Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 14, No. 3, 2025

Research on Crop Planting Strategies Based on Optimization Algorithms

Zhiwei Wen[1, †, *], Chao Liang[1, †], Yurui Yao[2]

[1]College of Mathematics Science, Chengdu University of Technology, Chengdu, China
[2]College of Management Science, Chengdu University of Technology, Chengdu, China

[†] These authors also contributed equally to this work

*Corresponding author: baire314@outlook.com

Abstract: With the development of organic planting industry, choosing suitable crops and optimizing planting strategies are of great significance to the sustainable development of rural economy. In order to help a village design the optimal planting plan for crops in 2024~2030, this paper first builds an objective function based on the treatment of expected sales exceeding the situation in multiple scenarios, sets constraints according to the requirements, models and solves to obtain the optimal planting strategy[1], and comprehensively examines the model's stability and and reasonableness based on the sensitivity analysis. Secondly, Monte Carlo algorithm is introduced to search for the optimal planting strategy of many simulated scenarios, and particle swarm algorithm is used to analyze the uncertainty and planting risk of global optimization of planting strategy. Pearson's correlation coefficient is then used to introduce linear regression model and nonlinear regression model to analyze the substitutability and complementarity between crops, and the best planting plan is obtained through comparison.

Keywords: Crop Planting Strategies, Monte Carlo Simulation, Particle Swarm Algorithm, Pearson Related Analysis.

1. Introduction

A village in the mountainous region of northern China is located in a low-temperature environment, and most of the arable land can only be cultivated for one crop season per year. The village has 1,201 acres of open cultivated land, which is scattered on 34 plots of different sizes, including four types of land: flat dry land, terraced land, hillside land and watered land. There are also 16 ordinary greenhouses and 4 smart greenhouses in the countryside, each with an area of 0.6 acres. With the development of the organic farming industry, it is particularly important to select suitable crops and optimize the planting strategy. The expected sales volume, planting cost, acreage, and sales price of each crop, as well as the effect of the uncertainty factor “climate” on acreage, and the substitutability and complementarity of crops are all factors that will influence the planting plan for the region in 2024~2030. The planning modeling will play a key role in the development of optimal cropping scenarios to improve the profitability of farming.

2. Correlation Analysis

2.1. Patterns of area utilization and yield distribution

The total area of crops grown on each plot in 2023 was counted and then compared with the total area of the plot, so the utilization rate on each plot could be obtained, and it was found that the distribution of the area utilization rate of the plots was relatively simple, with only two cases, 50%, 66.7%, and 100%, and the specific quantities could be derived from Figure 1:

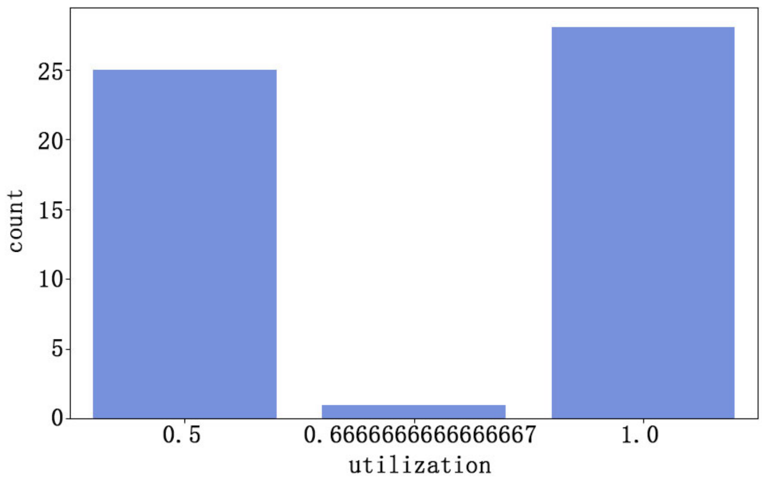

Figure 1. 2023 parcel area utilization rate

The pattern of yield distribution varies according to different crop types and different plot types and is visualized through Figure 2 and Figure 3:

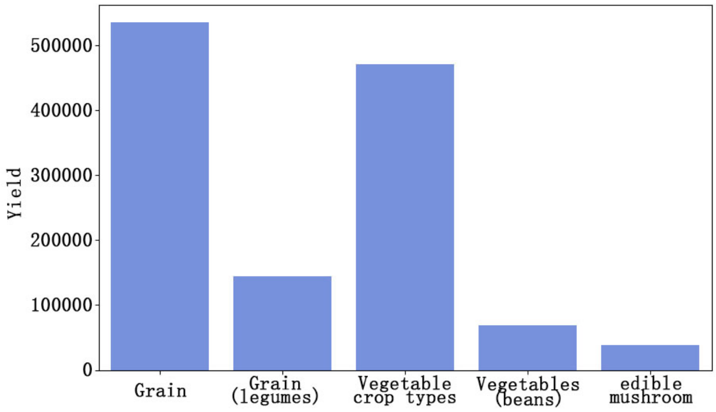

Figure 2. Yield by type for different crops

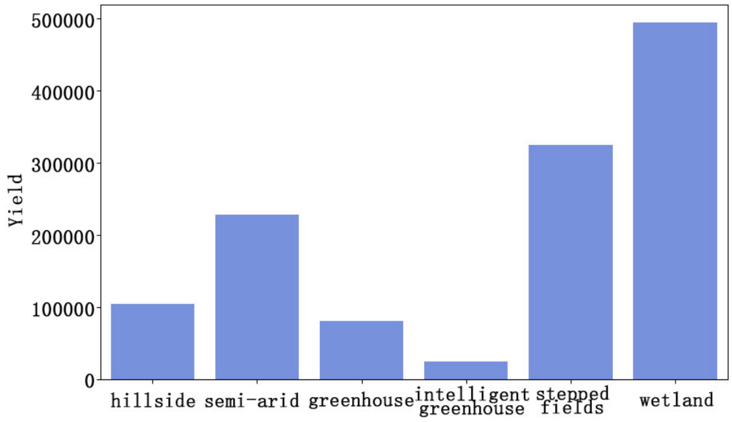

Figure 3. Yield of different plot types

The following conclusions can be drawn from Figure 2: the yield of food grains is the largest among all crop types, followed by vegetables, and legumes are less productive than non-legumes, but more productive than edibles. Observation of Figure 3 shows that the total crop yield is highest in irrigated land, followed by crop yield in terraced land and lowest in greenhouses.

2.2. Pearson correlation analysis

Pearson correlation analysis is a statistical method used to measure the linear correlation between two variables. Based on the planting data of 2023, production, selling price and cost were taken as the research objects and correlation analysis was conducted using Pearson's correlation coefficient, and in order to visualize the results of the analysis, the correlation heat Figure 4 was plotted as follows:

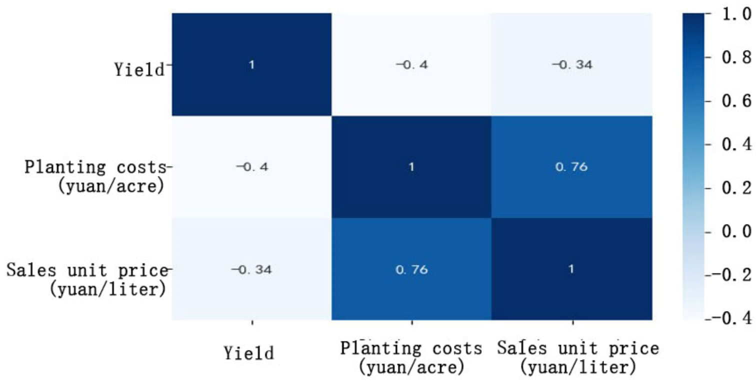

Figure 4. Heat map for correlation analysis of production, selling price and cost.

By observing the correlation heat map 4 , we can find that: the correlation coefficient between planting cost and selling price is 0.76, indicating that there is a strong positive correlation between planting cost and selling price, that is to say, the selling price of the species with higher planting cost tends to be higher; the correlation coefficient between planting cost and yield is -0.40, indicating that there exists a certain negative correlation between planting cost and yield; there is a certain negative correlation between selling price and yield as well. The correlation coefficient between planting cost and yield is -0.40, indicating that there is a negative correlation between planting cost and yield.

So overall, the correlation between the three is more significant.

3. Strategies for Handling Excess Sales Projections in Multiple Scenarios

3.1. Proceeds from different expected sales volumes

When the yield of a crop exceeds the marketing expectations, the excess crop will not be marketable and the objective function is to maximize the return on the marketable crop:

$$\max \sum_{t=1}^{6} \min\left(\sum_{k=1}^{2} \sum_{i} \sum_{j} (P_i \cdot Y_{ijt} \cdot x_{ijtk} - C_{ijt} \cdot x_{ijtk}), W_i 0 \right) \quad (1)$$

Where $W_i 0$ denotes the crop yield from sales in 2023.

When the yield of a crop exceeds the sales expectation, the excess is sold at a 50% discount with an objective function:

$$\max \sum_{i,j,k,t} (P_{ijk} \cdot Y_{ijk} \cdot x_{ijok} - C_{ijk} \cdot x_{ijtk}) + \max(0, \frac{1}{2} \sum_{i,j,k,t} P_{ijk} Y_{ijk} (x_{ijtk} - x_{ijok}) \quad (2)$$

Different types of plots are suitable for different crops, and the planted area of each plot cannot exceed its maximum usable area per year.

Various objective functions are added by setting the constraints of plot type and planting season limit, crop rotation limit, planting cycle of legume crops, minimum planting area, and finally the model is summarized.

3.2. Rationalization of results

The utilization rates of all parcels are counted and the distribution of the results can be analyzed with the help of Figure 5. According to the graph, the utilization rate of the plots varies from 0.6 to 1, and most of the utilization rates are located above 0.9, which is more similar to the situation of the utilization rate in 2023, which indicates that the area utilization rate of the plots is reasonable.

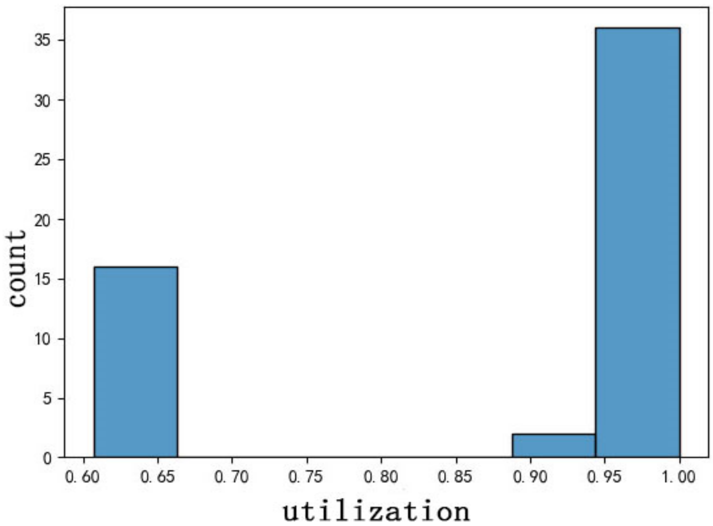

Figure 5. Distribution of area utilization rates

The stability of the model is tested and a sensitivity analysis is performed. Since the calculation of the unit price of sales is obtained by averaging the endpoints of the intervals directly, and the fluctuation of the price in the real scenario is not sufficiently taken into account, we consider randomly selecting 50% of the crops and simulating the price of these crops at the level of 1% ~ 5% for seven consecutive times, and then re-optimizing the model by introducing the new price data, and then observing the change of the objective function (total profit), and the results are shown in Figure. 6. Based on the graph, it can be seen that the optimization algorithm is able to make timely adjustments to the planting strategy when continuous fluctuations in price occur, and makes the percentage change in profit fluctuation vary only between ~1% and ~4.5% at each time.

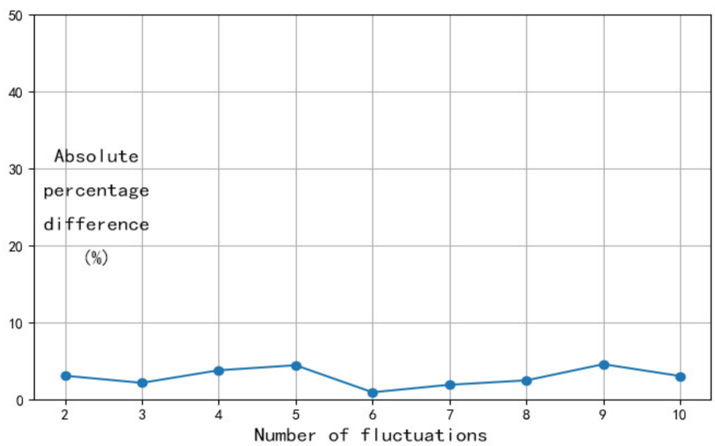

Figure 6. Sensitivity analysis results

4. Strategies for Analyzing Uncertainties and Potential Planting Risks

4.1. Monte-Carlo stochastic simulation

Monte Carlo Simulation (MCS)[2] is a numerical computation method based on the generation of random numbers for estimating the outcome of complex systems or solving mathematical problems that are difficult to analyze. The method obtains a solution to the problem by introducing randomness in a large number of repeated experiments and analyzing the statistical properties of the results of these experiments, generating multiple possible market scenarios covering a variety of uncertainties such as fluctuations in yields, sales volumes, cultivation costs, and sales prices:

Sales Volume Perturbation: generates multiple possible market scenarios covering various uncertainties such as fluctuations in production, sales volume, growing costs and sales price:

$$S_{it} = \sum_{k=1}^{2} \sum_{j} x_{ijtk} \cdot Y_{ijk}, \forall t, \forall j \tag{3}$$

And then the volume of a crop sold in 2023 is:

$$S_{i0} = \sum_{k=1}^{2} \sum_{j} x_{ij0k} \cdot Y_{ijk}, \forall j \tag{4}$$

Define the change in sales volume $\delta_{Sit}, i \in \{6,7\}$, denoting the numbering of wheat and corn. Wheat and maize: sales volume is expected to grow between 5% and 10% per year:

$$\delta_{S_{it}} \sim U(0.05, 0.10), \forall i \in \{6,7\} \tag{5}$$

For other crops: ±5% change in sales volume per year relative to 2023:

$$\delta_{S_{it}} \sim U(-0.05, 0.05), \forall i \in \{1,2,...,5,8,...,41\} \tag{6}$$

Change in acreage: denote by Y_{ijkt} the acreage of crop i grown on plot j in season k of year t, and denote by Y_{ijkt} the acreage of crop i grown on plot j in season k of year 2023, defining the change in acreage, $\delta_{Y_{ijkt}}$, with acreage of all crops fluctuating in the range of ±10% per year due to climate change:

$$\delta_{Y_{ijkt}} \sim U(-0.10, 0.10), \forall i, \forall j, \forall k \tag{7}$$

Change in planting costs: denote by C_{ijkt} the cost of planting the i th crop in the k th season of year t in plot j, and by C_{ijk0} the cost of planting the i th crop in the k th season of year 2023 in plot j. Define the change in planting costs, $\delta_{C_{ijkt}}$, with an average annual increase in planting costs of all crops of about 5 percent:

$$\delta_{C_{ijkt}} \sim N(0.05, 0.01), \forall i, \forall j, \forall k \tag{8}$$

Change in sales price: Denote by P_{ijkt} the sales price of planting the i th crop on the j th plot in the k th season of the t th year, and denote by P_{ijk0} the sales price of planting the i th crop on the j th plot in the k the season of the 2023 year, and define the change in sales price $\delta_{P_{ijkt}}$, where:

Grain crops (number i, $i \in \{1,2,... ,16\}$) have largely stable sales prices with no significant changes:

$$\delta_{P_{ijkt}} = 0, \forall i \in \{1,2,...,16\}, \forall j, \forall k \tag{9}$$

Vegetable crops (number i, $i \in \{17,18,... ,37\}$) numbered sales prices increase on average by about 5% per year:

$$\delta_{P_{ijkt}} \sim N(0.05, 0.01), i \in \{17,18,...,37\}, \forall j, \forall k \tag{10}$$

Edible mushrooms (number i, $i \in \{38,39,40,41\}$ show annual declines of between 1% and 5%, with a fixed decline of 5% for morel mushrooms (number i = 41) in particular:

$$\delta_{P_{ijkt}} \sim U(-0.05, -0.01), i \in \{38,39,40\}, \forall j, \forall k \tag{11}$$

$$\delta_{P_{41jkt}} = -0.05, \forall j, \forall k \tag{12}$$

Monte Carlo simulation to generate multiple market scenarios. The parameters in each scenario are calculated based on the random variables described above and are chosen to generate 100 different market scenarios, each representing a possible future market state.

4.2. Summary of non-linear optimization models

The objective is to maximize total revenue between 2024 and 2030, and production in excess of expected sales is treated as stagnant and disregarded, then the objective function is calculated as follows:

$$\max \sum_{i,j,k,t} \left(P_{ijkt} \cdot Y_{ijkt} \cdot x_{ijtk} - C_{ijkt} \cdot x_{ijtk} \right) \tag{13}$$

The constraints are consistent with the above multi-class scenarios in which the expected sales volume exceeds the handling of the situation model, it is worth noting that the limitations of the expected sales volume is different from the former, in order to avoid the risk of exceeding the expected sales volume of stagnant sales, the requirements of the year, the production of a crop should not be more than the corresponding year's expected sales volume, then:

$$S_{it} \geq \sum_{k=1}^{2} \sum_{j} x_{ijtk} \cdot Y_{ijkt}, \forall t, \forall j \tag{14}$$

For each market scenario, we use nonlinear programming to optimize the cropping strategy to maximize the total return to obtain the final nonlinear optimization summary model.

4.3. Particle Swarm (PSO) Optimization Algorithm Solution

Particle Swarm Optimization (PSO) is an optimization algorithm based on group intelligence. Based on the high dimensionality and dynamic scenario of the above-mentioned planning model, PSO performs well in high-dimensional optimization problems, and can handle a large number of variables. Its strong adaptivity as well as global search ability can effectively explore in the solution space to avoid falling into local optimums, which is suitable for solving optimization problems that change over time. Accordingly, PSO is used in this paper to solve the above nonlinear planning model.

Each particle represents a possible planting scheme, which is represented by a vector containing the planting area x_{ijkt} of all crops on all plots.

The fitness function is to maximize the total return over the period 2024 to 2030:

$$\max \sum_{i,j,k,t} \left(P_{ijkt} \cdot Y_{ijkt} \cdot x_{ijtk} - C_{ijkt} \cdot x_{ijtk} \right) \tag{15}$$

The velocity and position update equations for each particle i are as follows:

$$v_i^{(t+1)} = w \cdot v_i^{(t)} + c_1 \cdot r_1 \cdot \left(p_i^{\text{best}} - x_i^{(t)} \right) + c_2 \cdot r_2 \cdot \left(g^{\text{best}} - x_i^{(t)} \right) \tag{16}$$

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)} \tag{17}$$

When the maximum number of iterations is reached or the global optimal solution converges to maximize profit at the time, the algorithm stops and outputs the optimal solution. The plot utilization rates are all above 0.9, indicating that each plot is quite fully utilized and the results are good, suggesting that this perspective is reasonable[3].

5. Strategies for Analyzing Substitutability and Complementarity Among Crops

5.1. Crop correlation analysis

First, substitution and complementarity are explained[4]: substitution refers to the fact that if there are many substitutes for a good, an increase in price may lead to a decrease in sales as consumers may switch to other options; complementarity refers to the fact that if a good is complementary to other goods, an increase in price may reduce sales of the complementary good and thus overall sales. The impact on sales of changes in the prices of complementary commodities needs to be considered.

In order to investigate the substitutability and complementarity between crops, this paper firstly judges the correlation between mu yield and sales price and planting cost in different crop types, including grain (beans), grain, vegetables (beans), vegetables and edible fungi. Pearson's correlation coefficient was used to calculate the correlation between different variables, and the heat map of correlation was obtained and some of the results were visualized and displayed in Figure 7.

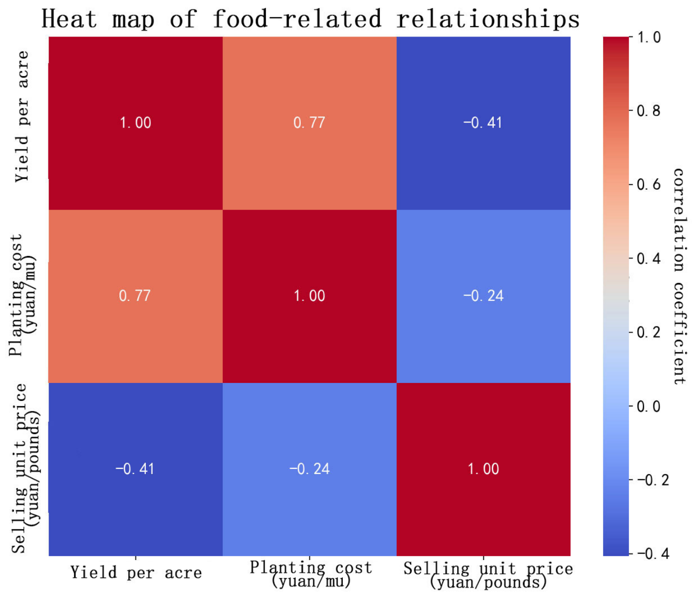

Figure 7. Heat map of food-related relationships

Accordingly, we obtained the correlation coefficients of different crop types of grain (beans), grain, vegetables (beans), vegetables, edible mushrooms in the expected sales volume and sales price, planting costs, in order to obtain the correlation matrix, the closer the correlation coefficient is to 1, indicating that the stronger the positive correlation, i.e., the stronger the complementarity; the closer the correlation coefficient is to -1, indicating that the stronger the negative correlation, i.e., the stronger the substitutability.

5.2. Linear regression models

Multiple linear regression models were used to predict the relationship between acreage value and selling price and cost of planting. Stepwise Regression is a variable selection method used to select the optimal set of variables from multiple candidate independent variables to construct a regression model. Stepwise regression optimizes the predictive and explanatory power of a model by gradually adding or removing variables. Using stepwise regression, the results of solving the regression model are shown in Table 1:

Table 1. Results of linear regression analysis

Linear regression analysis results							
	Non-standardized coefficient		Standardized coefficient	T	P	R2	F
	B	Standard error	Beta				
Constant	20112.495	1484.847	0	13.545	0.000***	0.12	F=28.724, P=0.000***
Planting Costs (yuan/mu)	-2.784	0.519	-0.346	-5.36	0.000***		

From the analysis of the results in Table 1, it can be obtained that the significance p-value is 0.000***, the level is significant, and the original hypothesis that the regression coefficient is 0 is rejected. And p-value 0.000****, indicating that the planting cost has a significant effect on the expected sales volume, the model is well constructed, and collated to get the fitting effect Figure 8. and this paper uses a common nonlinear regression model for fitting, and analyzed using SPSSPRO to get the nonlinear fitting effect Figure 9:

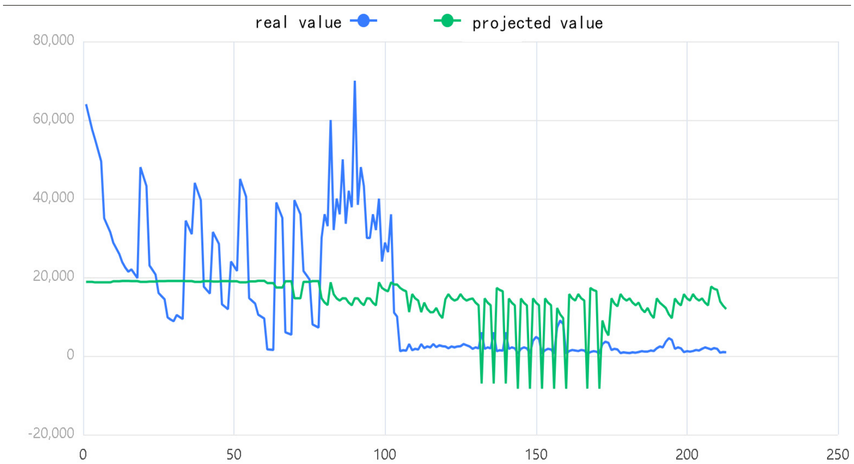

Figure 8. Plot of linear fitting effects

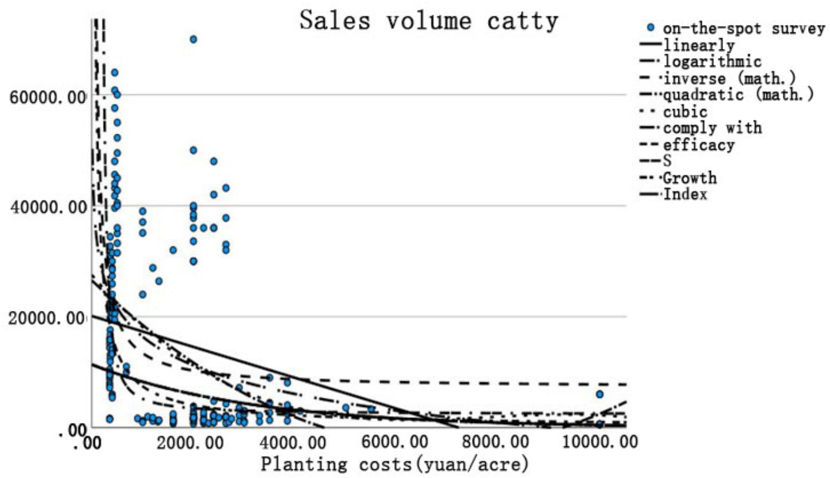

Figure 9. Plot of non-linear fitting effects

5.3. Summary of optimization models based on uncertainty and potential planting risk analysis strategies

Substitutability and complementarity among crops and correlation of expected sales volume, sales price, and planting cost were considered. A multivariate correlation model was developed by combining the matrix of correlation coefficients for different crop types and the regression equations for sales price, and planting cost[5].

Based on the uncertainty and potential planting risk analysis strategy, stochastic planting cost, acreage, and unit sales price parameters for a given crop and values in the correlation matrix were used to adjust these stochastic parameters to simulate new market scenarios under substitutability and complementarities (constraints were consistent and were solved using the pso particle swarm algorithm[6] as well.). Again, through simulation, 100 different market scenarios are chosen to be generated, each representing a possible future market state, resulting in an optimized model summary.

5.4. Comparative analysis of model summary results

For the two planting strategies, the total yield of each crop was calculated separately and compared and analyzed using line graphs, as shown in Figure 10, there is a large discrepancy between the distribution of the total yields of the different crops planted in the two questions, with the former having a more centralized distribution of yields and the latter having a more decentralized distribution of yields.

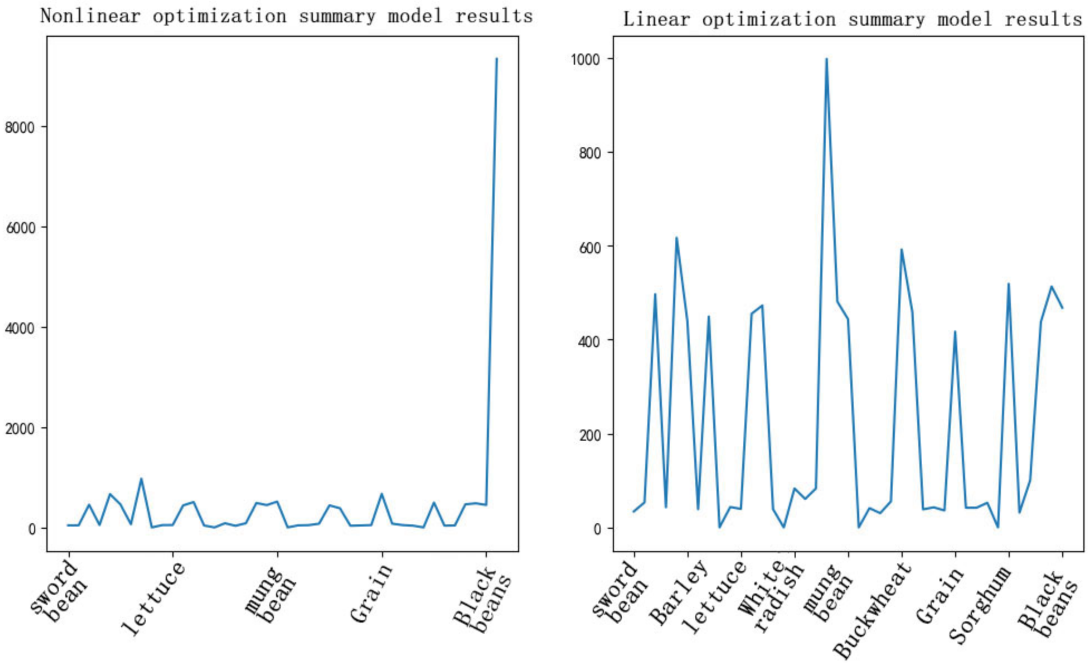

Figure 10. Comparison of total yields considering uncertainty and considering substitutability and complementarity between crops

6. Conclusion

In this paper, through the analysis of experimental results and sensitivity test, the stability and reasonableness of the model based on the handling of the expected sales exceeding the situation in different situations are effectively confirmed. At the same time, based on the uncertainty of influencing factors and potential planting risks, and considering the substitutability and complementarity between various crops, Monte Carlo simulation and optimization algorithms were used to find the optimal planting strategy, and the obtained optimization model summarizes the strategy as an excellent planting strategy that can be adapted to a variety of real-world scenarios.

References

[1] Zhang Jianheng, Zhang Yixing, Hou Saixai et al. Research progress on the benefits of cover crops and their rational selection[J]. Journal of Agricultural Engineering, 2023.

[2] Zhang Mei, Yang Lili. A study on performance evaluation of risk management tools for farmers' cooperatives [J]. Insurance Research, 2022, (03): 58-69. DOI:10.13497/j.cnki.is.2022.03.005.

[3] Meng Zhihang. Research on multi-objective optimization system for limited resource scheduling [D]. Hebei University, 2024.

[4] Zhu Shouyin, YANG Dongxia, CHENG Hua. Research on the construction of crop seed production bases with high quality--Research and reflection based on the practice in four provinces and five places[J]. Seed, 2024, 43(04): 147-156. DOI: 10.16590/j.cnki.1001-4705.2024.04.147.

[5] Deng Chengjun. Research on key technology of intelligent service platform for crop production [D]. Chengdu University, 2022.DOI: 10.27917/d.cnki.gcxdy.2022.000330.

[6] Wei Zhi-Fang. Intelligent optimization algorithm for solving multimodal problems[D]. Xi'an University of Electronic Science and Technology, 2023.