

■ 松辽流域水资源保护系列丛书(六)

松花江流域典型河湖水质 评价与预测研究

郑国臣 秦雨 杨帆 等/著

Assessment and Prediction Research of Water Quality for
Typical Rivers and Lakes in Songhua River Basin



科学出版社

松辽流域水资源保护系列丛书（六）

松花江流域典型河湖 水质评价与预测研究

郑国臣 秦 雨 杨 帆 等 著

科学出版社

北 京

内 容 简 介

本书在对比和分析现行松花江流域典型河湖水质评价的典型方法(单因子评价法、污染指数法、主成分分析法、因子分析法、层次分析法、聚类分析法、物元分析法、模糊综合评价法、灰色分析法、云模型法、BP神经网络评价法)和水质预测方法(多元回归法、时间序列法、马尔可夫法、BP神经网络预测法、贝叶斯网络法)的基础上,将水质综合评价平台和智能预测平台融合,进行典型河流与水库水质的综合评价和智能预测。

本书所构建的可视化平台可实现水库水质综合评价的自动化、预测的智能化,且易于操作,对于进行合理的水质综合评价、预测具有重要的理论意义,也可以为松花江流域水环境的管理、水资源的开发利用提供重要的科学依据。

本书可作为水环境管理等相关领域的科研人员及管理人士的参考用书,也可作为大专院校环境、水利、生态专业教师和研究生的参考用书。

图书在版编目(CIP)数据

松花江流域典型河湖水质评价与预测研究/郑国臣等著. —北京:科学出版社, 2018.8

(松辽流域水资源保护系列丛书; 六)

ISBN 978-7-03-056646-1

I. ①松… II. ①郑… III. ①松花江-流域-河流水质-研究 IV. ①P343.1

中国版本图书馆 CIP 数据核字 (2018) 第 039898 号

责任编辑: 张 震 孟莹莹 / 责任校对: 赵桂芬

责任印制: 吴兆东 / 封面设计: 无极书装

科学出版社出版

北京东黄城根北街 16 号

邮政编码: 100717

<http://www.sciencep.com>

北京厚诚则铭印刷科技有限公司印刷

科学出版社发行 各地新华书店经销

*

2018 年 8 月第 一 版 开本: 720×1000 1/16

2018 年 8 月第一次印刷 印张: 16

字数: 320 000

定价: 116.00 元

(如有印装质量问题, 我社负责调换)

作者委员会

主 任:

郑国臣 (松辽流域水资源保护局)

秦 雨 (长春工程学院水利与环境工程学院)

副主任:

杨 帆 (松辽流域水资源保护局)

王 硕 (江南大学环境与土木工程学院)

张多英 (黑龙江大学建筑工程学院)

参加写作人员:

马品非 (塔城地区水利水电勘察设计院)

赵毓斌 (深圳先进技术研究院)

曹广丽 (城市水资源与水环境国家重点实验室
哈尔滨工业大学)

张 正 (松辽流域水资源保护局)

何佳吉 (松辽流域水资源保护局)

刘晓旭 (松辽流域水资源保护局)

田浩然 (松辽流域水资源保护局)

前 言

伴随着人们对流域水质的更高要求和社会经济的不断发展,找到更加科学合理的水质评价与预测方法是当务之急。松花江流域位于中国东北地区的北部,东西长 920km,南北宽 1070km,流域面积 55.68 万 km²,是中国重要工业(机械、石油、化工、制药等)基地和粮食主产区。松花江流域作为吉林省、黑龙江省、内蒙古自治区的工业、农业用水和饮用水的主要来源,水质状况与社会的发展密切相关。与中国其他流域相比,松花江流域地处寒冷的东北部,具有自身的流域特点,如化工产品污染突出,农业面源污染严重,且流域冰封期极长。随着东北老工业基地振兴战略和国家粮食安全战略的实施,针对松花江流域的特点,松花江流域水质评价与预测的实践工作仍然面临很多难题。因此,系统开展松花江流域典型河流、水库水质评价与预测相关工作,对进一步加强流域水资源保护与管理具有重要的意义。

流域尺度的水环境管理应从目前的单要素分散管理转向以生态系统为对象的综合管理。中国颁布的《地表水环境质量标准》(GB 3838—2002)将地表水水域环境功能和保护目标按功能高低依次划分为 5 类,并规定了 24 项水质基本项目对应于不同水域功能类别的浓度限值。对水质单项指标,将其实测值与不同功能类别对应的水质浓度限值相比较,可以判断出单项指标的水质类别。除了单项水质指标的评价,还要开展综合水质评价,即通过一个确定的数值,对一组水质指标所反映的水质予以总体评价。在流域水资源保护中,尤其需要通过对一组水质指标反映的整体水质状况进行评价,科学合理地分析水环境总体水质变化情况及趋势。水质变化的定量考核能够更客观、更具有操作性地对水资源保护效果予以评价与预测。综合水质评价结论既不能掩盖水质污染的现实,尤其是主要污染指标,也不能过度保护,以最差水质指标的水质代表综合水质。近年来,综合水质评价与预测在流域水资源保护中得到了越来越多的应用。但是,综合水质评价与预测方法仍有一些关键问题需要研究和探讨。

水环境质量的评价与预测方法有很多种,各方法之间研究目的与水质评价的侧重点均有所不同。在松花江典型河流与水库的水环境质量评价的研究中,所采用的水质评价方法通常比较单一,评价结果与实际水质状况存在偏差,致使水环境管理缺乏科学的理论依据,导致水资源不能够得到合理的开发和充分的利用。

为解决这一问题,本书在对比和分析现行松花江流域典型河湖水质评价的典型方法(单因子评价法、污染指数法、主成分分析法、因子分析法、层次分析法、聚类分析法、物元分析法、模糊综合评价法、灰色分析法、云模型法、BP神经网络评价法)和水质预测方法(多元回归法、时间序列法、马尔可夫法、BP神经网络预测法、贝叶斯网络法)的基础上,根据各水质评价、预测方法的适用条件、优缺点等,将水质综合评价平台和智能预测平台进行融合,采用 MATLAB 技术实现平台的可视化,利用该平台进行典型河流与水库水质的综合评价和智能预测。本书所构建的可视化平台可实现水库水质综合评价的自动化、预测的智能化,且易于操作,对于进行合理的水质综合评价、预测具有重要的理论意义;可以为松花江流域水环境的管理、水资源的开发利用提供重要的科学依据;对于科学合理的流域水质综合评价、预测具有重要的推广价值;对于松花江流域水质控制决策的制定、水资源的合理开发利用具有重要的现实意义。

针对上述问题,本书分四部分共 15 章介绍松花江流域水环境管理中的综合水质评价技术及应用。

第一部分:第 1 章,绪论。阐述了松花江流域水环境管理中的综合水质评价和预测问题研究进展。

第二部分:第 2 章~第 9 章。围绕松花江流域水质评价问题,介绍多种典型的综合水质评价方法,包括单因子评价法、污染指数评价法、主成分分析法、因子分析法、层次分析法、聚类分析法、物元分析法、模糊综合评价法、灰色分析法、云模型法等。其中涵盖各种评价方法的理论、计算、应用等。

第三部分:第 10 章~第 14 章。围绕着松花江流域水质预测问题,详细介绍了水质预测方法(多元回归法、时间序列法、马尔可夫法、BP 神经网络法、贝叶斯网络技术)的原理、计算方法与流程、应用实例等。对已有的典型水质评价方法和新提出的融合水质评价方法平台进行比较研究,并进一步比选验证。

第四部分:第 15 章。本章围绕着流域水环境管理的实际需求,回答如何在水环境质量评价工作中,融合水质评价与预测方法平台。除了选用科学合理的水质评价方法外,本章还明确融合的水质预测功能、水质定性评价、水质随时间或空间变化评价等关键技术问题。

本书由郑国臣、秦雨统稿,杨帆、王硕、张多英主笔。其中第 1、4 章由郑国臣、秦雨撰写,第 2、3 章由杨帆、赵毓斌、刘晓旭撰写,第 5、7 章由王硕、张多英撰写,第 6、8 章由马品非、杨帆撰写,第 9、10 章由曹广丽、张正撰写,第 11、12 章由何佳吉、田浩然撰写,第 13、14 章由王硕、张多英撰写,第 15 章由秦雨、郑国臣撰写。

在本书的撰写过程中,水利行业的领导和同行给予了大力的支持;同时,哈

尔滨工程大学刘雨、贺旭、黄昊、孙宇同、马小兵，长春工程学院王兆波，东北林业大学关博文、姜厚竹，江南大学钱凯、朱引、王玉莹，东北农业大学谷际岐、张梦琦等同学也为本书的撰写做了大量而繁琐的工作，在此一并表示感谢。为了系统梳理水环境评价与预测工作，加快完善松花江流域水环境评价体系，作者结合近年来开展的相关课题及工作积累撰写本书。经过三年的艰苦努力，在完成科研项目的时候，也顺利完成了此书的撰写工作，并得到了部分专家、学者和管理人员的宝贵建议，在此也对给予过我们无私帮助的所有人表示衷心的感谢。本书得到国家水体污染控制与治理科技重大专项“基于水环境风险防控的松花江水文过程调控技术及示范”项目（项目编号：2012ZX07201006）的资助。

由于作者水平有限，书中不当之处在所难免，望广大读者批评指正。

郑国臣

2018年5月

目 录

前言

第 1 章 绪论	1
1.1 松花江典型河湖水质状况分析	1
1.1.1 河流水资源质量状况	1
1.1.2 重要江河湖泊水功能区水质达标状况	1
1.1.3 省(区)界及其他重要水体水资源质量状况	2
1.1.4 松花江区重要大中型水库	2
1.2 河湖水质综合评价与预测情况分析	2
1.2.1 河湖水质评价研究情况	2
1.2.2 河湖水质预测研究情况	2
1.3 流域水环境管理中的水质评价法概况	3
1.3.1 水质指标的选择	3
1.3.2 水质指标的预处理	3
1.3.3 权重系数的确定	4
1.3.4 主成分分析法	6
1.3.5 聚类分析	6
1.3.6 综合值的处理	6
1.3.7 计算机智能模型法	7
1.4 流域水环境管理中的水质预测方法	8
1.4.1 多元回归预测法	9
1.4.2 时间序列法	9
1.4.3 马尔可夫法	9
1.4.4 人工神经网络预测法	9
1.4.5 贝叶斯网络法	10
1.4.6 其他预测方法	11
1.5 本章小结	11
第 2 章 基础评价方法	12
2.1 概述	12
2.1.1 单因子评价法	12
2.1.2 内梅罗污染指数法	13

2.2	单因子评价法应用	13
2.2.1	断面选取	13
2.2.2	水质评价指标	15
2.2.3	单因子评价法的常见应用	16
2.3	内梅罗污染指数法应用	20
2.3.1	断面选取	20
2.3.2	水质指标分析	21
2.3.3	内梅罗污染指数变化分析	25
2.4	省界缓冲区监管具体对策及建议	26
2.5	本章小结	27
第3章	主成分分析法与因子分析法	28
3.1	主成分分析法简介	28
3.2	SPSS 软件及其使用方法	28
3.3	主成分分析法在松花江流域省界缓冲区管理中的应用	30
3.3.1	监测数据标准化处理	31
3.3.2	检验是否符合主成分分析条件	32
3.3.3	特征根及方差贡献	32
3.3.4	主成分负荷及主成分得分	33
3.4	因子分析法简介	34
3.5	因子分析的使用方法	35
3.6	因子分析法在嫩江流域省界缓冲区水质分析的应用	36
3.6.1	2011 年嫩江流域省界缓冲区重要断面因子分析	36
3.6.2	2013 年嫩江流域省界缓冲区重要断面因子分析	40
3.6.3	2011~2015 年大河断面因子分析	43
3.6.4	大河断面污染情况综合分析	44
3.7	本章小结	45
第4章	层次分析法	46
4.1	概述	46
4.2	AHP 的指标体系构建原则	49
4.3	尼尔基水库指标体系构建	50
4.3.1	监测断面布置	50
4.3.2	监测项目和分析方法	51
4.3.3	评价指标体系建立	51
4.3.4	指标权重的确定	51

4.3.5 构建判断矩阵	51
4.4 AHP 的应用	53
4.4.1 评价方法	53
4.4.2 数据收集与分析	53
4.4.3 指标层评价值的确定	55
4.4.4 要素层评价	55
4.5 本章小结	56
第 5 章 聚类分析法	57
5.1 概述	57
5.1.1 K-均值聚类分析	57
5.1.2 模糊聚类分析	59
5.2 K-均值聚类法的应用	60
5.2.1 水质监测断面优化原则	60
5.2.2 K-均值聚类法优化水质监测断面	61
5.2.3 监测断面分类 K 值的确定	61
5.2.4 K-均值聚类优化断面基本步骤	62
5.2.5 K-均值聚类对断面优化结果与分析	63
5.3 模糊聚类法的应用	70
5.3.1 确定分类对象及标准化样本数据	70
5.3.2 建立模糊相似关系	72
5.3.3 建立模糊等价关系	72
5.3.4 对各监测断面进行聚类	73
5.3.5 模糊聚类评价结果分析	78
5.4 本章小结	79
第 6 章 物元分析法	80
6.1 概述	80
6.1.1 物元的确定	80
6.1.2 关联度的确定	81
6.1.3 计算权系数	81
6.1.4 水质类别的确定	82
6.2 物元分析法的应用	82
6.2.1 选取监测样本	82
6.2.2 确定物元矩阵	85
6.2.3 计算权系数	86

6.2.4 松花江流域省界缓冲区水质评价	87
6.3 本章小结	90
第7章 模糊综合评价法	91
7.1 概述	91
7.2 模糊综合评价法的计算方法与流程	91
7.2.1 确定模糊关系矩阵	91
7.2.2 确定权重向量	92
7.2.3 模糊综合评价	93
7.3 模糊综合评价法的应用	93
7.3.1 断面布置	93
7.3.2 选取评价指标	93
7.3.3 评价结果与分析	96
7.4 本章小结	100
第8章 灰色分析法	101
8.1 概述	101
8.2 灰色分析法计算方法与流程	102
8.2.1 确定聚类白化数	102
8.2.2 数据的标准化处理	102
8.2.3 确定白化函数及白化矩阵	102
8.2.4 求聚类权	103
8.2.5 求聚类系数	103
8.3 灰色分析法的应用	103
8.3.1 新立城水库概述	103
8.3.2 水质指标的选取	104
8.3.3 灰色分析法评价结果分析	105
8.4 本章小结	112
第9章 云模型方法	113
9.1 概述	113
9.1.1 数据的小失真预处理方法	113
9.1.2 云模型及确定度的计算	114
9.1.3 污染贡献率法及熵权法的选择	116
9.1.4 综合值的处理	117
9.2 云模型法综合评价法的应用	117
9.2.1 数据的筛选	117

9.2.2 新立城水库水质综合评价	117
9.2.3 评价结果分析	118
9.3 本章小结	119
第 10 章 多元回归法	120
10.1 概述	120
10.2 尼尔基水库情况及检测指标	121
10.2.1 尼尔基水库概况	121
10.2.2 检测项目	121
10.2.3 软件及使用方法	122
10.3 多元回归法建立尼尔基水库坝前叶绿素 a 线性回归模型	122
10.3.1 尼尔基水库叶绿素 a 线性回归模型-坝前汛期	122
10.3.2 尼尔基水库叶绿素 a 线性回归模型-坝前非汛期	125
10.3.3 回归模型预测分析-坝前汛期水质	126
10.3.4 回归模型预测分析-坝前非汛期	127
10.4 多元回归法建立尼尔基水库库末叶绿素 a 线性回归模型	128
10.4.1 尼尔基水库叶绿素 a 线性回归模型-库末汛期	128
10.4.2 尼尔基水库库末叶绿素 a 线性回归模型-库末非汛期	130
10.4.3 回归模型预测分析-库末汛期	132
10.4.4 回归模型预测分析-库末非汛期	133
10.5 本章小结	134
第 11 章 时间序列法	135
11.1 时间序列法简介	135
11.2 时间序列平滑法计算方法与流程	135
11.2.1 移动平均法	135
11.2.2 指数平滑法	136
11.3 时间序列法的应用	137
11.4 本章小结	143
第 12 章 马尔可夫法	144
12.1 马尔可夫法简介	144
12.2 马尔可夫法计算方法与流程	144
12.2.1 划分预测对象状态	144
12.2.2 计算初始概率 p_i	145
12.2.3 计算状态的一步转移概率 p_{ij}	145

12.2.4	预测	145
12.3	马尔可夫法的应用	146
12.3.1	磨盘山水库简介	146
12.3.2	叶绿素 a、氮、磷含量预测	146
12.4	本章小结	149
第 13 章	BP 神经网络法	151
13.1	概述	151
13.1.1	BP 神经网络的结构	151
13.1.2	标准 BP 神经网络算法	152
13.2	BP 神经网络法的应用	153
13.2.1	评价背景	153
13.2.2	评价指标与数据的选取	153
13.2.3	BP 神经网络模型的建立	156
13.2.4	评价结果分析	157
13.3	BP 神经网络模型预测尼尔基水库水质	158
13.3.1	预测背景	158
13.3.2	BP 神经网络的构建	158
13.3.3	数据的标准化处理	159
13.3.4	尼尔基水库水质预测	162
13.4	本章小结	163
第 14 章	贝叶斯网络评价法	165
14.1	概述	165
14.1.1	贝叶斯网络技术原理	165
14.1.2	贝叶斯网络技术发展阶段	165
14.1.3	贝叶斯网络技术的优势	166
14.2	贝叶斯网络评价法计算方法及功能	166
14.2.1	贝叶斯网络模型方法	166
14.2.2	基于贝叶斯理论的水质评价	167
14.2.3	基于贝叶斯统计推断的水环境模型参数识别	167
14.2.4	贝叶斯网络技术在流域水环境模型预测中的应用	167
14.2.5	贝叶斯网络的功能	168
14.3	贝叶斯网络算法的应用	168
14.3.1	设计思路	168
14.3.2	基于贝叶斯技术评价与预测尼尔基水库水质	169

14.3.3 贝叶斯网络模型模拟结果	171
14.4 水环境成分贝叶斯网络数据分析软件开发	180
14.4.1 软件功能	180
14.4.2 软件运行条件及过程	180
14.5 本章小结	184
第 15 章 典型河湖水质评价与预测平台	185
15.1 典型河湖水质评价与预测平台构建	185
15.2 软件模块及功能设计	186
15.2.1 典型河湖水质评价方法平台	186
15.2.2 典型河湖水质预测方法平台	187
15.3 典型河湖水质评价与预测平台应用	187
15.3.1 测试数据——以石头口门水库数据为例	187
15.3.2 水质评价模块操作步骤	188
15.3.3 水质预测模块操作步骤	192
15.4 方法改进向导及一键测评的介绍	196
15.4.1 评价方法改进向导及一键评价	196
15.4.2 预测方法改进向导及一键预测	198
15.4.3 实例应用	199
15.5 本章小结	200
参考文献	201
附录 水质综合评价与预测源代码	203

第 1 章 绪 论

本书通过松花江流域典型河湖水质分析,介绍国内外相关评价方法[主要包括单因子评价法、污染指数法、主成分分析法、因子分析法、层次分析法(analytic hierarchy process, AHP)、聚类分析法、物元分析法、模糊综合评价法、灰色分析法、云模型法、反向传播(back propagation, BP)神经网络法]、预测方法(主要包括多元回归法、时间序列法、马尔可夫法、BP 神经网络法、贝叶斯网络法)的研究进展,解析适用于松花江典型河流、水库的水质评价与预测方法,进而提出构建松花江流域典型河湖水质评价与预测综合平台的技术思想,为诊断松花江流域水质状况提供技术支持。

1.1 松花江典型河湖水质状况分析

1.1.1 河流水资源质量状况

2016 年,水利部对松花江区 16 239.5km 的河流水质状况进行了全年、汛期和非汛期评价,水质总体为中。全年 I~III 类、IV~V 类和劣 V 类水河长分别占 66.7%、25.7%和 7.6%。汛期 I~III 类、IV~V 类和劣 V 类水河长分别占 60.5%、34.1%和 5.4%。非汛期 I~III 类、IV~V 类和劣 V 类水河长分别占 69.5%、22.9%和 7.6%。全年主要超标项目为高锰酸盐指数、化学需氧量和氨氮。与 2015 年相比,全年 I~III 类水河长比例下降了 3.1%,IV~V 类水河长比例上升了 4.5%,劣 V 类水河长比例下降了 1.4%。

1.1.2 重要江河湖泊水功能区水质达标状况

2016 年,水利部对松花江区 319 个重要江河湖泊水功能区进行了评价。水功能区全因子评价达标 113 个,达标率为 35.4%,主要超标项目为高锰酸盐指数、氨氮和总磷。水功能区限制纳污红线主要控制项目评价达标 161 个,达标率为 50.5%,主要超标项目为高锰酸盐指数、氨氮和化学需氧量。与 2015 年相比,水功能区限制纳污红线主要控制项目评价达标率上升了 3.7%,水功能一级区达标率上升了 4.9%,二级区达标率上升了 3.0%。

1.1.3 省（区）界及其他重要水体水资源质量状况

2016 年，水利部对松花江区省（区）界及其他重要水体 20 条河流的 51 个断面进行了监测评价，水质为良。Ⅰ～Ⅲ类、Ⅳ～Ⅴ类和劣Ⅴ类水质监测断面比例分别为 92.1%、5.9%和 2.0%。劣Ⅴ类断面 1 个，为卡岔河的龙家亮子断面。全年主要超标项目为氨氮、化学需氧量和高锰酸盐指数。与 2015 年相比，全年Ⅰ～Ⅲ类水质监测断面比例上升 5.8%，Ⅳ～Ⅴ类比例下降 5.9%，劣Ⅴ类比例上升 0.1%。

1.1.4 松花江区重要大中型水库

2016 年，水利部对松花江区 18 座大中型水库（17 座大型，1 座中型）进行了水质和营养状态评价，全年水质为Ⅱ～Ⅲ类的水库 13 座，占 72.2%；Ⅳ～Ⅴ类 4 座，占 22.2%；劣Ⅴ类 1 座，占 5.6%。按营养状态评价，中营养水库 4 座，占 22.2%；富营养水库 14 座，占 77.8%。

1.2 河湖水质综合评价与预测情况分析

1.2.1 河湖水质评价研究情况

河湖水质评价是水环境保护与管理的重要内容。通过对水质状况的分析，才能有针对性地制定河湖水环境管理决策与规划；通过科学合理的水质评价，才能准确诊断河湖水质状况，为推进流域水资源保护奠定良好的基础（康晓风等，2014；Fan et al., 2016）。依据《地表水环境质量标准》（GB 3838—2002），地表水水域环境功能和保护目标按功能高低依次划分为 5 类，该标准还规定了 24 项基本项目对应水质类别的浓度限值。对水质单项指标，将实测值与不同功能类别对应的水质浓度限值相比较，可以判断出单项指标的水质类别（Ⅰ、Ⅱ、Ⅲ、Ⅳ、Ⅴ）以及水质浓度的时空变化，这是单项指标的水质评价。然而，影响水质评价的因素是复杂的，往往需要将反映评价水质的多项指标信息加以汇集，得到多指标综合评价方法。在流域水资源保护与管理中，尤其需要通过对一组水质指标反映的整体水质状况进行评估，分析水环境总体水质变化状况及趋势。近年来，综合水质评价在松花江流域水资源保护中得到了越来越多的应用，但要更好地服务于松花江流域水环境管理，仍有一些关键问题需要研究和探讨。

1.2.2 河湖水质预测研究情况

流域水质状况的预测，不仅包括综合水质类别的变化，还应包括同一水质类别和不同水质类别间综合水质变化的趋势分析，而综合水质预测更具有实际意义。

经过流域水环境治理后, 尽管水质会有所改善, 但由于评价方法的差异, 或者某一因子没有根本改善, 水质类别不一定好转。因此, 对流域水质变化的定量预测与分析应该更客观、更具有操作性。目前, 典型流域水质预测方法主要有多元回归法、时间序列法、马尔可夫法、BP 神经网络法、贝叶斯网络法等。本书分别对这些方法的原理、计算方法与流程等进行介绍和应用, 针对流域典型河湖水质预测方法的不足之处, 提出方法的改进方向, 对已有的典型方法和新提出的方法进行比较、遴选和验证, 进而针对流域水环境管理的实际应用, 构建流域典型河湖的水质预测平台。

1.3 流域水环境管理中的水质评价法概况

1.3.1 水质指标的选择

水质指标的选择是水质综合评价的首要工作。水质指标的选择包括定性与定量两方面, 其中全面性与代表性两方面难以兼顾, 是水质指标选择的主要问题。《地表水环境质量标准》(GB 3838—2002) 中项目共计 109 项, 其中地表水环境质量标准基本项目 24 项, 集中式生活饮用水地表水源地补充项目 5 项, 集中式生活饮用水地表水源地特定项目 80 项 (黄廷林等, 2009)。由于很难对所有的水质指标进行评价, 水质指标信息的冗余等因素决定了我们在进行水质综合评价时, 指标以及指标数量的选取至关重要。指标数量过少, 代表性较弱; 过多则会造成计算量加大、成本上升、大量信息冗余等问题。目前, 水质指标的选取一般依据主观定性的方法, 常见的有层次分析法、主成分分析法和聚类分析法等。

1.3.2 水质指标的预处理

由于各个水质指标的单位不同, 数据的数量级相差较大, 且溶解氧 (DO)、pH 等指标不像总氮 (TN) 等指标数值越大表示水质越差, 所以在进行评价工作时, 为了工作的方便以及结果的准确性, 需要对数据进行标准化。水质指标的标准化包含两个方面: 一是由于量纲的不同, 对水质指标进行无量纲化; 二是针对 DO、pH 等指标, 对数据进行同向化处理。水质指标的标准化方法多种多样, 本章只介绍三种常用的标准化方法。

1. 数据的极值化处理

此方法是将实测数据与水质标准的最大、最小值做商, 具体计算公式如下。

越小越好的指标: