Assessment of Regression Models for Surface Water Quality Modeling via Remote Sensing of a Water Body in the Mexican Highlands

Alejandro Cruz-Retana [1,2,*], Rocio Becerril-Piña [3], Carlos Roberto Fonseca [2,*], Miguel A. Gómez-Albores [2], Sandra Gaytán-Aguilar [4], Marivel Hernández-Téllez [2] and Carlos Alberto Mastachi-Loza [2]

1 Tecnológico de Estudios Superiores de Jocotitlán (TESJo), Jocotitlan 50700, Mexico
2 Instituto Interamericano de Tecnología y Ciencias del Agua (IITCA), Toluca 50295, Mexico; magomeza@uaemex.mx (M.A.G.-A.); mhernandezt@uaemex.mx (M.H.-T.); camastachil@uaemex.mx (C.A.M.-L.)
3 Red Lerma- IITCA, Toluca 50295, Mexico; rbecerrilp@uaemex.mx
4 Deltares, 5058 Delft, The Netherlands; sandra.gaytan@deltares.nl
* Correspondence: alejandro.cruz@tesjo.edu.mx (A.C.-R.); crfonsecao@uaemex.mx (C.R.F.)

Abstract: Remote sensing plays a crucial role in modeling surface water quality parameters (WQPs), which aids spatial and temporal variation assessment. However, existing models are often developed independently, leading to uncertainty regarding their applicability. This study focused on two primary objectives. First, it aimed to evaluate different models for chemical oxygen demand (COD), total phosphorus (TP), total nitrogen (TN), and total suspended solids (TSS) in a surface water body, the J. A. Alzate dam, in the Mexican highland region ($R^2 \geq 0.78$ and RMSE ≤ 16.1 mg/L). The models were estimated using multivariate regressions, with a focus on identifying dilution and dragging effects in inter-annual flow rate estimations, including runoff from precipitation and municipal discharges. Second, the study sought to analyze the potential scope of application for these models in other water bodies by comparing mean WQP values. Several models exhibited similarities, with minimal differences in mean values (ranging from -9.5 to 0.57 mg/L) for TSS, TN, and TP. These findings suggest that certain water bodies may be compatible enough to warrant the exploration of joint modeling in future research endeavors. By addressing these objectives, this research contributes to a better understanding of the suitability of remote sensing-based models for characterizing surface water quality, both within specific locations and across different water bodies.

Keywords: water quality parameter; regression model's scope; remote sensing; surface water

1. Introduction

Effective water resource management requires monitoring supply sources' water quality. Alterations in water resources' quality limit their intended functions, which include providing safe drinking water, recreation, ecosystems services, irrigation, and regional planning, among other vital uses [1]. Conventionally, surface water resources are monitored by collecting water samples and analyzing them in a laboratory, which is reliable, but time-consuming and costly [2–4]. Furthermore, this method has limitations regarding spatio-temporal variation identification [5]. The rapid development of remote sensing techniques has enabled the inference of water quality over time and space, even from inaccessible sampling sites [6]. Remote sensing for water quality monitoring is regarded as a promising and cost-effective monitoring tool that offers new opportunities to assess water quality, particularly in developing countries. Valuable reviews for consideration include those conducted by [5].

Numerous studies have elucidated the relationship between reflectance data acquired from remote sensors and water quality parameters (WQPs). Reflections are affected by

change in one or more of these parameters; this dynamic has allowed the development of various WQP estimation analysis methods. Empirical approaches have been developed using statistical regression techniques that establish a connection between in situ WQP measurements and the reflectance of the optimal band or a combination of bands. Their greatest advantages are simplicity, operability, and that their precision can be improved by selecting spectral bands. However, these models' applicability might be limited to the region and temporality for which they were originally developed, and their precision is based on the size and representativeness of field samples [7].

Some WQPs play a strategic role in decision-making processes regarding integrated water management. Two critical parameters are chemical oxygen demand (COD) and total suspended solids (TSS). These parameters offer valuable insights into water quality in a given environment, which aids the identification of potential pollution sources and overall aquatic ecosystem health. They are essential for evaluating the environmental impact of activities such as industrial discharges, agricultural runoff, and urban development.

In addition to COD and TSS, total nitrogen (TN) and total phosphorus (TP) are also significant WQPs that provide valuable information about the environmental and economic dynamics associated with a water body. Elevated levels of these nutrients can lead to eutrophication, resulting in harmful algal blooms and oxygen depletion. Monitoring TN and TP concentrations is crucial for implementing strategies to prevent or mitigate eutrophication. Many regions have established regulations and standards governing permissible concentrations of COD, TSS, TN, and TP in wastewater discharges and natural water bodies. Consequently, water treatment plants rely on data related to COD, TSS, and nutrient levels to optimize treatment processes, ensure pollutant removal, and produce safe drinking water [8].

Furthermore, these parameters' concentrations are closely linked to optical mechanisms that facilitate estimation. In the context of optical measurements, absorption and scattering are the primary mechanisms associated with TSS, while absorption is a dominant mechanism associated with COD, TN, and TP. Refraction plays a minor role in these measurements, primarily when light interacts with particles in the water.

Measuring concentrations relies on the absorption of specific wavelengths. For TSS measurement, the 600–700 nm wavelength range is particularly important, due to its influence on reflectance [9–11]. Reflectivity increases and spectral signature variations depend on particle size and solid properties. As particle size decreases, reflectivity increases [12,13]. However, if particles are of an organic origin, reflectivity is also influenced by their chromatic characteristics [14]. Reflectance in the 600–700 nm wavelength range has been linearly related to TSS concentrations ranging from 0 to 50 mg/L [15–17]. Beyond this range, the relationship may exhibit a curvilinear trend. In the context of optical sensors such as the Landsat 8 OLI, green (B3) and blue (B2) spectrum bands can be as effective as the red (B4) band for TSS estimation, especially when combined with a near-infrared band (B5 or B7) [18].

For COD, the absorption of energy occurs in wavelength bands from B1 (472 nm) to B4 (670 nm) [18,19]. For TP, absorption is prominent in bands B5 (880 nm) and B4 (670 nm). TN mainly absorbs energy at wavelengths close to 472 and 670 nm. TP can often be found at the water's surface (0.0–0.60 m) and can be represented by wavelengths within the visible spectrum, particularly in Landsat 8 OLI bands B2 and B3 [3].

Machine learning has the ability to process large amounts of information in non-linear frameworks [4], such as decision trees (DTs), support vector machines (SVMs), artificial neural networks (ANNs) [5], and genetic algorithms (GAs). However, the initial training phase can be a costly and time-consuming process that is difficult to apply if sufficient data are unavailable; therefore, non-linear models have been developed to a lesser extent [20–26].

Some bio-optical characterization studies [20–22,27–29] have reported satisfactory results for parameters such as chlorophyll-a, total suspended solids, and turbidity; multiple regression models have been the most used approach [30–33], while non-linear models have been developed to a lesser extent [30,34–37].

Linear regression analysis has allowed simplicity in the models' explanatory capacity. Some models consulted for this study (Table 1) had explanatory capacities in the range $0.69 \leq R^2 \leq 0.98$ of the TSS; most were above $R^2 \geq 0.69$. To attain these explanatory capacities, different studies' authors used different sampling universes. For example, other studies, such as [38], used a density of approximately 0.32 samples/km^2, whereas [39] used a density of 0.000216 samples/km^2. The study with the highest sampling density corresponded to 0.64 samples/km^2 for TSS and COD [17]. In [40], there were 14 samples (0.001 samples/km^2), and [3] used 18 samples (0.32 fields/km^2). Previous studies noted that it is advisable to collect water samples at an average depth of 0.1 m [27,35,37,41–43].

In the studies consulted, it was possible to highlight the range of concentration models estimated. For a coastal water body [22], a TSS range between 0 and 135 mg/L was observed; ranges between 0 and 386 mg/L were observed in continental water bodies [44,45]. For TN and TP, 0–26 mg/L concentrations were observed in intervals [3]; these WQPs are generally present in 0–30 mg/L intervals in rivers, dams, and lakes. Some studies [18,46] reported 0–86 mg/L COD concentrations, highlighting similar environmental, urban, and agricultural characteristics.

Regression models used to estimate water quality parameters (WQPs) were calibrated considering the specific environmental conditions of each water body and each sensor's characteristics [31,34–36]. Thus, when these models were applied outside their intended contexts, they tended to yield highly uncertain estimates [12]. Regarding environmental conditions, uncertainties stem from factors such as cloud cover, geographical coordinates (longitude, latitude, and altitude), atmospheric conditions, and concentrations of algae within the water body [4]. Regarding sensor characteristics, limitations often pertain to the platform type (height, inclination, etc.), as well as spatial, radiometric, and temporal resolutions [4,10,20,47]. Nevertheless, despite these differences, one can observe a degree of similarity in these models' structures, such regression models' band weights and the choice of linear or non-linear regression techniques [4,10,20,48].

Considering the above factors, there might be some compatibility among these models when applied to water bodies different from those for which they were initially designed. Confirming such compatibility could facilitate the development of integrated models. As a result, our study focused on two primary objectives. First, we aimed to assess various models for estimating chemical oxygen demand (COD), total phosphorus (TP), total nitrogen (TN), and total suspended solids (TSS) in a surface water body, the J. A. Alzate Dam, in the Mexican highland region. These models were derived using multivariate regression techniques; however, in addition to results from similar previous studies, we intended to identify dilution and dragging effects in inter-annual flow rate estimations, which included runoff from precipitation and municipal discharges.

Secondly, our study sought to explore these models' potential applicability in other water bodies by comparing mean WQP values. Importantly, this comparison was not intended to determine which model was superior, but rather to provide evidence of their compatibility. This comparison considered models (from [3,46,49,50]) developed in conditions as similar as possible with respect to sensor types, resolution, available information, and the choice of regression techniques.

In the context of remote-sensing water quality studies in Mexico, some available research focuses on chemical oxygen demand (COD), total dissolved solids, chlorophyll-a, total suspended solids (TSS), and temperature [36,42,44]. However, comprehensive data regarding reflectance and WQPs selected for this study were limited for replication and thorough analysis. Furthermore, in this investigation, a select number of viable water bodies were identified as similar for comparison purposes using the obtained models. Based on these findings, the studies listed in Table 1 were chosen to assess the applicability of the model derived from this study.

Table 1. Primary contributions of different authors to WQP estimation using Landsat 8 OLI.

WQP	Surface (km²)	Resolution (m)	Sample of Size	Bands of Normalization	R^2	Estimation Interval (mg/L)	Author
TSS	12,000	30 m	26	$\hat{Y} = -161.98\left(\frac{B5}{B4}\right)^3 + 713.478\left(\frac{B5}{B4}\right)^2 - 811.43\left(\frac{B5}{B4}\right) + 278.46$	0.98	0–386	[45]
		30 m	14	$\hat{Y} = (1.5212\left(\frac{LOG(B2)}{LOG(B3)}\right) - 0.3698)$	0.69	0–135	[22]
TN	53	30 m	18	$\hat{Y} = e^{(8.228 - 2.713*(In(\frac{B3}{B2})))}$	--------	0–36	[3]
TP	53	30 m	18	$\hat{Y} = e^{(-0.4081 - 8.659(In(\frac{B3}{B2})))}$	--------	0–26	[3]
COD	150	30 m	-------	$\hat{Y} = 2.76 - 17.27B1 + 72.15B2 - 12.11B3$	--------	0–19.3	[19]

Symbology: Chemical oxygen demand (COD), total phosphorus (TP), total nitrogen (TN), total suspended solids (TSS), remote sensing bands (Bn), adjusted coefficient of determination ($\overline{R}^2$), and no data (---).

2. Materials and Methods

In the study area, a WQP estimation using multispectral images was carried out through the development of regression models; the spatio-temporal distribution of estimated WQPs; and the assessment of the scope of its application relative to other water bodies. The WQPs considered were total suspended solids (TSS), total nitrogen (TN), chemical oxygen demand (COD), and total phosphorus (TP). The study area was the J. A. Alzate Dam, in Toluca, Mexico (Figure 1). Imagery from the Landsat 8 OLI sensor (Path 26 Raw 46) [47] was obtained from the United States Geological Survey database and coincided with dates of water quality monitoring campaigns. Notably, the Landsat 8 sensor was chosen for its appropriate spatial resolution (approximate 51,000 m² water body surface), and for the assessment of the model's scope in similar studies.
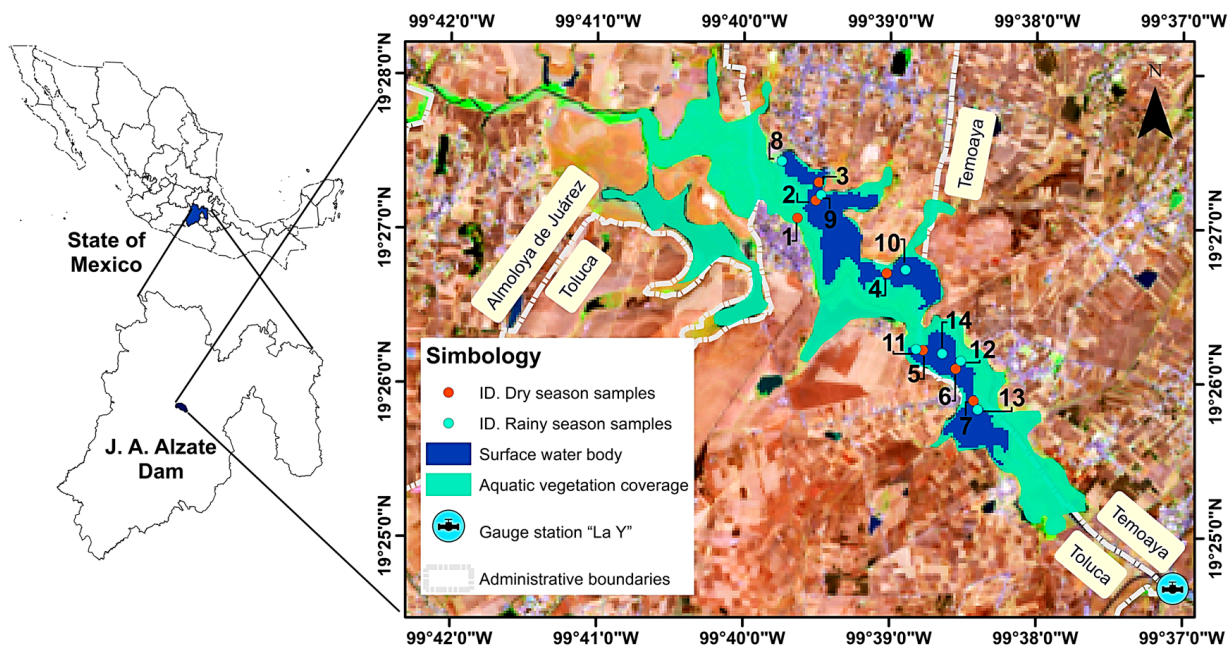

Figure 1. Delimitation of the study area and georeferencing of sampling for water quality parameters.

To develop regression models, measuring WQP in the laboratory from representative samples (14 sites) and pretreating satellite images were necessary. The WQP concentrations were obtained following standards for TN [51], TSS [52], TP [53], and COD [54,55]. According to some authors [37,48], samples should be collected under different seasonal conditions, with caution and appropriate validation; therefore, samples were geograph-

ically distributed, as indicated in Figure 1, and divided into two seasons. This research involved two distinct sampling campaigns for two main reasons: first, to ensure the field data robustness throughout the year; and second, to account for the varying sample densities observed in previous studies [3,19,22,45,56], which ranged from 0.0002 to 0.64 samples per square kilometer. Interestingly, these sample density variations did not appear to significantly affect results.

To ensure a representative sample for the finite population under study, we determined that a resolution of 0.3 samples per square kilometer was sufficient for the J. A. Alzate Dam. Consequently, we organized the sampling process into two campaigns, which were conducted both before and after the rainy season (19 May 2018 and 16 October 2018, respectively), and classified based on the standard Mexican normative [55].

The MODTRAN 4 module was used to pretreat satellite images [27,57], as it could consider study area characteristics such as altitude, latitude, proximity to the sea, aerosol type, atmosphere type, and image visibility in atmospheric correction. To accurately identify pixels in each satellite image corresponding to the water body, we employed the normalized difference water index (NDWI). The NDWI is a widely used remote sensing index for detecting the presence of water in satellite or aerial imagery [19,32,57]. Positive NDWI values generally indicate the presence of water; however, its specific water detection threshold can vary depending on the dataset and environmental conditions. After analyzing NDWI values in the J. A. Alzate Dam area, we observed a mean 0.65 NDWI value with a 0.17 standard deviation. Notably, some pixels situated along the water body's shoreline exhibited values below this standard deviation. Based on this observation, we established a reliable range for water detection (0.12 < NDWI < 1.0) within a 0.17 standard deviation in the reflectance for each multispectral band. Utilizing the NDWI not only aided in the delineation of pixels constituting the water body, but also restricted the occurrence of anomalous data, particularly with regard to TSS.

2.1. Statistical Analysis for Model Development

The first stage (development of regression models) was cross-validation, which has been commonly used in similar studies. For example, in [4] and [57] the authors considered 11 random combinations to model chlorophyll (Chl-a) and suspended particulate matters (SPM), with mean RMSEs of 1.6–1.7 mg/m^3 and 8.8 to 11.4 g/m^3, respectively. In [40], cross-validation was applied in a random forest-type study with a RMSE = 0.02–3.03 (mg/L), using 10 iterations that reached an explanatory capacity of R^2 = 95–99%. The developed regression models were evaluated during the cross-validation process, which included multiple iterations (i) between the testing and validation subsets. The primary metric used for evaluation was the root mean square error $\left(RMSE = \sqrt{\frac{1}{i}\sum_{i=1}^{i}(yi - \hat{y}i)^2} \right)$ calculated as the average of discrepancies between predicted values and actual observations. Additionally, we considered the adjusted coefficient of determination $\left(\overline{R}^2 = 1 - \left(\frac{n-1}{n-k-1} \right) * \left(1 - R^2 \right) \right)$ to assess the models' explanatory power, taking the number of multispectral bands employed into account. The coefficients' collinearity and heteroscedasticity were also considered as selection criteria.

A recommended range for the adjusted R^2 is between 0.6 and 0.8, which is considered suitable for estimating water quality parameters (WQPs), as per the reference. This range ensures that errors exhibit appropriate behavior. When $R^2 \geq 0.9$, estimates are regarded as both statistically significant and well fitting in relation to established values [58]. The reference's authors also analyzed the multiple linear regression from which they obtained a RMSE = 0.03–3.14 (mg/L) using 10 iterations, reaching an explanatory capacity of R^2 = 55–91%. The multivariate regression models proposed in the present study correspond to linear types ($\hat{y} = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_3 + \ldots + \beta_n * x_n + u$), such as those presented in [35] and [56]), exponential types ($\hat{y} = e^{(\beta_0 + \beta_1 * x_1 + \beta_2 * x_2 \ldots \beta_n * x_n)}$ such as those studied in [3,13,19,35]), and polynomials with the structure $\hat{y} = \beta_0 + \beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_1^2 + \beta_4 * x_2^2 + \beta_5 * x_1 x_2 + \ldots + \beta_n * x_n^n$ [36,45,48,59]. A SIG environment,

TerrSet® [59], tools such as IBM SPSS Statistics® [60,61], and add-ins for Excel®, including RISK Simulator, Analyse-It, and XLSTAT [59], were used. In cases of input value validation (reflectance and TSS sampling), the first procedure proposed in [62], it was verified that input values complied with homoscedasticity, adequate micro numerosity, omitted outliers (depending on the model), non-linearity (if applicable), normality, and the absence of multicollinearity. Wavelengths that depict regressive variables were selected individually for each WQP. For TSS, B1, B4, and B6 have shown acceptable behaviors both linearly and non-linearly [24,31,40]. For TN and TP, some authors used B1 and B3 from the Landsat and Sentinel sensors (non-linearly) because B3 captures chlorophyll-a pigment associated with aquatic vegetation [3]. COD has been linearly studied based on bands 1, 2, and 3 [46], and in a non-linear way based on bands 2, 6, and 7 [18].

2.2. Spatio-Temporal Distribution of Estimated WQPs

Once the model with the best fit was identified, its model equation could be applied to classify a water body according to intervals established by applicable regulations (in this case, Mexican regulations) [55]. This is how regions were evaluated; intervals were evaluated for agricultural use (whose permissible limits were TSS = 60–75 mg/L, TN = 40–60 mg/L, and COD = 60–75 mg/L).

In addition to representing the information's spatial distribution, it was possible to analyze the WQPs' temporal behaviors relative to the influent water. This study estimated municipal water discharges (depending on population size), as well as inflow (measured by [63]) to infer causal relationships for the WQP concentrations' variation throughout the year.

Furthermore, to evaluate the possibility of implementing a regression model in different water bodies, at this stage, the difference of means was analyzed based on regression models' estimates for other bodies with conditions similar to those of the J. A. Alzate Dam (sensor and model type). In this way, regression models whose mean difference ($\mu_1 - \mu_2$) were close to zero, or a tolerable under- or over-estimation, were considered as similar, assuming a t-distribution (Equation (1)). The studies considered as references were [3,19,45].

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 <= (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}\,, \tag{1}$$

where $\bar{x}_1$ is the study's sample mean using the authors' equation [mg/L]; $\bar{x}_2$ is the study's sample mean using the present study's equation [mg/L]; s_1^2 corresponds to the variance in the authors' data [mg/L]; s_2^2 is the variance in this study's data [mg/L]; and n_1 and n_2 are sample sizes.

3. Results and Discussion

To develop multivariate regression models (stage 1), each model type's reflectances from the considered bands and each sample's WQP concentrations were evaluated. Although it was difficult to observe a clear spatial trend, Table 2 shows that, in general, WQPs before the rainy season had, on average, 59.3% higher concentrations than after the rainy season, assuming a preliminary dilution effect.

In contrast, Table 3 shows the results of various statistical tests regarding the quality treatment of input data. In general terms, the data showed a normal trend, homoscedasticity (i.e., W-$_{pvalue}$ > 0.1 and χ^2 < V$_{crit}$), no globally identified outliers (≤1 sample), and no collinearity of sampling results (r ≤ 0.75 and VIF ≤ 4.0); most models did not present multi-collinearity (F > V$_{crit}$). Finally, the *p*-value was significant (P$_{value}$ < V$_{crit}$) for variables in exponential models for COD, TP, and TSS; in linear models for TN and TP; and in polynomial models for TP.

Table 2. Sampling in the J. A. Alzate Dam, Mexico.

ID	X (W Longitude)	Y (N Latitude)	Season	TN (mg/L) PL = 40–60	COD (mg/L) PL = 60–75	TP (mg/L) PL = 20–30	TSS (mg/L) PL = 12–75
				Chemistry Parameters			**Physics p.**
1	−99.660564	19.451172		33	173	99.76	92.5
2	−99.658486	19.453087		11	127	46.27	151.5
3	−99.658131	19.455031	Before the rain	16	66	44.34	198
4	−99.650393	19.445229	season	9	93	34.25	48
5	−99.646237	19.436919	18th May	28	107	70.91	33
6	−99.642504	19.434917		20	98	67.34	55
7	−99.640462	19.431472		10	109	74.92	42
8	−99.662355	19.457317		7.2	67.5	33.16	35
9	−99.657869	19.453669		6.3	64	34.46	34
10	−99.648234	19.445605	After the rain	3.9	80	23.64	15.2
11	−99.644043	19.436555	season	4.4	21	30.76	11
12	−99.641932	19.435785	26th October	3.5	30	32.79	16
13	−99.639985	19.430514		2.8	26	30.39	17
14	−99.647021	19.437027		3.3	39.5	33.71	19

Symbology: sample fields (ID), Total phosphorus (TP), total nitrogen (TN), chemical oxygen demand (COD), total suspended solids (TSS), permissible limits (PL).

Table 3. Validation assumptions in the input data.

	WQP	**Total Nitrogen (TN)**						**Chemistry Oxygen Demand (COD)**					
	Model Type	Exponential		Linear		Polynomial		Exponential		Linear		Polynomial	
Statistical Test	Independent Variables	B1, B3		B3, B6, (B3 + B7)		(B1/B3)		LN(B5), LN(B2/B3), LN(B7)		(B1/B6), (B7/B4), (B2/B1), (B2/B3)		(B3/B1), (B3/B5)	
		V_{crit}	V	V_{crit}	V	V_{crit}	V	V_{crit}	V	V_{crit}	V	V_{crit}	V
Homoscedasticity	$W_{\text{-pvalue}} > V_{crit}$	0.1	0.25	0.1	0.33	0.1	0.23	0.1	0.18	0.1	0.25	0.1	0.38
Square chi test	$\chi^2 < V_{crit}$	7.81	4.54	7.81	4.65	11.07	9.30	7.81	2.25	9.48	2.26	11.07	0.82
Atypical values	$V_{calc} \leq V_{crit}$	2	1	2	1	2	1	2	1	2	0	2	1
Collinearity	$VIF < V_{crit}$	4	0.21	4	0.16	4	4.0	4	0.21	4	2.6	4	1.17
	$r \leq V_{crit}$	0.75	0.74	0.75	0.75	0.75	0.86	0.75	0.74	0.75	0.78	0.75	0.78
Multicollinearity	$F > V_{crit}$	3.70	22.90	3.70	13.20	3.68	6.25	3.70	19.08	3.63	6.75	3.68	11.82
Normality	$D < V_{crit}$	0.22	0.12	0.22	0.17	0.17	0.22	0.22	0.19	0.22	0.08	0.22	0.16
Significance	$P_{value} \leq V_{crit}$	0.05	0.19	0.05	0.05	0.05	0.06	0.05	0.00	0.05	0.09	0.05	0.79
	WQP	**Total Phosphorus (TP)**						**Total Suspended Solids (TSS)**					
	Model Type	Exponential		Linear		Polynomial		Exponential		Linear		Polynomial	
Statistical Test	Independent Variables	LN(B5), LN(B6)		(B5/B4), (B5/B6)		(B5/B2), (B5/B2)		LN(B4), LN(B3+B5), LN(B5+B7), LN(B2/B3), LN(B2/B4)		(B7/B4), (B7/B5)		(B3/B5), B3	
		V_{crit}	V	V_{crit}	V	V_{crit}	V	V_{crit}	V	V_{crit}	V	V_{crit}	V
Homoscedasticity	$W_{\text{-pvalue}} > V_{crit}$	0.1	0.18	0.1	0.64	0.1	0.10	0.1	0.11	0.1	0.11	0.1	0.12
Square chi test	$\chi^2 < V_{crit}$	5.99	4.87	5.99	2.62	11.07	5.61	11.07	5.93	3.32	5.99	11.07	1.90
Atypical values	$V_{calc} \leq V_{crit}$	2	1	2	0	2	1	2	1	2	0	2	1
Collinearity	$FIV < V_{crit}$	4	2.63	4	6.31	4	2.88	4	4.1	4	1.28	4	1.38
	$r \leq V_{crit}$	0.75	0.78	0.75	0.9	0.75	0.80	0.75	0.87	0.75	0.46	0.75	0.52
Multicollinearity	$F > V_{crit}$	3.98	18.65	3.98	24.3	3.68	25.63	3.68	30.60	3.98	11.28	3.68	136.12
Normality	$D < V_{crit}$	0.22	0.16	0.23	0.22	0.22	0.12	0.22	0.15	0.22	0.20	0.22	0.15
Significance	$P_{value} \leq V_{crit}$	0.05	0.00	0.05	0.003	0.05	0.00	0.05	0.04	0.05	0.07	0.05	0.09

Symbology: Water quality parameter (WQP), significance for White test ($W_{\text{-}p\text{value}}$), variance inflation factor (VIF), calculated value (V), critical value (V_{crit}), and significance (P_{value}).

The development of the current regression models used various statistical criteria, including the adjusted determination coefficient, mean square error, *p*-values of coefficients,

and confidence intervals. Additionally, techniques such as cross-validation were employed to mitigate the risk of overfitting, while residual analysis was used to assess model assumptions and check for heteroscedasticity. The cross-validation results are presented in Table 4 and provide insights into each regression model's coefficients.

Table 4. Multiple regression models to different WQPs.

WQP	Type	Regression Model	i	RMSE	R^2
TN (mg/L)	Exp.	$e^{(-4.49-2.47LNB1+1.10LNB5-0.69LNB7)}$	5	3.82	0.79
	Linear	$30.71 + 1120.39B3 + 823.79B6 - 1269.80(B3+B7)$	7	4.24	0.73
	Pol.	$8.4 - 42.3\left(\frac{B7}{B1}\right) + 21.6\left(\frac{B5}{B4}\right) + 109.3\left(\frac{B7}{B1}\right)^2 + 34.5\left(\frac{B5}{B4}\right)^2 - 123\left(\frac{B7}{B1}\right)\left(\frac{B5}{B4}\right)$	14	31.13	0.68
COD (mg/L)	Exp.	$e^{(4.2069+0.9788234LNB5-2.4215LN(B2/B3)-0.5209LNB7)}$	6	16.1	0.80
	Linear	$-66.73 + 5.66\left(\frac{B1}{B6}\right) + 110.16\left(\frac{B7}{B4}\right) + 240.1259\left(\frac{B2}{B1}\right) - 222.73\left(\frac{B2}{B3}\right)$	5	21.4	0.62
	Pol.	$-129.04 + 612.5\left(\frac{B3}{B1}\right) - 396.95\left(\frac{B3}{B5}\right) - 266.95\left(\frac{B3}{B1}\right)^2 + 42.51\left(\frac{B3}{B5}\right)^2 + 173.42\left(\frac{B3}{B1}\right)\left(\frac{B3}{B5}\right)$	9	15.35	0.84
TP (mg/L)	Exp.	$e^{(5.1265551+1.154335LNB5-0.52206356LNB6)}$	8	10.25	0.74
	Linear	$LogTP = 1.3544 + 0.1240\left(\frac{B5}{B4}\right) + 0.04610\left(\frac{B5}{B6}\right)$	5	9.63	0.79
	Pol.	$14.79 + 62.96\left(\frac{B5}{B2}\right) - 1644.8B6 + 41.86\left(\frac{B5}{B2}\right)^2 + 71481.28B6^2 - 3642.34\left(\frac{B5}{B2}\right)B6$	5	5.49	0.92
TSS (mg/L)	Exp.	$e^{(-0.2663-4.05LNB4+7.49LN(B3+B5)-4.03LN(B5+B7)+3.99LN(B2/B3)+1.4LN(B2/B4))}$	5	14.37	0.90
	Linear	$LogTSS = 2.0848 + 0.339\left(\frac{B7}{B4}\right) - 1.316\left(\frac{B7}{B5}\right)$	5	43.9	0.61
	Pol.	$-72.79 + 1695.66B3 + 36.58\left(\frac{B3}{B5}\right) + 51578.99B3^2 + 143.87\left(\frac{B3}{B5}\right)^2 - 6246.61\left((B3)\left(\frac{B3}{B5}\right)\right)$	6	7.19	0.98

Symbology: total suspended solids (TSS), chemistry oxygen demand (COD), total phosphorus (TP), total nitrogen (TN), wavelength band (B_i), adjusted determination coefficient (R^2), number of iterations (i), root mean square error (RMSE), best model (---), and other models (---).

In the context of TSS, both linear and non-linear models indicated that the B3 and B5 wavelengths had the most substantial influence; this aligns with findings from previous studies [48,50]. In contrast to the model presented in [3], the exponential model for TN was primarily influenced by B1. Although the polynomial model exhibited a higher determination coefficient and a lower RMSE, it raised concerns related to the significance and collinearity of certain coefficients.

For TP, the ratio B5/B2 and B6 exerted significant influence and provided acceptable estimates within the 18 to 98 mg/L concentration range in the polynomial model, whereas in the model used in [3], this parameter was largely dependent on B2.

COD was primarily influenced by the B2/B3 relationship in the exponential model, consistent with findings in [46] and [63]. The exponential model demonstrated a superior explanatory capacity for TN and COD (as shown in Figure 2a,b) based on the determination coefficient and RMSE. However, the model might have slightly overestimated COD concentrations exceeding 150 mg/L.

All models (as depicted in Figure 2c,d) exhibited a good fit with determination coefficients greater than 0.92, particularly in the polynomial configuration (Table 4); notably, both linear and exponential models tended to underestimate concentrations in these cases.

The application of regression models (stage 2) was carried out according to a classification based on standard Mexican concentrations [51–55] permitted for discharges into rivers, streams, and drains (C1); reservoirs, lakes, and lagoons (C2); and outside permissible limits (C3). It was observed that before the rainy season, 92% of the surface (1.32 km^2) was below the permissible C1 limit for TN discharge (Figure 3a). After the rainy season, this area increased (especially in the southern zone) to 99.7% (1.42 km^2), inferring a dilution process (Figure 4a). COD presented a similar dilution behavior; in the dry season 73.2%

of the surface was below the permissible C2 limit; after the rainy season, it increased to 99.7%, mainly in the central zone (Figures 3b and 4b). The water body's surface presented 92.2% and 41.0% concentrations out of TP's range (C3) before and after the rainy season, respectively (Figures 3c and 4c). The surface outside permissible TSS limits decreased from 97.01% to 60.7% before and after the rainy season, respectively (Figures 3d and 4d). TSS' and COD's percentages presented a lower sensitivity to rain and can be explained by municipal and industrial discharges along the Lerma River. It should be noted that certain data points in Figures 3 and 4 may appear to be outside the established region's range due to the dots' size relative to the resolution of the value distribution image.
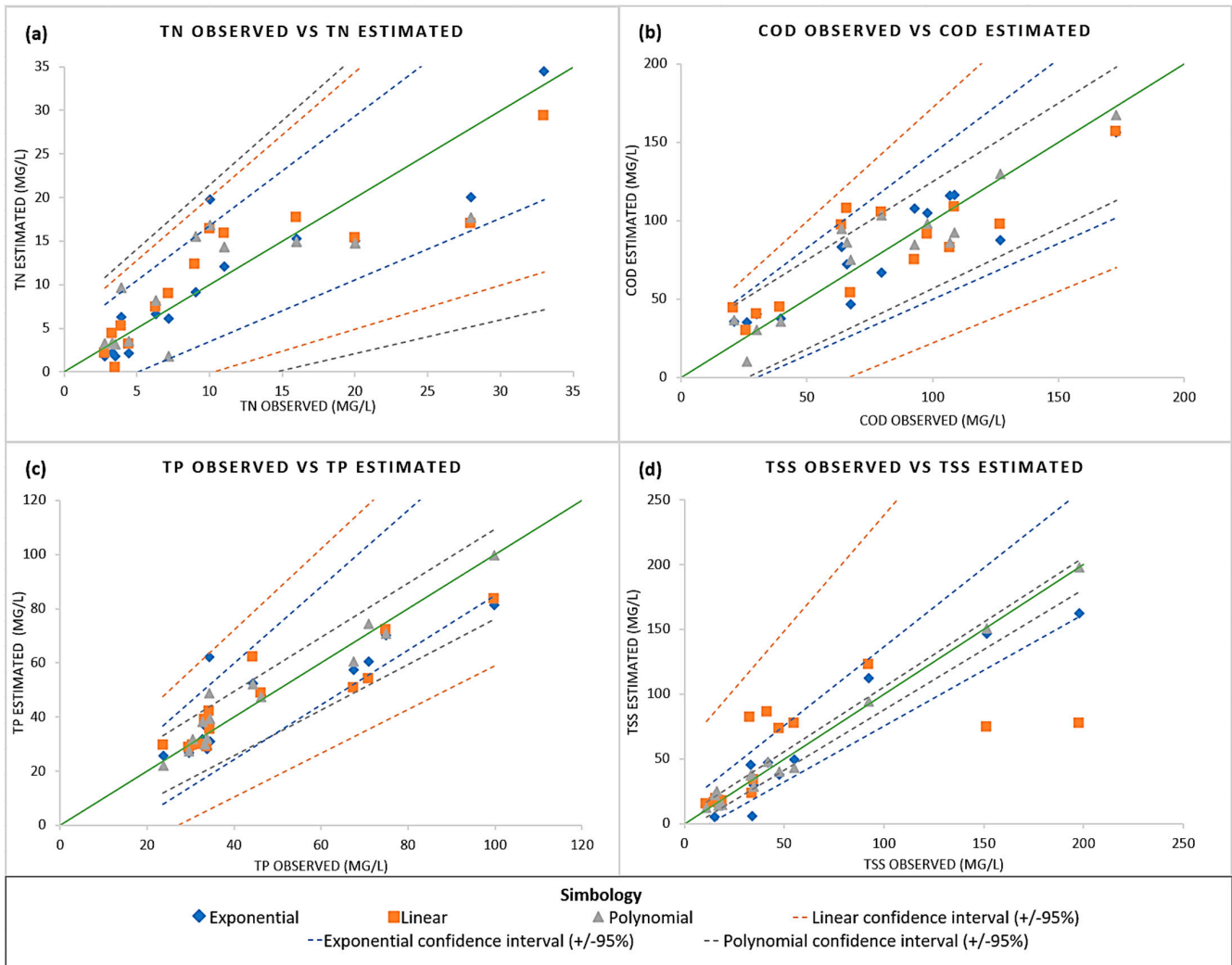
Figure 2. Comparison of WQPs measured in the laboratory vs. WQPs estimated through remote sensing.
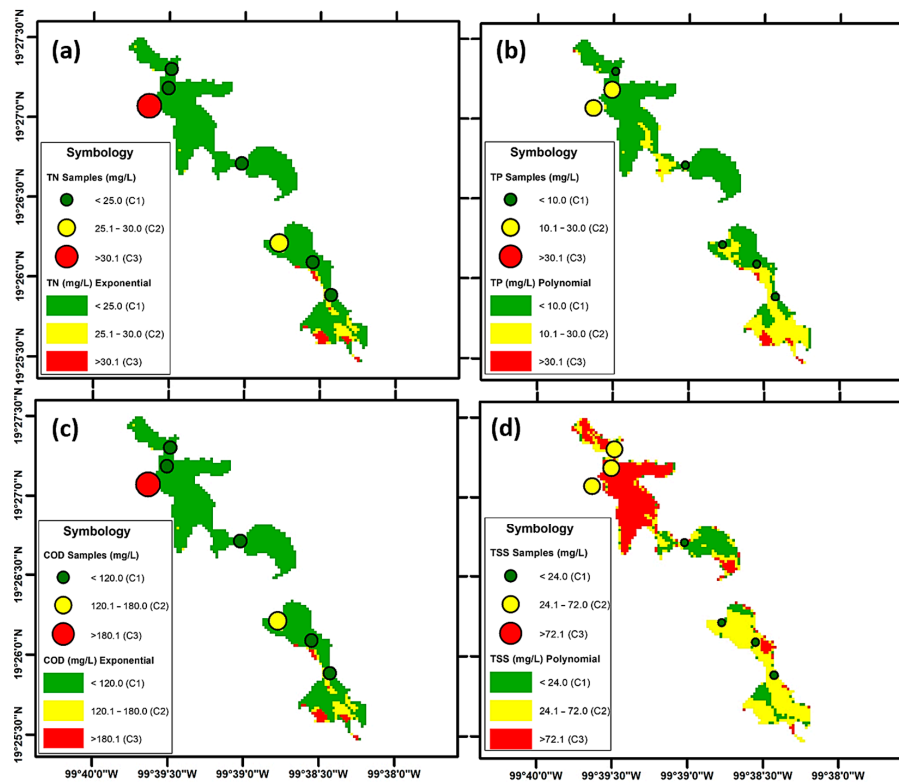
Figure 3. Pre-rainy season multiple regression model maps: (**a**) TN, (**b**) COD, (**c**) TP, and (**d**) TSS (mg/L).
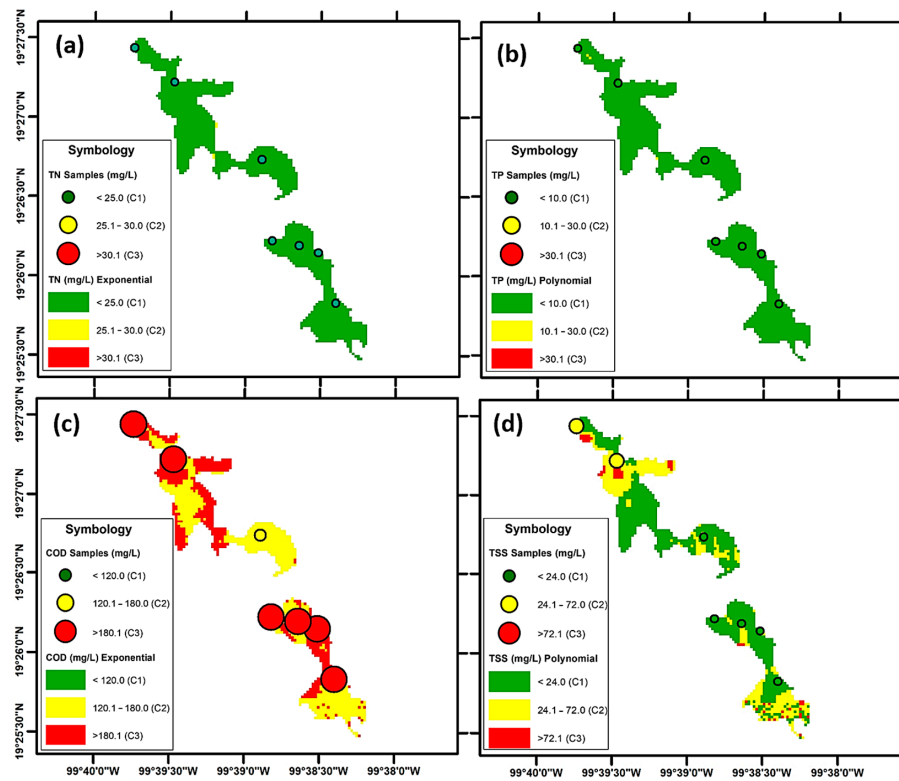
Figure 4. Post-rainy season multiple regression model maps: (**a**) TN, (**b**) COD, (**c**) TP, and (**d**) TSS (mg/L). C1: permissible limit concentration for any discharge; C2: permissible limit for discharge into reservoirs, lakes, and lagoons; C3: out of permissible limit.

In addition to examining the WQPs' spatial distributions, we also investigated their interannual variations by applying the developed models to imagery captured between

2019 and 2021. These variations were influenced by factors such as dilution and dragging, which were identified through changes in inflow and domestic discharge patterns. The analysis considered recorded inflows at gauge station 12,374, known as 'La Y' [48], and municipal wastewater discharges, which were estimated based on demographic data (as shown in Figure 5).
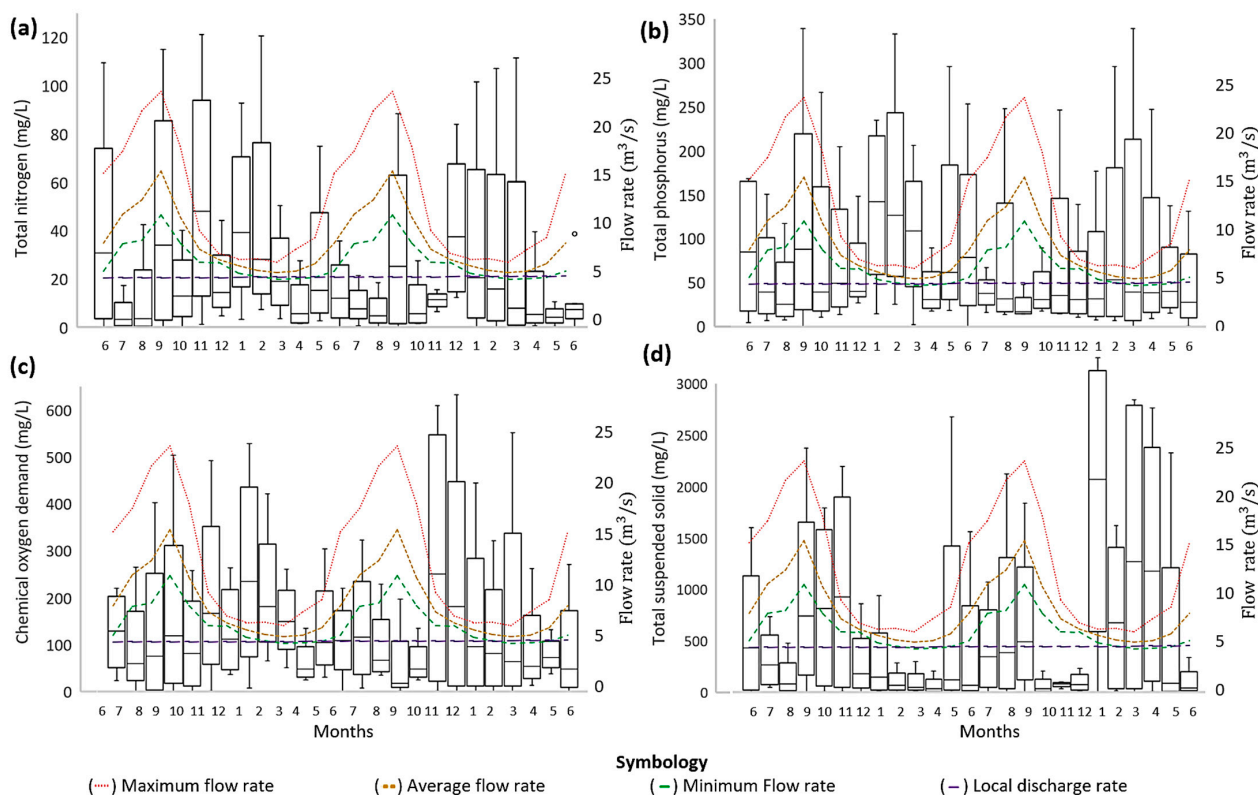
Figure 5. Box plot for time series (June 2019 to June 2021) for concentrations of (**a**) TN, (**b**) COD, (**c**) TP, and (**d**) TSS.

Through WQP estimation using selected models over a two-year period, we were able to elucidate patterns of concentration influenced by dilution and dragging effects. This analysis involved comparing variations in WQP concentrations with inflow variations, which encompassed both runoff and municipal discharges (based on observations from [64]).

Figure 5 illustrates the dynamic relationship between inflow and WQP concentrations. Notably, there was a nearly constant minimum flow of 5 m3/s stemming from municipal discharges. During the rainy season, this flow could increase up to fourfold. In contrast, concentrations of COD, TN, and TP exhibited their highest values during the dry season; on average, TN and COD concentrations doubled, while TP concentrations increased threefold. This behavior is indicative of a dilution effect resulting from increased water volume.

However, TSS concentrations followed a different pattern. As the inflow rate increased, TSS concentrations also increased, indicating a dragging effect. Notably, during the last few months, TSS experienced a significant increase, deviating from this typical trend. In summary, this analysis demonstrates the interplay between inflow dynamics and WQP concentrations. While TN, COD, and TP exhibited dilution patterns, TSS displayed a dragging effect, with exceptions noted during the later months. Understanding these effects is crucial for effective water quality management and environmental monitoring.

In the context of modeling water quality parameters, it is essential to explore the similarities and differences between various models. This discussion focuses on the TSS, TN, and TP models (Figure 6).
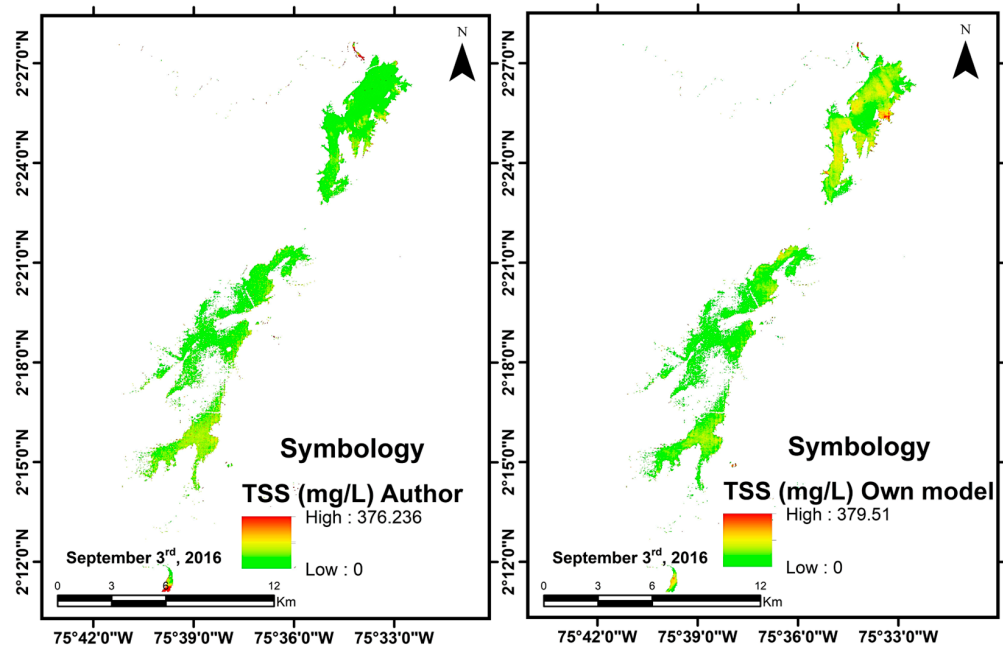
Figure 6. Spatio-temporal distribution of estimated WQPs for author model [45] and current model for TSS.

First, when examining [45]'s TSS model, we noted that it (along with one of the models developed for the J. A. Alzate Dam) employs a non-linear function based on bands 4 and 5. Both models predict TSS concentrations that reach up to 500 mg/L. Comparing these models' estimations to a stratified sample using the same reflectance images, we observed that the authors' TSS model appeared to exhibit more scattering, with a standard deviation 2.5 times higher than the current model's. However, on average, the models indicated the potential for providing similar estimations, as the difference between their means ranged from −9.51 to 0.57 (Table 5). In other words, the current model tended to slightly overestimate TSS concentrations compared to the authors' model. It is important to note that this comparison did not intend to determine which model was superior. Instead, it highlighted that certain models, even when developed under different climatological and geographical conditions, can exhibit close correlations, as demonstrated in Figure 6. The spatial distribution of estimations in this case not only helped identify the location of detected overestimations (northeastern shore) for high TSS values, but also indicated slight underestimations of concentrations ranging from 95 to 110 mg/L.

Table 5. Differences in means between regression models.

WQP	Mean in Samples (mg/L)		Standard Deviation (mg/L)		Confidence Interval for Difference in Means ($\mu_1 - \mu_2$) (mg/L)		Author
	A.M.	**C.M.**	**A.M.**	**C.M.**			
TSS	11.06	15.74	22.79	8.81	−9.51	0.57	[45]
TN	23.71	26.99	1.36	0.94	−3.62	−2.94	[3]
TP	33.21	36.14	8.15	5.32	−4.92	−0.92	[3]

Symbols: Author model (A.M.), current linear model (C.M.), total nitrogen (NT), total phosphorus (TP), total suspended solids (TSS). The images used for the estimation are the same as those used by the author.

Notably, the current TN and TP models tended to overestimate concentrations by 2.8% for TN and 14.8% for TP. However, the dispersion of these estimates appeared to follow a pattern similar to that observed for TSS when considering deviation standards. This suggests that it would be advisable to develop integrative models that take the nature and type of model specific to each water quality parameter (WQP) into account.

In summary, the comparison of these models underscores the need to carefully assess model performance and consider the variability associated with different environmental conditions. By doing so, we can make informed decisions about which models are suitable for particular applications, and whether integrated models should be developed to account for the nuances of each WQP.

4. Conclusions

As noted in similar studies, some regression models were satisfactorily adjusted using cross-validation to estimate (through remote sensing) some water quality parameters, such as total nitrogen (TN), total phosphorus (PT), chemical oxygen demand (COD), and total suspended solids (TSS), in a surface water body, the J. A. Alzate Dam, in the Mexican highlands. The proposed configurations of these models were polynomial, exponential, and linear for each WQP, according to a review of similar previous studies. Several model characteristics were identified. First, the polynomial model provided a better fit for COD than the exponential model did; however, it presented a slower cross-validation convergence. These models' respective terms might be subject to a lower significance and greater collinearity. Second, the exponential model presented a higher determination coefficient for TN; its terms had no problems of significance, homoscedasticity, or collinearity. Third, the polynomial models for TP and TSS presented a better goodness-of-fit, although their collinearity of terms could be present in any TP model. Therefore, given the present research, it is recommended that input data be evaluated under linear and non-linear schemes. Notably, using remote sensing requires the proper identification of pixels corresponding to the water body. In this case, in addition to using the normalized water index, the dispersion of reflectance values (represented by their standard deviations) was relevant, which helped significantly reduce outliers in the models' terms.

In addition to identifying the most appropriate models, it was possible to determine a spatial distribution of WQP concentrations and their temporal behaviors. On the one hand, spatial distribution serves as a tool for the classification of critical regions under some regulations. For example, in this case study, Mexican regulations were used to determine areas that exceeded maximum permissible limits for water discharges. On the other hand, the estimation of WQPs' temporal behavior was useful during the monitoring and retrospective evaluation of the water body. To do this, it was necessary to estimate influents (such as flows) using precipitation and discharge from human settlements. This provided evidence of dilution and drag effects, as observed in this case study. Dilution was associated with a decrease in concentrations of TN, TP, and COD, while the drag effect corresponded to TSS, which presented higher concentrations in the first months after the rain. In this sense, it is important to consider a strategic distribution of sampling throughout the year to enhance a model's scope.

Each model's scope was evaluated with respect to different regression types and its applicability to other water bodies under conditions similar to those in which the model was developed (regression and sensor types). In general, a low relationship was observed between results obtained using the current models and those of other authors. However, in three models differences between their means were identified as acceptable, indicating model compatibility. This supports the approach of studies that developed integrated models according to the characterization or regionalization of surface water bodies, e.g., [65,66].

References

1. Adjovu, G.E.; Stephen, H.; James, D.; Ahmad, S. Measurement of Total Dissolved Solids and Total Suspended Solids in Water Systems: A Review of the Issues, Conventional, and Remote Sensing Techniques. *Remote Sens.* **2023**, *15*, 3534. [CrossRef]
2. Zeinalzadeh, K.; Rezaei, E. Determining spatial and temporal changes in surface water quality using principal component analysis. *J. Hydrol. Reg. Stud.* **2017**, *13*, 1–10. [CrossRef]
3. Zeiny, A.; Kafrawy, S. Assessment of water pollution induced by human activities in Burullus Lake using Landsat 8 operational land imager and GIS. *Egypt. J. Remote Sens. Space Sci.* **2016**, *20*, 549–556.
4. Ouma, Y.; Noor, K.; Herbert, K. Modelling Reservoir Chlorophyll-a, TSS, and Turbidity Using Sentinel-2A MSI and Landsat-8 OLI Satellite Sensors with Empirical Multivariate Regression. *J. Sens.* **2020**, *2020*, 8858408. [CrossRef]
5. Gholizadeh, M.H.; Melesse, A.M.; Deddi, L. A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensor* **2016**, *16*, 1298. [CrossRef]
6. 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]
7. Yang, H.; Kong, J.; Hu, H.; Du, Y.; Gao, M.; Chen, F. A review of remote sensing for water quality retrieval: Progress and challenges. *Remote Sens.* **2022**, *14*, 1770. [CrossRef]
8. Chang, N.B.; Imen, S.; Vannah, B. Remote Sensing for Monitoring Surface Water Quality Status and Ecosystem State in Relation to the Nutrient Cycle: A 40-Year Perspective. *Crit. Rev. Environ. Sci. Technol.* **2015**, *45*, 101–166. [CrossRef]
9. Chang, N.B.; Bai, K.; Imen, S.; Chen, C.F.Y.; Gao, W. Fusión y creación de redes de imágenes satelitales multisensor para el monitoreo ambiental en todo clima. *IEEE Syst. J.* **2018**, *12*, 1341–1357. [CrossRef]
10. Fauzi, M.; Wicaksono, P. Total Suspended Solid (TSS) Mapping of Wadaslintang Reservoir Using Landsat 8 OLI. In *IOP Conference Series: Earth and Environmental Science-Proceedings of the 2nd International Conference of Indonesian Society for Remote Sensing (ICOIRS), Yogyakarta, Indonesia, 17–19 October 2016*; IOP Publishing: Bristol, UK, 2019; Volume 47, pp. 1–9.
11. Wang, H.; Wang, J.; Cui, Y.; Yan, S. Consistency of Suspended Particulate Matter Concentration in Turbid Water Retrieved from Sentinel-2 MSI and Landsat-8 OLI Sensors. *Sensor* **2021**, *21*, 1662. [CrossRef]
12. Gómez, J.L.; Dalence, J.S. Determinación del parámetro sólidos suspendidos totales (SST) mediante imágenes de sensores ópticos en un tramo de la cuenca media del río Bogotá (Colombia). *Rev. UD Geomática* **2014**, *9*, 19–27.
13. Torres Vera, M.A. Mapping of total suspended solids using Landsat imagery and machine learning. *Int. J. Environ. Sci. Technol.* **2023**, *20*, 11877–11890. [CrossRef]
14. Xu, H.; Xu, G.; Hu, X.; Wang, Y. Lockdown effects on total suspended solids concentrations in the Lower Min River (China) during COVID-19 using time-series remote. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *98*, 102301. [CrossRef] [PubMed]
15. Kumar, A.; Equeenuddin, S.M.; Mishra, D.R.; Acharya, B.C. Remote monitoring of sediment dynamics in a coastal lagoon: Long-term Spatio-temporal variability of suspended sediment in Chilika. *Estuar. Coast. Shelf Sci.* **2016**, *170*, 155–172. [CrossRef]
16. Li, W.; Yu, W. Modelling Reservoir Turbidity Using Landsat 8 Satellite Imagery by Gene Expression Programming. *Water* **2019**, *11*, 1479. [CrossRef]
17. Langhorst, T.; Pavelsky, T.; Eidam, E.; Cooper, L.; Davis, L.; Spellman, K.; Clement, S.; Arp, C.; Bondurant, A.; Friedmann, E.; et al. Increased scale and accessibility of sediment transport research in rivers through practical, open-source turbidity and depth sensors. *Res. Square* **2023**, *1*, 1–23. [CrossRef]

18. Hajigholizadeh, M.; Melesse, A.M. Assortment and spatiotemporal analysis of surface water quality using. *CATENA* **2016**, *151*, 247–258. [CrossRef]

19. Li, J.; Meng, Y.; Li, Y.; Cui, Q.; Yand, X.; Tao, C.; Wang, Z.; Li, l.; Zhang, W. Accurate water extraction using remote sensing imagery based on Normalized difference water index and unsupervised deep learning. *J. Hydrol.* **2022**, *612*, 128202. [CrossRef]

20. Zhang, Y.; Wu, L.; Ren, H.; Deng, L.; Zhan, P. Retrieval of Water Quality Parameters from Hyperspectral Images Using Hybrid Bayesian Probabilistic Neural Network. *Remote Sens.* **2020**, *12*, 1567. [CrossRef]

21. Chang, N.B.; Benjamin, W.; Jeffrey-Yang, Y.; Elovitz, M. Evaluation of dynamic linkages between evapotranspiration and land-use/land-cover changes with Landsat TM and ETM+ data. *Int. J. Remote Sens.* **2014**, *33*, 3733–3750.

22. Jaelani, L.M.; Ratnaningsih, R.Y. Spatial and Temporal Analysis of Water Quality Parameter using Sentinel-2A Data; Case Study: Lake Matano and Towuti. *Int. J. Adv. Sci. Eng. Inf. Technol.* **2018**, *8*, 547–553. [CrossRef]

23. Zheng, Z.; Wang, D.; Gong, F.; He, X.; Bai, Y. A Study on the Flux of Total Suspended Matter in the Padma River in Bangladesh Based on Remote-Sensing Data. *Water* **2021**, *13*, 2373. [CrossRef]

24. Abdelmalik, K.W. Role of statistical remote sensing for Inland water quality parameters prediction. *Egypt. J. Remote Sens. Space Sci.* **2016**, *21*, 193–200. [CrossRef]

25. Rahman, A.S.; Rahman, A. Application of Principal Component Analysis and Cluster Analysis in Regional Flood Frequency Analysis: A Case Study in New South Wales, Australia. *Water* **2020**, *12*, 781. [CrossRef]

26. Sagan, V.; Peterson, K.T.; Maimaitijiang, M.; Sidike, P.; Sloan, J.; Greeling, B.A.; Adams, C. Monitoring inland water quality using remote sensing: Potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Sci. Rev.* **2020**, *205*, 103187. [CrossRef]

27. Bernardo, N.; Watanabe, F.; Rodrigues, T.; Alcántara, E. Atmospheric correction issues for retrieving total suspended matter concentrations in inland waters using OLI/Landsat-8 image. *Adv. Space Res.* **2017**, *59*, 2335–2348. [CrossRef]

28. Chen, J.; Quan, W.; Duan, H.; Xing, Q.; Xu, N. An Improved Inherent Optical Properties Data Processing System for Residual Error Correction in Turbid Natural Waters. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 6596–6607. [CrossRef]

29. Wang, C.; Chen, S.; Li, D.; Wang, D.; Liu, W.; Yang, J. A Landsat-based model for retrieving total suspended concentration of estuaries and coasts in China. *Geoscientific Model Dev.* **2017**, *10*, 4347–4365. [CrossRef]

30. 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. [CrossRef]

31. Chongyang, W.; Weijiao, L.; Shuisen, C.; Dan, L.; Danni, W.; Jia, L. The spatial and temporal variation of total suspended solid concentration in Pearl River Estuary during 1987–2015 based on remote sensing. *Sci. Total Environ.* **2018**, *618*, 1125–1138.

32. Ghada, Y.E.; Marieke, A.E.; Meinte, B.; Kessel, T.; Gaytan, S.; Hendrik, J. Improving the Description of the Suspended Particulate Matter Concentrations in the Southern North Sea through Assimilating Remotely Sensed Data. *Ocean Sci. J.* **2011**, *46*, 179–204.

33. Cahyono, B.; Jamilah, U.L.; Nugroho, M.A.; Subekti, A. Analysis of Total Suspended Solids (TSS) at Bedadung River, Jember District of Indonesia Using Remote Sensing Sentinel 2A Data. *Singap. J. Sci. Res.* **2019**, *9*, 117–123.

34. Saberioon, M.; Brom, J.; Nedbal, V.; Soucek, P.; Cízar, P. Chlorophyll-a and total suspended solids retrieval and mapping using Sentinel-2A and machine learning for inland waters. *Ecol. Indic.* **2020**, *113*, 106236. [CrossRef]

35. Vakili, T.; Amanollahi, J. Determination of optically inactive water quality variables using Landsat 8 data: A case study in Geshlagh reservoir affected by agricultural land use. *J. Clean. Prod.* **2019**, *247*, 119134. [CrossRef]

36. 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]

37. Pizani, F.C.; Ferreira, A.F.; Amorim, C.C. Estimation of water quality in a reservoir from Sentinel-2 MSI and Landsat 8-OLI Sensor. *SPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *3*, 401–408. [CrossRef]

38. Ekercin, S. Water Quality Retrievals from High-Resolution Ikonos Multispectral Imagery: A Case Study in Istanbul, Turkey. *Water Air Soil Pollut.* **2007**, *183*, 239–251. [CrossRef]

39. Carrillo, I.D.; Medina, R.J. Multitemporal analysis of the flow of sediments using modis MYD09 and MOD09 images. *Cienc. Ing. Neogranadina-Univ. Mil. Nueva Guin.* **2019**, *29*, 69–86. [CrossRef]

40. Yeboah, Y.; Quaye-Ballard, J.; Amatey, A.; Appiah, A. Spatial prediction mapping of water quality of Owabi reservoir from satellite imageries and machine learning models. *Egypt. J. Remote Sens. Space Sci.* **2021**, *24*, 825–833.

41. Nguyen, T.H.; Phan, D.; Nguyen, H.T.; Tran, S.; Tran, T.; Tran, B.; Doan, T. Total Suspended Solid Distribution in au River Using Sentinel 2A Satellite Imagery. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *4*, 91–97. [CrossRef]

42. Markogianni, V.; Kalivas, D.; Petropoulos, G.; Dimitriou, E. Analysis on the Feasibility of Landsat 8 Imagery for Water Quality Parameters Assesment in an Oligotrophic Mediterranean Lake. *Int. J. Geol. Environ. Eng.* **2017**, *11*, 906–914.

43. Wang, D.; Ma, R.; Xue, K.; Loiselle, S.A. The Assessment of Landsat-8 OLI Atmospheric Correction Algorithms for Inland Waters. *Remote Sens.* **2019**, *11*, 169. [CrossRef]

44. Kavurmacı, M.; Karakuş, C.B. Evaluación de la calidad del agua de riego mediante análisis envolvente de datos e índices de calidad del agua basados en procesos de jerarquía analítica: El caso de la ciudad de Aksaray, Turquía. *Contam. Agua Aire Suelo* **2020**, *55*, 1–123.

45. Ruiz, D.C. Method of Estimating Total Suspended Solids as an Indicator of Water Quality Using Satellite Images. In *Reposiorio Institucional-Biblotteca Digital*; National University of Colombia: Bogotá, Colombia, 2017; pp. 1–47.

46. Fernández, A.; Moreira, J.M. *Methodology for the Multitemporal Monitoring of the Quality of Coastal Waters in Andalusia through Landsat-TM Image Processing*; Remote Sensing Uses and Applications, University of Seville; Deposito de Investigation Universidad de Sevilla: Seville, Spain, 2014; Volume 1, pp. 1–65.

47. Lu, H.; Liu, Q.; Liu, X.Y.; Zhang, Y. Un estudio sobre la construcción semántica y la aplicación de imágenes y datos de teledetección por satélite. *Rev. De Informática Organ. Usuario Final. (JOEUC)* **2021**, *33*, 1–20.

48. Hernandez, J. Methodology for the Evaluation of Volumetric and Energy Impacts Inflows by Transfer: Case Study Upper Course of the Lerma River. Master's Thesis, Inter-American Water Resources Center/UAEMex, Toluca, Mexico, 2018. (In Spanish).

49. Ciancia, E.; Campanelli, A.; Lacava, T.; Palombo, A.; Pascucci, S.; Pergola, N.; Tramutoli, V. Modeling and Multi-Temporal Characterization of Total Suspended Matter by the Combined Use of Sentinel 2-MSI and Landsat 8-OLI Data: The Pertusillo Lake Case Study (Italy). *Remote Sens.* **2020**, *12*, 2147. [CrossRef]

50. Doña, C. Monitoring water quality and hydrological patterns of wetlands using recent techniques in remote sensing. In *Departament de Física de la Terra i Termodinàmica*; Universitat Valencia: Valencia, Spain, 2016; pp. 1–93.

51. DOF. *Water Analysis—Measurement of Total Nitrogen Kjeldahl in Natural Water, Wastewater and Treated Wastewater—Test Method*; Ministry of Economy: Mexico City, Mexico, 2010. (In Spanish)

52. DOF. *Water Analysis—Measurement of Dissolved Solids and Salts in Natural Water, Wastewater, and Treated Wastewater—Test Method*; Ministry of Economy: Mexico City, Mexico, 2015. (In Spanish)

53. DOF. *Water Analysis—Measurement of Total Phosphorus in Natural Water, Wastewater, and Treated Wastewater—Test Method*; Ministry of Economy: Mexico City, Mexico, 2001. (In Spanish)

54. DOF. *Water Analysis—Measurement of Chemical Oxigen Demand in Natural Water, Wastewater, and Treated Wastewater—Test Method*; Ministry of Economy: Mexico City, Mexico, 2012. (In Spanish)

55. *NOM-001-SEMARNART-2021*; Official Mexican Standard. SEGOB: Mexico City, Mexico, 2023. (In Spanish)

56. Kim, Y.; Im, J.; Ha, H.K.; Choi, J.-K.; Ha, S. Machine learning approaches to coastal water quality monitoring using GOCI Satellite data. *GIS Sci. Remote Sens.* **2014**, *51*, 158–174. [CrossRef]

57. Li, C.; Rousta, I.; Olafsoon, H.; Zhang, H. Lake Water Quality and Dinamics Assesssment during 1990–2020 (A case Study: Chao Lake: China). *Atmosphere* **2023**, *14*, 382. [CrossRef]

58. Li, L.; Gu, M.; Gong, C.; Hu, Y.; Wang, X.; Yang, Z.; He, Z. An advanced remote sensing retrieval method for urban non-optically active water quality parameters: An example from Shanghai. *Sci. Total Environ.* **2023**, *880*, 163389. [CrossRef]

59. Mun, J. *Risk Simulator User Manual in Spanish*; R-Real Options Valuation: Dublin, Ireland, 2012.

60. Swain, R.; Sahoo, B. Improving river water quality monitoring using satellite data product and a genetic algorithm processing aproach. *Sustain. Water Qual. Ecol.* **2017**, *10*, 122–149.

61. IBM. *IBM SPSS Statistics 28 Brief Guide*; IBM Corporation: Endicott, NY, USA, 2021; Volume 1, pp. 1–90.

62. Zou, D.; Lloyd, J.V.; Baumbusch, J.L. Using SPSS to analyze Complex Survey Data: A Primer. *J. Mod. Appl. Stat. Methods* **2019**, *18*, 16. [CrossRef]

63. Aiman, M.; Mohosen, M.; Hossam, S. Statistical estimation of Rosetta Branch Water Quality using multi-spectral data. *Water Sci.* **2014**, *28*, 18–30.

64. CONAGUA. (1 July 2019). National Water Commission. Available online: https://app.conagua.gob.mx/bandas/ (accessed on 14 April 2021).

65. Tu, M.C.; Smith, P.; Filippi, A.M. Hybrid forward-selection method-based water-quality estimation via combining Landsat TM, ETM+, and OLI/TIRS images and ancillary environmental data. *PLoS ONE* **2018**, *13*, e0201255. [CrossRef] [PubMed]

66. Sundarabalan, V.; Pahlevan, N.; Smith, B.; Binding, C.; Schalles, J.; Loisel, H.; Boss, E. Robust algorithm for estimating total suspended solids (TSS) in inland and nearshore coastal waters. *Remote Sens. Environ.* **2020**, *246*, 111768.