

# A GENERALIZED WATER QUALITY INDEX FOR TAIWAN

SHIOW-MEY LIOU\*, SHANG-LIEN LO and SHAN-HSIEN WANG

*Graduate Institute of Environmental Engineering, National Taiwan University, Taipei, Taiwan*

*(\* author for correspondence, e-mail: d87541004@ntu.edu.tw)*

(Received 28 October 2002; accepted 23 June 2003)

**Abstract.** By exploring the behavior and limitations of conventional methods for quality evaluation, a better overall index for water quality in Taiwan and its application in Keya River are proposed. The thirteen variables are employed for the quality assessment. Numerical scales relating to the degree of quality are established for each variable to assess variations in quality and to convey findings in a comprehensive manner to others. The employed variables are categorized into three aspects referred to the principal components analysis, which are 'organics', 'particulates' and 'microorganisms'. The overall index consists of the geometric mean for the three defined aspects and multiplying by three coefficients, which reflect quality degrees of temperature, pH and toxicity, respectively. Besides, the arithmetic mean is addressed among the variables associated in the same aspect. Finally, a fresh, mixed aggregation function is developed for ensuring that the index score tells us something useful and does not hide important information. A case study of the Keya River is made to illustrate the application of this new index system. The new proposed index provides comprehensive information for water quality assessment, than the existing index used in Taiwan does.

**Keywords:** aggregation function, principal component analysis, rating curve, water quality index

## 1. Introduction

Landwehr (1979) points out that an index is a performance measurement that aggregates information into a usable form, which reflects the composite influence of significant physical, chemical and biological parameters of water quality conditions. House and Newsome (1989) states that the use of a Water Quality Index (WQI) allows 'good' and 'bad' water quality to be quantified by reducing a large quantity of data on a range of physico-chemical and biological variables to be a single number in a simple, objective and reproducible manner. Since 1965, when Horton (1965) proposed the first water quality index (WQI), a great deal of consideration has been given to the development of index methods. Indices define a unique rating curve for each variable, by which its values are interpreted using a questionnaire, in terms of conceptual quality units, or some set of standards. Various types of arithmetical methods used, have included aggregation of quality-monitoring data to yield an overall quality index. Examples include the weighted averaging methods of Brown *et al.* (1970), Ross *et al.* (1977), Ball *et al.* (1980), House and Ellis (1987), Stanbuk-Giljanovic (1999), Pesce and Wunderlin (2000) and Jonnalagadda and Mehre (2001); weighted geometric means (Dini-



TABLE I  
The classification ranks defined by the existing River Pollution Index (RPI)

Items/ranks	Good	Slightly polluted	Moderate polluted	Gross polluted
DO (mg L <sup>-1</sup> )	Above 6.5	4.6–6.5	2.0–4.5	Under 2.0
BOD <sub>5</sub> (mg L <sup>-1</sup> )	Under 3.0	3.0–4.9	5.0–15	Above 15
SS (mg L <sup>-1</sup> )	Under 20	20–49	50–100	Above 100
NH <sub>3</sub> -N (mg L <sup>-1</sup> )	Under 0.5	0.5–0.99	1.0–3.0	Above 3.0
Index Scores ( $S_i$ )	1	3	6	10
Sub-index	Under 2	2.0–3.0	3.1–6.0	Above 6.0

$$\text{Sub-index} = \frac{1}{4} \sum_{i=1}^4 S_i.$$

use, 1987); minimum operators (Smith, 1990); and hybrid methods (Dojlido *et al.*, 1994; Swamee and Tyagi, 2000). Aggregation can occasionally mask developing problems associated with trends in one of the quality measurements in the index. Ott (1978) provides an excellent treatise on exploring the characteristics of aggregations. Swamee and Tyagi (2000) states that even the minimum operator does not have an eclipsing problem, but it is not suitable as an aggregation, because it fails to give a composite picture of water quality.

The Principal Components Analysis (PCA) is used widely for identifying relevant groups of water and the factors that realize a change in their quality. Reisenhofer *et al.* (1998) investigates that the PCA supports the results of multivariate cluster analysis, which conforms the discriminating ability of the considered parameters of the Timavo River. Ceballos *et al.* (1998) summarizes that PCA shows water typology defined by 9 variables, grouped in two principal factors for three lakes, located in northeast Brazil. Morales *et al.* (1999) explores the quality of coastal water in the Gulf of Valencia (Spain) by factor analysis. The PCA is utilized to reduce the dimensionality of the data sets by explaining the variance among a large set of correlated variables in terms of a small number of Principal Components (PCs). Hence, the PCA is one of the most powerful and common techniques used for reducing the dimensionality of large sets of data without loss of information (Vega *et al.*, 1998; Gómez, 1999; Helena *et al.*, 2000; Pesce and Wunderlin, 2000).

The River Pollution Index (RPI) is an existing index, which is employed by EPA of Taiwan to explore monitor trends for both planning and day-to-day management of surface water quality for the public currently. The RPI involves four variables: dissolved oxygen (DO), biochemical oxygen demand (BOD<sub>5</sub>), suspended solids (SS), and ammonia nitrogen (NH<sub>3</sub>-N), each of which is ultimately converted to a four-state quality sub-index (1, 3, 6, and 10). The overall index is then divided into four pollution levels (non-polluted, lightly-polluted, moderately-polluted, and grossly-polluted) by averaging the four sub-indices (see Table I) (Liou *et al.*, 2003).

The existing index mainly focuses on the organic pollutions caused by the agriculture activities. For representing updated information from the dramatic changes of possible discharge sources, especially after the enormous economic developments and human activities throughout Taiwan in 1980's, the overall water quality index for Taiwan ought to be reviewed. A generalized pattern for water quality assessment is attempted to develop by linking the two mature methodologies mentioned above, WQI and PCA. Rating curves are used in data preprocessing. The PCA is proposed for categorizing the employed variables in accordance with common features. Finally, a fresh look at the aggregation process is proposed after investigating the philosophy of mathematical aggregation functions. The intention is to produce a generalized aggregation method for water quality, which could provide vital and usable information.

## 2. Study Site and Methods

### 2.1. STUDY AREA

The salient features of Taiwan impacted on water quality are as follows. First, the population density of Taiwan is the second highest in the world. The domestic sewage has been the major water pollution source for many urban rivers in Taiwan, because the sewer system around the Taiwan area is still below 10%. Second, wastewater from the swine industry has been a major water pollution source many years ago. The domestic pig farms are gradually reducing in size after joining the World Trade Organization (WTO), but there are still about 7 million pigs being raised in Taiwan. Third, the last three decades Taiwan has begun to develop itself into a big trading economy in the world. There are nearly 11 000 factories and wastewater, containing various contaminants, is usually produced by the manufacturing processes of these factories. For understanding the sources of river water pollution, assessing the water resources and establish long-term water quality data for controlling the polluted segments of the river, the program of water quality monitoring has been conducted for many years. The main goal of water quality inspection and analysis is to provide the information for checking out the cleanup effects as well. Currently the EPA has completed the investigation of the background information of 339 monitoring locations around the island and has put this information and monitoring data on the website to enhance circulation and sharing. Figure 1 shows the geographical location of 21 main rivers around Taiwan, to which regular monitoring station are distributed.

### 2.2. DATA TREATMENT AND DATA PROCESSING

The thirteen variables are employed, which are dissolved oxygen (DO,  $\text{mg L}^{-1}$ ), organics oxygen demand ( $\text{BOD}_5$ ,  $\text{mg L}^{-1}$ ), ammonia nitrogen ( $\text{NH}_3\text{-N}$ ,  $\text{mg L}^{-1}$ ), faecal coliforms (CFU/100 mL), turbidity (NTU), suspended solids (SS,  $\text{mg L}^{-1}$ ),

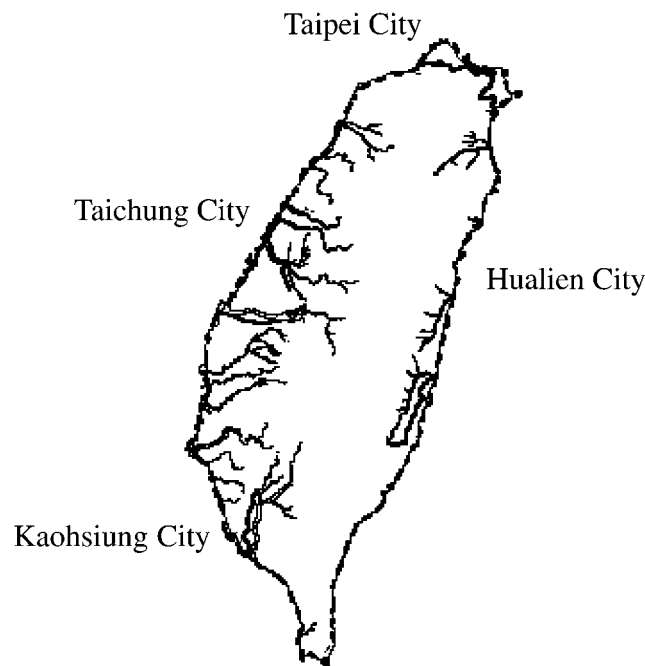


Figure 1. Geographical location of 21 main rivers around Taiwan, to which regular monitoring stations are distributed.

temperature ( $^{\circ}\text{C}$ ), pH, cadmium ( $\text{Cd}$ ,  $\text{mg L}^{-1}$ ), lead ( $\text{Pb}$ ,  $\text{mg L}^{-1}$ ), chromium ( $\text{Cr}$ ,  $\text{mg L}^{-1}$ ), copper ( $\text{Cu}$ ,  $\text{mg L}^{-1}$ ) and zinc ( $\text{Zn}$ ,  $\text{mg L}^{-1}$ ). In surface water quality,  $\text{DO}$ ,  $\text{BOD}_5$ ,  $\text{NH}_3\text{-N}$ , faecal coliforms, turbidity and  $\text{SS}$  are potential contaminants contributed by microbial communities and human activities.  $\text{Cd}$ ,  $\text{Pb}$ ,  $\text{Cr}$ ,  $\text{Cu}$  and  $\text{Zn}$  are toxic substances found to cause adverse effects on human health. Measurements of environmental monitoring variables encompass natural scales that may range from a few micrograms per litre to hundreds of grams per litre, or may

TABLE II

The source water criteria for heavy metals (Cd, As, Cr, Cu and Zn) in Taiwan

Item		Maximum permissible concentration (mg L <sup>-1</sup> )
Cadmium	(Cd)	0.01
Lead	(Pb)	0.1
Chromium	(Cr)	0.05
Copper	(Cu)	0.03
Zinc	(Zn)	0.03

require comparison among measurements with different natural units. A simple numerical scale relating to degree of quality would seem a feasible approach to assess variations in quality and to convey findings in a comprehensive manner to other (Brocon *et al.*, 1970; Stambuk-Giljanovič, 1999; Dinius, 1987; Smith, 1990; Alberto *et al.*, 2001). Measurements for each variable are converted to values on an interval scale ranging from 0 to 100, in accordance with the degree of water quality from worst to highest. The criteria for the subindices refer to the classification of national water, source water criteria adopted by other countries, the background data of water quality, the legislated standards of Taiwan, etc. Among the employed variables, Cd, Pb, Cr, Cu and Zn are regarded as toxic substance and rating curves for each of them is referred to the concentration ratio, which is calculated from the following equation (Dee *et al.*, 1973; Schierow and Chesters, 1988):

$$r_i = \frac{C_i}{S_i}$$

where  $r_i$  = concentration ratio of  $i$  substance,  $C_i$  = substance with concentration (mg L<sup>-1</sup>), and  $S_i$  = maximum permissible concentration (mg L<sup>-1</sup>). Each of them is compared to management criteria representing the upper limits of effluent concentration. Table II shows the surface water criteria for heavy metals in Taiwan.

Following the scaling of variables, the principal components analysis (PCA) is applied to recognize the common features among the variables. A principal component analysis is to account  $k$  new principal components (PCs), which are obtained as weighted linear combinations of  $p$  original components (Johnson and Wichern, 1998). The software of S-plus 4.5 is employed for PCA analysis in the study. The correlation coefficients between PCs and parameters, eigenvalues of PCs and the cumulative percentage of total variance are obtained.

TABLE III  
The range of variables and their key-points defined for rating curves

Score	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
value	Faecal coliform (MPN/ 100 mL)	Dissolved oxygen (mg L <sup>-1</sup> )	Biochemical oxygen demand (mg L <sup>-1</sup> )	Ammonia nitrogen (mg L <sup>-1</sup> )	Suspended solid (mg L <sup>-1</sup> )	Turbidity (NTU)	Tempera- ture (°C)	pH	Toxicity (r*) <sup>a</sup>
100	0	≥6.55	≤2.95	≤0.45	≤19.5	0	≤38	≥6; ≤9	0
90	6								
80	50					4			
70	5000	5.55	3.95	0.7	34.5	15			
50	10000				30				
30		3.25	10	2	75	50			
0	1000000	≤2.04	≥15.05	≥3.05	≥100	120	>38	>9; <6	≥1

<sup>a</sup> r\* means ration of the particular toxic substance to its maximum permissible concentration.

### 2.3. AGGREGATION FUNCTIONS

A mixed aggregation process is recommended to produce the final index score to which both of additive form and multiplicative form are appointed. Smith (1990) and Ott (1978) point out that the problem of eclipsing might be occurred when the additive form is applied in aggregation. Suppose that an overall index is formed consisting of just two subindices,  $I_1$  and  $I_2$ . Eclipsing can occur over a relatively wide range of value,  $10 < I \leq 55$  for the example of point  $I_1 = 10$  and  $I_2 = 100$ , giving  $I = (0.5)(10) + (0.5)(100) = 55$ . That is to say, when extremely poor water quality exists for at least one variable, but the overall index does not represent this fact (Swamee and Tyagi, 2000; Ott, 1978). It is plausible that the scruple of eclipsing could be eliminated and would not exist when applying the arithmetic mean to the positive correlated variables. The positive correlated variables are determined in having the dependent trends and it is unreasonable that only one variable would ever give a very low rating, when all others produced maximum ratings. Under this specific consideration, the arithmetic mean could be assigned to aggregation.

The multiplicative aggregation function is one of the most popular aggregation techniques. In multiplicative aggregation function, if any variable exhibits low score value, the overall index will exhibit poor environmental quality. Conversely, the overall index is zero if and only if at least one subindex is zero. Smith (1990) points out that, the more parameters used in multiplicative aggregation function, the lower still will be the sensitivity, because the individual weights will be smaller. If weighting is very close to zero, the subindex is transformed in such a matter that it can take on essentially just two states, 0 and 1. It becomes, in effect, a dichotomous subindex, a step function (Ott, 1978). Suppose, for example, that an overall index

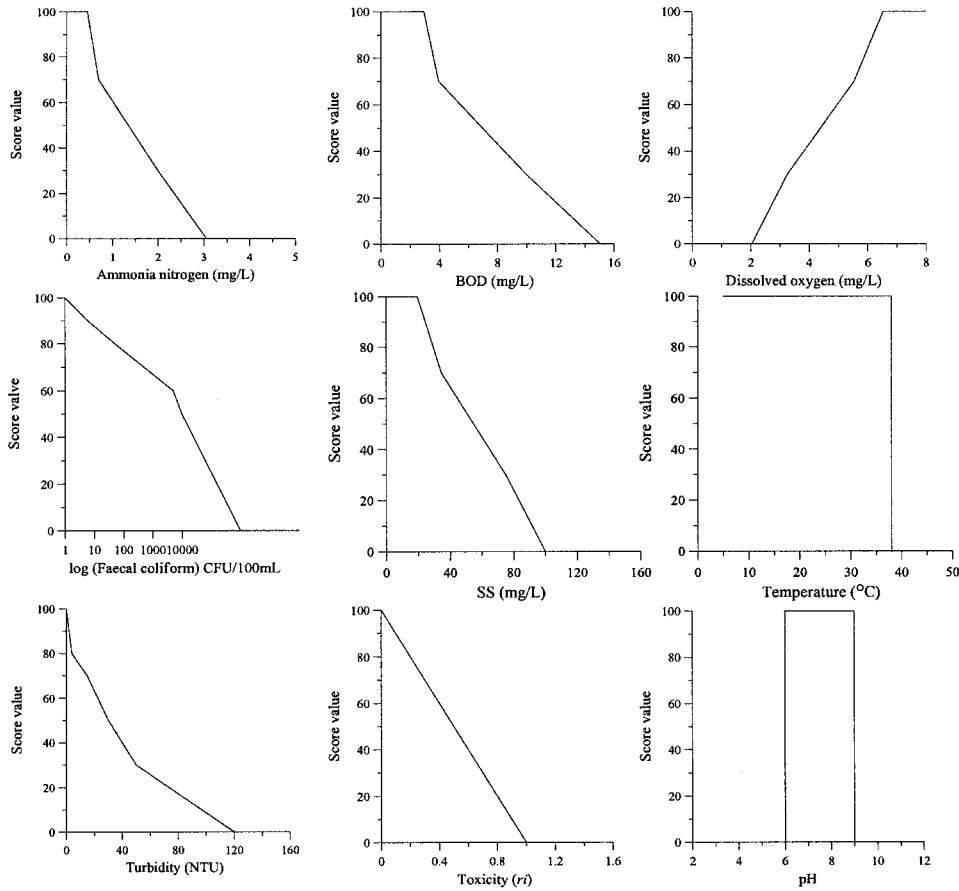


Figure 2. The assigned rating curves for the studied variables.

is developed with nine subindices ( $I_1, I_2, \dots, I_9$ ). Except of just one subindex,  $I_1$ , all the other subindices reflect good quality (that is,  $I_i = 100$  for  $i = 2, 3, \dots, 9$ ). The following relationship results:  $I = (100)^{8/9} (I_1)^{1/9}$ . If  $I_1 = 0$ , for example, then  $I = 0$ ; however, if  $I_1$  increases from 0 to only 2, the overall index  $I$  will increase from 0 to 64.8, a striking change. By contrast, if  $I_1$  increases from 20 to 100,  $I$  will increase rather gradually from approximately 80 to 100. Therefore, an increase of just a few units when  $I_1$  is zero causes  $I$  to increase over half its total range. Thus, addressing the geometric mean among fewer items could eliminate the ambiguous caused from smaller weightings.

### 3. Results and Discussion

Table III shows the range of possible investigations of studied variables and their score values for water quality. The rating curves for the employed variables are

TABLE IV  
The statistic summary of the 2530 studied data of Taiwan

Parameters	Analytical detection limit	Min	Max	Median	Mean	St. Dev. <sup>a</sup>	C.V. <sup>b</sup>
DO (mg L <sup>-1</sup> )	N.A. <sup>c</sup>	0	16	6.9	5.97	2.98	0.49
BOD (mg L <sup>-1</sup> )	N.A.	0.02	462	2.9	6.66	14.67	2.20
SS (mg L <sup>-1</sup> )	N.A.	1	15600	26	116.25	598.11	5.14
NH <sub>3</sub> -N (mg L <sup>-1</sup> )	N.A.	0	112	0.62	3.16	7.04	2.22
pH	N.A.	3.5	10.5	7.7	7.71	0.48	0.06
Faecal coliform (CFU/100 mL)	N.A.	0.01	7.1×10 <sup>7</sup>	12000	245962.3	1763593	7.17
Temperature (°C)	N.A.	2.8	35	24	23.66	4.42	0.18
Turbidity (NTU)	N.A.	0.065	72000	19	155.23	2258.26	14.54
Cd (mg L <sup>-1</sup> )	0.005	0.005	0.03	0.005	0.005	0.001	0.304
Pb (mg L <sup>-1</sup> )	0.01	0.01	1.45	0.01	0.02	0.04	2.25
Cr (mg L <sup>-1</sup> )	0.01	0.01	3.18	0.01	0.01	0.07	3.96
Cu (mg L <sup>-1</sup> )	0.01	0.01	1.2	0.01	0.02	0.05	2.39
Zn (mg L <sup>-1</sup> )	0.01	0.01	4.7	0.03	0.08	0.21	2.42

<sup>a</sup> Standard deviation.

<sup>b</sup> Coefficient of variation (St. Dev./Mean).

<sup>c</sup> N.A. = Not applicable.



TABLE V  
Correlation matrix for the studied data

	DO	BOD <sub>5</sub>	NH <sub>3</sub> -N	Faecal coliform	Turb.	SS	Temp.	pH	Toxicity
DO	1.0000								
BOD <sub>5</sub>	<b>0.6618</b>	1.0000							
NH <sub>3</sub> -N	<b>0.7258</b>	<b>0.6618</b>	1.0000						
Faecal coliform	0.3779	0.3686	0.3913	1.0000					
Turb.	-0.0120	0.0667	-0.0189	0.1454	1.0000				
SS	-0.0832	0.0285	-0.0850	0.1025	<b>0.8042</b>	1.0000			
Temperature	NA	NA	NA	NA	NA	NA	NA		
pH	0.0005	0.0249	-0.0262	-0.0049	-0.008	0.0079	NA	1.0000	
Toxicity	0.0035	0.0379	-0.0052	0.0638	0.1042	0.1246	NA	0.0004	1.0000

TABLE VI

The structure matrix, eigenvalues and cumulative percentage for the extracted principal components

		PC1	PC2	PC3	PC4	PC5
DO	(x <sub>1</sub> )	<b>0.535</b>				-0.190
BOD	(x <sub>2</sub> )	<b>0.526</b>				-0.258
SS	(x <sub>3</sub> )		<b>0.685</b>		0.129	-0.104
Ammonia nitrogen	(x <sub>4</sub> )	<b>0.542</b>				-0.182
pH	(x <sub>5</sub> )			<b>0.998</b>		
Turbidity	(x <sub>6</sub> )		<b>0.679</b>		0.158	
Faecal coliform	(x <sub>7</sub> )	0.375	0.132			<b>0.915</b>
Toxic substances	(x <sub>8</sub> )		0.189		<b>-0.976</b>	
Eigenvalues		1.6284	1.3686	1.0010	0.9845	0.8400
Cumulative percentage (%)		33.15	56.56	69.09	81.21	90.03

shown in Figure 2, which are the piecewise linear membership functions based on Table III. The segmented linear functions are recommended for pH and temperature because, unlike the others, the two variables can not readily to be rated to show gradations in quality, but fall more into the category of 'yes or no' indicators. Because of the variation in the natural temperature of rivers occurs in geographic location and seasonal interaction, it is impossible to assign rating values for different temperature levels. There is, however, an upper limit on temperature beyond which deterioration in quality is clearly recognizable. Temperature of the

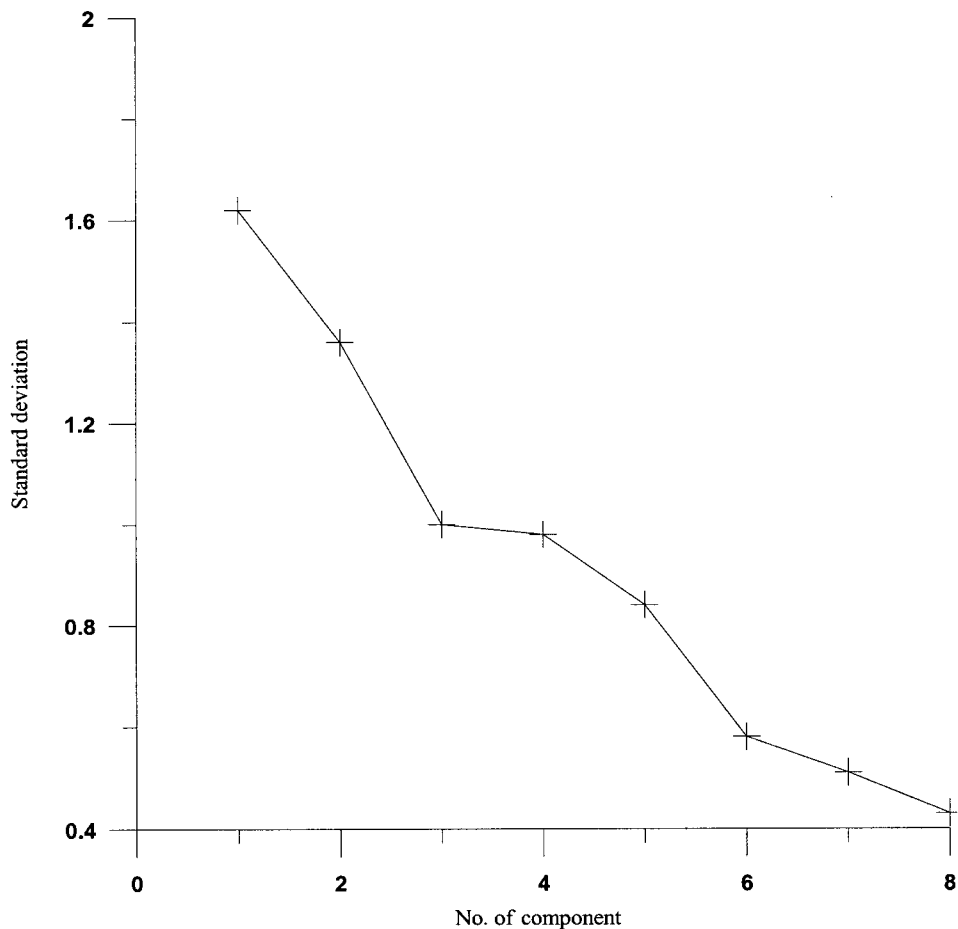


Figure 3. The screen plot of PCA for the studied data.

river should not be raised above 38 °C at any place or at any time in Taiwan. It is generally acknowledged that the value of pH from 6 to 9 is fine for the water quality, except when sharp deterioration arises. The overall subindex of toxic substances is derived by averaging the five scores of heavy metals but directly set to zero if any one of them gains zero score. Horton (1965) states that all of these rating values are arbitrary and, since they are, there will be differences of opinion as to specific choices. Swamee and Tyagi (2000) points out that the subindex function of a quality variable is not unique, but depends on the intended water use. Štambuk-giljanović (1999) indicates that the choices of the rated value of quality are sort of arbitrarily estimated, but with a base in the survey, and the application of this method to the evaluation of surface water quality, particularly with the intention of establishing an inventory of the quality of the water resources in a region or country, they can be modified according to their territorial circumstances elsewhere.

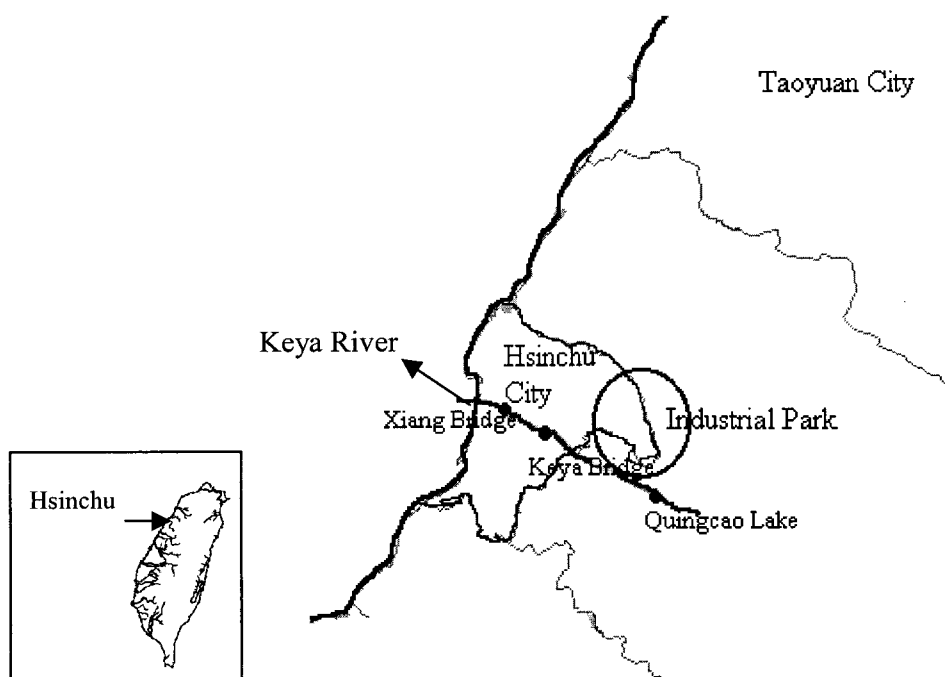


Figure 4. Geographical location and system environment of the Keya River.

A total of 2530 samples from 205 monitoring stations on 21 main rivers throughout Taiwan are employed, from 1994 to 2000. The summary of the employed variables is shown in Table IV. It is obvious that the value of pH is ranging narrowly, so does the measurement of temperature. The heavy metals are frequently under detection limits. The result of correlation matrix for the data set is shown in Table V, where some clear hydro-chemical relationships can be readily inferred. The values of temperature among the 2530 observations are uniform and thus the correlation with other variables could not be evaluated. High and positive correlation (underlined values) can be observed between DO, BOD<sub>5</sub>, ammonia nitrogen, SS and turbidity ( $r = 0.66$  to  $0.80$ ). The dissolved oxygen is strongly correlated with ammonia nitrogen ( $r = 0.73$ ) and BOD ( $r = 0.66$ ), which are closely related to organic contamination. Organic matter is partially oxidized by oxygen, while nutrients are responsible for the eutrophication of freshwater, thus causing a further increase in the concentration of organic matter and, hence, in oxygen demand and lowered water quality (Vega *et al.*, 1998). The variable of faecal coliforms shows slight correlation with DO, BOD<sub>5</sub> and ammonia nitrogen. The suspended solid and turbidity are significantly correlated ( $r = 0.80$ ), which are responsible for contamination of insoluble inorganic matter. It is plausible that the toxicity of heavy metals in Taiwan has not endangered to aquatics yet, because of weak correlation with the microbial related variables. In addition, non-related correlation between

TABLE VII  
The statistic summary of the studied data for Keya River

Parameters	Analytical detection limit	Station	Min	Max	Median	Mean	St. Dev. <sup>a</sup>	C.V. <sup>b</sup>
DO (mg L <sup>-1</sup> )	N.A. <sup>c</sup>	Qingcao Lake	1.7	10	6	5.89	1.76	0.30
		Keya Bridge	1.5	9.4	4.3	4.35	1.81	0.41
		Xiang Bridge	0	8.8	2.8	3.02	2.38	0.78
BOD (mg L <sup>-1</sup> )	N.A.	Qingcao Lake	1.3	35	6.6	8.27	6.23	0.75
		Keya Bridge	2.4	46	12	12.77	8.50	0.66
		Xiang Bridge	3.1	66	17	21.11	15.36	0.72
SS (mg L <sup>-1</sup> )	N.A.	Qingcao Lake	5.7	143	35	44.99	35.53	0.78
		Keya Bridge	5.1	189	34	43.46	33.07	0.76
		Xiang Bridge	3.9	217	34	42.81	41.42	0.96
NH <sub>3</sub> -N (mg L <sup>-1</sup> )	N.A.	Qingcao Lake	0.05	9.3	1.7	2.30	2.17	0.94
		Keya Bridge	0.12	41	7.5	11.18	10.22	0.91
		Xiang Bridge	0.09	46.9	10	14.38	12.34	0.85
pH	N.A.	Qingcao Lake	6.3	9	7.7	7.65	0.49	0.06
		Keya Bridge	6.9	8.7	7.5	7.47	0.30	0.04
		Xiang Bridge	6.7	9	7.5	7.49	0.36	0.04
Faecal coliform (CFU 100 mL)	N.A.	Qingcao Lake	160	6300	1615	2422	2903	1.19
		Keya Bridge	160	230000	220000	685040	1081739	1.57
		Xiang Bridge	140	270000	220000	177535	120865	0.68
Temperature (°C)	N.A.	Qingcao Lake	15	36	25	25.24	4.96	0.19
		Keya Bridge	14	34	25.5	25.01	4.66	0.18
		Xiang Bridge	13	32	25.2	24.80	4.71	0.19
Turbidity (NTU)	N.A.	Qingcao Lake	3	137	15	27.00	30.42	1.12
		Keya Bridge	4	194	19	27.35	37.58	1.37
		Xiang Bridge	8.8	206	17	27.59	33.07	1.19
Cd (mg L <sup>-1</sup> )	0.01	Qingcao Lake	0.01	0.02	0.01	0.01	0.001	0.14
		Keya Bridge	0.01	0.02	0.01	0.01	0.001	0.17
		Xiang Bridge	0.01	0.02	0.01	0.01	0.002	0.23
Pb (mg L <sup>-1</sup> )	0.05	Qingcao Lake	0.05	<b>0.06</b>	0.05	0.05	0.001	0.03
		Keya Bridge	0.05	<b>0.115</b>	0.05	0.05	0.01	0.27
		Xiang Bridge	0.05	<b>0.13</b>	0.05	0.05	0.01	0.32
Cr (mg L <sup>-1</sup> )	0.01	Qingcao Lake	0.01	0.01	0.01	0.01	0	0
		Keya Bridge	0.01	0.01	0.01	0.01	0	0
		Xiang Bridge	0.01	0.01	0.01	0.01	0	0
Cu (mg L <sup>-1</sup> )	0.01	Qingcao Lake	0.01	<b>0.06</b>	0.01	0.014	0.011	0.84
		Keya Bridge	0.01	<b>0.2</b>	0.02	0.03	0.03	1.12
		Xiang Bridge	0.01	<b>0.17</b>	0.01	0.02	0.02	1.28
Zn (mg L <sup>-1</sup> )	0.01	Qingcao Lake	0.01	<b>0.23</b>	0.01	0.02	0.04	1.93
		Keya Bridge	0.01	<b>0.4</b>	0.02	0.03	0.06	1.71
		Xiang Bridge	0.01	<b>0.15</b>	0.02	0.03	0.03	1.04

<sup>a</sup> Standard deviation.

<sup>b</sup> Coefficient of variation (St. Dev./Mean).

<sup>c</sup> N.A. = Not applicable.

toxicity with suspended solids and turbidity could be expected for the heavy metals are ionic existing in waters.

The screen plot of the principal components analysis is shown in Figure 3. Five PCs are extracted, together explaining more than 90% of the variance of information contained in the original data. Table VI summaries the PCA results including the loadings exceeded 0.1 and the eigenvalue of each PC. PC 1 explains 33.14% of the variance and is mainly contributed by ammonia nitrogen, BOD<sub>5</sub>, and DO. PC 1 is denoted as 'organics' component. PC 2 explains 23.41% of the variance and is

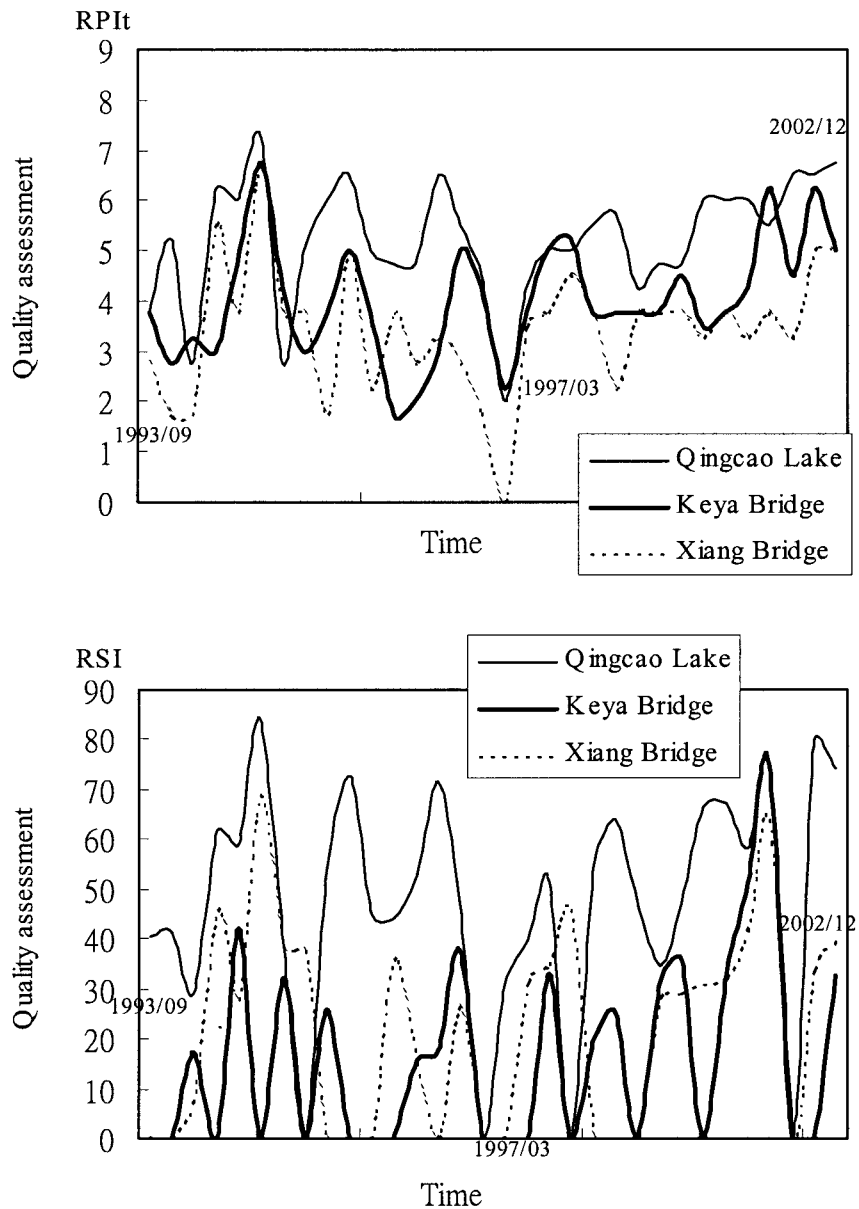


Figure 5. The quality assessment for Keya River based on the existing index (RPIt) and the new proposed index (RSI).

mainly participated by SS and turbidity. PC 2 is nominated as the ‘particulates’ component, which causes undesirable turbidity or muddiness in surface waters and objectionable bottom deposits. PC 3 is related to pH and PC 4 is for toxicity. PC 5 is contributed by faecal coliforms, which represents microorganisms. The results of

principal components are the inputs to the aggregation function for overall index. Finally, the overall River Status Index (RSI) is recommended as Equation (1).

$$RSI = C_{\text{tem}} C_{\text{pH}} C_{\text{tox}} \left[ \left( \sum_{i=1}^3 I_i w_i \right) \times \left( \sum_{j=1}^2 I_j w_j \right) \times \left( \sum_{k=1}^1 I_k \right) \right]^{1/3}, \quad (1)$$

where, the unweighted arithmetic mean is assigned for the variables associated in the same components.  $I_i$  denotes the subindex values for the ‘organics’ and three variables are contained that are DO, BOD<sub>5</sub> and ammonia nitrogen ( $I_1$  for the subindex of DO,  $I_2$  for the subindex of BOD<sub>5</sub> and  $I_3$  for the subindex of NH<sub>3</sub>-N);  $I_j$  represents the subindex values for the ‘particulates’, which consist of suspended solids and turbidity;  $I_k$  is the measurement of faecal coliform, which represents the ‘microorganisms’. In addition, the geometric mean is appointed on the extracted components. Three scaling coefficients are prefixed, which address the subindices of temperature ( $C_{\text{tem}}$ ), pH ( $C_{\text{pH}}$ ) and toxicity substances ( $C_{\text{tox}}$ ), respectively. The multipliers are commended with regards to the characteristics of effects and the rating structures of the three variables. In the majority of situations the three variables would take on a value of 100, but water is considered injurious to human, animals and aquatic life and makes the recovery of water quality difficult as lower ratings are evaluated. In addition, any assigned weights to them would cause fewer weightings are available for the rest of the variables, since the sum of the weights is constraint on 1. The lowered sensitivity of the overall index is inevitable, when smaller weightings are assigned to the variables, having dramatic changes with different circumstances. Finally, the overall water quality index consists of the arithmetic mean, geometric mean and multiplied by three coefficients. From the proposed aggregation function, if any one of the components closes to zero, the overall index RSI approaches to zero.

In principle, water quality characteristics may be classified in three board categories: physical, chemical, and biological. Within each category, a number of quality variables may be employed as the result shows in this study. The concerns of eclipsing and ambiguity occurring from aggregation, or large numbers of variables, have been solved because those variables are categorized and appropriate mathematic functions are addressed.

### 3.1. A CASE STUDY ON KEYA RIVER

The proposed method for assessing river quality has been applied in the Keya River, which is located in the vicinity of Hsinchu city, in the north-central part of Taiwan. Since the founding of Hsinchu Science-based Industrial Park, Hsinchu has attracted many high technology factories along with the accompanying threat to the quality of the Keya River. The regular monitoring stations of Qingcao Lake, Keya Bridge and Xiang Bridge are located in upriver, middle stream and downstream, respectively. A map for geographical location and system environment of Keya

TABLE VIII  
The data for the scores of RSI dropping into 'zero' in the station of Keya-Bridge

Date	DO (mg L <sup>-1</sup> )	BOD (mg L <sup>-1</sup> )	Amm. N (mg L <sup>-1</sup> )	SS (mg L <sup>-1</sup> )	Tur. (NTU)	Temp.	pH	Cd (ppm)	Pb (ppm)	Cr (ppm)	Cu (ppm)	Zn (ppm)
09/1998	4.8	23	16	31	20	28.5	7.3	-0.01 <sup>a</sup>	-0.05	-0.03	0.06	-0.01
03/1998	5.5	12	18	52	15	20	7.5	0.01	-0.05	-0.03	0.03	0.14
09/1997	4.4	11	0.27	86	32	23	7.2	-0.01	-0.05	-0.03	0.06	0.01
05/1997	2.6	13	8	30	19	28.5	7.6	-0.01	-0.05	-0.03	0.03	-0.01
03/1997	1.5	16	16	8	38	22.5	7.8	-0.01	0.11	-0.03	0.03	-0.01
01/1997	2.6	13	32	17	19	20	7.8	0.01	0.09	-0.03	0.03	-0.01
05/1996	4.8	18	4.7	110	30	26	7.8	-0.01	0.08	-0.03	0.03	0.03
03/1996	2.3	12	25.2	29	20	22.5	7.6	-0.01	-0.05	-0.03	0.09	0.03
01/1996	4	18	0.1	25	14	20.5	6.9	-0.01	-0.05	-0.03	0.11	0.02
09/1995	3.3	8	38.6	64	22	30	7.2	-0.01	-0.05	-0.03	0.03	0.02
11/1994	5.2	9.6	0.1	34	22	21	7.5	-0.01	-0.05	-0.03	0.05	0.02
07/1994	4.3	6.9	7.4	51	24	31	7.5	-0.01	-0.05	-0.03	0.04	0.04
01/1994	2.2	38	5.2	18	17	18	7.4	-0.01	-0.05	-0.03	0.03	0.01
11/1993	1.8	13	9.6	30	16	25	7.6	-0.01	-0.05	-0.03	0.2	0.01
09/1993	2.1	37	1.7	36	20	26	7.3	-0.01	-0.05	-0.03	0.04	0.04

a ' - ' = The concentration is below the detection limit.

River is shown in Figure 4. Historical data are collected from the EPA. Table VII summarizes the statistics of the employed data. The proposed index RSI and the prevailing index RPI are converse for an observation since the RPI describes river pollution index but the RSI denotes river quality index. A saturated polluted measurement would produce 0 for RSI and 10 for RPI; on the contrary, an absolute excellent measurement would come out 100 for RSI and 0 for RPI. Thus for easy comparison between the two indices, the RPI<sub>t</sub> ranges from 0 to 9 according to the river quality from bad to good, which is obtained by subtracting 10 from the value of RPI. Hence, both of the RSI and RPI<sub>t</sub> are quality indices for rivers.

The results of RSI and RPI<sub>t</sub> for the three monitoring stations are drawn in Figure 5. It is obvious that the quality is markedly decreased at Keya Bridge and the values are mostly below 40 and even more drop to zero with RSI index. Intensive industrial activities are carried out in this area. Keya River is located the industrial zone and surrounded by numerous industries, including chemical materials manufacturing, chemical manufacturing, non-metal mineral manufacturing, basic metal refining and electroplating factories. After the river flows through Hsinchu Science-based Industrial Park and Hsinchu City, the heavily polluted water containing of organics and hazardous discharges into the Keya River. The impacts on the river quality are inevitably considerable although the water treatment facilities and wastewater controlling systems are asked among the factories. Table VIII shows the data leading the values of quality dropping into zero, in which high concentration of Cu, Zn and Pb are measured. The index RSI obviously reflects the entrance of industrial pollutants in the river but the existing index RPI does not explore the same information clearly. The quality of Keya Bridge shows serious grossly polluted with RSI but moderately polluted with RPI. In Taiwan, most of the oyster mariculture beds are located along the west of the island. It is believed that copper and zinc concentrations in oysters could be the cause of the concentrations in the river (Lee *et al.*, 1996). It is very important to establish an index, showing remarkable information of the river quality and furthermore provides the warning of taking pollution control actions.

#### 4. Conclusions

An effective gradation index for diagnosing a generalized river quality has been developed and illustrated with the case study of the Keya River in Taiwan. Firstly, environmental monitoring measurements are proposed with rating curves relating the various levels of environmental quality. Secondly, analyses of principal components serve as intermediate steps for integrating an overall index. By the explorations of the correlation matrix and principal components analysis, 'organics' component, 'particulates' component and 'microorganism' are categorized from the employed variables. Three coefficients for the variables of pH, temperature and toxic substances are recommended for their characteristics of data structures and



quality effects. Finally, a fresh look at the aggregation process is proposed. An overall index is proposed based on the behavior and limitations of the aggregation functions. The arithmetic mean is employed in variables associated in the same component. In general, as more parameters are added, the distortion after aggregation tends to increase. Since the variables are categorized in advance, the problem of eclipsing and ambiguity caused by a large number of variables in aggregation could be delimited.

Because of the changes of human activities, the resources of contaminants in rivers are different than two or three decades before. The existing RPI index might be suitable in times of agriculture societies, but not good enough after various human activities are carried out. Compare to the existing index in Taiwan, the variables are comprehensive reviewed in the new proposed index, which contains the contaminants caused by farming, industrial and municipal discharges. The aggregation function is mathematically considered. The proposed measurement RSI, applied in the Keya River shows that the quality, affected by industrial activities, causing water quality problems in Taiwan, does not significant represent the existing index. Furthermore, the proposed index is established for evaluating river quality around the island in general.

## References

- Alberto, W. D., Pilar, D. M. D., Valeria, A. M., Fabiana, P. S., Cecilia, H. A. and Maria, B.: 2001, 'Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality – A case study: Suquia River basin (Cordoba, Argentina)', *Water Res.* **35**(12), 2881–2894.
- Ball, R. O. and Church, R. L.: 1980, 'Water quality indexing and scoring', *J. Environ. Engin. Div. ASCE* **106**(4), 757–771.
- Brown, R. M., McClelland, N. I., Deiningner, R. A. and Tozer, R. G.: 1970, *A Water Quality Index – Do we Dare?*, Water Sewage Works, October, pp. 339–343.
- Ceballos, B. S. O., König, A. and Oliveira, J. F.: 1998, 'Dam reservoir eutrophication: A simplified technique for a fast diagnosis of environmental degradation', *Water Res.* **32**(11), 3477–3483.
- Dee, N., Baker, J., Drobny, N., Duke, K., Whitman, I. and Fahringer, D.: 1973, 'An environmental evaluation system for water resource planning', *Water Resour. Res.* **9**(3), 523–535.
- Dinius, S. H.: 1987, 'Design of an index of water quality', *Water Res. Bull.* **23**(5), 833–843.
- Dojlido, J., Raniszewski, J. and Woyciechowska, J.: 1994, 'Water quality index-application for river in Vistula River Basin in Poland', *Water Sci. Technol.* **30**(10), 57–64.
- Gómez, N.: 1999, 'Epipellic diatoms from the Matanza-Riachuelo river (Argentina), a highly polluted basin from the pampean plain: Biotic indices and multivariate analysis', *Aqua. Ecosys. Health Manage.* **2**, 301–309.
- Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J. M. and Fernandez, L.: 2000, 'Temporal evaluation of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis', *Water Res.* **34**(3), 807–816.
- Horton, R. K.: 1965, 'An index number system for rating water quality', *J. Wat. Pollut. Con. Fed.* **37**(3), 300–305.
- House, M. A. and Ellis, J. B.: 1987, 'The development of water quality indices for operational management', *Water Sci. Technol.* **19**(9), 145–154.

- House, M. A. and Newsome, D. H.: 1989, 'Water quality indices for the management of surface water quality', *Water Sci. Technol.* **21**, 1137–1148.
- Johnson, R. A. and Wichern, D. W.: 1998, *Applied Multivariate Statistical Analysis*, Prentice-Hall Inc.
- Jonnalagadda, S. B. and Mhere, G.: 2001, 'Water quality of the Odzi River in the eastern highlands of Zimbabwe', *Water Res.* **35**(10), 2371–2376.
- Landwehr, J. M.: 1979, 'A statistic view of a class of water quality indices', *Water Resour. Res.* **15**(2), 460–468.
- Lee, C. L., Yi, H. and Chuang, M. Y.: 1996, 'Use of oyster, *Crassostrea gigas*, and ambient water to assess metal pollution status of the charting coastal area, Taiwan, after the 1986 green oyster incident', *Chemosphere* **33**(12), 2505–2532.
- Liou, S. M., Lo, S. L. and Hu, C. Y.: 2003, Application of two-stage fuzzy set theory to river quality evaluation in Taiwan, *Water Res.* **37**, 1406–1416.
- Morales, M. M., Martin, P., Llopis, A. Campos, L. and Sagrado, S.: 1999, An environmental study by factor analysis of surface seawater in the Gulf of Valencia (Western Mediterranean), *Anal. Chim. Acta* **394**, 109–117.
- Ott, W. R.: 1978, *Environmental Indices: Theory and Practice*, Ann Arbor Science Publishers.
- Pesce, S. F. and Wunderlin, D. A.: 2000, 'Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquia River', *Water Res.* **34**(11), 2915–2926.
- Pesce, S. F. and Wunderlin, D. A.: 2000, 'Use of water quality indices to verify the impact of Córdoba City (Argentina) on Suquia River', *Water Res.* **34**(11), 2915–2926.
- Reisenhofer, E., Adami, G. and Barbieri, P.: 1998, 'Using chemical and physical parameters to define the quality of Karstic freshwaters (Timavo River, north-eastern Italy): A chemometric approach', *Water Res.* **32**(4), 1193–1203.
- Ross, S. L., L.R.I.C., M.Sc. and M.I.W.E.S.: 1977, 'An index system for classifying river water quality', *Water Poll. Con.* **7**(1), 113–122.
- Schierow, L. J. and Chesters, G.: 1988, 'Evaluation of the Great Lakes Nearshore index', *Water Res.* **22**(3), 269–277.
- Smith, D. G.: 1990, 'A better water quality indexing system for rivers and streams', *Water Res.* **24**(10), 1237–1244.
- Štambuk-Giljanović, N.: 1999, 'Water quality evaluation by index in Dalmatia', *Water Res.* **33**(16), 3423–3440.
- Swamee, P. K. and Tyagi, A.: 2000, 'Describing water quality with aggregate index', *J. Environ. Engin.* **126**(5), 451–455.
- Vega, M., Pardo, R., Barrado, E. and Deban, L.: 1998, 'Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis', *Water Res.* **32**(12), 3581–3592.