

防风色原酮类成分的近红外光谱定量分析

薛 雪 熊 慧 刘梦楠 屈文佳 贾天颖 王海丽 辛洁萍 徐新房 李向日

(北京中医药大学中药学院,北京,102488)

摘要 目的:本文建立了能快速测定防风中升麻素、4 种色原酮含量之和的近红外光谱定量模型。方法:用高效液相色谱法测定 55 批防风饮片的 4 种色原酮的含量作为参考值,结合偏最小二乘回归法(PLS)建立回归模型。筛选了不同的光谱预处理方法、波段及最佳因子数,最终模型通过相关系数(R^2)、预测均方根误差(RMSEP)和校正均方差(RMSEC)来评估其稳健性。结果:升麻素的 RMSEP 值、RMSEC 值、 R^2 分别为 0.025 0、0.024 0、0.949 9,4 种色原酮含量总和的最佳 RMSEP 值、RMSEC 值、 R^2 分别为 0.2460、0.1720、0.9110。结论:本研究建立的近红外光谱定量模型较准确,可用于对防风质量的快速检测。

关键词 防风;近红外;质量;色原酮;升麻苷;升麻素;5-O-甲基维斯阿米醇苷;亥茅酚苷

Quantitative Analysis of Chromones in Saposhnikovia Radix by Near Infrared Spectroscopy

Xue Xue, Xiong Hui, Liu Mengnan, Qu Wenjia, Jia Tianying, Wang Haili, Xin Jieping, Xu Xinfang, Li Xiangri

(School of Chinese Pharmacy, Beijing University of Chinese Medicine, Beijing 102488, China)

Abstract Objective: To rapidly establish the detection method for cimifugin and the amounts of four chromones in Saposhnikovia Radix by near infrared spectroscopy. **Methods:** The contents of four chromones of 55 batches of Saposhnikovia Radix was determined by high performance liquid chromatography as reference value. Combined with partial least squares regression(PLS), the regression model was established. Different spectral pretreatment methods, wave ranges and optimal factors were selected. The stability of the models were evaluated by R^2 , RMSEP and RMSEC. **Results:** The optimum RMSEP values, RMSEC values and R^2 values of cimifugin were 0.025 0, 0.024 0 and 0.949 9, respectively. The optimum RMSEP values, RMSEC values and R^2 values of the amounts of four chromones were 0.2460, 0.1720 and 0.9110, respectively. **Conclusion:** The NIR quantitative models established in this study are accurate and can be used for rapid detection of quality of Saposhnikovia Radix.

Key Words Saposhnikovia Radix; Near-infrared; Quality; Chromone; Prim-O-glucosyl-cimifugin; Cimifugin; 4-O- β -D-glucosyl-5-O-methylvisamminol; Sec-O-glucosylhamaudol

中图分类号:R284 文献标识码:A doi:10.3969/j.issn.1673-7202.2019.03.011

防风为伞形科植物防风 *Saposhnikovia Divaricata* (Turcz.) Schischk. 的干燥根^[1],其药性辛、甘,微温;归膀胱、肝、脾经。具有祛风解表,胜湿止痛,止痉的功效。防风始载于《神农本草经》并被列为上品,现主要产于黑龙江、吉林、辽宁、内蒙古、河北等省区,是临床常用大品种,应用广泛。防风化学成分复杂,主要有色原酮类、香豆素类、多糖类、有机酸类、挥发油等成分,现代药理实验表明,色原酮类化合物具有良好的解热、抗炎等活性^[2],升麻苷、升麻素、5-O-甲基维斯阿米醇苷、亥茅酚苷(下文简称“4 种色原酮”)是防风中主要的色原酮类成分,常被作为指标性成分评价防风饮片的质量^[3]。其中升麻素是防风的主要药效成分及入血成分,也是升麻苷和 5-O-甲基维斯阿米醇苷的苷元,二者在血中转化为升麻素

从而发挥药效^[4],因此升麻素对于说明防风饮片的品质具有重要的意义。

《中华人民共和国药典》(2015 年版)中规定采用高效液相色谱法测定升麻苷和 5-O-甲基维斯阿米醇苷的含量之和,该方法需要对样品进行复杂的前处理,分析时间较长,成本较高,易造成污染,无法满足对样品快速测定和在线检测的要求。近红外光谱技术是近年来在中药领域应用广泛的新技术,它是将化学计量学技术、近红外光谱检测技术和基础测量技术相结合的新方法,具有绿色、无损、快速、无需前处理等优势。因此本文利用近红外光谱技术,结合偏最小二乘法,建立防风饮片中升麻素、4 种色原酮总和的定量分析模型,达到快速测定升麻素、4 种色原酮之和的含量的目的,为防风饮片的质量检测

基金项目:国家中药标准化项目(ZYBZH-Y-ZY45)——白术等 14 种中药饮片标准化建设;北京中医药大学在读研究生项目(2018-JYB-XS076)——基于近红外光谱技术的防风饮片等级划分研究

作者简介:薛雪(1994.08—),女,硕士研究生,研究方向:中药质量控制,E-mail:sherry.xue@bucm.edu.cn

通信作者:李向日(1972.11—),女,教授,博士研究生导师,研究方向:中药活性物质基础及质量控制研究,E-mail:lixiangri@sina.com

提供快捷简便的新方法。

1 仪器与试药

1.1 仪器 Waters 2695 高效液相色谱仪; Agilent ZORBAX SB-C₁₈ 色谱柱(5 μm, 4.6 mm × 250 mm); 电子分析天平; FW100 高速万能粉碎机(天津市泰斯特仪器有限公司); B-220 电热恒温水浴锅(上海亚荣生化仪器厂); Antaris 傅立叶变换近红外光谱仪(Thermo Nicolet 公司, 光源: 卤钨灯, 软件 Result 3.0 用于光谱采集, 软件 TQ Analyst 7.2 用于光谱分析), 配备 InGaAs 检测器, 光纤附件。

1.2 试药 防风样品有 49 批分布于黑龙江佳木斯市汤原县、吉林省洮南市军马场、内蒙古阿荣旗、河北安国市娄营村等 4 个省、10 个地区, 另有 6 批样品分布于蒙古国、俄罗斯 2 个国家, 共计 55 批, 均为课题组人员亲自采集, 经北京中医药大学李向日教授鉴定为伞形科植物防风 *Saposhnikovia divaricata* (Turcz.) Schischk. 的干燥根。将以上 55 批饮片样品粉碎, 过六号筛。

甲醇(色谱纯, Fisher 公司), 娃哈哈纯净水, 甲醇等其他试剂均为分析纯。对照品升麻苷(批号 0083-0025), 升麻素(批号 0084-0020)、5-O-甲基维斯阿米醇苷(批号 0081-0025)、亥茅酚苷(批号 0969-0025), 购自成都普思生物科技股份有限公司。

2 方法

2.1 NIR 光谱采集 光谱采集范围为 10 000 ~ 4 000 cm⁻¹, 分辨率 8 cm⁻¹, 扫描次数 64 次, 波长间隔为 2 cm⁻¹。以仪器内置背景为参比, 对 55 批防风饮片的粉末进行光谱采集, 为减小实验误差, 每份样品重复测定两次, 取平均光谱。55 批防风饮片样品的近红外光谱叠加图见图 1。

2.2 防风中 4 种色原酮含量的 HPLC 测定

2.2.1 色谱条件 采用 Agilent ZORBAX SB-C₁₈ 色谱柱(5 μm, 4.6 mm × 250 mm), 检测波长 254 nm, 流速 1.0 mL/min, 柱温 35 °C, 进样量 10 μL, 流动相为甲醇(A)-水(B)梯度洗脱(0 ~ 5 min, 15 ~ 35% A; 5 ~ 15 min, 35% A; 15 ~ 20 min, 35 ~ 45% A; 20 ~ 29 min, 45% A; 29 ~ 35 min, 45 ~ 55% A; 35 ~ 40 min, 55 ~ 63% A)。

2.2.2 供试品溶液的制备 取细粉约 0.25 g, 精密称定, 置具塞锥形瓶中, 精密加入甲醇 10 mL, 称定重量, 水浴回流 2 h, 放冷, 再称定重量, 用甲醇补足减失的重量, 摇匀, 滤过, 取续滤液, 即得。

2.2.3 对照品溶液的制备 取升麻苷、升麻素、5-O-甲基维斯阿米醇苷、亥茅酚苷对照品各适量, 加甲醇分别配制成浓度为 0.201 0 mg/mL、0.078 0 mg/mL、0.277 2 mg/mL、0.033 0 mg/mL 的对照品溶液, 即得。

2.2.4 方法学考察

2.2.4.1 标准曲线 精密称取各对照品溶液 0.1、0.5、1.0、2.0、5.0 mL, 分别用甲醇稀释于 10 mL 容量瓶中, 即得各系列浓度对照品溶液。不同浓度的 4 种对照品溶液分别进样 10 μL, 以浓度(mg/mL)为横坐标, 峰面积为纵坐标绘制标准曲线, 并计算回归方程。见表 1。

2.2.4.2 精密度试验 取升麻苷、升麻素、5-O-甲基维斯阿米醇苷、亥茅酚苷对照品各适量, 加甲醇分别配制成浓度为 0.016 1 mg/mL、0.006 2 mg/mL、0.015 8 mg/mL、0.013 2 mg/mL 的对照品溶液。精密吸取对照品溶液, 重复进样 6 次, 各对照品峰面积 RSD 均 < 1%, 符合要求, 表明仪器精密度良好。

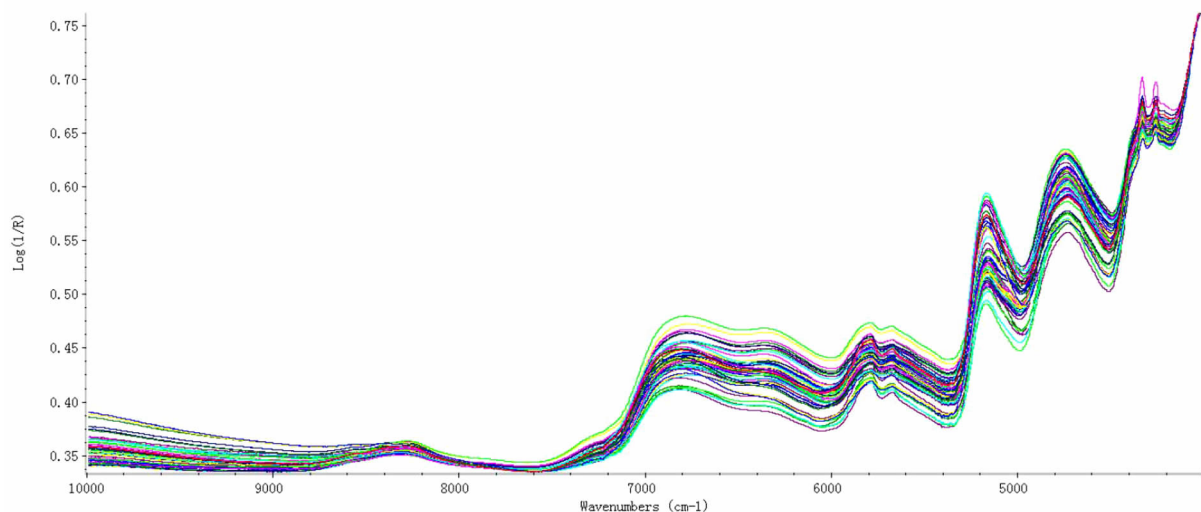


图 1 55 批防风饮片样品的近红外光谱叠加图

表 1 4 种色原酮化合物的线性方程、相关系数及线性范围

成分	线性方程	R ²	线性范围/(mg/mL)
升麻苷	y = 18 148 844. 966 8x - 4 716. 939 8	0. 999 7	0. 003 2 ~ 0. 201 0
升麻素	y = 30 768 557. 514 9x - 16 059. 936 4	0. 999 6	0. 001 2 ~ 0. 078 0
5-O-甲基维斯阿米醇苷	y = 19 683 336. 806 1x + 11 359. 004 3	0. 999 9	0. 003 2 ~ 0. 277 2
亥茅酚苷	y = 27 781 474. 174 1x - 8 778. 145 2	0. 999 9	0. 000 5 ~ 0. 033 0

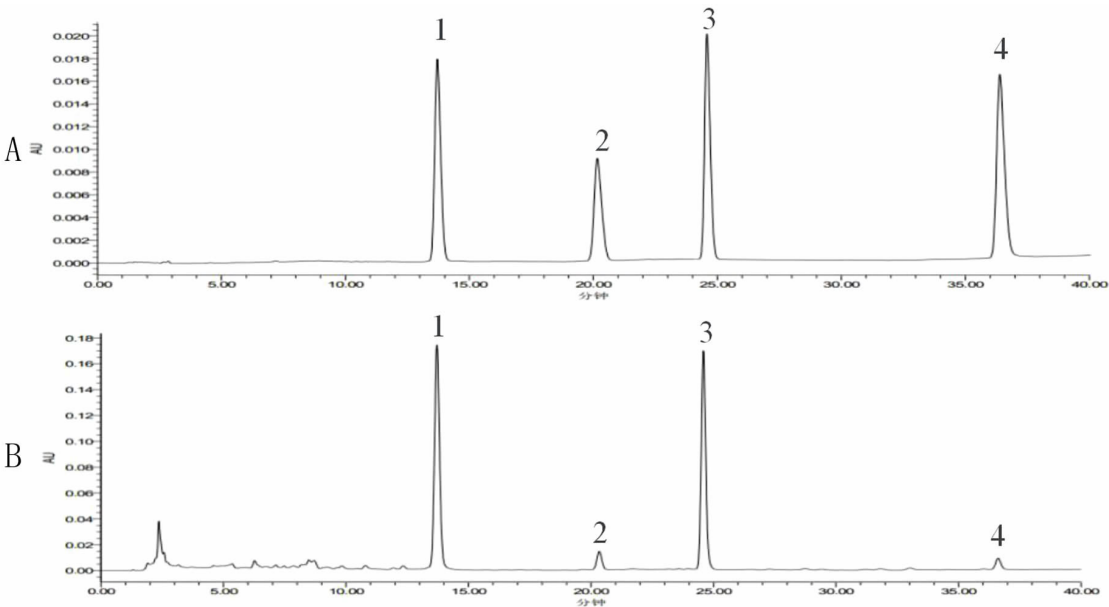


图 2 对照品(A)和防风样品(B)的 HPLC 图

注:1:升麻苷;2:升麻素;3:5-O-甲基维斯阿米醇苷;4:亥茅酚苷

2. 2. 4. 3 稳定性试验 取同一份供试品溶液,分别在 0、2、4、8、12、24 h 进样测定,所测 4 种成分的色谱峰的峰面积 RSD 均小于 2%,表明样品在 24 h 内稳定。

2. 2. 4. 4 重复性试验 取同一批样品细粉 6 份,按 2. 2. 2 项下方法制备供试品溶液,进样测定,所测 4 种成分的色谱峰的峰面积 RSD 均小于 3%,表明方法重复性良好。

2. 2. 4. 5 加样回收率 精密称定已知含量的防风样品 0. 125 g,精密加入各对照品适量,按供试品溶液制备及测定法操作,进行色谱分析。结果所测 4 种成分的平均回收率 (n = 6) 分别为 95. 22%、96. 78%、95. 55%、98. 53%,RSD 都在 2% 以内。

2. 2. 5 样品的测定 精密吸取供试品溶液 10 μL 注入液相色谱仪,记录 40 min 色谱图。见图 2。

2. 3 光谱预处理 光谱采集的过程会受多种因素的影响,除样品信息外,还有噪声、基线漂移和无用信息。在建立近红外定量模型之前,需要对近红外光谱进行预处理,同时提取光谱特征。常用的预处理方法有多元散射校正法 (MSC)、标准归一化法 (SNV)、一阶导数法 (First Derivative)、二阶导数法

(Second Derivative), SG 滤波平滑 (Savitzky-Golay Filter Smoothing), Norris 导数滤波平滑 (Norris Derivative Filter Smoothing) 等。

2. 4 模型的建立 本研究采用偏最小二乘回归法 (PLS) 建立定量模型。以校正均方差 (RMSEC)、交叉验证均方差 (RMSECV)、预测均方差 (RMSEP) 和 R² 等来评价最终 PLSR 模型的性能。

3 结果

3. 1 防风中 4 种色原酮成分的含量测定 见图 1。

3. 2 近红外定量模型的建立

3. 2. 1 校正集和验证集样品的选择 为了确保回归模型的稳健和精度,以 3:1 的比例将 55 批样品划分为校正集和验证集,包括校正集 41 批,验证集 14 批。用校正集建立近红外模型后,再将验证集对模型进行验证。校正集与验证集选择时,需要保证验证集样品的含量范围在校正集样品的含量范围之内。样品划分结果及含量的分布范围。见表 2。

表 2 样品划分批次及 4 种色原酮含量分布范围 (mg/g)

集合	样品总批次	升麻素	4 种色原酮含量之和
校正集	41	0 ~ 3. 4	4. 0 ~ 20
验证集	14	0 ~ 2. 7	7. 5 ~ 17

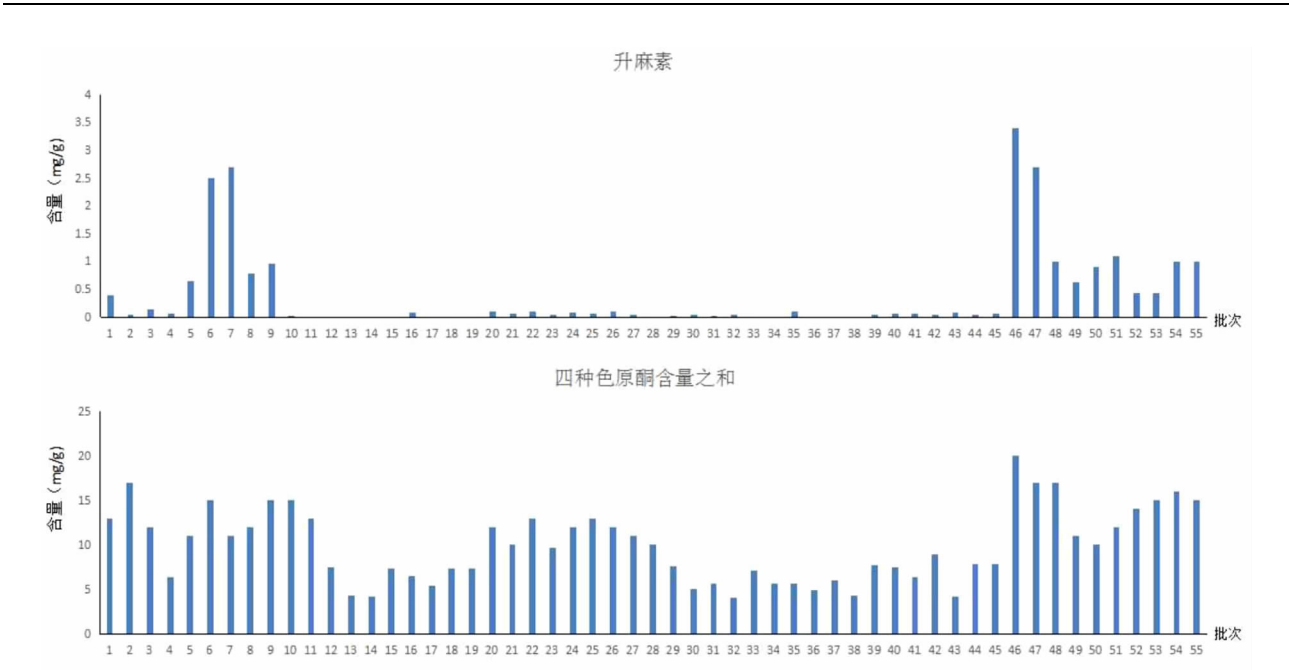


图3 防风4种色原酮成分含量测定结果 (mg/g, $n=2$)

表3 不同预处理方法对模型的影响

	预处理方法	RMSEP	RMSEC	R ²
升麻素	MSC + 不求导 + SG 平滑	0.044 1	0.021 1	0.961 4
	MSC + 一阶导数 + SG 平滑	0.034 6	0.027 6	0.932 8
	MSC + 二阶导数 + SG 平滑	0.037 6	0.023 6	0.951 5
	MSC + 一阶导数 + 不平滑	0.035 9	0.026 9	0.936 6
	MSC + 二阶导数 + 不平滑	0.058 0	0.029 5	0.922 8
	MSC + 一阶导数 + Norris 平滑	0.031 2	0.025 1	0.945 0
	MSC + 二阶导数 + Norris 平滑	0.033 0	0.030 3	0.918 4
	SNV + 不求导 + SG 平滑	0.044 0	0.019 7	0.966 5
	SNV + 一阶导数 + SG 平滑	0.034 5	0.027 6	0.933 1
	SNV + 二阶导数 + SG 平滑	0.037 6	0.023 6	0.951 3
	SNV + 一阶导数 + 不平滑	0.035 8	0.026 8	0.936 9
	SNV + 二阶导数 + 不平滑	0.063 4	0.014 7	0.981 4
	SNV + 一阶导数 + Norris 平滑	0.030 9	0.025 7	0.942 2
	SNV + 二阶导数 + Norris 平滑	0.032 9	0.030 4	0.917 9
4 种色原酮含量之和	MSC + 不求导 + SG 平滑	0.265 0	0.260 0	0.780 8
	MSC + 一阶导数 + SG 平滑	0.264 0	0.278 0	0.744 3
	MSC + 二阶导数 + SG 平滑	0.248 0	0.173 0	0.909 4
	MSC + 一阶导数 + 不平滑	0.263 0	0.278 0	0.744 9
	MSC + 二阶导数 + 不平滑	0.305 0	0.126 0	0.953 3
	MSC + 一阶导数 + Norris 平滑	0.253 0	0.281 0	0.737 8
	MSC + 二阶导数 + Norris 平滑	0.250 0	0.262 0	0.777 7
	SNV + 不求导 + SG 平滑	0.264 0	0.260 0	0.781 7
	SNV + 一阶导数 + SG 平滑	0.262 0	0.277 0	0.748 0
	SNV + 二阶导数 + SG 平滑	0.246 0	0.172 0	0.911 0
	SNV + 一阶导数 + 不平滑	0.262 0	0.276 0	0.748 6
	SNV + 二阶导数 + 不平滑	0.304 0	0.124 0	0.954 5
	SNV + 一阶导数 + Norris 平滑	0.278 0	0.255 0	0.791 7
	SNV + 二阶导数 + Norris 平滑	0.250 0	0.262 0	0.778 8

3.2.2 光谱预处理方法的选择 使用不同光谱预处理方法后,比较校正集内部的交叉验证相关系数R²、校正均方差(RMSEC)及验证集预测均方差(RMSEP)。由表3可以看出,采用PLS法建立近红外定量分析模型,升麻素的最佳预处理方法为SNV + 一阶导数 + Norris 平滑,定量分析模型参数为R²

为 0.942 2, RMSEP 为 0.030 9, RMSEC 为 0.025 7; 4 种色原酮含量之和的最佳预处理方法为 SNV + 二阶导数 + SG 平滑, 定量分析模型参数为 R^2 为 0.911 0, RMSEP 为 0.246 0, RMSEC 为 0.172 0。此时, RMSEP 和 RMSEC 均较小且接近^[5], $R^2 > 0.9$, 模型的预测结果较准确。

3.2.3 波段的优选 在建立模型时, 为了剔除近红外光谱中的谱带冗余信息, 调整模型的准确性, 需要对建模波段进行优选。以 RMSEP、RMSEC 和 R^2 为指标, 对不同波段进行优化比较。见表 4。最终确定, 升麻素的最佳建模波段为 5 727.55 ~ 7 251.03 cm^{-1} , 4 种色原酮含量之和的最佳建模波段为 4 000 ~ 10 000 cm^{-1} 。

表 4 不同波段对模型的影响

	波段(cm^{-1})	RMSEP	RMSEC	R^2
升麻素	4 000.00-10 000.00	0.0309	0.0257	0.9422
	4 000.00-7 500.00	0.0240	0.0044	0.9984
	4 069.06-6 371.65	0.0272	0.0264	0.9387
	5 727.55-7 251.03	0.0250	0.0240	0.9499
4 种色原酮 含量之和	4 000.00-10 000.00	0.2460	0.1720	0.9110
	6 645.50-7 443.88	0.2770	0.2170	0.8540
	4 069.06-6 371.65	0.2100	0.2040	0.8720
	4 350.62-5 947.39	0.2460	0.2590	0.7841

3.2.4 最佳因子数的选择 采用 PLS 建立定量分析模型, 选择的主成分对模型的预测能力有很大的影响。主成分数过少, 信息缺失导致“欠拟合”现象; 主成分数过多, 会出现“过拟合”现象而影响结果^[6]。最佳因子数由交叉验证均方差(RMSECV)决定。由图 4 可知, 升麻素、4 种色原酮含量之和的最佳因子数分别为 8 和 11, RMSECV 分别为 0.047 6 和 0.280 1。

3.2.5 近红外定量模型的建立 经过光谱预处理、波段的优选、最佳因子数筛选, 建立的最佳近红外定量模型的参数如表 5 所示, 2 种模型的真实值与 NIR 预测值之间的相关图。见图 5。可以观察到 2 种模型的相关系数都在 0.9 以上, 显示出近红外模型计算值和真实值之间良好的相关性。

表 5 最佳近红外定量模型的相关参数

指标	RMSEP	RMSEC	R^2
升麻素	0.025 0	0.024 0	0.949 9
4 种色原酮含量之和	0.246 0	0.172 0	0.911 0

4 讨论

升麻苷、升麻素、5-O-甲基维斯阿米醇苷、亥茅酚苷常被作为指标性成分评价防风饮片的质量, 4

种色原酮含量之和能间接地反映防风中色原酮类成分的含量高低, 同时能更多反应药材的成分信息。此外, 升麻素作为防风的主要药效成分、入血成分以及药典 2 种色原酮的苷元, 与防风饮片疗效发挥具有很强的相关性, 也是区分防风的野生与家种的重要指标。因此本研究建立的 2 个近红外定量模型对于快速评价防风饮片质量提供了依据。

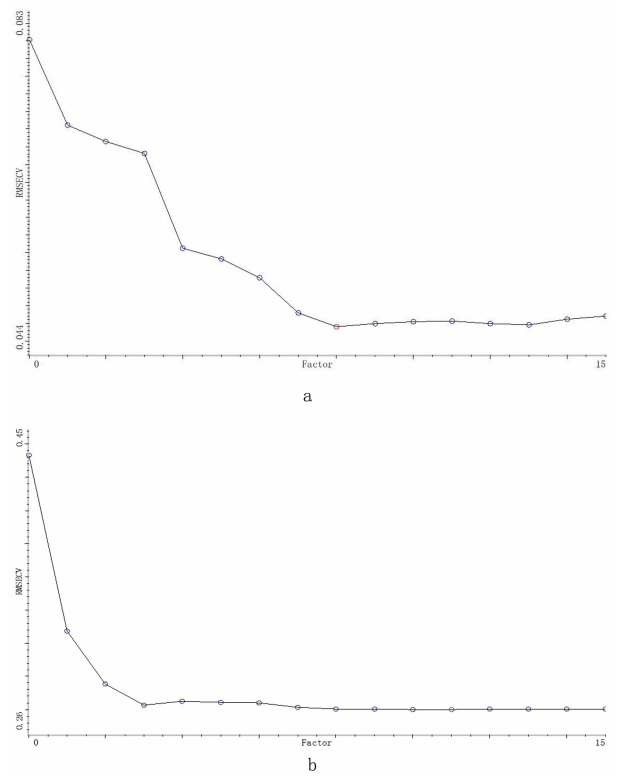


图 4 最佳因子数图

注: a. 升麻素; b. 4 种色原酮含量之和

近红外主要由 C-H、O-H、N-H 等含氧基团的倍频和合频吸收谱组成, 本文测定的 4 种色原酮只由 C、H、O 组成, 在近红外谱区均有吸收, 虽然 2 种测定的指标都是由色原酮类化合物组成, 但由于结构不同, 2 种模型的具体参数如波段、预处理方法、最佳因子数有所差异。根据 TQ Analyst 7.2 光谱分析软件中的算法程序筛选光谱范围, 得到 4 069.06 ~ 6 371.65 cm^{-1} 、6 645.50 ~ 7 443.88 cm^{-1} 、4 350.62 ~ 5 947.39 cm^{-1} 等光谱区间。优化以上波段, 在 RMSEP、RMSEC 较低且接近, R^2 趋近于 1 时, 模型的性能最优, 由此确定了 2 个模型的最佳波段。本研究表明近红外光谱法分别在 5 727.55 ~ 7 251.03 cm^{-1} 、4 000 ~ 10 000 cm^{-1} 波段范围内, 以偏最小二乘算法, 分别采用 SNV + 一阶导数 + Norris 平滑、SNV + 二阶导数 + SG 平滑的预处理方法可同时实现对防风中升麻素、4 种色原酮之和的含量的快速测定。

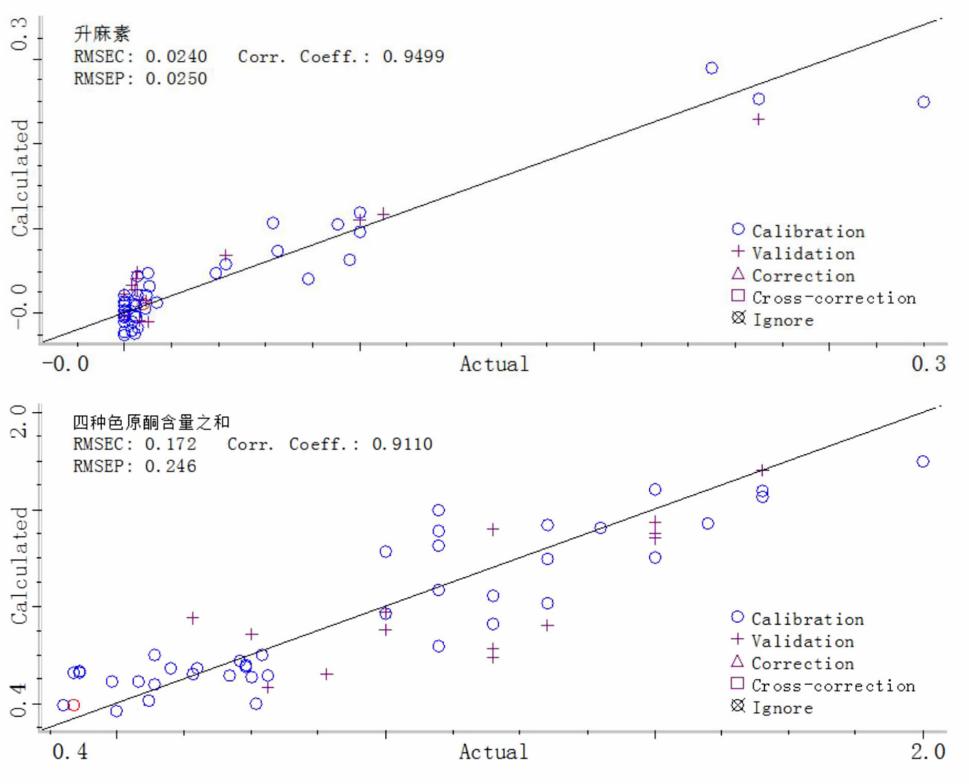


图5 近红外定量模型真实值和预测值的相关图

2 种模型的相关系数 R^2 均大于 0.9, RMSEP、RMSEC 都较小且接近,表明模型准确、强壮,可以准确地预测升麻素、4 种色原酮之和的含量。相比高效液相色谱法、薄层色谱法等传统方法,近红外光谱法拥有无损、快速、环保的优点,可以实现多种目的的快速检测,对于种植、生产、使用单位快速了解防风质量具有重要意义。

参考文献

[1] 国家药典委员会. 中华人民共和国药典[S]. 1 部. 北京: 中国医药科技出版社, 2015: 149.

[2] 薛宝云, 李文, 李丽, 等. 防风色原酮甙类成分的药理活性研究[J]. 中国中药杂志, 2000, 25(5): 297-299.

[3] 刘双利, 张春红, 张连学. 不同产地防风中 4 种色原酮成分比较研究[J]. 中草药, 2007, 38(5): 776-778.

[4] 戴锦娜. 防风药效物质基础和相关成分药代动力学研究[D]. 沈阳: 沈阳药科大学, 2009.

[5] 袁而文, 严新龙, 赵东辉. 烟梗原料 4 种常规化学组分的近红外快速分析[J]. 红外技术, 2015, 37(1): 82-86.

[6] 杨阳, 代涛, 汪学楷. 胆酸含量的近红外分析数学模型[J]. 化学研究与应用, 2011, 23(2): 204-207.

(2018-10-04 收稿 责任编辑: 徐颖)