Optimisation of crop planting strategies based on Monte Carlo algorithm

Yan Guo[1, *, #], Zhilin Su[1, #], Xize Lin[1, #]

[1]Faculty of Economics, Northwest University of Political Science and Law, Xi'an, China

*Corresponding author: 3091326193@qq.com

[#]These authors contributed equally to this work

Keywords: Monte Carlo; 0-1 planning; Binary trees; Planting strategy optimisation

Abstract: With the continuous development and progress of scientific knowledge and intelligent technology, the combination of agriculture and science and technology is one of the development trends of agricultural digitisation and intellectualisation. As an important primary industry of the country, the production of crops is one of the most important weights affecting the country's GDP, and dealing with the cultivation of all kinds of agricultural products plays a pivotal role in maximising the economic benefits, and is an issue that must be considered by the country's macro-control. It is an issue that must be considered by the national macro-control. In this paper, based on the sales data of some crops in 2023, firstly, the outliers are identified through box-and-line diagrams, and the identified outliers are rationalised. Then the optimisation problem based on static selling price and sales volume is considered, and 0-1 variables are introduced to constrain whether the crops are planted in the corresponding plots, and the optimal binary search tree is used to solve the corresponding model. Finally, a planning model based on Monte Carlo algorithm is developed to optimise the planting strategy of agricultural products for the period of 2024-2030, taking into account the effects of the market and the natural environment, so as to maximise the profit in the next seven years.

1. Introduction

With the requirements of sustainable development of the world economy, the rational use of arable land is becoming more and more important, how to make efficient use of limited arable land resources in order to rationally develop organic agriculture and enhance food production is the core issue to ensure sustainable development of rural economy. Three rural issues as the basis of people's livelihood, the excessive use of chemical products such as fertilisers and other chemical products will inevitably affect the health of people's livelihood, therefore, only reasonable and healthy planting can solve the problem of maximising the utility of land resources and food security [1].

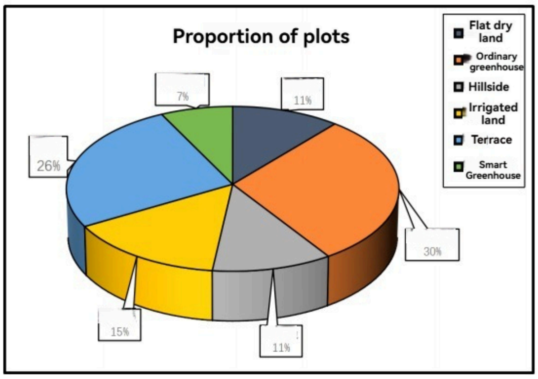
To produce organic and high-yield food, it is necessary to combine the traditional crop growth rules, use the knowledge of management, economics, agriculture and other disciplines, reasonably plan the area and types of crops to be cultivated on different arable land, combine the climate and the suitability of crops to the arable land, and adhere to the principle of not wasting food, so as to analyse the cultivated land in the rural areas of North China from 2024 to 2030, and give the optimal cultivation scheme for the crops. The optimal crop cultivation plan will be given [2].

2. Data preprocessing

2.1 Data visualisation

By analysing the number of parcels of the six major lands and the actual number of acres of different land types, as shown in Figures 1-2, the concentration of that land can be visualised. The higher the concentration of land, the easier it is for farmers to manage, and it also serves as a guide for crop cultivation on that type of land.

  DOI: 10.25236/icemeet.2024.032



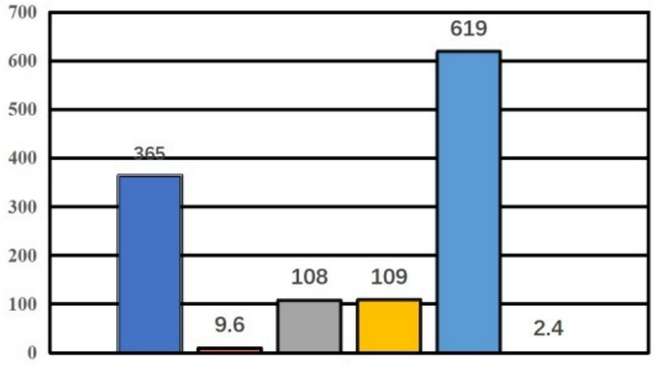


Fig. 1 Number of plots as a percentage Fig. 2 Acres of different plots

The common shed has a larger number of parcels, but its overall acreage is smaller, suggesting that the common shed land is more dispersed overall. In contrast, arid lands occupy fewer parcels, but their overall acreage is larger, indicating that arid lands are more concentrated. In the acres of land in the region, terraces accounted for a larger proportion of 619 acres of land, followed by arid land of 365 acres of land, meaning that food crops are grown in a larger amount in the countryside of North China, while in contrast, due to financial and technological constraints, greenhouses occupy fewer acres, and the production of some crops is affected by the climate and other seasonal or technological influences.

2.2 Outlier analysis

In this paper, through the box-and-line diagram to reflect the distribution of data intuitively, portraying the relevant characteristics of the data, mu yield, planting costs, unit price of sales of the upper quartile, median, and the lower quartile for Q_1, Q_2, Q_3, its quartile spacing as $IQR = Q_1 - Q_3$, by observing the distribution of values beyond the defined range of the box-and-line diagram to determine whether the data exist in the state of abnormality, remembering that The interval of normal distribution is: $[Q_3 - 1.5IQR, Q_1 + 1.5IQR]$.

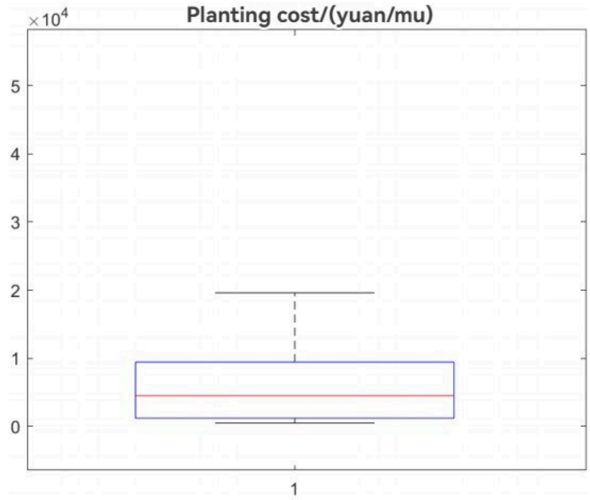


Fig. 3 Cost of planting (yuan/acre)

The Figure 3 shows that there are no outliers in the cost of planting/($/acre). Also the median in the box and line plot is closer to the lower quartile of the data, indicating some skewness distribution of the data [3].

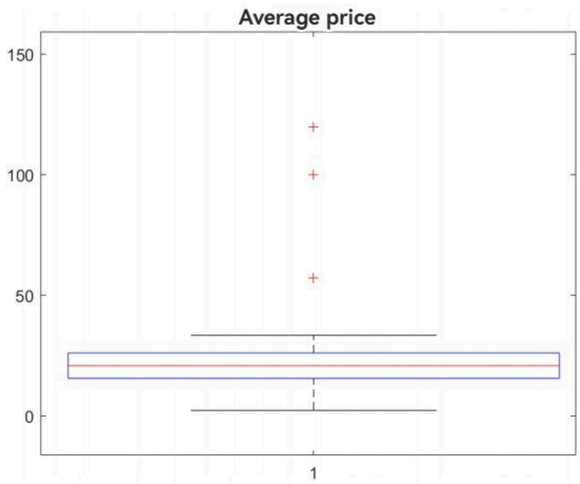


Fig. 4 Average price

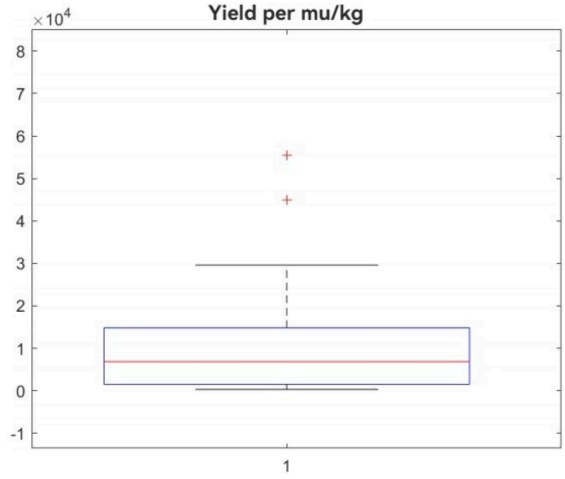


Fig. 5 Yield per acre/kg

Table. 1 Table of types of unit price outliers

Assortment	Average sales unit price
Morel mushrooms	100
Edible mushroom	57.5
Buckwheat	120

Through the Figures 4-5, this paper finds that there are three outliers in the unit sales price as follows. However, this paper does not consider this as a special outlier as a treatment. As shown in Table 1, because the "morel mushrooms" and "elm mushrooms" belong to rare mushrooms, "buckwheat" belongs to rare herbaceous plants, high nutritional value, supply is small, the market demand is large, so this paper believes that should not be excluded as outliers.

2.3 JB test to validate data distribution characteristics

Applicable conditions: test statistical samples for large samples (n>30) can be used, and the more data the better the test, mu yield, unit price, and planting costs of the number of samples are greater than 30 sample size, so this paper adopts the JB test to verify the characteristics of the data distribution.

Step1: First of all, the establishment of the original hypothesis and alternative hypothesis: H_0: the variable obeys the normal distribution, H_1: the variable does not obey the normal distribution.

Step2: Construct the test statistic: for any random variable X_i with kurtosis K and skewness S, construct the test statistic as a chi-square distribution with two degrees of freedom.

Step3: Calculate JB by substituting the data, and calculate the corresponding P-value through JB.

Step4: Draw a conclusion through the P-value test. (Generally take 0.05)

Table. 2 JB test p-value table

	Yield per acre	Cost of cultivation	Average sales price
Yield per acre	1	0	0
Cost of cultivation	0	1	0
Average sales price	0	0	1

As shown in Table 2, the h return value of mu yield, planting cost, and average sales price are all 1, and the p value returned is very small and basically 0, which is less than 5% of the significance level, rejecting the original hypothesis. As shown in Table 3, the distribution of the data of mu yield, planting cost, and average sales price do not conform to the normal distribution.

Table. 3 Normal distribution test table

	Yield per acre	Cost of cultivation	Average sales price
Whether or not it conforms to a normal distribution	no	no	no

3. Research on Crop Planting Strategies

3.1 Modelling

The optimal crop planting strategy is determined by the following equation:

$$MAXprofit = X_{ijky}Q_{ijky}P_{ijky} - X_{ijky}Q_{ijky}C_{ijky} \tag{1}$$

Where i stands for crop i, j stands for land j, k stands for crop k, y stands for year y, X_{ijky} stands for acreage of crop j in land i in season k in year y, C_{ijky} stands for cultivation cost of crop j in land i in season k in year y, Q_{ijky} stands for acreage production of crop j in land i in season k in year y, P_{ijky} stands for selling price of crop j in land i in season k in year y, where X_{ijky} is the core decision of the model. Q_{ijky} represents the acre yield of crop j in plot i in season k in year y, and P_{ijky} represents the selling price of crop j in plot i in season k in year y, where X_{ijky} is the core decision variable of the model.

3.2 Constraints and Objective Functions

This cropping decision needs to satisfy the following constraints:

(1) One season of food crops per year is grown on dry land, terraces, and slopes, and rice is excluded;

$$\begin{gathered} X_{ij1y} > 0 i \in (A_1, A_2, \dots, B_{14}, \dots, C_6) j \in (1,2 \dots ,15) y \in (1,2 \dots ,7) y, j \in z \\ else \;\; X_{ijky} = 0 \end{gathered} \tag{2}$$

(2) Watered land is suitable for growing one season of rice and two seasons of vegetables. If you choose to grow two seasons of vegetables, you cannot grow cabbage, carrot and white radish in the first season, and you can only grow three kinds of vegetables, namely, cabbage, carrot and white radish, in the second season;

When $i \in (D_1, D_2, \dots, D_8) j = 16$, there is

$$X_{i_{j1y}} > 0 X_{i162y} = 0 else X_{ijky} = 0 \tag{3}$$

When $i \in (D_1, D_2, \dots, D_8) j \in (17,18 \dots ,34)$,

$$X_{i_{j1y}} > 0 \tag{4}$$

When $j \in (35,36,37)$,

$$\begin{gathered} X_{i_{j2y}} > 0 \\ else \;\; X_{ijky} = 0 \end{gathered} \tag{5}$$

(3) Ordinary greenhouses can grow vegetables in the first season and edible mushrooms in the second season;

$$\begin{gathered} X_{ij1y} > 0 i \in (E_1, E_2, \dots, E_{16}) j \in (17,18 \dots ,34) y \in (1,2 \dots ,7) y, j \in z \\ else \;\; X_{ijky} = 0 \end{gathered} \tag{6}$$

$$X_{ij2y} > 0 i \in (E_1, E_2, \dots, E_{16}) j \in (38,39 \dots ,41) y \in (1,2 \dots ,7) y, j \in z \tag{7}$$

(4) Smart greenhouses are suitable for growing two seasons of vegetables per year (except cabbage, carrot and white radish);

$$\begin{gathered} X_{ijky} > 0 i \in (F_1, F_2, \dots, F_4) y \in (1,2 \dots ,7) y, j \in z k \in (1,2) \\ else \;\; X_{ijky} = 0 \end{gathered} \tag{8}$$

(5) No heavy cropping;

When $i \in (A_1, A_2, \dots, E_{16})$,

$$X_{ijky}X_{ijk(y+1)} = 0 \tag{9}$$

When $i \in (F_1, F_2, \dots, F_4)$,

$$X_{ij1y}X_{ij2y} = 0 \tag{10}$$

$$X_{ij2y}X_{ijk(y+1)} = 0 \tag{11}$$

(6) Legume rotation constraints;

$$\sum_{m=0}^{2} X_{ijk(y+m)} \geq 1 j \in (1,2,3,4,5,17,18,19) \tag{12}$$

(7) Total area constraints.

$$\sum_{i=1}^{41} X_{ijky} \leq 1213 k = 1 \tag{13}$$

If the total production per season is higher than the expected sales, the grain in excess of the expected sales is stagnant and wasted with no gain. The objective function is as follows.

$$MAXprofit = \sum_{i=1}^{54} \sum_{j=1}^{41} \sum_{y=1}^{7} \sum_{k=1,2} X_{ijky}Q_{ijky}P_{ijky} - X_{ijky}Q_{ijky}C_{ijky} - max\{Q_{ijky}X_{ijky} - f_{ijky}, 0\}(C_{ijky} + P_{ijky}) \tag{14}$$

In this paper, due to the excessive number of constraints and the complexity of some of the constraints, the 0-1 planning algorithm is used to introduce (0, 1) planning variables to speed up the solution, and the dynamic planning algorithm is introduced into the overall optimisation process and the optimal binary search tree is used to optimise the final planting scheme. An optimal binary search tree is a sequence $K = (k_1, k_2, k_3, \dots, k_n)$, where $k_1 < k_2 < k_3 < \cdots < k_n$, use these elements to construct a search binary tree, and for each keyword k_i being looked for, construct n+1 dummy keywords, $d_1, d_2, d_3, d_4, \dots, d_n$, where $d_0 < k_1, d_n > k_n, k_i < d_i < k_{(i+1)}$ where $i = 1,2, \dots, 1 - n$, for each keyword di, the probability of finding it is qi. for a given binary search tree, and define the cost of a search. For a given set of probabilities, construct a binary tree that is expected to minimise the cost of a search, this tree is called an optimal binary search tree [4].

3.3 Planting strategy

Based on the above traditional planning solution model, this paper solves the recommended planting plan for crops in the village in 2024-2030, and the specific planting results in 2024 are shown in the Figure 6. Seven-year expected sales profit is in Figure 7.

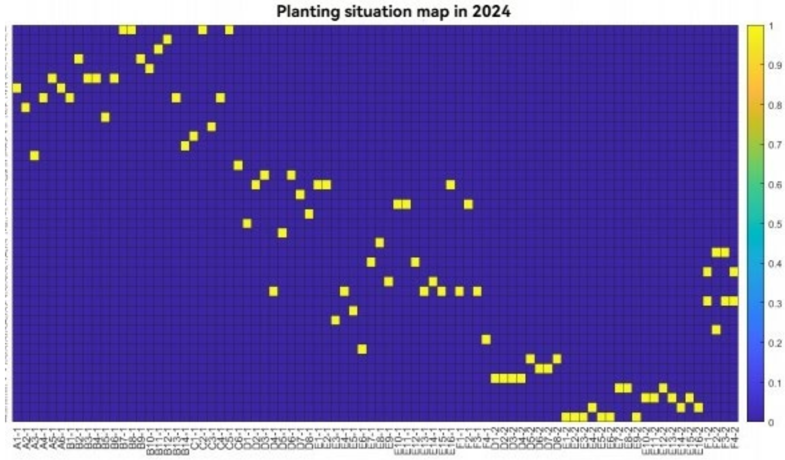


Fig. 6 Map of plantings in 2024

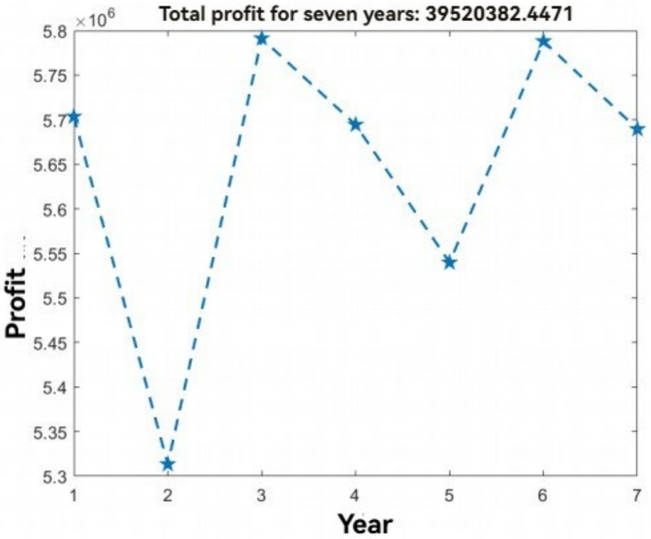


Fig. 7 Seven-year expected sales profit chart

4. Optimisation of planting strategies under market volatility

Based on farming experience, future sales of wheat and corn are expected to increase at an average annual growth rate of 5 to 10 per cent, except for other crops where future sales are expected to fluctuate up or down by about 5 per cent per year compared to 2023, crop yields will fluctuate up or down by 10 per cent per year due to climatic influences, and the cost of crops, taking into account market conditions, will increase by an average of about 5 per cent per year. The cost of growing crops, taking into account market conditions, has been increasing by an average of about 5 per cent per year. The selling price of food crops will remain stable, the selling price of vegetable crops will increase at an average rate of about 5% per year, and the selling price of edible mushrooms will decline slightly by 1% to 5% per year, with the selling price of morel mushrooms declining more sharply by 5% per year.

Monte Carlo simulation is a stochastic simulation method, which is based on probability and statistical theoretical methods and relates the problem to be solved to a certain probability model. An electronic computer is used to generate random numbers obeying a certain distribution to achieve statistical simulation or sampling, and when the sample capacity is large enough, i.e., when the number of simulations is sufficiently high, the mean of the generated random numbers will converge to the expected value, thus obtaining an approximate solution to the problem. In asset valuation, Monte Carlo simulation based on historical data using computer simulation can to a certain extent exclude the influence of subjective factors on asset valuation, thus making the valuation results more scientific and reasonable [5].

In order to consider the expected sales volume, mu yield, planting cost, uncertainty of selling price and potential planting risk of each type of crop, the planning model in the previous section was adjusted and the objective function was modified to the following form.

$$MAXprofit = \sum_{i=1}^{54}\sum_{j=1}^{41}\sum_{y=1}^{7}\sum_{k=1,2}\big[(IntP_{ijky} \times IntQ_{ijky} \times X_{ijky} - IntC_{ijky} \times IntQ_{ijky} \times X_{ijky}) \\ -W max\{IntQ_{ijky} \times X_{ijky} - Intf_{ijky}, 0\} \times (IntC_{ijky} + IntP_{ijky})\big] \tag{15}$$

Through the above solution method in this paper, the recommended planting plan of crops under the conditions of crop yield, cost and price fluctuation is solved, and the specific results are shown in the Figure 8. Seven-year expected sales profit is in Figure 9.

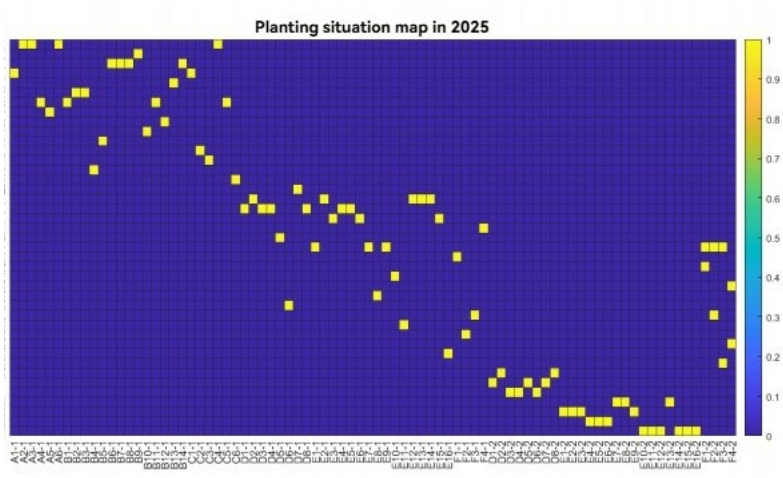


Fig. 8 Map of plantings in 2025

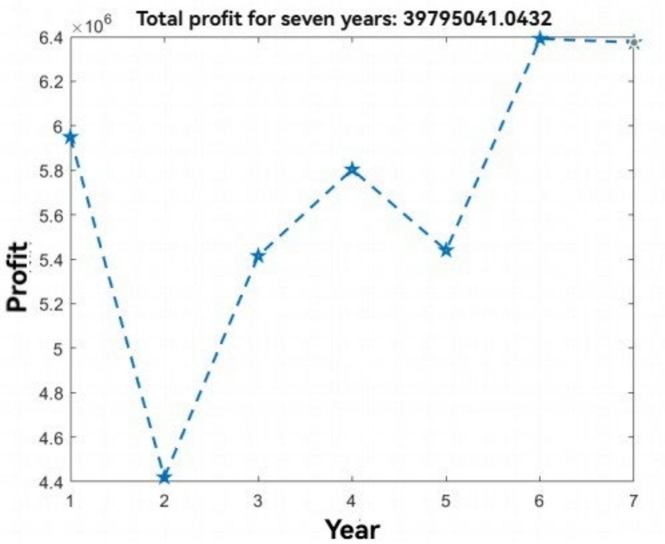


Fig. 9 Seven-year expected sales profit chart

5. Summary

In this paper, we first consider an optimisation problem based on static selling price and sales volume, assuming that the acre yield in 2023 is its sales volume. Two different optimisation objective functions are constructed by discussing whether to sell crops that exceed the expected sales volume. By introducing 0-1 variables to constrain whether the crops are planted in the corresponding plot or not, and by introducing the optimal binary search tree to solve the corresponding model, if all the products exceeding the expected sales volume are not sold, the total profit in seven years will be: 395,203,382.4471 Yuan.

Due to the expected future sales of wheat and corn have a growing trend, by the market and the natural environment, the mu yield and planting costs of crops and sales prices have interval fluctuation characteristics. Using Monte Carlo simulation to select random values in the interval values of the above variables in order to update the planning model and optimise the agricultural planting strategy for the years 2024-2030, the seven-year profit can be obtained as follows: 39795041.0432 Yuan.

References

[1] WU Jiuniu, GAO Decheng, JIANG Weidong, et al. Application of interpolation method based on box-and-line diagram in data processing of empty box barometer[J]. Industrial Instrumentation and Automation Device,2023(3):93-98.

[2] Coleman. Chapter 14: Dynamic programming algorithms[M]. U.S.: Mechanical Industry Press, 2013.

[3] Yang Jixiang and Ling Ling. A 0 - 1 integer programming solution model for the multicore cluster task allocation problem[J]. High Technology Newsletter,2016.

[4] J.H. Liao, S. Liu, M.V. Ma, et al. A binary tree-based synthesis algorithm for Verilog multiplexed branching statements[J]. Journal of Ningbo University (Science and Technology Edition), 2024,37(2):10-17.

[5] ZHANG Long,LIU Yaozong. Reliability prediction of measurement and control devices for aviation based on Monte Carlo simulation[J]. Electrical Measurement and Instrumentation, 2024,61(6):188-193,217.