

文章编号: 1002-2082 (2020) 04-0761-08

基于 PCA 和 BP 神经网络的硝酸盐氮浓度检测方法

陈 朋¹, 严宪泽², 韩洋洋², 吴晨阳², 咎 昊²

(1. 浙江工业大学 计算机科学与技术学院, 浙江 杭州 310023; 2. 浙江工业大学 信息工程学院, 浙江 杭州 310023)

摘 要: 针对紫外分光光度法 (UV 法) 检测混有干扰物质的硝酸盐氮溶液浓度精度不高的问题, 提出一种基于主成分分析 (principal component analysis, PCA) 和 BP 神经网络的硝酸盐氮浓度检测方法。通过微型光谱仪物质成分检测系统测得硝酸盐氮试剂在 196 nm~631 nm 波段的吸光度数据, 分为测试集和训练集。通过 PCA 计算训练集, 得到主成分。根据 BP 算法搭建三层人工神经网络。将所得主成分除以 8 后输入网络展开训练。训练过程中采用留一法交叉验证。用该模型计算训练集和测试集, 所得值与真实浓度的平均相对误差分别为 2.411 5% 和 1.553%。实验结果表明, 该方法能较好检测出混有干扰物质的硝酸盐氮溶液浓度。

关键词: 光谱分析; 浓度; 神经网络; 硝酸盐氮; 主成分分析

中图分类号: TN65; O657.32

文献标志码: A

DOI: 10.5768/JAO202041.0410002

Nitrate nitrogen concentration detection method based on principal component analysis and BP neural network

CHEN Peng¹, YAN Xianze², HAN Yangyang², WU Chenyang², ZAN Hao²

(1. College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou 310023, China;

2. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: Aiming at the problem of inaccurate detection of the nitrate nitrogen solution concentration with interfering substances in ultraviolet spectrophotometry (UV method), a nitrate nitrogen concentration detection method based on principal component analysis (PCA) and BP neural network was proposed. First, the absorbance of the nitrate nitrogen reagent at 196 nm~631 nm was measured by the material composition detection system of the micro-spectrometer, which was divided into test set and training set. Then, the PCA was used to calculate the training set to obtain the principal components. Finally, a three-layer artificial neural network was built based on the BP algorithm. The obtained principal components were divided by 8 and input into the network for training. During the training, the leaving-one method was adopted for cross-validation. This model was used to calculate the training set and test set, the mean relative error between the obtained results and the true concentration is 2.411 5% and 1.553% respectively. The experimental results show that the method can better detect the concentration of the nitrate nitrogen reagent with interfering substances.

Key words: spectral analysis; concentration; neural network; nitrate nitrogen; principal component analysis

引言

随着工农业的快速发展, 水污染问题越发严

重。长期饮用硝酸盐含量过高的水体容易引发人体各种疾病。作为含氮有机物的最终产物, 世界

收稿日期: 2019-11-01; 修回日期: 2020-03-12

基金项目: 浙江省属高校基本科研业务费专项资金资助 (RF-C2019001); 浙江省教育厅一般科研项目 (GZ18571030014); 浙江省重点研发计划 (2019C01007)

作者简介: 陈朋 (1981-), 男, 博士, 教授, 主要从事嵌入式系统设计、图像处理和模式识别研究。E-mail: chenpeng@zjut.edu.cn

卫生组织规定饮用水中硝酸盐氮浓度不能超过 10 mg/L^[1]。传统 UV 法^[2]检测硝酸盐氮的原理是利用硝酸根离子、水中有机物在 220 nm 波长处有吸收,而硝酸根离子在 275 nm 处没有吸收,水中有机物有吸收的特点,建立 $A_{220}-2A_{270}$ 校正标定模型。但 UV 法在检测硝酸盐氮溶液实际应用中,如果待测试剂浓度过高,存在颗粒物质、荧光或磷光等干扰,吸光度与物质浓度之间不再遵守线性关系,导致 UV 法失效^[3]。

为了能够精确测量水体中硝酸盐氮的浓度,目前国内外已经有多项研究。杨鹏程^[4]等人将偏最小二乘(partial least squares, PLS)应用于海水中硝酸盐氮的紫外光谱分析,可以有效分离重叠光谱,降低干扰物质影响,但是作为线性建模方法,PLS 具有局限性。刘思乡等人^[5]提出机器学习在 UV 法测定硝酸盐氮浓度中的应用,利用局部加权线性回归(locally weighted linear regression, LWLR)结合 PCA 建立数学模型,但 LWLR 存在模型复杂度低和模型表达能力不强的问题。俞禄等人^[6]针对光谱水质分析的软测量问题,提出了在紫外光谱条件下的主成分回归(principal components regression, PCR)和最小二乘支持向量机(least square support vector machine, LSSVM)建模方法。Jean Causse 等人^[7]提出了一种基于 UV 法测量原始硝酸盐氮溶液的方法,该方法将双光程耦合用于光谱采集,同时对信号二阶求导,但其不适合检测有色度、浊度干扰的溶液。

综上所述,本文针对传统光谱分析方法检测带有干扰物质的硝酸盐氮浓度精度不高的问题,提出了基于 PCA 和 BP 神经网络的硝酸盐氮浓度检测方法。

1 检测方法总体设计

本文提出的基于 PCA-BP 神经网络的硝酸盐氮浓度检测方法流程如图 1 所示,主要由采集吸光度数据、PCA 降维、BP 神经网络建模^[8]、数据保存 4 部分组成。首先,由微型光谱仪物质成分检测系统测量混有干扰物质的不同浓度硝酸盐氮试剂的吸光度数据;然后,将吸光度数据经 PCA 计算后得到数据的主成分;接着,在 TensorFlow 深度学习框架上搭建神经网络模型,建模过程采用留一法交叉验证;最后,保存训练结果,检验模型精度。

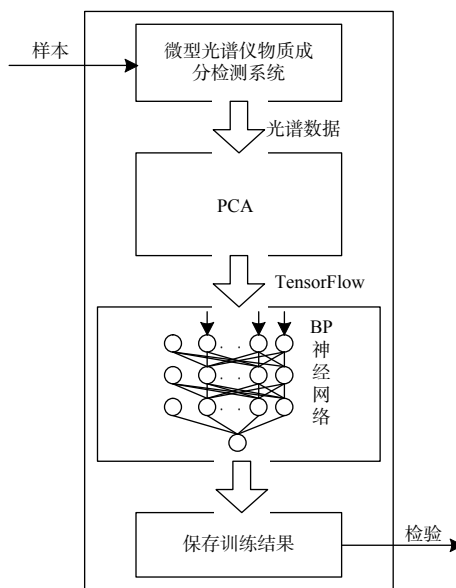


图 1 检测方法流程图

Fig. 1 Flow chart of detection method

2 基于 PCA 对光谱数据降维设计

PCA 作为多元统计中的数据压缩技术,广泛应用于化学实验数据的统计分析。PCA 在尽可能保留原始信息的基础上,将高维空间的样本映射到低维主成分空间^[9]。在分析不同浓度硝酸盐氮试剂的光谱数据时,如果将所有光谱数据同时输入 BP 神经网络,过多的迭代和计算可能会导致所建立的模型过拟合,从而导致精度下降。同时原始数据中也可能存在噪声干扰和冗余信息,所以将数据输入神经网络之前需要对其进行主成分分析处理^[10]。

设研究对象总共有 n 个样本,每个样本有 p 个数据。故样本原始数据可表示为

$$X = [X_1 \quad X_2 \quad \cdots \quad X_n] = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (1)$$

进一步计算均值,得:

$$\bar{x} = \frac{\sum_{j=1}^p x_{ij}}{p}, i = 1, 2, \cdots, n \quad (2)$$

样本原始数据均值化处理:

$$y_{ij} = x_{ij} - \bar{x}_i, i = 1, 2, \cdots, n; j = 1, 2, \cdots, p \quad (3)$$

均值化后的数据矩阵为 $Y = [Y_1, Y_2, \cdots, Y_n]$ 。

根据协方差公式:

$$\text{Cov}_{ij} = \frac{\sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)}{(p-1)}, j = 1, 2 \cdots p \quad (4)$$

建立协方差矩阵 C :

$$C = \begin{bmatrix} \text{Cov}_{11} & \text{Cov}_{12} & \cdots & \text{Cov}_{1p} \\ \text{Cov}_{21} & \text{Cov}_{22} & \cdots & \text{Cov}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}_{n1} & \text{Cov}_{n2} & \cdots & \text{Cov}_{np} \end{bmatrix} \quad (5)$$

依据特征分解法确定 C 的特征值 $e_1, e_2 \cdots e_p$ 和特征向量 $[u_1, u_2 \cdots u_p]$, 其中 $u_i = [u_{i1}, u_{i2} \cdots u_{ip}]^T$ 。

根据公式(6)确定每个主成分的累计贡献率:

$$\varphi(i) = \frac{\sum_{i=1}^i e_i}{\sum_{i=1}^p e_i} \quad (6)$$

对主成分累计贡献率从大到小进行排序, 依此选择主成分, 求第 k 个主成分 F_k :

$$F_k = a_k \times Y^T \quad (7)$$

3 基于 BP 神经网络搭建物质浓度标定模型

3.1 BP 神经网络介绍

BP 神经网络是采用误差反向传播算法并具有非线性连续转移函数的多层前馈网络, 由输入层、若干隐藏层和输出层组成^[1]。多层前馈网络利用全连接的方式连接前后相邻两层之间的神经元, 其中每一层有多个神经元, 但各神经元之间没有连接, 隐藏层和输出层中的神经元可以从前一层接收多个输入, 但只输出一个结果作为下一层神经元的输入。

3.2 BP 神经网络训练过程

本实验 BP 神经网络训练过程如下。

1) 初始化 BP 神经网络参数。初始化内容包括设置学习速率、激活函数、损失函数、初始值、偏置、最大迭代步数。

①设置 BP 神经网络学习速率为 0.05。

②针对输入层和隐藏层选择 Sigmoid 激活函数, 公式如下:

$$f_2(z) = \frac{1}{1 + e^{-z}} \quad (8)$$

为避免 relu 激活函数引起的梯度消失问题, 输出层使用 leak relu 激活函数, 公式如下:

$$f_3(z) = \begin{cases} \max(z, 0), & z > 0 \\ 0.2z, & z < 0 \end{cases} \quad (9)$$

式中: z 为神经元的输入。

③以均方根误差(RMSE)作为损失函数, 公式如下:

$$\text{RMSE}(y, y') = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \quad (10)$$

式中: y_i 为第 i 个样本的期望值; y'_i 为神经网络得出的预测值。

④本文中初始权值由计算机从满足正态分布的数值中随机抽取, 初始偏置全部设为 0。

⑤设置初始最大迭代步数为 1 800。

2) 建模过程采用留一法交叉验证, 每隔 k 次输出平均误差 e , 如公式(11)所示, 判断 e 是否符合要求或者是否达到预先设定的学习次数。如果是, 则结束本轮训练。

$$e = \frac{1}{k} \sum_{i=1}^k (y_i - y'_i) \quad (11)$$

3) 判断 n 个样本中每个样本是否都已被作为验证使用, 如果是, 则训练过程结束; 否则继续步骤 2)。

4 实验分析与结果

4.1 实验仪器和试剂配制

本文选用美国 Pyxis SQ-3000 型紫外波段微型光谱仪, 该款微型光谱仪可检测的波长范围为 196 nm~631 nm, 波长分辨率为 1.5 nm, 光源选用 Pyxis XE-10W 型氙灯光源。如图 2 所示, 整套检测系统主要由电源、氙灯光源、石英光纤、支架系统(含比色皿)、微型光谱仪和 PC 机组成, 其中微型光谱仪通过串口实现对 PC 机和氙灯光源的控制。

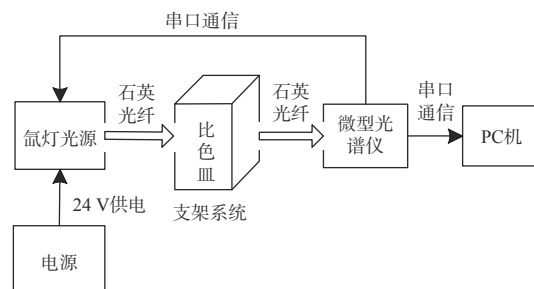


图2 微型光谱仪测量系统

Fig. 2 Micro-spectrometer measurement system

硝酸盐氮试剂的配置: 取 1 000 mg/L 标准硝酸盐氮溶液, 有色溶液(Fe^{3+} , 1 000 mg/L), 浊液 F(硫酸肼和六次甲基四胺, 1 000 NTU), 按照表 1 所示配置 68 份样品试剂。

表 1 样品配置方案

Table 1 Sample configuration scheme

mg·L⁻¹

编号	硝酸盐氮	有色溶液Fe ³⁺								浊液F								
		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	9
1	2	0.5	1	1.5	2	2.5	3	4	5	2.5	5	7.5	10	15	20	30	40	50
2	5	0.5	1	1.5	2	2.5	3	4	5	2.5	5	7.5	10	15	20	30	40	50
3	10	0.5	1	1.5	2	2.5	3	4	5	2.5	5	7.5	10	15	20	30	40	50
4	20	0.5	1	1.5	2	2.5	3	4	5	2.5	5	7.5	10	15	20	30	40	50

由表 1 中选取 S_(1, 6_Fe3+) (试剂中硝酸盐氮浓度为 2 mg/L, 有色溶液浓度为 3 mg/L, 以下同理)、S_(1, 7_F)、S_(2, 1_Fe3+)、S_(2, 2_F)、S_(3, 2_Fe3+)、S_(3, 6_F)、S_(4, 5_Fe3+) 和 S_(4, 4_F) 的共 8 组样品作为测试样本(测试集), 剩余的 60 份样品作为训练样本(训练集)。

4.2 数据采集

根据图 2 连接设备, 首先测量微型光谱仪空载时的光谱数据 $B(n)$, 然后将蒸馏水放入比色皿, 测量其光谱数据 $F(n)$, 接着以同样的方式测量样品的光谱数据 $S(n)$, 最后根据公式(12)测得样品吸光度。每个样品分别测试 3 次, 取其平均值并保存, 部分吸光度曲线如图 3 所示。

$$A(n) = \lg \left[\frac{F(n) - B(n)}{S(n) - B(n)} \right]$$

(12)

式中: $A(n)$ 为样品吸光度。

由图 3 分析可知: 样品中硝酸盐氮浓度, 色度和浊度都是影响吸光度的重要因素, 本文的重点

就是寻找在色度和浊度的干扰下硝酸盐氮浓度与吸光度之间的关系。

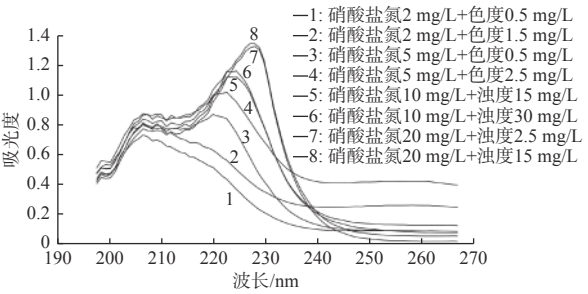


图 3 硝酸盐氮试剂的吸光度曲线

Fig. 3 Absorbance curve of nitrate nitrogen

4.3 建立物质成分回归模型

对训练集的吸光度数据通过 PCA 计算。首先将吸光度数据按照公式(1)表示为矩阵; 然后, 根据公式(2)计算每个样本的平均值; 其次, 根据公式(3)对每维数据均值化处理; 接着, 根据公式(4)、(5)建立协方差矩阵 C , 计算特征值和特征向量, 部分数据如表 2 所示。

表 2 PCA 中间计算结果

Table 2 PCA intermediate calculation results

维度	1	2	3	4	5	6	7	8	9
平均值 R	0.061 8	0.103 1	0.077 0	0.246 7	0.192 8	0.223 9	0.160 4	0.049 8	0.053 2
中间值 Y_{ij}	0.315 6	0.356 2	0.406 5	0.390 3	0.404 9	0.443 8	0.489 7	0.551 5	0.590 9
协方差 C_{ij}	0.009 2	0.011 2	0.012 7	0.012 4	0.012 7	0.014 1	0.016 1	0.018 1	0.019 9
特征值 E	2.783 7	0.257 7	0.089 9	0.006 5	0.001 1	6.284 6e-04	1.227 5e-04	4.776 3e-05	2.042 4e-05
特征向量 U_{ij}	-0.037 1	-0.006 4	-0.062 8	0.042 1	0.009 6	-0.002 7	-0.042 7	-0.044 5	0.072 2

表 2 中列举了前 9 个样本的平均值 R ; Y_{ij} 由第一个训练样本减去平均值后得到; 表中 C_{ij} 为协方差矩阵第一行第 j 列元素; 特征值 E , 表中列举了前 9 维数据; 特征向量 U_{ij} 表示第一行第 j 列元素, j 表示维度。

根据公式(6)计算输出的主成分贡献率和累积贡献率如表 3 所示。

由表 3 可知, 在输出数据的第一个维度上, 即第一个主成分的贡献率占有所有维度上贡献率之和的 89.253%, 前 9 个主成分的累积贡献率之和占全体的 99.989 9%。为了寻找模型输入主成分数量的最优值, 按照 1~9 的顺序选取数量不等的主成分输入模型, 最后, 测试集得到平均误差如表 4 所示。

表 3 主成分贡献率分布表

Table 3 Distribution of principal component contribution rate

主成分	主成分贡献率/%	累积贡献率/%
1	89.253	89.253
2	7.608	96.861
3	2.784	99.645
4	0.238	99.883
5	0.050 95	99.934
6	0.029 837	99.963 8
7	0.013 621	99.977 4
8	0.008 753 7	99.986 2
9	0.003 751 4	99.989 9

由表 4 可知, 当模型输入 8 个主成分时预测性能最好, 预测值的平均误差为 1.553%, 因此实验最终选择 8 个主成分作为神经网络的输入。同时为了满足激活函数的输入在 0~1 范围内的要求, 需将公式(7)所求主成分除以 8, 部分输入如表 5 所示。

表 5 中每个训练样本经 PCA 后降为 8 维, 对应的标签是以 40 为底浓度的对数。

本实验建模过程采用留一法交叉验证, 即在训练样本中抽取 59 个样本用于学习, 余下的 1 个样本用于验证。这一步骤持续到每个样本都被当作一次验证资料为止, 得到关于 RMSE 收敛曲线。如图 4 所示, 依据收敛曲线在后续的训练过程中可将前述最大迭代步数改为 100。

表 4 模型不同输入下测试集的平均误差

Table 4 Mean error of test set under different inputs of model

模型输入主成分数量	1	2	3	4	5	6	7	8	9
平均误差/%	89.723	9.055	4.056	4.038	4.054	4.097	2.674	1.553	1.652

表 5 神经网络的输入

Table 5 Neural network input

序号	训练数据								标签/mg·L ⁻¹
样本1	1.545 5e-03	-3.772 4e-04	-2.676 4e-03	-1.480 5e-02	3.521 9e-02	-0.150 6	3.001 1e-02	-0.450 6	0.187 9
样本2	2.582 2e-03	-1.026 4e-03	1.375 5e-03	-8.703 3e-03	-2.523 8e-02	-9.614 6e-02	-6.800 7e-02	-0.528 0	0.436 3
样本3	0.528 3e-03	2.248 4e-03	-1.033 2e-02	-1.144 9e-02	-3.446 8e-02	-1.005 6e-02	-0.146 4	-0.578 3	0.624 2
样本4	1.124 2e-03	-1.038 6e-04	4.109 1e-03	-1.160 3e-02	2.101 0e-02	0.133 2	-0.151 5	-0.671 9	0.812 1

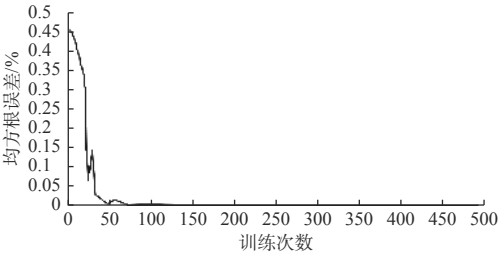


图 4 RMSE 收敛曲线

Fig. 4 RMSE convergence curve

最终训练所得模型结果如公式(13)所示。

$$y = 40[(XW^{(1)}+b^{(1)})W^{(2)}+b^{(2)}](W^{(3)})^T+b^{(3)} \tag{13}$$

式中: $W^{(1)}$ 、 $W^{(2)}$ 和 $W^{(3)}$ 分别表示第 1 个隐藏层、第 2 个隐藏层和输出层的连接权值; $b^{(1)}$ 、 $b^{(2)}$ 和 $b^{(3)}$ 分别表示输入层, 第 1 隐藏层, 第 2 隐藏层的偏置; X 为模型输入(测试样本的主成分除以 8); y 表示硝酸盐氮的预测浓度, 部分参数如表 6 所示。

表 6 模型部分参数

Table 6 Model partial parameters

维度	1	2	3	4	5	6	7	8
权值 $W_{ij}^{(1)}$	7.992 026 3	7.980 647	6.703 853 1	8.044 995 3	-7.389 650 8	8.452 992	7.845 027	6.633 517 3
权值 $W_{ij}^{(2)}$	-3.336 365 9	-0.233 432	0.191 100 3	-2.469 691 3	-2.978 089 3	-2.251 955	-1.279 54	-3.285 006 0
权值 $W^{(3)}$	-0.538 127 3	-1.048 994	0.191 781 9	-1.011 383 3	-1.525 026 1	-1.060 538	-2.099 07	-0.967 715
偏置 $b_j^{(1)}$	-2.809 857 4	-3.030 030	-3.057 474	-2.074 893 7	-0.707 169 1	0.082 838	-2.467 16	-3.537 668
偏置 $b_j^{(2)}$	-2.952 146 1	-0.106 216	-0.903 022	-1.182 433 4	-2.085 746 8	-1.792 576	-0.927 11	-2.404 344
偏置 $b^{(3)}$	0.801 382 4	\	\	\	\	\	\	\

表 6 中 $W_{ij}^{(1)}$ 表示矩阵第 1 行, 第 j 列权值 ($W_{ij}^{(2)}$ 同理), $W^{(3)}$ 为输出层权值, 共 8 个成员。由于本模型的目的是为实现预测, 因此输出层只有一个神经元, 故 $b^{(3)}$ 是一维变量。

4.4 实验结果

检测系统获取 8 个测试样本的吸光度数据后, 分别与主成分系数 P 运算得到测试样本的主成分。

$$PC_i = P \times X_i, i = 1, 2, \dots, 8 \tag{14}$$

式中: P 为前述训练样本 8 个主成分对应的特征向量。

将测试样本的主成分 $PC_i/8$ 作为公式(13)的输入, 求得预测值。同时引入 PCR^[12]、PLS^[13] 和 LSSVM^[14] 算法作对比实验, 然后分别用相同的训练集建立标定模型, 最后分别对 8 个测试样本进行预测, 得到如表 7 所示预测结果。

表 7 测试集硝酸盐氮浓度预测结果

Table 7 Prediction results of test set nitrate nitrogen concentration

建模方法	测试样本	$S_{(1,6_Fe3+)}$	$S_{(1,7_F)}$	$S_{(2,1_Fe3+)}$	$S_{(2,2_F)}$	$S_{(3,2_Fe3+)}$	$S_{(3,6_F)}$	$S_{(4,5_Fe3+)}$	$S_{(4,4_F)}$
	NO3-N浓度/mg·L ⁻¹	2	2	5	5	10	10	20	20
PCR	预测值/mg·L ⁻¹	2.130	1.834	5.101	4.992	10.467	10.231	19.429	20.366
PLS	预测值/mg·L ⁻¹	1.972	2.047	5.110	5.033	10.197	10.671	19.611	20.000
LSSVM	预测值/mg·L ⁻¹	2.090	2.657	5.053	4.928	10.706	13.290	19.843	19.674
PCA-BP	预测值/mg·L ⁻¹	1.972	1.995	5.011	5.072	10.462	10.172	19.875	20.429

为了验证 PCA-BP 神经网络的优越性, 引入平均相对误差对表 7 作进一步分析, 分析结果如表 8 所示。

经表 8 分析得出, 利用 PCA-BP 神经网络建立的硝酸盐氮物质标定模型, 相对于 PCR、PLS 和 LSSVM 平均相对误差分别减小了 56.64%、27.94% 和 84.90%, 由此可知基于 PCA-BP 神经网络的标定模型预测性能最好。

由于本文测试样本较少, 仅通过测试样本评价 4 种模型性能好坏可能存在偶然性, 即因个别测

试样本所得预测值的精度可能很高, 而使得平均相对误差很小, 因此还需考察模型对训练集样本的误差情况。现将所有训练集样本分别输入训练好的硝酸盐氮物质标定模型中, 计算结果如图 5 所示。

表 8 不同建模方法测量结果对比

Table 8 Comparison of measurement results with different modelling methods

算法	PCR	PLS	LSSVM	PCA-BP神经网络
平均相对误差/%	3.581	2.154	10.278	1.553

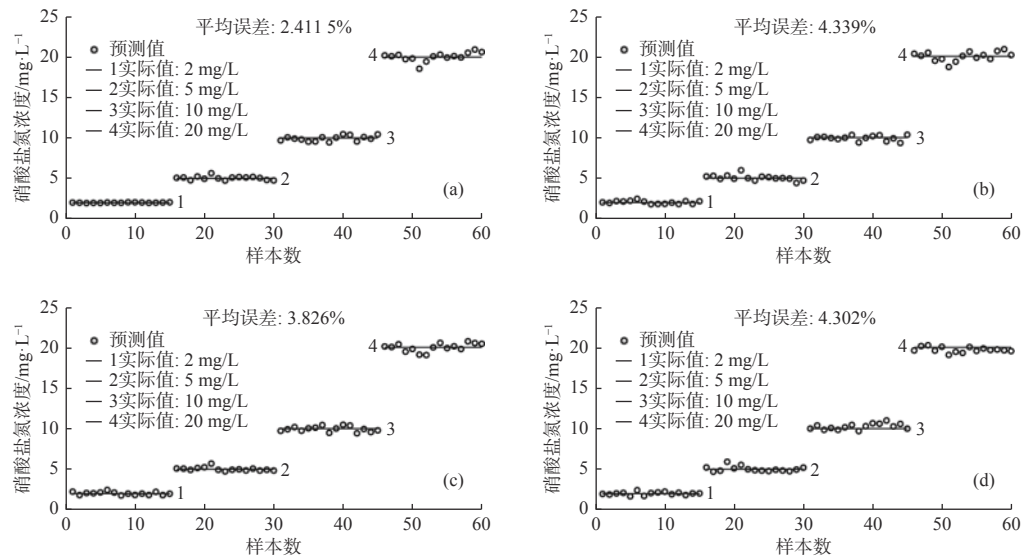


图 5 训练样本结果对比

Fig. 5 Comparison of training sample results

图 5 中(a)为 PCA-BP 神经网络模型的测试结果; (b)为 PCR 模型的测试结果; (c)为 PLS 模型的训练结果; (d)为 LSSVM 模型的测试结果。针对训练集数据, 4 种模型计算结果的平均相对误差分别为 2.4115%、4.339%、3.826% 和 4.302%, 由此可知基于 PCA-BP 神经网络的标定模型对训练集样本计算能力同样最优。

综上所述, 本文设计的 PCA-BP 神经网络在 4 种模型中性能表现最好, 能有效排除冗余信息影响和噪声干扰, 具有较好的非线性映射能力、泛化能力和容错能力^[15], 可以精确预测混有干扰物质的硝酸盐氮浓度的浓度。

5 结论

本文提出的基于 PCA 和 BP 神经网络的硝酸盐氮溶液浓度检测方法在检测混有干扰物质的硝酸盐氮溶液时效果较好, 针对测试集的平均相对误差相较于 PCR、PLS 和 LSSVM 算法减小了 56.64%、27.94% 和 84.90%。训练集的平均相对误差相较于 PCR、PLS 和 LSSVM 算法减小了 44.42%、36.97% 和 43.94%, 证明了该方法的准确性、可行性, 同时该方法也适用于对水体中其它物质成分的建模分析, 在工程实践中具有重要的指导意义。

参考文献:

- [1] ZHANG Yali, ZHANG Yizhang, ZHANG Yuan, et al. Characteristics of nitrate in surface water and groundwater in the Hun River Basin[J]. China Environmental Science, 2014, 34(1): 170-177.
张亚丽, 张依章, 张远, 等. 浑河流域地表水和地下水氮污染特征研究[J]. 中国环境科学, 2014, 34(1): 170-177.
- [2] State Environment Protection Administration. SL84-1994 Water quality-determination of nitrate-nitrogen-ultraviolet spectrophotometry[S]. Beijing: Chinese Standard Publishing, 2007.
国家环境保护总局. SL84-1994 水质硝酸盐氮的测定紫外分光光度法[S]. 北京: 中国标准出版, 2007.
- [3] SONG Ge, SUN Bo, JIAO Jianying. Comparison between ultraviolet spectrophotometry and other methods in determination of soil Nitrate-N[J]. Acta Pedologica Sinica, 2007, 44(2): 288-293.
宋哥, 孙波, 教剑英. 测定土壤硝态氮的紫外分光光度法与其他方法的比较[J]. 土壤学报, 2007, 44(2): 288-293.
- [4] YANG Pengcheng. Research on determination of nitrate in seawater based on ultraviolet spectra combined with PLS method[D]. Tianjin: National Ocean Technology Center, 2013.
杨鹏程. 紫外吸收光谱结合偏最小二乘法海水硝酸盐测量技术研究[D]. 天津: 国家海洋技术中心, 2013.
- [5] LIU Sixiang, FAN Weihua, GUO Hui, et al. Application of machine learning in determination of nitrate nitrogen based on ultraviolet spectrophotometry[J]. Spectroscopy and Spectral Analysis, 2017, 37(4): 1179-1182.
刘思乡, 范卫华, 郭慧, 等. 机器学习在紫外法测定硝酸盐氮浓度中的应用[J]. 光谱学与光谱分析, 2017, 37(4): 1179-1182.
- [6] YU Lu, WANG Xuejie, MING Qian, et al. Application and comparison of several modeling methods in spectral based water quality analysis[C]. Yantai: Technical Committee on Control Theory, Chinese Association of Automation, 2011: 5227-5230.
俞禄, 王雪洁, 明倩, 等. 几种建模方法在光谱水质分析中的应用和比较[C]. 烟台: 中国自动化学会控制理论专业委员会B卷, 2011: 5227-5230.
- [7] CAUSSE Jean, THOMAS Olivier, JUNG Audevalérie, et al. Direct DOC and nitrate determination in water using dual pathlength and second derivative UV spectrophotometry[J]. Water Research, 2017, 108: 312-319.
- [8] WANG Lei, QIAO Xiaoyan, ZHANG Shu, et al. Pesticide residue detection by fluorescence spectral analysis based on BP neural network[J]. Journal of Applied Optics, 2010, 31(3): 442-446.
王雷, 乔晓艳, 张姝, 等. 基于BP神经网络的荧光光谱法农药残留检测[J]. 应用光学, 2010, 31(3): 442-446.
- [9] JIANG Chuan, ZHANG Xiaohong, ZOU Xiaolong, et al. Identification of salmonella and listeria monocytogenes by near-infrared spectroscopy and PCA[J]. Meteorological and Environmental Research, 2019, 10(4): 77-80.
- [10] LI Chan, WAN Xiaoxia, XIE Wei, et al. Color filter design method for multi-channel spectral acquisition system[J]. Journal of Applied Optics, 2016, 37(5): 639-643.
李婵, 万晓霞, 谢伟, 等. 多通道光谱采集系统滤光片设计方法研究[J]. 应用光学, 2016, 37(5): 639-643.
- [11] WANG Zhuqing, CHEN Yashao, QI Bin, et al. Prediction of daytime variations of HO₂ radical concentrations in the marine boundary layer using BP network[J]. Science China(Chemistry), 2010, 53(12): 2612-2621.
- [12] ZHANG Jingyang, PAN Guangyou. Comparison and ap-

- plication of multiple regression and BP neural network prediction model[J]. Journal of Kunming University of Science and Science and Technology:Natural Science Edition, 2013, 38(6): 61-67.
- 张景阳, 潘光友. 多元线性回归与BP神经网络预测模型对比运用研究[J]. 昆明理工大学学报:自然科学版, 2013, 38(6): 61-67.
- [13] RIBEIRO C L, DENZIN T G H, REIS M F, et al. Artificial neural network and partial least square regressions for rapid estimation of cellulose pulp dryness based on near infrared spectroscopic data[J]. *Carbohydrate polymers*, 2019, 224: 115186.
- [14] BAGHBAN A, SASANIPOUR J, HABIBZADEH S, et al. Estimating solubility of supercritical H₂S in ionic liquids through a hybrid LSSVM chemical structure model[J]. *Chinese Journal of Chemical Engineering*, 2019, 27(3): 620-627.
- [15] ZHOU Lingao. Application of improved BP neural network in pattern recognition[D]. Nanjing: Nanjing University of Science and Technology, 2010.
- 周凌翱. 改进BP神经网络在模式识别中的应用及研究[D]. 南京: 南京理工大学, 2010.