

DOI: 10.19812/j.cnki.jfsq11-5956/ts.20240930006

引用格式: 胡孟凤, 操丽丽, 庞敏, 等. 近红外光谱技术定量分析小麦中黄曲霉毒素 B₁[J]. 食品安全质量检测学报, 2025, 16(4): 10–17.

HU MF, CAO LL, PANG M, *et al.* Quantitative analysis of aflatoxin B₁ in *Triticum aestivum* L. by near-infrared spectroscopy technology [J]. Journal of Food Safety & Quality, 2025, 16(4): 10–17. (in Chinese with English abstract).

近红外光谱技术定量分析小麦中黄曲霉毒素 B₁

胡孟凤^{1,2}, 操丽丽^{1,2,3}, 庞敏^{1,2,3}, 高春^{3,4}, 许丽^{3,4}, 姜绍通^{1,2,3}, 赵妍嫣^{1,2,3*}

(1. 合肥工业大学食品与生物工程学院, 合肥 230601; 2. 农产品现代加工安徽省重点实验室, 合肥 230601;

3. 农产品智能化绿色质选技术与装备安徽省联合共建学科重点实验室, 合肥 230012;

4. 安徽捷迅光电技术有限公司, 合肥 230012)

摘要: **目的** 基于近红外光谱技术建立定量预测模型, 实现快速无损测定小麦籽粒中黄曲霉毒素 B₁ (aflatoxin B₁, AFB₁)含量。**方法** 采集小麦样品在 900~1700 nm 波长范围内的反射光谱, 用高效液相色谱法测定小麦中 AFB₁ 含量, 将小麦样品原始光谱数据进行预处理, 提取特征波长, 分别通过反向传播神经网络(back propagation neural network, BPNN)、随机森林(random forest, RF)和支持向量机(support vector machine, SVM)建立 AFB₁ 含量预测模型, 并与全波段建模结果进行比较。**结果** 经多元散射校正(multiplicative scatter correction, MSC)和竞争性自适应加权算法(competitive adaptive reweighted sampling, CARS)处理后建立的 SVM 模型优于其他模型和全波段建模模型。**结论** 结合 CARS 算法有效提取了 AFB₁ 值的特征波长, MSC-CARS-SVM 模型能够用于 AFB₁ 含量的快速、无损检测, 利用近红外光谱技术实现对 AFB₁ 含量的定量分析是可行的, 可通过该方法实现储藏期间小麦品质的检测研究。

关键词: 近红外光谱技术; 黄曲霉毒素 B₁; 定量分析; 无损检测; 小麦

Quantitative analysis of aflatoxin B₁ in *Triticum aestivum* L. by near-infrared spectroscopy technology

HU Meng-Feng^{1,2}, CAO Li-Li^{1,2,3}, PANG Min^{1,2,3}, GAO Chun^{3,4}, XU Li^{3,4},
JIANG Shao-Tong^{1,2,3}, ZHAO Yan-Yan^{1,2,3*}

(1. School of Food and Biological Engineering, Hefei University of Technology, Hefei 230601, China; 2. Key Laboratory of Modern Processing of Agricultural Products of Anhui Province, Hefei 230601, China; 3. Intelligent Green Quality Selection Technology and Equipment for Agricultural Products Key Laboratory of Anhui Province Jointly Constructed Disciplines, Hefei 230601, China; 4. Jiexun Optoelectronic Technology Co of Anhui Province, Hefei 230012, China)

ABSTRACT: Objective To achieve rapid and non-destructive determination of aflatoxin B₁ (AFB₁) content in *Triticum aestivum* L. kernels by establishing a quantitative prediction model based on near-infrared spectroscopy technology. **Methods** The reflectance spectra of *Triticum aestivum* L. samples in the wavelength range of 900–1700 nm were collected, and the AFB₁ content in *Triticum aestivum* L. was determined by high performance liquid

收稿日期: 2024-09-30

基金项目: 安徽省重点研究与开发计划项目(2023n06020014)

第一作者: 胡孟凤(2000—), 女, 硕士, 主要研究方向为农产品无损检测。E-mail: hmengf@126.com

*通信作者: 赵妍嫣(1971—), 女, 博士, 副教授, 主要研究方向为农产品精深加工。E-mail: zhaoyanyan@hfut.edu.cn

chromatography. The raw spectral data of the *Triticum aestivum* L. samples were subjected to preprocessing, and the feature wavelengths were extracted in order to establish a prediction model. A model for predicting the AFB₁ content was developed using a back propagation neural network (BPNN), random forest (RF), and support vector machine (SVM), the results of this model were compared with those of a full-wavelength modelling approach. **Results** The SVM model constructed following the application of multiplicative scatter correction (MSC) and competitive adaptive reweighted sampling (CARS) processing demonstrates superior performance compared to the other models and the full-band modelling model. **Conclusion** The combination of the CARS algorithm and the MSC-CARS-SVM model allows for the rapid and non-destructive detection of AFB₁ content. The feasibility of using near-infrared spectroscopy for quantitative analysis of AFB₁ content has been demonstrated, and this approach can be employed to assess the quality of *Triticum aestivum* L. during storage.

KEY WORDS: near-infrared spectroscopy technology; aflatoxin B₁; quantitative analysis; non-destructive testing; *Triticum aestivum* L.

0 引言

小麦是中国三大粮食作物之一^[1], 是中国农业的支柱, 对国家的稳定 and 经济增长具有重要影响。小麦从收获到食用中间需经历长时间储存, 由于其具有吸湿性, 当仓储环境的温度和湿度发生变化时, 容易受到植物病原菌的侵染, 从而导致霉变, 有的还会产生有害的真菌毒素, 例如, 黄曲霉会生成黄曲霉毒素 B₁ (aflatoxin B₁, AFB₁), 而镰孢菌则可能产生脱氧雪腐镰刀菌烯醇(deoxynivalenol, DON)和玉米赤霉烯酮(zearalenone, ZEN)^[2-4]。其中黄曲霉的代谢产物 AFB₁ 是其所有代谢产物中毒性最强并易致癌的物质, 可对人类和动物造成严重危害^[5-8]。AFB₁ 的污染对人类健康构成了重大威胁。许多国家已采取严格措施, 以控制粮食中 AFB₁ 的含量, 降低其对健康的风险。按照 GB 2761—2017《食品安全国家标准 食品中真菌毒素限量》规定, AFB₁ 的浓度必须低于 10 μg/kg, 然而, 欧盟的规定则更为严格, 其规定小麦中 AFB₁ 的浓度必须低于 5 μg/kg。

目前, 用于检测 AFB₁ 的方法包括色谱技术、质谱技术和基于免疫测定的分子识别^[9-10]。尽管色谱法和质谱法具有很高的准确度和精密性, 但检测时操作烦琐不可逆、成本高、破坏种子状态且不适用于快速准确的大规模检测等缺点^[11], 已无法满足现代小麦 AFB₁ 污染状况快速无损检测的需求, 迫切需要一种快速准确的方法来检测小麦 AFB₁ 污染状况。

目前, 近红外光谱(near-infrared spectroscopy, NIRS)法^[12-14]、拉曼光谱法^[15-16]以及高光谱法^[17-19]等光谱检测技术由于高效简便且对样本无损等优点正日益广泛应用于食品安全检测领域^[20]。NIRS 检测技术利用不同化合物基团的运动具有各自固有的频率, 当电磁波的波长在 780~2526 nm 范围内时, NIRS 对其的吸收会有所不同^[21]。对于待测物质的化学结构, 可以通过吸收峰在光谱中的位置和形态来判断, 而待测物质中各成分的大小则由特征

吸收峰的强度来决定^[22]。在 NIRS 技术中, 待测样品不需要预处理, 而且检测耗时短、结果准确、效率高^[23-26], 因此, NIRS 已被广泛应用于现场检测中^[27]。当小麦等粮食作物发生霉变时, 霉菌及其代谢物会产生特征峰, 这些特征峰的含量对待测样品的光谱信息具有重要影响^[28]。由于 NIRS 数据中包含的信息量巨大且存在冗余, 因此去除无效信息并过滤特征波长以提高模型的效率和准确性至关重要。据此, 本研究利用高效液相色谱法测定小麦样品的 AFB₁ 值, 将采集的 NIRS 信息结合反向传播神经网络(back propagation neural network, BPNN)、随机森林(random forest, RF)和支持向量机(support vector machine, SVM)对光谱信息及化学值进行拟合, 从而建立一种高效、快速、无损的小麦中 AFB₁ 含量检测技术, 以期为小麦 AFB₁ 污染监测的便携式 NIRS 系统无损检测装置的设计与开发提供技术依据与方法参考, 同时也为粮食监管部门现场检测提供一种快速、准确、无损的技术方法。

1 材料与方法

1.1 材料与试剂

小麦: 产地河南商丘。

AFB₁ 标准品(纯度≥97.0%, 北京百奥莱博科技有限公司); 甲醇、乙腈(色谱纯, 德国 Merck 公司); 三氯甲烷(纯度≥99.0%)、乙酸(纯度≥99.5%)(国药集团化学试剂有限公司); 三氟乙酸(色谱纯, 上海麦克林生化科技股份有限公司)。

1.2 仪器与设备

S6000 高效液相色谱仪(北京华谱科仪科技有限公司); NIRQUEST+1.7-100 近红外光谱仪(上海蔚海光学仪器有限公司); TG-16E 博科高速台式离心机(山东高芯生物传感器研究院有限公司); DHP-9032T 生化培养箱(上海一恒科学仪器有限公司); BSA124S-CW 天平(精度 0.0001 g, 德国 Sartorius 公司)。

1.3 实验方法

1.3.1 制备黄曲霉菌孢子悬浮液

将购买的黄曲霉菌 2219 接种到麦芽汁琼脂培养基斜面上, 在 28 °C 下培养 3 d 后在斜面中加入无菌水制备孢子悬浮液, 将孢子悬浮液浓度调整为 10^6 CFU/mL。

1.3.2 不同黄曲霉菌污染水平的小麦样品制备

取 400 g 小麦放入培养皿中, 在样品表面喷洒一层无菌水并混匀, 加入 20 mL 稀释后的孢子悬液混匀, 在 28 °C 培养箱中培养。设置培养时间为 2、4、6、8 d, 每次取 5 g 污染样品(10 份)并另取 5 g 未被污染样品(5 份)于培养皿用无菌塑料封口膜封口后置于 -18 °C 冷冻备用。

1.3.3 NIRS 采集样品

在反射模式下使用 NIRQUEST+1.7-100 近红外光谱仪采集样本在 900~1700 nm 范围内的光谱数据(图 1), 由 128 个变量组成。每个光谱代表扫描 100 次的平均值, 积分时间为 10 ms, 光谱分辨率为 5.88 nm, 信噪比为 13000:1。



图 1 NIRS 采集样品
Fig.1 Sample collection by NIRS

1.3.4 不同黄曲霉菌污染水平的小麦 AFB₁ 含量测定

参考 GB 5009.22—2016《食品安全国家标准 食品中黄曲霉毒素 B 族和 G 族的测定》中的高效液相色谱-柱前衍生法。

1.4 模型建立方法

BPNN 是一种双向传播的网络结构, 具备自主学习和自我进化的能力, 它主要由输入层、隐藏层和输出层构成, 输入层负责接受自变量; 隐藏层用于提取自变量对网络的影响权重特征; 输出层则生成最终的目标或预测结果^[29]。在 BPNN 中输入数据通过正向传输从输入层流向输出层, 而误差则在反向传播过程中从输出层传回输入层, 通过这种反馈机制, 网络可以通过迭代调整权重和阈值, 优化结构, 减少误差, BPNN 能够广泛应用于各类情况, 如分类、回归、模式识别和函数逼近等^[30]。

RF 是一种基于集成学习思想的机器学习算法, 通过构建大量的独立决策树来进行预测, 每棵决策树的训练过程都使用特征和样本的随机子集, 这种随机性不仅增强了模型的多样性, 还有效地减少了不同树之间的相关性, 从而降低了过拟合的风险^[31]。RF 能够通过多棵树的投票机

制或平均值来做出最终预测, 处理具有高维特征、噪声较大或复杂数据的任务, 有效捕捉数据中的潜在规律。尽管 RF 在准确性和鲁棒性方面具有显著优势, 但是它的训练过程通常需要构建大量的决策树, 这导致其在计算上相对较为繁重, RF 的训练和预测速度会较慢, 尤其是在数据量较大时, 可能会消耗较多的计算资源和时间。

SVM 是一种基于统计学习理论的监督学习方法, 广泛应用于分类问题。其核心思想是通过构建一个最优超平面, 将不同类别的数据点分开, 为了提高分类的准确性和鲁棒性, SVM 不仅要求超平面能够正确地划分数据, 还通过最大化超平面与离它最近的样本点(即支持向量)之间的间隔, 从而实现最佳的分类效果^[32]。SVM 能够在高维特征空间中进行有效分类, 即使在数据中存在噪声和异常值时, 也能保持较好的性能。本研究所使用模型参数(BPNN 模型参数、RF 模型参数、SVM 模型参数)见表 1~3。

表 1 BPNN 模型参数
Table 1 Model parameters of BPNN

| 网络结构参数 | | | | 训练参数 | | |
|------------------|------------------|------------------|------------------|----------|-----------|------|
| 输入层 神经 元数量 | 输出层 神经 元数量 | 隐藏层 神经 元数量 | 隐藏层 神经 元数量 | 迭代 次数 | 目标 误差 | 学习率 |
| 57 | 1 | 1 | 5 | 1000 | 10^{-6} | 0.01 |

表 2 RF 模型参数
Table 2 Model parameters of RF

| 模型结构参数 | | 数据处理参数 | |
|--------|---------|--------|-------|
| 决策树数量 | 最小叶子节点数 | 输入特征数 | 输出目标数 |
| 100 | 5 | 36 | 1 |

表 3 SVM 模型参数
Table 3 Model parameters of SVM

| 模型结构参数 | | | 数据处理参数 | |
|--------|-------------|------|-----------|-----------|
| 惩罚因子 | 径向基函数 参数 | 容忍度 | 输入 特征数 | 输出 目标数 |
| 4.0 | 0.8 | 0.01 | 15 | 1 |

1.5 数据处理

以 45 份不同 AFB₁ 含量的小麦样品作为建立预测模型的样品, 每份样品在 NIRS 仪测定 3 次取平均值作为原始光谱数据, 对原始光谱数据进行预处理, 包括标准正态变化(standard normal variate, SNV)、多元散射矫正(multiplicative scatter correction, MSC)、一阶导数(first derivative, FD)和去中心化(mean detrending, MD), 使用连续投影算法(successive projections algorithm, SPA)以及竞争性自适应加权算法(competitive adaptive reweighted sampling, CARS)进行特征波长的提取, 可以增强分析的精确性和敏感度, 从而有效减少光谱数据的复杂度, 并提高模型运算的速度。将经预处理和特征波长提取后的样品按照 7:3 的比例

随机划分为训练集和测试集, 分别建立 BPNN、RF、SVM 回归预测模型, 并与全波段建模结果进行比较。

其中 NIRS 预处理、特征波长提取和建模在 MATLABR2023a 中进行, 绘图在 MATLABR2023a 和 Origin 2024b 中进行。

1.6 模型评价指标

AFB₁ 含量预测模型建立后, 用训练集和测试集的决定系数(determination coefficient, R^2)、均方根误差(root mean square error, RMSE)、平均绝对误差(mean absolute error, MAE)和残差预测偏差(relative percent difference, RPD)对定量模型的预测性能进行综合评价。 R^2 衡量所测的 AFB₁ 真实值与模型所给预测值之间的相关性, 值越接近 1 表示该预测模型拟合效果越好。RMSE 衡量所测 AFB₁ 真实值与预测值之间的差异, MAE 衡量所测 AFB₁ 真实值与预测值之间差异的平均大小, 两者值越小, 模型的预测性能越好, 误差越小。RPD 表示预测误差相对于 AFB₁ 真实值变异度的比值, $RPD < 1.4$, 认为所建模型不可靠; $1.4 < RPD < 2.0$, 认为所建模型较可靠; $RPD > 2.0$, 认为具备较高可靠性, 能够用于模型分析。各个指标的计算公式如(1)~(4)所示:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (3)$$

$$RPD = \frac{SD}{RMSE} \quad (4)$$

式中, y_i 表示第 i 个小麦样品的 AFB₁ 真实值, \hat{y}_i 表示第 i 个小麦样品的 AFB₁ 预测值, \bar{y} 表示真实值的平均值, SD 表示所有真实值 y_i 的标准差, n 为所有样本数。

2 结果与分析

2.1 小麦样品中 AFB₁ 的含量测定

采用高效液相色谱法测定得到 40 个污染样品的 AFB₁ 含量化学测量值的结果如图 2 所示, 5 个未污染样品含量为 0 $\mu\text{g}/\text{kg}$ 。高效液相色谱法的方法检出限(method detection limit, MDL)为 15.82 $\mu\text{g}/\text{kg}$, 污染样品的 AFB₁ 含量均高于 MDL。从图 2 中的中位线、均值可看出, 随着培养天数增加, 小麦 AFB₁ 含量呈现增加的趋势, 符合预期效果。

2.2 原始光谱分析

收集到样品的原始光谱如图 3a 所示, 原始光谱图存

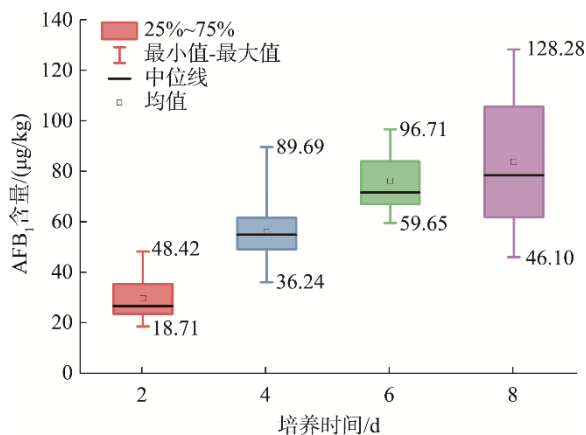


图 2 不同培养天数小麦中的 AFB₁ 含量
Fig.2 AFB₁ content in *Triticum aestivum* L. at different days of incubation

在噪声多、不够平滑并且基线不稳的问题^[33]。因此本研究将原始光谱图进行平滑处理, 能够去除噪声和基线漂移, 增强数据平稳性, 如图 3b 所示。同时从图 3c 中的平均光谱图能够看出, 未霉变籽粒的吸光度明显高于霉变籽粒, 随着霉变程度的加剧, 吸光度降低。不同霉变程度小麦的平均光谱之间有明显差别。这是由于在霉变过程中, 脂质被氧化, 蛋白质和其他营养物质则被降解, 这种变化引起了 C-H、N-H、O-H 等化学键的变化, 从而影响吸光度^[34]。为了消除光谱强度差异, 增强特征对比度还需要对平滑处理后的数据进行预处理。

2.3 光谱预处理方法分析

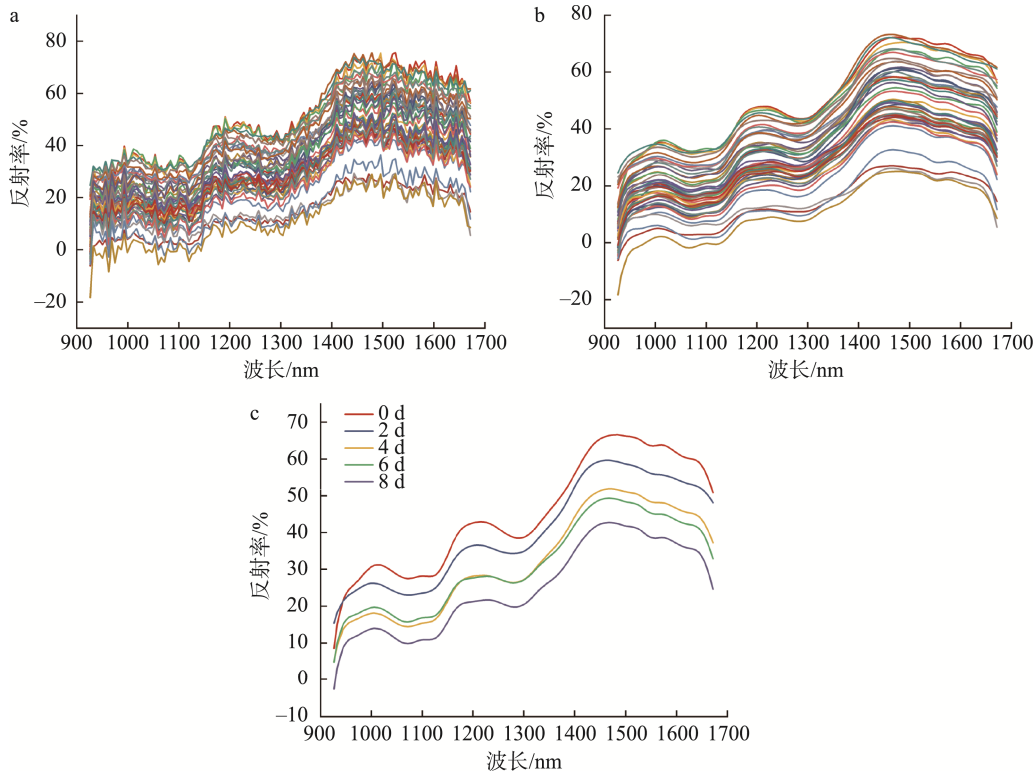
光谱经 MSC、SNV、MD、FD 处理后的结果如图 4 所示, 再分别将未经过预处理的光谱信息, 经 MSC、SNV、MD、FD 处理的光谱信息和 AFB₁ 含量结合, 建立 BPNN、RF、SVM 回归预测模型, 结果如表 4 所示。由表 4 可知, 光谱数据经过 MSC 处理后, 在 BPNN、RF 和 SVM 模型中都得到了优化, 并且优于其他预处理方法的建模结果指标。其他方法如 SNV 和 MD 虽然在训练集上表现良好, 但在测试集上的泛化能力不足, FD 的性能则显著较差。因此选择 MSC 处理光谱数据进行后续研究。

2.4 光谱特征波长的提取

特征提取的目的是从原始光谱数据中筛选出最具信息量的特征, 以减少冗余和降低数据维度, 从而提升模型性能和解释能力。本研究使用以下两种方法分别对经 MSC 处理后的光谱信息进行特征提取。

2.4.1 SPA 提取结果

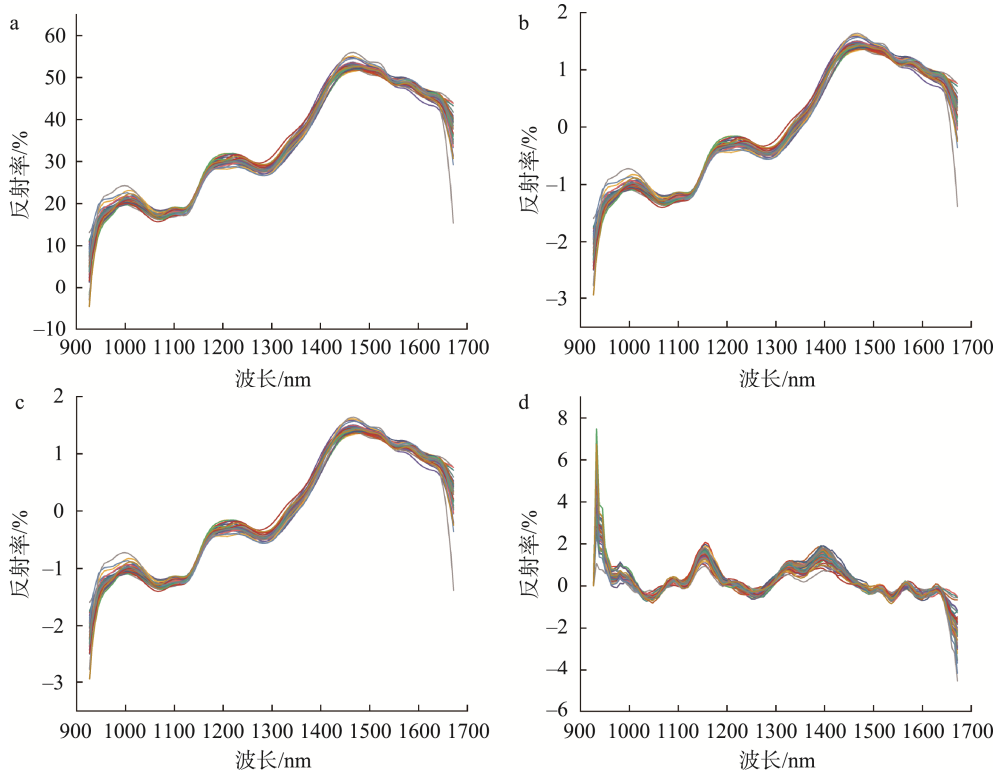
SPA 提取结果如图 5 所示, 由图 5a 可看出, 随着变量数的增加, RMSE 呈下降趋势, 在变量数 0 到 15 之间下降趋势较快, 变量数为 15 时最小(RMSE=27.9091), 在 15 之后趋于平缓, 说明变量数大于 15 时, RMSE 变化无明显差异, 最终选取 15 个变量用作回归预测模型的建立。选择的特征波长在全波段的分布情况如图 5b 所示。



注: a. 样品原始光谱图; b. 经平滑处理后的光谱图; c. 平均光谱图。

图 3 小麦样品光谱曲线

Fig.3 Spectral curves of *Triticum aestivum* L. samples



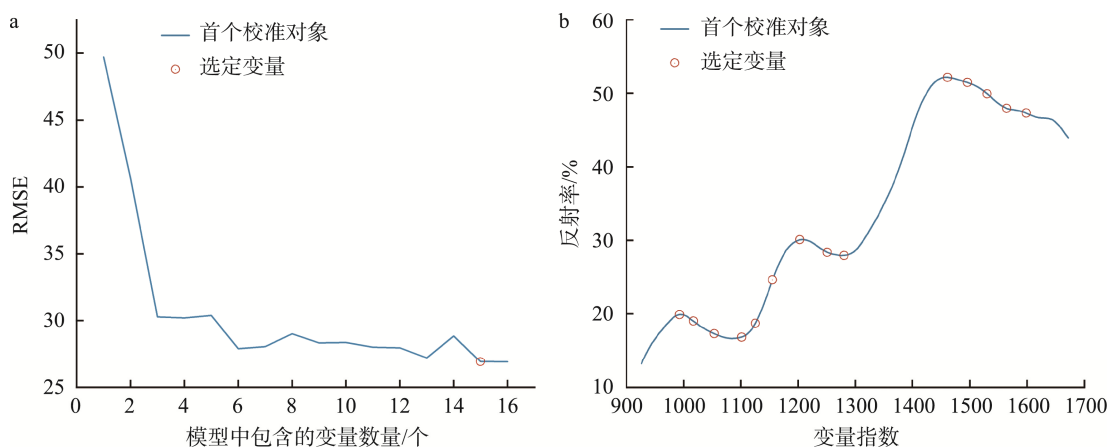
注: a. 经 MSC 处理后的光谱图; b. 经 SNV 处理后的光谱图; c. 经 MD 处理后的光谱图; d. 经 FD 处理后的光谱图。

图 4 小麦样品光谱预处理结果

Fig.4 Results of spectral pre-processing of *Triticum aestivum* L. samples

表 4 不同预处理方法下 3 种模型的统计结果
Table 4 Statistical results of 3 kinds of models with different preprocessing methods

| 模型 | 方法 | 训练集 | | | | 测试集 | | | |
|------|-----|----------------|---------|---------|---------|----------------|---------|---------|--------|
| | | R ² | RMSE | MAE | RPD | R ² | RMSE | MAE | RPD |
| BPNN | 无 | 0.6066 | 19.8077 | 14.3703 | 1.6207 | 0.1159 | 29.4942 | 21.9066 | 1.1037 |
| | MSC | 0.7827 | 15.4047 | 10.0418 | 2.1807 | 0.3407 | 23.2480 | 19.6825 | 1.2781 |
| | SNV | 0.7162 | 16.4689 | 11.3797 | 1.9080 | 0.2311 | 29.2436 | 19.9858 | 1.1835 |
| | MD | 0.7654 | 14.7944 | 6.6103 | 2.0989 | 0.1405 | 30.0240 | 25.8817 | 1.1194 |
| | FD | 0.5978 | 19.2963 | 10.3953 | 1.6029 | 0.1498 | 31.4603 | 23.9068 | 1.1254 |
| RF | 无 | 0.6384 | 18.0416 | 12.5366 | 1.6904 | 0.2420 | 30.6508 | 24.4341 | 1.1919 |
| | MSC | 0.7451 | 15.2823 | 11.7977 | 2.0134 | 0.4187 | 22.0455 | 19.0727 | 1.3611 |
| | SNV | 0.8208 | 14.1392 | 11.2590 | 2.4011 | 0.2212 | 24.2081 | 19.1020 | 1.1759 |
| | MD | 0.8097 | 13.8763 | 10.6132 | 2.3302 | 0.2215 | 26.7742 | 21.1363 | 1.1761 |
| | FD | 0.7573 | 15.1620 | 10.9401 | 2.0634 | 0.1825 | 29.3754 | 24.4149 | 1.1477 |
| SVM | 无 | 0.9180 | 9.3103 | 2.8784 | 3.5505 | 0.1158 | 28.0837 | 17.8915 | 1.1036 |
| | MSC | 0.9986 | 1.2601 | 1.2467 | 27.1807 | 0.48583 | 19.2709 | 15.8376 | 1.4472 |
| | SNV | 0.9984 | 1.2588 | 1.2467 | 25.1120 | 0.1909 | 29.1010 | 22.5854 | 1.1537 |
| | MD | 0.9983 | 1.2373 | 1.2168 | 24.9824 | 0.2104 | 30.0901 | 27.6135 | 1.1678 |
| | FD | 0.9980 | 1.2567 | 1.2414 | 22.8174 | 0.0265 | 37.7159 | 33.3029 | 1.0518 |



注: a. 最终选定的变量数: 15 (RMSE=27.9091); b. 选定的特征波长。

图 5 SPA 筛选结果

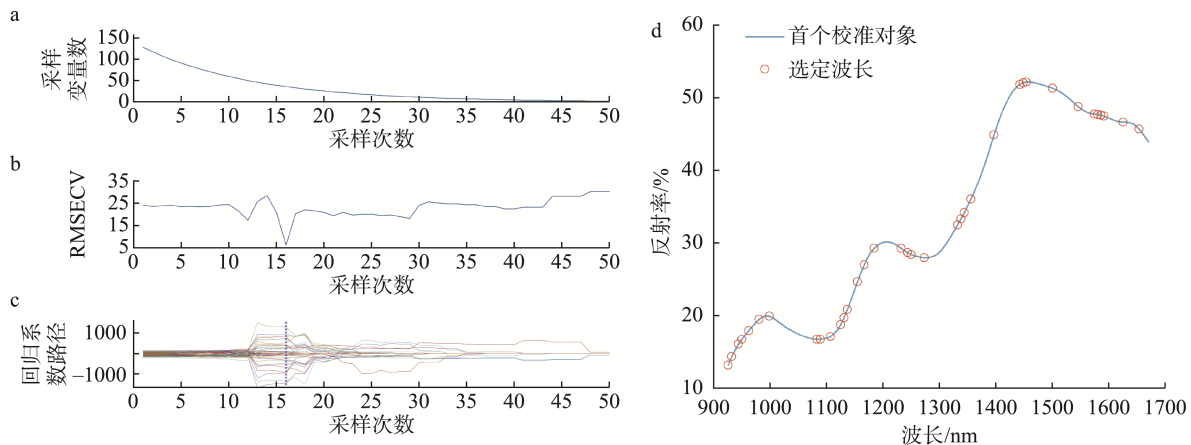
Fig.5 Screening results of SPA

2.4.2 CARS 提取结果

CARS 筛选特征波长结果如图 6 所示。从图 6a 和图 6b 可以看出, 前期抽样过程中, 变量数量逐渐减少, 同时均方根交叉验证误差 (root mean square error of cross-validation, RMSECV) 也在降低。RMSECV 是交叉验证中的关键指标, RMSECV 值越小, 表示模型对未知样本的预测能力越强。当抽样次数达到 16 时, RMSECV 降至最低值 6.4416, 此后 RMSECV 开始波动上升, 这表明此时剔除的变量包含了相关信息。由图 6c 可知, 最终选择 36 个特征波长。选择的特征波长在全波段的分布情况如图 6d 所示。

2.5 不同特征波长提取算法下的小麦 AFB₁ 含量预测模型建立

将 MSC-无(全波段建模)、MSC-SPA 和 MSC-CARS 分别建立 BPNN、RF 和 SVM 3 种预测模型, 分别得到训练集和预测集的 R²、RMSE、MAE 和 RPD 结果如表 5 所示。由表 5 可知, 两种特征波长提取算法相比全波段建模的性能有不同程度的提升, 通过 SPA 特征提取, BPNN 和 SVM 的预测能力有显著改善。CARS 特征提取对所有模型的性能都有提升, 其中 RF 和 SVM 的表现较 SPA 特征提取显著提高。总体而言, SVM 在所有特征提取方法中表现最优。



注: a. 波长保留趋势; b. RMSECV(交叉验证中的 RMSE)变化趋势; c. 回归系数路径; d. 选定的特征波长。

图 6 CARS 筛选结果

Fig.6 Screening results of CARS

表 5 不同特征提取方法下 3 种模型的统计结果

Table 5 Statistical results of 3 kinds of models with different feature extraction methods

| 处理方法 | 模型 | 训练集 | | | | 测试集 | | | |
|----------|------|--------|---------|---------|---------|--------|---------|---------|--------|
| | | R^2 | RMSE | MAE | RPD | R^2 | RMSE | MAE | RPD |
| MSC-无 | BPNN | 0.7827 | 15.4047 | 10.0418 | 2.1807 | 0.3407 | 23.248 | 19.6825 | 1.2781 |
| | RF | 0.7451 | 15.2823 | 11.7977 | 2.0134 | 0.4187 | 22.0455 | 19.0727 | 1.3611 |
| | SVM | 0.9986 | 1.2601 | 1.2467 | 27.1807 | 0.4858 | 19.2709 | 15.8376 | 1.4472 |
| MSC-SPA | BPNN | 0.8203 | 11.7795 | 9.0045 | 2.3977 | 0.4902 | 27.0661 | 21.3260 | 1.4534 |
| | RF | 0.7134 | 17.2045 | 13.8782 | 1.8988 | 0.4994 | 21.7225 | 17.1625 | 1.4668 |
| | SVM | 0.9809 | 4.8236 | 2.4971 | 7.3594 | 0.7571 | 11.0854 | 7.2084 | 2.1057 |
| MSC-CARS | BPNN | 0.8203 | 11.7795 | 9.0045 | 2.3977 | 0.4902 | 27.0661 | 21.3260 | 1.4534 |
| | RF | 0.7807 | 16.1473 | 12.9482 | 2.1706 | 0.6289 | 14.1473 | 12.1905 | 1.7036 |
| | SVM | 0.9977 | 0.0314 | 0.0293 | 21.0874 | 0.9426 | 0.1509 | 0.1160 | 4.3302 |

2.6 最优模型对小麦中 AFB₁ 含量的预测与分析

NIRS 数据经平滑处理和 MSC-CARS 后, SVM 模型在训练集上的 R^2 值为 0.9977, RMSE 为 0.0314, MAE 为 0.0293, RPD 值为 21.0874, 反映了模型在训练集上的预测能力很强。测试集的 R^2 值为 0.9426, RMSE 为 0.1509, MAE 为 0.1160, RPD 值为 4.3302, 表明模型在新数据上的拟合能力仍然很高, 预测集的 $RPD > 2.0$, 具有良好的性能。这些结果表明, 使用 NIRS 技术结合 SVM 回归模型来预测小麦中的 AFB₁ 含量是可行且有效的。模型在训练集和预测集上均表现出了较高的预测精度和可靠性, 在实际应用中, 模型能够提供准确的毒素含量预测。

3 结论

本研究通过采集被黄曲霉菌污染时间不同的小麦的 NIRS 信息和测定样品中 AFB₁ 含量化学值, 通过不同预处理方法、特征波长提取算法来构建不同 AFB₁ 定量预测模型, 从中选择较好的预处理方法、特征波长提取算法以及定量预测模型, 并对部分样品进行预测, 最后得到的定量预测模型测试集和训练集的相关系数均大于 0.90, 表明本

研究建立的 AFB₁ 含量预测模型能够对未知样品进行检测。总体来说, NIRS 无损检测方法在监测小麦中黄曲霉毒素含量方面具有显著的应用潜力和实际价值。

参考文献

- [1] WANG LY, ZHENG YS, DUAN LL, *et al.* Artificial selection trend of wheat varieties released in Huang-Huai-Hai Region in China evaluated using DUS testing characteristics [J]. *Frontiers in Plant Science*, 2022, 13: 898102.
- [2] 李欢, 梁晓艳, 张君, 等. 镰孢菌毒素的主要类型及其收获前后的生物防控方法[J]. *食品安全质量检测学报*, 2021, 12(9): 3431-3440.
LI H, LIANG XY, ZHANG J, *et al.* Main types of fusarium toxins and their biocontrol control methods from preharvest to postharvest [J]. *Journal of Food Safety & Quality*, 2021, 12(9): 3431-3440.
- [3] 孙宝胜. 新收获小麦中玉米赤霉烯酮监测与分析[J]. *现代面粉工业*, 2021, 35(1): 27-28.
SUN BS. Monitoring and analysis of zearalenone in newly harvested wheat [J]. *Modern Flour Milling Industry*, 2021, 35(1): 27-28.
- [4] 戴宝玲, 杨华, 戴贤君, 等. 小麦真菌结构分析及其与呕吐毒素含量的关系[J]. *江苏农业科学*, 2019, 47(12): 228-232.
DAI BL, YANG H, DAI XJ, *et al.* Analysis of wheat fungi structure and its relationship with vomiting toxins content [J]. *Jiangsu Agricultural Sciences*, 2019, 47(12): 228-232.
- [5] DU Y, YANG MY, ZHANG YM, *et al.* Development of a bioluminescent immunoassay based on Fc-specific conjugated antibody-nanoluciferase immunoreagents for determining aflatoxin B₁ [J]. *Food Chemistry*, 2024,

- 463 (Pt 2): 141220.
- [6] AMIR MB, IRFAN A. Preserving wheat grain quality and preventing aflatoxin accumulation during storage without pesticides using dry chain technology [J]. *Environmental Science and Pollution Research International*, 2020, 27(33): 1–8.
- [7] LOI M, FANELLI F, LIUZZI V, *et al.* Mycotoxin biotransformation by native and commercial enzymes: Present and future perspectives [J]. *Toxins*, 2017, 9(4): 111.
- [8] KADEMI IH, BABA AI, SAAD TF. Modelling the dynamics of toxicity associated with aflatoxins in foods and feeds [J]. *Toxicology Reports*, 2017, 4: 358–363.
- [9] WU WL, XIE S, ZHAO M, *et al.* Colorimetric liquid crystal-based assay for the ultrasensitive detection of AFB₁ assisted with rolling circle amplification [J]. *Analytica Chimica Acta*, 2022, 1220: 340065.
- [10] WU SW, XU JT, CHEN WX, *et al.* Protein nanoscaffold enables programmable nanobody-luciferase immunoassembly for sensitive and simultaneous detection of aflatoxin B₁ and ochratoxin A [J]. *Journal of Hazardous Materials*, 2024, 462: 132701.
- [11] WEI M, ZHAO F, XIE Y. A novel gold nanostars-based fluorescent aptasensor for aflatoxin B₁ detection [J]. *Talanta*, 2020, 209: 120599.
- [12] ZHENG SY, WEI ZS, LI S, *et al.* Near-infrared reflectance spectroscopy-based fast versicolorin A detection in maize for early aflatoxin warning and safety sorting [J]. *Food Chemistry*, 2020, 332: 127419.
- [13] 沈飞, 魏颖琪, 张斌, 等. 稻谷有害霉菌侵染的近红外光谱快速检测[J]. *光谱学与光谱分析*, 2018, 38(12): 3748–3752.
- SHEN F, WEI YQ, ZHANG B, *et al.* Rapid detection of harmful mold infection in rice by near infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2018, 38(12): 3748–3752.
- [14] 温冯睿, 关海鸣, 马晓丹, 等. 近红外光谱图像处理霉变稻谷检测方法[J]. *光谱学与光谱分析*, 2022, 42(2): 428–433.
- WEN FR, GUANG HOU, MA XD, *et al.* Moldy rice detection method based on near infrared spectroscopy image processing technology [J]. *Spectroscopy and Spectral Analysis*, 2022, 42(2): 428–433.
- [15] DENG JH, JIANG H, CHEN QS. Determination of aflatoxin B₁ (AFB₁) in maize based on a portable Raman spectroscopy system and multivariate analysis [J]. *Spectrochimica Acta. Part A, Molecular and Biomolecular Spectroscopy*, 2022, 275: 121148.
- [16] 任菲, 于治国, 陆峰, 薏苡仁中黄曲霉毒素 G₁ 的动态表面增强拉曼光谱检测[J]. *分析测试学报*, 2021, 40(4): 612–616.
- REN F, YU ZG, LU F. Detection of aflatoxin G₁ in *Coix* seed by dynamic surface-enhanced Raman spectroscopy [J]. *Journal of Instrumental Analysis*, 2021, 40(4): 612–616.
- [17] 李艳肖, 胡雪桃, 张芳, 等. 基于高光谱技术的菌落图像分割与计数[J]. *农业工程学报*, 2020, 36(20): 326–332.
- LI YX, HU XT, ZHANG F, *et al.* Colony image segmentation and counting based on hyperspectral technology [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2020, 36(20): 326–332.
- [18] GAYATRI M, KUMAR BP, ARIZA WR, *et al.* Application of SWIR hyperspectral imaging coupled with chemometrics for rapid and non-destructive prediction of aflatoxin B₁ in single kernel almonds [J]. *LWT*, 2022. DOI: 10.1016/J.LWT.2021.112954
- [19] 韩仲志, 刘杰. 高光谱亚像元分解预测花生中的黄曲霉毒素 B₁[J]. *中国食品学报*, 2020, 20(3): 244–250.
- HAN ZJ, LIU J. Detecting aflatoxin B₁ in peanuts by hyperspectral subpixel decomposition [J]. *Journal of Chinese Institute of Food Science and Technology*, 2020, 20(3): 244–250.
- [20] 杨承霖, 刘嘉祺, 郭芸成, 等. 结合太赫兹光谱与机器学习的小麦霉变程度判别[J]. *食品科学*, 2023, 44(12): 343–350.
- YANG CL, LIU JQ, GUO YC, *et al.* Detection of mildew degree of wheat using terahertz spectroscopy and machine learning [J]. *Food Science*, 2023, 44(12): 343–350.
- [21] PASQUINI C. Near infrared spectroscopy: A mature analytical technique with new perspectives—A review [J]. *Analytica Chimica Acta*, 2018, 1026: 8–36.
- [22] 凡霞. 酶转化淀粉接枝单体的表面施胶剂的合成[D]. 南京: 南京林业大学, 2010.
- FAN X. Synthesis of enzyme-converted starch grafted monomers surface sizing agent [D]. Nanjing: Nanjing Forestry University, 2010.
- [23] MARIN DCP, GARRIDDO-VARO A, GUERRERO JE, *et al.* External validation and transferability of NIRS models developed for detecting and quantifying MBM in intact compound feeding stuffs [J]. *Journal of Food Quality*, 2008, 31(1): 96–107.
- [24] SALGUERO-CHAPARRO L, BAETEN V, ABBAS O, *et al.* On-line analysis of intact olive fruits by vis-NIR spectroscopy: Optimisation of the acquisition parameters [J]. *Journal of Food Engineering*, 2012, 112(3): 152–157.
- [25] 易永艳, 李德荣, 张蕴薇, 等. 近红外技术在真菌生物量和真菌毒素检测中的研究与应用[J]. *光谱学与光谱分析*, 2009, 29(7): 1826–1829.
- YI YY, LI DR, ZHANG YW, *et al.* Applications of near infrared reflectance spectroscopy in detecting chitin, ergosterol and mycotoxins [J]. *Spectroscopy and Spectral Analysis*, 2009, 29(7): 1826–1829.
- [26] 曾子琦, 蒋立文, 刘霞, 等. 近红外光谱无损检测在食用油脂分析中的应用研究进展[J]. *中国油脂*, 2018, 43(8): 137–142.
- ZENG ZQ, JIANG LW, LIU X, *et al.* Advance in application of near infrared spectroscopy in nondestructive analysis of edible oil [J]. *China Oils and Fats*, 2018, 43(8): 137–142.
- [27] YANG YF, ZHAO CL, CHEN XD, *et al.* Development of a bioluminescent immunoassay based on Fc-specific conjugated antibody-nanoluciferase immunoreagents for determining aflatoxin B₁ [J]. *Food Chemistry*, 2024, 463 (Pt 2): 141220.
- [28] JIANG H, WANG JN, CHEN QS. Comparison of wavelength selected methods for prediction of prediction performance of PLS model to determine aflatoxin B₁ (AFB₁) in wheat samples during storage [J]. *Microchemical Journal*, 2021, 1: 170.
- [29] ZHANG J, WANG Y, NIU F, *et al.* Parameterization of multi-angle shaker based on PSO-BP neural network [J]. *Minerals*, 2023, 13(7): 929.
- [30] 甄贺. 基床系数的反演分析及其在工程中的应用研究[D]. 石家庄: 石家庄铁道大学, 2018.
- ZHEN H. Inverse analysis of coefficient of subgrade reaction and its application in engineering [D]. Shijiazhuang: Shijiazhuang Railway University, 2018.
- [31] 廖玉琼. 基于图神经网络的蛋白质突变效应预测研究[D]. 广州: 广州大学, 2024.
- LIAO YQ. Protein mutation effect prediction based on graph neural network [D]. Guangzhou: Guangzhou University, 2024.
- [32] 罗正飞, 龚正礼, 杨坚, 等. 基于近红外的多糖红茶快速无损检测方法[J]. *光谱学与光谱分析*, 2023, 43(8): 2649–2656.
- LUO ZF, GONG ZL, YANG J, *et al.* Rapid non-destructive detection method for black tea with exogenous sucrose based on near-infrared spectroscopy [J]. *Spectroscopy and Spectral Analysis*, 2023, 43(8): 2649–2656.
- [33] 张强, 刘成海, 孙井坤, 等. 基于支持向量机的稻谷黄曲霉毒素 B₁ 近红外无损检测[J]. *东北农业大学学报*, 2015, 46(5): 84–88.
- ZHANG Q, LIU CH, SUN JK, *et al.* Near-infrared spectroscopy nondestructive determination of aflatoxin B₁ in paddy rice based on support vector machine regression [J]. *Journal of Northeast Agricultural University*, 2015, 46(5): 84–88.
- [34] SUN JF, WANG GX, ZHANG H, *et al.* Detection of fat content in peanut kernels based on chemometrics and hyperspectral imaging technology [J]. *Infrared Physics & Technology*, 2020, 105: 103226.

(责任编辑: 于梦娇 韩晓红)