

近红外漫反射光谱法定量分析川芎中的阿魏酸含量^{*}

王小梅¹, 焦龙^{1,2}, 刘小丽³, 李华^{1**}

(1. 西北大学化学与材料科学学院, 西安 710069; 2. 西安石油大学化学化工学院, 西安 710065;

3. 西北大学生命科学学院, 西安 710069)

摘要 目的: 应用近红外光谱分析技术和化学计量学方法构建了川芎中阿魏酸含量的定量测定模型。**方法:** 通过偏最小二乘法建立数学模型, 并对预测集进行预测。**结果:** 34 个川芎样品经交叉验证建立校正模型, 交叉验证均方根误差 (RMSECV) 为 0.146%, 决定系数 (R^2) 为 0.9883。用 11 个川芎样品进行预测, 预测值与参考值的决定系数 (R^2) 达 0.9751, 预测均方根误差 (RMSEP) 为 0.251%。**结论:** 该方法简便快速, 结果准确, 可应用于对不同产地不同批次的川芎进行快速检查或质量控制。

关键词: 偏最小二乘法; 近红外; 漫反射光谱; 阿魏酸; 川芎; 快速检测; 中药材

中图分类号: R917

文献标识码: A

文章编号: 0254-1793(2011)06-1016-04

Quantitative analysis of ferulic acid in *Ligusticum chuanxiong* Hort. by near infrared diffuse reflectance spectroscopy^{*}

WANG Xiao-me¹, JIAO Long^{1,2}, LIU Xiao-li³, LI Hua^{1**}

(1. School of Chemistry and Materials Science, Northwestern University, Xi'an 710069, China;

2. College of Chemistry and Chemical Engineering, Xi'an Shiyou University, Xi'an 710065, China;

3. School of Life Sciences, Northwestern University, Xi'an 710069, China)

Abstract Objective: Near infrared (NIR) spectroscopy and chemometrics were applied to the quantitative analysis of ferulic acid in *Ligusticum chuanxiong* Hort. **Method:** Partial least squares (PLS) regression model was built using a calibration set (34 samples). The prediction set was established by 11 samples. And we applied this method to predict the data in the prediction set. **Results:** The coefficient of determination (R^2) was 0.9883 as indicated from the cross-validation, and the root mean square error of cross validation (RMSECV) was 0.146%. The coefficient of determination (R^2) of the reference and predicted values was 0.9751, and the root mean square of prediction (RMSEP) was 0.251%. NIR spectroscopy method was simple, rapid and accurate. **Conclusion:** The method can be applied in rapid qualitative and quantitative of different batches of *Ligusticum chuanxiong* Hort. from the different places of origin.

Key words: partial least squares; near infrared; diffuse reflectance; ferulic acid; *Ligusticum chuanxiong* Hort.; rapid testing; traditional Chinese medicine

川芎(*Ligusticum chuanxiong* Hort.) 为伞形科植物的干燥根茎^[1], 含有内酯类、生物碱类、甾体、酚类以及挥发油类等多种化合物^[2]。其味辛、微苦、性温, 具有活血祛瘀、行气开郁、祛风止痛之功效。临床上主要用于治疗心血管疾病、妇科疾病等。川芎中主要成分阿魏酸的化学名称为 3-(4-羟基-3-甲氧基苯基)-2-丙烯酸。阿魏酸的分子式为 $C_{10}H_{10}O_4$, 相对分子质量为 194.18, 其结构式如图 1

所示。

阿魏酸含量测定已有文献报道, 主要采用流动注射化学发光法^[3]、液相色谱法^[4]。这些方法预处理样品比较烦琐, 而且测定过程中耗时, 难以满足制药企业实施在线控制分析的要求。

近红外光谱(near-infrared spectroscopy, NIRS)分析技术是近年来发展极为迅速而且应用相当广泛的 1 种绿色分析技术。该方法不需对样品进行预处理

* 国家自然科学基金(No. 20975081); 教育部科学技术研究重点项目(204182); 陕西省教育厅产业化中试项目(No. 09JC04)

** 通讯作者 Tel: (029) 88302942; E-mail: huali@nwu.edu.cn

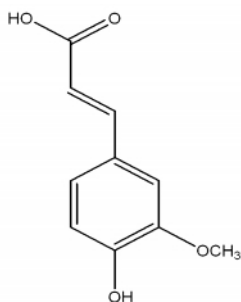


图1 阿魏酸结构式

Fig 1 The structure of ferulic acid

理,操作简便快速,已被广泛应用于农业^[5,6]、烟草^[7,8]、食品^[9~12]和药物^[13,14]的定性定量分析。目前,采用 NIR 光谱分析技术定量测定川芎中阿魏酸的文献尚未见报道。本文采用偏最小二乘法结合近红外光谱分析技术构建定量测定川芎中阿魏酸的数学模型,无需对样品进行繁杂的预处理,测定方法简单快速,结果准确可靠。

1 仪器和材料

VERTEX 70 型傅里叶变换红外光谱仪(德国 Bruker 公司),岛津 SCL-10AVP 高效液相色谱仪(日本岛津公司),KQ5200DB 型数控超声波清洗器(上海楚定分析仪器有限公司),玻璃仪器气流烘干机(河南太康科教仪器厂),Sartorius 电子天平(北京赛多利斯天平有限公司),石英亚沸高纯水蒸馏器(江苏省宜兴市勤华石英玻璃仪器厂)。

四川等不同产地不同批次的川芎样品共 45 个,阿魏酸对照品(批号 11773-200613,中国药品生物制品检定所),甲醇(色谱纯,西安化学试剂厂),水为二次蒸馏水。

2 阿魏酸含量测定

HPLC 测量条件^[15,16]:采用 Kromasil C₁₈ 色谱柱(4.6 mm × 150 mm, 5 μm),流动相为 1% 醋酸(A) - 甲醇(B),梯度洗脱(0~40 min, 10% B→50% B; 40~70 min, 50% B→100% B),进样体积为 10 μL,流速为 1 mL · min⁻¹,检测波长 280 nm,柱温 30 ℃。根据上述条件,测定所有川芎样品,将所得到的测定值作为参考值。

3 NIR 光谱采集

将川芎样品用粉碎机粉碎,并过 100 目筛,装在石英样品杯中,厚度大约 3 cm(光不透过样品),置于漫反射积分球上进行测定,设置扫描次数为 32 次,分辨率为 8 cm⁻¹,采集光谱范围 9000~4000 cm⁻¹,每份样品重复扫描 3 次,取其平均光谱。

4 数据处理方法

样品随机分为校正集和预测集,并保证校正集样品浓度范围大于预测集,建立 NIR 光谱校正模型。将 45 个样品随机分为校正集(34 个样品)和预测集(11 个样品),校正集样品用于校正模型的建立。将所得到的光谱数据导入 OPUS 6.5 定量分析软件中,选择用于建模的光谱区间,对光谱预处理方法进行比较,采用 PLS 建立 NIR 光谱定量校正模型。以校正集的交叉验证均方根误差(RMSECV)为指标选择建模参数优化模型结构,进而得到最优的校正模型。同时考察了 PLS 模型的潜变量数(latent variable number)和光谱预处理方法对 PLS 模型的影响,以交叉验证均方根误差(RMSECV)、预测均方根误差(RMSEP)和决定系数(R²)为优化参数对模型的稳健性进行了评价。RMSECV 计算公式:

$$\text{RMSECV} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (1)$$

式中, n 为校正集样品数, y_i 为校正集样品 i 的参考值, $\hat{y}_i$ 为建立模型与去除样品 i 时样品 i 的预测值。RMSECV 值越小,所构建的模型性能越好。对于预测集,预测均方根误差(RMSEP):

$$\text{RMSEP} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

式中, n 为预测集样品数, y_i 为预测集样品 i 的参考值, $\hat{y}_i$ 为所建模型对预测集样品 i 的预测值。

在校正集和预测集中,参考值与预测值之间的决定系数由如下公式计算所得:

$$R^2 = \left[1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] \times 100 \quad (3)$$

式中, $\bar{y}$ 为校正集和预测集中所有样品的参考值的平均值。在此,决定系数的含义为由解释变量引起的被解释变量的变化占被解释变量总变化的比重,用来判定回归直线拟合的优劣,该值越大说明拟合得越好。即决定系数的大小决定了相关的密切程度。当 R^2 越接近 1 时,表示相关性越高。

5 结果与讨论

5.1 川芎样品的 NIR 光谱图 图 2 是川芎样品的近红外漫反射谱图。川芎样品中所含的主要成分种类很多,谱峰复杂,而且相互重叠比较严重,单靠近红外谱图无法直接对样品中的阿魏酸含量进行定量,因此,只有将近红外光谱与相关化学计量学方法相结合,才能实现川芎样品中阿魏酸的定量分析。

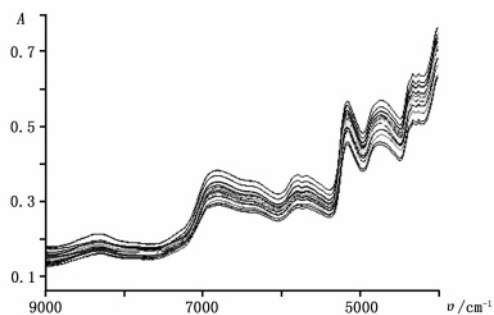


图2 川芎的 NIR 光谱图

Fig 2 The near infrared spectra of *Ligusticum chuanxiong* Hort. samples

5.2 建模区间的选择 通过比较 OPUS 6.5 软件光谱范围优化结果中的 RMSECV 值,考察了光谱范围对模型性能的影响,如表 1 所示。由表 1 的结果可看出,在 7501.9 ~ 6799.9 cm^{-1} 和 4424 ~ 4246.6 cm^{-1} 光谱范围内所建模型性能最好。因此,选择最优的建模区间为 7501.9 ~ 6799.9 cm^{-1} 和 4424 ~ 4246.6 cm^{-1} 。

表 1 光谱范围对 RMSECV 的影响

Tab 1 The effect of spectral coverage on RMSECV

序号 (No.)	光谱范围 (spectral region) / cm^{-1}	RMSECV /%
1	7501.9 ~ 6799.9; 4424 ~ 4246.6	0.146
2	7501.9 ~ 6799.9; 4601.4 ~ 4246.6	0.191
3	7501.9 ~ 6799.9; 4601.4 ~ 4424	0.236
4	7501.9 ~ 5450; 4601.4 ~ 4246.6	0.239
5	7501.9 ~ 5450	0.241
6	7501.9 ~ 6097.9; 4601.4 ~ 4424	0.245
7	6799.9 ~ 6097.9; 4601.4 ~ 4246.6	0.280

5.3 光谱预处理方法的选择 一般来说,样品的物理性质的不同会引起光谱基线和斜率的变化,在建立校正模型前需采用化学计量学方法对光谱进行预处理以消除这些影响。OPUS 6.5 软件中内建的光谱预处理方法有消除常数偏移量、多元散射校正(multiplicative scatter correction, MSC)、一阶导数法、二阶导数法、Savitzky - Golay 方法、标准正交变量变换(standard normal variate, SNV)等。本文在选定的光谱区间,采用不同的预处理方法进行比较寻求最优。通过交叉验证,采用 RMSECV 来考察校正模型的性能。通常情况下, RMSECV 值越小,模型中光谱信息与成分含量相关越好。阿魏酸预处理方法的筛选分析结果如表 2 所示。从表 2 可以看出,川芎样品光谱经消除常数偏移量、多元散射校正、标准正交变量变换、最大 - 最小归一化和一阶导数后,在波段 7501.9 ~ 6799.9 cm^{-1} 和 4424 ~ 4246.6 cm^{-1} 光谱的预测效

果中,一阶导数的预测效果最好, RMSECV 最小。因此,最适合该模型的预处理方法为一阶导数法。

表 2 不同光谱预处理方法对 PLS 模型参数的影响

Tab 2 The effect of different preprocessing methods on PLS model

序号 (No.)	预处理方法 (pretreatment)	RMSECV /%
1	一阶导数法(first order derivative)	0.146
2	多元散射校正(multiplicative scatter correction, MSC)	0.221
3	标准正交变量变换(standard normal variate, SNV)	0.227
4	无光谱预处理(no spectral preprocessing)	0.231
5	最大 - 最小归一化(max - min normalized)	0.235
6	消除常数偏移量(constant offset elimination)	0.240

5.4 由校正集样品建立数学模型

本文采用 PLS 建立定量校正模型时,采用不同的潜变量数,模型预测能力有较大的差异。在校正集样品一定的情况下,潜变量数取得太少,模型预测能力降低,出现不充分拟合现象;潜变量数过多,会导致模型过于复杂,并出现过拟合现象。因此,必须合理选择潜变量数。本文采用交叉验证法,考察了潜变量数对 RMSECV 值的影响。图 3 为 RMSECV 值随潜变量数变化的曲线图。由图 3 可看出,当 PLS 模型的潜变量数取 7 时,交互验证均方根误差 RMSECV 值最小,所选潜变量数最佳。

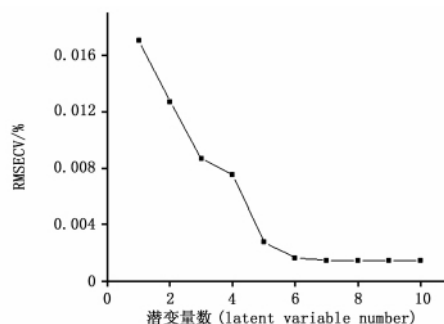


图 3 潜变量数对 RMSECV 的影响

Fig 3 The effect of latent variable number on RMSECV

因此,将 7501.9 ~ 6799.9 cm^{-1} 和 4424 ~ 4246.6 cm^{-1} 范围内的光谱,经过一阶导数预处理后,设定潜变量数为 7,采用 PLS 法建立用于计算阿魏酸含量的校正模型。如图 4 为全部样品的 leave - one - out 交叉验证结果。全部样品中阿魏酸含量的预测值与其参考值之间的 R^2 达到 0.9883,全部样品的 RMSECV 值为 0.146%。

5.5 预测集阿魏酸的分析 用建立的最佳校正模型,对预测集样品中的阿魏酸含量进行预测,预测结果见表 3。由表 3 可以看出,预测值与参考值之间的

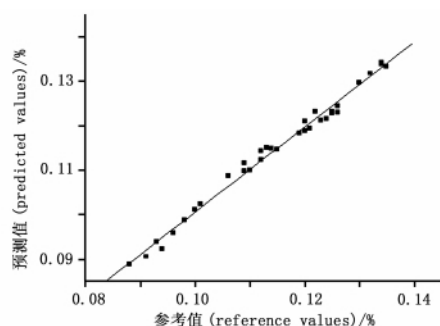


图4 校正模型预测值与参考值之间的相关图

Fig 4 The coefficient program of predicted and reference values for calibration set

R^2 达 0.9751, RMSEP 为 0.251%, 表明该校正模型预测的结果准确, 因此可以认为该数学模型比较可靠。

表3 预测集中阿魏酸的预测结果

Tab 3 The predicted results of ferulic acid in the prediction set

序号 (order number)	参考值 (reference values) / %	预测值 (predicted values) / %	误差 (error)	相对误差 (relative error) / %
1	0.092	0.089	-0.003	-3.2
2	0.124	0.125	0.001	0.8
3	0.115	0.115	0.000	0.0
4	0.102	0.106	0.004	3.9
5	0.108	0.109	0.001	0.9
6	0.117	0.113	-0.004	-3.4
7	0.115	0.116	0.001	0.9
8	0.119	0.120	0.002	1.7
9	0.118	0.119	0.001	0.8
10	0.096	0.096	0.000	-0.0
11	0.094	0.095	0.001	1.1

6 结论

通过应用近红外光谱分析技术和化学计量学方法构建了川芎中阿魏酸的定量分析模型, 最终选出阿魏酸含量分析的最佳波段, 同时考察了预处理方法和 PLS 模型的潜变量数对模型的影响, 并对 PLS 模型性能做出了评价。该方法简便快速, 预测结果准确可靠, 可应用于对不同产地不同批次的川芎进行快速检查或质量控制。

参考文献

- 1 ChP(中国药典). 2005. Vol I (一部): 28
- 2 WANG Wen-xiang(王文祥), GU Ming(顾明), JIANG Xiao-gang(蒋小岗), et al. Studies on chemical constituents of Ligusticum chuanxiong(川芎化学成分研究). *Chin Tradit Herb Drugs*(中草药), 2002, 33(1): 4
- 3 Wang JP, Li NB, Luo HQ. Chemiluminescence determination of fer-

- ulic acid by flow-injection analysis using cerium (IV) sensitized by rhodamine 6G. *Spectrochim Acta Part A: Molecul Biomolecul Spectrosc*, 2008, 71(1): 204
- 4 Philippe AG, Mathieu R, Denis B, et al. Quantitative analysis of plasma caffeic and ferulic acid equivalents by liquid chromatography tandem mass spectrometry. *J Chromatogr B*, 2009, 877(31): 3965
- 5 Li XL, He Y, Wu CQ, et al. Non-destructive discrimination of paddy seeds of different storage age based on Vis/NIR spectroscopy. *J Stored Products Res*, 2008, 44(3): 264
- 6 Giorgia F, Marina C, Mario LV, et al. Different feature selection strategies in the wavelet domain applied to NIR-based quality classification models of bread wheat flours. *Chemometr Intell Lab Syst*, 2009, 99(2): 91
- 7 Cai WS, Li YK, Shao XG. A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra. *Chemometr Intell Lab Syst*, 2008, 90(2): 188
- 8 Shao XG, Bian XH, Cai WS. An improved boosting partial least squares method for near-infrared spectroscopic quantitative analysis. *Anal Chim Acta*, 2010, 666(1-2): 32
- 9 Liu YD, Sun XD, Ouyang AG. Nondestructive measurement of soluble solid content of navel orange fruit by visible-NIR spectrometric technique with PLSR and PCA-BPNN. *LWT. Food Sci Technol*, 2010, 43(4): 602
- 10 Els B, Serban AS, Mihai N, et al. Postharvest quality of apple predicted by NIR-spectroscopy: Study of the effect of biological variability on spectra and model performance. *Postharvest Biol Technol*, 2010, 55(3): 133
- 11 Cecilia S, Luis RS. Application of NIR and MIR spectroscopy in quality control of potato chips. *J Food Compos Anal*, 2009, 22(6): 596
- 12 Yu HY, Niu XY, Lin HJ, et al. A feasibility study on on-line determination of rice wine composition by vis NIR spectroscopy and least-squares support vector machines. *Food Chem*, 2009, 113(1): 291
- 13 Puchert T, Lochmann D, Menezes JC, et al. Near-infrared chemical imaging (NIR-CI) for counterfeit drug identification-a four-stage concept with a novel approach of data processing (linear image signature). *J Pharm Biomed Anal*, 2010, 51(1): 138
- 14 Atul DK, Paul WSH, Celine VL. In-line quantification of micronized drug and excipients in tablets by near infrared (NIR) spectroscopy: Real time monitoring of tableting process. *Internat J Pharm*, 2010, 396(1-2): 63
- 15 HUANG Yan-ping(黄艳萍), HUANG Yong-hong(黄勇红). TLC of Angelica menstruation film and determination of ferulic acid (当归调经片的薄层鉴别和阿魏酸的测定). *Chin J Exp Tradit Med Form*(中国实验方剂学杂志), 2010, 16(6): 119
- 16 WANG Fang(王芳), KUANG Wei-hua(匡维华). RP-HPLC determination of ferulic acid of *Ligusticum chuanxiong* Hort. from different areas(反相高效液相色谱法对不同产地川芎中阿魏酸的测定). *Chin J Mod Appl Pharm*(中国现代应用药学), 2002, 19(4): 310

(本文于 2011 年 2 月 22 日收到)