Expert Systems with Applications 37 (2010) 2517–2521

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Using fuzzy theory and information entropy for water quality assessment in Three Gorges region, China

Li Liu [a], Jianzhong Zhou [b,*], Xueli An [c], Yongchuan Zhang [b], Li Yang [b]

[a] College of Hydroelectric and Digitalization Engineering, Huazhong University of Science and Technology, Hubei Province, 430074 Wuhan, China
[b] Dispatch & Communication Center, Hunan Electric Power Company, Hunan Province, Changsha 410007, China
[c] State Key Laboratory of Control and Simulation of Power System and Generation Equipment, Department of Thermal Engineering, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Keywords:
Water quality assessment
Fuzzy mathematics
Information entropy
Three Gorges region

ABSTRACT

Considering that the water quality assessment is a fuzzy concept with multiple indicators and classes, and there are still some limits of fuzzy comprehensive evaluation method, the fuzzy mathematics method and the information entropy theory are combined to establish an improved fuzzy comprehensive evaluation method for water quality assessment. In this method, the exponential membership function has been adopted to solve the zero-weight problem, and the information entropy has been used to modify the coefficients of weight in order to exploit the useful information of data to a maximum extent. In addition, the weighted average principle has been taken to replace the maximum membership principle for reserving the information in the assessment coefficients as much as possible. The water quality of Three Gorges region is taken as an example and the results show that the improved fuzzy comprehensive evaluation method is superior to the traditional model and worth to be recommended.

1. Introduction

As a kind of attribute recognition process, the water quality assessment plays a significant role in the water resource development and utilization, and the method for it is also an important theme in the research of foundation theory of modern environmental science. Considering that the water quality assessment for natural water resource is a fuzzy concept with multiple indicators and classes, the fuzzy comprehensive evaluation method, which has been proved effective in solving problems of fuzzy boundaries and controlling the effect of monitoring errors on assessment results (Guleda, Ibrahim, & Halil, 2004; Wang, 2002), has been studied for the water quality assessment extensively in recent years (Icaga, 2007; Karmakar & Mujumdar, 2006; Kung, Ying, & Liu, 1992; Lu & Lo, 2002; Shen, Lu, Wang, & Sun, 2005). These researches find that this model is fit to describe fuzzy character of classified bounds for water quality and can reflect the actual water quality on objectiveness.

However, there are still some limits when applying fuzzy comprehensive evaluation method in water quality assessment, such as when the method emphasizes extreme value action, more information is lost and the scientific character of weight value is not clear, etc. Especially, the weight value which usually contains the information of the individual indicator only but has nothing to do with the relationship between assessment objects.

To solve the problems, the information entropy has been introduced (Karmeshu, 2003). As a measurement of the disorder degree of a system, information entropy can measure the amount of useful information with the provided data and has been widely used in engineering, economy and finance, etc. (Chang, Chen, Wang, & Althouse, 1994; Cheng, Chen, & Li, 1998; Kannathal, Choo, Acharya, & Sadasivan, 2005; Piplani & Wetjens, 2007; Shuiabi, Thomson, & Bhuiyan, 2005). When the difference of the evaluating objects on the same indicator is large, the entropy is small, which illustrates that this indicator provides more useful information and this indicator's weight should be set relatively large. On the other hand, if the difference is small, the entropy is large, and the corresponding weight would be small. Hence, the entropy theory is a comparatively objective way to determine the weight (Zou, Yun, & Sun, 2006).

Based on the concept mentioned earlier, this paper combines the entropy method with the fuzzy mathematics method to establish an improved fuzzy comprehensive evaluation method, and applied the method in water quality evaluation of the Three Gorges reservoir area. The organization of this paper is as follows: Section 2 describes the general model of fuzzy comprehensive evaluation method and its improvement, Section 3 provides an application example of the Three Gorges region, which verifies its practicality while Section 4 presents the conclusions and discusses future research.

* Corresponding author. Tel.: +86 027 87543127.
E-mail addresses: liuli0324@gmail.com (L. Liu), jz.zhou@mail.hust.edu.cn (J. Zhou), anxl614@163.com (X. An), yczhang@mail.hust.edu.cn (Y. Zhang).

0957-4174/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.eswa.2009.08.004

2. Methodology

2.1. Original fuzzy comprehensive evaluation

Fuzzy comprehensive evaluation is the process of evaluating an objective utilizing the fuzzy set theory, which comprehensively considers the contributions of multiple related indicators according to weights and decreases the fuzziness by using membership functions (Chen & Wei, 2000; Jorge, Antunes, & Martins, 2000).

Usually, $i = 1, 2, \ldots, n$ stands for the assessment object set, $j = 1, 2, \ldots, m$ for the assessment indicator set, $k = 1, 2, \ldots, c$ for the assessment criteria set. The procedure of fuzzy comprehensive evaluation is described as follows:

(1) *Select assessment parameters*
It is crucial to select assessment parameters that are representative, rational and accurate to form an assessment indicator matrix U, which is based on the actual local situation, and can be expressed as:

$$U = (u_{i,j})_{n\times m}$$

where $u_{i,j}$ is the measured value from ith object for jth indicator.

(2) *Establish assessment criteria*
The assessment criteria matrix V is established and expressed as:

$$V = (v_{j,k})_{m\times c}$$

where $v_{j,k}$ is the value of the assessment criteria from jth indicator for kth class.

(3) *Establish membership functions*
The membership functions represent the degree to which the specified concentration belongs to the fuzzy set. The membership degrees of assessment parameters at each class can be described quantitatively by a set of formulae comprised of membership functions as follows:

$$g_{j,1}(u_{i,j}) = \begin{cases} 1 & u_{i,j} \in [0, v_{j,1}] \\ (v_{j,2} - u_{i,j})/(v_{j,2} - v_{j,1}) & u_{i,j} \in [v_{j,1}, v_{j,2}] \\ 0 & u_{i,j} \in [v_{j,2}, +\infty) \end{cases} \quad (1)$$

$$g_{j,k}(u_{i,j}) = \begin{cases} 0 & u_{i,j} \in [0, v_{j,k-1}] \\ (u_{i,j} - v_{j,k-1})/(v_{j,k} - v_{j,k-1}) & u_{i,j} \in [v_{j,k-1}, v_{j,k}] \\ (v_{j,k+1} - u_{i,j})/(v_{j,k+1} - v_{j,k}) & u_{i,j} \in [v_{j,k}, v_{j,k+1}] \\ 0 & u_{i,j} \in [v_{j,k+1}, +\infty) \end{cases} \quad (2)$$

$$g_{j,c}(u_{i,j}) = \begin{cases} 0 & u_{i,j} \in [0, v_{j,c-1}] \\ (u_{i,j} - v_{j,c-1})/(v_{j,c} - v_{j,c-1}) & u_{i,j} \in [v_{j,c-1}, v_{j,c}] \\ 1 & u_{i,j} \in [v_{j,c}, +\infty) \end{cases} \quad (3)$$

where $f_{j,k}$ is the membership function from jth indicator for kth class.

(4) *Calculate the weights matrix*
The weight is a kind of relative weight which reflects the dangerous degree of each indicator, and the value is always obtained as follows:

$$r_{i,j} = u_{i,j} / \frac{1}{K}\sum_{k=1}^{K} S_{j,k} \quad (4)$$

$$\omega_{i,j} = r_{i,j} / \sum_{j=1}^{m} r_{i,j} \quad (5)$$

where $S_{j,k}$ is the certified value from jth indicator for kth class, $\omega_{i,j}$ is the weight from ith object for jth indicator.

(5) *Calculate assessment coefficient set*
As a comprehensive evaluation mathematic model of major indicator dominating, the Zadeh operator $M(\vee, \bullet)$ takes "$\bullet$" and "$\vee$" to denotes intersection and union operators, respectively and is commonly used in the fuzzy comprehensive evaluation. However, this operator might neglect some useful information, especially the information of those non-main indicators, when there are many qualitative indicators and each weight is small. So, the weight average operator $M(+, \bullet)$ has been introduced, which takes "+" and "$\bullet$" to replace "$\vee$" and "$\wedge$", respectively:

$$\varepsilon_{i,k} = \sum_{j=1}^{m} \omega_{i,j} \bullet g_{j,k}(u_{i,j}) \quad (6)$$

where $\varepsilon_{i,k}$ is the assessment coefficient from ith object for kth class.
The assessment coefficient considers and maintains all the indicators through each indicator's weight. This offers a basis for the synthetic evaluation with multiple indicators.

(6) *Judgment of object class*
According to the maximum membership principle, the object class is the membership class to which the maximum assessment coefficient corresponds.

2.2. Exponential membership function

When calculating the actual weight, the influence between the non-neighboring classes has always been neglected in fuzzy comprehensive evaluation, which means that the zero-weight problem may appear, and it may get warped results. So, the exponential membership function is put forward in the method as follows:

$$g_{j,1}(u_{i,j}) = \begin{cases} 1 & u_{i,j} \in (0, v_{j,1}] \\ e^{p\times(v_{j,1}-u_{i,j})}/v_{j,1} & u_{i,j} \in (v_{j,1}, +\infty) \end{cases} \quad (7)$$

$$g_{j,k}(u_{i,j}) = \begin{cases} e^{p\times(u_{i,j}-v_{j,k})}/u_{i,j} & u_{i,j} \in (0, v_{j,k}] \\ e^{p\times(v_{j,k}-u_{i,j})/v_{j,k}} & u_{i,j} \in (v_{j,k}, +\infty) \end{cases} \quad (8)$$

$$g_{j,c}(u_{i,j}) = \begin{cases} e^{p\times(u_{i,j}-v_{j,c})/u_{i,j}} & u_{i,j} \in (0, v_{j,c}] \\ 1 & u_{i,j} \in (v_{j,c}, +\infty) \end{cases} \quad (9)$$

where p is an attenuation indicator and its value determines the attenuation velocity of correlation degree between in membership league (Fig. 1).

According to the exponential membership function, each value of certain indicator is unique corresponding with a nonzero weight of various classes, so that the zero-weight problem can be addressed.

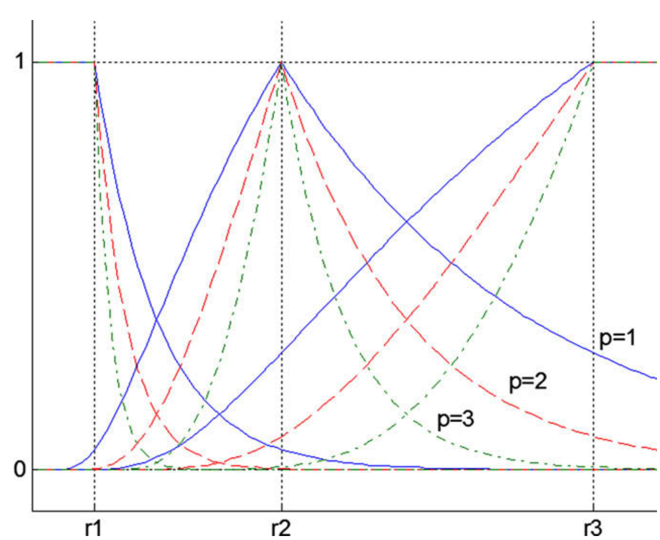

Fig. 1. Exponential membership function with different p.