A Comprehensive Method for Water Environment Assessment considering Trends of Water Quality

Sheng Bi,1,2,3 Li Wang,4 Yongrong Li,1 Zhenping Zhang,1 Zhimian Wang,1 Xiaoling Ding,1,5 and Jianzhong Zhou5

1Changjiang River Scientific Research Institute, Wuhan, Hubei 430010, China
2Key Lab of Basin Water Resource and Eco-Environmental Science, Hubei Province, Wuhan, Hubei 430010, China
3Hunan Provincial Key Laboratory of Key Technology on Hydropower Development, Changsha 410000, China
4South-to-North Water Diversion Project Midline Water Source Co., Ltd., Shiyan 442700, China
5School of Civil and Hydraulic Engineering, Huazhong University of Science and Technology, Wuhan, Hubei 430010, China

Correspondence should be addressed to Sheng Bi; bbs.hust@gmail.com

Received 8 January 2021; Revised 13 April 2021; Accepted 19 August 2021; Published 9 September 2021

1. Introduction

Sufficient access to freshwater resources is essential to the survival of human and natural ecosystems, and its role is inextricably coupled in economic sectors [1]. The actual state of water resources is a sensitive and critical issue in many countries, and the deterioration of water quality has become a global problem. An appraisal of available water resources and their quality is the first step to ensuring water security; thus, research on water quality evaluation has been an area of increasing interest and necessity [2].

Traditional methods for water quality assessment, which include measuring chemical parameters in water and normalized measurements against existing standards, are not sufficiently effective [3]. With the advantage of fully using water quality parameter information under a normative value, the water quality index (WQI), proposed by Horton [4] and Brown et al. [5], was widely applied to classify surface quality [1, 6–9]. WQI pairs concentrations of several water quality parameters such as dissolved oxygen (DO), pH, nitrate, phosphate, ammonia, chloride, hardness, and metals with associated weights [10]. The lower the WQI, the better the water quality, while a higher value indicates degraded water quality. The index value depends on monitored constituents and threshold values of specific domestic or local standards. In recent years, the WQI concept has evolved into other formulations including the National Sanitation Foundation WQI (NSFWQI) and the integrated water quality index (IWQI) [11, 12].

With the advent of environmental monitoring networks and chemometrics, there have been associated increases in the application of multivariate statistical techniques, such as cluster analysis (CA), principal component analysis (PCA), and factor analysis (FA) [13–16]. These techniques can reduce the dimensionality of a multivariate dataset while still maintaining the original structure, such as principal...
components or factors [14]. With the advantage of effective dimension reduction, these techniques have been frequently employed to classify water quality data and detect similarities among variables in many research studies [17], which help with the interpretation of results and in turn proffer a less subjective and more objective process.

Previous methods for assessing water quality in the environment were almost always based on historical data. When, in fact, the conveyance and storage systems were dynamic, water quality in the cleanest rivers could be in the process of deteriorating, while the opposite could be true for the most polluted rivers. A dynamic change in water quality could lead to inconsistencies between results of a one-time evaluation and reality.

In this paper, we proposed a water quality assessment method which takes into account water quality trends. First, a comprehensive pollution index (CPI) was calculated at several monitoring locations within a larger field site. Next, Spearman’s rank correlation analysis was used to determine CPI trends at each location. Finally, principal component analysis (PCA) was deployed to assess water quality at the field scale using CPI trends as one of the evaluation parameters. The process of the method is illustrated in Section 2.4.

### 2. Methods

#### 2.1. Comprehensive Pollution Index (CPI) Method

The CPI method is one of the most common approaches used in China to qualitatively evaluate the water quality [18–21]. The following equation calculates the index $P$ as

$$P = \frac{1}{n} \sum_{i=1}^{n} P_i = \frac{1}{n} \sum_{i=1}^{n} \frac{C_i}{S_i} \quad (i = 1, 2, \ldots, n), \quad (1)$$

where $C_i$ is the monitored value of water quality parameter $i$; $n$ is the number of water quality parameters monitored; and $S_i$ is the standard permissible limit on water quality parameter $i$ according to the environmental quality standards for surface waters (GB3838-2002) in China (Table 1) [22].

With this approach, $P$ increases proportionally to the concentration increase of each constituent normalized by its water quality standard; hence, the value of $P$ could be used to classify the degree of water contamination (Table 2).

#### 2.2. Spearman’s Rank Correlation Analysis

Spearman’s rank correlation coefficient is defined as the Pearson correlation coefficient between two ranked random variables. It is used when data are not normally distributed between two variables. Spearman’s rank correlation coefficient, $rs$, can be calculated with the following formula [23]:

$$rs = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n(n^2 - 1)} \quad (2)$$

where $d_i = X_i - Y_i$ is the difference between the two ranked variables and $n$ is the total number of observations. In its formulation, values of $rs$ are constrained within $-1 \leq rs \leq +1$. If $rs = +1$, it indicates a perfect positive correlation, while $rs = -1$ indicates a perfect negative correlation.

In this approach, $X_i$ is the ranking sequence according to listed $P$ values, and $Y_i$ is the ranking sequence according to the year of sampling. As $rs$ approaches +1, the more obvious the uptrend in the value of $P$ over time, hence indicating a progressive deterioration of water quality. Conversely, as $rs$ approaches −1, the downtrend in $P$ over time is more obvious as is the progressive improvement in water quality over time.

#### 2.3. Principal Component Analysis

In this framework, the main pollution sources were determined by principal component analysis (PCA). This method is widely used because it can extract data and reduce system dimensionality. In PCA, the dataset contains $P$ uncorrelated variables each of which is called a principal component (PC). The most important parameters are extracted in the first component (PC$_1$), while other parameters with relatively less importance fall within subsequent components (PC$_2$, \ldots, PC$_n$). Algebraically, for $n$ original variables, $x_1$, $x_2$, \ldots, $x_n$, the expression of the principal component PC$_i$ is determined as

---

**Table 1: Classification of water area functions in GB3838-2002.**

| Classification standard value items | Class I | Class II | Class III | Class IV | Class V |
|------------------------------------|---------|----------|-----------|---------|---------|
| pH                                 | 7.5     | 6        | 6–9       | 3       | 2       |
| DO                                 |         | 4        | 6         | 10      | 15      |
| COD$_{Mn}$                         | 15      | 15       | 20        | 30      | 40      |
| COD$_C$                            | 3       | 3        | 4         | 6       | 10      |
| BOD$_C$                            | 0.15    | 0.5      | 1         | 1.5     | 2       |
| NH$_3$-N                           | 0.2     | 0.2      | 0.2       | 0.3     | 0.4     |
| TP                                 | 0.02    | 0.1      | 0.1       | 0.2     | 0.3     |

**Table 2: Standard of surface water quality classification.**

| Comprehensive pollution index ($P$) | Water quality level |
|-------------------------------------|---------------------|
| ≤0.20                               | Clean               |
| 0.21–0.40                           | Less clean          |
| 0.41–1.00                           | Slight pollution    |
| 1.01–2.0                            | Moderate pollution  |
| ≥2.01                               | Severe pollution    |
\[ PC_1 = a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n, \]
\[ \vdots \]
\[ PC_n = a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n, \]

where \( a_{ij} \) are the eigenvectors extracted from the covariance or correlation matrix and the number of components \( n \) is equal to the number of variables. As stated above, most of the information found within the \( n \) original variables is captured within the first few principal components, thus reducing the practical dimensionality of the dataset. Further details can be found elsewhere [14, 24].

The score of each principal component can be obtained from
\[ F_i = \text{Fac}_i \sqrt{\lambda_i}, \]
where \( F_i \) is the score of \( PC_i \), \( \text{Fac}_i \) is the value of factor \( i \), and \( \lambda_i \) is the eigenvalue of \( PC_i \), and \( i \) is the index of the PCs. The comprehensive score can be calculated using the following equation:
\[ F = \frac{1}{\sum_{i=1}^{n} \alpha_i} \sum_{i=1}^{n} \alpha_i F_i, \]
where \( \alpha_i \) is the contribution rate of \( PC_i \) and \( n \) is the number of PCs. The relative level of pollution at any given location can be determined by ranking the comprehensive scores of those sampled sections. The higher the comprehensive value, the lower the water quality.

2.4. The Process of the Method. The process of the comprehensive method can be described as follows (Figure 1).

Step 1. Calculate the annual composite pollution index \( P \) for each monitoring location using equation (1) and at least 4 years of monitoring data to generate 4 yearly averages to be used in the next step.

Step 2. Use Spearman’s rank correlation coefficient method to calculate CPI trends \( (rs) \) at each monitoring location using equation (2).

Step 3. Using PCA and the original water quality measurements of \( \text{pH}, \text{DO}, \text{COD}_{\text{Mn}}, \text{COD}_{\text{Cr}}, \text{BOD}_5, \text{NH}_3-N, \text{TN}, \text{and TP} \), while the last variable was the CPI trend. Using PCA, the comprehensive scores of each tributary were obtained. Again, a higher comprehensive score equates to a tributary with lower water quality. PCA results were obtained using the commercial computer program SYSTAT version 19 from SPSS Inc.

3.3. P Value Trend Analysis. Using annual \( P \) values listed under each stream in Table 4, Spearman’s rank correlation coefficient, \( rs \), was calculated using equation (2), which equates the slope of the CPI trendline listed in Table 5. Of the 16 major tributaries of the Danjiangkou Reservoir, 14 showed CPI trends of water quality improvement; moreover, the Shengding River and the Quyuan River showed continuous improvement trends. The two exceptions were the Jiangu River and the Jianhe River.

3.4. Evaluation of Water Quality with Trends Considered. There are 9 original variables which were used in PCA. For each sampled tributary, the first eight variables were water quality indicators which included \( \text{pH}, \text{DO}, \text{COD}_{\text{Mn}}, \text{COD}_{\text{Cr}}, \text{BOD}_5, \text{NH}_3-N, \text{TN}, \text{and TP} \), while the last variable was the CPI trend. Using PCA, the comprehensive scores of each tributary were obtained. Again, a higher comprehensive score equates to a tributary with lower water quality. PCA results were obtained using the commercial computer program SYSTAT version 19 from SPSS Inc.

The 9 original variables were standardized (Table 6) and then subjected to PCA. The factor correlation coefficient considered significant is the one that is greater than 85.0%. PCA on 9 parameters yielded two principal components explaining 88.3% of the sample variance. Table 7 shows component 1 describing 76.24% of the sample variance and component 2 capturing 12.09%. Figure 3 represents the 2D plot of both principal components. \( \text{COD}_{\text{Mn}}, \text{COD}_{\text{Cr}}, \text{BOD}_5, \text{NH}_3-N, \text{TP} \), and \( \text{TN} \) had a positive influence on \( PC_1 \), while \( \text{pH} \) and \( \text{DO} \) had a negative impact. As for \( PC_2 \), it was dependent on \( rs \) coefficients, representing CPI trends.

According to the above analysis, scores from the first two principal components and the comprehensive score are obtained and shown in Table 8. As seen from this table, the pollution conditions in the tributaries were confirmed according to the comprehensive scores. When the CPI trends were considered, the three most polluted rivers were the Sihe River, Shending River, and Jianghe River, and the three least polluted were the Taohe River, Baihe River, and Qihe River.
Obtain the annual P value of each section by CPI method

Obtain the $r_s$ of each section by Spearman rank correlation analysis

Obtain and rank the scores of each section by PCA

The concentration of pollutant (PH, DO, COD$_{ML}$, COD$_{Cr}$, BOD$_5$, NH$_3$-N, TN, TP in this study) of each section

Determine the major pollution sections based on the score rankings

Figure 1: Process of the comprehensive method.

Figure 2: Location map showing the Danjiangkou Reservoir and monitoring sites.
### Table 3: Monitoring sites in the Danjiangkou Reservoir and its tributaries.

| Tributary  | Qihe River | Baihe River | Duhe River | Tianhe River | Langhe River | Shending River | Sihe River | Jianhe River |
|------------|------------|-------------|------------|--------------|--------------|----------------|------------|--------------|
| Measuring site | M1         | M2          | M3         | M4           | M5           | M6             | M7         | M8           |
| Tributary   | Jiangjun River | Taohe River | Taogou River | Quyuan River | Xianghe River | Laoguan River | Guanshan River | Jianghe River |
| Measuring site | M9          | M10         | M11        | M12          | M13          | M14            | M15        | M16          |

### Table 4: The annual $P$ value for each river section.

| Measuring site | Qihe River | Baihe River | Duhe River | Tianhe River | Langhe River | Shending River | Sihe River | Jianhe River |
|----------------|------------|-------------|------------|--------------|--------------|----------------|------------|--------------|
| P1             | 0.28       | 0.28        | 0.18       | 0.26         |              |                |            |              |
| P2             | 0.27       | 0.27        | 0.18       | 0.24         |              |                |            |              |
| P3             | 0.22       | 0.22        | 0.35       | 0.36         |              |                |            |              |
| P4             | 0.29       | 0.29        | 0.34       | 0.34         |              |                |            |              |

### Table 5: The $rs$ coefficient of all river sections.

| Tributary     | Qihe River | Baihe River | Duhe River | Tianhe River | Langhe River | Shending River | Sihe River | Jianhe River |
|---------------|------------|-------------|------------|--------------|--------------|----------------|------------|--------------|
| rs            | -0.80      | -0.80       | -0.60      | -0.80        | -0.40        | -1.00          | -0.80      | -0.80        |
| Tributary     | Jiangjun River | Taohe River | Taogou River | Quyuan River | Xianghe River | Laoguan River | Guanshan River | Jianghe River |
| rs            | 0.20       | -0.80       | -0.60       | -1.00        | -0.40        | -0.40          | -0.80      | 0.20         |

### Table 6: The normalized value of original variables.

| Tributary     | ZpH    | ZDO | ZCOD$_{Mg}$ | ZCOD$_{Cl}$ | ZBOD$_5$ | ZNH$_3$-N | ZTP | ZTN | Z$_{rs}$ |
|---------------|--------|-----|-------------|-------------|----------|------------|-----|-----|----------|
| Qihe River    | 0.03   | 0.37| -0.76       | -0.70       | -0.70    | -0.51      | -0.63| -0.62| -0.55    |
| Baihe River   | 0.33   | 0.39| -0.81       | -0.76       | -0.84    | -0.53      | -0.55| -0.72| -0.55    |
| Duhe River    | 0.09   | 0.23| -0.55       | -0.59       | -0.67    | -0.49      | -0.55| -0.65| 0.00     |
| Tianhe River  | 0.93   | 0.94| -0.44       | -0.44       | -0.32    | -0.27      | -0.05| -0.41| -0.55    |
| Langhe River  | -0.45  | -0.39| 0.00        | -0.01       | -0.32    | -0.43      | -0.55| -0.33| 0.55     |
| Shending River| -1.77  | -1.88| 1.98        | 2.07        | 2.49     | 2.60       | 2.29| 2.36| -1.10    |
| Sihe River    | -2.49  | -2.39| 1.96        | 2.05        | 1.89     | 2.44       | 2.63| 2.10| -0.55    |
| Jianhe River  | 1.53   | 0.73| 1.07        | 1.44        | 1.01     | -0.30      | -0.21| -0.12| -0.55    |
| Jiangjun River| 0.63   | 0.75| -0.52       | -0.62       | -0.54    | -0.49      | -0.38| -0.59| 2.19     |
| Taohe River   | 0.39   | -0.09| 1.05        | -0.80       | -0.71    | -0.51      | -0.72| -0.81| -0.55    |
| Taogou River  | 0.39   | 1.30| -0.69       | -0.66       | -0.64    | -0.52      | -0.63| -0.92| 0.00     |
| Quyuan River  | 0.39   | 0.32| -0.77       | -0.64       | -0.46    | -0.26      | -0.05| -0.04| -1.10    |
| Xianghe River | 0.87   | 0.31| -0.86       | -0.78       | -0.62    | -0.48      | -0.55| 0.07| 0.55     |
| Laoguan River | -0.33  | -0.80| 0.66        | 0.25        | 0.54     | 0.08       | -0.13| -0.03| 0.55     |
| Guanshan River| -0.63  | -0.52| -0.07       | -0.26       | -0.41    | -0.20      | -0.21| -0.45| -0.55    |
| Jianghe River | 0.15   | 0.72| 0.87        | 0.44        | 0.29     | -0.14      | 0.29| 1.16| 2.19     |
Using PCA again, we repeated our evaluation without CPI trends and compared results to those presented in Table 8 in Figure 4. As shown in the figure, pollution rankings for most rivers were different with the exception of the two most polluted rivers. For example, the Jianghe River changed its pollution rating from 4th to 3rd when its \( rs \)
coefficient was considered. The rs coefficient is greater than 0 which indicates that water quality was deteriorating. Therefore, particular attention should be paid to this tributary when the local government plans to carry out pollution mitigation efforts.

4. Conclusions

In this paper, a comprehensive method for water environment assessment considering water quality trends was proposed and employed to analyze pollution in tributaries of the Danjiangkou Reservoir. A preliminary assessment was conducted using the CPI method and annual averaged water quality data. Using the annual P values from each tributary, CPI trends were quantified under Spearman’s rank correlation analysis. CPI trends were shown to be important parameters in PCA and in the determination of comprehensive scores. By employing the comprehensive method developed herein, polluted tributaries were identified as candidates for future environmental management and pollution control.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors’ Contributions

Sheng Bi, the corresponding author of this article, contributed to the main contents for the abstract, Sections 1, 2.2, 2.4, and 3.4, and the references. Li Wang contributed to the suggestions for versions 2 and 3, including Sections 1, 3.3, and 3.4, and the conclusions. Yongrong Li contributed to the main contents for Sections 2.1 and 3.1. Zhenping Zhang and Zhimian Wang contributed to the main contents for Sections 2.2, 3.2, and 3.3. Xiaoling Ding and Jianzhong Zhou contributed to the main contents for Sections 2.3 and 3.4 together with the corresponding author.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (nos. 51709012, U1865202, and 51527809), Fundamental Research Funds for Central Public Welfare Research Institutes of China (nos. CKSF2020058/SZ and CKSF2019526/TG3), and Open Research Program of Hunan Provincial Key Laboratory of Key Technology on Hydropower Development (no. PKLHD201705).

References

[1] T. Ammar, L. Belkhiri, and L. Mouni, “Evaluation of surface water quality for drinking purposes using fuzzy inference system,” Groundwater for Sustainable Development, vol. 6, pp. 235–244, 2018.
[2] Z. Witek and A. Jarosiewicz, “Long-term changes in nutrient status of river water,” Polish Journal of Environmental Studies, vol. 18, no. 6, p. 1177, 2009.
[3] R. Noori, R. Berndtsson, M. Hosseinzadeh, J. F. Adamowski, and M. R. Abyaneh, “A critical review on the application of the national sanitation foundation water quality index,” Environmental Pollution, vol. 244, pp. 575–587, 2019.
[4] R. K. Horton, “An index number system for rating water quality,” Journal Water Pollution Control Federation, vol. 37, no. 3, pp. 300–306, 1965.
[5] R. M. Brown, N. I. McClelland, R. A. Deininger, and R. G. Tozer, “A water quality index—do we dare,” Water & Sewage Works, vol. 117, pp. 339–343, 1970.
[6] R. Reza and G. Singh, “Assessment of ground water quality status by using water quality index method in Orissa, India,” World Applied Sciences Journal, vol. 9, no. 12, pp. 1392–1397, 2010.
[7] S. M. Sadat-Noori, K. Ebrahimi, and A. M. Liaghat, “Groundwater quality assessment using the water quality index and GIS in saveh-nobaran aquifer, Iran,” Environmental Earth Sciences, vol. 71, no. 9, pp. 3827–3843, 2014.
[8] P. Tirkey, T. Bhattacharya, and S. Chakraborty, “Arsenic and other metals in the groundwater samples of Ranchi city, Jharkhand, India,” Current Science, vol. 110, no. 1, pp. 76–80, 2016.
[9] S. H. Ewaid, S. A. Abed, and S. A. Kadhum, "Predicting the Tigris River water quality within Baghdad, Iraq by using water quality index and regression analysis," *Environmental Technology & Innovation*, vol. 11, pp. 390–398, 2018.

[10] D. Machiwal, V. Cloutier, C. Guler, and N. Kazakis, "A review of GIS-integrated statistical techniques for groundwater quality evaluation and protection," *Environmental Earth Sciences*, vol. 77, no. 19, p. 681, 2018.

[11] F. Misaghi, F. Delgosha, M. Razzaghrmanesh, and B. Myers, "Introducing a water quality index for assessing water for irrigation purposes: a case study of the Ghezel Ozan River," *The Science of the Total Environment*, vol. 589, pp. 107–116, 2017.

[12] S. Mukate, V. Wagh, D. Panaskar, J. A. Jacobs, and A. Sawant, "Development of new integrated water quality index (IWQI) model to evaluate the drinking suitability of water," *Ecological Indicators*, vol. 101, pp. 348–354, 2019.

[13] S. Platikanov, D. Baquero, S. Gonzalez et al., "Chemometric analysis for river water quality assessment at the intake of drinking water treatment plants," *The Science of the Total Environment*, vol. 667, pp. 552–562, 2019.

[14] M. Tripathi and S. K. Singal, "Use of principal component analysis for parameter selection for development of a novel water quality index: a case study of river ganga India," *Ecological Indicators*, vol. 96, no. 1, pp. 430–436, 2019.

[15] C. Pastore, E. Barca, G. Del Moro, C. Iaconi, H. Singer, and G. Mascolo, "Comparison of different types of landfill leachate treatments by employment of nontarget screening to identify residual refractory organics and principal component analysis," *The Science of the Total Environment*, vol. 635, pp. 984–994, 2018.

[16] T. G. Kazi, M. B. Arain, M. K. Jamali et al., "Assessment of water quality of polluted lake using multivariate statistical techniques: a case study," *Ecotoxicology and Environmental Safety*, vol. 72, no. 2, pp. 301–309, 2009.

[17] K. Zeinalzadeh and E. Rezaei, "Determining spatial and temporal changes of surface water quality using principal component analysis," *Journal of Hydrology: Regional Studies*, vol. 13, pp. 1–10, 2017.

[18] G. Matta, A. Kumar, A. Nayak, P. Kumar, and A. K. Tiwari, "Water quality and planktonic composition of river henwal (India) using comprehensive pollution index and biotic-indices," *Transactions of the Indian National Academy of Engineering*, vol. 5, no. 3, pp. 541–553, 2020.

[19] Y. Zhao, X. H. Xia, Z. F. Yang, and F. Wang, "Assessment of water quality in Baiyangdian Lake using multivariate statistical techniques," *Procedia Environmental Sciences*, vol. 13, pp. 1213–1226, 2012.

[20] X. Y. Liu, R. R. Wan, and B. G. Zhou, "Analysis on the water quality changes in Guajinggang river," *Water Conservancy Science and Technology and Economy*, vol. 16, pp. 173–175, 2010.

[21] T. Zhang, J. L. Liu, and X. M. Wang, "Causal analysis of the spatial-temporal variation of water quality in Baiyangdian lake," *Acta Science Circumstantiae*, vol. 30, pp. 217–261, 2010.

[22] J. Xia, G. Xu, P. Guo, and H. Peng, "Tempo-Spatial analysis of water quality in the three gorges reservoir, China, after its 175-m experimental impoundment," *Water Resources Management*, vol. 32, no. 9, pp. 2937–2954, 2018.

[23] J. Ashok Kumar and S. Abirami, "Aspect-based opinion ranking framework for product reviews using a Spearman’s rank correlation coefficient method," *Information Sciences*, vol. 460, pp. 23–41, 2018.