Research on Crop Planting Scheme in North China Based on Agricultural Production Efficiency Model

Yuntao Xu, Jingmin Zhang, Yongyu Cai *

School of Space Information, Space Engineering University, Beijing, China, 101416

* Corresponding author: thegreatmasterc@hgd.edu.cn

Abstract. This paper addresses the key issues in agricultural economic development within the context of the rural revitalization strategy in North China, proposing an agricultural production efficiency model aimed at maximizing economic benefits through optimized planting strategies. The article reviews the shortcomings of existing research, poses three research questions, and builds a model based on these questions. By introducing Monte Carlo simulation and the Particle Swarm Optimization (PSO) algorithm, it resolves the uncertainties in agricultural production. Utilizing data preprocessing and visualization techniques, an optimal planting plan for the period from 2024 to 2030 is formulated. The research results demonstrate that the proposed model effectively copes with the uncertainties in production, providing a scientific basis for the sustainable development of agricultural economies in the mountainous areas of North China, while also offering important references and guidance for the implementation of the rural revitalization strategy in the region.

Keywords: Agricultural production efficiency model, Monte Carlo simulation, Particle swarm optimization, Inaccuracy.

1. Introduction

With the gradual advancement of the rural revitalization strategy, promoting the economic development of rural areas has become an important link in achieving this goal [1]. Due to the complex terrain and climate conditions in North China, the arable land resources are limited and scattered. Therefore, how to develop organic planting industries in accordance with local conditions and promote the sustainable development of rural economies has significant practical significance. To promote the economic development of rural agriculture in North China and improve economic benefits, it is necessary to formulate reasonable and efficient planting strategies. The main factors influencing planting strategies include water resources, demand, yield per mu, planting costs, and selling prices per unit.

Over the past few decades, there have been many studies on optimizing agricultural planting strategies based on various factors. In these studies, a wide range of optimization techniques have been used to solve multi-objective models [2]-[5]. For instance, Yang Gaiqiang et al. employed linear goal programming to address the problem of fuzzy objectives with fractional goals and determined the optimal planting areas for winter wheat and summer corn [2]. Wu Li et al. comprehensively considered irrigation water consumption, economic benefits, and crop yields, and established a multi-objective optimization model for the planting structure in irrigation areas using the Grey Wolf Optimization (GWO) algorithm, ensuring the coordinated development of resource, economic, and social systems [3]. Wang Fulin et al. used the BP neural network model and regression methods to fit field data, respectively, to optimize planting density and fertilizer application, and verified its feasibility through field experiments [4]. Mao Ye et al. constructed a two-level multi-objective optimization model based on TOPSIS, and obtained the optimal planting structure by integrating the upper-level decisions of the highest total planting income and the minimum total irrigation water consumption and the lower-level decisions of the highest planting income and the lowest planting cost [5]. Although these studies all used various optimization algorithms to solve multi-objective models, the factors considered in establishing the models were relatively few, deviating from the actual agricultural production situation, and the established agricultural production models and planting strategies were not universal.

In this study, we take a mountainous area in North China as the research object, classify the arable land types based on the amount of water resources, and propose the following three research questions (RQs) to address tBased on RQs, the main contributions of this paper are:

1. Proposing an agricultural production benefit model that aims to maximize economic benefits and comprehensively considers factors such as arable land area, planting seasons, crop adaptability and priority, expected sales volume, and crop rotation, closely linked to the actual situation in the mountainous areas of North China, which is conducive to promoting the development of agricultural economies.
2. On the basis of the above model, using the Monte Carlo simulation to reflect the fluctuations in expected sales volume, planting costs, yield per mu, and selling prices in agricultural production and incorporating them into the original model, improving the robustness of the model.
3. Using the particle swarm optimization algorithm to solve the agricultural production model, which has the advantages of strong global search ability and high computational efficiency compared with traditional linear programming.

he deficiencies in the above studies:

RQ1: How to determine the planting strategy to maximize economic benefits by comprehensively considering the arable land area, planting seasons, crop adaptability and priority, expected sales volume, and crop rotation?

RQ2: How to handle the fluctuations in expected sales volume, planting costs, yield per mu, and selling prices of crops to obtain the optimal planting strategy?

RQ3: What method should be used to solve the optimization model established in the above two questions?

2. Methodology

2.1. Description of the Dataset

The experimental data used in this article is from a public dataset of farmland and crops in a village in the mountainous area of North China, which can be obtained at https://www.mcm.edu.cn. This dataset includes the basic situation of the existing farmland and crops in the village, as well as the planting situation and related statistical data of the village's crops in 2023. The basic situation of the existing farmland and crops in the village describes the quantity and area of six types of plots (denoted by A to F for different types), as well as the types of 41 crops and the suitable types of farmlands for planting. The planting situation and related statistical data of the village's crops in 2023 describes the planting situation of different plots in 2023, including planting area, season, yield per mu, cost and selling price per unit, etc.

2.2. Agricultural production efficiency model

2.2.1. Decision variables of the model

In actual production and life, the planting cost, per mu yield and selling unit price of crops are different in different plots and seasons, so this paper sets decision variables $x_{i,j,k,t}$. Where i represents the cultivated land type, j represents the crop category, and k represents the planting season. t is the year of planting. Therefore, the decision variables represent the planting area of crop j in plot i in the k quarter of year t.

2.2.2. Constraints on crop cultivation

In agricultural production and life, crop planting decisions often take into account many factors, including planting season, cultivated land area, expected sales and planting rules. Accordingly, this paper makes the following constraints on the agricultural production benefit model:

(1) According to the literature, different types of cultivated land have different abilities to grow crops. After the investigation of the research objects in this paper, we found that the flat dry land,

terrace and slope land in this mountain area can only grow vegetables for one quarter, so the type of cultivated land that can only grow vegetables for one quarter is set S, and the following constraints are made:

$$x_{i,j,2,t} = 0, i \in S \tag{1}$$

(2) Since the total planted area of each plot i in each quarter cannot exceed the available area of the plot, the following constraints are determined:

$$\sum_{j \in R_i} x_{i,j,k,t} \le A_i, \forall i,k,t \tag{2}$$

Where, A_i represents the total plantable area of plot i, and R_i represents the number set of crops suitable for planting in plot i.

(3) Since each plot has corresponding crops suitable for planting, in order to ensure that plot i can only plant crops in the set, the following constraints are established:

$$x_{i,j,k,t} = 0, \forall j \notin R_i, \forall i,k,t \tag{3}$$

(4) In agricultural production and life, in order to ensure the growth quality of crops, each crop cannot be continuously planted in the same plot, and in order to maintain the nitrogen content, pulses should be planted at least once every three years in each plot. Therefore, the following constraints are established:

$$x_{i,j,k,t-1} \cdot x_{i,j,k,t} = 0, \forall i, j \in R_i, k, t \tag{4}$$

$$\sum_{t}^{t+2} \sum_{j \in \text{legume}} x_{i,j,k,t} > 0, \forall i,k \tag{5}$$

(5) As can be seen from the supply relationship, the total crop output exceeding the expected sales volume of each season should be within the controllable range; otherwise, the excess volume cannot be sold normally. Therefore, a controllable yield factor $\alpha_j(t,k)$ is set in this paper to represent the maximum acceptable range of the total output of crop j exceeding the expected sales volume $S_j(t,k)$ in the k quarter of year t. Accordingly, the following constraints are established:

$$\sum_{i} x_{i,j,k,t} \times Y_{i,j,k,t} \le S_j(t,k) \times (1+\alpha_j(t,k)), \quad \forall j,k,t \tag{6}$$

2.2.3. The objective function with uncertain conditions is introduced

As crops are subject to a series of influences during planting, growth and sales, such as climate, market and other factors, various indicators of crops are constantly changing. Therefore, combined with the relevant experience of actual agricultural production, this paper introduces some random variables to reflect the relevant changes of crops, as shown below:

(1) The sales volume of crops is often affected by market demand, supply, weather conditions and other factors, so the expected sales volume will fluctuate. In this paper, D_j is introduced to represent the change rate of the expected sales volume of crop j over time.

(2) The per mu yield of crops is often affected by climate and other factors, so the introduction of q_j represents the change rate of the per mu yield of crop j over time.

(3) Due to the influence of market conditions, the planting cost will also change constantly, so c_j is introduced to represent the changing rate of the planting cost of crop j over time.

(4) The unit price of sales changes with the changes of market supply and demand, and the degree of change of different types of crops is also different. Therefore, p_j is introduced in this paper to represent the change rate of the unit price of sales of crop j.

In actual agricultural life, the above random variables have a rough value range. In order to simplify the establishment of the model, the above random variables are taken as fixed values in this paper. Therefore, we can establish the objective function. Assuming that t_0 is the initial year of the agricultural production benefit model in this study, the formula for calculating the total output Q_{i,j,k,t_0} of crop j in plot i in the k quarter of year t_0 is as follows:

$$Q_{i,j,k,t_0} = x_{i,j,k,t_0} \times Y_{i,j,k,t_0} \tag{7}$$

Since the yield per acre of the crop changes from year to year, the expression for the total output in the future year t is as follows:

$$Q_{i,j,k,t} = Q_{i,j,k,t_0} \times (1+q_j)^{t-t_0} \tag{8}$$

C_{j,k,t_0} and P_{j,k,t_0} are introduced to represent the planting cost and selling unit price of crop j in the k quarter of year t_0 respectively, then the expression of planting cost and selling unit price of year t in the future is as follows:

$$C_{j,k,t} = C_{j,k,t_0} \times (1+c_j)^{t-t_0} \tag{9}$$

$$P_{j,k,t} = P_{j,k,t_0} \times (1+p_j)^{t-t_0} \tag{10}$$

By the same token, the expected sales in the initial year t_0 is S_{j,t_0}, so the formula for calculating the expected sales in the future year t is as follows:

$$S_{j,t} = S_{j,t_0}(1+D_j)^{t-t_0} \tag{11}$$

In actual agricultural production and sales, the part exceeding the expected sales volume cannot be sold normally, and generally needs to be sold at a reduced price. Therefore, this paper introduces discount coefficient m to reduce the price of the excess part. The value range of m is 0,1. Based on the above content, we can list the expression of sales $M_{j,k,t}$ and total cost $T_{j,k,t}$ of crop j in the agricultural production efficiency model, as follows:

$$M_{j,k,t} = P_{j,k,t} \times \min(\sum_i Q_{i,j,k,t}, S_j(t,k)) + m \times P_{j,k,t} \times \max(\sum_i Q_{i,j,k,t} - S_j(t,k), 0) \tag{12}$$

$$T_{j,k,t} = C_{j,k,t} \times \sum_i Q_{i,j,k,t} \tag{13}$$

Therefore, the objective function of introducing random variables is as follows:

$$F = \max \sum_t \sum_j \sum_k (M_{j,k,t} - T_{j,k,t}) \tag{14}$$

2.3. Particle swarm optimization combined with Monte Carlo simulation

The model in this paper has several random variables, and each random variable has a value range in the actual agricultural production experience. In order to make the objective function as large as possible, it is necessary to optimize the model parameters through continuous attempts, so this paper uses Monte Carlo simulation to simulate the value of random variables, and obtains a relatively ideal value through continuous optimization [6]-[8]. At the same time, when solving the model of agricultural production efficiency, if the traditional traversal method is used, it has a very high time cost. In order to improve the solving efficiency, particle swarm optimization (PSO) is used to solve the model [9]-[11]. Therefore, the specific solving steps of the model in this paper are as follows:

Step1: Monte Carlo algorithm simulates random variables

Set the number of iterations T, and use Monte Carlo simulation to randomly sample the values of random variables.

Step2: Use particle swarm optimization to solve the planting scheme

For the random variables determined by the current Step1, this paper adopts particle swarm optimization algorithm to solve. For the planting scheme obtained by a one-time simulation, this paper calculates its total production benefit. If it is greater than the current optimal total production benefit, it replaces it and records the planting scheme under the current optimal total production benefit.

Step3: Iterative optimization

The operation of Step1 and Step2 was repeated to get the optimal planting plan after iteration.

3. Result and Discussion

3.1. Data Preprocessing

After the natural connection of the planting situation and related statistical data of the village's crops in 2023, this paper finds abnormal data that are not easy to be found, such as the same food crops, the category is close, and most of the unit sales price is less than 10 yuan, while some crops have a unit sales price as high as 50 yuan, which obviously does not conform to the general pricing law of food crops. Therefore, after browsing all kinds of data, this paper respectively conducted box diagram inspection on the per mu yield, planting cost and unit sales.

(1) Per mu yield inspection

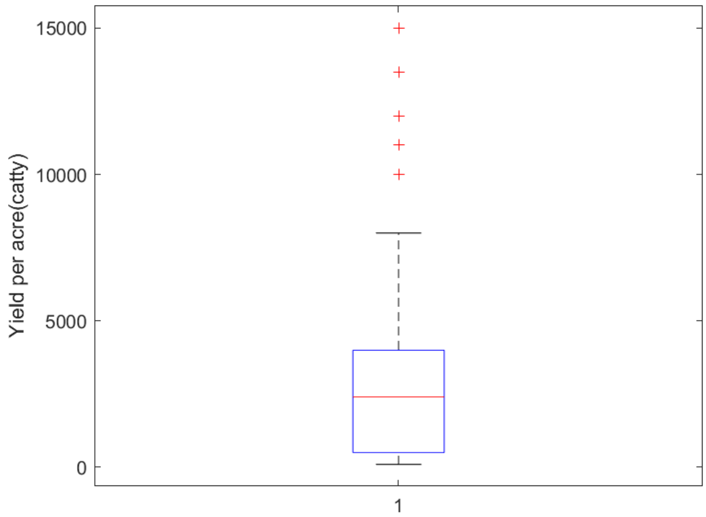

Figure 1. Box plot of yield per acre

Table 1. Yield outliers per acre

Crop	Yield per acre (catty)
Cucumber	12000
Water spinach	10000
Cucumber	15000
Water spinach	12000
White spirit mushroom	10000
Cucumber	13500
Water spinach	11000

Figure 1 is a box plot test of yield per mu, and the outliers derived from Figure 1 are shown in Table 1. The analysis shows that cucumber and water spinach belong to economic crops and are common vegetables. White spirit mushroom is similar to shiitake mushroom and elm mushroom, which are common edible fungi. According to the yield per mu of other similar crops and the actual agricultural production, the paper thinks that the normal yield per mu of cucumber and water spinach is in the range of 5000 to 8000 jin, and the yield of white spirit mushroom is in the range of 4000 to 5000 jin, so the paper takes half of all the yield per mu.

(2) Planting cost inspection

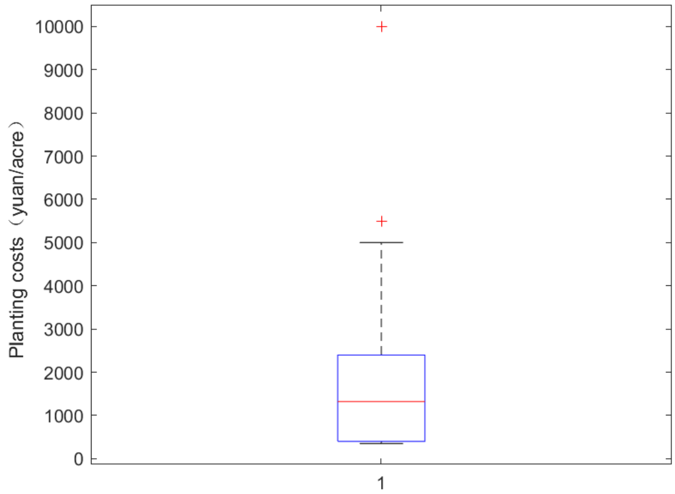

Figure 2. Box plot of Planting costs

Table 2. Planting cost outlier results

Crop	Yield per acre (catty)
Morel	10000
Water spinach	5500
White spirit mushroom	10000

Figure 2 is a box plot test of planting costs, and the outliers derived from Figure 2 are shown in Table 2. Morel is a rare fungus, the planting cost is higher than other fungi in a reasonable range, so there is no need to do outlier processing, and white spirit mushroom belongs to a more common edible fungus, the planting cost is similar to mushroom, so its planting cost is divided by 2 to maintain within a normal range. Water spinach is built in intelligent greenhouses, so the cost is higher than that of ordinary greenhouses, and there is no additional processing.

(3) Sales unit price inspection

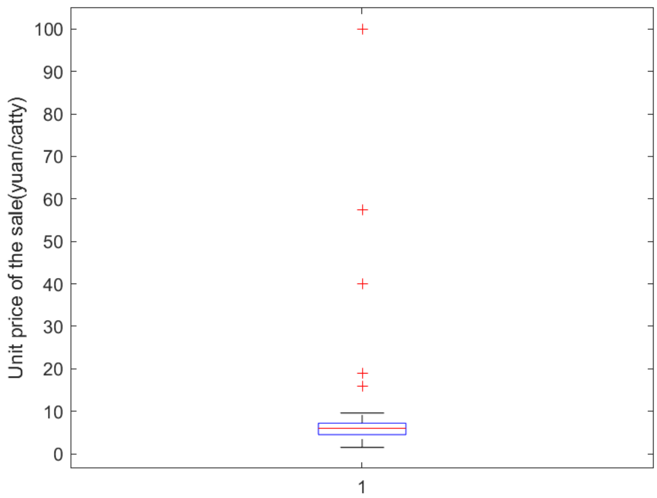

Figure 3. Box plot of Unit price of the sale

Table 3. Sales unit price outlier result

Crop	Yield per acre (catty)
Buckwheat	40
Elm mushroom	57.5
Lentinula edodes	19
White spirit mushroom	16
Morel	100

Figure 3 is a box plot test of the sales unit price, and the outliers derived from Figure 3 are shown in Table 3. Although the actual price of buckwheat may fluctuate due to changes in varieties, processing methods and market supply and demand, the price of 40 yuan/jin is much higher than the 3-5 yuan/jin in the information, so its price is adjusted to one-tenth of the normal value.

3.2. Optimal planting scheme

Through the above solution steps, this paper obtains the planting scheme that makes the annual agricultural production of 2024 to 2030 the most profitable. Taking 2024 as an example, the number of planting schemes is huge, which is visualized in this paper as follows:

(1) Black and white visualization

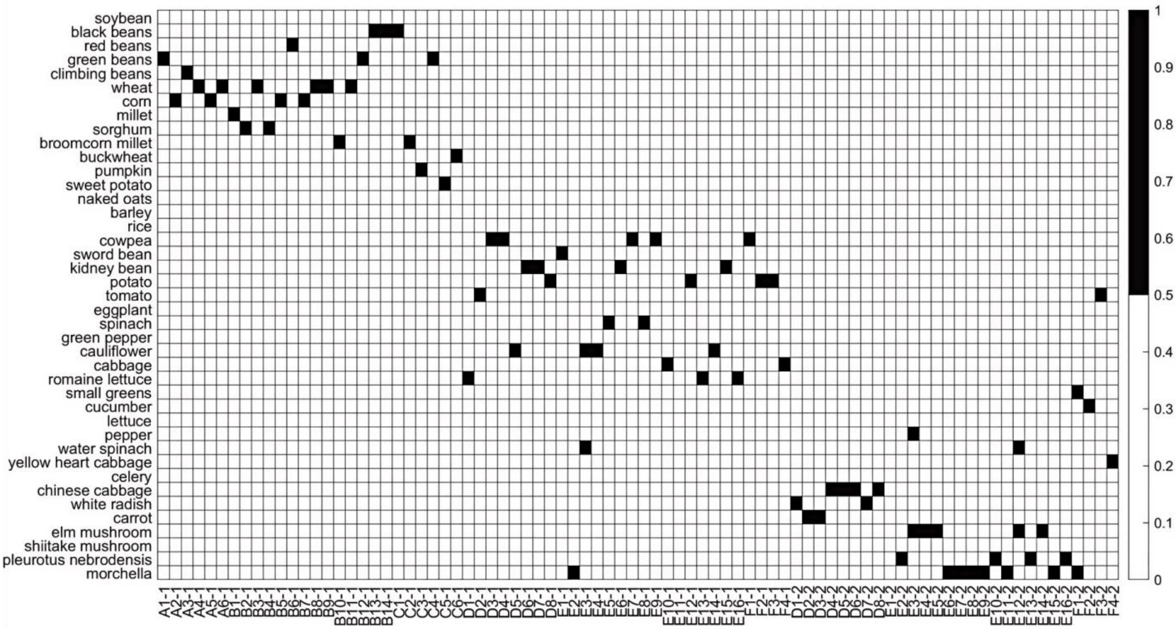

Figure 4. Plots taken for the optimal planting plan in 2024

The horizontal coordinate of Figure 4 reflects the planting situation of different types of plots in a certain quarter. For example, A1-1 represents the planting situation of flat dry land 1 in the first quarter. Through the analysis of the solution results, it can be seen that the planting area of different types of plots is significantly different. Large plots, such as 86 mu of soybeans planted on terraces and 80 mu of wheat planted on hillsides; On smaller plots, such as ordinary greenhouses, only 0.3 mu of pepper and water spinach were planted. Due to the large difference between the planting area of 0.3 mu and 80 mu, it is difficult to distinguish the color from the background color of 0 mu in the visualization. Therefore, this paper first carries out black and white visualization, and adopts the 0-1 classification method to unify the non-zero values into 1, thus generating a highly distinguishable black and white view.

By looking at Figure 4, you can clearly see the crops grown in each plot and their planting season. Crop planting in 2024 is roughly distributed from top left to bottom right. By comparing the X-axis and Y-axis in the image, it can be confirmed that the solution results conform to the constraint relationship of crop, cultivated land type and season.

(2) Gradient visualization

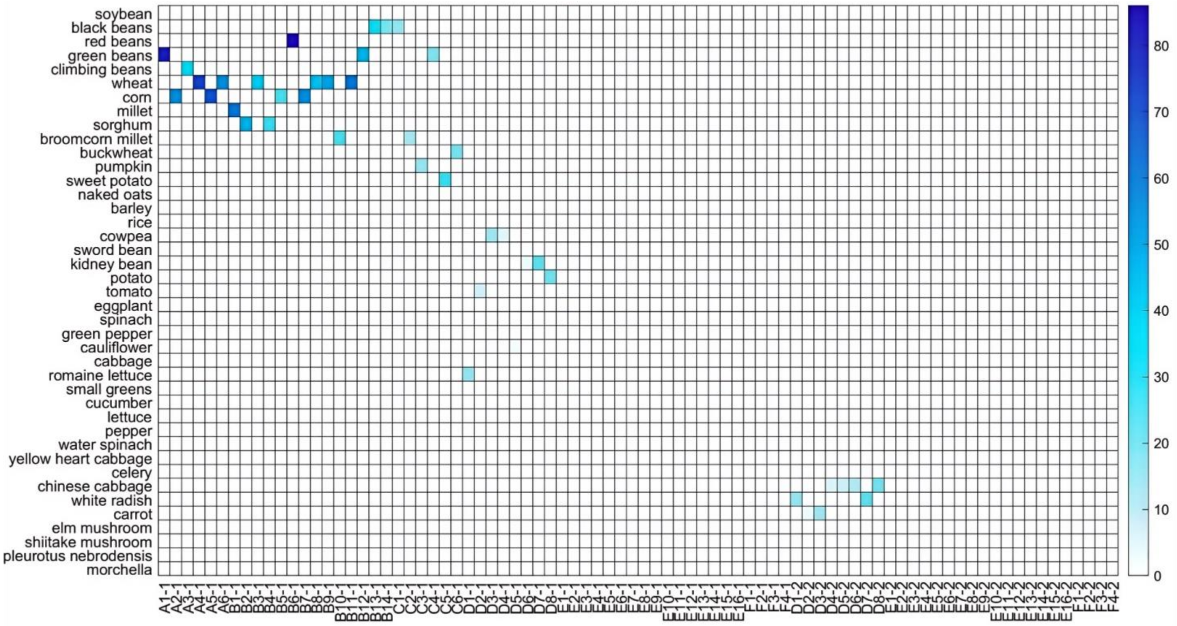

Figure 5. The amount of planting in each plot of the optimal scenario in 2024

By looking at Figure 5, the paper finds that the planting of crops in 2024 goes from the top left to the bottom right, and the color goes from dark to light, indicating a gradual decrease in planted area. Comparing the X axis and Y axis of the image, it can be seen that the darker color areas are concentrated in flat dry land, terraces, hillsides and irrigated land, while the color of the area representing the greenhouse is almost the same as the background color, which is difficult to distinguish. This is because the greenhouses can be planted in an area of 0.6 acres, while the area of open farmland is tens of acres, the difference is large. This further verifies the reliability of the solution results.

3.3. Optimal planting scheme

According to the above optimal planting scheme, this paper calculates the annual production benefit, as shown in the figure below:

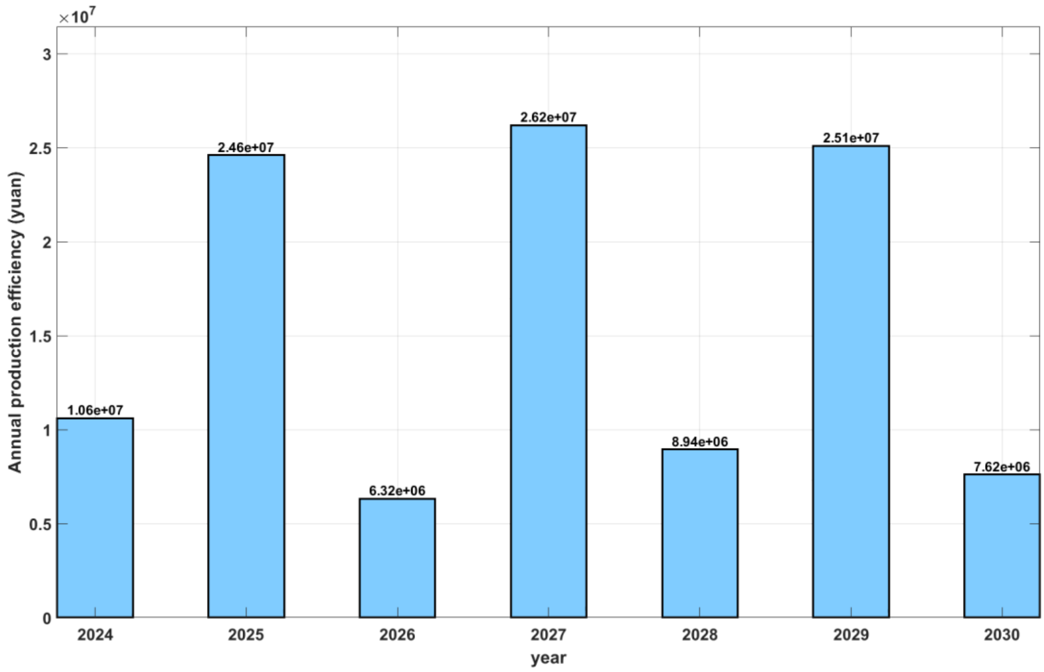

Figure 6. Annual production efficiency

By observing the images, this paper finds that there are large differences in the solution data in some years, which indicates that after adding uncertainty conditions, the model shows instability under the influence of random variables, resulting in large fluctuations in annual production benefits.

4. Conclusion

Based on Monte Carlo simulation and particle swarm optimization algorithm, a scientific planting strategy optimization method was proposed to maximize the economic benefits of agricultural production in North China. The innovation of this paper lies in the introduction of random variables, the optimization of the production benefit model, and the proposed planting scheme adapted to the geographical characteristics of the mountainous areas in North China, which provides theoretical support for the sustainable development of agriculture. However, the limitation of the study is that the data is only based on a specific region, which may not fully represent the whole region of North China, and environmental factors such as climate change are not fully considered. Future research can expand the scope of data and combine climate models to optimize planting strategies to cope with ecological protection and climate change, and promote agricultural economic and ecological sustainable development.

References

[1] Huang P. The peasant economy and social change in North China [M]. Stanford University Press, 1985.

[2] Yang G, Li X, Huo L, et al. A solving approach for fuzzy multi-objective linear fractional programming and application to an agricultural planting structure optimization problem [J]. Chaos, Solitons & Fractals, 2020, 141: 110352.

[3] Wu L, Tian J, Liu Y, et al. Multi-Objective Planting Structure Optimisation in an Irrigation Area Using a Grey Wolf Optimisation Algorithm [J]. Water, 2024, 16 (16): 2297.

[4] Wang F, Dong Z, Wu Z, et al. Optimization of maize planting density and fertilizer application rate based on BP neural network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2017, 33 (6): 92 - 99.

[5] Ye M, Wang L, Zhang H. Research on multi-objective optimization model of the planting structure based on TOPSIS [C]//Journal of Physics: Conference Series. IOP Publishing, 2021, 1848 (1): 012109.

[6] Liu J, Li Y P, Huang G H, et al. Assessment of uncertainty effects on crop planning and irrigation water supply using a Monte Carlo simulation based dual-interval stochastic programming method [J]. Journal of Cleaner Production, 2017, 149: 945 - 967.

[7] Maa H. Optimization of Crop Planting Strategies Based on Linear Programming and Monte Carlo Simulation [J]. 2024.

[8] Yang G Q, Li M, Guo P. Monte Carlo-Based Agricultural Water Management under Uncertainty: A Case Study of Shijin Irrigation District, China [J]. Journal of Environmental Informatics, 2022, 39 (2).

[9] Liu L, Chen T, Gao S, et al. Optimization of agricultural machinery allocation in heilongjiang reclamation area based on particle swarm optimization algorithm [J]. Tehnički vjesnik, 2021, 28 (6): 1885 - 1893.

[10] Kaleeswaran V, Dhamodharavadhani S, Rathipriya R. Multi-crop selection model using binary particle swarm optimization [C]//Innovative Data Communication Technologies and Application: Proceedings of ICIDCA 2020. Springer Singapore, 2021: 57 - 68.

[11] Mythili K, Rangaraj R. Deep learning with particle swarm based hyper parameter tuning based crop recommendation for better crop yield for precision agriculture [J]. Indian Journal of Science and Technology, 2021, 14 (17): 1325 - 1337.