Article

Temporal and Spatial Characteristics of River Water Quality and Its Influence Factors in the TAIHU Basin Plains, Lower Yangtze River, China

Bin Gao, Youpeng Xu *, Zhixin Lin, Miao Lu and Qiang Wang

School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China; gaogisstudy@163.com (B.G.); linzhixin115@163.com (Z.L.); giser_lm@163.com (M.L.); wqianghy@163.com (Q.W.)
* Correspondence: xypnju@163.com; Tel.: +86-025-89685069

Abstract: Water quality pollution has been a serious problem in the Taihu Basin plains, which is a highly urbanized area in China. This study aims to detect the interannual and seasonal changes and spatial patterns of water quality in this region. Based on cluster analysis, Moran’s I, and standard deviational ellipses, the site clusters, spatial heterogeneity of water quality characteristics and identified polluted regions were clarified. Results showed that (1) water quality improved since 2002, and nutrient concentrations were lower in summer and autumn than in winter and spring. (2) The monitoring sites were divided into six clusters according to the water quality during the period from 2010 to 2014. Water quality worsened from Cluster 1 to Cluster 4. Cluster 1 sites were mostly distributed beside the Yangtze River and Taihu Lake. Cluster 4 sites were mainly located along the southeast border near Shanghai, while the remaining sites were separately distributed in the main cities. (3) A polluted region of both total nitrogen (TN) and total phosphorus (TP) was present in the southeastern part of the study area near the border from 2010 to 2014. In addition, polluted regions were most likely to form near the junctions of main cities. (4) Anthropogenic factors had greater impacts on water quality than natural factors. More attention should be given to water quality protection around impervious surface areas due to the greatest considerable effect.

Keywords: river water quality; plain area of the Lower Yangtze River; the Taihu Basin; temporal change; spatial characteristics

1. Introduction

As the most developed region in China, the urbanization of the Taihu Basin, which is located in the lower reach of the Yangtze River, has continued at a rapid pace. Influenced by anthropogenic activities (e.g., urbanization), water quality has deteriorated and flood events have occurred frequently in this area [1–3]. Cyanobacteria pollution has long been a severe problem in the Taihu Lake. The algae bloom in 2007 resulted in a shortage of drinking water for over two million people and attracted considerable attention from society. Algal pollution is mainly caused by high concentrations of nitrogen and phosphorus that originate from excessive loadings from point and nonpoint sources. The release of wastewater was still uncontrolled in the plains area of the Taihu Basin [4]. Eutrophication in the Taihu Basin has threatened the sustainable development of the region. Although water purification projects have not achieved remarkable effects, water quality has improved in recent years [5]. The reasons for these improvements are complex, and the problem of lower water quality caused by pollution is difficult to solve.

Many studies have been conducted on the aquatic environment and other environmental problems in the Taihu Basin and other basins [6–10]. Researchers have detected changes in water quality [11,12] and pollution characteristics [13] as well as their influencing factors [14–17]. In this area, pollution can be found in all parts of the water cycle, such as rainwater, soil water, surface water, and groundwater. A similar previous study conducted...
Several statistical methods have been used to analyze the characteristics of water quality or identify the influencing factors of pollution, including factor/principal component analysis (FA/PCA) [13,19,20], cluster analysis (CA) [11,21], discriminant analysis (DA) and redundancy analysis (RDA) [22]. Xia et al. identified NO$_3$-N, TN and TP as the key pollutant indexes in the Three Gorges Reservoir [21]. Moreover, parameters contributing to river water quality variation may differ among different seasons [19]. Zeinalzadeh and Rezaei found that reduced river flows, increased pollution and agricultural drainage water depletion were important influencing factors affecting considerable changes in water quality downstream of the Shahr Chai River [20]. Sener et al. considered that anthropogenic pollutants increased the values of nutrients and trace metals in the Aksu River [23]. Similarly, Wu et al. recognized that anthropogenic factors and land use were most likely responsible for the spatial variations in water quality in the Taihu Basin [24]. Impervious surfaces expansion and secondary industries contribute to an increase in population, which has a great impact on water quality [25]. Machine Learning methods, such as artificial neural networks, have also been used in water quality modeling [26]. However, few studies have detected the spatial characteristics of polluted region changes. Nitrogen and phosphorus are the main limiting nutrient elements of algal growth in the study area. The TN concentration was far beyond the value specified in the Class V national standard of surface water. P drives phytoplankton assemblages under an N-rich environment in the rivers that flow into Taihu Lake [22]. High concentrations of TP and TN are more likely to cause cyanobacterial outbreaks [5], and regions with the highest TN and TP pollution should be identified and treated based on their specific characteristics.

This study aims to analyze the temporal variations in water quality in rivers in the plain area of the Taihu Basin. Additionally, water quality spatial pattern variations based on Moran’s I (the index is described in what follows), CA, and standard deviational ellipses are illustrated. Finally, the major factors affecting water quality among artificial and natural factors are explored. The results could improve the scientific understanding of river water quality and help improve and protect the aquatic environment by identifying polluted regions for the prevention and control of cyanobacteria.

2. Materials and Methods

2.1. Study Area

The Taihu Basin is located in the southern part of Jiangsu Province, the northern part of Zhejiang Province and Shanghai. Specifically, it is in the middle of the Yangtze River Delta Region and contains hilly and plain areas (Figure 1). The area of the Taihu Basin is approximately 36,900 km$^2$, and it contains a complex river network with a river density of approximately 3.24 km/km$^2$ [27]. The study area includes the Wuchengxiyu Water Conservancy Area (WCXY) and the Yangchengdianmao Water Conservancy Area (YCDM), which are typical plain areas in the Taihu Basin. It is adjacent to Shanghai in the southeast and is located between Taihu Lake and the Yangtze River which exchange water between them. This region has an average rainfall of 1078.72 mm, and approximately 62% of the precipitation is concentrated in summer and autumn; this average is based on monitoring data collected from 2002 to 2014 (Table 1). The amount of precipitation increases from April to August and reaches the highest value in August. Both temperature and precipitation are high in summer because this area is under the control of the East Asia Summer Monsoon, which contributes to more rain and higher temperatures during the summer season. The water level is high from June to October and shows delays relative to the timing of precipitation. The highest water level appears in August, although the water level is also high in July. Natural conditions, such as precipitation and temperature, may significantly affect water quality. Because of the natural conditions of the area, anthropogenic factors in the Taihu Basin greatly affect water quality in the corresponding season. The East Asian Summer Monsoon and flat terrain of the Taihu Plain make it a highly suitable area for
2.2. Water Quality and Influencing Factors

Water samples were monitored from 2002 to 2014 by the local hydrographic offices of Jiangsu Province located in this area to support water management decisions [29]. The monitoring sites were distributed along the main rivers. The number of sites increased annually as more attention was given to water quality in the plain area of the Taihu Basin. The numbers of sites were 16, 20, 23, 34, 32, 35, 48, 67, 118, 122, 123, 121, and 120 from 2002 to 2014. Some sites were monitored monthly, while some sites were monitored every two months. Water temperature data were also monitored every month at some key sites.

Biochemical indexes (e.g., dissolved oxygen (DO), percent saturation of oxygen, potassium permanganate index (COD$_{Mn}$)) and nutrients (e.g., TP, ammonia nitrogen (NH$_4^+$-N), and TN) were selected to describe the water quality characteristics in this area.

The impervious surface and paddy land rates were obtained by interpretation from Landsat 7 TM images. Population raster data for 2014 were obtained from WorldPop (https://www.worldpop.org, accessed on 11 July 2021). Precipitation data were obtained by the interpolation from monitoring data by the kriging method. All monitoring sites corresponded to the influencing factors within a 1000 m buffer radius. Deng [30] and Yang et al. [31] adapted this radius for use in this region, and research by Yang indicated a more significant correlation when the spatial scale was increased from 100 m to 1000 m. Factors in 500–2000 m buffers were obtained.

The Mann–Kendall (MK) test [32] was adopted for the trend analysis of the water quality from 2002 to 2014. Spearman’s rank correlation was adopted to determine the relationships between water quality and the factors.

2.3. Cluster Analysis (CA)

To distinguish sites by the indicators’ concentrations measured at different sites from 2010 to 2014, hierarchical clustering was adopted to analyze the monitored sites. The average linkage clustering method and the measured Euclidean distance were used. The
ranges of various indicators concentration values differed. To eliminate the effects of this difference, all values were first standardized. Clustering analysis was conducted using IBM SPSS Statistics 22 for Windows.

2.4. Moran’s I

To determine the spatial heterogeneity of water quality in this area, Moran’s I was adopted to reflect the clustering characteristics.

\[
I = \frac{n}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}
\]

where \( n \) is the number of geographic units; \( x_i \) and \( x_j \) are the spatial locations of \( i \) and \( j \), respectively; and \( w_{ij} \) is the spatial weight between units \( i \) and \( j \) which was all set as 1.

Moran’s I values range from \(-1\) to 1, with positive values indicating the spatial clustering characteristics of the geographic units, and negative values indicating the spatial dispersal characteristics of the units. A value of 0 indicates the lack of spatial clustering and spatial dispersal characteristics. The obtained pattern did not appear to be significantly different from that expected at random. Moran’s I values were calculated in ArcMap 10.3. The spatial patterns were determined using the z-score and \( p \)-value. When the z-score was higher than 1.65 and the \( p \)-value was lower than 0.10, a clustered pattern was more likely. When the z-score was lower than \(-1.65\) and the \( p \)-value was lower than 0.10, a dispersed pattern was more likely. In other situations, a random pattern was observed.

2.5. Distribution of Polluted Regions (Standard Deviational Ellipse)

The pollution evaluation for this area was based on the number of polluted sites. For example, water quality was divided into 5 levels according to the National Standard of Surface Water of China (GB3838-2002) (Table 2). Sites with mean TP concentrations lower than that of the Class IV water standard (0.3 mg/L) were assumed to be unpolluted or sub-polluted. In contrast, sites with a mean TP concentrations higher than that of the Class IV standard were assumed to be relatively polluted. TN is the most serious type of pollution in this area, and its concentrations in most sites were much higher than the Class V standard. Therefore, we used the tripled Class V concentration standard (6.0 mg/L) to identify relatively polluted sites.

Table 2. National Standard of Surface Water of China (GB3838-2002).

<table>
<thead>
<tr>
<th>Class</th>
<th>Class I</th>
<th>Class II</th>
<th>Class III</th>
<th>Class IV</th>
<th>Class V</th>
</tr>
</thead>
<tbody>
<tr>
<td>DO (mg/L)</td>
<td>7.5</td>
<td>6.0</td>
<td>5.0</td>
<td>3.0</td>
<td>2.0</td>
</tr>
<tr>
<td>COD_{Mn} (mg/L)</td>
<td>2.0</td>
<td>4.0</td>
<td>6.0</td>
<td>10.0</td>
<td>15.0</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>0.2</td>
<td>0.5</td>
<td>1.0</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>NH₄⁺-N (mg/L)</td>
<td>0.15</td>
<td>0.5</td>
<td>1.0</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>0.02</td>
<td>0.1</td>
<td>0.2</td>
<td>0.3</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Polluted sites were identified, and then, the spatial distributions and directions of these sites were calculated to reflect the spatial patterns of polluted regions. Thiessen polygons were calculated to separate regions around the monitoring sites. For areas with fewer than three neighboring polluted polygons, the monitoring sites corresponding to these polygons were used to identify polluted regions. The polluted region gravity centers were defined as the mean X and Y values of the sites belonging to each region. The spatial distribution directions were calculated using a standard deviational ellipse tool.
Standard deviational ellipses can reflect the spatial characteristics of pollution distributions, including the locations of polluted regions and their main directions. The standard deviational ellipses were calculated as follows:

\[
SDE_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{X})^2}{n}}
\]

\[
SDE_y = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \bar{Y})^2}{n}}
\]

where \((SDE_x, SDE_y)\) are the coordinates of the center of the standard deviational ellipse, and \((x_i, y_i)\) are the coordinates of point \(i\). \((\bar{X}, \bar{Y})\) represents the average center of all points. \(n\) represents the number of points.

The computational formula of the rotation angle \(\theta\) is as follows:

\[
\tan \theta = \frac{A + B}{C}
\]

\[
A = \left( \sum_{i=1}^{n} \bar{x}_i^2 - \sum_{i=1}^{n} \bar{y}_i^2 \right)
\]

\[
B = \sqrt{\left( \sum_{i=1}^{n} \bar{x}_i^2 - \sum_{i=1}^{n} \bar{y}_i^2 \right)^2 + 4 \left( \sum_{i=1}^{n} \bar{x}_i \bar{y}_i \right)^2}
\]

\[
C = 2 \sum_{i=1}^{n} \bar{x}_i \bar{y}_i
\]

where \(\bar{x}_i\) and \(\bar{y}_i\) are the differences values between the average values and point \(i\).

\[
\sigma_x = \sqrt{\frac{\sum_{i=1}^{n} (\bar{x}_i \cos \theta - \bar{y}_i \sin \theta)^2}{n}}
\]

\[
\sigma_y = \sqrt{\frac{\sum_{i=1}^{n} (\bar{x}_i \sin \theta - \bar{y}_i \cos \theta)^2}{n}}
\]

where \(\sigma_x\) and \(\sigma_y\) represent the long axis and short axis of the ellipse, respectively.

Standard ellipse analysis was conducted using the Directional Distribution Tool in ArcMap 10.3.

3. Results

This study aimed to detect the interannual and seasonal changes and spatial patterns of water quality in this region, especially focusing on clusters, the spatial heterogeneity of water quality characteristics and the identification of polluted regions. The correlation between influence factors and water quality indicators was also presented.

3.1. Interannual Variation

Boxplots of DO, percent saturation of oxygen and COD\(_{Mn}\) in different years are shown in Figure 2. The changes in the other examined indicators are shown in Figure 3. Boxplots based on 16 sites monitored from 2002 to 2014 were also plotted to provide a more reliable comparison (Figures 2 and 3).
The correlation between influence factors and water quality indicators was also presented.

3.1. Interannual Variation

Boxplots of DO, percent saturation of oxygen and CODMn in different years are shown in Figure 2. The changes in the other examined indicators are shown in Figure 3. Boxplots based on 16 sites monitored from 2002 to 2014 were also plotted to provide a more reliable comparison (Figures 2 and 3).

**Figure 2.** Variations in the mean values and boxplots of the annual average DO concentration, percent saturation of oxygen and COD\textsubscript{Mn} concentration from 2002 to 2014 based on different levels of the GB3838-2002 standards (the lines of the Class V standard to the Class I standard gradually change from black to 50% gray; the curves represent the mean values, the black ‘+’ symbols represent the mean values of the data, the red lines represent the median values of the data, the edges of the boxes are the 25th and 75th percentiles, the whiskers extend to the most extreme data points that were not considered outliers by the algorithm, and the outliers are plotted individually as red ‘+’ symbols; the boxplots in the left column are based on all sites, and the plots in the right column are based on 16 sites). The trends of mean values are estimated using the MK test, as shown as the texts of MK\_slope and \( p \)-value in all sub figures.
The results showed a well-established trend in water quality. Generally, the NH$_4^+$-N, TN, and COD$_{Mn}$ concentrations showed significant decreasing trends, whereas the DO concentration and percent saturation of oxygen showed significant increasing trends. When considering only the 16 sites monitored in 2002–2014, the water quality was better and the resulting trends were steeper than the trends corresponding to all sites. The trend
of the NH$_4^+\text{-N}$ concentration showed a steady decrease, while the other three indicators produced wavy curves. There were high outlier values for TN, COD$_{Mn}$, NH$_4^+\text{-N}$, and TP concentrations, which made the mean values larger than the median values. This result indicated that some rivers were highly polluted and were in a worse state than the average in this area in different years. Determining where these sites are located is critical for pollution abatement.

The DO and COD$_{Mn}$ indicators showed that the biochemical situation in this area was relatively good. The DO concentration at most sites reached the Class V standard. Thus, aquatic organisms can live in these environments without the risk of low DO, which improves water quality. The mean DO concentration generally increased from 2002 to 2014. As the temperature was stable from 2002 to 2014, the trend of percent saturation of oxygen was almost the same as that of DO. Overall, the water quality improved considerably, and the growth environment for aquatic plants and animals improved. The COD$_{Mn}$ situation was the best among these indicators examined in the study area. All sites achieved the Class V standard, except for one site in 2004. Compared with the other indicators, organic pollution was less serious. Moreover, the COD$_{Mn}$ concentration decreased from 2002 to 2014. The mean concentrations were lower than the Class IV standard before 2009 (6.16–7.14 mg/L), while they were lower than the Class III standard at most sites after 2010 (5.27–6.10 mg/L). Thus, the COD$_{Mn}$ pollution in this area still requires attention.

The TN and NH$_4^+\text{-N}$ indicators showed similar trends, and their mean concentrations showed decreasing trends from 2002 to 2014. Nitrogen pollution, especially TN pollution, was serious, and the effect of governance was not obvious. The TN concentration seriously exceeded the water standard. The mean TN concentrations were more than three times higher than the Class V standard before 2006 (6.08–6.52 mg/L). After 2007, the mean TN concentration was still more than double the Class V standard (4.32–5.01 mg/L). The mean NH$_4^+\text{-N}$ concentrations showed an obvious inflection point in 2006. They ranged from 3.39 to 4.20 mg/L before 2006 and from 1.77 to 2.91 mg/L after 2007. Finally, the mean values reached the Class V standard after 2010.

TP was the only indicator that did not show a significant trend when taking all sites into consideration. However, there was a decreasing trend. The mean concentration values were lower between 2010 and 2014 (0.23–0.27 mg/L) than between 2002 and 2009 (0.24–0.35 mg/L). When only considering the 16 sites monitored in 2002–2014, a significant decreasing trend was observed. However, the mean concentration increased from 2010 to 2014. From 2004 to 2014, the sites in this area produced many outlier values. Many regions in the study area remained seriously polluted. Overall, the level of TP pollution improved, but this led to an increased frequency of outliers.

The values of all indicators were low and showed stability after approximately 2009. Before approximately 2009, water quality was poor and quite variable. After this time, the observed fluctuations were much smaller, and the concentrations of indicators other than TP and DO were much lower than those in the previous period. NH$_4^+\text{-N}$ was the indicator which improved the most when considering all sites, and the mean concentration decreased from 3.31 mg/L before 2009 to 1.88 mg/L after 2010, which reached the Class V standard. After the cyanobacterial pollution event in 2007, great efforts were made to improve the aquatic environment in the study area.

### 3.2. Seasonal Variation

The variations in these five indicators in different seasons (2010–2014) are shown in Figure 4. To quantify the differences in the mean concentrations of indicators, the differences in the annual and seasonal mean concentrations are shown in Table 3. The DO concentration trend was similar to the trends of the other indicators. The mean DO concentration was highest in winter (7.73 mg/L) and lowest in summer (4.73 mg/L), and it showed an obvious increasing trend in autumn and a decreasing trend in spring (Figure 4). Temperature change was one of the factors that influenced this change. Therefore, the
percent saturation of oxygen was also calculated to avoid the influence of temperature (Figure 4). The mean value of percent saturation of oxygen was highest in spring, lower in summer and lowest in autumn.

Figure 4. Changes in the mean values and boxplots of the average concentrations of indicators in each season from 2010 to 2014 relative to different levels of the GB3838-2002 standards (the lines of the Class V standard to the Class I standard gradually change from black to 50% gray; the curves represent the mean values, the black ‘+’ symbols represent the mean values of the data, the red lines represent the median values of the data, the edges of the boxes are the 25th and 75th percentiles, the whiskers extend to the most extreme data points that were not considered outliers by the algorithm, and the outliers are plotted individually as red ‘+’ symbols).
Table 3. Mean concentrations and percent values in different seasons.

<table>
<thead>
<tr>
<th></th>
<th>Year</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td>DO (mg/L)</td>
<td>5.93</td>
<td>6.12 (+0.19)</td>
<td>4.73 (−1.20)</td>
<td>5.17 (−0.76)</td>
<td>7.73 (+1.80)</td>
</tr>
<tr>
<td>Percent saturation of oxygen (%)</td>
<td>61.84</td>
<td>66.73 (+4.89)</td>
<td>58.34 (−3.50)</td>
<td>58.14 (−3.70)</td>
<td>64.14 (+2.30)</td>
</tr>
<tr>
<td>COD$_{Mn}$ (mg/L)</td>
<td>5.54</td>
<td>5.87 (+0.33)</td>
<td>5.56 (+0.02)</td>
<td>5.13 (−0.41)</td>
<td>5.60 (+0.06)</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>4.45</td>
<td>5.17 (+0.72)</td>
<td>3.92 (−0.53)</td>
<td>3.55 (−0.90)</td>
<td>5.15 (+0.70)</td>
</tr>
<tr>
<td>NH$_4^+$-N (mg/L)</td>
<td>1.90</td>
<td>2.41 (+0.51)</td>
<td>1.65 (−0.25)</td>
<td>1.33 (−0.57)</td>
<td>2.23 (+0.33)</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>0.25</td>
<td>0.27 (+0.02)</td>
<td>0.26 (+0.01)</td>
<td>0.21 (−0.04)</td>
<td>0.25 (0.00)</td>
</tr>
</tbody>
</table>

Data inside parentheses are the differences between the mean values in different seasons and the mean values over the whole year.

The COD$_{Mn}$ concentration decreased from spring (5.87 mg/L) to autumn (5.13 mg/L) and increased in winter (5.60 mg/L). The concentrations were generally higher in the non-flood season than in the flood season. The concentrations of TN and NH$_4^+$-N showed similar changes. The TN and NH$_4^+$-N concentrations were higher in spring (5.17 and 2.41 mg/L) and winter (5.15 and 2.23 mg/L) than in summer (3.92 and 1.65 mg/L), and they reached a minimum in autumn (3.55 and 1.33 mg/L). The mean TP concentrations in spring (0.27 mg/L) declined to relatively low levels in summer (0.26 mg/L). Finally, they declined to the lowest annual level in autumn (0.21 mg/L) and increased again in winter (0.25 mg/L). Outliers were relatively scarce in summer and autumn, and their values were lower. Thus, TP pollution was more uniform in this area during these two seasons.

Generally, the water quality was better during the flood season (summer and autumn) than during the nonflood season (winter and spring).

3.3. Spatial Characteristics

3.3.1. Clusters of Water Quality

According to the water quality during 2010–2014, the monitoring sites were divided into six clusters. The site numbers and mean concentrations of the parameters are shown in Table 4. Clusters 5 and 6 only contained one site each, and these were the most polluted clusters in the study area. The water quality worsened from Cluster 1 to Cluster 4 according to almost all parameters, except for TP between Cluster 3 and Cluster 4. However, when assessing water quality using the worst parameter according to GB3838-2002, even the least polluted Cluster 1 did not reach the Class V standard due to the high TN concentration. Cluster 1 reached only the Class IV standard when TN was not considered.

Table 4. Site numbers and mean parameter concentrations of different clusters.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of sites</td>
<td>71</td>
<td>33</td>
<td>4</td>
<td>13</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>DO (mg/L)</td>
<td>6.64</td>
<td>5</td>
<td>4.33</td>
<td>4.19</td>
<td>3.09</td>
<td>3.94</td>
</tr>
<tr>
<td>NH$_4^+$-N (mg/L)</td>
<td>1.08</td>
<td>2.43</td>
<td>2.72</td>
<td>4.14</td>
<td>5.83</td>
<td>6.35</td>
</tr>
<tr>
<td>TP (mg/L)</td>
<td>0.17</td>
<td>0.29</td>
<td>0.53</td>
<td>0.47</td>
<td>0.73</td>
<td>0.37</td>
</tr>
<tr>
<td>TN (mg/L)</td>
<td>3.37</td>
<td>5.08</td>
<td>5.72</td>
<td>7.21</td>
<td>9.03</td>
<td>11.36</td>
</tr>
<tr>
<td>COD$_{Mn}$ (mg/L)</td>
<td>4.80</td>
<td>6.31</td>
<td>6.93</td>
<td>7.12</td>
<td>9.97</td>
<td>9.06</td>
</tr>
</tbody>
</table>

Cluster 1 contained 71 sites, and they were distributed near the Yangtze River, Taihu Lake, and the Wangyu River (Figure 5). The Wangyu River was the most important river for the Water Diversion Project from the Yangtze River to Taihu Lake and divided the study region into the YCDM and WCXY. The emission of pollutants into the Wangyu River and rivers near the Yangtze River and Taihu Lake was limited by the government. The water quality at sites in Cluster 1 was not classified as indicating polluted regions. Cluster 2 contained 33 sites. These sites were distributed in the middle of the WCXY and Taicang. Some sites in this cluster were categorized as part of polluted regions due to the instability of water quality among different years. Cluster 3 contained four sites. Three of these sites were in the WCXY, and one site was located in the northern part of Suzhou city. The sites
in Cluster 3 were more likely to be part of the polluted regions than those in Cluster 2. Cluster 4 contained 13 sites, most of which were distributed in Taicang and Kunshan. These sites should be given more attention. One site in Cluster 5 was located in Changshu, which is next to Gehu Lake. This was a severely polluted site, but sites surrounding it did not show signs of serious pollution. One site in Cluster 6 was located at the boundary of Taicang and Changshu.

Figure 5. Clusters divided by hierarchical clustering of sites in the Jiangsu Taihu Plain, with samples from 2010 to 2014.

In general, sites in Clusters 4–6 were included in polluted regions in most years. Sites in Clusters 2 and 3 were included in polluted regions in some years.

3.3.2. Spatial Heterogeneity of Water Quality

Moran’s I values of the TN and TP concentrations showed similar trends from 2010 to 2014 (Table 5). Both concentrations had high Moran’s I values and z-scores in 2010 and 2011. The z-scores were positive from 2012 to 2014, which means that the patterns in these years showed weak clustering characteristics. However, the values were small; thus, they showed random distribution patterns. The TN and TP pollution in the plain area of Taihu Basin showed clustered patterns in 2010 and 2011. These types of pollutants were clustered to indicate polluted regions. However, from 2012 to 2014, the patterns showed random distribution characteristics, which means that these types of pollutants did not obviously cluster during this time. Therefore, TN and TP pollution was not as concentrated as it was previously.

Table 5. Moran’s I of TN and TP concentrations at monitoring sites from 2010 to 2014.

<table>
<thead>
<tr>
<th>Year</th>
<th>TN</th>
<th>TP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Moran’s I</td>
<td>z-Score</td>
</tr>
<tr>
<td>2010</td>
<td>0.51</td>
<td>2.50</td>
</tr>
<tr>
<td>2011</td>
<td>0.58</td>
<td>3.52</td>
</tr>
<tr>
<td>2012</td>
<td>0.13</td>
<td>0.16</td>
</tr>
<tr>
<td>2013</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>2014</td>
<td>0.05</td>
<td>0.06</td>
</tr>
</tbody>
</table>
3.3.3. Water Quality Pollution Areas

To determine how the regions of TN and TP pollution specifically changed, standard deviational ellipses were calculated for the polluted regions.

Sites with high mean TP concentrations were distributed in three regions. The first region was around the boundary of Changzhou, Jiangyin, and Wuxi. The second region was around the boundary of Kunshan and Taicang. The third region was located near the northern part of Suzhou Old Town. Generally, almost every region decreased in size during these 5 years (Figure 6).

![Figure 6. Standard ellipses of TP polluted regions from 2010 to 2014.](image)

The first region was smaller in 2011 than in 2010, but it always comprised the boundary of Changzhou, Jiangyin, and Wuxi. The distribution at the boundary may be because all of these cities have high-pollution industries that are located far from the city centers where humans aggregate. The direction of this ellipse’s long axis is similar to the direction of the Jiangnan Canal, which was part of the Beijing–Hangzhou Canal in Wuxi in 2010 and 2011. During these two years, water pollution may have been considerably affected by the Jiangnan Canal. Thus, the distribution of polluted sites along the east of the Jiangnan Canal was likely. In 2012 and 2013, no polluted region was recorded in this area. There were two small polluted regions in the same area in 2014. One was in Changzhou, while the other was in Jiangyin/Zhangjiagang. The river water quality has improved in recent years, and the polluted regions have decreased in size.

The second polluted region was the worst and largest region, and the sites in this region were monitored most intensively. From 2010 to 2012, the polluted region was large, but it decreased in size in 2013 and 2014. The polluted region was near the boundary of Shanghai, and its direction was similar to that of the boundary. We can assume that the main reason for the observed trend was economic development. Shanghai is a large city, and its economy is very developed. To reduce economic costs, many economic activities, especially industrial activities, were relocated to Taicang and Kunshan. The increased economic activities and population sizes caused by economic growth led to more pollutant emissions in these two counties and subsequently higher TP concentrations than those in other regions. In 2014, the polluted region moved to the north, and its direction no longer followed the Shanghai municipal boundary. Thus, the distribution attribution changed.
The third polluted region was located in the northern part of Suzhou Old Town. The polluted sites were distributed in the river north of Suzhou in 2010 and expanded to the moat of northern Suzhou. However, in 2014, the TP concentration in the northern moat decreased to 0.189 mg/L, and this polluted region disappeared. The TP pollution problem greatly improved in this area.

The TN-polluted regions were located around Wuxi and Kunshan/Taicang and in some other areas in different years. The TN pollution distribution was similar to the TP pollution distribution (Figure 7).

Figure 7. Standard ellipses of TN polluted regions from 2010 to 2014.

The first polluted region was located between Changzhou and Wuxi, and it included the Jiangnan Canal, which linked these two cities. The direction of this ellipse’s long axis was similar to the direction of the canal. TN pollution was more likely affected by the canal than TP pollution in 2010. There were no polluted regions in this area in 2011 and 2012. However, in 2013 and 2014, polluted regions appeared in Wuxi/Jiangyin and Changzhou/Wuxi, respectively.

The second polluted region was relatively large in 2010 and 2011 and decreased in size in 2012 and 2013. The region eventually split into two regions in 2014. One region was distributed at the boundary of Changshu/Taicang, and the other was distributed at the boundary of Taicang/Kunshan. The direction of the polluted region was similar to the direction of the Shanghai boundary.

Other TN pollution regions appeared in Suzhou in 2011 and Zhangjiagang in 2012. The TN pollution was not stable and could not be easily controlled in the plain area of Taihu Basin. In 2012, both TN and TP showed high pollution attributes in Zhangjiagang, which suffered from considerable water pollution.

3.4. Influencing Factors of Water Quality

The factors influencing water quality include the impervious surface rate, paddy land rate, population and precipitation. Precipitation is a natural factor and the others are artificial factors. (Table 6)
Table 6. Correlations between different factors and TN and TP.

<table>
<thead>
<tr>
<th></th>
<th>Year</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
<th>Winter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TN</td>
<td>TP</td>
<td>TN</td>
<td>TP</td>
<td>TN</td>
</tr>
<tr>
<td>Impervious surface</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rate 500 m</td>
<td>0.197*</td>
<td>0.236*</td>
<td>0.202*</td>
<td>0.205*</td>
<td>0.167</td>
</tr>
<tr>
<td>1000 m</td>
<td>0.221*</td>
<td>0.223*</td>
<td>0.247**</td>
<td>0.208*</td>
<td>0.156</td>
</tr>
<tr>
<td>1500 m</td>
<td>0.243**</td>
<td>0.231*</td>
<td>0.259**</td>
<td>0.202*</td>
<td>0.159</td>
</tr>
<tr>
<td>2000 m</td>
<td>0.255**</td>
<td>0.238**</td>
<td>0.264**</td>
<td>0.190*</td>
<td>0.173</td>
</tr>
<tr>
<td>Paddy land rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500 m</td>
<td>−0.020</td>
<td>0.057</td>
<td>−0.063</td>
<td>0.026</td>
<td>0.049</td>
</tr>
<tr>
<td>1000 m</td>
<td>−0.026</td>
<td>0.086</td>
<td>−0.086</td>
<td>0.041</td>
<td>0.091</td>
</tr>
<tr>
<td>1500 m</td>
<td>−0.019</td>
<td>0.089</td>
<td>−0.074</td>
<td>0.070</td>
<td>0.138</td>
</tr>
<tr>
<td>2000 m</td>
<td>−0.009</td>
<td>0.092</td>
<td>−0.067</td>
<td>0.093</td>
<td>0.146</td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500 m</td>
<td>0.157</td>
<td>0.058</td>
<td>0.169</td>
<td>0.021</td>
<td>0.068</td>
</tr>
<tr>
<td>1000 m</td>
<td>0.179*</td>
<td>0.106</td>
<td>0.195*</td>
<td>0.067</td>
<td>0.131</td>
</tr>
<tr>
<td>1500 m</td>
<td>0.172</td>
<td>0.075</td>
<td>0.178</td>
<td>0.018</td>
<td>0.117</td>
</tr>
<tr>
<td>2000 m</td>
<td>0.158</td>
<td>0.07</td>
<td>0.167</td>
<td>0.014</td>
<td>0.093</td>
</tr>
<tr>
<td>Precipitation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>−0.002</td>
<td>−0.014</td>
<td>−0.109</td>
<td>−0.304**</td>
<td>0.191*</td>
</tr>
</tbody>
</table>

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

The directions of the correlations between artificial and natural factors and water quality parameters were opposite. The correlations between water quality parameters and the impervious surface rate were more significant than the correlations with other factors at all spatial scales. However, the correlation was not significant in summer. The paddy land rate presented a slight effect on water quality. It had a positive influence on the TP concentration in summer, with a statistically significant at the 0.05 level. Population exhibited positive correlation with TN, and the significant level of the correlation was highest in the 1000 m buffer in any seasons. The correlations between water quality parameters and precipitation were not significant. However, precipitation had significant positive influences on TN and TP concentration in summer.

4. Discussion

The variations and characteristics of river water quality in study area are influenced by many factors, including natural factors (e.g., meteorological elements and river order) and artificial factors (e.g., urbanization).

4.1. Variations of the Water Quality in Plain Area of the Taihu

The mean DO concentration was the highest in winter and the lowest in summer [22,33], and percent saturation of oxygen was lower in summer and autumn. The DO concentration is usually affected by the factors of aeration, photosynthesis, respiration, and oxidation from waste [34], which have relationships with temperature. The temperature was high in summer and led to less DO in the water. Cyanobacterial growth is another factor that affects DO and results in lower DO concentrations or percent saturation of oxygen. Under eutrophic conditions in the study area, cyanobacterial blooms easily occur. Although cyanobacteria will increase the DO in the surface water through photosynthesis, a high amount of cyanobacterial bloom particles floating on the surface of the river will reduce light transmittance and affect the depth of the water body in which phytoplankton photosynthesize to generate oxygen. High-density algae accumulation will cause a sharp drop in DO in the water when the algae decays and decomposes [35]. Meanwhile, microbiological activities in the water and bottom increased in summer and autumn. For example, the higher temperature results in the decomposition of organic matter in oxygen-consuming processes.

Similarly, Xiao thought that the CODMn concentration was slightly better during the flood season than during the low water period in the southern part of the Taihu Lake [33]. Similarly, the nitrogen concentrations were generally lower in summer than in winter in a study of rivers surrounding the Taihu Lake [22]. In the southern part of the Taihu Basin, there was more TN pollution during the low water period [33], which was almost the same as the situation during winter and spring. In the Taihu Lake, the TN concentration also
reached a maximum in winter–spring and reached a minimum in summer–autumn [36]. Based on the sources of nitrogen pollution, high nitrogen concentrations were expected to appear in the wet season. Because more rain and increased temperatures occur during the same period, rice is planted during the wet season, and fertilizer and pesticides are used to achieve higher yields. The lower concentrations of TN and NH$_4^+$-N indicate that the dilution effect from precipitation was more pronounced than the enhanced runoff effects and the influence of rice planting [37].

The river water quality near the Yangtze River and the Taihu Lake was relatively good. Taihu Lake and the Yangtze River are important sources of drinking water and industrial water for these cities. After the cyanobacterial bloom in 2007, the government focused more attention on the water quality of this lake. Therefore, pollutant emissions around Taihu Lake were strictly controlled, and the water diversion and drainage project were also conducted to accelerate changes in the water of Taihu Lake. The water quality in the rivers around the lake also greatly benefited from this project.

4.2. Influence Factors of the Surface Water Quality

4.2.1. Natural Factors

Natural factors change led to the seasonal variation, as the industrial point-source pollutants are relatively stable across different seasons. Precipitation has a negative correlation with water quality. The quality of rain water is much better than that of river water in the Taihu Plain [18,38]. Rain water would dilute the pollution in the surface water during the flood season. Meanwhile, the large amount of precipitation makes pollution accumulation difficult because of quicker water exchange [30]. More precipitation occurs during the flood season, resulting in a high water level which can dilute pollutants. Flow reduction can result in larger concentrations of pollutants in the Ganga River around the industrialized Kanpur Region [39]. Therefore, water quality during in the flood season is better than that during the nonflood season.

In summer, agricultural activities result in many nonpoint-source pollutants. More precipitation leads to higher pollution concentrations in rivers. The temperature is higher in summer than in winter. In summer, aquatic plants grow, absorb some pollution from the water and improve water quality. However, cyanobacterial outbreaks may worsen river water quality in summer in some regions. The natural factors mainly cause the seasonal variation of the river water quality.

4.2.2. Artificial Factors

The impervious surface rate influences the water quality the most among all factors. This occurs because factories are built on impervious surfaces, which has a greater impact on water quality in this area [25]. Water quality in the Taihu Basin was related not only to point sources but also nonpoint sources [40]. Point source pollutants such as domestic and municipal sewage and industrial wastewater are discharged into rivers. Meanwhile, impervious surfaces were also considered as the main source of nonpoint-source pollution, and the nonpoint output in urban land was even greater than that in paddy land [41]. Therefore, the impervious surface rate distribution makes river water around polluted. Meanwhile, this region has interconnected waterways and is agriculturally well developed. Nutrients can be generated by agricultural activities [42], which results in a significant positive correlation between the paddy land rate and TP. Water area changes to agricultural land through the reclamation of lakes and rivers. This may result in a negative influence of the water body on TP and TN [43]. Rapid urbanization and industrialization in this area have also resulted in tremendous changes in the structure of the river network [44]. Rivers were buried in the study area, influencing the transfer of water pollution and self-cleaning capacity. However, the flow velocity increased after the water diversion and drainage project was conducted, and the self-cleaning capacity of river water subsequently improved.

The influence of artificial factors on river water quality is very complex. To improve environmental conditions, relevant departments have focused attention on pollution problems
in river water. Point source pollutants were monitored at sewage discharge locations. Non-point source pollutants were intended to be controlled by existing policies and economics. Overall, artificial factors influence water quality more than natural factors.

5. Conclusions

This study aimed to detect the interannual and seasonal changes and spatial patterns of water quality in this region. Based on cluster analysis, Moran’s I, and standard deviational ellipses, the site clusters, spatial heterogeneity of water quality characteristics and identified polluted regions were clarified. Finally, factors influencing water quality were discussed.

1) Water quality in the plain area of Taihu Basin generally improved from 2002 to 2014. There was an obvious increasing trend in the DO (4.38 to 5.98 mg/L) concentration and decreasing trends in the COD\textsubscript{Mn} (6.88 to 5.35 mg/L), NH\textsubscript{4}\textsuperscript{+}-N (4.20 to 1.87 mg/L), and TN (6.48 to 4.39 mg/L) concentrations from 2002 to 2014. TN was the worst type of pollution in the plain area of Taihu Basin. The TP concentration did not show an effectively steady descending trend. However, its concentration was 0.25 mg/L in 2010–2014, which was lower than the previous concentration (0.29 mg/L). The DO concentration was high in winter and low in summer because of the temperature changes among different seasons. In terms of nutrients and COD\textsubscript{Mn}, the water quality was relatively good in summer and the best in autumn. Water quality should be paid more attention in winter and spring.

2) There were still relatively polluted regions, although the water quality improved. Polluted sites showed clustered patterns in 2010 and 2011, while they were randomly distributed from 2012 to 2014. The regions with TN pollution were distributed in Taicang/Kunshan, Wuxi/Changzhou/Jiangyin, and Suzhou. The regions with TP pollution were mainly distributed in Taicang/Kunshan and other areas. The region in Taicang/Kunshan was the most polluted. TN and TP pollution in this area lasted from 2010 to 2014. Pollution reduction programs should be conducted in these polluted regions.

3) Natural factors influence the seasonal variations of water quality. In summer, the precipitation influenced TP or TN more significantly than any other factor. Artificial factors had a more significant influence on water quality than natural factors. Among artificial factors, the impervious surface rate influences water quality the most. Pollution from impervious surface should be paid more attention for water quality improvement.

Author Contributions: Conceptualization, B.G.; Formal analysis, B.G. and Z.L.; Funding acquisition, Y.X.; Methodology, B.G. and Z.L.; Project administration, Y.X.; Resources, Y.X.; Supervision, Y.X.; Visualization, B.G.; Writing—original draft, B.G.; Writing—review and editing, Y.X., Z.L., M.L. and Q.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key Technology Support Program (No. 2018YFC1508201), National Natural Science Foundation of China (No. 41771032 and No. 41901026), and Water Conservancy Science and Technology Foundation of Jiangsu Province (No. 2021010). Data Availability Statement: Not applicable.

Acknowledgments: We would like to express our thanks to the Wuxi, Changzhou, Suzhou branches of hydrology and the Water Resources Survey Bureau of Jiangsu Province for their cooperation, monitoring efforts on water quality data. We acknowledge the data support from “Geographic Data Sharing Infrastructure, College of Urban and Environmental Science, Peking University (http://geodata.pku.edu.cn, accessed on 12 June 2021)”.

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


42. Edwards, A.C.; Withers, P. Transport and delivery of suspended solids, nitrogen and phosphorus from various sources to freshwaters in the UK. J. Hydro. 2008, 350, 144–153. [CrossRef]