

# 三种权重识别方法在同安湾水质模糊评价中的应用比较

陈克亮<sup>1</sup>, 王金坑<sup>1</sup>, 朱晓东<sup>2</sup>, 赵文珺<sup>2</sup>, 李杨帆<sup>2</sup>, 赵艳<sup>2</sup>

(1. 国家海洋局 第三海洋研究所, 福建 厦门 361005; 2. 污染控制与资源化研究国家重点实验室, 南京大学环境学院, 江苏 南京 210093)

**摘要:**以厦门同安湾为例, 利用 1995~2005 年的监测数据, 分别采取污染物浓度超标加权法、主成分分析法和对比排序法三种权重识别方法对近岸海域水质时间序列上的演变规律进行模糊评价的比较。结果表明, 三种权重识别方法的评价结果一致, 均可作为同安湾水质模糊评价的权重确定方法; 同时, 还分析了三种权重确定方法在应用过程中的优缺点, 并进行了比较, 为海湾水质模糊评价中的权重确定提供参考。

**关键词:** 权重识别方法; 水质; 模糊评价

中图分类号: X55

文献标识码: A

文章编号: 1000-3096(2010)05-0033-04

水质综合评价中的水质类别、分类界限是人为规定的明确概念, 而客观存在的水体污染程度是一定数值范围内的“模糊概念”, 模糊数学的突出特点在于对事物的辨别与评判, 应用模糊理论进行水质评价, 能更客观地反映水质的实际状况。因此有关模糊评判作为水环境评价常用的一种方法, 在水环境综合评价中得到广泛应用<sup>[1]</sup>。

用模糊数学方法进行水质分析首先要解决两个关键问题<sup>[2]</sup>: 一是评价指标的隶属度分析与计算; 二是各评价指标的权重分配。文献[3]应用模糊聚类和模糊贴近度的概念初步建立了模糊数学评价水质的理论和方法; 文献[4~7]考虑到各单项指标对环境污染的贡献不同, 采用污染物超标法确定各评价指标的权重系数, 此种方法是目前最为常用的一种确定水质各评价指标权重的方法。在上述研究的基础上, 作者根据陈守煜提出的模糊模式识别理论建立模糊评价矩阵, 模糊识别模型中的关键问题是怎样确定各个评价指标的权重。采用三种方法确定各水质指标的权重, 以改进和优化评价指标权重的确定方法, 最后用模糊识别模型判断水样所属的水质级别。

作者在分析厦门同安湾近岸海域水环境质量变化的基础上, 探讨应用模糊模式识别模型进行水质综合评价, 为近岸海域水质评价方法的优化和完善提供理论依据。

## 1 模糊识别模型与方法

### 1.1 模糊模式识别模型原理

根据文献[1,3,7], 建立模糊模式识别模型。

设系统的指标特征值向量为:  $X=(x_1, x_2, x_3, \dots, x_n)$   
 $n$  表示指标数。

系统指标按  $C$  个级别进行识别,  $C$  个级别的指标标准特征值为:  $Y=(y_1, y_2, \dots, y_c)$ , 系统的  $n$  个指标特征值的相对隶属度  $R$  与指标标准特征值的相对隶属度  $S$ , 将系统的  $n$  个指标特征值的相对隶属度  $R=(r_1, r_2, \dots, r_n)$ , 分别与  $S=(s_1, s_2, \dots, s_c)$  进行比较, 可得系统指标  $i$  的级别上限值  $b$ , 级别下限值  $a$ 。

根据目标函数、约束条件构造拉格朗日函数, 将等式约束求极值变为求无约束极值问题, 则最终系统对于级别  $h$  的相对隶属度公式为:

$$u_{hj} = \begin{cases} 0 & h < a \text{ 或 } h > b \\ \frac{1}{\left[ \sum_{i=1}^n w_i (r_{ij} - s_{ih}) \right]^2 \times \sum_{b=1}^b \left\{ \frac{1}{\left[ \sum_{i=1}^n w_i (r_{ij} - s_{ih}) \right]^2} \right\}} & a \leq h \leq b \\ 1 & d_h = 0 \end{cases} \quad (1)$$

则级别特征值公式为:

$$H=(1, 2, \dots, c) \times (u_1, u_2, \dots, u_c)^T \quad (2)$$

式中,  $H$  为系统级别特征值, 用以衡量系统的真实

收稿日期: 2009-04-25; 修回日期: 2009-08-02

基金项目: 国家海洋局海洋生态环境科学与工程重点实验室开放基金项目(MESE-2009-06); 国家海洋局海洋公益性行业科研专项经费项目(200805065); 国家自然科学基金项目(40976021, 40901081)

作者简介: 陈克亮(1978-), 男, 河南西平人, 博士, 助理研究员, 主要从事海洋环境管理研究, 电话: 0592-2195513, E-mail: klchen@yeah.net

状态。

## 1.2 权重的确定方法

污染浓度超标法需对每次样本进行权系数的运算;主成分分析法和对比排序法各指标采用固定的权系数,克服了权系数确定随实测水样不同而变化的缺点,便于对不同样本的评价结果进行对比分析,同时简化了多次重复确定权系数的工作量,符合一般评价方法的最基本要求。作者利用这三种权重确定方法进行水质模糊评价结果的比较。

### 1.2.1 污染物浓度超标加权法

根据文献[8],采用污染物浓度超标加权法确定权重:

$$p_i = C_i / C_{oi} \quad (3)$$

式(3)中,  $C_i$  为第  $i$  种污染物在水中的实测浓度,  $C_{oi}$  为第  $i$  种污染物作为某种用途时的浓度允许值。  $p_i$  表明了环境中污染物浓度超标的倍数,超标倍数越大,对环境污染贡献越大。加权就是根据这个理由引进的,其权重  $w_i = p_i$ ,  $w_{DO} = \overline{C_{oi}} / C_i$ 。

对于同种用途的水体分成若干个等级,因此,允许浓度就有若干个数值,这时在求  $w_i$  时,  $C_{oi}$  值可取平均值。如分  $m$  个等级等,则  $C_{oi} = \frac{1}{m} (C_{oi} + C_{oi} + C_{oi} + \dots + C_{oi}^m)$ , 对各因子所加的权值为超标倍数。对于所求得的权数,为了进行模糊变换,  $w_i$  应满足归一化要求。

### 1.2.2 主成分分析法

根据文献[9],假设有  $n$  个样本构成的一个环境系统,每个样本由  $p$  个变量  $x_1, x_2, x_3, \dots, x_p$  来表征,构成原始数据矩阵  $X$ , 即  $X = (x_{ij})_{n \times p}$ 。

(1) 计算有关统计量,建立变量的相关矩阵

在主成分分析之前,必须将原始数据进行标准化,为此需要计算出变量平均值  $\bar{x}_j$ 、标准差  $S_j$  与相关系数矩阵  $R$ 。

$$\bar{x}_j = \frac{1}{n} \sum_{k=1}^n x_{kj} \quad (j=1,2,\dots,p) \quad (4)$$

$$S_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2} \quad (j=1,2,\dots,p) \quad (5)$$

$$R = (r_{ij})_{p \times p} \quad (6)$$

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \quad (j=1,2,\dots,p) \quad (7)$$

(2) 求  $R$  矩阵的特征根  $\lambda_i$  及相应的标准正交特征向量  $u_i$

一般可采用 Jacobi 法求出  $R$  的所有特征根并依大小排列,即  $\lambda_1 \lambda_2 \dots \lambda_p > 0$ , 并且各特征根之和等于矩阵  $R$  的迹,即  $\sum_{i=1}^p \lambda_i = tr(R) = p$ 。则  $\lambda_i$  对应的标准正交特征向量为:

$$u_i = (u_{1i}, u_{2i}, \dots, u_{pi})' \quad (8)$$

(3) 求出负荷量

一般由  $l_{ij} = \sqrt{\lambda_i} u_{ji}$ , 可求出  $X$  的各个原变量  $x_j$  对主分量  $y_i$  的负荷量,列成矩阵的形式为:

$$L = U \Lambda^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\lambda_1} u_{11} & \dots & \sqrt{\lambda_p} u_{1p} \\ \vdots & \ddots & \vdots \\ \sqrt{\lambda_1} u_{p1} & \dots & \sqrt{\lambda_p} u_{pp} \end{pmatrix} \quad (9)$$

(4) 确定各变量的权重

在负荷阵  $L$  中,令  $h_i^2 = \sum_{j=1}^2 l_{ij}^2$ , 将  $h_i^2$  归一化,就

得到各变量的权重。

关于权重的确定采用能够自动生成信息量权数的主成分分析法,以避免人为因素的影响。

### 1.2.3 对比排序法

根据文献[10],采用对比排序法确定权重的方法,根据水质指标在专家心目中的重要性排序调查统计获得,最不重要的排在第 1 位,其次排第 2 位,.....,最重要排在第  $n$  位, 1, 2, ...,  $n$  为各因素的秩次,各因素的权重由下式计算:

$$w'_j = \frac{\sum_{k_j=1}^n (f_{k_j} \log m^{k_j})}{\sum_{k_j=1}^n f_{k_j}} \quad (10)$$

$$w_j = \frac{w'_j}{\sum w'_j} \quad (11)$$

式中,  $m$  为评价因素的个数;  $k_j$  为第  $j$  个因素的秩次;  $f_{k_j}$  为  $k_j$  出现的频数。

## 2 同安湾近岸海域水质的模糊模式识别分析

监测资料(表 1)表明,同安湾近岸海域主要污染物为溶解氧(DO)、无机氮(DIN)、石油类、化学耗氧量(COD)和活性磷酸盐(DRP),因此选择这五项指标进行海域水质分析。

表 1 同安湾近岸海域主要污染物监测数据

Tab. 1 The monitoring data of the main contaminants in Tong'an Bay (mg/L)

监测数据	年份										
	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
DO	7.04	6.75	6.89	8.39	7.35	6.6	7.27	5.96	6.61	6.19	6.99
COD <sub>Mn</sub>	1.34	1.16	0.85	1.34	1.17	0.78	0.72	1.09	0.69	0.84	0.54
DIN	0.200	0.179	0.240	0.414	0.305	0.234	0.38	0.44	0.407	0.531	0.412
DRP	0.013	0.016	0.020	0.023	0.017	0.018	0.034	0.046	0.029	0.033	0.031
石油类	0.012	0.043	0.034	0.031	0.030	0.035	0.040	0.01	0.009	0.024	0.044

评价标准中, 级别特征值(1, 2, 3, 4)对应于《中华人民共和国海水水质标准》(GB3097-1997)中第四类海水水质标准。采用模糊模式识别模型对同安湾近岸海域进行海域水质分析。

同安湾海域, 各指标特征值(1996~2005 年)对于标准级别的相对隶属度  $R_{ij}$ 、标准级别的相对隶属度  $S_{ih}$ , 计算为:

$$R_{ij} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0.987 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0.93 & 0.616 & 0.767 & 0.407 & 0.4 & 0.2 & 0.31 & 0 & 0.293 \\ 0.967 & 0.833 & 0.7 & 0.7 & 0.9 & 0.367 & 0 & 0.533 & 0.4 & 0.467 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$S_{ih} = \begin{bmatrix} 1 & 0.667 & 0.333 & 0 \\ 1 & 0.667 & 0.333 & 0 \\ 1 & 0.667 & 0.333 & 0 \\ 1 & 0.5 & 0.5 & 0 \\ 1 & 1 & 0.444 & 0 \end{bmatrix}$$

$R_{ij}$  为  $i$  指标对于第  $j$  年的相对隶属度( $i=1,2,3,4,5$ , 依次为 DO、COD、DIN、DRP、石油类;  $j = 1996, 1997, \dots, 2005$  年);

隶属度( $h=1,2,3,4$ ),

根据 1.2 中权重的确定方法, 经计算, DO、COD、DIN、DRP、石油类五项指标的权重依次为:

$S_{ih}$  为标准级别中  $i$  指标相对于  $h$  类海水的相对

(1) 超标污染法权重:

$$w = \begin{bmatrix} 0.50 & 0.47 & 0.39 & 0.41 & 0.42 & 0.38 & 0.30 & 0.38 & 0.32 & 0.38 \\ 0.10 & 0.08 & 0.08 & 0.08 & 0.06 & 0.05 & 0.07 & 0.05 & 0.06 & 0.04 \\ 0.16 & 0.19 & 0.25 & 0.22 & 0.30 & 0.26 & 0.28 & 0.30 & 0.35 & 0.29 \\ 0.18 & 0.20 & 0.22 & 0.23 & 0.17 & 0.27 & 0.34 & 0.25 & 0.25 & 0.25 \\ 0.06 & 0.05 & 0.05 & 0.06 & 0.05 & 0.04 & 0.01 & 0.01 & 0.02 & 0.05 \end{bmatrix}$$

(2) 对比排序法权重:  $w = (0.22, 0.20, 0.21, 0.24, 0.13)$

(3) 主成分分析权重:  $w = (0.17, 0.23, 0.23, 0.21, 0.16)$

根据各方法确定的权重, 计算环境级别的特征值分别为:

$$H_1 = (1,2,3,4) \times U_1 = (1.0017, 1.0732, 1.6407, 1.4043, 1.8736, 2.0877, 2.6090, 2.1246, 2.5751, 2.2991)$$

$$H_2 = (1,2,3,4) \times U_2 = (1.0048, 1.1767, 1.7297, 1.5709, 1.7651, 2.1007, 2.3562, 2.0681, 2.2844, 2.2381)$$

$$H_3 = (1,2,3,4) \times U_3 = (1.0038, 1.1320, 1.7400, 1.5437, 1.8452, 2.0787, 2.3166, 2.0745, 2.3034, 2.2420)$$

运用 3 种权重确定方法得出的环境级别特征值比较如图 1, 3 种方法计算结果差别不大, 从 1996~2001 年水质较好的年份, 超标法计算出来的环境级别特征值略低于其他两类, 从 2002~2006 年, 水质恶化趋势明显, 此时, 超标法计算出的级别特征值高于其他两种方法。从方法上来讲, 主成分分析法避免了确定评价指标权重的主观任意性, 且与其他两种方法比较来看, 没有太大的差异。通过相似性分析和显著性检验:  $H_1-H_2$ 、 $H_1-H_3$  和  $H_2-H_3$  组间相关系数( $R$ )分别为 0.991, 0.990 和 0.998。T 检验结果发现 3 组数据间无显著性差异,  $P=0.905>0.10$ , 说明结

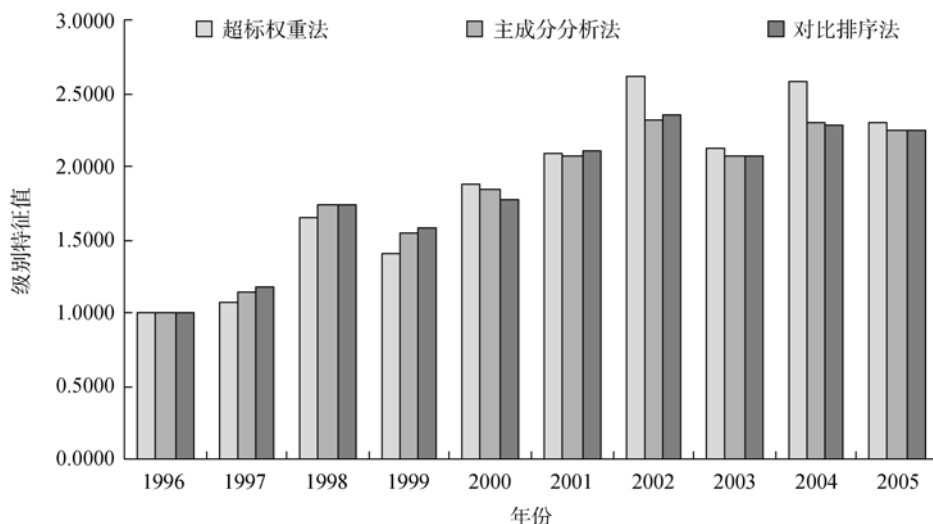


图 1 3 种权重确定方法下的水质级别特征值比较

Fig. 1 Comparison of water quality characteristics values obtained by three recognition methods of weight coefficient

果一致性较强。

从图 1 各环境级别特征值来看, 从 1996 年即同安湾近岸海域水质呈现逐年恶化的现象, 近几年越来越接近于三类海水, 从 1996 年的一类水到 2002 年超二类水, 逐年上升, 同安湾 2002 年海水级别特征值最大, 污染最严重。2003 年海水水质有所改善, 2004 年与 2005 年又反弹。从 2002 年到 2005 年基本变化不大, 级别值最高为 2.6090, 最低值为 2.0681。

根据标准级别的相对隶属度  $R_{ij}$  看出, DIN 和 DRP 是同安湾近岸海域的主要污染物, 在各年份中均有不同程度的超标现象。

### 3 结果与讨论

考虑到各级水质标准的界限模糊性, 根据模糊数学方法提出三种权重识别方法用于水质综合评价, 使评价更为接近客观实际。通过三种权重的识别方法得出同安湾水质模糊评价结果一致, 说明三种方法均可作为同安湾水质模糊评价权重确定的有效方法。

利用主成分分析法确定各指标的权重, 避免了确定评价指标权重的主观任意性; 污染浓度超标法是常用的水质评价方法, 采用超标污染浓度加权法来确定权重, 各指标的权系数随时间序列样本的不同而变化, 需要对同一时间所测样本进行权系数的运算来确定权重, 加大了工作量; 对比排序法增加了人为的主观因素。主成分分析法和对比排序法克服了权系数确定随实测水样不同而变化的缺点, 采

用各指标间固定的权系数, 符合一般评价方法的最基本要求, 便于对不同样本的评价结果进行对比分析, 同时也简化了多次重复确定权系数的工作量。

#### 参考文献:

- [1] 陈守煜, 赵英琪. 模糊模式识别理论模型与水质评价[J]. 水利学报, 1991, 6: 35-40.
- [2] 田景环, 邱林, 柴福鑫. 模糊识别在水质综合评价中的应用[J]. 环境科学学报, 2005, 25(7): 950-953.
- [3] 陈守煜. 模糊水文学与水资源系统模糊优化原理[M]. 大连: 大连理工大学出版社, 1990. 51-66.
- [4] 曾明荣, 王成海. 模糊数学在水质评价中的应用[J]. 福建环境, 1999, 16(5): 7-9.
- [5] 杨林, 李晓玲. 模糊数学在湟水流域水质综合评价中的应用[J]. 中国环境监测, 2000, 16(6): 49-52.
- [6] 梅学彬, 王福刚, 曹剑锋. 模糊综合评判法在水质评价中的应用及探讨[J]. 世界地质, 2000, 19(2): 172-177.
- [7] 陈守煜. 工程模糊集理论与应用[M]. 北京: 国防工业出版社, 1998. 24-35.
- [8] 付雁鹏, 高嘉瑞. 模糊数学在水质评价的应用[M]. 华中工学院出版社, 1986. 68-69.
- [9] 陈玉成, 吕宗清, 李章平. 环境数学分析[M]. 西南师范大学出版社, 1998. 109-113.
- [10] 胡永宏, 贺思辉. 综合评价方法[M]. 科学出版社, 2000. 188-189.

(下转第 47 页)