

文章编号: 0253-2239(2010)12-3637-06

一种基于蒙特卡罗方法的近红外波长选择算法

洪明坚^{1, 2, 3} 温泉³ 温志渝^{1, 2}

¹ 重庆大学新型微纳器件与系统技术国家重点学科实验室, 重庆 400030
² 重庆大学微系统研究中心, 重庆 400030
³ 重庆大学软件学院, 重庆 400030

摘要 针对近红外光谱数据的特点, 分析了基于偏最小二乘法(PLS)回归系数的波长选择方法, 指出了其存在的问题, 提出了一种新的波长选择算法。将 PLS 回归系数归一化为对应波长被选择的概率, 并进行蒙特卡罗(Monte-Carlo)计算, 即对不同的随机波长组合建立一系列 PLS 模型, 预测误差最小的模型所对应的波长组合将被选择。这个过程可以在前一次波长选择的基础上重复进行, 从而形成迭代算法。采用三个近红外数据集对提出的算法进行了验证, 同时与基于 PLS 的无信息波长剔除法(UVE-PLS)和遗传算法(GA)进行了比较分析。实验结果表明, 该方法能有效地提高波长选择的准确性和稳定性, 在选择波长点个数、模型的复杂度与预测误差方面, 达到甚至优于现有算法, 具有实用价值。

关键词 光谱学; 近红外; 波长选择; 蒙特卡罗

中图分类号 O657.3 文献标识码 A doi: 10.3788/AOS20103012.3637

New Near Infrared Wavelength Selection Algorithm Based on Monte-Carlo Method

Hong Mingjian^{1,2,3} Wen Quan³ Wen Zhiyu^{1,2}

¹ Key Laboratory of Micro/Nano Devices and System Technology, Chongqing University, Chongqing 400030, China
² Microsystem Research Center of Chongqing University, Chongqing 400030, China
³ School of Software Engineering, Chongqing University, Chongqing 400030, China

Abstract Based on the feature of the (NIR) spectra, this paper analyses the method of wavelength selection using the partial least squares (PLS) regression coefficients and points out the existing problems, then proposes a new method for selecting wavelengths. It normalizes the PLS regression coefficients into the probability of the selected corresponding wavelengths, then a Monte-Carlo simulation based on the aforementioned probability is calculated. Some PLS models are constructed and evaluated using different random wavelengths combinations. The model with minimum predictive error is retained and the corresponding wavelength combinations are selected. This procedure can be iterated using the previous selected wavelengths to select fewer and fewer wavelengths. This method is tested on 3 NIR datasets and compared with the PLS-based uninformative variable elimination (UVE-PLS) and genetic algorithm (GA). Experimental results show that this method could select fewer wavelengths without sacrificing the complexity and predictive ability of the PLS model and could effectively improve the accuracy and stability of the wavelength selection.

Key words spectroscopy; near infrared; wavelength selection; Monte-Carlo

收稿日期: 2010-02-08; 收到修改稿日期: 2010-04-03

基金项目: 国家科技部国际合作项目(2007DFC00040, 2005CR0814)和国家 863 计划(2007AA042101)资助课题。

作者简介: 洪明坚(1977—), 男, 博士研究生, 主要从事光谱分析和模式识别等方面的研究。

E-mail: hongmingjian@gmail.com

导师简介: 温志渝(1949—), 男, 博士, 教授, 主要从事 MEMS 方面的研究。E-mail: wzy@cqu.edu.cn

1 引 言

在仪器和计算机技术快速发展的推动下,近红外光谱分析作为一种新兴的分析技术,在如石油、化工、农业、制药和环境监测等多个领域取得了广泛应用^[1~4]。近红外光谱包含了丰富的、微弱的构效关系信息,具有波长点多和谱带归属难以确定等特点,而且波长点之间存在严重的共线性(collinearity)关系^[5]。为了从光谱中提取尽可能多的信息,针对共线性问题,一般采用偏最小二乘(PLS)等方法进行分析和建模^[5]。显然,参与建模的变量个数越少,模型受噪声影响也越小。但是,光谱中的某些波长点,不仅对于分析和建模没有任何贡献,甚至会影响模型质量,导致更加复杂的模型及其预测能力下降。因此,研究如何剔除这些波长点对于降低模型的复杂度、减小模型的预测误差和增强模型的可靠性具有重要意义^[6]。

常用的波长选择算法可以简单地分为两类。第一类算法对各个波长点逐一考察,决定剔除与否。传统的前向选择、后向剔除和逐步回归方法都属于这一类。文献[7]提出了基于 PLS 后向波长选择(BVSPLS),对所有波长点,它每次剔除一个(leave-one-variable-out),然后建立一系列 PLS 回归模型,从中选择预测误差最小的模型,真正剔除对应的波长点,并重复上述过程,直到预测误差上升为止。也有将波长选择问题看作全局优化问题,采用随机算法进行波长选择,如遗传算法(GA)^[8,9]和退火算法^[10]等。此外,Bayes 方法也可以用于波长选择^[11,12]。第二类算法将各个波长点根据某种指标进行排序,保留前面部分的波长点或剔除最后部分的波长点。常用的指标有回归系数,协相关系数和残差等。文献[13]提出采用 PLS 回归系数的绝对值作为排序指标。文献[14]通过比较发现,PLS 回归系数是最有效的指标。文献[15]提出了基于 PLS 无信息波剔除法(UVE-PLS),它引入回归系数的稳定性作为指标,通过加入人工随机变量,剔除稳定性低于随机变量的波长点。

然而,由于近红外光谱谱峰重叠严重^[1]以及噪声的影响,用 PLS 回归系数度量波长的重要性也常常出现偏差,从而误导波长选择。本文首先分析了严格基于 PLS 回归系数直接进行波长选择存在的问题,提出一种基于蒙特卡罗(Monte-Carlo)方法的波长选择算法。它将 PLS 回归系数归一化成为波长被选择的概率,并基于这个概率用 Monte-Carlo

方法进行迭代计算,然后根据模型的复杂度和预测性能来选择波长。最后,通过三个数据集对该方法进行了实验,验证了其有效性。

2 方 法

2.1 PLS 回归

一般来说,线性回归方程具有如下形式:

$$y = X\beta + \varepsilon,$$

式中 X 是 $n \times p$ 光谱矩阵, y 一般是浓度向量, β 是回归系数, ε 是误差项, n 是样本个数, p 是波长点个数。对于近红外光谱,一般有 $n \ll p$,因此不能用常规的最小二乘方法求解。PLS 是求解上式最常用的方法之一。本质上,它将 X 分解成 $X=WRU^T$,其中 $W_{n \times a}$ 和 $U_{p \times a}$ 的列向量互为正交, $R_{a \times a}$ 为上三角矩阵。从而得到 X 的广义逆矩阵 $X^- = UR^{-1}W^T$,并有 $\beta_a = X^- y$ 。若 $\text{rank}X$ 为 X 的规则 $1 \leq a \leq \text{rank}X$ 称为主成分(PC)或因子(factor)个数,它度量了模型的复杂度,值越大越复杂。如果 a 太小, X 中的信息没有充分被用于建模,导致模型欠拟合(Under-fitting);如果 a 太大, X 中的噪声和背景等无用信息也被用于建模,导致模型过拟合(Over-fitting)。一般都采用交叉验证(CV)^[16]方法确定 a ,例如常用的留一交叉验证法(LOOCV)。但是 LOOCV 倾向于选择大的 a ,常常导致轻微的过拟合^[17]。因此,本文使用 Q. S. Xu 等^[18]提出的蒙特卡罗交叉验证(MCCV)。

2.2 回归系数分析

回归系数 β_a 是模型最重要的参数,它的每一个元素给出了对应波长点在模型中的重要性的一种近似度量。文献[13]通过 LOOCV 确定 a ,将 $|\beta_a|$ 的各个元素从大到小排序,再根据交叉验证均方根误差(RMSECV)选择前面的部分波长点。但是,这种方法存在两个问题:1) a 难以确定,因为用交叉验证选择的 a 只是适合于建立模型,并不一定适合于选择波长;2) 因为近红外光谱谱峰重叠严重以及 X 中噪声的影响,严格基于某个 β_a 可能会误导波长选择。下面以 corn 数据集来举例说明这两个问题。图 1(a)是用 MCCV 计算的 RMSECV, a 可能取 5 或 6。图 1(b)中的曲线表明 β_5 和 β_6 在 1200,1600 和 2250 nm 附近有较大的差异。图 1(c)是分别采用 β_5 和 β_6 按文献^[13]做波长选择的结果,与回归系数的大小严格一致。图 1(d)给出了波长选择后模型的预测均方根误差(RMSEP),在模型复杂度和误差方面都有差异。

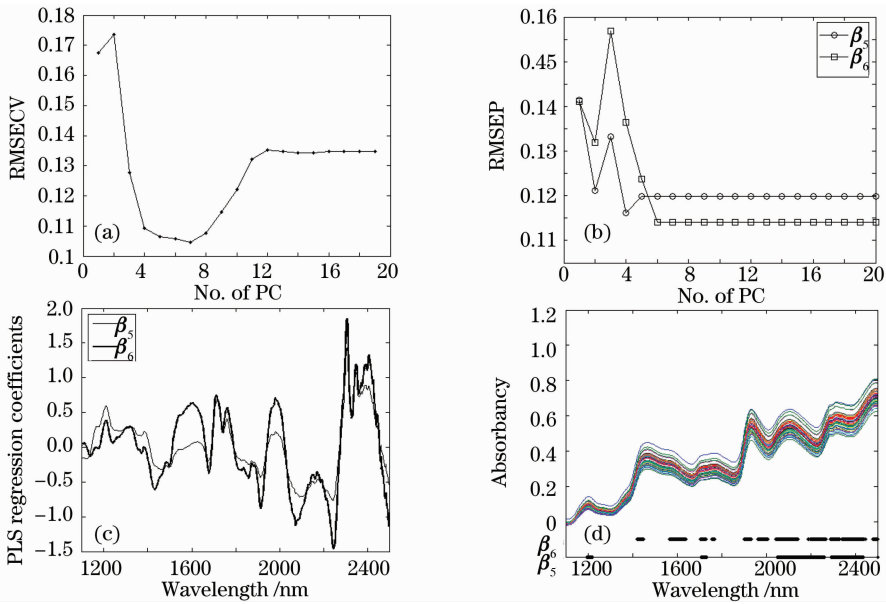


图 1 β_5 和 β_6 的 (a)RMSECV 与 (b)RMSEP 随主成份的变化, (c) PLS 消光系数与 (d)吸光度随波长的变化

Fig. 1 (a) RMSECV and (b) RMSEP of β_5 和 β_6 versus the No. of principal component (PC),

(c) PLS regression coefficient and (d) the absorbancy of β_5 和 β_6 versus wavelength

2.3 Monte-Carlo 波长选择方法

针对前面的问题,提出一种蒙特卡罗波长选择方法(MC-VS),基本思想是仍然将回归系数作为波长重要性的衡量指标。但是,考虑到 a 的选择和近红外光谱的特点,没有采用严格基于 β_a 的大小进行选择波长,而是允许一定程度的随机变化。具体做法是将 β_a 归一化成为波长被选择的概率,然后基于这个概率做一定次数的 Monte-Carlo 计算。因此系数数值越大,表示对应波长被选择的概率大,但不能保证一定会被选择;同样,系数值小,也只能表示它被选择的概率小,但是仍然有机会被选择。为了尽量选择少的波长,这个过程还可以迭代进行,在前一次迭代选出的波长的基础上,重复上述过程,直到波长个数不能再减少为止。MC-VS 的具体算法步骤:

步骤 1: $X_1 = X, i = 1$;

步骤 2: 对 X_i 和 y 做 MCCV, 选择成分个数 A_i ;

步骤 3: 建立 K 个成分的 PLS 模型, 得到回归系数 β_{A_i} ;

步骤 4: 将 β_{A_i} 归一化, 得到各个波长点被选择的

的概率 $p_{A_i} = \frac{|\beta_{A_i}|}{\max\{|\beta_{A_i,n}|\}}$;

步骤 5: 生成随机向量 $r \in R^p, r_j \in [0, 1]$;

步骤 6: 选择波长点 $\{j | p_{A_i,j} < r_j, j = 1, 2, \dots, p\}$, 建立 PLS 模型, 用测试集计算验证均方根误差(RMSEV) R_{RMSEV} ;

步骤 7: 重复步骤 5 和 6 共 N 次;

步骤 8: 从 N 个模型中选择 R_{RMSEV} 最小的一个, 记录误差为 $R_{RMSEV,i}$ 和所选波长 J_i ;

步骤 9: 如果 $R_{RMSEV,i} \geq R_{RMSEV,i-1}$ 或者 $(R_{RMSEV,i-1} - R_{RMSEV,i})R_{RMSEV,i-1} < \epsilon$,

算法结束;

步骤 10: 更新光谱矩阵, 使得 X_{i+1} 只包含 X_i 中的波长 J_i , 将 $i+1$ 赋值给 i ; 跳到 2。

算法说明: 1) 随着迭代进行, 波长个数将逐渐减少, 同时误差也逐渐下降。但是, 当波长个数继续减少的时候, 误差下降缓慢甚至会增加, 这表明波长个数太少, 已经不能满足建立模型的需要, 算法停止; 2) 绝对值最大的回归系数所对应的波长点, 一定会被选择; 3) 该算法有两个控制参数, 分别是步骤 7 中的 N 和步骤 9 中的 ϵ 。其中 N 既是 Monte-Carlo 计算的次数, 又表示了第 j 个波长点在 N 次 Monte-Carlo 计算过程中, 大约有 $N \times p_{A_i,j}$ 次被选择; ϵ 表示 RMSEV 下降幅度, 一般取 0.05, 即当 RMSEV 下降小于 5% 时, 算法停止。

3 实验部分

为了验证方法的有效性, 采用了 3 个数据集进行了实验。所有数据集的样本都采用双重(DUPLEX)方法^[19]分成训练集、测试集和验证集, 其中训练集用于建立模型和选择 A_i , 测试集用于计

算 RMSEV,以测试不同的波长组合,验证集用于计算 RMSEP,以验证波长选择后模型的预测误差。

3.1 数据集

(1)第一个数据集是玉米(corn)的近红外光谱与植物油(oil)含量。光谱采集范围 1100~2500 nm,步长 2 nm。训练集 50 个样本,测试集和验证集各 15 组样本。该数据集可以从 Eigenvector Research 公司的网站下载。

(2)汽油(gasoline)^[20]的近红外光谱和辛烷(octane)值。光谱采集范围是 900~1700 nm,步长 2 nm。前 100 个波长(即 900~1100 nm)几乎没有信息,事先把它们删除了。训练集 40 个样本,测试集和验证集各 10 组样本。

(3)葡萄糖溶液(glucose)的近红外光谱和葡萄糖含量。在相同条件下配制了 30 组已知浓度的葡萄糖溶液,用重庆大学微系统研究中心自行研制的近红外光谱仪^[21](256 pixel,光谱范围 850~1690 nm,分辨率小于 10 nm)采用透射方式采集光谱,其中只有 30~160 pixel(大约 949~1376 nm)包含信息,其他的波长点主要包含仪器噪声。同样用 DUPLEX 方法,选择 20 组作为训练集,验证和测试集各 5 组样本。

3.2 实验结果

对于 MC-VS,取参数 $N=200, \epsilon=0.05$ 。实验

选择了基于 PLS 的无信息波长剔除法(UVE-PLS)和 GA 两个典型的算法与 MC-VS 进行了比较。尽管三种方法都带有一定的随机性,每一次运行的结果都不完全相同,但是还是表现出了明显的可重复性。实验过程中,UVE-PLS 和 GA 没有用到测试集样本。

(1) corn 数据集

图 2(a)给出了三种波长选择方法所选择的波长及其个数左下角括号中。这次实验中,MC-VS 算法迭代 2 次后终止。从选择的波长点个数来说,GA 最少,只有 134 个,其次是 MC-VS,144 个,UVE-PLS 最多,291 个。可以看出,GA 和 MC-VS 所选择的波长点比较分散,覆盖了几个主要的吸收波段。而 UVE-PLS 没有选择 1200 nm 附近的波段,它对应于 C-H 键的二级倍频^[1]。图 2(b)给出了波长选择前后,PLS 模型的复杂度以及 RMSEP 的变化。波长选择前,PLS 模型的复杂度为 5, RMSEP 为 0.1221;经过 MC-VS,GA 和 UVE-PLS 三种波长选择算法后,PLS 模型的复杂度分别为 4,5 和 6, RMSEP 为 0.1204,0.1197 和 0.1127。对于这个数据集,MC-VS 的结果接近于 GA,而且降低了模型的复杂度。尽管 UVE-PLS 的 RMSEP 最小,但是它选择了最多的变量,而且模型变复杂了。

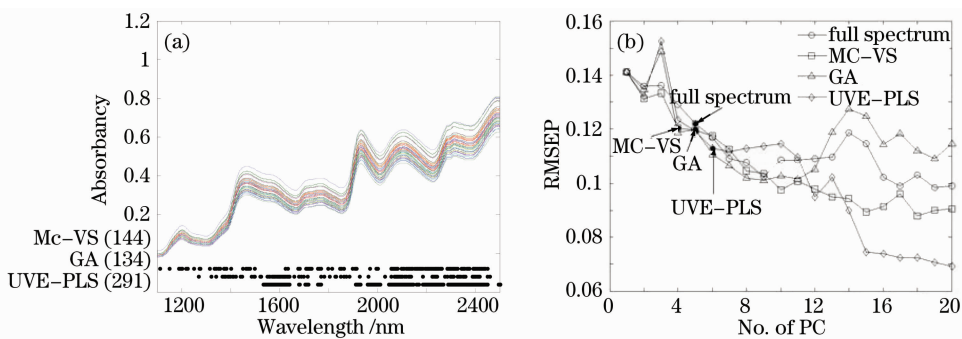


图 2 玉米集的(a)吸光度随波长变化,(b)RMSEP 随主成份数的变化

Fig. 2 Corn dataset's (a) absorbancy versus wavelength and (b) RMSEP versus the No. of principal component (PC)

(2) gasoline 数据集

图 3(a)给出了三种方法波长选择的结果。这次实验中,MC-VS 迭代 3 次停止,选择了最少的波长点,12 个,其次是 GA,68 个,UVE-PLS 最多,77 个。在模型复杂度和预测误差方面,如图 3(b)所示,经过 MC-VS,GA 和 UVE-PLS 三种波长选择方

法后,PLS 模型的复杂度从原来的 5 减少到 4, RMSEP 从原来的 0.2929 分别下降到 0.1695, 0.1671 和 0.1579。尽管 MC-VS 对应的 RMSEP 稍微大了一点,但是它选择的波长点个数明显小于另外两种方法。

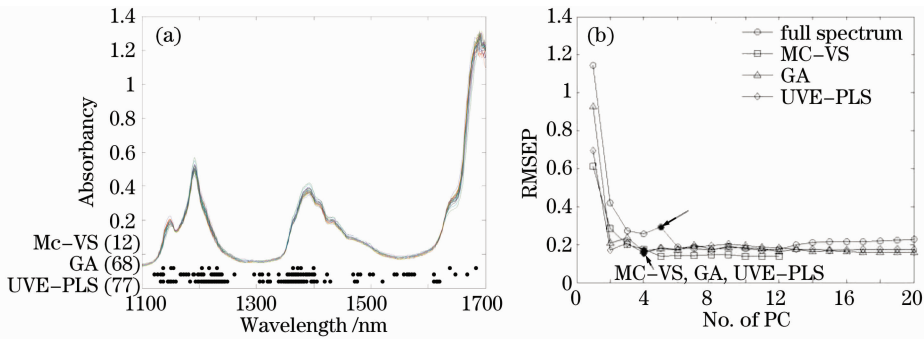


图 3 汽油集的(a)吸光度随波长变化,(b)RMSEP 随主成份数的变化

Fig. 3 Gasoline dataset's (a) absorbancy versus wavelength and (b) RMSEP versus the No. of principal component (PC)

(3) glucose 数据集

实验结果如图 4 所示。从图 4(a)的光谱曲线可以看出,光谱数据中包含了一定量的噪声。MC-VS算法迭代 2 次后终止,选择了 26 个波长点,而 GA 和 UVE-PLS 分别选择了 40 和 38 个波长点,如图 4(a)所示。图 4(b)给出了在波长选择前,PLS 模型的复杂度为 5, RMSEP 为 0.06173; MC-VS,

GA 和 UVE-PLS 波长选择后,PLS 模型的复杂度分别为 4, 5 和 4, 对应的 RMSEP 为 0.04358, 0.09826 和 0.07076。GA 和 UVE-PLS 受到了噪声的影响,所选择的波长导致了 RMSEP 反而有所增加。因此,对于这个数据集,MC-VS 不仅选择了最少的波长点,而且 RMSEP 也最小,说明了它具有比较好的稳定性。

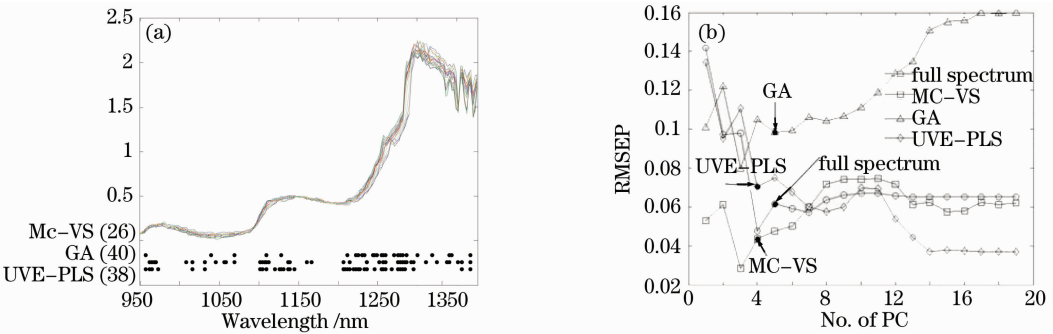


图 4 葡萄糖的(a)吸光度随波长的变化和(b)RMSEP 随主成份数的变化

Fig. 4 Glucose dataset's (a) absorbancy versus wavelength and (b) RMSEP versus the No. of principal component (PC)

3.3 分析与讨论

从 MC-VS 的算法可以看出,MC-VS 与 GA 都是在所有可能的波长组合中,随机搜索最优的波长组合。但是,它们之间有着重要的区别:GA 一开始为所有的波长点指定了一个相同的概率,然后随着迭代的进行,根据前面迭代的结果,动态更新每一个波长点被选择的概率,并在 RMSECV 的控制下选择波长。但是,这种概率指定和更新方法并没有理论上的依据。而 MC-VS 充分利用了 PLS 回归系数作为衡量波长点重要性最有效的指标这一结论,将 PLS 回归系数归一化为每一个波长点被选择的概率,然后在这个概率的指导下、在 RMSEV 的控制下进行波长选择,而且通过 Monte-Carlo 随机模拟,一定程度上消除了光谱数据中的噪声对于 PLS 回归系数的负面影响,同时保证了波长选择的准确性

和稳定性。从前面的实验结果来看,两种方法的结果比较接近,但是 MC-VS 算法更加稳定,而且比 GA-PLS 简单。而 UVE-PLS 算法通过引入人工随机变量,剔除稳定性指标低于人工变量的波长点,所以只能剔除光谱中的噪声波长点。然而,有些波长点既不是噪声,对于 y 也没有贡献,而 UVE-PLS 仍然保留了这些波长点。因此,从上面的实验结果可以看出,UVE-PLS 所选择的波长点比较多,是一种比较“保守”的波长选择方法。显然,与 GA-PLS 一样,UVE-PLS 也没有利用 PLS 回归系数来指导波长选择过程。因为 PLS 提取了光谱中与 y 最相关的信息,它的回归系数一定程度上度量了各个波长点对于 y 的贡献,所以 MC-VS 在 PLS 回归系数的指导下选择波长,所选择的波长点既不是噪声,又对 y 有一定的贡献。

4 结 论

分析了基于回归系数的近红外光谱波长选择方法,针对存在的问题,提出了一种新的波长选择方法。该方法的基本思路是将 PLS 回归系数归一化成为对应波长被选择的概率,并基于这个概率进行 Monte-Carlo 计算,并形成迭代算法。最后,通过实验验证了方法的有效性,并与两种典型的方法进行了比较。实验结果表明,该方法能够有效地克服存在的问题,提高了波长选择的准确性和稳定性,在选择波长点个数和模型的预测误差方面,达到甚至优于现有算法。

参 考 文 献

- Lu Wanzhen, Yuan Hongfu, Xu Guangtong *et al.*. Modern Analysis Techniques for Near Infrared Spectroscopy [M]. Beijing: China Petrochemical Press, 2000
- 陆婉珍, 袁洪福, 徐广通 等. 现代近红外光谱分析[M]. 北京: 中国石化出版社; 2000
- Nie Lixing, Wang Gangli, Li Zhiming *et al.*. Application of near infrared spectroscopy in process analysis of TCM manufacturing [J]. *Acta Optica Sinica*, 2009, **29**(2): 541~547
- 聂黎行, 王钢力, 李志猛 等. 近红外光谱法在中药生产过程分析中的应用[J]. *光学学报*, 2009, **29**(2): 541~547
- Zhang Haihong, Zhang Shujuan, Wang Fenghua *et al.*. Study on fast discrimination of seabuckthorn juice varieties using visible-NIR spectroscopy[J]. *Acta Optica Sinica*, 2010, **30**(2): 574~578
- 张海红, 张淑娟, 王凤花 等. 应用可见近红外光谱快速识别沙棘汁品牌[J]. *光学学报*, 2010, **30**(2): 574~578
- Cao Fang, Wu Di, He Yong *et al.*. Variety discrimination of grapes based on visible-near reflection infrared spectroscopy[J]. *Acta Optica Sinica*, 2009, **29**(2): 537~540
- 曹芳, 吴迪, 何勇 等. 基于可见-近红外反射光谱技术的葡萄品种鉴别方法的研究[J]. *光学学报*, 2009, **29**(2): 537~540
- Martens H, Naes T. Multivariate Calibration[M]. Chichester, UK: John Wiley & Sons, Inc., 1989
- Chu Xiaoli, Yuan Hongfu, Lu Wanzhen. Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique[J]. *Progress in Chemistry*, 2004, **16**(4): 528~542
- 褚小立, 袁洪福, 陆婉珍. 近红外分析中光谱预处理及波长选择方法进展与应用[J]. *化学进展*, 2004, **16**(4): 528~542
- J. A. F. Pierna, O. Abbas, V. Baeten *et al.*. A backward variable selection method for PLS regression (BVSPLS) [J]. *Analytica Chimica Acta*, 2009, **642**(1~2): 89~93
- R. Leardi, Application of genetic algorithm-PLS for feature selection in spectral data sets [J]. *Journal of Chemometrics*, 2000, **14**: 643~655
- Zou Xiaobo, Zhao Jiewen. Methods of characteristic wavelength region and wavelength selection based on genetic algorithm [J]. *Acta Optica Sinica*, 2007, **27**(7): 1316~1321
- 邹小波, 赵杰文. 用遗传算法快速提取近红外光谱特征区域和特征波长[J]. *光学学报*, 2007, **27**(7): 1316~1321
- Chen Xiaojing, Wu Di, Yu Jiajia *et al.*. A new choice method of characteristic wavelength of visible/near infrared spectroscopy [J]. *Acta Optica Sinica*, 2008, **28**(11): 2153~2158
- 陈孝敬, 吴迪, 虞佳佳 等. 一种用于可见-近红外光谱特征波长选择的新方法[J]. *光学学报*, 2008, **28**(11): 2153~2158
- T. J. Mitchell, J. J. Beauchamp. Bayesian variable selection in linear regression [J]. *Journal of the American Statistical Association*, 1988, **83**(404): 1023~1032
- Tao Chen, Elaine Martin. Bayesian linear regression and variable selection for spectroscopic calibration [J]. *Analytica Chimica Acta*, 2009, **631**: 13~21
- A. G. Frenich, D. Jouan-Rimbaud, D. L. Massart *et al.*. Wavelength selection method for multicomponent spectrophotometric determinations using partial least squares [J]. *Analyst*, 1995, **120**: 2787~2792
- R. F. Teófilo J. P. A. Martins, M. M. C., Ferreira MMC. Sorting variables by using informative vectors as a strategy for feature selection in multivariate regression [J]. *Journal of Chemometrics*, 2009, **23**(1): 32~48
- V. Centner, D. L. Massart, O. E de Noord *et al.*. Elimination of uninformative variables for multivariate calibration [J]. *Analytical Chemistry*, 1996, **68**(21): 3851~3858
- M. Stone, R. J. Brooks. Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression [J]. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1990, **52**(2): 237~269
- H. A. Martens, P. Dardenne. Validation and verification of regression in small data sets [J]. *Chemometrics and Intelligent Laboratory Systems*, 1998, **44**(1~2): 99~121
- Q. S. Xu, Y. Z. Liang. Monte Carlo cross validation [J]. *Chemometrics and Intelligent Laboratory Systems*, 2001, **56**(1): 1~11
- R. D. R. Validation of regression models: methods and examples [J]. *Technometrics*, 1977, **19**(4): 415~428
- J. H. Kalivas. Two data sets of near infrared spectra [J]. *Chemometrics and Intelligent Laboratory Systems*, 1997, **37**(2): 255~259
- Xiang Xianyi, Wen Zhiyu, Long Zaichuan *et al.*. Research on the near-infrared spectrometer system [J]. *Spectroscopy and Spectral Analysis*, 2009, **29**(8): 2286~2290
- 向贤毅, 温志渝, 龙再川 等. 微型近红外光谱仪分析系统的研制 [J]. *光谱学与光谱分析*, 2009, **29**(8): 2286~2290