Remote Sensing Applications: Society and Environment 36 (2024) 101317

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

journal homepage: www.elsevier.com/locate/rsase

Obtaining estimation algorithms for water quality variables in the Jaguari-Jacareí Reservoir using Sentinel-2 images

Zahia Catalina Merchan Camargo [a, *], Xavier Sòria-Perpinyà [b], Marcelo Pompêo [c], Viviane Moschini-Carlos [a, d], Maria Dolores Sendra [d]

[a] University of São Paulo State (UNESP), ICT, Sorocaba Campus, Av. Três de Março, 511 - Alto da Boa Vista, Sorocaba, SP, 18087-180, Brazil
[b] Image Processing Laboratory, University of València, Edifici E4, 4a planta, C/Catedrático Agustín Escardino, 9, 46980, Paterna, Spain
[c] Department of Ecology, University of São Paulo, R. do Matão, 321 - Butantã, São Paulo, SP, 05508-090, Brazil
[d] Cavanilles Institute of Biodiversity and Evolutionary Biology (ICBiBE), Universitat de València, C/ Catedrático José Beltrán Martínez, nº 2 46980, Paterna, València, Spain

ARTICLE INFO

Keywords:
Remote sensing
Reservoirs
Water quality
Sentinel-2
Eutrophication
Cyanobacteria

ABSTRACT

Satellite images are essential tools for monitoring aquatic ecosystems and assessing water quality, as they enable the measurement of parameters such as chlorophyll-*a* (Chl-*a*) concentration, phycocyanin (PC), and cyanobacteria density. These indicators aid in evaluating eutrophication processes and detecting cyanobacteria in aquatic ecosystems. This study utilized field data and images captured by the Sentinel-2 sensor from 2015 to 2022 to investigate the Jaguari-Jacareí reservoirs (JAG-JAC). Two atmospheric corrections from the Case 2 Regional Coast Color (C2RCC) processor, namely C2X and C2XC, were applied, and algorithms were developed to estimate the parameters using both *in situ* data measurements and reflectance data extracted from the images. For Chl-*a* concentration, the dataset was divided into two blocks: one for model calibration (70% of the data) and the other for validation (30% of the data). As for PC, the entire dataset was utilized to calibrate the model, and validation was conducted through cross-validation using the Automated Radiative Transfer Model Operator (ARTMO) software. Cyanobacteria density was indirectly estimated from the Chl-*a* concentrations determined in the field samples, as these variables exhibited a strong correlation, also validating the model previously proposed for the Cantareira system for estimating cyanobacteria density from Chl-*a* data. Additionally, the automatic chlorophyll-*a* products (con_chla) derived from the C2X and C2XC processors were validated. The findings revealed that the C2X processor exhibited the greatest potential for estimating water quality parameters. It was observed that the most effective algorithms were derived using the R705/R665 band ratio for Chl-*a* and the R705/R490 ratio for PC. For cyanobacteria density, the optimal algorithm was established based on the relationship between cyanobacteria density and Chl-*a* using the data obtained in this study.

1. Introduction

Reservoirs are impacted by various anthropogenic actions, such as land occupation and inadequate land use, as well as the discharge of untreated sanitary and industrial sewage. These actions interfere with the natural water flow, increase sedimentation rates,

* Corresponding author.
E-mail addresses: zahia.merchan@unesp.br (Z.C. Merchan Camargo), soperja@uv.es (X. Sòria-Perpinyà), mpompeo@ib.usp.br (M. Pompêo), viviane.moschini@unesp.br (V. Moschini-Carlos), sendrac@uv.es (M.D. Sendra).

https://doi.org/10.1016/j.rsase.2024.101317
Received 6 March 2024; Received in revised form 16 July 2024; Accepted 3 August 2024
Available online 8 August 2024

and modify nutrient cycling (Carvalho et al., 2013; Barbosa et al., 2012). Artificial nutrient enrichment, particularly phosphorus and nitrogen, is known as eutrophication and has historically been considered a major concern in the management of aquatic ecosystems (Yang et al., 2008; Pompeo and Moschini-Carlos, 2020; Torremorell et al., 2021). Eutrophication leads to changes in physical, chemical, and biological characteristics, promoting excessive algae growth and recurrent cyanobacterial blooms (Michalak et al., 2013; Chaffin and Bridgeman, 2013), with ecological consequences and direct impacts on water quality, supply, and human health (Zorzal-Almeida et al., 2018).

These algal blooms, including cyanobacterial species, are becoming increasingly frequent and widespread worldwide. Their presence is mainly associated with eutrophication and climate change, particularly global warming (Glibert, 2019). The massive growth of cyanobacteria can cause ecological changes in water bodies and produce toxic components (cyanotoxins), which are harmful to human health, aquatic life, and the environment, leading to negative environmental and economic impacts (Zanchett and Oliveira-Filho, 2013; Pearson et al., 2016).

Chlorophyll-*a* (Chl-*a*) concentration has been used to assess the magnitude of algal blooms and determine water quality and eutrophication levels (Carvalho et al., 2013; Moschini-Carlos et al., 2017). However, Chl-*a* is unable to distinguish between different phytoplankton groups (Damar et al., 2020). Therefore, other indicators, such as phycocyanin (PC) concentration, are used. PC is a functional protein and an accessory pigment of Chl-*a* commonly associated with cyanobacteria (Glazer, 1989). And it is a good indicator for estimating the abundance of cyanobacteria in inland waters (Brient et al., 2007).

Considering the importance of monitoring the ecological status and water quality of aquatic ecosystems, remote sensing tools, particularly satellite images, have been implemented. Remote techniques offer the advantage of obtaining data more frequently, at a lower cost, and with greater spatial and temporal coverage (Campbell et al., 2011). This makes large-scale observations of water quality possible, enabling comprehensive spatiotemporal analysis (Hou et al., 2022). The estimation of Chl-*a* and PC is one of the most established applications of remote sensing (Matthews, 2010; Yan et al., 2018). Several remote estimation approaches for Chl-*a* and PC are based on empirical relationships between reflectance in sensor bands or band ratios and *in situ* parameter measurements (Woźniak et al., 2016). This is because the optical properties of water and its components are related to spectral measurements (Bocharov et al., 2017). The development of empirical and specific algorithms for limnological variables based on *in situ* data has shown great potential in the study of aquatic ecosystems (Bramich et al., 2021; Toming et al., 2016).

There is a need for algorithm development to cover as many water characteristics as possible (Li and Song, 2017). Particularly, studies with remote sensing tools in tropical inland waters have been less extended (Llodrà-Llabrés et al., 2023). The region's climatic complexity and characteristics influenced by latitude and irradiance directly impact photosynthesis and productivity in aquatic ecosystem (Lewis, 2002). Additionally, tropical, and subtropical waters exhibit significant differences in remote spectral reflectance (Rrs) due to the variability of water components (Da Silva et al., 2020). However, estimating water quality variables from satellite images requires an accurate assessment of the radiance coming from the water. It is important to note that the radiation measured by sensors includes a significant contribution from the atmosphere, consisting of atmospheric gases and aerosols, which account for at least 90% of the signal measured by satellite sensors (Delegido et al., 2019).

Therefore, atmospheric correction is a crucial step in removing the atmospheric contribution. This underscores the ongoing need to validate different atmospheric correction approaches, considering the wide variety of water types, locations, and environmental conditions (Soriano-González et al., 2022; Li et al., 2023). Nevertheless, Virdis et al. (2022) have demonstrated the potential and effectiveness of the Case2 eXtreme (C2X) neural network in tropical inland waters with mesotrophic and eutrophic conditions.

In this context, this study aims to assess the performance of Sentinel-2 (S2) in a tropical reservoir for obtaining three water quality algorithms using S2 images in the Jaguari-Jacareí reservoir (JAG-JAC). First, Chl-a will be estimated by correlating *in situ* measurements with atmospherically corrected S2 image reflectance data with two neural networks, whose automatic water quality products will also be tested, and then the most accurate algorithm will be selected. Once the most appropriate neural network has been selected, it was used to retrieve algorithms for PC and cyanobacteria density, developed correlating *in situ* measurements with reflectance data.

2. Materials and methods

2.1. Study area

The JAG-JAC System is in the state of São Paulo, Brazil, between coordinates 22°53′20″ S and 46°24′49″ W (Fig. 1). It is in a tropical climate zone with an average annual rainfall of 1400 mm. May and June are typically drier and cooler, while November and January are the rainy season with higher temperatures (Pompêo et al., 2017).

The system has two reservoirs, Jaguari (JAG) and Jacareí (JAC), which are connected by a 25 m wide and 130 m long channel. Although each reservoir has its dam, they are considered a single body of water (Fig. 1).

Fig. 2 illustrates the workflow of the process used for estimating the three water quality parameters using S2 images.

2.2. Field data collection and laboratory measurements

This study used data from two sources. Firstly, water quality reports from the São Paulo State Environmental Company (CETESB. Companhia Ambiental do Estado de São Paulo. 2019-2022 available at https://cetesb.sp.gov.br/aguas-interiores/publicacoes-e-relatorios). Secondly, our own database included three field expeditions conducted in October 2021, June, and November 2022.

From the CETESB reports, 15 sampling dates were selected, covering the period from 2015 to 2021. These reports provided data from two sampling points, one in JAG and one in JAC, as shown in Fig. 1. CETESB monitors surface water in these reservoirs every two months, with collection points located in the central area of JAG (22°55′40″S, 46°25′27″ W) and JAC (22°58′16″S, 46°24′03″W).

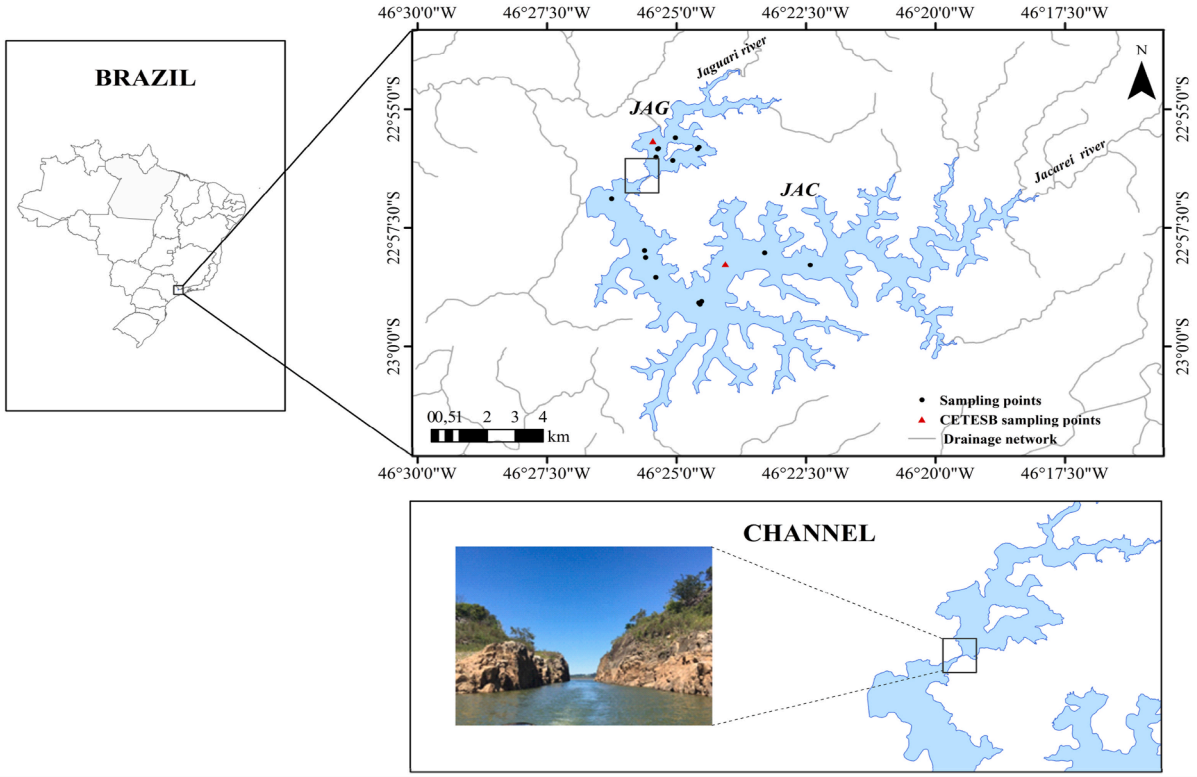

Fig. 1. Geographical location of the Jaguari-Jacareí reservoir system and location of the sampling points in the different campaigns and a view of the interconnection channel between the Jacareí and Jaguari reservoirs.

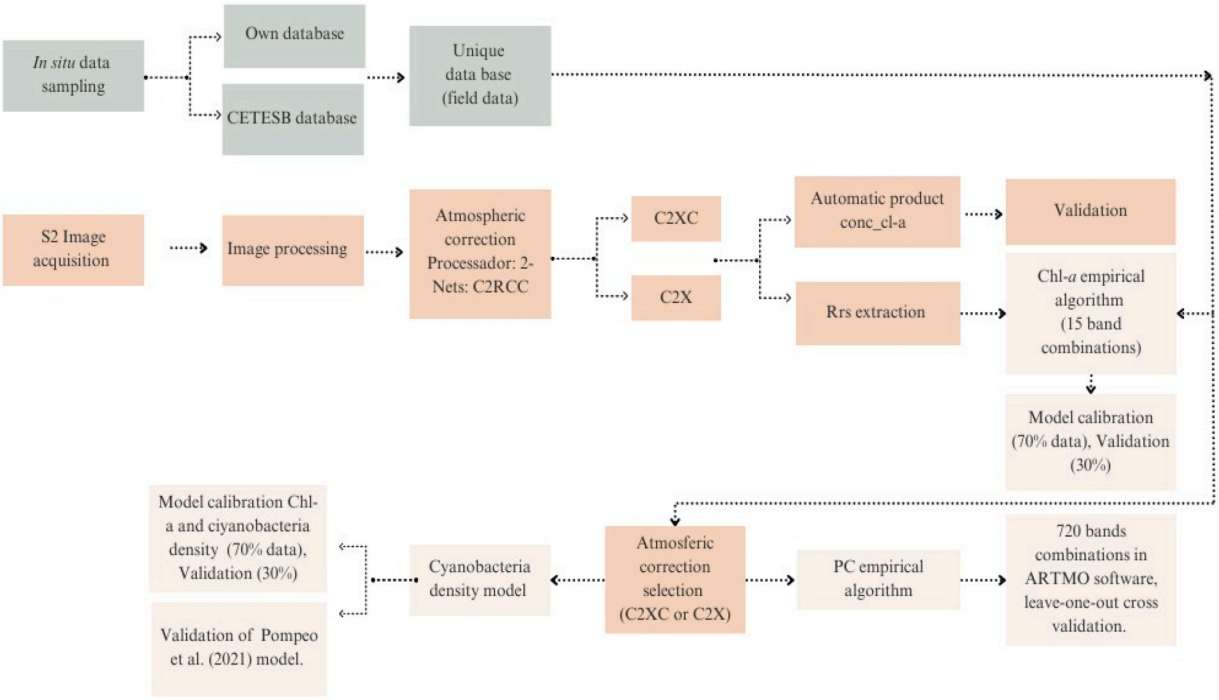

Fig. 2. Workflow of quality parameters estimation using S2.

The data include Secchi disk (SD) values, Chl-*a* concentration (measured spectrophotometrically using 90% acetone (Lorenzen, 1967)), and cyanobacteria density, determined by cell counting per ml following the methodology of Chorus and Bartram (1999) and CETESB Technical Standard L5.303 (CETESB. Companhia Ambiental do Estado de São Paulo, 2012).

In each field campaign, the different points collected in the reservoirs were georeferenced after anchoring. Subsequently, the depth of the Secchi disk (a 20 cm diameter white disk) was measured. Then, water samples were collected with a Van Dorn bottle to the depth of the Secchi disk, and subsequently stored at low temperature and in the dark. The concentration values of phycocyanin (PC) in the water column were also recorded using the submersible fluorometer TURNER C3 (Turner Design Instruments; San Jose, CA, USA).

The fluorescence intensity at the Ex/Em = 590 ± 30/≥645 nm pair was calibrated with a standard phycocyanin extract from Spirulina sp. (Sigma-Aldrich Chemicals). Thus, the PC concentration was calculated by multiplying the electrical signal value obtained by the C3 by a factor of 0.7. For each sampling station, the average PC concentration from the surface to the depth of the Secchi disk was used. Additionally, the Chl-*a* concentration was measured in the laboratory using the spectrophotometric extraction method (DMSO + acetone 90%) according to the method of Shoaf and Lium (1976), and the calculation was performed following Jeffrey and Humphrey (1975).

Phytoplankton samples were fixed in the field with Lugol's solution, sedimented, and counted using a Nikon-ECLIPSE TE-2000S inverted microscope in the laboratory according to the Utermöhl method(Utermöhl, 1958). Phytoplankton organisms were identified at the best possible taxonomic level, and biovolume was calculated by producing a simple or composite approximate geometric shape for each species (Hillebrand et al., 1999; Sun and Liu, 2003; Fonseca et al., 2014).

The Pearson correlation coefficient was determined to evaluate the association between variables, and normal distribution was verified using the Shapiro-Wilk normality test.

2.3. Satellite image processing - Sentinel 2

The S2 mission by the European Space Agency (ESA) consists of a constellation of two satellites, S2A and S2B, launched on June 23, 2015, and March 7, 2017, respectively. The satellites are equipped with the Multispectral Instrument (MSI), which comprises 13 spectral bands with a spatial resolution of 10 m, 20 m, and 60 m, and a revisit time of 10 days with one satellite and 5 days with both satellites (ESA - European Space Agency, 2021). The specifications of the MSI sensor are detailed in Table 1. Due to its radiometric quality and high spatial resolution and integration three additional near-infrared (NIR) bands S2 has proven suitable for estimating water quality variables and has enabled the monitoring of spatial and temporal dynamics in inland waters (Sòria-Perpinyà et al., 2019; Alvado et al., 2021; Pompêo et al., 2021; Li et al., 2023)).

The 18 S2 images closest to the field data acquisition and with low cloud cover. These images were downloaded from the European Space Agency's (ESA) Open Access Hub server (https://dataspace.copernicus.eu/) without atmospheric correction, at the L1C processing level (top of atmosphere; TOA) images are georeferenced without atmospheric correction). Due to unfavorable cloud cover conditions during the sampling campaigns, only images with a time window of no more than 6 days were selected (Pompêo et al., 2021), allowing for an increase in the database used to generate the models. Given the different spatial resolutions of the images, it was necessary to adjust the resolution of all the bands. Thus, using the SNAP (SeNtinel Application Platform) program developed by ESA, the images were resampled to a resolution of 10 m and cropped, retaining only the study area.

2.3.1. Atmospheric correction

Due to the characteristics and components of coastal and inland waters, the waters under study were classified as 'Case 2' waters. They are characterized by the significant or predominant contribution of inorganic and/or organic sediments to optical properties (Soriano-González et al., 2022). Despite the existence of different atmospheric correction methods such as ACOLITE, C2RCC, C2X, C2XC, iCOR, POLYMER, SeaDAS, and Sen2Cor, for this work, freshwater-specific atmospheric corrections, with a wide range of inherent optical properties, developed for S2 using the SNAP software by Brockmann et al. (2016) were applied.

These atmospheric corrections are two neural networks: (a) Case2 eXtreme (C2X) and (b) C2X-COMPLEX (C2XC). These networks use neural networks trained with data for radiative transfer simulations of reflectance leaving the water and TOA radiances, as well as

Table 1
Characteristics of the MSI sensor.

Band	Spectral region	Spatial resolution (m)	λS2A (nm)	λS2B (nm)	ID in the study
B1	Aerossol	60	442.7	442.2	R443
B2	Blue	10	492.4	492.1	R490
B3	Green	10	559.8	559	R560
B4	Red	10	664.6	664.9	R665
B5	Red-edge1	20	704.1	703.8	R705
B6	Red-edge2	20	740.5	739.1	R740
B7	Red-edge3	20	782.8	779.7	R783
B8	Infrared (NIR)	10	832.8	832.9	R842
B8A	NIR narrow	20	864.7	864	R865
B9	Water vapor	60	945.1	943.2	R945
B10	SWIR/Cirrus	60	1373.5	1376.9	R1376
B11	SWIR1	20	1613.7	1610.4	R1620
B12	SWIR2	20	2202.4	2185.7	R2200

specific optical properties of the water (Brockmann et al., 2016). Each processor was trained with different ranges of inherent optical properties (IOPs) (Soriano-González et al., 2022). This correction has been used based on the characteristics and trophic status of inland water bodies showing good performance (Warren et al., 2019; Pereira-Sandoval et al., 2019; Soriano-González et al., 2022; Renosh et al., 2020; Pompêo et al., 2021). The output data include the Rrs. The corrections were applied separately to the 18 images to compare and select the one that best suited the study site's characteristics and provided the best results in calibrating and validating algorithms for estimating Chl-*a* concentration, due to it is an automatic product of SNAP neural nets. Once the appropriate atmospheric correction was selected, it was used to estimate the algorithms for the other variables studied, such as PC and cyanobacteria density.

2.4. Model calibration

2.4.1. Chlorophyll-a

To determine the Chl-*a* algorithm, Rrs values of each pixel coinciding with the coordinates of the *in situ* sampling points were extracted from each corrected image. The Chl-*a* concentration dataset was divided into two, the first for model calibration (70% of the data) and the second for validation (30%), covering the range of data in both groups.

During the calibration phase, different relationships with two and three bands were tested. Gilerson et al. (2010), Cairo et al. (2020), Ha et al. (2017), and Sòria-Perpinyà et al. (2021) demonstrated their effectiveness in estimating Chl-*a* concentrations. In this study, bands associated with reflectances were considered: R490 (B2), R560 (B3), R665 (B4), R705 (B5), R740 (B6), R783 (B7), R842 (B8), and R865 (B8A). For each case, regressions were performed between *in situ* Chl-*a* data and Rrs, and the best fit of the function was determined (linear, potential, exponential, or polynomial). In total, 15 band combinations were tested for each atmospheric correction (C2X and C2XC).

2.4.2. Phycocyanin

For determining an algorithm for PC estimates, the same methodology described for Chl-*a* concentration was initially followed using 14 *in situ* sampling points. However, due to the reduced amount of data, the statistical results were not significant, leading to a complementary methodology, as described below.

To expand the model calibration database, the Automated Radiative Transfer Model Operator (ARTMO) software (available at https://artmotoolbox.com/) and its Spectral Index (SI) evaluation toolbox were used in this study. This index allows for defining possible band combinations, and correlating combinations of reflectance measured at different wavelength ranges with *in situ* PC values. Simple ratio (RS = b1/b2) and normalized difference (DN = (b2 - b1)/(b2 + b1)) combinations were used, as well as triband combinations derived from the model of Dall'Olmo et al. (2003), TBDO = b1* ((1/b2) - (1/b3)). Various curve-fitting functions were used to obtain all band combinations (linear, exponential, power, logarithmic, and polynomial).

ARTMO divides the data into two parts, one for training the model and the other for validation. In this case, validation was achieved by calculating the leave-one-out cross-validation (LOOCV) method. This cross-validation method can be used for algorithms with sparse or limited data, as in this case (Bonansea et al., 2018). In total, 720 band combinations were tested.

2.4.3. Cyanobacteria density

Cyanobacteria density cannot be directly observed using satellite images. However, density can be estimated indirectly, provided there are statistically significant relationships with elements that the satellite can directly measure, such as Chl-*a*.

Thus, for estimating cyanobacteria density (cell numbers/ml), two algorithms were tested. One algorithm correlated Chl-*a* concentration and cyanobacteria density determined in field samples, supplemented with the CETESB database. The dataset was divided into two parts: one for model calibration (70% of the data), whether linear, potential, exponential, or polynomial and the other part was used for validation (30%).

For the second algorithm, a polynomial model (Equation (1)) was evaluated, as proposed by Pompêo et al. (2021), also applied to the Cantareira System.

$$\text{Cyanobacteria density} = (133.75*[\text{Chl-}a]\ ^{\wedge}2) + (3089.9*[\text{Chl-}a]) + 3125.9 \quad \text{(Eq.1)}$$

In this case, validation was conducted by comparing the model-estimated values with the field data available in the database used in this study. It is important to emphasize that, due to the origin of the model calibration data being from CETESB, measurements related to cyanobacteria density and Chl-*a* concentration for the period from 2015 to 2018 were excluded from the analysis. Therefore, validation was strictly limited to the values collected in the years between 2019 and 2022. This procedure ensured a more precise and representative evaluation of the model's performance (Hastie et al., 2009).

2.5. Model validation

The algorithms were validated with five statistics using field values and model-estimated values comparing: coefficient of determination (R^2), root mean square error (RMSE) (Eq. (2)), normalized root mean square error (NRMSE) (Eq. (3)), relative root mean square error (RRMSE) (Eq. (4)), and bias (Eq. (5)). The coefficient of determination was calculated by linear regression. The selection of the better performance is based on higher values of R2 and lower values of bias, RMSE, RRMSE, indicating lower prediction error and improved accuracy in parameter estimation (Moghimi et al., 2024).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i^{estimated} - x'_i{}^{measured\ in\ situ}\right)^2} \quad \text{(Eq.2)}$$

$$NRMSE = \frac{RMSE}{x'_{max}{}^{estimated} - x'_{min}{}^{estimated}} \quad \text{(Eq.3)}$$

$$RRMSE = \frac{RMSE}{\sum_{i=1}^{n} x'_i{}^{estimated}/n} \times 100\% \quad \text{(Eq.4)}$$

$$bias = \frac{1}{n}\sum_{i=1}^{n}\left(x_i^{measured\ in\ situ} - x'_i{}^{estimated}\right) \quad \text{(Eq.5)}$$

2.6. Automatic products

The C2X and C2XC processors provide automatic products that allow for the extraction of Chl-*a* concentration (conc_chla). These values are derived from the application of a specific equation (Eq. (6)), which incorporates standard factors, where a_pig represents the absorption coefficient of phytoplankton pigments. For detailed information on the coefficients associated with each atmospheric correction, it is possible to consult Soriano-González et al. (2022).

With the objective of making an independent validation of automatic products of atmospheric correction C2RCC, these automatically estimated values were related to the measurements obtained in the field, corresponding to the coordinates of the *in situ* sampling points. It is important to note that these automatic products do not include data related to PC products or cyanobacteria density, and therefore, they were tested exclusively for Chl-*a* concentration. Previous research conducted by Pompêo et al. (2021) and Pompêo and Moschini-Carlos (2022) demonstrated the robustness of these automatic products in tropical reservoirs.

$$conc\_chla - a = 21 * a_{pig}{}^{1.04} \quad \text{(Eq.6)}$$

where a_{pig} is the absorption coefficient of phytoplankton pigments, with training ranges from 0 to 30.81 (742.09 μg/L) for C2XC and from 0 to 51 (1253.41 μg/L) for C2X (Soriano-González et al., 2022).

3. Results

3.1. Field and laboratory data

Regarding the CETESB reports, Table 2 presents the values of the dataset obtained for the period from 2015 to 2021. The amount of data included differs between JAG and JAC, as the reports have information gaps.

As observed, in most sampling dates, JAG exhibited higher values for Chl-*a* concentration and cyanobacteria density (cells per ml), resulting in lower values for transparency (SD). However, a One-Way ANOVA analysis demonstrated statistically significant differences only for Chl-*a* concentration ($F = 4.2$, $p = 0.04$), which was higher in JAG compared to JAC.

Regarding the proprietary database, Table 3 contains the data of the study variables collected in the field during the period from 2021 to 2022. One-way ANOVA analysis demonstrated statistically significant differences for the variables Chl-*a* ($F = 172.9$, $p = 0.00$) and PC ($F = 83.29$, $p = 0.00$), being higher in JAG compared to JAC. This suggests a marked difference in water quality between the analyzed reservoirs. However, for the other parameters, the analysis of variance did not indicate statistically significant differences. It is important to highlight that, in all sampling locations, the [PC]: [Chl-*a*] ratio presented values above 0.5, suggesting,

Table 2
Values of the parameters obtained from the CETESB reports and descriptive statistics for each of the JAG-JAC reservoirs.

Date	JAG			JAC		
	DS (m)	Chl-*a* (μg/L)	Cyanobacteria density (Cells/ml)	DS (m)	Chl-*a* (μg/L)	Cyanobacteria density (Cells/ml)
06/07/2017	1.90	17.61	40,065	2.20	7.29	20,420
09/11/2017	1.40	19.78	81,272	1.20	10.69	45,373
22/05/2018	1.50	9.80	42,650	1.40	14.35	25,120
18/07/2018	1.20	8.22	36,522	1.80	4.01	17,928
03/01/2019	1.20	23.26	68,980	***0.80***	10.42	58,034
24/07/2019	**2.80**	10.42	35,796	**2.50**	***1.73***	11,015
07/07/2020	–	18.71	104,181	1.50	–	–
02/09/2020	–	13.70	59,685	1.10	–	–
04/11/2020	***0.80***	**40.10**	118,105	0.50	**18.27**	**123,315**
22/02/2021	1.00	23.68	**193,719**	0.90	17.95	163,638
27/05/2021	1.70	11.51	94,774	1.80	8.32	119,855
24/08/2021	1.30	15.37	132,924	1.10	5.35	21,856
10/11/2021	1.10	25.73	111,193	0.60	9.45	48,323

Table 3
Values of the parameters obtained during the campaigns from 2021 to 2022 for each of the JAG and JAC reservoirs.

Data	Station	DS (m)	Chl-*a* (μg/L)	Cyanobacteria density (Cells/ml)	PC (μg/L)	[PC]: [Chl-*a*]
6/10/2021	JAG1	0.65	38.9	**321,410**	**212.14**	5.45
	JAC1	**2.10**	***5.10***	78,417	***20.59***	4.04
	JAC2	1.90	7.77	72,339	22.52	2.90
	JAC3	1.40	6.43	***48,620***	22.34	3.47
9/06/2022	JAG1	**1.50**	***32.64***	***37,982***	140.17	4.29
	JAG2	1.53	29.44[a]	66,434	36.01[a]	1.22
	JAG3	***0.50***	317.38[a]	197,788	309.43[a]	0.97
	JAC1	1.60	13.46	134,247	25.44	1.89
	JAC2	***1.45***	9.37	77,945	26.20	2.80
	JAC3	1.58	10.49	**103,459**	42.91	4.09
17/11/2022	JAG1	0.62	**40.47**	319,782	136.81	3.38
	JAG2	0.64	39.85	301,760	143.37	3.60
	JAG3	0.71	36.27	309,501	114.40	3.15
	JAC1	1.51	14.61	160,378	42.78	2.93
	JAC2	1.51	**15.14**	190,915	42.21	2.79
	JAC3	1.43	14.78	247,173	**45.45**	3.08

[a] Excluded values from the models.

indirectly, the dominance of cyanobacteria over the total phytoplankton biomass (Shi et al., 2015; Hunter et al., 2009). It is noteworthy that, in the case of Chl-*a* and PC concentrations, two outliers were excluded.

From the databases described earlier, a single database was created by merging the datasets. This approach allowed for an increase in the amount of available information, enabling the incorporation of variations in the evaluated parameters in the reservoir during different periods and limnological conditions. This diversity of data plays an important role in the process of model calibration and validation (Cairo et al., 2020). Table 4 presents the descriptive statistics of the unique database (mean, median, minimum, maximum, standard deviation (σ), and coefficient of variation (CV)).

Regarding the correlations among variables in the unique dataset (Table 5), a strong positive correlation between Chl-*a* and PC ($r = 0.94$, $p < 0.001$), and Chl-*a* and cyanobacteria cell density (0.74, $p < 0.001$) is evident. This suggests that cyanobacteria are significant contributors to the total phytoplankton, both in terms of chlorophyll and density.

3.2. Empirical model calibration

3.2.1. Chlorophyll-a

Table 6 presents the results of calibration and validation of algorithms for estimating Chl-*a* concentrations generated from different combinations of bands and for the two atmospheric corrections employed, displaying only the models that achieved an R^2 value greater than 0.60 in calibration.

In general, polynomial models exhibited more robust results compared to other types of regressions, as demonstrated in prior studies (Pereira Sandoval et al., 2019; Sòria-Perpinyà et al., 2021). Among the tested combinations, the best results were obtained for algorithms that included the R665, R705, and R740 bands.

Table 4
Descriptive statistics of the parameters obtained from the unique database dataset for each of the JAG and JAC reservoirs.

	JAG				JAC			
	DS	Chl-*a*	Cyanobacteria density (Cells/ml)	PC	DS	Chl-*a*	Cyanobacteria density (Cells/ml)	PC
n	19	23	22	5	21	19	22	9
Mean	1.21	20.8	121,864	149	1.5	9.53	84,802	32.3
Median	1.1	17.6	88,023	140	1.45	9.37	75,142	26.2
Mín.	0.5	1.07	240	114.40	0.8	1.73	3865	20.6
Máx.	2.5	40.5	321,410	212.14	2.8	18.3	247,173	45.5
σ	0.579	11.9	105,338	36.9	0.432	4.59	64,888	10.7
CV (%)	47.77	57.08	86.44	24.69	28.82	48.17	76.52	33.04

Table 5
Pearson correlation coefficients, the probability value of the unique dataset.

Parameter	Chl-*a* (μg/L)	Cyanobacteria density (Cells/ml)	PC (μg/L)
Chl-*a* (μg/L)	–		
Cyanobacteria density (Cells/ml)	0.74**	–	
PC (μg/L)	0.94 **	0.66[a]	–

[a] $p < 0.01$. **$p < 0.001$.

Table 6
Calibration and validation results for Chl-*a* algorithms. Regression: linear (l.); polynomial order 2 (pl.); logarithmic (log.).

Atmospheric correction	Bands relation	Calibration			Validation					
		n	R^2		N	R^2	RMSE	RRMSE	NRMSE	Bias
C2X	Automatic C2X				42	0.75	6.79	43.17	17.22	−4.08
C2XC	Automatic C2XC				42	0.73	7.30	46.50	18.55	−4.26
C2X	**R705/R665**	30	0.63	(l)	12	0.81	5.61	27.92	16.49	1.09
C2X[a]			**0.67**	**(pl)**		**0.84**	**5.07**	**25.21**	**14.89**	**0.26**
C2XC			0.68	(l)		0.81	5.78	28.75	16.98	−1.35
C2XC			0.71	(pl)		0.74	9.68	48.14	28.43	−3.53
C2X	(1/R665-1/R705) [a]R740	30	0.63	(l)	12	0.83	5.3	26.37	15.58	0.72
C2X			0.65	(pl)		0.84	5.37	26.7	15.77	−0.24
C2XC			0.67	(l)		0.83	5.32	26.44	15.62	−0.95
C2XC			0.69	(pl)		0.80	6.89	34.28	20.24	−2.23
C2X	R665/R705	30	0.66	(pl)	12	0.87	4.59	22.84	13.49	0.30
C2XC			0.66	(log)		0.81	5.59	27.79	16.41	−0.66
C2XC			0.69	(pl)		0.81	6.08	30.23	17.85	−1.74
C2XC	(R705+R740)/R665	30	0.64	(l)	12	0.81	5.56	27.64	16.32	−0.58
C2XC			0.66	(pl)		0.77	7.56	37.62	22.22	−2.18
C2X	(((35.75[a](R705/R665)) −19.3) ^1.124	30	0.64	(pl)	12	0.82	5.418	26.95	15.92	0.89
C2X			0.68	(pl)		0.85	5.11	25.42	15.01	−0.26
C2XC			0.71	(pl)		0.73	10.181	50.64	29.92	−3.73
C2XC			0.68	(l)		0.80	5.934	29.51	29.51	−1.54

[a] Selected bands relation.

The most favorable outcomes, as indicated by various validation statistics, were associated with the red and near-red edge bands processed by the C2X algorithm. Specifically, the relationships R705/R665(pl) and R665/R705(pl) demonstrated similar results, showing robustness and precision. For the R705/R665(pl) relationship, a coefficient of determination (R^2) of 0.84, a root mean square error (RMSE) of 5.07 μg/L, a relative RMSE (RRMSE) of 25.21%, a normalized RMSE (NRMSE) of 14.89%, and a bias of 0.26 μg/L were observed. Meanwhile, the R665/R705(pl) relationship achieved an R2 of 0.87, an RMSE of 4.59 μg/L, an RRMSE of 22.84%, an NRMSE of 13.49%, and a bias of 0.30. Given the similarity in performance, the physically based relationship was selected, defining the ratio between the wavelength associated with maximum scattering (minimum Chl-*a* absorption) and the wavelength associated with the maximum absorption of this pigment (Gitelson et al., 1999; Morel and Prieur, 1977). This relationship is represented by the R705/R665 ratio for S2.

Thus, employing the R705/R665 ratio, Fig. 3a presents the regression calculated and selected for Chl-*a* concentration estimation, given by Equation (7), and Fig. 3b shows validation through linear regression between the *insitu* measured Chl-*a* concentrations and those calculated with the algorithm (Eq. (7)).

$$Chl-a\left(\frac{\mu g}{L}\right) = 23.96 * \frac{R705^{2}}{R665} - 7.37 * \frac{R705}{R665} \quad \text{(Eq. 7)}$$

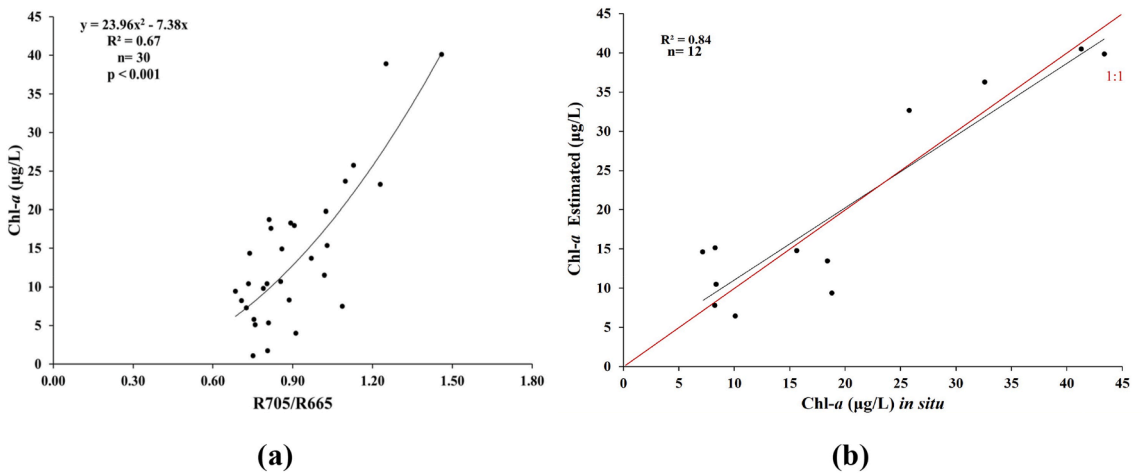

Fig. 3. (a) Algorithm calculated for Chl-*a* concentration from field-measured data and the reflectance ratio of bands R705/R665. **(b)** Validation: linear regression between estimated values of calculated Chl-*a* and *in situ* measured Chl-*a* concentration.

3.2.2. Phycocyanin

For the calibration and validation of PC algorithms, the atmospheric correction derived from C2X was utilized, as detailed in the methodology. The results from the analysis conducted using the ARTMO software (Table 7) showed that exponential and power regression models yielded the best outcomes. Only 10 out of the tested combinations exhibited $R^2 \geq 0.6$. Overall, RMSE ranged from 37.11 to 39.77 μg/L, NRMSE ranged from 19.37 to 20.75%, and RRMSE remained above 50%, suggesting significant variability in the results. It is noteworthy that all estimated algorithms demonstrated a tendency to underestimate PC values, as indicated by bias ranging from −6.0 μg/L to −8.0 μg/L.

The best performance in estimating PC concentration was observed for the algorithm utilizing the relationship between the two bands, R490/R705 (e.), as it yielded the highest coefficient of determination (R^2) of 0.63, along with lower values for RMSE of 37.11 μg/L, RRMSE of 50.08%, and NRMSE of 19.37%. These results highlight the potential of this algorithm for PC estimation. The algorithm equation is presented in Equation (8), with its validation depicted in Fig. 4.

Table 7
Calibration and validation results for phycocyanin algorithms. Regression: exponential (e.); potential (p.).

Relation	Regression	bands	R^2	RMSE	RRMSE	NRMSE	Bias
RS	e.	**R705, R490**[a]	**0.63**	**37.11**	**50.08**	**19.37**	**−7.20**
TBDO	e.	R740, R705, R490	0.61	37.99	51.28	19.84	−7.00
TBDO	e.	R740, R490, R705	0.61	37.99	51.28	19.84	−7.00
TBDO	e.	R783, R705, R490	0.61	38.15	51.49	19.92	−6.00
TBDO	e.	R783, R490, R705	0.61	38.15	51.49	19.92	−6.00
RS	e.	R783, R865	0.60	39.19	52.89	20.46	−6.00
ND	p.	R865, R783	0.60	39.15	52.84	20.44	−7.00
ND	p.	R783, R865	0.60	39.15	52.84	20.44	−7.00
TBDO	e.	R865, R705, R490	0.60	38.51	51.98	20.11	−7.20
TBDO	e.	R865, R490, R705	0.60	38.51	51.98	20.11	−7.20
TBDO	e.	R665, R705, R490	0.59	39.92	53.88	20.84	−7.20
TBDO	e.	R665, R490, R705	0.59	39.92	53.88	20.84	−7.20
TBDO	p.	R865, R740, R490	0.59	40.06	54.07	20.92	−6.70
TBDO	p.	R865, R490, R740	0.59	40.06	54.07	20.92	−6.70
TBDO	p.	R783, R740, R865	0.58	40.02	54.00	20.89	−7.20
TBDO	e.	R783, R560, R490	0.58	39.77	53.68	20.76	−8.00
TBDO	e.	R783, R490, R560	0.58	39.77	53.68	20.76	−8.00

RS: Simple Ratio, ND: Normalized Difference, TBDO Three-Band Algorithm.
[a] Selected band's relation.

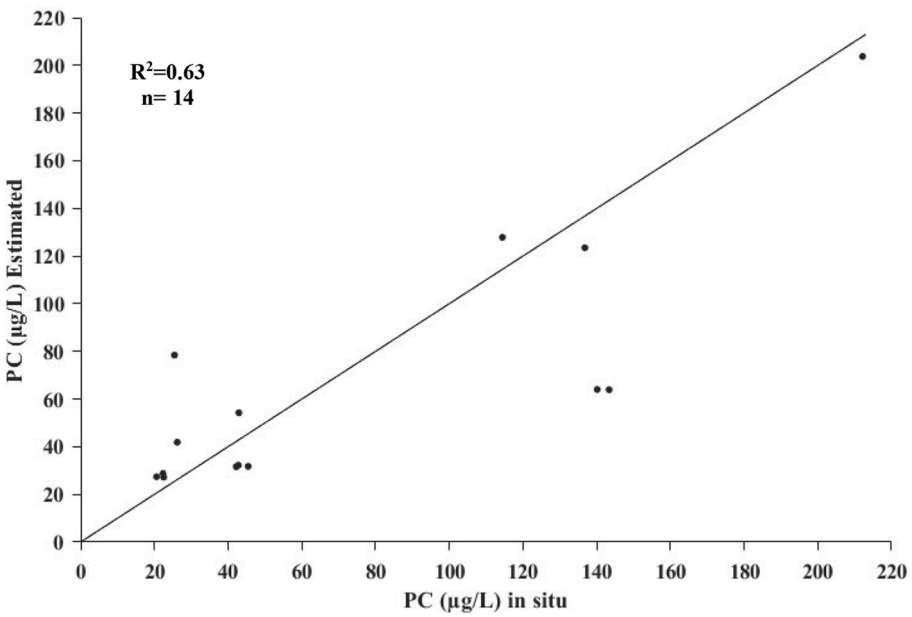

Fig. 4. Validation of the calculated algorithm for PC concentration from in-situ measured data and estimated values by the model from reflectance of the band ratio R705/R490.

$$PC\left(\frac{\mu g}{L}\right) = 17.23 * e^{1.17*\left(\frac{R705}{R490}\right)} \quad \text{(Eq.8)}$$

3.2.3. Cyanobacteria density

Table 8 presents the calibration and validation results for algorithms developed to estimate cyanobacteria density. In comparison to the model proposed by Pompêo et al. (2021), it exhibited favorable results during validation. Fig. 5 depicts the validation of this model through linear regression between in-situ measured Chl-*a* concentration values and those calculated by the algorithm.

On the other hand, the algorithm derived from the dataset of the two databases exhibited a more precise fit. This was achieved through a polynomial regression model, highlighting a strong correlation between the variables, with an R^2 of 0.75. During validation, R^2 values of 0.84 and an RMSE of 19.53 cells/ml were observed, the latter being like that obtained in the validation of the model proposed by Pompêo et al. (2021), however, the RMSE and RRMSE are lower, suggesting a more robust model for the data in this study.

Fig. 6a presents the calculated regression for estimating cyanobacteria density based on the selected model, while Fig. 6b displays validation through linear regression between values of Chl-*a* concentration measured *in situ* and those calculated by the algorithm. Therefore, the algorithm selected is presented by Equation (9) (Eq. (9)).

$$121.98 * \text{Chl} - a^2 + 3027.4 * \text{Chl} - a + 9991.8 \quad \text{(Eq. 9)}$$

3.3. Automatic products

According to the validation of the automatic products, it was determined that the concentration values of conc_cl-a showed a clear correlation with the *in situ* concentration for both C2X and C2XC (Fig. 7a and b), with a coefficient of determination (R^2) of 0.75 and 0.73, respectively. However, the C2X product demonstrated the best fit for estimating conc_cl-a concentration. Additionally, the C2XC automatic product showed an even greater tendency to underestimate conc_cl-a concentration values (see 1:1 line - Fig. 7a and b).

Table 8
Calibration and validation results for cyanobacteria density algorithms based on Chl-*a* concentration values. Regression: polynomial (pl.).

Source of data	Calibration			Validation					
	n	R^2	Regression	n	R^2	RMSE	RRMSE	NRMSE	Bias
Pompêo et al. (2021)	90	0,84	(pl.)	31	0,83	62,338,12	55,8	19,63	17,114,94
Own	**25**	**0,75**	**(pl.)** [a]	**11**	**0,84**	52,952,11	21,52	19,53	28,009,82

[a] Selected model.

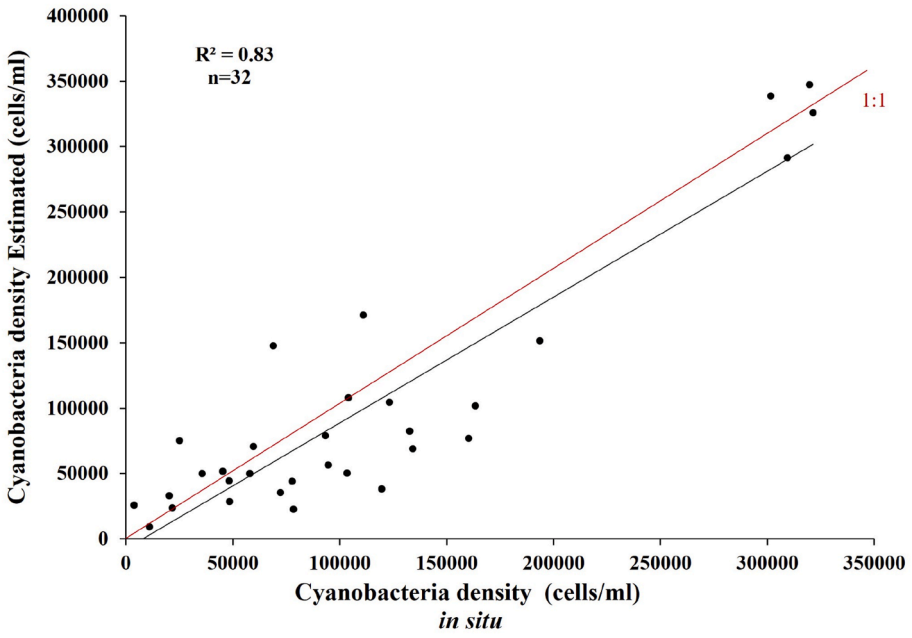

Fig. 5. Validation of the model proposed by Pompêo et al. (2021), linear regression between estimated values of calculated cyanobacteria density and cyanobacteria density measured *in situ*.

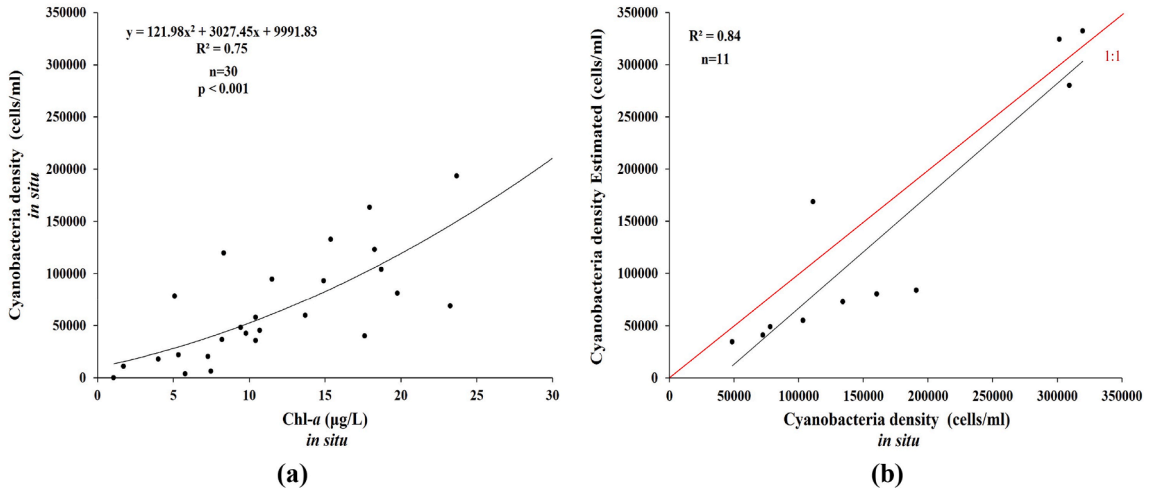

Fig. 6. **(a)** Algorithm calculated for estimating cyanobacteria density from field-measured data **(b)** Validation, linear regression between estimated values of calculated cyanobacteria density and cyanobacteria density measured *in situ*.

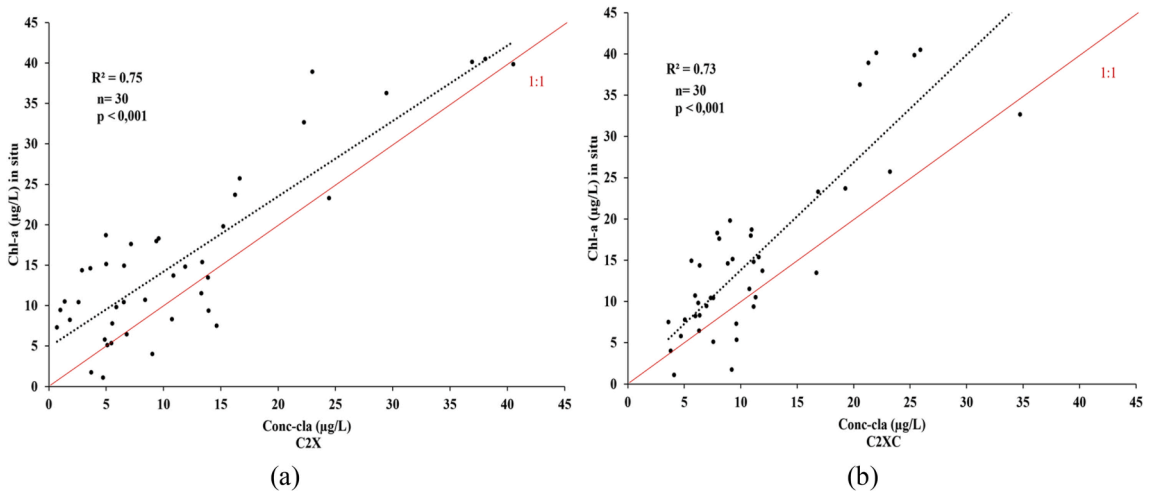

Fig. 7. Validation of the estimated values by the automatic products (Conc-cl-a) vs. *in situ* Chl-a values (a) C2X (b) C2XC.

Table 6 presents the validation statistics for the automatic products generated with C2X and C2XC atmospheric corrections. Both atmospheric corrections exhibited similar performance. However, the values obtained for the C2X correction indicate superior model performance, with an RMSE of 6.79 µg/L, an RRMSE of 43.17%, and an NRMSE of 17.22%. These results suggest that the automatic conc_cl-a products from the C2X atmospheric correction may have greater applicability in the study reservoir.

4. Discussion

In this study, it was observed that Sentinel-2 (S2) showed good performance in estimating parameters such as Chl-*a*, PC, and cyanobacteria density. This reinforces the potential of remote sensing as a support tool for *in situ* techniques in the study area, allowing for more frequent and synoptic monitoring (Ioannou et al., 2014). Besides enabling ecological and water quality analyses, especially for eutrophication assessment and the possibility of future application conducting spatial and temporal studies in tropical inland aquatic ecosystems.

Although the field-collected dataset is valuable, its representativeness is limited due to the restricted number of data points and sampling dates, which may affect the robustness of water quality parameter estimates. An approach to overcome this limitation is the incorporation of public and freely available databases, such as those maintained by CETESB, which regularly samples every two months in the JAG-JAC reservoir. This can also be replicated not only for the study reservoirs but also in other reservoirs in the state of São Paulo, as demonstrated by Pompêo et al. (2021) and Pompêo and Moschini-Carlos (2022).

However, it is important to note that in the water quality monitoring carried out by CETESB, PC is not included in the analyzed variables, and has never been monitored in this reservoir. The algorithm proposed in this work has the potential to fill this gap by allowing its estimation. However, it is emphasized that there is a need for this agency to advance in PC monitoring given the impor-

tance of assessing the presence of cyanobacteria and the associated risks in aquatic ecosystems (Pamula et al., 2023), as cyanobacteria are important components of the total phytoplankton in the JAG-JA System.

In satellite image usage, it is important to recognize that due to the low reflectance of water, about 90% of the radiance received by remote sensors is influenced by the atmosphere (gases, aerosols, scattering). Therefore, it is evident that atmospheric correction is essential to obtain accurate estimates of water reflectance and water quality variables (Delegido et al., 2019). Thus, the C2X processor showed better results compared to C2XC when *in situ* and calculated Chl-a was compared. The C2X processor has shown acceptable results for the analysis of the variables considered in this study, performing well in mesotrophic to eutrophic waters in temperate (Soria-Perpinyà et al., 2022) and tropical (Virdis et al., 2022) ecosystems. Moreover, this can probably be attributed to the fact that C2X has shown a good correlation in radiometric validation with red and infrared bands (bands used in Chl-*a* concentration estimation), especially in eutrophic waters such as JAC-JAC and in turbid waters (Pereira-Sandoval et al., 2019).

Algorithms based on spectral bands associated with the red edge and NIR regions of the electromagnetic spectrum are recommended for the study of productive and turbid waters, particularly in waters where Chl-*a* concentration is above 10 μg/L (Mishra and Mishra, 2020). This is due to the maximum absorption characteristics of Chl-*a* in the red region around 670 nm and the peak reflectance in the NIR region between 670 and 750 nm (Gilerson et al., 2010). Especially, the red-edge band is important, as phytoplankton causes a peak in spectral reflectance near 700 nm due to the combined minimum absorption of water and phytoplankton (Bramich et al., 2021).

Comparing this study with the results of other works, it can be observed that Soria-Perpinyà et al. (2021) obtained a stronger correlation between in-situ data and model data, with an R^2 of 0.9, but with higher values for RMSE (39.74 μg/L) and RMSE = 48%, with a bias = 2.22 μg/L, for 296 data points. In another study with three bands (R665 nm, R705 nm, and R740 nm), conducted by Aranha et al. (2022) in five Brazilian reservoirs with trophic states from mesotrophic to hypereutrophic, a lower R2 value of 0.77 and an NRMSE of 28% were obtained for 242 data points (70% calibration data and 30% validation data).

For studies considering a similar number of data points to this work, Chen et al. (2017) presented an R^2 = 0.9, an RMSE = 9.972 μg/L, and an NRMSE of 48.47% (n = 42), and Moses et al. (2019), using 15 calibration data points and 34 validation data points, obtained an RMSE = 6.53 μg/L and an NRMSE = 8%, R^2 = 0.95. The statistical results of these studies are like the present study, confirming that the selected model is robust in estimating Chl-*a* concentration.

Empirical algorithms generally produce robust results for the specific areas and datasets from which they are derived, and they present advantages over other types of bio-optical models (Matthews, 2010).

Few studies have validated the use of automatic Chl-*a* products offered by the C2RCC processor. However, these products may present an acceptable correlation. For example, in the study by Alvado et al., 2021, on temperate waters, the statistical values were (R^2 = 0.55), (RMSE = 14.41 mg/L; NRMSE = 30%), demonstrating that the validation results shown in this study are more robust. In the study by Soriano-González et al., 2022), conducted in temperate waters with different water characteristics, C2X showed better performance compared to previous studies, with validation statistics of RMSE = 17.3, bias = 5.85, R^2 = 0.66, and for C2XC, RMSE = 48.1, bias = 12.1, and R^2 = 0.88. This case demonstrated that for different types of waters, C2X is a better processor for estimating Chl-*a* in eutrophic-hypertrophic waters. The results in this study are also superior. The use of automatic products in tropical reservoirs is very limited; only Pompêo et al. (2021 e 2022) tested the C2X processor, where the product showed good accuracy for Brazilian reservoirs.

In the case of the algorithm for PC estimates, in this work, the two-band ratio (R460/R705) showed the best fit (RMSE = 37.11 μg/L, RRMSE = 50.08%, NRMSE = 19.37%, bias = −7.20, and R^2 = 0.63). Other studies also used combinations of bands from the visible and infrared spectrum regions for PC estimates in turbid and productive waters, using sensors such as Sentinel 3 and MERIS. (Qi et al., 2014)Qi et al. (2014) used bands R550, R620, and R665, and obtained an R^2 value of 0.64 and NRMSE of 85.4% (n = 28), the coefficient of determination is higher than that determined in this study; however, the NRMSE is much higher, indicating that the model proposed in this study could be an initial model for monitoring the JAG and JAC system, given that this variable has not been historically monitored.

PC commonly exhibits maximum absorption around R620; for this reason, studies suggest using this spectral region to obtain suitable algorithms (Yan et al., 2018). Nevertheless, this band wavelength is available only in sensors such as MERIS and Sentinel3. Although S2 is useful and accurate in PC quantification, which is why the most common algorithms use R665 as an alternative (Sòria-Perpinyà et al., 2019). Most of the consulted studies determined algorithms with spectral regions ranging from R600 to R720 (Woźniak et al., 2016; Sòria-Perpinyà et al., 2019; Pérez-González et al., 2021).

When there is insufficient data or a scarcity of PC information, an algorithm based on Chl-*a* concentration can be a viable alternative to infer cyanobacteria density, as has been done in other studies (Drozd et al., 2019; Pompêo et al., 2021; Pompêo and Moschini-Carlos, 2022). This is because Chl-*a* is used as an indirect indicator for estimating this parameter (Randolph et al., 2008). This is feasible when cyanobacteria dominate the phytoplankton, as described in this study.

5. Conclusions

The importance of establishing a monitoring protocol for continental aquatic ecosystems is evident, aiming to allow integrated spatial and temporal analysis and to optimize decision-making and management actions for reservoirs sensitive to eutrophication processes, especially in scenarios of climate change and increased frequency of cyanobacterial blooms. In this context, the use remote sensing shows great potential for estimating water quality parameters such as Chl-*a*, phycocyanin, and cyanobacteria density through empirical algorithms. As evidenced in this study, the use of Sentinel-2 (S2) demonstrated good performance and accuracy for this purpose.

The objective of this study was to generate algorithms for Chl-*a*, PC and cyanobacteria density estimation using S2 images for the JAC-JAG reservoir. These selected algorithms show great potential for estimating this parameter, assessing water quality variables, and evaluating the ecological state of the reservoir. They can also be useful for future applications in continuous monitoring of spatial and temporal variations, serving as a tool for the Brazilian environmental agency CETESB.

The algorithm for chlorophyll- Chl-*a* a suggests using the ratio between the red band and red edge, a relationship associated with maximum scattering (minimum chlorophyll-a absorption) and the wavelength associated with the maximum absorption of this pigment. This ratio is recommended for eutrophic waters. The algorithm for PC suggests using the red edge and blue bands. This is the first algorithm obtained for PC in this reservoir, where PC has never been monitored, marking an initiative for this purpose. Finally, the cyanobacteria density algorithm was derived from a relationship with *in situ* Chl-*a* data, which is feasible in scenarios of cyanobacteria dominance.

The automatic conc_cla products derived from the atmospheric corrections C2X and C2XC provided good results. The best performance was derived from the C2X processor, which was more suitable and supported the study of the JAC-JAG reservoir. These automatic products have great potential for the rapid estimation of Chl-*a*. However, this study showed that empirical algorithms are more robust, using specific algorithms for the study area always reduces uncertainty.

Funding

His work was supported by Brazilian FAPESP agency (processes 2021/11283-0, 2020/11759-1 and 2021/10637-2).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

CRediT authorship contribution statement

Zahia Catalina Merchan Camargo: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Xavier Sòria-Perpinyà:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Marcelo Pompêo:** Writing – review & editing, Methodology, Investigation, Funding acquisition, Conceptualization. **Viviane Moschini-Carlos:** Writing – review & editing, Methodology, Investigation, Funding acquisition. **Maria Dolores Sendra:** Writing – review & editing, Methodology, Investigation.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT 3.5 to translate the manuscript into English. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors are very grateful to Juan Pablo Rivera and Gabriel Caballero for their support in the use of the Artmo software. We are also very thankful to Maria Elicia Mac Donagh for her comments and feedback.

Data availability

Data will be made available on request.

References

Alvado, B., Sòria-Perpinyà, X., Vicente, E., Delegido, J., Urrego, P., Ruíz-Verdú, A., Soria, J.M., Moreno, J., 2021. Estimating organic and inorganic parts of suspended solids from sentinel 2 in different inland waters. Water 13, 2453. https://doi.org/10.3390/w13182453.

Aranha, T., Martinez, J.-M., Souza, E., Barros, M., Martins, E., 2022. Remote analysis of the chlorophyll-a concentration using sentinel-2 MSI images in a semiarid environment in Northeastern Brazil. Water 14 (3), 451. https://doi.org/10.3390/w14030451.

Barbosa, J., Medeiros, E., Brasil, J., Cordeiro, R., Crispim, M.C., Silva, G., 2012. Aquatic systems in semi-arid Brazil: limnology and management. Acta Limnol. Bras. 24, 103–118. https://doi.org/10.1590/S2179-975X2012005000030.

Bocharov, A., Tikhomirov, O., Khizhnyak, S., 2017. Monitoring of chlorophyll in water reservoirs using satellite data. J. Appl. Spectrosc. 84 (2), 291–295. https://doi.org/10.1007/s10812-017-0466-7.

Bonansea, M., Ledesma, M., Rodriguez, C., Pinotti, L., 2018. Using new remote sensing satellites for assessing water quality in a reservoir. Hydrol. Sci. J. 64 (1), 34–44. https://doi.org/10.1080/02626667.2018.1552001.

Bramich, J., Fischer, A., Bolch, C., 2021. Improved red-edge chlorophyll-a detection for Sentinel 2. Ecol. Indicat. . 120.10.1016/j.ecolind.2020.106876.

Brient, L., Lengronne, M., Bertrand, E., Rolland, D., Sipel, A., Steinmann, D., Baudin, I., Legeas, M., Rouzic, B., Bormans, M., 2007. A phycocyanin probe as a tool for monitoring cyanobacteria in freshwater bodies. J. Environ. Monit.: JEM 10, 248–255. https://doi.org/10.1039/b714238b.

Brockmann, C., Doerffer, R., Peters, M., Kerstin, S., Embacher, S., Ruescas, A., 2016. Evolution of the C2RCC neural network for sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters. ESASP 740, 54.

Cairo, C., Barbosa, C., Lobo, F., Novo, E., Carlos, F., Maciel, D., Flores Júnior, R., Silva, E., Curtarelli, V., 2020. Hybrid chlorophyll-a algorithm for assessing trophic states of a tropical Brazilian reservoir based on MSI/Sentinel-2 data. Rem. Sens. 12, 40. https://doi.org/10.3390/rs12010040.

Campbell, G., Phinn, S.R., Dekker, A.G., Brando, V.E., 2011. Remote sensing of water quality in Australia. Tropical freshwater impoundment using matrix inversion and MERIS images. Rem. Sens. Environ. 115, 2402–2414. https://doi.org/10.1016/j.rse.2011.05.003.
Carvalho, L., Poikane, S., Lyche Solheim, A., Phillips, G., Borics, G., Catalan, J., De Hoyos, C., Drakare, S., Dudley, B.J., Jarvinen, M., Laplace-Treyture, C., Maileht, K., McDonald, C., Mischke, U., Moe, J., Morabito, G., Noges, P., Noges, T., Ott, I., Pasztaleniec, A., Skjelbred, B., Thackeray, S.J., 2013. Strength and uncertainty of phytoplankton metrics for assessing eutrophication impacts in lakes. Hydrobiologia 704 (1), 127–140. https://doi.org/10.1007/s10750-012-1344-1.
CETESB. Companhia Ambiental do Estado de São Paulo, 2012. L5.303: Fitoplâncton de água doce: métodos qualitativo e quantitativo. São Paulo. No prelo.
CETESB. Companhia Ambiental do Estado de São Paulo. 2019-2022. Publicações e relatórios – Águas interiores. Águas Interiores - CETESB. Retrieved from https://cetesb.sp.gov.br/aguas-interiores/publicacoes-e-relatorios (Accessed 17 March 2023).
Chaffin, J.D., Bridgeman, T.B., 2013. Organic and inorganic nitrogen utilization by nitrogen-stressed cyanobacteria during bloom conditions. J. Appl. Phycol. 26 (1), 299–309. https://doi.org/10.1007/s10811-013-0118-0.
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. J. Appl. Remote Sens. 11 (3), 036007. https://doi.org/10.1117/1.JRS.11.036007.
Chorus, I., Bartram, J., 1999. Toxic Cyanobacteria in Water: a Guide to Their Public Health Consequences, Monitoring and Management. E&FN Spon: WHO, London, GB, p. 416.
Dall'Olmo, G., Rundquist, D., Gitelson, A., 2003. Towards a unified approach for remote estimation of chlorophyll-a in both terrestrial vegetation and turbid productive waters. Geophys. Res. Lett. 30, 1938. https://doi.org/10.1029/2003GL018065.
Damar, A., Colijn, F., Hesse, K.-J., Kurniawan, F., 2020. Coastal phytoplankton pigments composition in three tropical estuaries of Indonesia. J. Mar. Sci. Eng. 8 (5), 311. https://doi.org/10.3390/jmse8050311.
Delegido, J., Urrego, P., Vicente, E., Sòria-Perpinyà, X., Soria, J.M., Pereira-Sandoval, M., Ruiz-Verdú, A., Peña, R., Moreno, J., 2019. Turbidity and Secchi disc depth with sentinel-2 in different trophic status reservoirs at the comunidad valenciana. Rev. de Teledetec. 54, 15–24. https://doi.org/10.4995/raet.2019.12603.
Drozd, A., de Tezanos Pinto, P., Fernández, V., Bazzalo, M., Bordet, F., Ibañez, G., 2019. Hyperspectral remote sensing monitoring of cyanobacteria blooms in a large South American reservoir: high- and medium-spatial resolution satellite algorithm simulation. Mar. Freshw. Res. https://doi.org/10.1071/mf18429.
ESA - European Space Agency, 2021. Sentinel-2 mission. Retrieved from. https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2. (Accessed 10 September 2023).
Fonseca, B.M., Ferragut, C., Tucci, A., Crossetti, L.O., Ferrari, F., de Campos Bicudo, D., de Mattos Bicudo, C.E., 2014. Biovolume de cianobactérias e algas de reservatórios tropicais do Brasil com diferentes estados tróficos. HOEHNEA 41 (1), 9–30. https://doi.org/10.1590/S2236-89062014000100002.
Gilerson, A., Gitelson, J., Zhou, D., Gurlin, M., Wesley, I., Iannou, S. Ahmed, 2010. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near-infrared bands. Opt. Express 18 (23), 24109–24125. https://doi.org/10.1364/OE.18.024109.
Gitelson, A.A., Schalles, J.F., Rundquist, D., Schiebe, F.R., Yacobi, Y.Z., 1999. Comparative reflectance properties of algal cultures with manipulated densities. J. Appl. Phycol. 11, 345–354.
Glazer, A.N., 1989. Light guides. Directional energy transfer in a photosynthetic antenna. J. Biol. Chem. 264, 1–4. https://doi.org/10.1016/S0021-9258(17)31212-7.
Glibert, P.M., 2019. Harmful algae at the complex nexus of eutrophication and climate change. Harmful Algae 101583. . 10.1016/j.hal.2019.03.001.
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 Int. J. Geo-Inf. 6 (9), 290. https://doi.org/10.3390/ijgi6090290.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, second ed. Springer, New York. https://doi.org/10.1007/978-0-387-84858-7.
Hillebrand, H., Dürselen, C.D., Kirschtel, D., Pollingher, U., Zohary, T., 1999. Biovolume calculation for pelagic and benthic microalgae. J. Phycol. 35 (2), 403–424. https://doi.org/10.1046/J.1529-8817.1999.3520403.X.
Hou, X., Feng, L., Dai, Y., Hu, C., Gibson, L., Tang, J., Lee, Z., Wang, Y., Cai, X., Liu, J., Zheng, Y., Zheng, C., 2022. Global mapping reveals increase in lacustrine algal blooms over the past decade. Nat. Geosci. 15 (2), 130–134. https://doi.org/10.1038/s41561-021-00887-x.
Hunter, P.D., Tyler, A.N., Gilvear, D.J., Willby, N.J., 2009. Using remote sensing to aid the assessment of human health risks from blooms of potentially toxic cyanobacteria. Environ. Sci. Technol. 43 (7), 2627–2633. https://doi.org/10.1021/es802977u.
Ioannou, I., Gilerson, A., Ondrusek, M., Foster, R., El-Habashi, A., Bastani, K., Ahmed, S., 2014. Algorithms for the remote estimation of chlorophyll-a in the Chesapeake Bay. Ocean Sens. Monit. 9111, 911118. https://doi.org/10.1117/12.2053753.
Jeffrey, S.W., Humphrey, G.F., 1975. New spectrophotometric equations for determining chlorophylls a, b, c and c2 in higher plants. Algae Natur. Phytoplankton. Biochem. Phys. der Pflanzen 167, 191–194.
Lewis, Jr, W.M., 2002. Tropical lakes: how latitude makes a difference. In: Department of Environmental, Population, and Organismic Biology. University of Colorado.
Li, L., Song, K., 2017. Bio-optical modeling of phycocyanin. Bio-Opt. Model. Remote Sens. Inland Water. 233–262. https://doi.org/10.1016/b978-0-12-804644-9.00008-2.
Li, S., Song, K., Li, Y., Liu, G., Wen, Z., Shang, Y., Lyu, L., Fang, C., 2023. Performances of atmospheric correction processors for Sentinel-2 MSI imagery over typical lakes across China. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 16, 2065. https://doi.org/10.1109/JSTARS.2023.3238713.
Llodrà-Llabrés, J., Martínez-López, J., Postma, T., Pérez-Martínez, C., Alcaraz-Segura, D., 2023. Retrieving water chlorophyll-a concentration in inland waters from Sentinel-2 imagery: review of operability, performance and ways forward. Int. J. Appl. Earth Obs. Geoinf. 125, 103605. https://doi.org/10.1016/j.jag.2023.103605.
Lorenzen, C.J., 1967. Determination of chlorophyll and pheopigments: spectrophotometric equations. Limnol. Oceanogr. 12, 343–346. https://doi.org/10.4319/lo.1967.12.2.0343.
Matthews, M.W., 2010. Empirical remote sensing in inland and near-coastal transitional waters. Int. J. Rem. Sens. 31 (7), 1827–1837.
Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., Chaffin, J.D., Cho, K., Confesor, R., Daloglu, I., Depinto, J.V., Evans, M.A., Fahnenstiel, G.L., He, L., Ho, J.C., Jenkins, L., Johengen, T.H., Kuo, K.C., Laporte, E., Liu, X., et al., 2013. Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. Proc. Natl. Acad. Sci. U.S.A. 110 (16), 6448–6452. https://doi.org/10.1073/pnas.1216006110.
Mishra, S., Mishra, D.R., 2020. Normalized difference chlorophyll index: a novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. Remote Sens. Environ. 117, 394–406. https://doi.org/10.1016/j.rse.2011.10.016.
Moghimi, A., Tavakoli Darestani, A., Mostofi, N., Fathi, M., Amani, M., 2024. Improving forest above-ground biomass estimation using genetic-based feature selection from Sentinel-1 and Sentinel-2 data (case study of the Noor forest area in Iran). Kuwait J. Sci. 51, 100159. https://doi.org/10.1016/j.kjs.2023.100159.
Morel, A., Prieur, L., 1977. Analysis of variation in ocean color. Limnol. Oceanogr. 37, 147–149. https://doi.org/10.4319/lo.1977.22.4.0709.
Moschini-Carlos, V., Pompêo, M., Nishimura, P.Y., Armengol, J., 2017. Phytoplankton as trophic descriptors of a series of Mediterranean reservoirs (Catalonia, Spain). Fund. Appl. Limnol. 191, 37–52.
Moses, W., Saprygin, V., Gerasyuk, V., Povazhnyi, V., Berdnikov, S., Gitelson, A., 2019. OLCI-based NIR-red models for estimating chlorophyll-a concentration in productive coastal waters—a preliminary evaluation. Environ. Res. Commun. 1, 011002. https://doi.org/10.1088/2515-7620/aaf53c.
Pamula, A., Siva, P., Gholizadeh, H., Krzmarzick, M., Mausbach, W., Lampert, D., 2023. A remote sensing tool for near real-time monitoring of harmful algal blooms and turbidity in reservoirs. JAWRA. J. Am. Water Resour. Asso. 1–21. https://doi.org/10.1111/1752-1688.13121.
Pearson, L.A., Dittmann, E., Mazmouz, R., Ongley, S.E., D'Agostino, P.M., Neilan, B.A., 2016. The genetics, biosynthesis, and regulation of toxic specialized metabolites of cyanobacteria. Harmful Algae 54, 98–111. https://doi.org/10.1016/j.hal.2015.11.002.
Pereira-Sandoval, M., Urrego, E.P., Ruiz-Verdú, A., Tenjo, C., Delegido, J., Soria-Perpinyà, X., Vicente, E., Sória, J., Moreno, J., 2019. Calibration and validation of algorithms for the estimation of chlorophyll-a and Secchi depth in inland waters with Sentinel-2. Limnética 38 (1), 471–487. https://doi.org/10.23818/limn.38.27.
Pérez-González, R., Sòria-Perpinyà, X., Soria, J.M., Delegido, J., Urrego, P., Sendra, M.D., Ruíz-Verdú, A., Vicente, E., Moreno, J., 2021. Phycocyanin monitoring in some Spanish water bodies with sentinel-2 imagery. Water 13, 2866. https://doi.org/10.3390/w13202866.
Pompeo, M., Moschini-Carlos, V., 2020. Reservatórios que abastecem São Paulo: problemas e perspectivas. Instituto de Biociências, São Paulo. Retrieved from. http://ecologia.ib.usp.br/portal/publicacoes/.

Pompêo, M., Moschini-Carlos, V., 2022. Avaliação da degradação da qualidade da água do reservatório Carlos Botelho em Itirapina, São Paulo, Brasil, por meio de imagens do satélite Sentinel 2. Eng. Sanitária Ambient. 27 (2), 279–290. https://doi.org/10.1590/s1413-415220210002.

Pompêo, M., Moschini-Carlos, V., López-Doval, J.C., Abdalla-Martins, N., Cardoso-Silva, S., Freire, R.H.F., de Souza Beghelli, F.G., Brandimarte, A.L., Rosa, A.H., López, P., 2017. Nitrogen and phosphorus in cascade multi-system tropical reservoirs: water and sediment. Limnol. Rev. 17, 133–150. https://doi.org/10.1515/limre-2017-0013.

Pompêo, M., Moschini-Carlos, V., Bitencourt, M.D., Sòria-Perpinyà, X., Vicente, E., Delegido, J., 2021. Water quality assessment using Sentinel-2 imagery with estimates of chlorophyll-a, Secchi disk depth, and Cyanobacteria cell number: the Cantareira System reservoirs (São Paulo, Brazil). Environ. Sci. Pollut. Control Ser. 28 (26), 34990–35011. https://doi.org/10.1007/s11356-021-12975-x.

Qi, L., Hu, C., Duan, H., Cannizzaro, J., Ma, R., 2014. A novel MERIS algorithm to derive cyanobacterial phycocyanin pigment concentrations in a eutrophic lake: theoretical basis and practical considerations. Remote Sen. Environ. 154, 298–317. https://doi.org/10.1016/j.rse.2014.08.026.

Randolph, Kaylan, Wilson, Jeffrey, Tedesco, L.P., Li, Lin, Pascual, Denise, Soyeux, Emmanuel, 2008. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll-a and phycocyanin. Rem. Sens. Environ. 112, 4009–4019. https://doi.org/10.1016/j.rse.2008.06.002.

Renosh, P.R., Doxaran, D., Keukelaere, L.D., Gossn, J.I., 2020. Evaluation of atmospheric correction algorithms for sentinel-2-MSI and sentinel-3-OLCI in highly turbid estuarine waters. Rem. Sens. 12, 1285. https://doi.org/10.3390/rs12081285.

Shi, K., Zhang, Y., Li, Y., Li, L., Lv, H., Liu, X., 2015. Remote estimation of cyanobacteria-dominance in inland waters. Water Res. 1 (68), 217–226. https://doi.org/10.1016/j.watres.2014.10.019.

Shoaf, W.T., Lium, B.W., 1976. Improved extraction of chlorophyll a and b from algae using dimethyl sulphoxide. Limnol. Oceanogr. 21, 926–928.

Sòria-Perpinyà, X., Urrego, P., Pereira-Sandoval, M., Ruiz-Verdú, A., Peña, R., Soria, J.M., Delegido, J., Vicente, E., Moreno, J., 2019. Monitoring the ecological state of a hypertrophic lake (Albufera of València, Spain) using multitemporal sentinel-2 images. Limnética 38 (1), 457–469. https://doi.org/10.23818/limn.38.26.

Sòria-Perpinyà, X., Vicente, E., Urrego, P., Pereira-Sandoval, M., Tenjo, C., Ruíz-Verdú, A., Delegido, J., Soria, J.M., Peña, R., Moreno, J., 2021. Validation of water quality monitoring algorithms for sentinel-2 and sentinel-3 in mediterranean inland waters with *in situ* reflectance data. Water 13, 686. https://doi.org/10.3390/w13050686.

Sòria-Perpinyà, X., Delegido, J., Urrego, E.P., Ruíz-Verdú, A., Soria, J.M., Vicente, E., Moreno, J., 2022. Assessment of sentinel-2-MSI atmospheric correction processors and in situ spectrometry waters quality algorithms. Rem. Sens. 14, 4794. https://doi.org/10.3390/rs14194794.

Soriano-González, J., Urrego, E.P., Sòria-Perpinyà, X., Angelats, E., Alcaraz, C., Delegido, J., Ruíz-Verdú, A., Tenjo, C., Vicente, E., Moreno, J., 2022. Towards the combination of C2RCC processors for improving water quality retrieval in inland and coastal areas. Rem. Sens. 14 (5), 1124. https://doi.org/10.3390/rs14051124.

Sun, J., Liu, D., 2003. Geometric models for calculating cell biovolume and surface area for phytoplankton. J. Plankton Res. 25, 1331–1346. https://doi.org/10.1093/plankt/fbg096.

Toming, K., Kutser, T., Laas, A., Sepp, M., Paavel, B., Noges, T., 2016. First experiences in mapping lake water quality parameters with sentinel-2 MSI imagery. Rem. Sens. 8 (8), 1–14. https://doi.org/10.3390/rs8080640.

Torremorell, A., Hegoburu, C., Brandimarte, A.L., Rodrigues, E.H., Pompeo, M., Cardoso-Silva, S., Moschini- Carlos, V., Caputo, L., Fierro, P., Mojica, I., Matta, A.P.L., Donato, C., Jiménez, P., Molinero, J., Rios- Rouma, B., Goyenola, G., Iglesias, C., López-Rodriguez, A., Meerhoff, M., Pacheco, J.P., Mello, F.T., Rodrígues-Olarte, D., Gomez, M.B., Montoya, J.V., López-Doval, J.C., Navarro, E., 2021. Present and future threats for the ecological quality management of South American freshwater ecosystems. Inland Waters 11 (2), 125–140. https://doi.org/10.1080/20442041.2019.1608115.

Utermöhl, Hans, 1958. Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. SIL Commun. 9 (1), 1–38. . 10.1080/05384680.1958.11904091.

Virdis, S., Xue, W., Winijkul, E., Nitivattananon, V., Punpukdee, P., 2022. Remote sensing of tropical riverine water quality using sentinel-2 MSI and field observations. Ecol. Indic. 144, 109472. https://doi.org/10.1016/j.ecolind.2022.109472.

Warren, M.A., Simis, S.G.H., 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. Rem. Sens. Environ. 225, 267–289. https://doi.org/10.1016/j.rse.2019.03.018.

Woźniak, M., Bradtke, K., Darecki, M., Krężel, A., 2016. Empirical model for phycocyanin concentration estimation as an indicator of cyanobacterial bloom in the optically complex coastal waters of the Baltic Sea. Rem. Sens. 8 (3), 212. https://doi.org/10.3390/rs8030212.

Yan, Y., Bao, Z., Shao, J., 2018. Phycocyanin concentration retrieval in inland waters: a comparative review of the remote sensing techniques and algorithms. J. Great Lake. Res. 44 (4), 748–755. . 10.1016/j.jglr.2018.05.004.

Yang, X.E., Wu, X., Hao, H.L., He, Z.L., 2008. Mechanisms and assessment of water eutrophication. J. Zhejiang Univ. - Sci. B 9 (3), 197–209. https://doi.org/10.1631/jzus.B0710626.

Zanchett, G., Oliveira-Filho, E.C., 2013. Cyanobacteria and cyanotoxins: from impacts on aquatic ecosystems and human health to anticarcinogenic effects. Toxins 5, 1896–1917. https://doi.org/10.3390/toxins5101896.

Zorzal-Almeida, S., Salim, A., Andrade, M.R.M., Nascimento, M. de N., Bini, L.M., Bicudo, D.C., 2018. Effects of land use and spatial processes in water and surface sediment of tropical reservoirs at local and regional scales. Sci. Total Environ. 644, 237–246. https://doi.org/10.1016/j.scitotenv.2018.06.361.

Da Silva, E.F.F., Novo, E.M.L. de M., Lobo, F. de L., Barbosa, C.C.F., Noernberg, M.A., Rotta, L.H. da S., Cairo, C.T., Maciel, D.A., Flores Júnior, R., 2020. Optical water types found in Brazilian waters. Limnology 22, 57–68. https://doi.org/10.1007/s10201-020-00633-z.