

文章编号:1672-3961(2009)03-0139-05

模糊聚类分析在区域水资源承载能力评价中的应用

王好芳¹, 吴美², 陈文艳¹

(1. 山东大学土建与水利学院, 山东 济南 250061; 2. 济南市科学技术馆, 山东 济南 250002)

摘要:介绍了模糊聚类分析的基本原理和计算步骤,并利用该分析方法对山东省17地市的水资源承载能力进行评价.评价结果表明:淄博市、威海市、青岛市、潍坊市、济宁市、日照市、枣庄市、莱芜市、烟台市和临沂市等10个地市的水资源承载能力较高,而其它7地市的水资源承载能力相对较低.

关键词:模糊聚类分析;水资源承载能力;区域

中图分类号:TV213 **文献标志码:**A

Application of fuzzy cluster analysis to regional water resources carrying capacity evaluation

WANG Hao-fang¹, WU Mei², CHEN Wen-yan¹

(1. School of Civil and Engineering, Shandong University, Jinan 250061, China;

2. Jinan Science and Technology Museum, Jinan 250002, China)

Abstract: The fuzzy cluster analysis principle and calculation approaches were introduced and adopted to evaluate the water resource carrying capacity of seventeen regions in Shandong Province. The evaluation results showed that Zibo City, Weihai City, Qingdao City, Weifang City, Jining City, Rizhao City, Zaozhuang City, Laiwu City, Yantai City, and Linyi City were in the higher capacity of water resources, while the water resources carrying capacity of the other 7 cities were low.

Key words: fuzzy cluster analysis; water resource carrying capacity; region

0 引言

水资源是社会经济持续发展的战略资源,水资源承载能力直接影响一个区域社会经济与生态环境的健康持续发展.水资源承载能力的大小,不仅取决于水资源系统本身的水资源状况,而且更决定于社会经济系统对水资源系统的开发利用以及对水环境的保护与治理.因此,它是水资源系统、社会经济系统与水环境系统中多种影响因素共同耦合的结果,具有模糊性.另外,一个区域往往包括许多行政区域,每个行政区域的水资源状况、社会经济发展规模、水资源开发利用模式以及水环境保护与治理力度的不同,也对该区域的水资源承载能力产生综合影响,而每个地区影响水承载能力的各主要因素又具有相似性,于是本文引进模糊聚类分析法,综合评价区域水资源承载能力状况,利用模糊聚类分析法能得到较客观的、合理的评价结果,目前,该方法在水资源、水环境研究中已有较多应用^[1-3].

1 模糊聚类分析的原理及计算步骤

模糊聚类分析是数理统计中研究“物以类聚”的一种多元分析方法,依据客观事物间的特征、亲疏程度

和相似性,通过建立模糊相似矩阵对客观事物进行分类的方法^[4-5].常用的模糊聚类算法有基于相似性关系和模糊关系的方法、基于模糊等价关系的传递闭包、基于模糊图论的最大树法、基于目标函数法^[6].本文采用基于模糊等价关系的传递闭包法,其分析计算步骤如下:

1.1 建立原始数据矩阵

设论域 $X = (x_1, x_2, \dots, x_n)$ 为 n 个待分类对象,且每个对象都有 m 个特征统计指标,即 $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$, $i = 1, 2, \dots, n$. 由此可得到论域的原始数据矩阵:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1m} \\ x_{21} & \cdots & x_{2m} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nm} \end{bmatrix}. \quad (1)$$

1.2 数据标准化

由于指标的量纲和数量级不同,直接利用原始数据进行计算,就可能突出某些数量级特别大的特性指标对分类的作用,而降低甚至排斥某些数量级较小的特性的作用^[7].因此,必须对原始数据进行量纲-处理,以便分析和比较.常用的方法有:标准差法、极差法.

(1) 标准差法

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j}, \quad (2)$$

式中, $\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij}$, $S_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}$.

(2) 极差法

$$x'_{ij} = \frac{x_{ij} - x_{j\min}}{x_{j\max} - x_{j\min}}, \quad (3)$$

式中, $x_{j\max} = \max\{x_{1j}, x_{2j}, \dots, x_{nj}\}$, $x_{j\min} = \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}$.

1.3 构造模糊相似矩阵

通过计算分类对象间的相似程度 $r_{ij} \in [0, 1]$,构造模糊相似矩阵 $R = (r_{ij})_{n \times n}$,式中 $r_{ij} = r_{ji}$, $r_{ii} = 1$ ($i, j = 1, 2, \dots, n$).计算 r_{ij} 的方法主要有:距离法、相似系数法和贴近度法三大类.究竟选用3类中的哪种方法,视问题的实际情况而定,具体原则可参考文献^[7].在本文的应用示例中,选用了最大最小法其数学模型为

$$r_{ij} = \frac{\sum_{k=1}^m \min(x'_{ik}, x'_{jk})}{\sum_{k=1}^m \max(x'_{ik}, x'_{jk})}. \quad (4)$$

1.4 构造模糊等价矩阵

在多数情况下, R 只满足自反性和对称性,而不满足传递性.因此,若要对被分类对象集进行分类,必须由模糊相似矩阵 R 出发构造一个满足传递性要求的模糊等价矩阵.而模糊相似矩阵 R 的传递闭包 $t(R)$ 就是一个模糊等价矩阵,通常采用平方法求 $t(R)$,即采用平方法计算传递闭包: $R \rightarrow R^2 \rightarrow R^4 \rightarrow \dots \rightarrow R^{2^k}$,经过有限次运算后存在 k 使 $R^{2^k} = R^{2^{(k+1)}}$,于是 $t(R) = R^{2^k}$.

1.5 聚类分析

对满足传递性的模糊等价矩阵 $t(R)$ 进行聚类分析,给定不同置信水平的 λ ,计算 $[t(R)_\lambda]$ 的矩阵.找出 $t(R)$ 的 λ 显示,得到普通的分类关系^[4-5,7].然后根据实例中模糊等价矩阵中各指标及分类要求,按不同的 λ 值进行分类,将得到不同的聚类结果.

一般情况下对给定模糊等价矩阵 $t(R) = (r_{ij})$,对任意 $\lambda \in [0, 1]$ 记为 $[t(R)_\lambda] = (\lambda r_{ij})$,其中:当 $r_{ij} \geq \lambda$ 时, $\lambda r_{ij} = 1$;当 $r_{ij} < \lambda$ 时, $\lambda r_{ij} = 0$.令 λ 由 1 降至 0,写出 $[t(R)_\lambda]$,按 $[t(R)_\lambda]$ 分类元素 x_i 与 x_j 归为同一类的条件是 $R_\lambda(x_i, x_j) = 1$, ($i, j = 1, 2, 3, 4, 5$).

3 结语

本文介绍了模糊聚类分析的基本理论和计算步骤,并将该方法应用于区域水资源承载能力评价中.根据聚类分析结果可知,由于淄博市、威海市、青岛市、潍坊市、济宁市、日照市、枣庄市、莱芜市、烟台市和临沂市等10个地市的人均水资源量、单位耕地占有水资源量和单位面积水资源量较高,因此,其水资源承载能力较高,但其单位GDP耗水量也相对较高,若合理优化、调整其产业结构,降低单位GDP耗水量,其水资源承载能力潜力更大.其它7地市的水资源量、单位耕地占有水资源量和单位面积水资源量较低,水资源短缺,而其单位GDP耗水量相对较低,可通过开源节流、提高水资源重复利用率等措施,提高其水资源承载能力.

参考文献:

- [1] 王莹, 李安贵, 刘钦圣. 模糊聚类分析在全国主要地区供水用水情况中的应用[J]. 广西民族大学学报:自然科学版, 2007, 13(1):57-60.
WANG Ying, LI Angui, LIU Qincheng. A study of fuzzy clustering analysis in water supply and water use of Chinese great region[J]. Journal of Guangxi University for Nationalities: Natural Science Edition, 2007, 13(1):57-60.
- [2] 王立新, 刘华民. 模糊聚类法在我国城市内河水质污染分类研究中的应用[J]. 内蒙古大学学报:自然科学版, 2004, 35(6):710-715.
WANG Lixing, LIU Huamin. Application of fuzzy clustering analysis for studying water pollution classification of the inland river in cities of China[J]. Acta Scientiarum Naturalium Universitatis NeiMongol, 2004, 35(6):710-715.
- [3] 许延生, 刘兴芳. 模糊聚类迭代模型在水资源承载能力评价中的应用[J]. 山东大学学报:工学版, 2007, 37(3):100-105.
XU Yansheng, LIU Xingfang. Application of the fuzzy clustering iterative model to the evaluation of water resource carrying capacity[J]. Journal of Shandong University: Engineering Science, 2007, 37(3):100-105.
- [4] 贺仲雄. 模糊数学及其应用[M]. 天津:天津科学技术出版社, 1983.
HE Zhongxiong. Fuzzy mathematics and its applications[M]. Tianjin: Publishing Company of Tianjin Science and Technology, 1983.
- [5] 高新波. 模糊聚类分析及其应用[M]. 西安:西安电子科技大学出版社, 2004.
GAO Xingbo. Fuzzy cluster analysis and its application[M]. Xian: Publishing Company of Xi'an University of Electronic Science and Technology, 2004.
- [6] 扶名福, 谢明祥. 聚类分析法在抽风机故障诊断中应用[J]. 工程力学, 2007, 24(增刊1):36-40.
FU Mingfu, XIE Mingxiang. Application of fuzzy cluster analysis to the fault diagnosis of suction fan[J]. Engineering Mechanics, 2007, 24(Sup. I):36-40.
- [7] 庄恒扬, 沈新平, 陆建飞, 等. 模糊聚类计算方法的理论分析[J]. 江苏农学院学报, 1998, 19(3):37-41.
ZHUANG Hengyang, SHEN Xiping, LU Jianfei, et al. Theoretical analysis of the computing method of fuzzy cluster[J]. Journal of Jiangsu Agricultural College, 1998, 19(3):37-41.

(编辑:孙培芹)