

Article

Four-Decade CDOM Dynamics in Amur River Basin Lakes from Landsat and Machine Learning

Ye Wang¹, Pengfei Han^{1,*} , Chi Zhang¹, Zhuohang Xin¹, Lu Zhang², Xixin Lu³ and Jinkun Huang³¹ School of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China; 12006072@mail.dlut.edu.cn (Y.W.)² Science and Technology Research Institute, China Three Gorges Corporation, Beijing 101199, China³ China MCC17 Group Co., Ltd., Maanshan 243000, China

* Correspondence: hanpf22@dlut.edu.cn

Highlights

What are the main findings?

- A 40-year CDOM dataset (1984–2023) was constructed for 69 large lakes within the Amur River Basin based on support vector regression (SVR) model and multi-decadal Landsat imagery, enabling basin-wide spatiotemporal analysis;
- Results reveal a significant increasing trend of CDOM in 27 lakes, mainly located in the Mongolian Plateau and Northeast Plain, while four lakes show significant declines;
- Hydro-climatic drivers (wind speed, temperature) and anthropogenic pressures (irrigation, grazing) exert regionally distinct influences on lake CDOM dynamics.

What is the implication of the main finding?

- The multi-decadal CDOM dataset offers critical evidence for guiding basin-level monitoring priorities;
- The revealed combined roles of climate forcing and human activities offer guidance for developing adaptive water-quality management strategies at the watershed level.

Abstract

Lakes in the Amur River Basin (ARB) are increasingly influenced by climate variability and human activities, yet long-term basin-scale patterns of colored dissolved organic matter (CDOM) remain unclear. In this study, we developed a support vector regression (SVR) model to retrieve lake CDOM from Landsat 5/7/8 imagery and generated a 40-year (1984–2023) CDOM dataset for 69 large lakes. The model provides a reliable tool for multi-decadal, large-area water quality monitoring considering its robust performance ($R^2 = 0.88$, $rRMSE = 22.4\%$, $MAE = 2.63 \text{ m}^{-1}$). Trend analysis revealed a significant rise in CDOM since 1999, particularly across the Mongolian Plateau and Northeast China Plain. Among the 69 lakes, 27 exhibited increasing CDOM, while 4 showed declines, highlighting pronounced regional variability. Variance partitioning indicated that human activities, especially irrigation and grazing, account for $\sim 30\%$ of CDOM variation, exceeding the contribution of any single climatic driver, whereas temperature represents the dominant climate driver (12.8%). Shallow systems were more sensitive to external disturbances, while deep lakes responded more strongly to thermal conditions. This study delivers the first long-term satellite-based CDOM assessment in the ARB and underscores the combined impacts of climate change and land-use pressures on lake optical dynamics.

Keywords: CDOM; lakes; Amur River Basin; climatic factors; human activities

Academic Editor: Nicolas Baghdadi

Received: 31 October 2025

Revised: 20 December 2025

Accepted: 27 December 2025

Published: 29 December 2025

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article

distributed under the terms and

conditions of the [Creative Commons](https://creativecommons.org/licenses/by/4.0/)[Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Colored dissolved organic matter (CDOM) is a key parameter for assessing water quality and is commonly present in natural aquatic environments [1,2]. CDOM originates primarily from the decomposition of terrestrial vegetation and leaching of soil organic matter, with additional anthropogenic inputs such as industrial effluents, agricultural wastewater, and domestic sewage [3,4]. Elevated CDOM reduces light penetration depth and constrains phytoplankton production due to the strong optical absorption of CDOM [5,6]. Moreover, CDOM forms complexes with heavy metals and organic pollutants that enhance their mobility and ecological risks [7,8]. As a major carrier and reactive fraction of aquatic carbon, the spatiotemporal variability of CDOM influences catchment-to-ocean carbon fluxes and regional carbon budgets [9]. Therefore, accurate characterization of CDOM distribution and drivers is essential for advancing carbon peaking and carbon neutrality goals, particularly in vulnerable high-latitude systems such as the Amur River Basin (ARB). Arctic amplification and intensified human activities jointly accelerate carbon cycling feedbacks and pose challenges to ecological sustainability [10,11].

The ARB is one of the largest transboundary basins in Northeast Asia, spanning China, Russia, and Mongolia. It hosts numerous lakes from large systems such as Xingkai Lake, Hulun Lake, and Chagan Lake to many smaller water bodies, which play vital roles in water supply, biodiversity conservation, and regional ecological security [12,13]. However, lake water quality and ecosystem stability in the basin have been seriously affected due to Arctic amplification and intensified human activities (e.g., agricultural expansion, aquaculture) [14–16]. Although hydrology, eutrophication, and carbon cycling in the ARB have been investigated [10,17,18], CDOM dynamics remain underexplored. Existing work has typically focused on single lakes or river reaches [19–21] and rarely adopts a basin-wide, transboundary perspective, particularly for lakes outside China. These gaps underscore the need for systematic, basin-scale CDOM monitoring to resolve spatiotemporal patterns and their drivers across the ARB.

Satellite remote sensing enables basin-scale, temporally consistent observation and thus supports long-term CDOM monitoring at regional to global scales [8,22]. Sensors such as MODIS (daily) and Sentinel-2 (~5-day) provide high observation frequency, while the Landsat series (5/7/8/9) offers unique advantages for inland waters: 30 m spatial resolution, spectral sensitivity to CDOM absorption (blue-to-NIR ratios), and a continuous 40-year archive spanning 1984–2023 that facilitates robust long-term trend analysis [23,24]. However, retrieving CDOM from inland turbid waters remains challenging due to its strong absorption, which reduces water-leaving radiance, especially in the blue to UV bands [1,25].

To address the challenges, numerous approaches have been developed, from simple band ratio algorithms to semi-analytical models like QAAs (quasi-analytical algorithms) and matrix inversion techniques [19,26,27]. Although these models offer computational simplicity and ease of evaluation [19], their accuracy and robustness are often limited when applied to lakes beyond the specific conditions for which they were developed [8,28]. These limitations arise primarily from the substantial optical heterogeneity of inland water bodies, where the co-variation and spectral overlap of other optically active constituents (e.g., chlorophyll-a, total suspended matter) confound the isolation of CDOM-specific signals [29,30].

Such challenges underscore the value of data-driven approaches, particularly machine learning (ML), which can capture complex and nonlinear relationships between CDOM and remote sensing reflectance [2,31,32]. Recent studies have demonstrated the effectiveness of ML methods. Feng et al. [19] reported that Support Vector Regression (SVR) outperformed four alternative ML algorithms for CDOM retrieval from Sentinel-2 imagery in the Songhua

River Basin, achieving $R^2 = 0.91$ and $rRMSE = 18.2\%$. Keller et al. evaluated ten ML techniques in the Elbe River and found R^2 values exceeding 0.7 [31]. Sun et al. showed that four ML models, including backpropagation neural networks (BP), SVR, random forest regression (RFR), and Gaussian process regression (GPR), applied to Landsat 8 OLI data, achieved accuracies above 70% in inland lakes [24]. Collectively, these advances indicate that integrating multi-year Landsat observations with advanced inversion algorithms provide a robust methodological foundation for systematically characterizing CDOM dynamics in the ARB.

Building on recent advances in satellite retrieval and modelling, this study addresses basin-wide knowledge gaps on CDOM by characterizing long-term patterns in ARB lakes and disentangling the relative contributions of hydro-climatic and socio-economic drivers. Specifically, (1) we develop a CDOM inversion model tailored to the ARB lakes; (2) map and quantify multi-decadal spatiotemporal changes of lake CDOM from Landsat imagery; and (3) apportion driver contributions across scales for key hydro-climatic factors and socio-economic stressors. The results fill a critical gap for CDOM monitoring in the ARB lakes and provide actionable information for transboundary water management and aquatic ecosystem protection.

2. Study Area and Data

2.1. Study Area

The ARB is located in northeast Asia and encompasses a drainage area of approximately 2.08 million km². It is one of the largest transboundary river basins in the world, spanning between about 42–52°N and 108–141°E (Figure 1). Its latitudinal extent drives diverse climates (temperate to subarctic) and heterogeneous land cover, including boreal forests, wetlands underlain by discontinuous permafrost, and intensively farmed plains [10]. Reflecting the basin's multinational character (China, Russia, Mongolia), population density and anthropogenic pressure display marked gradients across the region. Southern regions generally experience higher population density and intensified agricultural/industrial activities compared to the more sparsely populated northern sections [33,34]. Land cover is correspondingly diverse, including extensive cropland, forest, grassland, and wetlands with large permafrost-affected areas in the north.

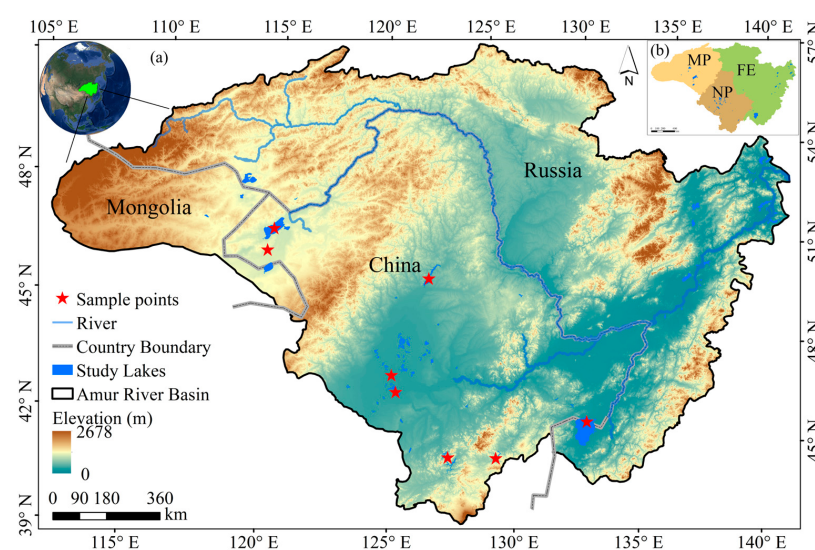


Figure 1. (a) Map of the Amur River Basin (ARB) showing the locations of the 69 lakes larger than 20 km² analyzed in this study. Blue polygons indicate all study lakes and red stars highlight the sample lakes. The Amur River Basin was divided into three sub-regions (b), namely, NP: the Northeast Plain region, MP: Mongolian Plateau region, and FE: Far East Federal region.

The ARB contains 1138 lakes larger than 1 km² [17]. These lakes are essential components of the regional ecosystem since they regulate water storage, sustain unique biodiversity (e.g., migratory waterbirds and cold-adapted fish), and mediate terrestrial-aquatic carbon fluxes [12,13]. Among these, medium-to-giant lakes (>20 km²) disproportionately dominate hydrological and biogeochemical functions, storing > 80% of total lake water volume while serving as sentinels of Arctic-amplified climate change. Hence, this study focuses on lakes larger than 20 km² in the ARB. A total of 69 lakes were selected (Figure 1) which span a broad latitudinal gradient, ranging from northern permafrost-affected lakes with limited human disturbance to southern reservoirs exposed to intensive agricultural activities. This spatial coverage enables us to capture the heterogeneity of hydro-climatic stressors and anthropogenic pressures. The basic characteristics of the lakes [17] are provided in Table S1.

2.2. Sampling Sites

As illustrated in Figure 1, field surveys were carried out in eight lakes, including Chagan Lake, Yueliang Lake, Xiaoxingkai Lake, Jingpo Lake, Nierji Lake, Hulun Lake, Songhua Lake, and Wulannuoer Lake, during the period from September 2020 to June 2023. These lakes represent diverse typologies across environmental gradients, ranging from shallow lowland lakes to deep mountainous reservoirs. The selection ensures representative coverage of lake and reservoir water quality conditions across the ARB. Altogether, 203 water samples were collected at a depth of approximately 0.5 m below the surface using 1 L acid-washed plastic bottles. Immediately after collection, the samples were stored in lightproof coolers with ice packs to maintain a temperature below 4 °C, which minimized the risks of photodegradation and microbial alteration of CDOM. After transportation to the laboratory, the samples were kept at 4 °C in a darkness until filtration and subsequent analysis.

Each water sample was filtered through a 0.45 µm polycarbonate filter membrane, followed by an additional filtration through a 0.22 µm polycarbonate filter membrane. This methodology was employed to prepare the water samples specifically for the analysis of CDOM. The absorbance of CDOM was measured using a UV-2600 PC spectrophotometer (Beijing, China), covering a wavelength range of 200 to 800 nm at 1 nm intervals. Subsequently, the absorption coefficient was calculated from the absorbance values at each wavelength using Equation (1), and corrections for scattering effects were applied by subtracting the absorbance at 700 nm as described in Equation (2) [2].

$$\alpha^*_{\text{CDOM}}(\lambda) = 2.303 \times A(\lambda)/L \quad (1)$$

$$\alpha_{\text{CDOM}}(\lambda) = \alpha^*_{\text{CDOM}}(\lambda) - \alpha^*_{\text{CDOM}}(700) \times \lambda/700 \quad (2)$$

where λ is the wavelength (nm); the length of the cuvette L is 1 cm; $A(\lambda)$ is the absorbance; $\alpha_{\text{CDOM}}(\lambda)$ represents the corrected absorption coefficient of the CDOM (m^{−1}); $\alpha^*_{\text{CDOM}}(\lambda)$ represents the uncorrected absorption coefficient at a given wavelength. In prior research, the absorption coefficients at 355 nm have typically been employed as indicators of CDOM in inland lakes and rivers [19,35]. Consequently, this study selects the absorption coefficient at 355 nm to characterize the CDOM present in lakes within the ARB. Table S2 compiles field observation data for CDOM absorption coefficients, where $\alpha_{\text{CDOM}}(355)$ values range from 1.64 to 11.57 m^{−1} (mean ± SD: 5.49 ± 2.72 m^{−1}), reflecting significant spatial heterogeneity across ARB lakes.

2.3. Satellite Data and Processing

Previous studies have widely applied Landsat imagery to assess lake water quality [24,36]. In this study, we used Level-2 surface reflectance (SR) products from Landsat-5 TM, Landsat-7 ETM+, and Landsat-8 OLI, which were accessed from Google

Earth Engine (GEE) (<https://earthengine.google.com/>) on 13 July 2024. These datasets correspond to USGS Collection 2 Surface Reflectance, which are generated using the Landsat Surface Reflectance Code (LaSRC) atmospheric correction algorithm. The LaSRC algorithm incorporates a coastal aerosol model and represents the current operational standard for producing Landsat surface reflectance data [37]. All three datasets have a spatial resolution of 30 m and a temporal resolution of 16 days [38]. To avoid the influence of ice and snow, only images acquired during the ice-free season from May to October were used [39]. In addition, the CFMask algorithm implemented in GEE was applied to remove pixels contaminated by clouds, cloud shadows, or snow [40,41].

To establish a reliable dataset for model development, in situ CDOM measurements were matched with Landsat-8 OLI surface reflectance in space and time. For each field sample, a 3×3 pixel window centered on the corresponding OLI location was extracted, and the mean reflectance of these nine pixels was used to represent the spectral signal at the sampling site. This averaging approach reduces the influence of geolocation uncertainty, mixed pixels, and sensor noise, thereby providing a more robust match between in situ measurements and satellite observations [42]. A temporal window of ± 7 days between the field sampling date and the satellite overpass was adopted to balance data availability and accuracy as recommended in previous work [43]. After applying quality control (e.g., removing pixels with high cloud probability or adjacency effects from land), a total of 194 valid spatiotemporally matched pairs were obtained for subsequent algorithm calibration and validation.

Consistency across different Landsat sensors is critical for reliable long-term retrieval of CDOM. To achieve this, additional Landsat-7 images from 2009–2011 and 2018–2020 were collected to facilitate cross-calibration with Landsat-5 and Landsat-8. Specifically, random sampling of more than 1000 lake pixels for each period was conducted where images were spatiotemporally matched. Correlations were then established between $\alpha_{\text{CDOM}}(355)$ estimated from synchronous Landsat-7 and $\alpha_{\text{CDOM}}(355)$ estimated from Landsat-5 TM or Landsat-8 OLI. Based on the regression relationships, all CDOM values derived from Landsat-5 and Landsat-7 were converted to the Landsat-8 equivalent scale using a transfer function.

2.4. Auxiliary Data

To explore the mechanism and the primary drivers of CDOM dynamics in lakes across the ARB, we compiled multi-source datasets encompassing hydro-climatic factors (e.g., temperature, precipitation), socio-economic pressures (e.g., grazing, irrigation), and land cover properties (NDVI). Hydro-climatic data, including precipitation (Prep), wind speed (WS), runoff, and temperature (Temp), are recognized as key determinants of water quality dynamics in aquatic systems [3,44]. These data were obtained from the enhanced global dataset of the fifth generation European Reanalysis Land Segment (ERA5-Land) provided by the Copernicus Climate Data Store (<https://cds.climate.copernicus.eu/>, accessed on 23 October 2024). ERA5-Land is a global reanalysis product that offers high accuracy and fine spatial resolution ($0.1^\circ \times 0.1^\circ$, approximately 10 km at the equator) [45,46]. A 5 km riparian buffer was selected to extract hydro-climatic variables to ensure consistency across lakes with highly heterogeneous watershed boundaries.

In addition to natural drivers, socio-economic factors were considered to evaluate human influences on CDOM dynamics. Grazing intensity (Gr, number of sheep and goats) and irrigation extent (Ir, area of irrigated croplands) were selected as indicators of anthropogenic pressures because they can significantly alter lake biogeochemistry [5,47]. Grazing and irrigation data were collected from the National Bureau of Statistics of China (<http://www.stats.gov.cn/>) on 26 October 2024 and the Food and Agriculture Organization

of the United Nations (<https://www.fao.org/faostat/>) on 26 October 2024. County-level statistics were downscaled to lake watersheds by weighting with the proportional area of cropland, grassland, and rural land cover.

Vegetation conditions were also included as potential regulators of CDOM. Vegetation can increase organic matter inputs through plant production while simultaneously reducing terrestrial inputs to lakes by improving soil erosion control [48]. We quantified land greenness and vegetation density using the Normalized Difference Vegetation Index (NDVI) derived from Landsat images [49]. The same 5 km riparian buffers applied for hydro-climatic variables were also used to extract NDVI values.

3. Methodology

3.1. Algorithm Development

Remote sensing provides unique advantages for large-scale monitoring of CDOM [8]. To identify the most suitable CDOM retrieval models for ARB lakes, we evaluated a set of established Landsat-based algorithms that include empirical and semi-empirical models [4,50–52], as well as semi-analytical approaches [53]. Although we recalibrated model coefficients using 194 in situ CDOM measurements matched with Landsat reflectance, the overall performance of these algorithms remained unsatisfactory. A summary of the tested models, including the input bands or band ratios and their mathematical formulations, is shown in Table 1. Empirical band ratio models showed poor performance under conditions of strong optical heterogeneity [54]. In addition, the spectral overlap between CDOM and suspended sediment absorption in the blue band (450–500 nm), together with persistently high turbidity in ARB lakes [44], further limited the accuracy of traditional retrieval methods. To overcome these constraints, we applied machine learning (ML) approaches including SLR (Stepwise Linear Regression), BP (Backpropagation Neural Network), and SVR (Support Vector Regression) for estimating CDOM in ARB lakes.

The machine learning framework was trained using the Landsat-8 OLI surface reflectance values in the blue (480 nm), green (560 nm), red (655 nm), and near-infrared (865 nm) bands as input features. The corresponding, coincidentally measured in situ $\alpha_{\text{CDOM}}(355)$ values served as the target output. The underlying principle is that the concentration of CDOM, which strongly absorbs in the ultraviolet–blue region, alters the spectral shape and magnitude of water-leaving reflectance across the visible spectrum. This approach allows for the estimation of $\alpha_{\text{CDOM}}(355)$ from satellite imagery despite the sensor's lack of spectral bands in the 355 nm region, provided a robust empirical relationship is established through representative ground-truth data.

SLR is an iterative variable selection method that dynamically adds or removes predictors according to statistical criteria (e.g., $p < 0.05$, Akaike Information Criterion (AIC) reduction). This approach maximizes model explanatory power while reducing the risk of overfitting. The integration of algorithmic filtering with domain-specific knowledge makes SLR a reliable tool for identifying dominant drivers in ecological datasets [18,55].

The BP is a multilayer feedforward neural network trained via error backpropagation [56]. It consists of fully connected input, hidden, and output layers that enable hierarchical feature extraction. Nonlinear activation functions, such as the sigmoid function, are applied in hidden layers, whereas linear functions are often used in the output layer. The BP model is trained by iteratively optimizing synaptic weights and neuronal thresholds.

SVR, proposed by Cortes and Vapnik [57] (1995), is a supervised kernel-based algorithm grounded in statistical learning theory and the principle of structural risk mini-

mization. By mapping input features into high-dimensional feature spaces through kernel functions (e.g., radial basis function, polynomial kernel), SVR constructs an optimal regression hyperplane that balances model simplicity with predictive accuracy. This shallow learning method is well suited to address challenges associated with limited sample sizes, nonlinear relationships, high-dimensional feature spaces, and overfitting. For the SVR implementation, we employed the radial basis function (RBF) kernel, which effectively maps input features into a higher-dimensional space to capture nonlinear relationships. The key hyperparameters—the regularization coefficient (C), the kernel coefficient (gamma, γ), and the epsilon-tube width (ϵ)—were optimized through a grid search with 5-fold cross-validation. The search space encompassed: C values of [0.1, 1, 10, 100, 1000], γ values of ['scale', 'auto', 0.001, 0.01, 0.1, 1], and ϵ values of [0.01, 0.1, 0.5]. The optimal hyperparameter combination minimized the root mean square error (RMSE) on the validation dataset.

For ML-based models, reflectance values from Landsat visible and near-infrared bands (centered at 480, 560, 655, and 865 nm) were used as input features. Reflectance normalization was applied to reduce interference from sediment scattering [58]. Model training and testing were conducted using the 194 paired Rrs spectra and CDOM observations. The dataset was randomly partitioned, with 70% of the samples (136) assigned to the training set and 30% (58 samples) reserved for validation. Moreover, both the samples employed for training and validation include all sampled lakes to preserve spatial representativeness. This retrospective application assumes that the empirical relationship between remote sensing reflectance and CDOM remains temporally stable, a common assumption in long-term water quality reconstruction using historical Landsat imagery [18,36]. Model performance was evaluated using three metrics: r RMSE (relative Root Mean Square Error), MAE (Mean Absolute Error), and R^2 (Coefficient of Determination). The flowchart for CDOM quantitative retrieval models is depicted in Figure S1.

Table 1. CDOM quantitative retrieval models (LR: Linear Regression, ER: Exponential Regression, Power: Power Regression, QAA-CDOM: Quasi-Analytical Algorithm for CDOM). The coefficients were adjusted using in situ and Landsat match-ups of this study. In this table, R_{blue} , R_{green} , R_{red} , and R_{nir} are the Landsat-8 surface reflectance values at 480, 560, 655, and 865 nm, respectively.

Statistical Technique	Band/Ratio	Adjusted Model	Reference	rRMSE (%)	R^2	MAE
LR	$x = R_{\text{red}}/R_{\text{green}}$	$y = -0.18x + 6.12$	[59]	37.6	0.23	6.54
LR	$x = (R_{\text{green}} + R_{\text{nir}})/(R_{\text{blue}}/R_{\text{nir}})$	$y = 0.28x + 4.76$	[51]	42.3	0.39	5.27
Power	$x = R_{\text{blue}}/R_{\text{red}}$	$y = 5.65x^{0.021}$	[52]	41.8	0.43	4.52
ER	$x = R_{\text{red}}$	$y = 5.32e^{2.14x}$	[60]	33.7	0.35	4.93
LR	$x = R_{\text{blue}}/R_{\text{green}}$	$y = 0.67\ln(x) + 6.03$	[50]	32.2	0.33	5.31
QAA-CDOM		/	[53]	30.2	0.59	4.37
BP	$x = R_{\text{red}}, R_{\text{green}}, R_{\text{nir}}, R_{\text{blue}}$	/	[4]	26.5	0.81	3.79
SLR		/	[61]	27.2	0.82	2.85
SVR		/	[18]	22.4	0.88	2.63

3.2. Statistical Analysis

The interannual variations in lake CDOM from 1984 to 2023 were evaluated using linear regression analysis against time (year). The slope of the regression line was used to indicate the long-term rate of change. To further assess the role of environmental drivers, correlation analysis was applied to quantify the relationships between CDOM and potential

explanatory variables. Statistical significance was evaluated at the 0.05 level, where $p < 0.05$ was considered significant and $p > 0.05$ indicated no significant relationship.

Attribution analysis of CDOM variability was conducted using multivariate General Linear Modeling (GLM), which is widely applied to disentangle the combined effects of multiple drivers [62,63]. This statistical framework enabled the quantitative partitioning of variance explained by two categories of drivers: hydro-climatic factors (precipitation, runoff, temperature, NDVI, and windspeed) and anthropogenic factors (grazing and irrigation). The model formulation is expressed as:

$$Y = \beta_0 + \sum_{i=1}^k \beta_i X_i + \varepsilon \quad (3)$$

where Y is the dependent variable, β_0 is a constant, X_i and β_i are the i -th independent variable and its regression coefficients; ε represents the random error term. The relative contribution of each independent variable is defined as the proportion of its variance in the total variance, as follows:

$$C_i = F_i / (\sum_{i=1}^k F_i + F_\varepsilon) \quad (4)$$

where C_i is the contribution of the i -th independent variable, F_i is the variance of the i -th independent variable, and F_ε is the variance of the random error term.

3.3. Accuracy Assessment

Model accuracy was evaluated by comparing in situ CDOM absorption coefficients with estimated values using three statistical indicators: the coefficient of determination (R^2), mean absolute error (MAE), and relative root mean square error ($rRMSE$). These metrics are defined as:

$$MAE = \frac{\sum |X_i - Y_i|}{n} \quad (5)$$

$$RMSE = \sqrt{\frac{\sum (X_i - Y_i)^2}{n}} \quad (6)$$

$$rRMSE = \frac{RMSE}{\bar{X}} \times 100\% \quad (7)$$

where X_i and Y_i are the in situ and model-predicted CDOM absorption coefficients for the i -th sample, respectively, n is the total sample number, and $\bar{X}$ is the mean of in situ CDOM absorption coefficients.

4. Results

4.1. Performance of CDOM Estimation Models

Table 1 shows that the overall predictive performance of traditional models was limited. Across all models, R^2 values were below 0.5, $rRMSE$ exceeded 30%, and MAE was greater than 4.7 m^{-1} . These results indicate that these models were not able to reliably capture spatiotemporal dynamics of CDOM in ARB lakes. The quasi-analytical algorithm (QAA-CDOM) demonstrated slightly better performance compared with other traditional methods, but its accuracy remained insufficient for large-scale and long-term remote sensing applications for CDOM monitoring ($R^2 < 0.6$).

In contrast, machine learning methods achieved substantially higher predictive when applied to Landsat data (Table 1). The performance of these models improved markedly, with R^2 values greater than 0.8, $rRMSE$ values below 30%, and MAE below 4 m^{-1} (Figure 2). These results highlight that ML models effectively capture key spectral features associated with CDOM variability, thereby enabling accurate quantitative retrieval across both spatial and temporal dimensions.

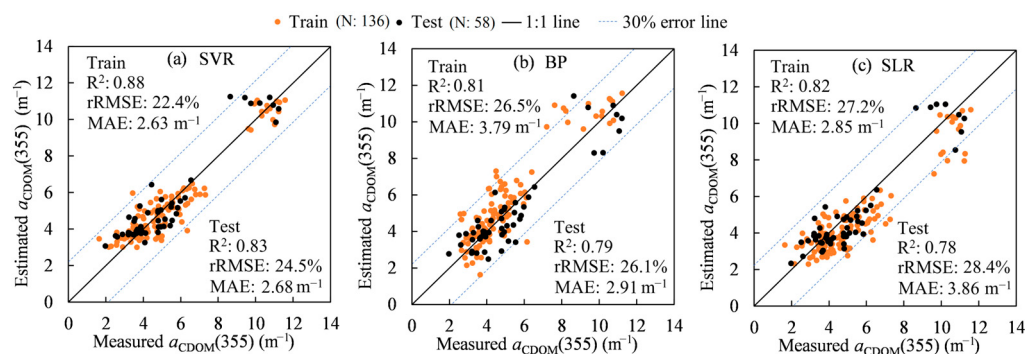


Figure 2. Comparison of field-measured CDOM absorption coefficient $\alpha_{\text{CDOM}}(355)$ with the estimated results based on the SVR model (a), BP (b), and SLR (c).

Among the three machine learning approaches, SVR achieved the best overall performance. This model reached the highest R^2 of 0.88, the lowest $r\text{RMSE}$ of 22.4%, and the lowest MAE of 2.63 m^{-1} (Figure 2). These results suggest that SVR possesses strong generalization capability under conditions of limited sample size and moderate data quality, which are common challenges in large-scale limnological remote sensing [18]. The SVR model demonstrated consistently strong performance across different CDOM ranges, including <5.49 and $>5.49 \text{ m}^{-1}$ (Table S3). A perturbation sensitivity analysis using Hulun Lake showed that the 40-year CDOM trend remained unchanged across all scenarios, indicating strong robustness to calibration uncertainty (Figure S3).

To ensure the consistency of the results, we conducted comparative experiments across different satellite sensors. The estimated values of $\alpha_{\text{CDOM}}(355)$ derived from TM, OLI and ETM+ during overlapping observation periods (within seven days) showed strong agreement, with a determination coefficient (R^2) greater than 0.9 (Figure 3). Specifically, the correlation between $\alpha_{\text{CDOM}}(355)$ from OLI and ETM+ reached 0.93 ($p < 0.01$, $N = 1025$, Figure 3), while the correlation between $\alpha_{\text{CDOM}}(355)$ from TM and ETM+ was 0.91 ($p < 0.01$, $N = 1225$). These results confirm that OLI-based estimates of $\alpha_{\text{CDOM}}(355)$ can serve as a reference for calibrating TM- and ETM+-based values. The calibration formula is expressed as follows:

$$\alpha_{\text{CDOM}}(355)_{\text{OLI}} = 0.93 \times \alpha_{\text{CDOM}}(355)_{\text{TM}} + 0.35 \quad (8)$$

$$\alpha_{\text{CDOM}}(355)_{\text{OLI}} = 0.99 \times \alpha_{\text{CDOM}}(355)_{\text{ETM+}} + 0.01 \quad (9)$$

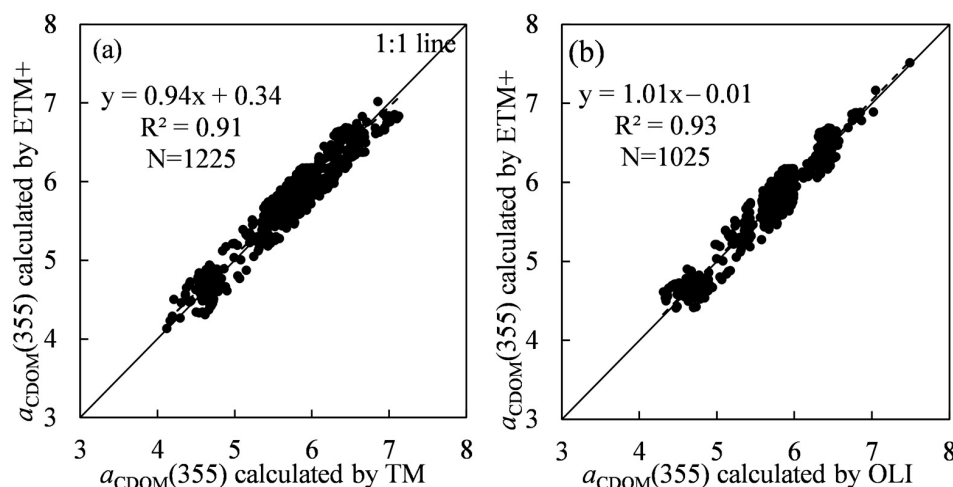


Figure 3. Cross-sensor consistency of CDOM absorption at 355 nm ($\alpha_{\text{CDOM}}(355)$) estimated from Landsat imagery. (a) Comparison between Landsat-7 ETM+ and Landsat-5 TM. (b) Comparison between Landsat-7 ETM+ and Landsat-8 OLI.

4.2. Spatial-Temporal Variation of CDOM in the ARB

The SVR-based CDOM retrieval model was applied to Landsat imagery to construct a long-term CDOM dataset for ARB lakes covering the period 1984–2023. Given the ARB's broad geographic extent and heterogeneous environmental gradients shaped by both natural and anthropogenic processes (as detailed in Section 2.1), a sub-regional framework was necessary to capture spatially differentiated CDOM dynamics. Accordingly, the ARB was classified into three major sub-regions (Figure 1b): the Northeast Plain (45 lakes), Mongolia Plateau (8 lakes), and Far East Federal District (16 lakes). This stratification enabled targeted assessment of lake CDOM variations across distinct geomorphological and socio-ecological zones. Therefore, Figure 4 presents the long-term mean $\alpha_{\text{CDOM}(355)}$ for both the whole ARB and each sub-region.

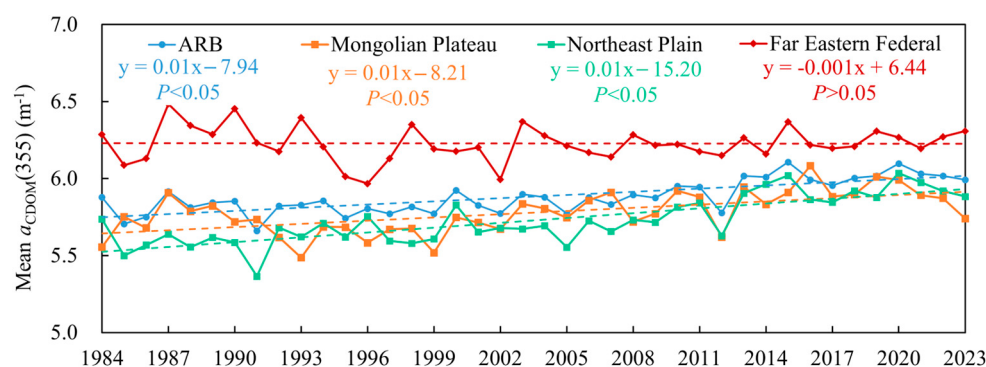


Figure 4. Long-term trend of mean $\alpha_{\text{CDOM}(355)}$ values in the ARB and three sub-regions from 1984 to 2023. The fitted lines show the interannual trends.

Results showed that the average $\alpha_{\text{CDOM}(355)}$ across ARB lakes increased significantly during the past four decades, with a trend of 0.01 m^{-1} per year ($p < 0.05$) (95% CI: 0.008 to 0.012). The basin-wide mean $\alpha_{\text{CDOM}(355)}$ was 5.92 m^{-1} , which is consistent with values reported by Zhao [52]. Annual means ranged from 5.67 m^{-1} in 1992 to 6.11 m^{-1} in 2015. Sub-regional analysis further revealed divergent trends that lakes in the Mongolia Plateau (95% CI: 0.006 to 0.013) and the Northeast Plain (95% CI: 0.008 to 0.013) exhibited significant increases in CDOM, with rates of approximately 0.1 m^{-1} per decade. In contrast, lakes in the Far East Federal District displayed a declining tendency (95% CI: -0.003 to 0.003), with CDOM levels remaining relatively stable between 6.0 – 6.8 m^{-1} (Figure 4).

The annual mean CDOM absorption coefficient of lakes were classified into five ranges: 4.5 – 5.0 m^{-1} , 5.0 – 5.5 m^{-1} , 5.5 – 6.0 m^{-1} , 6.0 – 6.5 m^{-1} , and 6.5 – 7.0 m^{-1} . These categories are represented in Figure 5, where circle size denotes the corresponding CDOM class. The number of lakes in each class is summarized in Figure 6a. Specifically, 7 lakes (10%) fall within the 4.5 – 5.0 m^{-1} range, 12 lakes (17%) within 5.0 – 5.5 m^{-1} , 17 lakes (25%) within 5.5 – 6.0 m^{-1} , 18 lakes (26%) within the 6.0 – 6.5 m^{-1} range, and 15 lakes (22%) within the 6.5 – 7.0 m^{-1} range. These results show that most lakes have an annual average CDOM above 5.5 m^{-1} , a condition that may constrain the growth of submerged vegetation. Lakes in the Far East region exhibit a highest average CDOM at 6.23 m^{-1} , exceeding those in the other two sub-regions. Interannual CDOM trends were further classified into four categories: increase, significant increase, decrease, and significant decrease, as shown with color coding in Figure 5. The distribution of lakes across these categories is presented in Figure 6b. Overall, 42 lakes (61%) showed an increasing trend in annual CDOM, among which 21 lakes experienced a significant increase and the majority of these lakes are located in the Northeast Plain (16 lakes). This pattern may be linked to agricultural expansion and permafrost degradation in this region [17]. In contrast, 11 lakes in the Far Eastern Federal

showed no significant decreasing trend, contributing to the overall decline of CDOM in that sub-region.

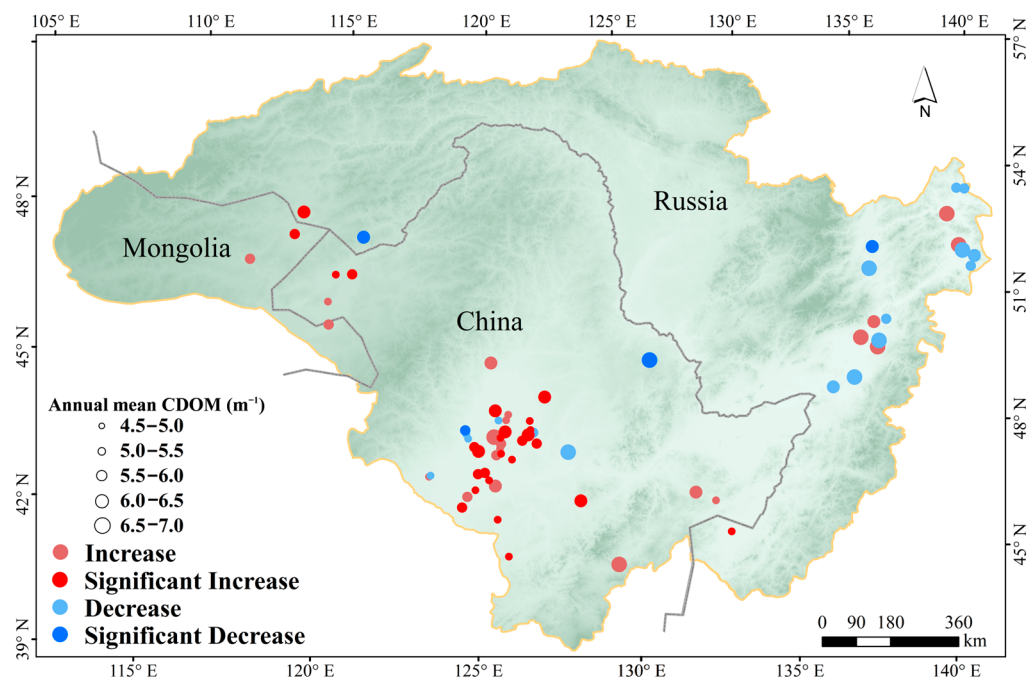


Figure 5. Locations of 69 lakes distributed across three sub-regions of the ARB, showing CDOM levels and long-term trends. CDOM levels are categorized into five classes ($4.5\text{--}5.0\text{ m}^{-1}$, $5.0\text{--}5.5\text{ m}^{-1}$, $5.5\text{--}6.0\text{ m}^{-1}$, $6.0\text{--}6.5\text{ m}^{-1}$, and $6.5\text{--}7.0\text{ m}^{-1}$) and are represented by different circle sizes. Changing trends are classified into four categories (Increase, Significant Increase, Decrease, and Significant Decrease) and are indicated by color.

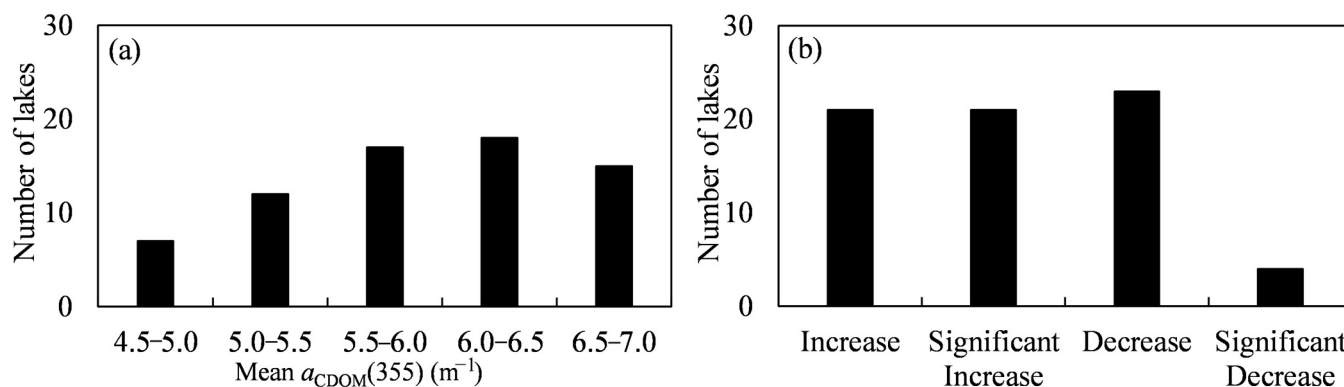


Figure 6. Distribution of lakes across CDOM categories and long-term trends. (a) Histogram showing the number of lakes within each CDOM level class. (b) Histogram showing the number of lakes in each trend category.

4.3. Spatial-Temporal Variation of CDOM in Three Representative Lakes

To further investigate the spatiotemporal variability of CDOM within the ARB, three representative lakes (Hulun Lake, Sikeye Lake, and Yuebing Lake) were selected following a comprehensive basin-wide analysis. Each of these lakes represents one of the three sub-regions. By analyzing the long-term trends and spatial distribution of CDOM, we provide insight into lake-level dynamics of CDOM.

Figure 7 shows the spatial patterns and temporal variations of CDOM in the three lakes from 1984 to 2023, divided into eight time periods and visualized with box plots and line graphs. In Hulun Lake, CDOM exhibited a significant increasing trend ($y = 0.03x + 5.42$, $p < 0.05$, Figure 7b), with most areas of the lake showing an upward

tendency (Figure 7a). High-value zones were concentrated along the southern and eastern shores (Figure 7c), likely reflecting terrestrial inputs from inflowing rivers [64]. In contrast, Sikeye Lake displayed a significant decline in CDOM ($y = -0.1x + 6.63$, $p < 0.05$, Figure 7e), which was consistent with its spatial variation (Figure 7d). This reduction may be linked to the slowing of wetland degradation and land use changes such as the conversion of farmland to forested areas [10]. High CDOM values in Sikeye Lake were mainly located along the shoreline (Figure 7f), possibly due to sediment resuspension [65]. Yuebing Lake showed a contrasting pattern, with CDOM increasing significantly across the entire lake region from 1984 to 2023 ($y = 0.13x + 5.39$, $p < 0.05$, Figure 7g,i). This trend is likely associated with intensive human activities in the region [39].

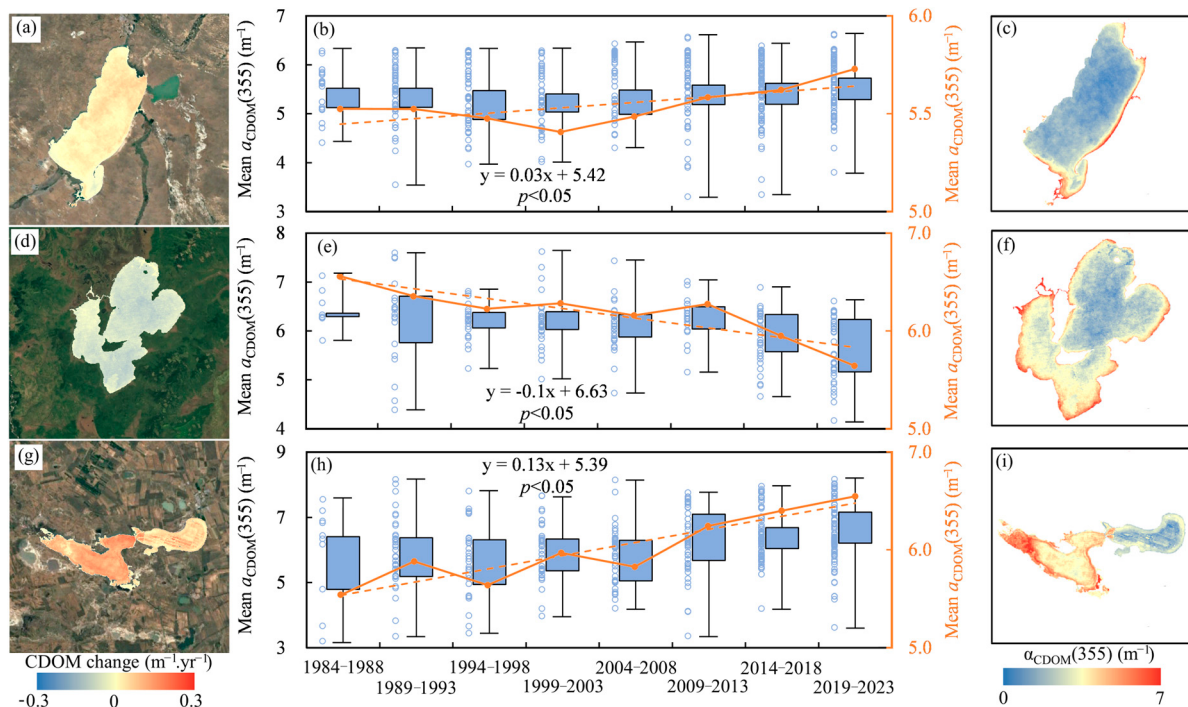


Figure 7. Spatial and temporal variations of CDOM in three representative lakes of the ARB from 1984 to 2023. (a–c) Hulun Lake in the Mongolian Plateau, (d–f) Sikeye Lake in the Far East, and (g–i) Yuebing Lake in the Northeast Plain. Box plots show mean $\alpha_{\text{CDOM}}(355)$ values for eight periods (I: 1984–1988, II: 1989–1993, III: 1994–1998, IV: 1999–2003, V: 2004–2008, VI: 2009–2013, VII: 2014–2018, VIII: 2019–2023). The orange line represents the changing trends of mean $\alpha_{\text{CDOM}}(355)$ (5-year) values in eight time periods.

4.4. Potential Driving Forces of CDOM Changes

Lake ecosystems are shaped by complex interactions between natural processes and anthropogenic activities, with climate variability and anthropogenic pressures playing increasingly important roles. Rapid urbanization and agricultural intensification have contributed to the degradation of aquatic environments, often through nutrient enrichment and pollutant accumulation. However, scientific understanding of the combined effects of climatic fluctuations, watershed characteristics, and socio-economic pressures on CDOM dynamics in inland waters remains limited. To explore these relationships, we conducted correlation analysis between the annual mean CDOM of each lake and potential explanatory variables, namely, precipitation, temperature, runoff, wind speed, NDVI, irrigation area, and grazing, covering the period from 1984 to 2023 (Table S1). The results showed that in 69 lakes, more than 80% exhibited a significant correlation with at least one factor, and the direction of correlation varied among lakes, with both positive and negative relationships observed for all seven factors. This highlights the heterogeneity in CDOM responses

across different lakes. Specifically, precipitation was significantly correlated with CDOM in 15 lakes, temperature in 10 lakes, runoff in 8 lakes, wind speed in 13 lakes, NDVI in 9 lakes, irrigated in 26 lakes, and grazing in 16 lakes. These results indicate that hydro-climatic and anthropogenic variables provide substantial explanatory power for CDOM variability across the watershed, although their impacts differ among individual lakes.

To further assess the role of lake morphology, we examined whether the effects of driving factors varied with lake size and depth. Lakes were grouped into two categories by size ($>100 \text{ km}^2$ and $<100 \text{ km}^2$) and by depth ($>3 \text{ m}$ and $<3 \text{ m}$). The proportion of lakes in each category that exhibited a significant correlation with CDOM, as indicated by an asterisk in Table S1, was then calculated and summarized in Figure 8.

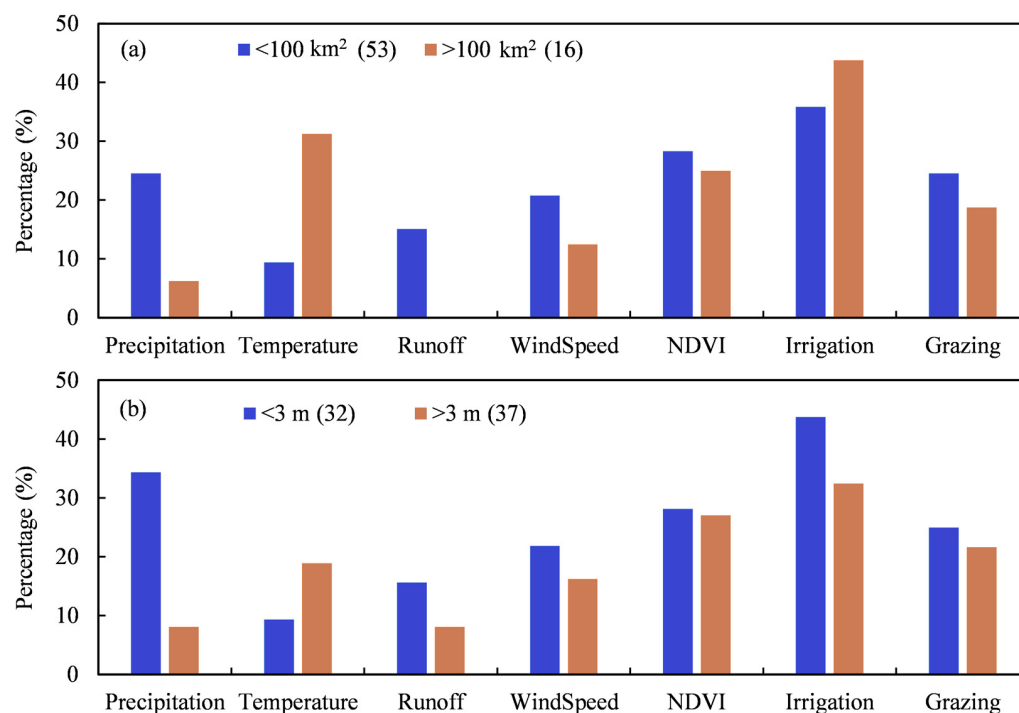


Figure 8. Proportion of lakes showing significant correlations between CDOM and environmental drivers. Results are presented according to (a) lake area and (b) lake depth. Percentages indicate the fraction of lakes in each group with significant correlations for precipitation, temperature, runoff, wind speed, NDVI, irrigation, and grazing.

Figure 8a shows that lake size had a minor influence on the correlation between CDOM and most examined factors, including wind speed, NDVI, irrigation, and grazing. Both small lakes ($<100 \text{ km}^2$, $N = 53$) and large lakes ($>100 \text{ km}^2$, $N = 16$) exhibited similar proportions of significant correlations, with differences generally within 10%. Specifically, the proportions were as follows: wind speed (20.75% for small lakes vs. 12.50% for large lakes), NDVI (28.30% vs. 25.00%), irrigation (35.85% vs. 43.75%), and grazing (24.53% vs. 18.75%). In contrast, lake size had a significant influence on the correlation between CDOM and precipitation, temperature, and runoff. The proportion of small lakes correlated with precipitation was 24.53%, much higher than that of large lakes (6.25%). Similarly, 15.09% of small lakes were significantly correlated with runoff, whereas no significant correlation was found for large lakes. Conversely, the proportion of lakes correlated with temperature was much lower for small lakes (9.43%) compared with large lakes (31.25%).

Figure 8b shows that the influence of lake depth on the proportion of significant correlations between CDOM and its driving factors. Overall, shallow lakes ($<3 \text{ m}$, $N = 32$) are more sensitive to external forcing, with six out of seven driving factors showing

higher proportions of significant correlations. Specifically, the significant correlations of shallow lakes with precipitation (34.38%) and irrigation (43.75%) are markedly higher than those of deep lakes (8.11% and 32.43%, respectively), indicating that shallow lakes are more susceptible to external inputs and hydrological fluctuations. Temperature is the only exception, with deep lakes exhibiting a higher proportion of significant correlations (18.92%) compared to shallow lakes (9.38%). These results suggest that shallow lakes, due to their smaller volume and limited buffering capacity, are more directly affected by climatic and anthropogenic disturbances [66].

The stronger correlation between CDOM and temperature in deeper lakes (Figure 8b) likely reflects the influence of thermal stratification. In deep lakes, summer stratification creates a cold, isolated bottom layer (hypolimnion) with low oxygen levels [67]. This anoxic environment inhibits the microbial degradation of organic matter, potentially leading to the accumulation and preservation of CDOM in the water column. In contrast, shallow lakes remain well-mixed, exposing CDOM uniformly to light and oxygen at the surface, which promotes photochemical and microbial breakdown, and tightens its coupling to immediate external inputs like runoff [68]. Thus, depth mediates how a lake responds to climatic warming.

Considering the diverse hydro-climatic conditions across sub-regions (Section 2.1), we further analyzed the correlations between CDOM and potential drivers at both the basin and sub-regional scales from 1984 to 2023 (Figure 9). The analysis revealed distinct spatial patterns in CDOM drivers across the ARB. At the basin scale, CDOM was most strongly correlated with anthropogenic factors, including grazing ($R = 0.60$, $p < 0.05$) and irrigation ($R = 0.76$, $p < 0.05$). In agricultural regions (e.g., the Northeast Plain), irrigation return flows act as conduits, transporting fertilizers, pesticides, and dissolved organic matter leached from soils into lakes [17]. In pastoral areas (e.g., the Mongolian Plateau), grazing intensity increases the loading of manure-derived nutrients and organic matter, either through direct deposition or surface runoff. These inputs alter the composition and quantity of allochthonous (externally sourced) CDOM, explaining its strong statistical link to human drivers. Hydro-climatic influences were secondary, with temperature ($R = 0.53$, $p < 0.05$) and wind speed ($R = 0.51$, $p < 0.05$) showing moderate effects. Marked differences emerged among the three sub-regions. On the Mongolian Plateau, both grazing intensity ($R = 0.78$, $p < 0.05$) and wind speed ($R = 0.45$, $p < 0.05$) exerted joint control over CDOM variability. Livestock excrement transported by surface runoff elevated CDOM [17], while wind-driven turbulence promoted resuspension of CDOM-rich sediments into the epilimnion [65]. In the Far Eastern Federal District, CDOM dynamics were correlated only with wind speed ($R = 0.51$, $p < 0.05$), with no significant relationships to anthropogenic drivers. This pattern highlights the minimal human interference in the region and emphasizes wind-driven sediment resuspension as the dominant CDOM source. In the Northeast Plain, irrigation intensity emerges as the strongest driver of CDOM variability ($R = 0.77$, $p < 0.05$), followed by temperature ($R = 0.58$, $p < 0.05$). Agricultural drainage transports terrestrial organic matter (e.g., fertilizer residues and crop debris) into aquatic systems, significantly enhancing allochthonous CDOM inputs [69]. Furthermore, the observed correlation between temperature and CDOM is consistent with established mechanisms whereby climate warming may enhance autochthonous CDOM production [70]. Specifically, warming can stimulate phytoplankton growth and bacterial degradation of organic substrates, especially under prolonged ice-free conditions and elevated microbial activity [71,72].

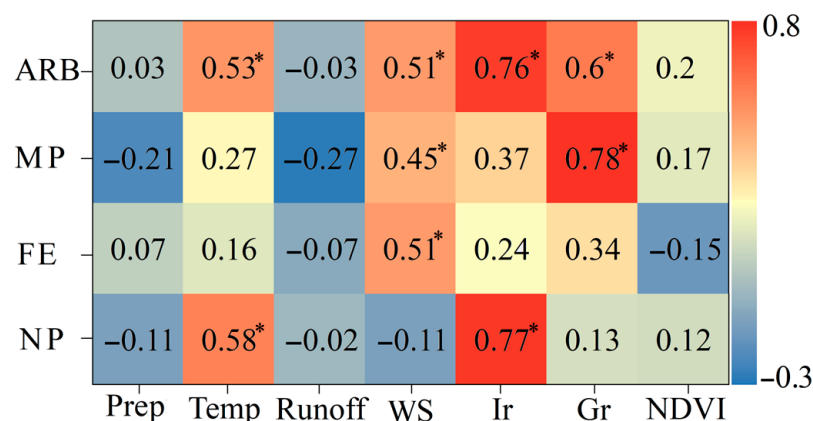


Figure 9. Correlations between lake CDOM and environmental drivers in the ARB and its three sub-regions: the Mongolian Plateau (MP), the Far East (FE), and the Northeast Plain (NP), from 1984 to 2023. Statistically significant correlations ($p < 0.05$) are marked with asterisks.

The relative contributions of hydro-climatic and anthropogenic factors to the inter-annual variability of CDOM were assessed using multiple Generalized Linear Models (GLMs), with the results shown in Figure 10. The cumulative contribution of the seven parameters analyzed ranged from 66.7% to 98.4%, with 48 lakes exhibiting contributions exceeding 85%, thereby indicating that these factors predominantly influence the inter-annual variability of CDOM. Conversely, the residuals from the multiple GLM analysis for the studied lakes varied between 1.6% and 33.3%, suggesting the potential influence of unidentified factors on the clarity of water bodies in certain lakes. Statistically, irrigation emerged as a significant contributor, accounting for 17% of the variability, while the contributions of the other six factors were as follows: grazing (13.7%), precipitation (11.1%), NDVI (11.2%), runoff (10.4%), wind speed (11.4%), and temperature (12.8%). Notably, the interannual fluctuations in CDOM for 45 lakes were significantly correlated with one or more of the seven explanatory variables ($p < 0.05$). However, the interannual changes in CDOM for 24 lakes could not be adequately explained by these factors. This lack of explanation may be attributed to two primary reasons: (1) the alterations in CDOM may result from the interplay of multiple factors, and (2) other unidentified driving factors, represented by the residuals in Figure 10, may exert a more substantial influence on the CDOM of these lakes. In Section 5.2, we explore the impact of several additional potential influencing factors on the variations in CDOM within water bodies.

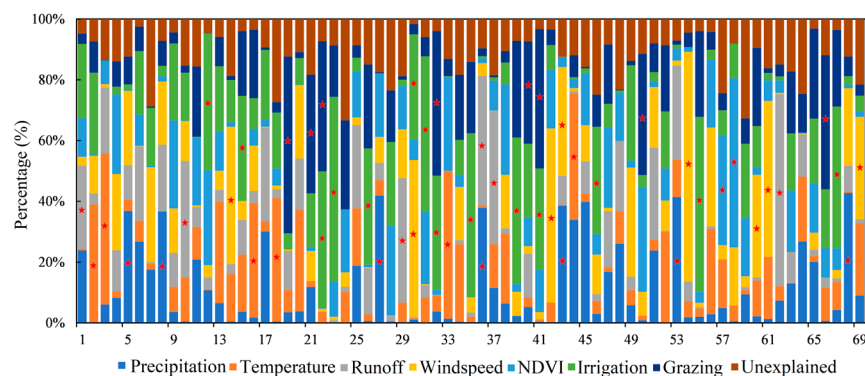


Figure 10. The relative contributions of seven explanatory factors to lake CDOM in the ARB. Statistically significant contributions are annotated with a red star.

5. Discussion

5.1. Uncertainties of the Estimated CDOM

The accurate retrieval of CDOM in inland waters is challenging because its optical signals strongly overlap with those of other optically active substances such as chlorophyll a (Chl-a) and total suspended matter (TSM) [1]. The absorption peaks of Chl-a at 440 nm and 675 nm and the backscattering of TSM between 550 and 750 nm interfere with pronounced absorption properties of CDOM in the ultraviolet–visible spectrum (250–550 nm) [73,74]. Furthermore, the strong absorption of CDOM in the blue band [48], together with uncertainties in atmospheric correction within this wavelength range [75,76], makes the development of robust inversion methods particularly challenging. Hence, we adopted a direct modeling approach that relates synchronous satellite-derived remote sensing reflectance (R_{rs}) to in situ CDOM measurements. This method avoids the accumulation of systematic errors from atmospheric correction and implicitly compensates for related uncertainties, thereby improving the reliability of CDOM retrieval [77].

The comparison of different models (Table 1) showed that SVR applied to Landsat 8 OLI data outperformed traditional empirical models, with RMSE reduced by 9.8–19.9% and R^2 increased by 0.45–0.65. This result indicates the feasibility of extending the approach to Landsat series data for long-term CDOM monitoring. By systematically collecting CDOM measurements across diverse water body types (e.g., elevated, deep, and shallow lakes), this study further enhanced the generalization ability and regional adaptability of the SVR model ($rRMSE < 25\%$, $R^2 > 0.85$). However, its performance remains constrained by the representativeness of the training dataset, and inversion accuracy may decline in regions with hydro-optical conditions that differ substantially from those in the training set [78]. These findings highlight both the promise and limitations of SVR algorithms: while they provide a robust framework for CDOM remote sensing estimation, their effectiveness ultimately depends on the quality and diversity of the training data, which remains a key source of uncertainty. Therefore, uncertainties in CDOM retrieval arise not only from intrinsic optical interferences and atmospheric correction errors, but also from the data- and model-dependent nature of statistical approaches, underscoring the importance of addressing both observational and methodological limitations.

An additional and important source of uncertainty concerns the temporal transferability of the model. The SVR was trained using 194 in situ–satellite matchups collected during 2020–2023, representing a limited temporal window. Applying this model retrospectively to Landsat data from 1984–2019 implicitly assumes stability in lake hydro-optical properties, CDOM composition, and sensor characteristics over time. While this assumption is common in long-term water quality reconstructions [18,39], it may not fully account for multi-decadal environmental changes such as shifting water constituents, land-use transitions, or climate-driven alterations in lake metabolism. Differences among Landsat sensors (TM, ETM+, and OLI) further add to potential uncertainties [52]. Therefore, although the reconstructed time series is internally consistent and suitable for trend analysis, absolute CDOM values may vary across decades.

Sensor characteristics also constrain CDOM estimation. Landsat imagery, with a spatial resolution of 30 m, has shown considerable potential for the monitoring of bio-optical parameters in inland waters [52]. The visible band configuration of the Landsat-8 OLI (443, 482, 561, and 655 nm) provides a useful basis for CDOM retrieval [6,79], but the lack of shorter UV bands (<400 nm) limits the ability to fully characterize CDOM absorption features [24]. Furthermore, the 16-day revisit cycle of Landsat restricts the ability to capture short-term CDOM fluctuations caused by events such as heavy precipitation

and typhoons [80,81]. Compared with Landsat, sensors such as Sentinel-2 MSI (13 bands, 5-day revisit) and Gaofen-6 (8 bands, 4-day revisit) offer clear advantages in terms of temporal and spectral resolution [82]. Future research should therefore focus on integrating multi-source satellite data to produce high-resolution dynamic maps of CDOM in the ARB lakes and other global inland waters, enabling improved assessments of both seasonal variability and long-term CDOM dynamics in inland waters.

5.2. Additional Potential Factors in Lake CDOM

This study constructed a multivariate driving framework to explain the changes in lake CDOM in the ARB and its sub-regions by selecting five hydro-climatic factors, including Precipitation, wind speed, temperature, runoff, and NDVI, as well as two anthropogenic factors, including grazing intensity and irrigation area (Section 4.4, Figure 10). The study showed that these variables can explain 66.7% to 98.4% of the changes in lake CDOM in the basin, but it is worth noting that some potential driving mechanisms have not yet been included in the model system. Also, It should be noted that our driver analysis using GLM is exploratory in nature. While it successfully identifies the broad categories of factors influencing CDOM, future studies would benefit from more sophisticated approaches, such as structural equation modeling or machine learning-based feature selection, to better handle predictor collinearity and elucidate causal pathways. For example, permafrost degradation, Gross Domestic Product (GDP), and Gross Industrial Product (GIP) (regional industrial water use intensity represented by GIP and GDP) [83,84]. Specifically, permafrost will release a large amount of dissolved organic carbon during the melting process, thereby increasing the exogenous input of lake CDOM [85]. At the same time, industrial wastewater discharge is considered to be one of the important exogenous inputs of lake organic matter [86]. However, due to limitations in data availability, this study could not quantitatively assess the contributions of these factors to lake CDOM changes. Therefore, future research should expand observational datasets and explicitly incorporate key physical processes (e.g., permafrost degradation and industrial water use) into CDOM analysis.

In addition, changes in lake area exert a critical influence on CDOM. Hydrodynamic-sediment interactions in shallow lakes (less than 5 m deep) promote sediment resuspension [82], thereby increasing the flux of self-generated organic matter [87]. To examine this relationship, we analyzed trends in CDOM and surface from 1999 to 2021 for three representative lakes: Hulun Lake (>100 km²), Jingpo Lake (50–100 km²), and Wulacao Lake (20–50 km²).

Analysis of Hulun Lake shows that the significant increase in CDOM is independent of fluctuations in lake area ($r = 0.15$, $p > 0.05$). This is mainly because CDOM in this lake is dominated by terrestrial inputs [64], so reductions in lake area do not significantly alter its concentration (Figure 11a). A similar pattern was observed in Jingpo Lake, where CDOM remained stable despite area variations ($r = 0.12$, $p > 0.05$) (Figure 11b). The 40 m depth of Jingpo Lake effectively restrained sediment resuspension, thereby decoupling CDOM dynamics from lake area [88]. In contrast, Wulacao Lake (a shallow lake) exhibited synchronous oscillations between CDOM and lake area ($r = -0.45$, $p < 0.05$) (Figure 11c). In contrast, Wulacao Lake exhibited synchronous oscillations ($r = -0.45$, $p < 0.05$) (Figure 11c). In this case, reductions in water level enhanced sediment resuspension, leading to increased CDOM [89]. Collectively, these findings indicate that depth exerts a buffering effect, weakening the influence of area fluctuations on CDOM in deep lakes, whereas in shallow lakes (<5 m), reduced water depth amplifies the coupling between lake area and CDOM dynamics.

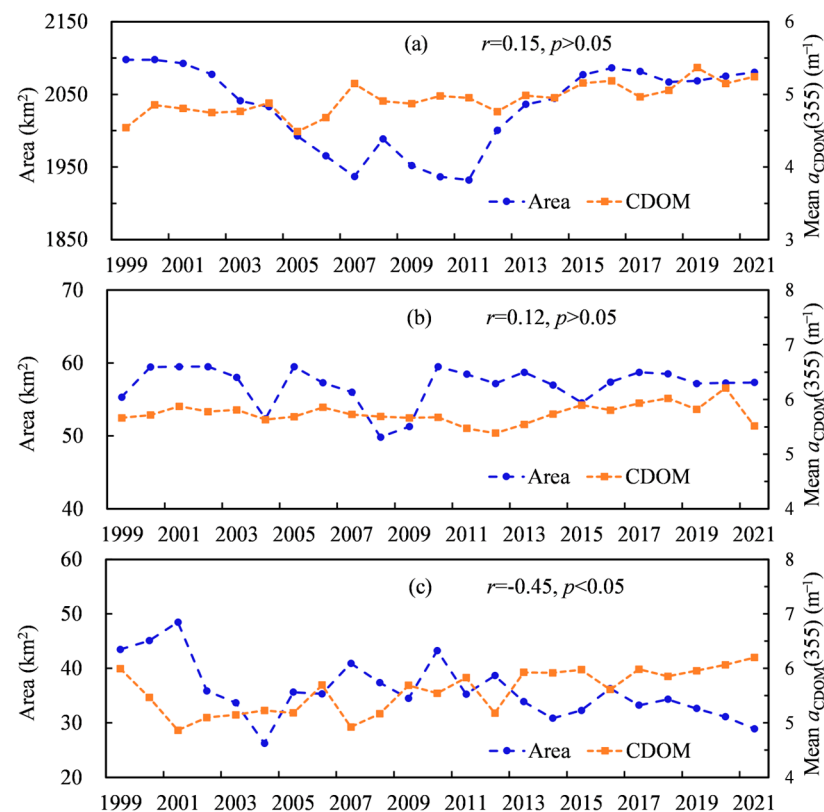


Figure 11. Trends of lake area and mean CDOM for three lakes: (a) Hulun Lake, (b) Jingpo Lake, and (c) Wulacao Lake.

5.3. Drivers of the Observed Spatial and Temporal Heterogeneity in CDOM

The pronounced spatial and temporal heterogeneity of CDOM across the Amur River Basin reflects the combined influence of regional climate regimes, landscape characteristics, and human activities. Long-term climatic trends provide an important context for understanding CDOM variability. As shown in Figure S2, the Mongolian Plateau has undergone significant warming and declining precipitation since 1984, indicating intensified aridification. These drying conditions likely elevate CDOM concentrations by shrinking lake surface areas, concentrating dissolved organic matter, and promoting the mobilization of organic-rich littoral soils during episodic rainfall events [90]. In contrast, the Far East Federal District exhibits increasing precipitation and decreasing wind speeds, which favor CDOM dilution and reduce resuspension-driven inputs [91]. These opposing climatic patterns—drying in the west and moistening in the east—suggest that basin-scale atmospheric processes, including East Asian monsoon influences, shape the temporal trajectories of CDOM across the ARB.

Beyond temporal variability, the baseline spatial pattern of multi-year mean CDOM (Figure 5) reflects fundamental differences in terrestrial carbon sources among sub-regions. Lakes in the Far Eastern Federal District and northern ARB exhibit consistently high CDOM levels due to extensive wetlands and discontinuous permafrost, which release large quantities of labile dissolved organic matter upon thaw [10,17]. In contrast, lakes on the Mongolian Plateau generally show moderate baseline CDOM, owing to semi-arid grassland landscapes with lower soil organic carbon and episodic hydrological connectivity. Meanwhile, lakes in the Northeast China Plain display lower baseline CDOM due to historical agricultural conversion that reduced natural organic carbon storage, although recent increases indicate growing anthropogenic contributions.

Overall, both the baseline spatial template and long-term temporal changes in CDOM are jointly shaped by region-specific interactions among climate, hydrology, soil carbon pools, and land use. This integrated perspective explains why different sub-regions respond to distinct combinations of climatic and anthropogenic drivers, as further corroborated by the attribution analysis (Figures 9 and 10).

5.4. Implications for Safeguarding Lake Ecology

This study shows that there is significant spatial heterogeneity in the driving mechanism of changes in CDOM in lakes in different regions (Sections 4.2–4.4), and it is necessary to implement zoning and classification management strategies based on differences in dominant factors. On the Mongolian Plateau (MP), the synergistic effect of grazing activities and wind erosion processes dominates CDOM input (total contribution 58.97%). It is recommended to reduce soil erosion and organic matter loss through optimization of rotational grazing systems, dynamic regulation of stock carrying capacity, and grassland vegetation restoration, thereby reducing terrestrial CDOM input. In view of the human disturbance centered on irrigation drive in the NP (contribution rate of 50.19%), it is urgent to promote water-saving irrigation technology (such as drip irrigation and intelligent water–fertilizer integration) to reduce the input of dissolved organic matter (DOM) in agricultural drainage, and to construct ecological buffer zones (such as riverbank wetlands and ditch plant barriers) to intercept CDOM migrating to lakes. In addition, the implementation of wind erosion control projects (such as windbreaks) can effectively slow down the fluctuation of CDOM caused by the resuspension of suspended matter.

Given that lakes in the ARB basin are affected by the combined stress of climate change and human activities [17], remote sensing images with better temporal, spatial, and spectral resolution need to be used for frequent and continuous monitoring of lake CDOM, such as Sentinel-2A and OLCI [18]. At the same time, cross-border basins (such as the lakes at the border between China, Russia, and Mongolia) need to coordinate water environment governance and pollution prevention and control standards through international cooperation frameworks, and build a CDOM source analysis and collaborative governance system across administrative boundaries [10]. The framework proposed in this study can provide a reference for the CDOM management of global lakes, especially for the sustainable management of intensive agricultural areas, ecologically fragile zones, and cross-border waters.

6. Conclusions

This study developed a CDOM inversion model tailored for lakes in the Amur River Basin (ARB) using the support vector regression (SVR) machine learning algorithm. Based on this model, a 40-year CDOM database was constructed for 69 large lakes in the ARB using Landsat-5, Landsat-7, and Landsat-8 imagery from 1984 to 2023. The results revealed an overall increasing trend of CDOM across the ARB lakes, with most increases occurring in the Mongolian Plateau (MP) and the Northeast Plain (NP). In total, 27 lakes showed significant increases in CDOM, whereas four lakes exhibited significant decreases. The analysis further demonstrated that the dominant drivers of CDOM vary regionally, with hydro-climatic factors such as wind speed and temperature and anthropogenic pressures including irrigation and grazing playing key roles. These findings improve the understanding of the spatiotemporal dynamics of CDOM in ARB lakes and provide an important reference for long-term monitoring and conservation of inland water quality.

This study provides the first comprehensive, satellite-based perspective on the long-term dynamics of CDOM across the ARB. The identified spatiotemporal patterns and their attributed drivers offer clear pathways for practical application. Specifically, the

regional dominance of particular drivers can guide targeted interventions within national lake protection programs, such as China's Lake Chief System. This enables managers to prioritize measures like agricultural best practices in irrigation-dominated basins or climate adaptation strategies in temperature-sensitive lakes. Moreover, CDOM can be proposed as a supplementary water quality indicator within existing trilateral dialogue frameworks, serving as a proxy for terrestrial organic matter inputs that complement conventional parameters. Finally, the established SVR model, combined with near-real-time satellite data, lays the groundwork for an operational early-warning system to detect sudden CDOM increases following extreme events. In summary, this work not only fills a critical knowledge gap but also delivers a scalable, evidence-based framework to support integrated transboundary management and ecological conservation of the ARB's freshwater ecosystems.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs18010125/s1>. Figure S1. Flow chart for CDOM quantitative retrieval. Figure S2. Trends in wind speed (a), precipitation (b), and temperature (c) across the Amur River Basin from 1984 to 2023. Figure S3. Sensitivity of Hulun Lake CDOM trends (1984–2023) to calibration uncertainty. (a) Baseline CDOM time series; (b) CDOM series under the -1 RMSE perturbation scenario; (c) CDOM series under the $+1$ RMSE perturbation scenario. Table S1: The fundamental information of the lakes under investigation, along with the correlation coefficient (r) between the annual mean CDOM and seven explanatory variables throughout the study period. Table S2: Statistical information on CDOM in sampled lakes. Table S3. Comparison of the performance of SVR at different CDOM levels.

Author Contributions: All the authors participated in editing and reviewing this manuscript. Conceptualization, Y.W. and P.H.; methodology, Y.W., P.H. and C.Z.; validation, Y.W. and P.H.; data curation, Y.W., X.L. and J.H.; writing—original draft preparation, Y.W. and P.H.; writing—review and editing, Y.W., P.H., C.Z. and Z.X.; funding acquisition, P.H., Z.X. and L.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Key Research and Development Program of China (grant numbers 2023YFC3208402 and 2023YFF0804902), the National Natural Science Foundation of China (grant numbers 52422904, 52409021, and 52309007), Science and Technology Program of Liaoning Province (Grant No. 2023-BSBA-070), and the Fundamental Research Funds for the Central Universities (Grant No. DUT24BS050). Comments from reviewers and editors are valuable in improving this study and are highly appreciated.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: Author Xixin Lu and author Jinkun Huang are employed by China MCC17 Group. Author Lu Zhang is employed by China Three Gorges Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Brezonik, P.L.; Olmanson, L.G.; Finlay, J.C.; Bauer, M.E. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sens. Environ.* **2015**, *157*, 199–215. [[CrossRef](#)]
2. Kim, J.; Kim, J.H.; Jang, W.; Pyo, J.; Lee, H.; Byeon, S.; Lee, H.; Park, Y.; Kim, S. Enhancing Machine Learning Performance in Estimating CDOM Absorption Coefficient via Data Resampling. *Remote Sens.* **2024**, *16*, 2313. [[CrossRef](#)]
3. Li, J.; Yu, Q.; Tian, Y.Q.; Becker, B.L.; Siqueira, P.; Torbick, N. Spatio-temporal variations of CDOM in shallow inland waters from a semi-analytical inversion of Landsat-8. *Remote Sens. Environ.* **2018**, *218*, 189–200. [[CrossRef](#)]
4. Shao, T.; Song, K.; Du, J.; Zhao, Y.; Liu, Z.; Zhang, B. Retrieval of CDOM and DOC using in situ hyperspectral data: A case study for potable waters in Northeast China. *J. Indian Soc. Remote Sens.* **2016**, *44*, 77–89. [[CrossRef](#)]

5. Shang, Y.; Song, K.; Jacinthe, P.A.; Wen, Z.; Lyu, L.; Fang, C.; Liu, G. Characterization of CDOM in reservoirs and its linkage to trophic status assessment across China using spectroscopic analysis. *J. Hydrol.* **2019**, *576*, 1–11. [[CrossRef](#)]
6. Wei, J.; Lee, Z.; Shang, S.; Yu, X. Semianalytical derivation of phytoplankton, CDOM, and detritus absorption coefficients from the Landsat 8/OLI reflectance in coastal waters. *J. Geophys. Res. Ocean.* **2019**, *124*, 3682–3699. [[CrossRef](#)]
7. Fan, T.; Yao, X.; Sang, D.; Liu, L.; Sun, Z.; Deng, H.; Zhang, Y.; Sun, X. Composition characteristics and metal binding behavior of macrophyte-derived DOM (MDOM) under microbial combined photodegradation: A state closer to actual macrophytic lakes. *J. Hazard. Mater.* **2024**, *465*, 133124. [[CrossRef](#)]
8. Griffin, C.; McClelland, J.; Frey, K.; Fiske, G.; Holmes, R. Quantifying CDOM and DOC in major Arctic rivers during ice-free conditions using Landsat TM and ETM+ data. *Remote Sens. Environ.* **2018**, *209*, 395–409. [[CrossRef](#)]
9. Lønborg, C.; Carreira, C.; Jickells, T.; Álvarez-Salgado, X.A. Impacts of Global Change on Ocean Dissolved Organic Carbon (DOC) Cycling. *Front. Mar. Sci.* **2020**, *7*, 466. [[CrossRef](#)]
10. Mao, D.; Tian, Y.; Wang, Z.; Jia, M.; Du, J.; Song, C. Wetland changes in the Amur River Basin: Differing trends and proximate causes on the Chinese and Russian sides. *J. Environ. Manag.* **2021**, *280*, 111670. [[CrossRef](#)]
11. Borsch, S.; Khristoforov, A.; Krovotyntsev, V.; Leontieva, E.; Simonov, Y.; Zatyagalova, V. A basin approach to a hydrological service delivery system in the Amur River Basin. *Geosciences* **2018**, *8*, 93. [[CrossRef](#)]
12. Yi, K.; Zhang, J.; Batbayar, N.; Higuchi, H.; Natsagdorj, T.; Bysykatova, I.P. Using tracking data to identify gaps in knowledge and conservation of the critically endangered Siberian crane (*Leucogeranus leucogeranus*). *Remote Sens.* **2022**, *14*, 5101.
13. Zhang, C.; Xiao, X.; Wang, X.; Yi, S.; Meng, C.; Qin, Y.; Yao, Y.; Yin, L.; Celis, J.; Pan, L. Climate-induced losses of surface water and total water storage in Northeast Asia. *Commun. Earth Environ.* **2025**, *6*, 479. [[CrossRef](#)]
14. Liu, H.; Yin, Y.; Piao, S.; Zhao, F.; Engels, M.; Ciais, P. Disappearing lakes in semiarid northern China: Drivers and environmental impact. *Environ. Sci. Technol.* **2013**, *47*, 12107–12114. [[CrossRef](#)]
15. Peng, X.; Zhang, T.; Frauenfeld, O.W.; Wang, S.; Qiao, L.; Du, R.; Mu, C. Northern Hemisphere greening in association with warming permafrost. *J. Geophys. Res. Biogeosci.* **2020**, *125*, e2019JG005086.
16. Zhou, S.; Zhang, W.; Guo, Y. Impacts of climate and land-use changes on the hydrological processes in the Amur River Basin. *Water* **2019**, *12*, 76. [[CrossRef](#)]
17. Wang, Y.; Xin, Z.; Zhang, C.; Han, P.; Pi, X.; Song, C. Revealing lake dynamics across the Amur River Basin over the past two decades using multi-source remote sensing datasets. *J. Hydrol. Reg. Stud.* **2024**, *55*, 101928. [[CrossRef](#)]
18. Zhang, L.; Xin, Z.; Guan, Q.; Feng, L.; Hu, C.; Zhang, C.; Zhou, H. Monitoring and understanding chlorophyll-a concentration changes in lakes in northeastern China using MERIS and OLCI satellite data. *GIScience Remote Sens.* **2024**, *61*, 2285166. [[CrossRef](#)]
19. Feng, P.; Song, K.; Wen, Z.; Tao, H.; Yu, X.; Shang, Y. Remote Sensing Estimation of CDOM for Songhua River of China: Distributions and Implications. *Remote Sens.* **2024**, *16*, 4608. [[CrossRef](#)]
20. Shi, J.; Zhao, Y.; Wei, D.; Zhang, D.; Wei, Z.; Wu, J. Insight into transformation of dissolved organic matter in the Heilongjiang River. *Environ. Sci. Pollut. Res.* **2019**, *26*, 3340–3349.
21. Zhu, L.; Zhao, Y.; Bai, S.; Zhou, H.; Chen, X.; Wei, Z. New insights into the variation of dissolved organic matter components in different latitudinal lakes of northeast China. *Limnol. Oceanogr.* **2020**, *65*, 471–481. [[CrossRef](#)]
22. Zhang, Y.; van Dijk, M.A.; Liu, M.; Zhu, G.; Qin, B. The contribution of phytoplankton degradation to chromophoric dissolved organic matter (CDOM) in eutrophic shallow lakes: Field and experimental evidence. *Water Res.* **2009**, *43*, 4685–4697. [[CrossRef](#)] [[PubMed](#)]
23. Olmanson, L.G.; Brezonik, P.L.; Finlay, J.C.; Bauer, M.E. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. *Remote Sens. Environ.* **2016**, *185*, 119–128. [[CrossRef](#)]
24. Sun, X.; Zhang, Y.; Zhang, Y.; Shi, K.; Zhou, Y.; Li, N. Machine learning algorithms for chromophoric dissolved organic matter (CDOM) estimation based on Landsat 8 images. *Remote Sens.* **2021**, *13*, 3560. [[CrossRef](#)]
25. Li, J.; Yu, Q.; Tian, Y.Q.; Becker, B.L. Remote sensing estimation of colored dissolved organic matter (CDOM) in optically shallow waters. *ISPRS J. Photogramm. Remote Sens.* **2017**, *128*, 98–110. [[CrossRef](#)]
26. Hoge, F.E.; Lyon, P.E. Satellite retrieval of inherent optical properties by linear matrix inversion of oceanic radiance models: An analysis of model and radiance measurement errors. *J. Geophys. Res. Ocean.* **1996**, *101*, 16631–16648. [[CrossRef](#)]
27. Ogashawara, I.; Mishra, D.R.; Nascimento, R.F.; Alcântara, E.H.; Kampel, M.; Stech, J.L. Re-parameterization of a quasi-analytical algorithm for colored dissolved organic matter dominant inland waters. *Int. J. Appl. Earth Obs. Geoinf* **2016**, *53*, 128–145. [[CrossRef](#)]
28. Chen, J.; de Hoogh, K.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Bauwelinck, M.; Van Donkelaar, A.; Hvidtfeldt, U.A.; Katsouyanni, K. A comparison of linear regression, regularization, and machine learning algorithms to develop Europe-wide spatial models of fine particles and nitrogen dioxide. *Environ. Int.* **2019**, *130*, 104934. [[CrossRef](#)]
29. Ciancia, E.; Campanelli, A.; Colonna, R.; Palombo, A.; Pascucci, S.; Pignatti, S.; Pergola, N. Improving Colored Dissolved Organic Matter (CDOM) Retrievals by Sentinel2-MSI Data through a Total Suspended Matter (TSM)-Driven Classification: The Case of Pertusillo Lake (Southern Italy). *Remote Sens.* **2023**, *15*, 5718. [[CrossRef](#)]

30. Menken, K.; Brezonik, P.; Bauer, M. Influence of chlorophyll and humic color on reflectance spectra of lakes: Implications for measurement of lake-water properties by remote sensing. *Lake Reserv. Manage* **2006**, *22*, 179–190. [\[CrossRef\]](#)
31. Keller, S.; Maier, P.M.; Riese, F.M.; Norra, S.; Holbach, A.; Börsig, N.; Wilhelms, A.; Moldaenke, C.; Zaaake, A.; Hinz, S. Hyperspectral data and machine learning for estimating CDOM, chlorophyll a, diatoms, green algae and turbidity. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1881. [\[CrossRef\]](#)
32. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3. *Remote Sens. Environ.* **2012**, *118*, 127–139. [\[CrossRef\]](#)
33. Yan, B.; Xia, Z.; Huang, F.; Guo, L.; Zhang, X. Climate change detection and annual extreme temperature analysis of the amur river basin. *Adv. Meteorol.* **2016**, *2016*, 6268938. [\[CrossRef\]](#)
34. Zhou, S.; Zhang, W.; Wang, S.; Zhang, B.; Xu, Q. Spatial–Temporal Vegetation Dynamics and Their Relationships with Climatic, Anthropogenic, and Hydrological Factors in the Amur River Basin. *Remote Sens.* **2021**, *13*, 684. [\[CrossRef\]](#)
35. Kim, J.; Jang, W.; Kim, J.H.; Lee, J.; Cho, K.H.; Lee, Y.-G.; Chon, K.; Park, S.; Pyo, J.; Park, Y. Application of airborne hyperspectral imagery to retrieve spatiotemporal CDOM distribution using machine learning in a reservoir. *Int. J. Appl. Earth Obs. Geoinf* **2022**, *114*, 103053. [\[CrossRef\]](#)
36. Xu, Y.; Feng, L.; Hou, X.; Wang, J.; Tang, J. Four-decade dynamics of the water color in 61 large lakes on the Yangtze Plain and the impacts of reclaimed aquaculture zones. *Sci. Total Environ.* **2021**, *781*, 146688. [\[CrossRef\]](#)
37. Skakun, S.; Vermote, E.F.; Roger, J.-C.; Justice, C.O.; Masek, J.G. Validation of the LaSRC cloud detection algorithm for Landsat 8 images. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2019**, *12*, 2439–2446. [\[CrossRef\]](#)
38. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [\[CrossRef\]](#)
39. Wang, X.; Song, K.; Wen, Z.; Liu, G.; Shang, Y.; Fang, C.; Lyu, L.; Wang, Q. Quantifying turbidity variation for lakes in Daqing of Northeast China using Landsat images from 1984 to 2018. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 8884–8897. [\[CrossRef\]](#)
40. Skakun, S.; Wevers, J.; Brockmann, C.; Doxani, G.; Aleksandrov, M.; Batič, M.; Frantz, D.; Gascon, F.; Gómez-Chova, L.; Hagolle, O. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. *Remote Sens. Environ.* **2022**, *274*, 112990. [\[CrossRef\]](#)
41. Wei, J.; Huang, W.; Li, Z.; Sun, L.; Zhu, X.; Yuan, Q.; Liu, L.; Cribb, M. Cloud detection for Landsat imagery by combining the random forest and superpixels extracted via energy-driven sampling segmentation approaches. *Remote Sens. Environ.* **2020**, *248*, 112005. [\[CrossRef\]](#)
42. Feng, L.; Hou, X.; Zheng, Y. Monitoring and understanding the water transparency changes of fifty large lakes on the Yangtze Plain based on long-term MODIS observations. *Remote Sens. Environ.* **2019**, *221*, 675–686. [\[CrossRef\]](#)
43. Hou, X.; Feng, L.; Dai, Y.; Hu, C.; Gibson, L.; Tang, J.; Lee, Z.; Wang, Y.; Cai, X.; Liu, J.; et al. Global mapping reveals increase in lacustrine algal blooms over the past decade. *Nat. Geosci.* **2022**, *15*, 130–134. [\[CrossRef\]](#)
44. Zhang, H.; Li, Y.; Yao, B.; Huang, Y.; Wang, S.; Ni, S. Untangling the coupling effect of water quality and quantity on lake algal blooms in Lake Hulun from a dual perspective of remote sensing and sediment cores. *J. Hydrol.* **2024**, *645*, 132141. [\[CrossRef\]](#)
45. Dee, D.P.; Uppala, S.M.; Simmons, A.J.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.; Balsamo, G.; Bauer, d.P. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [\[CrossRef\]](#)
46. Xu, Y.; Gun, Z.; Zhao, J.; Cheng, X. Variations in lake water storage over Inner Mongolia during recent three decades based on multi-mission satellites. *J. Hydrol.* **2022**, *609*, 127719. [\[CrossRef\]](#)
47. Zhang, Y.; Zhang, E.; Yin, Y.; Van Dijk, M.A.; Feng, L.; Shi, Z.; Liu, M.; Qina, B. Characteristics and sources of chromophoric dissolved organic matter in lakes of the Yungui Plateau, China, differing in trophic state and altitude. *Limnol. Oceanogr.* **2010**, *55*, 2645–2659. [\[CrossRef\]](#)
48. Zhao, D.; Feng, L.; He, X. Global gridded aerosol models established for atmospheric correction over inland and nearshore coastal waters. *J. Geophys. Res. Atmos.* **2023**, *128*, e2023JD038815. [\[CrossRef\]](#)
49. Huang, S.; Tang, L.; Hupy, J.P.; Wang, Y.; Shao, G. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. *J. For. Res.* **2021**, *32*, 1–6. [\[CrossRef\]](#)
50. Mannino, A.; Russ, M.E.; Hooker, S.B. Algorithm development and validation for satellite-derived distributions of DOC and CDOM in the US Middle Atlantic Bight. *J. Geophys. Res. Ocean.* **2008**, *113*. [\[CrossRef\]](#)
51. Qiang, S.; Song, K.; Shang, Y.; Lai, F.; Wen, Z.; Liu, G.; Tao, H.; Lyu, Y. Remote sensing estimation of CDOM and DOC with the environmental implications for Lake Khanka. *Remote Sens.* **2023**, *15*, 5707. [\[CrossRef\]](#)
52. Zhao, Z.; Shi, K.; Peng, Y.; Wang, W.; Lai, L.; Zhang, Y.; Zhou, Y.; Zhang, Y.; Qin, B. Widespread decrease in chromophoric dissolved organic matter in Chinese lakes derived from satellite observations. *Remote Sens. Environ.* **2023**, *298*, 113848. [\[CrossRef\]](#)

53. Zhu, W.; Yu, Q. Inversion of Chromophoric Dissolved Organic Matter From EO-1 Hyperion Imagery for Turbid Estuarine and Coastal Waters. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 3286–3298. [[CrossRef](#)]
54. Zhu, W.; Yu, Q.; Tian, Y.Q.; Becker, B.L.; Zheng, T.; Carrick, H.J. An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments. *Remote Sens. Environ.* **2014**, *140*, 766–778. [[CrossRef](#)]
55. Hu, C. A novel ocean color index to detect floating algae in the global oceans. *Remote Sens. Environ.* **2009**, *113*, 2118–2129. [[CrossRef](#)]
56. Rumelhart, D.E.; Hinton, G.E.; McClelland, J.L. A general framework for parallel distributed processing. *Parallel Distrib. Process. Explor. Microstruct. Cogn.* **1986**, *1*, 26.
57. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
58. Zhang, B.; Li, J.; Shen, Q.; Chen, D. A bio-optical model based method of estimating total suspended matter of Lake Taihu from near-infrared remote sensing reflectance. *Environ. Monit. Assess.* **2008**, *145*, 339–347.
59. Ficek, D.; Zapadka, T.; Dera, J. Remote sensing reflectance of Pomeranian lakes and the Baltic. *Oceanologia* **2011**, *53*, 959–970. [[CrossRef](#)]
60. Ling, Z.; Sun, D.; Wang, S.; Qiu, Z.; Huan, Y.; Mao, Z.; He, Y. Remote sensing estimation of colored dissolved organic matter (CDOM) from GOCI measurements in the Bohai Sea and Yellow Sea. *Environ. Sci. Pollut. Res.* **2020**, *27*, 6872–6885. [[CrossRef](#)] [[PubMed](#)]
61. Massicotte, P.; Assani, A.A.; Gratton, D.; Frenette, J.-J. Relationship between water color, water levels, and climate indices in large rivers: Case of the St. Lawrence River (Canada). *Water Resour. Res.* **2013**, *49*, 2303–2307. [[CrossRef](#)]
62. Tong, Y.; Zhang, W.; Wang, X.; Couture, R.-M.; Larssen, T.; Zhao, Y.; Li, J.; Liang, H.; Liu, X.; Bu, X. Decline in Chinese lake phosphorus concentration accompanied by shift in sources since 2006. *Nat. Geosci.* **2017**, *10*, 507–511. [[CrossRef](#)]
63. Wang, L.; Zhao, D.; Yang, J.; Chen, Y. Retrieval of total suspended matter from MODIS 250 m imagery in the Bohai Sea of China. *J. Oceanogr.* **2012**, *68*, 719–725. [[CrossRef](#)]
64. Song, W.; Yinglan, A.; Wang, Y.; Fang, Q.; Tang, R. Study on remote sensing inversion and temporal-spatial variation of Hulun lake water quality based on machine learning. *J. Contam. Hydrol.* **2024**, *260*, 104282. [[CrossRef](#)] [[PubMed](#)]
65. Zhao, J.; Cao, W.; Xu, Z.; Ai, B.; Yang, Y.; Jin, G.; Wang, G.; Zhou, W.; Chen, Y.; Chen, H. Estimating CDOM concentration in highly turbid estuarine coastal waters. *J. Geophys. Res. Ocean.* **2018**, *123*, 5856–5873. [[CrossRef](#)]
66. Zhou, J.; Leavitt, P.R.; Zhang, Y.; Qin, B. Anthropogenic eutrophication of shallow lakes: Is it occasional? *Water Res.* **2022**, *221*, 118728. [[CrossRef](#)]
67. Phillips, A.J.; Govedich, F.R.; Moser, W.E. Leeches in the extreme: Morphological, physiological, and behavioral adaptations to inhospitable habitats. *Int. J. Parasitol. Parasites Wildl.* **2020**, *12*, 318–325. [[CrossRef](#)]
68. Lv, Z.; Liu, C.; Wang, J.; Li, C.; Teng, X.; Tan, Y. Spectral diversities of chromophoric dissolved organic matter in paddy field water adjacent to black soil regions of Northeast China. *RSC Adv.* **2025**, *15*, 18732–18741. [[CrossRef](#)]
69. Hu, Y.; Lu, Y.; Edmonds, J.; Liu, C.; Zhang, Q.; Zheng, C. Irrigation alters source-composition characteristics of groundwater dissolved organic matter in a large arid river basin, Northwestern China. *Sci. Total Environ.* **2021**, *767*, 144372. [[CrossRef](#)]
70. Du, Y.; Chen, F.; Zhang, Y.; He, H.; Wen, S.; Huang, X.; Song, C.; Li, K.; Wang, J.; Keellings, D.; et al. Human Activity Coupled With Climate Change Strengthens the Role of Lakes as an Active Pipe of Dissolved Organic Matter. *Earth's Future* **2023**, *11*, e2022EF003412. [[CrossRef](#)]
71. Vione, D.; Minero, C.; Carena, L. Fluorophores in surface freshwaters: Importance, likely structures, and possible impacts of climate change. *Environ. Sci. Process. Impacts* **2021**, *23*, 1429–1442. [[CrossRef](#)]
72. Hamdan, M.; Byström, P.; Hotchkiss, E.R.; Al-Haidarey, M.J.; Karlsson, J. An experimental test of climate change effects in northern lakes: Increasing allochthonous organic matter and warming alters autumn primary production. *Freshw. Biol.* **2021**, *66*, 815–825. [[CrossRef](#)]
73. Mahlasi, C. *Remote Sensing of Water Quality in Inland Water Bodies*; University of Johannesburg: Johannesburg, South Africa, 2017.
74. Wiangwang, N. *Hyperspectral Data Modeling for Water Quality Studies in Michigan's Inland Lakes*; Michigan State University: East Lansing, MI, USA, 2006.
75. Wang, M.; Jiang, L. Atmospheric correction using the information from the short blue band. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6224–6237. [[CrossRef](#)]
76. Zhang, M.; Hu, C.; Cannizzaro, J.; English, D.; Barnes, B.B.; Carlson, P.; Yarbrow, L. Comparison of two atmospheric correction approaches applied to MODIS measurements over North American waters. *Remote Sens. Environ.* **2018**, *216*, 442–455. [[CrossRef](#)]
77. Hou, X.; Feng, L.; Duan, H.; Chen, X.; Sun, D.; Shi, K. Fifteen-year monitoring of the turbidity dynamics in large lakes and reservoirs in the middle and lower basin of the Yangtze River, China. *Remote Sens. Environ.* **2017**, *190*, 107–121. [[CrossRef](#)]
78. Cao, Z.; Ma, R.; Duan, H.; Pahlevan, N.; Melack, J.; Shen, M.; Xue, K. A machine learning approach to estimate chlorophyll-a from Landsat-8 measurements in inland lakes. *Remote Sens. Environ.* **2020**, *248*, 111974. [[CrossRef](#)]
79. Chen, J.; Zhu, W.; Tian, Y.Q.; Yu, Q. Monitoring dissolved organic carbon by combining Landsat-8 and Sentinel-2 satellites: Case study in Saginaw River estuary, Lake Huron. *Sci. Total Environ.* **2020**, *718*, 137374.

80. Wu, T.; Wang, Z.; Niu, C.; Zhang, Y.; Li, B.; Li, P. The effect of intense hydrodynamic disturbance on chromophoric dissolved organic matter in a shallow eutrophic lake. *J. Freshw. Ecol.* **2015**, *30*, 143–156.
81. Kieber, R.J.; Adams, M.B.; Willey, J.D.; Whitehead, R.F.; Avery, G.B., Jr.; Mullaugh, K.M.; Mead, R.N. Short term temporal variability in the photochemically mediated alteration of chromophoric dissolved organic matter (CDOM) in rainwater. *Atmos. Environ.* **2012**, *50*, 112–119. [[CrossRef](#)]
82. Wang, Y.; Feng, L.; Liu, J.; Hou, X.; Chen, D. Changes of inundation area and water turbidity of Tonle Sap Lake: Responses to climate changes or upstream dam construction? *Environ. Res. Lett.* **2020**, *15*, 0940a0941. [[CrossRef](#)]
83. Wang, B.; Xu, G.; Li, P.; Li, Z.; Zhang, Y.; Cheng, Y.; Jia, L.; Zhang, J. Vegetation dynamics and their relationships with climatic factors in the Qinling Mountains of China. *Ecol. Indic.* **2020**, *108*, 105719. [[CrossRef](#)]
84. Wauthy, M.; Rautio, M.; Christoffersen, K.S.; Forsström, L.; Laurion, I.; Mariash, H.L.; Peura, S.; Vincent, W.F. Increasing dominance of terrigenous organic matter in circumpolar freshwaters due to permafrost thaw. *Limnol. Oceanogr. Lett.* **2018**, *3*, 186–198. [[CrossRef](#)]
85. Grunert, B.K.; Tzortziou, M.; Neale, P.; Menendez, A.; Hernes, P. DOM degradation by light and microbes along the Yukon River-coastal ocean continuum. *Sci. Rep.* **2021**, *11*, 10236. [[CrossRef](#)]
86. Liu, D.; Du, Y.; Yu, S.; Luo, J.; Duan, H. Human activities determine quantity and composition of dissolved organic matter in lakes along the Yangtze River. *Water Res.* **2020**, *168*, 115132. [[CrossRef](#)] [[PubMed](#)]
87. Dupouy, C.; Röttgers, R.; Tedetti, M.; Frouin, R.; Lantoiné, F.; Rodier, M.; Martias, C.; Goutx, M. Impact of contrasted weather conditions on CDOM absorption/fluorescence and biogeochemistry in the eastern lagoon of New Caledonia. *Front. Earth Sci.* **2020**, *8*, 54. [[CrossRef](#)]
88. Cai, Y.; Qi, L.; Shan, T.; Liu, Y.; Zhang, N.; Lu, X.; Fan, Y. Application of phytoplankton taxonomic α -diversity indices to assess trophic states in Barrier Lake: A case of Jingpo Lake. *Diversity* **2022**, *14*, 1003. [[CrossRef](#)]
89. Bao, T.; Wang, P.; Hu, B.; Wang, X.; Qian, J. Mobilization of colloids during sediment resuspension and its effect on the release of heavy metals and dissolved organic matter. *Sci. Total Environ.* **2023**, *861*, 160678. [[CrossRef](#)]
90. Zhou, Y.; Dong, J.; Xiao, X.; Liu, R.; Zou, Z.; Zhao, G.; Ge, Q. Continuous monitoring of lake dynamics on the Mongolian Plateau using all available Landsat imagery and Google Earth Engine. *Sci. Total Environ.* **2019**, *689*, 366–380. [[CrossRef](#)]
91. Wen, Z.; Song, K.; Zhao, Y.; Du, J.; Ma, J. Influence of environmental factors on spectral characteristics of chromophoric dissolved organic matter (CDOM) in Inner Mongolia Plateau, China. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 787–801. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.