doi:10.6041/j.issn.1000-1298.2026.13.032

鲜食玉米果穗不同部位含水率近红外光谱建模研究

韩太林[1]　王美蟠[1,2]　张永立[2,3]　孙　静[2,3]　邢　斌[4]　刘　轩[1]

(1.长春理工大学电子信息工程学院, 长春 130013; 2.农业农村部农产品产地初加工重点实验室, 北京 100125;
3.农业农村部规划设计研究院, 北京 100125; 4.北京市农林科学院信息技术研究中心, 北京 100097)

摘要: 含水率是评价鲜食玉米果穗品质的重要指标之一,影响鲜食玉米的品质分级。由于鲜食玉米果穗具有棒状且粗细不一的形态,表面籽粒呈现排布凹凸不平的特殊物理特性,对鲜食玉米果穗的近红外光谱获取和含水率预测建模分析产生了极大影响,因此有必要开展鲜食玉米果穗不同部位及点位含水率近红外光谱建模研究。首先,使用自制的近红外无损检测综合实验装置采集鲜食玉米果穗首段、中段和尾段 3 个区域的 360°光谱,每个区域内间隔 60°采集 6 个点位。其次,通过 Z-score 进行异常值剔除,结合无预处理(NONE)、标准正态变换(Standard normal variate,SNV)、多元散射校正(Multiplicative scatter correction, MSC)、一阶导数(First derivative,1D)、二阶导数(Second derivative,2D)和 Savitzky-Golay 卷积平滑(Savitzky-Golay smoothing,SG)等进行预处理。然后,采用光谱-理化值共生距离(Sample set portion based on joint x-y distance,SPXY)算法将其划分为校正集和预测集。最后,采用偏最小二乘法(Partial least squares regression,PLSR)分别建立包含所有区域数据的全局预测模型和不同区域数据的局部预测模型。进一步探究了果穗中段不同采集点位数量对预测模型的影响,并分别构建了不同采集点位数量(1、2、3、4、5、6)下的预测模型。结果显示,全局预测模型的最佳建模结果 R_p^2、RMSEP 和 RPD 分别为 0.905、0.011% 和 3.270;局部预测模型中果穗中段建模效果最佳且泛化能力更强,其最佳建模结果 R_p^2、RMSEP 和 RPD 分别为 0.955、0.007% 和 4.884;当采集点位数量为 5 时,模型预测精度达到最佳,其 R_p^2 和 R_c^2 分别为 0.967 和 0.974。研究表明,选择鲜食玉米果穗中段及 5 个采集点位的方案能够建立最佳鲜食玉米果穗含水率预测模型。

关键词: 鲜食玉米; 含水率; 近红外光谱; 定量分析模型

中图分类号: O65　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)13-0377-09　　**OSID**:

Near Infrared Spectroscopy Modeling of Moisture Content in Different Parts of Fresh Corn Ear

Han Tailin[1]　Wang Meipan[1,2]　Zhang Yongli[2,3]　Sun Jing[2,3]　Xing Bin[4]　Liu Xuan[1]

(1. *School of Electronic and Information Engineering*, *Changchun University of Science and Technology*, *Changchun* 130013, *China*
2. *Key Laboratory of Primary Processing of Agricultural Products at Origin*, *Ministry of Agriculture and Rural Affairs*, *Beijing* 100125, *China*
3. *Academy of Agricultural Planning and Engineering*, *Ministry of Agriculture and Rural Affairs*, *Beijing* 100125, *China*
4. *Information Technology Research Center*, *Beijing Academy of Agriculture and Forestry Sciences*, *Beijing* 100097, *China*)

Abstract: Water content is one of the most important indicators for evaluating the quality of fresh corn cobs, which affects the quality grading of fresh corn. Because of the special physical characteristics of fresh corn cobs, such as the rod-like shape with different thickness and the unevenness of kernel rows on the surface, the near-infrared spectroscopic (NIRS) acquisition of fresh corn cobs and the prediction of water content modeling analysis have a great impact on the quality of fresh corn cobs, and therefore it is necessary to carry out the NIRS modeling study of the water content of fresh corn cobs in different collection areas and points. Firstly, the 360° spectra of the first, middle and last regions of the cob were collected by using a homemade NIR NDT device, with 60° intervals between each region and six point locations. Secondly, the outliers were rejected by Z-score, and combined with no preprocessing

收稿日期: 2025-03-04　修回日期: 2025-03-18

基金项目: 国家重点研发计划项目(2023YFD2001301)和农业农村部重点实验室开放课题(KLAPPP2024-01)

作者简介: 韩太林(1968—),男,教授,博士,主要从事信号检测与信息处理研究,E-mail: hantl@cust.edu.cn

通信作者: 张永立(1986—),男,工程师,博士,主要从事农产品品质检测研究,E-mail: Lalizhang@163.com

(NONE), standard normal variate (SNV), multiplicative scatter correction (MSC), first order derivative (1D), second order derivative (2D) and Savitzky - Golay smoothing (SG). Then, the sample set portion based on joint x - y distance (SPXY) algorithm was used to divide it into correction set and prediction set. Finally, the partial least squares regression (PLSR) was used to establish a global prediction model containing data from all regions and a local prediction model for data from different regions, respectively. The effects of different numbers of collection sites in the middle part of the cob on the prediction model were further explored, and the prediction models were constructed under different numbers of collection sites (1, 2, 3, 4, 5 and 6), respectively. The results showed that the modeling results of the global prediction model, R_p^2, RMSEP and RPD, were 0.905, 0.011% and 3.270, respectively; the local prediction model had the best modeling effect and stronger generalization ability in the mid-section of the cob, with the modeling results, R_p^2, RMSEP and RPD, being 0.955, 0.007% and 4.884, respectively; when the number of collection sites was 5, the predictive accuracy of the model was optimal, with R_p^2 and R_c^2 values of 0.967 and 0.974, respectively. The research result showed that the scheme of choosing the mid-section of fresh maize cob and five collection sites could establish the best predictive model.

Key words: fresh corn; moisture content; near-infrared spectroscopy; quantitative analysis model

0 引言

鲜食玉米营养丰富,包括淀粉、糖、蛋白质、维生素和微量元素铁、钙等物质,且不含胆固醇[1],深受消费者喜爱,因此促进了鲜食玉米产业的发展。据统计,我国鲜食玉米种植面积已增加到 1.34×10^6 hm^2[2],位居世界前列。然而鲜食玉米收获后籽粒呼吸依然旺盛,在储存时品质易发生劣变,使鲜食玉米的市场流通受到极大限制[3]。含水率是判断鲜食玉米成熟度的重要指标[4]。此外,鲜食玉米的品质与含水率密切相关,二者呈显著正相关关系[5]。目前含水率检测用传统理化值检测方法,繁琐耗时,难以及时测定鲜食玉米含水率,且破坏样本,严重时还会对生产加工造成一定程度的经济损失[6]。因此开展鲜食玉米果穗含水率近红外光谱检测可为鲜食玉米品质保障提供支撑。

近红外光谱技术(NIRS)通过透射或者反射光谱特性反映内部组织的物理或者化学信息[7-10],在农产品内部检测应用比较广泛,如苹果、菠萝、柑橘等大众水果。研究表明,近红外光谱结合 PLSR 模型可较准确地实现水果品质参数的无损检测。其中,菠萝可溶性固形物检测模型的 R_p^2、RMSEP 分别为 0.88 和 0.010 4%,苹果和柑橘等品质参数预测结果也均表现出较好的准确性与稳定性[11-13]。果蔬在生长过程中果实的形状、组织结构、成分分布以及光学特性等方面的差异所反映的物理特性可变性都会影响光谱信息的获取和预测模型的建立[14-15]。含水率越高的样品其散射越强,导致平均光程越大,光穿透深度越小[16]。穿透深度通常用来描述光与果实相互作用的有效范围或者敏感区域,通常与光谱采集方式、结构组织和光学特性有关[17]。文献[18-19]研究结果表明,不同部位采集的光谱信息存在明显差异,因此检测部位对近红外光谱预测模型性能产生重要影响。此外,姜小刚等[21]基于 NIRS 结合 PLSR 模型建立了不同尺寸苹果糖度的通用模型,该模型降低了苹果尺寸对苹果糖度模型的影响。

鲜食玉米果穗的独特物理特征与苹果、西瓜等存在显著差异,现有技术难以适用。鲜食玉米果穗由籽粒和穗轴构成,其果穗表面呈现纵向纹理,且籽粒的排列呈现出凹凸不平的状态。此外,果穗的首段、中段和尾段在尺寸上存在差异,表现出明显的形态差异。文献[22-23]研究表明,近红外技术在玉米籽粒品质检测中具有较大潜力,但在玉米果穗检测中仍存在挑战。张永立等[24]针对鲜食玉米果穗中部,研究了近红外光谱无损检测影响因素,匹配出最佳采集参数。然而,目前缺乏针对鲜食玉米果穗不同部位含水率的近红外光谱建模研究,对其进行深入研究可促进鲜食玉米果穗品质无损检测技术发展。

本文基于近红外光谱技术,结合 PLSR 建模方法,以鲜食糯玉米果穗作为研究对象,系统构建鲜食玉米果穗含水率的全局预测模型和不同部位局部预测模型。深入探讨鲜食玉米近红外光谱检测技术的关键影响因素,并确定最佳的采集部位及采集点数量,以期为鲜食玉米果穗的近红外光谱检测提供理论依据和技术支持。

1 材料与方法

1.1 实验材料

选用万糯 2000 玉米作为研究对象,于 2024 年 11 月 8 日清晨采摘并运送至实验室。在带苞叶的

状态下，样本在室温（25℃）下静置 2 h 后，人工剥开苞叶，去除玉米须和穗柄，选取大小均匀、籽粒饱满且无病虫害的玉米果穗 90 根，进行编号作为样本。

1.2　实验方法

1.2.1　数据获取

实验采用自行搭建的漫反射式近红外光谱无损检测装置对鲜食玉米样品进行光谱测量。实验在室温（25℃）下进行，测量前将光谱仪预热 30 min，光谱数据由 SPECview_7_1 软件采集。为避免环境干扰，测量时采用标准漫反射板（型号 WS－60，材料为聚四氟乙烯）作为背景以放置样本。采集装置由卤素灯光源、暗室、近红外模块、光纤探头和支架组成。本次实验共选取 90 个样品，玉米果穗平均长度为 22.26 cm，将其平均分为 3 段，分别定义为首段、中段和尾段，如图 1b 所示。果穗的首段、中段和尾段局部原始光谱如图 1a 所示。每根样品旋转 360°，并在旋转周期内，将果穗各分段间隔划分 6 个不同的采集点区域，中段采集点示意如图 1c 所示，采集点位 1、2、3、4、5、6 原始波形如图 1f 所示，μ、2μ、3μ、4μ、5μ、6μ 表示不同采集点位数量的平均光谱，如图 1e 所示。每根果穗采集 18 次光谱，所有样本各分段采集 540 次光谱，首段、中段和尾段共计采集 1 620 次光谱，全局原始光谱如图 1d 所示。使用 Matlab R2022b 建立 PLSR 鲜食玉米果穗的含水率预测模型。

获取光谱数据后，使用 WP－UC200 型相机和 CAM－MS 软件对玉米果穗图像进行采集，图像分辨率为 1 920 像素 ×1 200 像素。首先用标准钢板尺搭载物料盒，采集尺寸校准图像，保持采集参数不变，对样本玉米果穗图像依次进行采集，计算果穗首段、中段、尾段的长度像素尺寸，将像素尺寸转换为果穗不同分段的实际尺寸，公式为

$$I_{size} = I_P \frac{\sum_{i=1}^{3} \frac{P_i}{T_i}}{3} \tag{1}$$

式中　I_{size}——样本实际尺寸

I_P——样本像素尺寸

P_i——第 i 幅尺寸校准图像中钢板尺像素尺寸

T_i——第 i 幅尺寸校准图像中实际尺寸

1.2.2　含水率测定

在光谱采集完成后，分别测定鲜食玉米果穗首段、中段和尾段的含水率。通过人工手剥，从对应采集区域获取 7～10 粒完整籽粒进行含水率测量。将籽粒称量后放置于鼓风干燥机中进行循环

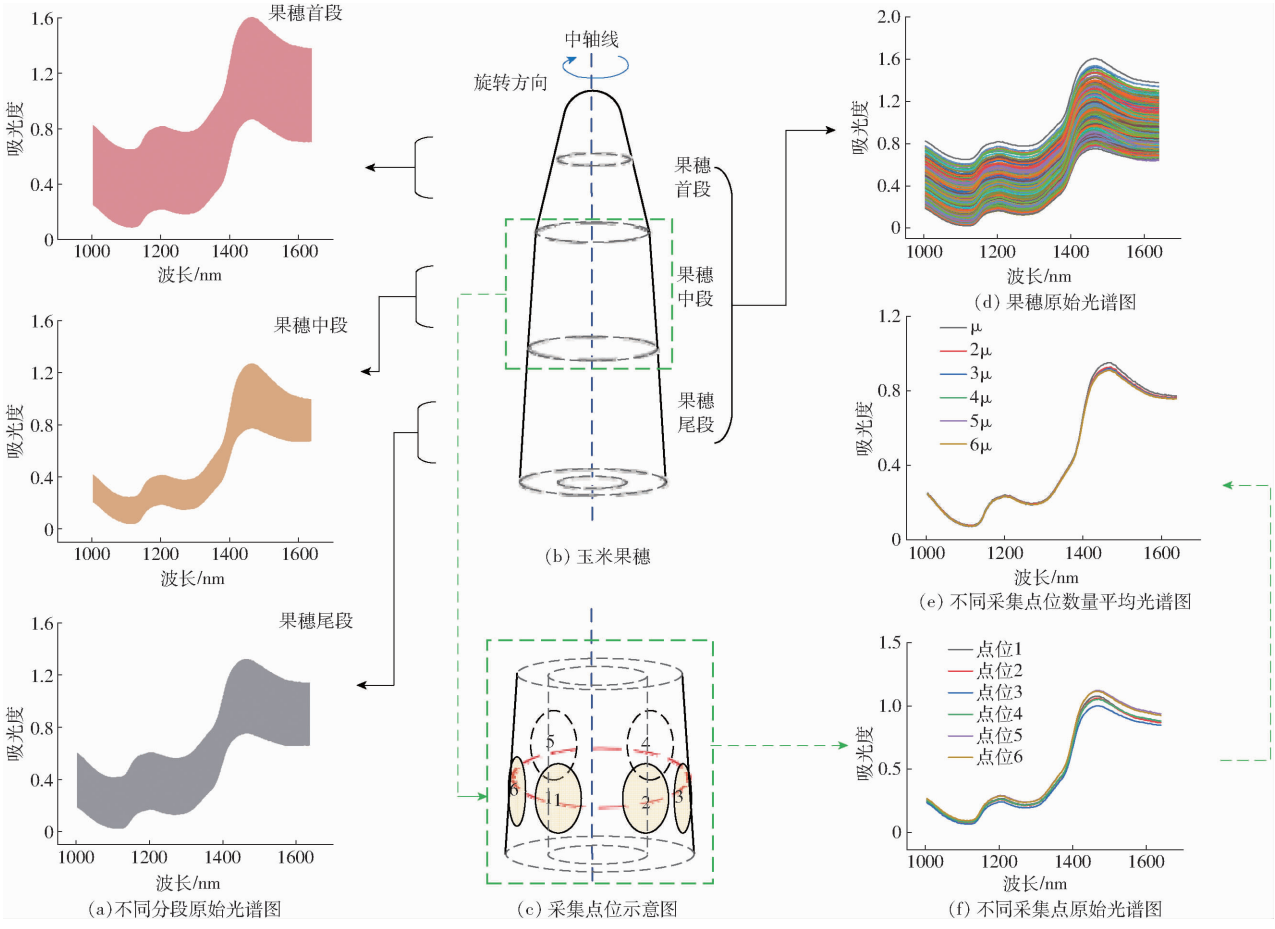

图 1　玉米果穗光谱采集示意图

Fig. 1　Schematics of corn ear spectral collection

干燥处理,采用直接干燥法称量,直至质量基本保持恒定。随后,取出所有籽粒再次称量,计算含水率[25]。

1.3 数据分析与建模方法

1.3.1 异常样本剔除

Z-score 异常值检测方法[26]是基于统计学中的标准化分数(Z-score)对数据集中的异常值进行识别与剔除,公式为

$$Z_i = \frac{x_i - \bar{x}_i}{\sigma} \tag{2}$$

式中 Z_i——第 i 个鲜食玉米籽粒含水率的标准化分数

x_i——第 i 个鲜食玉米籽粒含水率

$\bar{x}_i$——鲜食玉米籽粒含水率均值

σ——鲜食玉米籽粒含水率标准差

Z-score 值越大,表明数据点与数据集整体均值之间的偏差越显著,从而能够有效地识别出与数据分布差异较大的异常值。通过剔除数据中的异常值,保证了数据集的质量,有助于提高模型的准确性。

1.3.2 样本数据集划分

采用 SPXY 算法将鲜食玉米样品划分为校正集和预测集。SPXY 算法最早由 GALVAO 等[27]提出,该算法同时使用理化值和光谱来计算样本之间的距离,以确保最大程度地表征样本分布,并有效覆盖多维向量空间[28]。距离 d_{xy} 计算公式为

$$d_{xy}(i,j) = \frac{d_x(i,j)}{\max\limits_{i,j\in(1,z)}(d_x(i,j))} + \frac{d_y(i,j)}{\max\limits_{i,j\in(1,z)}(d_y(i,j))} \tag{3}$$

式中 $d_x(i,j)$——光谱特征的样本距离

$d_y(i,j)$——理化值特征的样本距离

z——样本总数

i、j——样本编号

1.3.3 光谱预处理方法

鲜食玉米果穗原始光谱数据存在散射、基线漂移、随机噪声,为了降低这些效应对模型稳健性的影响,提高模型预测精度,需要对光谱数据进行预处理。本研究采用了无预处理(NONE)、SNV、MSC、1D、2D 和 SG 等预处理方法[29-32]。

1.3.4 含水率建模方法

使用偏最小二乘法(Partial least squares regression,PLSR)进行鲜食玉米含水率建模。在使用 PLSR 建立校正模型时[33-34],为避免出现欠拟合或过拟合现象,选取参与建模的最佳潜变量(Latent variables,LVs)数。本研究采用交叉验证法,依据预测残差平方和达到最小值时对应的 LVs 数为最优主成分数。

1.3.5 模型评价参数

通过校正集决定系数 R_c^2、预测集决定系数 R_p^2、校正集均方根误差(Root mean square error of calibration,RMSEC)、预测集均方根误差(Root mean square error of prediction set,RMSEP)和剩余预测偏差(Residual predictive deviation,RPD)来评价模型性能[35-36]。

2 结果与分析

2.1 果穗形态分析

通过采集玉米果穗图像,计算出果穗分段形态参数,见表 1。统计结果显示,果穗的首段直径范围为 4.24 ~ 5.60 cm,果穗中段直径范围为 5.35 ~ 6.40 cm,果穗尾段直径范围为 5.86 ~ 7.03 cm,呈现出上细下粗的圆锥体形态,各果穗分段之间存在明显的粗细差异。

表 1 样本果穗各分段形态参数

Tab.1 Morphological parameters of corn ear segments

cm

分段	平均直径	最大直径	最小直径	直径标准差	直径均方根
首段	4.97	5.60	4.24	0.28	4.98
中段	5.87	6.40	5.35	0.26	5.88
尾段	6.40	7.03	5.86	0.29	6.41

2.2 含水率数据分析

为确保在剔除异常样本的同时尽可能保留足够的有效样本量,将 Z-score 的阈值设定为 2。即当数据点的 Z-score 绝对值超过 2 时,该数据点被判定为异常值,并从数据集中剔除。在果穗首段数据中,剔除 3 个样本,分别是样本 46、64 和 71;在果穗中段数据中,剔除了 5 个样本,分别是样本 24、29、33、64 和 71;在果穗尾段中,剔除 3 个样本,分别是样本 6、29 和 64,如图 2 所示。

在剔除异常值后,利用 SPXY 算法将果穗的首段、中段和尾段的数据集按比例 4∶1 划分为校正集和预测集,并分别进行统计分析,见表 2。样本集划分后,各果穗分段的校正集和预测集的含水率较为相近,说明通过 SPXY 划分的数据集分布较为合理,有助于提升预测模型的稳定性。由表 2 可知,不同果穗分段的含水率标准偏差、偏度和变异系数存在明显的差异,其中首段最高,中段最低,尾段在首段和中段之间,表明中段的含水率分布最为对称且离散程度最小。此外,不同果穗分段的含水率也存在差异,首段的含水率最高,中段和尾段的含水率相近。

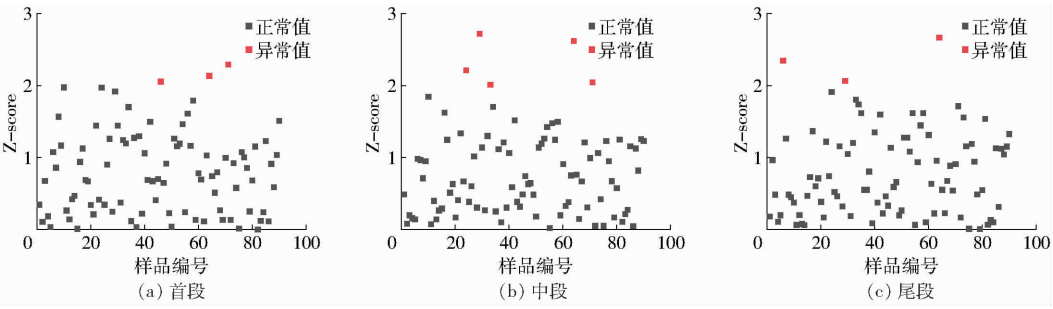

图 2　含水率异常值分布图

Fig. 2　Outlier distribution of moisture content

表 2　样本含水率统计

Tab. 2　Sample moisture content statistics

分段	数据集	数量/个	平均值/%	最大值/%	最小值/%	标准偏差/%	偏度	变异系数/%
首段	总数据集	87	64.98	75.56	56.10	4.80	0.47	7.39
	校正集	70	65.33	75.56	56.10	5.09	0.33	7.79
	预测集	17	63.54	70.73	59.19	3.06	0.75	4.81
中段	总数据集	85	63.87	72.56	56.34	3.96	0.35	6.21
	校正集	68	64.11	72.87	56.34	4.07	0.23	6.35
	预测集	17	62.91	70.92	58.39	3.45	0.95	5.48
尾段	总数据集	87	63.94	74.58	54.91	4.30	0.42	6.72
	校正集	70	64.59	74.58	54.91	4.35	0.20	6.74
	预测集	17	61.27	69.59	58.05	2.84	1.48	4.63

2.3　鲜食玉米光谱分析

选取波长范围为 1 000 ~ 1 640 nm 内不同果穗分段的漫反射光谱，如图 3 所示。在 1 190 nm 和 1 450 nm 处存在较为明显的水吸收峰[37]，波长 1 190 nm 附近的波峰主要是 C—H 官能团的 2 倍频弯曲振动吸收所导致，而波长 1 450 nm 附近的波峰则主要是 O—H 官能团的 1 倍频弯曲振动吸收引起的[38]。由图 3 可知，果穗各分段的平均光谱吸光度差异较大，果穗首段的平均光谱吸光度明显高于中段和尾段，而中段和尾段的平均光谱吸光度较为接近，推测其由含水率分布不均所导致。根据表 2 可知，果穗首段含水率最高，果穗中段与果穗尾段的含水率相近。图 3 分别展示了果穗首段、果穗中段和果穗尾段的原始光谱，光谱整体变化趋势一致，特征波峰一致，但其吸光度存在显著差异性。在波长 1 190 nm 处，果穗首段的吸光度为 0.200 ~ 0.811，果穗中段的吸光度为 0.160 ~ 0.451，果穗尾段的吸光度为 0.158 ~ 0.600；在波长 1 450 nm 处，果穗首段的吸光度为 0.856 ~ 1.593，果穗中段的吸光度为 0.763 ~ 1.258，果穗尾段的吸光度为 0.740 ~ 1.311。原因可能在于果穗首段、中段和尾段在直径及籽粒排布上的差异，导致光的散射和衰减程度各异。此外，各分段含水率的差异性也会影响吸收特性，对携带出来的样本内部信息产生影响，从而影响检测结果。同时可以发现，果穗中段不同采集点位和不同采集点位数量的平均光谱曲线特征基本一致，吸光度同样表现出显著差别，如图 1e、1f 所示。进一步表明鲜食玉米果穗表面籽粒排布凹凸不平的结构特性对光谱特性的影响。

2.4　全局预测建模分析

原始光谱数据不仅包含待测样品内部各种化学成分的相关信息，还受到随机噪声、杂散光和基线漂移等干扰信息，为减少杂散光等因素对建模结果的影响，提高模型的鲁棒性，采用 NONE、SNV、MSC、SG、1D 和 2D 等预处理方法结合 PLSR 模型建立鲜食玉米果穗包含全部区域（首段、中段和尾段）的全局模型，预测结果如表 3 所示。在全局建模中，经 1D 预处理后效果最好，建模结果 R_p^2、RMSEP 和 RPD 分别为 0.905、0.011% 和 3.270，LVs 数为 15。可以发现，在全局建模中 LVs 数较高，表明全局模型对数据适应性不足，模型结构复杂，泛化能力较弱[39]。为提升建模效果，进一步探讨不同区域（首段、中

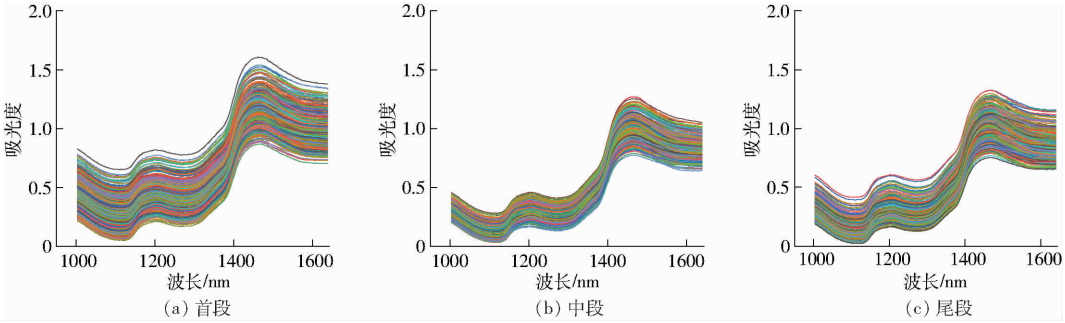

图 3　不同分段果穗光谱

Fig. 3　Spectra of different segmented ears

段、尾段)建模可行性,从而增强预测模型的稳定性和预测精度。

表 3 鲜食玉米果穗含水率全局模型的预测结果

Tab.3 Prediction results of global model for water content of fresh corn cob

预处理方法	LVs 数	R_c^2	RMSEC/%	R_p^2	RMSEP/%	RPD
NONE	15	0.938	0.011	0.893	0.011	3.087
SNV	17	0.949	0.010	0.889	0.011	3.028
MSC	18	0.950	0.010	0.891	0.011	3.057
1D	15	0.951	0.010	0.905	0.011	3.270
2D	17	0.957	0.009	0.860	0.012	2.699
SG	16	0.946	0.010	0.893	0.011	3.082

2.5 不同部位预测模型分析

为进一步提高检测精度,建立不同果穗部位的局部预测模型。从表 4 可知,在果穗首段建模中,经 1D 预处理后的模型精度最高,建模结果 R_p^2、RMSEP 和 RPD 分别为 0.937、0.009% 和 4.114,LVs 数为 9;在果穗中段建模中,经 SG 平滑预处理后的模型精度最高,建模结果 R_p^2、RMSEP 和 RPD 分别为 0.955、0.007% 和 4.884,LVs 数为 8。在果穗尾段建模中,经 1D 预处理后模型精度最高,建模结果 R_p^2、RMSEP 和 RPD 分别为 0.840、0.011% 和 2.579,LVs 数为 6。通过对比分析模型预测结果发现,果穗中段的建模效果最好,略优于果穗首段建模,明显优于果穗尾段建模,如图 4 所示。分析认为,果穗中段籽粒的排列结构均匀,其尺寸、形态和排列具有相当的一致性,且籽粒间紧密分布,能够降低由于物理特性变化引起的随机噪声,增强光谱信号的稳定性,同时携带出更多有用光谱信息。果穗首段的籽粒排列虽然较均匀,但由于果穗直径变小,光路间隙较大,从而引入了较多无用信息。果穗尾段的籽粒排列杂乱且空白增大,不仅增加了无效光谱信号的比例,还可能导致玉米芯反射信号的干扰,降低检测精度。

表 4 不同预处理方法下 PLSR 模型预测结果

Tab.4 Prediction results of PLSR model by different preprocessing methods

分段	预处理方法	LVs 数	R_c^2	RMSEC/%	R_p^2	RMSEP/%	RPD
首段	NONE	12	0.970	0.009	0.896	0.010	3.202
	SNV	9	0.967	0.009	0.896	0.010	3.192
	MSC	9	0.969	0.009	0.900	0.010	3.255
	1D	9	0.972	0.008	0.937	0.009	4.114
	2D	7	0.964	0.009	0.920	0.011	3.645
	SG	9	0.971	0.009	0.898	0.012	3.117
中段	NONE	11	0.970	0.007	0.955	0.008	4.878
	SNV	8	0.961	0.008	0.932	0.009	3.892
	MSC	8	0.961	0.008	0.934	0.009	3.951
	1D	7	0.969	0.007	0.945	0.008	4.406
	2D	7	0.956	0.008	0.908	0.009	3.404
	SG	8	0.968	0.007	0.955	0.007	4.884
尾段	NONE	10	0.901	0.013	0.815	0.014	2.361
	SNV	9	0.905	0.012	0.516	0.014	1.481
	MSC	9	0.905	0.012	0.522	0.014	1.490
	1D	6	0.872	0.010	0.840	0.011	2.579
	2D	5	0.882	0.013	0.746	0.018	2.045
	SG	8	0.916	0.012	0.707	0.015	1.904

为了进一步说明在果穗中部建立含水率预测模型的优越性,分别建立果穗首段、中段和尾段的预测模型,并对不同果穗分段的模型进行交叉验证,如表 5 所示。结果表明,各段模型在预测各自分段的含水率时效果最佳,而在预测其它分段时,预测结果均有所下降。分别计算了 3 个模型在预测不同果穗分段时平均 RMSEP,结果分别为 0.016 7%、0.014 0%、0.016 0%。结果显示,果穗中段预测模型预测首段、中段和尾段时的 RMSEP 最低,说明在果穗中部建立的含水率预测模型具有更好的泛化能力。

2.6 不同采集点位预测模型分析

针对鲜食玉米果穗籽粒排布及含水率分布不均的问题,本研究系统探讨果穗中部采集点位数量对鲜食玉米果穗含水率预测模型的影响。基于 SG-PLSR 模型,分别构建了果穗中段在不同采集点位数量平均光谱的预测模型,如表 6 所示。实验结果表明,当采集点位数量为 1 时,模型预测精度最低,R_p^2 仅为 0.535;随着采集点位数量的增加,预测集的预

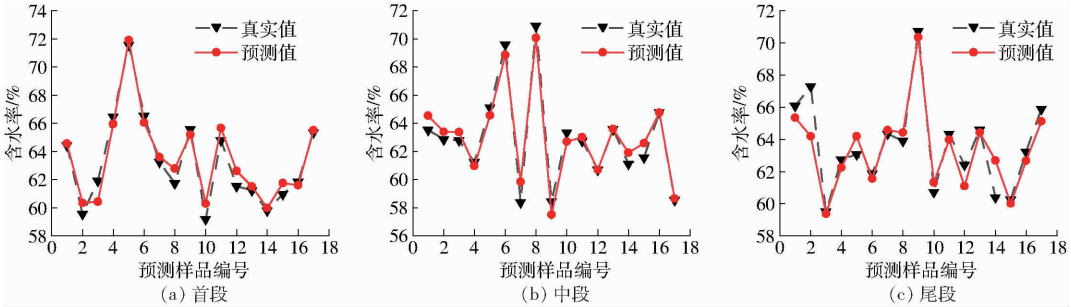

图 4 果穗不同分段预测值与真实值对比曲线

Fig.4 Modeling results of different segments of fruit ear

表 5　不同果穗分段交叉验证

Tab. 5　Cross validation of different ear segments

模型建立	预测部位	预处理方法	LVs 数	R_p^2	RMSEP/%
首段	首段	NONE	9	0.946	0.007
	中段	N0NE	9	0.926	0.011
	尾段	NONE	9	0.306	0.032
中段	首段	SG	8	0.936	0.012
	中段	SG	8	0.955	0.007
	尾段	SG	8	0.648	0.023
尾段	首段	1D	6	0.837	0.019
	中段	1D	6	0.785	0.018
	尾段	1D	6	0.840	0.011

表 6　果穗中段不同采集点数量建模结果

Tab. 6　Results of modeling number of different collection regions in middle part of ear

采集点位数量	LVs 数	R_c^2	RMSEC/%	R_p^2	RMSEP/%	RPD
1	8	0.980	0.009	0.535	0.015	1.531
2	10	0.955	0.009	0.896	0.009	3.221
3	10	0.970	0.007	0.901	0.008	3.271
4	10	0.977	0.006	0.908	0.009	3.394
5	9	0.974	0.006	0.967	0.007	5.967
6	8	0.968	0.007	0.955	0.007	4.884

测精度呈现逐步提升的趋势。当采集点位数量增加至 5 时，模型的预测精度达到最佳，其 R_p^2、RMSEP 和 RPD 分别为 0.967、0.007% 和 5.967；采集点数量继续增加至 6 时，模型预测精度都呈现下降趋势。

图 5 进一步展示了不同采集点位数量对鲜食玉米果穗含水率预测精度的影响趋势。可以看出，采集点数量从 5 增加到 6 后，模型预测精度出现小幅度下降。其原因在于，新增采集点与已有采集点光谱特征相似，难以提供更多有效信息，反而可能引入测量误差或环境噪声。同时，鲜食玉米果穗的物理特性本就较为复杂，增加采集点会进一步提升数据复杂度，进而影响模型的稳定性和泛化能力。综合来看，采集点数量为 5 时，在信息获取与噪声引入之间达到了较好的平衡。

3　结论

（1）本文系统分析了鲜食玉米果穗不同分段的形态和含水率分布特征。统计结果显示，果穗呈现出上细下粗的圆锥体形态，各段之间存在明显的粗细差异。果穗首段的含水率最高，而果穗中段和果穗尾段的含水率相近，表明整根鲜食玉米果穗含水率分布不均，其果穗中段的含水率分布较为理想。这一发现揭示了鲜食玉米果穗存在形态差异性和含水率分布不均的现象，直接影响近红外光谱的吸收和散射特性，为后续的模型分析提供了重要依据。

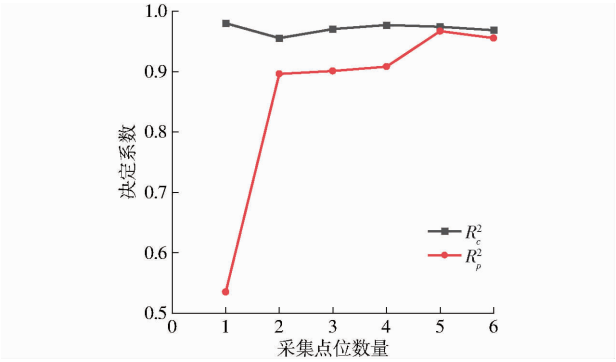

图 5　不同采集点位数量建模精度变化曲线

Fig. 5　Trend of modeling accuracy for different number of acquisition points

（2）采用 6 种预处理方法结合 PLSR 模型，建立了融合鲜食玉米果穗首段、中段和尾段全部数据的全局模型，以及分别针对首段、中段和尾段建立的局部模型。结果显示，在全局模型中 LVs 数较高，表明全局模型对数据适应性不足，模型结构复杂且泛化能力较弱。因此，进一步分析果穗不同分段局部建模，并对果穗不同分段模型进行交叉验证。对比首段、中段和尾段各自精度最高建模方法发现，果穗中段预测决定系数 R_p^2 比果穗首段高 1.8%，比果穗尾段高 11.5%。因此，果穗中段建模效果最优，其最佳预测模型为 SG－PLSR，建模结果 R_p^2、RMSEP 和 RPD 分别为 0.955、0.007% 和 4.884，LVs 数为 8。同时，果穗中段的模型交叉验证的平均 RMSEP 最低，为 0.014%，优于果穗首段和果穗尾段的模型表现。综合对比分析，果穗中段建模效果最优且具有更好的泛化能力。

（3）基于 SG－PLSR 模型，分别构建了果穗中段在不同采集点位数量下平均光谱的预测模型。结果发现，当采集点位数量为 5 时，模型的预测精度达到最佳，其 R_p^2、RMSEP 和 RPD 值分别为 0.967、0.007% 和 5.967。

参　考　文　献

[1]　王娟紫，乔勇进，王春芳，等. 鲜食糯玉米采后生理与保鲜技术的研究进展[J]. 食品与发酵工业，2023，49(16)：356－351.
Wang Juanzi，Qiao Yongjin，Wang Chunfang，et al. Progress of postharvest physiology and preservation technology of fresh glutinous corn[J]. Food and Fermentation Industry，2023，49(16)：356－351.（in Chinese）

[2]　徐丽，赵久然，卢柏山，等. 我国鲜食玉米种业现状及发展趋势[J]. 中国种业，2020(10)：14－18.

Xu Li, Zhao Jiuran, Lu Baishan, et al. Current situation and development trend of fresh corn seed industry in China[J]. China Seed Industry, 2020(10): 14 - 18. (in Chinese)

[3] 司婉芳, 杨福馨, 王金鑫. 鲜食玉米常温保鲜工艺研究[J]. 食品与发酵工业, 2019,45(6):159 - 164.
Si Wanfang, Yang Fuxin, Wang Jinxin. Research on room temperature preservation process of fresh corn[J]. Food and Fermentation Industry, 2019,45(6):159 - 164. (in Chinese)

[4] 杨明花, 刘强, 冯国瑞, 等. 鲜食糯玉米适宜采收期与籽粒含水量分析[J]. 新疆农业科技, 2024, 61(7),1626 - 1630.
Yang Minghua, Liu Qiang, Feng Guorui, et al. Analysis of suitable harvesting period and grain moisture content of fresh glutinous maize[J]. Xinjiang Agricultural Sciences, 2024, 61(7),1626 - 1630. (in Chinese)

[5] Wang H, Liu J S, Min W H, et al. Changes of moisture distri-bution and migration in fresh ear corn during storage[J]. Integrative Agriculture,2019,18(11):2644 - 2651.

[6] Ali M M M, Hashim N. Non-destructive methods for detection of food quality[M]. Amsterdam: Future Foods Academic Press, 2022: 645 - 667.

[7] Duckena L, Alksnis R, Erdberga I, et al. Non-destructive quality evaluation of 80 tomato varieties using Vis - NIR spectroscopy [J]. Foods, 2023, 12(10): 1990.

[8] Zhang X, Zhou C, Sun Q, et al. Prediction of solid soluble content of green plum based on improved CatBoost[J]. Agriculture, 2023, 13(6): 1122.

[9] 李鸿强,孙红,李民赞. 基于可见/短波近红外光谱检测结球甘蓝维生素 C 含量[J]. 农业工程学报, 2018, 34(8): 296 - 275.
Li Hongqiang, Sun Hong, Li Minzan. Detection of vitamin C content in nodular kale based on visible/short-wave near-infrared spectroscopy[J]. Transactions of the CSAE, 2018, 34(8): 296 - 275. (in Chinese)

[10] 赵娟, 全朋坤, 马敏娟, 等. 富士苹果采收成熟度光谱无损预测模型对比分析[J]. 农业机械学报, 2018, 49(12): 347 - 354.
Zhao Juan, Quan Pengkun, Ma Minjuan, et al. Comparative analysis of spectral nondestructive prediction models for Fuji apple harvesting maturity[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(12): 347 - 354. (in Chinese)

[11] Zhang Y, Chen Y, Wu Y, et al. Accurate and nondestructive detection of apple brix and acidity based on visible and near-infrared spectroscopy[J]. Applied Optics, 2021, 60(13): 4021 - 4028.

[12] Semyalo D, Kwon O, Wakholi C, et al. Nondestructive online measurement of pineapple maturity and soluble solids content using visible and near-infrared spectral analysis[J]. Postharvest Biology and Technology, 2024, 209: 112706.

[13] Santos C S P, Cruz R, Goncalves D B, et al. Non-destructive measurement of the internal quality of citrus fruits using a portable NIR device[J]. AOAC International, 2021, 104(1): 61 - 67.

[14] Zhang B, Gu B, Tian G, et al. Challenges and solutions of optical-based nondestructive quality inspection for robotic fruit and vegetable grading systems: a technical review[J]. Trends in Food Science & Technology, 2018, 81: 213 - 231.

[15] Hong F W, Chia K S. A review on recent near infrared spectroscopic measurement setups and their challenges[J]. Measurement, 2021, 171: 108732.

[16] 黄积微, 李洋, 袁迪, 等. 基于便携式近红外光谱仪的蓝莓热风干燥过程模型的建立[J]. 食品与发酵工业, 2023, 49(16): 283 - 290.
Huang Jiwei, Li Yang, Yuan Di, et al. Modeling of hot air drying process of blueberry based on portable near infrared spectrometer[J]. Food and Fermentation Industry, 2023, 49(16): 283 - 290. (in Chinese)

[17] Lu R, Van B R, Saevs W, et al. Measurement of optical properties of fruits and vegetables: a review[J]. Postharvest Biology and Technology, 2020, 159: 111003.

[18] Fan S, Zhang B, Li J, et al. Effect of spectrum measurement position variation on the robustness of NIR spectroscopy models for soluble solids content of apple[J]. Biosystems Engineering, 2016, 143: 9 - 19.

[19] Greensill C V, Walsh K B. A remote acceptance probe and illumination configuration for spectral assessment of internal attributes of intact fruit[J]. Measurement Science and Technology, 2000, 11(12): 1674.

[20] 黄文倩, 田喜, 张艺飞, 等. 基于全透射近红外光谱的西瓜不同部位可溶性固形物含量在线检测研究[J]. 光谱学与光谱分析, 2023, 43(6): 1800 - 1808.
Hang Wenqian, Tian Xi, Zhang Yifei, et al. On-line determination of soluble solids in different parts of watermelon based on full transmission near infrared spectroscopy[J]. Spectroscopy and Spectral Analysis, 2023, 43(6): 1800 - 1808. (in Chinese)

[21] 姜小刚,朱明旺,姚金良,等 .基于近红外光谱技术苹果尺寸差异对糖度模型适用性的影响[J]. 华中农业大学学报, 2024,43(1):242 - 248.
Jiang Xiaogang, Zhu Mingwang, Yao Jinliang, et al . Effect of apple size difference on the applicability of brix model based on near infrared spectroscopy[J]. Journal of Huazhong Agricultural University, 2024, 43(1): 242 - 248. (in Chinese)

[22] Fan C, Liu Y, Cui T, et al. Quantitative prediction of protein content in corn kernel based on near-infrared spectroscopy[J]. Foods, 2024, 13(24): 4173.

[23] Cataltas O, Tutuncu K. Detection of protein, starch, oil, and moisture content of corn kernels using one-dimensional convolutional autoencoder and near-infrared spectroscopy[J]. PeerJ Computer Science, 2023, 9: e1266.

[24] 张永立, 杨光辉, 王美蟠, 等. 鲜食玉米果穗含水率近红外光谱无损检测影响因素[J]. 农业工程学报, 2024, 40(15): 262 – 270.
Zhang Yongli, Yang Guanghui, Wang Meipan, et al. Factors affecting water content of fresh corn cobs by near-infrared spectroscopy for nondestructive testing[J]. Transactions of the CSAE, 2024, 40(15): 262 – 270. (in Chinese)

[25] 廉孟茹, 张淑娟, 任锐. 基于高光谱技术的鲜食水果玉米含水率无损检测[J]. 食品工业, 2021(9):127 – 132.
Lian Mengru, Zhang Shujuan, Ren Rui. Non-destructive detection of water content in fresh fruit corn based on hyperspectral technology[J]. Food & Machinery, 2021(9):127 – 132. (in Chinese)

[26] 张晨, 朱玉杰, 冯国红. 基于集成学习和近红外光谱的蓝莓贮藏品质检测方法研究[J]. 食品与发酵工业, 2023, 49(18): 306 – 314.
Zhang Chen, Zhu Yujie, Feng Guohong. Research on blueberry storage quality detection method based on integrated learning and near infrared spectroscopy[J]. Food and Fermentation Industry, 2023, 49(18): 306 – 314. (in Chinese)

[27] Galvao R K H, Araujo M C U, Jose G E, et al. A method for calibration and validation subset partitioning[J]. Talanta, 2005, 67(4): 736 – 740.

[28] Wang R, Song J, Liu J, et al. The combination of near-infrared spectroscopy with chemometrics in achieving rapid and accurate determination of rice mildew[J]. Food Measurement and Characterization, 2024,18(7):5437 – 5449.

[29] 汪六三, 黄子良, 王儒敬. 基于近红外光谱和机器学习的大豆种皮裂纹识别研究[J]. 农业机械学报, 2021, 52(6): 361 – 368.
Wang Liusan, Huang Ziliang, Wang Rujing. Research on soybean seed coat crack recognition based on near infrared spectroscopy and machine learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(6): 361 – 368. (in Chinese)

[30] 彭彦昆, 杨清华, 王文秀. 基于近红外光谱的猪肉水分在线检测与分级[J]. 农业机械学报, 2018, 49(3): 347 – 353.
Peng Yankun, Yang Qinghua, Wang Wenxiu. Online detection and grading of pork moisture based on near infrared spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(3): 347 – 353. (in Chinese)

[31] 张进, 胡芸, 周罗雄, 等. 近红外光谱分析中的化学计量学算法研究新进展[J]. 分析测试学报, 2023, 39(10): 1196 – 1203.
Zhang Jin, Hu Yun, Zhou Luoxiong, et al. New advances in chemometric algorithms for near-infrared spectral analysis[J]. Journal of Analytical Testing, 2023, 39(10): 1196 – 1203. (in Chinese)

[32] 赵娟, 沈懋生, 浦育歌, 等. 基于近红外光谱与多品质指标的苹果出库评价模型研究[J]. 农业机械学报, 2022, 54(2): 386 – 395.
Zhao Juan, Shen Maosheng, Pu Yuge, et al. Research on the evaluation model of apple leaving the warehouse based on near-infrared spectroscopy and multiple quality indicators[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 54(2): 386 – 395. (in Chinese)

[33] 闫雨桐,史策,韩帅,等.基于电子鼻的玫瑰香葡萄货架期品质预测及新鲜度判别研究[J].山地农业生物学报,2024, 43(6):1 – 7.
Yan Yutong, Shi Ce, Han Shuai, et al. Research on shelf-life quality prediction and freshness discrimination of rosy grapes based on electronic nose[J]. Journal of Mountain Agricultural Biology,2024,43(6):1 – 7. (in Chinese)

[34] Li S, Li J, Wang Q, et al. Determination of soluble solids content of multiple varieties of tomatoes by full transmission visible-near infrared spectroscopy[J]. Frontiers in Plant Science, 2024, 15: 1324753.

[35] Liu W, Han Y, Wang N, et al. Apple sugar content non-destructive detection device based on near-infrared multi-characteristic wavelength[C]// Physics: Conference Series. Bristol: IOP Publishing, 2022: 012012.

[36] 刘文政, 周雪健, 平凤娇, 等. 基于可见-近红外光谱的鲜食葡萄成熟品质关键指标检测[J]. 农业机械学报, 2024, 55(2): 372 – 383.
Liu Wenzheng, Zhou Xuejian, Ping Fengjiao, et al. Detection of key quality indicators of maturity in fresh table grapes based on visible-near infrared spectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(2): 372 – 383. (in Chinese)

[37] Hu Y, Gao X, Shen Z, et al. Estimating fertility index by using field-measured vis-nir spectroscopy in the Huanghui River basin[J]. Chinese Journal of Soil Science, 2021, 52(3): 575 – 584.

[38] 毛欣然, 夏静静, 徐惟馨, 等. 手持式近红外光谱仪测定梨三种品质指标通用模型建模方法研究[J]. 光谱学与光谱分析, 2024, 44(2): 406 – 412.
Mao Xinran, Xia Jingjing, Xu Weixin, et al. A general modeling approach for the determination of three pear quality indicators by handheld near infrared spectroscopy[J]. Spectroscopy and Spectral Analysis, 2024, 44(2): 406 – 412. (in Chinese)

[39] Gong B. Study of PLSR – BP model for stability assessment of loess slope based on particle swarm optimization[J]. Scientific Reports, 2021, 11(1): 17888.