基于近、中红外光谱的热毒宁注射液制剂过程投料和二次热处理工序快速 检测方法研究

童 枫1,徐芳芳 2.3*,张 欣 2.3,李执栋 2.3,吴 云 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 应用近红外光谱(near-infrared reflectance spectroscopy, NIRS)与中红外光谱(mid-infrared reflectance spectroscopy, MIRS)技术,对热毒宁注射液(Reduning Injection, RI)制剂过程的投料和二次热处理工序中6种质控指标进行快速检测,提高制剂过程的质量控制水平。方法 利用 NIRS 透射技术与 MIRS 衰减全反射技术,结合偏最小二乘法(partial least squares, PLS),经过光谱预处理方法的优选以及波段筛选,分别建立绿原酸、新绿原酸、隐绿原酸、栀子苷、断氧化马钱子苷和固含量的快速预测模型,以校正集相关系数(*r*cal)、验证集相关系数(*r*pre)、校正均方根误差(root mean square error of calibration, RMSEC)、交叉验证均方根误差(root mean square error of cross validation, RMSECV)、预测相对误差(relative standard error of prediction, RSEP)为评价指标,评价模型性能。结果 NIRS 预测模型的 6个质控指标的 RMSEC 和 RMSECV 均小于 0.3, RSEP 小于 4.0%; MIRS 预测模型的 6个质控指标的 RMSEC 和 RMSECV 均小于 0.4, RSEP 均小于 5.0%,建立的 PLS 模型具有模型性能好、预测精度高的优点。结论 NIRS 及 MIRS 分析技术,均可用于 RI 投料和二次热处理工序中 6 种质控指标的快速检测,模型 RSEP 在 5%以内,方法操作简单,结果可靠。 关键词:近红外光谱;中红外光谱;热毒宁注射液;制剂过程;光谱预处理;波段筛选;偏最小二乘法;绿原酸;新绿原酸;

隐绿原酸; 栀子苷; 断氧化马钱子苷; 固含量 中图分类号: R283.6 文献标志码: A

```
DOI: 10.7501/j.issn.0253-2670.2022.21.007
```

22.21.007

文章编号: 0253 - 2670(2022)21 - 6706 - 10

Research on fast detection method of feeding and secondary heat treatment process of Reduning Injection preparation process by near-infrared and mid-infrared reflectance spectroscopy

TONG Feng¹, XU Fang-fang^{2, 3}, ZHANG Xin^{2, 3}, LI Zhi-dong^{2, 3}, WU Yun^{2, 3}, ZHANG Chen-feng^{2, 3}, WANG Zhen-zhong^{2, 3}

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

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

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

Abstract: Objective The near-infrared reflectance spectroscopy (NIRS) and mid-infrared reflectance spectroscopy (MIRS) techniques were used to detect six quality control indexes in the feeding and secondary heat treatment process of Reduning Injection (热毒宁注射液, RI) preparation process, so as to improve the quality control level of preparation process. Methods The quantitative correction models of chlorogenic acid, neochlorogenic acid, cryptochlorogenic acid, gardenoside, secoxyloganin and solid content were established by using NIRS technology and attenuated total MIRS technology, combined with partial least squares (PLS) method, after spectral pretreatment method selection and band selection respectively. Calibration set correlation coefficient (*r*_{pre}), root mean square error of calibration (RMSEC), root mean square error of cross

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: 童 枫 (1998—), 男, 硕士研究生, 研究方向为中药制药过程新技术。Tel: 18851193208 E-mail: tongfeng5324@126.com *通信作者: 王振中, 研究员级高级工程师, 硕士生导师, 研究方向为中药制药过程新技术。E-mail: kanionlunwen@163.com 徐芳芳, 女, 博士, 主管药师, 研究方向为中药制药过程新技术。E-mail: 879164331@qq.com

• 6706 •

收稿日期: 2022-05-16

validation (RMSECV) and relative standard error of prediction (RSEP) were used as evaluation indexes to evaluate the model performance. **Results** The RMSEC and RMSECV of the six quality control indexes of the NIRS prediction model were both less than 0.3 and close to each other, and the RSEP value were all less than 4.0%. The RMSEC and RMSECV of the six quality control indexes of the MIRS prediction model were all less than 0.4 and close to each other, and the RSEP value were all less than 5.0%. These PLS models had the advantages of good model performance and high prediction accuracy. **Conclusion** Both NIRS and MIRS technology can be used for rapid detection of six quality control indexes in RI feeding-secondary heat treatment process, and the RSEP value is within 5%, and the method is simple and reliable.

Key words: near-infrared reflectance spectroscopy; mid-infrared reflectance spectrum; Reduning Injection; preparation process; spectral pretreatment; band screening; partial least squares; chlorogenic acid; neochlorogenic acid; cryptochlorogenic acid; gardenoside; secoxyloganin; solid content

热毒宁注射液(Reduning Injection, RI)是江 苏康缘药业股份有限公司的独家品种,由青蒿 Artemisiae Annuae Herba、金银花Lonicerae Japonicae Flos、栀子 Gardeniae Fructus 3味药材制得,具有 清热、疏风、解毒的功效,常用于上呼吸道感染所 致的高热、微恶风寒、头身痛、咳嗽、痰黄等症状 及抑制流感病毒^[1-2]。在中华人民共和国国家卫生健 康委员会发布的《新型冠状病毒感染的肺炎诊疗方 案(试行第九版)》中,针对重型、危重型新型冠状 病毒肺炎(coronavirus disease 2019, COVID-19) 患者推荐的中成药中, RI 也名列其中^[3]。

RI 的生产过程可以分为提取、精制和制剂过 程,目前,利用近红外光谱(near-infrared reflectance spectroscopy, NIRS) 或中红外光谱(mid-infrared reflectance spectroscopy, MIRS) 技术结合化学计量 学,针对 RI 提取、精制过程的快速检测已较为成熟, 并取得一定的应用效果^[4-8]。但是,基于 NIRS 和 MIRS 的 RI 制剂过程的快速检测研究尚未见报道。 因此, 基于 RI 的制剂过程, 研究合适的快速检测方 法,可为 RI 制剂过程的质量控制水平的提升提供技 术支撑。NIRS 以及 MIRS 作为过程分析技术的典型 代表,近年来发展迅速。与传统的化学分析方法相 比,这2种光谱分析方法具有以下优点:(1)可同 时用于定性和定量分析;(2)分析速度快,只需扫 描得出光谱即可测定样品中的各种成分; (3) 不使 用化学试剂,不损害样品,对环境友好;(4)分析 成本低,操作简单[9-11]。

投料和二次热处理工序指的是投料和二次热处 理2个连续工序,即在一定的温湿度以及压力下, 依次将栀子提取物以及青金提取物投入到合并一定 比例的注射用水和盐酸的配液罐中,待冷却后转移 至冷藏罐;冷藏罐内药液通过板框过滤器滤过至热 配间热配罐中,调节蒸气压力以及药液温度,一段

时间后对药液进行冷却。

投料和二次热处理工序作为 RI 制剂过程的初始工序,在实际生产中,会出现提取物溶解不充分等情况,导致有效成分的损失,直接影响最后的成品质量,因此,亟需引入1种快速检测方法,降低生产过程控制风险。本实验将 NIRS 和 MIRS 技术与化学计量学结合,首次应用于 RI 制剂过程,建立投料和二次热处理工序 6 种关键质控指标的偏最小二乘(partial least squares, PLS)定量校正模型,实现对绿原酸、新绿原酸、隐绿原酸、栀子苷、断氧化马钱子苷及固含量的快速测定。

1 材料与仪器

1.1 材料

59 批次的投料工序样本(批号 Z220123~ Z2201240、Z220201~Z220232、Z220301~Z220309) 及对应批次的 59 批次的二次热处理工序样本,共 118 批次样品,由江苏康缘药业股份有限公司水针 车间提供。对照品绿原酸(批号 110753-202018, 质量分数 96.1%)、栀子苷(批号 110749-201919, 质量分数 97.1%)购自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%)购自成都乐美天医药-德思特生 物技术有限公司;隐绿原酸对照品(批号 ST07850120,质量分数 98.04%)购自上海诗丹德标 准技术服务有限公司。

1.2 仪器

UltiMate 3000 型高效液相色谱仪、Antaris II 型 傅立叶变换 NIRS 仪,配有透射检测器、采样软件 以及 Result、CAMO 等数据处理软件,(赛默飞世 尔科技(中国)有限公司);React IR 702L 型 MIRS 仪、XP6 型电子天平,梅特勒-托利多仪器(上海) 有限公司;BSA 224S-CW 型万分之一电子天平,德 国 Sartorius 公司; KQ-500DB 型数控超声波清洗器, 昆山市超声仪器有限公司; Milli-Q Academic 型纯 水机,美国 Millipore 公司。

1.3 数据处理软件

采用 Unscramble X 10.4 (挪威 Camo Analytics 公司)软件进行光谱预处理;采用 GraphPad Prism 8.0 (美国 GraphPad Software)绘图;采用 Matlab R2018a(美国 MathWorks 公司)软件进行样本划分、 波段筛选及模型构建。

2 方法与结果

2.1 NIRS 的采集

以空气为扫描背景,采用 NIRS 技术于室温下 采集光谱,选择 1 mm 比色皿,光谱扫描范围 4000~ 10 000 cm⁻¹,扫描次数为 64 次,分辨率 4 cm⁻¹,衰 减器选择 B 模块,1 倍增益,每个样品采集 3 次,取 平均光谱,得到 59 批投料工序样本及 59 批二次热 处理工序样本,共计 118 个样本。NIRS 结果见图 1。

2.2 MIRS 的采集

在室温条件下,使用光纤探头扫描样品,以空 气为扫描背景;光谱扫描范围 3000~650 cm⁻¹;扫 描次数 32 次;分辨率 8 cm⁻¹;增益选择"low",每







个样品扫描 2 次,取其一作为样本光谱,得到 59 批投料工序样本及 59 批二次热处理工序样本,共计 118 个样本。MIRS 结果见图 2。





2.3 5种质控指标含量的测定

2.3.1 HPLC 分析色谱条件 参照崔小弟等^[12]的实验结果,色谱条件设为 Kromasil C₁₈(150 mm×4.6 mm, 6 µm)色谱柱;流动相为乙腈-0.1%磷酸水溶液,梯度洗脱: 0~5 min,15%~18%乙腈;5~15 min,18%乙腈;15~38 min,18%~31%乙腈;38~38.5 min,31%~90%乙腈;38.5~40 min,90%乙腈;体积流量0.7 mL/min;进样量10 µL;柱温20 ℃;检测波长238、327 nm。

2.3.2 供试品溶液的制备 精密吸取 1 mL 样品置 于 20 mL 量瓶中,用 50%甲醇溶解并稀释至刻度, 摇匀, 0.45 μm 微孔滤膜滤过,即得。

2.3.3 样品测定 取 118 批次样品,按照"2.3.2" 项方法制备供试品溶液,按照"2.3.1"项色谱条件 进样测定,每份样品平行 2 次测定,取平均值,作 为建模的质控指标含量实测值。6 种质控指标的含 量测定结果见表 1。

2.4 固含量的测定

参照文献方法^[13-14],称取约5g样品至已烘干

	表1	样本中 6 种质控指标统计值 (n = 2)
Table 1	Statistical va	lues of six quality control indicators in samples $(n = 2)$

医按比杆		田会旱/0/				
则	绿原酸	新绿原酸	隐绿原酸	栀子苷	断氧化马钱子苷	——————————————————————————————————————
最小值	8.768 8	2.066 8	2.304 3	12.285 4	1.116 2	7.295 3
下四分位数	9.405 6	2.349 4	2.601 2	12.791 3	1.178 7	7.765 5
上四分位数	10.279 1	2.619 4	2.898 0	14.254 1	1.432 7	8.221 8
最大值	10.606 1	2.809 4	3.052 0	15.311 1	1.506 9	8.725 7
均值	9.851 8	2.478 3	2.740 8	13.633 5	1.302 3	7.995 1

至恒定质量的称量瓶 (X_0) 中,称定质量 (X_1) ,置 烘箱 105 °C条件下烘干 5h 至恒定质量,计为 X_2 。

固含量= $(X_2-X_0)/(X_1-X_0)$

2.5 校正集与验证集的划分

采用 Kennard-Stone (K-S) 划分法^[15],将收集的 59 批投料工序样品以及对应批次的 59 批二次热处理工序样品,共计 118 批样品,以 4:1 的比例划分为校正集和验证集,最终得到校正集样本 94 个,验证集样本 24 个。

2.6 光谱预处理

采集光谱时,由于受到环境温度、湿度、仪器 状态、杂散光和随机噪声等因素的影响,红外光谱 会出现基线漂移和噪声干扰等问题,因此,在建模 前对光谱进行适当的预处理是有必要的,可以有效 地减少噪声,净化无用信息,提高模型的精度和预 测效果。常见的预处理方式有:移动平均法(moving average, MA)、Savitzky-Golay(S-G)平滑法、导 数法 (一阶导数、二阶导数)、矢量归一化法、基线 校正、标准正态变换(standard normal transformation, SNV)、多元散射校正 (multivariate scattering correction, MSC)以及上述方法的结合。MA和S-G 平滑法可以消除光谱信息中的随机噪声,提高信噪 比; 一阶导数可以消除背景的常数平移, 二阶导数 可以消除背景的线性平移, 消除背景漂移造成的影 响;矢量归一化法通过增强光谱数据间的差异,校 正由光程变化或者样品浓度改变引起的光谱变化; 基线校正可以消除仪器背景漂移对信号的影响: SNV 和 MSC 可以消除由于样品分布不均匀对光谱 造成的影响[16]。

本研究采用 MA、S-G 平滑、S-G 1st、基线校 正、归一化法以及 SNV 分别对 NIRS 和 MIRS 进行 预处理,并分别建立投料和二次热处理工序的 6 种 质控指标的 PLS 定量校正模型。根据模型评价指标 筛选出最佳的光谱预处理方法,使用留一交叉验证 法确定校正模型中的最佳的潜变量数(latent variables, LVs),以预测相对误差(relative standard error of prediction, RSEP)为主要评价指标,RSEP 值越小越好。一般而言,性能优异的预测模型应具 有较高的校正集相关系数(*r*_{cal})及验证集相关系数 (*r*_{pre})、较小且接近的校正均方根误差(root mean square error of calibration, RMSEC)和交叉验证均 方根误差(root mean square error of cross validation, RMSECV)及较低的 RSEP^[17]。上述评价指标的相 关公式为式(1)~(4)。

 $r_{\rm pre} = [1 - \sum (y - y_i)^2 / \sum (y - y_n)^2]^{1/2}$ (2)

RMSEC= $[\sum (y-y_i)^2/m]^{1/2}$ (3)

$$RSEP = [\sum (y - y_i)^2 / \sum y^2]^{1/2}$$
(4)

*m、n*为校正集、验证集样本数, y为实测值, y_i为预测值,*y_m、y_n*分别为校正集和验证集实测值的平均值

NIRS 和 MIRS 预处理结果见表 2、3,可以发 现:基于 NIRS,当绿原酸、新绿原酸、隐绿原酸、 栀子苷、断氧化马钱子苷及固含量分别以 SNV、 MA、MA、无预处理、SNV 和归一化法为预处理方 法时,RSEP 最小,且都在 4.0%以下,此时,*r*_{cal} 与 *r*_{pre} 值较大,RMSEC 与 RMSECV 值较小且接近; 基于 MIRS,当绿原酸、新绿原酸、隐绿原酸、栀 子苷、断氧化马钱子苷及固含量的预处理方法分别 为无预处理、MA、MA、MA、SNV 和无预处理时, 模型具有较低的 RSEP,均小于 5%,较大的 *r*_{cal}与 *r*_{pre} 值以及较小且接近的 RMSEC 与 RMSECV 值, 从预处理方法的比较中可以猜测,在 MIRS 中,MA 适合多数质控指标,可能是由于其减轻了光谱信息 自身携带的随机误差,提高信噪比。

2.7 光谱波段的筛选

全光谱模型的准确性和稳健性可能会因为一些 无关信息的存在而减弱,筛选光谱波段可以剔除无 用信息,提高模型的预测精度和稳定性。本研究在 上述筛选出的最佳预处理方法的基础上,对比全光 谱、间隔 PLS(interval PLS, iPLS)、组合间隔 PLS (synergy interval PLS, siPLS)及移动窗口 PLS (moving windows PLS, mwPLS)建模的效果。

2.7.1 iPLS 筛选最优波段 iPLS 是由 Nørgaard 等^[18] 提出的一种光谱区间筛选方法,通过将全光谱平均 分为若干子区间,在各子区间内以不同的潜变量数 建立独立的 PLS 回归模型^[19]。本研究将光谱区间均 分为 20 个子区间,并以 RMSECV 作为评价指标, 筛选出最优建模波段。

2.7.2 siPLS 筛选最优波段 siPLS 是在 iPLS 的基础上,对不同子区间进行任意组合,建立所有可能的 2、3或4个区间的 PLS 回归模型^[20]。本研究将光谱区间均分为 20 个子区间,以子区间组合数为 3 建立模型,以 RMSECV 作为评价指标,筛选出最优建模波段。

2.7.3 mwPLS 筛选最优波段 mwPLS 的基本思想 是将 1 个窗口沿着光谱轴连续移动,每移动 1 个波

质控指标	预处理方法	$r_{\rm cal}$	RMSEC	RMSECV	r _{pre}	RESP/%	LVs	质控指标	预处理方法	$r_{\rm cal}$	RMSEC	RMSECV	r _{pre}	RESP/%	LVs
绿原酸	无预处理	0.979 8	0.088 1	0.158 5	0.933 3	1.717 5	15	栀子苷	无预处理	0.975 4	0.187 4	0.257 1	0.974 6	1.100 9	12
	MA	0.976 0	0.096 0	0.164 5	0.924 0	1.828 1	16		MA	0.979 0	0.203 5	0.276 3	0.972 8	1.168 5	12
	S-G 平滑	0.975 7	0.096 6	0.159 7	0.926 4	1.788 7	15		S-G 平滑	0.972 5	0.197 9	0.269 6	0.974 4	1.120 0	12
	S-G 1st	0.982 5	0.082 2	0.186 0	0.922 2	1.818 2	10		S-G 1st	0.998 8	0.041 4	0.290 1	0.948 2	1.551 1	16
	基线校正	0.980 0	$0.087\ 8$	0.176 4	0.932 0	1.709 3	15		基线校正	0.974 7	0.189 9	0.274 4	0.975 0	1.132 6	12
	归一化法	0.971 3	0.104 7	0.153 8	0.945 5	1.544 1	12		归一化法	0.998 9	0.039 8	0.177 0	0.975 5	1.077 6	20
	SNV	0.969 9	0.107 2	0.152 5	0.951 9	1.473 4	12		SNV	0.999 0	0.038 7	0.175 4	0.974 7	1.086 6	20
新绿原酸	无预处理	0.834 3	0.096 8	0.127 0	0.829 9	3.934 6	10	断氧化马	无预处理	0.973 0	0.030 2	0.049 1	0.960 2	2.573 8	14
	MA	0.811 0	0.101 5	0.130 3	0.846 6	3.777 0	10	钱子苷	MA	0.984 7	0.023 0	0.048 7	0.947 5	2.860 5	18
	S-G 平滑	0.818 5	0.099 6	0.128 8	0.843 1	3.820 9	10		S-G 平滑	0.984 8	0.023 0	0.050 2	0.956 5	$2.708\ 0$	17
	S-G 1st	0.957 1	0.050 2	0.118 1	0.6864	5.0844	10		S-G 1st	0.993 8	0.014 7	0.046 8	0.965 9	2.546 3	12
	基线校正	0.799 0	0.104 3	0.134 5	0.846 9	3.807 1	9		基线校正	0.993 7	0.014 8	0.045 5	0.963 1	2.556 4	18
	归一化法	0.849 1	0.094 3	0.123 9	0.837 4	3.847 0	10		归一化法	0.974 4	0.029 7	0.047 3	0.963 8	2.533 1	14
	SNV	0.838 3	0.103 1	0.130 8	0.840 1	3.850 6	9		SNV	0.973 0	0.030 5	0.047 1	0.968 2	2.294 6	14
隐绿原酸	无预处理	0.872 4	0.085 1	0.111 6	0.942 4	2.591 1	10	固含量	无预处理	0.973 4	0.077~7	0.131 6	0.941 2	1.378 4	15
	MA	0.865 5	$0.087\ 0$	0.111 5	0.951 1	2.507 3	10		MA	0.970 5	0.081 8	0.133 2	0.936 1	1.423 9	17
	S-G 平滑	0.868 0	0.086 5	0.111 6	0.950 2	2.527 3	10		S-G 平滑	0.9677	0.085 5	0.128 8	0.943 9	1.388 2	15
	S-G 1st	0.978 0	0.036 3	0.098 5	0.922 2	3.170 5	11		S-G 1st	0.995 5	0.031 9	0.151 7	0.906 2	1.704 7	14
	基线校正	0.882 1	0.082 1	0.110 8	0.941 2	2.725 0	10		基线校正	0.975 1	0.075 3	0.138 6	0.927 7	1.488 5	15
	归一化法	0.881 6	0.082 1	0.108 1	0.949 8	2.548 2	10		归一化法	0.978 2	0.0704	0.120 6	0.952 9	1.259 1	15
	SNV	0.885 4	0.080 8	0.106 4	0.945 3	2.551 6	10		SNV	0.978 5	0.069 9	0.119 9	0.948 6	1.309 9	15

表 2 NIRS 不同预处理方法建模结果分析

 Table 2 Analysis of modeling results of different NIRS pretreatment methods

除固含量的 RMSEC 与 RMSECV 的单位为 "%"外,其余 5 种指标的 RMSEC 与 RMSECV 的单位均为 "mg·mL⁻¹",下表同 except the unit of RMSEC and RMSECV of solid content is "%", the units of RMSEC and RMSECV of the other five indicators are all "mg·mL⁻¹", same as the following table

表 3 MIRS 不同预处理方法建模结果分析

Table 3	Analysis of	f modeling	results of	different MIRS	pretreatment	methods

质控指标	预处理方法	r _{cal}	RMSEC	RMSECV	rpre	RESP/%	LVs	质控指标	预处理方法	r _{cal}	RMSEC	RMSECV	rpre	RESP/%	LVs
绿原酸	无预处理	0.924 5	0.168 2	0.267 1	0.8764	2.324 4	6	栀子苷	无预处理	0.911 4	0.350 0	0.594 3	0.723 1	3.883 1	6
	MA	0.9164	0.176 6	0.268 2	0.870 0	2.374 8	9		MA	0.968 1	0.212 9	0.577 8	0.822 9	3.093 0	19
	S-G 平滑	0.938 2	0.152 6	0.280 3	0.866 6	2.352 6	9		S-G 平滑	0.908 1	0.356 2	0.602 9	0.704 3	4.062 3	8
	S-G 1st	0.929 2	0.163 1	0.313 2	0.849 6	2.529 3	8		S-G 1st	0.963 3	0.228 1	0.667 6	0.745 6	4.248 7	16
	基线校正	0.925 9	0.166 4	0.293 3	0.861 0	2.390 2	7		基线校正	0.912 7	0.347 4	0.608 1	0.720 0	3.761 0	6
	归一化法	0.900 1	0.191 9	0.308 8	0.841 8	2.623 8	6		归一化法	0.922 0	0.329 1	0.657 6	0.696 9	4.207 2	8
	SNV	0.909 9	0.182 7	0.312 6	0.836 8	2.639 7	6		SNV	0.897 6	0.374 7	0.655 2	0.652 6	4.355 5	6
新绿原酸	无预处理	0.802 1	0.103 5	0.150 1	0.738 5	4.691 8	5	断氧化马	无预处理	0.8707	0.0650	0.1133	0.780 1	5.630 6	6
	MA	0.806 4	0.107 1	0.160 4	0.765 1	4.447 5	8	钱子苷	MA	0.8367	0.0724	0.0972	0.7364	6.110 4	7
	S-G 平滑	0.744 8	0.115 7	0.156 4	0.701 7	4.935 9	5		S-G 平滑	0.869 9	0.0652	0.1112	0.739 8	6.084 8	8
	S-G 1st	0.798 3	0.104 4	0.163 1	0.774 1	4.728 1	6		S-G 1st	0.967 3	0.0335	0.0977	0.727 3	7.128 5	16
	基线校正	0.815 6	0.100 4	0.154 9	0.721 8	4.877 8	5		基线校正	0.900 6	0.0576	0.1121	0.802 8	5.308 2	7
	归一化法	0.812 9	0.100 9	0.144 1	0.7507	4.566 5	5		归一化法	0.883 0	0.0620	0.1080	0.787 8	5.455 5	6
	SNV	0.808 9	0.101 9	0.146 2	0.743 3	4.634 8	5		SNV	0.897 0	0.0584	0.1085	0.806 8	5.247 9	6
隐绿原酸	无预处理	0.893 6	0.087 8	0.134 2	0.799 6	4.848 5	5	固含量	无预处理	0.914 6	0.138 5	0.230 2	0.784 2	2.547 0	6
	MA	0.830 8	0.096 9	0.145 0	0.870 2	4.172 6	8		MA	0.959 3	0.095 7	0.221 9	0.796 6	2.822 9	16
	S-G 平滑	0.863 5	0.087 7	0.148 1	0.802 0	4.798 0	7		S-G 平滑	0.882 0	0.159 9	0.239 1	0.718 9	2.882 6	6
	S-G 1st	0.821 6	0.099 1	0.1514	0.831 9	4.441 3	5		S-G 1st	0.922 7	0.121 3	0.265 9	0.775 1	2.748 0	10
	基线校正	0.842 7	0.093 7	0.145 1	0.801 7	4.770 0	5		基线校正	0.901 3	0.147 0	0.241 6	0.765 8	2.596 6	6
	归一化法	0.852 3	0.091 0	0.134 1	0.813 2	4.855 3	5		归一化法	0.908 8	0.141 4	0.256 2	0.725 1	2.931 1	7
	SNV	0.847 0	0.092 4	0.137 8	0.822 9	4.587 0	5		SNV	0.910 8	0.139 9	0.253 8	0.707 8	3.010 7	7

长点,采用交叉验证法确定最佳潜变量数并建立相应的 PLS 最优模型^[21]。本研究以全光谱波数的 10% 作为窗口宽度,以 RMSECV 作为评价指标,筛选出最优建模波段。

2.7.4 最优波段筛选结果 由 3 种波段筛选方法建 立的模型与全光谱模型的性能参数对比结果见表 4,以 RMSEC、RMSECV 和 RSEP 作为评价指标, 综合评价,筛选出最优建模波段。结果表明,基于 NIRS,绿原酸和断氧化马钱子苷的预测模型经 3 种 方法筛选波段后,虽然建模变量数减少,但 RMSEC、 RMSECV 普遍增大,且 RSEP 均增大,模型预测性 能下降,故选择全光谱区间(3 999.64~9 999.10 cm⁻¹)建模;新绿原酸、隐绿原酸及栀子苷预测模 型,采用 siPLS 筛选波段后,建模变量数降为 465 个,*r*cal、*r*pre 均增大,RMSEC、RMSECV 均减小, 且 RSEP 均减小,模型预测性能提升,故新绿原酸、 隐绿原酸及栀子苷分别选择 4 601.32~4 900.23、 5 503.84~5 802.76、7 609.73~7 906.71 cm⁻¹, 5 503.84~5 802.76、5 804.69~6 103.60、7 308.89~ 7 607.80 cm⁻¹及 4 300.48~4 599.39、5 203.00~

表 4 不同波段筛选方法对 NIRS 及 MIRS 的影响 Table 4 Influences of different band screening methods on NIRS and MIRS spectra

质控指标	光谱	筛选方法	v/cm^{-1}	rcal	RMSEC	RMSECV	rpre	RESP/%	LVs
绿原酸	NIRS	全光谱	3 999.64~9 999.10	0.969 9	0.107 2	0.152 5	0.951 9	1.473 4	12
		siPLS	4 300.48~4 599.39, 7 008.05~7 306.96,	0.978 4	0.091 1	0.138 9	0.940 8	1.606 2	13
			7 609.73~7 906.71						
		iPLS	4 300.48~4 599.39	0.974 1	0.099 6	0.162 8	0.903 7	2.014 6	10
		mwPLS	4 387.26~4 987.02	0.922 7	0.169 7	0.194 3	0.923 9	1.865 7	8
	MIRS	全光谱	3 000.00~648.00	0.924 5	0.168 2	0.267 1	0.876 4	2.324 4	6
		siPLS	1 572.00~1 460.00, 1 224.00~1 112.00,	0.940 6	0.149 6	0.194 3	0.902 4	2.093 7	12
			1 108.00~996.00						
		iPLS	1 224.00~1 112.00	0.886 3	0.205 9	0.244 9	0.826 2	2.743 9	10
		mwPLS	1 176.00~944.00	0.865 5	0.220 6	0.296 3	0.863 3	2.453 4	11
新绿原酸	NIRS	全光谱	3 999.64~9 999.10	0.811 0	0.101 5	0.130 3	0.846 6	3.777 0	10
		siPLS	4 601.32~4 900.23, 5 503.84~5 802.76,	0.877 3	0.083 3	0.108 9	0.870 5	3.431 8	12
			7 609.73~7 906.71						
		iPLS	4 601.32~4 900.23	0.767 2	0.111 9	0.147 5	0.890 0	3.192 0	6
		mwPLS	6 067.60~6 676.35	0.673 0	0.128 3	0.145 6	0.762 1	4.545 6	6
	MIRS	全光谱	3 000.00~648.00	0.806 4	0.107 1	0.1604	0.765 1	4.447 5	8
		siPLS	1 456.00~1 344.00, 1 340.00~1 228.00,	0.800 8	0.103 8	0.121 9	0.738 5	4.661 2	10
			1 108.00~996.00						
		iPLS	1 224.00~1 112.00	0.645 4	0.132 7	0.1467	0.717 2	4.811 1	8
		mwPLS	1 401.00~1 172.00	0.767 6	0.129 9	0.147 9	0.790 3	4.317 9	9
隐绿原酸	NIRS	全光谱	3 999.64~9 999.10	0.865 5	0.087~0	0.111 5	0.951 1	2.507 3	10
		siPLS	5 503.84~5 802.76, 5 804.69~6 103.60,	0.900 6	0.075 7	0.099 3	0.952 6	2.589 5	10
			7 308.89~7 607.80						
		iPLS	4 601.32~4 900.23	0.844 4	0.094 4	0.121 1	0.921 7	3.433 0	7
		mwPLS	4 697.75~5 297.50	0.872 4	0.085 6	0.169 7	0.855 7	4.340 0	7
	MIRS	全光谱	3 000.00~648.00	0.830 8	0.096 9	0.145 0	0.870 2	4.172 6	8
		siPLS	1 340.00~1 228.00, 1 108.00~996.00,	0.763 8	0.112 3	0.135 1	0.829 4	4.464 4	6
			992.00~880.00						
		iPLS	1 224.00~1 112.00	0.777 9	0.109 5	0.129 8	0.834 2	4.453 0	9
		mwPLS	1 412.00~1 180.00	0.828 7	0.097 2	0.115 2	0.848 4	4.283 2	11

续表 4									
质控指标	光谱	筛选方法	v/cm^{-1}	rcal	RMSEC	RMSECV	rpre	RESP/%	LVs
栀子苷	NIRS	全光谱	3 999.64~9 999.10	0.975 4	0.187 4	0.257 1	0.974 6	1.100 9	12
		siPLS	4 300.48~4 599.39, 5 203.00~5 501.92,	0.989 7	0.125 6	0.252 4	0.975 7	1.070 6	15
			5 804.69~6 103.60						
		iPLS	5 804.69~6 103.60	0.948 0	0.272 7	0.356 5	0.933 4	2.068 6	9
		mwPLS	5 866.40~6 466.15	0.970 9	0.203 8	0.294 6	0.944 8	1.683 7	10
	MIRS	全光谱	3 000.00~648.00	0.968 1	0.212 9	0.577 8	0.822 9	3.093 0	19
		siPLS	1 572.00~1 460.00, 1 340.00~1 228.00,	0.966 5	0.218 3	0.324 5	0.888 3	2.320 5	17
			1 108.00~996.00						
		iPLS	1 108.00~996.00	0.873 2	0.414 3	0.5596	0.718 7	3.555 7	16
		mwPLS	1 292.00~1 060.00	0.965 8	0.220 3	0.325 9	0.863 6	2.538 9	20
断氧化马	NIRS	全光谱	3 999.64~9 999.10	0.973 0	0.030 5	0.046 8	0.968 2	2.294 6	14
钱子苷		siPLS	4 300.48~4 599.39, 5 804.69~6 103.60,	0.994 8	0.013 4	0.040 6	0.957 4	2.434 0	16
			7 908.64~8 205.63						
		iPLS	8 207.56~8 504.54	0.965 8	0.034 2	0.050 0	0.953 5	2.884 1	8
		mwPLS	9 144.79~9 744.54	0.969 0	0.032 6	0.054 1	0.943 3	2.992 6	7
	MIRS	全光谱	3 000.00~648.00	0.897 0	0.058 4	0.108 5	0.806 8	5.247 9	6
		siPLS	1 688.00~1 576.00, 1 224.00~1 112.00,	0.930 0	0.048 6	0.069 1	0.892 7	3.8164	11
			1 108.00~996.00						
		iPLS	1 572.00~1 460.00	0.784 9	0.055 3	0.081 8	0.750 8	5.5700	9
		mwPLS	1 300.00~1 068.00	0.847 8	0.070 9	0.089 8	0.855 8	4.859 0	10
固含量	NIRS	全光谱	3 999.64~9 999.10	0.978 2	0.070 4	0.120 6	0.952 9	1.259 1	15
		siPLS	4 300.48~4 599.39, 5 503.84~5 802.76,	0.970 2	0.082 2	0.108 1	0.948 3	1.329 8	13
			7 008.05~7 306.96						
		iPLS	5 804.69~6 103.60	0.941 3	0.114 5	0.148 2	0.867 3	1.942 7	9
		mwPLS	6 057.31~6 657.07	0.970 5	0.081 7	0.128 3	0.939 3	1.362 2	9
	MIRS	全光谱	3 000.00~648.00	0.914 6	0.138 5	0.230 2	0.784 2	2.547 0	6
		siPLS	1 456.00~1 344.00, 1 224.00~1 112.00,	0.903 1	0.145 6	0.1667	0.899 2	1.720 3	8
			1 108.00~996.00						
		iPLS	1 224.00~1 112.00	0.816 1	0.198 8	0.246 9	0.790 0	2.479 5	13
		mwPLS	1 396.00~1 164.00	0.877 6	0.163 3	0.0.188 2	0.852 6	2.128 1	8

5 501.92、5 804.69~6 103.60 cm⁻¹ 作为最优建模波段; 固含量预测模型采用 siPLS 筛选波段后,虽然 *r*cal、*r*pre 略微减小,RSEP 略微增大,但是 RMSEC、 RMSECV 减小且接近,故固含量选择 4 300.48~4 599.39、5 503.84~5 802.76、7 008.05~7 306.96 cm⁻¹ 作为最优建模波段。

基于 MIRS, 隐绿原酸预测模型经 3 种方法筛 选波段后,虽然建模变量数减少,但 r_{cal}、r_{pre}减小, RMSEC、RMSECV 普遍增大,且 RSEP 增大,模 型预测性能下降,故选择全光谱区间(3 000.00~ 648.00 cm⁻¹)建模;绿原酸、栀子苷、断氧化马钱 子苷和固含量的预测模型经采用 siPLS 筛选最优波段后,建模变量数降为 87个,r_{cal}、r_{pre} 普遍增大,RMSEC、RMSECV 较小且接近,RSEP 均增大,模型预测性能提升,故绿原酸、栀子苷、断氧化马钱子苷和固含量分别选择 1 572.00~1 460.00、1 224.00~1 112.00、1 108.00~996.00 cm⁻¹,1 572.00~1 460.00、1 340.00~1 228.00、1 108.00~ 996.00 cm⁻¹,1 688.00~1 576.00、1 224.00~ 1 112.00、1 108.00~996.00 cm⁻¹及 1 456.00~ 1 344.00、1 224.00~1 112.00、1 108.00~996.00 cm⁻¹ 作为最优建模波段;新绿原酸预测模型采用 mwPLS 筛选最优波段后,建模变量数降为59个,RMSEC、 RMSECV较小且接近,且RSEP略微减小,模型预 测性能提升,故新绿原酸选择1401.00~1172.00 cm⁻¹作为最优建模波段。

2.8 PLS 模型的建立

模型经过光谱预处理方法的选择、最佳的潜变 量数的确定及最优建模波段的确定后,运用 PLS 分 别建立了 6 个质控指标的 NIRS 和 MIRS 最佳定量 校正模型,6个质控指标的 NIRS 模型预测效果均 强于 MIRS 模型,故选择6个质控指标的 NIRS 预 测模型作为最优模型,结果见表5。图3为对应的6 种质控指标的最佳模型的预测值与实测值的相关 性。从表5可见,对于 RI制剂过程的投料和二次热 处理工序的6种质控指标绿原酸、新绿原酸、隐绿 原酸、栀子苷、断氧化马钱子苷及固含量,利用 NIRS 建模,其 RSEP 分别为1.473 4%、3.431 8%、2.589 5%、

表 5 6 种质控指标的最优模型性能参数 Table 5 Optimal model performance parameters of six quality control indexes

			-			1 1			
质控指标	光谱	预处理方法	波段筛选	LVs	$r_{\rm cal}$	RMSEC	RMSECV	rpre	RSEP/%
绿原酸	NIRS	SNV	全光谱	12	0.969 9	0.107 2	0.152 5	0.951 9	1.473 4
	MIRS	无预处理	siPLS	6	0.940 6	0.149 6	0.194 3	0.902 4	2.093 7
新绿原酸	NIRS	MA	siPLS	12	0.877 3	0.083 3	0.108 9	0.870 5	3.431 8
	MIRS	MA	mwPLS	9	0.767 6	0.129 9	0.147 9	0.790 3	4.317 9
隐绿原酸	NIRS	MA	siPLS	10	0.900 6	0.075 7	0.099 3	0.952 6	2.589 5
	MIRS	MA	全光谱	8	0.830 8	0.096 9	0.145 0	0.870 2	4.172 6
栀子苷	NIRS	无预处理	siPLS	15	0.989 7	0.125 6	0.252 4	0.975 7	1.070 6
	MIRS	MA	siPLS	17	0.966 5	0.218 3	0.324 5	0.888 3	2.320 5
断氧化马钱子苷	NIRS	SNV	全光谱	14	0.973 0	0.030 5	0.047 1	0.968 2	2.294 6
	MIRS	SNV	siPLS	11	0.930 0	0.048 6	0.069 1	0.892 7	3.816 4
固含量	NIRS	归一化法	siPLS	13	0.970 2	0.082 2	0.108 1	0.948 3	1.329 8
	MIRS	无预处理	siPLS	8	0.903 1	0.145 6	0.1667	0.899 2	1.720 3





1.070 6%、2.294 6%、1.329 8%,均小于4%。而利用 MIRS 建模,其 RSEP 分别为2.093 7%、4.317 9%、4.172 6%、2.320 5%、3.816 4%、1.720 3%,均小于5%。鉴于 NIRS 及 MIRS 模型的 RMSEC 及 RMSECV 也均令人满意,可以认为 NIRS 及 MIRS 模型均满足实际运用的需要,可以快速有效地预测投料和二次热处理工序的6种质控指标的含量预测。

将验证集的光谱数据导入用已建立的最佳模型 中,即 NIRS 预测模型。根据样品的实测值和模型 预测值,计算其相对误差。绿原酸、新绿原酸、隐 绿原酸、栀子苷、断氧化马钱子苷和固含量的模型 预测值与样本实测值的平均绝对偏差(mean absolute, MAD)及平均相对偏差(mean relative bais, MRB)见表 6,结果显示这 6 种质控指标的 MRB 均小于 3.00%。

$$MAD = \sum_{i=1}^{n} (y - y_i)/n \tag{5}$$

$$MRB = MAD / \overline{y}_n \tag{6}$$

n 为验证集样本数, y 为实测值, y_i 为预测值, \overline{y}_n 为样本实测值的平均值

表 6 验证集样本在 NIRS 预测模型中的预测值与实测值的 对比

 Table 6
 Comparison of predicted and measured values in validation set samples

医坎比尔	平均绝对偏差/	亚均相对伯兰/0/
灰1至1日小	$(mg \cdot mL^{-1})$	1均相利佣左/%
绿原酸	0.114 1	1.20
新绿原酸	0.070 8	2.90
隐绿原酸	0.055 3	2.11
栀子苷	0.118 1	0.89
断氧化马钱子苷	0.024 8	1.85
固含量	0.080 2	1.00

3 讨论

本研究首次以 RI 制剂过程为研究对象,采用 NIRS 和 MIRS 技术,结合 PLS 法,经过光谱预处 理方法及建模波段筛选,建立了投料和二次热处理 工序的绿原酸、新绿原酸、隐绿原酸、栀子苷、断 氧化马钱子苷和固含量的定量预测模型。对比 NIRS 和 MIRS 模型对 6 种质控指标的预测效果,可以发 现,NIRS 模型均优于 MIRS 模型,这可能与质控指 标的质量浓度有关系。5 种质控指标成分的平均质 量浓度分别为 9.851 8、2.478 3、2.740 8、13.633 5、 1.302 3 mg/mL,均高于 1 mg/mL,可以认为均为高 质量浓度分析物,平均固含量 7.995 1%亦属于高质 量分数。这与 Tao 等^[22]的研究结果相符合,即对于 高质量浓度的分析物,NIRS 表现出比 MIRS 更好的 预测性能。

根据平均质量浓度对 6 种质控指标进一步分 组,将新绿原酸、隐绿原酸及断氧化马钱子苷分为 高质量浓度组,绿原酸、栀子苷及固含量分为超高 质量浓度组,以预测模型的 RSEP 作为决定评价指 标,可以发现,超高质量浓度组预测模型的 RSEP 均小于较高质量浓度组,可以猜测,NIRS 模型对 高质量浓度样品具有优异的预测效果,并且质量浓 度越高,预测效果越好,后续可以通过增大样本量 来验证。NIRS 模型与 MIRS 模型预测效果存在差 异,也有可能是由于 NIRS 相较于 MIRS,其波数更 多,建模变量数更多,数据中包含的信息更加丰富, 使得其最终模型预测效果更佳,这也需要后续的研 究证明。

本研究表明, NIRS 与 MIRS 模型都适用于这 6 种质控指标的预测, 2 种光谱技术在离线或者在线 检测方面各有所长, NIRS 模型预测精度高、检测 速度快、稳定性好^[23]; MIRS 建模参数少、仪器小 巧便携,而且可以适用于混悬体系样品的检测,减 少了样品的预处理过程,扩大了红外光谱的应用范 围^[24],但同时 MIRS 对环境的要求更高,仪器位置 的改变、温湿度及压力的变化、光纤的弯折程度等 都会对光谱产生影响。与传统化学分析方法相比, 这 2 种光谱分析技术更环保,耗时更少,均可以对 RI 投料和二次热处理工序的质控指标进行快速检 测,也为后续的 RI 制剂过程的全工序快速检测研究 奠定了基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- 孙兰,周军,王振中. 热毒宁注射液药理作用与临床研 究进展 [J]. 中国中医急症, 2014, 23(12): 2247-2249.
- [2] 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.
- [3] 新型冠状病毒肺炎诊疗方案(试行第九版)[J]. 中国医药, 2022, 17(4): 481-487.
- [4] Sinelli N, Cerretani L, Egidio V D, *et al.* Application of near (NIR) infrared and mid (MIR) infrared spectroscopy as a rapid tool to classify extra virgin olive oil on the

basis of fruity attribute intensity [J]. *Food Res Int*, 2010, 43(1): 369-375.

- [5] Uríčková V, Sádecká J. Determination of geographical origin of alcoholic beverages using ultraviolet, visible and infrared spectroscopy: A review [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2015, 148: 131-137.
- [6] Nenadis N, Tsimidou M Z. Perspective of vibrational spectroscopy analytical methods in on-field/official control of olives and virgin olive oil [J]. *Eur J Lipid Sci Technol*, 2017, 119(1): 1600148.
- [7] 耿姝,陈勇,金叶,等.基于近红外技术的金银花药材 多指标成分快速检测 [J].中国现代应用药学,2016, 33(2):144-150.
- [8] 王磊,杨越,潘红烨,等.热毒宁注射液金银花提取过 程轨迹研究 [J].世界科学技术一中医药现代化,2017, 19(8):1277-1282.
- [9] 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.
- [10] 吴莎, 刘启安, 吴亚男, 等. 近红外光谱对热毒宁注射 液栀子萃取过程中的可行性分析 [J]. 中国实验方剂学 杂志, 2015, 21(1): 53-56.
- [11] 王璇, 王永香, 毕宇安, 等. 热毒宁注射液中金银花和
 青蒿干燥工艺过程的批放行标准 [J]. 中成药, 2017, 39(10): 2059-2063.
- [12] 崔小弟,李执栋,徐芳芳,等. 热毒宁注射液多指标成分的一测多评法同时测定 [J]. 时珍国医国药, 2019, 30(5): 1116-1119.
- [13] 金叶,丁海樱,吴永江,等.近红外光谱技术用于血必 净注射液提取过程的在线检测研究 [J].药物分析杂 志,2012,32(7):1214-1221.
- [14] 杜文俊,刘雪松,陶玲艳,等. 热毒宁注射液金银花和 青蒿(金青)醇沉过程中多指标的近红外快速检测 [J].

中草药, 2015, 46(1): 61-66.

- [15] 陈奕云,赵瑞瑛,齐天赐,等.结合光谱变换和 Kennard-Stone 算法的水稻土全氮光谱估算模型校正集 构建策略研究 [J].光谱学与光谱分析,2017,37(7): 2133-2139.
- [16] 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.
- [17] 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.
- [18] Nørgaard L, Hahn M T, Knudsen L B, *et al.* Multivariate near-infrared and Raman spectroscopic quantifications of the crystallinity of lactose in whey permeate powder [J]. *Int Dairy J*, 2004, 15(12): 1261-1270.
- [19] 冼瑞仪,黄富荣,黎远鹏,等.可见和近红外透射光谱 结合区间偏最小二乘法(iPLS)用于橄榄油中掺杂煎炸 老油的定量分析 [J].光谱学与光谱分析, 2016, 36(8): 2462-2467.
- [20] 王彩虹, 黄林, 刘木华, 等. 基于 SiPLS 模型的稻壳中 重金属铬 LIBS 检测 [J]. 激光与光电子学进展, 2016, 53(11): 298-305.
- [21] 史智佳,田寒友,邹昊,等. MWPLS 法在羊肉中 TVB-N近红外定量分析中的应用 [J]. 食品科学, 2015, 36(20): 218-221.
- [22] Tao L Y, Lin Z L, Chen J S, et al. Mid-infrared and near-infrared spectroscopy for rapid detection of *Gardeniae Fructus* by a liquid-liquid extraction process [J]. J Pharm Biomed Anal, 2017, 145: 1-9.
- [23] 汪东升. 在线近红外检测系统的研发及在油菜籽含油 率检测中的应用 [D]. 镇江: 江苏大学, 2020.
- [24] 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.

[责任编辑 郑礼胜]