

BP 神经网络在渭河水环境质量评价中的应用

郭庆春^{1,3}, 何振芳², 李力³, 寇立群¹

(1. 陕西广播电视大学 教务处, 陕西 西安 710068; 2. 中国科学院 寒区旱区环境
与工程研究所, 甘肃 兰州 730000; 3. 中国科学院 地球环境研究所, 陕西 西安 710075)

摘要: 为准确和客观地评价渭河水环境质量状况, 将改进算法的 BP 神经网络引入地表水环境质量评价领域, 采用渭河 2010 年(9 月 7 日至 10 月 11 日)水质指标监测数据, 构建了渭河水环境质量评价模型。通过 3 个水质监测项目(氨氮、高锰酸盐指数、溶解氧)对渭河水质进行了评价, 评价结果与环境保护部公布结果完全一致。结果表明, BP 神经网络应用于水环境质量评价具有客观性、通用性和实用性, 可以很好地解决评价因子与水质等级间复杂的非线性关系, 评价方法简便可靠, 预测精度高。

关键词: BP 神经网络; 地表水; 环境质量评价; 渭河

文献标识码: B

文章编号: 1000-288X(2011)04-0112-04

中图分类号: X824

Application of BP Neural Network on Water Environmental Quality Evaluation of Weihe River

GUO Qing-chun^{1,3}, HE Zhen-fang², LI Li³, Kou Li-qun¹

(1. Teaching Affairs Office, Shaanxi Radio & TV University, Xi'an,

Shaanxi 710068, China; 2. Cold and Arid Regions Environmental and Engineering

Research Institute, Chinese Academy of Sciences, Lanzhou, Gansu 730000, China;

3. Institute of Earth Environment Research, Chinese Academy of Sciences, Xi'an, Shaanxi 710075, China)

Abstract: In order to evaluate the surface water environment quality accurately and objectively, the BP neural network with improved algorithms is introduced in water environmental quality evaluation. Water environment quality evaluation model of the Weihe River was constructed with recorded data of water quality during the period from Sep. 7 to Oct. 11 in 2010. Water quality of the Weihe River was evaluated through 3 water quality monitoring projects ($\text{NH}_3\text{-N}$, potassium permanganate index, DO), and the evaluation results by the BP neural network and by the environmental protection department were compared, and both evaluation results are identical. The results showed that the BP neural network model could well solve complex nonlinear relation between the evaluation factors and water quality levels, and the method was simple and reliable, and the forecast precision was high. The BP neural network for water environmental quality evaluation is objective, universal and practical.

Keywords: BP neural network; surface water; environmental quality evaluation; Weihe River

渭河是黄河最大的一级支流, 被誉为关中平原的“母亲河”。然而近年来, 随着渭河水量减少以及关中地区经济的快速发展和人口膨胀, 渭河流域出现了水资源供需矛盾突出、中下游防洪形势严峻、水污染日趋严重、水土流失治理进展缓慢等情况, 下游地区问题尤为突出。加之该地区缺乏完善的排水体系, 入渭污水量呈逐年上升的趋势。据陕西省环保厅 2007 年水质报告, 渭河的污染程度已高居陕西省 6 大江河之首, 成为黄河流域污染最严重的河流之一, 下游河段

均为 V 类(污染严重, 农业和景观用水)或劣 V 类(污染最严重, 无法使用水)水, 已基本丧失生态功能, 严重阻碍了区域经济和社会的发展。因此, 对渭河流域进行合理有效地水质预测、评价和治理是刻不容缓的。近年来, 水质评价方法的发展突飞猛进, 而且水质评价方法很多, 如灰色评价法、模糊评价法、指数评价法, 但都不能很好地解决评价因子与水质等级间复杂的非线性关系, 人工神经网络能够模仿人脑的功能, 具有对信息进行并行处理、分布式存储, 以及良好

的自适应性、自组织性、自学习与推理的能力,表现出容错性、非线性、非局域性、非凸性等特点,适于对模糊信息或复杂的非线性关系进行识别与映射。神经网络技术不断成熟和应用的日益广泛也为解决渭河严峻的水环境问题提供了强有力的理论和技术支持。

人工神经网络已被广泛地运用到水环境质量评价中^[1-12]但水环境质量评价中使用的大都是没有改进的BP神经网络,由于BP网络自身训练速度慢、容易陷入局部极小点等局限性,使得在网络设计中需要反复试凑和训练。针对渭河下游地区日益恶劣的水质污染状况,尝试将评价标准范围进行针对性的修改,利用rand函数对评价标准插入足够的训练样本,并将训练和测试样本进行归一化处理。运用改进的BP神经网络对渭河下游水质进行环境质量评价,取得了较为理想的评价结果。以此为渭河流域的重点治理和规划提供研究方法和依据,也为合理的利用流域资源提供理论参考。

1 数据与方法

1.1 实验观测数据

为了说明BP神经网络模型在地表水环境质量评价中的应用,本研究是对陕西省渭南—潼关吊桥断面的水质状况进行评价。数据来源于我国环境保护部数据中心,观测时段为2010年9月7日至10月11日,由于水质状况的改变主要是由于水体中pH、溶解氧(DO)、高锰酸盐指数和氨氮(NH₃-N)浓度变化造成的,陕西渭南—潼关吊桥断面主要污染指标为氨氮、溶解氧和高锰酸盐指数,且环境保护部数据中心目前只提供这些数据,而环境保护部也采用这些指标对水质进行评价,为了和环境保护部的结果对比,所以采用监测项目为pH值、溶解氧(DO)、氨氮(NH₃-N)、高锰酸盐指数等4项(表1)。

表1 2010年陕西渭南潼关吊桥断面水质状况

时间	pH值	溶解氧/ (mg·L ⁻¹)	高锰酸盐指数/ (mg·L ⁻¹)	NH ₃ -N/ (mg·L ⁻¹)	水质 等级
第41周	7.60	5.09	4.4	0.84	III
第40周	7.76	4.21	9.2	0.44	IV
第39周	7.86	5.14	8.6	0.91	IV
第38周	7.67	5.28	7.2	0.77	IV
第37周	7.88	2.23	8.5	1.02	V
第36周	7.61	3.69	8.6	1.21	IV

评价标准采用《地表水环境质量标准(GB3838—2002)》作为水质评价分级标准,pH值的评价标准在6~9之间,且由于监测数据中的pH值都在标准值

之内,因此在进行学习和评价时没有将pH值考虑在内,此次评价因子选用溶解氧(DO)、氨氮(NH₃-N)、高锰酸盐指数等3项评价指标(表2)。

表2 地表水环境质量标准

分类	mg/L		
	溶解氧	高锰酸盐指数	NH ₃ -N
I	≥7.5	≤2	≤0.15
II	≥6.0	≤4	≤0.50
III	≥5.0	≤6	≤1.00
IV	≥3.0	≤10	≤1.50
V	≥2.0	≤15	≤2.00

1.2 神经网络原理

BP算法即误差反向传播算法是神经网络学习中最常用的学习方法之一,它具有归纳、分类、容错性、冗余性、非线性处理、自组织及自适应性、学习能力、联想能力等特点。BP算法通过对神经计算得到的输出和样本值进行误差分析,不断反复修正神经网络中各权重系数和偏置量,从而使网络的输出接近所希望的输出,最终误差满足要求,即确定学习结束,获得学习成功后的权重系数和偏置量,由于在BP算法中误差计算是由输出层向输入层的方向进行,因此称为误差反向传播算法。

1.3 数据预处理

由于训练数据的样本数目偏少,而BP神经网络则需要较多的输入输出数据来训练网络。利用rand函数在各级评价标准中间按随机均匀分布方式内插训练数据的方法,针对训练样本过少的情况,在I—V类水质标准中,每2个相邻标准之间内插生成10组训练样本。

一般来说,在实际应用中,由于所采集的数据跨度较大,为提高训练速度和灵敏性以及有效避开Sigmoid函数的饱和区,防止计算过程出现“过拟合”等问题,需先将数据进行“标准化”处理,标准化方法可采用零均值标准差标准化方法或归一化等方法,在实践中发现,采用归一化方法效果较好,因此,要对输入数据进行预处理。为保证建立的模型具有一定的外推能力,最好使数据预处理后的值在0.2~0.8之间,把0~0.2和0.8~1.0的空间预留。数据预处理方法:

$$y = (0.8 - 0.2) \times \frac{(x - \min(x))}{(\max(x) - \min(x))} + 0.2$$

式中: x ——水质原始数据; y ——预处理后的数据; $\max(x)$, $\min(x)$ ——原始数据取值范围的最大值和最小值。

1.4 输入层和输出层神经元数目的选取

由于本次评价指标选择了3个评价因子,因此输

入层神经元节点数为 3 个。地表水环境质量标准中将地表水分为 5 类, 用于地表水评价的神经网络的学习样本为 40 个, 因而将该网络的输出神经元节点数定为 5, 每个学习样本的期望输出分别为: I 级(0. 9, 0. 1, 0. 1, 0. 1, 0. 1); II 级(0. 1, 0. 9, 0. 1, 0. 1, 0. 1); III 级(0. 1, 0. 1, 0. 9, 0. 1, 0. 1); IV 级(0. 1, 0. 1, 0. 1, 0. 9, 0. 1); V 级(0. 1, 0. 1, 0. 1, 0. 1, 0. 9)。

1.5 隐含层神经元数目的选取

地表水环境质量评价是一个非线性关系较为复杂的问题, 神经网络具有逼近任意非线性函数的能力, 且一般的 3 层 BP 神经网络模型就能够满足大部分非线性系统的要求, 因此本研究采用 1 个隐含层。分别组建了隐层节点数从 1~ 25 的 BP 网络, 最后根据试训效果, 确定了较为理想的隐层神经元数目为 7。

1.6 转换函数的确定

转移函数常用的有两种: sigmoid 函数和线性函数。本研究穷尽了这二种转移函数的搭配组合, 测试不同组合对网络性能的影响。结果表明, 当输入隐层为 tansig 函数、隐层—输出层为线性函数时, 效果最好。

1.7 训练函数的确定

traingdm (动量梯度下降反向传播算法)在对权重和阈值更新时不仅考虑当前的梯度方向, 而且还考虑了前一时刻的梯度方向, 从而降低了网络性能对参数调整的敏感性, 有效地抑制了局部极小, 但收敛速度较慢, 30 000 次的训练并把期望误差增大为 0. 05 时, 仍然难以找出最优的网络权值和阈值。

trainlm (Levenberg—Marquardt 优化方法)能够根据网络训练误差变化情况, 自动调节网络训练参数 μ , 使网络实时采取适宜的训练方法。当 μ 较小, 网络训练过程主要依据 Gauss—Newton 法; 当 μ 较大, 网络训练过程主要依据梯度下降算法, 该训练函数的效率优于梯度下降法, 但占用内存较大, 适用于规模较小的网络。

trainbr (Levenberg—Marquardt 优化方法与 Bayesian 正则化方法)函数采用 Levenberg—Marquardt 优化方法进行网络权值和阈值的最优化搜索, 并采用 Bayesian 正则化方法在网络训练过程中自适应地调节性能函数比例系数 γ 的大小, 使其达到最优, 并且采用 trainbr 函数训练后 BP 网络具有较好的推广能力。 trainlm 和 trainbr 函数均采用 Levenberg—Marquardt 优化方法, 能够快速获得最优的网络权值和阈值, 收敛速度较快。 traingdx (自适应学习速率动量梯度下降反向传播训练函数), 对于梯度

下降法, 训练成功与否与学习率的选取有很大关系, 而自适应学习算法能够自适应调整学习率来增加稳定性, 提高速度和精度。 trainbfg (BFGS 拟牛顿回退法)收敛速度快, 但需要更多内存, 与共轭梯度法训练参数相同, 适合规模较小的网络。研究中将 trainbr , trainlm , traingdm , traingdx , trainbfg 等方式做了对比分析, 最后确定训练方式为 traingbr 。 traingbr 收敛速度较快, 拟合精度高, 而且通过修正神经网络的训练性能函数来提高推广能力, 防止出现“过度训练”, 采用 trainbr 函数训练的 BP 网络稳定性要优于 trainlm 函数和 traingdm 函数等其它函数, 且通常只需几百次的训练就能使网络收敛, 找到最优的网络权值和阈值。

2 结果与分析

利用《地表水环境质量标准(GB3838—2002)》中的溶解氧(DO)、氨氮($\text{NH}_3\text{—N}$)、高锰酸盐指数等 3 项评价指标及其插值形成的数据作为输入向量, 即训练样本, 把学习样本的期望(相应的评价等级)作为输出向量。例如, 把地表水环境质量标准中的 I 类指标的数据作为输入向量, 那么输出变量就是 I 级(0. 9, 0. 1, 0. 1, 0. 1, 0. 1); 如果把地表水环境质量标准中的 I 类和 II 类之间的插值作为输入向量, 那么输出变量是 II 级(0. 1, 0. 9, 0. 1, 0. 1, 0. 1); 其它类似操作。随机赋予初始权重系数和偏置量, 均方差为 0. 005, 当学习迭代次数为 547 步时, 网络收敛并达到所需精度, 连接权重系数和偏置量就确定了。此时网络模型已建立了该评价标准与评价级别之间的复杂对应关系, 该模型已能够作为地表水水质评价模型。经过上述训练, 用调整后的权重系数和偏置量, 再把陕西省渭南市潼关吊桥断面 6 个检验样本作为 BP 神经网络输入向量, 用已掌握知识信息的网络模型进行评价, 按输出值与期望值的接近程度去判定其归属哪一级, 其结果如表 3 所示。

表 3 2010 年陕西渭南潼关吊桥断面数据的训练结果

时间	BP 神经网络训练结果					等级
第 41 周	0. 213	0. 152	0. 864	0. 374	0. 051	III
第 40 周	0. 243	0. 091	0. 085	0. 834	0. 229	IV
第 39 周	0. 022	0. 144	0. 124	0. 875	0. 021	IV
第 38 周	0. 124	0. 231	0. 142	0. 845	0. 054	IV
第 37 周	0. 216	0. 302	0. 158	0. 312	0. 886	V
第 36 周	0. 168	0. 278	0. 148	0. 891	0. 035	IV

从实验结果看, 2010 年第 41 周水质等级为 III 类; 2010 年第 40, 39, 38, 36 周水质等级为 IV 类; 2010

年第37周水质等级为V类。为验证水质评价结果的可靠性,将BP神经网络评价结果和我国环境保护部公布的结果对比,两者评价结果完全一致。改进后的BP神经网络评价结果的准确性和精度都非常高,结果更为实际、客观、合理,很好地解决了评价因子与水质等级间复杂的非线性关系,而且泛化能力好,还可选用更多的参数指标进行学习,程序具有广泛的适用性。

环境保护部数据中心采用单因子评价方法,评价结果为2005—2009年的同一时期(第36周至第41周)水质等级基本为劣V类,污染严重,基本丧失生态功能。2010年第36周至第41周水质主要为IV类,说明水质有所好转。

3 结论

(1) 应用改进算法的BP神经网络,利用《地表水环境质量标准(GB3838—2002)》中的溶解氧(DO)、氨氮(NH₃-N)、高锰酸盐指数等3项评价指标及其插值形成的数据作为训练样本,利用渭河下游的数据作为检验样本,得到合理的评价等级,其评价结果和我国环境保护部公布的结果完全一致。

(2) 利用BP神经网络模型评价地表水环境质量是可行的,该模型具有很强的学习和联想功能,计算速度快,泛化能力较好,其分析结果和过程都接近人脑的思维过程和分析方法。

(3) 应用改进算法的BP神经网络进行地表水环境质量评价,只需将地表水环境质量标准作为训练样本来进行网络的训练,当训练达到误差要求后,就可以用来进行地表水环境质量评价。改进后的BP神经网络计算速度快、方法简单、评价结果客观准确,有推广应用前景,为水环境质量评价开辟了一条新途径。

(4) 应用改进算法的BP神经网络对地表水水环境质量进行评价,所得结果客观、准确、可靠,反映了应用当代最新技术成果的益处。随着计算机技术的发展,利用神经网络来解决一些模糊性和不确定性问题,具有简便、易行、易于普及推广等特点。

(5) 本研究只针对陕西省渭南市潼关吊桥断面

2010年(第36周至第41周)水质做了评价分析,由于数据搜集存在较大困难,因此下一步工作希望获取更新更全面的数据,这样才能更准确地反映水质状况;同时希望将研究范围向支流扩展,形成完整的区域水环境研究体系,并搜集渭河中上游河段历年详细的水质水文资料,将研究范围从空间和时间上向上推移,全面地研究渭河水环境演变的规律和特征,进一步分析污染成因及制定对策,为水质分析及治理提供借鉴。

[参 考 文 献]

- [1] 李兴旺,董曼玲.地面水质评价的RBF神经网络方法[J].水土保持通报,2002,22(3):51-54.
- [2] 刘金生,周焕银,刘金辉.基于BP神经网络的抚河水环境质量评价研究[J].东华理工大学学报:自然科学版,2008,31(1):85-88.
- [3] 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.
- [4] 冯梅.水环境质量评价、预测的数学模型探讨及其应用[J].运筹与管理,2008,17(6):163-169.
- [5] 朱长军,李文耀,张普.人工神经网络在水环境质量评价中的应用[J].工业安全与环保,2005,31(2):27-29.
- [6] 杨国栋,王肖娟,尹向辉.人工神经网络在水环境质量评价和预测中的应用[J].干旱区资源与环境,2004,18(6):10-14.
- [7] 刘连芳,魏砾宏,李爱民,等.神经网络模型在辽河水环境质量评价中的应用[J].城市环境与城市生态,2003,16(6):251-253.
- [8] 闫同海,张百奇.基于BP神经网络的地表水环境质量评价[J].中国环境管理,2003,22(6):56-57.
- [9] 张文艺.基于人工神经网络的地面水环境质量评价模型[J].环境保护科学,2000,26(4):40-42.
- [10] 芦春梅,马成有,曹剑锋.BP神经网络在地下水环境质量评价中的应用[J].检验检疫科学,2008,18(6):48-51.
- [11] 李磊,孙卉,翟秋,等.RBF神经网络在平顶山市地表水评价中的应用[J].安徽农业科学,2008,36(26):11514-11516.
- [12] 陈丽华,马德山.黄河水质综合评价的人工神经网络模型[J].西北民族学院学报:自然科学版,2002,23(3):28-32.