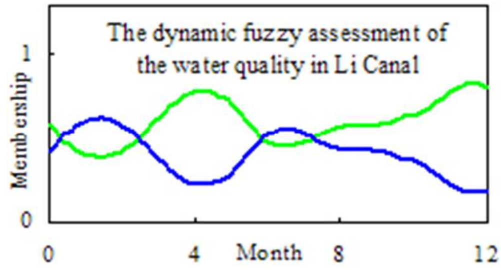

87x46mm (72 x 72 DPI)

This study initially establishes a functional fuzzy synthetic evaluation (FFSE) method to make dynamic fuzzy assessment of water quality.

In this article, we think that the water quality condition, which can be generalized to other environmental processes or impacts, is a dynamic fuzzy process. In addition, we think that missing values and different sampling times are common troubles in the conventional environmental processes or impacts research. To measure this dynamic fuzzy process more generally and intuitively, and to deal with the observation with missing values and different sampling times, we initially introduce functional data analysis (FDA) theory into the conventional fuzzy synthetic evaluation (FSE) model to generate the functional fuzzy synthetic evaluation (FFSE) method in this study.

Water quality assessment of Li Canal using functional fuzzy synthetic evaluation model

Yan Feng, Liu Ling

(State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing, 210098, China)

Abstract: Through introducing functional data analysis (FDA) theory into the conventional fuzzy synthetic evaluation (FSE) method, the functional fuzzy synthetic evaluation (FFSE) model is established. FFSE keeps the property of the conventional FSE that the fuzziness in the water quality condition can be well measured. Furthermore, compared with FSE, FFSE has the following advantages: (1) FFSE requires much less conditions in observation, for example, pollutants can be monitored at different time, and data missing is accepted; (2) the dynamic variation of the water quality condition can be represented more comprehensively and intuitively. The procedure of FFSE is discussed and the water quality of Li Canal in 2012 is evaluated as an illustration. The synthetic classification of Li Canal is “II” in January, February and July, and “I” in other months, which can satisfy the requirement of the Chinese South-to-North Water Diversion Project.

Key words: water quality assessment; functional data analysis; functional fuzzy synthetic evaluation; Li Canal

1. Introduction

The South-to-North Water Diversion Project is a strategic measure in the optimization of water resources distribution of China [1]. There are three diversion routes: Eastern Route Project (ERP), Middle Route Project (MRP), and Western Route Project (WRP) [1]. ERP diverts water from the Changjiang River, the longest river in China, and supplies water for the eastern Huang-Huai-Hai Plain [1]. ERP is the first diversion route to be completed, and it plays an important role in the social development and ecology protection in Jiangsu, Anhui, Shandong, Hebei and Tianjin.

The trunk diversion channel of ERP is Beijing-Hangzhou Grand Canal, and the key problem in ERP is pollution control [1]. Li Canal, which is a section of Beijing-Hangzhou Grand Canal, is the initial diversion interval of ERP. Therefore the water quality assessment of Li Canal is crucial for the management of ERP.

The water pollution situation is a vague concept, for the imprecision and fuzziness in classification criteria are inevitable [2]. Therefore, fuzzy synthetic evaluation (FSE), has been widely used in recent literatures. Lu firstly introduced FSE into the water quality assessment of Fei-Tsui Reservoir in 1999 [3]; Chang made a fuzzy classification of the water quality in Twseng-Wen River in 2001 [4]; Lu made a further research of Fei-Tsui Reservoir with FSE and self-organizing maps in 2002 [5]; William and Zou assessed the environment condition of Ebro River and Gorges Reservoir using fuzzy sets in 2006 respectively [6,7]; Yilmaz made a fuzzy classification of the Eber Lake in 2007 [8]; Wang evaluated the water quality of Naoli River with FSE and parameter correlation analysis in 2008 [9]; Andre improved the conventional water quality index method with FSE and proposed a fuzzy water quality index in 2009 [10]; Lu assessed the environment of Wei River using FSE in 2010 [11]. FSE represents the water quality with membership vectors or matrixes, thus fuzzy uncertainties can be well described [2-11].

In the conventional FSE, observed values are of time series, as a result, membership vectors and matrixes are discrete. This method often incorporates some shortcomings such as: (1) the monitoring time of each indicator needs to be identical [12]; (2) no missing values should occur [12]; (3) it is difficult to capture the dynamic changing rules with discrete vectors or matrixes series.

Functional data analysis (FDA), a new statistics theory emerged in 1980s, provides a way to solve the problems described above. In FDA, data are formatted into continuous function curves instead of discrete values; therefore the time-varying process of pollutants can be described quantitatively and intuitively [12-14]. Furthermore, FDA requires less condition in data, for example, the indicator's monitoring time series can be different from each other, and data missing is accepted [12-14]. Therefore, FDA has been widely used in many statistics problems in iatrology, economy and ecology [14]. In the water environment research, Champely [12], Henderson [15], and Haggarty [16] have successfully taken FDA in place of the conventional multivariate method into the water quality trends analysis.

In this paper, we introduce the FDA theory into the FSE method to establish a new fuzzy assessment model – functional fuzzy synthetic evaluation model. And as an illustration, we apply this model in the water quality assessment of Li Canal.

2. Materials and methods

2.1 Study area

Li Canal, the initial diversion interval of ERP, is an artificial canal with two thousand years history. Its origin is Sanjiangying (119°42’53”E, 32°18’45”N), a village located in the north bank of Changjiang River; and terminal is Yangzhuang village (118°56’02”E, 33°33’27”N), which is the demarcation point of Li Canal and Zhong Canal. In ERP, approximately 8.9 billion m^3 of water is diverted from Changjiang River and supplied through Li Canal to Northern China every year [1]. The length of Li Canal is 141.8 km, and it flows through two cities: Yangzhou and Huaian. The major purposes of Li Canal are water diversion, navigation and supplying municipal, industrial and agricultural water for cities along the river. To monitor the concentrations of pollutants, 7 observation sites are laid along the river (Fig1).

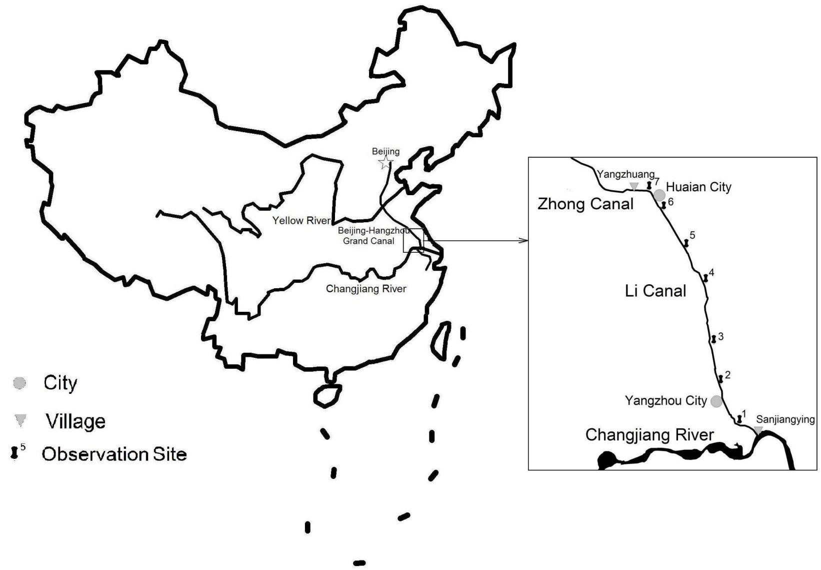

Fig1. Map of the study area

In the national surface water environmental quality standards of China, the water quality is classified into five classifications {I, II, III, IV, V} [17], which corresponds to the criterion set {Excellent, Good, Ordinary, Poor, Bad}. According to purpose of the water body, a minimum allowable classification is designed in the Chinese environment protection. For example, the water quality should reach classification "I" in the national nature reserve; if the river is drinking water sources; its water quality condition is often required to be no less classification "II" or "III"; if the water is just supplied for industry, its minimum allowable classification is "IV"; and if the water is just supplied for agriculture, its minimum allowable classification is "V" [17]. In planning of ERP, the water quality condition of Li Canal is required to be no less than classification "III" throughout 2012 [1].

According to the research of Gao Mingyuan [18], the major pollutants in Li Canal are oxygen-consuming pollutants and ammonia nitrogen. Industrial, domestic and agricultural non-point source pollution along the river had significant effects on the water quality. In 2008, DO, COD_{Mn} and NH_3-N could not stratify the requirement of the water quality protection in the flood period; while the other pollutants could meet the demand of protection. Therefore we take {DO, COD_{Mn}, NH_3-N} as the indicator vector in this study. The criterions and threshold values of assessment variables are listed in table 1.

Table 1. National surface water environmental quality standards of China

Classification	Criterion	DO (mg/L)	COD_{Mn} (mg/L)	NH_3-N (mg/L)
I	Excellent	>7.5	<2	<0.015
II	Good	6–7.5	2–4	0.015–0.5
III	Ordinary	5–6	4–6	0.5–1
IV	Poor	3–5	6–10	1–1.5
V	Bad	2–3	10–15	1.5–2

2.2 Fuzzy synthetic evaluation (FSE) model

In fuzzy synthetic evaluation model, the information of indicators is represented by the membership vector $\mathbf{u}_i = \{u_{i1}, u_{i2}, u_{i3}, u_{i4}, u_{i5}\}$, which is generated from observed data through membership functions [2-10].

The forms of membership function can be formulated according to the condition of the study area and thresholds. For example, trapezoidal membership function is used in the fuzzy analysis of Tseng-Wen River [4], Ebro River [6], Eber Lake [8], Three Gorges Reservoir [7]; Lu designed a nonlinear membership function in the research of Fei-Tsui Reservoir [3,5].

There is little literature on which membership function should be chosen in Li Canal. But in the water quality fuzzy assessment of Changjiang, which is the major water source of Li Canal, trapezoidal membership function is widely used and proved to be suitable [19, 20]. Therefore we chose trapezoidal membership function to measure the fuzziness in assessment variables in this study.

For the indicator which is the smaller the better:

$$u_{i1} = \begin{cases} 1 & x_i \le c_{i1} \\ \dfrac{c_{i2} - x_i}{c_{i2} - c_{i1}} & c_{i1} \le x_i < c_{i2} \\ 0 & x_i > c_{i2} \end{cases} \tag{1}$$

$$u_{ij}=\begin{cases}0 & x_i \le c_{i(j-1)} \\ \dfrac{x_i-c_{ij}}{c_{i(j-1)}-c_{ij}} & c_{i(j-1)} \le x_i < c_{ij} \\ \dfrac{c_{i(j+1)}-x_i}{c_{i(j+1)}-c_{ij}} & c_{ij} \le x_i < c_{i(j+1)} \\ 0 & x_i > c_{i(j+1)}\end{cases} \quad j=2,3,4 \tag{2}$$

$$u_{i5}=\begin{cases}0 & x_i \le c_{i4} \\ \dfrac{x_i-c_{i4}}{c_{i5}-c_{i4}} & c_{i4} < x_i \le c_{i5} \\ 1 & x_i > c_{i5}\end{cases}. \tag{3}$$

In formula (1) ~ (3), u_{ij} is the membership; c_{ij} is the threshold of the *j*th classification for the *i*th indicator.

For the indicator which is the bigger the better, just replace the "$<$", "$\le$" and "$>$" in formula (1) ~ (3) with "$>$", "$\ge$" and "$<$", respectively.

The membership matrix $\mathbf{U}$ is derived from synthesizing the membership vectors of indicators [2-10]:

$$\mathbf{U}=\begin{pmatrix}\mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3\end{pmatrix}=\begin{pmatrix}u_{11} & u_{12} & \cdots & u_{15} \\ u_{21} & u_{22} & \cdots & u_{25} \\ u_{31} & u_{32} & \cdots & u_{35}\end{pmatrix}. \tag{4}$$

The comprehensive condition of the water quality is represented by the synthetic membership vector $\mathbf{v}=\{v_1,v_2,v_3,v_4,v_5\}$, which is got by:

$$\mathbf{v}=\boldsymbol{\omega}\mathbf{U}. \tag{5}$$

Here $\boldsymbol{\omega}=\{\omega_1,\omega_2,\omega_3\}$ is the weight vector.

And the classification is determined by the largest membership principle [7]:

If

$$v_j=\max\{v_1,v_2,v_3,v_4,v_5\}, \tag{6}$$

the classification is *j*.

There are two disadvantages in conventional FSE model in dealing with dynamic assessment problems:

(1) Too many conditions are required in observation, for example, sampling time

series should be identified for every indicator and no missing value should occur.

If pollutants are monitored at different time, the membership vectors of indicators will represent the water quality conditions at different time point, and the membership matrix will be meaningless for it is impossible to identify which time matrix correspond to. If the value of the *i*th pollutant is missing, the membership matrix will be invalid for the losing of the *i*th membership vector. As a result, to keep the membership matrix meaningful, sampling time series should be identified for every indicator and no missing value should occur in FSE.

(2) The result of FSE can not represent the variation of the water quality comprehensively or intuitively.

The concentrations of pollutants are varied continuously over a time domain, but the result of FSE is a discrete membership matrix or vector series, which only represents water quality situations at finite time points. Furthermore, the intuitiveness of discrete vector or matrix series is quite poor; therefore it is often difficult to capture the dynamic rules of the water quality.

2.3 Functional Data Analysis (FDA)

FDA is a new statistic theory with a basic idea that the information is represented in form of functional curves [14]. In recent years, FDA is increasingly being used to better analyze, model and predict dynamic variables [14]. Compared with the conventional discrete data analysis methods, FDA approach has the following advantages [13. 14]:

(1) Many objectives of analysis, such as concentrations of pollutants, are continuously varied with time in nature. Therefore it is more intuitive, actual and comprehensive to represent these dynamic processes with continuous curves than discrete values.

(2) FDA requires much less conditions in observation. For example, missing values and different sampling times, which often cause invalid or meaningless calculation in conventional discrete data analysis methods, are accepted in FDA.

(3) FDA provides many distinctive tools in research, such as derivative analysis, which can extract additional information contained in dynamic curves.

The first step in FDA is to generate function curves from discrete monitoring values [13-15]. If the sampling is well controlled and concentrations of pollutants are accurately measured, interpolation suffices in curves generation [15]. But in most cases, monitoring errors are inevitable; therefore smoothing is the most common method to convert the raw discrete data with noises and errors into smoothly varying function in literatures [14].

The basic idea of smoothing is to represent the smooth curve over time t through linear combinations of basis functions [15]. That is,

$$x_i(t) = \sum_{m=1}^{M} d_{im}\varphi_m(t) \quad i = 1,2,3\,. \tag{7}$$

Here $x_i(t)$ is the concentration curve of the ith pollutant; $\varphi_m(t)$ is basis function; d_{im} is the coefficient.

The choice of basis function is dependent on the underlying behavior of the data being researched [13-15]. For example, Fourier basis are traditional used when the data is periodic; B-splines are often chosen to analysis nonperiodic data; and wavelet bases is used when there are discontinuities and/or rapid changes in behaviors of data [13-15].

In the literatures of FDA, the least square principle is often used to generate coefficients d_{im} [15].

$$\min\left\{\sum_{k=1}^{K_i}\left(\sum_{m=1}^{M} d_{im}\varphi(t_{ik}) - x_i(t_{ik})\right)^2\right\} \tag{8}$$

Here t_{ik} is the time of the kth observation for the ith indicator; K_i is the total number of times of observation for the ith indicator.

Note the sampling time series $\{t_{i1}, t_{i2}, \cdots t_{iK_i}\}$ as t_i^*. t_i^* and K_i are not required to be consistent for every I in formula (7) ~ (8). There fore the sampling time series not necessarily to be same for every indicator and data missing is accepted.

If there are multiple observation sites in the research section, the concentration curves of pollutants are generated by

$$\bar{x}_i(t) = \frac{\sum_{m=1}^{M} l_m \cdot x_{im}(t)}{\sum_{m=1}^{M} l_m} \quad i = 1, 2, 3. \tag{9}$$

Here M is the amount of the observation sites; l_m is the representative length of the mth site; $x_{im}(t)$ is the concentration curve of the ith pollutant at the mth site.

Furthermore, the monitoring time series for each site are not necessary to be same either.

The uncertainty in concentration curves is introduced by two factors: how accurate the observation is and how well the concentration curve fits the observation data. To minimize this uncertainty, the following procedure is suggested be adopted in concentration curves generation:

(1) If the error in observation is regarded to as negligible, the concentration curve should be generated by interpolation, the result of which passes through all observation data accurately [15].

(2) If the error in observation is non-negligible, while the roughness of the raw data set is not great, the smoothing method based on least square principle to calculate coefficients should be adopted to generate curves with a good smoothness and a minimum mean-squared error [15].

(3) If the roughness and noises in the raw data set is great, a penalty operator should be added in the smoothing formula to keep the smoothness of the curve and reduce the effect of outliers [13-15].

In this study we concentrate on establishing the functional fuzzy synthetic evaluation model though combining FDA method and the conventional FSE model. More theoretical content of FDA can be acquired in the works of Ramsay [13] and Ulah [14]; and some applications of FDA in water quality trend analysis are illustrated in the articles of Champely [12], Henderson [15] and Haggarty [16].

2.4 Functional fuzzy synthetic evaluation (FFSE) model

Introducing FDA method into fuzzy synthetic evaluation theory, we establish the Functional fuzzy synthetic evaluation (FFSE) model.

The first step in FFSE is to generate the concentration curve $x_i(t)$ from discrete monitoring values according to formula (7) and (8).

If there are multiple observation sites in the research section, the average concentration curves of pollutants in the research section are generated according to formula (9)

Then replace the x_i with $x_i(t)$ in formula (1)~(3) and generalize the membership vector curves $\mathbf{u_i}(t)=\{u_{i1}(t),u_{i2}(t),\cdots,u_{iN}(t)\}$

The membership $u_{ij}(t)$ is the function of time in FFSE, which represents the dynamic process of the *i*th indicator's approaching degree to the *j*th classification.

The calculation method for membership matrix curves **U(*t*)** is:

$$\mathbf{U}(t)=\begin{pmatrix}\mathbf{u}_1(t)\\ \mathbf{u}_2(t)\\ \mathbf{u}_3(t)\end{pmatrix}=\begin{pmatrix}u_{11}(t) & u_{12}(t) & \cdots & u_{15}(t)\\ u_{21}(t) & u_{22}(t) & \cdots & u_{25}(t)\\ u_{31}(t) & u_{32}(t) & \cdots & u_{35}(t)\end{pmatrix}. \quad (10)$$

The synthetic membership vector curves $\mathbf{v}(t)=\left\{v_1(t),v_2(t),v_3(t),v_4(t),v_5(t)\right\}$ are generated by:

$$\mathbf{v}(t)=\boldsymbol{\omega}\mathbf{U}(t). \quad (11)$$

The expression of the membership vector in FFSE is a family of curves, which is more intuitively than vector series in the dynamic research of water quality.

3. Result and discussion

In this section, the water quality condition of Li Canal in 2012 is assessed based on the FFSE model established in section 2.4, and a comparison between the results of FFSE and the conventional FSE is briefly discussed.

3.1 Concentration curves

The first step in FFSE is to generate concentration curves of pollutants at every observation site. The observation data set of Li Canal, which contains the monitoring data at the beginning and ending of 2012 and the average values in every month (the data of COD_{Mn} in April is missing), are provided by the hydrology office of Jiangsu province in China .

In this study, the basis function is chosen as 4-orded B-spline basis, for there is no periodic phenomenon in the intra-annual variation of pollutants. But if the research object is the inter-annual variation of pollutants, Fourier basis may be a better choice.

With the smoothing method discussed in section 2.3, the concentration curves of pollutants are generated. And as an illustration, concentration curves of pollutants at the fifth monitoring site are exhibited in Fig 2.

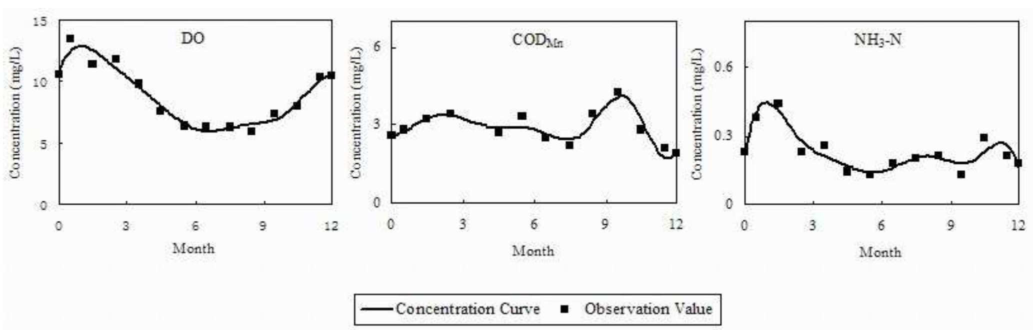

Fig 2. Concentration curves of pollutants at the fifth monitoring site

As is shown in Fig 2, the discrete data are converted into concentration curves over a same time domain, though the observation value of COD_{Mn} in April is missing.

In the smoothing method, monitoring errors are considered as inevitable and the optimal fitting curve is regarded as one with good smoothness and minimal squared distance to the observation values [13, 15]. Therefore concentration curves generated by smoothing method generally do not pass through all observation data. But this difference is not significant in most cases.

When sampling is well controlled and concentrations of pollutants are accurately measured, interpolation could be chosen to convert concentration curves too [17]. The curves generated by interpolation method can pass through every observation value accurately. But if there is error in the raw monitoring data, the smoothness of these curves is often poor; therefore interpolation is rarely used in FDA literatures.

To extract more information, the water quality conditions in Yangzhou section, Huaian section and the whole river of Li Canal are assessed separately. The concentration curves of pollutants in these three research objects are generated according to formula (9) and exhibited in Fig3.

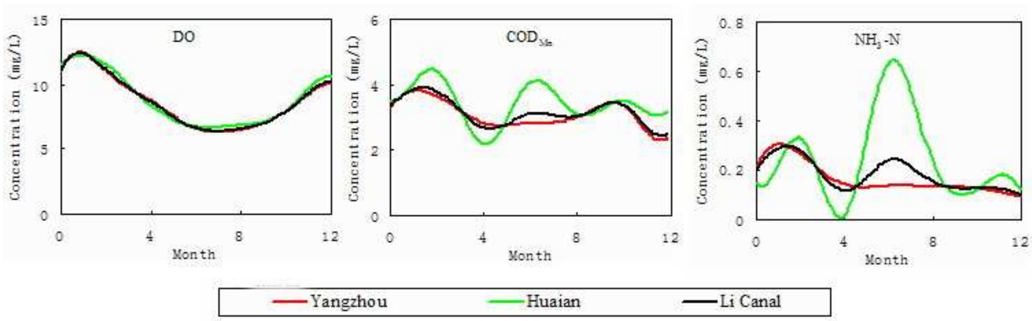

Fig 3. The concentration curves of pollutants in Yangzhou section, Huaian section and Li Canal

Note the DO concentration curves in Yangzhou, Huaian and Li Canal as $x_{1Y}(t)$, $x_{1H}(t)$ and $x_{1L}(t)$ respectively. As is shown in Fig 3, all of $x_{1Y}(t)$, $x_{1H}(t)$ and $x_{1L}(t)$ show similar traces, which is decreasing from January to July and increasing from August to December. This phenomenon is caused by the variation of temperature in the basin, which rises before July and decreases after August.

Note the COD_{Mn} and NH_3-N concentration curves in Huaian as $x_{2H}(t)$ and $x_{3H}(t)$ respectively. As is exhibited in Fig 3, both of $x_{2H}(t)$ and $x_{3H}(t)$ show similar fluctuation tends, which is induced by the agricultural non-point source pollution. The agricultural operating system in Huaian is rice-wheat rotation. In the wheat growth period, January and February is the winter top dressing period; April is the panicle fertilizer period; and May is the harvest time. In the rice growth period, July is the panicle fertilizer period; September is the granular fertilizer period; and

October is the harvest time. In these months, the redundant fertilizer and the rotten straws after harvest are discharged into the river with the agricultural runoff; as a result, COD_{Mn} and NH_3-N concentrations rise in these periods.

Note the COD_{Mn} and NH_3-N concentration curves in Yangzhou as $x_{2Y}(t)$ and $x_{3Y}(t)$ respectively. January is the rising period for both $x_{2Y}(t)$ and $x_{3Y}(t)$, because the dilution and degradation effects is poor in this month for a low temperature and water level. $x_{2Y}(t)$ shows another increasing trend from June to August, and this is majorly caused by the industrial and domestic oxygen-consuming pollutants discharged into the river with the urban flood in these months, which is the rainy season in the basin. $x_{3Y}(t)$ shows decreasing trends in the other months, which is quite different from that in Huaian. Furthermore, in most period of 2012, $x_{2Y}(t)$ and $x_{3Y}(t)$ are much less than $x_{2H}(t)$ and $x_{3H}(t)$ respectively. The reason for these phenomenons is that the water level of Li Canal is generally higher than the farms along the canal in Yangzhou, thus the agricultural non-point source pollution is not as serious that in Huaian.

Note the average COD_{Mn} and NH_3-N concentration curves in the Li Canal as $x_{2L}(t)$ and $x_{3L}(t)$ respectively. As is shown in Fig, the dynamic pattern of $x_{2L}(t)$ and $x_{3L}(t)$ are very similar with that of $x_{2H}(t)$ and $x_{3H}(t)$, indicating that the variation of pollutants in Huaian plays the crucial role in the synthetic water quality condition of Li Canal.

3.2 Membership curves of pollutants

The second step in FFSE is to generate membership curves of pollutants. With the method discussed in section 2.4, these membership curves are converted and exhibited in Fig 4.

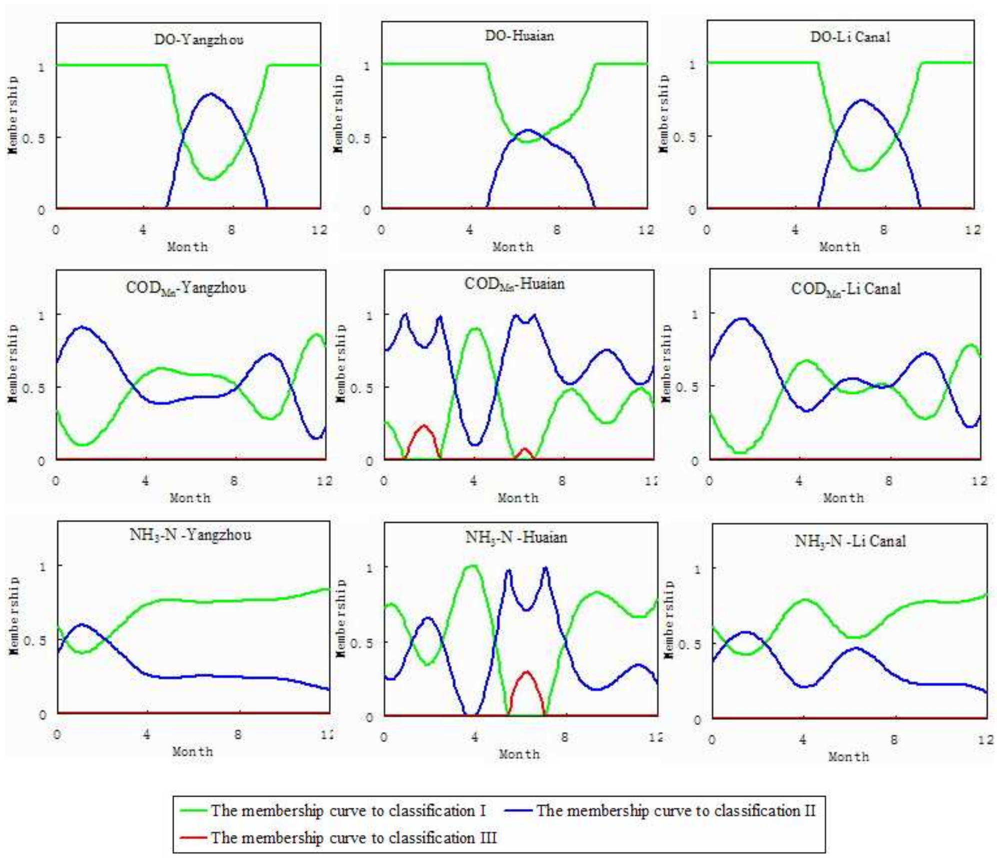

Fig 4. Membership curves of pollutants.

Note the membership curve of the ith indicator to the jth classification in Yangzhou, Huaian and Li Canal as $u_{ijY}(t)$, $u_{ijH}(t)$ and $u_{ijL}(t)$ respectively.

As is shown in Fig 4, the memberships of DO to classification "III", "IV" and "V" are 0 through out the year in Yangzhou, Huaian and Li Canal. All of $u_{11Y}(t)$, $u_{11H}(t)$ and $u_{11L}(t)$ show a similar variation principle, which is decreasing from May to July, increasing from August to October, and equaling 1 in the other months. From May to October, the variations of $u_{ijY}(t)$, $u_{ijH}(t)$ and $u_{ijL}(t)$ are induced by the changes of $x_{1Y}(t)$, $x_{1H}(t)$ and $x_{1L}(t)$. However, $x_{1Y}(t)$, $x_{1H}(t)$ and $x_{1L}(t)$ are more than 7.5 in other months, therefore the conditions of DO absolutely belong to classification "I" in this period. Because the sum of the all memberships equals 1, $u_{12Y}(t)$, $u_{12H}(t)$ and $u_{12L}(t)$ show a contrary variation principle with $u_{11Y}(t)$, $u_{11H}(t)$ and $u_{11L}(t)$ respectively.

As is illustrated in Fig 4, the memberships of COD_{Mn} and NH_3-N to classification “III”, “IV” and “V” are 0 through out the year in both Yangzhou and Li Canal. The variations of $u_{21Y}(t)$, $u_{21L}(t)$, $u_{31Y}(t)$ and $u_{31L}(t)$ are in correspondence with $x_{2Y}(t)$, $x_{2L}(t)$, $x_{3Y}(t)$ and $x_{3L}(t)$ respectively, while the variation trends of $u_{22Y}(t)$, $u_{22L}(t)$, $u_{32Y}(t)$ and $u_{32L}(t)$ are just contrary. This is caused by the concentration variations of COD_{Mn} and NH_3-N. As is illustrated in Fig 3, $x_{2Y}(t)$ and $x_{2L}(t)$ are varied in the domain [2,4], and $x_{3Y}(t)$ and $x_{3L}(t)$ are varied in the domain [0.015,0.5]. In this case, the memberships of these two pollutants to classification “I” are increasing functions to their concentrations , while the memberships to classification “II” are decreasing functions to concentrations according to formula (1)~(3).

The memberships of COD_{Mn} and NH_3-N to classification “IV” and “V” are 0 through out the year in Huaian. As is exhibited in Fig 4, in March, August and November, which is not the fertilizer or harvest period, $u_{21H}(t)$ and $u_{31H}(t)$ increase while $u_{22H}(t)$ and $u_{32H}(t)$ decrease with the concentration reducing of COD_{Mn} and NH_3-N. In January, May, September and October, $u_{21H}(t)$ and $u_{31H}(t)$ decrease while $u_{22H}(t)$ and $u_{32H}(t)$ increase with the concentration rising of COD_{Mn} and NH_3-N. Furthermore, according to the trend of $u_{23H}(t)$ and $u_{33H}(t)$, February and July are the worst time of COD_{Mn} and NH_3-N respectively.

What should be noticed is that none of three indicators belongs to classification “IV” or “V” at any time in 2012, indicating that all indicators can satisfy the requirement of the water quality protection. This is different from the research of Gao Mingyuan in 2008[18], and the reason is that lots of sewage treatment regulations and projects have been developed in these years, which reduces the concentrations of pollutants significantly.

3.3 Synthetic membership curves

The third step in FFSE is to generate synthetic membership curves. According to the method discussed in section 2.4 and the weight vector {0.2, 0.4, 0.4}, which is generated based on AHP method [21], the synthetic membership curves are calculated and exhibited in Fig3.

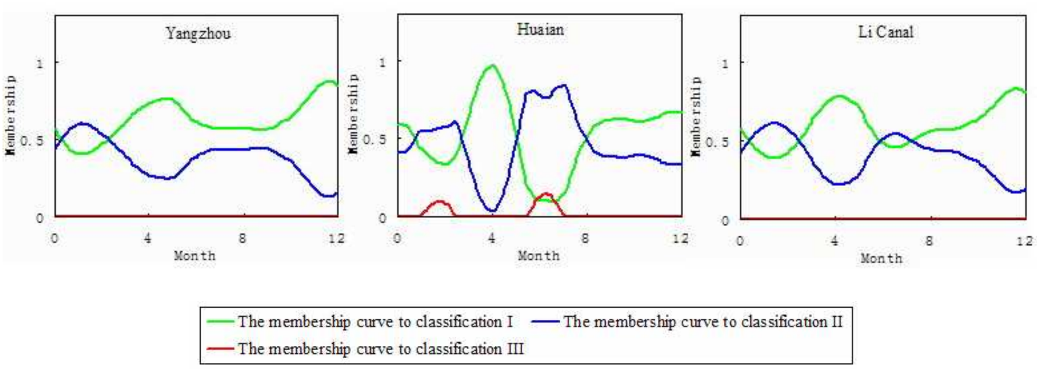

Fig 5. The synthetic membership curves

Note the synthetic membership curve of to the *j*th classification in Yangzhou, Huaian and Li Canal as $v_{jY}(t)$, $v_{jH}(t)$ and $v_{jL}(t)$ respectively. As is shown in Fig 5, the classification of Li Canal in Yangzhou is "II" from January to March and "I in other months. The ameliorating periods are from February to May for the decreasing of COD_{Mn} and NH_3-N, and from September to December for the increasing of DO.

From February to March and from May to August, the classification of Li Canal in Huaian is "II", and in other months, the classification is "I". The deteriorating periods are from January to February and from April to May, when the concentration of COD_{Mn} and NH_3-N rises and DO reduces.

The synthetic classification of Li Canal is "II" in January, February and July, and "I" in other months, indicating that the water quality condition in Li Canal could satisfy the requirement of the water quality protection in ERP in 2012. The synthetic dynamic rule of the water quality in Li Canal is in correspondence with that in Huaian.

3.4 Comparison between FFSE and FSE

Using the conventional FSE, the water quality of Li Canal is assessed and exhibited in Fig 6.

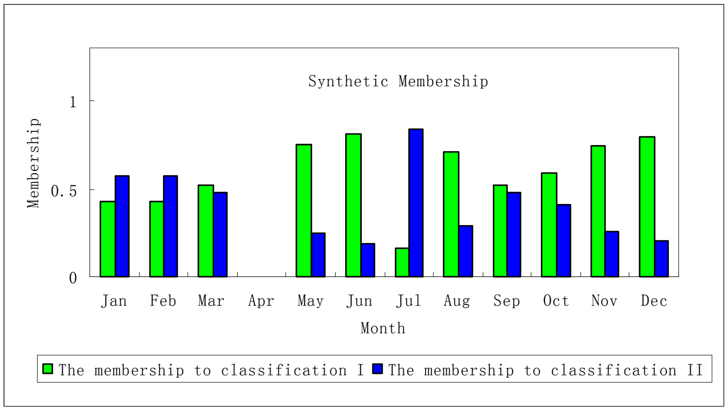

Fig 6. The assessment result of Li Canal based on FSE

As is illustrated in Fig 6, the water quality condition in April can not be assessed for the missing of COD_{Mn} values, while in other months, the classifications identified by FSE and FFSE are same. Moreover, membership vector series of FSE is much less intuitively than the membership curves of FFSE, for too many membership values of different classifications are crowed in a single figure.

As is discussed above, the FFSE established in this study can be regarded as the generalization of the conventional FSE, and FSE can be taken as a case of FFSE at a special time point. FFSE keeps the property of the conventional FSE that the fuzziness in the water quality condition can be well measured. Furthermore, compared with FSE, FFSE has the following advantages:

(1) FFSE requires fewer conditions in observation, for example, different sampling time and missing value are accepted.

In FFSE, the first step is to convert raw discrete data into continuous curves over the same time domain, no matter which sampling time series is adopted. Therefore there will not be invalid generation in the following calculation of membership matrix or synthetic membership vector caused by the inconsistence of time or missing of value.

(2) FFSE represents the variation of the water quality more comprehensively and intuitively.

The result of FFSE is membership curves instead of discrete values; therefore the dynamic rules of water quality condition, which is varied continuously in nature, can be represented comprehensively. Furthermore, the curve graph is much concise than the bar graph, especially when the sampling time series is long, thus the result of FFSE is more intuitively.

4. Conclusion

The functional fuzzy synthetic evaluation (FFSE) model is established through introducing functional data analysis (FDA) theory into the conventional fuzzy synthetic evaluation (FSE) method.

There are three steps in functional fuzzy synthetic evaluation model: (1) generating concentration curves of pollutants; (2) calculating the membership curves of pollutants; (3) generating the synthetic membership curves.

FFSE can be regarded as the generalization of the conventional FSE, and FSE can be taken as a case of FFSE at a special time point. FFSE keeps the property of the conventional FSE that the fuzziness in the water quality condition can be well measured. Compared with FSE, FFSE has the following advantages: (1) FFSE requires much less conditions in observation, for example, pollutants can be monitored at different time, and data missing is accepted; (2) the dynamic variation of the water quality condition can be represented more comprehensively and intuitively.

The dynamic water quality situation of Li Canal is assessed using FFSE, and the result shows that from February to March and from May to August, the classification is "II", and in other months, the classification is "I", which can satisfy the requirement of ERP. The deteriorating periods are from January to February and from April to May, when the concentration of COD_{Mn} and NH_3-N rises and DO reduces. To further improve the water quality condition, more regulations and projects have been developed to control the agricultural non-point source pollution in Huaian.

Acknowledgment: This work is supported by the National Natural Science Foundation of China under the contract No. 51279060.

References

1. Office of the South-to-North water Diversion Project Commission of the State Council. General Layout of South-to-North Water Diversion. 2001.
2. Lu X W, Li L Y, Lei K, Wang L J, Zhai Y X, Zhai M. Water quality assessment of Wei River, China using fuzzy synthetic evaluation. Environmental Earth Sciences, 2010, 60(8): 1693–1699.
3. R-S Lu, S-L Lo, J-Y Hu. Analysis of Reservoir Water Quality Using Fuzzy Synthetic Evaluation. *Stochastic Environmental Research and Risk Assessment,* 13 (1999), 327-36.
4. Ni-Bin Chang, Ho-Wen Chen, Shu-Kuang Ning. Identification of River Water Quality Using the Fuzzy Synthetic Evaluation Approach. *Journal of environmental management,* 63 (2001), 293-305.
5. Ruei-Shan Lu, Shang-Lien Lo. Diagnosing Reservoir Water Quality Using Self-Organizing Maps and Fuzzy Theory. *Water Research,* 36 (2002), 2265-74.
6. William Ocampo-Duque, Nuria Ferre-Huguet, José L Domingo. Assessing Water Quality in Rivers with Fuzzy Inference Systems: A Case Study. Environment International, 32 (2006), 733-42.
7. Zhi-Hong Zou, Yi Yun, Jing-Nan Sun. Entropy Method for Determination of Weight of Evaluating Indicators in Fuzzy Synthetic Evaluation for Water Quality Assessment. Journal of Environmental Sciences, 18 (2006), 1020-23.
8. Yilmaz Icaga. Fuzzy Evaluation of Water Quality Classification. Ecological Indicators, 7 (2007), 710-18
9. Jianhua Wang, Xianguo Lu, Jinghan Tian. Fuzzy Synthetic Evaluation of Water Quality of Naoli River Using Parameter Correlation Analysis. *Chinese*

Geographical Science, 18 (2008), 361-68.

10. André Lermontov, Lídia Yokoyama, Mihail Lermontov. River Quality Analysis Using Fuzzy Water Quality Index: Ribeira Do Iguape River Watershed, Brazil. *Ecological Indicators,* 9 (2009), 1188-97.
11. Zuzeng Peng, Yunyu Sun. Fuzzy Maths and Application. *Publishing Company of WuHan University in Chinese* (2002), 7-16.
12. Stephane Champely, Sylvain Doledec. How to Separate Long-Term Trends from Periodic Variation in Water Quality Monitoring. *Water Research,* 31 (1997), 2849-57.
13. Ramsay J.O, Silverman B.W. Applied Functional Data Analysis: Methods and Case Studies. Springer, New York. 2002.
14. Shahid Ullah, Caroline F Finch. Applications of Functional Data Analysis: A Systematic Review. *BMC medical research methodology,* 13 (2013), 43.
15. Brent Henderson. Exploring between Site Differences in Water Quality Trends: A Functional Data Analysis Approach. *Environmetrics,* 17 (2006), 65-80.
16. RA Haggarty, CA Miller, EM Scott. Functional Clustering of Water Quality Data in Scotland. *Environmetrics,* 23 (2012), 685-95.
17. Chinese Environmental Protection Agency. National surface water environmental quality standards of China (GB3838-2002). China Standards Press, Beijing. 2002
18. M-Y Gao. Analysis of Water Quality in Water-Receiving Areas in Jiangsu Province on Eastern Route of South-to-North Water Diversion Project. *Water Resources Protection,* 28 (2012), 64-66.
19. Shou-yu Chen, Ya-wei Li. Water Quality Evaluation Based on Fuzzy Artificial Neural Network. *Advances in water science,* 16 (2005), 88-91.
20. Da-zhi PAN, Yong DU. Multistage Fuzzy Diagnosis Model and Its Application for Assessment of Yangtze River's Water Quality. *Sichuan Environment,* 4 (2009), 15.
21. Da-Yong Chang. Applications of the Extent Analysis Method on Fuzzy AHP. *European journal of operational research,* 95 (1996), 649-55.