

基于MATLAB的农作物种植策略优化研究

王乔仪, 蔡昌友

辽宁科技大学电子与信息工程学院, 辽宁 鞍山

收稿日期: 2024年12月25日; 录用日期: 2025年1月18日; 发布日期: 2025年1月24日

摘要

本研究聚焦于现代农业生产中农作物种植策略的优化问题, 目的在于通过科学配置作物与土地资源, 结合MATLAB工具构建优化模型, 以实现种植收益的最大化。收集的作物数据, 运用线性回归模型对关键变量进行了相关性分析, 针对农作物市场状况不变的前提, 构建了基于静态销售价格的种植优化模型, 考虑了超出预期销售量部分作物滞销或降价销售的两种策略, 并运用优化算法为每块土地分配最优种植方案。研究发现, 降价销售策略能够有效提升整体收益。同时, 研究还构建了基于不确定性的多阶段优化模型, 以适应市场和气候的不确定性, 动态优化种植方案。本研究提出的模型能够为不同土地在未来数年内提供最优种植方案, 同时考虑了市场变化、作物轮作等多重现实约束。为农业种植策略的优化提供理论支持和实践指导。

关键词

多元线性回归, 粒子群算法, 农作物种植优化模型, 多阶段优化

Optimization of Crop Planting Strategy Based on MATLAB

Qiaoyi Wang, Changyou Cai

School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan Liaoning

Received: Dec. 25th, 2024; accepted: Jan. 18th, 2025; published: Jan. 24th, 2025

Abstract

This study focuses on the optimization of crop planting strategies in modern agricultural production. The purpose is to construct an optimization model by scientifically allocating crop and land resources and combining MATLAB tools to maximize the benefits of planting. Based on the collected crop data, the linear regression model was used to analyze the correlation of key variables. Aiming

at the premise of unchanged crop market conditions, a planting optimization model based on static sales price was constructed. Two strategies of unsalable or reduced-price sales of some crops beyond the expected sales volume were considered, and the optimization algorithm was used to allocate the optimal planting plan for each land. The study found that the price reduction sales strategy can effectively improve the overall revenue. At the same time, the study also constructed a multi-stage optimization model based on uncertainty to adapt to the uncertainty of market and climate and dynamically optimize the planting plan. The model proposed in this study can provide optimal planting schemes for different land in the next few years, taking into account multiple realistic constraints such as market changes and crop rotation. It provides theoretical support and practical guidance for the optimization of agricultural planting strategies.

Keywords

Multiple Linear Regression, Particle Swarm Optimization, Optimization Model for Crop Planting, Multi-Stage Optimization

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

目前, 农作物种植策略优化领域的研究正日益受到重视。随着农业集约化和智能化趋势的显著增强, 如何科学配置作物与土地资源, 以实现种植收益的最大化, 已成为农业科学研究的关键。当前研究主要集中在利用数学模型和算法对种植策略进行优化。例如, 通过线性规划、粒子群优化算法等, 结合作物的生长周期、市场需求、气候条件等多维度信息, 为每块土地分配最优种植方案[1]。这些研究不仅考虑了作物的产量和种植成本, 还深入分析了市场波动和气候变化对农作物经济效益的影响。该领域的研究也存在一些争议。例如, 在模型构建过程中, 如何准确预测作物的产量、销售价格和种植成本等关键参数, 以及如何处理这些参数的不确定性, 是当前研究的难点和争议点[2][3]。此外, 不同地区的土壤、气候和市场需求等条件不同, 如何制定具有普遍性的种植策略也是研究者们需要面临的挑战。通过运用 Matlab 软件进行农作物种植策略的优化分析, 提出了一个基于不确定性和作物关联性的多阶段优化模型。未来, 农作物种植策略优化研究的发展方向将更加注重模型的实用性和精准性[4][5]。

2. 数据来源与分析

本研究的数据主要来源于 2023 年辽宁省数学建模大赛中的信息, 包含了地块信息、作物信息、市场销售与成本数据以及历史种植数据四个方面。地块信息详细记录了研究乡村的 34 块露天耕地及 20 个大棚的具体信息, 并根据地形特征分类, 确保了种植策略与地块特性的匹配; 作物信息则涵盖了每种作物的亩产量、种植成本、销售单价及适宜种植的地块类型和季节等关键属性, 为分析作物的经济效益和适宜种植条件提供了基础; 市场销售与成本数据反映了 2023 年市场对不同作物的需求和种植的经济压力, 对制定未来种植策略至关重要; 历史种植数据整合了 2023 年的实际种植情况, 包括亩产量、种植面积、总产量及收益等, 作为后续模型构建和优化的基础。在进行模型构建之前, 对数据进行了彻底的清洗和预处理。包括地块信息的整合、作物数据的清洗、异常值的检测与处理等。

为了清晰直观地观察异常的数据情况, 将所给的不同农作物每亩地的产量值, 种植所需的成本以及销售的单价情况进行可视化的分析, 得到直观的柱状图如图 1 所示:

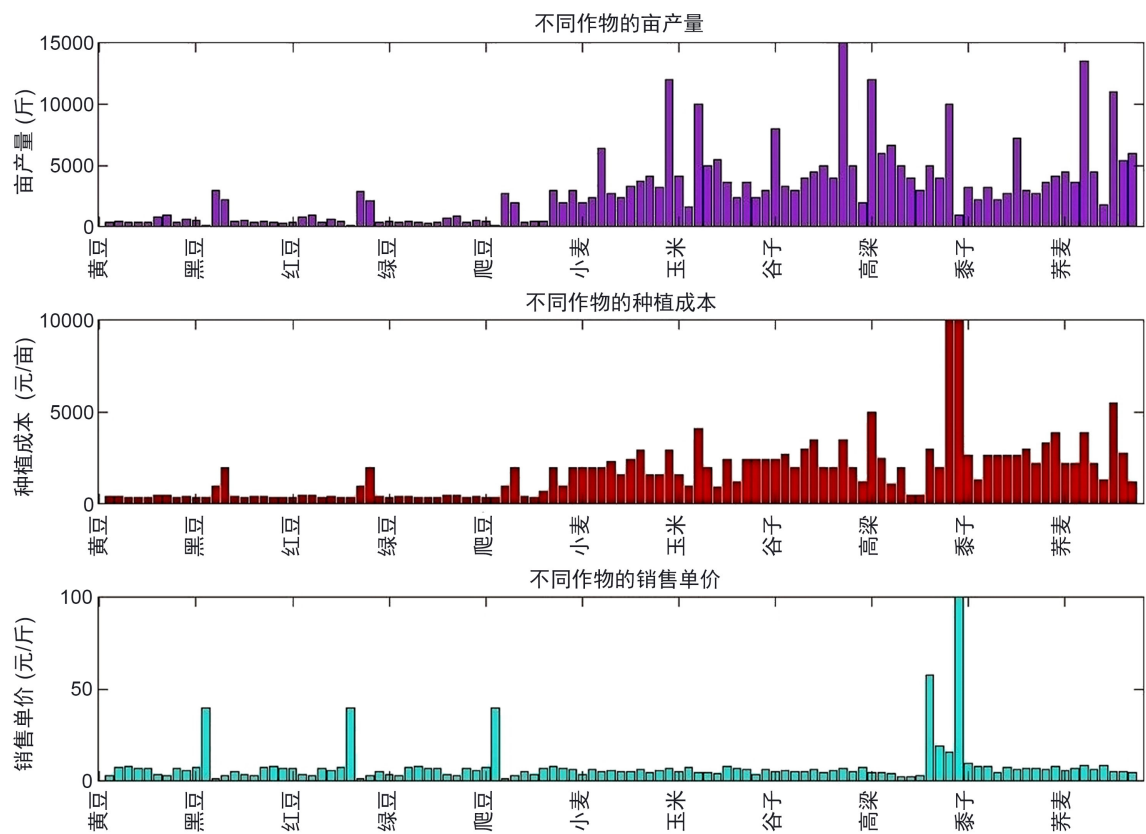


Figure 1. Visualization of crop yield per mu, planting cost and sales price
图 1. 作物亩产量、种植成本和销售价格的可视化结果

对所有可视化的数据进行分布方式的判定, 采用 Kolmogorov-Smirnov 检验。Kolmogorov-Smirnov 检验是一种非参数的统计检验方法, 是一种用于检验数据集是否服从某种分布的统计方法, 其中最常用的是检验数据集是否服从正态分布。通过检验, 亩产量、种植成本以及销售单价三大指标服从的分布形式如表 1 所示:

Table 1. Distribution patterns of the three major indicators
表 1. 三大指标分布规律

指标	亩产量	种植成本	销售单价
分布规律	非正态分布	非正态分布	非正态分布

由于三大指标均不服从正态分布, 因此异常值的判定引入箱型图进行判定, 用 Matlab 进行仿真得到的箱型图判定结果如图 2 所示。

在检测结果判定之后, 存在异常值, 对异常值的结果进行可视化的处理, 具体结果如下表 2 所示。
在异常值的检测里面属于菌类的是榆黄菇和白灵菇, 但二者在销售单价上存在明显的差价, 这可能与它们的产量和营养价值有关。农作物的售价是否合理取决于其品种、市场供需、品质。从初步结果上来看, 香菇、白灵菇和羊肚菌的售价处于正常的售价范围, 因此不需要对其进行处理。

计算了各项数据的均值、中位数、标准差、峰度和偏度等统计量, 以了解数据的集中趋势、波动情况以及分布形态。分析的结果如表 3 所示。

Table 2. Outlier detection results
表 2. 异常值检测结果

作物名称	销售单价
芥麦	40
芥麦	40
芥麦	40
榆黄菇	57.5
香菇	19
白灵菇	16
羊肚菌	100

Table 3. Data characteristics of final descriptive analysis results
表 3. 最终描述性分析结果数据特征

	均值	中位数	标准差	峰度	偏度
亩产量	2990.8	2400	3047.9	6.44	1.78
种植成本	1686.1	1320	1636.14	13.47	2.64
销售单价	8.2	6	11.91	37.85	5.54

作物亩产量、种植成本和销售单价的异常值检测

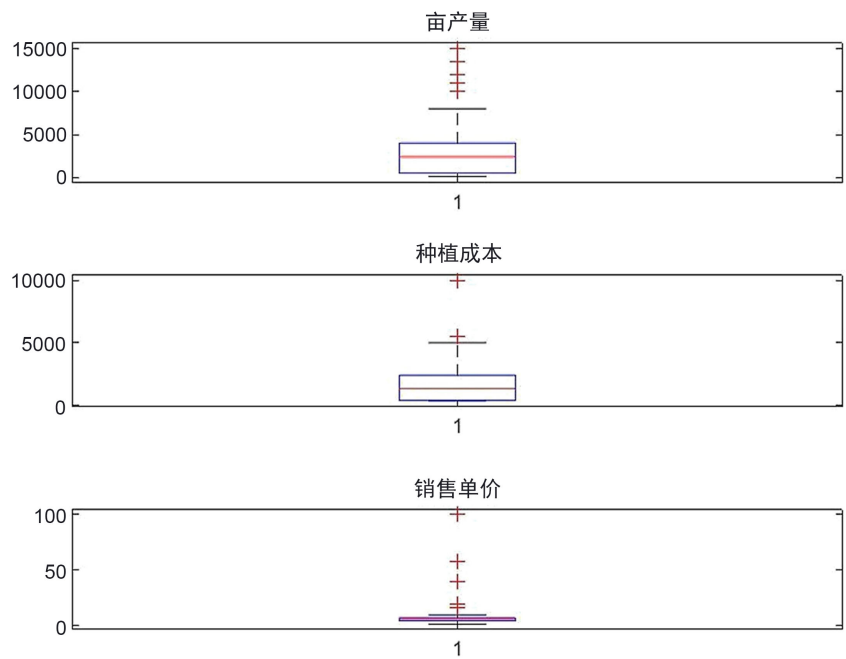


Figure 2. Box plot test results
图 2. 箱型图检测结果图

亩产量：由于中位数低于平均值，中位数比均值低，表明分布具有右偏(正偏)。标准差高达 3048 斤，这表明了不同作物间的产量存在显著差异。偏度为 1.78 揭示了亩产量的分布有显著的右偏(正偏)，即大部分数据倾向于较低的值，而高值则出现较少。峰度为 6.44 则表明分布的尖峰程度较高，数据点集中在均值

附近的情况比正态分布更加显著。

种植成本: 标准差为 1636.14 元/亩, 说明种植成本波动很大, 不同作物间的成本差异显著。偏度为 2.64, 表明种植成本分布有更为明显的右偏(正偏), 意味着大多数作物的成本相对较低, 少部分作物的种植成本非常高。峰度为 13.47, 表示分布更加尖锐, 较多数据集中在均值附近, 同时存在一些极端值。

销售单价: 标准差为 11.91 元/斤, 这指出作物销售价格在不同种类间存在显著差异, 部分作物的售价远高于平均水平。偏度为 5.54, 这一数值突出了销售价格的右偏性, 即大多数作物的售价偏低, 而少数作物的售价异常高。峰度为 37.85, 表明销售价格分布具有极大的尖峰, 尤其是在较低价格区间, 同时少数高价位作物构成了数据的离群点。

3. 模型的建立与求解

3.1. 模型的构建

基于对数据的预处理, 针对农作物种植决策中的一系列复杂问题, 构建了一个静态销售价格、考虑未来农作物产量、成本和销售价格的波动以及作物关联性的多阶段优化模型。

3.1.1. 静态销售价格模型

若某作物每季的总产量超过相应的预期销售, 且超出部分将面临销售难题, 可按照滞销部分浪费和超过部分按 50% 降价出售两种情况进行计算, 从而寻找最优的种植策略。

第一种情况, 当某种作物的实际产量超过预期销售量时, 超出部分无法正常销售, 目的是 2024 年到 2030 年间的种植策略, 在考虑作物的产量、成本、销售价格、预期销售量、作物间的轮作约束、以及豆类作物的轮作要求等因素下使得利益最大化。目标函数: 使得 2024~2030 年收益最大化:

$$\max \sum_{t=1}^T \sum_{i=1}^n \sum_{j=1}^m \left[\min(Y_j \times A_{i,j,t} \times E_{i,j,t}) \times P_j^s - C_j^c \times A_{i,j,t} \right] \quad (1.1)$$

约束条件:

(1) 设每一种农作物的预期销售量均为 80%, 即

$$E_{expect} = 0.8 \times Y_j \times A_i \quad (1.2)$$

(2) 最小种植面积为了避免种植不集中的问题, 设每种农作物在单个地板上最小的种植面积为 0.1 亩:

$$A_i \geq 0.1 \quad (1.3)$$

(3) 不重茬种植约束

每块地连续种植可能会导致减产, 因此不能连续两年种植同一种农作物:

$$C_{i,t-1} \neq C_{i,t} \quad (1.4)$$

(4) 豆类作物轮作要求

根据资料, 每块地三年内至少需要种一次豆类。豆类作物的集合记为 S , 若某地最近三年一直没有种植豆类作物, 则下年必须种植豆类: (这里用 1 表示种植豆类)

$$\sum_{k=t-2}^t 1\{C_{i,k} \in S\} \geq 1 \quad (1.5)$$

第二种情况, 当某种作物的实际产量超过预期销售量时, 考虑超过部分作物产量按 50% 降价出售的情形。模型的目标是为 2024 至 2030 年的种植策略进行优化, 在满足一定的种植约束条件下, 最大化种植收益。

目标函数：2024 至 2030 年的总种植收益最大化。

总收益公式：对于第 i 块地在第 t 年种植作物 C_j ，它的收益由以下公式表示：

$$R_{i,j,t} = \min(Y_j \times A_{i,j,t}, E_{i,j,t}) \times P_j + \max(0, Y_j \times A_{i,j,t} - E_{i,j,t}) \times 0.5 \times P_j - C_j^c \times A_{i,j,t} \quad (1.6)$$

总的生产量 A : $A = Y_j \times A_{i,j,t}$;

预期的销售量: $E_{i,j,t} = 0.8 \times Y_j \times A_{i,j,t}$;

正常的销售部分 B : $B = \min(Y_j \times A_{i,j,t}, E_{i,j,t}) \times P_j$;

降价的销售部分 C : $C = \max(0, Y_j \times A_{i,j,t} - E_{i,j,t}) \times 0.5 \times P_j$;

总收益 W : $W = B + C - D$

约束条件与轮作要求：

(1) 地块面积限制：每块地的总种植面积不能超过地块面积。

(2) 不重茬约束：每块地连续种植可能会导致减产，因此不能连续两年种植同一种农作物。

(3) 豆类轮作约束：根据资料，每块地三年内至少需要种一次豆类。豆类作物的集合记为 S ，若某地最近三年一直没有种植豆类作物，则下年必须种植豆类(这里用 1 表示种植豆类)：

$$\sum_{k=t-2}^t 1\{C_{i,k} \in S\} \geq 1, \forall i, \forall t$$

3.1.2. 基于不确定性因素的优化模型

考虑未来作物产量、成本和销售价格的波动，以及不同作物的销售增长趋势等不确定性的因素，建立优化模型。为寻找农作物的最优种植方案，以下是模型建立的过程：

(1) 未来 7 年内每块地的种植方案，以最大化每年的种植总收益。需要考虑作物的产量、种植成本、销售价格、预期销售量等随时间的变化，同时满足作物轮作和重茬限制。为了模拟农作物的价格、产量、种植成本和预期销售量随时间变化，代码实现了 `adjust parameters` 函数。该函数根据作物种类和年份，调整这些参数，参数所需要的公式。

(2) 目标函数为通过选择最优作物，从而达到最大化种植收益。接着与上文同理设置种植约束条件，包括每块地的种植面积不能超过地块面积的约束，每个地方三年保修种植一次，连续两年不能种植同一类农作物。

(3) 利用 `find best crop` 函数在每一年的每一块地上选择收益最高的作物，该函数的目标是为每个地块每年确定最具经济效益的作物种植方案。这一过程涉及一系列细致的步骤，确保选择的作物能够在特定的环境和市场条件下带来最大的回报。

3.1.3. 作物关联优化模型

前面所描述的优化种植方案基础上，研究种植成本、亩产量和销售单价之间的相互影响关系，可建立以下两种模型：第一种方法是将亩产量作为因变量，种植成本和销售单价为自变量，分析这两个因素是如何对亩产量产生影响，并得出因变量与自变量之间的线性关系第二种方法是将销售单价看作因变量，种植成本作为自变量，研究种植成本对销售单价的影响并分析出二者之间的线性关系。

通过计算三者之间的相关性矩阵，可以初步判断其线性相关性，具体的热力图结果如图 3 所示。

建立模型目的是在优化的种植方案基础上，以农作物的经济收益最大化为目标函数，同时受豆类作物轮作，避免重茬种植，季节限制的约束。经济收益是在基于每季作物的产量、销售单价、种植成本等因素。总收益为种植的经济收益减去种植成本。其基本公式如下：

$$R_{ij} = \text{收益} = (P_j \times A_j \times S_j) - (C_j - A_j) \quad (1.7)$$

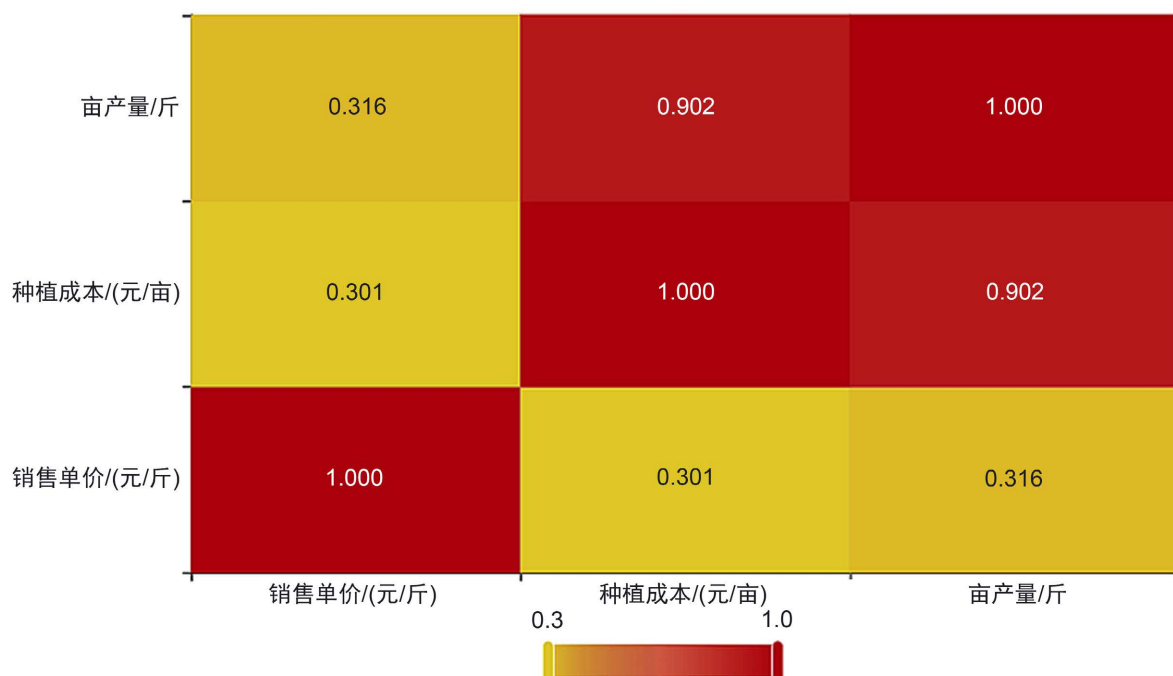


Figure 3. Heat map
图 3. 热力图

设农作物的预期销售量为其产量的 80%，超过的 20%将以降价 50%销售，因此完整的收益公式为：

$$R_{ij} = \min(P_j \times A_{ij}, 0.8 \times P_j \times A_{ij}) \times S_j + \max(0, P_j \times A_{ij} - 0.8 \times P_j \times A_{ij}) \times 0.5 \quad (1.8)$$

同样是对场地面积、轮作、重茬种植三个方面的约束，约束条件也同上。

模型中采用固定的公式根据成本计算销售单价和亩产量：

销售单价：

$$S_j = 2.5 + 0.0032 \times C_j \quad (1.9)$$

亩产量：

$$P_j = 300 + 0.045 \times C_j \quad (1.10)$$

3.2. 算法的选择

针对未来作物产量、成本和销售价格波动，以及不同作物的销售增长趋势等不确定因素建立优化模型，结合粒子群优化算法进行求解，优化未来 7 年内的种植方案，以每年的收益最大化为目标函数，每块地的种植方案设计需要通过选择最优的农作物，从而达到最大化的收益。在考虑上述不确定性因素的基础上，本文引入研究了农作物之间的替代性和互补性，构建了一个分析种植成本、销售单价和亩产量之间相互关系的多阶段模型，通过多元线性回归方法进行求解。

4. 结果与讨论

4.1. 静态销售价格模型结果

通过求解建立的静态销售价格模型，第一种情况下，每年一季度和第二季度，都会根据地块来决定种植方案，地块的面积会根据其总面积和可以使用的农作物的产量和收益来进行相应的调整。每个地块在每

个季度的种植, 其种植面积不能超过地块的总面积, 如果种植的总和超过了给定的目标值, 也就是超过了土地的面积的上限, 则会根据缩放种植面积的办法使其满足约束。第二种情况下, 从 2024 年到 2030 年每年都会生成一个种植方案, 对于每块地的每一季, 会按照收益最大化的原则分配作物。最终的优化目标是通过 2024 至 2030 年的种植策略, 使每年、每块地的作物种植收益最大化, 同时满足作物的轮作约束。

4.2. 不确定性因素优化模型结果

通过采用粒子群的优化算法进行连续优化处理, 将粒子群优化问题用于处理农作物种植方案的优化问题, 在每一年的每一块地选择合适的作物以及合适的种植面积, 来获取最大化的利益, 通过粒子群的算法迭代进行搜索, 寻找全局的最优解对应的种植方案, 所找到的种植方案即是每个地块在每年每一个季度的最优种植策略, 同时优化之后产生的种植方案也包括农作物和种类和面积, 同时满足所有的约束情况, 进而得到的该乡村 2024~2030 年农作物在诸多不确定因素情况下的最优种植方案。

4.3. 作物关联性的优化结果

在不确定性因素优化模型基础上进一步引入了农作物之间的替代性和互补性, 构建了一个更为复杂的多阶段关联优化模型, 通过线性回归模型对不同作物之间的相关性进行了判定。针对每一年每一个季度, 模型都需要对所有的可种植的地块以及可种植的农作物逐步进行评估, 最终得出一个最优的种植方案, 通过优化模型, 能够合理分配地块和作物资源, 确保在满足轮作要求、避免重茬种植等约束下, 实现种植收益的最大化。

5. 结论

本文使用 Matlab 通过数据预处理、优化模型构建, 对现代农作物种植策略进行了探讨与优化。构建了一个多阶段优化模型, 并结合粒子群算法与多元线性回归方法, 针对农作物种植决策中的复杂问题提出解决方案, 为农业决策者提供了有力的理论支持和实践指导。研究结果表明, 在考虑农作物轮作、避免连作等生态学原则的基础上, 通过合理调配作物与土地资源, 可以显著提升种植收益。引入了不确定性因素和作物间的关联性分析。模型能够根据实际情况动态调整种植方案, 从而有效应对市场波动和气候变化带来的挑战。同时, 考虑作物间的替代性和互补性, 进一步优化了种植策略, 提高土地利用效率, 增强经济效益。当产量超出预期销售量时, 应该采取降价销售策略提升整体收益。通过模型对多年的动态优化, 能够为不同土地提供最优的种植方案。未来, 可以进一步探索极端气候条件下的种植策略优化, 并考虑结合大数据和人工智能等先进技术, 以进一步提高模型对市场波动的预测精度和适应性。

基金项目

辽宁科技大学大学生创新创业训练计划专项经费资助(X20251014600X)。

参考文献

- [1] 胡树峰. 生物技术在现代农作物种植方面的应用[J]. 河北农机, 2022(16): 133-135.
- [2] 符翠芬. 农作物种植中土肥技术的优化与提升[J]. 种子科技, 2024, 42(10): 84-86.
- [3] 杨心怡, 杨铁军, 徐阳, 等. 基于粒子群算法相位寻优的同步定相振动控制仿真研究[C]//中国振动工程学会. 第十五届全国振动理论及应用学术会议摘要集. 2023: 1. <https://doi.org/10.26914/c.cnkihy.2023.111761>
- [4] 冯玮榕. 水稻种植技术优化及种植效益提升策略研究[J]. 种子科技, 2024, 42(6): 131-133.
- [5] 鲍树忠. 农业种植因素对种植结构的影响及优化策略[J]. 农业与技术, 2020, 40(13): 90-91.