

PAPER • OPEN ACCESS

## Research on comprehensive evaluation of drinking water source based on two principal component analysis for water body classification and time-space comparability

To cite this article: Yongchen Zong *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **227** 062025

View the [article online](#) for updates and enhancements.

# Research on comprehensive evaluation of drinking water source based on two principal component analysis for water body classification and time-space comparability

Yongchen Zong<sup>1,2,3,4,5,6</sup>, Decai Huang<sup>6</sup>, Yongheng Zhang<sup>6</sup> and Guanghua Lu<sup>1,2,3,4,5,6,7</sup>

<sup>1</sup> Res. Institute of Tibet Plateau Ecology, Tibet Agriculture & Animal Husbandry University, Linzhi 860000, China;

<sup>2</sup> Tibet Key Laboratory of Forest Ecology in Plateau Area, Ministry of Education, Linzhi 860000, China;

<sup>3</sup> National Key Station of Field Scientific Observation & Experiment, Linzhi 860000, China;

<sup>4</sup> Key Laboratory of Forest Ecology in Plateau Area, Tibet Autonomous Region, Linzhi 860000, China;

<sup>5</sup> United Key Laboratories of Ecological Security, Tibet Autonomous Region, Linzhi 860000, China;

<sup>6</sup> Water Conservancy Project & Civil Engineering College, Tibet Agriculture & Animal Husbandry University, Linzhi 860000, China.

<sup>7</sup> Email: ghlu@hhu.edu.cn

**Abstract.** Based on the monitoring data of 15 indicators of 7 typical drinking water sources in Niyang River, the principal component analysis was carried out by the standard values of the corresponding surface water environmental quality standards, and first principal component analysis and second principal component analysis were constructed successively with the standardized mean value, standard deviation and the cumulative variance, eigenvector and component matrix of the principal component analysis. The models of two principal component analysis for the comprehensive evaluation of drinking water sources were constructed on the basis of a principal component analysis and two principal component analysis. The results showed that the water quality is good in the drinking water sources in Niyang River, and all the drinking water sources belong to the class II water bodies through the status evaluation, eutrophication evaluation, health risk evaluation and comprehensive evaluation. The comprehensive evaluation results based on the two principal component analysis method were completely consistent with those obtained by the single factor evaluation method, indicating that the two principal component analysis method is feasible for comprehensive evaluation of water quality. In addition, the consistency of the principal component analysis matrices ensured the comparability of the temporal and spatial variation in water quality.

## 1. Introduction

Principal component analysis, also known as principal component analysis, aims to use the idea of dimensionality reduction to transform multiple indicators into a few comprehensive indicators (i.e.



principal components), in which each principal component can reflect most of the information of the original variables, and the information contained in each other does not repeat. This method can simplify the problem and obtain more scientific and effective data information by introducing multiple variables and attributing complex factors to several principal components. It has been widely used in many different disciplines and fields, with the promotion of its application in the corresponding fields of defects are constantly found. For example, in the field of water environment assessment, the evaluation results cannot achieve the function of water grading, and the evaluation results of different time and space are not comparable. In the field of water environment assessment, the evaluation results can not realize the function of water classification, and the evaluation results of different time and space are not comparable. In this paper, the comprehensive evaluation model of drinking water source is studied aiming at the function of water body classification, space-time comparability and complex two evaluations.

## 2. Data and methods

Located in Linzhi, Tibet, among the first tributaries of the Yarlung Zangbo River, Niyang River is bordered by the Mila Mountain as the watershed in the west, the Yarlung Zangbo River and the Sejila Mountain to the east, and the eastern Himalayas as the watershed in the south, while bounded by the eastern Nyenchen Tanglha to the north[1], with 17815 km<sup>2</sup> basin area, 23.072 billion m<sup>3</sup> annual precipitation, and 17.229 billion m<sup>3</sup> annual runoff. It covers 16 relatively backward organic villages and towns in 2 counties (districts), leading to slowly-developing supporting infrastructure. The surface water is the main drinking water source in the most organic villages and towns, typically including [2] Songduo Township, Jinda Town, Kongpo Gyamda, Zhongsha Township, Bahe Town, Changzhang Township, Bayi Town (the seat of Linzhi Government). In recent years, with the further economic development, the quality of surface drinking water has suffered a certain deterioration, followed by the comprehensive evaluation of drinking water sources becoming an important task in water source selection.

### 2.1. Spot location of water quality monitoring

**Table 1.** Typical township list.

Section name	Water source	Water source locations	Water source elevation (m)
Songduo Township	Natural water	N29°53'28.44" W92°28'27.64"	4378
Jinda Town	Natural water	N30°01'10.04" W92°54'16.23"	3888
Kongpo Gyamda	Natural water + groundwater	N29°52'00.64" W93°14'56.21"	3601
Zhongsha Township	Natural water	N29°52'26.66" W94°20'14.78"	3404
Bahe Town	Natural water	N29°51'57.79" W94°39'10.32"	3605
Changzhang Township	Natural water	N29°44'24.09" W94°02'42.85"	3098
Bayi Town	Natural water + groundwater	N29°48'04.09" W94°25'18.12"	4103

According to the status of the drinking water source, population and water intake, a total of 7 monitoring points were set up in the study (see Table 1 for details) to monthly monitor 15 [3] index values including TN, TP,  $\text{NH}_4^+\text{-N}$ , COD, DO, SD, chl<sub>a</sub>,  $\text{Cr}^{6+}$ , Cd, As, Pb, Hg, volatile phenol, cyanide and fluoride from June to August (wet season) and from November to December (dry season) in 2017, and the mean values of each status parameter in the wet season or the dry season were used in this study, and the mean values of each status parameter in the wet season or the dry season were used in this study. The monitored items were analyzed based on GB 3838-2002 Environmental Quality Standard for Surface Water and Water and Exhausted Water Monitoring Analysis Method [4].

## 2.2. Evaluation methods

In accordance with GB 3838-2002 *Environmental Quality Standard for Surface Water*, the pollution degree was divided into I, II, III, IV, V and inferior V in the evaluation for the status of five water quality indicators comprising TN, TP,  $\text{NH}_4^+\text{-N}$ , COD and DO [5], for the eutrophication of TN, TP, SD, COD and SD [6,7], for the health risk of  $\text{Cr}^{6+}$ , Cd, As, Pb, Hg, volatile phenol, CN and F [8,9] and for comprehensive studies on water quality status, eutrophication indicators, health risks indicators [10,11].

## 3. Research methods

### 3.1. Principal component analysis

It was proposed to carry out comprehensive evaluation research with two principal component analysis[12] which is the second principal component analysis based on the principal component analysis and an extension of the principal component analysis[13]. In the study, the principal component analysis was adopted to perform the status evaluation, the eutrophication evaluation, and the health risk evaluation, whose results were used into the quadratic component analysis to make the corresponding comprehensive evaluation.

Principal component analysis [14] lies in statistical analysis that converts multiple indicators into a few unrelated comprehensive indicators, which demonstrate most of the information provided by the source indicators, thus leading to dimensionality reduction and source identification, and enjoying 85% or more cumulative variance contribution rate [15] since the eigenvalue is not more than 1.

In order to ensure the uniformity and comparability of the data processing of the evaluation object, the study was to conduct standardization and principal component analysis for the standard value for basic items in environmental quality standards for surface water, and with the information obtained (standardized mean, standard deviation, and cumulative variance, eigenvectors, and component matrices of the principal component analysis), to evaluate and calculate the monitoring data of each section for relevant evaluation.

### 3.2. Raw Data

Refer to Table 2 and Table 3 for the original water quality indicators of drinking water sources in 7 typical towns in wet and dry seasons, while the standard values for 15 water quality indicators in environmental quality standards for surface water are shown in Table 4.

The status evaluation was carried out for the above-mentioned TN, TP,  $\text{NH}_4^+\text{-N}$ , COD, and DO, the eutrophication evaluation for TN, TP, SD, COD and SD, and the health risk evaluation for  $\text{Cr}^{6+}$ , Cd, As, Pb, Hg, volatile phenol, CN and F. According to the status parameters during the wet and dry seasons, TN,  $\text{NH}_4^+\text{-N}$ , SD, Pb, Hg, and volatile phenol of each section (except Kongpo Gyamda in the wet season) were among Class II water bodies while others fell into Class I, that is, the single-index evaluation corresponded to Class II water bodies, whose corresponding pollutants were the above water quality indicators.

**Table 2.** Status parameter values in the wet season.

Indicator(mg L <sup>-1</sup> )	Songduo Township	Jinda Town	Kongpo Gyamda	Zhongsha Township	Bahe Town	Gengzhang Township	Bayi Town
DO	8.8	8.3	8.4	8.3	8.6	8.4	8.4
COD	1.7	1.6	1.5	1.7	1.8	1.8	1.9
TP	0.003	0.003	0.004	0.003	0.004	0.005	0.006
TN	0.31	0.31	0.18	0.25	0.28	0.35	0.41
NH <sub>4</sub> <sup>+</sup> -N	0.22	0.21	0.16	0.2	0.21	0.23	0.25
Chla	5.2	5.5	5.1	5.4	5.6	5.5	5.6
SD(m)	2.6	2.3	2.2	2.3	1.8	1.9	2
Cr <sup>6+</sup>	0.006	0.006	0.007	0.005	0.008	0.009	0.008
Cd	0.002	0.002	0.003	0.003	0.004	0.005	0.003
As	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Pb	0.019	0.017	0.016	0.012	0.013	0.014	0.015
Hg	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
CN	0.002	0.002	0.002	0.002	0.002	0.002	0.002
F <sup>-</sup>	0.03	0.03	0.03	0.03	0.03	0.03	0.03
Volatile phenol	0.002	0.002	0.002	0.002	0.002	0.002	0.002

**Table 3.** Status parameter values in the dry season.

Indicator(mg L <sup>-1</sup> )	Songduo Township	Jinda Town	Kongpo Gyamda	Zhongsha Township	Bahe Town	Gengzhang Township	Bayi Town
DO	8.5	8.5	8.3	8.3	8.2	8.1	7.9
COD	1.5	1.5	1.4	1.5	1.6	1.2	1
TP	0.007	0.006	0.009	0.007	0.011	0.008	0.012
TN	0.34	0.35	0.36	0.34	0.32	0.37	0.43
NH <sub>4</sub> <sup>+</sup> -N	0.21	0.22	0.3	0.25	0.26	0.26	0.27
Chla	5.1	5.3	5.2	5.6	5.8	5.6	5.7
SD(m)	2.5	2.3	2	2.3	1.9	1.8	1.7
Cr <sup>6+</sup>	0.005	0.006	0.005	0.005	0.006	0.007	0.006
Cd	0.002	0.002	0.001	0.001	0.001	0.001	0.002
As	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Pb	0.018	0.017	0.008	0.009	0.009	0.009	0.007
Hg	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
CN	0.002	0.002	0.002	0.002	0.002	0.002	0.002
F <sup>-</sup>	0.02	0.02	0.02	0.03	0.02	0.02	0.02
Volatile phenol	0.002	0.002	0.002	0.002	0.002	0.002	0.002

**Table 4.** Standard value for basic items in environmental quality standards for surface water.

Indicator(mg L <sup>-1</sup> )	I	II	III	IV	V
DO	7.5	6	5	3	2
COD	15	15	20	30	40
TP	0.02	0.1	0.2	0.3	0.4
TN	0.2	0.5	1	1.5	2
NH <sub>4</sub> <sup>+</sup> -N	0.15	0.5	1	1.5	2
Chla	1.58	9.99	25.1	63.03	63.03
SD(m)	2.979	1.063	0.635	0.379	0.379
Cr <sup>6+</sup>	0.01	0.05	0.05	0.05	0.1
Cd	0.001	0.005	0.005	0.005	0.01
As	0.05	0.05	0.05	0.1	0.1
Pb	0.01	0.01	0.05	0.05	0.1
Hg	0.00005	0.00005	0.0001	0.001	0.001
CN	0.005	0.05	0.2	0.2	0.2
F <sup>-</sup>	1.0	1.0	1.0	1.5	1.5
Volatile phenol	0.002	0.002	0.005	0.01	0.10

#### 4. Results and discussion

##### 4.1. First principal component analysis

According to the relevant data in Table 4, the standardization on the standard limit values of routine status parameters, eutrophication parameters and health risk parameters were carried out, and the results were presented in Tables 5-7.

**Table 5.** Standardization on the standard limit values of routine status parameters.

Category	DO	COD	TP	TN	NH <sub>4</sub> <sup>+</sup> -N
I	1.25851	-0.83028	-1.21116	-1.15058	-1.18177
II	0.58431	-0.83028	-0.68457	-0.73966	-0.71175
III	0.13484	-0.36901	-0.02633	-0.05479	-0.04029
IV	-0.76409	0.55352	0.63191	0.63008	0.63117
V	-1.21356	1.47605	1.29014	1.31495	1.30263

**Table 6.** Standardization on the standard limit values of eutrophication parameters.

Category	TN	TP	Chla	COD	SD
I	-1.15058	-0.87264	-1.065	-1.0431	1.72948
II	-0.73966	-0.67727	-0.77576	-0.65677	-0.02194
III	-0.05479	-0.35166	-0.25609	-0.27043	-0.41317
IV	0.63008	0.29956	1.04842	0.50223	-0.64718
V	1.31495	1.602	1.04842	1.46807	-0.64718

**Table 7.** Standardization on the standard limit values of health risk parameters.

Category	Cr <sup>6+</sup>	Cd	As	Pb	Hg	CN	F <sup>-</sup>	Volatile phenol
I	-1.31507	-1.31507	-0.7303	-0.91525	-0.76229	-1.31507	-0.7303	-0.51027
II	-0.06262	-0.06262	-0.7303	-0.91525	-0.76229	-0.8454	-0.7303	-0.51027
III	-0.06262	-0.06262	-0.7303	0.16151	-0.66456	0.72016	-0.7303	-0.44005
IV	-0.06262	-0.06262	1.09545	0.16151	1.09457	0.72016	1.09545	-0.32302
V	1.50294	1.50294	1.09545	1.50747	1.09457	0.72016	1.09545	1.78361

The evaluation criteria for five types of water body were constructed for routine status parameters, eutrophication parameters and health risk parameters values, respectively, and the results were shown in Table 8.

**Table 8.** Standard value of limit value of basic project standard for surface water environmental quality standard.

Category	Standard value of Status	Standard value of Eutrophication	Standard value of Health risk
I	$\leq -2.5215$	$\leq -2.7412$	$\leq -2.1658$
II	$\leq -1.5869$ & $> -2.5215$	$\leq -1.3003$ & $> 2.7412$	$\leq -1.2202$ & $> -2.1658$
III	$\leq -0.2771$ & $> -1.5869$	$\leq -0.1108$ & $> -1.3003$	$\leq -0.4435$ & $> -1.2202$
IV	$\leq 1.4365$ & $> -0.2771$	$\leq 1.5455$ & $> -0.1108$	$\leq 0.8579$ & $> 0.4435$
V	$\leq 2.9490$ & $> 1.4365$	$\leq 2.6068$ & $> 1.5455$	$\leq 2.9716$ & $> 0.8579$

The data normalization and initial eigenvalue conversion were completed on the original monitoring data listed in Tables 2 and 3, then the cumulative variance contribution rates were calculated, and the evaluation values at different sections in different seasons were obtained finally (Table 9).

**Table 9.** Evaluation value of each section.

Section name	Wet season			Dry season		
	Status evaluation	Eutrophication evaluation	Health risk evaluation	Status evaluation	Eutrophication evaluation	Health risk evaluation
Songduo Township	-1.7713	-2.3258	-1.4996	-1.8478	-2.5550	-1.4636
Jinda Town	-1.7621	-2.2506	-1.4978	-1.7574	-2.4496	-1.4808
Kongpo Gyamda	-1.6626	-2.1301	-2.0072	-1.8890	-2.5074	-1.3754
Zhongsha Township	-1.7070	-2.2465	-1.9901	-1.7964	-2.4804	-1.4306
Bahe Town	-1.6773	-2.0811	-1.9882	-1.8251	-2.2638	-1.2872
Gengzhang Township	-1.6514	-2.0699	-1.9778	-1.7268	-2.2544	-1.2646
Bayi Town	-1.5945	-1.9885	-1.5839	-1.6707	-2.2395	-1.3737

Form Table 8 and Table 9, it was found that 1) in addition to the status evaluation, the eutrophication evaluation and health risk evaluation of all sections belonged to Class II water bodies; 2) in terms of time, the status evaluation, eutrophication evaluation and health risk evaluation in dry season overmatched those in the wet season; and 3) for space, the upstream monitoring sections were exposed to generally better status evaluation, eutrophication evaluation and health risk evaluation compared with the downstream sections.

#### 4.2. Second principal component analysis

The comprehensive evaluation standard values in environmental quality standards for surface water (-2.0262, -1.1219, -0.2328, 1.0426, 2.3383) followed second principal component analysis and calculation made based on the data in Table 5, as shown in Table 11.

The second standardization for the routine status parameters, eutrophication parameters and health risk parameters were completed based on the data in Tables 8, and the results were presented in Table 10.

**Table 10.** The second standardization values for the routine status parameters, eutrophication parameters and health risk parameters.

Category	standardization values of Status	standardization values of Eutrophication	standardization values of Health risk
I	-1.13582	-1.21798	-1.08483
II	-.71484	-.61514	-.61121
III	-.12480	-.12507	-.22216
IV	.64707	.65879	.42973
V	1.32840	1.29941	1.48847

The data normalization and initial eigenvalue conversion were completed on the evaluation values at different sections in different seasons listed in Table 9, then the cumulative variance contribution rates were calculated, and the comprehensive evaluation values of water quality were obtained for different sections and different seasons (Table 11).

**Table 11.** Comprehensive evaluation value of cross section.

Section name	Comprehensive evaluation value in wet season	Comprehensive evaluation value in dry season
Songduo Township	-1.5275	-1.5993
Jinda Town	-1.5042	-1.5520
Kongpo Gyamda	-1.5926	-1.5717
Zhongsha Township	-1.6308	-1.5561
Bahe Town	-1.5776	-1.4634
Gengzhang Township	-1.5648	-1.3997
Bayi Town	-1.4063	-1.4414

It can be seen from Table 11 that all the drinking water sources in both wet and dry seasons belong to the class II water bodies based on the comprehensive evaluation results by using the second principal component analysis method. The overall water quality is good in Niyang River and can meet the requirements of drinking water sources. In addition, the correlation between different standardization values of is significant, and the correlation coefficients are all large than 0.99.



In summary, two principal component analysis led the comprehensive evaluation values of all sections in the wet and dry seasons to Class II water bodies, indicating the fine overall water quality of each section to be used as a drinking water source, which lay in the consistency with the results of the single factor status evaluation [16], and shared similarity with literature research achievements [2]. As for time span, in the comprehensive evaluation conclusion [17], four sections performed better in the wet season than in the dry season, three sections in the wet season better than in the dry season, while in the spatial span, the overall comprehensive evaluation of upstream sections is superior to that of downstream section [18].

## 5. Conclusions

The water quality status, eutrophication and health risks were evaluated for typical drinking water sources in Niyang River by using the principal component analysis method in the present study. On this basis, the comprehensive status evaluation were completed by using the second principal component analysis, and the following conclusions were obtained:

1) By taking the standard limit values of the surface water environmental quality standard as the principal component analysis matrix, we constructed the assessment models of the status, eutrophication and health risk for the drinking water sources, and their scores were calculated at different sections in the wet and dry seasons by using the standardized matrix. The consistency of the principal component analysis matrices ensured the temporal and spatial comparability of the obtained evaluation results.

2) The overall water quality is good in the drinking water sources in Niyang River based on the two principal component analysis method. All the drinking water sources belong to the class II water bodies through the status evaluation, eutrophication evaluation, health risk evaluation and comprehensive evaluation. On the whole, the water quality of the upstream section is better than that of the downstream section, and it's better in the wet season than in the dry season.

3) It is confirmed that the two principal component analysis method is feasible for comprehensive evaluation of drinking water sources. The comprehensive evaluation results based on the two principal component analysis method are completely consistent with those obtained by the single factor evaluation method.

## Acknowledgment

This study was supported by grants from National Natural Science Foundation of Tibet (XZ 2018 ZR G-20), National Natural Science Foundation of China (51868069, 51769034), the Program for Scientific Research Innovation Team in Colleges and Universities of Tibet Autonomous Region, Postgraduate Innovation Project of XIZANG Agriculture and Animal Husbandry College(STX2018-04).

## References

- [1] LI Y Q, YANG Y H, ZONG Y C, LI L J 2011 Analysis of Water Environment Quality Status and Protection Countermeasures in Niyang River Basin China Population, Resources and Environment **21** 417-420
- [2] ZONG Y C 2017 Comprehensive Evaluation on drinking water source in Niyang River basin in Tibet Tibet Science and Technology **8** 22-25
- [3] CHANG C P, ZHENG N, ZHENG H P 2009 Groundstatus evaluation and analyses in Hanzhong City Journal of Shaanxi University of Technology **25** 78-81
- [4] 2003 Ministry of Environmental Protection of the People's Republic of China Water and Wastewater Monitoring Analysis Method China Environmental Science
- [5] CHEN L H, YU J X, LI L, LV W Q, ZHANG J L 2017 Water Quality Assessment of Caohai Lake in Guizhou based on TLI and PCA Journal of Yangtze River Scientific Research Institute
- [6] Gladstone-Gallagher R V, Hughes R W, Douglas E J, Pilditch CA 2018 Biomass-dependent

- seagrass resilience to sediment eutrophication *Journal of Experimental Marine Biology & Ecology* **501** 54–64
- [7] ZHU G, YUAN C, DI G, ZHANG M, NI L, CAO T, FANG R, Wu G 2018 Morphological and biomechanical response to eutrophication and hydrodynamic stresses *Science of the Total Environment* **622–623** 421–435
- [8] JIANG L, ZHONG M, JIA X, XIA T 2018 Advance in Health Risk Assessment Methodology of Brownfield Sites in China
- [9] Kadi M W, Ali N, Albar H 2018 Phthalates and polycyclic aromatic hydrocarbons (PAHs) in the indoor settled carpet dust of mosques, health risk assessment for public *Science of the Total Environment* **627** 134–140
- [10] Cicalese M P, Ferrua F, Castagnaro L, Rolfe R, De Boever E, Reinhardt RR, Appleby J, Roncarolo MG, Aiuti A 2018 Gene Therapy for Adenosine Deaminase Deficiency: A Comprehensive Evaluation of Short- and Medium-Term Safety *Molecular Therapy*
- [11] ZHANG P, FENG G 2018 Application of fuzzy comprehensive evaluation to evaluate the effect of water flooding development *Journal of Petroleum Exploration & Production Technology* **2** 1–9
- [12] LOU Z, SHEN D, WANG Y 2018 Two - step principal component analysis for dynamic processes monitoring *Canadian Journal of Chemical Engineering* **96** 1
- [13] Cserhádi T, Illés Z 1991 Comparison of two principal component analysis methods to evaluate reversed-phase retention data *Journal of Pharmaceutical & Biomedical Analysis* **9** 685–691
- [14] Comon P 2015 Independent component analysis, a new concept? Elsevier North-Holland, Inc
- [15] Moore B 1981 Principal component analysis in linear systems: Controllability, observability, and model reduction *Automatic Control IEEE Transactions on* **26** 17–32
- [16] XIA F, HU S, GONG Z J, ZUO H H 2017 Comparative study on application of different status evaluation methods: A case study of inflow rivers in Danjiangkou reservoir area *Yangtze River* **17** 11–15
- [17] GAO D D, ZHAO L Y, L I C, CHENG C 2016 Evaluation of Water Quality in the Yuncheng Section of Fenhe River Based on Principal Component Analysis *Journal of Anhui Agricultural Sciences* **44** 69–71
- [18] CHEN Y 2012 Research of Status evaluation of the Sanhe Rivers in the Central Area of Chengdu based on BP-ANN Southwest Jiaotong University