

基于近红外光谱法的水稻秸秆可溶性糖快速检测

付苗苗 刘梅英 牛智有 杨芳 向喻

华中农业大学工学院, 武汉 430070

摘要 采集并制备不同地域、不同品种的水稻秸秆样本 288 个, 根据浓度梯度法, 按照 3:1 的比例划分校正集与验证集。采用蒽酮硫酸比色法测定试验样本中可溶性糖含量, 并采集在近红外全波段(10 000~4 000 cm^{-1}) 范围内样本的近红外光谱信息。采用多元散射校正(MSC)、标准正态变量变换(SNV)、导数、S-G 平滑及其组合方法对光谱进行预处理, 分别运用逐步多元线性回归(SMLR)、偏最小二乘回归(PLS)和主成分回归(PCR)化学计量学算法, 建立基于近红外光谱的逐步多元线性回归(SMLR)、偏最小二乘回归(PLS)和主成分回归(PCR)定量分析模型。通过比较分析, 对光谱进行一阶导数预处理, 建立的 PLS 模型效果最优, 校正集实测值与预测值之间的决定系数 R^2_c 达到 0.880 6, 交互验证决定系数(R^2_{cv})和验证集决定系数(R^2_v)分别为 0.771 1、0.857 8, 均方根差 RMSEC、RMSECV、RMSEP 分别为 0.318%、0.440%、0.404%, 校正集相对分析误差(RPD_c)和验证集相对分析误差(RPD_v)均大于 2.5。结果表明, 采用近红外光谱法建立的 PLS 模型基本可以实现水稻秸秆中可溶性糖含量的快速检测。

关键词 水稻; 秸秆; 水稻秸秆; 近红外光谱; 偏最小二乘回归; 可溶性糖; 快速检测

中图分类号 S 216.2 **文献标识码** A **文章编号** 1000-2421(2016)02-0115-07

中国水稻年产量达 2 亿 t 以上, 占世界总产量的 1/3^[1]。水稻秸秆是一种重要的生物质能源, 与石油、煤炭、天然气等不可再生能源相比, 它具有品种丰富、资源分布广、价格低、无污染和可再生等优点。现代生物质能源利用技术的开发对替代或部分替代化石能源, 保护生态环境, 实现人类社会的可持续发展具有非常重要的意义^[2]。对水稻生物质秸秆的有效合理利用, 需要大量的基础性数据。目前农作物秸秆的生物质特性基础性的数据严重匮乏, 数据使用性和代表性较差^[3]。国内外对秸秆的研究多集中在纤维素和半纤维素上, 对秸秆中可溶性糖的研究较少, 而可溶性糖是农作物秸秆中最基本的有机物之一, 关系到秸秆的饲料化、肥料化、燃料化、原料化和基料化等“五化”利用^[4]。秸秆是草食动物的粗饲料来源之一, 秸秆中所含的可溶性糖作为植物自身的主要营养物质, 不但能够为动物生长提供营养所需, 而且肥料化后它也可作为呼吸基质, 为植物的各项生命活动及合成过程提供能量; 此外, 它对植物中多种有机物的合成也具有重要影响。因此, 研

究秸秆中可溶性糖的含量对于生物质秸秆的开发利用具有重要意义。近红外光谱(NIRS)技术在生物质秸秆化学组成测定中的应用已经有报道^[5-8], 如皇才进等^[5]利用近红外光谱技术预测了秸秆中灰分、固定碳含量; 郜书静等^[7]构建了玉米秸秆体外干物质消化率、酸性洗涤纤维、中性洗涤纤维和可溶性糖含量的近红外光谱分析模型。但关于水稻秸秆可溶性糖含量方面的研究并不多。采用传统的化学方法测定水稻秸秆中的可溶性糖含量, 不但耗时、耗力、成本高, 对实验技术人员也有一定的要求。应用近红外光谱(NIRS)分析技术结合化学计量学方法, 建立水稻秸秆可溶性糖含量的快速分析模型, 探索生物质秸秆中可溶性糖的快速检测方法具有重要的现实意义。

1 材料与amp;方法

1.1 样本采集与制备

采集湖北、湖南、重庆、四川、贵州、云南等 6 个地区的水稻秸秆总计 288 个, 涵盖不同区域、不同气

收稿日期: 2015-06-30

基金项目: 国家公益性行业(农业)科研专项(201003063-04)

付苗苗, 硕士研究生, 研究方向: 农产品加工技术与装备。E-mail: Yirenxiao@webmail.hzau.edu.cn

通信作者: 牛智有, 教授, 研究方向: 农产品加工技术与装备。E-mail: nzhy@mail.hzau.edu.cn

候条件和不同品种。样本的采集方法依据《GB/T21923—2008 固体生物质燃料检验通则》要求进行,采集的样本经过晾干、粉碎等一系列规范化处理,使样本全部通过孔径 $425\ \mu\text{m}$ 筛,将研磨好的样本装入封口袋后贴上标签并编号,置于室内常温 ($25\ ^\circ\text{C}$ 左右) 下保存,供光谱扫描和可溶性糖含量的测定。

1.2 光谱采集

采用 Nicolet 公司生产的 Antaris II 型傅里叶变换近红外光谱仪(配备卤钨灯光源、漫反射积分球附

件、InGaAs 检测器)采集样本光谱,采集方式为漫反射静态采集。在使用前预热 $30\ \text{min}$ 以上,实验室环境温度在 $25\ ^\circ\text{C}$ 左右,以确保仪器的稳定性。

在近红外全波段($10\ 000\sim 4\ 000\ \text{cm}^{-1}$)范围内对样本进行光谱扫描,扫描次数 32,分辨率 $4\ \text{cm}^{-1}$ 。数据间隔设置为 $2\ \text{nm}$,最后将采集的光谱数据以吸光度的形式存储。每个样本重复装填 3 次,取 3 次测定结果的平均值为最终处理时所需的光谱数据,288 个水稻秸秆样本的原始近红外光谱见图 1 所示。

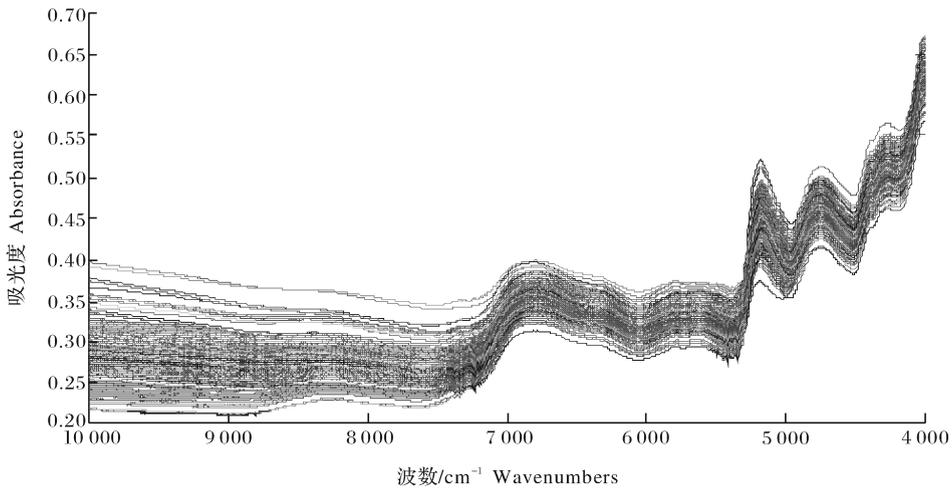


图 1 样本原始近红外漫反射光谱

Fig. 1 Samples original near infrared diffuse reflection spectra

1.3 可溶性糖测定

水稻秸秆样本中的可溶性糖含量的测定采用蒽酮硫酸比色法。

1.4 数据处理与建模

利用 TQ Analyst 6.0 软件对样本原始光谱进行 S-G 平滑、多元散射校正(MSC)、标准正态变量变换(SNV)及其组合的预处理方法,分别采用逐步多元线性回归(SMLR)、偏最小二乘法(PLS)和主成分回归法(PCR)的化学计量方法,建立水稻秸秆可溶性糖的校正模型,并对模型进行验证。

为了评价模型的定标效果,以决定系数(校正集决定系数 R_c^2 、交叉验证决定系数 R_{cv}^2 、验证集决定系数 R_v^2)、均方根差(校正集为 RMSEC、交互验证集为 RMSECV、验证集为 RMSEP)、相对分析误差(校正集为 RPD_c 、验证集为 RPD_v)和模型的性能指数(performance index)为评价指标。在优化模型时,对于同一样本集的同组分,相关系数越大、均方根差越小,说明建模所提取的光谱信息与分析组分的相关性越好,所建模型的预测能力和稳健性越

高^[9];性能指数最大值为 100,其值越大说明模型预测效果越好。同时参考文献[10-12]的方法对建模结果进行评价:若 $R^2 > 0.95$ 、 $RPD > 4$,可以认为校正模型建模非常成功;若 $0.90 \leq R^2 < 0.95$ 、 $3 \leq RPD < 4$,可以认为校正模型建模成功;若 $0.80 \leq R^2 < 0.90$ 、 $2.25 \leq RPD < 3$,可以认为校正模型建模基本成功;若 $0.70 \leq R^2 < 0.80$ 、 $1.75 \leq RPD < 2.25$,可以认为校正模型建模基本可用;若 $R^2 < 0.70$ 、 $RPD < 1.75$ 可以认为校正模型建模不可用。

2 结果与分析

2.1 可溶性糖测定及异常样本的剔除

288 个水稻秸秆样本可溶性糖含量的最小值为 1.04% ,最大值为 5.88% ,平均值为 2.45% ,标准差为 0.96% 。上述结果表明,可溶性糖的含量变化幅度较大,基本覆盖了文献中出现的值,说明样本具有很好的代表性,可以用于水稻秸秆可溶性糖的近红外光谱分析。

采用 Dixon 检验法和光谱杠杆值(leverage)分

布图来剔除异常样本。TQ Analyst 软件能够直接进行 Dixon 检验。图 2 为检验得到的马氏距离分布图,图 2 表明没有异常样本。

计算出样本可溶性糖含量的杠杆值和学学生残差

(student residual),得到样本光谱杠杆值分布如图 3 所示。设定光谱杠杆值的阈值为各样本平均杠杆值的 2 倍,大于该阈值说明光谱异常;学生残差设定为样本平均误差的 3 倍,大于该阈值说明样本化学值异常^[13]。



图 2 样本马氏距离分布

Fig. 2 Samples Mahalanobis distance distribution

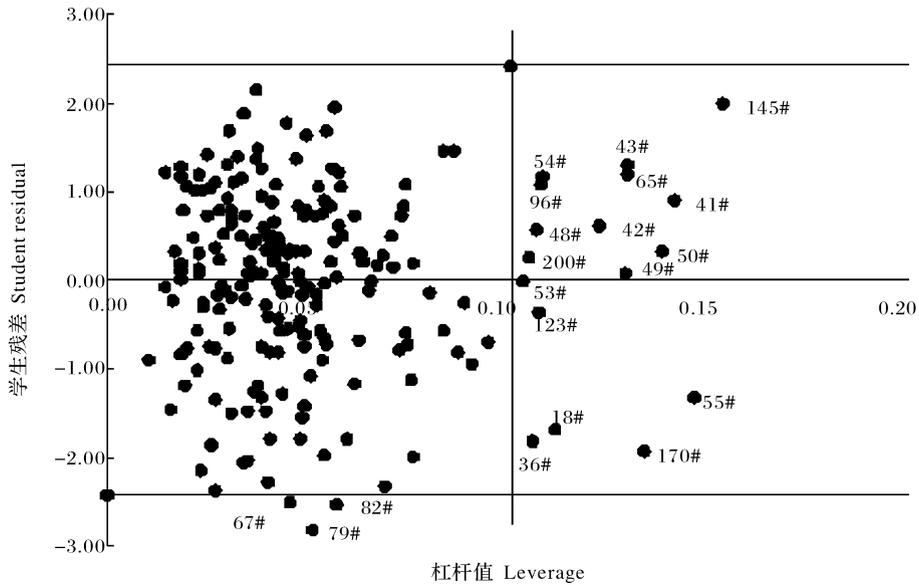


图 3 样本光谱杠杆值分布图

Fig. 3 Sample spectrum leverage value distribution

经计算得到水稻秸秆可溶性糖含量光谱杠杆值的均值为 0.050 9,故其阈值设定为 0.101 8;样本学生残差绝对值的均值为 0.809 3,则其阈值设定为 2.427 8。由图 2 可知,67#、79# 和 82# 学生残差的绝对值均大于 2.454 9,由此可以得知这 3 个样本的光谱值正常而化学值异常;18#、36#、41#、42#、43#、48#、49#、50#、53#、54#、65#、96#、123#、145#、170#、200# 这 16 个样本的光谱杠杆值均大于其阈值,可以得知他们光谱值异常而化学值正常。综上所述,以上 19 个样本均为“好”的异常样本,所以 288 个样本均参与模型的建立。

2.2 样本集的划分

根据样本可溶性糖含量化学测定值的浓度梯度分布,将样本集按 3 : 1 划分为校正集和验证集。校正集样本用来建立定标模型,而验证集样本则用来检验所建模型的预测效果。表 1 所示为校正集和验证集样本的可溶性糖含量统计结果。

2.3 模型的建立

1) 逐步多元线性回归。对样本光谱分别进行 MSC 和 SNV 2 种光程散射校正处理,结合一阶导数、二阶导数和 S-G 平滑 3 种数学处理方法及其组合对光谱进行预处理,采用 SMLR 方法,建立了水稻秸秆可溶性糖的 SMLR 定标模型,定标结果见表 2。

表 1 校正集和验证集的可溶性糖含量统计结果

Table 1 The statistical results of calibration and validation sets soluble sugar content

%

| 样本集 Sample set | 样本数 Sample number | 最大值 Maximum | 最小值 Minimum | 平均值 Average | 标准差 SD |
|---------------------|----------------------|----------------|----------------|----------------|--------|
| 校正集 Calibration set | 216 | 5.88 | 1.04 | 2.43 | 0.92 |
| 验证集 Validation set | 72 | 5.81 | 1.17 | 2.50 | 1.08 |

表 2 不同光谱预处理方法得到的 SMLR 模型建模结果

Table 2 The SMLR model modeling results of different spectral preprocessing methods

| 预处理 Pretreatment | 校正集 Calibration set | | 交互验证 Cross-validation | | 验证集 Validation set | | 性能指数 Performance index |
|--|------------------------|---------|--------------------------|----------|-----------------------|---------|------------------------------|
| | R_c^2 | RMSEC/% | R_{CV}^2 | RMSECV/% | R_v^2 | RMSEP/% | |
| 无 None | 0.328 6 | 0.753 | 0.301 5 | 0.769 | 0.461 9 | 0.827 | 60.1 |
| S-G 平滑 S-G Filter | 0.329 1 | 0.753 | 0.302 1 | 0.768 | 0.045 0 | 0.827 | 60.1 |
| 一阶导数 First derivative | 0.610 4 | 0.574 | 0.596 4 | 0.584 | 0.679 6 | 0.608 | 70.6 |
| S-G 平滑+MSC S-G filter+MSC | 0.461 3 | 0.675 | 0.437 3 | 0.690 | 0.615 8 | 0.678 | 67.3 |
| S-G 平滑+SNV S-G filter+SNV | 0.480 4 | 0.680 | 0.455 1 | 0.694 | 0.621 7 | 0.732 | 64.6 |
| S-G 平滑+一阶导数 S-G filter+First derivative | 0.613 2 | 0.572 | 0.598 1 | 0.583 | 0.682 6 | 0.607 | 70.7 |
| MSC+一阶导数 MSC+First derivative | 0.511 7 | 0.643 | 0.488 7 | 0.658 | 0.509 8 | 0.755 | 63.5 |
| SNV+一阶导数 SNV+First derivative | 0.505 8 | 0.663 | 0.486 5 | 0.674 | 0.568 2 | 0.763 | 63.1 |
| S-G 平滑+MSC+一阶导数 S-G filter+MSC+First derivative | 0.593 8 | 0.586 | 0.577 9 | 0.597 | 0.641 1 | 0.645 | 68.9 |
| S-G 平滑+SNV+一阶导数 S-G filter+SNV+First derivative | 0.587 5 | 0.629 | 0.567 0 | 0.641 | 0.627 3 | 0.723 | 65.1 |

由表 2 可知,采用不同的光程校正和数学处理方法建模时预测效果有明显差异。通过比较可知,采用 SMLR 的建模方法,对样本原始光谱进行 S-G 平滑和一阶导数组合预处理,建模效果最优。 R_c^2 和 R_{CV}^2 分别为 0.613 2 和 0.598 1, RMSEC 和 RMSECV 分别为 0.572% 和 0.583%; R_v^2 为 0.682 6, RMSEP 为 0.607%; 模型的性能指数仅为 70.7。但总体来说,采用 SMLR 建立的模型预测性能较差。

2) 偏最小二乘回归。对样本光谱分别进行 MSC 和 SNV 2 种光程散射校正处理,结合一阶导数、二阶导数和 S-G 平滑 3 种数学处理方法及其组合对光谱进行预处理,采用 PLS 法建立了水稻秸秆可溶性糖的 PLS 定标模型,定标结果见表 3。

综合比较各个评价指标,由表 3 可以看出,对光谱进行一阶导数预处理,建模效果最优。PLS 模型的校正集决定系数 R_c^2 达到 0.880 6;交互验证决定

表 3 不同光谱预处理方法得到的 PLS 模型建模结果

Table 3 The PLS model modeling results of different spectral preprocessing methods

| 预处理 Pretreatment | 校正集 Calibration set | | 交互验证 Cross-validation | | 验证集 Validation set | | 性能指数 Performance index |
|--|------------------------|---------|--------------------------|----------|-----------------------|---------|------------------------------|
| | R_c^2 | RMSEC/% | R_{CV}^2 | RMSECV/% | R_v^2 | RMSEP/% | |
| 无 None | 0.798 7 | 0.413 | 0.742 9 | 0.467 | 0.850 1 | 0.418 | 79.8 |
| S-G 平滑 S-G filter | 0.788 0 | 0.423 | 0.725 9 | 0.482 | 0.854 3 | 0.410 | 80.2 |
| 一阶导数 First derivative | 0.880 6 | 0.318 | 0.771 1 | 0.440 | 0.857 8 | 0.404 | 80.5 |
| S-G 平滑+MSC S-G filter+MSC | 0.806 2 | 0.405 | 0.747 5 | 0.463 | 0.861 9 | 0.404 | 80.5 |
| S-G 平滑+SNV S-G filter+SNV | 0.792 1 | 0.419 | 0.732 2 | 0.476 | 0.850 8 | 0.415 | 79.9 |
| S-G 平滑+一阶导数 S-G filter+First derivative | 0.852 7 | 0.353 | 0.749 8 | 0.461 | 0.825 2 | 0.449 | 78.3 |
| MSC+一阶导数 MSC+First derivative | 0.890 8 | 0.304 | 0.768 1 | 0.445 | 0.851 0 | 0.414 | 80.0 |
| SNV+一阶导数 SNV+First derivative | 0.890 4 | 0.304 | 0.785 2 | 0.427 | 0.856 7 | 0.405 | 80.4 |
| S-G 平滑+MSC+一阶导数 S-G filter+MSC+First derivative | 0.877 2 | 0.322 | 0.753 6 | 0.459 | 0.833 4 | 0.437 | 78.9 |
| S-G 平滑+SNV+一阶导数 S-G filter+SNV+First derivative | 0.882 8 | 0.315 | 0.769 1 | 0.443 | 0.842 4 | 0.425 | 79.5 |

系数 R_{CV}^2 和验证集决定系数 R_V^2 分别为 0.771 1、0.857 8,均方根差 RMSEC、RMSECV、RMSEP 分别为 0.318%、0.440%、0.404%,均在 0.5% 以下,模型的性能指数达到 80.5,校正集和验证集都表现出较好的预测效果,可见模型具有一定的预测能力。对比定标结果可以得知,采用 PLS 的建模效果比采用 SMLR 的建模效果好。

3)主成分回归。在近红外全波段范围内,对样本光谱分别进行 MSC 和 SNV 2 种光程散射校正处理,结合一阶导数、二阶导数和 S-G 平滑 3 种数学处理方法及其组合对光谱进行预处理,采用 PCR 法,建立了水稻秸秆可溶性糖的 PCR 定标模型,定标结果见表 4。

由表 4 可知,对样本原始光谱进行 S-G 平滑预处理,建模效果最优,决定系数都达到最大值, R_C^2 、

R_{CV}^2 分别为 0.665 0、0.621 9, RMSEC 和 RMSECV 分别为 0.532%、0.566%; R_V^2 为 0.735 5, RMSEP 为 0.552%,模型的性能指数为 73.3。总体来说,校正集和验证集的化学测定值与 NIRS 预测值之间都存在一定的线性关系,但是模型的预测性能一般。

2.4 模型比较分析

比较在不同的光谱预处理条件下,分别采用 SMLR、PLS、PCR 建模的 3 种最优模型的建模结果,对比结果见表 5。

由表 5 可知,采用一阶导数的光谱预处理,建立的 PLS 模型 R_C^2 最高,达到 0.880 6, R_V^2 也达到 0.857 8, RPD_V 和 RPD_V 大于 2.5,说明模型的预测值比较接近实测值,所建立的 PLS 定量分析模型预测结果比较理想,可用于定量分析。而采用 SMLR

表 4 不同光谱预处理方法得到的 PCR 模型建模结果

Table 4 The PCR model modeling results of different spectral preprocessing methods

| 预处理 Pretreatment | 校正集 Calibration set | | 交互验证 Cross-validation | | 验证集 Validation set | | 性能指数 Performance index |
|--|------------------------|---------|--------------------------|----------|-----------------------|---------|---------------------------|
| | R_C^2 | RMSEC/% | R_{CV}^2 | RMSECV/% | R_V^2 | RMSEP/% | |
| 无 None | 0.650 8 | 0.543 | 0.611 7 | 0.573 | 0.714 0 | 0.543 | 72.3 |
| S-G 平滑 S-G filter | 0.665 0 | 0.532 | 0.621 9 | 0.566 | 0.735 5 | 0.552 | 73.3 |
| 一阶导数 First derivative | 0.639 4 | 0.552 | 0.589 7 | 0.589 | 0.692 7 | 0.604 | 70.8 |
| S-G 平滑+MSC S-G filter+MSC | 0.583 5 | 0.593 | 0.526 9 | 0.633 | 0.522 6 | 0.679 | 67.2 |
| S-G 平滑+SNV S-G filter+SNV | 0.604 7 | 0.578 | 0.551 0 | 0.617 | 0.642 6 | 0.643 | 69.0 |
| S-G 平滑+一阶导数 S-G filter+First derivative | 0.637 6 | 0.554 | 0.588 0 | 0.591 | 0.691 1 | 0.607 | 70.7 |
| MSC+一阶导数 MSC+First derivative | 0.606 5 | 0.577 | 0.555 9 | 0.613 | 0.662 3 | 0.633 | 69.4 |
| SNV+一阶导数 SNV+First derivative | 0.612 0 | 0.573 | 0.562 1 | 0.609 | 0.669 8 | 0.625 | 69.8 |
| S-G 平滑+MSC+一阶导数 S-G filter+MSC+First derivative | 0.599 9 | 0.582 | 0.553 7 | 0.615 | 0.648 7 | 0.644 | 68.9 |
| S-G 平滑+SNV+一阶导数 S-G filter+SNV+First derivative | 0.604 8 | 0.578 | 0.559 1 | 0.611 | 0.654 8 | 0.638 | 69.2 |

表 5 3 种最优模型建模结果比较

Table 5 Contrast of three optimal models modeling results

| 建模方法 Modeling method | 预处理 Pretreatment | 校正集 Calibration set | | | 交互验证 Cross-validation | | 验证集 Validation set | | |
|-------------------------|-----------------------------|------------------------|---------|------------------|--------------------------|----------|-----------------------|---------|------------------|
| | | R_C^2 | RMSEC/% | RPD _C | R_{CV}^2 | RMSECV/% | R_V^2 | RMSEP/% | RPD _V |
| SMLR | S-G 平滑+一阶导数 | 0.613 2 | 0.572 | 1.608 | 0.598 1 | 0.583 | 0.682 6 | 0.607 | 1.779 |
| | S-G filter+First derivative | | | | | | | | |
| PLS | 一阶导数 | 0.880 6 | 0.318 | 2.893 | 0.771 1 | 0.440 | 0.857 8 | 0.404 | 2.673 |
| | First derivative | | | | | | | | |
| PCR | S-G 平滑 S-G filter | 0.665 0 | 0.532 | 1.729 | 0.621 9 | 0.566 | 0.735 5 | 0.552 | 1.957 |

和 PCR 建模时, R_C^2 和 R_V^2 均小于 0.7,说明化学测定值与 NIRS 预测值之间的相关性较差,所建模型不可以用于样本可溶性糖含量的定量分析。

为了进一步说明 PLS 定量分析模型的精度与预测效果,分别将模型预测值与实测值进行回归分析,得到校正集和验证集样本可溶性糖含量的实测

值与预测值相关关系(图 4)。从图 4 可以看出,样本可溶性糖含量的实测值与 PLS 模型预测值散点在回归线附近的集中程度较高,校正集和验证集实测值与预测值之间回归方程的斜率 slope 分别为 0.880 3和 0.862 3,说明预测值与实测值有较好的拟合精度。

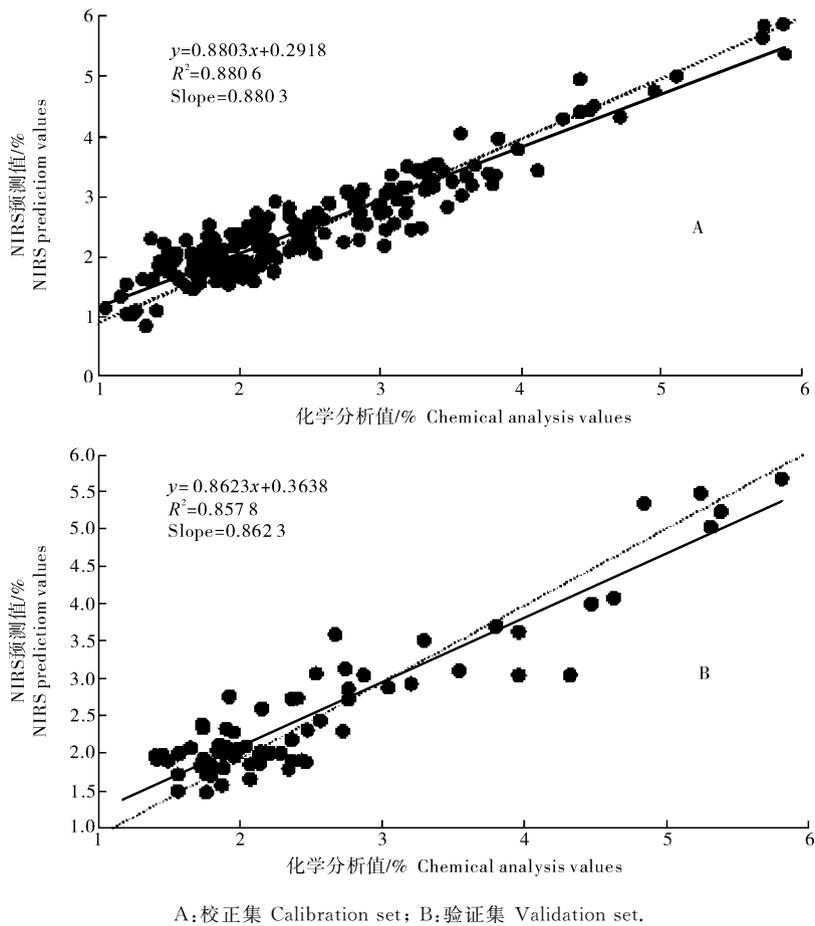


图 4 秸秆可溶性糖含量实测值与 NIRS 预测值之间的关系

Fig. 4 Correlation between the chemical analysis values and the NIRS prediction values of soluble sugar content in straw

3 讨论

采用近红外光谱分析技术,结合 SMLR、PLS、PCR 3 种化学计量学方法以及多种光谱预处理手段,分别建立水稻秸秆样品可溶性糖含量的定量分析模型。探讨应用近红外光谱技术对于水稻秸秆样品中可溶性糖含量进行定量分析的可行性。SMLR 的优点是计算速度快,同时它也能够很好地避免参与回归变量数(波长数)过多时共线性问题和过拟合问题的出现^[14]。PLS 则是目前在光谱多元校正中应用最广泛的一种化学计量学方法。它结合了主成分分析、相关性分析和多元线性回归的特点,对自变量存在多重相关的回归问题具有良好的解决能力^[15]。而 PCR 的基本思想则是对多变量的数据表进行最佳的综合简化,然后利用多元线性回归对提取的成分进行回归分析^[16]。本文利用近红外光谱分析技术,采用原始光谱进行 S-G 平滑、MSC、SNV 及其组合的预处理方法,结合 SMLR、PLS、PCR 3 种建模方法,分别建立了水稻秸秆中可溶性糖含

量的 SMLR 模型、PLS 模型和 PCR 模型。由于异常样本会对所建模型的准确度和稳健性造成严重影响,因此,在对样本光谱数据建模前,应对异常样本进行剔除^[17]。通过比较分析,对原始光谱进行一阶导数预处理方法,采用 PLS 方法建立的水稻秸秆可溶性糖含量定量分析模型具有较好的稳定性和预测效果, R_c^2 达到 0.880 6, RMSEC 为 0.318%, RPD_c 为 2.893; R_v^2 达到 0.857 8, RMSEP 为 0.404%, RPD_v 为 2.673。结果表明,利用近红外光谱分析技术实现水稻秸秆可溶性糖含量的定量分析模型是可行的。

另外,秸秆样本在保存过程受到外界环境影响,如湿度、温度等,由于秸秆粉体在环境中容易产生吸湿或解吸,最终造成秸秆可溶性糖含量发生一定的变化,导致对测试结果产生一定的影响,会导致 NIRS 定量分析模型的定标效果和预测精度受到一定影响。因此,在规范试验过程、确保测试仪器稳定和有效控制实验室环境的条件下,可以进一步提高

PLS 定量分析模型的预测精度。

参 考 文 献

- [1] 李信文,孙晓惠. 中国水稻生态化种植方法[J]. 北京农业, 2013(33):69.
- [2] 吴剑之,周肇秋,阴秀丽,等. 我国生物质能源发展现状与思考[J]. 农业机械学报, 2009, 40(1):91-99.
- [3] 谢庚楠. 基于 NIRS 的水稻和油菜秸秆干物质和可溶性糖快速测定方法及模型建立[D]. 武汉:华中农业大学图书馆, 2013.
- [4] 牛文娟,肖卫华,刘贤,等. 秸秆可溶性糖测定时样品前处理条件的优化[J]. 中国农业大学学报, 2012, 17(5):99-104.
- [5] 皇才进,韩鲁佳,刘贤,等. 基于近红外光谱技术的秸秆工业分析[J]. 光谱学与光谱分析, 2009, 29(4):960-963.
- [6] 刘贤,董苏晓,韩鲁佳,等. 青贮饲料近红外光谱分析模型转移研究[J]. 农业机械学报, 2009, 40(5):153-157.
- [7] 邵书静,张仁和,史俊通,等. 近红外光谱法测定玉米秸秆饲用品质[J]. 农业工程学报, 2009, 25(12):151-155.
- [8] HUANG C J, HAN L J, YANG Z L, et al. Exploring the use of near infrared reflectance spectroscopy to predict minerals in straw[J]. Fuel, 2009, 88(1):163-168.
- [9] 夏俊芳,李培武,李小昱,等. 不同预处理对近红外光谱检测脐

- 橙 V_c 含量的影响[J]. 农业机械学报, 2007, 38(6):107-111.
- [10] WILLIAMS P C, NORRIS K H, SOBERING D C. Determination of protein and moisture in wheat and barley by near-infrared transmission[J]. Journal of Agricultural and Food Chemistry, 1985, 33(2):239-244.
- [11] MALLEY D F, MCCLURE C, MATIN P D, et al. Compositional analysis of cattle manure during composting using a field-portable near-infrared spectrometer [J]. Communications in Soil Science and Plant Analysis, 2005, 36(4):455-475.
- [12] 李晓金,朱凯,牛智有,等. 基于 PLS 算法的生物质秸秆元素分析 NIRS 快速检测[J]. 华中农业大学学报, 2015, 34(2):131-135.
- [13] 闵顺耕,李宁,张明洋. 近红外光谱分析中异常值的判别与定量模型优化[J]. 光谱学与光谱分析, 2004, 24(10):1205-1209.
- [14] 徐惠荣. 基于可见/近红外光谱的水果糖度检测模型优化及应用研究[D]. 杭州:浙江大学图书馆, 2010.
- [15] 肖锋. 基于偏最小二乘的近红外煤质分析[D]. 杭州:浙江大学图书馆, 2013.
- [16] RAJALAHTI T, KVALHEIM O M. Multivariate data analysis in pharmaceutics: a tutorial review [J]. International Journal of Pharmaceutics, 2011, 417(1):280-290.
- [17] 牛智有. 鱼粉、精料补充料及其中肉骨粉含量的近红外漫反射光谱分析[D]. 北京:中国农业大学图书馆, 2005.

Rapidly detecting the content of soluble sugar in rice straw with near infrared reflectance spectroscopy

FU Miaomiao LIU Meiyang NIU Zhiyou YANG Fang XIANG Yu

College of Engineering, Huazhong Agricultural University, Wuhan 430070, China

Abstract 288 rice straw samples collected from different regions and varieties were used to study the feasibility of rapidly detecting the content of soluble sugar in rice straw with the near infrared reflectance spectroscopy (NIRS) technique. The near infrared spectral information of samples were collected within the near infrared wavelength range ($10\ 000 - 4\ 000\ \text{cm}^{-1}$). The models of quantitative analysis based on near infrared spectra of stepwise multiple linear regression (SMLR), partial least squares regression (PLS) and principal component regression (PCR) were established using stoichiometric algorithm SMLR, PLS and PCR, combined with different spectral pretreatments including multiplicative scatter correction (MSC), standard normal variation transformation (SNV), derivative, S-G smoothing and their combinations. Through comparison and analysis, the optimal effect of PLS model which established by using first derivative spectra pretreatment had a determination coefficient R_C^2 between the calibration set chemical analysis values and predicted values of 0.880 6, the determination coefficient R_{CV}^2 and R_V^2 of 0.771 1, 0.857 8, and root mean square difference RMSEC, RMSECV, RMSEP of 0.318%, 0.440%, 0.404%, respectively. Both of the relative analysis error RPD_C and RPD_V are greater than 2.5. The results show that establishing the model by using near infrared spectroscopy, combined with PLS modeling method can quickly detect the content of soluble sugar in rice straw.

Keywords rice; straw; rice straw; near infrared reflectance spectroscopy; partial least squares regression; soluble sugar; fast detection

(责任编辑:陆文昌)