

人工神经网络在苏州空气污染治理中的应用

白晓平¹, 张启明¹, 方 栋¹, 洪维民², 刘峰磊², COSTABILE Francesca³, 王芬娟¹

1. 清华大学核能与新能源研究院, 北京 100084

2. 苏州环境监测中心站, 江苏苏州 215004

3. 意大利国家研究委员会大气污染研究所, 罗马, 意大利

[摘要] 人工神经网络在预测预报领域的应用越来越广泛。简单介绍了 BP 神经网络的基本原理, 较详细地回顾了国内 BP 神经网络在空气污染预报领域的研究应用情况, 并建立了苏州市区 SO₂ 浓度预报的 BP 神经网络, 预报结果表明: 该模型具有简便、快速、准确的优点, 可推广用于其它空气污染物的预报。

[关键词] 空气污染预报; 神经网络; BP 神经网络

[中图分类号] X823

[文献标识码] A

[文章编号] 1000-7857(2007)03-0045-05

Application of Artificial Neural Network to Air Pollution Prediction in Suzhou City

BAI Xiaoping¹, ZHANG Qiming¹, FANG Dong¹, HONG Weimin², LIU Fenglei², COSTABILE Francesca³, WANG Fenjuan¹

1. Institute of Nuclear and Energy Technology, Tsinghua University, Beijing 100084, China

2. Suzhou Environmental Monitoring Centre, Suzhou 215004, Jiangsu Province, China

3. Institute of Atmospheric Pollution, Italian National Research Council, Rome, Italy

Abstract: The artificial neural network has been widely used in the forecast and prediction due to the development of its theoretical basis and computer technology. The basic principle of BP neural network is briefly introduced, and the applications of BP neural network to air pollution prediction in China are reviewed in detail. It is applied to predict SO₂ concentration in Suzhou, and it is shown that BP neural network is easy to implement and gives quite accurate results, and that it can also be used to predict other air pollutants such as PM₁₀, NO_x, and so on.

Key Words: air pollution prediction; artificial neural network; BP neural network

CLC Number: X823

Document Code: A

Article ID: 1000-7857(2007)03-0045-05

空气污染物浓度的时空分布受到气象场、排放源、复杂下垫面、理化生过程的耦合等多种因素的影响, 具有较强的非线性特性^[1]。人工神经网络(Artificial Neural Network, ANN)简称神经网络, 是用大量的简单处理单元(神经元)组成的非线性动力学系统, 它具有自学习、自组织、自适应和较强的容错性等特点, 是描述和刻画

非线性现象的一种有效的工具, 已经广泛应用于信号处理、目标跟踪、模式识别、预测、运输与通信等众多领域。神经网络用于空气污染预报领域, 始于 20 世纪 90 年代初。目前, 神经网络主要用于空气污染物浓度的短期预报、空气污染指数的预报, 尤其适用于当前无法开展空气污染数值预报的城市。

收稿日期: 2006-11-24

作者简介: 白晓平, 女, 北京市海淀区清华大学核能与新能源技术研究院, 博士生, 主要研究方向为空气污染预报;

E-mail: bxp03@mails.tsinghua.edu.cn

方 栋(通讯作者), 男, 北京市海淀区清华大学核能与新能源技术研究院, 研究员, 主要研究方向为辐射防护与环境保护;

E-mail: dfang@tsinghua.edu.cn

神经网络的实现方法主要有软件和硬件两种, 其中, 软件是比较常用的方法, 特别是一些仿真语言和工具(如 MATLAB) 的出现, 提高了软件实现的效率。本文将较详细地回顾国内 BP(Back Propagation) 网络在空气污染预报中的研究应用情况, 并以苏州市的 SO_2 浓度为研究对象, 建立基于 MATLAB 的 BP 网络, 并进行预报试验。

1 BP 网络基本原理

BP 网络(即误差反向传播神经网络)是神经网络模型中使用最广泛的一类。从结构上讲, BP 网络属于多层网络, 包含输入层、隐含层和输出层, 各层之间实行全连接, 前层单元的输出不能反馈到更前层, 同层单元之间也没有连接, 如图 1 所示。BP 网络的基本原理是: 当给定网络一个输入模式时, 它由输入层单元传递到隐含层单元, 经隐含层单元逐层处理后再传递到输出层单元, 由输出层单元处理后产生一个输出模式。这是一个逐层状态更新过程, 称为前向传播; 如果输出响应与期望输出模式有误差, 不满足要求, 则转入误差反向传播, 将误差值沿连接通路逐层反向传送并修正各层连接权值, 当每个训练模式都满足要求时, 学习结束。

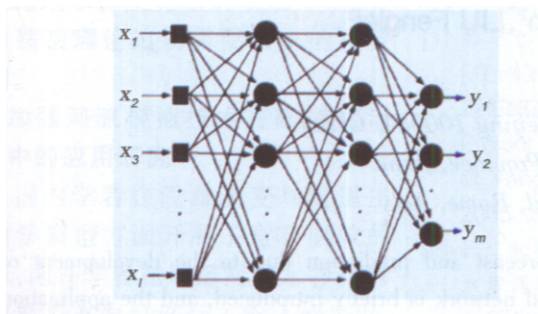


图 1 4 层 BP 网络

Fig. 1 BP neural network with four layers

2 BP 网络在国内空气污染预报中的应用研究进展

在国内, 多以 BP 网络展开空气污染预报研究。1997 年, 李祚泳等^[2]分析了某市 SO_2 浓度与工业耗煤量、人口密度、交通密度、生活服务点等 4 个因子的相关性, 并建立了带有 4 个输入节点、3 个隐节点、1 个输出节点的 SO_2 浓度预报的 BP 网络。实例证明: 只要学习样本具有代表性, 网络经训练后就能自动获得预测因子和预测量之间的合理规则。1998 年, Mok 等^[3]建立了 3 层 BP 网络用于澳门短期 SO_2 浓度预测, 两个测试期模型的误差分别为 14.45% 和 13.71%, 结果表明: 即使使用非常有限的训练数据, 网络仍可得到准确的预测结果。2001 年, 赵慧宏^[4]对 1995-2000 年兰州市工业 SO_2 排放总量进行分析后发现: SO_2 排放总量年变化数据具有非线性特性, 并以 MATLAB 软件建立 BP 网络进行仿

真试验, 结果较为理想。2002 年, 王俭等^[5]根据沈阳市 1999 年秋季 NO_x 与气象要素的关系建立了 NO_x 小时预报的 BP 网络, 网络的输入因子为大气稳定度、进行观测的北京时间、温度、云量、风向、风速、前一时刻 NO_x 含量, 网络的隐含层节点个数为 11, 预报试验表明该网络能够很好捕捉气象要素与空气污染物之间内在的规律性。针对传统 BP 网络存在训练速度较慢、局部极值以及最佳网络结构无法准确确定的问题, 王俭等^[6]对传统 BP 网络进行了改进研究, 即采用变步长和加动量项的方法来加速网络的收敛和避免振荡现象, 采用遗传算法来确定最佳网络结构, 并根据沈阳市 1999 年冬季的气象数据和污染物浓度数据建立改进的 BP 网络, 得到了令人满意的结果。2003 年, 马雁军等^[7]用本溪市 1995 年和 1996 年冬季的数据建立了冬季 TSP (Total Suspended Particulate) 和 NO_x 浓度日均值预报的 BP 网络, 网络的输入因子为源强、初始浓度、风速、风向、天气云况、日照、温度、相对湿度, 隐含层含有 8 个神经元; 采用 1997 年的数据进行预报试验, TSP 和 NO_x 的计算值与监测值的相关系数分别为 0.768 和 0.785。王灿星等^[8]采用 1999-2001 年杭州市区 TSP 实测数据作为 BP 网络的训练样本, 气压、湿度、温度、风速、雨量等为网络的输入因子, 建立 BP 网络, TSP 预报结果与实测值的性能函数均方差小于 0.2%。杨树平等^[9]根据气象部门和环境部门的数据分别建立了 SO_2 、 NO_2 和 PM_{10} 的 BP 网络, 网络的输入节点数、隐层节点数、输出节点数分别为 10, 5, 1, 将 BP 网络的预测值与实测值进行比较后发现二者具有较高的逼近精度, SO_2 、 NO_2 和 PM_{10} 平均绝对误差分别为 0.006, 0.004, 0.009。万显列等^[10]建立了大连市 BP 网络进行 O_3 浓度的预报, 网络的输入因子为风速、风向、相对湿度、云量、平均气温、最高气温, 网络的隐含层节点数为 5, 输出节点数为 1; 网络的预测值与实测值的平均相对误差为 21.49%, 相关系数为 0.837。2004 年, 蒋大和等^[11]采用即时训练、较小的学习率以及使用验证数据组来改进传统 BP 网络, 应用改进后的 BP 网络预报上海市 TSP、 SO_2 和 NO_x 的污染指数, 发现经改进后的 BP 网络的预报效果大大提高。2005 年, 黄世芹^[12]用泛化改进后的 BP 网络分别建立贵阳市 4 个季节的 SO_2 、 NO_2 、 PM_{10} 日均浓度预报模式, 结果表明: 该方法很好地解决了在传统 BP 网络训练误差很小时, 一个新的输入与对应的目标输出具有较大误差的问题。吴小红等^[13]针对大气环境参数预测精确性和可靠性的要求提出一种基于 BP 网络快速收敛算法, 该算法根据学习误差、误差梯度及梯度率自动适时地调节动量因子 α 的值, 使得过去权重的变化对当前权重变化的影响程度实现自适应, 仿真试验表明: 该算法切实可行。黄海洪等^[14]采用 BP 网络与主成分分析相结合的方法建立南宁市 SO_2 浓度预报模型, 结果表明: 主成分 BP 网络模型通

过浓缩信息、降维去噪可提高预报泛化性能及预报准确率。李少华等^[15]分别建立 NO_x 小时预测的 RBF (Radial Basic Function) 网络和 BP 网络, 将预报结果进行比较后发现 RBF 网络具有更好的预报结果。刘永等^[16]利用 BP 网络对城市的 SO_2 浓度进行预测, 并与多元统计预测方法和模糊识别预测方法的预测结果进行了比较分析。周秀杰等^[17]在确定了空气污染指数影响因子的基础上, 综合考虑 BP 网络的逼近能力和泛化能力, 提出空气污染指数 BP 网络, 并将 BP 网络和逐步回归法进行对比试验, 结果表明: BP 网络的预报准确度明显高于逐步回归法, 特别是对骤升骤降趋势也能得到准确度较高的预报结果。

3 苏州空气污染预报的神经网络方法

3.1 项目的提出

意大利环境与领土部和中国国家环保总局于 2002 年 6 月 19 日就“中意环保合作示范项目——苏州空气质量实时监测系统”签署了协议, 该协议旨在苏州首先建立一套空气质量实时监测系统 (Air Quality Monitoring System, AQMS), 应用成功后在中国其他城市推广普及。该系统包括 1 个数据质量控制中心、6 个先进的移动监测站、20 个饱和监测站、9 个固定监测站和 1 个传统移动站。2004 年 5 月 24 日, 苏州空气质量监测系统项目通过验收, 该项目的完成大幅增加了苏州市区空气自动监测网的覆盖面积, 加强了监测数据的分析能力, 提高了苏州市的空气质量监测水平, 为当地政府在环境管理、环境规划和污染控制等方面提供了更有效的决策依据。

为了充分利用 AQMS 监测数据对苏州空气质量进行评价和预报, 意大利国家大气污染研究所、苏州环境监测中心站和清华大学于 2005 年开始合作, 利用神经网络进行苏州空气质量评价和预报的应用研究。该合作项目由苏州环境监测中心站负责提供气象数据和空气污染物监测数据, 清华大学负责苏州空气质量评价、预报的方法和技术的研究。

3.2 数据来源与处理

3.2.1 数据来源

本项目所使用的数据主要有 3 个来源: 苏州市环境监测中心站、苏州市气象台、AQMS。其中, 苏州市环境监测中心站的数据有两种形式: 每月一份/每日一份的污染数据, 包括 PM_{10} 、 CO 、 NO_2 、 SO_2 , 目前所开展的试验都是以 SO_2 为研究对象进行的; 苏州市气象台的数据为每日一份的气象数据, 包括当天 2:00、8:00、14:00、20:00 4 个时刻的气温、露点、风速、风向、云量等数据; AQMS 监测系统的数据为每日一份的 FoxPro 文件, 在仪器正常情况下, 每隔 5 min 记录一组数据, 每天共计 288 组数据, 每组数据中包括温度、湿度、气压和太阳辐

射 4 项。

3.2.2 数据处理

由于仪表和记录的原因, 一些数据存在不完整性和误差, 此外, 一部分所需的数据不是以数值的方式给出, 因此还需要对这些数据进行量化和处理。

风向的输入采用赋权输入, 即

$$p(\theta) = \cos(\theta - 3\pi/4)$$

其中, θ 为风向的弧度值, $p(\theta)$ 为风向赋值后的对应值。在后期的训练过程中, 事实证明该赋权方法可以有效地减少训练曲线的振荡, 对提高预测模型的精确度有较大的帮助。对于降水情况, 由于数据中缺少具体的降水量值, 所以将降水直接赋值 0 (无降雨) 和 1 (有降雨, 包括降雪); 经相关度分析后发现 SO_2 浓度与当天的降水情况的相关度并不高, 但与前两天的降水有较高的相关度; 经计算比较, 当取预测日前两天的降水情况的均值时, 可以得到最大的相关度, 故取降水的输入项为前两天降水情况的均值。大气稳定度采用 P-T 方法判定, 该方法使用云量、风速、太阳高度角计算得到大气稳定度数据。混合层高度采用国标法计算, 在大气稳定度为 A、B、C、D 时, 混合层高 $h = a_s u_{10}/f$; 在大气稳定度为 E、F 类时, $h = b_s u_{10}/f$, 其中, u_{10} 为 10 m 处的平均风速, 直接取地面风速值; a_s 、 b_s 为边界层系数; f 为地转参数, $f = 2\Omega \sin\psi$ Ω 为地转角速度, 取为 $7.8 \times 10^{-5} \text{ rad/s}$, ψ 为地理纬度。此外, 由于所使用的数据中既包括气象数据又包括污染数据, 这就涉及到数据间的量级、量纲会有一定差异, 所以还需对获得的数据进行归一化处理, 使数据处于 0~1 的区间范围。

3.3 SO_2 浓度预报的 BP 模型的建立

在分析了各气象因子及污染因子与 SO_2 浓度的相关度之后, 去除相关度过低的影响因子, 最终将输入因子确定为如下 12 个: 预报目标日当天的湿度、最大温度、最大温差、主导风力、主导风向、混合层高度、大气稳定度、预报目标日前一天的湿度、最高温度、最大温差、 SO_2 浓度和降雨情况。可以看到, 所建立的模型既考虑了预报日当天的气象预报数据, 又考虑了预报日前一天的气象数据, 这样建立的模型具有合理性, 可以增加预报的准确度。此外, 苏州的气象数据和污染数据在 1-3 月、4-6 月、7-9 月、10-12 月各季度内的数据变化具有相对一致的特点, 因此采用分 4 个季度分别建立 BP 网络预报模型, 以捕捉不同季节气候污染物相应的变化特征。

为了比较不同隐层节点数对网络性能的影响, 取每一种网络节点数对应的网络结构做多次运算, 比较最终的训练效果后得出: 在网络节点为 30~35 时, 网络结构本身已经足够储存预测所需要的信息。

将 Variable Learning Rate Backpropagation, Resilient Backpropagation, Fletcher-Powell Conjugate Gra-

dient, Conjugate Gradient with Powell/Beale Restarts, Polak-Ribière Conjugate Gradient (CGP), Scaled Conjugate Gradient, One-Step Secant, Levenberg-Marquardt, BFGS Quasi-Newton 等 9 种典型算法对训练结果与真实值的最佳相关度、最佳相关时训练截止时的相对误差、训练时间这些参数进行比较后发现, 对于该项目, CGP 算法在大多数情况下可以实现较好的误差收敛。

建立模型时, 首先在 MATLAB 中获得最优网络数据。为了方便使用, 之后在 VC 环境下开发了可视化的输入界面, 如图 2 所示, 用面向对象的方法采用 C++ 语言重新实现了网络的模型; 使用了 C 语言和 MATLAB 的接口库, 将在 MATLAB 中获得的最优网络数据读取到新实现的网络模型中, 再结合从接口界面得到的数据进行预测。

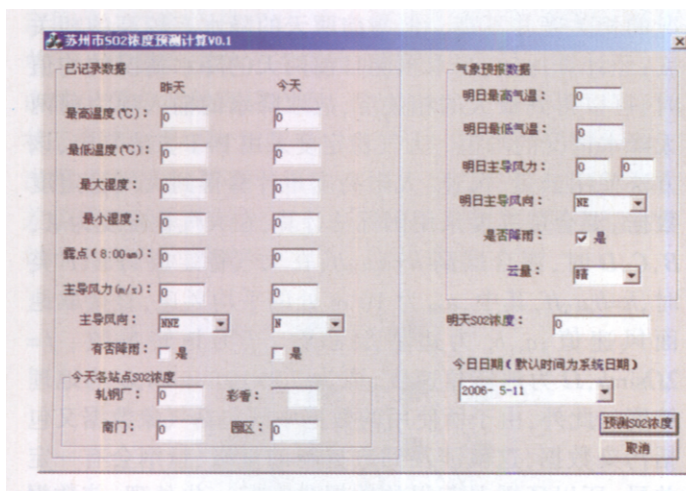


图 2 VC 可视化输入界面
Fig. 2 VC inputting interface

3.4 预报结果

将苏州 2004 年数据分成 4 组, 每组的 90 个数据中, 筛选其中 70 个作为训练数据、10 个作为验证数据、10 个作为测试数据。经过训练选优之后, 网络实现了较好的预测效果。对于未参加训练的测试数据集, 平均相对误差如下: 第 1 季度为 13.97%、第 2 季度为 11.54%、第 3 季度为 14.43%、第 4 季度为 10.64%, 各季度的预测效果如图 3~图 6 所示。此外, 由全年比较可以得出, 苏州 SO_2 的预测浓度与实测浓度的相关度为 0.819 5, 全年平均误差为 13.52%。

4 结论

1) 用 MATLAB 编写了基于 BP 网络的 SO_2 浓度预报程序, 计算结果表明, 预测值与实测值吻合较好, 所建立的模型具有较好的实用性。

2) 所建立的 BP 网络虽以 SO_2 为研究对象, 但由于大气污染物扩散和计算原理具有一般性, 因此, 该方法可推广应用于 NO_2 , PM_{10} 等大气污染物的预报研究。

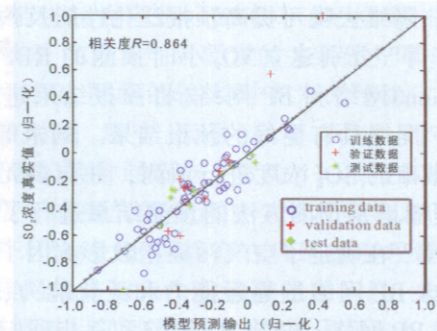


图 3 第 1 季度的预测效果
Fig. 3 Forecast for first quarter

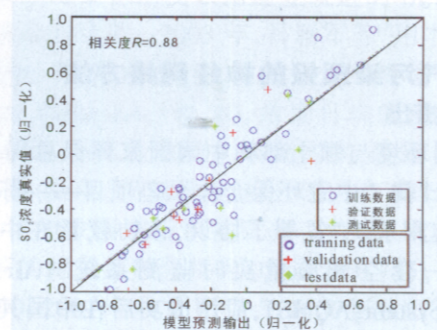


图 4 第 2 季度的预测效果
Fig. 4 Forecast for second quarter

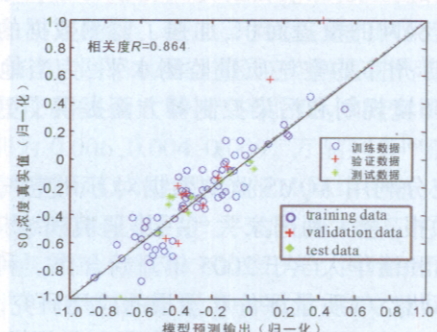


图 5 第 3 季度的预测效果
Fig. 5 Forecast for third quarter

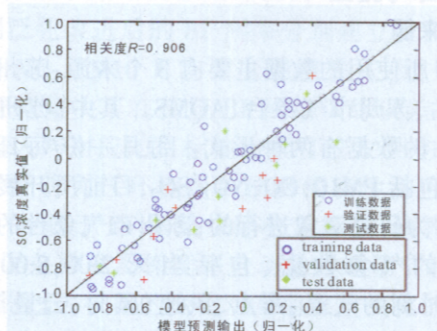


图 6 第 4 季度的预测效果
Fig. 6 Forecast for fourth quarter

3) 相对于空气污染数值预报而言, 神经网络应用于空气污染预报中, 计算简单, 所需数据要求相对较低, 尤其适用于无法开展数值预报的城市和地区。

4) 目前所进行的预报试验, 数据的收集、处理以及数据与程序的接口处理等一系列问题都是单独分开进行的, 因此, 降低了工作效率。今后应将整个预报系统中的各个部分动态联系起来, 实现预报的自动化和业务化。

参考文献 (References)

- [1] RAGA G B, MOYNE L L E. On the nature of air pollution dynamics in Mexico City - I. Nonlinear analysis[J]. Atmos Environ, 1996, 30(23): 3987- 3993.
- [2] 李祚泳, 邓新民. 环境污染预测的人工神经网络模型[J]. 成都气象学院学报, 1997, 4(12): 279- 283.
LI Zuoyong, DENG Xinmin. Prediction model of environmental pollution using artificial neural network[J]. Journal of Chengdu Institute of Meteorology, 1997, 4(12): 279- 283.
- [3] MOK K M, TAM S C. Short - term prediction of SO₂ concentration in Macau with artificial neural networks[J]. Energy and Buildings, 1998(28): 279- 286.
- [4] 赵慧宏. 兰州市 SO₂ 排放总量预测的 BP 人工神经网络模型[J]. 甘肃环境研究与监测, 2001, 14(3): 157- 159.
ZHAO Huihong. BP artificial neural networks for the prediction of Sulfur Dioxide total emission amounts in Lanzhou [J]. Gansu Environmental Study and Monitoriny, 2001, 14(3): 157- 159.
- [5] 王 俭, 胡筱敏, 郑龙熙, 等. 基于 BP 模型的大气污染预报方法的研究[J]. 环境科学研究, 2002, 15(5): 62- 64.
WANG Jian, HU Xiaomin, ZHENG Longxi, et al. Study on forecasting of air pollution based on BP model [J]. Research of Environmental Sciences, 2002, 15(5): 62- 64.
- [6] 王 俭, 胡筱敏, 郑龙熙, 等. BP 模型的改进及其在大气污染预报中的应用 [J]. 城市环境与城市生态, 2002, 15 (5): 17- 19.
WANG Jian, HU Xiaomin, ZHENG Longxi, et al. Improvement and application of BP model in prediction of atmospheric pollution [J]. Urban Environment & Urban Ecology, 2002, 15(5): 17- 19.
- [7] 马雁军, 杨洪斌, 张云海. BP 神经网络法在大气污染预报中的应用研究[J]. 气象, 2003(7): 49- 52.
MA Yanjun, YANG Hongbin, ZHANG Yunhai. Study on prediction of atmosphere pollution concentration based on BP model[J]. Meteorological Monthly, 2003(7): 49- 52.
- [8] 王灿星, 祁国伟, 何 曦, 等. BP 神经网络用于大气中颗粒物 (TSP) 预测的研究[J]. 仪器仪表学报, 2003, 24(4 增刊): 529- 530.
WANG Canxing, QI Guowei, HE Xi, et al. Research on application of BP neural network to predict TSP pollution [J]. Chinese Journal of Scientific Instrument, 2003, 24 (suppl): 529- 530.
- [9] 杨树平, 冯 晖, 黄 晓. B- P 神经网络在环境空气质量预报中的应用[J]. 云南环境科学, 2003(22): 52- 54.
YANG Shuping, FENG Hui, HUANG Xiao. Application of B- P nerve network on environment air quality forecast[J]. Yunnan Environmental Science, 2003(22): 52- 54.
- [10] 万显列, 杨凤林, 王慧卿. 利用人工神经网络对空气中 O₃ 浓度进行预测[J]. 中国环境科学, 2003, 23(1): 110- 112.
WAN Xianlie, YANG Fenglin, WANG Huiqin. The approach of artificial neural network applied in ambient ozone forecast [J]. China Environmental Science, 2003, 23(1): 110- 112.
- [11] JIANG Dahe, ZHANG Yang, HU Xiang, et al. Progress in developing an ANN model for air pollution index forecast [J]. Atmospheric Environment, 2004(38): 7055- 7064.
- [12] 黄世芹. 改进 BP 神经网络在城市环境大气污染分季节预报中的应用[J]. 贵州气象, 2005, 29(3): 6- 8.
HUANG Shiqin. Application of improved BP neural network on seasonal prediction of urban air pollution[J]. Journal of Guizhou Meteorology, 2005, 29(3): 6- 8.
- [13] 吴小红, 康海燕, 任德官. 基于神经网络中小城市空气污染指数数预估值的设计 [J]. 数学的实践与认识, 2005, 35(2): 87- 91.
WU Xiaohong, KANG Haiyan, REN Deguan. Forecast of air pollution index in one city based on the artificial neural net[J]. Mathematics in Practice and Theory, 2005, 35(2): 87- 91.
- [14] 黄海洪, 孙崇智. SO₂ 浓度主成分神经网络预报方法初探[J]. 上海环境科学, 2003(增刊): 171- 174.
HUANG Haihong, SUN Chongzhi. Preliminary approach on sulfur dioxide concentration prediction by principal component neural network method [J]. Shanghai Environmental Science, 2003(suppl): 171- 174.
- [15] 李少华, 董向元, 于静梅. 基于神经网络的大气污染预报方法的研究[J]. 东北电力学院学报, 2003, 23(4): 2- 4.
LI Shaohua, DONG Xiangyuan, YU Jingmei. Study on prediction of air pollution based on artificial neural network [J]. Journal of Northeast China Institute of Electric Power Engineering, 2003, 23(4): 2- 4.
- [16] 刘 永, 郭怀成. 城市大气污染物浓度预测方法研究[J]. 安全与环境学报, 2004, 4(4): 60- 62.
LIU Yong, GUO Huaicheng. Approach to more efficient forecasting of concentration of air pollutants in urban areas [J]. Journal of Safety and Environment, 2004, 4(4): 60- 62.
- [17] 周秀杰, 苏小红, 袁美英. 基于 BP 网络的空气污染指数预报研究[J]. 哈尔滨工业大学学报, 2004, 36(5): 582- 585.
ZHOU Xiujie, SU Xiaohong, YUAN Meiyong. Forecast of air pollution index based on BP neural network[J]. Journal of Harbin Institute of Technology, 2004, 36(5): 582- 585.

(责任编辑 王 芷)