Optimal Planting Strategy Model for Crops in Mountainous Areas of North China Based on Robust Optimization

Shenyang Li [*,†], Kehan Li [†], Yao Sun [†], Lixiang Gu [†], Jingjing Fu [†]

School of Software, Xinjiang University, Urumqi, China, 830091

* Corresponding Author Email: lishenyang229@163.com

[†] These authors contributed equally.

Abstract. With the rural revitalization strategy's deepening, North China's mountainous rural economy needs to shift from traditional to modern efficient agriculture. This paper proposes a robust optimization-based optimal crop planting strategy model for the region's unique conditions. It first introduces the area's topography, climate, and arable land resources, then presents a combined linear programming and robust optimization model. Model performance is validated through data preprocessing and market analysis, and its effectiveness is shown by comparison with traditional strategies. The study also analyzes crop substitutability and complementarity, offering practical recommendations for North China's mountainous areas. Results indicate that robust optimization can handle market and climate uncertainties, ensuring planting scheme efficiency. This research provides scientific support for agricultural decision-making in these regions and offers theoretical and practical references for similar agricultural optimization issues.

Keywords: Robust Optimization, Correlation Analysis, Mathematical Planning, Crop Complementarity, Optimal Planting Strategy.

1. Introduction

With the national rural revitalization strategy, the rural economy's sustainable development is crucial. The shift from traditional to modern, efficient agriculture is urgent. Organic planting can enhance product value, meet market demands for quality and green products, and improve the rural ecological environment, achieving economic and ecological benefits. In North China's mountainous areas, unique geographic and climatic conditions, such as low temperatures and diverse arable land types, require tailored agricultural practices. The limited land necessitates efficient management. Developing organic planting according to local conditions is key to sustainable rural economic development. Growing consumer interest in food safety and healthy diets also presents a vast market potential for organic products.

Developing the organic planting industry in rural areas involves complex decision-making. Maximizing yield, minimizing risk, and optimizing ecological benefits under limited land resources is a multi-objective optimization problem. Traditional planting methods are insufficient, while operations research optimization algorithms offer a scientific approach. By considering crop growth patterns, land constraints, and market demands, these algorithms can find optimal planting plans, supporting rural organic planting industry development.

In the study of optimal planting scheme in North China, optimization method and model design are key. In recent years, researchers have used a variety of approaches to cope with the complexity and uncertainty of agricultural systems. Jiaying Wang and Yiquan Wang [1] proposed a multi-stage crop planting strategy optimization model based on particle swarm optimization algorithm (PSO), which dynamically adjusts the strategy by integrating factors such as yields, market prices, and crop rotations, in order to maximize the economic benefits and maintain a sustainable agriculture. In this paper, we present a multi-stage crop planting strategy optimization model based on particle swarm optimization algorithm (PSO) to maximize the economic benefits and maintain a sustainable agriculture. Jose Miguel F. Custodio et al. [2] used linear programming to optimize crop harvesting schedules and land allocation to handle large-scale problems and provide exact solutions.

Haitao Ye et al. [3] combined mixed integer programming with Monte Carlo simulation to consider the effects of climate and market conditions. Kun Liang et al. [4] proposed a linear programming model based on an improved genetic algorithm to increase the solution efficiency and introduced an orthogonal experimental design to reduce the consumption of computational resources. Shi Chen and Kengtai Peng [5] proposed a linear programming model based on the Sample Average Approximation (SAA) dynamic planning model to improve the robustness of the scheme. Changcai Guo et al. [6] solved the single-objective profit maximization problem through a mixed integer programming approach emphasizing the effects of irrigation and greenhouse planting area.

Smart agriculture and prediction techniques are important in planting scheme optimization for precision agriculture management through data-driven models and Internet of Things (IoT) technology. Patel et al. [7] proposed Deep Belief Network (DBN)-based optimization model for predicting crop yields. Shankar and Moorthi [8] developed an IoT-based crop yield prediction framework that combines optimal integrated learning and hybrid optimization models. Pitakaso et al. [9] proposed a decision support system based on Artificial Multiple Intelligence System (AMIS) to balance yield, resource utilization and economic efficiency.

In this paper, based on the analysis of crop planting strategies in the mountainous regions of North China, we propose an optimal planting strategy model based on robust optimization and apply it to agricultural production decision-making in the mountainous regions of North China after verifying the superiority of the model. The main contents of the whole paper can be summarized as follows: firstly, the topography, climate and arable land resources in the mountainous areas of North China are introduced, and a crop planting strategy model based on linear programming and robust optimization is proposed; the performance of the model is verified through data preprocessing and analysis of the market demand, and compared with the traditional planting strategy in order to validate the validity and reasonableness of the improved model; and the subsequent analysis of the crop Substitutability and complementarity are analyzed to construct a comprehensive planting strategy model, which is applied to the actual problems of crop planting in the mountainous areas of North China, and further specific suggestions and recommendations are made.

2. Method

2.1. Linear programming model

Linear programming is a mathematical optimization technique for solving problems that maximize or minimize a linear objective function under given constraints. It is widely used in a variety of fields, including production planning, resource allocation, transportation, and financial investment.

A linear programming problem consists of a linear objective function and a set of linear constraints. The objective function represents the quantity to be optimized (maximized or minimized), while the constraints define the space of feasible solutions. The objective of a linear programming problem is to find a feasible solution that satisfies all constraints and optimizes the objective function.

2.2. Pearson correlation coefficient

Pearson's correlation coefficient is a statistical measure of the degree of linear correlation between two variables, with a value between -1 and 1. When the coefficient is 1, it indicates a perfect positive correlation; when it is -1, it indicates a perfect negative correlation; and when it is 0, it indicates that there is no linear correlation between the two variables.

$$\rho_{XY}=\frac{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)\left(Y_i-\bar{Y}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i-\bar{X}\right)^2}\cdot\sqrt{\sum_{i=1}^{n}\left(Y_i-\bar{Y}\right)^2}} \tag{1}$$

2.3. Robust optimization

Robust Optimization (RO) is an optimization method for dealing with uncertainty, which aims to find solutions that maintain performance in the worst-case scenario. Unlike traditional optimization methods, Robust Optimization does not attempt to predict specific values for uncertain parameters, but rather describes the range of possible values for these parameters by defining the set of uncertainties. In this way, robust optimization ensures that the resulting solution is stable and reliable in the face of uncertainty [10].

Using Robust Optimization to Solve Crop Optimization Problems under Market Volatility. Define the uncertainty set: based on the volatility of expected crop sales, acreage, planting costs and prices, the range of possible variations of these parameters is described. Build a robust optimization model: with the uncertainty set and constraints, an objective function is built and the model is solved using the built-in solver in Python's PuLP library to ensure that the planting scheme remains economically efficient in the face of market fluctuations and climate change. Model solution and analysis: The optimal planting scheme from 2024 to 2030 under the influence of uncertainties was obtained by solving the robust optimization model, and the results were compared and analyzed with the results of the no-volatility scenario to verify the validity and robustness of the model.

3. Results

3.1. Determination of unit sales price

Prior to addressing the problem, it is imperative to accurately ascertain the unit price and sales volume of the 41 crops under consideration. Given the inherent volatility of crop-related data, particularly across different seasons, the data must be subjected to rigorous preprocessing to ensure its reliability and applicability for subsequent analysis.

(1) Crop sales in 2023: After reviewing the relevant literature, it was found that in 2023 China's North China did not appear crop supply exceeds demand and in 2023 North China experienced a long period of high temperatures and droughts, crop yields have been greatly affected, assuming that the total production of various crops in 2023 for the year's crop sales. The calculation method is shown in the formula:

$$sales = crop\ acreage \times production\ from\ diferent\ types\ of\ plots \tag{2}$$

(2) Determination of the exact unit sales price: Based on the range of unit sales prices of crops throughout the year, and taking into account the existence of certain price fluctuations of crops in different periods, we use the results of the following formula in place of the unit sales price of a specific crop:

$$unit\ price\ of\ sales = 2023\min + random\ number \tag{3}$$

3.2. Establishment of a linear programming model

3.2.1 Restrictive condition

The model needs to satisfy the following constraints:

(1) Crop rotation constraint: for all plots, the same crop cannot be grown consecutively in two adjacent years.

$$x_{tij} + x_{(t-1)ij} \leqslant 1, \forall t \in \{2,\ 3,\ 4,\ 5,\ 6,\ 7\}, \forall i, \forall j \tag{4}$$

Where x_{tij} and $x_{(t-1)ij}$ indicate whether crop j is planted in plot i in year t and (t-1), respectively.

(2) Legume planting constraint every three years: each plot is planted with a legume crop at least once in three years.

$$\sum_{t}^{t+2} x_{tij} \geqslant 1, \forall t \in \{1, 2, 3, 4, 5\}, \forall i, \forall j \tag{5}$$

(3) Maximum plot area constraint: for all plots, the total area of each type of crop planted in the plot should not exceed the area of the plot.

$$\sum_{j=1}^{41} x_{tij} \leqslant A_i, \forall t, \forall i \tag{6}$$

Where A_i denotes the area of plot i.

(4) Flat and dry land, terraced land, and hillside land constraint: flat and dry land, terraced land, and hillside land can only be planted with food crops other than rice for one season.

$$x_{ij} = 0, \forall t, \forall i \in \{i | 1 \leqslant i \leqslant 26, i \in Z^*\}, \forall j \in \{j | 1 \leqslant j \leqslant 15, j \in Z^*\} \tag{7}$$

Indicates x_{ij} crop j for the i-th plot of land.

(5) One season of rice on watered land constraint: only one season of rice can be grown on watered land.

$$x_{ij} \times x_{(i+17)j} = 0, \forall i \in \{i | 27 \leqslant i \leqslant 34, i \in Z^*\}, j = 16 \tag{8}$$

Where $x_{(i+17)j}$ indicates whether the second season j crop of the ith piece of watered land is planted or not.

(6) Two season vegetable constraint for watered land:

Vegetables other than cabbage, white radish and red radish can be grown in the first season of the watered land.

$$x_{ij} = 0, \forall i \in \{i | 27 \leqslant i \leqslant 34, i \in Z^*\}, \forall j \notin \{j | 16 \leqslant j \leqslant 34, j \in Z^*\} \tag{9}$$

Only cabbage, white radish and red radish can be planted in the second season of the watered land.

$$x_{ij} = 0, \forall i \in \{i | 56 \leqslant i \leqslant 63, i \in Z^*\}, \forall j \notin \{35, 36, 37\} \tag{10}$$

(7) Common shed first season constraint: only vegetables other than cabbage, white radish and carrot may be grown in the first season of a common shed.

$$x_{ij} = 0, \forall i \in \{i | 36 \leqslant i \leqslant 51, i \in Z^*\}, \forall j \notin \{j | 17 \leqslant j \leqslant 34, j \in Z^*\} \tag{11}$$

(8) Second quarter constraints for ordinary greenhouses: only edible crops may be grown in the second quarter of ordinary greenhouses.

$$x_{ij} = 0, \forall i \in \{i | 36 \leqslant i \leqslant 51, i \in Z^*\}, \forall j \notin \{j | 17 \leqslant j \leqslant 34, j \in Z^*\} \tag{12}$$

(9) Intelligent greenhouse constraint: intelligent greenhouses may only grow vegetables other than cabbage, white radish and carrot in both seasons.

$$x_{ij} = 0, \forall i, \forall j \notin \{j | 17 \leqslant j \leqslant 34, j \in Z^*\} \tag{13}$$

3.2.2 Linear Programming Modeling

For the part of the linear programming model that only considers constraints, we chose linprog's highs algorithm because it can efficiently handle large-scale linear programming problems and ensure optimal solutions under complex constraints. highs algorithm combines the advantages of the simplex and interior-point methods, which is robust and suitable for our crop cultivation optimization problem; it is also suitable for crop cultivation optimization.

Finally, the maximum return value of the following linear programming model is obtained using the interior point method:

$$maxP_0 = \sum_{i=1}^{82} \sum_{j=1}^{41} (Y_{ij} \times U_{ij} - C_{ij}) \times S_{ij} - \max\left(0, \sum_{j=1}^{41} \left(\sum_{i=1}^{82} S_{ij} - Q_j\right) \times U_j\right) \times \beta$$

$$s.t. \begin{cases} x_{tij} + x_{(t-1)ij} \leqslant 1, \forall t \in \{2,3,4,5,6,7\}, \forall i, \forall j \\ \sum_{t}^{t+3} x_{tij} \geqslant 1, \forall t \in \{1,2,3,4,5\}, \forall i, \forall j \\ \sum_{j=1}^{41} x_{tij} \leqslant A_i, \forall t, \forall i \\ x_{ij} = 0, \forall t, \forall i \in \{i \mid 1 \leqslant i \leqslant 26, i \in Z^*\}, \forall j \in \{j \mid 1 \leqslant j \leqslant 15, j \in Z^*\} \\ x_{ij} \times x_{(i+17)j} = 0, \forall i \in \{i \mid 27 \leqslant i \leqslant 34, i \in Z^*\}, j = 16 \\ x_{ij} = 0, \forall i \in \{i \mid 56 \leqslant i \leqslant 63, i \in Z^*\}, \forall j \notin \{35,36,37\} \\ x_{ij} = 0, \forall i \in \{i \mid 36 \leqslant i \leqslant 51, i \in Z^*\}, \forall j \notin \{j \mid 17 \leqslant j \leqslant 34, j \in Z^*\} \\ x_{ij} = 0, \forall i \in \{i \mid 63 \leqslant i \leqslant 78, i \in Z^*\}, \forall j \notin \{38, 39, 40, 41\} \\ x_{ij} = 0, \forall i, \forall j \notin \{j \mid 17 \leqslant j \leqslant 34, j \in Z^*\} \end{cases} \quad (14)$$

3.2.3 Defining the uncertainty set

The expected annual sales volume of various crops, planting costs and the growth rate of mu yield are all an uncertain interval, while the annual changes in the sales price of vegetables and edible fungi are also uncertain, after considering the uncertainty of the above factors and the potential risk, a robust optimization model is established to find the optimal planting scheme.

Based on the potential risks in the actual planting, define the random variables, establish the uncertainty set, the uncertainty set has the following six random variables in the parameter variation range:

(1) Average annual growth rate of expected sales volume of wheat and corn: 5% to 10%:

$$q_{tij} \in [0.05, 0.1], \forall t, \forall j \in \{6, 7\} \quad (15)$$

q_{tij} denotes the average annual growth rate of expected sales volume of crop j in year t.

(2) Expected sales of crops other than wheat and corn relative to sales in 2023: ±5% change:

$$q_{tij} \in [0.05, 0.1], \forall t, \forall j \in \{6, 7\} \quad (16)$$

(3) Annual acreage production of crops: ±10% change:

$$y_{tij} \in [-0.1, 0.1], \forall t, \forall i, \forall j \quad (17)$$

y_{tij} indicates the average annual growth rate of the annual acre yield of crop j in plot i in year t.

(4) Average annual growth rate of cost of cultivation of crops: about 5%:

$$c_{tij} \in [-0.045, 0.055], \forall t, \forall i, \forall j \quad (18)$$

c_{tij} denotes the average annual growth rate of cultivation cost of crop j in plot i in year t.

(5) Average annual growth rate of selling price of vegetable crops: around 5%:

$$u_{tj} \in [0.045, 0.045], \forall t, \forall j \in \{j \mid 17 \leqslant j \ll 37\} \quad (19)$$

u_{tj} indicates the annual evaluated growth rate of sales price of crop j in year t.

(6) Average annual growth rate of sales price of edible mushrooms other than morel mushrooms: -5% to -1%:

$$u_{tj} \in [-0.05, -0.01], \forall t, \forall j \in \{38, 39, 40\} \quad (20)$$

3.2.4 Objective function

The relationship between the constraints and the uncertainty set is synthesized to finally define the objective function:

$$\begin{aligned} maxP_0 = & \sum_{t=1}^{7}\sum_{i=1}^{82}\sum_{j=1}^{41} S_{tij\times}\left[Y_{ij}\times(1+y_{ij})^{t}\times U_j\times(1+u_j)^{t} - C_{ij}\times(1+C_{ij})^{t}\right] \\ & - \max\left(0, \sum_{t=1}^{7}\sum_{j=1}^{41}\left(\sum_{i=1}^{82} S_{ij} - Q_j\times(1+q_j)^{t}\right)\times U_j\right)\times\beta \end{aligned} \tag{21}$$

3.2.5 Solving the model

The results obtained from the robust optimization model are calculated to find the net profit for each year from 2024 to 2030, compared to the data without fluctuations, and the data comparison is displayed as shown in Table 1:

Table 1. Net profit for each of the years 2024 through 2030

Year	Annual earnings without uncertainty/(yuan million)	Annual earnings with uncertainty/(yuan million)
2024	8.34	10.04
2025	8.04	9.86
2026	8.39	10.18
2027	8.12	10.17
2028	7.76	10.40
2029	7.95	10.53
2030	8.53	10.47

It can be clearly seen that after adding uncertainties such as the average annual growth rate of crop sales, the total annual profit for the village is significantly higher than the total annual profit for each indicator relative to the 2023 stabilization scenario. After analyzing the crops, it is concluded that the sales and acreage of all crops, except for corn and wheat, which have some positive sales growth, move up and down in a controlled manner within a range influenced by the market and natural climate, and that the sales price of vegetable crops has positive growth and that of edible mushroom crops has negative growth, but overall, the profit on sales of vegetable crops is much higher than the profit on edible mushroom crops.

3.3. Study of substitutability and complementarity

3.3.1 Analysis of correlation

After calculating the correlation coefficients based on the three main variables of each type of crops: planting cost, selling unit price and selling volume, the corresponding heat map is visualized (the darker the color, the stronger the correlation), and the correlation results are shown in Figures 1, 2 and 3:

Figure 1. Heat map of correlation coefficients (left grain, right grain (legume))

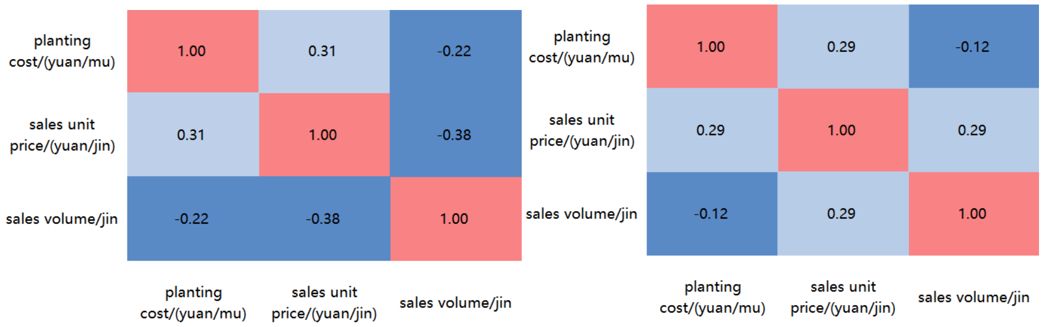

Figure 2. Heat map of correlation coefficients (left vegetable, right vegetable (legume))

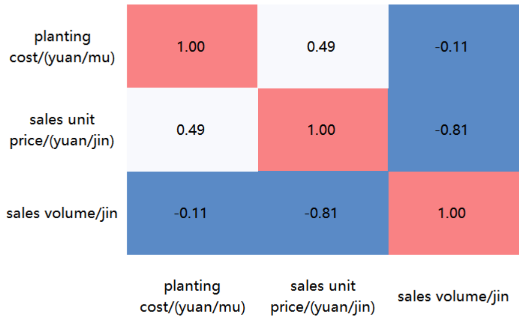

Figure 3. Heat map of correlation coefficients (edible mushrooms)

The combination of sales volume and planting cost of grain crops, the combination of sales volume and planting cost of grain (bean) crops, the combination of unit price and planting cost of vegetable crops, the combination of unit price and planting cost of vegetable (bean) crops, the combination of unit price and sales volume, the combination of unit price and sales volume, and the combination of unit price and planting cost of edible fungus crops are all positively correlated; most of the correlation coefficients are at about 0.25, and the correlation coefficient of the correlation coefficient between sales volume and planting cost of grain (bean) crops is 0.43, and the correlation coefficient between sales unit price and planting cost of edible fungus crops is 0.49, which is a strong positive correlation, and it can be prioritized in actual planting, and more resources can be invested to achieve the maximization of the rational allocation of resources.

While the rest of the combinations show negative correlation, most of them have correlation coefficients greater than -0.5, showing a weak negative correlation, while the correlation coefficient of the combination of sales unit price and planting cost of grain (bean) crops is -0.75, and the correlation coefficient of the combination of sales unit price and sales volume of edible fungus crops is -0.81, with a strong negative correlation, which needs to be introduced as a restriction in the model, or else it may lead to a strong negative correlation for a single crop in the planting. may lead to overdependence on a single crop.

3.3.2 Analysis of complementarity

Crop complementarities were analyzed on the basis of correlation, as shown in Figure 4:

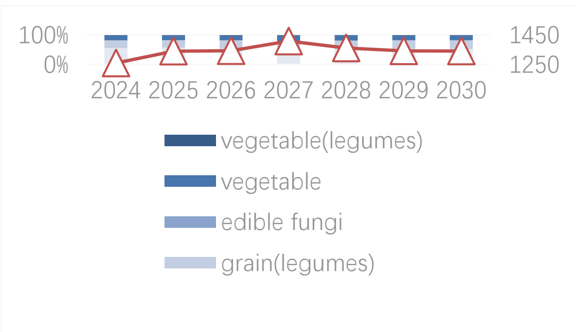

Figure 4. Proportion of area under cultivation

It can be seen from the image that the sum of the arable land occupied by grain and food (legumes) crops fluctuates up and down around 80% of the total area with minimal fluctuations, irrespective of changes in the area under cultivation. This indicates that there is a degree of complementarity between grain crops and grain (legumes) crops.

4. Conclusion

North China, a key grain-producing region in China, shoulders significant national food security responsibilities. However, its mountainous areas face complex agricultural development challenges due to varied topography, climate, and limited, diverse arable land. Existing studies often focus on single-cropping, neglecting the region's unique conditions. Thus, in-depth research to address these issues and propose integrated strategies is highly practical.

The model developed in this study considers various real-world agricultural constraints, enhancing its applicability and providing a feasible optimization scheme for North China's crop cultivation. It incorporates key data like sales prices and volumes of 41 crops, and reveals inter-crop substitutability and complementarity through correlation analysis, supporting optimization decisions. The robust optimization model effectively handles market and natural uncertainties, ensuring decision-making effectiveness even in adverse conditions, demonstrating stability and reliability.

The model's concept extends beyond agriculture to fields like water resource management and forestry planning. By adjusting parameters and constraints, it can address broader resource optimization problems. Future research could introduce dynamic factors like real-time market prices and climate change to enhance the model's timeliness and adaptability. Additionally, integrating the model with smart agriculture concepts could lead to a decision support system offering farmers real-time planting optimization solutions.

References

[1] Wang J, Wang Y. Multi-stage Crop Planting Strategy optimization Model Based on PSO Algorithm [C] // 2024 3rd International Conference on Electronics and Information Technology (EIT). IEEE, 2024: 915-919.

[2] Custodio J M, Billones R K, Concepcion R, et al. Optimization of Crop Harvesting Schedules and Land Allocation Through Linear Programming [J]. Process Integration and Optimization for Sustainability, 2024, 8 (1): 123-134.

[3] Esteso A, Alemany M M E, Ortiz Á, et al. Crop planting and harvesting planning: Conceptual framework and sustainable multi-objective optimization for plants with variable molecule concentrations and minimum time between harvests[J]. Applied Mathematical Modelling, 2022, 112: 136-155.

[4] Liang K, Lei J, Wu F, et al. Analysis of optimal planting scheme based on the North China region [C] // Journal of Physics: Conference Series. IOP Publishing, 2024, 2898 (1): 012004.

[5] Reis S A, Leal J E, Thomé A M T. A Two-Stage Stochastic Linear Programming Model for Tactical Planning in the Soybean Supply Chain [J]. Logistics, 2023, 7 (3): 49.

[6] Azab R, Mahmoud R S, Elbehery R, et al. A Bi-Objective Mixed-Integer Linear Programming Model for a Sustainable Agro-Food Supply Chain with Product Perishability and Environmental Considerations [J]. Logistics, 2023, 7 (3): 46.

[7] Patel J, Chandel S T, Sonker A. An optimization driven deep belief network model for crop yield prediction in IoT based smart agriculture [J]. Multimedia Tools and Applications, 2024: 1-23.

[8] Shankar K, Moorthi M. An efficient IoT-based crop yield prediction framework using optimal ensemble learning and hybridized optimization model [J]. Earth Science Informatics, 2025, 18 (1): 156.

[9] Kalimuthu T, Kalpana P, Kuppusamy S, et al. Intelligent decision-making framework for agriculture supply chain in emerging economies: Research opportunities and challenges [J]. Computers and Electronics in Agriculture, 2024, 219: 108766.

[10] Beyer H G, Sendhoff B. Robust optimization–a comprehensive survey [J]. Computer methods in applied mechanics and engineering, 2007, 196 (33-34): 3190-3218.