

DOI: 10.3969/j.issn.1007-5461.2023.03.010

基于IRIV-SA的乙烯FTIR光谱波数优选

张梦思¹, 鞠薇^{2*}, 程志友², 任慧东¹

(1 安徽大学电子信息工程学院, 安徽 合肥 230601;
2 安徽大学互联网学院, 安徽 合肥 230039)

摘要: 利用傅里叶变换红外光谱 (FTIR) 技术能够对有机物的成分及含量进行快速无损测量。作为重要的有机化工原料, 乙烯在塑料、醇类和纤维等大宗化学品的制造中有着广泛的应用, 但同时由于其易挥发性, 乙烯对环境 and 人体有着潜在的危害。为提高 FTIR 技术检测乙烯浓度模型的精度, 综合迭代保留信息变量法 (IRIV) 和模拟退火算法 (SA) 的优点, 提出了改进的 IRIV-SA 红外光谱波数优选算法。该方法在 IRIV 算法稳定选取大量光谱特征波数的基础上, 利用 SA 进一步筛选少量有效特征波数, 从而降低模型复杂度, 提高有机物光谱检测精度。实验首先利用 IRIV-SA 对乙烯红外光谱的浓度进行波数选取, 最终获取的特征波数由全光谱的 271 个变量降低至 5 个变量, 再利用特征波数进行建模, 结果表明其验证集相关系数、均方根误差为 0.9989 和 0.3943, 预测集相关系数、均方根误差为 0.9978 和 0.6652, 较全光谱建模精度有大幅提高。为进一步验证该算法的有效性, 同时建立 IRIV-SA-CARS (自适应重加权采样算法)、SPA (连续投影算法) 以及 IRIV-CARS、IRIV-SPA 波数选取模型对相同数据集进行对比实验, 对比结果表明 IRIV-SA 算法优于上述 6 种波数选取方法, 是一种更有效的特征波数优选方法。

关键词: 傅里叶变换红外光谱; 乙烯; 波数优选; 迭代保留信息变量; 模拟退火算法

中图分类号: O657.33 文献标识码: A 文章编号: 1007-5461(2023)03-00383-09

FTIR spectral wavenumber optimization for ethylene based on IRIV-SA

ZHANG Mengsi¹, JU Wei^{2*}, CHENG Zhiyou², REN Huidong¹

(1 School of Electronic and Information Engineering, Anhui University, Hefei 230601, China;
2 School of Internet, Anhui University, Hefei 230039, China)

Abstract: Fourier transform infrared spectroscopy (FTIR) can rapidly and nondestructively measure the composition and content of organic matter. As an important organic chemical raw material, ethylene is widely used in the manufacturing of bulk chemicals such as plastics, alcohols and fibers. However, due to its volatility, ethylene is harmful to the environment and human body. To improve the accuracy of FTIR in the detection model of ethylene concentration, an improved IRIV-SA infrared spectral wavenumber optimization algorithm is proposed based on the advantages of iteratively retains informative variables method (IRIV) and simulated annealing algorithm (SA). On the basis of stable selection of a large number of spectral characteristic wavenumbers by IRIV algorithm, this method uses SA to further screen a small

基金项目: 国家自然科学基金 (61672032), 安徽省教育厅自然科学基金 (KJ2021A0026)

作者简介: 张梦思 (1995 -), 女, 研究生, 主要从事近红外光谱定性和定量分析方面的研究。E-mail: p20201060@stu.ahu.edu.cn

导师简介: 鞠薇 (1986 -), 女, 博士, 讲师, 硕士生导师, 主要从事光谱信号特征提取及识别、混合信号盲分离、医学图像处理、机器视觉等方面的研究。E-mail: 20029@ahu.edu.cn

程志友 (1972 -), 博士, 教授, 博士生导师, 主要从事电能质量分析、电力负荷检测方面的研究。E-mail: 363870358@qq.com

收稿日期: 2022-04-16; 修改日期: 2022-05-12

*通信作者。

number of effective characteristic wavenumbers, so as to reduce the complexity of the model and improve the detection accuracy of organic matter spectrum. In the experiment, IRIV-SA is used to select the wavenumber of the concentration of ethylene infrared spectrum at first, and the number of characteristic wavenumbers obtained is reduced from 271 to 5, then the characteristic wavenumber is used for modeling. The results show that the correlation coefficient and root mean square error of validation set are 0.9989 and 0.3943, and the correlation coefficient and root mean square error of prediction set are 0.9978 and 0.6652, which indicates that the modeling accuracy of the proposed algorithm is significantly improved compared with that of the whole spectrum modeling. To further verify the effectiveness of the improved algorithm, IRIV, SA, CARS (competitive adaptive reweighted sampling algorithm), SPA (successive projections algorithm), IRIV-CARS and IRIV-SPA wavenumber selection models are established for comparative experiments on the same data set. The comparison results show that IRIV-SA algorithm is superior to the above six wavenumber selection methods, and is an effective feature wavenumber selection method.

Key words: Fourier transform infrared spectroscopy; ethylene; wavenumber selection; iteratively retains informative variables; simulated annealing algorithm

0 引言

乙烯是重要的有机化工原料,被广泛地用于塑料、醇类和纤维等各种大宗化学品的制造中^[1]。目前,全球乙烯年产量已超过 1.8 亿吨^[2],其生产过程中会产生大量二氧化碳和其他副产物,不仅会导致全球变暖,而且对水资源造成严重污染^[3]。同时,乙烯也是挥发性有机物 (VOCs) 的重要组成成分之一,会促进光化学臭氧 (O_3)、二次有机气溶胶的形成^[4,5],影响空气质量,并增加人体患哮喘、肺癌等的概率^[6]。因此,检测大气中乙烯的浓度对环境保护和人体健康都具有深远意义。

检测乙烯浓度的主要方法是气相色谱法,但其存在分析周期长、操作繁琐等问题^[7]。傅里叶变换红外光谱 (FTIR) 技术具有分析速度快、操作简单、稳定强和无损性等特点^[8,9],FTIR 光谱仪能获得几乎所有类型样品的光谱,且 FTIR 技术适用于样品的定性和定量分析^[10],FTIR 技术也可用于气体的多组分测量^[11,12]。在红外光谱分析中,为了不丢失样本细节信息,通常采用全光谱建模。但随着精密测量仪器的快速发展,测出的红外光谱在包含大量信息的同时,也存在全光谱吸收带较宽、部分区域与待测物质的相关性较弱、相邻波数间相关性较强、冗余信息过多等问题,因此对其进行波数变量优选以减少建模所需的光谱信息,从而降低模型复杂度,提高检测效率。

现已有使用特征变量优选算法用于甲烷、乙烷、丙烷、正丁烷、异丁烯、正戊烷、异戊烷和二氧化碳样品浓度方面的检测^[7]。常见的特征变量优选算法有:遗传算法 (GA)^[13]、蚁群算法 (ACO)^[14]、模拟退火算法 (SA)^[15]、迭代保留信息变量 (IRIV)^[16]、连续投影算法 (SPA)^[17]等。全局搜索能力强的 GA 和具有正反馈机制的 ACO 都易陷于局部最优解; SPA 选择的投影长度最大的变量不能确定为有效变量^[18]; SA 运行效率高,在保证局部寻优精度的同时避免陷入局部最优^[19]。2014 年由 Yun 等^[20]提出迭代保留信息变量 (IRIV) 算法,根据重要性将变量分为四类:强信息变量、弱信息变量、无信息变量和干扰变量。将对模型无用的无信息变量和干扰变量剔除,在强信息变量和弱信息变量的基础上,IRIV 执行反向消除策略,进一步消减变量的数

目, 最终保留下来的变量即是最优变量。IRIV因其筛选的变量数少、精度高的特点逐渐被应用于物质的定量分析。Zhang等^[21]利用联合区间偏最小二乘法(SiPLS)筛选特征区间, 接着以GA、竞争自适应重加权采样法(CARS)、IRIV进一步提取特征波长变量, 建立乙醇和还原糖含量的偏最小二乘(PLS)模型, SiPLS-IRIV-PLS模型预测精度均高于其他方法, 选出的特征波长数量也少于其他方法, 表明IRIV算法可以有效地优选特征波长。Cheng等^[22]提出了一种改进的IRIV-SPA算法, 在IRIV选出强、弱信息波长变量的基础上使用SPA算法进一步选择波长, 解决了SPA算法选出的投影长度最大的变量并不一定就是有效变量的不足, 在柴油50%回收率下的沸点、土壤有机质含量的近红外光谱数据集上建立IRIV-SPA-MLR模型并与SPA、IRIV、随机蛙跳(RF)特征波长选择算法的模型相比较, 验证了IRIV-SPA-MLR预测精度最优, 相关系数(R)达到了0.9845(沸点指标)、0.9745(有机质含量指标)。An等^[23]采用变量组合集群分析法(VCPA)联合IRIV变量筛选算法建立红参总皂苷PLS定量分析模型, 模型的相关系数为0.928, 均方根误差(RMSE)为7.99, 精度优于CARS-PLS、RF-PLS、CARS-PLS等模型, VCPA-IRIV在考虑变量组合效应的同时排除了干扰变量。以上IRIV算法及其改进算法提高了红外光谱选取模型的预测精度。

本文在对比以上IRIV算法变量选取模型优缺点的基础上, 提出了改进的IRIV-SA算法。首先研究了IRIV从全FTIR光谱中选出强信息波数变量和弱信息波数变量后不需要进行反向消除波数变量这一特点, 利用模拟退火算法(SA)进行二次波数变量优选, 该算法可以减少IRIV算法筛选得到的变量数目, 最大程度地搜索全局近似最优解, 同时IRIV算法可以弥补SA参数设置不合理造成波数变量冗余的问题。实验建立了FULL-PLS、IRIV-PLS、SA-PLS、CARS-PLS、SPA-PLS、IRIV-CARS-PLS、IRIV-SPA-PLS、IRIV-SA-PLS模型, 输入样本数据并根据相关系数和均方根误差评估模型运算效率。本研究以乙烯浓度的傅里叶变换红外光谱作为样本数据, 比较几种模型的相关系数和均方根误差, 得出IRIV-SA-PLS模型预测精度最优, IRIV-SA算法能有效地筛选出FTIR光谱特征波数, 为大气中乙烯浓度的检测提供了理论支持。

1 样本与评价指标

1.1 样本的制备

乙烯FTIR光谱数据采集系统由配气平台、FTIR光谱仪以及相关配套设备组成, 如图1所示。其中: 初始样本是由南京特种气体有限公司提供的存储于4 L容积钢气瓶中的纯净乙烯气体; 所需浓度乙烯气体的配置使用中国科学院合肥物质科学研究院自主研发的四通道高精度配气平台调整辅助气体氮气和乙烯的体积比来完成, 配气平台的精度为气体初始浓度的0.1%, 误差范围为0.05%~0.1%; 乙烯红外光谱的采集由中国科学院安徽光学精密机械研究所自主研发的FTIR气体光谱分析仪获得。

实验步骤: 首先确定储气瓶所在空间具有良好的通风性, 通过聚四氟乙烯采样管将储气瓶、配气平台以及FTIR光谱仪密接起来; 接着通过储气瓶阀门和配气平台上的压力计来控制乙烯气体的流速, 当气体流速稳定并且分析仪气体取样池内压力值为一个大气压时, 观察光谱仪显示屏上光谱是否有变化, 当光谱数据稳定后, 记录当前的光谱数据和乙烯气体的浓度。一共制备了24种不同浓度的乙烯样本, 浓度范围为75.349~133.116 mg/m³, 每份样本经由系统扫描20次并输出平均光谱, 光谱测量范围为700~5000 cm⁻¹, 采样间隔约为16 cm⁻¹, 全谱段包含271个波数点, 乙烯的FTIR光谱图如图2所示。

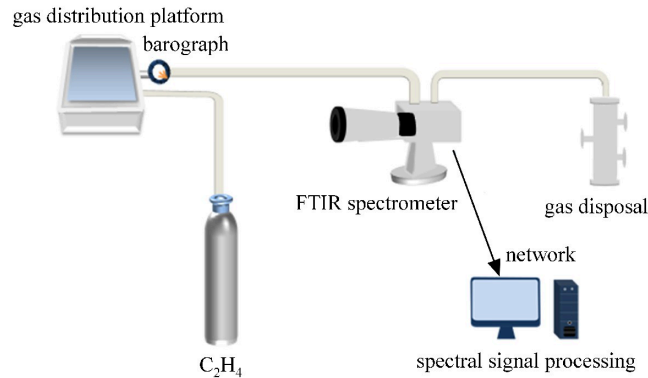


图 1 乙烯 FTIR 光谱采集系统示意图

Fig.1 Schematic diagram of ethylene FTIR spectra acquisition system

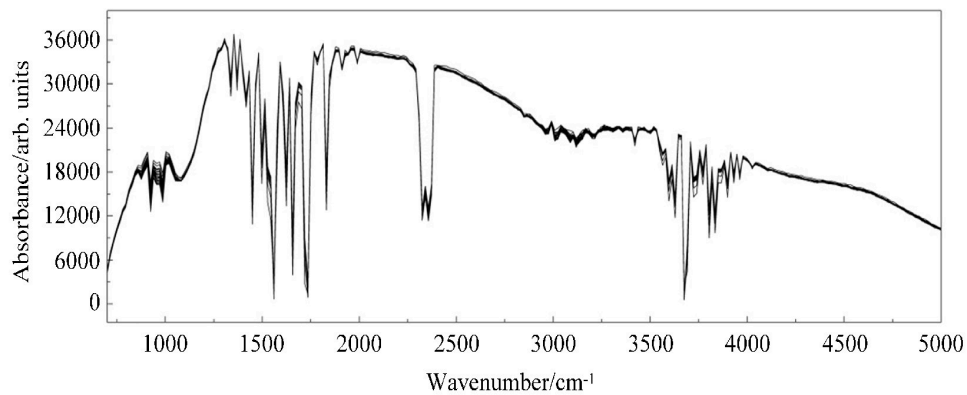


图 2 乙烯 FTIR 光谱图

Fig. 2 FTIR spectra of ethylene

1.2 评价指标

模型评价指标由相关系数 (R) 和均方根误差 (E_{RMS}) 决定, 相关系数越大、均方根误差越小, 模型预测更准确。验证集相关系数、均方根误差分别为 R_{C} 、 $E_{\text{RMS-CV}}$, 预测集相关系数、均方根误差分别为 R_{P} 、 $E_{\text{RMS-P}}$ 。以下实验均在 Python、Origin 中进行数据分析。

2 实验结果

2.1 样本划分

采用 SPXY (Sample set portioning based on joint x - y distance) 算法对样本按照 6:2:2 划分为校正集、验证集、预测集。SPXY 相较于 KS (Kennard-Stone) 样本划分算法在计算样品间的距离时同时考虑了光谱间的距离和因变量值间的距离, 因此 SPXY 划分的样本更具代表性, 且已有研究表明由 SPXY 采样方法得到的训练集建立的模型优于 KS 方法^[24]。表 1 是经 SPXY 算法划分后的样本集, 乙烯样本验证集和预测集的最大值、最小值都在校正集值的区间内, 说明了 SPXY 划分样本的合理性。

2.2 IRIV 波数选择

IRIV 采用二进制矩阵重排过滤器方法获得波数变量的随机组合, 对每个组合进行建模。从一个波数变量开始, 根据该变量存在波数组合中不存在时的交叉验证均方根误差平均值之差和 Mann-Whitney U 检验

的 p 值来决定该变量的去留, 依次重复这个操作直至变量组合中只余强信息变量和弱信息变量^[25]。对强信息变量和弱信息变量执行反向消除 (Reverse elimination) 策略, 最终保留的变量即最优变量。图 3 是 IRIV 保留变量的过程, IRIV 采用最大主成分数为 10、交叉验证次数为 5, 迭代了 5 次, 第 4 次迭代保留了 32 个强信息变量和弱信息变量, 第 5 次反向消除了 8 个变量。

表 1 乙烯样本划分结果

Table 1 Sample partition results of ethylene

	Number of samples	Concentration/ (mg·m ⁻³)			
		Maximum	Minimum	Average	Standard deviation
Calibration set	14	133.116	75.349	101.900	19.473
Validation set	5	115.535	85.395	98.456	11.700
Prediction set	5	130.604	95.442	116.539	13.939

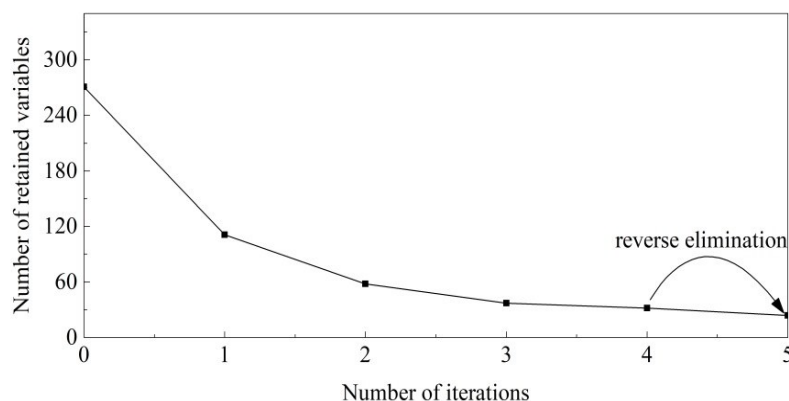


图 3 IRIV 保留变量过程

Fig. 3 IRIV retain variable process

本研究把 IRIV 算法第一步所保留的强、弱信息波数变量 [即未进行反向消除 (No reverse elimination) 的变量] 和进行反向消除后保留的变量分别用于 PLS 建模, 表 2 是 PLS 建模效果。由表 2 中验证集、预测集的相关系数和均方根误差可知, 进行反向消除后保留的变量较少但建模效果不如未进行反向消除的情况, 说明反向消除剔除了一些有效波数变量。实验中与 IRIV 算法相结合的波数选择都是基于 IRIV 选择的强、弱信息波数变量。

表 2 IRIV 的 PLS 建模效果

Table 2 PLS modeling effect of IRIV

	R_c	E_{RMS-CV}	R_p	E_{RMS-P}
No reverse elimination	0.9947	0.8612	0.9937	1.1154
Reverse elimination	0.9915	1.0856	0.9917	1.2792

2.3 IRIV-SA 波数优选

SA 是基于蒙特卡洛思想的求解近似最优化问题的方法, 可以获得全局近似最优解。经 IRIV 选择后的变量数仍较多, 因此使用 SA 对其选取的波数变量进行二次特征波数优选。首先以 IRIV 选取的强、弱信息变量为基础, 分别对 SA 的初始温度和退温速率 α [$\alpha=T(k+1)/T(k)$, 式中 T 为温度, k 为玻尔兹曼常量] 进行分

析, 图4是IRIV-SA算法中初始温度和退温速率的选择过程, 最终选择100 °C的初始温度和0.93的退温速率作为IRIV-SA模型的参数; 接下来, 从强、弱信息变量中随机抽取5个波数的光谱信息变量作为SA初始解, 以IRIV-SA模型的RMSECV作为SA目标函数, SA从100 °C开始降温, 同一温度下迭代300次, 根据Metropolis准则记录当前温度下IRIV-SA模型的最优解和最优目标函数值, 以0.93的退温速率进行降温, 直至温度降低为0.00001 °C, 此时记录的最优变量即为全局最优变量。经IRIV-SA提取的波数量由全光谱271降为5, 选择的波数点为716.1600、923.0506、1002.6240、2164.3946、2578.1759 cm^{-1} , 如图5所示。

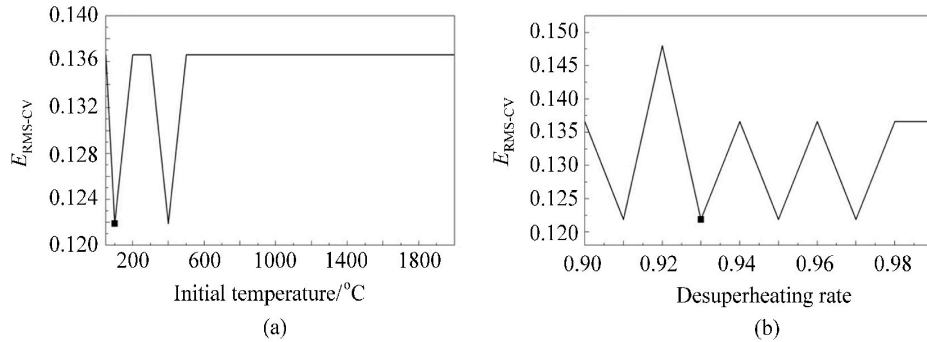


图4 IRIV-SA参数选择过程。(a) 初始温度的选择; (b) 退温速率的选择

Fig.4 IRIV-SA parameter selection process. (a) Selection of initial temperature; (b) Selection of desuperheating rate

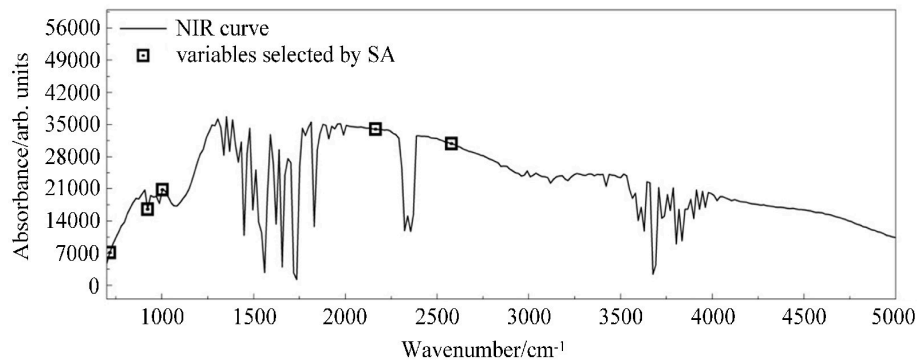


图5 IRIV-SA选择的波数点

Fig.5 The wavenumber points selected by IRIV-SA

2.4 模型预测

为验证IRIV-SA算法的有效性, 本研究使用了对比方法: FULL (全波长)、IRIV、SA、CARS、SPA、IRIV-CARS、IRIV-SPA。其中, CARS基于达尔文进化论的“适者生存”思想, 利用蒙特卡洛 (Monte Carlo) 采样技术和PLS模型的指数衰减函数, 选择交叉验证均方根误差最小值的变量作为最优变量; SPA是一种前向循环特征变量选择方法, 从一个变量开始, 每次循环都计算其在未入选变量上的投影, 提取投影向量最大的变量并引入到变量组合中, 这样每一个新入选的变量都与前一个线性关系最小^[17]。

IRIV-SA、FULL、IRIV、SA、CARS、SPA、IRIV-CARS、IRIV-SPA结合PLS建模的乙烯浓度预测结果如表3所示。IRIV-SPA是几种算法中模型获取特征波数最少的, 波数量为2, 但其特征波数建模效果不如IRIV-SA; IRIV-SA算法建模效果最优, 验证集的相关系数为0.9989, 预测集的相关系数为0.9978, 保留的变量数也从271减少到了5, 说明该方法可以准确地筛选与乙烯浓度相关性强的波数同时降低模型复杂度。图6是不同建模方法下乙烯浓度预测集实测值和预测值的散点图, 更为直观地表明了IRIV-SA算法最优。

表 3 乙烯浓度 PLS 建模结果

Table 3 PLS modeling results of ethylene concentration

Method	Selected wavenumber points	R_C	E_{RMS-CV}	R_P	E_{RMS-P}
FULL	271	0.9776	1.7593	0.9564	2.9072
IRIV	32	0.9947	0.8612	0.9937	1.1154
CARS	24	0.9918	1.0704	0.9899	1.4107
SPA	12	0.9923	1.0360	0.9754	2.1928
SA	14	0.9950	0.8352	0.9918	1.2719
IRIV-CARS	17	0.9948	0.8530	0.9937	1.1120
IRIV-SPA	2	0.9979	0.5389	0.9956	0.9183
IRIV-SA	5	0.9989	0.3943	0.9978	0.6652

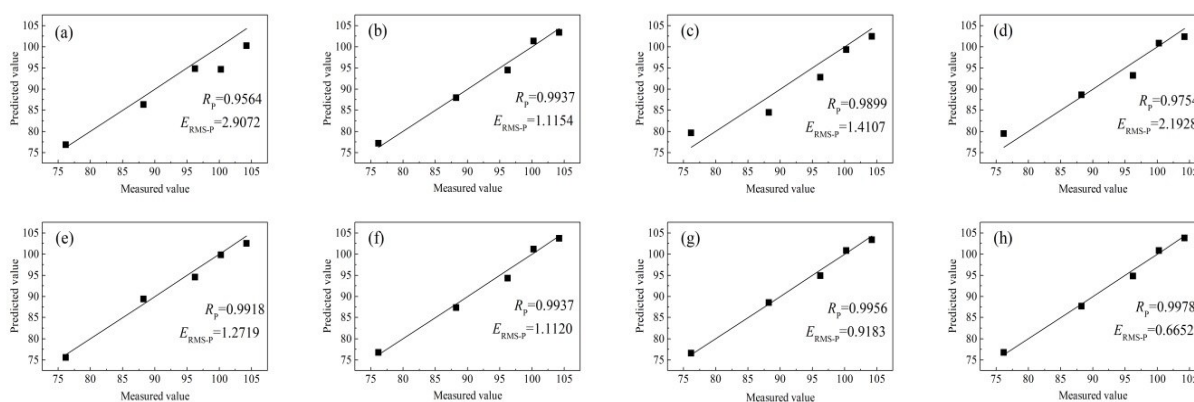


图 6 乙烯浓度预测集定量模型散点图。(a) FULL; (b) IRIV; (c) CARS; (d) SPA; (e) SA; (f) IRIV-CARS; (g) IRIV-SPA; (h) IRIV-SA

Fig.6 Scatter plots for ethylene concentration prediction set quantitative model. (a) FULL; (b) IRIV; (c) CARS; (d) SPA; (e) SA; (f) IRIV-CARS; (g) IRIV-SPA; (h) IRIV-SA

3 总结

提出了一种用于乙烯 FTIR 光谱波数优选的改进的 IRIV-SA 算法, 在 IRIV 选择的与乙烯浓度相关的强、弱信息波数变量基础上, 利用 SA 方法进行特征波数提取, 剔除光谱中冗余的波数信息, 精简经 IRIV 提取的变量数目。实验结果表明, 与文献 [22] 提出的 IRIV-SPA 算法相比, IRIV-SA 建模精度更高; 与其他波数优选算法 (IRIV, SA, CARS, SPA, IRIV-CARS) 相比, IRIV-SA 建模效果最佳。总之, IRIV-SA 可以有效提取特征表征能力强的 FTIR 光谱吸光度变量, 进而快速地预测乙烯气体浓度, 对改善空气质量、提高人体健康度有重要意义。

参考文献:

- [1] Wang Z P, Shi R, Zhang T R. Three-phase electrochemistry for green ethylene production [J]. *Current Opinion in Electrochemistry*, 2021, 30: 100789.
- [2] He C H, Wang Y, Chen Y, *et al.* An ethane-favored metal-organic framework with tailored pore environment used for efficient ethylene separation [J]. *Microporous and Mesoporous Materials*, 2021, 320: 111096.

- [3] Ren T, Patel M K, Blok K. Steam cracking and methane to olefins: Energy use, CO₂ emissions and production costs [J]. *Energy*, 2008, 33(5): 817-833.
- [4] Niu H, Mo Z W, Shao M, *et al.* Screening the emission sources of volatile organic compounds (VOCs) in China by multi-effects evaluation [J]. *Frontiers of Environmental Science & Engineering*, 2016, 10(5): 25-35.
- [5] Wei W, Ren Y T, Yang G, *et al.* Characteristics and source apportionment of atmospheric volatile organic compounds in Beijing, China [J]. *Environmental Monitoring and Assessment*, 2019, 191(12): 762.
- [6] Akhter F, Siddiquei H R, Mukhopadhyay S C. A graphene-based composite for selective detection of ethylene at ambient environment [C]. 2021 IEEE Sensors, 2021: 1-4.
- [7] Li Z B, Pang W, Liang H B, *et al.* A novel method to realize multicomponent infrared spectroscopy gas logging based on PSO-split peak fitting-SVM [J]. *Journal of Natural Gas Science and Engineering*, 2022, 101(7): 104496.
- [8] Radica F, Della V G, Malfatti L, *et al.* Real-time quantitative detection of styrene in atmosphere in presence of other volatile-organic compounds using a portable device [J]. *Talanta*, 2021, 233: 122510.
- [9] Barra I, Kharbach M, Qannari E M, *et al.* Predicting cetane number in diesel fuels using FTIR spectroscopy and PLS regression [J]. *Vibrational Spectroscopy*, 2020, 111: 103157.
- [10] Fehelbom K M, Saleh A, Al-Tabakha M M A, *et al.* Recent applications of quantitative analytical FTIR spectroscopy in pharmaceutical, biomedical, and clinical fields: A brief review [J]. *Reviews in Analytical Chemistry*, 2022, 41(1): 21-33.
- [11] Qu L G, Liu J G, Xu L, *et al.* Vehicle exhaust detection method based on portable FTIR [J]. *Spectroscopy and Spectral Analysis*, 2021, 41(6): 1751-1757.
曲立国, 刘建国, 徐亮, 等. 便携式FTIR的机动车尾气检测方法 [J]. *光谱学与光谱分析*, 2021, 41(6): 1751-1757.
- [12] Gibier M, Girods P, Rogaume Y. Development of an original and easy method for isocyanic acid (HNCO) calibration on FTIR Spectrometer [J]. *Vibrational Spectroscopy*, 2021, 116: 103290.
- [13] Zhang J, Yuan S, Zhang J. Detection of sugar degree in strawberry based on wavelet transform-genetic algorithm-partial least squares [J]. *Journal of Analytical Science*, 2020, 36(1): 111-115.
张娟, 原帅, 张骏. 基于小波变换-遗传算法-偏最小二乘的草莓糖度检测研究 [J]. *分析科学学报*, 2020, 36(1): 111-115.
- [14] Deng H R, Lü C W, Chen D L, *et al.* Application of ant colony optimization in hyperspectral prediction modeling of soil texture [J]. *Chinese Journal of Soil Science*, 2021, 52(5): 1063-1068.
邓浩然, 吕成文, 陈东来, 等. 蚁群算法在土壤质地高光谱预测建模中的应用 [J]. *土壤通报*, 2021, 52(5): 1063-1068.
- [15] Xu L, Zhu W H, Yao H B, *et al.* Prediction of acetic acid concentration in Chinese liquors based on fluorescence spectrum and simulated annealing algorithm [J]. *Spectroscopy and Spectral Analysis*, 2021, 41(7): 2159-2165.
许蕾, 朱卫华, 姚红兵, 等. 基于荧光光谱-模拟退火法年份白酒中乙酸浓度预测研究 [J]. *光谱学与光谱分析*, 2021, 41(7): 2159-2165.
- [16] Huang J S, Wang D X, Xiong A H, *et al.* A hybrid variable selection strategy of near infrared spectroscopy for detection the ratio of tea polyphenols to amino acids in green tea infusion [J]. *Acta Agriculturae Universitatis Jiangxiensis*, 2020, 42(6): 1270-1276.
黄俊仕, 王冬欣, 熊爱华, 等. 基于混合变量选择的绿茶酚氨比近红外光谱检测方法 [J]. *江西农业大学学报*, 2020, 42(6): 1270-1276.
- [17] Zhao J Y, Xiong Z X, Ning J M, *et al.* Wavelet transform combined with SPA to optimize the near-infrared analysis model of caffeine in tea [J]. *Journal of Analytical Science*, 2021, 37(5): 611-617.
赵静远, 熊智新, 宁井铭, 等. 小波变换结合连续投影算法优化茶叶中咖啡碱的近红外分析模型 [J]. *分析科学学报*, 2021, 37(5): 611-617.

- [18] Li C T, Zhang Y Y, Chen H Z, *et al.* Near-infrared spectral waveband selection for soil potassium content based on simulated annealing [J]. *Laser & Optoelectronics Progress*, 2022, 59(13): 1330002.
李春婷, 张优优, 陈华舟, 等. 基于模拟退火的土壤钾含量近红外光谱波段优选 [J]. 激光与光电子学进展, 2022, 59(13): 1330002.
- [19] Sun X R, Zhou Z J, Liu C L, *et al.* Near infrared spectroscopic detection of gluten content in wheat flour based on spectral pretreatment and simulated annealing algorithm [J]. *Food Science*, 2018, 39(2): 222-226.
孙晓荣, 周子健, 刘翠玲, 等. 光谱预处理结合模拟退火算法的小麦粉面筋含量检测 [J]. 食品科学, 2018, 39(2): 222-226.
- [20] Yun Y H, Wang W T, Tan M L, *et al.* A strategy that iteratively retains informative variables for selecting optimal variable subset in multivariate calibration [J]. *Analytica Chimica Acta*, 2014, 807(1): 36-43.
- [21] Zhang H, Liu G H, Jiang H, *et al.* Quantitative detection of ethanol solid-state fermentation process parameters based on near infrared spectroscopy [J]. *Laser & Optoelectronics Progress*, 2017, 54(2): 320-326.
张航, 刘国海, 江辉, 等. 基于近红外光谱技术的乙醇固态发酵过程参数定量检测 [J]. 激光与光电子学进展, 2017, 54(2): 320-326.
- [22] Cheng J H, Chen Z G. Wavelength selection method for near infrared spectroscopy based on iteratively retains informative variables and successive projections algorithm [J]. *Chinese Journal of Analytical Chemistry*, 2021, 49(8): 1402-1409.
程介虹, 陈争光. 基于迭代保留信息变量和连续投影的近红外光谱波长选择方法 [J]. 分析化学, 2021, 49(8): 1402-1409.
- [23] An S Y, Zhang L, Shang X Z, *et al.* Variable selection method in the NIR quantitative analysis model of total saponins in red ginseng extract [J]. *Spectroscopy and Spectral Analysis*, 2021, 41(1): 206-209.
安思宇, 张磊, 尚献召, 等. 红参提取物总皂苷近红外定量分析建模中的变量筛选 [J]. 光谱学与光谱分析, 2021, 41(1): 206-209.
- [24] Zhan X R, Zhu X R, Shi X Y, *et al.* Determination of hesperidin in tangerine leaf by near-infrared spectroscopy with SPXY algorithm for sample subset partitioning and Monte Carlo cross validation [J]. *Spectroscopy and Spectral Analysis*, 2009, 29(4): 964-968.
展晓日, 朱向荣, 史新元, 等. SPXY样本划分法及蒙特卡罗交叉验证结合近红外光谱用于橘叶中橙皮苷的含量测定 [J]. 光谱学与光谱分析, 2009, 29(4): 964-968.
- [25] Ji R S, Chen X Y, Liu S Z, *et al.* Nondestructive testing of volatile oil of zanthoxylum bungeanum based on hyperspectral technique and IRIV-FOA-ELM algorithm [J]. *Laser & Optoelectronics Progress*, 2020, 57(20): 394-400.
纪然仕, 陈晓燕, 刘素珍, 等. 基于高光谱技术和IRIV-FOA-ELM算法的花椒挥发油无损检测 [J]. 激光与光电子学进展, 2020, 57(20): 394-400.