

基于 BP 神经网络的辽河源头区水质评价研究

卞建民¹, 胡昱欣¹, 李育松¹, 马永祥², 边 静³

(1. 吉林大学 环境与资源学院, 长春 130021;

2. 宁夏回族自治区煤田地质局, 银川 750004; 3. 吉林省地质环境监测总站, 长春 130021)

摘 要: 基于辽河源头区水环境问题日益突出的现状, 该文开展了辽河源头区水环境质量的研究, 旨在对区内的水体质量进行分析评价。通过资料收集与汇总, 基于 BP 人工神经网络结构的思想和理论, 利用研究区内 13 个控制断面的水质监测数据, 建立了包括 pH、溶解氧、氨氮、化学需氧量、五日生化需氧量、高锰酸盐指数的水质综合评价模型, 并应用训练好的模型进行仿真运算及水质综合评价。结果显示, 在选取的 13 个断面中, 约 76.92% 的断面为 V 类—劣 V 类水质, 仅有 23.08% 的断面水质级别在 II—III 类之间, 研究区上游断面的水质状况较好, 中下游的水质较差。将该结果与《环境公报》公布的主要断面水质结果进行对比, 81.25% 的评价结果相同, 采用 BP 神经网络对研究区水质进行综合评价具有较强的适用性和可靠性。

关键词: 辽河源头区; BP 神经网络; 网络训练; 水质评价

中图分类号: X524; TP183

文献标识码: A

文章编号: 1005-3409(2014)01-0147-05

Water Quality Assessment in Source Area of Liao River Based on BP Neural Network

BIAN Jian-min¹, HU Yu-xin¹, LI Yu-song¹, MA Yong-xiang², BIAN Jing³

(1. College of Environment and Resources, Jilin University, Changchun 130021, China;

2. Ningxia Hui Autonomous Region Bureau of Coal Geological Exploration, Yinchuan 750004,

China; 3. Geo-Environmental Monitoring Central Station of Jilin Province, Changchun 130021, China)

Abstract: Based on the present situation that the water environmental issues in source area of Liao River have become increasingly prominent, the research for water environment quality was carried out to evaluate and analyze the regional water quality. Through data collection and summary, a comprehensive water quality evaluation model was established based on the thought and theory of BP artificial neural network with including pH, DO, ammonia nitrogen, COD, BOD₅, potassium permanganate index, and finished with water quality monitoring data of the 13 sections in the study area. After training well, it can just be applied in model simulating operation and water quality comprehensive evaluation. The results have showed that in the selected sections, approximately 76.92% of the sections are between class V and worse than class V, leaving only 23.08% of the sections whose water quality levels are between class II and class III in the selected 13 sections. The sections located in the upper reaches have a better water quality than that in the downstream. Compared the evaluation results with the results of main sections published in Environment Communique, 81.25% of the evaluation results are identical. It has strong applicability and reliability that BP neural network was used to comprehensively evaluate water quality in the study area.

Key words: source area of Liao River; BP neural network; network training; water quality assessment

近年来,随着经济的发展,水体环境问题日益突出,而河流水体质量是对河流环境的一个较为综合的体现,水质的优劣严重影响了以河流水体为供水水源

的工业、企业的发展,尤以生活在沿岸的百姓为重。因此,对于水质优劣情况的了解显得尤为重要,通常判断水质优劣的最有效的方法是对水体进行质量评

价。有关地表水水质进行评价的方法很多,以模糊数学法及层次分析法等单因子评价方法应用广泛^[1-5],且具有理论成熟、可操作性较强等特点,但对多因素多属性的综合评价问题,采用单因子评价法则表现出局限性及欠合理性,此时,采用多因子评价方法对水体质量进行评价,体现出其客观性和真实性。因此,提出一种能够评价水体质量的多因子综合评价体系十分重要,而选取合理的评价指标,建立一种可靠的模型,是解决复杂评价问题的基础,亦是衡量水质评价准确与否的关键所在。

辽河源头区内的东辽河、条子河及招苏台河是沿岸工业、企业等的重要水源,随着区域经济的发展,水污染问题逐渐成为研究区关注的重点。然而,水污染日趋严重,不仅降低了水体的使用功能,也进一步加剧了水资源短缺的矛盾,严重威胁城市及乡村居民的饮水安全和人民群众的健康^[1-3]。因此,对辽河源头区水质的评价研究非常重要。对于河流水质的评价,一般采用的方法如内梅罗指数法、综合指数法、试用概率法、主成分分析法、模糊模式识别法等^[6-9]。而人工神经网络是目前广泛应用于多因子水质评价的一种较为成熟的模型,以BP(Error Back Propagation,误差反向传播)神经网络最为常用^[10]。BP神经网络除应用于水质综合评价外,还可对需水量、地表水分等水文过程进行预测^[11-12],本次主要对其水质评价方面的应用展开研究。针对研究区水环境问题突出现状,采用BP神经网络方法,通过对水质数据的处理、构建模型,实现对研究区内河流水质的综合评价。依据评价结果,对水质优劣的影响因素展开研究,为有关部门进行河流水体污染治理提供可靠的依据。

1 研究区概况

辽河源头区位于吉林省的西南部,松辽平原中部,地处东经 123° 42′—125° 31′,北纬 42° 34′—44° 08′,流域面积 11 283 km²,占全省总土地面积的 6.55%。研究区地势由东南向西北缓慢下降,海拔高程 611~120 m,区内三条主要河流均发源于吉林省境内,并在辽宁省境内与其他河流汇合。河流流量年内分布不均,季节变化明显。6—9 月为丰水期,径流量占年径流量的 60%~70%,3—5 月、10 月份为平水期,11 月—次年 2 月为枯水期,进入枯水期后,河水流量明显减少。

近年来,辽河水体污染问题突出,据资料统计,2010 年辽河源头区废水排放总量为 9 941.88 万 t。其中,工业废水排放量为 8 458.22 万 t,COD 排放量为 7.05 万 t,氨氮排放量为 0.65 万 t。如此大量的工

农业及生活污水的入河,使环境容量本已饱和的水体水质严重恶化,甚至出现黑臭等现象^[1-3]。而污水中 COD 和氨氮的大量排入,不仅造成了地表水的污染,并在一定程度上污染了地下水。

2 数据来源及方法

2.1 数据来源

本研究采用辽河源头区 2010 年的水质监测数据,以杨木水库、拦河闸、气象站、河清、城子上、周家河口、四双大桥、汇合口、林家、四台子、新立屯等 13 个主要监测断面水质月检测数据中的 pH、溶解氧(DO)、氨氮(NH₃-N)、化学需氧量(COD)、五日生化需氧量(BOD₅)、高锰酸盐指数(COD_{Mn})为指标,对研究区内各主要断面的水体质量展开评价研究,各监测断面分布见图 1。

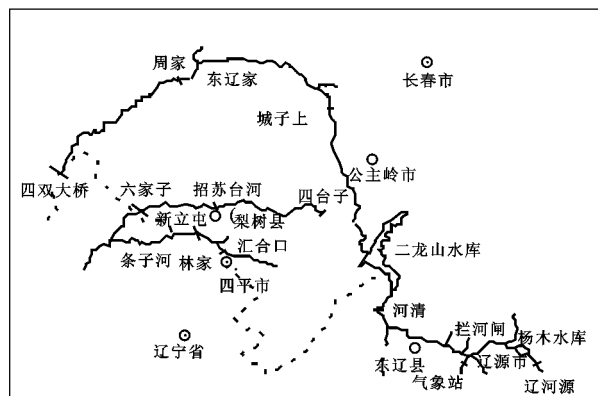


图1 研究区各监测断面位置

2.2 BP神经网络模型

BP神经网络包括输入层、隐含层和输出层^[13],采用BP神经网络对研究区水质进行评价,其问题实质就是通过设计一个BP神经网络,利用所给的有限样本的训练来表达逼近样本所蕴含的内在规律,最大限度地减少训练误差。模型的运行由信息正向传播和误差反向传播两个过程组成,这种正反传播来调整各层权值的过程是反复进行的^[14]。本次通过选取典型的3层BP神经网络结构,将各项待评价指标实测的水质检测值作为输入层单元,输出层为水质级别值,从而对研究区水质进行综合评价。

2.2.1 数据预处理 为建立和训练水质评价BP神经网络分析水质,首先搜集研究区各断面的水质监测数据,且必须保证样本量足够,降低网络误差,达到训练和测试的目的。将各监测断面2010年各月份检测数据中所选水质指标在丰水期、枯水期、平水期的平均值及年平均值作为水质评价的初始样本数据。由于各监测指标具有不同的量纲,难以进行直接比较。在评价前需把这些指标归一化到某一无量纲区间,包

括成本型和效益型两种标准化处理方法,对实测的水质数据也需要进行标准化处理。

以《地表水环境质量标准》(GB3838-2002)中所研究指标Ⅰ类—劣Ⅴ类的类别值作为评价网络的标准训练样本,标准化处理后采取等间隔均匀插值的方法来增加样本量,便于 BP 神经网络运算。并将水质

标准的前 5 个标准类别在[0,1]数值区间上划分,得到相对应 BP 神经网络模型输出的级别值,见表 1,以上处理可使水环境评价问题数值化。本次将研究指标的分级标准值进行标准化处理后作为输入,而输出的级别值与表 1 中的评价等级依次对应,这种内在规律就是我们寻找的水环境质量评价的函数^[15]。

表 1 与标准化后的地表水分级标准对应的水质级别

水质级别	Ⅰ类	Ⅱ类	Ⅲ类	Ⅳ类	Ⅴ类	劣Ⅴ类
区间	0.0~0.2	0.2~0.4	0.4~0.6	0.6~0.8	0.8~1.0	[0,1]以外

2.2.2 网络结构与训练函数的确定 构建 BP 神经网络时,网络层数、隐含层节点数同时影响着训练精度,网络层数的增加可进一步降低误差,但同时也提高了网络复杂性,增加了网络权值的训练时间^[16]。考虑通过增加隐含层中神经元的数目亦可提高训练精度,其训练效果较增加层数更易观察和调整。因此,一般选择增加隐含层中神经元的方法进行研究。在设计 BP 神经网络时,通常优先考虑 3 层典型结构,而对隐含层中神经元数目的选择,往往通过对选取不同神经元数目的网络训练效果进行对比后确定,在满足精度的要求下取尽量紧凑、高效的训练结果,以获得可靠的训练网络,本次选取典型的 3 层 BP 神经网络结构。

在确定网络算法和训练函数时,为了改进标准 BP 算法的不足,提高训练精度,本文选择应用较为广泛的 Levenberg-Marguardt 数值优化算法^[17],简称 L-M 优化算法。L-M 优化算法可以根据网络训练误差变化自动调节网络参数,它是梯度下降法与高斯牛顿法的结合^[18]。通过研究和实验积累得出,目前 LM 算法的训练函数 trainlm 具有收敛速度快和精度高等优点,由于 trainlm 函数的训练网络对目标输出的逼近能力很好,网络的误差亦非常小,在函数应用上更合理,本次研究即采用 trainlm 作为训练函数。

2.2.3 仿真运算及评价 对已经建立的 BP 神经网络进行训练,应用误差反传原理不断调整网络权值使得网络模型输出值与已知的训练样本输出值之间的误差平方和达到最小或小于某一期望值^[19],此时,将标准化后的实测样本数据带入训练好的网络进行仿真运算,获得相应的评价结果。

3 结果与分析

BP 神经网络的输入单元个数与选取的指标个数相同,即本次 BP 神经网络的输入单元为 6;输出单元即为每个断面对应的水质综合评价结果个数,因而网络的输出单元为 1,研究过程中建立的 BP 神经网络水质评价模型如图 2 所示。采用 trainlm 函数进行反复训练,在经过 15 000 次的训练后,经对比后确定隐含层神经元个数为 5 的 BP 网络对函数的逼近能力较强,误差较小^[20]。综合考虑网络性能和训练速度,将网络隐含层的神经元数目定为 5 个,即水质评价的 3 层 BP 神经网络结构确定为 6→5→1。

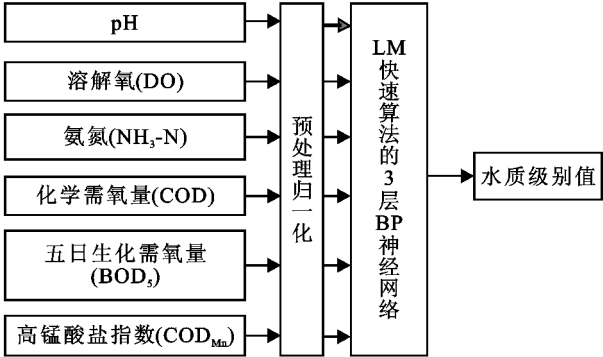


图 2 研究区水质评价的 BP 神经网络模型示意图

按照前述方法,将各监测断面实测的水质样本数据进行标准化处理后,输入已训练好的 BP 神经网络模型,进行评价网络仿真运算,获得网络输出各断面各水期的类别值,如表 2 所示。

将表 2 中评价网络输出的 13 个断面各水期的评价结果类别值,按照表 1 对应的区间值进行水质级别归类,获得相应的水质综合评价结果,如表 3 所示。

表 2 研究区 13 个断面经评价网络输出的水质类别值

各水期	辽河源	杨木水库	拦河闸	气象站	河清	城子上	周家河口	四双大桥	汇合口	林家	四台子	新立屯	六家子
丰水期	0.267	0.298	0.447	1.16	1.37	0.594	0.791	0.687	1.236	1.098	1.008	1.231	0.787
枯水期	0.497	0.523	0.588	1.98	1.99	0.792	1.111	1.365	1.768	1.899	1.989	1.999	1.768
平水期	0.327	0.478	0.543	1.34	1.47	0.818	0.878	0.978	1.567	1.634	1.592	1.654	1.432
年平均	0.335	0.508	0.513	1.46	1.58	0.833	0.915	0.996	1.548	1.612	1.745	1.456	1.365

表 3 2010 年辽河源头区水质综合评价结果

各水期 水质评价 级别	各监测断面												
	辽河源	杨木 水库	拦河闸	气象站	河清	城子上	周家 河口	四双 大桥	汇合口	林家	四台子	新立屯	六家子
丰水期	Ⅱ类	Ⅱ类	Ⅲ类	劣Ⅴ类	劣Ⅴ类	Ⅲ类	Ⅳ类	Ⅳ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	Ⅴ类	Ⅳ类
枯水期	Ⅲ类	Ⅲ类	Ⅲ类	劣Ⅴ类	劣Ⅴ类	Ⅳ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类
平水期	Ⅱ类	Ⅲ类	Ⅲ类	劣Ⅴ类	劣Ⅴ类	Ⅴ类	Ⅴ类	Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类
年平均	Ⅱ类	Ⅲ类	Ⅲ类	劣Ⅴ类	劣Ⅴ类	Ⅴ类	Ⅴ类	Ⅴ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类	Ⅴ类	劣Ⅴ类

从表 3 可知,辽河源、杨木水库、城子上、周家河口、四双大桥、新立屯、六家子断面在丰水期水质依次为Ⅱ类、Ⅱ类、Ⅲ类、Ⅳ类、Ⅳ类、Ⅴ类、Ⅳ类;在平水期依次为Ⅱ类、Ⅲ类、Ⅴ类、Ⅴ类、Ⅴ类、劣Ⅴ类、劣Ⅴ类;在枯水期依次为Ⅲ类、Ⅲ类、Ⅴ类、Ⅴ类、Ⅴ类、劣Ⅴ类、劣Ⅴ类。可见,较平、枯水期相比,各监测断面处于丰水期时水质状况最好,由于水质状况受降水的影响较大,丰水期时辽河源头区水量大,河流的流动性好,对污染物的冲刷及稀释作用越大,污染物的含量及浓度随之降低;年平均水质评价结果与平水期相同,平水期水质状况一定程度上反映了年内水质的平均情况。

从辽河源至四双大桥的 8 个断面均位于东辽河干流上,汇合口、林家断面位于条子河上,四台子、新立屯、六家子位于招苏台河上。通过对表 3 进行统计分析,从河流断面水质状况来看,13 个断面的年均值水质评价结果中,10 个断面水质为Ⅴ类—劣Ⅴ类之间,约占 76.92%;3 个断面的水质级别在Ⅱ—Ⅲ类之间,仅占 23.08%,可见,研究区水体已严重受到污染。

从断面所处的位置特点来看,四双大桥作为东辽

河下游的一个主要断面,由于处于人类活动中心,接受了来自上游工业等的入河排污,因而其水质较差;汇合口、林家作为条子河的两个主要断面,其水质均为劣Ⅴ类,汇合口断面测得 COD、BOD₅、氨氮分别达到 290 mg/L,3.9 mg/L,20.63 mg/L,主要超标污染因子为 COD、BOD₅、高锰酸盐指数、氨氮。条子河承担着四平城区生活及工业、企业的污水排放,因而水体中 COD、氨氮超标最为严重。此外,河流污染严重亦由于条子河自身天然径流量小,自净能力较差有关;四台子、新立屯、六家子作为招苏台河主要监测断面,其水质级别劣Ⅴ类居多,超标因子主要为 COD、BOD₅、高锰酸盐指数、氨氮,经过调查发现,招苏台河径流量小,加之汇入了吉林省梨树县排放的工业及生活污水,致其污染较重。从水体中污染物浓度上对比可知,招苏台河水质略好于条子河。但从河流水体状况上看,两条河流污染均非常严重。

在采用 BP 神经网络方法评价水质的基础上,参考 2010 年吉林省、辽宁省环境公报中关于辽河流域吉林省境内主要代表断面的水质评价结果,进行对比,见表 4。

表 4 水质评价结果对比

各水期	辽河源		拦河闸		四双大桥		汇合口	
	A	B	A	B	A	B	A	B
丰水期	Ⅱ类	Ⅱ类	Ⅲ类	Ⅲ类	Ⅳ类	Ⅳ类	劣Ⅴ类	劣Ⅴ类
枯水期	Ⅱ类	Ⅱ类	Ⅲ类	Ⅲ类	Ⅳ类	劣Ⅴ类	劣Ⅴ类	劣Ⅴ类
平水期	Ⅱ类	Ⅱ类	Ⅲ类	Ⅲ类	Ⅳ类	Ⅴ类	劣Ⅴ类	劣Ⅴ类
年平均	Ⅱ类	Ⅱ类	Ⅲ类	Ⅲ类	Ⅳ类	Ⅴ类	劣Ⅴ类	劣Ⅴ类

注:A——环境公报中代表断面评价结果;B——BP 神经网络的水质评价结果。

对表 4 中评价结果统计可知,两者相同的占 81.25%,不同的仅占 18.75%,说明应用 BP 人工神经网络方法对研究区水体进行质量评价,具有较高的可靠性。且该方法简单便捷,运算较快,减少了大量人工运算,具有较强的通用性和适用性。与单因子水质评价方法相比,该方法考虑了多种水质指标的综合影响,评价结果较为客观,在多因子水质评价方法中亦存在较大优势。

4 结论

(1) 对年平均评价结果进行统计后可知,辽河源

头区的 13 个监测断面中,Ⅴ类—劣Ⅴ类水质所占比例最大,为 76.92%,仅有 23.08%的断面为Ⅱ—Ⅲ类水质,且东辽河干流的上游断面水质相对较好,而下游及支流断面的水质较差。

(2) 分析三条河流的区内水质评价结果可知,东辽河下游污染较为严重;条子河上的两个主要断面的水质均为劣Ⅴ类,由于该河接受四平城区生活、工业等污水的排放,加之河流天然流量小,使条子河水体反复污染的状况加重。招苏台河流经地区的污废水大都未经处理即排入该河中,河流水体严重污染,其流经断面的水质劣Ⅴ类居多。

(3) 将 BP 评价结果与环境公报中研究区内主要代表断面的水质评价结果进行对比后发现,81.25% 的断面评价结果相同,说明应用 BP 人工神经网络方法评价水质,具有较高的可靠性。

(4) 应用 BP 人工神经网络对辽河源头区域进行水质评价,以地表水环境评价标准作为标准训练样本,将水质级别数值化,评价结果客观、合理。与其它传统的水质综合评判方法相比,其优势在于避免了确定权重和隶属函数时,存在人为确定的主观因素,评价结果的精度较高,具有普适性和操作性强的特点。

参考文献:

- [1] 张利静,余麟,刘红琴,等.辽河源头区跨界污染输入响应模型的建立[J].科学技术与工程,2012,12(23):5952-5955.
- [2] 葛淑芳.辽河吉林省段水环境容量研究[D].长春:吉林大学,2012.
- [3] 马溪平,吕晓飞,张利红,等.辽河流域水质现状评价及其污染源解析[J].水资源保护,2011,27(4):1-4.
- [4] 孙维.基于 GIS 的辽河源头区地表水评价研究[D].长春:吉林大学,2011.
- [5] 郭华.不同评价法在清河水质评价中的应用[J].人民长江,2012,43(1):132-134.
- [6] 鲁斐,李磊.主成分分析法在辽河水质评价中的应用[J].水利科技与经济,2006,12(10):660-662.
- [7] 侯淑艳,曲晓晶,刘建新,等.基于综合评价法的天津市主要排污河流水质评价[J].水土保持研究,2011,18(4):61-64.
- [8] 周林飞,赵崧,孙佳竹,等.基于模糊模式识别的浑河水水质评价研究[J].水土保持研究,2012,19(5):163-166.
- [9] 刘登峰,齐实,韩小杰,等.缙云山不同土地利用类型地表径流水质评价[J].水土保持研究,2009,16(1):126-130.
- [10] 郭庆春,何振芳,李力,等.BP 神经网络在渭河水环境质量评价中的应用[J].水土保持通报,2011,31(4):112-115.
- [11] 张丽,吴金亮,杨国范.基于 BP 神经网络的东港灌区作物需水量预测研究[J].水土保持研究,2012,19(6):207-211.
- [12] 许秀英,黄操军,杨秋梦,等.基于遗传神经网络的黑龙江浅表地层水分预测[J].水土保持研究,2013,20(1):269-272.
- [13] 杜富芝,傅瓦利,杜小红,等.基于 BP 神经网络的三峡库区小流域水质评价[J].节水灌溉,2009(1):8-10.
- [14] 陈东彦.数学建模[M].北京:科学出版社,2007.
- [15] 杨芳,原松.基于 BP 神经网络的水环境质量评价模型的研建[J].人民长江,2008,39(23):46-48.
- [16] 马莉.MATLAB 数学实验与建模[M].北京:清华大学出版社,2010.
- [17] 杨洁,吴贻名,万飏,等.汉江水质综合评价的 BP 网络方法[J].武汉大学学报:工学版,2004,37(1):51-54.
- [18] 王李,刘志斌,常欢.水质评价的 BP 神经网络应用[C]//第三届全国农业环境科学学术研讨会论文集.2009.
- [19] 王庚,王敏生.现代数学建模方法[M].北京:科学出版社,2008.
- [20] 楼文高,王延政.基于 BP 网络的水质综合评价模型及其应用[J].环境污染治理技术与设备,2003,4(8):23-26.

(上接第 146 页)

- [5] 王钧,蒙吉军.黑河流域近 60 年来径流量的变化及影响因素[J].地理科学,2008,28(1):83-88.
- [6] 凌红波,徐海量,张青青,等.新疆塔里木河三源径流量变化趋势分析[J].地理科学,2011,31(6):728-733.
- [7] 王建,李硕.气候变化对中国内陆干旱区山区融雪径流的影响[J].中国科学:D 辑,2005,35(7):664-670.
- [8] 曹丽青,葛朝霞,薛梅.1951—2006 年新疆地区气候变化特征及其与水资源的关系[J].河海大学学报:自然科学版,2009,37(3):281-283.
- [9] 崔步礼,李小雁,李岳坦,等.青海湖流域河川径流特征及其对降水的滞后效应[J].中国沙漠,2011,31(1):247-253.
- [10] 陈忠升,陈亚宁,李卫红,等.塔里木河干流径流损耗及其人类活动影响强度变化[J].地理学报,2011,66(1):90-97.
- [11] 陈亚宁,郝兴明,徐长春,等.新疆塔里木河流域径流变化趋势分析[J].自然科学进展,2007,17(2):205-210.
- [12] 李红军,江志红,刘新春,等.阿克苏河径流变化与北大西洋涛动的关系[J].地理学报,2008,63(5):491-501.
- [13] 麦麦提吐尔逊·艾则孜,海米提·依米提,祖皮艳木·买买提,等.焉耆盆地土地利用变化对生态服务价值的影响[J].水土保持研究,2012,19(6):138-141.
- [14] 郭高轩,辛宝东.基于小波变换的北京地区 1724—2009 年降水量多尺度分析[J].水文,2012,32(3):29-33.
- [15] 桑燕芳,王栋.水文序列小波分析中小波函数选择方法[J].水利学报,2008,39(3):295-306.
- [16] 程正兴.小波分析算法与应用[M].西安:西安交通大学出版社,1998.
- [17] 王文圣,丁晶,李跃清,等.水文小波分析[M].北京:化学工业出版社,2005.
- [18] 毕雪丽.新疆博斯腾湖流域气候变化及地表径流响应[D].上海:华东师范大学,2012.