayagamoorthy, S.K., Wang, Y.H., Wang, J., Yang, Y.J., Paluszkiewicz, T., (David) Tang, T.Y., 2015. The formation and fate of internal waves in the South China Sea. Nature 521(7550), 65-69.

Asmala, E., Stedmon, C.A., Thomas, D.N., 2012. Linking CDOM spectral absorption to dissolved organic carbon concentrations and loadings in boreal estuaries. Estuar. Coast. Shelf Sci. 111, 107-117.

Banoub, M.W., 1973. Ultra violet absorption as a measure of organic matter in natural waters in Bodensee.

Birkel, C., Soulsby, C., Tetzlaff, D., 2014. Integrating parsimonious models of hydrological connectivity and soil biogeochemistry to simulate stream DOC dynamics. J. Geophys. Res. 119(5), 1030-1047.

Boss, E., Pegau, W.S., Zaneveld, J.R.V., Barnard, A.H., 2001. Spatial and temporal variability of absorption by dissolved material at a continental shelf. J. Geophys. Res. 106(C5), 9499-9507.

Butman, D., and P.A. Raymond (2011), Significant efflux of carbon dioxide from streams and rivers in the United States, Nat. Geosci., 4(12), 839-842.

Cai, P., Zhao, D., Wang, L., Huang, B., Dai, M., 2015. Role of particle stock and phytoplankton community structure in regulating particulate organic carbon export in a large marginal sea. J. Geophys. Res. Solid Earth 120(3), 2063-2095.

Cai, W.J., 2011. Estuarine and coastal ocean carbon paradox: CO_2 sinks or sites of terrestrial carbon incineration? Ann Rev Mar Sci 3, 123-145.

Callahan, J., Dai, M., Chen, R.F., Li, X., Lu, Z., Huang, W., 2004. Distribution of dissolved organic matter in the Pearl River Estuary, China. Mar Chem 89(1-4), 211-224.

Carlson, C.A., Ducklow, H.W., Michaels, A.F., 1994. Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. Nature 371(6496), 405-408.

Carter, H.T., Tipping, E., Koprivnjak, J.F., Miller, M.P., Cookson, B., Hamilton-Taylor, J., 2012. Freshwater DOM quantity and quality from a two-component model of UV absorbance. Water Res. 46(14), 4532-4542.

Catalá, T.S., Álvarez-Salgado, X.A., Otero, J., Iuculano, F., Companys, B., Horstkotte, B., et al., 2016. Drivers of fluorescent dissolved organic matter in the global epipelagic ocean. Limnol. Oceanogr. 61(3), 1101-1119.

Chen, C.T.A., Wang, S.L., Wang, B.J., Pai, S.C., 2001. Nutrient budgets for the South China Sea basin. Mar Chem 75(4), 281-300.

Chu, P.C., Fan, C., Lozano, C.J., Kerling, J.L., 1998. An airborne expendable bathythermograph survey of the South China Sea, May 1995. J. Geophys. Res. Oceans 103(C10), 21637-21652.

Cole, J.J., et al., 2007), Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. Ecosystems 10(1), 171-184.

Cui, W., Pan, D., Bai, Y., He, X., Zhu, Q., Hu, Z., Li, T., Gong, F., Zhang, L., 2017. Satellite views of the seasonal and interannual variations of the particulate organic carbon in the northern South China Sea. Acta Oceanologica Sinica 39(3), 122-134.

Dai, M., Cao, Z., Guo, X., Zhai, W., Liu, Z., Yin, Z., … Du, C., 2013a. Why are some marginal seas sources of atmospheric CO_2? Geophys. Res. Lett. 40(10), 2154-2158.

Dai, M., Liu, H., Liao, X., Li, K., Yan, J., Qi, Z., Huang, H., 2013b. Characteristics of net phytoplankton community and their relationships to environmental factors in the waters around Nansha Islands. Chinese Journal of Applied Ecology 24(12).

Dai, M., Meng, F., Tang, T., Kao, S.J., Lin, J., Chen, J., Huang, J., Tian, J., Gan, J., Yang, S., 2009. Excess total organic carbon in the intermediate water of the South China Sea and its export to the North Pacific. Geochemistry, Geophys. Geosystems 10(12).

Diesing, M., Kröger, S., Parker, R., Jenkins, C., Mason, C., Weston, K., 2017. Predicting the standing stock of organic carbon in surface sediments of the North–West European continental shelf. Biogeochemistry 135(1), 183-200.
Doval, M., Hansell, D.A., 2000. Organic carbon an apparent oxygen utilization in the western South Pacific and central Indian Oceans, Mar. Chem. 68, 249-264.
Ducklow, H.W., Steinberg, D.K., Buesseler, K.O., 2001. Upper ocean carbon export and the biological pump. Oceanography 14(4), 50-58.
Dunne, J.P., Sarmiento, J.L., Gnanadesikan, A., 2007. A synthesis of global particle export from the surface ocean and cycling through the ocean interior and on the seafloor. Global Biogeochem Cycles 21(4).
Dzierzbicka-Głowacka, L., Kuliński, K., Maciejewska, A., Jakacki, J., Pempkowiak, J., 2011. Numerical modelling of POC dynamics in the southern Baltic under possible future conditions determined by nutrients, light and temperature. Oceanologia 53(4), 971-992.
Emerson, S., 2014. Annual net community production and the biological carbon flux in the ocean. Global Biogeochem Cycles 28(1), 14-28.
Eppley, R.W., Peterson, B.J., 1979. Particulate organic matter flux and planktonic new production in the deep ocean. Nature 282(5740), 677-680.
Ferrari, G.M., 2000. The relationship between chromophoric dissolved organic matter and dissolved organic carbon in the European Atlantic coastal area and in the West Mediterranean Sea (Gulf of Lions). Mar Chem 70(4), 339-357.
Fichot, C.G., Benner, R., 2011. A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. Geophys. Res. Lett. 38, L03610.
Fu, Y., Zhao, Y., Zhang, Y., Guo, T., He, Z., Chen, J., 2013. GIS and ANN-based spatial prediction of DOC in river networks: a case study in Dongjiang, Southern China. Environ. Earth Sci. 68(5), 1495-1505.
Gardner, W.D., Mishonov, A.V., Richardson, M.J., 2006. Global POC concentrations from in-situ and satellite data. Deep Sea Res. Part II Top. Stud. Oceanogr. 53(5-7), 718-740.
Hansell, D.A., Carlson, C.A., 1998. Net community production of dissolved organic carbon. Global Biogeochem Cycles 12(3), 443-453.
Hansell, D.A., Carlson, C.A., 2001. Marine dissolved organic matter and the carbon cycle. Oceanography 14(4), 41-49.
Helms J.R., Stubbins A., Perdue E.M., Green N.W., Chen H., Mopper K., 2013. Photochemical bleaching of oceanic dissolved organic matter and its effect on absorption spectral slope and fluorescence. Mar. Chem. 155, 81-91.
Hestir, E.L., Brando, V., Campbell, G., Dekker, A., Malthus, T., 2015. The relationship between dissolved organic matter absorption and dissolved organic carbon in reservoirs along a temperate to tropical gradient. Remote Sens Environ 156, 395-402.
Ho, T.Y., Pan, X., Yang, H.H., George, T.F., Shiah, F.K., 2015. Controls on temporal and spatial variations of phytoplankton pigment distribution in the Northern South China Sea. Deep Sea Res. Part II Top. Stud. Oceanogr. 117, 65-85.
Hoge, F.E., Williams, M.E., Swift, R.N., Yungel, J.K., Vodacek, A., 1995. Satellite retrieval of the absorption coefficient of chromophoric dissolved organic matter in continental margins. J. Geophys. Res. Oceans 100(C12), 24847-24854.
Hu, W.J., Jiang, W.W., Hao, T.Y., Xu, Y., Zhao, B.M., 2011. Integrated Geophysical Research on the Distribution of Pre-Cenozoic Residual Basins in the South China Sea. Chinese J. Phys. 54(6), 1138-1148.
Huang, J.J., Wang, S.M., Chen, Y.L., 2007. Biogeochemical controls on distributions and fluxes of dissolved and particulate organic carbon in the Northern South China Sea. Deep Sea Res. Part II Top. Stud. Oceanogr. 54(14-15), 1486-1503.
Iuculano, F., Álverez-Salgado, X.A., Otero, J., Catalá, T.S., Sobrino, C., Duarte, C.M., Agustí, S., 2019. Patterns and drivers of UV absorbing chromophoric dissolved organic matter in the euphotic layer of the open ocean. Front. Mar. Sci. 6, 320.
Jeandel, C., Rutgers, V.D.L.M., Lam, P.J., Roy-Barman, M., Sherrell, R.M., Kretschmer, S., German, C., Dehairs, F., 2015. What did we learn about ocean particle dynamics in the geosecs–jgofs era? Prog. Oceanogr. 133(apr.), 6-16.
Johannessen, S.C., Miller, W.L., Cullen, J.J., 2003. Calculation of UV attenuation and colored dissolved organic matter absorption spectra from measurements of ocean color. J. Geophys. Res. Oceans 108(C9).
Kim, J., Jang, W., Kim, J.H., Lee, J., Cho, K.H., Lee, Y.G., Chon, K., Park, S., Pyo, J., Park, Y., Kim, S., 2022. Application of airborne hyperspectral imagery to retrieve spatiotemporal CDOM distribution using machine learning in a reservoir. Int J Appl Earth Obs Geoinf 114, 103053.
Kowalczuk, P., Cooper, W.J., Durako, M.J., Kahn, A.E., Gonsior, M., Young, H., 2010. Characterization of dissolved organic matter fluorescence in the South Atlantic Bight with use of PARAFAC model: Relationships between fluorescence and its components, absorption coefficients and organic carbon concentrations. Mar Chem 118(1-2), 22-36.
Kowalczuk, P., Tilstone, G.H., Zabłocka, M., Röttgers, R., Thomas, R., 2013. Composition of dissolved organic matter along an Atlantic meridional transect from fluorescence spectroscopy and parallel factor analysis. Mar Chem 157, 170-184.
Lee, D., Son, S., Joo, H., Kim, K., Kim, M.J., Jang, H.K., Yun, M.S., Kang, C.K., Lee, S.H., 2020. Estimation of the particulate organic carbon to chlorophyll-a ratio using MODIS-Aqua in the East/Japan Sea, South Korea. Remote Sens. 12(5), 840.
Li, D., Pan, B., Han, X., Li, J., Zhu, Q., Li, M., 2021. Assessing the potential to use CDOM as an indicator of water quality for the sediment-laden Yellow river, China. Environ. Pollut. 289, 117970.
Li, S., Chen, Y.N., Zhang, J., Song, K., Mu, G., Sun, C., Ju, H.Y., Ji, M., 2018. The relationship of chromophoric dissolved organic matter parallel factor analysis fluorescence and polycyclic aromatic hydrocarbons in natural surface waters. Environ. Sci. Pollut. Res. 25(2), 1428-1438.
Li, Y., Song, G., Massicotte, P., Yang, F., Li, R., Xie, H., 2019. Distribution, seasonality, and fluxes of dissolved organic matter in the Pearl River (Zhujiang) estuary, China. Biogeosciences 16(13), 2751-2770.

Liu, K.K., Kao, S.J., Hu, H.C., Chou, W.C., Hung, G.W., Tseng, C.M., 2007. Carbon isotopic composition of suspended and sinking particulate organic matter in the northern South China Sea—from production to deposition. Deep Sea Res. Part II Top. Stud. Oceanogr. 54(14-15), 1504-1527.

Liu, N., Chen, S., Cheng, Z., Wang, X., Xiao, Y., Xiao, L., Gong, Y., Wang, T., Zhang, X., Liu, S., 2022. Long-term prediction of sea surface chlorophyll-a concentration based on the combination of spatio-temporal features. Water Res. 211, 118040.

Liu, Q., Kandasamy, S., Wang, H., Wang, L., Lin, B., Gao, A., Chen, C.T.A., 2019. Impact of hydrological conditions on the biogeochemical dynamics of suspended particulate organic matter in the upper mixed layer of the Southern East China Sea. J. Geophys. Res. Oceans 124(8), 6120-6140.

Lorenzoni, L., Hu, C., Varela, R., Arias, G., Guzmán, L., Muller-Karger, F., 2011. Bio-optical characteristics of cariaco basin (caribbean sea) waters. Cont Shelf Res 31(6), 582–593.

Ma, H., Yang, W., Zhang, L., Zhang, R., Chen, M., Qiu, Y., Zheng, M., 2017. Utilizing 210Po deficit to constrain particle dynamics in mesopelagic water, western South China Sea. Geochemistry, Geophys. Geosystems 18(4), 1594-1607.

Ma, J., Zhan, H., Du, Y., 2011. Seasonal and interannual variability of surface CDOM in the South China Sea associated with El Niño. J Mar Syst 85(3-4), 86-95.

Mannino, A., Russ, M.E., Hooker, S.B., 2008. Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the US Middle Atlantic Bight. J. Geophys. Res. Oceans 113(C7).

Marhaba, T.F., Bengraïne, K., Pu, Y., Aragó, J., 2003. Spectral fluorescence signatures and partial least squares regression: model to predict dissolved organic carbon in water. J. Hazard. Mater. 97(1-3), 83-97.

Marwick, T.R., Tamooh, F., Teodoru, C.R., Borges, A.V., Darchambeau, F., Bouillon, S., 2015. The age of river-transported carbon: A global perspective. Global Biogeochem Cycles 29(2), 122-137.

Massicotte, P., Asmala, E., Stedmon, C., Markager, S., 2017. Global distribution of dissolved organic matter along the aquatic continuum: Across rivers, lakes and oceans. Sci. Total Environ. 609, 180-191.

Mopper, K., Zhou, X., Kieber, R.J., Kieber, D.J., Sikorski, R.J., Jones, R.D., 1991. Photochemical degradation of dissolved organic carbon and its impact on the oceanic carbon cycle. Nature 353(6339), 60-62.

Moran, M.A., Zepp, R.G., 1997. Role of photoreactions in the formation of biologically labile compounds from dissolved organic matter. Limnol. Oceanogr. 42(6), 1307-1316.

Nelson, N.B., Siegel, D.A., 2013. The global distribution and dynamics of chromophoric dissolved organic matter. Ann Rev Mar Sci 5(1), 447-476.

Nelson, N.B., Siegel, D.A., Michaels, A.F., 1998. Seasonal dynamics of colored dissolved material in the Sargasso Sea. Deep Sea Res. Part I Oceanogr. Res. Pap. 45(6), 931-957.

Pan, X., Wong, G.T., 2015. An improved algorithm for remotely sensing marine dissolved organic carbon: Climatology in the northern South China Sea shelf-sea and adjacent waters. Deep Sea Res. Part II Top. Stud. Oceanogr. 117, 131-142.

Pan, X., Wong, G.T., Ho, T.Y., Shiah, F.K., Liu, H., 2013. Remote sensing of picophytoplankton distribution in the northern South China Sea. Remote Sens Environ 128, 162-175.

Pan, X., Wong, G.T., Shiah, F.K., Ho, T.Y., 2012. Enhancement of biological productivity by internal waves: observations in the summertime in the northern South China Sea. J Oceanogr 68(3), 427-437.

Qu, T., Song, Y.T., Yamagata, T., 2009. An introduction to the South China Sea throughflow: Its dynamics, variability, and application for climate. Dyn. Atmospheres Oceans 47(1-3), 3-14.

Ren, W., Tian, H., Cai, W., Lohrenz, S.E., Hopkinson, C.S., Huang, W., Yang, J., Tao, B., Pan, S., He, R., 2016. Century-long increasing trend and variability of dissolved organic carbon export from the Mississippi River basin driven by natural and anthropogenic forcing. Global Biogeochem Cycles 30(9), 1288-1299.

Rochelle-Newall E.J., Fisher T.R., 2002. Production of chromophoric dissolved organic matter fluorescence in marine and estuarine environments: An investigation into the role of phytoplankton. Mar. Chem. 77, 7-21.

Savini, A., Vertino, A., Marchese, F., Beuck, L., Freiwald, A., 2014. Mapping cold-water coral habitats at different scales within the Northern Ionian Sea (Central Mediterranean): an assessment of coral coverage and associated vulnerability. PLoS One 9(1), e87108.

Schleyer, M.H., Kruger, A., Celliers, L., 2008. Long-term community changes on a high-latitude coral reef in the Greater St Lucia Wetland Park, South Africa. Mar. Pollut. Bull. 56(3), 493-502.

Sexton, P.F., Norris, R.D., Wilson, P.A., Pälike, H., Westerhold, T., Röhl, U., Bolton, C.T., Gibbs, S., 2011. Eocene global warming events driven by ventilation of oceanic dissolved organic carbon. Nature 471(7338), 349-352.

Shao, T., Wang, T., 2020. Effects of land use on the characteristics and composition of fluvial chromophoric dissolved organic matter (CDOM) in the Yiluo River watershed, China. Ecol. Indic. 114, 106332.

Sharpless, C.M., Blough, N.V., 2014. The importance of charge-transfer interactions in determining chromophoric dissolved organic matter (CDOM) optical and photochemical properties. Environ Sci Process Impacts 16(4), 654-671.

Shen, Z., Yang, H., Liu, Q., 2020. Particulate organic carbon and its composition in Jiaozhou Bay. In Studies of the biogeochemistry of typical estuaries and bays in China. (pp. 227-235). Springer, Berlin, Heidelberg.

Shih, Y.Y., Hung, C.C., Tuo, S.H., Shao, H.J., Chow, C.H., Muller, F.L., Cai, Y.H., 2020. The impact of eddies on nutrient supply, diatom biomass and carbon export in the Northern South China Sea. Front. Earth Sci. 8, 537332.

Siegel, D.A., Maritorena, S., Nelson, N.B., Hansell, D.A., Lorenzi-Kayser, M., 2002. Global distribution and dynamics of colored dissolved and detrital organic materials. J. Geophys. Res. 107(C12), 3228.

Siegenthaler, U., Sarmiento, J.L., 1993. Atmospheric carbon dioxide and the ocean. Nature 365(6442), 119-125.
Spencer, R.G., Ahad, J.M., Baker, A., Cowie, G.L., Ganeshram, R., Upstill-Goddard, R.C., Uher, G., 2007. The estuarine mixing behaviour of peatland derived dissolved organic carbon and its relationship to chromophoric dissolved organic matter in two North Sea estuaries (UK). Estuar. Coast. Shelf Sci. 74(1-2), 131-144.
Stewart, A.J., Wetzel, R.G., 1981. Asymmetrical relationships between absorbance, fluorescence, and dissolved organic carbon 1. Limnol. Oceanogr. 26(3), 590-597.
Stramski, D., Reynolds, R.A., Babin, M., Kaczmarek, S., Lewis, M.R., Röttgers, R., Sciandra, A., Stramska, M., Twardowski, M.S., Franz, B.A., Claustre, H., 2008. Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans. Biogeosciences 5(1), 171-201.
Sun, Q., Tang, D., Legendre, L., Shi, P., 2014. Enhanced sea-air CO_2 exchange influenced by a tropical depression in the S outh C hina S ea. J. Geophys. Res. Oceans 119(10), 6792-6804.
Tassan, S., 1994. Local algorithms using SeaWiFS data for the retrieval of phytoplankton, pigments, suspended sediment, and yellow substance in coastal waters. Appl. Opt. 33(12), 2369-2378.
Tehrani, N.C., D'Sa, E.J., Osburn, C.L., Bianchi, T.S., Schaeffer, B.A., 2013. Chromophoric dissolved organic matter and dissolved organic carbon from sea-viewing wide field-of-view sensor (SeaWiFS), moderate resolution imaging spectroradiometer (MODIS) and MERIS sensors: Case study for the northern gulf of mexico. Remote Sens. 5(3), 1439-1464.
Van Pinxteren, M., Barthel, S., Fomba, K.W., Müller, K., Von Tümpling, W., Herrmann, H., 2017. The influence of environmental drivers on the enrichment of organic carbon in the sea surface microlayer and in submicron aerosol particles–measurements from the Atlantic Ocean. Elementa-SCI Anthrop 5.
Vodacek, A., Blough, N.V., DeGrandpre, M.D., DeGrandpre, M.D., Nelson, R.K., 1997. Seasonal variation of CDOM and DOC in the Middle Atlantic Bight: Terrestrial inputs and photooxidation. Limnol. Oceanogr. 42(4), 674-686.
Walsh, J.J., 1991. Importance of continental margins in the marine biogeochemical cycling of carbon and nitrogen. Nature 350(6313), 53-55.
Wang, C., Guo, W., Li, Y., Stubbins, A., Li, Y., Song, G., Wang, L., Cheng, Y., 2017. Hydrological and biogeochemical controls on absorption and fluorescence of dissolved organic matter in the Northern South China Sea. J. Geophys. Res. Biogeosci. 122(12), 3405-3418.
Wang, C., Li, Y., Li, Y., Zhou, H., Stubbins, A., Dahlgren, R.A., Wang, Z., Guo, W., 2021. Dissolved organic matter dynamics in the epipelagic Northwest Pacific low-latitude western boundary current system: insights from optical analyses. J. Geophys. Res. 9(126), e2021JC017458.
Wong, G.T., Ku, T.L., Mulholland, M., Tseng, C.M., Wang, D.P., 2007. The SouthEast Asian time-series study (SEATS) and the biogeochemistry of the South China Sea—an overview. Deep Sea Res. Part II Top. Stud. Oceanogr. 54(14-15), 1434-1447.
Wu, K., Dai, M., Chen, J., Meng, F., Li, X., Liu, Z., Du, C., Gan, J., 2015. Dissolved organic carbon in the South China Sea and its exchange with the Western Pacific Ocean. Deep Sea Res. Part II Top. Stud. Oceanogr. 122, 41-51.
Yan, H., Yu, K., Shi, Q., Tan, Y., Zhang, H., Zhao, M., Li, S., Chen, T., Huang, L., Wang, P., 2011. Coral reef ecosystems in the South China Sea as a source of atmospheric CO_2 in summer. Chinese Science Bulletin 56(7), 676-684.
Yang, F., Song, G., Massicotte, P., Wei, H., Xie, H., 2020. Depth-resolved photochemical lability of dissolved organic matter in the western tropical Pacific Ocean. J. Geophys. Res. Biogeosci. 125(3), e2019JG005425.
Zhang, G., Chen, S., Meng, F., Fan, X., 2017. Inversion of yellow substance concentration and its spatialtemporal variation analysis in Ledong waters, South China Sea. Global Geology, 36(01), 305-310 (In Chinese).
Zhu, Z., Wei, H., Guan, Y., Zhang, L., Jia, R., Sun, P., Yang, Y., Zhang, Z., Zhang, Q., 2022. Distribution, spectral characteristics, and seasonal variation of dissolved organic matter in the northern Beibu Gulf, South China Sea. Front. Mar. Sci. 9.