Optimal Crop Planting Strategy for a Rural Village Considering Uncertain Factors

Yunze Wang [1, #], Jinyi Zhu [1, #], Yutong Jiao [2, *, #]

[1] School of Communication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an, China, 710121

[2] School of Computer, Xi'an University of Posts and Telecommunications, Xi'an, China, 710121

* Corresponding author: yujiao@stu.xupt.edu.cn

[#]These authors contributed equally.

Abstract. With the development of China's crop planting industry and rural economy, it is especially important to improve the planting efficiency. This paper presents a genetic algorithm-based decision model to maximize profits in North China's mountainous crop planting. Through data preprocessing, details about the data can be obtained from the following source: agdata.cn/dataManual/dataTable/MTE5Njkx.html. the median floating unit price is determined, and the planting returns at different times and land types are analyzed to clarify the expected sales volume of each type of crop in 2023. Aiming to maximize profits from 2024 to 2030, an optimization model with 11 constraints is developed. Constraints are simplified using four methods, and a genetic algorithm optimizes the objective function, with dynamic programming enhancing the fitness function to find the best planting strategy. The results of the study provide a scientific basis and decision support for crop cultivation in the mountainous areas of North China, and provide a reference for solving similar problems in other regions, which has important application value.

Keywords: Multi-objective Combinatorial Optimization, Genetic Algorithm, Dynamic Programming Algorithm.

1. Introduction

With the rapid development of agricultural planting and rural economy, agricultural planting needs to be more efficient. Therefore, the selection of regional agricultural planting schemes should fully utilize the limited land resources [1] and choose crop types suitable for the land. This paper establishes an optimized decision-making model for planting crops in mountainous areas from 2023 to 2030, aiming to maximize profits [2].

In the past, most researchers have established a relatively fast and scalable greedy algorithm model [3], mainly by selecting the best choice at the moment at each step, but without considering the future impact, analyzing the planting strategy of crops in each year under different planting conditions [4], planting according to the planting rules of the current year each year, optimizing the planting method to obtain maximum profits, and constantly adjusting the planting strategy [5] to obtain crops in different regions and different time periods.

The greedy algorithm model only considers the current optimal choice for each step, and cannot guarantee the global optimal solution [6]. It is also poorly adapted to a large number of constraints. After analyzing the results of the existing problems, we established a genetic algorithm model [7] that can guarantee the global optimal solution, and preprocessed [8] the information provided, determined the floating unit price interval as the median, and systematically analyzed the planting and income of various types of crops at different times and different land types. 11 constraints were obtained, and 7 constraints were simplified by dynamic programming [9], and feasible combinations (quarter, crop type, and land type) were selected. The genetic algorithm [10] was used to solve the objective function under the remaining three constraints, and the decision variables were encoded to solve the optimal planting strategy.

2. Modeling Multi-Conditional Optimization Problems

At present, there are two major problems in crop cultivation in the mountainous areas of North China: problem one is the waste of land for crop cultivation, and problem two is the low income generated by crops. To increase the planting profit rate of the difficult-to-cultivate land in the mountainous areas, through the data analysis on the above website, we have established a planting strategy model that can predict the planting strategy of crops in the mountainous areas for the next six years, maintaining the optimal planting decision for the maximum yield and the highest rate of return.

In reality, the total production per acre per year needs to be compared with the expected sales volume to ensure that the value of the different [9] crops produced each year is as small as possible, and for the portion of the actual production that exceeds the expected sales volume, the two most common methods used in the marketplace are set up, to sell all of them or to sell the excess portion of them at a 50% reduction in price.

It has been assumed that the expected sales yield is the value of the yield per acre in 2023. Have categorized the types of crops that will be available in different terrains and what crops will be grown in which plot in 23 years and extracted the data i.e. expected sales volume.

Establish the objective function of the optimization model, respectively, for the part of the expected sales volume exceeded and not exceeded the expected sales volume part of the sum of the most money, is the best economic efficiency.

To find out the profit, we need to set the expected sales volume as ZS_i , which means the value of the production of different crops in each piece of land given in the appendix in 2023, X_{eik} is the required decision variable, which means the area of that piece of land divided by that species in that season, ZO_{dik} means the production per acre, ZQ_{dik} is the amount of the price of the sales volume, which is the pricing that is taken as the middle of the interval, and ZP_{dik} is the amount of the sales volume, which is the pricing that takes the middle of the interval, and ZP_{dik} is the amount of the sales volume. ZP_{dik} is set as the required cost of each piece of land.

The profit of the over part is multiplied by the quantity and price of all the un-exceeded part, minus the money lost due to the unsold over part, minus the cost price contained in the original planting of the land, which is given in the following formula:

$$(\mathrm{ZS}_i - \mathrm{X}_{eik} \cdot \mathrm{ZO}_{dik}) \cdot \mathrm{ZO}_{dik} - \mathrm{ZP}_{dik} \tag{1}$$

Where, $e = 1, 2, ..., 14$, indicates that there are a total of 14 quarters for seven years after the prediction, $i = 1, 2, ..., 41$, represents a total of 41 crops, $k = 1, 2, ..., 54$, d denotes the quarter 1, 2 and denotes the numbering of each of the corresponding plot of land, numbered downwards from A1, that Contains barns.

When the production does not exceed the expected sales volume, the total profit value when it is not exceeded is obtained by multiplying the actual acreage of each future crop under different seasonal plots by the yield per acre and multiplying it by the selling price minus the cost of cultivation.

$$X_{eik} \cdot ZQ_{dik} \cdot ZQ_{dik} - ZP_{dik} \tag{2}$$

So Profit maximization is achieved through the objective function as shown in equations (1) and (2). Values that exceed the expected sales output are considered stagnant sales, as indicated in equation (3):

$$\mathrm{MAX}\begin{cases} \sum_{e=1}^{14}\sum_{i=1}^{41}\sum_{k=1}^{54} X_{eik} \cdot ZO_{dik} \cdot ZQ_{dik} - ZP_{dik}, \sum_{i=1}^{41}(\sum_{k=1}^{54} X_{eik} \cdot ZQ_{dik} - S_i) \le 0 \\ \sum_{e=1}^{14}\sum_{i=1}^{41}\left[(ZS_i - X_{eik} \cdot ZO_{dik}) \cdot \sum_{k=1}^{54} ZQ_{dik} - \sum_{i=1}^{54} ZP_{dik}\right], \sum_{i=1}^{41}(\sum_{k=1}^{54} X_{eik} \cdot ZQ_{dik} - S_i) > 0 \end{cases} \tag{3}$$

According to equations (1) and (2), the optimization objective function requires adjustment to more accurately reflect profitability calculations for sales exceeding the expected volume. Additionally, a 50% price reduction on the excess goods should be taken into account:

$$\sum_{k=1}^{54} X_{eik} \cdot ZO_{dik} - S_i \cdot \sum_{k=1}^{54} ZQ_{dik} \cdot (1-50\%) - \sum_{k=1}^{54} ZP_{dik}) \tag{4}$$

The simplification is then carried into the objective function as:

$$\begin{cases} \sum_{e=1}^{14}\sum_{i=1}^{41}\sum_{k=1}^{54} X_{eik} \cdot ZO_{dik} \cdot ZQ_{dik} - ZP_{dik}, \sum_{i=1}^{41}(\sum_{k=1}^{54} X_{eik} \cdot ZQ_{dik} - S_i) \le 0 \\ \sum_{e=1}^{14}\sum_{i=1}^{41}\left[\left(1.5S_i - 0.5\sum_{k=1}^{54} X_{eik} \cdot ZO_{dik}\right) \cdot \sum_{k=1}^{54} ZQ_{dik} - \sum_{i=1}^{54} ZP_{dik}\right], \sum_{i=1}^{41}(\sum_{k=1}^{54} X_{eik} \cdot ZQ_{dik} - S_i) > 0 \end{cases} \tag{5}$$

There are very many constraints in determining the actual crop planting strategies, We established the following constraints by reviewing the relevant documentation, as detailed below:

(1) It is predicted that the later years will be relatively stable compared to the production acreage, sales prices, and other factors from 2023. This assumes that there are no external influences causing disruptions. The production acreage, planting costs, and unit sales prices for different crops in various regions and seasons will be modeled using the data from 2023 as a reference, structured as follows for each quarter.

$$\begin{cases} ZO_{dik} = ZO_{1ik}, (e \bmod 2 = 1) \\ ZO_{dik} = ZO_{2ik}, (e \bmod 2 = 0) \end{cases} \tag{6}$$

$$\begin{cases} ZP_{dik} = ZP_{1ik}, (e \bmod 2 = 1) \\ ZP_{dik} = ZP_{2ik}, (e \bmod 2 = 0) \end{cases} \tag{7}$$

$$\begin{cases} ZQ_{dik} = ZQ_{1ik}, (e \bmod 2 = 1) \\ ZQ_{dik} = ZQ_{2ik}, (e \bmod 2 = 0) \end{cases} \tag{8}$$

(2) In order to make the decision-making results in line with the reality of the planting area, it is necessary for each type of crop corresponding to each piece of land area sum does not exceed the area of the land, so it is necessary to planting decision-making variables for each piece of land should be constraints. Let the total area of flat dry land, terraced land, hillside land, watered land, common greenhouse, and intelligent greenhouse be Al_f, Bl_g, Cl_h, Dl_t, El, and Fl, respectively.

$$\begin{cases} \sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=1}^{6} X_{eik} - \sum_{f=1}^{6} Al_f) \le 0 \\ \sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=7}^{20} X_{eik} - \sum_{g=1}^{4} Al_g) \le 0 \\ \sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=21}^{26} X_{eik} - \sum_{h=1}^{6} Al_h) \le 0 \\ \sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=27}^{34} X_{eik} - \sum_{t=1}^{8} Al_t) \le 0 \\ \sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=35}^{50} X_{eik} - \sum_{f=1}^{6} Al) \le 0 \\ \sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=50}^{54} X_{eik} - \sum_{f=1}^{6} Al) \le 0 \end{cases} \tag{9}$$

(3) Flat dry land, terraced land, hillside land planted one season of food crops per year, 41 kinds of crops categorized and organized to get 1-15 numbering corresponds to food crops except rice, 17-37 numbering corresponds to vegetables, 38-41 corresponds to edible fungi, assuming that $i=1,2,...,41$, respectively, corresponds to the corresponding crop numbering, you can list the set of the first three plots of land planted only with grain.

$$\begin{cases} (e \bmod 2=1)\wedge(1\le k\le 6)\wedge(1\le i\le 15) \\ (e \bmod 2=1)\wedge(7\le k\le 20)\wedge(1\le i\le 15) \\ (e \bmod 2=1)\wedge(21\le k\le 26)\wedge(1\le i\le 15) \end{cases} \tag{10}$$

(4) The watered land is planted with one season of rice or two seasons of vegetables per year, along the variable i in Constraint III, which denotes the set of demands.

$$\begin{aligned} &\left[(e \bmod 2=1)\wedge(27\le k\le 34)\wedge(i=16)\right] \\ &\vee\left[(e \bmod 2=1)\wedge(27\le k\le 34)\wedge(17\le i\le 34)\right] \\ &\wedge\left[(e \bmod 2=0)\wedge(27\le k+1\le 34)\wedge(35\le i+1\le 37)\right] \end{aligned} \tag{11}$$

(5) Cabbage, white radishes and carrots can only be grown in the second season in a watered field, after listing the conditions under which they can be grown using the meaning of i above.

$$35\le i\le 37\to(e \bmod 2=0)\wedge(27\le k\le 34) \tag{12}$$

(6) Ordinary greenhouse first season vegetables, second season edible fungi cultivation according to the seasonal difference condition limitations to build constraints.

$$\begin{aligned} &(e \bmod 2=1)\wedge(35\le k\le 50)\wedge(17\le i\le 34)\wedge(e \bmod 2=1) \\ &\wedge(35\le k+1\le 50)\wedge(38\le i\le 41) \end{aligned} \tag{13}$$

(7) Edible mushrooms are grown under the condition that they can only be grown in the second season in ordinary greenhouses.

$$38\le i\le 1\to(e \bmod 2=0)\wedge(35\le k\le 50) \tag{14}$$

(8) The constraints for growing crop species in smart greenhouses are:

$$(17\le i\le 34)\wedge(51\le k\le 50) \tag{15}$$

(9) The area allocated for each crop planted in a single plot should not be too small. It is important to define decision variables that limit the area constraints of different plots. The minimum planting area in 2023 can be calculated similarly to the corresponding regions from previous years.

$$\begin{cases} X_{eik}>35, 1\le k\le 6 \\ X_{eik}>20, 7\le k\le 20 \\ X_{eik}>13, 21\le k\le 26 \\ X_{eik}>12, 27\le k\le 34 \\ (X_{eik}=0.3)\vee(X_{eik}=0.6), 35\le k\le 54 \end{cases} \tag{16}$$

(10) Each crop cannot be planted consecutively in the same plot, comparing the planting limit of each plot in the current year with that of last year, and if the variety was planted last year it cannot be planted in that plot this year.

$$C(e,i,k)\Rightarrow\neg C(e-1,i,k) \tag{17}$$

(11) Legumes need to be planted at least once in three years.

$$\begin{aligned}
& C\left(e,\left(5<i<17\right)\vee\left(19<i\leq 41\right),k\right) \\
& \Rightarrow C\left(e+1,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right) \\
& \vee C\left(e+2,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right) \\
& \vee C\left(e+3,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right) \\
& \vee C\left(e+4,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right) \\
& \vee C\left(e+5,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right) \\
& \vee C\left(e+6,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right)
\end{aligned} \tag{18}$$

For the objective function sum of the above constraints, the topic shows a relatively large number of constraints, the objective function and constraints established to organize and summarize the final objective function and all the constraints can be derived from the expression of the final objective function and all the constraints:

$$\begin{cases}
MAX\begin{cases}
\sum_{e=1}^{14}\sum_{i=1}^{41}\sum_{k=1}^{54}X_{eik}\cdot ZO_{dik}\cdot ZQ_{dik}-ZP_{dik},\sum_{i=1}^{41}(\sum_{k=1}^{54}X_{eik}\cdot ZQ_{dik}-S_i)\leq 0 \\
\sum_{e=1}^{14}\sum_{i=1}^{41}\left[(ZS_i-X_{eik}\cdot ZO_{dik})\cdot\sum_{k=1}^{54}ZQ_{dik}-\sum_{i=1}^{54}ZP_{dik}\right],\sum_{i=1}^{41}(\sum_{k=1}^{54}X_{eik}\cdot ZQ_{dik}-S_i)>0 \\
\sum_{e=1}^{14}\sum_{i=1}^{41}\left[\left(1.5S_i-0.5\sum_{k=1}^{54}X_{eik}\cdot ZO_{dik}\right)\cdot\sum_{k=1}^{54}ZQ_{dik}-\sum_{i=1}^{54}ZP_{dik}\right],\sum_{i=1}^{41}(\sum_{k=1}^{54}X_{eik}\cdot ZQ_{dik}-S_i)>0
\end{cases} \\
s.t.\begin{cases}
\begin{cases} ZO_{dik}=ZO_{1ik},(e\bmod 2=1) \\ ZO_{dik}=ZO_{2ik},(e\bmod 2=0)\end{cases} \\
\begin{cases} ZP_{dik}=ZP_{1ik},(e\bmod 2=1) \\ ZP_{dik}=ZP_{2ik},(e\bmod 2=0)\end{cases} \\
\begin{cases} ZQ_{dik}=ZQ_{1ik},(e\bmod 2=1) \\ ZQ_{dik}=ZQ_{2ik},(e\bmod 2=0)\end{cases} \\
\begin{cases}
\sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=1}^{6}X_{eik}-\sum_{f=1}^{6}Al_f)\leq 0 \\
\sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=7}^{20}X_{eik}-\sum_{g=1}^{4}Al_g)\leq 0 \\
\sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=21}^{26}X_{eik}-\sum_{h=1}^{6}Al_h)\leq 0 \\
\sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=27}^{34}X_{eik}-\sum_{t=1}^{8}Al_t)\leq 0 \\
\sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=35}^{50}X_{eik}-\sum_{f=1}^{6}Al)\leq 0 \\
\sum_{e=1}^{14}(\sum_{i=1}^{41}\sum_{k=50}^{54}X_{eik}-\sum_{f=1}^{6}Al)\leq 0
\end{cases} \\
\left(e\bmod 2=1\right)\wedge\left(1\leq k\leq 6\right)\wedge\left(1\leq i\leq 15\right) \\
\left(e\bmod 2=1\right)\wedge\left(7\leq k\leq 20\right)\wedge\left(1\leq i\leq 15\right) \\
\left(e\bmod 2=1\right)\wedge\left(21\leq k\leq 26\right)\wedge\left(1\leq i\leq 15\right) \\
\left[\left(e\bmod 2=1\right)\wedge\left(27\leq k\leq 34\right)\wedge\left(i=16\right)\right]\vee\left[\left(e\bmod 2=1\right)\wedge\left(27\leq k\leq 34\right)\wedge\left(17\leq i\leq 34\right)\right]\wedge \\
\left[\left(e\bmod 2=0\right)\wedge\left(27\leq k+1\leq 34\right)\wedge\left(35\leq i+1\leq 37\right)\right] \\
35\leq i\leq 37\rightarrow\left(e\bmod 2=0\right)\wedge\left(27\leq k\leq 34\right) \\
\left(e\bmod 2=1\right)\wedge\left(35\leq k\leq 50\right)\wedge\left(17\leq i\leq 34\right)\wedge\left(e\bmod 2=1\right)\wedge\left(35\leq k+1\leq 50\right)\wedge\left(38\leq i\leq 41\right) \\
38\leq i\leq 1\rightarrow\left(e\bmod 2=0\right)\wedge\left(35\leq k\leq 50\right) \\
\left(17\leq i\leq 34\right)\wedge\left(51\leq k\leq 50\right) \\
\begin{cases}
X_{eik}>35,1\leq k\leq 6 \\
X_{eik}>20,7\leq k\leq 20 \\
X_{eik}>13,21\leq k\leq 26 \\
X_{eik}>12,27\leq k\leq 34 \\
\left(X_{eik}=0.3\right)\vee\left(X_{eik}=0.6\right),35\leq k\leq 54
\end{cases} \\
C\left(e,\left(5<i<17\right)\vee\left(19<i\leq 41\right),k\right)\Rightarrow C\left(e+1,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right)\vee \\
C\left(e+2,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right)\vee C\left(e+3,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right)\vee \\
C\left(e+4,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right)\vee C\left(e+5,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right)\vee \\
C\left(e+6,\left(1<i<5\right)\vee\left(17<i\leq 19\right),k\right)
\end{cases}
\end{cases} \tag{19}$$

3. Results

3.1. Variable simplification and dimensionality reduction

It can be seen that the optimization problem has a total of 11 constraints, and the constraints are all very complex, so first of all, the constraints in the model building can be divided into two categories of direct constraints on the decision variables and constraints on the decision scheme, and first of all, the second category of constraints are solved and simplified. Constraints (3), (4), (5), (6), (8), (9), (10), (11) are all constraints on e, i, k in X_{eik}.

Method 1: Directly generate the scheme space containing all the strategies and perform mathematical calculations to obtain the size of the space:

$$SpaceSize = size(e) \times size(i) \times size(k) \tag{20}$$

Solve for $SpaceSize = 30996$. Construct the data structure of 30996×3 to store all the programs, and fill the whole data structure by traversing the value ranges of e, i, and k respectively, then each specific program is the random combination set of e, i, and k variables $Space$. Each constraint on the program conditions will generate the corresponding combination of data, through the combination of comparison and screening and merger composition to achieve the purpose of constraints.

Method 2: The constraints generated based on the specific crop cultivation specifications stipulated by the 4 types of terrain are (3), (4), (9), and (11), all of which are logically represented, and these 4 types of constraints are merged and constituted through logical operations to generate a program space that meets the requirements of all the specific planting sites, and the original $Space$ set is screened using a comparative screening method.

Method 3: For constraints (8) and (10), implication relations are introduced for expression, and transformations are performed when solving, e.g., (8) is transformed into two conditional equations expressed only by logical operations:

$$condition1 = (35 \le i \le 37) \wedge (e \bmod 2 = 0) \wedge (27 \le k \le 34) \tag{21}$$

$$condition2 = (1 \le i \le 34) \wedge (38 \le i \le 41) \tag{22}$$

Then simplify the problem, using the aforementioned merged composition method to solve $FinalCondition = condition1 \,\&\&\, condition2$, and finally compare the screening to realize the constraints, constraints (10) with (8), will not be repeated here.

Method 4: Constraints (5) and (6) increase the consideration of time dimension and use the expression of implication relation, because the combined program of all years in the comprehensive calculation of the combination of all years in the strong constraints of (5) will appear in the combination of irrational sieve, so choose to solve the problem step-by-step, year by year. Directly introduce the planting situation in 2023 as known information, directly construct the corresponding specific program space $Space_{2023}(e_{2023}, i_{2023}, k_{2023})$, where e_{2023}, i_{2023}, k_{2023} all represent the statistical values obtained by the data preprocessing, directly materialize the time dimension of 2024 $e_{2024} = 1, 2$, and then use the method three processing to obtain the constraints of the program after the constraints of the year 2024, and similarly solve the program for the two seasons of 2025. 2023-2025 constitutes the planting cycle of the legume crop, the first combined with the planting situation in 2023 to determine the planting situation in all years, and then use the implicit relationship to express the relationship. First, combined with the 2023 planting situation to determine the situation of this cycle, to give the unit cycle of bean planting decision-making adjustment program, pseudo-code is as follows:

$$\begin{aligned} &if\,(dou_{2023} == 1) dou_{2024} == 1 \,\|\, dou_{2025} == 1 \,\|\, dou_{2024} == 0 \,\|\, dou_{2025} == 0 \\ &if\,(dou_{2023} == 0) dou_{2024} == 1 \,\|\, dou_{2025} == 1 \end{aligned} \tag{23}$$

Where dou_{2023}, dou_{2024}, dou_{2025} represent the legume crop planted in each plot in 2023-2025.

3.2. Solving Optimization Models by Genetic Algorithms

3.2.1. Model selection

After simplifying the constraints still belongs to the strategy optimization model with complex constraints, which is difficult to solve using traditional algorithms, and the intelligent optimization genetic algorithm with heuristic information is chosen to solve the problem.

3.2.2. Genetic algorithm model solving

Step 1: Initialize the stock

Individuals consist of a combination of filtered combination schemes and decision variables, such as $a_{(1,1)} = (e, i, k, X_{eik})$,where (e,i,k) is a reasonable combination, combined with constraints (7) and randomly generated initialization of the value of X_{eik} , so that the length of the individual is 4, and the size of the population is the number of reasonable constraint schemes in a fixed quarter $size((e,i,k))$.

During the coding process, $a(:,1:3)$ are the optimal cases that have been solved, into the initial value of the decision variable X_{eik} is binary coded. The empirical $X_{eik} \in [12,86]$ is obtained from Annex 1.xlsx and the value mapping intervals are calculated:

$$rg = \max X_{eik} - \min X_{eik} + 1 \tag{24}$$

To illustrate the encoding process using X_{1ik} (first quarter of 2024) as an example, solve for $rg = 75$, because $2^6 = 64$, $2^7 = 128$, so perform a 7-bit binary encoding.

Table 1. Binary coding table

season (sports)	Crop number	Lot No.	acres planted	season (sports)	Crop number	Lot No.	encodings
1	1	4	49	1	1	4	0111011
1	1	17	55	1	1	17	1000100
1	1	23	15	1	1	23	0000010
1	2	8	60	1	2	8	1000101
...	...	...	...	...	...	...	...

Table 1 represents the coding of acres planted for different crops corresponding to different land in different seasons, the above table represents only a part of the results, the rest is available by running the code.

The filtered decision variablesX_{eik}Initial values corresponding to the visualization were obtained, binary code shown in Figure 1.

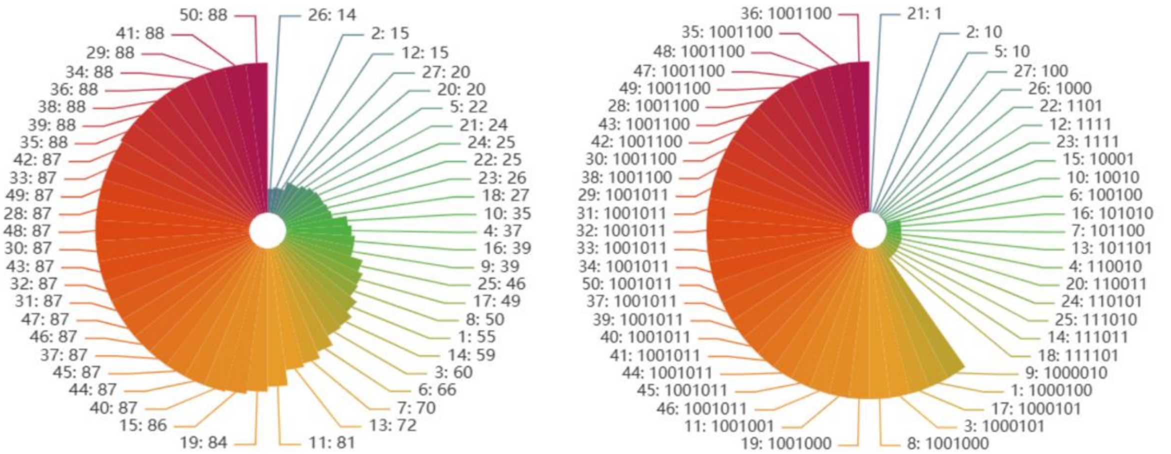

Figure 1. Coding results display diagram

The uncoded 50 combination scenarios on the left. Figure 1 On the right is a visualization of the binary coding results for each scheme, turning each scheme into a binary code.

Step 2: Calculation of the objective function

In the decoding process, the greenhouse crops are specialized, the range of X_{eik} is increased to $[12,88]$, and it is stipulated that when the binary code is converted to decimal, it is 87 and 88, which represent the minimum planting area and the maximum planting area of the common greenhouse and the intelligent greenhouse, respectively. Following the idea of method 4 in variable simplification and dimensionality reduction, the problem is simplified into step-by-step solution, matching the ZO_{dik}, ZP_{dik}, ZQ_{dik} corresponding to X_{eik} after decoding, and then listing the segmented objective function.

Step 3: Dynamic programming algorithms simplify the constraints to compute the degree of adaptation

The different state intervals of constraint (2) are processed using dynamic programming algorithms, and each interval is set up as a state space, e.g., $Z(k)=(k,X_{eik})$, where the range of values of k denotes the state, and the state transfer equation is:

$$Z(k_s+1)=(k_{s+1},X_{eik_{s+1}}) \tag{25}$$

The transfer boundary is defined as:

$$s[i]=k\in\begin{bmatrix}1 & 6\\ 7 & 20\\ 21 & 26\\ 27 & 34\\ 35 & 50\\ 51 & 54\end{bmatrix} \tag{26}$$

The unsatisfied terms in the constraints are constructed into a penalty function by introducing adjustable penalty coefficients C, bounded by each area $A1_f$, $B1_g$, $C1_h$, $D1_t$, $E1$, $F1$, and fitness function = objective function - C.

Step 4: Crossover, mutation

Genetic operations are performed only within the same crop, i.e., if $a(x,2)==a(x+1,2)$ then coding changes to $a(x,4)$ and $a(x+1,4)$. Constraints (7) are again introduced for solving the optimal individual storage to the required cropping strategy. The result is shown in Table 2 and Table 3

Table 2. Results of planting strategy in 2024

Plot Name	Soybean	Black Bean	Red Bean	...	Mushroom	White Maitake Mushroom	Cordyceps Sinensis
A1	/	/	/	...	/	/	/
A2	/	21.8714	/	...	/	/	/
A3	/	21.8285	13.1714	...	/	/	/
…	…	…	…	…	…	…	/
F2	/	/	/	...	/	/	/
F3	/	/	/	...	/	/	/
F4	/	/	/	...	/	/	/

Table 3. Results of planting strategy in 2025

Plot Name	Soybean	Black Bean	Red Bean	...	Mushroom	White Maitake Mushroom	Cordyceps Sinensis
A1	/	43.7	36.3	...	/	/	/
A2	/	/	/	...	/	/	/
A3	/	/	/	...	/	/	/
…	…	…	…	…	…	…	/
F2	/	/	/	...	/	/	/
F3	/	/	/	...	/	/	/
F4	/	/	/	...	/	/	/

3.2.3. Analysis of results

The annual profit appears to fluctuate first and then stabilize as shown in Table 2 and Table 3, which is in line with the characteristics of the genetic algorithm, in which the planting strategy is more dispersed, but also presents a certain cyclicity, such as the planting strategy of parcel B13-C6 in the six-year period is hovering between soybeans and grains, and it is presumed that the real overproduction and pre-sale condition is more likely to be taken out of the priority queue for planting more parcels due to the higher expected yields of soybeans and grains, and at the same time Because soybeans are legumes, they are able to fit well into the three-year-one-legume constraint, resulting in a soybean-grain-soybean cycle.

4. Conclusions and outlooks

This paper establishes a profit-maximizing decision model for crop cultivation in mountainous areas of North China based on data. By analyzing the planting and profit under different time and land types, the mu yield, planting cost and selling price in 2023 are compiled, and the expected sales volume of each crop is set as its total output. Combining this information, 11 constraints are analyzed and an optimization model is constructed with the objective of maximizing overall planting profit. Four methods are used to simplify the constraints, including generating a complete decision space, filtering solutions by logical operations, transforming implication relations into simple operations, and solving year by year. Finally, genetic algorithm and dynamic programming algorithm were used to solve the objective function and obtain the optimal planting strategy. This study can provide a reference for crop planting problems in other regions.

Although the model developed in this study effectively solves the problem, there are still some drawbacks and improvements that can be made, the constraints in the model do not take into account all the actual market fluctuations, and more optimization algorithms can be explored in the future to enhance the robustness and adaptability of the model.

References

[1] Xie Y L, Xia D X, Ji L, et al. An inexact stochastic-fuzzy optimization model for agricultural water allocation and land resources utilization management under considering effective rainfall [J]. Ecological indicators, 2018, 92: 301 - 311.

[2] Anitha S, Chandrakanth M G. Do Farmers Gain by More Crop Per Drop? [J]. Indian Journal of Agricultural Economics, 2022, 77 (4): 592 - 609.

[3] Hu X, Li K, Li J, et al. Load forecasting model consisting of data mining based orthogonal greedy algorithm and long short-term memory network [J]. Energy Reports, 2022, 8: 235 - 242.

[4] Vayabari D A G, Ilham Z, Md Saad N, et al. Cultivation strategies of kenaf (Hibiscus cannabinus L.) as a future approach in Malaysian agriculture industry [J]. Horticulturae, 2023, 9 (8): 925.

[5] Feng X, Zhao F, Jiang G, et al. A tabu memory based iterated greedy algorithm for the distributed heterogeneous permutation flowshop scheduling problem with the total tardiness criterion [J]. Expert Systems with Applications, 2024, 238: 121790.

[6] Chang C, Liao Z, Costa A L H, et al. Globally optimal design of intensified shell and tube heat exchangers using complete set trimming [J]. Computers & Chemical Engineering, 2022, 158: 107644.

[7] Gen M, Lin L. Genetic algorithms and their applications [M]//Springer handbook of engineering statistics. London: Springer London, 2023: 635 - 674.

[8] Çetin V, Yıldız O. A comprehensive review on data preprocessing techniques in data analysis [J]. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 2022, 28 (2): 299 - 312.

[9] Bertsekas D. Abstract dynamic programming [M]. Athena Scientific, 2022.

[10] Mirjalili S, Mirjalili S. Genetic algorithm [J]. Evolutionary algorithms and neural networks: theory and applications, 2019: 43 - 55.