

博斯腾湖水质评价的属性识别模型

胡安焱

(长安大学环境科学与工程学院, 陕西 西安 710054)

摘要 根据博斯腾湖水质资料,应用属性识别模型进行博斯腾湖的水质评价,结果表明博斯腾湖水质基本良好,影响湖泊水质的主要原因是农田排水,有排水口的区域水质较差,没有排水口的湖区水质均为良好。

关键词 博斯腾湖;水质评价;熵权;属性识别

中图分类号:X824 文献标识码:A 文章编号:1004-693X(2006)06-0025-03

Water quality evaluation model of Boston Lake based on attribute distinction theory

HU An-yan

(College of Environmental Science and Engineering, Chang'an University, Xi'an 710054, China)

Abstract Based on the water quality data of Boston Lake, an attribute distinction model was established to evaluate the water quality. The results indicate that the water quality is good in most regions of Boston Lake. Agricultural drainage is a main factor influencing the water quality, as the water quality is bad in region with effluents while good without effluents.

Key words Boston Lake; water quality evaluation; entropy weight; attribute distinction

博斯腾湖属中生代断陷湖,湖泊东西长 55 km,南北平均宽 20 km,是我国最大的内陆淡水湖。博斯腾湖既是开都河的尾闾,又是孔雀河的源头,兼有开都河来水的水资源调控、孔雀河流域农田灌溉、工业及城乡生活用水、流域生态环境保护和向塔里木河中下游紧急调水等多种功能。在水资源紧缺的新疆南部地区,有着十分重要的作用。随着上游工农业的发展和人口的增加,工业废水、农田排水和生活污水对博斯腾湖水质的影响加大,因此,加强湖泊水质的科学监测与评价具有重要的现实意义。

常用的水质评价数学模型有简单指数法、分级评分法、概率统计法、模糊数学法等,本文利用属性识别理论进行水质评价。一方面水质评价中的污染程度、水质类别等都是对水质属性的一种描述,对水质进行评价即是对水质属性进行确定和识别,水质评价过程本质上是一个对水质属性识别的过程;另一方面,水质状况受控于众多水质指标,各个指标对水质属性影响程度需要给出合理的定量表达,以往常用的定权重法、专家打分法和层次分析法一定程

度上都包含人为的主观影响,如何客观准确给出各指标的权重系数是值得关注的问题。

1 属性识别模型

1.1 属性空间矩阵的建立

在评价对象空间 X 中取得 n 个样本 x_1, x_2, \dots, x_n 对于每一个样本有 m 个评价指标 I_1, I_2, \dots, I_m , 第 i 个样本的第 j 个指标的测度值为 x_{ij} , 因此,每一个样本可以表示为一个向量 $x_i = (x_{i1}, x_{i2}, \dots, x_{im})$, $i = 1, 2, \dots, n$, n 个样本构成了 $n \times m$ 的样本空间矩阵:

$$X_{ij} = \begin{bmatrix} I_1 & x_{11} & x_{12} & \dots & x_{1n} \\ I_2 & x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ I_m & x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix} \quad (1)$$

设 F 为 x 上某类属性空间, C_1, C_2, \dots, C_k 为属性空间的有序分割类,且满足 $C_1 > C_2 > \dots > C_k$ 。每一个评价指标的属性分类已知,写成属性分类标准矩阵如下:

$$A = \begin{matrix} & C_1 & C_2 & \dots & C_k \\ \begin{matrix} I_1 \\ I_2 \\ \vdots \\ I_m \end{matrix} & \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mk} \end{bmatrix} \end{matrix} \quad (2)$$

$a_{j1} < a_{j2} < \dots < a_{jk}$ 或 $a_{j1} > a_{j2} > \dots > a_{jk}$

1.2 权重的确定

为了避免人为主观因素的干扰,权重的确定采用熵的方法,从实测数据出发,充分利用数据自身信息,客观地确定权重。评价指标根据其不同属性,可分为递增型和递减型。在计算各指标权重之前有必要对每个样本进行归一化处理,对递增型指标

$$x'_{ij} = x_{ij} / \max x_{ij}$$

对递减型指标

$$x'_{ij} = \min x_{ij} / x_{ij}$$

从而得到新评价样本矩阵:

$$X'_{ij} = \begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1m} \\ x'_{21} & x'_{22} & \dots & x'_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x'_{n1} & x'_{n2} & \dots & x'_{nm} \end{bmatrix} \quad (3)$$

再设第 j 个评价指标下第 i 个待评价监测点评价指标比重为

$$P_{ij} = x'_{ij} / \sum_{i=1}^n x'_{ij} \quad (4)$$

计算第 j 个评价指标的熵为

$$e_j = - \frac{1}{mn} \sum_{i=1}^n P_{ij} \ln P_{ij} \quad (5)$$

则第 j 个评价指标的权重为

$$\omega_j = (1 - e_j) / \sum_{i=1}^n (1 - e_j) \quad (6)$$

1.3 属性测度的计算和属性识别

第 i 个样本第 j 个指标 x_{ij} 具有属性 C_k 的属性测度 $\mu_{ijk} = \mu(x_{ij} \in C_k)$ 。假设

$$a_{j1} < a_{j2} < \dots < a_{jk}$$

当 $x_{ij} \leq a_{j1}$ 时, $\mu_{ij1} = 1, \mu_{ij2} = \dots = \mu_{ijk} = 0$

当 $x_{ij} \geq a_{jk}$ 时, $\mu_{ijk} = 1, \mu_{ij1} = \dots = \mu_{ijk-1} = 0$

当 $a_{jq} \leq x_{ij} \leq a_{jq+1}$ 时,

$$\mu_{ijq} = \text{ab}[(x_{ij} - a_{jq+1}) / (a_{jq} - a_{jq+1})]$$

$$\mu_{ijq+1} = \text{ab}[(x_{ij} - a_{jq}) / (a_{jq} - a_{jq+1})]$$

$$\mu_{ijk} = 0$$

$$k < 1 \text{ 或 } k > q + 1$$

得每个样本的属性测度矩阵:

$$\mu_{ik} = \mu(x_i \in C_k) = \sum_j \omega_j \mu_{ijk} \quad (7)$$

其中 $1 \leq i \leq n, 1 \leq q \leq k, 1 \leq j \leq m$ 。

有了评价指标权重和样本属性测度,就可以进行样本的属性识别。按照置信度准则,一般情况下,

置信度 λ 取值在 0.7 左右,识别模型为

$$k_i = \min \left\{ k : \sum_{q=1}^k \mu_{x_i}(C_q) \geq \lambda, 1 \leq q \leq k \right\} \quad (8)$$

则可根据 k_i 的大小对 x_i 进行排序和比较评价,得到样本 x_i 属于 C_{k_i} 类。

2 计算实例

根据波斯腾湖水质的多年资料的分析,选择 5 个主要的代表性评价指标为氯离子、硫酸根离子、矿化度、COD、氨氮。各测站分布见图 1,实测水质资料见表 1。



图 1 波斯腾湖水水质监测各测站分布

表 1 波斯腾湖各测站水质实测值 mg/L

测站	$\rho(\text{Cl}^-)$	$\rho(\text{SO}_4^{2-})$	矿化度	$\rho(\text{COD})$	$\rho(\text{NH}_3\text{-N})$
开都河	14	44	276	2.52	0.30
河口区	14	45	285	0.52	0.64
黄水区	689	1105	3128	13.0	0.50
出流区	199	270	1024	5.1	0.24
孔雀河	126	156	828	20.0	0.42
团结总干排	1206	1581	4580	64.0	0.42
东风总干排	809	1418	3810	50.0	0.49
胜利总干排	915	993	3790	169.0	5.57
黄水总干排	689	1102	3128	13.0	0.50

由波斯腾湖各测站实测水质资料得出评价样本的空间矩阵为

$$X = \begin{bmatrix} I_1 & 14 & 14 & 689 & 199 & 126 & 1206 & 809 & 915 & 689 \\ I_2 & 44 & 45 & 1105 & 270 & 156 & 1581 & 1418 & 993 & 1102 \\ I_3 & 276 & 285 & 3128 & 1024 & 828 & 4580 & 3810 & 3790 & 3128 \\ I_4 & 2.52 & 0.52 & 13.0 & 5.1 & 20.0 & 64.0 & 50.0 & 169.0 & 13.0 \\ I_5 & 0.30 & 0.64 & 0.5 & 0.24 & 0.42 & 0.42 & 0.49 & 5.57 & 0.50 \end{bmatrix}$$

部分地表水环境质量标准(GB 3838—2002)见表 2。

表 2 水质评价标准值 mg/L

水质类型	$\rho(\text{Cl}^-)$	$\rho(\text{SO}_4^{2-})$	矿化度	$\rho(\text{COD})$	$\rho(\text{NH}_3\text{-N})$
I	250	250	1000	15	0.15
II	250	250	3000	15	0.50
III	400	400	5000	20	1.00
IV	600	600	7000	30	1.50
V	800	800	10000	40	2.00

属性空间矩阵为

$$A = \begin{matrix} I_1 \\ I_2 \\ I_3 \\ I_4 \\ I_5 \end{matrix} \begin{bmatrix} 250 & 250 & 400 & 600 & 800 \\ 250 & 250 & 400 & 600 & 800 \\ 1000 & 3000 & 5000 & 7000 & 10000 \\ 15 & 15 & 20 & 30 & 40 \\ 0.15 & 0.5 & 1.0 & 1.5 & 2.0 \end{bmatrix}$$

通过计算评价指标的权重见表 3,各测站水质属性测度见表 4。

表 3 水质类型的权重

水质类型	I	II	III	IV	V
权重	0.1987	0.1986	0.1982	0.2019	0.2025

表 4 属性测度值

测 站	I	II	III	IV	V
开都河	0.9132	0.0868	0	0	0
河口区	0.7975	0.1431	0.0556	0	0
黄水区	0	0.5536	0.0425	0.1103	0.2871
出流区	0.7470	0.2257	0.0264	0	0
孔雀河	0.6419	0.1532	0	0.1987	0
团结总干排	0.0464	0.1949	0.1570	0	0.5962
东风总干排	0.0059	0.3112	0.0805	0	0.5962
胜利总干排	0	0.1202	0.0785	0	0.7949
黄水总干排	0	0.5536	0.0425	0.1103	0.2871

取置信度 $\lambda = 0.7$ 进行属性识别,得出水质评价结果见表 5。

表 5 博斯腾湖水质评价结果

测站	开都河	河口区	黄水区	出流区	孔雀河	团结总干排	东风总干排	胜利总干排	黄水总干排
水质类型	I	I	IV	I	II	V	V	V	IV

由评价结果看,博斯腾湖的源流开都河及其河口水质良好为 I 类,博斯腾湖水出流到孔雀河的水质也为 I 类,说明博斯腾湖水质总体是好的;黄水

区、黄水总干排向湖区排水水质为 IV 类,团结总干排、东风干排、胜利干排的排水口附近水质为 V 类,水质较差,污染物主要是农田排水带来的,其中包括农田废水、工业废水和城市生活污水。博斯腾湖水质评价结果与文献资料 and 实际水质基本状况是一致的,说明属性识别模型是可行的,可用于水质评价。

3 结 论

博斯腾湖水质状况总体上是好的,但在有排水口处,水污染严重,水质较差,今后应加强有排水区域的水质监测和综合治理。利用熵权来确定权重的方法更符合客观实际,可提高水质评价模型结果的精度。属性识别模型是根据样本信息本身来评价水质,能减少人为的主观因素对评价结果的影响,评价过程简单,评价结果可靠,是进行湖泊水质评价的一种实用的方法。

参考文献:

- [1] 门宝辉,梁川,付强. 湖库富营养化综合评价的属性识别模型[J]. 四川大学学报:自然科学版, 2002, 34(6):109-111.
- [2] 程乾生. 属性识别理论模型及其应用[J]. 北京大学学报:自然科学版, 1997, 33(1):12-20.
- [3] 刘开弟,庞彦军,张博文. 水环境质量评价的未确知测度模型[J]. 环境工程, 2000, 18(2):56-60.
- [4] 夏军,左其亭,邵民诚. 博斯腾湖水资源可持续利用[M]. 北京:科学出版社, 2003:200-400.
- [5] 胡汝骥,马虹,樊自立. 近期新疆湖泊变化所示的气候趋势[J]. 干旱区资源与环境, 2002, 17(1):22-26.
- [6] 刘文祥,李喜俊,郭海燕. 新疆博斯腾湖水环境容量研究[J]. 环境科学研究, 1999, 12(1):35-38.

(收稿日期 2005-04-25 编辑 舒 建)

(上接第 20 页)

- [2] 韩宇平,阮本清. 区域水安全评价指标体系初步研究[J]. 环境科学学报, 2003, 23(2):267-272.
- [3] 贾绍凤,张军岩,张士锋. 区域水资源压力指数与水资源安全评价指标体系[J]. 地理科学进展, 2002, 21(6):538-545.
- [4] 张巧显,欧阳志云,王如松,等. 中国水安全系统模拟及对策比较研究[J]. 水科学进展, 2002, 13(5):569-577.
- [5] 汪松年,阮仁良. 上海市水资源普查报告[M]. 上海:上海科学技术出版社, 2001.
- [6] 陈东景,徐中民. 西北内陆河流域生态安全评价研究——以黑河流域中游张掖地区为例[J]. 干旱区地理, 2002, 25(3):219-224.
- [7] 左伟,王桥,王文杰,等. 区域生态安全评价指标与标准

研究[J]. 地理学与国土研究, 2002, 18(1):67-71.

- [8] 上海市统计局. 上海统计年鉴 1999-2003 年[M]. 北京:中国统计出版社, 2004.

(收稿日期 2005-06-08 编辑 舒 建)

