

DOI: 10.20098/j.cnki.1001-5132.2023.0334

基于犹豫模糊集的水质评价方法

曹磊^{1,2}, 戴俊杰², 毛小燕^{2*}, 蒋吉³, 刘锋琴³

(1. 宁波大学 信息科学与工程学院, 浙江 宁波 315211; 2. 宁波大学科学技术学院 信息工程学院, 浙江 宁波 315300;
3. 慈溪市水利局, 浙江 宁波 315399)

摘要: 为了更全面反映水体综合状态, 本文基于单因子评价水质等级的结果, 计算指标与等级的相关系数, 筛选出强相关的指标. 采用专家打分构造犹豫模糊水质评价初始矩阵, 建立基于犹豫模糊指数熵和线性插值悲观升维的加权犹豫模糊兰氏距离测度多属性评价方法. 最后, 结合浙江省部分水库水质数据进行实例分析. 结果表明, 所提出的基于犹豫模糊集距离测度的水质综合评价方法更为合理有效.

关键词: 水质评价; 犹豫模糊集; 距离测度; 升维

中图分类号: X524

文献标志码: A

文章编号: 1001-5132 (2024) 01-0051-07

随着人类社会经济的飞速发展, 水环境污染以及水生态退化等问题日益明显, 严重违反了可持续发展理念, 背离了习近平同志“绿水青山就是金山银山”的重要论断. 因此加强水环境治理已经成为各省市的当务之急, 而其中对水质的评价是所有事项中的重中之重, 只有对水质观测数据进行合理有效的评价才能引导制定科学有益的防治对策. 从这方面来说, 水质综合评价的合理性将直接影响管理者的决策^[1-2].

目前, 国内外常见的水质评价方法有单因子评价、神经网络评价、模糊综合评价等. 单因子评价是指采用目标水体中所有参与综合评价的水质指标, 选取其中最差的水质单项指标所属的类别作为该目标水体的类别^[3], 其对水质的要求过于保护, 无法客观地反映水体的综合状态. 人工神经网络(Artificial Neural Networks)是一种模仿动物神经网络行为特征, 通过调节内部大量节点参数及其权重信息从而实现其拟合函数、处理信息等功能的算法. 早在 20 世纪末, 就已经有学者将模型结构运用到水质评价任务中^[4-5]. 虽然神经网络具有很好的客观性, 精度也相较于其他方法要高, 但其训

练过程需要大量带标签样本数据, 且特征选取会极大地影响最终模型的效果. 近几年也有学者对神经网络予以改进, 用于水质评价研究^[6-7], 但也不能达到令人十分满意的效果. 由于水质评价是一个多指标影响的复杂过程, 且对污染程度的分级界限较为模糊, 传统方法很难刻画这种不确定性, 而模糊集理论为解决此类不确定性问题提供了新思路. 模糊综合评价^[8]关键在于隶属函数和指标权重的确定, 常通过构造高斯隶属函数^[9]或梯形隶属函数^[10]来确定监测数据与各等级间的隶属度, 有效地解决了评价标准边界模糊问题; 而权重的确立方法主要有熵权赋值法^[11]、超标倍数法^[12]及层次分析法^[13]. 在实际应用中, 决策者在确定各属性隶属度值时, 常会遇到多个数值之间难以取舍的问题. 为处理这类情况, Torra 等^[14-15]提出了犹豫模糊集概念, 其允许选定多个数值作为隶属度, 能够更真实地体现决策者犹豫不决的心理. 因此本文也首次将犹豫模糊集应用到水质评价领域.

本文采用一种基于犹豫模糊集的水质评价模型, 并提出一种新的基于线性插值的悲观补齐方案, 根据距离测度来评价水质等级. 通过浙江省部

收稿日期: 2023-03-21.

宁波大学学报(理工版)网址: <http://journalg.nbu.edu.cn/>

基金项目: 浙江省自然科学基金(LY20A010012).

第一作者: 曹磊, 硕士研究生, 主要研究方向: 数据挖掘与决策分析. E-mail: 1160842475@qq.com

*通信作者: 毛小燕, 副教授, 主要研究方向: 代数学、复杂信息决策理论与方法. E-mail: maoxiaoyan@nbu.edu.cn

分水库的实际算例,进行稳定性和有效性检验,以期丰富水质评价方法,为水资源监管提供客观科学依据。

1 指标筛选

研究数据来源于浙江省生态环境厅公布的地表水水质自动检测数据。根据国家《地表水环境质量标准》(GB 3838—2002)规定,基本的水质评价指标有 24 项。但由于指标过多且存在对评价结果参考意义较小的指标,因此需先对指标进行筛选和处理。

挑选慈溪市 7 处水体,采集共计 10 个月的监测样本,并收集相应的水质等级数据。水体成分和外界环境因素(下文称为特征)具体包含:水温、pH 值、溶解氧、高锰酸盐、总磷、总氮、硫化物等;对应的水质等级分别为 1 级至 6 级,等级越低表示水质越好,其中,6 级表示水质指标严重超标,又称劣 V 类,即超出地表水环境质量标准规定的五类水质。首先初步剔除那些存在缺失值以及无任何参考价值的特征(如气温、水温等),接着计算每个特征与水体等级之间的相关系数。水质特征及等级具体结果见表 1。

表 1 水质特征与水质等级的相关分析

特征	相关系数	对应 P 值
pH 值	-0.047	0.70
溶解氧	-0.247	0.04*
高锰酸盐	0.644	<0.01*
五日生化需氧量	0.746	<0.01*
氨氮	0.710	<0.01*
总磷	0.538	<0.01*
总氮	0.857	<0.01*
硝酸盐氮	0.352	0.01*
铁	-0.048	0.74
锰	-0.017	0.91
叶绿素	0.273	0.06
粪大肠菌群	0.083	0.57

注: (1) $P<0.05$ 在统计学上意味着存在显著的相关性, $P<0.01$ 则为极显著; (2)*表示与水质等级显著相关的特征。

通过相关性分析筛选出 7 项显著相关特征。特征与水质等级之间存在相关性,各特征之间也同样存在高度相关性,比如“溶解氧”与“五日生化

需氧量”以及“氨氮”与“总氮”。为避免选择的特征之间相关性过高,通过查阅相关文献,并结合《浙江省水资源公报》分析各水库近年来的水质状况,最终选取溶解氧(DO)、高锰酸盐指数(COD_{Mn})、总磷(TP)和氨氮含量(NH₃-N) 4 项监测数据作为本文水质评价指标,并且以《地表水环境质量标准》(GB 3838—2002)作为评价依据(表 2)。

表 2 地表水环境质量标准 $\text{mg}\cdot\text{L}^{-1}$

水质等级	评价指标			
	DO	COD _{Mn}	TP	NH ₃ -N
I	7.5	2	0.01	0.15
II	6	4	0.025	0.5
III	5	6	0.05	1.0
IV	3	10	0.1	1.5
V	2	15	0.2	2.0

2 评价方法

首先构建水质的犹豫模糊标准评价矩阵,在《地表水环境质量标准》(GB 3838—2002)基础上,采用专家打分值建立犹豫模糊矩阵,再通过犹豫模糊指数熵来计算各指标间权重,采用升维补齐方式统一犹豫模糊数的维度,最后使用加权犹豫模糊兰氏距离测度方法计算参评水域犹豫模糊矩阵与标准评价矩阵各等级的距离,最终根据距离最小原则划定水质等级,从而得出水质评价结论。

2.1 水质评价标准矩阵的建立

犹豫模糊集定义: 设论域 $X = \{x_1, x_2, \dots, x_n\}$, 则 X 上的犹豫模糊集(Hesitant Fuzzy Set, HFS)定义为 $M = \{\langle x, h_M(x) \rangle | x \in X\}$, 其中 $h_M(x)$ 为犹豫模糊数, 是 $[0, 1]$ 中一些数值的集合, 表示集合 X 中的任一元素 x 对集合 M 的隶属度。特别地, 如果各犹豫模糊数有且仅有一个隶属度值, 那么犹豫模糊集退化为普通的模糊集。

水质评价指标的属性集合 $X = \{\text{溶解氧, 高锰酸盐指数, 总磷, 氨氮}\}$, 并设置 5 个不同等级, 即等级属性集合为 $G = \{\text{优, 良, 一般, 较差, 差}\}$ 。邀请 3 名水环境评价方面专家组成水质评价小组, 参照表 2 规定的各水质指标值, 对溶解氧、高锰酸盐指数等评价指标进行打分, 分值在 0~1, 且分值越高代表水质等级越高。最后对评价小组的分值进行合并, 形成水质评价标准矩阵。由于不同专家可

能有不同的专业知识、背景和偏好,导致分值可能存在差异;但每位专家的意见都需被考虑,因此水质评价标准矩阵中的每个指标值用犹豫模糊数表示为专家打分值组成的集合(表3).

表3 水质评价标准矩阵

水质等级	DO	COD _{Mn}
优(I)	{0.7, 0.8, 0.9}	{0.8, 0.9}
良(II)	{0.6, 0.7, 0.8}	{0.7, 0.8}
一般(III)	{0.4, 0.5}	{0.5, 0.6}
较差(IV)	{0.2, 0.3}	{0.3, 0.4, 0.5}
差(V)	{0.1}	{0.1, 0.2}
水质等级	TP	NH ₃ -N
优(I)	{0.8, 0.9}	{0.8, 0.9}
良(II)	{0.6, 0.7, 0.8}	{0.7, 0.8}
一般(III)	{0.4, 0.6}	{0.5, 0.6, 0.7}
较差(IV)	{0.2, 0.3}	{0.3, 0.4}
差(V)	{0.1}	{0.2}

注:表中数据是专家打分值组成的集合,该集合可能是单个值,也可能是多个不同数值的集合.

2.2 指标权重计算

在指标权重信息未知的条件下,通常需要人为主观给出.本文通过实际数据信息构造犹豫模糊指数熵的方式来确定指标权重,避免了过多人为主观因素的影响.

定义一组犹豫模糊数:

$$h_M(x_i) = \{h_M^{\sigma(1)}(x_i), h_M^{\sigma(2)}(x_i), \dots, h_M^{\sigma(l_{x_i})}(x_i)\},$$

式中: l_{x_i} 为 $h_M(x_i)$ 中隶属度的个数,则犹豫模糊数 $h_M(x_i)$ 的指数熵定义为^[16]:

$$E(h_M(x_i)) = \frac{1}{l_{x_i}(\sqrt{e}-1)} \sum_{j=1}^{l_{x_i}} \left(h_M^{\sigma(j)}(x_i) e^{1-h_M^{\sigma(j)}(x_i)} + (1-h_M^{\sigma(j)}(x_i)) e^{h_M^{\sigma(j)}(x_i)} - 1 \right). \quad (1)$$

根据定义计算犹豫模糊指数熵,运用信息熵最小化原则计算确定属性权重:

$$w_i = \frac{1-E_i}{n-\sum_{i=1}^n E_i}, \quad (2)$$

式中: $E_i = \frac{1}{m} \sum_{k=1}^m E(h_{ki}), i=1, 2, \dots, n$.

2.3 距离测度计算

对于任意集合 $X = \{x_1, x_2, \dots, x_n\}$ 上的犹豫模糊

集 M 和 N , 距离测度 $d(M, N)$ 需满足以下性质:

性质一: $0 \leq d(M, N) \leq 1$;

性质二: $d(M, N) = 0$, 当且仅当 $M = N$;

性质三: $d(M, N) = d(N, M)$.

常用距离测度的方式有多种,如豪斯道夫距离测度、海明距离测度、兰氏距离测度^[17].本文水质评价中,不同专家对参评水库打分可能会出现一定差异性,如当遇到评价指标存在偏倚较大的数据,就会导致最终的评价结果偏离真实情况.因此本文将采用兰氏距离进行测度分析,因其是无量纲,并且是采用比值方式来衡量数据之间的距离,减少了极端值的影响,能有效避免上述问题对评价结果的影响.

在进行距离测度计算时,常会遇到犹豫模糊评价矩阵与水质评价标准矩阵在各指标下隶属度不等长的情况,导致无法直接进行距离度量.此时通常都会采用升维或降维的方式将犹豫模糊数长度统一,但降维方式会丢失过多的专家打分信息,因此这里选择采用升维补齐方案,将长度略短的犹豫模糊数补齐至更长的犹豫模糊数.补齐数据的方式也有多种,如对情境乐观的乐观补齐法,可选用犹豫数中最大的隶属度进行补齐;偏向于规避风险的悲观补齐法,可挑选犹豫数中最小的隶属度进行补齐^[18].本文认为虽然补齐方式可以很大程度上保留专家们的可靠评价,但是无论是乐观补齐法还是悲观补齐法都未能较好地权衡所有专家之间的意见,鉴于水质评价是一个需要严格把关的问题,因此悲观倾向是能够被采纳的.

所以本文提出基于线性插值的悲观补齐法,它能很好地全面考虑各个专家之间的态度.具体操作如下:设定一组犹豫模糊数 $h_M(x_i) = \{h_M^{\sigma(1)}(x_i), h_M^{\sigma(2)}(x_i), \dots, h_M^{\sigma(l_{x_i})}(x_i)\}$ 和升维大小 d , 其中 $l_{x_i} \leq d$. 为保证补齐方法偏向于悲观,从 $h_M(x_i)$ 中选取最小的两个隶属度 $h_M^{\sigma(m)}(x_i)$ 、 $h_M^{\sigma(n)}(x_i)$, 且 $h_M^{\sigma(m)}(x_i) < h_M^{\sigma(n)}(x_i)$, 并且在两者之间按照一维线性插值法插入共 $d-l_{x_i}$ 个隶属度,从而达到补齐的效果.

在水质评价的实际运用中,经常会出现不同参评水库有着相同犹豫模糊数的情况.例如一个6名专家组成的评价小组对甲、乙水库的某项水质指标等级进行打分,假设其中专家小组对甲水库的评分情况为4名专家评分是0.8,另外2名专家评

分是 0.6, 用犹豫模糊数 $\{0.6, 0.8\}$ 表示评分值; 专家小组对乙水库评分情况为 3 名专家评分是 0.8, 另外 3 名专家评分是 0.6, 用犹豫模糊数 $\{0.6, 0.8\}$ 表示评分值. 虽然甲、乙水库的评价结果完全相同, 但显然不符合实际的评分情况. 因此, 需引入加权犹豫模糊数概念, 对犹豫模糊数中每个隶属度值都进行加权处理, 且权重大小能反映出每个隶属度的重要程度. 但加权的犹豫模糊数也分为专家权重已知和未知两种情形. 在专家权重已知情况下, 构造加权犹豫模糊数, 设专家 $e_1, e_2, \dots, e_p$ 对应的权重向量 $w_q = (w_1, w_2, \dots, w_p)^T$, 其中, $w_k \geq 0$, 且 $\sum_{k=1}^p w_k = 1$.

构造加权犹豫模糊数为:

$$h_M^{w_q}(x_i) = \left\{ \left(h_M^{\sigma(j)}(x_i), w_j \right) \middle| h_M^{\sigma(j)}(x_i) \in h_M(x_i), \right. \\ \left. w_j = \sum_{k \in N(h_M^{\sigma(j)}(x_i))} w_k, i=1, 2, \dots, n \right\},$$

式中: $N(h_M^{\sigma(j)}(x_i))$ 表示给出隶属值为 $h_M^{\sigma(j)}(x_i)$ 的专家数.

但在绝大多数情况下, 专家权重是未知的, 此时则令每位专家权重都相同, 结合加权犹豫模糊数概念, 可给出加权犹豫模糊兰氏距离公式.

假设任意 $x_i \in X$ 的权重为 w_i , 其中 $i=1, 2, \dots, n$, 并满足 $w_i \in [0, 1]$, $\sum_{i=1}^n w_i = 1$, 则加权犹豫模糊兰氏距离^[19]:

$$d_{\text{whl}}(M, N) = \sum_{i=1}^n w_i \left[\frac{1}{l_{x_i}} \sum_{j=1}^{l_{x_i}} \left| \frac{h_M^{\sigma(j)}(x_i) - h_N^{\sigma(j)}(x_i)}{h_M^{\sigma(j)}(x_i) + h_N^{\sigma(j)}(x_i)} \right| \right]. \quad (3)$$

由此运用加权犹豫模糊兰氏距离测度方法计算犹豫模糊矩阵与标准评价矩阵的距离, 并根据距离最小原则可以得出最终结论.

3 实际应用

使用上述方法对浙江省部分水库水体进行水质评价, 并对评价结果进行有效性检验, 最后再与其他评价方法进行对比检验.

3.1 算例

出于保护水资源的战略选择, 现对浙江省内 5 处(编号为 A、B、C、D、E)水体进行各自的水质

量评价. 为保证评价结果的可靠性和专业性, 邀请 8 名水环境评价方面的专家, 并根据 5 处水体的各自情况分别进行打分评价, 整理得到各水体的犹豫模糊评价矩阵, 其相关数据见表 4.

表 4 犹豫模糊评价矩阵

水体编号	DO	COD _{Mn}
A	{0.30, 0.45, 0.40}	{0.80, 0.85, 0.75}
B	{0.90, 0.80, 0.75}	{0.70, 0.75}
C	{0.80, 0.65, 0.70}	{0.85, 0.75, 0.80}
D	{0.90, 0.95, 0.85}	{0.85, 0.90}
E	{0.65, 0.55, 0.50}	{0.65, 0.70}
水体编号	TP	NH ₃ -N
A	{0.20, 0.25, 0.30}	{0.95, 0.90}
B	{0.60, 0.45, 0.50}	{0.75, 0.85, 0.80}
C	{0.30, 0.45, 0.50}	{0.60, 0.75, 0.80}
D	{0.95, 0.90}	{0.95}
E	{0.25, 0.30}	{0.80, 0.70, 0.75}

根据表 4 的犹豫模糊评价矩阵, 利用式(1)计算出各水体指标的犹豫模糊指数熵. 其中, 以水体 A 为例:

$$E(h_{M,DO}) = \frac{1}{3(\sqrt{e}-1)} (0.7e^{0.3} + 0.3e^{0.7} - 1 + \\ 0.55e^{0.45} + 0.45e^{0.55} - 1 + 0.6e^{0.4} + \\ 0.4e^{0.6} - 1) = 0.9329.$$

在得到水体 A 在 DO 指标上的犹豫模糊指数熵后, 同理可构建出各水体相关指数的犹豫模糊指数熵矩阵, 其相关数据见表 5.

表 5 犹豫模糊指数熵矩阵

水体编号	DO	COD _{Mn}	TP	NH ₃ -N
A	0.932 9	0.643 8	0.752 0	0.284 1
B	0.593 5	0.802 6	0.984 1	0.643 8
C	0.803 7	0.643 8	0.945 6	0.790 5
D	0.363 3	0.446 4	0.284 1	0.197 3
E	0.968 1	0.880 1	0.802 6	0.752 0

接着可进一步计算出每个水体在单个指标上的平均模糊熵. 再以表 5 中 DO 指标为例:

$$E_1 = (0.9329 + 0.5935 + 0.8037 + \\ 0.3633 + 0.9681) / 5 = 0.732,$$

其中, E_1 表示第 1 个指标 DO 的平均指数熵. 同理可得其余 3 项指标的平均指数熵分别为 $E_2 = 0.683$ (COD_{Mn}), $E_3 = 0.754$ (TP), $E_4 = 0.534$ (NH₃-N).

表 6 升维后的水质评价标准矩阵

水质等级	DO	COD _{Mn}	TP	NH ₃ -N
优(I)	{0.70, 0.73, 0.75, 0.78, 0.80, 0.90}	{0.80, 0.82, 0.84, 0.86, 0.88, 0.90}	{0.80, 0.82, 0.84, 0.86, 0.88, 0.90}	{0.80, 0.82, 0.84, 0.86, 0.88, 0.90}
良(II)	{0.60, 0.63, 0.65, 0.68, 0.70, 0.80}	{0.70, 0.72, 0.74, 0.76, 0.78, 0.80}	{0.60, 0.63, 0.65, 0.68, 0.70, 0.80}	{0.70, 0.72, 0.74, 0.76, 0.78, 0.80}
一般(III)	{0.40, 0.42, 0.44, 0.46, 0.48, 0.50}	{0.50, 0.52, 0.54, 0.56, 0.58, 0.60}	{0.40, 0.44, 0.48, 0.52, 0.56, 0.60}	{0.50, 0.53, 0.55, 0.58, 0.60, 0.70}
较差(IV)	{0.20, 0.22, 0.24, 0.26, 0.28, 0.30}	{0.30, 0.33, 0.35, 0.38, 0.40, 0.50}	{0.20, 0.22, 0.24, 0.26, 0.28, 0.30}	{0.30, 0.32, 0.34, 0.36, 0.38, 0.40}
差(V)	{0.10, 0.10, 0.10, 0.10, 0.10, 0.10}	{0.10, 0.12, 0.14, 0.16, 0.18, 0.20}	{0.10, 0.10, 0.10, 0.10, 0.10, 0.10}	{0.20, 0.20, 0.20, 0.20, 0.20, 0.20}

表 7 升维后的犹豫模糊评价矩阵

水体编号	DO	COD _{Mn}	TP	NH ₃ -N
A	{0.30, 0.33, 0.35, 0.38, 0.40, 0.45}	{0.75, 0.76, 0.78, 0.79, 0.80, 0.85}	{0.20, 0.21, 0.23, 0.24, 0.25, 0.30}	{0.90, 0.91, 0.92, 0.93, 0.94, 0.95}
B	{0.75, 0.76, 0.78, 0.79, 0.80, 0.90}	{0.70, 0.71, 0.72, 0.73, 0.74, 0.75}	{0.45, 0.46, 0.48, 0.49, 0.50, 0.60}	{0.75, 0.76, 0.78, 0.79, 0.80, 0.85}
C	{0.65, 0.66, 0.68, 0.69, 0.70, 0.80}	{0.75, 0.76, 0.78, 0.79, 0.80, 0.85}	{0.30, 0.34, 0.38, 0.41, 0.45, 0.50}	{0.60, 0.64, 0.68, 0.71, 0.75, 0.80}
D	{0.85, 0.86, 0.88, 0.89, 0.90, 0.95}	{0.85, 0.86, 0.88, 0.89, 0.90, 0.95}	{0.90, 0.91, 0.92, 0.93, 0.94, 0.95}	{0.95, 0.95, 0.95, 0.95, 0.95, 0.95}
E	{0.50, 0.51, 0.53, 0.54, 0.55, 0.65}	{0.65, 0.66, 0.67, 0.68, 0.69, 0.70}	{0.25, 0.26, 0.27, 0.28, 0.29, 0.30}	{0.70, 0.71, 0.73, 0.74, 0.75, 0.8}

通过每个指标模糊熵在总体上的占比, 可以得到各个指标的具体权重数值. 同样, 以 DO 指标为例, 并依照式(2)可得:

$$w_1 = \frac{1 - E_1}{4 - \sum_{i=1}^4 E_i} = \frac{1 - 0.732}{4 - 2.703} = 0.21.$$

同理可得, $w_2 = 0.24$ (COD_{Mn}), $w_3 = 0.19$ (TP), $w_4 = 0.36$ (NH₃-N).

由于 5 处水体的犹豫模糊评价矩阵与水质评价标准矩阵在各指标下的隶属度不等长, 导致无法直接进行距离测度计算, 因此需要分别对表 3 和表 4 进行升维处理. 升维后的矩阵数据分别见表 6 和表 7.

最后根据式(3)计算各水体与各评价等级之间的距离测度, 其结果见表 8. 由最小距离原则, 本文将测度最小的等级作为该水体的等级评价, 最终各水体的水质评价结果见表 9. 从表 9 可以看出

表 8 各水体与各评价等级之间的距离测度

评价等级	A	B	C	D	E
优(I)	0.268 0	0.106 0	0.149 6	0.045 4*	0.231 5
良(II)	0.239 6	0.073 5*	0.093 6*	0.124 3	0.159 4
一般(III)	0.223 0*	0.146 9	0.151 9	0.278 4	0.158 1*
较差(IV)	0.244 7	0.389 6	0.342 9	0.505 2	0.258 9
差(V)	0.566 2	0.675 0	0.643 0	0.745 4	0.587 8

注: *为各水体最小距离测度.

表 9 各水体的水质评价结果

水体编号	官方评价等级	本文评价等级
A	较差(IV)	一般(III)
B	良(II)	良(II)
C	良(II)	良(II)
D	优(I)	优(I)
E	一般(III)	一般(III)

表 10 各水体连续 3 d 的水质评价结果

水体编号	第 1 天		第 2 天		第 3 天	
	官方	本文	官方	本文	官方	本文
A	较差(IV)	一般(III)	较差(IV)	一般(III)	较差(IV)	一般(III)
B	良(II)	良(II)	良(II)	良(II)	良(II)	良(II)
C	良(II)	良(II)	良(II)	良(II)	一般(III)	一般(III)
D	优(I)	优(I)	优(I)	优(I)	优(I)	优(I)
E	一般(III)	一般(III)	一般(III)	一般(III)	良(II)	良(II)

本文评价结果与官方评价结果基本一致,仅 A 水体的评价等级与官方出现差异,但对于 A 水体来说,从表 8 可以看出,距离“良(II)”等级的距离为 0.239 6,距离“一般(III)”等级的距离为 0.223 0,两者相差仅 0.016 6,因此不论从后续改进还是总体准确性来说,本文方法都具备较好的可靠性。

3.2 稳定性分析

为验证本文评价方法的稳定性,现将上述 5 个水体另外相邻 3 d 的监测数据和对应的官方评价等级依次进行上述步骤处理,最终得到本文的评价等级,结果见表 10。

从表 10 可以明显看出,水体在相邻时间点的水质变化不明显,并且本文的评价等级与官方评价等级基本一致,还能在不同时候保持两者的高度一致性,这也同样说明了本文方法的稳定性。

3.3 有效性分析

采用水质评价领域常用的 BP 神经网络和模糊综合评价方法来验证本文方法的有效性,其最终对比结果见表 11。

表 11 不同方法水质评价结果

水体编号	本文评价	BP 神经网络	模糊综合评价	官方评价
A	一般(III)	较差(IV)	优(I)	较差(IV)
B	良(II)	良(II)	一般(III)	良(II)
C	良(II)	一般(III)	良(II)	良(II)
D	优(I)	优(I)	优(I)	优(I)
E	一般(III)	一般(III)	良(II)	一般(III)

根据表 11 的对比结果可以看出, BP 神经网络在水质评价中的结果与官方之间存在一定差异,并且神经网络的训练需要大量带标签数据,同时网络输出值也并非离散值,而是连续值。因此如果取舍不恰当,容易造成最终结果的偏差。而模糊

综合评价则与官方评价出入较大,因其仅考虑一位专家的评价,虽简单方便,但也容易导致大量极端值的出现,从而导致结果可信度的缺失。综上所述,本文方法还是具有一定的优越性。

4 结论

水体环境监测是重要的民生大事,是关乎推动建设美丽社会开展的极大事宜,科学全面的水质量评价是其中的重要环节。本文通过筛选有效的水质评价指标,建立水体质量评价体系,并采用线性插值的悲观补齐方案和兰氏距离测度方法,提出一种针对水体质量评价的方法,通过自身和对比检验验证了其稳定性和有效性,以期对日后的水质评价提供一定的参考。

参考文献:

- [1] 徐祖信. 我国河流单因子水质标识指数评价方法研究[J]. 同济大学学报(自然科学版), 2005, 33(3):321-325.
- [2] 徐祖信. 我国河流综合水质标识指数评价方法研究[J]. 同济大学学报(自然科学版), 2005, 33(4):482-488.
- [3] 朱灵峰, 王燕, 王阳阳, 等. 基于单因子指数法的海浪河水质评价[J]. 江苏农业科学, 2012, 40(3):326-327.
- [4] 李祚泳. BP 网络用于水质综合评价方法的研究[J]. 环境工程, 1995, 13(2):51-53.
- [5] 薛建军, 姚桂基. 人工神经网络在水质评价中的应用[J]. 水文, 1997, 17(3):38-40.
- [6] 赵颖, 王建英, 孙燕, 等. 改进人工神经网络在河南部分河流的水质评价中的应用[J]. 环境与发展, 2018, 30(3):216-217.
- [7] 朱永军, 吴琼, 湛忠宇. 基于主成分分析法与人工神经网络耦合模型的水质评价[J]. 江苏水利, 2021(8):48-54.
- [8] 梁广林, 贾振宇, 高艳妮, 等. 基于模糊隶属度模型的辽河干流水质评价[J]. 环境工程技术学报, 2021, 11(4):693-700.

- [9] 邢驰, 聂光成, 唐上钦, 等. 基于犹豫模糊集的毁伤效能评估方法[J]. 兵器装备工程学报, 2021, 42(10):120-129.
- [10] 孙永泉, 刘旭, 陈吉, 等. 基于模糊 BP 神经网络的苏南 Y 湖水质评价[J]. 安徽农业科学, 2014, 42(34): 12253-12255.
- [11] 邹志红, 孙靖南, 任广平. 模糊评价因子的熵权法赋权及其在水质评价中的应用[J]. 环境科学学报, 2005, 25(4):552-556.
- [12] 裴廷权, 王里奥, 韩勇, 等. 三峡库区小江流域水体富营养化的模糊评价[J]. 农业环境科学学报, 2008, 27(4): 1427-1431.
- [13] 张丛. 环境评价教程[M]. 北京: 中国环境科学出版社, 2003.
- [14] TORRA V, NARUKAWA Y. On hesitant fuzzy sets and decision[C]//2009 IEEE International Conference on Fuzzy Systems, Jeju Island, Korea. IEEE, 2009:1378-1382.
- [15] TORRA V. Hesitant fuzzy sets[J]. International Journal of Intelligent Systems, 2010, 25(6):529-539.
- [16] 谭吉玉, 刘高常. 犹豫模糊指数熵及其应用[J]. 统计与决策, 2017(24):66-69.
- [17] 沙秀艳, 尹传存. 基于隶属度偏差加权的改进犹豫模糊距离测度及应用[J]. 统计与决策, 2022, 38(4):180-184.
- [18] XU Z S, XIA M M. Distance and similarity measures for hesitant fuzzy sets[J]. Information Sciences, 2011, 181(11):2128-2138.
- [19] 沙秀艳, 尹传存, 徐泽水. 基于指数熵加权的降维犹豫模糊兰氏距离测度及应用[J]. 控制与决策, 2020, 35(3): 728-734.

Hesitant fuzzy sets based water quality assessment

CAO Lei^{1,2}, DAI Junjie², MAO Xiaoyan², JIANG Ji³, LIU Fengqin³

(1.Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China;

2.College of Information Engineering, College of Science and Technology of Ningbo University, Ningbo 315300, China;

3.Cixi Municipal Bureau of Water Resources, Ningbo 315399, China)

Abstract: In order to assess the overall state of water bodies, based on the results of single factor evaluation of water quality grades, the correlation coefficient between the indicators and grades is calculated, and strong correlation indicators are selected. An initial matrix of hesitation fuzzy water quality assessment is constructed using expert scoring, and a weighted hesitation fuzzy Lancet distance measure multiple attribute evaluation method is established based on hesitation fuzzy exponential entropy and linear interpolation pessimistic dimensionality. Finally, an example analysis is conducted based on the water quality data of some reservoirs in Zhejiang Province. The results show that the proposed method for overall assessment of water quality is more reasonable and efficient.

Key words: water quality evaluation; hesitant fuzzy set; distance measure; dimension increase

(责任编辑 章践立)