Research on Crop Planting Strategy Based on Monte Carlo Simulation and Genetic Algorithm

Bin Li*

School of Economics and Management, China University of Mining and Technology, Xuzhou, China, 221116

*Corresponding author: 15137667561@163.com

Abstract. Along with the gradual progress of urbanization, rural areas are facing challenges such as limited land resources, market fluctuations, and environmental changes, and how to maximize economic benefits by optimizing crop planting structure and increasing yields has become a key issue in current agricultural development. In this study, we first established a planning model for the target area by optimizing planting strategies and introducing relevant constraints such as land, climate, and market. Subsequently, Monte Carlo simulation was used to generate random samples to simulate crop planting scenarios under different scenarios. The model was optimized by the Pareto multi-objective genetic algorithm to find out the best trade-off between profit maximization and risk minimization. Finally, the optimal solution was solved using the ideal point method. The analysis results show that the scenario has good applicability in the target area, and the expected annual profit can reach 2.17 million yuan, which has a significant economic reference value.

Keywords: Planning Model, Genetic Algorithm, Monte Carlo Method, Pareto Optimization.

1. Introduction

With the growing global population and increasing demand for food, the importance of agricultural production in guaranteeing food security has become increasingly prominent. How to increase crop yields, optimize planting structures, and enhance farmers' incomes under limited land resources has become a core issue in current agricultural development. Especially in rural areas, due to limited land resources, fluctuating market prices, and changes in the natural environment, rational planning of crop planting layout and maximization of economic benefits is an urgent problem to be solved.

In recent years, the application of intelligent and algorithmic optimization technologies in the field of agriculture has provided new solutions for crop planting planning. Among them, genetic algorithms are gradually playing a key role in the field of agriculture due to their powerful optimization ability, helping to optimize the planting and breeding process. Zhang et al. used genetic algorithms to solve a single-objective planning model and solved the problem of logistics center location[1]. Peng et al. used genetic algorithms to optimize the control parameters of an intelligent irrigation controller for farmland to realize intelligent irrigation control[2]. Zhang et al. used a genetic algorithm to optimally solve the wireless sensor node deployment problem and came up with a deployment scheme with higher coverage under a certain number of sensors[3]. With the modernization and mechanization of agriculture, genetic algorithms can effectively help farmers make reasonable and optimal decisions in an environment with many uncertainties.

In this study, it is necessary to solve the problem of how to effectively utilize arable land resources, rationally allocate planting areas, and ensure the maximization of revenue in the case of stable market demand. Based on this, a planning model is established in this study, and eight constraints (e.g., the total area of crops planted on each plot does not exceed the total area of the plot, and each crop can not be planted in the same plot (including greenhouses) continuously, etc.) are introduced to make the model more relevant to the reality, to ensure the reasonableness and feasibility of the model. Considering the uncertainty of many factors in the actual situation, this study utilized Monte Carlo simulation to generate random samples through which the effects of uncertainty were captured. Then, the Pareto multi-objective genetic algorithm was used to optimize and find multiple possible optimal solutions by simulating planting scenarios under different market conditions. Finally, a sensitivity analysis of these possible optimal solutions was performed to select the one that achieves the best balance between profit maximization and risk minimization.

2. Data Sources and Visualization

The data for this study was obtained from www.mcm.edu.cn. Figure 1 shows the total area under different categories of crops in 2023 and it can be seen that the major crop grown in the village in the last year was food grains which accounted for more than 85% of the total area under cultivation, vegetables accounted for 10% and edible mushrooms accounted for only one percent of the area under cultivation. Moreover, the area under non-leguminous crops is lower than that of non-leguminous crops.

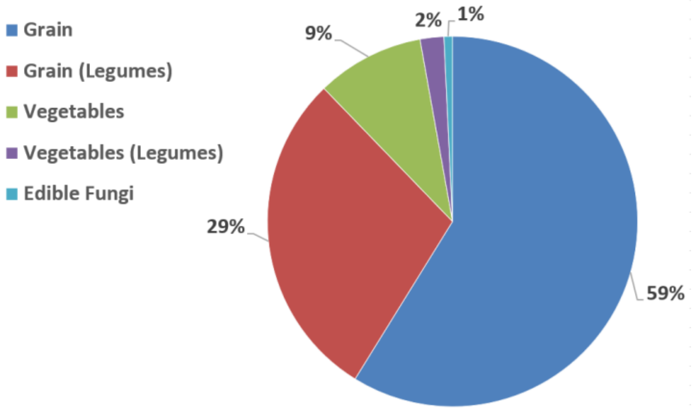

Figure 1. Map of total acreage of different crop types in 2023

In Figure 2, which shows the total sales and profits of different crops in 2023, it can be seen that the total sales and profits of non-legume grains are the highest, indicating that the annual revenue of the region is mainly derived from the cultivation of non-legume grains. Moreover, the total sales of non-leguminous vegetables in the region have exceeded that of leguminous grains, indicating that non-leguminous vegetable crops are also the major crops in the region and occupy an important share of the local market. In addition, edible fungi crops have a much lower annual production than the other categories, but have a higher total profit, indicating that the profitability of this category is very high.

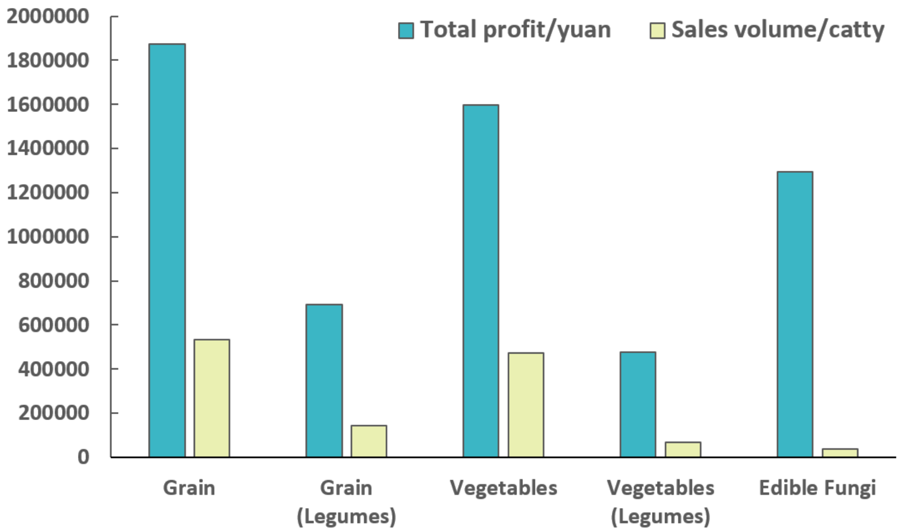

Figure 2. Map of Sales and Profit of Different Crops by Category in 2023

As Figure 3 shows the data of total profit and profit margin of different crops in 2023 (the data are arranged in descending order of total profit, and only the first 8 are shown due to space limitation), it can be seen from the bar part of the figure that the total profit of edible fungi crops such as elm mushrooms, wheat, cereals, and morel mushrooms in this region is higher, which is the main source of income for agribusinesses in this region; and from the folded part of the figure, it can be seen that elm mushrooms have a significantly higher profit margin than the other crops, suggesting that the crop has great commercial value.

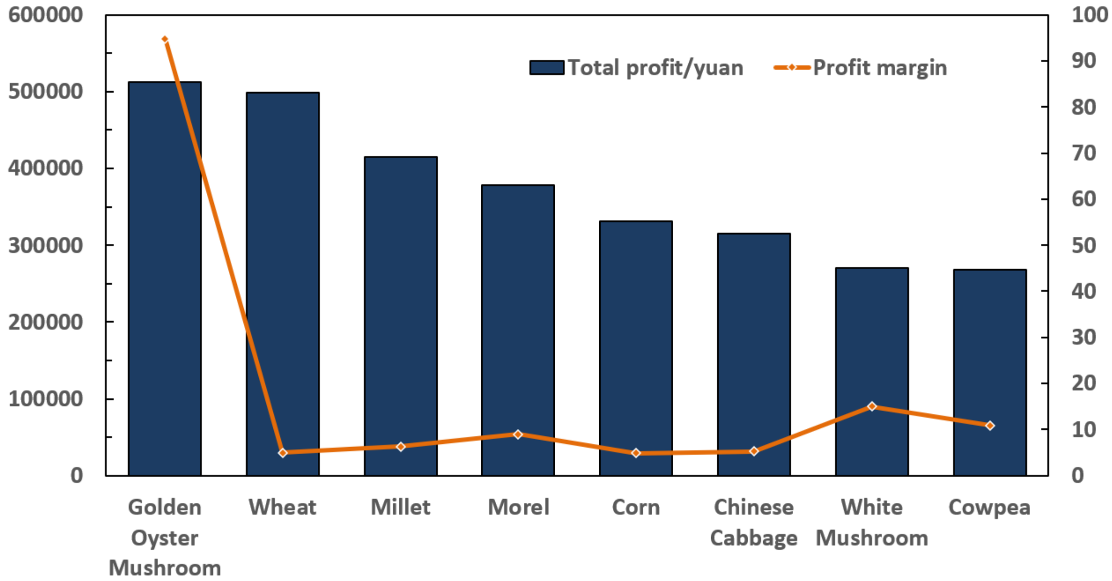

Figure 3. Map of Sales and Profit of Different Crops by Category in 2023

3. Crop cultivation optimization model

Since the expected future sales volume, planting cost, acre yield, and sales price of various crops remain stable relative to 2023, the cost, sales price, and acre yield of each crop planted in different plots can be obtained by data cleansing and data processing of the available data. Eventually, the expected sales volume, planting cost, acre yield, and sales price of each crop in 2024~and 2030 can be obtained.

After obtaining specific data on the expected sales volume, planting cost, acre yield, and sales price of various crops for the next seven years, a planning model was developed in this study. In this planning model, the area under crop cultivation is selected as the decision variable, which is represented by the set X . The decision variable is the area under crop cultivation. It is specified as follows:

$$X=\{X_{ijtk}\,|\,i\in\{1,2,...,41\},j\in\{1,2,...,54\},t\in\{2024,...,2030\},k\in\{1,2\}\} \quad (1)$$

Where X_{ijtk} denotes the area of the j th plot planted with the i th crop in the k th season of the year t , i denotes the crop number, j denotes the plot number, t denotes the year, and k denotes the quarter.

The model has the objective function of maximizing the total profit of the village over the next 7 years. It is shown below:

$$maxE(Z_1)=\frac{\sum_{t=2024}^{2030}\sum_{k=1}^{2}\sum_{i=1}^{41}\left[P_{it}\cdot min(Q_{itk},D_{it})-C_{it}\cdot\sum_{j=1}^{54}X_{ijtk}\right]}{n} \quad (2)$$

Where n is the number of predictions, q_{it} is the expected mu yield of the i th crop in a year t , D_{it} the expected sales volume of the i th crop in a year t , P_{it} the expected sales unit price of the

i th crop in a year t, C_{it} is the planting cost of the i th crop in a year t, and Q_{itk} is the expected total production of the i th crop in the k th quarter of the t year.

Considering the complexity of the geographic environment of the village and the planting characteristics of various crops, this study formulated eight constraints in the model to ensure that the final planting scheme obtained can fit the complex actual situation and have high feasibility[4].

The total area of crops grown on each plot does not exceed the total area of the plot.

Ensure that the total area of crops planted on each plot in each season of the year is less than or equal to the total area of the plot. The expression is as follows:

$$\sum_{i=1}^{41} X_{ijtk} \leqslant S_j, \forall j, k \in \{1, 2\}, t \in \{2024, 2025, \cdots, 2030\} \tag{3}$$

Each crop cannot be planted on the same plot (including greenhouses) in consecutive crops[5]. Then at least one of the two adjacent years of the same crop on the same plot will be 0. The expression is given below:

$$X_{ijtk} \cdot X_{ij(t+1)k} = 0 \tag{4}$$

All land in each plot (including sheds) is planted with a legume crop at least once in three years[6]. Then each plot will be planted with legume crop not less than 1 times in three adjacent years with the following expression:

$$\sum_{t}^{t+2} \sum_{i \in \zeta_l} y_{ijtk} \geqslant 1, \forall j, k \in \{1, 2\}, t \in \{2024, 2025, \cdots, 2030\} \tag{5}$$

$$\begin{cases} y_{ijtk} = 0 & if\, X_{ijtk} = 0 \\ y_{ijtk} = 1 & if\, X_{ijtk} > 0 \end{cases} \tag{6}$$

where y_{ijtk} is a 0-1 variable indicating whether the j th plot in year t grows the i th crop in the k th season, and a value of 0 means no crop is grown, and vice versa; ζ_l is the numbered set of legume crops, i.e., $\zeta_l = \{1, 2, 3, 4, 5, 17, 18, 19\}$

Flat drylands, terraces, and hillsides are suitable for growing food crops (except rice) in a single season each year.

The expression is given below:

$$\sum_{i \in \zeta_{gts}} \sum_{k=1}^{2} X_{ijtk} \leqslant S_j, j \in \{1, 2, 3, \cdots, 26\} \tag{7}$$

where ζ_{gts} is the numbered set of crops suitable for planting in flat dry land, terraced land, and hillside land each year, i.e., $\zeta_{gts} = \{1, 2, 3, \cdots, 15\}$

Area constraints for growing rice in a single season on irrigated land.

Here first, only the case of growing rice in a single season is considered, and the following constraints will consider the case of growing vegetables in two seasons with the following expressions:

$$\sum_{k=1}^{2} X_{ijtk} \leqslant S_j, i = 16, j \in \{27, 28, \cdots, 34\} \tag{8}$$

where 16 is the crop number of rice, then $i = 16$ indicates that the crop is rice.

When two seasons of vegetables are planted on a watered site, the first season can be planted with a variety of vegetables (except cabbage, white radish, and carrot); the second season can only be planted with one of cabbage, white radish, and carrot.

When a watered land grows vegetables in the first season, the total area planted with cabbage, white radish, and carrot is 0; when the watered land grows cabbage, white radish, or carrot in the second season, the total area planted with other vegetables except these three is 0. The expression is as follows:

$$\begin{cases} \sum_{i \in \zeta_{s1}} X_{ijtk} \leqslant S_j, \sum_{i \in \zeta_{s2}} y_{ijtk} = 0, k = 1 \\ \sum_{i \in \zeta_{s2}} X_{ijtk} \leqslant S_j, \sum_{i \in \zeta_{s1}} y_{ijtk} = 0, k = 2 \end{cases} \tag{9}$$

where ζ_{s1} is the numbered set of crops that can be grown in the first season of the watered land, i.e., $\zeta_{s1} = \{17, 18, \cdots, 34\}$ ζ_{s2} is the numbered set of crops that can be grown in the second season of the watered land, i.e., $\zeta_{s1} = \{35, 36, 37\}$

Ordinary greenhouses grow two crops per year; in the first season, a variety of vegetables can be grown (except cabbage, white radishes, and carrots), and in the second season only edible mushrooms can be grown.

The expression is as follows:

$$\begin{cases} \sum_{i \in \zeta_{p1}} X_{ijtk} \leqslant S_j, \sum_{i \in \zeta_{p2}} y_{ijtk} = 0, j \in \{35, 36, \cdots, 50\}, k = 1 \\ \sum_{i \in \zeta_{p2}} X_{ijtk} \leqslant S_j, \sum_{i \in \zeta_{p1}} y_{ijtk} = 0, j \in \{35, 36, \cdots, 50\}, k = 2 \end{cases} \tag{10}$$

Where ζ_{p1} is the numbered set of crops that can be grown in the first quarter of the common shed, i.e., $\zeta_{p1} = \{17, 18, \cdots, 34\}$ ζ_{p2} is the numbered set of crops that can be grown in the second quarter of the common shed, i.e., $\zeta_{p2} = \{38, 39, 40, 41\}$

Intelligent greenhouses can grow two seasons of vegetables (except cabbage, white radish, and carrot) every year.

The expression is given below:

$$\sum_{i \in \zeta_z} X_{ijtk \leqslant S_j}, \sum_{i \in \zeta_{s1}} y_{ijtk = 0}, j \in \{51, 52, 53, 54\}, k \in \{1, 2\} \tag{11}$$

where ζ_z is the numbered set of crops that can be grown in the smart shed, that is, $\zeta_z = \{17, 18, \cdots, 34\}$

4. Risk and Uncertainty Management in Planting Optimization

4.1. Optimization of models

Considering the reality of carrying out planting, variables such as expected sales volume, acreage, planting costs, and sales price are uncertain, there is a potential risk of planting. The planting of crops has a large period, a large number of crops, and the planting program is irreversible, so the sensitivity is low[7]. Therefore, this study adds the objective function of risk minimization based on the original optimization model as follows:

$$min\sigma = \sqrt{\frac{1}{n-1}\sum_{n_0=1}^{n}(\rho_n-\rho_0)^2} \tag{12}$$

Where σ is the standard deviation of the predicted profit, ρ_n the n th predicted profit, and ρ_0 the expected profit.

Empirically, future sales volumes, acreage, costs, and prices for a wide range of crops will fluctuate within a certain range. Based on known information, ranges of uncertainty can be determined for sales volumes, yields, market prices, and planting costs.

Changes in sales volume

Expected future sales of wheat and maize tend to increase, with an average annual growth rate r_{it} between 5 percent and 10 percent, and expected future sales of other crops have a rate of change of ±5% per year relative to 2023, denoted here by η_{it}. The formula is as follows. The formula is as follows:

$$D_{it}=D_{i,2023}\cdot(1+r_{it})^{t-2023}, r_{it}\in[0.05,0.1], i\in\{6,7\}, t\in\{2024,2025,\cdots,2030\} \tag{13}$$

$$D_{it}=D_{i,2023}\cdot(1+\eta_{it})^{t-2023}, \eta_{it}\in[-0.05,+0.05],\ i\in\{1,2,\cdots,5,8,9,\cdots,41\}, \\ t\in\{2024,2025,\cdots,2030\} \tag{14}$$

Changes in mu yield

Crop yields vary by ±10% per year and the rate of change is denoted by μ_{it}. The rate of change is given in the following formula. The formula is given below:

$$q_{it}=q_{i,2023}\cdot(1+\mu_{it})^{t-2023}, \mu_{it}\in\{-0.1,+0.1\}, \forall i, t\in\{2024,2025,\cdots,2030\} \tag{15}$$

Changes in cost

The cost of growing crops increases on average by about 5% per year. The formula is given below:

$$C_{it}=C_{i,2023}\cdot(1+0.05)^{t-2023}, \forall i, t\in\{2024,2025,\cdots,2030\} \tag{16}$$

Changes in selling price

The sales price of food crops is stable; the sales price of vegetable crops tends to increase, with an average increase of about 5% per year; the sales price of edible mushrooms is stable and decreasing, and can decrease by about 1% ~ 5% per year, with the sales price of morel mushrooms decreasing by 5% per year. The specific formula is as follows:

$$P_{it}=P_{i,2023}, i\in\{1,2,\cdots,16\}, t\in\{2024,2025,\cdots,2030\} \tag{17}$$

$$P_{it}=P_{i,2023}\cdot(1+0.05)^{t-2023}, i\in\{17,18,\cdots,37\}, t\in\{2024,2025,\cdots,2030\} \tag{18}$$

$$P_{it}=P_{i,2023}\cdot(1+\varphi_{it})^{t-2023}, \varphi_{it}\in[-0.05,-0.01], i\in\{38,39,40\}, \\ t\in\{2024,2025,\cdots,2030\} \tag{19}$$

$$P_{it}=P_{i,2023}\cdot(1-0.05)^{t-2023}, i=41, t\in\{2024,2025,\cdots,2030\} \tag{20}$$

A large number of random samples were generated through Monte Carlo simulation, and in this study, 1000 simulations were conducted, each simulating a different market price, production, and cost[8]. For each random sample, the expected profit and the standard deviation of profit were calculated separately and the results were recorded for all samples.

Based on Monte Carlo simulation, this study used the Pareto optimization algorithm to construct a model of the optimal trade-off between expected profit and risk[9]. By analyzing the results of multiple simulations, the Pareto frontier, i.e. the set of undominated solutions under the two objectives of profit maximization and risk minimization (minimization of the standard deviation of profit), is obtained[10]. The Pareto frontier demonstrates the distribution of risk at different levels of profit, providing a diverse choice of planting strategies. These strategies cover the full range from low risk and low profit to high risk and high profit, achieving a dynamic balance between return and risk.

To obtain the optimal planting scheme that simultaneously satisfies the two objective functions of profit maximization and risk minimization, this study uses the ideal point method[11]. The ideal point is the point consisting of the optimal solution for each objective. For a multi-objective optimization problem, the optimal values of all the objectives can be calculated separately, and assuming that these optimal values can be achieved at the same time, then the point is the ideal point. For this study, the ideal point is the combination of maximum profit and minimum risk.

In the actual optimization process, the ideal point may not be achieved at the same time. Therefore, the closest solution to the ideal point can be found by calculating the distance from each solution in the solution space to the ideal point. The Euclidean distance is used here to measure the distance between a solution and the ideal point d [12] :

$$d=\sqrt{(P-P_{\text{ideal}})^2+(R-R_{\text{ideal}})^2} \tag{21}$$

P R Where and are the profit and risk of the solution, and P_{ideal} R_{ideal} are the profit and risk of the ideal point.

Let the planting scheme be x, profit be $P(x)$ and risk be $R(x)$, then the optimization objective is to maximize $P(x)$ as well as minimize $R(x)$. The ideal point $(P_{\text{ideal}},R_{\text{ideal}})$ is:

$$P_{\text{ideal}}=\max P(x) \tag{22}$$

$$R_{\text{ideal}}=\min R(x) \tag{23}$$

For each Pareto front solution x_i, calculate its distance from the ideal point:

$$d(x_i)=\sqrt{(P(x_i)-P_{\text{ideal}})^2+(R(x_i)-R_{\text{ideal}})^2} \tag{24}$$

Choose the solution with the smallest $d(x_i)$ as the final solution.

4.2. Solution of the model and analysis of results

The above model was solved and finally, the optimal cropping scheme for crops in this village for the period 2024～2030 was obtained. The planting scheme after clustering is shown in the following table (2024 partial data). The scheme maximizes the total profit based on the comprehensive consideration of market sales, planting cost, and mu yield.

Table 1. Table of optimal planting programmes for 2024 (partial data)

Plot Name	Soybeans	Black Beans	Red Beans	Mung Beans	Cowpeas	Wheat	Corn	Millet
A1	0	0	35.44	0	0	0	0	0
A2	0	0	0	0	0	0	0	0
A3	0	3.491	0	0	0	0	0	0
A4	0	0	0	28.45	0	0	0	0
A5	0	0	0	0	0	0	0	0
A6	0	0	0	0	0	0	0	0
B1	0	0	0	0	0	0	0	0
B2	0	0	0	0	0	0	0	15.95
B3	0	0	26.86	0	0	0	0	0
B4	0	0	0	0	0	0	0	0
B5	0	0	0	0	0	0	0	0
B6	0	0	0	0	0	0	0	0
B7	0	0	18.47	0	0	0	0	0

To facilitate the observation of the distribution of the area and location of the different crops cultivated, we visualized the data from the resulting scenarios, as shown in Figures 4 (partial data for 2024) and 5 (partial data for 2024～2030).

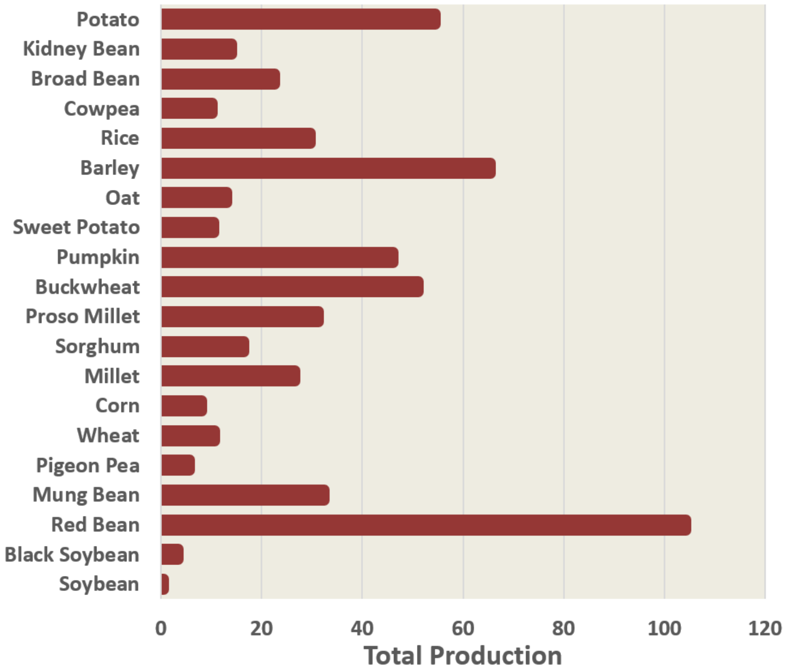

Figure 4. Map of optimal planting by crop for 2024

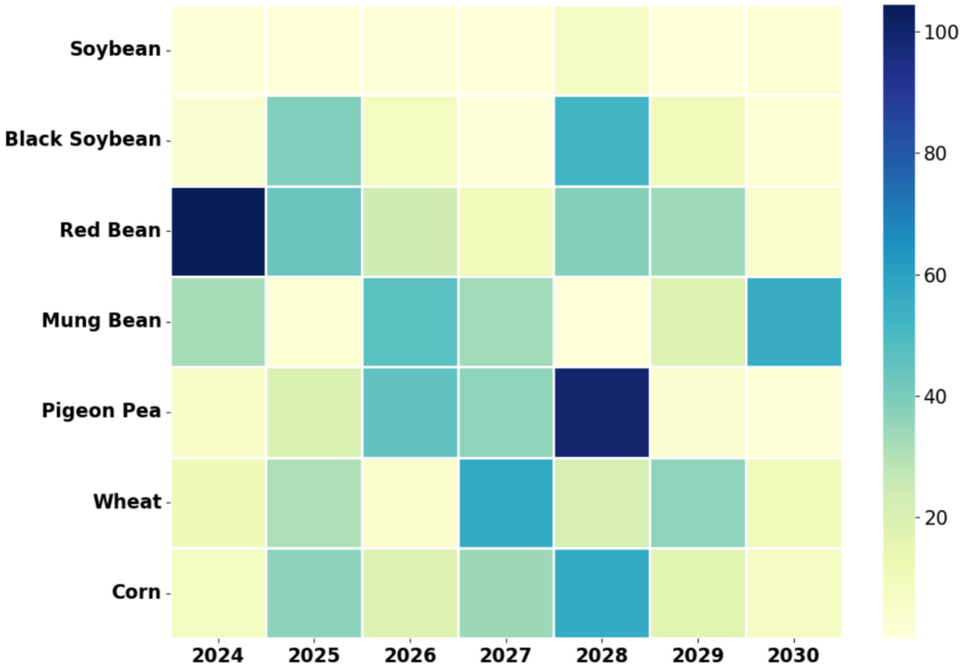

Figure 5. Heat map of the optimal scheme for 2024-2030

As can be seen from Figure 4, the total planting area of each crop in the program is reasonably distributed, and important food crops such as barley, red beans, and potatoes, occupy a larger area, while the rest vary according to their profitability as well as planting conditions. As can be seen from Figure 5, the planting area of each crop in each year has a certain degree of volatility, indicating that the program has fully taken into account the impact of uncertainty factors, so that the optimization results are more realistic. Overall, it can be seen that the optimal solution has a reasonable distribution of crops, makes full use of land resources, and meets various planting constraints. After calculating the projected income as well as the cost of the program, the projected total profit of the optimal program solved by the model for seven years reaches RMB 15218206, and the average annual profit is RMB 2174029. It can be seen that the model also achieves a very considerable profit under the premise of satisfying the production reality, which can effectively promote the economic development of local agro-industry.

5. Conclusions

In this study, the crop planting decision of a rural village in the mountainous region of North China was optimized based on a genetic algorithm, a multi-objective optimization model with profit maximization and risk minimization as the objectives were constructed, and the optimal planting scheme was obtained through Monte Carlo simulation and Pareto optimization. The results show that the scheme can maximally meet the planting requirements and production needs of the region, and the expected total profit from 2024 to 2030 reaches 15.21 million yuan, with an average annual profit of about 2.17 million yuan.

This paper provides a research idea and framework applied to the field related to agricultural economic optimization, based on genetic algorithms, multi-objective optimization, Monte Carlo simulation, and other techniques, and constructs an optimization model with the objectives of profit maximization and risk minimization. Through in-depth analysis and practical verification of crop planting decisions, the feasibility of the method in coping with planting uncertainty and optimizing agricultural economic benefits is demonstrated, which helps to provide effective solutions to similar problems.

References

[1] Donghua Z, Gang S. Research on the Location of Logistics Distribution Center Based on Genetic Algorithm [J]. Logistics Sci-tech, 2024, 47(19): 33-5.

[2] Xiaoli P, Suhong Z, Zheng Y, et al. Design of intelligent irrigation control system for farmland based on genetic algorithm [J]. Technical Supervision in Water Resources, 2024, (03): 38-40+55.

[3] Boyan Z, Tengyue Z. Design of Agricultural Wireless Sensor Node Deployment Algorithm Based on Genetic Algorithm [J]. Agricultural Engineering, 2022, 12(01): 30-4.

[4] Alemany M, Esteso A, Ortiz Á, et al. Centralized and distributed optimization models for the multi-farmer crop planning problem under uncertainty: Application to a fresh tomato Argentinean supply chain case study [J]. Computers & Industrial Engineering, 2021, 153: 107048.

[5] WU H, LIN W. A commentary and development perspective on the consecutive monoculture problems of medicinal plants [J]. Chinese Journal of Eco-Agriculture, 2020, 28(6): 775-93.

[6] ZHAO A, BAI Y. The dialectical relationship between "sourcing food from croplandstoring grain in the land" and "sourcing food from grasslandstoring grain in the grass" from the perspective of agricultural ethics——On the "Farmland Protection" in chapter Chapter 2 of the Food Security Law [J]. Acta Prataculturae Sinica, 2024, 33(10): 183-93.

[7] Hao L, Kelin H, Baoguo L. Parameter optimization and sensitivity analysis of soil-crop system model using PEST [J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(3).

[8] Lin S-S, Shen S-L, Zhou A, et al. Approach based on TOPSIS and Monte Carlo simulation methods to evaluate lake eutrophication levels [J]. Water Research, 2020, 187: 116437.

[9] Zhang X, Meng D, Li W, et al. Evolutionary based Pareto optimization algorithms for bi-objective PV array reconfiguration under partial shading conditions [J]. Energy Conversion and Management, 2022, 271: 116308.

[10] Panichella A. An improved Pareto front modeling algorithm for large-scale many-objective optimization; proceedings of the Proceedings of the Genetic and Evolutionary Computation Conference, F, 2022 [C].

[11] Weixiong Z, Yifei R, Jian L, et al. High-speed train stop schedule optimization based on ideal point method [J]. JOURNAL OF BEIJING JIAOTONG UNIVERSITY, 2022, 46(04): 23-30.

[12] Songtao W. Application of euclidean distance clustering in site optimization [J]. Environmental and Development, 2020, 32(11): 144-5.