

塔里木河水质模糊模式识别评价

周维博 郭小砾

(长安大学环境科学与工程学院, 陕西 西安 710054)

摘要 :以塔里木河水质为研究对象,采用模糊模式识别评价方法对污染程度、水质类别、分类界限等多层次、多因素的模糊边界问题进行了综合评价。评价结果表明,将相对隶属度应用到河流水质模糊评价中,克服了以往用单一绝对数据进行水质类别划分的缺点,能比较科学、实际、有效地利用监测数据对地表水的污染程度进行综合评价。

关键词 :水质评价;相对隶属度;模糊模式;识别模型;塔里木河

中图分类号 :X824 文献标识码 :A 文章编号 :1004-693X(2007)04-0033-04

Fuzzy pattern recognition of water quality in Tarim River

ZHOU Wei-bo, GUO Xiao-li

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

Abstract :Taking Tarim River as a case, the pollution degree, water quality grade, grade bound, and other fuzzy boundary problems of different hierarchy and factors were assessed by fuzzy pattern recognition model. Results showed that the application of relative membership degree to water quality evaluation avoided the shortcomings of water quality classification by a single absolute data. Therefore, such approach is more scientific, practical and effective in the comprehensive evaluation of surface water pollution using monitoring data.

Key words :water quality evaluation; relative grade of membership; fuzzy pattern recognition model; Tarim River

随着社会、经济的发展,加上人类活动等综合影响,塔里木河干流水质已受到不同程度的污染,并且污染的程度正在不断加剧,河流中化学需氧量(COD)、氟化物(以 F^- 计)等指标均出现超标,有时甚至超标数倍。在传统的水质评价中,往往采用绝对数据对水质类别进行划分,从而直接得出最劣的即为水质类别的判断结果。但实际的水质监测数据分析中,水质污染程度、水质类别、分类界限等方面均存在模糊边界问题,用传统绝对数据进行判断,依此确定水质级别,恰好忽略了这些模糊边界问题的客观存在。本文运用模糊模式识别评价方法对污染程度、水质类别、分类界限等客观存在的模糊概念与模糊现象进行水质综合评价。应用模糊集理论进行河流水质评价,能够客观地反映水质的实际状况,充分体现模糊模式识别评价的科学性、客观性。

1 模糊模式识别原理

1.1 确定目标特征值矩阵和指标标准特征值矩阵

假设需要对一模糊集进行识别的 n 个样本组成的集合,有 m 个指标特征值表示样本的整体特征,则有样本集的指标特征值矩阵 X ,若样本集依据 m 个指标按 c 个状态或级别的已知指标标准特征值进行识别,则有指标标准特征值矩阵 $Y^{[1]}$ 。

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} = (x_{ij}) \quad (1)$$

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1c} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mc} \end{bmatrix} = (y_{ih}) \quad (2)$$

式中 x_{ij} 为样本 j 指标 i 的特征值($i = 1, 2, \dots, m; j$

作者简介:周维博(1956—),男,陕西乾县人,教授,博士生导师,主要从事水资源与水环境及节水灌溉方面的教学与研究工作。

E-mail: zwbzyz823@163.com

$(i = 1, 2, \dots, m)$; y_{ic} 为状态或级别 h 指标 i 标准特征值 ($h = 1, 2, \dots, c$)

$$e_i = \frac{c_i}{s_i} \quad (9)$$

1.2 确定相对隶属度矩阵

若规定 I 级水质为优良水质, 其评价指标 I 级的标准浓度为 y_{i1} 对于污染程度的相对隶属度 $s_{i1} = 0$; c 级水质为重度污染, 其指标 i 的标准浓度 y_{ic} 对于污染程度的相对隶属度 $s_{ic} = 1$, 则河流水体 h 级污染程度指标 i 的标准浓度 y_{ih} 的相对隶属度 s_{ih} 可根据式 (3) 确定:

$$s_{ih} = \frac{y_{ih} - y_{i1}}{y_{ic} - y_{i1}} \quad (3)$$

类似地, 将指标值越大污染程度越重的指标实测浓度 x_{ij} 按式 (4) 变为对应的相对隶属度 r_{ij} , 将指标值越小而污染程度越重的指标实测浓度 x_{ij} 按式 (5) 变为对应的相对隶属度 r_{ij} 。

$$r_{ij} = \begin{cases} 1 & x_{ij} > y_{ic} \\ \frac{x_{ij} - y_{i1}}{y_{ic} - y_{i1}} & y_{i1} \leq x_{ij} \leq y_{ic} \\ 0 & x_{ij} < y_{i1} \end{cases} \quad (4)$$

$$r_{ij} = \begin{cases} 1 & x_{ij} < y_{ic} \\ \frac{x_{ij} - y_{i1}}{y_{ic} - y_{i1}} & y_{ic} \leq x_{ij} \leq y_{i1} \\ 0 & x_{ij} > y_{i1} \end{cases} \quad (5)$$

运用式 (3)、(4)、(5) 可将样本集的指标特征值矩阵 X 、指标标准特征值矩阵 Y 变换为对应的相对隶属度矩阵 R 、 S :

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix} = (r_{ij}) \quad (6)$$

$$S = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1c} \\ s_{m1} & s_{m2} & \dots & s_{mc} \end{bmatrix} = (s_{ih}) \quad (7)$$

1.3 确定综合权重矩阵

相对隶属度矩阵 R 只反映了各项参数对应评价等级的相对隶属程度, 并不能直接反映不同因子、因素对整体环境的影响, 因此要引入权重来进行判断。计算权重的方法很多, 例如层次分析法、专家评价法、加权法等。本文采用因子污染分担率^[2]和相关系数^[3]两种方法分别对各环境因素的污染因子权重进行计算。

1.3.1 因子污染分担率法

计算公式:

$$w_i = \frac{e_i}{\sum_{i=1}^m e_i} \quad (i = 1, 2, \dots, m) \quad (8)$$

$$s_i = \frac{1}{n} (s_{i1} + s_{i2} + \dots + s_{in}) \quad (10)$$

式中: w_i 为因素权重; e_i 为第 i 种污染物的污染指数; c_i 为第 i 种污染物浓度; s_i 为第 i 种污染物的平均标准值; s_1, s_2, \dots, s_n 为第 i 种污染物各级别的标准值(地表水水质分级一般分 5 级)。

1.3.2 相关系数法

计算公式:

$$v = (v_1, v_2, \dots, v_m) = \left[\frac{|r_1|}{\sum_{i=1}^m |r_i|}, \frac{|r_2|}{\sum_{i=1}^m |r_i|}, \dots, \frac{|r_m|}{\sum_{i=1}^m |r_i|} \right] \quad (11)$$

$$\sum_{i=1}^m v_i = 1$$

式中: v_1, v_2, \dots, v_m 为指标本身权重; r_1, r_2, \dots, r_m 为某种污染物指标与其余 m 个评价指标的相关系数。

综合考虑超标权重与指标权重, 建立 n 个样本 m 项指标的综合权重矩阵^[4]

$$A = \begin{bmatrix} v_1 & 0 & \dots & 0 \\ 0 & v_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & v_m \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix} = \begin{bmatrix} v_1 r_{11} & v_1 r_{12} & \dots & v_1 r_{1n} \\ v_2 r_{21} & v_2 r_{22} & \dots & v_2 r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_m r_{m1} & v_m r_{m2} & \dots & v_m r_{mn} \end{bmatrix} = (v_i r_{ij}) \quad (12)$$

将矩阵 A 的元素按列归一化, 得到

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{m1} & w_{m2} & \dots & w_{mn} \end{bmatrix} = (w_{ij}) \quad (13)$$

$$w_{ij} = \frac{v_i r_{ij}}{\sum_{i=1}^m v_i r_{ij}}$$

$$\sum_{i=1}^m w_{ij} = 1, \forall j$$

1.4 确定模糊模式识别模型

将 R 、 S 、 W 中数据代入模型, 可判断出研究对象属于哪一个模式或标准样本。模糊模式识别模型^[1]:

$$u_{hj} = \begin{cases} 0 & h < a_j \text{ 或 } h > b_j \\ \frac{1}{\sum_{k=a_j}^{b_j} \left\{ \frac{\sum_{i=1}^m [w_{ij}(r_{ij} - s_{ih})]^p}{\sum_{i=1}^m [w_{ij}(r_{ij} - s_{ik})]^p} \right\}^{\frac{2}{p}}} & a_j \leq h \leq b_j \\ & d_{hj} \neq 0 \\ 1 & d_{hj} = 0 \end{cases} \quad (14)$$

1.5 水质级别的确定

根据模糊概念在分级条件下的最大隶属度原则不适用性^[1],采用级别特征值^[1]的方法对水质类别进行判断:

$$H = (1 \ 2 \ 3 \ \dots \ c) U_{c \times n} \quad (15)$$

H 为样本的级别特征值矩阵。根据 H 对水体样本 j 进行综合评价,考虑到水质评价中的区间值类别类型,对上述结果做如下处理,对水体质量进行归类^[5]:

$c - 1 \leq H_j < c$, 令 $H_j = c - 1$, 即将样本 j 归于 $c - 1$ 类。

2 塔里木河水水质模糊模式识别评价

2.1 确定实测矩阵和标准矩阵

选取塔里木河比较有代表性的 4 个水质监测断面作为评价样本。由于从多次监测结果看,其超标污染物种类基本比较稳定,因此取部分代表性的评价指标进行综合评价,能够更充分体现塔里木河水水质实际情况,其评价结果比较有说服力。选取溶解氧(DO)、高锰酸盐指数(COD_{Mn})、化学需氧量(COD)、五日生化需氧量(BOD₅)、总氮(TN)、氟化物(以 F⁻ 计)的实测数据组成 X 矩阵。参照 GB 3838—2002《地表水分级标准环境质量标准》,建立评价标准矩阵 Y 。

2.2 确定相对隶属度矩阵

通过运用式(3)(4)(5),将实测数据矩阵 X 、标准矩阵 Y 分别转换为相应的相对隶属度矩阵 R 、 S 。矩阵 R 表示塔里木河 4 个监测断面的 6 个指标实测浓度值对于模糊概念“污染”的作用大小或影响程度,同时也表达了全体监测断面(样本)全部评价指标(指标)对于“污染”的相对隶属度。由于 I 级污染程度标准浓度值的相对隶属度为 0,故矩阵 R 给出了监测断面(样本集)的全体指标超过 I 级污染程度(优良水质)标准浓度的相对隶属度值。因此相对隶属度在模糊集合论中定义为权重,故矩阵 R 为超标权重矩阵,矩阵 S 则表示地表水标准级别的相对隶属度矩阵。

$$R = \begin{bmatrix} 0.313 & 0.295 & 0 & 0.356 \\ 0.008 & 0.002 & 0.176 & 0.190 \\ 0 & 0 & 0 & 0.358 \\ 0 & 0.066 & 0 & 0 \\ 0.167 & 0.028 & 0.328 & 0.406 \\ 0.6 & 0.6 & 0.4 & 0 \end{bmatrix}$$

$$S = \begin{bmatrix} 0 & 0.273 & 0.455 & 0.818 & 1 \\ 0 & 0.154 & 0.308 & 0.615 & 1 \\ 0 & 0 & 0.2 & 0.6 & 1 \\ 0 & 0 & 0.143 & 0.429 & 1 \\ 0 & 0.167 & 0.444 & 0.722 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

2.3 权重的计算

根据塔里木河的实测数据分析,所选取的部分评价指标对于塔里木河水水质污染程度的影响是不同的,因而还要对塔里木河不同监测断面、不同评价指标进行整体的综合性衡量,确定断面、评价指标的综合权重。 W 即为环境指标综合权重矩阵,它既考虑了指标的超标权重,同时也计入了指标本身的权重,综合表示了监测断面(样本集)各个指标对塔里木河流域水质污染作用和影响的程度。

2.3.1 污染分担率法

计算结果见表 1。

表 1 污染分担率法综合权重求解结果

污染物	断面 1		断面 2		断面 3		断面 4	
	e_i	a_i	e_i	a_i	e_i	a_i	e_i	a_i
DO	1.230	0.322	1.251	0.336	1.617	0.343	1.179	0.261
COD _{Mn}	0.285	0.075	0.274	0.074	0.580	0.123	0.604	0.134
COD	0.208	0.054	0.208	0.056	0.446	0.095	0.998	0.221
BOD ₅	0.535	0.140	0.665	0.179	0.310	0.066	0.427	0.094
TN	0.481	0.126	0.240	0.064	0.760	0.161	0.894	0.198
F ⁻	1.083	0.283	1.083	0.291	1	0.212	0.417	0.092
$\sum e_i$	3.822	1	3.721	1	4.713	1	4.519	1

由表 1 可知,综合权重矩阵

$$W = \begin{bmatrix} 0.322 & 0.336 & 0.343 & 0.261 \\ 0.075 & 0.074 & 0.123 & 0.134 \\ 0.054 & 0.056 & 0.095 & 0.221 \\ 0.140 & 0.179 & 0.066 & 0.094 \\ 0.126 & 0.064 & 0.161 & 0.198 \\ 0.283 & 0.291 & 0.212 & 0.092 \end{bmatrix}$$

2.3.2 相关系数法

若选定 $r_1 = 1$ (溶解氧 DO 与自己的相关系数为 1), $r_2 = 0.396$, $r_3 = 0.471$, $r_4 = 0.221$, $r_5 = 0.229$, $r_6 = 0.151$, 由式(11)得指标权向量 $v = (0.405, 0.160, 0.191, 0.090, 0.093, 0.061)$, 由式(12)得矩阵 A 。并按式(13)归一化,得指标综合权重矩阵 W 。

$$A = \begin{bmatrix} 0.1268 & 0.1195 & 0 & 0.1442 \\ 0.0013 & 0.0003 & 0.0282 & 0.0304 \\ 0 & 0 & 0 & 0.0684 \\ 0 & 0.0059 & 0 & 0 \\ 0.0155 & 0.0026 & 0.0305 & 0.0378 \\ 0.0366 & 0.0366 & 0.0244 & 0 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.7037 & 0.7246 & 0 & 0.5135 \\ 0.0072 & 0.0018 & 0.3394 & 0.1083 \\ 0 & 0 & 0 & 0.2436 \\ 0 & 0.0358 & 0 & 0 \\ 0.0860 & 0.0158 & 0.3670 & 0.1346 \\ 0.2031 & 0.2220 & 0.2936 & 0 \end{bmatrix}$$

2.4 模糊模式识别评价

根据模糊模式识别模型,即式(14),取 $p = \alpha$ (欧氏距离)^[1],可对塔里木河的实际水质状况进行模糊模式识别评价。运用2种不同的权重计算方法,均可计算出不同断面(样本)对于不同污染物种类(指标)的相对隶属程矩阵 U_1 、 U_2 。

$$U_1 = \begin{bmatrix} 0.232 & 0.236 & 0.489 & 0.075 \\ 0.314 & 0.309 & 0.307 & 0.183 \\ 0.275 & 0.278 & 0.157 & 0.694 \\ 0.179 & 0.177 & 0.047 & 0.048 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$U_2 = \begin{bmatrix} 0.1244 & 0.1435 & 0.1971 & 0.0637 \\ 0.5057 & 0.5053 & 0.3621 & 0.2590 \\ 0.3114 & 0.2913 & 0.3565 & 0.6374 \\ 0.0585 & 0.0599 & 0.0843 & 0.0429 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

根据式(15)计算得到样本集的级别特征值向量 H_1 、 H_2 ,从而对河流水质进行级别归类。

$$H_1 = (2.401, 2.396, 1.762, 2.715)$$

$$H_2 = (2.304, 2.2676, 2.328, 2.6655)$$

将采用不同方法计算权重所得的水质模糊模式识别评价进行比较,水质类别见表2。

表2 模糊模式识别评价结果

权重计算方法	断面1	断面2	断面3	断面4
因子污染分担率法	Ⅲ类水	Ⅲ类水	Ⅱ类水	Ⅲ类水
相关系数法	Ⅲ类水	Ⅲ类水	Ⅲ类水	Ⅲ类水

采用2种计算权重的方法进行塔里木河水质综合评价,其评价结果基本一致,仅个别断面略有差异。运用模糊模式识别模型对水质进行综合评价的优越性主要体现在以下几个方面:

a. 指出了仅用单一的绝对数据对水质类别进行划分的缺点,同时也弥补了运用最大隶属度对水质类别进行划分中存在的不足。

b. 不仅考虑了不同评价指标、不同监测断面对河流水质污染的相对隶属程度,同时也考虑了它们之间对污染程度的相互作用和影响。

c. 能够从宏观即河流的整体状况去把握水质类别及污染程度,从而能够对整个流域的水质进行综合评价。

3 结论

由于传统的水质评价方法不能够准确、客观处理模糊边界问题,从而对评价结果产生偏差和影响。有鉴于此,本研究将相对隶属度应用到河流水质模糊评价中。通过对评价河流监测断面不同指标的相对隶属度的判断,不仅能得出评价断面的水质类别,而且能反映出河流整体污染程度,从而得出水质类别。这样能够更加科学、实际、有效地利用监测数据对地表水的污染程度进行综合评价,同时也弥补了现行水质评价中用单一指标作为分界线的简单评价法中的缺陷。

参考文献:

- [1] 陈守煜. 工程模糊集理论与应用[M]. 北京: 国防工业出版社, 1998: 159-163.
- [2] 吴超, 廖国礼. 矿区总体环境质量模糊综合评价实践[J]. 矿业研究与开发, 2004(4): 60-63.
- [3] 盛骤, 谢士干, 潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 1979: 118-122.
- [4] 陈守煜, 熊德琪. 湖泊富营养化评价模糊集理论与模式[J]. 湖泊科学, 1993(2): 145-152.
- [5] 陈守煜, 李亚伟. 基于模糊神经网络识别的水质评价模型[J]. 水科学进展, 2005, 16(1): 88-91.

(收稿日期 2006-02-25 编辑: 傅伟群)

(上接第22页)

- [48] GOPAL B, GOEL U. Competition and allelopathy in aquatic plant communities[J]. Bot Rev, 1993, 59: 155-210.
- [49] GRACE J B, WETZEL R G. Habitat partitioning and competitive displacement in cattails (Typha): experimental field studies[J]. Am Nat, 1981, 118: 463-474.
- [50] GRACE J B, WETZEL R G. Long-term dynamics of Typha populations[J]. Aquat Bot, 1998, 61: 137-146.
- [51] 李文朝. 富营养水体中常绿水生植被组建及净化效果研究[J]. 中国环境科学, 1997, 17(1): 53-57.
- [52] 高吉喜. 水生植物对面源污水净化效率研究[J]. 中国环境科学, 1997, 17(3): 247-251.
- [53] 张运林, 秦伯强, 陈伟民, 等. 太湖水体透明度的分析、变化及相关性分析[J]. 海洋湖沼通报, 2003(2): 30-36.
- [54] 白峰青, 郑丙辉, 田自强. 水生植物在水污染控制中的生态效应[J]. 环境科学与技术, 2004, 27(4): 99-100.
- [55] 江行玉, 赵可夫. 植物重金属伤害及其抗性机理[J]. 应用与环境生物学报, 2001, 7(1): 92-99.
- [56] 林昱, 林荣澄. 厦门西港引发有害硅藻水华磷的阈值研究[J]. 海洋与湖沼, 1999, 30(4): 391-396.
- [57] 夏汉平. 垃圾污水的植物毒性与植物净化效果之研究[J]. 植物生态学报, 1999, 23(4): 289-301.

(收稿日期 2005-11-30 编辑: 傅伟群)