

# 基于 Vis/NIR 光谱技术的酿酒葡萄成熟期间 SSC 预测研究

张旭<sup>1</sup>, 张天罡<sup>2</sup>, 穆维松<sup>1</sup>, 傅泽田<sup>2,3</sup>, 张小栓<sup>2,3\*</sup>

1. 中国农业大学信息与电气工程学院, 北京 100083

2. 中国农业大学工学院, 北京 100083

3. 中国农业大学食品质量与安全北京实验室, 北京 100083

**摘要** 酿酒葡萄成熟度是确定葡萄采收期的重要品质指标, 针对酿酒葡萄大田中成熟度检测难度大的问题, 利用可见/近红外(Vis/NIR)光谱技术和化学计量学, 研究了酿酒葡萄可溶性固形物含量(SSC)与光谱数据之间的内在联系。采用 USB2000+光谱仪获取 5 种酿酒葡萄及其叶片在不同成熟时期的 Vis/NIR 光谱数据, 通过 OMNIC 8.0 软件提取光谱数据, 将化学值与光谱吸收率值通过 TQ Analyst 8.0 软件建立模型。选取信噪比高的 450~1 000 nm 波段, 利用 PCA 剔除异常光谱数据, 将一阶导数(FD)、Savitzky-Golay 卷积平滑(S-G)、多元散射校正(MSC)、标准正态变换(SNV)分别组合共 4 种方法用于光谱数据预处理。利用偏最小二乘(PLS)法分别建立了 5 种葡萄基于酿酒葡萄光谱数据的 SSC 预测模型, 建立了 5 种葡萄基于冠层叶片光谱数据的 SSC 预测模型, 对比了不同方式预处理后的建模效果, 并选择最优预处理方式建模。最后用外部样本分别验证了 SSC 预测模型。结果表明, 采用 S-G 平滑+FD+MSC 的预处理方法时大多数预测模型性能达到最好。5 种葡萄浆果校正集和验证集的 R 分别达到 0.93 和 0.86 以上, 最高均方根误差分别为 0.30 和 0.48, 5 种葡萄冠层叶片校正集和验证集的 R 分别达到 0.73 和 0.65 以上, 最大均方根误差分别为 0.95 和 0.75。5 种葡萄浆果外部试验样本预测值与真实值间的平均 RE 最高为 0.43%。基于酿酒葡萄浆果光谱的 SSC 预测模型具备良好的预测能力, 优于基于酿酒葡萄冠层叶片光谱的 SSC 预测模型, SSC 预测模型能够为酿酒葡萄成熟度评价研究提供理论参考。Vis/NIR 光谱技术适用于在酿酒葡萄大田中快速、无损检测 SSC。

**关键词** 可见/近红外光谱; 酿酒葡萄成熟度; 偏最小二乘法; 可溶性固形物

**中图分类号:** O657.33 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)01-0229-07

## 引言

以优质的酿酒葡萄为原料才能酿造优质葡萄酒。酿酒葡萄的成熟度影响着葡萄的质量, 同时成熟度的判别对葡萄园安排采摘和葡萄酒厂安排生产非常重要<sup>[1]</sup>。可溶性固形物含量(soluble solids content, SSC)是测定葡萄成熟度的关键参数<sup>[2]</sup>, 决定着葡萄酒酿造的种类及工艺。通常酿酒葡萄采摘前 SSC 的检测过程是在葡萄园中均匀选取植株, 并从每穗葡萄上摘取 1~3 粒葡萄果实, 破碎、取汁后采用折光仪测定 SSC。检测过程耗时费力, 需要消耗一定葡萄原料, 同时存在取样面积大、葡萄品种多样等难点。

可见/近红外(Vis/NIR)光谱技术具有检测快速、高效等特点, 是近年来快速发展的新兴无损检测技术, 越来越多地应用于农产品品质检测和评价领域。通过全光谱或特征波长进行定量或定性分析, 对被测对象的成熟度<sup>[3]</sup>、新鲜度<sup>[4-6]</sup>、霉变度、感官特性<sup>[7]</sup>等品质指标以及颜色<sup>[8]</sup>、淀粉、SSC<sup>[9-10]</sup>等理化指标进行检测。

本研究将针对葡萄园酿酒葡萄成熟度检测中存在的问题, 结合 Vis/NIR 光谱检测技术及化学计量学方法, 在葡萄大田中获取 5 个品种酿酒葡萄及其叶片的光谱数据, 讨论不同光谱预处理算法建模效果, 确立酿酒葡萄 SSC 的最优预测模型并进行验证, 为采用 Vis/NIR 技术预测酿酒葡萄成熟度判别提供理论依据。

收稿日期: 2019-12-24, 修订日期: 2020-05-10

基金项目: 现代农业产业技术体系建设专项(CARS-29), 国家自然科学基金面上项目(31371538)资助

作者简介: 张旭, 1991年生, 中国农业大学信息与电气工程学院博士研究生 e-mail: zhxu@cau.edu.cn

\* 通讯作者 e-mail: zhxshuan@cau.edu.cn

## 1 实验部分

### 1.1 材料与仪器

实验样本 5 个品种为赤霞珠、品丽珠、梅乐、霞多丽、小芒森, 样品均取自山东省烟台市君顶葡萄酒酒庄。

光谱数据由 Vis/NIR 光谱仪 USB2000+ (美国 Ocean Optics 公司) 测得。SSC 由手持式糖度仪 PAL-1 (日本 Atago 公司) 测得。

### 1.2 光谱数据获取

葡萄样品均来自君顶葡萄酒酒庄, 于 2016 年和 2017 年先后在 4 个不同成熟阶段分别对 5 种酿酒葡萄进行光谱数据采集, 根据葡萄园大小和葡萄藤生长状况均匀选取 12 个采样位点, 如图 1 所示。

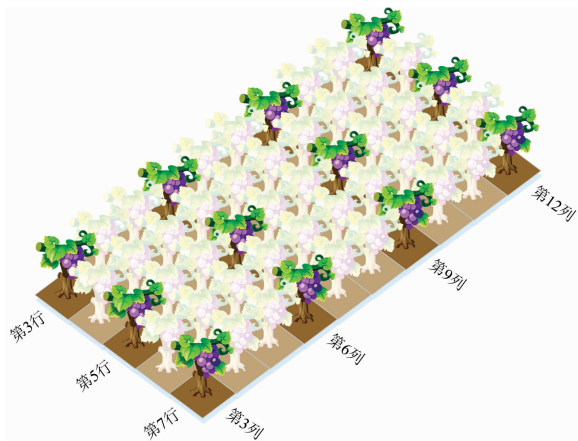


图 1 酿酒葡萄样本采集位点

Fig. 1 Sampling sites of wine grape

采集光谱数据前先用漫反射标准白板进行校正, 再采集暗电流。校正操作完成后, 设定仪器参数为: 波长范围 350~1 050 nm, 积分时间 10 ms, 扫描次数 32。随后进行样本采集, 首先选取采集位点生长状况良好的葡萄藤及葡萄果实采摘浆果样本。之后沿葡萄藤向上寻找生长良好的冠层枝叶采摘叶片样本。运送至实验室后进行光谱数据采集。采样设备为 ISP-REF 型反射用积分球, 内置卤钨灯光源, 采样孔直径为 10.32 mm。

### 1.3 光谱处理与数据分析

利用主成分分析法 (principle component analysis, PCA) 消除异常样本, 以提高酿酒葡萄可见/近红外光谱 SSC 预测模型的精确性和可靠性。

样品光谱数据往往受到随机噪声、基线漂移、样品均匀性等因素影响, 需要预处理光谱来消除这些干扰。一阶导数 (first derivative, FD) 可以降低低频噪声、放大高频噪声; Savitzky-Golay 卷积平滑 (S-G) 可有效消除基线漂移; 多元散射校正 (multiple scattering correction, MSC) 和标准正态变换 (standard normal variate, SNV) 可消除样品均匀性对漫反射光的影响。采用它们组合的 4 种不同方法对样本进行光谱预处理, 以找到最佳预处理方法。

应用偏最小二乘 (partial least squared, PLS) 方法建立校正模型。模型性能评估指标为相关系数  $R$ 、校正集均方根误差 RMSEC、验证集均方根误差 RMSEV, 其计算公式如式 (1)~式 (4) 所示

$$R_C = \sqrt{\frac{\sum_{i=1}^{n_C} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n_C} (\hat{y}_i - y_{\text{mean}})^2}} \quad (1)$$

$$R_V = \sqrt{\frac{\sum_{i=1}^{n_V} (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n_V} (\hat{y}_i - y_{\text{mean}})^2}} \quad (2)$$

$$\text{RMSEC} = \sqrt{\frac{\sum_{i=1}^{n_C} (\hat{y}_i - y_i)^2}{n_C}} \quad (3)$$

$$\text{RMSEV} = \sqrt{\frac{\sum_{i=1}^{n_V} (\hat{y}_i - y_i)^2}{n_V}} \quad (4)$$

其中,  $n_C$  及  $n_V$  分别为校正集和验证集样本个数,  $\hat{y}_i$  为第  $i$  个预测值,  $y_i$  为第  $i$  个测定值,  $y_{\text{mean}}$  为样本测定值的平均值。

将酿酒葡萄校正集样本在 400~1 000 nm 区间的 Vis/NIR 光谱数据导出, 应用 OMNIC 8.0 软件 (美国 Thermo Nicolet 公司) 进行光谱提取。光谱在 400~450 nm 处噪音明显, 若用此波段建模将降低模型的精度。因此选取 450~1 000 nm 波段的数据, 使用 TQ Analyst 8.0 (美国 Thermo Nicolet 公司) 光谱处理软件预处理光谱数据及建立校正模型。

## 2 结果与讨论

### 2.1 酿酒葡萄 Vis/NIR 光谱图

以赤霞珠葡萄果实为例, 酿酒葡萄的 Vis/NIR 吸收率原始光谱如图 2 所示, 各样本不同成熟阶段的光谱变化趋势基本一致, 呈现的吸收特征变化反映了葡萄内部不同组分信息含量的差异。光谱在小于 450 nm 和大于 1 000 nm 存在较多噪声。小于 780 nm 的吸收主要是样品在可见光区的电子跃

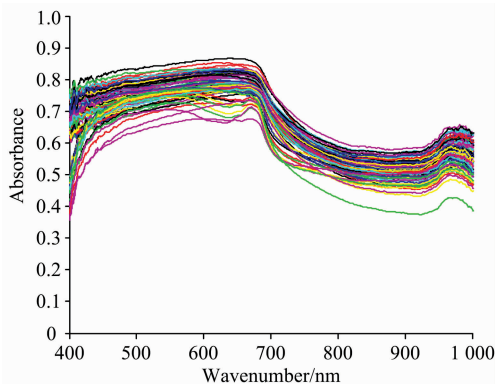


图 2 葡萄浆果样本吸收率光谱图

Fig. 2 The absorbance spectra of grape berries

迁,主要表征果皮颜色特征,超过 780 nm 属于近红外短波段。在波长 680 和 980 nm 附近均有明显吸收峰。由于赤霞珠果实的果皮呈紫黑色,对所有颜色均有吸收,因此可见光区无明显波谷。在 680 nm 附近的吸收峰可能是由果实中的叶绿素分子吸收引起的。960~1 000 nm 间的吸收峰归属于分子振动的三级倍频吸收带,主要是由水分振动引起的。

图 3 为霞多丽叶片原始光谱吸收率图,可以看出,不同生长时期的叶片样本的光谱吸收率变化趋势基本相同。在叶片光谱可见光波段,在 450 nm 附近出现蓝光波段吸收峰。在 520~570 nm 形成明显的吸收谷,由于叶绿素在绿光波段的强反射所致。在 650 nm 处为红光波段吸收峰。吸收率在 670~720 nm 波段快速下降,表现出绿色植物的红边特征。在 980 nm 附近出现水的弱吸收峰。

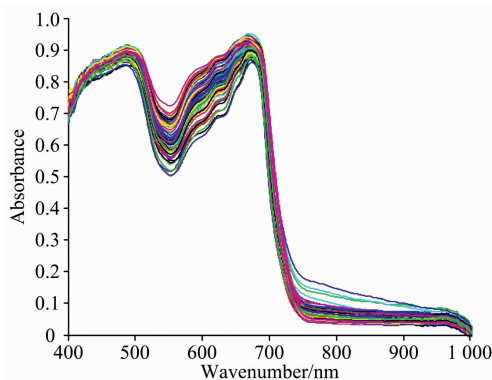


图 3 霞多丽冠层叶片样本吸收率光谱图

Fig. 3 The absorbance spectra of canopy leaves of grape

2.2 异常光谱数据剔除

采用 PCA 方法检测异常样本,对光谱的全部波段提取主成分,原始光谱信息被主成分替代,若某光谱样本得分因子与其他光谱差异大时,则认定该光谱样本为异常样本,予以剔除。如图 4 所示为霞多丽样本的第一、第二主成分的得分因子分布图,图中有 4 个样本差异较大位于置信区间以外,判定为异常样本,据此准则依次剔除 5 种酿酒葡萄浆果及其叶片的异常光谱数据。

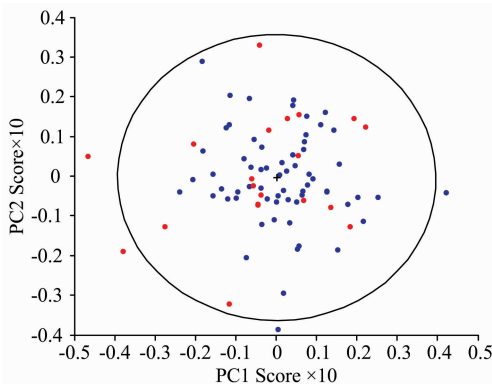


图 4 霞多丽浆果样本主成分得分图

Fig. 4 PCA score plot of Chardonnay grape

2.3 样本集划分

剔除光谱数据异常值后,SSC 数据集按校正集:验证集=3:1 的比例随机划分,然后进行建模分析。用于浆果光谱特征建模和冠层叶片光谱特征建模的酿酒葡萄 SSC 含量统计结果见表 1。

表 1 样品集 SSC 统计

Table 1 Statistics results of SSC of sample sets

样品名称	类别	样品子集	样本数量	范围/%	均值/%	标准误差/%
霞多丽	浆果	校正集	64	17.0~22.9	19.76	0.78
		验证集	21	18.4~20.8	19.72	0.66
小芒森	浆果	校正集	64	19.9~23.4	21.67	0.72
		验证集	20	20.5~22.6	21.55	0.63
梅洛	浆果	校正集	66	17.0~22.9	20.22	1.03
		验证集	22	18.3~22.2	20.50	1.00
赤霞珠	浆果	校正集	60	19.1~20.9	20.17	0.46
		验证集	20	18.8~21.3	20.05	0.64
品丽珠	浆果	校正集	66	17.2~21.3	19.67	0.80
		验证集	22	18.0~21.2	19.53	0.78
霞多丽	叶片	校正集	50	18.0~21.4	19.54	0.72
		验证集	16	18.5~21.3	19.71	0.56
小芒森	叶片	校正集	50	19.9~23.6	21.87	1.03
		验证集	17	20.8~23.0	21.81	0.56
梅洛	叶片	校正集	53	17.2~23.7	20.47	1.54
		验证集	16	19.5~22.8	20.56	0.91
赤霞珠	叶片	校正集	49	18.1~23.2	20.12	1.03
		验证集	15	18.6~21.4	19.83	0.79
品丽珠	叶片	校正集	53	18.1~21.2	19.61	0.73
		验证集	17	18.8~20.7	19.85	0.55

2.4 基于酿酒葡萄浆果光谱特征的 SSC 预测模型建立

分别采用 S-G 平滑、S-G 平滑+FD、S-G 平滑+FD+SNV、S-G 平滑+FD+MSC 等 4 种组合预处理光谱数据,并利用 PLS 进行建模分析。表 2 是 5 种酿酒葡萄浆果的光谱信息通过不同预处理后的 PLS 建模结果对比,可以看出,FD 的处理方法可以显著提高模型 R 值,降低均方根误差,求导后再使用 SNV 或 MSC 的方法可以进一步提高模型 R 值,其中赤霞珠样本光谱使用 SNV 处理时模型性能最佳,其余 4 种使用 MSC 时最佳,5 种酿酒葡萄验证集  $R_V$  均达到 0.86 以上,模型预测能力较强。

分别选取光谱数据最优预处理方法,并建立基于酿酒葡萄浆果光谱的 SSC 预测模型,图 5 至图 9 分别为霞多丽、小芒森、梅洛、赤霞珠、品丽珠的预测模型散点图,SSC 预测模型相关系数  $R$  分别达到了 0.93, 0.95, 0.96, 0.97 和 0.96,相关性很高, RMSEC 均不超过 0.30,散点图离散程度低。利用验证集样本评价 SSC 预测模型,5 种酿酒葡萄验证集  $R$  依次达到 0.86, 0.86, 0.88, 0.88 和 0.86,相关性较好,且 RMSEV 最高为 0.48,预测值与真实值离散程度低,模型效果优良。对比文献,陈辰等<sup>[2]</sup>研究中的模型性能

同样优良,但结果表现出差异性可能是由于品种以及建模方法的不同。

表 2 四种预处理方式的 PLS 预测模型比较

Table 2 Comparison of PLS prediction models with four different pretreatment methods

样品种类	预处理方法	PCS	校正集		验证集	
			$R_c$	RMSEC	$R_v$	RMSEV
霞多丽	S-G	2	0.72	0.76	0.73	0.64
	S-G+FD	2	0.78	0.49	0.83	0.40
	S-G+SNV+FD	2	0.83	0.44	0.85	0.36
	S-G+MSC+FD	3	0.93	0.30	0.86	0.36
小芒森	S-G	4	0.61	0.58	0.41	0.63
	S-G+FD	3	0.95	0.22	0.86	0.33
	S-G+SNV+FD	3	0.95	0.22	0.86	0.33
	S-G+MSC+FD	3	0.95	0.22	0.86	0.30
梅洛	S-G	7	0.73	0.70	0.76	0.70
	S-G+FD	7	1.00	0.09	0.83	0.54
	S-G+SNV+FD	4	0.96	0.30	0.88	0.48
	S-G+MSC+FD	5	0.96	0.29	0.88	0.48
赤霞珠	S-G	4	0.91	0.19	0.87	0.35
	S-G+FD	1	0.93	0.18	0.85	0.37
	S-G+SNV+FD	2	0.97	0.12	0.88	0.31
	S-G+MSC+FD	7	0.98	0.02	0.82	0.40
品丽珠	S-G	6	0.71	0.56	0.67	0.33
	S-G+FD	3	0.81	0.37	0.71	0.66
	S-G+SNV+FD	4	0.93	0.27	0.81	0.37
	S-G+MSC+FD	4	0.96	0.22	0.86	0.44

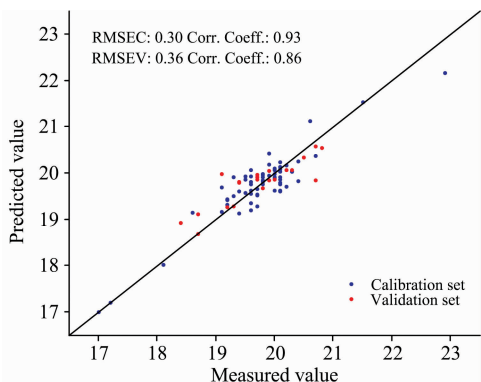


图 5 霞多丽葡萄浆果 SSC 模型校正与验证结果

Fig. 5 Results of calibration and validation of SSC prediction model for Chardonnay grape

2.5 基于酿酒葡萄叶片光谱特征的可溶性固型物(SSC)预测模型建立

表 3 是 5 种酿酒葡萄冠层叶片的样本光谱经过不同预处理后的 PLS 建模结果对比,从表 3 中可以看出,FD 的处理方法可以降低均方根误差,但对 R 的改善效果小于基于葡萄浆果光谱数据所建立的预测模型,原因可能是冠层叶片携带与 SSC 有关的信息有限,近红外波段漫反射信息采集较少,

更多的采集到了可见光部分光谱信息,而可见光波段数据较为稳定,噪声较少。求导后再使用 SNV 或 MSC 的方法可以进一步提高模型相关性,5 种葡萄叶片的预测模型均使用 MSC 处理效果最佳,验证集  $R_v$  均不低于 0.65,模型具有一定预测能力。

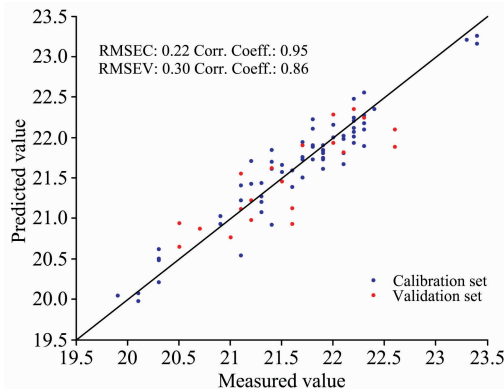


图 6 小芒森葡萄浆果 SSC 模型校正与验证结果

Fig. 6 Results of calibration and validation of SSC prediction model for Petit Manseng grape

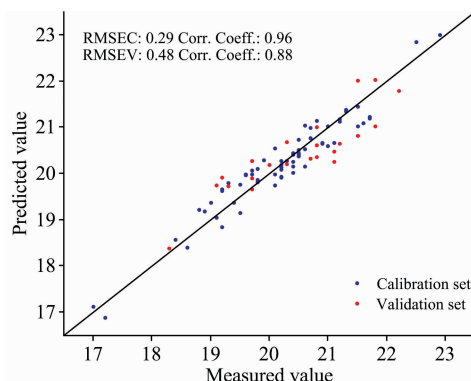


图 7 梅洛葡萄浆果 SSC 模型校正与验证结果

Fig. 7 Results of calibration and validation of SSC prediction model for Merlot grape

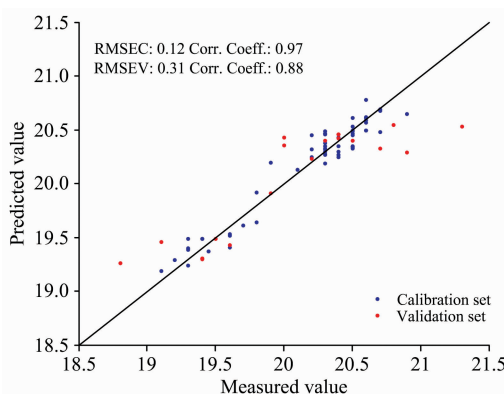


图 8 赤霞珠葡萄浆果 SSC 模型校正与验证结果

Fig. 8 Results of calibration and validation of SSC prediction model for Cabernet Sauvignon grape

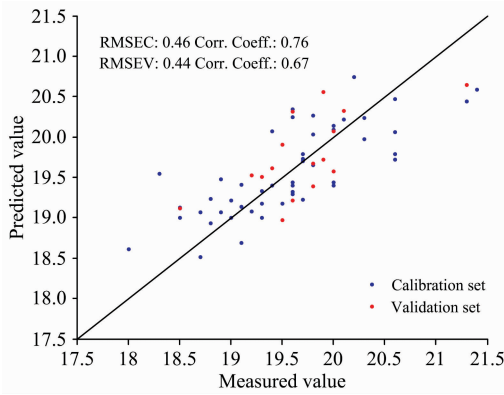


图 9 品丽珠葡萄浆果 SSC 模型校正与验证结果

Fig. 9 Results of calibration and validation of SSC prediction model for Cabernet Franc grape

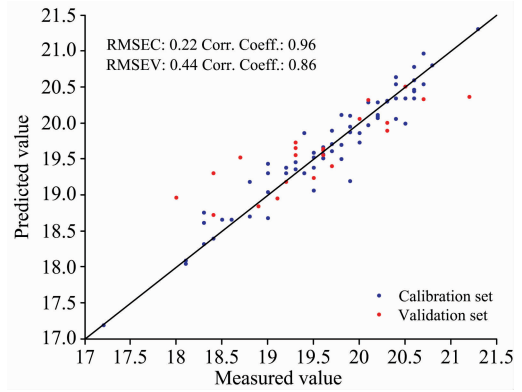


图 10 霞多丽葡萄叶片 SSC 模型校正与验证结果

Fig. 10 Results of calibration and validation of SSC prediction model for Chardonnay leaves

表 3 四种预处理方式的 PLS 预测模型比较

Table 3 Comparison of PLS prediction models with four different pretreatment methods

样品种类	预处理方法	PCS	校正集		验证集	
			$R_c$	RMSEC	$R_v$	RMSEV
霞多丽	S-G	1	0.69	0.65	0.36	0.79
	S-G+FD	1	0.70	0.52	0.55	0.68
	S-G+SNV+FD	1	0.71	0.66	0.61	0.55
	S-G+MSC+FD	1	0.76	0.46	0.67	0.44
小芒森	S-G	2	0.61	0.96	0.41	0.75
	S-G+FD	3	0.66	0.83	0.59	0.70
	S-G+SNV+FD	2	0.75	0.79	0.60	0.58
	S-G+MSC+FD	2	0.80	0.61	0.66	0.48
梅洛	S-G	1	0.62	0.70	0.52	0.97
	S-G+FD	2	0.72	0.69	0.61	0.86
	S-G+SNV+FD	2	0.72	0.90	0.61	0.80
	S-G+MSC+FD	2	0.78	0.96	0.66	0.75
赤霞珠	S-G	1	0.47	1.26	0.51	1.27
	S-G+FD	2	0.61	0.99	0.61	0.99
	S-G+SNV+FD	1	0.70	0.86	0.62	0.77
	S-G+MSC+FD	1	0.73	0.71	0.69	0.70
品丽珠	S-G	3	0.60	0.90	0.41	1.65
	S-G+FD	2	0.69	0.77	0.59	0.99
	S-G+SNV+FD	2	0.71	0.59	0.60	0.75
	S-G+MSC+FD	2	0.78	0.46	0.65	0.44

基于冠层叶片光谱数据采用最优预处理方法，分别建立了 SSC 的 PLS 预测模型，图 10 至 14 分别为霞多丽、小芒森、梅洛、赤霞珠、品丽珠叶片的预测模型散点图，SSC 预测模型  $R$  分别达到了 0.76, 0.80, 0.78, 0.73 和 0.78, 相关性较高, RMSEC 均低于 0.95, 预测值与真实值之间离散程度小。利用验证集样本对基于冠层叶片的 SSC 预测模型进行模型评价, 5 种冠层叶片验证集  $R_v$  依次达到 0.67, 0.66, 0.66, 0.69 和 0.65, 具有一定相关性, RMSEV 均在 0.75 以下, 散点图效果相对离散, 模型效果较好。

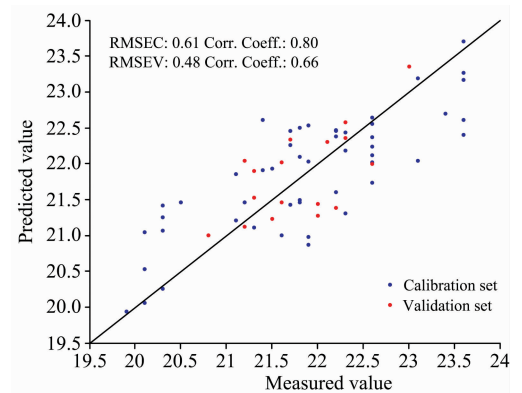


图 11 小芒森葡萄叶片 SSC 模型校正与验证结果

Fig. 11 Results of calibration and validation of SSC prediction model for Petit Manseng leaves

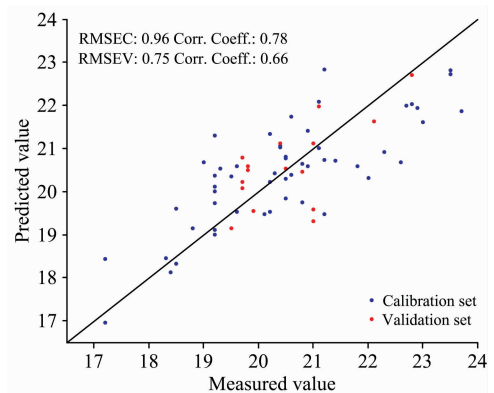


图 12 梅洛葡萄叶片 SSC 模型校正与验证结果

Fig. 12 Results of calibration and validation of SSC prediction model for Merlot leaves

2.6 外部试验验证

随机采集 5 种酿酒葡萄的浆果样本各 20 个, 采集方法与预测建模实验数据采集方法相同, 利用所建立的基于酿酒葡萄浆果光谱特征的 SSC 预测模型对外部未知样本进行

SSC 含量预测, 验证结果见表 4。从结果中可以看出, 葡萄浆果外部试验样本的 SSC 预测值与真实值之间 RE 较小, 验证样本 RE 均表现良好, 品丽珠品种的平均 RE 最高, 达到 0.43%, 小芒森和赤霞珠样本的预测结果表现优秀, RE 接

近于 0。

表 4 酿酒葡萄浆果叶片样本的外部验证结果

Table 4 External validation results of samples of berry and canopy leaves

样品名称	样品类别	数量	范围/%	绝对误差/%	相对误差/%
霞多丽	浆果	20	18.4~20.8	0.08	-0.44
小芒森	浆果	20	20.5~22.6	-0.02	-0.06
梅洛	浆果	20	19.1~22.2	-0.09	-0.35
赤霞珠	浆果	20	18.8~20.9	-0.01	-0.05
品丽珠	浆果	20	18.0~21.2	0.07	0.43

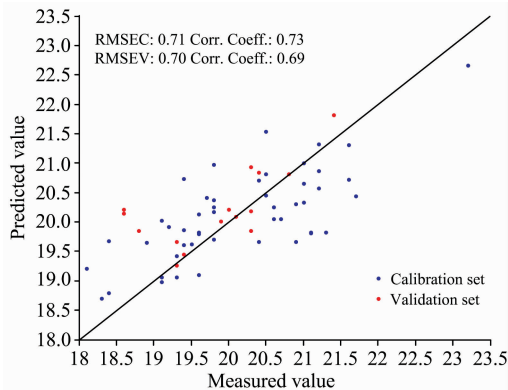


图 13 赤霞珠葡萄叶片 SSC 模型校正与验证结果

Fig. 13 Results of calibration and validation of SSC prediction model for Cabernet Sauvignon leaves

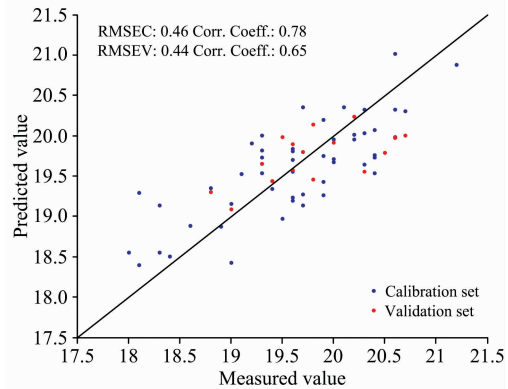


图 14 品丽珠葡萄叶片 SSC 模型校正与验证结果

Fig. 14 Results of calibration and validation of SSC prediction model for Cabernet Franc leaves

### 3 结论

采集了 5 种酿酒葡萄浆果及冠层叶片的 Vis/NIR 光谱数据, 利用 PLS 方法建立酿酒葡萄的 SSC 预测模型, 主要结论如下:

(1) 将酿酒葡萄光谱信息经过 PCA 方法剔除异常数据后, 对比四种光谱预处理方法的建模效果。结果表明大多数预测模型采用 S-G 平滑+FD+MSC 的预处理方法时效果最好。

(2) 选取最优光谱数据预处理方法, 分别建立了各品种基于酿酒葡萄浆果和冠层叶片光谱特征的 SSC 预测模型。5 种酿酒葡萄浆果的  $R_V$  最低为 0.86, 且 RMSEV 均不超过 0.48。5 种酿酒葡萄冠层叶片的  $R_V$  最高为 0.69, RMSEV 均在 0.75 以下, 模型性能均低于各品种浆果所建立模型的性能。

(3) 利用外部试验样品对模型总体预测能力进行评价。5 种葡萄浆果试验样本预测值与真实值间的平均 RE 在 0.44%~0.43% 之间。基于酿酒葡萄浆果光谱的 SSC 预测模型具备良好的预测能力, SSC 预测模型能够为酿酒葡萄成熟度评价提供理论参考。

### References

- [1] LIU Xu, CHEN Min, WU Xuan, et al(刘旭, 陈敏, 武轩, 等). Food Science(食品科学), 2016, 37(22): 230.
- [2] CHEN Chen, LU Xiao-xiang, ZHANG Peng, et al(陈辰, 鲁晓翔, 张鹏, 等). Food Science(食品科学), 2015, 36(20): 109.
- [3] Ignat T, Schmilovitch Z, Feföldi J, et al. Biosystems Engineering, 2013, 114(4): 414.
- [4] Dong X, Dong J, Li Y, et al. Computers and Electronics in Agriculture, 2019, 156: 669.
- [5] ZHANG Jing-jing, LIU Gui-shan, REN Ying-chun, et al(张晶晶, 刘贵珊, 任迎春, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2019, 39(6): 1909.
- [6] Qiao L, Tang X, Dong J. Food Chemistry, 2017, 237: 1179.
- [7] Ouyang Q, Liu Y, Chen Q, et al. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2017, 180: 91.
- [8] ZHENG Xiao-chun, LI Yong-yu, PENG Yan-kun, et al(郑晓春, 李永玉, 彭彦昆, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2016, 47(S1): 332.
- [9] Jie D, Xie L, Rao X, et al. Postharvest Biology and Technology, 2014, 90: 1.
- [10] Li M, Han D, Liu W. Biosystems Engineering, 2019, 188: 31.

# Prediction of Soluble Solids Content for Wine Grapes During Maturing Based on Visible and Near-Infrared Spectroscopy

ZHANG Xu<sup>1</sup>, ZHANG Tian-gang<sup>2</sup>, MU Wei-song<sup>1</sup>, FU Ze-tian<sup>2,3</sup>, ZHANG Xiao-shuan<sup>2,3\*</sup>

1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. College of Engineering, China Agricultural University, Beijing 100083, China

3. Beijing Laboratory of Food Quality and Safety, China Agricultural University, Beijing 100083, China

**Abstract** The maturity of wine grape is an important quality index to determine the harvest time of grape. Aiming at the problem that the maturity of wine grape is difficult to be detected in the field, the internal relationship between SSC and spectral data of wine grape was studied by Vis/NIR spectroscopy and chemometrics. The Vis/NIR spectral data of five varieties of grape and their leaves in different mature periods were obtained by USB2000+ spectrometer. The spectral data were extracted by OMNIC 8.0 software, and the chemical values and spectral absorption values were modeled by TQ analyst 8.0 software. The wave band 450~1 000 nm which had high signal-to-noise ratio was selected, and PCA was adopted to eliminate the abnormal spectral data. The first derivative (FD), Savitzky-Golay smoothing (S-G), multiple scattering correction (MSC) and standard normal variate (SNV) were combined into four methods to preprocess the spectral data. Based on the spectral data of five varieties of grape berry and the spectral data of five varieties of grape leaf, the prediction models of SSC were established by PLS. The model effects with different pretreatment methods were compared, and the optimal pretreatment method was selected for modeling. Finally, the prediction models of SSC were verified by external samples. The results show that the performance of most prediction models is the best when S-G smoothing+FD+MSC preprocessing method is applied. The correlation coefficient of calibration sets and validation sets of grape berries were above 0.93 and 0.86, respectively, and the maximum root means square error is 0.30 and 0.48, respectively. The correlation coefficient of calibration sets and validation sets of grape leaves were above 0.73 and 0.65, respectively, the maximum root mean square error is 0.95 and 0.75, respectively. The highest average relative error between the predicted value and the real value of grape berry samples was 0.43%. The SSC prediction model built by the spectra of grape berry has a good predictive ability, which is superior to the SSC prediction model built by the spectra of the grape leaf. The prediction model of SSC can provide a theoretical reference for the study of grape maturity evaluation. Therefore, Vis/NIR spectroscopy is suitable for rapid and non-destructive detection of solid soluble content in the wine grape field.

**Keywords** Visible/near infrared spectroscopy; Maturity of wine grape; Partial least square method; Soluble solids content

(Received Dec. 24, 2019; accepted May 10, 2020)

\* Corresponding author