Research on the Crop Planting Scheme in a Mountainous Village in North China

Xu Li, Jinlin Zhuo, Jiahao He, Zhenting Chen*
School of Artificial Intelligence, Guangzhou Huashang University, Guangzhou, Guangdong, China
**Corresponding Author.*

Abstract: This study focuses on the crop planting planning in a mountainous village in North China, aiming to optimize the planting strategies from 2024 to 2030. Utilizing the agricultural data of 2023, considering factors such as arable land resources, crop rotation requirements, sales situations, and various uncertainties, models are constructed through the Monte Carlo method, the greedy algorithm, and the genetic algorithm. In response to different sales situations (such as slow sales or sales at reduced prices) and changes in uncertain factors, the optimal planting schemes under various scenarios are solved to increase net profits and reduce risks. The results of the models show that reasonable planning can significantly improve yields, and the robustness of the models is verified through sensitivity analysis, providing strong support for rural planting decisions.

Keywords: Agricultural Planting Strategy; Monte Carlo Simulation; Genetic Algorithm; Greed Algorithm; Multi-Factor Optimization; Sensitivity Analysis

1. Introduction

1.1 Research Background

In today's world, climate change and resource shortages have become severe realities. How to efficiently utilize limited arable land resources to achieve sustainable crop planting is of great significance for ensuring food security and promoting rural economic growth. A village in the mountainous area of North China serves as a typical case, and the optimization of its crop planting strategies is crucial for the agricultural development of the local area and similar regions. [1]

1.2 Literature Review

Previous studies have achieved fruitful results in the field of agricultural planting strategies. Zhou Min focused on the adjustment of the rural rice planting structure and industrial upgrading, emphasizing the importance of conforming to market demands, providing an industrial development perspective for this paper. Zhu Lulu explored the Monte Carlo method, providing an effective way to handle uncertainty problems. Fang Wei applied Monte Carlo simulation to project risk management, inspiring ideas for the risk response strategies in this paper. Chang Youqu et al. discussed the greedy algorithm, elaborating on its principles and applications in optimization problems. Ge Jike et al. reviewed the genetic algorithm, laying a theoretical foundation for the algorithm selection in this paper. However, there is still room for improvement in existing studies in formulating long-term and comprehensive planting strategies by comprehensively considering various complex factors. This paper aims to fill this research gap.

2. Research and Analysis

2.1 Basic Situation

This study aims to formulate a crop planting plan for a village in the mountainous area of North China from 2024 to 2030, fully considering the limitations of arable land resources, crop rotation requirements (some crops cannot be continuously planted, and leguminous crops should be planted at least once every three years), as well as the convenience of planting operations and the efficiency of field management.

2.2 Specific Problems

2.2.1 Problem one

If the crop yield exceeds the expected sales volume, the surplus part cannot be sold, resulting in a waste of resources.

If the yield exceeds the expected sales volume, the excess part will be sold at 50% of the 2023 price.

2.2.2 Situation two

Considering the future changes in the sales volume, yield per unit area, planting cost, and selling price of crops, formulate a planting strategy that adapts to these uncertain factors. Among them, the annual growth rate of the sales volume of wheat and corn is 5% - 10%, the sales volume of other crops fluctuates by ±5%; the yield per unit area changes by ±10% annually; the planting cost increases by 5% annually; the selling price of grain crops is relatively stable, the selling price of vegetable crops increases by 5% annually, and the selling price of edible fungi, especially morel mushrooms, decreases by 5% annually.

2.2.3 Problem three

Based on Problem Two, further explore the substitution or complementary relationships between different crops as well as the interrelationships between sales volume, price, and cost, formulate a more comprehensive planting strategy, and conduct a comparative analysis with the strategy of Problem Two to determine the optimal scheme.

3. Model Establishment

3.1 Symbol Explanation

Table 1. Symbol Explanation

Symbol	Definition
x_{ij}	The planting area of crop i on plot j
Q_i	The unit selling price of crop i (yuan/jin)
P_{ij}	The yield per unit area of crop i on plot j
Y_{ij}	The loss due to slow sales, calculated by multiplying the part exceeding the expected sales volume by the unit price
Z_{ij}	The planting area of crop i on plot j
δ_{ij}	The planting cost of crop i on plot j
E	Net income
SV_i	The sales volume of crop i
$S_{i,j,t}$	The area of crop

See Table 1 above, clearly define the meanings of each symbol, such as the planting area of crops on a certain plot, the unit selling price, the yield per unit area, and the loss due to slow sales, etc., to provide clear variable definitions for subsequent model construction.

3.2 Model Assumptions

Assume that the rural climate conditions remain relatively stable from 2024 to 2030.

Expect that the market demand for agricultural products remains relatively stable without drastic fluctuations.

Assume that the policies of the state and local governments regarding agricultural planting, land use, and agricultural product sales are continuous and stable.

Assume that all the plots involved in the study are fully utilized, and there is no situation of idleness or underutilization.

4. Problem Analysis

4.1 Analysis of Problem One

This problem focuses on constructing a model aimed at maximizing net profits under the premise of considering the loss of some agricultural products due to overstocking. We have preliminarily processed the agricultural planting and related statistical data of 2023, which involves the collection, integration, and standardization of the planting area, yield, selling price, and cost of various plots, in order to provide necessary data support for model construction. There are two situations in this problem: In the first situation, given that excessive production will lead to overstocking and waste of products, we set decision variables, such as the planting area and yield of each crop, and aim to maximize net profits. The constraints of the model involve the types of crops on the plots, rotation regulations, and the limits of estimated sales volume. We use linear programming techniques to solve the model, including optimizing the planting strategy according to the expected yield and sales situation of seasonal crops to achieve an increase in net profits. The output part of the model shows the convergence process and proposes the optimal planting strategy to ensure the matching of yield and market demand. In the second situation, for the excess yield sold at half price, we adjusted the objective function on the basis of the original model to reflect the impact of sales at reduced prices.

4.2 Analysis of Problem Two

Problem Two aims to create an optimization model for crop planting planning that takes into account numerous uncertain factors. The model aims to rationally arrange crop planting to improve the economic benefits of rural areas while minimizing the risks brought by market and production fluctuations. Decision variables include the planting area, yield, and cost of each crop on different plots. The model introduces various random factors to simulate the possible fluctuations in the sales volume, yield per unit area, planting cost, and selling price of crops from 2024 to 2030. These uncertain factors are characterized by setting reasonable change intervals, such as the expected annual growth rate of the sales volume of wheat and corn is 5% - 10%, the change in the sales volume of other crops is within ±5%, the annual fluctuation of the yield per unit area is ±10%, and the annual increase in the planting cost is 5%, etc. We use the planning and solving algorithm, combined with the comprehensive consideration of crop income, cost, and market fluctuations, and use optimization techniques such as linear programming or nonlinear programming to solve the problem. The solving steps include determining the objective function, handling constraint conditions, and performing optimization calculations.

4.3 Analysis of Problem Three

The discussion of Problem Three focuses on the role of factors such as the interrelationships between crops in optimizing the planting strategy. First, we use the Pearson correlation coefficient to analyze the interactions between the planting cost, unit selling price, and sales volume of crops, thereby quantifying the connections between these variables and providing a data basis for the construction of subsequent optimization models. Based on these analysis results, we establish a dynamic multi-factor planting planning optimization model that takes into account the substitutability and complementary nature of crops. For example, some crops may have a substitution relationship, or some combined plantings may increase the overall income.

4.4 Model Solution

4.4.1 Data preprocessing

Utilizing the 2023 annual crop planting information and the statistical data of the same year provided in Annex 2, with the help of the merge function in Python, a left outer join operation is performed according to the crop number to merge these two datasets into a tabular dataset named dataset1 as shown below.

0	A1	6	Wheat	Grain	80.0	Single Season	Flat Dry Land	800	450	3.00-4.00	64000
1	A1	6	Wheat	Grain	80.0	Single Season	Terraced Field	760	450	3.00-4.00	60800
2	A1	6	Wheat	Grain	80.0	Single Season	Slope Land	720	450	3.00-4.00	57600
3	A2	7	Corn	Grain	55.0	Single Season	Flat Dry Land	1000	500	2.50-3.50	55000
4	A2	7	Corn	Grain	55.0	Single Season	Terraced Field	950	500	2.50-3.50	52250
...	...	...	...	...	...	...	...	...	...	...	...
208	F4	34	Celery	Vegetable	0.3	Second Season	Ordinary Greenhouse	6600	1100	3.20-4.80	1980
209	F4	34	Celery	Vegetable	0.3	Second Season	Intelligent Greenhouse	6000	1200	3.80-5.80	1800
210	F4	23	Spinach	Vegetable	0.3	Second Season	Irrigated Land	2700	2300	4.80-6.70	810
211	F4	23	Spinach	Vegetable	0.3	Second Season	Ordinary Greenhouse	3300	2700	4.80-6.70	990
212	F4	23	Spinach	Vegetable	0.3	Second Season	Intelligent Greenhouse	3000	3000	5.80-8.00	900

Figure 1. Dataset Dataset1 (Partial Display)

See Figure 1 above, processing the data in the "Unit Selling Price/(yuan/jin)" column in dataset1: First, divide the price range in this column into two limits. Then, write a function to randomly generate the specific selling price for each row, and this function will take values within the specified price range. Apply this function to the "Unit Selling Price/(yuan/jin)" column to obtain the specific selling price for each row. The following is the formula description for calculating the selling price: Let the lower limit of the price range be O_{min} and the upper limit be O_{max}. For each row of the data table, the formula for calculating the selling price is:

$$O = O_{min} + Random \times (O_{max} - O_{min}) \quad (1)$$

where $Random$ is a random number drawn from a uniform distribution, with a value between 0 and 1. After completing this step, aggregate the data according to the crop number to further calculate the total sales volume (i.e., total yield) of 2023 and calculate the average selling price. The formula for calculating the total sales volume is:

$$Total\,SV_i = \sum_i SV_i \quad (2)$$

where SV_i represents the sales volume of the i-th crop number.

The formula for calculating the average selling price is:

$$\mathrm{AVG\ SP_i} = \frac{\sum_i (SP_i \times SV_i)}{\sum_i SV_i} \quad (3)$$

Here, SP_i refers to the selling price of the i-th crop number, and SV_i is the corresponding sales volume.

Table 2. Sales Volume and Price of Each Crop After Polymerization

Crop Name	Unit Selling Price/(yuan/jin)	Sales Volume/jin
Rice	6.201604	21000
Lettuce	4.35216	4500
Climbing Bean	7.477958	9875
Corn	3.26384	132750
Lettuce	6.506156	1350
White Lingzhi Mushroom	16.95986	18000
White Radish	2.797871	100000

4.4.2 Crop-Suitable planting plot categories and seasons

See Table 2 above, the instructions on the rural crop planting table in Annex 1 led us to organize the following table, named dataset3. At the same time, we sorted out the existing cultivated land in multiple villages, removed the remarks, and obtained dataset4.

4.4.3 Solution to problem one

4.4.3.1 First situation

Model Establishment: In this model, our goal is to optimize the crop planting strategy to improve the overall net profit of the village. Decision variables include the types of crops planted in different seasons on each plot and their areas. Our aim is to calculate the sales income and planting cost of crops to maximize the overall income, taking into account the potential overstocking losses. Specifically, we set the objective function to maximize the total net profit, which is the crop sales income minus the planting cost, minus the overstocking loss of the part exceeding the market expected sales volume. The calculation of the overstocking loss is based on the selling price of the excess part of the crops.

a. Variable Definition: Decision variable x_{ij} : Represents the planting area of crop i on plot j.

b. Objective Function: The goal is to maximize the net profit, considering factors such as yield, selling price, and overstocking losses. Under this assumption, the overstocked crops have no income. The objective function is expressed as follows:

$$E_{max} = \sum_{i,j} (Q_i \cdot P_{i,j} \cdot Z_{i,j} - \delta_{i,j} \cdot Z_{i,j} - Y_{i,j}) \quad (4)$$

c. Constraint Conditions: Plot Area Limitation: For any plot, the total area of crops planted on it shall not exceed the total area of the plot.

$$\sum_i Z_{i,j} \le S_j \quad \forall i,j \quad (5)$$

Let S_j represent the total area of plot j .

Special Limitation for Irrigated Land: If rice is planted on irrigated land in the first season, no other crops shall be planted in the second season.

$$\begin{gathered} if\ Z_{rice,j} > 0, then \\ Z_{ij} = 0 \\ \forall i\ in\ the\ second\ season, \\ j\ is\ watering\ the\ field \end{gathered} \quad (6)$$

Minimum Planting Area Limitation: The planting area of each plot shall meet the minimum area requirement. Assume that the minimum area is 20% of the total area of the plot.

$$Z_{i,j} \ge 0.2 \times S_j \quad \forall i,j \quad (7)$$

Crop Rotation Rule Constraint: Each crop shall not be continuously planted on the same plot, and each plot shall be planted with leguminous crops at least once within two years.

$$\sum_{t=1}^{3} A_{legume,j,t} \ge 0\ \forall j \quad (8)$$

The Monte Carlo method is extremely practical in model solution. [2] It combines determinism and randomness to comprehensively simulate various schemes to deal with uncertainties. The execution steps include randomly generating planting strategies, calculating net profits, and repeating the simulation to select the optimal strategy.[3]

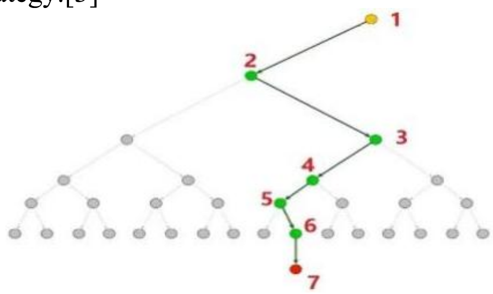

Figure 2. The Application Process of the Monte Carlo Algorithm

See Figure 2 above, as shown in the application process of the Monte Carlo algorithm, each iteration involves the random generation of the planting area, which in turn affects the calculation of the net profit. The specific solution steps are detailed as follows:

4.4.3.2 Initialization stage

Input the plot type, crop attributes, and planting rules, and set the total number of simulations N.

4.4.3.3 Iterative random simulation

In each simulation, based on the adaptability of crops, rotation rules, and plot area, randomly generate a planting strategy, that is, randomly assign a crop combination Z to each plot. For each plot, calculate the corresponding yield P and planting cost δ, and estimate the loss due to slow sales in combination with market demand limitations.

4.4.3.4 Objective function calculation

For each simulation result, calculate the net profit. The formula for calculating the net profit is: Net profit E = (Q × P × Z - δ × Z - Y).

$$E = \sum_{i,j} (Q_i \cdot P_{i,j} \cdot Z_{i,j} - \delta_{i,j} \cdot Z_{i,j} - Y_{i,j}) \quad (9)$$

The calculation method of the loss due to slow sales is: L = max (0, Z - D) × Q,where D is the expected sales volume of crop i.

$$Y_{i,j} = \max(0, P_{i,j} \cdot Z_{i,j} - D_i) \cdot Q_i \quad (10)$$

4.4.3.5 Result screening

Execute N random simulations, record the net profit E of each simulation, and select the planting strategy with the highest net profit.

4.4.3.6 Output optimal scheme

After random simulations, select the scheme with the highest net profit as the optimal planting scheme. By combining the Monte Carlo algorithm, we introduce randomness into the objective function, enhancing the ability to optimize the planting strategy in an uncertain environment and effectively assisting decision-making.

4.5 Model Iteration

Through Monte Carlo method simulation random sampling, optimize the planting strategy to ensure that the crop combination meets the requirements and improve economic benefits. After a large number of samplings, select the scheme with the highest net profit as the optimal solution to achieve the best economic benefits from 2024 to 2030. Multi-season planting schemes have been formulated and optimized through 200 iterations. Now, the iteration process chart will be displayed.

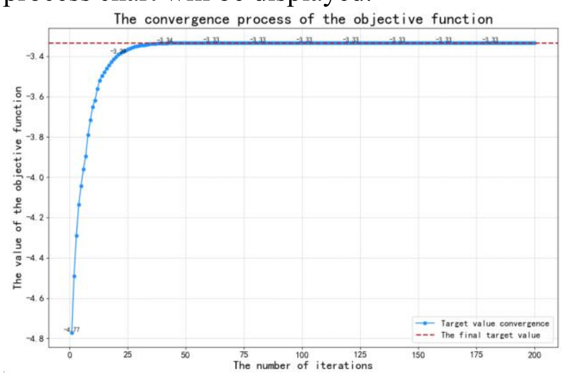

Figure 3. The Convergence Process of the Objective Function in Situation 1 of Problem 1 See Figure 3 above, in the iteration chart, the vertical axis represents the value of the objective function, and the horizontal axis represents the number of iterations. The chart clearly shows that the algorithm has become stable at the 25th iteration, and the subsequent changes are minimal, indicating that the algorithm has almost found the optimal solution or a very close solution.

2024 to 2030 Annual Optimal Crop Planting Schemes:

After optimizing the crop planting strategies from 2024 to 2030, we obtained a series of annual income results. The following is a summary of the target values for each year:

Table 3. Results of Crop Sector Earnings from 2024 to 2030

Results of Crop Yields from 2024 to 2030			
Year	**Target Yield**	**Additional Yield**	**Total Yield**
2024	5851076.72	3301525.27	9152601.99
2025	6232632.95	3114653.31	9347286.26
2026	5798432.11	3049959.76	8848391.86
2027	6405154.09	2992719.31	9397873.40
2028	5433856.36	3194906.03	8628762.39
2029	5793714.81	3133847.05	8927561.86
2030	5660099.07	3115208.65	8775307.72

See Table 3 above, it can be seen from these data that the total yield in 2027 was the highest, while that in 2028 was relatively low. Overall, the yields showed a trend of significant fluctuations every other year. Such fluctuations may be related to adjustments in planting strategies, fluctuations in market prices, and how to handle slow-moving products. These results provide us with important information for further optimizing planting strategies in the coming years to achieve more stable yields.

4.5.1 Second scenario

4.5.1.1 Model establishment

Optimize the crop planting strategies in a certain village from 2024 to 2030, aiming to increase the net profit of crop planting and properly handle the problem of slow-moving crops that exceed market demand. In the set scenario, that is, "the crops in excess will be sold at 50% of the 2023 selling price", the model not only takes into account the planting costs and yields but also includes the calculation of the revenue from selling the slow-moving part at a reduced price.

a. Variable Definition: Decision variable $X_{i,j}$: Represents the planting area of crop i on plot j in year t.

b. Objective Function: The objective is to maximize the net profit of the crops while taking into account the additional income from the slow-moving part. The net profit is calculated as the total revenue of the crops minus the planting cost, plus the income from selling the slow-moving crops at 50% of the price.

The specific formula is as follows:

$$E_{i,j} = \sum_{i,j} (Q_{i,t} \cdot P_{i,j,t} \cdot Z_{i,j,t} - \delta_{i,j,t} \cdot Z_{i,j,t} + \max(0, (P_{i,j,t} \cdot Z_{i,j,t} - S_{i,t}) \cdot 0.5 \cdot Q_{i,t} \tag{11}$$

The objective function calculates the net profit of the crops planted on each plot and calculates the income from slow-moving crops that exceed the sales volume at 50% of the price. During the optimization process of this model, both the regular sales revenue of the crops and the secondary sales revenue from the slow-moving part are considered. The reduced price income of the slow-moving crops is calculated using the max function to avoid negative values. The last term in the formula represents the income from the slow-moving part, and its price is 50% of the market price.

c. Constraint Conditions: The same as the constraint conditions in Figure 3.c Situation 1.

4.5.1.2 Model solution

In the implementation process of the Monte Carlo algorithm, the core of each iteration is to randomize the planting area, which in turn affects the calculation of the net profit. The specific solution steps are as follows:

Initialization Stage: Input the plot type, crop attributes, and planting rules. Set the total number of simulations N.

4.5.1.3 Iterative random simulation

In each simulation, based on the adaptability of the crops, rotation rules, and plot area, randomly generate a planting strategy, that is, randomly assign a crop combination Z to each plot. For each plot, calculate the corresponding yield P and planting cost δ, and estimate the income due to slow-moving in combination with market demand limitations.

4.5.1.4 Objective function calculation

For each simulation result, calculate the net profit. The formula for calculating the net profit is: Net profit E = ($Q \times P \times Z - \delta \times Z + Y$).

$$E = \sum_{i,j} (Q_i \cdot P_{i,j} \cdot Z_{i,j} - \delta_{i,j} \cdot Z_{i,j} + Y_{i,j} \times 0.5) \tag{12}$$

The calculation method of the slow-moving income is: $L = \max(0, Z - D) \times Q \times 0.5$, where D is the expected sales volume of crop i.

$$Y_{i,j} = \max(0, P_{i,j} \cdot Z_{i,j} - D_i) \cdot Q_i \times 0.5 \tag{13}$$

4.5.1.5 Result screening

Execute N random simulations, record the net profit E of each simulation, and select the planting strategy with the highest net profit.

4.5.1.6 Output optimal scheme

After numerous random simulations, select the scheme with the highest net profit E as the optimal crop planting scheme.

By integrating the Monte Carlo algorithm, we incorporate randomness simulation into the original objective function, thereby enhancing the ability to optimize the planting strategy in a complex and uncertain environment. This method can effectively assist decision-makers in selecting the crop planting scheme with the largest revenue among many potential situations.

4.5.2 Result display and output

4.5.2.1 Model iteration:

Adopt the Monte Carlo method for simulation sampling to optimize the planting strategy to increase yields and meet land requirements. After screening, select the scheme with the highest net profit to ensure the maximization of economic benefits from 2024 to 2030. Multi-season planting schemes have been formulated and iteratively optimized, and the iteration process chart of 200 iterations will be displayed.

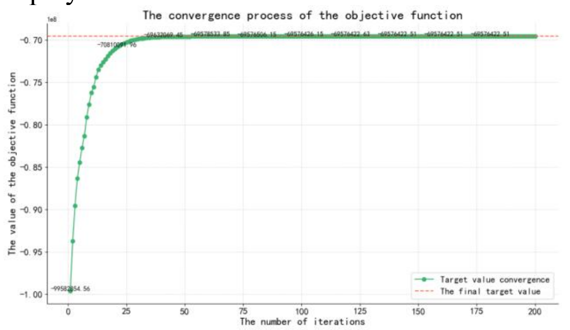

Figure 4. Convergence Process of the Objective Function in Situation 2 of Problem1

See Figure 4 above, in the iteration chart, the vertical axis represents the value of the objective function, and the horizontal axis represents the number of iterations. The chart clearly shows that the algorithm has become stable at the 25th iteration, and the subsequent changes are minimal, indicating that the algorithm has almost found the optimal solution or a very close solution. This stable convergence process fully demonstrates the stability and effectiveness of the Monte Carlo algorithm in problem optimization.

4.5.2.2. 2024 to 2030 Annual optimal crop planting schemes

After optimizing the crop planting strategies from 2024 to 2030, we obtained a series of annual income results. The following is a summary of the target values for each year.

Table 4. Results of Crop Sector Earnings from 2024 to 2030

Results of crop sector earnings from 2024 to 2030			
Year	**Target Income**	**Additional Income**	**Total Income**
2024	6143630.55	3275113.06	9418743.61
2025	6438309.83	3118079.42	9556389.25
2026	5821625.83	3254307.06	9075932.89
2027	7039264.34	3288998.52	10328262.86
2028	5792490.88	3300337.92	9092828.80
2029	5851651.95	3165185.52	9016837.47
2030	6225542.97	3423614.31	9649157.28

See Table 4 above, it can be seen from these data that the total income in 2027 was the highest, while that in 2029 was relatively low. Overall, the incomes showed a trend of significant fluctuations every other year. Such fluctuations may be related to adjustments in planting strategies, fluctuations in market prices, and how to handle slow-moving products. These results provide us with important information for further optimizing planting strategies in the future years to achieve more stable incomes.

4.6 Problem 2: Model Establishment and Solution

4.6.1 Multi-factor model establishment

The goal of this subsection is to optimize the crop planting strategies in a certain village from 2024 to 2030 to adapt to the future trends and uncertainties of the expected sales volume, yield per unit area, planting cost, and selling price of crops. The strategy takes into account changes in various factors, such as the expected increase in the sales volume of wheat and corn, the rise in the price of vegetables, the decline in the price of edible fungi, and the increase in the planting cost. To capture these uncertain factors, the model constructs a dynamic optimization scheme by randomly generating the main parameters of different crops each year, including the sales volume, yield per unit area, planting cost, and selling price. The model uses linear programming to find the optimal solution and simulates different risk scenarios through random variables.

a. Model Objective: The objective of the model is to adjust the proportion of the crop planting area to maximize the net profit of the crops while effectively dealing with the problem of crop surplus that exceeds market expectations. In addition, the model ensures that the crop combinations on different plots conform to their respective plot characteristics, seasonal planting restrictions, and rotation regulations.

b. Variable Definition: Decision variable $X_{i,j}$: Represents the planting area of crop i on plot j in year t. Each combination of plot and crop defines a decision variable, representing different planting schemes.

c. Objective Function: The objective is to minimize the net profit while taking into account the normal sales of crops and the reduced-price sales of the part that exceeds the expected sales volume. The formula for calculating the net profit is as follows:

$$E_{i,j} = \sum_{i,j} (Q_{i,t} \cdot P_{i,j,t} \cdot Z_{i,j,t} - \delta_{i,j,t} \cdot Z_{i,j,t} + \max(0, (P_{i,j,t} \cdot Z_{i,j,t} - S_{i,t}) \cdot 0.5 \cdot Q_{i,t} \tag{14}$$

If the planting area exceeds the expected sales volume, the last term in the formula represents the income from the slow-moving part, and its price is 50% of the market price.

d. Constraint Conditions: To ensure the feasibility of the planting scheme, the following constraint conditions are set in the model;

Plot Area Limitation: For any plot, the total area of crops planted on it shall not exceed the total area of the plot.

$$\sum_i Z_{i,j} \le S_j \quad \forall i,j \tag{15}$$

Let S_j represent the total area of plot j.

Special Limitation for Irrigated Land: If rice is planted on irrigated land in the first season, no other crops shall be planted in the second season.

$$\begin{gathered} if\ Z_{rice,j} > 0, then\ Z_{ij} = 0 \\ \forall i\ in\ the\ second\ season, \\ j\ is\ watering\ the\ field \end{gathered} \tag{16}$$

Minimum Planting Area Limitation: The planting area of each plot shall meet the minimum area requirement. Assume that the minimum area is 20% of the total area of the plot.

$$Z_{i,j} \ge 0.2 \times S_j \quad \forall i,j \tag{17}$$

Crop Rotation Rule Constraint: Each crop shall not be continuously planted on the same plot, and each plot shall be planted with leguminous crops at least once within two years.

$$\sum_{t=1}^{3} A_{legume,j,t} \ge 0\ \forall j \tag{18}$$

e. Introduction of Randomness Factors: In this model, we consider the annual random changes in the sales volume, yield per unit area, planting cost, and selling price of crops. The specific rules are as follows:

4.6.2 Expected sales volume

For grain crops such as wheat and corn, it is assumed that the expected annual sales volume has a growth rate of 5% - 10%. The formula for updating the annual sales volume is:

$$D_{i,t+1} = D_{i,t} \times (1 + r) \tag{19}$$

where r is the annual growth rate, randomly selected from the interval of 5% - 10%. For other crops, the annual sales volume will have a fluctuation of ±5% relative to 2023. The formula for calculating is:

$$D_{i,t+1} = D_{i,t} \times (1 \pm 0.05) \tag{20}$$

Randomly select a value within the range of ±5% fluctuations.

Planting Cost:

The planting cost of all crops increases by 5% annually. The formula for calculating is:

$$\delta_{i,j,t+1} = \delta_{i,j,t} \times (1 + 0.05) \tag{21}$$

Selling Price:

The price of grain crops is basically stable:

$$Q_{i,t+1} = Q_{i,t} \tag{22}$$

The price of vegetable crops increases by 5% annually. The formula for calculating is:

$$Q_{i,t+1} = Q_{i,t} \times (1 + 0.05) \tag{23}$$

The price of edible fungi crops decreases randomly by 1% - 5% annually. The formula for calculating is:

$$Q_{i,t+1} = Q_{i,t} \times (1 - r) \tag{24}$$

where r is randomly selected from the interval of 1% - 5%. For morel mushrooms, the price decreases by 5% annually.

Yield per Unit Area:

The yield per unit area of crops is affected by external factors such as climate. The annual yield per unit area will have a fluctuation of ±10%. The formula for calculating is:

$$P_{i,j,t+1} = P_{i,j,t} \times (1 \pm 0.10) \tag{25}$$

This fluctuation range can be dynamically generated by a random function.

4.6.3 Greedy algorithm solution

The greedy algorithm is a strategy of selecting the currently seemingly optimal solution in stages to gradually construct a global solution. [4]. This method selects the locally optimal solution in each step, hoping that through such local optimal selections, a solution close to the global optimal can finally be pieced together.

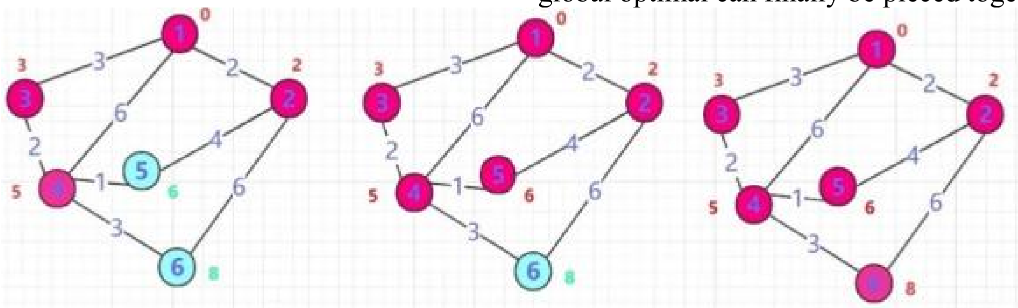

Figure 5. Example Diagram of the Greedy Algorithm

See Figure 5 above, using the greedy strategy, we can gradually improve the crop planting strategy to maximize the net profit of each plot of land. In the implementation process of the greedy algorithm, in each step, we will select the crop with the highest income under the existing conditions and arrange it on the plot suitable for planting this crop. This cycle continues until all plots have completed crop allocation. The model solution process of the greedy algorithm is described as follows:

a. Initialization: Import data related to crops, including information such as the annual expected sales volume, yield per unit area, planting cost, and selling price. Set each plot of land to an unplanted state, and compile a list of crops suitable for planting on each plot according to the land type and season characteristics.

b. Select the Optimal Crop: For each plot of land, select crops according to the following greedy strategy: Calculate the net profit of all crops suitable for planting on each plot of land:

$$E = Q_{i,j} \times P_{i,j,t} - \delta_{i,j,t} \tag{26}$$

If the crop yield exceeds the expected sales volume, further calculate the reduced-price income of the excess part:

$$Y = (Z_{i,j,t} \times P_{i,j,t} - D_{i,t}) \tag{27}$$

Combine the net profit of regular sales and the reduced-price sales part to calculate the total income:

$$E = E + \max(0, Y) \tag{28}$$

Select the crop with the highest total income: For each plot of land, select the crop with the highest total income as the planting option for this plot.

c. Allocate the Planting Area: Allocate crops according to the actual area of the plot to ensure that the total area allocated to crops on each plot does not exceed the capacity limit of the plot:

$$Z_{i,j,t} \leq S_j \ \forall j \tag{29}$$

When the expected sales volume of the selected crop exceeds, priority should be given to allocating more area to the crop with higher income. If a single crop cannot fully utilize the entire plot, continue to select the sub-optimal crop to fill the remaining space.

d. Consider Rotation and Seasonal Limitations:

When allocating crops, it is necessary to follow the conditions such as the type of plot, season changes, and the suitability of crops to ensure compliance with the rotation rules and seasonal planting requirements. Especially for paddy fields, if rice has been planted in the first season, other crops should be avoided in the second season:

$$Z_{rice,j,t} > 0 \ \Rightarrow Z_{i,j,t} = 0 \ \forall i, Season\, 2 \tag{30}$$

At the same time, it is necessary to comply with the constraint that leguminous crops should be rotated at least once every three years:

$$\sum_{t=1}^{3} Z_{rice\ldots,j,t} \geq 0 \ \forall j \tag{31}$$

e. Iterative Optimization Process: From 2024 to 2030, select the most suitable crop for each plot every year, following the seasonal and rotation requirements. The greedy algorithm selects the crop with the highest income each year, adjusts the plot allocation to maximize the total income. Although it may not be globally optimal, it is highly efficient for optimizing large-scale and complex planting combinations and effectively improves the net profit.

4.6.4 Result display and output

Model iteration convergence: after a series of iterative optimization processes, the greedy algorithm improves the crop planting strategy through iterative optimization, selects the crop combination with the highest income, and ensures compliance with the plot characteristics and seasonal conditions. After screening, the optimal solution is determined to ensure the optimization of the planting strategy from 2024 to 2030. Taking a certain village as an example, a multi-season planting maximum income scheme with 200 iterations has been customized. The following is an example of the iteration chart.

The iteration chart in Figure 6 shows that the vertical axis is the value of the objective function, and the horizontal axis is the number of iterations. It can be seen from the chart that the algorithm becomes stable after the 25th iteration, and the value of the objective function changes little, indicating that the algorithm converges and may find the optimal solution. The results confirm that the greedy algorithm is efficient and effective, and the iteration curve converges smoothly, showing the robustness of the algorithm.

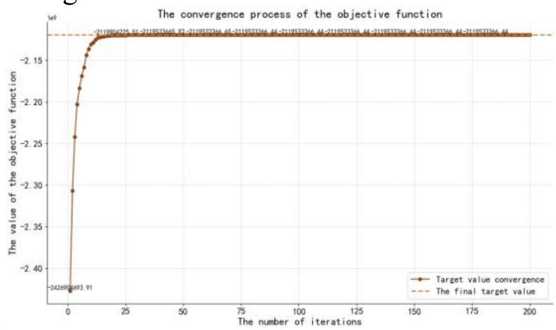

Figure 6. Convergence Process of the Objective Function in Problem 2

4.7 Problem 3

Based on the previously processed dataset1, we extract the data related to planting and sales, including the planting cost, sales price range, and sales volume of different crops.

4.7.1 Correlation analysis

4.7.1.1 Data preprocessing

Based on the previously processed Dataset1, we extracted data related to planting and sales, including information such as the planting costs, sales price ranges, and sales volumes of different crops.

4.7.1.2 Random generation of unit sales price

Since the sales price is given in the form of a range, representing the lowest and highest sales prices per catty of the crop, we need to generate a specific sales price. We simulate this process using a random number generator. Assuming the sales price range is [$Qmin$, $Qmax$], we randomly generate a value r (satisfying r $\in$ [0, 1]) and calculate the specific unit sales price Q using the following formula:

$$Q = Q_{max} + r \cdot (Q_{max} - Q_{min}) \tag{32}$$

Note:

ris a random number generated within the range of [0, 1].

$Qmin$ is the lower bound of the unit sales price range.

$Qmax$ is the upper bound of the unit sales price range.

Data Grouping and Mean Calculation

To analyze the planting and sales situations of different crop types and names more clearly, we group the data according to crop types and names. The dataset contains multiple crop types, and each type includes different crop names. For each crop type and name, we calculate the average values of related variables such as the planting cost, unit sales price, and sales volume:

$$X_{ij} = \frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} X_{ijk} \quad (33)$$

Note:
represents the number of data records for the jth name in the ith crop type.
represents the value of the kth record under the jth name in the ith crop type.
represents the value of the related variable (such as planting cost, unit sales price, and sales volume) for this crop type and name.

Variable Selection

We selected variables related to yield and sales for subsequent correlation analysis. These variables include:
δ: Planting cost per unit area (yuan/mu).
Q: Unit sales price of the crop (yuan/jin).
D: Sales volume of the crop (jin).
After processing, we obtained the following partial data (partial display):

Table 5. Partial Contents of the sem_data.csv File

Crop Type	Crop Name_x	Planting Cost/(yuan/mu)	Unit Sales Price/(yuan/jin)	Sales Volume/jin
Grain	Wheat	450	3.792824932	45600
Grain	Wheat	450	3.441505341	43200
Grain (Beans)	Black Beans	400	6.501856294	23000
Grain (Beans)	Black Beans	400	6.655795737	21850
Grain (Beans)	Black Beans	400	7.841350807	20700
Grain (Beans)	Red Beans	350	8.593731327	16000
Grain (Beans)	Red Beans	350	7.774234085	15200
Grain (Beans)	Red Beans	350	7.8402815	14400
Grain (Beans)	Green Beans	350	7.628216232	9800
Grain (Beans)	Green Beans	350	7.04618512	9240
Grain (Beans)	Green Beans	350	7.352277093	8820

4.7.1.3 Pearson correlation evaluation

See Table 5 above, for each crop category, we calculated the Pearson correlation coefficients between the planting cost, unit sales price, and sales volume. The Pearson correlation coefficient is a statistical indicator used to measure the strength of the linear relationship between two variables, and its calculation formula is as follows:

$$r(X,Y) = \frac{\rho(X,Y)}{\sigma_X \cdot \sigma_Y} \quad (34)$$

Note:
represents the Pearson correlation coefficient between variables X and Y;
represents the covariance between X and Y, which reflects the degree of joint variation of the two variables;
and are the standard deviations of X and Y respectively, which quantify the fluctuation levels of the respective variables.

The value range of the Pearson correlation coefficient is between -1 and 1:
When, it means there is a perfect positive correlation between X and Y;
When, it indicates there is a perfect negative correlation between X and Y;
When = 0, it means there is no linear correlation between X and Y.

Through analysis, we can understand the interdependence among the planting cost, unit sales price, and sales volume of different crop types, and further reveal the strength and direction of their linear relationships. This is of great significance for understanding crop market dynamics and formulating planting strategies.

Table 6. Correlation Matrix – Grain

Correlation Matrix – Grain			
	Planting Cost/(yuan/mu)	Unit Sales Price/(yuan/jin)	Sales Volume/jin
Sales Volume/jin	0.23	-0.51	1.00
Unit Sales Price/(yuan/jin)	-0.19	1.00	-0.51
Planting Cost/(yuan/mu)	1.00	-0.19	0.29

Table 7. Correlation Matrix – Grain (Beans)

Correlation Matrix – Grain (Beans)			
	Planting Cost/(yuan/mu)	Unit Sales Price/(yuan/jin)	Sales Volume/jin
Sales Volume/jin	0.43	-0.24	1.00

Unit Sales Price/(yuan/jin)	-0.73	1.00	-0.24
Planting Cost/(yuan/mu)	1.00	-0.73	0.43

Table 8. Correlation Matrix – Vegetables

Correlation Matrix – Vegetables			
	Planting Cost/(yuan/mu)	Unit Sales Price/(yuan/jin)	Sales Volume/jin
Sales Volume/jin	-0.22	-0.41	1.00
Unit Sales Price/(yuan/jin)	0.42	1.00	-0.41
Planting Cost/(yuan/mu)	1.00	0.42	-0.22

Table 9. Correlation Matrix – Vegetables (Beans)

Correlation Matrix – Vegetables (Beans)			
	Planting Cost/(yuan/mu)	Unit Sales Price/(yuan/jin)	Sales Volume/jin
Sales Volume/jin	-0.12	0.30	1.00
Unit Sales Price/(yuan/jin)	0.24	1.00	0.30
Planting Cost/(yuan/mu)	1.00	0.24	-0.12

Table 10. Correlation Matrix – Edible Fungi

Correlation Matrix – Edible Fungi			
	Planting Cost/(yuan/mu)	Unit Sales Price/(yuan/jin)	Sales Volume/jin
Sales Volume/jin	-0.11	-0.80	1.00
Unit Sales Price/(yuan/jin)	0.47	1.00	-0.80
Planting Cost/(yuan/mu)	1.00	0.47	-0.11

See Table 6, Table 7, Table 8, Table 9, Table 10 above, Through the correlation analysis of crop types, we found that:

For grain crops: The cost is slightly negatively correlated with the selling price, the sales volume is slightly positively correlated with the cost, but the selling price is moderately negatively correlated with the sales volume.

For bean grain crops: The cost is strongly negatively correlated with the selling price, the sales volume is moderately positively correlated with the cost, and the selling price has a small impact on the sales volume.

For edible fungi: The cost is positively correlated with the selling price, the sales volume has little impact, but the selling price is strongly negatively correlated with the sales volume.

For vegetables: The cost is positively correlated with the selling price, the sales volume is slightly negatively correlated with the cost, and the selling price is moderately negatively correlated with the sales volume.

For bean vegetables: The cost has a small impact on the selling price and the sales volume, and the selling price is slightly positively correlated with the sales volume.

Based on the comprehensive analysis, when making crop planting decisions, we need to consider the general inhibitory effect of price on sales volume as well as the substitution and complementarity among crops. The planting strategies from 2024 to 2030 should combine the interactions among crops, and through simulation optimization, formulate reasonable plans. Therefore, in the crop planting strategies from 2024 to 2030, we need to comprehensively consider the benefits of individual crops and the interactions among different crops. By simulating various scenarios, optimize the planting strategies, and compare and analyze the results with the plans considering relevant factors to seek more reasonable and feasible planting plans.

4.8 Establishment and Solution of Multi-factor Optimization Model

4.8.1 Model establishment

In Problem 3, based on the achievements of Problem 2, we further considered in depth the substitutability, complementarity among different crop types, as well as the interrelationships among the expected sales volume, selling price, and planting cost of crops. Through the analysis of the provided data, we found that there are certain correlations among the cost, price, and sales volume of crops such as grains, grains (beans), vegetables, vegetables (beans), and edible fungi. We incorporated these correlations into the model and constructed a more practical optimization scheme by adjusting parameters.

a. Objective Function: The objective of the model is still to maximize the net profit of crops in each year, that is:

$$E = \sum_{i,j,t} (Q_{i,t} \cdot P_{i,j,t} \cdot Z_{i,j,t} - \delta_{i,j,t} \cdot Z_{i,j,t} + \max\left(0, \left(P_{i,j,t} \cdot Z_{i,j,t} - D_{i,t}\right) \cdot 0.5 \cdot Q_{i,t}\right)) \quad (35)$$

represents the income from the part where the crop exceeds the expected sales volume and is sold at 50% of the 2023 price.

b. Introduction of Variable Correlations: Through the data analysis of the planting cost, unit sales price, and sales volume of different crop types, we can establish the following correlations:

For grain crops: The planting cost is positively correlated with the sales volume (0.230), indicating that as the demand increases, the cost rises; the unit sales price is negatively correlated with the sales volume (-0.508), meaning that an increase in the sales volume will lead to a decrease in the price.

$$Q_{i,t} = Q_{i,t-1} \times (1 - 0.508 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (36)$$

$$\delta_{i,t} = \delta_{i,t-1} \times (1 + 0.230 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (37)$$

For grain (beans) crops: The planting cost is negatively correlated with the unit sales price (-0.727), and is positively correlated with the sales volume (0.428).

$$Q_{i,t} = Q_{i,t-1} \times (1 - 0.727 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (38)$$

$$\delta_{i,t} = \delta_{i,t-1} \times (1 + 0.428 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (39)$$

For vegetable crops: The planting cost is positively correlated with the unit sales price (0.418), and is negatively correlated with the sales volume (-0.224); the unit sales price is also negatively correlated with the sales volume (-0.405).

$$Q_{i,t} = Q_{i,t-1} \times (1 - 0.224 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (40)$$

$$\delta_{i,t} = \delta_{i,t-1} \times (1 + 0.418 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (41)$$

For vegetable (beans) crops: The planting cost is positively correlated with the unit sales price (0.248), and is negatively correlated with the sales volume (-0.116); the unit sales price is positively correlated with the sales volume (0.299).

$$Q_{i,t} = Q_{i,t-1} \times (1 - 0.116 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (42)$$

$$\delta_{i,t} = \delta_{i,t-1} \times (1 + 0.248 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (43)$$

For edible fungi: The planting cost is positively correlated with the unit sales price (0.468), and is negatively correlated with the sales volume (-0.109); the unit sales price is also negatively correlated with the sales volume (-0.805).

$$Q_{i,t} = Q_{i,t-1} \times (1 - 0.109 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (44)$$

$$\delta_{i,t} = \delta_{i,t-1} \times (1 + 0.468 \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (45)$$

These correlations reflect the complex relationships among the planting cost, unit sales price, and sales volume of different crop types, providing an important basis for the establishment of the model. Through these correlations, we can more accurately predict and plan the crop planting strategies to achieve the goal of maximizing the net profit.

c. Incorporation of Substitutability and Complementarity into the Model: In the process of formulating the planting strategy, we take into account the substitutability and complementarity among crops and construct the corresponding model.

Substitutability: The model considers the degree of substitution in the crop market, simulates the changes in price and sales volume, and recommends crops with better sales performance.

Complementarity: The model gives priority to crop combinations with high synergy effects to improve economic benefits or reduce costs, such as resource sharing and cost sharing. Implementation methods:

Substitutability simulation: Dynamic market simulation predicts crop performance and adjusts the planting strategy in response to market changes.

Complementarity optimization: The algorithm evaluates the synergy effects of crop combinations to determine the best matching scheme and maximize the benefits. This model provides precise decision-making support for farmers and helps them formulate the optimal planting plan. Through such model construction, we can provide farmers with a more precise and practical decision-making support tool to help them formulate the optimal planting plan according to the dynamic changes in the market and the characteristics of the crops.

4.8.2 Solution by genetic algorithm

The genetic algorithm is a heuristic search technique that simulates the natural evolution process and is used to find the optimal solution in the problem space. [5] This study uses the genetic algorithm to optimize the crop planting strategies, aiming to maximize the net profit, considering the substitutability, complementarity, and correlations among crops to achieve the maximization of the planting income from 2024 to 2030.

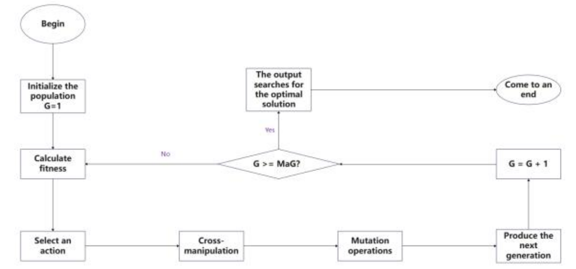

Figure 7. Genetic Algorithm Process

See Figure 7 above, in the constructed model, there are interrelationships among the planting cost, unit sales price, and sales volume of crops, and different crops also show substitutability and complementarity. To dynamically adjust the annual crop price and planting cost, we introduce the following formulas to simulate these correlations:

Negative correlation between sales volume and price:

$$Q_{i,t} = Q_{i,t-1} \times (1 - \alpha \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (46)$$

where, represents the sensitivity of sales volume to price. When the sales volume increases, the price will decrease accordingly.

Positive correlation between sales volume and planting cost:

$$\delta_{i,t} = \delta_{i,t-1} \times (1 + \beta \cdot \frac{S_{i,t} - S_{i,t-1}}{S_{i,t-1}}) \quad (47)$$

where, represents the sensitivity of sales volume to planting cost. An increase in the sales volume may lead to an increase in the planting cost.

The substitutability and complementarity among crops are reflected through the fitness function of the genetic algorithm. In terms of substitutability, the algorithm will tend to select substitute crops when the benefits are similar; for crops with complementarity, the fitness of their combined planting will be higher. The following are the specific implementation steps:

a. Individual Encoding: The individuals representing the crop planting schemes on the plots are encoded using binary or real numbers, with the crop combinations and areas on the plots as genes to generate diverse planting schemes.

b. Fitness Function: The fitness function in the genetic algorithm is used to evaluate the superiority or inferiority of the planting schemes, with the net profit as the fitness value, considering crop sales, costs, and substitutability and complementarity.

c. Selection Mechanism: Through the roulette wheel or tournament method, select high-fitness individuals to enter the next generation, so that the excellent genes can be passed on.

d. Cross-Operation: Generate new planting schemes through gene combination, maintain diversity, improve the quality of the solution, and increase the possibility of obtaining the global optimal solution.

e. Mutation-Operation: Randomly change the genes to increase the diversity of the schemes, avoid local optima, and explore new schemes.

f. Termination Conditions: The algorithm terminates when it reaches the maximum number of iterations or when the fitness no longer significantly improves. The optimal solution is the individual with the highest fitness.

The genetic algorithm efficiently searches for the optimal planting scheme through selection, cross- and mutation-operations, considering multiple factors, avoiding local optima, and achieving the maximization of the net profit and long-term sustainability.

4.8.3 Result output

Model iteration convergence: the genetic algorithm simulates natural evolution to optimize the crop combinations to achieve the maximum benefit. By iteratively selecting individuals with high fitness, passing on excellent genes, and determining the optimal planting scheme, it ensures the planting benefits from 2024 to 2030. We formulated a multi-season planting scheme with 200 iterations and showed the improvement and stability of the population fitness through the iteration graph, revealing the optimal solution.

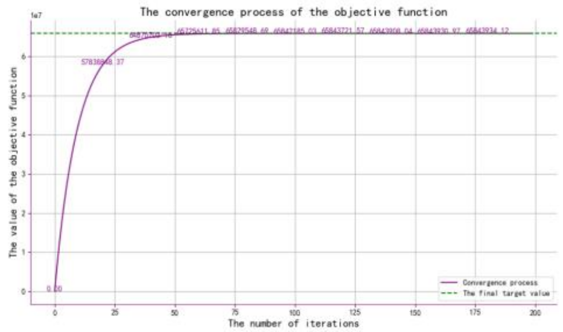

Figure 8. The Convergence Process of the Objective Function in Problem 3

See Figure 8 above, in the iteration graph, we can observe that the vertical axis represents the value of the objective function, and the horizontal axis represents the number of iterations. It can be clearly seen from the graph that the algorithm has entered a relatively stable state at about the 60th iteration, and the value of the objective function has hardly changed significantly in the subsequent iterations. This phenomenon indicates that the algorithm has converged and probably has found the optimal solution or a very close solution.

This shows the effectiveness of the genetic algorithm in dealing with such optimization problems. It not only successfully optimizes the problem but also finds a solution with a relatively high quality level. The stable convergence characteristic of the iteration curve further reveals the robustness and computational efficiency of the genetic algorithm.[6] This means that the algorithm can quickly converge to a satisfactory solution within a limited number of iterations, thus saving computational resources in practical applications and improving the efficiency of the solution process.

2024 to 2030 Annual Optimal Crop Planting Schemes

After optimizing the crop planting strategies from 2024 to 2030, we obtained a series of annual income results. The following is a summary of the target values for each year:

Table 11. 2024 to 2030 Annual Income Results of Crop Planting

2024 to 2030 Annual Income Results of Crop Planting			
Year	Target Income	Additional Income	Total Income
2024	7062443	3011847	10074290
2025	5965309	3231662	9196971
2026	6026631	3149235	9175866
2027	6422375	3030512	9452887
2028	6742523	3012254	9754778
2029	6658412	3023557	9681969
2030	7002385	3100452	10102838

See Table 11 above, the analysis shows that from 2024 to 2030, the values of the objective function generally present a stable trend, with little variation in the target values each year, mainly fluctuating between 9 million and 10 million. Specifically, the target values in 2024 and 2030 are relatively high, reaching 10,074,290 and 10,102,838 respectively, which implies that the planting strategies adopted in these two years may have generated higher economic benefits. Although the target values in other years are slightly lower, they still remain at a relatively stable level, indicating that the optimization results of the planting schemes during the entire research period, that is, from 2024 to 2030, are persistent and reliable.

Overall, these data demonstrate the stable performance of the optimized planting strategies in terms of economic benefits. Although there are certain differences in the annual revenues, these differences are not significant. This provides valuable reference information for strategy makers, helping them formulate more robust and long-term planting plans. This stability also reflects the effectiveness of the optimization model, which can maintain good economic benefits even under different market and environmental conditions.

Comparison of the Optimal Crop Planting Schemes for 2024 - 2030 between Problem 2 and Problem 3.

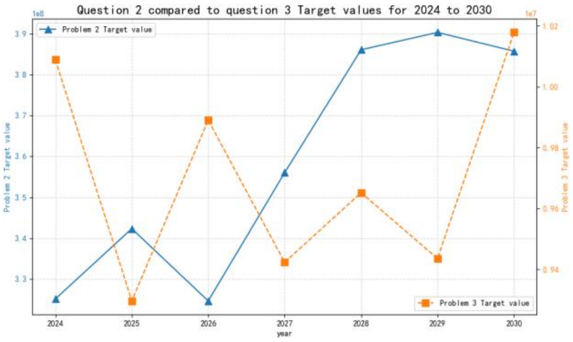

Figure 9. Comparison of the Optimal Crop Planting Schemes for 2024 - 2030 between Problem 2 and Problem 3

See Figure 9 above, the analysis reveals that during the period from 2024 to 2030, the revenues of Problem 2 are generally higher than those of Problem 3, as Problem 2 does not consider the substitutability, complementarity, and correlations among crops. The revenues of Problem 3 are more stable, focusing on risk management and alleviating market fluctuations. A two-dimensional kernel density estimation (KDE) graph is used to analyze the relationship and distribution characteristics between the two, providing precise support for decision-making. In the previous analysis, we have already noticed that the target values of Problem 2 are higher, while Problem 3 shows a more stable revenue curve. This phenomenon reveals the trade-off between revenue optimization and risk control between the two.

To further explore the relationship between the target values of Problem 2 and Problem 3, as well as their distribution characteristics in different intervals, we conducted an in-depth analysis using a two-dimensional kernel density estimation (KDE) graph [7]. This method helps us to more comprehensively understand the revenue characteristics and risk distributions of the two strategies, providing more precise decision-making support for decision-makers.

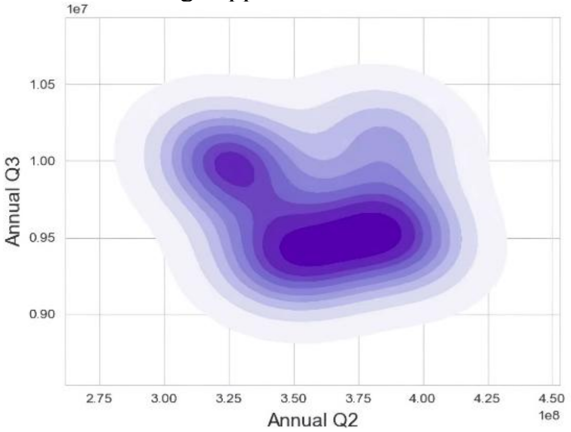

Figure 10. Two-Dimensional Kernel Density Estimation (KDE)

See Figure 10 above, the two-dimensional kernel density estimation (KDE) graph reveals the joint distribution of the target values of Problem 2 and Problem 3, with the dark areas indicating high correlations. The KDE graph shows that there is a high joint density in some intervals between the two, indicating that these planting strategies have good optimization effects. This trend provides a reference for strategy optimization, helping to find the balance between revenue and risk. The comparative analysis shows that the optimal planting schemes during the period from 2024 to 2030 combine the strategies of maximizing revenue and managing risks.

5. Result Analysis

In the process of constructing the crop planting optimization model [8], we introduced a subjective constraint condition, that is, the minimum planting area of each cultivated land cannot be less than 10% of the total area. Such a constraint may have a certain impact on the final results of the model. To evaluate the sensitivity of this constraint, we conducted a specific sensitivity analysis. In this analysis, we took the minimum planting area constraint as a variable and set the iteration step size to 0.02 to explore its impact on the revenue. The relevant sensitivity analysis results are shown in the following charts.

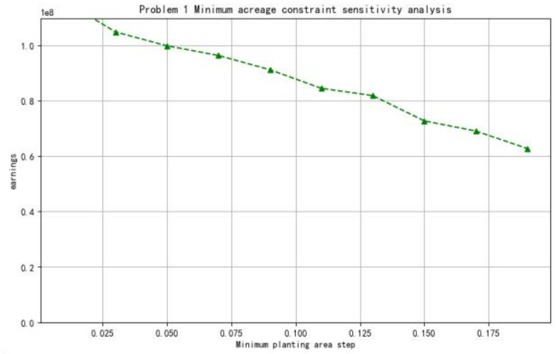

Figure 11. Problem 1 Minimum Planting Area

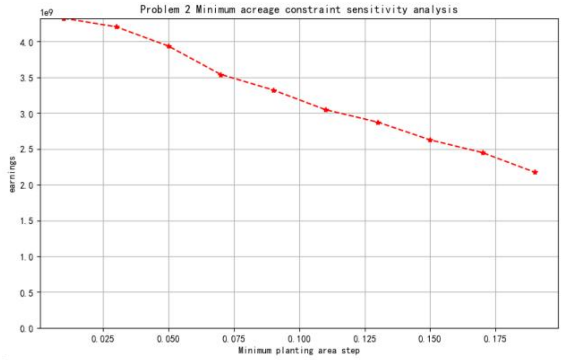

Figure 12. Problem 2 Minimum PlantingArea

For the model in Figure 11, the sensitivity analysis shows that as the minimum planting area [9] constraint is gradually relaxed, the overall revenue of crop planting shows an obvious downward trend. This indicates that when the minimum planting area constraint is relatively high, the model can better restrict uneconomical crop planting combinations, thereby achieving higher revenues. However, as the planting area constraint is relaxed, the flexibility of the model increases, and some inefficient crop planting schemes may be introduced, resulting in a decrease in the overall revenue.

From the sensitivity analysis charts of Figure 12, it can be seen that as the minimum planting area constraint is relaxed, the overall revenue of crop planting shows a gradual downward trend. This indicates that when the planting area constraint is relatively strict, the model can better optimize the crop planting combinations, thereby achieving higher revenues. However, as the planting area constraint is gradually relaxed, the degree of freedom of the model increases, causing inefficient or uneconomical crop combinations to enter the optimization scheme, resulting in a decrease in the revenue.

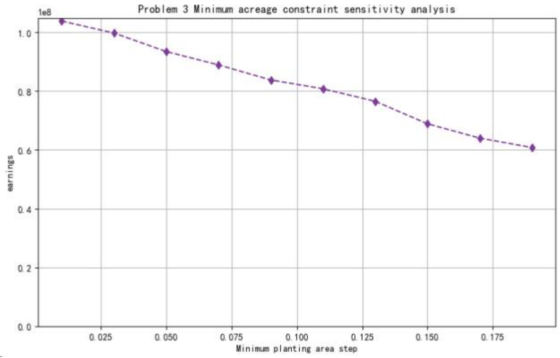

Figure 13. Sensitivity Analysis of MinimPlanting Area

When analyzing the Figure 13 model, the analysis shows that after relaxing the minimum planting area constraint, the decrease in revenue shows fluctuations, and different crops have different sensitivities to the constraint changes. By comparing different algorithm models, it is found that strict constraints help to concentrate resources on high-revenue crops, while excessive relaxation may reduce the revenue. The sensitivity analysis indicates that reasonably setting the minimum planting area constraint is crucial for optimizing the planting scheme [10], and a balance needs to be found between flexibility and maximizing revenue.

6. Conclusions

This study aims to optimize the crop planting strategies of a certain village in the mountainous area of North China from 2024 to 2030 to maximize the net profit and minimize the loss due to slow sales. Considerations include plot types, yields, costs, sales volumes, crop rotation, and market demand uncertainties. The economic benefits are evaluated through mathematical models, and the substitutability, complementarity, and price correlations among crops are explored to provide scientific planting plans for the village.

6.1 Model Advantages

The advantages of the model lie in comprehensively considering multiple factors, including costs, prices, yields, crop substitutability, complementarity, and price correlations, adapting to the complexity of agricultural production. The introduction of the time dimension enables dynamic adjustment of strategies, enhancing flexibility. Considering market uncertainties and price fluctuations reduces risks. The adoption of Monte Carlo simulation, particle swarm optimization, and genetic algorithms improves the solution efficiency. Based on historical and current data, the input accuracy is ensured, providing guidance for agricultural production and policy formulation.

6.2 Model Disadvantages

The disadvantages of the model lie in its dependence on data quality and integrity. The difficulty in data processing may affect its practicality. High computational resource requirements limit its application in resource-limited environments. It may not be able to fully cover real-life situations such as extreme weather and policy changes, affecting the prediction accuracy. The optimization algorithms have limitations and may not be applicable to all problems. Meanwhile, environmental protection and ecological balance may be overlooked.

6.3 Model Prospects

This study provides a comprehensive crop planting strategy planning for a certain village in the mountainous area of North China. The advantages of the model lie in comprehensively considering factors, adopting advanced algorithms and data-driven methods, but there are limitations such as data dependence and high computational resource requirements. Future research can expand data integration processing, algorithm innovation, and model expansion to improve data quality and model practicality, thereby facilitating agricultural development.

Acknowledgments

This study was supported by the Guangzhou Huashang College Mentorship Program Project (Project No. 2024HSDS12).

References

[1] Zhou Min. Adjustment of Rice Planting Structure and Industrial Upgrading under the Comprehensive Transformation of Rural Areas. Northern Rice,2024,54(04): 55-57. DOI: 10.16170/j.cnki.1673-6737.2024.04.028.
[2] Zhu Lulu. Monte Carlo method and application. Central China Normal University, 2014.
[3] Fang Wei. Research on Project Risk Management Method Based on Monte Carlo Simulation. Computer and Modernization,2012, (04):33-36.

[4] Chang Youqu, Xiao Guiyuan, Zeng Min. Discussion and Research on Greedy Algorithm. Journal of Chongqing Electric Power College,2008, (03):40-42+47.
[5] Ge Jike, Qiu Yuhui, Wu Chunming, et al. A Review of Genetic Algorithm Research. Computer Application Research,2008, (10):2911-2916.
[6] Zhao Wence, Shu Chuanhua, Wang Shengxi. Optimal Design of Low-inclination Satellite Constellation Using Genetic Algorithm. Aerospace Control,2024,42(05):76-82. DOI: 10.16804/j.cnki.issn1006-3242.2024.05.003.
[7] Qiu Kebin, Chen Weiguo, Zhang Zhiqiang, et al. Color Feature Extraction Method of Colored Fiber Based on Two-dimensional Gaussian Kernel Density Estimation. Journal of Textile Research,2024,45(05):85-93. DOI: 10.13475/j.fzxb.20221100901.
[8] Feng Guangliang. Research on Agricultural Planting Structure Allocation in the Lower Reaches of Tarim River Based on Double-layer Planting Optimization Model. Water conservancy technical supervision,2024, (05):213-217.
[9] Jin Man, Cao Jitao. Analysis of Centrifugal Compressor Application Model and Parameter Sensitivity. Today's Manufacturing and Upgrading,2024, (05):56-58.
[10]Zhang Rui, Yan Wenyi, Yang Shumei, et al. 2023 Crop Yield and Planting Area in Russia. Heilongjiang Grain, 2024, (09):19-21.

