

# 基于 MATLAB GUI 的近红外光谱定量分析系统的开发

## Development of near infrared spectral quantitative analysis system based on MATLAB GUI

帕提姑丽·赛买提, 李晓斌, 张馨之, 卞希慧\*

(天津工业大学化学工程与技术学院, 中国天津 300387)

**摘要:** 近红外光谱因其具有非破坏性、快速高效、适用性广泛等优势在复杂样品分析中得到了广泛应用。实验采集到的样品近红外光谱信号除包含与样本相关的有用信息外, 也掺杂着噪声、背景、杂散光引起的光谱差异等干扰信息。这将对定量模型的质量和未知样本预测的准确度有严重影响。虽然目前已经提出了很多化学计量学方法, 但这些方法的 MATLAB 程序代码对非专业人士晦涩难懂。为化学计量学方法的推广, 基于 MATLAB GUI 开发出一款可以实现数据导入、预处理、变量选择、保存、删除、建模于一体的定量分析系统。为光谱定量分析研究人员提供方便快捷的工作程序, 提高工作效率。

**关键词:** 近红外光谱; 数据预处理; 变量选择; 定量分析

## 1 引言

近红外光谱 (Near - Infrared Spectroscopy, NIR) 分析是一种重要的化学分析技术, 它与化学计量学紧密结合, 可以用于快速、非破坏性地获取物质的光谱信息, 并实现对其成分和属性的分析与预测。目前提出了很多的化学计量学方法<sup>[1-3]</sup>, 包括预处理、变量选择、多元校正等。这些方法多数采用 MATLAB, Python 等专业编程语言编写, 对非专业人士就很难使用, 所以, 研究出一款简单易于操作的定量分析系统对化学计量学方法的推广具有重大意义。

基于 MATLAB GUI 的近红外光谱分析方法具有一系列的优点。通过 GUI 界面, 用户可以直观地进行参数设置、数据可视化和结果分析, 无需编写复杂的代码, 降低了使用门槛。MATLAB 的广泛应用和强大的社区支持, 为研究人员提供了丰富的资源和交流平台, 促进了研究成果的共享和合作。近年来, 许多研究论文都聚焦于基于 MATLAB GUI 的近红外光谱分析方法及应用。比如, 刘等<sup>[4]</sup>设计了原油近红外光谱定量分析软件; 尹等<sup>[5]</sup>采用 C++ 与 MATLAB, 研制开发出一款食用植物油品质检测鉴别系统软件; 方等<sup>[6]</sup>利用 LabVIEW 和 MATLAB 软件, 开发出基于盲信号处理技术的近红外光谱分析仪。MATLAB GUI 设计

和实现的近红外光谱分析软件通过 GUI 界面下载数据、处理数据并可视化分析结果，为使用人员提供了方便且灵活的工具<sup>[7,8]</sup>。

本文的分析系统在单一预处理的基础上，增加了不进行预处理、两种预处理组合、三种预处理组合，预处理组合在同一个预处理模块中，通过点击菜单即可实现对光谱数据组合处理从而获得更高质量的光谱数据，减小误差，提高建模效果。系统还将新型的群体智能优化方法，如灰狼优化算法、鲸鱼优化算法等引入到近红外光谱变量选择。建立了包含数据导入、预处理、变量选择、建模的可视化界面。使得原本通过 Matlab 进行的复杂数据处理-建模过程变得更加简便、集中，极大提高工作人员工作效率。系统还将 MATLAB 环境下开发运行的 .m 格式文件的分析系统通过封装成为 .exe 格式文件，从而使分析系统从 MATLAB 工作环境中脱离出来<sup>[9-12]</sup>，提升了软件的流通性和商业性。

## 2 GUI 设计所使用的化学计量学方法

### 2.1 数据预处理方法

数据预处理方法有很多，按照预处理效果，可以将预处理方法分为基线扣除、散射校正、平滑处理和尺度缩放四大类<sup>[13]</sup>。其中每类预处理方法又包括多种具体算法。需要进行数据预处理通常包括以下情况：

消除各样本在自变量强度及范围上差异，使得数据具有标准的变化幅度或范围，同时保留自变量的相对比值信息。从化学分类学角度而言，复杂体系样本个体所携带的化合物组成种类和构成比例信息，通常要比单纯的化合物含量信息更稳定、也对分类更有价值<sup>[14]</sup>。

#### 2.1.1 一阶导数 (1st Derivative)

一阶导数法是一种常用的预处理方法，通过计算光谱中每个数据点的斜率来获得一阶导数光谱，用于去除基线漂移以及突出光谱特征。

#### 2.1.2 二阶导数 (2nd Derivative)

二阶导数法通过计算光谱的二阶导数来增强光谱特征并减少背景干扰。二阶导数法的基本原理是通过计算光谱中相邻数据点的差异来获取光谱的斜率信息。一阶导数可以提供光谱的斜率变化趋势，而二阶导数可以提供光谱的曲率信息。

#### 2.1.3 Savitzky-Golay 平滑

Savitzky-Golay 平滑法用于降低光谱中的噪声<sup>[15]</sup>和提取信号特征。它通过对光谱数据进

行多项式拟合和平滑处理，实现对光谱曲线的光滑化，从而改善数据的质量和可分析性。

Savitzky-Golay 平滑法的核心思想是在局部窗口内，通过多项式拟合来逼近原始光谱数据，并用该拟合曲线替代原始数据点，从而平滑光谱曲线<sup>[15]</sup>。

#### 2.1.4 标准正态变量 (Standard Normal Variate, SNV)

标准正态变量校正<sup>[16]</sup>方法可以减少光谱中的散射效应和光程差引起的波动，提高数据的可比性和分析精度。

SNV 法可以消除光谱中的整体偏移和尺度效应，使得光谱数据更加稳定和可比。

$$X_{i,SNV} = \frac{x_{i,k} - x_i}{\sqrt{\frac{\sum_{k=1}^m (x_{i,k} - x_i)^2}{m-1}}} \quad (1)$$

#### 2.1.5 乘性散射校正 (Multiplicative scattering correction, MSC)

乘性散射校正<sup>[16]</sup>主要用于消除光谱中的多次散射效应，提高数据的准确性和可比性。

MSC 法的基本原理是建立一个线性回归模型，将参考样品的光谱与其相应的散射光谱进行拟合。

$$X_i = a_i + b_i X_i + e_i \quad (2)$$

式中， $X_i$ 为第*i*个样品的光谱， $X_i$ 为该组样品的平均光谱， $e_i$ 为残差光谱，理想地认为它是与化学性质相关的信息， $a_i$ 和 $b_i$ 分别为线性回归的截距与斜率，可由样品光谱集线性回归获得，并用于校正每条光谱：

$$X_{i,MSC} = \frac{(X_i - a_i)}{b_i} \quad (3)$$

#### 2.1.6 组合预处理

不同的预处理方法对近红外光谱进行预处理，通过模型的组合来实现对不同预处理结果的集合预测。模型的组合可能被用于未知样品的集合预测。从统计上看，集成模型在稳定性和预测误差方面优于单一模型。光谱分析的发展过程中，样品的复杂程度一直在不断地改变，先是对样品的单一组分进行测量和分析，随着复杂样品体系的出现，对多组分进行测量，并在光谱进行定量和定性分析。

## 2.2 变量选择方法

### 2.2.1 灰狼优化算法 (Grey Wolf Optimization, GWO)

灰狼优化算法<sup>[17,18]</sup> (GWO) 是一种基于自然灰狼行为的优化算法, 最初由 Mirjalili 等人<sup>[19]</sup>于 2014 年提出。该算法受到灰狼群体在狩猎过程中的行为策略启发, 通过模拟灰狼群体的寻食行为来解决优化问题。灰狼优化算法具有快速收敛、全局搜索能力强等特点, 已广泛应用于函数优化、参数选择和模型训练等领域。

Chanda 等人<sup>[20]</sup>2017 年将灰狼优化算法应用于对茶叶近红外光谱中多酚含量的预测。以 PLS 的均方根误差(RMSE)作为适应度函数建立多酚含量预测模型。得到 RMSECV 为 0.1364, 预测集的 RMSEP 和相关系数分别为 0.3244 和 0.91。Zhang 等<sup>[21]</sup>2022 年将离散的灰狼优化算法作为变量选择方法引入到四元食用调和油的紫外光谱分析中。该方法在更新狼群的位置后利用逻辑函数对其进行离散化。在最佳参数的基础上建立 GWO-PLS 模型并预测样品中纯油的含量, 如图 1 所示。

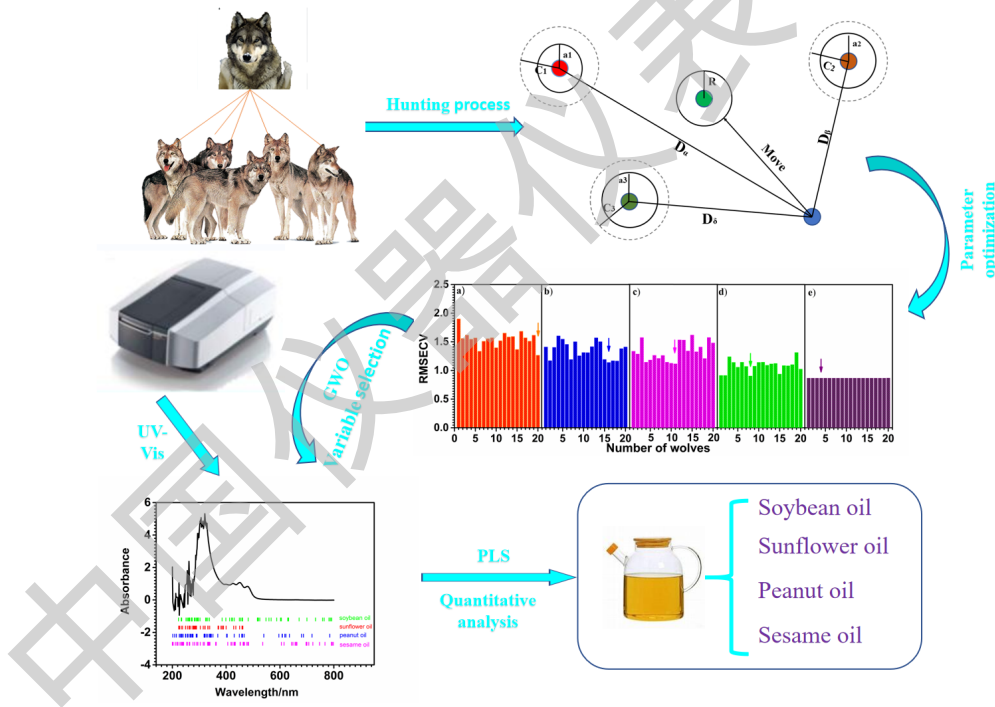


图 1 WOA 用于四元调和油中的不同组分预测的过程

### 2.2.2 鲸鱼优化算法 (Whale Optimization Algorithm, WOA)

鲸鱼优化算法<sup>[22,23]</sup> (Whale Optimization Algorithm) 是一种基于仿生学和群体智能的优化算法, 由 Seyedali Mirjalili 等人<sup>[24]</sup>于 2016 年提出。鲸鱼优化算法模拟了鲸鱼群体的游动行为, 通过个体间的合作和竞争, 寻找最优解。该算法在全局搜索和局部优化方面具有较强的能力<sup>[25]</sup>。此外, 参数较少、原理简单的优势被广泛应用于适用于各种复杂的优化问题,

尤其是在变量选择问题中有着广泛的应用。

Tang 等人<sup>[26]</sup>2021 年提出基于高光谱成像技术和竞争自适应重加权采样-鲸鱼优化算法-支持向量机的枸杞品种鉴定。首先采用 SNV 对原始光谱数据进行预处理，选择竞争性自适应权重取样（Competitive Adaptive Reweighted Sampling, CARS）提取特征波长，最后建立了 SVM 模型。为了提高分类精度，引入了 WOA。WOA-SVM 模型得到的训练集和预测集的准确率分别为 89.44%和 88.33%。

## 2.3 多元校正建模方法

### 2.3.1 偏最小二乘回归（Partial Least Squares, PLS）

偏最小二乘<sup>[27]</sup>是一种多元回归分析方法，它可以应用于建立预测模型、探索变量之间的关系以及解决高维数据问题<sup>[28]</sup>。PLS 建模在统计学和化学领域得到广泛应用，并逐渐在其他领域中被采用。

### 2.3.2 支持向量机（Support Vector Machine, SVM）

支持向量机<sup>[29-31]</sup>是一种强大的机器学习算法用于解决二分类问题的方法。广泛应用于模式识别、数据分类和回归分析等领域。SVM 建模通过在特征空间中找到最优的超平面来实现数据的划分和分类，其独特的优化目标使其在处理线性和非线性问题上具有出色的性能。

SVM 通过选择一组支持向量，即距离超平面最近的样本点，来描述决策边界，并利用这些支持向量进行分类和预测，如图 2 所示。

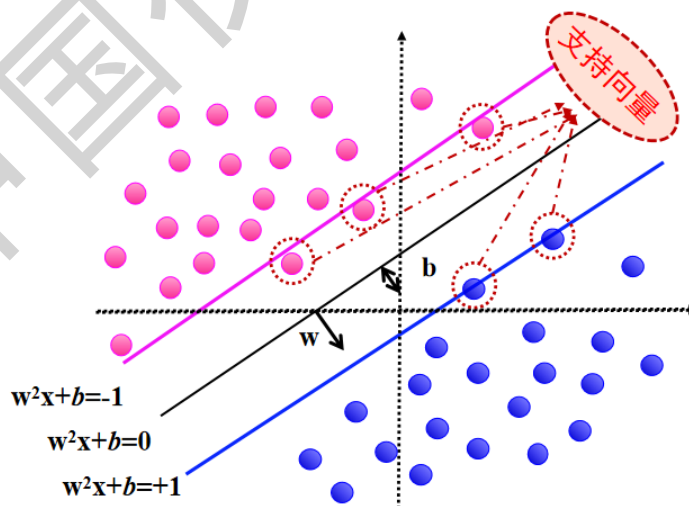


图 2 SVM 原理图

### 2.3.3 极限学习机 (Extreme Learning Machine, ELM)

极限学习机<sup>[32,33]</sup>是一种快速且有效的人工神经网络算法,广泛应用于模式识别、数据分类和回归分析等领域。ELM 建模以其快速的训练速度和良好的泛化能力而备受关注,并在大数据环境下展现了出色的性能。

极限学习机 (ELM) 是黄光斌在 2004 年提出的一种人工神经网络模型,它是一种解决单隐层前馈神经网络的学习算法。传统的前向神经网络,如 BP 神经网络<sup>[34]</sup>,需要手动设置许多网络参数,但这种方法只需要建立网络,不需要设置其他参数,所以使用方便。

## 2.4 模型性能评价参数

校正模型的预测性能通过相关系数 (correlation coefficient,  $R$ ), 交叉验证均方根误差 (Root mean Square Error of Cross-Validation, RMSECV) 和预测均方根误差 (Root Mean Square Error of Prediction, RMSEP 来评价。 $R$  越接近 1, 表示模型的预测精度越高, 性能越好。RMSECV 和 RMSEP 分别表示校正集和预测集的预测值和真实值的误差, 这两个值越小, 代表模型性能越好。

## 3 GUI 系统的设计与开发

### 3.1 系统编程语言的选择

MATLAB 是由 MathWork 公司发布的主要用于科学计算、可视化以及人机界面设计的一种开发的软件<sup>[35]</sup>。MATLAB 在数据处理和建模方面具有丰富的工具箱、高效的矩阵运算、友好的图形化界面和交互式环境、广泛的社区支持以及多领域应用等优点,使得在 MATLAB 中实现对近红外光谱的功能更加方便、高效和可靠。

### 3.2 MATLAB GUI 图形用户界面设计

MATLAB (MATrix LABoratory) 是一种多范式数值计算环境和第四代编程语言。MATLAB 允许进行矩阵操作、函数和数据绘图、算法实现、创建用户界面以及与用其他语言 (包括 C、C++、Java、Fortran 和 Python) 编写的程序连接。随着计算机技术的快速发展,开发者赋予其越来越多的附加工具箱, GUIDE (Graphical User Interface Developmeng Environment) 是众多附加工具箱的一种,其主要目的是设计和建立图形用户界面 (Graphical User Interface, GUI)。在 GUI 用户界面中,人们只需要通过简单操作鼠标、键盘等外部输入设备便能够执行内部相应的程序。GUI 编辑界面由状态栏、工具栏、控件面板以及图形设计区域等构成<sup>[36]</sup>。

图形用户界面 (GUI) 为采用图形方式显示的计算机操作用户界面, MATLAB 为设计

者设计图形界面用户提供了方便、快捷的制作开发环境。同时 MATLAB 有着大量库函数、极高的编程效率、数据可视化等优点，许多设计者都采用此模块进行设计开发软件<sup>[37]</sup>。总之，MATLAB GUI 界面设计的原则包括易用性、一致性、响应性、易于导航和组织、合理的布局和对齐、可访问性、异常处理和错误提示、可扩展性和灵活性，以及测试和优化。遵循这些原则，结合 MATLAB 提供的 GUI 设计工具，可以创建出功能强大、易于使用的 GUI 界面，提供良好的用户体验。

### 3.3 主界面设计

本系统在 Windows10 及 MATLAB R2018b 版本开发运行。主界面如图 3 所示。

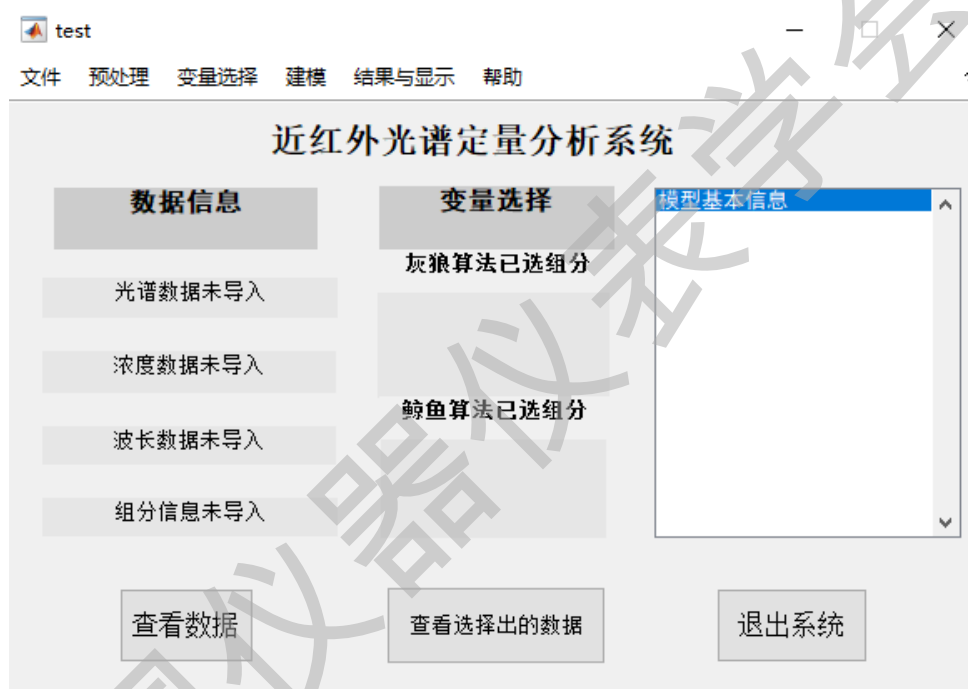


图 3 近红外光谱分析系统主界面

主界面上方菜单栏中有六个菜单按钮，菜单需在菜单编辑器中编辑<sup>[38]</sup>。菜单编辑器界面如图 4 所示。

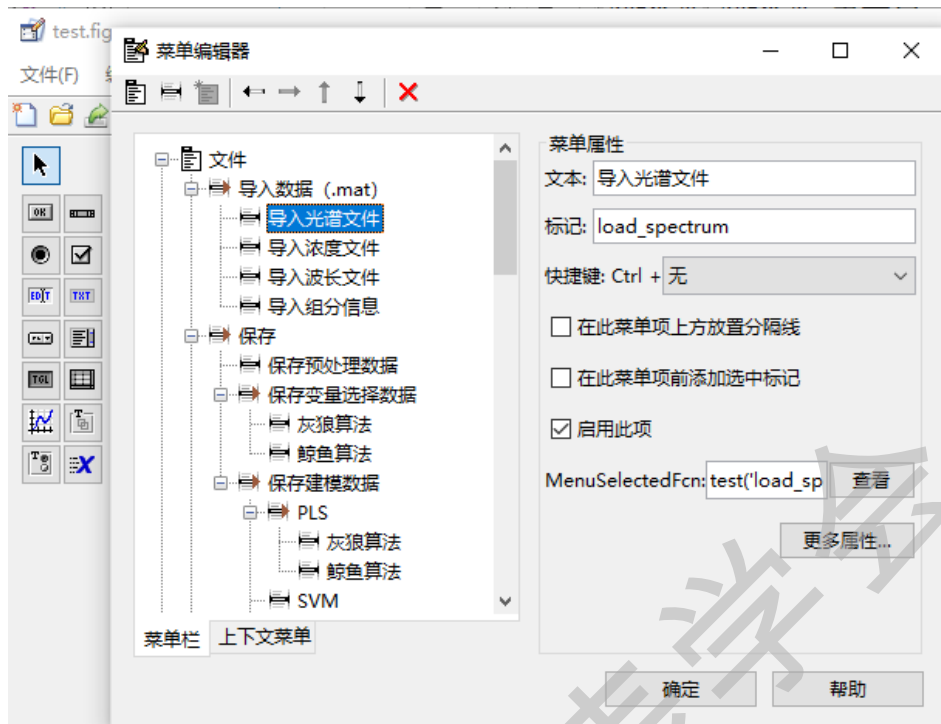


图 4 GUI 菜单编辑器

在 GUI 编辑器中编辑完成控件并设置好对应的 String、Tag 及其他属性后，点击保存按钮，MATLAB 会生成一个 .fig 格式的文件以及一个相同名称的 .m 格式文件。

### 3.4 光谱数据操作模块

光谱数据操作模块 String 属性为“文件”，位于菜单的最左边，该设计符合用户对办公软件的使用习惯。文件菜单有三个二级菜单，主要功能是对需要处理的数据进行导入、保存、删除操作。

在进行数据导入说明之前，需对数据的格式及构成进行说明。本系统支持对 .mat 格式文件的处理，使用的完整数据应由四部分构成：近红外光谱数据、组分浓度信息、波长数据、组分名称信息。组分浓度信息的数据行为各组分的浓度值，数据列为不同的组分。波长数据的行数与光谱数据的列数保持一致。组分名称信息的每列对应于组分浓度信息的每列数据。

实验的数据为天津超市购买的保质期以内的合格大豆油、芝麻油、葵花油、玉米油、花生油和稻米油。将这些油按质量等间隔、不重复制成 51 份六元食用调和油样品。数据采集使用 Vertex 70 近红外光谱仪(德国埃特林根布鲁克光谱仪器公司)测量，在  $12000-4000\text{cm}^{-1}$  的波长范围内测得 4148 个变量。

### 3.5 预处理模块

预处理模块位于菜单栏的第二个菜单，其 Tag 属性为“reprocessing”，包含有六个二级菜单，组合预处理下包含 22 个菜单。回调函数中主要使用 isfield()、msgbox()、inputdlg()、



sgolayfilt()、figure、plot()等函数。用户点击菜单后，先运行 isfield()对数据进行判断，是否已经导入了光谱数据，若未导入，则运行 msgbox()弹出对话框提醒用户先导入数据；若已导入，则运行 inputdlg()弹出输入对话框，以 Savitzky-Golay 平滑为例，需要输入窗口大小以及多项式级数，窗口大小必须为奇数，且在输入框已提示用户，当用户输入偶数则报错。如图 5 所示。

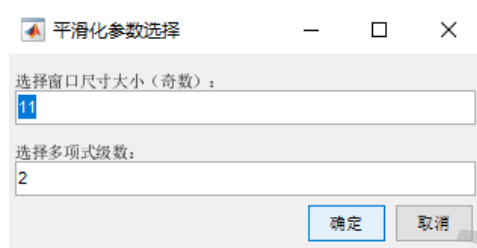


图 5 Savitzky-Golay 平滑化后生成的处理前后对比图形对话框

### 3.6 变量选择模块

变量选择模块位于菜单栏第三菜单，包含灰狼优化算法(GWO)、鲸鱼优化算法(WOA)两个二级菜单。以用户点击鲸鱼优化算法菜单为例，首先运行 isfield()函数，检查是否光谱数据已经预处理，若未进行预处理，则弹出对话框提示用户先对数据进行预处理，用户也可以选择不进行预处理，直接对原始光谱数据进行鲸鱼优化算法变量选择操作。

用户选择结束后，执行WOA()及gray\_wolf\_optimizer\_Callback()与鲸鱼优化算法相关的变量计算。鲸鱼优化算法与灰狼优化算法计算时间会随迭代次数增加而变长，故在执行鲸鱼优化算法之前，会先执行dialog()弹出对话框提示用户等待鲸鱼优化算法计算。

鲸鱼优化算法进行变量选择结束后，等待对话框关闭，弹出鲸鱼优化算法变量选择的点在原始光谱数据图形中的位置信息对话框，如图 6 所示。

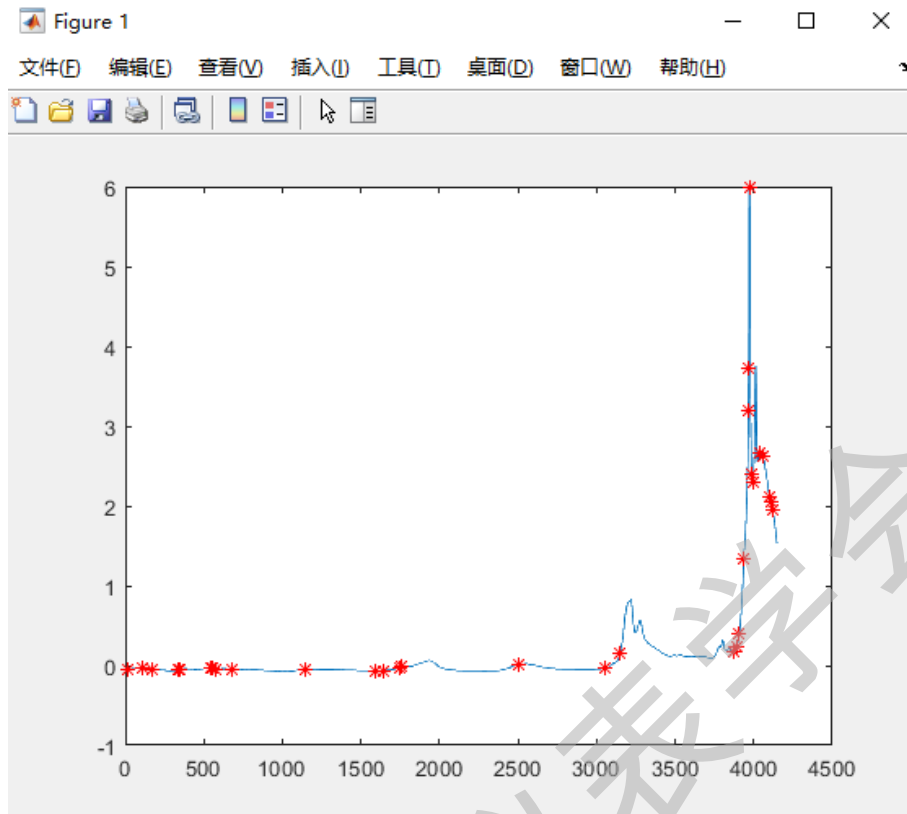


图 6 鲸鱼优化算法变量选择后弹出选择结果图

最后执行两次 set() 函数，用于将变量选择结果信息以及选择的组分信息设置更新到 text 静态文本控件显示，如图 7 所示。

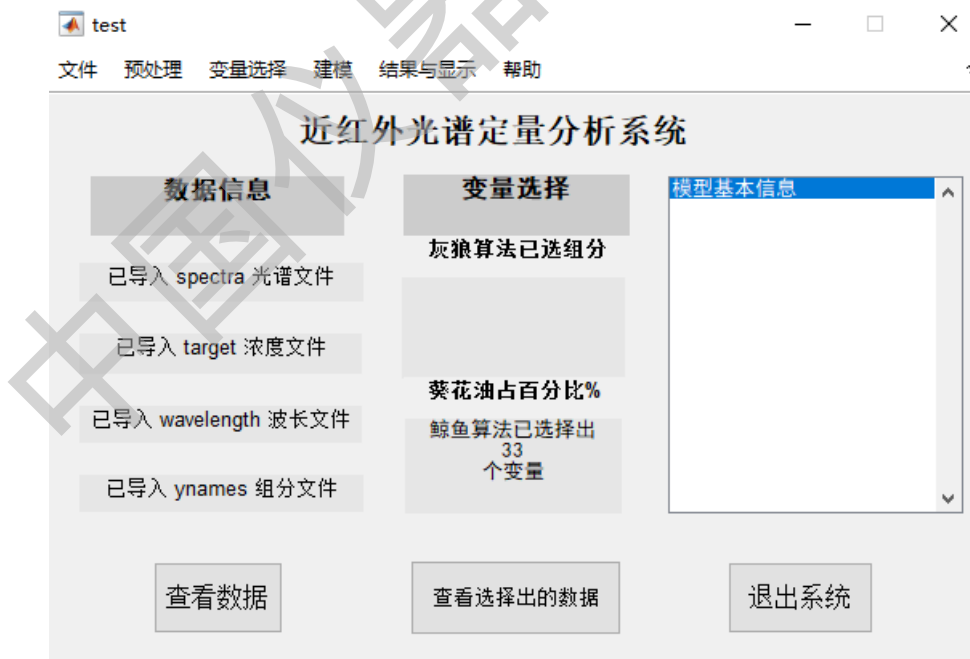


图 7 鲸鱼优化算法变量选择更新到主界面

此时点击查看选择出的数据按钮，会打开对话框供用户选择需要打开已选择出的数据，若未执行变量选择，则执行 msgbox() 弹出对话框提示用户先进行变量选择，若只进行了一

种变量选择, 则直接执行 `assignin()` 函数及 `openvar` 函数打开变量, 若灰狼优化算法和鲸鱼优化算法都已选择出变量, 则弹出 `listdlg` 对话框由用户选择需要查看的已选择出的数据变量。

### 3.7 建模模块

建模模块位于菜单栏的第四菜单, 其包含三个二级菜单。以偏最小二乘 PLS 建模, 选择灰狼优化算法结果变量选择为例, 此建模方式将同时对未进行变量选择及进行变量选择的数据建模, 并输出对应结果。用户点击菜单后, 先执行 `isfield()` 函数, 判断数据是否已进行变量选择, 若未进行, 则执行 `msgbox()` 弹出对话框提醒用户进行变量选择。

若已进行变量选择, 则执行建模相关函数, 建模结束后, 执行 `figure`、`plot()` 等绘图函数, 绘制出四张图, 如图 8 所示。

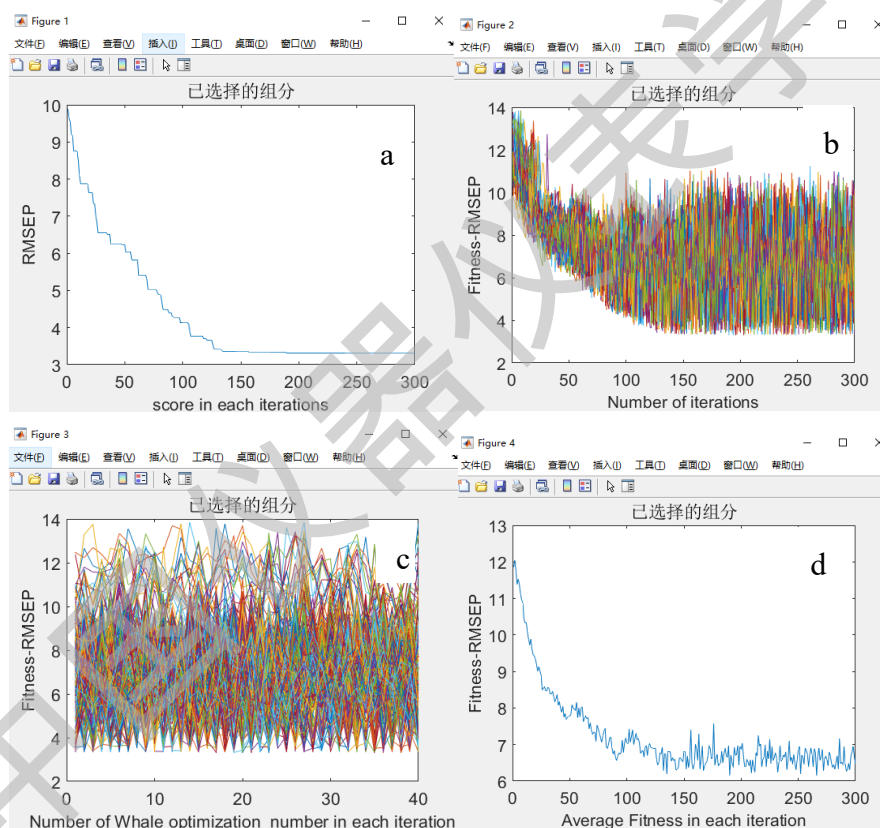


图 8 组分的 RMSEP 随迭代次数的变化(a) 300 次迭代下每只鲸鱼的适应度值 (b) 每支鲸鱼在 300 次迭代下的适应度值 (c) 300 次迭代下鲸鱼种群的平均适应度值 (d)

同时, `set()` 函数将建模产生的 `RMSECV`、`RMSEP`、训练集相关系数、预测集相关系数结果更新到主界面右侧列表框中, 如图 9 所示。

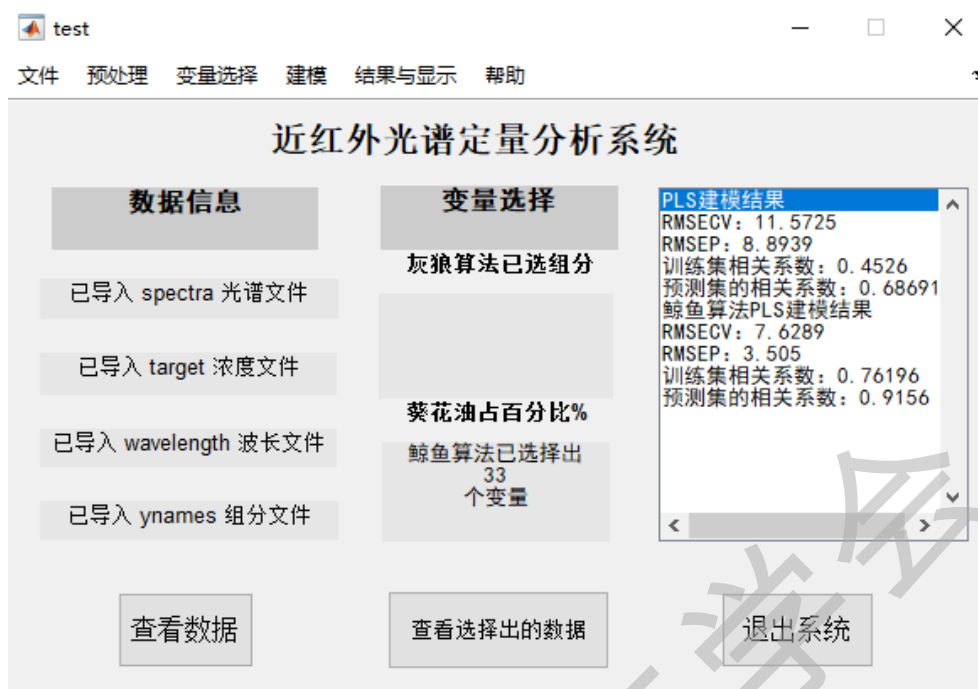


图 9 执行 set()函数将建模结果更新设置到列表框中

### 3.8 结果与显示模块

结果与显示模块位于菜单栏的第五菜单，其中包含三个二级菜单，每个二级菜单中包含灰狼优化算法和鲸鱼优化算法两个菜单，由于建模是基于不同变量选择结果进行的，所以需要建模结果做出区分。用户点击对应的建模结果菜单后，会打开 txt 文本，建模产生的图形结果，RMSECV、RMSEP、训练集相关系数、预测集相关系数等信息等都包含在打开的.txt 文件中。用户可以在.txt 文本文件中进行修改，另存为等操作。

### 3.9 帮助模块

帮助菜单为菜单栏中最后一个菜单，Tag 属性为值“help”，点击帮助菜单，弹出对话框，对话框对近红外建模系统处理所需数据的要求，操作注意事项，常见错误等进行说明解释，以帮助用户更好的使用此系统。

### 3.10 系统封装

系统封装是本课题的创新点之一，对近红外光谱建模系统进行封装，可以使其脱离 MATLAB 环境运行，并提高建模系统软件在网络及不同计算机中的流动性，增加商业属性，方便更多用户。

将封装的.exe 软件与 MATLAB Runtime 打包为一个文件夹，即可进行传输，建模系统可以在不具备 MATLAB 环境下 Windows 操作系统运行。

## 4 结论

建模分析系统最大亮点在于将多种数据预处理、变量选择、建模集成于一个软件系统,使得对数据能进行组合处理,最大限度的使建模结果更加准确,模型质量更高,预测结果更准确,并且能够对处理的数据进行保存等操作,极大地方便用户对于近红外光谱数据的定量分析,且能够使工作效率得到极大提升。

#### 参考文献:

- [1] Parastar H, Tauler R, Big (bio) chemical data mining using chemometric methods: a need for chemists, *Angew. Chem. Int. Ed.* 2022, 134(44): 201801134.
- [2] Kim Y S, Ha H J. Rapid determination of the geographical origin of kimchi by Fourier transform near-infrared spectroscopy coupled with chemometric techniques[J]. *Sci. Rep.* 2024,14(1):24581-24581.
- [3] 唐振豪, 张智勇, 沙鑫, 等. 近红外水光谱组学技术及其应用近十年研究进展[J]. *分析测试学报*, 2024, 43(07): 1086-1096.
- [4] 刘源华. 原油近红外光谱定量分析软件的设计与实现[D]. 重庆大学, 2021.
- [5] 尹成. 基于 C++与 MATLAB 食用油近红外光谱分析软件的开发与应用[D]. 武汉轻工大学, 2016.
- [6]方利民, 林敏. ICA 的近红外光谱分析软件的研制[J].*中国计量学院学报*, 2010, 21(01): 42-45+81.
- [7] 邓昌瑞. 基于 MATLAB GUI 的多功能计算系统设计及实现[D]. 南昌大学, 2013.
- [8] Zhao Y, Sun P P, Tan F L, et al. NIRS-ICA: A MATLAB toolbox for independent component analysis applied in fnirs studies[J]. *Front. Neuroinformatics*, 2021, 15: 683735.
- [9] Choi S, Kim W, Yoon S. Development of an encapsulator for interoperability of non-HLA based MATLAB programs (SIMULINK or M-file) to HLA base simulations[J]. *Int. J. Mach. Learn. Comput.* 2020, 10(1):176-181.
- [10] Eric W, Mark W, R A C. Hydra image processor: 5-D GPU image analysis library with MATLAB and python wrappers[J]. *Bioinformatics(Oxford England)*, 2019, 35(24): 5393-5395.
- [11] Jiménez J, Mosquera E M, Maestre J M. A simple framework for working with MATLAB and Home I/O[J]. *IFAC Pap Online*, 2023, 56(2):9588-9593.
- [12]魏建, 孙祥娥, 任睿杰, 等. 一种远程部署 MATLAB 应用的新方法[J]. *计算机仿真*, 2024, 41(07): 423-428.

- [13] 尼珍, 胡昌勤, 冯芳. 近红外光谱分析中光谱预处理方法的作用及其发展[J]. 药物分析杂志, 2008, 28(5): 824-829.
- [14] 罗华飞. MATLAB GUI 设计学习手记 (第二版) [M]. 北京: 北京航空航天大学出版社, 2011, 2, 80-390.
- [15] 雷林平. 基于 Savitzky-Golay 算法的曲线平滑去噪[J]. 电脑与信息技术, 2014, 22(05): 30-31.
- [16] 尹成诚, 康景, 刘建新, 等. 基于 SNV 和 MSC 结合遗传算法对羊肉葡萄糖含量可见—近红外光谱建模效果研究[J/OL]. 草业科学, 2024, 1, 1-11
- [17] Zhang R L, Wu X Y, Chen Y J, Xiang Y, Liu D, Bian X H, Grey wolf optimizer for variable selection in quantification of quaternary edible blend oil by ultraviolet-visible spectroscopy, *Molecules*, 2022, 27 (16), 5141.
- [18] 冯严冰, 王洪亮. 基于入侵杂草优化和灰狼优化算法的混合优化算法研究[J]. 信息技术与信息化, 2024, (09): 67-70.
- [19] Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer[J]. *Adv. Eng. Softw.* 2014, 69: 46-61.
- [20] Chanda S, Sing D, Majumder S, et al. NIR spectroscopy with grey wolf optimization algorithm for prediction of polyphenol content in inward tea leaves[C]//2017 IEEE Calcutta Conference (CALCON). IEEE, 2017: 392-396.
- [21] Zhang R L, Wu X Y, Chen Y J, et al. Grey Wolf Optimizer for Variable Selection in Quantification of Quaternary Edible Blend Oil by Ultraviolet-Visible Spectroscopy[J]. *Molecules*, 2022, 27(16): 5141.
- [22] 李梦悦, 那振宇, 林云. 基于改进鲸鱼优化算法的异构无人机群频谱分配[J]. 海军航空大学学报, 2024, <https://link.cnki.net/urlid/37.1537.V.20241025.1122.006>.
- [23] Liu Y X, Sun H, Zhao C Y, Ai C K, Bian X H, Extreme learning machine combined with whale optimization algorithm for spectral quantitative analysis of complex samples, *J. Chemometrics*, 2024, 38 (10): 3590.
- [24] Mirjalili S, Lewis A. The whale optimization algorithm[J]. *Adv. Eng. Softw.* 2016, 95: 51-67.
- [25] 温少晨. 鲸鱼优化和堆优化的混合算法研究[D]. 河南师范大学, 2022.
- [26] Tang N Q, Sun J, Yao K S, et al. Identification of lycium barbarum varieties based on hyperspectral imaging technique and competitive adaptive reweighted sampling-whale optimization algorithm-support vector machine[J]. *Journal of Food Process Engineering*, 2021, 44(1): e13603.

- [27] Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics[J]. Chemom. Intell. Lab. Syst. 2001, 58(2): 109-130.
- [28] 王惠文, 吴载斌, 孟洁. 偏最小二乘回归的线性与非线性方法[M]. 北京:国防工业出版社, 2006.
- [29] Vapnik V N, Lerner A Y. Recognition of patterns with help of generalized portraits[J]. Autom. Remote Control, 1963, 24(6): 774-780.
- [30] Thissen U, Pepers M, Üstün B, et al. Comparing support vector machines to PLS for spectral regression applications[J]. Chemom. Intell. Lab. Syst. , 2004, 73(2): 169-179.
- [31] Soares F, Anzanello M J. Support vector regression coupled with wavelength selection as a robust analytical method[J]. Chemom. Intell. Lab. Syst. 2018, 172: 167-173.
- [32] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks[C]//2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541). IEEE, 2004, 2: 985-990.
- [33] 徐睿, 梁循, 齐金山, 等. 极限学习机前沿进展与趋势[J]. 计算机学报, 2019, 42(07): 1640-1670.
- [34] 刘立轴, 卢毅勤, 林梦婧, 等. BP 神经网络在安全评估领域的研究进展文献研究综述[J]. 安全与健康, 2024, (01): 65-68.
- [35] 王璐. 基于 MATLAB 的车牌识别系统研究[D]. 上海交通大学, 2009.
- [36] 王泽刚, 陈照海, 吴泽, 等. 基于 MATLAB 的雷达探测数据处理图形界面设计[J]. 中国设备工程, 2024(02): 98-101.
- [37] 尹嵩杰. 基于 Matlab 平台开发红外光谱建模工具包及其在中药中的应用[D]. 广东药科大学, 2016.
- [38] Consonni V, Baccolo G, Gosetti F, et al. A MATLAB toolbox for multivariate regression coupled with variable selection[J]. Chemom. Intell. Lab. Syst. 2021, 213: 104313.

通讯作者简介

卞希慧, 女, 1983 年生, 天津工业大学化学工程与技术学院教授, 主要进行化学计量学算法研究及其在中药、食品、环境等方面的应用研究。

E-mail: [bianxihui@163.com](mailto:bianxihui@163.com)

第一作者简介

帕提姑丽·赛买提, 女, 2002 年生, 本科生。

E-mail: [13439382896@163.com](mailto:13439382896@163.com)

中国仪器仪表学会