

# 基于改进型极限学习机的电子鼻气体浓度检测\*

王洁,陶洋,梁志芳

(重庆邮电大学 通信与信息工程学院,重庆 400065)

**摘要:** 针对目前电子鼻应用于气体污染物浓度检测时难以达到理想精度的问题,提出基于粒子群算法与人工蜂群算法的极限学习机(Particle Swarm Optimization and Artificial Bee Colony algorithm based Extreme Learning Machine, PSOABC-ELM)算法,通过改进极限学习机输入层与隐含层权值和隐含层阈值随机的缺陷,提高电子鼻浓度检测的精度。将 PSOABC-ELM 算法与其他算法进行比较,并在公开数据集上进行验证。实验结果表明,PSOABC-ELM 算法用于电子鼻气体浓度检测时比其他算法精准度更高,检测结果误差更小,模型稳定性更强,为电子鼻气体浓度检测提供了一种新的方法。

**关键词:** 电子鼻;粒子群算法;人工蜂群算法;极限学习机;浓度检测

中图分类号: TN02;TP212

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.211309

中文引用格式: 王洁,陶洋,梁志芳. 基于改进型极限学习机的电子鼻气体浓度检测[J]. 电子技术应用, 2021, 47(10): 63-67.

英文引用格式: Wang Jie, Tao Yang, Liang Zhifang. Gas concentration detection of E-nose based on improved ELM [J]. Application of Electronic Technique, 2021, 47(10): 63-67.

## Gas concentration detection of E-nose based on improved ELM

Wang Jie, Tao Yang, Liang Zhifang

(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**Abstract:** Aiming at unsatisfied ideal accuracy of electronic nose while testing the concentration of gas pollutants, the particle swarm optimization and artificial bee colony algorithm based extreme learning machine (PSOABC-ELM) algorithm is proposed. The accuracy of electronic nose concentration detection is enhanced by improve extreme learning machine weights of input layer and hidden layer and hidden layer threshold random defects. PSOABC-ELM is compared with other algorithms and validated on the public data set. The results show that the PSOABC-ELM algorithm perform better than the others when Testing gas concentration of electronic nose, and the detection result error is smaller and the algorithm stability is stronger, which provides a new method for the detection of gas concentration of electronic nose.

**Key words:** electronic nose; particle swarm optimization; artificial bee colony algorithm; extreme learning machine; concentration detection

### 0 引言

电子鼻是一种仿生嗅觉系统,由气体传感器阵列和模式识别算法组成,主要用于气体识别<sup>[1]</sup>,在环境监测<sup>[2]</sup>、食品检测<sup>[3]</sup>和医疗诊断<sup>[4]</sup>等多个领域均有所应用。电子鼻系统通过其内部的气体传感器阵列对气体信息进行采集,将气体信号转变为电信号,再通过模式识别算法的处理输出对应气体的浓度检测结果。

针对电子鼻模式识别系统,目前提出了多种网络模型,其中极限学习机是由黄广斌提出的一种典型单隐层前馈神经网络(Single-hidden Layer Feedforward Networks,

SLFN)<sup>[5]</sup>,与其他神经网络(BP神经网络<sup>[6]</sup>、支持向量机<sup>[7]</sup>(Support Vector Machine, SVM))相比,其结构简单,不需要反复迭代,学习速度快,泛化性能好,具有良好的函数逼近能力,因此被广泛应用于解决各种分类和回归的问题。但由于 ELM 输入层与隐含层的权值以及隐含层的阈值是随机给定的,这将会降低网络模型对浓度的检测精度。

针对目前电子鼻在检测气体浓度精度不高的问题,本文利用粒子群算法的局部搜索能力和人工蜂群算法的全局搜索能力,将两个算法进行嵌入融合,并与极限学习机相结合,最终达到提高电子鼻气体浓度检测精度的目的。

\* 基金项目:重庆市基础研究与前沿探索项目(cstc2018jcyjAX0549);重庆市教育委员会科学技术研究项目(KJQN201800617)

1 基于 PSOABC-ELM 的检测算法

本文结合粒子群算法<sup>[8]</sup>(Particle Swarm Optimization, PSO)模型简单、局部搜索能力强和人工蜂群算法<sup>[9]</sup>(Artificial Bee Colony Algorithm, ABC)种群丰富、全局搜索能力强的优点,提出一种 PSOABC-ELM 算法。其具体思想是,以 PSO 算法为主框架,结合 ABC 算法中改进的跟随蜂和改进的侦察蜂阶段来提高粒子群算法的全局搜索能力。在该混合算法中,生成一个初始解,并计算每个个体的 pbest。还增加了 3 个控制参数(pbest<sub>Measure</sub>、Limit1 和 Limit2),以确定执行算法的哪个阶段。pbest<sub>Measure</sub> 用于通过记录 pbest 在迭代期间是否更新来管理种群中的个体,通过比较 pbest<sub>Measure</sub> 的值与 Limit1 和 Limit2 的大小,来判断计算粒子新位置所采用的算法模型。重复迭代,通过不断更新粒子的位置,直至达到最小误差或者最大迭代次数终止,得到算法的最优解,将最优解作为 ELM 神经网络的输入层与隐含层的权值和隐含层的阈值,降低随机给定权值阈值带来的误差。流程图如图 1 所示,算法具体步骤如下:

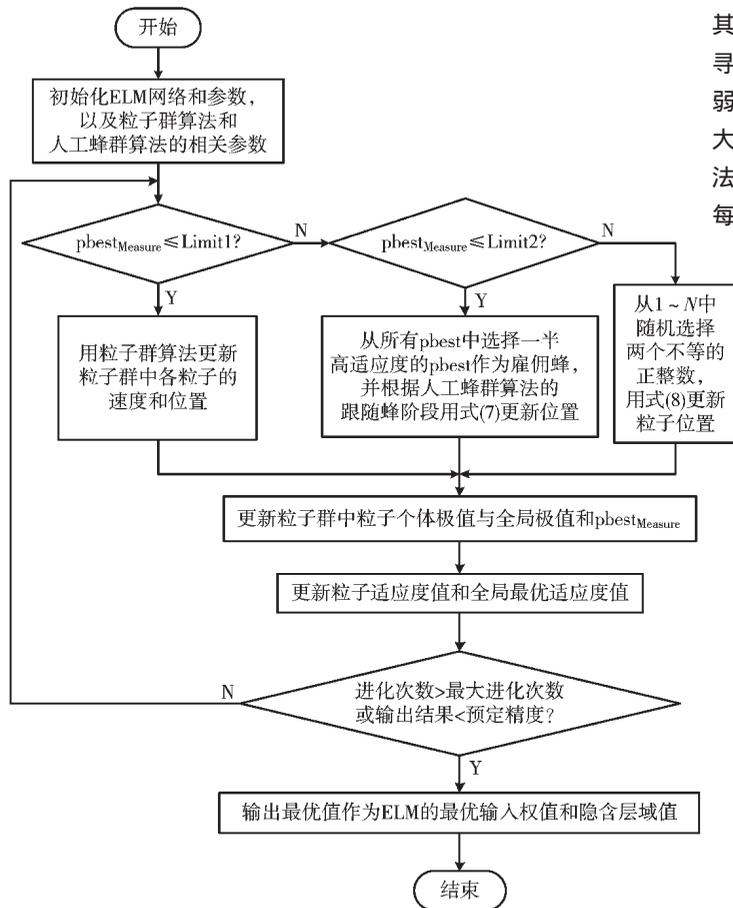


图 1 PSOABC-ELM 算法流程图

(1)确定神经网络拓扑结构。得到具有  $n$  个输入层节点、 $l$  个激活函数为  $g(\cdot)$  的隐含层节点、 $m$  个输出层节点的 ELM;形成  $N$  个任意不同的维度为  $D$  的训练样本集,

随机产生输入层与隐含层的权值和隐含层的阈值。

(2)参数初始化。对粒子群算法和人工蜂群算法中所有参数初始化,同时计算控制参数 Limit1 和 Limit2,其中  $l_1=0.01, l_2=0.05$ ;并将控制参数 pbest<sub>Measure</sub> 初始化为 0。

$$\text{Limit1} = l_1 \cdot \frac{D \times N}{2} \tag{1}$$

$$\text{Limit2} = l_2 \cdot \frac{D \times N}{2} \tag{2}$$

(3)选择适应度函数。本文选用均方误差作为最佳适应度函数,适应度值越小,个体越优秀。

(4)将处理后的数据输入模型,根据 pbest<sub>Measure</sub> 判断当前个体执行哪个阶段。

①当 pbest<sub>Measure</sub>( $i$ ) ≤ Limit1,用传统的 PSO 算法得到最优个体对应的粒子位置,算法的核心是对粒子的速度  $x$  和位置  $v$  进行更新,其更新公式如下:

$$v_{id} = \omega v_{id} + c_1 r_1 (P_{id} - x_{id}) + c_2 r_2 (P_{gd} - x_{id}) \tag{3}$$

$$x_{id} = x_{id} + v_{id} \tag{4}$$

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{t_{\max}} \times t \tag{5}$$

其中,  $\omega$  称为惯性因子,其值为非负,其值较大时,全局寻优能力强,局部寻优能力弱;较小时,全局寻优能力弱,局部寻优能力强<sup>[10]</sup>;  $\omega_{\max}$  和  $\omega_{\min}$  分别为惯性权重的最大值和最小值,通过将  $\omega$  调整为一个合适的值,使得算法的全局寻优能力和局部寻优能力达到一个平衡;  $c_1$  为每个粒子的个体学习因子,  $c_2$  为每个粒子的社会学习因子,通常取  $c_1=c_2=2$ ;  $r_1$  和  $r_2$  是  $[0, 1]$  内均匀分布的随机数,可增加搜索随机性;  $P_{id}$  表示第  $i$  个变量的个体极值的第  $d$  维,  $P_{gd}$  表示全局最优解的第  $d$  维。通常将位置和速度限制在  $[-x_{\max}, x_{\max}]$ 、 $[-v_{\max}, v_{\max}]$ 。

②如果 Limit1 < pbest<sub>Measure</sub>( $i$ ) ≤ Limit2,将从所有的 pbest 中选择一半高适应度的 pbest 作为雇佣蜂,利用 ABC 算法的跟随蜂阶段,跟随蜂根据位置信息,依概率选择适应度较高的蜜源信息,在雇佣蜂的基础上产生一个变化的位置,去寻找新的蜜源。其概率选择公式如下:

$$p_i = \text{fitness}(x_i) / \sum_{j=1}^{SN} \text{fitness}(x_j) \tag{6}$$

式中,  $\text{fitness}(x_i)$  为第  $i$  个跟随蜂的适应度值,  $p_i$  为第  $i$  个跟随蜂被选择的概率, SN 为蜜源数。一旦跟随蜂被选择,用式(6)计算  $x_i^{t+1}$ ,通过贪婪算法选择第  $i$  个跟随蜂原位置  $x_i^t$  和新位置  $x_i^{t+1}$  最优值,得到新的粒子位置;

$$x_i^{t+1} = x_i^t + \Psi \cdot (x_i^t - \text{pbest}_j^t) \tag{7}$$

式中,  $\Psi$  为区间  $[-1, 1]$  的随机值, pbest <sub>$j$</sub>  <sup>$t$</sup>  为第  $i$  个跟随蜂在  $t$  时刻的个体极值。

③如果 pbest<sub>Measure</sub>( $i$ ) ≥ Limit2,从  $\{1, 2, \dots, N\}$  中随机选

择两个正整数  $k_1 \neq k_2 \neq i$ , 用式(7)计算  $x_i^{t+1}$ , 通过贪婪算法选择  $x_i^t$  和  $x_i^{t+1}$  最优值得到新的粒子位置:

$$x_i^{t+1} = x_i^t + \Psi \cdot (\text{pbest}_{k_1}^t - \text{pbest}_{k_2}^t) \quad (8)$$

式中,  $\text{pbest}_{k_1}^t$  和  $\text{pbest}_{k_2}^t$  分别为第  $k_1$  和  $k_2$  个跟随蜂在  $t$  时刻的个体极值。

得到新的粒子位置后, 更新粒子个体极值  $\text{pbest}$  和全体极值  $\text{gbest}$ , 如果更新了  $\text{pbest}$  的值, 则  $\text{pbest}_{\text{Measure}}$  的设置为 0, 否则  $\text{pbest}_{\text{Measure}}$  将增加 1。

对算法进行重复迭代, 通过不断更新粒子的位置, 直至达到最小误差或者最大迭代次数终止, 得到算法的最优解, 将最优解作为 ELM 的最优输入权值和隐含层阈值。

(5) 用训练的模型对测试集进行检测, 并将检测结果与真实值进行比较, 得到两个值之间的误差, 通过对比误差得出结论。

## 2 实验仿真及结果分析

### 2.1 实验数据集介绍

该研究使用的数据来自加州大学欧文分校(University of California Irvine, UCI)机器学习数据库, 数据集为不同浓度下气体传感器阵列数据集。该数据集收集了分别来自 16 个化学传感器的 6 种不同气体不同浓度的测量数据, 这些数据提供了有关传感器在每次测量中暴露的浓

度水平信息, 可用于分析传感器回归问题<sup>[11-12]</sup>。

### 2.2 实验设置与仿真

为充分验证本文提出的算法的有效性, 本文选择 BP 神经网络<sup>[13]</sup>、SVM 神经网络<sup>[14]</sup>、ELM 神经网络<sup>[15]</sup>、PSO-ELM<sup>[16]</sup>神经网络作为对比算法。实验基于这 4 种对比算法以及本文提出的 PSOABC-ELM 算法对传感器收集到的污染物气体进行浓度检测, 为得到更好的网络检测性能, 首先对高维的原始数据进行归一化处理, 再利用主成分分析法(Principal Component Analysis, PCA)对数据进行降维, 通过大量实验验证, 最终确定提取 6 个主成分数据作为电子鼻模型的输入数据。各模型参数设置如表 1 所示。

对各算法进行多次实验取平均值, 各算法的检测效果可以由统计平均绝对误差(MAE)、均方根误差(RMSE)、相关系数( $R$ )衡量, 结果见表 2, 检测结果分别如 2~图 7 所示, 其中图横坐标为样本编号, 纵坐标为样本对应的气体浓度值, 单位为 ppmv。

### 2.3 实验结果分析

实验设置下各算法对不同气体的浓度检测效果如表 2 所示。通过表 2 可知, BP、SVM、ELM、PSO-ELM、PSOABC-ELM 算法对 6 种污染物浓度检测结果与监测值的  $R$  平均值分别为 0.853 9、0.987 8、0.901 3、0.994 2、0.998 4, 其中 PSOABC-ELM 算法对 6 项污染物的  $R$  值均达到了 0.99 以上。

表 1 各模型参数设置

| 模型         | 输入层节点数 | 隐含层节点数 | 隐含层激活函数 | 学习函数   | 迭代次数 | 粒子种群数 |
|------------|--------|--------|---------|--------|------|-------|
| BP         | 6      | 9      | tansing | loding | /    | /     |
| SVM        | 6      | /      | /       | /      | /    | /     |
| ELM        | 6      | 20     | sigmoid | /      | 100  |       |
| PSO-ELM    | 6      | 20     | sigmoid | /      | 100  | 40    |
| PSOABC-ELM | 6      | 20     | sigmoid | /      | 100  | 40    |

表 2 实验设置下各算法对不同气体的浓度检测效果比较

| 评价标准 | 检测模型       | 乙醇       | 乙烯       | 氨         | 乙醛       | 丙酮       | 甲苯      |
|------|------------|----------|----------|-----------|----------|----------|---------|
| MAE  | BP         | 13.583 2 | 29.839 6 | 89.330 6  | 40.825 7 | 36.606 5 | 3.213 3 |
|      | SVM        | 6.407 9  | 5.633 9  | 21.815 9  | 0.512 4  | 6.030 3  | 0.993 5 |
|      | ELM        | 26.020 5 | 21.898 5 | 100.544 9 | 11.563 7 | 34.854 1 | 7.138 6 |
|      | PSO-ELM    | 2.185 4  | 2.719 9  | 37.960 7  | 0.425 0  | 4.233 8  | 0.535 7 |
|      | PSOABC-ELM | 0.189 7  | 1.554 0  | 18.090 2  | 0.116 2  | 0.413 1  | 0.152 0 |
| RMSE | BP         | 15.422 1 | 31.740 8 | 216.657 7 | 49.124 2 | 42.400 7 | 3.971 1 |
|      | SVM        | 8.084 9  | 7.379 4  | 66.412 4  | 0.857 9  | 6.887 8  | 1.805 1 |
|      | ELM        | 28.025 1 | 26.075 3 | 209.283 6 | 8.760 4  | 41.976 4 | 9.244 8 |
|      | PSO-ELM    | 2.744 8  | 3.504 7  | 50.657 7  | 0.312 1  | 5.730 4  | 0.671 0 |
|      | PSOABC-ELM | 0.256 0  | 2.003 9  | 26.932 2  | 0.084 3  | 0.609 3  | 0.203 6 |
| $R$  | BP         | 0.982 9  | 0.970 6  | 0.528 2   | 0.697 0  | 0.960 7  | 0.984 5 |
|      | SVM        | 0.991 2  | 0.992 4  | 0.955 1   | 0.998 3  | 0.997 2  | 0.992 7 |
|      | ELM        | 0.979 0  | 0.955 2  | 0.605 2   | 0.978 8  | 0.916 2  | 0.973 6 |
|      | PSO-ELM    | 0.998 8  | 0.998 3  | 0.972 1   | 0.999 5  | 0.997 9  | 0.998 9 |
|      | PSOABC-ELM | 0.999 8  | 0.999 4  | 0.992 1   | 0.999 6  | 0.999 7  | 0.999 9 |

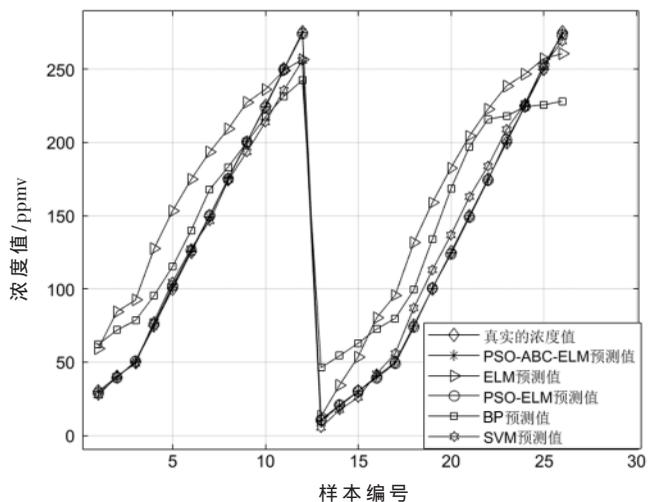


图2 乙醇预测值与真实值结果对比

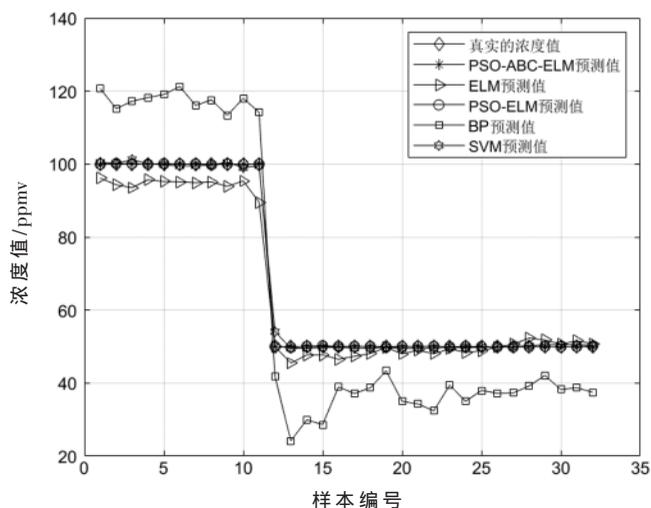


图5 乙醛预测值与真实值结果对比

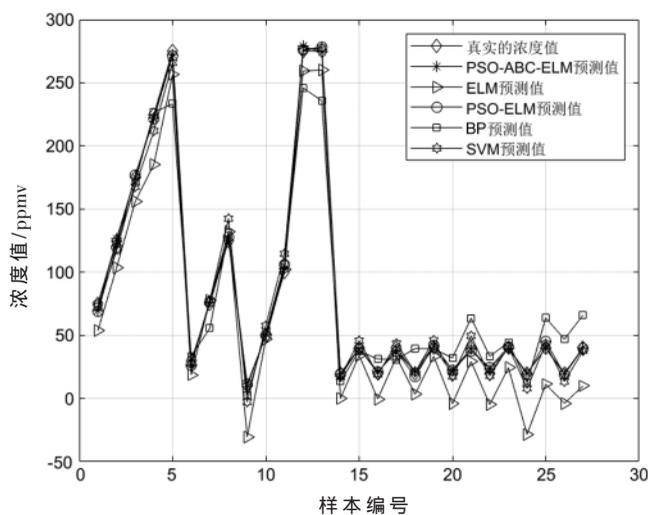


图3 乙烯预测值与真实值结果对比

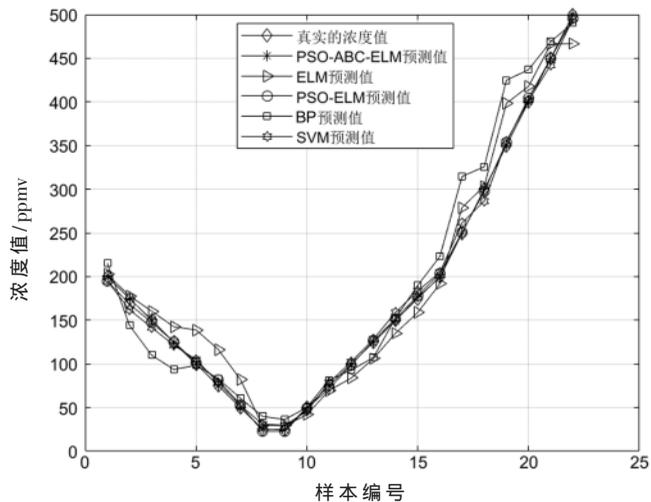


图6 丙酮预测值与真实值结果对比

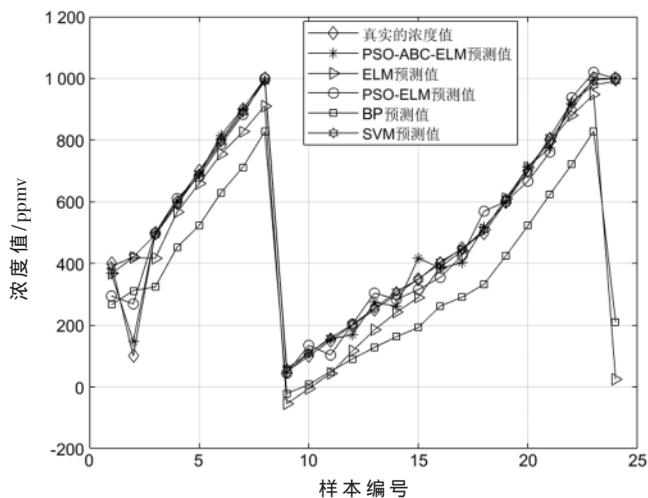


图4 氨预测值与真实值结果对比

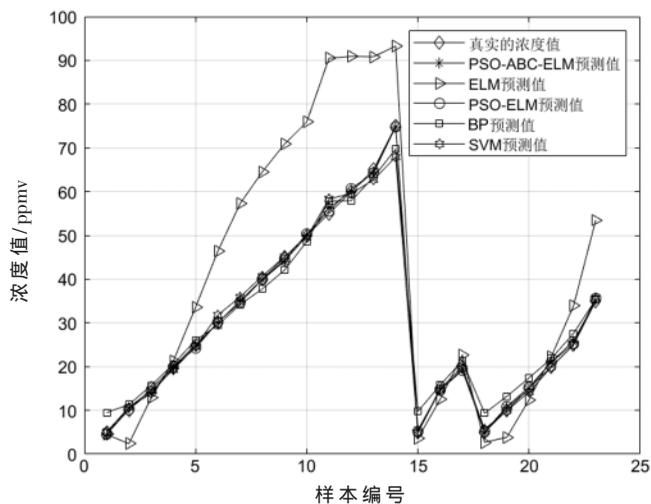


图7 甲苯预测值与真实值结果对比

对比5种算法对6种气体的MAE和RMSE,每种气体中PSOABC-ELM的MAE和RMSE均为最低值,这意

味着PSOABC-ELM检测值与真实的浓度值最接近,拥有最佳检测效果。PSOABC-ELM对6种气体的MAE除了氨的MAE大于2外,其余气体的MAE均小于2。

为方便跟踪每类气体浓度的检测值与真实值之间的差异,图2~图7分别给出了这6种气体通过这5种算法获得的气体浓度检测值与真实浓度值的跟踪曲线。由图可知,5种算法对6种气体的浓度检测值均可以较好地反映出污染物浓度变化的趋势,其中PSOABC-ELM对于6种气体检测结果均与监测值基本拟合。

综上所述,PSOABC-ELM算法检测精度高,性能稳定可靠,应用PSOABC-ELM算法的电子鼻系统可以成功地对空气污染物浓度进行检测。

### 3 结论

本文针对电子鼻应用于气体污染物浓度检测时难以达到理想精度的问题,提出了一种PSOABC-ELM算法,对电子鼻模式识别模块进行改进,并将其与BP、SVM、ELM和PSO-ELM算法进行比较。仿真结果表明,本文提出的算法有效提高了电子鼻对气体浓度检测的精度和可靠性。但其在气体浓度短时间内变化范围较大时误差会变大,今后的工作将针对本文提出的算法进行改进,提高在短时间内对变化范围较大的气体浓度检测精度,扩展应用范围。

### 参考文献

- [1] SCOTT S M, JAMES D, ALI Z. Data analysis for electronic nose systems[J]. *Microchimica Acta*, 2006, 156(3-4): 183-207.
- [2] TAŞ TAN M, GÖKOZAN H. Real-time monitoring of indoor air quality with Internet of Things-based E-nose[J]. *Applied Sciences*, 2019, 9(16): 3435.
- [3] LEGGIERI M C, MAZZONI M, FODIL S, et al. An electronic nose supported by an artificial neural network for the rapid detection of aflatoxin B1 and fumonisins in maize-Science-Direct[J]. *Food Control*, 2021, 123: 107722.
- [4] TOZLU B H, SIMSEK C, AYDEMİR O, et al. A high performance electronic nose system for the recognition of myocardial infarction and coronary artery diseases[J]. *Biomedical Signal Processing and Control*, 2021, 64: 102247.
- [5] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: a new learning scheme of feedforward neural networks[C]// *IEEE International Joint Conference on Neural Networks*. IEEE, 2005.
- [6] 闻新. 智能故障诊断技术: MATLAB应用[M]. 北京: 北京航空航天大学出版社, 2015.
- [7] 周志华. 机器学习[J]. *中国民商*, 2016, 3(21): 93.
- [8] EBERHART R C, KENNEDY J. A new optimizer using particle swarm theory[C]// *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 1995. MHS'95, 1995.
- [9] KARABOGA D. An idea based on honey bee swarm for numerical optimization[Z]. 2005.
- [10] Lo Chih-Hao, Yuan Jingxia, Ni Jun. Optimal temperature variable selection by grouping approach for thermal error modeling and compensation[J]. *International Journal of Machine Tools & Manufacture*, 1999, 39(9): 1383-1396.
- [11] VERGARA A, VEMBU S, AYHAN T, et al. Chemical gas sensor drift compensation using classifier ensembles[J]. *Sensors & Actuators B Chemical*, 2012, 166-167(none): 320-329.
- [12] RODRIGUEZ-LUJAN I, FONOLLOSA J, VERGARA A, et al. On the calibration of sensor arrays for pattern recognition using the minimal number of experiments[J]. *Chemometrics & Intelligent Laboratory Systems*, 2014, 130: 123-134.
- [13] HE J, XU L, WANG P, et al. A high precise E-nose for daily indoor air quality monitoring in living environment[J]. *Integration*, 2016, 58(jun): 286-294.
- [14] LAREF R, LOSSON E, SAVA A, et al. A comparison between SVM and PLS for E-nose based gas concentration monitoring[C]// *International Conference on Industrial Technology (ICIT)*, Lyon, France. 20-22 February, 2018.
- [15] QIU S, WANG J. The prediction of food additives in the fruit juice based on electronic nose with chemometrics[J]. *Food Chemistry*, 2017, 230(SEP.1): 208-214.
- [16] 庄玉册, 黎蔚. 基于 PSO 优化极限学习机神经网络的空气质量预报[J]. *沈阳工业大学学报*, 2020, 42(2): 213-217.

(收稿日期: 2021-01-19)

### 作者简介:

王洁(1996-),女,硕士研究生,主要研究方向:机器嗅觉、模式识别。

陶洋(1964-),男,博士后,教授,博士生导师,主要研究方向:机器学习、模式识别。

梁志芳(1989-),女,博士,讲师,主要研究方向:机器嗅觉、模式识别、迁移学习。



扫码下载电子文档

## 版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所