

基于近、中红外光谱技术的热毒宁注射液制备过程中金银花浓缩过程含量预测研究

候化蕊¹, 徐芳芳^{2,3*}, 张欣^{2,3}, 吴云^{2,3}, 张永超^{2,3}, 章晨峰^{2,3}, 王振中^{2,3*}

1. 南京中医药大学, 江苏 南京 210023

2. 江苏康缘药业股份有限公司, 江苏 连云港 222001

3. 中药制药过程新技术国家重点实验室, 江苏 连云港 222001

摘要: 目的 采用近红外光谱 (near infrared spectroscopy, NIRS)、中红外光谱 (mid-infrared spectroscopy, MIRS) 技术实现对热毒宁注射液制备过程中金银花浓缩过程绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量 8 个质控指标的含量预测, 并对比 2 种技术的预测效果。方法 收集热毒宁注射液制备过程中金银花浓缩过程样本, 进行 NIRS、MIRS 采集和含量测定, 优选最佳光谱预处理方法和特征波段, 采用偏最小二乘 (partial least square, PLS) 法建立 8 个质控指标的含量预测模型, 并比较 8 个质控指标的 NIRS、MIRS 模型性能, 得到 8 个最优含量预测模型, 并对其进行外部验证。结果 NIRS 技术对绿原酸、隐绿原酸、异绿原酸 C、断氧化马钱子苷、固含量的预测效果更好, 平均相对预测误差 (average relative prediction error, ARPE) 分别为 1.57%、1.88%、4.13%、3.79%、0.94%, 故选用 NIRS 模型作为这 5 个质控指标的最佳模型; MIRS 技术对新绿原酸、异绿原酸 A、异绿原酸 B 的预测效果更好, ARPE 分别为 4.00%、4.01%、2.32%, 故选用 MIRS 模型作为这 3 个质控指标的最佳模型。经外部验证, 绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量预测模型的 ARPE 分别为 6.22%、7.94%、5.44%、5.05%、5.37%、7.44%、6.00% 和 2.74%。结论 建立了 NIRS、MIRS 技术对热毒宁注射液制备过程中金银花浓缩过程的含量预测方法, 该方法快速简便、结果可靠, 且 MIRS 技术较 NIRS 技术有更低的检测限。

关键词: 热毒宁注射液; 金银花; 近红外光谱; 中红外光谱; 含量预测; 浓缩过程; 绿原酸; 新绿原酸; 隐绿原酸; 异绿原酸 A; 异绿原酸 B; 异绿原酸 C; 断氧化马钱子苷

中图分类号: R283.6 文献标志码: A 文章编号: 0253-2670(2023)02-0520-14

DOI: 10.7501/j.issn.0253-2670.2023.02.019

Prediction of *Lonicerae Japonicae Flos* concentration in preparation of Reduning Injection based on NIRS and MIRS technology

HOU Hua-rui¹, XU Fang-fang^{2,3}, ZHANG Xin^{2,3}, WU Yun^{2,3}, ZHANG Yong-chao^{2,3}, ZHANG Chen-feng^{2,3}, WANG Zhen-zhong^{2,3}

1. Nanjing University of Chinese Medicine, Nanjing 210023, China

2. Jiangsu Kanion Pharmaceutical Co., Ltd., Lianyungang 222001, China

3. State Key Laboratory of New-tech for Chinese Medicine Pharmaceutical Process, Lianyungang 222001, China

Abstract: Objective Using near infrared spectroscopy (NIRS) and mid-infrared spectroscopy (MIRS) technology to predict the content of eight quality control indicators during the concentration process of Jinyinhua (*Lonicerae Japonicae Flos*, LJF) in preparation of Reduning Injection (热毒宁注射液), namely chlorogenic acid, neochlorogenic acid, cryptochlorogenic acid, isochlorogenic acid A, isochlorogenic acid B, isochlorogenic acid C, secoxyloganin, solid content, and to compare the prediction results of the two techniques. **d. Methods** The samples collected during the concentrated process of Reduning Injection LJF were collected for NIRS and MIRS acquisition and content determination. The optimal spectral preprocessing method and characteristic

收稿日期: 2022-07-26

基金项目: 2022 年中央财政转移支付地方项目: 基于重点研究室研究领域的中医药多学科研究能力提升项目—中药提取精制新技术

作者简介: 候化蕊, 硕士研究生, 研究方向为中药制药过程新技术。E-mail: 2770198973@qq.com

*通信作者: 王振中, 研究员, 硕士生导师, 研究方向为中药新药研发及应用研究。E-mail: kyywzz@163.com

徐芳芳, 女, 博士, 研究方向为中药新药研发及过程控制研究。E-mail: 879164331@qq.com

band were selected, and the partial least squares method was used to establish the content prediction model of eight quality control indicators. And compare NIRS and MIRS model performance of eight quality control indicators. Eight optimal content prediction models were obtained and externally verified. **Results** The NIRS technique had better prediction effect on chlorogenic acid, cryptochlorogenic acid, isochlorogenic acid C, secoxyloganin and solid content, the average relative prediction error (ARPE) was 1.57%, 1.88%, 4.13%, 3.79% and 0.94%, respectively. The prediction effect of MIRS technology was better for neochlorogenic acid, isochlorogenic acid A and isochlorogenic acid B, the ARPE was 4.00%, 4.01% and 2.32%, respectively. Therefore, chlorogenic acid, cryptochlorogenic acid, isochlorogenic acid C, secoxyloganin and solid were selected by NIRS model; neochlorogenic acid, isochlorogenic acid A and isochlorogenic acid B were selected the MIRS model as the best model. Through external verification, the ARPEs of chlorogenic acid, neochlorogenic acid, cryptochlorogenic acid, isochlorogenic acid A, isochlorogenic acid B, isochlorogenic acid C, secoxyloganin and solid content prediction models were 6.22%, 7.94%, 5.44%, 5.05%, 5.37%, 7.44%, 6.00% and 2.74%, respectively. **Conclusion** The established method of NIRS and MIRS technology for the content prediction of LJF in Reduning Injection during the concentration process is fast, simple and reliable, and MIRS technology has lower detection limit than NIRS technology.

Key words: Reduning Injection; *Lonicerae Japonicae Flos*; near infrared spectroscopy; mid-infrared spectroscopy; content prediction; concentration process; chlorogenic acid; neochlorogenic acid; cryptochlorogenic acid; isochlorogenic acid A; isochlorogenic acid B; isochlorogenic acid C; secoxyloganin

热毒宁注射液 (Reduning Injection, RI) 是由青蒿、金银花、栀子 3 味中药组成的复方中药注射液, 具有清热、疏风、解毒的功效, 常用于治疗上呼吸道感染所致的咳嗽、高热、头身疼痛、痰色发黄等症状^[1-3]。金银花作为热毒宁注射液的臣药, 具有解热、抗炎、抗病毒的作用, 其含有大量的有机酸类、环烯醚萜类成分, 如绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷等^[4-5]。在热毒宁注射液的生产过程中, 金银花提取液的浓缩过程是不断挥发溶剂的过程, 除有效成分含量之外, 固含量也常常是中药质量控制指标^[6]。

中药成分复杂, 且影响因素众多, 过程分析技术应用中药制药过程质量控制如中药提取、浓缩、醇沉、干燥、混合、制粒、包衣等工序均取得了良好的质量控制效果^[7]。近红外光谱 (near infrared spectroscopy, NIRS) 技术和中红外光谱 (mid-infrared spectroscopy, MIRS) 技术属于过程分析技术中的光谱分析技术。目前, 通过 NIRS 技术建立定量预测模型对质控指标进行快速检测, 已在热毒宁注射液的质量控制有了大量应用^[8-11], 而 MIRS 技术的应用研究还较少。由于漫反射傅里叶变换红外光谱 (diffuse reflection Fourier transform infrared spectroscopy, DR-FTIR) 和衰减全反射傅里叶变换红外光谱 (attenuated total internal reflectance FTIR, ATR-FTIR) 的出现, MIRS 技术在中药质量控制也开始实现了应用, 如中药的定性鉴别^[12]、定量分析^[13]和在线监测^[14]。研究表明, ATR-FTIR 是通过样品

表面的反射信号获得样品表层有机成分的结构信息, 对样品的大小、形状、含水量没有特殊要求, 而且灵敏度更高, 特征性更强, 可以实现原位测试、实时跟踪^[15-16], Tao 等^[17]研究表明 MIRS 技术对低浓度的分析物有较好的预测效果。因此, 除 NIRS 技术外, 本实验同时使用 ATR-FTIR 技术建立绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量的含量预测模型, 并对比 NIRS 模型和 MIRS 模型对各个质控指标的预测效果。

1 仪器与材料

1.1 仪器

ReactIR 702L 型原位傅里叶变换红外光谱仪, 梅特勒-托利多仪器上海有限公司; Antaris II 型傅里叶近红外变换光谱仪、Ultimate3000 型高效液相色谱仪, 赛默飞世尔科技 (中国) 有限公司; Mettler Toledo 型电子天平, 梅特勒-托利多仪器上海有限公司; H1650-W 型台式高速离心机, 湖南湘仪实验室仪器开发有限公司; KQ-500DE 型数控超声波清洗器, 昆山市超声仪器有限公司。

1.2 材料

对照品绿原酸, 批号 110753-202018, 质量分数 96.1%, 中国食品药品检定研究院; 对照品新绿原酸 (批号 DSTDX001503, 质量分数 99.58%)、断氧化马钱子苷 (批号 DST211224-111, 质量分数 98.87%), 成都德思特生物技术有限公司; 对照品隐绿原酸 (批号 3208, 质量分数 98.04%)、异绿原酸 B (批号 3089, 质量分数 98.5%), 上海诗丹德标

准技术服务有限公司; 对照品异绿原酸 A (批号 MUST-15042518, 质量分数 98.82%)、异绿原酸 C (批号 MUST-15081414, 质量分数 99.84%), 成都曼思特生物科技有限公司; 甲醇, 色谱纯, 美国 Tedia 公司; 甲醇, 分析纯, 南京化学试剂有限公司; 磷酸, 分析纯, 国药集团化学试剂有限公司; 水为超纯水; 金银花浓缩过程样品来自江苏康缘药业股份有限公司。

2 方法与结果

2.1 样品收集

热毒宁注射液制备中金银花浓缩工序包括过程浓缩和终点浓缩 2 个部分, 整个浓缩过程在 6 h 以内。过程浓缩开始 10 min 后采集第 1 个样本, 每隔 10~15 min 取样 1 次; 当缓存罐液位达到超低限 80 mm 时, 将一效、二效内药液并入三效蒸发室内再进行终点浓缩, 终点浓缩开始后每隔 3~5 min 取样 1 次, 直至浓缩结束。采集 4 批样品 (批号 Z211204、Z220213、Z220302、Z220503) 数据, 4 个批次样品数量分别为 25、33、28、24 个, 共计 110 个金银花浓缩过程样本。

2.2 金银花浓缩过程质控指标的测定

2.2.1 有效成分的含量测定

(1) 色谱条件: 参考企业内控质量标准, 色谱柱为 Kromasil C₁₈ 柱 (250 mm×4.6 mm, 5 μm); 流动相为乙腈-0.1%磷酸水溶液, 梯度洗脱: 0~6 min, 15%~18%乙腈; 6~15 min, 18%乙腈; 15~38 min, 18%~31%乙腈; 38~40 min, 31%~90%乙腈; 40~41 min, 90%~15%乙腈; 41~45 min, 15%乙腈; 体积流量 1.0 mL/min; 柱温 30 °C; 检测波长为 327、238 nm (断氧化马钱子苷); 进样体积 10 μL。

(2) 混合对照品溶液的配制: 精密称取 7 种对照品适量, 分别加入 50%甲醇溶解并稀释制成含绿原酸 507.5 μg/mL、新绿原酸 307.1 μg/mL、隐绿原酸 705.6 μg/mL、异绿原酸 A 807.9 μg/mL、异绿原酸 B 509.2 μg/mL、异绿原酸 C 799.6 μg/mL、断氧化马钱子苷 740.3 μg/mL 的各对照品储备液。依次精密移取绿原酸对照品储备液 5 mL、新绿原酸对照品储备液 0.5 mL、隐绿原酸对照品储备液 0.5 mL、异绿原酸 A 对照品储备液 1 mL、异绿原酸 B 对照品储备液 0.5 mL、异绿原酸 C 对照品储备液 1 mL、断氧化马钱子苷对照品储备液 0.5 mL, 置 25 mL 量瓶中, 加 50%甲醇稀释, 定容, 摇匀, 即得混合对

照品溶液。

(3) 供试品溶液配制: 精密称取 0.3 g 浓缩液置 25 mL 量瓶中, 用 50%甲醇稀释, 50 Hz 超声 30 min, 定容至刻度, 摇匀, 12 000 r/min 条件下离心 5 min, 取上清液, 0.45 μm 有机微孔滤膜滤过, 即得供试品溶液。

(4) 线性关系考察: 分别精密称取新绿原酸 4.982 mg、绿原酸 17.984 mg、隐绿原酸 7.564 mg、异绿原酸 B 4.974 mg、异绿原酸 A 7.665 mg、异绿原酸 C 8.143 mg、断氧化马钱子苷 7.996 mg 置 10 mL 量瓶中, 加 50%甲醇溶解并稀释至刻度, 摇匀, 制成各对照品储备液。依次精密移取绿原酸 5 mL, 新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷各 2 mL 置 25 mL 量瓶中, 加 50%甲醇稀释, 定容, 摇匀, 即得混合对照品溶液。分别精密吸取混合对照品溶液 0.25、0.5、1.0、2.0、4.0、10.0 mL 置 10 mL 量瓶中, 加 50%甲醇稀释至刻度, 摇匀, 即得系列对照品溶液。分别吸取上述系列对照品溶液 10 μL 注入高效液相色谱仪, 进行测定, 以对照品质量浓度为横坐标 (X), 峰面积为纵坐标 (Y), 进行线性回归, 得回归方程: 新绿原酸 $Y=1019 X-3.262$, $r=0.999 8$, 线性范围 0.99~39.69 μg/mL; 绿原酸 $Y=721.2 X-21.50$, $r=0.999 9$, 线性范围 8.64~345.65 μg/mL; 隐绿原酸 $Y=692.5 X-3.126$, $r=0.999 7$, 线性范围 1.48~59.33 μg/mL; 异绿原酸 B $Y=380.7 X-3.775$, $r=0.999 9$, 线性范围 0.98~39.20 μg/mL; 异绿原酸 A $Y=173.4 X+13.68$, $r=0.999 9$, 线性范围 1.51~60.60 μg/mL; 异绿原酸 C $Y=3145 X-82.04$, $r=0.999 9$, 线性范围 1.63~65.04 μg/mL; 断氧化马钱子苷 $Y=508.0 X-3.774$, $r=0.999 7$, 线性范围 1.58~63.25 μg/mL。

(5) 精密度考察: 精密吸取混合对照品溶液, 按“(1)”色谱条件连续进样 6 次, 记录峰面积, 结果新绿原酸、绿原酸、隐绿原酸、异绿原酸 B、异绿原酸 A、异绿原酸 C 和断氧化马钱子苷峰面积的 RSD 分别为 0.8%、0.2%、0.2%、0.7%、0.3%、0.2%、0.6%, 表明在此条件下进样精密度良好。

(6) 重复性考察: 按“(3)”项下方法平行制备 6 份供试品溶液 (批号 Z211204-25), 按“(1)”项色谱条件测定, 计算样品中各成分的含量, 结果新绿原酸、绿原酸、隐绿原酸、异绿原酸 B、异绿原酸 A、异绿原酸 C 和断氧化马钱子苷含量的 RSD

分别为 1.0%、0.9%、0.9%、1.6%、1.7%、1.4%、0.8%，说明该方法重复性良好。

(7) 稳定性考察：按“(3)”项下方法制备供试品溶液（批号 Z211204-25），按“(1)”项色谱条件分别在不同时间点（0、2、4、8、16、24 h）进样，记录峰面积，结果新绿原酸、绿原酸、隐绿原酸、异绿原酸 B、异绿原酸 A、异绿原酸 C 和断氧化马钱子苷峰面积的 RSD 分别为 0.6%、0.5%、0.9%、0.7%、0.5%、0.8%、0.9%，说明供试品溶液在 24 h 内稳定性良好。

(8) 加样回收率考察：精密称取已知含量的样品（批号 Z211204-25）0.1 g 置 25 mL 量瓶中，精密加入混合对照品溶液 2 mL，加 50% 甲醇稀释至刻度，摇匀。平行制备 6 份，按“(1)”项下色谱条件进样，计算加样回收率，得到新绿原酸、绿原酸、隐绿原酸、异绿原酸 B、异绿原酸 A、异绿原酸 C 和断氧化马钱子苷的平均加样回收率依次为 98.1%、99.4%、101.2%、99.6%、98.6%、98.1%、97.5%，RSD 分别为 1.2%、0.9%、1.0%、1.2%、1.1%、1.3%、1.5%，符合定量分析要求。

2.2.2 固含量测定 依据《中国药典》2020 年版“固体总量测定法”，量取 5 g 金银花浓缩液至已烘干至恒定质量 (X_0) 的称量瓶中，称定质量 (X_1)，置烘箱中 105 °C 条件下烘干至恒定质量，称定质量 (X_2)。固含量按公式计算。

$$\text{固含量} = (X_2 - X_0) / (X_1 - X_0)$$

2.2.3 质控指标测定结果 110 个样品中 7 个有效成分的含量及固含量共 8 个指标的测定结果如表 1 所示。其中，绿原酸为 2.57~14.80 mg/g、新绿原酸为 0.28~2.54 mg/g、隐绿原酸为 0.45~3.11 mg/g、异绿原酸 A 为 0.33~2.97 mg/g、异绿原酸 B 为 0.19~

表 1 110 个样品质控指标测定结果

Table 1 Quality control indicator determination results of 110 samples

质控指标	单位	最小值	最大值	上四分位数	下四分位数	均值
绿原酸	mg·g ⁻¹	2.57	14.80	7.54	10.06	8.75
新绿原酸	mg·g ⁻¹	0.28	2.54	0.81	1.60	1.19
隐绿原酸	mg·g ⁻¹	0.45	3.11	1.26	2.13	1.70
异绿原酸 A	mg·g ⁻¹	0.33	2.97	1.43	2.03	1.73
异绿原酸 B	mg·g ⁻¹	0.19	1.79	0.74	1.34	1.06
异绿原酸 C	mg·g ⁻¹	0.44	3.73	1.89	2.76	2.33
断氧化马钱子苷	mg·g ⁻¹	0.48	2.86	1.04	1.55	1.36
固含量	%	5.11	32.98	17.47	25.26	21.12

1.79 mg/g、异绿原酸 C 为 0.44~3.73 mg/g、断氧化马钱子苷为 0.48~2.86 mg/g，固含量为 5.11%~32.98%。

2.3 光谱采集

2.3.1 NIRS 采集 在室温条件下，空气作为背景，使用 Antaris II 型傅里叶近红外变换光谱仪液体透射法（1 mm 比色皿），采集金银花浓缩过程中间体的 NIRS。光谱采集条件：光谱扫描范围 4000~10 000 cm⁻¹，扫描次数为 64 次，分辨率为 8 cm⁻¹。采集的金银花浓缩过程 NIRS 图见图 1。

2.3.2 MIRS 采集 在室温条件下，空气作为背景，使用 ReactIR 702L 型原位傅里叶变换红外光谱仪的 ATR 光纤探头采集金银花浓缩过程中间体的 MIRS。光谱条件为光谱扫描范围 648~3000 cm⁻¹，扫描次数为 64 次，分辨率为 4 cm⁻¹，扫描速度为 7 次/s。采集的金银花浓缩过程 MIRS 图见图 2。

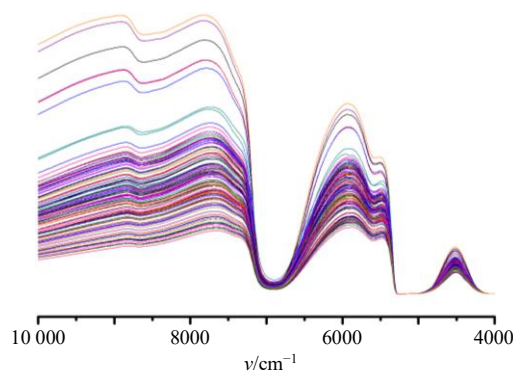


图 1 110 个样品 NIRS 原始光谱

Fig. 1 Original NIRS of 110 samples

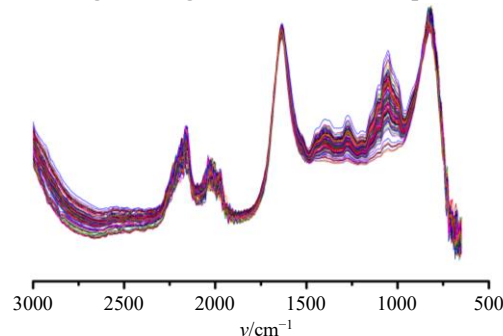


图 2 110 个样品 MIRS 原始光谱

Fig. 2 Original MIRS of 110 samples

2.4 样本的划分及预处理

4 批样本中，Z211204、Z220213、Z220302 3 个批次共 86 个样本用作定量模型的建立，为了避免样本分布不均，采用 Kennard-Stone^[18-19]算法划分 86 个金银花浓缩过程中间体，得到 17 个预测集样本，69 个校正集样本，Z220503 批次 24 个样本作

为外部验证集。

预处理采集的光谱可以减少杂散光、噪声、基线漂移等不必要因素对它的干扰,提升模型的精确度^[20]。本实验采用的光谱预处理方法有:Savitzky-Golay (SG) 卷积平滑与一阶导数结合、移动窗口平滑、基线校正、归一化法、多元散射校正法(multiplicative scatter correction, MSC)和标准正态变量校正法(standard normal variate, SNV)。

其中,移动窗口平滑法的基本思想是选择 1 个宽度为 $(2\omega+1)$ 的平滑窗口,计算窗口内的中心波长点 i 以及 i 点前后 ω 点处光谱测量值得平均值 x_i , x_i 用来代替波长点 i 处的测量值,依次改变 i 值来移动窗口,完成对所有波长点的平滑。

SG 卷积平滑法(多项式平滑),是一种在保留化学信息的同时有效去除光谱噪声尖峰的非常有用的方法,波长点 i 处经过平滑后的平均值按下列公式计算。

$$x_{i,\text{smooth}} = \bar{x} = \frac{1}{H} \sum_{k=-\omega}^{+\omega} x_{k+1} h_k, \quad H = \sum_{k=-\omega}^{+\omega} h_k$$

h_k 为平滑系数, H 为归一化因子

一阶导数能够消除基线偏移对光谱带来的干扰,对于 1 个离散光谱 x_i , 计算其波长点 i 处的一阶导数光谱, g 为窗口宽度,公式为 $x_{i,1st} = (x_i - x_{i+g})/g$ 。归一化法可以消除光程变化对光谱产生的影响,设原始样本的吸光度为 A_{ij} , 平均样本吸光度为 $\bar{A}$ 。用公式表示为 $\varphi = (x_{ij} - \bar{A}) / \sum x_{ij}^2$ 。

基线校正通过将数据调整到数据中的最小点来调整光谱偏移。 $f(x) = x - \min X$, 其中 x 是 1 个变量, X 表示该样本的所有选定变量。

MSC 主要是用来减小颗粒大小及分布不均匀产生的散射对光谱的影响,设样本光谱为 x 则光散射校正法首先求校正集光谱的平均光谱 $\bar{x}$ 。对 x 、 $\bar{x}$ 做线性回归,得到 $x = b_0 + b \bar{x}$, 利用最小二乘法计算出 b_0 、 b 。对原始样本光谱进行 MSC 变换,得到原始样本光谱的光散射方程为 $x_{\text{MSC}} = x - b_0/b$ 。

SNV 通过对每个单独的光谱进行居中和缩放来消除散射效应, x_i 表示原始样本光谱在 i 个波长点处的取得的吸光度。其公式如下 $x_{i,\text{SNV}} = (x_i - \bar{x}) / \text{SDev}(x)$ 。

2.5 特征波段筛选方法

使用组合间隔偏最小二乘法(synergy interval PLS, siPLS)、间隔偏最小二乘法(interval PLS, iPLS)、移动窗口偏最小二乘法(moving window

PLS, mwPLS) 筛选波段^[21]。其中, siPLS 将全光谱划分为 20 个子区间,组合数为 3。mwPLS 依据 NIRS、MIRS 中波数点的不同,将窗口宽度设置为 31、29。

2.6 预测模型的建立与验证

使用偏最小二乘(partial least square, PLS)法建立绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量的定量模型,依据模型评价指标选择最优的光谱预处理方法和波段筛选方法。再通过计算预测值与参考值的平均相对预测误差(average relative prediction error, ARPE),对比 NIRS 模型、MIRS 模型的预测效果,确定最优的含量预测模型。最后,取新的 1 批样本(Z220503)导入建好的模型中,对模型进行外部验证。

模型评价指标:校正集相关系数(R_{cal})、预测集相关系数(R_{pre})、校正集均方根误差(root mean square errors of calibration, RMSEC)、交叉验证均方根误差(root mean square errors of cross validation, RMSECV)、预测均方根误差(root mean square errors of prediction, RMSEP)、性能偏差比(ratio of performance to deviation, RPD)、预测相对偏差(relative standard error of prediction, RSEP)。其中 R_{cal} 与 R_{pre} 越大模型拟合效果越好, RMSEC、RMSECV、RMSEP、RSEP 越小, RPD 越大,模型的预测性能越好^[21]。相关计算公式如下所示。

$$R = [1 - \sum (C_i - \hat{C}_i)^2 / \sum (C_i - C_m)^2]^{1/2}$$

$$\text{RMSE} = [\sum_{i=1}^n (\hat{C}_i - C_i)^2 / n]^{1/2}$$

$$\text{RPD} = [\sum_{i=1}^n (\hat{C}_i - C_i)^2]^{1/2} / \text{RMSE}$$

$$\text{RSEP} = [\sum (\hat{C}_i - C_i)^2 / \sum C_i^2]^{1/2}$$

$$\text{ARPE} = \sum |\hat{C}_i - C_i| / n C_i$$

C_i 为标准方法测得的参考值, $\hat{C}_i$ 为定量模型得到的预测值, C_m 为 C_i 的平均值, n 为不同样品集的样品个数

2.7 数据分析软件

采用 Unscrambler X 10.4 (Camo software AS, Norway) 软件对光谱进行预处理,采用 Matlab 2015b (Mathwork Inc., USA) 软件进行波段筛选与建模,采用 GraphPad Prism 8.0 (GraphPad Software 公司)、Oringin 8.0 (OringinLab 公司) 绘图。

2.8 光谱预处理结果

2.8.1 NIRS 预处理结果 不同预处理方法对 NIRS

建模效果的影响如表 2 所示, 综合考虑各项评价指标, 如绿原酸和断氧化马钱子苷经过 S-G 1st 处理后 R_{cal} 、 R_{pre} 、RPD 均上升, RMSEC、RMSECV、RMSEP、RSEP 均下降, 建模效果最好, 其 RPD 分别为 4.65、3.27, RSEP 分别为 4.18%、8.05%。新绿原酸的未

预处理的原始光谱建模效果最好, 其 RPD 为 10.00, RSEP 为 3.42%。隐绿原酸经基线校正处理后的建模效果更好, RPD 为 14.67, RSEP 为 1.91%。异绿原酸 A、异绿原酸 B、异绿原酸 C 采用 SNV 预处理建模效果最好, 其 RPD 分别为 3.88、7.10、4.87,

表 2 不同预处理方法对 NIRS 模型的影响

Table 2 Effects of different preprocessing methods on NIRS models

指标	预处理方法	R_{cal}	RMSEC/(mg·g ⁻¹)	RMSECV/(mg·g ⁻¹)	R_{pre}	RMSEP/(mg·g ⁻¹)	RPD	RSEP/%
绿原酸	无预处理	0.980 0	0.479 9	0.541 2	0.946 6	0.545 0	3.09	6.10
	移动窗口平滑	0.980 0	0.479 5	0.540 3	0.947 2	0.542 3	3.10	6.07
	S-G 1 st	0.993 1	0.280 5	0.359 3	0.981 6	0.360 3	4.65	4.18
	归一化法	0.977 0	0.514 5	0.618 7	0.932 7	0.524 6	2.71	6.37
	基线校正	0.980 4	0.474 3	0.537 0	0.940 6	0.601 6	2.83	6.73
	SNV	0.985 7	0.402 1	0.574 6	0.966 5	0.474 0	3.29	5.76
	MSC	0.986 2	0.395 0	0.558 8	0.966 9	0.473 1	3.29	5.75
新绿原酸	无预处理	0.995 3	0.049 4	0.112 0	0.995 2	0.043 2	10.00	3.42
	移动窗口平滑	0.992 5	0.049 8	0.112 5	0.995 2	0.043 3	9.98	3.42
	S-G 1 st	0.988 9	0.076 7	0.118 2	0.986 3	0.073 2	5.16	6.72
	归一化法	0.980 8	0.100 1	0.153 2	0.964 6	0.103 9	3.51	10.08
	基线校正	0.985 3	0.087 7	0.130 1	0.962 7	0.117 2	3.56	9.57
	SNV	0.981 5	0.097 0	0.156 0	0.977 4	0.084 1	4.84	8.04
	MSC	0.980 9	0.098 6	0.159 2	0.976 5	0.085 9	4.74	8.20
隐绿原酸	无预处理	0.998 5	0.034 3	0.061 8	0.998 2	0.035 8	13.80	2.05
	移动窗口平滑	0.998 3	0.036 5	0.064 0	0.998 5	0.034 7	14.24	1.99
	S-G 1 st	0.994 1	0.067 7	0.081 3	0.992 3	0.055 8	8.07	3.56
	归一化法	0.996 9	0.048 6	0.066 3	0.990 2	0.065 9	6.34	4.44
	基线校正	0.998 8	0.030 2	0.054 7	0.997 6	0.032 5	14.67	1.91
	SNV	0.993 9	0.067 5	0.091 0	0.991 5	0.059 6	7.79	3.99
	MSC	0.994 5	0.064 3	0.090 0	0.992 2	0.057 3	8.10	3.84
异绿原酸 A	无预处理	0.983 6	0.100 1	0.126 8	0.976 6	0.113 8	3.93	6.71
	移动窗口平滑	0.983 8	0.099 5	0.127 6	0.978 4	0.111 8	4.00	6.59
	S-G 1 st	0.992 4	0.067 3	0.131 5	0.974 3	0.103 6	4.57	6.40
	归一化法	0.987 4	0.086 6	0.119 0	0.977 4	0.093 7	4.71	6.00
	基线校正	0.984 5	0.098 1	0.126 4	0.976 9	0.102 6	4.13	6.04
	SNV	0.987 1	0.088 5	0.121 8	0.966 7	0.083 2	3.88	5.79
	MSC	0.987 2	0.088 2	0.121 8	0.966 0	0.084 0	3.85	5.84
异绿原酸 B	无预处理	0.994 0	0.044 4	0.058 9	0.989 1	0.049 3	6.98	4.41
	移动窗口平滑	0.994 0	0.044 5	0.058 9	0.989 1	0.049 3	6.98	4.41
	S-G 1 st	0.985 1	0.069 5	0.085 9	0.988 6	0.055 5	6.00	5.55
	归一化法	0.994 9	0.040 8	0.056 7	0.985 8	0.052 0	5.87	5.53
	基线校正	0.993 2	0.047 7	0.062 0	0.989 6	0.047 1	6.94	4.28
	SNV	0.993 1	0.047 3	0.061 8	0.990 8	0.041 4	7.10	4.55
	MSC	0.993 1	0.047 1	0.061 7	0.990 7	0.041 7	7.06	4.58

续表 2

指标	预处理方法	R_{cal}	RMSEC/(mg·g ⁻¹)	RMSECV/(mg·g ⁻¹)	R_{pre}	RMSEP/(mg·g ⁻¹)	RPD	RSEP/%
异绿原酸 C	无预处理	0.993 3	0.083 7	0.131 4	0.987 8	0.106 9	5.43	4.65
	移动窗口平滑	0.992 7	0.086 8	0.132 1	0.988 1	0.111 7	5.33	4.73
	S-G 1 st	0.986 9	0.114 9	0.186 0	0.990 0	0.109 9	5.52	4.98
	归一化法	0.995 4	0.069 5	0.096 3	0.978 5	0.129 5	3.88	6.12
	基线校正	0.992 5	0.089 1	0.137 6	0.986 1	0.109 2	5.21	4.64
	SNV	0.988 5	0.109 1	0.134 5	0.985 2	0.090 3	4.87	4.48
	MSC	0.994 6	0.074 8	0.101 6	0.978 4	0.119 9	3.66	5.95
断氧化马钱子苷	无预处理	0.947 5	0.181 1	0.201 5	0.940 2	0.154 7	2.33	11.18
	移动窗口平滑	0.947 4	0.181 2	0.201 6	0.940 1	0.154 8	2.33	11.19
	S-G 1 st	0.968 5	0.140 5	0.168 2	0.982 6	0.107 7	3.27	8.05
	归一化法	0.947 8	0.179 9	0.196 0	0.882 0	0.156 2	2.04	12.25
	基线校正	0.945 9	0.184 8	0.205 5	0.930 7	0.134 3	2.36	10.08
	SNV	0.948 1	0.178 6	0.195 1	0.864 8	0.159 3	2.04	12.61
	MSC	0.947 2	0.180 1	0.196 8	0.863 8	0.160 0	2.03	12.67
固含量	无预处理	0.998 9	0.292 5	0.494 8	0.998 8	0.225 6	21.07	1.06
	移动窗口平滑	0.998 9	0.292 6	0.494 2	0.998 8	0.224 8	21.14	1.06
	S-G 1 st	0.995 4	0.580 2	0.733 5	0.998 5	0.428 5	10.81	2.18
	归一化法	0.999 2	0.237 7	0.316 1	0.999 0	0.190 5	22.06	1.01
	基线校正	0.998 8	0.297 3	0.467 3	0.998 3	0.269 7	17.16	1.28
	SNV	0.998 5	0.329 7	0.417 8	0.999 3	0.190 9	22.99	1.03
	MSC	0.999 2	0.241 9	0.323 2	0.998 8	0.219 7	19.97	1.19

RSEP 分别为 5.79%、4.55%、4.48%。固含量采用归一化法预处理后的建模效果更好, 其 RPD 为 22.06, RSEP 为 1.01%。

2.8.2 MIRS 预处理结果 不同预处理方法对 MIRS 建模效果的影响如表 3 所示。综合考虑各项评价指标, 如绿原酸的预处理方法中归一化法的

表 3 不同预处理方法对 MIRS 模型的影响

Table 3 Effects of different preprocessing methods on MIRS models

指标	预处理方法	R_{cal}	RMSEC/(mg·g ⁻¹)	RMSECV/(mg·g ⁻¹)	R_{pre}	RMSEP/(mg·g ⁻¹)	RPD	RSEP/%
绿原酸	无预处理	0.980 5	0.476 6	0.663 1	0.912 8	0.627 9	2.50	7.04
	移动窗口平滑	0.978 0	0.506 5	0.687 2	0.912 3	0.631 1	2.49	7.06
	S-G 1 st	0.963 2	0.647 7	0.801 2	0.924 9	0.640 3	2.62	7.23
	归一化法	0.986 5	0.402 1	0.611 5	0.900 9	0.647 4	2.17	7.21
	基线校正	0.972 3	0.574 6	0.708 5	0.870 4	0.581 6	2.09	6.88
	SNV	0.982 6	0.452 9	0.648 6	0.916 9	0.581 3	2.57	6.59
	MSC	0.982 3	0.455 9	0.653 5	0.914 5	0.587 5	2.55	6.66
新绿原酸	无预处理	0.988 1	0.080 2	0.182 7	0.943 9	0.117 6	3.06	10.66
	移动窗口平滑	0.985 9	0.087 1	0.186 1	0.944 5	0.118 6	3.12	10.63
	S-G 1 st	0.979 6	0.103 6	0.232 3	0.960 0	0.112 2	3.61	9.61
	归一化法	0.987 2	0.083 5	0.189 0	0.921 3	0.141 0	2.51	12.01
	基线校正	0.999 4	0.017 1	0.147 5	0.867 1	0.178 1	2.05	17.07
	SNV	0.989 5	0.074 7	0.184 0	0.914 8	0.164 6	2.41	14.20
	MSC	0.989 3	0.075 1	0.185 2	0.914 4	0.163 7	2.46	14.05

续表 3

指标	预处理方法	R_{cal}	RMSEC/(mg·g ⁻¹)	RMSECV/(mg·g ⁻¹)	R_{pre}	RMSEP/(mg·g ⁻¹)	RPD	RSEP/%
隐绿原酸	无预处理	0.995 9	0.057 1	0.117 4	0.966 1	0.122 6	3.55	7.64
	移动窗口平滑	0.994 6	0.064 9	0.123 8	0.964 0	0.123 9	3.59	7.66
	S-G 1 st	0.996 1	0.054 7	0.161 2	0.974 1	0.108 4	4.40	6.53
	归一化法	0.995 4	0.060 9	0.124 3	0.969 9	0.121 0	3.40	7.24
	基线校正	0.996 5	0.051 8	0.114 9	0.939 0	0.142 7	2.88	9.47
	SNV	0.996 8	0.050 2	0.118 1	0.954 9	0.145 2	3.18	8.82
	MSC	0.995 9	0.056 7	0.118 0	0.959 8	0.135 3	3.45	8.19
异绿原酸 A	无预处理	0.997 7	0.037 8	0.146 5	0.892 4	0.187 5	2.12	10.01
	移动窗口平滑	0.965 9	0.141 6	0.188 2	0.957 1	0.150 1	3.01	7.84
	S-G 1 st	0.958 4	0.153 7	0.202 6	0.894 9	0.270 3	1.85	14.11
	归一化法	0.997 3	0.039 5	0.146 9	0.883 2	0.196 7	2.13	9.85
	基线校正	0.997 5	0.038 6	0.158 5	0.934 4	0.174 8	2.69	9.80
	SNV	0.996 8	0.043 5	0.158 4	0.918 2	0.183 3	2.50	9.63
	MSC	0.996 9	0.043 2	0.158 6	0.914 8	0.187 4	2.45	9.84
异绿原酸 B	无预处理	0.989 1	0.061 0	0.084 9	0.942 4	0.111 5	2.70	10.31
	移动窗口平滑	0.992 3	0.050 7	0.082 5	0.952 0	0.106 5	3.12	9.60
	S-G 1 st	0.988 5	0.061 8	0.098 8	0.972 1	0.091 8	3.68	8.07
	归一化法	0.988 5	0.062 2	0.095 7	0.967 5	0.098 5	3.18	8.41
	基线校正	0.991 8	0.051 8	0.083 3	0.966 7	0.099 8	3.44	9.68
	SNV	0.921 0	0.051 2	0.087 6	0.950 0	0.102 3	3.30	9.07
	MSC	0.992 2	0.050 8	0.086 4	0.961 0	0.101 9	3.35	9.00
异绿原酸 C	无预处理	0.997 0	0.056 8	0.152 1	0.943 1	0.161 3	3.05	6.55
	移动窗口平滑	0.984 1	0.129 0	0.173 2	0.969 5	0.143 3	3.81	5.73
	S-G 1 st	0.989 7	0.102 6	0.179 3	0.966 7	0.205 4	2.90	8.17
	归一化法	0.986 7	0.117 8	0.163 3	0.955 8	0.155 3	3.25	5.94
	基线校正	0.988 5	0.109 5	0.168 0	0.971 1	0.141 2	4.11	6.03
	SNV	0.983 0	0.133 6	0.165 6	0.957 6	0.165 4	3.29	6.63
	MSC	0.982 7	0.134 5	0.166 7	0.960 2	0.166 2	3.28	6.64
断氧化马钱子苷	无预处理	0.961 6	0.153 3	0.211 9	0.848 1	0.223 8	1.76	16.52
	移动窗口平滑	0.962 6	0.151 6	0.223 1	0.890 2	0.184 3	2.12	13.49
	S-G 1 st	0.921 2	0.214 5	0.281 2	0.945 2	0.145 9	3.02	10.55
	归一化法	0.973 3	0.131 7	0.198 2	0.719 9	0.240 2	1.13	18.36
	基线校正	0.963 0	0.153 1	0.229 7	0.802 9	0.159 1	1.54	13.03
	SNV	0.965 8	0.145 1	0.228 8	0.901 3	0.179 4	2.17	13.17
	MSC	0.965 2	0.146 3	0.230 7	0.912 8	0.165 8	2.36	12.15
固含量	无预处理	0.996 4	0.530 9	0.624 4	0.987 3	0.590 6	6.38	2.78
	移动窗口平滑	0.996 1	0.550 3	0.639 1	0.987 9	0.600 0	6.50	2.81
	S-G 1 st	0.999 0	0.279 7	0.681 6	0.988 5	0.693 0	6.22	3.25
	归一化法	0.996 2	0.545 1	0.643 4	0.984 5	0.695 9	5.14	3.15
	基线校正	0.999 1	0.266 9	0.609 9	0.992 8	0.482 5	8.11	2.39
	SNV	0.999 8	0.131 6	0.498 9	0.992 1	0.479 3	8.11	2.25
	MSC	0.999 6	0.174 6	0.51 56	0.993 4	0.437 6	8.86	2.05

R_{cal} 、RMSEC、RMSECV 最优, 但 R_{pre} 、RMSEP、RPD 和 RSEP 均次于 SNV 方法。S-G 1st 方法 R_{pre} 、RPD 最优, 但 R_{cal} 、RMSEC、RMSECV、RMSEP、RSEP 均次于 SNV 方法。因此, 采用 SNV 预处理后建模效果最好, 其 RPD 为 2.57, RSEP 为 6.59%。新绿原酸、隐绿原酸、异绿原酸 B、断氧化马钱子苷采用 S-G 1st 方法处理后建模效果最好, 其 RPD 分别为 3.61、4.40、3.68、3.02, RSEP 分别为 9.61%、6.53%、8.07%、10.55%。异绿原酸 A 经移动窗口平滑方法处理后建模效果最好, 其 RPD 为 3.01, RSEP 为 7.84%。异绿原酸 C 采用基线校正方法处理后建模效果最好, 其 RPD 为 4.11, RSEP 为 6.03%。固含量经 MSC 方法处理后建模效果最好, 其 RPD 为 8.86, RSEP 为 2.05%。

2.9 特征波段筛选结果

2.9.1 NIRS 特征波段筛选结果 3 种波段筛选方法对 NIRS 建模效果的影响如表 4 所示。综合考虑各项评价指标, 如绿原酸采用 siPLS 筛选波段的各项评价指标均优于其他筛选方法, 其 RPD 为 11.70,

RSEP 为 1.66%, 最佳建模波段为 4 300.480~4 597.465、5 804.685~6 101.669、6 105.525~6 402.510 cm^{-1} 。异绿原酸 A、断氧化马钱子苷的波段筛选方法中采用 iPLS 筛选波段的建模效果最好, 其分别 RPD 为 4.39、6.25, RSEP 分别为 5.11%、4.21%, 最佳建模波段分别为 5 804.685~6 101.669 cm^{-1} ; 4 300.480~4 597.465 cm^{-1} 。新绿原酸、隐绿原酸、异绿原酸 B、异绿原酸 C、固含量经过波段筛选后, RPD 变小、RSEP 变大, 预测性能下降, 因此使用全光谱的建模效果最好, 其 RPD 分别为 10.00、14.67、7.10、4.87、22.06, RSEP 分别为 3.42%、1.91%、4.55%、4.48%、1.01%。除新绿原酸、隐绿原酸、异绿原酸 B、异绿原酸 C、固含量使用全光谱建模外, 绿原酸、异绿原酸 A、断氧化马钱子苷的建模波段在光谱上的位置如图 3 所示。

2.9.2 MIRS 特征波段筛选结果 3 种波段筛选方法对 NIRS 建模效果的影响如表 5 所示。综合考虑各项评价指标, 如绿原酸的筛选方法中, 虽然 iPLS 的 R_{pre} 最优, 但 R_{cal} 、RMSEC、RMSECV、RMSEP、

表 4 不同波段建模对 NIRS 模型的影响

Table 4 Effects of different waveband modeling on NIRS models

指标	筛选方法	建模波段/ cm^{-1}	R_{cal}	RMSEC/ ($\text{mg}\cdot\text{g}^{-1}$)	RMSECV/ ($\text{mg}\cdot\text{g}^{-1}$)	R_{pre}	RMSEP/ ($\text{mg}\cdot\text{g}^{-1}$)	RPD	RSEP/ %
绿原酸	全光谱	3 999.640~10 001.030	0.993 1	0.280 5	0.359 3	0.981 6	0.360 3	4.65	4.18
	siPLS	4 300.480~4 597.465, 5 804.685~6 101.669, 6 105.525~6 402.510	0.998 3	0.138 4	0.235 8	0.997 0	0.143 2	11.70	1.66
	iPLS	5 804.685~6 101.669	0.992 5	0.293 6	0.398 1	0.972 2	0.381 9	4.39	4.44
	mwPLS	5 847.111~6 086.241	0.988 9	0.359 0	0.473 7	0.979 1	0.334 0	5.02	3.88
新绿原酸	全光谱	3 999.640~10 001.030	0.995 3	0.049 4	0.112 0	0.995 2	0.043 2	10.00	3.42
	siPLS	5 203.003~5 499.987, 5 503.844~5 800.828, 7 308.889~7 605.873	0.998 5	0.028 4	0.081 2	0.992 0	0.055 0	7.86	4.34
	iPLS	9 113.934~9 407.061	0.935 0	0.181 2	0.275 8	0.958 0	0.136 8	3.16	10.80
	mwPLS	4 304.338~4 543.468	0.990 8	0.068 9	0.162 4	0.980 8	0.099 0	4.37	7.82
隐绿原酸	全光谱	3 999.640~10 001.030	0.998 8	0.030 2	0.054 7	0.997 6	0.032 5	14.67	1.91
	siPLS	3 999.640~4 296.624, 5 203.003~5 499.987, 5 804.685~6 101.669	0.998 2	0.037 8	0.067 8	0.991 8	0.064 7	7.37	3.79
	iPLS	7 008.048~7 305.032	0.982 5	0.116 5	0.156 5	0.951 6	0.152 0	3.14	8.90
	mwPLS	4 393.047~4 632.177	0.994 8	0.063 7	0.103 8	0.982 5	0.098 0	4.87	5.74
异绿原酸 A	全光谱	3 999.640~10 001.030	0.987 1	0.088 5	0.121 8	0.966 7	0.083 2	3.88	5.79
	siPLS	5 804.685~6 101.669, 9 113.934~9 407.061, 9 707.901~10 001.030	0.998 5	0.030 4	0.089 1	0.972 6	0.083 9	3.85	5.84
	iPLS	5 804.685~6 101.669	0.983 0	0.101 4	0.120 5	0.981 3	0.073 5	4.39	5.11
	mwPLS	7 285.747~7 524.877	0.985 2	0.094 9	0.127 1	0.937 7	0.122 8	2.63	8.54

续表 4

指标	筛选方法	建模波段/ cm^{-1}	R_{cal}	RMSEC/ ($\text{mg}\cdot\text{g}^{-1}$)	RMSECV/ ($\text{mg}\cdot\text{g}^{-1}$)	R_{pre}	RMSEP/ ($\text{mg}\cdot\text{g}^{-1}$)	RPD	RSEP/ %
异绿原酸 B	全光谱	3 999.640~10 001.030	0.993 1	0.047 3	0.061 8	0.990 8	0.041 4	7.10	4.55
	siPLS	6 707.207~7 004.191, 8 813.093~9 110.077, 9 113.934~9 407.061	0.999 7	0.009 8	0.050 2	0.984 3	0.057 7	5.10	6.34
	iPLS	5 503.844~5 800.828	0.991 9	0.051 2	0.064 2	0.990 6	0.041 8	7.03	4.60
	mwPLS	5 418.991~5 658.121	0.992 5	0.049 2	0.065 4	0.988 0	0.049 9	5.89	5.49
异绿原酸 C	全光谱	3 999.640~10 001.030	0.988 5	0.109 1	0.134 5	0.985 2	0.090 3	4.87	4.48
	siPLS	5 804.685~6 101.669, 9 113.934~9 407.061, 9 707.901~10 001.030	0.999 3	0.027 8	0.082 8	0.980 4	0.122 6	3.58	6.08
	iPLS	5 804.685~6 101.669	0.991 8	0.092 3	0.107 7	0.980 3	0.121 2	3.62	6.01
	mwPLS	5 600.268~5 839.397	0.993 6	0.081 9	0.116 0	0.972 4	0.156 9	2.80	7.78
断氧化马钱子苷	全光谱	3 999.640~10 001.030	0.968 5	0.140 5	0.168 2	0.982 6	0.107 7	3.27	8.05
	siPLS	3 999.640~4 296.624, 4 300.480~4 597.465, 5 203.003~5 499.987	0.991 2	0.074 6	0.090 3	0.980 2	0.072 0	4.89	5.38
	iPLS	4 300.480~4 597.465	0.993 3	0.065 5	0.075 4	0.990 5	0.056 3	6.25	4.21
	mwPLS	4 258.055~4 497.185	0.994 5	0.059 3	0.076 6	0.985 4	0.065 4	5.39	4.88
固含量	全光谱	3 999.640~10 001.030	0.999 2	0.237 7	0.316 1	0.999 0	0.190 5	22.06	1.01
	siPLS	5 203.003~5 499.987, 6 105.525~6 402.510, 7 910.570~8 207.555	0.999 3	0.228 9	0.280 9	0.998 5	0.225 0	18.68	1.19
	iPLS	7 308.889~7 605.873	0.999 3	0.225 4	0.330 3	0.998 3	0.257 7	16.31	1.36
	mwPLS	7 046.617~7 285.747	0.9986	0.321 9	0.402 4	0.998 0	0.275 2	15.28	1.45

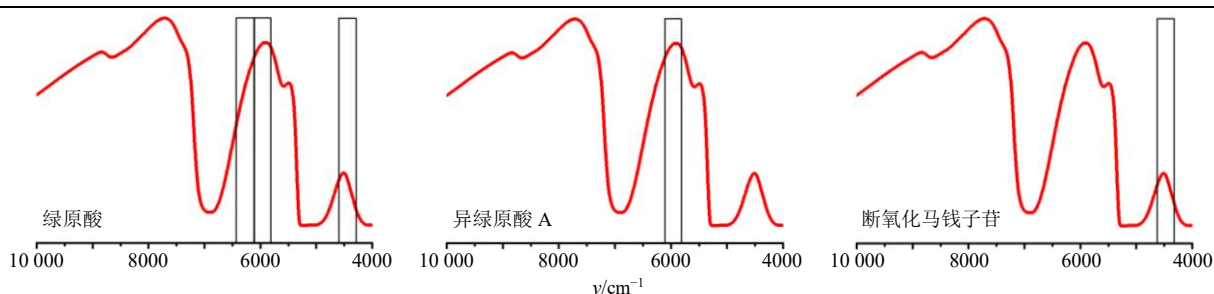


图 3 NIRS 特征波段筛选结果

Fig. 3 NIRS feature waveband screening results

RPD 和 RSEP 均为 siPLS 最优，因此选择了 siPLS。

从结果来看，绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、断氧化马钱子苷、固含量 7 个指标采用 siPLS 方法筛选出的特征波段建模效果最好。其中，绿原酸模型的 RPD 为 6.89，RSEP 为 2.46%，最佳建模波段为 996~1108、1576~1688、1692~1804 cm^{-1} 。新绿原酸模型的 RPD 为 7.51，RSEP 为 5.37%，最佳建模波段为 996~1108、1228~1340、1692~1804 cm^{-1} 。隐绿原酸模型的 RPD 为 9.74，RSEP 为 2.95%，最佳建模波段为 1112~1224、1344~1456、1692~1804 cm^{-1} 。异绿原酸 A 模型的

RPD 为 4.95，RSEP 为 4.77%，最佳建模波段为 1344~1456、1460~1572、1692~1804 cm^{-1} 。异绿原酸 B 模型的 RPD 为 8.98，RSEP 为 3.31%，最佳建模波段为 1112~1224、1460~1572、1692~1804 cm^{-1} 。断氧化马钱子苷模型的 RPD 为 6.50，RSEP 为 4.90%，最佳建模波段为 996~1108、1344~1456、1692~1804 cm^{-1} 。固含量模型的 RPD 为 12.05，RSEP 为 1.51%，最佳建模波段为 996~1108、1228~1340、1576~1688 cm^{-1} 。异绿原酸 C 采用 iPLS 方法筛选出的特征波段建模效果最好，其 RPD 为 4.73，RSEP 为 5.24%，最佳建模波段为 1344~1456 cm^{-1} 。

表 5 不同波段建模对 MIRS 模型的影响
Table 5 Effects of different waveband modeling on MIRS models

指标	筛选方法	建模波段/cm ⁻¹	<i>R</i> _{cal}	RMSEC/ (mg·g ⁻¹)	RMSECV/ (mg·g ⁻¹)	<i>R</i> _{pre}	RMSEP/ (mg·g ⁻¹)	RPD	RSEP/ %
绿原酸	全光谱	648~3000	0.982 6	0.452 9	0.648 6	0.919 7	0.581 3	2.57	6.59
	siPLS	996~1108, 1576~1688, 1692~1804	0.999 2	0.096 1	0.198 1	0.989 0	0.217 3	6.89	2.46
	iPLS	1460~1572	0.993 1	0.286 4	0.404 4	0.990 8	0.218 7	6.84	2.48
	mwPLS	896~1128	0.988 3	0.371 2	0.499 8	0.960 3	0.444 9	3.36	5.04
新绿原酸	全光谱	648~3000	0.979 6	0.103 6	0.232 3	0.960 0	0.112 2	3.61	9.61
	siPLS	996~1108, 1228~1340, 1692~1804	0.980 0	0.100 1	0.130 3	0.993 9	0.058 0	7.51	5.37
	iPLS	996~1108	0.963 9	0.137 2	0.160 5	0.965 0	0.127 4	3.18	10.91
	mwPLS	856~1088	0.977 7	0.108 2	0.181 2	0.967 9	0.119 3	3.39	10.22
隐绿原酸	全光谱	648~3000	0.996 1	0.054 7	0.161 2	0.974 1	0.108 4	4.40	6.53
	siPLS	1112~1224, 1344~1456, 1692~1804	0.999 1	0.026 7	0.052 2	0.994 8	0.049 0	9.74	2.95
	iPLS	1692~1804	0.984 4	0.109 6	0.145 7	0.965 2	0.130 9	3.64	7.89
	mwPLS	868~1110	0.993 9	0.068 7	0.113 3	0.967 8	0.125 4	3.80	7.55
异绿原酸 A	全光谱	648~3000	0.965 9	0.141 6	0.188 2	0.957 1	0.150 1	3.01	7.84
	siPLS	1344~1456, 1460~1572, 1692~1804	0.991 6	0.070 7	0.116 2	0.981 8	0.091 3	4.95	4.77
	iPLS	1344~1456	0.969 8	0.133 5	0.159 1	0.971 4	0.105 7	4.28	5.52
	mwPLS	1428~1660	0.970 2	0.096 7	0.154 7	0.932 1	0.165 6	2.73	8.65
异绿原酸 B	全光谱	648~3000	0.988 5	0.061 8	0.098 8	0.972 1	0.091 8	3.68	8.07
	siPLS	1112~1224, 1460~1572, 1692~1804	0.992 8	0.049 0	0.062 0	0.993 5	0.037 6	8.98	3.31
	iPLS	880~992	0.973 4	0.093 4	0.102 7	0.942 5	0.118 6	2.85	10.42
	mwPLS	880~1112	0.988 0	0.063 1	0.077 9	0.984 8	0.059 9	5.64	5.26
异绿原酸 C	全光谱	648~3000	0.988 5	0.109 5	0.168 0	0.971 1	0.141 2	4.11	6.03
	siPLS	880~992, 1460~1572, 1692~1804	0.997 4	0.051 8	0.1061	0.941 4	0.192 8	3.01	8.23
	iPLS	1344~1456	0.988 8	0.108 1	0.159 2	0.980 7	0.122 8	4.73	5.24
	mwPLS	1352~1584	0.989 3	0.105 9	0.125 0	0.989 3	0.132 0	4.40	5.63
断氧化马钱子苷	全光谱	648~3000	0.921 2	0.214 5	0.281 2	0.945 2	0.145 9	3.02	10.55
	siPLS	996~1108, 1344~1456, 1692~1804	0.997 6	0.038 1	0.074 1	0.988 1	0.067 7	6.50	4.90
	iPLS	996~1108	0.986 6	0.090 2	0.118 3	0.952 9	0.147 3	2.99	10.65
	mwPLS	860~1092	0.989 9	0.078 4	0.127 5	0.951 8	0.141 3	3.12	10.22
固含量	全光谱	648~3000	0.999 6	0.174 6	0.515 6	0.993 4	0.437 6	8.86	2.05
	siPLS	996~1108, 1228~1340, 1576~1688	0.999 0	0.283 9	0.414 9	0.996 9	0.322 1	12.05	1.51
	iPLS	1460~1572	0.997 9	0.402 7	0.536 3	0.984 2	0.699 1	5.55	3.28
	mwPLS	1384~1616	0.999 1	0.261 2	0.495 9	0.990 1	0.560 4	6.92	2.63

8 个质控指标的 MIRS 特征波段在全光谱上的位置如图 4 所示。

经过波段筛选,模型的预测性能均有大幅提升。这是由于 MIRS 属于分子振动-转动光谱,富含大量有机官能团信息,经过筛选得到的特征波段,可以对应有机物的特征吸收峰,排除一些相关性差的吸收峰,使得模型预测性能提升。如这 7 个有效成分

的化学结构式中均含有羰基,其伸缩振动产生的吸收峰是羧酸类成分的特征吸收峰,峰位为 1650~1740 cm⁻¹,这与筛选的 1692~1804 cm⁻¹ 波段相对应。而 NIRS 主要是 O-H、N-H 及 C-H 键倍频信息,有机物其他结构在此区域一般无特征信息,波段筛选后模型预测性能提升并不明显,有的成分甚至不如全光谱建模。

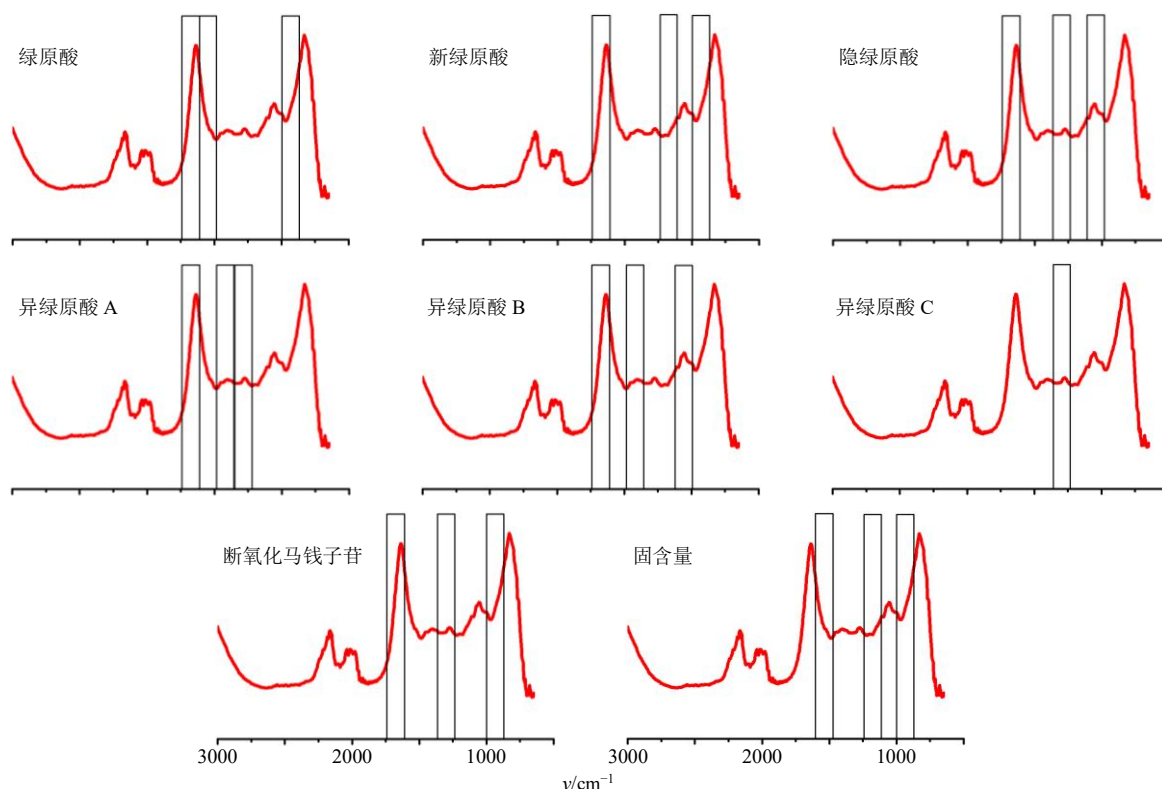


图 4 MIRS 特征波段筛选结果

Fig. 4 MIRS feature waveband screening results

2.10 含量预测模型的建立

依据上述步骤得到的最佳光谱预处理方法和最佳建模波段建立 PLS 模型，分别计算 NIRS 模型与 MIRS 模型对 8 个质控指标的预测值与参考值之间的 ARPE。其中，使用 NIRS 建模得到的绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量 8 个定量预测模型的 ARPE 分别为 1.57%、4.05%、1.88%、4.74%、4.39%、4.13%、3.79%、0.94%；使用 MIRS 建模得到的绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量 8 个定量预测模型的 ARPE 分别为 2.03%、4.00%、2.83%、4.01%、2.32%、4.28%、4.50%、1.42%。由此发现，16 个模型的 ARPE 均在 5% 以下，模型的预测性能很好，而且 NIRS 模型对绿原酸、隐绿原酸、异绿原酸 C、断氧化马钱子苷、固含量的预测效果更好，MIRS 模型对新绿原酸、异绿原酸 A、异绿原酸 B 的预测效果更好。因此，绿原酸、隐绿原酸、异绿原酸 C、断氧化马钱子苷、固含量 5 个指标应用 NIRS 建立的模型，新绿原酸、异绿原酸 A、异绿原酸 B 3 个指标应用 MIRS 建立的模型。

综上，8 种质控指标的最优含量预测模型采用

的光谱、预处理方法、建模波段以及模型评价参数如表 6 所示，8 个模型预测值与参考值的偏差如图 5 所示。

2.11 模型的外部验证

将 Z220503 批次的 24 个外部验证集样本的光谱数据导入 8 个模型中，计算得到绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量的 ARPE 分别为 6.22%、7.94%、5.44%、5.05%、5.37%、7.44%、6.00%、2.74%，均在 8% 以下。

3 讨论

本研究应用 HPLC 法和固含量测定法测定绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量为 2.57~14.80 mg/g、0.28~2.54 mg/g、0.45~3.11 mg/g、0.33~2.97 mg/g、0.19~1.79 mg/g、0.44~3.73 mg/g、0.48~2.86 mg/g、5.11%~32.98%。可以发现新绿原酸、异绿原酸 A、异绿原酸 B 的最低质量分数均在 0.4 mg/g 以下。其次，运用 NIRS 技术和 MIRS 技术对热毒宁注射液金银花浓缩过程绿原酸、新绿原酸、隐绿原酸、异绿原酸 A、异绿原酸 B、异绿原酸 C、断氧化马钱子苷和固含量共 8 个质控指标进

表 6 8 个质控指标最佳模型的评价参数

Table 6 Evaluation parameters of the best model for eight quality control indicators

指标	光谱	预处理	波段 筛选	R_{cal}	RMSEC/ ($\text{mg}\cdot\text{g}^{-1}$)	RMSECV/ ($\text{mg}\cdot\text{g}^{-1}$)	R_{pre}	RMSEP/ ($\text{mg}\cdot\text{g}^{-1}$)	RPD	RSEP/ %	ARPE/ %
绿原酸	NIRS	S-G 1 st	siPLS	0.998 3	0.138 4	0.235 8	0.997 0	0.143 2	11.70	1.66	1.57
新绿原酸	MIRS	S-G 1 st	siPLS	0.980 0	0.100 1	0.130 3	0.993 9	0.058 0	7.51	5.37	4.00
隐绿原酸	NIRS	基线校正	全光谱	0.998 8	0.030 2	0.054 7	0.997 6	0.032 5	14.67	1.91	1.88
异绿原酸 A	MIRS	移动窗口平滑	siPLS	0.991 6	0.070 7	0.116 2	0.981 8	0.091 3	4.95	4.77	4.01
异绿原酸 B	MIRS	S-G 1 st	siPLS	0.992 8	0.049 0	0.062 0	0.993 5	0.037 6	8.98	3.31	2.32
异绿原酸 C	NIRS	SNV	全光谱	0.988 5	0.109 1	0.134 5	0.985 2	0.090 3	4.87	4.48	4.13
断氧化马钱子苷	NIRS	S-G 1 st	iPLS	0.993 3	0.065 5	0.075 4	0.990 5	0.056 3	6.25	4.21	3.79
固含量	NIRS	归一化	全光谱	0.999 2	0.237 7	0.316 1	0.999 0	0.190 5	22.06	1.01	0.94

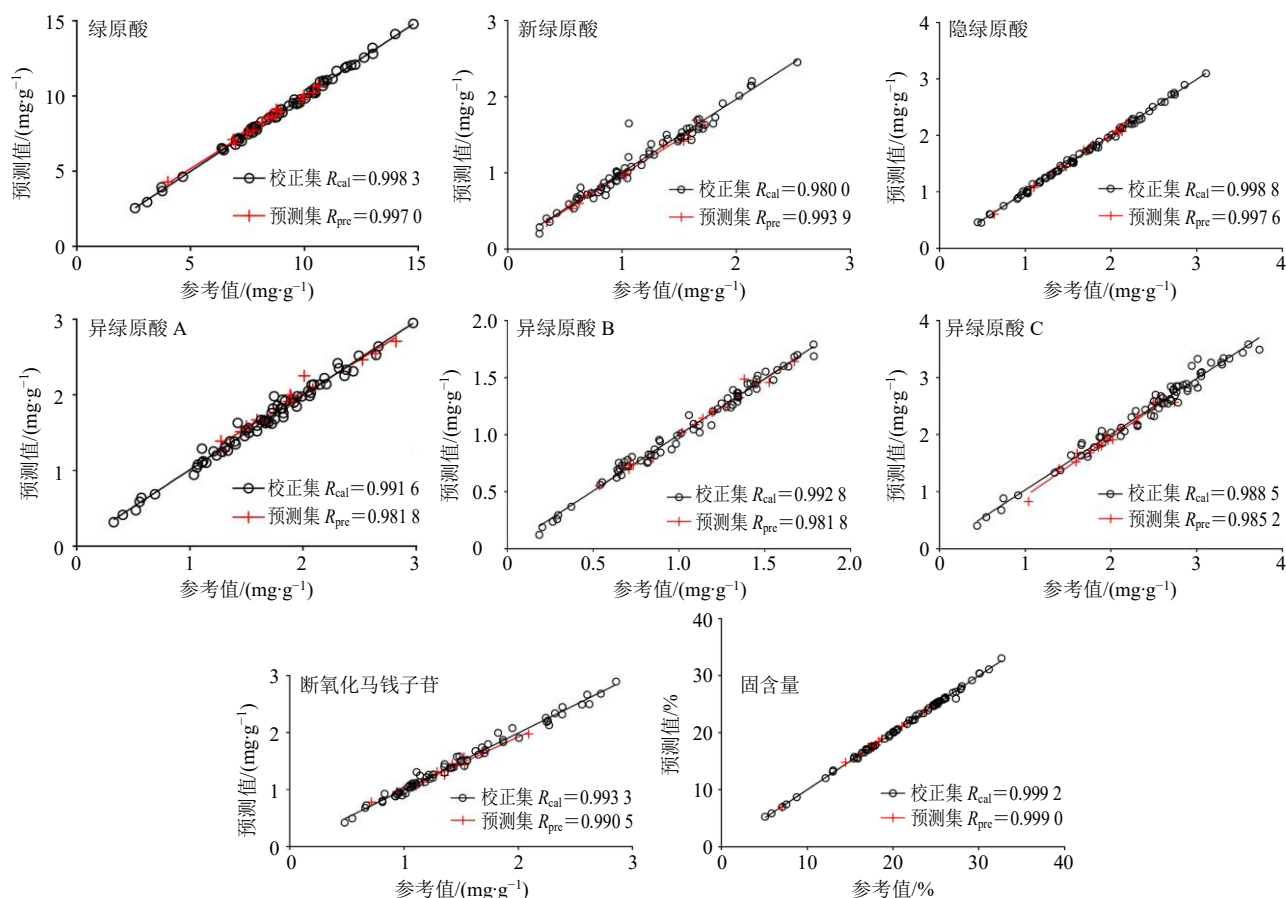


图 5 8 个质控指标预测值与参考值的相关性

Fig. 5 Correlation between predicted value and reference value of eight quality control indicators

行建模研究。结果表明, NIRS 对绿原酸、隐绿原酸、异绿原酸 C、断氧化马钱子苷、固含量的预测效果更好, ARPE 分别为 1.57%、1.88%、4.13%、3.79%、0.94%。MIRS 对新绿原酸、异绿原酸 A、异绿原酸 B 的预测效果更好, ARPE 分别为 4.00%、4.01%、2.32%。通过对比各指标成分含量预测模型验证后的 ARPE 后发现 0.4 mg/g 以下低含量的样品

会影响 NIRS 的检测效果, 而 MIRS 技术对热毒宁注射液金银花浓缩过程低含量指标预测结果更精确, 这与 MIRS 技术较 NIRS 技术灵敏度更高相对应。最后, 对 8 个模型进行外部验证, 计算得到的 ARPE 均在 8% 以下, 符合外部验证 ARPE 在 10% 的要求^[14], 模型效果良好可靠。

由此可见, 除现在广泛应用于热毒宁注射液生

产过程质控指标含量预测的 NIRS 技术外, MIRS ATR-FTIR 技术对热毒宁注射液生产过程质控指标含量预测也有良好的预测性能。但是, 同样作为一种简便、快速、无损的检测手段, NIRS 技术会受到样品粒度、水分含量、颜色、杂质含量的影响^[22]。而 MIRS ATR 附件基于光内反射原理而设计, MIRS ATR-FTIR 技术可以通过样品表面的反射信号获得样品表层有机成分的结构信息, 不会受到这一影响。并且从研究结果发现, 相较于 NIRS 技术, MIRS ATR-FTIR 技术对于含量较低的指标预测效果更好。因此, 在应用 NIRS 技术的同时, 可以考虑应用 MIRS ATR-FTIR 分析技术, 避免样品本身物理性质对分析结果的影响, 提高对低含量指标的检测灵敏度, 提高模型精度, 提升含量预测的准确度, 为中药生产过程的质量控制提供另一方案。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 余俭. 抗菌抗病毒新药: 热毒宁注射液 [J]. 中南药学, 2010, 8(7): 548-550.
- [2] 黄小民, 柳于介, 何煜舟, 等. 热毒宁注射液治疗急性上呼吸道感染的临床研究 [J]. 中国临床药理学与治疗学, 2006, 11(4): 470-473.
- [3] 王高举, 焦红军. 基于网络药理学和分子对接技术的热毒宁注射液抗 SARS、MERS 和 COVID-19 的潜在共性作用机制与活性成分研究 [J]. 药物评价研究, 2022, 45(5): 842-852.
- [4] 韩旭, 赵薇. 金银花化学成分研究进展 [J]. 农家参谋, 2018(22): 54.
- [5] 吴娇, 王聪, 于海川. 金银花中的化学成分及其药理作用研究进展 [J]. 中国实验方剂学杂志, 2019, 25(4): 225-234.
- [6] 王磊, 杨越, 李页瑞, 等. 热毒宁注射液金银花提取浓缩工段过程性能指数研究 [J]. 中草药, 2017, 48(14): 2864-2869.
- [7] 王馨, 徐冰, 徐翔, 等. 中药质量源于设计方法和应用: 过程分析技术 [J]. 世界中医药, 2018, 13(3): 527-534.
- [8] 刘雪松, 陶玲艳, 杜文俊, 等. 基于近红外光谱的热毒宁注射液萃取浓缩过程快速检测研究 [J]. 时珍国医国药, 2015, 26(9): 2144-2146.
- [9] 王永香, 米慧娟, 张传力, 等. 近红外光谱技术用于热毒宁注射液金银花青蒿醇沉过程在线监测研究 [J]. 中国中药杂志, 2014, 39(23): 4608-4614.
- [10] 徐芳芳, 冯双双, 李雪珂, 等. 青蒿浓缩过程在线近红外快速检测模型的建立 [J]. 中草药, 2016, 47(10): 1690-1695.
- [11] 吴莎, 刘启安, 王伟, 等. 近红外光谱技术在热毒宁注射液萃取工艺过程质量控制研究 [J]. 中国中药杂志, 2015, 40(3): 437-442.
- [12] 李梦, 刘杰, 吕晓娜, 等. ATR-FTIR 技术在进口血竭掺伪鉴别中的应用 [J]. 世界中医药, 2016, 11(10): 2123-2125, 2129.
- [13] 于喆源, 杨玺文, 王晓霞, 等. 中红外光谱技术对临泽小枣中多糖的定量分析 [J]. 食品工业, 2021, 42(5): 249-252.
- [14] 徐芳芳, 杜慧, 张欣, 等. 在线中红外光谱监测热毒宁注射液金银花与青蒿醇沉过程 7 种指标成分研究 [J]. 中草药, 2021, 52(10): 2909-2917.
- [15] 黄红英, 尹齐和. 傅里叶变换衰减全反射红外光谱法 (ATR-FTIR) 的原理与应用进展 [J]. 中山大学研究生学刊: 自然科学与医学版, 2011(1): 20-31.
- [16] 赵秀琴. 中红外光谱分析技术及研究进展 [J]. 安庆师范学院学报: 自然科学版, 2012, 18(4): 94-97.
- [17] Tao L Y, Lin Z L, Chen J S, *et al.* Mid-infrared and near-infrared spectroscopy for rapid detection of *Gardenia Fructus* by a liquid-liquid extraction process [J]. *J Pharm Biomed Anal*, 2017, 145: 1-9.
- [18] 吴静珠, 王一鸣, 张小超, 等. 近红外光谱分析中定标集样品挑选方法研究 [J]. 农业机械学报, 2006, 37(4): 80-82, 101.
- [19] 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.
- [20] 第五鹏瑶, 卞希慧, 王姿方, 等. 光谱预处理方法选择研究 [J]. 光谱学与光谱分析, 2019, 39(9): 2800-2806.
- [21] 张永超, 徐芳芳, 张欣, 等. 腰痹通胶囊 4 种中间体粒径的近红外光谱通用定量模型研究 [J]. 中草药, 2021, 52(1): 55-64.
- [22] 张恩阳, 夏维高, 闫晓剑. 近红外光谱的原理及应用 [J]. 信息记录材料, 2017, 18(6): 26-27.

[责任编辑 郑礼胜]