

Article

Remote Estimation of Water Quality Parameters of Medium- and Small-Sized Inland Rivers Using Sentinel-2 Imagery

Kuan Huangfu ¹, Jian Li ^{2,*}, Xinja Zhang ², Jinping Zhang ¹, Hao Cui ² and Quan Sun ¹

¹ School of Water Conservancy Science & Engineering, Zhengzhou University, Zhengzhou 450001, China; hfzzu2014@163.com (K.H.); zhjp@zzu.edu.cn (J.Z.); sunquan@gs.zzu.edu.cn (Q.S.)

² School of the Geo-Science & Technology, Zhengzhou University, Zhengzhou 450001, China; zhangxinjia@zzu.edu.cn (X.Z.); cuihao@zzu.edu.cn (H.C.)

* Correspondence: lijian5277@163.com

Received: 27 August 2020; Accepted: 5 November 2020; Published: 7 November 2020



Abstract: In the application of quantitative remote sensing in water quality monitoring, the existence of mixed pixels greatly affects the accuracy of water quality parameter inversion, especially for narrow inland rivers. Improving the image spatial resolution and weakening the interference of mixed pixels in the image are some of the urgent problems to be solved in the study of water quality monitoring of medium- and small-sized inland rivers. We processed Sentinel-2 multispectral images using the super-resolution algorithm and generated a set of 10 m spatial resolution images with basically unchanged reflection characteristics. Both qualitative and quantitative evaluation results show that the super-resolution algorithm can weaken the influence of mixed pixels while maintaining spectral invariance. Before the application of the super-resolution algorithm, the inversion accuracy of water quality parameters in this study were as follows: for NH₃-N, the R² was 0.61, the root mean squared error (RMSE) was 0.177 and the mean absolute percentage error (MAPE) was 29.33%; for Chemical Oxygen Demand (COD), the R² was 0.26, the RMSE was 0.756 and the MAPE was 4.62%; for Total Phosphorus (TP), the R² was 0.69, the RMSE was 0.032 and the MAPE was 30.58%. After the application of the super-resolution algorithm, the inversion accuracy of water quality parameters in this study were as follows: for NH₃-N, the R² was 0.67, the RMSE was 0.161 and the MAPE was 25.88%; for COD, the R² was 0.53, the RMSE was 0.546 and the MAPE was 3.36%; for TP, the R² was 0.60, the RMSE was 0.034 and the MAPE was 24.28%. Finally, the spatial distribution of NH₃-N, COD and TP was obtained by using a machine learning model. The results showed that the application of the super-resolution algorithm can effectively improve the retrieval accuracy of NH₃-N, COD and TP, which illustrates the application potential of the super-resolution algorithm in water quality remote sensing quantitative monitoring.

Keywords: super-resolution algorithm; total phosphorus; chemical oxygen demand; water quality monitoring; Sentinel-2; NH₃-N; remote sensing

1. Introduction

Water is one of the core elements of ecology and the environment and is an important resource for all life including human beings. However, with the rapid development of the economy, the problem of water pollution is becoming increasingly serious. At present, a large number of organic pollutants are produced in industrial production, livestock breeding and human daily life, which seriously pollute the surrounding water and endanger aquatic organisms [1,2]. With the expansion of the food chain, this affects human health [3]. Early, timely, accurate and complete monitoring of water quality parameters of inland waters is of great significance to assess the risks that may affect human

health. The traditional water quality monitoring method is time-consuming and laborious. Its biggest disadvantage is that it can only obtain water quality information by monitoring section points of a river, which is not enough to reflect the water quality information of the whole river. Remote sensing can realize large-scale synchronous observations and can be used to monitor and analyze the spatial heterogeneity of water quality of the whole river so as to provide a timely, accurate and reliable guarantee for water quality monitoring of inland rivers, which cannot be achieved by field sampling of the traditional monitoring sections.

Water quality monitoring is an important remote sensing application. As critical water quality parameters, $\text{NH}_3\text{-N}$, Chemical Oxygen Demand (COD) and Total Phosphorus (TP) are important reference indexes for water pollution prevention and control. Gong et al. found that nitrogen had the highest correlation coefficient for 404 nm and 447 nm [4]. Wang et al. found that $\text{NH}_3\text{-N}$ and COD have a high correlation with the SPOT5 satellite's red band, green band and near infrared (NIR) band [5]. Chen et al. analyzed the COD standard solution with a certain ratio and found that COD has a significant spectral response characteristic at 550 nm, 565 nm, 1016 nm and 1047 nm [6]. Sun et al. estimated the concentration of TP in inland waters with complex optical turbidity based on the near-infrared and green bands of the HJ-1 satellite [7]. Previous studies have proven that there is a certain relationship between $\text{NH}_3\text{-N}$, COD, TP and reflectance spectra [4–9], which lays a foundation for the use of remote sensing technology for water quality monitoring.

At present, the sensors used in water quality inversion include WorldView-2 [10], TM/ETM+ [11], OLI [9], MODIS [12], MERIS [13,14], HJ-1 CCD [15,16], SPOT [5,17], GOCI [18], GF-1 [19], Sentinel-2 [20,21], etc. Due to the limitations of the spatial resolution of satellite sensors, most of these studies are focused on lakes or coastal areas with large water areas [22–24], and less research is conducted on inland rivers, especially small and medium-sized rivers. The river surfaces of small and medium-sized inland rivers are usually narrow, and their narrow and long shape limits the application of some low-resolution remote sensing data. Compared with other media and high-resolution images, Sentinel-2, which is free and easy to obtain, has 13 multispectral bands with a spatial resolution of up to 10 m [25]. The combination of satellites A and B can revisit the site for up to five days, which has a strong advantage in water quality inversion and can provide reliable data support for accurate inversion of water quality parameters. Only four of Sentinel-2's 13 spectral bands have a spatial resolution of 10 m, so the spectral bands that can be effectively used for water quality monitoring in such a narrow and long water area as inland small and medium-sized rivers are limited. However, the existence of a 10 m high resolution band makes it possible to improve other relatively low spatial resolution bands. Super-resolution (SR) technology [26–28] can obtain the maximum output without increasing the cost of the satellite.

Super-resolution technology refers to the reconstruction of high-resolution images from the observed low-resolution images. It has important application value in the fields of monitoring equipment, satellite images and medical images. There is a lot of research on monitoring water quality parameters using remote sensing data, and few people pay attention to the application potential of super-resolution technology in this aspect. The early super-resolution algorithm is actually from the frequency domain signal processing technology [29,30], but in the super-resolution process, the spatial domain algorithm can better maintain the image details, edges and overall structure and has a relatively wide range of applications. Therefore, most of the later super-resolution algorithms are mainly developed in the spatial domain. The commonly used spatial domain algorithms include an imaging model [31,32], interpolation-based algorithm [33], Iterative Back Projection (IBP) [34], Maximum A Posterior (MAP) [35,36], Projection Onto Convex Sets (POCS) [37,38], etc. In addition, the learning-based algorithm is also used for image super-resolution reconstruction [26].

Sentinel-2 has the problem of uneven spatial resolution distribution (10 m, 20 m, 60 m). The existence of mixed pixels in low-resolution bands lead to a large error between the calculated pixel reflectance and the actual reflectance of the water surface. Multispectral images with high spatial resolution can provide more reliable data support for water quality inversion. In previous studies,

few researchers paid attention to the application of super-resolution technology in water quality monitoring. Moreover, for $\text{NH}_3\text{-N}$, COD and TP, which are three important parameters reflecting water quality, there is no corresponding inversion model for Sentinel-2 data. Therefore, the novelty of this paper is to apply super-resolution technology to water quality monitoring to weaken the impact of mixed pixels on water quality monitoring of small and medium-sized inland rivers and improve the utilization rate of multispectral bands of Sentinel-2 data. In addition, this paper focuses on establishing inversion models of $\text{NH}_3\text{-N}$, COD and TP based on Sentinel-2.

2. Materials and Methods

2.1. The Super-Resolution Algorithm

For satellite images, panchromatic sharpening is usually used to improve image resolution. The panchromatic sharpening method ranges from the simplest multispectral band renormalization to more advanced decomposition technology [39–41]. By estimating the contribution of each multispectral band to the panchromatic band and solving each correlation component, the spatial resolution of the multispectral image can be improved and the spectral distortion can be reduced simultaneously. However, Sentinel-2 data donot have panchromatic bands. The super-resolution algorithm used in this paper was proposed by Brodu [42] and is based on local consistency between neighboring pixels and geometric consistency of sub-pixel components in the multispectral band. The algorithm is designed for multispectral and multi-resolution images. In remote sensing images, most natural objects do not fall completely within the pixel boundaries. As shown in Figure 1, some content is often shared between adjacent pixel values, and the reflection of different wavebands is different. In remote sensing, the reflection of each pixel (H) is usually considered as a linear mixture of its components (such as the mixture of water and soil) [43]. Therefore, the mixing equation of each pixel can be expressed as follows (Equation (1)):

$$H_{x,y} = W_{x,y,0} \times S_{x,y} + W_{x,y,1} \times S_{x+1,y} + W_{x,y,2} \times S_{x,y+1} + W_{x,y,3} \times S_{x+1,y+1} \quad (1)$$

where S represents the reflection of elements shared across pixels and W is the proportion of mixed elements in the pixel.

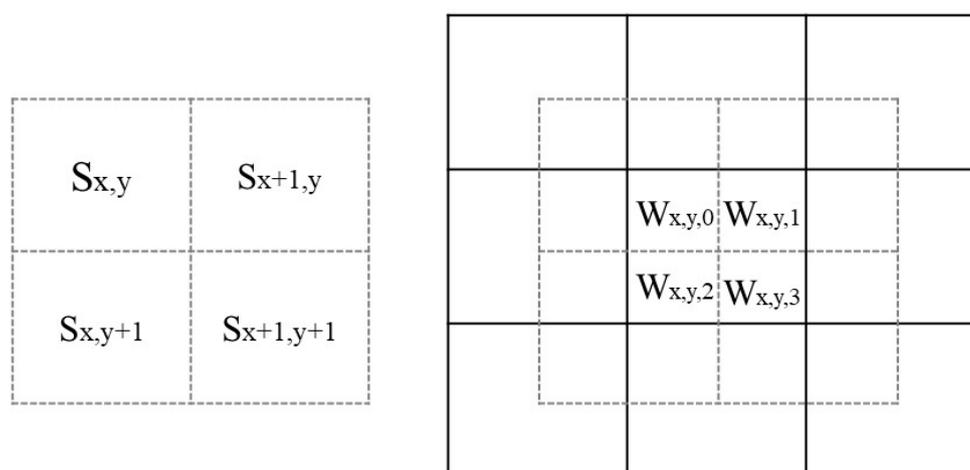


Figure 1. Values (S) are shared between neighbor pixels and combined by weights (W) to form the high-resolution pixels. The black solid line represents the distribution of high-resolution pixels, and the gray dotted line represents the distribution of shared values across pixels.

Therefore, the super-resolution pixel value with low resolution can be obtained only by calculating the proportion weight (W) of the mixing elements in the pixel and the shared value (S) of the low-resolution band. The specific operation is as follows:

For images with 20 m resolution:

- (a) High-resolution, band-specific information S is separated from common-band information W. Firstly, the 10 m resolution band H^o is subsampled to the 20 m version L^d . In this case, the optimal hybrid model is estimated by minimizing the difference between (a) the high-resolution observed pixel value H^o and (b) the resolution enhancement value H^r calculated from the subsampled data L^d (Equation (2)). The initial weight W^{opt} is set to 1/4. The initial shared value S^{opt} is the mean value of H^o of each high-resolution pixel it covers. The optimal weight value W^{opt} and shared value S^{opt} are solved iteratively.

$$\{S^{opt}, W^{opt}\} = \operatorname{argmin} \sum_{\beta \in \mathcal{H}} \sum_{x,y} \left\| H_{\beta,x,y}^o - H_{\beta,x,y}^r \right\|^2, \tag{2}$$

where \mathcal{H} represents the set of all 10 m resolution bands and β is one of the 10 m resolution bands.

- (b) Low-resolution, shared-value S is calculated.

$$V = \operatorname{argmin} \sum_{\beta \in \mathcal{H}} \sum_{x,y} \left\| S_{\beta,x,y}^{opt} - \sum_{n \in N(x,y)} v_{x,y,n} L_{\beta,n}^d \right\|^2, \tag{3}$$

$$S_{b,x,y}^{fit} = \sum_{n \in N(x,y)} v_{x,y,n} L_{b,n}, \tag{4}$$

$$S_{b,x,y}^{cor} = \bar{q}_{b,x,y} S_{b,x,y}^{fit}, \tag{5}$$

where N is the neighborhood location set of high-resolution shared value S relative to low-resolution pixels, V is the neighborhood coefficient set of low-resolution pixel position relative to the high-resolution shared value, $S_{b,x,y}^{fit}$ is the shared value of the low-resolution b-band and $\bar{q}_{b,x,y}$ is the overall average sharpening rate of the b-band, which is used to correct the low-resolution shared value.

- (c) Super-resolution pixel values H are calculated.

$$H_{b,x,y} = W_{x,y,0} \times S_{b,x,y}^{cor} + W_{x,y,1} \times S_{b,x+1,y}^{cor} + W_{x,y,1} \times S_{b,x,y+1}^{cor} + W_{x,y,3} \times S_{b,x+1,y+1}^{cor} \tag{6}$$

For images with 60 m resolution, the super-resolution method is similar. Firstly, the 60 m resolution band is converted to the 20 m super-resolution band, and then the image is processed according to the steps of 20 m to 10 m.

To quantify the performance of the super-resolution algorithm, Pearson’s correlation coefficient (R) [44], the root mean squared error (RMSE) [45] and estimation accuracy (EA) [46] were calculated for each band as follows (Equations (7) and (8)):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \tag{7}$$

$$EA = \left(1 - \frac{RMSE}{\bar{y}} \right) \times 100\% \tag{8}$$

where n is the number of pixels, \hat{y}_i is the estimated value pixel in the super-resolution image, y_i is the value pixel in the original image and \bar{y} is the average value of a pixel in the original image.

2.2. The Inversion Models

The traditional statistical regression model (TSR) has been widely used in the inversion of water quality parameters [47,48]. In this paper, a linear model, quadratic model, cubic model, exponential model, logarithmic model, reciprocal model and power function model were used for the construction of the inversion model (Table 1), and the model with the highest inversion accuracy was selected as the output of the traditional statistical regression model.

Table 1. The traditional statistical regression models.

TSR Models	Mathematical Expression (a, b, c and d Are Undetermined Parameters)
linear model	$y = a + b \times x^1$
quadratic model	$y = a + b \times x^1 + c \times x^2$
cubic model	$y = a + b \times x^1 + c \times x^2 + d \times x^3$
exponential model	$y = a \times e^{bx}$
logarithmic model	$y = a + b \times \log(x)$
reciprocal model	$y = a + \frac{b}{x}$
power model	$y = a \times x^b$

A machine learning algorithm is a method that can automatically obtain rules from data and predict unknown data according to the rules. Among many machine learning algorithms, the Back Propagation (BP) neural network model [49] has very good nonlinear fitting ability, and the Random Forest (RF) model [50] has a fast learning speed and can synthesize regression or classification functions based on discrete or continuous data sets. Therefore, based on MATLAB 2016a, a three-layer BP neural network model and an RF model were constructed for training, and the model with good performance was used as the output of the machine learning model for the inversion of water quality parameters. The accuracy of each model was mainly evaluated via a coefficient of determination (R^2), the mean absolute percentage error (MAPE) (Equation (9)) and the root mean squared error (RMSE).

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right|, \quad (9)$$

where n is the number of sampling points, y_i is the measured value and \hat{y}_i is the estimated value.

In the process of modeling, accurate selection of characteristic bands is the basis for obtaining high inversion accuracy. We calculated the correlation coefficients of reflectance and band combinations (such as addition, difference, ratio and NDVI pattern) with water quality parameters ($\text{NH}_3\text{-N}$, COD and TP) and selected the band or band combination with high correlation as parameters for modeling. For machine learning models, too many input variables may not necessarily improve the inversion accuracy [14,18,51]. After many attempts, 4~6 bands or band combinations with high correlation with $\text{NH}_3\text{-N}$, COD and TP were selected as parameters for modeling.

The experiment mainly included four modules: feature extraction, feature selection, regression and prediction. Before constructing the regression model, we split the matching data into 55% for training and 45% for testing. The specific methodological workflow is shown in Figure 2.

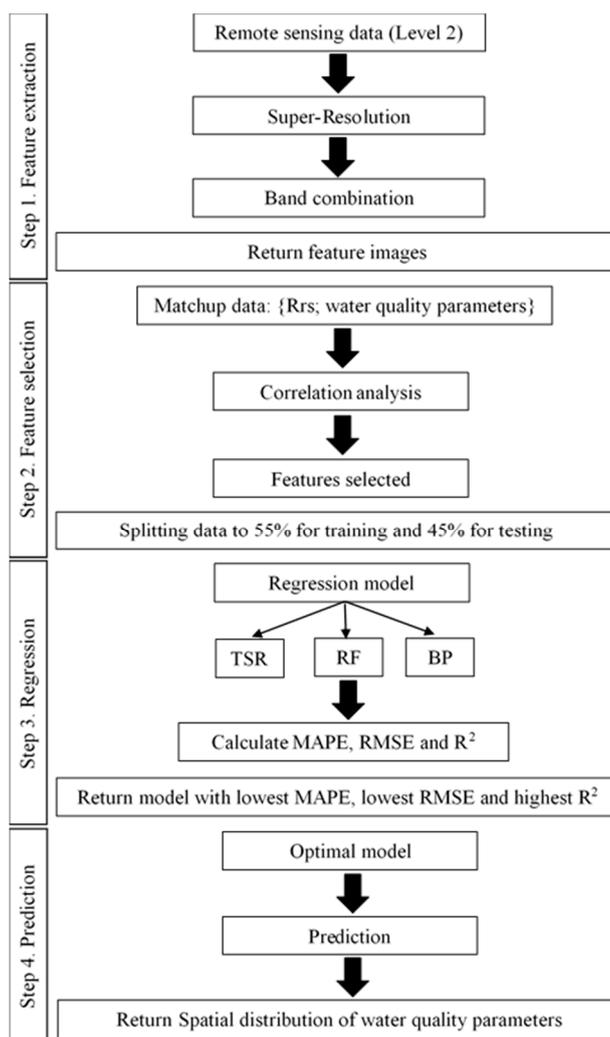


Figure 2. Estimation of water quality parameters of inland medium and small rivers based on the super-resolution algorithm.

2.3. Study Area

This paper's research area is part of the rivers in the Xinyang section of the Huaihe River Basin ($113^{\circ}45' E \sim 115^{\circ}55' E$ and $31^{\circ}23' N \sim 32^{\circ}37' N$). The study area is located on the boundary line between North and South China (Qinling–Huaihe Line), which belongs to the transition zone between subtropical and temperate monsoon climates and the transition zone between humid and semi-humid regions. The Xinyang Huaihe River is 363.5 km long. The Huaihe tributaries are dense. The tributaries on the south side of the Huaihe River account for 2/3 of the total number of tributaries, which flows into the Huaihe River in the southwest–northeast direction. The minimum width of the rivers is about ten meters. On 1 September 2016, the concentrations of $\text{NH}_3\text{-N}$, COD and TP of 41 water samples were collected in the study area. The distribution of the measured sampling points and the location of the study area are shown in Figure 3.

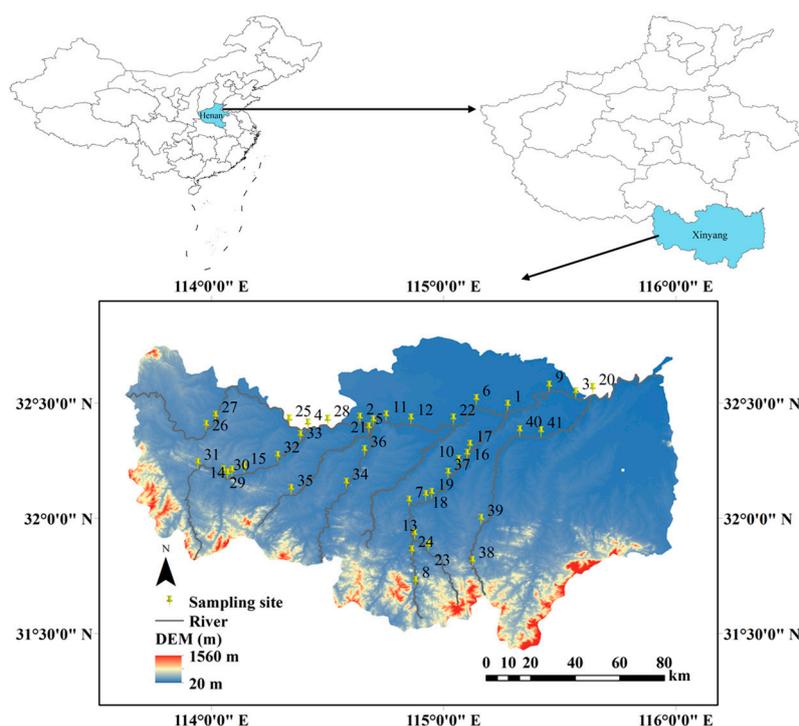


Figure 3. The study area.

2.4. Sentinel-2a Data

In this study, Sentinel-2a level 1C images obtained on 28 August 2016 were used (<https://scihub.copernicus.eu/dhus/#/home>). In this paper, the sampling time and the satellite transit time interval were short, and there was no event that obviously affected the experiment, such as heavy rainfall. Thus, the difference between the sampling time and satellite transit time had no obvious influence on the water quality parameter estimation in this paper [52]. The spatial resolution and spectral range used in this paper are shown in Table 2. To obtain the surface reflectance, Sencor-2.5.5 [53] was used to correct the Sentinel-2a image. To get a set of high-spatial-resolution images within 10 m, the super-resolution algorithm was used to enhance the atmospheric corrected data. To compare with the image enhanced by the super-resolution algorithm, the nearest-neighbor method was used to resample 60 m and 20 m images to 10 m to obtain the same size of pixels and the same reflectivity.

Table 2. Spectral bands of the Sentinel-2a L2A used in this study.

Band	Wavelength Range (nm)	Spatial Resolution (m)
Band1—Coastal aerosol	430.4~457.4	60
Band2—Blue	447.6~545.6	10
Band3—Green	537.5~582.5	10
Band4—Red	645.5~683.5	10
Band5—Vegetation Red edge	694.4~713.5	20
Band6—Vegetation Red edge	731.2~749.2	20
Band7—Vegetation Red edge	768.2~796.5	20
Band8—NIR	762.6~907.6	10
Band8a—Vegetation Red edge	848.3~881.3	20
Band9—water vapour	932.0~958.0	60
Band11—SWIR	1542.2~1756.7	20
Band12—SWIR	2081.4~2323.4	20

3. Experiments and Results

3.1. Consistency Evaluation of the Super-Resolution Image and Original Image

In order to evaluate the enhancement effect of the super-resolution algorithm, we verified the super-resolution image qualitatively and quantitatively. Figure 4 shows a detailed comparison of Band 8a and Band 9 between the original and enhanced Sentinel-2 reflection images. From the perspective of subjective vision, the resolution of images (Figure 4d,g) processed by the super-resolution algorithm was significantly improved, the image details were significantly enhanced and the edges of water bodies, buildings and bridges were further refined, which was particularly prominent in Band 9 with 60 m spatial resolution. The smaller the pixel size, the higher the pixel purity. This means the mixed pixels of the bridge across the river and the edge of the water on both sides of the river were reduced, and the spatial heterogeneity of the corresponding actual features was smaller. The images (Figure 4c,f) processed by the nearest-neighbor method were still low-resolution images in essence, although the image size was the same as that of the high-resolution image.

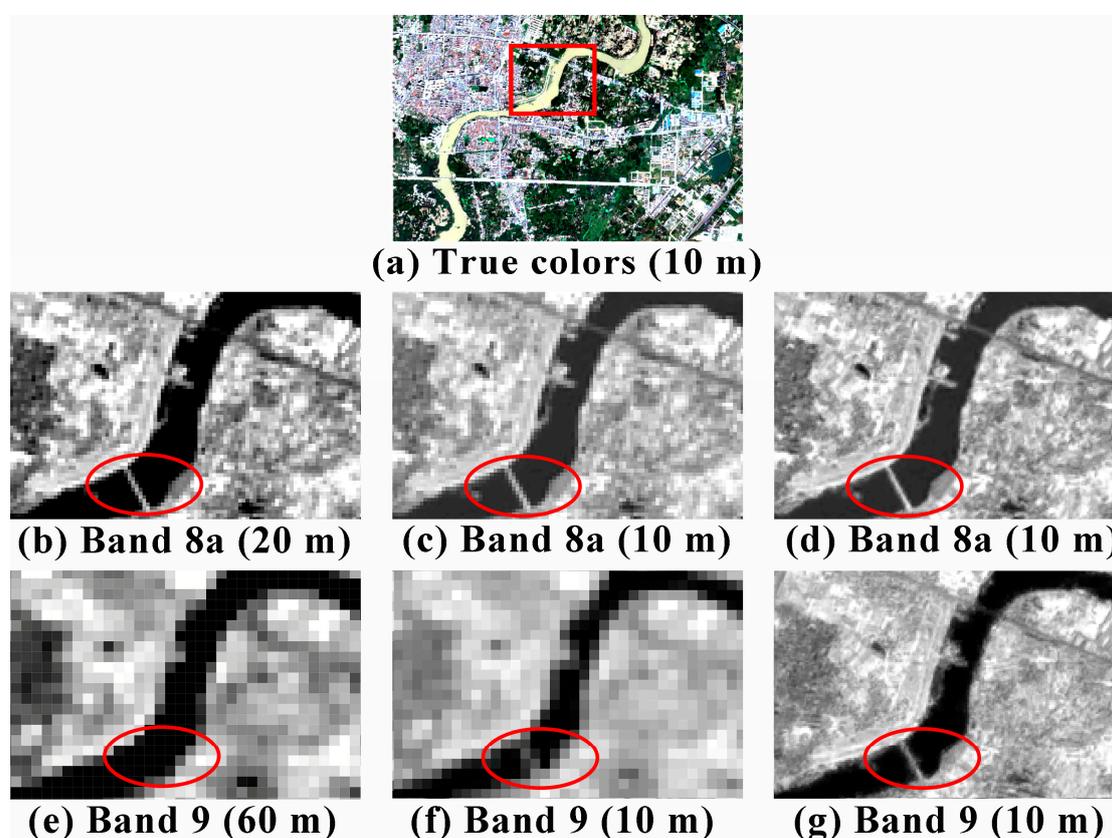


Figure 4. Comparison of the enhanced image and original image. (a) RGB true-color image, R/G/B = Band4/Band3/Band2; (b) Band 8a (20 m) of the original image; (c) Band 8a (10 m) of the resampling image by nearest neighbor method; (d) Band 8a (10 m) of the enhanced image by super-resolution algorithm; (e) Band 9 (60 m) of the original image; (f) Band 9 (10 m) of the resampling image by nearest neighbor method; (g) Band 9 (10 m) of the enhanced image by super-resolution algorithm.

We randomly selected 1000 samples from the study area to quantitatively analyze the changes in pixel values before and after the super-resolution algorithm, as shown in Figure 5. We found that the reflectivity of the super-resolution image is highly correlated with that of the original image. The R of each band before and after the super-resolution algorithm enhancement was above 0.89. The EA values of reflectance estimation accuracy of 20 m spatial resolution bands were all above 83%, and that of 60 m

spatial resolution bands was above 72%. This means that the surface reflectance in the super-resolution image is highly consistent with that of the original image.

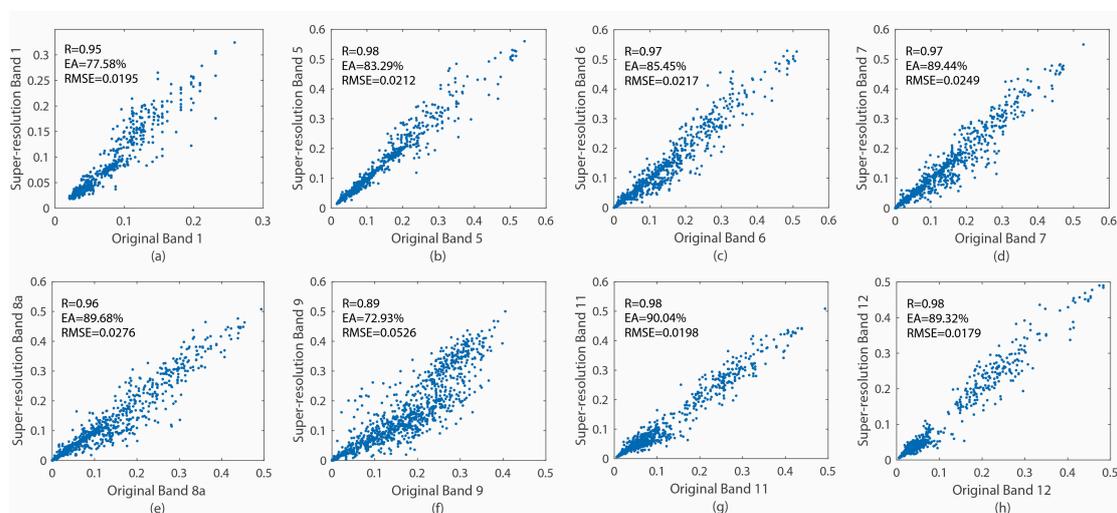


Figure 5. Reflectance consistency analysis of the super-resolution image and the original image. The X-axis is the original image reflectivity value, and the Y-axis is the super-resolution image reflectivity value.

We can therefore conclude that the super-resolution algorithm can not only enhance the details of the image and reduce the spatial heterogeneity of pixels, but also maintain the invariance of the spectrum, which provides the possibility to obtain higher-precision water quality inversion results.

3.2. Correlation Analysis among $\text{NH}_3\text{-N}$, COD and TP with Sentinel-2

The correlation among $\text{NH}_3\text{-N}$, COD and TP with 276 groups composed of a single band or a band combination was calculated. Four combination modes were adopted: (A/B), (A–B), (A+B) and (A–B)/(A+B), where A and B represent two different bands. The R value with the highest correlation coefficient among the four combinations was used as the R value between the two bands to draw the correlation matrix (Figure 6). From the statistical results of these three correlation coefficient matrices, it could be seen that the correlation coefficients of band combinations were quite different. We found that the correlation between reflectance and $\text{NH}_3\text{-N}$, COD and TP increased after super-resolution enhancement. It also showed that the inversion accuracy of $\text{NH}_3\text{-N}$, COD and TP could be improved by the enhancement of the Sentinel-2 image using the super-resolution algorithm.

From Figure 6a, it can be seen that the combinations containing Band 5 have a high correlation with $\text{NH}_3\text{-N}$, among which the combination of Band 5 and Band 8 has the highest correlation with $\text{NH}_3\text{-N}$ and has a correlation coefficient reaching 0.78. It is speculated that Band 2, Band 3, Band 4, Band 5, Band 6, Band 7, Band 8 and Band 8a are important bands for $\text{NH}_3\text{-N}$ inversion. The combination of the visible band and the red edge band has a high correlation with COD (Figure 6c). The combination of Band 4 and Band 5 can significantly improve the correlation between COD and spectral information, and its correlation coefficient can reach 0.53. It can be inferred that Band 2, Band 3, Band 4, Band 5, Band 6, Band 11 and Band 12 are important bands for COD inversion. The correlation coefficient between TP and spectral reflectance was generally high (Figure 6e). The correlation coefficient of TP with the combination of Band 1 and Band 4 was the highest, and the correlation coefficient was 0.81. It is speculated that Band 1, Band 2, Band 3, Band 4, Band 5, Band 7, Band 11 and Band 12 are important bands for TP inversion.

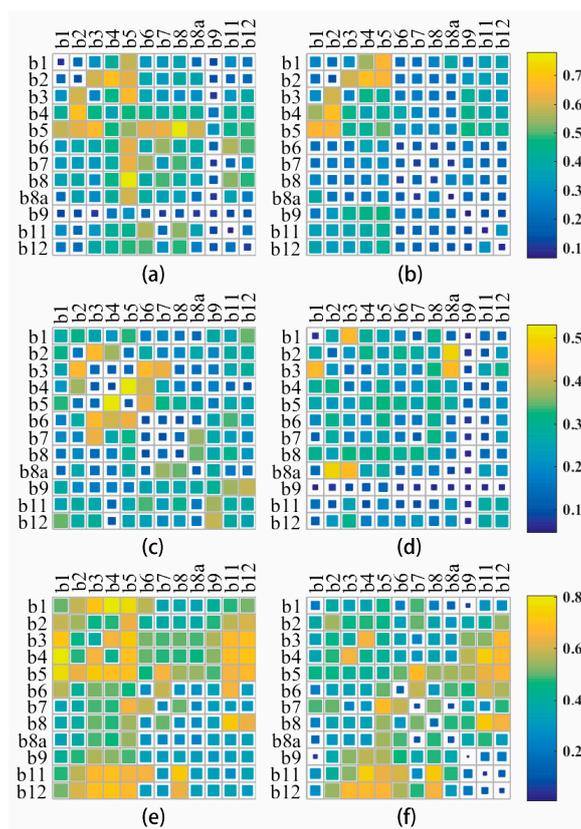


Figure 6. The correlation coefficient matrix of water quality parameters with bands and band combinations. The X and Y axes are the respective bands of the image: (a,c,e) super-resolution image; (b,d,f) resampling image by nearest-neighbor method. (a,b) NH₃-N; (c,d) Chemical Oxygen Demand (COD); (e,f) Total Phosphorus (TP). Relation of $0.6 > R > 0.4$ is significant at the 0.05 level, and the relation of $R > 0.6$ is significant at the 0.01 level.

3.3. Accuracy Comparison of Inversion Models

Based on the original image and the super-resolution image, the statistical regression models of NH₃-N, COD and TP were constructed. Twenty-three samples were used to train the models, and the remaining eighteen samples were selected for verification. The accuracy of each model method was evaluated by comparing the measured values with the estimated values. The detailed information on the optimal model inversion results is shown in Figure 7. For the super-resolution images, the optimal results obtained by the statistical regression method are shown in Figure 7a–c. The R² between the measured NH₃-N and the estimated NH₃-N using Band 5 and Band 8 was 0.67, the RMSE was 0.161 and the MAPE was 25.88%. The R² between the measured COD and the estimated COD using Band 4 and Band 5 was 0.53, the RMSE was 0.546 and the MAPE was 3.36%. The R² between the measured TP and the estimated TP using Band 1 and Band 4 was 0.60, the RMSE was 0.034, and the MAPE was 24.28%. For the original image, the optimal results based on statistical regression method are shown in Figure 7d–f. The R² of measured NH₃-N and estimated NH₃-N using Band 2 and Band 4 was 0.61, the RMSE was 0.177, and the MAPE was 29.33%. The R² of measured COD and estimated COD using Band 2 and Band 8a was 0.26, the RMSE was 0.756 and the MAPE was 4.62%. The R² of measured TP and estimated TP using Band 4 and Band 11 was 0.69, the RMSE was 0.032 and the MAPE was 30.58%. We found that when using the same inversion method, the accuracy of COD and NH₃-N retrieved from the super-resolution image was higher than from the original image. Although the coefficient of determination R² between the measured TP and the estimated TP was reduced, the average relative percentage error MAPE was significantly reduced by 6.3%. In general, the super-resolution image is better for water quality inversion than the original Sentinel-2 image. In addition, it was found that

in the process of constructing the statistical regression model, the correlation coefficient between the TP concentration value retrieved from the reciprocal model and the measured value data was higher, the error was smaller and the cubic function model was more suitable for the inversion of $\text{NH}_3\text{-N}$ and COD.

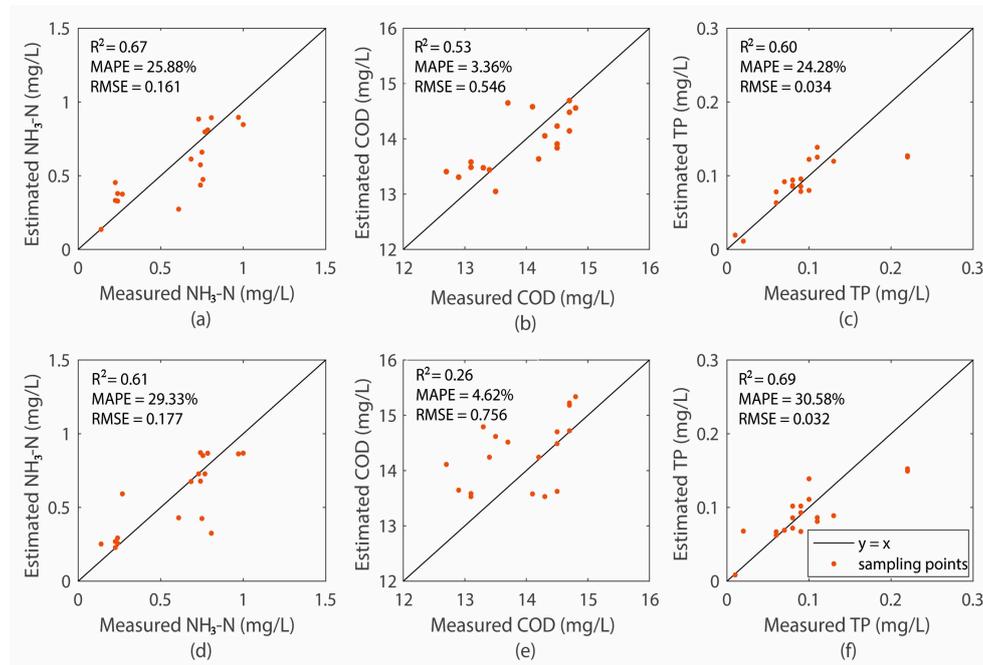


Figure 7. Accuracy assessment results of $\text{NH}_3\text{-N}$, COD and TP estimated by the statistical regression models using super-resolution images (a–c) and original images (d–f). The X-axis is measured data, and the Y-axis is estimated data. (a,b,d,f) used the cubic function model; (c,f) used the reciprocal model.

In this paper, a machine learning inversion model of $\text{NH}_3\text{-N}$, COD and TP was constructed based on super-resolution images. The optimal inversion results based on the machine learning method are shown in Figure 8. The R^2 between the measured $\text{NH}_3\text{-N}$ and estimated $\text{NH}_3\text{-N}$ using Band 2, Band 3, Band 4, Band 5 and Band 8 was 0.74, the RMSE was 0.149 and the MAPE was 22.68%. The R^2 between the measured COD and estimated COD values using Band 2, Band 3, Band 4, Band 5, Band 6 and Band 7 was 0.60, RMSE was 0.476 and MAPE was 2.99%. The R^2 between measured TP and estimated TP using Band 1, Band 3, Band 4, Band 5, Band 11 and Band 12 was 0.81, the RMSE was 0.028 and the MAPE was 21.93%. The results showed that the machine learning method has greater advantages in estimating the concentration of water quality parameters of complex inland rivers.

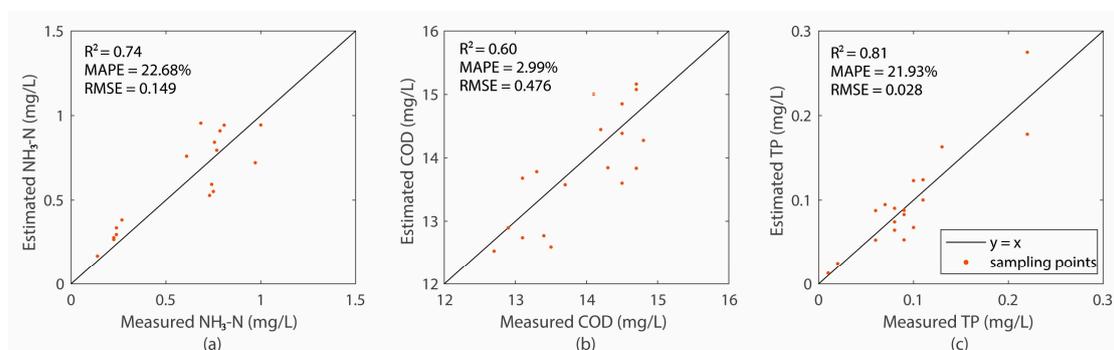


Figure 8. Accuracy assessment results of $\text{NH}_3\text{-N}$, COD and TP estimated based on the machine learning model. The X-axis is the measured data, and the Y-axis is the estimated data. (a) Comparison of measured and estimated $\text{NH}_3\text{-N}$; (b) comparison of measured and estimated COD; (c) comparison of measured and estimated TP.

3.4. Spatial Distribution of Water Quality Parameters

In order to further verify the applicability of the model, the machine learning model was applied to remote sensing images. The spatial distribution of NH₃-N, COD and TP in the Xinyang section of Huaihe River Basin is shown in Figure 9. Most of the water bodies in the study area were in a healthy state, which is consistent with the results of the local water resources bulletin. It can be seen that most of the ammonia nitrogen concentrations in the study area were within the limits of the three indicators (Figure 9a), and only the southeast part of the study area had mild COD pollution (Figure 9b). In addition, the concentration of total phosphorus in some sections of the Huanghe River was relatively high, and the pollution was serious (Figure 9c). The nutrient C:N:P required for the growth and reproduction of microorganisms is 100:5:1, and microorganisms can obtain C and N from nature, while P is almost completely supplied by human beings (there is no urban sewage with phosphorus removal or poor phosphorus removal and irrigation water flowing into water body on the land with phosphate fertilizer). It can be seen from the geographical location of the river section that the river passes through residential areas and that the discharge of domestic sewage directly affects the water quality of the river section, resulting in a TP concentration of the river section higher than that of other areas.

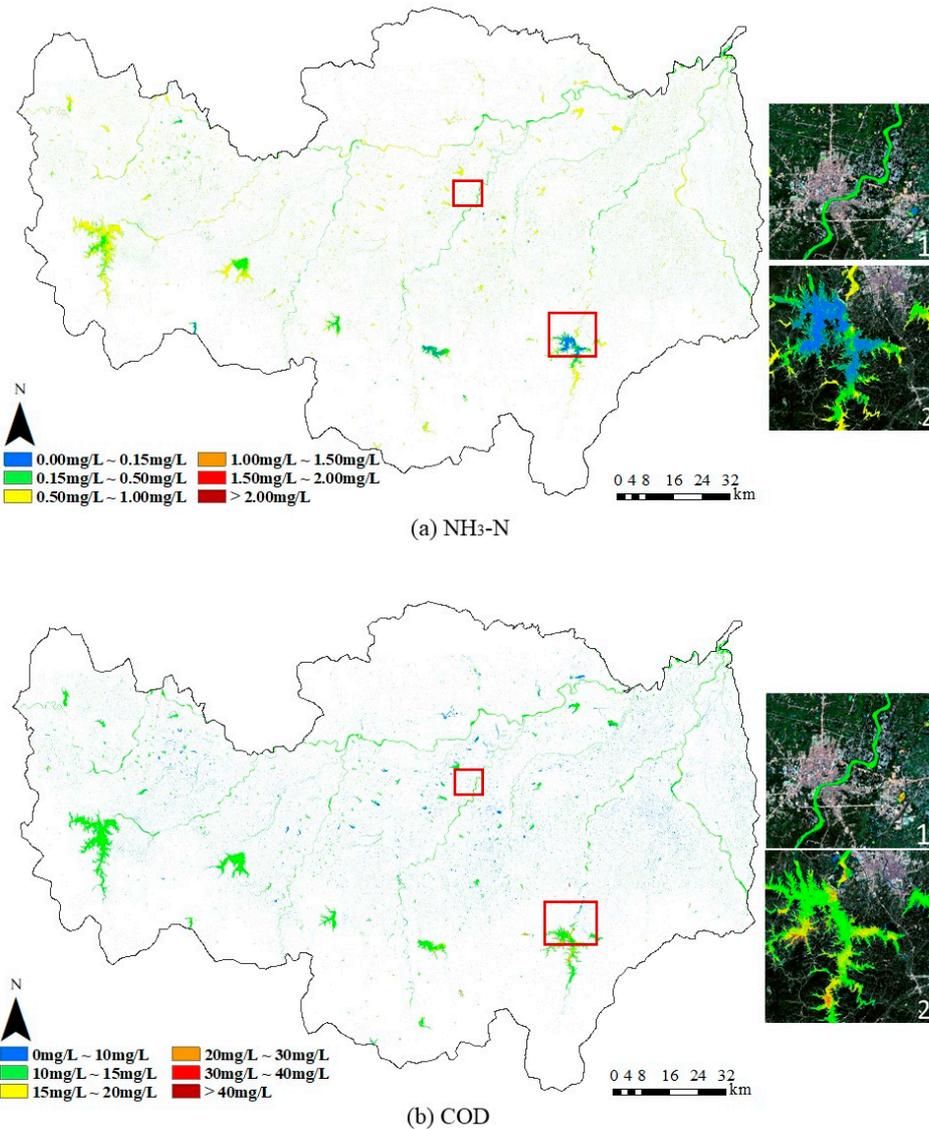


Figure 9. Cont.

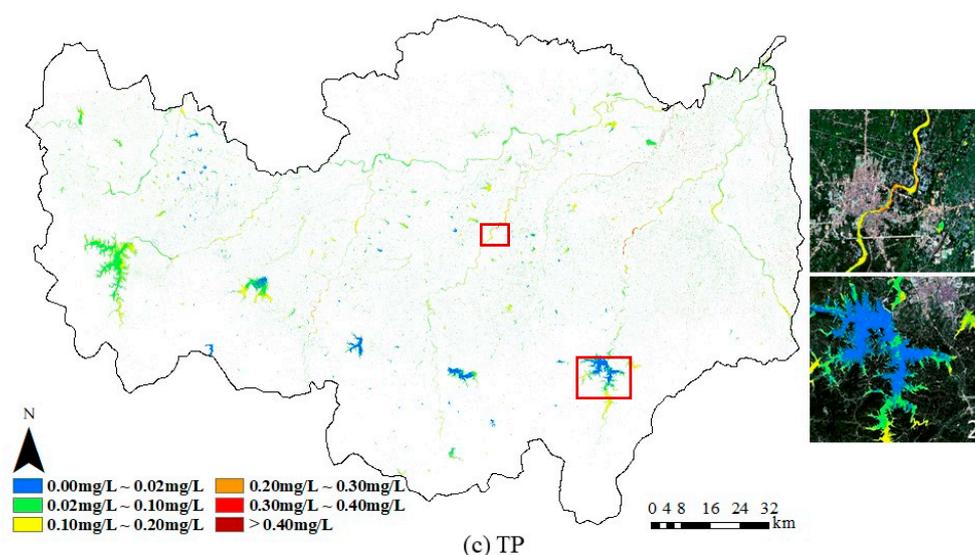


Figure 9. The spatial distribution of $\text{NH}_3\text{-N}$, COD and TP in the Xinyang section of Huaihe River Basin; 1 is part of the Huanghe River; 2 is part of the Nianyushan reservoir. (a) The spatial distribution of $\text{NH}_3\text{-N}$; (b) the spatial distribution of COD; (c) the spatial distribution of TP.

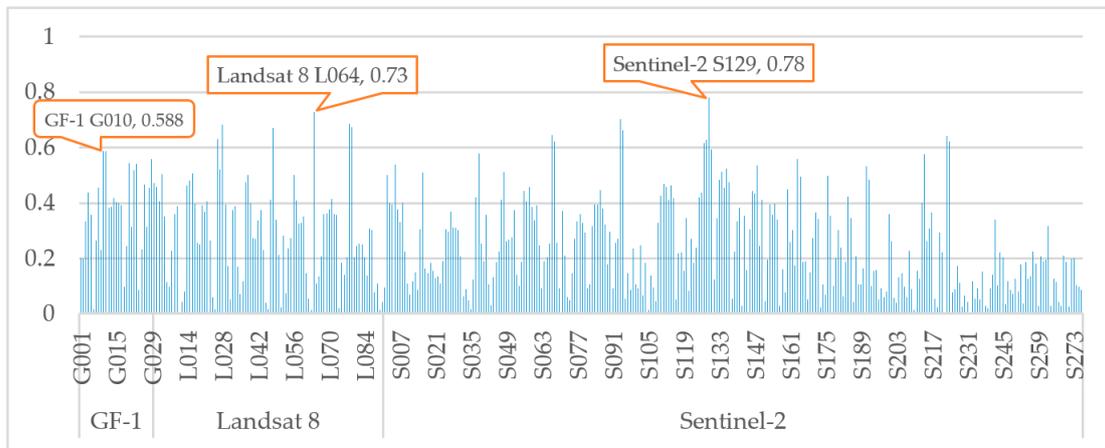
4. Discussion

Differences in water component content cause significant differences in reflectance in a certain wavelength range, which is the theoretical basis of quantitative inversion of water quality parameters by remote sensing. Whether the direct inversion method or the indirect inversion method is used, its essence is based on the correlation between water quality parameters and remote sensing reflectance to carry out quantitative inversion. Our research showed that Band 2, Band 4, Band 5 and Band 8 are sensitive bands for $\text{NH}_3\text{-N}$ (Figure 6a). Among them, the sensitivity of $\text{NH}_3\text{-N}$ to the near-infrared band shows the frequency-doubled or combined frequency characteristics of nitrogen-containing functional groups. In addition, COD has a high correlation with Band 2, Band 3, Band 4, Band 5 and Band 8a (Figure 6b). The sensitivity of COD to Band 3 and Band 8a shows the combined frequency of the vibrations of hydrogen-containing groups (O-H, C-H) in organic molecules and their frequency-doubled absorption characteristics. The red band and the red edge band are the sensitive bands of Chl-a [20], and the metabolism of phytoplankton in the water body will also release organic matter. Chl-a can reflect the level of phytoplankton inventory. From this process, COD is the passive factor of Chl-a. TP affects and controls the growth and reproduction of planktonic algae to different degrees, and some studies have shown that it is closely related to chlorophyll content [7]. In this paper, Band 1, Band 4 and Band 11 were used to estimate TP concentration in both the statistical regression method and the machine learning method, which provides a useful reference for monitoring TP concentration with band data of the multi-spectral imager. The rivers in the study area belong to a flowing water body, and the water composition is not uniform. Therefore, some sensitive bands of $\text{NH}_3\text{-N}$, COD and TP in Sentinel-2 satellite are slightly different from those of previous studies.

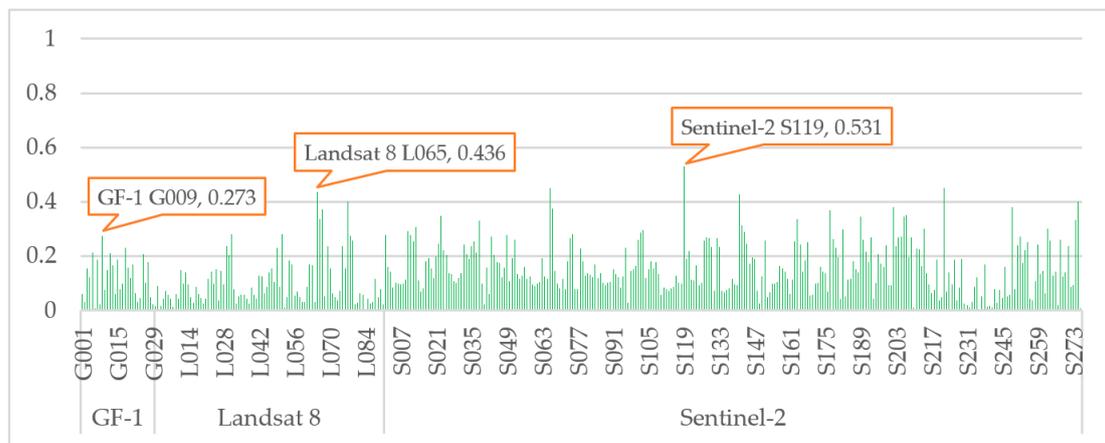
In general, the method for remote sensing inversion of water quality parameters using satellite images is to establish the remote sensing inversion model of water quality parameters by analyzing the relationship between water quality parameters and remote sensing reflectance. In the actual research process, it is often done by establishing an effective connection between the point data obtained from field sampling and the remote sensing pixel surface data with different spatial resolutions. The remote sensing observations value is a surface datum, and the sample point measurement value is a point datum. The conversion between them will inevitably cause certain errors, which is also the general uncertainty of quantitative remote sensing inversion. Using high-spatial-resolution images is one way to reduce this uncertainty. This is the significance of our research. We focus on fully

exploiting the potential of existing data without increasing the cost of satellites and producing a series of high-resolution image data by using super-resolution algorithm for water quality inversion.

To prove the advantages of Sentinel-2 satellite used in this paper, for water quality monitoring, we processed GF-1 satellite images and Landsat 8 satellite images, which were at a similar field water sample collection time. Distribution of the correlation coefficient between different band combinations and water quality parameters (NH₃-N, COD and TP) was calculated (Figure 10). The central wavelength and bandwidth of the same band of different data sources were different, so for different satellite images, the correlation between the same band or band combination and water quality parameters is different. Compared with other medium and high-resolution satellite data (GF-1 and Landsat 8), the band combination of Sentinel-2 data shows a relatively high correlation with water quality parameters (NH₃-N, COD, TP) and has great potential in water quality parameter inversion.

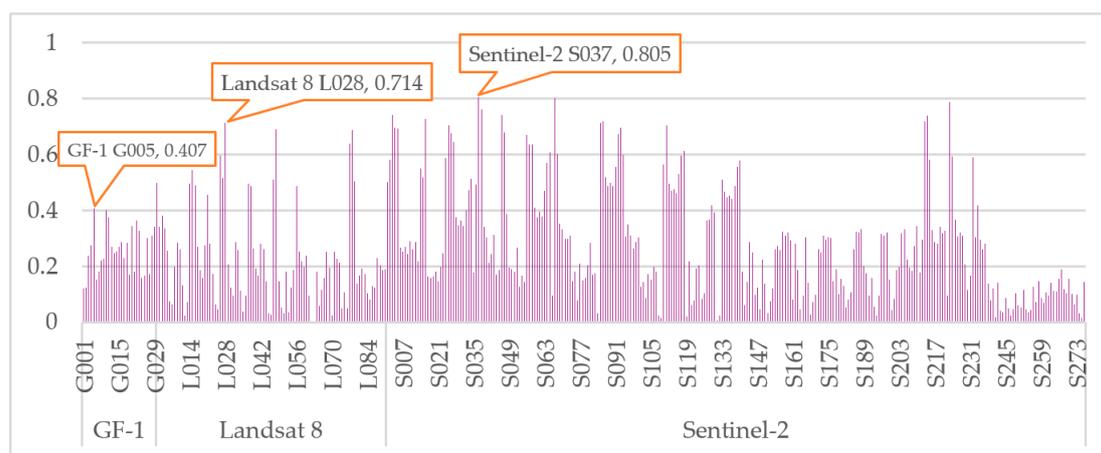


(a) NH₃-N



(b) COD

Figure 10. Cont.



(c) TP

Figure 10. Distribution of the correlation coefficient between different band combinations and water quality parameters ($\text{NH}_3\text{-N}$, COD and TP). The X-axis is the number of single bands and band combinations. G001, G002, \dots , G028 and G029 are the numbers of the four multispectral bands and band combinations of the GF-1 satellite. L001, L002, \dots , L090 and L091 are the numbers of the seven multispectral bands and band combinations of the Landsat 8 satellite. S001, S002, \dots , S275 and S276 are the numbers of the seven multispectral bands and band combinations of Sentinel-2 satellite. The Y-axis is the R value. The labels in the figure are the band combination numbers and R values of the highest correlation between different satellite sensors and $\text{NH}_3\text{-N}$, COD and TP. (a) $\text{NH}_3\text{-N}$; (b) COD; (c) TP.

Our results show that the Sentinel-2 processed by the super-resolution algorithm has advantages in the inversion of water quality parameters. Sentinel-2 provides the only satellite data that can obtain multispectral bands and resolution up to 10 m for free. The super-resolution algorithm can improve the spatial resolution of its low-resolution band to 10 m without increasing the cost and can improve the spatial resolution while maintaining the spectral consistency so that the product generated by the super-resolution algorithm has higher precision. In the application of quantitative remote sensing in water quality monitoring, the smallest monitoring unit is a pixel. The purity of a pixel has an important impact on the final results, especially when the monitoring range is inland rivers. Spatial heterogeneity of water quality parameters results in marked scaling effects, and it influences precision in the retrieval of water quality parameters from remote sensing images [54]. Therefore, the selection of high-resolution satellite data is helpful to eliminate the error caused by scale effects in the inversion process. Compared with lakes with large areas of water, inland rivers are usually narrow and long. The existence of the river edge or a bridge across the river will also reduce the purity of the corresponding pixels, and the existence of mixed pixels will bring more uncertainty to the river water quality monitoring. In addition, the water quality near the sewage outlet in the inland river is worse than that in other areas. The spatial heterogeneity of water quality parameters corresponding to the sewage outlet in the 60 m, 30 m and 20 m spatial-resolution remote-sensing data is greater than that in the 10 m spatial-resolution remote-sensing data. Therefore, 10 m spatial resolution images can reduce the uncertainty caused by these factors and then improve the accuracy of water quality parameter inversion and inland river water quality monitoring.

5. Conclusions

This study evaluated the performance of the super-resolution algorithm in monitoring several water quality parameters in medium and small-sized rivers, such as the Xinyang section of the Huaihe River Basin. The relationship between the field measured data and the Sentinel-2 band was studied, and three indexes ($\text{NH}_3\text{-N}$, COD and TP) were estimated from Sentinel-2 data. It can be concluded from the present study that Sentinel-2 is an effective tool to establish a cheap and effective routine

monitoring method for inland river water quality. The super-resolution algorithm can effectively improve the quality of Sentinel-2 data, weaken the negative impact of mixed pixels on water quality monitoring, and improve the accuracy of water quality prediction, which has great application potential in quantitative remote sensing of water quality. The machine learning model constructed in this paper can be used as the best method to predict water quality in the Xinyang section of Huaihe River Basin and can provide data support for water quality monitoring and water pollution prevention and control in Huaihe River Basin. Agencies can combine super-resolution technology with remote sensing data as an alternative to collecting and processing information on inland surface water quality. In the future, the super-resolution algorithm can be extended to other satellite remote sensing images with multispectral bands and different spatial resolutions. In terms of long-term dynamic water quality monitoring, multiple data sources can be combined.

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

Funding: This study was supported by the National Key R&D Program of China (Grant No. 2018YFC0406501), China Postdoctoral Science Foundation (Grant No. 2019M662534) and National natural science foundation (Grant No. 42001405).

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

References

1. Marin, S.; Erika, L.; Oana, C.; Ramona, S.L.; Marius, R.; Ferenc, P.; Mihaela, S. Assessment of Availability and Human Health Risk Posed by Arsenic Contaminated Well Waters from Timis-Bega Area, Romania. *J. Anal. Methods Chem.* **2017**, *2017*, 3037651.
2. Hussain, S.; Shaikh, S.; Farooqui, M. COD reduction of waste water streams of active pharmaceutical ingredient—Atenolol manufacturing unit by advanced oxidation-Fenton process. *J. Saudi Chem. Soc.* **2013**, *17*, 199–202. [[CrossRef](#)]
3. Shao, M.; Tang, X.; Zhang, Y.; Li, W. City clusters in China: Air and surface water pollution. *Front. Ecol. Environ.* **2006**, *4*, 353–361. [[CrossRef](#)]
4. Gong, S.; Huang, J.; Li, Y.; Lu, W.; Wang, H.; Wang, G. Preliminary Exploring of Hyperspectral Remote Sensing Experiment for Nitrogen and Phosphorus in Water. *Spectrosc. Spectr. Anal.* **2008**, *28*, 839–842.
5. Wang, X.; Fu, L.; He, C. Applying support vector regression to water quality modelling by remote sensing data. *J. Remote Sens.* **2011**, *32*, 8615–8627. [[CrossRef](#)]
6. Chen, Y.; Huang, C.; Zhang, L.; Qiao, N. Spectral Characteristics Analysis and Remote Sensing Retrieval of COD Concentration. *Spectrosc. Spectr. Anal.* **2020**, *40*, 824–830.
7. Sun, F.; Sun, W.; Chen, J.; Gong, P. Comparison and improvement of methods for identifying waterbodies in remotely sensed imagery. *Int. J. Remote Sens.* **2012**, *33*, 6854–6875. [[CrossRef](#)]
8. Huang, M.; Xing, X.; Qi, X.; Yu, W.; Zhang, Y. Identification mode of chemical oxygen demand in water based on remotely sensing technique and its application. In Proceedings of the IEEE International Geoscience & Remote Sensing Symposium, Boston, MA, USA, 6–11 July 2008.
9. Mushtaq, F.; Ghosh, M. Remote Estimation of Water Quality Parameters of Himalayan Lake (Kashmir) using Landsat 8 OLI Imagery. *Geocarto Int.* **2016**, *32*. [[CrossRef](#)]
10. Tao, J.; Zhang, Z.; Yu, W. Monitoring Taihu water quality by using high resolution satellite image. In Proceedings of the International conference on Geoinformatics, Prague, Czech Republic, 12 August 2011; pp. 1–4.
11. Tu, M.C.; Patricia, S.; Filippi, A.M.; Fei, L. Hybrid forward-selection method-based water-quality estimation via combining Landsat TM, ETM+, and OLI/TIRS images and ancillary environmental data. *PLoS ONE* **2018**, *13*, e0201255. [[CrossRef](#)]
12. Swain, R.; Sahoo, B. Mapping of heavy metal pollution in river water at daily time-scale using spatio-temporal fusion of MODIS-aqua and Landsat satellite imageries. *J. Environ. Manag.* **2017**, *192*, 1–14. [[CrossRef](#)]

13. Salem, S.; Marie, S.; Hiroto, H.; Hyungjun, K.; Komatsu, K.; Kazuo, O.; Taikan, O. Evaluation of MERIS Chlorophyll-a Retrieval Processors in a Complex Turbid Lake Kasumigaura over a 10-Year Mission. *Remote Sens.* **2017**, *9*, 1022. [[CrossRef](#)]
14. Blix, K.; Eltoft, T. Machine Learning Automatic Model Selection Algorithm for Oceanic Chlorophyll-a Content Retrieval. *Remote Sens.* **2018**, *10*, 775. [[CrossRef](#)]
15. Zhang, Y.; Shi, K.; Zhang, Y.; Moreno-Madriñán, M.J.; Zhu, G.; Zhou, Y.; Yao, X. Long-term change of total suspended matter in a deep-valley reservoir with HJ-1A/B: Implications for reservoir management. *Environ. Sci. Pollut. Res. Int.* **2019**, *26*, 3041–3054. [[CrossRef](#)]
16. Cao, Y.; Ye, Y.; Zhao, H.; Jiang, Y.; Wang, H.; Shang, Y.; Wang, J. Remote sensing of water quality based on HJ-1A HSI imagery with modified discrete binary particle swarm optimization-partial least squares (MDBPSO-PLS) in inland waters: A case in Weishan Lake. *Ecol. Inform.* **2018**, *44*, 21–32. [[CrossRef](#)]
17. Isidro, C.M.; McIntyre, N.; Lechner, A.M.; Callow, I. Quantifying suspended solids in small rivers using satellite data. *Sci. Total Environ.* **2018**, *634*, 1554–1562. [[CrossRef](#)]
18. Yong, H.K.; Im, J.; Ha, H.K.; Choi, J.K.; Ha, S. Machine learning approaches to coastal water quality monitoring using GOCI satellite data. *Mapp. Ence Remote Sens.* **2014**, *51*, 158–174.
19. Shang, P.; Shen, F. Atmospheric Correction of Satellite GF-1/WFV Imagery and Quantitative Estimation of Suspended Particulate Matter in the Yangtze Estuary. *Sensors* **2016**, *16*, 1997. [[CrossRef](#)]
20. Pahlevan, N.; Smith, B.; Schalles, J.; Binding, C.; Cao, Z.; Ma, R.; Alikas, K.; Kangro, K.; Gurlin, D.; Hà, N.; et al. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. *Remote Sens. Environ.* **2020**, *240*, 111604. [[CrossRef](#)]
21. Thu, H.N.T.; Phuong, T.N.T.; Katsuaki, K.; Trong, N.M. Selecting the Best Band Ratio to Estimate Chlorophyll-a Concentration in a Tropical Freshwater Lake Using Sentinel 2A Images from a Case Study of Lake Ba Be (Northern Vietnam). *Isprs. Int. J. Geo. Inf.* **2017**, *6*, 290.
22. Wang, C.; Li, W.; Chen, S.; Li, D.; Wang, D.; Liu, J. The spatial and temporal variation of total suspended solid concentration in Pearl River Estuary during 1987–2015 based on remote sensing. *Sci. Total Environ.* **2018**, *618*, 1125–1138. [[CrossRef](#)]
23. Wang, X.; Yang, W. Water quality monitoring and evaluation using remote sensing techniques in China: A systematic review. *Ecosyst. Health Sustain.* **2019**, *5*, 47–56. [[CrossRef](#)]
24. Gholizadeh, M.H.; Melesse, A.M.; Reddi, L.N. A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors* **2016**, *16*, 1298. [[CrossRef](#)]
25. Drusch, M.; Bello, U.D.; Carlier, S.; Colin, O.; Fernandez, V.; Gascon, F.; Hoersch, B.; Isola, C.; Laberinti, P.; Martimort, P. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote Sens. Environ.* **2012**, *120*, 25–36. [[CrossRef](#)]
26. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 295–307. [[CrossRef](#)] [[PubMed](#)]
27. Shen, H.F.; Li, P.X.; Zhang, L.P.; Wang, Y. Overview on super resolution image reconstruction. *Opt. Tech.* **2009**, *35*, 194–203.
28. Shen, M.; Xue, P.; Wang, C. Down-Sampling Based Video Coding Using Super-Resolution Technique. *Circuits Syst. Video Technol. IEEE Trans.* **2011**, *21*, 755–765. [[CrossRef](#)]
29. De Santis, P.; Gori, F. On an Iterative Method for Super-resolution. *Opt. Acta Int. J. Opt.* **1975**, *22*, 691–695. [[CrossRef](#)]
30. Gerchberg, R. Super-resolution through Error Energy Reduction. *Opt Acta Int. J. Opt.* **1974**, *21*, 709–720. [[CrossRef](#)]
31. Hardie, R.C.; Barnard, K.J.; Armstrong, E.E. Joint MAP registration and high-resolution image estimation using a sequence of undersampled images. *IEEE Trans. Image Process.* **1997**, *6*, 1621–1633. [[CrossRef](#)]
32. Xiao, C.; Yu, J.; Xue, Y. A Novel Fast Algorithm for MAP Super-Resolution Image Reconstruction. *J. Comput. Res. Dev.* **2009**, *46*, 872–880.
33. Nuno-Maganda, M.A.; Arias-Estrada, M.O. Real-time FPGA-based architecture for bicubic interpolation: An application for digital image scaling. In Proceedings of the International Conference on Reconfigurable Computing and FPGAs (ReConFig'05), Puebla City, Mexico, 28–30 September 2005.
34. Irani, M.; Peleg, S. Improving resolution by image registration. *Cogip Graph. Models Image Process.* **1991**, *53*, 231–239. [[CrossRef](#)]

35. Schultz, R.R.; Stevenson, R.L. Bayesian estimation of subpixel-resolution motion fields and high-resolution video stills. In Proceedings of the International Conference on Image Processing, Santa Barbara, CA, USA, 26–29 October 1997; pp. 62–65.
36. Zhang, H.; Zhang, Y.; Li, H.; Huang, T.S. Generative Bayesian Image Super Resolution With Natural Image Prior. *IEEE Trans. Image Process.* **2012**, *21*, 4054–4067. [[CrossRef](#)]
37. Stark, H.; Oskoui, P. High-resolution image recovery from image-plane arrays, using convex projections. *J. Opt. Soc. Am. A Opt. Image Sci.* **1989**, *6*, 1715–1726. [[CrossRef](#)] [[PubMed](#)]
38. Baker, S.; Kanade, T. Limits on super resolution and how to break them. *IEEE TPAMI. IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 1167–1183. [[CrossRef](#)]
39. Gillespie, A.R.; Kahle, A.B.; Walker, R.E. Color enhancement of highly correlated images. II-Channel ratio and ‘chromaticity’ transformation techniques. *Remote Sens. Environ.* **1987**, *22*, 343–365. [[CrossRef](#)]
40. Nikolakopoulos, K.G. Comparison of Nine Fusion Techniques for Very High Resolution Data. *Photogramm. Eng. Remote Sens.* **2008**, *74*, 647–659. [[CrossRef](#)]
41. Wu, B.; Fu, Q.; Sun, L.; Wang, X. Enhanced hyperspherical color space fusion technique preserving spectral and spatial content. *J. Appl. Remote Sens.* **2015**, *9*, 097291. [[CrossRef](#)]
42. Brodu, N. Super-resolving multiresolution images with band-independant geometry of multispectral pixels. *Arxiv Comput. Vis. Pattern Recognit.* **2017**, *55*, 4610–4617.
43. Settle, J.J.; Drake, N. Linear mixing and the estimation of ground cover proportions. *Int. J. Remote Sens.* **1993**, *14*, 1159–1177. [[CrossRef](#)]
44. Ahlgren, P.; Jarneving, B.; Rousseau, R. Requirements for a cocitation similarity measure, with special reference to Pearson’s correlation coefficient. *J. Assoc. Inf. Ence Technol.* **2014**, *54*, 550–560. [[CrossRef](#)]
45. Lin, C.H.; Bioucas-Dias, J.M. An Explicit and Scene-Adapted Definition of Convex Self-Similarity Prior With Application to Unsupervised Sentinel-2 Super-Resolution. *IEEE Trans. Geoenve Remote Sens.* **2020**, *58*, 3352–3365. [[CrossRef](#)]
46. Zhang, M.; Su, W.; Fu, Y.; Zhu, D.; Yao, C. Super-resolution enhancement of Sentinel-2 image for retrieving LAI and chlorophyll content of summer corn. *Eur. J. Agron.* **2019**, *111*, 125938. [[CrossRef](#)]
47. Torbick, N.; Hession, S.; Hagen, S.; Wiangwang, N.; Becker, B.L.; Qi, J. Mapping inland lake water quality across the Lower Peninsula of Michigan using Landsat TM imagery. *J. Remote Sens.* **2013**, *34*, 7607–7624. [[CrossRef](#)]
48. Kutser, T.; Arst, H.; Miller, T.; Kaarmann, L.; Milius, A. Telespectrometrical estimation of water transparency, chlorophyll-a and total phosphorus concentration of Lake Peipsi. *Int. J. Remote Sens.* **1995**, *16*, 3069–3085. [[CrossRef](#)]
49. Pineda, F.J. Generalization of back-propagation to recurrent neural networks. *Phys. Rev. Lett.* **1987**, *59*, 2229–2232. [[CrossRef](#)]
50. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
51. Kown, Y.; Baek, S.; Lim, Y.; Pyo, J.; Ligaray, M.; Park, Y.; Cho, K. Monitoring Coastal Chlorophyll-a Concentrations in Coastal Areas Using Machine Learning Models. *Water* **2018**, *10*, 1020. [[CrossRef](#)]
52. Olmanson, L.G.; Bauer, M.E.; Brezonik, P.L. A 20-year Landsat water clarity census of Minnesota’s 10,000 lakes. *Remote Sens. Environ.* **2008**, *112*, 4086–4097. [[CrossRef](#)]
53. Louis, J.; Debaecker, V.; Pflug, B.; Main-Knorn, M.; Gascon, F. SENTINEL-2 SEN2COR: L2A Processor for Users. In Proceedings of the Living Planet Symposium, Prague, Czech Republic, 9–13 May 2016.
54. Liu, X.; Duan, H.; Ma, R. The spatial heterogeneity of water quality variables in Lake Taihu, China. *J. Lake Sci.* **2010**, *03*, 367–374.

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



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).