

基于云端-互联便携式近红外技术现场快检西红花真伪

李庆^{1,2,3}, 闫晓剑⁴, 赵魁¹, 李兰^{2,3}, 彭善贵^{2,3}, 罗霄^{2,3}, 文永盛^{2,3}, 严铸云^{1*}

1. 成都中医药大学, 中药材标准化教育部重点实验室, 中药资源系统研究与开发利用省部共建国家重点实验室培育基地, 四川 成都 611137
2. 成都市食品药品检验研究院, 四川 成都 610045
3. 国家药品监督管理局中药材质量监测评价重点实验室, 四川 成都 610045
4. 成都虹微技术有限公司, 四川 成都 610041

摘要 利用云端-互联便携式近红外技术结合化学计量学对名贵药材西红花与其常见伪品(红花、玉米须、莲须、菊花、纸浆)和掺伪品进行现场快速真伪鉴别及掺伪量的定量预测。用移动手机控制的 PV500R-I 便携式近红外仪采集西红花与其伪品和掺伪品光谱数据。对原始光谱数据进行一阶导, 二阶导, 三阶导, 标准正态变量转换和光散射校正前处理。采用偏最小二乘判别分析分步建立西红花与其伪品、西红花与其掺伪品鉴别模型。结果表明, 一个最优模型可将西红花与其五类伪品彼此完全区分; 两个最优模型分步区分西红花与其五类掺伪品, 外部预测准确率最低为 93%, 西红花掺入红花、玉米须、莲须、菊花和纸浆的识别水平分别为 0.5%, 0.5%, 4.0%, 0.5% 和 0.5%。采用偏最小二乘回归对五类西红花掺伪品的掺伪量建立定量预测模型, 五个最优模型的外部预测相关系数范围为 0.920~0.999, RMSEP 范围为 0.005~0.044, 当西红花掺入红花、菊花、莲须、纸浆和玉米须的掺伪量大于 8% 时, 其外部预测相对误差分别低于 8%, 8%, 3%, 10% 和 5%, 表明最终模型能较好地预测西红花掺伪品的掺伪量。基于云端-互联便携式近红外光谱技术所建立的西红花真伪鉴别方法和掺伪品掺伪量预测方法快速准确, 经济环保, 能满足西红花现场快速无损真伪鉴别要求。

关键词 云端-互联便携式近红外技术; 化学计量学; 西红花; 伪品; 掺伪品

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)10-3029-09

引言

西红花为鸢尾科植物番红花 *Crocus sativus* L. 的干燥柱头, 具有活血化瘀、凉血解毒、解郁安神的功效^[1]。西红花是一种名贵中药材, 近年来, 存在较多的易混淆品和相似物染色后掺入西红花或冒充西红花销售^[2-4], 国内常见的西红花伪品主要有红花、菊花、莲须、玉米须、线状纸浆^[2]。目前其真伪鉴别方法主要包括显色反应^[5]、薄层色谱法^[5]、紫外分光光度法^[6]、高效液相色谱法^[5, 7]、质谱法^[6]和分子标记技术^[8]。现有的真伪鉴别方法存在前处理繁杂、使用有机溶剂、破坏样品、不能现场检验、检验成本高等弊端。急需开发一种经济、环保、无损伤样品的现场快检方法。

利用近红外光谱技术对西红花与其伪品进行鉴别的研究文献少见, 目前仅见 Eman Shawky 等利用台式近红外仪分

析西红花与其国外常见伪品的报道^[3], 尚未见利用该技术分析国内常见西红花伪品菊花、玉米须、莲须、线状纸浆的报道。云端-互联的便携式近红外技术是通过手机蓝牙将可随身携带的便携式近红外光谱仪与云端大数据相关联, 与台式近红外技术相比, 操作更加简单快速, 可用于现场快检, 因此研究中尝试利用该技术构建西红花与其国内常见伪品的定性鉴别和掺伪量的定量测定方法, 以期向西红花提供一种新的不损伤样品的现场快检方法。

1 实验部分

1.1 样品

共收集 60 份西红花样品, 其详细信息见表 1, 所有样品经成都中医药大学严铸云教授鉴定西红花 *Croci stigma* 为正品; 红花、菊花、莲须、玉米须和线状纸浆样品详细信息同

收稿日期: 2020-05-11, 修订日期: 2020-07-15

基金项目: 四川省重大科技专项(2018TZDZX0007)资助

作者简介: 李庆, 1980年生, 成都中医药大学、食品药品检验研究院在读博士研究生 e-mail: liqqin2001@outlook.com

* 通讯联系人 e-mail: yanzhuyun@cdutcm.edu.cn

样见附录表 1。西红花和常见染色伪品见附录图 1。

表 1 西红花及其伪品的来源和数量信息
Table 1 The source and number information of saffron and its adulterants

样品名称	来源	数量
西红花	伊朗	40
西红花	上海	5
西红花	浙江建德	5
西红花	安徽亳州	5
西红花	河南永城	5
红花	成都市荷花池中药材市场	43
菊花	成都市荷花池中药材市场	20
莲须	成都市荷花池中药材市场	20
玉米须	农贸市场	20
纸浆	自制	20

掺伪品制作：在西红花中按质量比掺入伪品，掺入百分比范围为：0.5%，1%，2%，3%，4%，5%，6%，7%，8%，9%，10%，20%，30%，40%和 50%，共 15 份，平行制备两份，一份为训练集，一份为测试集。

1.2 云端-互联 PV500R-I 便携式近红外光谱技术

云端-互联便携式近红外光谱系统主要由三部分构成：(1) 无线 PV500R-I 便携式近红外仪(长虹科技有限公司，中国)，性能参数：长宽高为 110 mm×70 mm×70 mm；重量为 400 g；光谱分辨率：20 nm；波长重复性为±2 nm；超大光斑 70 mm×70 mm。(2) 移动手机。(3) 云端数据库。其工作示意图见图 1，便携式近红外光谱仪通过蓝牙与移动手机连接，将扫描的样品光谱数据经移动手机上传至云端，在云端通过已建立的预测模型快速计算未知样本类别或含量，并将结果迅速反馈至手机端。

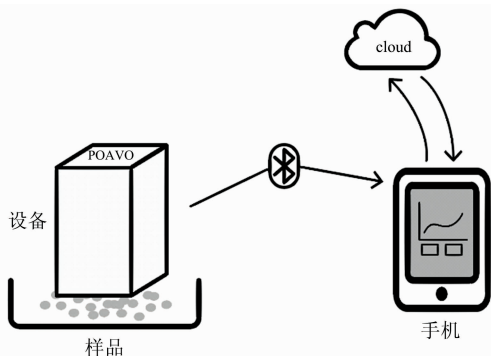


图 1 云端-互联 PV500R-I 便携式近红外仪主要工作示意图
Fig. 1 Main work diagram of cloud connected PV500R-I portable near infrared spectrometer

1.3 近红外光谱采集

便携式近红外光谱仪经校正后，将仪器直接贴于样品表面，在 1 350~1 850 nm 波长范围内采集光谱图。每个样品重复测定 6 次，求平均光谱用于建模。

1.4 样品的选择和光谱数据前处理

为使模型稳健，Kennard-Stone selection algorithm^[9]用

于选择训练集(三分之二的样本量)和预测集(三分之一的样本量)。由于各类别样本量有一定差异，为避免产生更大的不平衡样本，先对每类样本选择训练集和预测集，后将所有类别样本的训练集和预测集各自相加，得到最后的训练集和预测集。

标准正态变量转换(standard normal variable transformation, SNV)和光散射校正(multiplicative scatter correction, MSC)是散射校正前处理技术，常用于消除颗粒分布不均匀和粒径散射造成的影响，同时也可消除光谱扫描时样品引起的光谱平移和随机噪声的影响，从而提高模型预测能力^[10]。采用一、二、三阶导数滤波器提高光谱分辨率，消除原始红外光谱中的基线漂移和背景^[11]。

1.5 基于 PLS-DA 的西红花真伪鉴别模型的建立

在全波段(1 350~1 850 nm)条件下，利用偏最小二乘判别分析(PLS-DA)建立西红花与其常见伪品的真伪鉴别模型。用 7 折交叉验证的交叉验证均方根误差(RMSECV)的最小值确定最适隐变量数(LVs)。最优模型选取原则： $R^2 X$ ， $R^2 Y$ ， Q^2 ^[12]，内部预测准确率和外部预测准确率的值越大，模型性能越好。使用 Simca(version 13.0, Umetrics, Sweden)软件完成 PLS-DA 模型的建立。

1.6 基于 PLSR 的西红花掺伪品掺伪量的定量预测模型的建立

在全波段条件下，建立五种西红花掺伪品掺入量的偏最小二乘回归(PLSR)定量预测模型。根据变量与模型性能相关的回归系数大小，选择最适回归系数及相应的重要波段对模型性能进行改进。模型评价指标包括：决定系数(R^2)，均方根误差(RMSE)，RMSECV，预测均方根误差(RMSEP)， R^2 越大，RMSE，RMSECV 和 RMSEP 越小，同时 RMSECV 与 RMSEP 之间差异越小，表明模型性能越好。使用 Unscrambler(version 7.5, CAMO ASA, Norway)软件完成 PLSR 模型的建立。

2 结果与讨论

2.1 西红花与其伪品的光谱特征

图 2(a)和(b)分别是西红花与其伪品，西红花与其掺伪品在 1 350~1 850 nm 范围内的原始光谱图。该光谱范围为 C—H，O—H，N—H 等基团伸缩振动的一级倍频区域^[13]。直观上看，西红花与其伪品之间的光谱存在一定交叉，但各自有集中分布范围，由上至下依次为红花、莲须、西红花、菊花或玉米须，表明西红花与伪品之间的原始光谱之间彼此存在一定差异，这主要是由于各自的成分组成不一致导致。比较特殊的是纸浆的近红外光谱图(黄色曲线)，其光谱曲线在 1 560~1 660 nm 范围有一个突降的过程，产生原因可能是由于纸浆为工业产品，其成分与其他植物源样品差异明显。西红花与其伪品的光谱特点为西红花鉴别和定量分析提供了光谱基础。西红花与其掺伪品在光谱图上同样存在各自的集中区域，由上至下依次为红花、玉米须或西红花或纸浆、莲须、菊花的近红外光谱，但由于西红花在掺伪品中占比较大，各类间的光谱曲线彼此重叠更加严重。

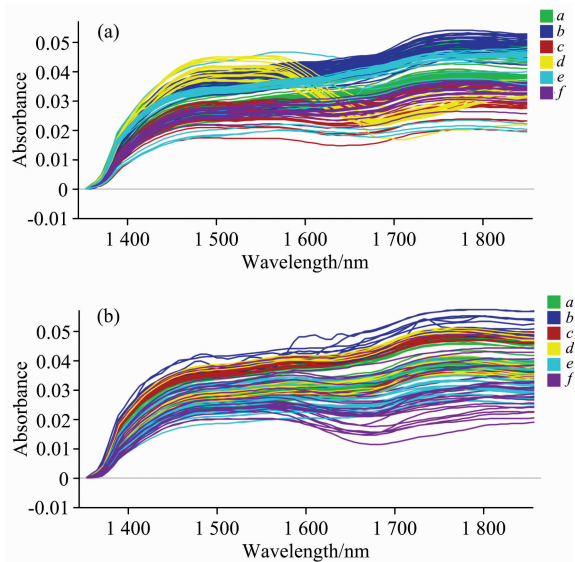


图 2 西红花与其伪品 (a), 西红花与其掺伪品 (b) 的原始光谱图

a, b, c, d, e, f 分别为西红花, 红花, 玉米须, 纸浆, 莲须, 菊花 (以下图表中 *a, b, c, d, e, f* 均代表相同类别)

Fig. 2 The raw spectra of saffron and its adulterants (a), saffron and its adulterated saffron (b)

The *a, b, c, d, e, f* are saffron, carthami flos, corn silk, pulp, nelumbinis stamen, chrysanthemi flos, respectively (*a, b, c, d, e, f* in the following chart represent the same category)

2.2 PLS-DA 真伪鉴别模型

建模过程中发现同时有效区分西红花与其伪品和掺伪品十分困难, 因此, 分步为西红花与其伪品, 西红花与其掺伪品建立 2 个以上的识别模型。

2.2.1 西红花与其伪品识别模型

在全波段条件下, 用西红花与其伪品的原始数据和经五种数据处理方法(一阶导, 二阶导, 三阶导, MSC, SNV)处理后的数据建立六个 PLS-DA 模型, 其结果见表 2, 图 3, 附录表 1 和附录表 2。由附录表 1 和表 2 可知, 原始数据所建

立的模型性能最优($R^2X=1$, $R^2Y=0.841$, $Q^2=0.733$, $LV=13$), 其对训练集和测试集的预测准确率均为 100%, 而数据经前处理后所建立的模型并未改善, 这可能是光谱预处理时虽然降低了噪音, 提高了信噪比, 但也丢失了更为重要的信息。附录表 2 为最优模型外部预测误分类表, 可知六类样本彼此间均能 100% 区分。图 3(a) 为主成分 1 和 2 绘制的二维得分图, 可知主成分 1 对西红花、纸浆、红花彼此之间的准确区分起主要作用, 且纸浆与其他五类样本均能明显区分, 而主成分 2 对红花与玉米须, 红花与菊花之间的准确区分起主要作用。同样, 由图 3(a), (c) 和 (d) 可知, 莲须与西红花、玉米须与菊花、红花与莲须均能有效区分。图 3(e) 为模型的置换检验结果($R^2=0.105$, $Q^2=-0.341$), 可知所有的蓝色 Q^2 值均处于绿色 R^2 值的下方, 表明模型可靠^[3]。

2.2.2 西红花与其掺伪品识别模型

用上述同样方法建立西红花与其掺伪品的六个 PLS-DA 模型, 结果见表 2、附录图 2、附录表 1 和附录表 2。附录表 1 和表 2 表明原始数据所建立的识别模型最优($R^2X=1$, $R^2Y=0.739$, $Q^2=0.527$, $LV=17$), 其对训练集和测试集的预测准确率为分别为 91% 和 89%。置换检验结果($R^2=0.237$, $Q^2=-0.663$) [见附录图 2(d)] 表明模型可靠。分析附录表 2 可知, 除莲须掺伪品外, 模型能同时将其其他四类掺伪品与西红花完全区分, 附录图 2(a) 表明主成分 1 与主成分 2 可将西红花与玉米须和纸浆完全区分, 附录图 2(c) 表明主成分 1 和主成分 4 可将西红花与菊花和红花完全区分, 表明四种西红花掺伪品的掺伪量识别水平可低至 0.5%, 低于文献掺伪量识别最低水平(1% 或 5%)^[3-4]。莲须有四个掺伪量低的样品误判为西红花 [见附录图 2(c) 和附录表 3], 其识别水平为 4%。附录图 2(b) 和附表 2 表明玉米须与纸浆之间可完全区分, 玉米须和纸浆掺伪品两者识别率高于 93%, 且其他三类掺伪品未与两者相混淆, 表明特异性高。但红花和莲须掺伪品之间互有误判, 两者的识别率分别为 80% 和 60%。另外, 尽管菊花掺伪品的识别率为 100%, 但有三个其他掺伪品误判为菊花, 在五类掺伪品中, 除莲须外, 菊花的特异性最差。

表 2 在全波段条件下, 西红花与其伪品, 西红花与其掺伪品, 三类西红花掺伪品的最优 PLS-DA 模型结果, 详细结果见附录表 1

Table 2 The best PLS-DA model results of saffron and its adulterants, saffron and its adulterated saffron, three kinds of adulterated saffron under the condition of full-spectra. Detailed results can be seen in table 1 of Supplementary Material

模型	数据	LV	R^2X	R^2Y	Q^2	内部预测准确率/%	外部预测准确率/%
西红花与其伪品	原始数据	13	1	0.841	0.733	100	100
西红花与其掺伪品	原始数据	17	1	0.739	0.527	91	89
三类西红花掺伪品	原始数据	5	1	0.632	0.554	98	96

进一步为菊花、红花和莲须三类西红花掺伪品建立识别模型。结果见表 2、附录图 3、附录表 1 和附录表 2。由附录表 1 和表 2 可知, 原始数据建立的模型最优($R^2X=1$, $R^2Y=0.632$, $Q^2=0.554$, $LV=5$), 其对训练集和测试集的预测准确率分别为 98% 和 96%, 错误分类表(见附录表 2)表明 3 类掺伪品的识别率均在 93% 以上。附录图 3(a) 表明菊花掺

伪品能与红花和莲须完全区分, 附录图 3(b) 表明有一个莲须掺伪品误判为红花, 与附录表 2 中的结果一致。置换检验结果($R^2=0.116$, $Q^2=-0.302$) [见附录图 3(c)] 表明模型可靠。

2.3 五类西红花掺伪品掺伪量的 PLSR 模型

在全波段条件下, 对红花、菊花、莲须、玉米须和纸浆五种西红花掺伪品掺伪量的原始数据和经五种数据前处理方

法处理后的数据各建立六个 PLSR 定量预测模型, 其结果见表 3、图 4 和附录表 3。由附录表 3 和表 3 可知, 仅红花的最优模型为原始光谱数据提供, 其预测集 R^2 为 0.920, RMSEP 为 0.044。其他四类掺伪品掺伪量的最优定量模型均由一阶导或二阶导提供, 其预测集 R^2 均高于 0.924, RMSEP 低于 0.041, 表明原始光谱数据经一阶导或二阶导前处理后, 更易获得最优模型。同时, 模型的 RMSECV 值与 RMSEP 值之间差异较小, 表明五个模型可靠, 不存在过拟合。图 4 表

明五类西红花掺伪品外部预测集的掺伪量参考值和预测值所对应的点均匀分布于外部预测曲线两侧, 但当掺伪量范围低于 9% 时, 红花、纸浆、菊花、玉米须和莲须的预测值与真实值相对误差分别高达 372%, 203%, 61%, 42% 和 23%, 表明模型不适合预测掺伪量低的掺伪品; 当掺伪品掺伪量大于 8% 时, 红花、菊花、莲须、纸浆和玉米须的相对误差分别低于 8%, 8%, 3%, 10% 和 5%, 表明模型对五类掺伪品掺伪量能较好或很好的预测。

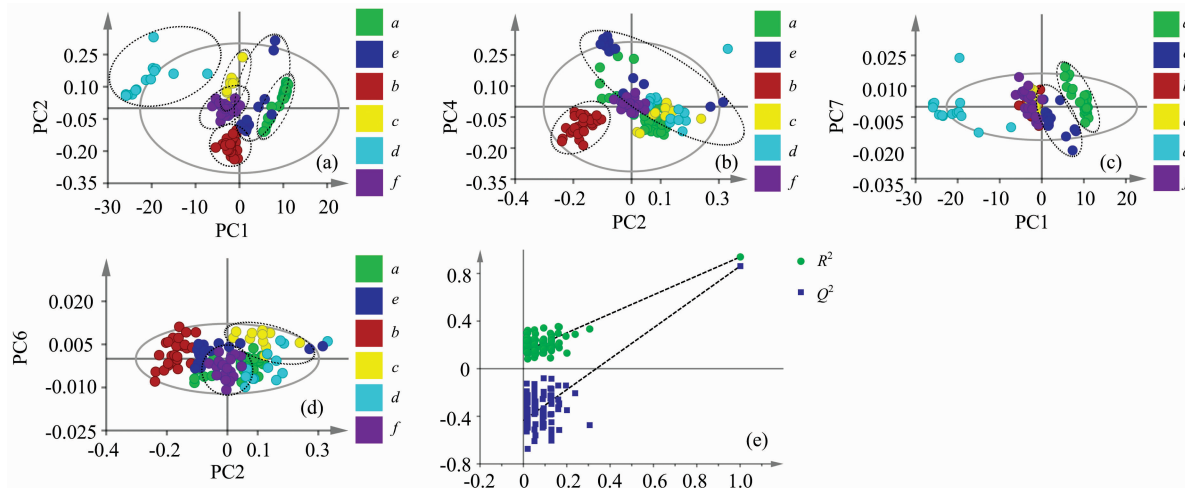


图 3 西红花与其伪品的真伪鉴别 PLS-DA 模型结果

(a), (b), (c), (d) 分别为主成分 1 和 2, 主成分 2 和 4, 主成分 1 和 7, 主成分 2 和 6 所绘制的二维得分图; (e) 为置换检验结果

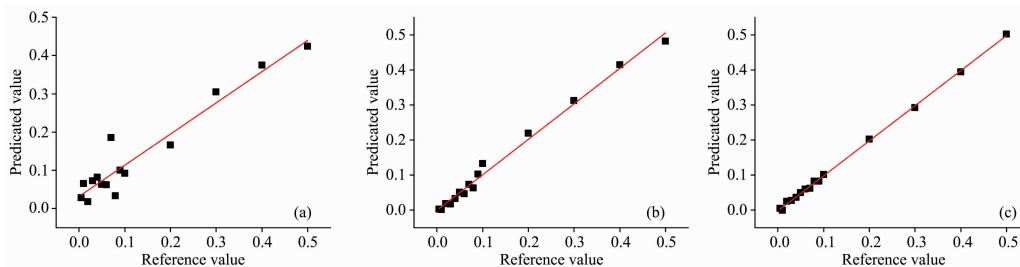
Fig. 3 The PLS-DA model results of saffron and its adulterants

(a), (b), (c), (d) are the two-dimensional score graph drawn by the principle components 1 and 2, 2 and 4, 1 and 7, 2 and 6, respectively, (e) is the result of permutation test

表 3 在全波段条件下, 红花、菊花、莲须、纸浆、玉米须五类西红花掺伪品的掺伪量的最优 PLSR 定量预测模型结果, 详细结果见附录表 3

Table 3 The best PLSR quantitative prediction model results of five kinds of adulterated saffron, i. e. carthami flos, chrysanthemi flos, nelumbinis stamen, pulp, corn silk, under the condition of full-spectra. Detailed results can be seen in table 3 of Supplementary Material

伪品	数据前处理	LV	校正集		交叉验证集		预测集	
			RMSCE	R^2	RMSECV	R^2	RMSEP	R^2
红花	原始数据	3	0.042	0.926	0.045	0.919	0.044	0.920
菊花	2 阶导	2	0.013	0.993	0.017	0.989	0.014	0.993
莲须	1 阶导	2	0.005	0.999	0.006	0.998	0.005	0.999
纸浆	1 阶导	3	0.021	0.979	0.036	0.948	0.041	0.924
玉米须	2 阶导	2	0.012	0.993	0.015	0.991	0.013	0.992



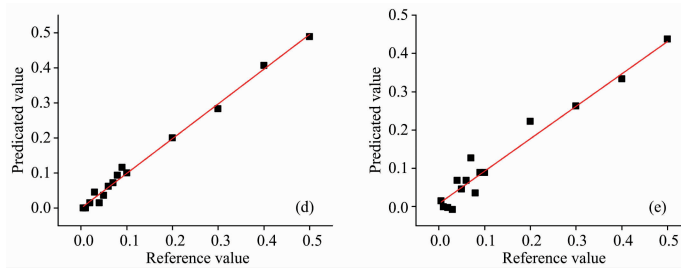


图 4 红花(a), 菊花(b), 莲央(c), 纸浆(d), 玉米须(e)五类西红花掺伪品的掺伪量的最优 PLSR 定量预测模型的外部预测结果

Fig. 4 The external prediction results of the optimal PLSR quantitative prediction model for the adulterants of carthami flos (a), chrysanthemi flos (b), nelumbinis stamen (c), pulp (d) and corn silk (e)

从全波段中选取重要变量建模通常能改善模型性能, 但仅红花掺伪品掺伪量的最优定量模型通过筛选重要变量, 其模型性能稍有改善, 其他四类的定量模型经变量筛选后所得模型性能反而变差, 原因可能是: (1) 筛选的重要变量缺乏代表性, 一些次要变量在建模过程中同样不可或缺; (2) 样品成分复杂, 导致一些重要特征变量被掩盖; (3) 便携式近红外仪波段较窄(1 350~1 850 nm), 缺失了其他特征区, 如组合频区。附录图 4(a)为红花的变量与模型性能之间的回归系数图, 可知以 ± 0.01 为界, 其重要变量为 1 350~1 057, 1 383~1 398, 1 653~1 749 和 1 797~1 850 nm, 其模型性能(RMSECV=0.043, RMSEP=0.043)较原始模型(RMSECV=0.045, RMSEP=0.044)略有改善。

3 结 论

应用云端-互联 PV500R-I 便携式近红外光谱技术结合化

学计量学, 建立一个最优 PLS-DA 模型能将西红花与其伪品完全区分; 两个最优 PLS-DA 模型可使西红花与其掺伪品预测准确率高于 93%, 掺伪量识别水平低至 0.5%~4%; 为五类掺伪品掺伪量建立的五个最优 PLSR 定量预测模型, 其外部预测相关系数范围为 0.920~0.999, RMSEP 范围为 0.005~0.044, 当掺伪量大于 8% 时, 定量预测模型能较好或很好地预测掺伪品掺伪量。基于云端-互联便携式近红外光谱技术建立的西红花真伪鉴别和掺伪量定量预测方法具有较好的准确性和可靠性, 与常用的台式近红外技术相比, 该技术将云端(大数据)-移动手机-便携式近红外仪一体化, 可使样品测试和数据分析接近同步完成, 操作更加智能化, 且可用于现场快检, 这也为其他名贵中药材的快速无损检验提供参考。

附 录

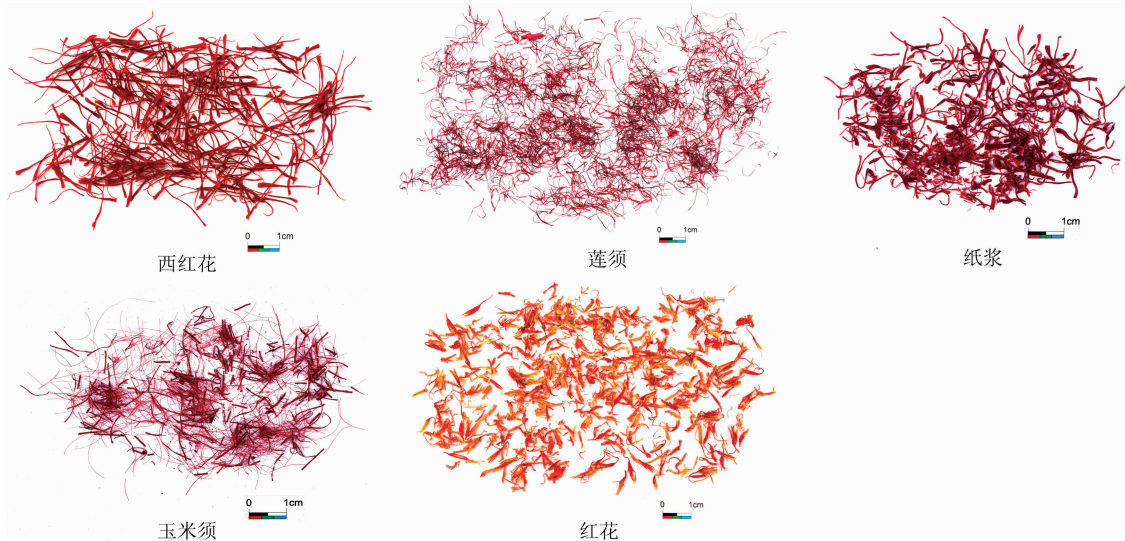


图 1 西红花和染色伪品照片

Fig. 1 Photos of saffron and dyed adulterants

表 1 在全波段条件下,使用原始数据和经 5 种不同的数据前处理所得数据建立的西红花与其伪品,西红花与其掺伪品,3 类西红花掺伪品的真伪鉴别 PLS-DA 模型结果,黑体字表示最优模型

Table 1 PLS-DA model results of saffron and its adulterants, saffron and its adulterated saffron, three kinds of adulterated saffron developed using original data and datas obtained by five different data preprocessing under the condition of full-spectra. Bold characters represent the optimal model

模型	数据	LV	$R^2 X$	$R^2 Y$	Q^2	内部预测准确率/%	外部预测准确率/%
西红花与其伪品	原始数据	13	0.1	0.841	0.733	100	100
	1 阶导	8	0.999	0.713	0.667	98	100
	2 阶导	6	0.965	0.650	0.580	97	100
	3 阶导	10	1	0.689	0.587	98	98
	MSC	7	0.878	0.698	0.578	98	98
	SNV	12	1	0.772	0.653	99	98
西红花与其掺伪品	原始数据	17	1	0.739	0.527	91	89
	1 阶导	7	0.989	0.460	0.368	81	79
	2 阶导	6	0.879	0.383	0.276	70	70
	3 阶导	10	0.936	0.480	0.284	79	74
	MSC	7	0.995	0.387	0.191	78	77
	SNV	7	0.991	0.396	0.259	77	74
三类西红花掺伪品	原始数据	5	1	0.632	0.554	98	96
	1 阶导	8	0.463	0.459	0.337	80	76
	2 阶导	4	0.632	0.675	0.539	93	90
	3 阶导	6	0.835	0.690	0.506	93	90
	MSC	5	0.654	$R^2 Y$	0.333	64	64
	SNV	4	0.633	0.841	0.336	67	64

表 2 西红花与其伪品,西红花与其掺伪品,三类西红花掺伪品的最优 PLS-DA 模型的误分类表

Table 2 The misclassification table of best PLS-DA model results of saffron and its adulterants, saffron and its adulterated saffron, three kinds of adulterated saffron

最优模型	Members	Correct	a	b	c	d	e	f	No class
西红花与伪品	a	20	100	20	0	0	0	0	0
	b	9	100	0	9	0	0	0	0
	c	14	100	0	0	14	0	0	0
	d	6	100	0	0	0	6	0	0
	e	6	100	0	0	0	0	6	0
	f	6	100	0	0	0	0	0	6
	No class	0	0	0	0	0	0	0	0
西红花与掺伪品	a	15	100	15	0	0	0	0	0
	b	15	80	0	12	0	0	2	1
	c	15	100	0	0	15	0	0	0
	d	15	93	0	0	0	14	0	1
	e	15	60	4	1	0	0	9	1
	f	15	100	0	0	0	0	0	15
	No class	0	0	0	0	0	0	0	0
三类西红花掺伪品 (菊花、红花和莲须)	a	/	/	/	/	/	/	/	/
	b	15	100	/	15	/	/	0	0
	c	/	/	/	/	/	/	/	/
	d	/	/	/	/	/	/	/	/
	e	15	93	/	1	/	/	14	0
	f	15	100	/	0	/	/	0	15
	No class	0	0	/	0	/	/0	0	0

注:“/”: a(西红花), c(玉米须)和 d(纸浆)三类样品未参与建模,其相应位置标注“/”

Note:“/”: The a (saffron), c (corn silk) and d (pulp) samples did not participate in modeling, their corresponding positions were marked “/”

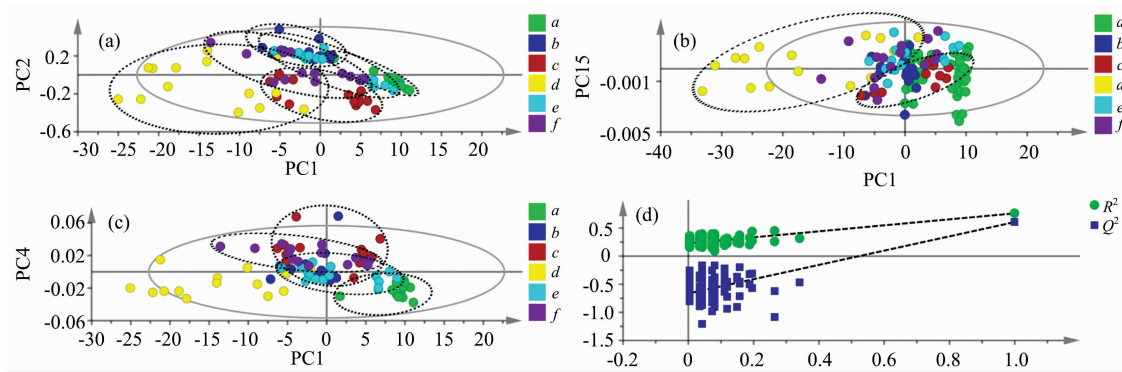


图 2 西红花与其掺伪品的真伪鉴别 PLS-DA 模型结果

(a), (b), (c) 分别为主成分 1 和 2, 主成分 1 和 15, 主成分 2 和 6 所绘制的二维得分图; (d) 为置换检验结果

Fig. 2 The PLS-DA model results of saffron and its adulterated saffron

(a), (b), (c) are the two-dimensional score graph drawn by the principle components 1 and 2, 1 and 15, 1 and 4, respectively, (d) is the result of permutation test

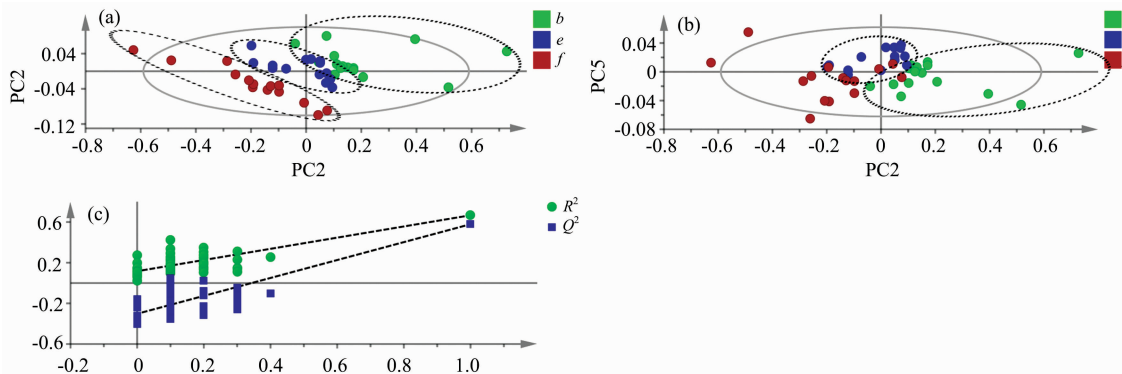


图 3 菊花、红花、莲须的三类西红花掺伪品之间的 PLS-DA 鉴别模型结果

(a), (b) 分别为主成分 1 和 2, 主成分 1 和 4 所绘制的二维得分图; (c) 为置换检验结果

Fig. 3 The PLS-DA model results of three kinds of adulterated saffron:

Chrysanthemum flos, *Carthami flos*, *Nelumbinis stamen*

(a) and (b) are the two-dimensional score graph drawn by the principle components 1 and 2, 1 and 4, respectively, (c) is the result of permutation test

表 3 在全波段条件下,使用原始数据和经 5 种不同的数据前处理所得数据建立的红花、菊花、莲须、纸浆、玉米须五类西红花掺伪品的掺伪量的 PLSR 定量预测模型结果,黑体字表示最优模型

Table 3 The PLSR quantitative prediction model results of five kinds of adulterated saffron, i. e. *carthami flos*, *chrysanthemum flos*, *nelumbinis stamen*, pulp, corn silk, were developed using original data and data obtained by five different data preprocessing under the condition of full-spectra. Bold characters represent the optimal model

伪品	数据前处理	LV	校正集		交叉验证集		预测集	
			RMSCE	R ²	RMSECV	R ²	RMSEP	R ²
红花	原始数据	3	0.042	0.926	0.045	0.919	0.044	0.920
	1 阶导	2	0.047	0.898	0.049	0.868	0.051	0.864
	2 阶导	6	0.049	0.875	0.057	0.812	0.0524	0.857
	3 阶导	2	0.054	0.849	0.056	0.821	0.055	0.822
	MSC	3	0.052	0.832	0.057	0.821	0.053	0.832
	SNV	3	0.052	0.831	0.054	0.828	0.053	0.831

续表 3

菊花	原始数据	2	0.013	0.992	0.019	0.986	0.013	0.992
	1 阶导	2	0.013	0.993	0.015	0.989	0.013	0.993
	2 阶导	2	0.013	0.993	0.017	0.988	0.014	0.991
	3 阶导	2	0.015	0.989	0.021	0.982	0.019	0.983
	MSC	6	0.013	0.992	0.042	0.928	0.047	0.977
	SNV	4	0.017	0.986	0.040	0.934	0.040	0.924
莲须	原始数据	3	0.049	0.890	0.064	0.834	0.058	0.845
	1 阶导	2	0.005	0.999	0.006	0.998	0.005	0.999
	2 阶导	2	0.005	0.999	0.007	0.998	0.006	0.998
	3 阶导	2	0.005	0.999	0.007	0.998	0.011	0.995
	MSC	3	0.062	0.822	0.126	0.363	0.064	0.811
	SNV	4	0.010	0.995	0.017	0.989	0.029	0.960
玉米须	原始数据	3	0.126	0.267	0.144	0.158	0.125	0.278
	1 阶导	2	0.015	0.990	0.017	0.988	0.014	0.990
	2 阶导	2	0.012	0.993	0.015	0.991	0.013	0.992
	3 阶导	2	0.016	0.988	0.019	0.984	0.016	0.988
	MSC	2	0.040	0.927	0.079	0.746	0.058	0.842
	SNV	5	0.017	0.986	0.070	0.801	0.035	0.941
纸浆	原始数据	5	0.060	0.831	0.136	0.258	0.095	0.580
	1 阶导	3	0.021	0.979	0.036	0.948	0.041	0.924
	2 阶导	3	0.036	0.937	0.046	0.916	0.051	0.877
	3 阶导	4	0.037	0.937	0.046	0.916	0.051	0.876
	MSC	2	0.073	0.755	0.079	0.746	0.103	0.505
	SNV	3	0.072	0.758	0.079	0.748	0.102	0.518

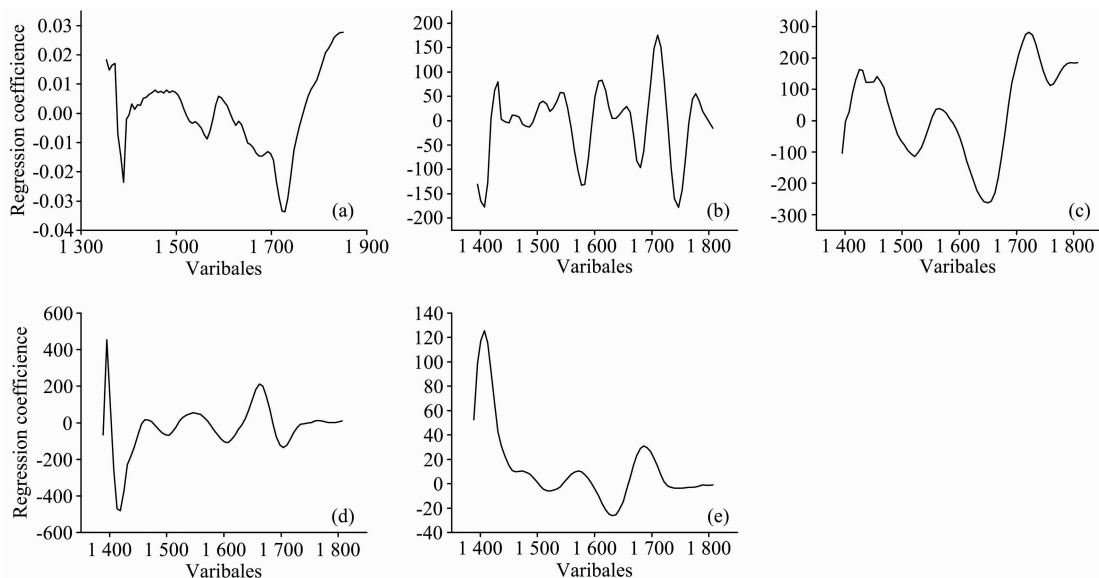


图 4 五类西红花掺伪品掺伪量的在全波段条件下的最优 PLSR 定量预测模型的变量与模型性能相关的回归系数

(a): 红花; (b): 菊花; (c): 莲须; (d): 纸浆; (e): 玉米须

Fig. 4 Regression coefficient between the variables and performance of the optimal PLSR quantitative prediction model of five kinds of adulterated saffron under the condition of full band

(a): Chrysanthemi flos; (b): nelumbinis; (c): Pulp; (d): Corn silk; (e): Under the condition of full band

References

- [1] Chinese Pharmacopoeia Commission (中国药典委员会). Pharmacopoeia of the People's Republic of China(中华人民共和国药典), Part One (第一部). Beijing: China Medical Science Press(北京: 中国医药科技出版社), 2015. 129.

- [2] LIU Jiang-di, OUYANG Zhen, YANG Bin(刘江弟, 欧阳臻, 杨 滨). *China Journal of Chinese Materia Medica(中国中药杂志)*, 2017, 42(3): 405.
- [3] Eman Shawky, Rasha M Abu El-Khair, Dina A Selim. *LWT-Food Science and Technology*, 2020, 122: 109032.
- [4] Eleftherios A Petrakis, Moschos G Polissiou. *Talanta*, 2016, 162: 558.
- [5] The International Organization for Standardization, *International Standard ISO 3632-2: Saffron (Crocus sativus L.) Test Methods*, Geneva, 2010.
- [6] Sabatino L, Scordino M, Gargano M, et al. *Natural Product Communications*, 2011, 6(12): 1873.
- [7] Masoum S, Gholami A, Hemmesi M, et al. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2015, 148: 389.
- [8] Soffritti G, Buscon M, Sánchez R A, et al. *Molecules*, 2016, 21: 343.
- [9] Daszykowski M, Walczak B, Massart D L. *Analytica Chimica Acta*, 2002, 468: 91.
- [10] Li Wenlong, Cheng Zhiwei, Wang Yuefei, et al. *Journal of Pharmaceutical and Biomedical Analysis*, 2013, 72: 33.
- [11] Shen Fei, Yang Danting, Ying Yibin, et al. *Food and Bioprocess Technology*, 2012, 5(2): 786.
- [12] MAO Hai-lei, XU Min, WANG Bin, et al(冒海蕾, 徐 旻, 王 斌, 等). *Acta Chimica Sinica(化学学报)*, 2007, 65(2): 152.
- [13] Yang Liming, Sun Qun. *Analytical Methods*, 2016, 8(8): 1914.

Fast Inspection of Saffron on the Spot Based on Cloud-Connected Portable Near-Infrared Technology

LI Qing^{1,2,3}, YAN Xiao-jian⁴, ZHAO Kui¹, LI Lan^{2,3}, PENG Shan-gui^{2,3}, LUO Xiao^{2,3}, WEN Yong-sheng^{2,3}, YAN Zhu-yun^{1*}

1. Chengdu University of Traditional Chinese Medicine, The Ministry of Education Key Laboratory of Standardization of Chinese Herbal Medicine State, Key Laboratory Breeding Base of Systematic Research Development and Utilization of Chinese Medicine Recourses, Chengdu 611137, China
2. Chengdu Institute for Food and Drug Control, Chengdu 610045, China
3. NMPA Key Laboratory for Quality Monitoring and Evaluation of Traditional Chinese Medicine (Chinese Materia Medica), Chengdu 610045, China
4. Panovasic Technology Co., Ltd., Chengdu 610041, China

Abstract The use of cloud-connected infrared spectroscopy technology combined with chemometrics to identify the rarest saffron and its commonly encountered adulterants (carthami flos, corn silk, nelumbinis stamen, chrysanthemi flos, pulp) and adulterated saffron, and quantitative determination of adulterant in saffron. Near-infrared spectra of saffron, adulterants, and adulterated saffron were collected by using the PV500R-I portable near-infrared instrument controlled by mobile phone. The first derivative, second derivative, third derivative, standard normal variable transformation and multiplicative scatter correction are used to preprocess the original spectral data. Partial Least Squares Discrimination Analysis was used to establish the identification model of saffron and its adulterant, and saffron and adulterated saffron. The results show that an optimal recognition model can distinguish saffron and its five kinds of adulterant from each other completely; the lowest 93% external prediction accuracy of saffron and five kinds of samples of adulterated saffron can be achieved step-by-step by two optimal recognition models, and the adulteration recognition level of saffron mixed with carthami flos, corn silk, nelumbinis stamen, chrysanthemi flos and pulp are 0.5%, 0.5%, 4.0%, 0.5% and 0.5%, respectively. Partial least squares regression was used to establish quantitative prediction models for the five kinds adulterant in saffron. The external prediction correlation coefficient range of the final model was 0.920~0.999, and RMSEP range was 0.005~0.044, and when the saffron mixed with carthami flos, chrysanthemi flos, nelumbinis stamen, pulp and corn silk are more than 8%, its external prediction relative error is less than 8%, 8%, 3%, 10% and 5% respectively, which indicated that the quantitative prediction model could be used to predict the amount of adulterant in saffron. To sum up, the identification method based on cloud connected portable near-infrared spectroscopy and the prediction method of the amount of adulterant is fast and accurate, economic and environmental protection, and can meet the requirements of quick and non-destructive identification of saffron on site.

Keywords Cloud-connected portable near-infrared technology; Chemometrics; Saffron; Adulterant; Adulterated samples

* Corresponding author

(Received May 11, 2020; accepted Jul. 15, 2020)