Research on Agricultural Planting Planning Optimization Based on Mixed Integer Programming and Genetic Algorithm

Tengyao Wang [1, *], Panyi Wang [2], Junye Feng [3]

[1] School of Information Engineering, Xinyang Agriculture and Forestry University, Xinyang, China, 464000

[2] School of Engineering, Ningbo University of Technology, Ningbo, China, 315211

[3] Department of Civil Engineering, Hunan City University, Yiyang, China, 413000

* Corresponding Author Email: wangtengyao666@outlook.com

Abstract. This study aims to optimize agricultural planting strategies in the mountainous regions of North China, maximizing economic benefits under the constraints of limited and fragmented arable land while addressing challenges such as overproduction, market demand fluctuations, and crop rotation requirements. While China's rural revitalization strategy accelerates agricultural modernization, mountainous agriculture still faces issues such as fragmented farmland, high production costs, and unstable market demand. Existing research primarily focuses on agricultural resource allocation and planning, but lacks comprehensive solutions to dynamic constraints, including overproduction risks, crop complementarities, and rotation. To address these challenges, this paper proposes an agricultural planting optimization model that integrates mixed integer programming (MIP) with a genetic algorithm (GA). Based on data collection and current state analysis, the model incorporates multiple-variable constraints and is optimized for two scenarios: one where excess production remains unsold, and another where excess produce is sold at a 50% discount. The model is solved, and the results are analyzed to derive the optimal planting strategy. The experimental results show that in the case of overproduction, the discounted sales strategy effectively reduces revenue fluctuations, offering higher revenue stability compared to the unsold scenario. Therefore, it is recommended to adopt a discounted sales strategy in cases of overproduction to stabilize agricultural income and enhance economic benefits. This paper provides theoretical support and practical references for the modernization of agriculture in the mountainous regions of North China.

Keywords: Genetic algorithm, mixed integer programming, planting optimization, Resource Allocation, Precision agriculture.

1. Introduction

With the deepening of the rural revitalization strategy, China's agricultural modernization has accelerated, transforming both production methods and rural landscapes. However, the complex distribution of agricultural resources, the diversity of arable land types, and climatic limitations pose significant challenges to achieving a win-win situation in terms of economic and ecological benefits under limited resources. This is particularly evident in the mountainous areas of North China, where fragmented farmland, frequent market demand fluctuations, and uncertain production costs render traditional agricultural models inadequate for modern development.

In recent years, many scholars have attempted to address issues in agricultural planning and resource allocation from various perspectives. For example, Zhou et al. (2023) proposed a path planning method combining the Floyd algorithm with an improved genetic algorithm to optimize agricultural robot routes in hilly regions, significantly enhancing machinery efficiency [1]. However, their approach does not extend to optimizing planting schemes. Similarly, Han et al. (2021) developed a mixed integer programming model that incorporates labor and capital constraints to solve a two-stage joint decision problem in the production and distribution of fresh agricultural products [2], yet it fails to consider constraints such as land rotation and specific crop cultivation (e.g., soybeans).

To address these shortcomings, this paper makes the following contributions: 1) It establishes an agricultural planting planning optimization model that integrates mixed integer programming with genetic algorithms, effectively handling multi-variable and dynamic constraints in planting decisions; 2) It innovatively incorporates overproduction risk—encompassing both unsold and discounted scenarios—with crop complementarities and rotation requirements, proposing an optimization scheme that balances economic and ecological benefits; and 3) It provides theoretical support and decision-making references for the coordinated development of traditional farmland and modern planting technologies in the mountainous regions of North China. The empirical analysis utilizes agricultural resource datasets from the official platform of China Undergraduate Mathematical Contest in Modeling (http://www.mcm.edu.cn/), with real competition cases validating the model's dual strengths in theoretical rigor and operational feasibility.

2. Related Theories

2.1. Mixed Integer Programming (MIP)

Mixed Integer Programming (MIP) is an extension of linear programming, enabling the inclusion of discrete decision variables. It optimizes a linear (or linearized) objective function subject to a set of linear constraints. MIP is widely used in resource allocation, scheduling, and strategic planning, where decisions often involve discrete choices, such as selecting which crops to plant on specific plots. In the context of agricultural planting optimization, MIP is particularly useful in modeling decisions about crop allocation, where land plots and crop types are discrete choices, and certain constraints, like planting area and crop adaptability, must be satisfied. The key components of a MIP model include the objective function, decision variables (which can be continuous or binary), and constraints that capture real-world limitations such as land availability and crop compatibility. The computational efficiency of solving MIP models is a major advantage when applying them to complex agricultural systems with multiple constraints. However, as the number of variables increases or when nonlinear constraints are present, MIP models may face computational challenges. This paper addresses such challenges by combining MIP with a genetic algorithm, thus improving the ability to solve complex dynamic optimization problems typical of agricultural planning [3].

2.2. Genetic Algorithm (GA)

Genetic algorithms (GA) are heuristic optimization methods inspired by the principles of natural evolution and genetics. GAs simulate natural processes such as selection, crossover, and mutation to evolve a population of candidate solutions toward an optimal or near-optimal solution. In agricultural planning, GAs are especially suited for handling complex, nonlinear relationships and multiple dynamic constraints. These include crop rotation requirements, resource limitations, and overproduction risks, such as unsold crops or excess produce sold at a discounted rate. GA's ability to explore diverse regions of the solution space and avoid premature convergence to local optima makes it highly effective in addressing agricultural optimization problems that cannot be easily tackled by traditional methods like linear programming. While GAs may not guarantee finding a globally optimal solution, they provide a reliable approach for solving problems with a high level of complexity and uncertainty, such as those found in agricultural planting strategies [4].

2.3. Theoretical Synergy

The integration of MIP with GA, as proposed in this paper, allows for the optimization of agricultural planting strategies in mountainous regions of North China, where constraints such as limited land availability, fluctuating market demand, and the need for crop rotation must be considered. Specifically, the research explores how the mixed integer approach can handle discrete decisions like crop planting choices and land allocation, while the genetic algorithm helps in finding optimal solutions that respect dynamic and nonlinear constraints, like overproduction and market demand fluctuations. This combined approach is a novel and effective method for optimizing planting strategies under uncertain and changing conditions, ultimately contributing to more sustainable agricultural practices in the region.

3. Experments

This study aims to optimize agricultural planting planning under the constraints of limited and fragmented arable land in the North China mountainous regions by formulating a mixed-integer programming (MIP) model and solving it using a genetic algorithm, with the objective of maximizing crop planting revenue. The overall experimental process is designed as follows: 1) Data collection and current state analysis; 2) formulation of a revenue maximization model based on mixed-integer programming (MIP), considering two scenarios: unsold excess production and discounted sale of excess production; 3) solving the models using a genetic algorithm; and 4) Comparative result analysis and strategy recommendations. Initially, a statistical analysis of the planting areas and yields of various crops in 2023 is conducted. Based on these data, a MIP model is constructed that incorporates both continuous decision variables (e.g., planting area) and binary variables (e.g., whether a crop is planted on a plot), along with corresponding objective functions and constraints. Thereafter, a genetic algorithm is employed to search for the optimal planting schemes. Finally, a series of conclusions regarding optimal planting strategies are drawn through detailed result analysis.

3.1. Current Planting Area and Yield Analysis

Before model formulation, it is essential to perform a comprehensive statistical analysis of the planting areas and actual yields per acre of various crops in 2023 (see Figures 1 and 2). These data provide a robust empirical foundation for the subsequent development of the mixed-integer programming model.

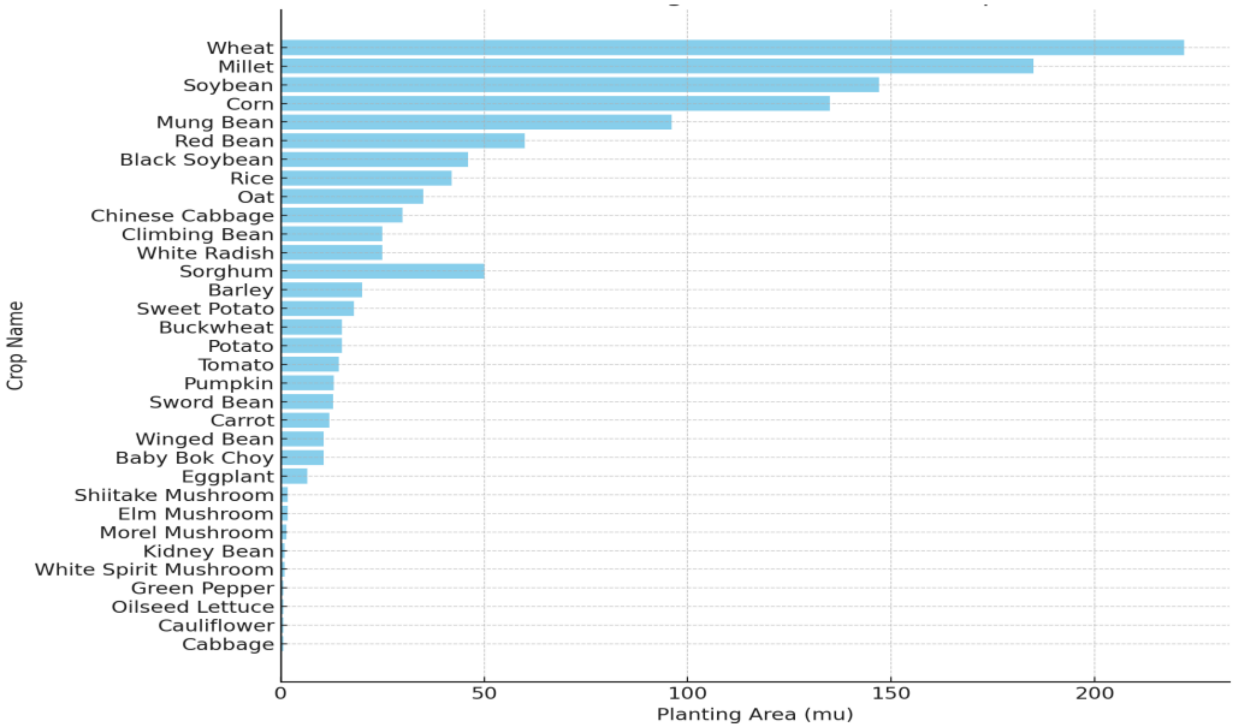

Figure 1. Distribution of Planting Areas for Various Crops in 2023.

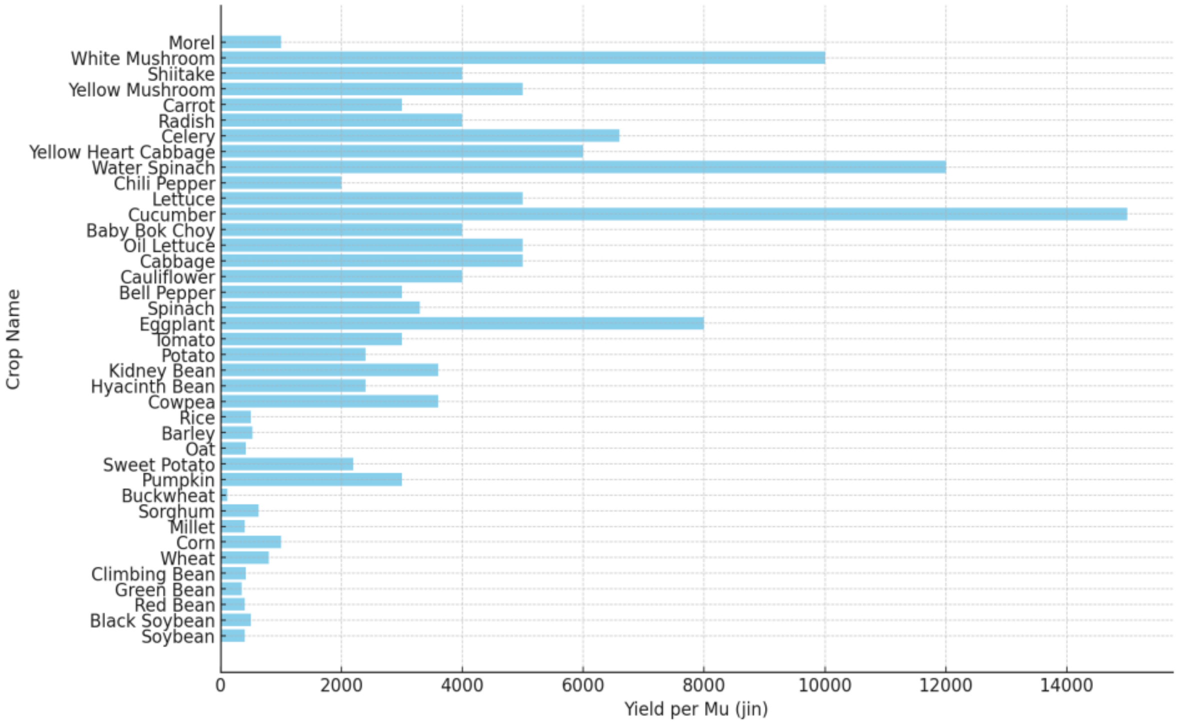

Figure 2. Distribution of Crop Yield per Acre in 2023.

For instance, the figures indicate that although certain leguminous crops cover large planting areas, their yield per acre is relatively low; whereas, crops such as corn and some cash crops, despite being cultivated on smaller areas, exhibit high yields per acre. This demonstrates that crop yield is influenced not only by the planting area but also by inherent growth characteristics and market demand. Consequently, the subsequent model will integrate these factors to achieve revenue maximization.

3.2. Scenario One: Revenue Maximization under Unsold Produce Conditions

In this experiment, we establish a revenue maximization model for the scenario in which the production exceeding the expected sales volume remains unsold. The model requires that the planting strategy strictly conforms to market demand in order to minimize waste.

3.2.1. Decision Variable

The decision variable is defined as the area allocated for planting crop c on plot i in season j of year t, denoted by x_{ijt}, where: i: Plot index (a total of 34 plots including flat lands, terraces, hillsides, and irrigated lands); j: Season index (either the first or second season); t: Year, ranging from 2024 to 2030; c: Crop type (e.g., staples, rice, vegetables, legumes, etc.).

3.2.2. Objective Function

In Scenario One, the objective function is formulated as the actual sales revenue minus the planting cost [5]. The formula is given as:

$$\text{Maximize } Z = \sum_{t=2024} \sum_i \sum_j \sum_c \min\left(y_{cjt} \cdot x_{ijt}, D_{cjt}\right) \cdot p_{cjt} - \sum_{t=2024}^{2030} \sum_i \sum_j \sum_c x_{ijt} \cdot k_{cjt} \tag{1}$$

Where: y_{cjt} is the yield per acre of crop c in season j of year t, assumed to be relatively stable based on 2023 data; p_{cjt} is the selling price of crop c in season j of year t, assumed to be the same as in 2023; D_{cjt} is the expected sales volume of crop c in season j of year t; k_{cjt} is the planting cost per acre of crop c in season j of year t, which includes labor, fertilizer, water resources, etc. The first term, min $(y_{cjt} \cdot x_{ijt}, D_{cjt}) \cdot p_{cjt}$, represents the revenue from the sold portion (capped by market demand), while the second term, $\sum x_{ijt} \cdot k_{cjt}$, and represents the total planting cost.

3.2.3. Constraints

To ensure the model's feasibility, the following constraints are set in the experiment [6]:

(1) Land Area Constraint [7]: The total planting area on each plot must not exceed its available area. In other words, the total area allocated for crops on each plot should not surpass the maximum plantable area A_i. Here, A_i represents the total area of plot i. Based on the above description, this constraint can be expressed by the following formula:

$$\sum_c x_{ijt} \leq A_i, \quad \forall i, j, t \tag{2}$$

(2) Crop Dispersion Constraint [8]: For the sake of facilitating more efficient management of the crops, there is a restriction on the number of plots on which each crop can be planted. To enforce this, we introduce a binary (0-1) variable z_{icjt}. Here, z_{icjt} equals 1 when crop c is planted on plot i in season t, and 0 otherwise. The variable N_{cjt} represents the upper limit on the number of plots for crop c in a given season. By setting this limit and using the variable z_{icjt}, we can ensure centralized management of each crop's planting across plots:

$$\sum_i z_{icjt} \leq N_{cjt}, \quad \forall c, j, t \tag{3}$$

(3) Crop Adaptability Constraint [9]: Only crops suitable for a given plot are allowed, if crop c is not suitable for planting in plot i:

$$x_{ijt} = 0, \quad \text{if crop } c \text{ is not suitable for planting in plot } i \tag{4}$$

(4) Crop Rotation Constraint [10]: To protect soil health, each plot must plant legumes at least once every three years, where ϵ is a very small value:

$$\sum_{t=t_1}^{t_1+2} \sum_j x_{ijt} \geq \epsilon, \quad \forall i \tag{5}$$

3.3. Scenario Two : Revenue Maximization Model with 50% Discount for Excess Production

In Scenario Two, the experiment considers that even if production exceeds the expected sales volume, the excess portion can be sold at a 50% discount. The model optimizes the total revenue by integrating normal sales and discounted sales. The objective function is modified as:

$$\text{Maximize} Z = \sum_{t=2024}^{2030} \sum_i \sum_j \sum_c \begin{pmatrix} \min(y_{cjt} \cdot x_{ijt}, D_{cjt}) \cdot p_{cjt} \\ +0.5 \cdot \max(0, (y_{cjt} \cdot x_{ijt} - D_{cjt}) \cdot p_{cjt}) \end{pmatrix} \tag{6}$$

This includes an additional term of 0.5 · max (0, ($y_{cjt} \cdot p_{cjt} - D_{cjt}$)), representing that the portion exceeding the expected sales volume is sold at half price, thereby partially offsetting losses due to overproduction.

3.4. Solving the Model Using a Genetic Algorithm

To solve the optimization problems in both scenarios, the experiment employs a Genetic Algorithm (GA) [11]. The steps are as follows:

First, initialize the population: Randomly generate several planting schemes as the initial population, where each scheme consists of crop planting areas on various plots, denoted by $X^k = \{x_{ijt}{}^k\}$.

Then, calculate the fitness function: Compute the fitness value $F(X^k)$ for each scheme. In Scenario One, the fitness function directly corresponds to the objective function:

$$\begin{aligned} F(x^k) = \sum_{t=2024}^{2030} \sum_i \sum_j \sum_c \min(y_{cjt} \cdot x_{ijt}, D_{cjt}) \cdot p_{cjt} \\ - \sum_{t=2024}^{2030} \sum_i \sum_j \sum_c x_{ijt} \cdot k_{cjt} \end{aligned} \tag{7}$$

In Scenario Two, discounted sales revenue is also considered:

$$F(X^k) = \sum_{t=2024}^{2030} \sum_i \sum_j \sum_c \begin{pmatrix} \min(y_{cjt} \cdot x_{ijt} D_{cjt}) \cdot p_{cjt} \\ +0.5 \cdot \max(0, (y_{cjt} \cdot x_{ijt} - D_{cjt}) \cdot p_{cjt}) \end{pmatrix} \tag{8}$$

Next, perform the selection operation: Use a roulette wheel selection mechanism to choose individuals for the next generation based on fitness values:

$$P(X^k) = \frac{F(x^k)}{\sum_{k=1}^{P} F(x^k)} \tag{9}$$

Subsequently, perform the crossover operation: Generate new offspring schemes through single-point or multi-point crossover to introduce new solution space:

$$X^{new} = \alpha X^A + (1-\alpha)X^B, \quad \alpha \in [0,1] \tag{10}$$

Afterwards, perform the mutation operation: Randomly adjust parts of the individual's genes (such as crop type or planting area) to prevent premature convergence to a local optimum, where:

$$x_{ijt}^{\text{mutated}} = x_{ijt}^k + \epsilon, \quad \epsilon \text{ is a small perturbation value} \tag{11}$$

Finally, iterate until termination: Repeat the above operations until the maximum number of iterations is reached or the fitness values converge.

4. Results

The experiment analyzed the optimal planting strategies under both scenarios: 1) Unsold Produce Scenario: The portion exceeding market demand generates no revenue, resulting in greater fluctuations in overall revenue; 2) Discount Scenario: The excess production is sold at a 50% discount, leading to more stable revenue. The comparison of revenue between unsold products and discount sales strategies is shown in Figure 3 below.

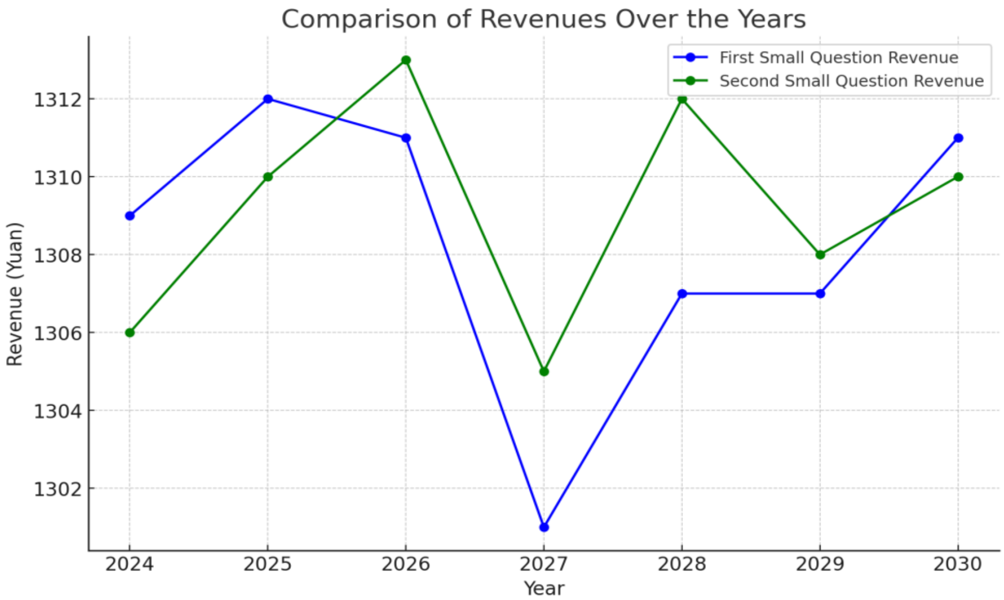

Figure 3. Revenue Comparison between Unsold Produce and Discounted Sales Scenarios.

Figure 3 illustrates the revenue comparison between the two scenarios from 2024 to 2030, where the blue line represents the unsold produce scenario and the orange dashed line represents the discount scenario. The results show that: 1) From 2025 to 2026, revenue in both scenarios increases, but in the unsold produce scenario, revenue peaks in 2026 and then sharply declines in 2027; 2) In 2027, revenue in the unsold scenario falls to a minimum (approximately 1301 yuan), reflecting the risk of overproduction, whereas the discount scenario maintains relative stability due to partial sales; 3) From 2028 to 2030, although revenue in the unsold scenario recovers somewhat, its overall stability is still inferior to that of the discount scenario.

During the experiment, we conducted a detailed analysis of the optimal planting strategies under both scenarios and provided numerical results for selected strategies in Table.1.

Table 1. Selected Planting Strategies (Excerpt).

Season	Field Name	Soybean	Black Soybean	Red Bean	Mung Bean	Climbing Bean	……
First	A1	0	0	8.0565	0	0	……
First	A2	0	5.4453	5.3032	0	5.2762	……
First	A3	0	3.6073	3.5965	0	3.5284	……
First	A4	0	7.4580	0	7.3386	6.9429	……
First	A5	6.8990	0	6.4563	0	6.9554	……
First	A6	5.6623	0	5.4803	5.2624	0	……
First	B1	5.8019	6.1337	0	6.1783	6.0433	……
First	B2	0	4.8144	4.7063	0	4.3814	……
First	B3	0	3.8077	3.9705	0	0	……
First	B4	2.7538	0	0	2.7818	0	……
……	……	……	……	……	……	……	……

In summary, the analysis indicates that the discount scenario, with its smaller revenue fluctuations, is a more advantageous strategy for coping with market variations. Based on the experimental results, it is recommended that in cases of overproduction, a discount sales strategy should be adopted to stabilize overall revenue and reduce losses caused by unsold produce.

5. Conclusions

In this study, we focused on optimizing revenue under conditions of overproduction by investigating two distinct scenarios: unsold produce and discounted sales. By formulating a mixed-integer programming model with a revenue maximization objective, we derived the optimal planting strategies for both cases. In the unsold scenario, any production exceeding the expected sales volume remains unsold, resulting in significant revenue fluctuations; indeed, our genetic algorithm solution revealed that revenue in 2027 dropped to a minimum of 1301 yuan. In contrast, under the discounted scenario, excess production is sold at 50% of the market price, a strategy that notably reduced revenue volatility, with revenue peaking at 1313 yuan in 2026. These results demonstrate that the discount strategy effectively mitigates the risks associated with unsold produce and enhances revenue stability.

References

[1] Zhou L G, Liu T, Lu J Z. Traversal Path Planning for Farmland in Hilly Areas Based on Floyd and Improved Genetic Algorithm [J]. Smart Agriculture, 2023, 5(4): 45-57.

[2] Han J, Lin N, Ruan J, Wang X, Wei W, Lu H. A Model for Joint Planning of Production and Distribution of Fresh Produce in Agricultural Internet of Things [J]. IEEE Internet of Things Journal, 2021, 8(12): 9683-9696.

[3] Wei Jiahui, Yin Limin. Optimal Capacity Configuration of Integrated Energy Systems Based on Linear Mixed-Integer Programming [J]. Jilin Electric Power, 2022, 50(05): 41-45.

[4] Shen Mingbin. Research on Agricultural Machinery Operation Path Optimization Based on Genetic Algorithm [J]. Agricultural Machinery Usage and Maintenance, 2025, (01): 15-18.

[5] Xia Tian, Shen Xinwei, Shang Yuwei. Multi-objective Optimization Decision and Sensitivity Analysis of Distribution Network Maintenance Based on Mixed-Integer Programming [J]. Power System Technology, 2024, 48(11): 4680-4689.

[6] Song Yan. Rational Planning and Utilization of Agricultural Land Resources [J]. Hebei Agricultural Machinery, 2024, (18): 69-71.

[7] Zhang Gaiqing, Feng Junwei. Research on the Choice Behavior of Grain Farmers in Outsourcing Agricultural Machinery Services under Heterogeneous Resource Constraints — An Analysis Based on the Perspectives of Land, Labor, and Ecological Resource Constraints [J]. Agricultural Economics and Management, 2023, (06): 44-58.

[8] Guan Haipeng, Ren Yan, Zhao Qiu Xia. Decentralized Agricultural Product Production and Sales Logistics Model Based on Blockchain [J]. Journal of Yuncheng University, 2021, 39(06): 36-41.

[9] Lü Xiao. Decoding the Seed 'Chip' to Safeguard Food Security at the State Key Laboratory of Crop Stress Adaptation and Improvement, Henan University [N]. Henan Science and Technology News, 2024-06-28(008).

[10] Sun Lei. Response of Soil Microbes to Fertilization and Crop Rotation in Typical Black Soil Regions [D]. Northeast Forestry University, 2023.

[11] Yu Qiuhong. Application of Improved Genetic Algorithm in Computer Mathematical Modeling [J]. Information Systems Engineering, 2024, (09): 59-62.