remote sensing

MDPI

Article

Ocean Colour Atmospheric Correction for Optically Complex Waters under High Solar Zenith Angles: Facilitating Frequent Diurnal Monitoring and Management

Yongquan Wang [1,2,3,4], Huizeng Liu [1,2,*], Zhengxin Zhang [1,3], Yanru Wang [1,3,5], Demei Zhao [1,3], Yu Zhang [1,2], Qingquan Li [1,2,3] and Guofeng Wu [1,3]

1 MNR Key Laboratory for Geo-Environmental Monitoring of Great Bay Area, Guangdong Key Laboratory of Urban Informatics, Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China; wangyongquan2021@email.szu.edu.cn (Y.W.); 2200325023@email.szu.edu.cn (Z.Z.); wangyr@szu.edu.cn (Y.W.); 2050322003@email.szu.edu.cn (D.Z.); yuzhang21@szu.edu.cn (Y.Z.); liqq@szu.edu.cn (Q.L.); guofeng.wu@szu.edu.cn (G.W.)
2 Institute for Advanced Study, Tiandu-Shenzhen University Deep Space Exploration Joint Laboratory, Space Science Center, Shenzhen University, Shenzhen 518060, China
3 School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518060, China
4 Light & Life Laboratory, Department of Physics, Stevens Institute of Technology, Hoboken, NJ 07030, USA
5 Shenzhen Marine Development & Promotion Center, Shenzhen 518034, China
* Correspondence: huizeng.liu@szu.edu.cn

Abstract: Accurate atmospheric correction (AC) is one fundamental and essential step for successful ocean colour remote-sensing applications. Currently, most ACs and the associated ocean colour remote-sensing applications are restricted to solar zenith angles (SZAs) lower than 70°. The ACs under high SZAs present degraded accuracy or even failure problems, rendering the satellite retrievals of water quality parameters more challenging. Additionally, the complexity of the bio-optical properties of the coastal waters and the presence of complex aerosols add to the difficulty of AC. To address this challenge, this study proposed an AC algorithm based on extreme gradient boosting (XGBoost) for optically complex waters under high SZAs. The algorithm presented in this research has been developed using pairs of Geostationary Ocean Colour Imager (GOCI) high-quality noontime remote-sensing reflectance (Rrs) and the Rayleigh-corrected reflectance (ρ_{rc}) derived from the Ocean Colour–Simultaneous Marine and Aerosol Retrieval Tool (OC-SMART) in the morning (08:55 LT) and at dusk (15:55 LT). The algorithm was further examined using the daily GOCI images acquired in the morning and at dusk, and the hourly (total suspended sediment) TSS concentration was also obtained based on the atmospherically corrected GOCI data. The results showed that: (i) the model produced an accurate fitting performance ($R^2 \geq 0.90$, RMSD ≤ 0.0034 sr^{-1}); (ii) the model had a high validation accuracy with an independent dataset (R^2 = 0.92–0.97, MAPD = 8.2–26.81% and quality assurance (QA) score = 0.9–1); and (iii) the model successfully retrieved more valid Rrs for GOCI images under high SZAs and enhanced the accuracy and coverage of TSS mapping. This algorithm has great potential to be applied to AC for optically complex waters under high SZAs, thus increasing the frequency of available observations in a day.

Keywords: atmospheric correction; ocean colour; optically complex waters; high solar zenith angles

Citation: Wang, Y.; Liu, H.; Zhang, Z.; Wang, Y.; Zhao, D.; Zhang, Y.; Li, Q.; Wu, G. Ocean Colour Atmospheric Correction for Optically Complex Waters under High Solar Zenith Angles: Facilitating Frequent Diurnal Monitoring and Management. *Remote Sens.* **2024**, *16*, 183. https://doi.org/10.3390/rs16010183

Academic Editor: Chung-Ru Ho

Received: 2 November 2023
Revised: 28 December 2023
Accepted: 29 December 2023
Published: 31 December 2023

1. Introduction

The ocean and coastal environments are subject to constant damage due to improper anthropogenic activities and natural factors, such as reclamation, sewage discharge, extreme weather and more [1,2]. Therefore, it is crucial to monitor ocean water to ensure the protection of the ocean and coastal aquatic ecosystems. Instant and reliable techniques are crucial in capturing the dynamic characteristics of oceanic and coastal ecosystems. Although field surveys are typically employed to address the issue, the methodologies

incur significant economic and temporal expenses. Field surveys for the management of ocean and coastal ecosystems may be restricted by real-time constraints in capturing ocean process characteristics, physical challenges in the field and the limitations of time, cost and labour [3]. Moreover, a few field survey points could not provide sufficient representation of large-scale oceanic changes [4]. To address these limitations, a combination of ocean colour remote-sensing data and field surveys has been employed as a complementary data source since the late 1980s [5]. Through synergistically combining remote sensing and in situ measurements, researchers could surmount the limitations specific to each approach, enabling a more comprehensive monitoring and management strategy for the marine ecosystem [6–8].

Ocean colour remote sensing provides long-term data on the distribution and dynamics of coastal waters in a cost-effective manner, far exceeding the spatio-temporal limitations of traditional in situ measurements. Many researchers are focusing on the use of ocean colour data from low-altitude polar orbiting satellites ocean colour data to monitor the aquatic ecological environment, including water quality and harmful algal blooms [9,10], estimating chlorophyll-a concentrations and particulate organic carbon concentrations [11,12], managing fisheries resources [13], which provide valuable information for ocean management, and understanding covariations between the ocean and coastal environment and productivity. However, one polar-orbiting ocean colour satellite can typically observe once or twice per day over most of the global region, making it difficult to adequately monitor the highly dynamic water column in areas such as inland and estuarine coastal zones. Compared with polar orbiting satellites, geostationary ocean colour satellites demonstrate the unique advantages of monitoring the diurnal variations in the materials in the upper ocean with a high sampling frequency [14]. More than 80% of the total radiation received by ocean colour sensors comes from atmospheric sources, while the water-leaving radiation that actually carries ocean colour information only accounts for about 10% [15]. The atmospheric correction (AC) is a crucial data-processing procedure that eliminates interfering signals from the atmospheric path radiance [16].

The accuracy of ocean colour AC has a direct impact on the subsequent retrieval of ocean colour elements and biogeochemical parameters [17]. However, AC often fails under high solar zenith angles (SZAs) [18], limiting the potential application of ocean colour satellite data in the aquatic environment monitoring to the noontime. The performance of standard AC algorithms is significantly influenced by the SZA [19]. Under high SZAs ($\geq$70°), the current AC algorithms usually fail to produce accurate ocean colour data products. This is primarily attributed to the inadequate consideration of the Earth's curvature effect in calculating Rayleigh radiance [18] and insufficient illumination [20,21]. Therefore, it is essential to develop appropriate AC models for ocean colour remote sensing under high SZAs. For examples, Li et al. [22–24] developed several AC models for Geostationary Ocean Colour Imager (GOCI) data and polar-orbiting satellite sensors for high-latitude seas in winter based on neural networks for the open ocean under high SZAs. Another challenge of AC is its applications over turbid coastal and inland waters. In the open oceans, the classical "black pixel" assumption is commonly employed to estimate aerosol scattering contributions. This assumption assumes that the water-leaving radiance or Rrs is zero in the near-infrared (NIR) band. Due to the complex influences of both the water constituents and aerosol conditions in coastal regions, particularly the presence of absorbing aerosols [25–27], the "black pixel" assumption at the near-infrared bands often fails in coastal areas [28]. Therefore, standard AC algorithms yield biased water-leaving radiances and overestimate the aerosol contribution [29,30], resulting in the failure of AC. To address the abovementioned problems, several AC algorithms for optically complex waters have been developed. One such algorithm is the spectral matching algorithm, which utilises the complete spectrum available and employs an iterative optimisation scheme to separate the radiometric contribution of water from the atmospheric and surface contributions [31]. Additionally, artificial neural network-based algorithms with powerful nonlinear fitting capabilities have also been applied to ocean colour AC [32–34].

Geostationary ocean colour satellites encounter a high-SZA observation environment when detecting optically complex coastal and inland water bodies in the morning or at dusk. Due to AC's imperfect performance or failure, the successful retrieval of remote-sensing reflectance (Rrs) and water quality parameter products from GOCI images for complex water bodies under high SZAs is challenging. In winter, the SZA for GOCI data obtained in the morning or at dusk usually exceeds 70°, so it cannot be effectively processed by standard AC algorithms or GOCI data-processing systems (GDPS). Figure 1 illustrates the GOCI hourly Raleigh corrected reflectance (ρ_{rc}) and Rrs obtained from the Korea Ocean Satellite Centre (KOSC) on 13 January 2021 at 15:55 local time (LT) (UTC+09:00, as adopted below). Notably, no ρ_{rc} values are retrieved when the SZA is higher than 80°, and few valid Rrs values are available in the meantime. Therefore, it is highly important to develop an effective AC algorithm for optically complex waters under high SZAs.

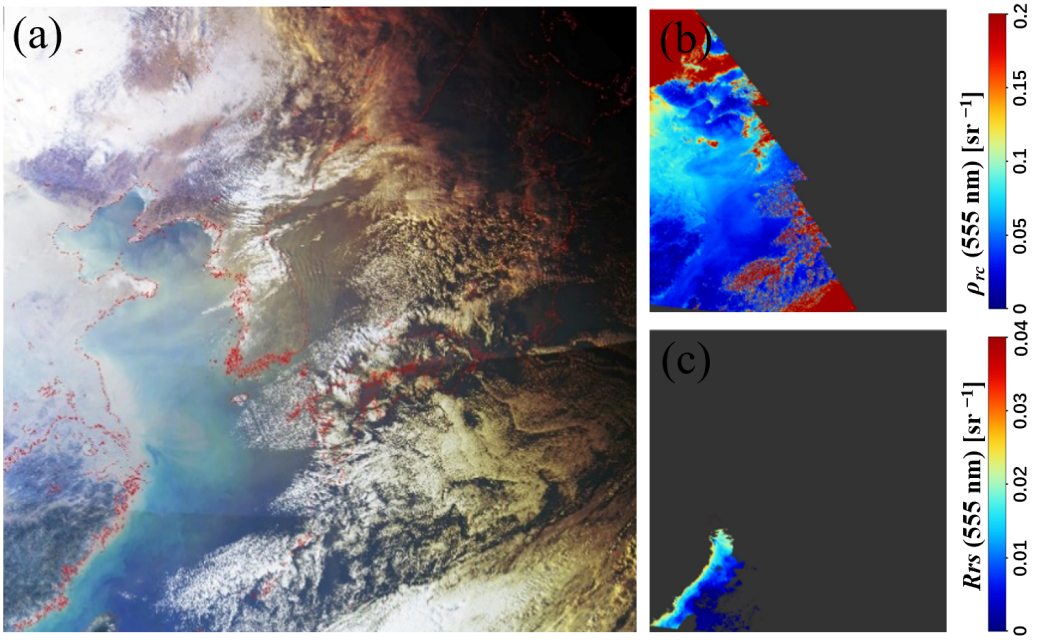

Figure 1. GOCI hourly RGB image (**a**), ρ_{rc} (555 nm) (**b**), and the corresponding Rrs (555 nm) (**c**) provided by the KOSC on 15:55 LT, 13 January 2021.

In this study, a novel AC algorithm (denoted as XGBAC) for optically complex waters under high SZAs was developed based on the extreme gradient boosting (XGBoost). Previous studies have demonstrated machine-learning capabilities in ocean colour AC. For example, based on the neural network, Li et al. [24] proposed a new AC algorithm for GOCI data under high SZA conditions in the open oceans, and Fan et al. [35,36] developed an Ocean Colour–Simultaneous Marine and Aerosol Retrieval Tool (OC-SMART) to retrieve ocean colour products for coastal and clear waters. Although these methods have been effectively applied in turbid Case 2 waters under low SZAs or in clear waters under high SZAs, the conditions of optically complex waters under high SZAs are not considered. Therefore, the new XGBAC model was established to retrieve the Rrs considering the optical complex water under high SZA by training a large diverse matchup dataset between GOCI high-quality noontime Rrs and ρ_{rc} derived from OC-SMART in the morning and at dusk. Evaluation datasets were used to validate the XGBAC algorithm. Furthermore, the application of the XGBAC algorithm to GOCI data processing and estimation of the total suspended sediment (TSS) concentration from atmospherically corrected GOCI Rrs data in the coastal waters of the Yangtze River Estuary (YRE) and Hangzhou Bay (HZB) is also presented. The proposed algorithm could be helpful in accurately deriving Rrs from GOCI images with high SZAs, thus improving the monitoring of the diurnal variations in complex inland and coastal waters.

2. Data and Methods

2.1. Satellite Images

In this study, the GOCI images were used to develop and test the XGBAC model for optically complex waters under high SZAs. The GOCI is the world's first ocean colour satellite sensor in geostationary orbit, launched by KOSC in 2010 [37]. It covers the coasts of Eastern and Northern China, which are characterised by a variety of optically complex waters. With its high spatial resolution and hourly observations from morning to dusk, GOCI could monitor ocean colour in near real-time, and has been successfully applied in the short-term dynamic monitoring of oceanic phenomena, such as planktonic algae, suspended particulate matter and chlorophyll-a [14,38].

To develop the algorithm, hourly GOCI-L1B products were downloaded from the KOSC website (http://kosc.kiost.ac.kr/, accessed on 16 March 2023). To ensure that the training dataset was representative of high SZAs, a total of 17 days of winter satellite images were acquired between 2010 and 2021, with 6 images per day. Specifically, four images captured during the midday period (between 11:00 LT and 14:00 LT) were utilised to identify stable waters, while two images captured in the morning and at dusk (at 9:00 LT and 16:00 LT, respectively) were used to obtain matching data for model training. GOCI-L1B data were processed with the OC-SMART to obtain the SZA, viewing zenith angle (VZA), relative azimuth angle (RAA), Rrs, ρ_{rc} and quality control flags. OC-SMART is a high-performance multi-sensor data analysis model based on machine learning, supporting heritage, current and possible future multi-spectral and hyper-spectral sensors. It was developed using extensive radiative transfer simulations and a comprehensive dataset. And it can generate higher quality global ocean colour products than GDPS, including Rrs, chlorophyll-a concentrations and inherent optical properties, even under complex environmental conditions [36]. The SZA threshold of the OC-SMART was set to 70°–88° to mask the pixels with low and super-high SZAs. To test the model performance, hourly GOCI-L1B data on 13 January 2021 were applied to retrieve the Rrs in the coasts of Eastern and Northern China.

2.2. Development of XGBAC Algorithm

The XGBoost is a machine-learning algorithm based on a distributed gradient-boosted decision tree and widely used in regression, classification and ranking problems [39]. Compared with the neural network and random forest model, XGBoost works by fitting a learner to the training dataset and then adding a regularisation term to the loss function to prevent overfitting, which can be represented as follows:

$$\text{obj} = \sum_{i}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{t} \Omega(f_k) \quad (1)$$

where $l(y_i, \hat{y}_i)$ is the loss function term that measures the difference between the actual value y_i and prediction $\hat{y}_i$ for each sample i. $\Omega(f_k)$ is a regularisation term used to penalise the complexity of each model to avoid overfitting. In order to approximate this objective function in the XGBoost model, a second-order Taylor expansion of the loss function is performed.

XGBoost can provide good model interpretability and has demonstrated excellent capability in previous ocean colour studies [12,40]. Therefore, XGBoost was used to train the XGBAC models in this study. XGBoost model was implemented using the Scikit-learn library in Python (version 3.9.13) [41]. The input factors include the SZA, VZA, RAA and ρ_{rc} of seven GOCI bands (412, 443, 490, 555, 660, 680 and 745 nm), and the output layer comprises the Rrs values at these seven bands. The parameter-setting scheme was based on comparing the best results of different setting parameters, i.e., the booster is set to gbtree, the number of regression trees is 500, the maximum depth of the trees is 6, the learning rate is 0.1, the weight of L2 regularization term is 0.01 and the maximum number of iterations is 100. A model was developed for each of the seven bands, respectively.

2.3. Training Dataset Selection

High-quality training datasets are essential for the development of good machine-learning models. To ensure the quality of the Rrs data for the development of the XGBAC model, the following data filtering criteria were applied.

(1) Pixels with any of the following quality flags were excluded: unavailable satellite L1B reflectance due to either saturation or missing values (flag 1), solar or sensor viewing angle out of range (flag 4), land (flag 16), cloud (flag 64), ρ_{rc} out of scope (flag 56) or negative ρ_{rc} (flag 1024). The numbers in brackets represent the 32-bit flag value;

(2) The percentage of pixels with effective values in a 5 × 5 pixel frame was calculated. To avoid adjacency effects, only the data with a 100% effective amount were selected for the next step [42]; Only the pixel frames with a coefficient of variation (CV) below 0.15 were kept for subsequent analysis [24];

(3) The temporal stability of valid pixel values was checked using four noontime (11:00 LT–14:00 LT) observations within one day. Pixels with a CV of multiple observations below 0.15 were kept.

(4) After the high-quality Rrs data were extracted based on the above criteria, the Rrs values at noontime with low SZAs were matched to the ρ_{rc} values in the morning and at dusk with high SZAs. Within this time range, the water bodies were assumed to be relatively stable. Finally, a total of 3,088,523 matchups of both turbid and clear waters were obtained, of which 70% were randomly selected as the training dataset and the remaining 30% were used as the validation dataset. The SZA ranges from 70° to 88° within the matchup dataset (shown in Figure 2).

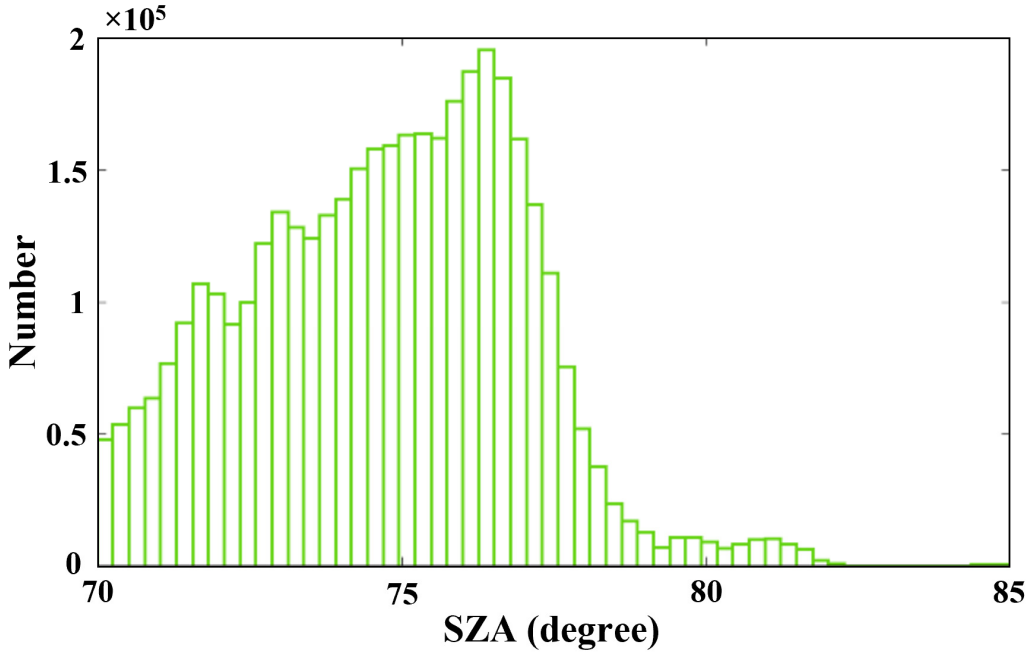

Figure 2. Frequency distribution of solar zenith angle in the matchup dataset.

2.4. Model Performance Evaluation

The model performance was evaluated with the validation dataset. Several metrics were used to determine the model performance, including the coefficient of determination (R^2), the mean absolute percentage deviation (MAPD), the mean relative percentage deviation (RPD) and the root mean square deviation (RMSD), which were calculated as follows:

$$R^2 = \frac{\sum_{i=1}^{N}(Y_i - \overline{X})^2}{\sum_{i=1}^{N}(X_i - \overline{X})^2} \quad (2)$$

$$\text{MAPD} = \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_i - X_i|}{X_i} \times 100\% \quad (3)$$

$$\text{RPD} = \frac{1}{N}\sum_{i=1}^{N}\frac{Y_i - X_i}{X_i} \times 100\% \quad (4)$$

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{N}(Y_i - X_i)^2}{N}} \quad (5)$$

where Y_i, X_i, $\overline{X}$ and N are the retrieved values, known values, the mean of known values and the sample number, respectively. To compare the performance of XGBAC and other popular models, the variations in the first quartile (Q1), median and third quartile (Q3) of the MAPD of Rrs obtained by the XGBAC and OC-SMART models were analysed and compared. Additionally, the Rrs retrievals at SZAs of up to 80°–88° from the XGBAC and OC-SMART were assessed with the quality assurance (QA) system [43]. The QA system employs a set of 23 optical water types with Rrs spectra, spanning from blue to yellow water, as references. By comparing the target Rrs with the reference Rrs, it assigns a score ranging from 0 to 1, with 0 indicating that the Rrs is unsuitable for use, and 1 denoting that the Rrs is in complete agreement with the reference spectra.

Furthermore, the XGBAC model was applied to retrieve Rrs products from GOCI data to verify its reliability and applicability. For the evaluation of the stability of Rrs retrieval, the Rrs (555) product in the morning and at dusk was generated by XGBAC and OC-SMART, respectively. The RPDs of Rrs retrieval values between the morning (08:55 LT)/dusk (15:55 LT) observation times and noontime (11:55 LT) were calculated, respectively. Based on the retrieved Rrs products after atmospheric correction by XGBAC, the TSS can be further derived. Because TSS is a crucial parameter for assessing the estuarine and coastal waters' quality, which is widely used in marine environmental management [44,45]. The YRE and HZB with turbid water were taken as an example, and the regional empirical TSS inversion algorithm from [14] was adopted as follows:

$$\begin{cases} TSS = 10^{1.075+1.1230\times Ratio}, \\ Ratio = \frac{\text{Rrs (745 nm)}}{\text{Rrs (490 nm)}} \end{cases} \quad (6)$$

3. Results

3.1. Performance of XGBAC Model

The calibration performance of the XGBAC model is shown in Figure 3. The models had a high fitting accuracy ($R^2 \geq 0.90$, RMSD ≤ 0.0034 sr^{-1}), with the Rrs values retrieved by XGBAC versus satellite Rrs values (referred to as reference values) scattering closely around the 1:1 line. Figure 4 shows scatterplots comparing the Rrs values in the validation dataset and the Rrs values derived by XGBAC at each GOCI band. The XGBAC-retrieved Rrs values were basically consistent with the reference values, with R^2 exceeding 0.92 and RMSD lower than 0.0033 sr^{-1} for validation datasets. The validation statistic results for XGBAC are summarised in Table 1. The XGBAC algorithm gave the best estimate for Rrs (490 nm) with an R^2 of 0.96, an RMSD of 0.0012, a MAPD of 8.2% and an RPD of −0.72%, respectively, followed by Rrs (443 nm) and Rrs (412 nm), while Rrs (745 nm) was found to perform the worst with an R^2 of 0.94, an RMSD of 0.0026 sr^{-1}, a MAPD of 26.81% and an RPD of −7.05%, respectively. It is noteworthy that the Rrs at 745 nm exhibited a low magnitude, resulting in relatively high MAPD and RPD values. The accuracy of the XGBAC model-derived Rrs (490 nm) was further evaluated and compared for three different SZA ranges (shown in Figure 5). A decreasing pattern of accuracy was observed with increasing SZAs (MAPD of 7.2% and 11.15% at the SZA of 70°–75° and 80°–90°, respectively). However, the XGBAC model still worked well at the SZA of up to 80°–88° with R^2 = 0.94 and RMSD = 0.0016 sr^{-1}. Overall, the Rrs values estimated by XGBAC exhibited good agreement with the reference values, and the small difference in statistical parameters between the training and validation datasets also indicates that the developed

XGBAC model can competently acquire knowledge from the training datasets, with a commendable level of stability and reliability.

Table 1. Statistic results of the validation performance for each GOCI band.

Parameters	R^2	RMSD (sr^{-1})	MAPD	RPD
412 nm	0.92	0.0010	9.29%	−0.89%
443 nm	0.94	0.0011	8.61%	−0.77%
490 nm	0.96	0.0013	8.2%	−0.72%
555 nm	0.97	0.0019	9.79%	−1.03%
660 nm	0.97	0.0032	19.2%	−3.71%
680 nm	0.96	0.0033	20.16%	−4.06%
745 nm	0.94	0.0026	26.81%	−7.05%

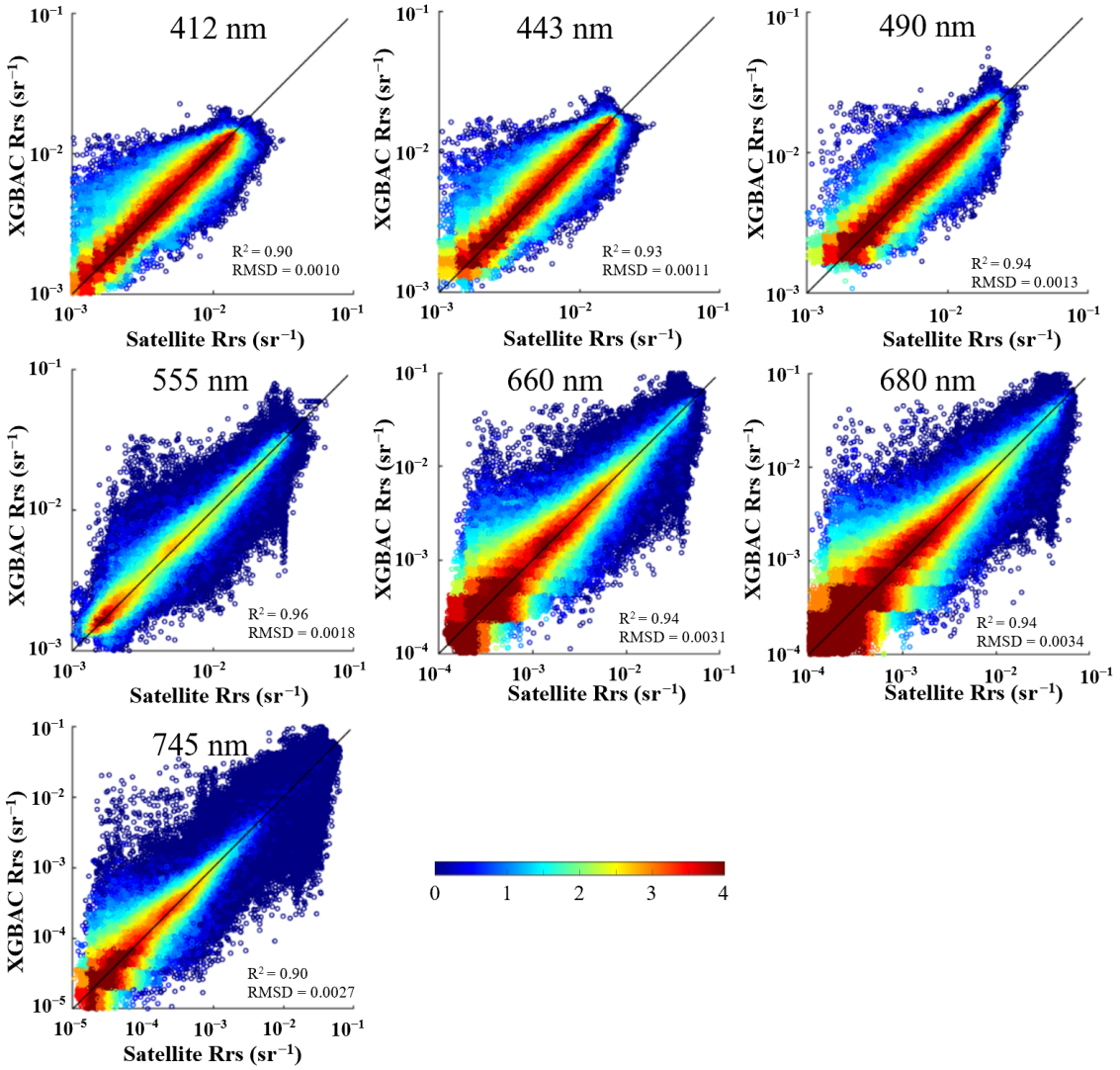

Figure 3. Scatterplots of Rrs (λ) retrieved by the XGBAC model vs. the reference values in the training dataset at each GOCI band. The numerical values along the colour bar correspond to the pixel density on a logarithmic scale.

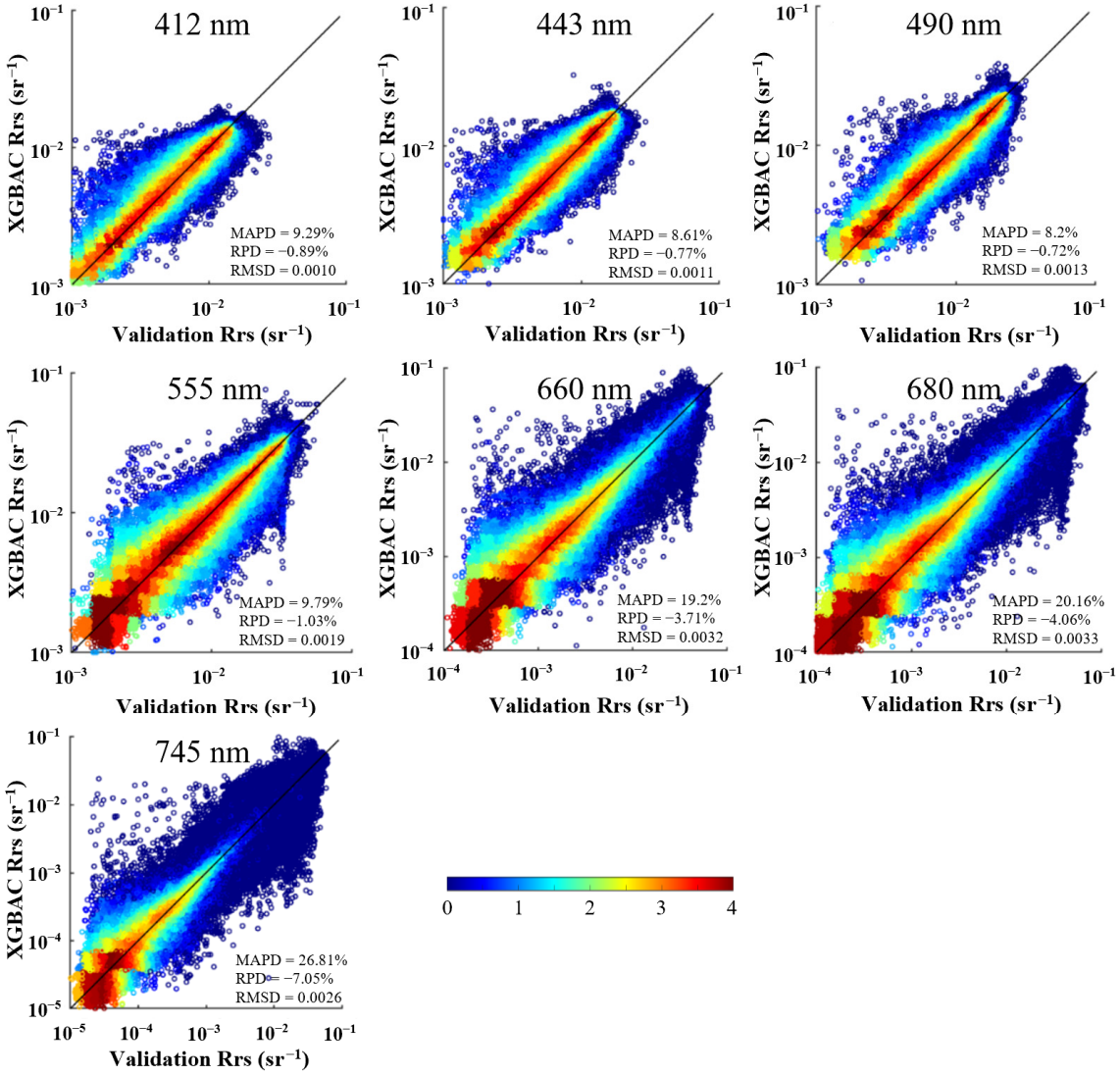

Figure 4. Scatterplots of Rrs from the XGBAC model and the reference Rrs values of the validation dataset at each GOCI band. The numerical values along the colour bar correspond to the pixel density on a logarithmic scale.

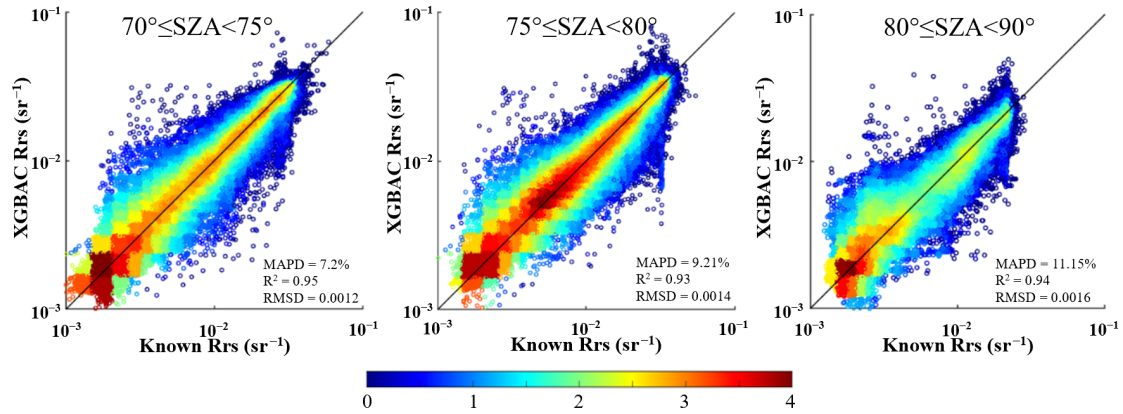

Figure 5. Scatterplots of Rrs (490 nm) retrieved from the XGBAC model vs. the values of the evaluation dataset for the different ranges of SZA. The numerical values along the colour bar correspond to the pixel density on a logarithmic scale.

Figure 6 shows the variations in the Q1, median and Q3 of the MAPD of the Rrs retrieved from XGBAC and OC-SMART as a function of SZA. The Q1, median and Q3 of the MAPD of the Rrs obtained from the XGBAC model were obviously lower than those from OC-SMART. The Q1, median and Q3 values of the MAPD in the Rrs (490 nm) using the XGBAC model exhibited a slight increase with the SZA, reaching approximately 10% (shown in Figure 6). On the contrary, these parameters from OC-SMART were significantly higher than those from the XGBAC model at high SZAs and increased sharply with increasing SZA. Based on the above results, it can be concluded that the XGBAC algorithm is better than OC-SMART and decreases slightly under super-high SZAs (>80°).

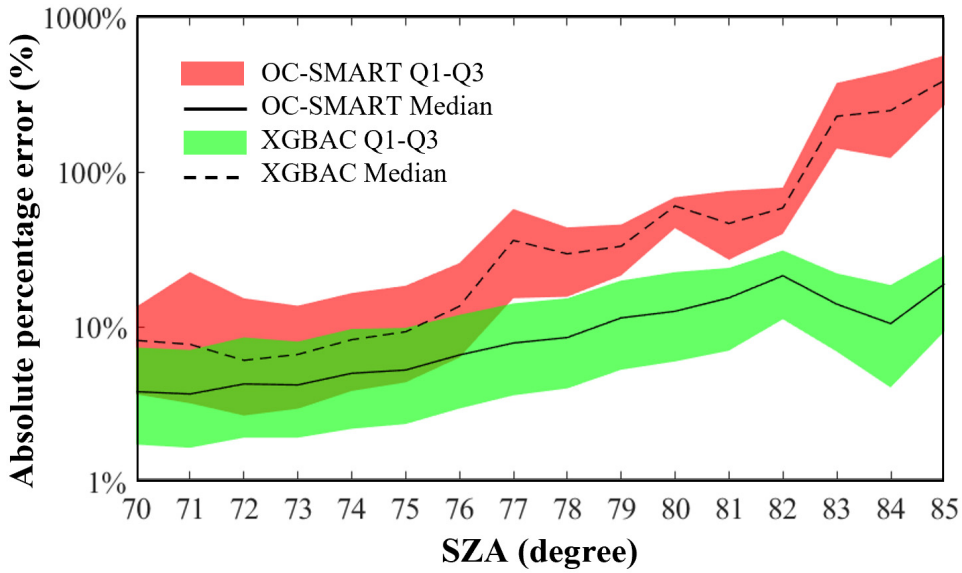

Figure 6. Variations in absolute percentage error of Rrs (490 nm) retrieved from the XGBAC and OC-SMART algorithms with the SZA.

The satellite and model-estimated Rrs were evaluated with the QA system. The frequency distribution and the statistical parameters of the QA scores obtained using the XGBAC are shown in Figure 7 and given in Table 2, respectively. Similar statistical parameters of the QA score were observed for the XGBAC model and OC-SMART (mean = 0.99, median = 0.99). The percentages of the QA score > 0.99 were 92%, 89% and 87% for reference Rrs, XGBAC-estimated Rrs and OC-SMART-estimated Rrs, respectively. The QA scores obtained by XGBAC were highly concentrated above the threshold of 0.99 with smaller standard deviations of 0.006, which indicated that the spectral shapes of validation dataset from XGBAC were more consistent than those from OC-SMART.

Table 2. Statistical parameters for the QA score obtained using XGBAC and OC-SMART.

Dataset	Mean	Median	Min	Max	Std.
XGBAC	0.99	0.99	0.90	1	0.006
OC-SAMRT	0.99	0.99	0.84	1	0.022
Reference *	0.99	0.99	0.93	1	0.005

Note: * the reference indicates the reference Rrs in the validation dataset.

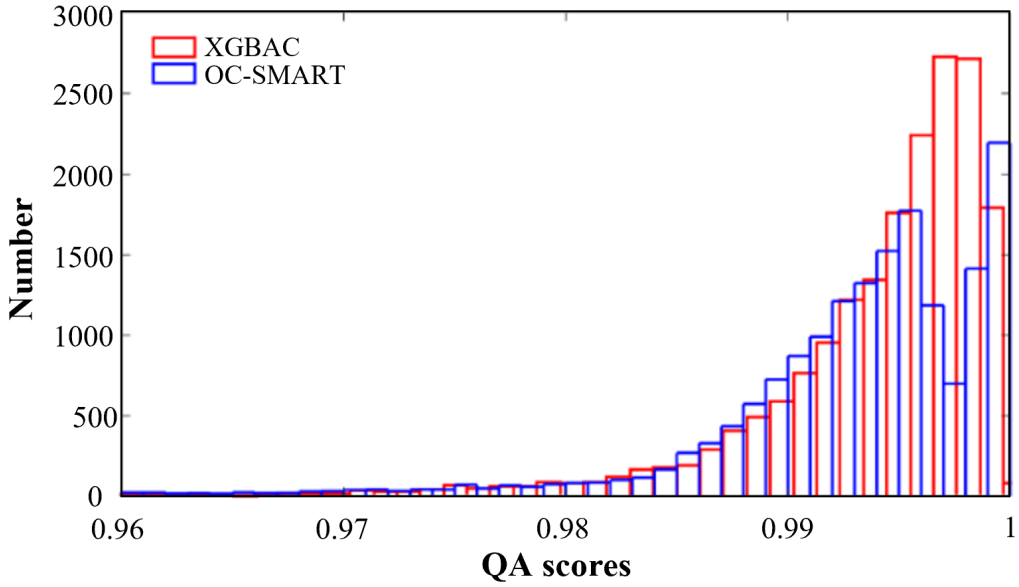

Figure 7. Frequency distribution of QA scores for the Rrs retrievals obtained using XGBAC and OC-SMART.

3.2. Application to GOCI Images

Figure 8 shows the Rrs maps at 443 nm, 555 nm and 680 nm retrieved using XGBAC from GOCI images in the morning and at dusk (08:55 LT and 15:55 LT), and the corresponding SZA values are also presented in Figure 8. Obviously, the XGBAC model could generate more valid data under high SZAs compared with the KOSC algorithm (shown in Figure 1). Furthermore, XGBAC demonstrates effective processing of the ocean colour remote-sensing data even when the SZA is higher than 80°. Figure 9 shows the RPD results at the 555 nm band retrieved by XGBAC and OC-SMART, respectively. It is evident that the Rrs (555 nm) values obtained using OC-SMART showed considerable variation over the two GOCI observation times, with an RPD exceeding 100% in the Yellow Sea, indicating large uncertainties in the Rrs in the morning and at dusk retrieved from OC-SMART. In contrast, XGBAC yielded more stable Rrs (555 nm) retrievals at morning/dusk. Although the Rrs retrievals from the two algorithms were consistent at noon, XGBAC produced more dependable Rrs retrievals in the morning and at dusk. However, there are also some underestimations in the Bohai Sea with high turbidity and SZAs, which might be attributed to the complex and unstable nature of the water conditions in inshore regions with a higher CV (shown in Figure 10).

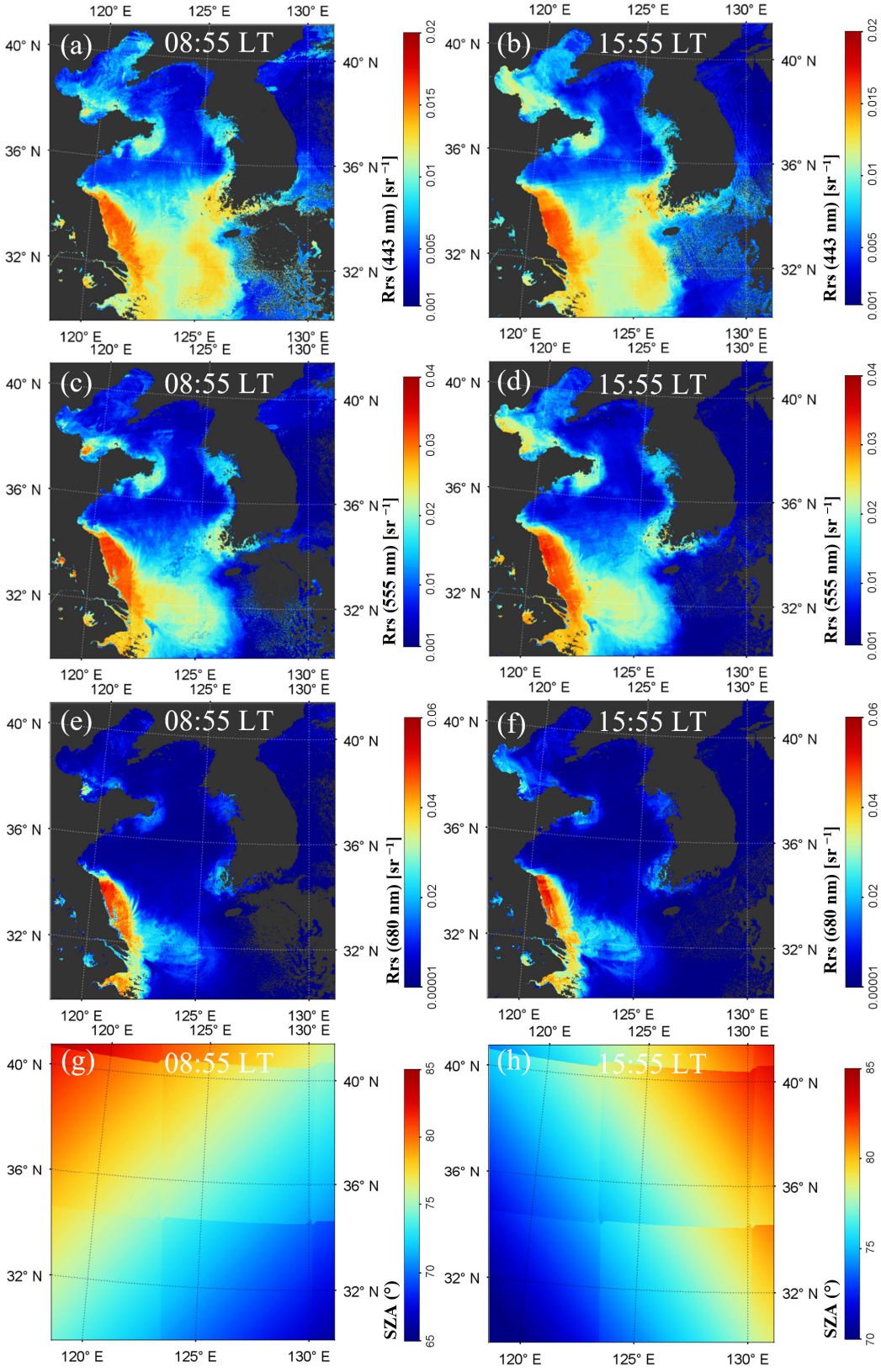

Figure 8. Rrs maps at 443 nm, 555 nm and 680 nm derived by XGBAC from the GOCI data sensed on 13 January 2021 at 08:55 LT (**a**,**c**,**e**) and 15:55 LT (**b**,**d**,**f**); SZA in the 08:55 LT (**g**) and 15:55 LT (**h**).

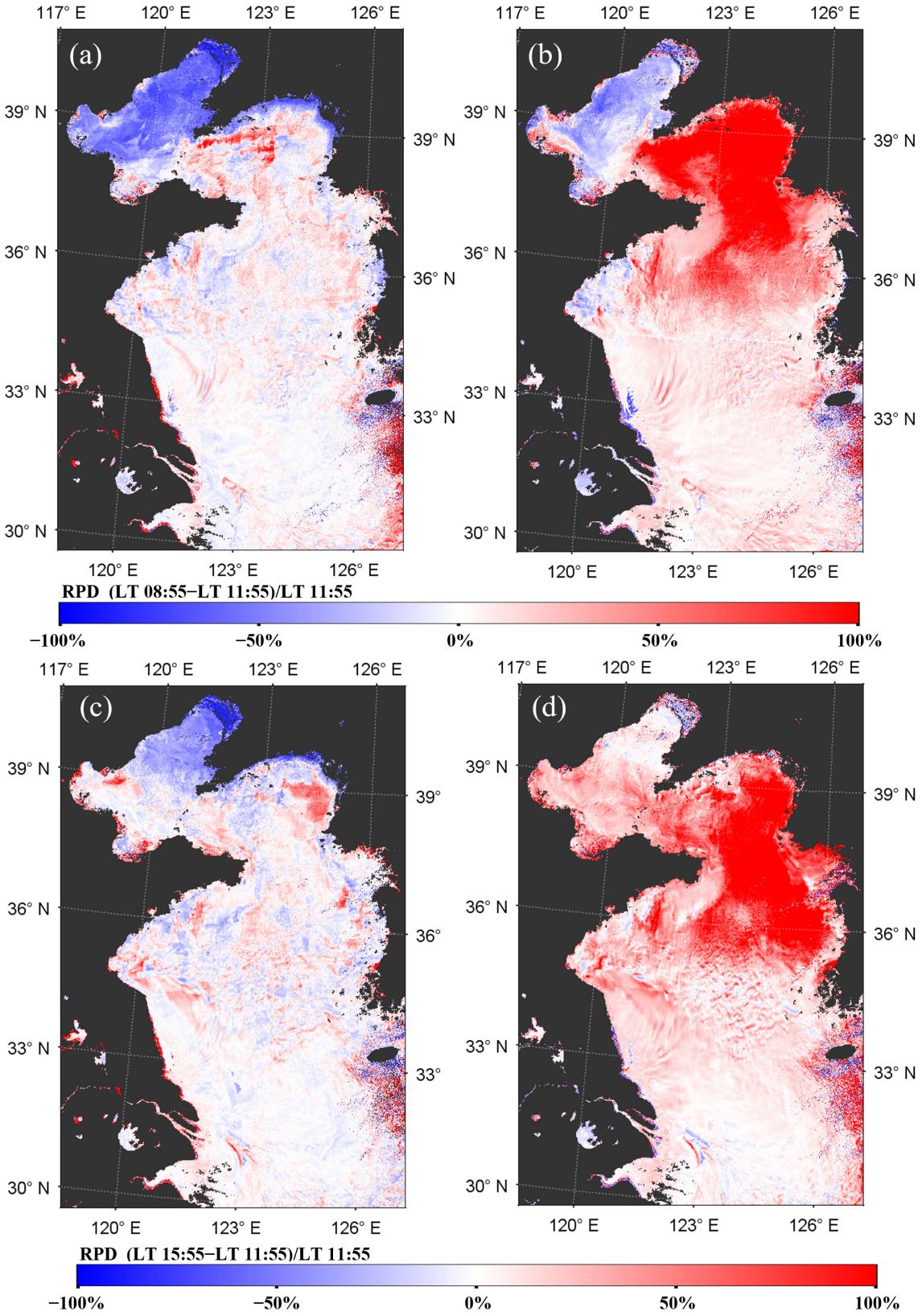

Figure 9. RPD between Rrs (555 nm) values in morning/afternoon hours and noontime observations at 12:55 LT using the XGBAC algorithm (**a**,**c**) and OC-SMART algorithm (**b**,**d**).

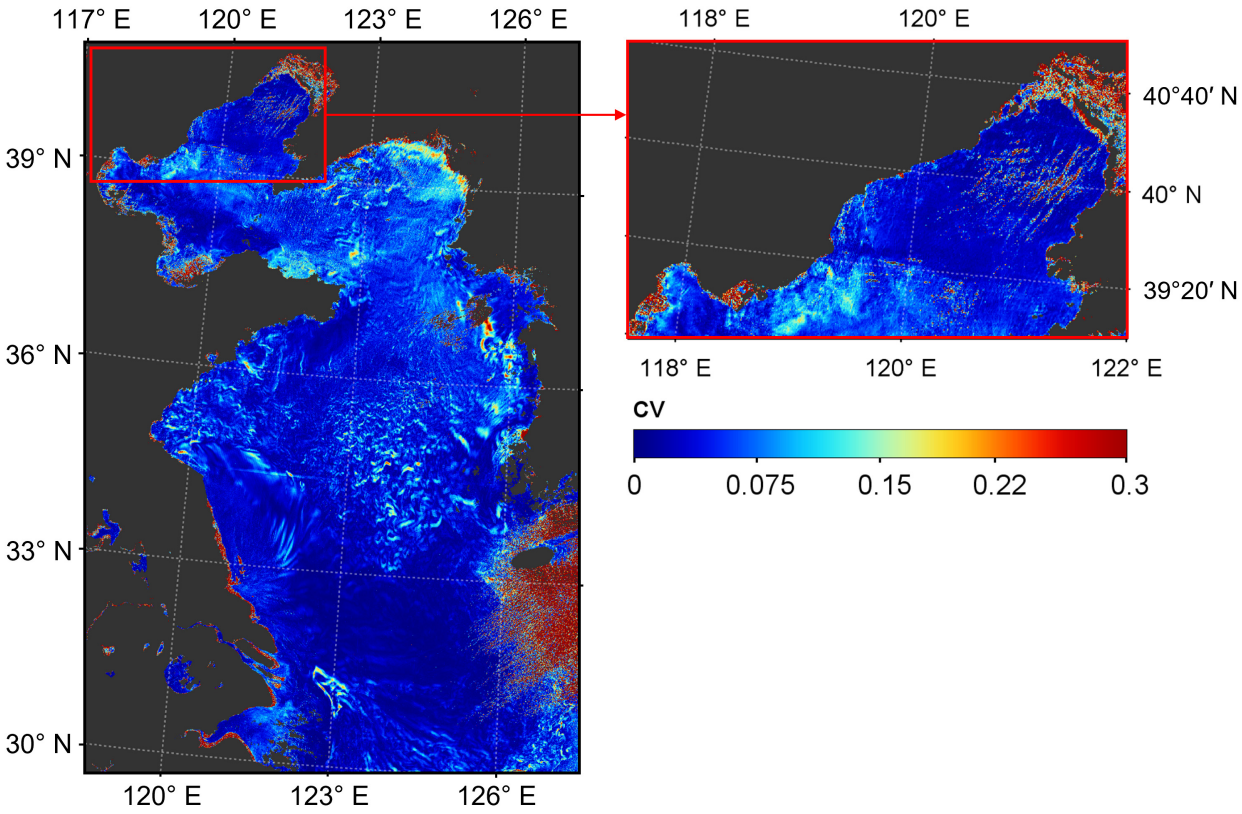

Figure 10. The temporal CV of the pixel values examined using multiple noontime (10:55 LT–13:55 LT) Rrs (555 nm) values within a day.

The YRE and HZB is one of the most turbid regions with sediment-laden waters in the world [46], making the standard AC algorithms invalid [47,48]. Using the Rrs data retrieved by the XGBAC and the regional TSS model following Equation (6), we obtained hourly TSS concentration maps in the YRE and HZB at every 1 h from 08:55 LT to 15:55 LT on 13 January 2021 (shown in Figure 11a–h). There are no retrievals in the YRE and HZB regions on the KOSC TSS products at 08:55 LT and 15:55 LT (shown in Figure 11i,j). With the XGBAC, the first noticeable aspect is the recovery of the area that had a previous algorithm failure in the GOCI-derived TSS, as well as significantly changed values in the regions after the correction. The TSS concentration throughout the HZB was over 100 mg/L, and even over 1000 mg/L in most areas. High TSS concentrations in the HZB were probably caused by the re-suspension of sediment by strong tidal currents and wind. The spatial distribution of TSS concentration in this area is similar to the previous study [14,49], where a higher TSS concentration was observed for coastal waters and decreased in moving further away from the outer shelf area. It is also evident that TSS values around the YRE are also significantly increased due to improved Rrs, and numerous plumes of TSS are also clearly present in the coastal region with TSS concentrations ranging from 50 to 500 mg/L. This result can assist environmental management agencies in understanding the intra-day fluctuations of TSS to make better decisions to safeguard coastal ecosystems. Compared to the KOSC standard products, there are obvious improvements in GOCI ocean colour products after applying the XGBAC.

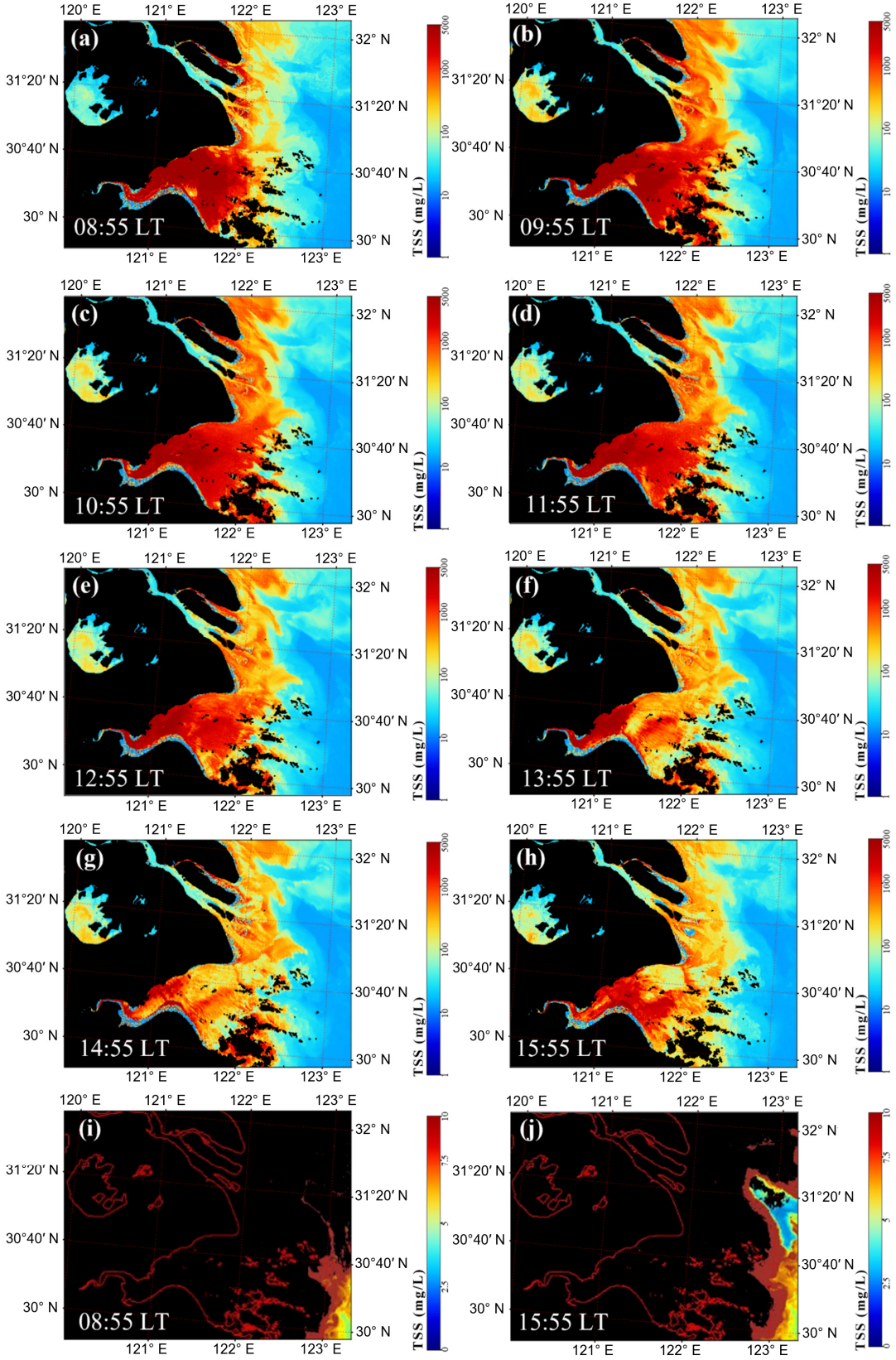

Figure 11. The hourly TSS maps in the YRE and HZB retrieved by Rrs product applying the XGBAC (**a**–**h**) and KOSC standard TSS products (**i**,**j**) on 13 January 2021.

4. Discussion

Over the past several decades, most remote-sensing algorithms have been established and validated based on satellite and in situ data under conditions of low-to-moderate SZA (<70°). In recent years, however, some studies showed the ability of satellites to detect ocean colour components under high SZAs [19,50]. Coastal waters are complex due to various factors, including changes in optical properties, absorbing aerosols effects, interactions with land and high SZAs in the morning and at dusk. Therefore, it is essential to develop specific AC algorithms for turbid coastal waters. However, the current AC algorithms only take into account either the optical complexity of coastal waters or high SZAs [24,36], without considering the joint effect of both. This study proposed a novel AC algorithm for optically complex waters under high SZAs. As more geostationary ocean colour satellite images become available over time, such as GOCI-II, Himawari-8 and Himawari-9, and the geostationary ocean colour satellite to be launched by China, the merits of XGBAC could be further demonstrated. The XGBAC can be easily applied to other remote-sensing data sources, leading to the better monitoring frequency and time coverage of the aquatic environment. This will facilitate the identification of patterns and underlying mechanisms responsible for diurnal changes in the aquatic ecosystem, providing decision support for ocean and coastal management.

The proposed XGBAC model is based on machine-learning models. To ensure the applicability of the model, a diverse dataset covering large Rrs and SZA variations is required. Therefore, the high quality Rrs–ρ_{rc} matchup data were collected to train and validate the proposed XGBAC model. For high-quality Rrs, the stable waters with a low CV were used for the extraction of high-quality Rrs. The model can also be well migrated to applications when higher quality Rrs data at lower SZAs is acquired through alternate AC algorithms. The successful retrieval of Rrs from the satellite-observed ρ_{rc} data in this study could be partially attributed to the competence of XGBoost in learning nonlinear relationships and executing intricate fitting tasks [51]. Nevertheless, it is important to note that as an ensemble tree-based method, XGBoost is known to be highly dependent on sample size during model training [52], and therefore has a limited generalisation ability for small sample sizes. Therefore, more in situ Rrs samples have the potential to further improve the performance of XGBoost. Particularly, for water quality assessment and AC in small regional water bodies, it is necessary to conduct a validation analysis of remote-sensing data with the model data and the in situ measurements. The MAPD and RPD of Rrs were used to assess the performance of the atmospheric correction, as they are a dimensionless metric to measure the estimation bias for each sample. A dimensionless metric is important considering the large-scale variation in Rrs from below 10^{-5} to above 10^{-1} sr^{-1}.

The results indicate that the proposed XGBAC algorithm can effectively retrieve Rrs in the morning or at dusk and outperforms the KOSC products. However, due to the opacity and difficulty in the explaining of machine-learning methods, understanding the mechanism of the relationship between Rrs and ρ_{rc} requires further investigation. The uncertainty of the noontime Rrs may have a limited effect on the accuracy of the XGBAC algorithm. And the complex variations in coastal turbid water might introduce uncertainties in extracting stable Rrs–ρ_{rc} matchup data, which could also limit the accuracy of the training model. Additionally, adjacency effects are a critical factor to consider when analysing coastal waters, can be challenging to correct at high SZAs and would contaminate larger areas at high SZAs and VZAs [53]. The accuracy of XGBAC would decrease slightly under super-high SZAs (>80°). This is due to the accuracy of the ρ_{rc} declining and the limited amount of effective Rrs data under high SZAs. The decrease in AC accuracy under high SZAs is mainly affected by the high total air mass, Earth curvature and Fresnel reflectance [54]. Furthermore, the AC procedure must consider the various components of the atmosphere, which significantly impact physical and chemical processes and serve as climatic and environmental factors. Therefore, there is still a need to

investigate the radiative transfer mechanisms that underlie the variations in Rrs due to the Earth's curvature and other high SZA processes.

Taking advantage of the hourly observations, GOCI data with XGBAC clearly and accurately capture the dynamic of TSS in the YRE and HZB. These products will be useful to researchers, modellers and those currently involved in assessing the impact of water quality. Many factors could impact the accuracy of TSS retrievals, including the natural high-frequency variability of the TSS, sensor noise and the uncertainty of the TSS retrieval algorithm [55,56]. Diurnal changes and drivers of water quality variability should be further investigated in the YRE and HZB using TSS products and data on coastal processes such as tides, winds and freshwater discharge.

5. Conclusions

To solve the problem of the number and magnitude of errors in the operation of the atmospheric-correction algorithm for optically complex waters under high SZAs, this study proposed a new XGBAC model. The proposed model utilised the XGBoost algorithm to retrieve the Rrs from the GOCI ρ_{rc} data directly. The retrieval of Rrs using the XGBAC model agreed well with the noontime high-quality Rrs ($R^2 \geq 0.92$, MAPD $\leq 26.81\%$ and QA score ≥ 0.9). Estimates of Rrs in the coasts of Eastern and Northern China are investigated using the newly proposed XGBAC method. Furthermore, the atmospherically corrected GOCI data were used to map the TSS concentration in the coastal waters of the YRE and the HZB. The results show that the XGBAC model obtained reasonable Rrs values in the optically complex water under high SZAs, which clearly worked better than the OC-SMART model and KOSC standard algorithms. The GOCI-derived ocean colour products are improved in spatio-temporal coverage after applying the new correction method, especially for TSS products that are derived using the Rrs ratio algorithms. The XGBAC model has the potential to extend the temporal coverage of effective ocean colour remote-sensing data, thereby enabling more frequent monitoring of aquatic ecosystems and their diurnal variations. This could provide valuable insights into the dynamics of oceanic processes, such as phytoplankton blooms, ocean currents and river plumes, which are critical for understanding and managing aquatic ecosystems. In the future, the integration of ocean colour remote-sensing data with government management data is expected to provide more effective support for ocean and coastal management. This research can potentially provide information to support the sustainable development and management of coastal and oceanic environments.

Author Contributions: Conceptualization, Y.W. (Yongquan Wang) and H.L.; methodology, Y.W. (Yongquan Wang); software, Z.Z.; validation, Y.W. (Yongquan Wang). and D.Z.; formal analysis, Y.W. (Yanru Wang); resources, H.L., Q.L. and G.W.; data curation, Y.Z.; writing—original draft preparation, Y.W. (Yanru Wang); writing—review and editing, H.L. and G.W.; visualization, Y.W. (Yongquan Wang); supervision, H.L. and G.W.; project administration, H.L.; funding acquisition, H.L. and G.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (Grant No. 42001281 and 42371337), the Guangdong Basic and Applied Basic Research Foundation (Grant No. 2023A1515011946), the Guangdong Basic and Applied Basic Research Foundation (Name: Quantitative retrieval of water environment using unmanned aerial vehicle hyperspectral remote sensing under low solar illumination conditions, PI: Huizeng Liu, duration: 2024.01–2026.12) and the Shenzhen Science and Technology Program (Grant No. JCYJ20230808105709020).

Data Availability Statement: Data will be made available on request.

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

References

1. Sagar, S.; Roberts, D.; Bala, B.; Lymburner, L. Extracting the intertidal extent and topography of the Australian coastline from a 28 year time series of Landsat observations. *Remote Sens. Environ.* **2017**, *195*, 153–169. [CrossRef]
2. Wang, Z.; Gao, Z.; Jiang, X. Analysis of the evolution and driving forces of tidal wetlands at the estuary of the Yellow River and Laizhou Bay based on remote sensing data cube. *Ocean Coast. Manag.* **2023**, *237*, 106535. [CrossRef]
3. Ryu, J.; Won, J.; Min, K. Waterline extraction from Landsat TM data in a tidal flat: A case study in Gomso Bay, Korea. *Remote Sens. Environ.* **2002**, *83*, 442–456. [CrossRef]
4. Boak, E.; Turner, I. Shoreline definition and detection: A review. *J. Coast. Res.* **2005**, *21*, 688–703. [CrossRef]
5. Cracknell, A. Remote sensing techniques in estuaries and coastal zones—An update. *Int. J. Remote Sens.* **1999**, *20*, 485–496. [CrossRef]
6. Liu, W.; Liu, S.; Zhao, J.; Duan, J.; Chen, Z.; Guo, R.; Chu, J.; Zhang, J.; Li, X.; Liu, J. A remote sensing data management system for sea area usage management in China. *Ocean Coast. Manag.* **2018**, *152*, 163–174. [CrossRef]
7. Wang, Y.; Liu, H.; Wu, G. Satellite retrieval of oceanic particulate organic nitrogen concentration. *Front. Mar. Sci.* **2022**, *9*, 943867. [CrossRef]
8. Liu, H.; Li, Q.; Zhu, P.; Hu, Z.; Yang, C.; Wang, Y.; Cui, A.; Wang, Z.; Wu, G. A Glimpse of Ocean Color Remote Sensing from Moon-Based Earth Observations. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4206811. [CrossRef]
9. Zhu, J.; Zhang, C.; Ji, H.; Hong, P.; Wu, X.; Zhang, J. Long-term characteristics of water quality around the Guangdong-Hong Kong-Macau Greater Bay Area: Spatiotemporal variations and meteorological effects. *Ocean Coast. Manag.* **2022**, *216*, 105962. [CrossRef]
10. Chen, C.; Liang, J.; Yang, G.; Sun, W. Spatio-temporal distribution of harmful algal blooms and their correlations with marine hydrological elements in offshore areas, China. *Ocean Coast. Manag.* **2023**, *238*, 106554. [CrossRef]
11. Nazeer, M.; Nichol, J. Development and application of a remote sensing-based Chlorophyll-a concentration prediction model for complex coastal waters of Hong Kong. *J. Hydrol.* **2016**, *532*, 80–89. [CrossRef]
12. Liu, H.; Li, Q.; Bai, Y.; Yang, C.; Wang, J.; Zhou, Q.; Hu, S.; Shi, T.; Liao, X.; Wu, G. Improving satellite retrieval of oceanic particulate organic carbon concentrations using machine learning methods. *Remote Sens. Environ.* **2021**, *256*, 112316. [CrossRef]
13. Klemas, V. Fisheries applications of remote sensing: An overview. *Fish. Res.* **2013**, *148*, 124–136. [CrossRef]
14. He, X.; Bai, Y.; Pan, D.; Huang, N.; Dong, X.; Chen, J.; Chen, C.; Cui, Q. Using geostationary satellite ocean color data to map the diurnal dynamics of suspended particulate matter in coastal waters. *Remote Sens. Environ.* **2013**, *133*, 225–239. [CrossRef]
15. Wang, M. Atmospheric correction for remotely-sensed ocean-colour products. In *Reports and Monographs of the International Ocean-Colour Coordinating Group (IOCCG)*; IOCCG: Dartmouth, NS, Canada, 2010. [CrossRef]
16. Fukushima, H.; Higurashi, A.; Mitomi, Y.; Nakajima, T.; Noguchi, T.; Tanaka, T.; Toratani, M. Correction of atmospheric effect on ADEOS/OCTS ocean color data: Algorithm description and evaluation of its performance. *J. Oceanogr.* **1998**, *54*, 417–430. [CrossRef]
17. Liu, H.; He, X.; Li, Q.; Hu, X.; Ishizaka, J.; Kratzer, S.; Yang, C.; Shi, T.; Hu, S.; Zhou, Q.; et al. Evaluation of Ocean Color Atmospheric Correction Methods for Sentinel-3 OLCI Using Global Automatic In Situ Observations. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4206319. [CrossRef]
18. He, X.; Stamnes, K.; Bai, Y.; Li, W.; Wang, D. Effects of Earth curvature on atmospheric correction for ocean color remote sensing. *Remote Sens. Environ.* **2018**, *209*, 118–133. [CrossRef]
19. Li, H.; He, X.; Shanmugam, P.; Bai, Y.; Wang, D.; Huang, H.Q.; Zhu, Q.K.; Gong, F. Semi-analytical algorithms of ocean color remote sensing under high solar zenith angles. *Opt. Express* **2019**, *27*, A800–A817. [CrossRef]
20. Gordon, H.; Wang, M. Surface-Roughness Considerations for Atmospheric Correction of Ocean Color Sensors. 2. Error in the Retrieved Water-Leaving Radiance. *Appl. Opt.* **1992**, *31*, 4261–4267. [CrossRef]
21. Wang, M. Effects of ocean surface reflectance variation with solar elevation on normalized water-leaving radiance. *Appl. Opt.* **2006**, *45*, 4122–4128. [CrossRef]
22. Li, H.; He, X.; Ding, J.; Bai, Y.; Wang, D.; Gong, F.; Li, T. The Inversion of HY-1C-COCTS Ocean Color Remote Sensing Products from High-Latitude Seas. *Remote Sens.* **2022**, *14*, 5722. [CrossRef]
23. Li, H.; He, X.; Bai, Y.; Gong, F.; Wang, D.; Li, T. Restoration of Wintertime Ocean Color Remote Sensing Products for the High-Latitude Oceans of the Southern Hemisphere. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4212312. [CrossRef]
24. Li, H.; He, X.; Bai, Y.; Shanmugam, P.; Park, Y.J.; Liu, J.; Zhu, Q.; Gong, F.; Wang, D.; Huang, H. Atmospheric correction of geostationary satellite ocean color data under high solar zenith angles in open oceans. *Remote Sens. Environ.* **2020**, *249*, 112022. [CrossRef]
25. Fedorov, S.; Molkov, A.; Kalinskaya, D. Aerosol Optical Properties above Productive Waters of Gorky Reservoir for Atmospheric Correction of Sentinel-3/OLCI Images. *Remote Sens.* **2022**, *14*, 6130. [CrossRef]
26. Kalinskaya, D.; Papkova, A. Variability of the Water-Leaving Radiance under the Conditions of Dust Transport by the Satellite Sentinel-3 Data on the Example of the Black Sea and Sevastopol. *Phys. Oceanogr.* **2023**, *30*, 370.

27. Kalinskaya, D.; Papkova, A. Why Is It Important to Consider Dust Aerosol in the Sevastopol and Black Sea Region during Remote Sensing Tasks? A Case Study. *Remote Sens.* **2022**, *14*, 1890. [CrossRef]
28. Gordon, H.; Wang, M. Retrieval Of Water-Leaving Radiance and Aerosol Optical-Thickness over the Oceans with Seawifs—A Preliminary Algorithm. *Appl. Opt.* **1994**, *33*, 443–452. [CrossRef]
29. Wang, M.; Jiang, L. Atmospheric Correction Using the Information from the Short Blue Band. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 6224–6237. [CrossRef]
30. Liu, H.; Zhou, Q.; Li, Q.; Hu, S.; Shi, T.; Wu, G. Determining switching threshold for NIR-SWIR combined atmospheric correction algorithm of ocean color remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2019**, *153*, 59–73. [CrossRef]
31. Zhang, M.; Hu, C.; Barnes, B.B. Performance of POLYMER Atmospheric Correction of Ocean Color Imagery in the Presence of Absorbing Aerosols. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 6666–6674. [CrossRef]
32. Wang, M. Remote sensing of the ocean contributions from ultraviolet to near-infrared using the shortwave infrared bands: Simulations. *Appl. Opt.* **2007**, *46*, 1535–1547. [CrossRef] [PubMed]
33. Brajard, J.; Santer, R.; Crepon, M.; Thiria, S. Atmospheric correction of MERIS data for case-2 waters using a neuro-variational inversion. *Remote Sens. Environ.* **2012**, *126*, 51–61. [CrossRef]
34. Diouf, D.; Niang, A.; Brajard, J.; Crepon, M.; Thiria, S. Retrieving aerosol characteristics and sea-surface chlorophyll from satellite ocean color multi-spectral sensors using a neural-variational method. *Remote Sens. Environ.* **2013**, *130*, 74–86. [CrossRef]
35. Fan, Y.; Li, W.; Gatebe, C.K.; Jamet, C.; Zibordi, G.; Schroeder, T.; Stamnes, K. Atmospheric correction over coastal waters using multilayer neural networks. *Remote Sens. Environ.* **2017**, *199*, 218–240. [CrossRef]
36. Fan, Y.; Li, W.; Chen, N.; Ahn, J.H.; Park, Y.J.; Kratzer, S.; Schroeder, T.; Ishizaka, J.; Chang, R.; Stamnes, K. OC-SMART: A machine learning based data analysis platform for satellite ocean color sensors. *Remote Sens. Environ.* **2021**, *253*, 112236. [CrossRef]
37. Ryu, J.-H.; Han, H.-J.; Cho, S.; Park, Y.-J.; Ahn, Y.-H. Overview of geostationary ocean color imager (GOCI) and GOCI data processing system (GDPS). *Ocean Sci. J.* **2012**, *47*, 223–233. [CrossRef]
38. Ryu, J.H.; Ishizaka, J. GOCI data processing and ocean applications. *Ocean Sci. J.* **2012**, *47*, 221. [CrossRef]
39. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794. [CrossRef]
40. Song, Z.; He, X.; Bai, Y.; Dong, X.; Wang, D.; Li, T.; Zhu, Q.; Gong, F. Atmospheric correction of absorbing aerosols for satellite ocean color remote sensing over coastal waters. *Remote Sens. Environ.* **2023**, *290*, 113552. [CrossRef]
41. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830. [CrossRef]
42. Feng, L.; Hu, C. Land adjacency effects on MODIS Aqua top-of-atmosphere radiance in the shortwave infrared: Statistical assessment and correction. *J. Geophys. Res. Ocean.* **2017**, *122*, 4802–4818. [CrossRef]
43. Wei, J.; Lee, Z.; Shang, S. A system to measure the data quality of spectral remote-sensing reflectance of aquatic environments. *J. Geophys. Res. Ocean.* **2016**, *121*, 8189–8207. [CrossRef]
44. Balasubramanian, S.V.; Pahlevan, N.; Smith, B.; Binding, C.; Schalles, J.; Loisel, H.; Gurlin, D.; Greb, S.; Alikas, K.; Randla, M.; et al. Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters. *Remote Sens. Environ.* **2020**, *246*, 111768. [CrossRef]
45. Chen, S.; Han, L.; Chen, X.; Li, D.; Sun, L.; Li, Y. Estimating wide range Total Suspended Solids concentrations from MODIS 250-m imageries: An improved method. *ISPRS J. Photogramm. Remote Sens.* **2015**, *99*, 58–69. [CrossRef]
46. Shi, W.; Wang, M. Characterization of global ocean turbidity from Moderate Resolution Imaging Spectroradiometer ocean color observations. *J. Geophys. Res. Ocean.* **2010**, *115*, C11022. [CrossRef]
47. Jiang, L.; Wang, M. Improved near-infrared ocean reflectance correction algorithm for satellite ocean color data processing. *Opt. Express* **2014**, *22*, 21657–21678. [CrossRef]
48. Shi, W.; Wang, M. Ocean reflectance spectra at the red, near-infrared, and shortwave infrared from highly turbid waters: A study in the Bohai Sea, Yellow Sea, and East China Sea. *Limnol. Oceanogr.* **2014**, *59*, 427–444. [CrossRef]
49. Ding, X.; He, X.; Bai, Y.; Zhu, Q.; Gong, F.; Li, H.; Li, J. High-frequency and tidal period observations of suspended particulate matter in coastal waters by AHI/Himawari-8. *Opt. Express* **2020**, *28*, 27387–27404. [CrossRef] [PubMed]
50. Li, H.; He, X.Q.; Bai, Y.; Chen, X.Y.; Gong, F.; Zhu, Q.K.; Hu, Z.F. Assessment of satellite-based chlorophyll-a retrieval algorithms for high solar zenith angle conditions. *J. Appl. Remote Sens.* **2017**, *11*, 012004. [CrossRef]
51. Ramraj, S.; Nishant, R.; Shatadeep, B. Experimenting XGBoost algorithm for prediction and classification of different datasets. *Int. J. Control. Theory Appl.* **2016**, *94*, 651–662.
52. Fassnacht, F.E.; Hartig, F.; Latifi, H.; Berger, C.; Hernández, J.; Corvalán, P.; Koch, B. Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sens. Environ.* **2014**, *154*, 102–114. [CrossRef]
53. Santer, R.; Schmechtig, C. Adjacency effects on water surfaces: Primary scattering approximation and sensitivity study. *Appl. Opt.* **2000**, *39*, 361–375. [CrossRef] [PubMed]
54. Ruddick, K.; Neukermans, G.; Vanhellemont, Q.; Jolivet, D. Challenges and opportunities for geostationary ocean colour remote sensing of regional seas: A review of recent results. *Remote Sens. Environ.* **2014**, *146*, 63–76. [CrossRef]

55. Su, J.; Wang, K. Changjiang River Plume and Suspended Sediment Transport in Hangzhou Bay. *Cont. Shelf Res.* **1989**, *9*, 93–111.
56. Dorji, P.; Fearns, P. Atmospheric correction of geostationary Himawari-8 satellite data for Total Suspended Sediment mapping: A case study in the Coastal Waters of Western Australia. *ISPRS J. Photogramm. Remote Sens.* **2018**, *144*, 81–93. [CrossRef]