

荧光光谱技术结合机器学习算法检测 大白菜中吡虫啉含量

陈 珏¹, 李佳琮^{1,2}, 刘翠玲^{1,2*}, 孙晓荣^{1,2}, 张善哲¹

(1. 北京工商大学人工智能学院, 北京 100048;
2. 北京工商大学北京市食品安全大数据技术重点实验室, 北京 100048)

摘要: **目的** 建立荧光光谱技术结合机器学习算法检测大白菜中吡虫啉含量的方法。**方法** 采集 400 nm 激发下的 130 个农药残留光谱数据, 经过数据预处理、光谱特征筛选, 构建基于支持向量机(support vector machine, SVM)的吡虫啉残留含量预测模型, 并利用麻雀搜索算法(sparrow search algorithm, SSA)对 SVM 的参数进行寻优。**结果** 卷积平滑(Savitzky-Golay smooth, S-G)与标准正态变量校正(standard normal variable, SNV)联用的预处理效果最好; 利用连续投影算法(successive projections algorithm, SPA)对遗传算法(genetic algorithm, GA)提取的特征波长进行二次特征降维能获得最优特征波段; SSA 寻优后构建的 SVM 模型精度最佳, 测试集决定系数为 0.9234, 均方根误差为 0.4129。**结论** 荧光光谱技术可以实现白菜中吡虫啉含量的检测, 为蔬菜中农药残留快速检测提供了新的思路。

关键词: 荧光光谱技术; 吡虫啉含量; 二次波段选择; 麻雀搜索算法; 支持向量机

Determination of imidacloprid in cabbage by fluorescence spectroscopy combined with machine learning algorithms

CHEN Jue¹, LI Jia-Cong^{1,2}, LIU Cui-Ling^{1,2*}, SUN Xiao-Rong^{1,2}, ZHANG Shan-Zhe¹

(1. School of Artificial Intelligence, Beijing Technology and Business University, Beijing 100048, China; 2. Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China)

ABSTRACT: Objective To establish a method for the determination of imidacloprid in cabbage by fluorescence spectroscopy combined with machine learning algorithm. **Methods** The spectral data of 130 agricultural residues under excitation at 400 nm were collected. After data preprocessing and spectral feature screening, the prediction model of imidacloprid residue content based on support vector machine (SVM) was constructed. The sparrow search algorithm (SSA) was used to optimize the parameters of SVM. **Results** The combination of Savitzky-Golay smooth (S-G) and standard normal variable (SNV) had the best pretreatment effect. The optimal feature band was obtained by secondary feature downscaling of genetic algorithm (GA) extracted feature wavelengths using the successive projections algorithm (SPA). The SVM model constructed after the SSA search achieved the best accuracy with a test set coefficient of determination of 0.9234 and a root mean square error of 0.4129. **Conclusion** Fluorescence

基金项目: 北京市自然科学基金项目(4222043)、北京工商大学 2023 科研能力提升计划项目

Fund: Supported by the Beijing Natural Science Foundation Project (4222043), and the Beijing Technology and Business University 2023 Research Capacity Enhancement Program

*通信作者: 刘翠玲, 教授, 主要研究方向为智能测量技术与数据处理研究。E-mail: liucl@btbu.edu.cn

*Corresponding author: LIU Cui-Ling, Professor, Beijing Key Laboratory of Big Data Technology for Food Safety, Beijing Technology and Business University, Beijing 100048, China. E-mail: liucl@btbu.edu.cn

spectroscopy enables the detection of imidacloprid in cabbage and provides a new idea for the rapid detection of pesticide residues in vegetables.

KEY WORDS: fluorescence spectroscopy; imidacloprid content; secondary band selection; sparrow search algorithm; support vector machine

0 引言

吡虫啉作为新烟碱类农药是目前我国使用量最大的杀虫剂之一, 具有触杀、胃毒、趋避等多重作用, 可以防治蚜虫、叶蝉等害虫, 但同时, 也会对非靶标生物产生毒害作用^[1-2]。不合理的施用使其残留在果蔬等食物和环境中, 会导致其在人体中富集, 危害人体健康^[3]。因此, 采用合理有效的检测方法检测农产品中吡虫啉的含量并对其进行施用管理对于人类健康至关重要。对于果蔬中农药残留量的检测, 国内外常用的色谱技术有高效液相色谱法(high performance liquid chromatography, HPLC)、气相色谱法(gas chromatography, GC)、薄层色谱法(thin layer chromatography, TLC)、超临界流体色谱法(supercritical fluid chromatography, SFC)等^[4], 这些技术虽然得到了广泛应用, 但是色谱法定性能力差, 且费用昂贵, 操作复杂, 容易受到样品浓度的影响, 从而导致检测结果不准确。

近年来出现的光谱检测技术, 广泛应用于农药残留物的检测当中。刘磊等^[5]以金纳米溶胶作为拉曼光谱增强基底对山药中倍硫磷、三唑磷、福美双农药残留进行了研究, 建立了决定系数 0.9 以上的检测模型。DU 等^[6]构建了梨表面毒死蜱农药残留含量的随机森林(random forest, RF)模型, 预测决定系数达 0.849。吕赫一^[7]采用近红外光谱技术来检测苹果中的农药残留和可溶性固形物含量, 分别建立了偏最小二乘、支持向量机(support vector machine, SVM)两种苹果可溶性固形物预测模型, 预测的相关系数均达 0.9 以上。CAO 等^[8]采用太赫兹时域谱法检测面粉中多菌灵残留含量, 预测集决定系数达 0.9 以上。然而, 针对农药残留样本的微量检测, 表面增强拉曼光谱大多需要借助添加增强剂等辅助手段, 仪器操作复杂, 大多费时费力且检测的灵敏度不高; 此外, 由于红外光谱易出现谱带重叠且水分子的红外信号极易覆盖其他待检信号, 因此不适宜对白菜等含水量大的蔬菜进行检测。

荧光光谱检测方法凭借其操作简单、快捷环保、灵敏度高、且分析速度快的优点, 在农产品检测领域得到了迅速的发展^[9]。HAO 等^[10]以 280 nm 作为荧光激发波长, 构建了多种蜂蜜种类的判别模型。NTAKATSANE 等^[11]利用荧光技术对牛奶中 3 种抗生素种类成功判别, 并构建了决定系数 0.9 以上的含量检测模型。JI 等^[12]确定了哒螨灵的荧光特征峰位于 356 nm 处并成功构建了该农药的含量预测模型。然而, 荧光光谱技术对农药的检测仍处于浅层探

索阶段, 现有研究多停留于纯农药溶液的检测, 为扩充该技术在农药残留检测领域的研究深度, 以实际蔬菜作为背景, 探究荧光技术在农药残留检测领域的可行性具有一定研究价值^[13]。

鉴于此, 本研究以白菜中吡虫啉农药残留为研究对象, 首先利用荧光光谱技术采集不同农药残留样本的光谱数据, 分析白菜中农药残留样本的光谱特征和农药残留浓度变化, 然后对白菜中农药残留的发射光谱数据进行预处理和特征波长筛选, 采用 SVM 建立白菜中农药残留含量的最优预测模型, 并采用群体智能优化算法对模型中的参数进行寻优, 最终实现对白菜中农药残留的快速检测, 为白菜的品质检测提供依据和技术支撑。

1 材料与方法

1.1 样本制备

选用于深圳诺普信农化股份有限公司的可溶性吡虫啉农药, 农药含量为 20%。将市购的普通大白菜用去离子水洗净晾干后粉碎, 并经过滤取其汁液为背景溶液, 按照 GB 2763—2021《食品安全国家标准 食品中农药最大残留限量》对大白菜中吡虫啉(最大残留量 0.2 mg/kg)的限定, 配制 0~5 mg/kg 的农药残留溶液, 总计 13 个浓度梯度, 每个梯度 10 个样本, 共 130 个。该质量分数范围分布都在 GB 2763—2021 最大残留量附近, 具有实际意义。

1.2 实验仪器及其采集参数

使用英国爱丁堡 FS5 荧光光谱仪, 设置激发波长间隔为 10 nm, 发射波长间隔为 2 nm, 激发波长 Ex 范围为 350~550 nm, 发射波长 Em 为 400~600 nm。荧光光谱的激发光源选用脉冲氙灯, 光电检测器选用 PMT-900 光电倍增管。三维荧光光谱采集, 荧光发射光谱采集, 设置采集步长为 3 nm, 采集范围为 430~610 nm。

1.3 数据分析方法

1.3.1 样本划分

在模型校准中, 样本被分为校正集和预测集, 使用 Kennard-Stone (K-S)法进行样本集的划分。该方法把所有的样本都看作校正集候选样本, 依次从中挑选样本进校正集^[14]。首先选择欧氏距离最远的两个样本进入校正集, 其后通过计算剩下的每一个样品到校正集内每一个已知样品的欧式距离, 找到距已选样本最远以及最近的两个样本, 并将这两个样本选入校正集, 重复上述步骤直到样本数量达到要求。

1.3.2 光谱预处理

采集光谱时,由于受到环境温湿度、仪器状态以及样本状态等各种因素的影响,会产生噪声信号以及杂散信息等情况,因此需要对光谱进行预处理,以消除干扰,提升模型稳健性^[15]。在全光谱范围内使用卷积(Savitzky-Golay, S-G)平滑、标准正态变换(standard normal variable, SNV)以及有限脉冲响应(finite impulse response, FIR) 3 种方法对原始光谱数据进行预处理。S-G 平滑法通过多项式来对移动窗口内的数据进行多项式最小二乘拟合,当平滑窗口宽度选取恰当时,可有效降低检测时仪器所引入的光谱数据噪声^[16]。SNV 可以用来校正样品间因散射而引起的光谱误差,通过单独对每个样本的光谱进行校正,使其每个波段的吸光度均符合正态分布^[17]。FIR 通过加权叠加的计算方式滤除光谱颗粒分布不均的噪声,留下特征增强后的光谱数据^[18]。

1.3.3 光谱特征提取

由于光谱信息中通常会包含冗余的变量,有时需要对光谱区间进行筛选以获取关键的信息。合适的变量筛选方法有利于从光谱数据中获取最有用的信息,舍弃不必要的信息,提高模型的性能^[19]。文章采用连续投影算法(successive projections algorithm, SPA)和遗传算法(genetic algorithm, GA)筛选出能够表征白菜中农药含量变化的特征波长。SPA 是在数据矩阵中应用变量投影操作寻找含冗余信息最低、共线性最少的光谱特征变量组,从而提高模型的稳定性和准确性。利用向量的投影分析,以某一单波长作为起点,循环计算未选入波长上的投影,使其寻找到在光谱信息中含有最少冗余信息的变量组,最终经 SPA 选出的波长都与前一个共线性最小^[20]。GA 是基于自然遗传和自然选择过程建立的一种优化算法。应用 GA 对建模波长进行优化选择是基于该算法能够不断地对模型进行简化以及有利于提高建模效果的波段或者波长点进行优化组合,从而达到最佳的预测效果^[21-22]。

1.3.4 麻雀搜索算法优化 SVM

SVM 是一种根据有限的样本信息在模型的复杂性和学习能力之间寻求最佳折中的理论,具备解决非线性、高维模式识别、及小样本等问题优势^[23]。其通过一个非线性变换将输入空间对应于一个特征空间,使得在输入空间中的决策超曲面模型对应于特征空间中的决策超平面模型。但是, SVM 理论在选取模型参数方面也具有一定弱点,惩罚系数 C 和核函数 σ 等相关参数对预测结果起着决定性作用,选取不慎则会影响预测精度,实践中依赖通过多轮实验和人工选取参数,容易偏向主观判断和经验主义,而且需要大量时间^[24-25]。

麻雀搜索算法(sparrow search algorithm, SSA)是一种新型的群智能优化算法,该算法主要是受麻雀觅食行为和逃避捕食者行为启发而设计^[26]。一个麻雀种群中存在着发现者、加入者和侦察者 3 种个体,发现者在搜索空间中拥有较好的位置并带领发现者觅食,侦察者随机产生,负责

侦察预警。算法起源于一个由 n 只麻雀组成的种群,种群 X 可由公式(1)所示矩阵表示:

$$X = \begin{pmatrix} x_1^1 & x_1^2 & \dots & x_1^d \\ x_2^1 & & \ddots & x_2^d \\ \vdots & & & \vdots \\ x_n^1 & x_n^2 & \dots & x_n^d \end{pmatrix} \quad (1)$$

其中, d 表示待优化问题变量的维数, n 则是麻雀的数量。那么,所有麻雀的适应度值可以表示为如公式(2)所示:

$$F_x = \begin{bmatrix} f([x_1^1 & x_1^2 & \dots & x_1^d]) \\ f([x_2^1 & x_2^2 & \dots & x_2^d]) \\ \dots & \dots & \dots & \dots \\ f([x_n^1 & x_n^2 & \dots & x_n^d]) \end{bmatrix} \quad (2)$$

其中, f 表示适应度值。同近年来新兴的群智能优化算法相比,麻雀搜索算法具有较好的全局搜索和局部开发的能力,在寻优过程中,能促进麻雀种群向全局最优值移动,有效避免易早熟收敛、收敛速度慢等缺点,具有良好的鲁棒性和收敛速度^[27]。

1.3.5 模型评价

选择决定系数(correlation coefficient of cross-validation, R^2)、均方根误差(root mean square error, RMSE)以及相对分析误差(relative percent deviation, RPD)作为白菜中吡虫啉农药残留模型的评价参数。其中, R_c^2 为校正集选择决定系数(correction set selection coefficient of determination, R_c^2), RMSEC 为校正集均方根误差(root-mean-square deviation of correction set, RMSEC), R_p^2 为测试集选择决定系数(test set selection coefficient of determination, R_p^2), RMSEP 为测试集均方根误差(root-mean-square deviation of test set, RMSEP), R^2 越接近 1, 表明荧光光谱信息与白菜中吡虫啉的相关性越好, RMSE 越小, 表明预测中产生的误差越小, 即光谱建模效果越好。RPD 小于 2.5, 表明模型预测能力差; RPD 在 2.5~3.0 之间, 表明模型预测能力较好; RPD 大于 3.0, 表明模型预测能力非常好^[28]。

1.4 数据处理

光谱数据处理及建模分析通过 Matlab 2020a 进行, 表格采用 Word 2021 软件制作, 图形采用 Origin 2021 绘制。

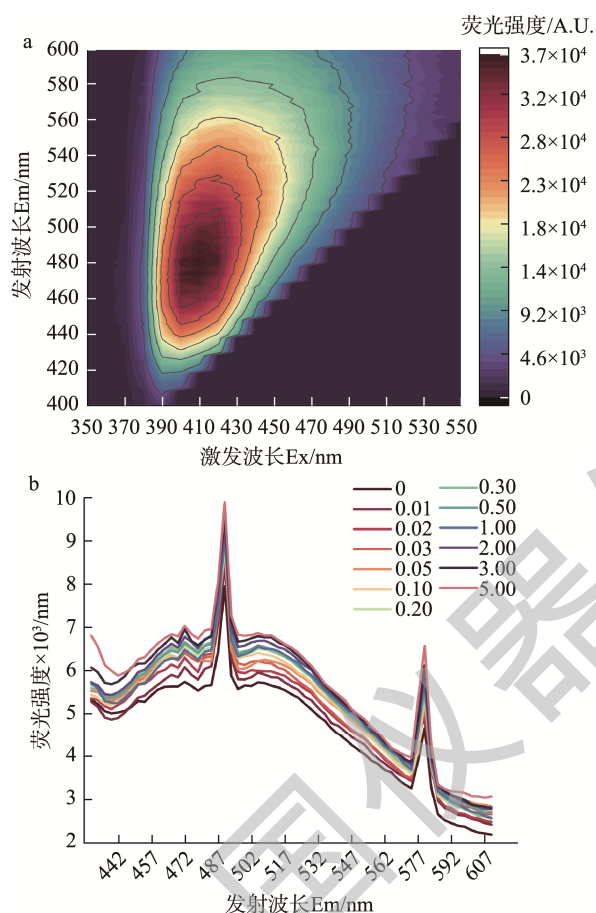
2 结果与分析

2.1 荧光光谱采集结果

为了得到吡虫啉农药的最佳激发波长, 首先需要扫描纯农药的三维荧光光谱, 绘制其对应的等高线图如图 1a 所示。可以看出, 在 $\lambda_{Ex}/\lambda_{Em}=400 \text{ nm}/480 \text{ nm}$ 存在明显荧光响应, 对应的荧光强度高达 36896 A.U., 因此可以确定出吡虫啉溶液的最佳激发波长在 400 nm 左右, 最佳发射波长在 480 nm 左右。

选用吡虫啉溶液的最佳激发波长 400 nm 扫描白菜中农

药残留溶液的荧光发射光谱, 共得到波段数为 61 维的光谱数据。为了更清晰地分析不同浓度农药残留样本的光谱差异, 将各浓度下的 10 组数据取平均, 并绘制成不同浓度变化对比曲线图, 如图 1b 所示。可以看出, 随着白菜汁中吡虫啉浓度的增加, 对应的荧光强度也随之增长。在 490 nm 和 580 nm 处分别存在两个尖峰, 对应表示为青光和黄光。通过分析发现荧光光谱信息与样本中农药含量存在规律变化, 因此理论上可以通过荧光数据表征白菜汁中吡虫啉的含量。



注: a 为纯吡虫啉农药溶液等高线图; b 为不同浓度农药样本的荧光强度图。

图 1 吡虫啉溶液的荧光强度图

Fig.1 Fluorescence intensity graph of imidacloprid solution

2.2 样本划分

将样本按照 2.5:1 的比例划分校正集和预测集。其中, 校正集与预测集分别包含 93 和 37 个样本用于构建白菜中吡虫啉含量预测模型。数据划分的具体情况如表 1 所示, 校正集的吡虫啉含量涵盖了预测集的范围, 说明校正集可以建立稳健的校正模型, 且预测集可以有效对模型进行预测。

2.3 数据预处理结果

采用 3 种预处理算法(SNV、S-G、FIR)对光谱数据进行处理。通过对比原始光谱与预处理后光谱的 SVM 建模

结果, 确定最佳预处理方法。表 2 为光谱数据的建模结果, 可以看出任意一种预处理都能提升模型精度, 其中 S-G 最能有效平滑数据, 相比原始光谱数据 R_p^2 提升了 0.0406。但由于 R_p^2 未达到 0.85, 还有提升空间, 尝试将 S-G 与其他预处理方法联用建模。由表 2 可以看出基于 S-G-SNV 与 S-G-FIR 预处理的建模 R_p^2 达 0.85 以上, 表明两种方法联用能够解决不同原因引起的干扰, 并有效提升建模精度。其中 S-G-SNV 的结果最佳, R_p^2 达 0.8639, RMSEP 达 0.5572。说明 S-G 能够平滑噪声并有效压缩数据集, SNV 能够校正样品间因散射而引起的光谱误差, 联用的效果最佳。因此选择 S-G-SNV 预处理后的光谱数据作为后续分析的基础。

表 1 样本的数据集划分
Table 1 Data set division of the sample

数据集	数量	白菜中吡虫啉含量/(mg/kg)		
		数据范围	平均含量	标准偏差
校正集	93	0.01~5.00	0.9196	1.463
预测集	37	0.01~5.00	0.9883	1.509

表 2 不同预处理的建模结果
Table 2 Modeling results with different pre-processing

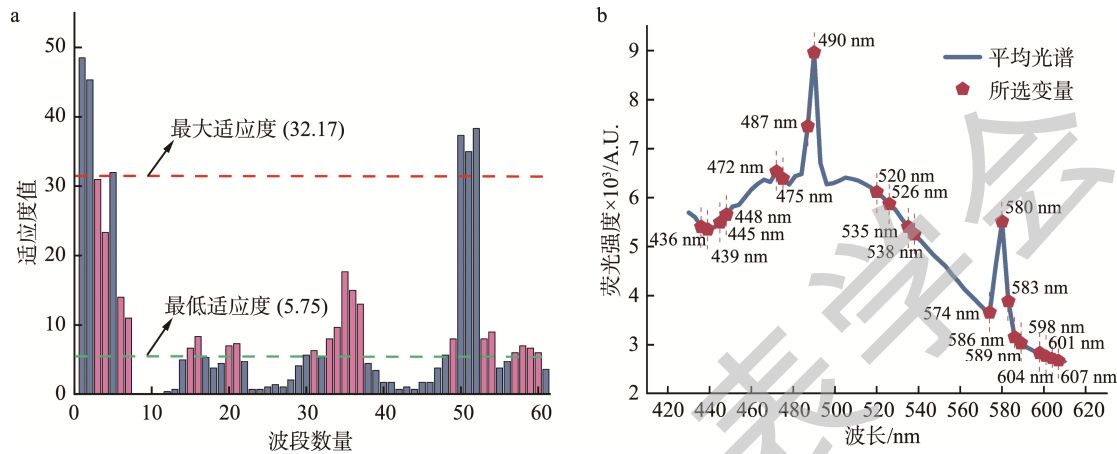
模型	校正集		测试集		
	R_c^2	RMSEC	R_p^2	RMSEP	RPD
无预处理	0.8384	0.5992	0.8093	0.6425	1.7024
SNV	0.8675	0.5504	0.8449	0.5914	1.8694
S-G	0.8712	0.5399	0.8499	0.5798	1.8989
FIR	0.8563	0.5771	0.8336	0.6021	1.8104
S-G-SNV	0.8847	0.5261	0.8639	0.5572	1.9855
S-G-FIR	0.8839	0.5297	0.8528	0.5739	1.9148

2.4 基于特征波长筛选的 SVM 建模结果

荧光光谱信息冗余和数据维度高导致了检测模型稳健性较差, 为解决农药残留样本光谱复杂问题、实现模型稳健监测, 基于数据量较小的有效波段进行建模是一种优先选择。GA 是一种进化算法, 随着迭代次数的增加, 每一代种群的适应度值逐渐收敛于局部最优解^[29]。图 2a 是 GA 算法计算出的每个光谱变量对应的适应度值, 两条水平虚线分别对应种群适应度的最大阈值(32.17)与最小阈值(5.75), 当适应度超出阈值区间, 剔除该光谱变量。图 2b 是 GA 选择出的特征变量在光谱曲线上的具体位置标记, 最终选择出 21 个特征波长。基于 GA 波长筛选算法的 SVM 的建模结果如表 3 所示。可以看出, 相较于全波段建模, 基于 GA 建模的 R_p^2 提升了 0.0168, RMSEP 减少了 0.0277, 表明特征波长筛选能够提升一定的精度。但由于 GA 算法筛选出的特征波长较多, 且建模结果精度不太理想, R_p^2 仍未达到 0.9, 进而采用 SPA 算法对其进行二次波段筛选。

SPA 通过计算原始光谱 130×61 矩阵中的任意一列向量在剩余列向量上的投影, 得到最大投影对应的特征波段数, 最终选择 RMSE 最小的对应的波长^[30]。图 3a 为 SPA 特征选择过程中, RMSE 的变化曲线, 当特征变量选择 8 维时, RMSE 最小且其后趋于平稳。图 3b 为 SPA 选择出的特征波长在光谱曲线上的具体位置标记。最终通过 SPA 选择了 8 个特征波

长。基于 GA-SPA 方法的 SVM 的建模结果如表 3 所示, 可以看到进行二次波长筛选后校正集和测试集的决定系数分别提升了 0.0176 和 0.0209, 测试集 R_p^2 达到了 0.9016, RMSEP 达到 0.4584, RPD 达到了 2.3118, 表明二次波长筛选有一定的效果。但由于 RPD 小于 2.5, 表明模型稳健性有待提升, 进而选择群体智能优化算法对 SVM 模型中的参数进行寻优。



注: a 为种群适应度计算过程图; b 为特征波长位置图。

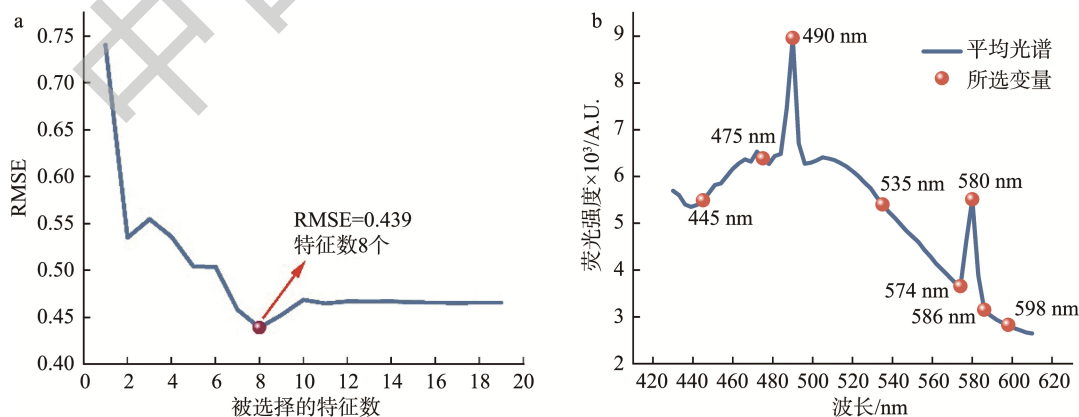
图 2 GA 选取特征波长过程图

Fig.2 Diagram of the GA selection process for feature wavelengths

表 3 SVM 建模结果对比表

Table 3 Comparison table of SVM modeling results

波长筛选方法	建模方法	特征数	校正集			测试集	
			R_c^2	RMSEC	R_p^2	RMSEP	RPD
无		61	0.8847	0.5261	0.8639	0.5572	1.9855
GA	SVM	21	0.9235	0.4102	0.8807	0.5295	2.1112
GA-SPA		8	0.9411	0.3482	0.9016	0.4584	2.3118
GA-SPA	SSA-SVM	8	0.9596	0.3275	0.9234	0.4129	2.6053



注: a 为选取的特征波长重要度排名; b 为特征波长位置图。

图 3 SPA 选取特征波长过程图

Fig.3 Diagram of the SPA selection process for feature wavelengths

2.5 基于 SSA 算法优化的 SVM 建模结果

为进一步提升吡虫啉含量预测模型的精度, 引入 SSA 算法对基于 GA-SPA 提取特征波长的 SVM 模型进行优化, 并继续比较分析。由表 3 可以看出, SSA-SVM 建模方法中模型 R_p^2 较 SVM 建模方法中模型的 R_p^2 提升了 0.0218, RMSEP 减少了 0.0455, RPD 增加了 0.2935, 这表明 SSA 算法优化后的模型具有更高的预测精度和可靠性。图 4 显示了最优模型的测试集结果, R_p^2 达 0.9234, RMSEP 达 0.4129。此外, 图中能够清晰地看出测试样本分布以及吡虫啉含量真实值和预测值的偏差, 当农药浓度低于按照 GB 2763—2021 规定 0.2 mg/kg 时, 预测值均分布在真实值附近, 预测偏差较小。当浓度远超出国标值时, 预测结果出现明显浮动, 模型稳定性有所下降, 但仍能正确判别出农残含量超标。分析 SSA 能够优化 SVM 的原因, 可能是由于该算法局部搜索能力较强^[31], 能够通过模拟麻雀捕食精准寻找到 SVM 的最优参数, 通过确定最佳惩罚系数 c 减少误差与过拟合, 调节核函数参数 g 影响数据映射到新的特征空间后的分布与训练速度。

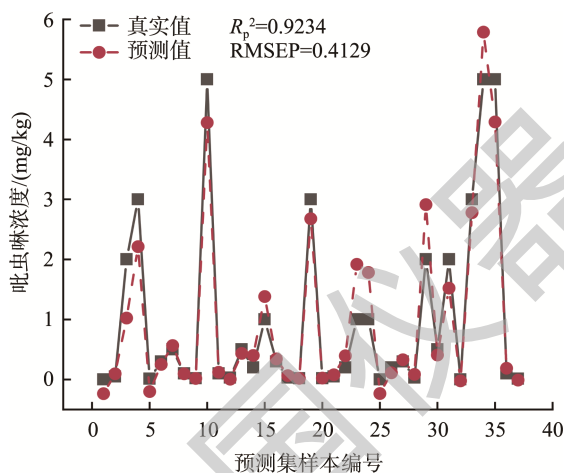


图 4 SSA-SVM 建模结果图

Fig.4 Diagram of SSA-SVM modeling results

3 结 论

本研究验证了荧光光谱技术检测白菜中农药残留含量的可行性, 制备并扫描出白菜汁中农药残留含量的荧光发射光谱数据后, 结合化学计量学算法构建了预测模型。结果表明, 利用三维荧光能够确定吡虫啉存在荧光特征峰; 相较于单一的预处理算法, S-G-SNV 联用能够消除不同原因产生的外部干扰, 预处理效果最佳; 此外, GA-SPA 二次波长筛选消除冗余变量的性能最佳, 有效剔除与建模相关度低的光谱变量后, 获得了 8 个特征波段; 为了提高模型精度, SSA 算法被用于对 SVM 的惩罚因子 c 和核函数参数 g 进行寻优, 相较于未经优化的 SVM 模型, 预测精度与稳

健度均得到提升, 最优结果的 R_p^2 达 0.9234, RMSEP 达 0.4129。该方法检测大白菜中吡虫啉含量预测精度达 0.9 以上, 验证了荧光光谱技术对农药残留含量检测的可行性。目前, 利用荧光光谱结合机器学习算法检测农药残留还属于探索阶段, 后续将检测更多的农药种类, 进一步提升模型精度, 尝试验证荧光光谱无损检测农药残留的可行性, 为保证农产品安全提供更为直观的针对性策略。

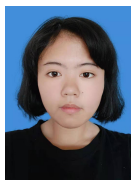
参考文献

- [1] 李成军, 郭梅燕, 李前进, 等. 吡虫啉对烟田东亚小花蝽的安全性评价[J]. 湖北农业科学, 2022, 61(24): 74-77.
LI CJ, GUO MY, LI QJ, *et al.* Safety assessment of imidacloprid against *Orius sauteri* in tobacco field [J]. Hubei Agric Sci, 2022, 61(24): 74-77.
- [2] 周兴藩, 孙谊然, 朱晓俊, 等. 基于非靶向代谢组学的吡虫啉暴露致人肝细胞 HepG2 毒性效应研究[J]. 环境与职业医学, 2023, 40(2): 216-223, 229.
ZHOU XF, SUN YR, ZHU XJ, *et al.* Study on the toxic effects of imidacloprid exposure on human liver cell HepG2 based on non targeted metabolomics [J]. Environ Occup Med, 2023, 40(2): 216-223, 229.
- [3] 张朝阳, 刘秀华, 王悦, 等. 吡虫啉在人体肝微粒体中的代谢产物和代谢途径研究[J]. 农产品质量与安全, 2022, 116(2): 42-46.
ZHANG CY, LIU XH, WANG Y, *et al.* Metabolites and metabolic pathways of imidacloprid in human liver microsomes [J]. Qual Saf Agro-Prod, 2022, 116(2): 42-46.
- [4] 李豪. 果蔬中农药残留检测方法研究[J]. 食品安全导刊, 2020, 289(30): 161.
LI H. Study of pesticide residue detection methods in fruits and vegetables [J]. China Food Saf Magaz, 2020, 289(30): 161.
- [5] 刘磊, 卞正兰, 董作人, 等. 山药中有机农药残留的表面增强拉曼光谱检测[J]. 激光与光电子学进展, 2022, 59(4): 355-362.
LIU L, BIAN ZL, DONG ZR, *et al.* Detection of residual organic pesticides in yam by surface enhanced Raman spectroscopy [J]. Laser Optoelectron Progress, 2022, 59(4): 355-362.
- [6] DU XF, WANG P, FU L, *et al.* Determination of chlorpyrifos in pears by Raman spectroscopy with random forest regression analysis [J]. Anal Lett, 2019, 53(6): 821-833.
- [7] 吕赫一. 基于近红外光谱的苹果品质快速无损检测方法研究[D]. 哈尔滨: 哈尔滨理工大学, 2021.
LV HY. Research on apple quality fast nondestructive testing method based on near infrared spectroscopy [D]. Harbin: Harbin Institute of Technology, 2021.
- [8] CAO BH, LI H, CAI E, *et al.* Determination of pesticides in flour by terahertz time-domain spectroscopy (THz-TDS) with voigt function fitting and partial least squares (PLS) analysis [J]. Anal Lett, 2020, 54(5): 830-841.
- [9] 顾丰颖, 丁雅楠, 朱金锦, 等. 基于稻谷 X 射线荧光光谱测定快速识别糙米和精米中的镉含量[J]. 食品安全质量检测学报, 2021, 12(20): 8018-8025.
GU FY, DING YN, ZHU JJ, *et al.* Rapid identification of cadmium content in brown and polished rice based on X-ray fluorescence spectroscopy [J]. J Food Saf Qual, 2021, 12(20): 8018-8025.
- [10] HAO SY, LI JY, LIU XY, *et al.* Authentication of acacia honey using

- fluorescence spectroscopy [J]. *Food Control*, 2021, 130: 108327.
- [11] NTAKATSANE N, CHEN P, LIU JS, *et al.* Multi-dimensional fluorescence spectroscopy coupled with chemometrics in rapid antibiotic detection and discrimination [J]. *J Food Meas Charact*, 2020, 14: 1892–1900.
- [12] JI RD, MA SC, BIAN HY, *et al.* Determination and modeling on ultraviolet light degradation of pyridaben based on fluorescence spectrum [J]. *Comb Chem High Throughput Screen*, 2020, 23(2): 141–147.
- [13] 张亚莉, 颜康婷, 王琳琳, 等. 基于荧光光谱分析的农药残留检测研究进展[J]. *光谱学与光谱分析*, 2021, 41(8): 2364–2371.
ZHANG YL, YAN KT, WANG LL, *et al.* Research progress of pesticide residue detection based on fluorescence spectrum analysis [J]. *Spectrosc Spectr Anal*, 2021, 41(8): 2364–2371.
- [14] ROBERTO KHG, MARIO CUA, JOS GE, *et al.* A method for calibration and validation subset partitioning [J]. *Talanta*, 2005, 67: 736–740.
- [15] 黄燕, 王璐, 关海鸣, 等. 基于优选 NIR 光谱波数的绿豆产地无损检测方法[J]. *光谱学与光谱分析*, 2021, 41(4): 1188–1193.
HUANG Y, WANG L, GUAN HOU, *et al.* Nondestructive detection method of mung bean origin based on optimized nir spectral wavenumber [J]. *Spectrosc Spectr Anal*, 2021, 41(4): 1188–1193.
- [16] ZHAO A, TANG X, ZHANG Z, *et al.* Optimizing Savitzky-Golay parameters and its smoothing pretreatment for FTIR gas spectra [J]. *J Instrum Anal*, 2016, 36(5): 1340–1344.
- [17] WANG QQ, LIU YH, XU Q, *et al.* Identification of mildew degrees in honeysuckle using hyperspectral imaging combined with variable selection [J]. *J Food Meas Charact*, 2019, 13(3): 2157–2166.
- [18] 张静, 李钰, 任舜文. 基于有限脉冲击应滤波器的实时小波算法及其在色谱信号解析中的应用[J]. *色谱*, 2017, 35(4): 368–374.
ZHANG J, LI Y, REN SW. Real-time wavelet algorithm based on finite impulse response filter and its application in chromatographic signal analysis [J]. *Chin J Chromatogr*, 2017, 35(4): 368–374.
- [19] NAZARLOO AS, SHARABIANI VR, GILANDEH YA, *et al.* Evaluation of different models for non-destructive detection of tomato pesticide residues based on near-infrared spectroscopy [J]. *Sensors*, 2021, 21(9): 3032.
- [20] 李速专, 童何馨, 袁雷明, 等. 间隔连续投影算法应用于近红外光谱苹果糖度模型的优化[J]. *食品安全质量检测学报*, 2019, 10(14): 4608–4612.
LI SZ, TONG HX, YUAN LM, *et al.* Optimization of apple sugar content model using interval continuous projection algorithm in near-infrared spectroscopy [J]. *J Food Saf Qual*, 2019, 10(14): 4608–4612.
- [21] 许召召, 申德荣, 聂铁铮, 等. 融合信息增益比和遗传算法的混合式特征选择算法[J]. *软件学报*, 2022, 33(3): 1128–1140.
XU ZZ, SHEN DR, NIE TZ, *et al.* Hybrid feature selection algorithm incorporating information gain ratio and genetic algorithm [J]. *J Software*, 2022, 33(3): 1128–1140.
- [22] 李艳肖, 黄晓玮, 邹小波, 等. 蚁群-遗传算法优化近红外光谱检测花茶花青素含量的研究[J]. *食品安全质量检测学报*, 2014, 5(6): 1679–1686.
LI YX, HUANG XW, ZOU XB, *et al.* Study on ant colony genetic algorithm optimized near infrared spectroscopy for the detection of anthocyanin content in *Camellia* [J]. *J Food Saf Qual*, 2014, 5(6): 1679–1686.
- [23] JAIR C, FAIRD GL, LISBETH RM, *et al.* A comprehensive survey on support vector machine classification: Applications, challenges and trends [J]. *Neurocomputing*, 2020, 408: 189–215.
- [24] NECATI C, CEVDET S. Rapid detection of total phenolics, antioxidant activity and ascorbic acid of dried apples by chemometric algorithms [J]. *Food Biosci*, 2022, 47: 101670.
- [25] 咎佳睿, 刘翠玲, 凌彩金, 等. 基于高光谱技术的红茶茶多酚可视化研究[J]. *食品安全质量检测学报*, 2023, 14(5): 37–44.
ZAN JR, LIU CL, LING CJ, *et al.* Visualization of tea polyphenols in black tea based on hyperspectral technology [J]. *J Food Saf Qual*, 2023, 14(5): 37–44.
- [26] WANG SM, SUN J, FU LH, *et al.* Identification of red jujube varieties based on hyperspectral imaging technology combined with CARS-IRIV and SSA-SVM [J]. *J Food Process Eng*, 2022, 45(10): 14137.
- [27] 曹敬椿, 卢敏. 基于麻雀搜索算法优化支持向量机的区域水资源安全评价[J]. *水电能源科学*, 2023, 41(5): 52–54, 129.
CAO JC, LU M. Regional water resource security evaluation based on sparrow search algorithm optimized support vector machine [J]. *Energy Eng Sci*, 2023, 41(5): 52–54, 129.
- [28] 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 XI. *Molecular spectroscopy analytical technology combined with chemometrics and its applications* [M]. Beijing: Chemical Industry Press, 2011.
- [29] 时鸣, 竞霞, 史晓亮. 基于高光谱和 GA-BP 神经网络模型的冬小麦叶绿素含量反演[J]. *江苏农业科学*, 2022, 50(20): 56–62.
SHI M, JING X, SHI XL. Inversion of chlorophyll content in winter wheat based on hyperspectral and GA-BP neural network models [J]. *Jiangsu Agric Sci*, 2022, 50(20): 56–62.
- [30] TANG RN, CHEN XP, LI C. Detection of nitrogen content in rubber leaves using near-infrared (NIR) spectroscopy with correlation-based successive projections algorithm (SPA) [J]. *Appl Spectrosc*, 2018, 72(5): 740–749.
- [31] CHEN Y, LIU ZY, XIAO CY, *et al.* Overlapping peak analysis of soil heavy metal X-ray fluorescence spectra based on sparrow search algorithm [J]. *Spectrosc Spectr Anal*, 2021, 41(7): 2175–2180.

(责任编辑: 韩晓红 于梦娇)

作者简介



陈 珺, 主要研究方向为食品检测技术与数据处理。

E-mail: chenjue@st.btbu.edu.cn



刘翠玲, 教授, 主要研究方向为智能测量技术与数据处理研究。

E-mail: liucl@btbu.edu.cn