

DOI:10.3788/JQO20243002.0501

BP神经网络和PLS方法在TDLAS定量分析 混合气体中的对比研究

张悦, 李勇, 李泽兵, 邬金妮, 赵刚[†], 马维光[†]

(山西大学激光光谱研究所 量子光学与光量子器件国家重点实验室, 山西 太原 030006)

摘要: 在基于激光吸收光谱技术的痕量气体检测过程中, 不同气体光谱重叠现象会对提取吸收光谱真实特征造成影响, 引入反演浓度的误差。设计了结合多通道池和直接吸收光谱技术的痕量气体检测装置, 分别应用BP神经网络和偏最小二乘法(PLS)消除干扰, 使用模拟光谱对BP神经网络模型和PLS模型进行训练, 从而大大简化了训练过程。为了保证模拟光谱的准确性, 利用F-P腔的透射峰进行定标, 获取了准确的激光波长变化, 继而带入模拟过程。使用实测的光谱数据作为测试集, 得到系统对浓度响应的线性拟合度大于0.99, 误差小于0.21%。最后, 分析了etalon噪声对两种算法的影响, 结果表明, PLS模型的浓度误差小于 4.4×10^{-7} , 比基于BP神经网络模型的结果优5倍以上。

关键词: 可调谐二极管激光吸收光谱; BP神经网络; 偏最小二乘法

中图分类号: O657.63

文献标识码: A

Comparative Study of the BP Neural Network and PLS Method in TDLAS Quantitative Analysis of Mixed Gases

ZHANG Yue, LI Yong, LI Ze-bing, WU Jin-ni, ZHAO Gang[†], MA Wei-guang[†]

(State Key Laboratory of Quantum Optics and Quantum Optics Devices, Institute of Laser Spectroscopy,
Shanxi University, Taiyuan 030006, China)

Abstract: In trace gas detection based on laser absorption spectroscopy, the overlap of different gas absorption spectra affects the extraction of the characteristics of the absorption spectrum, which consequently introduce the error of the deduced concentration. In this paper, a trace gas detector by the combination of a multipass cell and direct absorption spectroscopy is present. BP neural network model and PLS model are utilized, respectively, to restrict the spectral overlap. In order to simplify the training progress, simulated spectral models have been used as training set. The laser frequency are calibrated with the help of the transmission peaks of an F-P cavity and then introduced into the simulation model. As a result, the accuracy of the simulation has been

收稿日期: 2023-02-24; 接受日期: 2023-04-04

基金项目: 国家重点研发计划资助项目(2022YFC3703900); 国家自然科学基金(61875107; 62175139); 山西省留学人员科技活动择优资助项目(20220001)

作者简介: 张悦(1998-), 女, 山西太原人, 硕士研究生, 主要研究方向为光谱检测技术与数据处理算法。

[†]通信作者: 赵刚, E-mail: gangzhao@sxu.edu.cn; 马维光, E-mail: mwg@sxu.edu.cn

improved. Then, the measured data is acted as test set. The linearity of the system's response to the concentration is greater than 0.99, and the relative error is less than 0.21%. Finally, the influence of etalon noise to the two algorithm has been analyzed and the result shows the concentration error with PLS model is less than 4.4×10^{-7} , which is more than five times that by BP neural network.

Keywords: TDLAS; BP neural network; partial least squares method

0 引言

可调谐二极管激光吸收光谱技术 (Tunable Diode Laser Absorption Spectroscopy, TDLAS) 通过测量气体分子吸收光谱, 可以反演未知气体的种类和浓度, 其具有高分辨率、快响应度、高灵敏度、低探测限、非接触式等优势。近二十年来, 伴随着近红外波段相关光学器件的大规模商业化生产和价格降低, 基于 TDLAS 开发的痕量气体仪已得到广泛的应用^[1-4]。

灵敏度和选择性是影响 TDLAS 应用性的两个重要因素。为了提升灵敏度, 人们提出结合多通道池^[5-6]或者光学腔^[7-8]的方法, 利用激光在两个高反射镜之间的来回反射增长光与气体物质的作用路径, 可以将最小可探测气体浓度降低三个数量级以上。为了提升选择性, 通常需要对测量波长进行精心挑选, 从而避免多光谱的重叠。然而在红外波段, 由于分子丰富的振转能级跃迁, 经常很难找到单一光谱跃迁线。当不同气体吸收线存在重合时, 严重影响待测气体光谱的提取, 导致了检测系统准确性的降低。

为了解决这个问题, 人们将化学计量方法与红外光谱分析相结合, 通过建立混合气体浓度回归预测模型, 来解决多组分交叉干扰的问题。杨雅涵^[9]等应用偏最小二乘法 (Partial Least-Square, PLS) 检测了天然气中的硫化氢的直测光谱和差分光谱, 发现其预测结果均优于最小二乘法模型; 李越胜^[10]等用 BP 神经网络对燃煤热值进行预测, 在一定程度上减弱了试验参数波动和基体效应造成的影响; 张强^[11]等将深度神经网络 (Deep Neural Networks, DNN) 应用于傅里叶变换红外光谱 (Fourier Transform Infrared, FTIR), 对多组分易挥发性有机物进行浓度反演研究, 达到了较高精度。其他方法还包括人工神经网络 (Artificial Neural Networks, ANN)^[12-13]、支持向量机 (Support Vector Machine, SVM)^[14-15]等。

在这些算法里, 获得训练集数据通常需要对不

同状态下的气体光谱进行大规模测量, 通过训练得到满足精度的模型, 才可以应用于实际测试。这个过程不仅需要消耗大量时间进行测量, 而且当测试环境与训练环境改变时, 会产生误差影响模型精度。Ismail Bayrakli^[16]等人提出应用模拟光谱信号进行训练的方式, 利用从光谱数据库 (如 Hitran 数据库) 中获取的光谱参数, 结合光谱线型函数, 模拟不同浓度下的气体光谱信号, 将其作为训练集。这样避免了对训练集进行大量测量采集的需求, 也提高了模型精度。然而, 在该方式中, 模拟光谱与测量结果的一致性是影响训练效果的决定性因素, 这就要求模拟中使用与实际实验过程完全一致的光学参数, 特别是激光的波长, 其作为光谱信号的横坐标影响着整个光谱线形。

本文设计了结合多光程池和 TDLAS 技术的痕量气体检测装置, 可以实现对氢气和乙炔两种气体的同时检测, 并且结合 BP 神经网络和 PLS 算法, 对混叠干扰的两种气体光谱进行浓度反演。其中, 相较于其他算法, BP 神经网络可以较好地描述非线性关系, 而 PLS 算法则可以在样本数少于变量数的情况下很好地建模。我们使用模拟光谱对 BP 神经网络和 PLS 模型进行训练, 为了提升预测结果的准确性, 利用 F-P 腔的透射峰实时测量了实际激光波长, 代入模拟过程获得准确的训练光谱。

1 基本原理

1.1 TDLAS 气体检测基本原理

当激光照射分子时, 特定波长的光波被选择性吸收, 构成气体的吸收光谱。一束单色光经过气体吸收后, 其光强的变化遵循比尔朗伯 (Beer-Lambert law) 定律^[17-18], 其表达式为:

$$I(\nu) = I_0(\nu)e^{-\alpha(\nu)cL}, \quad (1)$$

$$\alpha(\nu) = PS(T)\phi(\nu). \quad (2)$$

公式 (1) 中, $I(\nu)$ 为穿过待测气体后的透射光光强, $I_0(\nu)$ 为入射光强, $\alpha(\nu)$ 为吸收系数, 该系数

与气体的种类以及激光频率有关, c 为待测气体的浓度, L 为光所经过的待测气体的吸收路径长度。 $P[\text{atm}]$ 表示气室中气体的总压强, $S(T)$ 为该气体特征谱线的吸收线强度, 与温度有关, $\phi(\nu)$ 为线型函数, 它与温度、总压力和气体中的各成分含量有关。

1.2 BP神经网络基本原理

BP算法即误差反向传播算法应用于神经网络的思想最早是由Paul Werbos在1974年首次论证, 后由Hinton和David E. Rumelhart等人在1986年重新发明, 并广泛应用于完成分类、回归等任务中^[19-20]。

假设气体光谱的训练集数据为 $D = (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, 其中 $i = 1, 2, \dots, n$, $x_i \in R^d$, $y_i \in R^q$, 构建一个由输入层、隐藏层和输出层组成的BP神经网络, 其中输入层节点数 d , 隐藏层节点数 l , 输出层节点数 q , 确定隐藏层激活函数 f , 输入层与隐藏层之间的连接权值 w_{ij} 及隐藏层阈值 a , 由式(4)计算隐藏层输出 H_j

$$h_j = \sum_{i=1}^n (w_{ij} x_i + a_j), \quad j = 1, 2, \dots, l, \quad (3)$$

$$H_j = f(h_j). \quad (4)$$

初始化隐藏层与输出层之间的连接权值 w_{jk} 及输出层阈值 b , 进一步计算得到输出层输出 O_k , 见式(6)

$$o_k = \sum_{j=1}^l (w_{jk} H_j + b_k), \quad k = 1, 2, \dots, q, \quad (5)$$

$$O_k = f(o_k). \quad (6)$$

信息正向传输完成后, 接着计算输出值与实际值之差, 若两者误差相差较大, 则进行权重校正, 权重的校正是反向的, 即先校正输出层与隐藏层之间的连接权重, 再校正隐藏层与输入层之间的权重。计算输出层节点误差 σ_k 为

$$\sigma_k = (T_k - O_k) f'(o_k). \quad (7)$$

T_k 为期望输出, 由于误差反方向地传播, 我们可以根据输出层误差计算得到隐藏层误差 σ_j 为

$$\sigma_j = \left(\sum_k e_k w_{kj} \right) f'(h_j). \quad (8)$$

在误差的反向传播中, 权值的修正可由梯度下降法来实现, 见式(9)

$$\Delta w_{ji} = \eta e_j t_i. \quad (9)$$

确定学习率 η , e_j 为上一层节点 j 的误差, t_i 为下一层节点 i 的实际值, 以负梯度方向调整误差参数, 则可以实现权重的修正, 直到系统收敛。

构造的BP神经网络系统中, 输入层为气体光谱数据, 输出层为样本的浓度值。

1.3 偏最小二乘回归基本原理

PLS集线性回归分析、主成分分析和典型相关分析方法的基本功能于一体, 可以同时实现回归建模、数据结构简化以及两组变量之间的相关性分析^[21]。

假定有 q 个因变量 y_1, y_2, \dots, y_q 和 p 个自变量 x_1, x_2, \dots, x_p , 在观测 n 个样本点后, 构成自变量和因变量数据矩阵 $\mathbf{X} = [x_1, x_2, \dots, x_n]_{n \times p}$ 和 $\mathbf{Y} = [y_1, y_2, \dots, y_q]_{n \times q}$, 偏最小二乘回归分别在 \mathbf{X} 和 \mathbf{Y} 中提取成分 t_1 和 u_1 , 即 t_1 是 x_1, x_2, \dots, x_p 的线性组合, u_1 是 y_1, y_2, \dots, y_q 的线性组合。在提取成分时 t_1 和 u_1 必须满足以下两个条件:

(1) t_1 和 u_1 应尽可能多地表示各自数据表中的变异信息;

(2) t_1 和 u_1 达到最大相关性。

如果上述条件得到满足, 那么 t_1 和 u_1 就最大可能地包含了自变量 \mathbf{X} 和因变量 \mathbf{Y} 的信息, 同时自变量的成分 t_1 对因变量的成分 u_1 又具有最强的解释能力。在第一个成分 t_1 和 u_1 提取后, 分别实施 \mathbf{X} 对 t_1 的回归以及 \mathbf{Y} 对 u_1 的回归, 如果回归方程满足预设精度, 则算法停止; 否则, 利用 \mathbf{X} 被 t_1 解释后的残余信息以及 \mathbf{Y} 被 u_1 解释后的残余信息进行第二轮的成分提取。如此反复, 直到精度满足要求为止。若最终对 \mathbf{X} 提取 m 个成分 t_1, t_2, \dots, t_m , 最后通过实施 y_k 对 t_1, t_2, \dots, t_m 的回归, 然后再转化为 y_k 关于原变量 x_1, x_2, \dots, x_p 的回归方程, 其中 $k = 1, 2, \dots, q$, 这样就完成了偏最小二乘回归建模。

2 实验系统

实验系统由四部分组成, 主要包括激光频率定标系统、TDLAS测试系统、配气系统、数据采集系统, 如图1所示。

实验中选用中心波长为1531 nm的分布式反馈激光器(DFB)作为光源, 其工作温度和电流可以通过一个半导体激光器控制器(LDC501, Stanford Research System)进行控制。一个函数发生器(SDG2000X, SIGLENT)生成一个频率为5 Hz、幅度为800 mV的三角波, 对激光电流进行扫描, 从而调谐激光波长。输出激光经过一个隔离器和一

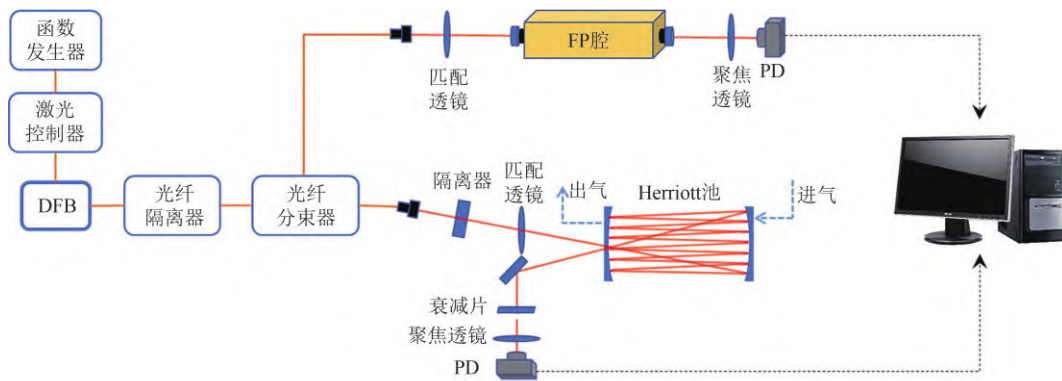
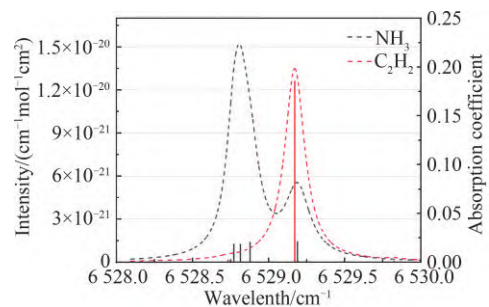


Fig. 1 TDLAS experimental system based on multi-path pools

图1 基于多光程池的TDLAS实验系统图

个9:1光纤分束器后分为两路。其中光功率较弱的一路激光依次通过准直器、光隔离器、匹配透镜后，耦合进入一个光学腔，腔的透射信号聚焦到一个铟镓砷光电探测器（PDA10CS2，Thorlab）上，用于激光的波长测量。该光学腔采用Fabry-Perot（F-P）光学腔结构，主要由两个高反镜组成，腔长为12 cm，根据 $FSR = \frac{c}{2nl}$ （ n 为折射率， l 为腔长），计算得出对应的自由光谱区（FSR）约为1 250 MHz。当激光频率与光学腔共振时，在腔内形成很强的光场，此时在腔的透射端将检测到腔模信号。而每两个腔模信号之间的激光频率间隔都等于腔的FSR。另外一束较强的光束通过匹配透镜进入Herriott吸收池，在吸收池中经过多次反射最后从原来的入射孔射出，激光在池内的光程达到22.86 m，经反射镜反射和聚焦透镜聚焦到另一个光电探测器（PDA10CF-EC，Thorlab）上。两个探测器的输出信号通过数据采集卡输入到电脑中，使用LabView程序处理腔模和直接吸收信号。采集卡以50 kHz的采样率采集气体吸收数据，每组相同浓度气体吸收数据采集15个样本。

激光扫描波长同时覆盖了 C_2H_2 和 NH_3 的吸收线，其吸收线位置如图2所示，图中红色竖线为 C_2H_2 在 $6\,529.171\,9\text{ cm}^{-1}$ 处的吸收线，吸收线强度为 $1.260 \times 10^{-20}\text{ cm}^{-1}/(\text{cm}^2 \cdot \text{mol})$ ，黑线为 NH_3 吸收线，主要包含在 $6\,529.190\,1\text{ cm}^{-1}$ 处的一条吸收线和在 $6\,528.77\text{ cm}^{-1}$ 处的三条吸收线，吸收线强度都在 $10^{-21}\text{ cm}^{-1}/(\text{cm}^2 \cdot \text{mol})$ 量级。我们对气压为1 atm、浓度分别为 2.96×10^{-4} 的 NH_3 和 7.5×10^{-5} 的 C_2H_2 的吸收光谱进行模拟，分别如图2中黑色和红色虚线所示。可以清楚地观察到两个气体的光谱存在严重的重叠

Fig. 2 Transition spectra and absorption lines of NH_3 and C_2H_2 图2 NH_3 和 C_2H_2 的跃迁谱线以及吸收线

现象，通过对重叠光谱进行分离可以更准确地反演两个气体的浓度值。

实验中，Herriott气室两端设计有进气口和出气口，将浓度为 1×10^{-3} 的氨气和浓度为 1×10^{-4} 乙炔的标准气体通入实验室自主装配的、精度优于1%的四通道气体混配器中，通过控制气体流速，进行不同浓度氨气和乙炔的混合气体的配比，同时采用复合真空计监测气室内压强。实验中测量的混合气体浓度如表1所示，浓度值根据流量配比值以及样气浓度换算而来。

3 结果与分析

3.1 激光频率定标

使用三角波对激光波长进行连续扫描，为了获取准确的激光波长变化，我们使用F-P腔对激光波长进行了测量，得到的腔透射信号如图3（a）中黑色曲线。从图中可以看出，当激光波长变化时，会不断出现透射尖峰，而每两个尖峰之间就表示激光

表1 实验设置浓度值
Tab.1 Experiment set concentration values

Group	1	2	3	4	5	6	7	8
NH ₃ concentration[×10 ⁻⁶]	250	270	296	300	333	364	400	416
C ₂ H ₂ concentration[×10 ⁻⁶]	85	82	75	79	75	72	66	65

波长变化了FSR值。将尖峰标出，如图中红点所示，利用尖峰位置画出激光波长变化结果如图3(b)所示，我们接着使用一个二次多项式对曲线进行拟

合，得到结果如图3(b)中红线所示，即为激光实时波长。根据图3(b)下图的拟合残差，我们得出拟合的误差小于1%。

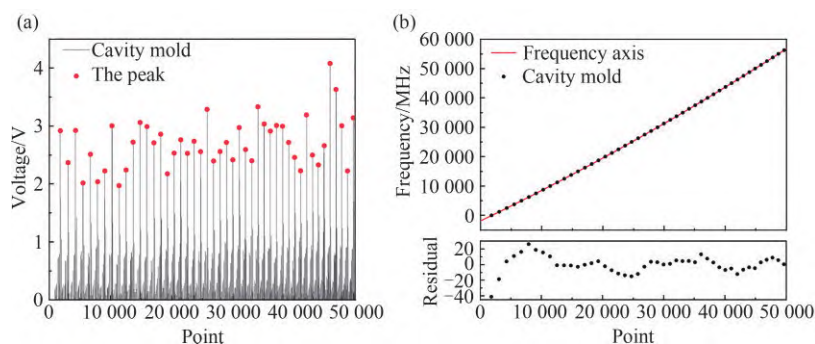


Fig. 3 Cavity mold and fitting wavelength

图3 腔模与拟合波长图

3.2 BP神经网络回归预测模型建立与应用

采用 HITRAN 数据库中氨气和乙炔气体在 1 531.3~1 531.9 nm 波段内的吸收光谱参数模拟获得吸收信号，设置氨气的浓度范围为 $2.5 \times 10^{-4} \sim 4.5 \times 10^{-4}$ ，乙炔浓度范围为 $6.2 \times 10^{-5} \sim 8.8 \times 10^{-5}$ ，按照浓度间隔约为 1×10^{-6} 各生成 290 个样本，其中随机选择 238 个样本作为训练集，剩下 52 个模拟样本结合八个实测样本，作为测试集。为了提高回归模型预测精度，采取不同种类气体独立建模的方法，分别建立了预测氨气浓度的神经网络模型和预测乙炔浓度的神经网络模型，记为 NH₃-model 和 C₂H₂-model，其输入变量为混合气体的光谱数据，输出变量分别为氨气浓度和乙炔浓度。

对训练神经网络的样本的输入输出进行归一化处理，所用归一化函数为：

$$y = (y_{\max} - y_{\min}) \times \frac{x - x_{\min}}{x_{\max} - x_{\min}} + y_{\min}, \quad (10)$$

其中 x 为待处理的输入矩阵， y 为处理过后的输出矩阵。

过多的隐藏层层数在提高网络预测结果精度的同时会增加结构的复杂性、降低训练速度，因此，

本文采用四层神经网络模型，即一个输入层、两个隐藏层和一个输出层。当隐藏层节点数逐层递减时，网络的预测误差较小，所以选择 100、10 作为两个隐藏层的节点数。输入层与隐藏层、隐藏层与隐藏层、隐藏层与输出层之间的传递函数分别采用 tansig、tansig 和 purelin 函数，设置最大迭代次数为 1 000，误差阈值为 1×10^{-6} ，学习速率为 0.01，用均方误差性能函数 (MSE) 来衡量网络性能，直到满足训练误差，网络自动停止。

对 NH₃-model 的训练集和测试集样本的预测结果如图 4 (a)、图 4 (b) 所示，C₂H₂-model 的训练集和测试集样本的预测结果如图 4 (c)、图 4 (d) 所示。

由图 4 可知对训练集样本本身作浓度回归预测时，NH₃-model 和 C₂H₂-model 训练集浓度设定值与浓度预测值均方根误差分别为 6.5×10^{-7} 和 1.7×10^{-7} ，误差小于 0.15%，决定系数 R^2 分别为 0.999 87 和 0.999 43；对测试集样本做浓度回归预测时，NH₃-model 和 C₂H₂-model 测试集浓度设定值与浓度预测值均方误差分别为 1.4×10^{-6} 和 2.2×10^{-7} ，误差小于 0.21%，决定系数 R^2 分别为 0.999 31 和 0.999 91。结果表明基于 BP 神经网络建立的 NH₃ 和 C₂H₂ 混合气体

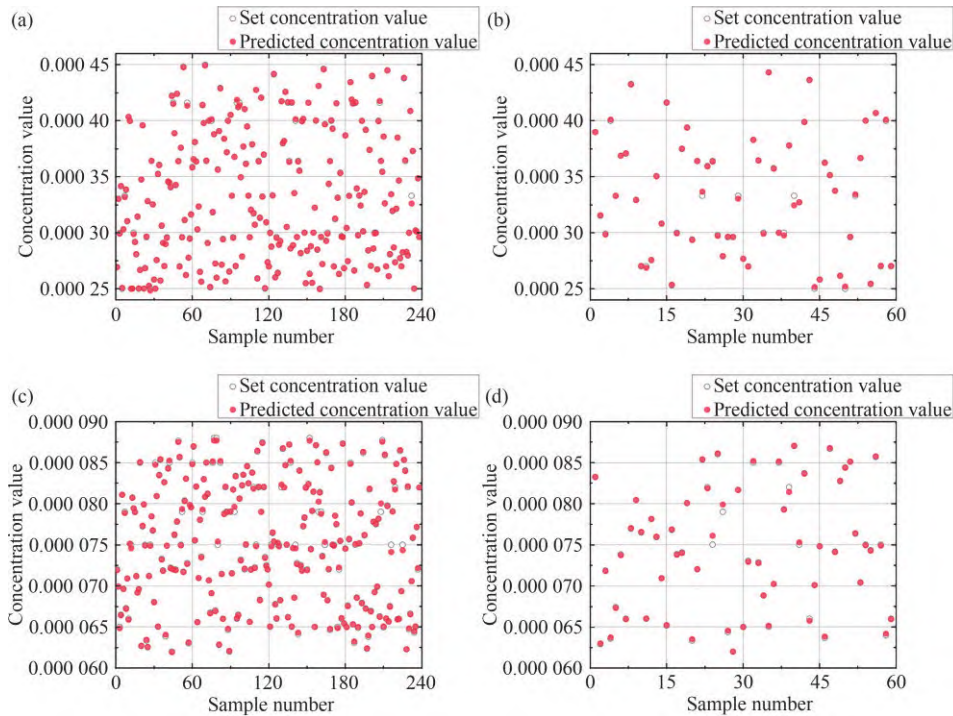


Fig. 4 Comparison of train set and test set results

(a) Comparison of set concentration and predicted concentration in NH_3 -model train set; (b) Comparison of set concentration and predicted concentration in NH_3 -model test set; (c) Comparison of set concentration and predicted concentration in C_2H_2 -model train set; (d) Comparison of set concentration and predicted concentration in C_2H_2 -model test set.

图4 训练集和测试集结果对比图

(a) NH_3 -model 训练集设定浓度和预测浓度对比图; (b) NH_3 -model 测试集设定浓度和预测浓度对比图;
(c) C_2H_2 -model 训练集设定浓度和预测浓度对比图; (d) C_2H_2 -model 测试集设定浓度和预测浓度对比图。

浓度回归预测模型的预测值和实际值可以很好地吻合, 有较好的预测能力, 拟合效果好, 误差小。

为了验证模型的准确性, 我们使用测量光谱对学习获得的模型进行了测试, 首先匹配模拟光谱与测量光谱, 应用之前计算得到的频率轴作为横坐标, 模拟绘制的吸收线图像与实际测量结果如图5所示。在

图5中, (a) (c) (e) 分别为乙炔浓度为 1×10^{-4} 、氨气浓度为0, 乙炔浓度为0、氨气浓度为 1×10^{-3} 以及乙炔浓度为 3.64×10^{-4} 、氨气浓度为 7.2×10^{-5} 的结果, (b) (d) (f) 为模拟和测量的残差图, 根据残差计算得信噪比分别为132.91、160.57和126.27, 可以说明模拟结果和测量结果实现了很好地匹配。

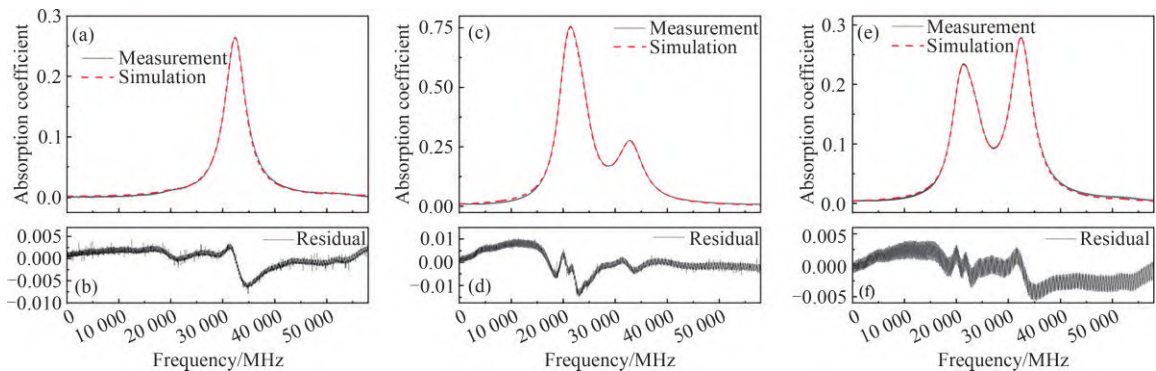


Fig. 5 Comparison of Simulation and Measurement

图5 模拟与测量对比图

训练好的神经网络模型用于预测各待测样本的输出,将之前实验测得的八组不同浓度混合气体的光谱数据代入训练好的预测模型中,对每组浓度的15个样本进行15次预测,取平均值作为预测样本的最终浓度值。神经网络预测浓度值与实际测得浓度值的对比如图6所示,其中图6(a)和图6(b)分别表示 NH_3 和 C_2H_2 两种气体,曲线的拟合度 R^2 均大于0.999,说明网络训练质量较高,对不同浓度的混

合气体预测有较好的结果。

3.3 PLS算法回归预测模型建立与应用

使用和BP神经网络训练集相同的数据来建立PLS回归模型,同样将实验测得的八组不同浓度的混合气体光谱数据代入建立好的PLS模型中,得到预测浓度值与实际浓度值的对比如图7所示,其中图7(a)和图7(b)分别表示 NH_3 和 C_2H_2 两种气体,曲线拟合度 R^2 均大于0.99,证实系统线性响应度较好。

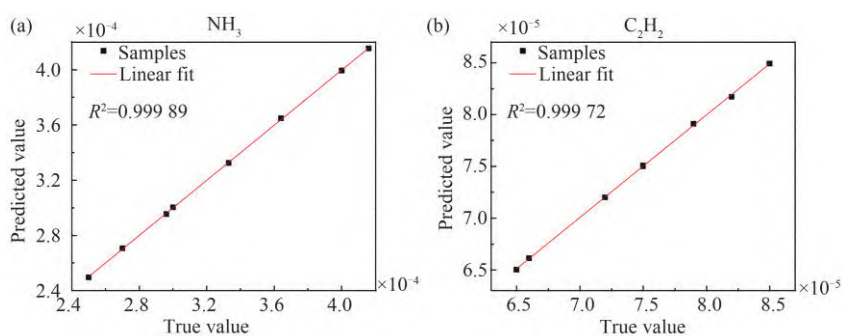


Fig. 6 Predicted results by BP neural network

图6 BP神经网络预测结果

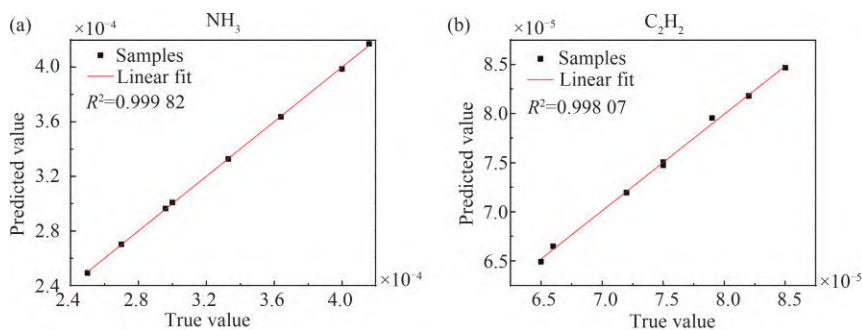


Fig. 7 Predicted results by PLS

图7 PLS回归预测结果

3.4 比较 etalon 噪声对 BP 神经网络和偏最小二乘回归的影响

etalon 噪声是影响光谱信号的随机噪声的主要来源之一,会导致反演浓度的误差和长期漂移。etalon 噪声主要由两个光学平面之间的干涉引起,其具有激光频率依赖性,主要表现为叠加在光谱信号上的正弦波。为了分析 etalon 噪声对两种算法的影响,我们人为地给实验测得数据叠加一个正弦信号,来模拟 etalon 噪声,模拟的 etalon 噪声对应于光学长度为 19 cm 的两个光学端面干涉,幅度为有用信号的 2.8%。将加上正弦的实验数据分别应用于神经网络和 PLS 模

型中,比较二者的预测结果,如图8所示。

由图8(a)和图8(b)可知,对加入模拟干涉噪声的光谱信号,BP神经网络的预测结果与真实值相差较大,甚至无法线性表示;PLS的预测结果和真实值可以线性拟合,计算得到预测氨气浓度的PLS模型相关系数为0.99979,预测乙炔浓度的PLS模型的相关系数为0.99975,基本能完成测量目标,测量准确度满足要求。

为了更好地对比两种算法对加入干涉噪声模拟信号的光谱的测量能力,计算两种算法对加正弦和不加正弦的测量误差,由图8可以看出,PLS算法

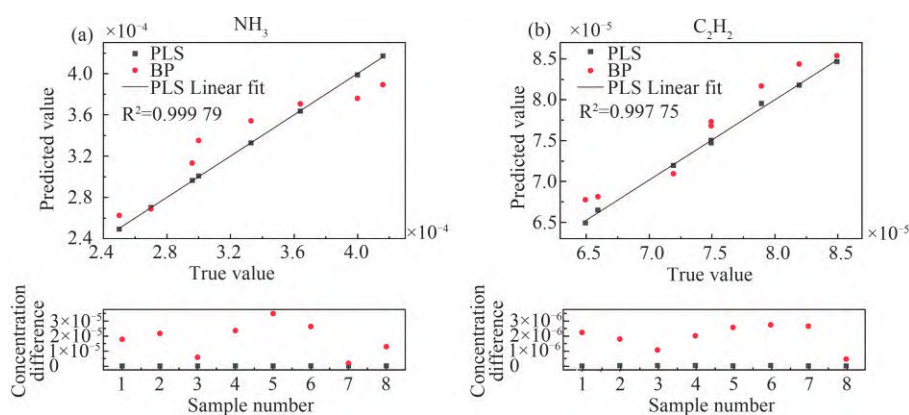


Fig. 8 Comparison between predicted value and true value of PLS and BP regression model

图8 PLS与BP回归模型的预测值与真实值对比

的测量准确度远高于BP神经网络，在测量氨气的模型中，BP神经网络的绝对误差在 3.5×10^{-5} 之内，PLS算法的绝对误差在 1.5×10^{-7} 之内；在测量乙炔的模型中，BP神经网络的绝对误差在 2.7×10^{-6} 之内，PLS算法的绝对误差在 4.4×10^{-7} 之内。这可以说明，PLS算法能够对系统中的信息进行分解和筛选，提取对因变量解释性最强的综合变量，辨识系统中的信息和噪声，对背景变换的容忍度强，能够在存在规律噪声的情况下工作，而BP神经网络由于在训练时未加入此类噪声，在遇到有噪声的样本时表现很差，无法对抗噪声，泛化能力差。

4 结论

本文发展了结合多通道池和TDLAS的痕量气体检测技术，采用基于BP神经网络和PLS算法的混合气体浓度定量分析的方法，根据HITRAN数据库提供的光谱参数模拟出不同浓度的气体吸收光谱作为

算法的训练集，避免了大规模数据测量。通过使用F-P腔实时测量激光频率变化，利用F-P腔的透射信号定标的激光频率进行了光谱模拟。之后我们利用该训练集训练得到预测氨气和乙炔浓度的BP神经网络模型和PLS模型，并且使用测量得到的光谱对BP神经网络模型和PLS模型进行了测试，对不同浓度结果进行线性拟合得到的相关系数 R^2 均大于0.99，预测值和实际值的误差小于0.34%。最后，针对光谱信号中最常见的非随机噪声—etalon噪声，我们研究了它对BP神经网络和PLS两种算法的影响，结果表明，PLS的表现远好于BP神经网络，可以抑制etalon噪声的影响，对乙炔和氨气获得的浓度误差分别小于 4.4×10^{-7} 和 1.5×10^{-7} 。因此，在混合气体浓度的定量分析中，BP神经网络和PLS算法在有随机噪声的情况下都有较大的应用潜力，而PLS对于干涉噪声的分辨能力较强。本研究为发展基于激光光谱的高灵敏、高分辨多气体检测提供了技术支撑。

参考文献

- [1] 聂伟, 阚瑞峰, 杨晨光, 等. 可调谐二极管激光吸收光谱技术的应用研究进展[J]. 中国激光, 2018, **45**(9):0911001. DOI: 10.3788/CJL201845.0911001.
- NIE W, KAN R F, YANG C G, *et al.* Progress in the application of tunable diode laser absorption spectroscopy[J]. *Chinese Journal of Lasers*, 2018, **45**(9):0911001. (in Chinese). DOI: 10.3788/CJL201845.0911001.
- [2] 姚路, 刘文清, 刘建国, 等. 基于TDLAS的长光程环境大气痕量CO监测方法研究[J]. 中国激光, 2015, **42**(2):0215003. DOI: 10.3788/CJL201542.0215003.
- YAO L, LIU W Q, LIU J G, *et al.* Study on the monitoring method of atmospheric trace CO in long optical path environment based on TDLAS[J]. *Chinese Journal of Lasers*, 2015, **42**(2):0215003. (in Chinese). DOI: 10.3788/CJL201542.0215003.
- [3] LI Y, WANG R Z, TITTEL F K, *et al.* Sensitive methane detection based on quartz-enhanced photoacoustic spectroscopy with a high-power diode laser and wavelet filtering[J]. *Optics and Lasers in Engineering*, 2020, **132**:106155. DOI: 10.1016/j.

- optlaseng.2020.106155.
- [4] 王铁云, 张雷, 董磊, 等. 激光遥测甲烷气体最低可探测浓度[J]. 中国激光, 2006, **33**(3):405–407.
WANG T Y, ZHANG L, DONG L, *et al.* Laser telemetry of the lowest detectable concentration of methane gas[J]. *Chinese Journal of Lasers*, 2006, **33**(3):405–407. (in Chinese).
- [5] LIU X, MA Y. Sensitive carbon monoxide detection based on light-induced thermoelastic spectroscopy with a fiber-coupled multipass cell[J]. *Chinese Optics Letters*, 2022, **20**(3):031201. DOI: 10.3788/COL202220.031201.
- [6] 杨牧, 李传亮, 魏计林. 基于 Herriott 型长程池的光学设计的研究[J]. 量子光学学报, 2013, **19**(2):189–194. DOI: 10.3788/ASQO2013902.0189.
YANG M, LI C L, WEI J L. Research on optical design based on Herriott long range cell[J]. *Journal of Quantum Optics*, 2013, **19**(2):189–194. (in Chinese). DOI: 10.3788/ASQO2013902.0189.
- [7] HE Q X, FENG Q B, LI J K. Long-term stable online acetylene detection by a CEAS system with suppression of cavity length drift[J]. *Sensors*, 2019, **19**(3):508. DOI: <https://doi.org/10.3390/s19030508>.
- [8] TIAN X, CHENG G, CAO Y, *et al.* Simultaneous detection of hydrogen sulfide and carbon dioxide based on off-axis integrated cavity output spectroscopy using a near-infrared distributed feedback diode laser[J]. *Microwave Optical Technology Letters*, 2021, **63**(8):2074–2078. DOI: <https://doi.org/10.1002/mop.32887>.
- [9] 杨雅涵, 李国林, 李小鹏, 等. 基于 TDLAS 技术的天然气中痕量硫化氢分析的 PLS 算法应用[J]. 光子学报, 2017, **46**(2):0230002. DOI: 10.3788/gzxb20174602.0230002.
YANG Y H, LI G L, LI X P, *et al.* Application of PLS algorithm for the analysis of trace hydrogen sulfide in natural gas based on TDLAS technology[J]. *Journal of Photonics*, 2017, **46**(2):0230002. (in Chinese). DOI: 10.3788/gzxb20174602.0230002.
- [10] 李越胜, 卢伟业, 赵静波, 等. 基于 BP 神经网络和激光诱导击穿光谱的燃煤热值快速测量方法研究[J]. 光谱学与光谱分析, 2017, **37**(8):2575–2579. DOI: 10.3964/j.issn.1000-0593(2017)08-2575-05.
LI Y S, LU W Y, ZHAO J B, *et al.* Research on fast measurement method of coal calorific value based on BP neural network and laser-induced breakdown spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2017, **37**(8):2575–2579. (in Chinese). DOI: 10.3964/j.issn.1000-0593(2017)08-2575-05.
- [11] 张强, 魏儒义, 严强强, 等. 深度神经网络在红外光谱定量分析 VOCs 中的应用[J]. 光谱学与光谱分析, 2020, **40**(4):1099–1106. DOI: 10.3964/j.issn.1000-0593(2020)04-1099-08.
ZHANG Q, WEI R Y, YAN Q Q, *et al.* Application of deep neural network in quantitative analysis of VOCs by infrared spectroscopy[J]. *Spectroscopy and Spectral Analysis*, 2020, **40**(4):1099–1106. (in Chinese). DOI: 10.3964/j.issn.1000-0593(2020)04-1099-08.
- [12] 刘丙萍, 李燕, 张琳, 等. 人工神经网络对 VOCs 的自动识别[J]. 光谱学与光谱分析, 2006, **26**(1):51–53.
LIU B P, LI Y, ZHANG L, *et al.* Automatic recognition of VOCs by artificial neural network[J]. *Spectroscopy and Spectral Analysis*, 2006, **26**(1):51–53. (in Chinese).
- [13] AL-ALAWI S M, ABDUL-WAHAB S A, BAKHEIT C S. Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone[J]. *Environmental Modelling and Software*, 2008, **23**(4):396–403. DOI: 10.1016/j.envsoft.2006.08.007.
- [14] 闪霁芳, 刘琨, 江俊峰, 等. 支持向量机在混合气体定量分析中的应用[J]. 光学学报, 2023, **43**(12):1206001. DOI: 10.3788/AOS221681.
SHAN J F, LIU K, JIANG J F, *et al.* Application of support vector machine in quantitative analysis of mixed gases[J]. *Journal of Optics*, 2023, **43**(12):1206001. (in Chinese). DOI: 10.3788/AOS221681.
- [15] ZHANG J, TITTEL F K, GONG L, *et al.* Support vector machine modeling using particle swarm optimization approach for the retrieval of atmospheric ammonia concentrations[J]. *Environmental Modelling and Assessment*, 2016, **21**(4):531–546. DOI: 10.1007/s10666-015-9495-x.
- [16] BAYRAKLI I, EKEN E. A novel breath molecule sensing system based on deep neural network employing multiple-line direct absorption spectroscopy[J]. *Optics & Laser Technology*, 2023, **158**:108918. DOI: 10.1016/j.optlastec.2022.108918.

- [17] WEBBER M E. Diode laser measurements of NH_3 and CO_2 for combustion and bioreactor applications[D]. Stanford: Stanford University, 2001.
- [18] BEER A. Bestimmung der absorption des rothen lichts in farbigen flussigkeiten[J]. *Ann Physik*, 1852, **162**:78–88.
- [19] 李航. 统计学习方法[M]. 北京: 清华大学出版社, 2012.
LI H. Statistical learning methods[M]. Beijing: Tsinghua University Press, 2012. (in Chinese).
- [20] 王巧云, 单鹏. 分子光谱检测及数据处理技术[M]. 北京: 科学出版社, 2019.
WANG Q Y, SHAN P. Molecular spectrum detection and data processing technology[M]. Beijing: Science Press, 2019. (in Chinese).
- [21] 王惠文. 偏最小二乘回归方法及其应用[M]. 北京: 国防工业出版社, 1999.
WANG H W. Partial least squares regression method and application[M]. Beijing: National Defense Industry Press, 1999. (in Chinese).