River Health Assessment Based on Entropy-Set Pair Analysis in Dalian City, China

Nan Zhang¹, Chunhui Li¹⁺, Yanpeng Cai²,³ and Xuan Wang²,³

¹ Key Laboratory for Water and Sediment Sciences of Ministry of Education, School of Environment, Beijing Normal University, Beijing, China
² State Key Laboratory of Water Environment Simulation, School of Environment, Beijing Normal University, Beijing, China
³ Beijing Engineering Research Center for Watershed Environmental Restoration & Integrated Ecological Regulation, School of Environment, Beijing Normal University, Beijing, China

Abstract. Rivers play a key role in both ecosystems and society. Regional development, especially urbanization has result in river system recession. Thus it is of great significance to evaluate and maintain the health of the river ecosystem for regional sustainable development. The evaluation of river ecological health is a complex problem for the ambiguous characteristics of river health and the incompatibility between the various elements of river. In this study, the health status of the six main rivers in Dalian city was assessed using entropy-set pair analysis method. The results indicated that that the river health state of S4, S9, S15 and S16 are the most sever. The sample site of S2, S3, S18 and S19 are the best in these sample sites.

Keywords: River health assessment, Set pair analysis (SPA), Dalian

1. Introduction

Rivers play a key role in ecosystems and society, and they provide a range of ecosystem functions such as shelter and food source for an array of biological species, aid in flood management and ecological refuge development [1]. Regional development, especially urbanization, changes the original functions of river profoundly, which results in a series of environmental and ecological problems, such as the river system recession, river connectivity decrease, river’s function decline and deterioration of water environment [2]. River health assessment is a complex problem for the ambiguous characteristics of river health and the incompatibility between the various elements. The evaluation of river ecological health is an analysis process combining certain evaluation index and evaluation standard with uncertain evaluation factors and its content changes.

Set pair analysis (SPA) proposed by Chinese scientist Zhao in 1989 [3] is an effective systems theory approach to process certain-uncertainty information. SPA has been increasingly used in many environment areas such as water resources planning [4], water quality assessment [5], urban ecosystem evaluation [6]. Entropy is a relatively ideal tool to gauge the weight of indexes in an index system and entropy weight method has the advantage in determining the index weight objectively [7].

In this research, Entropy-Set pair analysis is used to quantify the health state of six main rivers in Dalian city. The results would provide meaningful reference for river management and the sustainable development of Dalian city.

2. Study Area and Data Source

⁺ Corresponding author, Tel.: +86-10-58802756, Fax: +86-10-58802756.
E-mail address: chunhuili@bnu.edu.cn.
Dalian located in the south of the Liaodong Peninsula (Fig. 1), China (38°43’~40°10’N, 120°58’~123°31’E) covering an area of 12,573.85 km² and with a population of 5.91 million in 2013 [8]. With Northeast China Plain and the Inner Mongolia Autonomous Region in the north, Conjugating the Bohai Bay with Shandong Peninsula across the sea, Dalian known as the "gateway to Beijing and Tianjin" and the "windows of the Northeast of China".

Dalian is a typical water-stressed coastal city in China, the average gross quantity of water resources for many years is about 3.1 billion m³, and the quantity of per capita water resources in 2010 was only 566 m³, only roughly 1/16 of the world average (2100 m³) [9]. The six major rivers are the main surface water source in Dalian, covering more than 90% of the total amount of surface water supply. The exploitation degrees of the six main rivers are very high, some of them are close to or more than 50%.

Dalian as the Northeast Asian International Shipping Centre, the biggest coastal port city in the north-eastern part of China, the economy of it is rapidly developing. Under the influence of water consumption increases with economy developing, normal runoff may decrease with climate change and other factors, the situation of guaranteeing the river ecological water demand and maintaining the river ecosystem health become more severe.

In this study, the six main rivers were chosen as research object, each river chose three or four representative sample sites on the basis of the upstream, midstream and downstream. Through investigating the water quality, water volume and sewage outlet, 19 sample sites were chosen on total (Fig. 1.). The water quality data and aquatic biological diversity data are got on October in 2015, other data are from Dalian statistical yearbook.

![Study area with aquatic ecosystem monitoring stations](image)

3. Methodology

3.1. River Health Indicators

So far, there have been a lot of researches on river ecosystem health evaluation. Due to the particularity of the rivers in Dalian city, it is inappropriate to evaluate the river health state by using existed index system. The index system should be build based on the ecological and environment situation of rivers in Dalian.

Through investigation and analysis, the most prominent problem of the rivers in Dalian is: water quality deterioration; water quantity reduction; biological diversity decline. Considering independence and concise of indicators, availability of data, and studies of other researches [10], [11], 3 factors including water quality, water quantity, aquatic biological diversity to reflect the status of river health were selected (Table I).

3.2. Evaluation Standard

On the basis of The surface water environment quality standards GB3838-2002, the water quality pollution classification standard is shown in Table II.
The threshold of the exploitation degree of rivers should be determined based on the theory of human-water harmony, which means that it can not only support for the reasonable demand of economic and society, but also can be beneficial to the sustainable development of water resources and river ecosystem. Too high or too low of exploitation degree of rivers is not in conformity with the requirements of river health. Internationally, the recognized reasonable limit for the exploitation degree of rivers is 30%–40%, even if make full use of rain and flood resources; the exploitation degree of rivers should not be higher than 60%.

Table I: Assessment indicator system of river health

<table>
<thead>
<tr>
<th>Factor</th>
<th>Detailed indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>water quality</td>
<td>DO</td>
</tr>
<tr>
<td></td>
<td>COD&lt;sub&gt;Mn&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>BOD&lt;sub&gt;5&lt;/sub&gt;</td>
</tr>
<tr>
<td></td>
<td>NH&lt;sub&gt;3&lt;/sub&gt;-N</td>
</tr>
<tr>
<td></td>
<td>TP</td>
</tr>
<tr>
<td></td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>water quantity</td>
<td>Exploitation degree of the evaluated river</td>
</tr>
<tr>
<td></td>
<td>Amount of regional water resources</td>
</tr>
<tr>
<td>aquatic biological diversity</td>
<td>Diversity index of phytoplankton</td>
</tr>
<tr>
<td></td>
<td>Diversity index of zooplankton</td>
</tr>
<tr>
<td></td>
<td>Diversity index of macrobenthos</td>
</tr>
</tbody>
</table>

Table II: Water quality pollution classification standard

<table>
<thead>
<tr>
<th>Level</th>
<th>DO (mg/L)</th>
<th>NH&lt;sub&gt;3&lt;/sub&gt;-N (mg/L)</th>
<th>TP (mg/L)</th>
<th>COD&lt;sub&gt;Mn&lt;/sub&gt; (mg/L)</th>
<th>BOD&lt;sub&gt;5&lt;/sub&gt; (mg/L)</th>
<th>E.coli (MPN/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>≥ 7.5</td>
<td>≤ 0.15</td>
<td>≤ 0.02</td>
<td>≤ 2</td>
<td>≤ 3</td>
<td>≤ 200</td>
</tr>
<tr>
<td>II</td>
<td>6</td>
<td>0.5</td>
<td>0.1</td>
<td>4</td>
<td>3</td>
<td>2000</td>
</tr>
<tr>
<td>III</td>
<td>5</td>
<td>1</td>
<td>0.2</td>
<td>6</td>
<td>4</td>
<td>1000</td>
</tr>
<tr>
<td>IV</td>
<td>3</td>
<td>1.5</td>
<td>0.3</td>
<td>10</td>
<td>6</td>
<td>20000</td>
</tr>
<tr>
<td>V</td>
<td>2</td>
<td>2</td>
<td>0.4</td>
<td>15</td>
<td>10</td>
<td>40000</td>
</tr>
</tbody>
</table>

Therefore, the conceptual model of index score for the exploitation degree of the evaluated river is shown as Fig. 2. The model is in the form of parabola: the highest partition is in 30-40%, too high (over 60%) or too low (0%) of the exploitation degree will be given 0 score. Assume that when the score is 100~80, the health grade is I, so II: 80~60; III:60~40; IV:40~20; V:20~0.

Fig. 2: The conceptual model of index score for the exploitation degree of the evaluated river
The conceptual model formula is \( WRU_r = a(WRU)^2 + b*WRU \)

where, \( WRU_r \) is the index score for the exploitation degree of the evaluated river; \( WRU \) is the exploitation degree of the evaluated river; \( a \) and \( b \) are coefficients, \( a=-1111.11, b=666.67 \).

By calculating the biological diversity index, and based on the index evaluation standard to evaluate the health status of river. The diversity indexes include Shannon-Wiener diversity index \( (H') \), Pielou evenness index \( (E) \), Margalef index \( (M) \). The calculating formula of each index is shown below:

\[
H' = -\sum_{i=1}^{S} \frac{n_i}{N} \log_{10} \frac{n_i}{N} \\
E = -\sum_{i=1}^{S} \frac{n_i}{N} \log_{10} n_i \\
M = \frac{S-1}{\ln(N)}
\]

where \( S \) is the total number of species appeared in each sample point; \( N \) is the total number of individuals in each point; \( n_i \) is the number of \( i \) specie.

According to earlier researches [12, 13], the evaluation standard of each aquatic biological diversity index is shown on Table III.

### 3.3. Evaluation Methodology

The evaluation of river ecological health is a complex problem for the ambiguous characteristics of river health and the incompatibility between the various elements of river. Therefore in this research, the set pair analysis combine with entropy was used to solve the complex problem.

#### 3.3.1. Set pair analysis method of river ecosystem health evaluation

When use set pair analysis methods to evaluate river ecosystem health, according to the river ecosystem health evaluation index system the researcher constructed, the state of river health can be divided into five grades. Assume there are \( A \) evaluated indexes in \( Ⅰ \) grade, \( B_1, B_2, B_3 \) evaluated indexes in \( Ⅱ, Ⅲ, Ⅳ \) grades respectively, \( C \) evaluated indexes in \( Ⅴ \) grades. According to set pair analysis theory, the connection degree of each evaluated samples is performed as:

\[
\mu_{mn} = A + B_1 i_1 + B_2 i_2 + B_3 i_3 + C j = a + b_1 i_1 + b_2 i_2 + b_3 i_3 + c j
\]

By comparing the connection degree of \( a, b_1, b_2, b_3 \) and \( c \), the difference level of the evaluated samples can be conclude preliminarily. Then further set pair analysis should be conducted to get the same, difference and opposition of evaluation indexes values with evaluation standard. According to the characteristic of evaluation indexes, for indicators that correlates positively with standard, the connection degree \( \mu_{nn} \) is determined as:

\[
\mu_{mn} = \begin{cases} 
1 + 0i_1 + 0i_2 + 0i_3 + 0j, x_j \in [S_1, +\infty) \\
\frac{x-S_2}{S_2-S_1} + \frac{S_1-x}{S_2-S_1} i_1 + 0i_2 + 0i_3 + 0j, x_j \in [S_2, S_1) \\
0 + \frac{x-S_3}{S_3-S_2} i_1 + \frac{S_3-x}{S_3-S_2} i_2 + 0i_3 + 0j, x_j \in [S_3, S_2) \\
0 + 0i_1 + \frac{x-S_4}{S_4-S_3} i_2 + \frac{S_4-x}{S_4-S_3} i_3 + 0j, x_j \in [S_4, S_3) \\
0 + 0i_1 + 0i_2 + \frac{x-S_5}{S_5-S_4} i_3 + \frac{S_5-x}{S_5-S_4} j, x_j \in [S_5, S_4) \\
0 + 0i_1 + 0i_2 + 0i_3 + 1j, x_j \in [0, S_5)
\end{cases}
\]

For indicators that correlates negtively with standard, the connection degree \( \mu_{nn} \) is determined as
\[ \mu_{mn} = \begin{cases} 
1 + 0i_1 + 0i_2 + 0i_3 + 0j, x_j \in [0, S_1] \\
\frac{S_2 - x}{S_2 - S_1}i_1 + 0i_2 + 0i_3 + 0j, x_j \in (S_1, S_2] \\
0 + \frac{S_3 - x}{S_3 - S_2}i_1 + 0i_2 + 0i_3 + 0j, x_j \in (S_2, S_3] \\
0 + 0i_1 + \frac{S_4 - x}{S_4 - S_3}i_2 + 0i_3 + 0j, x_j \in (S_3, S_4] \\
0 + 0i_1 + 0i_2 + \frac{S_5 - x}{S_5 - S_4}i_3 + 0j, x_j \in (S_4, S_5] \\
0 + 0i_1 + 0i_2 + 0i_3 + 1j, x_j \in (S_5, +\infty) 
\end{cases} \]

where \( m \) is the number of evaluation sample; \( n \) is the number of evaluation index; \( x \) is the measured value; \( S_1, S_2, S_3, S_4, S_5 \) represent the upper limit of level standard respectively.

### Table III: Evaluation standard of aquatic biological diversity index

<table>
<thead>
<tr>
<th>Health level</th>
<th>Phytoplankton diversity index</th>
<th>Zooplankton diversity index</th>
<th>Macrobenthos diversity index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( H' )</td>
<td>( E )</td>
<td>( M )</td>
</tr>
<tr>
<td>I</td>
<td>( \geq )</td>
<td>( \geq )</td>
<td>( \geq )</td>
</tr>
<tr>
<td>II</td>
<td>3.0</td>
<td>0.8</td>
<td>4.0</td>
</tr>
<tr>
<td>III</td>
<td>2.5</td>
<td>0.6</td>
<td>3.0</td>
</tr>
<tr>
<td>IV</td>
<td>2.0</td>
<td>0.4</td>
<td>2.0</td>
</tr>
<tr>
<td>V</td>
<td>1.0</td>
<td>0.2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

### 3.3.2. The entropy weight method to determine the weight coefficient of the evaluation indexes

The weight is a measure of the influence of each index to the level of river ecosystem health. In this study, entropy weight method was adopted to determine the weight coefficient of the evaluation indexes. In information theory, entropy value reflects the degree of information disorder, which can be used to measure the amount of information. An index has more information, it plays a more important role to the final decision, and the entropy value is smaller, so is the information disorder degree of the system. Therefore entropy weight method can be used to determine the weight coefficient of the evaluation indexes. The main computational steps are as follows:

1. Assuming there are \( m \) evaluation objects, each evaluation objects have \( n \) evaluation indexes, the judgment matrix \( R \) is

\[
R = (r_{st})_{mn}, (s=1,2,\cdots, m; t=1,2,\cdots, n)
\]

where \( m \) is the number of evaluation sample; \( n \) is the number of evaluation index; \( r_{st} \) is the measured value.

2. Normalize the judgment matrix \( R \), the element of the normalized matrix \( B \) is

\[
b_{st} = \frac{r_{st} - r_{min}}{r_{max} - r_{min}}
\]

where \( r_{max} \), \( r_{min} \) are the most satisfying and the most dissatisfied value respectively (the bigger the better or the smaller the better) of different things in the same assessment index.

3. According to the traditional conception of entropy, the entropy value of each evaluation index is

\[
e_i = -\sum_{s=1}^{n} f_{st} \ln f_{st} / \ln m, (s=1,2,\cdots, m; t=1,2,\cdots, n)
\]

where \( f_{st} = b_{st} / \sum_{s=1}^{n} b_{st} \). When \( f_{st} = 0, \ e_i = 0 \).
3.3.3. River ecosystem health evaluation model based on entropy weight and set pair analysis

According to the index weight vector W and the connection degree \( \mu_m \) and \( \mu_{mn} \), the evaluation model based on entropy weight and set pair analysis is

\[
\bar{\mu}_m = \mu_m \times \sum_{i=1}^{n} (w_i \times \mu_{mi})
\]

where \( \bar{\mu}_m \) is the comprehensive connection degree of evaluation sample m, t is the number of evaluation index.

 Normalize all components in each \( \bar{\mu}_m \), the final connection degree of each evaluated sample can be get.

Finally the evaluated sample level can be get according to the following formula, the evaluated sample belong to l level based on \( h_l \).

\[
h_i = y_1 + y_2 + \cdots + y_t > \lambda, l = 1, 2, \cdots, 5,
\]

where \( y_1=a \), \( y_2=b_1 \), \( y_3=b_2 \), \( y_4=b_3 \), \( y_5=c \); \( \lambda \) is confidence coefficient, the general value is in \([0.5, 0.7]\). The evaluate result is more reliable if \( \lambda \) is bigger.

4. Results and Discussion

Evaluate the state of river health using the above method. First, based on the measured value of evaluation index in each monitoring station and the classification standard, preliminary calculated the set pair analysis connection degree of water quality situation of all the evaluation samples, such as:

\[
\mu_i = \frac{1}{6} + \frac{1}{6} i + \frac{3}{6} i + \frac{1}{6} j + \frac{1}{6} j
\]

\[
\mu_6 = \frac{3}{6} i + \frac{1}{6} i + \frac{1}{6} i + \frac{1}{6} j + \frac{1}{6} j
\]

According to the preliminary connection degree, the basic health status of rivers can be learned, such as \( S_{16} \) is at the same level with \( S_{17} \), \( S_2 \) is better than \( S_1 \).

To further analyze the relationship between the measured values with health standard, further set pair analysis is need. Calculate the connection degree between the measured values of each evaluation index and the healthy levels. Take \( S_1 \) for example:

\[
\mu_{DO} = 1 + 0i + 0i + 0i + 0j
\]

\[
\mu_{NH_3-N} = 0.5947 + 0.4053i + 0i + 0j + 0j
\]

\[
\mu_{TP} = 1 + 0i + 0i + 0i + 0j + 0j
\]

\[
\mu_{BOD_5} = 0 + 0.9383 + 0.9383i + 0i + 0j + 0j
\]

\[
\mu_{COD} = 1 + 0i + 0i + 0i + 0j + 0j
\]

The weight of each evaluation index is shown on Table IV based on entropy weight method.

<table>
<thead>
<tr>
<th>Index</th>
<th>DO</th>
<th>NH₃-N</th>
<th>TP</th>
<th>COD₅₉₀</th>
<th>BOD₅</th>
<th>Fecal coliform</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight</td>
<td>0.07</td>
<td>0.34</td>
<td>0.26</td>
<td>0.12</td>
<td>0.07</td>
<td>0.13</td>
</tr>
</tbody>
</table>

Finally, calculate the comprehensive connection degree of sample, the final connection degree are obtained after normalized processing (Table V). Take confidence coefficient \( \lambda \) for 0.6, the water quality health level are as follow (Table V).

The health level of aquatic biological diversity can also be obtained from the method. The results of river health evaluation in Dalian are shown as Fig. 3. The columns in the figure represents the health level of the rivers, the health level of water quality, water quantity and aquatic biological diversity were stacked together. Thus, the height of the columns can represent the whole health level of rivers. The higher of the columns, the health level of rivers are more severe. The researcher can get from the figure that the river health state of S4, S9, S15 and S16 are the most sever. The sample site of S2, S3, S18 and S19 are the best in these 19 sites.
### Table V: Final connection degree of each Sampling Sites and the health level

<table>
<thead>
<tr>
<th>Sampling Sites</th>
<th>Comprehensive connection degree</th>
<th>Health level</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>$\mu_1 = 0.091 + 0.270i_1 + 0.626i_2 + 0i_3 + 0.013j$</td>
<td>III</td>
</tr>
<tr>
<td>S2</td>
<td>$\mu_2 = 0.655 + 0.306i_1 + 0i_2 + 0i_3 + 0.039j$</td>
<td>I</td>
</tr>
<tr>
<td>S3</td>
<td>$\mu_3 = 0.659 + 0.34i_1 + 0i_2 + 0i_3 + 0j$</td>
<td>I</td>
</tr>
<tr>
<td>S4</td>
<td>$\mu_4 = 0.213 + 0.226i_1 + 0i_2 + 0i_3 + 0.561j$</td>
<td>V</td>
</tr>
<tr>
<td>S5</td>
<td>$\mu_5 = 0 + 0.502i_1 + 0i_2 + 0.023i_3 + 0.475j$</td>
<td>III</td>
</tr>
<tr>
<td>S6</td>
<td>$\mu_6 = 0.175 + 0.248i_1 + 0.414i_2 + 0i_3 + 0.163j$</td>
<td>V</td>
</tr>
<tr>
<td>S7</td>
<td>$\mu_7 = 0 + 0i_1 + 0.311i_2 + 0.081i_3 + 0.608j$</td>
<td>V</td>
</tr>
<tr>
<td>S8</td>
<td>$\mu_8 = 0.593 + 0i_1 + 0.043i_2 + 0.262i_3 + 0.102j$</td>
<td>III</td>
</tr>
<tr>
<td>S9</td>
<td>$\mu_9 = 0.184 + 0.142i_1 + 0.060i_2 + 0i_3 + 0.614j$</td>
<td>V</td>
</tr>
<tr>
<td>S10</td>
<td>$\mu_{10} = 0.036 + 0.386i_1 + 0.228i_2 + 0i_3 + 0.350j$</td>
<td>III</td>
</tr>
<tr>
<td>S11</td>
<td>$\mu_{11} = 0.661 + 0.208i_1 + 0i_2 + 0.110i_3 + 0.110j$</td>
<td>I</td>
</tr>
<tr>
<td>S12</td>
<td>$\mu_{12} = 0.601 + 0.245i_1 + 0i_2 + 0.066i_3 + 0.088j$</td>
<td>I</td>
</tr>
<tr>
<td>S13</td>
<td>$\mu_{13} = 0.511 + 0.482i_1 + 0.007i_2 + 0i_3 + 0j$</td>
<td>II</td>
</tr>
<tr>
<td>S14</td>
<td>$\mu_{14} = 0 + 0.602i_1 + 0.398i_2 + 0i_3 + 0j$</td>
<td>II</td>
</tr>
<tr>
<td>S15</td>
<td>$\mu_{15} = 0.530 + 0.008i_1 + 0i_2 + 0i_3 + 0.462j$</td>
<td>V</td>
</tr>
<tr>
<td>S16</td>
<td>$\mu_{16} = 0.194 + 0.214i_1 + 0.039i_2 + 0i_3 + 0.554j$</td>
<td>V</td>
</tr>
<tr>
<td>S17</td>
<td>$\mu_{17} = 0 + 0i_1 + 0.292i_2 + 0.232i_3 + 0.476j$</td>
<td>V</td>
</tr>
<tr>
<td>S18</td>
<td>$\mu_{18} = 0.805 + 0.179i_1 + 0.016i_2 + 0i_3 + 0j$</td>
<td>I</td>
</tr>
<tr>
<td>S19</td>
<td>$\mu_{19} = 0.493 + 0.495i_1 + 0i_2 + 0i_3 + 0.012j$</td>
<td>II</td>
</tr>
</tbody>
</table>

5. **Conclusions**

In this study, an assessment indicator system including water quality, water quantity and aquatic biological diversity 3 factors was build. Entropy-set pair analysis method was conducted to evaluate the status of river health in Dalian city. The results show that the river health state of S4, S9, S15 and S16 are the most severe. The sample site of S2, S3, S18 and S19 are the best in these 19 sites. How to improve and keep the river in the health state will be an important topic in the near future. The results would provide meaningful reference for river management and the sustainable development of Dalian city.
6. Acknowledgments

This research was supported by the National Science and Technology Pillar Program, China (No. 2012BAC05B02).

7. References


