Hybrid Integer Programming Model for Optimizing Crop Planting Strategies Considering Dynamic Market and Seasonal Factors

Weijie Ruan [#, *], Zihan Wu [#], Ziyan Yang [#]

School of Information Engineering, China University of Geosciences, Beijing, China, 100083

* Corresponding Author Email: 1004222124@email.cugb.edu.cn

[#]These authors contributed equally.

Abstract. This study explores the challenge of determining optimal crop allocation across various land types in the context of environmental degradation and resource scarcity. Focusing on the northern regions of China, this paper proposes a mixed-integer programming model designed to account for seasonal variations and market fluctuations, with the goal of maximizing cumulative planting income over multiple years. The model integrates a comprehensive analysis of both economic and environmental factors, including crop yield, costs, prices, and associated risks. In addition, the research employs entropy weighting and TOPSIS methods for effective risk assessment to address uncertainty. Validation of the model is conducted using datasets from several villages in China, highlighting its capability to predict income fluctuations while providing practical insights for agricultural planning. The findings indicate that this model not only stabilizes income but also improves resource efficiency and offers valuable references for effective agricultural market regulation. By addressing the complexities of crop allocation, this research contributes to sustainable agricultural practices and resource management strategies in the face of ongoing environmental challenges.

Keywords: Crop Planting Optimization, Dynamic Market Factors, Seasonal Variability, Hybrid Integer Programming.

1. Introduction

Under the current global challenges of environmental degradation, resource scarcity and escalating climate change [1], traditional agricultural land development practices have exposed significant issues related to soil quality, as well as problems associated with climates, pests and economic income. These severely impede farmers' ability to improve their living conditions and restrict the sustainable development of rural agricultural economies. Consequently, the efficient and organic utilization of limited arable land resources represents an urgent and pressing concern.

Organic farming refers to the cultivation of crops without the application of synthetic fertilizers or pesticides, utilizing methods such as soil enhancement and the promotion of soil biodiversity to maximize environmental protection and ecological integrity [2]. As of the end of 2023, data released by the National Bureau of Statistics indicates that China has approximately 0.315 billion acres of arable land, with the rural population constituting 33.8% of the total national population [3]. Given this context, the promotion of organic farming practices and the implementation of localized planting strategies are critical for enhancing the income of local farmers and facilitating rural prosperity.

This study aims to address the challenge of determining optimal crop distribution across different types of land, considering seasonal and market fluctuations, and aims to maximize overall planting income. The research will focus on North China, where agriculture is impacted by complex constraints, such as climate variability, fluctuating market demands, limited arable land, and the need for efficient land utilization to enhance productivity and minimize certain risks.

Esteso et al. proposed a centralized multi-objective mathematical programming model to support the sustainable crop planning definition for a specific region. This approach seeks to substantially reduce crop waste and the inequities among farmers while sacrificing only a small amount of supply chain profits [4]. Yang et al. utilized geographic distribution data for various crops within the Songhua

River Basin to develop a maximum entropy single-objective optimization model, thus optimizing the crop planting structure within the area [5]. In addressing the issue of irrational agricultural planting structures in northern China, Yu et al. designed a multi-objective optimization model that integrates the reduction of irrigation areas, control of crop scale, and conservation of irrigation water. Their work resulted in the construction of an optimal agricultural planting structure following the streamlining of constraints [6]. However, these existing studies concentrate primarily on objective optimization without fully addressing the variability in market demands and environmental constraints over multiple seasons. Additionally, some of the methods lack the robustness required for real-world agricultural applications, particularly in regions with varying land types and crop compatibility requirements. Therefore, a gap remains in achieving a model that integrates dynamic market and seasonal factors with specific constraints in agricultural settings.

To address the identified weaknesses, this research proposes the development of a hybrid integer programming model that aims to optimize crop planting strategies across various land types, including dry fields, terraces, irrigated lands, and greenhouses. The model's primary objective is to maximize overall income over the coming years while fulfilling crucial constraints related to crop rotation and yield stability. Additionally, the study will examine the impacts of potential overproduction and subsequent price adjustments, thereby providing insights into the influence of dynamic market factors on income variability.

This research is guided by two fundamental questions: how to determine the maximum potential planting income in the region over the next few years, assuming that excess production of local crops may render them unsellable if all indicators remain stable; and how to assess the maximum possible planting income while accounting for uncertainties in various crop indicators and potential local planting risks.To contribute meaningfully to this field, the study will present a multi-planting strategy model that integrates economic and environmental considerations, significantly enhancing its accuracy through risk assessment techniques that employ entropy and the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) algorithm. In doing so, this research will offer valuable insights for rural agricultural planning and policies aimed at achieving income stability and optimizing resource efficiency.

2. Methodology

2.1. Dataset Description

The dataset used in this study originates from Attachments 1 and 2 of Problem C in the 2024 China Undergraduate Mathematical Contest in Modeling hosted by Higher Education Press. Detailed information can be found at https://www.mcm.edu.cn/. The dataset includes the basic information of the existing cultivated land and crops in a village. And statistical data related to crop cultivation in the village in 2023. In the dataset, there are 1201 mu of open-air cultivated land in the village, which are scattered into 34 plots, including flat dry land, terraces, hillside land and irrigated land greenhouses. In addition, there are 16 ordinary greenhouses and 4 intelligent greenhouses in the countryside, each with an area of 0.6 mu. At the same time, the experimental data set also provides the planting situation and related statistical data of the rural crops in 2023. It includes detailed information such as planting area, yield per mu, total output, planting cost, selling price and sales volume of various crops. The crop data cover the yield and economic benefits of different plots of grain crops, vegetables, rice and edible fungi. It provides an important basis for planting planning and optimization scheme in 2024-2030.

2.2. Data Encoding

The data presented in this dataset have textual properties and are not suitable for input to mathematical models. Therefore, it is necessary to quantify the text-based indicators. The corresponding results of each variable after coding are shown in Table.1-3.

Table 1. Corresponding Table of Crop Planting Period t and Each Year.

Planting period	First quarter of 2024	The second season of 2024	First quarter of 2025	The second quarter of 2025	...	First quarter of 2030	Second quarter of 2030
Number	1	2	3	4	...	13	14

Table 2. Corresponding Table of Various Crops.

Crop name	Soybeans	Black beans	Red Bean	Mung bean	...	Pleurotus nebrodensis	Morchella
Number	1	2	3	4	...	40	41

Table 3. Corresponding Table of Plot Used for Crop Planting.

Planting plots	A1	A2	A3	A4		F4-1	F4-2
Number	1	2	3	4	...	73	74

2.3. Establishment of Mixed Integer Programming Planting Model

2.3.1. Construction of Decision Variables and Objective Function [7]

In this paper, a mixed programming model with single objective function is established. The maximum total income of each crop planting corresponding to the season time *t* in 2024-2030 is taken as the objective function to solve the optimal planting scheme. The decision variable to be solved is x_{ijt}, which represents the planting area of a certain crop *i* in a certain land *j* in the *t*-th season. The total revenue can be calculated as the difference between the total sales of the crop in each growing area and the total cost of growing the crop. Influenced by factors such as climate environment and market conditions [8], it is also necessary to consider changes in the expected sales volume and sales price of crops:

(1) The expected sales of wheat and corn in 2024-2030 show an increasing trend, with an average annual growth rate between 5% and 10%. The expected annual sales of other crops in 2024-2030 will vary by about ± 5% relative to 2023.

(2) Crop yields per acre tend to be erratic, varying by ± 10% per year.

(3) The cost of planting crops will increase by about 5% annually on average.

(4) In the future, the selling price of grain crops will be basically stable compared with 2023.The average annual growth rate of sales price of vegetable crops in the future is about 5%.In the future, the average annual decrease rate of the sales price of edible fungi will be 1% -5% (especially 5% for Morchella).

To simplify the solving process of the model, this paper assumes that the expected sales growth rate of the second quarter of each year can represent the expected sales growth rate of each year. The average quarterly change rate of each indicator is the same as the average annual change rate, and the average quarterly growth rate is the same as the average annual growth rate. This paper assumes that if the yield of a crop exceeds a given expected sales volume, the excess will be sold at an average selling price of 50%.

2.3.2. Entropy TOPSIS Method to Determine Additional Planting Risks [9]

The final yield per mu of various crops in each season usually has a certain randomness. It is related to the uncertainty of the expected sales volume, yield per mu, planting cost and sales price of crops. It is also related to the potential risk of local planting. This paper also considers different crop planting risk factors to determine a reasonable yield change rate per mu. And that model establish mode and the solving result are more close to real life. Natural risks include climate risk, biological risk and soil risk. Human risks include management, capital and technology risk [10]. Relative score of each planting risk factor:

$$\xi_i = \frac{w_i - (w_i)_{\min}}{(w_i)_{\max} - (w_i)_{\min}} \tag{1}$$

Obviously, the relative score of each planting risk factor is a very small index, and the score results should be positively and standardized. To simplify the solution of the model, it is assumed that the weight of each planting risk assessment object is the same. Converting ξ_i, the scores of all the planting risk factors into maximality indexes to obtain the converted score result. If the scores of the six-planting risk influencing factors are $\left[\xi_1', \xi_2', \xi_3', \xi_4', \xi_5', \xi_6'\right]^T$, the scores after standardization are:

$$\xi_i'' = \frac{\xi_i'}{\sqrt{(\xi_1')^2 + (\xi_2')^2 + (\xi_3')^2 + (\xi_4')^2 + (\xi_5')^2 + (\xi_6')^2}}, \ i = 1,2,3,\cdots,6 \tag{2}$$

In this model, there are six evaluation objects and one evaluation index, and the matrix is a one-dimensional column vector. Normalize it to get the matrix $Z = \left[z_{11}, z_{12}, z_{13}, z_{14}, z_{15}, z_{16}\right]^T$ .Then, the maximum and minimum values are selected from the indicators to form the maximum vector Z^+ and the minimum vector Z^-. For the i-th planting risk influencing factor, the distance D_i^+ of each index from the maximum value in the matrix Z and the distance D_i^- from the minimum value are calculated.

Then define the score of the object $S_i = \frac{D_i^-}{D_i^+ + D_i^-}$ and the result of normalization processing $S_i' = \frac{S_i}{\sum_{i=1}^{6} S_i}$, and the weight of each index can be obtained. The evaluation results after TOPSIS scoring and normalization are shown in Table.4.

Table 4. Risk Type Score after Normalized Treatment.

Factor number	Risk type I	Normalized evaluation S_i'
1	Climate risk	0.3530
2	Biological risk	0.1175
3	Soil risk	0.0000
4	Manage risk	0.1175
5	Capital risk	0.2354
6	Technical risk	0.1765

Based on the initial factor evaluation results in the table, combined with the weights of each planting risk factor, the additional reduction rate in yield due to potential planting risks in the area can be comprehensively calculated $m_i' = \sum_{i=1}^{6} m_i \cdot S_i'$, resulting in $m_i' \approx 0.06$. This indicates that, with the additional reduction rate included, the crop yield variation rate m_i should vary within the range of [-0.16, 0.04]. It is assumed that crops planted in both open-field and greenhouse areas will experience additional yield reductions in the future, and that the reduction rates are identical.

2.3.3. Determine the Constraint Conditions

According to the data set information, each plot (including cultivated land and greenhouse) can only grow crops suitable for the type of the plot. And there are detailed season and category requirements. Based on this, this paper formulates a more detailed constraint condition.

(1) Flat dry land, terraced fields and hillside land are only suitable for planting grain crops (except rice) in a single season (the first season) every year. The irrigated land can grow rice in a single season (the first season) every year.

(2) If two seasons of vegetables are planted in the irrigated area, a variety of vegetables (except Chinese cabbage, white radish and red radish) can be planted in the first season. In the second season, only one of Chinese cabbage, white radish and red radish can be planted.

(3) Irrigated land can be planted with rice in a single season (the first season) or vegetable crops in two seasons every year.

(4) Ordinary greenhouses grow two crops a year, and the first season can grow a variety of vegetables (except Chinese cabbage, white radish and carrot). In the second season, only edible fungi can be planted.

(5) Intelligent greenhouse can grow two seasons of vegetables every year (except Chinese cabbage, white radish and carrot). The same crop cannot be planted on the same plot in two consecutive seasons, and legumes should be planted at least once in each plot within three years. The planting area of each crop on a single plot should not be too small.

To establish the model conveniently, this paper takes the minimum value x_{min} of the planting area of a certain crop in the region as 0.3. That is, half of the minimum planting area of crops in the common greenhouse and intelligent greenhouse areas in 2023 in the data set. The variable y_{ijt} is set as 0-1, representing whether a certain crop *i* is planted in a certain field *j* area in the *t*-th season. If $y_{ijt}=1$, it means that the crop is planted in the planting area in a certain period. Otherwise, it is not planted.

2.3.4. Determine a Single-objective Mixed Integer Programming Model

Synthesizing the establishment of the optimization objective function and the constraint condition under the two different crop yield excess conditions, the single-objective mixed integer programming model established in this paper is as follows:

$$\max z=\sum_{i=1}^{41}\sum_{j=1}^{54}\sum_{t=1}^{14}(\min(P'_{it},D'_{it})\cdot p'_{it}+0.5\cdot\max(0,P'_{it}-D'_{it})\cdot p'_{it}-P'_{it}\cdot c'_{it}) \tag{3}$$

$$s.t.\begin{cases} y_{ijt}=0\ \ or\ 1,\ i=1,2,3,\cdots 41;\ \ j=1,2,3,\cdots 74;\ \ t=1,2,3,\cdots,14 \\ \sum_{i=1}^{41}x_{ijt}\le A_j,\ j=1,2,3,\cdots,74;\ \ t=1,2,3,\cdots,14 \\ \sum_{i=1}^{41}x_{ijt}\cdot x_{ij,t+1}=0,\ j=1,2,3,\cdots,66;\ \ t=1,3,5,\cdots,13 \\ \sum_{i=17}^{34}x_{ijt}\cdot x_{ij,t+1}\le 2A_j,\ j=67,68,69,\cdots,74;\ \ t=1,3,5,\cdots,13 \\ \sum_{i=1}^{41}(y_{ijt}+y_{ij,t+1})\le 1,\ j=1,2,3,\cdots,74;\ \ t=1,3,5,\cdots,13 \\ \sum_{i=16}y_{ijt}+\dfrac{\sum_{i=17}^{34}(y_{ijt}+y_{ij,t+1})+\sum_{i=35}^{37}(y_{ijt}+y_{ij,t+1})}{2}\le 1,\ j=27,28,29,\cdots 34;\ \ t=1,3,5,\cdots 13 \\ \sum_{t}^{t+6}y_{ijt}\ge 1,\ i\in\{1,2,3,4,5,17,18,19\}\ ;\ \ j=1,2,3,\cdots,74;\ \ t=1,2,3,\cdots,7 \\ x_{ijt}\ge\min\{0.3,y_{ijt}\}\ ,\ \ i=1,2,3,\cdots,41;\ \ j=1,2,3,\cdots,74;\ \ t=1,2,3,\cdots,14 \end{cases} \tag{4}$$

Among $P_{it}=x_{ijt}\cdot q_i$, where z represents the total sales revenue from crops, *i* is the number of crop types, *j* is the number of plots used for crop planting (including various types of arable land and greenhouse information from the attachments), and *t* is the number of seasons to be solved. c_i Represents the cost per mu for different crops, q_i denotes the yield per mu for each crop type, and A_j represents the different squares for different areas. Additionally, the relationship between the crop yield P_{it} and the given expected sales volume D_i needs to be considered. If $P_{it}>D_i$, the excess yield is not included in the total revenue. Conversely, the actual production volume is used for calculation when $P_{it} < D_i$.

2.4. Design of Randomized Greedy Algorithm [11]

Given the numerous constraints on decision variable x_{ijt} in the model developed in this paper, it is challenging to solve using traditional mixed integer programming models. Therefore, this paper adopts an improved randomized greedy algorithm for solution.

2.4.1. Randomized Greedy Algorithm Design Introduction

The randomized optimization algorithm aims to find the optimal planting combination by randomly trying different crop allocations. Specifically, for each plot and season, the algorithm randomly selects a crop from the available options and calculates the resulting profit. After many random attempts, the combination with the highest profit is retained.

In choosing the best crop combination, the greedy algorithm makes a local optimal choice at each step. It picks the crop that provides the highest immediate profit. During the profit calculation and combination selection process, the code compares the current profit and selects the most optimal crop to update the planting plan.

2.4.2. Constraint Condition Correction Strategy

The code includes several checks and correction steps for various planting constraints to ensure that the planting plan meets the real-world agricultural requirements. For example, there is a rule that mandates the planting of legume crops every six seasons. This is enforced using condition checks and coefficient adjustments, ensuring all constraints are satisfied.

2.4.3. Profit Calculation and Optimization

The objective function for total profit combines factors such as expected sales volume, sales price, and planting costs. By comparing actual and expected yields, the algorithm avoids waste while maximizing revenue. This step is essentially an economic optimization problem where the algorithm calculates the economic benefit of each possible scenario and selects the optimal one.

3. Results

3.1. The Establishment of Mixed Integer Programming Model

In this paper, the total income of crops from 2024 to 2030 is taken as the objective function. Considering the expected sales volume of various crops, the yield per acre, the uncertainty of planting cost and selling price, and the potential planting wind insurance, with computer aided calculation, the total income in seven years can reach 253162344.10 yuan. The details of some planting results are shown in Figure.1. and Table.5.

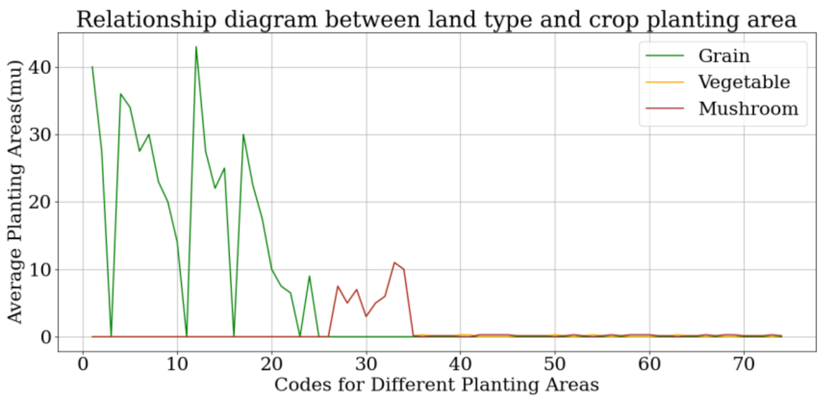

Figure 1. Average Planting Area of Various Crops in Different Plots.

Table 5. Planting Situation of First Season in 2024 (A1 ~ B9, Soybean ~ Sorghum).

Parcel name	Soy bean s	Black beans	Red beans	Mung beans	Climbin g beans	Whea t	Cor n	Mille t	Sorghu m
A1	0	0	0	0	0	0	80	0	0
A2	0	55	0	0	0	0	0	0	0
A3	0	35	0	0	0	0	0	0	0
A4	0	0	72	0	0	0	0	0	0
A5	0	0	0	0	0	0	0	0	68
A6	0	0	0	0	0	0	0	0	55
B1	0	0	0	0	0	0	0	0	60
B2	0	0	0	0	0	0	0	0	46
B3	0	40	0	0	0	0	0	0	0
B4	0	0	0	0	0	0	0	0	0
B5	0	0	0	0	0	0	0	0	25
B6	0	0	0	0	0	0	86	0	0
B7	0	0	0	0	0	0	55	0	0
B8	0	44	0	0	0	0	0	0	0
B9	0	0	0	0	0	0	0	0	50

3.2. Analysis of Experimental Results

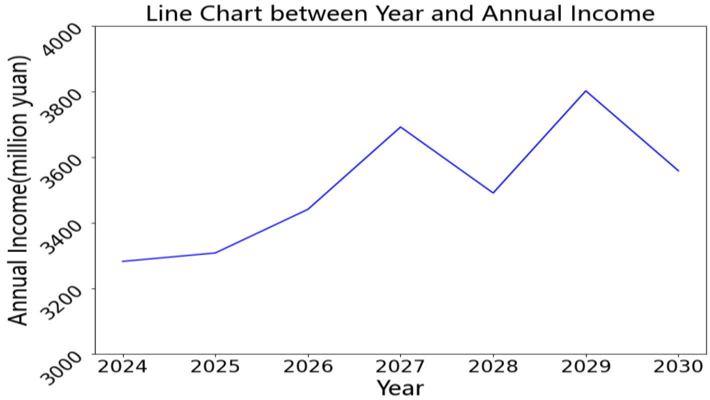

Figure 2. Total Income from Crop Cultivation in Each Year.

It can be seen from the total income of crop planting in each year in Figure.2. that the volatility of the model after adding the random item is smaller, which reflects good stability. The change in annual income was observed by adjusting the cost of growing all crops, keeping the other parameters constant. The results of the sensitivity analysis are shown in Figure.3. and Table.6. respectively.

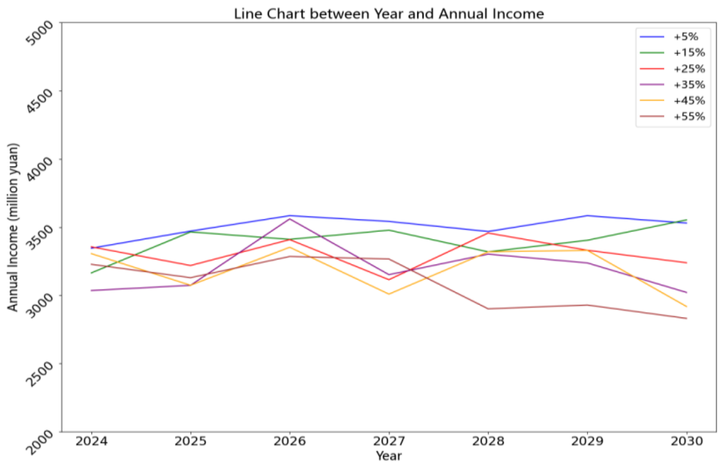

Figure.3. Sensitivity Analysis Visualization.

Table 6. Relationship between the Rate of Cost Increasing and the Average Value of Total income of Each Year.

Cost improvement rate (%)	Average annual profit (million yuan)
5%	35.0316
15%	33.9869
25%	33.0302
35%	31.9699
45%	31.8644
55%	30.8087

The impact of cost changes on revenue shows a linear relationship. For every 10% increasing in cost, the benefit drops by about 5-8%. Compared with the sensitivity of price and output, the impact of cost on revenue is relatively small, but it accounts for a larger proportion of the total revenue. The rising cost of planting will weaken the income of planting. Therefore, controlling the input of cost and optimizing the efficiency of agricultural use are important ways to improve the income. Especially for high-input crops (such as edible fungi), the impact of cost fluctuation on income is more obvious.

4. Conclusions

The mixed-integer programming model developed in this study is fundamentally based on the yield, cost, and pricing of agricultural crops. It thoroughly integrates practical constraints such as land type, planting seasons, and crop rotation, thereby accurately mirroring the agricultural production realities of the region while effectively simulating the influence of market fluctuations on planting returns. Although the model encompasses a variety of dynamic factors, including environmental and socio-economic elements, it necessitates an expansion of variable dimensions and a recalibration of constraints to adequately respond to unforeseen events, such as policy changes and fluctuations in sales channels. The optimized model has significant potential to enhance the predictive accuracy of supply trends for agricultural products within a specific region, offering invaluable insights for regulating supply and demand within the agricultural market. Consequently, this may contribute to mitigating market volatility arising from long-term supply-demand imbalances.

Looking to the future, research endeavors should prioritize enhancing the adaptability of the model through the incorporation of sophisticated machine learning techniques capable of accommodating real-time data inputs and responding to unexpected disruptions. Furthermore, investigating the integration of climate change impacts and socio-economic variables, such as labor availability and shifts in consumer behavior, may significantly refine the predictive capabilities of the model. Finally, examining the implications of alternative agricultural practices and technologies, including precision agriculture and sustainable farming practices, will be crucial for developing strategies that not only optimize crop yields but also promote environmental sustainability and resilience within agricultural systems.

References

[1] Ondrasek G, Horvatinec J, Kovačić M B, et al. Land Resources in organic Agriculture: Trends and challenges in the twenty-first century from global to Croatian contexts[J]. Agronomy, 2023, 13(6): 1544.

[2] Panday D, Bhusal N, Das S, et al. Rooted in nature: The rise, challenges, and potential of organic farming and fertilizers in agroecosystems[J]. Sustainability, 2024, 16(4): 1530.

[3] National Bureau of Statistics of China. Statistical Communiqué of the People’s Republic of China on the 2021 National Economic and Social Development, February 28, 2022.

[4] Esteso A, Alemany M M E, Ortiz A, et al. Optimization model to support sustainable crop planning for reducing unfairness among farmers[J]. Central European Journal of Operations Research, 2022, 30(3): 1101-1127.

[5] Yang S, Wang H, Tong J, et al. Impacts of environment and human activity on grid-scale land cropping suitability and optimization of planting structure, measured based on the MaxEnt model[J]. Science of the Total Environment, 2022, 836: 155356.

[6] Yu H, Liu K, Bai Y, et al. The agricultural planting structure adjustment based on water footprint and multi-objective optimisation models in China[J]. Journal of Cleaner Production, 2021, 297: 126646.

[7] Zhang J, Liu C, Li X, et al. A survey for solving mixed integer programming via machine learning[J]. Neurocomputing, 2023, 519: 205-217.

[8] Lahlali R, Taoussi M, Laasli S E, et al. Effects of climate change on plant pathogens and host-pathogen interactions[J]. Crop and Environment, 2024, 3(3): 159-170.

[9] Awodi N J, Liu Y, Ayo-Imoru R M, ET al.Fuzzy TOPSIS-based risk assessment model for effective nuclear decommissioning risk management [J]. Progress in Nuclear Energy, 2023, 155: 104524.

[10] Hou J, Chen L, Han B, ET al.Distribution characteristics and risk assessment of neonicotinoid insecticides in planting soils of mainland China [J]. Science of the Total Environment, 2023, 902: 166000.

[11] Gao W, Friedrich T, Neumann F, ET al.Randomized greedy algorithms for covering problems[C]. Proceedings of the Genetic and Evolutionary Computation Conference. 2018: 309-315.