Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2021.27758

238 (2021) 38–48
October

Spatio-temporal variability and water quality assessment of the Mudan River Watershed, Northern China: principal component analysis and water quality index

Liang Liu[a,†], Ting-ting Cao[a,b,†], Xiao-dong Wang[c], Zhou Dandan[b], Chong-wei Cui[a,*]

[a]*State Key Laboratory of Urban Water Resources Centre, School of Environment, Harbin Institute of Technology, Harbin 150090, China, Tel. 18845181188; emails: cuichongwei1991@126.com (C.-w. Cui), 303078961@qq.com (L. Liu), 15545009383@163.com (T.-t. Cao)*
[b]*Northeast Normal University, Changchun 130117, China, email: zhoudandan415@163.com (Z. Dandan)*
[c]*Jilin Jianzhu University, Municipal Environment Engineering Inst., 5088 XinCheng St., Changchun, Jilin Province, China, email: 369050446@qq.com*

Received 17 February 2021; Accepted 25 August 2021

ABSTRACT

The Mudan River Watershed, located in the northernmost region of China, is an important component of the Songhuajiang River System. The water quality of this watershed is of great importance because it is widely used for aquaculture production, hydraulic electro-generating, recreation, fishing, and irrigation. In this study, the water quality index (WQI) and principal component analysis (PCA) were employed to evaluate the Spatio-temporal variations in the surface water quality of the watershed. Ten monitoring sections were selected and monitored monthly for this study. Results indicated that the Chaihe bridge section exhibited the worst water quality with a WQI score of 32.81 due to large quantities of sewage discharge. Moreover, the water quality of the Mudan River Watershed showed seasonal variations. The WQI scores in the rainy season, non-rainy season, and icebound season were 69.38, 70.99, and 41.96, which was due to the surface water runoff in the rainy season as well as the low temperature and bacterial activity in the icebound season. Furthermore, the PCA technique revealed that permanganate index and ammonia nitrogen were the primary contaminants of this watershed regardless of the season. Therefore, it is urgent to restore the water quality of the Mudan River Watershed by adopting powerful management measures and improvement methods.

Keywords: Mudan River Watershed; Water quality; Spatio-temporal variations; Water quality index (WQI); Principal component analysis (PCA)

1. Introduction

The Mudan River Watershed, located in the northernmost region of China, is a crucial mainstream of the Songhuajiang River System. It plays an important role in water supply, flood control, shipping, fisheries, and power generation [1,2]. However, human activities in the basin have caused various pollution problems. Among them, human population growth, industrialization, intensive agricultural activities, and rapid urbanization are the major sources responsible for the deterioration of the surface water quality [3–5]. The Mudan River Watershed, a tributary of the Songhuajiang River, is one of the most seriously polluted areas due to the polluted water discharged into the water network around the watershed. Thus, pollution control and regional management in the Mudan River Watershed present significant challenges for China's

* Corresponding author.

† These authors contributed equally to this manuscript.

1944-3994/1944-3986

government. It is significant to identify the water quality and pollution sources to restore these fragile water ecosystems of the watershed.

Recently, numerous assessment methods are applied by researchers to evaluate water quality, such as the single-factor index, Nemerow index, organic contamination index, and water quality index (WQI) [6–13]. Among them, the WQI method, put forward by Horton in 1965, has been widely applied to assess water quality [14–17]. It mainly transforms large quantities of water quality data into a single number that represents the general quality of the surface water. For example, Hou et al. [18] applied the WQI method successfully in the water quality assessment of five typical reservoirs in the lower reaches of the Yellow River, China. They concluded that mercury was the main contaminant in five reservoirs, while TP (total phosphorus) and SO_4^{2-} were the other main contaminants in mountain and Yellow River reservoirs, respectively. Ma et al. [19] reported a classification ranging from "bad" to "medium" in the south coastal aquaculture area of Dalian, Liaoning Province, China by using the WQI method. Şener et al. [20] found that the north and south regions of the Aksu River had been heavily polluted by chemical oxygen demand and Mg assessed by the WQI method. Furthermore, several advanced multivariate statistical techniques are also employed to evaluate the water quality, identify the pollution sources, and understand the Spatio-temporal variations in water quality, such as fuzzy math, single factor way, fuzzy matrix, integrated pollution index, and neural network method [21–23]. These multivariate statistical techniques can help in better interpretation of the results and make the process less subjective. In particular, the principal component analysis (PCA) technique can reduce the dimensionality of a multivariate data set and simultaneously maintain its original variable information to the maximum extent possible [24–26]. Tripathi and Singal [27] applied the PCA method, which led to the reduction in the number of parameters from 28 to 9. They found this statistical technique was less biased and more objective in nature to appraise the water quality of river Ganga in India. Varol [28], Yang et al. [29], Unda-Calvo et al. [30], Camara et al. [31], and Sudhakaran et al. [32] applied different multivariate statistical techniques successfully to obtain the primary polluted section and the principal contamination indexes. In this research, the WQI and PCA method were used to provide integrated information on the Mudan River Watershed for managers.

Based on the dataset of ten state-controlled monitoring sections in 2018 during three periods (rainy season, non-rainy season, and icebound season), The objectives of the present study are to (1) appraise the variation of permanganate index (COD_{Mn}), ammonia nitrogen (NH_4^+–N), total nitrogen (TN) and total phosphorus (TP) and Spatio-temporal variations in water quality; (2) evaluate the water quality of the Mudan River Watershed using WQI method; (3) identify the severe pollution monitoring section through the PCA analysis. It is believed that the results of this research can provide a scientific basis for the Mudan River Watershed management and help to maintain and improve the water conditions of the watershed. Ultimately, this research will lead to the development of both effective management and conservation strategies for the Mudan River Watershed.

2. Methodology

2.1. Natural geographical situation

In this study, the water quality of the Mudan River Watershed, which is the main tributary of the Heilongjiang basin, was investigated by applying different multivariate statistical techniques and a water quality index. The longitude and latitude of this watershed are 43°15′N ~ 46°35′N and 126°39′E ~ 132°12′E. The average annual water temperature ranges from 1°C to 22°C. The annual rainfall in the entire area of Mudan River Watershed is generally about 535 mm, and the precipitation mainly occurs in June to September, accounting for 76% of the total. Due to the favorable climate, fertile soil, and abundant water sources, the land is dominated by agricultural activities. The location of the study area is depicted in Fig. 1.

2.2. Sampling and analysis

In this study, the water samples, collected by high-density polyethylene bottles (previously acid-washed), were taken seasonally from ten state-controlled monitoring sections in 2018. Data from these stations were collected at different months of the year for each parameter. And the corresponding data was divided into three major typical water quality periods, including the rainy season (June to September), non-rainy season (April, May, and October), and icebound season (November to March). The location of the monitoring sections is depicted in Fig. 1. The monitoring section layout and ribbon class of the Mudan River Watershed are listed in Table 1.

The permanganate index (COD_{Mn}), ammonia nitrogen (NH_4^+–N), total nitrogen (TN), and total phosphorus (TP) were chosen for measurement to evaluate the water quality of the Mudan River Watershed. Among them, COD_{Mn} was quantified using a titration method by acid digestion with potassium permanganate and oxalic acid [33]. NH_4^+–N was analyzed using Nessler's reagent spectrophotometric method [34]. TP and TN were measured by flow injection analysis (FIA, QC8500) [35]. All procedures of sampling, transportation, storage and chemical analysis of the surface water samples followed the national standard methods.

According to the national standard — "Environmental Quality Standards for Surface Water" (GB3838–2002) of China, the water body is classified into five types, namely Cases I~V. In this study, the Case III standard was chosen to evaluate the water quality, which is suitable for drinking water. The Case III standard values for COD_{Mn}, NH_4^+–N, TN, and TP are 6, 1, 1, and 0.2 mg/L, respectively. Furthermore, a one-way analysis of variance (ANOVA) and Pearson correlation analysis was conducted to determine if there were significant spatial and seasonal differences in water quality variables [28].

2.3. Water quality index calculations

The Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) is widely applied to appraise the water quality in recent years. It mainly calculates whether the water quality monitoring data exceeds the

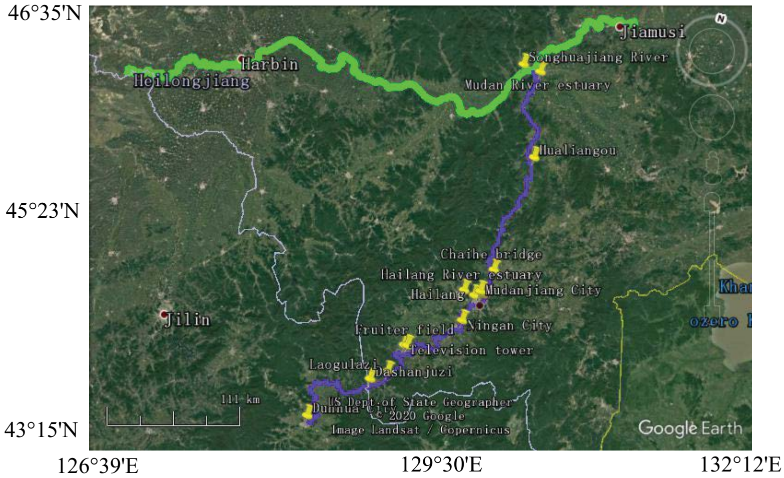

Fig. 1. Monitoring sections on the Mudan River Watershed.

Table 1
The monitoring section layout and ribbon class of the Mudan River Watershed

Water area	Monitoring section	Mark	Meaning of the monitoring section
Jingpo Lake	Laogulazi	S2	Representing the inflow water quality of the Jingpo Lake
	Television Tower	S3	Representing the water quality of the Jingpo Lake
	Fruiter Field	S4	Representing the effluent water quality of the Jingpo Lake
Hailang River	Hailang River estuary	S5	Representing the water quality from the Hailang River flowing into Mudan River
Mudan River mainstream	Hailang	S6	Representing the mixed water quality of the Hailang River and Mudan River
	Jiangbin bridge	S7	Representing the water quality of the industrial water
	Chaihe bridge	S8	Representing the water quality of the Mudan River
	Hualiangou	S9	Representing the effluent water quality of the Mudan River
	Mudan River estuary	S10	Representing the water quality from the Mudan River flowing into Songhua River
	Dashanjuzi	S1	Representing the inflow water quality of the Jilin Province

water quality standard limit based on the three dimensions of scope, frequency, and amplitude. The CCME WQI is an objective-based index that compares measured water quality values with standards to produce a score ranging from 0 to 100. The relevant computational formulas of CCME WQI are listed as follows [Eqs. (1)–(6)] [36–38]. The water quality index scoring system is presented in Table 2.

$$F_1 = \frac{P}{N} \times 100 \tag{1}$$

$$F_2 = \frac{q}{M} \times 100 \tag{2}$$

$$F_3 = \frac{Q}{0.01Q + 0.01} \tag{3}$$

$$Q = \frac{1}{M} \sum S \tag{4}$$

$$S = \frac{c_i}{c_s} - 1 \tag{5}$$

$$\text{WQI} = 100 - \frac{\sqrt{F_1^2 + F_2^2 + F_3^2}}{1.732} \tag{6}$$

where F_1 (scope) represents the percentage of the number of water quality parameters exceeding the standard limits in the total number of monitoring parameters; F_2 (frequency) represents the percentage of the number of monitoring data exceeding the standard limits in the total number of monitoring data; F_3 (amplitude) represents the amplitude; P represents the number of water quality parameters exceeding the standard; N represents the total number of water quality parameters; q represents the number of exceedances in all water quality monitoring data; M represents the total number of water quality monitoring data; S represents the multiple of the measured value of substandard water quality index deviating from the standard value; c_i represents the measured

Table 2
Water quality index scoring system of CCME WQI

Rank	Score	Description
Excellent	95~100	Water quality was not threatened or harmed by any pollution; Water quality conditions are very close to the natural pristine levels
Good	80~94	Water quality is only slightly threatened or impaired by pollution; Water quality conditions are basically maintained at a natural and satisfactory level
Acceptable	60~79	Water quality is occasionally threatened or damaged by pollution; Water quality conditions are sometimes unable to maintain the natural or satisfactory levels
Bad	45~59	Water quality is often threatened or damaged by pollution; Water quality conditions are often not maintained at the natural or satisfactory levels
Poor	0~44	Water quality is always threatened or damaged by pollution; Water quality conditions are usually not maintained at the natural or satisfactory levels

value of water quality index; c_s represents the standard limit of water quality index, and WQI represents the score of the comprehensive water quality index.

2.4. Multivariate statistical methods

Multivariate statistical techniques based on mathematical statistics are used to study multivariate problems. They mainly include linear regression analysis, discriminatory method, principal component analysis (PCA), cluster analysis, factor analysis, multivariate correspondence analysis, etc. Among them, PCA is usually applied to appraise the water quality. The core concept of PCA is to reduce the dimensionality of a dataset by converting multiple variables into a few comprehensive variables. Furthermore, each principal component that reflects most of the information is the linear combination of the original variables, and they are uncorrelated with each other among the principal components (PCs), simplifying the complex problems. Algebraically, these irrelevant PCs are ordered to the first variance (PC_1), the second variance (PC_2), etc. Generally, the matrix can be established as follows, with p representing the original variables, x_1, x_2, …, x_p [39,40].

$$PC_1 = a_{11}x_1 + a_{12}x_2 + ... + a_{1p}x_p = \sum_{p=1}^{p} a_{1p}x_p \quad (7)$$

$$PC_2 = a_{21}x_1 + a_{22}x_2 + ... + a_{2p}x_p = \sum_{p=1}^{p} a_{2p}x_p \quad (8)$$

$$PC_p = a_{p1}x_1 + a_{p2}x_2 + ... + a_{pp}x_p = \sum_{p=1}^{p} a_{pp}x_p \quad (9)$$

where $a_{m,p}$ is the component score coefficient for variable p on PC.

In this study, the water quality of the Mudanjiang River Watershed was evaluated by PCA with the aid of SPSS 19.0. The main analysis steps were as follows.

- Standardizing the original data and converting them to dimensionless data to eliminate the dimensional and order of magnitude effects between different indicators.
- Determining the correlation between indicators to evaluate whether the original variables are suitable for factor analysis.
- Calculating the eigenvalues (λ) and eigenvectors of the correlation coefficient matrix.
- Determining the number of principal components by assuming that the eigenvalue is greater than 1. If $\lambda \geq 1$, implies the explanatory power of the principal component can replace the original data. If $\lambda < 1$, it implies the explanatory power is insufficient.
- Forming the principal component coefficient matrix by dividing the eigenvector by the absolute value of the corresponding eigenroot.
- Multiplying the principal component coefficient and the standardized data to obtain the scores of each principal component.
- Obtain the comprehensive score value of the principal component. The higher the composite score is, the more severe is the pollution.

3. Results and discussion

3.1. Water quality variations of the Mudan River Watershed

3.1.1. Spatial variations

The mean values of the water quality variations measured at ten sampling stations on the Mudan River Watershed are presented in Fig. 2. According to the Case III standard of the "Environmental Quality Standards for Surface Water" (GB3838-2002) of China, the threshold values of COD_{Mn}, NH_4^+–N, TN, and TP are 6, 1, 1, and 0.2 mg/L, respectively. The majority of the stations met the requirements of the national standard, except for station S8 (Chaihe bridge). The mean values of COD_{Mn}, NH_4^+–N, TN, and TP at the station S8 were 6.34, 0.96, 1.13, and 0.135 mg/L. Compared to the other monitoring sections, the S8 station had the highest values, which indicated the inferior water quality at this section. This is mainly due to the numerous factories distributed around the Chaihe bridge, which discharge large quantities of sewage into the Mudan River. Moreover, agricultural activities, anthropogenic activities, and municipal wastewater also play an important role in causing severe pollution of the Chaihe bridge section. To further throw light on whether COD_{Mn}, NH_4^+–N, TN, and TP have significant variation among different sampling stations or not, data were analyzed by one-way ANOVA and Pearson correlation analysis. The results are shown in Table 3, which indicate that there are no significant differences among the ten sampling stations for COD_{Mn} (P_{ANOVA} = 0.307 > 0.05, $P_{Pearson}$ = 0.278 > 0.01), NH_4^+–N (P_{ANOVA} = 0.053 > 0.05, $P_{Pearson}$ = 0.095 > 0.01) and TP (P_{ANOVA} = 0.086 > 0.05, $P_{Pearson}$ = 0.091 > 0.01). However, only TN shows a significant difference at different sampling sites (P_{ANOVA} = 0.000 < 0.05, $P_{Pearson}$ = 0.000 < 0.01), which indicates there is a remarkable station-variable correlation between the sampling stations and the concentration of TN. It means that nitrogen pollution cannot be ignored by the government of the Mudan River Watershed, especially the Chaihe bridge section.

3.1.2. Seasonal variations

The mean values of the four water quality variables measured during the three seasons (rainy season, non-rainy season, and icebound season) are depicted in Fig. 3. All four water quality variables showed significant seasonal differences. Only the COD_{Mn} had higher values in the rainy season for all sampling sites. The average value of the COD_{Mn} was 6.28 mg/L, which significantly exceeded the Case III standard. This is largely due to the abundant rainfall causing the organic pollutants to flow into the river through surface water runoff. Especially at the sampling sites of S1, S2, S8, S9, and S10, sewage discharge from the surrounding factories also made a great contribution. However, in the non-rainy season and icebound season, the high COD_{Mn} load was relieved. All in all, the COD_{Mn} variation reaches a peak from June to September, while tends to be stable in the other period.

For NH_4^+–N, TN and TP, there existed an interesting variation compared to that of COD_{Mn}. NH_4^+–N, TN, and TP displayed high values in the icebound season. This is largely due to the low dissolved oxygen content at low temperatures, which results in the low activity of bacteria. The pollution is mainly from the release of bottom sludge and the weak dilution capacity in the icebound season. Especially in the middle sections of the Mudan River Watershed, such as S4~S9, the NH_4^+–N, TN, and TP

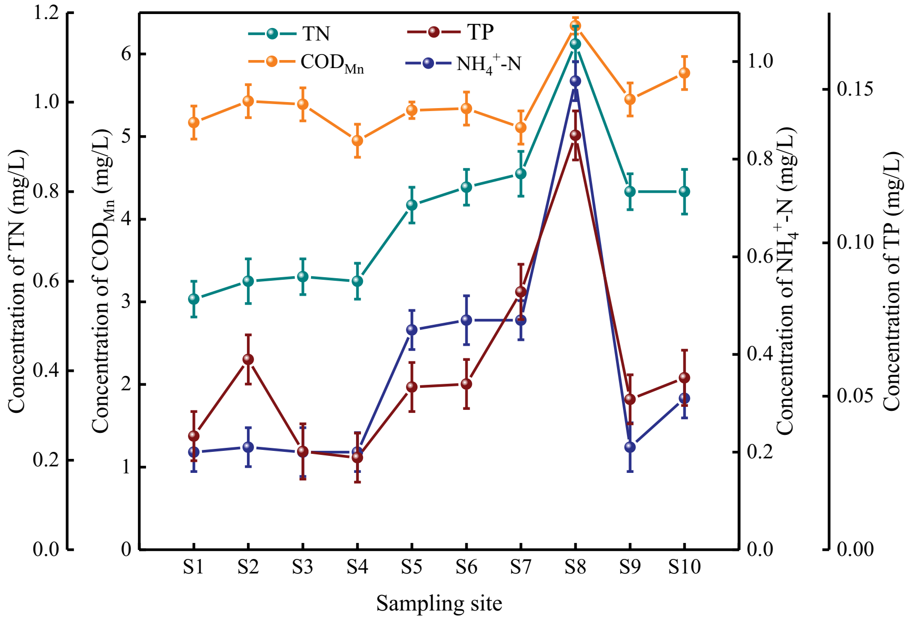

Fig. 2. The water quality variations at different sampling stations.

Table 3
One-way ANOVA (p < 0.05) and Pearson correlation (p < 0.01) for station-water quality variables and season-water quality variables

Water quality indexes	Station				Season			
	One-way ANOVA		Pearson correlation		One-way ANOVA		Pearson correlation	
	P	*F*	*P*	*r*	*P*	*F*	*P*	*r*
COD_{Mn}	0.307	1.279	0.278	0.205	0.000	11.834	0.000	–0.662
NH_4^+–N	0.053	2.357	0.095	0.310	0.173	1.872	0.073	0.333
TN	0.000	21.544	0.000	0.714	1.000	0.000	1.000	0.000
TP	0.086	2.056	0.091	0.314	0.353	1.083	0.224	0.229

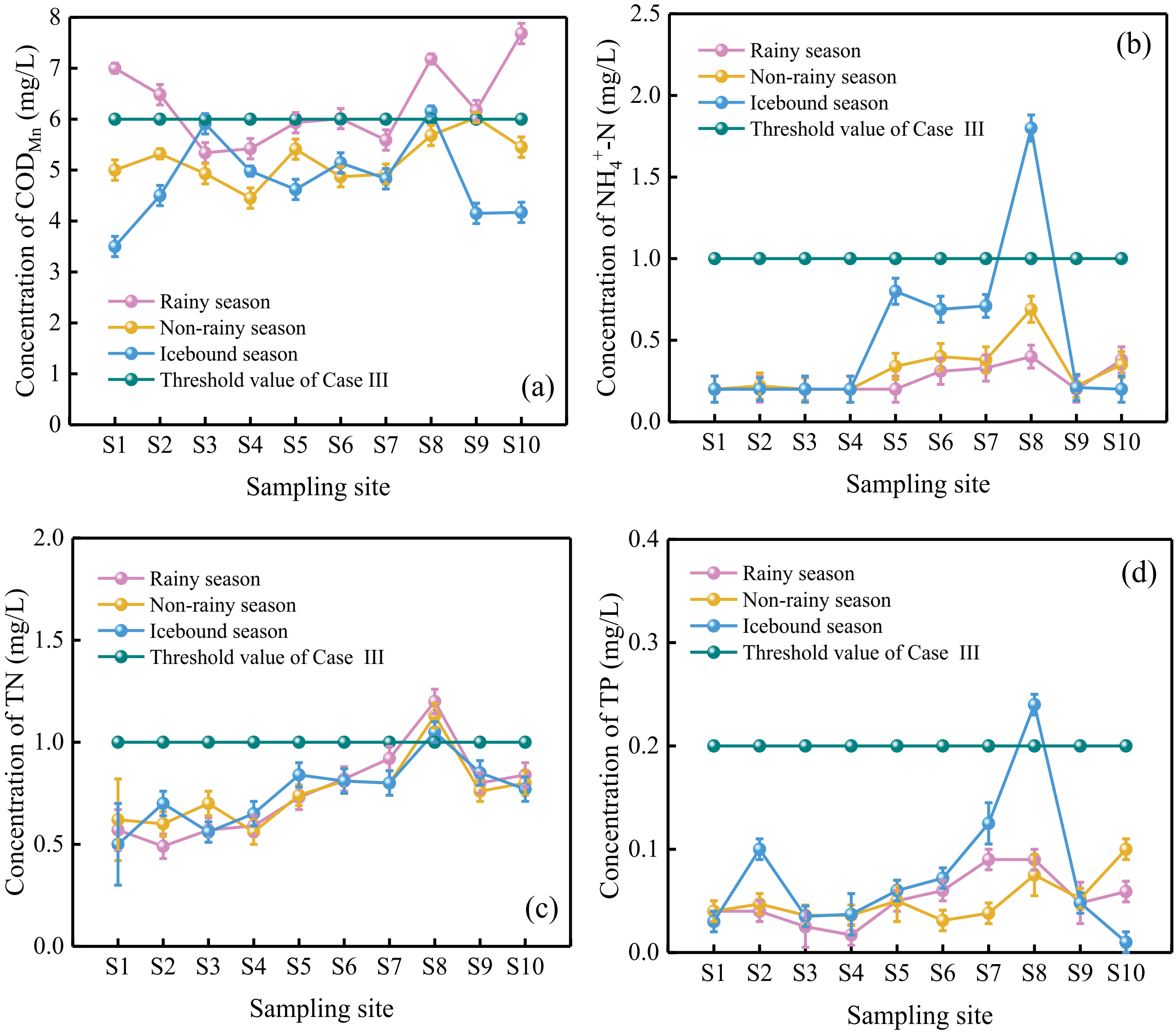

Fig. 3. The seasonal variations of the water quality at different sampling stations: (a) COD_{Mn}, (b) NH_4^+–N, (c) TN, and (d) TP.

variation showed higher values in comparison with the other sampling sites. For example, NH_4^+–N, TN, and TP displayed high values of 1.80, 1.05, and 0.24 mg/L at S8 in the icebound season, which are evidently beyond the limit of the Case III standard. On one hand, the strong dilution capacity of sampling sites S1 and S10 plays an important role in their positions as the head and tail catchment of the Mudan River Watershed. On the other hand, the sewage discharge of the surrounding factories and human activities increased the background values.

To illustrate the seasonal variations of the four water quality variables, data are also evaluated through season variable correlation analysis (Table 3). Results show that only COD_{Mn} shows pronounced correlation with $P_{ANOVA} = 0.000 < 0.05$, $P_{Pearson} = 0.000 < 0.01$, compared to that of NH_4^+–N, TN, and TP. It means that attention should be paid to organic pollution in the rainy season.

3.2. Assessment of the water quality index

It is very meaningful to comprehensively evaluate the water quality. In this work, CCME WQI was applied to distinguish the water quality of the Mudan River Watershed from the two aspects of season and sampling sites. The results are presented in Fig. 4. It can be seen that the WQI scores of the Mudan River Watershed in the rainy season, non-rainy season, and icebound season are 69.4, 71.0, and 42.0, which are rated as “Acceptable”, “Acceptable” and “Poor”, respectively. The river exhibits poor water quality in the icebound season because NH_4^+–N, TN, and TP are the primary control water quality indexes, while the main water quality index is only COD_{Mn} in the rainy season and non-rainy season. This indicates that enhanced treatment measures should be carried out to improve the water quality of the watershed.

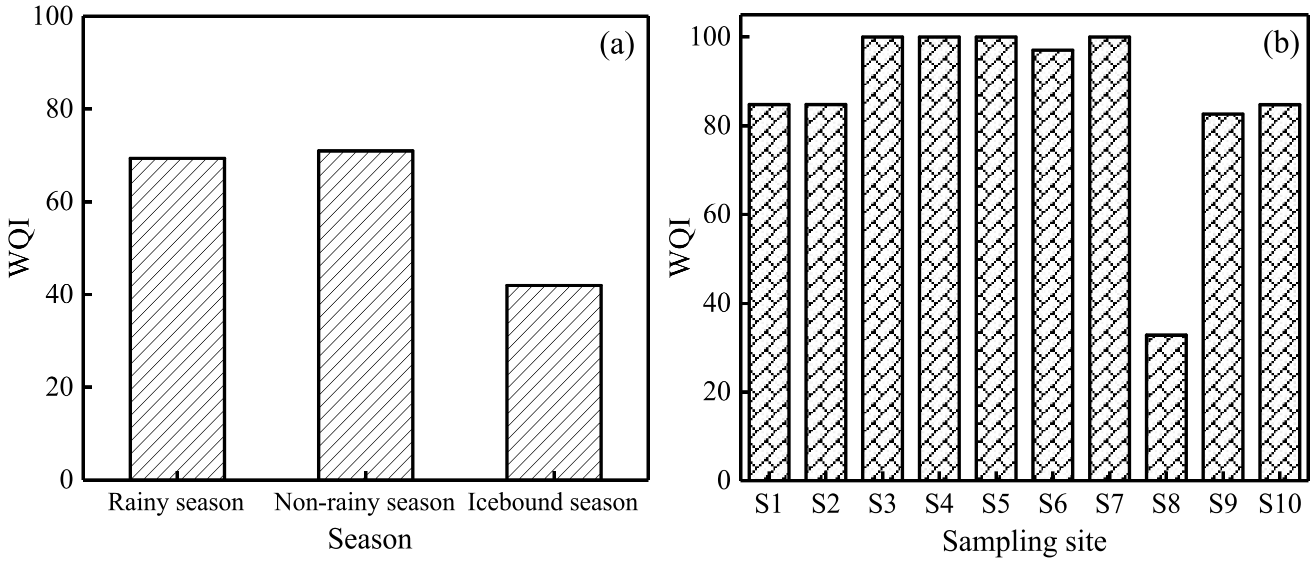

Fig. 4. Water quality index (WQI) for the Mudan River Watershed: (a) season and (b) sampling site.

The WQI scores at different sampling sites are shown in Fig. 4b. Stations S3, S4, S5, S6, and S7 showed scores of 100, 100, 100, 97.04, and 100, categorized as "Excellent", which indicated that the water quality of these regions had no anthropogenic impacts. The WQI scores of S1, S2, S9, and S10 were 84.8, 84.8. 82.7, and 84.8, rated as "Good", which indicated that the water quality of these regions was slightly polluted. Interestingly, the WQI score of S8 was 32.8, rated as "Poor", which indicated that the water quality of this region was severely damaged. A detailed explanation is given above. Therefore, it is urgent to curb the pollution of the Chaihe bridge section of the Mudan River Watershed.

3.3. *Principal components analysis for the assessment of water quality*

In this work, The PCA method was conducted to identify the correlated parameters used in WQI calculations and ascertain which one was responsible for most of the variance observed in the water quality data. After standardizing the original data, the correlation coefficients of COD_{Mn}, NH_4^+–N, TN, and TP in the rainy season, non-rainy season, and icebound season were calculated, as listed in Tables 4–6. The correlation coefficient between the majority of pollutant indexes is high. The correlation coefficients between NH_4^+–N and COD_{Mn}, TN, and TP are 0.525, 0.840, and 0.810. It indicates that the direct correlation between many variables is very strong, verifying that there is information overlap between these variables, making them suitable for the PCA method.

According to the calculation, there are two principal components screened out by the assumption that the eigenvalue is greater than 1, which are labeled as F_1 and F_2. The eigenvalue, variance contribution and the accumulative variance contribution of the studied indexes in the three seasons are presented in Table 7. The accumulative variance contributions in the three seasons were 92.919, 88.583, and 91.114, which imply that the principal components were COD_{Mn} and NH_4^+–N regardless of the season. It indicates that the Mudan River Watershed has the characteristics of organic pollution. According to the

Table 4
Correlation coefficient matrix of the indexes in the rainy season

	COD_{Mn}	NH_4^+–N	TN	TP
COD_{Mn}	1.000	0.525	0.345	0.355
NH_4^+–N	0.525	1.000	0.840	0.810
TN	0.345	0.840	1.000	0.873
TP	0.355	0.810	0.873	1.000

Table 5
Correlation coefficient matrix of the indexes in the non-rainy season

	COD_{Mn}	NH_4^+–N	TN	TP
COD_{Mn}	1.000	0.311	0.499	0.583
NH_4^+–N	0.311	1.000	0.941	0.446
TN	0.499	0.941	1.000	0.498
TP	0.583	0.446	0.498	1.000

Table 6
Correlation coefficient matrix of the indexes in the icebound season

	COD_{Mn}	NH_4^+–N	TN	TP
COD_{Mn}	1.000	0.603	0.389	0.586
NH_4^+–N	0.603	1.000	0.790	0.888
TN	0.389	0.790	1.000	0.709
TP	0.586	0.888	0.709	1.000

survey, there are numerous factories distributed on both sides of the Mudan River Watershed, which discharged large quantities of industrial effluents. Moreover, domestic wastewater and agricultural wastewater containing nitrogen and phosphorus also play an important role in the deterioration of water quality. Thus, it is urgent to vigorously develop the treatment of industrial effluents, enhance the treatment capacity of domestic sewage and agricultural wastewater, and improve the treatment facilities.

The component score coefficient matrix for the three seasons is shown in Table 8. According to these coefficients, the corresponding principal component expressions are as follows.

In the rainy season:

$$F_1 = 0.100x_1 + 0.160x_2 + 0.158x_3 + 0.156x_4 \quad (10)$$

$$F_2 = 0.158x_1 - 0.001x_2 - 0.169x_3 - 0.160x_4 \quad (11)$$

$$F = \frac{73.508}{92.919}F_1 + \frac{19.411}{92.919}F_2 \quad (12)$$

In the non-rainy season:

$$F_1 = 0.132x_1 + 0.161x_2 + 0.174x_3 + 0.142x_4 \quad (13)$$

$$F_2 = 0.327x_1 - 0.281x_2 - 0.190x_3 + 0.247x_4 \quad (14)$$

$$F = \frac{66.513}{88.583}F_1 + \frac{22.070}{88.583}F_2 \quad (15)$$

In the icebound season:

$$F_1 = 0.120x_1 + 0.159x_2 + 0.140x_3 + 0.154x_4 \quad (16)$$

$$F_2 = 0.531x_1 - 0.074x_2 - 0.325x_3 - 0.040x_4 \quad (17)$$

$$F = \frac{75.288}{91.114}F_1 + \frac{15.827}{91.114}F_2 \quad (18)$$

Based on the above comprehensive evaluation function, the comprehensive scores of water pollution in all monitoring sections are presented in Fig. 5. The results indicated that the comprehensive ranking of the water pollution was *S*8 > *S*10 > *S*5 > *S*6 > *S*9 > *S*7 > *S*1 > *S*3 > *S*4 > *S*2 in the rainy season, *S*8 > *S*10 > *S*9 > *S*1 > *S*5 > *S*2 > *S*7 > *S*6 > *S*3 > *S*4 in the non-rainy season, and *S*8 > *S*7 > *S*6 > *S*3 > *S*5 > *S*2 > *S*4 > *S*10 > *S*9 > *S*1 in the icebound season. It is obvious that sampling site *S*8-Chaihe bridge section is the most polluted monitoring section regardless of the season. This is mainly due to the numerous nearby factories, which discharge sewage into the river. In the rainy season, the F_1 score of *S*8 (2.269) was higher than that of the other sites, which indicated that the primary pollution of Chaihe bridge section was based on COD_{Mn}. The F_2 score of *S*2 (0.278) was maximum at all the sites, which indicated that NH_4^+–N was the primary pollution indicator of the Fruiter Field section. Hence, COD_{Mn} of the Chaihe bridge section and NH_4^+–N of the Fruiter Field section was the major control indexes in the rainy season. In the non-rainy season, the F_1 score of *S*8 (2.566) was the maximum among the studied sites, which showed COD_{Mn} was the primary pollutant of Chaihe bridge section. The F_2 score of *S*10 (0.425)

Table 7
The eigenvalue, variance contribution and the accumulative variance contribution of the studied indexes in the three season

Season	Principal component	Eigenvalue (λ)	Variance contribution (%)	Accumulative variance contribution (%)
Rainy season	F_1	5.881	73.508	73.508
	F_2	1.553	19.411	92.919
Non-rainy season	F_1	5.321	66.513	66.513
	F_2	1.766	22.070	88.583
Icebound season	F_1	6.023	75.288	75.288
	F_2	1.266	15.827	91.114

Table 8
Component score coefficient matrix in the three seasons

Index	Season					
	Rainy season		Non-rainy season		Icebound season	
	F_1	F_2	F_1	F_2	F_1	F_2
COD_{Mn}	0.100	0.158	0.132	0.327	0.120	0.531
NH_4^+–N	0.160	–0.001	0.161	–0.281	0.159	–0.074
TN	0.158	–0.169	0.174	–0.190	0.140	–0.325
TP	0.156	–0.160	0.142	0.247	0.154	–0.040

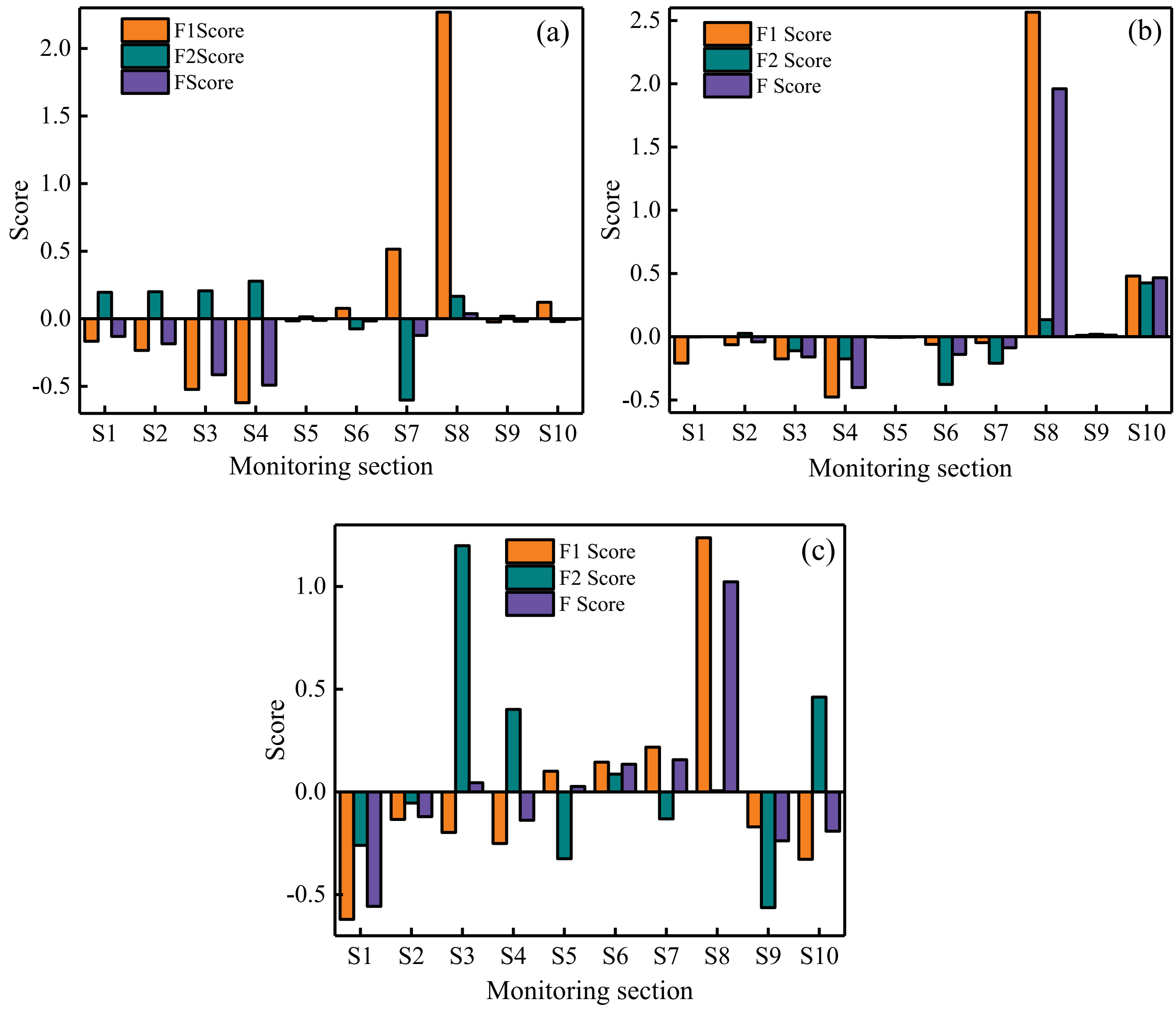

Fig. 5. The comprehensive evaluation result of water quality of each monitoring section: (a) in the rainy season, (b) in the non-rainy season, and (c) in the icebound season.

was higher than all the sites, which implied NH_4^+–N was the main pollution indicator of the Mudan River estuary section. Therefore, COD_{Mn} of the Chaihe bridge section and NH_4^+–N of the Mudan River estuary section was the major control indexes in the non-rainy season. Furthermore, in the icebound season, the F_1 score of *S*8 (1.237) was the peak value of the studied sites, which showed COD_{Mn} was the key pollutant of the Chaihe bridge section. The F_2 score of *S*3 (1.199) was the maximum among all the sites, which indicated NH_4^+–N was the primary pollution indicator of the Television Tower section. These results revealed that COD_{Mn} of the Chaihe bridge section and NH_4^+–N of the Television Tower section was the major control indexes in the icebound season. In addition, the comprehensive scores in the rainy season, non-rainy season, and icebound season were 0.038, 1.96, and 1.023, respectively. This indicated that enhanced treatment should be considered in the non-rainy season and icebound season due to the low water body self-purification and bacterial activity. Therefore, COD_{Mn} of the Chaihe bridge section and NH_4^+–N of the beginning and end sections of the watershed are the major control objectives to improve the water quality of the Mudan River Watershed, especially in the non-rainy season and icebound season.

Overall, the results obtained by the PCA method are consistent with the actual results and reflect the actual water quality of the Mudan River Watershed. Moreover, the pollution degree of different sections can also be more intuitively acquired by the PCA method, which can provide a more accurate basis for the protection and treatment of the Mudan River Watershed in the future.

4. Conclusions

This study was undertaken to evaluate the water quality of major monitoring sections located in the main stem of the Mudan River Watershed. Four water quality indexes (COD_{Mn}, NH_4^+–N, TN, and TP) were considered in this work, which may be useful for watershed management and aquatic body monitoring. The findings showed that there existed spatial and seasonal variations in the water quality of the different sampling stations on the Mudan River Watershed. Organic pollutants were the main contaminants of this watershed regardless of the monitoring section or season because of the numerous industrial effluents and surface water runoff. The WQI values in the rainy season, non-rainy season, and icebound season were calculated as 69.4, 71.0, and 42.0, which indicated that the water quality was "Acceptable", "Acceptable", and "Poor", respectively. Surface water runoff in the rainy season and low temperature and bacterial activity in the icebound season plays an important role. However, among the ten monitoring sections, the Chaihe bridge section had the lowest WQI value of 32.8, rated as "Poor" water quality. It is clear that industrial effluents, domestic discharge, and agricultural activities were major threats to the water quality of the watershed. Furthermore, the PCA was also applied to identify pollution sources and determine Spatio-temporal changes in the water quality of the watershed and make the data more objective in nature.

All the results reveal that the Mudan River Watershed is in a state of severe pollution, especially the Chaihe bridge section. Powerful management measures and improvement methods are necessary to ensure the water quality of this watershed.

Authorship conformation statement

Liang Liu and Ting-ting Cao contributed equally to this manuscript. All authors read and approved the final manuscript.

Authors disclosure statement

No competing financial interests exist.

Geolocation information

The Mudan River Watershed is located in the northeast of China (43°15′N~46°35′N, 126°39′E~132°12′E).

Acknowledgment

The research was financially supported by the Open Project of State Key Laboratory of Urban Water Resource and Environment, Harbin Institute of Technology (grant No. HCK201804).

References

[1] F. Yue, C. Liu, S. Li, Z. Zhao, X. Liu, H. Ding, B. Liu, J. Zhong, Analysis of $\delta^{15}N$ and $\delta^{18}O$ to identify nitrate sources and transformations in Songhua River, Northeast China, J. Hydrol., 519 (2014) 329–339.

[2] S. Yu, L. He, H. Lu, An environmental fairness based optimisation model for the decision-support of joint control over the water quantity and quality of a river basin, J. Hydrol., 535 (2016) 366–376.

[3] J. Liu, H. Dong, X. Tang, X. Sun, X. Han, B. Chen, C. Sun, B. Yang, Genotoxicity of water from the Songhua River, China, in 1994–1995 and 2002–2003: potential risks for human health, Environ. Pollut., 157 (2009) 357–364.

[4] Y. Wang, P. Wang, Y. Bai, Z. Tian, J. Li, X. Shao, L.F. Mustavich, B. Li, Assessment of surface water quality via multivariate statistical techniques: a case study of the Songhua River Harbin region, China, J. Hydro-environ. Res., 7 (2013) 30–40.

[5] Y. Zhai, X. Xia, G. Yang, H. Lu, G. Ma, G. Wang, Y. Teng, W. Yuan, S. Shrestha, Trend, seasonality and relationships of aquatic environmental quality indicators and implications: an experience from Songhua River, NE China, Ecol. Eng., 145 (2020) 105706–105715.

[6] O. Ighalo, A.G. Adeniyi, A comprehensive review of water quality monitoring and assessment in Nigeria, Chemosphere, 260 (2020) 127569–127595.

[7] E.R. Kelly, R. Cronk, E. Kumpel, G. Howard, J. Bartram, How we assess water safety: a critical review of sanitary inspection and water quality analysis, Sci. Total Environ., 718 (2020) 137237–137245.

[8] A. Seifi, M. Dehghani, V.P. Singh, Uncertainty analysis of water quality index (WQI) for groundwater quality evaluation: application of Monte–Carlo method for weight allocation, Ecol. Indic., 117 (2020) 106653–106667.

[9] S.K. Gautam, S.K. Singh, K.S. Rawat, Intrinsic vulnerability evaluation of groundwater nitrate pollution along a course of the Subarnarekha River in Jharkhand, India, Water Conserv. Sci. Eng., 6 (2021) 55–66.

[10] V.K. Singh, D. Kumar, S.K. Singh, Q.B. Pham, N.T.T. Linh, S. Mohammed, D.T. Anh, Development of fuzzy analytic hierarchy process based water quality model of Upper Ganga river basin, India, J. Environ. Manage., 284 (2021) 111985–112003.

[11] S.K. Gautam, J.K. Tripathi, S.K. Singh, Assessing the suitability of Ghaghra River water for irrigation purpose in India, Agric. Water Manage., (2021) 67–81, doi: 10.1016/B978-0-12-812362-1.00005-9.

[12] E. Maliqi, K. Jusufi, S.K. Singh, Assessment and spatial mapping of groundwater quality parameters using metal pollution indices, graphical methods and geoinformatics, Anal. Chem. Lett., 10 (2020) 152–180.

[13] S.K. Gautam, T. Evangelos, S.K. Singh, J.K. Tripathi, A.K. Singh, Environmental monitoring of water resources with the use of PoS index: a case study from Subarnarekha River basin, India, Environ. Earth Sci., 77 (2018) 70–86.

[14] T. Srebotnjak, G. Carr, A. de Sherbinin, C. Rickwood, A global water quality index and hot-deck imputation of missing data, Ecol. Indic., 17 (2012) 108–119.

[15] X. Nong, D. Shao, H. Zhong, J. Liang, Evaluation of water quality in the South-to-North Water Diversion Project of China using the water quality index (WQI) method, Water Res., 178 (2020) 115781–115795.

[16] H. Singh, D. Singh, S.K. Singh, D.N. Shukla, Assessment of river water quality and ecological diversity through multivariate statistical techniques, and earth observation dataset of rivers Ghaghara and Gandak, India, Int. J. River Basin Manage., 15 (2017) 347–360.

[17] H. Singh, R. Pandey, S.K. Singh, D.N. Shukla, Assessment of heavy metal contamination in the sediment of the River Ghaghara, a major tributary of the River Ganga in Northern India, Appl. Water Sci., 7 (2017) 4133–4149.

[18] W. Hou, S. Sun, M. Wang, X. Li, N. Zhang, X. Xin, L. Sun, W. Li, R. Jia, Assessing water quality of five typical reservoirs in lower reaches of Yellow River, China: using a water quality index method, Ecol. Indic., 61 (2016) 309–316.

[19] Z. Ma, H. Li, Z. Ye, J. Wen, Y. Hu, Y. Liu, Application of modified water quality index (WQI) in the assessment of coastal water quality in main aquaculture areas of Dalian, China, Mar. Pollut. Bull., 157 (2020) 111285–111292.

[20] S. Şener, E. Şener, A. Davraz, Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey), Sci. Total Environ., 584–585 (2017) 131–144.

[21] K.P. Singh, A. Malik, D. Mohan, S. Sinha, Multivariate statistical techniques for the evaluation of spatial and temporal

variations in water quality of Gomti River (India)—a case study, Water Res., 38 (2004) 3980–3992.
[22] T. Li, S. Li, C.T. Liang, R. Bush, L. Xiong, Y. Jiang, A comparative assessment of Australia's Lower Lakes water quality under extreme drought and post-drought conditions using multivariate statistical techniques, J. Cleaner Prod., 190 (2018) 1–11.
[23] R. Trabelsi, K. Zouari, Coupled geochemical modeling and multivariate statistical analysis approach for the assessment of groundwater quality in irrigated areas: a study from North Eastern of Tunisia, Groundwater Sustainable Dev., 8 (2019) 413–427.
[24] S. Platikanov, D. Baquero, S. González, J. Martín-Alonso, M. Paraira, J.L. Cortina, R. Tauler, Chemometric analysis for river water quality assessment at the intake of drinking water treatment plants, Sci. Total Environ., 667 (2019) 552–562.
[25] M. Varol, Spatio-temporal changes in surface water quality and sediment phosphorus content of a large reservoir in Turkey, Environ. Pollut., 259 (2020) 113860–113868.
[26] S.K. Singh, C.K. Singh, K.S. Kumar, R. Gupta, S. Mukherjee, Spatial-temporal monitoring of groundwater using multivariate statistical techniques in Bareilly district of Uttar Pradesh, India, J. Hydrol. Hydromech., 57 (2009) 45–54.
[27] M. Tripathi, 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, Ecol. Indic., 96 (2019) 430–436.
[28] M. Varol, Use of water quality index and multivariate statistical methods for the evaluation of water quality of a stream affected by multiple stressors: a case study, Environ. Pollut., 266 (2020) 115417–115426.
[29] K. Yang, Z. Yu, Y. Luo, Y. Yang, L. Zhao, X. Zhou, Spatial and temporal variations in the relationship between lake water surface temperatures and water quality – a case study of Dianchi Lake, Sci. Total Environ., 624 (2018) 859–871.
[30] J. Unda-Calvo, E. Ruiz-Romera, M. Martínez-Santos, M. Vidal, I. Antigüedad, Multivariate statistical analyses for water and sediment quality index development: a study of susceptibility in an urban river, Sci. Total Environ., 711 (2020) 135026–135038.
[31] M. Camara, N.R. Jamil, A.F.B. Abdullah, R. Hashim, A.G. Aliyu, Economic and efficiency based optimisation of water quality monitoring network for land use impact assessment, Sci. Total Environ., 737 (2020) 139800–138813.
[32] S. Sudhakaran, H. Mahadevan, V. Arun, A.P. Krishnakumar, K.A. Krishnan, A multivariate statistical approach in assessing the quality of potable and irrigation water environs of the Netravati River basin (India), Groundwater Sustainable Dev., 11 (2020) 100462–100471.
[33] J. Zhang, S. Li, C. Jiang, Effects of land use on water quality in a River Basin (Daning) of the Three Gorges Reservoir Area, China: watershed versus riparian zone, Ecol. Indic., 113 (2020) 106226–106236.
[34] Z. Li, J. Ma, J. Guo, H.W. Paerl, J.D. Brookes, Y. Xiao, F. Fang, W. Ouyang, L. Lu, Water quality trends in the Three Gorges Reservoir region before and after impoundment (1992–2016), Ecohydrol. Hydrobiol., 19 (2019) 317–327.
[35] M. Varol, B. Gokot, A. Bekleyen, B. Sen, Spatial and temporal variations in surface water quality of the dam reservoirs in the Tigris River basin, Turkey, Catena, 92 (2012) 11–21.
[36] A. Akkoyunlu, M.E. Akiner, Pollution evaluation in streams using water quality indices: a case study from Turkey's Sapanca Lake Basin, Ecol. Indic., 18 (2012) 501–511.
[37] T. Hurley, R. Sadiq, A. Mazumder, Adaptation and evaluation of the Canadian Council of Ministers of the Environment Water Quality Index (CCME WQI) for use as an effective tool to characterize drinking source water quality, Water Res., 46 (2012) 3544–3552.
[38] G.D. Gikas, G.K. Sylaios, V.A. Tsihrintzis, I.K. Konstantinou, T. Albanis, I. Boskidis, Comparative evaluation of river chemical status based on WFD methodology and CCME water quality index, Sci. Total Environ., 745 (2020) 140849–140863.
[39] R.L. Olsen, R.W. Chappell, J.C. Loftis, Water quality sample collection, data treatment and results presentation for principal components analysis-literature review and Illinois River watershed case study, Water Res., 46 (2012) 3110–3122.
[40] J.M. Paca, F.M. Santos, J.C.M. Pires, A.A. Leitão, R.A.R. Boaventura, Quality assessment of water intended for human consumption from Kwanza, Dande and Bengo rivers (Angola), Environ. Pollut., 254 (2019) 113037–113044.