Optimisation Of Crop Planting Strategies Based on Matrix Genetic Algorithm

Kunwei Wu, Liangshun Wang *

School of Mechanical and Electrical Engineering, Hainan University, Hainan, China, 570228

* Corresponding Author Email: 993576@hainanu.edu.cn

Abstract. With the increasing demand for sustainable agricultural development, it has become crucial to develop crop planting optimisation strategies, which are important for enhancing production efficiency, reducing planting risks, and using resources efficiently. This study focuses on the optimisation strategy of rural crop cultivation in the mountainous areas of North China, and adopts matrix genetic algorithm to take into account the actual influencing factors such as sales volume, yield per unit area, cultivation cost and sales price, and establishes a multi-period dynamic optimisation model to explore the optimal cultivation strategy for the next seven years under two scenarios, namely, the excess yield stagnation and the price reduction, and the optimal cultivation strategy for the next seven years under two scenarios, namely, the excess yield stagnation and the price reduction. In this way, the optimal planting strategies for the next seven years were investigated under the two scenarios of excess production and sales. The average annual profits for the next seven years under these two scenarios reached RMB 2,527,837 and RMB 2,684,160, respectively. Based on the results, it can be concluded that the optimisation of planting strategies carried out by this model is very considerable and has practical applicability, which solves the current situation that the optimisation of crop planting strategies fails to integrate the actual situation.

Keywords: MGA, Dynamic Optimisation Models, Crop Strategy Optimization, Rural Mountain Areas.

1. Introduction

A certain countryside located in the mountainous region of North China has abundant arable land resources such as flat dry land, terraced land, hillside land and agricultural greenhouses, but due to the cold climate and rugged terrain in the region[1], most of the arable land can only be planted with one season of crops per year, and at the same time, different arable land types are also suitable for planting different crops, so it is necessary to reasonably optimise the crop planting strategy in order to obtain the maximum profit. In the past, Zhang Hao[2] and others studied the optimisation of planting structure in tropical grassland climate based on stochastic dynamic programming model, which contributed to the sustainable development of agriculture, but the research targeted more extreme climate types, which is difficult to be promoted and applied in China; Bi Xiaoyang[3] and others combined linear programming and entropy weight TOPSIS method to determine the structure of agriculture, forestry, animal husbandry and fishery based on the comprehensive consideration of economic, ecological and social benefit. The optimisation model has a strong global nature, but does not take into account future market fluctuations, planting cost changes and other uncertainties; WU Feng[4] and others studied the crop conversion optimisation model, simulated a variety of scenarios in order to reduce the impact of resources and environmental factors, while maximizing the income of farmers, but the complexity of the model limits its direct application in the actual agricultural production process; CHEN Man[5] and others used the theory of comparative advantage to establish a model to measure the comprehensive advantage index of crops in Southwest China, which is of great significance for regional planting structure adjustment, but the study was based on historical data and did not predict the future situation.

In summary, although there have been many cutting-edge developments in the field of crop planting optimisation strategies, the consideration of the actual planting of rural crops in mountainous areas has not been deepened. Therefore, this paper proposes a matrix genetic algorithm for this kind of problem, establishes a multi-period dynamic optimisation model and applies it for the first time to the optimisation of crop planting strategy, which has a small demand for computational resources, accurate prediction results, and has practical applicability. This paper is divided into five chapters, Chapter 1 Introduction introduces the research background and related work, Chapter 2 Theory introduces the matrix genetic algorithm and multi-period dynamic optimisation model, Chapter 3 Experiment introduces the overall experimental process, including the establishment of the model to improve and validate, Chapter 4 Conclusion introduces the optimal crop planting strategy and the average annual profit based on the research of this paper, Chapter 5 Summary introduces the overall situation of this paper and the room for model improvement.

2. The basic funamenta

2.1. Multi-period dynamic optimisation model

Dynamic optimisation models have a wide range of applications in dealing with complex systems that change over time, including Dynamic Multi-Objective Optimisation (DMO) [6-7], Adaptive Dynamic Programming (ADP)[8-9] and modern optimisation theory and applications[10]. The objective of dynamic optimisation is to maximise (or minimise) an objective function, which is usually the cumulative utility or cost in time. The general formula for the objective function is as follows:

$$\max J = \int_{t_0}^{t_f} f(x_t, u_t, t)dt + g(x_{t_f}, t_f) \tag{1}$$

Among them:

J: cumulative utility or objective value. f (x_t, u_t, t): immediate utility or cost function at each point in time. g (x_{tf}, t_f): utility or cost function of the final state. x_t: state variable describing the dynamic behaviour of the system. u_t: control variable, which is a variable that can be adjusted by the decision maker. t_0, t_f: initial and termination time.

Dynamic optimisation models usually contain the following constraints:

State equations (dynamic transfer equations):

$$x_t = h(x_t, u_t, t) \tag{2}$$

Describe how the state variable x_t varies with time t and the control variable u_t.

Control Constraints:

$$u_t \in U(x_t, t) \tag{3}$$

Indicates the range of values of the control variable.

Boundary conditions:

$$x_{t_0} = x_0, x_{t_f} \in X_f \tag{4}$$

Give the constraints of the initial state and termination state of the system.

In the research process of this paper, in order to be able to effectively deal with the uncertainty and complexity factors such as seasonal changes, market price fluctuations and production cost changes in the process of agricultural production, to provide scientific decision-making support, to help farmers to rationally allocate resources in long-term planning and to reduce the risk of planting, it is necessary to determine the optimal planting scheme year by year, so a multi-period dynamic optimization model is established so as to maximize the benefits or minimize the risk of.

2.2. Matrix Genetic Algorithm

Genetic algorithm is a heuristic optimisation algorithm based on natural selection and genetic mechanism, which gradually optimises the solutions in the population by simulating the biological evolution process with operations such as selection, crossover and mutation. Because of its powerful finding ability and adaptability to complex constraints, genetic algorithm is widely used to solve multi-objective and multi-constraint optimisation problems of crop cultivation [11-12]. In this paper, the standard genetic algorithm is improved in threefold: (1) introducing matrices as variables [13] and using matrix optimisation algorithm for model building. Compared with the use of vectors to express populations in the standard genetic algorithm, the improved algorithm can search for matrix variables in the solution space more efficiently and handle complex constraints better. (2) With the help of the introduction of the Delicacy distribution [14], the algorithm achieves dynamic planting ratio adjustment, thus ensuring the diversity of planting schemes. (3) By rewriting the crossover and mutation functions, the adaptive crossover and mutation operations enhance the reasonableness of the planting scheme, making the optimisation process more in line with the practical application requirements. The above threefold improvement is used to describe the relationship between planting area and yield, sales volume and economic benefits of crops in a more suitable way.

3. Experiment

3.1. Overall experimental design

In this paper, the dataset is first pre-processed, then the optimisation model is built, then the results of the model are validated, and finally the conclusions are drawn. The overall flow design diagram of the experiment is shown in Figure 1.

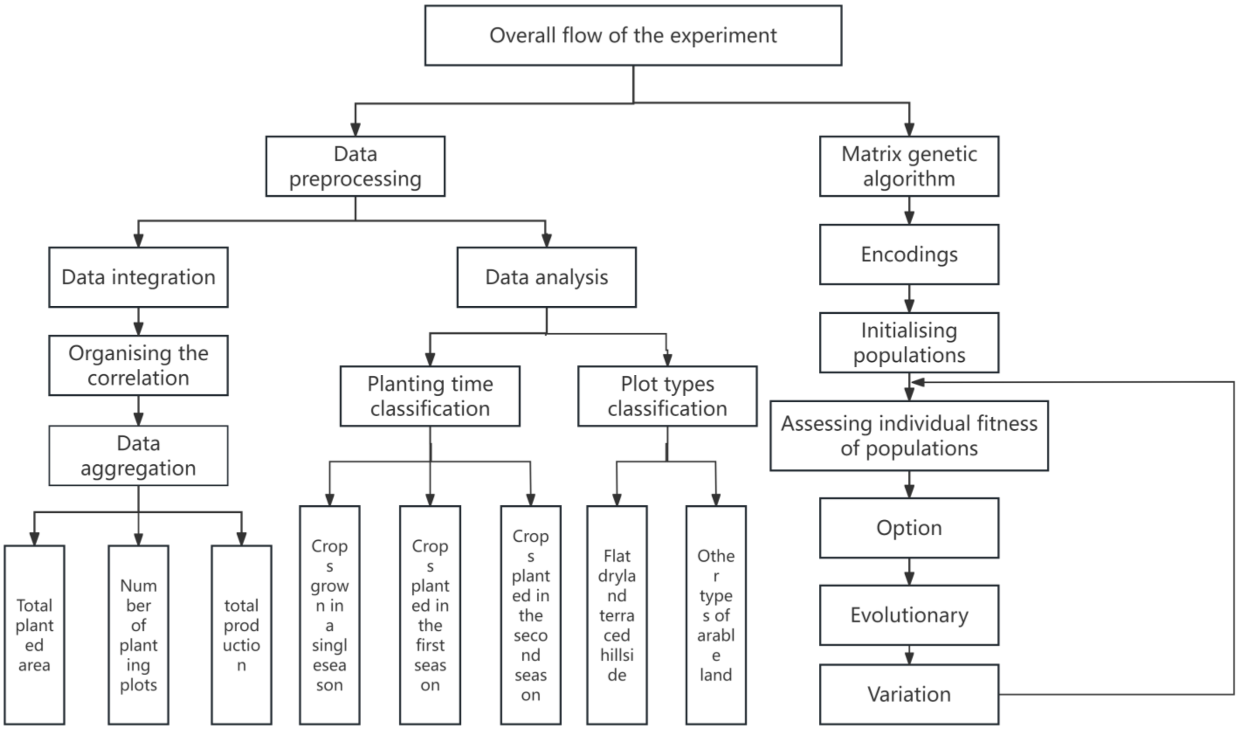

Figure 1. Design diagram of the overall flow of the experiment

3.2. Data Preprocessing

According to the Chinese Statistical Society and Kaggle public dataset data volume is huge and contains a lot of details, which is typical of the actual production and sales data. By performing data collation and correlation analysis, as well as preliminary data aggregation, it can effectively facilitate the subsequent model building, analysis and prediction work. In this paper, we first ensure the accuracy and consistency of the data by identifying and cleaning invalid data in the dataset. Then, using common fields such as crop name and planting plot as key links, the VLOOKUP function of Excel was used to correlate and merge the data, and then the results were stored in the cleaned dataset. This approach ensured that the same categories of information from different data sources were correctly matched. The merged data were validated again to ensure logical consistency and integrity. Then, based on the crop categories, the area of each crop in 2023 was aggregated with the following aggregation rules:

$$\sum_{x}(Q(x\,,y\,,z)) \tag{5}$$

Among them:

Q(x,y,z) denotes the area of the yth crop planted on the xth plot name of the zth plot type.

Subsequently, data analysis was carried out. Based on the cleaned dataset, this paper classifies all crops into three categories according to the time of planting: single-season crops, first-season crops, and second-season crops. Based on this classification, this paper aggregates the total amount of cultivated land for each type of crop in their respective cropping seasons, and the corresponding data are shown in Figure 2.

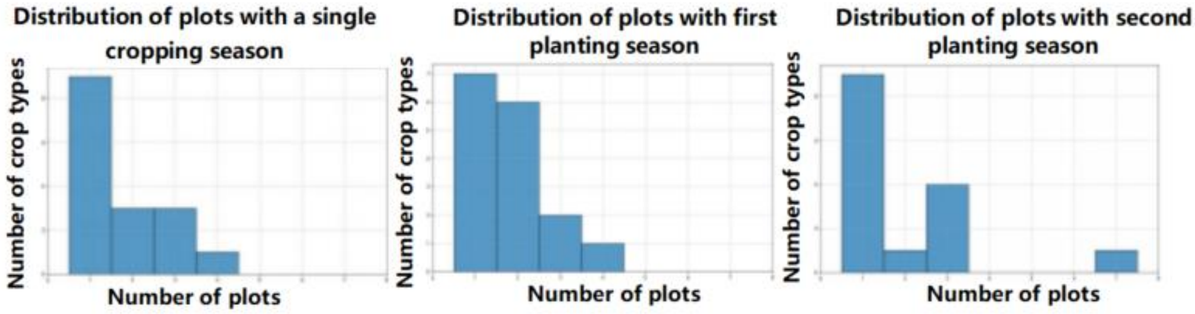

Figure 2. Number of planting plots

Secondly, according to the data in the dataset, except for rice, which can be grown on irrigated land, the other food crops are only suitable for cultivation on flat dry land, terraced land and hillside land. Therefore, in order to facilitate the calculation and management, this paper reclassifies the plot types into two main categories: (1) flat drylands, terraces and hillsides: planting all food crops except rice; (2) other cultivated land: planting rice and all other non-food crops.

3.3. Modelling

According to the actual situation in daily life, when the seasonal crop production exceeds the sales volume, farmers have two corresponding solutions: (1) the excess part is not sold, resulting in waste; (2) the excess part is sold at a reduced price of 50% of the sales price. Therefore, in this paper, we introduce the objective function of the multi-period dynamic optimisation model for each of the above two solutions.

Corresponding to the output in excess of the sales volume of the part of the stagnant sales of the objective function 1:

$$\max F_1 = \sum_i \sum_j (S(j) \times Ss(j) - Sp(j) \times Sc(j)) \tag{6}$$

Objective function 2 that corresponds to the excess of production over sales volume sold at a reduced price of 50 per cent of the sales price

$$\max F_2 = \begin{cases} \sum_i \sum_j (S(j) \times Ss(j) - Sp(j) \times Sc(j) + (Sc(j) - Se(j) \times S(j) \times 0.5))\,, Sc(j) > Se(j) \\ \sum_i \sum_j (S(j) \times Ss(j) - Sp(j) \times Sc(j))\;, Sc(j) \le Se(j) \end{cases} \tag{7}$$

Among them:

M(i,j): acres of the jth crop on the ith plot name. N(i,j): number of acres of the jth crop on the ith plot name. S(j): Unit price of the jth crop sold. Sc(j): Production of the jth crop in the current year. Sp(j): Cost of growing the jth crop. Se(j): Production of crop j in 2023, i.e., expected sales. Ss(j): Actual sales of crop j in that year.

The following can be derived:

$$\mathrm{Ss(j)} \ = \ \textstyle\sum_{\mathrm{j}}(\mathrm{M(i,j)} \times \mathrm{N(i,j)}) \tag{8}$$

$$\mathrm{Ss(j)} \ = \ \min\,(\mathrm{Sc(j)}\,,\mathrm{Se(j)}) \tag{9}$$

The corresponding decision variables of the model are represented:

$A_{26\times15}$: represents a 26×15 matrix. Where 26 is the total number of plots in the first broad category and 15 is the number of crops that can be grown on the first broad category of plots.

$AS_{26\times15\times8}$: represents the cultivation of the first major category of plots (i.e., flat dry land, terraced land and hillside land) over an 8-year period from 2023-2030.

$B_{28\times26\times2}$: represents a 28×26×2 matrix. Where 28 is the total number of plots in the second major category, 26 is the number of crops that can be grown on the second major category, and 2 denotes the seasonal situation. In this paper, the single season and the first season are denoted by 0, and the second season is denoted by 1.

$BS_{28\times26\times2\times8}$: denotes the cultivation of the second major category of plots in the 8-year period from 2023-2030.

The constraints corresponding to the model are as follows:

Constraint 1: The planting scheme should take into account the farmers' convenience of farming operations and field management, so the plots of each crop per season should not be too scattered. According to the data shown in Fig. 2, the maximum number of plots is seven, regardless of whether the planting is done in a single season, the first season or the second season. Therefore, if the total number of plots cultivated for any crop in each planting season is greater than or equal to 7 plots, the crop is considered in this paper to be too dispersed.

Constraint 2: Considering the economic efficiency, the cultivation area of each crop in a single plot should not be too small. Therefore, in the course of this paper, when the proportion of planted area of a crop on any individual plot of the first major plot type is less than 25 per cent, it is considered to be too small, and is expressed by the following formula:

$$0.25 \leq \ \mathrm{A[i\,,j]} \leq \ 1 \tag{10}$$

Constraint 3: A crop is considered to be appropriately planted when the proportion of its area planted on any individual plot in the second major plot type is limited to 50 per cent and 100 per cent; otherwise, it is considered to be underplanted and is expressed by the following formula:

$$\mathrm{B[i\,,j\,,m]} \ = \ 0.5 \text{ or } 1 \tag{11}$$

Constraint 4: The suitability of flat drylands, terraces and hillsides for growing food crops (except rice) in a single season each year. Due to the definition of the constraint variable $A_{26\times15}$, it can be ensured that flat early land, terraces and hillsides are planted with food crops only in a single season.

Constraint 5: Water-consumed land can be planted with rice in a single season or with vegetable crops in two seasons per year. Thus for i $\in$ {0,1,...... ,7}, if j = 0 and m = 0 and B[i, 0, 0] $\neq$ 0, we have:

$$\mathrm{B[i\,,j'\,,m']} \ = \ 0, \text{when } \mathrm{j'} \ \neq \ 0 \text{ or } \mathrm{m'} \ \neq \ 0 \tag{12}$$

The above equation shows that if in matrix B, the element B[i,0,0] in row i is not equal to zero, then all other elements in this row (in other columns j′and layers m′) will be set to zero. Here j′and m′denote all columns and layers except the first column and the first layer. Their physical meaning indicates that no other plants will be grown after planting rice.

Constraint 6: According to the seasonal requirement of crop planting, cabbage, white radish and carrot can only be planted in the second season of the watered field.Let matrix B, $B[i',j',0] = 0$ means no planting in the second season and $B[i',j',1] \neq 0$ means planting in the second season. For $i' \in$ {watered field} and $j' \in$ {cabbage,white radish,carrot}, the following conditions must be satisfied to comply with the above requirements:

$$B[i',j',0] = 0 \text{ and } B[i',j',1] \neq 0 \tag{13}$$

Constraint 7: Ordinary greenhouses grow two crops per year, in the first season a variety of vegetables can be grown (except cabbage, white radish and carrot) and in the second season only edible mushrooms can be grown. The constraint variable variable $B_{28\times26\times2}$ can satisfy the above requirements.

Constraint 8: Since edible mushrooms are adapted to grow in a lower and suitable temperature and humidity environment, they can only be grown in ordinary greenhouses in autumn and winter. Constraint variable $B_{28\times26\times2}$ can meet the above requirements.

Constraint 9: Smart greenhouses can grow two seasons of vegetables per year (except cabbage, white radish and carrot). Constraint variable $B_{28\times26\times2}$ can meet the above requirements.

Constraint 10: Ordinary greenhouses are suitable for growing one season of vegetables and one season of edible fungi per year, while intelligent greenhouses are suitable for growing two seasons of vegetables per year. Variable $B_{28\times26\times2}$ can meet the above requirements.

Constraint 11: According to the growth law of crops, each crop cannot be planted continuously in the same plot (including greenhouses), otherwise the yield will be reduced. The following formula can be used to ensure that the same plants will not be grown on the same plot for two consecutive years:

$$AS[i,j,t-1] \times A[i,j] = 0 \tag{14}$$

$$BS[i,j,m,t-1] \times B[i,j,m] = 0 \tag{15}$$

Constraint 12: All land in each plot is planted with legumes at least once in three years, as soils containing legume rhizobacteria favour the growth of other crops. This can be shown by the following equation:

$$BS[i,j',m,t-2] \times BS[i,j,m,t-1] = 0 \text{, when } j' \text{is the legume} \tag{16}$$

$$BS[i,j',m,t-2] \times BS[i,j,m,t-1] \times B[i,j,m] \neq 0 \text{, when } j' \text{is the legume} \tag{17}$$

Constraint 13: In accordance with life experience, there can be cases where there is empty space in each plot, i.e., no planting, provided that the planted area does not exceed the total area of the plot. This can be shown by the following formula:

$$\sum_j (A[i,j]) \leq 1 \tag{18}$$

$$\sum_j (B[i,j,m]) \leq 1 \tag{19}$$

At this point, the experimental design part of this paper has been presented. Through the above experimental analysis, it can be seen that all the planting profits obtained using the multi-period dynamic optimisation model will be improved, as shown in the conclusion section below.

4. Results

After the above model establishment, this research process has initially completed the establishment of the whole theoretical model. The validation of the effect of the multi-period dynamic optimisation model will be carried out in the following. The profits for each of the next seven years under the specific two solutions and the average annual profits are shown in Table 1:

Table.1. Profitability of optimal planting programmes, 2024-2030

Vintages	2024	2025	2026	2027	2028	2029	2030	Average annual profit
Situation 1 Profit ($)	2533999	2516439	2489189	2531590	2563162	2533280	2518199	2527837
Situation 2 Profit ($)	2685915	2701733	2650397	2695607	2663162	2684366	2697937	2684160

Through the above study, it can be concluded that the optimal planting strategy derived from the multi-period dynamic optimization model studied in this paper has higher profits than the current crop planting scheme every year, which significantly demonstrates the advantages of the model algorithm studied in this paper in the field of planting strategy optimization of crops; at the same time, the profit stabilisation point and the maximum value of the profit are different in each year, reflecting the fact that the optimal planting strategy may change in different years due to Meanwhile, the profit stability point and profit maximum value are different in each year, reflecting that the optimal planting strategy may change in different years due to the uncertainty and complexity of seasonal changes, market price fluctuations and changes in production costs, which is in line with the actual situation of agricultural production, and has the practical applicability of promotion.

5. Conclusions

In this paper, we focus on the problem that the existing dynamic optimisation models cannot well combine the actual situation of China's rural mountainous areas to formulate the optimisation of crop planting strategies, and combine the uncertainty and complexity of seasonal changes, market price fluctuations and changes in production costs that may be encountered in the process of actual agricultural production, and innovatively apply the multi-period dynamic optimisation model and matrix genetic algorithm to the optimisation of crop planting strategies in the mountainous areas of North China. We innovatively applied the multi-period dynamic optimisation model and matrix genetic algorithm to the optimisation of rural crop planting strategies in the mountainous areas of North China, and explored the optimal planting strategies for the next seven years under the two scenarios of excess production stagnation and price reduction sales respectively. The expected average annual profits for the next seven years under the two scenarios reached RMB 2,527,837 and RMB 2,684,160, respectively, which verified the reliability, sophistication and practical applicability of the multi-period dynamic optimisation model and matrix genetic algorithm in the field of crop planting optimisation strategies. In the future, the solution efficiency of the model can be further improved by the combination of parallel computing and pruning algorithm to shorten the computation time and reduce the waste of computational resources.

Acknowledgements

The authors gratefully acknowledge the financial support from Hainan Provincial Natural Science Foundation of China (Grant No. 621RC512) funds.

References

[1] LI Shaoting,MU Na,REN Yanjun,Thomas GLAUBEN.Spatiotemporal characteristics of cultivated land use eco-efficiency and its influencing factors in China from 2000 to 2020[J].Journal of Arid Land,2024,16(3):396-414

[2] ZHANG Hao, ZHAO Shengwei, QIAN Jun, MAO Jinqiao, ZHANG Peipei, GONG Yiqing.Optimisation of planting structure in irrigation areas with savannah climate based on stochastic dynamic planning[J]. Water Saving Irrigation,2024(10):15-21

[3] Bi Xiaoyang,Fang Lan.Research on the structural optimisation model of agriculture, forestry, animal husbandry and fishery based on linear programming and entropy weight TOPSIS method - A case study of Shandong Province[J].Advances in Applied Mathematics,2024,13(5):2399-2408

[4] Crop Switching for Improved Agricultural Sustainability in China[J]. Bulletin of the Chinese Academy of Sciences,2023,37(01):6.

[5] Chen Man,Liu Shiwei.Research on the optimisation of crop planting structure in Southwest China[J].Journal of Jiangxi Agricultural University,2022,44(01):12-20.

[6] Ke Wei,Keke Huang,Chunhua Yang,Weihua Gui.Multi-Objective Adaptive Optimization Model Predictive Control:Decreasing Carbon Emissions from a Zinc Oxide Rotary Kiln[J].Engineering,2023(8):96-105

[7] ZOU, FEI, YEN, GARY G., TANG, LIXIN, et al. A reinforcement learning approach for dynamic multi-objective optimization[J]. Information Sciences: An International Journal,2021,546815-834.

[8] DERONG LIU, SHAN XUE, BO ZHAO, et al. Adaptive Dynamic Programming for Control: A Survey and Recent Advances[J]. 2021,51(1):142-160.

[9] FEI-YUE WANG, HUAGUANG ZHANG, DERONG LIU. Adaptive Dynamic Programming: An Introduction[J]. IEEE computational intelligence magazine,2009,4(2):39-47.

[10] Deng Qi,Gao Jianjun,Ge Dongdong,et al.Modern optimisation theory and application[J].Chinese Science:Mathematics,2020,50(07):899-968.

[11] GUO, YUNLI. Integrating genetic algorithm with ARIMA and reinforced random forest models to improve agriculture economy and yield forecasting[J]. Soft computing: A fusion of foundations, methodologies and applications,2024,28(2):1685-1706.

[12] XIUMING SONG. Research on genetic algorithm optimization for agricultural machinery operation path planning[J]. Applied Mathematics and Nonlinear Sciences,2024,9(1).

[13] QIU J., CAI X., ZHANG L., et al. A loss matrix-based alternating optimization method for sparse PU learning[J]. Swarm and Evolutionary Computation,2022,75101174-1-101174-15.

[14] Geng JM, Li W. Historical Exploration of the Delicacy Principle [J]. Journal of Guangxi University for Nationalities (Natural Science Edition),2024,30(01):59-65+73.