

Article

Surface Water Quality Assessment through Remote Sensing Based on the Box–Cox Transformation and Linear Regression

Juan G. Loaiza ¹, Jesús Gabriel Rangel-Peraza ^{1,*} , Sergio Alberto Monjardín-Armenta ² , Yaneth A. Bustos-Terrones ³ , Erick R. Bandala ⁴ , Antonio J. Sanhouse-García ¹ and Sergio A. Rentería-Guevara ⁵ 

¹ Tecnológico Nacional de México/Instituto Tecnológico de Culiacán, Juan de Dios Bátiz 310, Col. Guadalupe, Culiacán 80220, Sinaloa, Mexico; d18170809@culiacan.tecnm.mx (J.G.L.); antonio.sg@culiacan.tecnm.mx (A.J.S.-G.)

² Facultad de Ciencias de la Tierra y el Espacio, Universidad Autónoma de Sinaloa, Circuito Interior Oriente, Cd Universitaria, Culiacán 80040, Sinaloa, Mexico; sergiomonjardin@uas.edu.mx

³ CONAHCYT-Instituto Tecnológico de Culiacán, Juan de Dios Bátiz 310, Col. Guadalupe, Culiacán 80220, Sinaloa, Mexico; yaneth.bt@culiacan.tecnm.mx

⁴ Division of Hydrologic Sciences, Desert Research Institute, 755 Flamingo Road, Las Vegas, NV 89119, USA; erick.bandala@dri.edu

⁵ Facultad de Ingeniería, Universidad Autónoma de Sinaloa, Circuito Interior Oriente, Cd Universitaria, Culiacán 80040, Sinaloa, Mexico; sergiorenteria@uas.edu.mx

* Correspondence: jesus.rp@culiacan.tecnm.mx

Abstract: A methodology to estimate surface water quality using remote sensing is presented based on Landsat satellite imagery and in situ measurements taken every six months at four separate sampling locations in a tropical reservoir from 2015 to 2019. The remote sensing methodology uses the Box–Cox transformation model to normalize data on three water quality parameters: total organic carbon (TOC), total dissolved solids (TDS), and chlorophyll a (Chl-a). After the Box–Cox transformation, a mathematical model was generated for every parameter using multiple linear regression to correlate normalized data and spectral reflectance from Landsat 8 imagery. Then, significant testing was conducted to discard spectral bands that did not show a statistically significant response ($\alpha = 0.05$) from the different water quality models. The r^2 values achieved for TOC, TDS, and Chl-a water quality models after the band discrimination process were found 0.926, 0.875, and 0.810, respectively, achieving a fair fitting to real water quality data measurements. Finally, a comparison between estimated and measured water quality values not previously used for model development was carried out to validate these models. In this validation process, a good fit of 98% and 93% was obtained for TDS and TOC, respectively, whereas an acceptable fit of 81% was obtained for Chl-a. This study proposes an interesting alternative for ordered and standardized steps applied to generate mathematical models for the estimation of TOC, TDS, and Chl-a based on water quality parameters measured in the field and using satellite images.

Keywords: surface water quality; remote sensing; Box–Cox optimization; linear modeling; Landsat imagery



Citation: Loaiza, J.G.; Rangel-Peraza, J.G.; Monjardín-Armenta, S.A.; Bustos-Terrones, Y.A.; Bandala, E.R.; Sanhouse-García, A.J.; Rentería-Guevara, S.A. Surface Water Quality Assessment through Remote Sensing Based on the Box–Cox Transformation and Linear Regression. *Water* **2023**, *15*, 2606. <https://doi.org/10.3390/w15142606>

Academic Editors: Mou Leong Tan, Fei Zhang, Xiaoping Wang and Chenfeng Wang

Received: 15 June 2023

Revised: 6 July 2023

Accepted: 12 July 2023

Published: 18 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Surface water quality monitoring is essential to assess the impacts of anthropogenic activities and natural phenomena [1,2], but it is labor-intensive, time-consuming, and costly [3]. Several studies have demonstrated that using remote sensing for water quality evaluation has significant advantages for surface water quality monitoring [4–7]. Remote sensing for water quality evaluation is based on measuring the radiance emerging from the water related to electromagnetic radiation that interacts with both suspended and dissolved matter through absorptive, refractive, and scattering mechanisms [8,9]. Specific

imagery bands are required to measure water quality parameters, but the remotely sensed reflectance may be influenced by external conditions, such as atmospheric and air–water interface effects, illumination conditions, or instrument characteristics [10]. To avoid these interferences, some authors have suggested observing the scattering and absorption characteristics of optically active constituents (OACs) to obtain accurate inherent optical properties [11]. However, identifying specific wavelengths for water quality estimation is a complex task because water constituents absorb and scatter light across the entire visible spectral range, which complicates their estimation from optical measurements [10].

The most employed methodologies to estimate surface water quality uses empirical approaches with multispectral sensors [12]. Water quality modeling using remote sensing is often carried out using normalized difference indices and spectral band ratios. Normalization can remove brightness variations, reducing the influence of atmospheric, and air–water surface effects [13–15]. Other water quality studies using remote sensing are based on water quality parameters and spectral reflectance multiple regression [16–20] consisting of obtaining correlations between water leaving radiance (L_w) and several optically active parameters such as chlorophyll-*a*, total suspended solids, and turbidity. Despite some studies having achieved satisfying results using broadband sensors, others report less accurate results because of the presence of suspended material in turbid and/or eutrophic water bodies [21].

Although significant advancements in mathematical models for surface water quality using reflectance values from satellite imagery are given, improving existing models using multiple linear regression to estimate water quality using Landsat 8 imagery remains an interesting pending research task. Recently, Sharaf El Din and Zhang [22] have proposed a regression-based technique to estimate surface water quality parameters using Landsat 8 OLI imagery. They propose a stepwise regression (SWR) to minimize the number of predictor variables and to maximize the precision of the water quality estimation. Highly accurate results were achieved when using the Landsat 8-based-SWR approach ($r^2 > 85\%$).

The present study proposes a series of ordered and standardized steps to generate mathematical models from a multiple linear regression analysis. The multiple regression analysis was carried out between the reflectance values of Landsat 8 images and the normalized concentrations of water quality parameters. Normalizing water quality data was used to eliminate the effects of certain errors that may be still present after the dataset validation procedures, such as outliers, censored values, seasonality trends, or serial correlations, and that can affect the accuracy of the data, making it more consistent, reliable, and suitable for further processing and analysis.

Some studies have normalized the dataset, but they were not used for the estimation of water quality from satellite imagery. For instance, Feng [23] developed normalized water quality indexes using band combination, and Qi et al. [24] normalized reflectance data for water quality estimation. Asadollahfardi et al. [25] suggested using the Box–Cox transformation for the normalization of water quality data.

The methodology proposes the use of the Box–Cox transformation to normalize data on three water quality parameters: total organic carbon (TOC), total dissolved solids (TDS), and chlorophyll *a* (Chl-*a*). After the Box–Cox transformation, a mathematical model was generated for every parameter using multiple linear regression between normalized data and spectral reflectance from Landsat 8 imagery. Using the proposed methodology, a surface water quality assessment was carried out in the Adolfo López Mateos (ALM) reservoir in Culiacan, Mexico. These mathematical models could be considered a crucial tool for decision-making since they could be used to estimate water quality during periods when field monitoring is not conducted.

2. Materials and Methods

2.1. Study Area

The ALM reservoir (Culiacan, Mexico) is in the Humaya River basin (25°05′25″, 25°20′15″ North, 107°33′00″, 107°15′00″ West) at 186.5 m above the sea level (m.a.s.l.). The

ALM dam is 105.5 m high, and 765 m long, considered one of the main sources of water for agricultural irrigation, power generation, fishing, and tourist activities [26–28], covering 11,354 ha and ranked tenth in Mexico according to its storage capacity [29] (Figure 1).

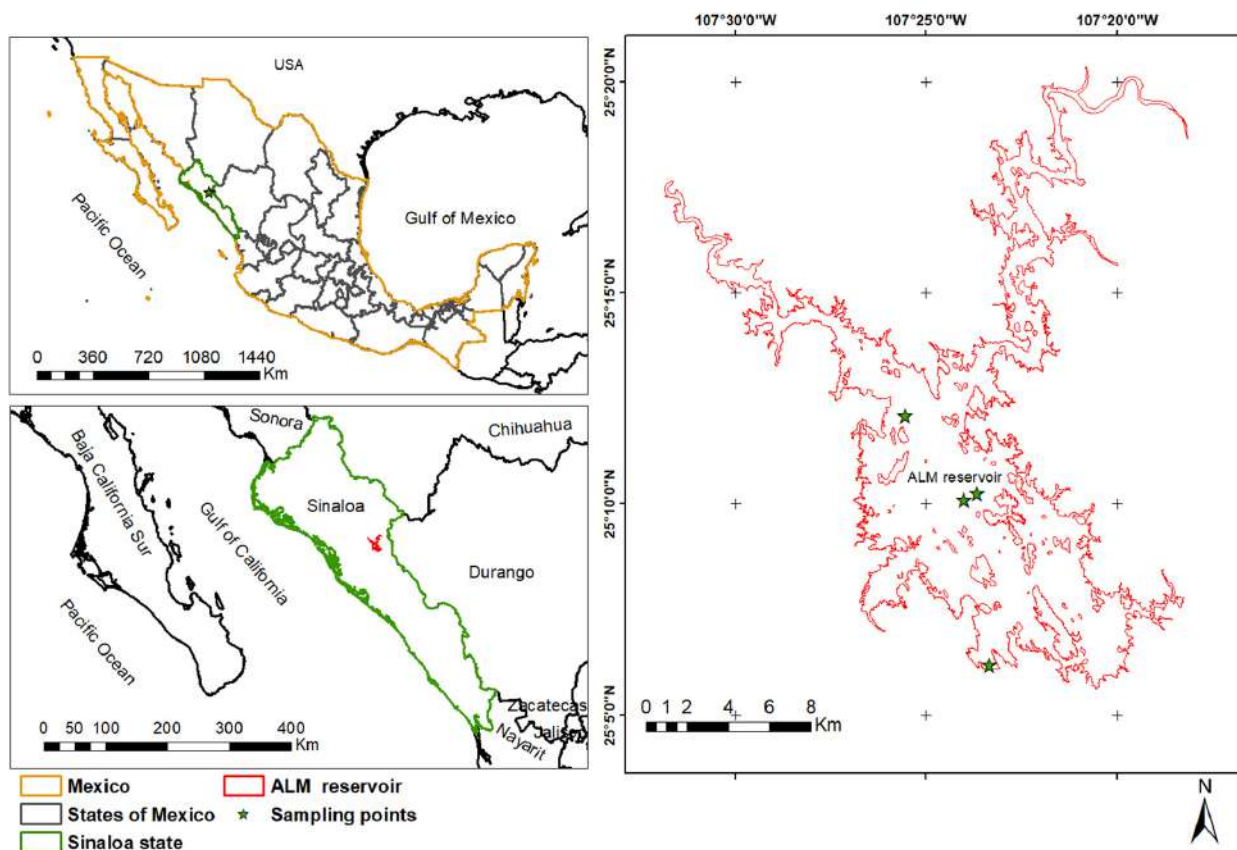


Figure 1. Geographic location of ALM reservoir.

The Humaya River basin is characterized by a warm humid climate toward the center and south, with summer rains. From the center toward the north, the climate is semi-warm sub-humid. The average annual temperature is 24.5 °C and the mean annual rainfall is 698.9 mm per year. The basin has a mountainous geography, with deep canyons, low mountains, highlands, and large plateaus with ravines. The basin elevation varies between 150 and 2300 m.a.s.l. [30].

The predominant vegetation is tropical deciduous forest, with small areas of pine-oak and pine forests toward the northwestern part [26]. Since high productivity is observed in the basin, a large proportion of the land is intended for agricultural and livestock activities [31]. According to Sanhouse-Garcia et al. [30] and Monjardin et al. [32], the Humaya River basin is affected by both natural (fires and frequent frosts) and anthropogenic (deforestation) factors that cause continuous and rapid changes in land use and aquatic ecosystems.

2.2. Methodology

Figure 2 sketches the methodology for water quality parameters estimation using satellite imagery, which was carried out in three phases: (i) processing Landsat 8 sensor imagery and reflectance data extraction, (ii) obtaining water quality data, and (iii) developing mathematical models for water quality parameters.

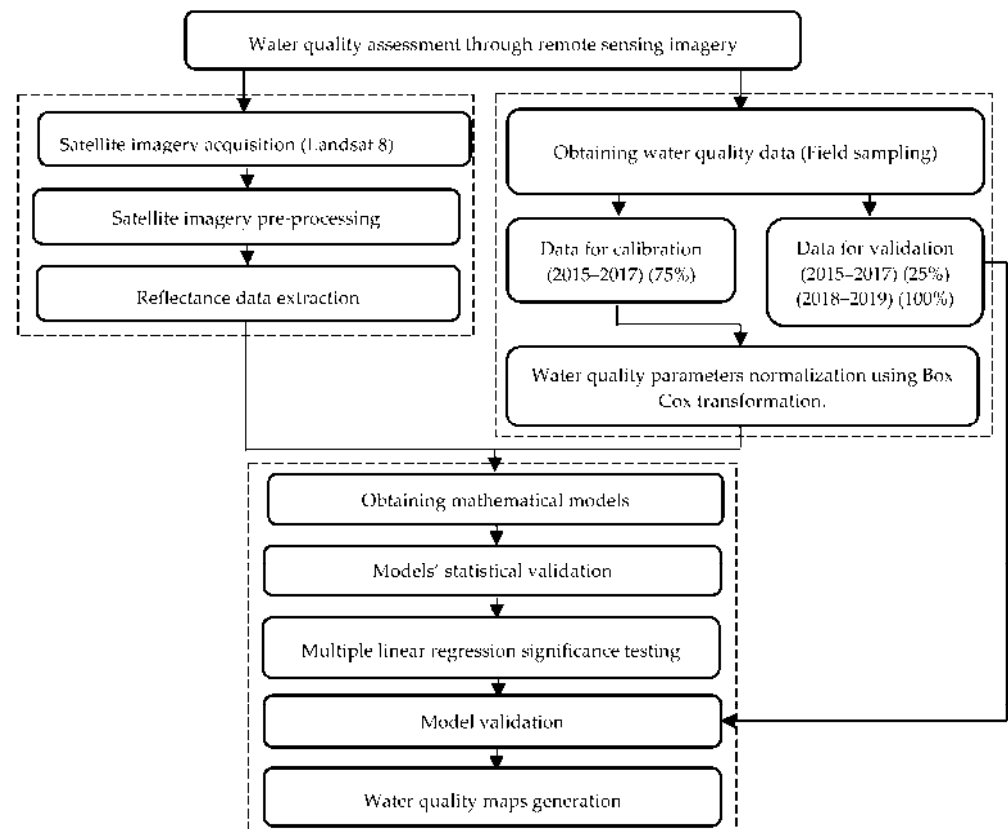


Figure 2. Proposed methodology to estimate water quality using satellite imagery.

2.2.1. Satellite Imagery Acquisition

Satellite imagery was obtained through the United States Geological Survey database [33] and coincided with the dates of water quality monitoring campaigns. GeoTiff level 1 (L1T) images from Landsat 8 were used. The L1T images are terrain-corrected; hence, these images already provide a radiometric and geodetic accuracy in a cartographic projection UTM (Universal Transversal of Mercator), referenced in WGS84 (World Geodetic System 1984). The images correspond to Path 32, Row 43 of the Landsat 8 sensor, which covered the reservoir surface during the study period (January 2015 to June 2019). Table 1 shows the acquisition dates of images from Landsat 8.

Table 1. Dates of acquisition of satellite images.

Sensor	Year	Acquisition Date	Path/Row
Landsat 8 OLI	2015	May 4th	32/43
		October 27th	
	2016	May 22nd	
		September 11th	
	2017	March 6th	
		September 30th	
	2018	February 2nd	
		October 2nd	
	2019	January 17th	

2.2.2. Imagery Pre-Processing

Landsat-8 imagery at level 1T was rescaled to the top of atmosphere (TOA) reflectance using radiometric rescaling coefficients [22,33]. This radiometric rescaling was performed in QGIS software based on TOA reflectance (Equations (1) and (2)) and using the semi-automatic classification plug-in [34].

$$\rho^* = M_\rho * Q_{cal} + A_\rho \quad (1)$$

where ρ^* is the TOA planetary reflectance, without correction for solar angle; M_ρ is the band-specific multiplicative rescaling factor from the metadata; Q_{cal} is the quantized and calibrated standard product pixel values (DN); and A_ρ is the band-specific additive rescaling factor from the metadata.

Since the reflectance obtained from the Landsat 8 data is not corrected for the solar zenith angle, the provided reflectance is generally too low and this error increases at high latitudes and in the cold season [22]. Therefore, a TOA reflectance correction for the solar zenith angle was performed using Equation (2).

$$\rho = \frac{\rho^*}{\cos(\theta_{SZ})} = \frac{\rho^*}{\sin(\theta_{SE})} \quad (2)$$

where ρ is the TOA planetary reflectance; θ_{SE} is the local sun elevation angle; θ_{SZ} local solar zenith angle; $\theta_{SZ} = 90^\circ - \theta_{SE}$.

Atmospheric correction processes were carried out using dark object subtraction (DOS). The basic assumption of the DOS method is that within the image some pixels are in complete shadow and their radiances measured at the satellite are due to atmospheric scattering (path radiance), selecting the spectral-band haze values that are correlated to each other [35]. This process was performed in QGIS by using Equations (3)–(7) [22].

$$\rho_{surface} = \left[\pi * (L_\lambda - L_P) * d^2 \right] / (T_V * \{ [E_{Sun\lambda} * \cos \theta_{SZ}] * T_Z \} + E_{down}) \quad (3)$$

$$L_P = L_{\lambda min} - L_{DO1\%} \quad (4)$$

$$L_{\lambda min} = M_L * DN_{min} + A_L \quad (5)$$

$$L_{DO1\%} = (0.01 * T_V * \{ [E_{Sun\lambda} * \cos \theta_{SZ}] * T_Z \} + E_{down}) / \left[\pi * d^2 \right] \quad (6)$$

$$E_{Sun\lambda} = \left[\pi * d^2 * Radiance_{max} \right] / [Reflectance_{max}] \quad (7)$$

where $\rho_{surface}$ is the surface reflectance; L_λ is the spectral radiance at the sensor's aperture; L_P is the path radiance due to atmospheric effects; d is the Earth–Sun distance in astronomical units; T_V is the atmospheric transmittance in the viewing direction; $E_{Sun\lambda}$ is the mean solar radiation; T_Z is the atmospheric transmittance in the illumination direction; E_{down} is the downwelling diffuse irradiance; $L_{\lambda min}$ is the radiance values correspond to the minimum pixel values; $L_{DO1\%}$ is the radiance of dark object; M_L is the radiance band-specific multiplicative rescaling factor; DN_{min} is the minimum pixel value; and A_L is the radiance band-specific additive rescaling factor.

2.2.3. Reflectance Data Extraction

The study area was delimited from satellite imagery with a polygon mask using QGIS software and a semi-automatic extraction utility. Then, reflectance data of corrected bands were extracted using the QGIS point sampling tools [36]. Landsat 8 has eleven bands, but in this study, the extraction process was carried out only for the bands B1, B2, B3, B4, B5, B6, and B7. Since the TOC, TDS, and Chl-a are known as optically active parameters and their

spectral responses are mainly in visible and near-infrared domains, B9 (cirrus band), B10, and B11 (infrared thermal bands) were excluded [37,38]. The B8 panchromatic band was also excluded from the extraction process since this band combines blue (B2), green (B3), and red (B4) bands with a greater spatial resolution and does not contain any additional wavelength-specific information.

2.2.4. Water Quality Monitoring

Water quality is monitored at the ALM reservoir every six months through sampling campaigns at four sampling sites by Mexico's National Water Commission (CONAGUA). CONAGUA is responsible for implementing processes to guarantee quality assurance/control (QA/QC). Hence, the sampling, transportation, and preservation of samples meet the appropriate Mexican standards, and the samples are analyzed in triplicate by an accredited laboratory, based on international standard methods for water analysis [39]. Official water quality information from CONAGUA has been used to assess water quality by other studies [26,28,40]. In this study, water quality data from 2015 to 2019 were used. To demonstrate the appropriateness of the proposed methodology, three water quality parameters were evaluated: total dissolved solids (TDS), chlorophyll (Chl-a), and total organic carbon (TOC). TOC, TDS, and Chl-a were measured based on APHA Methods 5310, 2510, and 10200H, respectively [39]. Past studies have shown that these parameters respond to the energy spectrum changes of reflected solar radiation from waterbodies [41,42].

2.2.5. Box–Cox Transformation of Water Quality Parameters

The Box–Cox transformation is a statistical technique to stabilize the variance of a certain dataset and ensure normal distribution of deviations around the model [43]. The main goal of data normalization is to adjust values to a common scale, achieving a standardized data format, which may facilitate comparison and analysis.

The Box–Cox transformation was only applied to the water quality parameters. For transformation, every data is raised to the λ_1 power after changing it to a certain amount λ_2 (often equal to 0). These transformations could be square roots, logarithms, reciprocals, and/or other common transformations (Table 2) [43]. Hence, the Box–Cox transformation (Equation (3)) is defined as a continuous function that varies as a function of power (λ) [44].

$$y' = (y + \lambda_2)^{\lambda_1}, \text{ for } \lambda_1 \neq 0 \quad (8)$$

where y' is the normalized water quality parameter, y is the originally measured water quality data, λ_1 and λ_2 are values that, when substituted in Equation (3), the standard deviation of y' will be zero.

Table 2. Box–Cox power transformation approaches.

Power	Transformation	Description
$\lambda_1 = 2$	$y' = y^2$	Square
$\lambda_1 = 1$	$y' = y$	Untransformed data
$\lambda_1 = 0.5$	$y' = \sqrt{y}$	Square root
$\lambda_1 = 0.33$	$y' = \sqrt[3]{y}$	Cube root
$\lambda_1 = 0^*$	$y' = \ln(y)$	Logarithm
$\lambda_1 = -0.5$	$y' = \frac{1}{\sqrt{y}}$	Inverse square root
$\lambda_1 = -1$	$y' = \frac{1}{y}$	Reciprocal

Note: * Note that as $\lambda_1 \rightarrow 0$, the power transformation approaches a logarithm.

The Statgraphics Centurion XVI software was used to perform the Box–Cox transformation of water quality data. Initially, the λ_1 values of 2, 1, 0.5, 0.33, 0, -0.5 , and 1 shown in Table 2 were investigated to determine which, if any, is most suitable. The software was used to solve for the optimum value of λ_1 using maximum likelihood estimation. Once the

Box–Cox transformation was performed, the normality of the data was evaluated using the Kolmogorov–Smirnov goodness of fit test.

2.2.6. Multiple Linear Regression

Multiple linear regression was carried out to correlate normalized water quality data and reflectance values from Landsat 8 imagery, where a fitted (or estimated) value is calculated using Equation (4) [45]:

$$y_i = b_0 + b_1x_1 + b_2x_2 + \dots + b_ix_i + \varepsilon_i \quad \text{foreach observation } i = 1, 2, \dots, n \quad (9)$$

where y_i is the estimated Box–Cox normalized value, $x_1, x_2, \dots, x_i$ are Landsat 8 imagery bands, b_0 is the intercept when all the predictors $x_1, x_2, \dots, x_i$ are all zero, $b_1, b_2, \dots, b_i$ are the linear regression coefficients obtained from the fitted values and ε is a random error corresponding to the n observations that are also assumed to be uncorrelated random variables [46].

2.2.7. Model Performance Evaluation

Two indicators were selected to evaluate model performance in estimating water quality: the coefficient of determination (r^2) and the root-mean-square error (RMSE). Equation (5) was used to estimate r^2 , which is a number between 0 and 1 that measures how well a model estimates an outcome [47]:

$$r^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (10)$$

where y_i is the measured water quality parameter, $\bar{y}_i$ is the average water quality parameter, $\hat{y}_i$ is the estimated water quality values, and n is the number of available data.

RMSE statistically assesses differences between values observed and estimated by the model, the higher the RMSE, the greater the difference between estimated and observed values. RMSE was calculated using Equation (6) [47,48]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{n}} \quad (11)$$

2.2.8. Multiple Linear Regression Significance Testing

The significance of individual regression coefficients in the multiple linear regression model was carried out using a t-test. This test measures the contribution of an independent variable while the remaining variables are still included in the model. For the model $y_i = b_0 + b_1x_1 + b_2x_2 + \dots + b_ix_i$ if the test is performed for b_1 , the significance of the variable x_1 is evaluated while controlling for the presence of the variables $x_2, \dots, x_i$ (i.e., the model $y_i = b_0 + b_2x_2 + \dots + b_ix_i$).

To determine whether $x_1, x_2, \dots, x_i$ variables are useful predictors in this model, the following null and alternative hypotheses were tested:

$$\begin{aligned} H_0: & b_i = 0 \\ H_1: & b_i \neq 0 \end{aligned}$$

To carry out this hypothesis test, a p -value was obtained for all coefficients in the model. Each p -value is based on a t-statistic calculated as:

$$t = \frac{b_i}{s_{b_i}} \quad (12)$$

where s_{bi} is the standard error of the regression coefficient b_i , calculated using Equation (8):

$$s_{bi} = \sqrt{\frac{\sum y_i^2 - b_0 \sum y_i - b_i \sum x_i y_i}{n - 2}} \quad (13)$$

The p -value is then compared with a significance level ($\alpha = 0.05$). This critical value is typically set for hypothesis testing.

2.2.9. Water Quality Model Validation

A model validation procedure was performed comparing estimated and measured water quality values not previously used for model development. In this study, the data used for model validation were 25% of the total field data for the 2015–2017 period and the total field data for 2018, as shown in Figure 2. RMSE and r^2 were used to estimate the models' fitness to field water quality measurements.

2.2.10. Water Quality Mapping

Estimated water quality parameters were used to generate simplified models resulting from the significant testing and validation process. These models were represented spatially and temporally using GIS tools (raster calculator) and employing QGIS software for TOC, TDS, and Chl-a estimation.

3. Results and Discussion

3.1. Water Quality from Field Sampling

Data of water quality parameters (TOC, TDS, and Chl-a) measured in the field are summarized in Figure 3. Figure 3a shows that TOC has similar spatial distribution in the reservoir remaining within the 3.4 to 5.4 mg/L range, with a slight increase in 2017. These results agree with values reported by Zhou et al. [49], where TOC concentrations of around 2.5 mg/L were found in a northeast China reservoir. Few studies have been carried out on the organic matter in the study area. Gonzalez-Farias [50] (2006) reported that the mean concentration of particulate organic carbon in the Culiacan River was 1.73 mg C/L. TDS also showed slight spatial variation (Figure 3b) where 92.4, 93.8, and 117.8 mg/L mean concentrations were observed in 2015, 2016, and 2017, respectively. These TDS concentrations were similar to those reported in other reservoirs located close to the ALM reservoir, such as Huites (124 mg/L), José Ortiz (137 mg/L), and Miguel Hidalgo (132 mg/L) reservoirs [51]. Chl-a values in the ALM reservoir ranged from 0.1 to 6.1 mg/m³. These values are within the reported by Fregoso-López et al. [52] who found a maximum Chl-a concentration of 3.4 mg/m³ and a minimum concentration of 0.3 mg/m³ in the Miguel Hidalgo y Costilla reservoir located in El Fuerte, Sinaloa.

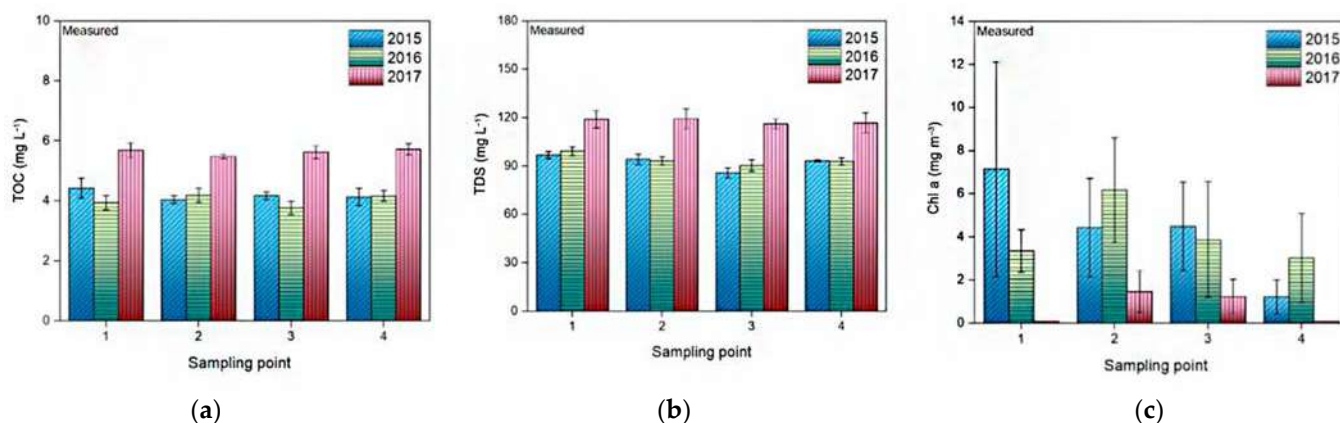


Figure 3. Descriptive analysis of the spatial and temporal distribution of water quality parameters (a) TOC, (b) TDS, and (c) Chl-a in the studied reservoir.

Figure 3 also shows that Chl-a levels in 2015 and 2016 were higher than those observed in 2017. This behavior is contrary to the behavior observed for TOC and TDS, where the highest concentration was found in 2017 when Chl-a was significantly reduced. Normally, there should be a correlation between these parameters because an excess of nitrogen and phosphorus in the water can lead to an overgrowth of algae or phytoplankton, resulting in higher Chl-a levels and an increase in organic matter and dissolved solids [53]. However, if the nutrients are depleted faster than they are being replenished, the algae eventually die and a decrease in Chl-a levels can be observed, and the decomposition of the excessive organic matter produced by the algae can increase TOC and TDS levels [54].

3.2. Box–Cox Transformation

Table 3 shows the resulting algorithms (a three-year normalized equation, from January 2015 to December 2017) obtained using the Box–Cox transformation for the water quality parameters studied. The mathematical models provided in Table 3 were rearranged for their later use to transform the estimated values to their original water quality units. The models showed a good fit, with r^2 values greater than 0.85 in all cases. Table 4 shows the results of the Kolmogorov–Smirnov goodness of fit test using the normalized water quality parameters. According to the Kolmogorov one-sample statistic (Dn) values and their respective p -values, the Box–Cox transformation was a good tool to normalize the water quality parameters.

Table 3. Mathematical models and r^2 values obtained from the Box–Cox transformation.

Parameter	Box–Cox Optimized Mathematical Model	r^2
TOC	Box–Cox (TOC) = $1 + (\text{TOC}^{1.3294} - 1) / (1.3294 \times 4.57379^{0.329397})$	0.96
TDS	Box–Cox (TDS) = $1 + (\text{TDS}^{4.16779} - 1) / (4.16779 \times 97.6453^{3.16779})$	0.88
Chl-a	Box–Cox (Chl-a) = $1 + (\text{Chl-a}^{0.333508} - 1) / (0.333508 \times 1.43584^{0.666492})$	0.85

Table 4. Kolmogorov–Smirnov test for the normalized water quality parameters.

Water Quality Normalized Parameter	Kolmogorov–Smirnov Test	
	Dn Value	p -Value
Chl-a	0.2393	0.2544
TDS	0.1644	0.7149
COT	0.1554	0.7769

Note: p -value > 0.05 suggests that there is not sufficient evidence to conclude that the data is not normally distributed.

3.3. Multiple Linear Regression Modeling and Discriminant Analysis

Multiple linear regression was performed between normalized water quality parameters and band reflectance values extracted from satellite imagery, using B1, B2, B3, B4, B5, B6, and B7 Landsat 8 bands. In this method, values from the Box–Cox transformation were considered dependent variables whereas band reflectance values from imagery captured by the sensor were considered independent variables. Table 5 shows the multiple linear regression models for the water quality parameters of the ALM reservoir. The multiple regression models showed fair fitting with r^2 greater than 0.80, considered satisfactory compared with the results of other empirical models used to estimate water quality through remote sensing [55,56].

A discriminant analysis was then performed to reduce the number of bands used in each model. A Student's t -test was carried out to assess whether each band has a significant effect on the water quality variables. p -values greater than or equal to 0.05 were considered not significant [57]. After these variables were eliminated, the hypothesis test was performed again using the simplified model, and the cycle was repeated until the model included only significant ($p < 0.05$) independent variables.

Table 5. Multiple linear regression models for the water quality parameters of the ALM reservoir based on the reflectance values of the Landsat 8 satellite images.

Parameter	Multiple Linear Regression Model	r^2	RMSE
TOC	Box–Cox (TOC) = $9.61963 - 700.238 \times B1 + 707.462 \times B2 - 39.2047 \times B3 - 25.1903 \times B4 - 18.2743 \times B5 + 216.704 \times B6 - 243.629 \times B7$	0.95	0.165
TDS	Box–Cox (TDS) = $34.849 - 3057.55 \times B1 + 4137.63 \times B2 - 2526.38 \times B3 + 2696.15 \times B4 + 1827.6 \times B5 - 6080.39 \times B6 + 2858.29 \times B7$	0.88	3.867
Chl-a	Box–Cox (Chl-a) = $-38.8501 + 212.068 \times B1 + 1213.14 \times B2 + 1207.01 \times B3 - 2935.1 \times B4 + 261.245 \times B5 - 2468.64 \times B6 + 3907.26 \times B7$	0.87	3.430

Table 6 shows the band discrimination process for TOC. As shown, six iterations were used to eliminate non-significant bands without significantly altering the fitting of the model. Table 7 shows the different p -value for TOC model parameters in each iteration. The initial mathematical model without band elimination (iteration 1) is the same as presented in Table 3. The student's t -test showed that the highest p -value was observed for B4 (Table 7), so this band was eliminated, and another hypothesis test was then performed generating a simplified model (Table 7, iteration 2) with a similar fit to the original. The band elimination process from TOC was repeated until only bands with a p -value less than 0.05 were included (this value is the established significance level for hypothesis tests). Through this process, B4, B5, B7, B6, and B3 were discriminated against for the TOC model. Despite the many bands removed, the simplified model presented a similar r^2 value compared to the initial one. This same methodology was carried out for each of the water quality parameters considered in this study and the results of the simplified models are shown in Table 8.

Table 6. Models resulting from the discriminant analysis for TOC and their degree of fit.

Iteration	Model	Discriminated Bands	r^2	RMSE
1	Box–Cox (TOC) = $9.61963 - 700.238 \times B1 + 707.462 \times B2 - 39.2047 \times B3 - 25.1903 \times B4 - 18.2743 \times B5 + 216.704 \times B6 - 243.629 \times B7$	0	0.9608	0.1658
2	Box–Cox (TOC) = $9.82457 - 711.379 \times B1 + 705.351 \times B2 - 48.6016 \times B3 - 25.9899 \times B5 + 245.128 \times B6 - 273.573 \times B7$	B4	0.9611	0.1676
3	Box–Cox (TOC) = $9.03939 - 661.472 \times B1 + 667.836 \times B2 - 45.8407 \times B3 + 147.039 \times B6 - 191.869 \times B7$	B4, B5	0.9423	0.1694
4	Box–Cox (TOC) = $8.92542 - 682.488 \times B1 + 677.616 \times B2 - 38.4808 \times B3 + 3.95873 \times B6$	B4, B5, B7	0.9521	0.1829
5	Box–Cox (TOC) = $8.87165 - 688.128 \times B1 + 688.322 \times B2 - 40.7919 \times B3$	B4, B5, B7, B6	0.9520	0.1835
6	Box–Cox (TOC) = $9.15197 - 620.429 \times B1 + 587.138 \times B2$	B4, B5, B7, B6, B3	0.9350	0.2024

Table 7. Statistical analysis for the discrimination of variables (bands).

Iteration	Parameter	Estimate	Standard Error	t-Statistic	p -Value
TOC model with all bands					
1	Constant	9.61963	2.15541	4.46302	0.0012
	B1	−700.238	100.765	−6.94922	<0.0000
	B2	707.462	109.424	6.4653	0.0001
	B3	−39.2047	53.3205	−0.735266	0.4791
	B4	−25.1903	136.359	−0.184735	0.8571
	B5	−18.2743	56.7695	−0.321903	0.7542
	B6	216.704	237.583	0.912119	0.3832
	B7	−243.629	244.321	−0.997167	0.3422
TOC model after discriminating B4					

Table 7. Cont.

Iteration	Parameter	Estimate	Standard Error	t-Statistic	p-Value
2	Constant	9.82457	1.46585	6.70228	<0.0000
	B1	−711.379	91.8242	−7.74718	<0.0000
	B2	705.351	97.8209	7.21064	<0.0000
	B3	−48.6016	23.9244	−2.03147	0.0671
	B5	−25.9899	40.4164	−0.643053	0.5334
	B6	245.128	185.017	1.32489	0.2121
	B7	−273.573	187.063	−1.46247	0.1716
TOC model after discriminating B4 and B5					
3	Constant	9.03939	0.526595	17.1657	<0.0000
	B1	−661.472	56.2812	−11.753	<0.0000
	B2	667.836	80.4374	8.30256	<0.0000
	B3	−45.8407	23.0887	−1.98542	0.0704
	B6	147.039	102.673	1.43211	0.1776
	B7	−191.869	134.751	−1.42388	0.18
TOC model after discriminating B4, B5, and B7					
4	Constant	8.92542	0.531479	16.7936	<0.0000
	B1	−682.488	55.898	−12.2095	<0.0000
	B2	677.616	83.0813	8.15606	<0.0000
	B3	−38.4808	23.4277	−1.64254	0.1244
	B6	3.95873	21.4212	0.184804	0.8562
TOC model after discriminating B4, B5, B7, and B6					
5	Constant	8.87165	0.47848	18.5413	<0.0000
	B1	−688.128	47.2176	−14.5735	<0.0000
	B2	688.322	60.099	11.4531	<0.0000
	B3	−40.7919	19.3641	−2.10657	0.0537
TOC model after discriminating B4, B5, B7, B6, and B3					
6	Constant	9.15197	0.510097	17.9416	<0.0000
	B1	−620.429	42.8204	−14.4891	<0.0000
	B2	587.138	44.5094	13.1913	<0.0000

Table 8. Final models after variable discrimination for the different parameters tested.

Parameter	Final Model	Bands Used	r^2
TOC	Box–Cox (TOC) = $9.15197 - 620.429 \times B1 + 587.138 \times B2$	B1, B2	0.9263
TDS	Box–Cox (TDS) = $55.7042 - 3387.46 \times B1 + 4108.64 \times B2 - 2874.84 \times B3 + 3514.37 \times B4 + 1386.56 \times B5 - 3490.39 \times B6$	B1, B2, B3, B4, B6	0.8753
Chl-a	Box–Cox (Chl-a) = $-24.4586 + 1204.69 \times B2 + 956.358 \times B3 - 2506.71 \times B4 + 996.356 \times B7$	B2, B3, B4, B7	0.8100

Figure 4a shows the estimated and measured TDS values in the ALM reservoir. The final TDS model accuracy ($r^2 = 0.875$; RMSE = 3.2613) can be considered satisfactory showing a better fit compared to other studies [58]. This could be attributed to the bands used for TDS estimation since low model accuracies have been reported in several studies that have only used B3, B4, and B5 bands (530 to 890 nm) of Landsat 8 [59,60]. According to Zhao et al. [61] (2020), the B3–B5 wavelength range (530–890 nm) can be used to characterize whether the water body contains phytoplankton chlorophyll (560–590 nm), cyanobacteria (620 nm), phycocyanin (650 nm), algae chlorophyll (675 nm), and suspended inorganic matter (810 nm). However, in this study, the discriminant analysis demonstrated that TDS estimation should be carried out using the bands B1, B2, B3, B4, B5, and B6 of Landsat 8. The use of a wider wavelength range could explain the satisfactory fit obtained since higher dissolved content of inorganic and organic substances could be detected, such as

the colored dissolved organic matter (CDOM) (420–555 nm). Our results agree with Maliki et al. [62], who successfully predicted the TDS of surface water in Bangladesh using Landsat 8 OLI and multiple linear models ($r^2 = 0.95$).

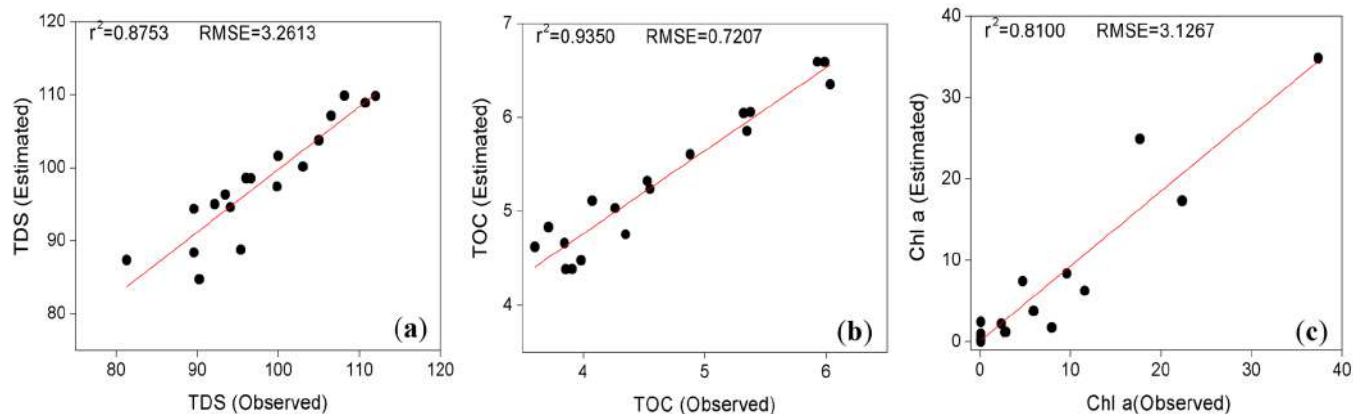


Figure 4. Estimated and observed values for (a) TOC, (b) TDS, and (c) Chl-a using the simplified models.

Chl-a is the most studied water quality parameter through remote sensing. Results for Chl-a showed a lower fit than TOC and TDS probably because of the normalization of water quality data. Several limitations were observed when Chl-a was normalized using the Box–Cox transformation, which generated the lowest r^2 value (see Table 3) in comparison with TOC and TDS, likely because Chl-a is a biological parameter showing exponential growth. In addition, Chl-a is more susceptible to seasonal variations related to physical, chemical, and climatic factors [63,64].

Mohsen et al. [65] used a multi-linear regression technique for the estimation of Chl-a through remote sensing using Landsat 7 bands B1 and B3 at Lake Burullus, Egypt, obtaining $r^2 = 0.86$ ($RMSE = 34.6$). Bohn et al. [66] reported $r^2 = 0.83$ estimating Chl-a using Landsat 7 bands B3 and B4. The accuracy of these models was similar to the results of this study ($r^2 = 0.81$, $RMSE = 3.1267$) (Figure 4c). In the Chl-a model, some bands (such as B2 and B7) appear in the final model generated and do not appear in other studies, such as the one performed by Bohn et al. [66]. This is because these studies estimate Chl-a by calculating predetermined indices such as the normalized difference vegetation index (NDVI), normalized area vegetation index (NAVI), enhanced vegetation index (EVI), and ratio vegetation index (RVI).

The results obtained in this study can be considered low compared to those reported by Tyler et al. [67], ($r^2 = 0.95$) for a linear mixture model used to estimate Chl-a in Lake Balaton, Hungary, using Landsat TM imagery. The accuracy of the water quality models can be improved by removing image interferences. For instance, in this study, the DOS atmospheric correction method was used which assumes that there are dark targets in the image, such as water and dense vegetation. But when the water body is turbid, such as the reservoir in this study, the reflection of water in the near-infrared band is close to 0, which leads to uncertainties of the atmospheric correction over water [68]. Other atmospheric correction methods have been proven to be effective for turbid waters, such as ACOLITE [69,70], ACIX-Aqua [71,72], iCOR [73], POLYMER [74], or MDM [68]. Thus, the performance of these algorithms on the regression models should be investigated in depth in further studies.

3.4. Model Validation

The remaining 25% of field water quality data (randomly selected data not previously used for model development) during the 2015–2017 period and 100% of the data obtained from January 2018 to June 2019 were used to validate the simplified models, showing a fair fit between estimated and observed data. The estimated and observed water quality data for the model development and validation is shown in Table S1. A good fit of 93%

and 81% was obtained for TOC and Chl-a, respectively (Figure 5b,c). TDS showed a good adjustment with 98% (Figure 5a). Therefore, this study suggests the high feasibility to develop mathematical models based on water quality parameters measured in the field and using satellite images.

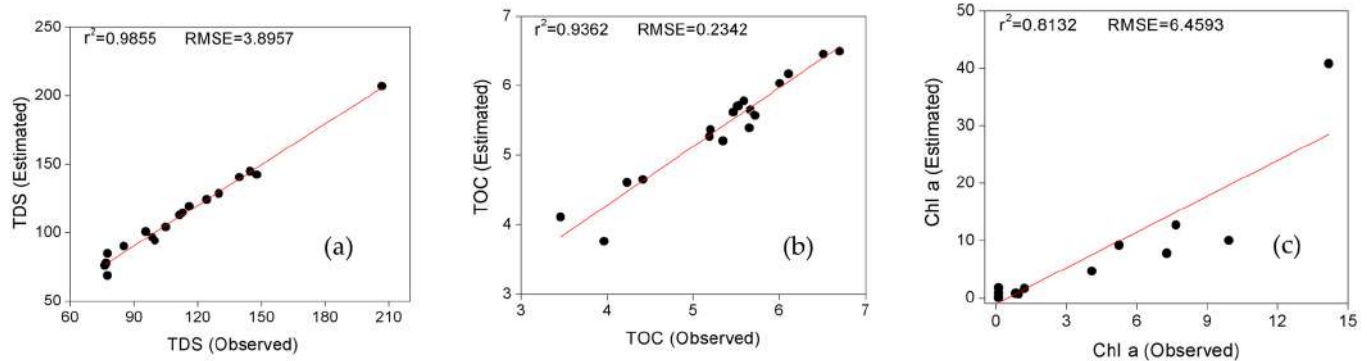


Figure 5. Validation of TDS (a), TOC (b), and Chl-a (c) with randomly selected data not previously used for model development.

3.5. Spatial and Temporal Distribution of Water Quality Parameters from Optimized Models

Simplified models were used to evaluate the spatial and temporal distribution of water quality parameters (TOC, TDS, and Chl-a) in the study area (Figures 6 and 7). Figure 6 shows the temporal behavior of TOC, TDS, and Chl-a in the ALM reservoir. In this figure, a time series comparison between the observed and estimated water quality values is shown. Only two observations are shown per year because water quality monitoring was carried out semi-annually. These observations represent the mean value of the four sample sites in the reservoir.

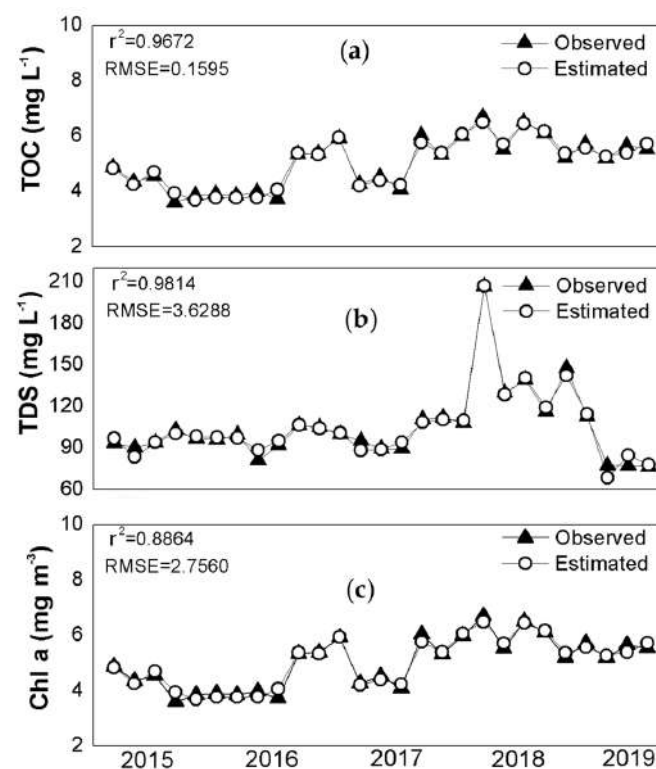


Figure 6. Temporal estimation of the estimated and observed values TOC (a), TDS (b), and Chl-a (c).

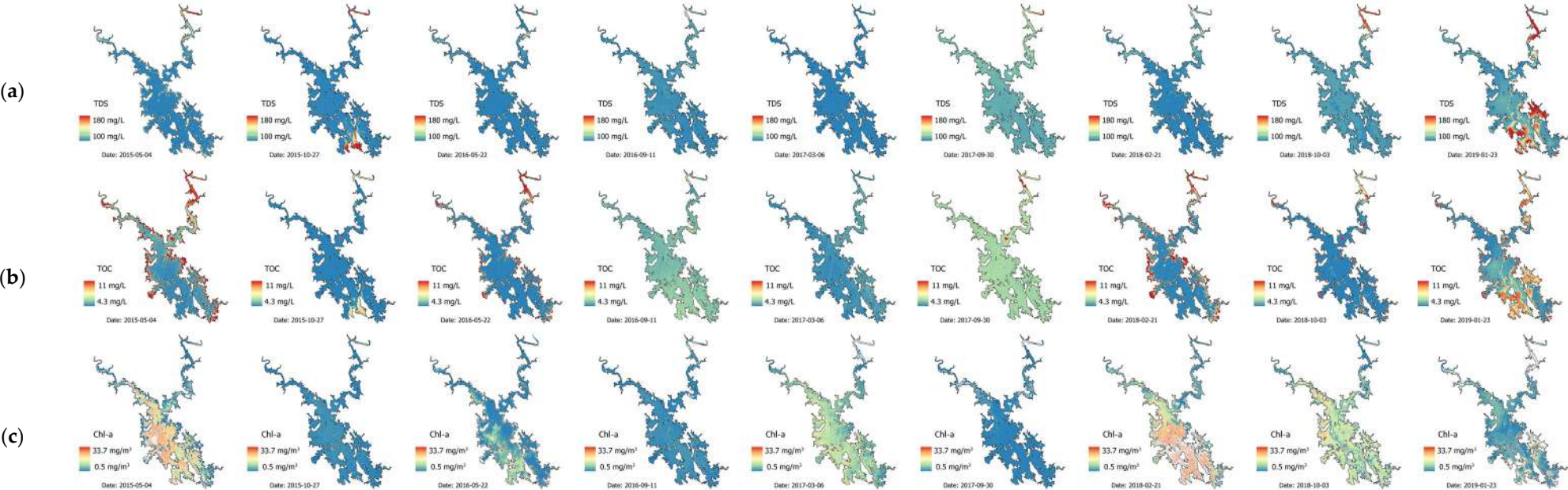


Figure 7. Maps generated from the simplified model for the estimation of water quality parameters.

The similarity between the values estimated using simplified models derived from remote sensing and field measurements was explained using the RMSE and coefficient of determination (r^2) indicators. A very low RMSE value was obtained when TOC observed and estimated concentrations were contrasted (Figure 6b). This figure also shows a fair estimation for TDS and Chl-a from the optimized models and satellite imagery (Figure 6a,c). The r^2 values obtained for TDS and Chl-a were higher than 0.81, which indicated a low variation between the observed and estimated water quality parameters.

This study estimated TOC, TDS, and Chl-a in the ALM reservoir on a bimonthly basis, despite the water quality information was available every six months. One of the main problems with empirical models is that they can generate unreliable results when applied at sites where they were not generated or on dates different from those used for their generation. The results demonstrated that the estimated water quality data agreed with the data observed in the ALM reservoir. These models were validated and suggest the feasibility of using Landsat imagery to estimate TDS, TOC, and Chl-a, which can be used as a decision-support tool for water quality management and policy analysis.

Figure 7 shows the spatial behavior of the TOC, TDS, and Chl-a through time (using the ISO 8601 date format YYYY-MM-DD). In this figure, a linear color gradient was used based on the RGB color model, where red and blue colors correspond to the highest and lowest concentrations of the water quality parameters depicted, respectively. Figure 7b shows that the TOC concentration in the ALM reservoir is higher during September and October, corresponding to the rainy season, associated with the entry of organic matter into the waterbody by runoff. Similarly, TOC concentrations of the nearshore area were higher than those within the ALM reservoir. According to this figure, the maximum TOC content occurred in the ALM upper reaches during May, when the water level in the reservoir is very low. Therefore, TOC behavior in the ALM reservoir is highly related to the biogeochemical processes of organic carbon. Hence, continuous monitoring of TOC using remote sensing could provide a quantitative basis for the estimation of carbon dioxide emission and sediment accumulation.

TDS showed a slight increase during the rainy season as runoff incorporates mineral salts into the reservoir (Figure 7a). Chl-a showed a slight increase in the first months of the year (Figure 7c) probably related to the lentic regime of the reservoir and the absence of rain. The spatial water quality variation observed in this study corresponds to water characteristics observed in waterbodies located in tropical regions [26,27,40,75].

4. Conclusions

This study proposed a methodology to estimate water quality parameters using satellite images. We proposed a methodology for band selection, discrimination, and water quality modeling based on ordered and standardized steps. However, it is important to highlight that this methodology was only validated for TOC, TDS, and Chl-a in the ALM reservoir in Mexico. Further studies should be focused on obtaining data from other water bodies to verify whether the methodology could be generalized.

The Box–Cox normalization proved to be effective in normalizing field water quality data, which was then used to find an optimal relationship with reflectance data from satellite bands. The models proposed were found robust since high coefficient of determination (r^2) values were obtained for the different water quality parameters estimated at the different stages (model development, discrimination, and validation). The obtained models were then used to estimate water quality parameters during periods where field monitoring was not conducted, which represents a crucial tool for decision-making.

This study provides an economical and effective alternative to monitor the water quality of a large water body in a short time based on a standardized repetitive basis. The methodology provides the spatial and temporal behavior of surface water quality, which can be used for water resources management. In this sense, this tool could contribute to improving the monitoring frameworks in many developing countries in the world, which

are limited by the expensive and time-consuming traditional methods for assessing and monitoring water quality.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/w15142606/s1>, Table S1. Estimated and observed water quality data for the model development and validation.

Author Contributions: J.G.L.: Conceptualization, Methodology, Investigation, Writing—Original draft preparation, and Writing—Reviewing and Editing. J.G.R.-P.: Conceptualization, Writing—Original draft preparation, Writing—Reviewing and Editing, and Project administration. S.A.M.-A.: Writing—Original draft preparation, Software, Formal analysis, Visualization, Resources, and Supervision. Y.A.B.-T.: Resources, Data Curation, Supervision, and Writing—Reviewing and Editing. E.R.B.: Software, Visualization, Validation, and Writing—Reviewing and Editing. A.J.S.-G.: Conceptualization, Investigation, Validation, and Supervision. S.A.R.-G.: Formal analysis, Visualization, and Writing—Reviewing and Editing. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Autonomous University of Sinaloa (PROFAPI—PRO_A1_012) and by the Tecnológico Nacional de México (Convocatoria Proyectos de Investigación Científica, Desarrollo Tecnológico e Innovación 2023—17107.23-P).

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank TecNM/Instituto Tecnológico de Culiacán and the Autonomous University of Sinaloa for providing the infrastructure to carry out this work and CONAH-CYT for the scholarship provided to the first author.

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

References

1. Ziemińska-Stolarska, A.; Kempa, M. Modeling and Monitoring of Hydrodynamics and Surface Water Quality in the Sulejów Dam Reservoir, Poland. *Water* **2021**, *13*, 296. [\[CrossRef\]](#)
2. Posthuma, L.; Zijp, M.C.; De Zwart, D.; Van de Meent, D.; Globevnik, L.; Koprivsek, M.; Birk, S. Chemical pollution imposes limitations to the ecological status of European surface waters. *Sci. Rep.* **2020**, *10*, 14825. [\[CrossRef\]](#)
3. Ryu, J.H. UAS-based real-time water quality monitoring, sampling, and visualization platform (UASWQP). *HardwareX* **2022**, *11*, e00277. [\[CrossRef\]](#)
4. Schaeffer, B.A.; Schaeffer, K.G.; Keith, D.; Lunetta, R.S.; Conmy, R.; Gould, R.W. Barriers to adopting satellite remote sensing for water quality management. *Int. J. Remote Sens.* **2013**, *34*, 7534–7544. [\[CrossRef\]](#)
5. Sayers, M.J.; Bosse, K.R.; Shuchman, R.A.; Ruberg, S.A.; Fahnenstiel, G.L.; Leshkevich, G.A.; Stuart, D.G.; Johengen, T.H.; Burtner, A.M.; Palladino, D. Spatial and temporal variability of inherent and apparent optical properties in western Lake Erie: Implications for water quality remote sensing. *J. Great Lakes Res.* **2019**, *45*, 490–507. [\[CrossRef\]](#)
6. Zhang, Y.; Kong, X.; Deng, L.; Liu, Y. Monitor water quality through retrieving water quality parameters from hyperspectral images using graph convolution network with superposition of multi-point effect: A case study in Maozhou River. *J. Environ. Manag.* **2023**, *342*, 118283. [\[CrossRef\]](#)
7. Zhang, D.; Li, X.; Huang, Y.; Zhang, L.; Zhu, Z.; Sun, X.; Lan, Z.; Guo, W. Hyperspectral remote sensing technology for water quality monitoring: Knowledge graph analysis and frontier trend. *Front. Environ. Sci.* **2023**, *11*, 1133325.
8. Dev, P.J.; Shanmugam, P. A new theory and its application to remove the effect of surface-reflected light in above-surface radiance data from clear and turbid waters. *J. Quant. Spectrosc. Radiat. Transf.* **2014**, *142*, 75–92. [\[CrossRef\]](#)
9. Mascarenhas, V.; Keck, T. Marine Optics and Ocean Color Remote Sensing. In *YOUWARES 8—Oceans across Boundaries: Learning from Each Other*; Jungblut, S., Liebich, V., Bode, M., Eds.; Conference Paper; Springer: Cham, Switzerland, 2018.
10. Gholizadeh, M.H.; Melesse, A.M.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors* **2016**, *16*, 1298. [\[CrossRef\]](#)
11. Chuvieco, E. Digital Image Processing (I): From Raw to Corrected Data. In *Fundamentals of Satellite Remote Sensing*; Chuvieco, E., Ed.; CRC Press: Boca Raton, FL, USA; London, UK, 2020; pp. 153–234.
12. Markogianni, V.; Kalivas, D.; Petropoulos, G.; Dimitriou, E. Analysis on the feasibility of Landsat 8 imagery for water quality parameters assessment in an oligotrophic Mediterranean lake. *J. Geotech. Eng.* **2017**, *11*, 906–914.
13. Ahmed, M.; Mumtaz, R.; Baig, S.; Zaidi, S.M.H. Assessment of correlation amongst physico-chemical, topographical, geological, lithological and soil type parameters for measuring water quality of Rawal watershed using remote sensing. *Water Supply* **2022**, *22*, 3645–3660. [\[CrossRef\]](#)

14. Khalil, M.T.; Saad, A.; Ahmed, M.; El Kafrawy, S.B.; Emam, W.W. Integrated field study, remote sensing and GIS approach for assessing and monitoring some chemical water quality parameters in Bardawil lagoon, Egypt. *Int. J. Innov. Res. Sci. Eng. Technol.* **2016**, *5*, 10–15680.
15. Theologou, I.; Patelaki, M.; Karantzas, K. Can single empirical algorithms accurately predict inland shallow water quality status from high resolution, multi-sensor, multi-temporal satellite data? *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 1511. [\[CrossRef\]](#)
16. Bonansea, M.; Bazán, R.; Ledesma, C.; Rodriguez, C.; Pinotti, L. Monitoring of regional lake water clarity using Landsat imagery. *Hydrol. Res.* **2015**, *46*, 661–670. [\[CrossRef\]](#)
17. Valentini, M.; dos Santos, G.B.; Muller, B. Multiple linear regression analysis (MLR) applied for modeling a new WQI equation for monitoring the water quality of Mirim Lagoon, in the state of Rio Grande do Sul—Brazil. *SN Appl. Sci.* **2021**, *3*, 70. [\[CrossRef\]](#)
18. Najafzadeh, M.; Homaei, F.; Farhadi, H. Reliability assessment of water quality index based on guidelines of national sanitation foundation in natural streams: Integration of remote sensing and data-driven models. *Artif. Intell. Rev.* **2021**, *54*, 4619–4651. [\[CrossRef\]](#)
19. Kadam, A.K.; Wagh, V.M.; Muley, A.A.; Umrikar, B.N.; Sankhua, R.N. Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India. *Model. Earth Syst. Environ.* **2019**, *5*, 951–962. [\[CrossRef\]](#)
20. Ewaid, S.H.; Abed, S.A.; Kadhum, S.A. Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis. *Environ. Technol. Innov.* **2018**, *11*, 390–398. [\[CrossRef\]](#)
21. Wang, X.; Yang, W. Water quality monitoring and evaluation using remote sensing techniques in China: A systematic review. *Ecosyst. Health Sustain.* **2019**, *5*, 47–56. [\[CrossRef\]](#)
22. Sharaf El Din, E.; Zhang, Y. Estimation of both optical and nonoptical surface water quality parameters using Landsat 8 OLI imagery and statistical techniques. *J. Appl. Remote Sens.* **2017**, *11*, 046008. [\[CrossRef\]](#)
23. Feng, L.; Hu, C.; Han, X.; Chen, X.; Qi, L. Long-term distribution patterns of chlorophyll-a concentration in China's largest freshwater lake: MERIS full-resolution observations with a practical approach. *Remote Sens.* **2014**, *7*, 275–299. [\[CrossRef\]](#)
24. Qi, L.; Hu, C.; Duan, H.; Barnes, B.B.; Ma, R. An EOF-based algorithm to estimate chlorophyll a concentrations in Taihu Lake from MODIS land-band measurements: Implications for near real-time applications and forecasting models. *Remote Sens.* **2014**, *6*, 10694–10715. [\[CrossRef\]](#)
25. Asadollahfardi, G.; Heidarzadeh, N.; Mosalli, A.; Sekhavati, A. Optimization of water quality monitoring stations using genetic algorithm, a case study, Sefid-Rud River, Iran. *Adv. Environ. Res.* **2018**, *7*, 87–107.
26. Quevedo-Castro, A.; Rangel-Peraza, J.G.; Bandala, E.; Amabilis-Sosa, L.; Rodríguez-Mata, A.; Bustos-Terrones, Y. Developing a water quality index in a tropical reservoir using a measure of multiparameters. *J. Water Sanit. Hyg. Dev.* **2018**, *8*, 752–766. [\[CrossRef\]](#)
27. Quevedo-Castro, A.; Lopez, J.L.; Rangel-Peraza, J.G.; Bandala, E.; Bustos-Terrones, Y. Study of the water quality of a tropical reservoir. *Environments* **2019**, *6*, 7. [\[CrossRef\]](#)
28. Quevedo-Castro, A.; Bandala, E.R.; Rangel-Peraza, J.G.; Amabilis-Sosa, L.E.; Sanhouse-García, A.; Bustos-Terrones, Y.A. Temporal and spatial study of water quality and trophic evaluation of a large tropical reservoir. *Environments* **2019**, *6*, 61. [\[CrossRef\]](#)
29. CONAGUA. Subdirección General Técnica. Presas de México. Available online: <http://sina.conagua.gob.mx/sina/tema.php?tema=presasPrincipalesandver=reporteando=2andn=nacional> (accessed on 13 November 2022).
30. Sanhouse-Garcia, A.J.; Bustos-Terrones, Y.; Rangel-Peraza, J.G.; Quevedo-Castro, A.; Pacheco, C. Multi-temporal analysis for land use and land cover changes in an agricultural region using open source tools. *Remote Sens. Appl. Soc. Environ.* **2017**, *8*, 278–290. [\[CrossRef\]](#)
31. INEGI. *Compendio de Información Geográfica Municipal 2010*; Badiraguato; Instituto Nacional de Estadística y Geografía: Aguascalientes, Mexico, 2015; Available online: https://www.inegi.org.mx/contenidos/app/mexicocifras/datos_geograficos/25/25003.pdf (accessed on 26 February 2023).
32. Monjardin-Armenta, S.A.; Plata-Rocha, W.; Pacheco-Angulo, C.E.; Franco-Ochoa, C.; Rangel-Peraza, J.G. Geospatial Simulation Model of Deforestation and Reforestation Using Multicriteria Evaluation. *Sustainability* **2020**, *12*, 10387. [\[CrossRef\]](#)
33. USGS. United States Geological Survey. Earth Explorer. 2021. Available online: <https://www.usgs.gov/landsat-missions/using-usgs-landsat-level-1-data-product> (accessed on 2 July 2023).
34. Congedo, L. Semi-automatic classification plugin documentation. *Release* **2016**, *4*, 29.
35. Chavez, P.S. Image-based atmospheric corrections—revisited and improved. *Photogramm. Eng. Remote Sens.* **1996**, *62*, 1025–1036.
36. Prieto-Amparan, J.A.; Villarreal-Guerrero, F.; Martinez-Salvador, M.; Manjarrez-Domínguez, C.; Santellano-Estrada, E.; Pinedo-Alvarez, A. Atmospheric and radiometric correction algorithms for the multitemporal assessment of grasslands productivity. *Remote Sens.* **2018**, *10*, 219. [\[CrossRef\]](#)
37. Huguet, A.; Vacher, L.; Relexans, S.; Saubusse, S.; Froidefond, J.M.; Parlanti, E. Properties of fluorescent dissolved organic matter in the Gironde Estuary. *Org. Geochem.* **2009**, *40*, 706–719. [\[CrossRef\]](#)
38. Hansen, C.H.; Williams, G.P.; Adjei, Z.; Barlow, A.; Nelson, E.J.; Miller, A.W. Reservoir water quality monitoring using remote sensing with seasonal models: Case study of five central-Utah reservoirs. *Lake Reserv. Manag.* **2015**, *31*, 225–240. [\[CrossRef\]](#)
39. APHA. *Standard Methods for the Examination of Water and Wastewater*, 18th ed.; American Public Health Association: Washington, DC, USA, 1992.

40. Loaiza, J.G.; Rangel-Peraza, J.G.; Sanhouse-García, A.J.; Monjardín-Armenta, S.A.; Mora-Félix, Z.D.; Bustos-Terrones, Y.A. Assessment of Water Quality in A Tropical Reservoir in Mexico: Seasonal, Spatial and Multivariable Analysis. *Int. J. Environ. Res. Public Health* **2021**, *18*, 7456. [\[CrossRef\]](#)
41. Ritchie, J.C.; Zimba, P.V.; Everitt, J.H. Remote sensing techniques to assess water quality. *Photogramm. Eng. Remote Sens.* **2003**, *69*, 695–704. [\[CrossRef\]](#)
42. Elkorashey, R.M. Utilizing chemometric techniques to evaluate water quality spatial and temporal variation. A case study: Bahr El-Baqar drain-Egypt. *Environ. Technol. Innov.* **2022**, *26*, 102332. [\[CrossRef\]](#)
43. Vélez, J.I.; Correa, J.C.; Marmolejo-Ramos, F. A new approach to the Box–Cox transformation. *Front. Appl. Math. Stat.* **2015**, *1*, 12. [\[CrossRef\]](#)
44. Peterson, R.A. Finding Optimal Normalizing Transformations via best Normalize. *R J.* **2021**, *13*, 3010–3329. [\[CrossRef\]](#)
45. Etemadi, S.; Khashei, M. Etemadi multiple linear regression. *Measurement* **2021**, *186*, 110080. [\[CrossRef\]](#)
46. Ouma, Y.O.; Okuku, C.O.; Njau, E.N. Use of artificial neural networks and multiple linear regression model for the prediction of dissolved oxygen in rivers: Case study of hydrographic basin of River Nyando, Kenya. *Complexity* **2020**, 9570789, 23. [\[CrossRef\]](#)
47. Abunama, T.; Othman, F.; Ansari, M.; El-Shafie, A. Leachate generation rate modeling using artificial intelligence algorithms aided by input optimization method for an MSW landfill. *Environ. Sci. Pollut. Res. Int.* **2019**, *26*, 3368–3381. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Jahani, A.; Rayegani, B. Forest landscape visual quality evaluation using artificial intelligence techniques as a decision support system. *Stoch. Environ. Res. Risk Assess.* **2020**, *34*, 1473–1486. [\[CrossRef\]](#)
49. Zhou, Z.; Huang, T.; Ma, W.; Li, Y.; Zeng, K. Impacts of water quality variation and rainfall runoff on Jinpen Dam, in Northwest China. *Water Sci. Eng.* **2015**, *8*, 301–308. [\[CrossRef\]](#)
50. Gonzalez-Farias, F.A.; Hernandez-Garza, M.d.R.; Gonzalez, G.D. Organic carbon and pesticide pollution in a tropical coastal lagoon-estuarine system in Northwest Mexico. *Int. J. Environ. Pollut.* **2006**, *26*, 234. [\[CrossRef\]](#)
51. Rodríguez, H.B.; González, L.C.; Trigueros, J.A.; Ávila, J.A.; Arciniega, M.A. Calidad del agua: Caracterización espacial en época de sequía en el río Fuerte, Sinaloa, México. *Rev. Cienc. Desde Occident.* **2016**, *3*, 35–47.
52. Fregoso-López, M.G.; Armienta-Hernández, M.A.; Alarcón-Silvas, S.G.; Ramírez-Rochín, J.; Fierro-Sañudo, J.F.; Páez-Osuna, F. Assessment of nutrient contamination in the waters of the El Fuerte River, southern Gulf of California, Mexico. *Environ. Monit. Assess.* **2020**, *192*, 417. [\[CrossRef\]](#)
53. Zhang, H.; Richardson, P.A.; Belayneh, B.E.; Ristvey, A.; Lea-Cox, J.; Copes, W.E.; Moorman, G.W.; Hong, C. Comparative Analysis of Water Quality between the Runoff Entrance and Middle of Recycling Irrigation Reservoirs. *Water* **2015**, *7*, 3861–3877. [\[CrossRef\]](#)
54. Fang, J.; Wu, F.; Xiong, Y.; Wang, S. A comparison of the distribution and sources of organic matter in surface sediments collected from northwestern and southwestern plateau lakes in China. *J. Limnol.* **2017**, *76*, 571–580. [\[CrossRef\]](#)
55. Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Estimating chlorophyll-a of inland water bodies in Greece based on landsat data. *Remote Sens.* **2020**, *12*, 2087. [\[CrossRef\]](#)
56. Abbas, M.M.; Melesse, A.M.; Scinto, L.J.; Rehage, J.S. Satellite estimation of chlorophyll-a using moderate resolution imaging spectroradiometer (MODIS) sensor in shallow coastal water bodies: Validation and improvement. *Water* **2019**, *11*, 1621. [\[CrossRef\]](#)
57. Shinmura, S.; Shinmura, S. *New Theory of Discriminant Analysis*; Springer: Singapore, 2016; pp. 81–97.
58. Batur, E.; Maktav, D. Assessment of Surface Water Quality by Using Satellite Images Fusion Based on PCA Method in the Lake Gala, Turkey. *IEEE Trans. Geosci. Remote Sens.* **2018**, *57*, 2983–2989. [\[CrossRef\]](#)
59. Kumar, M.; Kumar, M.; Denis, D.M.; Verma, O.P.; Mahato, L.L.; Pandey, K. Investigating water quality of an urban water body using ground and space observations. *Spat. Inf. Res.* **2021**, *29*, 897–906. [\[CrossRef\]](#)
60. Mejía Ávila, D.; Torres-Bejarano, F.; Martínez Lara, Z. Spectral indices for estimating total dissolved solids in freshwater wetlands using semi-empirical models. A case study of Guartínaja and Momil wetlands. *Int. J. Remote Sens.* **2022**, *43*, 2156–2184. [\[CrossRef\]](#)
61. Zhao, J.; Zhang, F.; Chen, S.; Wang, C.; Chen, J.; Zhou, H.; Xue, Y. Remote Sensing Evaluation of Total Suspended Solids Dynamic with Markov Model: A Case Study of Inland Reservoir across Administrative Boundary in South China. *Sensors* **2020**, *20*, 6911. [\[CrossRef\]](#) [\[PubMed\]](#)
62. Maliki, A.A.A.; Chabuk, A.; Sultan, M.A.; Hashim, B.M.; Hussain, H.M.; Al-Ansari, N. Estimation of Total Dissolved Solids in Water Bodies by Spectral Indices Case Study: Shatt al-Arab River. *Water Air Soil Pollut.* **2020**, *231*, 482. [\[CrossRef\]](#)
63. Obaid, A.A.; Ali, K.A.; Abiye, T.A.; Adam, E.M. Assessing the utility of using current generation high-resolution satellites (Sentinel 2 and Landsat 8) to monitor large water supply dam in South Africa. *Remote Sens. Appl. Soc. Environ.* **2021**, *22*, 100521. [\[CrossRef\]](#)
64. Lin, L.; Wang, F.; Chen, H.; Fang, H.; Zhang, T.; Cao, W. Ecological health assessments of rivers with multiple dams based on the biological integrity of phytoplankton: A case study of North Creek of Jiulong River. *Ecol. Indic.* **2021**, *121*, 106998. [\[CrossRef\]](#)
65. Mohsen, A.; Elshemy, M.; Zeidan, B. Water quality monitoring of Lake Burullus (Egypt) using Landsat satellite imagery. *Environ. Sci. Pollut. Res.* **2020**, *28*, 15687–15700. [\[CrossRef\]](#)
66. Bohn, V.Y.; Carmona, F.; Rivas, R.; Lagomarsino, L.; Diovisalvi, N.; Zagarese, H.E. Development of an empirical model for chlorophyll-a and Secchi Disk Depth estimation for a Pampean shallow lake (Argentina). *Egypt. J. Remote Sens. Space Sci.* **2018**, *21*, 183–191. [\[CrossRef\]](#)

67. Tyler, A.N.; Svab, E.; Preston, T.; Présing, M.; Kovács, W.A. Remote sensing of the water quality of shallow lakes: A mixture modelling approach to quantifying phytoplankton in water characterized by high-suspended sediment. *Int. J. Remote Sens.* **2006**, *27*, 1521–1537. [[CrossRef](#)]
68. Pahlevan, N.; Smith, B.; Schalles, J.; Binding, C.; Cao, Z.; Ma, R.; Alikas, K.; Kangro, K.; Gurlin, D.; Hà, N.; et al. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine learning approach. *Remote Sens. Environ.* **2020**, *240*, 111604. [[CrossRef](#)]
69. Vanhellemont, Q. Adaptation of the dark spectrum fitting atmospheric correction for aquatic applications of the Landsat and Sentinel-2 archives. *Remote Sens. Environ.* **2019**, *225*, 175–192. [[CrossRef](#)]
70. Maciel, F.P.; Pedocchi, F. Evaluation of ACOLITE atmospheric correction methods for Landsat-8 and Sentinel-2 in the Río de la Plata turbid coastal waters. *Int. J. Remote Sens.* **2022**, *43*, 215–240. [[CrossRef](#)]
71. Pahlevan, N.; Mangin, A.; Balasubramanian, S.; Smith, B.; Alikas, K.; Arai, K.; Barbosa, C.; Bélanger, S.; Binding, C.; Bresciani, M.; et al. ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters. *Remote Sens. Environ.* **2021**, *258*, 112366. [[CrossRef](#)]
72. Zolfaghari, K.; Pahlevan, N.; Simis, S.G.; O'Shea, R.E.; Duguay, C.R. Sensitivity of remotely sensed pigment concentration via Mixture Density Networks (MDNs) to uncertainties from atmospheric correction. *J. Great Lakes Res.* **2022**, *49*, 341–356. [[CrossRef](#)]
73. de Keukelaere, L.; Sterckx, S.; Adriaensen, S.; Knaeps, E.; Reusen, I.; Giardino, C.; Bresciani, M.; Hunter, P.; Neil, C.; van der Zande, D.; et al. Atmospheric correction of Landsat-8/OLI and Sentinel-2/MSI data using iCOR algorithm: Validation for coastal and inland waters. *Eur. J. Remote Sens.* **2018**, *51*, 525–542. [[CrossRef](#)]
74. Warren, M.A.; Simis, S.G.H.; Selmes, N. Complementary water quality observations from high and medium resolution Sentinel sensors by aligning chlorophyll-a and turbidity algorithms. *Remote Sens. Environ.* **2021**, *265*, 112651. [[CrossRef](#)]
75. Rangel-Peraza, J.G.; De Anda, J.; González-Farías, F.; Erickson, D. Statistical assessment of water quality seasonality in large tropical reservoirs. *Lakes Reserv. Res. Manag.* **2009**, *14*, 315–323. [[CrossRef](#)]

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