

近红外光谱结合 BP 神经网络法快速分析薰衣草精油中主要成分

唐军¹, 廖享¹, 韩凯乐²

(1. 新疆大学理化测试中心, 新疆乌鲁木齐 830046; 2. 新疆伊犁紫苏丽人生物科技有限公司, 新疆伊宁 835000)

摘要 [目的]建立薰衣草精油主体成分的快速测定模型,以满足中国薰衣草精油质量标准体系中对化学成分分析的要求。[方法]采用近红外光谱方法结合 BP 神经网络快速测定薰衣草精油中芳樟醇和乙酸芳樟酯的含量,在 7 000~4 000 cm^{-1} 波数范围内,薰衣草精油化学信息量比较丰富且噪音低,可作为分析区间。通过间隔偏最小二乘方法(iPLS)选择与测定指标相关性高的近红外波长段集合,分别得到乙酸芳樟酯和芳樟醇的 185 个定量分析用波长点数据集,对 BP 神经网络参数优化后,建立新疆薰衣草精油中芳樟醇和乙酸芳樟酯的 iPLS-BP 神经网络快速近红外定量分析模型。[结果]采用方法的芳樟醇和乙酸芳樟酯测定系数分别为 0.972 91 和 0.946 90,预测结果与 GC-MS 分析的相对误差分别小于 2.45% 和 2.85%。[结论]定量分析模型能够很好地快速测定薰衣草精油中乙酸芳樟酯、芳樟醇含量,具有良好的预测能力,可为新疆薰衣草精油主要成分的快速定量分析提供一种新的有效方法。

关键词 薰衣草精油;傅利叶近红外光谱;BP 神经网络

中图分类号 S 573⁺.9 文献标识码 A

文章编号 0517-6611(2019)03-0189-04

doi:10.3969/j.issn.0517-6611.2019.03.059



开放科学(资源服务)标识码(OSID):

Rapid Analysis of Main Components in Lavender Essential Oil by Near Infrared Spectroscopy Combined with BP Neural Network
TANG Jun¹, LIAO Xiang¹, HAN Kai-le² (1. Physics and Chemistry Detect Center, Xinjiang University, Urumqi, Xinjiang 830046; 2. Perilla beauty Biotechnology Co., Ltd., Yining, Xinjiang 835000)

Abstract [Objective] The research aimed to establish a rapid determination model of lavender essential oil body composition, it can meet the requirements of chemical components in the quality standard system of Chinese lavender essential oil. [Method] We can quickly determine the content of linalool and linalyl acetate in the essential oil of Xinjiang lavender by using near infrared spectra combining BP neural network. On the process of this study measured the lavender essential oil samples of near infrared absorption spectrum, by analysis of near infrared spectral absorption peak by the number of wavelengths, 7 000~4 000 cm^{-1} within the scope of lavender essential oil chemical information is abundant in Xinjiang and the wavelength section of low noise, so choose the wave number range for analysis. We obtained 185 data sets of wavelength points for quantitative analysis of linalool and linalyl acetate respectively, through interval partial least squares method (iPLS) selection high correlation of the near infrared wavelength section. After optimizing the parameters of BP neural network, the model of fast near infrared quantitative by iPLS-BP neural network method is established. [Result] The measurement coefficient of linalool and linalyl acetate by iPLS-ANN method was 0.972 91 and 0.946 90 respectively, and the relative error between the prediction and GC-MS analysis was less than 2.45% and 2.85%. [Conclusion] The quantitative analysis model can quickly determine the content of linalool and linalyl acetate in lavender essential oil, and has good predictive ability, which can provide a new effective method for the rapid quantitative analysis of main components of lavender essential oil in Xinjiang.

Key words Lavender oil; FTNIR; BP-ANN

薰衣草(Lavender)属唇形科,是一种广泛种植的香料植物,主要用于提取香料和观赏,其中提取的薰衣草精油含有多种天然香气和生物活性组分,具有抑菌、助睡眠、安神和帮助消化等功效^[1],在医药、化妆品及食品添加剂方面有着广泛应用^[2]。依据现行薰衣草精油的国家标准^[3],其化学组分以芳樟醇、乙酸芳樟酯、乙酸薰衣草酯、樟脑 4 种成分含量为准入指标,测定方法主要通过气相色谱法、气相色谱质谱联用法等,试验测试人员需要经过专业系统培训才能操作,测定过程时间相对比较长,不适宜现场实时快速检测。而随着计算机辅助的化学计量学快速发展,近红外光谱定量分析技术得到了迅速普及和推广,其方法具有无化学试剂污染、不需要处理样品、快速实时在线测定等优点,目前已经在农业、食品、石油及医药等行业有着广泛的应用^[4-7],并逐渐出台颁布实施了多项国家标准,所研究的薰衣草精油中,所有的组分结构都富含由氢组成官能团,非常适合用近红外光谱来进行定量分析测试。笔者采用间隔偏最小二乘(iPLS)法^[8-9]选

择精油特征成分的敏感近红外光谱波长段,再通过 BP 神经网络^[10-12]方法建立近红外的薰衣草精油中芳樟醇、乙酸芳樟酯定量分析模型。

1 材料与方法

1.1 仪器与材料 BRUKER VERTEX 70 型傅利叶近红外光谱仪,配备铟镓砷(InGaAs)检测器,OPUS 6.5 光谱采集和分析软件;岛津(Shimadzu)QP2010 气质联用仪,测试条件:RTX-50MS 石英毛细色谱柱(30 m×0.32 mm×1.0 μm)、载气为 99.99% He 气,流速 1.0 mL/min, EI 离子源;电子能量 70 eV;传输线温度 250 $^{\circ}\text{C}$;离子源温度 200 $^{\circ}\text{C}$ 。

试验所用的薰衣草采自伊犁地区的 65 团、69 团和 70 团,品种分别为 C-197(2)、法国蓝和 H-701。通过水蒸气蒸馏法提取薰衣草精油:称取晾干薰衣草花 20~50 g 于 1 000 mL 烧瓶中,以 1:12(g:mL)的料液比加入蒸馏水,提取时间 4 h,得到有香气薰衣草精油,置于 -20 $^{\circ}\text{C}$ 冷藏备用。

1.2 化学计量学建模 先对近红外光谱进行多元散射(MSC)校正、S-G 方法平滑、光谱数据的标准化处理后建立分析用近红外吸收度数据矩阵。采用间隔偏最小二乘(iPLS)方法进行波长段选择,最后通过 BP 神经网络法(BP-ANN)建立选择后的特征波长段薰衣草精油中芳樟醇、乙酸

基金项目 新疆自治区重点研发资助项目(2017B03016)。

作者简介 唐军(1971—),男,四川绵阳人,高级工程师,硕士,从事分析化学研究。

收稿日期 2018-08-15

芳樟酯含量的多元校正模型。所在数据处理都在 Matlab 8.0 版软件环境中进行,其中 iPLS 采用 Lars Nørgaard 的 iToolbox 软件包完成,iPLS 波长段选择采用交叉验证均方根差(RMSECV)进行评价。神经网络(ANN)定量分析模型采用 Matlab 自带的 nntool 工具箱完成,BP-ANN 模型质量以均方差(MSE)和测定系数 R 为评价指标指标。

2 结果与分析

2.1 薰衣草精油芳樟醇和乙酸芳樟酯的测定 依据目前薰衣草精油的国家标准 GB 1886.38—2015^[3],采用 GC-MS 对薰衣草精油的化学成分进行分析,每份样品平行测定 3 次,最后以国标方法中的峰面积归一化进行定量分析,计算芳樟醇和乙酸芳樟酯的相对百分含量。薰衣草精油样品的芳樟醇和乙酸芳樟酯分布如表 1 所示,从所有样品的平均值和相对标准偏差(RSD)来看,样本点分布范围合理,基本包含了国标要求的分析范围。

表 1 薰衣草精油样品化学成分分布

化合物 Compound	浓度 Concentration	平均值 Average	RSD
芳樟醇 Linalool	21.13~44.07	28.947 4	0.137 7
乙酸芳樟酯 Linalyl acetate	10.82~26.83	21.626 1	0.161 4

2.2 近红外谱数据采集 采用透射模式,选择光程为 1 mm 的石英比色皿,光谱测定范围 4 000~10 000 cm^{-1} ;分辨率 8 cm^{-1} ;扫描累加次数 32 次;并在相同的试验条件、温度、湿度下以空比色皿池进行背景扫描,然后对每份样品平行测定 3 次,取平均值,采集样品的近红外光谱图,最终选择吸收峰信息丰富的 7 000~4 000 cm^{-1} 波数区间的吸光度作为分析用原始数据集,所有样品分析区间近红外光谱叠加图如图 1 所示。在薰衣草精油的近红外光谱中,没有出现饱和和吸收现象,说明选择的比色皿厚度合适,全光谱可进行后期分析用,在波数段 4 000~4 500 cm^{-1} 为合频吸收带;4 680 cm^{-1} 处为烯烃类化合物的组合频吸收峰;5 260 cm^{-1} 处为碳基的二级倍频峰;5 800 cm^{-1} 附近为 C—H 的对称与反对称组合频振动和一级倍频吸收带;6 800 cm^{-1} 附近处的宽峰为键合的羟基一级倍频峰^[13]。

2.3 iPLS 选择光谱波长段集合 iPLS 波长选择以偏最小二乘(PLS)模型的交互验证均方根误差(RMSECV)作为衡量评价指标,当子区间的 RMSECV 小于全光谱的 RMSECV 时,就认为是包含与测定指标相关信息最高的区间,以此最终寻

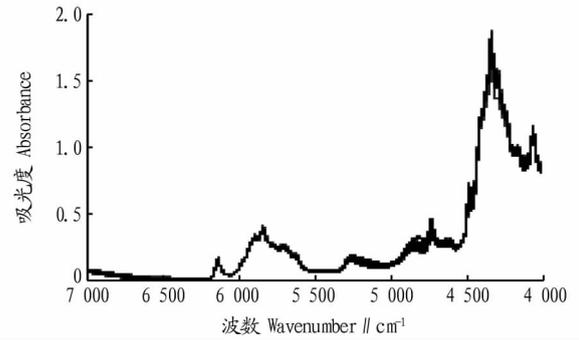


图 1 近红外光谱图

Fig. 1 Near infrared spectrogram

找到所有最优子区间的光谱段,组合成神经网络预测模型的数据子集。首先对原始光谱先进行多元散射校正、S-G 平滑和中心化预处理后,将 7 000~4 000 cm^{-1} 光谱区间划分为等间隔的 20 个子区间,每个子区间包含有 37 个波数点的吸光度。采用 PLS 方法优化选择后确定最佳主成分数为 6, iPLS 波长选择以 PLS 模型的 RMSECV 作为衡量评价指标,最终寻找到预测能力最优的数据子集。自动选取最佳建模子区间,其中芳樟醇和乙酸芳樟酯优化波长子区间结果如图 2 所示,从图中可直观地得到较优的区间,从而可剔除了一些与浓度向量相关性小的吸光值变量,提高模型的稳健性和预测能力。选取 iPLS 方法中 RMSECV 值小于虚横线值的所有波长段区间作为相应薰衣草精油组分的敏感分析数据集。采用间隔偏最小二乘法,优选出芳樟醇和乙酸芳樟酯光谱子区间和波数段如表 2 所示。

2.4 人工神经网络法(ANN)模型的建立与评价 ANN 适合处理原因和结果关系不确定的非线性测量数据,模拟神经元的工作原理来建立模型,具有抗噪声、抗干扰和强大的非线性转换能力等特点。采用 160 个样本用于建立 iPLS-BP 神经网络分析模型,将 iPLS 筛选的光谱子区间吸光度为 BP 神经网络的 X 输入节点变量,GC-MS 测定值作为输出响应 Y 变量,在 Matlab 环境下,使用 nntool 神经网络工具箱,建立芳樟醇、乙酸芳樟酯神经网络定量模型,输入节点都为 185,输出节点都为 1,通过多次比较,确定最优 iPLS-BP 神经网络校正模型参数如表 3 所示。在 Matlab 神经网络工具箱直接把数据划分成 3 个集合,分别为 training(训练)、validation(验证)、test(测试),只有 training 数据参加训练,其他两部分数据不参加训练,只用于检验。在 Matlab 的 nntool 工具箱中,网络模型以均方差(MSE)和测定系数(R)为评价指标指标,从表 3 的 MSE 值和图 3、4 中的回归曲

表 2 iPLS 薰衣草精油芳樟醇和乙酸芳樟酯光谱范围优化结果
Table 2 iPLS optimization results of spectral ranges of linalool and linalyl acetate

化合物 Compound	全波数 RMSECV	Ipls 优选区间号 Ipls preferred interval number	区间对应的波数段 Wavelength segment corresponding to the interval/ cm^{-1}	优选波数点 Preferred wavelength points
芳樟醇 Linalool	0.546 9	6、8、11、14、18	6 245~6 098、5 944~5 797、5 493~5 346、5 041~4 895、 4 440~4 293	185
乙酸芳樟酯 Linalyl acetate	0.114 2	2、6、12、16、18	6 847~6 700、6 245~6 098、5 342~5 196、4 740~4 592、 4 440~4 293	185

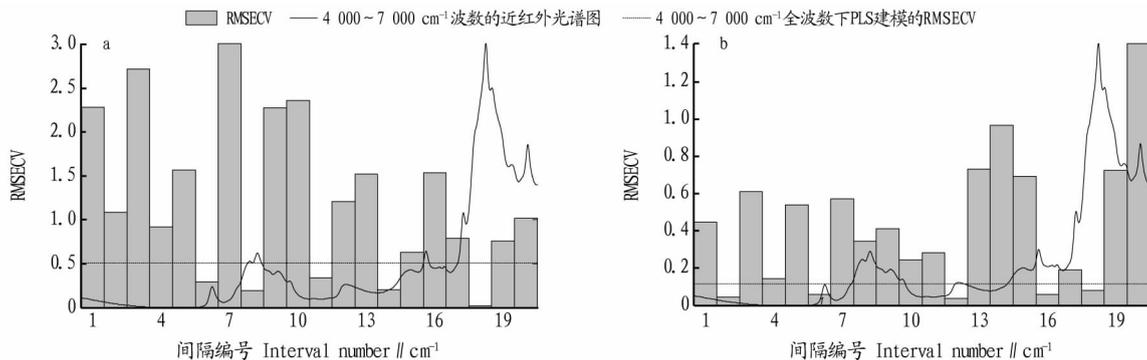


图 2 芳樟醇 (a) 和乙酸芳樟酯 (b) iPLS 优化

Fig. 2 iPLS optimization of linalool (a) and linalyl acetate (b)

线可以看出,其回归曲线的测定系数 R 均大于 0.91。芳樟醇、乙酸芳樟酯的 MSE 值较小,建立的模型具有良好的预测

能力,iPLS-BP 人工神经网络法建立的薰衣草精油代表性和特征性组分定量分析模型效果良好。

表 3 iPLS-BP 最优定量校正模型参数

Table 3 Optimal parameters of iPLS-BP quantitative calibration model

化合物 Compound	iPLS-BP-ANN 网络参数 iPLS-BP-ANN network parameters				逼近误差 Approximation error	
	隐含层神经元个数 Number of neurons in the hidden layer	训练函数 Training function	隐含层传递函数 Implicit layer transfer function	输出层传递函数 Output layer transfer function	MSE	迭代次数 Number of iterations
芳樟醇 Linalool	10	TRAINLM	TANSIG	PURELIN	0.897 7	33
乙酸芳樟酯 Linalyl acetate	13	TRAINSCG	TANSIG	PURELIN	1.156 6	12

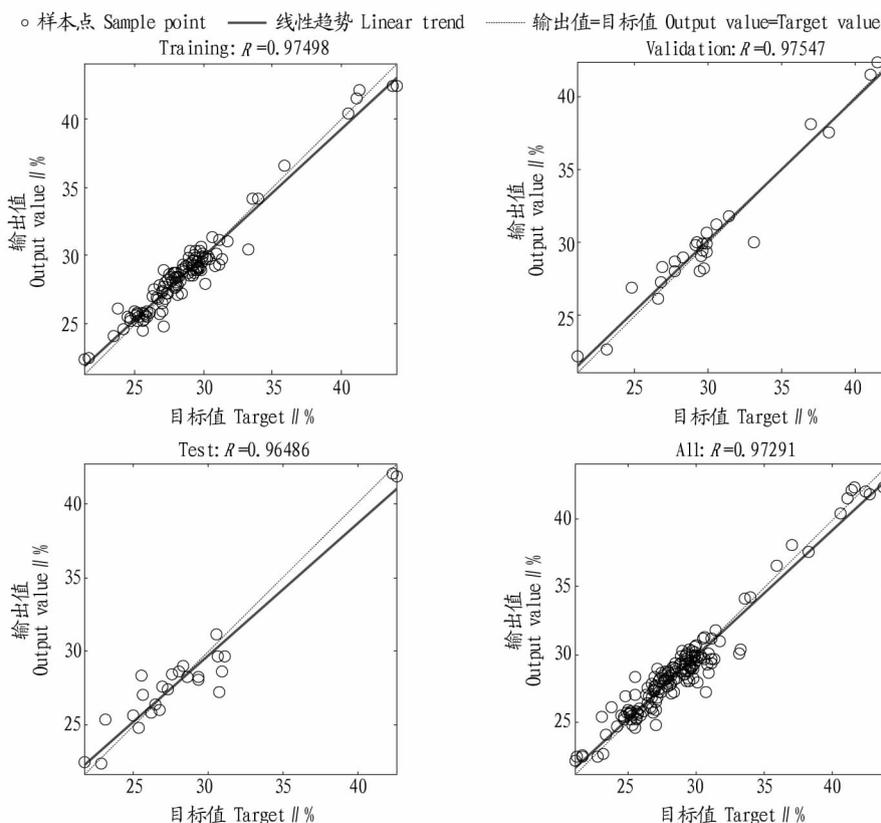


图 3 芳樟醇 BP 人工神经网络建模回归

Fig. 3 BP artificial neural network modeling regression of linalool

2.5 iPLS-ANN 人工神经网络模型预测 为进一步评价 iPLS-BP 神经网络的预测能力,采用建立的 BP 神经网络模

型对 16 组实际薰衣草精油样品中的芳樟醇和乙酸芳樟酯含量进行预测,对实际样品采用相同的前处理过程后,输入 185

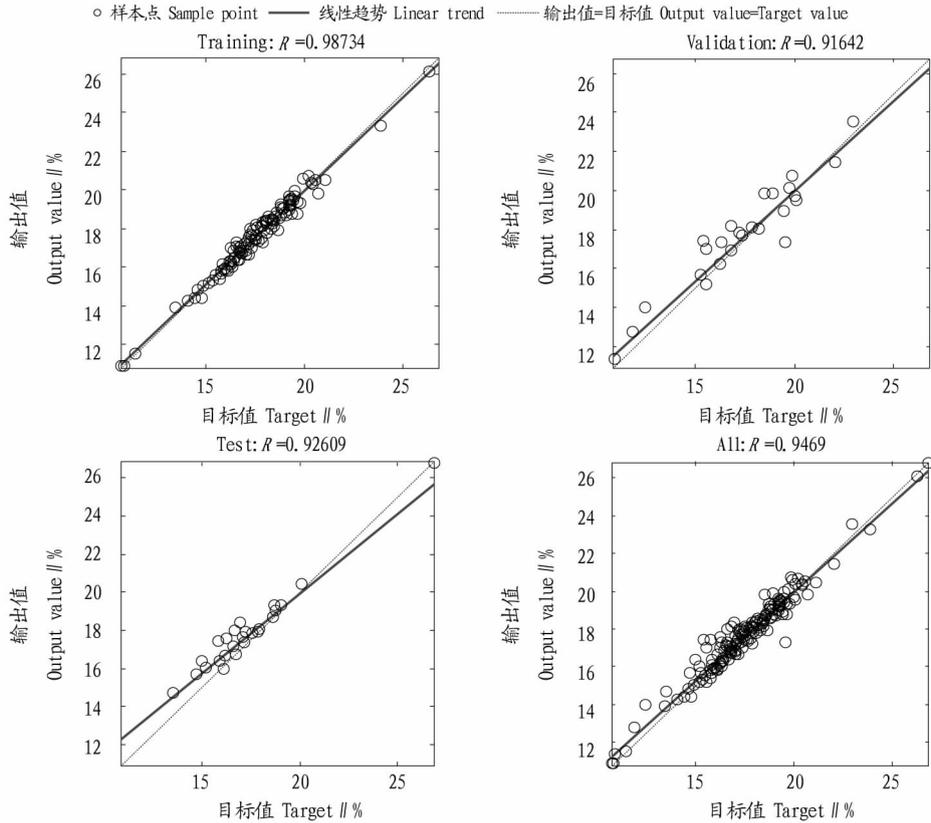


图 4 乙酸芳樟酯 BP 神经网络建模回归

Fig. 4 BP artificial neural network modeling regression of linalyl acetate

个特征波长点的近红外吸光值即可得到薰衣草精中 2 个组分的含量,操作简单。测定值和相对误差分析如表 4 所示,从预测结果来看,乙酸芳樟酯的相对误差在 2.45% 以内,芳樟醇的相对误差在 2.85% 以内,均满足实际分析要求,表明

采用 iPLS 选择出所有敏感波长段,再结合 BP 神经网络对薰衣草精油中的芳樟醇和乙酸芳樟酯预测有着良好的准确度,能够实现近红外方法快速测定薰衣草中芳樟醇和乙酸芳樟酯的要求,可满足实际分析需要。

表 4 iPLS-BP 模型的芳樟醇和乙酸芳樟酯的预测

Table 4 Prediction of linalool and linalyl acetate in the iPLS-BP model

样本 Sample	乙酸芳樟酯 Linalyl acetate			芳樟醇 Linalool		
	GC-MS 测定值 GC-MS measurement	NIR 预测值 NIR predictive value	相对误差 Relative error	GC-MS 测定值 GC-MS measurement	NIR 预测值 NIR predictive value	相对误差 Relative error
1	15.46	15.36	0.58	21.13	20.92	0.96
2	16.82	16.72	0.56	22.84	23.49	-2.84
3	17.77	17.94	-1.00	24.73	24.88	-0.63
5	18.09	17.77	1.74	24.94	25.07	-0.55
6	18.85	18.56	1.52	25.24	25.39	-0.59
7	19.29	19.76	-2.44	25.61	25.89	-1.11
8	20.02	19.96	0.29	25.89	26.31	-1.65
9	21.55	21.32	1.06	25.93	25.57	1.36
10	21.68	21.77	-0.44	26.45	26.32	0.45
11	22.02	21.85	0.74	27.19	27.33	-0.54
12	22.15	22.36	-0.97	27.81	28.02	-0.78
13	22.61	22.85	-1.09	28.35	28.06	0.99
14	23.79	23.57	0.89	29.25	29.16	0.28
15	24.36	24.69	-1.37	30.33	30.51	-0.59
16	25.57	25.67	-0.41	33.57	34.06	-1.48

和精密度均符合痕量分析的要求,能够满足实验室对进出口水果有机磷农残的检测需求。

2.4 水果样品检测 应用该试验所建立的处理方法和条件,对实验室常检样品葡萄、石榴、苹果、梨和桔子中 6 种有机磷进行检测,结果显示,葡萄、苹果和梨中毒死蜱有检出,毒死蜱检测结果均在食品中农药最大残留限量标准允许范围,其他 5 种农药残留无检出,石榴和桔子中均无这 6 种有机磷农药残留检出。

表 2 方法的加标回收率和精密度

Table 2 Calibration recovery and precision of the method

农药 Pesticides	添加水平 Adding level mg/kg	平均测定值 Average measured value//mg/kg	平均回收率 Average recovery rate//%	RSD %
敌敌畏	0.05	0.040 7	81.4	2.94
DDVP	0.10	0.105 0	105.0	2.23
	0.20	0.198 0	99.0	2.56
乐果	0.05	0.035 6	71.2	3.45
Dimethoate	0.10	0.081 0	81.0	2.43
	0.20	0.209 0	104.5	2.35
甲基对硫磷	0.05	0.041 5	83.0	2.38
Parathion-methyl	0.10	0.090 6	90.6	3.07
	0.20	0.210 0	105.0	1.66
毒死蜱	0.05	0.043 6	87.2	3.45
Chlorpyrifos	0.10	0.095 1	95.1	2.83
	0.20	0.202 0	101.0	2.94
马拉硫磷	0.05	0.039 5	79.0	5.23
Malathion	0.10	0.088 2	88.2	4.21
	0.20	0.203 0	101.5	2.18
杀扑磷	0.05	0.036 6	73.2	4.32
Methidathion	0.10	0.084 3	84.3	2.69
	0.20	0.216 0	108.0	2.89

(上接第 192 页)

3 结论

相对于薰衣草精油的传统气相和气质分析方法,红外光谱结合现代计量学方法能快速、实时和准确地测定精油中主要成分的百分比含量,该研究先以 iPLS 方法对近红外全波长吸光度进行筛选,选择出所有与测定指标相关性高的近红外波段作为 BP 神经网络模型的输入节点值,对 BP 神经网络参数优化后,建立新疆薰衣草精油中芳樟醇和乙酸芳樟酯的 iPLS-BP 神经网络快速近红外定量分析模型。结果表明,采用 iPLS-BP 神经网络建立定量模型的测定系数分析均大于 0.91,预测结果与 GC-MS 分析的相对误差小于 3%。定量分析模型能够很好地快速测定薰衣草精油中乙酸芳樟酯、芳樟醇含量,具有良好的预测能力,可为新疆薰衣草精油主要成分的快速定量分析提供一种新的有效方法。

参考文献

- [1] CAVANAGH H M A, WILKINSON J M. Biological activities of lavender essential oil [J]. *Phytother Res*, 2002, 16(4): 301-308.
- [2] 徐洁华, 文首文. 薰衣草挥发性有机物及其药理功效研究进展 [J]. *时珍国医国药*, 2010, 21(4): 979-980, 983.

3 结论

该研究通过超声波系统辅助提取水果中农药残留,毛细管柱气相色谱法(GC-FPD)测定水果中 6 种有机磷农药含量。此方法在水果样品浸渍提取的同时,辅以超声波提取,在合适的温度下完成样品的前处理工序,操作简单,提取效果好,快速节能、节省样品前处理的溶剂,污染小,可同时处理多组样品,省时高效,显著提高了样品前处理效率,进一步提高了实验室检测工作效率,适合进出口水果中有机磷农残的快速检测工作,能够更好地为检验检疫执法提供技术支持。

参考文献

- [1] 吴海燕, 陈建军. 蔬菜中多种有机磷农药残留气相色谱分析法 [J]. *食品与机械*, 2010, 26(4): 76-77, 83.
- [2] 中华人民共和国国家卫生和计划生育委员会, 中华人民共和国农业部, 国家食品药品监督管理局. 食品中农药最大残留限量: GB 2763—2016 [S]. 北京: 中国标准出版社, 2017.
- [3] 唐波, 张金娥, 齐力汇, 等. 气相色谱法测定农药残留量新进展 [J]. *山东师范大学学报(自然科学版)*, 2002, 17(4): 35-39.
- [4] 张福金, 刘建平, 冯小慧, 等. 气相色谱法检测蔬菜和水果多组分农药残留中若干问题的研究 [J]. *内蒙古农业科技*, 2006(3): 31-32.
- [5] 陆小磊, 吴慧明, 金绍强, 等. 毒死蜱在 4 种作物中残留量的气相色谱分析方法 [J]. *浙江农业科学*, 2009(1): 168-172.
- [6] 陆勋元, 蒋永祥. 气相色谱法同时检测和分析蔬菜中毒死蜱、氯氰菊酯和氰戊菊酯的残留量 [J]. *理化检验(化学分册)*, 2006, 42(9): 737-739.
- [7] 喻龙, 李光义, 邓晓, 等. 蔬菜中残留毒死蜱的检测方法 [J]. *中国农学通报*, 2007, 23(12): 338-340.
- [8] 汪军. GC-MS 定性定量分析出口西兰花样品中毒死蜱的农药残留量 [J]. *现代农药*, 2002(1): 23-24.
- [9] 徐颖洁. 气相色谱-质谱法测定农药中 14 种有机氯·有机磷成分 [J]. *安徽农业科学*, 2014, 42(17): 5470-5471, 5646.
- [10] 凌云, 王茵, 雍炜, 等. 气相色谱-质谱/质谱法检测蔬菜中的毒死蜱及其代谢物 [J]. *色谱*, 2009, 27(1): 78-81.
- [11] 于世林, 杜振霞. 化验员读本: 下册(仪器分析) [M]. 北京: 化学工业出版社, 2004.
- [12] 中华人民共和国农业部. 蔬菜和水果中有机磷、有机氯、拟除虫菊酯和氨基甲酸酯类农药多残留的测定: NY/T 761—2008 [S]. 北京: 中国农业出版社, 2008.

- [3] 国家卫生和计划生育委员会. 食品添加剂 薰衣草油: GB 1886. 38—2015 [S]. 北京: 中国标准出版社, 2016.
- [4] 王菊香, 邢志娜, 李伟, 等. BP 神经网络法结合红外光谱快速测定在用润滑油胺类抗氧化剂含量 [J]. *计算机与应用化学*, 2016, 33(2): 197-199.
- [5] 韩晓, 王菊香, 刘洁. 基于 BP-神经网络的航空煤油总酸值近红外光谱快速检测 [J]. *分析科学学报*, 2011, 27(6): 751-754.
- [6] 李智, 王圣毫, 郑维平, 等. 基于傅立叶变换的人工神经网络近红外光谱定量分析法 [J]. *分析测试学报*, 2012, 31(3): 343-346.
- [7] 王书涛, 陈东营, 魏蒙, 等. 荧光光谱法和 PSO-BP 神经网络在山梨酸钾浓度检测中的应用 [J]. *中国激光*, 2015, 42(5): 304-310.
- [8] 邹小波, 朱曾, 赵杰文. 基于间隔偏最小二乘法的农产品近红外光谱谱区选择方法 [J]. *现代科学仪器*, 2007(1): 86-88.
- [9] 迟亮, 张贺龙, 车英, 等. 基于 iPLS 的玉米近红外光谱数据处理方法研究 [J]. *长春理工大学学报(自然科学版)*, 2011, 34(2): 117-119.
- [10] 李强, 杨天邦, 涂公平. GA-BP 神经网络模型应用于岩芯扫描仪测定海洋沉积物中多种组分的半定量分析 [J]. *分析仪器*, 2018(1): 75-79.
- [11] 曹家兴, 陆建平. 遗传算法-贝叶斯正则化 BP 神经网络拟合滴定糖蜜中有机酸 [J]. *分析化学*, 2011, 39(5): 743-747.
- [12] 徐立鹏, 葛良全, 谷懿, 等. 基于 PCA-BP 神经网络的 EDXRF 分析测定地质样品中铁、钛元素含量的应用研究 [J]. *光谱学与光谱分析*, 2013, 33(5): 1392-1396.
- [13] 杰尔·沃克曼, 洛伊斯·文依. 近红外光谱解析实用指南 [M]. 褚小立, 许育鹏, 田高友, 译. 北京: 化学工业出版社, 2009.