

文章编号 1004-924X(2014)02-0281-08

# 基于最小二乘支持向量机的辣椒可溶性固形物和 维生素 C 含量近红外光谱检测

刘燕德\*, 周延睿, 潘圆媛

(华东交通大学 机电工程学院, 江西 南昌 330013)

**摘要:**应用傅里叶变换近红外光谱技术实现了鲜辣椒中可溶性固形物(SSC)和维生素 C(Vc)含量的快速无损检测。分别采用 7 种预处理方法对原始光谱进行处理后,建立了 SSC 和 Vc 预测的偏最小二乘法(PLS)模型。将利用最小二乘法(PLS)提取的主成分(PC)和蒙特卡罗无信息变量消除法(MC-UVE)提取的有效波长作为最小二乘支持向量机(LS-SVM)的输入变量,分别建立了 PC-LS-SVM 和 MC-UVE-LS-SVM 模型,并与 MC-UVE-PLS 模型进行了比较。采用优化后的模型对 27 个预测集未知样品进行了预测。结果表明,对鲜辣椒中 SSC 含量预测最优的为 MC-UVE-PLS 模型,其预测集相关系数( $r_p$ )为 0.971,预测集均方根误差(RMSEP)为 0.382 °Brix;对鲜辣椒中 Vc 含量预测最优的为 MC-UVE-LS-SVM 模型,其  $r_p$  为 0.899, RMSEP 为 21.022 mg/100 g。研究表明:鲜辣椒中 SSC 和 Vc 的含量与近红外光谱具有显著的相关性。

**关键词:**近红外光谱;最小二乘支持向量机;鲜辣椒;可溶性固形物;维生素 C

**中图分类号:** O657.33; S123 **文献标识码:** A **doi:** 10.3788/OPE.20142202.0281

## Determination of soluble solid contents and vitamin C of fresh peppers based on NIR spectrometry and least square support vector machines

LIU Yan-de\*, ZHOU Yan-rui, PAN Yuan-yuan

(School of Mechatronics Engineering, East China Jiaotong University, Nanchang 330013, China)

\* Corresponding author, E-mail: jxliuyd@163.com

**Abstract:** Fourier transform Near-infrared(NIR) spectroscopy was applied to the fast and nondestructive determination of Soluble Solid Contents (SSC) and Vitamin C (Vc) contents in fresh peppers. Seven kinds of pretreatment methods were used to the original spectral processing and the predicted Partial Least Square (PLS) models of SSC and Vc were established. The Principal Components (PC) selected by PLS and effective wavelengths selected by Monte Carlo Uninformative Variable Elimination (MC-UVE) method were used as the inputs of Least Square Support Vector Machine (LS-SVM), the PC-LS-SVM and MC-UVE-LS-SVM models were developed, and they were compared with the MC-UVE-PLS models. Twenty-seven unknown samples were predicted using the optimized models. The results show that MC-UVE-PLS model obtains the best result for SSC prediction with a correlation coefficient of prediction ( $r_p$ ) of 0.971 and Root Mean Square Error of Prediction (RMSEP) of

收稿日期:2013-04-09;修订日期:2013-05-13.

基金项目:科技部农业科技成果转化资金资助项目(No. 2011GB2C500008);国家 863 高技术研究发展计划资助项目(No. 2012AA101906);赣鄱英才 555 工程领军人才培养计划资助项目(No. 2011-64);赣鄱英才 555 工程高端柔性人才引进计划资助项目;江西省光电检测工程技术研究中心资助项目(赣科发财字[2012]155 号)

0.382 °Brix. The MC-UVE-LS-SVM model obtains the best result for Vc content prediction with  $r_p$  of 0.899 and RMSEP of 21.022 mg/100 g. The research results indicate that SSC and Vc contents in fresh peppers have a significant correlation with the NIR spectroscopy.

**Key words:** near infrared spectrum; Least Square Support Vector Machine (LS-SVM); fresh pepper; Soluble Solid Content (SSC); Vitamin C (Vc)

## 1 引言

辣椒是我国仅次于大白菜的第二大蔬菜作物,它营养丰富,含有大量的辣椒素、辣椒红素、胡萝卜素、碳水化合物、矿物质等,尤其是维生素 C (Vc) 的含量更是高居各类蔬菜的榜首,既可鲜食、调味,也可入药,具有重要的经济价值和食疗保健作用<sup>[1-3]</sup>。可溶性固形物 (Soluble Solids Content, SSC) 和维生素 C (Vc) 的含量测定是目前育种、农产品加工和食品营养研究中的基础问题之一。测定 SSC 含量的传统方法为蒽酮比色法或高效液相色谱分析,而测定 Vc 含量的传统方法为 2,6-二氯酚滴定法<sup>[4-7]</sup>。这些方法虽然检测精度较高,但测量时间长,操作流程繁琐,不适合现场的快速检测;而且由于一些化学物质不稳定,测量过程中会产生较大的误差。因此,研究开发一种简便、无损的 SSC 和 Vc 含量快速检测方法对监测辣椒的生长状况,管理辣椒的加工过程和提高辣椒的营养品质具有重要意义。

近红外光谱技术 (Near Infrared Reflectance Spectrum, NIRS) 因具有无损、快速、准确等特点在农业、石油工业、医药等行业得到了广泛的应用<sup>[8-10]</sup>,特别在农产品无损检测领域,已成功用于检测蔬菜的淀粉、膳食纤维、维生素、 $\beta$ 胡萝卜素、还原糖、粗蛋白、干物质、可溶性固形物、有机酸等<sup>[11-15]</sup>,其检测品种主要为大白菜、洋葱、马铃薯、西红柿等。然而,目前借助 NIRS 检测鲜辣椒中 SSC 和 Vc 含量的报道还很少。本文通过近红外光谱分析比较了不同的光谱预处理方法,并结合有效波长选择及 2 种不同的校正模型建立方法建立了鲜辣椒中 SSC 和 Vc 含量的快速无损检测模型,为检测鲜辣椒中 SSC 和 Vc 含量提供了一种无损、快速、准确的傅里叶变换近红外光谱方法。

## 2 材料与方 法

### 2.1 实验样品及分类

从南昌农贸市场购买了不同硬度、不同颜色的 115 个羊角型鲜辣椒样品,并做简单的表面清洗,在空调房 (20 °C) 放置 1 h,用于实验研究。鲜辣椒样品共 115 个,随机分为校正集、验证集及预测集。校正集用于建立鲜辣椒中 SSC 和 Vc 的含量的预测模型,验证集用于评价模型的稳健性,预测集用于评价模型的预测准确性。为保证校正模型的稳定性和适用性,SSC 和 Vc 含量最高和最低的样品被选入校正集。其中,校正集有 66 个样品,验证集有 22 个样品,预测集有 27 个样品。

### 2.2 漫反射光谱采集

鲜辣椒样品采用德国 Bruker 光学仪器公司的 Tensor37 型 FT-NIR 光谱仪及漫反射附件 (镀金积分球) 采集近红外光谱。仪器设置扫描次数为 32 次,扫描范围为 4 000~12 000  $\text{cm}^{-1}$ ,分辨率为 4  $\text{cm}^{-1}$ 。由于测量到的光谱在 4 000~4 200  $\text{cm}^{-1}$  及 10 600~12 000  $\text{cm}^{-1}$  存在较大的噪声,因此选用 4 200~10 600  $\text{cm}^{-1}$  进行分析。采用 MATLAB7.1、Origin7.0 和 Unscrambler X 10.1 化学计量学分析软件处理样品的光谱。将每一个鲜辣椒样品的中部凸表面点作为 NIR 漫透射光谱的采集点,并采集光谱 3 次,取平均值作为样品的光谱数据。

### 2.3 化学值参考测量

取对应光谱采集部位直径为 20 mm 的辣椒片,压汁并采用折射式数字糖度计 (PR-101 Cat. No3442, ATGO, Japan) 测量 SSC 含量,采用 2,6-二氯酚滴定法测定 Vc 含量。

### 2.4 数据处理及分析

为了消除光谱中无关信息和噪音的影响,本文比较了平滑 (Smoothing) 算法、一阶导数 (1st D)、二阶导数 (2nd D)、标准归一化 (SNV)、去趋

势算法(de-trending)、多元散射校正(MSC)及一阶导数+多元散射校正(1st D+MSC)7种预处理方法的优化效果。

为简化模型,增强模型的预测能力及稳健性,本文采用了蒙特卡罗无信息变量消除法(MC-UVE)提取有效波长<sup>[16-17]</sup>。该方法是在无信息变量消除法(UVE)的基础上,结合蒙特卡罗(MC)法,引入稳定性值来评价模型中每个变量的可靠性,从而决定每个变量的取舍。在 UVE 过程中,蒙特卡罗方法取代了留一法交叉验证(LOOCV)。为了比较分析,本文还采用偏最小二乘法(PLS)提取的光谱数据的主成分(PC)作为输入变量进行建模分析。MC-UVE 的算法步骤为:

(1) 利用 MC 技术从建模集中随机选取一定数目( $N_i$ )个样本建构 PLS 子模型,该步骤被重复  $M$  次,共建立  $M$  个 PLS 模型。然后,计算出 PLS 回归系数  $b_i$ ,根据变量稳定性计算公式计算得到每个变量的稳定性值  $C_i$ 。

(2) 从高到低排列所有变量的稳定性数值,设定第  $N_j$  个变量的稳定性值为阈值,低于此阈值的稳定性值所对应的变量被舍去。

(3) 根据选择出的变量,用校正集样本建立 PLS 模型,对预测集样本的相关指标进行预测。

本文同时采用了一种应用广泛的线性建模方法——偏最小二乘法(PLS)<sup>[18-19]</sup>和一种新型的统计学习方法——最小二乘支持向量机(LS-SVM)<sup>[20-21]</sup>进行建模。LS-SVM 是改进算法,其优势在于采用了等式约束,将经验风险改为二次方,并使用求解线性方程组的方法得出最优化结果,占用内存小,求解速度快。在使用 LS-SVM 建模方法时,必须选择适当的核函数。本研究采用径向基核函数(RBF):

$$K(x, y) = \exp \frac{-(x-y)^2}{2\sigma^2},$$

其中: $x$  为输入量; $y$  为  $x$  对应的目标值; $\sigma^2$  为 RBF 的参数。

在 LS-SVM 模型中,核函数参数的选择直接决定了其性能的优劣,在确定核函数后还要对核函数的参数进行优化。以 RBF 作为核函数的 LS-SVM 模型,需要调节两个参数  $\sigma^2$  (RBF 参数)和  $\gamma$  (回归误差的权重)。 $\sigma^2$  和  $\gamma$  的取值设定为

$10^{-3} \sim 10^8$ 。LS-SVM 采用基于交叉验证的网格搜索选择  $\gamma$  和  $\sigma^2$  的最优组合,即在  $\sigma^2$  和  $\gamma$  组成的空间内,利用交叉验证评价算法找到结果最好的一组参数  $\sigma^2$  和  $\gamma$ ;然后以该组参数为中心选定一个小的搜索区域再进行搜索,从而确定使 LS-SVM 结果最优的一组参数  $\sigma^2$  和  $\gamma$ 。将寻优得到的  $\sigma^2$  和  $\gamma$  输入 LS-SVM 模型,即得到鲜辣椒中 SSC 和 Vc 含量的校正模型<sup>[22-24]</sup>。

模型性能采用校正集相关系数  $r_c$  和校正集均方根误差 RMSEC,验证集或预测集的相关系数  $r_p$  和验证集或预测集的均方根误差 RMSEP 以及预测集相对分析误差 RPD<sup>[25]</sup> 作为评价指标。其中,RPD=SD/RMSEP,SD 为预测集真实值的标准偏差,RMSEP 为预测集预测值的均方根误差。如果 RPD $\geq 3$ ,说明建模效果良好,建立的模型可以用于实际检测;如果  $2.5 < \text{RPD} < 3$ ,说明利用近红外光谱技术对待测成分进行定量分析是可行的,但预测精度有待提高;如果 RPD $< 2.5$ ,则说明待测成分难以进行近红外定量检测分析,需进一步优化模型<sup>[26]</sup>。

### 3 结果与分析

#### 3.1 辣椒漫反射光谱

图 1 为 115 个鲜辣椒样本的原始近红外反射吸光度光谱图。鲜辣椒的光谱曲线趋势相似,但波段不同上升和下降的趋势也不同。鲜辣椒反射吸光度的不同说明其内部化学成分存在差异。从图中可以看出,在波长为 10 400,8 300,7 100 及

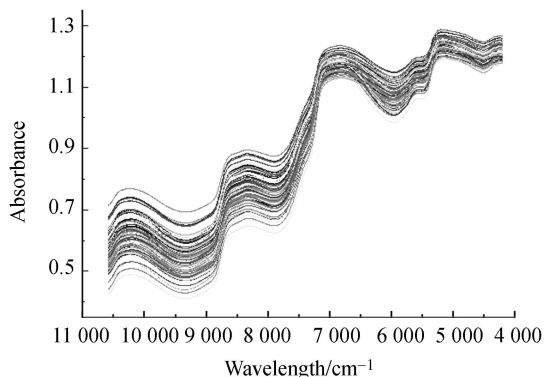


图 1 鲜辣椒的原始近红外反射吸光度光谱图  
Fig. 1 Original near infrared reflectance spectra of fresh chiles

5 300  $\text{cm}^{-1}$  附近有明显的吸收谱峰。因此,可根据这些吸收峰的位置和强度所对应的化学键,利用 NIRS 定量地描述鲜辣椒中各成分的含量。

### 3.2 化学参考值测定

鲜辣椒中 SSC 和 Vc 的统计值如表 1 所示。

表 1 鲜辣椒中可溶性固形物和维生素 C 的统计值  
Tab. 1 Statistical values of SSC and Vc in fresh chiles

测量参数	数据集	样本数	最大值	最小值	平均值	标准偏差	变异系数/%
SSC ( $^{\circ}\text{Brix}$ )	建模集	66	9.000	2.900	4.585	0.504	0.11
	验证集	22	8.100	3.067	4.326	0.389	0.09
	预测集	27	7.670	3.033	4.631	0.417	0.09
	全部	115	9.000	2.900	4.581	0.458	0.10
Vc ( $\text{mg}/100\text{ g}$ )	建模集	66	168.960	5.280	72.569	9.434	0.13
	验证集	22	158.280	9.680	71.867	10.061	0.14
	预测集	27	158.400	7.040	78.016	9.362	0.12
	全部	115	168.960	5.280	73.713	9.583	0.13

### 3.3 光谱预处理

鲜辣椒的光谱数据经过 Smoothing、1st D、2nd D、SNV、de-trending、MSC 及 1st D+MSC 7 种不同预处理后,作为 PLS 模型的输入变量,建立鲜辣椒 SSC 和 Vc 含量的预测 PLS 模型,结果分别如表 2 和表 3 所示。由表可知,最优 SSC 模型所对应的预处理方法为 1st D+MSC,其校正

鲜辣椒的 SSC 值为 2.900~9.000  $^{\circ}\text{Brix}$ ,平均值为 4.581  $^{\circ}\text{Brix}$ ;Vc 值为 5.280~168.960  $\text{mg}/100\text{ g}$ ,平均值为 73.713  $\text{mg}/100\text{ g}$ 。建模、验证及预测样本的 SSC 和 Vc 值的范围较大,这使模型具有代表性并有益于其稳定性。

集相关系数  $r_c=0.994$ ,校正集均方根误差 RMSEC=0.119  $^{\circ}\text{Brix}$ ;验证集相关系数  $r_p=0.956$ ,验证集均方根误差 RMSEP=0.309  $^{\circ}\text{Brix}$ ;最优 Vc 模型所对应的预处理方法为 SNV,其校正集相关系数  $r_c=0.949$ ,校正集均方根误差 RMSEC=10.696  $\text{mg}/100\text{ g}$ ,验证集相关系数  $r_p=0.812$ ,验证集均方根误差 RMSEP=20.189  $\text{mg}/100\text{ g}$ 。

表 2 基于 PLS 的 SSC 含量不同预处理结果  
Tab. 2 Prediction results of SSC by PLS with different pretreatment methods

光谱预处理	最佳主成分数 PCs	校正集相关系数 $r_c$	校正集均方根误差 RMSEC	验证集相关系数 $r_p$	验证集均方根误差 RMSEP
None	15	0.989	0.157	0.942	0.357
Smoothing	15	0.985	0.182	0.944	0.350
1st D	13	0.991	0.140	0.949	0.295
2nd D	10	0.985	0.185	0.887	0.498
MSC	13	0.989	0.161	0.943	0.352
SNV	14	0.991	0.140	0.949	0.336
de-trending	14	0.994	0.121	0.950	0.331
1st D+MSC	12	0.994	0.119	0.956	0.309

表 3 基于 PLS 的 V<sub>c</sub> 含量不同预处理结果

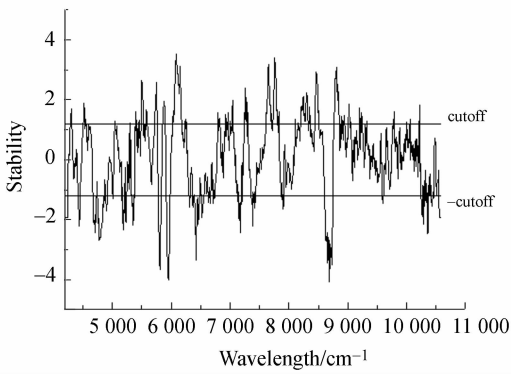
Tab. 3 Prediction results of V<sub>c</sub> by PLS with different pretreatment methods

光谱预处理	最佳主成分数 PCs	校正集相 关系数 $r_c$	校正集均 方根误差 RMSEC	验证集相 关系数 $r_p$	验证集均 方根误差 RMSEP
None	14	0.925	12.952	0.796	21.036
Smoothing	16	0.943	11.276	0.808 8	20.365
1st D	6	0.780	22.149	0.698	25.580
2nd D	7	0.847	18.509	0.703	25.385
MSC	7	0.796	21.355	0.754	23.118
SNV	12	0.949	10.696	0.812	20.189
de-trending	12	0.926	12.827	0.789	21.409
1st D+MSC	4	0.768	22.791	0.718	24.741

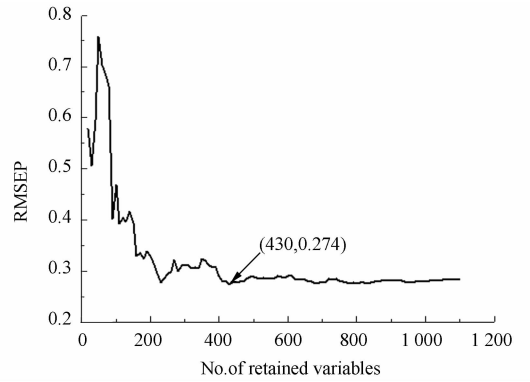
### 3.4 MC-UVE 有效波段选择

应用 MC-UVE 对上述两种最优预处理后的光谱数据进行有效波长提取。图 2(a) 和 3(b) 分别给出了波长为 4 200~10 600  $\text{cm}^{-1}$  时,鲜辣椒的 SSC 和 V<sub>c</sub> 通过 MC-UVE 方法得到的每个变

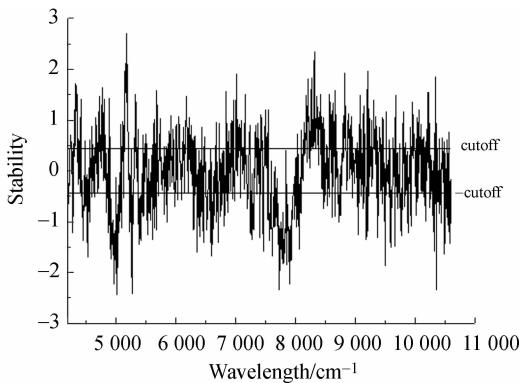
量的稳定性值。图中标示线表示变量筛选的阈值(cutoff),保留稳定性值位于标示线之外的变量用于 PLS 模型的构建。比较图 2(a) 和 2(b) 可知,当浓度指标不同时,被选择的建模变量也不同。



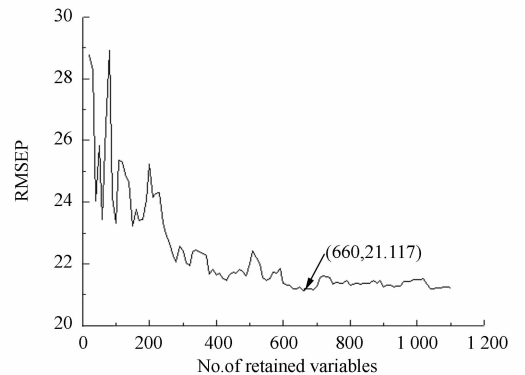
(a) 可溶性固形物  
(a) SSC



(a) 可溶性固形物  
(a) SSC



(b) 维生素 C  
(b) V<sub>c</sub>



(b) 维生素 C  
(b) V<sub>c</sub>

图 2 基于 MC-UVE 方法的预测稳定度分布图

Fig. 2 Stability distributions of variables based on MC-UVE method

图 3 所选波长下的预测均方根误差变化图

Fig. 3 Variations of RMSEP at wavelengths selected

在 MC-UVE 方法中, 阈值的选择非常重要, 直接影响模型预测的稳定性和精确度。图 3(a) 和 3(b) 给出了变量数目从 20~1 100 每隔 10 个变量计算出的 RMSEP 值。按稳定性值的绝对值从大到小选择变量, 采用每一组保留的变量建立一个 PLS 模型, 用来计算测试集的 RMSEP。从图 3(a) 可以看出, 当保留变数为 430 时, SSC 模型的 RMSEP 值最小为 0.274。从图 3(b) 可以看出, 当保留变数为 660 时, Vc 模型的 RMSEP 值最小为 20.117。

### 3.5 MC-UVE-PLS, MC-UVE-LS-SVM 和 PC-LS-SVM 模型对比

采用 1st D+MSC 及 SNV 预处理方法下 MC-UVE 提取的有效波长作为输入变量, 分别建立鲜辣椒 SSC 及 Vc 含量预测的 MC-UVE-PLS

和 MC-UVE-LS-SVM 模型。为了比较分析, 将 PLS 提取的特征向量作为 LS-SVM 模型的输入变量, 建立鲜辣椒 SSC 及 Vc 含量预测的 PC-LS-SVM 模型。

不同模型的验证结果如表 4 所示。在 MC-UVE-LS-SVM 和 PC-LS-SVM 中, 核函数为径向基函数(RBF), 模型参数( $\gamma, \sigma^2$ )的取值设定为  $10^{-3} \sim 10^8$ 。从表中可以看出, 对 SSC 而言, MC-UVE-PLS 模型的预测效果最优, 其验证集相关系数  $r_p$  为 0.987, 验证集均方根误差 RMSEP 为 0.274 °Brix; 而对于 Vc, MC-UVE-LS-SVM 模型的效果最优, 其验证集相关系数  $r_p$  为 0.911, 验证集均方根误差 RMSEP 为 19.271 mg/100 g; 此时, 模型参数的最优组合为  $\gamma = 8.4 \times 10^4, \sigma^2 = 2.2 \times 10^4$ 。

表 4 不同建模方法结果的对比

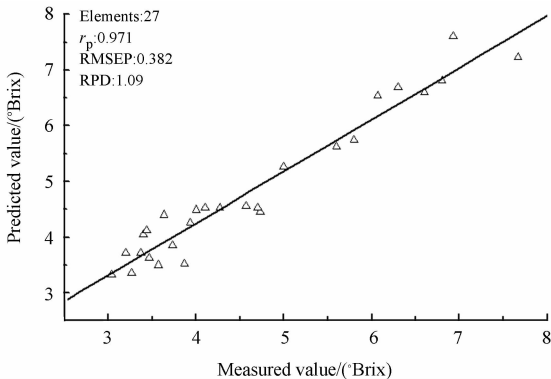
Tab. 4 Comparison of results with different modeling approaches

模型	SSC			Vc		
	最佳主成分数 PCs/( $\gamma, \sigma^2$ )	验证集相 关系数 $r_p$	验证集均 方根误差 RMSEP	最佳主成分数 PCs/( $\gamma, \sigma^2$ )	验证集相 关系数 $r_p$	验证集均 方根误差 RMSEP
MC-UVE-PLS	11	0.987	0.274	11	0.892	21.117
MC-UVE-LS-SVM	( $4.4 \times 10^3, 4.3 \times 10^4$ )	0.976	0.369	( $8.4 \times 10^4, 2.2 \times 10^4$ )	0.911	19.271
PC-LS-SVM	( $1.2 \times 10^5, 2.9 \times 10^4$ )	0.981	0.308	( $2.0 \times 10^4, 3.2 \times 10^4$ )	0.907	20.285

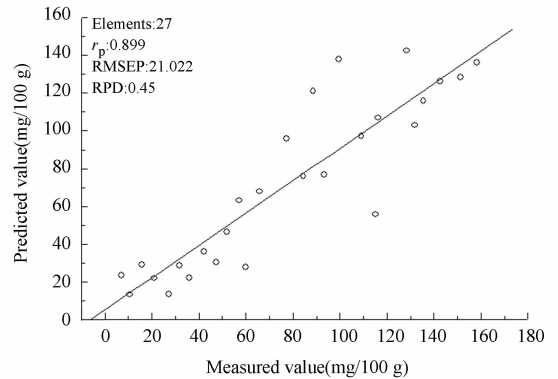
### 3.6 未知样品预测

本研究采用一阶导数结合多元散射(1st D+MSC)校正作为光谱的预处理方法, 以 MC-UVE 法确定模型的有效波段, 以 PLS 方法建立鲜辣椒中 SSC 含量的模型, 所建模型对 SSC 含量的预

测结果如图 4(a) 所示。采用 SNV 作为光谱的预处理方法, 以 MC-UVE 法确定模型的有效波段, 以 LS-SVM 方法建立鲜辣椒中 Vc 含量模型, 所建模型对 Vc 含量的预测结果如图 4(b) 所示。



(a) 可溶性固形物  
(a) SSC



(b) 维生素 C  
(b) Vc

图 4 建模效果最佳的模型预测结果

Fig. 4 Predicted results of prediction set in optimized model

## 4 结 论

本文应用傅里叶变换近红外光谱(FT-NIRS)技术实现了鲜辣椒中SSC和Vc含量的快速无损检测。通过对7种不同预处理方法、二种不同校正模型建立方法(PLS, LS-SVM)的比较,并结合MC-UVE有效波段选择方法得出:对SSC最优的预处理方法为1st D+MSC,最优的模型为MC-UVE-PLS模型;对Vc最优的预处理

方法为SNV,最优的模型为MC-UVE-LS-SVM模型。采用优化后模型对27个预测集未知样品进行了预测,预测结果为:SSC的预测集相关系数 $r_p$ 为0.971,预测集均方根误差RMSEP为0.382 $^{\circ}$ Brix,预测集相对分析误差RPD为1.09;Vc的 $r_p$ 为0.899, RMSEP为21.022 mg/100 g, RPD为0.45。虽然RPD偏小,但是鲜辣椒中SSC和Vc含量与它们的近红外光谱具有明显的相关关系。如何提高鲜辣椒中SSC和Vc含量NIR无损检测模型的实用性将是下一步研究的重点。

## 参考文献:

- [1] 方樟彩,徐国华,钱超,等. 辣椒红色素制备新工艺研究[J]. 高校化学工程学报,2012,26(1):105-111.  
FANG ZH C, XU G H, QIAN CH, *et al.*. Extraction and separation of paprika red from paprika [J]. *Journal of Chemical Engineering of Chinese Universities*, 2012, 26(1):105-111. (in Chinese)
- [2] 游国叶,陈洪轩. HPLC测定辣椒油中辣椒素含量[J]. 河南大学学报:医学版,2011,30(1):27-31.  
YOU G Y, CHEN H X. Determination of capsaicin in chili oil with HPLC [J]. *Journal of Henan University: Medical Science*, 2011, 30(1):27-31. (in Chinese)
- [3] IGNAT T, SCHMILOVITCH Z, FEFOLDI J, *et al.*. Non-destructive measurement of ascorbic acid content in bell peppers by VIS-NIR and SWIR spectrometry [J]. *Postharvest Biology and Technology*, 2012, 74: 91-99.
- [4] 张万明,王志明,陈开陆,等. 蒽酮比色法测定马铃薯淀粉深加工工艺废液总糖含量的研究[J]. 光谱实验,2010,27(2):435-440.  
ZHANG W M, WANG ZH M, CHEN K L, *et al.*. Determination of total sugar in potato starch deep-processing waste liquid using anthrone colorimetry [J]. *Chinese Journal of Spectroscopy Laboratory*, 2012, 27(2): 435-440. (in Chinese)
- [5] 李玉明. 柚子果肉中维生素C的超声波提取及高效液相色谱分析[J]. 计算机与应用化学,2011,28(5):580-582.  
LI Y M. Ultrasonic extraction of vitamin C in pomelo pulp and determination by high performance liquid chromatography [J]. *Computers and Applied Chemistry*, 2011, 28(5):580-582. (in Chinese)
- [6] 李书静,李可,姚新建,等. 2,6-二氯酚钠法测定果汁饮料中的维生素C[J]. 光谱实验室,2011,28(5):2391-2394.  
LI SH J, LI K, YAO X J, *et al.*. Determination of vitamin C in the fruit-juices by 2, 6-dichlorindophenol sodium method [J]. *Chinese Journal of Spectroscopy Laboratory*, 2011, 28(5):2391-2394. (in Chinese)
- [7] 覃方丽,闵顺耕,石正强,等. 鲜辣椒中糖份和维生素C含量的近红外光谱非破坏性测定[J]. 分析实验室,2003,33(4):59-61.  
QIN F L, MIN SH G, SHI ZH Q, *et al.*. Non-destructive determination of sugar and vitamin C in raw pepper by near-infrared spectroscopy [J]. *Chinese Journal of Analysis Laboratory*, 2003, 33(4): 59-61. (in Chinese)
- [8] 陆婉珍. 现代近红外光谱分析技术 [M]. 北京:中国石化出版社,2007.  
LU W ZH. *Modern Near Infrared Spectroscopy Analytical Technology* [M]. Beijing: China Petrochemical Press, 2007. (in Chinese)
- [9] 孙旭东,郝勇,高荣杰,等. 脐橙糖度近红外光谱在线检测数学模型优化研究[J]. 光谱学与光谱分析,2011,31(5):1230-1235.  
SUN X D, HAO Y, GAO R J, *et al.*. Research on optimization of model for detecting sugar content of navel orange by online near infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2011, 31(5):1230-1235. (in Chinese)
- [10] 刘振尧,潘涛. 可见-近红外光谱测定血红蛋白的等效波段选择[J]. 光学精密工程,2012,20(10):2170-2175.  
LIU ZH Y, PAN T. Equivalent waveband selection of VIS-NIR spectroscopic measurement for hemoglobin [J]. *Opt. Precision Eng.*, 2012, 20(10):2170-2175. (in Chinese)
- [11] PEDRO A M K, FERREIRA M M C. Simultaneously calibrating solids, sugars and acidity of tomato products using PLS2 and NIR spectroscopy [J]. *Analytica Chimica Acta*, 2007, 595: 221-227.
- [12] CLEMENT A, DORAIS M, VERNON M. Multivariate approach to the measurement of tomato maturity and gustatory attributes and their rapid assessment by Vis-NIR spectroscopy [J]. *Journal of Agricultural and Food Chemistry*, 2008, 56(5):1538-1544.
- [13] PENCHAIYA P, BOBELYN E, VERLINDEN B E, *et al.*. Non-destructive measurement of firm-

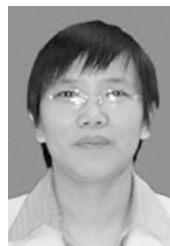
- ness and soluble solids content in bell pepper using NIR spectroscopy [J]. *Journal of Food Engineering*, 2009, 94:267-273.
- [14] SANCHEZ M T, FLORES R K, GUERRERO J E, *et al.*. Measurement of pesticide residues in peppers by near-infrared reflectance spectroscopy [J]. *Pest Management Science*, 2010, 66:580-586.
- [15] XIE L J, YING Y B, YING T J, *et al.*. Discrimination of transgenic tomatoes based on visible/near-infrared spectra [J]. *Analytica Chimica Acta*, 2007, 2(584):379-384.
- [16] 郝勇, 孙旭东, 潘圆媛, 等. 蒙特卡罗无信息变量消除方法用于近红外光谱预测果品硬度和表面色泽的研究[J]. *光谱学与光谱分析*, 2011, 31(5):1225-1229.
- HAO Y, SUN X D, PAN Y Y, *et al.*. Detection of firmness and surface color of pear by near infrared spectroscopy based on Monte Carlo uninformative variables elimination method [J]. *Spectroscopy and Spectral Analysis*, 2011, 31(5):1225-1229. (in Chinese)
- [17] CAI W S, LI Y K, SHAO X G. A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra [J]. *Chemometrics and Intelligent Laboratory Systems*, 2008, 90(2):188-194.
- [18] GELADI P, KOWALSKI B R. Partial least-squares regression: a tutorial [J]. *Analytica Chimica Acta*, 1986, 185:1-17.
- [19] 郭志明, 赵杰文, 陈全胜, 等. 特征谱区筛选在近红外光谱检测茶叶游离氨基酸含量中的应用[J]. *光学精密工程*, 2009, 17(8):1839-1844.
- GUO ZH M, ZHAO J W, CHEN Q SH, *et al.*. Application of selecting wavelength regions to determination of free amino acid content in tea by FT-NIR spectroscopy [J]. *Opt. Precision Eng.*, 2009, 17(8):1839-1844. (in Chinese)
- [20] SUYKENS J A K, GESTEL T V, BRABANTER J D, *et al.*. *Least Squares Support Vector Machines* [M]. Singapore: World Scientific, 2002.
- [21] LIU F, HE Y, WANG L. Comparison of calibrations for the determination of soluble solids content and pH of rice vinegars using visible and short-wave near infrared spectroscopy [J]. *Analytica Chimica Acta*, 2008, 610(2):196-204.
- [22] 彭彦昆, 黄慧, 王伟, 等. 基于 LS-SVM 和高光谱技术的玉米叶片叶绿素含量检测 [J]. *江苏大学学报: 自然科学版*, 2011, 32(2):125-128.
- PENG Y K, HUANG H, WANG W, *et al.*. Rapid detection of chlorophyll content in corn leaves by using least squares-support vector machines and hyperspectral images [J]. *Journal of Jiangsu University: Natural Science Edition*, 2011, 32(2):125-128. (in Chinese)
- [23] 孙旭东, 郝勇, 高荣杰, 等. 脐橙糖度近红外光谱在线检测数学模型优化研究 [J]. *光谱学与光谱分析*, 2011, 31(5):1230-1235.
- SUN X D, HAO Y, GAO R J, *et al.*. Research on optimization of model for detecting sugar content of navel range by online near infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2011, 31(5):1230-1235. (in Chinese)
- [24] 朱向荣, 李娜, 史新元, 等. 最小二乘支持向量机算法与紫外光谱法用于鉴别清开灵注射液四混中间体 [J]. *分析化学*, 2008, 36(6):770-774.
- ZHU X R, LI N, SHI X Y, *et al.*. Study on Chinese medicinal qingkailing injection intermediate by least squares support vector machines and ultraviolet spectrometry [J]. *Chinese Journal of Analytical Chemistry*, 2008, 36(6):770-774. (in Chinese)
- [25] 褚小立. *化学计量学方法与分子光谱分析技术* [M]. 北京: 化学工业出版社, 2011.
- CHU X L. *Molecular Spectroscopy Analytical Technology Combined with Chemometrics and Its Application* [M]. Beijing: Chemical Industry Press, 2011. (in Chinese)
- [26] 李晓云, 王加华, 黄亚伟, 等. 便携式近红外仪检测牛奶中脂肪、蛋白质及干物质含量 [J]. *光谱学与光谱分析*, 2011, 31(3):665-668.
- LI X Y, WANG J H, HUANG Y W, *et al.*. Determination of fat, protein and dm in raw milk by portable short-wave near infrared spectrometer [J]. *Spectroscopy and Spectral Analysis*, 2011, 31(3):665-668. (in Chinese)

#### 作者简介:



周延睿(1987-),男,山东德州人,硕士研究生,2011年于华东交通大学获学士学位,主要从事近红外光谱无损检测方面的研究。E-mail: zyr3211@163.com

#### 导师简介:



刘燕德(1967-),女,江西泰和人,博士,教授,博士生导师,2006于浙江大学获得博士学位,主要从事农产品智能无损检测技术的研究。E-mail: jxliuyd@163.com