DOI:10.12441/spyswjs.20240919002 [文章编号] 1673-1689(2025)09-0153-10

基于近红外光谱的铁观音乌龙茶烘焙程度判别及品质预测

黄丽竹[1] 卞 芸[1] 李臻峰[1] 李先峰[2] 宋春芳[1*]

(1. 江南大学智能制造学院,江苏 无锡 214122; 2. 山东碧海机械科技有限公司,山东 临沂 276600)

摘要:【目的】鉴定烘焙过程中铁观音乌龙茶的品质。【方法】使用近红外光谱技术对烘焙过程中的铁观音乌龙茶进行无损检测,并使用化学测量法对铁观音乌龙茶内在品质的变化进行分析。基于近红外光谱数据,对光谱进行预处理和降维分析,使用支持向量机(support vector machine,SVM)和反向传播神经网络(back propagation neural network,BPNN)构建铁观音乌龙茶烘焙程度判别模型。基于近红外光谱数据结合内在品质,使用偏最小二乘回归(partial least squares regression,PLSR)构建品质预测模型。【结果】通过茶叶内在品质的变化和感官评审将铁观音乌龙茶的烘焙程度划分为烘焙不足、烘焙适度和烘焙过度。在全波段光谱数据模型下,用多元散射校正(multiple scatter correction,MSC)进行预处理,用 BPNN 构建判别模型时,判别效果最优,测试集准确率为 100.00%。对 4 个内在品质进行预测时,连续投影变换(successive projections algorithm,SPA)算法结合 PLSR 的预测模型的精度均为最优。游离氨基酸、茶多酚、儿茶素和咖啡碱的最佳预测模型的测试集决定系数(determination coefficient of prediction,R_P^2)分别为 0.949 6、0.944 3、0.950 8 和 0.740 0。【结论】该研究实现了铁观音乌龙茶烘焙程度的精准判别和品质的快速预测,为乌龙茶烘焙程度的精准判别和控制提供了理论依据。

关键词:铁观音乌龙茶烘焙;近红外光谱;程度判别;品质预测

Baking Degree Discrimination and Quality Prediction of Tieguanyin Oolong Tea Based on Near-infrared Spectroscopy

HUANG Lizhu[1] *BIAN Yun*[1] *Li Zhenfeng*[1] *Li Xianfeng*[2] *SONG Chunfang*[1*]

(1. School of Intelligent Manufacturing, Jiangnan University, Wuxi 214122, China; 2. Shandong Bihai Machinery Technology Co., Ltd., Linyi 276600, China)

Abstract: [**Objective**] This study aims to assess the quality of Tieguanyin oolong tea during the baking process. [**Method**] Near-infrared spectroscopy was used for non-destructive detection of Tieguanyin oolong tea during the baking process and chemometrics was employed to analyze the changes in the intrinsic quality of the Tieguanyin oolong tea. Spectral preprocessing and dimensionality reduction were conducted for the near-infrared spectral data. Support vector machine (SVM) and back propagation neural network (BPNN) were adopted to construct discrimination models for the baking degree of Tieguanyin oolong tea. Combining the near-infrared spectral data with intrinsic quality data, partial least squares regression (PLSR) was employed to build quality prediction models. [**Result**] According to the changes in intrinsic quality and sensory evaluation results, the baking degree of Tieguanyin oolong tea can be graded into under baking, moderate baking, and over baking. Under the full-

基金项目:紫金县市级农业科技园区(茶叶)科技计划项目。

通信作者:宋春芳(1974—),女,博士,教授,博士研究生导师,主要从事农产品无损检测与控制研究。E-mail:songcf@jiangnan.edu.cn

收稿日期:2024-09-19 **修回日期:**2024-12-25

spectrum spectral data model, with preprocessing as multiple scatter correction (MSC) and the discriminant model as BPNN, the discriminant effect was the best, and the accuracy rate of the test set was 100.00%. When predicting the four intrinsic quality properties, the prediction model combining the successive projections algorithm (SPA) with PLSR had the best accuracy. The determination coefficients of prediction (R_P^2) of the best prediction model for free amino acids, tea polyphenols, catechins, and caffeine were 0.949 6, 0.944 3, 0.950 8, and 0.740 0, respectively. [**Conclusion**] This study realized the accurate discrimination of the baking degree and the rapid prediction of quality of Tieguanyin oolong tea, providing a theoretical foundation for the accurate discrimination and control of oolong tea baking.

Keywords: Tieguanyin oolong tea baking; near-infrared spectroscopy; degree discrimination; quality prediction

茶是中国的一种源远流长的传统民俗饮品，有几千年的发展历史[1]。根据加工工艺的差异，茶叶可划分为六大类，乌龙茶便是其中之一，它属于绿茶与红茶之间的半发酵茶种[2]。乌龙茶的口感别具一格，既保留了传统绿茶的鲜爽，又融合了传统红茶的香醇，故而享有“绿叶红镶边，七泡有余香”的美誉。乌龙茶主要产自中国的福建、广东及台湾等地区，以其独特的香气与丰富的口感享誉国内外。铁观音是乌龙茶的典型代表，起源于福建省安溪县[3]。铁观音以其独特的兰花香和甘甜回味著称，茶汤清澈金黄、口感圆润，且具有较高的抗氧化活性和多种健康功效。铁观音的制作工艺复杂，包括采摘、萎凋、摇青、杀青、揉捻、烘焙等多个环节，其中烘焙尤为关键，它对茶叶的风味、香气及最终品质有重要影响[4]。茶叶内的化学成分（如游离氨基酸、茶多酚类、咖啡碱等）在烘焙过程中会随着烘焙程度的不同而发生变化[5]。烘焙适度可以增添茶叶的香气，使茶汤口感更加醇厚；烘焙过度可能会导致茶叶苦涩感增加、香气损失；烘焙不足则可能无法充分激发茶叶的固有风味。因此，精确判别和控制乌龙茶的烘焙程度对于确保茶叶的最佳品质至关重要[6]，这不仅影响了茶叶的感官体验，还决定了茶叶的整体风味和市场接受度。

传统的茶叶品质的检测主要依赖制茶师的感官评审，这种方法具有一定的主观性，容易受到评估者个人经验和状态的影响，导致结果的可靠性和一致性不足。为了克服感官评审的局限性，通过无损检测技术检测食品质量逐渐成为研究重点，如计算机视觉技术、电子鼻技术、近红外光谱技术、电子舌技术、高光谱技术等广泛应用于茶叶[7-9]、肉类[10-11]、牛奶[12-13]等食品的质量检测。近红外光谱技术因快速、准确且无损的优势而备受关注。近红外光谱技术能够通过分析茶叶中化学成分的特征吸收光谱图，客观地评估茶叶的品质[14]。Tian 等[15]证明了将主成分分析（principal component analysis，PCA）和近红外光谱相结合可以区分不同地理位置的甜茶，利用不同的光谱预处理方法建立光谱信息与成分含量之间的 PLSR 模型，能够有效预测黄酮类化合物含量。Wang 等[16]针对鲜茶叶在采收后不断失水的问题，建立了基于近红外光谱的鲜茶采后水分在线检测系统。Ren 等[17]为实现茶叶中灰分的快速、有效定量分析，提出了利用国产微型近红外光谱仪结合多变量分析快速检测红茶中的灰分含量。Ren 等[18]发现便携式近红外光谱技术与化学测量学方法相结合，可提高滇红品质预测性能。Jia 等[19]建立了基于近红外光谱的越州龙井茶不同品质等级感官质量评分、总儿茶素和咖啡因的定量预测模型。Wu 等[20]发现可见-近红外光谱技术与多变量分析相结合，可快速、无损评估抹茶的粒径以及茶多酚与游离氨基酸的比例。Ding 等[21]提出了一种基于近红外光谱的茶叶品质等级划分方法。上述研究表明，基于近红外光谱技术建立判别模型对茶叶质量、品种、等级等进行鉴别，可实现对茶叶的定性分析。此外，结合茶叶品质指标可构建茶叶品质成分定量预测模型。

因此，该研究利用近红外光谱技术结合化学测量法，对烘焙过程中铁观音乌龙茶的内在品质进行检测，基于机器学习构建了铁观音乌龙茶烘焙程度判别模型及品质预测模型，实现了对烘焙乌龙茶的定性和定量分析。

1 材料与方法

1.1 样品

清香型安溪铁观音乌龙茶：购于福建省泉州

市雁石峰农家茶品牌店。铁观音乌龙茶茶叶已经经过茶农前期制作加工，热炒后茶叶的含水率为5%~8%。采用分阶段式烘焙法对铁观音乌龙茶进行烘焙，参照文献[22]，设置烘焙条件为：90 ℃烘焙2 h，110 ℃烘焙2 h，130 ℃烘焙2 h（烘焙笼中茶叶质量为200 g，摊叶厚度为1 cm）。每隔30 min对茶叶进行取样，[0，180] min时每次取10个茶叶样品，（180，300] min时每次取15个茶叶样品，（300，360] min时每次取30个茶叶样品。取出的茶叶立马进行研磨过筛处理，置于近红外光谱仪下进行测定，然后真空保存这些茶叶粉末并冷藏于−4 ℃冰箱，用于后续内在品质的测定。

1.2 仪器与设备

烘焙笼：AK-HP020型，安溪县泰旺机械设备有限公司；研磨机：JYL-C012型，九阳股份有限公司；电子天平：FA124C型，上海力辰邦西仪器科技有限公司；离心机：TG-16G型，上海川一实验仪器有限公司；数显恒温水浴锅：HH-1型，上海力辰邦西仪器科技有限公司；紫外可见分光光度计：UV-1800型，日本岛津公司；分样筛：50目型，上海言锦丝网加工厂；近红外光谱仪：NIR2500型，上海复享光学股份有限公司。

1.3 内在品质测定及感官评审

对研磨并过筛之后的茶叶粉末进行内在品质测定。含水率参照Song等[22]的方法测定；游离氨基酸质量分数依据GB/T 8314—2013《茶 游离氨基酸总量的测定》测定；茶多酚质量分数和儿茶素质量分数依据GB/T 8313—2018《茶叶中茶多酚和儿茶素类含量的检测方法》测定；咖啡碱质量分数依据GB/T 8312—2013《茶 咖啡碱测定》测定。酚氨比（茶多酚质量分数与游离氨基酸质量分数的比值）参照葛诗蓓等[23]的方法计算。感官评审由5名专业评茶员根据GB/T 23776—2018《茶叶感官审评方法》对茶叶的5个项目打分：外形（20%）、汤色（5%）、香气（35%）、滋味（35%）以及叶底（5%），每项均按照百分制进行。计算5位评茶员的各项分数的平均值，再根据各项得分占比计算该茶叶的感官评审总得分。

1.4 近红外光谱测定

室温下，在光谱采集前对仪器进行预热，设置近红外光谱仪测量模式为吸收模式，并设置近红外光谱仪参数，积分时长为100 ms，每条光谱都经过32次扫描，最后保存的光谱数据为32次的平均光谱数据，每次扫描时间3.2 s。光谱平滑处理采用默认窗口设置，即窗口参数为0。于900~2 500 nm波段对研磨过筛后的茶叶粉末进行光谱信息采集，每个茶叶样品采集3条曲线，3次光谱平均值为该份样品的光谱曲线，共180条近红外光谱曲线。

1.5 近红外光谱的预处理与降维处理

采集到的近红外光谱中往往含有仪器和外界环境引起的噪声，如基线漂移和杂散光，这些噪声会干扰分析。因此通常对近红外光谱进行预处理分析，具体方法为：用Savitzky-Golay（S-G）平滑去除高频噪声并保留信号的特征细节；用标准变量变换（standard normal variate，SNV）减少因仪器响应或样品差异导致的基线漂移和强度变化；用MSC减少散射引起的光谱失真；用一阶导数（first derivative，FD）识别重叠峰或消除基线漂移；用二阶导数（second derivative，SD）识别精细结构和提高分辨率。

由于光谱数据中变量较多，且包含了许多与烘焙茶叶品质无直接关联的多余信息，因此有必要对光谱数据进行降维处理，以精简信息，提高分析效率。使用PCA、竞争性自适应重加权（competitive adaptive reweighed sampling，CARS）算法、SPA算法和随机蛙跳（random frog，RF）算法对近红外光谱数据进行降维处理。

1.6 模型建立

SVM作为一种监督学习算法，其核心原理为：在特征空间中寻找一个最优超平面，该超平面是最大化2类数据点之间的分隔间隔，同时能确保那些最接近此分隔边界的数据点（即支持向量）与超平面之间的距离最大化。BPNN是一种用于神经网络训练的模型，它通过计算损失函数相对于各个网络权重的梯度信息来调整这些权重，从而最小化预测误差。作为监督学习技术中的核心优化手段之一，BPNN模型在深度学习和机器学习模型的训练过程中发挥着举足轻重的作用，广泛应用于各类模型的性能提升与优化。为了区分铁观音乌龙茶的烘焙程度，本文使用这2种判别模型进行判别。

PLSR作为一种多变量统计分析方法，广泛应用于数据分析领域，尤其擅长处理包含众多自变量与因变量的复杂数据集。因此，该研究利用PLSR

构建了烘焙乌龙茶品质的预测模型。均方根误差（root mean square error，RMSE）和决定系数（coefficient of determination，R^2）是2种常用的模型评估指标，分别衡量了模型预测值与实际值之间的偏差程度以及模型对数据的拟合度，R^2越大、RMSE越小表明模型预测性能越好、预测效果越佳。通过计算训练集均方根误差（root mean square error of cross-validation，RMSECV）和训练集决定系数（determination coefficient of cross-validation，R^2_{CV}），可以评估预测模型的训练性能。对于测试集，采用测试集均方根误差（root mean square error of prediction，RMSEP）、R^2_P和剩余预测偏差（relative percentage deviation，RPD）作为评价指标。

1.7 数据处理与分析

使用Morpho 3.2软件采集近红外光谱数据，使用Excel和SPSS软件进行数据统计分析，使用Origin软件绘制图表，使用Matlab软件进行数据处理与建模。

2 结果与讨论

2.1 内在品质分析

随着乌龙茶烘焙程度的加深，茶叶内在品质会发生一定变化。游离氨基酸是茶叶中鲜味的主要来源；茶多酚是茶叶中的主要抗氧化物质，影响茶叶的颜色、口感和健康效应；儿茶素是一类主要的茶多酚，是导致茶叶带有苦涩口感的关键因素之一；而茶叶中的咖啡碱，作为一种生物碱，能够激发中枢神经系统的活性，具有提神醒脑的效果，并赋予茶叶轻微的苦味[24]。如图1所示，随着烘焙时间的延长，游离氨基酸、茶多酚及儿茶素的质量分数均呈下降趋势。此外，烘焙温度的升高及时间的延长加快了这些内在品质的下降速度。茶叶中游离氨基酸质量分数下降会导致茶汤的鲜味减弱，甘甜感降低。茶多酚和儿茶素质量分数下降会让茶汤的苦涩味和收敛性减弱，使得茶汤更柔和，但也可能失去一些风味的层次感。陈加友等[25]在研究三段式烘焙工艺对安溪铁观音茶叶品质的影响时，发现随着烘焙温度的升高及时间的延长，茶叶内氨基酸、咖啡碱及茶多酚均呈下降趋势，其中氨基酸质量分数的减少尤为突出，咖啡碱质量分数在烘焙过程中的变化趋势较小，为2.0%~2.4%。酚氨比是衡量茶叶风味与品质的关键指标之一，对于乌龙茶而言，酚氨比对维持口感平衡尤为重要。具体而言，酚氨比越高，茶叶的苦涩味与收敛性愈发显著；相反，酚氨比较低能增强茶叶的鲜甜度与柔和感。不同的酚氨比适合于不同的消费群体和饮用习惯，因此在乌龙茶的加工和品评中，调节酚氨比是关键环节[26]。当烘焙时间>300 min时，酚氨比急剧上升，是由于烘焙后期游离氨基酸质量分数急剧下降，茶叶品质受到影响。因此，根据所测内在品质的动态变化，将铁观音乌龙茶烘焙时间为（0，180］min、（180，300］min、（300，360］min时分别判定为烘焙不足、烘焙适度和烘焙过度。如表1所示，游离氨基酸、茶多酚、儿茶素和咖啡碱的组间/组内差异统计分析表明，各内在品质的组间差异较大，说明烘焙程度有区分效果，P<0.001说明具有统计学意义。如表2所示，不同烘焙程度的铁观音乌龙茶的感官评审结果表明，烘焙适度的铁观音乌龙茶的感官评审总得分最高，风味品质最优。

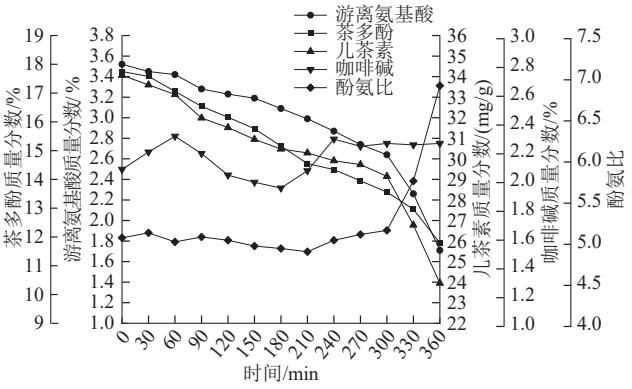


图1 烘焙期间茶叶内在品质的变化

Fig. 1 Changes in the intrinsic quality of tea leaves during baking

表1 3种烘焙程度茶叶品质ANOVA分析

Table 1 ANOVA analysis of tea leaves quality under three baking degrees

指标	组间/组内	均方	P
游离氨基酸	组间	29.042	<0.001
	组内	0.031	<0.001
茶多酚	组间	262.457	<0.001
	组内	0.354	<0.001
儿茶素	组间	707.092	<0.001
	组内	0.879	<0.001
咖啡碱	组间	0.309	<0.001
	组内	0.008	<0.001

表2　不同烘焙程度的铁观音乌龙茶的感官评审结果

Table 2　Sensory evaluation results of Tieguanyin oolong tea with different baking degrees　　单位：分

烘焙程度	外形	汤色	香气	滋味	叶底	总得分
烘焙不足	85	83	82	80	84	82.05
烘焙适度	91	89	92	90	81	90.40
烘焙过度	79	84	84	86	85	83.75

2.2　近红外光谱数据分析

在采集样品的近红外光谱数据时，获取了900~2 500 nm波段的近红外光谱曲线。然而，由于光谱两端区域受到噪声的显著干扰，需排除这些受影响的部分，因此选择了985~2 487 nm波段的近红外光谱曲线进行分析。不同烘焙程度的铁观音乌龙茶的原始近红外光谱曲线见图2(a)，其在1 200、1 400~1 800 nm和1 950 nm以及2 100 nm和2 300 nm处出现明显吸收峰，分别对应游离氨基酸(R—NH基团三级倍频)、咖啡碱(—OH)、儿茶素类(═C—H)的特征吸收，说明这些物质对铁观音乌龙茶品质的贡献较大[27]。

在光谱数据的收集过程中，会因环境因素(光线条件、温度状况、湿度水平)及人为因素的介入，导致原始红外光谱曲线呈现出众多不平滑的高峰与噪声杂点。因此分别使用S-G、SNV、MSC、FD和SD对原始近红外光谱曲线进行预处理，结果见图2(b)~(f)。

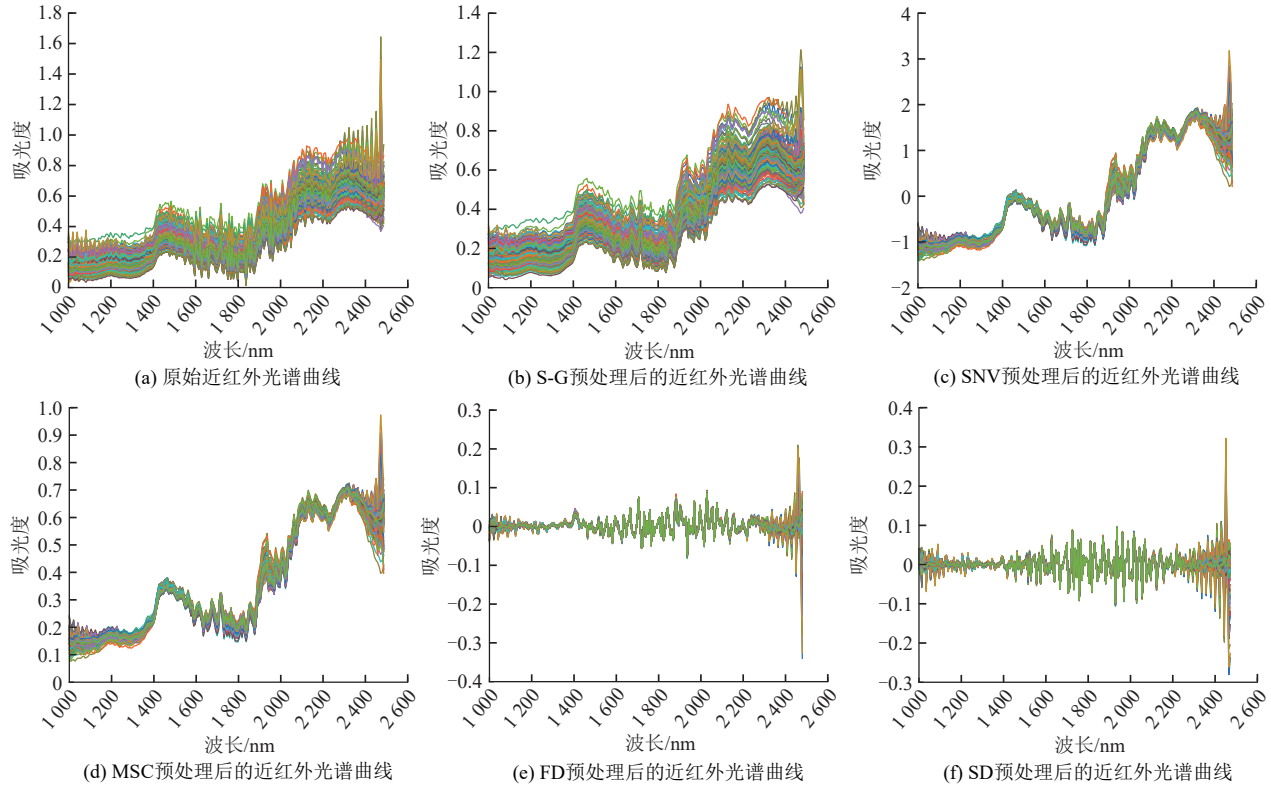


图2　原始近红外光谱曲线和不同方法预处理后的近红外光谱曲线

Fig. 2　Original near-infrared spectroscopy spectra and near-infrared spectroscopy spectra preprocessed with different methods

2.3　基于近红外光谱对铁观音乌龙茶进行不同烘焙程度的判别

由于原始近红外光谱曲线比较粗糙，因此选取经过S-G、SNV和MSC预处理后的近红外光谱曲线。在建立判别模型前，将180条近红外光谱曲线数据集以2∶1的比例划分为训练集和测试集，并使用SVM和BPNN构建判别模型对其进行比较，训练集和测试集的模型判别结果见表3。对于全波段近红外光谱，对近红外光谱进行预处理后，当使用SVM构建判别模型时，S-G预处理后的训练集和测试集准确率分别为98.33%和90.00%，SNV预处理后的训练集和测试集准确率分别为100.00%和96.67%，MSC预处理后的训练集和测试集准确率分别为100.00%和95.00%，SNV预处

理后的测试集准确率最高；当使用BPNN模型时，S-G预处理后的训练集和测试集准确率分别为99.44%和97.78%，SNV预处理后的训练集和测试集准确率分别为99.72%和98.33%，MSC预处理后的训练集和测试集准确率分别为99.17%和100.00%，MSC预处理后的测试集准确率最高。结果表明，SNV和MSC的预处理效果较好，这可能是由于SNV和MSC预处理能够有效校正近红外光谱数据中的各种系统误差、提高信噪比以及减少非线性效应，从而显著提升模型的判别效果，特别是在面对复杂、多变的近红外光谱数据时。付光明等[28]在对不同油分档次烤烟进行快速光谱鉴别时，发现MSC预处理结合最小二乘支持向量机（least squares support vector machine，LSSVM）构建的油分档次判别模型效果最好，提高了烤烟油分档次的判别效率。因此，后续进行降维分析前优先采用SNV和MSC这2种预处理方式。此外，使用SVM和BPNN判别模型的测试集准确率均在90.00%以上，但BPNN判别模型效果更优。

表3 全波段近红外光谱下的烘焙程度判别准确率

Table 3 Discrimination accuracy of baking degree based on full-band near-infared spectroscopy spectra 单位：%

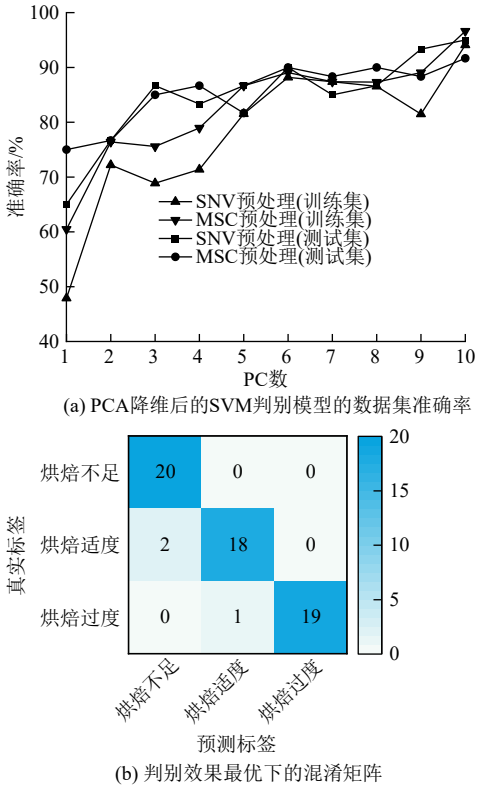
判别模型	预处理	训练集准确率	测试集准确率
SVM	S-G	98.33	90.00
	SNV	100.00	96.67
	MSC	100.00	95.00
BPNN	S-G	99.44	97.78
	SNV	99.72	98.33
	MSC	99.17	100.00

由于近红外光谱变量众多，且其中包含了与烘焙茶叶信息不相关的冗余变量，因此有必要对光谱数据进行降维处理，以优化其信息含量。针对铁观音乌龙茶烘焙程度的判别，使用PCA对光谱数据进行降维处理。对经过SNV和MSC预处理后的近红外光谱数据进行PCA降维处理，主成分（principal component，PC）数选择1~10，然后分别导入SVM模型进行判别，训练集与测试集的比值为2∶1，得出判别准确率，见图3（a）。从总体趋势上看，近红外光谱数据经SNV和MSC预处理后，当PC数为1~10时，SVM判别模型的测试集准确率均呈递增趋势；且当PC数到达5之后，其测试集准确率变化趋势比之前更平缓。当PC数为5时，经过SNV预处理后的测试集准确率为86.67%，经过MSC预处理后的测试集准确率为81.67%；当PC数为10时，经过SNV预处理后的测试集准确率为95.00%，经过MSC预处理后的测试集准确率为91.67%。

可以看出，经过SNV预处理，PC数为10时，茶叶烘焙程度判定的准确率最高。此时的测试集混淆矩阵见图3（b），在对烘焙适度的茶样进行判别时，发现有2个样品被误判为烘焙不足；在对烘焙过度的茶样进行判别时，发现有1个样品被误判为烘焙适度。这可能是由于烘焙过程中存在不均衡性，以及存在内在品质相近的样品，导致在个别样品中出现了误判的现象[29]。

当PC数为5和10时，使用BPNN判别模型对数据集进行分析，见表4。相较于SVM判别模型，PC为5时，BPNN判别模型的测试集准确率提高；PC为10时，经过SNV预处理的测试集准确率下降，而经过MSC预处理的测试集准确率提高。

总体而言，基于近红外光谱数据对铁观音乌龙茶烘焙程度进行判别，在全波段近红外光谱数据模型下，预处理为MSC、判别模型为BPNN时，



(a) PCA降维后的SVM判别模型的数据集准确率

(b) 判别效果最优下的混淆矩阵

图3 PCA降维后SVM判别模型性能

Fig. 3 Discrimination performance of the SVM model after dimensionality reduction by PCA

判别效果最优，测试集准确率为100.00%；而在使用PCA进行光谱数据降维后，预处理为SNV、PC数为10、判别模型为SVM以及预处理为MSC、PC数为5、判别模型为BPNN的2种模型处理下的判别效果均最优，测试集准确率为95.00%。

表4　PC数为5和10时BPNN判别模型的准确率

Table 4　Discrimination accuracy of the BPNN model with 5 and 10 PCs

判别模型	预处理	PC数	训练集准确率/%	测试集准确率/%
BPNN	SNV	5	87.50	91.67
		10	91.67	88.33
	MSC	5	93.33	95.00
		10	97.50	93.33

2.4　基于近红外光谱对铁观音乌龙茶进行品质预测

基于近红外光谱数据结合内在品质构建PLSR预测模型，对游离氨基酸、茶多酚、儿茶素和咖啡碱的质量分数进行预测。为了对比原始光谱和不同光谱预处理后PLSR预测模型的性能，采用S-G、SNV、MSC、FD和SD这5种光谱预处理方法，确定每个内在品质的最佳光谱预处理方法。PLSR作为一种多变量统计分析方法，在茶叶品质预测中得到了广泛应用[30-32]。PLSR预测模型的广泛适用性主要体现在它可以处理多变量间的复杂关系，特别是在结合近红外光谱等无损检测技术时，能够实现茶叶品质的快速、准确预测。PLSR预测模型展现出诸多优势，它不仅对数据噪声和变量共线性展现出高度的稳健性，而且降低了对样本量的要求，同时该模型还具备出色的可解释性。这些特性共同赋予了PLSR预测模型在茶叶品质评估研究中极高的应用价值。

如图4所示，游离氨基酸、茶多酚和儿茶素在原始光谱下的建模效果较优，游离氨基酸预测模型的R_P^2为0.947 3，茶多酚预测模型的R_P^2为0.941 3，儿茶素预测模型的R_P^2达到了0.948 4。咖啡碱在FD预处理后的建模较优，咖啡碱预测模型的R_P^2为0.709 5。这可能是因为经过平滑处理后的近红外光谱已经包含预测指标的大部分信息，使用一些预处理可能将有用信息筛选掉。

为了进一步提高模型的预测精度，可通过特征波长筛选算法滤除近红外光谱数据中的冗余变量[33-35]，因此采用3种算法（SPA、CARS和RF）进行特征波长的筛选。

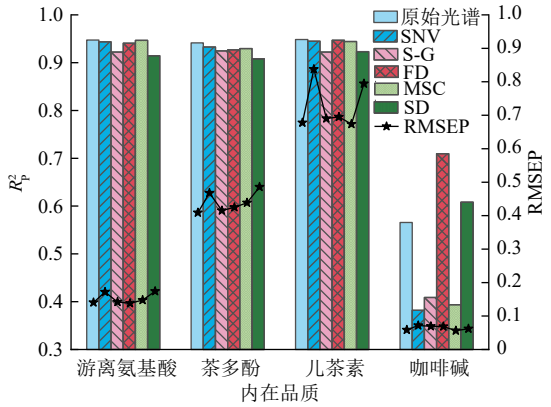


图4　不同预处理方法下的预测模型性能

Fig. 4　Performance of prediction models preprocessed with different methods

近红外光谱经最佳预处理后，使用不同特征波长筛选算法滤除冗余变量后的游离氨基酸、茶多酚、儿茶素和咖啡碱的PLSR预测模型性能见表5。3种特征波长筛选中，SPA算法均提高了PLSR预测模型的精度。游离氨基酸的最佳预测模型为原始光谱结合SPA算法，R_P^2为0.949 6，RMSEP为0.137 4；茶多酚的最佳预测模型为原始光谱结合SPA算法，R_P^2为0.944 3，RMSEP为0.398 3；儿茶素的最佳预测模型为原始光谱结合SPA算法，R_P^2为0.950 8，RMSEP为0.662 0；咖啡碱的最佳预测模型为FD预处理结合SPA算法，R_P^2为0.740 0，RMSEP为0.053 2。与全波段光谱构建的预测模型相比，使用最佳特征波长筛选算法筛选后的游离氨基酸、茶多酚、儿茶素和咖啡碱的预测模型的R_P^2提高，RMSEP降低。R_P^2分别提高了0.24%、0.32%、0.25%、4.30%，RMSEP分别降低了2.21%、2.59%、2.33%、5.03%。针对不同内在品质，选用最优光谱预处理方法及最优特征波长筛选算法组合，PLSR模型的预测性能得到提高，同时最大程度地保留了与内在品质紧密相关的特征波长信息。各个内在品质的最佳预测效果分析见图5。黑直线为$y=x$对照线，数据点距离该直线越近，说明预测效果越好，黑虚线为最佳预测模型下测试集预测方程的拟合线。因此，在最佳预测模型下，游离氨基酸的测试集预测方程为$y_1=0.950\ 0x_1+0.151\ 4$，相关系数为0.952 9；茶多酚的测试集预测方程为$y_2=1.004\ 1x_2+0.082\ 2$，相关系数为0.953 5；儿茶素的测试集预测方程为$y_3=0.960\ 8x_3+1.267\ 4$，相关系数为0.955 2；咖啡碱测试集预测方

表 5 不同特征波长筛选算法下的模型性能

Table 5 Performance of models with different characteristic wavelength screening algorithms

内在品质	最佳预处理	特征波长筛选算法	PC 数	训练集		测试集		
				R^2_{CV}	RMSECV	R^2_P	RMSEP	RPD
游离氨基酸	原始光谱	全波段光谱	15	0.942 1	0.121 8	0.947 3	0.140 5	4.394 0
		CARS	11	0.929 7	0.134 3	0.941 1	0.148 6	4.156 1
		SPA	13	0.914 0	0.148 4	0.949 6	0.137 4	4.493 9
		RF	13	0.944 9	0.118 9	0.947 6	0.140 2	4.403 8
茶多酚	原始光谱	全波段光谱	15	0.951 6	0.368 9	0.941 3	0.408 9	4.161 5
		CARS	12	0.948 5	0.380 2	0.931 5	0.441 7	3.851 9
		SPA	13	0.942 7	0.401 4	0.944 3	0.398 3	4.272 3
		RF	7	0.866 2	0.613 2	0.920 8	0.474 8	3.583 2
儿茶素	原始光谱	全波段光谱	15	0.951 2	0.562 2	0.948 4	0.677 8	4.439 9
		CARS	13	0.945 6	0.593 7	0.934 8	0.761 9	3.949 6
		SPA	15	0.951 9	0.558 1	0.950 8	0.662 0	4.545 5
		RF	12	0.955 0	0.539 8	0.942 9	0.713 2	4.219 3
咖啡碱	FD	全波段光谱	16	0.740 7	0.054 6	0.709 5	0.056 2	1.871 1
		CARS	18	0.671 0	0.061 4	0.660 3	0.060 8	1.730 1
		SPA	19	0.748 5	0.053 7	0.740 0	0.053 2	1.977 7
		RF	20	0.735 2	0.055 1	0.707 6	0.056 4	1.865 7

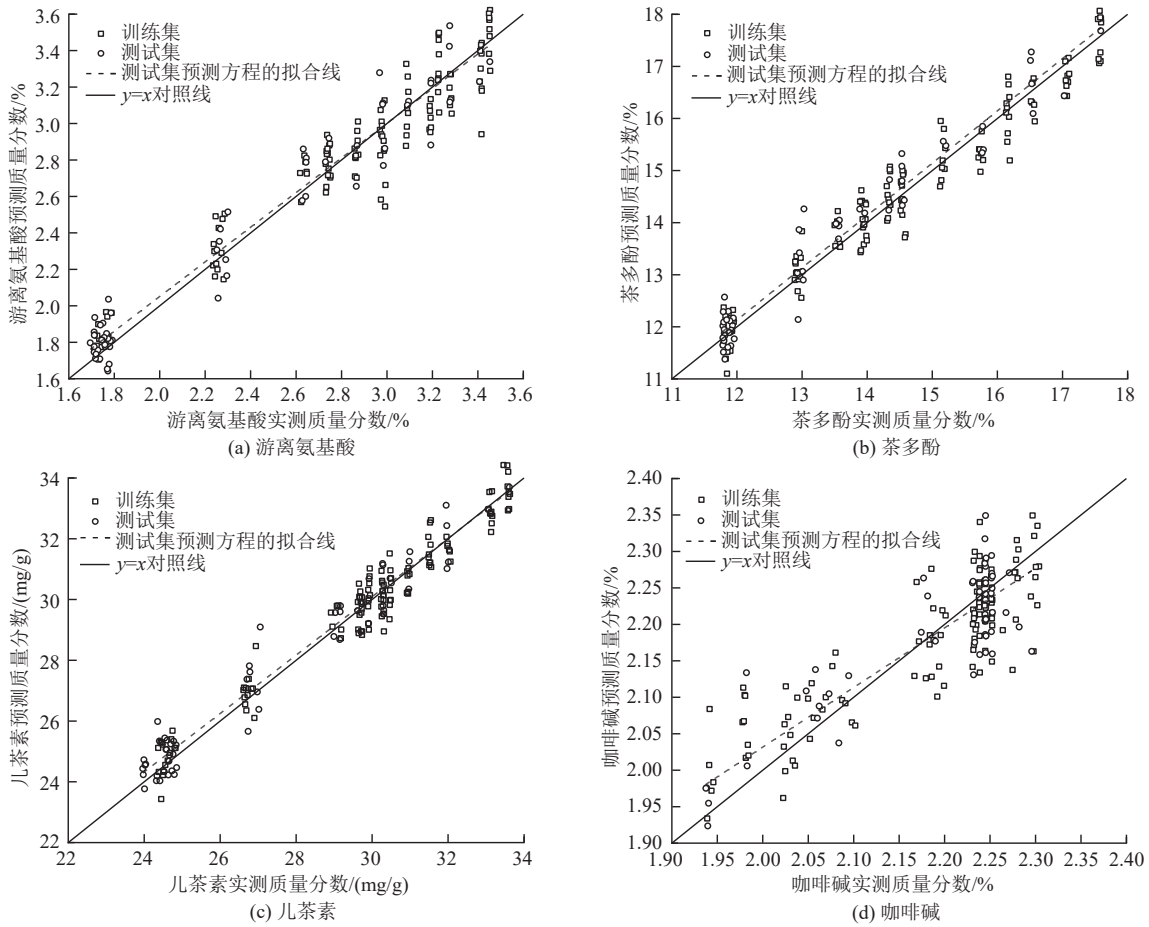


图 5 最佳预测模型下的各内在品质预测散点图

Fig. 5 Scatter plots of intrinsic quality properties predicted by the best prediction model

程为y_4=0.817 3x_4+0.397 1，相关系数为0.747 2。

3 结论

该研究利用近红外光谱技术对铁观音乌龙茶烘焙程度判别和内在品质预测进行可行性分析。结果表明，随着烘焙时间的延长和烘焙温度的升高，茶叶游离氨基酸、茶多酚和儿茶素的质量分数不断减少，咖啡碱质量分数在一定范围内波动，根据所测铁观音乌龙茶内在品质的动态变化以及感官评审结果（烘焙适度的铁观音乌龙茶感官评审总得分最高），铁观音乌龙茶烘焙程度可分为烘焙不足、烘焙适度、烘焙过度。基于近红外光谱对铁观音乌龙茶烘焙程度进行判别，在全波段光谱数据模型下，预处理为MSC、判别模型为BPNN时，判别效果均最优，测试集准确率均为100.00%。当预处理为SNV、PC数为10、判别模型为SVM以及预处理为MSC、PC数为5、判别模型为BPNN时，再使用PCA对近红外光谱数据进行降维时的判别效果均最优，测试集准确率均为95.00%。基于近红外光谱数据结合茶叶内在品质，构建PLSR预测模型。对比发现游离氨基酸的最佳预测模型为原始光谱结合SPA算法，其R_P^2为0.949 6；茶多酚的最佳预测模型为原始光谱结合SPA算法，其R_P^2为0.944 3；儿茶素的最佳预测模型为原始光谱结合SPA算法，其R_P^2为0.950 8；咖啡碱的最佳预测模型为FD预处理结合SPA算法，其R_P^2为0.740 0，表明SPA算法能有效提高模型预测的精度。

精确判断和控制乌龙茶的烘焙程度不仅对提升茶叶品质至关重要，还具有显著的经济和市场价值。利用近红外光谱技术结合机器学习，为烘焙乌龙茶的定性和定量分析提供了可靠的技术手段，也为茶叶品质预测开辟了研究方向。这种无损检测技术的应用，不仅能够减少生产过程中的人为误差，还为乌龙茶的标准化生产提供了科学依据。

参考文献

[1] BAG S，MONDAL A，MAJUMDER A，et al. Tea and its phytochemicals：hidden health benefits & modulation of signaling cascade by phytochemicals[J]. Food Chemistry，2022，371：131098.

[2] 黄福平，叶乃兴. 乌龙茶市场的现状与前景[J]. 茶叶科学简报，1993，34(4)：24-28.
HUANG F P，YE N X. Status and prospects of oolong tea market [J]. Acta Tea Sinica，1993，34(4)：24-28.(in Chinese)

[3] 柳镇章. 安溪铁观音产品品质及拼配、烘焙和贮存关键技术的研究[D]. 福州：福建农林大学，2022.

[4] CAO Q Q，FU Y Q，WANG J Q，et al. Sensory and chemical characteristics of Tieguanyin oolong tea after roasting[J]. Food Chemistry：X，2021，12：100178.

[5] 王飞，孙云，周子维，等. 不同焙火程度对闽北乌龙茶化学品质的影响[J]. 宁德师范学院学报(自然科学版)，2023，35(4)：407-413.
WANG F，SUN Y，ZHOU Z W，et al. Effects of different roasting degrees on the chemical quality of Minbei oolong tea[J]. Journal of Ningde Normal University (Natural Science)，2023，35(4)：407-413.(in Chinese)

[6] 谢善锦. 乌龙茶烘焙工艺现状及发展分析[J]. 福建茶叶，2022，44(5)：18-20.
XIE S J. Analysis on present situation and development of oolong tea baking technology [J]. Tea in Fujian，2022，44(5)：18-20.(in Chinese)

[7] DIAZ-OLIVARES J A，VAN NUENEN A，GOTE M J，et al. Near-infrared spectra dataset of milk composition in transmittance mode[J]. Data in Brief，2023，51：109767.

[8] WU X H，FANG Y H，WU B，et al. Application of near-infrared spectroscopy and fuzzy improved null linear discriminant analysis for rapid discrimination of milk brands[J]. Foods，2023，12(21)：3929.

[9] XIA H L，CHEN W，HU D，et al. Rapid discrimination of quality grade of black tea based on near-infrared spectroscopy (NIRS)，electronic nose (E-nose) and data fusion[J]. Food Chemistry，2024，440：138242.

[10] ZHANG Z Q，ZANG M W，ZHANG K H，et al. Effect of two types of thermal processing methods on the aroma and taste profiles of three commercial plant-based beef analogues and beef by GC-MS，E-nose，E-tongue，and sensory evaluation[J]. Food Control，2023，146：109551.

[11] LI H H，SUN X，PAN W X，et al. Feasibility study on nondestructively sensing meat's freshness using light scattering imaging technique[J]. Meat Science，2016，119：102-109.

[12] ZHANG Y，LI Z，WANG Q Y，et al. Rapid and visual evaluation the internal corruption of meat tissue by a designed near-infrared fluorescence probe with a broad pH response range[J]. Spectrochimica Acta Part A：Molecular and Biomolecular Spectroscopy，2023，302：123035.

[13] OUACHA E H，LEGHRIB R，KHATIB N，et al. Ultrasound monitoring of water content in ultra high temperature milk [J]. Instrumentation Mesure Métrologie，2023，22(4)：177-183.

[14] YAN S H. Evaluation of the composition and sensory

properties of tea using near infrared spectroscopy and principal component analysis [J]. Journal of Near Infrared Spectroscopy, 2005, 13(6): 313-325.

[15] TIAN Z X, TAN Z F, LI Y J, et al. Rapid monitoring of flavonoid content in sweet tea (*Lithocarpus litseifolius* (Hance) Chun) leaves using NIR spectroscopy[J]. Plant Methods, 2022, 18(1): 44.

[16] WANG S S, WU Z M, CAO C M, et al. Design and experiment of online detection system for water content of fresh tea leaves after harvesting based on near infra-red spectroscopy[J]. Sensors, 2023, 23(2): 666.

[17] REN G X, YIN L L, WU R, et al. Rapid detection of ash content in black tea using a homemade miniature near-infrared spectroscopy [J]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2024, 308: 123740.

[18] 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]. Infrared Physics & Technology, 2022, 127: 104418.

[19] JIA J M, ZHOU X F, LI Y, et al. Establishment of a rapid detection model for the sensory quality and components of Yuezhou Longjing tea using near-infrared spectroscopy [J]. LWT-Food Science and Technology, 2022, 164: 113625.

[20] WU J Z, ZAREEF M, CHEN Q S, et al. Application of visible-near infrared spectroscopy in tandem with multivariate analysis for the rapid evaluation of matcha physicochemical indicators[J]. Food Chemistry, 2023, 421: 136185.

[21] DING Y H, YAN Y L, LI J, et al. Classification of tea quality levels using near-infrared spectroscopy based on CLPSO-SVM[J]. Foods, 2022, 11(11): 1658.

[22] SONG F H, HAO X, LI Z F, et al. Monitoring the baking quality of Tieguanyin *via* electronic nose combined with GC-MS[J]. Food Research International, 2023, 165: 112513.

[23] 葛诗蓓,金迪迪,杨明来,等. 不同比例红蓝光对茶叶品质成分的影响与相关调控机理研究[J]. 浙江农业学报,2022,34(10):2105-2111.
GE S B, JIN D D, YANG M L, et al. Effects and regulation mechanism of different proportions of red and blue light on quality components in tea (*Camellia sinensis* L.) plant[J]. Acta Agriculturae Zhejiangensis, 2022, 34(10): 2105-2111.(in Chinese)

[24] 王秀英. 福建乌龙茶咖啡碱含量分析及烘焙工艺研究[D]. 福州:福建农林大学,2009.

[25] 陈加友,高飞,冉琴,等. 三段式烘焙工艺对安溪铁观音成茶品质的影响[J]. 宁德师范学院学报(自然科学版),2022,34(4):396-402.
CHEN J Y, GAO F, RAN Q, et al. Effects of different three-stage baking processes on the quality of Anxi Tieguanyin tea[J]. Journal of Ningde Teachers College (Natural Science), 2022, 34(4): 396-402.(in Chinese)

[26] LIU X B, LIU Y W, LI P, et al. Chemical characterization of Wuyi rock tea with different roasting degrees and their discrimination based on volatile profiles [J]. RSC Advances, 2021, 11(20): 12074-12085.

[27] 唐雪平. 基于化学分析与机器学习的铁观音茶叶品质评价体系[D]. 泉州:华侨大学,2020.

[28] 付光明,高子婷,杨建新,等. 基于近红外光谱的烤烟油分识别研究[J]. 河南农业大学学报,2024,58(4):583-591.
FU G M, GAO Z T, YANG J X, et al. Research on oil levels identification of flue-cured tobacco based on near infrared spectroscopy[J]. Journal of Henan Agricultural University, 2024, 58(4): 583-591.(in Chinese)

[29] ZHANG B, LI Z F, SONG F H, et al. Discrimination of black tea fermentation degree based on multi-data fusion of near-infrared spectroscopy and machine vision [J]. Journal of Food Measurement and Characterization, 2023, 17(4): 4149-4160.

[30] WU T, ZHANG H M, XIAO Y X, et al. Quantification of multiple elements in Anji white tea using hyperspectral imaging combined with machine learning regression [J]. Journal of Food Composition and Analysis, 2025, 142: 107520.

[31] WANG Y J, JIN S S, LI M H, et al. Onsite nutritional diagnosis of tea plants using micro near-infrared spectrometer coupled with chemometrics[J]. Computers and Electronics in Agriculture, 2020, 175: 105538.

[32] GUO Z M, CHEN Q S, CHEN L P, et al. Optimization of informative spectral variables for the quantification of EGCG in green tea using Fourier transform near-infrared (FT-NIR) spectroscopy and multivariate calibration [J]. Applied Spectroscopy, 2011, 65(9): 1062-1067.

[33] HUANG Y F, DONG W T, SANAEIFAR A, et al. Development of simple identification models for four main catechins and caffeine in fresh green tea leaf based on visible and near-infrared spectroscopy[J]. Computers and Electronics in Agriculture, 2020, 173: 105388.

[34] LUO W, TIAN P, FAN G Z, et al. Non-destructive determination of four tea polyphenols in fresh tea using visible and near-infrared spectroscopy [J]. Infrared Physics & Technology, 2022, 123: 104037.

[35] CHANDA S, HAZARIKA A K, CHOUDHURY N, et al. Support vector machine regression on selected wavelength regions for quantitative analysis of caffeine in tea leaves by near infrared spectroscopy [J]. Journal of Chemometrics, 2019, 33(10): e3172.

(责任编辑:闫林红)