

基于 EXCEL 的 BP 网络法在水质评价中的应用

胡 晋

(常州市环境监测中心, 江苏 常州 213001)

摘 要: 阐述了在 Excel 平台上建立的 BP 网络模型水质评价系统的方法和要点, 以及网络模型结构的设计、训练数据的处理、网络的训练和评价结果的仿真情况。实例表明 Excel 建立的 BP 网络模型水质评价系统的评价结果客观, 模型使用方便, 有推广价值。

关键词: BP 网络; Excel; 水质评价

中图分类号: X824

文献标识码: A

文章编号: 1006-2009(2005)03-0018-04

Application of BP Network Based on EXCEL in Water Quality Evaluation

HU Jin

(Changzhou Environmental Monitoring Center, Changzhou, Jiangsu 213001, China)

Abstract: To establish BP network model water quality evaluation system based on Excel platform. The design of network model structure, treat of train data and simulation of evaluation result were discussed. The evaluation result was objective.

Key words: BP network; Excel; Water quality evaluation

近年来, 国内外用于水质综合评价的方法很多, 如模糊综合评价^[1]、灰色关联分析^[2], 以及物元模型法等^[3]。由于影响水质的因素很多, 且因素与水质类别之间通常存在复杂的非线性关系, 所以迄今还没有一种统一和公认的方法。在模式识别领域中, 人工神经网络以其自学习和较好的容错性, 以及优势的非线性逼近能力而表现出很好的特性, 将其运用于水质综合评价中也是可行的, 在实际应用中, 80%~90%的人工神经网络模型是采用误差反传算法或其变化形式的网络模型(简称 BP 网络)。

目前, 可以使用的神经网络主要应用 MATLAB 6.5 或者自己编写神经网络的算法程序, 但由于 MATLAB 6.5 来源比较困难, 对于环境监测人员来说, 自己编写程序也十分繁琐。而 Microsoft Excel 是环境监测人员常用的工具软件, 它在数据综合管理和分析方面具有功能强大、技术先进和使用方便等特点, 再加上有 Youngfan (<http://www.aitech.cn/>, 可以下载使用) 编写的 BP 网络模型在 Excel 中的宏插件, 故可以运用 Excel 实现 BP 网络方法

在水质综合评价中的应用。

1 BP 神经网络原理

BP 网络是一种具有 3 层或 3 层以上的阶层型神经网络。它的特点是各层神经元之间无反馈连接, 各层内神经元之间无任何连接, 仅相邻层神经元之间有连接。

典型 BP 网络是 3 层前馈网络, 即输入层、隐层和输出层, 各层之间实行全连接, 见图 1。

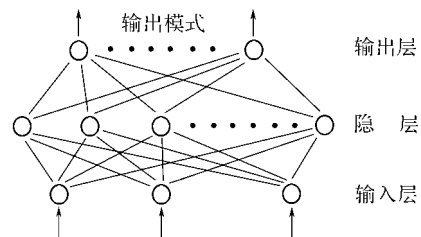


图 1 典型 BP 网络

收稿日期: 2004-06-23; 修订日期: 2005-03-10

作者简介: 胡 晋 (1980—), 男, 江苏常州人, 助理工程师, 学士, 从事环境监测工作。

BP 网络法的主要思想可概括为训练本集 P ($P^1, P^2 \dots P^K$, K 为样本数) 和已知的输出样本集 T ($T^1, T^2, \dots T^K$)。训练的目的在于求出网络的模拟输出 A ($A_1, A_2, \dots A_K$), 通过减少 A 与 T 之间的误差来修改模拟过程的权值, 使网络模拟输出值与实际样本值之间的误差达到最小值。每一次的权值变化和偏差都与网络误差的变化成正比, 并以反相传播的方式传递到每一层。BP 网络法由信息的前向传递和误差的正相传播两部分组成^[4]。

2 水质综合评价的 BP 网络方法

2.1 各个神经元的确定

在 BP 神经网络中, 一般把实测的水质参数作为输入层神经元, 输出层即为水质类别用一个神经元表示, 隐层要反复比较确定, 建立的网络模型见图 2。

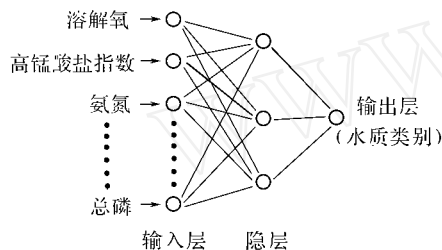


图 2 水质评价的 BP 网络结构

在水质评价中, 需要根据评价的目的、要求和水的用途选择评价标准, 通常采用的评价标准为 GB 3838 - 2002《地表水环境质量标准》^[5]。现将水质类别用数值表示, 并由输出层的数据来决定水质类别。

2.2 标准化处理

对收集到的训练样本, 一般不直接拿来使用, 要先进行一定的标准化处理。但由于各分指标具有不同的量纲, 且类型不同, 故指标间具有的不可公度性难以直接进行比较。因此, 在综合评价前, 必须把这些分指标按某种规则归一化到某一无量纲区^[6]。

标准化处理的方法很多, 可采用初值化、级差化或等比例变换方法, 因为 BP 网络中 Sigmoid 函数的输出范围在 0 ~ 1 之间, 所以一般也把数据归一到 [0, 1]。

对于成本型因子 (如 BOD₅、COD 等), 其数值越大表示水质越差, 可用下式标准化:

$$z_{ij} = \begin{cases} 1, & x_{ij} > s_{ik} \\ x_{ij} / s_{ik}, & 0 \leq x_{ij} \leq s_{ik} \end{cases} \quad (1)$$

式中:

x_{ij} ——实测数据;

s_{ik} ——水质评价标准中最差的一级数值。

对于效益型因子 (如 DO), 数值越大表示水质越好, 可用下式标准化:

$$z_{ij} = \begin{cases} 0, & x_{ij} > s_{il} \\ 1 - x_{ij} / s_{il}, & 0 \leq x_{ij} \leq s_{il} \end{cases} \quad (2)$$

式中:

x_{ij} ——实测数据;

s_{il} ——水质标准中最好一级数值。

3 应用

选取 2003 年京杭运河常州段 7 个断面的污染物年平均值, 并根据常州市工业特点, 以溶解氧 (DO)、高锰酸盐指数 (M_n)、生化需氧量 (BOD₅)、氨氮 (NH₃ - N)、总磷 (TP) 和石油类为评价指标, 监测结果和评价标准见表 1、表 2。

表 1 2003 年京杭运河常州段水质监测结果 mg/L

项目	DO	M_n	BOD ₅	NH ₃ - N	TP	石油类
新河口	6.9	4.8	4.4	1.03	0.242	0.24
连江桥	6.6	4.9	3.9	0.75	0.197	0.37
河水厂	6.4	5.4	3.5	0.93	0.167	0.08
水门桥	5.3	6.1	6.8	1.69	0.277	0.44
石化厂	4.8	6.7	6.4	1.84	0.258	0.35
戚墅堰	3.1	7.4	6.6	2.29	0.177	0.51
横洛间	2.7	7.4	5.5	1.94	0.251	0.31

表 2 地表水分级评价标准 mg/L

级别	DO	M_n	BOD ₅	NH ₃ - N	TP	石油类
DO	7.5	6	5	3	2	
M_n	2	4	6	10	15	
BOD ₅	3	3	4	6	10	
NH ₃ - N	0.15	0.5	1.0	1.5	2.0	
TP	0.02	0.1	0.2	0.3	0.4	
石油类	0.05	0.05	0.05	0.5	1.0	

为了保证有足够的训练样本和训练精度, 需将分级标准进行插值加密, 可采用线性插值的方法将分级标准加密 1 倍, 即把标准分为更小的级别。在 Excel 中可以方便地使用其公式编辑功能, 进行线性插值计算。由于 BP 网络中 Sigmoid 函数的输出范围在 0 ~ 1 之间, 所以对模型的输出值也要在

[0, 1] 上划分, 对应于水质标准的 5 个级别, 可得到与模型输出相对应的分级环境标准, 见表 3。

表 3 模型的水质分级指标

水质级别
数据区间
0~0.2
0.2~0.4
0.4~0.6
0.6~0.8
0.8~1.0

根据表 3 的模型水质分级指标, 对通过插值得到的分级标准数值进行标准化处理, 参照公式 (1)、公式 (2), 在 Excel 中编写计算公式, 得到标准化后的数据, 再将相对应的输出值也写在其中, 便于 BP 网络运算, 见表 4。

表 4 标准化后的水质分级标准

DO	I_{Mn}	BOD ₅	NH ₃ - N	石油类	TP	类别值
0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.0
0.000 0	0.066 7	0.150 0	0.037 5	0.025 0	0.025 0	0.1
0.000 0	0.133 3	0.300 0	0.075 0	0.050 0	0.050 0	0.2
0.100 0	0.200 0	0.300 0	0.162 5	0.050 0	0.150 0	0.3
0.200 0	0.266 7	0.300 0	0.250 0	0.050 0	0.250 0	0.4
0.266 7	0.333 3	0.350 0	0.375 0	0.050 0	0.375 0	0.5
0.333 3	0.400 0	0.400 0	0.500 0	0.050 0	0.500 0	0.6
0.466 7	0.533 3	0.500 0	0.625 0	0.275 0	0.625 0	0.7
0.600 0	0.666 7	0.600 0	0.750 0	0.500 0	0.750 0	0.8
0.666 7	0.833 3	0.800 0	0.875 0	0.750 0	0.875 0	0.9
0.733 3	1.000 0	1.000 0	1.000 0	1.000 0	1.000 0	1.0

通过应用 Youngfan 编写的神经网络 Office 插件 NNX, 按照说明在 Excel 中进行 BP 网络训练。由于评价因子有 7 个, 可以确定网络的输入节点为 7 个, 并且每个断面的水质评价结果只有 1 个, 所以网络的输出节点数也为 1 个, 隐层节点数经过反复训练、比较, 参照网络训练的误差情况, 最后确定为 5。其他的计算参数是学习率为 0.01, 算法采用常规 BP 法, 训练次数为 15 000 次。训练过程中的学习误差推移见图 3。

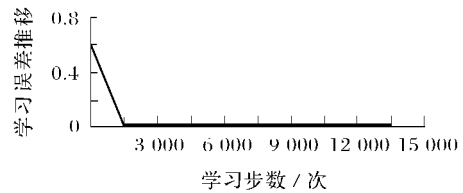


图 3 训练过程中学习误差推移

为了便于以后仿真计算使用, 将训练后的 BP 网络保存为 Excel 文件, 然后将监测数据标准化, 见表 5。

表 5 标准化后的水质监测数据

类 别	DO	I_{Mn}	BOD ₅	NH ₃ - N	TP	石油类
新河口	0.041 7	0.320 0	0.44	0.515	0.605 0	0.24
连江桥	0.083 3	0.326 7	0.39	0.375	0.492 5	0.37
河水厂	0.111 1	0.360 0	0.35	0.465	0.417 5	0.08
水门桥	0.263 9	0.406 7	0.68	0.845	0.692 5	0.44
石化厂	0.333 3	0.446 7	0.64	0.920	0.645 0	0.35
戚墅堰	0.569 4	0.493 3	0.66	1.14	0.442 5	0.51
横洛间	0.625 0	0.493 3	0.55	0.970	0.627 5	0.31

将监测数据输入训练好的网络进行仿真计算, 得到的输出值, 以及对照模型分类标准得到的水质综合评价结果见表 6。

表 6 水质综合评价结果

类 别	新河口	连江桥	河水厂	水门桥	石化厂	戚墅堰	横洛间
输出值	0.57	0.53	0.52	0.79	0.80	0.90	0.82
水质级别							

按照水质综合评价结果,结合运河断面顺序和运用 Excel 建立的运河断面水质情况见图 4。

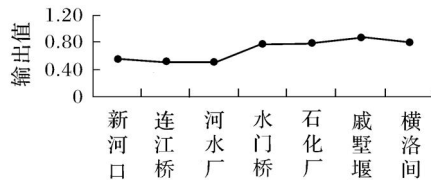


图 4 常州市运河断面沿程水质情况

由水质综合评价表及运河断面水质图可知,京杭运河常州段水质情况不容乐观,虽然上游水质情况较好,但进入市区后的运河水质开始恶化,石化厂段水质超过标准的水体功能区要求,说明市区向运河的排污量已经超过了水体的承受能力。

4 结语

(1)神经网络评价方法是一种比较接近人类思维模式的定量与定性结合的综合评价模型,不需要对各种评价指标权值做出认为的规定,在学习过程中会自适应调整,评价结果具有客观性。

(2)用 Excel 建立的 BP 神经网络模型使用方便,只要将网络模型训练好,然后保存,应用时只需将新测的数据标准化后输入模型,即可得出综合评价结果,且评价结果非常客观、合理,完全可以应用于实际的水质综合评价工作。

(3)建立的评价模型方法既适应于水环境质量评价,对其他一些环境质量评价也同样适用,如

大气环境质量评价、富营养化程度评价等。特别是对参数和等级较多的综合评价,只需在标准样本训练时改变输入节点数和隐层节点数即可。因此,该方法有较广的适用范围。

(4)运用 Excel 建立的 BP 神经网络系统评价环境质量,不需要繁琐的编程,只需了解神经网络 Office 插件 NNX 的使用方法,从而节省了程序设计、调试及网络学习所要的时间,可以集中精力思考解决问题的模型,以提高环境质量评价的效率和质量。

(5)基于 Excel 建立的 BP 网络模型水质评价系统,软件来源方便,使用简单,有很大的推广价值。

[参考文献]

- [1] 国家环境保护总局监督管理司. 中国环境影响评价 [M]. 北京: 化学工业出版社, 2000. 217 - 243.
- [2] 吴业文, 戈建民, 黄奕龙. 应用灰色关联分析进行城市地表水环境质量评价 [J]. 世界地质, 2000, 19 (1): 53 - 56, 65.
- [3] 朱继业, 龚贻俭, 方红松. 动态系统物元模型在综合水质预报中的研究和应用 [J]. 城市环境与城市生态, 1999, (2): 51 - 53.
- [4] 曹焕光. 神经网络原理 [M]. 北京: 气象出版社, 1992. 67 - 82.
- [5] 王蜀南, 王鸣周. 环境水文学 [M]. 北京: 中国水利出版社, 1996.
- [6] 戴文战. 基于三层 BP 网络的多指标综合评估方法及应用 [J]. 系统工程理论和实践, 1999, (5): 29 - 40.

本栏目责任编辑 张启萍

· 简讯 ·

南京市环境噪声在线自动监测试点工作正式启动

根据国家环境监测总站和江苏省环境监测中心的部署,已将南京市环境监测中心站列为环境噪声在线自动监测试点单位之一。目前由哈尔滨市环境信息中心提供的“环境噪声在线自动监测系统”已安装在典型工业噪声类型的南京扬巴工业区,经一个月时间的运行实验,系统工作正常。

该系统是利用传感、通讯、计算机及网络技术对噪声污染进行实时监测,以图形、报表形式通过网络及时、准确地传达到环境管理部门实施监管和处理。

该项试点将为《全国环境噪声在线自动监测规范》的制定作技术准备,同时还将在全市典型交通、施工、生活噪声区域及机动监测方面进行试点。

南京市环保网