

文章编号: 0253-2239(2009)10-2795-05

# 近红外光谱在快速测定蛹虫草有效成分含量中的应用

王 迪 张媛莉 孟庆繁 逯家辉 金 璐 滕利荣

(吉林大学生命科学学院, 吉林 长春 130012)

**摘要** 采用偏最小二乘法(PLS)建立测定蛹虫草菌丝体中多糖、蛋白、虫草酸和腺苷含量的近红外光谱定量分析模型。选择了原始光谱建立蛹虫草菌丝体中腺苷和多糖含量的 PLS 定量分析模型,选择卷积光滑预处理后的光谱建立蛹虫草菌丝体中蛋白和虫草酸含量的 PLS 定量分析模型。测定蛹虫草菌丝体中多糖、蛋白、虫草酸和腺苷含量的最优 PLS 定量分析模型的交叉验证均方根误差(RMSECV)分别为 0.0091,0.0222,0.0088 和 0.6520,预测均方根误差(RMSEP)分别为 0.0079,0.0196,0.0087 和 0.5780。结果表明,测定蛹虫草菌丝体四种有效成分含量的定量分析模型稳健性好和预测精度高,在中药有效成分定量分析方面有很大的应用前景。

**关键词** 测量; 近红外光谱; 偏最小二乘法; 蛹虫草

中图分类号 O657.33; R927.2 文献标识码 A doi: 10.3788/AOS20092910.2795

## Application of Near Infrared Spectroscopy on Rapid Determination of Essential Components for Cordyceps Militaris

Wang Di Zhang Aili Meng Qingfan Lu Jiahui Jin Lu Teng Lirong

(College of Science, Jilin University, Changchun, Jilin 130012, China)

**Abstract** Partial least squares (PLS) method was applied to establish models with near infrared spectroscopy (NIRS) for quantitative analysis of polysaccharides, protein, cordycepic acid, adenosine content in cordyceps militaris mycelium. The origin spectra were used to establish the PLS models for quantitative analysis of adenosine and polysaccharides content in cordyceps militaris mycelium. The Savitzky-Golay smoothing spectra were applied to establish the PLS models for quantitative analysis of protein and cordycepic acid content. The values of root mean square error of cross-validation (RMSECV) of the most optimum PLS quantitative analysis models of polysaccharides, protein, cordycepic acid and adenosine were 0.0091, 0.0222, 0.0088 and 0.6520, the values of root mean square error of prediction (RMSEP) were 0.0079, 0.0196, 0.0087 and 0.5780. The results demonstrated that the quantitative analysis model for determining the four contents of cordyceps militaris mycelium possessed good stability and high precision. This method has great foreground in quantitative analysis of Chinese traditional medicine.

**Key words** measurement; near infrared spectroscopy; partial least squares; cordyceps militaris

## 1 引 言

蛹虫草[Cordyceps militaris(L. ex Fr.) Link]又名北冬虫夏草、北虫草等,与冬虫夏草同属异种。作为一种具有多种药理功能、抗癌活性的药用真菌蛹虫草,随着人们逐渐发现和认识蛹虫草的滋补疗效和提高人体免疫功能的功效,它的开发利用倍受世人的极大关注。液体深层发酵的蛹虫草菌丝体中含有蛋白、多糖、虫草酸和腺苷等有效成分,具有增

强人体免疫功能及抗肿瘤、抗氧化,抗衰老、降低血脂及血压<sup>[1~3]</sup>等作用。但现行常用的腺苷测定方法高效液相法<sup>[4]</sup>、蛋白测定方法凯氏定氮法<sup>[5]</sup>、多糖测定方法蒽酮-硫酸法<sup>[6]</sup>和虫草酸测定方法比色法<sup>[7]</sup>都具有材料损耗大,测定过程复杂,化学试剂消耗量大且分析时间长等缺点。

近红外光谱介于可见光区和中红外光区之间,波长范围为 780 ~ 2526 nm,波数范围为

收稿日期: 2008-11-17; 收到修改稿日期: 2008-12-31

基金项目: 中国医学基金会新药发展基金(20061108)资助项目。

作者简介: 王 迪(1984—),男,硕士研究生,主要从事生物药物的非破坏在线质量控制的方面的研究。

E-mail: wangdi4093@email.jlu.edu.cn

导师简介: 滕利荣(1954—),男,博士生导师,主要从事生物药物有效成分的分离,新药开发等方面的研究。

E-mail: tenglr@jlu.edu.cn

4000~13330  $\text{cm}^{-1}$ 。其光谱信息来源于有机分子中的含氢基团(C-H, O-H, N-H)振动的合频与各级倍频的吸收,是样品中有机分子含氢基团的特征振动信息,化学信息量相当丰富<sup>[8]</sup>。蛹虫草中的多糖、蛋白、腺苷、虫草酸等都含有各种氢基团,所以通过对蛹虫草的近红外光谱分析可以测定这些成份的含量。近红外光谱技术在建立可靠的校正模型的基础上,可直接对样品进行无损检测,具有分析速度快、效率高、成本低和多组分同时测定等优点,因此应用非常广泛<sup>[9-11]</sup>。

偏最小二乘法(PLS)是目前化学计量学中最有效的分析方法之一。它从自变量矩阵和因变量矩阵中提取偏最小二成分,有效地降维,并消除自变量间可能存在的复共线关系,明显改善数据结果的可靠性和准确度<sup>[12]</sup>。本文应用近红外光谱法结合 PLS(NIR-PLS)建立蛹虫草菌丝体中多糖、蛋白、虫草酸和腺苷含量的定量分析模型,并用所建模型对预测集样品进行预测,得到较好的结果,该方法有望成为一种代替现行真菌活性成分测定的快速安全分析方法。

## 2 实 验

### 2.1 材料、试剂与设备

蛹虫草(CGMCC 5.699),购于中国科学院微生物研究所。

日本岛津 UV-3150 型紫外可见近红外分光光度计,日本岛津 ISR-3100 积分球附件,联想家悦 E3030 微型计算机,全自动在位灭菌发酵系统,日本岛津 LC-10AT 高效液相色谱仪,全自动凯氏定氮仪。

### 2.2 样品收集及光谱测定

通过化学诱变法获得 389 株诱变蛹虫草菌株,分别通过液体深层发酵获得蛹虫草菌丝体,冻干粉碎过 60 目筛,制备蛹虫粉备用。

应用积分球,光谱通带宽度为 12 nm,扫描波长

范围 800~2500 nm,每个样品进行近红外光谱扫描 3 次,取平均值作为该样品的近红外光谱。

### 2.3 蛹虫草菌丝体四种有效成分含量的测定

不同批次蛹虫草菌丝体样品中腺苷、蛋白、多糖和虫草酸的含量测定分别采用高效液相色谱法<sup>[4]</sup>、凯氏定氮法<sup>[5]</sup>、蒽酮-硫酸法<sup>[6]</sup>和比色法<sup>[7]</sup>。

### 2.4 PLS 定量分析模型的建立

根据样品近红外(NIR)光谱的第一主成分(PC<sub>1</sub>)和第二主成分得分(PC<sub>2</sub>)得分分布将样品分成校正集和预测集。校正集样品用于建立校正模型,采用交互验证的方法检验模型的内部稳健性和拟合效果,评价参数为交互验证均方根误差(RMSECV)和交互验证预测值与真实值间的相关系数(R<sub>v</sub>);采用预测集检验模型的预测能力,防止模型过拟合,评价参数为预测均方根误差(RMSEP)和预测集中预测值与真实值间的相关系数(R<sub>p</sub>)。RMSECV 和 RMSEP 的计算方法参考文献[13]。

实验中分别采用卷积平滑(Savitzky-Golay smoothing),快速傅立叶变换(FFT),一阶导数(First derivative),二阶导数(Second derivative)和光谱小波变换(WT)分别对样品的 NIR 光谱进行预处理,原始光谱和各预处理后的光谱分别建立 PLS 定量分析模型,所建模型经过最适主因子数的选择,比较模型的内部稳健性、拟合效果以及预测性能,确定最优的蛹虫草菌丝体有效成分近红外定量分析模型。

## 3 结果与讨论

### 3.1 校正集与预测集的选择

多元校正方法要求适当设计校正集样品与预测集样品,以期得较好的预测结果。本实验以 307 个蛹虫草样品作为校正集,以 82 个蛹虫草样品作为预测集,预测集样品较均匀地分布于校正集样品之中。校正集与预测集样品中各有效成分测定含量统计结果如表 1 所示。

表 1 蛹虫草有效成分的质量含量统计

Table 1 Statistical mass contents of cordyceps militaris

	Components	Samples numbers	Average	Ranges
calibration sets	Protein	307	0.5352 g/g	0.1531~0.7686 g/g
	Adenosine		2.8964 mg/g	0.4082~6.2683 mg/g
	Cordycepic acid		0.08269 g/g	0.0384~0.1417 g/g
	Polysaccharides		0.0618 g/g	0.0305~0.1263 g/g
prediction sets	Protein	82	0.5231 g/g	0.1567~0.6934 g/g
	Adenosine		2.9997 mg/g	0.3840~5.9159 mg/g
	Cordycepic acid		0.0839 g/g	0.0460~0.1373 g/g
	Polysaccharides		0.0642 g/g	0.0348~0.1248 g/g

### 3.2 最优模型的建立

由于采集的近红外原始光谱含有杂噪音和基线飘移的干扰,本文采用不同的处理方法对蛹虫草菌丝体样品的 NIR 光谱进行预处理,采用 FFT 和卷积平滑光谱预处理方法时,考察了不同窗口大小对处理效果的影响;采用一阶导数和二阶导数光谱预

表 2 不同预处理方法和主因子数对 PLS 模型的影响

Table 2 Influence of different pretreating methods and factors on capabilities of PLS models

Pretreating method	Components	Window size/ Scale size/ Width difference	Factor	RMSECV	$R_v$	RMSEP
Origin spectra	<b>Adenosine</b>	—	<b>10</b>	<b>0.6520</b>	<b>0.9014</b>	<b>0.5780</b>
	Protein	—	9	0.0227	0.9763	0.0231
	<b>Polysaccharides</b>	—	<b>10</b>	<b>0.0091</b>	<b>0.8871</b>	<b>0.0079</b>
FFT	Cordycepic acid	—	9	0.0090	0.8774	0.0087
	Adenosine	10	12	0.6467	0.9057	0.6080
	Protein	20	11	0.0225	0.9770	0.0199
Savitzky— Golay smoothing	Polysaccharides	10	8	0.0094	0.8655	0.0091
	Cordycepic acid	20	11	0.0087	0.8798	0.0092
	Adenosine	20	11	0.6499	0.9052	0.6130
First derivative	<b>Protein</b>	<b>40</b>	<b>11</b>	<b>0.0222</b>	<b>0.9774</b>	<b>0.0196</b>
	Polysaccharides	10	8	0.0094	0.8676	0.0090
	Cordycepic acid	20	10	0.0088	0.8798	0.0086
Second derivative	Adenosine	40	9	0.7087	0.8987	0.6520
	Protein	40	9	0.0217	0.9771	0.0211
	Polysaccharides	20	5	0.0092	0.8765	0.0083
Wavelets transform	Cordycepic acid	50	10	0.0091	0.8803	0.0096
	Adenosine	40	3	0.8060	0.8320	0.7850
	Protein	50	6	0.0246	0.9701	0.0230
Wavelets transform	Polysaccharides	40	5	0.0094	0.8723	0.0086
	Cordycepic acid	50	6	0.0110	0.8034	0.0126
	Adenosine	2	12	0.6469	0.9078	0.5910
Wavelets transform	Protein	4	12	0.0226	0.9777	0.0204
	Polysaccharides	1	10	0.0091	0.8873	0.0079
	Cordycepic acid	4	10	0.0088	0.8807	0.0087

通过对表 2 分析可知,原始光谱所建立的多糖含量 PLS 定量分析模型,在稳定性上优于其它模型, RMSEP 值为 0.0079, 预测能力较强;虽然 db2 函数尺度为 1 的小波变换处理光谱建模后拟合度上略高,但考虑到复杂的处理过程,本实验选择原始光谱建立多糖含量定量分析模型。移动窗口为 20 的卷积平滑处理后的光谱所建立的蛋白含量 PLS 定量分析模型的拟合度和预测能力最佳,其相关系数  $R_v$  达到 0.9774。移动窗口为 20 的 FFT 处理后的光谱所建立的虫草酸含量 PLS 定量分析模型虽然 RMSECV 值为 0.0087 较 db2 函数尺度为 4 的小波变换处理后的光谱建立的虫草酸含量定量分析模型的 RMSECV 略低,但似合度和预测能力较差,选择小波变换处理后的光谱所建立虫草酸含量 PLS 分析模型。db2 函数下 2 尺度小波变换处理后的光谱所建立的腺苷含量定量分析模型虽然在模型的稳定

处理方法时,考察了不同差分宽度对处理效果的影响;采用光谱小波变换预处理方法时,选择 db2 函数考察了不同尺度对处理效果的影响。分别采用预处理后的光谱建立定量分析模型,每个模型经过内部交互验证和外部预测集的验证,选择最适主因子数,比较结果见表 2。

性上略优于原始光谱建立的模型,但考虑到该模型预测能力(RMSEP 值为 0.5910)不如原始光谱建立的模型(RMSEP 值为 0.5780),选择原始光谱建立腺苷含量定量分析模型。

经过以上分析,本实验最终选择原始光谱建立蛹虫草中腺苷和多糖含量的 PLS 定量分析模型,选择的主因子数均为 10,主成分得分为 99.94%;选择卷积平滑处理后的光谱建立蛹虫草中蛋白含量的 PLS 定量分析模型,最适主因子数为 11,主成分得分为 99.98%;选择小波变换处理后的光谱建立虫草酸含量的 PLS 模型,最适主因子数为 10,主成分得分为 99.97%。

### 3.3 最优模型的内部交互验证

采用最优的条件,建立测定蛹虫草样品中四种主要成分含量的最优模型,为了评价模型的稳定性和拟合效果,对最优模型进行内部交互验证,结果如

图 1 所示,多糖、蛋白、虫草酸和腺苷含量的交互验证预测值与实验测定值间的相关系数( $R_v$ )分别为 0.8871,0.9774,0.8807 和 0.9014, RMSECV 值分

别为 0.0091,0.0222,0.0088 和 0.6520,说明交互验证预测值与实验测定值接近,说明模型的稳定性和拟合效果非常好。

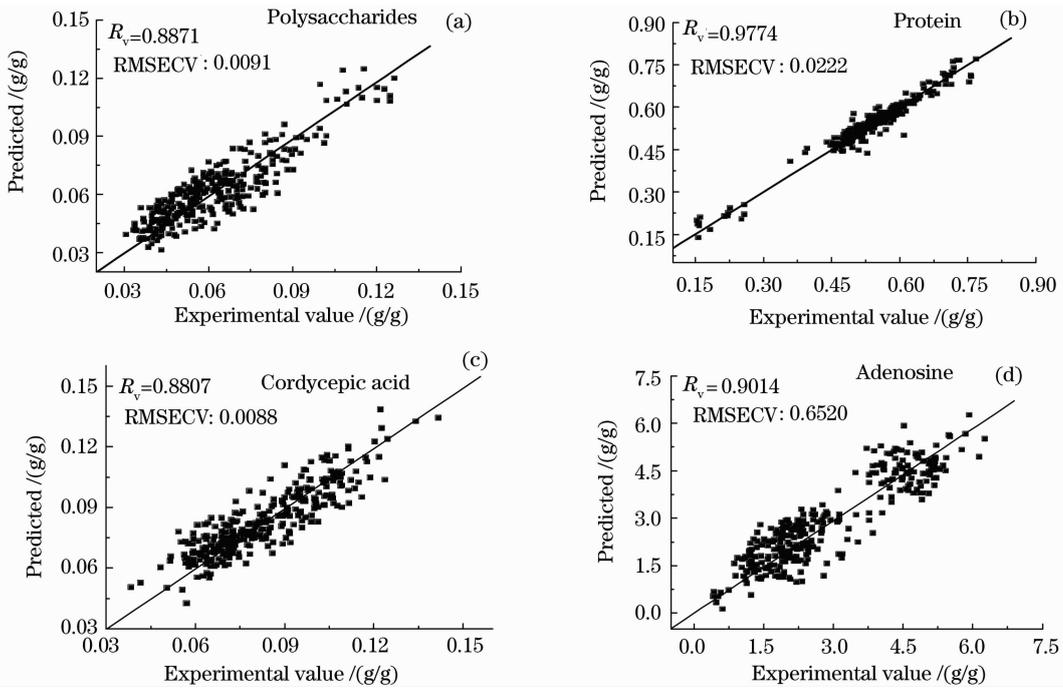


图 1 内部交互验证的模型预测值与实验测定值关系图

Fig. 1 Correlation between experimental values and NIR predicted values obtained by cross-validation

### 3.4 最优模型的外部验证

外部验证一般采用未参与定标,但样品性质与参与定标的样品性质相似的样品,通过比较这些未参与定标样品的预测值与实验测定值的差异来判断模型的预测准确性。预测集中蛹虫草四种有效成分的

预测值与实验测定值的相关关系图见图 2。由图 2 可见,测定多糖、蛋白、虫草酸和腺苷含量的定量分析模型的 RMSEP 分别为 0.0079,0.0196,0.0087 和 0.5780,实验结果表明,所建的校正模型具有良好的预测能力。

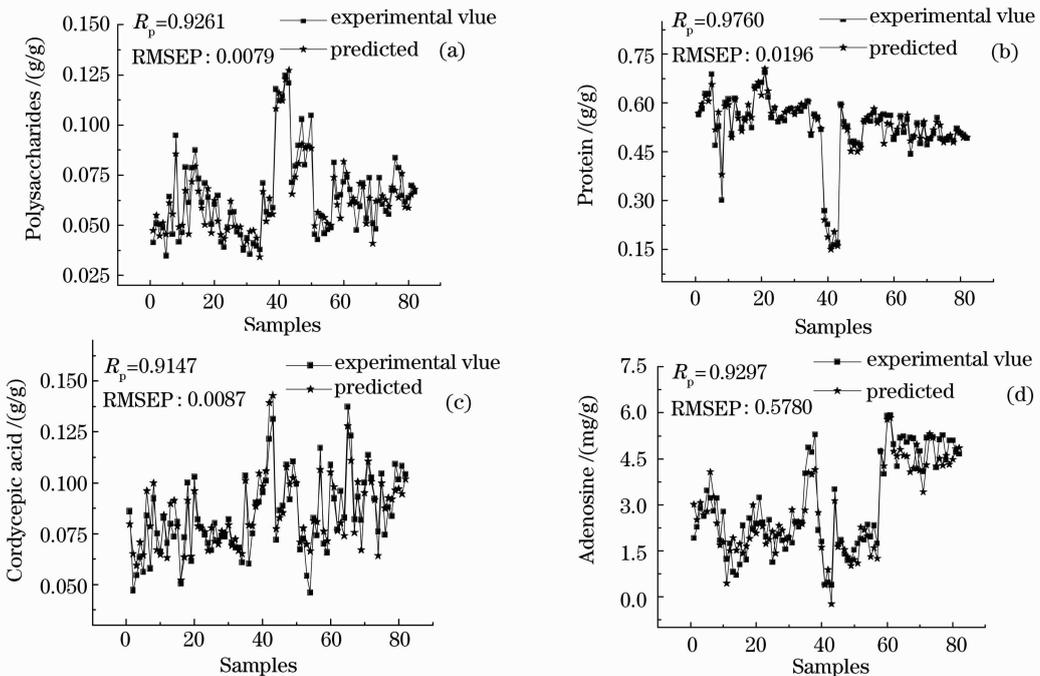


图 2 预测集中蛹虫草有效成分模型预测值与实验测定值关系图

Fig. 2 Relationship between prediction and experimental values of cordyceps militaris contents

本实验将预测集样品各组分预测值与实验测定值进行成对数据双边 t 检验(见表 3), 各组分实验

测定值和模型预测值间的 t 检验值均未达到显著水平, 表明预测值与实验测定值间无显著差异。

表 3 模型预测集结果评价

Table 3 Evaluation of the prediction models with prediction sets

Components	Samples	Actual average	Predicted average	t <sub>Values</sub>	t <sub>0.05/n=60</sub>
Polysaccharides		0.0642 g/g	0.0634 g/g	0.3990	2.000
Protein		0.5231 g/g	0.5243 g/g	0.0740	2.000
Cordycepic acid	82	0.0839 g/g	0.0850 g/g	0.3748	2.000
Adenosine		2.9997 mg/g	2.8527 mg/g	0.6308	2.000

## 4 结 论

(1) 本实验以药用真菌蛹虫草的四种活性成分为研究对象, 采用多种预处理方法对光谱进行处理。通过对预处理方法的筛选和主因子数的选择, 建立了最佳的测定蛹虫草菌丝体中多糖、蛋白、虫草酸和腺苷含量的 PLS 定量分析模型, 模型的 RMSECV 值分别为 0.0091, 0.0222, 0.0088 和 0.6520, RMSEP 值分别为 0.0079, 0.0196, 0.087 和 0.5780, 通过 t 检验, 进一步说明了预测值与实验值间差异不显著。结果表明, 所建立的最佳模型内部稳定性好, 预测精密度高。

(2) 本法无须对样品进行预处理, 可同时分析多个组分, 分析速度快, 结果准确, 具有其它传统化学分析方法所不具有的显著优点, 很适合对组成复杂的中药进行快速分析, 可用于蛹虫草的有效成分测定和质量监控。

## 参 考 文 献

- Ahn Y. J., Park S. J., Lee S. G *et al.*. Cordycepin: selective growth inhibitor derived from liquid culture of *Cordyceps militaris* against *Clostridium* spp [J]. *J. Agric. Food Chem.* 2000, **48**: 2744~2748
- Zhou X. X., Meyer C. U., Schmidtke P *et al.*. Effect of cordycepin on interleukin-10 production of human peripheral blood mononuclear cells [J]. *Eur. J. Pharmacol.* 2002, **453**: 309~317
- De Julian-Ortiz J. V., Galvez J., Munoz-Collado C *et al.*. Virtual combinatorial syntheses and computational screening of new potential anti-herpes compounds [J]. *J. Med. Chem.* 1999, **17**: 3308~3314
- Xia Min. To Determine the contents of the adenosine in cordyceps militaris link by high performance liquid chromatography [J]. *J. Huaiyin Teachers College (Natural Science Edition)*. 2004, **3**(3): 226~232  
夏敏. 高效液相色谱法测定蛹虫草中腺苷的含量[J]. 淮阴师范学院学报(自然科学版). 2004, **3**(3): 226~232
- Yu Jianliang, Li Kaixiong, Lu Shiling *et al.*. Determination of the TVB-N of mutton by KDY-9820 type automatic azotometer [J]. *Food research and development.*, 2008, **29**(2): 141~143

- 于见亮, 李开雄, 卢士玲等. 应用 KDY-9820 凯氏定氮仪测定羊肉中挥发性盐基氮 [J]. *食品研究与开发.* 2008, **29**(2): 141~143
- Wang Liming, Xia Wenshui. Determination of TPS by Improvement of Anthrone-sulfuric acid method [J]. *Food Science*. 2005, **26**(7): 185~188  
王黎明, 夏文水. 蒽酮-硫酸法测定茶多糖含量的研究 [J]. *食品科学*, 2005, **26**(7): 185~188
- Cai Youhua, Fan Wenxia, Liu Xueming *et al.*. Determination of cordycepic acid content in the mycelia of *cordyceps brasiliensis* henn by colorimetry [J]. *Modern Food Science and Technol.*, 2008, **24**(1): 76~79  
蔡友华, 范文霞, 刘学铭等. 比色法测定巴西虫草菌丝体中虫草酸的含量 [J]. *现代食品科技*, 2008, **24**(1): 76~79
- Wang Duoia, Zhou Xiangyang, Jin Tongming *et al.*. Application of near 2 infrared spectroscopy to agriculture and food analysis [J]. *Spectroscopy and Spectral Analysis*, 2004, **24**(4): 447  
王多加, 周向阳, 金同铭等. 近红外光谱检测技术在农业和食品分析上的应用 [J]. *光谱学与光谱分析*, 2004, **24**(4): 447
- Feng Fengqin, Zhang Hui, Wang Li *et al.*. Rapid detection of the quality indicators in glycerol monolaurate products using fourier transform infrared spectroscopy [J]. *Acta Optica Sinica*, 2008, **28**(2): 326~330  
冯凤琴, 张辉, 王莉等. 应用红外光谱技术快速检测月桂酸单甘油脂的品质指标 [J]. *光学学报*, 2008, **28**(2): 326~330
- Liu Fei, He Yong, Wang Li. Methods for the prediction of sugar content of rice wine using visible-near infrared spectroscopy [J]. *Acta Optica Sinica*, 2007, **27**(11): 2054~2058  
刘飞, 何勇, 王莉. 黄酒糖度预测的可见近红外光谱方法研究 [J]. *光学学报*, 2007, **27**(11): 2054~2058
- Shi Youming, Liu Gang, Liu Jianhong *et al.*. Identification of auricularia auricula from different regions by fourier transform infrared spectroscopy [J]. *Acta Optica Sinica*, 2007, **27**(1): 129~132  
时有明, 刘刚, 刘剑虹等. 不同产地黑木耳的傅里叶变换红外光谱鉴别 [J]. *光学学报*, 2007, **27**(1): 129~132
- Liu Liying, Chen Hongzhang. Prediction of maize stover components with near infrared reflectance [J]. *Spectroscopy Spectroscopy and Spectral Analysis*, 2007, **27**(2): 275  
刘丽英, 陈洪章. 玉米秸秆组分近红外漫反射光谱(NIRS)测定方法的建立 [J]. *光谱学与光谱分析*, 2007, **27**(2): 275
- Blanco M. Coello J. Iturriaga H *et al.*. NIR calibration in non-linear systems: different PLS approaches and artificial neural networks. *Chemom. Intell. Lab. Syst.* 2000, **50**(1): 75~82