remote sensing

MDPI

Article

An Empirical Algorithm for Estimating the Absorption of Colored Dissolved Organic Matter from Sentinel-2 (MSI) and Landsat-8 (OLI) Observations of Coastal Waters

Vu Son Nguyen [1], Hubert Loisel [1,*], Vincent Vantrepotte [1], Xavier Mériaux [1] and Dinh Lan Tran [2]

[1] Univ. Littoral Côte d'Opale, Univ. Lille, CNRS, IRD, UMR 8187, LOG, Laboratoire d'Océanologie et de Géosciences, F 62930 Wimereux, France; vu-son.nguyen@univ-littoral.fr (V.S.N.); vincent.vantrepotte@univ-littoral.fr (V.V.); xavier.meriaux@univ-littoral.fr (X.M.)
[2] Institute of Marine Environment and Resources (IMER), Vietnam Academy of Science and Technology (VAST), Hai Phong 04000, Vietnam; lantd@imer.vast.vn
* Correspondence: hubert.loisel@univ-littoral.fr

Abstract: Sentinel-2/MSI and Landsat-8/OLI sensors enable the mapping of ocean color-related bio-optical parameters of surface coastal and inland waters. While many algorithms have been developed to estimate the Chlorophyll-a concentration, Chl-a, and the suspended particulate matter, SPM, from OLI and MSI data, the absorption by colored dissolved organic matter, a_{cdom}, a key parameter to monitor the concentration of dissolved organic matter, has received less attention. Herein we present an inverse model (hereafter referred to as AquaCDOM) for estimating a_{cdom} at the wavelength 412 nm (a_{cdom} (412)), within the surface layer of coastal waters, from measurements of ocean remote sensing reflectance, R_{rs} (λ), for these two high spatial resolution (around 20 m) sensors. Combined with a water class-based approach, several empirical algorithms were tested on a mixed dataset of synthetic and in situ data collected from global coastal waters. The selection of the final algorithms was performed with an independent validation dataset, using in situ, synthetic, and satellite R_{rs} (λ) measurements, but also by testing their respective sensitivity to typical noise introduced by atmospheric correction algorithms. It was found that the proposed algorithms could estimate a_{cdom} (412) with a median absolute percentage difference of ~30% and a median bias of 0.002 m^{-1} from the in situ and synthetic datasets. While similar performances have been shown with two other algorithms based on different methodological developments, we have shown that AquaCDOM is much less sensitive to atmospheric correction uncertainties, mainly due to the use of band ratios in its formulation. After the application of the top-of-atmosphere gains and of the same atmospheric correction algorithm, excellent agreement has been found between the OLI- and MSI-derived a_{cdom} (412) values for various coastal areas, enabling the application of these algorithms for time series analysis. An example application of our algorithms for the time series analysis of a_{cdom} (412) is provided for a coastal transect in the south of Vietnam.

Keywords: a_{cdom} (412); Landsat-8; Sentinel-2; coastal water; aquaculture

check for updates

Citation: Nguyen, V.S.; Loisel, H.; Vantrepotte, V.; Mériaux, X.; Tran, D.L. An Empirical Algorithm for Estimating the Absorption of Colored Dissolved Organic Matter from Sentinel-2 (MSI) and Landsat-8 (OLI) Observations of Coastal Waters. *Remote Sens.* **2024**, *16*, 4061. https://doi.org/10.3390/rs16214061

Academic Editor: SeungHyun Son

Received: 23 September 2024
Revised: 26 October 2024
Accepted: 29 October 2024
Published: 31 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Colored dissolved organic matter (CDOM) is a photoactive component of dissolved organic matter (DOM) that occurs naturally in aquatic systems from the degradation of detritus and organic materials [1]. The sources of CDOM are either autochthonous (phytoplankton and submerged vegetation) [2], allochthonous (of terrestrial origin via surface runoff) [3], or anthropogenic (e.g., agriculture activities) [4]. As organic matter, CDOM plays a significant role in the carbon cycle [5], and its nitrogen and phosphorus content indirectly regulate the nutrient availability for algae and aquatic plant growth [6,7]. At the same time, its decomposition boosts the microbial consumption of dissolved oxygen [8]. Furthermore, CDOM absorbs light strongly in the ultraviolet (UV) and visible

wavelengths [9], therefore regulating light availability in the water column. In a water column, a high CDOM concentration limits the penetration of photosynthetically active radiation (PAR), essential for primary production, and protects light-sensitive species from UV radiation [1]. Due to its importance to aquatic environments, numerous studies have focused on investigating the variability of CDOM over large spatial scales using ocean color radiometry (OCR) [10–13].

Different algorithms have been developed to estimate the absorption coefficient of CDOM, a_{cdom} (λ), from standard ocean color sensors including the MEdium Resolution Imaging Spectrometer (MERIS), the Moderate-Resolution Imaging Spectroradiometer (MODIS), the Sea-viewing Wide Field-of-view Sensor (SeaWiFS), and the Ocean and Land Color Instrument (OLCI), among others [13–17]. These sensors have a relatively short revisit time (up to one day) that is well adapted to high-frequency monitoring. However, their spatial resolution, around 1 km^2, is unsuitable for the observation of small-to-medium estuaries and highly diverse coastal waters. In such cases, an enhanced resolution of lower than 100 m is required. Moreover, sub-pixel variations can also induce an underestimation of space-retrieved ocean color parameters from coarse spatial resolution observations [18]. For small-scale inland and nearshore waters, the use of more recent high spatial resolution ($\leq$30 m) sensors like the Operational Land Imager (OLI) onboard Landsat-8, and the Multi-Spectral Instrument (MSI) aboard the Sentinel-2A and 2B (hereafter referred to collectively as Sentinel-2) constellation are, therefore, preferable [19,20]. Due to their much lower revisit time (i.e., 16 days for OLI), high spatial resolution sensors have to be combined to reduce their sampling interval. The combination boosts their potential for monitoring very nearshore aquatic environments, where high-frequency processes generally occur. For this reason, OLI and MSI on Sentinel-2 A and B are usually combined [20,21] to theoretically reduce the sampling interval to 2 to 3 days [22]. For a seamless merging for temporal studies, it is also essential to ensure the consistency between the two sensors at the radiometric [19] and bio-optical product levels.

Several studies have attempted to develop inverse algorithms to estimate a_{cdom} (λ) from OLI and MSI observations. These models can be empirical using simple statistical relationships [23–27], based on neural networks [28–30], or semi-analytical [31–34]. These different algorithms were developed for both inland and coastal waters using in situ measurements [23–25] or only simulated data [28]. While these algorithms have been developed for aquatic environments for which atmospheric correction (AC) is still challenging due to their bio-optical complexity and to the immediate proximity of land (i.e., the adjacency effect), the impact of AC-related uncertainties on their performance has not necessarily been considered. Furthermore, except for the semi-analytical algorithm proposed by [33], all these algorithms aim at estimating a_{cdom} at 443 nm, a wavelength where the absorption of Chlorophyll-a, Chl-a, also peaks [35], which can potentially impede the separation of the absorption between CDOM and Chl-a. For this reason, a_{cdom} algorithms developed for traditional ocean color sensors (MERIS, MODIS, and SeaWiFS) estimate a_{cdom} at the wavelength 412 nm or shorter to avoid such interference [14–17,36]. The preference for 443 nm in current algorithms aiming at estimating a_{cdom} from OLI and MSI observations could be attributed to the lack of the 412 nm band on these two sensors, a fact that was addressed by [33] who proposed a method to generate a virtual 412 nm spectral band. However, the performance of [33] depends on the accuracy of the virtual band estimator which could be affected by atmospheric correction noise.

This study has two main objectives:

(1) To create new empirical algorithms to retrieve, from visible bands, a_{cdom} (λ) at the wavelength 412 nm, a_{cdom} (412), from OLI and MSI, using an extensive in situ dataset collected in bio-optically contrasted coastal waters, mixed with a synthetic dataset to expand the retrievable range of a_{cdom} (λ). Different empirical formulations will be tested against an independent in situ and synthetic dataset and their performances will be quantified from a match-up exercise for both OLI and MSI. The selection of the final empirical formulations will be based on these latter validation exercises, but

also, through the examination of their sensitivity to the noise induced by atmospheric corrections. For that purpose, the different datasets used in this study and the methodology for algorithm development and noise sensitivity analysis will first be presented. Then, the validation and selection of the final algorithms will be provided, and their performances will be compared to two other recently developed algorithms based on different methodologies.

(2) To demonstrate, through a case study, the full methodology to obtain a consistent bio-optical product from combined OLI and MSI observations. This case study analyzes the temporal variability of a_{cdom} (412) obtained from OLI and MSI observations over a Vietnamese coastal area impacted by aquaculture activities.

2. Materials and Methods

2.1. A Description of the Different Datasets

2.1.1. The Development of the Mixed Dataset

A mixed dataset (hereafter referred to as DSM) of synthetic and in situ data was created to develop the inverse algorithm for retrieving a_{cdom} (412) from OLI and MSI observations. The synthetic data were generated from radiative transfer simulations and are available in the fifth report of the International Ocean Color Coordinating Group (IOCCG) [37]. While IOCCG published the synthetic dataset (available at http://www.ioccg.org/groups/OCAG_data.html (accessed on 29 April 2021)) for two zenith angles, 30° and 60°, we only used the synthetic data generated at 30° in the present study (500 data points). The spectral resolution of the IOCCG dataset is 10 nm, which requires interpolation to obtain a_{cdom} (412) and the R_{rs} (λ) values at the four visible central bands of OLI and MSI. For a_{cdom} (λ), we used the provided slope of the a_{cdom} (λ) spectra described by an exponential function [38]; for R_{rs} (λ), we used a spline interpolation following [39]. The in situ measurements (890 data points) included in DSM were gathered from various coastal zones (Figure 1) around Europe, French Guiana, and Vietnam and have already been described in previous studies [40–44]. The in situ measurement includes hyperspectral R_{rs} data (from 310 nm to 950 nm), a_{cdom} (412), and for some data points, the full a_{cdom} (λ) spectrum is also available from 250 nm to 850 nm. In total, DSM has 1390 data points and covers a wide range of a_{cdom} (412) values (0.0037–6.0611 m^{-1}). The statistics of DSM and its synthetic and in situ a_{cdom} (412) subsets are shown in Table 1.

Table 1. Summary of DSM and its synthetic and in situ subsets (N is the number of data points) for a_{cdom} (412).

Dataset	N	Mean (m^{-1})	Standard Deviation (m^{-1})	Min (m^{-1})	Max (m^{-1})
Synthetic dataset	500	0.4999	0.7229	0.0037	3.5439
In situ dataset	890	0.4487	0.5851	0.0142	6.0611
Mixed dataset (DSM)	1390	0.4671	0.6383	0.0037	6.0611

2.1.2. Development of the Satellite Dataset

The OLI Level-1 Collection-2 and MSI Level-1C images used in this study were downloaded using the United States Geological Survey (USGS) Machine-to-Machine (M2M) API (https://m2m.cr.usgs.gov/ (accessed on 29 June 2022)) and Google's Sentinel-2 Public Cloud Storage (https://cloud.google.com/storage/docs/public-datasets/sentinel-2 (accessed on 15 February 2023)), respectively. OLI has a spatial resolution of 30 m, while MSI has three spatial resolutions (10, 20, and 60 m) depending on the spectral band. Nonetheless, for MSI, we decided to perform atmospheric correction at a spatial resolution of 10 m to fully evaluate the algorithm's potential for the monitoring of waters in estuaries. The satellite scenes cover the match-up dataset for a_{cdom} (412) and R_{rs} (λ) described in Section 2.1.3. In addition, 11 near-simultaneous nadir-captured Landsat-8 and Sentinel-2 image pairs were randomly selected to evaluate the coherence of the a_{cdom} (412) products estimated

from OLI and MSI for future time series analysis. These image pairs cover various coastal areas in different latitude zones and with various sun zenith angles (Figure 1). Finally, to demonstrate the potential of combining Landsat-8 and Sentinel-2, we applied our composite algorithm to the Ganh Hao River estuary in the Mekong Delta region in Vietnam (Figure 2). This area was chosen due to the intensive aquaculture activities (especially shrimp farming), which could contribute to the coastal variability of a_{cdom} (412) [45]. A transect in front of the river mouth was chosen for analysis of the variability of a_{cdom} (412) from September 2018 to April 2020. The 80 Landsat-8 and 148 Sentinel-2 images used for the transect analysis have a maximum cloud coverage of 30%, and the transect pixels have at most 50% invalid masked pixels. For all the selected images, the R_{rs} (λ) values were retrieved from the top-of-atmosphere (TOA) signal measured by OLI and MSI using the ACOLITE (version 20221114.0) atmospheric correction algorithm [46,47], and non-water pixels were masked using the WiPE algorithm [48].

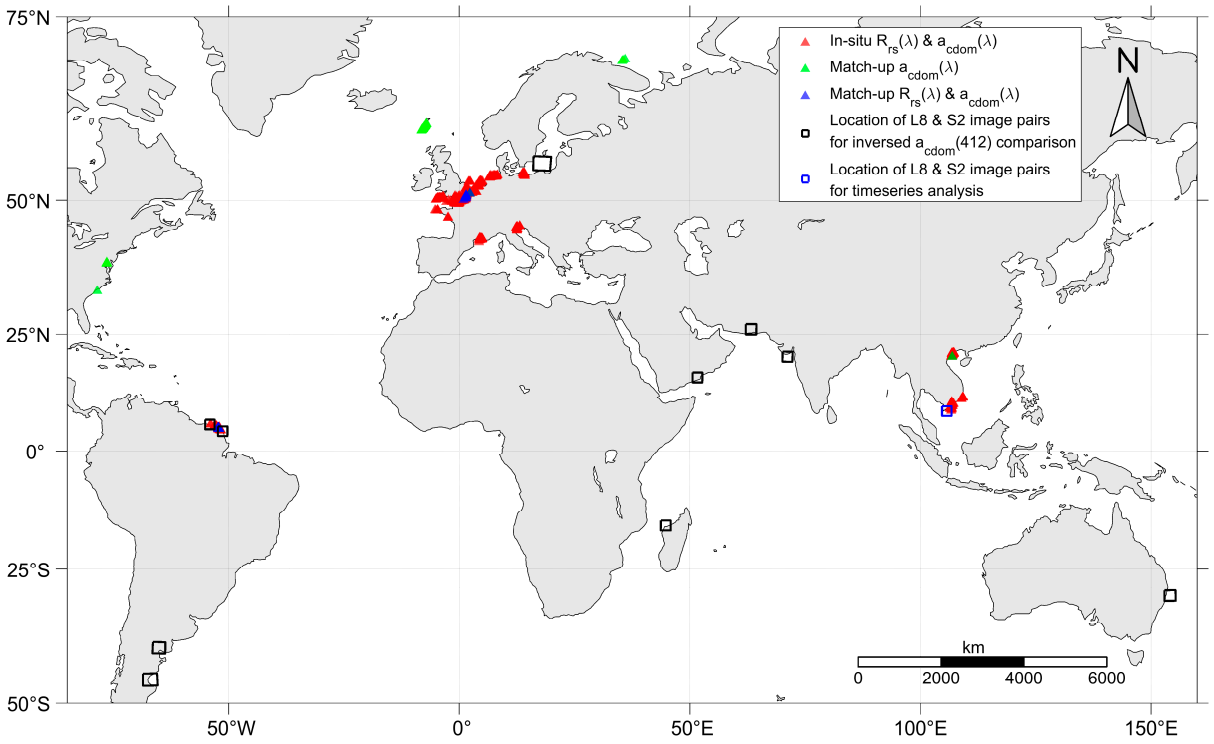


Figure 1. Global distribution of the in situ (red triangles) and match-up dataset for a_{cdom} (λ) (green triangles) and for both R_{rs} (λ) and a_{cdom} (λ) (blue triangles) data points. The squares represent the locations of Landsat-8 and Sentinel-2 near-simultaneous nadir-captured image pairs used for the calibration of a_{cdom} (412) product (black squares) and the time series analysis (blue squares).

2.1.3. Development of the Match-Up Dataset for a_{cdom} (412) and R_{rs} (λ)

A set of a_{cdom} (412) and R_{rs} (λ) match-up data points for OLI and MSI was extracted from the in situ dataset presented in Section 2.1 as well as from a few stations (see the green and red triangles in Figure 1) gathered in Chesapeake Bay, the Florida coast, the Barents Sea, and near the Faroe Islands [49]. To select the satellite images for match-up analysis, the acceptable time difference between in situ and satellite measurements should be within ±0.5 h for complex dynamic water [50–52] and up to ±12 h [52] for open ocean water. However, since our in situ dataset is limited, the maximum time difference is set at ±24 h for the current study. The extraction of the match-up data points from the selected atmosphere-corrected water pixels (see Section 3.1) was performed following the procedure described in [53]. In practice, daily match-ups were produced using a 3 × 3 pixel window, in which the coefficient of variation of R_{rs} (λ) needs to be below 0.15 with a minimum of 5 valid pixels. It should be noted that the same 3 × 3 pixel window was used for both OLI and

MSI, despite them having different spatial resolutions (30 and 10 m, respectively), because (1) we were assessing the performance of the algorithms on each sensor independently, and (2) we did not want to introduce resampling uncertainties to the match-ups. The a_{cdom} (412) algorithms were then applied to the mean of R_{rs} (λ) value calculated on the 3 × 3 pixel box. The resulting match-up datasets are hereafter referred to as Mu-R_{rs} (24 and 30 data points for OLI and MSI, respectively) and Mu-CDOM (61 and 24 data points for OLI and MSI, respectively). While the time gap condition to search for match-up images is set to less than ±24 h, the maximum time gaps of Mu-R_{rs} and Mu-CDOM for water Class 1 and 2 (see definitions in Section 2.2.1) are 11.8 and 1.2 h for OLI, and 1.3 and 4.1 h for MSI, respectively.

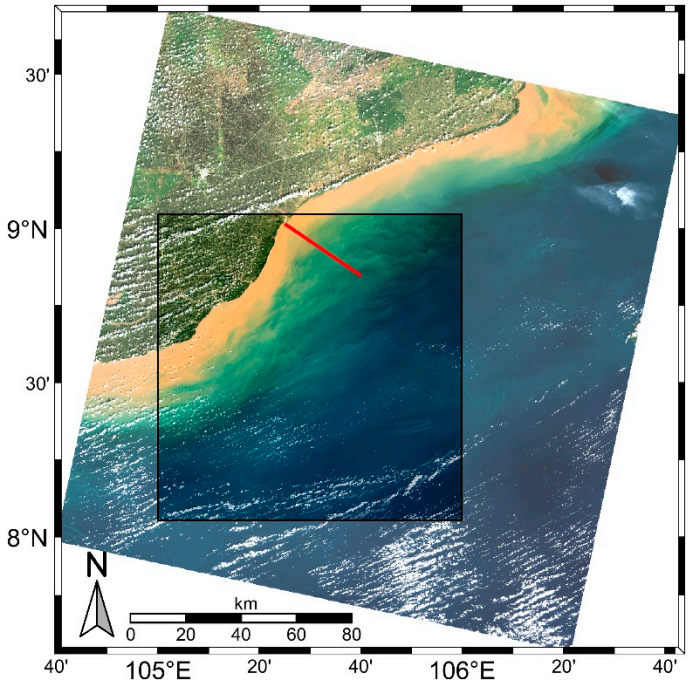


Figure 2. The RGB of Landsat-8 (dimmed image) and Sentinel-2 (bounded by black box) images used for algorithm sensitivity analysis. The red line is the transect in front of the Ganh Hao River in Vietnam.

2.2. The Development of the Empirical Algorithm

2.2.1. Classification of Low and High a_{cdom} (412) Conditions

Jiang et al. (2021) [54] proposed a water classification scheme that compares the R_{rs} (λ) values at 490, 560, 620, and 754 nm to identify four water types based on the level of turbidity. We follow a similar approach using the R_{rs} (λ) values at the first three bands of OLI and MSI. Unlike Jiang et al. (2021) [54], we only define two water classes based on the a_{cdom} (412) value, where Class 1 waters are characterized by low values (hereafter referred to as DSM-1), and Class 2 waters (hereafter referred to as DSM-2) by medium-to-high a_{cdom} (412) values. While Jiang et al. (2021) used a purely synthetic dataset to formulate their classification, the present one is based on DSM including both synthetic and in situ data. Based on different tests, a simple condition using the Blue 1 (B1), Blue 2 (B2), and Green (G) bands (centering at 443, 490, and 560 nm, respectively) of both MSI and OLI was formulated. A flow chart of our water classification process is presented in Figure 3, while Figure 4 shows the distribution of a_{cdom} (412) for DSM, DSM-1, and DSM-2. A pixel is considered Class 1 if both the reflectance values at B1 and B2 are higher than the reflectance value at the G band, otherwise, it belongs to Class 2. The median values of the two water classes, 0.0275 m^{-1} for Class 1 and 0.34 m^{-1} for Class 2, confirm that the condition successfully separates two ranges of a_{cdom} (412) values. It is also noted that the distribution of the two water classes overlaps in the log10 (a_{cdom} (412)) range between −1 and −0.5, indicating

that the algorithms for each class can reasonably estimate each other over this common range. This creates a soft boundary between the water classes on the a_{cdom} (412) map, of which an example is shown in Figure A3 using a sample OLI image.

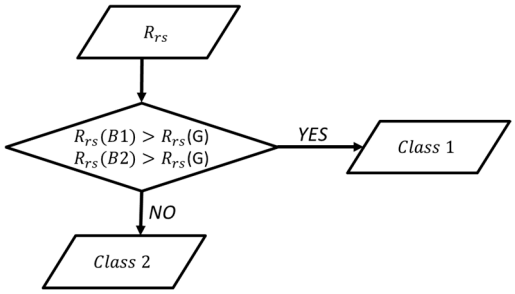


Figure 3. Flowchart of the water classification method proposed in this study.

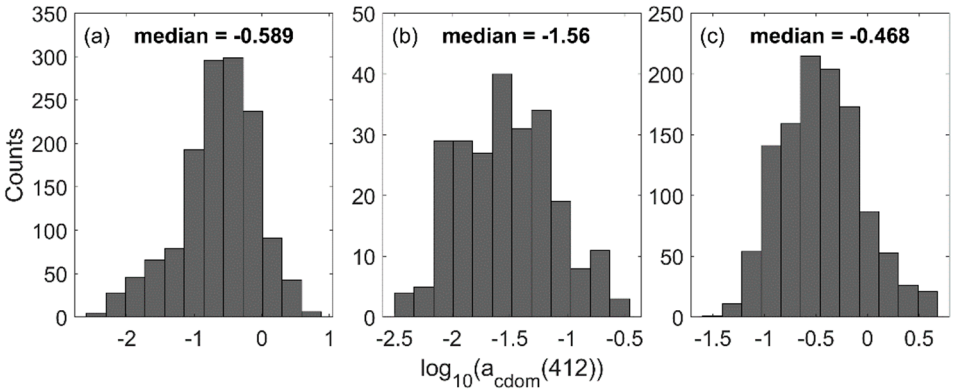


Figure 4. Histograms of the a_{cdom} (412) values for (**a**) DSM, (**b**) DSM-1 (i.e., Class 1), and (**c**) DSM-2 (i.e., Class 2).

2.2.2. The Development of an Empirical Algorithm for Each Water Class

For each water class, we have tested different algorithms, based on four different mathematical formulations, to estimate a_{cdom} (412) from R_{rs} (λ). The first one is the Multilinear Regression Model (MLR) to predict $\log_{10}$ of a_{cdom} (412) from $\log_{10}$ of R_{rs} (λ) at four visible bands. The second type comprises a set of four polynomial functions from the 1st to the 4th order (Poly1, Poly2, Poly3, and Poly4) using one of the five $\log_{10}$ of R_{rs} (λ) band ratios. The band ratios considered are Blue 1/Red (B1R), Blue 2/Red (B2R), Blue 1/Green (B1G), Blue 2/Green (B2G), and Green/Red (GR). Finally, the power law (Power) and exponential (EXP) functions of a_{cdom} (412) using the same R_{rs} (λ) band ratios were tested. In total, 31 algorithms were tested for each water class in this study.

To identify the optimal version of each algorithm, we performed 10,000 iterations of fitting. Each iteration consists of the following processes:

1. First, 30% of Class 2 data from DSM were randomly removed from the 5% to the 95% quantile of the distribution of $\log_{10}$ of a_{cdom} (412) to reduce data scattering. In contrast, no random removal was applied to Class 1 data due to their limited size.
2. The remaining data were then divided into two sub-datasets: one for development (referred to as DSM-D with 70% of data points) and one for validation (referred to as DSM-V with 30% of data points); without significantly modifying DSM's original $\log_{10}$ (a_{cdom} (412)) distribution.
3. The algorithms were fitted using DSM-D and their statistics (see Equations (1)–(11)) were calculated for both DSM-D and DSM-V sub-dataset.
4. From each water class, only the algorithms having an R^2 value of at least 0.65 for both development and validation sub-datasets were kept in each iteration.

The best iteration was finally identified as the one where the highest number of algorithm formulations remained (having $R^2 \geq 0.65$) and where the sum of the polygon area indicator calculated from DSM-V was the lowest.

2.3. Sensitivity Analysis of the a_{cdom} (412) Algorithm to Atmospheric Correction Noise

We performed a sensitivity analysis of each a_{cdom} (412) algorithm formulation in response to a random, band-specific noise introduced by the atmospheric correction algorithm. For that purpose, we selected a Landsat-8 and Sentinel-2 near-simultaneous nadir-captured image pair over a contrasted coastal site in the south of Vietnam (see Figure 1). Atmospheric correction (AC) of the downloaded Landsat-8 and Sentinel-2 images was performed using ACOLITE (version 20221114.0). The noise level of each satellite band was defined based on the results of the Atmospheric Correction Inter-Comparison Exercise (ACIX-Aqua) for OLI and MSI, where the performances of 8 ACs were compared [55]. The band-specific noise is here quantified using the MAPD values obtained from their match-up exercise at each band which are: 62.10%, 28.72%, 19.56%, and 21.46% at 443, 490, 560, and 664 nm, respectively [55]. The noise at each band was generated 100 times, forming a normal distribution with a mean average value of 0 and a standard deviation equal to the MAPD values. We applied the noise distribution to each R_{rs} (λ) value to create a set of 100 noisy R_{rs} (λ) images, from which the a_{cdom} (λ) algorithm formulations are applied to obtain an ensemble of 100 a_{cdom} (λ) images. The standard deviation maps of a_{cdom} (λ) were finally taken as the proxy for the uncertainty of each algorithm. For MSI, due to limited computation memory, the sensitivity analysis was performed using the 60 m spatial resolution data.

2.4. Statistical Indicators for Algorithm Performance Assessment

The performance of the models was assessed using several standard statistical indicators, which are the root mean square deviation (RMSD), the median absolute percentage difference (MAPD) in %, the mean bias (MB) in the unit of the parameter, and the median ratio MR. The significance of these parameters is discussed in Loisel et al. (2018) [56]. These parameters are calculated as follows

$$\mathrm{RMSD_{log}} = \sqrt{\frac{\sum_{i=1}^{N}\left(\log_{10}\left(X_i^{\mathrm{estimated}}\right) - \log_{10}\left(X_i^{\mathrm{measured}}\right)\right)^2}{N}} \quad (1)$$

$$\mathrm{MAPD} = \mathrm{median}\left(\frac{\left|X_i^{\mathrm{estimated}} - X_i^{\mathrm{measured}}\right|}{X_i^{\mathrm{measured}}}\right) \times 100 \quad (2)$$

$$\mathrm{MB} = \mathrm{median}\left(X_i^{\mathrm{estimated}} - X_i^{\mathrm{measured}}\right) \quad (3)$$

$$\mathrm{MR} = \mathrm{median}\left(\frac{X_i^{\mathrm{estimated}}}{X_i^{\mathrm{measured}}}\right) \quad (4)$$

where N is the number of data points and X_i stands either for a_{cdom} (412) or for R_{rs} (λ). The slope and R-square of the linear regression between the estimated and in situ measured values (in a $\log_{10}$ unit for a_{cdom} (412)) are also used. Following Tran et al. (2023) [57], these statistical parameters were summarized in radar/spider plots where smaller polygon areas indicate improved model performance. The polygons are calculated from the normalized values of the six following statistical indicators:

$$\mathrm{RMSD_{log}}^{\mathrm{norm}} = \frac{\mathrm{RMSD_{log}}}{\max\left(\mathrm{RMSD_{log}}\right)} \quad (5)$$

$$\mathrm{MAPD}^{\mathrm{norm}} = \frac{\mathrm{MAPD}}{\max(\mathrm{MAPD})} \tag{6}$$

$$\mathrm{MB}^{\mathrm{norm}} = \frac{\mathrm{abs}(\mathrm{MB})}{\max(\mathrm{abs}(\mathrm{MB}))} \tag{7}$$

$$\mathrm{MR}^{\mathrm{norm}} = \frac{\mathrm{MR}}{\max(\mathrm{MR})} \tag{8}$$

$$\mathrm{Slope}^{\mathrm{norm}} = \frac{\mathrm{abs}(1-\mathrm{Slope})}{\max(\mathrm{abs}(1-\mathrm{Slope}))} \tag{9}$$

$$\mathrm{R}_{\mathrm{squared}}^{\mathrm{norm}} = \frac{\mathrm{abs}\left(1-\mathrm{R}_{\mathrm{squared}}\right)}{\max\left(\mathrm{abs}\left(1-\mathrm{R}_{\mathrm{squared}}\right)\right)} \tag{10}$$

The formula for the unitless area of a polygon is calculated as follows

$$\begin{aligned} Area = \tfrac{1}{2}\cdot\sin\tfrac{\pi}{6}\cdot(&{RMSD_{log}}^{norm}\cdot MAPD^{norm} + MAPD^{norm}\cdot MB^{norm} \\ &+MB^{norm}\cdot MR^{norm} + MR^{norm}\cdot Slope^{norm} + Slope^{norm} \\ &\cdot R_{squared}^{norm} + R_{squared}^{norm}\cdot RMSD_{l}og^{norm}) \end{aligned} \tag{11}$$

3. Results

3.1. Atmospheric Correction of MSI and OLI Observations

One objective of this study is to showcase the potential of combining OLI and MSI to capture the dynamics of coastal CDOM at the relatively high temporal resolution of ~2.9 days near the equator [22]. For that purpose, it is essential to ensure the consistency of the Level-1 TOA reflectance [19] via TOA band-specific gain application before atmospheric correction. The gains for the Sentinel-2 constellation were taken from [19] and those for Landsat-8 from [58]. Atmospheric correction (AC) of the Level-1 images of both sensors was performed using the ACOLITE program, which uses the Dark Spectrum Fitting (DSF) method [47,59] initially developed for the correction of Landsat-8 products over turbid to extremely turbid water [60,61]. We enabled ACOLITE's optional image-based glint correction, which uses the surface reflectance at 1.6 and 2.1 µm, for our images since most of our match-ups are located in the tropics (French Guiana and Vietnam), largely impacted by permanent sun glint [62]. The detection of water pixels was performed using the WiPE algorithm [48], instead of ACOLITE's default water detection, which is based on a simple threshold method [59]. The application of these different gains and ACOLITE's atmospheric correction to Mu-R_{rs} shows that R_{rs} (λ) can be retrieved with a relatively good performance for both OLI (Figure 5a) and MSI (Figure 5b). Nonetheless, for both sensors, a slight overestimation of R_{rs} (λ) is still observed in the lower range. Note that we also tested POLYMER [63,64], which presents lower performance on our match-up dataset. The choice of ACOLITE here is motivated by the fact that, since R_{rs} (λ) retrieval is relatively good, the validation using the Mu-CDOM dataset, which includes all the data points of Mu-R_{rs}, will mainly reflect the performance of the bio-optical algorithm.

3.2. Validation of the Algorithms on DSM-V for OLI and MSI

The performance of the different algorithms has been evaluated on DSM-V for the two water classes and each considered sensor (Figure 6). Among the four distinct families of algorithms, MLR shows the best performance (i.e., the smallest area on the radar plot) for Classes 1 and 2, except for Class 2 and OLI where its performance is slightly lower than Poly 2 and 3 (Figure 6d). For Class 1 waters, Poly1_B1G and Power_B2G also present, while slightly lower than MLR, excellent performances for both MSI (Figure 6a) and OLI (Figure 6c). The selection of the final algorithm for both Class 1 and 2 was performed by testing not only MLR, but also other algorithms which present only slightly lower performances (Figure 6), on the match-up dataset (Mu-CDOM), and by examining their respective sensitivity to noise propagation. These algorithms are (1) for MSI

and Class 1, Poly1_B1G, Power_B2G, and Exp_B2G; (2) MSI and Class 2, Poly2_B1R, Poly3_B2R, Power_B1R, Poly1_B1R, and Poly4_B1R; (3) for OLI, and Class 1 Power_B1G, and Poly1_B1G; and (4) OLI for Class 2, Poly3_B1R, Poly2_B1R, Poly4_B1R, and Power_B1R.

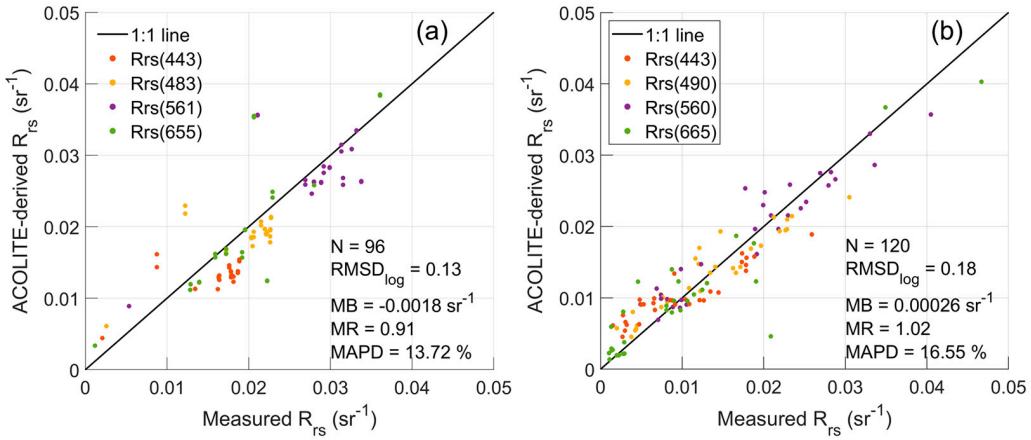


Figure 5. Comparison of ACOLITE-derived and measured values of R_{rs} (λ) from the (**a**) OLI and (**b**) MSI 10 m spatial resolution (right) match-up dataset. The different statistical indicators calculated for the data of ACOLITE-derived versus measured R_{rs} (λ) are provided (see text for details). The solid line is the 1:1 line. The different colors stand for a given wavelength.

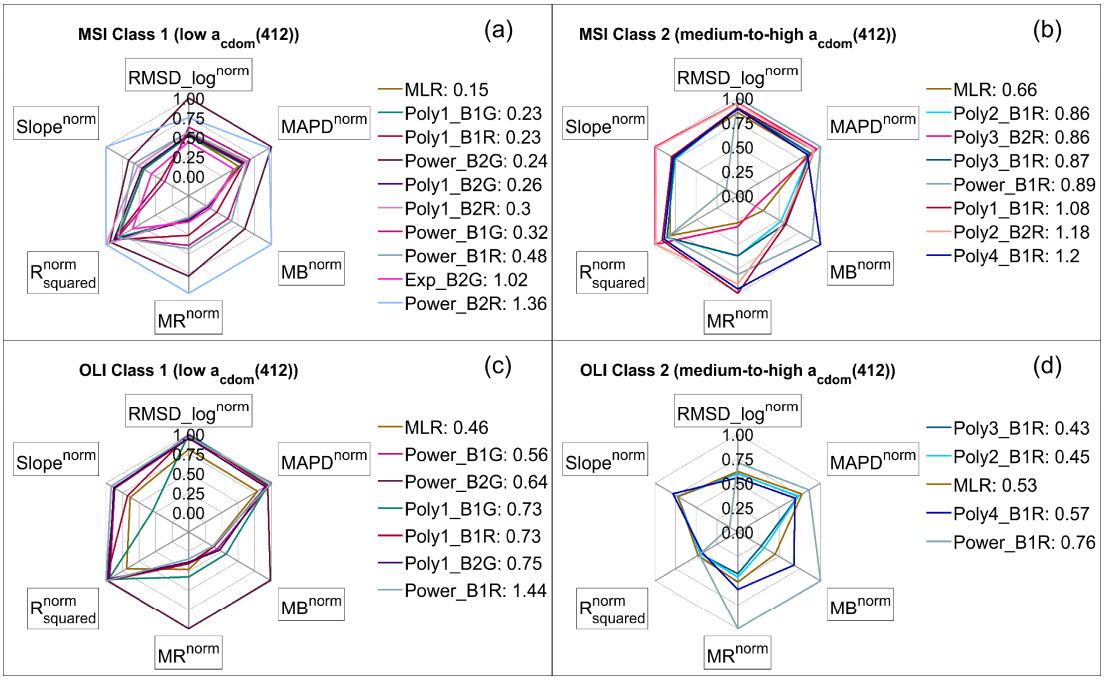


Figure 6. Radar plots summarizing the performances of the different a_{cdom} (412) algorithms as indicated for (**a**) MSI and Class 1 water, (**b**) MSI and Class 2 water, (**c**) OLI and Class 1 waters, and (**d**) OLI and Class 2 waters. The number following the algorithm name represents the value of the surface of the polygon (the smaller the value the better the algorithm performs).

3.3. Selection of Final Algorithms for OLI

The many different combinations of algorithms for Class 1 and 2 selected in Section 3.2 were tested on the OLI's Mu-CDOM match-up dataset. While combinations with an R^2 less than 0.65 were discarded, the remaining combinations were ranked based on the area indicator, and the best combinations are shown in Table 2. Based on this exercise, it appears

that Poly1_B1G for Class 1 in combination with any of the five Class 2 algorithms presents the best match-up performances, i.e., having the lowest polygon area. In contrast, although MLR for Class 1 shows the best performance on DSM-V, it presents the lowest retrieval accuracy of a_{cdom} (412) over Mu-CDOM compared to the other empirical formulations. This could be explained by atmospheric correction errors which impact the different absolute values of R_{rs} (λ) used in MLR, especially in relatively clear waters (i.e., low signal-to-noise ratio (SNR)). In contrast, the use of R_{rs} (λ) band ratios by the other formulations may help to mitigate the atmospheric correction errors.

Table 2. The statistical indicators of the evaluation of different Class 1 and Class 2 algorithm combinations on Mu-CDOM for OLI.

Class 1	Class 2	$RMSD_{log}$	MAPD	MB	MR	Slope	R^2	Area
Poly1_B1G	**MLR**	0.268	70.908	0.0256	1.613	0.698	0.89	0.0364
Poly1_B1G	**Poly2_B1R**	0.273	71.297	0.0256	1.662	0.689	0.88	0.039
Poly1_B1G	**Poly3_B1R**	0.2738	71.29	0.0256	1.648	0.6878	0.88	0.0388
Poly1_B1G	**Poly4_B1R**	0.274	71.581	0.0252	1.648	0.686	0.88	0.039
Poly1_B1G	**Power_B1R**	0.275	73.478	0.0249	1.691	0.708	0.87	0.0399

The assessment of the sensitivity of the a_{cdom} (412) algorithm to atmospheric correction is performed through the examination of the standard deviation maps of a_{cdom} (412) (see Figure 7) calculated, as previously described in Section 2.3, for each combination of algorithms selected on the match-up dataset and listed in Table 2. All standard deviation maps are characterized by the same spatial patterns, with the lowest standard deviation values in offshore waters and the highest ones in the coastal band. These results clearly show that Poly4_B1R (Figure 7d), when used for Class 2 (i.e., medium-to-high a_{cdom} (412) waters), presents the lowest standard deviation compared to the other formulations. In contrast, the three best algorithms for Class 1 selected through the match-up exercise present very similar patterns in terms of standard deviation over the offshore water. Based on this sensitivity analysis as well as on the match-up exercise the Poly1_B1G/Poly4_B1R combination was selected to derive a_{cdom} (412) for OLI.

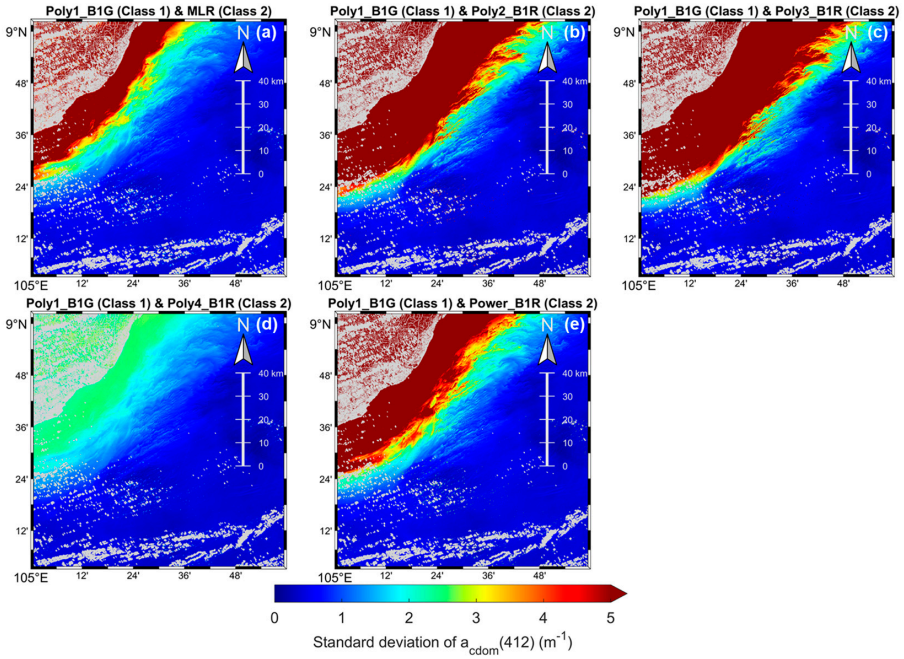


Figure 7. The OLI maps of the standard deviation of modeled a_{cdom} (412) for the algorithm combinations listed in Table 2. The name of the algorithm for each water class in each combination is displayed at the top of each map from (**a**–**e**).

3.4. Selection of Final Algorithms for MSI

Similarly to OLI, the performance of the combination of algorithms selected in Section 3.2, based on the validation dataset DSM-V, was then evaluated using the match-up dataset Mu-CDOM (Table 3) and the standard deviation maps for MSI (Figure 8). The same selection criteria as Section 3.3 were employed for the match-up and the top combinations are shown in Table 3. The best results are obtained when Poly1_B1G (Class 1) and Poly3_B2R (Class 2) are combined. In contrast to the results observed for OLI, the MAPD calculated with the match-up dataset was much lower. This difference, although the number of match-up data points is not the same, may be due to atmospheric correction errors. The standard deviation maps, obtained with the best algorithm for each sensor, show that the final a_{cdom} (412) product is more sensitive to atmospheric correction errors for OLI (Figure 7d) than for MSI (Figure 8d) for both open and coastal waters.

Table 3. The statistical indicators of the evaluation of different Class 1 and Class 2 algorithm combinations on Mu-CDOM for MSI.

Class 1	Class 2	$RMSD_{log}$	MAPD	MB	MR	Slope	R^2	Area
Poly1_B1G	**MLR**	0.262	41.585	0.182	0.665	0.773	0.606	0.678
Poly1_B1G	**Poly1_B1R**	0.211	29.142	0.136	0.763	0.751	0.65	0.679
Poly1_B1G	**Poly2_B1R**	0.216	33.24	0.169	0.709	0.721	0.669	0.716
Poly1_B1G	**Poly3_B2R**	0.165	18.757	0.054	0.897	0.676	0.7	0.716
Poly1_B1G	**Poly4_B1R**	0.216	31.725	0.16	0.716	0.745	0.662	0.758
Poly1_B1G	**Power_B1R**	0.241	36.616	0.105	0.754	0.876	0.614	0.777

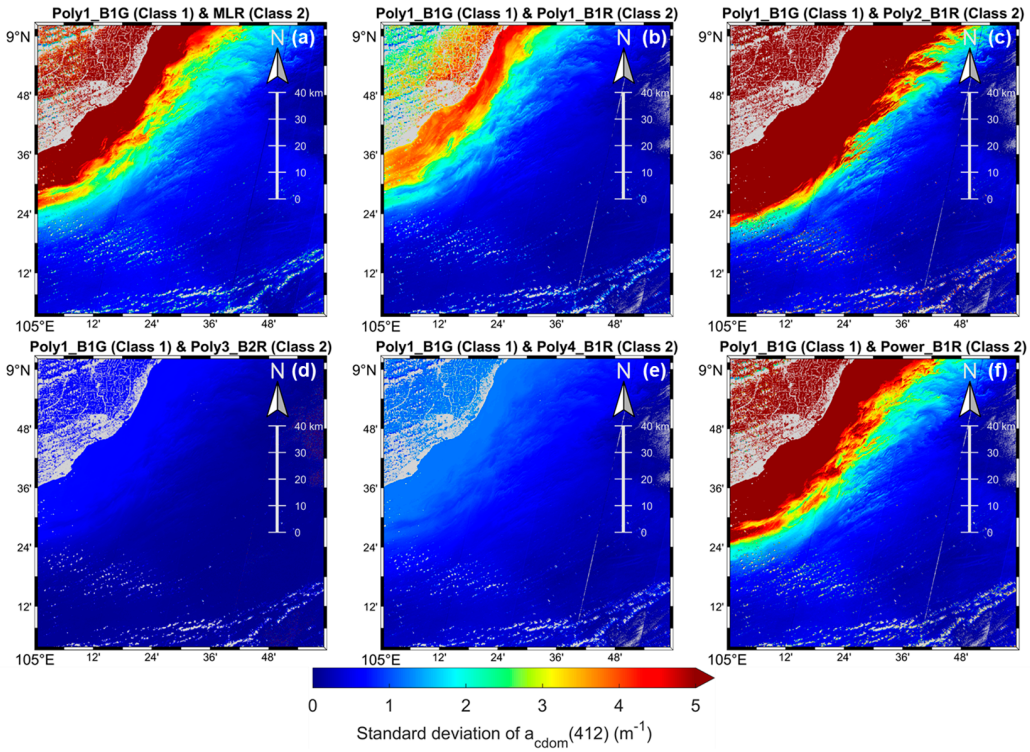


Figure 8. The MSI maps of the standard deviation of modeled a_{cdom} (412) at a 60 m spatial resolution for the algorithm combinations listed in Table 3. The name of the algorithm for each water class in each combination is displayed at the top of each individual map from (**a**–**f**).

3.5. Summary of the Different Acdom (412) Algorithms for OLI and MSI

The selected algorithms (referred to as AquaCDOM), based on the two validation steps performed on the DSM-V and Mu-CDOM datasets, and using the noise propagation exercise are provided in Table 4. The final evaluation of these algorithms for OLI and MSI on DSM-V shows that a_{cdom} (412) can be retrieved with an MAPD of 30% and a very low median bias value (between -0.0026 m^{-1} and -0.0035 m^{-1}) (see Figure 9).

Table 4. Equations of the composite inverse algorithms used to derive a_{cdom} (412) from OLI and MSI imagery for Class 1 and Class 2 waters.

	OLI	MSI
Class 1	$\log_{10} a_{cdom}(412) = p1.\log_{10} X + p2$ with $X = \frac{R_{rs}(443)}{R_{rs}(560)}$, $p1 = -1.516$, $p2 = -0.897$	$\log_{10} a_{cdom}(412) = p1.\log_{10} X + p2$ with $X = \frac{R_{rs}(443)}{R_{rs}(560)}$, $p1 = -1.481$, $p2 = -0.918$
Class 2	$\log_{10} a_{cdom}(412) = p1.\log_{10} X^4 + p2.\log_{10} X^3 + p3.\log_{10} X^2 + p4.\log_{10} X + p5$ with $X = \frac{R_{rs}(443)}{R_{rs}(655)}$, $p1 = -0.173$, $p2 = -0.0232$, $p3 = 0.385$, $p4 = -0.956$, $p5 = -0.440$	$\log_{10} a_{cdom}(412) = p1.\log_{10} X^3 + p2.\log_{10} X^2 + p3.\log_{10} X + p4$ with $X = \frac{R_{rs}(490)}{R_{rs}(665)}$, $p1 = 0.400$, $p2 = -0.142$, $p3 = -1.138$, $p4 = -0.0876$

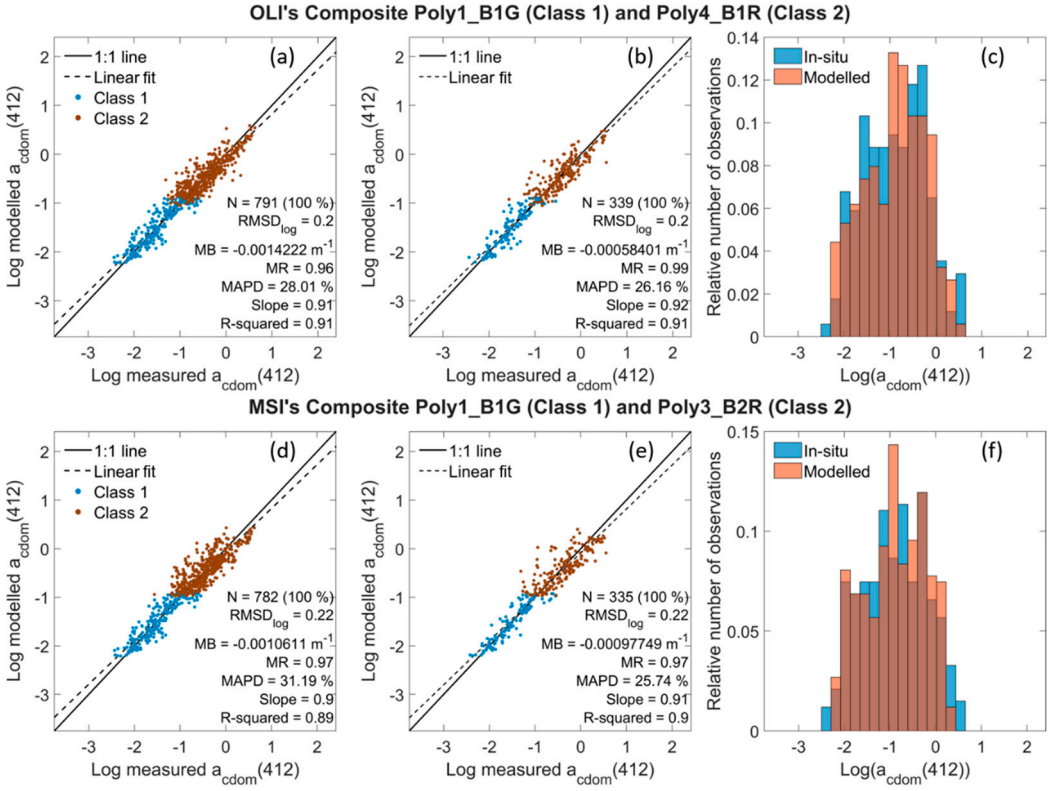


Figure 9. Comparison of AquaCDOM-derived and measured values of a_{cdom} (412) for OLI (top row) and MSI (bottom row) from the DSM-D (panels (**a**,**d**)) and DSM-V (panels (**b**,**e**)) datasets. Histograms of the derived and measured a_{cdom} (412) values from DSM-V for OLI and MSI are shown in (**c**,**f**), respectively. The different statistical indicators calculated for the data of AquaCDOM-derived and measured a_{cdom} (412) are provided (see text for details). The percentage of retrieved data points is indicated in brackets. The solid line is the 1:1 line. The different colors stand for different water classes as indicated.

4. Discussion

4.1. Comparison with Existing Algorithms

The performance of our algorithm is compared with two recently developed algorithms for estimating a_{cdom} (λ) from Landsat-8/OLI and Sentinel-2/MSI observations. The choice of these algorithms is motivated by the fact that they are not regional (i.e., they can be applied in all coastal water environments) and are based on different methodologies. The first one is the Semi-Analytical algorithm with a Virtual-band Estimator (SAVE), which estimates the R_{rs} (412) band from the other visible bands, to derive, from a modified version of the Quasi-Analytical Algorithm (QAA) [65], a_{cdom} (412) and other inherent optical properties [33]. The second algorithm is a machine learning algorithm using the Mixed Density Network approach (MDN) developed by Pahlevan et al. (2020), making it possible to derive, among other parameters, a_{cdom} (440) [29]. Since our final algorithms were developed for a_{cdom} (412), we re-developed them using the same methodology to estimate a_{cdom} (440) for comparison with the MDN model's output (see Appendix A for the coefficients of AquaCDOM for a_{cdom} (440)). This inter-comparison exercise was performed using DSM-V, Mu-CDOM, and the sensitivity analysis to atmospheric correction errors. Note that the a_{cdom} (λ) range of variability of the datasets used for the development of SAVE and NDM cover the ones of DSM-V and Mu-CDOM.

4.1.1. Comparison Between AquaCDOM and SAVE

For both OLI (Figure 10) and MSI (Figure 11), the two algorithms can estimate a_{cdom} (412) with a similar performance on DSM-V with, however, a slightly better retrieval in the higher range of a_{cdom} (412) by AquaCDOM, which also retrieves 100% of data points against about 96% to 93% for SAVE. The better performance of AquaCDOM is reinforced on the match-up dataset, especially for the higher range of a_{cdom} (412) values which are underestimated by SAVE for both OLI and MSI. Note however that SAVE was not developed for MSI which has slightly different bands, especially in the red (i.e., 10 nm of difference).

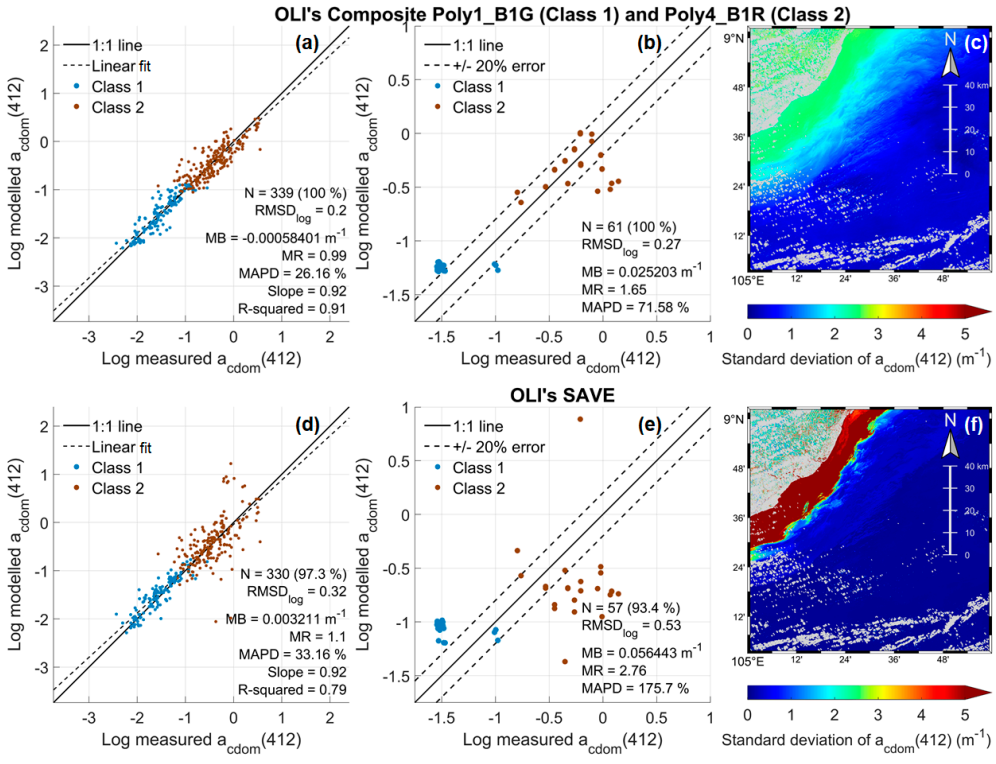


Figure 10. Comparison of AquaCDOM-derived and measured values of a_{cdom} (412) for OLI from the (**a**) DSM-V and (**b**) Mu-CDOM datasets. (**c**) Map of the standard deviation of the AquaCDOM-derived

a_{cdom} (412) values. Comparison of SAVE-derived and measured values of a_{cdom} (412) for OLI from the (**d**) DSM-V and (**e**) Mu-CDOM datasets. (**f**) Map of the standard deviation of the SAVE-derived a_{cdom} (412) values. The different statistical indicators calculated for the data of model-derived versus measured a_{cdom} (412) values are provided (see text for details). The percentage of retrieved data points is indicated in brackets. The solid line is the 1:1 line. The different colors stand for different water classes as indicated.

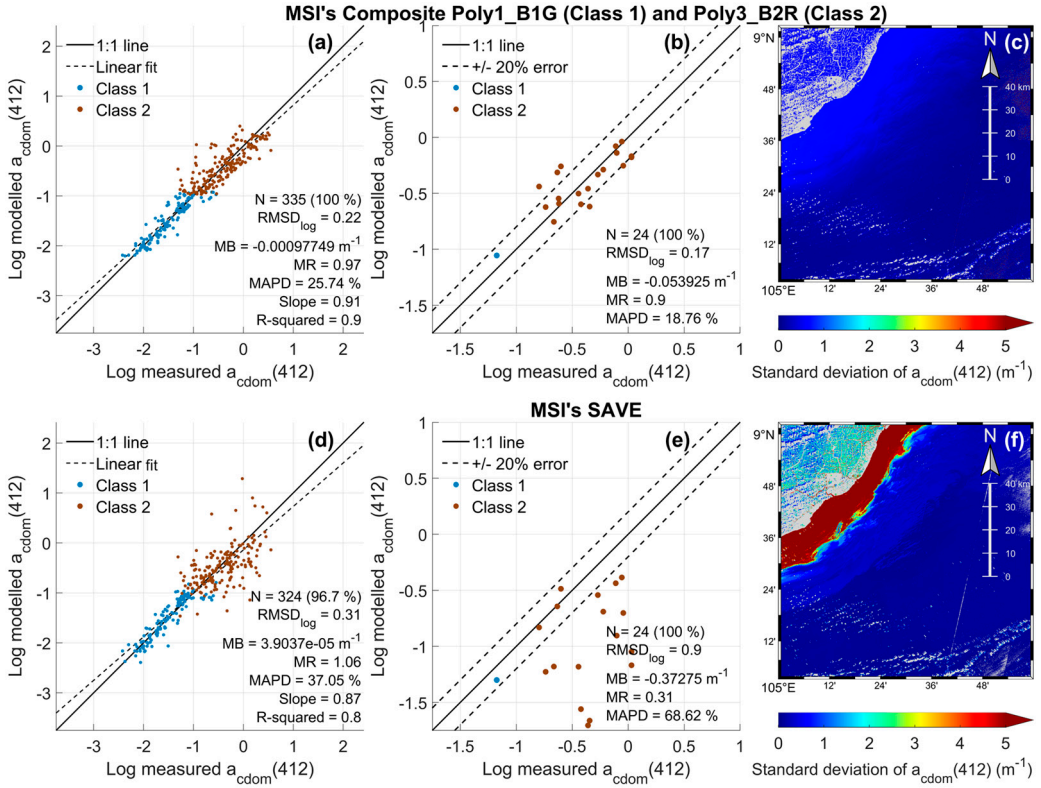


Figure 11. Same as Figure 10 but for MSI.

The uncertainty maps (Figures 10c and 11c) indicate that our algorithm has lower sensitivity to noise than SAVE over nearshore water characterized by high CDOM concentrations (standard deviation < 2.5 m^{-1} for AquaCDOM and ≥5 m^{-1} for SAVE). This is particularly notable for MSI (Figure 10f). Over open water, the two algorithms have about the same sensitivity to noise, with however a slightly lower standard deviation for SAVE (difference < 0.5 m^{-1}) when applied to OLI. Additionally, there is a pronounced separation between nearshore and offshore waters using SAVE despite using the same algorithm for the whole image, indicating possible artificial boundaries over the contrasted water area. The lower performance of SAVE compared to AquaCDOM on both the match-up exercise and noise sensitivity analysis could be due to the use in SAVE of a virtual 412 nm spectral band, estimated from the available R_{rs} (λ) spectral shape, which could be impacted by atmospheric correction. The use of band ratios in AquaCDOM can however mitigate the noise introduced by atmospheric correction.

4.1.2. Comparison Between AquaCDOM and MDN

The same inter-comparison exercise was performed between AquaCDOM and MDN for OLI (Figure 12) and MSI (Figure 13). Very similar retrieval performances are achieved by the two algorithms on the DSM-V and Mu-CDOM datasets for OLI. A slight overestimation of a_{cdom} (440) by MDN can however be observed for a_{cdom} (440) values lower than about 0.05 m^{-1} because this range of values is scarcely represented in the in situ dataset used to train the MDN neural networks. With regard to their respective sensitivity to atmospheric corrections, the two algorithms applied to OLI display a spatially even

distribution of the standard deviation of a_{cdom} (440), with a more pronounced sensitivity of MDN (standard deviation around 1.6 m^{-1}) compared to AquaCDOM (standard deviation around 0.8 m^{-1}). Note that despite using the same raw R_{rs} (λ) bands for OLI as SAVE, MDN is much less sensitive to AC uncertainties supporting the previous notion that SAVE's virtual 412 nm spectral band could increase its sensitivity to AC uncertainties. The same results are obtained on MSI with, however, two differences. First, the number of retrieved data points using MDN for MSI is lowered by 8.4% for DSM-V, and, more importantly, MDN seems to be much more sensitive to atmospheric correction in the nearshore waters. While such a pattern was not observed from OLI, it may be due to the different SNR between these two sensors over water areas [66,67]. In addition, MDN uses more bands within the 400 nm to 800 nm range for MSI compared to only the visible range for OLI. The additional NIR bands are non-negligible for inland and coastal waters and carry more AC error than the visible bands [55].

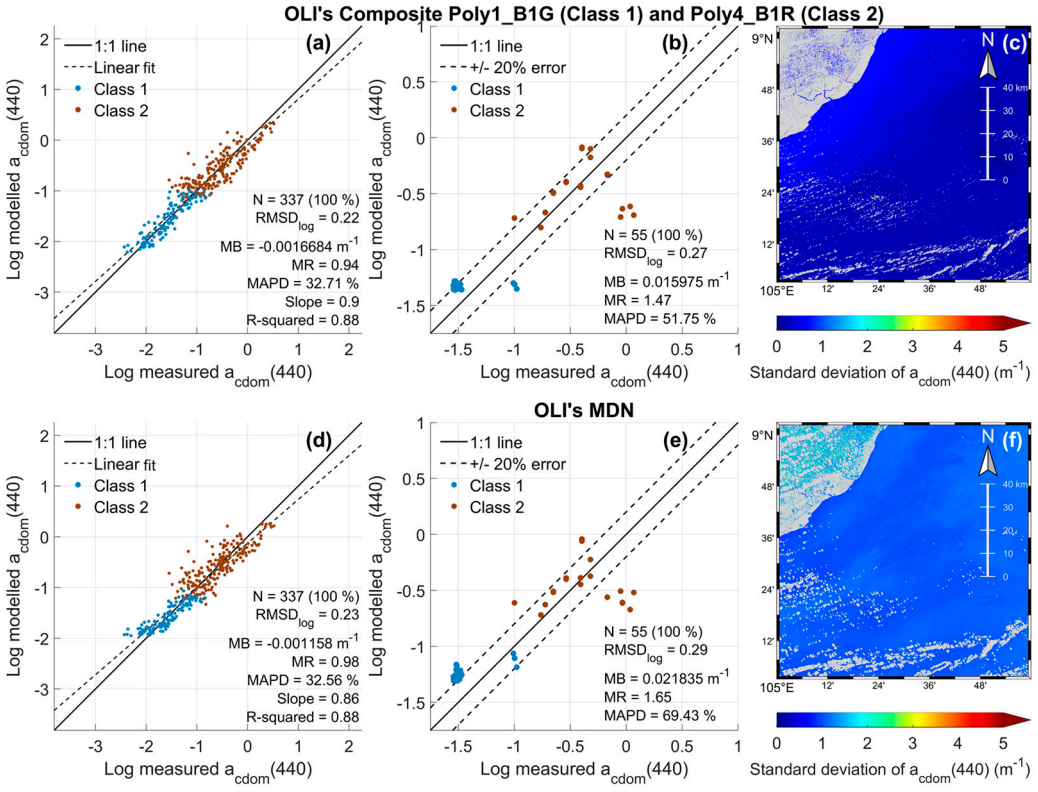


Figure 12. Comparison of AquaCDOM-derived and measured values of a_{cdom} (440) for OLI from the (**a**) DSM-V and (**b**) Mu-CDOM datasets. (**c**) Map of the standard deviation of the AquaCDOM-derived a_{cdom} (440) values. Comparison of MDN-derived and measured values of a_{cdom} (440) for OLI from the (**d**) DSM-V and (**e**) Mu-CDOM datasets. (**f**) Map of the standard deviation of the MDN-derived a_{cdom} (440) values. The different statistical indicators calculated for the data of model-derived versus measured a_{cdom} (440) are provided (see text for details). The percentage of retrieved data points is indicated in brackets. The solid line is the 1:1 line. The different colors stand for different water classes as indicated.

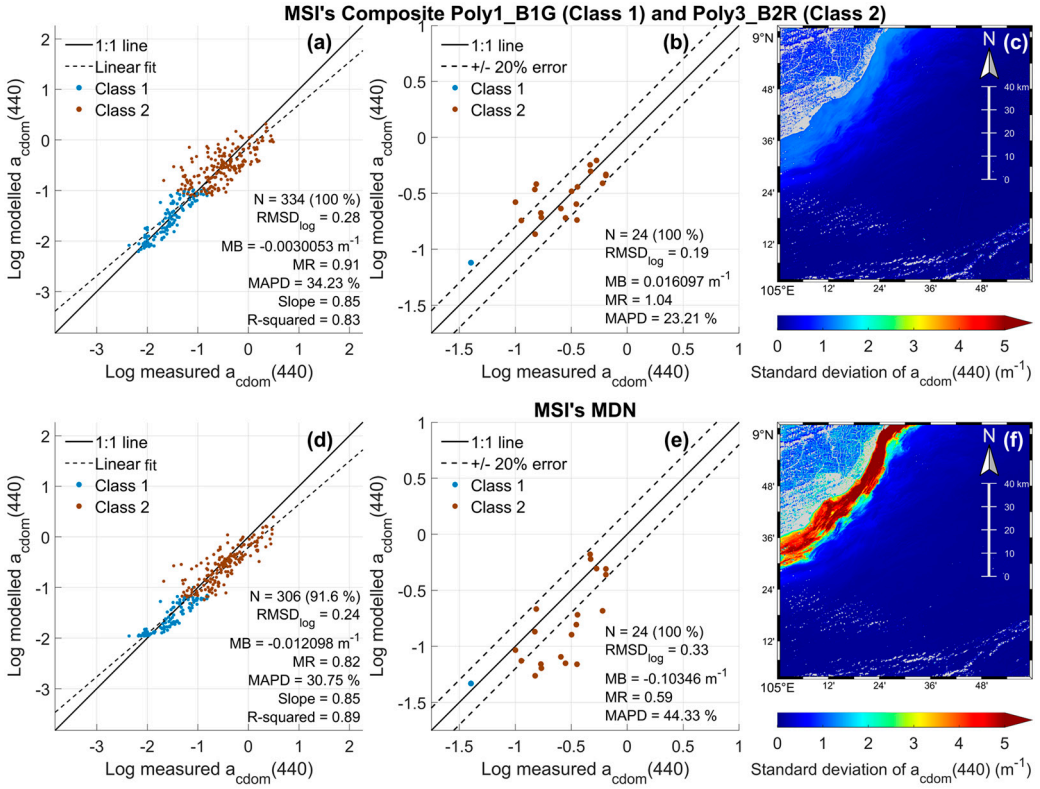


Figure 13. Same as Figure 12 but for MSI.

4.2. Landsat-8/Sentinel-2 a_{cdom} (412) Calibration

The temporal analysis of a_{cdom} (412) over a selected area from both OLI and MSI requires inter-calibration of their a_{cdom} (412) products. For that purpose, a global sample of 11 near-simultaneous- nadir-captured cloud-free image pairs of Landsat-8 and Sentinel-2 observations were selected covering different ranges of latitude and solar zenith angle. Since the selected area presents relatively high CDOM values, this inter-calibration exercise is only performed on the Class-2 algorithms. The a_{cdom} (412) values derived from OLI are here chosen as a reference because OLI has longer operation time and, therefore, is more advantageous for historical time series analysis. The selection process of the pixels to perform correlation analysis between OLI and MSI is as follows:

1. The overlapping subsets of the Landsat-8/Sentinel-2 image pair were extracted. This step ensures the complete overlap of the image pair.
2. The subset image was divided into blocks 990 m × 990 m in size. This step ensures the data selected for the correlation analysis (Figure 14) are sufficiently different to avoid an artificially strong correlation.
3. Blocks with zero invalid pixels were selected, i.e., the block is free from land or cloud interference.
4. For each block, the centers, 3-by-3 pixels and 9-by-9 pixels of OLI and MSI, respectively, were extracted to calculate the match-up data between the two sensors. The different sizes ensure that the extraction was carried out for the same area.

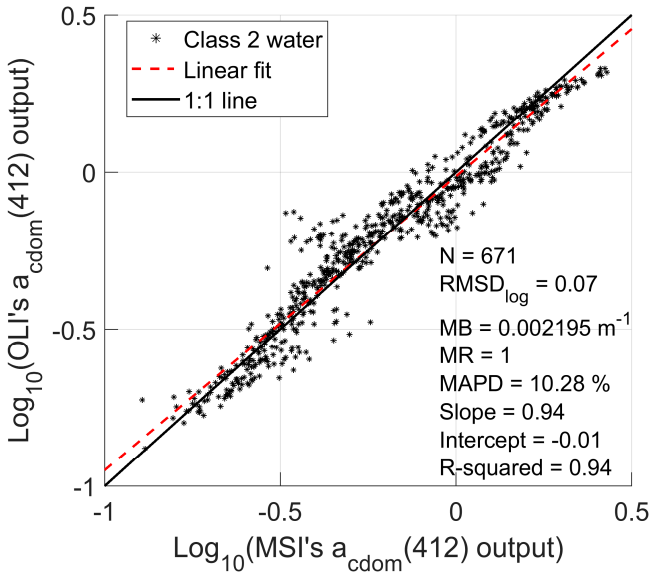


Figure 14. Comparison of OLI-derived and MSI-derived a_{cdom} (412) values over the 11 near-simultaneous acquisition scenes. The different statistical indicators calculated for the data of OLI-derived versus MSI-derived a_{cdom} (412) are provided.

From the cloud-free overlapping pixels of the 11 image pairs, the relationship between the two sensors' Class 2 a_{cdom} (412) outputs is shown in Figure 14. This relationship is used to convert a_{cdom} (412) measured by MSI to an a_{cdom} (412) measured by OLI. Over the whole Class 2 a_{cdom} (412) range, the outputs of both sensors are very similar (R^2 = 0.95). The MAPD, $RMSD_{log}$, and bias values are of 10.3%, 0.07, and 0.002 m^{-1}, respectively. This consistency could be attributed to the application of the TOA radiance gains, as well as to the application of the same atmospheric corrections. The correction equation is $y = 10^{0.94 \times \log_{10} x - 0.01}$, where x is MSI's a_{cdom} (412) output and y is OLI's. It should be considered that this correction was derived using a small global sample of coastal zones, and the method described could be applied to a more regional scale for better efficiency.

4.3. Application to a Case Study in the Mekong Delta, Vietnam

We applied AquaCDOM to analyze the temporal variability of a_{cdom} (412) along a transect in the Mekong Delta, Vietnam, which spans from the outlet of the Ganh Hao River to offshore waters. The variability was analyzed in comparison with the monthly accumulated rainfall and the monthly mean intraday high and low tide height difference; all of which are shown in Figure 15. The rainfall data were taken from the remote sensing-derived NASA's Global Precipitation Measurements (GPMs) (https://gpm.nasa.gov/ (accessed on 23 August 2022)), while the tide data come from the modeled tidal calendars of Vietnam (published yearly by the Vietnam Administration of Seas and Islands). The a_{cdom} (412) values peak during the first few months of each year and spread up to about 20 km offshore. Relatively low a_{cdom} (412) values were observed for the rest of the year, especially from April 2019 to December 2019. A second maximum, of much lower intensity, was also observed in October when the Mekong River outflow was at its maximum [68]. Interestingly, these extrema of a_{cdom} (412) are not in phase with the monthly rainfall stressing that the a_{cdom} (412) variability at the outlet during this time period is not driven by meteorological factors, specifically from land washing by rainfall. However, the CDOM content during this period may also be due to dilution processes during the rainy season. Besides this general seasonal pattern, other environmental forcings also act on the regional CDOM dynamics. Tidal forcing explains the export of CDOM offshore in January 2019 when the rainfall was low. Specifically, the higher gap observed between high and low tides during this period results in a stronger tidal export. Other exceptional events can also be identified, emphasizing the impact of human pressure on CDOM distribution. The CDOM temporal

pattern, in this area characterized by large aquaculture activities [69], matches the stocking period of shrimp ponds, which are located along the Ganh Hao River near the coast. During the stocking time (typically from February to May), the pond water is generally replaced before the installation of new shrimp batches, and the treated wastewater, which contains high organic matter loading, is discarded into the river. In 2020, COVID-19 disrupted the global distribution of aquaculture goods forcing global farmers to keep their harvest or sell at a low price [70]. In addition to being affected by COVID-19, the Mekong River Delta in Vietnam suffered a drought at the beginning of 2020, which is visible in the precipitation plot. This forced the farmers to only partially stock their ponds waiting for better weather conditions and market stability [71]. This could explain the second peak of a_{cdom} (412) in March 2020 despite low rainfall and weak tidal forcing, while only one was visible in 2019.

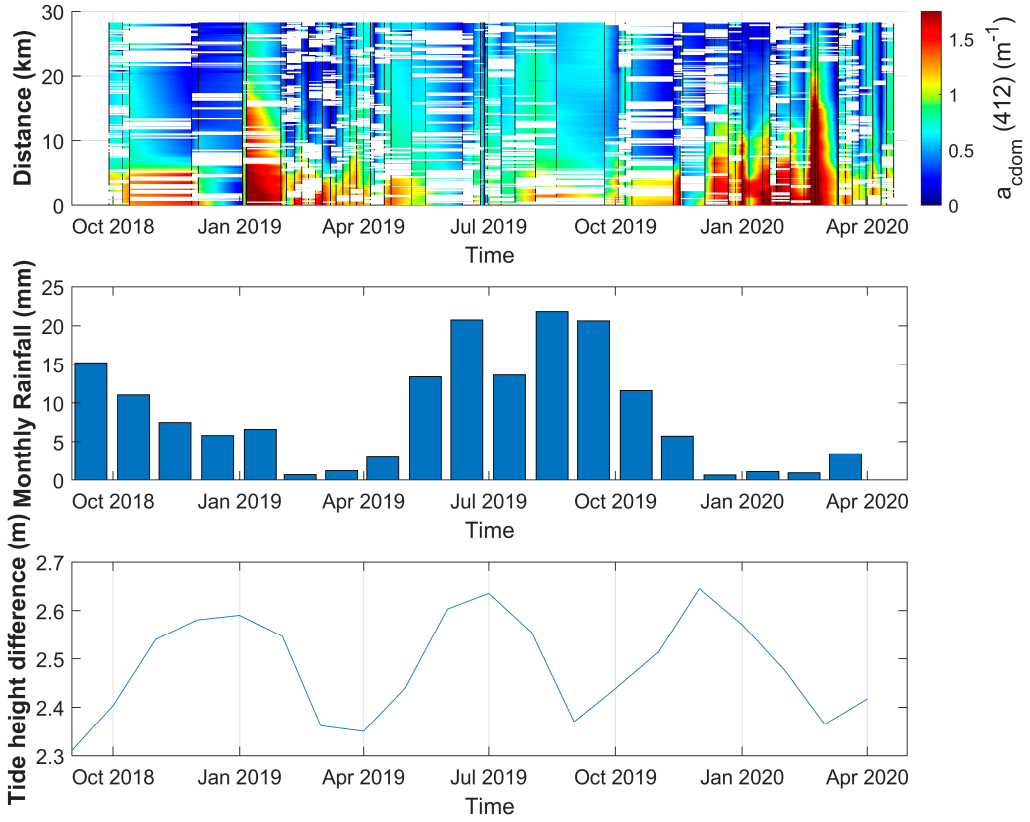


Figure 15. (**a**) Temporal variability of a_{cdom} (412) in Ganh Hao transect (Figure 2) in relation to the (**b**) monthly accumulated rainfall (in mm) and (**c**) the monthly mean intraday high and low tide height difference (in m).

5. Concluding Remarks

Two a_{cdom} (412) inverse empirical models were developed for OLI and MSI using a mixed dataset of in situ and synthetic a_{cdom} (λ) and R_{rs} (λ) data. AquaCDOM utilizes a simple water classification scheme and empirical formulations for each water class. Note that these parameterizations have also been adapted to assess a_{cdom} (440). Due to the overlapping of the higher range of a_{cdom} (λ) values used for the development of the Class 1 algorithm, with the lower range of a_{cdom} (λ) values used for the development of the Class 2 algorithm, no artificial discontinuity is observed on the inversed a_{cdom} (λ) spatial patterns (see Figure A3). For both sensors, the input parameters of the algorithm are R_{rs} (443), R_{rs} (560), and R_{rs} (665) with the addition of R_{rs} (490) for MSI. The new model can assess a_{cdom} (412) with an MAPD of 30% and a negligible bias of 0.003 m^{-1} for both sensors. The AquaCDOM retrieval accuracy was compared with that obtained from a semi-analytical model (SAVE) and a machine learning approach (MDN). While all models showed similar retrieval accuracy, the new model is much less sensitive to noise

introduced by atmospheric correction. This sensitivity analysis performed in the frame of this study paves the way for a new metric to be applied in the selection process of a given inverse bio-optical algorithm. Since all the spectral bands used in AquaCDOM are available on OLCI onboard Sentinel-3, this algorithm is potentially applicable to this sensor. A preliminary test performed on standard OLCI $R_{rs}(\lambda)$ measurements collected over coastal water where in situ a_{cdom} (412) measurements are available (see Figure A1 for the locations) demonstrates that a_{cdom} (412) can be estimated with a MAPD of 44% (see Figure A2). These numbers should be confirmed based on a more extensive match-up dataset. The application of AquaCDOM in the time series analysis of a water transect in the south of Vietnam shows good CDOM variability that matches the changes in local shrimp farming activities related to the COVID-19 pandemic that started at the end of 2019. The temporal analysis demonstrated the potential of combining Landsat-8/MSI and MSI to monitor, at high spatial resolution (around 20 m) and enhanced temporal coverage, the changes in coastal environments.

Author Contributions: Conceptualization, V.S.N., H.L. and V.V.; methodology, V.S.N., V.V. and H.L.; formal analysis, V.S.N. and H.L.; in situ data collection, X.M.; satellite data processing, V.S.N.; writing—original draft preparation, V.S.N., H.L., D.L.T. and V.V.; supervision, H.L. and D.L.T.; project administration, H.L.; funding acquisition, H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Institute of Research and Development (IRD) ARTS program (Ref. PDG-MApS-2019-465) and the French Embassy in Hanoi Excellence Scholarship, as well as by the ANR CO2COAST (Grant Nos. ANR-20-CE01-0021) projects.

Data Availability Statement: The database used in this study encompasses the SeaBASS dataset (https://seabass.gsfc.nasa.gov/wiki/NOMAD accessed on 21 May 2021), the LOG's dataset (included in GLORIA dataset: https://doi.org/10.1594/PANGAEA.948492 accessed on 15 March 2024).

Acknowledgments: We gratefully acknowledge all scientists and supporting personnel involved in the collection, processing, and dissemination of the in situ and satellite data used in this study as well as all agencies (ESA and NASA) that provided support for these activities.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Equations of the composite inverse algorithms used to derive a_{cdom} (440) from OLI and MSI imagery for Class 1 and Class 2 waters.

	OLI	MSI
Class 1	$\log_{10} a_{cdom}(440) = p1.\log_{10} X + p2$ with $X = \frac{R_{rs}(443)}{R_{rs}(560)}$, $p1 = -1.370$, $p2 = -1.012$	$\log_{10} a_{cdom}(440) = p1.\log_{10} X + p2$ with $X = \frac{R_{rs}(443)}{R_{rs}(560)}$, $p1 = -1.420$, $p2 = -0.988$
Class 2	$\log_{10} a_{cdom}(440) = p1.\log_{10} X^4 + p2.\log_{10} X^3 + p3.\log_{10} X^2 + p4.\log_{10} X + p5$ with $X = \frac{R_{rs}(443)}{R_{rs}(655)}$, $p1 = -0.807$, $p2 = 0.201$, $p3 = 0.845$, $p4 = -1.04$, $p5 = -0.608$	$\log_{10} a_{cdom}(440) = p1.\log_{10} X^3 + p2.\log_{10} X^2 + p3.\log_{10} X + p4$ with $X = \frac{R_{rs}(490)}{R_{rs}(665)}$, $p1 = 0.137$, $p2 = 0.113$, $p3 = -1.102$, $p4 = -0.255$

Appendix B

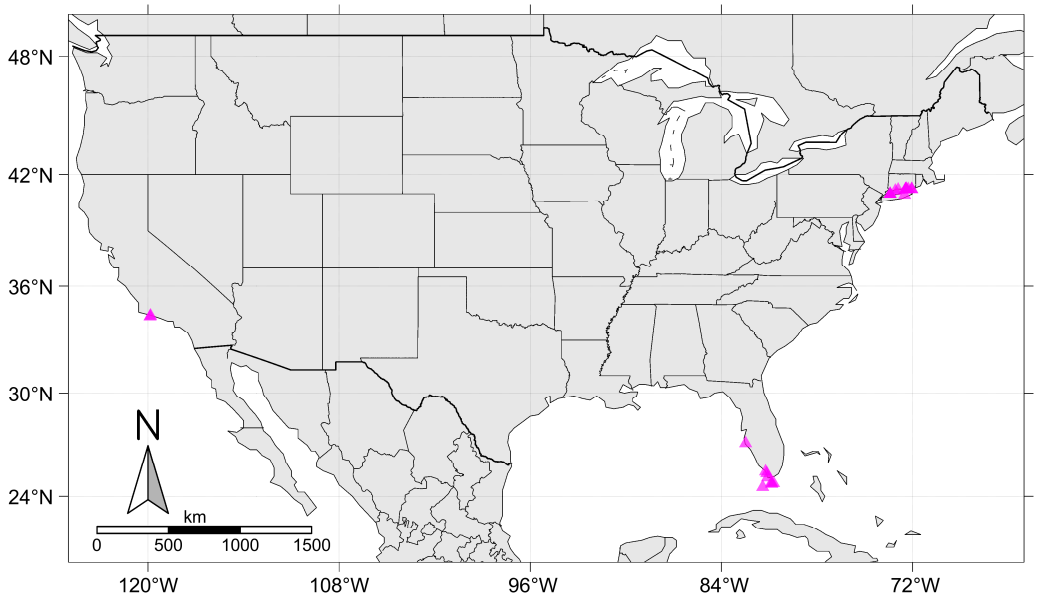


Figure A1. Location of the a_{cdom} (412) match-up data points (purple triangles) for OLCI on Sentinel-3. In situ a_{cdom} (412) measurements were extracted from SeaBASS.

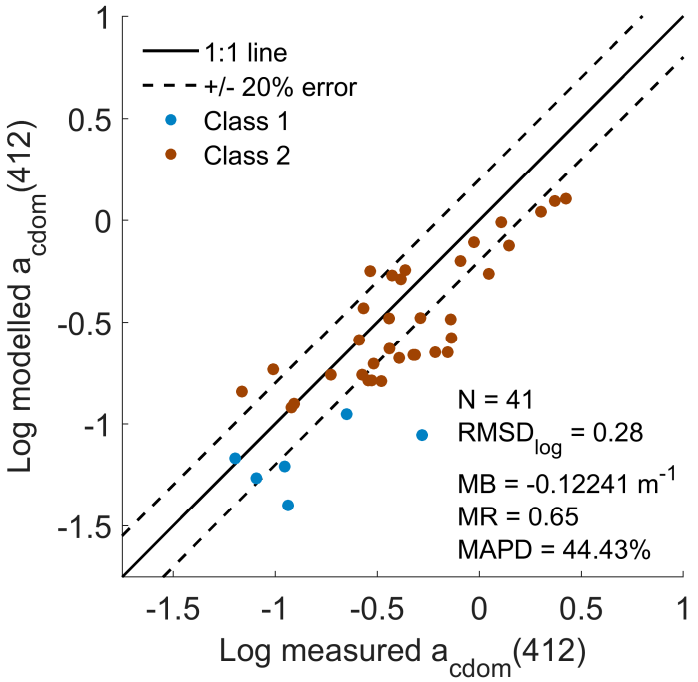


Figure A2. Comparison of AquaCDOM-derived and measured values of a_{cdom} (412) for the OLCI match-up dataset presented in Figure A1.

Appendix C

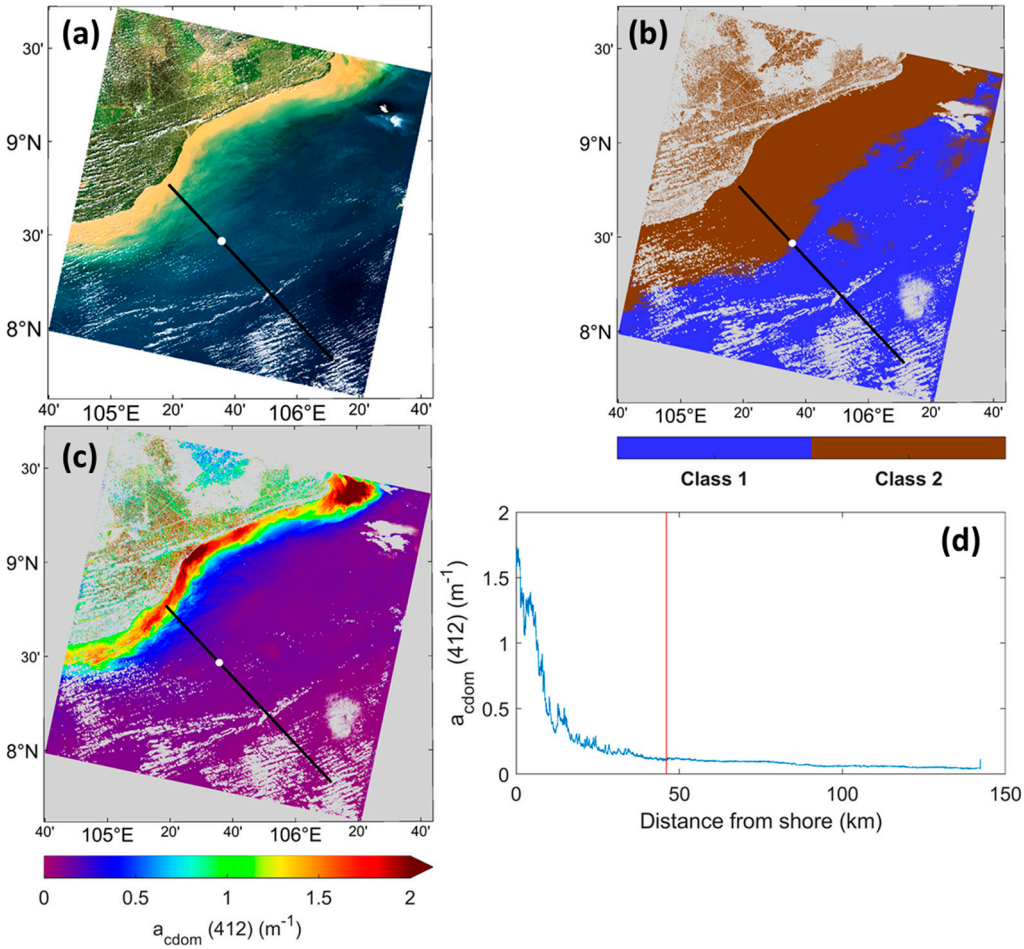


Figure A3. (**a**) An RGB composite of an OLI image over a Vietnam coastal area. The black line represents the transect along which the a_{cdom} (412) value estimated from AquaCDOM is extracted. The location of the boundary between the two water classes is shown by a white dot. (**b**) The spatial distribution of the two water classes used in this study. (**c**) The a_{cdom} (412) spatial distribution estimated from the AquaCDOM algorithm. (**d**) The a_{cdom} (412) values along the cross-shore transect. The red line shows the spatial delimitation between the two water classes along the cross-shore transect.

References

1. Coble, P.G. Marine Optical Biogeochemistry: The Chemistry of Ocean Color. *Chem. Rev.* **2007**, *107*, 402–418. [CrossRef] [PubMed]
2. 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]
3. Kieber, R.J.; Whitehead, R.F.; Reid, S.N.; Willey, J.D.; Seaton, P.J. Chromophoric Dissolved Organic Matter (CDOM) In Rainwater, Southeastern North Carolina, USA. *J. Atmos. Chem.* **2006**, *54*, 21–41. [CrossRef]
4. Wilson, H.F.; Xenopoulos, M.A. Effects of Agricultural Land Use on the Composition of Fluvial Dissolved Organic Matter. *Nat. Geosci.* **2008**, *2*, 37–41. [CrossRef]
5. Hartnett, H.E. Dissolved Organic Matter (DOM). In *Encyclopedia of Earth Sciences Series*; Springer: Cham, Switzerland, 2018; pp. 375–378. [CrossRef]
6. Fork, M.L.; Karlsson, J.; Sponseller, R.A. Dissolved Organic Matter Regulates Nutrient Limitation and Growth of Benthic Algae in Northern Lakes through Interacting Effects on Nutrient and Light Availability. *Limnol. Oceanogr. Lett.* **2020**, *5*, 417–424. [CrossRef]
7. Wymore, A.S.; Rodríguez-Cardona, B.; McDowell, W.H. Understanding Dissolved Organic Matter Biogeochemistry Through In Situ Nutrient Manipulations in Stream Ecosystems. *J. Vis. Exp.* **2016**, *2016*, 54704. [CrossRef]
8. Chen, X.; Liu, J.; Chen, J.; Wang, J.; Xiao, X.; He, C.; Shi, Q.; Li, G.; Jiao, N. Oxygen Availability Driven Trends in DOM Molecular Composition and Reactivity in a Seasonally Stratified Fjord. *Water Res.* **2022**, *220*, 118690. [CrossRef]

9. Coble, P.G. Characterization of Marine and Terrestrial DOM in Seawater Using Excitation-Emission Matrix Spectroscopy. *Mar. Chem.* **1996**, *51*, 325–346. [CrossRef]
10. Massicotte, P.; Asmala, E.; Stedmon, C.; Markager, S. Global Distribution of Dissolved Organic Matter along the Aquatic Continuum: Across Rivers, Lakes and Oceans. *Sci. Total Environ.* **2017**, *609*, 180–191. [CrossRef]
11. Keith, D.J.; Yoder, J.A.; Freeman, S.A. Spatial and Temporal Distribution of Coloured Dissolved Organic Matter (CDOM) in Narragansett Bay, Rhode Island: Implications for Phytoplankton in Coastal Waters. *Estuar. Coast. Shelf Sci.* **2002**, *55*, 705–717. [CrossRef]
12. Zhang, Y.; Zhou, Y.; Shi, K.; Qin, B.; Yao, X.; Zhang, Y. Optical Properties and Composition Changes in Chromophoric Dissolved Organic Matter along Trophic Gradients: Implications for Monitoring and Assessing Lake Eutrophication. *Water Res.* **2018**, *131*, 255–263. [CrossRef] [PubMed]
13. Bonelli, A.G.; Vantrepotte, V.; Jorge, D.S.F.; Demaria, J.; Jamet, C.; Dessailly, D.; Mangin, A.; Fanton d'Andon, O.; Kwiatkowska, E.; Loisel, H. Colored Dissolved Organic Matter Absorption at Global Scale from Ocean Color Radiometry Observation: Spatio-Temporal Variability and Contribution to the Absorption Budget. *Remote Sens. Environ.* **2021**, *265*, 112637. [CrossRef]
14. Mannino, A.; Novak, M.G.; Hooker, S.B.; Hyde, K.; Aurin, D. Algorithm Development and Validation of CDOM Properties for Estuarine and Continental Shelf Waters along the Northeastern U.S. Coast. *Remote Sens. Environ.* **2014**, *152*, 576–602. [CrossRef]
15. Cao, F.; Tzortziou, M.; Hu, C.; Mannino, A.; Fichot, C.G.; del Vecchio, R.; Najjar, R.G.; Novak, M. Remote Sensing Retrievals of Colored Dissolved Organic Matter and Dissolved Organic Carbon Dynamics in North American Estuaries and Their Margins. *Remote Sens. Environ.* **2018**, *205*, 151–165. [CrossRef]
16. Mannino, A.; Russ, M.E.; Hooker, S.B. Algorithm Development and Validation for Satellite-Derived Distributions of DOC and CDOM in the U.S. Middle Atlantic Bight. *J. Geophys. Res. Ocean.* **2008**, *113*, 7051. [CrossRef]
17. Loisel, H.; Vantrepotte, V.; Dessailly, D.; Mériaux, X. Assessment of the Colored Dissolved Organic Matter in Coastal Waters from Ocean Color Remote Sensing. *Opt. Express* **2014**, *22*, 13109–13124. [CrossRef]
18. Lee, Z.; Hu, C.; Arnone, R.; Liu, Z. Impact of Sub-Pixel Variations on Ocean Color Remote Sensing Products. *Opt. Express* **2012**, *20*, 20844–20854. [CrossRef]
19. Pahlevan, N.; Chittimalli, S.K.; Balasubramanian, S.v.; Vellucci, V. Sentinel-2/Landsat-8 Product Consistency and Implications for Monitoring Aquatic Systems. *Remote Sens. Environ.* **2019**, *220*, 19–29. [CrossRef]
20. Osadchiev, A.; Sedakov, R. Spreading Dynamics of Small River Plumes off the Northeastern Coast of the Black Sea Observed by Landsat 8 and Sentinel-2. *Remote Sens. Environ.* **2019**, *221*, 522–533. [CrossRef]
21. Li, P.; Chen, S.; Ji, H.; Ke, Y.; Fu, Y. Combining Landsat-8 and Sentinel-2 to Investigate Seasonal Changes of Suspended Particulate Matter off the Abandoned Distributary Mouths of Yellow River Delta. *Mar. Geol.* **2021**, *441*, 106622. [CrossRef]
22. Li, J.; Roy, D.P.; Atzberger, C.; Zhou, G. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sens.* **2017**, *9*, 902. [CrossRef]
23. 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]
24. Chen, J.; Zhu, W.N.; Tian, Y.Q.; Yu, Q. Estimation of Colored Dissolved Organic Matter from Landsat-8 Imagery for Complex Inland Water: Case Study of Lake Huron. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2201–2212. [CrossRef]
25. Xu, J.; Fang, C.; Gao, D.; Zhang, H.; Gao, C.; Xu, Z.; Wang, Y. Optical Models for Remote Sensing of Chromophoric Dissolved Organic Matter (CDOM) Absorption in Poyang Lake. *ISPRS J. Photogramm. Remote Sens.* **2018**, *142*, 124–136. [CrossRef]
26. Mabit, R.; Araújo, C.A.S.; Singh, R.K.; Bélanger, S. Empirical Remote Sensing Algorithms to Retrieve SPM and CDOM in Québec Coastal Waters. *Front. Remote Sens.* **2022**, *3*, 834908. [CrossRef]
27. Al-Kharusi, E.S.; Tenenbaum, D.E.; Abdi, A.M.; Kutser, T.; Karlsson, J.; Bergström, A.K.; Berggren, M. Large-Scale Retrieval of Coloured Dissolved Organic Matter in Northern Lakes Using Sentinel-2 Data. *Remote Sens.* **2020**, *12*, 157. [CrossRef]
28. Ruescas, A.B.; Hieronymi, M.; Mateo-Garcia, G.; Koponen, S.; Kallio, K.; Camps-Valls, G. Machine Learning Regression Approaches for Colored Dissolved Organic Matter (CDOM) Retrieval with S2-MSI and S3-OLCI Simulated Data. *Remote Sens.* **2018**, *10*, 786. [CrossRef]
29. Pahlevan, N.; Smith, B.; Schalles, J.; Binding, C.; Cao, Z.; Ma, R.; Alikas, K.; Kangro, K.; Gurlin, D.; Hà, N.; et al. Seamless Retrievals of Chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in Inland and Coastal Waters: A Machine-Learning Approach. *Remote Sens. Environ.* **2020**, *240*, 111604. [CrossRef]
30. 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]
31. 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]
32. 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]
33. Wei, J.; Lee, Z.P.; 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]
34. Wang, Y.; Shen, F.; Sokoletsky, L.; Sun, X. Validation and Calibration of QAA Algorithm for CDOM Absorption Retrieval in the Changjiang (Yangtze) Estuarine and Coastal Waters. *Remote Sens.* **2017**, *9*, 1192. [CrossRef]

35. Ylöstalo, P.; Kallio, K.; Seppälä, J. Absorption Properties of In-Water Constituents and Their Variation among Various Lake Types in the Boreal Region. *Remote Sens. Environ.* **2014**, *148*, 190–205. [CrossRef]
36. Shanmugam, P. New Models for Retrieving and Partitioning the Colored Dissolved Organic Matter in the Global Ocean: Implications for Remote Sensing. *Remote Sens. Environ.* **2011**, *115*, 1501–1521. [CrossRef]
37. IOCCG. *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications*; Lee, Z.-P., Ed.; Reports of the International Ocean Colour Coordinating Group; IOCCG: Dartmouth, NS, Canada, 2006; Volume 5.
38. Bricaud, A.; Morel, A.; Prieur, L. Absorption by Dissolved Organic Matter of the Sea (Yellow Substance) in the UV and Visible Domains. *Limnol. Oceanogr.* **1981**, *26*, 43–53. [CrossRef]
39. Jamet, C.; Loisel, H.; Dessailly, D. Retrieval of the Spectral Diffuse Attenuation Coefficient $K_d(\lambda)$ in Open and Coastal Ocean Waters Using a Neural Network Inversion. *J. Geophys. Res. Ocean.* **2012**, *117*, C10023. [CrossRef]
40. Vantrepotte, V.; Danhiez, F.-P.; Loisel, H.; Ouillon, S.; Meriaux, X.; Cauvin, A.; Dessailly, D. CDOM-DOC Relationship in Contrasted Coastal Waters: Implication for DOC Retrieval from Ocean Color Remote Sensing Observation. *Opt. Express* **2015**, *23*, 33–54. [CrossRef]
41. Loisel, H.; Mériaux, X.; Berthon, J.F.; Poteau, A. Investigation of the Optical Backscattering to Scattering Ratio of Marine Particles in Relation to Their Biogeochemical Composition in the Eastern English Channel and Southern North Sea. *Limnol. Oceanogr.* **2007**, *52*, 739–752. [CrossRef]
42. Lubac, B.; Loisel, H.; Guiselin, N.; Astoreca, R.; Artigas, L.F.; Mériaux, X. Hyperspectral and Multispectral Ocean Color Inversions to Detect Phaeocystis Globosa Blooms in Coastal Waters. *J. Geophys. Res. Ocean.* **2008**, *113*, 6026. [CrossRef]
43. Mograne, M.; Jamet, C.; Loisel, H.; Vantrepotte, V.; Mériaux, X.; Cauvin, A. Evaluation of Five Atmospheric Correction Algorithms over French Optically-Complex Waters for the Sentinel-3A OLCI Ocean Color Sensor. *Remote Sens.* **2019**, *11*, 668. [CrossRef]
44. Massicotte, P.; Babin, M.; Fell, F.; Fournier-Sicre, V.; Doxaran, D. The Coastal Surveillance Through Observation of Ocean Color (COASTlOOC) Dataset. *Earth Syst. Sci. Data* **2023**, *15*, 3529–3545. [CrossRef]
45. Kim, J.; Kim, Y.; Park, S.E.; Kim, T.H.; Kim, B.G.; Kang, D.J.; Rho, T.K. Impact of Aquaculture on Distribution of Dissolved Organic Matter in Coastal Jeju Island, Korea, Based on Absorption and Fluorescence Spectroscopy. *Environ. Sci. Pollut. Res.* **2022**, *29*, 553–563. [CrossRef] [PubMed]
46. Vanhellemont, Q. Sensitivity Analysis of the Dark Spectrum Fitting Atmospheric Correction for Metre- and Decametre-Scale Satellite Imagery Using Autonomous Hyperspectral Radiometry. *Opt. Express* **2020**, *28*, 29948–29965. [CrossRef]
47. Vanhellemont, Q. Adaptation of the Dark Spectrum Fitting Atmospheric Correction for Aquatic Applications of the Landsat and Sentinel-2 Archives. *Remote Sens. Environ.* **2019**, *225*, 175–192. [CrossRef]
48. Ngoc, D.D.; Loisel, H.; Jamet, C.; Vantrepotte, V.; Duforêt-Gaurier, L.; Minh, C.D.; Mangin, A. Coastal and Inland Water Pixels Extraction Algorithm (WiPE) from Spectral Shape Analysis and HSV Transformation Applied to Landsat 8 OLI and Sentinel-2 MSI. *Remote Sens. Environ.* **2019**, *223*, 208–228. [CrossRef]
49. Werdell, P.J.; Bailey, S.; Fargion, G.; Pietras, C.; Knobelspiesse, K.; Feidman, G.; Mcclain, C. Unique Data Repository Facilitates Ocean Color Satellite Validation. *Eos Trans. Am. Geophys. Union.* **2003**, *84*, 377–387. [CrossRef]
50. Caballero, I.; Steinmetz, F.; Navarro, G. Evaluation of the First Year of Operational Sentinel-2A Data for Retrieval of Suspended Solids in Medium- to High-Turbidity Waters. *Remote Sens.* **2018**, *10*, 982. [CrossRef]
51. Ilori, C.O.; Pahlevan, N.; Knudby, A. Analyzing Performances of Different Atmospheric Correction Techniques for Landsat 8: Application for Coastal Remote Sensing. *Remote Sens.* **2019**, *11*, 469. [CrossRef]
52. Qin, P.; Simis, S.G.H.; Tilstone, G.H. Radiometric Validation of Atmospheric Correction for MERIS in the Baltic Sea Based on Continuous Observations from Ships and AERONET-OC. *Remote Sens. Environ.* **2017**, *200*, 263–280. [CrossRef]
53. Jamet, C.; Loisel, H.; Kuchinke, C.P.; Ruddick, K.; Zibordi, G.; Feng, H. Comparison of Three SeaWiFS Atmospheric Correction Algorithms for Turbid Waters Using AERONET-OC Measurements. *Remote Sens. Environ.* **2011**, *115*, 1955–1965. [CrossRef]
54. Jiang, D.; Matsushita, B.; Pahlevan, N.; Gurlin, D.; Lehmann, M.K.; Fichot, C.G.; Schalles, J.; Loisel, H.; Binding, C.; Zhang, Y.; et al. Remotely Estimating Total Suspended Solids Concentration in Clear to Extremely Turbid Waters Using a Novel Semi-Analytical Method. *Remote Sens. Environ.* **2021**, *258*, 112386. [CrossRef]
55. Pahlevan, N.; Mangin, A.; Balasubramanian, S.V.; Smith, B.; Alikas, K.; Arai, K.; Barbosa, C.; Bélanger, S.; Binding, C.; Bresciani, M.; et al. ACIX-Aqua: A Global Assessment of Atmospheric Correction Methods for Landsat-8 and Sentinel-2 over Lakes, Rivers, and Coastal Waters. *Remote Sens. Environ.* **2021**, *258*, 112366. [CrossRef]
56. Loisel, H.; Stramski, D.; Dessailly, D.; Jamet, C.; Li, L.; Reynolds, R.A. An Inverse Model for Estimating the Optical Absorption and Backscattering Coefficients of Seawater From Remote-Sensing Reflectance Over a Broad Range of Oceanic and Coastal Marine Environments. *J. Geophys. Res. Ocean.* **2018**, *123*, 2141–2171. [CrossRef]
57. Tran, M.D.; Vantrepotte, V.; Loisel, H.; Oliveira, E.N.; Tran, K.T.; Jorge, D.; Mériaux, X.; Paranhos, R. Band Ratios Combination for Estimating Chlorophyll-a from Sentinel-2 and Sentinel-3 in Coastal Waters. *Remote Sens.* **2023**, *15*, 1653. [CrossRef]
58. Pahlevan, N.; Lee, Z.; Wei, J.; Schaaf, C.B.; Schott, J.R.; Berk, A. On-Orbit Radiometric Characterization of OLI (Landsat-8) for Applications in Aquatic Remote Sensing. *Remote Sens. Environ.* **2014**, *154*, 272–284. [CrossRef]
59. Vanhellemont, Q.; Ruddick, K. Atmospheric Correction of Metre-Scale Optical Satellite Data for Inland and Coastal Water Applications. *Remote Sens. Environ.* **2018**, *216*, 586–597. [CrossRef]
60. Vanhellemont, Q.; Ruddick, K. Advantages of High Quality SWIR Bands for Ocean Colour Processing: Examples from Landsat-8. *Remote Sens. Environ.* **2015**, *161*, 89–106. [CrossRef]

61. Vanhellemont, Q.; Ruddick, K. Turbid Wakes Associated with Offshore Wind Turbines Observed with Landsat 8. *Remote Sens. Environ.* **2014**, *145*, 105–115. [CrossRef]
62. Zorrilla, N.A.; Vantrepotte, V.; Ngoc, D.D.; Huybrechts, N.; Gardel, A. Automated SWIR Based Empirical Sun Glint Correction of Landsat 8-OLI Data over Coastal Turbid Water. *Opt. Express* **2019**, *27*, A294–A318. [CrossRef]
63. Steinmetz, F.; Ramon, D. Sentinel-2 MSI and Sentinel-3 OLCI Consistent Ocean Colour Products Using POLYMER. In Proceedings of the Conference on Remote Sensing of the Open and Coastal Ocean and Inland Waters, Honolulu, HI, USA, 24–25 September 2018; Volume 10778, pp. 46–55. [CrossRef]
64. Ramon, D.; Steinmetz, F.; Deschamps, P.-Y. Atmospheric Correction in Presence of Sun Glint: Application to MERIS. *Opt. Express* **2011**, *19*, 9783–9800. [CrossRef]
65. Lee, Z.; Carder, K.L.; Arnone, R.A. Deriving Inherent Optical Properties from Water Color: A Multiband Quasi-Analytical Algorithm for Optically Deep Waters. *Appl. Opt.* **2002**, *41*, 5755–5772. [CrossRef] [PubMed]
66. Pahlevan, N.; Schott, J.R.; Franz, B.A.; Zibordi, G.; Markham, B.; Bailey, S.; Schaaf, C.B.; Ondrusek, M.; Greb, S.; Strait, C.M. Landsat 8 Remote Sensing Reflectance (Rrs) Products: Evaluations, Intercomparisons, and Enhancements. *Remote Sens. Environ.* **2017**, *190*, 289–301. [CrossRef]
67. Pahlevan, N.; Sarkar, S.; Franz, B.A.; Balasubramanian, S.V.; He, J. Sentinel-2 MultiSpectral Instrument (MSI) Data Processing for Aquatic Science Applications: Demonstrations and Validations. *Remote Sens. Environ.* **2017**, *201*, 47–56. [CrossRef]
68. Loisel, H.; Mangin, A.; Vantrepotte, V.; Dessailly, D.; Ngoc Dinh, D.; Garnesson, P.; Ouillon, S.; Lefebvre, J.P.; Mériaux, X.; Minh Phan, T. Variability of Suspended Particulate Matter Concentration in Coastal Waters under the Mekong's Influence from Ocean Color (MERIS) Remote Sensing over the Last Decade. *Remote Sens. Environ.* **2014**, *150*, 218–230. [CrossRef]
69. FAO. Viet Nam—National Aquaculture Sector Overview. Available online: https://www.fao.org/fishery/en/countrysector/naso_vietnam (accessed on 23 July 2024).
70. FAO. *The Effect of COVID-19 on Fisheries and Aquaculture in Asia*; FAO: Rome, Italy, 2020.
71. Trung, K. Triển Khai Kế Hoạch Ngành Tôm Năm 2020: Nhiều Thách Thức và Cơ Hội. Available online: https://www.baobaclieu.vn/nong-nghiep-nong-dan-nong-thon/trien-khai-ke-hoach-nganh-tom-nam-2020-nhieu-thach-thuc-va-co-hoi-64831.html (accessed on 14 October 2022).

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.