

基于多目标优化的农作物种植策略研究

宋瑞雪 柯妤菡 黄涛 陈保周*

武汉轻工大学 数学与计算机学院

DOI:10.12238/as.v8i2.2708

[摘要] 在农业现代化与可持续发展背景下,耕地有限、气候多变、市场需求波动等带来挑战,使得优化农作物种植策略至关重要。本研究基于2024年全国大学生数学建模竞赛C题数据,创新构建多元回归与双目标线性规划模型,通过引入NSGA-II遗传算法对该模型进行深度优化。最终,该算法生成帕累托前沿,避免过收敛,提升优化效率。实验结果表明,本文所提出的最优种植方案即便在销售超预期且遇停滞期时,也能实现最高利润。这一结论为农业种植决策提供了科学有效的参考依据。

[关键词] 多元回归分析; 双目标线性规划模型; NSGA-II算法; Pareto前沿

中图分类号: F223 文献标识码: A

Research on crop planting strategy based on multi-objective optimization

Ruixue Song Yuhan Ke Tao Huang Baozhou Chen*

School of Mathematics and Computer Science, Wuhan Polytechnic University

[Abstract] In the context of agricultural modernization and sustainable development, challenges such as limited arable land, variable climate, and fluctuating market demand make it imperative to optimize crop planting strategies. This study innovatively constructs multiple regression and dual objective linear programming models based on the data of question C in the 2024 National College Student Mathematical Modeling Competition. The model is deeply optimized by introducing NSGA-II genetic algorithm. Ultimately, the algorithm gives rise to Pareto front, avoiding over convergence and improving optimization efficiency. The experimental results demonstrate that the optimal planting scheme proposed in this article can achieve the highest profit even when sales exceed expectations and encounter stagnant periods. This conclusion provides a scientific and helpful reference for agricultural planting decisions.

[Key words] Multiple Regression Analysis; Bi-objective Linear Programming; Non-dominated Sorting Genetic Algorithm II; Pareto Frontier

引言

面对气候变化和人口增长挑战,科学制定农作物种植策略对提升农业效率、保障粮食安全和增加农民收入至关重要。2024年全国大学生数学建模竞赛C题聚焦华北地区乡村种植问题,要求制定最优方案。传统策略依赖经验,而现代数学建模技术(如多元回归、双目标线性规划)结合NSGA-II遗传算法,为优化种植策略提供了新途径,其Pareto前沿生成能力和避免过早收敛的特性在多目标优化中表现优异。李忠峰等^[1]在研究煤化问题时,采用改进的NSGA-II多目标优化方法,综合考虑预期销售量、种植成本和气候变化等因素,设计最优种植方案并预测不同销售情况下的最大利润。成果为当地农民提供数据驱动的种植决策支持,助力农业可持续发展。

1 方法

本文基于2024年全国大学生数学建模竞赛C题数据,研究多

目标优化的农作物种植策略,分为三种情况:

(1) 超出预期销售量部分滞销。(2) 超出部分按2023年销售价格的50%降价出售。(3) 在第二种情况基础上,考虑农作物相关变量的波动性。

1.1 空间预测模型建立

本文通过蒙特卡洛法模拟2023年销售量,将季节、地块和作物等因素转化为变量,基于情况(1)和(2)建立线性回归模型,并将问题要求转换为约束条件。最后,利用遗传算法优化这些模型^[2],经过多次迭代求出最优解,得出最优种植方案。

同时,本文考虑农作物的可替代性和互补性,设定约束条件并构建多元回归模型,通过F检验验证其有效性。随后,利用拉丁超立方抽样方法^[3]生成多组不确定性场景,并采用快速前代消除法精简出4组代表性场景。最终,通过Pareto优化方法^[4]得出最优解,为决策提供了科学依据,降低过拟合风险。

其次,本文建立了双目标线性规划模型(Bi-objective Linear Programming, BLP),通过正交实验^[5](16次实验,215个水平)分析农作物波动性的影响。基于滞销或降价销售情况建立目标函数,并设定种植风险的约束条件。采用NSGA-II算法^[6]优化两个目标:最大化平均利润和最小利润。通过快速非支配排序分层筛选最优解集,并利用Pareto优化方法得出最佳方案。

1.2 目标函数的建立

本研究已知农作物的种植成本、亩产量和销售价格,并假设这些参数未来保持恒定。基于决策变量(不同农作物的种植面积)与参数的线性关系,本文针对三种情况构建目标函数和约束条件。情况(1)和(2)的目标函数分别如式(1)和(2)所示,约束条件如式(3)~(7)所示。

$$\max f_1 = \sum_t \sum_i \sum_j (P_{ijt} \times Y_{ijt} - B_{ij} \times x_{ijt}) \quad (1)$$

$$\max f_2 = \sum_i \sum_j [P_{ijt} \times Y_{ijt} - \frac{P_{ijt}}{2} \max(Y_{ijt} - T_{ijt}, 0) - B_{ij} \times x_{ijt}] \quad (2)$$

(1)作物轮换的约束:每块地三年内必须要种植一次豆类作物。

$$S_j \leq \sum_{t=1}^{t+2} x_{i,j,t} < 2S_j, \quad \forall i \quad (3)$$

(2)种植密度的约束:为了方便地块之间的管理,每种农作物的种植面积不宜过小。

$$x_{ij} \leq x_{\min}, \quad \forall i, j \quad (4)$$

(3)地块间的约束:平旱地、梯田和山坡地种植单季粮食作物如式(5-1)所示,水浇地种植单季水稻如式(5-2)所示,水浇地两季蔬菜种植,第一季可种多种蔬菜(除大白菜、红萝卜和胡萝卜),第二季仅限这三种蔬菜如式(5-3) (5-4)所示。

$$\begin{cases} \sum_i x_{i,j,t} \leq A_j, \forall j \in \{1, 2, \dots, 25, 26\} & (5-1) \\ \sum_i x_{i,j,t} \leq A_j, \forall j \in \{27, 28, \dots, 34\} & (5-2) \\ \sum_{i \in \{17, \dots, 34\}} y_{i,j,t} \leq 1, \forall j \in \{27, 28, \dots, 34\} & (5-3) \\ \sum_{i \in \{35, \dots, 37\}} y_{i,j,t} \leq 1, \forall j \in \{35, 36, 37\} & (5-4) \end{cases} \quad (5)$$

$$\sum_{i=16} y_{i,j,t} + \frac{1}{2} \left(\sum_{i \in \{17 \sim 34\}} y_{i,j,t} + \sum_{i \in \{35 \sim 37\}} y_{i,j,t} \right) \leq 1 \quad (6)$$

其中, y_{ijt} 表示于 t 年在 j 地块上是否种植第 i 种作物。

普通大棚第一季种植多种蔬菜(除大白菜、红萝卜和胡萝卜),第二季种植食用菌。

$$\begin{cases} \sum_{i \in \{17, \dots, 34\}} x_{i,j,t} \leq A_j, \forall j \in \{E^{1 \sim 16}\} \\ \sum_{i \in \{38, \dots, 41\}} x_{i,j,t} \leq A_j, \forall j \in \{E^{1 \sim 16}\} \\ \sum_{i \in \{17, \dots, 34\}} y_{i,j,t} \leq 1, \forall j \in \{E^{1 \sim 16}\} \\ \sum_{i \in \{38, \dots, 41\}} y_{i,j,t} \leq 1, \forall j \in \{E^{1 \sim 16}\} \end{cases} \quad (7)$$

其中, $x_{(i,j,t)}$ 表示乡村某种农作物的种植面积(单位:亩), i 表示农作物的种类(=1, 2, 3..., 41), j 表示种植地块的编号(=1, 2, 3..., 82), t 表示种植该作物的年份, S_i 表示 A_i 的面积。 P_i 是第 i 种作物的销售价格(单位:元/斤); Y_i 是 i 作物的预期销售量(单位:斤); B_i 是第 i 种作物的种植成本(单位:元/亩)。

本文设定迭代次数为800次,交叉率为0.7,变异率为0.3,种群规模为30,以平衡计算效率与搜索广度。

表1 正交实验方案

小麦、玉米销量年增长率	其他作物销售量波动	亩产量	食用菌价格年缩减率(除羊肚菌)
5%	-5%	-10%	1%
6%	-4%	-9%	2%
7%	-3%	-8%	3%
8%	-2%	-7%	4%
9%	-1%	-6%	5%
10%	0%	-5%	1%
5%	1%	-4%	2%
6%	2%	-3%	3%
7%	3%	-2%	4%
8%	4%	-1%	5%
9%	5%	0%	1%
10%	-5%	1%	2%
5%	-4%	2%	3%
6%	-3%	3%	4%
7%	-2%	4%	5%
8%	-1%	5%	1%

本文利用L16(215)正交表进行正交实验,其中16表示实验次数,215表示各因素的水平数。通过16次实验,均匀全面地测试

各因素的不同水平组合, 确保每个水平组合出现次数相等, 保证实验的均匀性和平衡性。结果如表1所示。

本文将农作物种植成本和蔬菜、羊肚菌的价格年增长率设为5%, 基于两种情况及相同约束条件进行分析。为降低未来七年种植风险, 引入第二标准: 每年种植方案在所有正交情况下的利润是否最大。仅满足利润最大化的方案视为最优。

目标函数如式(8)和(9)所示。

$$\max f_3 = \frac{1}{t} \sum_{t=1}^7 profit_t - W_t \quad (8)$$

$$\max f_4 = \min \{profit_t - W_t\}, t \in \{1, \dots, N\} \quad (9)$$

其中, f_3 表示在所有正交情况中的平均利润, $profit_t$ 表示在第 t 年的收入, W_t 表示第 t 年的输出, f_4 表示在所有的正交情况下最小利润达到最大。

在商品市场中, 存在众多可替代商品时, 价格变动可能导致销售量降低, 消费者更倾向于选择性价比高的替代品。因此, 预测销售量时需考虑农作物间的可替代性, 并将其作为约束条件。假设在第 i 年第 k 个季度具有可替代性的两种作物的总产量分别为 $L_{i,k}^{(1)}$ 、 $L_{i,k}^{(2)}$, 应该满足式(10)。

$$\min \left\{ \frac{L_{i,k}^{(1)}}{L_{i,k}^{(1)} + L_{i,k}^{(2)}}, \frac{L_{i,k}^{(2)}}{L_{i,k}^{(1)} + L_{i,k}^{(2)}} \right\} \geq L_{\min} \quad (10)$$

在同一季度, 互补性农作物的销售收益更高。将互补性作为约束条件加入总利润函数, 假设两种互补农作物的收益因子为(不同农作物收益因子不同, 无互补种植时为0)。因此, 目标函数为:

$$W = \sum_{i,j,t} [P_{ijt} \times x_{ijt} \times r - 0.5P_{ijt} \max(x_{ijt})] \quad (11)$$

由此建立如下多元回归模型:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon \quad (12)$$

其中, X_1 、 X_2 为销售价格和种植成本, Y 为预期销售量。

原假设: $H_0: \beta_1 = \beta_2 = 0$, 备择假设: $H_1: \beta_1$ 、 β_2 不全为0。检验统计量为:

$$F = \frac{S_R / p}{S_E / (n - p)} \sim F(p, n - p) \quad (13)$$

进行回归分析后得到回归方程。 $\hat{y} = 0.496x_1 + 0.009x_2$

+ 3.453 检验后得到统计量 $F=15.8$, 且 $P<0.0001$ 。

表2 方差分析

	平方和	自由度	均方	F	显著性
回归	37920	3	12640		
残差	54400	68	800	15.8	0
统计	92320	71			

由于 β_1 、 β_2 不全为0, 所以回归效果显著。进一步对系数进行检验, 得到 X_1 、 X_2 的显著性均小于0.05, X_1 、 X_2 对 Y 的回归效果显著。

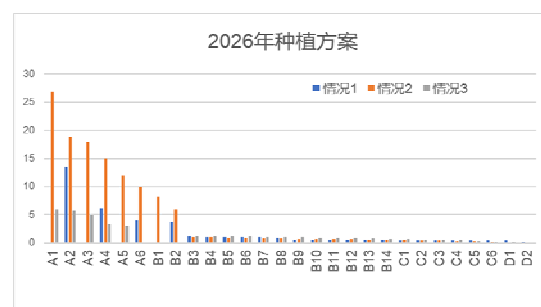
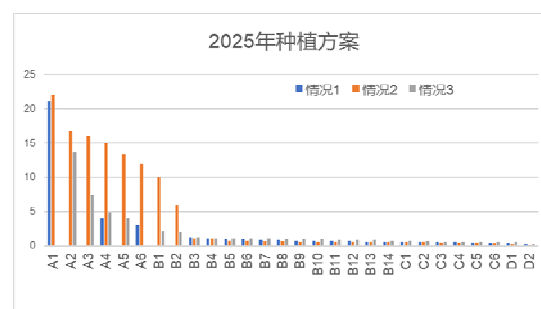
为了减少不确定因素的场参数, 本文采用拉丁超立方抽样的方法来提高采样效率。拉丁超立方抽样可以确保每个变量在其取值范围内均匀分布, 在每一维上, 抽样点不会重复, 并且每个变量的取值范围都被采样点所覆盖^[3]。

1.3 问题步骤

针对问题一, 本文使用SPSS处理附件一、二数据, 删除异常值并获取描述统计量。通过Shapiro-Wilk检验(显著水平 $P=0.05$), 得出预期销售量不符合正态分布($P<0.05$)。在遗传算法优化中, 简化多维决策变量处理, 设计交叉策略生成新种群后代, 并修正不满足条件的解以符合农业种植约束(如季节性限制、土地使用政策和种植面积上限)。

2 结果分析

本文发现: 情况(1)在迭代740次后收敛, 最优解为300万; 情况(2)在600次迭代后稳定, 最优解收益为1500万。目标函数最大值对应的种植方案即为最优方案。2025-2030年各地块种植面积见图1。



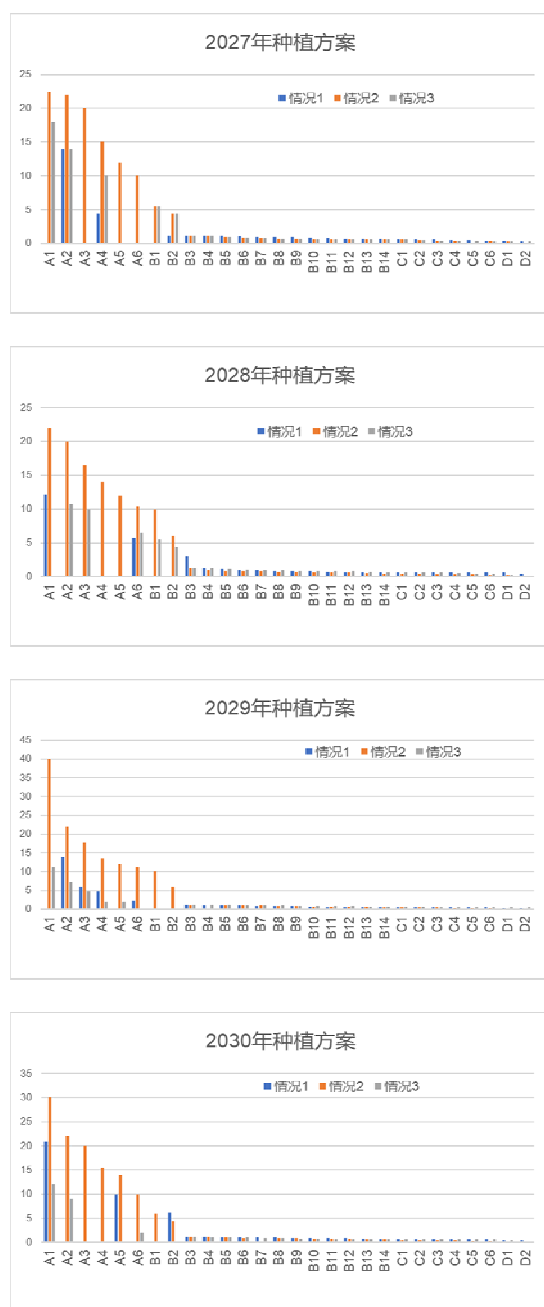


图1 不同情况下2025年-2030年的最优种植面积图

情况(1): 算法初期收敛快, 500次迭代后利润稳定在340万左右, 平均利润与最小利润曲线接近, 表明种群性能差异小, 遗传算法有效促进了最优解传播。

情况(2): 算法在460代迭代后收敛至1800万左右, 优化过程

高效且适应多目标复杂关系。

3 总结

本文通过多元回归与双目标线性规划模型给出了华北地区某乡村的农作物种植方案, 得出以下结论:

(1) 不考虑不确定性及风险, 销售量超出预期滞销时最大收益约300万; 超出预期折半销售时最大收益约1500万。(2) 考虑市场风险和不确定性后, 超出预期折半销售的利润显著高于滞销情况, 最高可达1800万。(3) 综合考虑作物互补性、可替代性及销售量、价格和成本的相关性后, 最优种植方案的收益超越情况(2), 这表明改进方案更全面地适应了农业复杂性和市场导向, 提升了经济效益。

[项目来源]

湖北省教育科学规划2022年度一般课题“基于多模态案例库建设的高等数学课程思政教学改革创新研究与实践”(NO.2022GB046); 武汉轻工大学引进(培养)人才科研启动项目“基于矩阵低秩逼近的GPS坐标时间序列分析”(NO.2024K2029); 武汉轻工大学校级教研项目“基于LinkWall智慧教室的公共数学类课程教学资源库的建设”(NO.XM202313)。

[参考文献]

[1]李忠峰,刘俊,金辉,等.改进NSGA-II的炼焦配煤多目标优化方法[J].制造业自动化,2024,46(05):103-108.

[2]方红,张伟.用于多元线性回归分析的遗传算法[J].计算与理论纳米科学杂志,2019,15(1):1-5.

[3]郑鹏,刘健,宋维,等.拉丁超立方抽样评估方法改进研究[J].核电子学与探测技术,2017,37(7):7-12.

[4]曾三友,蔡振华,张青,等.一种评估近似Pareto前沿多样性的方法[J].软件学报,2008,19(6):1301-1308.

[5]滕海英,祝国强,黄平,等.正交试验设计实例分析[J].药学服务与研究,2008,08(1):75-76.

[6]向芷恒,王秉哲,雪景州,等.基于NSGA-II算法解决多目标优化实际应用的研究[J].应用数学进展,2023,12(10):4195-4207.

作者简介:

宋瑞雪(2003--),女,汉族,河北邯郸人,软件工程专业本科在读。

*通讯作者:

陈保周(1984--),男,汉族,河南灵宝人,博士,讲师。近期主要研究方向为时间序列分析、机器学习等。