Vietnam Journal of Earth Sciences, 46(3), 360-380, https://doi.org/10.15625/2615-9783/20714

Vietnam Journal of Earth Sciences
https://vjs.ac.vn/index.php/jse

A multivariate linear regression model for estimating chlorophyll-a concentration in Quan Son Reservoir (Hanoi, Vietnam) using Sentinel-2B Imagery

Nguyen Thien Phuong Thao[1], Nguyen Thi Thu Ha[1*], Pham Quang Vinh[2], Tran Thi Hien[1], Dinh Xuan Thanh[1]

[1]*Faculty of Geology, VNU University of Science, Vietnam National University, Hanoi, Vietnam*
[2]*Institute of Geography, Vietnam Academy of Science and Technology, Hanoi, Vietnam*

Received 11 October 2023; Received in revised form 25 February 2024; Accepted 03 May 2024

ABSTRACT

Monitoring chlorophyll-a concentration (Chla) in inland waters is vital for environmental assessment. This study develops an empirical multivariate linear regression (MLR) model to directly estimate Chla in Quan Son Reservoir using Sentinel-2B (S2B) Level 2A images. Regression analysis of a 68-point *in-situ* Chla dataset measured in Quan Son Reservoir between 2021 and 2023, in conjunction with the corresponding S2B reflectance data, reveals a significant correlation between Chla and a combination of the blue (B2), green (B3), and red (B4) bands (coefficient of determination, R^2 = 0.95). The Chla estimation model is validated using a 30-point *in-situ* dataset collected on various dates (R^2 = 0.87; the root-mean-squared error RMSE < 5%). Subsequently, the model is applied to ten S2B images acquired from 2021 to 2023, revealing Chla's spatio-temporal distribution across the reservoir. Two key trends emerge: (1) Chla is lower during winter (November and December) than in summer and early autumn (July and September), and (2) The distribution of Chla undergoes noticeable spatial changes, particularly in July, with elevated levels observed in areas characterized by tourist hotspots. This approach shows promise for monitoring Chla in similar inland waters.

Keywords: Empirical model, Sen2Cor, Sentinel-2B images, reservoirs, trophic state.

1. Introduction

The concentration of chlorophyll-a (Chla) serves as a vital indicator for assessing the abundance and quality of phytoplankton biomass in aquatic ecosystems (Ha et al., 2017; Chen et al., 2017). It provides valuable insights into water quality, biophysical conditions, and notably, the level of eutrophication in a water body. Consequently, Chla is widely used to assess the trophic status of lakes and reservoirs (Carlson, 1977; Kasprzak et al., 2008) and to evaluate water quality (USEPA, 2009; UNEP, 2014).

In recent decades, satellite remote sensing has emerged as a powerful tool, providing valuable insights into monitoring water quality in lakes and reservoirs. Among the numerous applications, the estimation of Chla stands out as one of the most extensively researched and commonly employed remote sensing techniques, especially for assessing aquatic ecosystems and closely tracking the phenomenon of eutrophication (Schalles, 2006). Eminent scientists from diverse fields have made significant contributions to the

*Corresponding author, Email: hantt_kdc@vnu.edu.vn

endeavor of extracting Chla information from remote sensing satellites. Their efforts have spanned a wide range of environmental assessments, including the study of oceanic waters (Tehrani et al., 2021; Wu et al., 2020; Le et al., 2013a), coastal zones and estuarine regions (Le et al., 2013b; Vilas et al., 2011; Loisel et al., 2017), as well as inland lakes (Liu et al., 2021; Chu et al., 2021; Werther et al., 2022). Despite the critical importance of Chla monitoring in inland lakes through remote sensing, such applications have been hindered by numerous challenges, including the scientific complexities of retrieving water's physical and biogeochemical properties and the limitations imposed by sensor specifications (Palmer et al., 2015).

Currently, various optical satellite imaging systems have been utilized for estimating Chla, including the Coastal Zone Colour Scanner (CZCS) (Conkright et al., 2003), Sea-viewing Wide Field Sensor (SeaWiFS) (Gholizadeh et al., 2016), Medium Resolution Imaging Spectrometer (MERIS) (Moses et al., 2012; Augusto-Silva et al., 2014), Moderate Resolution Imaging Spectroradiometer (MODIS) (Ogashawara et al., 2014; Li et al., 2019), Geostationary Ocean Colour Imager (GOCI) (Kim et al., 2016), Ocean and Land Colour Instrument (OLCI) (Werther et al., 2021; Kravitz et al., 2020), Landsat 8 OLI (Kuhn et al., 2019; Pu et al., 2019; Cao et al., 2020), and Sentinel-2 (S2) (Sent et al., 2021; Ogashawara et al., 2021; Niroumand-Jadidi et al., 2021). Among the numerous optical satellite datasets available, Sentinel-2B MSI (S2B) data have garnered recognition as a promising tool for Chla retrieval in inland waters (Niroumand-Jadidi et al., 2021; Perrone et al., 2021). The selection of this sensor is primarily motivated by its temporal coverage (10 days), spatial resolution (up to 10 meters), and easy accessibility. Additionally, S2B's spectral bands span all major color segments within the visible spectrum, which has a proven track record in providing high-quality science products for monitoring coastal and inland waters, particularly for Chla retrievals in turbid waters (Pahlevan et al., 2017). Particularly for monitoring water quality in small waterbodies, the optimal spatial and spectral resolution of S2B demonstrates its high capacity.

Various algorithms have been proposed to quantify Chla using empirical, analytical, and machine learning approaches. While machine learning models require a large dataset and are often applied for monitoring Chla in multiple or large waterbodies (Cao et al., 2020), they may not be appropriate for estimating Chla in small lakes with limited datasets. Empirical algorithms, though straightforward, are sensitive to local conditions, which limits their generalizability. On the other hand, analytical algorithms offer scalability but require accurate optical parameter estimates, making them less effective when applied in inland turbid waters (Jiang et al., 2019; Liu et al., 2020). In other words, analytical models demonstrate proficiency in well-defined waters; however, for complex inland waters with varying optical properties (Palmer et al., 2015), empirical approaches often perform better in capturing Chla patterns. Consequently, an empirical algorithm is suitable for monitoring Chla in reservoirs where water optical features are complex due to multiple influencing factors from river discharges and surrounding human activities.

Empirical algorithms rely on statistical relationships between remotely sensed data and *in-situ* Chla, often employing straightforward mathematical functions of the band values of band ratios. While no single-band algorithm has consistently demonstrated superior performance in quantifying Chla in inland waters (Moore et al., 2014; Ylöstalo et al., 2014; Kim et al., 2022), various multi-band ratios have been proposed, including the two-band (Mishra and Mishra, 2012; Vinh et al., 2019), three-band (Gitelson et al., 2011),

and four-band ratio algorithms (Le et al., 2013b). These algorithms typically utilize the difference between the absorption peak of Chla at the red band (675 nm) and other bands, helping to mitigate the influence of other water constituents on Chla estimation. However, band-ratio algorithms perform well in coastal waters but exhibit uncertainties when applied to highly turbid water bodies (Dall'Olmo and Gitelson, 2005). Furthermore, recent literature reviews have highlighted uncertainties associated with current atmospheric corrections for S2 data, particularly regarding the derived surface reflectance, $R_{\mathrm{rs}}(\lambda)$ in near-infrared (NIR) bands (Warren et al., 2019; Pahlevan et al., 2021). Consequently, traditional band-ratio algorithms, which often rely on NIR band reflectances (Gitelson et al., 2011), may not be optimal for accurately estimating Chla from S2 data. Therefore, there is a pressing need to develop new algorithmic approaches to accurately estimate Chla in small and complex inland waters using S2 data to fully leverage the potential of these datasets for monitoring water quality.

To address the challenges stemming from uncertainties in atmospheric correction and to develop a universal and optimized model for Chla based on lake datasets, several studies have employed the multivariate linear regression (MLR) approach (Cho et al., 2009; Matus-Hernández et al., 2018; Lins et al., 2018; Ouma et al., 2020; Franklin et al., 2020; Zhang et al., 2023). These investigations highlight MLR as a potent statistical technique, providing a flexible framework for elucidating the relationships between Chla and satellite-derived reflectances while considering the influence of multiple independent variables. Noteworthy is the effectiveness of MLR in estimating Chla from S2 level 2 data, consistently achieving high accuracy with R^2 values exceeding 0.7 (Ouma et al., 2020; Barraza-Moraga et al., 2022; Latwal et al., 2023). These findings strongly indicate that the MLR algorithm holds promise for accurately estimating Chla in the Quan Son Reservoir using S2B data.

This study aims to develop a Chla estimation model for Quan Son Reservoir in Hanoi, Vietnam, using S2B imagery and the MLR approach. To achieve this, a dataset comprising 68 *in-situ* Chla measurements was cross-regressed with $\mathrm{R_{rs}}(\lambda)$ values extracted from five different S2B scenes captured at various times over the reservoir. After validation using another 30-point dataset, the model was applied to ten cloud-free scenes, enabling the exploration of Chla's spatial-temporal dynamics from July 2021 to July 2023. In the study, we also further evaluated the performance of various atmospheric corrections for S2B images, including the Image correction for atmospheric effects (ICOR), dark object subtraction (DOS), and Sentinel 2 Correction (Sen2Cor) using *in-situ* water surface reflectances (N = 18) obtained on November 14, 2021.

2. Materials and Methods

2.1. Study area

Quan Son Reservoir, located in the Hop Tien communes of My Duc district, southern Hanoi City (Fig. 1), is an artificial waterbody established in 1960. Its primary purposes include irrigation, aquaculture, and regional ecological enhancement. The reservoir spans approximately 80 hectares, averaging 2 to 4 meters deep. Its operation involves storing water during the rainy season (April to September) and releasing it in the dry season (October to March of the subsequent year). Multiple water sources contribute to the reservoir, including the Cau Duong Stream, local precipitation, and smaller streams from the surrounding limestone mountains. Annually, Quan Son Reservoir provides 180 tons of fish and other aquatic products (Nguyen et al., 2010).

With its captivating and diverse landscape, Quan Son Reservoir has been designated as an ecotourism destination (Ha Tay Provincial Committee, 2008). Consequently, a tourist area has been developed, featuring restaurants and other tourist amenities, situated in the southeastern corner of the reservoir (Fig. 1). Additionally, several rest stops and bird-watching spots have been established on small islands within the reservoir to enhance tourism experiences. During the summer, the lotus ponds in the southern part of the reservoir bloom with vibrant flowers, attracting a flurry of tourist activities. According to statistical data from 2015, the number of visitors to Quan Son Reservoir reached approximately 10,000, nearly doubling the figures from 2007. However, from the end of 2019 to mid-2022, tourism and service activities surrounding the reservoir were interrupted due to the impact of the COVID-19 epidemic. Fortunately, by the end of 2022, these activities resumed and flourished again.

It's worth noting that food service restaurants along the reservoir often discharge wastewater directly into the reservoir, posing a direct threat to water quality. Moreover, along the southern shore of the reservoir, a residential community thrives, engaging in livelihood activities such as livestock raising. Unfortunately, domestic wastewater from these households is also directly discharged into the reservoir, further exacerbating concerns about water quality and environmental health.

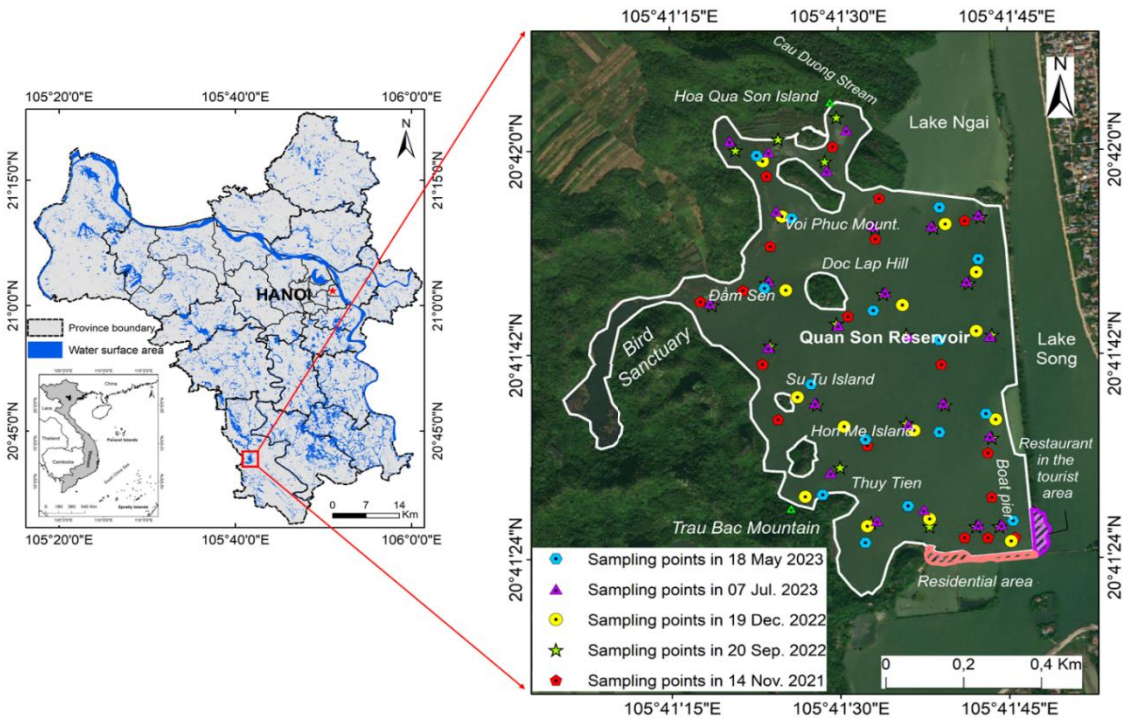

Figure 1. Location of Quan Son Reservoir in Hanoi city and water sampling points over the reservoir water on five surveyed dates

Positioned at the mountain base and distant from residential and agricultural zones, natural hydrological processes primarily influence the reservoir's water quality. However, the onset of tourism and aquacultural activities in the area has also

affected the water quality (Quan Sơn Fisheries and Tourism Joint Stock Company, 2013). Based on *in-situ* data, Thao et al. (2023) identified Quan Son Reservoir as a eutrophic water body, consistently exhibiting a Carlson's Trophic State Index (TSI) value exceeding 60 at all measurement points. High TSI values were particularly noticeable in areas near tourism service facilities, signifying the impact of human activities on the reservoir's trophic level.

2.2. Water sampling and measurement

In-situ data were collected at 98 points across the entire reservoir's surface during five field campaigns conducted on the following dates: November 14, 2021; September 20, 2022; December 19, 2022; May 18, 2023; and July 07, 2023. These points were accurately located using the Garmin eTrex Summit HC handheld Global Positioning System (GPS), as illustrated in Figure 1. The fieldwork dates were chosen to align with S2B satellite overpasses during periods of low cloud cover. Data collection occurred within one hour before and after the S2B satellite acquired the local scene.

At each designated point, water samples were collected at a depth of 30 cm using a Van Dorn water sampler and stored in 1-L dark-colored bottles. The water samples underwent analysis in the laboratory to determine Chla, following the standard method coded SMEWW 10200H: 2012 of the American Public Health Association (APHA, 1998). This involved filtering the water samples through a 0.45 µm pore size, 47 mm diameter Whatman Cellulose Nitrate filter to capture all algal cells in the water. Subsequently, the material retained on the filter paper was extracted using 90% acetone. Chla in the extracts was then determined spectrophotometrically using a DR 6000 UV-VIS Laboratory Spectrophotometer (Hach, USA) equipped with a 1 nm spectral bandwidth and optically matched 13 mm diameter cuvettes. The Chla was calculated using the following equations:

$$\mathrm{Chl_a(\mu g/L)} = \frac{C_a \times V_1}{V_o} \quad (1)$$

where V_1 is the extract volume in litter (L), V_o is the sample volume in cubic meters (m^3), C_a is Chla of pigment in the extract, which was calculated by the following equation:

$$C_a = 11.85 \cdot D_{664} - 1.54 \cdot D_{647} - 0.08 \cdot D_{630} \quad (2)$$

where D_{664}, D_{647}, D_{630} are optical density at wavelengths 664, 647, and 630, respectively.

Above-water surface reflectance measurements were conducted at 18 water-sampled sites on November 14, 2021, across Quan Son Reservoir using the GER1500 spectroradiometer (Spectra Vista Corporation, New York, U.S), following the protocol method (Mueller et al., 2003). *In-situ* reflectance at each measurement point was calculated using the following equation:

$$R_\omega(\lambda) = R_p \cdot \left(\frac{L_w(\lambda) - r \cdot L_{sky}(\lambda)}{\pi \cdot L_p(\lambda)} \right) \quad (3)$$

where R_p represents the reflectance of the standard reference panel, $L_W(\lambda)$ signifies water-leaving radiance, $L_{sky}(\lambda)$ represents the radiance of the sky measured sequentially at 40-45 degrees from the nadir and zenith, respectively, and 135 degrees from the Sun in azimuth (Mobley, 1999), $L_p(\lambda)$ corresponds to the radiance of the reference panel, and r is the air-water interface reflectance, with a value of 0.022 for the calm weather (Tang et al. 2020) as the condition at the reservoir on November 14, 2021. The unit of $R_\omega(\lambda)$ is sr^{-1}. The *in situ* $\mathrm{R_{rs}(\lambda)}$ then was obtained by multiple $R_\omega(\lambda)$ by pi (Lehmann et al., 2023).

Concurrently with water sampling, water clarity was assessed using a standard 20-cm plastic Secchi disk (Wildco, Yulee, FL, U.S). The Secchi depth (SD) data served as a reference in the field to select sampling points while mitigating the effects of the reservoir bottom on water reflectance data.

2.3. Sentinel-2B image processing

Concurrently with the field observations, we utilized $\mathrm{R_{rs}(\lambda)}$, the water-leaving reflectance,

derived from five S2B level 2 images, which were acquired on the same dates as the water measurements to calibrate and validate the Chla estimation model. The image's detailed information is provided in Table 1. All these five images were directly downloaded from the Copernicus Open Access Hub website (https://scihub.copernicus.eu/dhus/#/home). The S2 level 2 provides bottom-of-atmosphere (BOA) reflectance, also known as surface reflectance, using the Sentinel-2 Correction (Sen2Cor) atmosphere correction processor. The Sen2Cor performs comprehensive atmospheric, terrain, and cirrus corrections on the Top-Of-Atmosphere Level-1C data to generate a Level-2A BOA product (Louis et al., 2016; Pflug et al., 2020). It is worth noting that while Sen2Cor was initially designed for land applications, it demonstrated reasonable accuracy for water retrievals in inland waters (Al-Kharusi et al., 2020; Grendaitė and Stonevičius, 2022; Martins et al., 2017). However, further evaluation of the performance of this product for water retrievals in typical inland waters, such as highly eutrophic and turbid waters, is still warranted.

Table 1. Information of S2B Level-2A images used in this study

No.	Scene Identifier	Acquisition Date
1	S2B_MSIL2A_20210717T032539_N0500_R018_T48QWH_20230225T022304	July 17. 2021
2	S2B_MSIL2A_20211114T033009_N0301_R018_T48QWH_20211114T054001	November 14 2021
3	S2B_MSIL2A_20211204T033119_N9999_R018_T48QWH_20221116T072100	December 4 2021
4	S2B_MSIL2A_20220403T032539_N0400_R018_T48QWH_20220403T080206	April 3 2022
5	S2B_MSIL2A_20220503T032529_N0400_R018_T48QWH_20220503T064952	May 3 2022
6	S2B_MSIL2A_20220702T032519_N0400_R018_T48QWH_20220702T064119	July 2 2022
7	S2B_MSIL2A_20220920T032519_N0400_R018_T48QWH_20220920T063937	September 20. 2022
8	S2B_MSIL2A_20221219T033139_N0509_R018_T48QWH_20221219T060245	December 19 2022
9	S2B_MSIL2A_20230518T032519_N0509_R018_T48QWH_20230518T075517	May 18 2023
10	S2B_MSIL2A_20230707T032519_N0509_R018_T48QWH_20230707T064529	July 7 2023

Additionally, five other cloud-free images acquired between July 2021 and July 2023 were employed to detect the temporal variations of Chla across the reservoir. Accordingly, the BOA reflectance of water pixels across the reservoir in the images was then extracted to develop the Chla estimation model and evaluate the performance of various atmospheric correction methods. In extracting the lake water surface, pixels near the shore are excluded because the reflectance spectrum of these pixels is affected by the bottom surface, owing to the shallow water level along the shore and the composition of the shore material.

In this study, two other atmospheric correction methods, namely iCOR and DOS, were further evaluated and compared with the output of the Sen2Cor processors. iCOR is a scene-generic atmospheric correction approach designed for various environments, including land, coastal, inland, or transitional waters (De Keukelaere et al., 2018; Sterckx et al., 2015). The iCOR atmospheric correction process involves four main steps (De Keukelaere et al., 2018): All pixels with land and water values are identified and separated. Next, the Aerosol Optical Thickness (AOT) values are extracted from the soil pixels based on an improved version of the method developed by Guanter (2007). Then, the SIMEC method is utilized to implement the adjacency correction approach for water pixels and fixed background land (Sterckx et al., 2015). Finally, the radiative transfer equation is solved to obtain the surface reflectance. iCOR employs the MODTRAN5 Look-Up Table (LUT) for performing the atmospheric correction process mentioned above (Berk et al., 2006). In this study, S2B images were processed using default parameters with iCOR (v.0.3) through SNAP software.

DOS (Chavez 1988) is an image-based atmospheric correction method that operates under the assumption that particular objects

were in complete shadow during image acquisition and, thus should exhibit zero reflectance. However, due to atmospheric scattering and absorption, the imaging system records a non-zero digital number (DN) value for these dark objects. This constant non-zero DN value, DN haze, can be identified from the image histogram and subtracted from all bands. This study applied DOS correction to L1C images using the atmospheric correction module in the QGIS 3.28 software.

2.4. Multivariate Linear Regression Approach

In this study, we employed a MLR approach to construct a model for estimating Chla. The model assumes that the relationship between the dependent variable (Chla) and the independent variables ($R_{rs}(\lambda)$s derived from five S2B) is linear. This model leveraged both *in-situ* Chla data and the spectral reflectance values from S2B Level-2A bands 2 through 8A, as illustrated in equation (4). To establish the model, we incorporated all reflectance values from the RGB bands (2, 3, and 4), the red-edge bands (5, 6, and 7), as well as the NIR bands (8 and 8A) as predictor variables. Subsequently, non-significant predictor bands with $p > 0.05$ were systematically excluded from the model to enhance its predictive accuracy and effectiveness.

$$\text{Chla}\ (\mu g/L) = \alpha_0 + \alpha_1 B_1 + \alpha_2 B_2 + \alpha_3 B_3 + \cdots + \alpha_n B_n + \varepsilon \quad (4)$$

Where $Chla$ is the dependent variable, α are coefficients related to the independent or predictor variables (which in this case are the reflectivity of the bands (B)), and ε is the residual error associated with the regression.

The model underwent calibration using a 68-point dataset collected on November 14, 2021, September 20, 2022, and July 7, 2023, followed by validation using the remaining 30-point dataset obtained on December 19, 2022, and May 18, 2023. The optimal model selection hinged on several critical criteria, including the coefficient of determination (R^2), the *F*-test, and *p*-values. Specifically, a threshold of $R^2 > 0.7$ was set, indicating that at least 70% of the dataset's residuals should conform to the linear relationship established by the model. The model's fit was rigorously tested at a significance level of $p = 0.05$, with a removal significance level of 0.10. The *F*-test was conducted at a 95% confidence level to assess the model's acceptance. Additionally, validation was performed using the root mean square error (RMSE) indicator.

3. Results

3.1. In-situ data

Table 2 displays the *in-situ* data collected at Quan Son Reservoir on five distinct measurement dates. Accordingly, the Chla exhibited substantial variability across the surveyed dates, ranging from 22.6 µg/L on November 14, 2021, to 63.2 µg/L on September 20, 2022. The mean Chla were 28.8 µg/L (November 14, 2021), 58.5 µg/L (September 20, 2022), 43.5 µg/L (December 19, 2022), 45.8 µg/L (May 18, 2023), and 41.5 µg/L (July 7, 2023). In contrast, *in-situ* SD measurements at 98 points showed insignificant variation, ranging from 0.45 m to 0.76 m, with a mean value of 0.64 m (November 14, 2021), 0.57 m (September 20, 2022), 0.62 m (December 19, 2022), 0.6 m (May 18, 2023), and 0.58 m (July 7, 2023). The five-time measurements revealed no significant correlation between Chla and SD (the Pearson correlation coefficient, *r*, ranged from 0.33 to 0.44). Chla is relatively correlated with SD (r = –0.44) in the entire dataset, suggesting that factors other than algal density contribute to reducing water clarity within Quan Son Reservoir. The variation of Chla in different survey months in Table 2 aligns with the understanding that phytoplankton growth is influenced by factors such as water temperature, nutrient concentrations and forms, and photosynthetic conditions (Cloern et al., 2014). As a result,

the Chla content in Quan Son Reservoir and in water bodies worldwide (Adams et al., 2021; Brown et al., 1998), fluctuates over time, requiring continuous monitoring.

Table 2. Descriptive statistics of water parameters obtained in five surveys

Date	Parameter	No. of sample	Minimum	Maximum	Mean	Standard Deviation	*r*
14 Nov. 2021*	Chla (µg/L)	18	22.6	39	28.8	4.5	0.33
	SD (m)	18	0.52	0.76	0.64	0.07	
20 Sep. 2022*	Chla (µg/L)	25	51	63.2	58.5	2.7	0.40
	SD (m)	25	0.5	0.61	0.57	0.05	
19 Dec. 2022**	Chla (µg/L)	15	36.5	50.7	43.5	3.8	0.36
	SD (m)	15	0.51	0.67	0.62	0.04	
May 18 2023**	Chla (µg/L)	15	34.2	55.3	45.8	6.5	0.42
	SD (m)	15	0.55	0.68	0.60	0.02	
07 Jul. 2023*	Chla (µg/L)	25	35.1	46.5	41.5	3.4	0.44
	SD (m)	25	0.45	0.69	0.58	0.05	

(*): data belongs to the calibration dataset. (**) are components of the validation dataset

The measured water reflectance, $R_\omega(\lambda)$, spanning 400–900 nm, and recorded at 18 sites across Quan Son Reservoir (Fig. 1) on November 14, 2021, is depicted in Fig. 2. These spectra curves reveal three distinct peaks: one centered around 555 nm in the green region and two in the NIR region at approximately 710 nm and 810 nm. The first two peaks in the reflectance spectra, occurring at 555 nm and 710 nm wavelengths, are associated with the absorption of Chla in the water (Rundquist et al., 1996). The peak at 810 nm is linked to variations in suspended matter within the water (Chen et al., 1992; Ma et al., 2007). An additional small peak, appearing near 650 nm in the red region, is considered to be associated with phycocyanin fluorescence (Hunter et al., 2008). Two notable troughs are observed clearly between these peaks, around 625 nm and 675 nm, respectively. These spectral features align with the typical characteristics of eutrophic waters and closely resemble previously reported spectra (Gitelson, 1992; Bennet and Bogorad, 1973; Schalles et al., 1998; Ma and Dai, 2005).

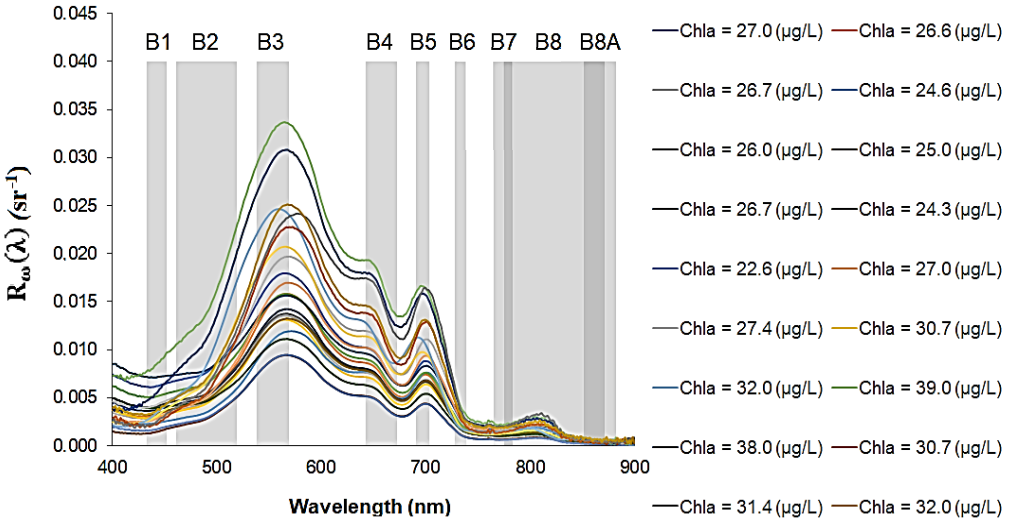

Figure 2. Reflectance spectra curves of water sampling points over Quan Son Reservoir measured on November 14, 2021, overlaid S2B visible to NIR bands (band 1: B1 to 8A: B8A) locations

3.2. Performance of Sentinel 2 Level 2 data

The *in-situ* water reflectance spectra shown in Fig. 2 were transformed into S2B band reflectance spectra using the averaging spectral band response function (Barsi et al., 2014), as presented in Fig. 3a. In this transformation, the reflectance value at each S2B band is computed by averaging the $R_{rs}(\lambda)$s values within the spectral range of the band. Figure 3b displays the top-of-atmosphere (TOA) reflectances at 18 measured points extracted from the S2B image acquired on November 14, 2021. It is evident that TOA reflectance is significantly higher and exhibits a different variation across the spectral bands than the water surface reflectance in Fig. 3a. The satellite-based surface reflectances, $R_{rs}(\lambda)$ at these 18 measured points, derived from TOA reflectances (Fig. 3b) and processed using DOS, ICOR, and Sen2Cor atmospheric correction methods, are depicted in Fig. 3c-e, respectively. Notably, all three atmospheric correction methods yield higher $R_{rs}(\lambda)$ values in the blue (band 1) and NIR bands (bands 6 to 8A) compared to *in-situ* $R_{rs}(\lambda)$. It's worth highlighting that the variations in satellite-based $R_{rs}(\lambda)$ values in the NIR bands are non-uniform and distinctly different from the variations in *in-situ* $R_{rs}(\lambda)$ values. In other words, all three atmospheric correction methods exhibit less accuracy in the NIR bands. This observation aligns with the findings of Pahlevan et al. (2021) and Pereira-Sandoval et al. (2019).

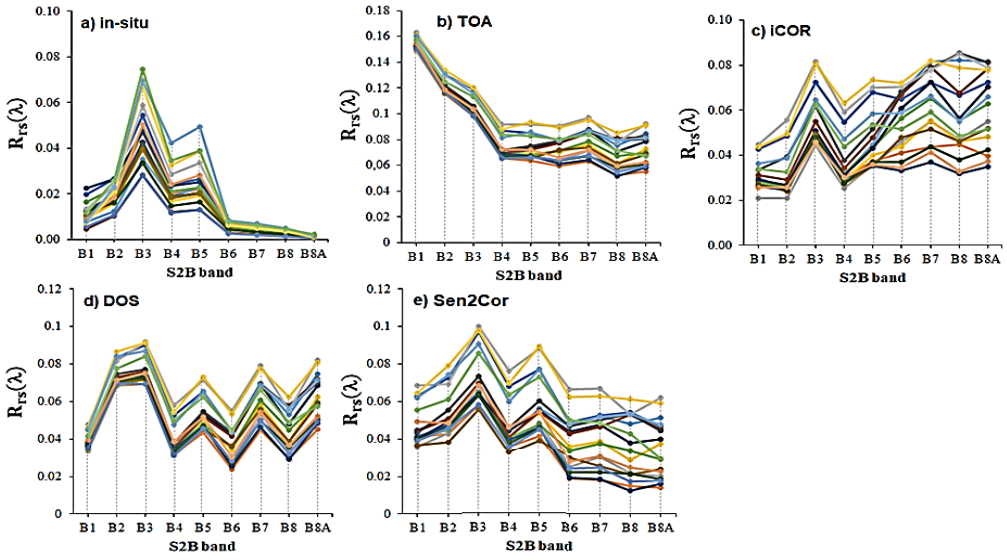

Figure 3. The *in-situ* $R_{rs}(\lambda)$ (a) compared to those derived from the S2B image (November 14, 2021) level 1 and level 2 using different atmospheric corrections: iCOR (c), DOS (d), and Sen2Cor (e)

To gain a comprehensive understanding of the differences between band-based $R_{rs}(\lambda)$ and *in-situ* $R_{rs}(\lambda)$, cross-comparisons were conducted and presented in Fig. 4. The computed metrics, including the correlation coefficient (r) and RMSE, reveal significant overestimations of $R_{rs}(\lambda)$ in the NIR bands by the DOS, iCOR, and Sen2Cor methods (Fig. 4f-i). However, on average, these methods yield reasonable $R_{rs}(\lambda)$ in bands 3, 4, and 5 (B3, B4, and B5) (Fig. 4c-e). ICOR demonstrates the best performance among the three methods, while DOS and Sen2Cor exhibit similar RMSE values in B3, B4, and B5 (Fig. 4c-e). Specifically, $R_{rs}(\lambda)$ derived from Sen2Cor exhibits stronger correlations with *in-situ* $R_{rs}(\lambda)$ than those from DOS, with values of r = 0.75, 0.74, and 0.75 for B3, B4,

and B5, respectively. All methods display acceptable correlations for band 2 (B2) (r = 0.52). The results indicate that there is not a significant difference in using $R_{rs}(\lambda)$ in the visible bands (B2 to B5) processed by ICOR or Sen2Cor. While ICOR demands more infrastructure and technical conditions during processing, particularly when dealing with images with a high percentage of cloud cover, Sen2Cor offers time-saving advantages in image processing (Bui et al., 2022; De Keukelaere et al., 2018). Consequently, Sen2Cor is deemed suitable for estimating Chla in Quan Son Reservoir.

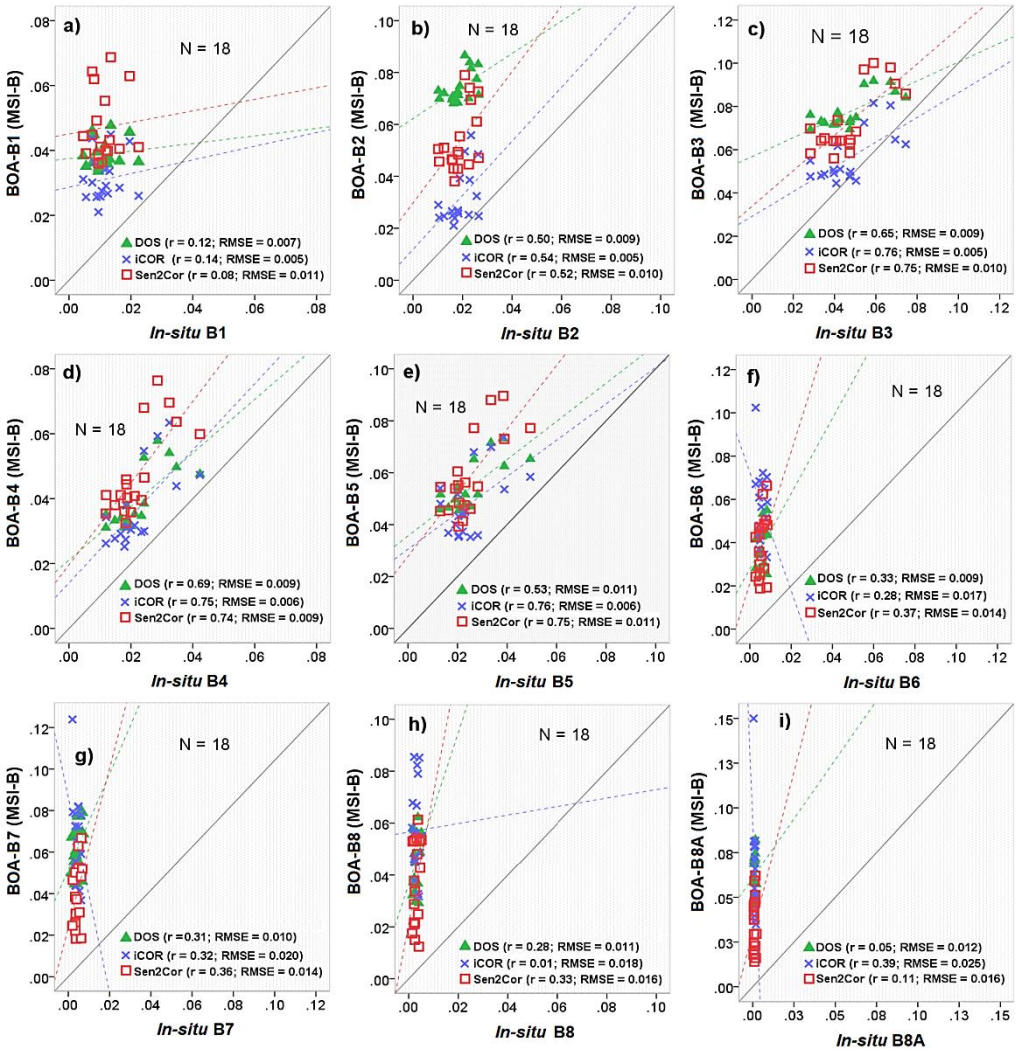

Figure 4. Scatter plots depicting the comparison between the bottom of the atmosphere (BOA) $R_{rs}(\lambda)$ and *in-situ* $R_{rs}(\lambda)$ values for S2B data using three different atmospheric correction algorithms at various spectral bands (band 1 to band 8A: B1 to B8A)

3.3. Model for estimating Chla from S2B imagery

This study examined the relationships between *in-situ* Chla and the $R_{rs}(\lambda)$ values at individual bands of S2B Level 2A imagery. The findings revealed that Chla exhibited poor and unstable correlations with the S2B band data overall (Fig. 5a). While Chla measured on July 7, 2023, demonstrated moderate

correlations with all nine S2B bands (r = 0.68 to 0.72), the *in-situ* Chla measured on September 20, 2022, and November 14, 2021, did not exhibit significant correlations with any of the S2B bands (r = –0.13 to 0.42). These results suggest that the single-band algorithm is unsuitable for estimating Chla in Quan Son Reservoir, and utilizing multiple spectral S2B bands would be more appropriate. Further analysis involved investigating the correlations between *in-situ* Chla and S2B band ratios, and the results are presented in Fig. 5b. Notably, Chla showed negative correlations with band ratios of B3/B2 (r = –0.94), B3/B4 (r = –0.79), B5/B2 (r = –0.73), and B5/B4 (r = –0.70). These negative correlations contradicted the expected behavior based on the absorption characteristics of Chla, which typically exhibits high absorption (low reflectance) in the blue and red regions (corresponding to B2 and B4) and low absorption (high reflectance) in the green and NIR regions (B3 and B5) (Mitchell & Kiefer, 1988; Fujiki, & Taguchi, 2002). As a result, these correlations between S2B band ratios and Chla do not adhere to physical laws but instead appear random, indicating that these correlations are unsuitable for developing a Chla estimation model.

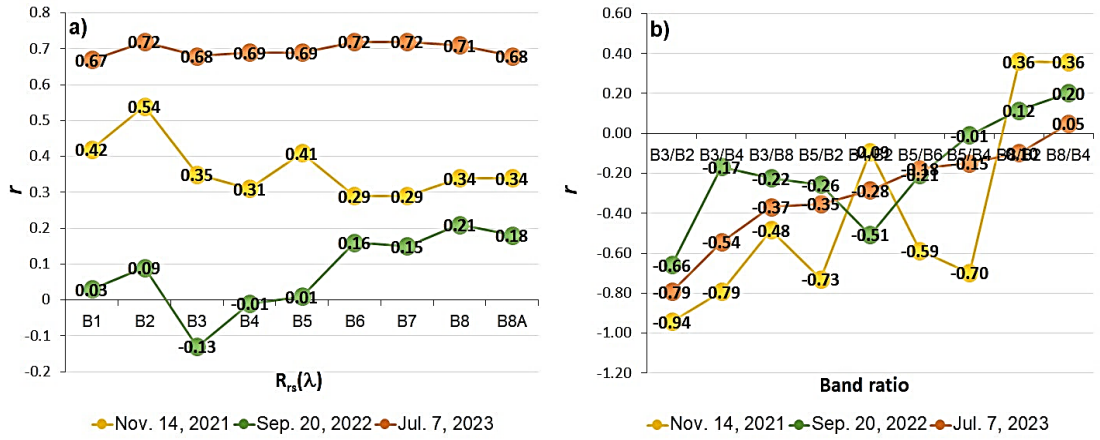

Figure 5. Correlations between *in situ* Chla with S2B band's $R_{rs}(\lambda)$s (a) and with band-ratios (b). *r* denotes the Pearson coefficient

To improve Chla estimation accuracy and develop the optimal regression model, we analyzed all $R_{rs}(\lambda)$s of eight S2B bands as explanatory variables (B2 to B8A) in a multivariate analysis using the calibration dataset of 68 data points. Results of the analysis are shown in Tables 3 and 4. Our analysis revealed that bands B2, B3, and B4 had significant regression coefficients (B = 0.075, –0.043, –0.028) with p-values less than 0.05, indicating their statistical validity as explanatory variables (Table 4). While B5 of S2 level 2 imagery demonstrated a relatively high correlation value with the *in situ* $R_{rs}(\lambda)$ (r = 0.75, Fig. 4e), it does not hold a significant value in estimating Chla within the multivariate regression model. This is because its p-value exceeds 0.05 (p = 0.223), so it is not selected as a variable for the Chla estimation model. We then evaluated MLR models using these three significant variables (B2, B3, and B4), resulting in the best-fit model with a maximum adjusted R^2 of 0.95, explaining 95% of Chla variation. The RMSE was 2.76 µg/L (Table 3), which is considered acceptable given the mean Chla of 48.5 µg/L at the 68 *in-situ* points.

Table 3. Summary of the multivariate linear regression model for estimating Chla from selected exploratory variables, B2, B3, and B4

Variable	r	R^2	Adjusted R^2	RMSE µg/L	F	p
B2, B3, B4	0.97	0.95	0.95	2.76	434.4	0.000

F and ρ denote the $F-ratio$ and statistical significance of the coefficient, respectively

The Chla prediction model summarized in Table 3 and 4 demonstrated statistical significance ($p < 0.05$) as the significance value (p) was 0.00, lower than the alpha value of 0.05. Additionally, the model exhibited an overall solid fit to the data, with F (3, 68) = 434.4 at $p < 0.05$. The constant value of the model, which is 26.76 ($p = 0.00$), indicates that the estimated TSI value of the model is 26.76 when reflectance values at bands B2, B3, and B4 of the images are simultaneously equal to zero (Smith, 2015). However, this situation may not occur in Quan Son Reservoir because the water body of the reservoir does not behave as an absolute blackbody that absorbs all of the light's energy in the visible bands. The model's coefficients (β) indicated that Chla increased with higher B2 values ($t = 11.16$; $p = 0.00$) and decreased with lower B3 ($t = –2.73$; $p = 0.08$) and B4 values ($t = –3.79$; $p = 0.00$). Among these variables, B2 had the most substantial impact on Chla estimation (Table 4). The MLR model for estimating Chla in Quan Son Reservoir can be expressed as:

$$\text{Chla}\ (\mu g/L) = 26.76 + 0.084 * B2 - 0.039 * B3 - 0.031 * B4 \quad (5)$$

Here, Chla is in units of µg/L, while B2, B3, and B4 represent $R_{rs}(\lambda)$ derived from the blue (B2), green (B3), and red (B4) bands of S2B level 2 images. Figure 6 illustrates the performance of Eq. (5) in estimating Chla using a 30-point validated dataset measured on December 19, 2022, and May 18, 2023. The results demonstrate that Eq. (5) is appropriate for estimating Chla in Quan Son Reservoir, with a low error value (RMSE = 1.68), corresponding to approximately 5% of the *in-situ* mean Chla.

Table 4. Coefficients of the model for estimating Chla in Quan Son Reservoir

	Unstandardized Coefficients		Standardized Coefficients	t	p
	B	Standard error	β		
(Constant)	26.76	2.984		8.967	0.000
B2	0.084	0.008	4.578	11.163	0.000
B3	–0.039	0.014	–2.180	–2.732	0.008
B4	–0.031	0.008	–1.651	–3.797	0.000

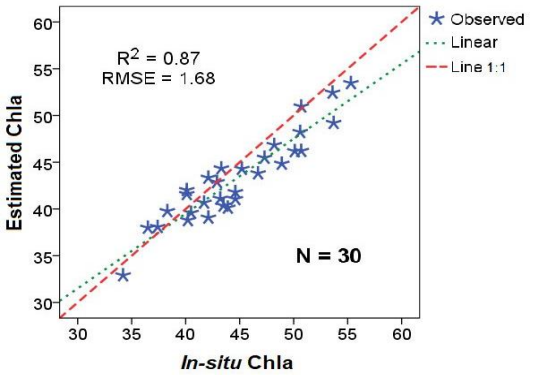

Figure 6. Scatterplot comparing *in-situ* Chla to *estimated* Chla values using Eq. (5) and the calibration dataset

3.4. Variations of Chla in Quan Son Reservoir in space and time

Equation (5) was then applied to assess the spatial and seasonal variations in Chla distribution across Quan Son Reservoir using ten S2B images captured from July 2021 to July 2023. The results, displayed in Fig. 7, reveal that Chla in Quan Son Reservoir ranged from 10.2 µg/L (on December 4, 2021) to 73.9 µg/L (on July 2, 2022), with mean values ranging from 26.4 µg/L (on December 4, 2021) to 56.6 µg/L (on September 20, 2022). Spatially, the distribution of Chla appeared relatively homogeneous, except during May and July. Chla distribution patterns became more pronounced in warmer months with increased precipitation. Higher Chla were observed around Doc Lap Hill, Trau Bac Mountain, and the eastern part of the reservoir, areas known for their higher human

activity, residential areas, and tourism hotspots. Conversely, Chla were comparatively lower in the reservoir's northern and central parts, where is less interfered by human activities. Of particular concern is the estimates in September 2022 when Chla exceeded 60 µg/L, covering the entire reservoir's water surface. According to Carlson and Simpson (1996), Chla levels exceeding 56 µg/L indicate a hypertrophic state, posing a risk of algal blooms. As a result, increasing the frequency of Chla monitoring in the reservoir is imperative during the summer and autumn seasons.

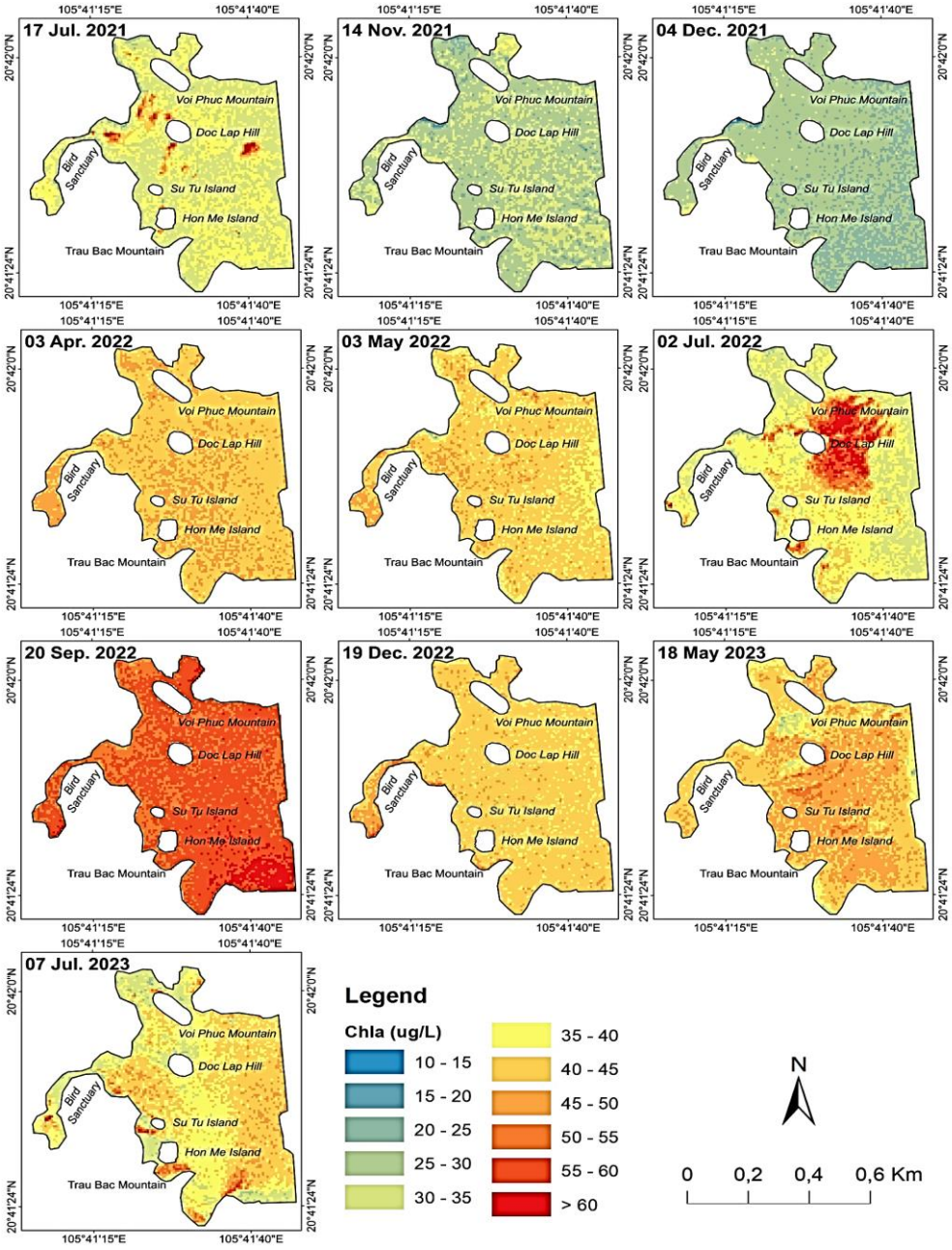

Figure 7. Changes in Chla distribution in Quan Son Reservoir from July 2021 to July 2023

Figure 8 illustrates the temporal variations in mean Chla distribution across Quan Son Reservoir. A discernible trend emerges, with Chla levels being lower in November and December, slightly increasing in April, and reaching peak levels in mid-summer (July) and early autumn (September). This seasonal pattern aligns with the observed Chla patterns in tropical irrigation reservoirs (León et al., 2016). The wide range of *in-situ* Chla in both calibration (ranging from 22.6 to 63.2 µg/L) and validation (ranging from 34.2 to 55.3 µg/L) datasets, taken across different months over an annual cycle, underscores the suitability of the model for estimating Chla in Quan Son Reservoir throughout different periods.

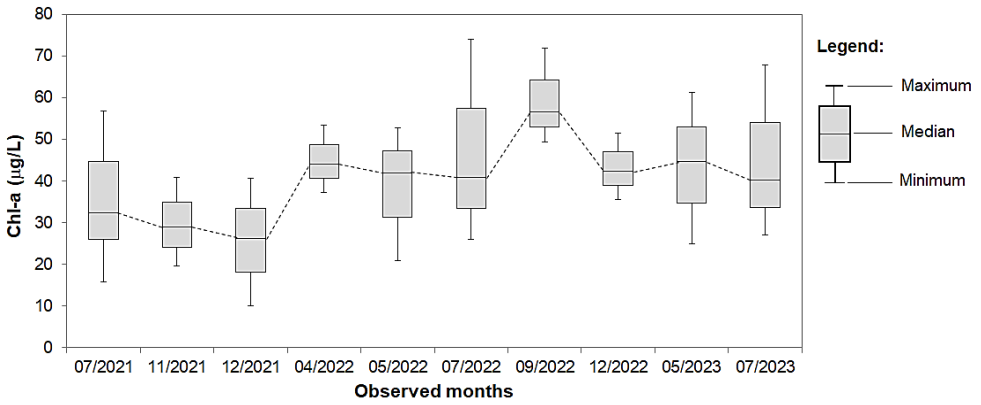

Figure 8. Temporal variations of Chla obtained from ten S2B images

4. Discussions

Our study employed MLR analysis to develop a robust model for estimating Chla in Quan Son Reservoir using S2B imagery and 68-point data. This model, as expressed in Eq. (5), was based on three bands in the visible region of S2B imagery (B2, B3, and B4) and exhibited excellent predictive capabilities, as evidenced by the high correlations and small error in calibration (R^2 = 0.95 and RMSE = 2.76, respectively). The validity of our proposed model was rigorously tested using a separate validation dataset, comprising 30-point data collected on December 19, 2022, and May 18, 2023. The validation results demonstrated a strong agreement between estimated Chla values and *in-situ* measurements (R^2 = 0.87), as indicated by the low RMSE value of 1.68, affirming the model's high performance. The high accuracy and reliability of the Chla estimation model using MLR in this study are aligned with other studies using similar methods and data (Ouma et al., 2020; Barraza-Moraga et al., 2022; Latwal et al., 2023). The superior R^2 value of our proposed model (R^2 = 0.95) is similar to those (R^2 = 0.93–0.97) proposed by Barraza-Moraga et al. (2022). The selection of bands 2, 3, and 4 of S2B level 2 data as variables of the MLR model for estimating Chla in this study demonstrates the great importance of reflectances within the visible region for quantifying Chla in lakes, as demonstrated in many previous works (e.g., Brezonik et al., 2005; Lim and Choi, 2015; Matus-Hernández et al., 2018). Particularly in cases where reflectances retrieved from the S2 NIR bands are often lower than ground-truth data (Warren et al., 2019). The success of our MLR model for estimating Chla in Quan Son Reservoir using S2B imagery carries significant implications. This model is a valuable tool for monitoring water quality within the reservoir, enabling timely responses to changes in Chla, which are essential for managing algal blooms and preserving the

reservoir's ecological health. Our approach aligns with previous studies that have highlighted the effectiveness of MLR models in estimating Chla in various aquatic environments (Brezonik et al., 2005; Kim et al., 2016; Patra et al., 2017; Lim and Choi, 2015; Matus-Hernández et al., 2018). However, it is essential to note a fundamental limitation of our study: the relatively small number of *in-situ* measurements used for calibration and validation (N = 68 and 30, respectively). To enhance the robustness of the model, it would be beneficial to expand validation efforts by incorporating independent data sources.

In this study, we further evaluated the performance of existing atmospheric correction methods for water retrievals from S2 images using *in-situ* $R_{rs}(\lambda)$ measured at 18 points across the reservoir. Three commonly used atmospheric correction methods were assessed, including DOS, iCOR, and Sen2cor. Our results align with previous findings regarding applying these methods for correcting images over subtropical regions (Warren et al., 2019; Ogashawara et al., 2021; Pahlevan et al., 2021). For instance, all three methods failed to retrieve *in-situ* $R_{rs}(\lambda)$s in S2's NIR bands from band 6 to band 8A ($r \approx$ 0.05–0.33; Figure 4). Similarly, our study also found a moderate relationship between *in-situ* $R_{rs}(\lambda)$ and BOA-$R_{rs}(\lambda)$ obtained at band 2 (blue), band 3 (green), band 4 (red), and band 5 (NIR band centered at 705 nm) from these methods ($r \approx$ 0.50–0.76; Fig. 4). This finding indicates that any algorithms relying on the NIR region beyond 740 nm using water surface reflectances are inappropriate for application to the S2 images. Additionally, the reflectances obtained from bands 2, 3, 4, and 5 of S2 data after applying these atmospheric correction methods are promising for water quality monitoring purpose.

The findings of this study highlight the need for more frequent and vigilant monitoring of water quality during the summer-autumn period when Chla tends to increase. Elevated Chla levels, significantly when exceeding 56 µg/L (indicative of a hypertrophic state), raise concerns about potential algal blooms. Such blooms can negatively impact water quality, disrupt ecosystems, and pose health risks to humans and aquatic life.

However, it is essential to note that our Chla estimation model is tailored to Quan Son Reservoir and may not directly apply to other water bodies with different optical properties and geographic conditions. Future research should explore the model's adaptability in diverse settings. To further enhance our understanding of Chla dynamics in Quan Son Reservoir, future studies could investigate the specific sources of nutrient inputs and their seasonal variations. Additionally, assessing the impact of algal blooms on water quality and ecosystem health within the reservoir would provide valuable insights for management and conservation efforts.

S2B data in this study provides valuable insights into the estimation and mapping of Chla in Quan Son Reservoir. However, its revisit frequency of ten days limits the temporal coverage available for this study. Future studies could consider integrating data from other satellite sensors, such as S2A, Landsat 8 and 9, to address this limitation to detect more detailed Chla dynamics. Additionally, continued monitoring with multiple Earth-observing satellites can provide more extended time series data for a more comprehensive analysis of water quality dynamics.

5. Conclusions

This study has successfully developed an empirical MLR model to estimate and map Chla in Quan Son, a small tropical eutrophic reservoir, using S2B Level 2A images. The comprehensive evaluation of all S2B bands

against a calibration dataset of 68-point data of *in-situ* Chla demonstrated the effectiveness of using a combination of the blue (B2), green (B3), and red (B4) bands of S2B for Chla estimation (R^2 = 0.95, RMSE = 2.7 µg/L). Importantly, this MLR model showed strong agreement with 30-point data of *in-situ* Chla collected on December 19, 2022, and May 18, 2023, with RMSE values lower than 5% of the mean *in-situ* Chla value, reaffirming the reliability of our proposed method.

The spatial distribution of Chla in Quan Son Reservoir, derived from ten S2B images acquired on days with minimal cloud cover since July 2021, was effectively mapped. This revealed that the spatial distribution of Chla exhibits variability, particularly during the summer months. Specifically, Chla levels were elevated at sites with tourist activities on the reservoir. Despite the limited number of available S2B images in this study, we were able to elucidate the seasonal, temporal dynamics of Chla in Quan Son Reservoir: Chla levels tend to peak during the summer months (July) and gradually decrease from autumn to lower values in the winter months (November and December).

The resulting spatial distribution maps of Chla over the reservoir underscore the high applicability of our proposed method for monitoring Chla in Quan Son Reservoir. This model contributes to the effective management of water quality and the timely identification of changes in Chla, which is particularly critical in mitigating the risks associated with algal blooms and ensuring the ecological health of the reservoir. However, it is essential to acknowledge that the model's applicability beyond Quan Son Reservoir to water bodies with different optical properties and geographic conditions warrants further investigation in future research. Future studies could also explore specific sources of nutrient inputs and their seasonal variations to enhance further our understanding of Chla dynamics in Quan Son Reservoir.

Acknowledgments

This research is funded by VNU University of Science, Vietnam National University Hanoi, under project number TN.22.12. The author thanks the ESA for providing the satellite data.

References

Adams H., Ye J., Persaud B., Slowinski S., Kheyrollah Pour H., Van Cappellen P., 2021. Chlorophyll-a growth rates and related environmental variables in global temperate and cold-temperate lakes. Earth System Science Data Discussions, 1–30.

Al-Kharusi E.S., Tenenbaum D.E., Abdi A.M., Kutser T., Karlsson J., Bergström A.K., Berggren M., 2020. Large-scale retrieval of coloured dissolved organic matter in northern lakes using Sentinel-2 data. Remote Sensing, 12(1), 157.

APHA, 1998. WPCF, 1998. Standard methods for the examination of water and wastewater, 20, 39–42.

Augusto-Silva P.B., Ogashawara I., Barbosa C.C., De Carvalho L.A., Jorge D.S., Fornari C.I., Stech J.L., 2014. Analysis of MERIS reflectance algorithms for estimating chlorophyll-a concentration in a Brazilian reservoir. Remote Sensing, 6(12), 11689–11707.

Barraza-Moraga F., Alcayaga H., Pizarro A., Félez-Bernal J., Urrutia R., 2022. Estimation of chlorophyll-a concentrations in Lanalhue Lake using Sentinel-2 MSI satellite images. Remote Sensing, 14(22), 5647.

Barsi J.A., Lee K., Kvaran G., Markham B.L., Pedelty J.A., 2014. The spectral response of the Landsat-8 operational land imager. Remote sensing, 6(10), 10232–10251.

Bennett A., Bogorad L., 1973. Complementary chromatic adaptation in a filamentous blue-green alga. The Journal of cell biology, 58(2), 419–435.

Berk A., Anderson G.P., Acharya P.K., Bernstein L.S., Muratov L., Lee J., Shettle E.P., 2006. MODTRAN5: 2006 Update, Defense and Security Symposium. Proceedings of the SPIE, Orlando, FL, USA, 6233.

Brezonik P., Menken K.D., Bauer M., 2005. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored

dissolved organic matter (CDOM). Lake and Reservoir Management, 21(4), 373–382.

Brown C.D., Canfield Jr D.E., Bachmann R.W., Hoyer M.V., 1998. Seasonal patterns of chlorophyll, nutrient concentrations and Secchi disk transparency in Florida lakes. Lake and Reservoir Management, 14(1), 60–76.

Bui Q.T., Jamet C., Vantrepotte V., Mériaux X., Cauvin A., Mograne M.A., 2022. Evaluation of sentinel-2/MSI atmospheric correction algorithms over two contrasted French coastal waters. Remote Sensing, 14(5), 1099.

Cao Z., Ma R., Duan H., Pahlevan N., Melack J., Shen M., Xue K., 2020. A machine learning approach to estimate chlorophyll-a from Landsat-8 measurements in inland lakes. Remote Sensing of Environment, 248, 111974.

Carlson R.E., 1977. A trophic state index for lakes 1. Limnology and oceanography, 22(2), 361-369.

Carlson R.E., 1996. A coordinator's guide to volunteer lake monitoring methods. North American Lake Management Society, 96, 305.

Chavez Jr P.S., 1988. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. Remote sensing of environment, 24(3), 459–479.

Chen J., Zhu W., Tian Y.Q., Yu Q., Zheng Y., Huang L., 2017. Remote estimation of colored dissolved organic matter and chlorophyll-a in Lake Huron using Sentinel-2 measurements. Journal of Applied Remote Sensing, 11(3), 036007–036007.

Chen Z., Curran P.J., Hansom J.D., 1992. Derivative reflectance spectroscopy to estimate suspended sediment concentration. Remote Sensing of Environment, 40(1), 67–77.

Cho K.H., Kang J.H., Ki S.J., Park Y., Cha S.M., Kim J.H., 2009. Determination of the optimal parameters in regression models for the prediction of chlorophyll-a: A case study of the Yeongsan Reservoir, Korea. Science of the total environment, 407(8), 2536–2545.

Chu H.J., He Y.C., Chusnah W.N.U., Jaelani L.M., Chang C.H., 2021. Multi-reservoir water quality mapping from remote sensing using spatial regression. Sustainability, 13(11), 6416.

Cloern J.E., Foster S.Q., Kleckner A.E., 2014. Phytoplankton primary production in the world's estuarine-coastal ecosystems. Biogeosciences, 11(9), 2477–2501.

Conkright M.E., Gregg W.W., 2003. Comparison of global chlorophyll climatologies: In situ, CZCS, Blended in situ-CZCS and SeaWiFS. International Journal of Remote Sensing, 24(5), 969–991.

Dall'Olmo G., Gitelson A.A., 2005. Effect of bio-optical parameter variability on the remote estimation of chlorophyll-a concentration in turbid productive waters: experimental results. Applied optics, 44(3), 412–422.

De Keukelaere L., Sterckx S., Adriaensen S., Knaeps E., Reusen I., Giardino C., Vaiciute D., 2018. Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: validation for coastal and inland waters. European Journal of Remote Sensing, 51(1), 525–542.

Franklin J.B., Sathish T., Vinithkumar N.V., Kirubagaran R., 2020. A novel approach to predict chlorophyll-a in coastal-marine ecosystems using multiple linear regression and principal component scores. Marine pollution bulletin, 152, 110902.

Fujiki T., Taguchi S., 2002. Variability in chlorophyll a specific absorption coefficient in marine phytoplankton as a function of cell size and irradiance. Journal of Plankton Research, 24(9), 859–874.

Gholizadeh M.H., Melesse A.M., Reddi L., 2016. A comprehensive review on water quality parameters estimation using remote sensing techniques. Sensors, 16(8), 1298.

Gitelson A., 1992. The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. International Journal of Remote Sensing, 13(17), 3367–3373.

Gitelson A.A., Gao B.C., Li R.R., Berdnikov S., Saprygin V., 2011. Estimation of chlorophyll-a concentration in productive turbid waters using a Hyperspectral Imager for the Coastal Ocean the Azov Sea case study. Environmental Research Letters, 6(2), 024023.

Grendaitė D., Stonevičius E., 2022. Uncertainty of atmospheric correction algorithms for chlorophyll α concentration retrieval in lakes from Sentinel-2 data. Geocarto International, 37(23), 6867–6891.

Guanter L., 2023. New Algorithms for Atmospheric Correction and Retrieval of Biophysical Parameters in Earth Observation Application to ENVISAT/MERIS Data. In Departament de Física de la Terra i Termodinàmica; Universitat de

Valéncia: Valéncia, Spain, 2007; Available online: https://www.tesisenred.net/handle/10803/9877;jsessionid=402C46F935666E5FF2CEF46F057954A7.tdx2#page=1 (accessed on August 6 2023).

Ha N.T.T., Thao N.T.P., Koike K., Nhuan M.T., 2017. Selecting the best band ratio to estimate chlorophyll-a concentration in a tropical freshwater lake using sentinel 2A images from a case study of Lake Ba Be (Northern Vietnam). ISPRS International Journal of Geo-Information, 6(9), 290.

Ha Tay Provincal Committee, 2008. Decision No. 462/QD-UBND approving detailed planning for construction at 1/2000 scale of Quan Son Reservoir urban eco-tourism area, My Duc district.

Hunter P.D., Tyler A.N., Présing M., Kovács A.W., Preston T., 2008. Spectral discrimination of phytoplankton colour groups: The effect of suspended particulate matter and sensor spectral resolution. Remote Sensing of Environment, 112(4), 1527–1544.

Jiang D., Matsushita B., Setiawan F., Vundo A., 2019. An improved algorithm for estimating the Secchi disk depth from remote sensing data based on the new underwater visibility theory. ISPRS journal of photogrammetry and remote sensing, 152, 13–23.

Kasprzak P., Padisák J., Koschel R., Krienitz L., Gervais F., 2008. Chlorophyll a concentration across a trophic gradient of lakes: An estimator of phytoplankton biomass?. Limnologica, 38(3–4), 327–338.

Kim H.H., Ko B.C., Nam J.Y., 2016b. Predicting chlorophyll-a using Landsat 8 OLI sensor data and the non-linear RANSAC method a case study of Nakdong River, South Korea. International Journal of Remote Sensing, 37(14), 3255–3271.

Kim W., Moon J.E., Park Y.J., Ishizaka J., 2016a. Evaluation of chlorophyll retrievals from Geostationary Ocean color imager (GOCI) for the north-east Asian region. Remote Sensing of Environment, 184, 482–495.

Kim Y.W., Kim T., Shin J., Lee D.S., Park Y.S., Kim Y., Cha Y., 2022. Validity evaluation of a machine-learning model for chlorophyll a retrieval using Sentinel-2 from inland and coastal waters. Ecological Indicators, 137, 108737.

Kravitz J., Matthews M., Bernard S., Griffith D., 2020. Application of Sentinel 3 OLCI for chl-a retrieval over small inland water targets: Successes and challenges. Remote Sensing of Environment, 237, 111562.

Kuhn C., de Matos Valerio A., Ward N., Loken L., Sawakuchi H.O., Kampel M., Richey J., Stadler P., Crawford J., Striegl R., Vermote E., Pahlevan N., Butman D., 2019. Performance of Landsat-8 and Sentinel-2 surface reflectance products for river remote sensing retrievals of chlorophyll-a and turbidity. Remote Sensing of Environment, 224, 104–118.

Latwal A., Rehana S., Rajan K.S., 2023. Detection and mapping of water and chlorophyll-a spread using Sentinel-2 satellite imagery for water quality assessment of inland water bodies. Environmental Monitoring and Assessment, 195(11), 1304.

Le C., Hu C., Cannizzaro J., Duan H., 2013a. Long-term distribution patterns of remotely sensed water quality parameters in Chesapeake Bay. Estuarine, Coastal and Shelf Science, 128, 93–103.

Le C., Hu C., Cannizzaro J., English D., Muller-Karger F., Lee Z., 2013b. Evaluation of chlorophyll-a remote sensing algorithms for an optically complex estuary. Remote Sensing of Environment, 129, 75–89.

Lehmann M.K., Gurlin D., Pahlevan N., Alikas K., Conroy T., Anstee J., Balasubramanian S.V., Barbosa C.C.F., Binding C., Bracher A., Bresciani M., Burtner A., Cao Z., Dekker A.G., Vittorio C.D., Drayson N., Errera R.M., Fernandez V., Ficek D., Fichot C.G., Gege P., Giardino C., Gitelson A.A., Greb S.R., Henderson H., Higa H., Rahaghi A.I., Jamet C., Jiang D., Jordan T., Kangro K., Kravitz J.A., Kristoffersen A.S., Kudela R., Li L., Ligi M., Loisel H., Lohrenz S., Ma R., Maciel D.A., Malthus T.J., Matsushita B., Matthews M., Minaudo C., Mishra D.R., Mishra S., Moore T., Moses W.J., Hà N., Novo E.M.L.M, Novoa S., Odermatt D., O'Donnell D.M., Olmanson L.G., Ondrusek M., Oppelt N., Ouillon S., Filho W.P., Plattner S., Verdú A.R., Salem S.I., Schalles J.F., Simis S.G.H., Siswanto E., Smith B., Somlai-Schweiger I., Soppa M.A., Spyrakos E., Tessin E., van der Woerd H.J., Woude A.V., Vandermeulen R.A., Vantrepotte V., Wernand M.R., Werther M., Young K., Yue L., 2023. GLORIA-A globally representative hyperspectral in situ dataset for optical sensing of water quality. Scientific Data, 10(1), 100.

León J.G., Beamud S.G., Temporetti P.F., Atencio A.G., Diaz M.M., Pedrozo F.L., 2016. Stratification and

residence time as factors controlling the seasonal variation and the vertical distribution of chlorophyll-a in a subtropical irrigation reservoir. International Review of Hydrobiology, 101(1–2), 36–47.

Li J., Gao M., Feng L., Zhao H., Shen Q., Zhang F., Wang S., Zhang B., 2019. Estimation of chlorophyll-a concentrations in a highly turbid eutrophic lake using a classification-based MODIS land-band algorithm. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(10), 3769–3783.

Lim J., Choi M., 2015. Assessment of water quality based on Landsat 8 operational land imager associated with human activities in Korea. Environmental monitoring and assessment, 187, 1–17.

Lins R.C., Martinez J.M., Motta Marques D.D., Cirilo J.A., Medeiros P.R.P., Fragoso Júnior C.R., 2018. A multivariate analysis framework to detect key environmental factors affecting spatiotemporal variability of chlorophyll-a in a tropical productive estuarine-lagoon system. Remote Sensing, 10(6), 853.

Liu X., Steele C., Simis S., Warren M., Tyler A., Spyrakos E., Selmes N., Hunter P., 2021. Retrieval of Chlorophyll-a concentration and associated product uncertainty in optically diverse lakes and reservoirs. Remote Sensing of Environment, 267, 112710.

Liu Y., Xiao C., Li J., Zhang F., Wang S., 2020. Secchi disk depth estimation from China's new generation of GF-5 hyperspectral observations using a semi-analytical scheme. Remote Sensing, 12(11), 1849.

Loisel H., Vantrepotte V., Ouillon S., Ngoc D.D., Herrmann M., Tran V., Mériaux X., Dessailly D., Jamet C., Duhaut T., Nguyen H.H., Van Nguyen T., 2017. Assessment and analysis of the chlorophyll-a concentration variability over the Vietnamese coastal waters from the MERIS ocean color sensor (2002–2012). Remote sensing of environment, 190, 217–232.

Louis J., Debaecker V., Pflug B., Main-Knorn M., Bieniarz J., Mueller-Wilm U., Cadau E., Gascon F. Sentinel-2 Sen2Cor: L2A processor for users. In Proceedings living planet symposium 2016. Spacebooks Online. 9–13 May 2016, Prague, Czech Republic, 1–8.

Ma R., Ma X., Dai J., 2007. Hyperspectral feature analysis of chlorophyll a and suspended solids using field measurements from Taihu Lake, eastern China. Hydrological Sciences Journal, 52(4), 808–824.

Ma R.H., Dai J.F., 2005. Chlorophyll-a concentration estimation with field spectra of waterbody near Meiliang Bayou in Taihu Lake. Journal of remote sensing, 9(1), 78–86.

Martins V.S., Barbosa C.C.F., De Carvalho L.A.S., Jorge D.S.F., Lobo F.D.L., Novo E.M.L.D.M., 2017. Assessment of atmospheric correction methods for Sentinel-2 MSI images applied to Amazon floodplain lakes. Remote Sensing, 9(4), 322.

Matus-Hernández M.Á., Hernández-Saavedra N.Y., Martínez-Rincón R.O., 2018. Predictive performance of regression models to estimate Chlorophyll-a concentration based on Landsat imagery. Plos One, 13(10), e0205682.

Mishra S., Mishra D.R., 2012. Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. Remote Sensing of Environment, 117, 394–406.

Mitchell B.G., Kiefer D.A., 1988. Chlorophyll α specific absorption and fluorescence excitation spectra for light-limited phytoplankton. Deep Sea Research Part A. Oceanographic Research Papers, 35(5), 639–663.

Mobley C.D., 1999. Estimation of the remote-sensing reflectance from above-surface measurements. Applied optics, 38(36), 7442–7455.

Moore T.S., Dowell M.D., Bradt S., Verdu A.R., 2014. An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. Remote sensing of Environment, 143, 97–111.

Moses W.J., Gitelson A.A., Berdnikov S., Saprygin V., Povazhnyi V., 2012. Operational MERIS-based NIR-red algorithms for estimating chlorophyll-a concentrations in coastal waters the Azov Sea case study. Remote Sensing of Environment, 121, 118–124.

Mueller J.L., Morel A., Frouin R., Davis C., Arnone R., Carder K., Lee Z.P., 2003. Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocols. Goddard Space Flight Center: Greenbelt, USA.

Nguyen X.H., Dao T.N., Nguyen T.N., 2010. The fish species composition in the area of Quan Son reservoir in My Duc district, Ha Noi. Journal of

Science, Natural Sciences and Technology, VNU, Hanoi, 26, 531–536.

Niroumand-Jadidi M., Bovolo F., Bruzzone L., Gege P., 2021. Inter-comparison of methods for chlorophyll-a retrieval: Sentinel-2 time-series analysis in Italian lakes. Remote Sensing, 13(12), 2381.

Ogashawara I., Alcântara E.H., Curtarelli M.P., Adami M., Nascimento R.F., Souza A.F., Stech J.L., Kampel M., 2014. Performance analysis of MODIS 500-m spatial resolution products for estimating chlorophyll-a concentrations in oligo-to meso-trophic waters case study: Itumbiara Reservoir, Brazil. Remote Sensing, 6(2), 1634–1653.

Ogashawara I., Kiel C., Jechow A., Kohnert K., Ruhtz T., Grossart H.P., Hölker F., Nejstgaard J.C., Berger S.A., Wollrab S., 2021. The use of Sentinel-2 for chlorophyll-a spatial dynamics assessment: A comparative study on different lakes in northern Germany. Remote Sensing, 13(8), 1542.

Ouma Y.O., Noor K., Herbert K., 2020. Modelling reservoir chlorophyll-a, TSS, and turbidity using Sentinel-2A MSI and Landsat-8 OLI satellite sensors with empirical multivariate regression. Journal of Sensors, 1–21.

Pahlevan N., Mangin A., Balasubramanian S.V., Smith B., Alikas K., Arai K., Barbosa C., Bélanger S., Binding C., Bresciani M., Giardino C., Gurlin D., Fan Y., Harmel T., Hunter P., Ishikaza J., Kratzer S., Lehmann M.K., Ligi M., Ma R., Martin-Lauzer F.R., Olmanson L., Oppelt N., Pan Y., Peters S., Reynaud N., Sander de Carvalho L.A., Simis S., Spyrakos E., Steinmetz F., Stelzer K., Sterckx S., Tormos T., Tyler A., Vanhellemont Q., Warren M., 2021. ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. Remote Sensing of Environment, 258, 112366.

Pahlevan N., Sarkar S., Franz B.A., Balasubramanian S.V., He J., 2017. Sentinel-2 MultiSpectral Instrument (MSI) data processing for aquatic science applications: Demonstrations and validations. Remote sensing of environment, 201, 47–56.

Palmer S.C., Kutser T., Hunter P.D., 2015. Remote sensing of inland waters: Challenges, progress and future directions. Remote sensing of Environment, 157, 1–8.

Patra P.P., Dubey S.K., Trivedi R.K., Sahu S.K., Rout S.K., 2017. Estimation of chlorophyll-a concentration and trophic states in Nalban Lake of East Kolkata Wetland, India from Landsat 8 OLI data. Spatial Information Research, 25, 75–87.

Pereira-Sandoval M., Ruescas A., Urrego P., Ruiz-Verdú A., Delegido J., Tenjo C., Soria-Perpinyà X., Vicente E., Soria J., Moreno J., 2019. Evaluation of atmospheric correction algorithms over Spanish inland waters for Sentinel-2 multi spectral imagery data. Remote Sensing, 11(12), 1469.

Perrone M., Scalici M., Conti L., Moravec D., Kropáček J., Sighicelli M., Lecce F., Malavasi M., 2021. Water mixing conditions influence Sentinel-2 monitoring of chlorophyll content in monomictic lakes. Remote Sensing, 13(14), 2699.

Pflug B., Louis J., Debaecker V., Mueller-Wilm U., Quang C., Gascon F., Boccia V., 2020. Next updates of atmospheric correction processor Sen2Cor. In Proceedings Volume 11533, Image and Signal Processing for Remote Sensing XXVI. SPIE Remote Sensing, 20 September 2020, 1153304.

Pu F., Ding C., Chao Z., Yu Y., Xu X., 2019. Water-quality classification of inland lakes using Landsat8 images by convolutional neural networks. Remote Sensing, 11(14), 1674.

Quan Sơn Fisheries and Tourism Joint Stock Company, 2013. Environmental Protection Project Report.

Rundquist D.C., Han L., Schalles J.F., Peake J.S., 1996. Remote measurement of algal chlorophyll in surface waters: the case for the first derivative of reflectance near 690 nm. Photogrammetric Engineering and Remote Sensing, 62(2), 195–200.

Schalles J.F., 2006. Optical remote sensing techniques to estimate phytoplankton chlorophyll a concentrations in coastal. In Remote sensing of aquatic coastal ecosystem processes. Dordrecht: Springer Netherlands, 27, 27–79.

Schalles J.F., Sheil A.T., Tycast J.F., Alberts J.J., Yacobi Y.Z., 1998. Detection of chlorophyll, seston, and dissolved organic matter in the estuarine mixing zone of Georgia coastal plain rivers. In 5th International Conference on Remote Sensing for Marine and Coastal Environments, 2, 315–324.

Sent G., Biguino B., Favareto L., Cruz J., Sá C., Dogliotti A.I., Palma C., Brotas V., Brito A.C., 2021. Deriving water quality parameters using sentinel-2 imagery: A case study in the Sado Estuary, Portugal. Remote sensing, 13(5), 1043.

Smith G., 2015. Essential statistics, regression, and econometrics. Academic Press, 386p.

Sterckx S., Knaeps S., Kratzer S., Ruddick K., 2015. SIMilarity Environment Correction (SIMEC) applied to MERIS data over inland and coastal waters. Remote Sensing of Environment, 157, 96–110.

Tang F., Ohto T., Sun S., Rouxel J.R., Imoto S., Backus E.H., Mukamel S., Bonn M., Nagata Y., 2020. Molecular structure and modeling of water-air and ice-air interfaces monitored by sum-frequency generation. Chemical reviews, 120(8), 3633–3667.

Tehrani N.A., Janalipour M., Babaei H., 2021. Estimating water surface chlorophyll-a concentration by big remote sensing data in the Persian Gulf, Bushehr. Remote Sensing in Earth Systems Sciences, 4(1-2), 87–95.

Thao N.T.P., Thang P.D., Hien T.T., Ha N.T.T., Vinh P.Q., 2023. Assessing and modelling the trophic state of Quan Son Reservoir in space and time. Vietnam Journal of Hydro-Meteorology, 748, 32–41.

USEPA, 2009. Environmental Protection Agency (USEPA). National Lakes Assessment: A Collaborative Survey of the Nation's Lakes; EPA 841-R-09-001; U.S. Environmental Protection Agency, Office of Water and Office of Research and Development: Washington, DC, USA, 2009; 103p. Available online: https://www.epa.gov/sites/production/files/2013-11/documents/nla_newlowres_fullrpt.pdf (assessed on July 4 2023).

UNEP, 2014. Review of Existing Water Quality Guidelines for Freshwater Ecosystems and Application of Water Quality Guidelines on Basin Level to Protect Ecosystems. In Proceedings of the First International Environment Forum for Basin Organizations towards Sustainable Freshwater Governance, Nairobi, Kenya, 26–28 November 2014; Available online: https://wedocs.unep.org/rest/bitstreams/35090/retrieve (assessed on July 4 2023).

Vilas L.G., Spyrakos E., Palenzuela J.M.T., 2011. Neural network estimation of chlorophyll a from MERIS full resolution data for the coastal waters of Galician rias (NW Spain). Remote Sensing of Environment, 115(2), 524–535.

Vinh P.Q., Ha N.T.T., Binh N.T., Thang N.N., Oanh L.T., Thao N.T.P., 2019. Developing algorithm for estimating chlorophyll-a concentration in the Thac Ba Reservoir surface water using Landsat 8 Imagery. Vietnam J. Earth Sci., 41(1), 10–20. https://doi.org/10.15625/0866-7187/41/1/13542.

Warren M.A., Simis S.G., Martinez-Vicente V., Poser K., Bresciani M., Alikas K., Spyrakos E., Giardino C., Ansper A., 2019. Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters. Remote Sensing of Environment, 225, 267–289.

Werther M., Spyrakos E., Simis S.G., Odermatt D., Stelzer K., Krawczyk H., Berlage O., Hunter P., Tyler A., 2021. Meta-classification of remote sensing reflectance to estimate trophic status of inland and nearshore waters. ISPRS Journal of Photogrammetry and Remote Sensing, 176, 109–126.

Werther M., Odermatt D., Simis S.G., Gurlin D., Lehmann M.K., Kutser T., Gupana R., Varley A., Hunter P., Tyler A., Spyrakos E., 2022. A Bayesian approach for remote sensing of chlorophyll-a and associated retrieval uncertainty in oligotrophic and mesotrophic lakes. Remote Sensing of Environment, 283, 113295.

Wu M., Zhao Y., Sun L.E., Huang J., Wang X., Ma Y., 2020. Remote sensing of spatial-temporal variation of chlorophyll-a in the Jiaozhou bay using 32 years Landsat data. Journal of Coastal Research, 102(SI), 271–279.

Ylöstalo P., Kallio K., Seppälä J., 2014. Absorption properties of in-water constituents and their variation among various lake types in the boreal region. Remote sensing of Environment, 148, 190–205.

Zhang S., Liu N., Luo M., Jiang T., Chan T.O., Yau C.S.T., Sun Y, 2023. Downscaling Sentinel-3 Chlorophyll-a Concentration for Inland Lakes based on Multivariate Analysis and Gradient Boosting Decision Trees Regression. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 16, 7850–7865.