

NIRS 法快速测定复方阿胶浆中总黄酮、总皂苷和可溶性固形物

韩海帆¹, 张路², 张淹², 李文龙^{1*}, 瞿海斌¹

1. 浙江大学 药物信息学研究所, 浙江 杭州 310058

2. 国家胶类中药工程技术研究中心, 山东 聊城 252201

摘要: 目的 利用近红外光谱(NIRS)快速测定复方阿胶浆中总黄酮、总皂苷和可溶性固形物。方法 采用比色法测定复方阿胶浆中总黄酮和总皂苷的量, 采用水分测定仪测定可溶性固形物的量, 作为参考值。运用偏最小二乘回归(PLSR)法建立 NIRS 与各项指标的量参考值之间的多元校正模型。结果 校正模型的预测相关系数(R^2)和预测误差均方根(RMSEP)分别为总黄酮 0.979 8, 69.6 $\mu\text{g/mL}$; 总皂苷 0.966 0, 17.6 $\mu\text{g/mL}$; 可溶性固形物 0.997 8, 0.394%。结论 本法简捷快速, 结果准确, 可用于复方阿胶浆成品的快速检测和质量控制。

关键词: 近红外光谱; 总黄酮; 总皂苷; 可溶性固形物; 复方阿胶浆

中图分类号: R286.02 文献标志码: A 文章编号: 0253-2670(2013)17-2397-07

DOI: 10.7501/j.issn.0253-2670.2013.17.011

Rapid determination of total flavonoids, total saponins, and soluble solid content in Fufang Ejiao Syrup by NIRS

HAN Hai-fan¹, ZHANG Lu², ZHANG Yan², LI Wen-long¹, QU Hai-bin¹

1. Pharmaceutical Informatics Institute, Zhejiang University, Hangzhou 310058, China

2. National Engineering Technology Research Center of Rubber Medicine, Liaocheng 252201, China

Abstract: Objective To develop an expeditious method for the simultaneous determination of main active compounds in Fufang Ejiao Syrup (FES, compound Ejiao oral liquid) by near infrared spectroscopy (NIRS) in transreflective mode. **Methods** The contents of total flavonoids and total saponins were determined by colorimetric assays, while the soluble solid content was determined using a moisture analyzer as reference value. Partial least squares regression (PLSR) was employed for the calibration of regression models between NIRS and the indexes. **Results** The correlation coefficients (R^2) of the constructed models for total flavonoids, total saponins, and soluble solid content were 0.979 8, 0.966 0, and 0.997 8, respectively, and the root mean square errors of prediction (RMSEP) were 69.6, 17.6 $\mu\text{g/mL}$, and 0.394%, respectively. **Conclusion** The proposed method is fast and accurate, and facilitating the rapid analysis of main active components in FES.

Key words: near-infrared spectroscopy; total flavonoids; total saponins; soluble solid content; Fufang Ejiao Syrup

成品检验是药品在进入市场前必经的最后一道质量控制程序,直接关系到患者的用药安全。目前,中药成品的检测方法多为色谱法,这类方法常需较为复杂的样品预处理过程,且1次分析只能测定1个或少数几个指标成分的量,分析时间较长,增加了批量生产过程的总耗时。近红外光谱(near-infrared spectroscopy, NIRS)用于中药质量分析能从整体上反映其化学组成信息,具有样品无需或仅需极少的预处理、操作简便、不消耗化学试剂以及

可实现在线过程控制等优势^[1],近年来已被广泛应用于中药质量控制领域,在定性^[2]和定量^[3-4]分析中都显示了巨大的潜力。

复方阿胶浆(Fufang Ejiao Syrup, FES)是山东东阿阿胶股份有限公司独家生产的中药保护品种,是根据明代医家张介宾《景岳全书》中的两仪膏(熟地黄、人参),加阿胶、党参等中药制成,主要用于气血两虚引起的头晕目眩、心悸失眠、食欲不振、贫血、白细胞减少症及放化疗的增效减毒^[5]。

收稿日期: 2013-04-11

基金项目: 国家“重大新药创制”科技重大专项(2011ZX09201-201-10); 浙江省公益技术研究项目(2012C21102)

作者简介: 韩海帆(1988—),男,硕士研究生,主要从事中药制药过程分析技术研究。E-mail: hhf_0202@126.com

*通信作者 李文龙 Tel: (0571)88208426 E-mail: wshlwl@zju.edu.cn

《中国药典》2010 年版中对于复方阿胶浆的定量测定仅有总氮量一项^[6], 不足以反映复方阿胶浆成品中有效成分的整体状况, 文献中也未见关于复方阿胶浆质量控制方法的研究报道。因此, 本实验以复方阿胶浆为研究对象, 运用 NIRS 同时测定复方阿胶浆成品中的总黄酮、总皂苷和可溶性固形物的量, 为复方阿胶浆成品提供了一种快速定量分析方法, 同时也为在线质量控制的实现奠定了基础。

1 仪器与材料

Antaris II 型傅里叶变换近红外光谱仪 (美国 Thermo Fisher 公司), 配有 SabIR 透反射配件和 Result 3.0 光谱采集软件; HG63 水分测定仪 (美国 Mettler Toledo 公司); TU-1900 型紫外可见分光光度计 (北京普析通用仪器公司)。

对照品人参皂苷 Re (批号 120220, 质量分数 99.31%) 和芦丁 (批号 100080-200707, 质量分数 92.5%) 均购自中国药品生物制品检定所。60 批复方阿胶浆成品由山东东阿阿胶股份有限公司提供, 包括 2012 年 4 月~5 月的成品样品, 批号分别为 1204054~1205408。

2 方法与结果

2.1 样品制备

用于建立 NIRS 定量分析模型的样本需有较强的代表性。复方阿胶浆成品的各项定量指标变化范围较窄, 若直接用于建立 NIRS 定量分析模型, 则会导致模型稳健性较差, 适用性不强。因此, 本研究通过实验室配制不同质量浓度的复方阿胶浆样品, 与成品样本一同建模, 从而增加了建模样本的代表性。

本实验将 60 批复方阿胶浆成品样品等分为 2 份。其中 30 批成品直接进行定量测定; 剩余 30 批成品随机分组合并, 得到 6 份样品。将这 6 份样品在 70 °C 下减压浓缩至体积约为原体积的一半, 再用超纯水进行逐级稀释, 并在稀释过程中采集 NIRS 并取样, 共得到样品 58 份。复方阿胶浆成品与浓缩自配样品总共 88 份。

2.2 样品定量测定

以人参皂苷 Re 作为对照, 采用香草醛-高氯酸比色法^[7]在 544 nm 处测定样品中总皂苷质量浓度; 以芦丁作为对照, 采用亚硝酸钠-硝酸铝比色法^[8]在 500 nm 处测定样品中总黄酮质量浓度。样品分析前, 分别建立回归曲线, 并进行方法学考察, 确认可用于复方阿胶浆样品的定量测定。比色法的各项

方法学考察结果: 总皂苷 $Y=24.809 X-0.018 1$, $R=0.999 7$, 重复性 RSD 为 2.44%, 1 h 内室温稳定性 RSD 为 2.14%; 总黄酮 $Y=13.381 X-0.008 6$, $R=0.999 7$, 重复性 RSD 为 1.12%, 1 h 内室温稳定性 RSD 为 1.55%。

样品中可溶性固形物的量, 使用水分测定仪测定^[9]。仪器加热程序设定为升温至 140 °C 后降至 90 °C, 直至样品质量不再改变, 可溶性固形物的量以质量分数表示。

2.3 光谱采集

采用 SabIR 透反射模式采集复方阿胶浆的 NIRS 图。光谱采集条件为以仪器内置背景为参比, 光谱采集波数范围: 4 000~10 000 cm^{-1} , 分辨率 4.0 cm^{-1} , 扫描次数 128 次。所采集的复方阿胶浆成品近红外光谱见图 1。

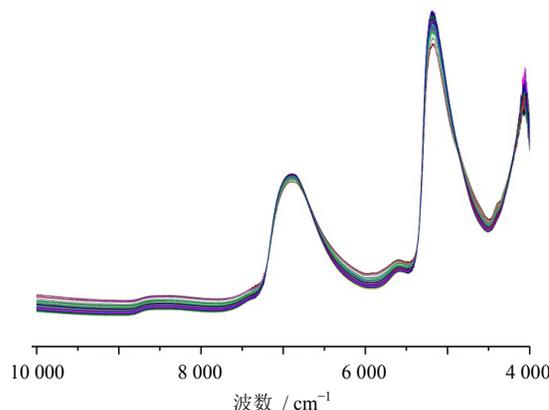


图 1 复方阿胶浆的 NIRS

Fig. 1 NIRS of FES

2.4 数据预处理和校正模型建立

采用偏最小二乘回归 (PLSR) 法建立 NIRS 与总皂苷、总黄酮和可溶性固形物的定量校正模型。首先采用 Chauvenet 检验结合学生化残差和杠杆值图鉴别并剔除异常样本 (Outliers)^[10], 采用 SPXY 法将样本集划分为校正集和验证集。之后, 选择合适的建模波段和光谱预处理方法, 运用留一法交叉验证 (Leave-One-Out Cross Validation, LOOCV), 以交叉验证误差均方根 (RMSECV) 为指标确定最佳因子数。采用校正集误差均方根 (RMSEC) 及其相关系数 R_c 、交叉验证预测误差均方根及其相关系数 R_{cv} 和验证集预测误差均方根 (RMSEP) 及其相关系数 R_p 评价模型的拟合能力及对未知样本的预测能力。数据处理由 TQ Analyst 光谱数据处理软件 (Version 8.5, 美国 Thermo Fisher 公司) 和 Matlab

(Version 7.13, 美国 Math Works 公司) 软件完成。

2.5 定量测定结果

样本总黄酮、总皂苷和可溶性固形物的化学值测定结果：总皂苷 0.499 5~0.095 6 mg/mL，均值 0.231 5 mg/mL，RSD 为 35.64%；总黄酮 3.237 8~0.388 6 mg/mL，均值 1.284 0 mg/mL，RSD 为 44.53%；可溶性固形物 0.397 7~0.079 0 mg/mL，均值 0.196 8 mg/mL，RSD 为 37.09%。样本的覆盖质量浓度范围较宽，具有较强的代表性，适于建立定量分析模型。

2.6 异常样本剔除

用 NIRS 数据建立光谱预测模型时，异常样本的存在将严重影响模型的预测准确性，必须将其从样本集中剔除。异常样本通常包括光谱异常或化学值异常 2 类。本实验采用 Chauvenet 检验法和杠杆值与学生化残差值相结合的方法进行异常样本的剔除，兼顾了化学值和光谱数据的异常，有助于提高模型的预测效果。

Chauvenet 检验法首先计算所有样品光谱的平均光谱，然后计算每个样品光谱与平均光谱之间的马氏距离，将距离值从小到大的顺序排列，根据 Chauvenet 判别准则判定距离值最大的样品光谱是否为异常，若是则继续判别距离值第二大的样品光谱是否为异常，以此类推，直至某一样品光谱被判定为正常。Chauvenet 检验结果见图 2，编号为 24 和 73 的样品与样品集的平均光谱有显著性差异，说明这 2 个样品的光谱数据与组内其余光谱差异较大，因此将其作为异常样品剔除。

杠杆值和学生化残差值是异常样本鉴别中的 2 项重要指标：前者反映了样品对模型的重要程度，

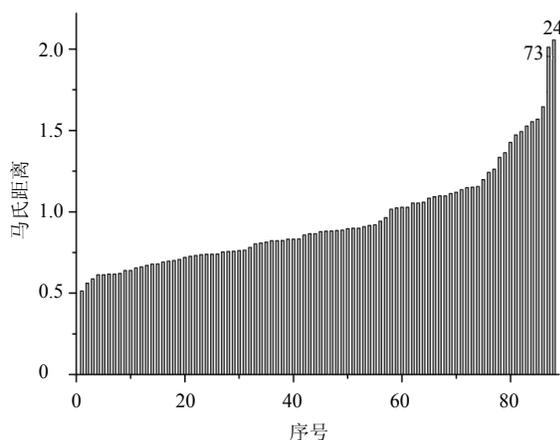


图 2 全部样本的马氏距离值

Fig. 2 Mahalanobis distance values of all samples

后者则反映了样品指标含量值的预测偏差。当一个样品的杠杆值或学生化残差值比较高时，将其暂列为异常样品。以总黄酮定量模型为例，图 3 为 86 份样品的杠杆值与学生化残差分布图，从图中可以得出：编号为 6、31、40、50 和 79 的样品的杠杆值较大，编号为 39 和 60 的样品的学生化残差值较大，因此，将这些样品暂列为异常样品。

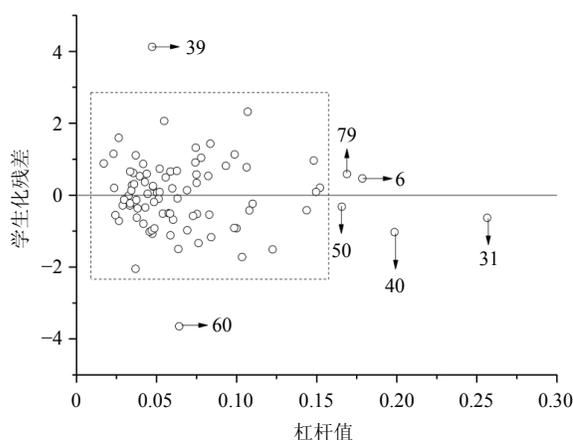


图 3 学生化残差与杠杆值的散点图

Fig. 3 Scatterplot of studentized residuals vs leverage value

如果直接剔除上述异常样品，则有可能将非异常样品误当作异常样品剔除。为避免发生这样的错误，需对被判定为异常的样品进行逐一回收，根据回收后的模型性能决定样品的去留。此法不仅可以有效避免误判，同时也保留了较多的样品，保证了模型的代表性。逐一回收上述暂列异常的 7 个样品，比较未剔除、全部剔除和逐个回收多种情况下的模型结果，从中选出最优的模型。由于尚未进行样本集的划分，所有的样品均用作校正集样品，采用 PLSR 对原始光谱进行建模。分析结果见表 1。由表可知，回收样品 31、39、40、50 和 60 使模型性能不同程度下降，因而将这些样品定为异常样品并将其从样品集中剔除。回收样品 6 和 79 后模型性能略有改善，因此将这些样品重新归入样品集。

综合以上分析，对于总黄酮定量模型，共剔除 7 个样品（编号为 24、31、39、40、50、60 和 73），剩余样品用于下一步建模。

2.7 样本集划分

如何挑选具有代表性的样本建立模型，即校正集样本的代表性问题，是近红外分析技术的关键问题之一。有代表性的校正集样品不但可以减少建模的工作量，而且直接影响所建模型的适用性和准确性。本实验在运用 PLSR 建模过程中，比较了采用随

表 1 总黄酮模型回收剔除样本后的建模结果

Table 1 Performance parameters of total flavonoid models with rejecting samples reinstated

剔除的样品	因子数	R_c	RMSEC / ($\mu\text{g}\cdot\text{mL}^{-1}$)	R_{cv}	RMSECV / ($\mu\text{g}\cdot\text{mL}^{-1}$)
未剔除	5	0.989 3	84.5	0.981 7	110.0
7 个全部剔除	5	0.988 0	67.7	0.979 9	90.6
剔除 6 个, 回收 6	5	0.988 9	67.1	0.979 5	90.8
剔除 6 个, 回收 31	5	0.990 7	68.8	0.983 1	92.3
剔除 6 个, 回收 40	5	0.988 9	69.3	0.980 7	91.2
剔除 6 个, 回收 50	4	0.987 5	77.8	0.982 1	93.1
剔除 6 个, 回收 79	5	0.990 0	68.4	0.982 5	90.5
剔除 6 个, 回收 39	4	0.982 1	86.6	0.974 7	103.0
剔除 6 个, 回收 60	5	0.986 3	77.1	0.976 1	102.0
剔除 5 个, 回收 6 和 79	5	0.990 1	67.8	0.982 3	90.7

机 (random sampling, RS) 法、Kennard-Stone (KS) 法、浓度梯度法、Sample set Partitioning based on joint X-Y distance (SPXY) 法以及 Duplex 法进行校正集和验证集的划分对所建立模型性能的影响。

随机法^[11]即随机选取一定数量的样本组成校正集。组成方法简单, 无需数据挑选, 但每次组成校正集的样本可能差异很大, 不能保证所选样本具有代表性以及模型的外推能力。浓度梯度法^[12]是一种较常用的选择方法, 是将样品按组分化学值的顺序排列, 然后按序抽取样品组成校正集和验证集。此法简单直观, 但是由于仅考虑了化学值的差异, 难以保证校正集样品的代表性。KS 法^[13]是在候选样本中迭代挑选最短欧式距离最大的样本向量对进入校正集, 直至达到预设样本数。Duplex 法^[14]与 KS 法类似, 同样是通过光谱差异来挑选校正集样本, 两法皆是将光谱差异较大的样品选入校正集, 虽然在一定程度上保证了校正集样品分布的均匀, 但却都未考虑变量 Y 的信息。SPXY 法是由 Galvão 等^[11]首先提出的, 其逐步选择的过程和 KS 法相似, 区别是在计算样品间距离时同时考虑了变量 X (光谱数据) 和变量 Y (化学值), 并通过将样本间距离

分别除以其在数据集中的最大距离以保证样本在 X 空间和 Y 空间具有相等的权重。SPXY 法优点在于能够有效地覆盖多维向量空间, 从而改善所建模型的预测能力。

分别采用随机法、浓度梯度法、KS 法、Duplex 法和 SPXY 法选取校正集的样本后, 以 PLSR 建立各指标的量 and NIRS 之间的定量校正模型, 然后分别对 4 种方法所获得的模型结果进行评价。以总黄酮定量模型为例, 为了便于比较, 各种划分方法的校正集样本数都定为 60, 建模结果见表 2。可见, 随机法和浓度梯度法选取的校正集样本所建立的模型预测性能较差, RMSEP 值较高。Duplex 法所建模型的校正性能良好, 但预测误差较大, 说明校正集的样本代表性不足。KS 法和 SPXY 法选取后所建模型更佳, 其中 KS 法的相关系数最大、RMSECV 最小, 而 SPXY 法所建模型的 RMSEP 最小, 模型的预测能力最好。综合以上比较, 最终选择 SPXY 法对样本集进行划分。其它指标模型也存在类似结果。

2.8 波段选择

模型所用的光谱范围直接影响到模型的模型性能和运算效率。由于 NIRS 数据通常存在严重的谱

表 2 不同样本集划分方法的总黄酮模型结果比较

Table 2 Performance parameters of total flavonoid models via different sample set partitioning strategies

样本集划分方法	因子数	校正集和验证集				交叉验证	
		R_c	RMSEC / ($\mu\text{g}\cdot\text{mL}^{-1}$)	R_p	RMSEP / ($\mu\text{g}\cdot\text{mL}^{-1}$)	R_{cv}	RMSECV / ($\mu\text{g}\cdot\text{mL}^{-1}$)
随机法	5	0.991 4	60.3	0.963 2	95.8	0.980 2	99.9
浓度梯度法	4	0.987 2	75.8	0.975 5	90.0	0.978 5	98.1
KS 法	5	0.992 1	64.3	0.981 4	82.1	0.984 0	91.7
Duplex 法	5	0.992 4	58.2	0.979 4	112.0	0.984 3	83.3
SPXY 法	5	0.991 0	67.7	0.979 9	69.6	0.980 8	98.8

峰重叠和共线性, 因此需要选择包含分析物有用的吸收信息光谱区域, 排除干扰信息^[15]。

透反射采样模式综合了透射和反射的光学特性, 因此可以选择将短波波段和长波波段相结合。由原始光谱图(图1)可见, 在4 000~4 400 cm⁻¹波段噪声较为明显, 这是光纤吸收造成的干扰, 因此建模时舍弃了这一波段。以总黄酮定量模型为例, 表3为不同建模波段范围下复方阿胶浆中总黄酮的建模量的比较。可知, 4 429~8 000 cm⁻¹波段所建

模型性能最优; 4 429~4 900 cm⁻¹和6 469~7 377 cm⁻¹波段虽然相关系数较高, 且RMSEC较小, 但RMSECV较高, 说明模型可能存在过拟合现象。综合比较各波段模型性能, 选择4 429~8 000 cm⁻¹波段建模。

2.9 光谱预处理方法

由于环境及样本本身因素的影响, NIRS的测定中存在噪音, 一定的光谱预处理能够减少此类非目标信息带来的干扰。对原始光谱分别采用多元散

表3 不同波段范围的总黄酮模型结果比较

Table 3 Performance parameters of total flavonoid models with different wavebands

波段范围 / (cm ⁻¹)	因子数	校正集和验证集				交叉验证	
		R _c	RMSEC / (μg·mL ⁻¹)	R _p	RMSEP / (μg·mL ⁻¹)	R _{cv}	RMSECV / (μg·mL ⁻¹)
4 429~8 000	5	0.983 5	91.6	0.971 7	83.6	0.969 2	125.0
4 429~4 900	5	0.991 2	66.9	0.955 9	106.0	0.960 3	141.0
6 469~7 377	5	0.983 8	90.7	0.970 5	78.4	0.957 1	147.0
7 377~8 000	5	0.976 3	109.0	0.979 5	84.0	0.954 0	152.0
4 429~4 900, 7 377~8 000	5	0.983 3	92.1	0.972 9	81.5	0.962 1	147.0

射校正(MSC)、标准正则变换(SNV)、一阶导数、二阶导数、Savitsky-Golay(SG)滤波平滑和Norris导数滤波平滑等预处理方法, 比较所建模型的各项性能参数。以总黄酮定量模型为例, 经各种预处理方法所建模型的性能参数见表4(表中SG平滑为7点3阶平滑, Norris平滑段长为15, 段间距为5)。可见, 相比原始光谱模型, MSC和SNV模型相关系数均增大, RMSEP和RMSECV均减小, 说明模型的预测能力有所提高。其中MSC的建模结果又略好于SNV。经过导数和平滑处理后的光谱建立的模型各项参数均有不同程度的下降, 其中经过SG+1stD和MSC+SG+1stD处理后的模型交叉验证相

关系数显著减小, RMSECV和RMSEP显著增大, 模型预测性能降低明显。综上分析, 选择MSC对原始光谱进行预处理。

2.10 因子数的确定

采用NIRS建立定量校正模型时, 需要对因子数(即潜变量数)进行合理选择。当因子数过小时会导致模型所包含的信息不足, 因子数过大则容易导致过拟合现象的产生。因此, 选择合理的因子数对定量模型预测稳定性十分关键。本实验采用留一法交叉验证确定最佳因子数。总黄酮定量分析模型的最佳因子数的优选方法见图4。按此方法确定总黄酮、总皂苷和可溶性固形物定量模型的最佳因子

表4 不同预处理方法的总黄酮模型结果比较

Table 4 Performance parameters of total flavonoid models via different spectral preprocessing methods

预处理方法	因子数	校正集和验证集				交叉验证	
		R _c	RMSEC / (μg·mL ⁻¹)	R _p	RMSEP / (μg·mL ⁻¹)	R _{cv}	RMSECV / (μg·mL ⁻¹)
原始光谱	5	0.983 5	91.6	0.971 7	83.6	0.969 2	125.0
MSC	5	0.991 0	67.7	0.979 8	69.6	0.980 8	98.8
SNV	5	0.991 0	67.8	0.980 1	69.7	0.980 0	101.0
SG+1 st D	5	0.987 8	78.8	0.954 0	135.0	0.886 7	237.0
Norris+2 nd D	5	0.979 8	101.0	0.975 4	82.1	0.956 6	148.0
MSC+SG+1 st D	5	0.991 8	64.7	0.933 9	164.0	0.884 8	237.0
SNV+Norris+2 nd D	5	0.983 6	91.1	0.977 4	74.8	0.967 1	129.0

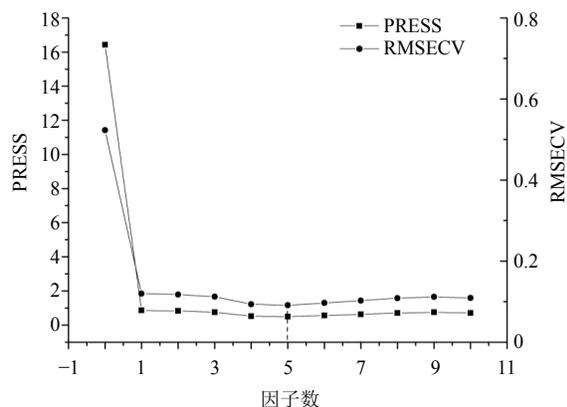


图 4 总黄酮模型最佳 PLS 因子数的优选

Fig. 4 Optimization of PLS factors for total flavonoid models

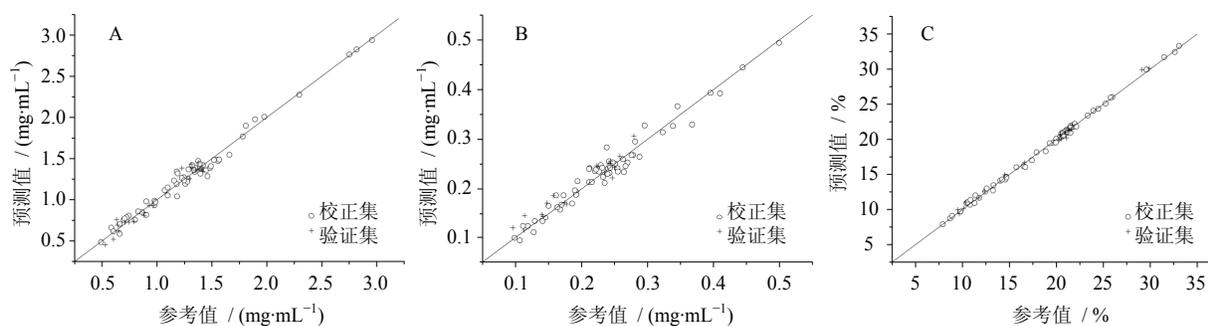


图 5 总黄酮 (A)、总皂苷 (B) 和可溶性固形物 (C) 的参考值与预测值的相关图

Fig. 5 Correlation between predicted values and measured values of total flavonoids (A), total saponins (B), and soluble solid content (C)

表 5 总皂苷、总黄酮和可溶性固形物模型的性能参数

Table 5 Performance parameters of calibration models for total saponins, total flavonoids, and soluble solid content

指标	因子数	校正集		交叉验证		验证集	
		R_c	RMSEC	R_{cv}	RMSECV	R_p	RMSEP
总黄酮	5	0.991 0	67.7 $\mu\text{g}/\text{mL}$	0.980 8	98.8 $\mu\text{g}/\text{mL}$	0.979 8	69.6 $\mu\text{g}/\text{mL}$
总皂苷	5	0.979 7	16.1 $\mu\text{g}/\text{mL}$	0.964 1	21.4 $\mu\text{g}/\text{mL}$	0.966 0	17.6 $\mu\text{g}/\text{mL}$
可溶性固形物	5	0.998 8	0.284%	0.996 3	0.504%	0.997 8	0.394%

以复方阿胶浆成品与实验室配制的浓缩稀释样品的近红外透反射光谱数据为基础，建立了复方阿胶浆成品的 NIRS 指纹图谱。

2.12.1 样本划分方法 从 30 批复方阿胶浆成品中随机选取 20 份样本建立标准库，并从这 20 份样本中随机选择 10 份采用相似度匹配算法建立相似度匹配模型，剩余 10 份标准样本用来计算相似度阈值。剩余 10 批复方阿胶浆成品与 30 份实验室配制的浓缩稀释样品一起作为验证样本，利用所建立的相似度匹配模型，以未知样品与标准样品光谱信息之间的相似度匹配值 (similarity matching values, SMV) 为依据，对这 40 份样品进行评价。SMV 值

数均为 5。

2.11 模型的建立

根据上述模型优化结果，运用 PLSR 法建立 NIRS 与复方阿胶浆中的总皂苷、总黄酮和可溶性固形物的校正模型。由相关图 (图 5) 可知，各模型的预测值与参考值之间具有良好的相关性。模型的各项性能参数见表 5，可见，所有预测相关系数均在 0.96 以上，表明各项指标的校正模型皆具有较好的拟合效果及预测能力。

2.12 基于相似度的复方阿胶浆成品指纹图谱

为了能够快速准确地鉴别复方阿胶浆成品，应用 NIRS 分析技术结合相似度匹配的定性分析方法，

的范围为 0~100，其值越高说明待测样品与校正集样品的光谱信号越相似。

2.12.2 相似度匹配模型的建立 选择波段 4 429~8 000 cm^{-1} 为建模波段，光谱预处理采用二阶导数和 Norris 平滑，以 10 批复方阿胶浆成品建立相似度匹配模型。

2.12.3 模型阈值的设定 相似度阈值的设定采用统计过程控制 3σ 原理^[16]。首先对样本的 SMV 值进行检验，若满足正态分布，则阈值 = SMV 均值 - $3 \times \text{SD}$ (标准偏差)。然后计算验证样本的 SMV 值，并与阈值比较，若大于或等于此阈值，则检验通过，反之则检验不通过。

应用 3σ 原理的前提是数据满足正态分布,因此首先对 10 份标准库样本的 SMV 值进行正态性检验 ($\alpha=0.05$), $P=0.490>0.05$, 说明 SMV 值分布符合正态分布。计算 SMV 阈值为 $99.753-3\times 0.071=99.54$ 。

2.12.4 验证样本 SMV 值计算结果 采用建立的相似度匹配模型计算 10 批复方阿胶浆成品与 30 份实验室配制的稀释样品的 SMV 值, 结果见图 6。编号 31~40 的样本为复方阿胶浆成品, SMV 值较为稳定, 且均达到或高于所设定阈值; 编号 1~30 的样本为实验室配制的浓缩稀释样品, SMV 值波动较为明显, 且均低于所设定阈值。另外, 与标准库样本浓度差距越大的样本, 其 SMV 值越低, 这也进一步说明了定性模型的准确性。验证样本的 SMV 值计算结果说明应用 NIRS 相似度匹配模型能够快速、准确地鉴别复方阿胶浆成品, 可辅助生产质量控制。

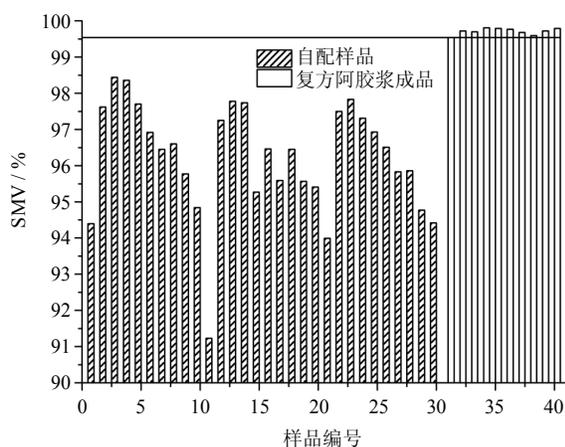


图 6 全部样本的相似度匹配值

Fig. 6 SMV graph for all samples

3 讨论

本实验研究了复方阿胶浆成品中总黄酮、总皂苷和可溶性固形物 3 项指标的快速测定方法, 并从异常样本剔除、样本集划分、光谱预处理方法等方面对所建立的定量分析模型进行了全面的优化, 所得模型的预测结果令人满意。另外, 本实验还采用 NIRS 结合相似度匹配算法建立了复方阿胶浆的 NIRS 指纹图谱, 从样品整体上反映其质量特征的全貌, 避免了采用 HPLC 等中药指纹图谱研究方法时常见的繁琐耗时等问题, 为中药指纹图谱分析提供了新的思路。本研究可减轻成品检验的工作量, 缩短批生产过程的总耗时; 另一方面也可适当提高抽检比例, 以增强成品检验结果的可靠性。

参考文献

- [1] 李文龙, 瞿海斌. 黄芪提取过程总皂苷质量浓度的在线监测 [J]. 中草药, 2012, 43(8): 1531-1535.
- [2] 李文龙, 邢丽红, 薛东升, 等. 一种基于近红外光谱技术的熊胆粉鉴别方法 [J]. 光谱学与光谱分析, 2011, 31(3): 673-676.
- [3] Li W L, Wang Y F, Qu H B. Near infrared spectroscopy as a tool for the rapid analysis of the *Honeysuckle* extracts [J]. *Vib Spectrosc*, 2012, 62: 159-164.
- [4] 吕琳昂, 师涛, 杨辉华, 等. NIR 在线检测监控安神补脑液水提过程的研究 [J]. 中草药, 2009, 40(2): 224-228.
- [5] 张荻, 侯丽, 孙韬, 等. 复方阿胶浆改善化疗相关性贫血的临床研究 [J]. 北京中医药大学学报: 中医临床版, 2012, 19(3): 15-18.
- [6] 中国药典 [S]. 一部. 2010.
- [7] 王静, 莫必琪, 李斌, 等. 近红外光谱法预测红参醇提过程中总皂苷的变化研究 [J]. 中草药, 2007, 38(9): 1323-1326.
- [8] Shi J Y, Zou X B, Zhao J W, et al. Determination of total flavonoids content in fresh *Ginkgo biloba* leaf with different colors using near infrared spectroscopy [J]. *Spectrochim Acta A*, 2012, 94: 271-276.
- [9] Rosas J G, Blanco M, González J M, et al. Real-time determination of critical quality attributes using near-infrared spectroscopy: A contribution for process analytical technology (PAT) [J]. *Talanta*, 2012, 97: 163-170.
- [10] Chen Y, Xie M Y, Zhang H, et al. Quantification of total polysaccharides and triterpenoids in *Ganoderma lucidum* and *Ganoderma atrum* by near infrared spectroscopy and chemometrics [J]. *Food Chem*, 2012, 135(1): 268-275.
- [11] Galvão R K H, Araujo M C U, José G E, et al. A method for calibration and validation subset partitioning [J]. *Talanta*, 2005, 67(4): 736-740.
- [12] 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.
- [13] Kennard R W, Stone L A. Computer aided design of experiments [J]. *Technometrics*, 1969, 11(1): 137-148.
- [14] Snee R D. Validation of regression models: Methods and examples [J]. *Technometrics*, 1977, 19(4): 415-428.
- [15] 杨皓旻, 卢启鹏, 黄富荣. 近红外光谱分析中建模光谱宽度的选择 [J]. 红外与毫米波学报, 2011, 30(6): 522-525.
- [16] 张峰, 刘江生, 刘泽春, 等. 近红外透射光谱技术用于烟用香精的品质控制 [J]. 中国烟草学报, 2009, 15(3): 12-16.