CrossMark

Assessing Water Poverty in China Using Holistic and Dynamic Principal Component Analysis

Ane Pan[1] · Darrell Bosch[2] · Huimin Ma[1]

Abstract   The Water Poverty Index (WPI) expands the analysis of China's water crises from hydrology to a broader focus on integrated water resources management including economic and social factors. This index was revised by principal component analysis (PCA) to avoid arbitrariness of weights and collinearity between variables. However, the traditional PCA is primarily oriented for static data, and it fails to reveal the evolutionary trend of data over time. Moreover, the conventional normalization methods are not adequate when the dimension of time is added to the data. In this study, the transformation of centralized logarithm of initial variable and holistic and dynamic principal component analysis are firstly proposed, then the improved methods are applied to assess water poverty in China using panel data from 2004 to 2012. The estimated WPI shows the growing scale and the clustering trend of regional water poverty. The analysis of influential factors reveals that aquatic environmental pollution is a vital driver of water poverty. Water resource endowment is the second important factor concerning regional water poverty. Inability to adapt to water scarcity, which leads to weak physical water access and low efficiency of water use, is still a critical driver of regional water poverty. Finally, the regional disparities and alleviation strategies of water poverty are discussed.

✉ Ane Pan
panane@126.com

Darrell Bosch
bosch@vt.edu

1   School of Management, Wuhan University of Technology, 205 Luoshi Road, Hongshan District, Wuhan 430070, China

2   Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA 24061, USA

Springer

1 Introduction

Water is an indispensable and fundamental natural resource for human welfare and social development (Gleick 1996), especially in China, which has more than twenty percent of the world's population but less than seven percent of its freshwater resources (Todd 2010). Water scarcity is a question of means rather than resource availability (Sullivan 2002; Rijsberman 2006), and so the social environmental and economic factors need be integrated into the water resource management model (Jemmali and Matoussi 2013). The theory of water poverty expands the analysis of water scarcity from the field of hydrology to integrated water resources management (Feitelson and Chenoweth 2002; Rijsberman 2006). However, the calculation method of WPI has some shortcomings when it was applied in the Chinese case. Thus the aims of this paper are to revise the method of the calculation of Water Poverty Index (WPI) and then use it to analyze the underlying complexities of water issues in China.

Water poverty is defined as not having sufficient water to cover basic needs. Water poverty might be caused by water unavailability, income poverty or other reasons for inadequate access (Lawrence et al. 2002; Komnenic et al. 2009). The measurement of water poverty enables water resource policymakers and managers to track progress and evaluate the effectiveness of actions. The Water Poverty Index (WPI) was created by Sullivan (2002) to reconcile measures of water availability with measures of people's capacity to access water. There are four methods for the calculation of WPI: conventional composite index approach, gap approach, matrix approach and the simple time-analysis approach (Sullivan 2002). The first one is the focus of this paper because it is based on identifying the multi-dimensional nature of water poverty. The underlying framework was developed by Lawrence et al. (2002) covering five key components: water availability, access to water service, capacity for sustaining access, use of water, and the environmental factors influencing water quality and the ecology. Each of the components is made up of a number of sub-components identified to capture a wide range of water problems. They are firstly calculated by min–max normalization of initial variables and a weighted arithmetic average method. Then the components are combined using the same weighted average method in order to obtain the final value of WPI (Sullivan 2002; Lawrence et al. 2002).

The composite WPI method has been successfully applied through its final value or in the form of its components as a monitoring tool to express the water situation at national scale (Lawrence et al. 2002; Komnenic et al. 2009), regional scale (Heidecke 2006), local scale (Sullivan et al. 2003; Sullivan et al. 2007; Giné Garriga and Pérez Foguet 2011) and basin scale (Pérez Foguet and Giné Garriga 2011). However, the index has been criticized on several weaknesses involving the quality of data, arbitrariness of weights, high correlations between dimensions and variables, and loss of information in the aggregation. Sullivan et al. (2003, 2007) encouraged the use of existing data to reduce costs and promote the calculation of the index. Molle and Mollinga (2003) criticized the index for conflating disparate and correlated pieces of information with arbitrary weights. Nardo et al. (2005) argued that the subject weights assigned to WPI components affect the statistical proprieties and interpretabilities of the final values of the index, and that the additive aggregation implies the possibility of offsetting poor performance in some indicators by sufficiently high values of other indicators. Heidecke (2006) recommended further investigation of the equal weighting scheme due to its inadequate explanation.

On the basis of previous criticism, principal component analysis (PCA) was developed by Cho et al. (2010), Gine Garriga and Perez Foguet (2010), 2011), Jemmali and Matoussi (2013) and Jemmali and Sullivan (2014) to revise the method of calculation of composite WPI. Cho et al. (2010) used PCA to simplify WPI from five components comprised of Resources, Access, Capacity, Use and Environment to three components comprised of Access, Capacity and Environment. Gine Garriga and Perez Foguet (2010) and Jemmali and Sullivan (2014) adopted PCA to calculate each component of WPI, and then aggregated the components to obtain the final value of WPI. Jemmali and Matoussi (2013) use PCA to give more weight to components with large variance and to discard components with smaller ones. Their results revealed that PCA can determine the weights objectively and avoid the problem of ambiguity among components. However, in the previous studies, PCA is used only in the calculation of each component, not oriented to all indicators. Therefore ambiguity may still exist in the holistic indicators. Moreover, water is highly variable both on a spatial and temporal scale (Sullivan et al. 2007), which adds two dimensions into the analysis of water poverty so that the variable matrix tends to be very large and the results of dimension reduction by the traditional PCA is not very good. To strengthen the interpretability of PCA and overcome the limitations of the traditional PCA, the transformation of centralized logarithm of initial variable and the methodology of holistic and dynamic PCA are proposed in Sect. 2. In Sect. 3, a step-by-step procedure for developing the composite WPI for Chinese case is provided. In Sect. 4, the improved holistic and dynamic PCA is applied to analyze regional water poverty trends and influential factors in China. A discussion about the derived policy implications of the obtained results and a conclusion constitute the final section.

2 Improved Holistic and Dynamic PCA

2.1 Limitations of Traditional PCA

Principal components analysis (PCA) is used to transform a large set of correlated variables into a smaller set of uncorrelated components, called principal components, which account for most of the variation in the original set of variables (Dunteman 1989). Thus, the complex relationship of initial variables can be simplified due to the uncorrelated principal components. The traditional PCA is primarily oriented to static data, which consists of indicators and samples, and it fails to reveal the evolutionary trend of data over time.

The traditional PCA is a linear dimensional reduction technique because the principal component is the linear combination of the initial variables with characteristic vectors of the correlation or covariance (Jolliffe 1986; Johnson and Wichern 2007). In the process of PCA, the initial data needs to be normalized in order to eliminate the influence of various units of measure and scale of parameters, especially for the multidimensional indicators of WPI in this study. There are some limitations of the conventional normalization methods. For examples, the mean–variance normalization method fails to retain the information of variance among the initial data due to the normalization, and the min- max normalization, minimum normalization, maximum normalization, and mean normalization hinder comparability of the results because the maximum, minimum and mean values are highly dependent on the sample. Moreover, these methods are based on the coefficient matrix of the correlation or covariance subject to the two-dimensional uniform distribution (0,1), and

not adequate when the dimension of time is added to the data as the relationship between indicators and sample data is nonlinear over time (Ye 2001; Liu et al. 2012).

Regarding a set of p-dimensional component data, $\sum_{i=1}^{p} x_i = 1$, then for the component of x_i,

$$\sum_{l=1}^{p} Cov(x_i, x_l) = 0, \sum_{i \neq l} Cov(x_i, x_l) = -Var(x_i) \tag{1}$$

Because $Var(x_i) \succ 0$, $(Var(x_i) = 0$ when x_i is constant), there is at least a negative covariance in $(p-1)Cov(x_i, x_l), i \neq l$. In other words, there are at least p negative covariances in the theoretical matrix of covariance $V = (Cov(x_i, x_l)), (i \neq l)$. Their correlation coefficients don't comply with the uniform distribution of $(0, 1)$. For the case of p-dimensional (x_1, x_2), $Cov(x_1, x_2) = -Var(x_1) = -Var(x_2)$, the correlation coefficient between x_1 and x_2 will be

$$\rho = Cov(x_1, x_2)\{Var(x_1)Var(x_2)\}^{-\frac{1}{2}} = -1 \tag{2}$$

This indicates the correlation coefficient of p-dimensional component data is certain to be -1, rather than fitting the uniform distribution of $(0, 1)$. The negative skewness of the coefficient matrix of the correlation or covariance is so significant in the component data that it is difficult to explain the p-dimensional component data by using the traditional normalized PCA, such as min–max normalized PCA, which is based on the coefficient matrix of correlation or covariance in the hypothesis of the two-dimensional uniform distribution of $(0, 1)$.

2.2 The p-dimensional component data transformation of centralized logarithm

According to Zhang and Chen (1996), Ye (2001) and Liu et al. (2012), assuming there is a p-dimensional component data vector $X = (x_1, x_2, \ldots, x_p)$, the method of centralized logarithm can be adopted to transform the vector into

$$y_i = \ln(x_i/g(x)), g(x) = (x_1 x_2 \ldots x_p)^{-\frac{1}{p}} \tag{3}$$

The random vector $Y = (y_1, y_2, \ldots, y_p)$ is in the p-dimensional real space R^p, because the space $\left\{(x_1, x_2, \ldots, x_p) \sum_{i=1}^{p} x_i = 1, x_i \succ 0 \right\}$ corresponds with R^p by the transformation of the centralized logarithm. The negative skewness of the covariance matrix of the component data is eliminated by entering the real space of R^p. The principal component analysis of $X = (x_1, x_2, \ldots, x_p)$ is transformed into analysis of $Y = (y_1, y_2, \ldots, y_p)$.

The principal component analysis of $Y = (y_1, y_2, \ldots, y_p)$ focuses on finding the maximum value of the variance $Var(a'Y)$ in the condition of $a'a = 1$ and $a = (a_1, a_2, \ldots, a_p)$. It is easy to obtain the equation that

$$Var(a'Y) = Var(a' \ln(x/g(x))) = a'\tau a \tag{4}$$

Then the variance matrix of the centralized logarithm is

$$\tau = \left[Cov(\ln(x_i/g(x)), \ln(x_j/g(x))) \right] \tag{5}$$

Assuming that p characteristic roots of τ are $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p$, the corresponding characteristic vectors are $a_1, a_2,\ldots, a_p$, and $(\tau - \lambda_i I)a_i = 0, \quad i = 1, 2, \ldots p$, then the ith principal component is $a_i' \ln(x/g(x))$ which is a nonlinear combination of x.

2.3 Holistic and Dynamic Principal Component Analysis

In the traditional PCA, the static data $R_{n\times p}$ consists of samples $(e_1, e_2,\ldots, e_n)$ and indicators $(x_1, x_2,\ldots, x_p)$. When time is taken into consideration, X_t is the array of time-series data of $R_{n\times p}$, which can be denoted as $K = \{X_t \in R_{n\times p}, t = 1, 2,\ldots, T\}$. Assuming that in the time-series data X_t, N_t are the values of indicators $(x_1, x_2, \ldots, x_p)$ of samples $(e_1, e_2,\ldots, e_n)$ at time t, then $N = \overset{T}{\underset{t=1}{U}} N_t$ is the cluster of the values of indicators $(x_1, x_2,\ldots, x_p)$ of samples $(e_1, e_2,\ldots, e_n)$ in period T. The cluster N is the objective of the holistic and dynamic principal component analysis.

Based on what we have discussed in Sect. 2.2, assuming that the p-dimensional component vector $X = (x_1, x_2,\ldots, x_p)$ has the initial variable $(x_{ijt})_{p\times n\times T}$ which can be simplified as $(x_{ik})_{p\times nT}$, the holistic and dynamic PCA involves four key steps:

Firstly, transform the initial variable into the centralized logarithm

$$y_{ik} = \ln x_{ik} - \frac{1}{nT}\sum_{k=1}^{nT} \ln x_{ik} \tag{6}$$

Secondly, compute the covariance matrix of the centralized logarithm sample

$$S = (S_{ef})_{p\times p} \tag{7}$$

where $S_{ef} = \frac{1}{nT-1}\sum_{k=1}^{nT}(y_{ek} - \overline{y_e})(y_{fk} - \overline{y_f})$, $\overline{y_e} = \frac{1}{nT}\sum_{k=1}^{nT} y_{ek}, \overline{y_f} = \frac{1}{nT}\sum_{k=1}^{nT} y_{fk}$.

Thirdly, assuming that $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p$ are p characteristic roots of S, and $a_1, a_2, \ldots, a_p$ are the corresponding characteristic vectors, then the value of the ith principal component is

$$F_i = \sum_{j=1}^{p} a_{ij} \ln x_{ij} \tag{8}$$

Finally, the principal components are weighted with the corresponding proportion of variance in the original set of variables explained by that particular principal component, and the comprehensive index is computed using the following formula:

$$Z = \sum_{i=1}^{m}(\omega_i F_i) \bigg/ \sum_{i=1}^{m} \omega_i \tag{9}$$

where Z is the value of the complex index, and ω_i is the proportion of variance explained by the ith principal component. The bigger the value of Z, the poorer the regional water situation, and vice versa.

3 Developing the Water Poverty Index for the Chinese Case

3.1 Study Area

With rapid economic growth, the gap between water resource availability and water demand for ecological protection is gradually widening in China. The Chinese government has been aware of the water scarcity problem and began to reform its water resource management system in the late 1990s. China's Water Law has been modified several times, and related water management regulations number in the hundreds or thousands. Yet the problems of water shortages and degraded water quality remain severe. The annual per capita freshwater resource is 2156 m^3, which is less than a quarter of the global average (Jiang 2009). The direct impact of water pollution is costing China about 1 percentage of its GDP each year (OECD 2007). Water resource management is a top priority in the government's policy agenda to promote natural resource conservation (SC 2006).

China has 22 provinces, four municipalities directly under the control of the central government (Beijing, Tianjin, Shanghai and Chongqing) and five autonomous regions (Guangxi, Xinjiang, Qinghai, Ningxia, and Tibet). In the national statistical process, the provinces, autonomous regions and municipalities directly under the control of the central government are investigated as observations at the same level because they own the equal administrative status in China. This study area covers all provinces, municipalities directly under the control of the central government and autonomous regions except Tibet, which is

Table 1 Eight regions in China

Region	Province/autonomous region/municipalities directly under the control of the central government	Features
Northeast	Liaoning, Jilin, Heilongjiang	Songhuajiang river, specialized agriculture, heavy industry, energy industry
North Coast	Beijing, Tianjin, Hebei, Shandong	Haihe river, convenient transportation, high-tech R&D and production
East Coast	Shanghai, Jiangsu, Zhejiang	Yangtze river, open-economy, the most competitive region, multi-function manufacturing
South Coast	Fujian, Guangdong, Hainan	Southeastern rivers, highly open-economy, foreign technology assimilation, high-tech production
Central Yellow River Delta	Shaanxi, Shanxi, Henan, Inner Mongolia	Huaihe river, coal mining and refining, natural gas, water energy, steel, metal, cereal, dairy products
Central Yangtze River Delta	Hubei, Hunan, Jiangxi, Anhui	Yangtze river, rice and cotton cultivation, steel and metal, various industries
Great Southwest	Yunnan, Guizhou, Sichuan, Chongqing, Guangxi	Southwestern rivers, tourism, heavy industry and light industry
Great Northwest	Gansu, Qinghai, Ningxia, Xinjiang	Northwestern rivers, cotton, fruit, and rice cultivation, livestock, energy industry, trade, tourism

not considered due to its insufficient data from 2004 to 2012. The endowment of natural resources and economic characteristics of study areas are presented based on the eight regions proposed by the State Council in 2006 (SC 2006), which are Northeast, North Coast, East Coast, South Coast, Central Yellow River Delta, Central Yangtze River Delta, Great Southwest and Great Northwest (shown in Table 1).

3.2 Selection of the indicators

In order to minimize the potential for ad hoc selection of indicators (Cho et al. 2010), a comprehensive review of indicators used in regional water poverty studies was firstly developed, which provides relevant indicators to represent the five components of WPI adopted in the study (Table 2). Secondly, the China Statistical Yearbook and the China Environmental Yearbook were inspected to select appropriate indicators for the measurement of each dimension of WPI to confirm the availability, reliability, understandability, regular updatability (Feitelson and Chenoweth 2002, Zhou et al. 2006, Nardo et al. 2005) of the panel data from the year of 2004–2012. As shown in Table 3, 26 indicators were obtained.

The third step is to classify appropriate indicators based on the WPI framework from Lawrence et al. (2002) because these five components are considered to integrate the complexity of water sector. The indicators of Resources include the amount of groundwater and surface water resources in order to assess total water availability. The indicators of Access cover access to drinking water and to sewage disposal. The indicators of Capacity focus on water institutional capacity by combining a set of regional development indicators. The indicators of Use combine the stress of water use and efficiency of water use in agricultural and industrial production, discarding domestic water use per capita and industrial water use per capita because increasing water use is good up to a certain point but it becomes bad if inefficiency results from waste water (Lawrence et al. 2002). The indicators of Environment cover aquatic environmental pollution and aquatic environmental protection. 26 indicators were classified, as shown in Table 3.

The fourth step is to pretreat the initial data by the min–max normalization method which is widely used in many WPI studies (adapted from Lawrence et al. 2002, Sullivan 2005, Komnenic et al. 2009). In the case that the higher the value of the variable, the less the regional water poverty, the variables are transformed into indicators using the following formula:

$$x_i^* = \frac{x_{\max} - x_i}{x_{\max} - x_{\min}} \tag{10}$$

Table 2 Basic steps in composite index design

1st: selection of indicators	1.a. Compile of the relevant indicators
	1.b. Select the appropriate indicators
	1.c. Classify appropriate indicator
	1.d. Pretreat the initial data
	1.e. Filter the indicators
2nd: construction of composite index.	2.a. Compute the covariance matrix of the centralized logarithm sample
	2.b. Calculate the value of each principal component
	2.c. Calculate the final value of the comprehensive index

Table 3 Specification of indicators used in the WPI framework

Component	Sub-component	Indicator	Formula (Unit)	Assumed functional relationship
Resource (R$_1$)	WR availability (R$_{11}$)	WR per unit area (R$_{111}$)	Total WR/land (M^3/km^3)	Higher R$_{111}$-less water poverty
		WR production coefficient (R$_{112}$)	Total WR/precipitation	Higher R$_{112}$-less water poverty
		WR per capita	*Total WR/population (M^3/person)*	*Higher the value, less water poverty*
Access (R$_2$)	Access to supply coverage (R$_{21}$)	Rural tap water access rate (R$_{211}$)	Rural population using tap water/all rural population	Higher R$_{211}$-less water poverty
		Urban water access rate (R$_{212}$)	Urban population using tap water/all urban population	Higher R$_{212}$-less water poverty
		Urban drainage pipe access rate (R$_{213}$)	Length of urban drainage pipes/all urban population (m/person)	Higher R$_{213}$-less water poverty
	Access to sanitation coverage (R$_{22}$)	Percentage of urban sewage disposal (R$_{221}$)	Quantity of urban waste water treatment/total quantity of urban sewage discharged	Higher R$_{221}$-less water poverty
		Proportion of industrial waste water meeting discharge standards in total volume (R$_{222}$)	Quantity of industrial waste water meeting discharge standards/total quantity of industrial waste water discharged	Higher R$_{222}$-less water poverty
Capacity (R$_3$)	Governance capacity (R$_{31}$)	Execution level of environmental projects (R$_{311}$)	The environmental protection investment in the environmental protection acceptance projects/total investment for the environmental protection acceptance projects	Higher R$_{311}$-less water poverty
		Proportion of environmental protection investment in GDP	*Investment for environmental protection projects/total GDP*	*Higher the value, less water poverty*
	Economic capacity (R$_{32}$)	Proportion of tertiary industry in GDP (R$_{321}$)	The added value of tertiary industry/GDP	Higher R$_{321}$-less water poverty
		GDP per capita	*Total GDP/total population (10,000 yuan/person)*	*Higher the value, less water poverty*
	Literary capacity (R$_{33}$)	Number of college students per 10,000 people (R$_{331}$)	Number of College Students/population/10,000 (person)	Higher R$_{331}$-less water poverty

Table 3 continued

Component	Sub-component	Indicator	Formula (Unit)	Assumed functional relationship
	Residents' welfare (R_{34})	Engel's coefficient of households (R_{341})	Food's expenses/total living expenses of households	Higher R_{341}-higher water poverty
		Income ratio between the urban and the rural households	Ratio between the income of urban households and the income of rural households	*Higher the value, higher water poverty.*
Use (R_4)	Pressure of water use (R_{41})	Percentage of agricultural water (R_{411})	Agricultural water use/total water use	Higher R_{411}-higher water poverty
		Water use intensity (R_{412})	Actual water use/total water supply	Higher R_{412}-higher water poverty
	Efficiency of water use (R_{42})	Water-saving irrigation rate (R_{421})	Water-saving irrigation area/cultivated area	Higher R_{421}-less water poverty
		Rate of recycling use for industrial water (R_{422})	The volume of recycling use for industrial water/total industrial water	Higher R_{422}- less water poverty
		Water consumption per 10,000 yuan of GDP	*The amount of water use/GDP/10,000*	*Higher the value, less water poverty*
		Water consumption per 10,000 yuan of industrial value added	*The amount of water use/industrial value added/10,000*	*Higher the value, less water poverty*
Environment (R_5)	Environmental stress (R_{51})	Rate of chemical oxygen demand discharged (R_{511})	The amount of chemical oxygen demand discharged/waste water discharged (g/m³)	Higher R_{511}-higher water poverty
		Rate of ammonia nitrogen discharged (R_{512})	The amount of ammonia nitrogen discharged/ waste water discharged (g/m³)	Higher R_{512}-higher water poverty
		Fertilizer and pesticide amount per unit cultivated land (R_{513})	The amount of fertilizer and pesticide applied/ cultivated land (Kg/hm²)	Higher R_{513}-higher water poverty
	Environmental protection(R_{52})	Public green areas per capita (R_{521})	Public green areas/total population	Higher R_{521}-less water poverty
		Proportion of wetland in total area of territory	*Wetland areas/total area of territory*	*Higher the value, less water poverty*

The italic fonts are discarded during the final selection of indicators based on the collinear analysis and balance of indicators

In the case that the higher the value of the variable, the poorer the regional water situation, the variables are transformed using the following equation:

$$x_i^* = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{11}$$

The descriptive statistical analysis of 26 indicators, covering 22 provinces, four municipalities directly under the control of the central government and four autonomous regions from 2004–2012, indicated the negatively skewed distribution of most data, which justify the conclusion in the section of 2.2 that the correlation coefficient of p-dimensional component data is negative, rather than fitting the uniform distribution of (0, 1).

Thus the initial variables were rescaled using the centralized logarithm as presented in Eq. (6) if the higher the value of the variable, the poorer the regional water situation. The bottom of the initial variable was applied to the centralized logarithm if the higher the value of the variable, the less the regional water poverty.

The fifth step is to filter the indicators for least eclipsing through comparison of the correlation coefficients between any two indicators and significance levels. Based on the comparison of the correlation coefficients (above 0.7) between any two indicators and significance levels (<0.01), seven redundant indicators were removed: water resource per capita, proportion of environmental protection investment, GDP per capita, income ratio between the urban and the rural households, water consumption per 10,000 yuan of GDP, water consumption per 10,000 yuan of industrial value added, and proportion of wetland in total area of territory. For instance, two variables, GDP per capita and proportion of tertiary industry in GDP, are used to measure the economic capacity, but GDP per capita is highly correlated with the number of college students per 10,000 people, Engel's coefficient of households, rural tap water access rate, urban water access rate, and percentage of urban sewage disposed (the correlation coefficient is above 0.70 and the significance level is <0.01). The variable GDP per capita was then discarded and proportion of tertiary industry in GDP was retained in the sub-component of economic capacity.

Bartlett's spherical test on the retained indicators is significant at the 0.01 level ($\chi^2 = 171$; P value = 0.000), indicating significant correlations. The Kaiser–Meyer–Olkin Measure of Sampling Adequacy is 0.685, which falls in the acceptable range. These test results support that the remaining 19 indicators are suitable for PCA. Thus, the indicators were finally selected such that they were measurable, reliable, valid and comparable, stayed independent with each other, and could be associated with the practices and policies of water resources management in China. The specification of parameters and indicators used in the WPI framework are presented in Table 3.

3.3 Construction of the Composite Index

As mentioned in the Sect. 2.3, compute the covariance matrix S of the centralized logarithm sample according to the Eq. (7). Then assuming that $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_p$ are p characteristic roots of S, and $a_1, a_2, \ldots, a_p$ are the corresponding characteristic vectors, then the value of the ith principal component is calculated according to the Eq. (8). Finally the principal components are weighted with the corresponding proportion of variance in the original set of variables explained by that particular principal component, and the comprehensive index is computed using the Eq. (9). The basic steps of composite index design is presented in Table 2.

3.4 Comparison of CLHDPCA and MMHDPCA

Factors contributing to the Water Poverty Index based on the min–max normalized holistic and dynamic PCA (MMHDPCA) and the central-logarithm normalized holistic and dynamic PCA (CLHDPCA), are constructed using SPSS17.0 for Windows (Table 4).

The cumulative variance explained by four components in CLHDPCA is 76.90 %, compared to 57.72 % in MMHDPCA. Four components in CLHDPCA are enough to account for more than 75 % of the variation while seven components are required in MMHDPCA, demonstrating the greater dimensional reduction in CLHDPCA compared to MMHDPCA. Therefore in the following discussion, four components are extracted, and based on CLHDPCA the comprehensive index (Z) is computed using Eq. (9) where $m = 4$.

4 Empirical Results

4.1 Water Poverty Trends in China

The WPI results based on the central-logarithm normalized holistic and dynamic PCA for China's provinces, autonomous regions, and municipalities directly under the control of the central government are exhibited in Table 5 and Fig. 1. In Table 5, Z_{2012} and Z_{2004} respectively represent the WPI values in 2012 and in 2004, while $Z_{2004–2012}$ represents the

Table 4 Total variance explained by centralized logarithm (CL) holistic and dynamic PCA and min–max normalized (MM) holistic and dynamic PCA

Component	Initial Eigenvalues (by CLHDPCA)			Initial Eigenvalues (by MMHDPCA)		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.104855	43.69193	43.69193	4.75699	25.03679	25.03678896
2	1.343936	14.30481	57.99674	2.974846	15.65709	40.69387515
3	1.110739	11.82266	69.8194	1.805647	9.503403	50.1972782
4	0.665357	7.08204	76.90144	1.429118	7.521673	57.71895135
5	0.456494	4.858903	81.76035	1.32965	6.998157	64.7171085
6	0.416912	4.437601	86.19795	1.024333	5.391229	70.10833712
7	0.287389	3.058962	89.25691	0.962414	5.065337	75.17367401
8	0.249184	2.652307	91.90922	0.885409	4.66005	79.83372399
9	0.211779	2.254171	94.16339	0.605874	3.188808	83.02253218
10	0.13321	1.417877	95.58126	0.572852	3.015012	86.03754442
11	0.120232	1.279745	96.86101	0.536837	2.825458	88.86300276
12	0.094885	1.00995	97.87096	0.454067	2.389826	91.25282857
13	0.067547	0.718967	98.58993	0.428975	2.257761	93.51059003
14	0.054656	0.581759	99.17169	0.371497	1.955246	95.46583609
15	0.036994	0.393765	99.56545	0.323429	1.702256	97.16809219
16	0.024663	0.262507	99.82796	0.23387	1.230893	98.39898492
17	0.008003	0.085184	99.91314	0.150165	0.790342	99.18932711
18	0.004698	0.05	99.96314	0.095736	0.503874	99.69320095
19	0.003463	0.036858	100	0.058292	0.306799	100

Table 5 Water poverty index (Z) and its rank (R) based on centralized logarithm holistic and dynamic PCA

Region	Z_{2012}	R_{2012}	$Z_{2004-2012}$	$R_{2004-2012}$	Z_{2004}	R_{2004}
Inner Mongolia	1.64	1	1.08	1	1.08	3
Xinjiang	1.53	2	0.55	5	−4.65	30
Gansu	1.48	3	0.65	4	−1.72	27
Heilongjiang	1.44	4	0.09	14	−0.51	23
Qinghai	1.22	5	1.03	2	1.11	2
Ningxia	1.2	6	0.76	3	0.95	5
Shaanxi	1.16	7	0.39	8	0.39	9
Shanxi	1.16	8	0.51	6	0.51	6
Hainan	1.15	9	0.36	9	1.05	4
Jilin	1.13	10	0.13	12	0.13	12
Hunan	1.06	11	0.12	13	−0.14	18
Liaoning	1.03	12	0.02	16	−0.35	21
Hebei	0.99	13	0.18	11	0.08	13
Guizhou	0.92	14	−0.43	23	−0.09	16
Shandong	0.9	15	0.23	10	0.27	10
Henan	0.83	16	−0.02	17	−0.04	15
Sichuan	0.77	17	−0.27	20	−1.48	26
Anhui	0.73	18	−0.28	21	−0.33	20
Hubei	0.72	19	−0.07	18	0.04	14
Yunnan	0.71	20	−0.32	22	−3.18	29
Jiangxi	0.58	21	−0.45	24	0.49	8
Guangxi	0.54	22	−0.22	19	−2.18	28
Tianjin	0.53	23	0.5	7	1.24	1
Chongqing	0.5	24	0.05	15	0.49	7
Jiangsu	0.38	25	−0.53	26	−0.26	19
Guangdong	0.35	26	−0.47	25	0.13	11
Fujian	0.33	27	−0.87	29	−0.73	25
Beijing	0.13	28	−1	30	−0.49	22
Zhejiang	0.11	29	−0.54	27	−0.57	24
Shanghai	0.03	30	−0.86	28	−0.11	17

median value between 2004 and 2012. R_{2012} and R_{2004}, respectively, represent the rank of the WPI in 2012 and 2004, while $R_{2004-2012}$ represents the rank of the median value between 2004 and 2012. Taking the region of Shanghai as an example, the 19 initial variables are normalized using centralized logarithm, 4 principal components are extracted, and the final values of the WPI from 2004 to 2012 are calculated according to Eq. (9). The results are $Z_{2004} = -0.11$, $Z_{2004-2012} = -0.86$ and $Z_{2012} = 0.03$. Shanghai's rank (R) of the final values (closer to the top rank means more water poverty) of the complex index of water poverty yearly from 2004 to 2012 can be obtained after all other regions' final values of the complex index of water poverty are computed. The ranks of the final values of the complex index of water poverty in Shanghai are $R_{2004} = 17$, $R_{2004-2012} = 28$ and $R_{2012} = 30$. Comparing the value in 2004 with that in 2012 (shown in Table 5), the value
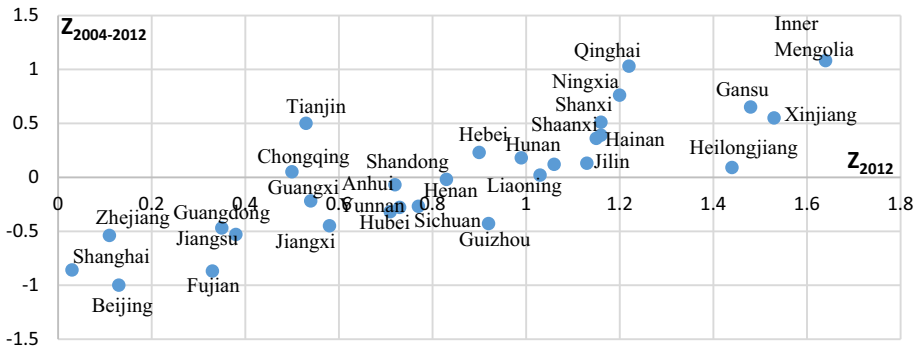
Fig. 1 Regional WPI values for 2012 and median values for 2004–2012

of the complex index (Z) of water poverty in nearly each region (except Tianjin) increased, which revealed that the general situation of water poverty was worsening.

Although the general trend of regional water poverty is similar, we find differing trajectories of the value of the complex index (Z) in different regions. Some provinces or autonomous regions located in Central Yellow River Delta, Great Northwest and South Coast, including Inner Mongolia, Qinghai, Ningxia, Gansu, Xinjiang, Shaanxi, Shanxi, and Hainan remain in the worst water poverty for recent years, whereas the level of water poverty in Xinjiang was the lowest in 2004. Four provinces and two municipalities directly under the control of the central government, located in East Coast, South Coast and North Coast, including Beijing, Shanghai, Fujian, Zhejiang, Jiangsu, and Guangdong, had the lowest water poverty in 2012. The level of water poverty in Tianjin, Chongqing, Hubei, Shandong, and Hebei has undergone the process of first declining and then rising. In Heilongjiang, Jilin, Liaoning, Hunan, Guizhou, Henan, Sichuan, Anhui, Yunnan, Guangxi, and Jiangxi the level of water poverty has increased steadily for recent years.

The results are presented further according to the ranks of comprehensive scores of WPI in different regions between 2004 and 2012 as shown in Table 4. In 2004, the regions of the worst water poverty (ranking from 1 to 8) were scattered across the Great Northwest (such as Qinghai and Ningxia), Central Yellow River Delta (as Inner Mongolia, and Shanxi), South Coast (as Hainan), and Southwest (as Chongqing). After several years, the municipalities directly under the control of the central government including Beijing, Tianjin, Chongqing, and Shanghai, and some developed provinces of southeastern coast, such as Guangdong, Fujian, Zhejiang, and Jiangsu, have dropped in the ranks of WPI. In 2012, the worst water poverty areas (ranking from 1 to 8) were clustered in the whole Great Northwest covering Qinghai, Ningxia, Xinjiang and Gansu, and Central Yellow River Delta such as Inner Mongolia, and Shanxi. The changing trajectories of the worst water poverty regions (ranking from 1 to 8) are exhibited as the darkest areas in Fig. 2, which indicates that regional water poverty has become more spatially clustered in the Great Northwest for recent years.

4.2 Influential Factors Associated with Water Poverty

Which factors drive the WPI for China? The principal components whose eigenvalues are more than one, which explain the majority of the variance in the original data, are the primary influential factors producing the final results. There are three principal components
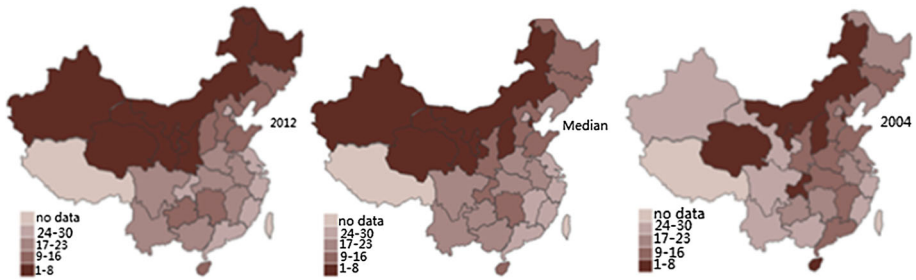
Fig. 2 The comprehensive ranks of WPI in different regions between 2004 and 2012

Table 6 Factor loading pattern

	P (water pollution factor)	R (water resource availability factor)	C(capacity-access-use efficiency factor)
Number of college students per 10,000 people (R_{321})	−0.25988	0.011221	0.688864
Urban drainage pipe access rate (R_{213})	−0.11793	0.011428	0.676526
Rural tap water access rate (R_{211})	0.011101	0.142988	0.610954
Percentage of urban sewage disposed (R_{212})	−0.09809	0.042005	0.581906
Water-saving irrigation rate (R_{421})	0.032597	−0.37252	0.747167
Proportion of industrial waste water meeting discharge standards in total volume (R_{222})	−0.02762	0.334746	0.505324
Urban water access rate (R_{212})	−0.08795	0.016213	0.511511
Proportion of tertiary industry in GDP (R_{351})	−0.3263	0.091077	0.403536
WR production coefficient (R_{112})	0.257235	0.760056	0.02462
WR per unit area(R_{111})	0.170142	0.929508	0.001397
Fertilizer and pesticide amount per unit cultivated land (R_{523})	−0.16985	−0.43559	−0.6321
Water use intensity (R_{412})	0.103005	0.533345	0.157082
Engel's coefficient (R_{341})	−0.13987	−0.47073	0.222411
Percentage of total water used by agriculture (R_{411})	0.115661	0.542303	0.39544
Rate of ammonia nitrogen discharged (R_{522})	0.964831	−0.10247	−0.10658
Rate of chemical oxygen demand discharged (R_{521})	0.913704	0.003212	0.215403
Rate of recycling use for industrial water (R_{422})	−0.14103	−0.20589	0.445497
Execution level of environmental projects (R_{311})	−0.04563	−0.17463	−0.1804
Public green areas per capita (R_{521})	−0.08016	0.136157	0.161472

P, R, and C correspond to components 1, 2, and 3, respectively, in Table 3, based on the CL PCA

(whose eigenvalues are more than one) to be analyzed as shown in Table 4 and 6. Among them, component 1 (Table 4, CLHDPCA) was characterized by a high positive loading for rate of ammonia nitrogen discharged (loading of 0.965, Table 6) and rate of chemical

oxygen demand discharged (loading of 0.914, Table 6). This component was interpreted as "pollution factor (P)", and explained 43.69 % of the variance in the data (Table 4). Obviously, water pollution is the current dominant factor contributing to regional water poverty in China. Together with the regional scores of principal components in Table 7, we can find that the score range of water pollution dimension is varying from [−11.19, 2.99] in 2004 to [0.72, 3.09] in 2012 with [−1.87, 1.48] as the range for the median. The lower limit in each range keeps rising, which indicates that the general level of regional water pollution is increasing. As water pollution is the heaviest loading component in all principal components, these results illustrate that aquatic environmental pollution is the main cause of China's water poverty and it has been worsening in the past 9 years. And furthermore, as exhibited as the darkest areas in Fig. 3, the changing trajectories of the heaviest water pollution regions (ranking from 1 to 8) indicate that regional water pollution has become more spatially clustered in recent years.

Component 2 was characterized by high positive loading for water resource per unit area, and water resource production coefficient. This component was interpreted as "resource availability factor (R)". This factor explained 14.30 % of the variance in the data (Table 4). This result reveals that water resource endowment is still an important endogenous factor concerning regional water poverty in China. In Table 7, the score range of water resource availability dimension is narrowing from [−2.67, 2.85] in 2004 to [−2, 1.92] in 2012 with [−2.65, 2.05] as the median of these 9 years, which indicates that the regionally imbalanced distribution of water resources is decreasing through the reform of the water resource distribution system covering South-to-North Water Diversion Project, interregional water rights transfer, and integrated watershed management. Beijing, Tianjin, Hebei, Shaanxi, Liaoning, and many other northern areas directly benefited from the South-to-North Water Diversion Project. The interregional water rights transfer and integrated watershed management strategy are gradually moving forward especially in the East Coast and Great Northwest. And yet for all that, the trajectories of the poorest water resource availability regions (ranking from 1 to 8) indicate that the scarcest area of water resources has still concentrated in the northern China, as exhibited as the darkest areas in Fig. 4.

Component 3 was characterized by high positive loading for the number of college students per 10,000 people, rural tap water access rate, urban drainage pipe access rate, water-saving irrigation rate, percentage of urban sewage disposed, rate of recycling use for industrial water, and proportion of tertiary industry in GDP. This component was interpreted as "capacity-access-use efficiency factor (C)", which explained 11.82 % of the variance in the data (Table 6). Some information is presented from the indicators of higher negative loading in each principal component. There is high negative loading for proportion of tertiary industry in GDP, and the number of college students per 10,000 people in Component 1, suggesting a negative relation between pollution and a more educated, service-oriented workforce. There is a high negative loading for Engel's coefficient, fertilizer and pesticide amount per unit cultivated land, and water-saving irrigation rate on component 2, suggesting water resource endowment is negatively related with poverty, agricultural chemical input use, and water-saving technology. There is a high negative loading for fertilizer and pesticide amount per unit cultivated land in Component 3, suggesting capacity-access is negatively related with agricultural chemical input use. Furthermore, as shown as the darkest areas in Fig. 5, the trajectories of the poorest capacity-access-use efficiency regions (ranking from 1 to 8) indicate that these areas have changed from the eastern coast to southwestern coast in recent years.

Table 7 Factors' scores (S) and ranks (R)

Region	P (water pollution factor)						R (water resource availability factor)					
	S_{2012}	R_{2012}	$S_{2004–2012}$	$R_{2004–2012}$	S_{2004}	R_{2004}	S_{2012}	R_{2012}	$S_{2004–2012}$	$R_{2004–2012}$	S_{2004}	R_{2004}
Inner Mongolia	3.09	1	1.48	1	1.48	5	1.97	1	2.05	1	2.02	3
Xinjiang	2.73	4	0.27	11	−11.19	30	1.32	2	1.5	7	2.85	1
Gansu	2.97	3	0.64	6	−5.05	28	1.2	4	1.78	2	2.64	2
Heilongjiang	2.98	2	−0.13	18	−1.63	25	0.42	10	0.64	12	0.76	10
Qinghai	2.47	10	1.42	3	1.8	3	0.48	9	1.02	7	0.83	7
Ningxia	2.58	7	1.33	4	1.54	4	0.75	7	0.96	8	0.79	9
Shaanxi	2.48	9	0.4	10	0.4	11	1.05	5	1.61	3	1.36	5
Shanxi	2.36	12	0.72	5	0.56	9	1.21	3	1.52	4	1.47	4
Hainan	2.7	5	0.61	8	2.01	2	−0.19	14	0.58	13	0.8	8
Jilin	2.62	6	−0.06	16	−0.06	15	−0.44	17	−0.13	16	0.05	16
Hunan	2.43	11	0	15	−0.59	19	−0.12	12	−0.13	15	0.05	15
Liaoning	2.5	8	−0.24	19	−0.18	16	−0.13	13	0.7	10	0.7	11
Hebei	2.25	14	0.18	13	−0.21	17	0.78	6	1.46	6	1.23	6
Guizhou	1.9	21	−0.84	26	−0.18	16	−1.69	28	−0.94	22	−0.98	27
Shandong	2.16	15	0.62	7	0.71	8	0.53	8	0.33	14	−0.01	17
Henan	2.03	18	0.1	14	−0.04	14	0.05	11	−0.17	17	−0.17	19
Sichuan	2.26	13	−0.29	22	−3.17	26	−1.2	22	−0.79	20	−0.71	24
Anhui	2.01	19	−0.24	20	−0.62	20	−1.11	21	−0.93	21	−0.6	22
Hubei	2.04	16	0.27	12	0.53	10	−1.54	25	−1.53	29	−1.56	29
Yunnan	1.98	20	−0.61	24	−7.31	29	−0.96	20	−0.94	23	−0.13	18
Jiangxi	2.03	17	0.44	9	1.32	6	−1.89	29	−1.45	28	−1.45	28
Guangxi	1.66	24	−0.26	21	−4.85	27	−1.52	24	−1.26	27	−0.51	20
Tianjin	1.81	23	1.43	2	2.99	1	−0.21	15	0.73	9	0.43	13
Chongqing	1.83	22	−0.08	17	0.87	7	−2.02	30	−2.65	30	−2.67	30

Table 7 continued

Region	P (water pollution factor)						R (water resource availability factor)					
	S_{2012}	R_{2012}	$S_{2004-2012}$	$R_{2004-2012}$	S_{2004}	R_{2004}	S_{2012}	R_{2012}	$S_{2004-2012}$	$R_{2004-2012}$	S_{2004}	R_{2004}
Jiangsu	1.32	27	−1.28	28	−0.46	18	−0.52	18	−0.52	18	0.13	14
Guangdong	1.37	26	−0.83	25	0.21	13	−1.3	23	−0.73	19	−0.52	21
Fujian	1.61	25	−1.41	29	−1.41	24	−1.63	27	−1.13	26	−0.6	23
Beijing	0.72	30	−1.87	30	−0.84	22	−0.25	16	0.66	11	0.68	12
Zhejiang	1.21	28	−0.57	23	−0.8	21	−1.56	26	−1.13	25	−0.87	25
Shanghai	0.84	29	−1.04	27	0.28	12	−0.61	19	−0.94	24	−0.94	26

Region	C (capacity-access-use efficiency factor)					
	S_{2012}	R_{2012}	$S_{2004-2012}$	$R_{2004-2012}$	S_{2004}	R_{2004}
Inner Mongolia	0.02	11	0.83	7	1.33	6
Xinjiang	−0.21	18	0.59	10	−0.99	27
Gansu	0.19	7	0.27	14	1.21	7
Heilongjiang	0.58	2	0.79	8	1.05	10
Qinghai	0.36	4	1.25	2	1.9	2
Ningxia	−0.19	16	0.56	11	1.5	5
Shaanxi	−0.55	21	0.02	18	0.32	20
Shanxi	−0.1	15	−0.07	20	0.7	14
Hainan	−0.84	23	−0.01	19	0.05	21
Jilin	0.49	3	1.2	3	1.56	4
Hunan	0.3	5	1.03	5	1.2	8
Liaoning	0.05	10	0.37	13	0.81	13
Hebei	−0.72	22	−0.49	23	−0.09	22
Guizhou	1.57	1	0.9	6	1.7	3
Shandong	−0.87	24	−0.76	24	−0.14	23

Table 7 continued

Region	S$_{2012}$	C (capacity-access-use efficiency factor)				
		R$_{2012}$	S$_{2004–2012}$	R$_{2004–2012}$	S$_{2004}$	R$_{2004}$
Henan	−0.41	20	−0.08	21	0.38	19
Sichuan	−0.21	17	0.21	16	0.55	17
Anhui	0.01	12	0.11	17	0.61	15
Hubei	−0.04	13	−0.2	22	0.51	18
Yunnan	0.21	6	1.08	4	1.19	9
Jiangxi	−0.04	14	0.66	9	1	11
Guangxi	0.17	8	0.43	12	0.57	16
Tianjin	−1.94	29	−1.81	29	−1.05	28
Chongqing	0.08	9	1.61	1	2.6	1
Jiangsu	−1.51	26	−1.07	27	−0.69	25
Guangdong	−0.35	19	0.26	15	0.97	12
Fujian	−1.2	25	−0.9	25	−0.16	24
Beijing	−1.92	28	−1.73	28	−1.98	30
Zhejiang	−1.53	27	−1.03	26	−0.74	26
Shanghai	−2.51	30	−2.13	30	−1.26	29
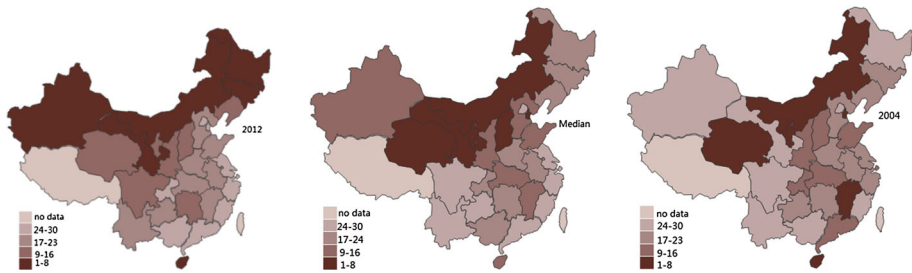
Fig. 3 The ranks of the water pollution factor in different regions between 2004 and 2012
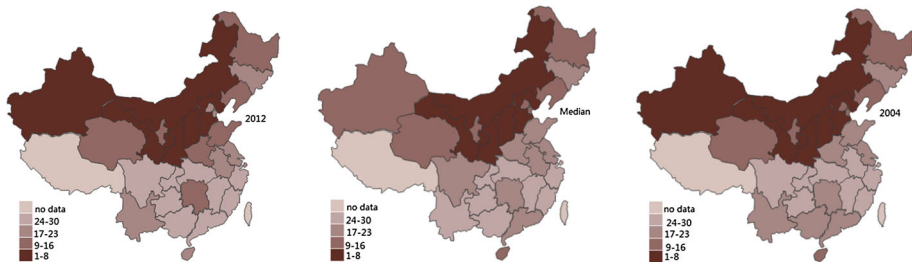
Fig. 4 The ranks of the water resource availability factor in different regions between 2004 and 2012
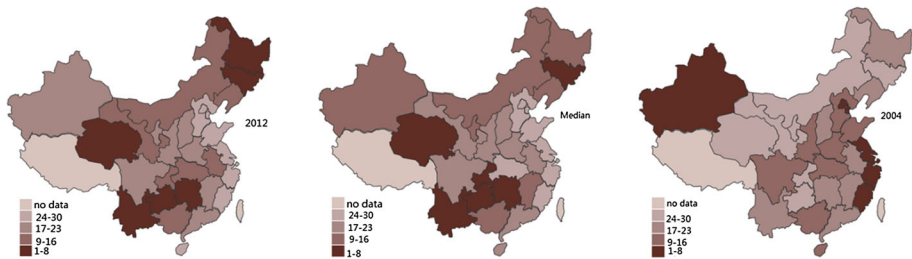
Fig. 5 The ranks of the capacity-access-use efficiency factor in different regions between 2004 and 2012

4.3 Categories of Water Poverty

Based on the performance of three principal components, the study areas are sorted and analyzed in order to understand the factors driving the results for each region (Table 8). The categories of the level of components' performance were defined: best (ranking 21th through 30th), second best (ranking 11th through 20rd), and worst (ranking 1st through 10th). If one or more component(s) category (categories) was (were) in the worst category, the component(s) was (were) included in its overall description. If three components were all in the second best category, the component(s) which ranked the lowest would be included in the overall description. If a component was in the best category, that component was left out. In cases where a region was in the best category for each component, that region was not included.

Table 8 Categories of regional water poverty

	In 2012	Median for 9 years	In 2004
P	Hainan, Sichuan	Hainan, Shandong, Henan, Hubei, Zhejiang	Shandong, Henan, Hubei, Jiangxi, Tianjin, Zhejiang, Shanghai
R	Shanghai, Beijing, Tianjin, Jiangsu, Shandong, Shanxi, Hebei, Henan	Liaoning, Hebei, Beijing,	Xinjiang, Shaanxi, Liaoning, Hebei, Jiangsu, Fujian, Beijing
C	Guizhou, Anhui, Hubei, Hunan, Yunnan, Jiangxi, Guangxi, Chongqing, Guangdong	Heilongjiang, Jilin, Yunnan, Guangxi, Chongqing, Guangdong, Guizhou, Sichuan, Hunan, Anhui	Jilin, Hunan, Guizhou, Sichuan, Anhui, Yunnan, Guangxi, Guangdong
P–R	Inner Mongolia, Shaanxi, Ningxia, Xinjiang	Gansu, Ningxia, Tianjin, Shanxi, Shaanxi	Hainan, Shanxi
P–C	Jilin, Liaoning	Jiangxi	Chongqing
R–C		Xinjiang	Heilongjiang, Gansu
P–R–C	Gansu, Qinghai, Heilongjiang,	Qinghai, Inner Mongolia	Qinghai, Ningxia, Inner Mongolia

For example, in 2012, Gansu's ranks of three principal components are respectively the 3rd, 4th and 7th, placing it in the worst category for each individual component, and thus its primary causes of water poverty are water pollution (P), water resource availability (R) and capacity-access-use efficiency (C). Hainan's ranks for three principal components in 2012 are respectively the 5th (worst) for P, 14th (second best) for R and 23rd (best) for C, therefore the component P was included in the overall description as its primary causes of water poverty. Liaoning's ranks of three principal components in 2004 are respectively the 16th for P, 11th for R and 13th for C, placing in the second best category for each individual component. The component R ranked the lowest, and is labeled as its primary cause of water poverty. Fujian's ranks of three principal components in 2012 are respectively the 25th, 27th and 25th, placing in the best category for each individual component. In this case, the region was not included in Table 8.

The first category is P, water poverty, characterized by high pollution emissions to water bodies. Hainan is typical of this type, having been in high water pollution for recent years. Hainan (latitude 3°30′N–20°7′N and longitude 108°15′E–120°5′E) is one of five Special Economic Zones which have had the priorities to attract foreign capital by exempting such capital investments from taxes and regulations since the early 1980s. From 1980 to 2012, Hainan's GDP increased ten-fold with an annual growth rate of 7.45 %. As a tropical island province, Hainan's marine fishery and marine aquaculture industries contribute greatly to local economic development but also pollution leading to severe stresses on the aquatic environment. Due to lack of environmental protection and wastewater treatment facilities, the annual rate of ammonia nitrogen discharged, rate of chemical oxygen demand discharged, and fertilizer and pesticide amount per unit cultivated land are 0.64, 2.16 and 0.09 respectively, much higher than the national average 0.26, 1.75 and 0.04 respectively. There is an urgent need for water pollution control and natural ecosystem restoration in Hainan.

The second category is R water poverty, featured by poor water resource availability, mainly covering the water-scarce North Coast, Central Yellow River Delta and Great Northwest. However, this category also includes some water-abundant developed regions

located in East Coast and South Coast including Shanghai, Jiangsu and Fujian, where rapid economic growth has given rise to increasing demand for water and stressed water resource availability. Shanghai is located in the downstream of Yangtze River and the Lake Tai basin, and has the polluted water from the upstream and local area. Jiangsu and Fujian face the similar problems. Water poverty in these areas not only derives from limited water resource availability relative to increasing requirements, but also the poor quality renders water unusable.

The third category is C water poverty embracing some water-abundant regions in Great Southwest, Central Yangtze River Delta, and Northeast. The low capacity-access-use efficiency reflects the fact that these regions do not use available water resources efficiently. For example, in Guangdong (latitude $23°02'$N–$23°38'$N and longitude $116°14'$E–$117°19'$E) the average water-saving irrigation rate and the average rate of recycling use for industrial water are 6.1 and 37 % respectively, compared to the national average of 24.61 and 60 %, respectively. Improvements in water-saving are urgently needed to forestall future scarcities as water needs rise.

The fourth category, P–R water poverty, characterized by high pollution and poor availability of water resources includes Xinjiang (latitude $42°45'$N–$44°08'$N and longitude $86°37'$E–$88°58'$E) located in northwestern China. In 2000, Xinjiang was involved in the West China Development Program (WCDP) which was launched to stimulate economic growth in western areas. The region's GDP grew at an annual rate of 14.6 % from 2000 to 2012. Meanwhile, Xinjiang has suffered seriously environmental degradation owing to its natural-resource focused industries. The underground water has been polluted by oil and natural gas overexploitation and poor waste disposal. The R water poverty of Xinjiang in 2004 converted into P–R in 2012 as a consequence of the increasing contamination of underground water.

The fifth category is P–C water poverty, characterized by high pollution and low capacity-access-use efficiency. Chongqing (latitude $28°10'$N–$32°13'$N and longitude $105°11'$E–$110°11'$E), is located upstream of the Three Gorges Reservoir Area. It is critically significant for the safety of Three Gorges Reservoir Area and the middle and lower Yangtze River to treat water pollution in Chongqing. In 2001, Water Pollution Prevention and Control Planning of the Three Gorges Reservoir Area was enacted, and more than 21 billion yuan from the central government was invested in water pollution treatment over 10 years including construction of urban sewage disposal plants, refuse landfill, and sewage pipe networks. As a result, Chongqing has moved from the P–C to the C category since 2011.

The sixth category, R–C water poverty, is characterized by poor water resources and weak capacity-access-use efficiency. For example, Gansu (latitude $35°5'$N–$38°$N and longitude $102°30'$E–$104°30'$E) located in the upper reach of Yellow River, the inland arid area of northwestern China, has water resource per capita less than one-third of the national average, and less than one-eighth of the world's average. Underground water in Gansu has been overexploited in recent years due to the scarcity of surface water. In 2000, Gansu was involved in WCDP for development of its abundant reserves of minerals and coal. Gansu's underground water and ecological environment was further damaged as a result of heavy industry development. As a result, water poverty in Gansu deteriorated from R–C in 2004 to P–R–C in 2012.

Finally, P–R–C water poverty features heavy pollution, poor water resource availability and low capacity-access-use efficiency. P–R–C water poverty is largely found in water-scarce and undeveloped northern China. For an instance, Qinghai (latitude $31°9'$N–$39°19'$N and longitude $89°35'$E–$103°04'$E) has been in this category for all of 2004–2012. Fragile

ecological environment, heavy water pollution, poor availability of water resources, weak access to water supply and sanitation facilities, and relatively low capacity to manage water resources and use efficiency are mainly responsible for the high water poverty of Qinghai.

5 Discussion and Conclusion

WPI results have some implications for water resource planning and management. China has experienced an unprecedented economic boom with real GDP rising at an annual average rate of nearly ten percentage in the past three decades. The improved holistic and dynamic principal component analysis results of WPI in this study show the economic development has improved water resource availability, physical access to water, water-saving facilities, sewage treatment utilities, and the regional distribution of water resources, but it does not automatically translate into less water poverty due to the rising consumption of water and increasing serious pollution discharge. The comprehensive scores and ranks of WPI from the use of panel data demonstrate that the general situation of water poverty is worsening in China, and the water poverty areas have become more spatially clustered in the Great Northwest for recent years. We highlight that it is a top priority for the policy-maker to consider the impact on the aquatic environment when setting any strategy of booming economy because water scarcity and poor water quality will directly threaten Chinese food security, economic development and the quality of life. It is believed that government policies play a critical role to drive sustainable water resource development in China, and the problem-specific policy interventions and planning would help improve the regional water poverty situation.

Environmental pollution is the most important driver of water poverty (P water poverty). Increased agricultural fertilizer and pesticide application, livestock waste, domestic sewage and industrial wastewater discharge contribute to the accelerated eutrophication of main rivers and lakes in China. Meanwhile, water scarcity and poor water quality interact with each other. The contaminated water is threatening the water availability even in some water-abundant areas, which results in the water unusable. The similarity of the trajectories of the poorest water resource availability regions and the heaviest water pollution areas in current years (shown as the darkest areas in Figs. 3, 4) might be a case in point. The social economic development is also under threat because the households, industries and agriculture are force to cut back their water use with a lack of clean water. And the cancer mortality rates related to poor water quality in China has been well above the world average (WB 2007). Each year, an estimate 190 million people fall ill and about 60,000 die from water pollution (Qiu 2011). Thus there is an urgent need to improve the public environmental protection awareness, devise sufficient incentives to the investment in the water environment protection and sewage treatment facilities, and formulate effective regulations to control the wastewater discharge and diverse pollution sources.

R water poverty characterized by lack of water resource endowment has spread from northern China to the South Coast and the Eastern Coast indicating that limited water resource availability cannot satisfy rising demand, even in the regions where ever rich in water resources. New approaches and strategies will be required to use and manage water resources more effectively. These strategies should include ways to slow the growth in water consumption, popularize the water-saving technologies, and promote the trade of

virtual water to relieve the pressure on water resource through importing water-intensive products and export water-extensive products and services.

Southwestern China has the best endowment of water resources in all study areas, but the region is suffering from water poverty characterized as C through the lack of socio-economic capacity, the weakness of water use efficiency and heavy water pollution. The policies are stressed to stimulate public and private investment in developing the water-saving, socially inclusive innovations and technologies to drive sustainable development in the regions.

It would be beneficial if further management interventions are prioritized as indicated by the study results. For instance, in the regions characterized as P–R, more sewage networks should be constructed, more effective regulations and/or incentives aimed at reducing pollution, and further improvement of the efficiency of water use should be the main priorities. Regions with P–C water poverty should focus on aquatic environmental pollution control, adoption of water-saving technologies, and economic and technological development aimed at improving water access. Regions characterized as R–C need to strengthen socio-economic capacity, improve the access to water and the efficiency of water use. The provision of safe water and hygienic facilities of sewage disposal, people's ability and capacity for sustaining access to clean water are all needed to diminish the water poverty of those regions featured by P–R–C.

To analyze the underlying complexities of water issues in China, the calculation method of WPI is revised through the transformation of centralized logarithm of initial variable and holistic and dynamic PCA. The comparison based on the traditional PCA and improved PCA shows that the proposed method achieves drastic dimension reduction. Although we use the improved holistic and dynamic PCA to assess the water poverty status in China, this study is based on the constructed WPI for Chinese case and available reliable data from the year of 2004–2012, which may impose certain limitations on the analysis of our results. For an example, other important indicators of water poverty were not captured in the study due to lack of reliable data and multi-collinearity among the variables. For instance, the variable of biodiversity is replaced by public green areas per capita, the Gini coefficient of income distribution is replaced by household's income ratio between the urban and the rural, and three variables including precipitation variance, mortality rate of children under five, and rate of soil erosion by area are left out of the analysis, due to incomplete and insufficient data. Seven variables including water resource per capita, proportion of environmental protection investment, GDP per capita, water use per 10,000 yuan GDP, water consumption per 10,000 yuan of industrial value added, percentage of investment in anti-pollution projects, and households income ratio between the urban and the rural are discarded owing to multi-collinearity among the variables. Notwithstanding the limitations, this study has produced some important findings, which might help improve the calculation method of Water Poverty Index, raise the awareness of the public and government about the aquatic environmental protection and sustainable water resource development in China, and devise better policies to alleviate the regional water poverty. It would be advantageous if future studies can identify the impacts of those omitted variables and examine the interrelationship amongst water poverty, urbanization, industrialization and human welfare. Such analysis would further help policy makers and planners to capture the possible impacts of alternative planning and development interventions in China.

References

Cho, I. D., Ogwang, T., & Opio, C. (2010). Simplifying the Water Poverty Index. *Social Indicators Research, 97*, 257–267.

Dunteman, G. H. (1989). *Principal components analysis*. Newbury Park, CA: Sage.

Feitelson, E., & Chenoweth, J. (2002). Water poverty: Towards a meaningful indicator. *Water Policy, 4*, 263–281.

Gine Garriga, R., & Perez Foguet, A. (2010). Improved method to calculate a Water Poverty Index at local scale. *Journal of Environmental Engineering, 11*, 1287–1300.

Gine Garriga, R., & Perez Foguet, A. (2011). Application of a revised Water Poverty Index to target the water poor. *Water Science and Technology, 6*, 1099–1112.

Gleick, P. (1996). Basic water requirements for human activities: Meeting basic needs. *Water International, 21*, 83–92.

Heidecke, C. (2006). Development and evaluation of a regional water poverty index for Benin. EPT Discussion Paper 145, Environment and Production Technology Division, International Food Policy Research Institute, January. http://www.ifpri.org/divs/EPTD/DP/papers/eptdp145.pdf

Jemmali, H., & Matoussi, M. S. (2013). A multidimensional analysis of water poverty at local scale: Application of improved water poverty index for Tunisia. *Water Policy, 15*, 98–115.

Jemmali, H., & Sullivan, C. (2014). Multidimensional analysis of water poverty in MENA region: An empirical comparison with physical indicators. *Social Indicator Research, 115*, 253–277.

Jiang, Y. (2009). China's water scarcity. *Journal of Environmental Management, 90*, 3185–3196.

Johnson, R. A., & Wichern, D. W. (2007). *Applied multivariate statistical analysis [M]* (6th ed.). Upper Saddle River: Pearson Education Incorporation, publishing as Prentice Hall.

Jolliffe, I. T. (1986). *Principal component analysis*. New York: Springer.

Komnenic, V., Ahlers, R., & Zaag, P. (2009). Assessing the usefulness of water poverty index by applying it to a special case: Can one be water poor with high levels of access? *Physics and Chemistry of the Earth, 34*, 219–224.

Lawrence, P., Meigh, J., Sullivan, C. (2002). *The water poverty index: An international comparison*. Keele Economics Research Paper, 19. Staffordshire, UK.

Liu, Z., Wang, X., & Chen, X. (2012). The influence of data transformation to extracting principal component. *Science Technology and Engineering, 12*(6), 1348–1351. (**In Chinese**).

Molle, F., & Mollinga, P. (2003). Water poverty indicators: Conceptual problems and policy issues. *Water Policy, 5*(5), 529–544.

Nardo, M., et al. (2005). *Handbook on constructing composite indicators: Methodology and user guide*. OECD Statistics Working Papers.

Nardo, M., Saisana, M., Saltelli, A., Tarantola, S. (2005). Tools for composite-indicator building. In: *Institute for the Protection and Security of the Citizen, European Commission*

National Bureau of Statistics, Ministry of Environmental Protection. China Statistical Yearbook on Environment. (2005–2013). China Statistic Press, Beijing, China (in Chinese). http://www.stats.gov.cn/ staticreq.htm

National Bureau of Statistics of China. China Statistical Yearbook. (2005–2013). China Statistic Press, Beijing, China (in Chinese). http://www.stats.gov.cn/ easyquery.htm?cn=E0103

OECD (Organization for Economic Cooperation and Development). (2007). *Environmental performance reviews: China*. Paris, France: OECD.

Perez Foguet, A., & Gine Garriga, R. (2011). Analyzing water poverty in basins. *Water Resources Management, 25*(14), 3595–3612.

Qiu, J. (2011). China to spend billions cleaning up groundwater. *Science, 334*, 745.

Rijsberman, F. R. (2006). Water scarcity: Fact or fiction? *Agricultural Water Management, 80*(1–3), 5–22.

SC (State Council, P.R. China). (2006). *The 11th Five-Year National Plan for Social and Economic Development*. Beijing, China: People's Publishing House.

Sullivan, C. (2002). Calculating a water poverty index. *World Development, 30*(7), 1195–1210.

Sullivan, C. (2005). *Method to develop and describe community level water poverty index scores*. UK: CEH Wallingford.

Sullivan, C., Meigh, J., et al. (2003). The Water Poverty Index: Development and application at the community scale. *Natural Resources Forum, 27*, 189–199.

Sullivan, C., Meigh, J., et al. (2007). Integration of the biophysical and social sciences using an indicator approach: Addressing water problems at different scales. *Water Resources Management, 21*(1), 111–128.

Todd, H. (2010). China's water scarcity and its implications for domestic and international stability. *Asian Affairs: An American Review, 37*, 71–83.

WB (World Bank). (2007). *Water pollution emergencies in China*. Washington, DC, USA: World Bank.

Ye, S. (2001). The application and consideration about principal component analysis. *Data Statistics and Management, 20*(2), 52–56. **(in Chinese)**.

Zhang, C., & Chen, S. (1996). The principal component analysis on component data and its applications. *Mathematical Statistics and Management, 15*(3), 11–15. **(in Chinese)**.

Zhou, O., Ang, B. W., & Poh, K. L. (2006). Comparing aggregation methods for constructing the composite environmental index: An objective measure. *Ecological Economics, 59*(3), 305–311.