Inversion and Monitoring of the TP Concentration in Taihu Lake Using the Landsat-8 and Sentinel-2 Images

Yongchun Liang [1,†], Fang Yin [2,†], Danni Xie [2], Lei Liu [1,*], Yang Zhang [3] and Tariq Ashraf [1]

1 School of Earth Science and Resources, Chang'an University, Xi'an 710054, China
2 Shaanxi Key Laboratory of Land Consolidation, School of Land Engineering, Chang'an University, Xi'an 710054, China
3 School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, Xi'an 710049, China
* Correspondence: liul@chd.edu.cn; Tel.: +86-29-82339936
† These authors contributed equally to this work.

Abstract: Eutrophication is a significant factor that damages the water ecosystem's species balance. The total phosphorus (TP) concentration is a vital water quality indicator in assessing surface water eutrophication. This paper predicts the spatial distribution of TP concentration using remote sensing, measured data, and the partial least squares regression (PLSR) method. Based on the correlation analysis, the models were built and tested using the TP concentration and Sentinel-2 Multispectral Instrument (MSI) and Landsat-8 Operational Land Imager (OLI) image spectra. The results demonstrated that the best technique based on band combinations of the Sentinel-2 and Landsat-8 images achieved good precision. The coefficient of determination (R^2), root mean square error of prediction (RMSEP), and residual prediction deviation (RPD) were 0.771, 0.023 mg/L, and 2.086 for Sentinel-2 images and 0.630, 0.032 mg/L, and 1.644 for Landsat-8 images, respectively. The TP concentration maps were interpolated using the inverse distance weighting method, and the inversion results obtained from the images were in good agreement. The western and northwestern regions of Taihu Lake, where significant cyanobacterial blooms occurred, had TP concentrations greater than 0.20 mg/L; nevertheless, the central and eastern regions had amounts ranging from 0.05 to 0.20 mg/L. In order to prove the extensibility of the model, the optimal algorithm was applied to the Sentinel-2 and Landsat-8 images in 2017. The optimal algorithm based on Landsat-8 images has a better verification effect (RMSEP = 0.027 mg/L, and R = 0.879 for one Landsat-8 image), and the optimal algorithm based on Sentinel-2 images has moderate verification effect (RMSEP = 0.054 mg/L and 0.045 mg/L, and R = 0.771 and 0.787 for two Sentinel-2 images). The interpolation and inversion maps are in good agreement, indicating that the model is suitable for the Landsat-8 and Sentinel-2 images, which can be complementary for higher temporal resolutions. Monitoring water quality using multiple remote sensing images can provide the scientific basis for water quality dynamic monitoring and prevention in China.

Keywords: eutrophication; total phosphorus; remote sensing; partial least squares regression; correlation analysis; inversion

1. Introduction

Water is a vital foundation for supporting human survival and social development, and freshwater lakes are critical natural resources that serve as drinking water sources, tourism destinations, and habitats within the biosphere [1,2]. However, pollution emissions increased with rapid economic development and the increased intensity of land development and human activities [3]. Lakes face severe problems, including shrinking lake areas and eutrophication [4]. Water eutrophication has been identified as a significant cause of freshwater resources since the 1960s [5], and 54%, 53%, 48%, 41%, and 28% of the lakes in Asia, Europe, North America, South America, and Africa are eutrophic [6]. Eutrophication

leads to the rapid growth of algae and other planktonic organisms and thus the reduction of dissolved oxygen (DO) contents [7,8] and the ecological health and balance in the water environment [9]. The mechanisms of water eutrophication are not fully understood, but excessive nutrient loading into the surface water system is considered one of the major factors [10]. Phosphorus (P) is the crucial element in determining the algae growth in inland waters [2], and 80% of water eutrophication is only restricted to P [11], which is consistent with previous research [12]. Excess nutrient elements encourage algae to overgrow, resulting in algal blooms with high nutrient concentrations and ideal temperatures [13].

Taihu Lake, the third-largest freshwater lake in China, has a water surface area of 2338 km^2 and supplies drinking water to more than 10 million people within its watershed. Due to continuous industrialization and urbanization, industrial effluent and domestic sewage discharges have deteriorated the water quality in Taihu Lake since the 1980s [1–3]. The management of Taihu Lake has been strengthened since the emergence of cyanobacteria in 2007, while the water eutrophication remains moderate [4]. Thus, it is vital to monitor Taihu Lake's total phosphorus (TP) content to aid in the water quality analysis of temporal and spatial changes and eutrophication therapy [5,6].

The commonly used methods for monitoring TP concentration include traditional chemical analysis and satellite remote sensing monitoring methods. The chemical sampling method consists of routine water quality sampling and laboratory analysis processes [7], which provide accurate data for water quality assessment and subsequent water quality studies. However, the method is limited by time consumption and complex sampling, and it is not easy to obtain information about the entire lake [8]. As a tool to supplement the traditional field-based sampling method [9], satellite remote sensing images are widely applied to water quality monitoring and management [10–12]. Although the influence of the TP concentration on the measured spectrum and remote sensing spectral bands is not yet evident, some studies have found a relationship between TP concentration and spectral reflectance [8]. The satellite remote sensing images used to retrieve the TP concentration involved multispectral images, such as the Landsat Thematic Mapper (TM) [13], Landsat Enhanced Thematic Mapper Plus (ETM+) [8], Landsat 8 Operational Land Imager (OLI) [6], Sentinel-2 [14,15], Sentinel-3 [16], HJ-1A satellites Charge Coupled Device (CCD) imagery [17], GaoFen-1(GF-1) [18], Geostationary Ocean Color Imager (GOCI) [19], and Moderate Resolution Imaging Spectroradiometer (MODIS) [20,21]. In addition to the spectral bands of various images, retrieval studies of water TP concentration can also be carried out using the measured spectra collected using a spectrometer [18,22].

In previous investigations, hyperspectral and multispectral imagery and measured spectra successfully monitored and simulated water TP concentration. Hyperspectral images typically have a low spatial resolution, which significantly hinders the expansion of hyperspectral remote sensing technology. The measured spectra, although with a higher spectral resolution than images, are taken at sites that would be time-consuming, expensive, and limited to a small scale. Landsat-8 and Sentinel-2 multispectral images have been widely used for water monitoring due to their low cost, high spatial and spectral resolution, frequent revisits, and extensive coverage [23–25]. Thus, multispectral images are more practicable for water quality parameter inversion analyses. The Landsat-8 image has a spatial resolution of 30 m and is widely used in water quality monitoring [6,26,27]. The deficiency of available Landsat imagery for some months due to cloud cover and a longer revisit period of 16 days can be offset by Sentinel-2 imagery with a short revisit time of 5 days. Sentinel-2 images have 13 bands ranging from visible to shortwave infrared with a higher spatial resolution of 10 m, 20 m, and 60 m. Water quality indicators [28,29] such as chlorophyll-a (Chl-a) [25,30], suspended matter [27], colored dissolved organic matter (CDOM) [24], turbidity, chemical oxygen demand (COD), DO, TP, and other water quality parameters [14,31] can be measured using Sentinel-2 images. In addition, the Sentinel-2 images have been used in various investigations such as phytoplankton bloom mapping [32] and mapping lake water quality [33]. Sentinel-2 and Landsat-8 imagery are widely used in water quality monitoring and parameter retrieving.

Studies found that water quality remote sensing monitoring methods have developed through multivariate data analysis [34–36], machine learning [12,31,37], and integrated methods. Machine learning methods work well in remote sensing studies of water quality parameters, but accurate models require many training samples. The partial least squares (PLS) method was developed in the 1960s by Herman Wold [38] for constructing predictive models when many factors are collinear. Partial least squares regression (PLSR) is one of the multivariate statistical analysis methods that focuses on regression model with multiple dependent variables and multiple independent variables. The basic algorithm of PLSR is through successive least squares fit on variables to find linear combinations of the original regression variables [38]. The method was used in water quality monitoring, such as the Chl-a concentration of Shitoukoumen Reservoir in China [39] and the TP concentration of Morse Reservoir in the USA [40].

The objectives of this study are to (1) construct the PLSR model by combining the measured TP concentration with Sentinel-2 and Landsat-8 multispectral images, (2) compare different maps of interpolated TP concentration with the retrieved results of remote sensing images to determine the spatial and temporal distribution of TP concentration in Taihu Lake, and (3) discuss the related factors of the TP variations. Using the results of TP remote sensing inversion to accurately depict the fluctuations in eutrophication levels and the status of water quality conditions in Taihu Lake, the study also aims to offer environmental managers vital information.

2. Materials and Methods

2.1. Study Site and Experimental Design

Taihu Lake (30°55′–31°33′N, 119°52′–120°37′E) is mainly located in the southern Jiangsu province and covers parts of the Jiangsu and Zhejiang provinces, including some cities such as Wuxi, Suzhou, Changzhou, and Huzhou. The water area, total shoreline length, average water depth, maximum water depth, and total water storage of Taihu Lake are 2428 km^2 (the water surface area is 2338 km^2, and the island area is 90 km^2), 393.2 km, 1.90 m, 3.34 m [41], and 4.40 × 10^9 m^3 [35], respectively [19]. Most of the runoff comes from the southwestern mountainous watershed and flows out at the east side of the lake after regulation and storage [5]. It has a subtropical monsoon climate, with an average temperature of 14.90–16.20 °C and an annual precipitation of 1000–1400 mm [13,42].

Twenty sampling sites were evenly spaced throughout Taihu Lake for this investigation, with no sampling sites in east Taihu Lake (Figure 1). Water samples were collected from monitoring sites monthly from January to December 2016, totaling 240 samples. Some of the measured data used in this study are shown in Table S1. The water samples were collected using 500 mL plastic bottles and stored at 4 °C before subsequent laboratory analysis.

2.2. Sample Collection and Analysis

2.2.1. Water Quality Parameters and Environmental Factors

The water quality monitoring process complies with the China Inspection Body and Laboratory Mandatory Approval (CMA) metrology certification. The water samples were filtered through a 0.45 μm membrane and left to settle for 30 min. The supernatant was taken by a siphon into sampling bottles and stored with an H_2SO_4 fixative to make the pH of the sample ≤ 1. The water quality parameters to be measured include TP, total nitrogen (TN), ammonia nitrogen (NH_3-N), permanganate index (COD_{Mn}), biochemical oxygen demand (BOD_5), chemical oxygen demand (COD), DO, and potential of Hydrogen (pH). The detection methods of all water quality parameters are based on the environmental quality standards for surface water (State Environmental Protection Administration, 2002). All parameters except for pH are expressed in milligrams per liter (mg/L). All water quality indicators are assessed according to the standard Chinese method, Standard Methods for the Examination of Water and Wastewater Editorial Board, published in 2002. The data on the TP concentration in 2017 were downloaded from a paper published previously [43],

with 15 sampling sites in Taihu Lake. Twelve groups of TP concentration data were obtained once a month by extraction. The monthly precipitation and temperature data in 2016 were provided by Loess Plateau Science Data Center, National Earth System Science Data Sharing Infrastructure, and National Science & Technology Infrastructure of China (http://loess.geodata.cn (accessed on 27 November 2021)). The spatial resolution of the precipitation and temperature data is 0.008333°. The average precipitation in August was 57.695 mm, and that in September was 238.87 mm. The downloaded temperature and precipitation data were vector files and processed by ARCGIS 10.3 software.
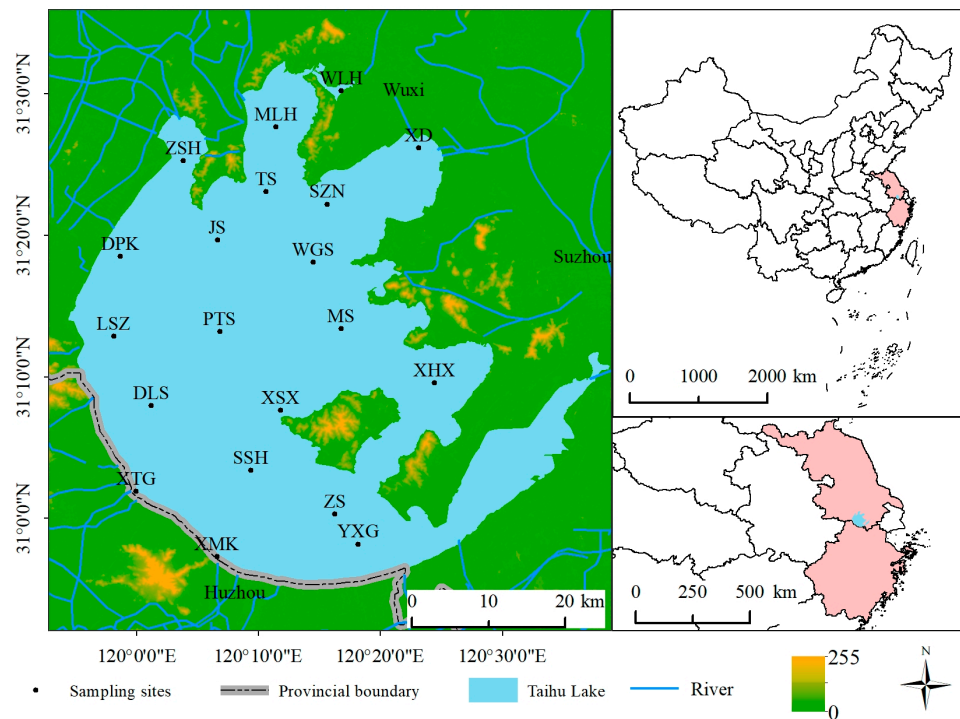

Figure 1. The location of sampling sites.

2.2.2. Remote Sensing Images Collection and Correction

The Landsat-8 and Sentinel-2 images were analyzed and compared in this research. The Landsat-8, including two sensors, the OLI and the Thermal Infrared Sensor (TIRS), was launched in 2013. The OLI has eight multispectral bands with a 30 m spatial resolution and a panchromatic band with a 15 m spatial resolution [44]. The MSI optical sensor on Sentinel-2 includes 13 bands with three different spatial resolutions: the visible and Near Infrared (NIR) bands with a 10 m resolution; the red edge, Narrow NIR, and Shortwave Infrared (SWIR) bands with a 20 m resolution; and the coastal aerosol, water vapor, and SWIR-Cirrus bands with a 60 m resolution [30]. Band1 to Band7 of the Landsat-8 image were selected as the corresponding bands of Band1, Band2, Band3, Band4, Band8, Band11, and Band12 of the Sentinel-2 image. The Landsat 8 image (Path: 119; Row: 038) and the Sentinel-2 image (Relative Orbit Number R089) were from the USGS Landsat Look Viewer https://landlook.usgs.gov/landlook/viewer.html/ (accessed on 4 September 2020) and the Copernicus Open Access Hub https://scihub.copernicus.eu/ (accessed on 21 November 2020), respectively. Two Landsat-8 images acquired on 27 July and 13 September 2016, along with two Sentinel-2 images acquired on 23 July and 1 September 2016, were chosen for further analysis due to their acquisition times being close to the sampling times. One Landsat-8 image acquired on 30 July 2017, along with two Sentinel-2 images acquired on 28 July and 27 August 2017, were chosen for verification. The time interval between the images and sampling sites is shown in Table S2.

The pre-processing of the Landsat-8 images included radiance calibration and atmospheric correction through the Fast Line-of-sight Atmospheric Analysis of the Spectral

Hypercubes (FLASH) model [45], which could remove atmospheric effects and convert the images to surface reflectance [46]. The sensor height, surface elevation, and atmospheric model parameters for atmospheric correction were similar to those in other studies [45,47]. The mean value and standard deviation of the corrected image were smaller than those of the pre-corrected image, but the signal-to-noise ratio of the corrected image was larger than that of the pre-corrected image. The pre-processing of the Sentinel-2 image was mainly the atmospheric correction and resampling. The raw Sentinel-2 images (L1C level) were atmospherically corrected by the Sentinel-2 toolbox, Sen2Cor v2.5.5, developed by the European Space Agency (ESA), and converted to the reflectance data (L2A level). After all Sentinel-2 images were atmospherically corrected, Band 10 became unavailable. All bands of the atmospheric correction image were resampled by nearest-neighbor interpolation to a 10 m resolution using the Sentinel Application Platform (SNAP) software provided by ESA and were saved as the Environment for Visualizing Images (ENVI) standard format.

The image spectra of the 20 samples were extracted by ArcGIS 10.3 based on the coordinates of the samples. The correlation between the image spectra (single-band reflectance and band combination reflectance) and the measured TP concentration was analyzed, and sensitive bands were selected based on the correlation coefficients.

2.3. Model Simulation and Assessment

The PLSR model was chosen to retrieve the measured TP concentration based on the image spectra. PLSR is a method developed in the 1960s by Herman Wold for constructing predictive models when the factors are many and highly collinear [38]. The PLSR model was developed with sensitive bands serving as the independent variables and measured concentrations serving as the dependent variable. The model coefficients were acquired using the MATLAB (R2018a, USA) program during the modeling phase. The accuracy and validity of the model were evaluated based on the validation results. The validation parameters of the PLSR model were the root mean square error of prediction (*RMSEP*), the coefficient of determination (R^2), and the residual prediction deviation (*RPD*).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y} - y_i)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \tag{1}$$

$$RMSEP = \sqrt{\frac{\sum_{i-1}^{n}(\hat{y} - \bar{y})^2}{N}} \tag{2}$$

$$RPD = \frac{SD}{RMSEP} \tag{3}$$

The model with a lower RMSEP and a higher R^2 and RPD is suitable for prediction. The R^2 is an essential factor in evaluating the prediction model with the following classifications: excellent prediction ($R^2 > 0.90$), good prediction ($R^2 = 0.82$–0.90), approximate quantitative prediction ($R^2 = 0.66$–0.81), prediction that can possibly distinguish between high and low values ($R^2 = 0.50$–0.65), and unsuccessful prediction ($R^2 < 0.50$) [48]. In addition, the model has poor results and cannot obtain accurate values when the $RPD \leq 1.5$, the model is considered moderately effective when the $RPD > 1.5$ and the $RPD < 2$, and the model has an excellent predictive ability when $RPD \geq 2$.

2.4. Correlation Analysis

Correlation analysis is a term used to denote the association or relationship between two (or more) quantitative variables [49]. The experimental data were analyzed using statistical product service solutions (SPSS 22.0) and Origin 2022. Band reflectance and TP concentration were subjected to correlation analysis to investigate the relationship between the spectral reflectance and TP concentration. Pearson's correlation coefficient was calculated to reveal the relationship between TP and various environmental factors

(COD$_{Mn}$, BOD$_5$, NH$_3$-N, TN, DO, pH, COD) and meteorological factors (precipitation and temperature). The equation for the correlation coefficient is:

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2}\sqrt{\sum (y - \bar{y})^2}} \tag{4}$$

Here, R is the correlation coefficient, and x and y are the mean values of variables x and y, respectively.

3. Results

3.1. The Variation in the TP Concentration in Taihu Lake

The TP concentration changed from 0.02 mg/L to 0.27 mg/L, and the average was 0.07 mg/L (Figure 2). The mean TP concentration in summer and autumn (0.07 mg/L) was higher than that in spring and winter (0.06 mg/L). The TP concentration at all sampling sites is higher in February, May, June, July, August, and September. In months except for February, the most significant changes in TP concentration in Taihu Lake were observed in the middle of each year. The suitable water temperature in the summer and fall provides better environmental conditions for algae growth [50], while Taihu Lake is hypereutrophic. Therefore, the TP concentration in the warm months was selected for the study.
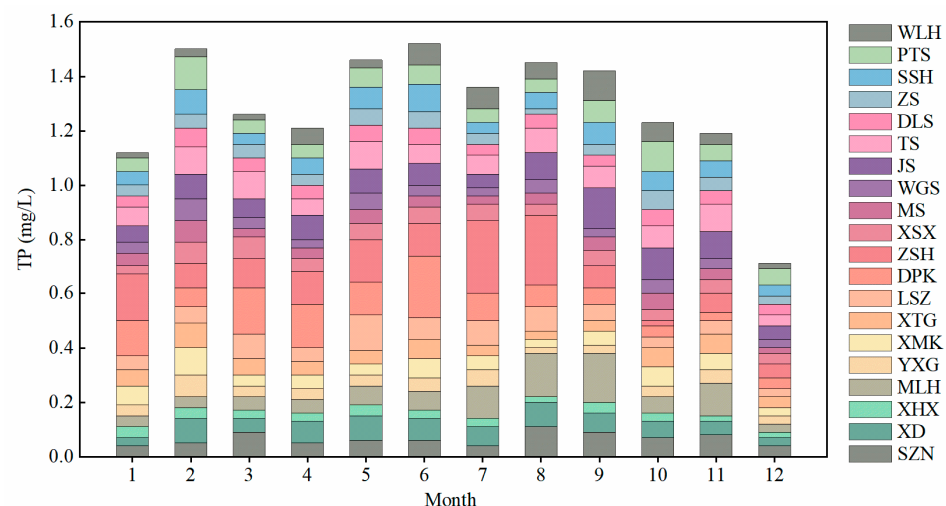

Figure 2. Monthly variation in the TP concentration of Taihu Lake in 2016.

Most of the water enters Taihu Lake from the west and flows towards the east; therefore, a large amount of domestic and industrial wastewater is discharged into Taihu Lake from the upstream rivers in the northwestern and western regions [51]. The rainy season in the study area is from May to September [52], and the center of rainfall in 2016 was in the western part of the lake, resulting in a high TP concentration [53]. Comparing the different sites, TP concentrations on MLH, ZSH, DPK, and JS located in the northwestern and northern parts of Taihu Lake are higher than others (Figure 2). Particularly, at ZSH and DPK, the average TP concentration was 0.127 mg/L and 0.103 mg/L, respectively, which exceed the limit for type IV water, showing that the P pollution in the western part of Taihu Lake was more severe than that of other regions [54] (Figure 2). Based on the prior analysis results, combined with the proportional contribution of P in eutrophication [11], TP concentrations in the middle of 2016 were selected for analysis in this study.

3.2. Modeling and Monitoring Using Sentinel-2 Image

3.2.1. Correlation between TP Concentration and Sentinel-2 Image Spectra

Band1, Band2, Band3, Band4, Band8, Band11, and Band12 of the Sentinel-2 images were selected, and the image spectra of the sampling sites were collected for further analysis.

Table 1 shows all the correlation coefficients between the single-band reflectance and TP concentration and some high-correlation coefficients between the band combinations and TP concentration. The single-bands and TP concentration were positively correlated, with the correlation between Band8 and the TP concentration being the highest (0.484). The Sentinel-2 image spectra were combined in several methods, including the sum, difference, ratio, and ratio combination $(r_i + r_j)/(r_i - r_j)$, where r_i and r_j denote the reflectance of bands i and j ($i \neq j$). The band combination correlation coefficients ranged from -0.533 to 0.540. Then, the selection of the optimal band combination was performed by the exhaustive method, which allowed us to establish a high-accuracy regression model of the water quality parameters in Taihu Lake [6]. It was found that the R^2 of the PLSR model constructed using a single band and the TP concentration was about 0.30. The results of several trials indicated that the R^2 of the models constructed on sensitive spectral variables and the TP concentration ranged from 0.20 to 0.77, and the sensitive spectral variables, including $((B_1 + B_8)/(B_1 - B_8))$, $((B_2 + B_8)/(B_2 - B_8))$, $((B_3 + B_8)/(B_3 - B_8))$, and $((B_4 + B_8)/(B_4 - B_8))$, had a better effect. Thus, the ratio combinations were used to construct the PLSR model.

Table 1. The correction between band combinations and TP concentration on Sentinel-2 and Landsat8.

Sensors	Band	R	Band Combination	R	Band Combination	R	Band Combination	R
Sentinel-2	B1	0.122	B8 + B11	0.514 **	B8/B1	0.505 **	(B3 + B8)/(B3−B8)	0.464 **
	B2	0.094	B8 + B12	0.510 **	B8/B2	0.513 **		
	B3	0.115	B11−B12	0.415 **	B8/B3	0.511 **		
	B4	0.111	B1/B8	−0.491 **	B8/B4	0.540 **		
	B8	0.484 **	B2/B8	−0.505 **				
	B11	0.392 *	B3/B8	−0.476 **				
	B12	0.361 *	B4/B8	−0.533 **				
Landsat-8	B1	−0.013	B1−B4	−0.439 **	B1/B5	−0.484 **	B5/B2	0.519 **
	B2	−0.004	B1−B5	−0.518 **	B2/B4	−0.496 **	B5/B4	0.427 **
	B3	0.135	B2−B3	−0.450 **	B2/B5	−0.473 **	(B1 + B5)/(B1−B5)	0.422 **
	B4	0.247	B2−B4	−0.490 **	B3/B5	−0.437 **	(B2 + B5)/(B2−B5)	−0.576 **
	B5	0.572 **	B2−B5	−0.433 **	B4/B1	0.448 **		
	B6	0.124	B5−B6	0.590 **	B5/B1	0.538 **		
	B7	0.082	B1/B4	−0.471 **	B4/B2	0.488 **		

* Correlation is significant at the 0.05 level. ** Correlation is significant at the 0.01 level.

To eliminate the effects of cloud, land, and aquatic vegetation from the images, three sampling sites (DLS) on 23 July 2016 and (MLH, JS) on 1 September 2016 were removed from the experiment. The measured data was divided with the 3:1 ratio into the training dataset (26 sample sites) and testing dataset (11 sample sites). The model was constructed by combining the Sentinel-2 image spectra on 23 July 2016 and 1 September 2016 with the measured TP concentration on 1–4 August 2016 and 1–8 September 2016. The optimal model was constructed as follows (formula 5):

$$TP = 0.0527 + 0.0231 * \left(\frac{B_1 + B_8}{B_1 - B_8} \right) + 0.0039 * \left(\frac{B_2 + B_8}{B_2 - B_8} \right) - 0.0536 * \left(\frac{B_3 + B_8}{B_3 - B_8} \right) + 0.0149 * \left(\frac{B_4 + B_8}{B_4 - B_8} \right) \quad (5)$$

where, B_1, B_2, B_3, B_4, B_8 are the spectral reflectance of Sentinel-2 Band1, Band2, Band3, Band4, and Band8, respectively.

3.2.2. Verification of the Optimal Algorithm

A relatively satisfactory correlation is observed between the estimated and measured TP concentration (Figure 3). Figure 3a is the scatter plot of the training dataset, which has an RMSEP of 0.023 mg/L, an R^2 of 0.771, and an RPD of 2.086, indicating that it is suitable for predicting the TP concentration in Taihu Lake. The scatter plot of the testing dataset (Figure 3b) has a moderate correlation coefficient of 0.782 between the estimated and measured TP concentration and a modest RMSEP of 0.020 mg/L, which indicates that the model has a medium prediction effect.
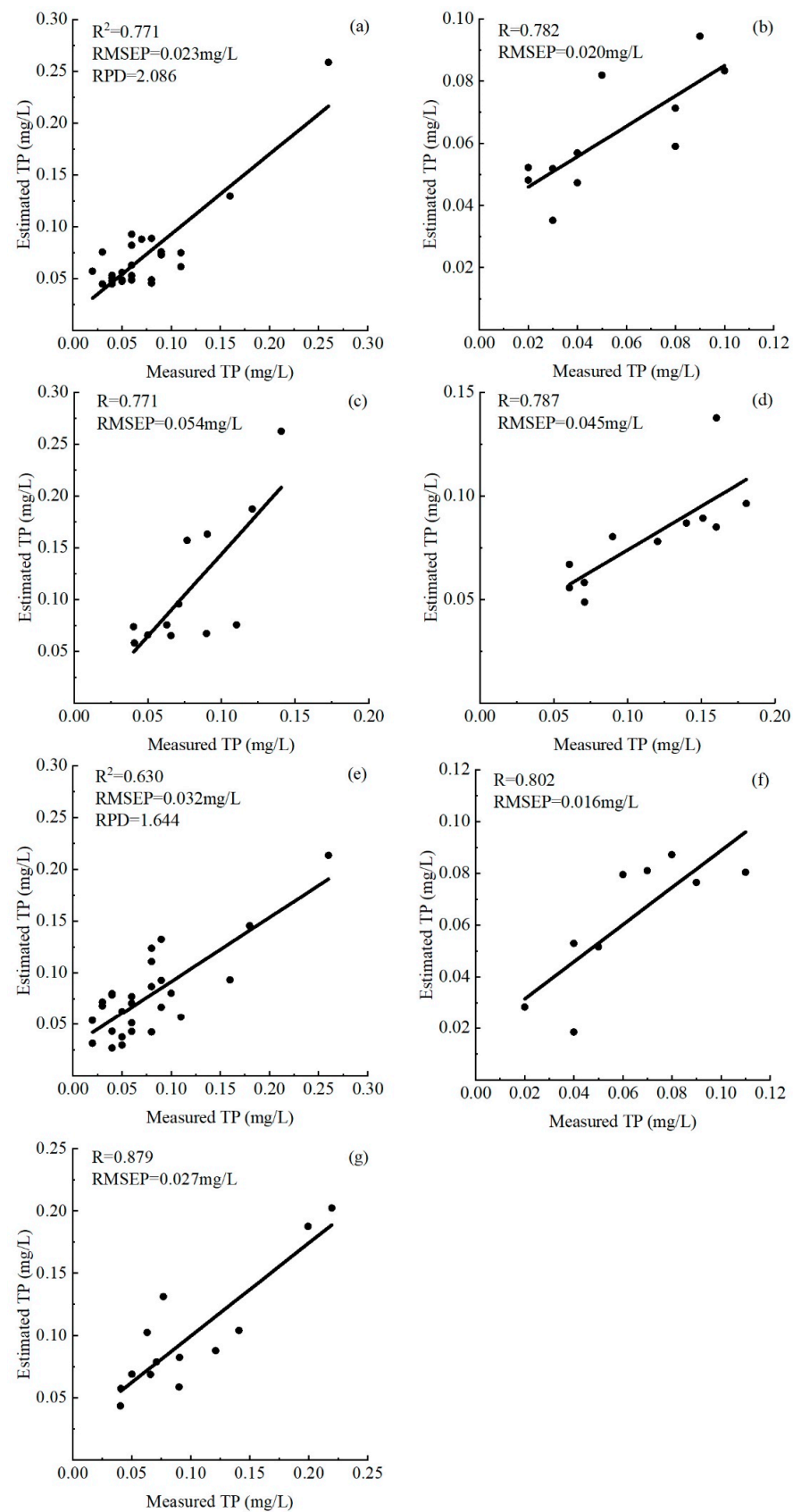
Figure 3. The scatter of predicted and measured TP using Sentinel-2 ((**a**) training dataset; (**b**) testing dataset; (**c**) verifying dataset (20170728); (**d**) verifying dataset (20170827)) and Landsat-8 ((**e**) training dataset; (**f**) testing dataset; (**g**) verifying dataset (20170730)).

To demonstrate the model's accuracy, the optimal algorithm was validated using data on the TP concentration for the same period in 2017. The algorithm was calculated for the same band combinations on the Sentinel-2 images acquired on 28 July 2017 and 27 August 2017. Similarly, several sampling sites were removed from the experiment to eliminate the effects of cloud, land, and aquatic vegetation from the images (three and four sites on 28 July and 27 August 2017, respectively). The algorithm has a moderate verification effect, where the R and RMSEP of the verification set (28 July 2017) reached 0.771 and 0.054 mg/L (Figure 3c), and the R and RMSEP of the verification set (27 August 2017) ran 0.787 and 0.045 mg/L (Figure 3d). Figure 3c shows that low and medium values have better predictions, but the high values are overestimated. Figure 3d shows that the low and medium values have better predictions, but the high values are underestimated.

3.2.3. Temporal and Spatial Distribution of TP Concentration Based on Sentinel-2 Images

The inverse distance weighting (IDW) approach was applied to spatially interpolate the TP concentration of 1–4 August and 1–8 September 2016. Figure 4a,d showed that the TP concentration gradually increased from the southeast to the northwest, with a high TP concentration in the west and northwest of Taihu Lake. The model was constructed with a two-month dataset representing different quality water conditions. Formula 5 was applied to Sentinel-2 images on 23 July and 1 September 2016 to obtain the inversion maps of TP concentration (Figure 4b,e). The inversion maps show an increasing trend of TP concentration from the southeast to the north and northwest, with a higher TP concentration in Taihu Lake's west and southwest parts. The TP concentration in the south and some areas of the north of Taihu Lake was low, with a concentration of about 0.00–0.05 mg/L, which belongs to type II and type III water based on the environmental quality standards for surface water published by the State Environmental Protection Administration in 2002; the TP concentration in the central and north regions was higher, with a concentration of 0.05–0.10 mg/L, which belongs to type IV water. The highest TP concentration exceeds 0.10 mg/L in the northwest and southeast regions, belonging to type V and low V water. Specifically, a series of anomalies with a significantly high TP concentration occurred in the central to the west part of Taihu Lake in the inversion map of 23 July 2016 (Figure 4b).

3.3. Modeling Using the Landsat 8 Image

3.3.1. Correlation between TP Concentration and Landsat-8 Image Spectra

Table 1 shows all the correlation coefficients between the image spectra of Landsat-8 and the measured TP concentration. Similarly, the correlation between the single band and TP concentration was lower, except for Band5, with a correlation coefficient of 0.572. The sensitive bands were selected from band combinations based on the relatively high correlation coefficients. With the same band combination method as the Sentinel-2 image, correlation coefficients of band combinations ranged from −0.576 to 0.590. An exhaustive process was used to select the optimal band combinations. The results of several trials showed that models constructed using sensitive spectral variables and the TP concentration have an R^2 ranging from 0.30 to 0.63, and band combinations including B_1-B_4, B_1-B_5, B_2-B_3, B_2-B_4, B_2-B_5, and B_5-B_6 indicate a better effect, with an R^2 of 0.63. Therefore, these band subtractions were used to construct the PLSR model.

Four sampling sites, XMK and DLS on 27 July 2016 and JS and YXG on September 13, 2016, were removed from the experiment. The rest data were divided with a 3:1 ratio into a training dataset (27 sites) and testing datasets (9 sites). The formula 6 was the optimal model of the TP concentration:

$$TP = 0.07082 + 0.00004*(B_1 - B_4) - 0.00007*(B_1 - B_5) - 0.0001*(B_2 - B_3) - 0.00001*(B_2 - B_4) - 0.00009*(B_2 - B_5) + 0.00021*(B_5 - B_6) \quad (6)$$

where, B_1, B_2, B_3, B_4, B_5, and B_6 are the spectral reflectance of Band1, Band2, Band3, Band4, Band5, and Band6 on the Landsat-8 image, respectively.
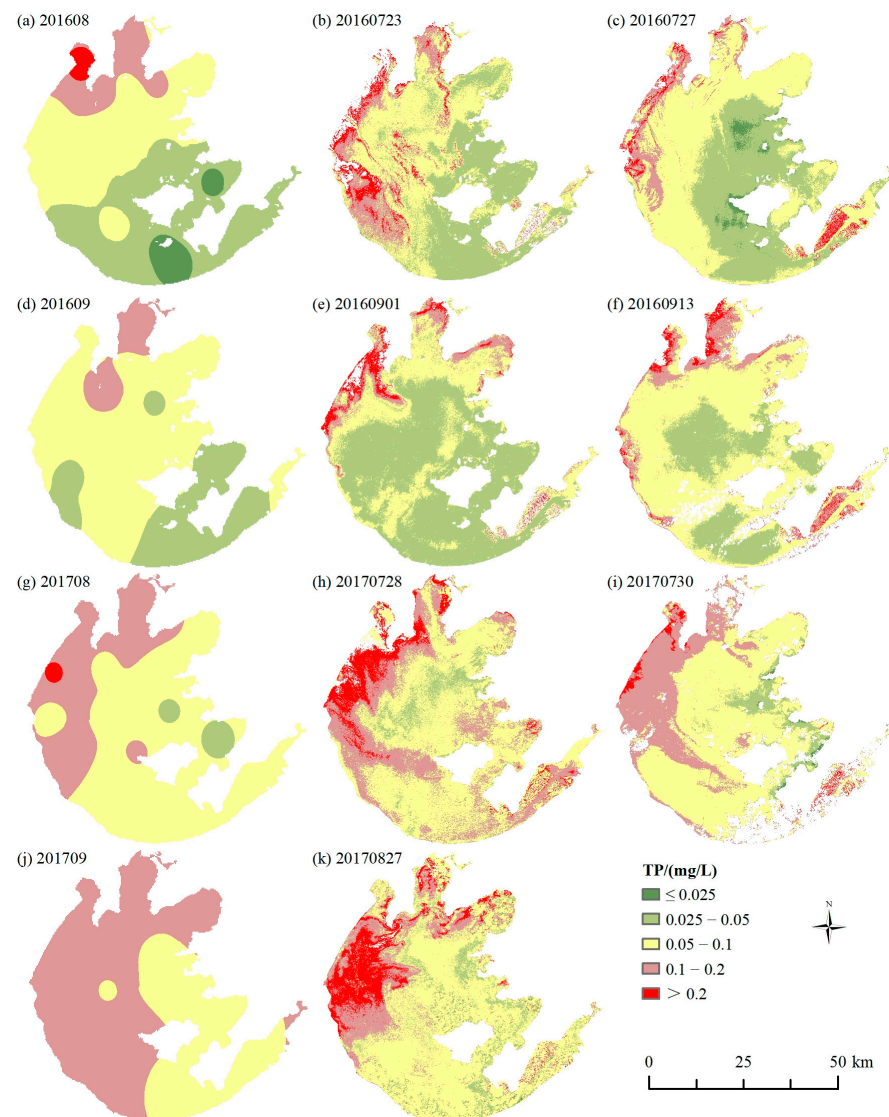
Figure 4. Spatial variation in the TP concentration in Taihu Lake. (**a**,**d**,**g**,**j**) were interpolated maps, and (**b**,**e**,**h**,**k**) of the Sentinel-2 image and (**c**,**f**,**i**) of the Landsat-8 image were inversed maps.

3.3.2. Verification of the Optimal Model

A relatively satisfactory correlation is observed between the estimated and measured TP concentration (Figure 3). Figure 3e shows the ideal model of the training dataset, which had an R^2, RPD, and RMSEP of 0.630, 1.644, and 0.032 mg/L, respectively, indicating the algorithm could be used to estimate the TP concentration. The scatter plot of the testing dataset revealed that the correlation coefficient and the RMSEP between the estimated and measured TP concentrations were 0.802 and 0.016 mg/L, respectively (Figure 3f). To prove the accuracy and extensibility of the model, formula 6 was applied to the Landsat-8 images acquired on 30 July 2017 2017. After eliminating the sampling sites related to the effects of cloud, land, and aquatic vegetation from the images, satisfactory results were achieved from the optimal algorithm, where the R and RMSEP of the verification set (30 July 2017) reached 0.879 and 0.027 mg/L (Figure 3g).

3.3.3. Temporal and Spatial Distribution of the TP Concentration

Formula 6 was applied to the Landsat-8 images acquired on 27 July and 13 September 2016 to obtain the remote sensing inversion maps of the TP concentration (Figure 4c,f), and the TP concentration distribution conditions on inversion results are in good agreement

with the spatial interpolation results. Based on the TP concentration of inversion results, the water quality for most areas of Taihu Lake belongs to type II and type III, which was mainly distributed in the south and central regions of Taihu Lake, with a value of about 0.00–0.05 mg/L. The high TP concentration, with a value exceeding 0.05 mg/L, was primarily distributed in the northwestern and southeastern regions of type IV, type V, and inferior V water.

3.4. Validation and Extension of the Optimal Model

The optimal models were applied to the Sentinel-2 and Landsat-8 images in 2017 to prove the accuracy and extensibility of the models. Figure 4g,j show the interpolation maps of the TP concentration for August and September 2017 using the IDW method. The interpolation maps showed that the TP concentration was higher, with a value of 0.10–0.20 mg/L, in the northwestern regions and lower, with a value of 0.025–0.10 mg/L, in the other areas. Formula 5 and 6 were applied to the Sentinel-2 image (28 July 2017) and Landsat-8 image (30 July 2017), respectively, to estimate the TP concentration. According to Figure 4h, TP concentration is higher in the northwestern and southeastern parts of Lake Taihu (higher than 0.20 mg/L) and low in the east and south of the lake (less than 0.10 mg/L). The TP concentration derived from the Landsat-8 image (Figure 4i) shows values ranging from 0.10 mg/L to 0.20 mg/L in the northwestern regions and values ranging from 0.025 mg/L to 0.10 mg/L in other areas.

Similarly, the optimal models were applied to the Sentinel-2 image acquired on 27 August 2017, to map the TP concentration of late August 2017. In Figure 4k, the low TP concentration is mainly located in the eastern and southern regions, while the high concentration (higher than 0.20 mg/L) is dominant in northwestern areas.

3.5. Correlation Analysis between TP Concentration and Influencing Factors

To delve further into the factors influencing TP concentration changes in Taihu Lake, the environmental factors (COD_{Mn}, BOD_5, NH_3-N, TN, DO, pH, COD) and meteorological factors (precipitation and temperature) in this study were correlated with the TP concentration, and the correlation analysis is shown in Figure 5. Most of the variables have a favorable relationship with the TP concentration. At the 0.01 level, the relationship between TP and the factors, including COD_{Mn}, BOD_5, NH_3-N, TN, DO, and pH, is significant. At the 0.05 level, the relationship between TP and COD is substantial. COD_{Mn}, BOD_5, and TP had the best association, with correlation coefficients of 0.671 and 0.677. The correlation coefficients between NH_3-N, DO, pH, COD, and TP concentration are 0.468, 0.443, 0.557, and 0.332, respectively. The correlation between TN and TP concentration is poor, with a correlation coefficient of 0.244. The meteorological factors, including precipitation and temperature, do not correlate with TP concentration.
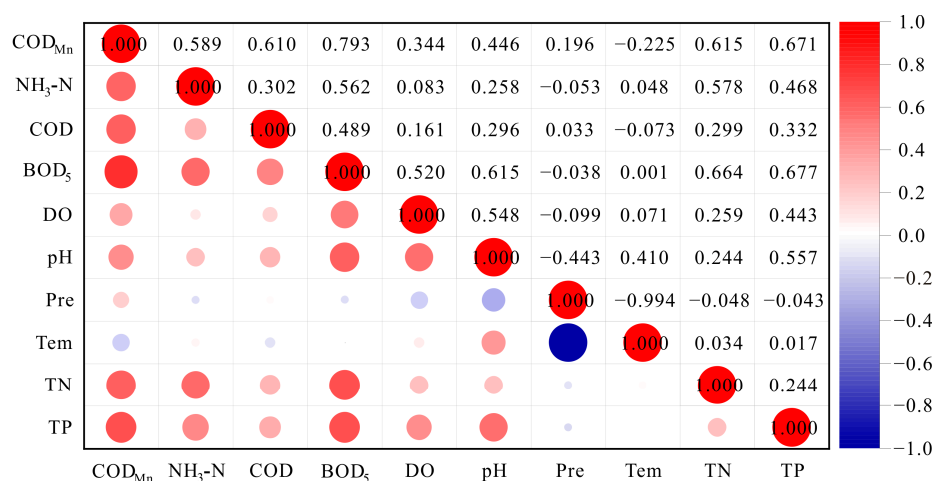

Figure 5. Correlation analysis of environmental and meteorological factors in Taihu Lake.

4. Discussion

4.1. Validation of Model Accuracy and Inversion Results on Sentinel-2 and Landsat-8 Images

Table 2 shows the related studies of TP remote sensing inversion. Multispectral images were extensively used for TP inversion studies [6,8,13,14,26,51,55,56]. Although hyperspectral images usually have a higher accuracy than multispectral images, the low temporal resolution hinders practicability. Landsat TM, ETM+, OLI imagery, and Sentinel-2 were widely used to estimate the TP concentration in freshwater lakes [6,8,13–15,56], and the obtained R^2 mainly ranged from 0.582 to 0.810. The inversion studies of TP concentration using Landsat-8 images were mostly in Taihu Lake ($R^2 = 0.712$), Poyang Lake ($R^2 = 0.582$), Dongting Lake ($R^2 = 0.758$) [6], and Xin'an River ($R^2 = 0.660$) [26], and the study using Sentinel-2 images was mainly in Huaihe River ($R^2 = 0.690$ and 0.810) [14]. This study's models derived from the TP concentration and remote sensing images (Sentinel-2 and Landsat-8) show a higher accuracy and could be used to predict the TP concentration in Taihu Lake. The R^2, RPD, and RMSEP of the PLSR model built using the modeling dataset of Sentinel-2 images were 0.771, 2.086, and 0.023 mg/L, respectively, whereas the validation dataset had a medium R of 0.782 and a small RMSEP of 0.020 mg/L between the predicted and observed TP concentration. The R^2, RPD, and RMSEP training datasets on Landsat-8 were 0.630, 1.644, and 0.032 mg/L, respectively, and the R and RMSEP of the validation dataset were 0.802 and 0.016 mg/L, respectively.

The Landsat-8 and Sentinel-2 images used in this study were applied to verify the model accuracy in Table 2. The predictions of the different models were compared with the measured TP concentration in this study, and the correlation coefficients and root mean square errors were calculated (Table 2). Because the bands of Landsat-8 and Sentinel-2 images are relatively similar, the prediction accuracy based on these two images is similar, which further indicates that the two images are complementary. The prediction accuracy of both Landsat-8 and Sentinel-2 images for Taihu Lake is relatively high compared to other lakes. The model has the best validation results (R = 0.720) [6] with the data collected from Taihu Lake in July and August 2018. The study [6] had the same sampling locations and dates as this study, so the accuracy of predictions was relatively high. This demonstrates that stable regional models can estimate TP concentrations in different years.

Previous studies mainly used the TP concentration of a single time to construct remote sensing models in freshwater lakes [8,13,17,55,57,58], reservoirs [22], and rivers [56], while its applicability was greatly limited by study area and data. Moreover, an analysis of the TP concentration in Taihu Lake (Figure 1) revealed that the TP concentrations were relatively stable in the summer and fall. The suitable water temperature provides better environmental conditions for algae growth [50]. The sources of pollutants discharged from the same area are similar; therefore, the same model can be used for TP monitoring in the same area at different times. In this paper, forty sampling sites from August and September were used to develop the TP estimation model to enhance the time availability in this research. The optimal model applied to Landsat-8 and Sentinel-2 images achieved satisfactory results, with both high and low values in good agreement with the measured values. There was a relatively large error in predicting high values due to the small amount of dissolved phosphorus (DP), which was in agreement with another study [16]. The optimal algorithm based on Landsat-8 images has a better verification effect (RMSEP = 0.027 mg/L, and R = 0.879 for one Landsat-8 image), and the optimal algorithm based on Sentinel-2 images has moderate verification effect (RMSEP = 0.054 mg/L and 0.045 mg/L, and R = 0.771 and 0.787 for two Sentinel-2 images).

Table 2. The comparison of remote sensing monitoring studies of TP concentration.

Reference	Lakes/ Reservoirs/ Rivers	Imagery	Models	Model Accuracy		Verification Accuracy Based on Landsat-8		Verification Accuracy Based on Sentinel-2	
				R^2	RMSE (mg/L)	R	RMSE (mg/L)	R	RMSE (mg/L)
(2006) [55]	Chagan Lake	Landsat TM	$TP = 0.0421*B1 + 0.0162*B2 + 0.0001*B3 - 0.175$	0.653	0.027	0.032	50.029	0.107	59.320
(2010) [56]	Qiantang River	Landsat TM	$\ln(TP) = -21.45\left(\frac{B3}{B2}\right) - 14.42\left(\frac{B1}{B3}\right) + 42.99(B1) + 27.1$	0.770	/	0.260	0.068	0.235	0.0732
(2012) [13]	Taihu Lake	Landsat TM	$TP = -10.11*B1 + 0.85*B2 + 7.34*B3 + 2.04*B4 + 0.17$	0.630	/	0.585	0.062	0.309	0.080
(2014) [8]	Champlain Lake	Landsat ETM+	$TP = 1.84 + 354*B3 - 444*B7$	0.610		0.120	23.219	0.073	24.291
(2014) [57]	Akkulam–Veli Lake	Indian Remote Sensing satellite (IRS P6- LISS III)	$TP = 1.861 - 0.647*B_{red}$	0.500	/	0.121	1.313	0.138	1.736
(2014) [39]	Reservoirs (ECR, MR, GR)	Airborne Hyperspectral	/	0.910	0.040				
(2015) [17]	Western Chaohu Lake	HJ-1A CCD	/	0.989	/				
	Eastern Chaohu Lake		/	0.996	/				
(2017) [26]	Qiandaohu Lake	Landsat-8 OLI	$TP = -0.0203 + 0.0055*e^{(0.67098*((B1+B3+B4)/B2))}$	0.660	0.008	0.445	0.062	0.158	0.069
(2017) [51]	Taihu Lake	GOCI	$TP = 0.2428*e^{-2.554*(B1/B8)}$	0.716		0.650	0.084	0.578	0.084
(2020) [18]	The Pearl River	GF-1 WFV, Measured spectra	$TP = -0.531*B1 - 0.9224*B2 - 5.4182*B3 + 12.638*B4 + 0.4083$	0.759	/	0.575	0.324	0.548	0.351
			$TP = C1*Chl - a + C2*SS + C3*COD_{Mn} + C0$ [1]	0.906	/				
(2020) [14]	Xinyang section of Huaihe River	Sentinel-2A	Back Propagation (BP) neural network model	0.810	0.028				
			Random Forest (RF) model	0.690	0.032				
			Statistical regression model						
(2021) [6]	Poyang Lake	Landsat-8 OLI	$TP = -5.3248*\frac{\ln(B4)}{B4} + 0.0885$	0.758	0.005	0.297	0.058	0.208	0.055
	Dongting Lake		$TP = 0.0038*\left(\frac{B4-B3}{B5}\right)^2 + 0.0146*\frac{B4-B3}{B5} + 0.0772$	0.582	0.004	0.456	0.047	0.352	0.047
	Taihu Lake		$TP = 0.0292*\left(\frac{B5}{B1}\right)^2 + 0.0979*\left(\frac{B5}{B1}\right) + 0.0332$	0.712	0.033	0.720	0.039	0.578	0.044
This paper	Taihu Lake	Landsat-8 OLI	$TP = 0.07082 + 0.00004*(B_1 - B_4) - 0.00007*(B_1 - B_5) - 0.0001*(B_2 - B_3) - 0.00001*(B_2 - B_4) - 0.00009*(B_2 - B_5) + 0.00021*(B_5 - B_6)$	0.630	0.032				
		Sentinel-2A	$TP = 0.0527 + 0.0231*\left(\frac{B_1+B_8}{B_1-B_8}\right) + 0.0039*\left(\frac{B_2+B_8}{B_2-B_8}\right) - 0.0536*\left(\frac{B_3+B_8}{B_3-B_8}\right) + 0.0149*\left(\frac{B_4+B_8}{B_4-B_8}\right)$	0.771	0.023				

[1] C0, C1, C2, C3 are the retrieval coefficients. SS is the suspended solids.

4.2. Stability of the Models

In order to prove the accuracy of the model, the optimal models of formula 5 and formula 6 were verified and applied to the Sentinel-2 and Landsat-8 images in 2017. The scatter plots derived from Sentinel-2 images (Figure 3c,d) show that the model predicted moderately well. There was a relatively large error in predicting high values, which agreed with the inversion results (Figure 4b,e). The reason for the large error in the scatter plot was the small sample points in the verification datasets (Figure 3c,d). The optimal model applied to Landsat-8 image provided better results, with high and low values in good agreement with the measured values (Figure 3g). According to Figure 3g, the optimal algorithm based on Landsat-8 images and in situ data has a better verification effect (RMSEP = 0.27 mg/L, and R = 0.879). Therefore, the model constructed based on the band combinations of Landsat-8 images had a better accuracy than that of Sentinel-2 images.

To better verify the accuracy of the model and the complementarity of the two images, 160 sampling points (Figure 6a) were uniformly laid in Taihu Lake, and the remote sensing inversion values of different images were extracted to draw the scatter plots (Figure 6). Figure 6b,c show the prediction plots of Landsat-8 and Sentinel-2 for July and September 2016, with a relatively small mean square error (RMSE) of about 0.020 mg/L between the estimated values. Figure 6d shows a comparison of the inversion results of the two images in July 2017. The linear fit of the scatter plot is good, and the scatter points are uniformly distributed on both sides of the trend line, with a correlation coefficient of 0.810 and a mean square error of 0.020 mg/L. All these prove that the two images could achieve similar results and are highly complementary.
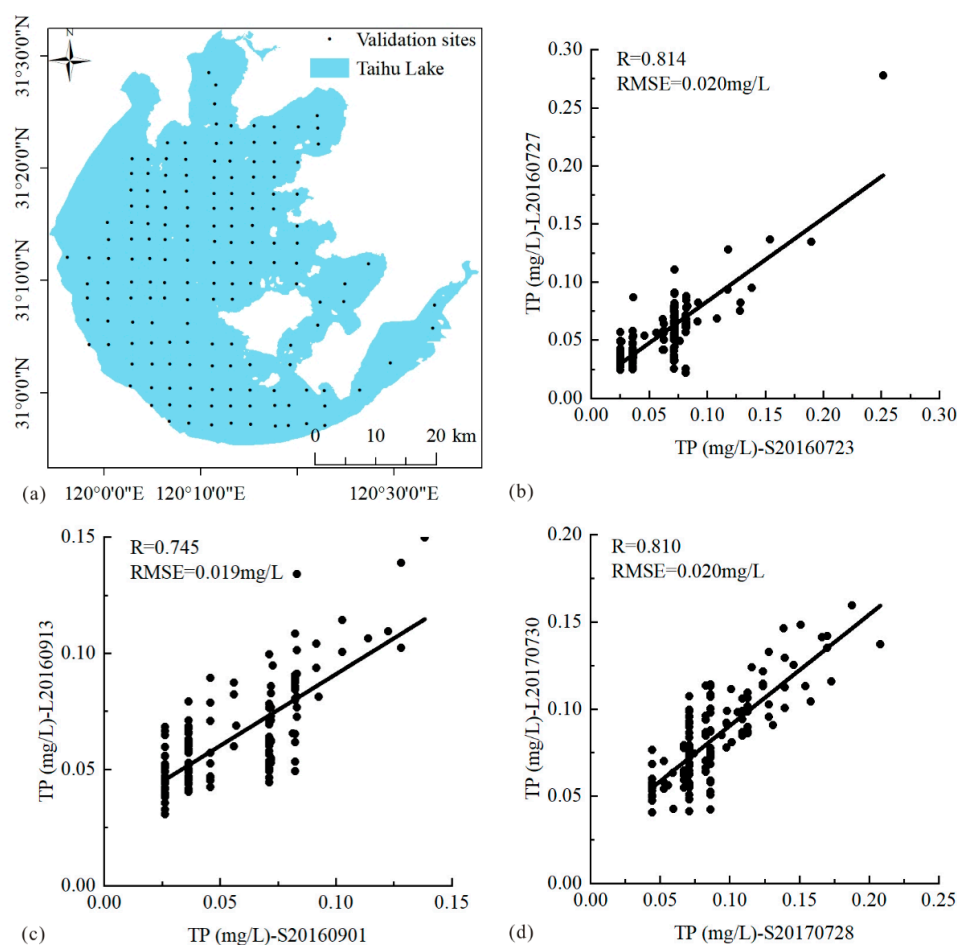

Figure 6. (a) The validation sites in Taihu Lake; **(b–d)** the comparison of the TP results derived from the Landsat-8 and Sentinel-2 images in July (2016), September (2016), and July (2017). (L: Landsat-8; S: Sentinel-2).

A comparison of the inversion results revealed that the distribution of TP concentration in the interpolated and inverse results in August 2017 was consistent, with the high- and low-value areas coinciding well with each other, proving that the model had good prediction results for the 2017 images. According to the distribution of TP concentration in the inverse results (Figure 4k) in September 2017, the TP concentration is lower in eastern and southern regions, which is in agreement with the interpolated map, and the TP concentration is higher (greater than 0.20 mg/L) in northwestern areas, which are different from the interpolated map. In the central part of Taihu Lake, the TP concentration is higher in the interpolated map than it is in the inversed map. By comparing the interpolation plots (Figure 4a,d,g) and Landsat-8 inversion plots (Figure 4c,f,i), we find that Landsat-8 predictions are relatively effective, but the Landsat-8 image has a long revisit period of 16 days and is susceptible to missing images affected by the cloud. There was a lack of available Landsat-8 imagery in September 2017. Without Landsat-8 images, Sentinel-2 concurrent images can be selected for TP concentration monitoring.

4.3. Causes of the TP Concentration Distribution

This study's results showed that high TP concentration values are located in the northwest, north, and southeast regions, while the low values are located in the east and central areas. The water quality in the northwest of Taihu Lake is type V class and inferior V class. Furthermore, the northern region close to the MLH and WLH and the southeast region of Taihu Lake are seriously polluted (Figure 4).

The P in Taihu Lake mainly comes from both internal and external sources. The emissions of farmland drainage and domestic wastewater are the primary external sources. The industrial waste discharges in the Taihu Lake area have been effectively controlled since the implementation of "Operation Zero" in 1998 [52]. P is mainly discharged into Taihu Lake by runoff and drainage processes, and the driving forces are the precipitation and landform. The rainy season of the Taihu Lake region is from May to September, and the TP concentration from May to September was higher than that from November to March (Figure 2). The northwest region is the Wujin District of Changzhou city, a large agricultural area in Jiangsu Province. The pesticide residue problem adversely affected the water quality condition of Wujin District and even the whole Taihu Lake basin [59]. Hence, the TP concentration is higher in the northwest of Taihu Lake. The water pollution in MLH, ZSH, and WLH in the north of Taihu Lake is severe. MLH and ZSH, located in the north of Taihu Lake, where the TP content was more significant than other regions, is one of the most seriously polluted areas, which is consistent with other studies [51,60]. The estuaries of WLH are located in Wuxi city, with a population of more than seven million, and domestic sewage is discharged into Taihu Lake. The topography of Taihu Lake is high in the west and low in the east (Figure 1); most runoff comes from the western mountainous watershed and flows out at the east side of the lake. Therefore, the TP concentration was higher in Taihu Lake's western and northwestern regions [54].

The internal sources are the P release from sediment and the P decomposition of cyanobacteria. The sediment P release is affected by various factors such as light, temperature, pH value, oxygen concentration, biological activity, Microcystis blooms [61], and sediment resuspension [62]. Various factors indirectly affect P release through biological activity. The light has adverse effects on P release. Light gives additional energy to the algal growth [63], and more oxygen is excreted, which is harmful to the anaerobic bacteria and thus limits the cyanobacterial decomposition. Meanwhile, some of the P could be assimilated by the growing algae, which would further reduce the amount of TP concentration in the systems [64]. The temperature could remarkably improve P release, and higher temperatures increase biological activity, thereby promoting the P release from the sediment. The peak of TP concentration generally occurs in July, April, or September (Figure 2). Increased water column stability resulting from increased temperatures is likely to increase cyanobacteria dominance by relieving light limitation [65]. The correlation analysis showed that the pH value is positively correlated with TP concentration (Figure 5),

and the correlation coefficient is 0.446. Higher pH values affect the growth of algae in lakes, especially cyanobacteria, and even lead to the outbreak of blooms [66]. The biological action of the water column promotes the P release. The correlation analysis showed that TP was positively correlated with COD_{Mn} and BOD_5, which were probably related to organic pollution. Cities surround the northwest of Taihu Lake, and household and industrial wastes are discharged into the water through the rivers [67]. Organic matter and nutrients in household and industrial wastes uptake the DO concentration [68]. In the anoxic and anaerobic system, P can be quickly released from the sediments to the overlying water [69]. Taihu Lake is a shallow lake; wind waves exert apparent disturbances and lead to a large amount of sediment resuspension [70]. Studies have shown that wind and wave disturbances can increase P levels in the water column to about 10 times, or even 20–30 times, in the short term. Taihu Lake experiences a monsoon climate, with prevailing southeast winds in the summer and northwest winds in the winter. The monsoon climate will cause the accumulation of pollutants, and the degree of pollution is more severe in the marginal areas than it is in the lake's center [51].

The enclosure culture area in the southeast region of Taihu was affected by residual bait, which increased the nutrients such as P in the lake, causing severe eutrophication of the water in East Taihu Lake. The TP concentration in the marginal zone of Taihu Lake is affected by drainages of domestic wastewater because Taihu Lake is surrounded by five cities [67].

5. Conclusions

This paper predicts the spatial distribution of TP concentration using remote sensing and measured data to monitor the water quality of Taihu Lake. Sentinel-2 and Landsat-8 images acquired in 2016 were used for modeling and inversion TP concentration. It was found that the band combinations of $(B_1 + B_8)/(B_1 - B_8)$, $(B_2 + B_8)/(B_2 - B_8)$, $(B_3 + B_8)/(B_3 - B_8)$, and $(B_4 + B_8)/(B_4 - B_8)$ of the Sentinel-2 images and the band combinations of B_1-B_4, B_1-B_5, B_2-B_3, B_2-B_4, B_2-B_5, and B_5-B_6 of the Landsat-8 images had a satisfactory effect. The optimal model was applied to the remote sensing images of 2016 and 2017 to obtain the inversion maps, which were comparable with the interpolation results. Therefore, it can be concluded that it is feasible to monitor the TP concentration using high-resolution Landsat-8 and Sentinel-2 multispectral imagery. The inversion maps based on Landsat-8 and Sentinel-2 images are in good agreement, indicating that two remote sensing images can be complementary for water quality monitoring. The inversion results analyzed the spatial variation in TP concentration. Comparing all the inversion maps, pollution in the northwestern part of Taihu Lake was more severe, whereas the TP concentration in the south and east of the lake was relatively low. This method can potentially allow for monitoring the TP concentration on a large scale and contributes to a more targeted treatment of water pollution.

Author Contributions: Y.L.: Investigation, Writing—Original draft preparation, Methodology, Formal analysis, Validation, Data Curation. F.Y.: Methodology, Writing—Original draft preparation, Conceptualization, Supervision, Funding acquisition. L.L.: Conceptualization, Resources, Writing—Review & editing, Supervision, Project administration. D.X.: Supervision, Writing—Review & editing, Funding acquisition. Y.Z.: Investigation, Supervision. T.A.: Writing—Review & editing. All authors have read and agreed to the published version of the manuscript.

References

1. Yuwei Chen, C.F.; Teubner, K.; Dokulil, M. Changes of nutrients and phytoplankton chlorophyll-a in a large shallow lake, Taihu, China an 8-year investigation. *Hydrobiologia* **2003**, *506*, 273–279. [CrossRef]
2. Qin, B.; Zhu, G.; Zhang, L.; Luo, L.; Gao, G.; Gu, B. Estimation of internal nutrient release in large shallow Lake Taihu, China. *Sci. China Ser. D Earth Sci.* **2006**, *49* (Suppl. S1), 38–50. [CrossRef]
3. Gao, G.; Zhu, G.; Qin, B.; Chen, J.; Wang, K. Alkaline phosphatase activity and the phosphorus mineralization rate of Lake Taihu. *Sci. China Ser. D Earth Sci.* **2006**, *49* (Suppl. S1), 176–185. [CrossRef]
4. Wu, Z.; Wang, X.; Chen, Y.; Cai, Y.; Deng, J. Assessing river water quality using water quality index in Lake Taihu Basin, China. *Sci. Total Environ.* **2018**, *612*, 914–920. [CrossRef]
5. Liu, L.; Dong, Y.; Kong, M.; Zhou, J.; Zhao, H.; Wang, Y.; Zhang, M.; Wang, Z. Towards the comprehensive water quality control in Lake Taihu: Correlating chlorphyll a and water quality parameters with generalized additive model. *Sci. Total Environ.* **2020**, *705*, 135993. [CrossRef] [PubMed]
6. Shang, W.; Jin, S.; He, Y.; Zhang, Y.; Li, J. Spatial–Temporal Variations of Total Nitrogen and Phosphorus in Poyang, Dongting and Taihu Lakes from Landsat-8 Data. *Water* **2021**, *13*, 1704. [CrossRef]
7. Song, K.; Li, L.; Wang, Z.; Liu, D.; Zhang, B.; Xu, J.; Du, J.; Li, L.; Li, S.; Wang, Y. Retrieval of total suspended matter (TSM) and chlorophyll-a (Chl-a) concentration from remote-sensing data for drinking water resources. *Environ. Monit. Assess.* **2011**, *184*, 1449–1470. [CrossRef]
8. Isenstein, E.M.; Park, M.-H. Assessment of nutrient distributions in Lake Champlain using satellite remote sensing. *J. Environ. Sci.* **2014**, *26*, 1831–1836. [CrossRef]
9. Li, L.; Sengpiel, R.E.; Pascual, D.L.; Tedesco, L.P.; Wilson, J.S.; Soyeux, E. Using hyperspectral remote sensing to estimate chlorophyll-*a* and phycocyanin in a mesotrophic reservoir. *Int. J. Remote Sens.* **2010**, *31*, 4147–4162. [CrossRef]
10. Choi, J.-K.; Park, Y.J.; Ahn, J.H.; Lim, H.-S.; Eom, J.; Ryu, J.-H. GOCI, the world's first geostationary ocean color observation satellite, for the monitoring of temporal variability in coastal water turbidity. *J. Geophys. Res. Oceans* **2012**, *117*, C9. [CrossRef]
11. Kim, H.-C.; Son, S.; Kim, Y.H.; Khim, J.S.; Nam, J.; Chang, W.K.; Lee, J.-H.; Lee, C.-H.; Ryu, J. Remote sensing and water quality indicators in the Korean West coast: Spatio-temporal structures of MODIS-derived chlorophyll-a and total suspended solids. *Mar. Pollut. Bull.* **2017**, *121*, 425–434. [CrossRef] [PubMed]
12. Kim, Y.H.; Im, J.; Ha, H.K.; Choi, J.-K.; Ha, S. Machine learning approaches to coastal water quality monitoring using GOCI satellite data. *GIScience Remote Sens.* **2014**, *51*, 158–174. [CrossRef]
13. Chen, J.; Quan, W. Using Landsat/TM Imagery to Estimate Nitrogen and Phosphorus Concentration in Taihu Lake, China. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 273–280. [CrossRef]
14. Huangfu, K.; Li, J.; Zhang, X.; Zhang, J.; Cui, H.; Sun, Q. Remote Estimation of Water Quality Parameters of Medium- and Small-Sized Inland Rivers Using Sentinel-2 Imagery. *Water* **2020**, *12*, 3124. [CrossRef]
15. Guo, H.; Huang, J.J.; Chen, B.; Guo, X.; Singh, V.P. A machine learning-based strategy for estimating non-optically active water quality parameters using Sentinel-2 imagery. *Int. J. Remote Sens.* **2020**, *42*, 1841–1866. [CrossRef]
16. Du, C.; Wang, Q.; Li, Y.; Lyu, H.; Zhu, L.; Zheng, Z.; Wen, S.; Liu, G.; Guo, Y. Estimation of total phosphorus concentration using a water classification method in inland water. *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *71*, 29–42. [CrossRef]
17. Gao, Y.; Gao, J.; Yin, H.; Liu, C.; Xia, T.; Wang, J.; Huang, Q. Remote sensing estimation of the total phosphorus concentration in a large lake using band combinations and regional multivariate statistical modeling techniques. *J. Environ. Manag.* **2015**, *151*, 33–43. [CrossRef]
18. Lu, S.; Deng, R.; Liang, Y.; Xiong, L.; Ai, X.; Qin, Y. Remote Sensing Retrieval of Total Phosphorus in the Pearl River Channels Based on the GF-1 Remote Sensing Data. *Remote Sens.* **2020**, *12*, 1420. [CrossRef]
19. Huang, C.; Guo, Y.; Yang, H.; Li, Y.; Zou, J.; Zhang, M.; Lyu, H.; Zhu, A.; Huang, T. Using Remote Sensing to Track Variation in Phosphorus and Its Interaction with Chlorophyll-a and Suspended Sediment. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 4171–4180. [CrossRef]
20. Xiong, J.; Lin, C.; Ma, R.; Cao, Z. Remote Sensing Estimation of Lake Total Phosphorus Concentration Based on MODIS: A Case Study of Lake Hongze. *Remote Sens.* **2019**, *11*, 2068. [CrossRef]
21. Ding, C.; Pu, F.; Li, C.; Xu, X.; Zou, T.; Li, X. Combining Artificial Neural Networks with Causal Inference for Total Phosphorus Concentration Estimation and Sensitive Spectral Bands Exploration Using MODIS. *Water* **2020**, *12*, 2372. [CrossRef]

22. Song, K.; Li, L.; Li, S.; Tedesco, L.; Hall, B.; Li, L. Hyperspectral Remote Sensing of Total Phosphorus (TP) in Three Central Indiana Water Supply Reservoirs. *Water Air Soil Pollut.* **2012**, *223*, 1481–1502. [CrossRef]
23. Chen, J.; Zhu, W.-N.; Tian, Y.Q.; Yu, Q. Estimation of Colored Dissolved Organic Matter from Landsat-8 Imagery for Complex Inland Water: Case Study of Lake Huron. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 2201–2212. [CrossRef]
24. Chen, J.; Zhu, W.; Tian, Y.Q.; Yu, Q.; Zheng, Y.; Huang, L. Remote estimation of colored dissolved organic matter and chlorophyll-a in Lake Huron using Sentinel-2 measurements. *J. Appl. Remote Sens.* **2017**, *11*, 036007. [CrossRef]
25. Kuhn, C.; De Matos Valerio, A.; Ward, N.; Loken, L.; Sawakuchi, H.O.; Kampel, M.; Richey, J.; Stadler, P.; Crawford, J.; Striegl, R.; et al. Performance of Landsat-8 and Sentinel-2 surface reflectance products for river remote sensing retrievals of chlorophyll-a and turbidity. *Remote Sens. Environ.* **2019**, *224*, 104–118. [CrossRef]
26. Li, Y.; Zhang, Y.; Shi, K.; Zhu, G.; Zhou, Y.; Zhang, Y.; Guo, Y. Monitoring spatiotemporal variations in nutrients in a large drinking water reservoir and their relationships with hydrological and meteorological conditions based on Landsat 8 imagery. *Sci. Total Environ.* **2017**, *599*, 1705–1717. [CrossRef] [PubMed]
27. Li, P.; Ke, Y.; Bai, J.; Zhang, S.; Chen, M.; Zhou, D. Spatiotemporal dynamics of suspended particulate matter in the Yellow River Estuary, China during the past two decades based on time-series Landsat and Sentinel-2 data. *Mar. Pollut. Bull.* **2019**, *149*, 110518. [CrossRef]
28. Ambrose-Igho, G.; Seyoum, W.M.; Perry, W.L.; O'Reilly, C.M. Spatiotemporal Analysis of Water Quality Indicators in Small Lakes Using Sentinel-2 Satellite Data: Lake Bloomington and Evergreen Lake, Central Illinois, USA. *Environ. Process.* **2021**, *8*, 637–660. [CrossRef]
29. Niroumand-Jadidi, M.; Bovolo, F.; Bruzzone, L. Water Quality Retrieval from PRISMA Hyperspectral Images: First Experience in a Turbid Lake and Comparison with Sentinel-2. *Remote Sens.* **2020**, *12*, 3984. [CrossRef]
30. Delegido, J.; Verrelst, J.; Alonso, L.; Moreno, J. Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors* **2011**, *11*, 7063–7081. [CrossRef]
31. Kupssinskü, L.S.; Guimarães, T.T.; de Souza, E.M.; Zanotta, D.C.; Veronez, M.R.; Gonzaga, L., Jr.; Mauad, F.F. A Method for Chlorophyll-a and Suspended Solids Prediction through Remote Sensing and Machine Learning. *Sensors* **2020**, *20*, 2125. [CrossRef] [PubMed]
32. Bresciani, M.; Cazzaniga, I.; Austoni, M.; Sforzi, T.; Buzzi, F.; Morabito, G.; Giardino, C. Mapping phytoplankton blooms in deep subalpine lakes from Sentinel-2A and Landsat-8. *Hydrobiologia* **2018**, *824*, 197–214. [CrossRef]
33. Toming, K.; Kutser, T.; Laas, A.; Sepp, M.; Paavel, B.; Nõges, T. First Experiences in Mapping Lake Water Quality Parameters with Sentinel-2 MSI Imagery. *Remote Sens.* **2016**, *8*, 640. [CrossRef]
34. Nazeer, M.; Bilal, M.; Alsahli, M.M.M.; Shahzad, M.I.; Waqas, A. Evaluation of Empirical and Machine Learning Algorithms for Estimation of Coastal Water Quality Parameters. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 360. [CrossRef]
35. Du, C.; Li, Y.; Lyu, H.; Liu, N.; Zheng, Z.; Li, Y. Remote estimation of total phosphorus concentration in the Taihu Lake using a semi-analytical model. *Int. J. Remote Sens.* **2020**, *41*, 7993–8013. [CrossRef]
36. 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.* **2014**, *45*, 101–166. [CrossRef]
37. Verrelst, J.; Muñoz, J.; Alonso, L.; Delegido, J.; Rivera, J.P.; Camps-Valls, G.; Moreno, J. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* **2012**, *118*, 127–139. [CrossRef]
38. Wold, H. Estimation of principal components and related models by iterative least squares. *Multivar. Anal.* **1966**, 391–420.
39. Song, K.; Li, L.; Tedesco, L.; Li, S.; Shi, K.; Hall, B. Remote Estimation of Nutrients for a Drinking Water Source Through Adaptive Modeling. *Water Resour. Manag.* **2014**, *28*, 2563–2581. [CrossRef]
40. Nakar, A.; Schmilovitch, Z.; Vaizel-Ohayon, D.; Kroupitski, Y.; Borisover, M.; Saldinger, S.S. Quantification of bacteria in water using PLS analysis of emission spectra of fluorescence and excitation-emission matrices. *Water Res.* **2020**, *169*, 115197. [CrossRef]
41. Xiong, J.; Lin, C.; Cao, Z.; Hu, M.; Xue, K.; Chen, X.; Ma, R. Development of remote sensing algorithm for total phosphorus concentration in eutrophic lakes: Conventional or machine learning? *Water Res.* **2022**, *215*, 118213. [CrossRef] [PubMed]
42. Duan, H.; Ma, R.; Zhang, Y.; Loiselle, S.A. Are algal blooms occurring later in Lake Taihu? Climate local effects outcompete mitigation prevention. *J. Plankton Res.* **2014**, *36*, 866–871. [CrossRef]
43. Xu, R.; Pang, Y.; Hu, Z.; Zhu, T.; Kaisam, J.P. Influence of water diversion on spatial and temporal distribution of flow field and total phosphorus (TP) concentration field in Taihu Lake. *Water Supply* **2020**, *20*, 1059–1071. [CrossRef]
44. Roy, D.P.; Wulder, M.A.; Loveland, T.R.; Woodcock, C.E.; Allen, R.G.; Anderson, M.C.; Helder, D.; Irons, J.R.; Johnson, D.M.; Kennedy, R.; et al. Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens. Environ.* **2014**, *145*, 154–172. [CrossRef]
45. Muchsin, F.; Dirghayu, D.; Prasasti, I.; Rahayu, M.I.; Fibriawati, L.; Pradono, K.A.; Hendayani; Mahatmanto, B. Comparison of atmospheric correction models: FLAASH and 6S code and their impact on vegetation indices (case study: Paddy field in Subang District, West Java). *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *280*, 012034. [CrossRef]
46. Pahlevan, N.; Mangin, A.; Balasubramanian, S.V.; 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]
47. 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]

48. Vohland, M.; Besold, J.; Hill, J.; Fründ, H.-C. Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy. *Geoderma* **2011**, *166*, 198–205. [CrossRef]

49. Gogtay, N.J.; Thatte, U.M. Principles of correlation analysis. *J. Assoc. Physicians India* **2017**, *65*, 78–81.

50. Nalley, J.O.; O'Donnell, D.R.; Litchman, E. Temperature effects on growth rates and fatty acid content in freshwater algae and cyanobacteria. *Algal Res.* **2018**, *35*, 500–507. [CrossRef]

51. Du, C.; Li, Y.; Wang, Q.; Liu, G.; Zheng, Z.; Mu, M.; Li, Y. Tempo-spatial dynamics of water quality and its response to river flow in estuary of Taihu Lake based on GOCI imagery. *Environ. Sci. Pollut. Res.* **2017**, *24*, 28079–28101. [CrossRef] [PubMed]

52. Chen, X.; Wo, F.; Chen, C.; Fang, K. Seasonal changes in the concentrations of nitrogen and phosphorus in farmland drainage and groundwater of the Taihu Lake region of China. *Environ. Monit. Assess.* **2009**, *169*, 159–168. [CrossRef] [PubMed]

53. Haoyun, W.; Genghua, J.; Bin, X.; Yanting, S.; Xiaoqing, Z. Analysis of variation and driving factors of total phosphorus in Lake Taihu, 1980-2020. *J. Lake Sci.* **2021**, *33*, 974–991. [CrossRef]

54. Kong, M.; Chao, J.; Zhuang, W.; Wang, P.; Wang, C.; Hou, J.; Wu, Z.; Wang, L.; Gao, G.; Wang, Y. Spatial and Temporal Distribution of Particulate Phosphorus and Their Correlation with Environmental Factors in a Shallow Eutrophic Chinese Lake (Lake Taihu). *Int. J. Environ. Res. Public Heal* **2018**, *15*, 2355. [CrossRef] [PubMed]

55. Song, K.; Zhang, B.; Wang, Z.; Li, F.; Duan, H.; Guo, Y. Water TOC and TP concentration estimation using Landsat TM data with empirical algorithms in Chagan lake, China. In Proceedings of the IEEE International Conference, Geosciences and Remote Sensing Symposium, IGRASS, Denver, CO, USA, 31 July–4 August 2006.

56. Wu, C.; Wu, J.; Qi, J.; Zhang, L.; Huang, H.; Lou, L.; Chen, Y. Empirical estimation of total phosphorus concentration in the mainstream of the Qiantang River in China using Landsat TM data. *Int. J. Remote Sens.* **2010**, *31*, 2309–2324. [CrossRef]

57. Moses, S.A.; Janaki, L.; Joseph, S.; Kandathil, R.K. Determining the spatial variation of phosphorus in a lake system using remote sensing techniques. *Lakes Reserv. Sci. Policy Manag. Sustain. Use* **2014**, *19*, 24–36.

58. Sun, D.; Qiu, Z.; Li, Y.; Shi, K.; Gong, S. Detection of Total Phosphorus Concentrations of Turbid Inland Waters Using a Remote Sensing Method. *Water Air Soil Pollut.* **2014**, *225*, 1953. [CrossRef]

59. Zhou, Y.; Wu, J.; Wang, B.; Duan, L.; Zhang, Y.; Zhao, W.; Wang, F.; Sui, Q.; Chen, Z.; Xu, D.; et al. Occurrence, source and ecotoxicological risk assessment of pesticides in surface water of Wujin District (northwest of Taihu Lake), China. *Environ. Pollut.* **2020**, *265*, 114953. [CrossRef]

60. Yan, C.; Che, F.; Zeng, L.; Wang, Z.; Du, M.; Wei, Q.; Wang, Z.; Wang, D.; Zhen, Z. Spatial and seasonal changes of arsenic species in Lake Taihu in relation to eutrophication. *Sci. Total Environ.* **2016**, *563*, 496–505. [CrossRef]

61. Jiang, X.; Jin, X.; Yao, Y.; Li, L.; Wu, F. Effects of biological activity, light, temperature and oxygen on phosphorus release processes at the sediment and water interface of Taihu Lake, China. *Water Res.* **2008**, *42*, 2251–2259. [CrossRef]

62. Cyr, H.; McCabe, S.K.; Nürnberg, G.K. Phosphorus sorption experiments and the potential for internal phosphorus loading in littoral areas of a stratified lake. *Water Res.* **2009**, *43*, 1654–1666. [CrossRef] [PubMed]

63. Ahmadi, A.; Riahi, H.; Noori, M. Studies of the effects of environmental factors on the seasonal change of phytoplankton population in municipal waste water stabilization ponds. *Toxicol. Environ. Chem.* **2005**, *87*, 543–550. [CrossRef]

64. Chuai, X.; Ding, W.; Chen, X.; Wang, X.; Miao, A.; Xi, B.; He, L.; Yang, L. Phosphorus release from cyanobacterial blooms in Meiliang Bay of Lake Taihu, China. *Ecol. Eng.* **2011**, *37*, 842–849. [CrossRef]

65. Paerl, H.W.; Huisman, J. Climate. Blooms like it hot. *Science* **2008**, *320*, 57–58. [CrossRef] [PubMed]

66. Hörnström, E. Phytoplankton in 63 limed lakes in comparison with the distribution in 500 untreated lakes with varying pH. *Hydrobiologia* **2002**, *470*, 115–126. [CrossRef]

67. Xu, J.; Jin, G.; Mo, Y.; Tang, H.; Li, L. Assessing Anthropogenic Impacts on Chemical and Biochemical Oxygen Demand in Different Spatial Scales with Bayesian Networks. *Water* **2020**, *12*, 246. [CrossRef]

68. Shukla, J.; Misra, A.; Chandra, P. Mathematical modeling and analysis of the depletion of dissolved oxygen in eutrophied water bodies affected by organic pollutants. *Nonlinear Anal. Real World Appl.* **2008**, *9*, 1851–1865. [CrossRef]

69. Gomez, E.; Durillon, C.; Rofes, G.; Picot, B. Phosphate adsorption and release from sediments of brackish lagoons: pH, O_2 and loading influence. *Water Res.* **1999**, *33*, 2437–2447. [CrossRef]

70. Lu, Y.; Li, S.; Zuo, L.; Liu, H.; Roelvink, J. Advances in sediment transport under combined action of waves and currents. *Int. J. Sediment Res.* **2015**, *30*, 351–360. [CrossRef]