

Article

Water Quality Index (WQI) as a Potential Proxy for Remote Sensing Evaluation of Water Quality in Arid Areas

Fei Zhang ^{1,*}, Ngai Weng Chan ², Changjiang Liu ¹, Xiaoping Wang ¹, Jingchao Shi ³, Hsiang-Te Kung ³, Xinguo Li ⁴, Tao Guo ⁵, Weiwei Wang ¹ and Naixin Cao ¹

- ¹ Key Laboratory of Wisdom City and Environment Modeling of Resources and Environmental Science College, Xinjiang University, Urumqi 830046, China; liuchangjiang3s@163.com (C.L.); wxp4911@163.com (X.W.); wangweiwei0820@163.com (W.W.); cnx19971214@163.com (N.C.)
- ² School of Humanities, Universiti Sains Malaysia, George Town 11800, Malaysia; nwchan1@gmail.com
- ³ Departments of Earth Sciences, The University of Memphis, Memphis, TN 38152, USA; jshi@memphis.edu (J.S.); hkung@memphis.edu (H.-T.K.)
- ⁴ School of Geography and Tourism, Xinjiang Normal University, Urumqi 830046, China; onlinelxg@163.com
- ⁵ Institute of Remote Sensing and Digital Agriculture, Sichuan Academy of Agricultural Sciences, Chengdu 610066, China; guotao3s@163.com
- * Correspondence: zhangfei3s@xju.edu.cn or zhangfei3s@163.com

Abstract: Water Resource Sustainability Management plays a vitally important role in ensuring sustainable development, especially in water-stressed arid regions throughout the world. In order to achieve sustainable development, it is necessary to study and monitor the water quality in the arid region of Central Asia, an area that is increasingly affected by climate change. In recent decades, the rapid deterioration of water quality in the Ebinur Lake basin in Xinjiang (China) has severely threatened sustainable economic development. This study selected the Ebinur Lake basin as the study target, with the purpose of revealing the response between the water quality index and water body reflectivity, and to describe the relationship between the water quality index and water reflectivity. The methodology employed remote sensing techniques that establish a water quality index monitoring model to monitor water quality. The results of our study include: (1) the Water Quality Index (WQI) that was used to evaluate the water environment in Ebinur Lake indicates a lower water quality of Ebinur Lake, with a WQI value as high as 4000; (2) an introduction of the spectral derivative method that realizes the extraction of spectral information from a water body to better mine the information of spectral data through remote sensing, and the results also prove that the spectral derivative method can improve the relationship between the water body spectral and WQI, whereby R^2 is 0.6 at the most sensitive wavelengths; (3) the correlation between the spectral sensitivity index and WQI was greater than 0.6 at the significance level of 0.01 when multi-source spectral data were integrated with the spectral index (DI, RI and NDI) and fluorescence baseline; and (4) the distribution map of WQI in Ebinur Lake was obtained by the optimal model, which was constructed based on the third derivative data of Sentinel 2 data. We concluded that the water quality in the northwest of Ebinur Lake was the lowest in the region. In conclusion, we found that remote sensing techniques were highly effective and laid a foundation for water quality detection in arid areas.

Keywords: Water Quality Index (WQI); Ebinur Lake; remote sensing



Citation: Zhang, F.; Chan, N.W.; Liu, C.; Wang, X.; Shi, J.; Kung, H.-T.; Li, X.; Guo, T.; Wang, W.; Cao, N. Water Quality Index (WQI) as a Potential Proxy for Remote Sensing Evaluation of Water Quality in Arid Areas. *Water* **2021**, *13*, 3250. <https://doi.org/10.3390/w13223250>

Academic Editor: Dimitrios E. Alexakis

Received: 25 September 2021
Accepted: 14 November 2021
Published: 17 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Water problems can be a great barrier to economic development in any corner of the world [1–3], especially in such arid regions as Xinjiang, China, where water shortages (and other water issues) aggravate ecological environment deterioration. Therefore, studying and monitoring water quality is very important to reduce the potential negative impacts on the ecological environment in Xinjiang. However, traditional water quality monitoring methods are time-consuming, cumbersome, and limited to a small scale. Therefore, they

can no longer meet the needs of water quality monitoring in terms of speed, large areas, or a long time series. In order to have more accurate estimates, new data sources and new methods need to be introduced in the monitoring of comprehensive water quality indicators [4,5]. The development of multi-source observation and the monitoring of remote sensing technology increasingly brings huge opportunities for speedy, high-precision water environment monitoring and evaluation over large areas.

Satellite remote sensing technology has developed very quickly since 1970. Consequently, more water resource researchers started to apply remote sensing technology in their research, and the water quality within remote sensing monitoring mechanisms has also gradually improved. In recent years, remote sensing satellites have been widely used to observe pollutants in rivers and lakes. As a result, the detectable types of pollutants retrieved from satellite images have greatly increased, and the inversion accuracy has been further improved, as well [6].

Remote sensing applied to water quality monitoring is mainly used to map the water quality indexes of rivers and lakes through the relationship between water quality indexes and spectral data with satellite image data such as Landsat, MODIS, ENVISAT, and SPOT data [6–8]. However, the spatial resolution of the above remote sensing data is greater than 10 m, which makes it difficult to meet monitoring requirements. Only a few water quality parameters can be monitored by these remote sensing data, such as Chlorophyll, SS, NTU, and CDOM. [4,9,10]. Other chemical indicators of water quality, such as COD, BOD₅, TN, TP, NH₃-N, DO, etc., cannot be directly monitored by remote sensing. The indirect monitoring accuracy is low, and the mechanism is unclear. Hence, introducing a new technology that makes up for the deficiency of remote sensing water quality monitoring is essential.

In fact, the process of water pollution follows a nonlinear regression that fluctuates with many factors, and the accuracy of the water quality inversion result is limited by the traditional linear inversion model [11]. However, machine learning has a good nonlinear approximation ability, and the application of machine learning in water quality monitoring provides a new idea to improve the accuracy of water quality monitoring. Alves simplified the input variables of the feed forward neural network through principal component analysis, thus accurately inverting the water quality index (WQI) [12]. Gogu proved that there is a good potential in using a neural network to invert the salt content of river water through experiments [13]. Wang [11] estimated the WQI of water quality in the Ebinur Lake basin based on the support vector machine (SVR) model by using near-surface spectroscopy technology, and found that the nonlinear model has great potential in water quality observation.

Although the water quality parameter estimation model provides relatively highly accurate data, the result is uncertain due to the complex and changeable water environment. The reason is that the water spectrum shows the entire water environment rather than a single water quality parameter. Many scholars have developed a single water quality parameter estimation model based on water spectral data [14–16]. Therefore, the estimation model of individual water parameters introduces a certain degree of uncertainty. At this point, the establishment of the water quality index reflecting the whole water environment to evaluate the whole water environment is necessary. Moreover, a good water quality evaluation method should not only accurately reflect the spatial change of the water quality but also conveniently monitor the water quality level. Data on the Water Quality Index (WQI) is compiled by the Ministry of Water Resources and the Water Environment Monitoring and Evaluation Center to evaluate the quality of drinking water [17,18]. The WQI was originally proposed by Horton and Brown [19,20]. Scholars have devised various methods to calculate the water quality Index (WQI) [21,22], which is a mathematical tool of converting large amounts of water quality data into a single value that represents the water environment and reflects the overall water quality level [23]. However, it is impossible to identify the temporal and spatial variation of water quality, which is crucial for the

comprehensive evaluation and management of water quality, even though the WQI method can provide reasonable accuracy of the water quality of a single sample.

In this paper, the relationship between the water quality index, water optical characteristics, and water reflectance is quantitatively analyzed. The specific research objectives in this paper include: (1) to better mine the information of remote sensing data by using a series of technologies, such as the remote sensing image differential algorithm, which are introduced to realize the extraction of water remote sensing information; (2) to construct a remote sensing spectral index (DI, RI and NDI) and fluorescence baseline height for monitoring water quality in arid areas; and (3) to establish a WQI model based on machine learning technology (particle swarm optimization algorithm) to achieve water quality monitoring. This study will provide an effective method for rapid, quantitative, and sustainable water quality management in arid areas, as well as a typical example for ecological conservation in arid areas, and it will also effectively contribute to the health of the ecological environment in arid areas.

2. Study Area

The Ebinur Lake watershed (43°38′–45°52′ N and 79°53′–85°02′ E) is located in northwest Xinjiang, China (Figure 1). The study area is 50,621 km², comprising Bortala River Valley, Jinghe oasis, Wusu Oasis, Dandagai desert, and the Mutetaer desert zone of the lower reaches of the Akeqisu-Kuitun River. The Ebinur Lake is in the lowest elevation of the watershed and is the largest saltwater lake in Xinjiang. It has all the typical characteristics that all other lakes do in the arid region of Central Asia. The area experiences a typical arid continental climate in the middle temperate zone and is characterized by drought, low rainfall, drastic temperature variations, and severe soil salinization. The average lake depth is merely 1.4–1.6 m, with a water density of about 1.079 g/cm³, pH 8.49, and mineralization of 112.4 g/L. The watershed is one of the key areas of China's Silk Road Economic Belt, and can be divided into three sub-basins, namely, the Jinghe River basin, Boltala River basin, and Kuitun River basin. The Ebinur Lake basin consists of a varied landscape of mountain, desert, and oasis, where land is mainly use for agricultural. The annual average temperature is 7.2 °C, with the highest 9.1 °C and the lowest 5.3 °C. The annual extreme high and low temperature is 41 °C and −34.7 °C, respectively. The annual average precipitation is only 149 mm, but the potential evapotranspiration reaches up to 2281 mm.

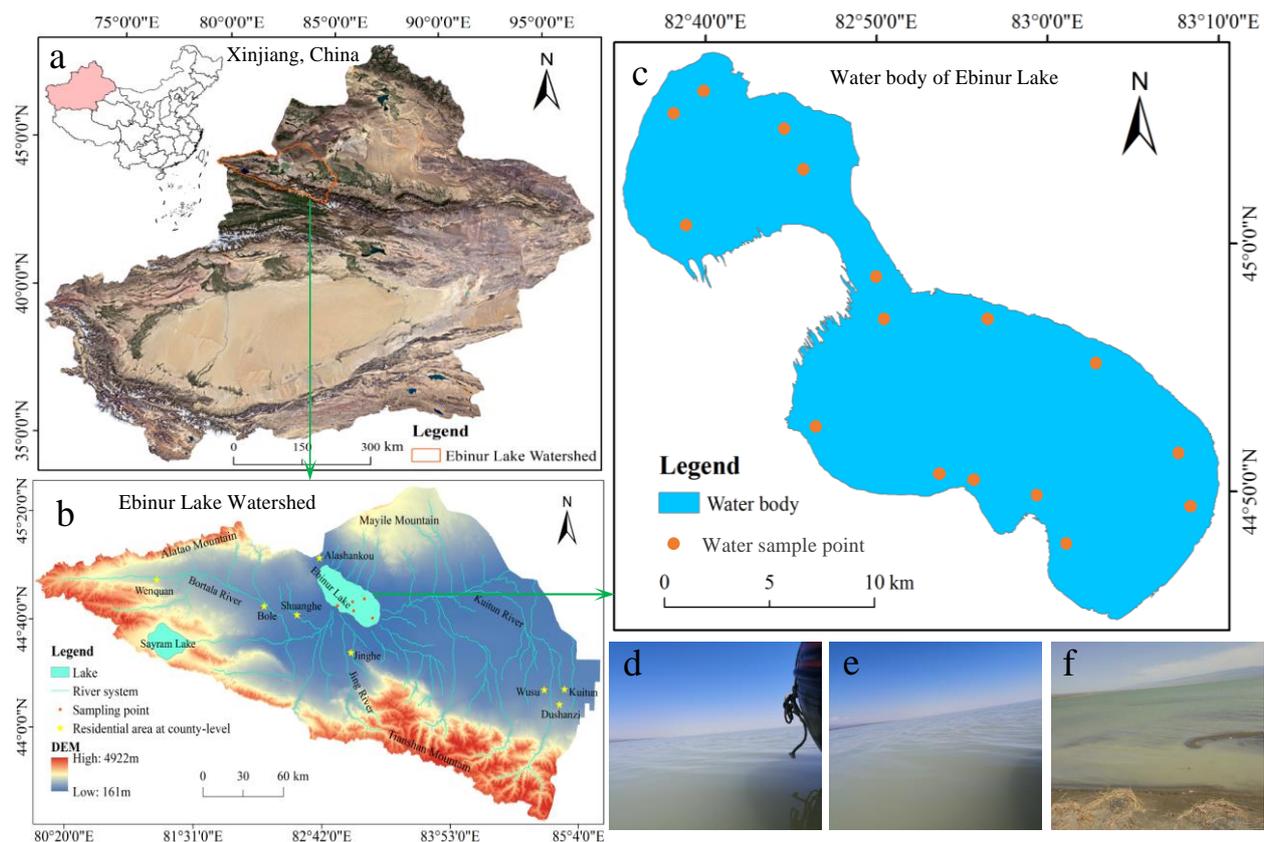


Figure 1. The study area: (a) The Xinjiang Uyghur Autonomous Region in northwestern China; (b) The Ebinur Lake Watershed with elevation and drainage information; (c) The water body and 16 sampling points at Ebinur lake extracted on October 2017; (d) A typical view of Ebinur Lake with sunny weather; (e) Ebinur Lake surface water landscape; and (f) The unique visual effect of Ebinur Lake.

3. Data and Methods

3.1. Data Collection

Water Quality Data Collection

The field investigations and water quality sample collections for this study were conducted in October 2015. They were integrated with the main body of the experiment, which included three parts of water sampling, including water surface spectral measurement, GPS record location, and other auxiliary information. Spectroscopy was measured by a FieldSpec® ProFR (wavelength range: 350–2500 nm), a portable ASD spectrometer (Analytical Spectral Devices, Boulder, CO, USA). Water samples were sent to the laboratory for analysis within a specified time frame.

The researchers collected a total of 16 water samples. For each sampling point, a water sample collector was used to collect water samples at 0.5 cm depth, just below the water surface, with 1000 mL of water samples collected at each sampling point. Samples were stored in Teflon plastic bottles (for standard and easy transportation). Teflon plastic bottles were washed several times with collected water before each collection. After the samples were collected, they were immediately put into the benzene board incubator with ice and transported to the laboratory where the water quality index was determined as soon as possible.

3.2. Remote Sensing Data Collection

The European Space Agency (ESA) recently launched the Copernicus Project, which is expected to improve the monitoring of forest conditions and land use, as well as enhance disaster management through the launch of Sentinel satellites. The Sentinel-2 Satellite

Multispectral Imager covers 13 spectral segments (443–2190 nm), a width of 290 km, with a spatial resolution of 10 m (4 visible spectral segments and 1 near infrared spectral segment), 20 m (6 red edge spectral segments and short-wave infrared spectral segments), and 60 m (3 atmospheric correction spectral segments). The Sentinel-3 was launched on 16 February 2016 [24]. The Sentinel-3-3 satellite has two payloads: one is the OLCI (Sea-Land Colorimeter) and the other is the SLSTR (Sea-Land Surface Temperature Radiometer). The OLCI is an optical instrument designed to provide data continuity for ENVISAT’s MERIS. The OLCI is a push-sweep imaging spectrometer that measures solar radiation reflected from the Earth in 21 spectral bands with a ground-based spatial resolution of 300 m [25]. Multispectral remote-sensing data of the Sentinel-2 MSI and Sentinel-3 OLCI data were obtained from the ESA (2 October 2021, <https://Sentinel.esa.int/web/Sentinel/home>). In this study, only ENVI (ENVI5.4.1) soft data were used for preprocessing, including radiometric calibration and FLAASH atmospheric correction.

3.3. Methods

3.3.1. Construction of Spectral Index

The information from the ground objects observed by remote sensing data is mainly displayed by the difference and change of the spectral characteristics of the ground objects [26]. The ground features obtained by the different spectral channels have different correlations with different elements or some characteristic states of ground features. However, complex remote sensing data can only be represented by a single channel or multi-channel spectral combination [11]. Therefore, further mining with very limited remote sensing signals is necessary to represent ground object information through remote sensing data. In this study, the combination of multi-spectral remote sensing data (such as linear and non-linear combination, subtraction, multiplication, and division) was selected to achieve the effective expression of spectral information and to lay a foundation for the qualitative and quantitative evaluation of water body information. The optimal remote sensing indices (RI, DI, and NDI) were selected for the estimation of WQI, in which multiband remote sensing data were used as variable factors. Subsequently, a combined operation was conducted for various bands and the sensitivity of WQI information, which was obviously better than that of the single-band models, highlighting the advantages of using band combinations. The remote sensing index of water quality in arid area was constructed by Formulas (1)–(3):

$$RI(i, j) = \frac{R_i}{R_j} \quad (1)$$

$$RI(i, j) = \frac{(R_i - R_j)}{(R_i + R_j)} \quad (2)$$

$$DI(i, j) = R_i - R_j \quad (3)$$

where RI (i, j) is the ratio remote sensing index, NDI (i, j) is the water body normalized remote sensing index, DI (i, j) is the water body difference remote sensing index, and i, j is any band of the data of any two bands of the 350–2500 nm band.

3.3.2. Fluorescence Line Height

The statistical algorithm, based on the correlation between fluorescence line height (FLH) and chlorophyll concentration, is called the fluorescence baseline height method. The general algorithm is derived based on three wavelengths, including the central wavelength which is the maximum value of chlorophyll fluorescence (around 685 nm, which varies with the concentration of water components), and the other two baseline bands which are located on both sides of the fluorescence peak, as shown in Figure 2 [27]. The fluorescence line height (FLH) was calculated as follows: where C was the concentration of chlorophyll

on the water surface (unit: mg/m^3); and FLH is the fluorescence baseline height (unit: $\text{mW}/(\text{cm}^2 \cdot \text{Sr} \cdot \text{nm})$). a , b , and k are the coefficients.

$$\text{FLH} = K + \frac{a \times C}{1 + b \times C} \quad (4)$$

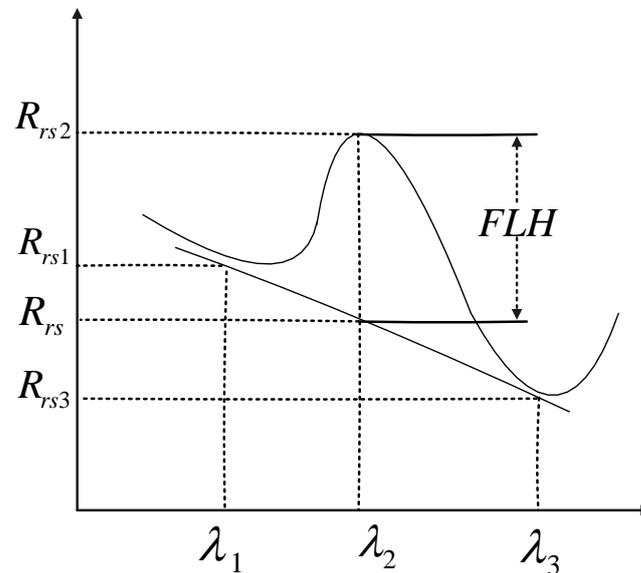


Figure 2. The principle of fluorescent line height.

The calculation formula of FLH is shown in Formula (4), where λ_2 is the central wavelength, and λ_1 and λ_3 are the selected baseline wavelengths. L_1 , L_2 , and L_3 are the radiance values of corresponding wavebands (unit: $\text{mW}/(\text{cm}^2 \cdot \text{Sr} \cdot \text{nm})$). The fluorescence channel designs are 665, 681.25, and 709 nm.

$$\text{FLH} = L_2 - \left[L_1 + (L_1 - L_3) \times \frac{\lambda_2 - \lambda_1}{\lambda_1 - \lambda_3} \right] \quad (5)$$

3.3.3. Water Quality Index (WQI)

The WQI is a comprehensive water environment index, which can reasonably quantify the degree of water pollution [28–30]. The method was first proposed by Horton and Brown [19,20], leading to the development of many water quality indices thereafter [21,22]. WQI can effectively reflect the water quality according to research objectives. Consequently, the WQI has been widely used in water environment assessments [31,32]. The smaller the WQI, the better the water quality. The researchers chose the water quality index constructed by Wang [11] for calculation. The index is constructed by using the measured water quality data of the Ebinur Lake basin, which meet the needs of water quality evaluation in arid areas. The water quality index scale is shown in Table 1.

Table 1. Water Quality Index scale.

Class	Threshold Value	Water Quality
I	≥ 50	Excellent water
II	(50–100)	Good water
III	[100–200)	Poor water
IV	[200–300)	Very poor water
V	≥ 300	Unsuitable for drinking

3.3.4. SVM Model

The Support Vector Machine (SVM) is a kind of machine learning technology based on the principle of structural risk minimization. It can solve the problems of small sample, nonlinear, high dimension, and local minimum well. It has an excellent prediction and generalization ability. The penalty factor C and the kernel function parameter σ in a support vector machine directly affect the prediction accuracy of the model. According to previous studies, the following three optimization algorithms can improve the accuracy of the SVM algorithm: Cross-validation selecting the optimal parameter (CV_cg); Genetic Algorithm (GA); and Particle Swarm Optimization (PSO) [33,34]. In this study, particle swarm optimization was selected for parameter optimization, as Wang proved that particle swarm optimization was more suitable for Ebinur Lake [11].

3.3.5. Estimate the Evaluation Index of the Model

In the establishment of the estimation model and the evaluation of accuracy, the fitting coefficient R^2 , standard deviation SD, and root mean square error RMSE were selected in this study. R^2 is the determination coefficient. RPD refers to relative analysis error. $RPD < 1.4$ indicates that the model is unreliable; $1.4 < RPD < 2$ indicates that the model has a general accuracy; and $RPD > 2$ indicates that the model has a high prediction ability [11].

4. Results and Analysis

4.1. Analysis of Spatial Variation Trend of WQI

Figure 3 shows the spatial distribution pattern of the WQI in Ebinur Lake, whereby the maximum value of WQI is 5678.35 and the minimum value is 1066.65. Overall, the degree of water pollution of Ebinur Lake is very high, and the salt content in Ebinur Lake is at a high level as well. However, different parts of Ebinur Lake are polluted at differing degrees. Specifically, the northwestern part of Ebinur Lake is the most polluted area. Similarly, the water environment and ecological environment safety of the Junggar Basin in northern Xinjiang are threatened by water quality issues. Therefore, efficient digital management of water quality is particularly important to ensure water sustainability in these areas.

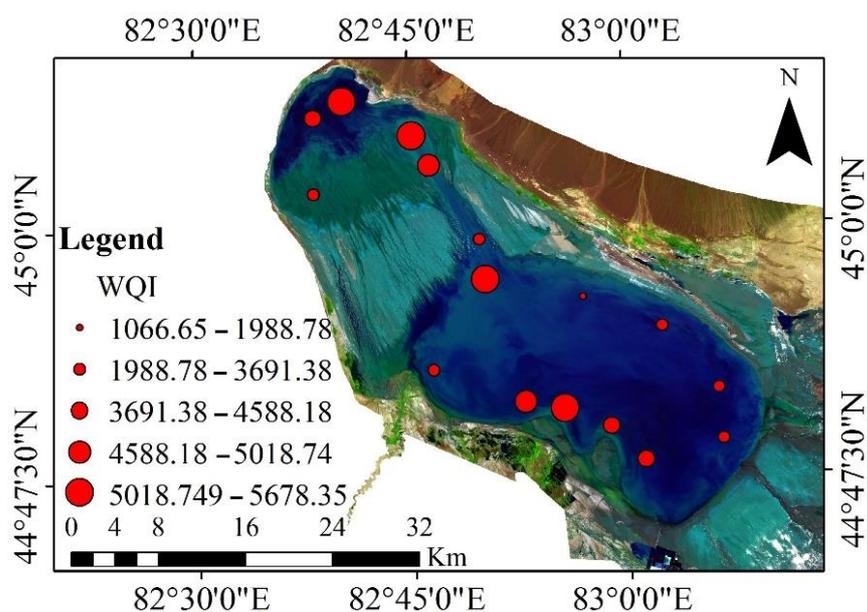


Figure 3. The spatial distribution of water quality index (WQI) in Ebinur Lake.

4.2. Study on Reflectance Spectral Characteristics of Water in Ebinur Lake

4.2.1. Spectral Characteristics of Water Based on Sentinel 3 Data

To obtain the most sensitive and effective water quality monitoring information, Sentinel 3 images were processed with the 1st, 2nd, and 3rd derivatives. However, the pixel reflectance value obtained by the 3rd derivative processing was the same due to the coarse spatial scale resolution. Thus, the pixel reflectance value was not considered in this study. The fluorescence baseline height (FHL) of the watercolor sensor was one of the main parameters examined in this study. The FHL calculated values in this paper are shown in Figure 4.

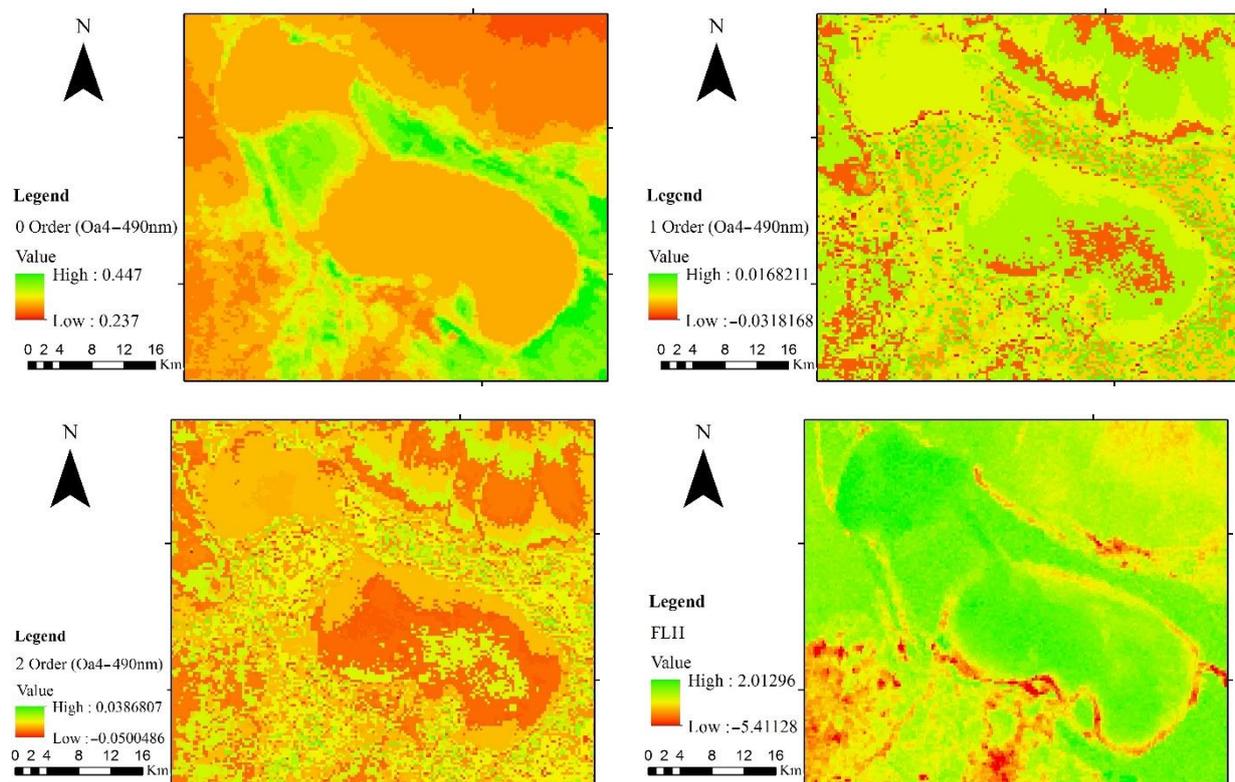


Figure 4. Spectral characteristics of water of Sentinel 3 (0 order, 1 order, 2 Order, and FHL).

To intuitively study the remote sensing data mining by image derivatives algorithm, we demonstrate the images of the fourth band (Oa4) of Sentinel 3 data with a central wavelength of 490 nm, shown in Figure 3. The raw data show that the lowest reflectance is 0.237 and the highest is 0.447. In the wetlands around the lake, salt spills out and forms salt shells on the surface due to the high degree of salinization. Therefore, the maximum reflectivity is in the salt crust around the lake, and the land-water boundary is very clear, but the difference between the land-water boundary is not as clear in the shallow lake depth. In the first derivative data, the lowest reflectance is -0.0318 and the highest is 0.0168. The boundary between land and water disappears. In terms of color, the reflectance of the surrounding mountains is in the same range as that of the center of the lake, but for the lake as a whole, the spectrum of the water body is different. In the second derivative data, the lowest reflectance is -0.05 and the highest is 0.0386. In terms of reflectance values, the second derivative amplifies the difference in reflectance values better than the first derivative. Although the boundary between water and land is blurred, it is still distinguishable. The reflectance of the surrounding mountains is in the same range as that of the center of the lake, but the spectral of the water body is different for the whole lake. In the fluorescence line height (FLH) image, the lowest value is -5.41128 and the highest value is 2.01296, where the reflectance value increases several times, the land-water

boundary is clear, the color difference in the lake is obvious, and the spectral difference of water body is distinguishable. The results show that the derivative algorithm can amplify the reflectivity difference, but it cannot separate the land-water boundary.

4.2.2. Spectral Characteristics of Water Based on Sentinel 2 Data

We also used the derivative method to process Sentinel 2 data and showed the data of the fourth band (B2) with a central wavelength of 490 nm in Figure 5. The raw data show the reflectivity ranged from 0.0003 to 0.6848. Furthermore, the maximum reflectivity is in the salt crust around the lake. The land-water boundary is very clear, but the difference is not distinguishable in the shallow water around the lake. In the first derivative data, the lowest reflectance is 0.00605 and the highest reflectance is 0.1872. The land-water boundary is clear, with the surrounding mountains and land almost distorted, but the land-water boundary cannot be clearly distinguished. In the second derivative data, the lowest reflectance is -0.3296 and the highest reflectance is 0.3591. The second derivative over the first derivative and the original image data magnify the difference in reflectivity values. The boundary between land and water is very clear, and the small lakes in the southwest can also be distinguished. The spectral difference between the surrounding plain land and vegetation cover area is clear, but the spectral difference between the water body in the lake is not significant. The third derivative image data shows that the lowest reflectivity is -0.209225 and the highest reflectivity is 0.1361. In the reflectance value, the difference of the reflectance value can be reduced by the third-order derivative image data compared with the first-order derivative and second-order derivative image data. The boundary between water and land is very clear.

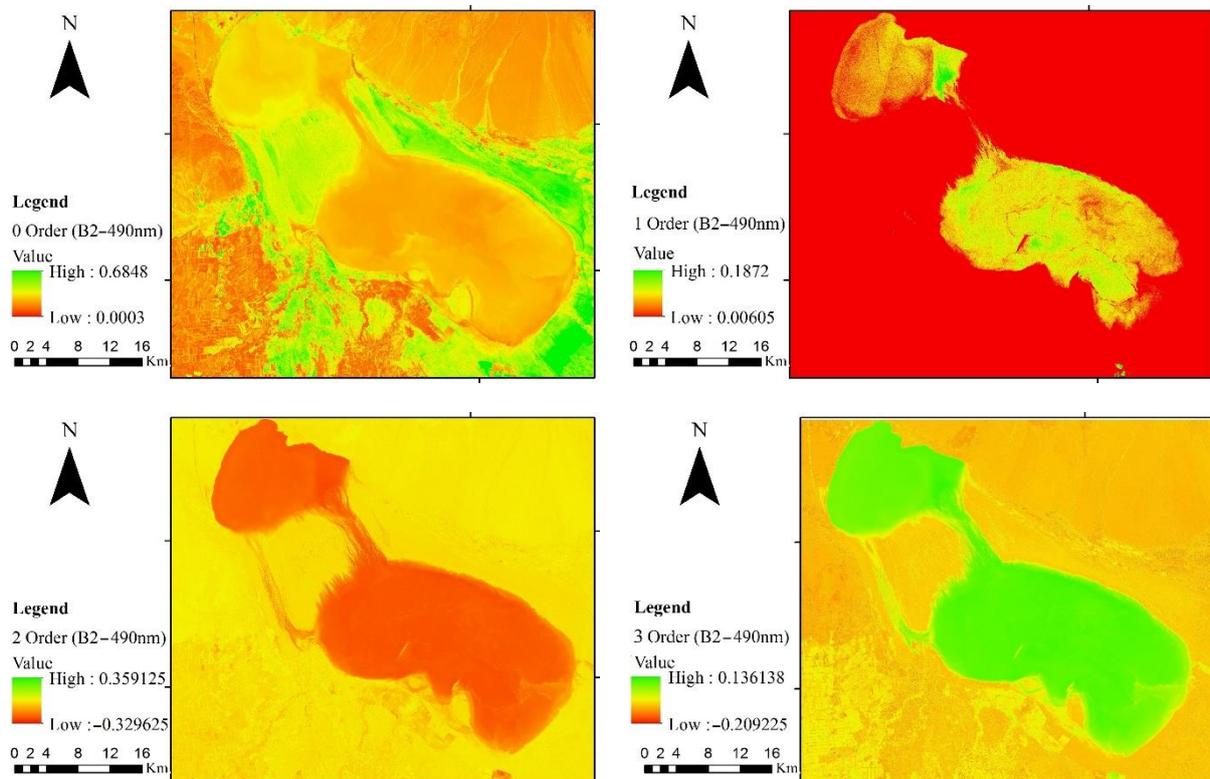


Figure 5. 0 order, 1 order, 2 Order, and 3 Order image of B14.

4.3. Relationship between Spectral Parameters and WQI

4.3.1. Relationship between WQI and Spectral Parameters from Sentinel 3 Data

(1) Relationship between Single Band Reflectance and WQI

The correlation coefficients between the WQI and the spectral reflectance from the raw image, and the first and second order derivative spectral values of the Sentinel-3 OLCI image data were calculated in this study. The results are shown in Figure 6. These correlation coefficients were tested at the 0.01 significance level. As the derivative order increases, the number of bands passing the significance test also increases, and the correlation coefficient also increases. The bands Oa4, Oa5, and Oa21 in the first-order differential passed the significance test, with the bands Oa3, Oa4, Oa5, Oa11, and Oa21 in the second-order differential also passing the significance test. The results further show that the differential method is helpful in remote sensing spectral data mining.

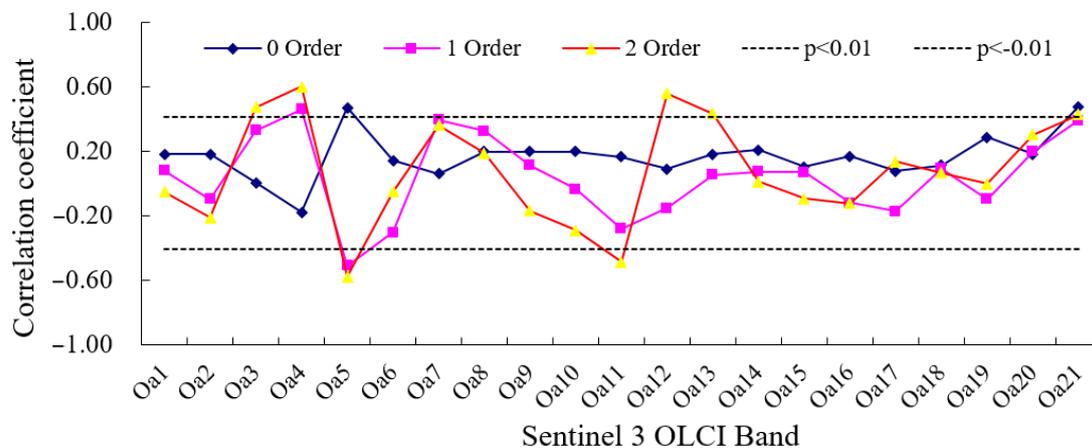


Figure 6. Relationship between single band reflectance and water quality WQI in Sentinel 3 data.

(2) Relationship between spectral index from Sentinel 3 data and WQI

To enhance the spectral difference between a water body and other ground objects, we have constructed the water spectral index. In this study, NDI, DI, and RI were selected as the combination methods of spectral indexes, and the relationship between the WQI and spectral indexes was studied through Sentinel 3 data, as shown in Figure 7, providing a basis for the further construction of a water quality evaluation model. The correlation coefficient between the WQI and spectrum index of water is shown in Table 2.

We found that DI and NDI chose the same band in the same derivative order, such as the 0 derivative. For RI, the highest correlation coefficient between the raw spectral reflectance, derivative spectral reflectance value, and WQI is 0.701 at the 0-order derivative. The combined band is Oa13 and Oa17, the lowest correlation coefficient is 0.602, and the combined band is Oa5 and Oa20 at the second order derivative. For DI, the highest correlation coefficient between the raw spectral reflectance, derivative spectral reflectance value, and WQI is 0.705 at the 0-order derivative. The combined band is Oa3 and Oa8, with the lowest correlation coefficient 0.602, and the combined band is Oa5 and Oa21 at the second order derivative. For NDI, the highest correlation coefficient between the raw spectral reflectance, derivative spectral reflectance value, and WQI is 0.701 at the 0-order derivative. The combined band is Oa4 and Oa5, the lowest correlation coefficient is 0.592, and the combined band is Oa5 and Oa21 at the second order derivative. The study found that the derivative algorithm for Sentinel 3 data did not significantly improve the relationship between the spectral index and WQI, because the relationship between the spectral index and water quality index (WQI) constructed from Sentinel 3 raw data was the best.

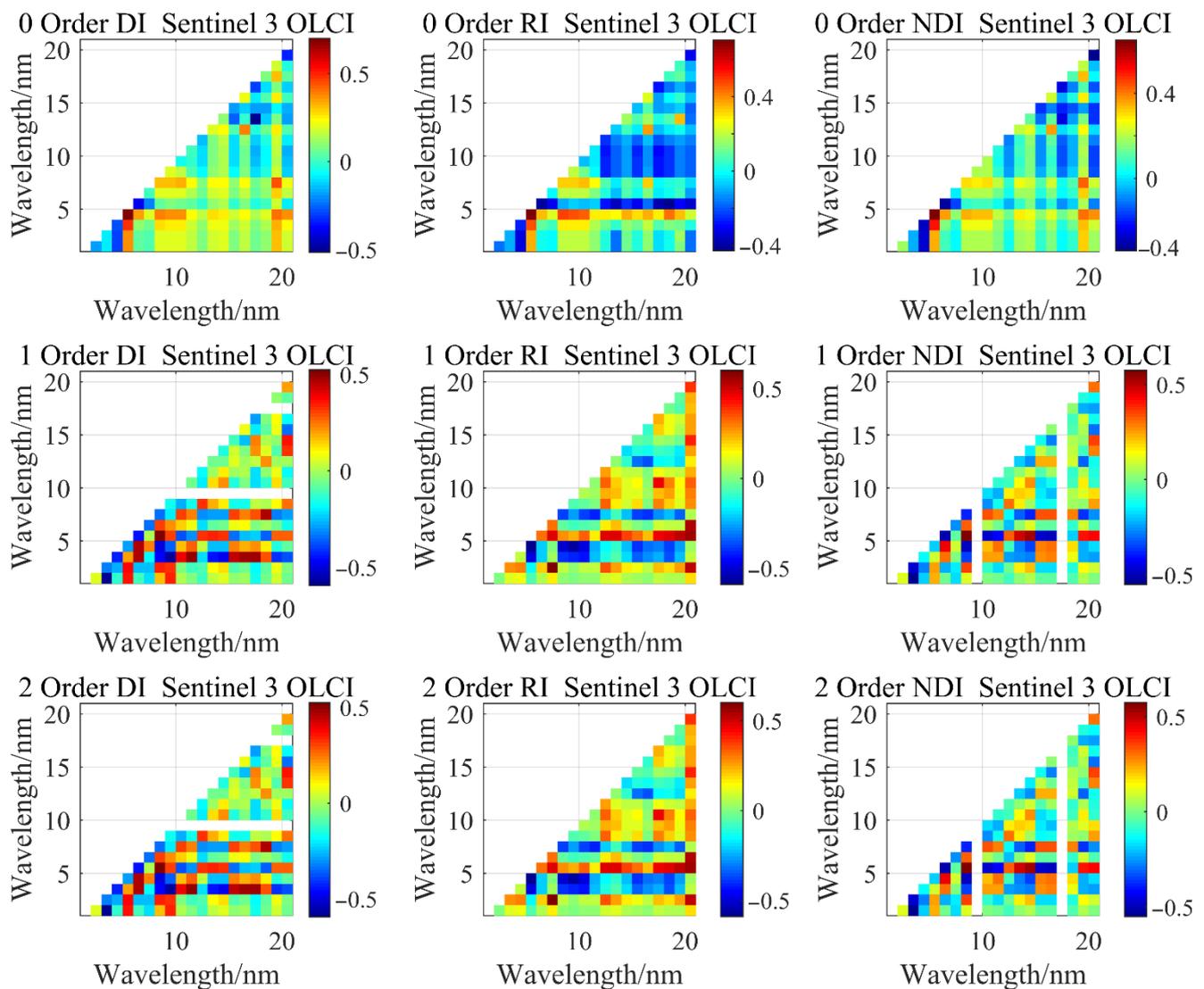


Figure 7. Relationship between WQI and spectrum index from Sentinel 3.

Table 2. Correlation coefficient between WQI and spectrum index of water.

Derivative Order	RI		DI		NDI	
	Band	R	Band	R	Band	R
0	Oa13/Oa17	0.701	Oa4 – Oa5	0.705	$(Oa4 - Oa5)/(Oa4 + Oa5)$	0.701
1	Oa5/Oa20	0.695	Oa3 – Oa8	0.662	$(Oa3 - Oa8)/(Oa3 + Oa8)$	0.622
2	Oa1/Oa3	0.602	Oa5 – Oa21	0.602	$(Oa5 - Oa21)/(Oa5 + Oa21)$	0.592

4.3.2. Relationship between WQI and Spectral Parameters from Sentinel 2 Data

(1) Relationship between single band reflectance and WQI

The correlation coefficients between the WQI and the spectral reflectance from the raw image, and the first and second order derivative spectral values of Sentinel-2 MSI image data were calculated in this study. The results are shown in Figure 5, in which correlation coefficients were tested at the 0.01 significance level. The correlation coefficient curves of Sentinel-2 MSI original spectral reflectance, derivative spectral values of order 1, 2, and 3, and water quality index WQI calculation are shown in Figure 8.

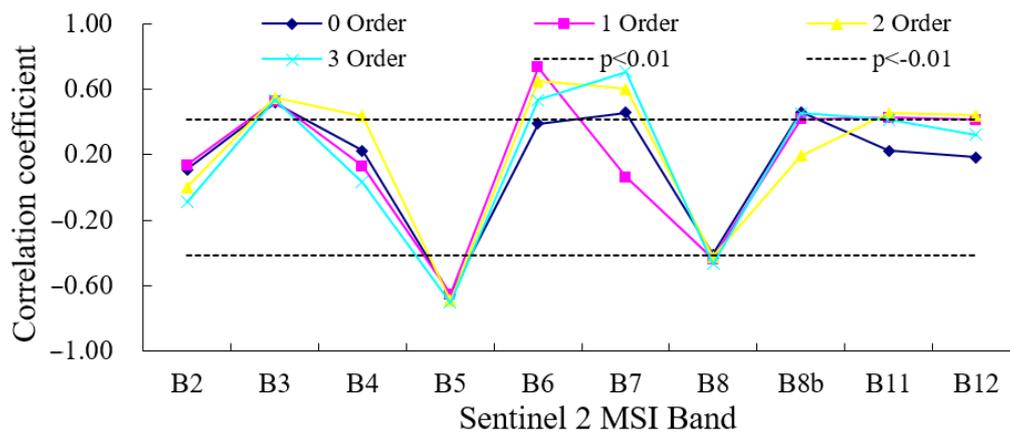


Figure 8. Relationships between WQI and spectral from sentinel-2 MSI.

The relationships between the raw reflectance of Sentinel-2 MSI image data and WQI was significantly correlated in four bands: B3, B5, B7, and B8b. The number of bands passing the significance test increased and the correlation coefficient also increased, with the derivative order increasing as well. The first order derivative was significant in the bands B3, B5, B7, and B8b, and the second order derivative was significant in the bands B3, B4, B5, B7, and B8b. The Order 2 derivative was significant in the bands B3, B5, B6, B7, and B8b. Although the bands' number of significance tests passed varied, the trend of the curves from the phase values in Figure 8 was consistent.

(2) Relationship between Spectral Index of Sentinel 2 Data and WQI

In this study, NDI, DI, and RI were selected as spectral indices, and the relationship between the spectral index and WQI was explored, as shown in Figure 9 and Table 3. For RI, the relationship between the RI spectral indices and the WQI was significant at the first-order derivative, with a R value of 0.763. For DI, the relationship between the DI spectral indices and the WQI was significant at the second-order derivative, with a R value of 0.778. For NDI, the relationship between the NDI spectral indices and the WQI was significant at the first-order derivative, with a R value of 0.776. We found that the derivative algorithm of Sentinel 2 MSI data improves the relationship between the spectral index and WQI.

4.4. Verification and Precision Analysis of Water Quality Estimation Model

4.4.1. Validation of WQI Estimation Model by Sentinel 2 Data

We used 15 groups of field sample data to train the SVR model, input images of Ebinur Lake to calculate WQI, and then extract the WQI of sampling points as the predicted WQI for model precision analysis. The predicted WQI is represented by WQI_P , and the measured WQI is represented by WQI_M . The relationship between the two is shown in Table 4. We found that the optimal model was Sentinel 2 MSI data based on the third derivative data. The R^2 and RPD of the model were 0.81 and 1.86, respectively. These results indicate that the model has a strong stability.

Table 3. Correlation coefficient between WQI and spectrum index of water.

Derivative Order	RI		DI		NDI	
	Band	R	Band	R	Band	R
0	B2/B4	0.706	B5 – B6	0.741	$(B2 - B4)/(B2 + B4)$	0.704
1	B3/B5	0.763	B3 – B6	0.763	$(B3 - B5)/(B3 + B5)$	0.776
2	B3/B4	0.741	B4 – B11	0.778	$(B3 - B4)/(B3 - B4)$	0.731
3	B5 – B8	0.736	B5 – B7	0.741	$(B4 - B5)/(B4 - B5)$	0.735

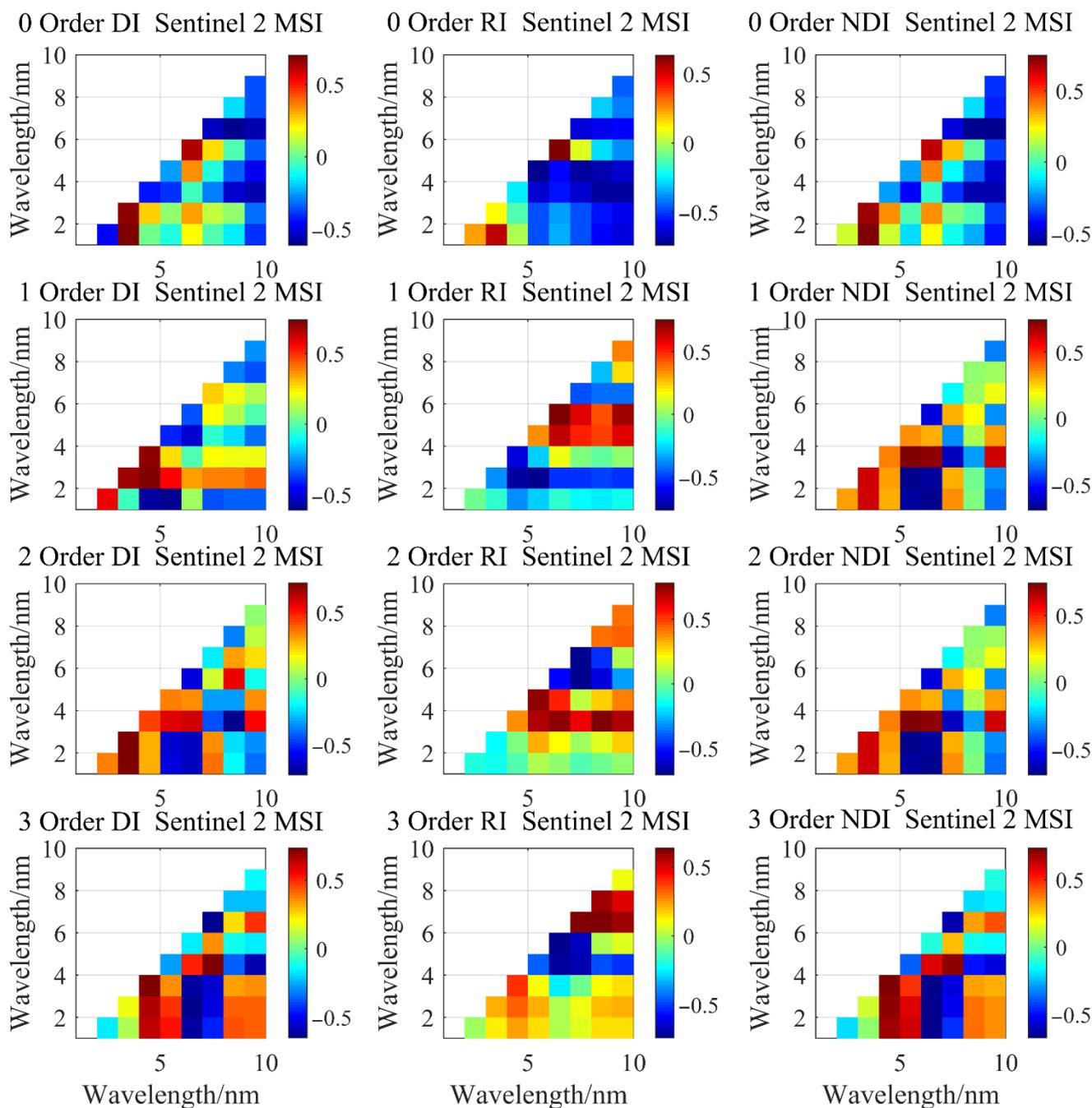


Figure 9. Correlation coefficient between WQI and spectrum index of water.

Table 4. Summary of relationship between the measured values and predicted values.

Order	X	Y	PSO-SVR					
			R ²	RMSE	SD	RPD	Slope	N
0	WQI _M	WQI _P	0.69	344.67	503.07	1.45	0.73	16
1	WQI _M	WQI _P	0.73	302.18	492.36	1.62	0.77	16
2	WQI _M	WQI _P	0.79	245.69	398.06	1.62	0.81	16
3	WQI _M	WQI _P	0.81	213.41	398.72	1.86	0.84	16

4.4.2. Validation of WQI Estimation Model Supported by Sentinel 3 Data

Similarly, we used 15 groups of field sample data and corresponding Sentinel 3 OLCI data for SVR model training, to input images of Ebinur Lake to calculate WQI, and then

to extract the WQI of sampling points as the predicted WQI for precision analysis. The predicted WQI is represented by WQI_P , and the measured WQI is represented by WQI_M . The relationship between the two is shown in Table 5. The best model was the fluorescence baseline data of Sentinel 3 OLCI data. The R^2 and RPD of the model were 0.80 and 1.79, respectively, showing that the model has a strong stability.

Table 5. Summary of relationship between the measured values and predicted values.

Order	X	Y	PSO-SVR					
			R^2	RMSE	SD	RPD	Slope	N
0	WQI_M	WQI_P	0.76	233.14	412.38	1.76	0.76	16
1	WQI_M	WQI_P	0.73	342.72	521.09	1.52	0.72	16
2	WQI_M	WQI_P	0.69	354.47	519.84	1.46	0.71	16
FLH	WQI_M	WQI_P	0.80	200.78	359.28	1.79	0.84	16

4.5. Spatial Distribution Map of WQI in Ebinur Lake

A spatial distribution map of WQI based on an optimal model constructed from Sentinel 2–3 derivative data is presented, showing that the water quality in the northwest of Ebinur Lake is the lowest in that region. The northwest of Ebinur Lake is eroded by the Alashan Pass gale, and the water depth is less than 1 m. The water quality in the northeast of Ebinur Lake was the second highest, but the water quality was deteriorated by the salinization of large saline-alkali land and soil around the lake. The deterioration of water quality in the northeast of Ebinur Lake is closely related to human activities in the north, which is one of the largest halogen insect production bases in China. The distribution of WQI in Ebinur Lake is shown in Figure 10.

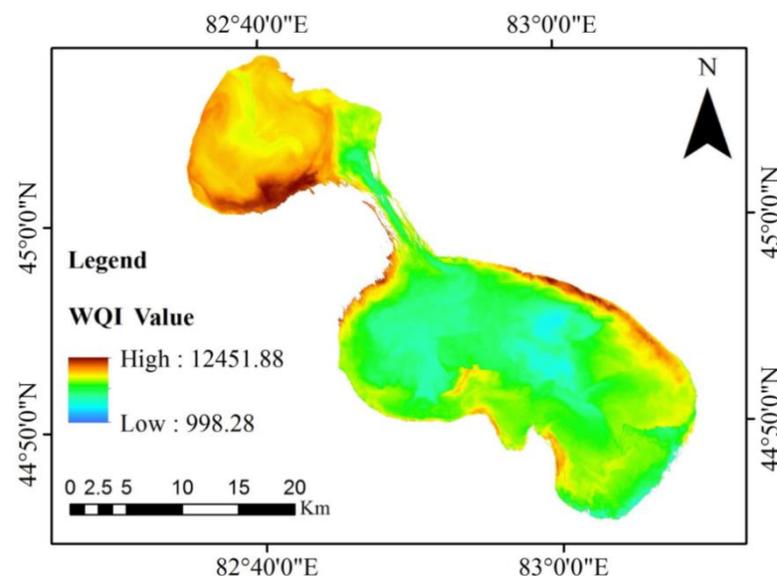


Figure 10. Inverted water quality index (WQI) in Ebinur Lake.

5. Discussion

5.1. Water Quality Index (WQI) as a Potential Proxy for Water Environment

Overall, the results of this study are very indicative, and in agreement with [11] our prediction, proving that remote sensing is a very useful potential tool for water quality monitoring. However, it should be noted that the uncertainty of the WQI remote sensing monitoring model for lake water quality was analyzed from the perspective of time and space. (1) In terms of time, this experiment was limited to the Ebinur Lake watershed during the dry season, aiming to clarify the relationship between WQI and spectral. Although WQI has seasonal variability, WQI also has great variability in the same period and within

the same watershed. Therefore, the precision of the WQI model is not limited by season and has its portability in time. (2) The spatial WQI is mainly affected by water in the watershed, whereby spectral also reflects the integration of the whole water environment. The WQI estimation model was established based on the relationship between the spectral index and WQI. In the space under the influence of the study area, portability needs further validation for the model. However, the Ebinur Lake watershed is a typical area of arid area in Central Asia, and its model has certain portability in Central Asia. The extension of a wider range needs further verification. In short, it should be noted that the WQI estimation model is spatially uncertain.

5.2. Spectral Derivative Method and Spectral Indices as Useful Tools for Remote Sensing Modeling of Water Quality

To better mine the information of spectral data from remote sensing, we introduced the spectral derivative method to realize the extraction of spectral information of a water body. The results show that the spectral derivative method can improve the relationship between water body spectral and WQI, whereby the R^2 value of 0.6 is at the most sensitive wavelengths. The derivative technology is not only a powerful tool for analyzing spectra, but also improves multiple collinearity problems considerably [35]. The derivative technology has a strong effect on the peak of the micro spectrum; therefore, it can be used to improve the spectral resolution and sensitivity of the analysis. To some extent, it has the function of removing noise. Fractional derivatives can reduce the intense peak deformation and effectively retain the structure of the original curve, which is more advantageous than other integer derivatives.

Spectral indices are useful for remote sensing modeling of water quality: the optimal remote sensing indices (RI, DI and NDI) were selected for the estimation of WQI, in which multiband remote sensing data were used as variable factors; a combined operation was conducted for various bands, and the sensitivity of WQI information, which was obviously better than that of the single-band models, highlights the advantages of using band combinations. Fernández-Buces et al. used a combined spectral response index to map the soil salinity of bare soil and vegetation. They found a correlation between the normalized difference vegetation indices (NDVI) and electrical conductivity [36]. Therefore, we applied this method, as well as a formula that uses the DI, RI, and NDI of the reflectance values, to establish a new spectral index for estimating WQI.

6. Conclusions

In this paper, the Ebinur Lake basin was selected as the study area, with the aims of revealing the response between water quality index and water body reflectivity, as well as to describe the relationship between water quality index and water reflectivity. A remote sensing monitoring model of WQI was further established, and the water quality of the lake was evaluated by remote sensing. The results indicate:

- (1) A Water Quality Index (WQI), based on remote sensing techniques, effectively evaluated the water environment in Ebinur Lake. The Water quality of Ebinur Lake is the lowest, with a WQI value as high as 4000;
- (2) To better mine the information of spectral data from remote sensing, we introduced the spectral derivative method to realize the extraction of spectral information from a water body. The results show that the spectral derivative method can improve the relationship between the water body spectral and WQI, whereby the R^2 value of 0.6 is at the most sensitive wavelengths;
- (3) When multi-source spectral data were integrated through the spectral index (DI, RI, and NDI) and fluorescence baseline, the correlation between the spectral sensitivity index and WQI was found to be greater than 0.6 at the significance level of 0.01;
- (4) The distribution map of WQI in Ebinur Lake was obtained by the optimal model, which was constructed based on the third derivative data of Sentinel 2 data. Results indicate that the water quality in the northwest of Ebinur Lake was the lowest in the

region. In conclusion, remote sensing techniques were found to be highly effective and lay a foundation for water quality detection in arid areas.

Author Contributions: Conceptualization, F.Z. and X.W.; methodology, N.W.C.; software, X.W.; validation, C.L., T.G. and X.L.; formal analysis, X.W.; investigation, C.L.; resources, F.Z.; data curation, X.W.; writing—original draft preparation, X.W.; writing—review and editing, H.-T.K. and J.S.; visualization, C.L., W.W. and N.C.; supervision, F.Z.; project administration, F.Z.; funding acquisition, F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (U1603241), the National Natural Science Foundation of China (Xinjiang Local Outstanding Young Talent Cultivation) (Grant No. U1503302), and the Tianshan Talent Project (Phase III) of the Xinjiang Uygur Autonomous Region. We want to thank the editor and anonymous reviewers for their valuable comments and suggestions to this paper.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to confidentiality.

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

References

- Cardinale, B.J. Biodiversity improves water quality through niche partitioning. *Nature* **2011**, *472*, 86–89. [[CrossRef](#)] [[PubMed](#)]
- Pekel, J.F.; Cottam, A.; Gorelick, N.; Belward, A.S. High-resolution mapping of global surface water and its long-term changes. *Nature* **2016**, *540*, 418–422. [[CrossRef](#)] [[PubMed](#)]
- Michalak, A.M. Study role of climate change in extreme threats to water quality. *Nature* **2016**, *535*, 349–350. [[CrossRef](#)] [[PubMed](#)]
- Feyisa, G.L.; Meilby, H.; Fensholt, R.; Proud, S.R. Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery. *Remote Sens. Environ.* **2014**, *140*, 23–35. [[CrossRef](#)]
- Kumar, V.; Sharma, A.; Chawla, A.; Bhardwaj, R.; Thukral, A.K. Water quality assessment of river Beas, India, using multivariate and remote sensing techniques. *Environ. Monit. Assess.* **2016**, *188*, 137. [[CrossRef](#)] [[PubMed](#)]
- Dörnhöfer, K.; Oppelt, N. Remote sensing for lake research and monitoring—Recent advances. *Ecol. Indic.* **2016**, *64*, 105–122. [[CrossRef](#)]
- Cheng, P.; Wang, X.L. The design and implementation of Remote-sensing Water Quality Monitoring System based on SPOT-5. In Proceedings of the Second IITA International Conference on Geoscience & Remote Sensing, Qingdao, China, 28–31 August 2010.
- Pullianen, J.; Kallio, K.; Eloheimo, K.; Koponen, S.; Hallikainen, M. A semi-operative approach to lake water quality retrieval from remote sensing data. *Sci. Total Environ.* **2001**, *268*, 79–93. [[CrossRef](#)]
- Su, T.C. A study of a matching pixel by pixel (MPP) algorithm to establish an empirical model of water quality mapping, as based on unmanned aerial vehicle (UAV) images. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 213–224. [[CrossRef](#)]
- Burns, P.; Nolin, A. Using atmospherically-corrected Landsat imagery to measure glacier area change in the Cordillera Blanca, Peru from 1987 to 2010. *Remote Sens. Environ.* **2014**, *140*, 165–178. [[CrossRef](#)]
- Wang, X.P.; Zhang, F.; Ding, J.L. Evaluation of water quality based on a machine learning algorithm and water quality index for the Ebinur Lake Watershed, China. *Sci. Rep.* **2017**, *7*, 12858. [[CrossRef](#)]
- Alves, E.M.; Rodrigues, R.J.; Corrêa, C.S.; Fidemann, T.; Rocha, J.C.; Buzzo, J.L.L.; Note, P.O.; Núñez, G.F. Use of ultraviolet–visible spectrophotometry associated with artificial neural networks as an alternative for determining the water quality index. *Environ. Monit. Assess.* **2018**, *190*, 319. [[CrossRef](#)]
- Gogu, R.; Carabin, G.; Hallet, V.; Peters, V.; Dassargues, A. GIS-based hydrogeological databases and groundwater modelling. *Hydrogeol. J.* **2001**, *9*, 555–569. [[CrossRef](#)]
- Gholizadeh, M.H.; Melesse, A.M.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors* **2016**, *16*, 1298. [[CrossRef](#)] [[PubMed](#)]
- Kutser, T.; Paavel, B.; Verpoorter, C.; Ligi, M.; Soomets, T.; Toming, K.; Casal, G. Remote sensing of black lakes and using 810 nm reflectance peak for retrieving water quality parameters of optically complex waters. *Remote Sens.* **2016**, *8*, 497. [[CrossRef](#)]
- Giardino, C.; Bresciani, M.; Cazzaniga, I.; Schenk, K. Evaluation of multi-resolution satellite sensors for assessing water quality and bottom depth of lake garda. *Sensors* **2014**, *14*, 24116–24131. [[CrossRef](#)] [[PubMed](#)]
- Farhad, M.; Fatemeh, D.; Mostafa, R.; Baden, M. Introducing a water quality index for assessing water for irrigation purposes: A case study of the ghezel ozan river. *Sci. Total Environ.* **2017**, *589*, 107–116.
- Nuraslinda, A.; Pauzi, A.M.; Bakar, A.A.A. Methodology of water quality index (WQI) development for filtrated water using irradiated basic filter elements. *Math. Sci. Its Appl.* **2017**, *1799*, 040010.
- Horton, R.K. An index number system for rating water quality. *J. Water Pollut. Control Fed.* **1965**, *37*, 300–306.
- Brown, R.M.; McClelland, N.I.; Deininger, R.A.; Tozer, R.G. A water quality index—do we dare? *Water Sew. Work.* **1970**, *117*, 339–343.

21. Debels, P.; Figueroa, R.; Urrutia, R.; Barra, R.; Niell, X. Evaluation of water quality in the Chilla'n river (Central Chile) using physicochemical parameters and a modified water quality index. *Environ. Monit. Assess.* **2005**, *110*, 301–322. [[CrossRef](#)]
22. Saeedi, M.; Abessi, O.; Sharifi, F.; Maraji, H. Development of groundwater quality index. *Environ. Monit. Assess.* **2009**, *163*, 327–335. [[CrossRef](#)] [[PubMed](#)]
23. Şener, Ş.; Şener, E.; Davraz, A. Evaluation of water quality using water quality index (WQI) method and GIS in Aksu river (Sw-turkey). *Sci. Total Environ.* **2017**, *584–585*, 131–144. [[CrossRef](#)]
24. D'Amico, F. Prompt-Particle-Events in ESA's Envisat/MERIS and Sentinel-3/OLCI Data: Observations, Analysis and Recommendations. Master's Thesis, University of Pisa, Pisa, Italy, 2015.
25. Sibanda, M.; Mutanga, O.; Rouget, M. Examining the potential of Sentinel-2 MSI spectral resolution in quantifying above ground biomass across different fertilizer treatments. *ISPRS J. Photogramm. Remote Sens.* **2015**, *110*, 55–65. [[CrossRef](#)]
26. Liu, C.J.; Duan, P.; Zhang, F.; Jim, C.Y.; Tan, M.L.; Chan, N.W. Feasibility of the Spatiotemporal Fusion Model in Monitoring Ebinur Lake's Suspended Particulate Matter under the Missing-Data Scenario. *Remote Sens.* **2021**, *13*, 3952. [[CrossRef](#)]
27. Hoge, F.E.; Lyon, P.E.; Swift, R.N.; Yungel, J.K.; Abbott, M.R.; Letelier, R.M.; Esaias, W.E. Validation of Terra-MODIS phytoplankton chlorophyll fluorescence line height. I. Initial airborne lidar results. *Appl. Opt.* **2003**, *42*, 2767. [[CrossRef](#)] [[PubMed](#)]
28. Brown, R.M.; McClelland, N.I.; Deining, R.A.; O'Connor, M.F. A Water Quality Index -Crashing the Psychological Barrier. In *Indicators of Environmental Quality*; Thomas, W.A., Ed.; Springer: Boston, MA, USA, 1972. [[CrossRef](#)]
29. Dede, O.T.; Telci, I.T.; Aral, M.M. The Use of Water Quality Index Models for the Evaluation of Surface Water Quality: A case study for Kirmir Basin, Ankara, Turkey. *Water Qual. Expo. Health* **2013**, *5*, 41–56.
30. Harkins, R.D. An Objective Water Quality Index. *J. (Water Pollut. Control. Fed.)* **1974**, *46*, 58591.
31. Wu, Z.S.; Wang, X.L.; Chen, Y.W.; Cai, Y.J.; Deng, J. Assessing river water quality using water quality index in Lake Taihu Basin, China. *Sci. Total Environ.* **2018**, *612*, 914–922. [[CrossRef](#)]
32. Mukate, S.; Wagh, V.; Panaskar, D.; Jacobs, J.A.; Sawant, A. Development of new integrated water quality index (IWQI) model to evaluate the drinking suitability of water. *Ecol. Indic.* **2019**, *101*, 348–354. [[CrossRef](#)]
33. Wiley, E.O.; Mcnyset, K.M.; Peterson, A.T.; Robins, C.R.; Stewart, A.M. Niche modeling and geographic range predictions in the marine environment using a machine-learning algorithm. *Oceanography* **2003**, *16*, 120–127. [[CrossRef](#)]
34. Marjanović, M.; Kovačević, M.; Bajat, B.; Voženilek, V. Landslide susceptibility assessment using SVM machine learning algorithm. *Eng. Geol.* **2011**, *123*, 225–234. [[CrossRef](#)]
35. Peón, J.; Fernández, S.; Recondo, C.; Calleja, J.F. Evaluation of the spectral characteristics of five hyperspectral and multispectral sensors for soil organic carbon estimation in burned areas. *Int. J. Wildland Fire* **2017**, *26*, 230–239. [[CrossRef](#)]
36. Fernández-Buces, N.; Siebe, C.; Cram, S.; Palacio, J.L. Mapping soil salinity using a combined spectral response index for bare soil and vegetation: A case study in the former Lake Texcoco, Mexico. *J. Arid Environ.* **2006**, *65*, 644–667. [[CrossRef](#)]