

任广鑫,张银凤,尹畅,等.基于微型近红外光谱的黑茶地理溯源研究[J].华中农业大学学报,2025,44(6):59-66.  
DOI:10.13300/j.cnki.hnlkxb.2025.06.006

## 基于微型近红外光谱的黑茶地理溯源研究

任广鑫<sup>1,2</sup>,张银凤<sup>1</sup>,尹畅<sup>1</sup>,甄亭<sup>1</sup>,陈璐<sup>1</sup>,宁井铭<sup>2</sup>,张正竹<sup>3</sup>

1.淮南师范学院安徽省酿造工业微生物资源开发与应用工程研究中心/数字生态与健康研究所,淮南 232038;  
2.安徽农业大学茶树种质创新与资源利用全国重点实验室,合肥 230036; 3.安徽省农业科学院,合肥 230001

**摘要** 为规范茶叶市场秩序,保障消费者权益,本研究提出基于微型近红外光谱结合化学计量学的黑茶地理溯源的快速甄别方法,以提供透明、可信的黑茶产地溯源信息。以安徽安茶、四川雅安藏茶、湖南安化黑茶和湖北青砖茶为研究对象,采用微型近红外光谱原位采集样品光谱信息,利用模拟退火算法、粒子群优化算法、蚁群优化(ant colony optimization, ACO)算法与迭代变量子集优化法结合线性判别分析和极限学习机(extreme learning machine, ELM)方法构建黑茶地理溯源分类模型。结果显示,采用ACO提取10个与黑茶地理溯源有关的特征波长信息,基于上述10个波长信息的ELM预测模型的正确判别率为97.5%。研究结果表明,基于微型近红外光谱结合化学计量学方法对国内主要产区的黑茶地理溯源具有良好的实用价值。

**关键词** 微型近红外光谱; 化学计量学算法; 黑茶; 地理溯源; 蚁群优化

**中图分类号** S571.1 **文献标识码** A **文章编号** 1000-2421(2025)06-0059-08

黑茶作为后发酵茶的代表,是全球消费量最大的特种茶类之一<sup>[1]</sup>,随着其消费市场的持续扩张,质量控制变得愈发关键。为此,许多国家及国际机构正致力于建立黑茶质量评价与鉴别的标准体系<sup>[2]</sup>。湖南安化黑茶、云南普洱熟茶、湖北青砖茶、四川雅安藏茶与安徽安茶等均为我国历史名茶,拥有稳定的消费群体。然而,除云南普洱熟茶多以饼形压制外,其余几种产品形态均近似砖形,这种形态上的高度相似性,为茶叶的质量监控与地理来源鉴别带来了巨大挑战。茶叶的地理溯源显著影响其内含物质的积累,并与最终品质密切相关<sup>[2]</sup>。传统的黑茶质量评估高度依赖感官评价,但该方法易受品鉴者主观状态影响,导致结果缺乏可重复性与客观公正性<sup>[3-4]</sup>。因此,开发一种快速、客观的分析方法以实现黑茶产地的准确鉴别势在必行。

过去几十年间,研究人员采用了多种化学分析技术以解析黑茶的主要成分并进行菌种鉴定。Yuan等<sup>[1]</sup>利用电子鼻与多元统计分析法对黑茶等级质量进行了评价。Lv等<sup>[3]</sup>基于不同加工工艺制备方法研究了黑茶的植物化学成分谱与抗氧化活性。然

而,这些方法通常流程复杂、耗时较长。近红外光谱(NIRS)技术具有快速、无损、高效的特点,已在农业与食品科学领域得到广泛应用<sup>[5]</sup>,成为传统耗时化学分析方法的理想替代方案<sup>[6]</sup>,其原理在于,近红外光谱能够捕获样品中C-H、O-H、N-H和S-H等分子键的泛音与组合振动信号,这些信号差异源于样品化学组成的不同,从而能够反映其内部化学结构信息<sup>[7]</sup>。目前,利用NIRS进行茶叶地理溯源的研究已有诸多报道,但多数集中于红茶及茶鲜叶。将微型NIRS设备与化学计量学结合,专门用于黑茶地理溯源的研究,则仍鲜见文献记载。

模拟退火算法(simulated annealing, SA)是一种基于优化蒙特卡罗方法的进化元启发式算法,该方法引入了合适的随机因素和物理系统退火过程的自然机制<sup>[8]</sup>。在这个过程中,SA接受最优解使目标函数更好,以接受概率的最小温度值为目标函数,来实现函数的全局优化<sup>[9]</sup>。在SA算法运行中,最大种群数设置为20。该算法具有良好的局部搜索能力,已广泛应用于食品行业的特征波段筛选和组合优化问题<sup>[10]</sup>。粒子群优化算法(particle swarm optimization)

收稿日期:2025-07-10

基金项目:国家重点研发计划项目(2021YFD1601102);安徽农业大学茶树种质创新与资源利用全国重点实验室开放基金(SKLT0F20220127);安徽省高校自然科学研究重点项目(2022AH051590);淮南市科技计划项目(2023A314)

任广鑫, E-mail: rgx@hnnu.edu.cn

通信作者:张正竹, E-mail: zzz@ahau.edu.cn

tion, PSO)作为一种相对新颖的优化技术被应用于模拟鸟群、鱼群等群集的社会行为<sup>[11]</sup>。将每只鸟看作目标空间中优化算法的每一个解,  $N$ 个粒子群随着每次迭代而演化,并根据优化方法的原理向问题的最优解移动,粒子在前一个最优位置和全局最优位置的路径优化中运动,探索整个种群的最优解<sup>[12]</sup>。蚁群优化(ant colony optimization, ACO)算法是一种基于蚁群自然进化觅食行为的群优化算法,当科学家们研究蚂蚁的觅食行为时,发现蚂蚁并没有形成一个完整的视觉系统,它们可以通过传播信息素来相互交流<sup>[13]</sup>,该算法根据蚁群所采用的信息反馈机制来寻找最短路径,实现智能搜索、参数优化等功能<sup>[14]</sup>。迭代变量子集优化法(iteratively variable subset optimization, IVSO)是由Wang等<sup>[15]</sup>提出的一种光谱变量筛选方法,它是基于偏最小二乘模型的回归系数而提出的一种算法,并且考虑了交叉验证产生的潜变量的最优数量会对回归系数产生很大的影响,有时这种影响甚至会相差几个数量级<sup>[16-17]</sup>。LDA是一种广为人知的机器学习算法,用于分类识别<sup>[18]</sup>。这种建模方法已被应用于多种领域,包括食品鉴别等需要分析高维数据的场景<sup>[19-20]</sup>。ELM是一种单隐藏层前馈神经网络,它克服了前馈神经网络的训练速度慢等缺点,运行时易于捕捉局部最优值<sup>[21]</sup>。

本研究以安徽安茶、四川雅安藏茶、湖南安化黑茶和湖北青砖茶为研究对象,运用微型NIRS结合SA、PSO、ACO与IVSO等算法以及线性判别分析(linear discriminant analysis, LDA)和极限学习机(extreme learning machine, ELM)方法,开发地理溯源识别模型,以期实现对黑茶产区的有效甄别,提供透明、可信的黑茶产地溯源信息。

## 1 材料与方 法

### 1.1 试验材料

从茶叶产区分别收集了不同等级的安徽安茶、四川雅安藏茶、湖南安化黑茶和湖北青砖茶等外形相似的黑茶产品,样品分别来源于合肥祥源茶业股份有限公司、雅安藏茶坊茶业有限公司、湖南安化黑茶集团有限公司和湖北赵李桥茶业有限公司,共计120份,每种茶叶30份。为了确保黑茶样品的品质,将其保存在 $-20\text{ }^{\circ}\text{C}$ 冰柜中,供进一步分析。

### 1.2 试验仪器

Horner10GT型Android智能手机,深圳华为技

术有限公司;NIR-S-R2型微型近红外光谱仪,中国台湾InnoSpectra公司;AR224CN型高精度电子天平,上海奥豪斯仪器有限公司;BD/BC-203KMD(E)型美的冷柜,美的集团股份有限公司;BHL-125B型电热恒温干燥箱,上海本亭仪器有限公司。

### 1.3 光谱信号获取

本研究的技术路线是利用微型NIRS仪采集黑茶样品光谱信息,采用NIRS预处理方法对其进行预处理,通过特征波长选择算法和模式识别算法建立简化分类模型。该微型光谱仪的长宽高分别约75、58、26.5 mm,尺寸大小与鼠标相近。在采集数据前,打开光谱仪预热约10 min,之后打开智能手机的蓝牙功能,通过蓝牙与光谱仪进行连接。同时在智能手机上根据仪器厂商的指示安装数据采集软件,并对采集过程中的关键参数和文件名进行设置。采集光谱数据时,对每份样品的随机5个点进行扫描和保存,之后以5次扫描保存数据的平均值作为该样品的代表光谱数据。得到的数据以CSV格式存储在智能手机中,再将其传输到计算机中进行处理。

### 1.4 光谱预处理

在NIRS分析中,对原始光谱数据的预处理有助于提升谱图质量,甄别分析均在此基础上展开。微型光谱仪器获取的原始NIRS容易受到噪声干扰与基线漂移等的影响<sup>[22]</sup>。选用Savitzky-Golay(SG)平滑预处理对样品原始光谱数据进行处理,以增强样品靶向特征属性<sup>[23]</sup>。SG平滑通过多项式的方式对移动窗口中的特征数据进行最小二乘的信息拟合的一种加权平均法,用于消除噪音与基线漂移现象<sup>[24]</sup>。样品的原始光谱与SG预处理光谱图见图1。图1中上述2种曲线没有表现出明显不同的特征吸收峰,预处理后的谱图比原始谱图曲线更集中与平滑,表明SG预处理消除了噪音和基线漂移,谱图预处理方法有利于提高光谱质量。

### 1.5 模型性能评价

为评价黑茶地理溯源模型性能,用校正集和预测集的正确判别率(correct discrimination rate, CDR)来评估模型的预测能力。一般来说,一个优秀的判别模型应有较高的CDR值(公式中用 $C_{DR}$ 表示),其计算方法如下:

$$C_{DR} = \frac{N_{CDR}}{N_t} \times 100\% \quad (1)$$

其中, $N_{CDR}$ 为校正/预测样本的正确估计数, $N_t$ 为校正/预测样本的总数。上述算法均由Matlab

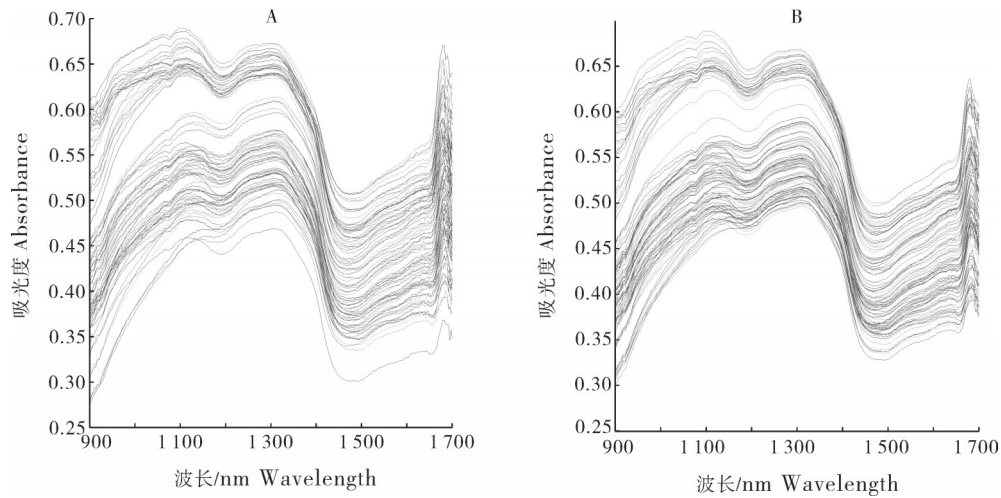


图1 样品原始光谱图(A)与SG预处理光谱图(B)

Fig.1 Original spectrum of the sample(A) and spectrum after SG pretreatment(B)

2019b软件在 Windows 8平台下自主开发。

## 2 结果与分析

### 2.1 样品集划分与主成分分析

采用 Kennard Stone 算法,按照 2:1 比例划分,依次得到校正样品集和预测样品集样品数分别为 80、40 个。由图 2A 可以看出,黑茶样品的 NIRS 数据特

征的样品集在空间分布上相对离散,校正集样品的分布变幅覆盖了预测集样品的分布,表明校正样品集与预测样品集的样品分布是合理的。图 2B 显示,不同地理来源的样品在三维空间上有部分重叠现象,基于 NIRS 特征数据无法简单地将不同地域的样品区分开,需要融入适当的化学计量学手段,以达到对不同产区黑茶样品的有效溯源。

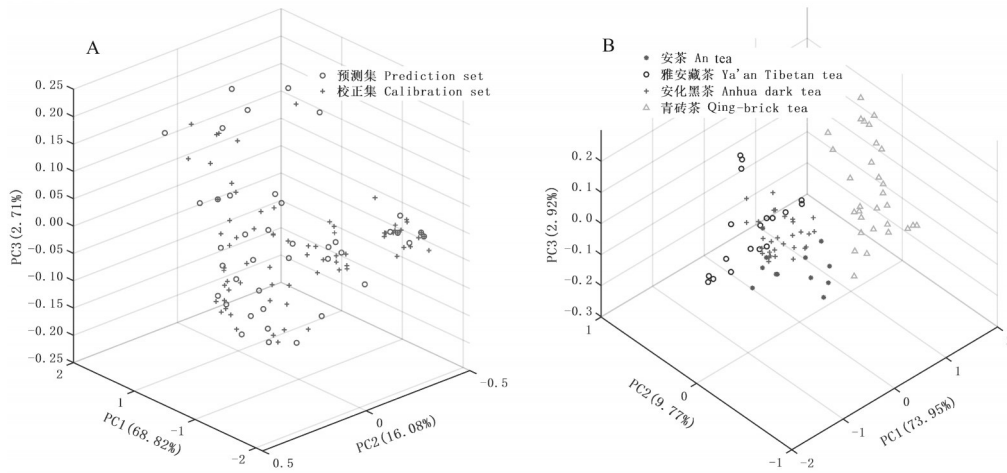


图2 不同地理来源样品的三维空间分布(A)与三维主成分分布(B)

Fig.2 Three-dimensional spatial distribution(A) and three-dimensional principal component distribution(B) of samples from different geographical origins

### 2.2 特征波长选择

采用特征波长选择算法(SA、PSO、ACO 和 IV-SO)对微型 NIRS 仪器采集的不同地理来源黑茶样品的光谱数据进行特征波长的筛选。上述特征波长优化选择算法经多次迭代,最终运行结束。上述 4 种特征波长选择算法在多次迭代过程中的特征波长统计结果见图 3 和表 1。表 1 结果显示,SA、PSO、ACO

和 IVSO 算法获得的黑茶样品特征波长数量依次为 20、13、10 和 58,获得的波长占总波长数目的比例依次为 8.77%、5.70%、4.38% 和 25.43%。特征波长可视化图见图 4,由图 4 可知,所选波长范围多集中在 1 200~1 700 nm,即在长波近红外区域。通过对黑茶样品光谱进行特征波段选择,可对模型进一步简化与提升运行效率,获得高质量模型。

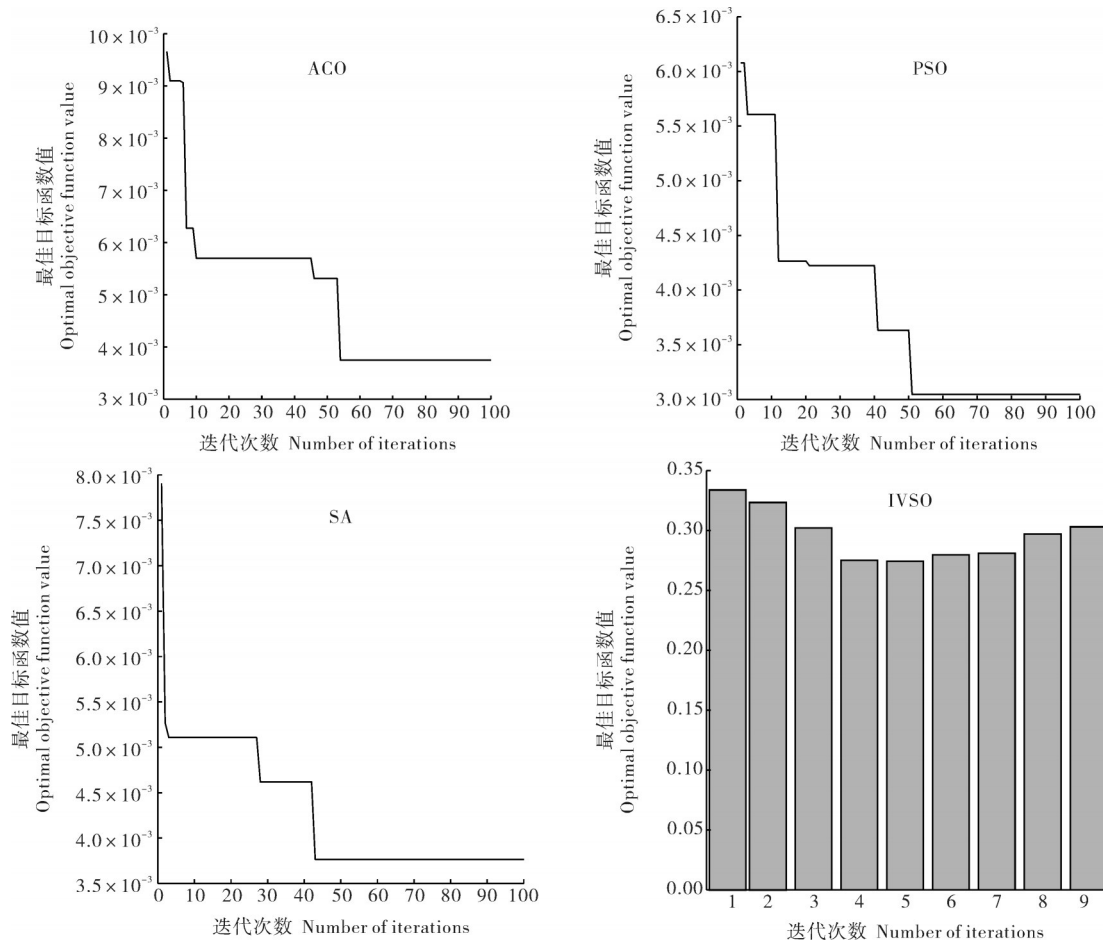


图3 特征波长选择运行统计图

Fig.3 Characteristic wavelength selection operation statistics chart

表1 不同特征波长选择方法的优化结果

Table 1 Optimization results of different feature wavelength selection methods

方法 Method	波长个数 Wavelength number	最佳目标函数值 Optimal objective function value	选择的波长/nm Selected wavelength
SA	20	0.003 7	985.61, 1 029.38, 1 033.11, 1 189.16, 1 254.42, 1 272.79, 1 284.21, 1 294.44, 1 326.04, 1 358.37, 1 403.39, 1 406.65, 1 416.42, 1 436.90, 1 477.34, 1 593.51, 1 618.39, 1 625.30, 1 689.31, 1 695.03
PSO	13	0.003 1	1 024.41, 1 067.65, 1 083.57, 1 094.54, 1 151.19, 1 199.75, 1 276.22, 1 297.85, 1 322.67, 1 422.90, 1 532.59, 1 679.75, 1 697.88
ACO	10	0.003 7	957.80, 978.05, 1 060.28, 1 083.57, 1 214.97, 1 497.30, 1 574.41, 1 615.42, 1 640.03, 1 697.88
IVSO	58	0.308 2	901.49, 905.36, 937.43, 941.26, 961.61, 965.41, 985.61, 989.39, 1 060.28, 1 121.19, 1 211.47, 1 228.96, 1 232.44, 1 247.50, 1 262.47, 1 272.79, 1 284.21, 1 291.04, 1 294.44, 1 315.93, 1 333.88, 1 343.93, 1 358.37, 1 399.04, 1 403.39, 1 406.65, 1 409.91, 1 413.17, 1 430.45, 1 454.01, 1 477.34, 1 480.50, 1 483.66, 1 503.57, 1 506.70, 1 509.82, 1 526.40, 1 545.94, 1 565.30, 1 568.34, 1 574.41, 1 584.49, 1 587.50, 1 590.51, 1 593.51, 1 599.51, 1 603.50, 1 609.47, 1 615.42, 1 631.21, 1 642.97, 1 646.87, 1 649.79, 1 655.63, 1 664.35, 1 679.75, 1 695.03, 1 697.88

2.3 基于光谱数据的黑茶地理溯源模型构建

基于SA、PSO、ACO和IVSO算法优化SG预处理

理光谱特征的LDA和ELM分类模型统计结果见表

2。表2结果显示,在校正集和预测集中所有ELM模

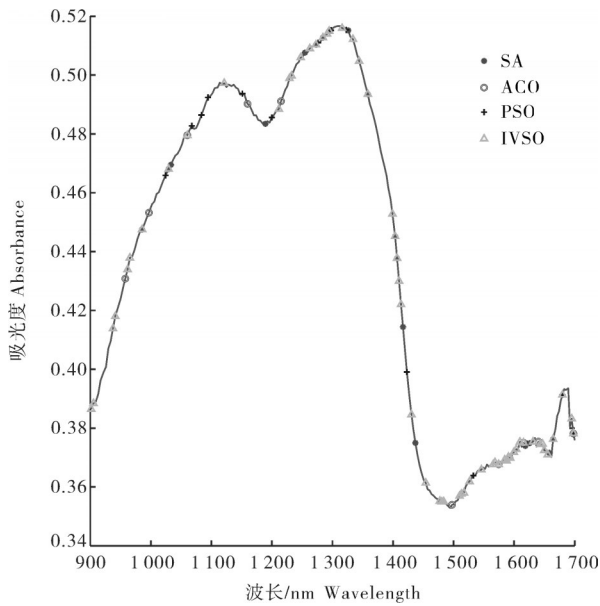


图4 特征波长可视化图

Fig.4 Feature wavelength visualization diagram

型的CDR普遍高于相应的LDA模型的CDR值。在LDA模型中,最佳性能为IVSO-LDA=ACO-LDA>PSO-LDA>SA-LDA, IVSO-LDA模型和ACO-LDA模型的预测集CDR均为92.5%。ACO优选的波段较IVSO获得的波段少,故ACO-LDA模型性能更好。在预测过程中,SA-LDA模型的CDR为87.5%,精度最低。

表2 基于光谱特征和不同化学计量学所建分类模型的正确判别率

Table 2 Discrimination accuracy of classification models constructed based on spectral features and different chemometric approaches

模型 Model	参数 Parameter	校正集/% Calibration set (n=80)	预测集/% Prediction set (n=40)
SA-ELM	nn=32	100.00	97.5
PSO-ELM	nn=13	97.50	95.0
ACO-ELM	nn=16	100.00	97.5
IVSO-ELM	nn=9	100.00	97.5
SA-LDA	PC=5	90.00	87.5
PSO-LDA	PC=3	93.75	90.0
ACO-LDA	PC=3	93.75	92.5
IVSO-LDA	PC=2	96.25	92.5

在ELM模型中,隐含层神经元个数对ELM性能的影响如图5所示。图5结果显示,PSO-ELM模型性能最低,其CDR为95%。SA、ACO和IVSO结合ELM模型的效果均为97.5%。然而,ACO提取的特征波长为10个,较SA和IVSO优化得到的波长

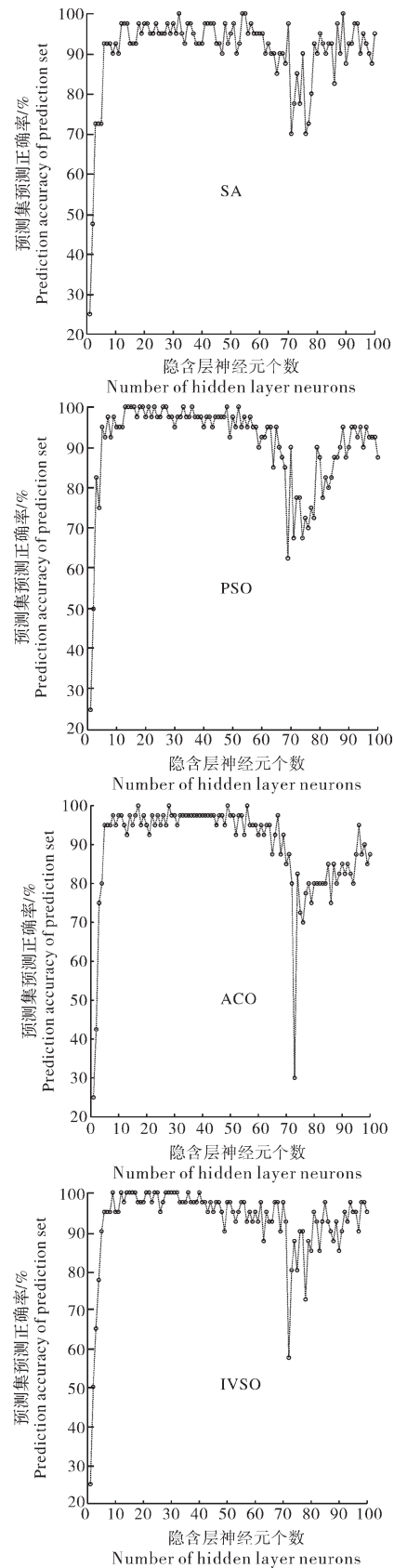


图5 隐含层神经元个数对ELM性能的影响  
Fig.5 The effect of the number of hidden layer neurons on ELM performance

少。因此,ACO-ELM算法对不同地理来源黑茶的分类模型进行了简化,在保证预测精度的前提下,ACO-ELM模型的预测效果最佳。

### 3 讨论

不同地理来源的黑茶样品在900~1700 nm波段内有228个波长点。SA、PSO、ACO和IVSO等4种特征波长优化方法选择的波段多集中于NIRS的长波段,该区域的光谱含有丰富的不同地理来源的茶叶样品信息,是NIRS的主要吸收带<sup>[25]</sup>。ACO算法为最佳波长选择方法,其最优波段是957.80、978.05、1060.28、1083.57、1214.97、1497.30、1574.41、1615.42、1640.03、1697.88 nm,其中,957.80、978.05 nm对应—CH和—CH<sub>2</sub>的第三倍频区域;1060.28、1083.57 nm对应Ar—CH的第三倍频区域;1214.97 nm对应—CH<sub>3</sub>的第二倍频区域;1497.30 nm和1574.41 nm对应—RNH<sub>2</sub>的第二倍频区域;1615.42、1640.03、1697.88 nm对应—CH和—CH<sub>2</sub>的第一倍频区域,与—CH和—OH的伸缩振动有关<sup>[26]</sup>。

LDA算法的分类准确率较ELM算法的低,其主要原因可能是LDA为线性算法,仅支持线性投影,无法处理非线性关系,可能限制复杂数据的分类能力,对类间散度矩阵产生显著影响,降低其模型的稳定性<sup>[27]</sup>。因此,在解决非线性的地理溯源问题上效果不是很好。ELM是一种单层前馈神经网络,被证明在产地预测领域中具有良好的性能。ELM通过随机初始化输入权重和偏置,以及解析地计算输出权重,实现了快速的学习速度<sup>[28]</sup>。

目前,有关茶叶地理溯源的NIRS研究主要针对红茶和鲜叶类产品。笔者所在团队开发一种新的NIRS和因子化法,构建了不同地理来源(中国、印度、肯尼亚、斯里兰卡和缅甸)的红茶同步识别模型,识别准确率达94.3%,且采纳上百个光谱波段<sup>[29]</sup>。Wang等<sup>[30]</sup>应用NIRS结合协同区间偏最小二乘法和遗传算法,用于追溯恩施玉露鲜茶叶的地理来源,该模型在预测集样本上的预测准确率为96.67%。本研究基于微型NIRS数据的最佳ACO-ELM模型的CDR为97.5%,且仅提取10个特征波长。可见,不同地理来源黑茶的NIRS分类模型性能更佳。未来,可将该模型扩展到其他茶叶品种或不同地理区域,为该模型与技术的推广应用提供可能。

### 参考文献 References

- [1] YUAN H B, CHEN X Q, SHAO Y D, et al. Quality evaluation of green and dark tea grade using electronic nose and multivariate statistical analysis[J]. Journal of food science, 2019, 84(12): 3411-3417.
- [2] ZOU Y, MA W J, TANG Q, et al. A high-precision method evaluating color quality of Sichuan dark tea based on colorimeter combined with multi-layer perceptron [J/OL]. Journal of food process engineering, 2020, 43(8): e13444 [2025-07-10]. <https://doi.org/10.1111/jfpe.13444>.
- [3] LV H P, ZHANG Y, SHI J, et al. Phytochemical profiles and antioxidant activities of Chinese dark teas obtained by different processing technologies[J]. Food research international, 2017, 100: 486-493.
- [4] ZHANG H, WANG J J, ZHANG D D, et al. Aged fragrance formed during the post-fermentation process of dark tea at an industrial scale[J/OL]. Food chemistry, 2021, 342: 128175 [2025-07-10]. <https://doi.org/10.1016/j.foodchem.2020.128175>.
- [5] 王林, 郑明明, 王翀, 等. 基于近红外光谱的卷烟配方模块香型预测[J]. 华中农业大学学报, 2024, 43(1): 226-231. WANG L, ZHENG M M, WANG C, et al. Effects of apple size on applicability of model for predicting content of sugar based on near infrared spectroscopy [J]. Journal of Huazhong Agricultural University, 2024, 43(1): 226-231 (in Chinese with English abstract).
- [6] 杨凯, 何显廷, 李沃霖, 等. 基于近红外光谱技术检测全蛋粉掺假[J]. 华中农业大学学报, 2024, 43(2): 264-272. YANG K, HE Y T, LI W L, et al. Detection of adulteration in whole egg powder based on near-infrared spectroscopy[J]. Journal of Huazhong Agricultural University, 2024, 43(2): 264-272 (in Chinese with English abstract).
- [7] REN G X, ZHANG X S, WU R, et al. Rapid characterization of black tea taste quality using miniature NIR spectroscopy and electronic tongue sensors[J/OL]. Biosensors, 2023, 13(1): 92 [2025-07-10]. <https://doi.org/10.3390/bios13010092>.
- [8] WANG S P, ALTANER C, FENG L, et al. A review: integration of NIRS and chemometric methods for tea quality control-principles, spectral preprocessing methods, machine learning algorithms, research progress, and future directions [J/OL]. Food research international, 2025, 205: 115870 [2025-07-10]. <https://doi.org/10.1016/j.foodres.2025.115870>.
- [9] 任广鑫, 高钰敏, 张雨茹, 等. 基于电子舌与近红外光谱融合的滇红工夫红茶滋味品质评价[J]. 农业机械学报, 2025, 56(1): 102-109. REN G X, GAO Y M, ZHANG Y R, et al. Evaluation of taste quality of Dianhong Congou black tea based on fusion of electronic tongue and near-infrared spectroscopy[J]. Transactions of the CSAM, 2025, 56(1): 102-109 (in Chinese with English abstract).
- [10] REN G X, SUN Y M, LI M H, et al. Cognitive spectroscopy

- for evaluating Chinese black tea grades (*Camellia sinensis*): near-infrared spectroscopy and evolutionary algorithms [J]. *Journal of the science of food and agriculture*, 2020, 100(10): 3950-3959.
- [11] REN G X, NING J M, ZHANG Z Z. Multi-variable selection strategy based on near-infrared spectra for the rapid description of Dianhong black tea quality [J/OL]. *Spectrochimica acta part A: molecular and biomolecular spectroscopy*, 2021, 245: 118918 [2025-07-10]. <https://doi.org/10.1016/j.saa.2020.118918>.
- [12] JIANG H, LIU T, HE P H, et al. Rapid measurement of fatty acid content during flour storage using a color-sensitive gas sensor array: comparing the effects of swarm intelligence optimization algorithms on sensor features [J/OL]. *Food chemistry*, 2021, 338: 127828 [2025-07-10]. <https://doi.org/10.1016/j.foodchem.2020.127828>.
- [13] WANG B N, AHMAD W, CHEN Q S, et al. Development of a dual-mode upconversion nanoparticles-3-aminophenol nanosystem based on inner filter effect for sensitive detection of nitrite and its application on test strips [J/OL]. *Sensors and actuators B: chemical*, 2023, 374: 132740 [2025-07-10]. <https://doi.org/10.1016/j.snb.2022.132740>.
- [14] XU W D, HE Y C, LI J H, et al. Portable beef-freshness detection platform based on colorimetric sensor array technology and bionic algorithms for total volatile basic nitrogen (TVB-N) determination [J/OL]. *Food control*, 2023, 150: 109741 [2025-07-10]. <https://doi.org/10.1016/j.foodcont.2023.109741>.
- [15] WANG W T, YUN Y H, DENG B C, et al. Iteratively variable subset optimization for multivariate calibration [J]. *RSC advances*, 2015, 5(116): 95771-95780.
- [16] WANG Y J, LI M H, LI L Q, et al. Green analytical assay for the quality assessment of tea by using pocket-sized NIR spectrometer [J/OL]. *Food chemistry*, 2021, 345: 128816 [2025-07-10]. <https://doi.org/10.1016/j.foodchem.2020.128816>.
- [17] LIU L H, ZAREEF M, WANG Z, et al. Monitoring chlorophyll changes during Tencha processing using portable near-infrared spectroscopy [J/OL]. *Food chemistry*, 2023, 412: 135505 [2025-07-10]. <https://doi.org/10.1016/j.foodchem.2023.135505>.
- [18] JIANG H, XU W D, CHEN Q S. High precision qualitative identification of yeast growth phases using molecular fusion spectra [J/OL]. *Microchemical journal*, 2019, 151: 104211 [2025-07-10]. <https://doi.org/10.1016/j.microc.2019.104211>.
- [19] GAGGIOTTI S, SHKEMBI B, SACCHETTI G, et al. Study on volatile markers of Pasta quality using GC-MS and a peptide based gas sensor array [J/OL]. *LWT*, 2019, 114: 108364 [2025-07-10]. <https://doi.org/10.1016/j.lwt.2019.108364>.
- [20] LIN H, MAN Z X, KANG W C, et al. A novel colorimetric sensor array based on boron-dipyromethene dyes for monitoring the storage time of rice [J]. *Food chemistry*, 2018, 268: 300-306.
- [21] LI H H, LUO X F, HARUNA S A, et al. Au-Ag OHCs-based SERS sensor coupled with deep learning CNN algorithm to quantify thiram and pymetrozine in tea [J/OL]. *Food chemistry*, 2023, 428: 136798 [2025-07-10]. <https://doi.org/10.1016/j.foodchem.2023.136798>.
- [22] REN G X, ZHANG X S, WU R, et al. Digital depiction of the quality of Dianhong black tea based on pocket-sized near infrared spectroscopy [J/OL]. *Infrared physics & technology*, 2022, 127: 104418 [2025-07-10]. <https://doi.org/10.1016/j.infrared.2022.104418>.
- [23] CHEN Y, GUO M Q, CHEN K, et al. Predictive models for sensory score and physicochemical composition of Yuezhou Longjing tea using near-infrared spectroscopy and data fusion [J/OL]. *Talanta*, 2024, 273: 125892 [2025-07-10]. <https://doi.org/10.1016/j.talanta.2024.125892>.
- [24] SONG Y, YI W Q, LIU Y, et al. A robust deep learning model for predicting green tea moisture content during fixation using near-infrared spectroscopy: integration of multi-scale feature fusion and attention mechanisms [J/OL]. *Food research international*, 2025, 203: 115874 [2025-07-10]. <https://doi.org/10.1016/j.foodres.2025.115874>.
- [25] LAN W J, JAILLAIS B, LECA A, et al. A new application of NIR spectroscopy to describe and predict purees quality from the non-destructive apple measurements [J/OL]. *Food chemistry*, 2020, 310: 125944 [2025-07-10]. <https://doi.org/10.1016/j.foodchem.2019.125944>.
- [26] 荣艳娜, 柳新荣, 邢志强, 等. 抹茶品质指标的可见近红外光谱检测研究 [J]. *食品安全质量检测学报*, 2024, 15(3): 125-132. RONG Y N, LIU X R, XING Z Q, et al. Detection of quality indicators in matcha using visible-near infrared spectroscopy [J]. *Journal of food safety & quality*, 2024, 15(3): 125-132 (in Chinese with English abstract).
- [27] MA G C, ZHANG Y B, ZHANG J Y, et al. Determining the geographical origin of Chinese green tea by linear discriminant analysis of trace metals and rare earth elements: taking Dongting Biluochun as an example [J]. *Food control*, 2016, 59: 714-720.
- [28] LI J C, ZHANG S Z, LIU C L, et al. Characterization of ash content in wheat flour using data fusion [J/OL]. *Infrared physics & technology*, 2023, 133: 104792 [2025-07-10]. <https://doi.org/10.1016/j.infrared.2023.104792>.
- [29] REN G X, WANG S P, NING J M, et al. Quantitative analysis and geographical traceability of black tea using Fourier transform near-infrared spectroscopy (FT-NIRS) [J]. *Food research international*, 2013, 53(2): 822-826.
- [30] WANG S P, FENG L, LIU P P, et al. An efficient method for tracing the geographic origin of Enshi Yulu fresh tea leaves based on near infrared spectroscopy combined with synergy interval PLS and genetic algorithm [J/OL]. *LWT*, 2024, 203: 116372 [2025-07-10]. <https://doi.org/10.1016/j.lwt.2024.116372>.

## Geographical traceability of dark tea based on miniature near infrared spectroscopy

REN Guangxin<sup>1,2</sup>, ZHANG Yinfeng<sup>1</sup>, YIN Chang<sup>1</sup>, ZHEN Ting<sup>1</sup>,  
CHEN Lu<sup>1</sup>, NING Jingming<sup>2</sup>, ZHANG Zhengzhu<sup>3</sup>

1. *Anhui Province Brewing Industry Microbial Resources Development and Application Engineering Research Center/Institute of Digital Ecology and Health,*

*Huainan Normal University, Huainan 232038, China;*

2. *State Key Laboratory of Tea Plant Germplasm Innovation and Resource Utilization, Anhui Agricultural University, Hefei 230036, China;*

3. *Anhui Province Academy of Agricultural Sciences, Hefei 230001, China*

**Abstract** A rapid identification method for geographical traceability of dark tea based on the miniature near-infrared spectroscopy combined with chemometrics was proposed to provide transparent and reliable traceability information about the areas of producing dark tea, standardize the order of tea market and protect rights of consumers. The miniature near-infrared spectrometer was used to collect spectral information *in situ* of dark tea including the An-tea in Anhui Province, Tibetan tea in Ya'an City of Sichuan Province, Anhua dark tea in Hunan Province, and Chin-brick tea in Hubei Province. The simulated annealing algorithms, particle swarm optimization algorithms, ant colony optimization (ACO) algorithm, and iterative variable set optimization method combined with linear discriminant analysis and extreme learning machine (ELM) method were used to construct a geographical traceability classification model for dark tea. The results showed that the accuracy of the ELM prediction model based on the 10 characteristic wavelength information related to the geographical traceability of dark tea extracted with ACO was 97.5%. It is indicated that the combination of the miniature near-infrared spectroscopy and chemometric methods has good practical value for the geographical traceability of dark tea from major production areas in China.

**Keywords** miniature near infrared spectroscopy; chemometric algorithms; dark tea; geographical traceability; ant colony optimization (ACO)

(责任编辑:陆文昌)