Research on Crop Planting Decision Model Based on Linear Programming and Genetic Algorithm

Huiyang Jiang [1, *, #], Yihan Tian [2, #]

[1] Software College of Northeastern University, Northeastern University, Shenyang, China, 110169

[2] School of Computer Science and Engineering Northeastern University, Northeastern University, Shenyang, China, 110169

* Corresponding Author Email: huiyangjiang0725@163.com

[#]These authors contributed equally.

Abstract. This study presents an optimization of crop planting strategies for a rural area in the mountainous region of North China, covering the period from 2024 to 2030, with the objective of maximizing profits amid resource constraints and market volatility. We employed a linear programming model to identify profit-maximizing planting strategies while incorporating multiple constraints. Utilizing an enhanced genetic algorithm, the model projected maximum profits of 53 million and 58 million yuan under different scenarios. Furthermore, we developed a bi-objective robust optimization model informed by orthogonal experimental design, utilizing an improved NSGA-II algorithm to address market uncertainties. This approach yielded average profits of 54.4 million and 61 million yuan, with minimum profit guarantees of 54.25 million and 59.5 million yuan in adverse conditions, effectively managing planting risks. The findings provide a framework for optimizing crop strategies in mountainous regions of North China, promoting economic growth and sustainable agricultural practices while enhancing decision-making and market resilience.

Keywords: Linear Programming, Improved Genetic Algorithm, NSGA-II, Orthogonal Experiment.

1. Introduction

The North China Mountain area, an important grain production base, plays a crucial role in the country's economic and social development. To promote sustainable rural economic development, it's essential to utilize arable land resources wisely and develop organic planting industries tailored to local conditions. With its low annual temperature, semi-arid and semi-humid monsoon climate, and abundant arable land, this region supports diverse crop planting, especially with the introduction of smart greenhouses. Proposing a regional agricultural planting layout plan is vital for the healthy development of agriculture in this area.

Numerous scholars have studied optimizing regional arable land planting structures, from rational land use to developing organic industries. Sun applied dynamic programming to determine optimal crop planting plans, focusing on multi-objective and indivisible optimization problems [1]. Yu et al. used grey correlation and principal component analysis to optimize the planting structure of main crops in Manas County [2]. Randall et al. employed genetic algorithms to create a multi-objective optimization model for Manas County, maximizing benefits under resource and environmental constraints [3]. Shi also used genetic algorithms to solve resource allocation and economic benefit maximization issues [4].

Despite advancements in optimizing crop planting structures, existing studies have limitations, such as insufficient consideration of uncertainty factors and overly complex models that hinder quick application. To address these issues, we propose a study on optimizing crop planting in the North China Mountain area using linear programming and genetic algorithms. We developed a linear programming model that accounts for planting season, crop type, and rotation rules to maximize planting profit. Additionally, we created a dual-target robust optimization model using orthogonal test design to balance economic benefits and risk control in uncertain environments. Our improved genetic algorithms successfully solved both models, demonstrating good convergence and stable maximum profit under varying conditions.

2. The Establishment of the Models

2.1. Linear Programming Model

Linear programming optimizes problems by defining a linear objective function and constraints, aiming to find the optimal decision variable that maximizes or minimizes the objective [5]. In this study, we use a linear programming model to optimize crop planting strategies and maximize profit for a rural area in North China from 2024 to 2030. The flowchart of the model is shown in Figure 1.

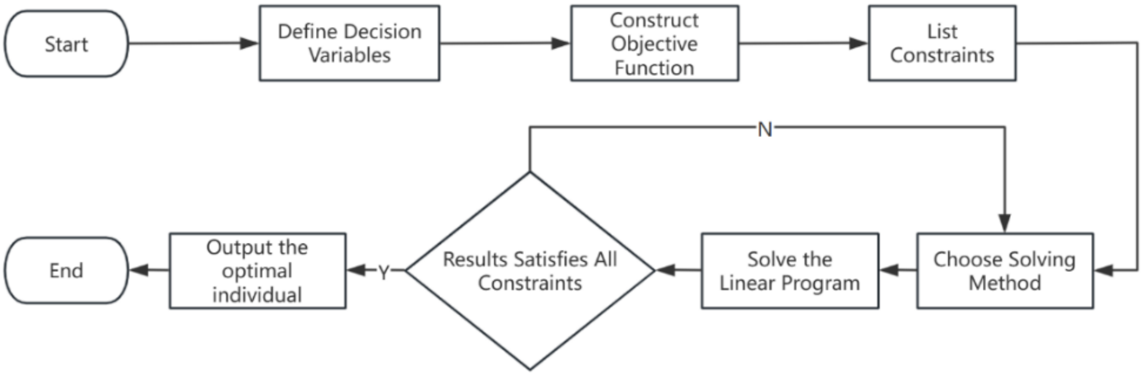

Figure 1. Linear programming model flowchart

Consider the decision variables as follows: n represents the name of the plot, y represents the year, s represents the planting season, and p represents the crop number. Therefore, $x_{n,y,s,p}$ represent the area in acres of crop p planted in season s on plot n in year y.

To optimize planting profits from 2024 to 2030, we define income as total revenue and expense as total costs over this period. $price_{y,s,p}$ is the selling price per unit of crop p in season s of year y, and $E_{n,y,s,p}$ is the cost of planting crop p on plot n in season s of year y. The objective is to maximize profit z.

$$maxz = income - expense \tag{1}$$

$$income = \sum_y \sum_s \sum_p \left(price_{y,s,p} * \sum_n x_{n,y,s,p}\right) \tag{2}$$

$$expense = \sum_y \sum_n \sum_p \sum_s E_{n,y,s,p} * x_{n,y,s,p} \tag{3}$$

We constrain the planting season s based on the plot type n: $s \leq 1$ for flat dry land, terraced fields, and sloping land, and $s \leq 2$ for both ordinary and smart greenhouses.

$$\begin{cases} s \leq 1, if n \in A1, \dots, A6, B1, \dots, B14, C1, \dots, C6 \\ s \leq 2,\ else \end{cases} \tag{4}$$

For irrigated land, the planting season s is limited to 1 if crop p is rice (16), and to 2 for other crops.

$$\begin{cases} s \leq 1, if p \in 16 \\ s \leq 2,\ else \end{cases}, n \in D1, \dots, D8 \tag{5}$$

On flat dry land, terraced fields, and sloping land, only one grain crop season is allowed annually, constraining the crop types based on the plot name n.

$$\begin{cases} x_{n,y,s,p} \geq 0, if n \in A1, \dots, A6, B1, \dots, B14, C1, \dots, C6, p \in 1,2, \dots, 15 \\ x_{n,y,s,p} = 0, else \end{cases} \tag{6}$$

We limit crop types in irrigated land to one rice or two vegetables per year, in ordinary greenhouses to one vegetable and one edible fungus, and in smart greenhouses to two vegetable crops.

$$\begin{cases} x_{n,y,s,p} \geq 0, if n \in D1, \ldots, D8, E1, \ldots, E16, s = 1 \\ x_{n,y,s,p} \geq 0, if n \in F1, F2, F3, F4 \\ x_{n,y,s,p} = 0, else \end{cases} \quad (7)$$

The restrictions on the crop types planted in the second season of irrigated land are as follows:

$$\begin{cases} x_{n,y,s,p} \geq 0, if n \in D1, \ldots, D8, s = 2 \\ x_{n,y,s,p} = 0, else \end{cases} \quad (8)$$

The restrictions on the crop types for the second season in ordinary greenhouses are as follows:

$$\begin{cases} x_{n,y,s,p} \geq 0, if n \in E1, \ldots, E16, s = 2 \\ x_{n,y,s,p} = 0, else \end{cases} \quad (9)$$

Due to the prohibition of continuous cropping of the same crop on the same plot (including greenhouses), single-season crops cannot be planted consecutively for two years, and other crops cannot be planted consecutively for two seasons. Therefore, we impose the following constraints:

$$\begin{cases} x_{n,y,s,p} + x_{n,y+1,s,p} \leq 1 \ , if\ s \leq 1 \\ x_{n,y,1,p} + x_{n,y,2,p}\ or\ x_{n,y,2,p} + x_{n,y+1,1,p} \ \leq 1 \ , else \end{cases} \quad (10)$$

From 2023, each plot must ensure leguminous crops are planted at least once every three years, with a minimum cumulative area of 1 acre over any three-year span.

$$\sum_{i=0}^{2} x_{n,y+i,s,p} \geq 1, p \in 1,2,3,4,5,17,18,19 \quad (11)$$

To ensure that the planting of each crop is not too dispersed, and based on a literature review that suggests a constraint, we limit the number of plots (including greenhouses) where a single crop can be planted in a season to no more than 5. Therefore, the constraint is as follows:

$$\sum_n x_{n,y,s,p} \leq 5 \quad (12)$$

To avoid excessive fragmentation in crop planting and ensure adequate rotation, we set a minimum planting area of 5% for each crop on any plot n, without exceeding the total area A_n.

$$x_{n,y,s,p} > 0.05 * A_n \quad (13)$$

$$\sum_p x_{n,y,s,p} \leq A_n \quad (14)$$

Incorporating all the aforementioned constraints on the decision variables, the model for Problem One is established.

2.2. Improved Genetic Algorithm

The improved genetic algorithm is an optimization algorithm inspired by natural selection, which solves optimization problems by simulating the genetic mechanisms of biological evolution. The structure of improved genetic algorithm is shown in Figure 2 and its core flowchart is shown in Figure 3.

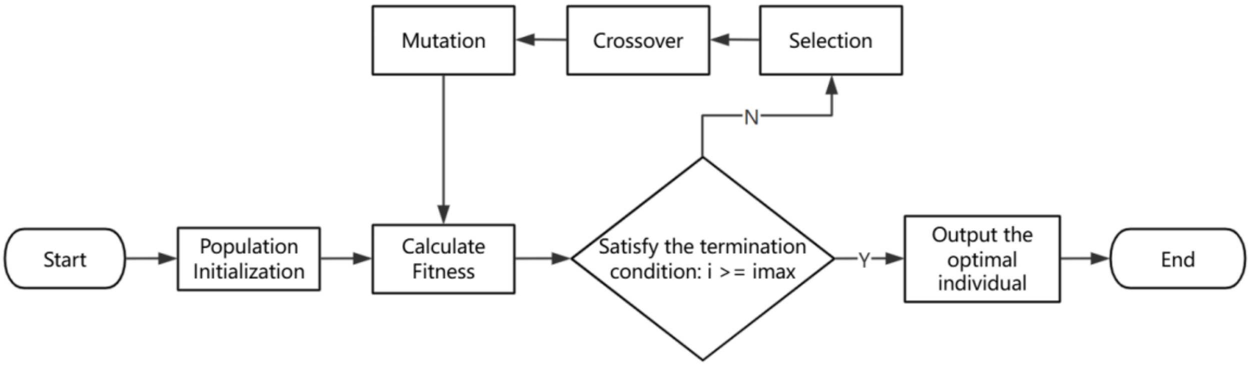

Figure 2. The structure of improved genetic algorithm

Due to the genetic algorithm's strong global search capabilities and excellent adaptability, we apply this model to solve linear programming problems [6]. Through extensive iterations, we ultimately obtain the optimal planting strategy. The key components of the genetic algorithm include encoding and decoding, population initialization, crossover operations, infeasible solution repair, mutation, and operator selection.

As indicated by the linear programming model, the decision variable $X_{n,y,s,p}$ encompasses multiple dimensions such as plot, year, season, and crop, which makes it difficult for traditional encoding methods to effectively represent these complex variables. To address this issue, a multi-matrix chromosome encoding is used to more intuitively map these decision variables.

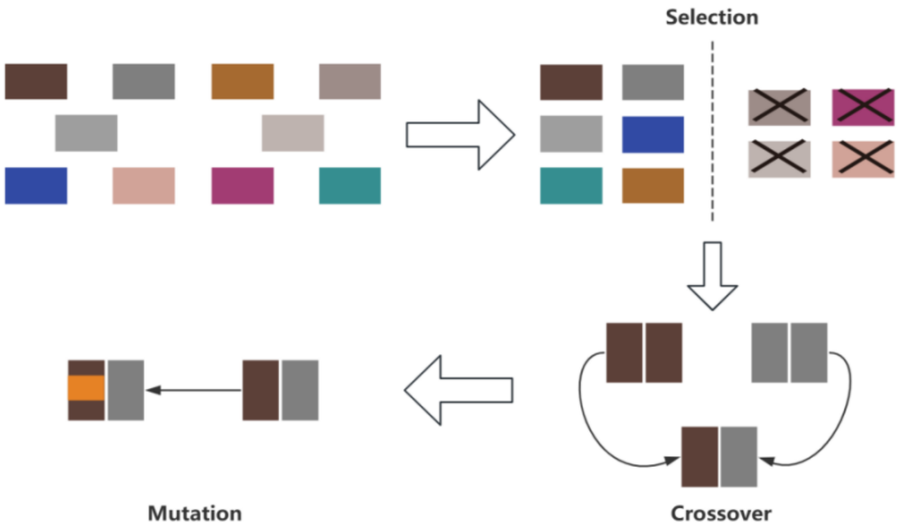

Figure 3. Improved genetic algorithm core flowchart

After randomly initializing the population, we employ an innovative chromosome crossover scheme to enhance solution diversity. Mutation operations are introduced to maintain genetic diversity within the population. Additionally, an infeasible solution repair step is incorporated to ensure the feasibility of the solutions. This approach effectively avoids local optima during the global search, gradually converging on the optimal solution to the problem. In this study, chromosome mutation is performed using a random perturbation approach. The mutation strategy flowchart is shown in Figure 4.

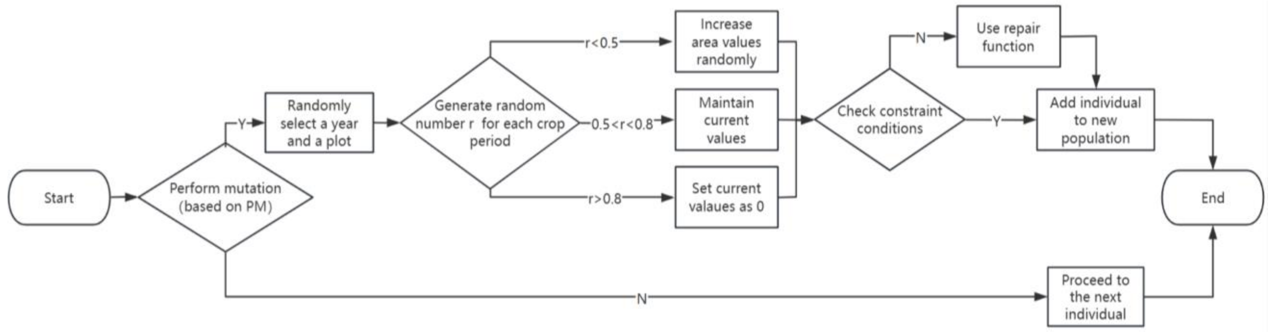

Figure 4. The mutation strategy flowchart

2.3. Bi-objective Robust Optimization Model

The bi-objective robust optimization model is a mathematical optimization model that considers two typically conflicting objectives: robustness and cost. This model aims to find a balance point that minimizes cost while ensuring a certain level of robustness, or maximizes robustness under a certain cost constraint [7].

Variations in crop sales volume, yield, costs, and prices prompt us to use orthogonal experimental design for multi-factor analysis. We set levels for each factor and conducted 11 orthogonal tests to form a compact experimental matrix.

To ensure representativeness, we calculate the average profit a for all orthogonal scenarios, using $income_y$ for total income and $expense_y$ for total expenses in year y, following the same formulas as previously mentioned. The average profit a is then given by:

$$max a = \frac{1}{y}\sum_{y=1}^{Y}\left(income_y - expense_y\right) \tag{15}$$

To find the best planting strategy, we maximize the minimum profit m among all orthogonal scenarios to ensure all profits are satisfactory. The formula for m is:

$$max m = min(income_i - cost_i), i \in 1,2,3, \dots, N \tag{16}$$

2.4. NSGA-II

NSGA-II is a highly efficient multi-objective genetic algorithm that addresses optimization challenges with competing goals. It improves population diversity using fast non-dominated sorting and crowding distance, applies an elitist strategy to preserve top individuals, and locates a well-distributed set of Pareto optimal solutions, all without manual parameter tuning [8].

We employ Pareto optimality to solve the aforementioned model. Assuming a and b are two feasible solutions in the decision space, if a dominates b ($a \prec b$), the following conditions must be met:

$$f_i(a) \le f_i(b), \forall i \in [1,2, \dots, n] \tag{17}$$

$$f_j(a) \le f_j(b), \exists j \in [1,2, \dots, n] \tag{18}$$

If no solution exists in the entire decision space that can dominate a, then a is referred to as a Pareto optimal solution. The set of Pareto optimal solutions is known as the Pareto front [9]. The Pareto solution for bi-objective optimization is shown in Figure 5.

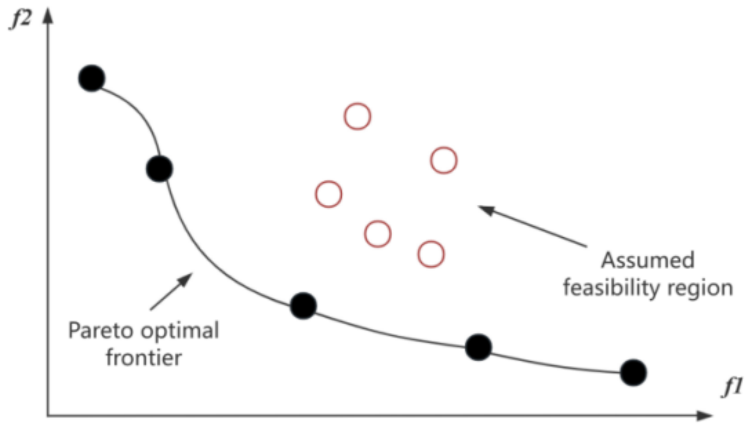

Figure 5. The Pareto solution for bi-objective optimization

However, in the solution process, if the dominance and subordination relationships are calculated for every individual, it would greatly increase the computational cost. Therefore, we use fast non-dominated sorting to decompose the solution set into Pareto fronts of different orders.

Incorporating fast non-dominated sorting into genetic algorithms led to the development of NSGA-II, which reduced the complexity of calculating non-dominance and expanded the sampling space [10]. However, beyond this, we also employ the crowding distance method to define and rank the selection priority among individuals at the same level. Employing the aforementioned algorithms enables the solution of multi-objective optimization problems.

3. Results

3.1. Model Establishment

3.1.1. The Establishment of a Linear Programming Model

Initially, a linear programming model was formulated to optimize the agricultural crop planning for the period spanning from 2024 to 2030. Subsequently, an enhanced genetic algorithm was employed to solve this model. This algorithm, which emulates the principles of natural selection and genetic inheritance, incorporates multi-matrix chromosome encoding, refined crossover and mutation strategies, as well as a mechanism for repairing infeasible solutions. These advancements have significantly enhanced the efficiency and quality of the solutions obtained.

3.1.2. The Establishment of a Bi-objective Robust Optimization Model

Focusing on the uncertainty brought by market environment changes in the field of crop planting, a bi-objective robust optimization model based on orthogonal experimental design was constructed. To effectively address the extraction of the Pareto front and the maintenance of solution diversity in multi-objective optimization, an improved Non-dominated Sorting Genetic Algorithm (NSGA-II) was employed. This algorithm simulates natural selection and genetic mechanisms, utilizing multi-matrix chromosome encoding, optimized crossover and mutation strategies, and infeasible solution repair mechanisms.

3.2. Results Analysis

3.2.1. Analysis of the Results for Experiment One

In response to the genetic algorithm we utilized, we proceeded to solve the aforementioned established planning model.

In Scenario 1, characterized by "excess inventory resulting in unsold goods and waste," the profit stabilizes at around 53,000,000, and the convergence curve of the model is depicted in Figure 6.

In Scenario Two, where the excess is sold at a 50% discount based on the 2023 sales price, the profit stabilizes at approximately 58,000,000, which is an increase of about 5,000,000 compared to Scenario One. The convergence curve of the model is depicted in Figure 6.

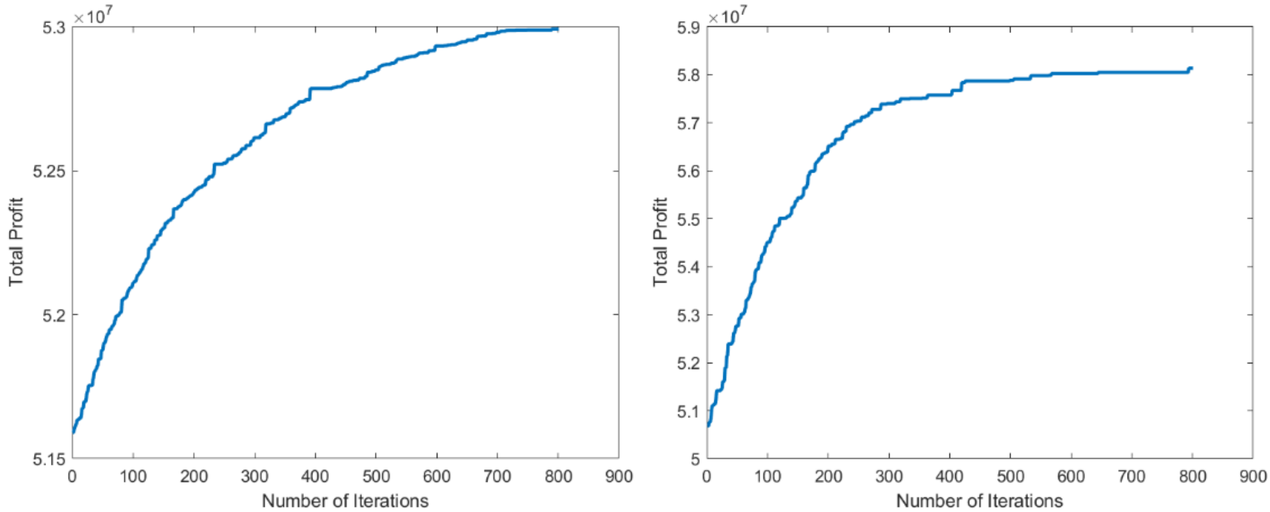

Figure 6. Total Profit Iteration Curves for Scenarios One and Two, 2024-2030.

The convergence curves show that as iterations increase, profit stabilizes with minimal fluctuations, indicating strong convergence. This suggests the algorithm has likely reached an optimal or near-optimal solution and maintains stability post-optimization. The results demonstrate the algorithm's effectiveness in finding optimal solutions and its reliability in solution stability.

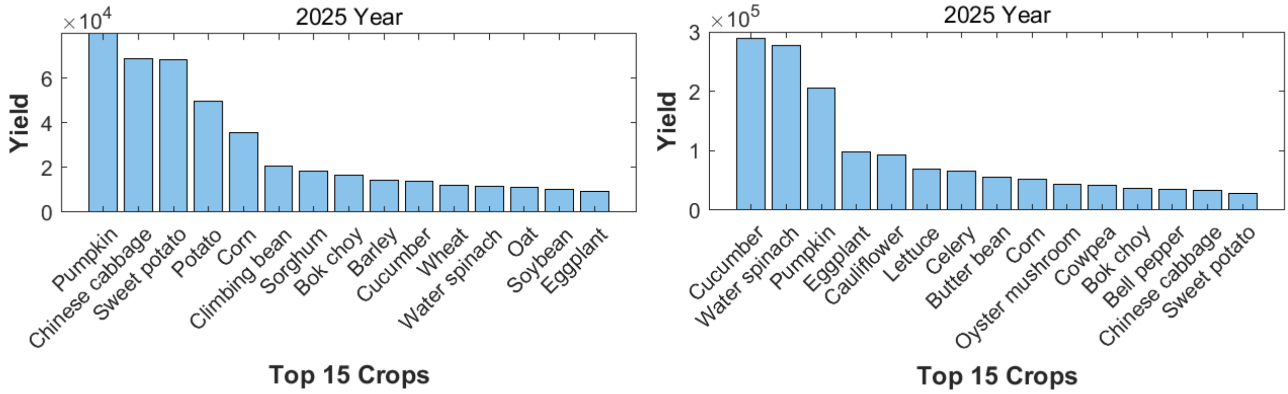

Figure 7. Crop Planting Yield for Scenarios One and Two in 2025

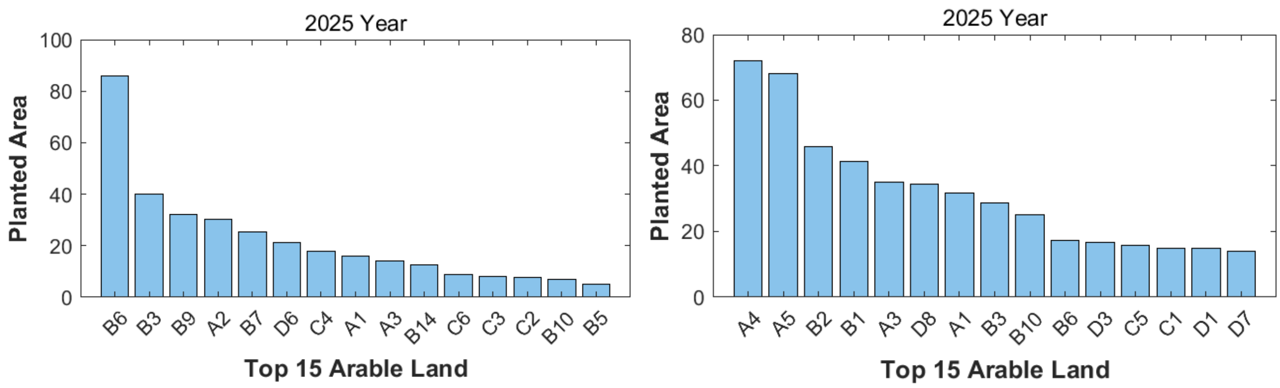

Figure 8. Cultivated Land Area for Crops in Scenarios One and Two in 2025

Due to space limitations, we have randomly selected the top fifteen crop planting yields and cultivated land areas for the year 2025 under the optimal planting strategy (i.e., the one with the highest total profit) for display. In Scenario One, where "excess leads to unsold goods and waste," Figures 7 and 8 illustrate the results. In Scenario Two, where "excess is sold at a 50% discount based on the 2023 sales price," Figures 7 and 8 show the outcomes.

Upon comprehensive analysis, it is observed that the annual production of various crops varies, with essentially all types of crops being cultivated. Regarding the cultivated land area, different types of arable land are fully utilized. The planting area neither exceeds the maximum area nor falls below 20% of the total land area, indicating a rational and efficient use, which is considered a reasonable outcome.

3.2.2. Analysis of the Results for Experiment Two

In Scenario One, where "excess leads to unsold goods and waste," the average profit from all feasible optimal planting schemes for the years 2024-2030 stabilizes at around 54,400,000, with the minimum profit stabilizing at approximately 54,250,000. The convergence curve is illustrated in Figure 9 below:

In Scenario Two, where "excess is sold at a 50% discount based on the 2023 sales price," the average profit from all feasible optimal planting schemes for the years 2024-2030 stabilizes at around 61,000,000, an increase of about 7,000,000 compared to Scenario One, with the minimum profit stabilizing at approximately 59,500,000, an increase of about 5,000,000 compared to Scenario One. The convergence curve is depicted in Figure 9 below:

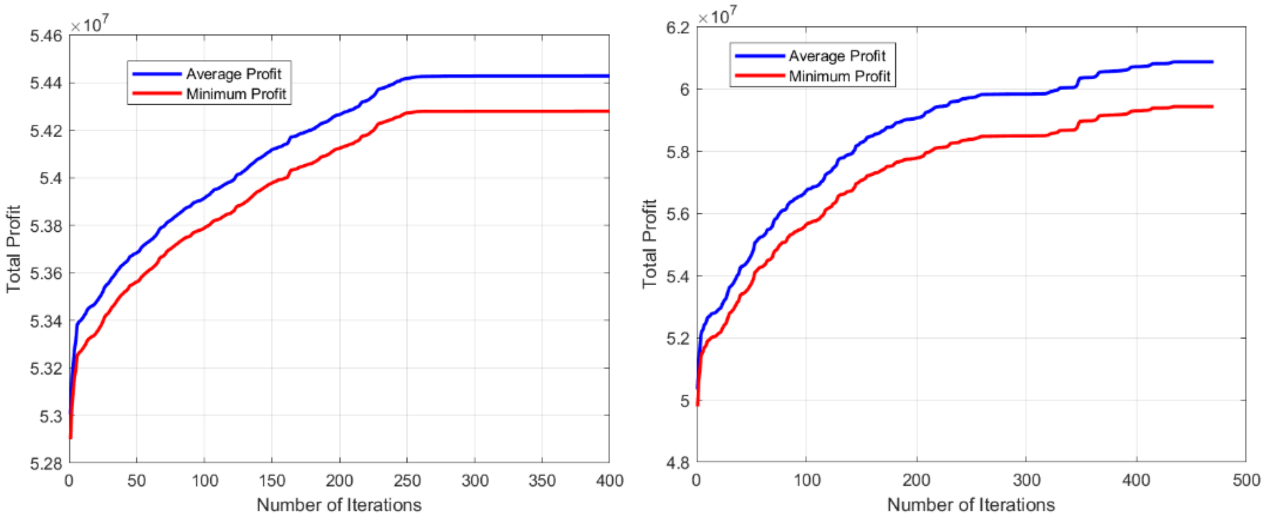

Figure 9. Average and Minimum Profits of All Orthogonal Plans

As can be seen from Figure 9, the curves overall show an upward trend, with profits continually increasing as the number of iterations grows, and gradually approaching stability. Eventually, at around the 250th and 390th iterations, respectively, the rate of profit increase began to flatten out, with profits essentially converging to an optimal solution.

Similar to the aforementioned context, we have randomly selected the top fifteen crop planting yields and cultivated land areas for the year 2029 under the optimal planting strategy (i.e., the one yielding the highest total profit) for display. Under Scenario One, where "excess leads to unsold goods and waste," Figures 10 and 11 illustrate the results. Under Scenario Two, where "excess is sold at a 50% discount based on the 2023 sales price," Figures 10 and 11 show the outcomes.

Rising vegetable crop prices have increased production, especially for potatoes, Chinese cabbage, and lettuce. Data shows diverse crop cultivation with full utilization of arable land, maintaining planting between 20% and 100% of the total area. The highest area for these crops is on Class D irrigated land, reflecting efficient and sustainable agricultural use.

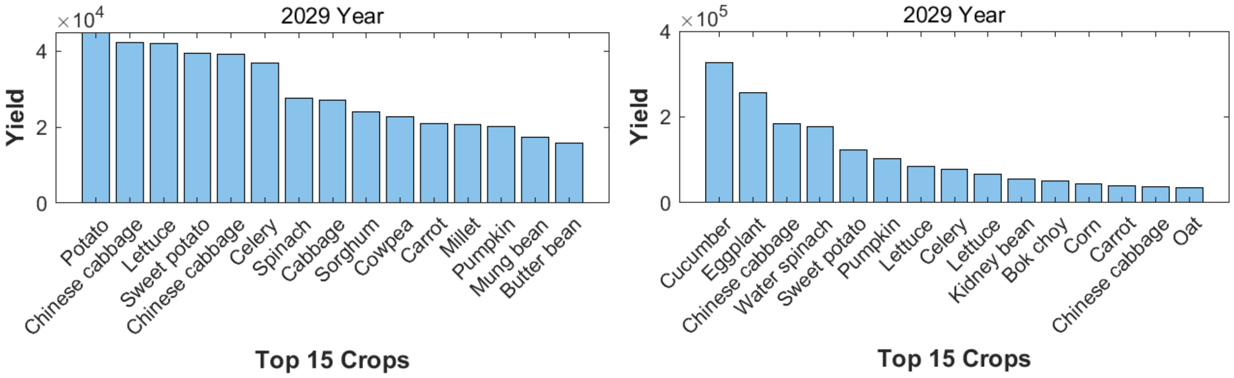

Figure 10. Crop Planting Yield for Scenarios One and Two in 2029

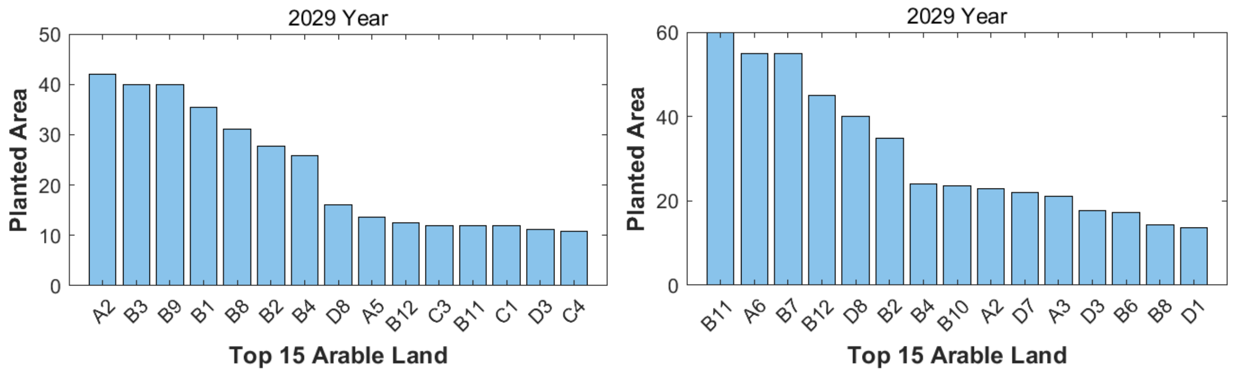

Figure 11. Cultivated Land Area for Crops in Scenarios One and Two in 2029

4. Conclusion

This study optimizes the crop planting strategy for a rural area in the North China mountain region, aiming to maximize planting profits from 2024 to 2030. The research takes into account factors such as land constraints, crop growth, market supply and demand, costs, and price fluctuations. It develops decision-making plans to cope with market uncertainty through linear programming and bi-objective robust optimization models. The study not only provides a long-term planting plan for the local area but is also applicable to other rural areas and can be extended to production scheduling and logistics path optimization. Although there are achievements, there is still room for improvement in the research. In the future, a more comprehensive model can be established to consider the relationship between crops and the impact of market demand. With technological progress, models and algorithms can be further optimized to support the sustainable development of rural economies and agricultural modernization.

References

[1] Sun L. Application of Dynamic Programming Model in Determining the Best Crop Planting Plan [J]. Jilin Agriculture, 2015, (08): 84.

[2] Yu H, Liu K, Bai Y, Luo Y, Wang T, Zhong J, Liu S, Bai Z. The agricultural planting structure adjustment based on water footprint and multi-objective optimization models in China [J]. Journal of Cleaner Production, 2021, 297: 126646.

[3] Randall M, Schiller K, Lewis A, et al. A Systematic Review of Crop Planning Optimisation Under Climate Change [J]. Agricultural Water Management, 2024, 10.1007/s11269 - 024 - 03758 - 3.

[4] Shi W. Research on Optimization of Crop Production Plan Based on Genetic Algorithm [D]. Northeast Agricultural University, 2005.

[5] Dubey C S. Biosecurity and biosafety challenges and strategies in relation to plant protection [J]. Indian Phytopathology, 2024, (prepublish): 1 - 12.

[6] Mirjalili S, Lewis A. The whale optimization algorithm [J]. Advances in Engineering Software, 2020, 95: 51 - 67.

[7] Fister Jr I, Fister D, Yang X S. A hybrid optimization algorithm for complex systems [J]. Mathematical Problems in Engineering, 2022: 1 - 11.

[8] Kesireddy A, Medrano F A. Elite multi-criteria decision making—Pareto front optimization in multi-objective optimization [J]. Algorithms, 2024, 17 (5): 206.

[9] Zheng X, Tang X, Wu Y, et al. An efficient CRISPR-Cas12a-mediated MicroRNA knockout strategy in plants. [J]. Plant biotechnology journal, 2024.

[10] Gao Y, Wang J. A multi-objective optimization model for sustainable agricultural land use planning under uncertainty [J]. Land Use Policy, 2021, 108: 105416.