AGU ADVANCING EARTH AND SPACE SCIENCES

Geophysical Research Letters

RESEARCH LETTER
10.1029/2025GL119009

Special Collection:
Advancing Interpretable AI/ML Methods for Deeper Insights and Mechanistic Understanding in Earth Sciences: Beyond Predictive Capabilities

Predicting Dissolved Black Carbon Concentration From Chromophoric Dissolved Organic Matter Along the Land-Ocean Continuum

Hongyan Bao[1,2], Yuanbi Yi[3], Weiqiang Zhao[1], Jutta Niggemann[4], Julian Merder[5], Keyuan Li[1], Shuwen Li[1], Lejin Peng[1], Xiaoqian Zhan[1], Ziming Fang[1], Weifeng Yang[1], Penghui Li[6,7], Ding He[3,8], Thorsten Dittmar[4,9], and Shuh-Ji Kao[10]

[1]State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen, China, [2]Fujian Provincial Key Laboratory for Coastal Ecology and Environmental Studies, Xiamen University, Xiamen, China, [3]Department of Ocean Science, Center for Ocean Research in Hong Kong and Macau, The Hong Kong University of Science and Technology, Hong Kong, China, [4]Research Group for Marine Geochemistry (ICBM-MPI Bridging Group), Institute for Chemistry and Biology of the Marine Environment (ICBM), School of Mathematics and Science, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany, [5]Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA, [6]School of Marine Sciences, Sun Yat-sen University, Zhuhai, China, [7]Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China, [8]State Key Laboratory of Marine Pollution, City University of Hong Kong, Kowloon, China, [9]Helmholtz Institute for Functional Marine Biodiversity (HIFMB) at the Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany, [10]School of Marine Sciences, State Key Laboratory of Marine Resources Utilization in South China Sea, Hainan University, Haikou, China

Abstract Dissolved black carbon (DBC) plays a key role in global carbon cycle and pollutant transport. However, the time-consuming and labor-intensive chemical analysis limits its spatiotemporal resolution. Here, we developed models to predict DBC from chromophoric dissolved organic matter (CDOM) measurements across the land-to-ocean continuum. We found that the mean ratio of DBC to light absorbance at 254 nm (a_{254}) changed <20% among different environments. However, a single-wavelength model is inadequate for precise prediction due to microbial production of CDOM. Incorporating longer wavelengths using multiple linear regression improves model performance. Random Forest Regression using the full spectral range performed even better at all environments, including the open ocean, achieving a root mean square logarithmic error of <0.15, median symmetric accuracy of <10%, and R^2 of >0.85. This study demonstrates the feasibility of using CDOM to predict DBC concentrations and highlights the potential for in situ monitoring and remote sensing applications.

Plain Language Summary Dissolved black carbon (DBC), one of the most persistent organic components in aquatic systems, plays a key role in the global carbon cycle and pollutant transport. However, the time-consuming and labor-intensive chemical analysis limits the spatiotemporal resolution of DBC concentrations, hindering our understanding of its geochemical behavior. In this study, we developed two new models based on UV-Vis light absorption spectra to predict DBC concentrations across a broad land-to-ocean continuum. Our results show that the random forest regression model effectively predicted DBC concentrations and captured key trends in all studied aquatic systems. Additionally, environment-specific multiple linear regressions using light absorbance at two to three wavelengths also performed well. This study provides a valuable tool for predicting DBC concentrations across diverse aquatic systems. With further refinement, this approach could be integrated into in situ instruments or remote sensing technologies.

Key Points:

- Dissolved black carbon (DBC) concentration was highly significantly correlated with a_{254} across the land-to-ocean continuum
- The random forest model can predict DBC with a median symmetric accuracy of <10% in aquatic systems along the land-to-ocean continuum
- Both linear and machine learning models highlight absorbance in the 260–280 nm and 400–500 nm ranges as predictors of DBC concentration

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:
H. Bao,
baohy@xmu.edu.cn

Citation:
Bao, H., Yi, Y., Zhao, W., Niggemann, J., Merder, J., Li, K., et al. (2025). Predicting dissolved black carbon concentration from chromophoric dissolved organic matter along the land-ocean continuum. *Geophysical Research Letters*, *52*, e2025GL119009. https://doi.org/10.1029/2025GL119009

Received 3 SEP 2025
Accepted 7 NOV 2025

Author Contributions:
Conceptualization: Hongyan Bao
Data curation: Hongyan Bao, Weiqiang Zhao, Jutta Niggemann, Keyuan Li, Shuwen Li, Xiaoqian Zhan, Ziming Fang, Weifeng Yang
Formal analysis: Yuanbi Yi
Funding acquisition: Thorsten Dittmar, Shuh-Ji Kao

1. Introduction

The incomplete combustion of biomass and fossil fuels releases a continuum of thermally altered organic matter, encompassing both biolabile and bio-refractory fractions (Bostick et al., 2021; Martinot et al., 2023; Myers-Pigg et al., 2015; Zhao et al., 2025). Organic carbon formed at higher temperatures is relatively resistant to biodegradation due to its condensed aromatic structure, and this fraction is referred to as black carbon (BC). Dissolved BC (DBC) represents the largest known slow-cycling pool of reduced carbon in the ocean (Dittmar & Paeng, 2009; Ziolkowski & Druffel, 2010) and influences pollutant and nutrient transport, as well as by-product formation with human health implications (Chen et al., 2022; Qu et al., 2016). Understanding DBC transport is

Methodology: Hongyan Bao, Yuanbi Yi, Julian Merder
Validation: Yuanbi Yi
Visualization: Hongyan Bao
Writing – original draft: Hongyan Bao
Writing – review & editing: Yuanbi Yi, Weiqiang Zhao, Jutta Niggemann, Julian Merder, Keyuan Li, Shuwen Li, Lejin Peng, Ziming Fang, Weifeng Yang, Penghui Li, Ding He, Thorsten Dittmar, Shuh-Ji Kao

therefore essential for predicting responses to anthropogenic disturbance and climate change (Coppola et al., 2022).

Current methods for quantifying the refractory fraction of DBC primarily include the benzene polycarboxylic acids (BPCA) method and chemo-thermal oxidation (CTO) (Dittmar, 2008; Wagner et al., 2018). CTO can potentially overestimate concentrations by converting non-black carbon compounds into BC, especially in the dissolved organic matter (DOM) pool (Dittmar, 2008). As a result, the BPCA method is currently the most advanced, yet it is labor-intensive and requires specialized equipment (Dittmar, 2008), limiting spatiotemporal coverage.

Chromophoric DOM (CDOM) absorption offers a rapid, inexpensive, and low-volume alternative. Given the aromatic structure of DBC, CDOM absorption correlates strongly with DBC. For example, in Arctic rivers, DBC showed a strong linear correlation with light absorbance across 254–400 nm (a_{254}–a_{400}), peaking at 254 nm (Stubbins et al., 2015), making CDOM a promising proxy. However, this relationship may not apply universally because CDOM absorption also comes from lignin, tannins (Fichot & Benner, 2012; Fichot et al., 2016), and microbial products (Bao, Qiao, et al., 2023; Nelson et al., 2004). For instance, in the Pacific Ocean, the correlation between a_{254} and DBC was weaker ($R^2 = 0.58$, $p < 0.01$) (Nakane et al., 2017). Furthermore, the slope between DBC and a_{254} can vary across different regions; for example, it was ∼0.0094 in Arctic rivers (Stubbins et al., 2015), but almost half (∼0.0050) in the Arctic Ocean (Fang et al., 2021). These differences indicate that while DBC, as aromatic organic matter, is generally well reflected by a_{254}, local microbial contributions can alter CDOM, reducing correlation and slope in some environments.

Light absorption at other wavelengths reflects different chemical compositions. For instance, absorption near 412 ± 27 nm has been linked to humic substances in polar regions (Granskog et al., 2015). Consequently, ratios such as a_{254}/a_{436} are commonly used to distinguish terrestrial versus autochthonous sources and assess humification (Battin, 1998). Thus, we hypothesize that considering additional wavelengths can help account for variations in microbial and humification processes and improve DBC predictions. We tested this hypothesis by measuring and collecting data across the land-to-ocean aquatic continuum, including rivers, reservoirs, mangrove porewater, estuaries, coastal seas, and the open ocean. Using CDOM absorption spectra, we developed multiple linear regression (MLR) and machine learning (ML) models to predict DBC concentrations. The model performance was evaluated across environments to determine their potential as optical proxies. These models could enable high spatiotemporal resolution and precise prediction of DBC in diverse environments, thereby advancing our understanding of the BC cycle.

2. Materials and Methods

2.1. Field Sampling

Samples were collected from subtropical reservoirs with seasonal stratification, rivers, sediment porewater, coastal sea, and the Northwestern Pacific Ocean (Figure 1a). Water samples from the reservoirs were pumped into clean PC bottles using a Flojet (Xylem) at various depths. Surface water of reservoirs and rivers was collected with a clean bucket, and deep samples in the estuaries, coastal seas and open ocean were obtained using Niskin bottles equipped with CTD. Water samples were then filtered through 0.7 μm GF/F filters (Whatman, pre-combusted at 450°C for 4 hr). Filtrate was subsampled into pre-combusted brown glass bottles and stored at −20°C for CDOM analysis. The remaining filtrate was placed in 2 L clean PC bottles, acidified to pH 2 with concentrated hydrochloric acid, and stored at 4°C for DOM extraction.

2.2. DOM Extraction

Acidified filtrates were concentrated using solid-phase extraction following Dittmar et al. (2008). Cartridges were eluted with 6–8 mL methanol, yielding SPE-DOM, which was stored at −20°C until analysis. Extraction efficiency averaged 49% ± 10% (mean ± standard deviation (sd)).

2.3. CDOM Analysis

Filtered water samples were equilibrated at room temperature and CDOM absorbance was measured from 200 to 800 nm using a Shimadzu UV-1800 spectrophotometer with 1 or 5 cm quartz cuvettes, depending on estimated CDOM abundance (Data set for details (Bao et al., 2025)). Absorbance coefficients (a_λ, m^{-1}) were calculated using the following formula (Helms et al., 2008):

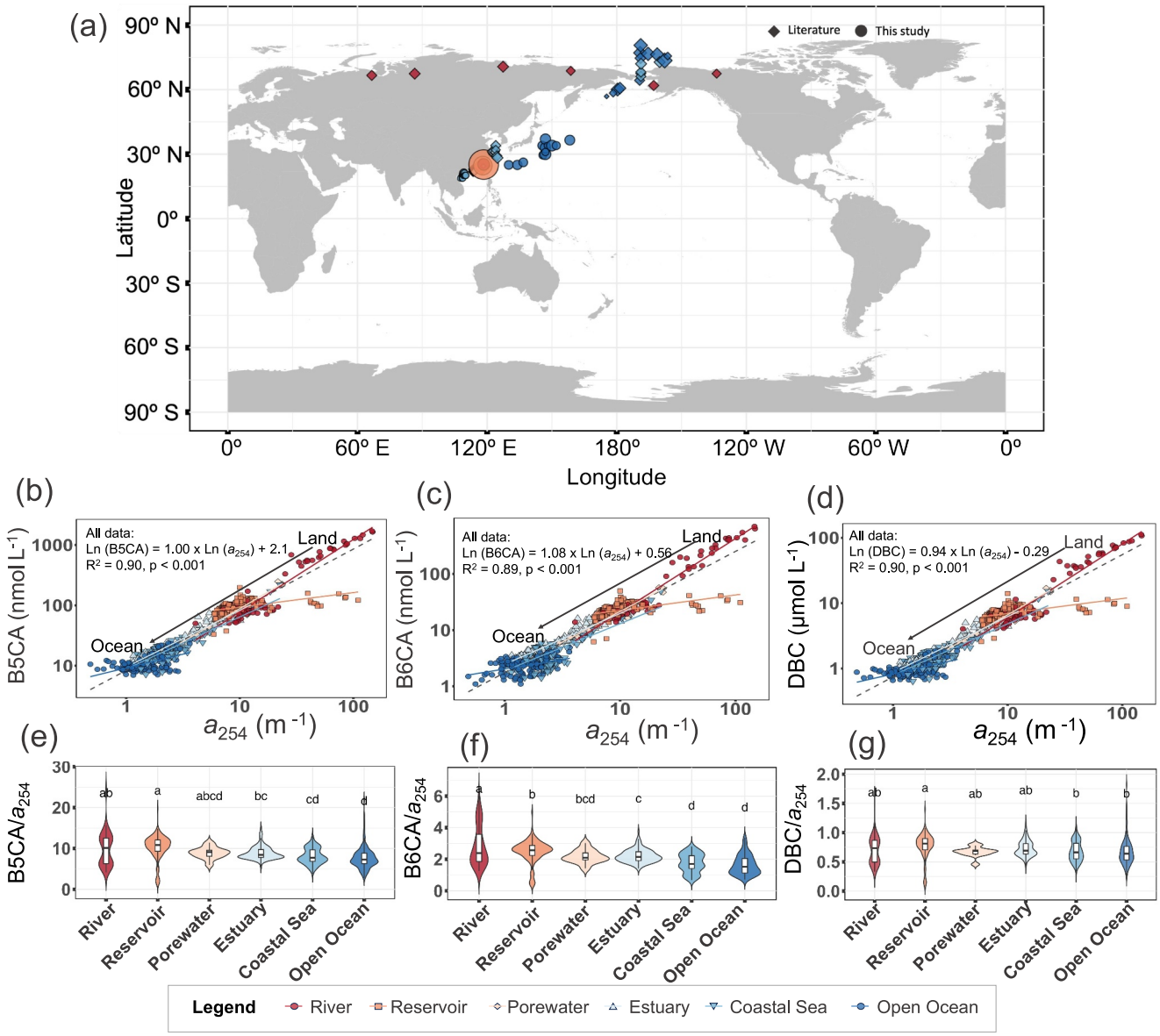

Figure 1. The distribution of sampling sites, correlations between different parameters and the ratios of benzene polycarboxylic acids to a_{254}. (a) Sampling sites, Diamonds: samples from literature; circles: this study. The size of the symbols indicates the number of data, which ranged from 1 to 110; (b)–(d) Relationships between a_{254} and B5CA, B6CA, and Dissolved black carbon (DBC), respectively. Solid lines show environment-specific linear correlations (ln-transformed), and the dashed gray line represents the overall trend. Data sources (Bao et al., 2019, Bao, Niggemann, et al., 2023; Fang et al., 2021; Stubbins et al., 2015; Zhao et al., 2023, 2025), see supplementary data set for the detailed source for each data (Bao et al., 2025). (e) B5CA/a_{254} (nmol L^{-1} (m^{-1})); (f) B6CA/a_{254} (nmol L^{-1} (m^{-1})); (g) DBC/a_{254} (μmol L^{-1} (m^{-1})).

$$a_\lambda = 2.303 \times \frac{A}{l} \quad (1)$$

where A is baseline-corrected absorbance and l the path length. Corrections followed Helms et al. (2008). Testing results showed negligible differences between 1 and 5 cm pathlengths (Figure S1 in Supporting Information S1). Instrument detection limits were 0.04 m^{-1} (5 cm) and 0.2 m^{-1} (1 cm).

2.4. DBC Quantification

DBC in this study was quantified using the BPCA method (Bao et al., 2017; Dittmar, 2008). Analytical error from duplicates was 1%–7%. DBC concentrations were derived from robustly quantified B5CA(1,2,3,4,5-benzenepolycarboxylic acid) and B6CA (1,2,3,4,5,6-benzenepolycarboxylic acid) following Stubbins et al. (2015):

$$\mathrm{DBC}(\mu M) = 0.0891 \times (\mathrm{B5CA}(nM) + \mathrm{B6CA}(nM))^{0.9175} \quad (2)$$

2.5. Data Collection

Additional BPCA data were obtained from published studies spanning the Jiulong River (Bao et al., 2019), Chinese estuaries (Zhao et al., 2023, 2025), marginal seas (Bao, Niggemann, et al., 2023), Arctic rivers (Stubbins et al., 2015), and the Arctic Ocean (Fang et al., 2021). UV-Vis absorbance data for Arctic rivers were sourced from The Arctic Great Rivers Observatory (2024) for Arctic rivers, and from Fang et al. (2021) for the Arctic Ocean. The combined data set covers 20°–80°N and diverse environments, including 158 reservoir, 120 estuarine, 70 river, 12 sediment porewater, 109 coastal sea, and 150 open ocean samples (Bao et al., 2025).

2.6. Model Building

Since BPCAs are the parameters that were directly measured from the instruments, we modeled the concentrations of the two monomers, in addition to DBC. Therefore, B5CA, B6CA and DBC were all modeled in this study. The conditional distributions of DBC, B5CA, and B6CA were best approximated by lognormal distributions (see example for DBC in Figure S2 in Supporting Information S1). Accordingly, we modeled their natural logarithms —ln (DBC), ln (B5CA), and ln (B6CA) for linear models, assuming Gaussian errors, which improved model fit and ensured positive predictions when back-transformed.

Both linear and machine learning (ML) models focused on 260–600 nm, excluding <260 nm (interference by salts) and >600 nm (no peaks) (Johnson & Coletti, 2002). Nitrate absorption near 302 nm was considered but shown to primarily reflect humic substances (Catalá et al., 2016; Sarmiento et al., 2007). Additionally, all absorbance values below a certain value were set to the same value for filtering noise in CDOM absorption, and we tested different thresholds (0.1, 0.01, 0.001) for both ML and MLR models. Although our primary goal is to build a model that could apply to the aquatic systems along the land-to-ocean continuum (global model), we also tested whether environment-specific linear models would improve prediction accuracy. We didn't test the environment-specific ML model to avoid over-fitting due to reduced data size in each category.

2.6.1. Single Linear Regression (a_{254}-Based Model)

Given the known correlation between DBC and a_{254} in Arctic rivers (Stubbins et al., 2015), we tested the correlation between ln (DBC) and ln (a_{254}) across the entire data set to explore the potential for predicting DBC based solely on a_{254}.

2.6.2. Multiple Linear Regression (MLR) Models

For the MLR model, forward selection based on the Bayesian Information Criterion (BIC) (Schwarz, 1978) was used to identify the most relevant wavelengths as predictors. The BIC was chosen over, for example, the Akaike Information Criterion because it applies a stricter penalty for model complexity, making it the more conservative choice for predictor selection and favoring more parsimonious models, which was the goal with our linear model. Linear two-way interactions were included only when the main effects were present (Nelder, 1977). Multicollinearity was checked by confirming that variance inflation factors were below 10 (O'brien, 2007).

2.6.3. Machine Learning (ML) Models

Five classic ML algorithms were tested: Decision Tree Regression (DTR), Gradient Boosting Regression (GBR), Random Forest Regression (RFR), K-Nearest Neighbor Regression, and Support Vector Regression (SVR) (Text S1 in Supporting Information S1). In these models, B5CA, B6CA and DBC concentration were the response variables, and absorption coefficients from 260 to 600 nm were the predictors. All predictors were normalized to have a mean of 0 and a standard deviation of 1 to reduce multicollinearity. The data set was randomly split into 75% training and 25% testing, with tenfold cross-validation used to optimize hyperparameters.

2.6.4. Model Performance Evaluation

Model performance was evaluated using the Root Mean Square Logarithmic Error (RMSLE) and Median Symmetric Accuracy (MSA), calculated as follows:

$$\mathrm{RMSLE}(\mathrm{y},\tilde{\mathrm{y}}) = \sqrt{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}\left(\ln\left(y_i\right)-\ln\left(\tilde{\mathrm{y}}_{\mathrm{i}}\right)\right)^2} \tag{3}$$

$$\mathrm{MSA} = 100(\exp(\mathrm{M}(|\ln \mathrm{Qi}|)) - 1) \tag{4}$$

where $\tilde{\mathrm{y}}_{\mathrm{i}}$ is the measured B5CA and B6CA as well as DBC calculated from them based on Equation 2, y_{i} is the predicted B5CA, B6CA and DBC concentration from our models. M(.) represents the median function, and $Q_i = \hat{\mathrm{m}}_{\mathrm{i}}/\,\tilde{\mathrm{y}}_{\mathrm{i}}$, where $\hat{\mathrm{m}}$ and $\tilde{\mathrm{y}}_{\mathrm{i}}$ is the median estimate and the observed DBC value, respectively, i is the data point index. The RMSLE indicates the mean natural log-transformed error, while the MSA can be interpreted as the median percent increase from the smaller value of the estimated and observed DBC pair to the larger of these two and is frequently used as a superior error metric assessing remote sensing algorithms, such as those for chlorophyll-a (Merder et al., 2024).

The coefficient of determination (R^2) between the predicted and measured values (both natural log-transformed) was also calculated. The Shapley additive explanation (SHAP) from ML models was further calculated to obtain each predictor's positive and negative contribution to the model (Lundberg et al., 2017), while for the MLR the coefficients of the selected wavelengths can be interpreted directly as a power-law relationship, given both response and predictors are log transformed. All statistical analyses, including linear regression, were performed in *R* (v4.2.2, https://www.rproject.org/) using RStudio (RCoreTeam, 2022).

3. Results and Discussion

3.1. The Characteristics of the Data Set

B5CA and B6CA concentrations in the data set ranged from 7.0 to 1699 nmol L^{-1} (mean ± sd: 82 ± 182 nmol L^{-1}) and 1.1–699 nmol L^{-1} (23 ± 66 nmol L^{-1}), respectively. Derived DBC concentrations ranged from 0.64 to 112 μmol L^{-1}, with a mean of 5.9 ± 12 μmol L^{-1} (Figure 1d). The highest DBC values occurred in rivers (24 ± 29 μmol L^{-1}, $n = 70$), particularly Arctic rivers (57 ± 28 μmol L^{-1}, $n = 25$) (Stubbins et al., 2015). Elevated levels were also observed in mangrove sediment porewater (7.8 ± 5.4 μmol L^{-1}, $n = 12$) and reservoirs (7.1 ± 1.6 μmol L^{-1}, $n = 158$) (Figure 1b). In contrast, the open ocean showed the lowest concentrations (1.0 ± 0.30 μmol L^{-1}, $n = 150$), consistent with earlier studies (Dittmar & Paeng, 2009; Zhang et al., 2024), though slightly higher than the North Pacific minimum of 0.2–0.4 μmol L^{-1} (Yamashita et al., 2023; Zhang et al., 2024).

a_{254} has been used as a proxy for DBC previously (Stubbins et al., 2015). In this data set, a_{254} values ranged from 0.56 to 147 m^{-1} for most samples and generally tracked DBC, B5CA, and B6CA concentrations (Figures 1b–1d). However, exceptionally high a_{254} values were observed in reservoirs, where the decoupling from DBC suggests additional sources of light-absorbing material beyond terrestrial inputs.

Ratios of B5CA/a_{254}, B6CA/a_{254}, and DBC/a_{254} decreased from land to ocean (Figures 1e–1g), reflecting reduced terrigenous DOM contribution. Across large environmental gradients, mean DBC/a_{254} ratios vary relatively little—for example, from 0.77 ± 0.22 in reservoirs to 0.67 ± 0.23 in the open ocean—while in contrast, mean DBC/DOC values differ by about fivefold (~0.10 in the global rivers (Jaffé et al., 2013), to ~0.02 in the open ocean (Dittmar & Paeng, 2009)). This relative stability suggests that a_{254} can serve as a robust first-order predictor of DBC along the land–ocean continuum. The strong relationship arises because a_{254} reflects a broader suite of aromatic compounds, while DBC is an aromatic fraction of DOM. Nevertheless, the standard deviations reveal some variability within each environment, likely reflecting differences in aromatic CDOM composition, including potential microbial contributions. Below, we assess how predictive models account for these variations across environments.

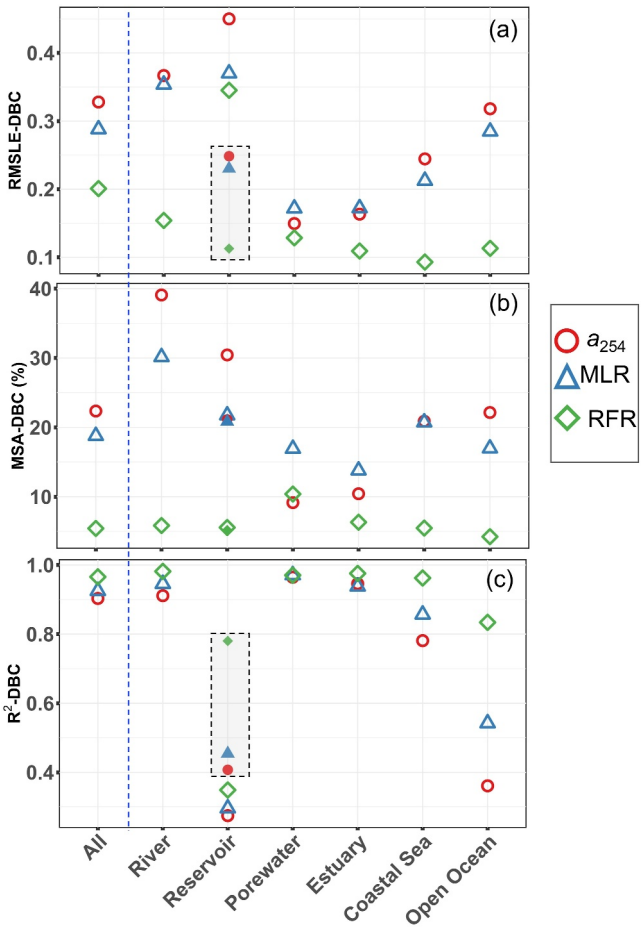

Figure 2. Comparison of a_{254}, multiple linear regression, and random forest regression models for dissolved black carbon prediction across ecosystems. (a) Root mean square log error (RMSLE); (b) median symmetric accuracy (MSA); and (c) coefficient of determination (R^2) between observed and predicted values. Solid symbols represent reservoir samples with outliers removed.

3.2. Modeling DBC From a_{254}

Strong linear relationships were observed between ln (a_{254}) and ln (B5CA), ln (B6CA) and ln (DBC), with $R^2 > 0.90$ ($n = 619$; Figure 1), suggesting similar sources or controlling factors. Residuals showed some deviations at extreme concentrations (Figure S3 in Supporting Information S1). For DBC, RMSLE ranged 0.15–0.45, MSA 9%–39%, and R^2 0.27–0.95 (Figure 2) with similar patterns for B5CA and B6CA (Figure S4 in Supporting Information S1). Yet, a single global a_{254}–DBC relationship produced higher errors in some settings. In rivers, R^2 exceeded 0.95, but RMSLE and MSA were still >0.3 and >30%, respectively. Reservoirs, coastal seas, and the open ocean showed even lower R^2 values (<0.5) and high errors.

Reservoirs had the poorest performance, likely due to bottom-water samples with elevated a_{254} (40–110 m^{-1}) but unchanged DBC levels (Figures 1b–1d; Figure S5 in Supporting Information S1), classified as "outliers." Removing these outliers reduced DBC RMSLE from 0.45 to 0.25, MSA from 30% to 21%, and improved R^2 from 0.27 to 0.41 (Figures 2a–2c). The extreme a_{254} values, observed in the stratified warm season, likely reflected microbially produced CDOM linked to high primary productivity (Berggren et al., 2019). Similarly, in mesopelagic open ocean, microbial CDOM may also influence the a_{254}-DBC relationship (Mo et al., 2025). In line with R^2 trends, lower B5CA/a_{254}, B6CA/a_{254}, and DBC/a_{254} ratios in coastal and oceanic waters compared to rivers and reservoirs suggest reduced soil- or DBC-derived CDOM along the continuum (Figure 1).

Environment-specific models did not increase R^2 compared to the global model (Table S2 in Supporting Information S1) but reduced errors in several cases. For example, MSA in rivers dropped from 39% to 24%, and RMSLE in the open ocean declined from 0.31 to 0.20 (Figure S6 in Supporting Information S1). Reservoirs showed the largest error reduction, likely because high-a_{254} samples influenced regression slopes, reducing outlier errors (Figure S7 in Supporting Information S1).

3.3. Multiple Linear Regression (MLR) and Machine Learning (ML) Models

We tested three thresholds (0.1, 0.01, 0.001) for filtering noise in CDOM absorption. For the MLR model, threshold changes caused only minor wavelength shifts: primary wavelengths consistently fell between 261 and 263 nm, while secondary wavelengths varied from 400 to 600 nm (Table S3 in Supporting Information S1). Predicted values remained consistent, with standard deviations <5% (Figure S8 in Supporting Information S1). Errors were highest at 0.1 and 0.001, while 0.01 yielded moderate, stable performance across environments (Figure S9 in Supporting Information S1). For ML models, predictions varied little across thresholds (e.g., RFR of DBC; Figure S8 in Supporting Information S1). To ensure consistency, we adopted 0.01 as the optimal threshold.

3.3.1. MLR Model

For the MLR model, two to three different wavelengths were chosen by the model for B5CA, B6CA and DBC, along with their linear interaction. The final model equations are as follows:

$$\ln(\mathrm{B5CA}) = (1.16 \pm 0.02) \times \ln(a_{263}) - (0.18 \pm 0.02) \times \ln(a_{468}) - (0.023 \pm 0.006) \times \ln(a_{263}) \times \mathrm{Ln}(a_{468}) + (1.66 \pm 0.06) + \epsilon \quad (5)$$

$$\ln(\mathrm{B6CA}) = (1.03 \pm 0.03) \times \ln(a_{260}) - (0.22 \pm 0.02) \times \ln(a_{586}) + (0.12 \pm 0.02) \times \ln(a_{360}) + (0.11 \pm 0.08) + \epsilon \quad (6)$$

$$\ln(\mathrm{DBC}) = (1.09 \pm 0.02) \times \ln(a_{261}) - (0.16 \pm 0.02) \times \ln(a_{468}) - (0.020 \pm 0.006) \times \ln(a_{261}) \times \ln(a_{468}) - (0.75 \pm 0.05) + \epsilon \quad (7)$$

B5CA and B6CA are in nmol L^{-1}, and DBC in μmol L^{-1}. QQ plots confirmed model performance, with minor low-concentration deviations due to absorbance limits, especially in open-ocean samples (Figure S3 in Supporting Information S1). Compared to the a_{254} model, the MLR improved RMSLE, MSA, and R^2 for all parameters, notably in reservoirs and the open ocean where a_{254} performed poorly (Figure 2). In porewaters and estuaries, MSA rose slightly while RMSLE remained stable. Outliers strongly affected reservoir results; excluding them reduced RMSLE from 0.37 to 0.21 and increased R^2 from 0.29 to 0.45. Incorporating longer wavelengths enhanced predictions overall, though rivers and reservoirs still showed RMSLE >0.3, MSA >20%, and $R^2 < 0.5$.

MLR was not applied to porewater due to the limited samples. Environment-specific MLR models (Table S4 in Supporting Information S1) markedly improved performance (Figure S10 in Supporting Information S1). For DBC, RMSLE declined from 0.17 to 0.37 to 0.12–0.21, and MSA from 14%–30% to 8.4%–14%. R^2 increased across environments except reservoirs. Although only one wavelength was chosen for reservoir and estuary, results confirmed the utility of multi-wavelength, environment-specific models and highlighted differing CDOM–DBC relationships among systems.

Both global and environment-specific models showed interaction terms, like ln (a_{263}) × ln (a_{468}) and ln (a_{261}) × ln (a_{468}) significantly improved prediction of B5CA and DBC ($p < 0.01$). Plots and coefficients indicated strong positive contributions from 260 to 263 nm absorption (Figure S11 in Supporting Information S1), consistent with benzene derivatives' absorption maxima near 254 nm. Substituted aromatic structures, including polycyclic aromatics, shift absorption into the 260–300 nm range, explaining their predictive power in aquatic environments (Chatzimichail et al., 2021; N. B. Nelson & Siegel, 2013). Conversely, 468 and 586 nm showed negative contributions. These wavelengths are linked to humic-like CDOM fluorescence (Granskog et al., 2015) and humified organic matter with absorption near 415 ± 27 nm (Catalá et al., 2016), suggesting microbial CDOM production influences longer-wavelength predictors.

The elevated a_{468}/a_{263} ratio observed in the open ocean and in reservoir "outlier" samples further supports this interpretation (Figure S12 in Supporting Information S1), suggesting that microbial production of CDOM may alter the optical properties of DOM, particularly by enhancing long-wavelength absorbance and thereby affecting the relation between DBC and CDOM in the shorter wavelength region.

3.3.2. ML Models

Except for the SVR model, the other ML approaches—DTR, GBR, RFR, and KNN—showed strong predictive skill for B5CA, B6CA, and DBC concentrations (DBC as example in Figures S13–S14 in Supporting Information S1). For these models, RMSLE ranged from 0.18 to 0.36, MSA from 10% to 25%, and R^2 from 0.96 to 0.98 (Figure S14 in Supporting Information S1). Among them, the RFR model performed best across all environments, showing the smallest training–testing differences for all error metrics (RMSLE, MSA, R^2). Importantly, it maintained accuracy even in the open ocean, where DBC concentrations are low and less variable (Dittmar & Paeng, 2009).

The superior performance of RFR arises from averaging many uncorrelated decision trees built from random subsets of data and features, which reduces variance and overfitting while handling noise and multicollinearity effectively (Breiman, 2001). It also requires little parameter tuning, making it well suited for heterogeneous environmental data sets. Based on these advantages, we selected RFR for subsequent analyses. Details on training and validation are provided in Text S2 in Supporting Information S1.

Across ecosystems, RFR model errors were low. For B5CA, B6CA, and DBC, RMSLE ranged 0.11–0.29, 0.14–0.32, and 0.09–0.34, respectively, with highest errors in reservoirs and lowest in marine systems (Figure 2a). MSA values were consistently small—4.8%–9.3%, 5.8%–10.2%, and 4.4%–10.2% (Figure 2b)—indicating reliable relative deviations. R^2 values generally exceeded 0.95, except in reservoirs (0.34–0.43) and the open ocean (0.76–0.87) (Figure 2c). As in linear models, removing reservoir outliers substantially improved accuracy: RMSLE fell to 0.11–0.14 and R^2 rose to 0.77–0.82, comparable to open-ocean levels (Figure 2; Figure S4 in Supporting

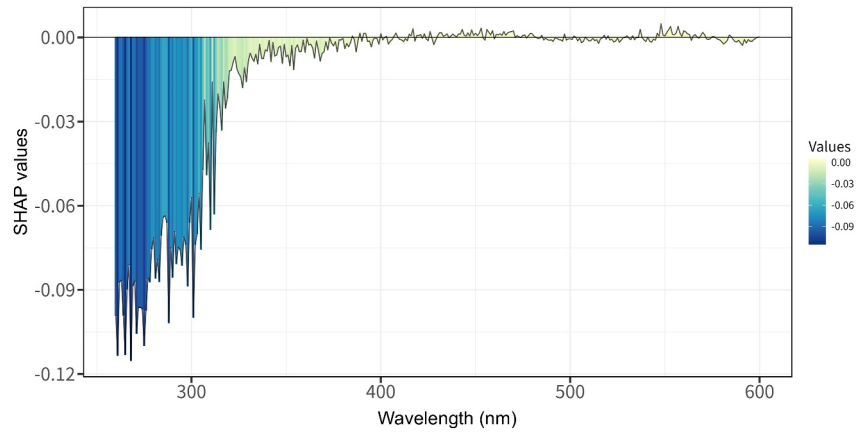

Figure 3. The SHAP value of light absorption at different wavelengths for Dissolved black carbon model. The colors indicate the SHAP values.

Information S1). These metrics align with prior studies where CDOM successfully predicted DOC (mean error: ~4%, Fichot and Benner (2011)), or lignin concentrations (mean error ~16% (Fichot & Benner, 2012)). Together, results confirm that RFR provides robust predictions of DBC dynamics across diverse aquatic environments, including difficult systems like reservoirs and the open ocean. To ensure an unbiased evaluation, model performance was independently validated using an external data set from Zhang et al. (2023). The results further confirmed the superior performance of the RFR model, yielding an MSA of 17%, an RMSLE of 0.24, and an R^2 of 0.90 (Text S3; Figure S15 in Supporting Information S1). Although the validation data were collected from the same estuary as part of the training data set, they represent a different sampling period, allowing an assessment of temporal generalizability. In addition, while the model was trained on samples spanning river-to-ocean gradients, the validation involved only estuarine samples. Despite these limitations, the strong agreement between predicted and observed values supports the robustness and transferability of our approach.

SHAP analysis clarified the main spectral drivers of predictions. Absorbance in the shorter UV range (260–300 nm) dominated, consistent with MLR results showing high coefficients at these wavelengths (Figure 3; Figure S16 in Supporting Information S1). The model also identified strong negative contributions when sub-300 nm values were exceptionally high, revealing a non-linear threshold effect. This likely reflects compositional heterogeneity, where high UV absorbance in some samples arises from lignin-rich or biogenic chromophores rather than BC. In contrast, longer wavelengths (>450 nm) provided weaker, independent positive signals that fine-tuned outputs slightly above the baseline established by shortwave features. This indicates a hierarchical decision structure: predictions are primarily governed by UV absorbance, with secondary variables adding contextual adjustments.

3.4. Potential Factors Influencing the Models

One key challenge in developing predictive models is the spatiotemporal coverage of the data set. In this study, we expanded the data set by adding 327 new DBC measurements and over 499 new full CDOM absorption spectra. Although sampling sites were limited in certain environments—for instance, only two reservoirs were included (see Supplementary Data set (Bao et al., 2025)), —we selected representative systems, such as subtropical reservoirs with seasonal stratification (Xu et al., 2024)). High-resolution vertical (4–10 samples across 40 m) and temporal (every two months over three years) sampling was conducted to better capture the effects of primary production and microbial transformation of DOM on the CDOM spectral, thus helping to mitigate spatial limitations (Li & Hur, 2017).

Sediment porewater samples were also included, but they were primarily collected from a single mangrove, indicating a need for further validation in other porewater environments (see Supplementary Data set (Bao et al., 2025)). Additionally, the training data set was dominated by subtropical samples, which typically exhibit a lower DBC-to-DOC ratio compared to tropical or high-latitude regions (Jones et al., 2020). Since CDOM trends

are often similar to DOC (Li & Hur, 2017), this may cause the model to overestimate DBC concentrations in tropical systems, necessitating regional refinement.

The models also face limitations in anoxic environments—especially in bottom waters of reservoirs, lakes, or coastal seas during stratification—due to the limited number of training samples (10 samples). These may not fully capture the distinctive CDOM characteristics under such conditions. Furthermore, the models were not tested with soil or charcoal leachate, or artificial waters such as industrial and agricultural effluents, which should be explored in future work.

3.5. Comparison of the Three Models and Recommendations

B5CA, B6CA and DBC were all modeled in this study. The DBC calculated from modeled B5CA and B6CA based on Equation 2 is almost identical to the DBC modeled directly (Figure S17 in Supporting Information S1), except for some reservoir samples in the RFR model. Thus, users are encouraged to use the modeled DBC results when interested only in DBC, and to apply the respective models for B5CA and B6CA when focusing on these individual monomers.

Our findings demonstrate that CDOM is a powerful tool for predicting DBC concentrations in aquatic environments. Depending on monitoring goals and frequency, different models can be applied. For long-term or high-frequency monitoring—such as tracking wildfire impacts or DBC variation in drinking water—environment-specific linear models using key wavelengths (e.g., a_{254}, a_{261}, a_{263}, a_{468}) are recommended. These a_{254}-and MLR-based models perform nearly as well as ML models (except in porewaters), while being computationally efficient and compatible with diode array sensors for continuous, low-storage monitoring. Environment-specific models are provided in Tables S2 and S4 in Supporting Information S1, and their accuracy can be further enhanced by adding regional data for parameter optimization. Model predictions can be further improved by incorporating regional data to fine-tune parameters.

For detailed geochemical studies requiring higher accuracy, we recommend using the ML model developed in this study. It is publicly accessible through the following web portal: https://yuanbiyi.shinyapps.io/Shiny_DBCs/. CDOM measurements require only a few milliliters of water, and portable full-spectrum UV-Vis spectrophotometers are now commercially available. With minor adjustments, in situ DBC prediction is feasible, significantly enhancing our capacity to study the biogeochemical cycling of DBC in diverse aquatic systems.

4. Conclusions and Future Overlook

This study provides the first comprehensive data set of DBC across the land-to-ocean continuum. We show that single-wavelength models (e.g., a_{254}) effectively predict B5CA, B6CA, and DBC in estuaries and mangrove porewaters dominated by terrestrial inputs. In more dynamic environments, however, microbial transformation and degradation reduce their accuracy. Incorporating longer wavelengths through MLR significantly improves model performance, consistent with microbial CDOM production signals. Using the full UV-Vis range (260–600 nm), ML models achieved strong predictions across all environments, including low-gradient systems like the open ocean and reservoirs ($R^2 > 0.8$, RMSLE <0.15, MSA <10%).

These findings demonstrate the potential of CDOM optical properties for estimating DBC across diverse aquatic environments. With the growing availability of wavelength-specific and full-spectrum sensors, regionally optimized in situ monitoring is now feasible. Given the strong correlations among CDOM, DOC, DBC and lignin phenols (Fichot & Benner, 2011; Fichot et al., 2016), multi-target ML and remote sensing approaches hold promise for mapping DBC variability and advancing understanding of DOM cycling.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

All training data, including DBC concentrations, full CDOM absorption spectra, predicted values, and data sources, were provided in a spreadsheet file. In addition, the *R* code for MLR models and a template for applying the ML model are also provided. They are all available on Figshare under a CC BY 4.0 license (Bao et al., 2025).

Acknowledgments
This study is funded by the National Key Research and Development Program of China (2024YFC2815800), the National Natural Science Foundation (Nos. 42076041, 92251306, 41721005), the Fundamental Research Funds for the Central Universities (20720250033). The authors acknowledge Wenbin Zou for assisting in the field sampling. Xinhuang Lin and Ina Ulber are acknowledged for assisting in the laboratory analysis of DBC. Dr. Hongyan Bao acknowledges the Alexander von Humboldt Foundation for financially supporting her stay in Germany.

References

Bao, H., Niggemann, J., Du, M., Zhao, W., Huang, D., Yi, Y., et al. (2023). Deciphering sources and processing of dissolved Black carbon in coastal seas. *Limnology & Oceanography*, *68*(11), 2562–2575. https://doi.org/10.1002/lno.12442

Bao, H., Niggemann, J., Huang, D., Dittmar, T., & Kao, S. J. (2019). Different responses of dissolved Black carbon and dissolved lignin to seasonal hydrological changes and an extreme rain event. *Journal of Geophysical Research: Biogeosciences*, *124*(3), 479–493. https://doi.org/10.1029/2018jg004822

Bao, H., Niggemann, J., Luo, L., Dittmar, T., & Kao, S. J. (2017). Aerosols as a source of dissolved Black carbon to the ocean. *Nature Communications*, *8*(1), 510. https://doi.org/10.1038/s41467-017-00437-3

Bao, H., Qiao, J., Huang, D., Niggemann, J., Yi, Y., Zhao, W., et al. (2023). Molecular level characterization of the biolability of rainwater dissolved organic matter. *Science of the Total Environment*, *862*, 160709. https://doi.org/10.1016/j.scitotenv.2022.160709

Bao, H., Yi, Y., Zhao, W., Niggemann, J., Merder, J., Li, K., et al. (2025). Supplementary dataset for predicting dissolved Black carbon concentration from chromophoric dissolved organic matter absorbance along the land-ocean continuum [Dataset]. *figshare*. https://doi.org/10.6084/m9.figshare.27932025

Battin, T. J. (1998). Dissolved organic matter and its optical properties in a blackwater tributary of the upper orinoco river, Venezuela. *Organic Geochemistry*, *28*(9), 561–569. https://doi.org/10.1016/S0146-6380(98)00028-X

Berggren, M., Gudasz, C., Guillemette, F., Hensgens, G., Ye, L., & Karlsson, J. (2019). Systematic microbial production of optically active dissolved organic matter in subarctic Lake water. *Limnology & Oceanography*, *65*(5), 951–961. https://doi.org/10.1002/lno.11362

Bostick, K. W., Zimmerman, A. R., Goranov, A. I., Mitra, S., Hatcher, P. G., & Wozniak, A. S. (2021). Biolability of fresh and photodegraded pyrogenic dissolved organic matter from laboratory-prepared chars. *Journal of Geophysical Research: Biogeosciences*, *126*(5), e2020JG005981. https://doi.org/10.1029/2020jg005981

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Catalá, T. S., Reche, I., Ramón, C. L., López-Sanz, À., Álvarez, M., Calvo, E., & Álvarez-Salgado, X. A. (2016). Chromophoric signatures of microbial by-products in the dark ocean. *Geophysical Research Letters*, *43*(14), 7639–7648. https://doi.org/10.1002/2016gl069878

Chatzimichail, S., Rahimi, F., Saifuddin, A., Surman, A. J., Taylor-Robinson, S. D., & Salehi-Reyhani, A. (2021). Hand-portable HPLC with broadband spectral detection enables analysis of complex polycyclic aromatic hydrocarbon mixtures. *Communications Chemistry*, *4*(1), 17. https://doi.org/10.1038/s42004-021-00457-7

Chen, H., Wang, J., Zhao, X., Wang, Y., Huang, Z., Gong, T., & Xian, Q. (2022). Occurrence of dissolved Black carbon in source water and disinfection byproducts formation during chlorination. *Journal of Hazardous Materials*, *435*, 129054. https://doi.org/10.1016/j.jhazmat.2022.129054

Coppola, A. I., Wagner, S., Lennartz, S. T., Seidel, M., Ward, N. D., Dittmar, T., et al. (2022). The Black carbon cycle and its role in the Earth system. *Nature Reviews Earth and Environment*, *3*(8), 516–532. https://doi.org/10.1038/s43017-022-00316-6

Dittmar, T. (2008). The molecular level determination of black carbon in marine dissolved organic matter. *Organic Geochemistry*, *39*(4), 396–407. https://doi.org/10.1016/j.orggeochem.2008.01.015

Dittmar, T., Koch, B., Hertkorn, N., & Kattner, G. (2008). A simple and efficient method for the solid-phase extraction of dissolved organic matter (SPE-DOM) from seawater. *Limnology and Oceanography: Methods*, *6*, 230–235. https://doi.org/10.4319/lom.2008.6.230

Dittmar, T., & Paeng, J. (2009). A heat-induced molecular signature in marine dissolved organic matter. *Nature Geoscience*, *2*(3), 175–179. https://doi.org/10.1038/ngeo440

Fang, Z., Yang, W., Stubbins, A., Chen, M., Li, J., Jia, R., et al. (2021). Spatial characteristics and removal of dissolved black carbon in the western Arctic Ocean and Bering Sea. *Geochimica et Cosmochimica Acta*, *304*, 178–190. https://doi.org/10.1016/j.gca.2021.04.024

Fichot, C. G., & Benner, R. (2011). A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. *Geophysical Research Letters*, *38*(3), L03610. https://doi.org/10.1029/2010gl046152

Fichot, C. G., & Benner, R. (2012). The spectral slope coefficient of chromophoric dissolved organic matter (S275-295) as a tracer of terrigenous dissolved organic carbon in river-influenced ocean margins. *Limnology & Oceanography*, *57*(5), 1453–1466. https://doi.org/10.4319/lo.2012.57.5.1453

Fichot, C. G., Benner, R., Kaiser, K., Shen, Y., Amon, R. M. W., Ogawa, H., & Lu, C.-J. (2016). Predicting dissolved lignin phenol concentrations in the coastal Ocean from chromophoric dissolved organic matter (CDOM) absorption coefficients. *Frontiers in Marine Science*, *3*. https://doi.org/10.3389/fmars.2016.00007

Granskog, M. A., Pavlov, A. K., Sagan, S., Kowalczuk, P., Raczkowska, A., & Stedmon, C. A. (2015). Effect of sea-ice melt on inherent optical properties and vertical distribution of solar radiant heating in arctic surface waters. *Journal of Geophysical Research: Oceans*, *120*(10), 7028–7039. https://doi.org/10.1002/2015jc011087

Helms, J. R., Stubbins, A., Ritchie, J. D., Minor, E. C., Kieber, D. J., & Mopper, K. (2008). Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnology & Oceanography*, *53*(3), 955–969. https://doi.org/10.4319/lo.2008.53.3.0955

Jaffé, R., Ding, Y., Niggemann, J., Vahatalo, A. V., Stubbins, A., Spencer, R. G., et al. (2013). Global charcoal mobilization from soils via dissolution and riverine transport to the oceans. *Science*, *340*(6130), 345–347. https://doi.org/10.1126/science.1231476

Johnson, K. S., & Coletti, L. J. (2002). In situ ultraviolet spectrophotometry for high resolution and long-term monitoring of nitrate, bromide and bisulfide in the ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, *49*(7), 1291–1305. https://doi.org/10.1016/S0967-0637(02)00020-1

Jones, M. W., Coppola, A. I., Santin, C., Dittmar, T., Jaffe, R., Doerr, S. H., & Quine, T. A. (2020). Fires prime terrestrial organic carbon for riverine export to the global oceans. *Nature Communications*, *11*(1), 2791. https://doi.org/10.1038/s41467-020-16576-z

Li, P., & Hur, J. (2017). Utilization of UV-Vis spectroscopy and related data analyses for dissolved organic matter (DOM) studies: A review. *Critical Reviews in Environmental Science and Technology*, *47*(3), 131–154. https://doi.org/10.1080/10643389.2017.1309186

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, *30*.

Martinot, P. L., Guigue, C., Chifflet, S., Cuny, P., Barani, A., Didry, M., et al. (2023). Assessing the bioavailability of Black carbon-derived dissolved organic matter for marine heterotrophic prokaryotes. *Science of the Total Environment*, *901*, 165802. https://doi.org/10.1016/j.scitotenv.2023.165802

Merder, J., Zhao, G., Pahlevan, N., Rigby, R. A., Stasinopoulos, D. M., & Michalak, A. M. (2024). A novel algorithm for ocean chlorophyll-a concentration using MODIS aqua data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *210*, 198–211. https://doi.org/10.1016/j.isprsjprs.2024.03.014

Mo, S., Liu, Z., Yuanhao, H., Hertkorn, N., Wang, H., Zhang, C., et al. (2025). Unveiling ongoing biogeochemical dynamics of CDOM from surface to deep ocean. *Nature Communications*, *16*(1), 5202. https://doi.org/10.1038/s41467-025-60510-0

Myers-Pigg, A. N., Louchouarn, P., Amon, R. M. W., Prokushkin, A., Pierce, K., & Rubtsov, A. (2015). Labile pyrogenic dissolved organic carbon in major Siberian arctic Rivers: Implications for wildfire-stream metabolic linkages. *Geophysical Research Letters*, *42*(2), 377–385. https://doi.org/10.1002/2014gl062762

Nakane, M., Ajioka, T., & Yamashita, Y. (2017). Distribution and sources of dissolved black carbon in surface waters of the Chukchi sea, Bering Sea, and the north Pacific Ocean. *Frontiers in Earth Science*, *5*, 34. https://doi.org/10.3389/feart.2017.00034

Nelder, J. A. (1977). A reformulation of linear models. *Journal of the Royal Statistical Society. Series A (General)*, *140*(1), 48. https://doi.org/10.2307/2344517

Nelson, N. B., Carlson, C. A., & Steinberg, D. K. (2004). Production of chromophoric dissolved organic matter by Sargasso Sea microbes. *Marine Chemistry*, *89*(1–4), 273–287. https://doi.org/10.1016/j.marchem.2004.02.017

Nelson, N. B., & Siegel, D. A. (2013). The global distribution and dynamics of chromophoric dissolved organic matter. *Annual Review of Marine Science*, *5*(1), 447–476. https://doi.org/10.1146/annurev-marine-120710-100751

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality and Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Qu, X., Fu, H., Mao, J., Ran, Y., Zhang, D., & Zhu, D. (2016). Chemical and structural properties of dissolved Black carbon released from biochars. *Carbon*, *96*, 759–767. https://doi.org/10.1016/j.carbon.2015.09.106

RCoreTeam. (2022). R: A language and environment for statistical computing. In *R foundation for statistical computing*.

Sarmiento, J. L., Simeon, J., Gnanadesikan, A., Gruber, N., Key, R. M., & Schlitzer, R. (2007). Deep ocean biogeochemistry of silicic acid and nitrate. *Global Biogeochemical Cycles*, *21*(1). https://doi.org/10.1029/2006gb002720

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Stubbins, A., Spencer, R. G. M., Mann, P. J., Holmes, R. M., McClelland, J. W., Niggemann, J., & Dittmar, T. (2015). Utilizing colored dissolved organic matter to derive dissolved black carbon export by arctic Rivers. *Frontiers in Earth Science*, *3*. https://doi.org/10.3389/feart.2015.00063

The Arctic Great Rivers Observatory. (2024). Absorbance dataset (version 20240402) [Dataset]. Retrieved from https://www.arcticrivers.org/data

Wagner, S., Jaffé, R., & Stubbins, A. (2018). Dissolved Black carbon in aquatic ecosystems. *Limnology and Oceanography Letters*, *3*(3), 168–185. https://doi.org/10.1002/lol2.10076

Xu, Z., Ge, L., Zou, W., Lv, B., Yang, J., Chai, Z., et al. (2024). The underestimated role of manganese in modulating the nutrient structure in a eutrophic seasonally-stratified reservoir. *Water Research*, *260*, 121940. https://doi.org/10.1016/j.watres.2024.121940

Yamashita, Y., Mori, Y., & Ogawa, H. (2023). Hydrothermal-derived Black carbon as a source of recalcitrant dissolved organic carbon in the ocean. *Science Advances*, *9*(6), eade3807. https://doi.org/10.1126/sciadv.ade3807

Zhang, Q., Zhou, J., Fang, Z., Yang, W., Chen, M., & Zheng, M. (2023). Sources and dynamics of dissolved Black carbon in the pearl river Estuary and shelf, northern South China Sea. *Journal of Oceanography*, *80*(1), 71–83. https://doi.org/10.1007/s10872-023-00708-2

Zhang, X., Wang, Y., Liu, Z., Liu, B., Wu, W., Liu, L., et al. (2024). Heterogeneous sources, distribution, and removal processes of dissolved Black carbon from East China Sea shelf to open ocean of northwest Pacific. *Progress in Oceanography*, *229*, 103374. https://doi.org/10.1016/j.pocean.2024.103374

Zhao, W., Bao, H., Huang, D., Niggemann, J., Dittmar, T., & Kao, S.-J. (2023). Evidence from molecular marker and FT-ICR-MS analyses for the source and transport of dissolved Black carbon under variable water discharge of a subtropical Estuary. *Biogeochemistry*, *162*(1), 43–55. https://doi.org/10.1007/s10533-022-00987-9

Zhao, W., Bao, H., Peng, L., Du, M., Li, K., Zhan, X., et al. (2025). Transport of dissolved black carbon in three estuaries in China: Roles of flood-Ebb tides and submarine groundwater discharge. *Global Biogeochemical Cycles*, *39*(6), e2025GB008532. https://doi.org/10.1029/2025gb008532

Ziolkowski, L. A., & Druffel, E. R. M. (2010). Aged black carbon identified in marine dissolved organic carbon. *Geophysical Research Letters*, *37*(16). https://doi.org/10.1029/2010GL043963